명세서

발명의 명칭: 유전자 발현 조절을 위한 인위적인 게놈 조작

기술분야

[1] 본 발명은 중복 유전자의 발현을 조절하기 위한 발현 조절용 조성물 또는 이를 이용한 방법에 관한 것이다. 보다 구체적으로, 중복 유전자의 전사 조절 영역을 표적화할 수 있는 가이드핵산을 포함하는 발현 조절용 조성물 및 이를 이용한 중복 유전자의 전사 조절 영역을 인위적으로 조작 및/또는 변형시켜 중복 유전자의 발현을 조절하는 방법에 관한 것이다. 또한 중복 유전자의 발현을 조절하기 위한 발현 조절용 조성물을 이용하여 유전자 중복에 의해 발생하는 질환을 치료 또는 개선하는 방법에 관한 것이다.

[2]

배경기술

[3] 유전자 중복은 염색체의 유전자 재조합 과정 중 발생하는 오류 중 하나로, 염색체의 일부 구간이 중복되는 복제 현상이다. 유전자 중복은 돌연변이의 한 종류이며, 다음 세대로 유전된다. 유전자 중복은 염색체의 일부 구간이 복제되지 않아 일어나는 유전자 결실과 함께 유전자 발현에 영향을 준다.

[4] 유전자 중복은 유전성 질환의 원인이 되기도 한다. 대표적으로 샤르코 마리 투스 질환 type 1A는 염색체의 특정 구간에서 일어나는 유전자 중복이 원인이며, 유전자 중복으로 인해 손과 발의 말초신경 발달에 관여하는 유전자의 과발현이 초래되며 이로 인해 손과 발에 기형을 유발하게 된다.

[5] 이렇듯 세포의 증식, 사멸, 노화 및 분화와 같은 생물학적 과정의 정상적인 수행을 위해 적절한 때에 적절한 위치에서 유전자가 발현하는 것이 중요하다. 부적절한 시기와 위치에서 유전자가 부적절하게 발현될 때, 특히 유전자 중복으로 인한 비정상적인 유전자의 발현은 질병으로 이어질 수 있기 때문에, 각각 유전자의 발현을 조절하는 분자 기작을 이해가 필요하며, 이를 위해서는 각각 유전자와 관련된 전사 조절 요소를 확인하는 것이 중요하다. 유전자의 발현을 정확하게 조절할 수 있도록하는 전사 조절 요소는 프로모터 및 원거리 조절 요소 등 다양하며, 또한 유전자 발현 조절에 관여하는 전사 인자, 활성제 및 보조 활성화 인자들도 있다.

[6] 전사 조절 요소의 변화를 통해 유전자 발현이 조절되고, 전사 조절 요소의 비정상적인 변화가 유전자의 비정상적인 발현을 초래하여 질병을 유발할 수 있다. 그러므로, 전사 조절 요소를 변화시키으로 다양한 질병을 초래할 수도, 또는 질병을 개선 및 치료할 수도 있을 것이다.

[7] 하지만 현재 전사 조절 요소를 조절하는 방법은 일시적인 유전자 발현을 조절할 뿐, 지속적인 유전자 발현 조절은 어려운 실정으로, 유전자 발현 이상 또는 장애로 인한 질환을 치료하기 위한 근본적인 치료 방법은 없는 상태이다.

따라서, 보다 지속적인 치료효과를 위해서는 전사 조절 요소의 유전적 교정 또는 변형을 통한 방법이 필요하다.

[8]

[9] [선행기술문헌]

[10] 비특허문헌

[11] 1. Hamdan, H., Kockara, N.T., Jolly, L.A., Haun, S., and Wight, P.A. (2015). Control of human PLP1 expression through transcriptional regulatory elements and alternatively spliced exons in intron 1. ASN Neuro *7*.

[12] 2. Hamdan, H., Patyal, P., Kockara, N.T., and Wight, P.A. (2018). The wmN1 enhancer region in intron 1 is required for expression of human PLP1. Glia.

[13] 3. Meng, F., Zolova, O., Kokorina, N.A., Dobretsova, A., and Wight, P.A. (2005). Characterization of an intronic enhancer that regulates myelin proteolipid protein (Plp) gene expression in oligodendrocytes. J Neurosci Res *82*, 346-356.

[14] 4. Tuason, M.C., Rastikerdar, A., Kuhlmann, T., Goujet-Zalc, C., Zalc, B., Dib, S., Friedman, H., and Peterson, A. (2008). Separate proteolipid protein/DM20 enhancers serve different lineages and stages of development. J Neurosci *28*, 6895-6903.

[15] 5. Wight, P.A. (2017). Effects of Intron 1 Sequences on Human PLP1 Expression: Implications for PLP1-Related Disorders. ASN Neuro *9*, 1759091417720583.

[16]

발명의 상세한 설명

기술적 과제

[17] 본 발명은 일 구체예로서 세포의 게놈 내 존재하는 중복 유전자의 발현을 조절하기 위한 발현 조절용 조성물을 제공하고자 한다.

[18] 본 발명은 일 구체예로서 진핵 세포의 게놈 내 존재하는 중복 유전자의 발현을 조절하는 방법을 제공하고자 한다.

[19] 본 발명은 일 구체예로서 발현 조절용 조성물을 이용한 유전자 중복 질환의 치료 방법을 제공하고자 한다.

[20]

과제 해결 수단

[21] 상기 과제를 해결하기 위하여, 본 발명은 세포의 게놈 내 존재하는 중복 유전자의 발현을 조절하기 위한 발현 조절용 조성물에 관한 것이다. 보다 구체적으로, 중복 유전자의 전사 조절 영역을 표적화할 수 있는 가이드핵산을 포함하는 발현 조절용 조성물 및 이를 이용한 중복 유전자의 전사 조절 영역을 인위적으로 조작 및/또는 변형시켜 중복 유전자의 발현을 조절하는 방법에 관한 것이다. 또한 중복 유전자의 발현을 조절하기 위한 발현 조절용 조성물을 이용하여 유전자 중복에 의해 발생하는 질환을 치료 또는 개선하는 방법에 관한 것이다.

[22]

[23] 본 발명은 세포의 게놈 내 존재하는 중복 유전자의 발현을 조절하기 위한 발현 조절용 조성물을 제공한다.

[24] 구현예에서, 발현 조절용 조성물은 다음을 포함할 수 있다:

[25] 중복 유전자의 전사 조절 영역 내 표적서열을 표적화할 수 있는 가이드핵산 또는 이를 암호화하는 핵산 서열; 및

[26] 하나 이상의 에디터단백질 또는 이를 암호화하는 핵산 서열.

[27] 상기 가이드핵산은 중복 유전자의 전사 조절 영역 내 표적 서열을 표적화할 수 있는 가이드 도메인을 포함할 수 있다.

[28] 이때, 가이드 도메인은 중복 유전자의 전사 조절 영역 내 표적 서열 중 가이드핵산 결합 서열과 상보적 결합을 형성할 수 있는 뉴클레오타이드 서열을 포함할 수 있다.

[29] 이때, 가이드 도메인은 중복 유전자의 전자 조절 영역 내 표적 서열 중 가이드핵산 결합 서열과 상보적 결합을 형성할 수 있다.

[30] 이때, 상기 상보적 결합은 0 내지 5의 미스매칭(mismatching) 결합을 포함할 수 있다.

[31] 상기 가이드핵산은 제 1 상보적 도메인, 연결 도메인, 제 2 상보적 도메인, 근위 도메인 및 꼬리 도메인으로 구성된 군에서 선택된 하나 이사의 도메인을 포함할 수 있다.

[32] 상기 에디터 단백질은 CRISPR 효소일 수 있다.

[33] 상기 가이드핵산 및 에디터단백질은 가이드핵산-에디터단백질 복합체를 형성할 수 있다.

[34] 이때, 상기 가이드핵산-에디터단백질 복합체는 가이드핵산의 일부 핵산과 에디터단백질의 일부 아미노산이 상호작용하여 형성될 수 있다.

[35] 상기 전사 조절 영역은 프로모터 영역, 인핸서 영역, 사일런서 영역, 인슐레이터 영역 및 로커스 컨트롤 영역(locus control region, LCR)으로 구성된 군에서 선택된 하나 이상의 영역일 수 있다.

[36] 상기 표적 서열은 중복 유전자의 전사 조절 영역 내 위치하는 연속하는 10 내지 25개의 뉴클레오타이드 서열일 수 있다.

[37] 상기 표적 서열은 중복 유전자의 프로모터 영역 내에 연속하는 10 내지 25개의 뉴클레오타이드 서열 또는 프로모터 영역에 인접한 연속하는 10 내지 25개의 뉴클레오타이드 서열일 수 있다.

[38] 이때, 상기 표적 서열은 중복 유전자의 코어 프로모터 영역 내에 연속하는 10 내지 25개의 뉴클레오타이드 서열 또는 코어 프로모터 영역에 인접한 연속하는 10 내지 25개의 뉴클레오타이드 서열일 수 있다.

[39] 이때, 표적 서열은 중복 유전자의 코어 프로모터 영역 중 TATA 박스 영역을 포함하는 연속하는 10 내지 25개의 뉴클레오타이드 서열 또는 TATA 박스 영역에 인접한 연속하는 10 내지 25개의 뉴클레오타이드 서열일 수 있다.

[40] 이때, 상기 표적 서열은 중복 유전자의 코어 프로모터 영역 내에 존재하는 5'-TATA-3' 서열의 전체 또는 일부를 포함하는 연속하는 10 내지 25개의 뉴클레오타이드 서열일 수 있다.

[41] 이때, 상기 표적 서열은 중복 유전자의 코어 프로모터 영역 내에 존재하는 5'-TATAWAW-3' 서열(W=A 또는 T)의 전체 또는 일부를 포함하는 연속하는 10 내지 25개의 뉴클레오타이드 서열일 수 있다.

[42] 이때, 상기 표적 서열은 중복 유전자의 코어 프로모터 영역 내에 존재하는 5'-TATAWAWR-3' 서열(W = A 또는 T, R = A 또는 G)의 전체 또는 일부를 포함하는 연속하는 10 내지 25개의 뉴클레오타이드 서열일 수 있다.

[43] 이때, 상기 표적 서열은 중복 유전자의 코어 프로모터 영역 내에 존재하는 5'-CATAAAA-3' 서열, 5'-CATAAAA-3' 서열, 5'-TATAA-3' 서열, 5'-TATAAAA-3' 서열, 5'-CATAAATA-3' 서열, 5'-TATATAA-3' 서열, 5'-TATATATATATATAA-3' 서열, 5'-TATATTATA-3' 서열, 5'-TATAAA-3' 서열, 5'-TATAAAATA-3' 서열, 5'-TATATA-3' 서열, 5'-GATTAAAAA-3' 서열, 5'-TATAAAAA-3' 서열, 5'-TTATAA-3' 서열, 5'-TTTTAAAA-3' 서열, 5'-TCTTTAAAA-3' 서열, 5'-GACATTTAA-3' 서열, 5'-TGATATCAA-3' 서열, 5'-TATAAATA-3' 서열, 5'-TATAAGA-3' 서열, 5'-AATAAA-3' 서열, 5'-TTTATA-3' 서열, 5'-CATAAAAA-3' 서열, 5'-TATACA-3' 서열, 5'-TTTAAGA-3' 서열, 5'-GATAAAG-3' 서열, 5'-TATAACA-3' 서열, 5'-TCTTATCTT-3' 서열, 5'-TTGTACTTT-3' 서열, 5'-CATATAA-3' 서열, 5'-TATAAAT-3' 서열, 5'-TATATATAAAAAAAA-3' 서열 및 5'-CATAAATAAAAAAAATTA-3' 서열로 구성된 군에서 선택된 서열의 전체 또는 일부를 포함하는 연속하는 10 내지 25개의 뉴클레오타이드 서열일 수 있다.

[44] 이때, 상기 표적 서열은 중복 유전자의 코어 프로모터 영역 내에 존재하는 5'-TATA-3' 서열, 5'-CATAAAA-3' 서열, 5'-CATAAAA-3' 서열, 5'-TATAA-3' 서열, 5'-TATAAAA-3' 서열, 5'-CATAAATA-3' 서열, 5'-TATATAA-3' 서열, 5'-TATATATATATATAA-3' 서열, 5'-TATATTATA-3' 서열, 5'-TATAAA-3' 서열, 5'-TATAAAATA-3' 서열, 5'-TATATA-3' 서열, 5'-GATTAAAAA-3' 서열, 5'-TATAAAAA-3' 서열, 5'-TTATAA-3' 서열, 5'-TTTTAAAA-3' 서열, 5'-TCTTTAAAA-3' 서열, 5'-GACATTTAA-3' 서열, 5'-TGATATCAA-3' 서열, 5'-TATAAATA-3' 서열, 5'-TATAAGA-3' 서열, 5'-AATAAA-3' 서열, 5'-TTTATA-3' 서열, 5'-CATAAAAA-3' 서열, 5'-TATACA-3' 서열, 5'-TTTAAGA-3' 서열, 5'-GATAAAG-3' 서열, 5'-TATAACA-3' 서열, 5'-TCTTATCTT-3' 서열, 5'-TTGTACTTT-3' 서열, 5'-CATATAA-3' 서열, 5'-TATAAAT-3' 서열, 5'-TATATATAAAAAAAA-3' 서열 및 5'-CATAAATAAAAAAAATTA-3' 서열로 구성된 군에서 선택된 서열의 5' 말단 또는3' 말단에 연속하는 10 내지 25개의 뉴클레오타이드 서열일 수 있다.

[45] 상기 표적 서열은 중복 유전자의 인핸서 영역 내에 연속하는 10 내지 25개의 뉴클레오타이드 서열일 수 있다.

[46] 상기 표적 서열은 중복 유전자의 인핸서 영역에 인접한 연속하는 10 내지 25개의 뉴클레오타이드 서열일 수 있다.

[47] 상기 표적 서열은 중복 유전자의 전사 조절 영역 내 핵산 서열 중 PAM(proto-spacer-adjacent Motif) 서열의 5' 말단 및/또는 3' 말단에 인접한 연속하는 10 내지 25개의 뉴클레오타이드 서열일 수 있다.

[48] 이때, 상기 PAM 서열은 CRISPR 효소에 따라 결정될 수 있다.

[49] 상기 CRISPR 효소는 Cas9 단백질 또는 Cpf1 단백질일 수 있다.

[50] 이때, 상기 Cas9 단백질은 스트렙토코커스 피요게네스(Streptococcus pyogenes) 유래의 Cas9 단백질, 캄필로박터 제주니 (Campylobacter jejuni) 유래의 Cas9 단백질, 스트렙토코커스 써모필러스 (Streptococcus thermophilus) 유래의 Cas9 단백질, 스타필로코커스 아우레우스 (Staphylococcus aureus) 유래의 Cas9 단백질 및 네이세리아 메닝기디티스 (Neisseria meningitidis)유래의 Cas9 단백질로 이루어진 군에서 선택된 하나 이상의 Cas9 단백질일 수 있다.

[51] 상기 중복 유전자는 PMP22 유전자, PLP1 유전자, MECP2 유전자, SOX3 유전자, RAI1 유전자, TBX1 유전자, ELN 유전자, JAGGED1 유전자, NSD1 유전자, MMP23 유전자, LMB1 유전자, SNCA 유전자 및 APP 유전자로 구성된 군에서 선택된 하나 이상의 유전자일 수 있다.

[52] 상기 중복 유전자는 암유전자(oncogene)일 수 있다.

[53] 이때, 상기 암유전자는 MYC 유전자, ERBB2(HER2) 유전자, CCND1(Cyclin D1) 유전자, FGFR1 유전자, FGFR2 유전자, HRAS 유전자, KRAS 유전자, MYB 유전자, MDM2 유전자, CCNE(Cyclin E) 유전자, MET 유전자, CDK4 유전자, ERBB1 유전자, MYCN 유전자 및 AKT2 유전자로 구성된 군에서 선택된 하나 이상의 유전자일 수 있다.

[54] 상기 세포는 진핵세포일 수 있다.

[55] 상기 진핵세포는 포유동물 세포일 수 있다.

[56] 상기 가이드핵산 및 에디터단백질은 각각 핵산 서열의 형태로 1 이상의 벡터에 존재할 수 있다.

[57] 이때, 상기 벡터는 플라스미드 또는 바이러스 벡터일 수 있다.

[58] 이때, 상기 바이러스 벡터는 레트로바이러스, 렌티바이러스, 아데노바이러스, 아데노-연관 바이러스(AAV), 백시니아바이러스, 폭스바이러스 및 단순포진 바이러스로 구성된 군에서 선택된 하나 이상의 바이러스 벡터일 수 있다.

[59] 상기 발현 조절용 조성물은 가이드핵산 및 에디터단백질을 가이드핵산-에디터단백질 복합체 형태로 포함할 수 있다.

[60] 상기 발현 조절용 조성물은 도너를 추가로 더 포함할 수 있다.

[61]

[62] 본 발명은 진핵 세포의 게놈 내 존재하는 중복 유전자의 발현을 조절하기 위한 방법을 제공한다.

[63] 구현예에서, 진핵 세포의 게놈 내 존재하는 중복 유전자의 발현을 조절하기

위한 방법은 발현 조절용 조성물을 진핵 세포에 도입하는 단계를 포함할 수 있다.

[64] 상기 발현 조절용 조성물은 다음을 포함할 수 있다:

[65] 중복 유전자의 전사 조절 영역 내 표적서열을 표적화할 수 있는 가이드핵산 또는 이를 암호화하는 핵산 서열; 및

[66] 하나 이상의 에디터단백질 또는 이를 암호화하는 핵산 서열.

[67] 상기 진핵세포는 포유동물 세포일 수 있다.

[68] 상기 가이드핵산은 중복 유전자의 전사 조절 영역 내 표적 서열을 표적화할 수 있는 가이드 도메인을 포함할 수 있다.

[69] 이때, 가이드 도메인은 중복 유전자의 전사 조절 영역 내 표적 서열 중 가이드핵산 결합 서열과 상보적 결합을 형성할 수 있는 뉴클레오타이드 서열을 포함할 수 있다.

[70] 이때, 가이드 도메인은 중복 유전자의 전자 조절 영역 내 표적 서열 중 가이드핵산 결합 서열과 상보적 결합을 형성할 수 있다.

[71] 이때, 상기 상보적 결합은 0 내지 5의 미스매칭(mismatching) 결합을 포함할 수 있다.

[72] 상기 가이드핵산은 제 1 상보적 도메인, 연결 도메인, 제 2 상보적 도메인, 근위 도메인 및 꼬리 도메인으로 구성된 군에서 선택된 하나 이사의 도메인을 포함할 수 있다.

[73] 상기 에디터 단백질은 CRISPR 효소일 수 있다.

[74] 상기 가이드핵산 및 에디터단백질은 가이드핵산-에디터단백질 복합체를 형성할 수 있다.

[75] 이때, 상기 가이드핵산-에디터단백질 복합체는 가이드핵산의 일부 핵산과 에디터단백질의 일부 아미노산이 상호작용하여 형성될 수 있다.

[76] 상기 발현 조절용 조성물은 가이드핵산 및 에디터단백질을 가이드핵산-에디터단백질 복합체 형태로 포함할 수 있다.

[77] 상기 발현 조절용 조성물은 가이드핵산 및 에디터단백질을 각각 핵산 서열로 포함하는 1 이상의 벡터를 포함할 수 있다.

[78] 상기 도입은 전기천공법 (electroporation), 리포좀, 플라스미드, 바이러스 벡터, 나노파티클(nanoparticles) 및 PTD (Protein translocation domain) 융합 단백질 방법 중 선택되는 1 이상의 방법으로 수행될 수 있다.

[79]

[80] 본 발명은 유전자 중복 질환을 치료하기 위한 방법을 제공한다.

[81] 구현예에서, 유전자 중복 질환을 치료하기 위한 방법은 발현 조절용 조성물을 치료 대상에 투여하는 단계를 포함할 수 있다.

[82] 상기 발현 조절용 조성물은 다음을 포함할 수 있다:

[83] 중복 유전자의 전사 조절 영역 내 표적서열을 표적화할 수 있는 가이드핵산 또는 이를 암호화하는 핵산 서열; 및

[84] 하나 이상의 에디터단백질 또는 이를 암호화하는 핵산 서열.

[85] 상기 가이드핵산은 중복 유전자의 전사 조절 영역 내 표적 서열을 표적화할 수 있는 가이드 도메인을 포함할 수 있다.

[86] 이때, 가이드 도메인은 중복 유전자의 전사 조절 영역 내 표적 서열 중 가이드핵산 결합 서열과 상보적 결합을 형성할 수 있는 뉴클레오타이드 서열을 포함할 수 있다.

[87] 이때, 가이드 도메인은 중복 유전자의 전자 조절 영역 내 표적 서열 중 가이드핵산 결합 서열과 상보적 결합을 형성할 수 있다.

[88] 이때, 상기 상보적 결합은 0 내지 5의 미스매칭(mismatching) 결합을 포함할 수 있다.

[89] 상기 가이드핵산은 제 1 상보적 도메인, 연결 도메인, 제 2 상보적 도메인, 근위 도메인 및 꼬리 도메인으로 구성된 군에서 선택된 하나 이사의 도메인을 포함할 수 있다.

[90] 상기 에디터 단백질은 CRISPR 효소일 수 있다.

[91] 상기 가이드핵산 및 에디터단백질은 가이드핵산-에디터단백질 복합체를 형성할 수 있다.

[92] 이때, 상기 가이드핵산-에디터단백질 복합체는 가이드핵산의 일부 핵산과 에디터단백질의 일부 아미노산이 상호작용하여 형성될 수 있다.

[93] 상기 유전자 중복 질환은 샤르코마리투스병 1A형(Charcot-Marie-Tooth 1A, CMT1A), 데제린-소타스병(Dejerine-Sottas disease, DSD), 선천성 저수초형성 신경병(Congenital Hypomyelination Neuropathy, CHN), 루시-레비 증후군(Roussy-Levy syndrome, RLS), 펠리제우스-메르츠바하병(Pelizaeus-Merzbacher disease, PMD), MECP2 중복 증후군(MECP2 duplication syndrome), X-연관 뇌하수체기능저하증(X-linked hypopituitarism, XLHP), Potocki-Lupski 증후군(Potocki-Lupski syndrome, PTLS), 구개심장안면증후군(Velocardiofacial syndrome, VCFS), 윌리엄스 보이렌 증후군(Williams Beuren syndrome, WBS), 알라질증후군(Alagille syndrome, AS), 발육지연 증후군(Growth retardation syndrome), 조기 폐쇄 두개봉합(Premature closure cranial sutures), 보통염색체우성 백질이영양증(Autosomal dominant leukodystrophy, ADLD), 파킨슨병(Parkinson disease) 또는 알츠하이머(Alzheimer disease)일 수 있다.

[94] 상기 유전자 중복 질환은 암유전자(oncogene) 중복에 의한 암일 수 있다.

[95] 이때, 상기 암유전자(oncogene) 중복에 의한 암은 유방암(Breast cancer), 자궁경부암(Cervical cancer), 대장암(Colorectal cancer), 식도암(Esophageal cancer), 위암(Gastric cancer), 교모세포종(Glioblastoma), 두경부암(Head and neck cancer), 간암(Hepatocellular cancer), 신경아세포종(Neuroblastoma), 난소암(Ovarian cancer), 육종(Sarcoma) 또는 소세포폐암(Small cell lung cancer)일 수 있다.

[96] 상기 치료 대상은 인간, 원숭이, 마우스, 래트를 포함하는 포유동물일 수 있다.

[97] 상기 투여는 주사(injection), 수혈(transfusion), 삽입(implantation) 또는 이식(transplantation)으로 수행될 수 있다.

[98]

발명의 효과

[99] 본 발명은 발현 조절용 조성물을 통해 중복 유전자의 발현을 조절할 수 있다. 보다 구체적으로, 중복 유전자의 전사 조절 영역을 표적화하는 가이드핵산을 포함하는 발현 조절용 조성물을 이용해 중복 유전자의 전사 조절 영역을 인위적으로 조작 및/또는 변형하여 중복 유전자의 발현을 조절할 수 있도록 한다. 또한, 중복 유전자의 발현을 조절하기 위한 발현 조절용 조성물을 이용하여 유전자 중복에 의해 발생하는 질환을 개선 또는 치료할 수 있다.

[100]

도면의 간단한 설명

[101] 도 1은 SpCas9-sgRNA 매개 유전자 조작으로 인한 인델 빈도(%)를 나타낸 결과로, sgRNA의 표적 부위를 (a) TATA-box, (b) Enhancer로 나누어 각각 인델 빈도를 나타낸다.

[102] 도 2는 CjCas9-sgRNA 매개 유전자 조작으로 인한 인델 빈도(%)를 나타낸 결과로, sgRNA의 표적 부위를 (a) TATA-box, (b) Enhancer로 나누어 각각 인델 빈도를 나타낸다.

[103] 도 3은 슈반 유사 세포에서 인간 PMP22 유전자의 조절 인자(regulatory elements)를 표적한 SpCas9-sgRNA에 의한 유전자 조작 효과를 나타낸다.

[104] 도 4는 인간 PMP22 유전자의 CDS 부위를 표적한 SpCas9-sgRNA에 의한 프래임시프트 돌연변이 비율을 보여준다.

[105] 도 5는 이중 sgRNA를 이용한 인간 PMP22 유전자의 작은 일부분 삭제를 보여준다.

[106] 도 6은 인간 슈반 유사 세포에서 SpCas9-sgRNA에 의한 인간 PMP22의 mRNA 발현감소를 나타낸 그래프이다.

[107] 도 7은 인간 슈반세포에서 인간 PMP22 유전자의 각 표적 부위별 SpCas9-sgRNA에 의한 PMP22의 효과적 및 특이적 발현감소를 나타낸 그래프로, (a)는 각 표적 부위별 SpCas9-sgRNA에 의한 인델 빈도 측정결과, (b)는 수초화 신호 인자 및 각 표적 부위별 RNP 복합체의 처리 유무에 따른 qRT-PCR로 측정한 상대적인 PMP22의 mRNA 발현 비교결과(n=3, One-way ANOVA 및 Tukey post-hoc tests: * $p < 0.05$), (c)는 원거리 인해서 부위(distal enhancer region) B 및 C를 표적하는 SpCas9-sgRNA에 의한 인델 빈도 측정결과를 나타낸다.

[108] 도 8은 in vitro에서 인간 PMP22 유전자의 TATA-box 부위를 표적하는 CRISPR-Cas9을 통한 PMP22의 효과적 및 특이적 발현감소를 나타낸 그래프로, (a)는 인간 PMP22 위치의 프로모터 부위를 표적하는 표적서열, (b)의 가장 좌측의 그래프는 인간 슈반세포에서 표적 딥 시퀀싱을 이용한 인델 빈도

측정결과, 가운데 그래프는 전체 인델 빈도 중 TATA-box 1 돌연변이 빈도 측정결과(n=3), 가장 우측의 그래프는 인간 슈반세포에서 수초화 신호 인자 및 RNP 복합체의 처리 유무에 따른 qRT-PCR로 측정한 상대적인 PMP22의 mRNA 발현 비교결과(n=3, One-way ANOVA 및 Tukey post-hoc tests: * $p < 0.05$)를 나타낸다.

[109] 도 9은 인간 슈반세포에서 표적 딥 시퀀싱에 의한 in silico 오프-타겟 분석을 통해 발견한 오프-타겟 및 온-타겟에서 PMP22-TATA RNP에 의한 인델 빈도를 나타낸 것으로, (a)는 인델 빈도를 나타낸 그래프이고, (b)는 빈도가 높은 인델 패턴을 보여주며, (c)는 in silico 오프-타겟 분석을 통해 발견한 오프-타겟 위치를 나타낸다.

[110] 도 10는 인간 전체 게놈에서의 PMP22-TATA RNP에 의해 절단된 위치는 보여주는 결과로, (a)는 Genome-wide Circos plot을 보여주고, (b)는 in silico 오프-타겟 분석을 통해 발견한 오프-타겟 위치 중 Digenome-seq에 의해 나타난 오프-타겟 위치를 보여주며, (c)는 오프-타겟 위치에서 인델 빈도를 나타낸 그래프이다.

[111] 도 11은 C22 마우스에서 PMP22-TATA RNA 치료법을 이용한 치료 접근법을 도식화하여 나타낸다.

[112] 도 12은 CMT1A 마우스에서 CRISPR/Cas9에 의한 PMP22의 발현저해를 통한 병증 표현형 완화를 보여주는 결과로, (a)는 mRosa26 또는 PMP22-TATA RNP 복합체를 처리한 좌골 신경에서 표적 딥 시퀀싱을 이용한 인델 빈도를 나타낸 그래프이고(n=3), (b)는 도 11a의 전체 인델 빈도 중 TATA-box 1 돌연변이 빈도 측정결과(n=3)이며, (c)는 mRosa26 또는 PMP22-TATA RNP 복합체를 처리한 좌골 신경으로부터 qRT-PCR을 이용한 상대적인 PMP22의 mRNA 발현양을 비교한 그래프이다.

[113] 도 13는 in silico 분석에 의한 마우스 게놈에서의 PMP22-TATA sgRNA의 오프-타겟 위치 및 인델 빈도를 보여주는 결과로, (a)는 오프-타겟 위치를 보여주며, (b)는 각 오프-타겟의 위치에서 인델 빈도를 나타낸 그래프이다.

[114] 도 14은 CMT1A 마우스에서 CRISPR/Cas9에 의한 PMP22의 발현저해를 통한 병증 표현형 완화를 보여주는 결과로, (a)는 mRosa26 또는 PMP22-TATA RNP 복합체를 처리한 좌골 신경 조직의 semithin section의 이미지이며, (b)의 상단 그래프는 PMP22-TATA RNP를 처리한 마위스에서 g-ratio가 증가함을 보여주는 산점도(scatter plot)이고, 하단 그래프는 MP22-TATA RNP를 처리한 마위스에서 수초화된 축산의 직경이 증가함을 보여주는 그래프이다.

[115] 도 15는 CMT1A 마우스에서 CRISPR/Cas9에 의한 PMP22의 발현저해를 통한 전기 생리학적 변화를 보여주는 결과로, (a)는 말단 지연속도(Distal latency, DL)의 변화를 나타낸 그래프이고, (b)는 운동 신경 전도 속도(nerve conduction velocity, NCV)의 변화를 나타낸 그래프이며, (c)는 복합 근육 활동 전위(compound muscle action potential, CMAP)의 변화를 나타낸 그래프이다(n=7

for mRosa26 RNP; n=10 for PMP22-TATA RNP).

[116] 도 16는 CMT1A 마우스에서 CRISPR/Cas9에 의한 PMP22의 발현저해로 인한 이주 행동(locomotor behaviour)을 분석 결과로, (a)의 상단 그래프는 로타로드 테스트 결과(n=7 for mRosa26 RNP, n=11 for PMP22-TATA)이고, 하단 그래프는 8주령부터 16주령이 될 때까지 매주마다 측정한 로타로드 테스트 결과(n=7 for mRosa26 RNP, n=11 for PMP22-TATA)이며, (b)의 상단 그래프는 mRosa26 또는 PMP22-TATA RNP 복합체를 처리한 C22 마우스의 비장근(gastrocnemius muscle) 무게/체중 비율을 나타낸 그래프이고, 하단 이미지는 mRosa26 또는 PMP22-TATA RNP 복합체를 처리한 C22 마우스의 비장근(gastrocnemius muscle) 이미지이다.

[117] 도 17은 PMD 치료 전략을 도식화한 이미지로, PLP1 유전자의 TATA 박스 영역 및 인핸서 영역을 표적하는 sgRNA를 설계하였다. 인핸서 영역을 표적하는 sgRNAs의 경우, 두 개의 sgRNAs를 이용해 인핸서를 제거하는 전략을 보여주며, 이때, 인핸서 영역의 상단(upstream)을 표적하는 sgRNA는 Up으로, 하단(downstream)을 표적하는 sgRNA는 down으로 표시하였고, 이는 표 5 및 6에도 위치(location)에 따라 Up 및 Down으로 기재하였다.

[118] 도 18은 실시예에 이용된 CjCas9 플라스미드의 도시화한 이미지이다.

[119] 도 19는 mPlp1의 TATA 박스 영역을 표적하는 SpCas9-sgRNA의 스크리닝 결과를 나타낸 그래프이다. (a)는 NIH-3T3 세포에서 확인한 인델(%)을 보여주며, (b)는 N20.1 세포에서 확인한 인델을 보여준다. 이때, 사용된 sgRNA는 mPlp1-TATA-Sp-01을 표적하는 sgRNA이고, 그래프에는 표적 서열의 뒤의 숫자로 표시하였다.

[120] 도 20은 mPlp1의 TATA 박스 영역을 표적하는 CjCas9-sgRNA의 스크리닝 결과를 나타낸 그래프이다. (a)는 NIH-3T3 세포에서 확인한 인델(%)을 보여주며, (b)는 N20.1 세포에서 확인한 인델을 보여준다. 이때, 사용된 sgRNA는 mPlp1-TATA-Cj-01 내지 mPlp1-TATA-Cj-04이고, 그래프에는 표적 서열의 뒤의 숫자로 표시하였다.

[121] 도 21은 mPlp1의 인핸서(wMN1 인핸서) 영역을 표적하는 SpCas9-sgRNA의 스크리닝 결과를 나타낸 그래프이다. (a)는 NIH-3T3 세포에서 확인한 인델(%)을 보여주며, (b)는 N20.1 세포에서 확인한 인델을 보여준다. 이때, 사용된 sgRNA는 mPlp1-wMN1-Sp-01 내지 mPlp1-wMN1-Sp-36이고, 그래프에는 표적 서열의 뒤의 숫자로 표시하였다.

[122] 도 22는 mPlp1의 인핸서(wMN1 인핸서) 영역을 표적하는 CjCas9-sgRNA의 스크리닝 결과를 나타낸 그래프이다. (a)는 NIH-3T3 세포에서 확인한 인델(%)을 보여주며, (b)는 N20.1 세포에서 확인한 인델을 보여준다. 이때, 사용된 sgRNA는 mPlp1-wMN1-Cj-01 내지 mPlp1-wMN1-Sp-28이고, 그래프에는 표적 서열의 뒤의 숫자로 표시하였다.

[123] 도 23은 mPlp1의 TATA 박스 및 인핸서(wMN1 인핸서) 영역을 표적하는

SpCas9-sgRNA 및 CjCas9-sgRNA에 의한 Plp의 mRNA 발현량을 나타낸 그래프이다. (a)는 SpCas9-sgRNA에 의한 Plp의 mRNA 발현량을 확인하였고, 이때, TATA 박스 영역을 표적하는 mPlp1-TATA-Sp-01 및 인핸서를 표적하는 mPlp1-wMN1-Sp-07 + mPlp1-wMN1-Sp-27 및 mPlp1-wMN1-Sp-08 + mPlp1-wMN1-Sp-27을 sgRNA로 사용하였다. (b)는 CjCas9-sgRNA에 의한 Plp의 mRNA 발현량을 확인하였고, 이때, TATA 박스 영역을 표적하는 mPlp1-TATA-Cj-02 및 mPlp1-TATA-Cj-03; 및 인핸서를 표적하는 mPlp1-wMN1-Cj-06 + mPlp1-wMN1-Cj-09, mPlp1-wMN1-Cj-06 + mPlp1-wMN1-Cj-10 및 mPlp1-wMN1-Cj-06 + mPlp1-wMN1-Cj-19를 sgRNA로 사용하였다. mRosa26은 대조군으로 이용하였다.

[124] 도 24는 hPLP1의 인핸서(wMN1 인핸서) 영역을 표적하는 SpCas9-sgRNA의 스크리닝 결과를 나타낸 그래프로, Jurkat 세포에서 확인한 인델(%)를 보여주며, 이때, 사용된 sgRNA는 hPLP1-wMN1-Sp-01 내지 hPLP1-wMN1-Sp-36이고, 그래프에는 표적 서열의 뒤의 숫자로 표시하였다.

[125] 도 25는 hPLP1의 인핸서(wMN1 인핸서) 영역을 표적하는 CjCas9-sgRNA의 스크리닝 결과를 나타낸 그래프로, 293T 세포에서 확인한 인델(%)를 보여주며, 이때, 사용된 sgRNA는 hPLP1-wMN1-Cj-01 내지 hPLP1-wMN1-Cj-36이고, 그래프에는 표적 서열의 뒤의 숫자로 표시하였다.

[126]

발명의 실시를 위한 최선의 형태

[127] 달리 정의되지 않는 한, 본 명세서에서 사용되는 모든 기술적 및 과학적 용어는 본 발명이 속하는 기술분야의 당업자에 의해 통상적으로 이해되는 것과 동일한 의미를 가진다. 본 명세서에 기재된 것과 유사 또는 동일한 방법 및 물질이 본 발명의 실행 또는 시험에서 사용될 수 있지만, 적합한 방법 및 물질이 이하에 기재된다. 본 명세서에 언급된 모든 간행물, 특허 출원, 특허 및 기타 다른 참고문헌은 전체가 참고로 포함된다. 추가로, 물질, 방법 및 실시예는 단지 예시적이며, 제한하는 것으로 의도되지 않는다.

[128]

[129] **본 명세서에 의해 개시되는 내용의 일 태양은 발현 조절용 조성물에 관한 것이다.**

[130] 발현 조절용 조성물은 유전자 중복으로 인한 중복 유전자의 발현을 조절하기 위한 조성물이다.

[131] "유전자 중복(gene duplication)"은 게놈 내 동일한 유전자가 2개 또는 그 이상 존재하는 것을 의미하며, 또한, 유전자 중복은 게놈 내 동일한 유전자의 일부가 2개 또는 그 이상 존재하는 것도 모두 포함한다. 유전자 중복은, 예를 들어, A 유전자가 게놈 내 A 유전자 전장이 2개 또는 그 이상 존재하는 것이거나, 또는 게놈 내 하나의 A 유전자 전장 및 하나 이상의 A 유전자의 일부, 예로, 엑손 1이

존재하는 것일 수 있다. 예를 들어, 유전자 중복은 게놈 내에 2개의 B 유전자 전장 및 하나 이상의 B 유전자의 일부, 예로, 엑손 1와 엑손 2이 존재하는 것일 수 있다. 유전자 중복의 형태는 다양할 수 있으며, 유전자 중복은 게놈 내에 유전자 전장 및/또는 유전자의 일부 서열이 중복(즉, 2개 이상)된 형태를 모두 포함한다.

[132] 또한, 유전자 중복은 염색체의 유전자 재조합 과정 중 발생하는 염색체의 일부 구간이 중복되는 복제 현상도 포함한다. 이러한 유전자 중복은 유전자 돌연변이의 한 종류이며, 다음 세대로 유전된다. 유전자 중복은 염색체의 일부 구간이 복제되지 않아 일어나는 유전자 결실과 함께 유전자 발현에 영향을 준다.

[133] 이때, 유전자 중복의 대상, 즉, 2개 또는 그 이상 존재하는 유전자는 "중복 유전자(duplication gene)"로 지칭된다.

[134] 상기 중복 유전자는 유전자 중복으로 인해 게놈 내 총 카피 수가 증가된 유전자일 수 있다.

[135] 상기 중복 유전자는 유전자 중복으로 인해 일부 영역만 중복된 돌연변이 유전자일 수 있다. 이때, 돌연변이 유전자는 유전자 중복으로 인해 유전자 전체 서열 중 하나 이상의 뉴클레오타이드 서열이 중복된 유전자일 수 있다. 또는 돌연변이 유전자는 유전자 중복으로 인해 유전자의 일부 핵산 조각(fragment)이 중복된 유전자일 수 있다. 이때, 상기 핵산 조각은 50bp 이상의 뉴클레오타이드 서열로 이루어진 것일 수 있다.

[136] 유전자 중복은 전체 게놈 중복(whole genome duplication)을 포함한다.

[137] 유전자 중복은 표적 유전자 중복(target gene duplication)을 포함한다. 이때, 표적 유전자 중복은 환경 변화에 따른 새로운 종이 분화되고 적등되는 과정 중에 특정 환경에 적합하도록 해당 유전자가 증폭되거나 사라지는 식의 유전자 중복으로, 대부분 transposon에 의한 복제가 진행된다.

[138] 유전자 중복은 이소성 재조합(ectopic recombination)을 포함한다. 이때, 이소성 재조합은 상동염색체의 감수분열 중 발생하는 불균등한 교차로 인해 발생되는 복제로 두 염색체 사이의 반복 염기서열(repeat sequence)의 정도에 따라 발생된다. 교차점에서의 중복과 상호간 삭제(reciprocal deletion)가 형성된다. 이소성 재조합은 전이성 유전인자(Transposable element)와 같은 전형적인 반복적인 유전인자(repetitive genetic element)에 의해 매개되는 것으로 재조합에 의한 복제로 진행된다.

[139] 유전자 중복은 복제 미끄러짐(replication slippage)을 포함한다. 이때, 복제 미끄러짐은 DNA 복제 과정의 오류로 인해 짧은 유전적 서열(genetic sequence)이 복제되는 것으로, DNA 폴리머라아제가 해리된 DNA 가닥에 잘못 부착되면서 다시 복제되어 발생된다. 복제 미끄러짐도 반복적인 유전인자(repetitive genetic element)에 의해 매개되는 경우가 많다.

[140] 유전자 중복은 레트로트랜스포지션(retrotransposition)을 포함한다. 이때, 레트로트랜스포지션은 세포에 침입된 레트로바이러스 또는 레트로인자(retroelement)에 의해 복제 되는 것으로, 레트로유전자(retrogene)를

만들어내기 위해 유전자의 역전사를 수행되고 이들의 재조합으로 인해 유전자의 복제가 수행된다. 레트로트랜스포지션은 레트로전이성 유전인자(retrotransposable element)와 같은 유전인자(genetic element)에 의해 매개된다.

[141]

[142] 유전자 중복은 중복 유전자가 전사된 mRNA의 발현을 증가시킬 수 있다. 이때, 전사된 mRNA의 발현은 유전자 중복이 존재하지 않는 상태에 비해 증가할 수 있다.

[143] 유전자 중복은 중복 유전자가 암호화하는 단백질의 발현을 증가시킬 수 있다. 이때, 단백질의 발현은 유전자 중복이 존재하지 않는 상태에 비해 증가할 수 있다.

[144] 유전자 중복은 중복 유전자가 암호화하는 단백질의 기능이상을 초래할 수 있다.

[145] 이때, 상기 기능이상은 단백질의 정상 기능의 과잉, 저해 및 제 3의 기능일 수 있다.

[146]

[147] 유전자 중복은 유전자 중복 질환을 야기할 수 있다.

[148] "유전자 중복 질환(gene duplication disease)"는 유전자 중복에 발생하는 질환으로, 중복 유전자의 비정상적인 증폭에 의해 유전적 이상을 초래하고, 이로 인해 과발현된 단백질 또는 비정상적으로 생성된 단백질로 인해 병리학적 특징을 야기하는 모든 질환 또는 질병을 포함한다. 이때, "병리학적 특징"은 질병으로 인한 유기체의 세포 수준 및 조직, 장기, 개체 수준에서의 변화를 의미한다.

[149] 유전자 중복 질환은 샤르코마리투스병 1A형(Charcot-Marie-Tooth 1A, CMT1A), 데제린-소타스병(Dejerine-Sottas disease, DSD), 선천성 저수초형성 신경병(Congenital Hypomyelination Neuropathy, CHN), 루시-레비 증후군(Roussy-Levy syndrome, RLS), 펠리제우스-메르츠바하병(Pelizaeus-Merzbacher disease, PMD), MECP2 중복 증후군(MECP2 duplication syndrome), X-연관 뇌하수체기능저하증(X-linked hypopituitarism, XLHP), Potocki-Lupski 증후군(Potocki-Lupski syndrome, PTLS), 구개심장안면증후군(Velocardiofacial syndrome, VCFS), 윌리엄스 보이렌 증후군(Williams Beuren syndrome, WBS), 알라질증후군(Alagille syndrome, AS), 발육지연 증후군(Growth retardation syndrome), 조기 폐쇄 두개봉합(Premature closure cranial sutures), 보통염색체우성 백질이영양증(Autosomal dominant leukodystrophy, ADLD), 파킨슨병(Parkinson disease) 또는 알츠하이머(Alzheimer disease)일 수 있다.

[150] 유전자 중복 질환은 암유전자(oncogene) 중복에 의한 암일 수 있다.

[151] 이때, 상기 암은 유방암(Breast cancer), 자궁경부암(Cervical cancer),

대장암(Colorectal cancer), 식도암(Esophageal cancer), 위암(Gastric cancer), 교모세포종(Glioblastoma), 두경부암(Head and neck cancer), 간암(Hepatocellular cancer), 신경아세포종(Neuroblastoma), 난소암(Ovarian cancer), 육종(Sarcoma) 또는 소세포폐암(Small cell lung cancer)일 수 있다.

[152] 유전자 중복 질환은 PMP22 유전자, PLP1 유전자, MECP2 유전자, SOX3 유전자, RAI1 유전자, TBX1 유전자, ELN 유전자, JAGGED1 유전자, NSD1 유전자, MMP23 유전자, LMB1 유전자, SNCA 유전자 또는 APP 유전자의 중복에 의해 발생한 질환일 수 있다.

[153] 유전자 중복 질환은 MYC 유전자, ERBB2(HER2) 유전자, CCND1(Cyclin D1) 유전자, FGFR1 유전자, FGFR2 유전자, HRAS 유전자, KRAS 유전자, MYB 유전자, MDM2 유전자, CCNE(Cyclin E) 유전자, MET 유전자, CDK4 유전자, ERBB1 유전자, MYCN 유전자 또는 AKT2 유전자의 중복에 의해 발생한 질환일 수 있다.

[154] 유전자 중복 질환은 중복 유전자의 전사된 mRNA의 발현이 비정상적으로 증가함에 의해 발생한 질환일 수 있다.

[155] 유전자 중복 질환은 중복 유전자가 암호화하는 단백질의 발현이 비정상적으로 증가함에 의해 발생한 질환일 수 있다.

[156]

[157] 발현 조절용 조성물은 중복 유전자의 전사에 의해 생성되는 mRNA의 발현을 조절에 이용될 수 있다.

[158] 발현 조절용 조성물은 중복 유전자가 암호화하는 단백질의 발현을 조절에 이용될 수 있다.

[159] 발현 조절용 조성물은 중복 유전자의 인위적인 조작 또는 변형을 위해 이용될 수 있다.

[160] 이때, “인위적인 조작 또는 변형(artificially modified or engineered or artificially engineered)”이라는 용어는 자연상태에서 일어나는 존재 그대로의 상태가 아닌, 인위적으로 변형을 가한 상태를 의미한다. 이하에서, 비자연적인 인위적으로 조작된 또는 변형된 중복 유전자는 인위적인 중복 유전자라는 용어와 혼용되어 사용될 수 있다.

[161] “발현 조절 시스템”은 인위적으로 조작된 중복 유전자의 발현 조절에 의해 발생되는 모든 현상을 포함하는 용어로, 이러한 발현 조절 시스템에 직접적으로 또는 간접적으로 관여하는 모든 물질, 조성물, 방법 및 용도를 포함한다.

[162] 발현 조절용 조성물은 중복 유전자의 전사 조절 영역의 인위적인 조작 또는 변형을 위해 이용될 수 있다.

[163] 이때, “전사 조절 영역(transcription control region)”은 유전자의 DNA를 기반으로 RNA가 합성되는 전반적인 과정을 조절하는 영역으로, 유전자의 DNA 서열 내에 및/또는 유전자의 근접한 DNA 서열 중 전사 인자(transcription factor)와 상호작용하는 모든 영역을 포함한다. 이때, 전사 인자는 활성화되면

유전자 근처의 반응요소(response element)라는 DNA 특정 부위에 결합하여 유전자의 발현을 촉진 또는 억제하는 단백질을 말하며, 상기 반응요소는 전사 조절 영역에 포함된다. 전자 조절 영역은 유전자에 따라 종류와 위치가 다양하며, 동일한 종이라도 개체간 핵산 서열에 차이가 있을 수 있다.

[164] 전사 조절 영역은 프로모터, 인핸서, 사일런서, 인슐레이터 및/또는 로커스 컨트롤 영역(locus control region, LCR)일 수 있다.

[165] 상기 프로모터는 코어 프로모터(core promoter), 근접 프로모터(proximal promoter) 및/또는 원위 프로모터(Distal promoter)일 수 있다.

[166] 이때, 코어 프로모터는 전사 시작 사이트(transcription start site, TSS), RNA 폴리머라아제(RNA polymerase) 결합 사이트, 전사인자 결합 사이트 및/또는 TATA 박스를 포함할 수 있다.

[167] 상기 TATA 박스는 Rpb4/Rbp7이 전사를 개시하기 위해 사용되는 시작 부위의 25 염기쌍 상류(upsteam)에 위치하는 영역일 수 있다.

[168] 상기 TATA 박스는 전사 시작 사이트(TSS)의 30 염기쌍 상류(upsteam)에 위치하는 영역일 수 있다.

[169] 상기 TATA 박스는 전사 시작 사이트(TSS)의 40 내지 100 염기쌍 상류(upsteam)에 위치하는 영역일 수 있다.

[170] 예를 들어, 상기 TATA 박스는 프로모터 및/또는 코어 프로모터 내에 존재하는 5'-TATA(A/T)A(A/T)-3' 서열을 포함하는 영역일 수 있다. 또는 상기 TATA 박스는 프로모터 및/또는 코어 프로모터 내에 존재하는 5'-TATA(A/T)A(A/T)(A/G)-3' 서열을 포함하는 영역일 수 있다.

[171] 예를 들어, 상기 TATA 박스는 프로모터 및/또는 코어 프로모터 내에 존재하는 5'-CATAAAA-3' 서열, 5'-CATAAAA-3' 서열, 5'-TATAA-3' 서열, 5'-TATAAAA-3' 서열, 5'-CATAAATA-3' 서열, 5'-TATATAA-3' 서열, 5'-TATATATATATATAA-3' 서열, 5'-TATATTATA-3' 서열, 5'-TATAAA-3' 서열, 5'-TATAAAATA-3' 서열, 5'-TATATA-3' 서열, 5'-GATTAAAAA-3' 서열, 5'-TATAAAAA-3' 서열, 5'-TTATAA-3' 서열, 5'-TTTTAAAA-3' 서열, 5'-TCTTTAAAA-3' 서열, 5'-GACATTTAA-3' 서열, 5'-TGATATCAA-3' 서열, 5'-TATAAATA-3' 서열, 5'-TATAAGA-3' 서열, 5'-AATAAA-3' 서열, 5'-TTTATA-3' 서열, 5'-CATAAAAA-3' 서열, 5'-TATACA-3' 서열, 5'-TTTAAGA-3' 서열, 5'-GATAAAG-3' 서열, 5'-TATAACA-3' 서열, 5'-TCTTATCTT-3' 서열, 5'-TTGTACTTT-3' 서열, 5'-CATATAA-3' 서열, 5'-TATAAAT-3' 서열, 5'-TATATATAAAAAAAA-3' 서열 및 5'-CATAAATAAAAAAAATTA-3' 서열로 구성된 군에서 선택된 하나 이상의 서열을 포함하는 영역일 수 있다.

[172] 예를 들어, 상기 TATA 박스는 프로모터 및/또는 코어 프로모터 내에 존재하는 TATA-결합 단백질(TATA-binding protein, TBP)이 결합하는 영역일 수 있다.

[173] 이때, 상기 근접 프로모터는 전사 시작 사이트의 1 내지 300bp 상류(upstream) 영역, CpG 사이트 및/또는 특정 전사인자 결합 사이트를 포함할 수 있다.

[174] 상기 인핸서는 인핸서-박스(enhancer-box, E-box)를 포함할 수 있다.

[175] 상기 인슐레이터는 인핸서와 프로모터 사이의 상호작용을 억제하거나 또는 억제된 상태의 크로마틴이 확장되는 것을 방지하는 영역일 수 있다.

[176] 상기 로커스 컨트롤 영역(locus control region, LCR)은 인핸서, 사일런서, 인슐레이터, MAR, SAR 등과 같은 다수의 cis-acting 요소들이 존재하는 영역일 수 있다.

[177]

[178] **본 명세서에 의해 개시되는 내용의 일 태양으로, 발현 조절용 조성물은 가이드핵산을 포함할 수 있다.**

[179] 발현 조절용 조성물은 중복 유전자를 표적하는 가이드핵산 또는 이를 암호화하는 핵산 서열을 포함할 수 있다.

[180] "가이드핵산"은 표적 핵산, 유전자 또는 염색체를 인지하고, 및 에디터단백질과 상호작용할 수 있는 뉴클레오타이드 서열을 의미한다. 이때, 상기 가이드핵산은 표적 핵산, 유전자 또는 염색체 내의 일부 뉴클레오타이드 서열과 상보적인 결합할 수 있다. 또한 상기 가이드핵산의 일부 뉴클레오타이드 서열은 에디터단백질의 일부 아미노산과 상호작용하여 가이드핵산-에디터단백질 복합체를 형성할 수 있다.

[181] 상기 가이드핵산은 가이드핵산-에디터단백질 복합체가 표적 핵산, 유전자 또는 염색체의 표적 영역에 위치할 수 있도록 유도하는 기능을 수행할 수 있다.

