



**(19) 대한민국특허청(KR)**  
**(12) 등록특허공보(B1)**

(45) 공고일자 2019년09월24일  
 (11) 등록번호 10-2023577  
 (24) 등록일자 2019년09월16일

(51) 국제특허분류(Int. Cl.)  
**A61K 38/46** (2006.01) **A61K 48/00** (2006.01)  
**A61P 35/00** (2006.01)  
 (52) CPC특허분류  
**A61K 38/465** (2013.01)  
**A61K 48/00** (2013.01)  
 (21) 출원번호 10-2019-0035288  
 (22) 출원일자 2019년03월27일  
 심사청구일자 2019년03월27일  
 (30) 우선권주장  
 1020180035298 2018년03월27일 대한민국(KR)  
 (56) 선행기술조사문헌  
 KR1020180020125 A\*  
 (뒷면에 계속)

(73) 특허권자  
**(주)지플러스 생명과학**  
 서울특별시 관악구 낙성대로 38, 1층(봉천동)  
 (72) 발명자  
**최성화**  
 서울특별시 서초구 양재대로2길 90, 204동 1703호(우면동, 서초힐스)  
**박미진**  
 서울특별시 동작구 상도로38길 73, 202호 (상도동)  
 (뒷면에 계속)  
 (74) 대리인  
**제일특허법인(유)**

전체 청구항 수 : 총 13 항

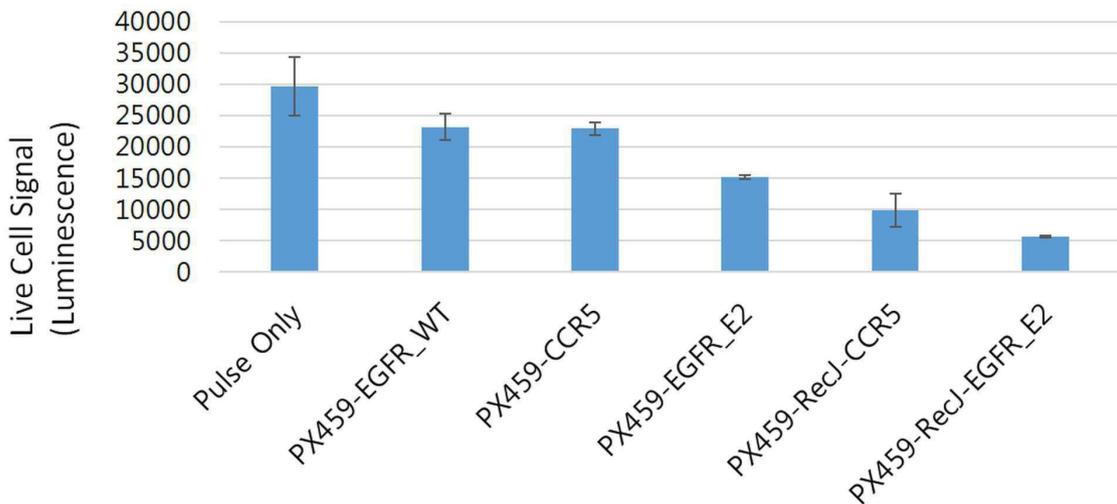
심사관 : 이예리

(54) 발명의 명칭 **가이드 RNA 및 엔도뉴클레아제를 유효성분으로 포함하는 암 치료용 약학적 조성물**

**(57) 요약**

본 발명은 crRNA 및 엔도뉴클레아제를 유효성분으로 포함하는 암 치료용 약학적 조성물에 관한 것이다. 본 발명은 DNA와 RNA가 특이적으로 결합하는 성질에 기초하여 암세포를 특이적으로 치료함으로써 환자 또는 세포 유형의 필요에 따라 맞춤형될 수 있다. 본 발명에 따른 CRISPR PLUS 시스템의 뉴클레아제 활성은 crRNA와 암세포에 특이적으로 존재하는 유전자와의 결합에 의해 활성화될 수 있다. 따라서, 본 발명의 암 치료 효과는 현재 개발된 다른 항암제보다 더욱 특이적이다.

**대표도** - 도26



(52) CPC특허분류

**A61P 35/00** (2018.01)

(72) 발명자

**박영훈**

서울특별시 관악구 낙성대역4길 12-22, 105호(봉천동)

**임정학**

서울특별시 관악구 관악로14길 23, 303호 (봉천동, 현대원룸)

**김동욱**

서울특별시 관악구 남현7길 3, 806호 (남현동, 예성그린캐슬2차아파트)

**인성용**

서울특별시 관악구 낙성대로 38, 1층 (봉천동)

**박종진**

미국 47906 인디애나주 웨스트라피엣시 코트하우스 드라이브 이스트 3010 아파트 1비

(56) 선행기술조사문헌

KR1020180020929 A\*

US20160186152 A1\*

US20170145405 A1\*

US20170247690 A1\*

WO2018009525 A1\*

\*는 심사관에 의하여 인용된 문헌

**명세서**

**청구범위**

**청구항 1**

삭제

**청구항 2**

삭제

**청구항 3**

삭제

**청구항 4**

삭제

**청구항 5**

삭제

**청구항 6**

삭제

**청구항 7**

삭제

**청구항 8**

삭제

**청구항 9**

삭제

**청구항 10**

삭제

**청구항 11**

삭제

**청구항 12**

삭제

**청구항 13**

암세포에 특이적으로 존재하는 핵산에 상보적으로 결합하는 폴리뉴클레오티드 및 엔도뉴클레아제와 엑소뉴클레아제가 결합된 융합단백질을 유효성분으로 포함하는 종양 세포 살해용 조성물로서,

상기 폴리뉴클레오티드는 crRNA 또는 gRNA 이며,

상기 엔도뉴클레아제는 Cas9이며,

상기 엑소뉴클레아제는 RecJ인 것인, 종양 세포 살해용 조성물.

**청구항 14**

제13항에 있어서,

상기 엔도뉴클레아제와 엑소뉴클레아제는 링커를 통해 결합된 것인, 종양 세포 살해용 조성물.

**청구항 15**

제13항에 있어서,

상기 암세포에 특이적으로 존재하는 핵산은 단일염기다형성(single nucleotide polymorphism, SNP), 유전자복제수변이(copy number variation, CNV), 구조적 변이(structural variations, SV), 유전자의 삽입(insertion) 또는 유전자의 결실(deletion)의 특징을 갖는 것인, 종양 세포 살해용 조성물.

**청구항 16**

제15항에 있어서,

상기 구조적 변이는 역위(inversion), 전좌(translocation) 또는 반복염기서열의 과다증폭(short nucleotide repeat expansion)의 특징을 갖는 것인, 종양 세포 살해용 조성물.

**청구항 17**

제13항에 있어서,

상기 암세포에 특이적으로 존재하는 핵산은 유전자복제수변이(CNV)가 적어도 4인 유전자에서 선택되는 것인, 종양 세포 살해용 조성물.

**청구항 18**

제17항에 있어서,

상기 암세포에 특이적으로 존재하는 핵산은 유전자복제수변이(CNV)가 적어도 7인 유전자에서 선택되는 것인, 종양 세포 살해용 조성물.

**청구항 19**

삭제

**청구항 20**

제13항에 있어서,

상기 암세포에 특이적으로 존재하는 핵산은 p53, PTEN, APC, MSH2, HBV, HCV 및 EGFR로 구성된 군에서 선택되는 어느 하나의 유전자의 돌연변이인 것인, 종양 세포 살해용 조성물.

**청구항 21**

삭제

**청구항 22**

제13항 내지 제18항 및 제20항 중 어느 한 항의 조성물을 포함하는 암 치료용 약학적 조성물.

**청구항 23**

제22항에 있어서,

상기 암은 방광암, 골암, 혈액암, 유방암, 흑색종양, 갑상선암, 부갑상선암, 골수암, 직장암, 인후암, 후두암, 폐암, 식도암, 췌장암, 대장암, 위암, 설암, 피부암, 뇌종양, 자궁암, 두부 또는 경부암, 담낭암, 구강암, 결장암, 항문 부근암, 중추신경계 종양, 간암 및 대장암으로 구성된 군으로부터 선택되는 어느 하나인 것인, 약학적 조성물.

**청구항 24**

암세포에 특이적으로 존재하는 핵산 서열에 상보적으로 결합하는 폴리뉴클레오티드, 엑소뉴클레아제 및 엔도뉴클레아제를 코딩하는 폴리뉴클레오티드가 적재된 벡터를 유효성분으로 포함하는 종양 세포 살해용 조성물로서, 상기 폴리뉴클레오티드는 crRNA 또는 gRNA 이며, 상기 엔도뉴클레아제는 Cas9이며, 상기 엑소뉴클레아제는 RecJ인 것인, 종양 세포 살해용 조성물.

**청구항 25**

제24항에 있어서, 상기 벡터는 바이러스 벡터 또는 플라스미드 벡터인 것인, 종양 세포 살해용 조성물.

**청구항 26**

삭제

**청구항 27**

제13항 내지 제18항, 제20항, 제24항 및 제25항 중 어느 하나의 조성물을 암에 걸린 인간을 제외한 동물에 투여하는 단계를 포함하는 암을 치료하는 방법.

**청구항 28**

제27항에 있어서, 상기 투여 경로는 종양내 투여, 정맥내, 근육내, 피내, 피하, 복강내, 소동맥내, 심실내, 병변내, 척추강내, 국소 및 이들의 조합으로 구성된 군으로부터 선택되는 어느 하나인 것을 특징으로 하는, 암을 치료하는 방법.

**발명의 설명**

**기술 분야**

[0001] 본 발명은 가이드 RNA 및 엔도뉴클레아제를 유효성분으로 포함하는 항암용 조성물에 관한 것이다.

**배경 기술**

[0002] 항암 치료 기술의 혁신적인 발전에도 불구하고 암은 여전히 인간에게 가장 위협적인 질병이다. 암은 다양한 발암물질에 의해 촉발될 수 있으며, 염색체 구조나 DNA 염기 서열상의 변이에 의해 유발될 수 있는 질병이다. 암의 가장 두드러진 특징 중 하나는 끊임없는 세포 증식이다. 현재 가장 많이 사용되는 항암 치료법은 세포를 효율적으로 사멸시키는 방사선이나, 특정 암세포를 타겟하는 화합물 또는 항체를 이용한 치료법이 있다. 그러나, 화학/방사선 치료법을 포함한 항암 1차 치료법은 머리카락과 면역 세포와 같이 체내에서 증식성이 높은 정상 세포를 함께 사멸시킴으로써 환자에게 심각한 부작용과 고통을 유발할 수 있다. 따라서, 체내에 있는 암세포만을 선택적으로 사멸시킬 수 있는 항암제의 개발이 요구되고 있는 실정이다.

[0003] 이러한 니즈에 부합하여 표적 치료 항암제에 대한 연구가 활발하게 진행되고 있다. 표적 치료 항암제는 암 발병과 관련된 특정 단백질이나 암 발병과 관련된 경로를 조절함으로써 암을 치료하는 항암제이다. 암세포에 특이적인 표적을 확인하기 위해 암세포의 총 단백질량과 정상 세포의 총 단백질량을 비교함으로써 표적을 확인할 수 있다. 즉, 암세포에 특이적으로 존재하거나, 암세포에 더 풍부한 단백질은 잠재적 표적이 될 수 있다. 표적 단백질의 예로는 인간 표피성장인자 수용체 2 단백질(HER-2)이 있다. HER-2를 과발현하는 유방암 및 위암을 치료하기 위해, 트라스투즈맵(Herceptin®)을 포함한 HER-2에 대한 항체를 이용한 여러 가지 표적 치료제가 개발되었다. 또 다른 접근법은, 암의 진행을 유발하는 돌연변이 단백질을 타겟으로 하는 것이다. 예를 들어, 세포 증식 신호 단백질 *BRCA1* 또는 *BRAF*는 많은 유방암 및 흑색종에서 각각 변형된 형태로 존재한다. 많은 표적 치료제가 이러한 돌연변이 형태를 타겟으로 하여 개발되었으며, 이렇게 개발된 표적 치료제는 수술이 불가능하거나 전이성인 암 환자를 치료할 수 있도록 승인되었다.

[0004] 최근 개발된 표적 치료제 및 면역치료제는 암세포에서 특이적으로 발현되는 바이오마커 단백질을 이용하였다. 그러나 암세포에 특이적으로 존재하는 바이오마커 단백질과 항암제 사이의 상호 결합의 강도는 특이적이지 않아 간혹 부작용이 나타나기도 한다. 또한, 표적 치료제 및 면역치료제는 전통적인 화합물 항암제에 비해 독성이 적음에도 여전히 많은 부작용이 보고되었다.

[0005] 따라서, 항암 치료 분야에서 체내 암세포를 특이적으로 타겟팅하는 항암 치료제의 개발이 여전히 요구되고 있는 실정이다. 특히, 암세포를 차별화 하는 가장 큰 특징인 DNA 염기 서열상의 차이를 이용한 항암제의 개발은 인류의 오랜 숙원이었다.

**선행기술문헌**

**비특허문헌**

[0006] (비특허문헌 0001) Praveen Sridhar and Fabio Petrocca, Regional Delivery of Chimeric Antigen Receptor(CAR) T-Cells for Cancer Therapy. Cancers 2017, 9, 92.

**발명의 내용**

**해결하려는 과제**

[0007] 암세포에 존재하는 염색체 구조 이상이나 정상세포와 상이한 암세포 내의 염기서열은 암세포와 정상세포를 구분하는 중요한 판단 기준이 될 수 있다. 따라서, 이러한 염색체나 염기서열의 차이점은 암세포 치료를 위한 중요한 표적이 될 수 있다. 따라서, 본 발명의 목적은 암세포에 특이적으로 존재하는 서열을 타겟팅하여 항암효과를 나타내는 암 치료용 약학적 조성물을 제공하는 것이다.

**과제의 해결 수단**

[0008] 상기 목적을 달성하기 위해, 암세포에 특이적으로 존재하는 핵산 서열에 상보적으로 결합하는 폴리뉴클레오티드 및 뉴클레아제를 유효성분으로 포함하는 종양 세포 살해용 조성물을 제공한다.

[0009] 또한, 상기 조성물을 포함하는 암 치료용 약학적 조성물을 제공한다.

[0010] 또한, 암세포에 특이적으로 존재하는 핵산 서열에 상보적으로 결합하는 폴리뉴클레오티드, 엔도뉴클레아제 및 엑소뉴클레아제를 코딩하는 폴리뉴클레오티드가 적재된 벡터를 유효성분으로 포함하는 종양 세포 살해용 조성물을 제공한다.

[0011] 또한, 본 발명의 조성물을 암에 걸린 개체에 투여하는 단계를 포함하는 암을 치료하는 방법을 제공한다.

**발명의 효과**

[0012] 본 발명의 약학 조성물은 DNA와 RNA 간의 높은 특이성에 기반한 약물로서, 암세포만을 특이적으로 타겟팅하고 살해할 수 있으므로, 환자별, 암종별로 맞춤형될 수 있다. 특히, 암세포에만 존재하는 단일염기다형성(SNP) 및/또는 유전자복제수변이(CNV)를 갖는 유전자를 선별함으로써 환자의 암세포만을 효율적으로 살해할 수 있다. 또한, 정상세포에는 타겟팅하는 유전자가 존재하지 않아 암세포만을 효율적으로 제거할 수 있으므로 종래 항암제에 비해 안정성이 우수하다. 특히, 크리스퍼 연관 단백질에 엑소뉴클레아제 중 RecJ가 결합된 융합 단백질을 이용할 경우 더욱 우수한 항암제를 제공할 수 있다.

**도면의 간단한 설명**

[0013] 도 1은 CRISPR 연관 단백질 중 crRNA 가이드된 타겟 부위 결합에 의해 활성화되어 비특이적 엑소뉴클레아제 기능이 활성화되어 타겟 암세포의 핵산을 분해하는 과정을 모식도로 나타낸 것이다.

도 2는 일 실시예인 crRNA, CRISPR 연관 단백질 및/또는 엑소뉴클레아제가 결합된 조성물(이하 CRISPR PLUS 시스템)이 암세포에서 특이적으로 활성화되어 항암제로 작용하는 것을 나타내는 모식도이다.

도 3은 일 실시예인 CRISPR Cas12a 단백질이 서열 특이적 엔도뉴클레아제 효소 활성화에 의존적으로 비특이적 엑소뉴클레아제 활성을 가질 수 있다는 것을 확인한 도면이다.

도 4는 인간 유래 암세포인 HEK293(도 4a)과 HeLa(도 4b)에 CRISPR Cas12a(Cpf1) 뉴클레아제, crRNA 혹은 두 분자의 결합체(RNP complex)를 각각 형질감염시킨 후, 24, 48, 72시간 후에 세포의 생존도를 측정된 결과를 나타낸 것이다.

도 5는 pulse only, EGFR\_WT 및 EGFR 돌연변이를 가지는 HCC827 세포주 실험군을 이용하여, EGFR 돌연변이에 특이적으로 결합하는 가이드 RNA 및 Cas9에 의한 세포 사멸 유도를 확인한 것이다.

도 6은 도 5에서 실험한 pulse only, EGFR\_WT 및 EGFR 돌연변이를 가지는 HCC827 세포주 실험군(EGFR\_E2)에서 생존한 세포수를 측정된 것이다.

도 7은 폐암 세포주인 H1299에서 타겟 유전자인 CCR5, HPRT1, MT2, SMIM11, GNPDA2, SLC15A5 및 KCNE2에 상보적인 가이드 RNA에 의한 폐암 세포의 사멸을 확인한 것이다. 이때, 가이드 RNA와 Cas9을 코딩하는 핵산을 도입하기 위하여 lipofection을 수행하였다. NT1은 폐암 세포에 일치하는 서열이 없는 가이드 RNA를 의미하며 음성 대조군으로 이용되었다. 특히, MT2, SMIM11, GNPDA2, SLC15A5 및 KCNE2에 상보적으로 결합 가능한 가이드 RNA는 폐암 세포를 효율적으로 사멸시킬 수 있음을 확인하였다.

도 8은 도 7의 나타낸 세포 사멸 결과를 확인하기 위하여 살아있는 세포 수를 측정된 것이다.

도 9는 폐암 세포주인 H1299에서 타겟 유전자인 CCR5, HPRT1, MT2, GNPDA2, SLC15A5 및 KCNE2에 상보적인 가이드 RNA에 의한 폐암 세포의 사멸을 확인한 것이다. 이때, DNA의 도입 없이 리포펙타민(Lipo)만 수행한 대조군을 포함시켰다.

도 10은 NucBlue Live ReadyProbes Reagent 및 Propidium Iodide ReadyProbes Reagent를 이용하여 NT1 대조군 및 GNPDA2 실험군의 살아있는 세포들을 현미경을 이미징한 것이다. 이때, NucBlue Live ReadyProbes Reagent는 살아있는 세포와 죽은 세포를 모두 염색하는 파란색 형광 염색 물질이다. 또한, Propidium Iodide ReadyProbes Reagent는 죽은 세포만 염색하는 빨간색 형광 염색 물질이다.

도 11은 폐암 세포주 H1563에서 타겟 유전자인 CCR5, HPRT1, MT2, IRX1 및 ADAMTS16에 상보적인 가이드 RNA에 의한 폐암 세포의 사멸을 확인하기 위해, 살아있는 세포 수를 측정된 것이다.

도 12는 폐암 세포주 A549에서 타겟 유전자인 HPRT1, CCR5 및 MT2에 상보적인 가이드 RNA에 의한 폐암 세포의 사멸을 확인하기 위해, 시간에 따른 살아있는 세포 수를 발광도를 이용하여 측정된 것이다.

도 13은 폐암 세포주 A549에서 타겟 유전자인 HPRT1, CCR5 및 MT2에 상보적인 가이드 RNA에 의한 폐암 세포의 사멸을 확인하기 위해, 살아있는 세포 수를 측정된 것이다.

도 14는 도 13의 결과를 현미경으로 관찰한 것이다.

도 15는 폐암 세포주 A549에서 타겟 유전자인 MT2, CD68, DACH2, HERC2P2 및 SHBG에 상보적인 가이드 RNA에 의한 폐암 세포의 사멸을 확인하기 위해 시간에 따른 살아있는 세포 수를 발광도를 이용하여 측정된 것이다.

도 16은 유방암 세포주 SKBR3에서 타겟 유전자인 MT2, ERBB2 및 KRT16에 상보적인 가이드 RNA에 의한 유방암 세포의 사멸을 확인하기 위해 시간에 따른 살아있는 세포 수를 발광도를 이용하여 측정된 것이다.

도 17은 유방암 세포주 SKBR3에서 타겟 유전자인 MT2, ERBB2 및 KRT16에 상보적인 가이드 RNA에 의한 유방암 세포의 사멸을 확인하기 위해 시간에 따른 살아있는 세포 수를 측정된 것이다.

도 18은 자궁경부암 세포주 HeLa에서 타겟 유전자인 CCR5, MT2 및 PRDM9에 상보적인 가이드 RNA에 의한 자궁경부암 세포의 사멸을 확인하기 위해 시간에 따른 살아있는 세포 수를 측정된 것이다. 이때, 각 유전자의 CNV는 CCR5는 2, MT2는 100 이상, PRDM9는 8 이상이었다.

도 19는 자궁경부암 세포주 HeLa에서 타겟 유전자인 CCR5, MT2, PRDM9 및 HPV\_1에 상보적인 가이드 RNA에 의한 자궁경부암 세포의 사멸을 확인하기 위해 시간에 따른 살아있는 세포 수를 측정된 것이다. 이때, 각 유전자의 CNV는 CCR5는 2, MT2는 100 이상, PRDM9는 8 이상, HPV\_1은 30이었다.

도 20a는 자궁경부암 세포 HeLa에서 타겟 유전자인 CCR5, MT2, HPV\_1 및 PRDM9에 상보적인 가이드 RNA에 의한 자궁경부암 세포의 사멸을 확인하기 위해 시간에 따른 살아있는 세포 수를 발광도를 이용하여 측정된 것이다.

도 20b는 자궁경부암 세포 HeLa에서 타겟 유전자인 CCR5, MT2 및 HPV\_1에 상보적인 가이드 RNA에 의한 자궁경부암 세포의 사멸을 확인하기 위해 세포 수를 발광도를 이용하여 측정된 것이다. 이때, 인간 유전체에 없는 영역

을 타겟하는 NT 서열인 NT1, NT2 및 NT3을 음성 대조군으로 포함시켰다.

도 21은 대장암 세포주 HT29에서 타겟 유전자인 CCR5, HPRT1, MT2, TRAPPC, LINC00536, TRPS1 및 CDK8에 상보적인 가이드 RNA에 의한 대장암 세포의 사멸을 확인하기 위해 살아있는 세포 수를 측정하는 것이다.

도 22는 대장암 세포 HT29에서 타겟 유전자인 MT2, CDK8, LINC00536, TRPS1 및 TRAPPC9에 상보적인 가이드 RNA에 의한 대장암 세포의 사멸을 확인하기 위해 살아있는 세포 수를 발광도를 이용하여 측정하는 것이다.

도 23은 폐암 세포주인 H1299에서 효과적인 타겟 유전자에 상보적인 가이드 RNA가 다른 폐암 세포주인 H1563에서도 효과가 있는지 확인하였다. 이를 위해, 타겟 유전자 SMIM11, GNPDA2, SLS15A5 및 KCNE2에 상보적인 가이드 RNA에 의한 H1563 폐암 세포의 사멸을 확인하기 위해 살아있는 세포 수를 측정하는 것이다.

도 24는 도 23의 실험 결과를 현미경으로 관찰한 것이다.

도 25는 폐암 세포주인 H1299에서 CNV에 의한 효과를 확인하기 위하여, CCR5, KCNE2, GNPDA2, SMIM11 및 SLS15A5에 상보적인 가이드 RNA에 의한 폐암 세포의 사멸을 확인한 것이다. 이때, AnnV reagent를 넣어 측정하였으므로, 사멸한 세포에서 높은 발광이 검출되었다.

도 26은 엔도뉴클레아제 기능이 있는 CRISPR 단백질인 Cas9과 엑소뉴클레아제 기능이 있는 단백질인 RecJ를 융합한 단백질의 암 세포 살상 효능을 측정하기 위하여, 폐암 세포주 HCC827 내의 EGFR 돌연변이 및 CCR5에 상보적인 가이드 RNA와 Cas9 및 Cas9-RecJ(서열번호 87)를 이용하였다. 그 결과, 엑소뉴클레아제가 결합된 융합단백질에서 폐암 세포의 살상 능력이 현저히 증가됨을 확인하였다.

도 27은 도 26의 실험 결과를 현미경으로 관찰한 것이다.

도 28은 전달 시스템에 따른 효과를 확인하기 위한 실험으로, 가이드 RNA와 엔도뉴클레아제 단백질인 Cas9이 결합된 형태인 RNP(ribonucleoprotein)가 세포 사멸 효능이 있는지를 확인한 것이다. 구체적으로, 폐암 세포주인 HCC827에 EGFR 돌연변이에 상보적으로 결합 가능한 RNA와 Cas9을 혼합한 RNP가 효율적으로 폐암 세포를 사멸시키는 것을 확인하였다. 반면, 음성 대조군으로 사용한 Cas9 단백질 및 가이드 RNA는 세포를 사멸시키지 못함을 확인하였다.

도 29는 전달 시스템에 따른 효과를 확인하기 위한 실험으로, RNP가 효과적으로 세포 사멸을 유도할 수 있는지를 확인한 것이다. 대조군으로는 sgRNA를 사용하였고, 실험군으로는 MT2에 상보적인 RNA 및 Cas9로 이루어진 RNP를 이용하였다. 세포주는 폐암 세포주인 H1563을 이용하였다.

도 30은 전달 시스템에 따른 효과를 확인하기 위한 실험으로, RNP가 효과적으로 세포 사멸을 유도할 수 있는지를 확인한 것이다. 대조군으로는 sgRNA 및 Cas9 단백질을 사용하였고, 실험군으로는 CCR5, GNPDA2 및 SMIM11에 상보적인 RNA 및 Cas9로 이루어진 RNP를 이용하였다. 세포주는 폐암 세포주인 H1299를 이용하였다.

도 31은 전달 시스템에 따른 효과를 확인하기 위한 실험으로, RNP가 효과적으로 세포 사멸을 유도할 수 있는지를 확인한 것이다. 대조군으로는 sgRNA 및 Cas9 단백질을 사용하였고, 실험군으로는 CCR5 및 MT2에 상보적인 RNA 및 Cas9로 이루어진 RNP를 이용하였다. 세포주는 폐암 세포주인 H1299를 이용하였다.

도 32는 도 30의 실험 결과를 현미경으로 관찰한 것이다.

도 33은 타겟 유전자인 CCR5 및 EGFR\_E2에 상보적인 가이드 RNA와 Cas12a를 이용하여 세포 사멸 유도를 확인한 것이다. 세포주는 폐암 세포주인 HCC827을 이용하였다.

도 34는 도 33의 실험 결과를 현미경으로 관찰한 것이다.

**발명을 실시하기 위한 구체적인 내용**

[0014] 본 발명은 일 측면으로, 암세포에 특이적으로 존재하는 핵산에 상보적으로 결합하는 폴리뉴클레오티드 및 뉴클레아제를 유효성분으로 포함하는 종양 세포 살해용 조성물을 제공한다.

[0015] 이때, 상기 폴리뉴클레오티드는 crRNA 또는 gRNA 일 수 있다. crRNA는 CRISPR RNA라 불리운다. 또한, gRNA는 가이드 RNA를 지칭한다. crRNA 및 gRNA는 단일 가닥 RNA(single strand RNA) 일 수 있다. 또한, crRNA는 tracrRNA와 결합하여 크리스퍼 연관 단백질을 작동하게 할 수 있으며, crRNA는 tracrRNA와 결합된 형태로 이용될 수 있다. 이때, crRNA는 타겟이 되는 암세포에 특이적으로 존재하는 유전자 서열에 상보적인 서열을 가질 수 있다. 또한, gRNA도 타겟이 되는 암세포에 특이적으로 존재하는 유전자 서열과 결합하여 크리스퍼 연관 단백질

이 활성을 나타내게 할 수 있다. 상기 crRNA 또는 gRNA는 15 내지 40개의 핵산으로 구성된 RNA 일 수 있다. 이때, 상기 폴리뉴클레오티드는 18 내지 30개, 20 내지 25개의 핵산으로 구성될 수 있다. 일 구체예로 crRNA 또는 gRNA는 20개의 핵산으로 구성될 수 있다. 또한, crRNA 또는 gRNA는 Cas9과 같은 크리스퍼 연관 단백질이 활성을 갖게 하기 위해 3'에 추가적인 서열을 포함할 수 있다. 일 실시예로 상기 gRNA는 서열번호 88 내지 서열번호 129로 표시되는 DNA에 의해 생산되는 RNA 일 수 있다.

[0016] 본 명세서에서 사용된 용어, “뉴클레아제”는 엔도뉴클레아제 일 수 있다. 이때, 상기 뉴클레아제는 CRISPR 연관 단백질일 수 있다. 본 발명에서 사용된 용어, CRISPR 연관 단백질은 DNA 또는 RNA와 같은 핵산이 이중 가닥 혹은 단일 가닥을 가질 경우(dsDNA/RNA 및 ssDNA/RNA) 이를 인식하여 절단할 수 있는 효소이다. 구체적으로, 이들은 crRNA 또는 가이드 RNA와 결합된 이중가닥 혹은 단일가닥 핵산을 인식하여 이를 절단할 수 있다.

[0017] 본 발명의 뉴클레아제의 일 구체예는, crRNA가 표적 부위와 결합된 것을 인식하여 엔도뉴클레아제 기능이 활성화되는 것일 수 있다. 또한, 엔도뉴클레아제 기능이 활성화됨에 따라, 이중가닥 및/또는 단일가닥 DNA 및/또는 RNA를 비특이적 절단할 수 있는 엑소뉴클레아제 활성을 가지는 것일 수 있다. 또한, 상기 Cas12a와 같은 크리스퍼 연관 단백질은 일단 활성화되면, 비특이적 엑소뉴클레아제 활성이 나타날 수 있다. 이러한 경우 DNA 및 RNA를 비특이적으로 절단할 수 있다.

[0018] 따라서, 본 발명의 일 구체예의 조성물은 암세포에 존재하는 특정 표적 부위에 crRNA 또는 gRNA가 결합되고, 이에 의해 활성화되는 비특이적 뉴클레아제에 의해 암세포를 특이적으로 사멸시킬 수 있다. 따라서, 본 발명의 조성물은 암세포만을 특이적으로 살해할 수 있으므로, 항암제로 활용될 수 있다.

[0019] 구체적으로, 상기 CRISPR 연관 단백질의 일 구체예는 Cas1, Cas1B, Cas2, Cas3, Cas4, Cas5, Cas6, Cas7, Cas8, Cas9, Cas10, Cas12a, Cas12b, Cas12c, Cas12d, Cas12e, Cas12g, Cas12h, Cas12i, Cas13a, Cas13b, Cas13c, Cas13d, Cas14, Csy1, Csy2, Csy3, Cse1, Cse2, Csc1, Csc2, Csa5, Csn2, CsMT2, Csm3, Csm4, Csm5, Csm6, Cmr1, Cmr3, Cmr4, Cmr5, Cmr6, Csb1, Csb2, Csb3, Csx17, Csx14, Csx10, Csx16, CsaX, Csx3, Csx1, Csx15, Csf1, Csf2, Csf3 및 Csf4로 구성된 군으로부터 선택되는 어느 하나의 뉴클레아제일 수 있다. 상기 CRISPR 연관 단백질의 일 실시예는 Cas9, Cas12a(Cpf1), Cas13a(C2c2)의 뉴클레아제 단백질일 수 있다.

[0020] 이때, 상기 CRISPR 연관 단백질의 일 구체예로 상기 Cas9 단백질은 서열번호 20의 아미노산 서열을 가질 수 있다. 이때, 상기 Cas9 단백질은 서열번호 19의 서열을 갖는 핵산에 의해 코딩되는 것일 수 있다. 또한, 상기 Cpf1 단백질은 서열번호 22의 아미노산 서열을 가질 수 있다. 이때, 상기 Cpf1 단백질은 서열번호 21의 서열을 갖는 핵산에 의해 코딩되는 것일 수 있다. 또한, 상기 C2c2 단백질은 서열번호 24의 아미노산 서열을 가질 수 있다. 이때, 상기 C2c2 단백질은 서열번호 23의 서열을 갖는 핵산에 의해 코딩되는 것일 수 있다.

[0021] 본 명세서에서 사용된 용어, “암세포에 특이적으로 존재하는 핵산”은 정상세포와 차별화되는 암세포에만 존재하는 핵산을 의미한다. 즉, 정상 세포와 상이한 서열을 의미할 수도 있으며, 적어도 1개의 핵산이 상이할 수 있다. 또한, 유전자의 일부가 치환, 결실된 것일 수 있다. 또한, 특정 시퀀스가 반복된 서열을 가질 수 있다. 이때, 반복되는 서열은 세포내에 존재하였던 서열일 수 있으며, 외부에서 삽입된 서열일 수 있다.

[0022] 일 구체예로, 암세포에 특이적으로 존재하는 핵산은 암세포에 특이적으로 존재하는 핵산은 단일염기다형성(single nucleotide polymorphism, SNP), 유전자복제수변이(copy number variation, CNV), 구조적 변이(structural variations, SV), 유전자의 삽입(insertion) 또는 유전자의 결실(deletion)의 특징을 갖는 것일 수 있다.

[0023] 구체적으로, 암세포에 특이적으로 존재하는 서열은 암세포에 존재하는 SNP일 수 있다. 상기 서열을 갖는 암세포 내에 존재하는 타겟 DNA와 타겟 RNA에 상보적인 서열을 갖는 crRNA 또는 가이드 RNA는 서로 특이적으로 결합할 수 있다. 따라서, 암세포 내에 특이적으로 존재하는 핵산은 본원 발명에 개시된 종양 세포 살해용 조성물의 특이성을 부여할 수 있다. 특히, 암세포에 특이적으로 존재하는 핵산은 여러 암조직의 유전체 서열분석으로부터 암세포에만 존재하는 특이적 SNP를 찾아서 이를 이용해서 crRNA 또는 gRNA를 제작할 수 있다. 따라서, 암세포 특이적 독성을 나타내는 방식이므로, 환자 맞춤형 항암치료제 개발이 가능할 수 있다.

[0024] 또한, 암세포에 특이적으로 존재하는 서열은 암세포에 존재하는 유전자복제수변이(CNV)일 수 있다. CNV는 유전체의 일부분들이 반복되어 나타나는 변이를 의미한다. 반복되는 유전자의 수는 암종류나 개체별로 서로 다르게 나타날 수 있다. 통상적으로, CNV는 2개의 카피 수(copy number)로 존재하는 일반적인 유전자들과는 달리 결실, 증폭되는 등 인간의 표준 참조 게놈과 비교해 반복되는 서열의 숫자의 차이를 보이는 핵산 조각을 의미한다. 일 구체예로, 정상세포에서는 CNV가 2이나, 암세포에서 CNV가 4 이상인 유전자는 종양 세포 살해용 조성물의 특이

성을 부여할 수 있다. 이때, CNV는 4, 8, 10, 12, 14, 16, 18, 20, 24, 30, 40, 50, 60, 70, 80, 90 또는 100 이상일 수 있다. 구체적으로 카피 수가 7개 이상인 경우 CNV로 판정할 수 있다. 암세포주별 유전자에 따른 카피 수의 일 구체예를 하기 표 1에 나타내었다.

표 1

Type	Cell line	Gene name	gRNA	Copy number
Lung	HCC827	EGFR	서열번호 88	>16
		VSTMT2A	서열번호 89	>13
		KIF5A	서열번호 90	>14
	H1563	IRX1	서열번호 91	>8
		ADAMTS16	서열번호 92	>7
	H1299	GNPDA2	서열번호 93	>12
		KCNE2	서열번호 94	>40
		SLC15A5	서열번호 95	>12
		SMIM11	서열번호 96	>40
	A549	DACH2	서열번호 97	>18
		HERC2P2	서열번호 98	>8
		CD68	서열번호 99	>10
		SHBG	서열번호 100	>9
Breast	SKBR3	ERBB2	서열번호 101-108	>33
		KRT16	서열번호 109, 110	>8
	Colon	HT29	LINC00536	서열번호 111
		TRPS1	서열번호 112	>8
		CDK8	서열번호 113	>18
		TRAPPC9	서열번호 114	>13
		HERC2P2	서열번호 98	>8
Pancreas	Capan2	SIRPB1	서열번호 115	>24
		MRC1	서열번호 116	>18
		ATP11A	서열번호 117	>8
		POTEB	서열번호 118	>7
		HERC2P2	서열번호 98	>7
Cervix	HeLa	PRDM9	서열번호 119	>8
		CDKN2B	서열번호 120	>10
		HPV	서열번호 121-124	30
ETC		LINE2(mt2)	서열번호 125	>100
		CCR5	서열번호 126	2
		HPRT1	서열번호 127	2
		NT	서열번호 128-130	0

상기 CNV는 암(cancer), 지적 장애(intellectual disability), 뇌전증(epilepsy), 조현병(schizophrenia), 소아비만(obesity) 등과 같은 인간의 질병과 연관성이 있는 매우 중요한 변이 유형 중 하나이다. 대부분의 암세포주(cancer cell line)는 CNV를 가지고 있으며, CNV 내에 존재하는 타겟 서열에 상보적인 서열을 갖는 가이드 RNA는 서로 특이적으로 결합한다. 따라서, 암세포 내에 특이적으로 존재하는 CNV는 본원 발명에 개시된 항암제의 특이성을 부여할 수 있다. 특히, CNV의 수가 높을수록 크리스퍼 연관 단백질에 의해 절단되는 유전자가 많아, 암세포의 핵이 크게 손상될 수 있다. 특히, 암세포에 특이적으로 존재하는 CNV의 데이터는 마이크로어레이(microarray), FISH(fluorescent in situ hybridization) 기법 등을 통해 쉽게 확인할 수 있으며, 이를 이용해서 crRNA 또는 gRNA를 제작할 수 있다. 암세포의 CNV 내에 존재하는 특정 서열을 타겟하는 crRNA 또는 gRNA와 CRISPR 연관 단백질을 함께 처리할 경우 암세포 특이적으로 세포 사멸이 유도될 수 있다. 따라서, 암세포 특이적 독성을 나타내는 방식이므로, 환자 맞춤형 항암치료제 개발이 가능할 수 있다.

본 발명의 일 실시예에서는, CNV의 예상 copy number(N)는 copy number value(V)로부터 하기의 식을 이용하여 계산될 수 있다:  $N=2x2^V$ .

- [0028] 본 발명의 일 실시예에서는 암세포에서만 특이적으로 높은 CNV를 가지는 유전자에 상보적이 결합을 하는 폴리뉴클레오티드를 크리스퍼 연관 단백질 혹은 크리스퍼 연관 단백질과 엑소뉴클레아제 단백질을 융합한 크리스퍼 플러스 단백질을 함께 이용될 경우, 암세포만을 효과적으로 사멸시킬 수 있음을 확인하였다. 또한, 암세포에만 특이적으로 존재하는 유전자 돌연변이 중에서 CNV가 높은 유전자를 선별할 경우 더욱 효과적으로 암세포를 사멸시킬 수 있음을 확인하였다.
- [0029] 또한, 그 외에도 암세포에 특이적으로 존재하는 서열은 암세포에 존재하는 구조적 변이(structural variations, SV)일 수 있으며, 상기 SV는 역위(inversion), 전좌(translocation), 반복염기서열의 과다증폭(short nucleotide repeat expansion) 등일 수 있다.
- [0030] 상기 역위는 유전자의 일부분이 뒤집어진 돌연변이로서, 혈우병, 폐암 등과 같은 질병과 연관성이 있는 변이 유형 중 하나이다. 상기 전좌는 염색체의 일부분이 떨어져나와 다른 염색체에 결합하는 돌연변이이다. 상기 반복염기서열의 과다증폭은 같은 서열이 계속 반복되어 과다증폭된 돌연변이이다.
- [0031] 대부분의 암세포주는 유전자 역위, 전좌, 반복염기서열의 과다증폭 등과 같은 SV를 가지고 있으며, SV가 일어난 서열 말단과 정상세포주 서열 말단 사이의 접합 영역 서열(junction 서열)은 해당 암세포주에만 존재하는 서열이다. 암세포 내에 존재하는 SV 접합 영역의 서열에 상보적인 서열을 갖는 가이드 RNA는 서로 특이적으로 결합한다. 따라서, 암세포 내에 특이적으로 존재하는 SV 접합 영역의 서열은 본원 발명에 개시된 항암제의 특이성을 부여할 수 있다.
- [0032] 특히, SV 접합 영역의 서열은 CNV와 다르게 정상세포에는 없는 서열이므로 암세포에 특이적이다. 따라서, 암세포 내에 존재하는 SV 접합 영역의 서열에 상보적으로 결합하는 폴리뉴클레오티드와 CRISPR 연관 단백질을 처리할 경우 특이적으로 암세포 사멸이 유도될 수 있다. 따라서, 특정 암세포에만 특이적으로 독성을 나타내는 방식이므로, 환자 맞춤형 항암치료제 개발이 가능할 수 있다.
- [0033] 또한, 암세포에 특이적으로 존재하는 서열은 유전자의 삽입(insertion) 또는 유전자의 결실(deletion)에 의해 생성될 수 있다. 이때, 상기 삽입 또는 결실되어 돌연변이가 일어난 서열에 상보적인 서열을 갖는 가이드 RNA는 효과적으로 암세포를 사멸시킬 수 있다. 따라서, 암세포 내에 특이적으로 존재하는 유전자의 삽입 또는 유전자의 결실은 본원 발명에 개시된 항암제의 특이성을 부여할 수 있다.
- [0034] 특히, 삽입 및 결실의 경우, SV 접합 영역의 서열과 동일하게 정상세포에는 없는 서열이므로 특이적이고, 암세포 내에 존재하는 삽입 및 결실 돌연변이 서열에 CRISPR 연관 단백질을 처리할 경우 특이적으로 세포 사멸이 유도될 수 있다. 그러나, 삽입 및 결실 돌연변이 서열을 타겟으로 하는 경우에는, Indel(insertion 및 deletion) 주위에 유전자가 인식할 수 있는 PAM(protospacer adjacent motif) 서열을 필요로할 수 있다.
- [0035] 또한, 삽입되는 유전자는 세포내에 존재하는 핵산 서열일 수도 있으나, 외부로부터 도입된 유전자 서열일 수 있다. 특히, 바이러스 감염 등에 의해 유발되는 암세포의 경우 바이러스 유전자가 세포내에 삽입될 수 있다. 이러한 경우 바이러스의 핵산 서열이 본 발명의 암세포에 특이적으로 존재하는 핵산으로 이용될 수 있다. 특히, 상기와 같이 삽입이 발생하는 암은 흔지 않지만, 정상세포에는 없는 서열을 타겟으로 하기 때문, 암세포에 특이적이다. 또한, viral 서열이 여러 카피(copy)로 통합되어 있을 경우, CNV가 높아 확실히 암세포의 사멸을 유도할 수 있다. 이러한 암의 일 구체예는 파필로마바이러스(papillomavirus)에 의해 유발되는 자궁경부암일 수 있다. 이러한, 외부에서 도입된 유전자를 타겟팅하는 gRNA의 일 실시예는 서열번호 121 내지 서열번호 124의 DNA에 의해 생성되는 gRNA일 수 있다.
- [0036] 또한, 암세포에 특이적인 유전자를 선별하고, 상기 유전자에 상보적인 폴리뉴클레오티드 서열을 선별하기 위하여, PAM(protospacer associated motif) 서열인 5'-NGG-3' 서열을 함께 고려할 수 있다. 일 구체예로, 암세포주 특이적인 서열 주변에 5'-NGG-3' 서열이 있을 경우, 3' 방향으로 20 뉴클레오타이드를 표적으로 지정할 수 있다. 또한, 표적 유전자의 선별시, 유전자의 삽입, 유전자의 결실 및 접합 영역 등의 서열 정보가 명확한 유전자 및 CNV의 카피수가 높은 유전자를 우선하여 선별할 수 있다. 또한, 표적 서열을 선별하기 위하여, 서열의 5'과 3' 말단에 G 또는 C 존재여부, 서열 전체의 GC contents(%)가 40 내지 60%에 포함되는지 여부 및 크리스퍼 연관 단백질인 엔도뉴클레아제가 절단하는 서열인 PAM의 3' 방향으로 3번째 - 4번째 염기 부분이 A 또는 T인지를 확인할 수 있다. 상기 서열의 5'과 3' 말단에 G 또는 C 존재하는 경우와 서열 전체의 GC contents(%)가 40 내지 60%에 포함되는 경우, sgRNA의 결합 친화성(binding affinity)이 높아질 수 있다. 특히, Streptococcus pyogenes Cas9(SpCas9)이 서열을 절단하는 부위인 PAM의 3' 방향으로 3번째 - 4번째 염기 부분이 A 또는 T인 경우, SpCas9의 절단 효율(cleavage efficiency)이 높아질 수 있다.

[0037] 상기 암의 일 구체예는 방광암, 골암, 혈액암, 유방암, 흑색종양, 갑상선암, 부갑상선암, 골수암, 직장암, 인후암, 후두암, 폐암, 식도암, 췌장암, 대장암, 위암, 설암, 피부암, 뇌종양, 자궁암, 두부 또는 경부암, 담낭암, 구강암, 결장암, 항문 부근암, 중추신경계 종양, 간암 및 대장암으로 구성된 군으로부터 선택되는 어느 하나 일 수 있다. 특히, 우리나라 5대 암으로 알려진 위암, 대장암, 간암, 폐암, 유방암 일 수 있다.

[0038] 상기 암에 존재하는 특이적으로 존재하는 핵산은 p53, PTEN, APC, MSH2, HBV, HCV 및 EGFR로 구성된 군에서 선택되는 어느 하나의 유전자의 돌연변이일 수 있으나, 이에 한정되는 것은 아니다. 구체적으로 위암의 경우 종양 억제 유전자로 알려진 p53과 PTEN의 돌연변이 일 수 있다. 또한, 대장암의 경우 APC와 MSH2 유전자의 돌연변이 일 수 있다. 또한, 간암은 HBV, HCV 바이러스의 감염이 주요 원인이므로, HBV 또는 HCV의 핵산이 타겟이 될 수 있다. 또한, 폐암은 EGFR 유전자의 돌연변이가 타겟이 될 수 있으며, 유방암은 BRCA1/2 유전자의 변이가 주요 타겟일 수 있다.

[0039] 상술한 바와 같이 암의 발생과 밀접하게 연관된 유전자의 돌연변이 및 바이러스 유전자는 암세포에 특이적으로 존재하는 핵산 서열로 활용되어, 상기 유전자는 crRNA 또는 gRNA의 제작을 위해 사용될 수 있다. 이때, 암세포에 특이적으로 존재하는 DNA의 SNP라면 어떤 것이라도 사용 가능하다. 암세포에 특이적으로 존재하는 핵산 서열의 일 구체예는 하기의 표 2에 기재된 서열일 수 있으나 이에 제한되는 것은 아니다.

표 2

유전자	암 종류	정상세포	암세포	단백질 변형
BRCA1 Exon 7	난소암/유방암	608:CAAAGTATGGGCTACAGAAACCGTGCCAAAAG(서열번호 26)	608:CAAAGTATGGGCTTCAGAAACCGTGCCAAAAG(서열번호 27)	p.Tyr130->Phe
BRCA1 Exon 10		1615:TGGGAAAACCTATCGGAAGAAGGCAAGCCTCC(서열번호 28)	1615:TGGGAAAACCTATCGGTAGAAGGCAAGCCTCC(서열번호 29)	p.Lys467->non-sense
BRCA1 Exon 11		3845:GGGGCCAAGAAA-TTAGAGTCCTCAGAAGAG(서열번호 30)	3845:GGGGCCAAGAAAATTAGAGTCCTCAGAAGAG(서열번호 31)	p.Leu1209->Ile
BRCA1 Exon 11		4260:ATGATGAAGAAAGAGGAACGGGCTTGAAGA(서열번호 32)	4260:ATGATGAAGAAAG--GAACGGGCTTGAAGA(서열번호 33)	p.Gly1348->Asn
BRCA1 Exon 11		3657:CATCTCAGGTTTGTCTGAGACACCTGATGACC(서열번호 34)	3657:CATCTCAGGTTTGTCT-AGACACCTGATGACC(서열번호 35)	p.Glu1148->Arg
BRCA1 Exon 15		7466:ATATACAGGATATGCGAATTAAGAAGAAACAAA(서열번호 36)	7466:ATATACAGGATATGTGAATTAAGAAGAAACAAA(서열번호 37)	p.Arg2494->Thr
TP53	위암	125:TAGGAGGCCGAGCTCTGTTGCTTCGAACTCCA(서열번호 38)	125:TAGGAGGCCGAGCTCT-TTGTCTCGAACTCCA(서열번호 39)	p.Leu20->Cys
MSH2	대장암	126:TGAGGAGGTTTCGACATGGCGGTGCAGCCGA(서열번호 40)	126:TGAGGAGGTTTCGACCTGGCGGTGCAGCCGA(서열번호 41)	p.Met1->Leu
EGFR	폐암	2137:AAAAAGATCAAAGTGCTGGCTCCGGTGCCTT(서열번호 42)	2137:AAAAAGATCAAAGTGCTGAGCTCCGGTGCCTT(서열번호 43)	p.Gly719->Ser
FGFR3	간암	1771:ATCCTCTCTGAAATCACTGAGCAGGAGAAAG(서열번호 44)	1771:ATCCTCTCTGAAATCACTGCGCAGGAGAAAG(서열번호 45)	p.Glu545->Ala

[0041] 이때, 상기 암세포에 특이적으로 존재하는 핵산 서열을 타겟팅하는 CRISPR RNA는 하나 이상의 crRNA 또는 gRNA 서열을 포함할 수 있다. 예를 들어, 난소암 또는 유방암에 존재하는 BRCA1의 엑손 10 또는 11을 동시에 타겟팅할 수 있는 crRNA 또는 gRNA를 사용할 수 있다. 또한, BRCA1의 엑손 11을 타겟팅 하는 두개 이상의 crRNA 또는 gRNA를 사용할 수 있다. 이와 같이 crRNA 또는 gRNA의 조합은 암 치료 목적 및 암의 종류에 따라 적합하게 선택될 수 있다. 즉, 서로 다른 gRNA를 선택하여 사용할 수 있다.

[0042] 상기 본 발명의 종양 살해용 조성물은 엑소뉴클레아제를 더 포함할 수 있다.

[0043] 본 명세서에서 사용된 용어, “엑소뉴클레아제”는 핵산 분자의 5' 말단 또는 3' 말단 방향으로부터 뉴클레오티드를 절단해나가는 효소이다. 따라서, 상기 엑소뉴클레아제는 핵산을 5'에서 3' 방향으로 분해하는 5'->3' 뉴클레아제일 수 있다. 또한, 상기 엑소뉴클레아제는 핵산을 3'에서 5' 방향으로 분해하는 3'->5' 뉴클레아제일 수 있다.

[0044] 구체적으로, 상기 5'->3' 뉴클레아제의 일 구체예는 대장균 유래의 RecE 또는 RecJ 일 수 있다. 또한, 박테리오파지 T5 유래의 T5일 수 있다. 또한, 상기 3'->5' 뉴클레아제의 일 구체예는 진핵세포 또는 원핵세포 유래의 Exo I 일 수 있다. 또한, 대장균 유래의 Exo III 일 수 있다. 또한, 인간 유래의 Trex1 또는 Trex2일 수 있다. 그 외에도 5'->3' 및 3'->5' 양방향 절단 활성이 있는 뉴클레아제로는 대장균 유래의 Exo VII 또는 RecBCD 일

수 있다. 또한, 일 구체예로 대장균 유래의 5'→3' 람다 엑소뉴클레아제 일 수 있다. 또한, 단일 사슬 DNA를 절단할 수 있는 *Vigna radiata* 유래의 Mungbean 일 수 있다.

- [0045] 이때, 상기 엑소뉴클레아제의 구체예로는 Exoribonuclease T, TREX2, TREX1, RecBCD, Exodeoxyribonuclease I, Exodeoxyribonuclease III, Mungbean exonuclease, RecE, RecJ, T5, Lambda exonuclease, Exonuclease VII small unit, Exonuclease VII large unit, Exo I, Exo III, Exo VII 및 Lexo로 구성된 군으로부터 선택되는 어느 하나일 수 있다.
- [0046] 구체적으로, 상기 엑소뉴클레아제는 Exoribonuclease T(서열번호 4), TREX2(서열번호 5), TREX1(서열번호 6), RecBCD\_RecB(서열번호 7), RecBCD\_RecC(서열번호 8), RecBCD\_RecD(서열번호 9), Exodeoxyribonuclease I(서열번호 10), Exodeoxyribonuclease III(서열번호 11), Mungbean exonuclease(서열번호 12), RecJ(서열번호 13), RecE(서열번호 14), T5(서열번호 15), Lambda exonuclease(서열번호 16), Exonuclease VII small unit(서열번호 17), 및 Exonuclease VII large unit(서열번호 18)으로 이루어진 군에서 선택되는 어느 하나일 수 있다.
- [0047] 본 발명은 또 다른 측면으로, 암세포에 특이적으로 존재하는 핵산에 상보적으로 결합하는 폴리뉴클레오티드 및 엔도뉴클레아제와 엑소뉴클레아제가 결합된 융합단백질을 유효성분으로 포함하는 종양 세포 살해용 조성물을 제공한다.
- [0048] 이때, "암세포에 특이적으로 존재하는 핵산", "암세포에 특이적으로 존재하는 핵산에 상보적으로 결합하는 "폴리뉴클레오티드", "엔도뉴클레아제" 및 "엑소뉴클레아제"는 상술한 바와 같다.
- [0049] 일 구체예로 crRNA 및 CRISPR 연관 단백질에 상기 엑소뉴클레아제를 결합하여 원하는 핵산서열이 존재하는 세포를 효과적으로 사멸시킬 수 있는 CRISPR/CAS 시스템을 CRISPR PLUS라 명명하였다.
- [0050] 일 구체예로, 상기 엔도뉴클레아제 및 엑소뉴클레아제가 결합된 융합단백질은 Cas9-Exoribonuclease T, Cas9-REX2, Cas9-TREX1, Cas9-RecBCD\_RecB, Cas9-RecBCD\_RecC, Cas9-RecBCD\_RecD, Cas9-Exodeoxyribonuclease I, Cas9-Exodeoxyribonuclease III, Cas9-Mungbean, Cas9-RecJ, Cas9-RecE, Cas9-T5, Cas9-Lambda, Cas9-Exonuclease VII small unit, Cas9-Exonuclease VII large unit, Cpf1-Exoribonuclease T, Cpf1-REX2, Cpf1-TREX1, Cpf1-RecBCD\_RecB, Cpf1-RecBCD\_RecC, Cpf1-RecBCD\_RecD, Cpf1-Exodeoxyribonuclease I, Cpf1-Exodeoxyribonuclease III, Cpf1-Mungbean, Cpf1-RecJ, Cpf1-RecE, Cpf1-T5, Cpf1-Lambda, Cpf1-Exonuclease VII small unit 또는 Cpf1-Exonuclease VII large unit일 수 있으며, 바람직하게는 Cas9-RecJ 또는 Cpf1-RecJ일 수 있으나 이에 제한되는 것은 아니다.
- [0051] 본 발명의 일 실시예에서, 엔도뉴클레아제 및 엑소뉴클레아제가 결합된 CRISPR PLUS 단백질을 이용하는 경우, CNV가 2인 유전자인 CCR5와 같이 적은 수의 CNV에서도 세포사멸률이 증가하였으며, 엔도뉴클레아제만을 이용하였을 때보다 더 우수한 세포사멸 효과가 관찰되었다.
- [0052] 이때, 상기 엔도뉴클레아제와 엑소뉴클레아제는 링커를 통해 결합된 것일 수 있다. 상기 링커는 알부민 링커 또는 펩티드 링커일 수 있다. 이때, 상기 링커는 1 내지 50개의 아미노산, 3 내지 40개의 아미노산, 10 내지 30개의 아미노산으로 이루어질 수 있다. 또한, 상기 펩티드 링커는 Gly 및 Ser 잔기로 구성된 펩티드일 수 있다. 또한, 상기 펩티드 링커는 루신(Leu, L), 이소루신(Ile, I), 알라닌(Ala, A), 발린(Val, V), 프롤린(Pro, P), 라이신(Lys, K), 아르기닌(Arg, R), 아스파라진(Asn, N), 세린(Ser, S) 및 글루타민(Gln, Q)으로 구성된 군에서 선택되는 1 내지 10개의 아미노산으로 구성된 펩티드일 수 있다. 또한, 상기 링커는 글라이신(Gly, G) 및 세린(Ser, S) 잔기로 구성된 3 내지 15개의 아미노산으로 구성된 폴리펩티드가 될 수 있으며, 6 내지 11개로 구성될 수 있다.
- [0053] 본 발명의 다른 측면으로, 상술한 일 구체예의 종양 세포 살해용 조성물을 포함하는 암 치료용 약학적 조성물을 제공한다.
- [0054] 이때, 상기 종양 또는 암은 방광암, 골암, 혈액암, 유방암, 흑색종양, 갑상선암, 부갑상선암, 골수암, 직장암, 인후암, 후두암, 폐암, 식도암, 췌장암, 대장암, 위암, 설암, 피부암, 뇌종양, 자궁암, 두부 또는 경부암, 담낭암, 구강암, 결장암, 항문 부근암, 중추신경계 종양, 간암 및 대장암으로 구성된 군으로부터 선택되는 어느 하나일 수 있다.
- [0055] 본 발명의 약학적 조성물의 제형은 비경구용일 수 있다. 제제화할 경우에는 보통 사용하는 충전제, 증량제, 결합제, 습윤제, 붕해제, 계면활성제 등의 희석제 또는 부형제를 사용하여 조제된다. 특히, 비경구 투여를 위한 제제에는 멸균된 수용액, 비수성용제, 현탁제, 유제, 동결건조제, 좌제가 포함된다. 비수성용제 및 현탁용제

로는 프로필렌글리콜(propylene glycol), 폴리에틸렌 글리콜, 올리브 오일과 같은 식물성 기름, 에틸올레이트와 같은 주사 가능한 에스테르 등이 사용될 수 있다.