[182] 상기 가이드핵산은 DNA, RNA 또는 DNA/RNA 혼합의 형태일 수 있고, 5 내지 150개의 핵산서열을 가질 수 있다.

[183] 상기 가이드핵산은 하나의 연속된 핵산서열일 수 있다.

[184] 예를 들어, 하나의 연속된 핵산 서열은 $(N)_m$ 일 수 있고, 이때 N은 A, T, C 또는 G, 또는 A, U, C 또는 G이며, m은 1 내지 150의 정수를 의미한다.

[185] 상기 가이드핵산은 연속된 핵산서열이 두 개 이상일 수 있다.

[186] 예를 들어, 두 개 이상의 연속된 핵산서열은 $(N)_m$과 $(N)_o$ 일 수 있고, 이때 N은 A, T, C 또는 G, 또는 A, U, C 또는 G이며, m 및 o는 1 내지 150의 정수를 의미하며, m과 o는 서로 같거나 다를 수 있다.

[187] 상기 가이드핵산은 하나 이상의 도메인을 포함할 수 있다.

[188] 상기 도메인은 가이드 도메인, 제 1 상보적 도메인, 연결 도메인, 제 2 상보적 도메인, 근위 도메인, 꼬리 도메인 등의 기능적 도메인일 수 있으며, 이에 제한되지 않는다.

[189] 이때, 하나의 가이드핵산은 2 이상의 기능적 도메인을 가질 수 있다. 이때, 상기 2 이상의 기능적 도메인은 서로 상이할 수 있다. 예를 들어, 하나의 가이드핵산은 가이드 도메인 및 제 1 상보적 도메인을 가질 수 있고, 다른 예를 들어, 하나의 가이드핵산은 제 2 상보적 도메인, 근위 도메인 및 꼬리 도메인을 가질 수 있으며, 또 다른 예를 들어, 하나의 가이드핵산은 가이드 도메인, 제 1 상보적 도메인, 연결 도메인, 제 2 상보적 도메인, 근위 도메인 및 꼬리 도메인을 가질 수

있다. 또는 하나의 가이드핵산에 포함된 2 이상의 기능적 도메인은 서로 동일할 수도 있다. 예를 들어, 하나의 가이드핵산은 2 이상의 근위 도메인을 가질 수 있고, 다른 예를 들어, 하나의 가이드핵산은 2 이상의 꼬리도메인을 가질 수 있다. 다만, 하나의 가이드핵산에 포함되어 있는 기능적 도메인이 서로 동일한 도메인이라는 말은 두 기능적 도메인의 시퀀스가 동일하다는 의미는 아니며, 시퀀스가 상이하여도 기능적으로 동일한 기능을 수행하고 있으면 동일한 도메인이라고 할 수 있다.

[190] 기능적 도메인에 대해서 이하에서 구체적으로 설명한다.

[191]

[192] **i) 가이드 도메인**

[193] "가이드 도메인"은 표적 유전자의 전사 조절 영역 내 핵산의 이중 가닥 중 어느 하나 가닥의 일부 서열에 상보적인 결합을 할 수 있는 도메인으로, 표적 유전자의 전사 조절 영역 내 핵산과의 특이적인 상호작용을 위해 역할한다. 예를 들어, 가이드 도메인은 표적 유전자의 전사 조절 영역 내 핵산의 특정 뉴클레오타이드 서열을 가지는 위치로 가이드핵산-에디터단백질 복합체를 유도하는 기능을 수행할 수 있다.

[194] 상기 가이드 도메인은 10 내지 35개의 뉴클레오타이드 서열일 수 있다.

[195] 일 예로, 상기 가이드 도메인은 10 내지 35개의 뉴클레오타이드 서열, 15 내지 35개의 뉴클레오타이드서열, 20 내지 35개의 뉴클레오타이드 서열, 25 내지 35개의 뉴클레오타이드 서열 또는 30 내지 35개의 뉴클레오타이드 서열일 수 있다.

[196] 다른 일 예로, 상기 가이드 도메인은 10 내지 15개의 뉴클레오타이드 서열, 15 내지 20개의 뉴클레오타이드 서열, 20 내지 25개의 뉴클레오타이드 서열, 25 내지 30개의 뉴클레오타이드 서열 또는 30 내지 35개의 뉴클레오타이드 서열일 수 있다.

[197] 상기 가이드 도메인은 가이드 서열을 포함할 수 있다.

[198] "가이드 서열"은 표적 유전자의 전사 조절 영역 내 핵산의 이중 가닥 중 어느 하나 가닥의 일부 서열에 상보적인 뉴클레오타이드 서열이며, 이때 상기 가이드 서열은 최소한 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90% 또는 95% 이상의 상보성을 가지거나 또는 완전한 상보성을 가지는 뉴클레오타이드 서열일 수 있다.

[199] 상기 가이드 서열은 10 내지 25개의 뉴클레오타이드 서열일 수 있다.

[200] 일 예로, 상기 가이드 서열은 10 내지 25개의 뉴클레오타이드 서열, 15 내지 25개의 뉴클레오타이드 서열 또는 20 내지 25개의 뉴클레오타이드 서열일 수 있다.

[201] 다른 일 예로, 상기 가이드 서열은 10 내지 15개의 뉴클레오타이드 서열, 15 내지 20개의 뉴클레오타이드 서열 또는 20 내지 25개의 뉴클레오타이드 서열일 수 있다.

[202] 또한, 상기 가이드 도메인은 추가 뉴클레오타이드 서열을 더 포함할 수 있다.

[203] 상기 추가 뉴클레오타이드 서열은 가이드 도메인의 기능 향상 또는 저하를 위한 것일 수 있다.

[204] 상기 추가 뉴클레오타이드 서열은 가이드 서열의 기능 향상 또는 저하를 위한 것일 수 있다.

[205] 상기 추가 뉴클레오타이드 서열은 1 내지 10개의 뉴클레오타이드 서열일 수 있다.

[206] 일 예로, 상기 추가 뉴클레오타이드 서열은 2내지 10개의 뉴클레오타이드 서열, 4 내지 10개의 뉴클레오타이드 서열, 6 내지 10개의 뉴클레오타이드 서열 또는 8 내지 10개의 뉴클레오타이드 서열일 수 있다.

[207] 다른 일 예로, 상기 추가 뉴클레오타이드 서열은 1 내지 3개의 뉴클레오타이드 서열, 3 내지 6개의 뉴클레오타이드 서열 또는 7 내지 10개의 뉴클레오타이드 서열일 수 있다.

[208] 일 구체예로서, 상기 추가 뉴클레오타이드 서열은 1개의 뉴클레오타이드 서열, 2개의 뉴클레오타이드 서열, 3개의 뉴클레오타이드 서열, 4개의 뉴클레오타이드 서열, 5개의 뉴클레오타이드 서열, 6개의 뉴클레오타이드 서열, 7개의 뉴클레오타이드 서열, 8개의 뉴클레오타이드 서열, 9개의 뉴클레오타이드 서열 또는 10개의 뉴클레오타이드 서열일 수 있다.

[209] 예를 들어, 상기 추가 뉴클레오타이드 서열은 1개의 뉴클레오타이드 서열 G(구아닌)일 수 있으며, 또는 2개의 뉴클레오타이드 서열 GG일 수 있다.

[210] 상기 추가 뉴클레오타이드 서열은 상기 가이드 서열의 5' 말단에 위치할 수 있다.

[211] 상기 추가 뉴클레오타이드 서열은 상기 가이드 서열의 3' 말단에 위치할 수 있다.

[212]

[213] **ii) 제 1 상보적 도메인**

[214] "제 1 상보적 도메인"은 이하에서 설명하는 제 2 상보적 도메인에 대해 상보적 뉴클레오타이드 서열을 포함하는 도메인으로, 제 2 상보적 도메인과 이중가닥을 형성할 수 있을 정도로 상보성을 가진다. 예를 들어, 제 1 상보적 도메인은 제 2 상보적 도메인에 대해 최소한 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90% 또는 95% 이상의 상보성을 가지거나 또는 완전한 상보성을 가지는 뉴클레오타이드 서열일 수 있다.

[215] 상기 제 1 상보적 도메인은 제 2 상보적 도메인과 상보적 결합을 통해 이중가닥을 형성할 수 있다. 이때, 상기 형성된 이중가닥은 에디터단백질의 일부 아미노산과 상호작용하여 가이드핵산-에디터단백질 복합체를 형성하는 역할을 할 수 있다.

[216] 상기 제 1 상보적 도메인은 5 내지 35개의 뉴클레오타이드 서열일 수 있다.

[217] 일 예로, 제 1 상보적 도메인은 5 내지 35개의 뉴클레오타이드 서열, 10 내지

35개의 뉴클레오타이드 서열, 15 내지 35개의 뉴클레오타이드 서열, 20 내지 35개의 뉴클레오타이드 서열, 25 내지 35개의 뉴클레오타이드 서열 또는 30 내지 35개의 뉴클레오타이드 서열일 수 있다.

[218] 다른 일 예로서, 상기 제 1 상보적 도메인은 1 내지 5개의 뉴클레오타이드 서열, 5 내지 10개의 뉴클레오타이드 서열, 10 내지 15개의 뉴클레오타이드 서열, 15 내지 20개의 뉴클레오타이드 서열, 20 내지 25개의 뉴클레오타이드 서열, 25 내지 30개의 뉴클레오타이드 서열 또는 30 내지 35개의 뉴클레오타이드 서열일 수 있다.

[219]

[220] **iii) 연결 도메인**

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

[222] 상기 연결 도메인은 1 내지 30개의 뉴클레오타이드 서열일 수 있다.

[223] 일 예로서, 상기 연결 도메인은 1 내지 5개의 뉴클레오타이드 서열, 5 내지 10개의 뉴클레오타이드 서열, 10 내지 15개의 뉴클레오타이드 서열, 15 내지 20개의 뉴클레오타이드 서열, 20 내지 25개의 뉴클레오타이드 서열 또는 25 내지 30개의 뉴클레오타이드 서열일 수 있다.

[224] 다른 일 예로서, 상기 연결 도메인은 1 내지 30개의 뉴클레오타이드 서열, 5 내지 30개의 뉴클레오타이드 서열, 10 내지 30개의 뉴클레오타이드 서열, 15 내지 30개의 뉴클레오타이드 서열, 20 내지 30개의 뉴클레오타이드 서열 또는 25 내지 30개의 뉴클레오타이드 서열일 수 있다.

[225]

[226] **iv) 제 2 상보적 도메인**

[227] "제 2 상보적 도메인"은 전술한 제 1 상보적 도메인과 상보적 핵산서열을 포함하는 뉴클레오타이드 서열을 포함하는 도메인으로, 제 1 상보적 도메인과 이중가닥을 형성할 수 있을 정도로 상보성을 가진다. 예를 들어, 제 2 상보적 도메인은 제 1 상보적 도메인에 대해 최소한 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90% 또는 95% 이상의 상보성을 가지거나 또는 완전한 상보성을 가지는 뉴클레오타이드 서열일 수 있다.

[228] 상기 제 2 상보적 도메인은 제 1 상보적 도메인과 상보적 결합을 통해 이중가닥을 형성할 수 있다. 이때, 상기 형성된 이중가닥은 에디터단백질의 일부 아미노산과 상호작용하여 가이드핵산-에디터단백질 복합체를 형성하는 역할을 할 수 있다.

[229] 제 2 상보적 도메인은 제 1 상보적 도메인과 상보적 뉴클레오타이드 서열 및 제 1 상보적 도메인과의 상보성이 없는 뉴클레오타이드 서열, 예를 들어, 제 1 상보적 도메인과 이중가닥을 형성하지 않는 뉴클레오타이드 서열을 포함할 수

있으며, 제 1 상보적 도메인보다 뉴클레오타이드 서열의 길이가 길 수 있다.

[230] 상기 제 2 상보적 도메인은 5 내지 35개의 뉴클레오타이드 서열일 수 있다.

[231] 일 예로, 상기 제 2 상보적 도메인은 1 내지 35개의 뉴클레오타이드 서열, 5 내지 35개의 뉴클레오타이드 서열, 10 내지 35개의 뉴클레오타이드 서열, 15 내지 35개의 뉴클레오타이드서열, 20 내지 35개의 뉴클레오타이드 서열, 25 내지 35개의 뉴클레오타이드 서열 또는 30 내지 35 뉴클레오타이드 서열일 수 있다.

[232] 다른 일 예로, 상기 제 2 상보적 도메인은 1 내지 5개의 뉴클레오타이드 서열, 5 내지 10개의 뉴클레오타이드 서열, 10 내지 15개의 뉴클레오타이드 서열, 15 내지 20개의 뉴클레오타이드 서열, 20 내지 25개의 뉴클레오타이드 서열, 25 내지 30개의 뉴클레오타이드 서열 또는 30 내지 35개의 뉴클레오타이드 서열일 수 있다.

[233]

[234] **v) 근위 도메인(proximal domain)**

[235] "근위 도메인"은 제 2 상보적 도메인에 근접하게 위치하는 뉴클레오타이드 서열이다.

[236] 근위 도메인은 근위 도메인 내의 상보적인 뉴클레오타이드 서열을 포함할 수 있으며, 상보적인 뉴클레오타이드 서열에 의해 이중가닥을 형성할 수 있다.

[237] 상기 근위 도메인은 1 내지 20개의 뉴클레오타이드 서열일 수 있다.

[238] 일 예로서, 상기 근위 도메인은 1 내지 20개의 뉴클레오타이드 서열, 5 내지 20개의 뉴클레오타이드 서열, 10 내지 20개의 뉴클레오타이드 서열 또는 15 내지 20개의 뉴클레오타이드 서열일 수 있다.

[239] 다른 일 예로서, 상기 근위 도메인은 1 내지 5개의 뉴클레오타이드 서열, 5 내지 10개의 뉴클레오타이드 서열, 10 내지 15개의 뉴클레오타이드 서열 또는 15 내지 20개의 뉴클레오타이드 서열일 수 있다.

[240]

[241] **vi) 꼬리 도메인**

[242] "꼬리 도메인"은 가이드핵산의 양 말단 중 어느 하나 이상의 말단에 위치하는 뉴클레오타이드 서열이다.

[243] 꼬리 도메인은 꼬리 도메인 내의 상보적인 뉴클레오타이드 서열을 포함할 수 있으며, 상보적인 뉴클레오타이드 서열에 의해 이중가닥을 형성할 수 있다.

[244] 상기 꼬리 도메인은 1 내지 50개의 뉴클레오타이드 서열일 수 있다.

[245] 일 예로, 상기 꼬리 도메인은 5 내지 50개의 뉴클레오타이드 서열, 10 내지 50개의 뉴클레오타이드 서열, 15 내지 50개의 뉴클레오타이드 서열, 20 내지 50개의 뉴클레오타이드 서열, 25 내지 50개의 뉴클레오타이드 서열, 30 내지 50개의 뉴클레오타이드 서열, 35 내지 50개의 뉴클레오타이드 서열, 40 내지 50개의 뉴클레오타이드 서열 또는 45 내지 50개의 뉴클레오타이드 서열일 수 있다.

[246] 다른 일 예로, 상기 꼬리 도메인은 1 내지 5개의 뉴클레오타이드 서열, 5 내지

10개의 뉴클레오타이드 서열, 10 내지 15개의 뉴클레오타이드 서열, 15 내지 20개의 뉴클레오타이드 서열, 20 내지 25개의 뉴클레오타이드 서열, 25 내지 30개의 뉴클레오타이드 서열, 30 내지 35개의 뉴클레오타이드 서열, 35 내지 40개의 뉴클레오타이드 서열, 40 내지 45개의 뉴클레오타이드 서열 또는 45 내지 50개의 뉴클레오타이드 서열일 수 있다.

[247]

[248] 한편, 상기 도메인들, 즉, 가이드 도메인, 제 1 상보적 도메인, 연결 도메인, 제 2 상보적 도메인, 근위 도메인 및 꼬리 도메인이 포함하는 핵산 서열의 일부 또는 전부는 선택적 또는 추가적으로 화학적 변형을 포함할 수 있다.

[249] 상기 화학적 변형은 methylation, acetylation, phosphorylation, phosphorothioate linkage, locked nucleic acid(LNA), 2'-O-methyl 3'phosphorothioate(MS) 또는 2'-O-methyl 3'thioPACE(MSP)일 수 있으나, 이에 제한되지 않는다.

[250]

[251] 가이드핵산은 하나 이상의 도메인을 포함한다.

[252] 상기 가이드핵산은 가이드 도메인을 포함할 수 있다.

[253] 상기 가이드핵산은 제 1 상보적 도메인을 포함할 수 있다.

[254] 상기 가이드핵산은 연결 도메인을 포함할 수 있다.

[255] 상기 가이드핵산은 제 2 상보적 도메인을 포함할 수 있다.

[256] 상기 가이드핵산은 근위 도메인을 포함할 수 있다.

[257] 상기 가이드핵산은 꼬리 도메인을 포함할 수 있다.

[258] 이때, 상기 도메인의 개수는 1, 2, 3, 4, 5, 6 또는 그 이상일 수 있다.

[259] 상기 가이드핵산은 1, 2, 3, 4, 5, 6 또는 그 이상의 가이드 도메인을 포함할 수 있다.

[260] 상기 가이드핵산은 1, 2, 3, 4, 5, 6 또는 그 이상의 제 1 상보적 도메인을 포함할 수 있다.

[261] 상기 가이드핵산은 1, 2, 3, 4, 5, 6 또는 그 이상의 연결 도메인을 포함할 수 있다.

[262] 상기 가이드핵산은 1, 2, 3, 4, 5, 6 또는 그 이상의 제 2 상보적 도메인을 포함할 수 있다.

[263] 상기 가이드핵산은 1, 2, 3, 4, 5, 6 또는 그 이상의 근위 도메인을 포함할 수 있다.

[264] 상기 가이드핵산은 1, 2, 3, 4, 5, 6 또는 그 이상의 꼬리 도메인을 포함할 수 있다.

[265]

[266] 이때, 상기 가이드핵산은 하나의 도메인이 중복되어 포함될 수 있다.

[267] 상기 가이드핵산은 여러 도메인을 중복 또는 중복시키지 않고 포함할 수 있다.

[268] 상기 가이드핵산은 같은 종류의 도메인을 포함할 수 있으며, 이때, 같은 종류의 도메인은 동일한 핵산서열을 가지거나 또는 서로 다른 핵산서열을 가질 수 있다.

[269] 상기 가이드핵산은 두 종류의 도메인을 포함할 수 있으며, 이때, 다른 두 종류의 도메인은 서로 다른 핵산서열을 가지거나 또는 동일한 핵산서열을 가질 수 있다.

[270] 상기 가이드핵산은 세 종류의 도메인을 포함할 수 있으며, 이때, 다른 세 종류의

도메인은 서로 다른 핵산서열을 가지거나 또는 동일한 핵산서열을 가질 수 있다.

[271] 상기 가이드핵산은 네 종류의 도메인을 포함할 수 있으며, 이때, 다른 네 종류의 도메인은 서로 다른 핵산서열을 가지거나 또는 동일한 핵산서열을 가질 수 있다.

[272] 상기 가이드핵산은 다섯 종류의 도메인을 포함할 수 있으며, 이때, 다른 다섯 종류의 도메인은 서로 다른 핵산서열을 가지거나 또는 동일한 핵산서열을 가질 수 있다.

[273] 상기 가이드핵산은 여섯 종류의 도메인을 포함할 수 있으며, 이때, 다른 여섯 종류의 도메인은 서로 다른 핵산서열을 가지거나 또는 동일한 핵산서열을 가질 수 있다.

[274] 예를 들면, 가이드핵산은 [가이드 도메인]-[제 1 상보적 도메인]-[연결 도메인]-[제 2 상보적 도메인]-[연결 도메인]-[가이드 도메인]-[제 1 상보적 도메인]-[연결 도메인]-[제 2 상보적 도메인]으로 구성될 수 있으며, 이때, 두 개의 가이드 도메인은 서로 다른 또는 동일한 표적을 위한 가이드 서열을 포함할 수 있으며, 상기 두 개의 제 1 상보적 도메인과 두 개의 제 2 상보적 도메인 동일한 핵산서열을 가지거나 다른 핵산서열을 가질 수 있다. 가이드 도메인이 서로 다른 표적을 위한 가이드 서열을 포함하는 경우, 상기 가이드핵산은 두 개의 표적에 특이적으로 결합할 수 있으며, 이때, 특이적 결합을 동시에 일어나거나 순차적으로 일어날 수 있다. 또한, 상기 연결 도메인은 특정 효소에 의해 절단될 수 있으며, 특정 효소의 존재 하에서 상기 가이드핵산은 두 부분 또는 세 부분으로 나누어질 수 있다.

[275]

[276] **본 명세서에 의해 개시되는 내용의 일 구체예로서, 가이드핵산은 gRNA일 수 있다.**

[277] "gRNA"는 표적 유전자의 전사 조절 영역 내 핵산에 대한 gRNA-CRISPR 효소 복합체, 즉, CRISPR 복합체를 특이적 표적화할 수 있는 RNA를 지칭한다. 또한, 상기 gRNA는 표적 유전자의 전사 조절 영역 내 핵산 특이적 RNA를 의미하며, CRISPR 효소과 결합하여 CRISPR 효소를 표적 유전자의 전사 조절 영역 내 핵산으로 인도할 수 있다.

[278] 상기 gRNA는 다수의 도메인을 포함할 수 있다. 각각의 도메인에 의해 3차원 행태 또는 gRNA의 활성 형태의 가닥내 또는 가닥간 상호작용을 할 수 있다.

[279] gRNA는 단일가닥 gRNA(단일 RNA 분자; single gRNA; sgRNA); 또는 이중 gRNA(하나 초과의 통상적으로 2개의 별개의 RNA 분자를 포함함)로서 지칭될 수 있다.

[280] 일 구체예에서, 단일가닥 gRNA는 5'으로부터 3' 방향으로 가이드 도메인, 즉 표적 유전자의 전사 조절 영역 내 핵산에 상보적인 결합을 할 수 있는 가이드 서열(guide sequence)를 포함하는 도메인; 제 1 상보적 도메인; 연결 도메인; 제 2 상보적 도메인, 상기 제 1 상보적 도메인 서열에 상보적인 서열을 가지므로 제 1 상보적 도메인과 이중가닥 핵산을 형성할 수 있는 도메인; 근위 도메인(proximal

domain); 및 선택적으로 꼬리 도메인을 포함할 수 있다.

[281] 다른 일 구체예로서, 이중 gRNA는 5'으로부터 3' 방향으로 가이드 도메인, 즉 표적 유전자의 전사 조절 영역 내 핵산에 상보적인 결합을 할 수 있는 가이드 서열(guide sequence)를 포함하는 도메인 및 제 1 상보적 도메인을 포함하는 제 1가닥; 및 제 2 상보적 도메인, 상기 제 1 상보적 도메인 서열에 상보적인 서열을 가지므로 제 1 상보적 도메인과 이중가닥 핵산을 형성할 수 있는 도메인, 근위 도메인(proximal domain); 및 선택적으로 꼬리 도메인을 포함하는 제 2 가닥을 포함할 수 있다.

[282] 이때, 상기 제 1가닥은 crRNA라고 지칭될 수 있고, 상기 제 2가닥은 tracrRNA로 지칭될 수 있다. 상기 crRNA는 가이드 도메인과 제 1 상보적 도메인을 포함할 수 있으며, 상기 tracrRNA는 제 2 상보적 도메인, 근위 도메인 및 선택적으로 꼬리 도메인을 포함할 수 있다.

[283] 또 다른 일 구체예로서, 단일가닥 gRNA는 3'으로부터 5' 방향으로 가이드 도메인, 즉 표적 유전자의 전사 조절 영역 내 핵산에 상보적인 결합을 할 수 있는 가이드 서열(guide sequence)를 포함하는 도메인; 제 1 상보적 도메인; 및 제 2 상보적 도메인, 상기 제 1 상보적 도메인 서열에 상보적인 서열을 가지므로 제 1 상보적 도메인과 이중가닥 핵산을 형성할 수 있는 도메인을 포함할 수 있다.

[284]

[285] 이때, 상기 제 1 상보적 도메인은 자연유래의 제 1 상보적 도메인과 상동성을 가지거나, 또는 자연유래의 제 1 상보적 도메인으로부터 유래될 수 있다. 또한, 상기 제 1 상보적 도메인은 자연에 존재하는 종에 따라 제 1 상보적 도메인의 뉴클레오타이드서열에 차이가 존재할 수 있으며, 자연에 존재하는 종이 포함하는 제 1 상보적 도메인으로부터 유래될 수 있고, 또는 자연에 존재하는 종이 포함하는 제 1 상보적 도메인과 일부 또는 완전한 상동성을 가질 수 있다.

[286] 일 구체예로서, 상기 제 1 상보적 도메인은 스트렙토코커스 피요게네스(*Streptococcus pyogenes*), 캄필로박터 제주니(*Campylobacter jejuni*), 스트렙토코커스 써모필러스(*Streptococcus thermophiles*), 스타필로코커스 아우레우스(*Staphylococcus aureus*) 또는 네이세리아 메닝기디티스(*Neisseria meningitides*)의 제 1 상보적 도메인 또는 유래된 제 1 상보적 도메인과 일부, 최소 50%이상, 또는 완전한 상동성을 가질 수 있다.

[287] 예를 들어, 상기 제 1 상보적 도메인이 스트렙토코커스 피요게네스의 제 1 상보적 도메인 또는 스트렙토코커스 피요게네스 유래 제 1 상보적 도메인인 경우, 상기 제 1 상보적 도메인은 5'-GUUUUAGAGCUA-3'일 수 있고, 또는 5'-GUUUUAGAGCUA-3'와 일부, 적어도 50% 이상의 상동성을 가지는 뉴클레오타이드서열일 수 있다. 이때, 상기 제 1 상보적 도메인은 추가로 $(X)_n$을 포함, 즉, 5'-GUUUUAGAGCUA$(X)_n$-3', 할 수 있다. 상기 X는 뉴클레오타이드 A, T, U 및 G로 이루어진 군에서 선택될 수 있으며, 상기 n은 뉴클레오타이드서열의 개수로, 5 내지 15의 정수일 수 있다. 이때, $(X)_n$은 동일한 뉴클레오타이드서열의

정수 n개만큼의 반복일 수 있고, 또는 뉴클레오타이드 A, T, U 및 G가 혼합된 정수 n개의 뉴클레오타이드서열일 수 있다.

[288] 또 다른 예를 들어, 상기 제 1 상보적 도메인이 캄필로박터 제주니의 제 1 상보적 도메인 또는 캄필로박터 제주니 유래 제 1 상보적 도메인인 경우, 상기 제 1 상보적 도메인은 5'-GUUUUAGUCCCUUUUUAAAUUUCUU-3' 또는 5'-GUUUUAGUCCCUU-3'일 수 있고, 또는 5'-GUUUUAGUCCCUUUUUAAAUUUCUU-3' 또는 5'-GUUUUAGUCCCUU-3'와 일부, 적어도 50% 이상의 상동성을 가지는 뉴클레오타이드서열일 수 있다. 이때, 상기 제 1 상보적 도메인은 추가로 $(X)_n$을 포함, 즉, 5'-GUUUUAGUCCCUUUUUAAAUUUCUU$(X)_n$-3' 또는 5'-GUUUUAGUCCCUU$(X)_n$-3', 할 수 있다. 상기 X는 뉴클레오타이드 A, T, U 및 G로 이루어진 군에서 선택될 수 있으며, 상기 n은 뉴클레오타이드서열의 개수로, 5 내지 15의 정수일 수 있다. 이때, $(X)_n$은 동일한 뉴클레오타이드서열의 정수 n개만큼의 반복일 수 있고, 또는 뉴클레오타이드 A, T, U 및 G가 혼합된 정수 n개의 뉴클레오타이드서열일 수 있다.

[289] 다른 일 구체예로서, 상기 제 1 상보적 도메인은 팔쿠박테리아 박테리움(*Parcubacteria bacterium* (GWC2011_GWC2_44_17)), 라츠노스피라세애 박테리움(*Lachnospiraceae bacterium* (MC2017)), 부티리비브리오 프로테오클라시커스(*Butyrivibrio proteoclasiicus*), 페레그리니박테리아 박테리움(*Peregrinibacteria bacterium* (GW2011_GWA_33_10)), 액시다미노코커스 에스피(*Acidaminococcus sp.* (BV3L6)), 포르피로모나스 마카캐(*Porphyromonas macacae*), 라츠노피라세애 박테리움(*Lachnospiraceae bacterium* (ND2006)), 포르피로모나스 크레비오리카니스(*Porphyromonas crevioricanis*), 프레보텔라 디이엔스(*Prevotella disiens*), 모라셀라 보보쿨리(*Moraxella bovoculi* (237)), 스미이헬라 에스피(*Smiihella sp.* (SC_KO8D17)), 렙포스피라 이나다이(*Leptospira inadai*), 라츠노스피라세애 박테리움(*Lachnospiraceae bacterium* (MA2020)), 프란시셀라 노비시다(*Francisella novicida* (U112)), 캔디다투스 메타노플라즈마 털미툼(*Candidatus Methanoplasma termitum*) 또는 에유박테리움 엘리겐스(*Eubacterium eligens*)의 제 1 상보적 도메인 또는 유래된 제 1 상보적 도메인과 일부, 최소 50%이상, 또는 완전한 상동성을 가질 수 있다.

[290] 예를 들어, 상기 제 1 상보적 도메인이 팔쿠박테리아 박테리움의 제 1 상보적 도메인 또는 팔쿠박테리아 박테리움 유래 제 1 상보적 도메인인 경우, 상기 제 1 상보적 도메인은 5'-UUUGUAGAU-3' 일 수 있고, 또는 5'-UUUGUAGAU-3'와 일부, 적어도 50% 이상의 상동성을 가지는 뉴클레오타이드 서열일 수 있다. 이때, 상기 제 1 상보적 도메인은 추가로 $(X)_n$을 포함, 즉, 5'-$(X)_n$UUUGUAGAU-3' 할 수 있다. 상기 X는 뉴클레오타이드 A, T, U 및 G로 이루어진 군에서 선택될 수 있으며, 상기 n은 뉴클레오타이드 서열의 개수로, 1 내지 5의 정수일 수 있다.

이때, $(X)_n$은 동일한 뉴클레오타이드 서열의 정수 n개만큼의 반복일 수 있고, 또는 뉴클레오타이드 A, T, U 및 G가 혼합된 정수 n개의 뉴클레오타이드 서열일 수 있다.

[291]

[292] 이때, 상기 연결 도메인은 제 1 상보적 도메인과 제 2 상보적 도메인을 연결하는 역할을 하는 뉴클레오타이드 서열일 수 있다.

[293] 상기 연결 도메인은 제 1 상보적 도메인 및 제 2 상보적 도메인과 각각 공유결합 또는 비공유결합을 할 수 있다.

[294] 상기 연결 도메인은 제 1 상보적 도메인과 제 2 상보적 도메인을 공유적 또는 비공유적으로 연결할 수 있다.

[295] 상기 연결 도메인은 단일가닥 gRNA 분자에 사용하기에 적합하며, 이중 gRNA의 제 1 가닥 및 제 2 가닥과 공유결합 또는 비공유결합 하거나, 또는 제 1 가닥 및 제 2 가닥을 공유적 또는 비공유적으로 연결하여 단일가닥 gRNA을 생성에 사용될 수 있다.

[296] 상기 연결 도메인은 이중 gRNA의 crRNA 및 tracrRNA과 공유결합 또는 비공유결합 하거나, 또는 crRNA 및 tracrRNA를 공유적 또는 비공유적으로 연결하여 단일가닥 gRNA를 생성에 사용될 수 있다.

[297]

[298] 이때, 상기 제 2 상보적 도메인은 자연유래의 제 2 상보적 도메인과 상동성을 가지거나, 또는 자연유래의 제 2 상보적 도메인으로부터 유래될 수 있다. 또한, 상기 제 2 상보적 도메인은 자연에 존재하는 종에 따라 제 2 상보적 도메인의 뉴클레오타이드서열에 차이가 존재할 수 있으며, 자연에 존재하는 종이 포함하는 제 2 상보적 도메인으로부터 유래될 수 있고, 또는 자연에 존재하는 종이 포함하는 제 2 상보적 도메인과 일부 또는 완전한 상동성을 가질 수 있다.

[299] 일 구체예로서, 상기 제 2 상보적 도메인은 스트렙토코커스 피요게네스(*Streptococcus pyogenes*), 캄필로박터 제주니(*Campylobacter jejuni*), 스트렙토코커스 써모필러스(*Streptococcus thermophiles*), 스타필로코커스 아우레우스(*Staphylococcus aureus*) 또는 네이세리아 메닝기디티스(*Neisseria meningitides*)의 제 2 상보적 도메인 또는 유래된 제 2 상보적 도메인과 일부, 최소 50%이상, 또는 완전한 상동성을 가질 수 있다.

[300] 예를 들어, 상기 제 2 상보적 도메인이 스트렙토코커스 피요게네스의 제 2 상보적 도메인 또는 스트렙토코커스 피요게네스 유래 제 2 상보적 도메인인 경우, 상기 제 2 상보적 도메인은 5'-UAGCAAGUUAAAAU-3'일 수 있고, 또는 5'-UAGCAAGUUAAAAU-3'와 일부, 적어도 50% 이상의 상동성을 가지는 뉴클레오타이드서열일 수 있다(밑줄 표시는 제 1 상보적 도메인과 이중가닥을 형성하는 뉴클레오타이드서열). 이때, 상기 제 2 상보적 도메인은 추가로 $(X)_n$ 또는/및 $(X)_m$을 포함, 즉, 5'-$(X)_n$ UAGCAAGUUAAAAU$(X)_m$-3', 할 수 있다. 상기 X는 뉴클레오타이드 A, T, U 및 G로 이루어진 군에서 선택될 수 있으며, 상기 n

및 m은 뉴클레오타이드서열의 개수로, 상기 n은 1 내지 15의 정수일 수 있고, 상기 m은 1 내지 6일 수 있다. 이때, $(X)_n$은 동일한 뉴클레오타이드서열의 정수 n개만큼의 반복일 수 있고, 또는 뉴클레오타이드 A, T, U 및 G가 혼합된 정수 n개의 뉴클레오타이드서열일 수 있다. 또한 $(X)_m$은 동일한 뉴클레오타이드서열의 정수 m개만큼의 반복일 수 있고, 또는 뉴클레오타이드 A, T, U 및 G가 혼합된 정수 m개의 뉴클레오타이드서열일 수 있다.

[301] 또 다른 예를 들어, 상기 제 2 상보적 도메인이 캄필로박터 제주니의 제 2 상보적 도메인 또는 캄필로박터 제주니 유래 제 2 상보적 도메인인 경우, 상기 제 2 상보적 도메인은 5'-AAGAAAUUUAAAAAGGGACUAAAAU-3' 또는 5'-AAGGGACUAAAAU-3'일 수 있고, 또는 5'-AAGAAAUUUAAAAAGGGACUAAAAU-3' 또는 5'-AAGGGACUAAAAU-3'와 일부, 적어도 50% 이상의 상동성을 가지는 뉴클레오타이드서열일 수 있다(밑줄 표시는 제 1 상보적 도메인과 이중가닥을 형성하는 뉴클레오타이드서열). 이때, 상기 제 2 상보적 도메인은 추가로 $(X)_n$ 또는/및 $(X)_m$을 포함, 즉, 5'-$(X)_n$ AAGAAAUUUAAAAAGGGACUAAAAU$(X)_m$-3' 또는 5'-$(X)_n$ AAGGGACUAAAAU$(X)_m$-3', 할 수 있다. 상기 X는 뉴클레오타이드 A, T, U 및 G로 이루어진 군에서 선택될 수 있으며, 상기 n은 1 내지 15의 정수일 수 있고, 상기 m은 1 내지 6일 수 있다. 이때, $(X)_n$은 동일한 뉴클레오타이드서열의 정수 n개만큼의 반복일 수 있고, 또는 뉴클레오타이드 A, T, U 및 G가 혼합된 정수 n개의 뉴클레오타이드서열일 수 있다. 또한 $(X)_m$은 동일한 뉴클레오타이드서열의 정수 m개만큼의 반복일 수 있고, 또는 뉴클레오타이드 A, T, U 및 G가 혼합된 정수 m개의 뉴클레오타이드서열일 수 있다.

[302] 다른 일 구체예로서, 상기 제 2 상보적 도메인은 팔쿠박테리아 박테리움(*Parcubacteria bacterium* (GWC2011_GWC2_44_17)), 라츠노스피라세애 박테리움(*Lachnospiraceae bacterium* (MC2017)), 부티리비브리오 프로테오클라시커스(*Butyrivibrio proteoclasiicus*), 페레그리니박테리아 박테리움(*Peregrinibacteria bacterium* (GW2011_GWA_33_10)), 액시다미노코커스 에스피(*Acidaminococcus sp.* (BV3L6)), 포르피로모나스 마카캐(*Porphyromonas macacae*), 라츠노피라세애 박테리움(*Lachnospiraceae bacterium* (ND2006)), 포르피로모나스 크레비오리카니스(*Porphyromonas crevioricanis*), 프레보텔라 디이엔스(*Prevotella disiens*), 모라셀라 보보쿨리(*Moraxella bovoculi* (237)), 스미이헬라 에스피(*Smiihella sp.* (SC_KO8D17)), 렙포스피라 이나다이(*Leptospira inadai*), 라츠노스피라세애 박테리움(*Lachnospiraceae bacterium* (MA2020)), 프란시셀라 노비시다(*Francisella novicida* (U112)), 캔디다투스 메타노플라즈마 털미툼(*Candidatus Methanoplasma termitum*) 또는 에유박테리움 엘리겐스(*Eubacterium eligens*)의 제 2 상보적 도메인 또는 유래된 제 2 상보적 도메인과 일부, 최소 50%이상, 또는 완전한 상동성을 가질 수 있다.

[303] 예를 들어, 상기 제 2 상보적 도메인이 팔쿠박테리아 박테리움의 제 2 상보적 도메인 또는 팔쿠박테리아 박테리움 유래 제 2 상보적 도메인인 경우, 상기 제 2 상보적 도메인은 5'-AAAUUUCUACU-3' 일 수 있고, 또는 5'-AAAUUUCUACU-3'와 일부, 적어도 50% 이상의 상동성을 가지는 뉴클레오타이드 서열일 수 있다(밑줄 표시는 제 1 상보적 도메인과 이중가닥을 형성하는 뉴클레오타이드 서열). 이때, 상기 제 2 상보적 도메인은 추가로 $(X)_n$ 또는/및 $(X)_m$을 포함, 즉, 5'-$(X)_n$AAAUUUCUACU $(X)_m$-3' 할 수 있다. 상기 X는 뉴클레오타이드 A, T, U 및 G로 이루어진 군에서 선택될 수 있으며, 상기 n 및 m은 뉴클레오타이드 서열의 개수로, 상기 n은 1 내지 10의 정수일 수 있고, 상기 m은 1 내지 6일 수 있다. 이때, $(X)_n$은 동일한 뉴클레오타이드 서열의 정수 n개만큼의 반복일 수 있고, 또는 뉴클레오타이드 A, T, U 및 G가 혼합된 정수 n개의 뉴클레오타이드 서열일 수 있다. 또한 $(X)_m$은 동일한 뉴클레오타이드 서열의 정수 m개만큼의 반복일 수 있고, 또는 뉴클레오타이드 A, T, U 및 G가 혼합된 정수 m개의 뉴클레오타이드 서열일 수 있다.

[304]

[305] 이때, 상기 제 1 상보적 도메인과 상기 제 2 상보적 도메인은 상보적 결합을 할 수 있다.

[306] 상기 제 1 상보적 도메인과 상기 제 2 상보적 도메인은 상기 상보적 결합을 통해 이중가닥을 형성할 수 있다.

[307] 상기 형성된 이중가닥은 CRISPR 효소와 상호작용할 수 있다.

[308] 선택적으로, 상기 제 1 상보적 도메인은 제 2 가닥의 제 2 상보적 도메인과 상보적 결합을 하지않는 추가 뉴클레오타이드 서열을 포함할 수 있다.

[309] 이때, 상기 추가 뉴클레오타이드 서열은 1 내지 15개의 뉴클레오타이드 서열일 수 있다. 예를 들어, 상기 추가 뉴클레오타이드 서열은 1 내지 5개의 뉴클레오타이드 서열, 5 내지 10개의 뉴클레오타이드 서열, 또는 10 내지 15개의 뉴클레오타이드 서열일 수 있다.

[310]

[311] 이때, 상기 근위 도메인은 제 2 상보적 도메인의 3' 방향에 위치하는 도메인일 수 있다.

[312] 상기 근위 도메인은 자연유래의 근위 도메인과 상동성을 가지거나, 또는 자연유래의 근위 도메인으로부터 유래될 수 있다. 또한, 상기 근위 도메인은 자연에 존재하는 종에 따라 근위 도메인의 뉴클레오타이드서열에 차이가 존재할 수 있으며, 자연에 존재하는 종이 포함하는 근위 도메인으로부터 유래될 수 있고, 또는 자연에 존재하는 종이 포함하는 근위 도메인과 일부 또는 완전한 상동성을 가질 수 있다.

[313] 일 구체예로서, 상기 근위 도메인은 스트렙토코커스 피요게네스(*Streptococcus pyogenes*), 캄필로박터 제주니(*Campylobacter jejuni*), 스트렙토코커스 써모필러스(*Streptococcus thermophiles*), 스타필로코커스 아우레우스(

Staphylococcus aureus) 또는 네이세리아 메닝기디티스(*Neisseria meningitides*)의 근위 도메인 또는 유래된 근위 도메인과 일부, 최소 50%이상, 또는 완전한 상동성을 가질 수 있다.

[314] 예를 들어, 상기 근위 도메인이 스트렙토코커스 피요게네스의 근위 도메인 또는 스트렙토코커스 피요게네스 유래 근위 도메인인 경우, 상기 근위 도메인은 5'-AAGGCUAGUCCG-3'일 수 있고, 또는 5'-AAGGCUAGUCCG-3'와 일부, 적어도 50% 이상의 상동성을 가지는 뉴클레오타이드 서열일 수 있다. 이때, 상기 근위 도메인은 추가로 $(X)_n$을 포함, 즉, 5'-AAGGCUAGUCCG$(X)_n$-3', 할 수 있다. 상기 X는 뉴클레오타이드 A, T, U 및 G로 이루어진 군에서 선택될 수 있으며, 상기 n은 뉴클레오타이드 서열의 개수로, 1 내지 15의 정수일 수 있다. 이때, $(X)_n$은 동일한 뉴클레오타이드 서열의 정수 n개만큼의 반복일 수 있고, 또는 뉴클레오타이드 A, T, U 및 G가 혼합된 정수 n개의 뉴클레오타이드 서열일 수 있다.

[315] 또 다른 예를 들어, 상기 근위 도메인이 캄필로박터 제주니의 근위 도메인 또는 캄필로박터 제주니 유래 근위 도메인인 경우, 상기 근위 도메인은 5'-AAAGAGUUUGC-3'일 수 있고, 또는 5'-AAAGAGUUUGC-3'와 일부, 적어도 50% 이상의 상동성을 가지는 뉴클레오타이드 서열일 수 있다. 이때, 상기 근위 도메인은 추가로 $(X)_n$을 포함, 즉, 5'-AAAGAGUUUGC$(X)_n$-3', 할 수 있다. 상기 X는 뉴클레오타이드 A, T, U 및 G로 이루어진 군에서 선택될 수 있으며, 상기 n은 뉴클레오타이드서열의 개수로, 1 내지 40의 정수일 수 있다. 이때, $(X)_n$은 동일한 뉴클레오타이드 서열의 정수 n개만큼의 반복일 수 있고, 또는 뉴클레오타이드 A, T, U 및 G가 혼합된 정수 n개의 뉴클레오타이드 서열일 수 있다.

[316]

[317] 이때, 상기 꼬리 도메인은 단일가닥 gRNA 또는 이중 gRNA의 제 1 가닥 또는 제 2 가닥의 3' 말단에 선택적으로 추가될 수 있다.

[318] 상기 꼬리 도메인은 자연유래의 꼬리 도메인과 상동성을 가지거나, 또는 자연유래의 꼬리 도메인으로부터 유래될 수 있다. 또한, 상기 꼬리 도메인은 자연에 존재하는 종에 따라 꼬리 도메인의 뉴클레오타이드서열에 차이가 존재할 수 있으며, 자연에 존재하는 종이 포함하는 꼬리 도메인으로부터 유래될 수 있고, 또는 자연에 존재하는 종이 포함하는 꼬리 도메인과 일부 또는 완전한 상동성을 가질 수 있다.

[319] 일 구체예로서, 상기 꼬리 도메인은 스트렙토코커스 피요게네스(*Streptococcus pyogenes*), 캄필로박터 제주니(*Campylobacter jejuni*), 스트렙토코커스 써모필러스(*Streptococcus thermophiles*), 스타필로코커스 아우레우스(*Staphylococcus aureus*) 또는 네이세리아 메닝기디티스(*Neisseria meningitides*)의 꼬리 도메인 또는 유래된 꼬리 도메인과 일부, 최소 50%이상, 또는 완전한 상동성을 가질 수 있다.

[320] 예를 들어, 상기 꼬리 도메인이 스트렙토코커스 피요게네스의 꼬리 도메인 또는 스트렙토코커스 피요게네스 유래 꼬리 도메인인 경우, 상기 꼬리 도메인은 5'-UUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGC-3'일 수 있고, 또는 5'-UUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGC-3'와 일부, 적어도 50% 이상의 상동성을 가지는 뉴클레오타이드서열일 수 있다. 이때, 상기 꼬리 도메인은 추가로 $(X)_n$을 포함, 즉, 5'-UUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGC$(X)_n$-3', 할 수 있다. 상기 X는 뉴클레오타이드 A, T, U 및 G로 이루어진 군에서 선택될 수 있으며, 상기 n은 뉴클레오타이드서열의 개수로, 1 내지 15의 정수일 수 있다. 이때, $(X)_n$은 동일한 뉴클레오타이드 서열의 정수 n개만큼의 반복일 수 있고, 또는 뉴클레오타이드 A, T, U 및 G가 혼합된 정수 n개의 뉴클레오타이드 서열일 수 있다.

[321] 또 다른 예를 들어, 상기 꼬리 도메인이 캄필로박터 제주니의 꼬리 도메인 또는 캄필로박터 제주니 유래 꼬리 도메인인 경우, 상기 꼬리 도메인은 5'-GGGACUCUGCGGGGUUACAAUCCCCUAAAACCGCUUUU-3'일 수 있고, 또는 5'-GGGACUCUGCGGGGUUACAAUCCCCUAAAACCGCUUUU-3'와 일부, 적어도 50% 이상의 상동성을 가지는 뉴클레오타이드 서열일 수 있다. 이때, 상기 꼬리 도메인은 추가로 $(X)_n$을 포함, 즉, 5'-GGGACUCUGCGGGGUUACAAUCCCCUAAAACCGCUUUU$(X)_n$-3', 할 수 있다. 상기 X는 뉴클레오타이드 A, T, U 및 G로 이루어진 군에서 선택될 수 있으며, 상기 n은 뉴클레오타이드 서열의 개수로, 1 내지 15의 정수일 수 있다. 이때, $(X)_n$은 동일한 뉴클레오타이드 서열의 정수 n개만큼의 반복일 수 있고, 또는 뉴클레오타이드 A, T, U 및 G가 혼합된 정수 n개의 뉴클레오타이드 서열일 수 있다.

[322] 다른 일 구체예에서, 상기 꼬리 도메인은 시험관내 또는 생체내 전사 방법과 관련된 3' 말단에 1 내지 10개의 뉴클레오타이드 서열을 포함할 수 있다.

[323] 예를 들어, T7 프로모터가 gRNA의 시험관내 전사를 위해 사용될 때, 상기 꼬리 도메인은 DNA 주형의 3' 말단에 존재하는 임의의 뉴클레오타이드 서열일 수 있다. 또한, U6 프로모터가 생체내 전사를 위해 사용되는 경우, 상기 꼬리 도메인은 UUUUUU일 수 있으며, H1 프로모터가 전사를 위해 사용되는 경우, 상기 꼬리 도메인은 UUUU일 수 있고, pol-III 프로모터를 사용하는 경우에는, 상기 꼬리 도메인은 여러 개의 우라실 뉴클레오타이드거나 또는 대안 될 수 있는 뉴클레오타이드를 포함할 수 있다.

[324]

[325] gRNA는 상기에 기재된 바와 같이 다수의 도메인을 포함할 수 있어, gRNA가 포함하는 도메인의 종류 및 개수에 따라 핵산 서열의 길이를 조절할 수 있으며, 각각의 도메인에 의해 3차원 형태 또는 gRNA의 활성 형태의 가닥내 또는 가닥간 상호작용을 할 수 있다.

[326] gRNA는 단일가닥 gRNA(단일 RNA 분자); 또는 이중 gRNA(하나 초과의 통상적으로 2개의 별개의 RNA 분자를 포함함)로서 지칭될 수 있다.

[327]

[328] **이중 gRNA**

[329] 이중 gRNA는 제 1 가닥 및 제 2 가닥으로 구성된다.

[330] 이때, 상기 제 1 가닥은

[331] 5'-[가이드 도메인]-[제 1 상보적 도메인]-3'으로 구성될 수 있고,

[332] 상기 제 2 가닥은

[333] 5'-[제 2 상보적 도메인]-[근위 도메인(proximal domain)]-3' 또는

[334] 5'-[제 2 상보적 도메인]-[근위 도메인(proximal domain)]-[꼬리 도메인]-3'으로 구성될 수 있다.

[335] 이때, 상기 제 1가닥은 crRNA라고 지칭될 수 있고, 상기 제 2가닥은 tracrRNA로 지칭될 수 있다.

[336] 이때, 상기 제 1가닥 및 제 2가닥은 선택적으로 추가적인 뉴클레오타이드 서열을 포함할 수 있다.

[337]

[338] 일 구체예로서, 상기 제 1 가닥은

[339] 5'-(N_{target})-$(Q)_m$-3'; 또는

[340] 5'-$(X)_a$-(N_{target})-$(X)_b$-$(Q)_m$-$(X)_c$-3'일 수 있다.

[341] 이때, 상기 N_{target}은 표적 유전자의 전사 조절 영역 내 핵산의 이중 가닥 중 어느 하나 가닥의 일부 서열에 상보적인 뉴클레오타이드 서열로서, 상기 N_{target}은 표적 유전자의 전사 조절 영역 내 핵산 상의 표적 서열에 따라 변할 수 있는 뉴클레오타이드 서열 부위이다.

[342] 이때, 상기 $(Q)_m$은 제 1 상보적 도메인을 포함하는 뉴클레오타이드 서열로, 제 2 가닥의 제 2 상보적 도메인과 상보적 결합을 할 수 있는 뉴클레오타이드 서열을 포함한다. 상기 $(Q)_m$은 자연에 존재하는 종의 제 1 상보적 도메인과 일부 또는 완전한 상동성을 가지는 서열일 수 있으며, 유래된 종에 따라 상기 제 1 상보적 도메인의 뉴클레오타이드 서열은 변경될 수 있다. 상기 Q는 A, U, C 및 G로 이루어진 군에서 각각 독립적으로 선택될 수 있으며, 상기 m은 뉴클레오타이드 서열의 개수로, 5 내지 35의 정수일 수 있다.

[343] 예를 들어, 상기 제 1 상보적 도메인이 스트렙토코커스 피요게네스의 제 1 상보적 도메인 또는 스트렙토코커스 피요게네스 유래 제 1 상보적 도메인과 일부 또는 완전한 상동성을 가질 경우에, 상기 $(Q)_m$은 5'-GUUUUAGAGCUA-3'일 수 있고, 또는 5'-GUUUUAGAGCUA-3'와 적어도 50% 이상의 상동성을 가지는 뉴클레오타이드 서열일 수 있다.

[344] 다른 예로, 상기 제 1 상보적 도메인이 캄필로박터 제주니의 제 1 상보적 도메인 또는 캄필로박터 제주니 유래 제 1 상보적 도메인과 일부 또는 완전한 상동성을 가질 경우에, 상기 $(Q)_m$은 5'-GUUUUAGUCCCUUUUUAAAUUUCUU-3' 또는

5'-GUUUUAGUCCCUU-3'일 수 있고, 또는
5'-GUUUUAGUCCCUUUUUAAAUUUCUU-3' 또는 5'-GUUUUAGUCCCUU-3'와 적어도 50% 이상의 상동성을 가지는 뉴클레오타이드 서열일 수 있다.

[345] 또 다른 예로, 상기 제 1 상보적 도메인이 스트렙토코커스 써모필러스의 제 1 상보적 도메인 또는 스트렙토코커스 써모필러스 유래 제 1 상보적 도메인과 일부 또는 완전한 상동성을 가질 경우에, 상기 $(Q)_m$은 5'-GUUUUAGAGCUGUGUUGUUUCG-3'일 수 있고, 또는 5'-GUUUUAGAGCUGUGUUGUUUCG-3'와 적어도 50% 이상의 상동성을 가지는 뉴클레오타이드 서열일 수 있다.

[346] 또한, 상기 $(X)_a$, $(X)_b$ 및 $(X)_c$는 선택적으로 추가할 수 있는 뉴클레오타이드 서열로, 상기 X는 A, U, C 및 G로 이루어진 군에서 각각 독립적으로 선택될 수 있으며, 상기 a, b 및 c는 뉴클레오타이드 서열의 개수로, 0 또는 1 내지 20의 정수일 수 있다.

[347]

[348] 일 구체예로서, 상기 제 2 가닥은

[349] $5'-(Z)_h-(P)_k-3'$; 또는

[350] $5'-(X)_d-(Z)_h-(X)_e-(P)_k-(X)_f-3'$ 일 수 있다.

[351] 다른 일 구체예로서, 상기 제 2 가닥은

[352] $5'-(Z)_h-(P)_k-(F)_i-3'$; 또는

[353] $5'-(X)_d-(Z)_h-(X)_e-(P)_k-(X)_f-(F)_i-3'$ 일 수 있다.

[354] 이때, 상기 $(Z)_h$는 제 2 상보적 도메인을 포함하는 뉴클레오타이드 서열로, 제 1 가닥의 제 1 상보적 도메인과 상보적 결합을 할 수 있는 뉴클레오타이드 서열을 포함한다. 상기 $(Z)_h$은 자연에 존재하는 종의 제 2 상보적 도메인과 일부 또는 완전한 상동성을 가지는 서열일 수 있으며, 유래된 종에 따라 상기 제 2 상보적 도메인의 뉴클레오타이드 서열은 변경될 수 있다. 상기 Z는 A, U, C 및 G로 이루어진 군에서 각각 독립적으로 선택될 수 있으며, 상기 h은 뉴클레오타이드 서열의 개수로, 5 내지 50의 정수일 수 있다.

[355] 예를 들어, 상기 제 2 상보적 도메인이 스트렙토코커스 피요게네스의 제 2 상보적 도메인 또는 스트렙토코커스 피요게네스 유래 제 2 상보적 도메인과 일부 또는 완전한 상동성을 가질 경우에, 상기 $(Z)_h$은 5'-UAGCAAGUUAAAAU-3'일 수 있고, 또는 5'-UAGCAAGUUAAAAU-3'와 적어도 50% 이상의 상동성을 가지는 뉴클레오타이드 서열일 수 있다.

[356] 다른 예로, 상기 제 2 상보적 도메인이 캄필로박터 제주니의 제 2 상보적 도메인 또는 캄필로박터 제주니 유래 제 2 상보적 도메인과 일부 또는 완전한 상동성을 가질 경우에, 상기 $(Z)_h$은 5'-AAGAAAUUUAAAAAGGGACUAAAAU-3' 또는 5'-AAGGGACUAAAAU-3'일 수 있고, 또는 5'-AAGAAAUUUAAAAAGGGACUAAAAU-3' 또는 5'-AAGGGACUAAAAU-3'와 적어도 50% 이상의 상동성을 가지는 뉴클레오타이드 서열일 수 있다.

[357] 또 다른 예로, 상기 제 2 상보적 도메인이 스트렙토코커스 써모필러스의 제 2 상보적 도메인 또는 스트렙토코커스 써모필러스 유래 제 2 상보적 도메인과 일부 또는 완전한 상동성을 가질 경우에, 상기 $(Z)_h$은 5'-CGAAACAACACAGCGAGUUAAAAU-3'일 수 있고, 또는 5'-CGAAACAACACAGCGAGUUAAAAU-3'와 적어도 50% 이상의 상동성을 가지는 뉴클레오타이드 서열일 수 있다.

[358] 상기 $(P)_k$는 근위 도메인을 포함하는 뉴클레오타이드 서열로, 자연에 존재하는 종의 근위 도메인과 일부 또는 완전한 상동성을 가지는 서열일 수 있으며, 유래된 종에 따라 상기 근위 도메인의 뉴클레오타이드 서열은 변경될 수 있다. 상기 P는 A, U, C 및 G로 이루어진 군에서 각각 독립적으로 선택될 수 있으며, 상기 k은 뉴클레오타이드 서열의 개수로, 1 내지 20의 정수일 수 있다.

[359] 예를 들어, 상기 근위 도메인이 스트렙토코커스 피요게네스의 근위 도메인 또는 스트렙토코커스 피요게네스 유래 근위 도메인과 일부 또는 완전한 상동성을 가질 경우에, 상기 $(P)_k$는 5'-AAGGCUAGUCCG-3'일 수 있고, 또는 5'-AAGGCUAGUCCG-3'와 적어도 50% 이상의 상동성을 가지는 뉴클레오타이드 서열일 수 있다.

[360] 다른 예로, 상기 근위 도메인이 캄필로박터 제주니의 근위 도메인 또는 캄필로박터 제주니 유래 근위 도메인과 일부 또는 완전한 상동성을 가질 경우에, 상기 $(P)_k$는 5'-AAAGAGUUUGC-3'일 수 있고, 또는 5'-AAAGAGUUUGC-3'와 적어도 50% 이상의 상동성을 가지는 뉴클레오타이드 서열일 수 있다.

[361] 또 다른 예로, 상기 근위 도메인이 스트렙토코커스 써모필러스의 근위 도메인 또는 스트렙토코커스 써모필러스 유래 근위 도메인과 일부 또는 완전한 상동성을 가질 경우에, 상기 $(P)_k$는 5'-AAGGCUUAGUCCG-3'일 수 있고, 또는 5'-AAGGCUUAGUCCG-3'와 적어도 50% 이상의 상동성을 가지는 뉴클레오타이드 서열일 수 있다.

[362] 상기 $(F)_i$는 꼬리 도메인을 포함하는 뉴클레오타이드 서열로, 자연에 존재하는 종의 꼬리 도메인과 일부 또는 완전한 상동성을 가지는 서열일 수 있으며, 유래된 종에 따라 상기 꼬리 도메인의 뉴클레오타이드 서열은 변경될 수 있다. 상기 F는 A, U, C 및 G로 이루어진 군에서 각각 독립적으로 선택될 수 있으며, 상기 i은 뉴클레오타이드 서열의 개수로, 1 내지 50의 정수일 수 있다.

[363] 예를 들어, 상기 꼬리 도메인이 스트렙토코커스 피요게네스의 꼬리 도메인 또는 스트렙토코커스 피요게네스 유래 꼬리 도메인과 일부 또는 완전한 상동성을 가질 경우에, 상기 $(F)_i$는 5'-UUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGC-3'일 수 있고, 또는 5'-UUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGC-3'와 적어도 50% 이상의 상동성을 가지는 뉴클레오타이드 서열일 수 있다.

[364] 다른 예로, 상기 꼬리 도메인이 캄필로박터 제주니의 꼬리 도메인 또는 캄필로박터 제주니 유래 꼬리 도메인과 일부 또는 완전한 상동성을 가질 경우에,

상기 $(F)_i$는 5'-GGGACUCUGCGGGGUUACAAUCCCCUAAAACCGCUUUU-3'일 수 있고, 또는 5'-GGGACUCUGCGGGGUUACAAUCCCCUAAAACCGCUUUU-3'와 적어도 50% 이상의 상동성을 가지는 뉴클레오타이드 서열일 수 있다.

[365] 또 다른 예로, 상기 꼬리 도메인이 스트렙토코커스 써모필러스의 꼬리 도메인 또는 스트렙토코커스 써모필러스 유래 꼬리 도메인과 일부 또는 완전한 상동성을 가질 경우에, 상기 $(F)_i$는 5'-UACUCAACUUGAAAAGGUGGCACCGAUUCGGUGUUUUU-3'일 수 있고, 또는 5'-UACUCAACUUGAAAAGGUGGCACCGAUUCGGUGUUUUU-3'와 적어도 50% 이상의 상동성을 가지는 뉴클레오타이드 서열일 수 있다.

[366] 또한, 상기 $(F)_i$는 시험관내 또는 생체내 전사 방법과 관련된 3' 말단에 1 내지 10개의 뉴클레오타이드 서열을 포함할 수 있다.

[367] 예를 들어, T7 프로모터가 gRNA의 시험관내 전사를 위해 사용될 때, 상기 꼬리 도메인은 DNA 주형의 3' 말단에 존재하는 임의의 뉴클레오타이드 서열일 수 있다. 또한, U6 프로모터가 생체내 전사를 위해 사용되는 경우, 상기 꼬리 도메인은 UUUUUU일 수 있으며, H1 프로모터가 전사를 위해 사용되는 경우, 상기 꼬리 도메인은 UUUU일 수 있고, pol-III 프로모터를 사용하는 경우에는, 상기 꼬리 도메인은 여러 개의 우라실 뉴클레오타이드거나 또는 대안될 수 있는 뉴클레오타이드를 포함할 수 있다.

[368] 또한, 상기 $(X)_d$, $(X)_e$ 및 $(X)_f$는 선택적으로 추가할 수 있는 뉴클레오타이드 서열로, 상기 X는 A, U, C 및 G로 이루어진 군에서 각각 독립적으로 선택될 수 있으며, 상기 d, e 및 f는 뉴클레오타이드 서열의 개수로, 0 또는 1 내지 20의 정수일 수 있다.

[369]

[370] **<u>단일가닥 gRNA</u>**

[371] 단일가닥 gRNA는 제 1 단일가닥 gRNA 및 제 2 단일가닥 gRNA로 나뉠 수 있다.

[372] ***제 1 단일가닥 gRNA***

[373] 제 1 단일가닥 gRNA는 상기 이중 gRNA의 제 1 가닥과 제 2 가닥을 연결 도메인으로 연결한 단일가닥 gRNA이다.

[374] 구체적으로, 상기 단일가닥 gRNA는

[375] 5'-[가이드 도메인]-[제 1 상보적 도메인]-[연결 도메인]-[제 2 상보적 도메인]-3',

[376] 5'-[가이드 도메인]-[제 1 상보적 도메인]-[연결 도메인]-[제 2 상보적 도메인]-[근위 도메인(proximal domain)]-3' 또는

[377] 5'-[가이드 도메인]-[제 1 상보적 도메인]-[연결 도메인]-[제 2 상보적 도메인]-[근위 도메인(proximal domain)]-[꼬리 도메인]-3'으로 구성될 수 있다.

[378] 상기 제 1 단일가닥 gRNA는 선택적으로 추가적인 뉴클레오타이드 서열을 포함할 수 있다.

[379]

[380] 일 구체예로서, 상기 제 1 단일가닥 gRNA는

[381] 5'-(N_{target})-$(Q)_m$-$(L)_j$-$(Z)_h$-3';

[382] 5'-(N_{target})-$(Q)_m$-$(L)_j$-$(Z)_h$-$(P)_k$-3'; 또는

[383] 5'-(N_{target})-$(Q)_m$-$(L)_j$-$(Z)_h$-$(P)_k$-$(F)_i$-3'일 수 있다.

[384] 다른 일 구체예로서, 상기 단일가닥 gRNA는

[385] 5'-$(X)_a$-(N_{target})-$(X)_b$-$(Q)_m$-$(X)_c$-$(L)_j$-$(X)_d$-$(Z)_h$-$(X)_e$-3';

[386] 5'-$(X)_a$-(N_{target})-$(X)_b$-$(Q)_m$-$(X)_c$-$(L)_j$-$(X)_d$-$(Z)_h$-$(X)_e$-$(P)_k$-$(X)_f$-3'; 또는

[387] 5'-$(X)_a$-(N_{target})-$(X)_b$-$(Q)_m$-$(X)_c$-$(L)_j$-$(X)_d$-$(Z)_h$-$(X)_e$-$(P)_k$-$(X)_f$-$(F)_i$-3'일 수 있다.

[388] 이때, 상기 N_{target}은 표적 유전자의 전사 조절 영역 내 핵산의 이중 가닥 중 어느 하나 가닥의 일부 서열에 상보적인 뉴클레오타이드 서열로서, 상기 N_{target}은 표적 유전자의 전사 조절 영역 내 핵산 상의 표적 서열에 따라 변할 수 있는 뉴클레오타이드 서열 부위이다.

[389] 상기 $(Q)_m$은 제 1 상보적 도메인을 포함하는 뉴클레오타이드 서열로, 제 2 상보적 도메인과 상보적 결합을 할 수 있는 뉴클레오타이드 서열을 포함한다. 상기 $(Q)_m$은 자연에 존재하는 종의 제 1 상보적 도메인과 일부 또는 완전한 상동성을 가지는 서열일 수 있으며, 유래된 종에 따라 상기 제 1 상보적 도메인의 뉴클레오타이드 서열은 변경될 수 있다. 상기 Q는 A, U, C 및 G로 이루어진 군에서 각각 독립적으로 선택될 수 있으며, 상기 m은 뉴클레오타이드 서열의 개수로, 5 내지 35의 정수일 수 있다.

[390] 예를 들어, 상기 제 1 상보적 도메인이 스트렙토코커스 피요게네스의 제 1 상보적 도메인 또는 스트렙토코커스 피요게네스 유래 제 1 상보적 도메인과 일부 또는 완전한 상동성을 가질 경우에, 상기 $(Q)_m$은 5'-GUUUUAGAGCUA-3'일 수 있고, 또는 5'-GUUUUAGAGCUA-3'와 적어도 50% 이상의 상동성을 가지는 뉴클레오타이드 서열일 수 있다.

[391] 다른 예로, 상기 제 1 상보적 도메인이 캄필로박터 제주니의 제 1 상보적 도메인 또는 캄필로박터 제주니 유래 제 1 상보적 도메인과 일부 또는 완전한 상동성을 가질 경우에, 상기 $(Q)_m$은 5'-GUUUUAGUCCCUUUUUAAAUUUCUU-3' 또는 5'-GUUUUAGUCCCUU-3' 일 수 있고, 또는 5'-GUUUUAGUCCCUUUUUAAAUUUCUU-3' 또는 5'-GUUUUAGUCCCUU-3'와 적어도 50% 이상의 상동성을 가지는 뉴클레오타이드 서열일 수 있다.

[392] 또 다른 예로, 상기 제 1 상보적 도메인이 스트렙토코커스 써모필러스의 제 1 상보적 도메인 또는 스트렙토코커스 써모필러스 유래 제 1 상보적 도메인과 일부 또는 완전한 상동성을 가질 경우에, 상기 $(Q)_m$은 5'-GUUUUAGAGCUGUGUUGUUUCG-3'일 수 있고, 또는 5'-GUUUUAGAGCUGUGUUGUUUCG-3'와 적어도 50% 이상의 상동성을 가지는 뉴클레오타이드 서열일 수 있다.

[393] 또한, 상기 $(L)_j$는 연결 도메인을 포함하는 뉴클레오타이드 서열로, 제 1 상보적

도메인과 제 2 상보적 도메인을 연결하여 단일가닥 gRNA을 생성할 수 있도록 하는 뉴클레오타이드 서열이다. 이때, 상기 L은 A, U, C 및 G로 이루어진 군에서 각각 독립적으로 선택될 수 있으며, 상기 j은 뉴클레오타이드 서열의 개수로, 1 내지 30의 정수일 수 있다.

[394] 상기 $(Z)_h$는 제 2 상보적 도메인을 포함하는 뉴클레오타이드 서열로, 제 1 상보적 도메인과 상보적 결합을 할 수 있는 뉴클레오타이드 서열을 포함한다. 상기 $(Z)_h$은 자연에 존재하는 종의 제 2 상보적 도메인과 일부 또는 완전한 상동성을 가지는 서열일 수 있으며, 유래된 종에 따라 상기 제 2 상보적 도메인의 뉴클레오타이드 서열은 변경될 수 있다. 상기 Z는 A, U, C 및 G로 이루어진 군에서 각각 독립적으로 선택될 수 있으며, 상기 h은 뉴클레오타이드 서열의 개수로, 5 내지 50의 정수일 수 있다.

[395] 예를 들어, 상기 제 2 상보적 도메인이 스트렙토코커스 피요게네스의 제 2 상보적 도메인 또는 스트렙토코커스 피요게네스 유래 제 2 상보적 도메인과 일부 또는 완전한 상동성을 가질 경우에, 상기 $(Z)_h$은 5'-UAGCAAGUUAAAAU-3'일 수 있고, 또는 5'-UAGCAAGUUAAAAU-3'와 적어도 50% 이상의 상동성을 가지는 뉴클레오타이드 서열일 수 있다.

[396] 다른 예로, 상기 제 2 상보적 도메인이 캄필로박터 제주니의 제 2 상보적 도메인 또는 캄필로박터 제주니 유래 제 2 상보적 도메인과 일부 또는 완전한 상동성을 가질 경우에, 상기 $(Z)_h$은 5'-AAGAAAUUUAAAAAGGGACUAAAAU-3' 또는 5'-AAGGGACUAAAAU-3'일 수 있고, 또는 5'-AAGAAAUUUAAAAAGGGACUAAAAU-3' 또는 5'-AAGGGACUAAAAU-3'와 적어도 50% 이상의 상동성을 가지는 뉴클레오타이드 서열일 수 있다.

[397] 또 다른 예로, 상기 제 2 상보적 도메인이 스트렙토코커스 써모필러스의 제 2 상보적 도메인 또는 스트렙토코커스 써모필러스 유래 제 2 상보적 도메인과 일부 또는 완전한 상동성을 가질 경우에, 상기 $(Z)_h$은 5'-CGAAACAACACAGCGAGUUAAAAU-3'일 수 있고, 또는 5'-CGAAACAACACAGCGAGUUAAAAU-3'와 적어도 50% 이상의 상동성을 가지는 뉴클레오타이드 서열일 수 있다.

[398] 상기 $(P)_k$는 근위 도메인을 포함하는 뉴클레오타이드 서열로, 자연에 존재하는 종의 근위 도메인과 일부 또는 완전한 상동성을 가지는 서열일 수 있으며, 유래된 종에 따라 상기 근위 도메인의 뉴클레오타이드 서열은 변경될 수 있다. 상기 P는 A, U, C 및 G로 이루어진 군에서 각각 독립적으로 선택될 수 있으며, 상기 k은 뉴클레오타이드 서열의 개수로, 1 내지 20의 정수일 수 있다.

[399] 예를 들어, 상기 근위 도메인이 스트렙토코커스 피요게네스의 근위 도메인 또는 스트렙토코커스 피요게네스 유래 근위 도메인과 일부 또는 완전한 상동성을 가질 경우에, 상기 $(P)_k$는 5'-AAGGCUAGUCCG-3'일 수 있고, 또는 5'-AAGGCUAGUCCG-3'와 적어도 50% 이상의 상동성을 가지는 뉴클레오타이드 서열일 수 있다.

[400] 다른 예로, 상기 근위 도메인이 캄필로박터 제주니의 근위 도메인 또는 캄필로박터 제주니 유래 근위 도메인과 일부 또는 완전한 상동성을 가질 경우에, 상기 $(P)_k$는 5'-AAAGAGUUUGC-3'일 수 있고, 또는 5'-AAAGAGUUUGC-3'와 적어도 50% 이상의 상동성을 가지는 뉴클레오타이드 서열일 수 있다.

[401] 또 다른 예로, 상기 근위 도메인이 스트렙토코커스 써모필러스의 근위 도메인 또는 스트렙토코커스 써모필러스 유래 근위 도메인과 일부 또는 완전한 상동성을 가질 경우에, 상기 $(P)_k$는 5'-AAGGCUUAGUCCG-3'일 수 있고, 또는 5'-AAGGCUUAGUCCG-3'와 적어도 50% 이상의 상동성을 가지는 뉴클레오타이드 서열일 수 있다.

[402] 상기 $(F)_i$는 꼬리 도메인을 포함하는 뉴클레오타이드 서열로, 자연에 존재하는 종의 꼬리 도메인과 일부 또는 완전한 상동성을 가지는 서열일 수 있으며, 유래된 종에 따라 상기 꼬리 도메인의 뉴클레오타이드 서열은 변경될 수 있다. 상기 F는 A, U, C 및 G로 이루어진 군에서 각각 독립적으로 선택될 수 있으며, 상기 i은 뉴클레오타이드 서열의 개수로, 1 내지 50의 정수일 수 있다.

[403] 예를 들어, 상기 꼬리 도메인이 스트렙토코커스 피요게네스의 꼬리 도메인 또는 스트렙토코커스 피요게네스 유래 꼬리 도메인과 일부 또는 완전한 상동성을 가질 경우에, 상기 $(F)_i$는 5'-UUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGC-3'일 수 있고, 또는 5'-UUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGC-3'와 적어도 50% 이상의 상동성을 가지는 뉴클레오타이드 서열일 수 있다.

[404] 다른 예로, 상기 꼬리 도메인이 캄필로박터 제주니의 꼬리 도메인 또는 캄필로박터 제주니 유래 꼬리 도메인과 일부 또는 완전한 상동성을 가질 경우에, 상기 $(F)_i$는 5'-GGGACUCUGCGGGGUUACAAUCCCCUAAAACCGCUUUU-3'일 수 있고, 또는 5'-GGGACUCUGCGGGGUUACAAUCCCCUAAAACCGCUUUU-3'와 적어도 50% 이상의 상동성을 가지는 뉴클레오타이드 서열일 수 있다.

[405] 또 다른 예로, 상기 꼬리 도메인이 스트렙토코커스 써모필러스의 꼬리 도메인 또는 스트렙토코커스 써모필러스 유래 꼬리 도메인과 일부 또는 완전한 상동성을 가질 경우에, 상기 $(F)_i$는 5'-UACUCAACUUGAAAAGGUGGCACCGAUUCGGUGUUUUU-3'일 수 있고, 또는 5'-UACUCAACUUGAAAAGGUGGCACCGAUUCGGUGUUUUU-3'와 적어도 50% 이상의 상동성을 가지는 뉴클레오타이드 서열일 수 있다.

[406] 또한, 상기 $(F)_i$는 시험관내 또는 생체내 전사 방법과 관련된 3' 말단에 1 내지 10개의 뉴클레오타이드 서열을 포함할 수 있다.

[407] 예를 들어, T7 프로모터가 gRNA의 시험관내 전사를 위해 사용될 때, 상기 꼬리 도메인은 DNA 주형의 3' 말단에 존재하는 임의의 뉴클레오타이드 서열일 수 있다. 또한, U6 프로모터가 생체내 전사를 위해 사용되는 경우, 상기 꼬리 도메인은 UUUUUU일 수 있으며, H1 프로모터가 전사를 위해 사용되는 경우,

상기 꼬리 도메인은 UUUU일 수 있고, pol-III 프로모터를 사용하는 경우에는, 상기 꼬리 도메인은 여러 개의 우라실 뉴클레오타이드거나 또는 대안될 수 있는 뉴클레오타이드를 포함할 수 있다.

[408] 또한, 상기 $(X)_a$, $(X)_b$, $(X)_c$, $(X)_d$, $(X)_e$ 및 $(X)_f$는 선택적으로 추가할 수 있는 뉴클레오타이드 서열로, 상기 X는 A, U, C 및 G로 이루어진 군에서 각각 독립적으로 선택될 수 있으며, 상기 a, b, c, d, e 및 f는 뉴클레오타이드 서열의 개수로, 0 또는 1 내지 20의 정수일 수 있다.

[409]

[410] ***제 2 단일가닥 gRNA***

[411] 제 2 단일가닥 gRNA는 가이드 도메인, 제 1 상보적 도메인 및 제 2 상보적 도메인으로 구성되는 단일가닥 gRNA일 수 있다.

[412] 이때, 상기 제 2 단일가닥 gRNA는

[413] 5'-[제 2 상보적 도메인]-[제 1 상보적 도메인]-[가이드 도메인]-3'; 또는

[414] 5'-[제 2 상보적 도메인]-[연결 도메인]-[제 1 상보적 도메인]-[가이드 도메인]-3'으로 구성될 수 있다.

[415] 상기 제 2 단일가닥 gRNA는 선택적으로 추가적인 뉴클레오타이드 서열을 포함할 수 있다.

[416]

[417] 일 구체예로서, 상기 제 2 단일가닥 gRNA는

[418] 5'-$(Z)_h$-$(Q)_m$-(N_{target})-3'; 또는

[419] 5'-$(X)_a$-$(Z)_h$-$(X)_b$-$(Q)_m$-$(X)_c$-(N_{target})-3'일 수 있다.

[420] 다른 일 구체예로서, 상기 단일가닥 gRNA는

[421] 5'-$(Z)_h$-$(L)_j$-$(Q)_m$-(N_{target})-3'; 또는

[422] 5'-$(X)_a$-$(Z)_h$-$(L)_j$-$(Q)_m$-$(X)_c$-(N_{target})-3'일 수 있다.

[423] 이때, 상기 N_{target}은 표적 유전자의 전사 조절 영역 내 핵산의 이중 가닥 중 어느 하나 가닥의 일부 서열에 상보적인 뉴클레오타이드 서열로서, 상기 N_{target}은 표적 유전자의 전사 조절 영역 내 핵산 상의 표적 서열에 따라 변할 수 있는 뉴클레오타이드 서열 부위이다.

[424] 상기 $(Q)_m$은 제 1 상보적 도메인을 포함하는 뉴클레오타이드 서열로, 제 2 상보적 도메인과 상보적 결합을 할 수 있는 뉴클레오타이드 서열을 포함한다. 상기 $(Q)_m$은 자연에 존재하는 종의 제 1 상보적 도메인과 일부 또는 완전한 상동성을 가지는 서열일 수 있으며, 유래된 종에 따라 상기 제 1 상보적 도메인의 뉴클레오타이드 서열은 변경될 수 있다. 상기 Q는 A, U, C 및 G로 이루어진 군에서 각각 독립적으로 선택될 수 있으며, 상기 m은 뉴클레오타이드 서열의 개수로, 5 내지 35의 정수일 수 있다.

[425] 예를 들어, 상기 제 1 상보적 도메인이 팔쿠박테리아 박테리움의 제 1 상보적 도메인 또는 팔쿠박테리아 박테리움 유래 제 1 상보적 도메인과 일부 또는 완전한 상동성을 가질 경우에, 상기 $(Q)_m$은 5'-UUUGUAGAU-3'일 수 있고, 또는

5'-UUUGUAGAU-3'와 적어도 50% 이상의 상동성을 가지는 뉴클레오타이드 서열일 수 있다.

[426] 상기 $(Z)_h$는 제 2 상보적 도메인을 포함하는 뉴클레오타이드 서열로, 제 1 상보적 도메인과 상보적 결합을 할 수 있는 뉴클레오타이드 서열을 포함한다. 상기 $(Z)_h$은 자연에 존재하는 종의 제 2 상보적 도메인과 일부 또는 완전한 상동성을 가지는 서열일 수 있으며, 유래된 종에 따라 상기 제 2 상보적 도메인의 뉴클레오타이드 서열은 변경될 수 있다. 상기 Z는 A, U, C 및 G로 이루어진 군에서 각각 독립적으로 선택될 수 있으며, 상기 h은 뉴클레오타이드 서열의 개수로, 5 내지 50의 정수일 수 있다.

[427] 예를 들어, 상기 제 2 상보적 도메인이 팔쿠박테리아 박테리움의 제 2 상보적 도메인 또는 팔쿠박테리아 박테리움 유래 제 2 상보적 도메인과 일부 또는 완전한 상동성을 가질 경우에, 상기 $(Z)_h$은 5'-AAAUUUCUACU-3'일 수 있고, 또는 5'-AAAUUUCUACU-3'와 적어도 50% 이상의 상동성을 가지는 뉴클레오타이드 서열일 수 있다.

[428] 또한, 상기 $(L)_j$는 연결 도메인을 포함하는 뉴클레오타이드 서열로, 제 1 상보적 도메인과 제 2 상보적 도메인을 연결하는 뉴클레오타이드 서열이다. 이때, 상기 L은 A, U, C 및 G로 이루어진 군에서 각각 독립적으로 선택될 수 있으며, 상기 j은 뉴클레오타이드 서열의 개수로, 1 내지 30의 정수일 수 있다.

[429] 또한, 상기 $(X)_a$, $(X)_b$ 및 $(X)_c$는 선택적으로 추가할 수 있는 뉴클레오타이드 서열로, 상기 X는 A, U, C 및 G로 이루어진 군에서 각각 독립적으로 선택될 수 있으며, 상기 a, b 및 c는 뉴클레오타이드 서열의 개수로, 0 또는 1 내지 20의 정수일 수 있다.

[430]

[431] **본 명세서에 의해 개시되는 내용의 일 구현예에서, 가이드핵산은 중복 유전자의 전사 조절 영역 내 표적 서열에 상보적으로 결합하는 gRNA일 수 있다.**

[432] "표적 서열"은 표적 유전자의 전사 조절 영역 내에 존재하는 뉴클레오타이드 서열로, 구체적으로는 표적 유전자의 전사 조절 영역 내에 표적 영역의 일부 뉴클레오타이드 서열이며, 이때 "표적 영역"은 표적 유전자의 전사 조절 영역 내에 가이드핵산-에디터단백질에 의해 변형될 수 있는 부위이다.

[433] 이하에서, 표적 서열이라 함은 두 가지의 뉴클레오타이드 서열 정보 모두를 의미하는 용어로 사용될 수 있다. 예를 들어, 표적 유전자의 경우, 표적 서열은 표적 유전자 DNA의 transcribed strand의 서열 정보를 의미하는 것일 수도 있고, 또는 non-transcribed strand의 뉴클레오타이드 서열 정보를 의미하는 것일 수도 있다.

[434] 예를 들어, 표적 서열은 표적 유전자 A의 표적 영역 중 일부 뉴클레오타이드 서열(transcribed strand)인 5'-ATCATTGGCAGACTAGTTCG-3'을 의미할 수도 있으며, 이에 상보적인 뉴클레오타이드 서열(non-transcribed strand)인 5'-CGAACTAGTCTGCCAATGAT-3' 을 의미할 수도 있다.

[435] 표적 서열은 5 내지 50개의 뉴클레오타이드 서열일 수 있다.

[436] 일 구체예로서 상기 표적 서열은 16개의 뉴클레오타이드 서열, 17개의 뉴클레오타이드 서열, 18개의 뉴클레오타이드 서열, 19개의 뉴클레오타이드 서열, 20개의 뉴클레오타이드 서열, 21개의 뉴클레오타이드 서열, 22개의 뉴클레오타이드 서열, 23개의 뉴클레오타이드 서열, 24개의 뉴클레오타이드 서열 또는 25개의 뉴클레오타이드 서열일 수 있다.

[437] 표적 서열은 가이드핵산 결합 서열 혹은 가이드핵산 비결합 서열을 포함한다.

[438] "가이드핵산 결합 서열"은 가이드핵산의 가이드 도메인에 포함되는 가이드 서열과 일부 또는 완전한 상보성을 가지는 뉴클레오타이드 서열로, 가이드핵산의 가이드 도메인에 포함되는 가이드 서열과 상보적인 결합을 할 수 있다. 표적 서열 및 가이드핵산 결합 서열은 표적 유전자의 전사 조절 영역에 따라, 즉 유전자 조작 또는 교정하고자 하는 대상에 따라 달라질 수 있는 뉴클레오타이드 서열로, 표적 유전자의 전사 조절 영역 내 핵산 서열에 따라 다양하게 설계될 수 있다.

[439] "가이드핵산 비결합 서열"은 가이드핵산의 가이드 도메인에 포함되는 가이드 서열과 일부 또는 완전한 상동성을 가지는 뉴클레오타이드 서열로, 가이드핵산의 가이드 도메인에 포함되는 가이드 서열과 상보적인 결합을 할 수 없다. 또한, 가이드핵산 비결합 서열은 가이드핵산 결합 서열과 상보성을 가지는 뉴클레오타이드 서열로, 가이드핵산 결합 서열과 상보적인 결합을 할 수 있다.

[440] 가이드핵산 결합 서열은 표적 서열 중 일부 뉴클레오타이드 서열로, 표적 서열의 두 가지 서로 다른 서열순서를 가지는 뉴클레오타이드 서열, 즉, 서로 상보적인 결합을 할 수 있는 두 가지의 뉴클레오타이드 서열 중 한 가지 뉴클레오타이드 서열일 수 있다. 이때, 가이드핵산 비결합 서열은 표적 서열 중 가이드핵산 결합 서열을 제외한 나머지 뉴클레오타이드 서열일 수 있다.

[441] 예를 들면, 표적 유전자 A의 전사 조절 영역 내에 표적 영역 중 일부 뉴클레오타이드 서열인 5'-ATCATTGGCAGACTAGTTCG-3'과 이에 상보적인 뉴클레오타이드 서열인 5'-CGAACTAGTCTGCCAATGAT-3'을 표적 서열로 할 때, 가이드핵산 결합 서열은 두 개의 표적 서열 중 하나, 즉, 5'-ATCATTGGCAGACTAGTTCG-3' 또는 5'-CGAACTAGTCTGCCAATGAT-3'일 수 있다. 이때, 가이드핵산 비결합 서열은, 가이드핵산 결합 서열이 5'-ATCATTGGCAGACTAGTTCG-3'인 경우, 5'-CGAACTAGTCTGCCAATGAT-3'일 수 있고, 또는 가이드핵산 결합 서열이 5'-CGAACTAGTCTGCCAATGAT-3'인 경우 가이드핵산 비결합 서열은 5'-ATCATTGGCAGACTAGTTCG-3'일 수 있다.

[442] 가이드핵산 결합 서열은 표적 서열, 즉, transcribed strand와 동일한 뉴클레오타이드 서열 및 non-transcribed strand와 동일한 뉴클레오타이드 서열 중 선택된 하나의 뉴클레오타이드 서열일 수 있다. 이때, 가이드핵산 비결합 서열은 표적 서열 중 가이드핵산 결합 서열, 즉, transcribed strand와 동일한

뉴클레오타이드 서열 및 non-transcribed strand와 동일한 뉴클레오타이드 서열 중 선택된 하나의 뉴클레오타이드 서열을 제외한 나머지 뉴클레오타이드 서열일 수 있다.

[443] 가이드핵산 결합 서열은 표적 서열의 길이와 동일할 수 있다.

[444] 가이드핵산 비결합 서열은 표적 서열 또는 가이드핵산 결합 서열의 길이와 동일할 수 있다.

[445] 가이드핵산 결합 서열은 5 내지 50개의 뉴클레오타이드 서열일 수 있다.

[446] 일 구체예로서 상기 가이드핵산 결합 서열은 16개의 뉴클레오타이드 서열, 17개의 뉴클레오타이드 서열, 18개의 뉴클레오타이드 서열, 19개의 뉴클레오타이드 서열, 20개의 뉴클레오타이드 서열, 21개의 뉴클레오타이드 서열, 22개의 뉴클레오타이드 서열, 23개의 뉴클레오타이드 서열, 24개의 뉴클레오타이드 서열 또는 25개의 뉴클레오타이드 서열일 수 있다.

[447] 가이드핵산 비결합 서열은 5 내지 50개의 뉴클레오타이드 서열일 수 있다.

[448] 일 구체예로서 상기 가이드핵산 비결합 서열은 16개의 뉴클레오타이드 서열, 17개의 뉴클레오타이드 서열, 18개의 뉴클레오타이드 서열, 19개의 뉴클레오타이드 서열, 20개의 뉴클레오타이드 서열, 21개의 뉴클레오타이드 서열, 22개의 뉴클레오타이드 서열, 23개의 뉴클레오타이드 서열, 24개의 뉴클레오타이드 서열 또는 25개의 뉴클레오타이드 서열일 수 있다.

[449] 가이드핵산 결합 서열은 가이드핵산의 가이드 도메인에 포함된 가이드 서열과 일부 또는 완전한상보적인 결합을 할 수 있으며, 상기 가이드핵산 결합 서열의 길이는 가이드 서열의 길이와 동일할 수 있다.

[450] 상기 가이드핵산 결합 서열은 가이드핵산의 가이드 도메인에 포함된 가이드 서열에 상보적인 뉴클레오타이드 서열일 수 있으며, 예를 들어 최소한 70%, 75%, 80%, 85%, 90% 또는 95% 이상의 상보적이거나 또는 완전하게 상보적인 뉴클레오타이드 서열일 수 있다.

[451] 일 예로, 상기 가이드핵산 결합 서열은 가이드핵산의 가이드 도메인에 포함된 가이드 서열에 상보적이지 않은 1 내지 8개의 뉴클레오타이드 서열을 가지거나 또는 포함할 수 있다.

[452] 가이드핵산 비결합 서열은 가이드핵산의 가이드 도메인에 포함된 가이드 서열과 일부 또는 완전한 상동성을 가질 수 있으며, 상기 가이드핵산 비결합 서열의 길이는 가이드 서열의 길이와 동일할 수 있다.

[453] 상기 가이드핵산 비결합 서열은 가이드핵산의 가이드 도메인에 포함된 가이드 서열에 상동성을 가진 뉴클레오타이드 서열일 수 있으며, 예를 들어 최소한 70%, 75%, 80%, 85%, 90% 또는 95% 이상의 상동성이거나 또는 완전하게 상동성인 뉴클레오타이드 서열일 수 있다.

[454] 일 예로, 상기 가이드핵산 비결합 서열은 가이드핵산의 가이드 도메인에 포함된 가이드 서열에 상동적이 않은 1 내지 8개의 뉴클레오타이드 서열을 가지거나 포함할 수 있다.

[455] 가이드핵산 비결합 서열은 가이드핵산 결합 서열과 상보적 결합을 할 수 있으며, 상기 가이드핵산 비결합 서열은 가이드핵산 결합 서열의 길이와 동일할 수 있다.

[456] 상기 가이드핵산 비결합 서열은 가이드핵산 결합서열에 상보적인 뉴클레오타이드 서열일 수 있으며, 예를 들어 최소한 90% 또는 95% 이상의 상보적이거나 또는 완전하게 상보적인 뉴클레오타이드 서열일 수 있다.

[457] 일 예로, 상기 가이드핵산 비결합 서열은 가이드핵산 결합 서열에 상보적이지 않은 1 내지 2개의 뉴클레오타이드 서열을 가지거나 포함할 수 있다.

[458] 또한, 상기 가이드핵산 결합 서열은 에디터단백질이 인식할 수 있는 뉴클레오타이드 서열에 근접한 위치에 위치한 뉴클레오타이드 서열일 수 있다.

[459] 일 예로, 상기 가이드핵산 결합 서열은 에디터단백질이 인식할 수 있는 뉴클레오타이드 서열의 5' 말단 또는/및 3' 말단에 인접하여 위치하는 연속하는 5 내지 50개의 뉴클레오타이드 서열일 수 있다.

[460] 또한, 상기 가이드핵산 비결합 서열은 에디터단백질이 인식할 수 있는 뉴클레오타이드 서열에 근접한 위치에 뉴클레오타이드 서열일 수 있다.

[461] 일 예로, 상기 가이드핵산 비결합 서열은 에디터단백질이 인식할 수 있는 뉴클레오타이드 서열의 5' 말단 또는/및 3' 말단에 인접하여 위치하는 연속하는 5 내지 50개의 뉴클레오타이드 서열일 수 있다.

[462] "표적화(targeting)"는 표적 유전자의 전사 조절 영역 내에 존재하는 표적 서열 중 가이드핵산 결합 서열과 상보적 결합을 하는 것을 의미한다. 이때, 상기 상보적 결합은 100%의 완전한 상보적 결합일 수 있고, 또는 70% 이상 100% 미만의 불완전한 상보적 결합일 수 있다. 따라서, "표적화하는 gRNA"는 표적 유전자의 전사 조절 영역 내에 존재하는 표적 서열 중 가이드핵산 결합 서열과 상보적 결합을 하는 gRNA를 의미한다.

[463] 본 명세서에 의해 개시되는 표적 유전자는 중복 유전자일 수 있다.

[464] 본 명세서에 의해 개시되는 표적 유전자는 PMP22 유전자, PLP1 유전자, MECP2 유전자, SOX3 유전자, RAI1 유전자, TBX1 유전자, ELN 유전자, JAGGED1 유전자, NSD1 유전자, MMP23 유전자, LMB1 유전자, SNCA 유전자 및/또는 APP 유전자일 수 있다.

[465] 본 명세서에 의해 개시되는 표적 유전자는 암유전자(oncogene)일 수 있다.

[466] 이때, 상기 암유전자는 MYC 유전자, ERBB2(HER2) 유전자, CCND1(Cyclin D1) 유전자, FGFR1 유전자, FGFR2 유전자, HRAS 유전자, KRAS 유전자, MYB 유전자, MDM2 유전자, CCNE(Cyclin E) 유전자, MET 유전자, CDK4 유전자, ERBB1 유전자, MYCN 유전자 및/또는 AKT2 유전자일 수 있다.

[467]

[468] 구현예에서,

[469] 본 명세서에 의해 개시되는 표적 서열은 중복 유전자의 프로모터 영역에 위치하는 연속하는 10 내지 35개의 뉴클레오타이드 서열일 수 있다.

[470] 이때, 상기 중복 유전자는 PMP22 유전자, PLP1 유전자, MECP2 유전자, SOX3 유전자, RAI1 유전자, TBX1 유전자, ELN 유전자, JAGGED1 유전자, NSD1 유전자, MMP23 유전자, LMB1 유전자, SNCA 유전자 및 APP 유전자로 구성된 군에서 선택된 하나 이상의 유전자일 있다.

[471] 또는, 상기 중복 유전자는 MYC 유전자, ERBB2(HER2) 유전자, CCND1(Cyclin D1) 유전자, FGFR1 유전자, FGFR2 유전자, HRAS 유전자, KRAS 유전자, MYB 유전자, MDM2 유전자, CCNE(Cyclin E) 유전자, MET 유전자, CDK4 유전자, ERBB1 유전자, MYCN 유전자 및 AKT2 유전자로 구성된 군에서 선택된 하나 이상의 유전자일 수 있다.

[472] 상기 표적 서열은 10 내지 35개의 뉴클레오타이드 서열, 15 내지 35개의 뉴클레오타이드 서열, 20 내지 35개의 뉴클레오타이드 서열, 25 내지 35개의 뉴클레오타이드 서열 또는 30 내지 35개의 뉴클레오타이드 서열일 수 있다.

[473] 또는 상기 표적 서열은 10 내지 15개의 뉴클레오타이드 서열, 15 내지 20개의 뉴클레오타이드 서열, 20 내지 25개의 뉴클레오타이드 서열, 25 내지 30개의 뉴클레오타이드 서열 또는 30 내지 35개의 뉴클레오타이드 서열일 수 있다.

[474] 일 예로, 상기 표적 서열은 중복 유전자의 코어 프로모터 영역에 위치한 연속하는 10 내지 25개의 뉴클레오타이드 서열일 수 있다.

[475] 예를 들어, 상기 표적 서열은 중복 유전자의 TSS을 포함하거나 또는 근접한 영역에 위치한 연속하는 10 내지 25개의 뉴클레오타이드 서열일 수 있다.

[476] 예를 들어, 상기 표적 서열은 중복 유전자의 RNA 폴리머라아제(RNA polymerase) 결합 사이트를 포함하거나 또는 근접한 영역에 위치한 연속하는 10 내지 25개의 뉴클레오타이드 서열일 수 있다.

[477] 예를 들어, 상기 표적 서열은 중복 유전자의 전사인자 결합 사이트를 포함하거나 또는 근접한 영역에 위치한 연속하는 10 내지 25개의 뉴클레오타이드 서열일 수 있다.

[478] 예를 들어, 상기 표적 서열은 중복 유전자의 TATA 박스를 포함하거나 또는 근접한 영역에 위치한 연속하는 10 내지 25개의 뉴클레오타이드 서열일 수 있다.

[479] 예를 들어, 상기 표적 서열은 중복 유전자의 코어 프로모터 영역 내에 존재하는 5'-TATA(A/T)A(A/T)-3' 서열의 전체 또는 일부을 포함하는 연속하는 10 내지 25개의 뉴클레오타이드 서열일 수 있다.

[480] 예를 들어, 상기 표적 서열은 중복 유전자의 코어 프로모터 영역 내에 존재하는 5'-TATA(A/T)A(A/T)(A/G)-3' 서열의 전체 또는 일부을 포함하는 연속하는 10 내지 25개의 뉴클레오타이드 서열일 수 있다.

[481] 예를 들어, 상기 표적 서열은 중복 유전자의 코어 프로모터 영역 내에 존재하는 5'-CATAAAA-3' 서열, 5'-CATAAAA-3' 서열, 5'-TATAA-3' 서열, 5'-TATAAAA-3' 서열, 5'-CATAAATA-3' 서열, 5'-TATATAA-3' 서열, 5'-TATATATATATATAA-3' 서열, 5'-TATATTATA-3' 서열, 5'-TATAAA-3' 서열, 5'-TATAAAATA-3' 서열, 5'-TATATA-3' 서열, 5'-GATTAAAAA-3' 서열, 5'-TATAAAAA-3' 서열,

5'-TTATAA-3' 서열, 5'-TTTTAAAA-3' 서열, 5'-TCTTTAAAA-3' 서열, 5'-GACATTTAA-3' 서열, 5'-TGATATCAA-3' 서열, 5'-TATAAATA-3' 서열, 5'-TATAAGA-3' 서열, 5'-AATAAA-3' 서열, 5'-TTTATA-3' 서열, 5'-CATAAAAA-3' 서열, 5'-TATACA-3' 서열, 5'-TTTAAGA-3' 서열, 5'-GATAAAG-3' 서열, 5'-TATAACA-3' 서열, 5'-TCTTATCTT-3' 서열, 5'-TTGTACTTT-3' 서열, 5'-CATATAA-3' 서열, 5'-TATAAAT-3' 서열, 5'-TATATATAAAAAAAA-3' 서열 및 5'-CATAAATAAAAAAAATTA-3' 서열로 구성된 군에서 선택된 하나 이상의 서열의 전체 또는 일부을 포함하는 연속하는 10 내지 25개의 뉴클레오타이드 서열일 수 있다.

[482] 예를 들어, 상기 표적 서열은 중복 유전자의 코어 프로모터 영역 내에 존재하는 TATA-결합 단백질(TATA-binding protein, TBP)이 결합하는 핵산 서열의 전체 또는 일부를 포함하는 연속하는 10 내지 25개의 뉴클레오타이드 서열일 수 있다.

[483] 다른 일 예로, 상기 표적 서열은 중복 유전자의 근접 프로모터 영역에 위치한 연속하는 10 내지 25개의 뉴클레오타이드 서열일 수 있다.

[484] 예를 들어, 상기 표적 서열은 중복 유전자의 TSS의 1 내지 300bp 상류(upstream) 영역에 위치한 연속하는 10 내지 25개의 뉴클레오타이드 서열일 수 있다.

[485] 또 다른 일 예로, 상기 표적 서열은 중복 유전자의 원위 프로모터 영역에 위치한 연속하는 10 내지 25개의 뉴클레오타이드 서열일 수 있다.

[486]

[487] 본 명세서에 의해 개시되는 표적 서열은 중복 유전자의 인핸서 영역에 위치하는 연속하는 10 내지 35개의 뉴클레오타이드 서열일 수 있다.

[488] 이때, 상기 중복 유전자는 PMP22 유전자, PLP1 유전자, MECP2 유전자, SOX3 유전자, RAI1 유전자, TBX1 유전자, ELN 유전자, JAGGED1 유전자, NSD1 유전자, MMP23 유전자, LMB1 유전자, SNCA 유전자 및 APP 유전자로 구성된 군에서 선택된 하나 이상의 유전자일 있다.

[489] 또는, 상기 중복 유전자는 MYC 유전자, ERBB2(HER2) 유전자, CCND1(Cyclin D1) 유전자, FGFR1 유전자, FGFR2 유전자, HRAS 유전자, KRAS 유전자, MYB 유전자, MDM2 유전자, CCNE(Cyclin E) 유전자, MET 유전자, CDK4 유전자, ERBB1 유전자, MYCN 유전자 및 AKT2 유전자로 구성된 군에서 선택된 하나 이상의 유전자일 수 있다.

[490] 상기 표적 서열은 10 내지 35개의 뉴클레오타이드 서열, 15 내지 35개의 뉴클레오타이드 서열, 20 내지 35개의 뉴클레오타이드 서열, 25 내지 35개의 뉴클레오타이드 서열 또는 30 내지 35개의 뉴클레오타이드 서열일 수 있다.

[491] 또는 상기 표적 서열은 10 내지 15개의 뉴클레오타이드 서열, 15 내지 20개의 뉴클레오타이드 서열, 20 내지 25개의 뉴클레오타이드 서열, 25 내지 30개의 뉴클레오타이드 서열 또는 30 내지 35개의 뉴클레오타이드 서열일 수 있다.

[492] 예를 들어, 상기 표적 서열은 중복 유전자의 인핸서-박스(E-box)를 포함하거나 또는 근접한 영역에 위치한 연속하는 10 내지 25개의 뉴클레오타이드 서열일 수

있다.

[493] 예를 들어, 상기 표적 서열은 중복 유전자의 인트론 내 존재하는 인핸서 영역에 위치하는 연속하는 10 내지 35개의 뉴클레오타이드 서열일 수 있다.

[494]

[495] 본 명세서에 의해 개시되는 표적 서열은 중복 유전자의 인슐레이터 영역에 위치하는 연속하는 10 내지 35개의 뉴클레오타이드 서열일 수 있다.

[496] 이때, 상기 중복 유전자는 PMP22 유전자, PLP1 유전자, MECP2 유전자, SOX3 유전자, RAI1 유전자, TBX1 유전자, ELN 유전자, JAGGED1 유전자, NSD1 유전자, MMP23 유전자, LMB1 유전자, SNCA 유전자 및 APP 유전자로 구성된 군에서 선택된 하나 이상의 유전자일 있다.

[497] 또는, 상기 중복 유전자는 MYC 유전자, ERBB2(HER2) 유전자, CCND1(Cyclin D1) 유전자, FGFR1 유전자, FGFR2 유전자, HRAS 유전자, KRAS 유전자, MYB 유전자, MDM2 유전자, CCNE(Cyclin E) 유전자, MET 유전자, CDK4 유전자, ERBB1 유전자, MYCN 유전자 및 AKT2 유전자로 구성된 군에서 선택된 하나 이상의 유전자일 수 있다.

[498] 상기 표적 서열은 10 내지 35개의 뉴클레오타이드 서열, 15 내지 35개의 뉴클레오타이드 서열, 20 내지 35개의 뉴클레오타이드 서열, 25 내지 35개의 뉴클레오타이드 서열 또는 30 내지 35개의 뉴클레오타이드 서열일 수 있다.

[499] 또는 상기 표적 서열은 10 내지 15개의 뉴클레오타이드 서열, 15 내지 20개의 뉴클레오타이드 서열, 20 내지 25개의 뉴클레오타이드 서열, 25 내지 30개의 뉴클레오타이드 서열 또는 30 내지 35개의 뉴클레오타이드 서열일 수 있다.

[500]

[501] 본 명세서에 의해 개시되는 표적 서열은 중복 유전자의 사일런서 영역에 위치하는 연속하는 10 내지 35개의 뉴클레오타이드 서열일 수 있다.

[502] 이때, 상기 중복 유전자는 PMP22 유전자, PLP1 유전자, MECP2 유전자, SOX3 유전자, RAI1 유전자, TBX1 유전자, ELN 유전자, JAGGED1 유전자, NSD1 유전자, MMP23 유전자, LMB1 유전자, SNCA 유전자 및 APP 유전자로 구성된 군에서 선택된 하나 이상의 유전자일 있다.

[503] 또는, 상기 중복 유전자는 MYC 유전자, ERBB2(HER2) 유전자, CCND1(Cyclin D1) 유전자, FGFR1 유전자, FGFR2 유전자, HRAS 유전자, KRAS 유전자, MYB 유전자, MDM2 유전자, CCNE(Cyclin E) 유전자, MET 유전자, CDK4 유전자, ERBB1 유전자, MYCN 유전자 및 AKT2 유전자로 구성된 군에서 선택된 하나 이상의 유전자일 수 있다.

[504] 상기 표적 서열은 10 내지 35개의 뉴클레오타이드 서열, 15 내지 35개의 뉴클레오타이드 서열, 20 내지 35개의 뉴클레오타이드 서열, 25 내지 35개의 뉴클레오타이드 서열 또는 30 내지 35개의 뉴클레오타이드 서열일 수 있다.

[505] 또는 상기 표적 서열은 10 내지 15개의 뉴클레오타이드 서열, 15 내지 20개의 뉴클레오타이드 서열, 20 내지 25개의 뉴클레오타이드 서열, 25 내지 30개의

뉴클레오타이드 서열 또는 30 내지 35개의 뉴클레오타이드 서열일 수 있다.

[506]

[507] 본 명세서에 의해 개시되는 표적 서열은 중복 유전자의 로커스 컨트롤 영역(LCR)에 위치하는 연속하는 10 내지 35개의 뉴클레오타이드 서열일 수 있다.

[508] 이때, 상기 중복 유전자는 PMP22 유전자, PLP1 유전자, MECP2 유전자, SOX3 유전자, RAI1 유전자, TBX1 유전자, ELN 유전자, JAGGED1 유전자, NSD1 유전자, MMP23 유전자, LMB1 유전자, SNCA 유전자 및 APP 유전자로 구성된 군에서 선택된 하나 이상의 유전자일 있다.