[0056] 본 발명의 약학적 조성물은 비경구로 투여될 수 있으며, 종양내 투여, 정맥내, 근육내, 피내, 피하, 복강내, 소동맥내, 심실내, 병변내, 척추강내, 국소, 및 이들의 조합으로 구성된 군으로부터 선택되는 어느 하나의 경로로 투여될 수 있다.

[0057] 본 발명의 약학적 조성물의 투여량은 환자의 체중, 연령, 성별, 건강상태, 식이, 투여시간, 투여방법, 배설율 및 질환의 중증도에 따라 그 범위가 다양하며, 당업자에 의해 적절하게 선택될 수 있다. 바람직한 효과를 위해서, 본 발명의 약학적 조성물을 1일 0.01 ug/kg 내지 100 mg/kg으로, 구체적으로는 1 ug/kg 내지 1 mg/kg으로 투여할 수 있다. 투여는 하루에 한번 투여할 수도 있고, 수회 나누어 투여할 수 있다. 따라서, 상기 투여량은 어떠한 면으로든 본 발명의 범위를 한정하는 것은 아니다.

[0058] 또한, 본 발명은 다른 측면으로, 암세포에 특이적으로 존재하는 핵산 서열에 상보적으로 결합하는 폴리뉴클레오티드 및 엔도뉴클레아제를 코딩하는 폴리뉴클레오티드가 적재된 벡터를 유효성분으로 포함하는 종양 세포 살해용 조성물을 제공한다.

[0059] 이때, "암세포에 특이적으로 존재하는 핵산", "엔도뉴클레아제" 및 "엑소뉴클레아제"는 상술한 바와 같다. 또한, "암세포에 특이적으로 존재하는 핵산에 상보적으로 결합하는 폴리뉴클레오티드"는 DNA이다. 상기 DNA 핵산은 암세포에 특이적으로 존재하는 핵산 서열에 상보적으로 결합할 수 있는 crRNA 또는 gRNA를 생산할 수 있다. 이때, "crRNA" 또는 "gRNA"는 상술한 바와 같다.

[0060] 상기 조성물은 i) 암세포에 특이적으로 존재하는 핵산 서열에 상보적으로 결합하는 폴리뉴클레오티드 및 ii) 엔도뉴클레아제를 코딩하는 폴리뉴클레오티드가 하나의 벡터에 적재될 수 있는 것이다. 필요에 따라, 상기 i) 암세포에 특이적으로 존재하는 핵산 서열에 상보적으로 결합하는 폴리뉴클레오티드 및 ii) 엔도뉴클레아제를 코딩하는 폴리뉴클레오티드가 각각의 벡터에 적재되어 사용될 수 있다.

[0061] 또한, 상기 조성물은 엑소뉴클레아제를 코딩하는 폴리뉴클레오티드가 상기 벡터에 추가적으로 포함할 수 있다. 또한, 상기 i) 암세포에 특이적으로 존재하는 핵산 서열에 상보적으로 결합하는 폴리뉴클레오티드, ii) 엔도뉴클레아제를 코딩하는 폴리뉴클레오티드, 및 iii) 엑소뉴클레아제를 코딩하는 폴리뉴클레오티드가 필요에 따라 각각의 벡터에 적재되어 사용될 수 있다.

[0062] 일 구체예에서는 암세포에 특이적으로 존재하는 핵산 서열에 상보적으로 결합하는 폴리뉴클레오티드, CRISPR 연관 단백질 및 엑소뉴클레아제가 결합된 융합단백질을 코딩하는 폴리뉴클레오티드가 적재된 벡터를 유효성분으로 포함하는 종양 세포 살해용 조성물일 수 있다. 본 발명의 일 실시예에서는 크리스퍼 연관 단백질인 엔도뉴클레아제인 Cas9과 엑소뉴클레아제인 RecJ가 융합된 Cas9-RecJ 융합단백질을 코딩하는 폴리뉴클레오티드를 적재하여 사용하였다.

[0063] 상기 벡터는 바이러스 벡터 또는 플라스미드 벡터일 수 있으나, 이에 제한되는 것은 아니다. 상기 암세포에 특이적으로 존재하는 핵산 서열에 상보적으로 결합하는 폴리뉴클레오티드, CRISPR 연관 단백질을 코딩하는 폴리뉴클레오티드 및/또는 엑소뉴클레아제를 코딩하는 폴리뉴클레오티드가 적재된 벡터는 당업계에 공지된 클로닝 방법으로 제조될 수 있으며, 그 방법에 특별히 제한되는 것은 아니다.

[0064] 또한, 본 발명은 상술한 종양 세포 살해용 조성물을 개체에 투여하는 단계를 포함하는 암을 치료하는 방법을 제공한다.

[0065] 이때, 암세포에 특이적으로 존재하는 핵산 서열에 상보적으로 결합하는 폴리뉴클레오티드와 뉴클레아제 단백질이 결합된 RNP 형태로 암에 걸린 개체에 투여될 수 있다. 또한, 엑소뉴클레아제를 추가하여 RNP 형태로 암에 걸린 개체에 투여할 수 있다. 본 발명의 약학 조성물은 가축, 인간 등의 포유동물에 다양한 경로로 투여될 수 있다. 투여의 모든 방식은 예상될 수 있는데, 예를 들면, 종양내 투여, 정맥내, 근육내, 피내, 피하, 복강내, 소동맥내, 심실내, 병변내, 척추강내, 국소, 및 이들의 조합으로 구성된 군으로부터 선택되는 어느 하나의 경로로 투여될 수 있다.

[0066] 이하, 본 발명을 하기 실시예에 의하여 더욱 상세하게 설명한다. 단, 하기 실시예는 본 발명을 예시하기 위한 것일 뿐, 본 발명의 범위가 이들만으로 한정되는 것은 아니다.

[0067] **제조예 1. Cas9을 위한 crRNA 제작**

- [0068] 타겟팅하는 유전자의 정확한 염기서열을 유전자 서열분석을 통하여 확보하였다. 타겟팅하는 유전자는 엑손(exon) 부분에서 protospacer adjacent motif(PAM, 5'-NGG-3') 서열을 찾은 후, 그의 상위부분 20 mer 서열을 protospacer 서열로 정하였다. 목표 서열의 위치에 따라 세포 내에서 편집 효율이 차이가 나기 때문에, protospacer를 적어도 세 종류로 디자인하였다.
- [0069] 20 mer 서열 5' 부분에 5'-TAGG-3' 염기를 결합하고, 상보서열의 3' 부분에는 5'-AAAC-3' 결합하여 각각의 올리고뉴클레오티드(올리고머)를 합성하였다. 합성된 두 서열 올리고머를 100 uM로 맞추고 각각 2 uL씩 취해 46 uL의 정제수에 희석하였다.
- [0070] Thermocycler를 이용하여 어닐링(annealing)을 진행하였고, 95°C에서 5분 처리하고, 4°C/sec의 속도로 55°C까지 온도를 낮춰 10분 동안 처리하였다. In vitro 전사를 위한 T7 promoter(서열번호 1: TAATACGACTCACTATAGG)와 crRNA 스캐폴드 서열(서열번호 2: GTTTTAGAGCTAGAAATAGCAAGTTAAATAAGGCTAGTCCGTTATCAACTGAAAAAGTGGCACCGAGTCGGTGC)을 포함하는 pUC19 벡터 약 5 내지 10 ug을 BsaI 제한효소로 하룻밤 동안 절단 후, 정제하였다. 이의 서열은 5'-TAATACGACTCACTATAGGTGAGACCGcAGGCTCGGTTTTAGAGCTAGAAATAGCAAGTTAAATAAGGCTAGTCCGTTATCAACTGAAAAAGTGGCACCGAGTCGGTGC-3'(서열번호 3)로 표시하였다.
- [0071] 이후, 제한효소가 처리된 벡터(100 내지 200 ng/uL)에 라이게이션(ligation)을 진행하였다. 5번의 어닐링 혼합물 6 uL, 벡터 2 uL, T4 DNA 리가아제 10X Buffer(Promega C126B) 1 uL, 및 T4 DNA 리가아제(Promega M180A) 1 uL를 1.5 uL 에펜도르프 튜브에 넣고 탭핑으로 섞어준 후, 4°C에서 하룻밤 동안 인큐베이션하였다. 대장균 DH5a에 라이게이션 혼합물을 넣어 형질전환하였다. M13 프라이머를 이용한 생어 시퀀싱을 통해 하기 서열로 클로닝 여부를 확인하였다:
- [0072] 5'-TAATACGACTCACTATAGG(서열번호 1)-20 mer protospacer 서열-GTTTTAGAGCTAGAAATAGCAAGTTAAATAAGGCTAGTCCGTTATCAACTGAAAAAGTGGCACCGAGTCGGTGC(서열번호 2) -3'.
- [0073] 이와 같이 만들어진 DNA 서열을 이용하여, T7 promoter, protospacer 및 crRNA 스캐폴드를 포함한 부위를 증폭효소 연쇄반응(PCR)을 통해 증폭하고, 아가로오즈겔 전기영동으로 확인한 후, 정제하였다. MEGAshortscript™ T7 Kit(Invitrogen, AM1354)를 이용하여 10번의 결과물 약 800 ng를 50 uL 반응 볼륨에서 약 4 내지 8시간 반응시켜 crRNA를 in vitro에서 전사하였다.
- [0074] 전사의 결과물을 MEGAclean™ Kit(Invitrogen, AM1908)를 이용하여 정제한 순수한 crRNA를 확보하였고, 분광 광도계로 농도를 측정하였다. 일반적으로 약 1 내지 2 ug/uL 이상의 crRNA를 수득하였다. 이때, RNase 오염에 주의하였다. 정제된 crRNA를 필요한 농도와 양으로 희석 및 분주하여 -80°C에서 보관하였고, 온도 변화 및 충격에 주의하였다.
- [0075] **제조예 2. Cpf1(Cas12a)을 위한 crRNA 제작**
- [0076] 타겟팅하는 유전자의 정확한 염기서열을 유전자 서열분석을 통하여 확보하였다. 타겟 부위를 정하고, 엑손 부분에서 protospacer adjacent motif(PAM, 5'-TTTN-3') 서열을 찾아 그의 하위부분 24 mer 서열을 protospacer 서열로 정하였다. 이때, protospacer 서열의 구아닌과 사이토신의 함량이 약 50% 정도 되는 곳을 타겟으로 정하였다. 목표 서열의 위치에 따라 세포 내에서 편집 효율이 차이가 나기 때문에, protospacer를 적어도 세 종류를 디자인하였다.
- [0077] in vitro 전사를 위한 T7 프로모터와 crRNA를 포함한 서열과 상보 서열에 대해 각각 올리고머를 합성하였다:
- [0078] 5'-AATTCTAATACGACTCACTATAGGAAATTCTACTGTTGAT(서열번호 25) - 24 mer protospacer 서열-3'. 합성된 올리고머 각각 2 ug씩이 들어있는 최종 볼륨 20 uL의 혼합물을 만들었다. 이때, 뉴클레아제가 없는 정제수를 이용하였다.
- [0079] Thermocycler를 이용하여 어닐링을 진행하였고, 95°C에서 5분 처리하고, 4°C/sec의 속도로 55°C까지 온도를 낮춰 10분 동안 처리하였다. MEGAshortscript™ T7 Kit(Invitrogen, AM1354)를 이용하여 5번의 결과물 4 uL(800 ng)를 50 uL 반응 볼륨에서 약 4 내지 8시간 반응시켜 crRNA를 in vitro에서 전사하였다. 전사의 결과물인 crRNA를 에탄올 침적 방법으로 정제하여 확보하였고, 분광 광도계로 농도를 구하였다. 보통 약 1 내지 2 ug/uL 이상의 crRNA를 얻게 되었고, RNase 오염에 주의하였다. 정제된 crRNA를 필요한 농도와 양으로 희석 및 분주하여 -80°C에서 보관하였고, 온도 변화 및 충격에 주의하였다.

- [0080] **제조예 3. 기질 DNA 제작**
- [0081] Cas9 혹은 Cpf1의 protospacer가 타겟팅하는 염기서열을 가운데 부위에 위치하도록 하여, 약 1 내지 1.5 kbp의 dsDNA를 중합효소연쇄 반응을 이용하여 타겟 유전자를 포함한 template로부터 증폭하였다. 이때, 상기 protospacer란 gRNA가 숙주세포내의 DNA에 상보적으로 결합할 수 있는 타겟팅 염기서열을 의미한다.
- [0082] 클로닝을 통해 pUC19, 혹은 pGEM 벡터에 삽입한 후, M13 프라이머를 이용해 시퀀싱을 함으로써, protospacer 서열을 검정하였다. M13 프라이머를 이용한 중합효소 연쇄반응으로 기질 DNA를 증폭한 후 정제하여 100 ng/uL의 농도로 -20℃에서 보관하였다.
- [0083] **실시예 1. crRNA에 의한 CRISPR/Cas 단백질의 비특이적 뉴클레아제 기능 활성화 확인**
- [0084] CRISPR/Cas12a를 포함하는 CRISPR/Cas 단백질의 게놈 편집 기능이 crRNA-가이드된 타겟 시퀀스 결합에 의해 활성화되어 DNA 또는 RNA 분자를 분해하는 비특이적 뉴클레아제의 기능을 활성시킴을 확인하였다. 이의 개략적인 모식도를 도 1에 나타내었다.
- [0085] **실시예 2. CRISPR PLUS의 암세포 특이적인 SNP 인식 및 암세포 사멸**
- [0086] 암세포는 유래 조직과 암의 유형에 따라, 정상 세포에는 존재하지 않는 특정 게놈 부위에 유전자 변이를 포함한 자체 염색체 변이 또는 단일 염기 다형성(SNPs)을 가지고 있다. 이러한 암-특이 SNP는 본 발명의 암세포 특이적인 마커로 사용되었다. 암세포 특이적인 SNP는 SNP에 상보적인 서열을 포함하는 crRNA의 합성에 사용되었고, 이 crRNA를 포함하는 CRISPR 연관 단백질 및 엑소뉴클레아제를 포함하는 CRISPR PLUS 단백질에 의해 인식되었다. 암세포의 게놈에서 SNP와 crRNA 사이의 서열 특이적인 결합은 CRISPR PLUS 단백질의 게놈 편집 기능을 활성화시켜 타겟 DNA/RNA 파손을 일으켰다. 이후, 상기 활성화는 암세포 내 ds 및 ss DNA/RNA 분자를 비가역적으로 파괴하는 CRISPR PLUS의 내제적인 비특이적 뉴클레아제 기능을 활성화시켜, 세포 사멸을 초래하였다. 이러한 결과를 도 2에 모식도로 나타내었다.
- [0087] **실시예 3. SpyCas9(서열번호 46)의 엔도뉴클레아제 기능 확인(In vitro)**
- [0088] 뉴클레아제가 제거된 정제수에 NEBuffer™ 3.1을 최종 1X 농도로 희석하고, 뉴클레아제 120 nM 및 crRNA 120 nM를 넣어준 후, RNP complex 형성을 유도하였다. 혼합 후 상온에서 약 15분간 인큐베이션하였다. 약 200 ng의 기질 DNA를 추가하고 탭핑한 후 37℃에서 반응시켰다. 목표 서열이 없는 기질 DNA에 대한 반응을 함께 확인하기 위해, 이 단계에서 타겟팅이 되지 않는 DNA를 넣어주었다. 반응액의 최종 볼륨은 20 uL로 조절하였고, 반응 후 젤 로딩 염색액을 넣고 잘 섞어주었다. 2% 아가로오즈 겔을 만든 후(Agarose, Sepro, GenDEPOT, A0224-050), 염색이 된 반응물 12 uL를 1kb DNA 마커(Thermo Scientific, SM0311)와 함께 전기영동하였다. 이후, 뉴클리아제의 활성화에 의해 잘려진 기질 DNA 밴드를 관찰하였다.
- [0089] 그 결과, 목표 서열이 없는 기질 DNA만을 반응시킬 경우 DNA가 절단되지 않음을 확인하였다. 그러나, 타겟팅이 되는 DNA를 함께 넣어주면, 모든 DNA가 절단되는 것을 확인하였다.
- [0090] **실시예 4. Cpf1의 비특이적 엑소뉴클레아제 기능 확인(In vitro)**
- [0091] 뉴클레아제가 제거된 정제수에 NEBuffer™ 1.1을 final 1X 농도로 희석하고, CRISPR/Cas12a 120 nM 및 crRNA<sup>DHCR7</sup> 120 nM를 넣어준 후, RNP complex 형성을 유도하였다. 230 nM의 RNP 결합체를 200 ng의 기질 DNA를 조심스럽게 추가 및 탭핑한 후 37℃에서 반응시켰다. 이때, 기질 DNA는 특이 혹은 비특이 DNA 기질 단독, 혹은 특이 DNA와 비특이 DNA를 함께 섞은 후, NEBuffer 1.1 버퍼에서 1.5 시간 혹은 24시간 동안 37℃에서 인큐베이션하였다. 반응액의 최종 볼륨은 20 uL로 조절하였다.
- [0092] 원하는 반응 시간 후, 젤 로딩 염색액을 넣고 잘 섞어주었다. 2% 아가로오즈 겔을 만든 후(Agarose, Sepro, GenDEPOT, A0224-050), 염색이 된 반응물 12 uL를 1kb DNA 마커(Thermo Scientific, SM0311)와 함께 전기영동하였다. 뉴클리아제의 활성화에 의해 잘려진 기질 DNA 밴드를 관찰하였다. 그 결과는 도 3에 나타내었다.
- [0093] 크리스퍼 뉴클레아제가 보유한 서열 비특이적 엑소뉴클레아제 활성을 증명하기 위하여, 크리스퍼 뉴클레아제와 특이적, 그리고 비특이적 DNA 기질을 이용한 in vitro DNA 절단 실험을 수행한 결과, 서열 특이적 엔도뉴클레아제 효소 활성화에 의존적으로 비특이적 엑소뉴클레아제 활성을 가지고 있음을 확인하였다.
- [0094] 구체적으로, CRISPR/Cas12a와 인간 DHCR7 유전자를 타겟팅하는 crRNA, 그리고, crRNA가 타겟팅하는 서열을 가지는 특이 DNA 기질(DNA #1, 1.5 kb) 혹은 crRNA의 타겟팅 서열이 없는 비특이 DNA(DNA #2, 0.5 kb)를 1.5 시간

혹은 24 시간 동안 인큐베이션하여 DNA 절단을 유도하고, 이를 아가로즈 젤에서 확인하였다(도 3 참조). 이때, DHCR7에 상보적으로 결합하는 crRNA, dsDNA1(1431bp) 및 dsDNA2(544bp) 서열을 각각 서열번호 132, 서열번호 133 및 서열번호 134로 나타내었다.

[0095] 특이 DNA 기질을 뉴클레아제 및 crRNA와 1.5 시간 동안 인큐베이션 하였을 때, 기질은 서열 특이적으로 절단되어 약 0.7 kb의 절편으로 되었다(윗 패널 라인 3), 그러나 crRNA가 없을 경우 절단은 일어나지 않은 것(윗 패널 라인 4)을 확인하였다. 비특이 DNA를 동일한 조건에서 인큐베이션 했을 때는 crRNA 존재에 관계없이 DNA의 절단이 일어나지 않았다(윗 패널 라인 5, 6).

[0096] 특이 DNA 기질과 비특이 기질을 동시에 뉴클레아제로 처리했을 경우에는, 예상대로 특이 DNA 기질만이 절단되는 것이 관찰되었다(윗 패널 라인 7, 8). 또한, 특이 DNA 기질과 뉴클레아제 및 crRNA의 인큐베이션 시간을 24 시간으로 늘였을 때, 특이 DNA 기질과 그 절편 모두 사라지는 것이 관찰 되었다(아래 패널 라인 3). 이는 DNA가 크리스퍼 뉴클레아제의 엑소뉴클레아제 활성화에 의해 잘려졌다는 것을 의미한다.

[0097] 같은 실험을 crRNA가 없는 상태에서 했을 때, DNA가 사라지지 않는 것을 볼 때(아래 패널 라인 4), 이 결과는 엑소뉴클레아제 활성화는 CRISPR/Cas12a의 서열 특이적 효소 활성화에 의존적임을 의미한다. 또한, 비특이 DNA를 뉴클레아제 및 crRNA로 24시간 동안 처리 했을때, DNA가 사라지지 않고 유지되는 것에서도 증명되었다(아래 패널 라인 5, 6).

[0098] 또한, 특이 DNA 기질과 비특이 기질을 동시에 뉴클레아제 및 crRNA로 24 시간 처리했을 때, 특이, 비특이 DNA 기질 모두가 분해되어 사라지며(아래 패널 라인 7), 이것은 crRNA의 존재 하에서만 일어나는 것임이 관찰되었다(아래 패널 라인 8). 이는 서열특이적 엔도뉴클레아제 기능의 활성화에 의해 유도된 CRISPR/Cas12a의 엑소뉴클레아제 기능이 서열 비특이적으로 작동한다는 것을 의미한다.

[0099] 따라서, 이 실험 결과는 CRISPR/Cas12a는 서열특이적 효소 활성화에 의존적으로, 비특이적 엑소뉴클레아제의 활성을 가지고 있음을 의미한다.

[0100] **실시예 5. 세포 독성 분석**

[0101] 인간 암 유래 세포인, HeLa 세포는 DMEM/10% FBS 성장 배지를 이용하여 37°C의 온도로 5% CO<sub>2</sub> 인큐베이터에서 배양하였다. 형질전환 하루 전, 2.5 X 10<sup>4</sup> 세포를 100 ul 배양액에 풀어 96웰에 플레이팅하였다. Blank(background control)는 100 ul의 media만 넣어주었다. 다음날, CRISPR/Cas 뉴클레아제와 crRNA complex(RNP)를 이용한 형질감염을 하기 표 3과 같은 조건별로 수행하였다. 상기 crRNA 중 하나는 인간 DHCR7 유전자에 서열 특이성을 가지며, 다른 하나는 비의 DWARF5 유전자에 서열 특이성을 가지는 것이다.

표 3

[0102]

조건	1	2	3	4	5	6	7
Lipofectamine	X	0	0	0	0	0	0
CRISPR/Cas nuclease(CRISPR/cas 12a)	X	0	X	0(1.2 nM)	0 (1.2 nM)	0 (2.4 nM)	0 (2.4 nM)
Targeting crRNA	X	X	0	0(1.2 nM)	X	0(2.4 nM)	X
Non-targeting crRNA	X	X	X	X	0(1.2 nM)	X	0(2.4 nM)

[0103] 각 웰당 1.5 ml 튜브에 Opti-MEM media 5 ul, CRISPR/CAS 2.4 nM, crRNA 2.4 nM을 섞은 후, 상온에서 10분간 인큐베이션하였다. 같은 튜브에 0.17 ul의 Lipofectamine Cas Plus Reagent를 넣고 상온에서 5분간 인큐베이션 하였다. 상기 튜브를 인큐베이션 하는 동안 다른 튜브를 준비하였다. Opti-MEM 5 ul, Lipofectamin CRISPRMAX Reagent 0.3 ul를 섞고, 상온에서 5분간 인큐베이션하였다. 상기 두 튜브를 섞어준 후, 상온에서 10분간 인큐베이션하였다. 섞어준 튜브 용액을 세포가 자라는 각 웰에 방울방울 넣어주었다. 이후, 세포를 37°C의 온도로 5% CO<sub>2</sub> 인큐베이터에서 배양하였다.

[0104] 24, 48 및 72시간 후, 깨끗한 벤치(bench)에서 WST-1(Cell proliferation Reagent, Roche 0501594401)를 10 ul 씩 각 웰에 첨가한 후, 37°C의 온도로, 5% CO<sub>2</sub> 인큐베이터에 넣고 색깔 변화를 관찰하였다(열은 빨강 -> 진한 빨강). 10분 후, FLUOstar Omega ELISA reader(BMG Labtech)를 이용하여, 420 내지 480 nm와 690 nm에서

background와 샘플의 흡광도를 측정하였다. 타겟 crRNA와 비타겟 crRNA에서의 CRISPR/Cas 뉴클레아제의 세포 독성을 분석하였다.

[0105] 인간 유래 암세포인 HEK293(도 4a)와 HeLa(도 4b)에 CRISPR/Cas12a 뉴클레아제, crRNA 혹은 두 분자의 결합체(RNP complex)를 각각 형질감염시킨 후, 24, 48, 72 시간 후에 세포의 생존도를 WST-1을 이용한 생존도 assay로 측정하였다. 그 결과, 뉴클레아제 혹은 crRNA 만을 형질감염시킨 세포는 24, 48, 72시간 모두에서 생존도의 변화가 거의 없음을 나타내었다(도 4 참조). 이러한 결과는 뉴클레아제와 crRNA 자체는 세포에 독성이 없음을 의미한다.

[0106] 반면에, 서열 특이적 효소 활성을 보유한 결합체에 의해 형질감염된 세포는 72시간에서 그 생존도가 현저히 떨어지는 현상을 보였다. 이 결과는 CRISPR/Cas12a 뉴클레아제가 서열 특이적 효소 활성에 의존적으로 세포에 독성을 나타냄을 의미한다. 같은 결합체로 형질감염된 세포가 24, 48시간에서는 생존도에 변화가 없든지(HEK293) 혹은 경미하게 떨어지는(HeLa) 현상을 보였다. 이는 CRISPR/Cas12a 뉴클레아제의 독성이 세포에 영향을 미치기 위하여 어느 정도의 시간이 필요함을 의미한다.

[0107] 일반적으로, CRISPR/Cas 뉴클레아제의 서열 특이적 효소 활성이 세포내에서 24시간부터 48시간까지 꾸준히 진행됨이 알려져 있다. 또한, 이 활성에 의해서 뒤이어 세포 독성을 나타내는 비특이적 뉴클레아제 활성이 유도됨을 유추해 볼 때, 72시간 후에 나타나는 세포독성은, 뉴클레아제 활성과 무관한 간접적 효과가 아니라, 뉴클레아제의 서열 특이적 효소활성과 연관된 기능에 의해 야기 되는 것이라고 할 수 있다. 그럼에도, 특이적 crRNA를 사용한 경우에 비특이적인 crRNA를 사용한 경우에 비해 세포독성이 높음을 확인할 수 있었다.

[0108] 그러므로, 이 실험의 결과는 CRISPR/Cas12a 뉴클레아제는 서열 특이적 효소활성에 의존적으로 세포에 독성을 나타내어 생존도를 떨어뜨리는 기능을 가지고 있음을 보여준다.

[0109] **실시예 6. 암세포 특이적 독성 분석**

[0110] 인간유래 폐암세포와 정상세포의 유전자 서열을 분석하여 암세포에서만 특이적으로 존재하는 SNP를 찾아내고, 이를 타게팅할 수 있는 crRNA를 합성하였다. 크리스퍼 뉴클레아제와 crRNA를 혼합하여 RNP 결합체를 만든 후, 암세포와 정상세포에 형질감염 시키고, 암세포 특이적 사멸 효과를 WST-1 기반 세포 생존도 assay를 이용하여 분석하였다. 이때, 사용한 crRNA는 폐암의 EGFR에 특이적으로 존재하는 서열(서열번호 43)을 타게팅하도록 제조하였다.

[0111] 형질전환 하루 전,  $2.5 \times 10^4$  세포를 100 ul 배양액에 풀어 96웰에 플레이팅하였다. Blank(background control)는 100 ul의 media만 넣어주었다. 다음 날, CRISPR/CAS 뉴클레아제와 crRNA complex(RNP)를 이용한 형질감염을 하기 표 4와 같은 조건으로 수행하였다.

**표 4**

[0112]

조건	1	2	3	4	5	6	7	8	9
Lipofectamine	X	0	0	0	0	0	0	0	0
CRISPR/Cas nuclease	X	X	0	X	X	0(1.2nM)	0(1.2nM)	0 (2.4nM)	0 (2.4nM)
Targeting crRNA	X	X	X	0	X	0(1.2 nM)	X	0(2.4nM)	X
Non-targeting crRNA	X	X	X	X	0	X	0(1.2nM)	X	0(2.4nM)

[0113] 각 웰당 1.5 ml 튜브에 Opti-MEM media 5 ul, CRISPR/CAS 2.4 nM, crRNA 2.4 nM을 섞은 후, 상온에서 10분간 인큐베이션하였다. 같은 튜브에 0.17 ul의 Lipofectamine Cas Plus Reagent를 넣고 상온에서 5분간 인큐베이션하였다. 상기 튜브를 인큐베이션 하는 동안 다른 튜브를 준비하였다. Opti-MEM 5 ul, Lipofectamin CRISPRMAX Reagent 0.3 ul를 섞고, 상온에서 5분 간 인큐베이션하였다. 상기 두 튜브를 섞어준 후, 상온에서 10분간 인큐베이션하였다. 섞어준 튜브 용액을 세포가 자라는 각 웰에 방울방울 넣어주었다. 이후, 세포를 37℃의 온도로 5% CO<sub>2</sub> 인큐베이터에서 배양하였다.

[0114] 24, 48 및 72시간 후, 깨끗한 벤치(bench)에서 WST-1(Cell proliferation Reagent, Roche 0501594401)를 10 ul씩 각 웰에 첨가한 후, 37℃의 온도로, 5% CO<sub>2</sub> 인큐베이터에 넣고 색깔 변화를 관찰하였다(열은 빨강 -> 진한 빨강). 10분 후, FLUOstar Omega ELISA reader를 이용하여, 420 내지 480 nm와 690 nm에서 background와 샘플

의 흡광도를 측정하였다. 타겟 crRNA와 비타겟 crRNA에서의 CRISPR/Cas 뉴클레아제의 세포 독성을 분석하였다.

[0115] 그 결과, 폐암세포에서 타겟 crRNA를 처리한 군에서만 특이적으로 폐암세포가 사멸하는 것을 확인하였다.

[0116] **실시예 7. EGFR 돌연변이 서열 특이적 가이드 RNA에 의한 세포 사멸 효과 확인**

[0117] 본 실시예에서는 Cas9 단백질 발현 벡터(PX459, Addgene plasmid #62988)를 이용해 폐암 세포인 HCC827 세포의 유전체에 multi-cleavage를 일으켜 세포 사멸을 일으킬 수 있음을 증명하였다. 전기천공법(electroporation)으로 세포 내부에 벡터들을 전달 하였으며, 유전체에 multi-cleavage를 일으키기 위해 HCC827세포 EGFR 유전자의 E2 돌연변이 서열을 타겟하는 가이드 RNA를 사용하였다. EGFR 유전자의 E2 돌연변이 서열은 약 18개 이상의 multi-copy로 존재하는 것으로 알려져 있다.

[0118] HCC827 세포는 75T 플라스크에서 RPMI-1640(10% 소태아혈청(fetal bovine serum)) 배지로 약 50% 밀도의 양으로 배양한 후, 트립신화(trypsinization) 및 PBS 세척을 거쳐 최종적으로 Neon Electroporation Buffer R에 리서스펜드(resuspend) 하였다. 10 μL Neon 파이프 팁(tip)에 150,000개의 세포와 500 ng의 벡터를 로딩한 후, 1300 V, 20 ms 및 2 pulses의 조건으로 전기천공법을 진행하였다. 이후, RPMI-1640(10% 소태아혈청(fetal bovine serum)) 배지에서 회복시키고, 6일 후 세포들을 수확하여 카운팅(counting)을 실시하였다. 그 결과를 도 5 및 도 6에 나타내었다.

[0119] 도 5 및 도 6에 나타난 바와 같이, electric pulse 군과 비교하였을 때, 타겟 서열이 존재하지 않는 EGFR\_WT 실험군(pSpCas9(BB)-2A-Puro(PX459)V2.0-EGFR\_WT)에 비해 EGFR\_E2 실험군(pSpCas9(BB)-2A-Puro(PX459)V2.0-EGFR\_E2)의 세포수가 현저하게 감소한 것을 확인하였다. 구체적으로, 상기 EGFR\_E2 실험군에서 83%의 세포사멸이 유도되었다. 이를 통해, 암세포에 Cas9 단백질과 세포 유전체 서열 특이적 multi-target 가이드 RNA를 넣어 주었을 때 세포 사멸을 일으킨다는 것을 확인할 수 있었다.

[0120] **실시예 8. 폐암 세포 H1299에서 표적 위치에 따른 세포 사멸 효과 확인**

[0121] **실시예 8.1. Lipofection을 통한 gRNA 및 크리스퍼 단백질 도입**

[0122] 24 웰 플레이트에서, 웰 1개당 폐암 세포인 H1299를 1.5x10<sup>5</sup>개 만큼 깔아주고, 24시간 후에 각 웰에 하기 표 5와 같은 종류의 DNA(CCR5, HPRT1, MT2, SMIM11, GNPDA2, SLC15A5 및 KCNE2)를 도입하였다. 도입을 위해, Lipofectamine 3000 시약을 사용하였고, 제조사에서 제공하는 매뉴얼을 따랐으며, DNA를 500 ng씩 사용하였다.

**표 5**

[0123]

조건	NT1	CCR5	HPRT1	MT2	SMIM11	GNPDA2	SLC15A5	KCNE2
Cas9	500 ng	500 ng						

[0124] 상기 DNA 도입 즉, 트랜스펙션(transfection) 시점을 기점으로 하여 72시간 후, 각 웰의 배양액을 모두 석션으로 제거하고 1X PBS를 500 μL만큼 넣어 한 번 세척하였다. 이후, Trypsin-EDTA를 처리하여 세포를 모두 떼어냈다. 이들을 Trypan blue 염색약을 이용하여 염색한 후, 살아있는 세포 수를 측정하였다. 그 결과를 도 7 및 도 8에 나타내었다.

[0125] 도 8에 나타난 바와 같이, 실험 결과 H1299에서 자르는 표적 서열이 존재하지 않는다고 알려진 NT(non target)와 한 쌍의 표적 위치만을 가지는 CCR5 및 HPRT1은 서로 비슷한 양의 세포 사멸 효과를 보였다. 반면, 100 군데 이상을 자른다고 알려진 MT2는 이들에 비해 살아있는 세포의 양이 약 50% 수준으로 떨어짐으로써 세포 사멸 효과를 나타내었다. 또한, 폐암 세포주의 표적 위치인 SMIM11(약 74%), GNPDA2(약 58%), SLC15A5(약 45%) 및 KCNE2(약 77%)에서 모두 MT2만큼 높은 사멸 효과를 확인하였으며, 이중 SMIM11과 KCNE2는 MT2보다 사멸 효과가 우수하였다. 각 실험군의 Copy number, Essentiality를 하기 표 6에 나타내었으며, 각 실험군에서 세포의 형태(morphology)를 현미경으로 관찰하여 도 7에 나타내었다.

**표 6**

[0126]

조건	NT1	CCR5	HPRT1	MT2	SMIM11	GNPDA2	SLC15A5	KCNE2
Copy number	0	2	2	>100	>40	>12	>12	>40
Essentiality	N/A	Nonessential	House keeping	Non essential	Non essential	oncogene	oncogene	Non essential

[0127] 실시예 8.2. 리포펙타민을 이용한 gRNA 및 크리스퍼 단백질 도입

[0128] 24 웰 플레이트에서, 웰 1개당 폐암 세포인 H1299 세포를  $1.5 \times 10^5$  개 만큼 갈아주고, 24시간 후에 각 웰에 하기 표 7과 같은 종류의 DNA를 도입하였다. 도입을 위해 Lipofectamine3000 reagent를 사용하였고, 제조사에서 제공하는 매뉴얼을 따랐으며, DNA를 500 ng씩 사용하였다.

표 7

[0129]	조건	Lipo only	NT1	CCR5	HPRT1	MT2	GNPDA2	SLC15A5	KCNE2
	cut number	N/A	0	2	2	>100	>12	>12	>40
	essentiality	N/A	N/A	Non-essential	House-keeping	Non-essential	oncogene	oncogene	Non-essential

[0130] 상기 DNA 도입 즉, 트랜스펙션 시점을 기점으로 하여 48시간 후, 각 웰의 배양액을 모두 석션으로 제거하고 1X PBS를 500  $\mu$ L 만큼 넣어 한 번 세척하였다. 이후, Trypsin-EDTA를 처리하여 세포를 모두 떼어냈다. 이들을 Trypan blue 염색약을 이용하여 염색한 후, 살아있는 세포 수를 측정하였다. 그 결과를 도 9 및 도 10에 나타내었다. 본 실시예에서는 DNA의 도입 없이 Lipofection만 수행한 실험군을 포함시켰다.

[0131] 도 9에 나타난 바와 같이, Lipofection only(Lipo only)에 비해 CCR5에서 살아있는 세포의 수가 80% 정도 수준으로 떨어짐을 관찰하였다. NT1, HPRT1의 경우 CCR5 대비 70% 정도로 나타났다. MT2에서는 NT1에 비해 25% 수준으로 세포 수가 떨어졌으며, CNV를 표적으로 한 3종(GNPDA2, SLC15A5 및 KCNE2)에서는 더욱 많은 세포가 사멸하여 Trypan blue를 이용한 세포 수 측정법으로는 살아있는 세포를 유의미하게 찾아낼 수 없었다. 구체적으로, NT1, CCR5, HPRT1, MT2, GNPDA2, SLC15A5 및 KCNE2에서 각각 약 43%, 약 23%, 약 50%, 약 86%, 약 99%, 약 99% 및 약 99%의 세포사멸률을 나타내었다.

[0132] NT1 및 GNPDA2 실험군을 각각 세포들이 많이 살아있는 실험군과 많이 죽은 실험군의 대표로 선정하였다. 이후, NucBlue Live ReadyProbes Reagent 및 Propidium Iodide ReadyProbes Reagent를 이용하여 각 웰에서 현미경으로 이미징하였다. 이를 도 10에 나타내었다. 형광으로 살아있는 세포의 비율을 분석한 결과, 도 9에 나타난 결과와 유사하였다.

[0133] 실시예 9. 폐암 세포 H1563에서 표적 위치에 따른 세포 사멸 효과 확인

[0134] 24 웰 플레이트에서, 웰 1개당 H1563 세포를  $1.5 \times 10^5$  개 만큼 갈아주고, 24시간 후에 각 웰에 하기 표 8과 같은 종류의 DNA를 도입하였다. 도입을 위해 Lipofectamine3000 reagent를 사용하였고, 제조사에서 제공하는 매뉴얼을 따랐으며, DNA를 500 ng씩 사용하였다.

표 8

[0135]	조건	CCR5	HPRT1	MT2	IRX1	ADAMTS16
	Copy number	2	2	>100	>8	>7
	Essentiality	Non essential	House keeping	Non essential	Non essential	Non essential

[0136] 상기 DNA 도입 즉, 트랜스펙션 시점을 기점으로 하여 24시간 후에 1  $\mu$ g/mL 농도의 puromycin을 처리하여 72시간 동안 셀렉션하였다. 이후, 각 웰의 배양액을 모두 석션으로 제거하고, 1X PBS를 500  $\mu$ L만큼 넣어 한 번 세척하였다. 이후, Trypsin-EDTA를 처리하여 세포를 모두 떼어냈다. 이들을 Trypan blue 염색약을 이용하여 염색한 후, 살아있는 세포 수를 측정하였다. 그 결과를 도 11에 나타내었다.

[0137] 도 11에 나타난 바와 같이, 단일 표적인 CCR5와 대비하여 HPRT1은 52% 정도 수준, MT2에서는 43% 수준, 그리고 CNV를 표적으로 한 2종(IRX1 및 ADAMTS16)에서는 각각 37% 및 40%로 적은 양의 세포들이 남아있었다. 이를 통해, 많은 양의 DSB(double strand breaker)를 유도함으로써 CCR5와 같은 단일 표적에 비해 세포 사멸 효과를 높일 수 있음을 확인하였다.

- [0138] **실시예 10. 폐암 세포 A549에서 표적 위치에 따른 세포 사멸 효과 확인**
- [0139] **실시예 10.1. multi 타겟에 의한 A549 세포 사멸 확인**
- [0140] 폐암 세포 A549에 Cas9 단백질과 gRNA가 발현되는 DNA 500 ng을 전기 천공법(Lonza)으로 도입하였다. 배양 중인 A549 세포를 1X PBS로 세척한 후, Trypsin-EDTA를 처리하여 바닥에서 떼어내었다. 필요한 세포 수를 덜어내어 1X PBS로 1번 세척한 후, SF buffer(Lonza)로 풀어준 후 각각의 DNA와 섞었다. 세포와 DNA 혼합액을 전기천공장치(Lonza)에 넣어 전기충격을 가하였다. 대조군으로서 인간 유전자서열에서 얼라인(aligned)되지 않는 NT1, 1 copy를 가지며 house keeping gene인 HPRT1, 1 copy의 유전자를 가지는 CCR5를 타겟하는 조건, 그리고 전기충격만 주는 조건(pulse only)을 추가하였다.
- [0141] 전기충격으로 DNA를 도입한 후, 24 웰에 2 반복, 그리고 96 웰에 3 반복으로 세포를 넣어주었다. DNA 도입 24, 42 및 72시간 후, 50  $\mu$ L의 CellTiter Glo reagent를 96 웰에 넣어주었다. FLUOstar omega reader 기계에 플레이트를 넣고 2분간 흔들어 주었다. 그리고, 10분 동안 상온에서 반응시킨 후, 발광도를 측정하였다. 상기 방법은 세포 내의 ATP양을 기반으로 대사과정을 하고 있는 살아있는 세포 양을 발광도를 통해 판별하는 방법이다. 그 결과를 도 12에 나타내었다.
- [0142] 도 12에 나타난 바와 같이, 3가지 대조군(NT1, HPRT1 및 CCR5)에 비해 100 곳 이상을 타겟하는 MT2 조건에서 시간이 지남에 따라 20% 내지 50%의 세포사멸이 유도되었다.
- [0143] DNA 도입 24시간 후, 24 웰에 1  $\mu$ g/mL의 puromycin을 넣고, 전기충격만 준 조건에서 세포가 90% 정도 사멸하는 시점에서 세포배양액을 바꿔주어 세포를 회복시켰다. 5일 내지 7일 회복시킨 후, 각 웰에서 세포를 떼어내어 trypan blue로 염색한 후, 살아있는 세포 수를 측정하였다. 그 결과를 도 13에 나타내었다. 한편, DNA 도입 후, puromycin으로 셀렉션(selection)한 후 세포 수를 측정하기 전 현미경으로 관찰한 사진을 도 14에 나타내었다.
- [0144] 그 결과 puromycin 저항성이 없는 전기충격만 준 조건에서는 세포가 살아남지 않아 컨트롤이 되었고, NT1, HPRT1조건에 비해 CCR5를 타겟했을때, 50%정도의 세포사멸이 보였다. 그리고 MT2 조건은 약 90%의 세포가 사멸한 것으로 확인되었다. DNA 도입 후 puromycin으로 selection 한 후 세포수를 측정하기전 현미경으로 관찰한 사진이다. 세포수 측정결과와 마찬가지로 NT1 컨트롤에 비해서 1 copy만 가지는 HPRT1, CCR5를 타겟할 경우 50% 이상의 세포가 회복 되는 반면 MT2 조건의 경우 90%정도의 세포는 사멸이 일어났고, 10% 세포만 회복되는 것을 관찰하였다(파란색 화살표: 회복된 세포 colony). 따라서 A549 세포에 Cas9 단백질을 이용하여 multi DNA Break를 유도하였을 때 세포사멸이 일어나는 것을 확인하였다.
- [0145] **실시예 10.2. CNV 타겟에 의한 A549 세포 사멸 확인**
- [0146] 폐암세포주인 A549세포에 CNV를 가지는 유전자를 타겟하여 Cas9 단백질로 DNA break를 유도했을때, 세포사멸 정도를 확인하였다. 구체적으로, 폐암 세포인 A549에 Cas9 단백질과 각 CNV 타겟 gRNA가 발현되는 DNA 500 ng을 전기 천공법(Lonza)으로 도입하였다. 배양 중인 A549 세포를 1X PBS로 세척한 후, Trypsin-EDTA를 처리하여 바닥에서 떼어내었다. 필요한 세포 수를 덜어내어 1X PBS로 1번 세척한 후, SF buffer(Lonza)로 풀어준 후 각 DNA와 섞었다. 세포와 DNA 혼합액을 전기천공장치(Lonza)에 넣어 전기충격을 가하였다. 대조군으로서 대장균에서 단백질 발현할 수 있는 pET21a 벡터(vector)를 도입한 조건, 전기충격만 주는 조건(pulse only), 및 아무런 처리도 하지 않은 조건(no pulse)을 추가하였다.
- [0147] 전기충격으로 DNA를 도입한 후, 96 웰에 각 조건당 3 반복으로 세포를 넣어주었다. DNA 도입 24, 44 및 51시간 후, 50  $\mu$ L의 CellTiter Glo reagent를 웰마다 넣어주었다. FLUOstar omega reader 기계에 플레이트를 넣고 2분간 흔들어 주었다. 그리고, 10분 동안 상온에서 반응시킨 후, 발광도를 측정하였다. 그 결과를 도 15에 나타내었다.
- [0148] 도 15에 나타난 바와 같이, 3가지 대조군(pulse only, pET21a 및 no pulse)에 비해 100곳 이상을 타겟하는 MT2 조건에서 시간이 지남에 따라 20% 내지 50%의 세포 사멸이 유도되었다. 또한, CD68, DACH2 및 HERC2P2 3종의 CNV타겟은 MT2와 유사한 세포사멸을 유도하였고, SHBG의 CNV타겟은 대조군에 비해 70% 내지 80%의 세포 사멸을 유도하였다. 따라서, A549 세포가 가지고 있는 CNV 4종(CD68, DACH2, HERC2P2 및 SHBG2)을 Cas9 단백질을 이용하여 타겟 DNA Break를 유도하였을 때 세포사멸이 일어나는 것을 확인하였다.
- [0149] **실시예 11. 유방암 세포 SKBR3에서 표적 위치에 따른 세포 사멸 효과 확인**
- [0150] **실시예 11.1. CNV 타겟에 의한 SKBR3 세포 사멸 확인**

- [0151] 유방암 세포인 SKBR3에 Cas9 단백질과 각 CNV 타겟 gRNA가 발현되는 DNA 500 ng을 전기천공법(Lonza)으로 도입하였다. 배양 중인 SKBR3 세포를 1X PBS로 세척한 후, Trypsin-EDTA를 처리하여 바닥에서 떼어내었다. 필요한 세포 수를 덜어 내어 1X PBS로 1번 세척한 후, SF buffer(Lonza)로 풀어준 후 각 DNA와 섞었다. 세포와 DNA 혼합액을 전기천공장치(Lonza)에 넣어 전기충격을 가하였다. 대조군으로서 인간 유전자서열에서 얼라인되지 않는 NT1(non target), 전기충격만 주는 조건(pulse only), 및 아무런 처리도 하지 않은 조건(no pulse)을 추가하였다. 전기충격으로 DNA를 도입한 후, 24 웰에 2 반복 및 96 웰에 3 반복으로 세포를 넣어주었다.
- [0152] DNA 도입 24, 42 및 48시간 후, 50  $\mu$ L의 CellTiter Glo reagent를 96 웰에 넣어주었다. FLUOstar omega reader 기계에 플레이트를 넣고 2분간 흔들어 주었다. 그리고, 10분 동안 상온에서 반응시킨 후, 발광도를 측정하였다. 그 결과를 도 16에 나타내었다.
- [0153] 도 16에 나타난 바와 같이, 3가지 대조군(NT1, pulse only 및 no pulse)에 비해 100곳 이상을 타겟하는 MT2 조건에서 30%의 세포사멸이 유도되었다. 또한 ERBB2 및 KRT16 2종의 CNV 타겟은 대조군에 비해 40% 내지 50%의 세포사멸을 유도하였다.
- [0154] **실시예 11.2. CNV 타겟에 의한 SKBR3 세포 사멸 확인**
- [0155] 유방암 세포인 SKBR3에 Cas9 단백질과 각 CNV 타겟 gRNA가 발현되는 DNA 500 ng을 전기천공법(Lonza)으로 도입하였다. 배양 중인 SKBR3 세포를 1X PBS로 세척한 후, Trypsin-EDTA를 처리하여 바닥에서 떼어내었다. 필요한 세포 수를 덜어 내어 1X PBS로 1번 세척한 후, SF buffer(Lonza)로 풀어준 후 각 DNA와 섞었다. 세포와 DNA 혼합액을 전기천공장치(Lonza)에 넣어 전기충격을 가하였다. 대조군으로서 인간 유전자서열에서 얼라인되지 않는 NT1, 1 copy를 가지며 house keeping gene인 HPRT1을 타겟하는 조건을 추가하였다. 전기충격으로 DNA를 도입한 후, 24 웰에 2 반복으로 세포를 넣어주었다.
- [0156] DNA 도입 48시간 후, 24 웰의 각 웰에서 세포를 떼어내어 trypan blue로 염색한 후 살아있는 세포수를 측정하였다. 그 결과를 도 17에 나타내었다. 도 17에 나타난 바와 같이, 대조군에 비해 100곳 이상을 타겟하는 MT2 조건에서 40%의 세포사멸을 유도하였고, ERBB2 및 KRT16 2종의 CNV 타겟은 대조군에 비해 40% 내지 50%의 세포사멸을 유도하였다. 따라서, SKBR3 세포가 가지고 있는 CNV 2종(ERBB2 및 KRT16)을 Cas9 단백질을 이용하여 타겟 DNA Break를 유도하였을 때 세포사멸이 일어나는 것을 확인하였다.
- [0157] **실시예 12. 자궁경부암 세포 HeLa에서 표적 위치에 따른 세포사멸 확인**
- [0158] **실시예 12.1. CNV 타겟과 HPV 유전자 타겟에 의한 세포 사멸 확인**
- [0159] 자궁경부암 세포주인 HeLa 세포에 Cas9 단백질과 gRNA가 발현되는 벡터 600 ng을 전기천공법으로 도입하였다. 배양 중인 HeLa 세포를 1X PBS로 세척한 후, Trypsin-EDTA를 처리하여 바닥에서 떼어내었다. 필요한 세포 수를 덜어 내어 1X PBS로 1번 세척한 후, SE buffer(Lonza)로 풀어준 후 각 벡터와 섞었다. 세포와 DNA 혼합액을 전기천공장치(Lonza)에 넣어 전기충격을 가하였다. 대조군으로서 1 copy의 non essential gene인 CCR5를 타겟하는 조건과 100개 이상의 non essential gene을 타겟하는 MT2 조건을 추가하였고, 실험군으로 HeLa에서 8 copy 이상 있는 것으로 알려진 PRDM9와 30 copy 이상 있는 것으로 알려진 Human PapillomaVirus(HPV) 유래의 유전자를 타겟하여 HeLa세포 특이적 세포사멸을 유도하는 실험을 진행하였다.
- [0160] 전기충격으로 DNA를 도입한 후, 24 웰에 2 반복으로 세포를 넣어준 후, 24시간이 지나 0.5  $\mu$ g/ml의 Puromycin을 넣어 DNA 벡터가 안 들어간 세포들을 죽이는 셀렉션 과정을 진행하였다. 셀렉션 3일 후, Puromycin이 안 들어간 세포배양액으로 바꾸주어 세포를 배양하였다. 3 내지 4일 후, 각 웰에서 세포를 떼어내어 trypan blue로 염색한 후 살아있는 세포 수를 측정하였다. 그 결과를 도 18 및 도 19에 나타내었다.
- [0161] 도 18 및 도 19에 나타난 바와 같이, HeLa 세포에서 1곳만 자르는 CCR5와 비교했을 때, 100곳 이상을 타겟하는 MT2 조건이 시간이 지남에 따라 약 50%의 세포사멸이 유도되는 것을 확인할 수 있었다. 또한, CNV 타겟인 PRDM9와 HPV 유전자 타겟인 HPV\_1의 경우 MT2 조건과 유사하거나 더 많은 세포사멸을 유도하는 것을 확인하였다. 구체적으로, 도 18에서, MT2 및 PRDM9에서 약 50% 및 약 80%의 세포사멸률을 나타내었다. 또한, 도 19에서, MT2, PRDM9 및 HPV\_1에서 약 50%, 약 40% 및 약 40%의 세포사멸률을 나타내었다. 따라서, HeLa 세포만 가지고 있는 CNV와 HPV 유전자를 Cas9 단백질을 이용하여 타겟 DNA Break를 유도하였을 때 세포사멸이 일어나는 것을 확인하였다.
- [0162] **실시예 12.2. CNV 타겟과 HPV 유전자 타겟에 의한 세포 사멸 확인**
- [0163] 전기충격으로 HeLa 세포에 DNA 벡터를 도입하는 과정은 상기 실시예 12.1의 방법과 동일하다. 전기충격으로 DNA

를 도입한 후, 96 웰에 2 반복으로 세포를 넣어주었다. 이후, 24, 48 및 72시간 간격으로 Cell Titer glo 2.0 을 사용하여 세포 내의 ATP양을 측정함으로써 대사과정을 하고 있는 살아있는 세포 정도를 발광도로 판별하였다. 대조군으로는 GFP를 발현하는 대조군 벡터, 1 copy의 non essential gene인 CCR5를 타겟하는 조건, 그리고 100개 이상의 non essential gene을 타겟하는 MT2 조건을 추가하였다. 실험군으로 HeLa에서 8 copy 이상 있는 것으로 알려진 PRDM9와 30 copy 이상 있는 것으로 알려진 Human Papilloma Virus(HPV) 유래의 유전자를 타겟하여 HeLa세포 특이적 세포사멸을 유도하는 실험을 진행하였다. 그 결과를 도 20a에 나타내었다.

[0164] 도 20a에 나타난 바와 같이, HeLa 세포에서 1곳만 자르는 CCR5 타겟과 GFP만 들어간 Pulse only는 비슷한 정도의 luminescence signal을 나타내었다. 또한, CCR5에 비하여 MT2 조건에서 약 50% 세포사멸이 유도되는 것을 확인할 수 있었다. 또한, CNV 타겟인 PRDM9와 HPV 유전자 타겟인 HPV\_1의 경우 MT2 조건보다 더 많은 세포사멸을 유도하는 것을 확인하였다. 구체적으로, MT2, PRDM9 및 HPV\_1에서 약 50%, 약 65% 및 약 65%의 세포사멸률을 나타내었다.

[0165] 따라서, HeLa 세포만 가지고 있는 CNV와 HPV 유전자를 Cas9 단백질을 이용하여 타겟 DNA Break를 유도하였을 때 세포사멸이 일어나는 것을 확인하였다.

[0166] **실시예 12.3. HPV 유전자 타겟에 의한 세포 사멸 확인**

[0167] 전기충격으로 HeLa 세포에 DNA 벡터를 도입하는 과정은 실시예 12.1의 방법과 동일하다. 기존의 1 copy의 non essential gene인 CCR5과 100개 이상의 non essential gene을 타겟하는 MT2 조건에 인간 유전체에 없는 영역을 타겟하는 NT 조건을 추가하였다. NT1의 경우 20 mer의 non-target sgRNA를 발현하는 조건, NT2 및 NT3은 non-target sgRNA의 spacer의 길이가 각각 10 mer, 5 mer인 조건이다. 실험군으로 HeLa에서 30 copy 이상 있는 것으로 알려진 Human Papilloma Virus(HPV) 유래의 유전자를 타겟하여 HeLa세포 특이적 세포사멸을 유도하는 실험을 진행하였다. 전기충격으로 DNA를 도입한 후, 24 웰에 2 반복으로 세포를 넣어준 후, 24시간이 지나 0.5 μg/mL Puromycin을 넣어 DNA 벡터가 안 들어간 세포들을 죽이는 셀렉션 과정을 진행하였다. 셀렉션 3일 이후, Puromycin이 안 들어간 세포배양액으로 바꿔주어 세포를 배양하고, 10일 후, 각 웰에서 세포를 떼어내어 Cell Titer Glo 방법으로 luminescence signal을 측정하였다. 그 결과를 도 20b에 나타내었다.