[509] 또는, 상기 중복 유전자는 MYC 유전자, ERBB2(HER2) 유전자, CCND1(Cyclin D1) 유전자, FGFR1 유전자, FGFR2 유전자, HRAS 유전자, KRAS 유전자, MYB 유전자, MDM2 유전자, CCNE(Cyclin E) 유전자, MET 유전자, CDK4 유전자, ERBB1 유전자, MYCN 유전자 및 AKT2 유전자로 구성된 군에서 선택된 하나 이상의 유전자일 수 있다.

[510] 상기 표적 서열은 10 내지 35개의 뉴클레오타이드 서열, 15 내지 35개의 뉴클레오타이드 서열, 20 내지 35개의 뉴클레오타이드 서열, 25 내지 35개의 뉴클레오타이드 서열 또는 30 내지 35개의 뉴클레오타이드 서열일 수 있다.

[511] 또는 상기 표적 서열은 10 내지 15개의 뉴클레오타이드 서열, 15 내지 20개의 뉴클레오타이드 서열, 20 내지 25개의 뉴클레오타이드 서열, 25 내지 30개의 뉴클레오타이드 서열 또는 30 내지 35개의 뉴클레오타이드 서열일 수 있다.

[512]

[513] 본 명세서에 의해 개시되는 표적 서열은 중복 유전의 전자 조절 영역 내의 위치한 PAM(proto-spacer-adjacent Motif) 서열의 5' 말단 및/또는 3' 말단에 인접한 연속하는 10 내지 35개의 뉴클레오타이드 서열일 수 있다.

[514] 이때, 상기 중복 유전자는 PMP22 유전자, PLP1 유전자, MECP2 유전자, SOX3 유전자, RAI1 유전자, TBX1 유전자, ELN 유전자, JAGGED1 유전자, NSD1 유전자, MMP23 유전자, LMB1 유전자, SNCA 유전자 및 APP 유전자로 구성된 군에서 선택된 하나 이상의 유전자일 있다.

[515] 또는, 상기 중복 유전자는 MYC 유전자, ERBB2(HER2) 유전자, CCND1(Cyclin D1) 유전자, FGFR1 유전자, FGFR2 유전자, HRAS 유전자, KRAS 유전자, MYB 유전자, MDM2 유전자, CCNE(Cyclin E) 유전자, MET 유전자, CDK4 유전자, ERBB1 유전자, MYCN 유전자 및 AKT2 유전자로 구성된 군에서 선택된 하나 이상의 유전자일 수 있다.

[516] 이때, 상기 중복 유전자의 전자 조절 영역은 중복 유전자의 프로모터, 인핸서, 사일런서, 인슐레이터 또는 로커스 컨트롤 영역(LCR)일 수 있다.

[517] "PAM(proto-spacer-adjacent Motif) 서열"은 에디터단백질이 인식할 수 있는 뉴클레오타이드 서열이다. 이때, PAM 서열은 에디터단백질의 종류 및 유래된 종에 따라 뉴클레오타이드 서열에 차이가 있을 수 있다.

[518] 이때, 상기 PAM 서열은 예를 들어, 하기의 서열 중 1 이상일 수 있다(5'에서 3'방향으로 기재함).

[519] NGG(N은 A, T, C 또는 G임);

[520] NNNNRYAC(N은 각각 독립적으로 A, T, C 또는 G이고, R은 A또는 G이고, Y는 C또는 T임);

[521] NNAGAAW(N은 각각 독립적으로 A, T, C 또는 G이고, W는 A 또는 T임);

[522] NNNNGATT(N은 각각 독립적으로 A, T, C 또는 G임);

[523] NNGRR(T)(N은 각각 독립적으로 A, T, C 또는 G이고, R은 A 또는 G임); 및

[524] TTN(N은 A, T, C 또는 G임).

[525] 상기 표적 서열은 10 내지 35개의 뉴클레오타이드 서열, 15 내지 35개의 뉴클레오타이드 서열, 20 내지 35개의 뉴클레오타이드 서열, 25 내지 35개의 뉴클레오타이드 서열 또는 30 내지 35개의 뉴클레오타이드 서열일 수 있다.

[526] 또는 상기 표적 서열은 10 내지 15개의 뉴클레오타이드 서열, 15 내지 20개의 뉴클레오타이드 서열, 20 내지 25개의 뉴클레오타이드 서열, 25 내지 30개의 뉴클레오타이드 서열 또는 30 내지 35개의 뉴클레오타이드 서열일 수 있다.

[527] 일 구체예로서, 에디터단백질이 인식하는 PAM 서열이 5'-NGG-3', 5'-NAG-3' 또는/및 5'-NGA-3' (N= A, T, G 또는 C; 또는 A, U, G 또는 C)인 경우, 상기 표적 서열은 중복 유전자의 전사 조절 영역 내의 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 10 내지 25개의 뉴클레오타이드 서열일 수 있다.

[528] 다른 일 구체예로서, 에디터단백질이 인식하는 PAM 서열이 5'-NGGNG-3' 또는/및 5'-NNAGAAW-3' (W = A 또는 T이며, N= A, T, G 또는 C; 또는 A, U, G 또는 C)인 경우, 상기 표적 서열은 중복 유전자의 전사 조절 영역 내의 5'-NGGNG-3' 또는/및 5'-NNAGAAW-3' (W = A 또는 T이며, N= A, T, G 또는 C; 또는 A, U, G 또는 C) 서열의 5' 말단 또는/및 3' 말단에 인접하여 위치하는 연속하는 10 내지 25개의 뉴클레오타이드 서열일 수 있다.

[529] 또 다른 일 구체예로서, 에디터단백질이 인식하는 PAM 서열이 5'-NNNNGATT-3' 또는/및 5'-NNNGCTT-3' (N= A, T, G 또는 C; 또는 A, U, G 또는 C)인 경우, 상기 표적 서열은 중복 유전자의 전사 조절 영역 내의 5'-NNNNGATT-3' 또는/및 5'-NNNGCTT-3' (N= A, T, G 또는 C; 또는 A, U, G 또는 C) 서열의 5' 말단 또는/및 3' 말단에 인접하여 위치하는 연속하는 10 내지 25개의 뉴클레오타이드 서열일 수 있다.

[530] 일 구체예로서, 에디터단백질이 인식하는 PAM 서열이 5'-NNNVRYAC-3' (V = G, C 또는 A; R = A 또는 G 이며, Y = C 또는 T 이고, N= A, T, G 또는 C; 또는 A, U, G 또는 C)인 경우, 상기 표적 서열은 중복 유전자의 전사 조절 영역 내의 5'-NNNVRYAC-3' (V = G, C 또는 A; R = A 또는 G 이며, Y = C 또는 T 이고, N= A, T, G 또는 C; 또는 A, U, G 또는 C) 서열의 5' 말단 또는/및 3' 말단에 인접하여

위치하는 연속하는 10 내지 25개의 뉴클레오타이드 서열일 수 있다.

[531] 다른 일 구체예로서, 에디터단백질이 인식하는 PAM 서열이 5'-NAAR-3'(R = A 또는 G이며, N= A, T, G 또는 C; 또는 A, U, G 또는 C)인 경우, 상기 표적 서열은 중복 유전자의 전사 조절 영역 내의 5'-NAAR-3'(R = A 또는 G이며, N= A, T, G 또는 C; 또는 A, U, G 또는 C) 서열의 5' 말단 또는/및 3' 말단에 인접하여 위치하는 연속하는 10 내지 25개의 뉴클레오타이드 서열일 수 있다.

[532] 또 다른 일 구체예로서, 에디터단백질이 인식하는 PAM 서열이 5'-NNGRR-3', 5'-NNGRRT-3' 또는/및 5'-NNGRRV-3' (R = A 또는 G이며, V = G, C 또는 A이고, N= A, T, G 또는 C; 또는 A, U, G 또는 C)인 경우, 상기 표적 서열은 중복 유전자의 전사 조절 영역 내의 5'-NNGRR-3', 5'-NNGRRT-3' 또는/및 5'-NNGRRV-3' (R = A 또는 G이며, V = G, C 또는 A이고, N= A, T, G 또는 C; 또는 A, U, G 또는 C) 서열의 5' 말단 또는/및 3' 말단에 인접하여 위치하는 연속하는 10 내지 25개의 뉴클레오타이드 서열일 수 있다.

[533] 일 구체예로서, 에디터단백질이 인식하는 PAM 서열이 5'-TTN-3' (N= A, T, G 또는 C; 또는 A, U, G 또는 C)인 경우, 상기 표적 서열은 중복 유전자의 전사 조절 영역 내의 5'-TTN-3' (N= A, T, G 또는 C; 또는 A, U, G 또는 C) 서열의 5' 말단 또는/및 3' 말단에 인접하여 위치하는 연속하는 10 내지 25개의 뉴클레오타이드 서열일 수 있다.

[534]

[535] 이하, 본 명세서에 의해 개시되는 일 구체예에서 사용될 수 있는 표적 서열들의 일 예들은 표 1, 표 2, 표 3, 표 4, 표 5 및 표 6에 정리하였으며, 표 1, 표 2, 표 3, 표 4, 표 5 및 표 6에 기재된 표적 서열은 가이드핵산 비결합 서열로, 기재한 서열을 통해 상보적인 서열, 즉, 가이드핵산 결합 서열은 예측될 수 있다. 또한, 표 1, 표 2, 표 3, 표 4, 표 5 및 표 6에 기재된 sgRNA 이름은 에디터단백질에 따라 SpCas9의 경우 Sp를 기재하였고, CjCas9의 경우 Cj를 기재하여 구분하였다.

[536] [표 1]

[537] SpCas9을 위한 인간 PMP22 유전자의 표적 서열(Target sequences of human PMP22 gene for SpCas9)

[538]

sgRNA	No.	Target (5' to 3')	SEQ ID NO
hPMP22-TATA-Sp#1	1	GGACCAGCCCCTGAATAAAC	SEQ ID NO: 1
hPMP22-TATA-Sp#2	2	GGCGTCTTTCCAGTTTATTC	SEQ ID NO: 2
hPMP22-TATA-Sp#3	3	GCGTCTTTCCAGTTTATTCA	SEQ ID NO: 3
hPMP22-TATA-Sp#4	4	CGTCTTTCCAGTTTATTCAG	SEQ ID NO: 4
hPMP22-TATA-Sp#5	5	TTCAGGGGCTGGTCCAATGC	SEQ ID NO: 5
hPMP22-TATA-Sp#6	6	TCAGGGGCTGGTCCAATGCT	SEQ ID NO: 6
hPMP22-TATA-Sp#7	7	ACCATGACATATCCCAGCAT	SEQ ID NO: 7
hPMP22-TATA-Sp#8	8	TTTCCAGTTTATTCAGGGGC	SEQ ID NO: 8
hPMP22-TATA-Sp#9	9	CAGTTACAGGGAGCACCACC	SEQ ID NO: 9
hPMP22-TATA-Sp#10	10	CTGGTCTGGCTTCAGTTACA	SEQ ID NO: 10
hPMP22-TATA-Sp#11	11	CCTGGTCTGGCTTCAGTTAC	SEQ ID NO: 11
hPMP22-TATA-Sp#12	12	AACTGGAAAGACGCCTGGTC	SEQ ID NO: 12
hPMP22-TATA-Sp#13	13	GAATAAACTGGAAAGACGCC	SEQ ID NO: 13
hPMP22-TATA-Sp#14	14	TCCAATGCTGGGATATGTCA	SEQ ID NO: 14
hPMP22-TATA-Sp#15	15	AATGCTGGGATATGTCATGG	SEQ ID NO: 15
hPMP22-TATA-Sp#16	16	ATAGAGGCTGAGAACCTCTC	SEQ ID NO: 16
hPMP22-Enh-Sp#1	17	TTGGGCATGTTTGAGCTGGT	SEQ ID NO: 17
hPMP22-Enh-Sp#2	18	TTTGGGCATGTTTGAGCTGG	SEQ ID NO: 18
hPMP22-Enh-Sp#3	19	GAGCTGGTGGGCGAAGCATA	SEQ ID NO: 19
hPMP22-Enh-Sp#4	20	AGCTGGTGGGCGAAGCATAT	SEQ ID NO: 20
hPMP22-Enh-Sp#5	21	TGGGCGAAGCATATGGGCAA	SEQ ID NO: 21
hPMP22-Enh-Sp#6	22	GGCCTCCATCCTAAACAATG	SEQ ID NO: 22
hPMP22-Enh-Sp#10	23	GGGTTGGGAGGTTTGGGCGT	SEQ ID NO: 23
hPMP22-Enh-Sp#11	24	AGGTTTGGGCGTGGGAGTCC	SEQ ID NO: 24
hPMP22-Enh-Sp#12	25	TTCAGAGACTCAGCTATTT	SEQ ID NO: 25
hPMP22-Enh-Sp#13	26	GGCCACATTGTTTAGGATG	SEQ ID NO: 26
hPMP22-Enh-Sp#14	27	GGCTTTGGGCATGTTTGAG	SEQ ID NO: 27
hPMP22-Enh-Sp#15	28	AACATGCCCAAAGCCCAGC	SEQ ID NO: 28
hPMP22-Enh-Sp#16	29	ACATGCCCAAAGCCCAGCG	SEQ ID NO: 29
hPMP22-CDS-Sp#1	30	CGATGATACTCAGCAACAGG	SEQ ID NO: 30
hPMP22-CDS-Sp#3	31	ATGGACACGCAACTGATCTC	SEQ ID NO: 31

[539]

[540] [표 2]

[541] CjCas9을 위한 인간 PMP22 유전자의 표적 서열(Target sequences of human PMP22 gene for CjCas9)

[542]

sgRNA	No.	Target (5' to 3')	SEQ ID NO
hPMP22-TATA-Cj#1	1	GCCCTCTGAATCTCCAGTCAAT	SEQ ID NO: 32
hPMP22-TATA-Cj#2	2	AATCTCCAGTCAATTCCAACAC	SEQ ID NO: 33
hPMP22-TATA-Cj#3	3	AATTAGGCAATTCTTGTAAAGC	SEQ ID NO: 34
hPMP22-TATA-Cj#4	4	TTAGGCAATTCTTGTAAAGCAT	SEQ ID NO: 35
hPMP22-TATA-Cj#5	5	AAAGCATAGGCACACATCACCC	SEQ ID NO: 36
hPMP22-TATA-Cj#6	6	GCCTGGTCTGGCTTCAGTTACA	SEQ ID NO: 37
hPMP22-TATA-Cj#7	7	GTGTCCAACTTTGTTTGCTTTC	SEQ ID NO: 38
hPMP22-TATA-Cj#8	8	GTATTCTGGAAAGCAAACAAAG	SEQ ID NO: 39
hPMP22-TATA-Cj#9	9	CAGTCTTGGCATCACAGGCTTC	SEQ ID NO: 40
hPMP22-TATA-Cj#10	10	GGACCTCTTGGCTATTACACAG	SEQ ID NO: 41
hPMP22-TATA-Cj#11	11	GGAGCCAGTGCGACCTCTTGGC	SEQ ID NO: 42
hPMP22-Enh-Cj#1	12	TAAATCACAGAGGCAAAGAGTT	SEQ ID NO: 43
hPMP22-Enh-Cj#2	13	TTGCATAGTGCTAGACTGTTTT	SEQ ID NO: 44
hPMP22-Enh-Cj#3	14	GGGTCATGTGTTTTGAAAACAG	SEQ ID NO: 45
hPMP22-Enh-Cj#4	15	CCCAAACCTCCCAACCCACAAC	SEQ ID NO: 46
hPMP22-Enh-Cj#5	16	ACTCAGCTATTTCTGGAATGAC	SEQ ID NO: 47
hPMP22-Enh-Cj#6	17	TCATCGCCTTTGTGAGCTCCAT	SEQ ID NO: 48
hPMP22-Enh-Cj#7	18	CAGACACAGGCTTTGCTCTAGC	SEQ ID NO: 49
hPMP22-Enh-Cj#8	19	CAAAGCCTGTGTCTGGCCACTA	SEQ ID NO: 50
hPMP22-Enh-Cj#9	20	AGCAGTTTGTGCCCACTAGTGG	SEQ ID NO: 51
hPMP22-Enh-Cj#10	21	ATGTCAAGGTATTCCAGCTAAC	SEQ ID NO: 52
hPMP22-Enh-Cj#11	22	GAATAACTGTATCAAAGTTAGC	SEQ ID NO: 53
hPMP22-Enh-Cj#12	23	TTCCTAATTAAGAGGCTTTGTG	SEQ ID NO: 54
hPMP22-Enh-Cj#13	24	GAGCTAGTTTGTCAGGGTCTAG	SEQ ID NO: 55

[543]

[544] [표 3]

[545] SpCas9을 위한 인간 PLP1 유전자의 표적 서열(Target sequences of human PLP1 gene for SpCas9)

[546]

sgRNA	No.	Target (5' to 3')	SEQ ID NO	Direction	Mismatch 0	Mismatch 1	Mismatch 2	location
hPLP1-TATA-Sp-01	1	GACTTTGGGAGCTAATATC TAGG	SEQ ID NO: 56	+	1	0	0	-
hPLP1-wMN1-Sp-01	2	CCCTTTCATCTTCCCATTCG TGG	SEQ ID NO: 57	+	1	0	0	Up
hPLP1-wMN1-Sp-02	3	CCTTTCATCTTCCCATTCGT GGG	SEQ ID NO: 58	+	1	0	0	Up
hPLP1-wMN1-Sp-03	4	CCCACGAATGGGAAGATG AAAGG	SEQ ID NO: 59	-	1	0	0	Up
hPLP1-wMN1-Sp-04	5	CATCTTCCCATTCGTGGGC AAGG	SEQ ID NO: 60	+	1	0	0	Up
hPLP1-wMN1-Sp-05	6	TCTCCACCTTGCCCACGAA TGGG	SEQ ID NO: 61	-	1	0	0	Up
hPLP1-wMN1-Sp-06	7	GTCTCCACCTTGCCCACGA ATGG	SEQ ID NO: 62	-	1	0	0	Up
hPLP1-wMN1-Sp-07	8	CCCAATGCTTGCACATAAA TTGG	SEQ ID NO: 63	+	1	0	0	Up
hPLP1-wMN1-Sp-08	9	CCAATTTATGTGCAAGCAT TGGG	SEQ ID NO: 64	-	1	0	0	Up
hPLP1-wMN1-Sp-09	10	TCCAATTTATGTGCAAGCA TTGG	SEQ ID NO: 65	-	1	0	0	Up
hPLP1-wMN1-Sp-10	11	TGTGCGCGTCTGAAGAGG AGTGG	SEQ ID NO: 66	+	1	0	0	Up
hPLP1-wMN1-Sp-11	12	GTGCGCGTCTGAAGAGGA GTGGG	SEQ ID NO: 67	+	1	0	0	Up
hPLP1-wMN1-Sp-12	13	TGCGCGTCTGAAGAGGAG TGGGG	SEQ ID NO: 68	+	1	0	0	Up
hPLP1-wMN1-Sp-13	14	TAGTCCAGATGCTGTTGCC GTGG	SEQ ID NO: 69	+	1	0	0	Up
hPLP1-wMN1-Sp-14	15	ATTACCACGGCAACAGCAT CTGG	SEQ ID NO: 70	-	1	0	0	Up
hPLP1-wMN1-Sp-15	16	GACACGATTTAGTATTACC ACGG	SEQ ID NO: 71	-	1	0	0	Up
hPLP1-wMN1-Sp-16	17	CTAAATCGTGTCCAAAGAG GAGG	SEQ ID NO: 72	+	1	0	0	Up
hPLP1-wMN1-Sp-17	18	AGGAATCTCAGCCTCCTCT TTGG	SEQ ID NO: 73	-	1	0	0	Up
hPLP1-wMN1-Sp-18	19	GTGGACAAGGTTAACTAAA AAGG	SEQ ID NO: 74	-	1	0	0	Up
hPLP1-wMN1-Sp-19	20	ATAGTCAAATCATGTGGAC AAGG	SEQ ID NO: 75	-	1	0	0	Up
hPLP1-wMN1-Sp-20	21	TGCTGGATAGTCAAATCAT GTGG	SEQ ID NO: 76	-	1	0	0	Up
hPLP1-wMN1-Sp-21	22	ACATGATTTGACTATCCAG CAGG	SEQ ID NO: 77	+	1	0	0	Up
hPLP1-wMN1-Sp-22	23	ATTTGACTATCCAGCAGGC TTGG	SEQ ID NO: 78	+	1	0	0	Up

[547]

hPLP1-wMN1-Sp-23	24	GTCCCGAAGTCTCTGGGGC CTGG	SEQ ID NO: 79	-	1	0	0	Up
hPLP1-wMN1-Sp-24	25	AAAACAGTCCCGAAGTCTC TGGG	SEQ ID NO: 80	-	1	0	0	Up
hPLP1-wMN1-Sp-25	26	GAAAACAGTCCCGAAGTC TCTGG	SEQ ID NO: 81	-	1	0	0	Up
hPLP1-wMN1-Sp-26	27	TATATACCACATTCAAGTG CTGG	SEQ ID NO: 82	-	1	0	0	Up
hPLP1-wMN1-Sp-27	28	TGGATATAACGAAGTTGTG TGGG	SEQ ID NO: 83	-	1	0	0	Down
hPLP1-wMN1-Sp-28	29	ATGGATATAACGAAGTTGT GTGG	SEQ ID NO: 84	-	1	0	0	Down
hPLP1-wMN1-Sp-29	30	ATATGTTTGTTCACCCCAA CAGG	SEQ ID NO: 85	-	1	0	0	Down
hPLP1-wMN1-Sp-30	31	GAAAACTTGAAATCCTGTT GGGG	SEQ ID NO: 86	-	1	0	0	Down
hPLP1-wMN1-Sp-31	32	TAGACATTAGGAGAAACA GAAGG	SEQ ID NO: 87	-	1	0	0	Down
hPLP1-wMN1-Sp-32	33	CTAGCAGTGACATAGACAT TAGG	SEQ ID NO: 88	-	1	0	0	Down
hPLP1-wMN1-Sp-33	34	AGCCACCTGACTTTGATGA AAGG	SEQ ID NO: 89	-	1	0	0	Down
hPLP1-wMN1-Sp-34	35	TGAGAAATGTTATTACTAT ATGG	SEQ ID NO: 90	+	1	0	0	Down
hPLP1-wMN1-Sp-35	36	AGACTGCGAGATGAGAGA GTTGG	SEQ ID NO: 91	-	1	0	0	Down
hPLP1-wMN1-Sp-36	37	CTCGCAGTCTGTACTTAGA CTGG	SEQ ID NO: 92	+	1	0	0	Down
hPLP1-wMN1-Sp-37	38	AATGTCTCTTGAGAGAGCC AAGG	SEQ ID NO: 93	+	1	0	0	Down

[548]

[549] [표 4]

[550] CjCas9을 위한 인간 PLP1 유전자의 표적 서열(Target sequences of human PLP1 gene for CjCas9)

[551]

sgRNA	No.	Target (5' to 3')	SEQ ID NO	Direction	Mismatch 0	Mismatch 1	Mismatch 2	location
hPLP1-wMN1-Cj-01	1	ATGGGAAGATGAAAGGGA AGTAACTGGTAC	SEQ ID NO: 94	-	1	0	0	Up
hPLP1-wMN1-Cj-02	2	ACTTTGATTGTTAAAACTTA TCCTTGGCAC	SEQ ID NO: 95	-	1	0	0	Up
hPLP1-wMN1-Cj-03	3	AGTCCTACCTCAGCTTCCC AATGCTTGCAC	SEQ ID NO: 96	+	1	0	0	Up
hPLP1-wMN1-Cj-04	4	CAATGCTTGCACATAAATT GGAATGTGTAC	SEQ ID NO: 97	+	1	0	0	Up
hPLP1-wMN1-Cj-05	5	ACACAGAGAGAGACAGAA TGAATGATGTAC	SEQ ID NO: 98	-	1	0	0	Up
hPLP1-wMN1-Cj-06	6	TCCTCTTCAGACGCGCACA CACACACACAC	SEQ ID NO: 99	-	1	0	0	Up
hPLP1-wMN1-Cj-07	7	ACTCCTCTTCAGACGCGCA CACACACACAC	SEQ ID NO: 100	-	1	0	0	Up
hPLP1-wMN1-Cj-08	8	CCACTCCTCTTCAGACGCG CACACACACAC	SEQ ID NO: 101	-	1	0	0	Up
hPLP1-wMN1-Cj-09	9	CCCCACTCCTCTTCAGACG CGCACACACAC	SEQ ID NO: 102	-	1	0	0	Up
hPLP1-wMN1-Cj-10	10	CTCCCCACTCCTCTTCAGA CGCGCACACAC	SEQ ID NO: 103	-	1	0	0	Up
hPLP1-wMN1-Cj-11	11	TACTCCCCACTCCTCTTCAG ACGCGCACAC	SEQ ID NO: 104	-	1	0	0	Up
hPLP1-wMN1-Cj-12	12	TATACTCCCCACTCCTCTTC AGACGCGCAC	SEQ ID NO: 105	-	1	0	0	Up
hPLP1-wMN1-Cj-13	13	ACAGCATCTGGACTATCTT GTTTCCTATAC	SEQ ID NO: 106	-	1	0	0	Up
hPLP1-wMN1-Cj-14	14	ATAGTCCAGATGCTGTTGC CGTGGTAATAC	SEQ ID NO: 107	+	1	0	0	Up
hPLP1-wMN1-Cj-15	15	AAAAGGAATCTCAGCCTCC TCTTTGGACAC	SEQ ID NO: 108	-	1	0	0	Up
hPLP1-wMN1-Cj-16	16	TGTCACTGCTAGTGTGCTT AATTCTTGTAC	SEQ ID NO: 109	+	1	0	0	Down
hPLP1-wMN1-Cj-17	17	ATGTGAATTCAGTACAAGA ATTAAGCACAC	SEQ ID NO: 110	-	1	0	0	Down
hPLP1-wMN1-Cj-18	18	TTATGTGAATTCAGTACAA GAATTAAGCAC	SEQ ID NO: 111	-	1	0	0	Down
hPLP1-wMN1-Cj-19	19	CTTTCATTTCTGTTTATGTG AATTCAGTAC	SEQ ID NO: 112	-	1	0	0	Down
hPLP1-wMN1-Cj-20	20	TTCACATAAACAGAAATGA AAGAAAAACAC	SEQ ID NO: 113	-	1	0	0	Down
hPLP1-wMN1-Cj-21	21	ATGCCAACTCTCTCATCTC GCAGTCTGTAC	SEQ ID NO: 114	+	1	0	0	Down
hPLP1-wMN1-Cj-22	22	GAGACATTCTCACATTTCC AGTCTAAGTAC	SEQ ID NO: 115	-	1	0	0	Down

[552]

[553] [표 5]

[554] SpCas9을 위한 마우스 Plp1 유전자의 표적 서열(Target sequences of mouse Plp1 gene for SpCas9)

[555]

sgRNA	No.	Target (5' to 3')	SEQ ID NO	Direction	Mismatch 0	Mismatch 1	Mismatch 2	location
mPlp1-TATA-Sp-01	1	TGTTTGGTAGTATAGTAAGTAGG	SEQ ID NO: 116	-	1	0	1	-
mPlp1-wMN1-Sp-01	2	GGTCTAGAAAAGATCAAGCCAGG	SEQ ID NO: 117	+	1	0	0	Up
mPlp1-wMN1-Sp-02	3	GCCAGGACTGTGACCTGATAAGG	SEQ ID NO: 118	+	1	0	0	Up
mPlp1-wMN1-Sp-03	4	TCACCTTCACACTTTAACCAAGG	SEQ ID NO: 119	+	1	0	0	Up
mPlp1-wMN1-Sp-04	5	CAAGGTTGAGACAATGTTCCAGG	SEQ ID NO: 120	+	1	0	0	Up
mPlp1-wMN1-Sp-05	6	CCAATTCATGTGCAAACATTTGG	SEQ ID NO: 121	-	1	0	0	Up
mPlp1-wMN1-Sp-06	7	CATCACAGTTTATACTTAGCTGG	SEQ ID NO: 122	+	1	0	0	Down
mPlp1-wMN1-Sp-07	8	ATCACAGTTTATACTTAGCTGGG	SEQ ID NO: 123	-	1	0	0	Down
mPlp1-wMN1-Sp-08	9	GGAATACCTCAGGCTCAACAGGG	SEQ ID NO: 124	-	1	0	0	Down
mPlp1-wMN1-Sp-09	10	TCTCTGTTTCGGAATACCTCAGG	SEQ ID NO: 125	-	1	0	0	Down
mPlp1-wMN1-Sp-10	11	CTGTCGACTACTTTGATGAAAGG	SEQ ID NO: 126	+	1	0	0	Down
mPlp1-wMN1-Sp-11	12	TGAACCAAGATGATTATTTGTGG	SEQ ID NO: 127	-	1	0	1	Up
mPlp1-wMN1-Sp-12	13	ATCTTCGTTCATAGAAATTTGGG	SEQ ID NO: 128	+	1	0	1	Up
mPlp1-wMN1-Sp-13	14	AGCCTTGCATGGCAGAGCTTGGG	SEQ ID NO: 129	-	1	0	1	Up
mPlp1-wMN1-Sp-14	15	ACACTTTAACCAAGGAAAGAGGG	SEQ ID NO: 130	+	1	0	1	Up
mPlp1-wMN1-Sp-15	16	TACCAGATCCCCTCTTTCCTTGG	SEQ ID NO: 131	-	1	0	1	Up
mPlp1-wMN1-Sp-16	17	CATTTGGAGGCCAAAATACAAGG	SEQ ID NO: 132	-	1	0	1	Up
mPlp1-wMN1-Sp-17	18	CCAAATGTTTGCACATGAATTGG	SEQ ID NO: 133	+	1	0	1	Up
mPlp1-wMN1-Sp-18	19	AGTCCAGATGCTGTCCCTGAAGG	SEQ ID NO: 134	+	1	0	1	Up
mPlp1-wMN1-Sp-19	20	CGCAAGCCATTCAAACACAAAGG	SEQ ID NO: 135	-	1	0	1	Down
mPlp1-wMN1-Sp-20	21	TCAAAACCCTGTTGAGCCTGAGG	SEQ ID NO: 136	+	1	0	1	Down
mPlp1-wMN1-Sp-21	22	CGGAATACCTCAGGCTCAACAGG	SEQ ID NO: 137	-	1	0	1	Down
mPlp1-wMN1-Sp-22	23	GTCAAAATGTGAATTCTAACAGG	SEQ ID NO: 138	-	1	0	1	Down

[556]

mPlp1-wMN1-Sp-23	24	TTATCTATTCTATTAGAGCTCGG	SEQ ID NO: 139	-	1	0	0	Down
mPlp1-wMN1-Sp-24	25	ATCAAGTAATGAAATGGACAAGG	SEQ ID NO: 140	-	1	0	1	Down
mPlp1-wMN1-Sp-25	26	CTCCCACTGCCTTATTAGGCAGG	SEQ ID NO: 141	-	1	0	0	Up
mPlp1-wMN1-Sp-26	27	AGAGCTCAAATGGGTTCTAAAGG	SEQ ID NO: 142	-	1	0	0	Up
mPlp1-wMN1-Sp-27	28	ACCACATTCAAGAGCTCAAATGG	SEQ ID NO: 143	-	1	0	0	Up
mPlp1-wMN1-Sp-28	29	TTACAGATTGGTTACACTTGGGG	SEQ ID NO: 144	+	1	0	0	Down
mPlp1-wMN1-Sp-29	30	ATCACTGCTGCTACTACTTATGG	SEQ ID NO: 145	-	1	0	0	Up
mPlp1-wMN1-Sp-30	31	ATACCTGCCTAATAAGGCAGTGG	SEQ ID NO: 146	+	1	0	0	Up
mPlp1-wMN1-Sp-31	32	GATCAGGAGAGTCAGTGGGATGG	SEQ ID NO: 147	+	1	0	1	Down
mPlp1-wMN1-Sp-32	33	CTATTGTGAGTCTCAGATTAAGG	SEQ ID NO: 148	-	1	0	1	Down
mPlp1-wMN1-Sp-33	34	TATTACAGATTGGTTACACTTGG	SEQ ID NO: 149	+	1	0	1	Down
mPlp1-wMN1-Sp-34	35	ATTACAGATTGGTTACACTTGGG	SEQ ID NO: 150	+	1	0	1	Down
mPlp1-wMN1-Sp-35	36	TACAGATTGGTTACACTTGGGGG	SEQ ID NO: 151	+	1	0	1	Down
mPlp1-wMN1-Sp-36	37	ACAGATTGGTTACACTTGGGGGG	SEQ ID NO: 152	+	1	0	1	Down

[557]

[558] [표 6]

[559] CjCas9을 위한 마우스 Plp1 유전자의 표적 서열(Target sequences of mouse Plp1 gene for CjCas9)

[560]

sgRNA	No.	Target (5' to 3')	SEQ ID NO	Direction	Mismatch 0	Mismatch 1	Mismatch 2	location
mPlp1-TATA-Cj-01	1	CTACTTACTATACTACCAA ACACACCGCAC	SEQ ID NO: 153	-	1	0	0	-
mPlp1-TATA-Cj-02	2	AAAGCCTACTTACTATACT ACCAAACACAC	SEQ ID NO: 154	-	1	0	0	-
mPlp1-TATA-Cj-03	3	CAAAAGCCTACTTACTATA C~ACCAAACAC	SEQ ID NO: 155	-	1	0	0	-
mPlp1-TATA-Cj-04	4	GGGTCTGAATCAAAAGCCT ACTTACTATAC	SEQ ID NO: 156	-	1	0	0	-
mPlp1-wMN1-Cj-01	5	AGAGTGGGATTCTACAAGT CACCTTCACAC	SEQ ID NO: 157	+	1	0	0	Up
mPlp1-wMN1-Cj-02	6	GGAAAGAGGGGATCTGGT AGCATAAAGTAC	SEQ ID NO: 158	+	1	0	0	Up
mPlp1-wMN1-Cj-03	7	GGGATCTGGTAGCATAAA GTACAGCTACAC	SEQ ID NO: 159	+	1	0	0	Up
mPlp1-wMN1-Cj-04	8	ATCTGTCACTAGCGACAAG TGTAGCTGTAC	SEQ ID NO: 160	-	1	0	0	Up
mPlp1-wMN1-Cj-05	9	TCATG~GCAAACATTTGGA GGCCAAAATAC	SEQ ID NO: 161	-	1	0	0	Up
mPlp1-wMN1-Cj-06	10	GACATACAGAGAGGGGGC GGAGAGAAATAC	SEQ ID NO: 162	-	1	0	0	Up
mPlp1-wMN1-Cj-07	11	ATACTGACGCCATCACATC ACAGTTTATAC	SEQ ID NO: 163	+	1	0	0	Down
mPlp1-wMN1-Cj-08	12	TAAAACTATAAGCTCTCTG T~TCGGAATAC	SEQ ID NO: 164	-	1	0	0	Down
mPlp1-wMN1-Cj-09	13	TCATCAAAGTAGTCGACAG TCAAAGCATAC	SEQ ID NO: 165	-	1	0	0	Down
mPlp1-wMN1-Cj-10	14	TGAATTCTAACAGGAAAAC TCAGAACATAC	SEQ ID NO: 166	-	1	0	0	Down
mPlp1-wMN1-Cj-11	15	ACTGCTGCTACTACTTATG GTGACTAGTAC	SEQ ID NO: 167	-	1	0	0	Up
mPlp1-wMN1-Cj-12	16	AGTCACCATAAGTAGTAGC AGCAGTGATAC	SEQ ID NO: 168	+	1	0	0	Up
mPlp1-wMN1-Cj-13	17	CATAAGTAGTAGCAGCAGT GATACTAATAC	SEQ ID NO: 169	+	1	0	0	Up
mPlp1-wMN1-Cj-14	18	TTGAATGGCTTGCGAACAA AGATTAAACAC	SEQ ID NO: 170	-	1	0	0	Down
mPlp1-wMN1-Cj-15	19	TTAATCTTTGTTCGCAAGCC ATTCAAACAC	SEQ ID NO: 171	+	1	0	0	Down
mPlp1-wMN1-Cj-16	20	TTGCTGCATCTCTAACGTG AACTCTAACAC	SEQ ID NO: 172	-	1	0	0	Down
mPlp1-wMN1-Cj-17	21	TTCACGTTAGAGATGCAGC AAAGTCTATAC	SEQ ID NO: 173	+	1	0	0	Down
mPlp1-wMN1-Cj-18	22	TGGAAGCAACTCTAAATCA CCACCCGATAC	SEQ ID NO: 174	-	1	0	0	Down
mPlp1-wMN1-Cj-19	23	TTCCAAAGTTCTGTCACCC AGTAAAAACAC	SEQ ID NO: 175	+	1	0	0	Down

[561]

mPlp1-wMN1-Cj-20	24	TTCAAGAGCTCAAATGGGT TCTAAAGGCAC	SEQ ID NO: 176	-	1	0	0	Up
mPlp1-wMN1-Cj-21	25	TTGAA GTGGTATAAGTGC TAATATCATAC	SEQ ID NO: 177	-	1	0	0	Up
mPlp1-wMN1-Cj-22	26	GTATAAGTGCTAATATCAT ACAGGAAACAC	SEQ ID NO: 178	+	1	0	0	Up
mPlp1-wMN1-Cj-23	27	GTGTTTCCTGTATGATATTA GCACTTATAC	SEQ ID NO: 179	-	1	0	0	Up
mPlp1-wMN1-Cj-24	28	GACTTTGTGTTTCCTGTATG ATATTAGCAC	SEQ ID NO: 180	-	1	0	0	Up
mPlp1-wMN1-Cj-25	29	AAAACAATTATCAGGCAGT GACAGAGACAC	SEQ ID NO: 181	-	1	0	0	Down
mPlp1-wMN1-Cj-26	30	CCAAGATACTAGAGTAGCT G GACTGGCAC	SEQ ID NO: 182	-	1	0	0	Down
mPlp1-wMN1-Cj-27	31	GGCCTATAGCCATTCAAAT GGCCAAGATAC	SEQ ID NO: 183	-	1	0	0	Down
mPlp1-wMN1-Cj-28	32	GTCCCATCTCCCTAAGTCT CGAATCTGCAC	SEQ ID NO: 184	-	1	0	0	Down

[562]

[563] **본 명세서에 의해 개시되는 내용의 일 태양으로, 발현 조절용 조성물은 가이드핵산 및 에디터단백질을 포함할 수 있다.**

[564] 발현 조절용 조성물은 다음을 포함할 수 있다:

[565] (a) 중복 유전자의 전사 조절 영역 내 위치한 표적 서열을 표적화할 수 있는 가이드핵산 또는 이를 암호화하는 핵산 서열; 및

[566] (b) 하나 이상의 에디터단백질 또는 이를 암호화는 핵산 서열.

[567] 상기 중복 유전자 관련 설명은 상기 기술한 바와 같다.

[568] 상기 전사 조절 영역 관련 설명을 상기 기술한 바와 같다.

[569] 상기 표적 서열 관련 설명은 상기 기술한 바와 같다.

[570]

[571] 발현 조절용 조성물은 가이드핵산-에디터단백질 복합체를 포함할 수 있다.

[572] "가이드핵산-에디터단백질 복합체"는 가이드핵산과 에디터단백질의 상호작용을 통해 형성된 복합체를 의미한다.

[573] 상기 가이드핵산 관련 설명은 상기 기술한 바와 같다.

[574] 상기 "에디터단백질"은 핵산과 직접적으로 결합하거나, 또는 직접 결합하지는 않지만 상호작용할 수 있는 펩타이드, 폴리펩타이드 또는 단백질을 의미한다.

[575] 이때, 상기 핵산은 표적 핵산, 유전자 또는 염색체에 포함된 핵산일 수 있다.

[576] 이때, 상기 핵산은 가이드핵산일 수 있다.

[577] 상기 에디터단백질은 효소일 수 있다.

[578] 이때, 상기 "효소"는 핵산, 유전자 또는 염색체를 절단할 수 있는 도메인을 포함하는 폴리펩타이드 또는 단백질을 의미한다.

[579] 상기 효소는 뉴클레아제 또는 제한효소일 수 있다.

[580]

[581] 상기 에디터단백질은 완전 활성 효소를 포함할 수 있다.

[582] 이때, 상기 "완전 활성 효소"는 야생형(wild type) 효소의 본래의 핵산, 유전자 또는 염색체 절단 기능과 동일한 기능을 가지는 효소를 의미한다. 예를 들면, DNA의 이중 가닥을 절단하는 야생형 효소는 DNA 이중 가닥을 모두 절단하는 완전한 활성 효소일 수 있다. 또 다른 예를 들면, DNA의 이중 가닥을 절단하는 야생형 효소가 인위적인 조작에 의해 아미노산 서열 중 일부 서열이 삭제(deletion) 또는 치환(substitution)된 경우, 인위적으로 조작된 효소 변이체가 야생형 효소와 동일하게 DNA의 이중 가닥을 절단한다면, 상기 인위적으로 조작된 효소 변이체는 완전 활성 효소일 수 있다.

[583] 또한, 상기 완전 활성 효소는 야생형의 효소의 기능보다 향상 된 기능을 가지고 있는 효소를 포함할 수 있다. 예를 들면, DNA의 이중 가닥을 절단하는 야생형 효소의 특정 변형 또는 조작된 형태는 야생형 효소보다 증가된 완전한 효소 활성, 즉, 증가된 DNA 이중 가닥을 절단하는 활성을 가질 수 있다.

[584] 상기 에디터단백질은 불완전 또는 부분 활성 효소를 포함할 수 있다.

[585] 이때, 상기 "불완전 또는 부분 활성 효소"는 야생형 효소의 본래의 핵산, 유전자 또는 염색체 절단 기능의 일부만을 가지는 효소를 의미한다. 예를 들면, DNA의 이중 가닥을 절단하는 야생형 효소의 특정 변형 또는 조작된 형태는 제1 기능을 가지는 형태 또는 제2 기능을 가지는 형태일 수 있다. 이때, 제1 기능은 DNA의 이중 가닥 중 제1 가닥을 절단하는 기능이고, 제2 기능은 DNA의 이중 가닥 중 제2 가닥을 절단하는 기능일 수 있다. 이때, 상기 제1 기능을 가지는 효소 또는

제2 기능을 가지는 효소는 불완전 또는 부분 활성 효소일 수 있다.

[586] 상기 에디터단백질은 불활성 효소를 포함할 수 있다.

[587] 이때, 상기 "불활성 효소"는 야생형 효소의 본래의 핵산, 유전자 또는 염색체 절단 기능이 모두 불활성화 된 효소를 의미한다. 예를 들면, 야생형 효소의 특정 변형 또는 조작된 형태는 제1 기능 및 제2 기능이 모두 상실된 형태, 즉, DNA의 이중 가닥 중 제1 가닥을 절단하는 제1 기능 및 제2 가닥을 절단하는 제2 기능이 모두 상실된 형태일 수 있다. 이때, 상기 제1 기능 및 제2 기능이 모두 상실된 효소는 불활성 효소일 수 있다.

[588]

[589] 상기 에디터단백질은 융합단백질일 수 있다.

[590] 이때, 상기 "융합 단백질"은 효소에 추가적인 도메인, 펩타이드, 폴리펩타이드 또는 단백질을 융합하여 생성한 단백질을 의미한다.

[591] 상기 추가적인 도메인, 펩타이드, 폴리펩타이드 또는 단백질은 상기 효소에 포함된 기능적 도메인, 펩타이드, 폴리펩타이드 또는 단백질과 동일하거나 다른 기능을 가지는 기능적 도메인, 펩타이드, 폴리펩타이드 또는 단백질일 수 있다.

[592] 상기 융합 단백질은 효소의 아미노 말단 또는 그 근처; 카르복시 말단 또는 그 근처; 효소의 중간부; 또는 이들 조합의 하나 이상에 상기 기능적 도메인, 펩타이드, 폴리펩타이드 또는 단백질이 부가된 형태일 수 있다.

[593] 이때, 상기 기능적 도메인, 펩타이드, 폴리펩타이드 또는 단백질은 메틸라아제(methylase) 활성, 디메틸라아제(demethylase) 활성, 전사촉진(transcription activation) 활성, 전사 저해(transcription repression) 활성, 전사 방출 인자(transcription release factor) 활성, 히스톤 변형(histone modification) 활성, RNA 절단(cleavage) 활성 또는 핵산 결합(nucleic acid binding) 활성을 가지는 도메인, 펩타이드, 폴리펩타이드 또는 단백질 일 수 있으며, 또는 단백질(펩타이드 포함)의 분리정제를 위한 태그(tag) 또는 리포터 유전자일 수 있으나, 이에 제한되지 않는다.

[594] 상기 기능적 도메인, 펩타이드, 폴리펩타이드 또는 단백질은 디아미네이즈(deaminase)일 수 있다.

[595] 상기 태그는 히스티딘(His) 태그, V5 태그, FLAG 태그, 인플루엔자 헤마글루티닌(HA) 태그, Myc 태그, VSV-G 태그 및 티오레독신(Trx) 태그 등을 포함하며, 상기 리포터 유전자는 글루타티온-S-트랜스 퍼라제(GST), 호스래디시(horseradish) 과산화효소(HRP), 클로람페니콜 아세틸트랜스퍼라제(CAT) 베타-갈락토시다제, 베타-글루쿠로니다제, 루시퍼라제, 녹색 형광 단백질(GFP), HcRed, DsRed, 청록색 형광 단백질(CFP), 황색 형광 단백질(YFP) 및 청색 형광 단백질(BFP)을 포함하는 자가형광 단백질을 포함하나, 이들에 한정되지 않는다.

[596] 또한, 상기 기능적 도메인, 펩타이드, 폴리펩타이드 또는 단백질은 NLS(nuclear localization sequence or signal) 또는 NES(nuclear export sequence or signal)일 수

있다.

[597] 상기 NLS는 아미노산 서열 PKKKRKV를 갖는 SV40 바이러스 대형 T-항원의 NLS; 뉴클레오플라스민(nucleoplasmin)으로부터의 NLS(예를 들어, 서열 KRPAATKKAGQAKKKK를 갖는 뉴클레오플라스민 이분(bipartite) NLS); 아미노산 서열 PAAKRVKLD 또는 RQRRNELKRSP를 갖는 c-myc NLS; 서열 NQSSNFGPMKGGNFGGRSSGPYGGGGQYFAKPRNQGGY를 갖는 hRNPA1 M9 NLS; 임포틴-알파로부터의 IBB 도메인의 서열 RMRIZFKNKGKDTAELRRRRVEVSVELRKAKKDEQILKRRNV; 마이오마(myoma) T 단백질의 서열 VSRKRPRP 및 PPKKARED; 인간 p53의 서열 POPKKKPL; 마우스 c-abl IV의 서열 SALIKKKKKMAP; 인플루엔자 바이러스 NS1의 서열 DRLRR 및 PKQKKRK; 간염 바이러스 델타 항원의 서열 RKLKKKIKKL; 마우스 Mx1 단백질의 서열 REKKKFLKRR; 인간 폴리(ADP-리보스) 중합효소의 서열 KRKGDEVDGVDEVAKKKSKK; 또는 스테로이드 호르몬 수용체(인간) 글루코코르티코이드의 서열 RKCLQAGMNLEARKTKK로부터 유래된 NLS 서열일 수 있으나, 이에 제한되지 않는다.

[598] 상기 추가적인 도메인, 펩타이드, 폴리펩타이드 또는 단백질은 특정 기능을 수행하지 않는 비기능적 도메인, 펩타이드, 폴리펩타이드 또는 단백질일 수 있다. 이때, 상기 비기능적 도메인, 펩타이드, 폴리펩타이드 또는 단백질은 상기 효소의 기능에 영향을 주지 않는 도메인, 펩타이드, 폴리펩타이드 또는 단백질일 수 있다.

[599] 상기 융합 단백질은 효소의 아미노 말단 또는 그 근처; 카르복시 말단 또는 그 근처; 효소의 중간부; 또는 이들 조합의 하나 이상에 상기 비기능적 도메인, 펩타이드, 폴리펩타이드 또는 단백질이 부가된 형태일 수 있다.

[600]

[601] 상기 에디터단백질은 자연 상태에 존재하는 효소 또는 융합 단백질일 수 있다.

[602] 상기 에디터단백질은 자연 상태에 존재하는 효소 또는 융합 단백질의 일부가 변형된 형태일 수 있다.

[603] 상기 에디터단백질은 자연 상태에 존재하지 않는 인위적으로 생성된 효소 또는 융합 단백질일 수 있다.

[604] 상기 에디터단백질은 자연 상태에 존재하지 않는 인위적으로 생성된 효소 또는 융합 단백질의 일부가 변형된 형태일 수 있다.

[605] 이때, 상기 변형은 에디터단백질에 포함된 아미노산의 치환, 제거, 부가 또는 이의 혼합일 수 있다.

[606] 또는 상기 변형은 에디터단백질을 암호화하는 뉴클레오타이드 서열 중 일부 뉴클레오타이드의 치환, 제거, 부가 또는 이의 혼합일 수 있다.

[607]

[608] 또한, 상기 발현 조절용 조성물은 삽입을 원하는 특정 뉴클레오타이드서열을

포함하는 도너 또는 이를 암호화하는 핵산 서열을 선택적으로 더 포함할 수 있다.

[609] 이때, 상기 삽입을 원하는 핵산 서열은 중복 유전자의 전사 조절 영역 내 일부 뉴클레타이드 서열일 수 있다.

[610] 이때, 상기 삽입을 원하는 핵산 서열은 중복 유전자의 전사 조절 영역 내 돌연변이를 도입하기 위한 핵산 서열일 수 있다. 이때, 상기 돌연변이는 중복 유전자의 전사를 방해하는 돌연변이일 수 있다.

[611] 상기 "도너(donor)"는 손상된 유전자 또는 핵산의 상동성 재조합에 의한 수복을 돕는 뉴클레오타이드서열을 의미한다.

[612] 상기 도너는 이중가닥 핵산 또는 단일가닥 핵산일 수 있다.

[613] 상기 도너는 선형 또는 원형일 수 있다.

[614] 상기 도너는 표적 유전자의 전사 조절 영역 내 핵산에 상동성을 가지는 뉴클레오타이드 서열을 포함할 수 있다.

[615] 예를 들어, 도너는 특정 뉴클레오타이드서열을 삽입하고자 하는 위치, 예를 들어, 손상된 핵산의 왼쪽(upstream)및 오른쪽(downstream)의 뉴클레오타이드서열과 각각 상동성을 가지는 뉴클레오타이드서열을 포함할 수 있다. 이때, 삽입하고자 하는 특정 뉴클레오타이드서열은 손상된 핵산의 오른쪽 뉴클레오타이드서열에 대해 상동성을 가지는 뉴클레오타이드서열과 손상된 핵산의 왼쪽 뉴클레오타이드서열에 대해 상동성을 가지는 뉴클레오타이드서열 사이에 위치할 수 있다. 이때, 상기 상동성을 가지는 뉴클레오타이드서열은 최소한 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90% 또는 95% 이상의 상동성을 갖거나 또는 완전하게 상동성을 가질 수 있다.

[616] 상기 도너는 특정 핵산 서열을 포함할 수 있다.

[617] 이때, 상기 특정 핵산 서열은 표적 유전자의 일부 뉴클레오타이드 서열 또는 일부 유사 뉴클레오타이드일 수 있다. 상기 표적 유전자의 일부 뉴클레오타이드 서열은, 예를 들어, 돌연변이를 포함하는 표적 유전자를 교정하기 위한 돌연변이가 교정된 정상 핵산 서열을 포함할 수 있다. 또는, 상기 표적 유전자의 일부 유사 뉴클레오타이드 서열은 정상 표적 유전자를 돌연변이 시키키 위한 표적 유전자의 일부 정상 핵산 서열 중 일부가 변형된 돌연변이 유발 핵산 서열을 포함할 수 있다.

[618] 이때, 상기 특정 핵산 서열은 외인성 핵산 서열일 수 있다. 예를 들어, 상기 외인성 핵산 서열은 표적 유전자를 포함하는 세포에서 발현하길 원하는 외인성 유전자일 수 있다.

[619] 이때, 상기 특정 핵산 서열은 표적 유전자를 포함하는 세포에서 발현시키길 원하는 핵산 서열일 수 있다. 예를 들어, 상기 핵산 서열은 표적 유전자를 포함하는 세포에서 발현하는 특정 유전자일 수 있으며, 이러한 경우, 도너를 포함하는 발현 조절용 조성물에 의해 상기 특정 유전자는 세포 내 카피수가 증가하여 발현이 증가할 수 있다.

[620] 상기 도너는 선택적으로 부수적인 뉴클레오타이드 서열을 포함할 수 있다. 이때, 부수적인 핵산서열은 도너의 안정성, 표적 내 삽입 효율 또는 상동성 재조합 효율을 높이는 역할을 하는 것일 수 있다.

[621] 예를 들어, 상기 부수적인 뉴클레오타이드 서열은 뉴클레오타이드 A, T가 풍부한 핵산서열, 즉, A-T 풍부 도메인(A-T rich domain)일 수 있다. 예를 들면, 상기 부수적인 뉴클레오타이드 서열은 SMAR(scaffold/matrix attachment region)일 수 있다.