[0168] 도 20b에 나타난 바와 같이, HeLa 세포에서 1곳만 자르는 CCR5 타겟은 5 mer의 spacer를 가지는 NT3와 비교하였을 때 75% 정도의 luminescence signal을 보였고, NT3에 비하여 MT2 조건에서 약 99.5% 세포사멸이 유도되는 것을 확인할 수 있었다. 또한, HPV 유전자 타겟인 HPV\_1의 경우 NT3 조건과 비교하였을 때, 약 90%의 세포를 죽이는 것을 확인할 수 있었다. 따라서, HeLa 세포만 가지고 있는 HPV 유전자를 Cas9 단백질을 이용하여 타겟 DNA Break를 유도하였을 때 세포사멸이 일어나는 것을 Cell Titer Glo 방법을 통해 확인하였다.

[0169] **실시예 13. 대장암 세포 HT29에서 표적 위치에 따른 세포 사멸 효과 확인**

[0170] **실시예 13.1. CNV 타겟에 의한 대장암 세포 사멸 확인**

[0171] 24 웰 플레이트에서, 웰 1개당 H1563 세포를 1.5x10<sup>5</sup>개 만큼 깔아주고, 24시간 후에 각 웰에 하기 표 9와 같은 종류의 DNA를 도입하였다. 도입을 위해 Lipofectamine3000 reagent를 사용하였고, 제조사에서 제공하는 매뉴얼을 따랐으며, DNA를 500 ng씩 사용하였다.

**표 9**

[0172]

조건	Lipo	GFP	NT1	CCR5	HPRT1	MT2	TRAPPC9	LINC00536	TRPS1	CDK8
Copynumber	N/A	0	2	2	2	>100	>13	>9	>8	>18
Essentiality	N/A	N/A	N/A	Non-essential	House-keeping	Non-essential	Non-essential	Non-essential	Non-essential	Oncogene

[0173] 상기 DNA 도입 즉, 트랜스펙션 시점을 기점으로 하여 24시간 후에 1 μg/mL 농도의 puromycin을 처리하여 90시간 동안 셀렉션하였다. 이후, 각 웰의 배양액을 모두 석션으로 제거하고, 12일 동안 normal media(McCOY+10% FBS, 1% P/S)로 리커버(recover)하였다. 이후, 1X PBS를 500 μL만큼 넣어 한 번 세척해준 후, Trypsin-EDTA을 처리하여 세포를 모두 떼어내었다. 이들을 Trypan blue 염색약을 이용하여 염색한 후, 살아있는 세포 수를 두 번 측정하여 평균 처리하였다. 그 결과를 도 21에 나타내었다.

[0174] 도 21에 나타난 바와 같이, 표적 시퀀스가 없다고 알려진 NT1과 단일 표적인 CCR5는 서로 오차 범위 이내의 차이를 나타내었다. HPRT1은 NT1 대비 47% 정도 수준의 세포가 남았으며, genome 전체에서 repeat sequence를 100 군데 이상 자르는 positive control인 MT2와 네 개의 CNV 표적(TRAPPC, LINC00536, TRPS1 및 CDK8)들은 차례대로 NT1 대비 2%, 3%, 18%, 20% 및 4.2%의 세포 생존율을 나타내었다. 즉, CNV 표적은 암세포를 효과적으로 사멸시킬 수 있음을 확인하였다.

[0175] **실시예 13.2. CNV 타겟에 의한 대장암 세포 사멸 확인**

[0176] CNV를 가지는 4종의 CDK8, LINC00536, TRPS1 및 TRAPPC9 유전자 및 MT2를 표적으로 하여 HT-29(결장암 세포주)의 특이적 사멸을 확인하기 위해 Cas9 단백질과 각 CNV 타겟 gRNA를 발현하는 DNA 500 ng을 전기천공법(Lonza)으로 HT-29 세포에 도입하였다. 배양중인 HT-29 세포를 1X PBS로 세척한 후, Trypsin-EDTA를 처리하여 바닥에서 떼어내었다. 필요한 세포 수를 덜어내어 1X PBS로 1번 세척한 후, SF buffer(Lonza)로 풀어준 후 각 DNA와 섞었다. 세포와 DNA 혼합물을 전기천공장치(Lonza)에 넣고 전기충격을 가하였다. 대조군으로 인간 게놈에서 정렬되지 않은 NT1(non target) 및 전기충격만주는 조건(pulse only)을 추가하였다. 전기충격 후, 96 웰에 각 조건마다 4 반복으로 플레이팅(plating)하였다. DNA 도입 24시간 후, 50 µL의 CellTiter Glo reagent를 96 웰에 넣어주었다. FLUOstar omega reader 기계에 플레이트를 넣고 2분간 흔들어 주고, 10분 동안 상온에서 반응시킨 후 발광도를 측정하였다. 그 결과를 도 22에 나타내었다.

[0177] 도 22에 나타난 바와 같이, 대조군에 비해 100곳 이상을 타겟하는 MT2 조건과 TRAPPC9 CNV 타겟에서 90%의 세포 사멸이 유도되었고, CDK8, LINC00536 및 TRPS1 3종의 CNV 타겟은 20% 내지 45%의 세포사멸을 유도하였다.

[0178] **실시예 14. 폐암 세포 H1563 세포에서 H1299에서 사용된 타겟 유전자에 의한 세포 사멸 효과 비교**

[0179] 24 웰 플레이트에서, 웰 1개당 H1563 세포를 1.5x10<sup>5</sup>개 만큼 깔아주고, 24시간 후에 각 웰에 하기 표 10과 같은 종류의 DNA를 도입하였다. 도입을 위해 Lipofectamine3000 reagent를 사용하였고, 제조사에서 제공하는 매뉴얼을 따랐으며, DNA를 500 ng씩 사용하였다.

표 10

조건	CCR5	SMIM11	GNPDA2	SLC15A5	KCNE2
cut number	2	2	2	2	2
essentiality	Non essential	Non essential	oncogene	oncogene	Non essential

[0181] 상기 DNA 도입 즉, 트랜스펙션 시점을 기점으로 하여 24시간 후에 1 µg/mL 농도의 puromycin을 처리하여 72시간 동안 셀렉션하였다. 이후, 각 웰의 배양액을 모두 석션으로 제거하고, 1X PBS를 500 µL만큼 넣어 한 번 세척해준 후 Trypsin-EDTA를 처리하여 세포를 모두 떼어내었다. 이들을 Trypan blue 염색약을 이용하여 염색한 후, 살아있는 세포 수를 측정하였다. 그 결과를 도 23에 나타내었다. 이때, 사용한 DNA 표적들은 CCR5를 제외하고 모두 H1299에서 증폭되어있는 표적들이므로 genome 상에서 12군데 이상을 잘라, 사멸 효과를 보았던 표적들(도 8 및 도 9 참조)이다. 그러나, H1563 세포에서는 이와 같이 증폭된 CNV가 존재하지 않으므로, 상기 표적들은 모두 CCR5와 같이 두 군데의 표적 위치만을 자르게 된다.

[0182] 도 23에 나타난 바와 같이, 폐암세포 H1563에서는 이들 표적들이 폐암세포 H1299에서와 같이 사멸 효과를 보이지 않았으며, CCR5에 대비했을 때 오히려 더 높은 세포 생존율을 보여주었다. 다만, KCNE2의 경우, H1563에서도 높은 세포 사멸 효과를 나타내는 것으로 관찰되었다. 이는 도 24의 현미경상 이미지에서 유추되는 세포 사멸 트렌드와도 일치하는 결과였다. KCNE2의 세포 사멸 효과는 알 수 없는 모종의 이유에서 일어났을 것으로 추측되며, 이를 제외한 나머지 H1299의 세 개의 CNV 표적들(SMIM11, GNPDA2 및 SLC15A5)이 H1563에서 세포 사멸을 일으키지 않았으므로 CNV의 세포 특이적 사멸 효과를 확인할 수 있었다.

[0183] **실시예 15. CNV 타겟에 의한 폐암세포 H1299 세포의 세포사멸 측정**

[0184] 형질전환 하루 전, 폐암세포주인 H1299 세포를 Trypsin-EDTA로 떼어낸 후, white 96 웰에 1.3x10<sup>4</sup>을 플레이팅하였다. 다음 날, Cas9 단백질과 각 CNV 타겟 gRNA가 발현되는 DNA 500 ng를 리포솜을 이용한 방법으로 도입하였다. 0.3 µL의 리포솜 reagent I와 5 µL의 opti-MEM을 섞어두었다(tube 1). 0.2 µL의 리포솜 reagent II, 5 µL의 opti-MEM, 및 각 조건의 DNA 500 ng를 섞어주어 tube 2를 제작하였다. 이후, tube 1과 혼합하여 상온에

서 15분 동안 방치하였다. 리포솜과 DNA 혼합액을 96 웰에 11  $\mu$ L씩 넣어주었다. 3시간 후, AnnV reagent를 넣고, 형질전환 24시간 후 발광 정도를 측정하였다. 그 결과를 도 25에 나타내었다. 상기 방법은 세포사멸이 일어날 때 바깥쪽 세포막에 노출되는 PS(phosphatidylserine) site에 AnnV가 붙어 발광하게 되면, 발광 정도에 따라 세포 사멸의 정도를 측정하는 방법이다.

[0185] 상기 실험 결과, KCNE2, GNPDA2, SMIM11 및 SLC15A5에서 약 30% 내지 40%의 세포사멸률을 나타내었다. 상기 결과를 통해, H1299 세포가 가지고 있는 CNV 4종(GNPDA2, KCNE2, SLC15A5 및 SMIM11)을 Cas9 단백질을 이용하여 타겟 DNA break를 유도했을 때, 세포막에 PS가 노출되면서 세포사멸이 일어나는 것을 확인할 수 있었다.

[0186] **실시예 16. EGFR 돌연변이 서열 특이적 가이드 RNA에 의한 세포 사멸 효과 및 Cas9-RecJ 융합단백질(CRISPR PLUS)을 이용한 세포사멸 효과 확인**

[0187] 본 실시예에서는 Cas9 단백질 발현 벡터(PX459, Addgene plasmid #62988)를 이용해 폐암세포인 HCC827 세포의 유전체에 multi-cleavage를 일으켜 세포 사멸을 일으킬 수 있음을 증명하였다. 또한, 인간 코돈 최적화된 Rec J 단백질을 PX459 벡터에 함께 발현시켜 세포 사멸 효과를 증폭시킬 수 있음을 확인하였다. 전기천공법으로 세포 내부에 벡터들을 전달하였으며, 유전체에 multi-cleavage를 일으키기 위해 HCC827세포 EGFR 유전자의 E2 돌연변이 서열을 타겟하는 가이드 RNA를 사용하였다.

[0188] HCC827 세포는 75T 플라스크에서 RPMI-1640(10% 소태아혈청) 배지로 약 50% 밀도의 양으로 배양한 후 트립신화 및 PBS 세척을 거쳐 최종적으로 Neon Electroporation Buffer R에 리서스펜드 하였다. 10  $\mu$ L Neon 파이펫 팁에 150,000개의 세포와 500 ng의 벡터를 로딩한 후, 1300 V, 20 ms 및 2 pulses 조건으로 전기천공법을 진행하였다. 이후, RPMI-1640(10% 소태아혈청) 배지에서 회복시키고, 4일 후 세포들을 수확해 카운팅을 실시하였다. 그 결과를 도 26 및 도 27에 나타내었다.

[0189] 도 26 및 도 27에 나타난 바와 같이, 대조군으로 사용한 electric pulse만 준 샘플(pulse only)과 비교해 봤을 때, 타겟 서열이 존재하지 않는 EGFR\_WT 실험군에 비해 EGFR\_E2 실험군의 세포수가 현저하게 감소하였다. 구체적으로, Cas9을 사용하여 CCR5를 타겟한 실험군의 경우 세포 수의 변화가 없었으나, EGFR\_E2를 타겟한 실험군의 경우 약 33%의 세포사멸률을 나타내었다. 또한, Rec J 단백질을 함께 발현 시켰을 시 multi-cleavage에 의한 세포 사멸 효과가 증폭 되는 것뿐만 아니라, single target을 절단했을 때에도 세포 수가 줄어드는 현상을 관찰하였다. 구체적으로, Cas9-RecJ를 사용하여 CCR5를 타겟한 실험군의 경우 약 50%의 세포사멸률을 나타내었고, Cas9-RecJ를 사용하여 EGFR\_E2 실험군의 경우 80%의 세포사멸률을 나타내었다. 상기 실험을 통해, 암세포에 Cas9 단백질과 세포 유전체 서열 특이적 multi-target 가이드 RNA를 넣어 주었을 때 세포 사멸을 일으키며 CRISPR PLUS 단백질로 그 효과를 조절할 수 있음을 알 수 있었다.

[0190] **실시예 17. EGFR 돌연변이 서열 특이적 가이드 RNA에 의한 세포 사멸 효과 및 RNP를 이용한 세포사멸 효과 확인**

[0191] 본 실시예에서는 Cas9/sgRNA ribonucleoprotein(Cas9 RNP)를 이용해 폐암세포인 HCC827 세포의 유전체에 multi-cleavage를 일으켜 세포 사멸을 일으킬 수 있음을 증명하였다. 전기천공법으로 세포 내부에 RNP들을 전달 하였으며, 유전체에 multi-cleavage를 일으키기 위해 HCC827세포 EGFR 유전자의 E2 돌연변이 서열을 타겟하는 가이드 RNA를 사용하였다.

[0192] HCC827 세포는 75T 플라스크에서 RPMI-1640(10% 소태아혈청) 배지로 약 50% 밀도의 양으로 배양한 후, 트립신화 및 PBS 세척을 거쳐 최종적으로 Neon Electroporation Buffer R에 리서스펜드 하였다. 10  $\mu$ L Neon 파이펫 팁에 150,000개의 세포와 1.2  $\mu$ M의 RNP를 로딩한 후, 1300 V, 20 ms 및 2 pulses의 조건으로 전기천공법을 진행 하였다. 이후, RPMI-1640(10% 소태아혈청) 배지에서 회복시키고, 1일 후 세포들을 수확해 카운팅을 실시하였다. 그 결과를 도 28에 나타내었다.

[0193] 도 28에 나타난 바와 같이, electric pulse만 준 샘플(pulse only)이나, Cas 단백질 혹은 가이드 RNA만 넣은 샘플들과 비교해 봤을 때, EGFR\_E2 서열을 타겟하는 RNP를 넣은 실험군의 세포수가 현저하게 감소하였다. 구체적으로, 상기 EGFR\_E2 서열을 타겟하는 RNP를 넣은 실험군의 경우, 약 35%의 세포사멸률을 나타내었다. 상기 실험을 통해, 암세포에 Cas9 단백질과 세포 유전체 서열 특이적 multi-target 가이드 RNA를 넣어 주었을 때 세포 사멸을 일으키며 Cas9 RNP로 그 효과를 조절할 수 있음을 확인할 수 있었다.

[0194] **실시예 18. MT2 서열 특이적 가이드 RNA에 의한 세포 사멸 효과 및 RNP를 이용한 세포사멸 효과 확인**

[0195] 본 실시예에서는 Cas9/sgRNA ribonucleoprotein(Cas9 RNP)를 이용해 폐암세포인 H1563 세포의 유전체에 multi-cleavage를 일으켜 세포 사멸을 일으킬 수 있음을 증명하였다. 전기천공법으로 세포 내부에 RNP들을 전달 하였

으며, 유전체에 multi-cleavage를 일으키기 위해 인간 유전체에서 약 100군데 이상을 타겟할 수 있는 가이드 RNA인 MT2를 사용하였다.

[0196] H1563 세포는 75T 플라스크에서 RPMI-1640(10% 소태아혈청) 배지로 약 50% 밀도의 양으로 배양한 후, 트립신화 및 PBS 세척을 거쳐 최종적으로 Neon Electroporation Buffer R에 리서스펜드 하였다. 10 μL Neon 파이펫 팁에 150,000개의 세포와 1.2 μM의 RNP를 로딩한 후, 1200 V, 20 ms 및 2 pulses의 조건으로 전기천공법을 진행하였다. 이후, RPMI-1640(10% 소태아혈청) 배지에서 회복시키고, 2일 후 세포들을 수확해 카운팅을 실시하였다. 그 결과를 도 29에 나타내었다.

[0197] 도 29에 나타난 바와 같이, electric pulse만 준 샘플(pulse only)이나, 가이드 RNA만 넣은 샘플들과 비교해 봤을 때, MT2 서열을 타겟하는 RNP를 넣은 실험군의 세포수가 현저하게 감소한 것을 확인 하였다. 구체적으로, 상기 MT2 서열을 타겟하는 RNP를 넣은 실험군의 경우, 약 35%의 세포사멸률을 나타내었다. 상기 실험을 통해, 암세포에 Cas9 단백질과 세포 유전체 서열 특이적 multi-target 가이드 RNA를 넣어 주었을 때 세포 사멸을 일으키며 RNP로 그 효과를 조절할 수 있음을 확인할 수 있었다.

[0198] **실시예 19. MT2, GNPDA2 및 SMIM11 서열 특이적 가이드 RNA에 의한 세포 사멸 효과 및 RNP를 이용한 세포사멸 효과 확인**

[0199] 본 실시예에서는 Cas9 단백질을 이용해 폐암세포인 H1299 세포의 유전체에 multi-cleavage를 일으켜 세포 사멸을 일으킬 수 있음을 증명하였다. 여러 암세포들 중에서 트랜스팩션 효율이 비교적 높은 H1299 세포를 실험에 이용하였으며 전기천공법으로 세포 내부에 RNP들을 전달하였다.

[0200] 유전체에 multi-cleavage를 일으키기 위한 가이드 RNA로 세 종류의 가이드 RNA를 사용하였다: MT2, GNPDA2 및 SMIM11. H1299 세포 특이적 유전체 서열 분석 정보를 토대로 copy-number variation(CNV) 이 높은 유전자들 중 약 12 copy 이상 존재하는 oncogene(GNPDA2)과 약 40 copy 이상 존재하는 non-essential 유전자(SMIM11)를 타겟하는 가이드 RNA를 제작하였다. 인간 유전체 서열에 존재하는 모든 Cas9 proto-spacer adjacent motif(PAM, 5'-NGG-3')를 분석해 100군데 이상을 타겟할 수 있는 가이드 RNA인 MT2를 제작하였다. 또한, 트랜스팩션 유무를 확인하기 위해 C-terminus 부분에 GFP가 달린 Cas9 단백질을 제작해 사용하였다.

[0201] H1299세포는 75T 플라스크에서 RPMI-1640(10% 소태아혈청) 배지로 약 50% 밀도의 양으로 배양한 후, 트립신화 및 PBS 세척을 거쳐 최종적으로 Neon Electroporation Buffer R에 리서스펜드 하였다. 10 μL Neon 파이펫 팁에 150,000개의 세포와 1.2 μM Cas9-GFP 단백질과 1.5 μM 가이드 RNA로 만든 RNP complement를 로딩한 후, 1300 V, 20 ms 및 2 pulses의 조건으로 전기천공법을 진행하였다. 이후, RPMI-1640(10% 소태아혈청) 배지에서 회복시키고, 이를 후 Promega사의 CellTiter-Glo 2.0을 이용해 살아있는 세포 신호(luminescence)를 이용하여 세포 수를 관찰해 비교하였다. 그 결과를 도 30 내지 도 32에 나타내었다.

[0202] 도 30 내지 도 32에 나타난 바와 같이, 단백질만 넣거나 가이드 RNA만 넣어준 샘플과 비교해 봤을 때, single-target이 존재하는 것으로 알려진 CCR5 타겟 실험군에서는 유의미한 세포 사멸이 관찰되지 않았다. 하지만, multi-cleavage를 일으키는 RNP를 넣어준 MT2, GNPDA2 및 SMIM11 실험군에서는 유의미한 세포사멸이 관찰되었다. 구체적으로, 상기 RNP를 넣어준 MT2, GNPDA2 및 SMIM11 실험군에서 각각 약 33%, 약 71% 및 약 40%의 세포사멸률을 나타내었다. 상기 실험을 통해 암세포에 Cas9 단백질과 세포 유전체 서열 특이적 multi-target 가이드 RNA를 넣어 주었을 때 세포 사멸을 일으키며 RNP로 그 효과를 조절할 수 있음을 확인할 수 있었다.

[0203] **실시예 20. Cas12a 단백질을 이용한 서열 특이적 세포 사멸 효과 및 돌연변이 CNV 서열 특이적 세포 사멸 효과 확인**

[0204] 본 실시예에서는 Cas12a 단백질 발현 박터를 이용해 폐암세포인 HCC827 세포의 유전체에 double 혹은 multi-cleavage를 일으켜 세포 사멸을 일으킬 수 있음을 증명하였다. 본 발명에서, Cas12a의 DNA 및 아미노산 서열을 서열번호 135 및 136으로 나타내었다. 전기천공법으로 세포 내부에 박터들을 전달하였으며, 야생형 non-essential 유전자인 CCR5(서열번호 138), 혹은 HCC827 세포에서 18 카피 이상이 존재하는 것으로 알려진 EGFR\_E2 돌연변이 서열을 타겟하는 crRNA(서열번호 139)를 사용하였다.

[0205] HCC827 세포는 75T 플라스크에서 RPMI-1640(10% 소태아혈청) 배지로 약 50% 밀도의 양으로 배양한 후 트립신화 및 PBS 세척을 거쳐 최종적으로 Neon Electroporation Buffer R에 리서스펜드 하였다. 10 μL Neon 파이펫 팁에 150,000개의 세포와 500 ng의 박터를 로딩 후 1300 V 20 ms 2 pulses 조건으로 전기천공법을 진행하였다. 이후 RPMI-1640(10% 소태아혈청) 배지에서 회복시키고, 6일 후 세포들을 수확해 발광 신호를 측정함으로써 세포

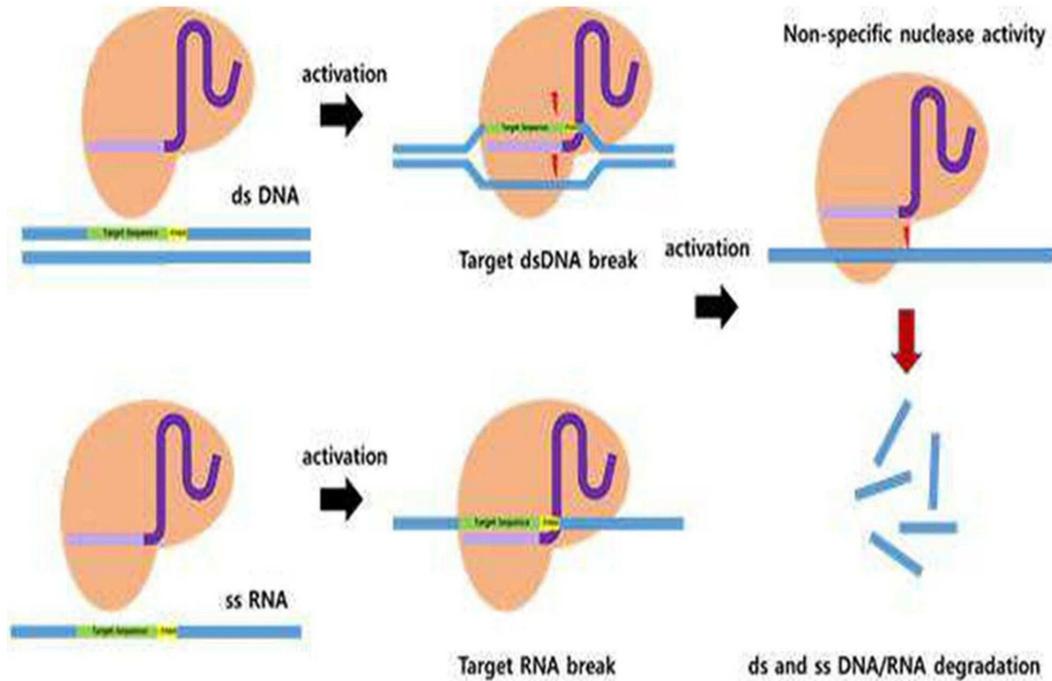
양을 정량하였다. 한편, 대조군으로 사용한 EGFR\_WT 서열(서열번호 137)은 HCC827 세포에는 존재하지 않는 것으로 알려진 서열이다. 그 결과를 도 33 및 도 34에 나타내었다.

[0206]

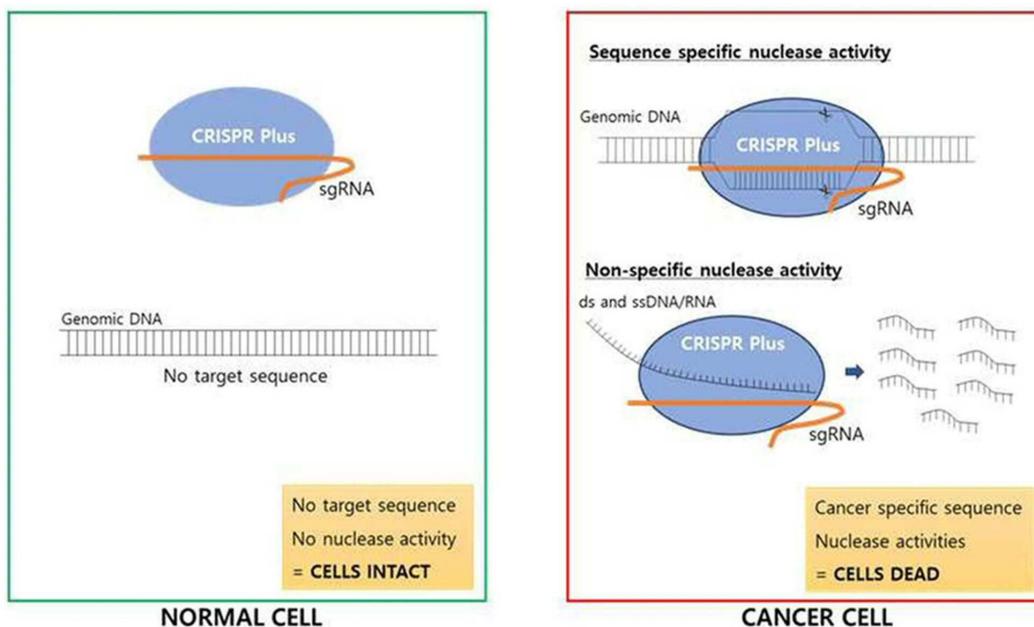
도 33 및 도 34에 나타난 바와 같이, 대조군과 비교해 봤을 때, CCR5을 타겟한 실험군에서는 세포가 약 76% 사멸하였고, EGFR\_E2를 타겟한 실험군에서는 세포가 약 83% 사멸하였다. 본 실험을 통해 copy number에 관계없이 암세포 특이적 서열을 Cas12a 단백질을 이용해 절단했을때, 타겟 특이적 세포사멸을 일으킬 수 있음을 확인할 수 있었다.