[622]

[623] 본 명세서에 의해 개시되는 가이드핵산, 에디터단백질 또는 가이드핵산-에디터단백질 복합체는 다양한 형태로 대상 내로 전달 또는 도입될 수 있다.

[624] 이때, "대상"은 가이드핵산, 에디터단백질 또는 가이드핵산-에디터단백질 복합체가 도입되는 유기체, 또는 가이드핵산, 에디터단백질 또는 가이드핵산-에디터단백질 복합체가 작동하는 유기체 또는 유기체로부터 획득한 검체 또는 시료를 의미한다.

[625] 상기 대상은 가이드핵산-에디터단백질 복합체의 표적 유전자 또는 염색체를 포함하는 유기체일 수 있다.

[626] 상기 유기체는 동물, 동물의 조직 또는 동물 세포일 수 있다.

[627] 상기 유기체는 인간, 인간의 조직 또는 인간 세포일 수 있다.

[628] 상기 조직은 안구, 피부, 간, 신장, 심장, 폐, 뇌, 근육 또는 혈액 등일 수 있다.

[629] 상기 세포는 섬유아세포, 슈반세포, 신경세포, 희소돌기아교세포(oligodendrocyte), 근아세포, 신경교세포, 대식세포, 면역세포, 간세포, 망막색소상피세포, 암세포 또는 줄기세포 등일 수 있다.

[630] 상기 검체 또는 시료는 침, 혈액, 망막조직, 뇌조직, 슈반세포, 희소돌기아교세포, 근아세포, 섬유아세포, 신경세포, 신경교세포, 대식세포, 간세포, 면역세포, 암세포 또는 줄기세포 등 표적 유전자 또는 염색체를 포함하는 유기체에서 획득한 것 일 수 있다.

[631] 바람직하게는, 상기 대상은 중복 유전자를 포함하는 유기체일 수 있다. 이때, 대상은 중복 유전자가 유전자 중복 상태로 포함된 유기체일 수 있다.

[632] 상기 가이드핵산, 에디터단백질 또는 가이드핵산-에디터단백질 복합체는 DNA, RNA 또는 이의 혼합의 형태로 대상 내로 전달 또는 도입될 수 있다.

[633] 이때, 가이드핵산 및/또는 에디터단백질을 암호화하는 DNA, RNA 또는 이의 혼합의 형태는 당업계에 공지된 방법에 의해 대상 내로 전달 또는 도입될 수 있다.

[634] 또는, 가이드핵산 및/또는 에디터단백질을 암호화하는 DNA, RNA 또는 이의 혼합의 형태는 벡터, 비벡터 또는 이들의 조합에 의해 대상 내로 전달 또는 도입될 수 있다.

[635] 상기 벡터는 바이러스 또는 비바이러스 벡터(예를 들어, 플라스미드)일 수

있다.

[636] 상기 비벡터는 네이키드 DNA, DNA 복합체 또는 mRNA일 수 있다.

[637] 일 구현예로, 상기 가이드핵산 및/또는 에디터단백질을 암호화하는 핵산서열은 벡터에 의해 대상 내로 전달 또는 도입될 수 있다.

[638] 상기 벡터는 가이드핵산 및/또는 에디터단백질을 암호화하는 핵산서열을 포함할 수 있다.

[639] 일 예로, 상기 벡터는 가이드핵산과 에디터단백질을 암호화하는 핵산서열을 동시에 포함할 수 있다.

[640] 다른 일 예로, 상기 벡터는 가이드핵산 암호화하는 핵산서열을 포함할 수 있다.

[641] 예를 들어, 상기 가이드핵산에 포함되는 도메인은 하나의 벡터에 모두 포함되거나 또는 각각의 도메인을 나누어 각각의 벡터에 포함시킬 수 있다.

[642] 다른 일 예로, 상기 벡터는 에디터단백질을 암호화하는 핵산서열을 포함할 수 있다.

[643] 예를 들어, 상기 에디터단백질의 경우, 에디터단백질을 암호화하는 핵산서열은 하나의 벡터에 포함되거나 또는 에디터단백질을 암호화하는 핵산서열이 분할되어 여러 개의 벡터에 포함될 수 있다.

[644] 상기 벡터는 하나 이상의 조절/제어 구성요소를 포함할 수 있다.

[645] 이때, 상기 조절/제어 구성요소는 포로모터, 인핸서, 인트론, 폴리아데닐화 신호, 코작 공통(Kozak consensus) 서열, 내부 리보솜 유입 부위(internal ribosome entry site, IRES), 스플라이스 억셉터 및/또는 2A 서열을 포함할 수 있다.

[646] 상기 프로모터는 RNA 중합효소 II에 의해 인식되는 프로모터일 수 있다.

[647] 상기 프로모터는 RNA 중합효소 III에 의해 인식되는 프로모터일 수 있다.

[648] 상기 프로모터는 유도성 프로모터일 수 있다.

[649] 상기 프로모터는 대상 특이적 프로모터일 수 있다.

[650] 상기 프로모터는 바이러스 또는 비바이러스 프로모터일 수 있다.

[651] 상기 프로모터는 제어 영역(즉, 가이드핵산 또는 에디터단백질을 암호화하는 핵산서열)에 따라 적합한 프로모터를 이용할 수 있다.

[652] 예를 들어, 가이드핵산을 위해 유용한 프로모터는 H1, EF-1a, tRNA 또는 U6 프로모터일 수 있다. 예를 들어, 에디터단백질을 위해 유용한 프로모터는 CMV, EF-1a, EFS, MSCV, PGK 또는 CAG 프로모터일 수 있다.

[653] 벡터는 바이러스 벡터 또는 재조합 바이러스 벡터일 수 있다.

[654] 상기 바이러스는 DNA 바이러스 또는 RNA 바이러스일 수 있다.

[655] 이때, 상기 DNA 바이러스는 이중가닥 DNA(dsDNA) 바이러스 또는 단일가닥 DNA(ssDNA) 바이러스 일 수 있다.

[656] 이때, 상기 RNA 바이러스는 단일가닥 RNA(ssRNA) 바이러스일 수 있다.

[657] 상기 바이러스는 레트로바이러스, 렌티바이러스, 아데노바이러스, 아데노-연관 바이러스(AAV), 백시니아바이러스, 폭스바이러스 또는 단순포진 바이러스일 수 있으나, 이에 제한되지 않는다.

[658] 일반적으로 바이러스는 숙주(예를 들면, 세포)를 감염시켜, 숙주 내에 바이러스의 유전정보를 암호화하는 핵산을 도입시키거나 숙주의 게놈 내로 유전정보를 암호화하는 핵산을 삽입시킬 수 있다. 이러한 특징을 가지는 바이러스를 이용하여 대상 내로 가이드핵산 및/또는 에디터단백질을 도입시킬 수 있다. 바이러스를 이용하여 도입된 가이드핵산 및/또는 에디터단백질은 대상(예를 들면, 세포)에서 일시적으로 발현될 수 있다. 또는 바이러스를 이용하여 도입된 가이드핵산 및/또는 에디터단백질은 대상(예를 들면, 세포)에서 장기간(예를 들면, 1주, 2주, 3주, 1개월, 2개월, 3개월, 6개월, 9개월, 1년, 2년 또는 영구적) 지속적으로 발현될 수 있다.

[659] 바이러스의 패키징 능력은 적어도 2kb 내지 50kb로 바이러스 종류에 따라 다를 수 있다. 이러한 패키징 능력에 따라 가이드핵산 또는 에디터단백질을 단독으로 포함하는 바이러스 벡터를 설계하거나 가이드핵산 및 에디터단백질을 모두 포함하는 바이러스 벡터를 설계할 수 있다. 또는 가이드핵산, 에디터단백질 및 추가 구성요소를 포함하는 바이러스 벡터를 설계할 수 있다.

[660] 일 예로, 가이드핵산 및/또는 에디터단백질을 암호화하는 핵산서열은 재조합 렌티바이러스에 의해 전달 또는 도입될 수 있다.

[661] 다른 일 예로, 가이드핵산 및/또는 에디터단백질을 암호화하는 핵산서열은 재조합 아데노바이러스에 의해 전달 또는 도입될 수 있다.

[662] 또 다른 일 예로, 가이드핵산 및/또는 에디터단백질을 암호화하는 핵산서열은 재조합 AAV에 의해 전달 또는 도입될 수 있다.

[663] 다른 일 예로, 가이드핵산 및/또는 에디터단백질을 암호화하는 핵산서열은 혼성 바이러스, 예를 들어 본 명세서에 기재한 바이러스 중 하나 이상의 혼성체에 의해 전달 또는 도입될 수 있다.

[664]

[665] 다른 일 구현예로, 상기 가이드핵산 및/또는 에디터단백질을 암호화하는 핵산서열은 비벡터로 대상 내로 전달 또는 도입될 수 있다.

[666] 비벡터는 가이드핵산 및/또는 에디터단백질을 암호화하는 핵산서열을 포함할 수 있다.

[667] 상기 비벡터는 네이키드 DNA, DNA 복합체, mRNA 또는 이의 혼합일 수 있다.

[668] 상기 비벡터는 전기천공법, 유전자총, 초음파천공법, 자기주입법(magnetofection), 일시적인 세포 압축 또는 스퀴징(예를 들어, 문헌[Lee, et al, (2012) Nano Lett., 12, 6322-6327]에 기재되어 있음), 지질-매개 형질감염, 덴드리머, 나노입자, 인산칼슘, 실리카, 실리케이트(오르모실) 또는 이의 조합에 의해 대상 내로 전달 또는 도입될 수 있다.

[669] 일 예로, 전기천공법을 통한 전달은 카트리지, 챔버 또는 큐벳 내에서 세포와 가이드핵산 및/또는 에디터단백질을 암호화하는 핵산서열을 혼합하고, 정해진 지속시간 및 진폭의 전기적 자극을 적용에 의해 수행될 수 있다.

[670] 다른 일 예로, 나노입자를 이용하여 비벡터를 전달할 수 있다. 상기 나노입자는

무기 나노입자(예를 들면, 자기 나노입자, 실리카 등) 또는 유기 나노입자(예를 들면, 폴리에틸렌 글리콜(PEG)로 코팅된 지질 등)일 수 있다. 상기 나노입자의 외면은 부착을 가능하게 하는 양 전하로 하전된 중합체(예를 들면, 폴리에틸렌이민, 폴리리신, 폴리세린 등)와 컨쥬케이팅될 수 있다.

[671] 임의의 구체예로, 지질 외피를 이용하여 전달할 수 있다.

[672] 임의의 구체예로, 엑소좀을 이용하여 전달할 수 있다. 엑소좀은 뇌 및 다른 표적 기관에 RNA를 전달할 수 있는, 단백질 및 RNA를 수송하는 내인성 나노-소낭이다.

[673] 임의의 구체예로, 리포좀을 이용하여 전달할 수 있다. 리포좀은 내부 수성 구획을 둘러싸는 단일 또는 다중 라멜라 지질 이중층 및 상대적으로 불투과성인 외측 친지성 인지질 이중층으로 구성된 구체 소낭 구조이다. 리포좀은 몇몇 상이한 유형의 지질로부터 만들어질 수 있지만; 약물 담체로서 리포좀을 생성하는데 인지질이 가장 통상적으로 사용된다.

[674] 또한, 비벡터의 전달을 위한 조성물은 기타 몇몇 다른 첨가제를 포함할 수 있다.

[675]

[676] 상기 에디터단백질은 펩타이드, 폴리펩타이드 또는 단백질의 형태로 대상 내로 전달 또는 도입될 수 있다.

[677] 상기 에디터단백질은 펩타이드, 폴리펩타이드 또는 단백질의 형태는 당업계에 공지된 방법에 의해 대상 내로 전달 또는 도입될 수 있다.

[678] 상기 펩타이드, 폴리펩타이드 또는 단백질의 형태는 전기천공법, 미량주사법, 일시적인 세포 압축 또는 스퀴징(예를 들어, 문헌[Lee, et al, (2012) Nano Lett., 12, 6322-6327]에 기재되어 있음), 지질-매개 형질감염, 나노파티클, 리포솜, 펩타이드-매개 전달 또는 이의 조합에 의해 대상 내로 전달 또는 도입될 수 있다.

[679] 상기 펩타이드, 폴리펩타이드 또는 단백질의 전달은 가이드핵산을 암호화하는 핵산서열과 함께 전달될 수 있다.

[680] 일 예로, 전기천공법을 통한 전달은 카트리지, 챔버 또는 큐벳 내에서 가이드핵산와 함께 또는 가이드핵산 없이 에디터단백질을 도입시킬 세포와 혼합하고, 정해진 지속시간 및 진폭의 전기적 자극을 적용에 의해 수행될 수 있다.

[681]

[682] 상기 가이드핵산 및 에디터단백질은 핵산-단백질 혼합의 형태로 대상 내로 전달 또는 도입될 수 있다.

[683] 상기 가이드핵산 및 에디터단백질은 가이드핵산-에디터단백질 복합체의 형태로 대상 내로 전달 또는 도입될 수 있다.

[684] 예를 들어, 상기 가이드핵산은 DNA, RNA 또는 이의 혼합 형태일 수 있다. 상기 에디터단백질은 펩타이드, 폴리펩타이드 또는 단백질의 형태일 수 있다.

[685] 일 예로, 가이드핵산 및 에디터단백질은 RNA 형태의 가이드핵산과 단백질 형태의 에디터단백질이 가이드핵산-에디터단백질 복합체, 즉

ribonucleoprotein(RNP)의 형태로 대상 내로 전달 또는 도입될 수 있다.

[686]

[687] 본 명세서에 의해 개시되는 가이드핵산-에디터단백질 복합체는 표적 핵산, 유전자 또는 염색체를 변형시킬 수 있다.

[688] 예를 들어, 가이드핵산-에디터단백질 복합체는 표적 핵산, 유전자 또는 염색체의 서열에 변형(modification)을 유도한다. 그 결과, 상기 표적 핵산, 유전자 또는 염색체에 의해 발현되는 단백질은 그 구조 및/또는 기능이 변형되거나, 그 단백질의 발현이 조절되거나, 그 단백질의 발현이 제거될 수 있다.

[689] 상기 가이드핵산-에디터단백질 복합체는 DNA, RNA, 유전자 또는 염색체 수준에서 작용할 수 있다.

[690] 일 예로, 상기 가이드핵산-에디터단백질 복합체는 표적 유전자의 전사 조절 영역을 조작 또는 변형하여 표적 유전자가 암호화하는 단백질의 발현을 조절(예를 들어, 억제, 저해, 감소, 증가 또는 촉진) 하거나, 또는 단백질 활성 조절(예를 들어, 억제, 저해, 감소, 증가 또는 촉진) 또는 변형된 단백질을 발현할 수 있다.

[691] 상기 가이드핵산-에디터단백질 복합체는 유전자의 전사와 번역의 단계에서 작용할 수 있다.

[692] 일 예로, 상기 가이드핵산-에디터단백질 복합체는 표적 유전자의 전사를 촉진 또는 저해하여 표적 유전자가 암호화하는 단백질의 발현을 조절(예를 들어, 억제, 저해, 감소, 증가 또는 촉진)할 수 있다.

[693] 다른 일 예로, 상기 가이드핵산-에디터단백질 복합체는 표적 유전자의 번역을 촉진 또는 저해하여 표적 유전자가 암호화하는 단백질의 발현을 조절(예를 들어, 억제, 저해, 감소, 증가 또는 촉진)할 수 있다.

[694]

[695] **본 명세서에 의해 개시되는 내용의 일 구체예로서, 발현 조절용 조성물은 gRNA 및 CRISPR 효소를 포함할 수 있다.**

[696] 발현 조절용 조성물은 다음을 포함할 수 있다:

[697] (a) 중복 유전자의 전사 조절 영역 내 위치한 표적 서열을 표적화할 수 있는 gRNA 또는 이를 암호화하는 핵산 서열; 및

[698] (b) 하나 이상의 CRISPR 효소 또는 이를 암호화는 핵산 서열.

[699] 상기 중복 유전자 관련 설명은 상기 기술한 바와 같다.

[700] 상기 전사 조절 영역 관련 설명을 상기 기술한 바와 같다.

[701] 상기 표적 서열 관련 설명은 상기 기술한 바와 같다.

[702]

[703] 발현 조절용 조성물은 gRNA-CRISPR 효소 복합체를 포함할 수 있다.

[704] "gRNA-CRISPR 효소 복합체"는 gRNA과 CRISPR 효소의 상호작용을 통해 형성된 복합체를 의미한다.

[705] 상기 gRNA 관련 설명은 상기 기술한 바와 같다.

[706] 상기 “CRISPR 효소”는 CRISPR-Cas 시스템의 주요 단백질 구성 요소로, gRNA와 복합체를 형성하여 CRISPR-Cas 시스템을 형성한다.

[707] 상기 CRISPR 효소는 CRISPR 효소를 암호화하는 서열을 가지는 핵산 또는 폴리펩타이드(또는 단백질)일 수 있다.

[708] 상기 CRISPR 효소는 Type II CRISPR 효소일 수 있다.

[709] Type II CRISPR 효소의 결정 구조는 2종 이상의 자연유래 미생물 Type II CRISPR 효소 분자에 대한 연구(Jinek et al., Science, 343(6176):1247997, 2014) 및 gRNA와 함께 복합체를 이루는 스트렙토코커스 피오게네스 Cas9(SpCas9)에 대한 연구(Nishimasu et al., Cell, 156:935-949, 2014; 및 Anders et al., Nature, 2014, doi: 10.1038/nature13579)를 통해 결정되었다.

[710] Type II CRISPR 효소는 2개의 로브, 즉, 인식(REC) 및 뉴클레아제(NUC) 로브를 포함하며, 각각의 로브는 여러 개의 도메인을 포함한다.

[711] 상기 REC 로브는 아르기닌-풍부 브릿지 나선(BH), REC1 도메인 및 REC2 도메인을 포함한다.

[712] 이때, 상기 BH 도메인은 긴 α-나선 및 아르기닌 풍부 영역이며, 상기 REC1 및 REC2 도메인은 gRNA 내의 형성되는 이중가닥의, 예를 들어, 단일가닥 gRNA, 이중 gRNA 또는 tracrRNA의 인식에 중요한 역할을 한다.

[713] 상기 NUC 로브는 RuvC 도메인, HNH 도메인 및 PAM-상호작용(PI) 도메인을 포함한다. 이때, 상기 RuvC 도메인은 RuvC-유사 도메인을 포괄하는 의미로 사용되고, 또한 상기 HNH 도메인은 HNH-유사 도메인을 포괄하는 의미로 사용된다.

[714] 이때, 상기 RuvC 도메인은 Type II CRISPR 효소를 포함하는 자연상태에 존재하는 미생물의 구성원에 대해 구조적으로 유사성을 공유하며, 단일가닥, 예를 들어 표적 유전자의 전사 조절 영역 내 핵산의 비상보성 가닥, 즉, gRNA와 상보적인 결합을 하지 않는 가닥을 절단한다. 상기 RuvC 도메인은 종종 당업계에서 RuvCI 도메인, RuvCII 도메인 및 RuvCIII 도메인으로서, 통상적으로 RuvC I, RuvCII 및 RuvCIII로 지칭된다.

[715] 상기 HNH 도메인은 HNH 엔도뉴클레아제와 구조적 유사성을 공유하며, 단일가닥, 예를 들어 표적 핵산 분자의 상보성 가닥, 즉, gRNA와 상보적인 결합을 하는 가닥을 절단한다. HNH 도메인은 RuvC II와 III 모티프 사이에 위치한다.

[716] 상기 PI 도메인은 표적 유전자의 전사 조절 영역 내의 특정 뉴클레오타이드서열, 즉, PAM(Protospacer adjacent motif)을 인식하거나 또는 PAM과 상호작용한다. 이때, 상기 PAM은 Type II CRISPR 효소의 유래(origin)에 따라 다를 수 있다. 예를 들어, CRISPR 효소가 SpCas9 인 경우 PAM은 5'-NGG-3'일 수 있고, 스트렙토코커스 써모필러스 Cas9(StCas9)인 경우 PAM은 5'-NNAGAAW-3'(W = A or T)일 수 있고, 네이세리아 메닝기디티스 Cas9(NmCas9)인 경우 PAM은 5'-NNNNGATT-3'일 수 있고, 캄필로박터 제주니 Cas9(CjCas9)의 경우 PAM은 5'-NNNVRYAC-3' (V = G or C or A, R = A or G, Y =

C or T)일 수 있으며, 이때 상기 N은 A, T, G 또는 C; 또는 A, U, G 또는 C일 수 있다. 다만, 전술한 효소의 유래에 따라 PAM이 결정되는 것으로 일반적으로 이해되고 있으나, 해당 유래의 효소의 돌연변이(mutant)에 대한 연구가 진행됨에 따라, 상기 PAM은 달라질 수도 있다.

[717] 상기 Type II CRISPR 효소는 Cas9일 수 있다.

[718] 상기 Cas9은 스트렙토코커스 피오게네스(Streptococcus pyogenes), 스트렙토코커스 써모필러스(Streptococcus thermophilus), 스트렙토코커스 속(Streptococcus sp.), 스타필로코커스 아우레스(Staphylococcus aureus), 캄필로박터 제주니 (Campylobacter jejuni), 노카르디옵시스 다손빌레이(Nocardiopsis dassonvillei), 스트렙토마이세스 프리스티네스피랄리스(Streptomyces pristinaespiralis), 스트렙토마이세스 비리도크로모게네스(Streptomyces viridochromogenes), 스트렙토마이세스 비리도크로모게네스(Streptomyces viridochromogenes), 스트렙토스포랑기움 로세움(Streptosporangium roseum), 스트렙토스포랑기움 로세움(Streptosporangium roseum), 알리사이클로바클루스 아시도칼다리우스(AlicyclobacHlus acidocaldarius), 바실러스 슈도마이코이데스(Bacillus pseudomycoides), 바실러스 셀레니티레두센스(Bacillus selenitireducens), 엑시구오박테리움 시비리쿰(Exiguobacterium sibiricum), 락토바실러스 델브루에키이(Lactobacillus delbrueckii), 락토바실러스 살리바리우스(Lactobacillus salivarius), 미크로스 킬라 마리나(Microscilla marina), 부르크홀데리아레스 박테리움(Burkholderiales bacterium), 폴라로모나스 나프탈레니보란스(Polaromonas naphthalenivorans), 폴라로모나스 속(Polaromonas sp.), 크로코스파에라 와트소니이(Crocosphaera watsonii), 시아노테세 속(Cyanothece sp.), 마이크로시스티스 아에루기노사(Microcystis aeruginosa), 시네코코커스 속(Synechococcus sp.), 아세토할로비움 아라바티쿰(Acetohalobium arabaticum), 암모니펙스 데겐시이(Ammonifex degensii), 칼디셀룰로시럽토 베시이(Caldicelulosiruptor bescii), 칸디다투스 데술포루디스(Candidatus Desulforudis), 클로스트리듐 보툴리눔(Clostridium botulinum), 클로스트리듐 디피실레(Clostridium difficile), 피네골디아 마그나(Finegoldia magna), 나트라나에로비우스 써모필러스 (Natranaerobius thermophilus), 펠로토마쿨럼 써모프로피오니쿰(Pelotomaculum thermopropionicum), 아시디티오바실러스 칼두스(Acidithiobacillus caldus), 아시디티오바실러스 페로옥시단스(Acidithiobacillus ferrooxidans), 알로크로마티움 비노숨(Allochromatium vinosum), 마리노박터 속(Marinobacter sp.), 니트로소코커스 할로필러스(Nitrosococcus halophilus), 니트로소코커스 와트소니(Nitrosococcus watsoni), 슈도알테로 모나스 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), 써모시포 아프리카누스(Thermosipho africanus) 또는 아카리오클로리스 마리나(Acaryochloris marina) 등 다양한 미생물 유래의 Cas9일 수 있다.

[719] 상기 Cas9은 gRNA와 결합하여 표적 유전자의 전사 조절 영역 내에서 표적 서열 또는 위치를 절단 또는 변형시키는 효소로서, gRNA가 상보적인 결합을 하는 핵산 가닥(strand)을 절단할 수 있는 HNH 도메인, gRNA와 비상보적인 결합을 하는 핵산 가닥(strand)을 절단할 수 있는 RuvC 도메인, 표적, 즉, 타겟과 상호작용하는 REC 도메인 및 PAM을 인식하는 PI 도메인으로 구성될 수 있다. 구체적인 Cas9의 구조적 특성은 Hiroshi Nishimasu et al. (2014) Cell 156:935-949를 참고할 수 있다.

[720] 상기 Cas9은 자연상태에서 존재하는 미생물에서 분리된 것 또는 재조합적 방법 또는 합성적 방법을 통해 비자연적으로 생산된 것일 수 있다.

[721]

[722] 또한, 상기 CRISPR 효소는 Type V CRISPR 효소일 수 있다.

[723] Type V CRISPR 효소는 Type II CRISPR 효소의 RuvC 도메인에 상응하는 유사한 RuvC 도메인이 있으며, Type II CRISPR 효소의 HNH 도메인은 결핍되어 있고, 대신에 Nuc 도메인을 포함하며, 표적과 상호작용하는 REC 도메인과 WED 도메인 및 PAM을 인식하는 PI 도메인으로 구성될 수 있다. 구체적인 Type V CRISPR 효소의 구조적 특성은 Takashi Yamano et al. (2016) Cell 165:949-962를 참고할 수 있다.

[724] Type V CRISPR 효소는 gRNA와 상호작용할 수 있으며, gRNA-CRISPR 효소 복합체, 즉, CRISPR 복합체를 형성할 수 있고, gRNA와 협력하여 가이드 서열을 및 PAM 서열을 포함하는 표적 서열로 근접시킬 수 있다. 이때, 표적 유전자의 전사 조절 영역 내 핵산과 상호작용하기 위한 Type V CRISPR 효소의 능력은 PAM 서열에 의존적이다.

[725] 상기 PAM 서열은 표적 유전자의 전사 조절 영역 내에 존재하는 서열로, Type V CRISPR 효소의 PI 도메인에 의해 인식될 수 있다. 상기 PAM 서열은 Type V CRISPR 효소의 유래에 따라 그 서열이 다를 수 있다. 즉, 종마다 특이적으로 인식할 수 있는 PAM 서열이 존재한다. 예를 들어, Cpf1이 인식하는 PAM 서열은 5'-TTN-3' (N은 A, T, C 또는 G)일 수 있다. 다만, 전술한 효소의 유래에 따라 PAM이 결정되는 것으로 일반적으로 이해되고 있으나, 해당 유래의 효소의 돌연변이(mutant)에 대한 연구가 진행됨에 따라, 상기 PAM은 달라질 수도 있다.

[726] 상기 Type V CRISPR 효소는 Cpf1일 수 있다.

[727] 상기 Cpf1은 Streptococcus, Campylobacter, Nitratifractor, 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, Brevibacilus, Methylobacterium 또는 Acidaminococcus 유래의 Cpf1일 수 있다.

[728] 상기 Cpf1은 Cas9의 RuvC 도메인에 상응하는 유사한 RuvC 도메인이 있으며, Cas9의 HNH 도메인은 결핍되어 있고, 대신에 Nuc 도메인을 포함하며, 타겟과 상호작용하는 REC 도메인과 WED 도메인 및 PAM을 인식하는 PI 도메인으로 구성될 수 있다. 구체적인 Cpf1의 구조적 특성은 Takashi Yamano et al. (2016) Cell 165:949-962를 참고할 수 있다.

[729] 상기 Cpf1는 자연상태에서 존재하는 미생물에서 분리된 것 또는 재조합적 방법 또는 합성적 방법을 통해 비자연적으로 생산된 것일 수 있다.

[730]

[731] 상기 CRISPR 효소는 표적 유전자의 전사 조절 영역 내 핵산의 이중가닥을 절단하는 기능을 가지는 뉴클레아제 또는 제한효소일 수 있다.

[732] 상기 CRISPR 효소는 완전 활성 CRISPR 효소일 수 있다.

[733] "완전 활성"는 야생형(wild type) CRISPR 효소의 기능과 동일한 기능을 가지고 있는 상태를 의미하며, 이러한 상태의 CRISPR 효소를 완전 활성 CRISPR 효소로 명칭한다. 이때, “야생형(wild type) CRISPR 효소의 기능”은 DNA의 이중 가닥을 절단하는 기능, 즉, DNA의 이중 가닥 중 제1 가닥을 절단하는 제1 기능 및 DNA의 이중 가닥 중 제2 가닥을 절단하는 제2 기능을 모두 가지는 상태를 말한다.

[734] 상기 완전 활성 CRISPR 효소는 DNA의 이중 가닥을 절단하는 야생형 CRISPR 효소일 수 있다.

[735] 상기 완전 활성 CRISPR 효소는 DNA의 이중 가닥을 절단하는 야생형 CRISPR 효소를 변형 또는 조작시킨 CRISPR 효소 변이체일 수 있다.

[736] 상기 CRISPR 효소 변이체는 야생형 CRISPR 효소의 아미노산 서열 중 하나 이상의 아미노산이 다른 아미노산으로 치환되거나 또는 하나 이상의 아미노산이 제거된 효소일 수 있다.

[737] 상기 CRISPR 효소 변이체는 야생형 CRISPR 효소의 아미노산 서열에 하나 이상의 아미노산이 부가된 효소일 수 있다. 이때, 부가되는 아미노산의 위치는 야생형 효소의 N 말단, C 말단 또는 아미노산 서열 내일 수 있다.

[738] 상기 CRISPR 효소 변이체는 야생형 CRISPR 효소보다 기능이 향상된 완전 활성 효소일 수 있다.

[739] 예를 들면, 야생형 CRISPR 효소의 특정 변형 또는 조작된 형태, 즉, CRISPR

효소 변이체는 절단해야 하는 DNA 이중 가닥과 결합하지 않거나 또는 일정한 거리 간격을 유지한 상태에서 DNA 이중 가닥을 절단할 수 있다. 이러한 경우, 상기 변형 또는 조작된 형태는 야생형 CRISPR 효소보다 기능 활성이 향상된 완전 활성 CRISPR 효소일 수 있다.

[740] 상기 CRISPR 효소 변이체는 야생형 CRISPR 효소보다 기능이 감소된 완전 활성 CRISPR 효소일 수 있다.

[741] 예를 들면, 야생형 CRISPR 효소의 특정 변형 또는 조작된 형태, 즉, CRISPR 효소 변이체는 절단해야 하는 DNA 이중 가닥과 일정 거리이상으로 근접한 상태 또는 특정 결합이 형성된 상태에서 DNA 이중 가닥을 절단할 수 있다. 이때, 특정 결합은, 예를 들어, 효소의 특정 위치의 아미노산과 절단 위치의 DNA 뉴클레오타이드서열과의 결합일 수 있다. 이러한 경우, 상기 변형 또는 조작된 형태는 야생형 CRISPR 효소보다 기능 활성이 감소된 완전 활성 CRISPR 효소일 수 있다.

[742] 상기 CRISPR 효소는 불완전 또는 부분 활성 CRISPR 효소일 수 있다.

[743] "불완전 또는 부분 활성"은 야생형 CRISPR 효소의 기능, 즉, DNA의 이중 가닥 중 제1 가닥을 절단하는 제1 기능 및 DNA의 이중 가닥 중 제2 가닥을 절단하는 제2 기능 중 선택된 하나의 기능을 가지는 상태를 의미한다. 이러한 상태의 CRISPR 효소는 불완전 또는 부분 활성 CRISPR 효소로 명칭한다. 또한 상기 불완전 또는 부분 활성 CRISPR 효소는 니카아제(nickase)로 지칭될 수 있다.

[744] "니카아제(nickase)"는 표적 유전자의 전사 조절 영역 내 핵산의 이중가닥 중 한 가닥만 절단되도록 조작 또는 변형된 CRISPR 효소를 의미하며, 상기 니카아제는 단일가닥, 예를 들어, 표적 유전자의 전사 조절 영역 내 핵산의 gRNA와 비상보성 가닥 또는 상보성 가닥을 절단하는 뉴클레아제 활성을 가진다. 따라서, 이중가닥을 절단하기 위해서는 2개의 니카아제의 뉴클레아제 활성이 필요하다.

[745] 상기 니카아제는 CRISPR 효소의 RuvC 도메인에 의한 뉴클레아제 활성을 가질 수 있다. 즉, 상기 니카아제는 CRISPR 효소의 HNH 도메인에 의한 뉴클레아제 활성을 포함하지 않을 수 있으며, 이를 위해 HNH 도메인은 조작 또는 변경될 수 있다.

[746] 일 예로, 상기 CRISPR 효소가 Type II CRISPR 효소일 때, 상기 니카아제는 변형된 HNH 도메인을 포함하는 Type II CRISPR 효소일 수 있다.

[747] 예를 들어, 상기 Type II CRISPR 효소가 야생형 SpCas9의 경우, 상기 니카아제는 야생형 SpCas9의 아미노산 서열 840번 히스티딘을 알라닌으로 변이(mutation)시켜 HNH 도메인의 뉴클레아제 활성이 불활성화된 SpCas9 변이체일 수 있다. 이때 생성된 니카아제, 즉, SpCas9 변이체는 RuvC 도메인에 의한 뉴클레아제 활성을 가지므로, 표적 유전자의 전사 조절 영역 내 핵산의 비상보성 가닥, 즉, gRNA와 상보적인 결합을 하지 않는 가닥을 절단할 수 있다.

[748] 또 다른 예를 들어, 상기 Type II CRISPR 효소가 야생형 CjCas9의 경우, 상기 니카아제는 야생형 CjCas9의 아미노산 서열 559번 히스티딘을 알라닌으로

변이(mutation)시켜 HNH 도메인의 뉴클레아제 활성이 불활성화된 CjCas9 변이체일 수 있다. 이때 생성된 니카아제, 즉, CjCas9 변이체는 RuvC 도메인에 의한 뉴클레아제 활성을 가지므로, 표적 유전자의 전사 조절 영역 내 핵산의 비상보성 가닥, 즉, gRNA와 상보적인 결합을 하지 않는 가닥을 절단할 수 있다.

[749] 또한, 상기 니카아제는 CRISPR 효소의 HNH 도메인에 의한 뉴클레아제 활성을 가질 수 있다. 즉, 상기 니카아제는 CRISPR 효소의 RuvC 도메인에 의한 뉴클레아제 활성을 포함하지 않을 수 있으며, 이를 위해 RuvC 도메인은 조작 또는 변경될 수 있다.

[750] 일 예로, 상기 CRISPR 효소가 Type II CRISPR 효소일 때, 상기 니카아제는 변형된 RuvC 도메인을 포함하는 Type II CRISPR 효소일 수 있다.

[751] 예를 들어, 상기 Type II CRISPR 효소가 야생형 SpCas9의 경우, 상기 니카아제는 야생형 SpCas9의 아미노산 서열 10번 아스파르트산을 알라닌으로 변이(mutation)시켜 RuvC 도메인의 뉴클레아제 활성이 불활성화된 SpCas9 변이체일 수 있다. 이때 생성된 니카아제, 즉, SpCas9 변이체는 HNH 도메인에 의한 뉴클레아제 활성을 가지므로, 표적 유전자의 전사 조절 영역 내 핵산의 상보성 가닥, 즉, gRNA와 상보적인 결합을 하는 가닥을 절단할 수 있다.

[752] 또 다른 예를 들어, 상기 Type II CRISPR 효소가 야생형 CjCas9의 경우, 상기 니카아제는 야생형 CjCas9의 아미노산 서열 8번 아스파르트산을 알라닌으로 변이(mutation)시켜 RuvC 도메인의 뉴클레아제 활성이 불활성화된 CjCas9 변이체일 수 있다. 이때 생성된 니카아제, 즉, CjCas9 변이체는 HNH 도메인에 의한 뉴클레아제 활성을 가지므로, 표적 유전자의 전사 조절 영역 내 핵산의 상보성 가닥, 즉, gRNA와 상보적인 결합을 하는 가닥을 절단할 수 있다.

[753] 상기 CRISPR 효소는 불활성 CRISPR 효소일 수 있다.

[754] "불활성"은 야생형 CRISPR 효소의 기능, 즉, DNA의 이중 가닥 중 제1 가닥을 절단하는 제1 기능 및 DNA의 이중 가닥 중 제2 가닥을 절단하는 제2 기능이 모두 상실된 상태를 의미한다. 이러한 상태의 CRISPR 효소는 불활성 CRISPR 효소로 명칭한다.

[755] 상기 불활성 CRISPR 효소는 야생형 CRISPR 효소의 뉴클레아제 활성을 가지는 도메인에 변이로 인한 뉴클레아제 불활성을 가질 수 있다.

[756] 상기 불활성 CRISPR 효소는 RuvC 도메인 및 HNH 도메인에 변이로 인한 뉴클레아제 불화성을 가질 수 있다. 즉, 상기 불활성 CRISPR 효소는 CRISPR 효소의 RuvC 도메인 및 HNH 도메인에 의한 뉴클레아제 활성을 포함하지 않을 수 있으며, 이를 위해 RuvC 도메인 및 HNH 도메인은 조작 또는 변경될 수 있다.

[757] 일 예로, 상기 CRISPR 효소가 Type II CRISPR 효소일 때, 상기 불활성 CRISPR 효소는 변형된 RuvC 도메인 및 HNH 도메인을 포함하는 Type II CRISPR 효소일 수 있다.

[758] 예를 들어, 상기 Type II CRISPR 효소가 야생형 SpCas9의 경우, 상기 불활성 CRISPR 효소는 야생형 SpCas9의 아미노산 서열 10번 아스파르트산과 840번

히스티딘을 모두 알라닌으로 변이(mutation)시켜 RuvC 도메인 및 HNH 도메인의 뉴클레아제 활성이 불활성화된 SpCas9 변이체일 수 있다. 이때 생성된 불활성 CRISPR 효소, 즉, SpCas9 변이체는 RuvC 도메인 및 HNH 도메인의 뉴클레아제 활성이 불활성 되므로, 표적 유전자의 전사 조절 영역 내 핵산의 이중가닥을 모두 절단할 수 없다.

[759] 또 다른 예를 들어, 상기 Type II CRISPR 효소가 야생형 CjCas9의 경우, 상기 불활성 CRISPR 효소는 야생형 CjCas9의 아미노산 서열 8번 아스파르트산과 559번 히스티딘을 모두 알라닌으로 변이(mutation)시켜 RuvC 도메인 및 HNH 도메인의 뉴클레아제 활성이 불활성화된 CjCas9 변이체일 수 있다. 이때 생성된 불활성 CRISPR 효소, 즉, CjCas9 변이체는 RuvC 도메인 및 HNH 도메인의 뉴클레아제 활성이 불활성 되므로, 표적 유전자의 전사 조절 영역 내 핵산의 이중가닥을 모두 절단할 수 없다.

[760]

[761] 상기 CRISPR 효소는 상기 기재된 뉴클레아제 활성 외에도 헬리카제 활성, 즉, 이중가닥 핵산의 나선 구조를 푸는 기능을 가질 수 있다.

[762] 또한, 상기 CRISPR 효소는 CRISPR 효소의 헬리카제 활성에 대해 완전 활성, 불완전 또는 부분 활성, 또는 불활성이 되도록 CRISPR 효소를 변형시킬 수 있다.

[763]

[764] 상기 CRISPR 효소는 야생형 CRISPR 효소를 인위적으로 조작 또는 변형시킨 CRISPR 효소 변이체일 수 있다.

[765] 상기 CRISPR 효소 변이체는 야생형 CRISPR 효소의 기능, 즉, DNA의 이중 가닥 중 제1 가닥을 절단하는 제1 기능 및/또는 DNA의 이중 가닥 중 제2 가닥을 절단하는 제2 기능을 변형시키기 위해 인위적으로 조작 또는 변형된 CRISPR 효소 변이체일 수 있다.

[766] 예를 들어, 상기 CRISPR 효소 변이체는 야생형 CRISPR 효소의 기능 중 제1 기능이 상실된 형태일 수 있다.

[767] 또는 상기 CRISPR 효소 변이체는 야생형 CRISPR 효소의 기능 중 제2 기능이 상실된 형태일 수 있다.

[768] 예를 들어, 상기 CRISPR 효소 변이체는 야생형 CRISPR 효소의 기능, 즉, 제1 기능 및 제2 기능이 모두 상실된 형태일 수 있다.

[769] 상기 CRISPR 효소 변이체는 gRNA와 상호작용을 통한 gRNA-CRISPR 효소 복합체 형성할 수 있다.

[770] 상기 CRISPR 효소 변이체는 야생형 CRISPR 효소의 gRNA와 상호작용하는 기능을 변형시키기 위해 인위적으로 조작 또는 변형된 CRISPR 효소 변이체일 수 있다.

[771] 예를 들어, 상기 CRISPR 효소 변이체는 야생형 CRISPR 효소에 비해 gRNA와의 상호작용이 감소된 형태일 수 있다.

[772] 또는 상기 CRISPR 효소 변이체는 야생형 CRISPR 효소에 비해 gRNA의

상호작용이 증가된 형태일 수 있다.

[773] 예를 들어, 상기 CRISPR 효소 변이체는 야생형 CRISPR 효소의 제1 기능을 가지면서 gRNA와의 상호작용이 감소된 형태일 수 있다.

[774] 또는 상기 CRISPR 효소 변이체는 야생형 CRISPR 효소의 제1 기능을 가지면서 gRNA와의 상호작용이 증가된 형태일 수 있다.

[775] 예를 들어, 상기 CRISPR 효소 변이체는 야생형 CRISPR 효소의 제2 기능을 가지면서 gRNA와의 상호작용이 감소된 형태일 수 있다.

[776] 또는 상기 CRISPR 효소 변이체는 야생형 CRISPR 효소의 제2 기능을 가지면서 gRNA와의 상호작용이 증가된 형태일 수 있다.

[777] 예를 들어, 상기 CRISPR 효소 변이체는 야생형 CRISPR 효소의 제1 기능 및 제2 기능을 가지지 않으면서 gRNA와의 상호작용이 감소된 형태일 수 있다.

[778] 또는 상기 CRISPR 효소 변이체는 야생형 CRISPR 효소의 제1 기능 및 제2 기능을 가지지 않으면서 gRNA와의 상호작용이 증가된 형태일 수 있다.

[779] 이때, gRNA와 CRISPR 효소 변이체의 상호작용 세기에 따라 다양한 gRNA-CRISPR 효소 복합체가 형성될 수 있고, CRISPR 효소 변이체에 따라 표적 서열에 접근 또는 절단하는 기능에 차이가 생길 수 있다.

[780] 예를 들어, gRNA와의 상호작용이 감소된 CRISPR 효소 변이체에 의해 형성된 gRNA-CRISPR 효소 복합체는 gRNA와 완전히 상보적 결합을 하는 표적 서열에 근접 또는 국소화되는 경우에만 오직 표적 서열의 이중가닥 또는 단일가닥을 절단할 수 있다.

[781] 상기 CRISPR 효소 변이체는 야생형 CRISPR 효소의 아미노산 서열 중 적어도 하나 이상의 아미노산이 변형된 것일 수 있다.

[782] 일 예로, 상기 CRISPR 효소 변이체는 야생형 CRISPR 효소의 아미노산 서열 중 적어도 하나 이상의 아미노산의 치환된 것일 수 있다.

[783] 다른 일 예로, 상기 CRISPR 효소 변이체는 야생형 CRISPR 효소의 아미노산 서열 중 적어도 하나 이상의 아미노산이 제거된 것일 수 있다.

[784] 또 다른 일 예로, 상기 CRISPR 효소 변이체는 야생형 CRISPR 효소의 아미노산 서열 내에 적어도 하나 이상의 아미노산이 부가된 것일 수 있다.

[785] 일 예로, 상기 CRISPR 효소 변이체는 야생형 CRISPR 효소의 아미노산 서열 중 적어도 하나 이상의 아미노산의 치환, 제거 및/또는 부가된 것일 수 있다.

[786] 또한, 상기 CRISPR 효소 변이체는 야생형 CRISPR 효소의 원래 기능, 즉, DNA의 이중 가닥 중 제1 가닥을 절단하는 제1 기능 및 DNA의 이중 가닥 중 제2 가닥을 절단하는 제2 기능, 이외에 선택적으로 기능적(functional) 도메인을 추가로 포함할 수 있다. 이때, 상기 CRISPR 효소 변이체는 야생형 CRISPR 효소의 원래 기능 이외에 부가적인 기능을 가질 수 있다.

[787] 상기 기능적 도메인은 메틸라아제(methylase) 활성, 디메틸라아제(demethylase) 활성, 전사촉진(transcription activation) 활성, 전사 저해(transcription repression) 활성, 전사 방출 인자(transcription release factor) 활성, 히스톤 변형(histone

modification) 활성, RNA 절단(cleavage) 활성 또는 핵산 결합(nucleic acid binding) 활성을 가지는 도메인일 수 있으며, 또는 단백질(펩타이드 포함)의 분리정제를 위한 태그(tag) 또는 리포터 유전자일 수 있으나, 이에 제한되지 않는다.

[788] 상기 태그는 히스티딘(His) 태그, V5 태그, FLAG 태그, 인플루엔자 헤마글루티닌(HA) 태그, Myc 태그, VSV-G 태그 및 티오레독신(Trx) 태그 등을 포함하며, 상기 리포터 유전자는 글루타티온-S-트랜스 퍼라제(GST), 호스라디시(horseradish) 과산화효소(HRP), 클로람페니콜 아세틸트랜스퍼라제(CAT) 베타-갈락토시다제, 베타-글루쿠로니다제, 루시퍼라제, 녹색 형광 단백질(GFP), HcRed, DsRed, 청록색 형광 단백질(CFP), 황색 형광 단백질(YFP) 및 청색 형광 단백질(BFP)을 포함하는 자가형광 단백질을 포함하나, 이들에 한정되지 않는다.

[789] 상기 기능적 도메인은 디아미네이즈(deaminase)일 수 있다.

[790] 예를 들어, 불완전 또는 부분 CRISPR 효소에 시티딘 디아미네이즈(cytidine deaminase)를 기능적 도메인으로 추가로 포함할 수 있다. 일 구체예로, SpCas9 니카아제에 시티딘 디아미네이즈, 예를 들면, APOBEC1(apolipoprotein B editing complex 1)를 추가하여 융합 단백질을 생성할 수 있다. 이렇게 형성된 [SpCas9 니카아제]-[APOBEC1]은 뉴클레오타이드 C를 T 또는 U로 뉴클레오타이드 교정 또는 편집에 이용되거나, 또는 뉴클레오타이드 G를 A로 뉴클레오타이드 교정 또는 편집에 이용될 수 있다.

[791] 또 다른 예를 들어, 불완전 또는 부분 CRISPR 효소에 아데닌 디아미네이즈(adenine deaminase)를 기능적 도메인으로 추가로 포함할 수 있다. 일 구체예로, SpCas9 니카아제에 아데닌 디아미네이즈, 예를 들면, TadA variants, ADAR2 variants, ADAT2 variants 등을 추가하여 융합 단백질을 생성할 수 있다. 이렇게 형성된 [SpCas9 니카아제]-[TadA variant], [SpCas9 니카아제]-[ADAR2 variant] 또는 [SpCas9 니카아제]-[ADAT2 variant]는 뉴클레오타이드 A를 inosine으로 변형시키며, 변형된 inosine은 polymerase에 의해 뉴클레오타이드 G로 인식되어 실질적으로 뉴클레오타이드 A를 G로 뉴클레오타이드 교정 또는 편집되는 효과를 보이므로, 뉴클레오타이드 A를 G로 뉴클레오타이드 교정 또는 편집에 이용되거나, 또는 뉴클레오타이드 T를 C로 뉴클레오타이드 교정 또는 편집에 이용될 수 있다.

[792] 상기 기능적 도메인은 NLS(nuclear localization sequence or signal) 또는 NES(nuclear export sequence or signal)일 수 있다.

[793] 일 예로, CRISPR 효소는 하나 이상의 NLS를 포함할 수 있다. 이때, 상기 NLS는 CRISPR 효소의 아미노 말단 또는 그 근처; 카르복시 말단 또는 그 근처; 또는 이들의 조합에 하나 이상의 NLS를 포함할 수 있다. 상기 NLS는 하기로부터 유래된 NLS 서열일 수 있으나, 이에 제한되지 않는다: 아미노산 서열 PKKKRKV를 갖는 SV40 바이러스 대형 T-항원의 NLS; 뉴클레오플라스민(nucleoplasmin)으로부터의 NLS(예를 들어, 서열

KRPAATKKAGQAKKKK를 갖는 뉴클레오플라스민 이분(bipartite) NLS); 아미노산 서열 PAAKRVKLD 또는 RQRRNELKRSP를 갖는 c-myc NLS; 서열 NQSSNFGPMKGGNFGGRSSGPYGGGGQYFAKPRNQGGY를 갖는 hRNPA1 M9 NLS; 임포틴-알파로부터의 IBB 도메인의 서열 RMRIZFKNKGKDTAELRRRRVEVSVELRKAKKDEQILKRRNV; 마이오마(myoma) T 단백질의 서열 VSRKRPRP 및 PPKKARED; 인간 p53의 서열 POPKKKPL; 마우스 c-abl IV의 서열 SALIKKKKKMAP; 인플루엔자 바이러스 NS1의 서열 DRLRR 및 PKQKKRK; 간염 바이러스 델타 항원의 서열 RKLKKKIKKL; 마우스 Mx1 단백질의 서열 REKKKFLKRR; 인간 폴리(ADP-리보스) 중합효소의 서열 KRKGDEVDGVDEVAKKKSKK; 및 스테로이드 호르몬 수용체(인간) 글루코코르티코이드의 서열 RKCLQAGMNLEARKTKK.

[794] 또한, 상기 CRISPR 효소 변이체는 CRISPR 효소를 분할하여 두 개 이상의 부분으로 나눈 스플릿(split) 형태의 CRISPR 효소를 포함할 수 있다. "스플릿(split)"은 단백질을 기능적으로 또는 구조적으로 또는 임의로 두 개 이상으로 분할하는 것을 의미한다.

[795] 상기 스플릿(split) 형태의 CRISPR 효소는 완전 활성 효소, 불완전 또는 부분 활성 효소 또는 불활성 효소일 수 있다.

[796] 예를 들어, 상기 CRISPR 효소가 SpCas9의 경우, 656번 타이로신과 657번 트레오닌 사이를 분할하여 두 개의 부분으로 나눈 스플릿 SpCas9을 생성할 수 있다.

[797] 상기 스플릿(split) 형태의 CRISPR 효소는 선택적으로 재구성(reconstitution)을 위한 추가 도메인, 펩타이드, 폴리펩타이드 또는 단백질을 포함할 수 있다.

[798] 상기 재구성을 위한 추가 도메인, 펩타이드 폴리펩타이드 또는 단백질은 스플릿(split) 형태의 CRISPR 효소가 구조적으로 야생형 CRISPR 효소와 동일하거나 유사하도록 조립될 수 있다.

[799] 상기 재구성(reconstitution)을 위한 추가 도메인, 펩타이드, 폴리펩타이드 또는 단백질은 FRB 및 FKBP dimerization domains; 인테인(intein); ERT 및 VPR domains 또는 특정 조건에서 이량이질체(heterodimer)를 형성하는 도메인일 수 있다.

[800] 예를 들어, 상기 CRISPR 효소가 SpCas9의 경우, 713번 세린과 714번 글라이신 사이를 분할하여 두 개의 부분으로 나눈 스플릿 SpCas9에 두 부분 중 하나에 FRB 도메인을 연결하고, 나머지 하나에 FKBP 도메인을 연결할 수 있다. 이렇게 생성된 스플릿 SpCas9은 라파마이신이 존재하는 환경에서 FRB 도메인과 FKBP 도메인이 다이머를 형성하여 재구성된 CRISPR 효소를 생성할 수 있다.

[801]

[802] 본 명세서에 의해 개시되는 CRISPR 효소 또는 CRISPR 효소 변이체는 폴리펩타이드, 단백질 또는 이를 암호화하는 서열을 가지는 핵산일 수 있으며,

상기 CRISPR 효소 또는 CRISPR 효소 변이체를 도입하고자 하는 대상에 맞추어 코돈 최적화(codon optimization)된 것일 수 있다.

[803] "코돈 최적화"는 고유 서열의 적어도 하나의 코돈을 숙주 세포의 유전자에 더욱 빈번하게 또는 가장 빈번하게 사용되는 코돈으로 대체하면서, 고유 아미노상 서열을 유지함으로써 관심 숙주 세포에서의 발현의 증진을 위해 핵산서열을 변형시키는 과정을 의미한다. 다양한 종은 특정 아미노산의 특정 코돈에 대한 특정 편향을 가지며, 코돈 편향(유기체 간의 코돈 사용의 차이)은 종종 mRNA의 번역의 효율과 상호관련 되며, 이는 번역되는 코돈의 특성 및 특정 tRNA 분자의 이용가능성에 의해 좌우되는 것을 여겨진다. 세포에서 선택된 tRNA의 우세는 일반적으로 펩타이드 합성에 가장 빈번하게 사용되는 코돈을 반영한 것이다. 따라서, 유전자는 코돈 최적화에 기초하여 주어진 유기체에서 최적의 유전자 발현을 위해 맞춤화될 수 있다.