도면

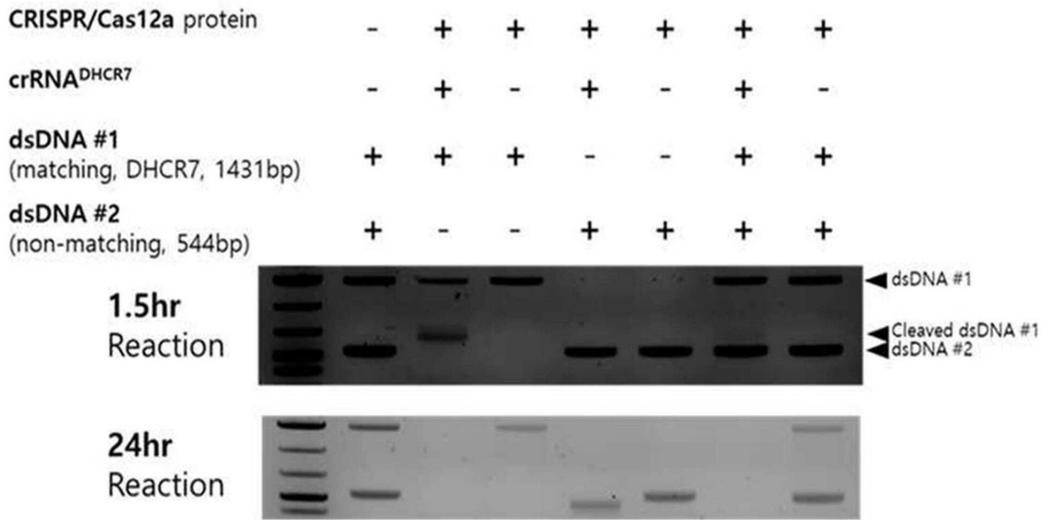
도면1



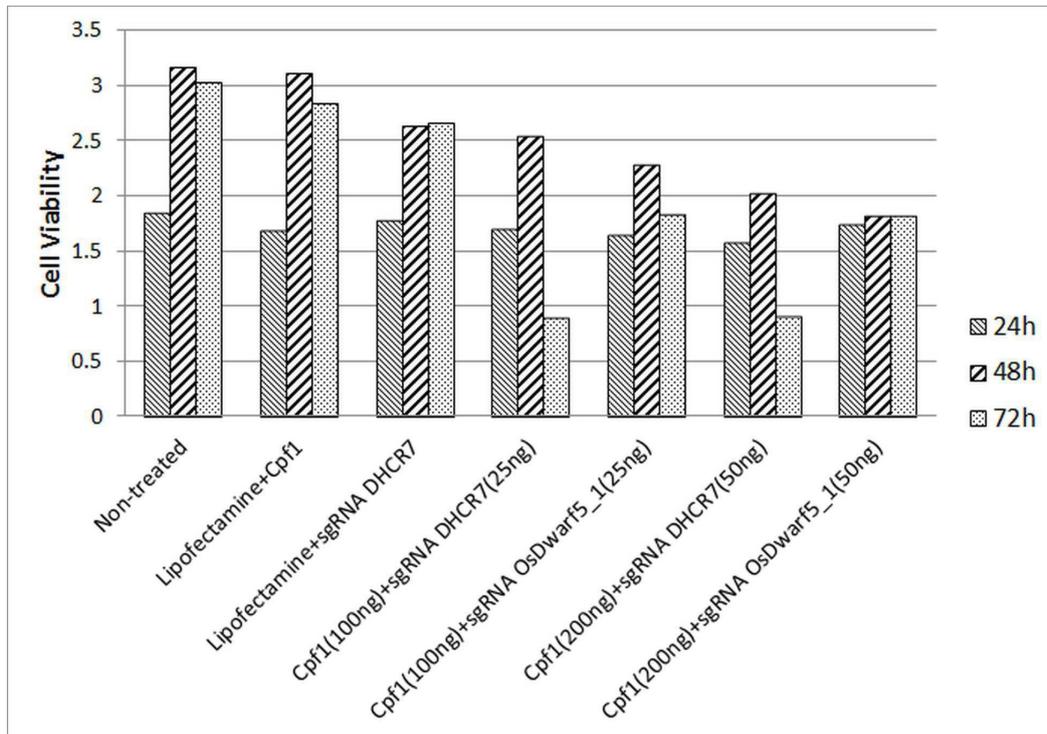
도면2



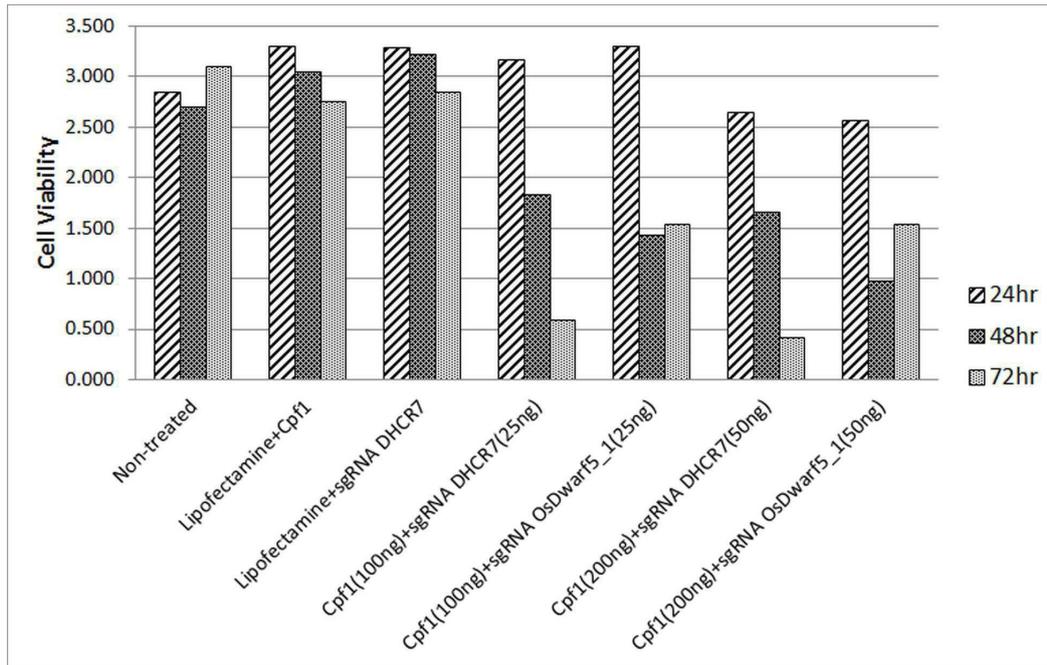
도면3



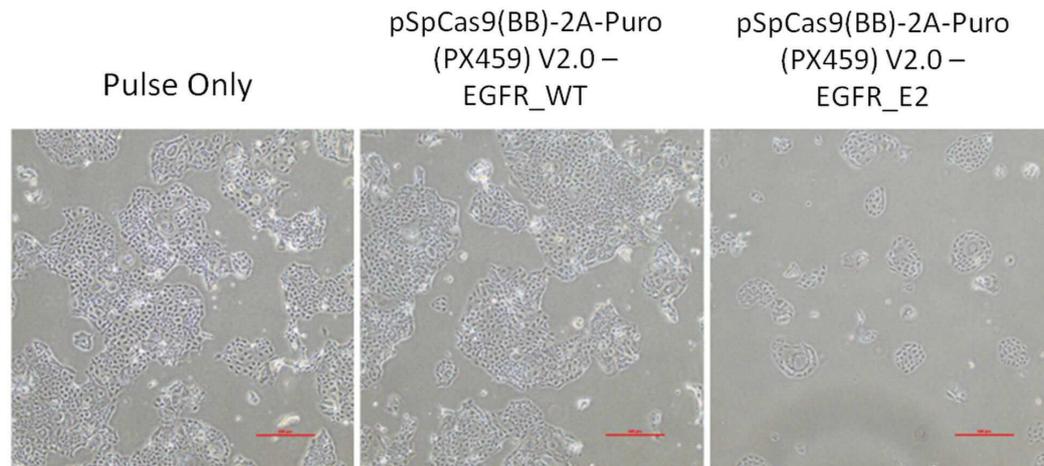
도면4a



도면4b

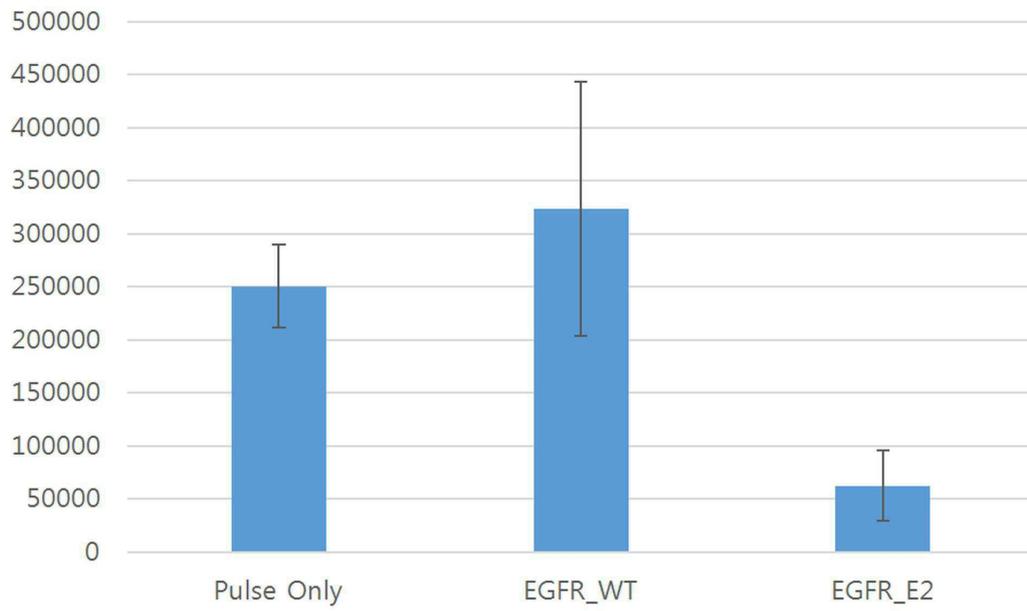


도면5

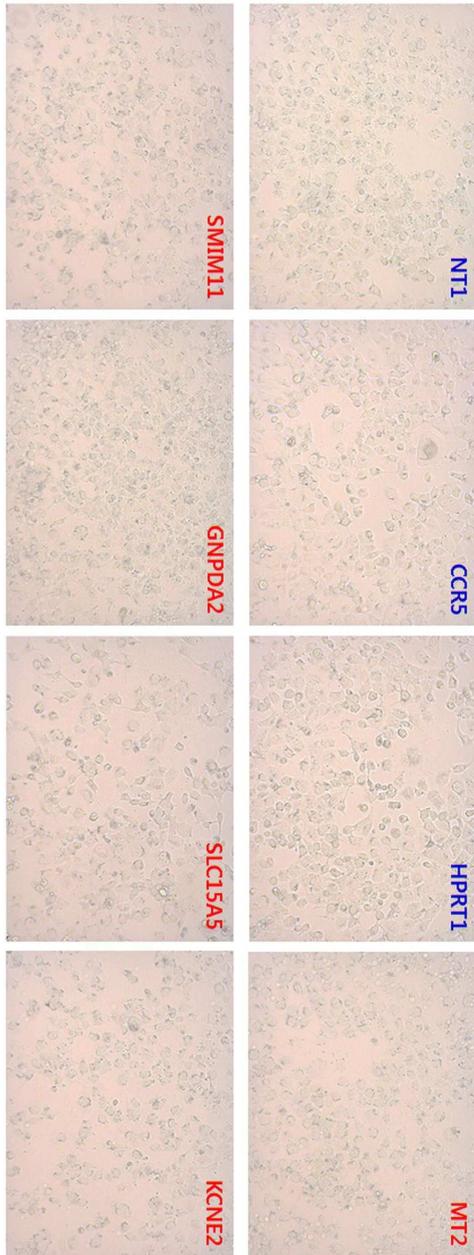


도면6

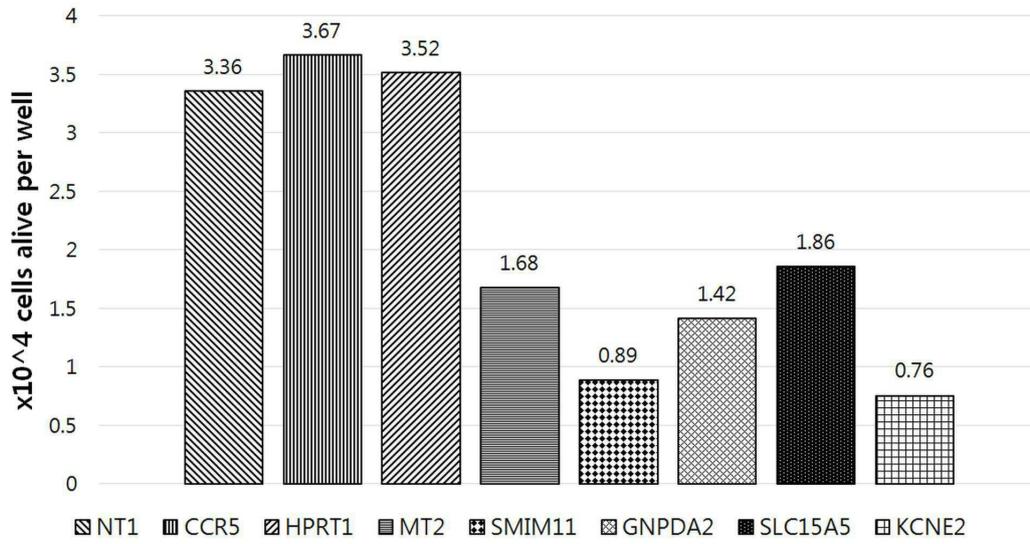
HCC827 세포 수 비교 (6dpt)