[804]

[805] 본 명세서에 의해 개시되는 gRNA, CRISPR 효소 또는 gRNA-CRISPR 효소 복합체는 다양한 형태로 대상 내로 전달 또는 도입될 수 있다.

[806] 상기 대상 관련 설명은 상기 기술한 바와 같다.

[807] 일 구현예로서, gRNA 및/또는 CRISPR 효소는 각각을 암호화하는 핵산서열을 포함하는 벡터에 의해 대상 내로 전달 또는 도입될 수 있다.

[808] 벡터는 gRNA 및/또는 CRISPR 효소를 암호화하는 핵산서열을 포함할 수 있다.

[809] 일 예로, 상기 벡터는 gRNA 및 CRISPR 효소를 암호화하는 핵산서열을 동시에 포함할 수 있다.

[810] 다른 일 예로, 상기 벡터는 gRNA 암호화하는 핵산서열을 포함할 수 있다.

[811] 예를 들어, 상기 gRNA에 포함되는 도메인은 하나의 벡터에 모두 포함되거나 또는 각각의 도메인을 나누어 각각의 벡터에 포함시킬 수 있다.

[812] 다른 일 예로, 상기 벡터는 CRISPR 효소를 암호화하는 핵산서열을 포함할 수 있다.

[813] 예를 들어, 상기 CRISPR 효소의 경우, CRISPR 효소를 암호화하는 핵산서열은 하나의 벡터에 포함되거나 또는 CRISPR 효소를 암호화하는 핵산서열은 분할되어 여러 개의 벡터에 포함될 수 있다.

[814] 상기 벡터는 하나 이상의 조절/제어 구성요소를 포함할 수 있다.

[815] 이때, 상기 조절/제어 구성요소는 포로모터, 인핸서, 인트론, 폴리아데닐화 신호, 코작 공통(Kozak consensus) 서열, 내부 리보솜 유입 부위(internal ribosome entry site, IRES), 스플라이스 억셉터 및/또는 2A 서열을 포함할 수 있다.

[816] 상기 프로모터는 RNA 중합효소 II에 의해 인식되는 프로모터일 수 있다.

[817] 상기 프로모터는 RNA 중합효소 III에 의해 인식되는 프로모터일 수 있다.

[818] 상기 프로모터는 유도성 프로모터일 수 있다.

[819] 상기 프로모터는 대상 특이적 프로모터일 수 있다.

[820] 상기 프로모터는 바이러스 또는 비바이러스 프로모터일 수 있다.

[821] 상기 프로모터는 제어 영역(즉, gRNA 및/또는 CRISPR 효소를 암호화하는 핵산서열)에 따라 적합한 프로모터를 이용할 수 있다.

[822] 예를 들어, gRNA를 위해 유용한 프로모터는 H1, EF-1a, tRNA 또는 U6 프로모터일 수 있다. 예를 들어, CRISPR 효소를 위해 유용한 프로모터는 CMV, EF-1a, EFS, MSCV, PGK 또는 CAG 프로모터일 수 있다.

[823] 상기 벡터는 바이러스 벡터 또는 재조합 바이러스 벡터일 수 있다.

[824] 상기 바이러스는 DNA 바이러스 또는 RNA 바이러스일 수 있다.

[825] 이때, 상기 DNA 바이러스는 이중가닥 DNA(dsDNA) 바이러스 또는 단일가닥 DNA(ssDNA) 바이러스 일 수 있다.

[826] 이때, 상기 RNA 바이러스는 단일가닥 RNA(ssRNA) 바이러스일 수 있다.

[827] 상기 바이러스는 레트로바이러스, 렌티바이러스, 아데노바이러스, 아데노-연관 바이러스(AAV), 백시니아바이러스, 폭스바이러스 또는 단순포진 바이러스일 수 있으나, 이에 제한되지 않는다.

[828] 일 예로, gRNA 및/또는 CRISPR 효소를 암호화하는 핵산서열은 재조합 렌티바이러스에 의해 전달 또는 도입될 수 있다.

[829] 다른 일 예로, gRNA 및/또는 CRISPR 효소를 암호화하는 핵산서열은 재조합 아데노바이러스에 의해 전달 또는 도입될 수 있다.

[830] 또 다른 일 예로, gRNA 및/또는 CRISPR 효소를 암호화하는 핵산서열은 재조합 AAV에 의해 전달 또는 도입될 수 있다.

[831] 다른 일 예로, gRNA 및/또는 CRISPR 효소를 암호화하는 핵산서열은 혼성 바이러스, 예를 들어 본 명세서에 기재한 바이러스 중 하나 이상의 혼성체에 의해 전달 또는 도입될 수 있다.

[832]

[833] 일 구현예로서, gRNA-CRISPR 효소 복합체의 형태로 대상 내로 전달 또는 도입될 수 있다.

[834] 예를 들어, 상기 gRNA는 DNA, RNA 또는 이의 혼합 형태일 수 있다. 상기 CRISPR 효소는 펩타이드, 폴리펩타이드 또는 단백질의 형태일 수 있다.

[835] 일 예로, gRNA 및 CRISPR 효소는 RNA 형태의 gRNA와 단백질 형태의 CRISPR gRNA-CRISPR 효소 복합체, 즉 ribonucleoprotein(RNP)의 형태로 대상 내로 전달 또는 도입될 수 있다.

[836] 상기 CRISPR gRNA-CRISPR 효소 복합체는 전기천공법, 미량주사법, 일시적인 세포 압축 또는 스퀴징(예를 들어, 문헌[Lee, et al, (2012) Nano Lett., 12, 6322-6327]에 기재되어 있음), 지질-매개 형질감염, 나노파티클, 리포솜, 펩타이드-매개 전달 또는 이의 조합에 의해 대상 내로 전달 또는 도입될 수 있다.

[837]

[838] **본 명세서에 의해 개시되는 gRNA-CRISPR 효소 복합체는 표적 유전자, 즉, 중복 유전자의 전사 조절 영역을 인위적으로 조작 또는 변형시키는데 이용될 수 있다.**

[839] 표적 유전자의 전사 조절 영역은 앞서 기재한 gRNS-CRISPR 효소 복합체, 즉, CRISPR 복합체를 이용하여 조작 또는 변형할 수 있다. 이때, 표적 유전자의 전사 조절 영역의 조작 또는 변형은 i) 표적 유전자의 전사 조절 영역의 절단 또는 손상과 ii) 손상된 전사 조절 영역의 수선 또는 수복하는 단계를 모두 포함한다.

[840] 상기 i) 표적 유전자의 전사 조절 영역의 절단 또는 손상은 CRISPR 복합체를 이용한 표적 유전자의 전사 조절 영역의 절단 또는 손상일 수 있으며, 구체적으로는 전사 조절 영역 내의 표적 서열의 절단 손상일 수 있다.

[841] 상기 표적 서열은 gRNA-CRISPR 효소 복합체의 타겟이 될 수 있고, 상기 표적 서열은 CRISPR 효소가 인식하는 PAM 서열을 포함하거나 또는 포함하지 않을 수 있다. 이러한 표적 서열은 실시자에게 gRNA 설계 단계에서 중요한 기준을 제공할 수 있다.

[842] 상기 표적 서열은 gRNA-CRISPR 효소 복합체의 gRNA에 의해 특이적으로 인식될 수 있으며, 이로 인해 gRNA-CRISPR 효소 복합체는 인식된 표적 서열에 근접하게 위치할 수 있다.

[843] 표적 부위의 "절단(cleavage)"은 폴리뉴클레오타이드의 공유결합(covalent backbone)의 파손(breakage)을 의미한다. 절단은 포스포다이에스터(phosphodiester) 결합의 효소적 또는 화학적 가수분해를 포함하나, 이에 제한되지 않으며, 이외의 다양한 여러 가지 방법들에 의하여 수행될 수 있다. 단일가닥의 절단 및 이중가닥의 절단 모두 가능하며, 이중가닥의 절단은 두 개의 구별되는(distinct) 단일-가닥의 절단의 결과로서 발생할 수 있다. 이중 가닥의 절단은 blunt ends 또는 staggered end(또는 sticky end)를 생성할 수 있다.

[844] 일 예로, 상기 CRISPR 복합체를 이용한 표적 유전자의 전사 조절 영역의 절단 또는 손상은 표적 서열의 이중가닥이 모두 절단 또는 손상되는 것일 수 있다.

[845] 일 구체예로서, CRISPR 효소가 야생형의 SpCas9인 경우, CRISPR 복합체는 gRNA와 상보적 결합을 하는 표적 서열의 이중가닥을 모두 절단시킬 수 있다.

[846] 다른 일 구체예로서, CRISPR 효소가 SpCas9 니카아제(D10A)와 SpCas9 니카아제(H840A)인 경우, 각각의 CRISPR 복합체는 gRNA와 상보적 결합을 하는 표적 서열의 두 개의 단일가닥을 각각 절단시킬 수 있다. 즉, gRNA와 상보적 결합을 하는 표적 서열의 상보성 단일가닥은 SpCas9 니카아제(D10A)에 의해 절단되고, gRNA와 상보적 결합을 하는 표적 서열의 비상보성 단일가닥은 SpCas9 니카아제(H840A)에 의해 절단될 수 있고, 각각의 절단은 순차적으로 또는 동시에 발생할 수 있다.

[847] 다른 일 예로, 상기 CRISPR 복합체를 이용한 표적 유전자의 전사 조절 영역의 절단 또는 손상은 표적 서열의 이중가닥 중 단일가닥만 절단 또는 손상되는 것일 수 있다. 이때, 단일가닥은 gRNA와 상보적 결합을 하는 표적 서열의 가이드핵산 결합 서열, 즉, 상보성 단일가닥일 수 있고, 또는 gRNA와 상보적 결합을 하지 않는 가이드핵산 비결합 서열, 즉, gRNA와 비상보성 단일가닥일 수 있다.

[848] 일 구체예로서, CRISPR 효소가 SpCas9 니카아제(D10A)인 경우, CRISPR 복합체는 gRNA와 상보적 결합을 하는 표적 서열의 가이드핵산 결합 서열, 즉, 상보성 단일가닥은 SpCas9 니카아제(D10A)에 의해 절단시키고, gRNA와 상보적 결합을 하지 않는 표적 서열의 가이드핵산 비결합 서열, 즉, gRNA와 비상보성 단일가닥은 절단시키지 않을 수 있다.

[849] 다른 일 구체예로서, CRISPR 효소가 SpCas9 니카아제(H840A)인 경우, CRISPR 복합체는 gRNA와 상보적 결합을 하지 않는 표적 서열의 가이드핵산 비결합 서열, 즉, gRNA와 비상보성 단일가닥은 SpCas9 니카아제(H840A)에 의해 절단시키고, gRNA와 상보적 결합을 하는 표적 서열의 가이드핵산 결합 서열, 즉, 상보성 단일가닥은 절단시키지 않을 수 있다.

[850] 또 다른 일 예로, 상기 CRISPR 복합체를 이용한 표적 유전자의 전사 조절 영역의 절단 또는 손상은 일부 핵산 조각(fragment)를 제거하는 것일 수 있다.

[851] 일 구체예로서, 상기 CRISPR 복합체가 서로 다른 표적 서열에 상보적 결합을 하는 두 개의 gRNA와 야생형 SpCas9에 의해 구성되는 경우, 제 1 gRNA와 상보적 결합을 하는 표적 서열의 이중가닥을 절단하고, 제 2 gRNA와 상보적 결합을 하는 표적 서열의 이중가닥을 절단하여 제 1 gRNA와 제 2 gRNA 및 SpCas9에 의해 핵산 조각을 제거할 수 있다.

[852] 예를 들어, 두개의 CRISPR 복합체가 서로 다른 표적 서열에 상보적 결합을 하는 두 개의 gRNA, 예를 들어, 인핸서의 상류(upstream)에 존재하는 표적 서열에 상보적 결합을 하는 하나의 gRNA 및 인핸서의 하류(downstream)에 존재하는 표적 서열에 상보적 결합을 하는 하나의 gRNA와 야생형 SpCas9에 의해 각각 구성되는 경우, 제 1 gRNA와 상보적 결합을 하는 인핸서의 상류(upstream)에 존재하는 표적 서열의 이중가닥을 절단하고, 제 2 gRNA와 상보적 결합을 하는 인핸서의 하류(downstream)에 존재하는 표적 서열의 이중가닥을 절단하여, 제 1 gRNA와 제 2 gRNA 및 SpCas9에 의해 핵산 조각, 즉, 인핸서 영역을 제거할 수 있다.

[853]

[854] 상기 ii) 손상된 전사 조절 영역의 수선 또는 수복은 비-상동성 말단-결합 (NHEJ) 및 상동 재조합 수리(HDR)을 통해 수선 또는 수복될 수 있다.

[855] 상기 비-상동성 말단-결합 (Non-homologous end joining, NHEJ)은 절단된 이중가닥 또는 단일가닥의 양 말단이 함께 결합함으로써 DNA 내 이중가닥 파손을 수복 또는 수선하는 방법으로, 일반적으로, 이중가닥의 파손(예를 들어, 절단)에 의해 형성된 2 개의 적합성 말단이 빈번한 접촉을 반복하여 2개의 말단이 완전히 결합되는 경우 파손된 이중가닥이 복구된다. NHEJ는 모든 세포주기에서 가능한 수복 방식으로, 주로 G1 시기와 같이 세포 내에 주형으로 쓸 상동유전체가 없을 때 발생한다.

[856] NHEJ를 이용한 손상된 전사 조절 영역의 수복 과정에서 NHEJ 수선 부위에 핵산 서열의 일부 삽입 및/또는 결실(삽입결실)을 초래하며, 이러한 삽입 및/또는

결실은 리딩 프레임을 변경시키고, 프레임쉬프트 된 전사체 mRNA를 만들어내고, 결과적으로 넌센스-매개 붕괴(nonsense mediated decay)를 겪거나 정상적인 단백질을 합성하는데 실패함으로써 본래의 기능을 상실하게 된다. 또는 추가적으로, 리딩 프레임을 유지하지만, 상당한 양의 서열을 삽입 또는 결실시키는 돌연변이를 초래해 단백질의 기능성을 파괴할 수 있다. 이는 중요한 기능적 도메인의 돌연변이가 단백질의 비중요 영역에서의 돌연변이보다 덜 용인될 가능성이 있기 때문에 좌위 의존적이다.

[857] NHEJ에 의해 생성된 삽입결실 돌연변이는 자연 상태에서 예측 불가능하지만, 주어진 파손 부위에서 특정 삽입 결실 서열이 선호되며, 이는 마이크로상동성의 작은 영역에 기인할 가능성이 있다. 통상적으로 결실 길이는 1 bp 내지 50 bp 범위이며, 삽입은 더 짧게 되는 경향이 있고, 종종 파손 부위를 바로 둘러싸는 짧은 중복 서열을 포함한다.

[858] 또한, NHEJ는 돌연변이를 유발하는 과정으로, 특이적 최종 서열의 생성이 필요하지 않은 경우, 작은 서열의 모티프를 결실시키는데 사용될 수 있다.

[859] 이러한 NHEJ를 이용하면, CRISPR 복합체에 의해 표적되는 전사 조절 영역에 의해 발현이 조절되는 표적 유전자의 특이적 넉아웃(knockout)할 수 있다. CRISPR 효소, 예를 들어, Cas9 또는 Cpf1을 이용하여 표적 유전자의 전사 조절 영역의 이중가닥 또는 두 개의 단일가닥의 절단하고, 파손된 전사 조절 영역의 이중가닥 또는 두 개의 단일가닥은 NHEJ에 의해 삽입결실이 생성되며, 이를 통해 전사 조절 영역에 의해 발현이 조절되는 표적 유전자의 특이적 넉아웃을 유도할 수 있다.

[860] 일 예로, CRISPR 복합체를 이용하여 표적 유전자의 전사 조절 영역의 이중가닥을 절단하고 NHEJ에 의해 수복되는 과정을 통해 수복 부위에 다양한 인델 (indel; insertion and deletion)이 발생할 수 있다.

[861] "인델(indel)"은 DNA의 뉴클레오타이드 배열에서 일부 뉴클레오타이드가 중간에 삽입 (insertion)되거나 결실 (deletion) 된 변이를 총칭한다. 인델은 상술한 바와 같이 gRNA-CRISPR 효소 복합체가 표적 유전자의 전사 조절 영역 내 표적 서열을 절단하는 경우, HDR 또는 NHEJ 기작에 의해 수선되는 과정에서 표적 서열에 도입되는 것일 수 있다.

[862] 상기 상동 재조합 수리(homology directed repairing, HDR)는 손상된 전사 조절 영역의 DNA를 수선 또는 수복하기 위해 상동성을 가진 서열을 주형으로 이용하는 방식으로 오류 없이 교정할 수 있는 방법으로, 일반적으로, 파손된 DNA을 수선 또는 수복하기 위해, 즉 세포가 가지고 있는 원래의 정보를 복원하기 위해, 변형이 이루어지지 않은 상보적인 뉴클레오타이드서열의 정보를 이용하거나 자매 염색분체의 정보를 이용하여 파손된 DNA를 수선 또는 수복한다. HDR의 가장 일반적인 형태는 상동성 재조합(homologous recombination, HR)이다. HDR은 통상적으로 활발하게 분열하는 세포의 S나 G2/M 시기에 주로 발생하는 수선 또는 수복 방식이다.

[863] HDR을 이용한 손상된 DNA 수선 또는 수복을 위해, 세포가 본래 가지는 상보적인 뉴클레오타이드서열 또는 자매 염색분체를 이용하는 대신에, 상보적인 뉴클레오타이드서열 또는 상동성 뉴클레오타이드서열 정보를 이용한 인공적으로 합성한 DNA 주형, 즉, 상보적인 뉴클레오타이드서열 또는 상동성 뉴클레오타이드서열을 포함하는 핵산 주형을 세포에 제공하여 파손된 DNA를 수선 또는 수복할 수 있다. 이때, 상기 핵산 주형에 추가로 핵산 서열 또는 핵산 조각을 포함시켜 파손된 DNA를 수선 또는 수복 할 때, 파손된 DNA에 추가로 포함시킨 핵산 서열 또는 핵산 조작을 삽입(Knock-In)할 수 있다. 추가로 포함시킨 핵산 서열 또는 핵산 조각은 돌연변이의 변형된 표적 유전자의 전사 조절 영역을 정상적으로 교정하기 위한 핵산 서열 또는 핵산 조각이거나 세포 내에서 발현을 원하는 유전자 또는 핵산일 수 있으나, 이에 제한되지 않는다.

[864] 일 예로, CRISPR 복합체를 이용하여 표적 유전자의 전사 조절 영역의 이중가닥 또는 단일가닥을 절단하고, 절단 위치와 근접한 뉴클레오타이드서열과 상보적인 뉴클레오타이드서열을 포함하는 핵산 주형을 세포에 제공하여 표적 유전자의 전사 조절 영역의 절단된 뉴클레오타이드서열을 HDR 방법으로 수선 또는 수복할 수 있다.

[865] 이때, 상기 상보적인 뉴클레오타이드서열을 포함하는 핵산 주형은 파손된 DNA, 즉 절단된 이중가닥 또는 단일가닥의 상보적인 뉴클레오타이드서열을 가지며, 추가로 파손된 DNA에 삽입하기 원하는 핵산 서열 또는 핵산 조각을 포함할 수 있다. 이와 같이 상보적인 뉴클레오타이드서열과 삽입하고자 하는 핵산 서열 또는 핵산 조각을 포함하는 핵산 주형을 이용하여 파손된 DNA, 즉 표적 유전자의 전사 조절 영역의 절단된 위치에 추가적인 핵산 서열 또는 핵산 조각을 삽입할 수 있다. 이때, 상기 삽입하고자 하는 핵산 서열 또는 핵산 조각 및 추가적인 핵산 서열 또는 핵산 조각은 돌연변이의 변형된 표적 유전자의 전사 조절 영역을 정상적으로 교정하기 위한 핵산 서열 또는 핵산 조각이거나 세포 내에서 발현을 원하는 유전자 또는 핵산일 수 있다. 상기 상보적인 뉴클레오타이드서열은 파손된 DNA, 즉 표적 유전자의 전사 조절 영역의 절단된 이중가닥 또는 단일가닥의 오른쪽 및 왼쪽의 뉴클레오타이드서열과 상보적인 결합을 하는 뉴클레오타이드서열일 수 있다. 또는 상기 상보적인 뉴클레오타이드서열은 파손된 DNA, 즉 전사 조절 영역의 절단된 이중가닥 또는 단일가닥의 3' 및 5' 말단과 상보적인 결합을 하는 뉴클레오타이드서열일 수 있다. 상기 상보적인 뉴클레오타이드서열은 15 내지 3000개의 뉴클레오타이드서열일 수 있으며, 핵산 주형의 크기 또는 표적 유전자의 전사 조절 영역에 따라 적절하게 상기 상보적인 뉴클레오타이드서열의 길이 또는 크기를 설계할 수 있다. 이때, 핵산 주형을 이중가닥 또는 단일가닥의 핵산일 수 있으며, 선형 또는 원형일 수 있으나, 이에 제한되지 않는다.

[866] 다른 일 예로, CRISPR 복합체를 이용하여 표적 유전자의 전사 조절 영역의 이중가닥 또는 단일가닥을 절단하고, 절단 위치와 근접한

뉴클레오타이드서열과 상동성 뉴클레오타이드서열을 포함하는 핵산 주형을 세포에 제공하여 표적 유전자의 전사 조절 영역의 절단된 뉴클레오타이드서열을 HDR 방법으로 수선 또는 수복할 수 있다.

[867] 이때, 상기 상동성 뉴클레오타이드 서열을 포함하는 핵산 주형은 파손된 DNA, 즉 절단된 이중가닥 또는 단일가닥의 상동성 뉴클레오타이드 서열을 가지며, 추가로 파손된 DNA에 삽입하기 원하는 핵산 서열 또는 핵산 조각을 포함할 수 있다. 이와 같이 상동성 뉴클레오타이드 서열과 삽입하고자 하는 핵산 서열 또는 핵산 조각을 포함하는 핵산 주형을 이용하여 파손된 DNA, 즉 전사 조절 영역의 절단된 위치에 추가적인 핵산 서열 또는 핵산 조각을 삽입할 수 있다. 이때, 상기 삽입하고자 하는 핵산 서열 또는 핵산 조각 및 추가적인 핵산 서열 또는 핵산 조각은 돌연변이의 변형된 표적 유전자의 전사 조절 영역을 정상적으로 교정하기 위한 핵산 서열 또는 핵산 조각이거나 세포 내에서 발현을 원하는 유전자 또는 핵산일 수 있다. 상기 상동성 뉴클레오타이드 서열은 파손된 DNA, 즉 전사 조절 영역의 절단된 이중가닥 또는 단일가닥의 오른쪽 및 왼쪽의 뉴클레오타이드 서열과 상동성을 가지는 뉴클레오타이드 서열일 수 있다. 또는 상기 상보적인 뉴클레오타이드 서열은 파손된 DNA, 즉 전사 조절 영역의 절단된 이중가닥 또는 단일가닥의 3' 및 5' 말단과 상동성을 가지는 뉴클레오타이드 서열일 수 있다. 상기 상동성 뉴클레오타이드 서열은 15 내지 3000개의 뉴클레오타이드 서열일 수 있으며, 핵산 주형의 크기 또는 표적 유전자의 전사 조절 영역에 따라 적절하게 상기 상동성 뉴클레오타이드 서열의 길이 또는 크기를 설계할 수 있다. 이때, 핵산 주형을 이중가닥 또는 단일가닥의 핵산일 수 있으며, 선형 또는 원형일 수 있으나, 이에 제한되지 않는다.

[868]

[869] 상기 NHEJ와 HDR 외에도 손상된 전사 조절 영역을 수선 또는 수복하는 다양한 방법이 존재한다. 예를 들어, 손상된 전사 조절 영역의 수선 또는 수복 방법은 단일가닥 어닐링, 단일가닥 파손 수선, 미스매치 수선, 뉴클레오타이드 절단 수선 또는 뉴클레오타이드 절단 수선을 이용한 방법일 수 있다.

[870] 상기 단일가닥 어닐링(Single-strand annealing, SSA)은 표적 핵산 중에 존재하는 2개의 반복부 서열 사이의 이중가닥 파손을 수선하는 방법으로, 일반적으로 30개 초과의 뉴클레오타이드서열의 반복부 서열을 이용할 수 있다. 파손 말단에서 표적 핵산의 이중가닥에 대해 반복부 서열을 각각의 단일가닥이 가지도록 절단(Sticky end가 되도록)되며, 절단 후 반복부 서열을 함유하는 단일가닥 돌출부는 반복체 서열이 자체적으로 부적절하게 어닐링되는 것을 방지하기 위해 RPA 단백질로 코팅된다. RAD52는 돌출부 상의 각각의 반복부 서열에 결합하고, 상보성 반복부 서열의 어닐링을 가능하게 하는 서열을 정렬한다. 어닐링 후에, 돌출부의 단일가닥 플랩(flap)은 절단되고, 새로운 DNA 합성이 임의의 갭을 채우면서 DNA 이중가닥을 복원한다. 이러한 수복 또는 수선의 결과로, 2개의 반복체 사이의 DNA 서열이 결실되며, 결실 길이는

이용되는 2개의 반복체의 위치를 포함하는 다수의 인자 및 절단 경로 또는 진행도에 의존될 수 있다.

[871] SSA는 HDR 방식과 유사하게는 상보성 서열, 즉 상보성 반복부 서열을 이용하고, 대조적으로는 표적 핵산 서열을 변경 또는 수정하기 위한 핵산 주형을 필요로 하지 않는다.

[872] 게놈 내 단일가닥 파손은 상기 논의된 수선 메커니즘과는 별도의 메커니즘인 단일가닥 파손 수선(Single-strand break repair, SSBA)을 통해 수선 또는 수복될 수 있다. DNA 파손의 형태가 단일가닥 파손일 때, PARP1 및/또는 PARP2는 파손을 인식하고 수선 기작을 동원한다. DAN 파손에서 PARP1의 결합 및 활성을 일시적이며, 손상부에 SSBR 단백질 복합체의 안정성을 촉진함으로써 SSBR을 촉진시킨다. SSBR 복합체에서 가장 중요한 단백질은 XRCC1으로, 이는 DNA의 3' 및 5' 말단 가공을 촉진하는 단백질과 상호작용하며, 안정화시킨다. 말단 가공은 일반적을 손상된 3' 말단을 하이드록실화 된 상태 및/또는 5' 말단을 인산염 모이어티로 복구하는 것을 수반하며, 말단이 가공되면, DNA 갭 채우기가 일어난다. DNA 갭 채우기에는 2가지 방법, 즉, 짧은 패치 수선 및 긴 패치 수선이 있으며, 이때 짧은 패치 수선은 빠져있는 단일 뉴클레오타이드의 삽입을 수반한다. DNA 갭 채우기 후, DNA 리가아제는 말단의 결합을 촉진한다.

[873] 상기 미스매치 수선(Mismatch repair, MMR)은 잘못 짝지어진 DNA 뉴클레오타이드상에서 작용할 수 있다. MSH2/6 또는 MSH2/3 복합체는 둘 다 미스매치 인식 및 수선의 개시에서 중요한 역할을 하는 ATP 분해효소 활성을 가지며, MSH2/6는 뉴클레오타이드-뉴클레오타이드 미스매치를 우선적으로 인식하고, 1개 또는 2개의 뉴클레오타이드의 미스매치를 동정하는 반면, MSH2/3는 더 큰 미스매치를 우선적으로 인식한다.

[874] 상기 뉴클레오타이드 절단 수선(Base excision repair, BER)은 세포주기 전체에서 활성이며, 게놈으로부터 작은 비-나선-뒤틀림 뉴클레오타이드 손상부를 제거하는데 이용되는 수선 방식이다. 손상된 DNA는 당 인산화 백본에 뉴클레오타이드를 연결하는 N-글리코사이드 결합을 절단하여 손상된 뉴클레오타이드를 절단하고, 이어서 포스포디에스테르 백본을 절단하여 DNA 단일가닥 파손을 생성한다. 이렇게 형성된 파손된 단일가닥 말단을 제거하고, 제거된 단일가닥에 의해 발생된 갭을 새로운 상보성 뉴클레오타이드로 채운 후, DNA 리가아제로 새로 채워진 상보성 뉴클레오타이드 말단과 백본을 결합시켜 손상된 DNA를 수선 또는 수복한다.

[875] 상기 뉴클레오타이드 절단 수선(Nucleotide excision repair, NER)은 DNA로부터 큰 나선-뒤틀림 손상을 제거하는 중요한 절단 메커니즘으로, 손상이 인식되면, 손상부를 함유하는 짧은 단일가닥 DNA 세그먼트를 제거하여, 22개 내지 30개 뉴클레오타이드의 단일가닥 갭을 생성한다. 생성된 갭은 새로운 상보성 뉴클레오타이드로 채운 후, DNA 리가아제로 새로 채워진 상보성 뉴클레오타이드 말단과 백본을 결합시켜 손상된 DNA를 수선 또는 수복한다.

[876]

[877] 상기 gRNA-CRISPR 효소 복합체에 의한 표적 유전자의 전사 조절 영역의 인위적인 조작 효과는 크게 넉아웃(knockout), 넉다운(knockdown), 넉인(knockin) 및 발현증가일 수 있다.

[878] "넉아웃(knockout)"은 표적 유전자 또는 핵산을 불활성화시키는 것을 의미하며, “표적 유전자 또는 핵산의 불활성화”는 표적 유전자 또는 핵산의 전사 및/또는 번역이 되지 못하는 상태를 의미한다. 넉아웃을 통해 질병을 유발하는 유전자 또는 비정상적 기능을 가지는 유전자의 전사 및 번역을 억제하여 단백질의 발현을 막을 수 있다.

[879] 예를 들어, gRNA-CRISPR 효소 복합체, 즉, CRISPR 복합체를 이용하여 표적 유전자의 전사 조절 영역을 편집 또는 교정하는 경우, CRISPR 복합체를 이용하여 표적 유전자의 전사 조절 영역을 절단할 수 있다. CRISPR 복합체를 이용하여 손상된 전사 조절 영역은 NHEJ를 통해 손상된 전사 조절 영역이 수복될 수 있다. 파손된 전사 조절 영역은 NHEJ에 의해 삽입결실이 생성되며, 이를 통해 전사 조절 영역이 불활성되어 표적 유전자 또는 염색체의 특이적 넉아웃을 유도할 수 있다.

[880] 다른 예를 들어, gRNA-CRISPR 효소 복합체, 즉, CRISPR 복합체 및 도너을 이용하여 표적 유전자의 전사 조절 영역을 편집 교정하는 경우, CRISPR 복합체를 이용하여 표적 유전자의 전사 조절 영역을 절단할 수 있다. CRISPR 복합체를 이용하여 손상된 전사 조절 영역은 도너를 이용해 HDR을 통해 손상된 전사 조절 영역이 수복될 수 있다. 이때, 도너는 상동성 뉴클레오타이드서열 및 삽입시키길 원하는 뉴클레오타이드서열을 포함한다. 이때, 삽입시키길 원하는 뉴클레오타이드 서열의 개수는 삽입 위치 또는 목적에 따라 다양하게 조절될 수 있다. 도너를 이용해 손상된 전자 조절 영역이 수복된 경우, 손상된 뉴클레오타이드서열 부위에 삽입시키길 원하는 뉴클레오타이드서열이 삽입되며, 이를 통해 전사 조절 영역이 불활성되어 표적 유전자 또는 염색체의 특이적 넉아웃을 유도할 수 있다.

[881] "넉다운(knockdown)"은 표적 유전자 또는 핵산의 전사 및/또는 번역을 감소시키거나 표적 단백질의 발현이 감소하는 것을 의미한다. 넉다운을 통해 과발현되는 유전자 또는 단백질의 발현을 조절하여 질병의 발생을 막거나 질병을 치료할 수 있다.

[882] 예를 들어, gRNA-CRISPR 효소 복합체, 즉, CRISPR 복합체를 이용하여 표적 유전자의 전사 조절 영역을 편집 또는 교정하는 경우, CRISPR 복합체를 이용하여 표적 유전자의 전사 조절 영역을 절단할 수 있다. CRISPR 복합체를 이용하여 손상된 전사 조절 영역은 NHEJ를 통해 손상된 전사 조절 영역이 수복될 수 있다. 파손된 전사 조절 영역은 NHEJ에 의해 삽입결실이 생성되며, 이를 통해 전사 조절 영역이 부분적으로 불활성되어 표적 유전자 또는 염색체의 특이적 넉다운을 유도할 수 있다.

[883] 다른 예를 들어, gRNA-CRISPR 효소 복합체, 즉, CRISPR 복합체 및 도너을 이용하여 표적 유전자의 전사 조절 영역을 편집 교정하는 경우, CRISPR 복합체를 이용하여 표적 유전자의 전사 조절 영역을 절단할 수 있다. CRISPR 복합체를 이용하여 손상된 전사 조절 영역은 도너를 이용해 HDR을 통해 손상된 전사 조절 영역이 수복될 수 있다. 이때, 도너는 상동성 뉴클레오타이드서열 및 삽입시키길 원하는 뉴클레오타이드서열을 포함한다. 이때, 삽입시키길 원하는 뉴클레오타이드 서열의 개수는 삽입 위치 또는 목적에 따라 다양하게 조절될 수 있다. 도너를 이용해 손상된 전자 조절 영역이 수복된 경우, 손상된 뉴클레오타이드서열 부위에 삽입시키길 원하는 뉴클레오타이드서열이 삽입되며, 이를 통해 전사 조절 영역이 부분적으로 불활성되어 표적 유전자 또는 염색체의 특이적 넉다운을 유도할 수 있다.

[884] 예를 들어, gRNA- CRISPR 불활성 효소-전사 저해 활성 도메인 복합체, 즉, 전사 저해 활성 도메인을 포함하는 CRISPR 불활성 복합체를 이용하여 표적 유전자의 전사 조절 영역을 편집 또는 교정하는 경우, 상기 CRISPR 불활성 복합체가 표적 유전자의 전사 조절 영역에 특이적으로 결합하고, CRISPR 불활성 복합체에 포함된 전사 저해 활성 도메인에 의해 전사 조절 영역의 활성이 저해되어 표적 유전자 또는 염색체의 발현이 저해되는 넉다운을 유도할 수 있다.

[885] "넉인(knockin)"은 표적 유전자 또는 핵산에 특정 핵산 또는 유전자를 삽입하는 것을 의미하며, 이때, "특정 핵산 또는 유전자"은 삽입하고자 하는 또는 발현시키기 원하는 핵산 또는 유전자를 의미한다. 넉인을 통해 질병을 유발하는 돌연변이 유전자를 올바르게 교정하거나, 정상 유전자를 삽입하여 정상 유전자의 발현을 유도하여 질병 치료에 이용할 수 있다.

[886] 더불어, 넉인은 추가적으로 도너(donor)를 필요로 할 수 있다.

[887] 예를 들어, gRNA-CRISPR 효소 복합체, 즉, CRISPR 복합체 및 도너를 이용하여 표적 유전자 또는 핵산을 편집 또는 교정하는 경우, CRISPR 복합체를 이용하여 표적 유전자 또는 핵산을 절단할 수 있다. CRISPR 복합체를 이용하여 손상된 표적 유전자 또는 핵산은 도너를 이용해 HDR를 통해 손상된 유전자 또는 핵산이 수복될 수 있다. 이때, 도너는 특정 핵산 또는 유전자를 포함하며, 이를 이용하여 손상된 유전자 또는 염색체에 특정 핵산 또는 유전자를 삽입할 수 있다. 이때, 삽입된 특정 핵산 또는 유전자는 단밸질로의 발현이 유도할 수 있다.

[888] “발현증가”는 인위적으로 조작되기 전과 비교하여 표적 유전자 또는 핵산의 전사 및/또는 번역을 증가시키거나 표적 단백질의 발현이 증가하는 것을 의미한다. 발현증가를 통해 저발현되거나 발현되지 않는 유전자 또는 단백질의 발현을 조절하여 질병의 발생을 막거나 질병을 치료할 수 있다.

[889] 예를 들어, gRNA-CRISPR 효소 복합체, 즉, CRISPR 복합체를 이용하여 표적 유전자의 전사 조절 영역을 편집 또는 교정하는 경우, CRISPR 복합체를 이용하여 표적 유전자의 전사 조절 영역을 절단할 수 있다. CRISPR 복합체를 이용하여 손상된 전사 조절 영역은 NHEJ를 통해 손상된 전사 조절 영역이

수복될 수 있다. 파손된 전사 조절 영역은 NHEJ에 의해 삽입결실이 생성되며, 이때 생성된 삽입결실로 인해 전사 조절 영역의 활성이 증가하여 정상적인 표적 유전자 또는 염색체의 발현을 유도할 수 있다.

[890]

[891] **본 명세서에 의해 개시되는 일 구체예로서, gRNA-CRISPR 효소 복합체는 중복 유전자의 전사 조절 영역에 인위적인 조작 또는 변형을 가할 수 있다.**

[892] 상기 gRNA-CRISPR 효소 복합체는 중복 유전자의 전사 조절 영역 내 표적 서열을 특이적으로 인식할 수 있다.

[893] 이때, 상기 중복 유전자는 PMP22 유전자, PLP1 유전자, MECP2 유전자, SOX3 유전자, RAI1 유전자, TBX1 유전자, ELN 유전자, JAGGED1 유전자, NSD1 유전자, MMP23 유전자, LMB1 유전자, SNCA 유전자 및 APP 유전자로 구성된 군에서 선택된 하나 이상의 유전자일 있다.

[894] 또는, 상기 중복 유전자는 MYC 유전자, ERBB2(HER2) 유전자, CCND1(Cyclin D1) 유전자, FGFR1 유전자, FGFR2 유전자, HRAS 유전자, KRAS 유전자, MYB 유전자, MDM2 유전자, CCNE(Cyclin E) 유전자, MET 유전자, CDK4 유전자, ERBB1 유전자, MYCN 유전자 및 AKT2 유전자로 구성된 군에서 선택된 하나 이상의 유전자일 수 있다.

[895] 상기 전사 조절 영역 관련 설명을 상기 기술한 바와 같다.

[896] 상기 표적 서열은 gRNA-CRISPR 효소 복합체의 gRNA에 의해 특이적으로 인식될 수 있으며, 이로 인해 gRNA-CRISPR 효소 복합체는 인식된 표적 서열에 근접하게 위치할 수 있다.

[897] 상기 표적 서열은 중복 유전자의 전사 조절 영역 내에 인위적인 변형이 일어나는 부위 또는 영역일 수 있다.

[898] 상기 표적 서열 관련 설명을 상기 기술한 바와 같다.

[899] 일 구체예에서, 상기 표적 서열은 표 1, 표 2, 표 3 및 표 4에 기재된 뉴클레오타이드 서열 중 선택된 하나 이상의 뉴클레오타이드 서열일 수 있다.

[900] 상기 gRNA-CRISPR 효소 복합체는 gRNA와 CRISPR 효소로 구성될 수 있다.

[901] 상기 gRNA는 중복 유전자의 전사 조절 영역 내 표적 서열 중 가이드핵산 결합 서열과 일부 또는 완전한 상보적 결합을 할 수 있는 가이드 도메인을 포함할 수 있다.

[902] 상기 가이드 도메인은 가이드핵산 결합 서열에 최소한 70%, 75%, 80%, 85%, 90% 또는 95% 이상의 상보성이거나 또는 완전하게 상보성일 수 있다.

[903] 상기 가이드 도메인은 중복 유전자의 전사 조절 영역 내 표적 서열 중 가이드핵산 결합 서열에 상보적인 뉴클레오타이드서열을 포함할 수 있다. 이때, 상기 상보적인 뉴클레오타이드서열은 0 내지 5, 0 내지 4, 0 내지 3, 0 내지 2개의 미스매치(mismatching)를 포함할 수 있다.

[904] 상기 gRNA는 제 1 상보적 도메인, 연결 도메인, 제 2 상보적 도메인, 근위 도메인 및 꼬리 도메인으로 구성된 군에서 선택된 하나 이상의 도메인을 포함할

수 있다.

[905] 상기 CRISPR 효소는 스트렙토코커스 피요게네스(Streptococcus pyogenes) 유래의 Cas9 단백질, 캄필로박터 제주니 (Campylobacter jejuni) 유래의 Cas9 단백질, 스트렙토코커스 써모필러스 (Streptococcus thermophilus) 유래의 Cas9 단백질, 스타필로코커스 아우레우스 (Staphylococcus aureus) 유래의 Cas9 단백질, 네이세리아 메닝기디티스 (Neisseria meningitidis)유래의 Cas9 단백질, 및 Cpf1 단백질로 이루어진 군에서 선택된 하나 이상일 수 있다. 일 예에서, 상기 에디터단백질은 캄필로박터 제주니 (Campylobacter jejuni) 유래의 Cas9 단백질 또는 스타필로코커스 아우레우스 (Staphylococcus aureus) 유래의 Cas9 단백질일 수 있다.

[906]

[907] 상기 gRNA-CRISPR 효소 복합체는 중복 유전자의 전사 조절 영역에 다양한 인위적인 조작 또는 변형을 가할 수 있다.

[908] 상기 인위적으로 조작 또는 중복 유전자의 전사 조절 영역은 표적 서열 내에 또는 표적 서열의 5' 말단 및/또는 3' 말단에 인접하여 위치하는 연속하는 1bp 내지 50bp의 뉴클레오타이드 서열 부위에 다음 중 하나 이상의 변형을 포함할 수 있다:

[909] i) 하나 이상의 뉴클레오타이드의 결실,

[910] ii) 하나 이상의 뉴클레오타이드의 야생형 유전자와 상이한 뉴클레오타이드로의 치환,

[911] iii) 하나 이상의 뉴클레오타이드의 삽입, 또는

[912] iv) 상기 i) 내지 iii) 중에서 선택된 2 가지 이상의 조합.

[913] 예를 들어, 인위적으로 조작 또는 변형된 중복 유전자의 전사 조절 영역은 표적 서열 내에 또는 표적 서열의 5' 말단 및/또는 3' 말단에 인접하여 위치하는 연속하는 1bp 내지 50bp의 뉴클레오타이드 서열 부위에 하나 이상의 뉴클레오타이드 결실을 포함할 수 있다. 일 예로, 결실되는 뉴클레오타이드는 연속되거나 연속되지 않는 1, 2, 3, 4 또는 5bp 일 수 있다. 다른 일 예로, 결실되는 뉴클레오타이드는 연속되는 2bp 이상의 뉴클레오타이드로 이루어진 뉴클레오타이드 조각일 수 있다. 이때, 뉴클레오타이드 조각은 2bp 내지 5bp, 6bp 내지 10bp, 11bp 내지 15bp, 16bp 내지 20bp, 21bp 내지 25bp, 26bp 내지 30bp, 31bp 내지 35bp, 36bp 내지 40bp, 41bp 내지 45bp 또는 46bp 내지 50bp일 수 있다. 또다른 일 예로, 결실되는 뉴클레오타이드는 2 이상의 뉴클레오타이드 조각일 수 있다. 이때, 2 이상의 뉴클레오타이드 조각은 뉴클레오타이드 서열이 연속되지 않는, 즉, 하나 이상의 뉴클레오타이드 서열 간격을 가지는 각각의 뉴클레오타이드 조각일 수 있으며, 결실된 2 이상의 뉴클레오타이드 조각에 의해 2 이상의 결실 부위가 발생할 수 있다.

[914] 또는, 예를 들어, 인위적으로 조작 또는 변형된 중복 유전자의 전사 조절 영역은 표적 서열 내에 또는 표적 서열의 5' 말단 및/또는 3' 말단에 인접하여 위치하는

연속하는 1bp 내지 50bp의 뉴클레오타이드 서열 부위에 하나 이상의 뉴클레오타이드 삽입을 포함할 수 있다. 일 예로, 삽입되는 뉴클레오타이드는 연속된 1, 2, 3, 4 또는 5bp 일 수 있다. 다른 일 예로, 삽입되는 뉴클레오타이드는 연속되는 5bp 이상의 뉴클레오타이드로 이루어진 뉴클레오타이드 조각일 수 있다. 이때, 뉴클레오타이드 조각은 5bp 내지 10bp, 11bp 내지 50bp, 50bp 내지 100bp, 100bp 내지 200bp, 200bp 내지 300bp, 300bp 내지 400bp, 400bp 내지 500bp, 500bp 내지 750bp 또는 750bp 내지 1000bp일 수 있다. 또 다른 일 예로, 삽입되는 뉴클레오타이드는 특정 유전자의 일부 또는 전체 뉴클레오타이드 서열일 수 있다. 이때, 특정 유전자는 중복 유전자를 포함하는 대상체, 예를 들어, 인간세포가 포함하지 않는 외부에서 유입된 유전자일 수 있다. 또는 특정 유전자는 중복 유전자를 포함하는 대상체, 예를 들어, 인간세포가 포함하는 유전자, 예를 들어, 인간세포 게놈에 존재하는 유전자일 수 있다.

[915] 또는, 예를 들어, 인위적으로 조작 또는 변형된 중복 유전자의 전사 조절 영역은 표적 서열 내에 또는 표적 서열의 5' 말단 및/또는 3' 말단에 인접하여 위치하는 연속하는 1bp 내지 50bp의 뉴클레오타이드 서열 부위에 하나 이상의 뉴클레오타이드 결실 및 삽입을 포함할 수 있다. 일 예로, 결실되는 뉴클레오타이드는 연속되거나 연속되지 않는 1, 2, 3, 4 또는 5bp 일 수 있다. 이때, 삽입되는 뉴클레오타이드는 1, 2, 3, 4 또는 5bp; 뉴클레오타이드 조각; 또는 특정 유전자의 일부 또는 전체 뉴클레오타이드 서열일 수 있으며, 상기 결실과 삽입은 순차적으로 일어나거나 동시에 발생할 수 있다. 이때, 삽입되는 뉴클레오타이드 조각은 5bp 내지 10bp, 11bp 내지 50bp, 50bp 내지 100bp, 100bp 내지 200bp, 200bp 내지 300bp, 300bp 내지 400bp, 400bp 내지 500bp, 500bp 내지 750bp 또는 750bp 내지 1000bp일 수 있다. 이때, 특정 유전자는 중복 유전자를 포함하는 대상체, 예를 들어, 인간세포가 포함하지 않는 외부에서 유입된 유전자일 수 있다. 또는 특정 유전자는 중복 유전자를 포함하는 대상체, 예를 들어, 인간세포가 포함하는 유전자, 예를 들어, 인간세포 게놈에 존재하는 유전자일 수 있다. 다른 일 예로, 결실되는 뉴클레오타이드는 2bp 이상의 뉴클레오타이드로 이루어진 뉴클레오타이드 조각일 수 있다. 이때, 결실되는 뉴클레오타이드 조각은 2bp 내지 5bp, 6bp 내지 10bp, 11bp 내지 15bp, 16bp 내지 20bp, 21bp 내지 25bp, 26bp 내지 30bp, 31bp 내지 35bp, 36bp 내지 40bp, 41bp 내지 45bp 또는 46bp 내지 50bp일 수 있다. 이때, 삽입되는 뉴클레오타이드는 1, 2, 3, 4 또는 5bp; 뉴클레오타이드 조각; 또는 특정 유전자의 일부 또는 전체 뉴클레오타이드 서열일 수 있으며, 상기 결실과 삽입은 순차적으로 일어나거나 동시에 발생할 수 있다. 또 다른 일 예로, 결실되는 뉴클레오타이드는 2 이상의 뉴클레오타이드 조각일 수 있다. 이때, 삽입되는 뉴클레오타이드는 1, 2, 3, 4 또는 5bp; 뉴클레오타이드 조각; 또는 특정 유전자의 일부 또는 전체 뉴클레오타이드 서열일 수 있으며, 상기 결실과 삽입은 순차적으로 일어나거나 동시에 발생할 수 있다. 또한, 상기 삽입은 결실된 2 이상의 부위에 일부 또는

전체 부위에서 발생할 수 있다.

[916]

[917] 상기 gRNA-CRISPR 효소 복합체는 gRNA 및 CRISPR 효소의 종류에 따라 중복 유전자의 전사 조절 영역에 다양한 인위적인 조작 또는 변형을 가할 수 있다.

[918] 일 예로, 상기 CRISPR 효소가 SpCas9 단백질인 경우, 상기 인위적인 조작 또는 변형된 중복 유전자의 전사 조절 영역은 표적 영역 내 존재하는 5'-NGG-3' (N은 A, T, G, 또는 C임) PAM 서열의 5' 말단 및/또는 3' 말단에 인접하여 위치하는 연속하는 1bp 내지 50bp, 1bp 내지 40bp, 1bp 내지 30bp, 바람직하게는 1bp 내지 25bp의 뉴클레오타이드 서열 부위 내에 다음 중 하나 이상의 변형을 포함할 수 있다:

[919] i) 하나 이상의 뉴클레오타이드의 결실,

[920] ii) 하나 이상의 뉴클레오타이드의 야생형 유전자와 상이한 뉴클레오타이드로의 치환,

[921] iii) 하나 이상의 뉴클레오타이드의 삽입, 또는

[922] iv) 상기 i) 내지 iii) 중에서 선택된 2 가지 이상의 조합.

[923] 다른 일 예로, 상기 CRISPR 효소가 CjCas9 단백질인 경우, 상기 인위적인 조작 또는 변형된 중복 유전자의 전사 조절 영역은 표적 영역 내 존재하는 5'-NNNNRYAC-3'(N은 각각 독립적으로 A, T, C 또는 G이고, R은 A또는 G이고, Y는 C 또는 T임) PAM 서열의 5' 말단 및/또는 3' 말단에 인접하여 위치하는 연속하는 1bp 내지 50bp, 1bp 내지 40bp, 1bp 내지 30bp, 바람직하게는 1bp 내지 25bp의 뉴클레오타이드 서열 부위 내에 다음 중 하나 이상의 변형을 포함할 수 있다:

[924] i) 하나 이상의 뉴클레오타이드의 결실,

[925] ii) 하나 이상의 뉴클레오타이드의 야생형 유전자와 상이한 뉴클레오타이드로의 치환,

[926] iii) 하나 이상의 뉴클레오타이드의 삽입, 또는

[927] iv) 상기 i) 내지 iii) 중에서 선택된 2 가지 이상의 조합.

[928] 또 다른 일 예로, 상기 CRISPR 효소가 StCas9 단백질인 경우, 상기 인위적인 조작 또는 변형된 중복 유전자의 전사 조절 영역은 표적 영역 내 존재하는 5'-NNAGAAW-3'(N은 각각 독립적으로 A, T, C 또는 G이고, W는 A 또는 T임) PAM 서열의 5' 말단 및/또는 3' 말단에 인접하여 위치하는 연속하는 1bp 내지 50bp, 1bp 내지 40bp, 1bp 내지 30bp, 바람직하게는 1bp 내지 25bp의 뉴클레오타이드 서열 부위 내에 다음 중 하나 이상의 변형을 포함할 수 있다:

[929] i) 하나 이상의 뉴클레오타이드의 결실,

[930] ii) 하나 이상의 뉴클레오타이드의 야생형 유전자와 상이한 뉴클레오타이드로의 치환,

[931] iii) 하나 이상의 뉴클레오타이드의 삽입, 또는

[932] iv) 상기 i) 내지 iii) 중에서 선택된 2 가지 이상의 조합.

[933] 일 예로, 상기 CRISPR 효소가 NmCas9 단백질인 경우, 상기 인위적인 조작 또는 변형된 중복 유전자의 전사 조절 영역은 표적 영역 내 존재하는 5'-NNNNGATT-3'(N은 각각 독립적으로 A, T, C 또는 G임) PAM 서열의 5' 말단 및/또는 3' 말단에 인접하여 위치하는 연속하는 1bp 내지 50bp, 1bp 내지 40bp, 1bp 내지 30bp, 바람직하게는 1bp 내지 25bp의 뉴클레오타이드 서열 부위 내에 다음 중 하나 이상의 변형을 포함할 수 있다:

[934] i) 하나 이상의 뉴클레오타이드의 결실,

[935] ii) 하나 이상의 뉴클레오타이드의 야생형 유전자와 상이한 뉴클레오타이드로의 치환,

[936] iii) 하나 이상의 뉴클레오타이드의 삽입, 또는

[937] iv) 상기 i) 내지 iii) 중에서 선택된 2 가지 이상의 조합.

[938] 다른 일 예로, 상기 CRISPR 효소가 SaCas9 단백질인 경우, 상기 인위적인 조작 또는 변형된 중복 유전자의 전사 조절 영역은 표적 영역 내 존재하는 5'-NNGRR(T)-3'(N은 각각 독립적으로 A, T, C 또는 G이고, R은 A또는 G이고, (T)는 임의로 포함가능한 서열을 의미함) PAM 서열의 5' 말단 및/또는 3' 말단에 인접하여 위치하는 연속하는 1bp 내지 50bp, 1bp 내지 40bp, 1bp 내지 30bp, 바람직하게는 1bp 내지 25bp의 뉴클레오타이드 서열 부위 내에 다음 중 하나 이상의 변형을 포함할 수 있다:

[939] i) 하나 이상의 뉴클레오타이드의 결실,

[940] ii) 하나 이상의 뉴클레오타이드의 야생형 유전자와 상이한 뉴클레오타이드로의 치환,

[941] iii) 하나 이상의 뉴클레오타이드의 삽입, 또는

[942] iv) 상기 i) 내지 iii) 중에서 선택된 2 가지 이상의 조합.

[943] 또 다른 일 예로, 상기 CRISPR 효소가 Cpf1 단백질인 경우, 상기 인위적인 조작 또는 변형된 중복 유전자의 전사 조절 영역은 표적 영역 내 존재하는 5'-TTN-3'(N은 A, T, C 또는 G임) PAM 서열의 5' 말단 및/또는 3' 말단에 인접하여 위치하는 연속하는 1bp 내지 50bp, 1bp 내지 40bp, 1bp 내지 30bp, 바람직하게는 1bp 내지 25bp의 뉴클레오타이드 서열 부위 내에 다음 중 하나 이상의 변형을 포함할 수 있다:

[944] i) 하나 이상의 뉴클레오타이드의 결실,

[945] ii) 하나 이상의 뉴클레오타이드의 야생형 유전자와 상이한 뉴클레오타이드로의 치환,

[946] iii) 하나 이상의 뉴클레오타이드의 삽입, 또는

[947] iv) 상기 i) 내지 iii) 중에서 선택된 2 가지 이상의 조합.

[948] 상기 gRNA-CRISPR 효소 복합체에 의한 중복 유전자의 전사 조절 영역의 인위적인 조작 효과는 넉아웃(knockout)일 수 있다.

[949] 상기 gRNA-CRISPR 효소 복합체에 의한 중복 유전자에 의해 각각 암호화되는 단백질의 발현이 억제될 수 있다.

[950] 상기 gRNA-CRISPR 효소 복합체에 의한 중복 유전자의 전사 조절 영역의 인위적인 조작 효과는 넉다운(knockdown)일 수 있다.

[951] 상기 gRNA-CRISPR 효소 복합체에 의한 중복 유전자에 의해 각각 암호화되는 단백질의 발현이 감소될 수 있다.

[952] 상기 gRNA-CRISPR 효소 복합체에 의한 중복 유전자의 전사 조절 영역의 인위적인 조작 효과는 넉인(knockin)일 수 있다.

[953] 이때, 상기 넉인 효과는 gRNA-CRISPR 효소 복합체 및 추가적으로 외부 유래 뉴클레오타이드서열 또는 유전자를 포함하는 도너에 의해 유도된 것일 수 있다.

[954] 상기 gRNA-CRISPR 효소 복합체 및 도너에 의한 중복 유전자의 전사 조절 영역의 인위적인 조작 효과는 외부 유래 뉴클레오타이드서열 또는 유전자에 의해 암호화되는 펩타이드 또는 단백질이 발현될 수 있다.

[955] 이때, 상기 넉인 효과는 gRNA-CRISPR 효소 복합체 및 삽입을 원하는 뉴클레오타이드 서열을 포함하는 도너에 의해 유도된 것일 수 있다.

[956]

[957] **본 명세서에 의해 개시되는 일 태양은 발현 조절 방법에 관한 것이다.**

[958] 본 명세서에 의해 개시되는 일 구체예는 중복 유전자의 발현을 조절하는 방법에 관한 것으로, 이것은 생체 내, 생체 외 또는 시험관 내에서 이루어질 수 있다.

[959] 일부 실시형태에서, 방법은 인간 또는 비인간 동물로부터 세포 또는 세포 집단을 시료 추출하는 단계, 및 세포 또는 세포들을 변형하는 단계를 포함한다. 배양은 생체 외에서 임의의 단계에서 일어날 수 있다. 세포 또는 세포들은 심지어 비인간 동물 또는 식물에 재도입될 수 있다.

[960] 상기 방법은 중복 유전자를 포함하는 진핵세포에 발현 조절용 조성물을 도입시키는 단계를 포함하는, 진핵세포를 인위적으로 조작하는 방법일 수 있다.

[961] 상기 발현 조절용 조성물 관련 설명은 상기 기술한 바와 같다.

[962] 일 구현예로, 상기 발현 조절용 조성물은 다음을 포함할 수 있다:

[963] (a) 중복 유전자의 전사 조절 영역 내 위치한 표적 서열을 표적화할 수 있는 가이드핵산 또는 이를 암호화하는 핵산 서열; 및

[964] (b) 스트렙토코커스 피요게네스(Streptococcus pyogenes) 유래의 Cas9 단백질, 캄필로박터 제주니 (Campylobacter jejuni) 유래의 Cas9 단백질, 스트렙토코커스 써모필러스 (Streptococcus thermophilus) 유래의 Cas9 단백질, 스타필로코커스 아우레우스 (Staphylococcus aureus) 유래의 Cas9 단백질, 네이세리아 메닝기디티스 (Neisseria meningitidis)유래의 Cas9 단백질, 및 Cpf1 단백질로 이루어진 군에서 선택된 하나 이상의 단백질을 포함하는 에디터단백질 또는 이를 암호화하는 핵산서열.

[965] 이때, 상기 중복 유전자는 PMP22 유전자, PLP1 유전자, MECP2 유전자, SOX3 유전자, RAI1 유전자, TBX1 유전자, ELN 유전자, JAGGED1 유전자, NSD1 유전자, MMP23 유전자, LMB1 유전자, SNCA 유전자 및 APP 유전자로 구성된

군에서 선택된 하나 이상의 유전자일 있다.

[966] 또는, 상기 중복 유전자는 MYC 유전자, ERBB2(HER2) 유전자, CCND1(Cyclin D1) 유전자, FGFR1 유전자, FGFR2 유전자, HRAS 유전자, KRAS 유전자, MYB 유전자, MDM2 유전자, CCNE(Cyclin E) 유전자, MET 유전자, CDK4 유전자, ERBB1 유전자, MYCN 유전자 및 AKT2 유전자로 구성된 군에서 선택된 하나 이상의 유전자일 수 있다.

[967] 상기 전사 조절 영역 관련 설명을 상기 기술한 바와 같다.

[968] 상기 가이드 핵산 및 에디터 단백질은, 각각 핵산 서열의 형태로 1 이상의 벡터에 존재하거나, 또는 가이드 핵산과 에디터 단백질의 결합으로 복합체를 형성하여 존재할 수 있다.

[969] 상기 발현 조절용 조성물은 삽입을 원하는 핵산서열을 포함하는 도너 또는 이를 암호화하는 핵산서열을 선택적으로 더 포함할 수 있다.

[970] 상기 가이드 핵산 및 에디터 단백질, 및/또는 도너는, 각각 핵산 서열의 형태로 1 이상의 벡터에 존재할 수 있다.

[971] 상기 도입 단계는 생체 내 또는 생체 외에서 수행될 수 있다.

[972] 예를 들어, 도입 단계는 전기천공법 (electroporation), 리포좀, 플라스미드, 바이러스벡터, 나노파티클(nanoparticles) 및 PTD (Protein translocation domain) 융합 단백질 방법 중 선택되는 1이상의 방법으로 수행될 수 있다.

[973] 예를 들어, 바이러스 벡터는 레트로바이러스, 렌티바이러스, 아데노바이러스, 아데노-연관 바이러스(AAV), 백시니아바이러스, 폭스바이러스 및 단순포진 바이러스로 구성된 군에서 선택되는 1이상일 수 있다.

[974]

[975] **본 명세서에 의해 개시되는 일 태양은 유전자 중복 질환을 치료하기 위한 발현 조절용 조성물을 이용한 유전자 중복 질환 치료방법에 관한 것이다.**

[976] 본 명세서에 의해 개시되는 일 구체예는 치료 대상에 중복 유전자의 전사 조절 영역을 인위적으로 조작하기 위한 발현 조절용 조성물의 투여를 포함하는 방법을 이용하여 유전자 중복 질환을 치료하기 위한 용도이다.

[977] 이때, 상기 치료 대상은 인간, 원숭이 등의 영장류, 마우스, 래트 등의 설치류 등을 포함하는 포유 동물일 수 있다.

[978] 상기 유전자 중복 질환 관련 설명을 상기 기술한 바와 같다.

[979] 일 실시예에서, 유전자 중복 질환은 PMP22 유전자의 중복에 의해 발생하는 질환일 수 있다.

[980] 일 예로, 상기 PMP22 유전자의 중복에 의해 발생하는 질환은 샤르코-마리-투스 질환 형태 1A(Charcot-Marie-Tooth disease type 1A, CMT1A), 데제린-소타스병(Dejerine-Sottas disease, DSS), 선천성 저수초형성 신경병(Congenital Hypomyelination Neuropathy, CHN) 또는 루시-레비 증후군(Roussy-Levy syndrome, RLS)일 수 있다.

[981] - 샤르코-마리-투스 질환(Charcot-Marie-Tooth disease, CMT)

[982] CMT 질환은 인간의 염색체에서 일어난 유전자 중복으로 인해 생기는 유전성 질환으로, 손과 발의 말초신경 발달에 관여하는 유전자가 돌연변이로 인해 중복되어 샴페인 병을 거꾸로 세운 것과 같은 모습의 기형을 유발한다. CMT 질환은 미국에서는 10만명 중 36명 꼴로 발생하는 비교적 흔한 신경계 유전질환으로, 환자수는 전 세계 280만명이며, 국내에서도 1만 7000명 정도로 추산되고 있다. CMT 질환은 유전되는 양상에 따라 크게 CMT1, CMT2, CMT3, CMT4 및 CMTX 총 5가지로 분류하며, CMT1, CMT2, CMT3는 우성으로 자녀에게 50%의 확률로 유전되고, CMT4는 열성으로 25%의 확률로 유전된다. 국내 환자의 대부분은 CMT1(80%)과 CMT2(20-40%)으로 우성유전이고, CMT3와 CMT4는 극히 드물다. CMTX는 X 염색체를 따라 모계로 유전되는데 빈도는 10-20%이다.

[983] CMT1은 신경 축삭돌기를 둘러싸는 신경 수초(myelin sheath)의 단백질 형성에 관여하는 유전자 중복으로 인해 정상적인 유전자 발현 과정을 수행하지 못하기 때문에 일어나는 질환으로, 3가지 타입이 있다. CMT1A는 상염색체 우성질환으로 17번 염색체 17p11.2-p12에 위치한 PMP22 유전자의 중복에 의한 것으로, 수초의 중요한 구성요소인 PMP22의 과발현으로 인한 수초의 구조와 기능 이상이 초래된다.

[984] CMT2는 축삭돌기 이상과 관련이 있고, 신경전도 속도는 정상에 가까우면서 운동감각 신경의 활동전위가 많이 저하된 신경병이며, CMT3은 극히 드물게 발생하는 상염색체 열성유전병으로 어린 나이에 발병하며 임상증상 및 신경전도 속도의 저하가 매우 심한 유형이다. CMT4도 상염색체 열성유전으로 발병 연령이 빠르고 임상 증상이 심한 유형이고, CMTX는 X 염색체와 연관되어 발병하며 여자보다 남자의 증상이 심하다.

[985] - 데제린-소타스병(Dejerine-Sottas disease, DSD)

[986] DSS는 어린 연령에서 발병하는 탈수초성 운동감각신경병증으로 주로 상염색체 우성으로 유전하지만 상염색체 열성으로도 유전되는 질환으로, 심한 탈수초성 신경병증을 보이고, 영아기 때부터 운동신경의 이상을 보이며, 매우 느린 신경전도와 뇌척수액에서의 특정 단백질의 증가를 보이는 특징을 가진다. 데제린-소타스병은 진행속도가 매우 빠르고 보행장애는 어린 나이에서부터 시작되며, 유전되기도 하지만, 산발적으로 발생하기도 하는 특징을 가진다. CMT1A와 유사하게 일부 DSS 환자 중 PMP22 중복이 발견되며, 더불어 해당 유전자의 미스센스 돌연변이가 존재함이 확인되었다.

[987] - 선천성 저수초형성 신경병(Congenital Hypomyelination Neuropathy, CHN)

[988] CHN은 출생 즉시 증상이 나타나는 신경계 질환으로, 주요 증상으로 호흡기계 부전, 근육 약화, 근육 운동 부조화증, 근육 긴장성 저하, 무반사(Areflexia), 조화운동불능증(운동실조증; Ataxia), 마비나 감각 장애가 나타나며 남녀 동일한 비율로 영향을 미친다. CHN은 유전질환으로, 운동신경 및 감각신경에서 수초형성에 장애가 발생한 것으로 수초의 손실(demyelination)과

재생(remyelination)이 반복되면서 수초형성(Myelination)이 저하되는 특징을 보인다.

[989] - 루시-레비 증후군(Roussy-Levy syndrome, RLS)

[990] RLS는 드문 유전운동감각신경병의 일종으로 1926년에 Roussy와 Levy 등에 의해 처음으로 기술되었으며, 수족의 떨림 및 보행 실족 등이 다른 유전 운동감각 신경병증에 비해 심한 경우이지만 후에 여러 유전 운동신경병증 아형들에서도 같은 증상들이 발견되어서 현재는 유전 감각신경병증에서 나타나는 하나의 증상으로 여겨지고 있다. RLS는 RLS로 처음 보고된 환자들의 유전자 검사에서 마이엘린 단백질 제로 유전자인 MPZ 유전자의 돌연변이가 발견되었고, 또 다른 환자들에서 말초 신경의 마이엘린 단백질 22의 유전자인 PMP22 유전자 중복이 있는 사례가 보고된 바 있다.

[991]

[992] 일 실시예에서, 유전자 중복 질환은 PLP1 유전자의 중복에 의해 발생하는 질환일 수 있다.

[993] 일 예로, 상기 PLP1 유전자의 중복에 의해 발생하는 질환은 펠리제우스-메르츠바하병(Pelizaeus-Merzbacher disease, PMD)일 수 있다.

[994] - 펠리제우스-메르츠바하병(Pelizaeus-Merzbacher disease, PMD)

[995] 펠리제우스-메르츠바하병(Pelizaeus-Merzbacher disease(PMD))은 중추신경계 백질의 수초 형성 부전으로 인하여 여러 신경학적 증상을 나타내는 매우 희귀한 수단친화성 백질이영양증(sudanophilic leukodystrophy)이다. 그 유병률은 1/400 000 정도로 추정되고 있다. 1885년 Pelizaeus는 X 염색체 의존으로 유전되며 진행성 대뇌 양측마비(cerebral diplegia)를 가지는 한 가족을 보고하였고 이 예들은 병의 초기에 나타나는 안진, 운동실조, 강직, 후천적 소두증을 특징으로 하였다. PMD의 임상증상은 영아기나 소아기의 초기에 나타나며, 특징적 임상 증상으로는 진자양 안진, 천명, 정신운동발달장애 또는 퇴행, 운동실조, 무정위운동, 불수의운동, 구음장애, 정신지체 등이 있다.

[996] Pelizaeus-Merzbacher병은 희소돌기아교세포(oligodendrocyte)의 감소와 proteolipid protein (PLP)의 합성장애에 의한 중추신경계 백질의 수초형성부전(dysmyelination)에 기인하는 신경변성(neurodegenerative) 질환 또는 백질이영양증(leukodystrophy)이다. Proteolipid protein (PLP)은 중추 신경계의 수초에 가장 풍부하게 존재하는 단백질로 X 염색체 장완에 위치한(Xq22) PLP1 유전자의 돌연변이로 인해 PLP가 비정상적으로 발현되거나 생성되어 중추 신경계에 수초 이형성을 일으키게 된다. PMD는 뇌조직 병리검사상 수단친화성을 가지며 이런 수단친화성은 지질과 반응하는 몇몇 아조화합물에 의해 나타나는데 이것은 수초의 파괴(breakdown)를 의미하며 대뇌의 난형중심(centrum semiovale), 소뇌, 뇌간에서 관찰된다. 하지만 수초 파괴물질(breakdown products)이 발견되지 않는 점으로 봐서는 병인이 탈수초화(demylination)라기보다 수초형성부전(dysmyelination 또는

hypomyelination)이라고 생각되고 있다. 대체로 선천형(connate form)은 전반적인 수초형성부전, 전형적인 형(classic form)은 부분수초형성부전이 나타나며 부분수초형성부전이 있을 때에는 정상 수초화 백질이 호반성 모양(tigroid appearance) 을 띠게 된다. 수초형성부전이 있는 병변의 축삭(axon)과 신경원(neuron)은 일반적으로 잘 보존되어 있고 희소돌기아교세포의 숫자는 감소되어 있으며 성상세포(astrocyte)의 증가와 섬유성 신경교증(fibrous gliosis)이 백질 내에서 발견되기도 하고 소회뇌증(micropolygyria)과 소뇌의 과립층에서 위축이 발견되기도 한다. 80% 이상의 남자환자들에서 X 염색체 장완에 위치한(Xq22) PLP1 유전자의 돌연변이가 발견된다. 이 중 10-30%는 유전자 내의 점돌연변이를 가지며, 이 경우 더 심한 임상증상을 나타낸다고 알려져 있다. PLP1 유전자 전체가 중복(duplication)되는 현상이 PMD 환자의 60~70% 이상으로 더 흔하게 발견된다. 최근에는, PLP1 유전자가 X 염색체에 위치하므로 대부분의 PMD는 X 염색체 의존적으로 유전되어 가족력을 가지고 대부분 남자에서 발생하는 특징이 있다. 하지만 PMD의 발병기전을 PLP1 유전자만으로 다 설명할 수는 없으며 때로 선천형에서 상염색체 열성으로 혹은 성인형에서 상염색체 우성으로 또는 가족력 없이 산발적으로 발생하는 경우가 있고 드물게 여자에서도 PMD의 증상이 발현되는 경우도 보고되고 있다.

[997]

[998] 일 실시예에서, 유전자 중복 질환은 MECP2 유전자의 중복에 의해 발생하는 질환일 수 있다.

[999] 일 예로, 상기 MECP2 유전자 중복에 의해 발생하는 질환은 MECP2 중복 증후군(MECP2 duplication syndrome)일 수 있다.

[1000] - MECP2 중복 증후군(MECP2 duplication syndrome)

[1001] MECP2 중복 증후군 (MECP2 duplication syndrome)이라 불리는 뇌 질환은 MECP2 유전자를 포함하고 있는 X 염색체의 특정 부위에 일어나는 유전 물질의 중복 (duplication of genetic material)이 원인이다. 이 질환은 다양한 증상들을 동반하며, 근긴장 저하 (low muscle tone), 발달지연 (developmental delays), 반복되는 호흡기 감염 (respiratory infection), 언어장애 (speech abnormalities), 발작 (seizures), 자폐행동 (autistic behaviors) 및 심각한 지적 장애 (intellectual disability) 등의 증상들이 포함된다.

[1002] 이 질환은 유전질환이지만, 가족력 없이 발생하기도 한다. MECP2 중복 증후군은 주로 남성에게서 나타나며, MECP2 유전자 결함으로 생기는 레트 증후군 (Rett Syndrome)은 주로 여성에게서 나타난다.

[1003]

[1004] 일 실시예에서, 유전자 중복 질환은 RAI1 유전자의 중복에 의해 발생하는 질환일 수 있다.

[1005] 일 예로, 상기 RAI1 유전자 중복에 의해 발생하는 질환은 Potocki-Lupski 증후군(Potocki-Lupski syndrome, PTLS)일 수 있다.

[1006] - Potocki-Lupski 증후군(Potocki-Lupski syndrome, PTLS)

[1007] Potocki-Lupski 증후군(Potocki-Lupski syndrome, PTLS)은 17번 염색체의 단완의 11.2 부위(17p11.2)의 미세중복(microduplication)을 포함하는 연속 유전자 증후군(contiguous gene syndrome)으로, 1996년에 첫 연구 사례가 보고되었다. PTLS는 RAI1(retinoic acid induced-1) 유전자를 포함하는 17p11.2에서 1.3 - 3.7 Mb의 중복에 의해 유발되는 것으로 알려져 있다. PTLS는 희귀 질환으로 간주되며, 발생 빈도는 출생아 2만명 중에 한 명 꼴인 것으로 예측된다. PTLS는 여러 선청성 이상과 정신 지체가 나타나는 것이 특징으로, 80%의 경우 자폐 스펙트럼 장애가 나타난다. 또한, PTLS의 다른 독특한 특징은 수면 무호흡증, 구조적 심혈관 이상, 비정상적인 사회적 행동, 학습 장애, 주의력 결핍 장애, 강박 관념의 행동, 짧은 키 등이 있다.

[1008]

[1009] 일 실시예에서, 유전자 중복 질환은 ELN 유전자의 중복에 의해 발생하는 질환일 수 있다.

[1010] 일 예로, 상기 ELN 유전자 중복에 의해 발생하는 질환은 윌리엄스 보이렌 증후군(Williams Beuren syndrome, WBS)일 수 있다.

[1011] - 윌리엄스 보이렌 증후군(Williams Beuren syndrome, WBS)

[1012] 윌리엄스-보이렌증후군은 특징적인 임상소견을 가진 7번 염색체이상과 관련된 근접 유전자증후군으로서, 출생아 2만 명당에 1명꼴의 발생빈도를 가진다. 7번 염색체의 장완 근위부(7q11.23)의 미세결실이 원인으로서, 이 부위에는 혈관벽 같은 탄성조직을 이루는 엘라스틴 단백질의 생성과 관계되는 엘라스틴 유전자와 인지능력과 관계되는 LIMK1 유전자를 비롯하여 여러 유전자들이 위치하고 있다. 이러한 유전자들의 결실로 인해 다양하고 특징적인 외형과 임상증상이 나타납니다. 7q11.23의 미세결실은 대부분 자연발생적으로 일어나며, 드물게 가족력이 있는 경우가 있다. WBS를 가진 아이들은 위로 솟은 작은 코끝, 긴 인중, 큰 입, 두툼한 입술, 작은 볼(협골형성부전), 부은 듯한 눈두덩이, 손톱의 형성부전, 엄지발가락의 외반증과 같은 외형적 특징을 가지고 있다.

[1013]

[1014] 일 실시예에서, 유전자 중복 질환은 JAGGED1 유전자의 중복에 의해 발생하는 질환일 수 있다.

[1015] 일 예로, 상기 JAGGED1 유전자 중복에 의해 발생하는 질환은 알라질증후군(Alagille syndrome, AS)일 수 있다.

[1016] - 알라질증후군(Alagille syndrome, AS)

[1017] 알라질 증후군(Alagille's syndrome)은 간내 담도의 수가 현저하게 감소하고 담즙정체를 나타내며 심혈관계, 골격계, 안구, 안면, 췌장, 신경발달의 장애를 동반할 수 있는 증후군이다. 이 질환의 유병률은 외국의 보고에 의하면 100,000명당 1명의 빈도이지만 질환의 특성상 증상이 경한 환자들까지

포함된다면 그 빈도는 더 클 것으로 예상이다. 알라질 증후군은 염색체 20번 단완에 위치하는 JAGGED1 유전자의 이상으로 인하여 발생합니다. 현재 유전학적 검사로 50-70% 정도에서 원인 돌연변이 또는 중복을 발견할 수 있다고 알려져 있다.

[1018] 알라질 증후의 임상 증상은 보통 생후 3개월 이내에 발현한다. 대부분 신생아 시기에 지속되는 황달과 담즙정체증으로 발견되고 소아기에서는 만성간질환으로 발견될 수 있으며 성인기에 뒤늦게 발견되기도 한다. 임상증상은 다양하게 나타나고 또한 불완전하게 유전될 수 있기 때문에 진단이 어려울 수 있습니다. 대부분의 환자들은 영아기에 황달과 담즙정체, 이로 인한 가려움증과 진행하는 간부전의 형태로 증상이 나타난다. 황달은 대부분의 환자에게서 관찰되며 이중 절반 이상의 환자에게서 늦은 소아기까지 지속되게 된다. 담즙정체로 인한 가려움증이 발생하며 일부 환아들은 피하조직에 황색종이 나타나기도 한다. 간의 합성기능은 비교적 잘 보존이 되는 편이지만 약 20%의 환자에서 간경화와 간부전으로 진행할 수 있다.

[1019]

[1020] 일 실시예에서, 유전자 중복 질환은 SNCA 유전자의 중복에 의해 발생하는 질환일 수 있다.

[1021] 일 예로, 상기 SNCA 유전자 중복에 의해 발생하는 질환은 파킨슨병(Parkinson disease)일 수 있다.

[1022] - 파킨슨병(Parkinson disease)

[1023] 파킨슨병 또는 파킨슨씨병(Parkinson's disease)이란 주로 진전(震顫, 떨림), 근육의 강직(剛直) 그리고 몸동작이 느려지는 서동(徐動) 등의 운동장애가 나타나는 질환이다. 파킨슨병은 적절한 치료를 받지 않으면, 운동장애가 점점 진행하여 걸음을 걷기가 어렵게 되고 일상생활을 전혀 수행할 수 없게 되기도 한다. 파킨슨병은 주로 노년층에서 발생하는 질환으로 연령이 증가할수록 이 병에 걸릴 위험은 점점 커지게 된다. 국내의 정확한 통계 자료는 없으나 인구 1,000 명 당 1-2명 정도의 비율로 발병하는 것으로 추정된다. 나이가 많이 들어서 발병하는 대부분의 파킨슨병의 경우는 여러 연구를 통해 유전적인 요소가 거의 영향이 없다고 알려져 있다. 하지만 40세 미만의 젊은 나이에 발생하는 파킨슨병의 경우에는 일부에서 유전적인 요소가 관련이 되는 것으로 알려져 있다.

[1024] 파킨슨병은 중뇌의 흑질(substantia nigra)에 존재하는 도파민 신경세포가 점차 사멸됨에 따라 도파민의 농도가 낮아져서 일어나는 질병이다. 파킨슨병의 또 다른 병리 특징은 뇌 부검에서 볼 수 있는 레비소체(Lewy body)라는 단백질 응집체의 생성이다. 레비소체는 알파시누클레인(α-synuclein)이라는 단백질이 주요 구성성분이며, 레비소체와 알파시누클레인은 레비소체치매, 시누클레인병(synucleinopathy)과 같은 다른 질병과도 연관된다. 알파시누클레인 응집체 형성은 파킨슨병 초기에 중뇌가 아닌 미주신경, 전후각신경핵(anterior

olfactory nucleus)에서 시작하여 병이 진행함에 따라 중뇌를 거쳐 말기에는 대뇌피질로 퍼져 나간다. 알파시누클레인이 병의 진행 정도에 따라 뇌의 여러 부위로 광범위하게 퍼져 나간다는 가설은 알파시누클레인이 한 세포에서 방출되어 다른 세포로 전달될 수 있다는 최근 보고들에 의해 지지를 받고 있다.

[1025] 파킨슨병의 유전 가능성은 레비소체의 주요 구성성분인 알파시누클레인의 돌연변이(A53T, A30P)가 파킨슨병을 일으킨다는 보고에 의해 처음으로 제기되었다. 그 이래로 알파시누클레인 유전자(SNCA)의 중복(duplication), 삼중중복(triplication)이 파킨슨병의 또 다른 원인으로 보고되었다. 이는 알파시누클레인 단백질의 변이뿐 아니라 정상 단백질의 과발현도 세포 내에 알파시누클레인을 축적하여 응집체를 형성함으로써 파킨슨병을 유발함을 의미한다.

[1026]

[1027] 일 실시예에서, 유전자 중복 질환은 APP 유전자의 중복에 의해 발생하는 질환일 수 있다.

[1028] 일 예로, 상기 APP 유전자 중복에 의해 발생하는 질환은 알츠하이머(Alzheimer disease)일 수 있다.

[1029] - 알츠하이머(Alzheimer disease)

[1030] 알츠하이머는 기억력의 점진적인 퇴행을 가져오는 뇌의 이상에서 오는 병이다. 또한 알츠하이머는 일상생활에 곤란을 겪을 정도의 심각한 (사고, 기억, 추론) 지적기능의 상실을 가져오는 치매(dementia)중에 이르게 된다. 대부분의 경우 알츠하이머병은 65세가 넘어 발병하지만, 드물게 그 이전에 발병할 수 있다. 미국에서는 65~74세 인구의 약 3%, 75~84세 인구의 약 19%, 85세 이상 인구의 50%가 이 병을 앓고 있다. 한국에서는 최근 한 농촌 지역을 중심으로 한 연구 보고에 의하면, 농촌 지역 60세 이상의 인구에서 약 21%가 치매 양상을 보이고, 이 중 63%가 알츠하이머형 치매인 것으로 보고되고 있다. 2006년 전 세계 26.6만명이 가진 질병이다. 2050년에는 85명중 1명꼴로 발병될 것으로 예측된다.

[1031] 질병의 특성은 개개인마다 다르게 나타나지만 일부 공통적인 증상을 가진다. 초기증상은 노화에 의한 단순한 증상이나 스트레스에 의한 증상으로 오해되는 경향이 있다. 발병 초기에는 이름·날짜·장소와 같은 것들이 기억에서 사라지는 공통된 단기 기억 상실을 겪는다. 질병이 악화되면, 혼란, 격한 행동, 조울증, 언어장애, 장기기억상실 등의 증상이 나타난다. 결과적으로 신체기능이 상실되고, 치사에 이른다. 개개인마다 증상이 다르기 때문에 질병이 어떻게 영향을 미칠지 예측하는 것은 어렵다. 알츠하이머병이 의심되면 보통 생각하거나 행동하는 능력을 검사하는 진단이 행해지고 가능한 경우 뇌 검사가 이루어진다. 하지만 정확한 진단을 위해서는 뇌신경에 대한 조사가 필요하다. 알츠하이머병이 발병해도 완전히 확인될 때까지는 보통 많은 시간이 소요되어, 진단이 내려지지 않은 채로 수년간 병이 진행될 수 있다. 발병할 경우 평균 기대 수명은 7년이며 3%미만이 진단 이후 14년을 넘긴다.

[1032] 알츠하이머병은 신경변성 질환으로 분류된다. 질병의 발병 원인은 완전히 이해되지 않은 상태이며 아밀로이드 플라크가 정상 알츠하이머 단백질을 변형시켜 플라크 덩어리를 형성해 고유기능을 잃어버리게 하는 것으로 추정된다. 조직병리학적으로는 뇌의 전반적인 위축, 뇌실의 확장, 신경섬유의 다발성 병변과 초로성 반점 등의 특징을 보인다.

[1033]

[1034] 일 실시예에서, 유전자 중복 질환은 SOX3 유전자, TBX1 유전자, NSD1 유전자, MMP23 유전자 또는 LMB1 유전자의 중복에 의해 발생하는 질환일 수 있다.

[1035] 일 예로, 상기 유전자 중복 질환은 X-연관 뇌하수체기능저하증(X-linked hypopituitarism, XLHP), 구개심장안면증후군(Velocardiofacial syndrome, VCFS), 발육지연 증후군(Growth retardation syndrome), 조기 폐쇄 두개봉합(Premature closure cranial sutures) 또는 백질이영양증(Autosomal dominant leukodystrophy, ADLD) 일 수 있다.

[1036]

[1037] 일 실시예에서, 유전자 중복 질환은 암유전자(oncogene)의 중복에 의해 발생하는 암일 수 있다.

[1038] 이때, 상기 암유전자는 MYC 유전자, ERBB2(HER2) 유전자, CCND1(Cyclin D1) 유전자, FGFR1 유전자, FGFR2 유전자, HRAS 유전자, KRAS 유전자, MYB 유전자, MDM2 유전자, CCNE(Cyclin E) 유전자, MET 유전자, CDK4 유전자, ERBB1 유전자, MYCN 유전자 또는 AKT2 유전자일 수 있다.

[1039] 일 예로, 상기 암은 유방암(Breast cancer), 자궁경부암(Cervical cancer), 대장암(Colorectal cancer), 식도암(Esophageal cancer), 위암(Gastric cancer), 교모세포종(Glioblastoma), 두경부암(Head and neck cancer), 간암(Hepatocellular cancer), 신경아세포종(Neuroblastoma), 난소암(Ovarian cancer), 육종(Sarcoma) 또는 소세포폐암(Small cell lung cancer)일 수 있다.

[1040]

[1041] 본 명세서에 의해 개시되는 내용의 일 구체예는 중복 유전자의 전사 조절 영역을 인위적으로 조작할 수 있는 발현 조절용 조성물을 포함하는 약학적 조성물이다.

[1042] 상기 발현 조절용 조성물 관련 설명은 상기 기술한 바와 같다.

[1043] 일 구현예로, 상기 발현 조절용 조성물은 다음을 포함할 수 있다:

[1044] (a) 중복 유전자의 전사 조절 영역 내 위치한 표적 서열을 표적화할 수 있는 가이드핵산 또는 이를 암호화하는 핵산 서열; 및

[1045] (b) 스트렙토코커스 피요게네스(Streptococcus pyogenes) 유래의 Cas9 단백질, 캄필로박터 제주니 (Campylobacter jejuni) 유래의 Cas9 단백질, 스트렙토코커스 써모필러스 (Streptococcus thermophilus) 유래의 Cas9 단백질, 스타필로코커스 아우레우스 (Staphylococcus aureus) 유래의 Cas9 단백질, 네이세리아 메닝기디티스 (Neisseria meningitidis)유래의 Cas9 단백질, 및 Cpf1 단백질로

이루어진 군에서 선택된 하나 이상의 단백질을 포함하는 에디터단백질 또는 이를 암호화하는 핵산서열.

[1046] 이때, 상기 중복 유전자는 PMP22 유전자, PLP1 유전자, MECP2 유전자, SOX3 유전자, RAI1 유전자, TBX1 유전자, ELN 유전자, JAGGED1 유전자, NSD1 유전자, MMP23 유전자, LMB1 유전자, SNCA 유전자 및 APP 유전자로 구성된 군에서 선택된 하나 이상의 유전자일 있다.

[1047] 또는, 상기 중복 유전자는 MYC 유전자, ERBB2(HER2) 유전자, CCND1(Cyclin D1) 유전자, FGFR1 유전자, FGFR2 유전자, HRAS 유전자, KRAS 유전자, MYB 유전자, MDM2 유전자, CCNE(Cyclin E) 유전자, MET 유전자, CDK4 유전자, ERBB1 유전자, MYCN 유전자 및 AKT2 유전자로 구성된 군에서 선택된 하나 이상의 유전자일 수 있다.

[1048] 상기 전사 조절 영역 관련 설명을 상기 기술한 바와 같다.

[1049] 상기 가이드 핵산 및 에디터 단백질은, 각각 핵산 서열의 형태로 1 이상의 벡터에 존재하거나, 또는 가이드 핵산과 에디터 단백질의 결합으로 복합체를 형성하여 존재할 수 있다.

[1050] 상기 발현 조절용 조성물은 삽입을 원하는 핵산서열을 포함하는 도너 또는 이를 암호화하는 핵산서열을 선택적으로 더 포함할 수 있다.

[1051] 상기 가이드 핵산 및 에디터 단백질, 및/또는 도너는, 각각 핵산 서열의 형태로 1 이상의 벡터에 존재할 수 있다.

[1052] 상기 약학적 조성물은 부가적 요소를 추가로 더 포함할 수 있다.

[1053] 상기 부가적 요소는 대상의 체내에 전달하기 위한 적절한 담체를 포함할 수 있다.

[1054]

[1055] 본 명세서에 의해 개시되는 내용의 일 구체예는 유전자 중복 질환을 치료하기 위해 유전자 중복 질환을 가지는 유기체에 유전자 조작용 조성물을 투여하는 단계를 포함하는, 유전자 중복 질환의 치료방법이다.

[1056] 상기 치료방법은 생체 내 존재하는 중복 유전자의 전사 조절 영역을 조작하여 중복 유전자의 발현을 조절하는 치료방법일 수 있다. 이러한 치료방법은 생체 내 존재하는 중복 유전자의 전사 조절 영역을 조작하기 위한 발현 조절용 조성물을 체내에 직접 주입하여 이루어질 수 있다.

[1057] 상기 발현 조절용 조성물 관련 설명은 상기 기술한 바와 같다.

[1058] 일 구현예로, 상기 발현 조절용 조성물은 다음을 포함할 수 있다:

[1059] (a) 중복 유전자의 전사 조절 영역 내 위치한 표적 서열을 표적화할 수 있는 가이드핵산 또는 이를 암호화하는 핵산 서열; 및

[1060] (b) 스트렙토코커스 피요게네스(Streptococcus pyogenes) 유래의 Cas9 단백질, 캄필로박터 제주니 (Campylobacter jejuni) 유래의 Cas9 단백질, 스트렙토코커스 써모필러스 (Streptococcus thermophilus) 유래의 Cas9 단백질, 스타필로코커스 아우레우스 (Staphylococcus aureus) 유래의 Cas9 단백질, 네이세리아

메닝기디티스 (Neisseria meningitidis)유래의 Cas9 단백질, 및 Cpf1 단백질로 이루어진 군에서 선택된 하나 이상의 단백질을 포함하는 에디터단백질 또는 이를 암호화하는 핵산서열.

[1061] 이때, 상기 중복 유전자는 PMP22 유전자, PLP1 유전자, MECP2 유전자, SOX3 유전자, RAI1 유전자, TBX1 유전자, ELN 유전자, JAGGED1 유전자, NSD1 유전자, MMP23 유전자, LMB1 유전자, SNCA 유전자 및 APP 유전자로 구성된 군에서 선택된 하나 이상의 유전자일 있다.

[1062] 또는, 상기 중복 유전자는 MYC 유전자, ERBB2(HER2) 유전자, CCND1(Cyclin D1) 유전자, FGFR1 유전자, FGFR2 유전자, HRAS 유전자, KRAS 유전자, MYB 유전자, MDM2 유전자, CCNE(Cyclin E) 유전자, MET 유전자, CDK4 유전자, ERBB1 유전자, MYCN 유전자 및 AKT2 유전자로 구성된 군에서 선택된 하나 이상의 유전자일 수 있다.

[1063] 상기 전사 조절 영역 관련 설명을 상기 기술한 바와 같다.

[1064] 상기 가이드 핵산 및 에디터 단백질은, 각각 핵산 서열의 형태로 1 이상의 벡터에 존재하거나, 또는 가이드 핵산과 에디터 단백질의 결합으로 복합체를 형성하여 존재할 수 있다.

[1065] 상기 발현 조절용 조성물은 삽입을 원하는 핵산서열을 포함하는 도너 또는 이를 암호화하는 핵산서열을 선택적으로 더 포함할 수 있다.

[1066] 상기 가이드 핵산 및 에디터 단백질, 및/또는 도너는, 각각 핵산 서열의 형태로 1 이상의 벡터에 존재할 수 있다.

[1067] 이때, 상기 벡터는 플라스미드 또는 바이러스벡터일 수 잇다.

[1068] 이때, 상기 상기 바이러스 벡터는 레트로바이러스, 렌티바이러스, 아데노바이러스, 아데노-연관 바이러스(AAV), 백시니아바이러스, 폭스바이러스 및 단순포진 바이러스로 구성된 군에서 선택되는 1이상일 수 있다.

[1069] 상기 유전자 중복 질환 관련 설명은 상기 기술한 바와 같다.

[1070] 유전자 중복 질환은 샤르코마리투스병 1A형(Charcot-Marie-Tooth 1A, CMT1A), 데제린-소타스병(Dejerine-Sottas disease, DSD), 선천성 저수초형성 신경병(Congenital Hypomyelination Neuropathy, CHN), 루시-레비 증후군(Roussy-Levy syndrome, RLS), 펠리제우스-메르츠바하병(Pelizaeus-Merzbacher disease, PMD), MECP2 중복 증후군(MECP2 duplication syndrome), X-연관 뇌하수체기능저하증(X-linked hypopituitarism, XLHP), Potocki-Lupski 증후군(Potocki-Lupski syndrome, PTLS), 구개심장안면증후군(Velocardiofacial syndrome, VCFS), 윌리엄스 보이렌 증후군(Williams Beuren syndrome, WBS), 알라질증후군(Alagille syndrome, AS), 발육지연 증후군(Growth retardation syndrome), 조기 폐쇄 두개봉합(Premature closure cranial sutures), 보통염색체우성 백질이영양증(Autosomal dominant leukodystrophy, ADLD), 파킨슨병(Parkinson disease) 또는 알츠하이머(Alzheimer disease)일 수 있다.

[1071] 또는, 상기 암은 유방암(Breast cancer), 자궁경부암(Cervical cancer), 대장암(Colorectal cancer), 식도암(Esophageal cancer), 위암(Gastric cancer), 교모세포종(Glioblastoma), 두경부암(Head and neck cancer), 간암(Hepatocellular cancer), 신경아세포종(Neuroblastoma), 난소암(Ovarian cancer), 육종(Sarcoma) 또는 소세포폐암(Small cell lung cancer)일 수 있다.

[1072] 상기 발현 조절용 조성물은 유전자 중복 질환을 가지는 치료 대사에 투여될 수 있다.

[1073] 상기 치료 대상은 인간, 원숭이 등의 영장류, 마우스, 래트 등의 설치류 등을 포함하는 포유동물일 수 있다.

[1074] 상기 발현 조절용 조성물은 치료 대상에 투여될 수 있다.

[1075] 상기 투여는 주사(injection), 수혈(transfusion), 삽입(implantation) 또는 이식(transplantation)으로 수행될 수 있다.

[1076] 상기 투여는 신경내(intraneural), 망막하(subretinal), 피하(subcutaneously), 피내(intradermaliy), 안구내(intraocularly), 유리체내(intravitreally) 종양내(intratumorally), 절내(intranodally), 골수내(intramedullary), 근육내(intramuscularly), 정맥내(intravenous), 림프액내(intralymphatic) 또는 복막내(intraperitoneally)에서 선택된 투여 경로로 수행될 수 있다.

[1077] 발현 조절용 조성물의 1회 투여량(소정의 소망하는 효과를 얻기 위한 약학적 유효량)은 투여 대상의 체중 kg 당 10^4-10^9 세포, 예컨대, 10^5 내지 10^6 세포/kg(체중) 정도로 상기 수치 범위들 내의 모든 정수값들 중에서 선택될 수 있으나, 이에 제한되는 것은 아니고, 투여 대상의 연령, 건강 및 체중, 동시에 받는 치료의 종류, 만약 있다면 치료의 빈도, 원하는 효과의 특성 등을 고려하여 적절히 처방될 수 있다.

[1078]

[1079] 본 명세서에 의해 개시되는 몇몇의 구현예들의 방법 및 조성물에 의해 중복 유전자의 전사 조절 영역을 인위적으로 조작할 경우, 중복 유전자의 mRNA 및/또는 단백질의 발현이 조절되고, 이를 통해 비정상적으로 발현되던 중복 유전자의 발현을 정상적으로 조절시키는 등의 효과를 얻을 수 있다.

[1080]

발명의 실시를 위한 형태

[1081] 이하, 실시예를 통하여 본 발명을 더욱 상세히 설명하고자 한다.

[1082] 이들 실시예는 오로지 본 발명을 보다 구체적으로 설명하기 위한 것으로, 본 발명의 범위가 이들의 실시예에 의해 제한되지 않는다는 것은 본 발명이 속하는 기술분야에서 통상의 지식을 가진 자에 있어 자명할 것이다.

[1083]

[1084] **실험방법**

[1085] **1. gRNA 설계**

[1086] CRISPR RGEN Tools (www.rgenome.net)을 사용하여 인간의 PMP22 유전자, 인간 PLP1 유전자 및마우스 PLP1 유전자의 CRISPR/Cas9 표적 부위를 선별하였다. PMP22 유전자 및 PLP1 유전자의 표적 부위는 CRISPR 효소의 종류에 따라 달라질 수 있다. SpCas9을 위한 인간 PMP22 유전자의 프로모터(TATA-box) 부위 및 인핸서 부위(예를 들어, EGR2-, SOX10- 또는 TEAD1-결합 부위; 또는 원거리 인핸서 부위(distal Enhancer region) B 또는 C의 표적 서열은 앞서 기재한 표 1에 정리하였고, CjCas9을 위한 인간 PMP22 유전자의 프로모터(TATA-box) 부위 및 인핸서 부위(예를 들어, EGR2- 또는 SOX10-결합 부위)의 표적 서열은 표 2에 정리하였다. 또한, SpCas9을 위한 인간 PLP1 유전자의 프로모터(TATA-box 부위) 및 인핸서 부위(예를 들어, wmN1 인핸서)의 표적 서열은 앞서 기재한 표 3에 정리하였고, CjCas9을 위한 인간 PLP1 유전자의 프로모터(TATA-box 부위) 및 인핸서 부위(예를 들어, wmN1 인핸서)의 표적 서열은 앞서 기재한 표 4에 정리하였다. SpCas9을 위한 마우스 PLP1 유전자의 프로모터(TATA-box 부위) 및 인핸서 부위(예를 들어, wmN1 인핸서)의 표적 서열은 앞서 기재한 표 5에 정리하였고, CjCas9을 위한 마우스 PLP1 유전자의 프로모터(TATA-box 부위) 및 인핸서 부위(예를 들어, wmN1 인핸서)의 표적 서열은 앞서 기재한 표 6에 정리하였다.

[1087] 모든 gRNA는 키메릭 단일가닥 RNA(chimeric single stranded RNA, sgRNA) 형태로 생성하였다. 표적 서열을 제외한 Cj- 및 Sp-특이적 sgRNA의 백본서열은 각

5'-GUUUUAGUCCCUGAAAAGGGACUAAAAUAAAGAGUUUGCGGGACUCUGCGGGGUUACAAUCCCCUAAAACCGCUUUU-3' 및

5'-GUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGC-3'이다.

[1088]

[1089] **2. gRNA의 구축 및 합성**

[1090] sgRNA는 AAV 벡터에 패키징하거나 또는 RNA로 합성하였다. 바이러스 벡터로 sgRNA를 삽입하기위해, sgRNA의 20 내지 22개의 염기서열에 상응하는 DNA 올리고뉴클레오타이드를 설계하여 어닐링하고, BsmBI 사이트를 이용해 pRGEN-Cas9(자체 개발한) 벡터내에 연결시켰다. Cas9과 5' 말단에 가변적인 표적 서열을 포함하는 sgRNA는 각각 CMV와 U6 프로모터를 통해 발현시켰다.

[1091] 또한, RNP에 의한 전달 시스템을 위해, sgRNA는 Phusion Taq 매개 중합에 의해 생성된 두 개의 부분적 상보성 올리고뉴클레오타이드의 어닐링에 의한 주형의 생성 후 T7 RNA 중합효소에 의해 전사시켰다. 전사된 sgRNA는 분광분석(spectrometry)을 이용해 정제 및 정량하였다.

[1092]

[1093] **3. Cas9 단백질 정제**

[1094] NLS 및 HA 에피토프를 포함하는 코돈-최적화된 Cas9 DNA 서열은 pET28

벡터로 서브 클로닝시키고, 최적의 배양 조건에서 IPTG를 이용해 BL21(DE3)에서 발현시켰다. 발현된 Cas9 단백질은 Ni-NTA 아가로스 비드를 이용해 정제하고, 적절한 버퍼로 투석시켰다. Cas9의 활성은 잘 알려진 효과적인 sgRNA를 이용한 in vitro 절단 시험을 통해 확인하였다.

[1095]

[1096] **4. 세포 배양**

[1097] 인간 슈반 유사 세포주(Human Schwann-like cell line)(ATCC) 및 인간 슈반세포(Human primary Schwann cell)(ScienCell)는 제조사의 매뉴얼에 따라 배양하였다. 인간 슈반 유사 세포는 1×페니실린/스트렙토마이신(WelGene) 및 10% 태아 송아지 혈청(WelGene)를 추가한 고농도 글루코스 포함 Dulbecco's modified Eagle's medium(DMEM)(WelGene)에서 배양하였다.

[1098] 인간 슈반세포는 판매사에서 제공하는 슈반세포 배양액(ScienCell)에서 유지시켰다. 분화를 위해, 세포를 1% 태아 송아지 혈청(WelGene)과 수초형성(myelination) 신호를 위한 100 ng/mL Nrg1(Peprotech) 및 100μM dbcAMP(Sigma-Aldrich)를 첨가한 저농도 글루코스 포함 DMEM(WleGene)에서 7일 동안 배양하였다.

[1099]

[1100] **5. 형질도입(Transfection)**

[1101] 형질도입(transfection)을 위해, 4㎍ Cas9 단백질(ToolGen)과 1㎍ sgRNA를 포함한 RNP 복합체를 실온에서 15분동안 인큐베이션하였다. 그 후, 2×10^5 세포에 상기 RNP 복합체를 10㎕ 전기천공용 팁과 Neon electroporator(ThermoFisher)를 이용해 전기천공하여 전달하였다. 표적 딥 시퀀싱을 위해, 형질도입 72시간 후에 형질도입 된 세포에서 게놈 DNA(gDNA)를 수집하였다.

[1102]

[1103] **6.** ***In vitro*** **Real time PCR(qRT-PCR)**

[1104] RNeasy mini kit(Qiagen)을 이용해 제조사의 프로토콜에 따라 인간 슈반세포로부터 mRNA를 추출하였다. 그 후, 100ng mRNA는 High-capacity cDNA reverse transcription kit(ThermoFisher)를 이용하여 역 전사시켰다. qRT-PCR은 QuantStudio 3(ThermoFisher)를 이용해 제조사의 프로토콜에 따라 Taqman Gene expression master mix 100ng으로 수행하였다. PMP22 발현 수준은 C_T 값을 이용해 계산하였고, GAPDH로 내인성 대조군(endogenous control)으로 사용하였다. 본 연구에서 사용한 Taqman 프로브(ThermoFisher)는 하단의 표 7에 정리하였다.

[1105] [표7]

Target Gene	Taqman Gene Experssion Assay	Accession number
PMP22	Hs00165556_m1	NM_000304.3
GAPDH	HS02786624_g1	NM_001256799.2

[1106]

[1107] **7. 표적 딥 시퀀싱(Targeted deep sequencing)**

[1108] 온-타겟 부위는 Phusion polymerase taq(New England BioLabs)을 이용하여 형질도입된 세포에서 추출한 gDNA로부터 PCR로 증폭하였다. 그 후, 상기 PCR 증폭물로 Mi-Seq(Illumina)을 이용하여 paired-end 딥 시퀀싱하였다. 딥 시퀀싱 결과는 온라인 Cas-Analyzer tool (www.rgenome.net)을 이용하여 분석하였다. PAM 서열로부터 3bp 업스트림에 인델이 Cas9에 의한 결과로 돌연변이가 발생하였는지 확인하였다. 본 연구에서 사용한 프라이머는 하단의 표 8에 정리하였다.