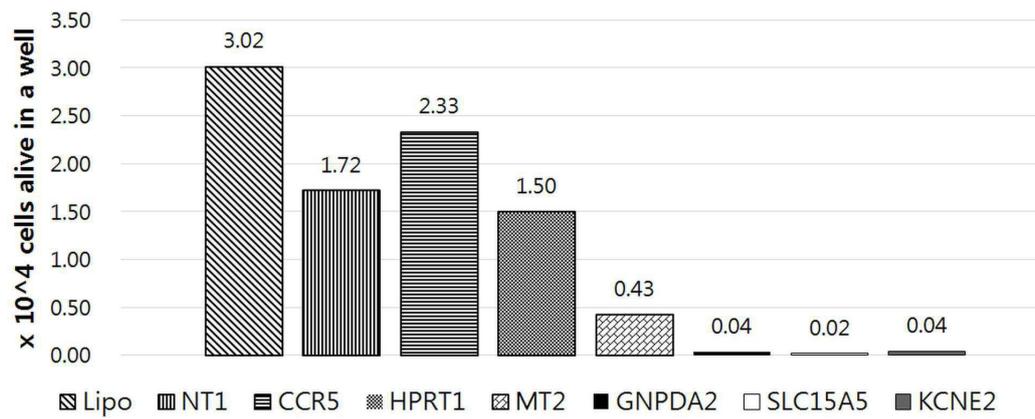
도면7



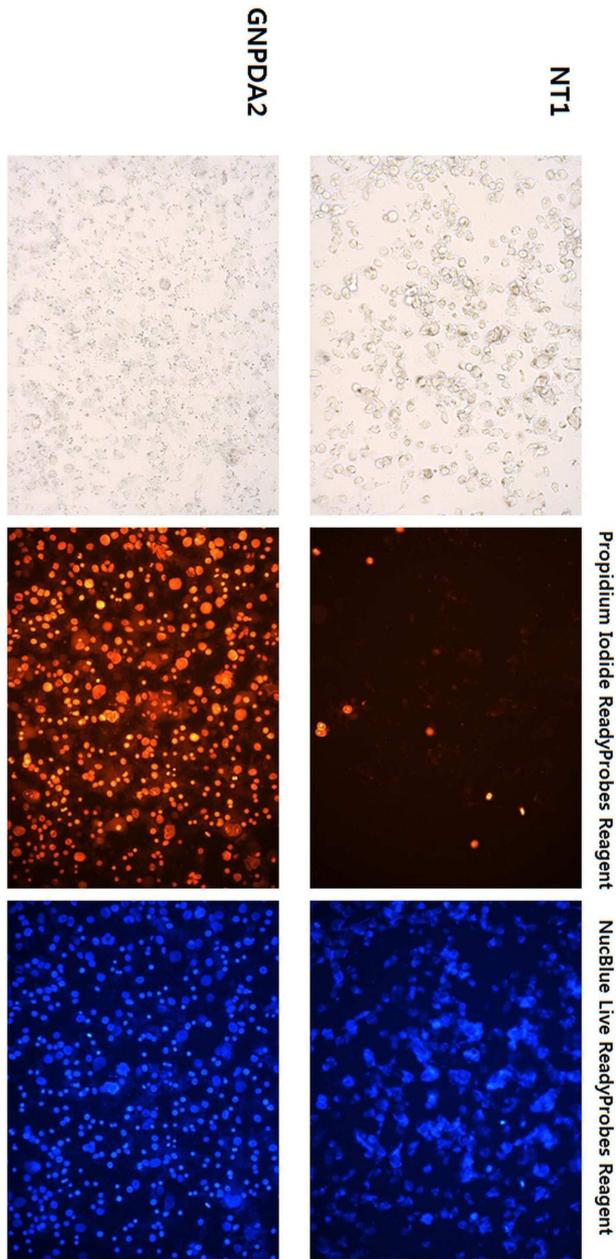
도면8



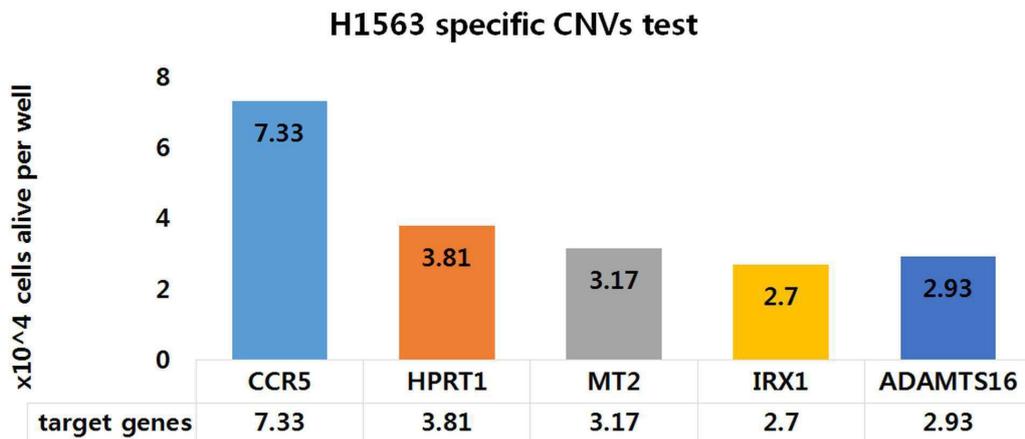
도면9



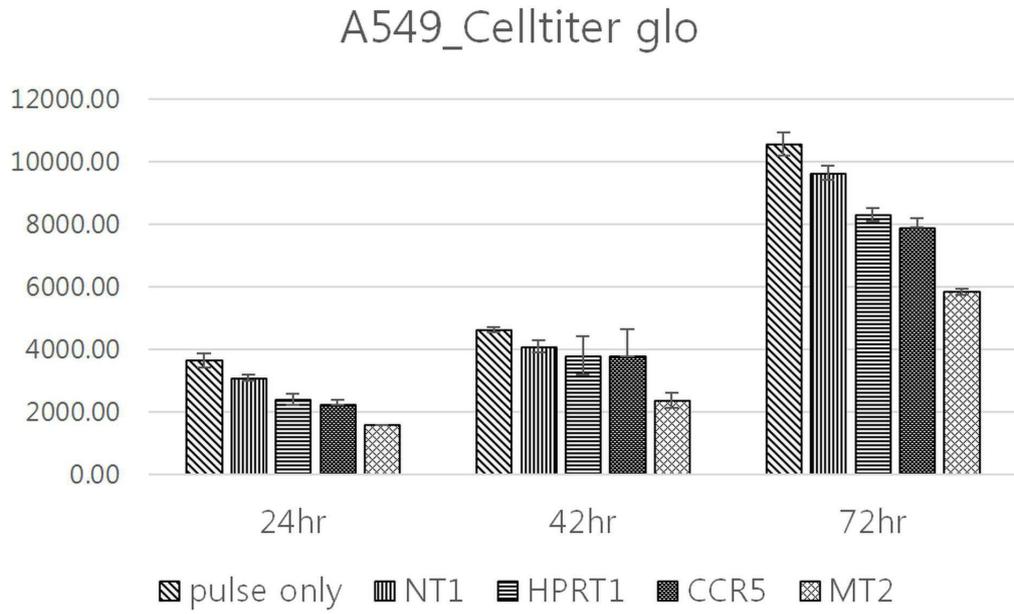
도면10



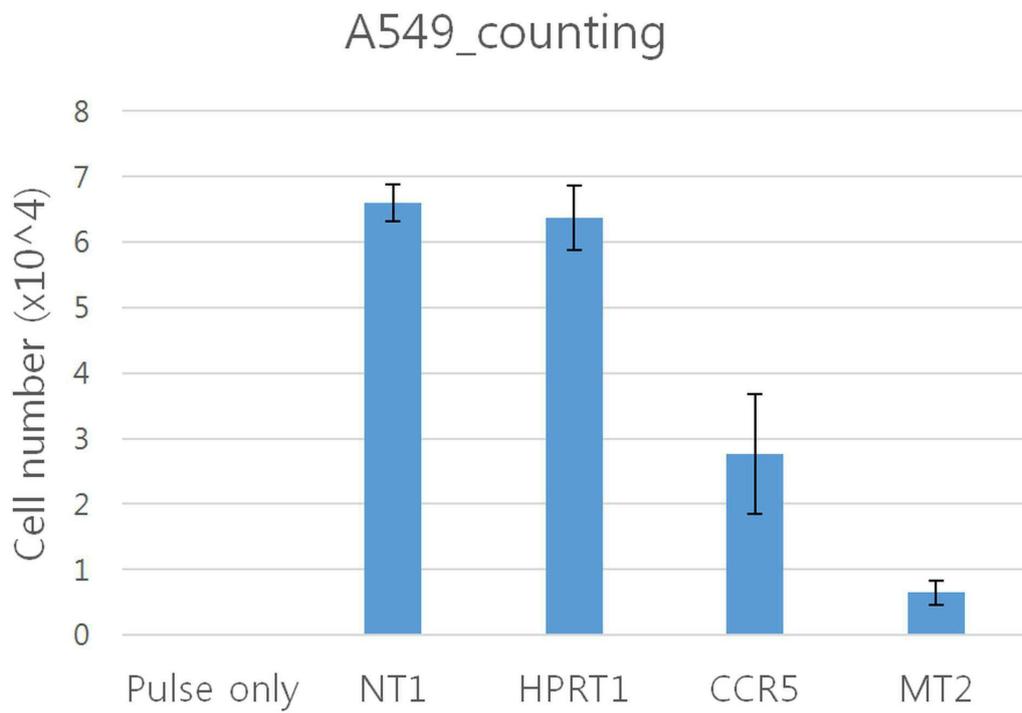
도면11



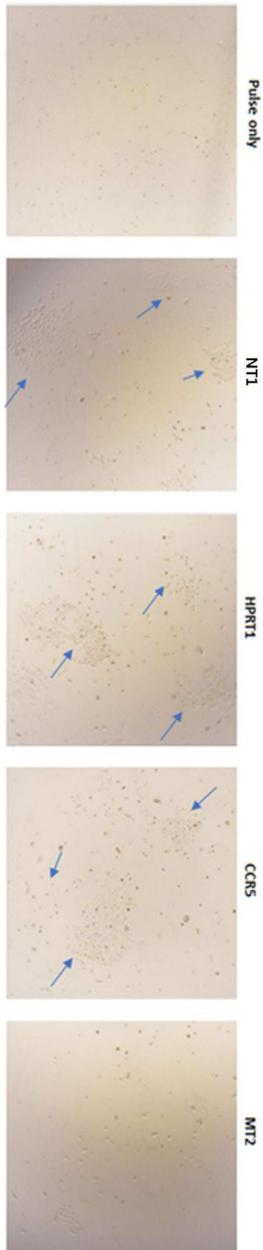
도면12



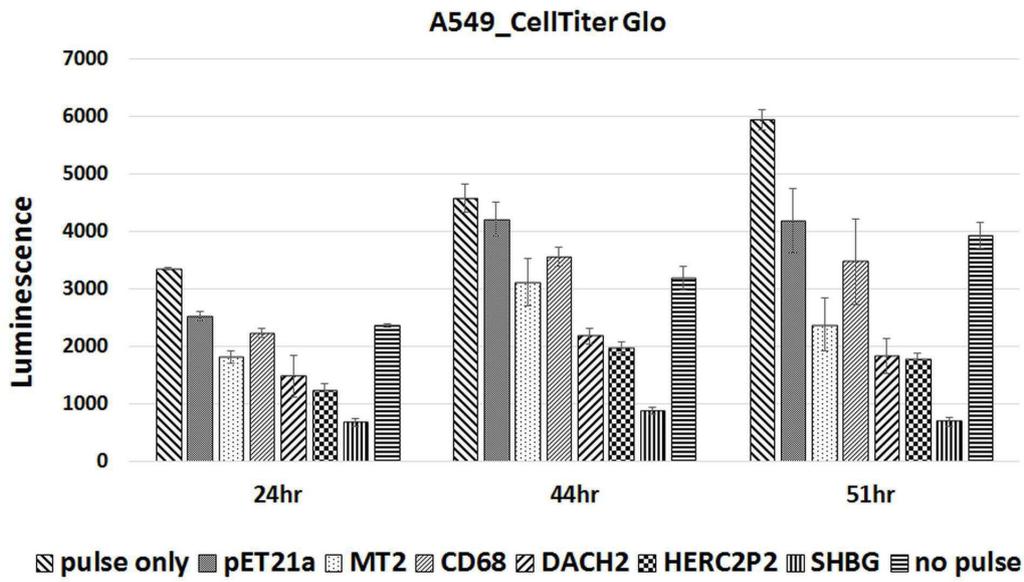
도면13



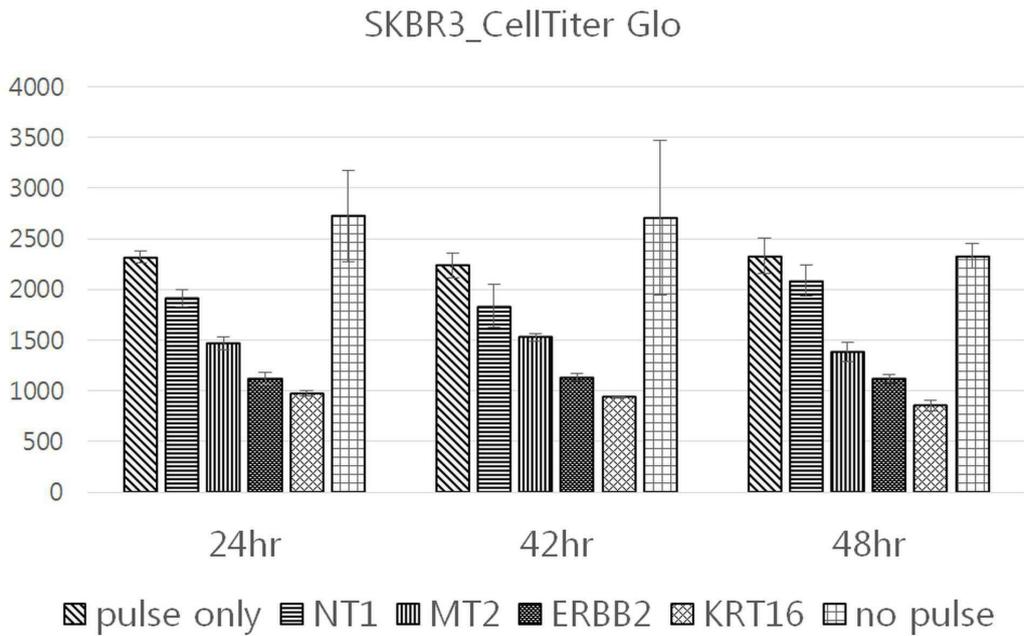
도면14



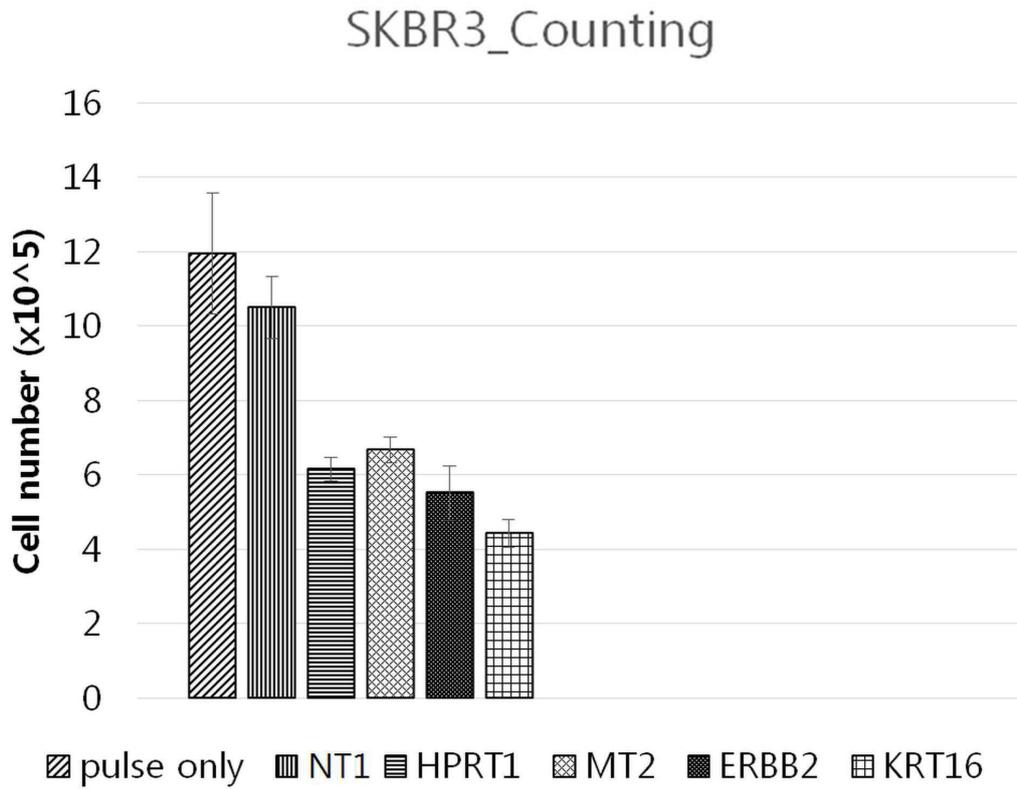
도면15



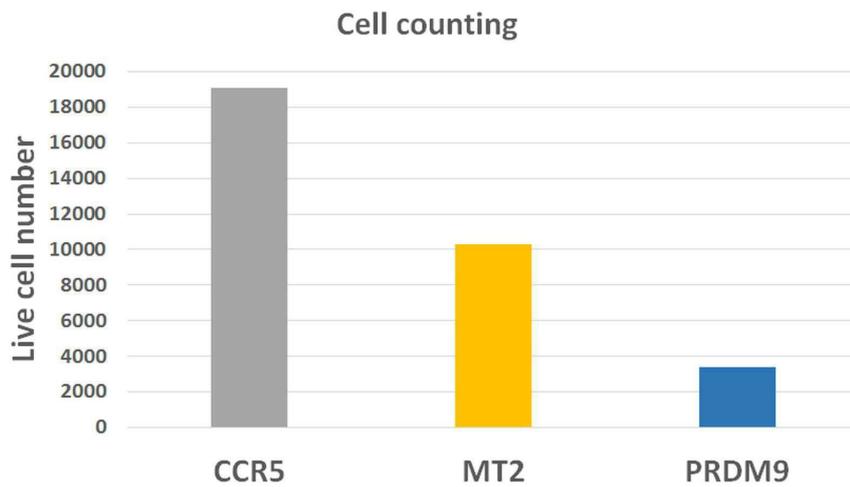
도면16



도면17

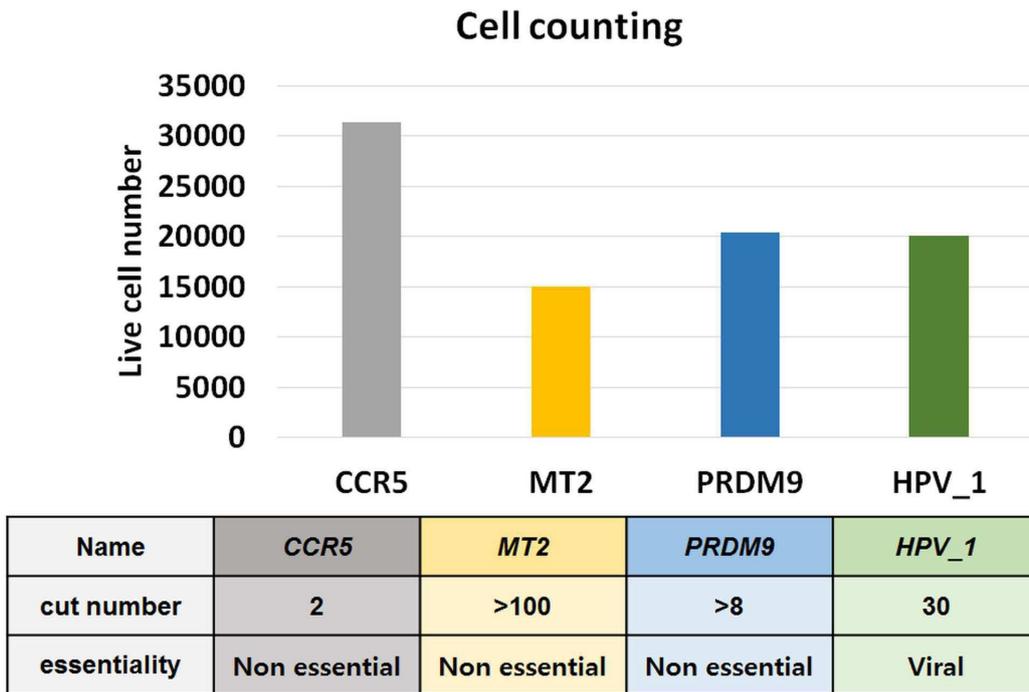


도면18

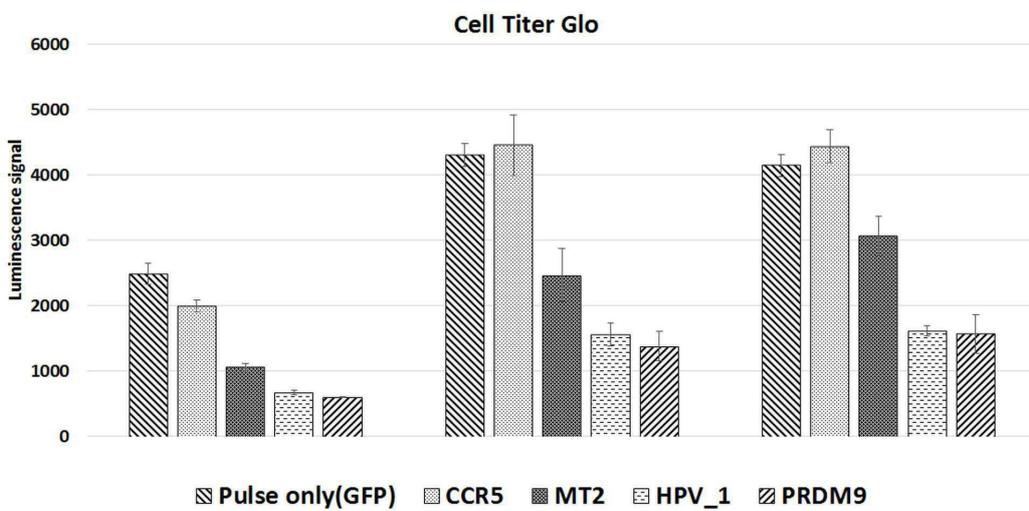


Name	CCR5	MT2	PRDM9
cut number	2	>100	>8
essentiality	Non essential	Non essential	Non essential