[1109] [표 8]

[1110]

	Target site	Primer-F (5' to 3')		Primer-R (5' to 3')	
On-Target	*hPMP22*-TATA	CACAGGGCAGTCAGAGACCC	SEQ ID NO: 185	GCAAACAAAGTTGGACACTG	SEQ ID NO: 186
	mRosa26	AGACTCCCGCCCATCTTCTAGAAA	SEQ ID NO: 187	AAGTCGCTCTGAGTTGTTATCAGT	SEQ ID NO: 188
	AAVS1	CAGTGAAACGCACCAGACG	SEQ ID NO: 189	AATCTGCCTAACAGGAGGTG	SEQ ID NO: 190
Off-target (*In silico, in vitro*)	*hPMP22*-TATA Off1	GAGGGAATGGGGACCAAAGGCATT	SEQ ID NO: 191	TCATGTGGGGTGATGTTCAGGAAG	SEQ ID NO: 192
	hPMP22-TATA Off2	AGAGCAGCTGACCTGAGGTCCAA	SEQ ID NO: 193	CCCAAGGGTAGAGTGCAAGTAAAC	SEQ ID NO: 194
	hPMP22-TATA Off3	GCATCCTAGCTCATTTGGTCTGCT	SEQ ID NO: 195	GAGAGGATTCCTCATGAATGGGAT	SEQ ID NO: 196
	hPMP22-TATA Off4	ACCAAACACTACACTTGGTTACTG	SEQ ID NO: 197	CTCCCACTAGCAATTTTAAAGTCT	SEQ ID NO: 198
	hPMP22-TATA Off5	GAATGTTCAGCACAGGTTTCCTTG	SEQ ID NO: 199	GGTCAAAAGGAGCTCCATATTTGA	SEQ ID NO: 200
	hPMP22-TATA Off6	CAGGACACCCATGGCCAAATCCAG	SEQ ID NO: 201	CAGAGCCTCCTGCAGGGATGTCAA	SEQ ID NO: 202
	hPMP22-TATA Off7	GCCTGCCAAGGTGACTCTCATCTA	SEQ ID NO: 203	TGCCCAGGCTGATCTTGAACTCCT	SEQ ID NO: 204
	hPMP22-TATA Off8	CCCAGAGTTAAGAGGTTCTTTCCT	SEQ ID NO: 205	GAAGCTACTCCAGTGCAACTAGCT	SEQ ID NO: 206
	hPMP22-TATA Off9	ACGCAGTCTGTTCTGTGCAGTGT	SEQ ID NO: 207	AGGCCTTCCCAAGGAAGACCCTGA	SEQ ID NO: 208
	hPMP22-TATA Off10	GCTGATCACTGGCCAAATCCAGCT	SEQ ID NO: 209	GGGAAACAATGGGATCAAGCTGCA	SEQ ID NO: 210
	hPMP22-TATA Off11	GCCCCTTTGTAAGTTGAGGAGCAT	SEQ ID NO: 211	CCCTCTACCTCTCTCAATGGGCTT	SEQ ID NO: 212
	hPMP22-TATA Off12	CAGACAAGCAAATGCTGAGAGATT	SEQ ID NO: 213	CCTGTCATTATGATGTTCGCTAGT	SEQ ID NO: 214
	hPMP22-TATA Off13	CCAGAGTTGGCCTCCTACAGAGAT	SEQ ID NO: 215	GTGGATGCCCCACTACTGTTCATT	SEQ ID NO: 216
	hPMP22-TATA Off14	TACCCAATTTGCCAGTCTGTGTCT	SEQ ID NO: 217	ACCACCAGGCCTGCCCTACAAGA	SEQ ID NO: 218
	hPMP22-TATA Off15	TGTGAATTTGATCCTGGCATTATG	SEQ ID NO: 219	TACAGACAAGCAGATGCTGAGAGA	SEQ ID NO: 220
	hPMP22-TATA Off16	CAGTCAACAGAGCTCTAACCTCCT	SEQ ID NO: 221	AGCACCTGGTTGCACATCAACTT	SEQ ID NO: 222
	hPMP22-TATA Off17	CATGTGGTCCCTGAACGTGAATGA	SEQ ID NO: 223	GTCTGTCGCTTGCCCTCTTCTCT	SEQ ID NO: 224
	hPMP22-TATA Off18	ATGCAGGGCCTCTAGACCATTTCA	SEQ ID NO: 225	CTCAGCCCTTTGTGCACTCACCT	SEQ ID NO: 226
Off-target (*Digenome-seq, in vitro*)	*hPMP22*-TATA Off1	TGCACATCGCAAACATTTCG	SEQ ID NO: 227	TGGGTATCGCACTGTGTCAG	SEQ ID NO: 228
	hPMP22-TATA Off2	AGGTTCACATGGCTTGTGGT	SEQ ID NO: 229	ATATCTGAAATGCCCGCAGG	SEQ ID NO: 230

[1111]

	hPMP22-TATA Off3	TGCACATCGCAAACATTTCG	SEQ ID NO: 231	TGGGTATCGCACTGTGTCAG	SEQ ID NO: 232
	hPMP22-TATA Off4	TCTTTAAAGGCCTTATCTCC	SEQ ID NO: 233	TTCTGCTTGAGAATTCATCC	SEQ ID NO: 234
	hPMP22-TATA Off5	CTCCTAATCTTTCACTTAGG	SEQ ID NO: 235	CAAAGCCTGGTATAACATAG	SEQ ID NO: 236
	hPMP22-TATA Off6	TCACTTCGAGCATCTGTGG	SEQ ID NO: 237	CCAAATGACAGGCTGAGCT	SEQ ID NO: 238
	hPMP22-TATA Off7	AGCAGGAAGTGAAGGCTAAG	SEQ ID NO: 239	ATGTAACGTGGCAACTCTGC	SEQ ID NO: 240
	hPMP22-TATA Off8	GTGTTGCTCTCGTCAATTAG	SEQ ID NO: 241	AGGTGTTGTACATGGAGAAG	SEQ ID NO: 242
	hPMP22-TATA Off9	TGTGAGCCACCATACCCAGC	SEQ ID NO: 243	CCTGCAGTCCTTTGCGGATC	SEQ ID NO: 244
Off-target (*In silico, In vivo*)	*hPMP22*-TATA Off1	TCGCTGCCAGTATAACATGC	SEQ ID NO: 245	AACTCCAGTCTCTAGACTCG	SEQ ID NO: 246
	hPMP22-TATA Off2	AATAGTTTGACGTTGGAGCC	SEQ ID NO: 247	ACTCCCAACATGTTCTCCTG	SEQ ID NO: 248
	hPMP22-TATA Off3	ATCATCGCTCACAGAGTCC	SEQ ID NO: 249	ACGACTGCAGGATCTTAATG	SEQ ID NO: 250
	hPMP22-TATA Off4	TGGATGGAGGTTGGGAATCC	SEQ ID NO: 251	TTGAGGCAGCAGCACTCTCC	SEQ ID NO: 252
	hPMP22-TATA Off5	AGTCTATCCTAGCAGCTCC	SEQ ID NO: 253	ACTGAGACCAGATAATGCAG	SEQ ID NO: 254
	hPMP22-TATA Off6	AAGAGATGCGAGTTGTTCC	SEQ ID NO: 255	CCTCTTCTACTCTGAGTGG	SEQ ID NO: 256
	hPMP22-TATA Off7	ACCTGGTTTATCACAAGCTA	SEQ ID NO: 257	AACGTGAACAGAAGGATTTC	SEQ ID NO: 258
	hPMP22-TATA Off8	ATCACTCCATCAGAGTCAGG	SEQ ID NO: 259	TGGCTCCTTCTATTCTCTCC	SEQ ID NO: 260

[1112]

[1113] **8.** *In silico* 오프-타겟 부위 설계

[1114] 오프-타겟 가능성 부위를 온라인 툴(www.rgenome.net)을 이용하여 *in silico*로 설계하였다. 최대 3bp의 미스매치를 오프-타겟 부위로 고려하였다.

[1115]

[1116] **9. Digenome-seq**

[1117] HeLa 세포의 게놈 DNA를 DNeasy Blood & Tissue Kit (Qiagen)를 이용하여 판매사의 프로토콜에 따라 정제하였다. 미리 인큐베이션한 Cas9 단백질(100nM)과 sgRNA(300nM)를 게놈 DNA(10㎍)와 1mL 반응액(100 mM NaCl, 50 mM Tris-HCl, 10 mM MgCl2, 100 μg/ml BSA, pH 7.9)에 37℃에서 8시간동안 혼합시켰다. 잘린 게놈 DNA에 RNase A(50㎍/mL)을 처리하고, DNeasy Tissue Kit (Qiagen)를 이용해 다시 정제하였다. 잘린 게놈 DNA의 1㎍을 Covaris system을 이용하여 조각으로 쪼개고, 조각 DNA에 라이브러리를 생성하기 위한 어댑터를 연결하였다. 그 후, 상기 라이브러리는 30-40×(Macrogen)의 시퀀싱 깊이(depth)에서 HiSeq × Ten Sequencer (Illumina)을 이용하여 전체 게놈 시퀀싱(WGS)을 수행하였다. In vitro 절단 스코어는 DNA 절단 스코어링 시스템에 의해 게놈에서 절단된 각각 염기서열 위치에서 계산되었다.

[1118]

[1119] **10. 마우스 및 신경내 주입(intraneural injection)**

[1120] 본 연구에 사용된 C22 마우스 주(B6;CBACa-Tg(PMP22)C22Clh/H)는 MRC Harwell(Oxfordshire, UK)에서 구입하였다. C22 마우스(수컷 4마리, 암컷 7마리)에 PMP22-TATA RNP를 처리하였다. 신경내 주입은 이전 연구(Daisuke Ino., J Vis Exp., (2016) 115)와 같은 방법으로 수행하였다. 6일령 마우스를 마취시키고, 수술로 마우스의 좌골신경을 노출시켰다. 신경 손상을 최소화하기 위해, 미량주입기(microinjector)에 부착된 인발 유리 마이크로피펫(pulled glass micropipettes)을 이용하여 좌골절흔(sciatic notch) 말단에 즉시 신경내 주입하였다. 마우스당 11㎍ Cas9 단백질과 2.75㎍ sgRNA의 RNP 복합체를 lipofectamine 3000 (Invitrogen, Carlsbad, CA, USA)과 함께 마우스에 주입시켰다. 본 연구에서 사용되는 모든 동물의 관리, 사용 및 치료는 실험 동물 관리 협회 평가 및 인증 협회에 따라 삼성 동물 관리 및 사용 위원회가 마련한 지침을 준수하여 수행되었다(SMC-20170206001).

[1121]

[1122] **11. 로타로드 실험(Rotarod test)**

[1123] 운동 협응력(Motor coordination)을 로타로드(rotarod) 장치(B.S. Technolab INC., Korea)를 이용하여 평가하였다. 이 실험은 마우스의 균형 및 운동 협응력를 평가하기 위해 수행하였다. 실험 전에, 마우스는 3일간의 훈련 기간을 거쳤다. 실험에서, 로타로드 실험을 위해 수평 회전 로드(21rpm)를 사용하였다. 회전하는 로드에서 마우스의 유지 시간을 측정하였고, 로드에서 최대 300초 동안 머무르도록 하였다.

[1124]

[1125] **12. 전기생리학(Electrophysiological) 검사**

[1126] 전기 생리학적 상태를 평가하기 위해, 신경 전도 검사(nerve conduction study, NCS)를 이전 연구(Jinho Lee., J biomed Sci., (2015) 22, 43)와 같은 방법으로

수행하였다. 간단히 요약하면, 마우스를 이산화탄소 가스로 마취시키고, 실험 수행 동안에 노즈 콘(nose cone)을 이용하여 1.5% 이소플루란(isoflurane)을 공급하여 마취를 유지하였다. 말단부터 뒷다리까지 털을 완전히 제거하였다. NCS는 Nicolet VikingQuest 장치(Natus Medical)를 이용하여 수행하였다. 좌골 신경의 운동 신경 전도 검사를 위해, 활성 기록 니들 전극을 힘줄에 기준 전극이 있는 비복근 근육(gastrocnemius muscle) 위에 놓고, 자극을 주는 음극을 둔두 안쪽에 몸 중심부 쪽으로 6mm 및 안후부 허벅 다리의 중간 선에서 기록 전극에 근접 위치에 배치하여 각각 원위부와 근위부의 반응을 측정하였다. 말단 지연속도(Distal latency, DL), 운동 신경 전도 속도(motor nerve conduction velocity, MNCV) 및 복합 근육 활동 전위(compound muscle action potential, CMAP)의 진폭을 측정하였다. CMAP는 최대 초과 자극에서 측정하였다.

[1127]

[1128] **13. 신경 조직학(Nerve histology) 및 이미지**

[1129] 좌골 신경을 마우스에서 생검하고 영향을 받은 검체의 병리학적 검사를 현미경으로 분석하였다. 검체들은 2% 글루타르알데히드(glutaraldehyde)를 포함하는 25mM 카코딜산염(cacodylate) 버퍼를 이용하여 각각 고정시켰다. 세미-틴 섹션(Semi-thin sections)은 톨루이딘블루(toluidine blue)로 염색하였다. 1% OsO4에서 한 시간동안 인큐베이션한 후에, 검체들은 에탄올 시리즈에서 탈수시킨 다음 프로필렌 옥사이드(propylene oxide)를 통과시키고 에폭시 수지(Epon 812, Oken, Nagano, Japan)에 묻었다. 세포는 Leica ultra-microtome (Leica Microsystems)을 이용해 어느정도 얇게(1㎛) 잘라내고, 30 - 45초간 톨루이딘블루로 염색하였다. g-ratio(axon 직경/fiber 직경)은 Zeiss Zen2 program (Carl Zeiss, Oberkochen, Germany)을 이용하여 마이엘린(myelin)의 내부 직경과 외부 직경을 측정하여 산출하였다.

[1130]

[1131] **14. 통계학적 분석**

[1132] mRNA 발현량 관련 데이터의 통계학적 유의성은 post-hoc Tukey의 다중 비교를 이용한 one-way ANOVA으로 평가되었다. 그 밖의 제시된 데이터는 Mann-Whitney U test (http://www.socscistatistics.com/tests/mannwhitney/Default2.aspx)를 이용해 계산되었다. 본 연구에서 생성된 데이터 및 그래프는 GraphPad Prism을 사용하여 분석되었다. 유의 수준은 0.05로 설정되었다.

[1133]

[1134] **15. Plp1 유전자의 표적화를 위한 sgRNA 스크리닝**

[1135] 마우스 fibroblast cells NIH-3T3(ATCC, CRL-1658), myoblast cells C2C12 line(ATCC, CRL-1772) 및 oligodendrocyte cells N20.1 (Cedarlane Laboratories, CLU108-P)는 제조사의 메뉴얼에 따라 배양하였다. 세포는 37℃ 및 5% CO_2 조건에서 1× 페니실린/스트렙토마이신(WelGene) 및 10% 태아 송아지

혈청(WelGene)를 추가한 고농도 글루코스 포함 Dulbecco's modified Eagle's medium(DMEM)로 배양하였다. CRISPR/Cas9 조성물은 형질도입(transfection)하기 위해, 4㎍의 Cas9 단백질과 1㎍의 sgRNA로 구성된 RNP 복합체(SpCas9) 또는 CjCas9 플라스미드(도 18)을 준비하였다. 그 후, 2×10^5 cells에 RNP 복합체 또는 CjCas9 플라스미드를 10㎕ 전기천공용 팁과 Neon electroporator (ThermoFisher)을 이용해 전기천공하여 전달하였다. 표적 딥 시퀀싱을 위해, 형질도입 72시간 후에 형질도입 된 세포에서 게놈 DNA(gDNA)를 수집하였다.

[1136]

[1137] **16. Plp1 유전자의 발현저해 시험(downregulation assay)**

[1138] RNeasy mini kit(Qiagen)을 이용해 제조사의 프로토콜에 따라 N20.1 cell lines으로부터 mRNA를 추출하였다. 그 후, 1㎍ mRNA는 High-capacity cDNA reverse transcription kit(ThermoFisher)를 이용하여 역 전사시켰다. qRT-PCR(Real-time quantitative reverse transcription polymerase chain reaction)은 QuantStudio 3(ThermoFisher)를 이용해 제조사의 프로토콜에 따라 Taqman Gene expression master mix 100ng으로 수행하였다. Plp1의 발현 수준은 C_T 값을 이용해 계산하였고, Gapdh로 내인성 대조군(endogenous control)으로 사용하였다. 본 연구에서 사용한 Taqman 프로브(ThermoFisher)는 하단의 표 9에 정리하였다.

[1139] [표9]

Target Gene	Taqman Gene Experssion Assay	Accession number
Plp1	Mm01297210_m1	NM_001290561.1
Gapdh	Mm99999915_g1	NM_001289726.1

[1140]

[1141] **17.PLP1 유전자의 표적화를 위한 sgRNA 스크리닝**

[1142] 인간 lymphoblast Jurkat cell line (ATCC, TIB-152) 및 인간 epithelial 293T cell line (ATCC, CRL-3216)는 제조사의 매뉴얼에 따라 배양하였다. 세포는 37℃ 및 5% CO_2 조건에서 1× 페니실린/스트렙토마이신(WelGene) 및 10% 태아 송아지 혈청(WelGene)를 추가한 고농도 글루코스 포함 Dulbecco's modified Eagle's medium(DMEM)로 배양하였다. CRISPR/Cas9 조성물은 형질도입(transfection)하기 위해, 4㎍의 Cas9 단백질과 1㎍의 sgRNA로 구성된 RNP 복합체(SpCas9) 또는 CjCas9 플라스미드(도 18)을 준비하였다. 그 후, 2×10^5 cells에 RNP 복합체 또는 CjCas9 플라스미드를 10㎕ 전기천공용 팁과 Neon electroporator (ThermoFisher)을 이용해 전기천공하여 전달하였다. 표적 딥 시퀀싱을 위해, 형질도입 72시간 후에 형질도입 된 세포에서 게놈 DNA(gDNA)를 수집하였다.

[1143]

[1144] **실시예 1. PMP22 유전자를 위한 sgRNA 스크리닝**

[1145] 인간 PMP22 발현을 정상범위로 줄일 수 있는 치료학적으로 효과적인 sgRNA 서열을 스크리닝하기 위해, PMP22 유전자의 프로모터(TATA-box) 및 인트로닉 인핸서 결합 부위를 표적기 위해 설계된 다양한 sgRNA와 Cas9을 인간 세포주로 형질도입하였다. 간단히, SpCas9 스크리닝을 위해 Jurkat 인간 T 세포를 사용하였고, CjCas9 스크리닝을 위해 HEK293T 세포를 사용하였다. gDNA를 상기 세포로부터 수집하고 표적 딥 시퀀싱을 수행하였다. NHEJ-매개 인델로 sgRNA 서열에 의해 유도된 다양한 패턴의 돌연변이를 확인하였다. 몇몇의 SpCas9-sgRNA은 두 조절 부위 모두에서 강하게 인델을 유도하였다(도 1). 특정 CjCas9-sgRNA에서 30 - 40% 인델이 유도됨을 확인하였다(도 2).

[1146]

[1147] **실시예 2. 슈반 유사 세포(Schwann-like cells)의 유전자 조작**

[1148] 비록 인간 세포에서 sgRNA로 인한 효과적인 인델 돌연변이를 확인하였지만, 이러한 효과가 슈반세포(Schwann cells)에서도 가능한지는 불확실하다. 이에 슈반세포에서 PMP22의 발현저해 및 유전자 조작 효과를 조사하기 위해, 슈반 유사 세포인 sNF02.0 세포를 이용하여 SpCas9-sgRNA 효과를 확인하였다. Jurkat 세포에서 확인한 효과적인 SpCas9-sgRNA를 sNF02.0 세포에서 반복 실험하였다. 형질도입 후 딥 시퀀싱 분석을 통해 동일한 sgRNA에 의한 동일하게 높은 인델 빈도를 확인하였다. 프로모터(TATA-box) 부위 및 인핸서-결합 위치를 표적하는 단일 sgRNA의 형질도입은 각각 31%, 59%의 인델을 유도하였다(도 3). 흥미롭게, 마이엘린 유전자의 주요 조절인자(예를 들면, EGR 또는 SOX10-결합 위치) 또는 중요한 TATA-box가 포함된 40 - 50bp의 작은 삭제는 이중 sgRNA를 처리한 세포의 매우 많은 세포에서 발견되었다(도 5).

[1149]

[1150] **실시예 3. 유전자 조작에 의한 PMP22의 발현 조절**

[1151] 효과적인 sgRNA에 의한 PMP22 발현의 변화를 평가하기 위해, 슈반 유사 세포를 분화시키고, qRT-PCR을 수행하였다. 그 결과, PMP22를 표적하는 대부분의 sgRNA가 효과적으로 PMP22의 발현을 저해하였다(도 6). 단일 sgRNA를 이용한 경우 Cas9만 처리한 대조군에 비해 PMP22의 발현이 30% 정도 감소하였고, 이중 sgRNA를 이용한 경우 PMP22의 발현이 50% 정도 감소하였다.

[1152]

[1153] **실시예 4. 슈반세포(Schwann cells)의 유전자 조작**

[1154] 앞서 슈반 유사 세포에서 PMP22의 발현저해 및 유전자 조작 효과를 확인한 후, 인간 슈반세포(human primary schwann cells)에서 앞선 결과가 유사한 효과를 나타내는지 확인하였다. 인간 슈반세포에 인간 PMP22 유전자의 각 표적 부위별 SpCas9-sgRNA을 이용하여 표적 부위에 따른 인델 빈도를 관찰하였다. 그 결과, PMP22 유전자의 TATA-box, 인핸서 및 암호화 서열을 표적하는 대부분의 sgRNA에서 표적 부위에 인델 빈도가 높게 확인되었다(도 7a). 또한, TATA-box,

인핸서를 각각 표적하는 이중 sgRNA를 이용한 경우에도 높은 인델 빈도를 보였다. 추가적으로 원거리 인핸서 부위 B 및 C를 암호화 하는 서열을 표적하는 sgRNA를 이용해 표적 부위가 인델이 발생함은 확인하였고(도 7c), 이때, 대조군으로 APOC3를 표적하는 sgRNA을 사용하였다.

[1155] 이와 더불어, 각 표적 부위별 SpCas9-sgRNA가 PMP22 유전자의 발현 감소를 야기하는지 확인하기 위해, qRT-PCR 분석을 수행하였다. PMP22는 슈반세포의 분화 마지막 단계에서 전사되기 때문에, 인간 슈반세포에 Nrg1(Neuregulin-1)과 dbcAMP(dibutryral cyclic AMP)를 포함하는 잘 알려진 분화 신호인자를 7일 동안 처리하였다. 그 결과, Nrg1 및 dbcAMP를 처리하지 않은 세포에 비해 처리한 세포에서 9 배 증가된 PMP22의 발현을 확인하였다. 반면에, 각 표적 부위별 SpCas9-sgRNA를 처리한 경우에는 4 ~ 6배로 유도된 PMP22의 발현을 확인하였다. 이는 각 표적 부위별 SpCas9-sgRNA에 의한 PMP22의 각 표적 부위 변형에 따른 PMP22의 발현 저해로 인한 것으로 판단된다(도 7b).

[1156]

[1157] **<u>실시예 5. 인간 PMP22 유전자의 TATA-box 부위를 표적하는 CRISPR/Cas9을 이용한 PMP22의 효과적 및 특이적 발현감소 효과(*in vitro*)</u>**

[1158] 앞서 스크리닝한 TATA-box 부위를 표적하는 sgRNA 중 높은 인델 효율을 보이며 TATA-box를 표적할 수 있는 sgRNA_TATA_Sp#1(이후 PMP22-TATA sgRNA로 기재함)를 선택하여 인간 슈반세포(human primary schwann cells)에서 실험하였다. 인간 슈반세포에 sgRNA와 Cas9 단백질을 포함하는 RNP 복합체를 형질도입하여 인델을 유도하였고(도 8b), 표적 딥 시퀀싱 분석을 통해 전체 인델 중 89.54±1.39%가 인간 PMP22의 TATA-box 부위에서 발생하였음을 확인하였다(도 8c).

[1159] 또한, PMP22의 TATA-box에 형성된 돌연변이가 PMP22 유전자의 발현 감소를 야기하는지 확인하기 위해, qRT-PCR 분석을 수행하였다. PMP22는 슈반세포의 분화 마지막 단계에서 전사되기 때문에, 인간 슈반세포에 Nrg1(Neuregulin-1)과 dbcAMP(dibutryral cyclic AMP)를 포함하는 잘 알려진 분화 신호인자를 7일 동안 처리하였다. 그 결과, Nrg1 및 dbcAMP를 처리하지 않은 세포에 비해 처리한 세포에서 9 배 증가된 PMP22의 발현을 확인하였다. 반면에, PMP22-TATA RNP를 같이 처리한 경우, 6배 유도된 PMP22의 발현을 확인하였다. 이는 CRISPR/Cas9에 의한 PMP22의 TATA 변형에 따른 PMP22의 발현 저해로 인한 것으로 판단된다(도 8d). 분화 신호인자와 AAVS1 표적 RNP를 함께 처리한 대조군에서는 PMP22 유전자의 발현의 차이를 확인할 수 없었다.

[1160] PMP22-TATA RNP의 특이성을 확인하기 위해, in silico 기반 오프-타겟 분석을 수행하였다. 표적 딥 시퀀싱을 통해 in silico 분석에 의해 확인된 오프-타겟 위치에서 시퀀싱 에러 비율(평균적으로 0.1%)을 넘는 인델 돌연변이는 확인되지 않았다(도 9). In silico 기반 오프-타겟 분석이 편향된 접근일 수 있기 때문에, 더불어 Digenome-seq(편향되지 않은 전체 게놈 시퀀싱 기반 오프-타겟 분석)을

수행하였다. 그 결과, in vitro에서 PMP22-TATA RNP에 의해 절단된 9개의 오프-타겟 위치를 확인할 수 있었다(도 10a, 도 10b). 하지만 표적 딥 시퀀싱을 통한 재분석 결과, 상기 오프-타겟 위치에서 비정상적 인델 돌연변이는 발견되지 않았다(도 10c).

[1161] 이러한 결과는 PMP22-TATA RNP에 의한 PMP22의 TATA-box의 효과적 및 특이적 변형이 인간 슈반세포에서 PMP22의 전사수준을 조절할 수 있음을 보여준다.

[1162]

[1163] **실시예 6. CMT1A 마우스에서 CRISPR/Cas9-매개 PMP22의 발현저해에 의한 질병 표현형 개선 효과**

[1164] In vivo에서 PMP22-TATA RNP에 의한 PMP22 전사조절을 시험하기 위해, C22 마우스의 좌골 신경으로 리포좀으로 둘러싸인 PMP22-TATA RNP를 직접 주입하였다(도 11). 이때, 대조군으로 Rosa26(mRosa26)를 표적하는 RNP 복합체로 사용하였다. mRosa26 RNP 또는 PMP22-TATA RNP는 6일령(p6) 마우스의 왼쪽 좌골 신경(ipsilateral)으로 신경내 주입하여 전달하였고, 오른쪽 좌골 신경은 내부 대조군(contralateral)으로 사용하였다. 주입 4주 후에 좌골 신경으로부터 게놈 DNA를 수집하여 표적 딥 시퀀싱을 통해 RNP 복합체의 신경내 전달 효율을 확인하였다. 그 결과, mRosa26 RNP 및 PMP22-TATA RNP를 각각 처리한 좌골 신경 모두 11% 정도의 인델 효율을 나타났다(도 12a). 또한, in vitro 결과와 일치하게 전체 인델 시퀀싱 리드 중 98.48±0.15%의 TATA-box 돌연변이가 확인되었다(도 12b).

[1165] 이와 더불어, in vivo에서 TATA-box 돌연변이에 의한 PMP22의 발현저해를 확인하기 위해, RNP를 처리한 좌골 신경에서 전체 좌골 신경으로부터 추출한 mRNA의 qRT-PCR 분석을 수행하였다. In vitro 결과와 유사하게, PMP22 유전자 발현이 대조군에 비해 38% 감소함을 확인하였다(도 12c).

[1166] PMP22-TATA RNP에 의해 좌골 신경에서 오프-타겟 돌연변이가 발생하였는지 확인하기 위해, in silico 기반 오프-타겟 분석을 수행하였다. 그 결과, 마우스 게놈에서 3bp 미스매치 이상을 포함하는 오프-타겟이 8개 확인되었고(도 13a), 표적 딥 시퀀싱을 수행한 결과, PMP22-TATA RNP를 처리한 신경(ipsilateral)에서 시퀀싱 에러 비율을 넘는 인델 돌연변이는 확인되지 않았다(도 13b).

[1167] PMP22-TATA RNP에 의한 PMP22 전사의 감소가 demyelination을 막을 수 있는지 시험하기 위해, PMP22-TATA RNP 또는 mRosa26 RNP를 처리한 C22 마우스의 좌골 신경을 수득하여 semi-thin cross sections하여 toluidine-blue(myelin 염색)로 염색하였다. 또한, g-ratio를 측정하기 위해, axon 직경과 fiber(myelin을 포함하는 axon) 직경을 측정하였다. 그 결과, PMP22-TATA RNP를 처리한 실험군에서 더 두꺼운 myelin 시트가 형성됨을 확인할 수 있었다(도 14a, 도 14b). 더불어, mRosa26 RNP를 처리한 대조군에 비해 PMP22-TATA RNP를 처리한 경우 큰 직경을 가지는 axon의 수가 많아짐을 발견하였다(도 14a, 도 14b).

대조군(2.6%, p<0.01)에 비해 PMP22-TATA RNP를 처리한 실험군(16.5%)에서 직경이 6㎛ 이상인 큰 myelinated fiber의 수를 측정한 결과가 더욱 명확한 치료적 효과를 보여준다.

[1168] Myelination의 조직학 분석에서 상당히 개선되었다는 것을 감안하여, 두 그룹의 전기 생리학적 프로파일을 조사하였다. 그 결과, mRosa26 RNP를 처리한 대조군에 비해 PMP22-TATA RNP를 처리한 실험군의 좌골 신경에서 감소된 말단 지연속도(distal latency, DL) 및 증가한 운동 신경 전도 속도(motor nerve conduction velocity, NCV)를 확인하였고(도 15a, 도 15b), 이러한 결과는 PMP22-TATA RNP 처리한 신경에서 myelin 두께 및 axon의 직경의 증가에 상응되는 결과이다. 또한, PMP22-TATA RNP를 처리한 신경에서 복합 근육 활동 전위(compound muscle action potential, CMAP)의 진폭이 상당히 증가하는 것을 확인하였으며(도 15c), 이는 앞선 결과와 상응된다.

[1169] PMP22-TATA RNP에 의한 조직학적 및 전기 생리학적 향상된 효과를 감안하여, 로타로드 실험으로 마우스의 이주 행동(locomotor behaviour)을 분석하였다. 그 결과, PMP22-TATA RNP를 처리한 마우스(11 - 16주령)가 mRosa26 RNP를 처리한 마우스(11 - 16주령)에 비해 더 오래 남아있는 것을 확인하였다(도 16a). 또한, MP22-TATA RNP를 처리한 마우스가 mRosa26 RNP를 처리한 마우스에 비해 근육이 증가됨을 확인하였다(도 16b).

[1170] 이러한 결과들은 CMT1A와 같은 PMP22의 과발현에 의한 demyelination를 개선 또는 치료하기 위한 PMP22-TATA RNP의 치료적 효과를 보여준다.

[1171]

[1172] 따라서, 상기의 결과들은 통해 PMP22의 프로모터 부위를 표적으로 하는 CRISPR/Cas9을 이용한 PMP22의 발현저해 효과를 보여준다. 또한, C22 마우스의 좌골 신경으로의 직접적인 PMP22-TATA RNP의 비바이러스 전달이 PMP22 과발현에 의해 야기되는 demyelination과 관련된 임상 및 신경병리학적 표현형을 개선시킬 수 있음을 보여준다. 그러므로, PMP22의 전사 조절 영역의 CRISPR/Cas9 매개한 인위적인 변형은 CMT1A 및 demyelinating neuropathies를 보이는 다른 질병의 치료를 위해 좋은 전략이 될 수 있다고 사료된다.

[1173]

[1174] **실시예 7. PLP 유전자의 발현 조절 효과**

[1175] PLP1 유전자가 중복되면 PLP1 유전자가 overexpress되고 이는 PMD 질환의 주요 원인이 된다. 따라서, PMD 질환의 치료를 위해 PLP1의 전사를 조절하기 위해 CRISPR/Cas9을 이용하여 PLP 유전자의 전사 조절 영역을 인위적으로 변형하여 그 효과를 확인하고자 하였다.

[1176] 이를 위해 mouse plp1의 프로모터 서열의 TATA-box, 인핸서 (wMN1)에 대한 SpCas9, CjCas9 스크리닝을 진행하였고, 가장 활성도가 높은 sgRNA를 선정한 뒤 qRT-PCR을 통해 plp1의 downregulation을 검증 하였다(도 17). 이때, Plp1의 인핸서는 ASE (Hamdan et al., 2015; Meng et al., 2005; Wight, 2017) 혹은 wMN1

(Hamdan et al., 2018; Tuason et al., 2008) 부위가 될 수 있다.

[1177] sgRNA의 스크리닝 결과를 바탕으로 indel 효율이 높은 SpCas9 및 CjCas9 타겟 sgRNA를 각각 선별하였고 (도 19-22 및 표 10), N20.1 이라는 Plp1유전자를 발현하는 oligodendrocyte cell line을 사용하여 Plp1의 TATA-box 및 wMN1 인핸서 부위를 타겟하였을 때 Plp1 유전자의 발현저해로 이어질 수 있는지에 대한 연구를 qRT-PCR을 통해 진행하였다.

[1178] [표 10]

[1179] 선별된 sgRNA 리스트(mPlp1-TATA, mPlp1-wmN1 SpCas9 and CjCas9 lead sgRNA list)

[1180]

mPlp1-TATA-SpCas9			
No.	#RGEN Target (5' to 3')	Indel ratio (%)	
1	TGTTTGGTAGTATAGTAAGTAGG (SEQ ID NO: 116)	74.6	
mPlp1-WmN1-SpCas9			
No.	#RGEN Target (5' to 3')	Indel ratio (%)	location
26	CTCCCACTGCCTTATTAGGCAGG (SEQ ID NO: 141)	98.9	Up
27	AGAGCTCAAATGGGTTCTAAAGG (SEQ ID NO: 142)	99.1	Up
28	ACCACATTCAAGAGCTCAAATGG (SEQ ID NO: 143)	98.6	Up
8	ATCACAGTTTATACTTAGCTGGG (SEQ ID NO: 123)	48.4	Down
9	GGAATACCTCAGGCTCAACAGGG (SEQ ID NO: 124)	66.6	Down
mPlp1-TATA-CjCas9			
No.	#RGEN Target (5' to 3')	Indel ratio (%)	
2	AAAGCCTACTTACTATACTACCAAACACAC (SEQ ID NO: 154)	27.9	
3	CAAAAGCCTACTTACTATACTACCAAACAC (SEQ ID NO: 155)	33.6	
mPlp1-wMN1-CjCas9			
No.	#RGEN Target (5' to 3')	Indel ratio (%)	location
10	GACATACAGAGAGGGGGCGGAGAGAAATAC (SEQ ID NO: 162)	28.5	Up
25	TTGAATGTGGTATAAGTGCTAATATCATAC (SEQ ID NO: 177)	33.7	Up
13	TCATCAAAGTAGTCGACAGTCAAAGCATAC (SEQ ID NO: 165)	13.8	Down
14	TGAATTCTAACAGGAAAACTCAGAACATAC (SEQ ID NO: 166)	15.7	Down
23	TTCCAAAGTTCTGTCACCCAGTAAAAACAC (SEQ ID NO: 175)	5.4	Down

[1181]

[1182] 그 결과, SpCas9 및 CjCas9을 이용하여 Plp1의 TATA 박스 또는 wmN1 인핸서 영역을 표적 하였을 때 유의미한 Plp1의 발현저해로 이어짐을 확인하였다(도 23). 또한, 인간 PLP1 유전자의 wmN1 인핸서 영역에 대한 SpCas9, CjCas9 스크리닝을 진행하여 인델(%)을 확인하였다(도 24, 도 25).

[1183] 따라서, PLP1의 전사 조절 영역의 CRISPR/Cas9 매개한 인위적인 변형은

PMD의 치료를 위해 좋은 전략이 될 수 있다고 사료된다.

[1184]

산업상 이용가능성

[1185] 중복 유전자의 발현을 조절하기 위한 발현 조절용 조성물을 통해 유전자 중복 질환을 위한 치료제를 수득할 수 있다. 예를 들어, 중복 유전자의 전사 조절 영역을 표적화할 수 있는 가이드핵산을 포함하는 발현 조절용 조성물은 중복 유전자의 전사 조절 영역을 인위적으로 조작 및/또는 변형시켜 중복 유전자의 발현을 조절함으로써 유전자 중복에 의해 발생하는 질환을 위한 치료제로 이용될 수 있다.

[1186]

서열목록 Free Text

[1187] 중복 유전자의 전사 조절 영역에 대한 표적 서열 및 실시예에 사용된 프라이머 서열

청구범위

[청구항 1] 세포의 게놈 내 존재하는 중복 유전자의 발현 조절용 조성물로,
상기 조성물은
중복 유전자의 전사 조절 영역 내 표적서열을 표적화할 수 있는 하나 이상의 가이드핵산 또는 이를 암호화하는 핵산 서열; 및
하나 이상의 에디터단백질 또는 이를 암호화하는 핵산 서열을 포함하고,
상기 가이드핵산은 중복 유전자의 전사 조절 영역 내 표적 서열을 표적화할 수 있는 가이드 도메인을 포함하고,
상기 가이드 도메인은 중복 유전자의 전사 조절 영역 내 표적 서열 중 가이드핵산 결합 서열과 상보적 결합을 형성할 수 있는 뉴클레오타이드 서열을 포함하며,
상기 가이드 도메인은 중복 유전자의 전사 조절 영역 내 표적 서열 중 가이드핵산 결합 서열과 상보적 결합을 할 수 있으며,
상기 상보적 결합은 0 내지 5개의 미스매칭(mismatching) 결합을 포함하며,
상기 가이드핵산은 제 1 상보적 도메인, 연결 도메인, 제 2 상보적 도메인, 근위 도메인 및 꼬리 도메인으로 구성된 군에서 선택된 하나 이상의 도메인을 포함하고,
상기 에디터단백질은 CRISPR 효소이며,
상기 가이드핵산 및 에디터단백질은 가이드핵산-에디터단백질 복합체를 형성할 수 있고,
상기 가이드핵산-에디터단백질 복합체는 가이드핵산의 일부 핵산과 에디터단백질의 일부 아미노산이 상호작용하여 형성될 수 있는 중복 유전자의 발현 조절용 조성물.

[청구항 2] 제 1항에 있어서,
상기 전사 조절 영역은 프로모터 영역, 인핸서 영역, 사일런서 영역, 인슐레이터 영역 및 로커스 컨트롤 영역(locus control region, LCR)으로 구성된 군에서 선택된 하나 이상의 영역인 것을 특징으로 하는 중복 유전자의 발현 조절용 조성물.

[청구항 3] 제 1항에 있어서,
상기 표적 서열은 중복 유전자의 전사 조절 영역 내 위치하는 연속하는 10 내지 25개의 뉴클레오타이드 서열인 것을 특징으로 하는 중복 유전자의 발현 조절용 조성물.

[청구항 4] 제 1항에 있어서,
상기 표적 서열은 중복 유전자의 프로모터 영역 내에 연속하는 10 내지 25개의 뉴클레오타이드 서열 또는 프로모터 영역에 인접한 연속하는 10 내지 25개의 뉴클레오타이드 서열인 것을 특징으로 하는 중복 유전자의

발현 조절용 조성물.

[청구항 5] 제 4항에 있어서,
상기 표적 서열은 중복 유전자의 코어 프로모터 영역 내에 연속하는 10 내지 25개의 뉴클레오타이드 서열 또는 코어 프로모터 영역에 인접한 연속하는 10 내지 25개의 뉴클레오타이드 서열인 것을 특징으로 하는 중복 유전자의 발현 조절용 조성물.

[청구항 6] 제 5항에 있어서,
상기 표적 서열은 중복 유전자의 코어 프로모터 영역 중 TATA 박스 영역을 포함하는 연속하는 10 내지 25개의 뉴클레오타이드 서열 또는 TATA 박스 영역에 인접한 연속하는 10 내지 25개의 뉴클레오타이드 서열 인 것을 특징으로 하는 중복 유전자의 발현 조절용 조성물.

[청구항 7] 제 5항에 있어서,
상기 표적 서열은 중복 유전자의 코어 프로모터 영역 내에 존재하는 5'-TATA-3' 서열의 전체 또는일부를 포함하는 연속하는 10 내지 25개의 뉴클레오타이드 서열인 것을 특징으로 하는 중복 유전자의 발현 조절용 조성물.

[청구항 8] 제 5항에 있어서,
상기 표적 서열은 중복 유전자의 코어 프로모터 영역 내에 존재하는 5'-TATAWAW-3' 서열(W=A 또는 T)의 전체 또는 일부를 포함하는 연속하는 10 내지 25개의 뉴클레오타이드 서열인 것을 특징으로 하는 중복 유전자의 발현 조절용 조성물.

[청구항 9] 제 5항에 있어서,
상기 표적 서열은 중복 유전자의 코어 프로모터 영역 내에 존재하는 5'-TATAWAWR-3' 서열(W = A 또는 T, R = A 또는 G)의 전체 또는 일부를 포함하는 연속하는 10 내지 25개의 뉴클레오타이드 서열인 것을 특징으로 하는 중복 유전자의 발현 조절용 조성물.

[청구항 10] 제 5항에 있어서,
상기 표적 서열은 중복 유전자의 코어 프로모터 영역 내에 존재하는 5'-CATAAAA-3' 서열, 5'-CATAAAA-3' 서열, 5'-TATAA-3' 서열, 5'-TATAAAA-3' 서열, 5'-CATAAATA-3' 서열, 5'-TATATAA-3' 서열, 5'-TATATATATATATAA-3' 서열, 5'-TATATTATA-3' 서열, 5'-TATAAA-3' 서열, 5'-TATAAAATA-3' 서열, 5'-TATATA-3' 서열, 5'-GATTAAAAA-3' 서열, 5'-TATAAAAA-3' 서열, 5'-TTATAA-3' 서열, 5'-TTTTAAAA-3' 서열, 5'-TCTTTAAAA-3' 서열, 5'-GACATTTAA-3' 서열, 5'-TGATATCAA-3' 서열, 5'-TATAAATA-3' 서열, 5'-TATAAGA-3' 서열, 5'-AATAAA-3' 서열, 5'-TTTATA-3' 서열, 5'-CATAAAAA-3' 서열, 5'-TATACA-3' 서열, 5'-TTTAAGA-3' 서열, 5'-GATAAAG-3' 서열, 5'-TATAACA-3' 서열, 5'-TCTTATCTT-3' 서열, 5'-TTGTACTTT-3' 서열, 5'-CATATAA-3' 서열,

5'-TATAAAT-3' 서열, 5'-TATATATAAAAAAAA-3' 서열 및 5'-CATAAATAAAAAAAATTA-3' 서열로 구성된 군에서 선택된 하나의 서열의 전체 또는 일부를 포함하는 연속하는 10 내지 25개의 뉴클레오타이드 서열인 것을 특징으로 하는 중복 유전자의 발현 조절용 조성물.

[청구항 11] 제 5항에 있어서,
상기 표적 서열은 중복 유전자의 코어 프로모터 영역 내에 존재하는 5'-TATA-3' 서열, 5'-CATAAAA-3' 서열, 5'-CATAAAA-3' 서열, 5'-TATAA-3' 서열, 5'-TATAAAA-3' 서열, 5'-CATAAATA-3' 서열, 5'-TATATAA-3' 서열, 5'-TATATATATATATAA-3' 서열, 5'-TATATTATA-3' 서열, 5'-TATAAA-3' 서열, 5'-TATAAAATA-3' 서열, 5'-TATATA-3' 서열, 5'-GATTAAAAA-3' 서열, 5'-TATAAAAA-3' 서열, 5'-TTATAA-3' 서열, 5'-TTTTAAAA-3' 서열, 5'-TCTTTAAAA-3' 서열, 5'-GACATTTAA-3' 서열, 5'-TGATATCAA-3' 서열, 5'-TATAAATA-3' 서열, 5'-TATAAGA-3' 서열, 5'-AATAAA-3' 서열, 5'-TTTATA-3' 서열, 5'-CATAAAAA-3' 서열, 5'-TATACA-3' 서열, 5'-TTTAAGA-3' 서열, 5'-GATAAAG-3' 서열, 5'-TATAACA-3' 서열, 5'-TCTTATCTT-3' 서열, 5'-TTGTACTTT-3' 서열, 5'-CATATAA-3' 서열, 5'-TATAAAT-3' 서열, 5'-TATATATAAAAAAAA-3' 서열 및 5'-CATAAATAAAAAAAATTA-3' 서열로 구성된 군에서 선택된 하나의 서열의 5' 말단 또는3' 말단에 연속하는 10 내지 25개의 뉴클레오타이드 서열인 것을 특징으로 하는 중복 유전자의 발현 조절용 조성물.

[청구항 12] 제 1항에 있어서,
상기 표적 서열은 중복 유전자의 인핸서 영역 내에 연속하는 10 내지 25개의 뉴클레오타이드 서열인 것을 특징으로 하는 중복 유전자의 발현 조절용 조성물.

[청구항 13] 제 1항에 있어서,
상기 표적 서열은 중복 유전자의 인핸서 영역에 인접한 연속하는 10 내지 25개의 뉴클레오타이드 서열인 것을 특징으로 하는 중복 유전자의 발현 조절용 조성물.

[청구항 14] 제 1항에 있어서,
상기 표적 서열은 중복 유전자의 전사 조절 영역 내 핵산 서열 중 PAM(proto-spacer-adjacent Motif) 서열의 5' 말단 및/또는 3' 말단에 인접한 연속하는 10 내지 25개의 뉴클레오타이드 서열인 것을 특징으로 하는 중복 유전자의 발현 조절용 조성물.

[청구항 15] 제 14항에 있어서,
상기 PAM 서열은 CRISPR 효소에 따라 결정되는 것을 특징으로 하는 중복 유전자의 발현 조절용 조성물.

[청구항 16] 제 1항에 있어서,
상기 CRISPR 효소는 Cas9 단백질 또는 Cpf1 단백질인 것을 특징으로 하는 중복 유전자의 발현 조절용 조성물.

[청구항 17] 제 16항에 있어서,
상기 Cas9 단백질은 스트렙토코커스 피요게네스(Streptococcus pyogenes) 유래의 Cas9 단백질, 캄필로박터 제주니 (Campylobacter jejuni) 유래의 Cas9 단백질, 스트렙토코커스 써모필러스 (Streptococcus thermophilus) 유래의 Cas9 단백질, 스타필로코커스 아우레우스 (Staphylococcus aureus) 유래의 Cas9 단백질 및 네이세리아 메닝기디티스 (Neisseria meningitidis)유래의 Cas9 단백질로 이루어진 군에서 선택된 하나 이상의 Cas9 단백질인 것을 특징으로 하는 중복 유전자의 발현 조절용 조성물.

[청구항 18] 제 1항에 있어서,
상기 중복 유전자는 PMP22 유전자, PLP1 유전자, MECP2 유전자, SOX3 유전자, RAI1 유전자, TBX1 유전자, ELN 유전자, JAGGED1 유전자, NSD1 유전자, MMP23 유전자, LMB1 유전자, SNCA 유전자 및 APP 유전자로 구성된 군에서 선택된 하나 이상의 유전자인 것을 특징으로 하는 중복 유전자의 발현 조절용 조성물.

[청구항 19] 제 1항에 있어서,
상기 세포는 진핵세포인 것을 특징으로 하는 중복 유전자의 발현 조절용 조성물.

[청구항 20] 제 19항에 있어서,
상기 진행세포는 포유동물 세포인 것을 특징으로 하는 중복 유전자의 발현 조절용 조성물.

[청구항 21] 제 1항에 있어서,
상기 가이드핵산 및 에디터단백질은 각각 핵산 서열의 형태로 1 이상의 벡터에 존재하는 것을 특징으로 하는 중복 유전자의 발현 조절용 조성물.

[청구항 22] 제 21항에 있어서,
상기 벡터는 플라스미드 또는 바이러스 벡터인 것을 특징으로하는 중복 유전자의 발현 조절용 조성물.

[청구항 23] 제 22항에 있어서,
상기 바이러스 벡터는 레트로바이러스, 렌티바이러스, 아데노바이러스, 아데노-연관 바이러스(AAV), 백시니아바이러스, 폭스바이러스 및 단순포진 바이러스로 구성된 군에서 선택된 하나 이상의 바이러스 벡터인 것을 특징으로 하는 중복 유전자의 발현 조절용 조성물.

[청구항 24] 제 1항에 있어서,
상기 중복 유전자의 발현 조절용 조성물은 가이드핵산 및 에디터단백질을 가이드핵산-에디터단백질 복합체 형태로 포함하는 것을 특징으로 하는 중복 유전자의 발현 조절용 조성물.

[청구항 25] 제 1항에 있어서,
상기 중복 유전자의 발현 조절용 조성물은 도너를 추가로 더 포함하는 것을 특징으로 하는 중복 유전자의 발현 조절용 조성물.

[청구항 26] 진핵 세포의 게놈 내 존재하는 중복 유전자의 발현을 조절하기 위한 방법으로,
상기 방법은 중복 유전자의 발현 조절용 조성물을 진핵 세포에 도입하는 단계를 포함하고,
상기 발현 조절용 조성물은
중복 유전자의 전사 조절 영역 내 표적서열을 표적화할 수 있는 하나 이상의 가이드핵산 또는 이를 암호화하는 핵산 서열; 및
하나 이상의 에디터단백질 또는 이를 암호화하는 핵산 서열을 포함하고,
상기 진핵 세포는 포유동물 세포이고,
상기 가이드핵산은 중복 유전자의 전사 조절 영역 내 표적 서열을 표적화할 수 있는 가이드 도메인을 포함하고,
상기 가이드 도메인은 중복 유전자의 전사 조절 영역 내 표적 서열 중 가이드핵산 결합 서열과 상보적 결합을 형성할 수 있는 뉴클레오타이드 서열을 포함하며,
상기 가이드 도메인은 중복 유전자의 전사 조절 영역 내 표적 서열 중 가이드핵산 결합 서열과 상보적 결합을 할 수 있으며,
상기 상보적 결합은 0 내지 5개의 미스매칭(mismatching) 결합을 포함하며,
상기 가이드핵산은 제 1 상보적 도메인, 연결 도메인, 제 2 상보적 도메인, 근위 도메인 및 꼬리 도메인으로 구성된 군에서 선택된 하나 이상의 도메인을 포함하고,
상기 에디터단백질은 CRISPR 효소이며,
상기 CRISPR 효소는 Cas9 단백질 또는 Cpf1 단백질이고,
상기 가이드핵산 및 에디터단백질은 가이드핵산-에디터단백질 복합체를 형성할 수 있고,
상기 가이드핵산-에디터단백질 복합체는 가이드핵산의 일부 핵산과 에디터단백질의 일부 아미노산이 상호작용하여 형성될 수 있는 중복 유전자의 발현을 조절하기 위한 방법.

[청구항 27] 제 26항에 있어서,
상기 발현 조절용 조성물은 가이드핵산 및 에디터단백질을 가이드핵산-에디터단백질 복합체 형태로 포함하는 것을 특징으로 하는 중복 유전자의 발현을 조절하기 위한 방법.

[청구항 28] 제 26항에 있어서,
상기 발현 조절용 조성물은 가이드핵산 및 에디터단백질을 각각 핵산 서열로 포함하는 1 이상의 벡터를 포함하는 것을 특징으로 하는 중복

유전자의 발현을 조절하기 위한 방법.

[청구항 29] 제 26항에 있어서,
상기 도입은 전기천공법 (electroporation), 리포좀, 플라스미드, 바이러스 벡터, 나노파티클(nanoparticles) 및 PTD (Protein translocation domain) 융합 단백질 방법 중 선택되는 1 이상의 방법으로 수행되는 것을 특징으로 하는 중복 유전자의 발현을 조절하기 위한 방법.

[청구항 30] 유전자 중복 질환을 치료하기 위한 방법으로,
상기 방법은 중복 유전자의 발현 조절용 조성물을 치료 대상에 투여하는 단계를 포함하고,
상기 발현 조절용 조성물은
중복 유전자의 전사 조절 영역 내 표적서열을 표적화할 수 있는 가이드핵산 또는 이를 암호화하는 핵산 서열; 및
하나 이상의 에디터단백질 또는 이를 암호화하는 핵산 서열을 포함하고,
상기 가이드핵산은 중복 유전자의 전사 조절 영역 내 표적 서열을 표적화할 수 있는 가이드 도메인을 포함하고,
상기 가이드 도메인은 중복 유전자의 전사 조절 영역 내 표적 서열 중 가이드핵산 결합 서열과 상보적 결합을 형성할 수 있는 뉴클레오타이드 서열을 포함하며,
상기 가이드 도메인은 중복 유전자의 전사 조절 영역 내 표적 서열 중 가이드핵산 결합 서열과 상보적 결합을 할 수 있으며,
상기 상보적 결합은 0 내지 5개의 미스매칭(mismatching) 결합을 포함하며,
상기 가이드핵산은 제 1 상보적 도메인, 연결 도메인, 제 2 상보적 도메인, 근위 도메인 및 꼬리 도메인으로 구성된 군에서 선택된 하나 이상의 도메인을 포함하고,
상기 에디터단백질은 CRISPR 효소이며,
상기 CRISPR 효소는 Cas9 단백질 또는 Cpf1 단백질이고,
상기 가이드핵산 및 에디터단백질은 가이드핵산-에디터단백질 복합체를 형성할 수 있고,
상기 가이드핵산-에디터단백질 복합체는 가이드핵산의 일부 핵산과 에디터단백질의 일부 아미노산이 상호작용하여 형성될 수 있는
유전자 중복 질환을 치료하기 위한 방법.

[청구항 31] 제 30항에 있어서,
상기 유전자 중복 질환은 샤르코마리투스병 1A형(Charcot-Marie-Tooth 1A, CMT1A), 데제린-소타스병(Dejerine-Sottas disease, DSD), 선천성 저수초형성 신경병(Congenital Hypomyelination Neuropathy, CHN), 루시-레비 증후군(Roussy-Levy syndrome, RLS), 펠리제우스-메르츠바하병(Pelizaeus-Merzbacher disease, PMD), MECP2

중복 증후군(MECP2 duplication syndrome), X-연관 뇌하수체기능저하증(X-linked hypopituitarism, XLHP), Potocki-Lupski 증후군(Potocki-Lupski syndrome, PTLS), 구개심장안면증후군(Velocardiofacial syndrome, VCFS), 윌리엄스 보이렌 증후군(Williams Beuren syndrome, WBS), 알라질증후군(Alagille syndrome, AS), 발육지연 증후군(Growth retardation syndrome), 조기 폐쇄 두개봉합(Premature closure cranial sutures), 보통염색체우성 백질이영양증(Autosomal dominant leukodystrophy, ADLD), 파킨슨병(Parkinson disease) 또는 알츠하이머(Alzheimer disease)인 것을 특징으로 하는 유전자 중복 질환을 치료하기 위한 방법.

[청구항 32] 제 30항에 있어서,
상기 치료 대상은 인간, 원숭이, 마우스, 래트를 포함하는 포유동물인 것을 특징으로 유전자 중복 질환을 치료하기 위한 방법.

[청구항 33] 제 30항에 있어서,
상기 투여는 주사(injection), 수혈(transfusion), 삽입(implantation) 또는 이식(transplantation)으로 수행되는 것을 특징으로 유전자 중복 질환을 치료하기 위한 방법.