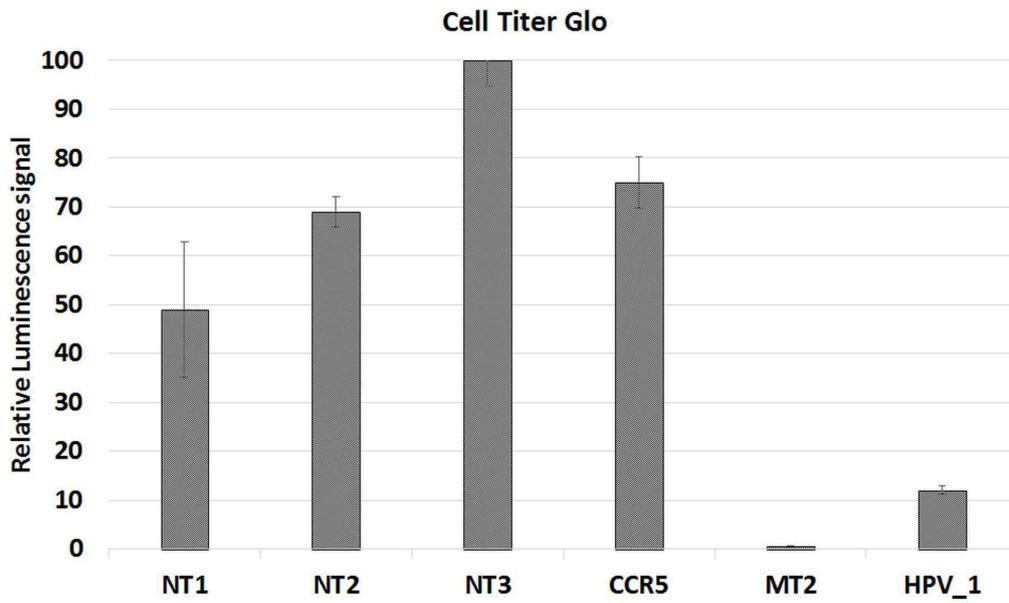
도면19



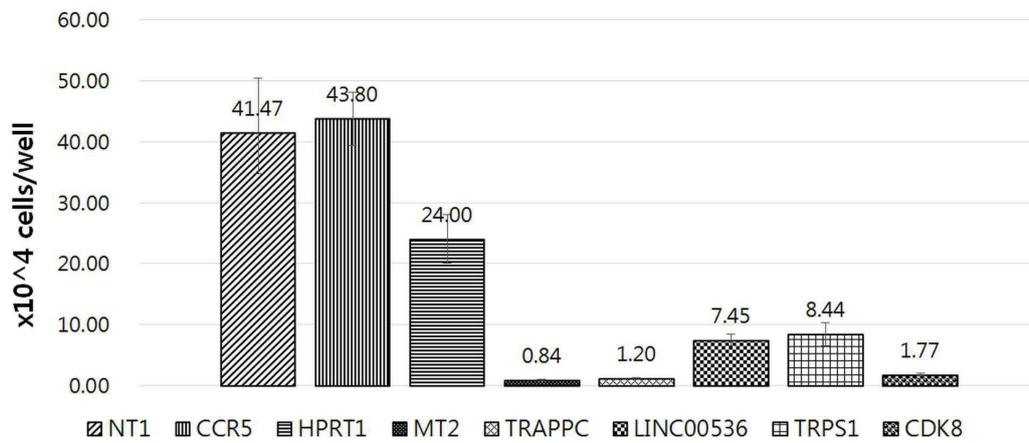
도면20a



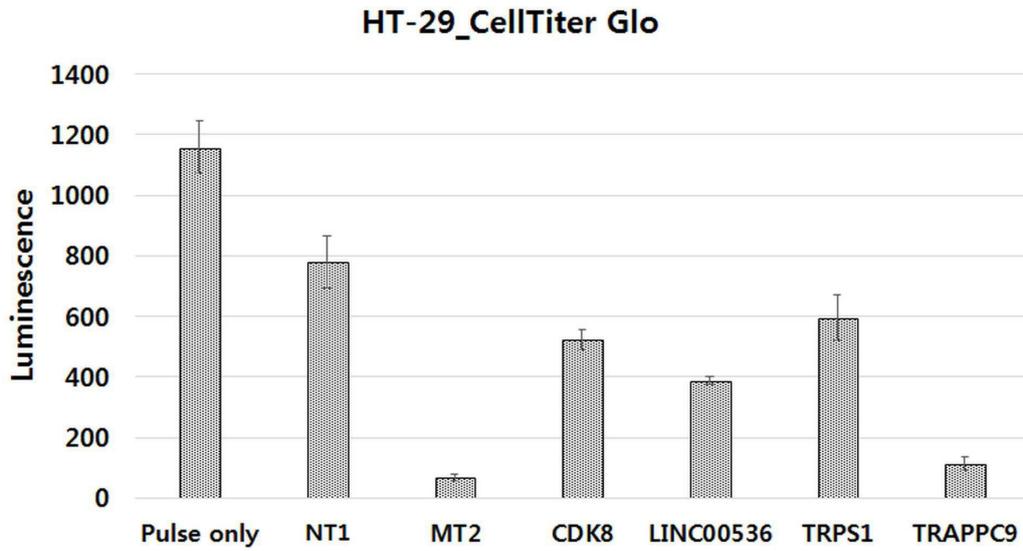
도면20b



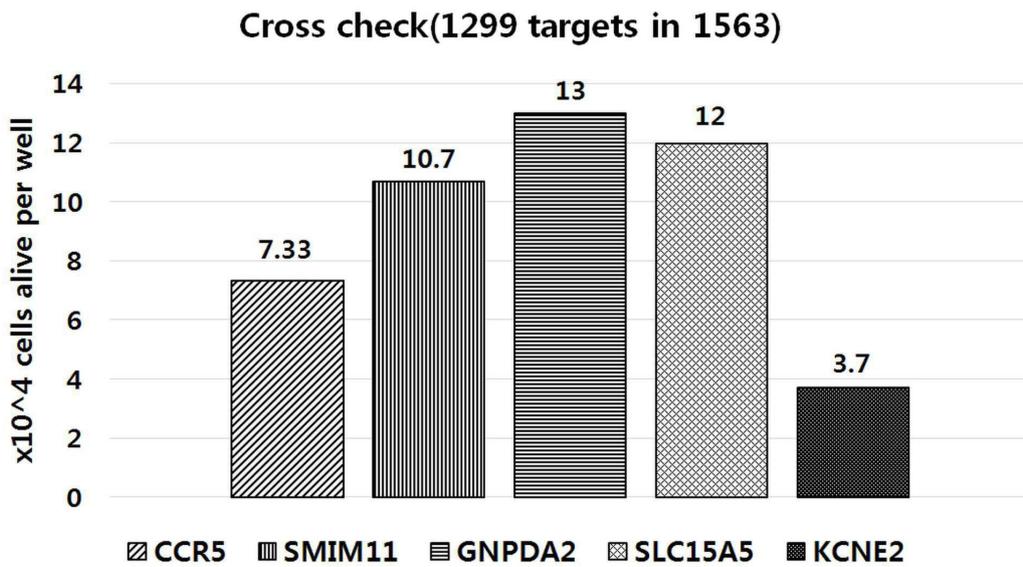
도면21



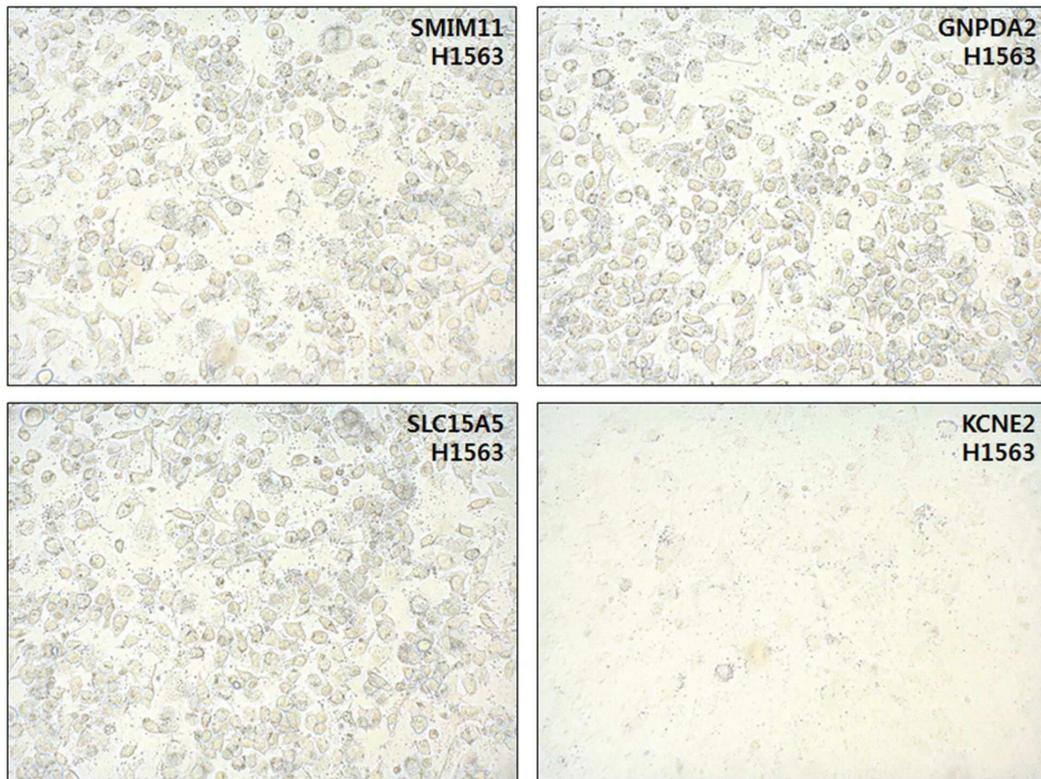
도면22



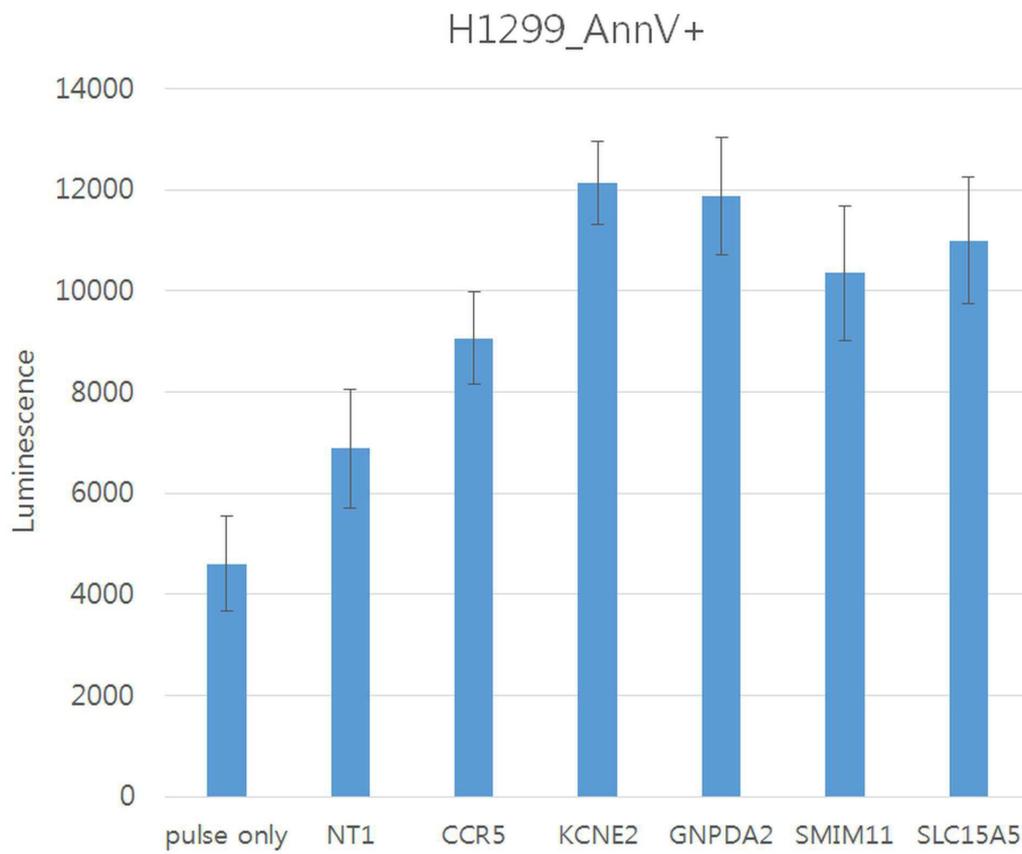
도면23



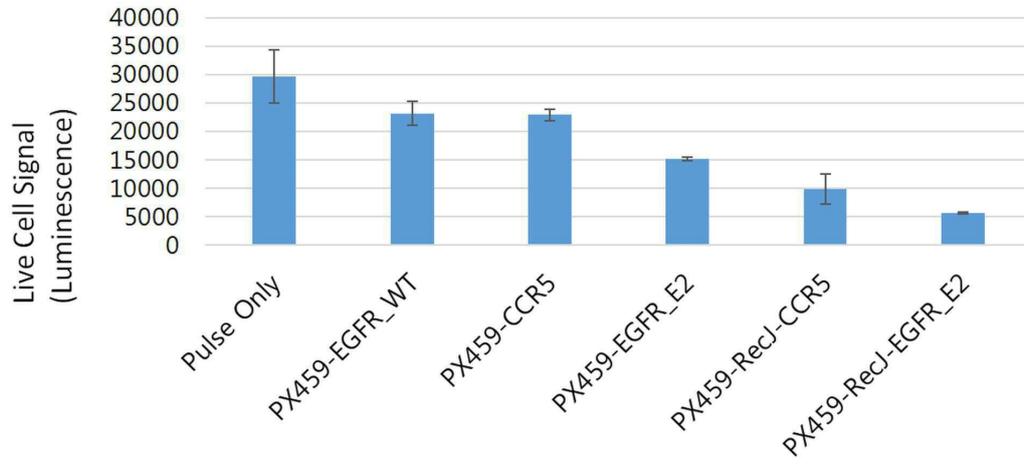
도면24



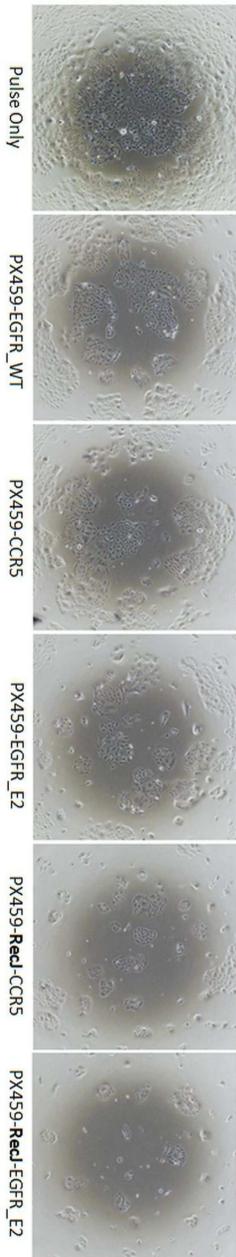
도면25



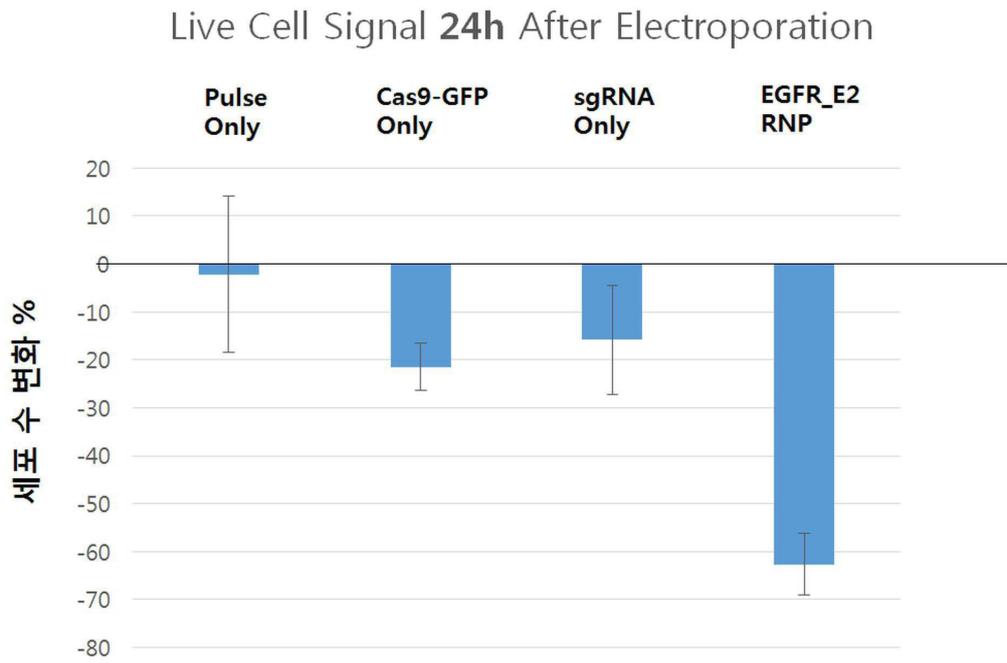
도면26



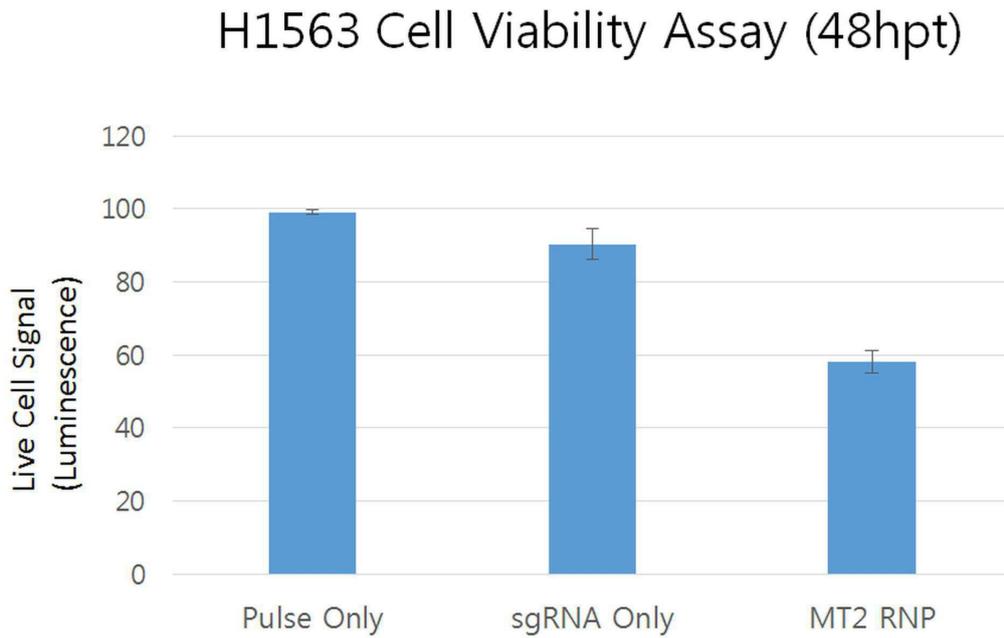
도면27



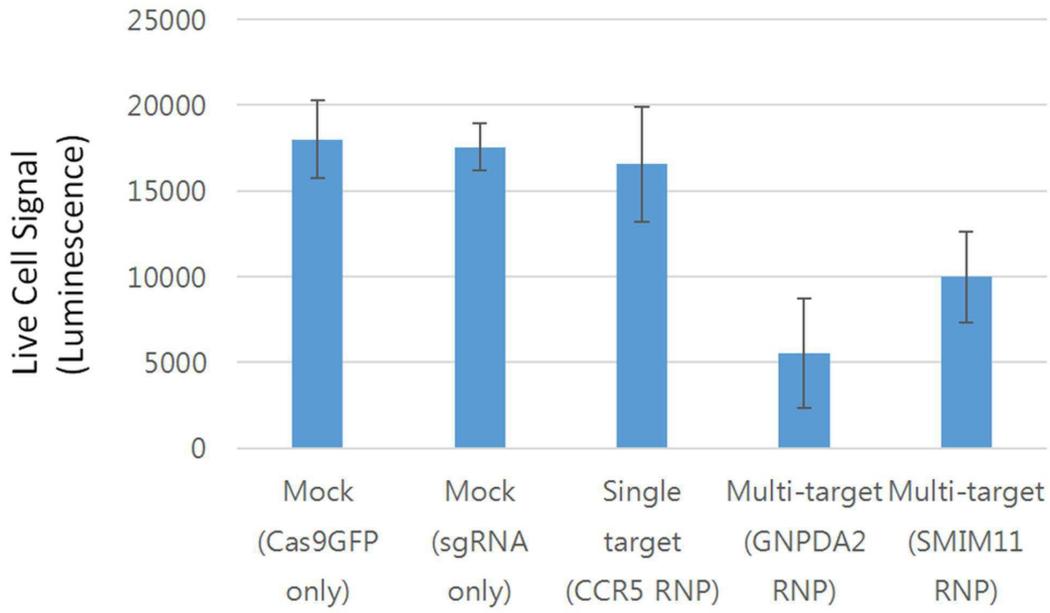
도면28



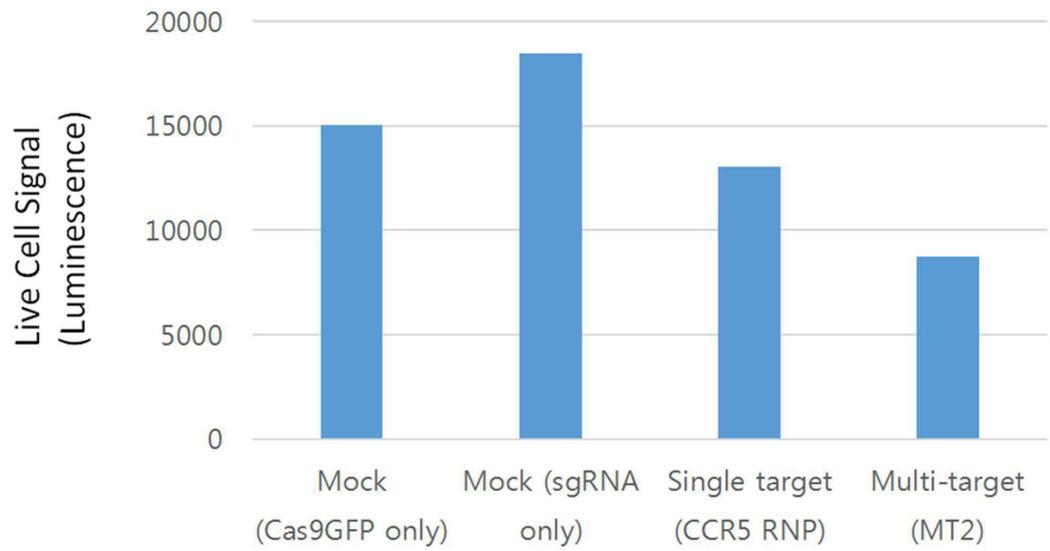
도면29



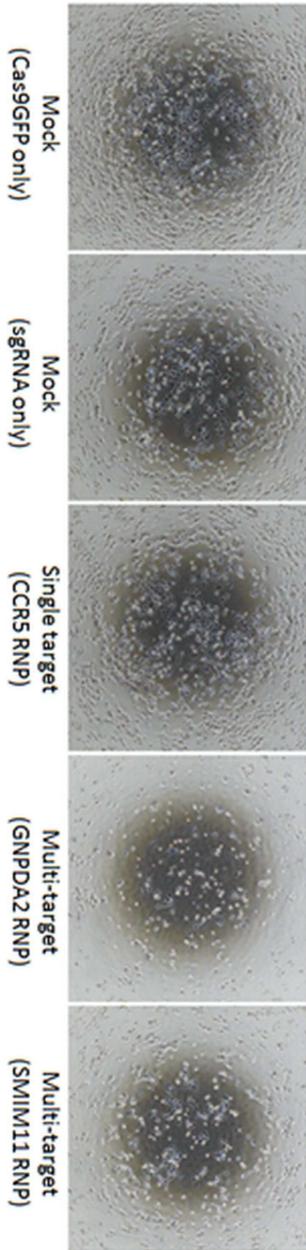
도면30



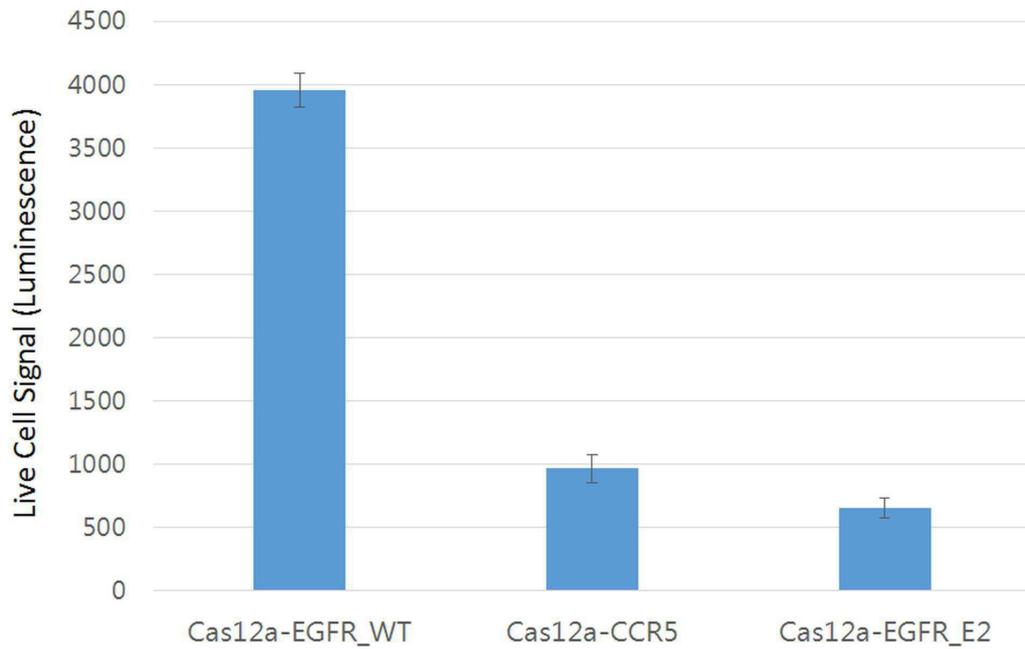
도면31



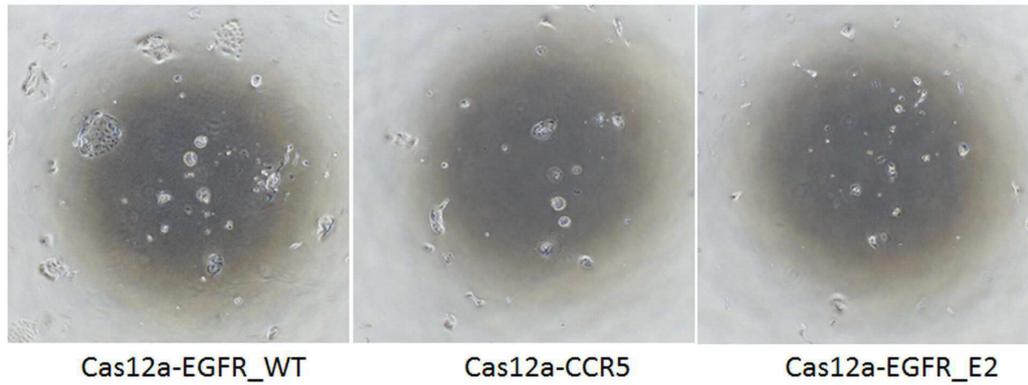
도면32



도면33



도면34



서열 목록

- <110> g-flas life sciences
- <120> PHARMACEUTICAL COMPOSITION FOR TREATING CANCERS COMPRISING GUIDE RNA AND ENDONUCLEASE
- <130> FPD/201903-0043
- <150> KR 10-2018-0035298
- <151> 2018-03-27
- <160> 139
- <170> Kopatent In 2.0
- <210> 1
- <211> 19

<212> DNA  
 <213> Artificial Sequence  
 <220><223> T7 promoter  
 <400> 1  
 taatacgact cactatagg 19  
 <210> 2  
 <211> 76  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> crRNA scaffold  
 <400> 2  
 gttttagagc tagaaatagc aagttaaaat aaggctagtc cgttatcaac ttgaaaaagt 60  
 ggcaccgagt cggtagc 76  
 <210> 3  
 <211> 111  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> T7 promoter-vector sequence-crRNA scaffold  
 <400> 3  
 taatacgact cactataggt gagaccgagg tctcggtttt agagctagaa atagcaagtt 60  
 aaaataaggc tagtccgtta tcaacttgaa aaagtggcac cgagtcggtg c 111  
 <210> 4  
 <211> 215  
 <212> PRT  
 <213> Artificial Sequence  
 <220><223> Exoribonuclease T  
 <400> 4  
 Met Ser Asp Asn Ala Gln Leu Thr Gly Leu Cys Asp Arg Phe Arg Gly  
 1 5 10 15  
 Phe Tyr Pro Val Val Ile Asp Val Glu Thr Ala Gly Phe Asn Ala Lys  
 20 25 30  
 Thr Asp Ala Leu Leu Glu Ile Ala Ala Ile Thr Leu Lys Met Asp Glu  
 35 40 45

Gln Gly Trp Leu Met Pro Asp Thr Thr Leu His Phe His Val Glu Pro

50 55 60

Phe Val Gly Ala Asn Leu Gln Pro Glu Ala Leu Ala Phe Asn Gly Ile

65 70 75 80

Asp Pro Asn Asp Pro Asp Arg Gly Ala Val Ser Glu Tyr Glu Ala Leu

85 90 95

His Glu Ile Phe Lys Val Val Arg Lys Gly Ile Lys Ala Ser Gly Cys

100 105 110

Asn Arg Ala Ile Met Val Ala His Asn Ala Asn Phe Asp His Ser Phe

115 120 125

Met Met Ala Ala Ala Glu Arg Ala Ser Leu Lys Arg Asn Pro Phe His

130 135 140

Pro Phe Ala Thr Phe Asp Thr Ala Ala Leu Ala Gly Leu Ala Leu Gly

145 150 155 160

Gln Thr Val Leu Ser Lys Ala Cys Gln Thr Ala Gly Met Asp Phe Asp

165 170 175

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

180 185 190

Leu Phe Cys Glu Ile Val Asn Arg Trp Lys Arg Leu Gly Gly Trp Pro

195 200 205

Leu Ser Ala Ala Glu Glu Val

210 215

<210> 5

<211> 236

<212> PRT

<213> Artificial Sequence

<220><223> TREX2

<400> 5

Met Ser Glu Ala Pro Arg Ala Glu Thr Phe Val Phe Leu Asp Leu Glu

1 5 10 15

Ala Thr Gly Leu Pro Ser Val Glu Pro Glu Ile Ala Glu Leu Ser Leu

20 25 30

Phe Ala Val His Arg Ser Ser Leu Glu Asn Pro Glu His Asp Glu Ser  
                   35                                  40                                  45  
 Gly Ala Leu Val Leu Pro Arg Val Leu Asp Lys Leu Thr Leu Cys Met  
           50                                  55                                  60  
 Cys Pro Glu Arg Pro Phe Thr Ala Lys Ala Ser Glu Ile Thr Gly Leu  
   65                                  70                                  75                                  80  
 Ser Ser Glu Gly Leu Ala Arg Cys Arg Lys Ala Gly Phe Asp Gly Ala  
                                   85                                  90                                  95  
 Val Val Arg Thr Leu Gln Ala Phe Leu Ser Arg Gln Ala Gly Pro Ile  
                   100                                  105                                  110  
  
 Cys Leu Val Ala His Asn Gly Phe Asp Tyr Asp Phe Pro Leu Leu Cys  
           115                                  120                                  125  
 Ala Glu Leu Arg Arg Leu Gly Ala Arg Leu Pro Arg Asp Thr Val Cys  
           130                                  135                                  140  
 Leu Asp Thr Leu Pro Ala Leu Arg Gly Leu Asp Arg Ala His Ser His  
   145                                  150                                  155                                  160  
 Gly Thr Arg Ala Arg Gly Arg Gln Gly Tyr Ser Leu Gly Ser Leu Phe  
                                   165                                  170                                  175  
 His Arg Tyr Phe Arg Ala Glu Pro Ser Ala Ala His Ser Ala Glu Gly  
  
                   180                                  185                                  190  
 Asp Val His Thr Leu Leu Leu Ile Phe Leu His Arg Ala Ala Glu Leu  
           195                                  200                                  205  
 Leu Ala Trp Ala Asp Glu Gln Ala Arg Gly Trp Ala His Ile Glu Pro  
           210                                  215                                  220  
 Met Tyr Leu Pro Pro Asp Asp Pro Ser Leu Glu Ala  
   225                                  230                                  235  
 <210> 6  
 <211> 304  
 <212> PRT  
 <213> Artificial Sequence  
 <220><223> TREX1  
 <400> 6

Met Gln Thr Leu Ile Phe Phe Asp Met Glu Ala Thr Gly Leu Pro Phe  
 1 5 10 15  
 Ser Gln Pro Lys Val Thr Glu Leu Cys Leu Leu Ala Val His Arg Cys  
 20 25 30  
 Ala Leu Glu Ser Pro Pro Thr Ser Gln Gly Pro Pro Pro Thr Val Pro  
 35 40 45  
 Pro Pro Pro Arg Val Val Asp Lys Leu Ser Leu Cys Val Ala Pro Gly  
 50 55 60  
 Lys Ala Cys Ser Pro Ala Ala Ser Glu Ile Thr Gly Leu Ser Thr Ala  
 65 70 75 80  
 Val Leu Ala Ala His Gly Arg Gln Cys Phe Asp Asp Asn Leu Ala Asn  
 85 90 95  
 Leu Leu Leu Ala Phe Leu Arg Arg Gln Pro Gln Pro Trp Cys Leu Val  
 100 105 110  
 Ala His Asn Gly Asp Arg Tyr Asp Phe Pro Leu Leu Gln Ala Glu Leu  
 115 120 125  
 Ala Met Leu Gly Leu Thr Ser Ala Leu Asp Gly Ala Phe Cys Val Asp  
 130 135 140  
 Ser Ile Thr Ala Leu Lys Ala Leu Glu Arg Ala Ser Ser Pro Ser Glu  
 145 150 155 160  
 His Gly Pro Arg Lys Ser Tyr Ser Leu Gly Ser Ile Tyr Thr Arg Leu  
 165 170 175  
 Tyr Gly Gln Ser Pro Pro Asp Ser His Thr Ala Glu Gly Asp Val Leu  
 180 185 190  
 Ala Leu Leu Ser Ile Cys Gln Trp Arg Pro Gln Ala Leu Leu Arg Trp  
 195 200 205  
 Val Asp Ala His Ala Arg Pro Phe Gly Thr Ile Arg Pro Met Tyr Gly  
 210 215 220  
 Val Thr Ala Ser Ala Arg Thr Lys Pro Arg Pro Ser Ala Val Thr Thr  
 225 230 235 240  
 Thr Ala His Leu Ala Thr Thr Arg Asn Thr Ser Pro Ser Leu Arg Glu

245 250 255  
 Ser Arg Gly Thr Lys Asp Leu Pro Pro Val Lys Asp Pro Gly Ala Leu  
 260 265 270  
 Ser Arg Glu Gly Leu Leu Ala Pro Leu Gly Leu Leu Ala Ile Leu Thr  
 275 280 285  
 Leu Ala Val Ala Thr Leu Tyr Gly Leu Ser Leu Ala Thr Pro Gly Asp  
  
 290 295 300  
 <210> 7  
 <211> 1180  
 <212> PRT  
 <213> Artificial Sequence  
 <220><223> RecBCD\_RecB  
 <400> 7  
 Met Ser Asp Val Ala Glu Thr Leu Asp Pro Leu Arg Leu Pro Leu Gln  
 1 5 10 15  
 Gly Glu Arg Leu Ile Glu Ala Ser Ala Gly Thr Gly Lys Thr Phe Thr  
 20 25 30  
 Ile Ala Ala Leu Tyr Leu Arg Leu Leu Leu Gly Leu Gly Gly Ser Ala  
 35 40 45  
  
 Ala Phe Pro Arg Pro Leu Thr Val Glu Glu Leu Leu Val Val Thr Phe  
 50 55 60  
 Thr Glu Ala Ala Thr Ala Glu Leu Arg Gly Arg Ile Arg Ser Asn Ile  
 65 70 75 80  
 His Glu Leu Arg Ile Ala Cys Leu Arg Glu Thr Thr Asp Asn Pro Leu  
 85 90 95  
 Tyr Glu Arg Leu Leu Glu Glu Ile Asp Asp Lys Ala Gln Ala Ala Gln  
 100 105 110  
 Trp Leu Leu Leu Ala Glu Arg Gln Met Asp Glu Ala Ala Val Phe Thr  
  
 115 120 125  
 Ile His Gly Phe Cys Gln Arg Met Leu Asn Leu Asn Ala Phe Glu Ser  
 130 135 140  
 Gly Met Leu Phe Glu Gln Gln Leu Ile Glu Asp Glu Ser Leu Leu Arg

145                      150                      155                      160  
 Tyr Gln Ala Cys Ala Asp Phe Trp Arg Arg His Cys Tyr Pro Leu Pro  
    165                      170                      175  
 Arg Glu Ile Ala Gln Val Val Phe Glu Thr Trp Lys Gly Pro Gln Ala  
    180                      185                      190  
  
 Leu Leu Arg Asp Ile Asn Arg Tyr Leu Gln Gly Glu Ala Pro Val Ile  
    195                      200                      205  
 Lys Ala Pro Pro Pro Asp Asp Glu Thr Leu Ala Ser Arg His Ala Gln  
    210                      215                      220  
 Ile Val Ala Arg Ile Asp Thr Val Lys Gln Gln Trp Arg Asp Ala Val  
 225                      230                      235                      240  
 Gly Glu Leu Asp Ala Leu Ile Glu Ser Ser Gly Ile Asp Arg Arg Lys  
    245                      250                      255  
 Phe Asn Arg Ser Asn Gln Ala Lys Trp Ile Asp Lys Ile Ser Ala Trp  
  
    260                      265                      270  
 Ala Glu Glu Glu Thr Asn Ser Tyr Gln Leu Pro Glu Ser Leu Glu Lys  
    275                      280                      285  
 Phe Ser Gln Arg Phe Leu Glu Asp Arg Thr Lys Ala Gly Gly Glu Thr  
    290                      295                      300  
 Pro Arg His Pro Leu Phe Glu Ala Ile Asp Gln Leu Leu Ala Glu Pro  
 305                      310                      315                      320  
 Leu Ser Ile Arg Asp Leu Val Ile Thr Arg Ala Leu Ala Glu Ile Arg  
    325                      330                      335  
  
 Glu Thr Val Ala Arg Glu Lys Arg Arg Arg Gly Glu Leu Gly Phe Asp  
    340                      345                      350  
 Asp Met Leu Ser Arg Leu Asp Ser Ala Leu Arg Ser Glu Ser Gly Glu  
    355                      360                      365  
 Val Leu Ala Ala Ala Ile Arg Thr Arg Phe Pro Val Ala Met Ile Asp  
    370                      375                      380  
 Glu Phe Gln Asp Thr Asp Pro Gln Gln Tyr Arg Ile Phe Arg Arg Ile  
 385                      390                      395                      400

Trp His His Gln Pro Glu Thr Ala Leu Leu Leu Ile Gly Asp Pro Lys  
 405 410 415  
 Gln Ala Ile Tyr Ala Phe Arg Gly Ala Asp Ile Phe Thr Tyr Met Lys  
 420 425 430  
 Ala Arg Ser Glu Val His Ala His Tyr Thr Leu Asp Thr Asn Trp Arg  
 435 440 445  
 Ser Ala Pro Gly Met Val Asn Ser Val Asn Lys Leu Phe Ser Gln Thr  
 450 455 460  
 Asp Asp Ala Phe Met Phe Arg Glu Ile Pro Phe Ile Pro Val Lys Ser  
 465 470 475 480  
  
 Ala Gly Lys Asn Gln Ala Leu Arg Phe Val Phe Lys Gly Glu Thr Gln  
 485 490 495  
 Pro Ala Met Lys Met Trp Leu Met Glu Gly Glu Ser Cys Gly Val Gly  
 500 505 510  
 Asp Tyr Gln Ser Thr Met Ala Gln Val Cys Ala Ala Gln Ile Arg Asp  
 515 520 525  
 Trp Leu Gln Ala Gly Gln Arg Gly Glu Ala Leu Leu Met Asn Gly Asp  
 530 535 540  
 Asp Ala Arg Pro Val Arg Ala Ser Asp Ile Ser Val Leu Val Arg Ser  
  
 545 550 555 560  
 Arg Gln Glu Ala Ala Gln Val Arg Asp Ala Leu Thr Leu Leu Glu Ile  
 565 570 575  
 Pro Ser Val Tyr Leu Ser Asn Arg Asp Ser Val Phe Glu Thr Leu Glu  
 580 585 590  
 Ala Gln Glu Met Leu Trp Leu Leu Gln Ala Val Met Thr Pro Glu Arg  
 595 600 605  
 Glu Asn Thr Leu Arg Ser Ala Leu Ala Thr Ser Met Met Gly Leu Asn  
 610 615 620  
  
 Ala Leu Asp Ile Glu Thr Leu Asn Asn Asp Glu His Ala Trp Asp Val  
 625 630 635 640  
 Val Val Glu Glu Phe Asp Gly Tyr Arg Gln Ile Trp Arg Lys Arg Gly

645                      650                      655  
 Val Met Pro Met Leu Arg Ala Leu Met Ser Ala Arg Asn Ile Ala Glu  
 660                      665                      670  
 Asn Leu Leu Ala Thr Ala Gly Gly Glu Arg Arg Leu Thr Asp Ile Leu  
 675                      680                      685  
 His Ile Ser Glu Leu Leu Gln Glu Ala Gly Thr Gln Leu Glu Ser Glu  
  
 690                      695                      700  
 His Ala Leu Val Arg Trp Leu Ser Gln His Ile Leu Glu Pro Asp Ser  
 705                      710                      715                      720  
 Asn Ala Ser Ser Gln Gln Met Arg Leu Glu Ser Asp Lys His Leu Val  
 725                      730                      735  
 Gln Ile Val Thr Ile His Lys Ser Lys Gly Leu Glu Tyr Pro Leu Val  
 740                      745                      750  
 Trp Leu Pro Phe Ile Thr Asn Phe Arg Val Gln Glu Gln Ala Phe Tyr  
 755                      760                      765  
  
 His Asp Arg His Ser Phe Glu Ala Val Leu Asp Leu Asn Ala Ala Pro  
 770                      775                      780  
 Glu Ser Val Asp Leu Ala Glu Ala Glu Arg Leu Ala Glu Asp Leu Arg  
 785                      790                      795                      800  
 Leu Leu Tyr Val Ala Leu Thr Arg Ser Val Trp His Cys Ser Leu Gly  
 805                      810                      815  
 Val Ala Pro Leu Val Arg Arg Arg Gly Asp Lys Lys Gly Asp Thr Asp  
 820                      825                      830  
 Val His Gln Ser Ala Leu Gly Arg Leu Leu Gln Lys Gly Glu Pro Gln  
  
 835                      840                      845  
 Asp Ala Ala Gly Leu Arg Thr Cys Ile Glu Ala Leu Cys Asp Asp Asp  
 850                      855                      860  
 Ile Ala Trp Gln Thr Ala Gln Thr Gly Asp Asn Gln Pro Trp Gln Val  
 865                      870                      875                      880  
 Asn Asp Val Ser Thr Ala Glu Leu Asn Ala Lys Thr Leu Gln Arg Leu  
 885                      890                      895  
 Pro Gly Asp Asn Trp Arg Val Thr Ser Tyr Ser Gly Leu Gln Gln Arg

900	905	910
Gly His Gly Ile Ala Gln Asp Leu Met Pro Arg Leu Asp Val Asp Ala		
915	920	925
Ala Gly Val Ala Ser Val Val Glu Glu Pro Thr Leu Thr Pro His Gln		
930	935	940
Phe Pro Arg Gly Ala Ser Pro Gly Thr Phe Leu His Ser Leu Phe Glu		
945	950	955
Asp Leu Asp Phe Thr Gln Pro Val Asp Pro Asn Trp Val Arg Glu Lys		
965	970	975
Leu Glu Leu Gly Gly Phe Glu Ser Gln Trp Glu Pro Val Leu Thr Glu		
980	985	990
Trp Ile Thr Ala Val Leu Gln Ala Pro Leu Asn Glu Thr Gly Val Ser		
995	1000	1005
Leu Ser Gln Leu Ser Ala Arg Asn Lys Gln Val Glu Met Glu Phe Tyr		
1010	1015	1020
Leu Pro Ile Ser Glu Pro Leu Ile Ala Ser Gln Leu Asp Thr Leu Ile		
1025	1030	1035
Arg Gln Phe Asp Pro Leu Ser Ala Gly Cys Pro Pro Leu Glu Phe Met		
1045	1050	1055
Gln Val Arg Gly Met Leu Lys Gly Phe Ile Asp Leu Val Phe Arg His		
1060	1065	1070
Glu Gly Arg Tyr Tyr Leu Leu Asp Tyr Lys Ser Asn Trp Leu Gly Glu		
1075	1080	1085
Asp Ser Ser Ala Tyr Thr Gln Gln Ala Met Ala Ala Ala Met Gln Ala		
1090	1095	1100
His Arg Tyr Asp Leu Gln Tyr Gln Leu Tyr Thr Leu Ala Leu His Arg		
1105	1110	1115
Tyr Leu Arg His Arg Ile Ala Asp Tyr Asp Tyr Glu His His Phe Gly		
1125	1130	1135
Gly Val Ile Tyr Leu Phe Leu Arg Gly Val Asp Lys Glu His Pro Gln		
1140	1145	1150

Gln Gly Ile Tyr Thr Thr Arg Pro Asn Ala Gly Leu Ile Ala Leu Met  
 1155 1160 1165  
 Asp Glu Met Phe Ala Gly Met Thr Leu Glu Glu Ala  
 1170 1175 1180  
 <210> 8  
 <211> 1122  
 <212> PRT  
 <213> Artificial Sequence  
 <220><223> RecBCD\_RecC  
 <400>  
 > 8  
 Met Leu Arg Val Tyr His Ser Asn Arg Leu Asp Val Leu Glu Ala Leu  
 1 5 10 15  
 Met Glu Phe Ile Val Glu Arg Glu Arg Leu Asp Asp Pro Phe Glu Pro  
 20 25 30  
 Glu Met Ile Leu Val Gln Ser Thr Gly Met Ala Gln Trp Leu Gln Met  
 35 40 45  
 Thr Leu Ser Gln Lys Phe Gly Ile Ala Ala Asn Ile Asp Phe Pro Leu  
 50 55 60  
 Pro Ala Ser Phe Ile Trp Asp Met Phe Val Arg Val Leu Pro Glu Ile  
 65 70 75 80  
 Pro Lys Glu Ser Ala Phe Asn Lys Gln Ser Met Ser Trp Lys Leu Met  
 85 90 95  
 Thr Leu Leu Pro Gln Leu Leu Glu Arg Glu Asp Phe Thr Leu Leu Arg  
 100 105 110  
 His Tyr Leu Thr Asp Asp Ser Asp Lys Arg Lys Leu Phe Gln Leu Ser  
 115 120 125  
 Ser Lys Ala Ala Asp Leu Phe Asp Gln Tyr Leu Val Tyr Arg Pro Asp  
 130 135 140  
 Trp Leu Ala Gln Trp Glu Thr Gly His Leu Val Glu Gly Leu Gly Glu  
 145 150 155 160  
 Ala Gln Ala Trp Gln Ala Pro Leu Trp Lys Ala Leu Val Glu Tyr Thr  
 165 170 175

His Gln Leu Gly Gln Pro Arg Trp His Arg Ala Asn Leu Tyr Gln Arg  
 180 185 190  
 Phe Ile Glu Thr Leu Glu Ser Ala Thr Thr Cys Pro Pro Gly Leu Pro  
 195 200 205  
 Ser Arg Val Phe Ile Cys Gly Ile Ser Ala Leu Pro Pro Val Tyr Leu  
 210 215 220  
 Gln Ala Leu Gln Ala Leu Gly Lys His Ile Glu Ile His Leu Leu Phe  
 225 230 235 240  
 Thr Asn Pro Cys Arg Tyr Tyr Trp Gly Asp Ile Lys Asp Pro Ala Tyr  
 245 250 255  
 Leu Ala Lys Leu Leu Thr Arg Gln Arg Arg His Ser Phe Glu Asp Arg  
 260 265 270  
 Glu Leu Pro Leu Phe Arg Asp Ser Glu Asn Ala Gly Gln Leu Phe Asn  
 275 280 285  
 Ser Asp Gly Glu Gln Asp Val Gly Asn Pro Leu Leu Ala Ser Trp Gly  
 290 295 300  
 Lys Leu Gly Arg Asp Tyr Ile Tyr Leu Leu Ser Asp Leu Glu Ser Ser  
 305 310 315 320  
 Gln Glu Leu Asp Ala Phe Val Asp Val Thr Pro Asp Asn Leu Leu His  
 325 330 335  
 Asn Ile Gln Ser Asp Ile Leu Glu Leu Glu Asn Arg Ala Val Ala Gly  
 340 345 350  
 Val Asn Ile Glu Glu Phe Ser Arg Ser Asp Asn Lys Arg Pro Leu Asp  
 355 360 365  
 Pro Leu Asp Ser Ser Ile Thr Phe His Val Cys His Ser Pro Gln Arg  
 370 375 380  
 Glu Val Glu Val Leu His Asp Arg Leu Leu Ala Met Leu Glu Glu Asp  
 385 390 395 400  
 Pro Thr Leu Thr Pro Arg Asp Ile Ile Val Met Val Ala Asp Ile Asp  
 405 410 415  
 Ser Tyr Ser Pro Phe Ile Gln Ala Val Phe Gly Ser Ala Pro Ala Asp



Lys Val Val Cys Leu Leu Gly Met Asn Asp Gly Val Tyr Pro Arg Gln  
 675 680 685  
 Leu Ala Pro Leu Gly Phe Asp Leu Met Ser Gln Lys Pro Lys Arg Gly  
 690 695 700  
 Asp Arg Ser Arg Arg Asp Asp Asp Arg Tyr Leu Phe Leu Glu Ala Leu  
 705 710 715 720  
  
 Ile Ser Ala Gln Gln Lys Leu Tyr Ile Ser Tyr Ile Gly Arg Ser Ile  
 725 730 735  
 Gln Asp Asn Ser Glu Arg Phe Pro Ser Val Leu Val Gln Glu Leu Ile  
 740 745 750  
 Asp Tyr Ile Gly Gln Ser His Tyr Leu Pro Gly Asp Glu Ala Leu Asn  
 755 760 765  
 Cys Asp Glu Ser Glu Ala Arg Val Lys Ala His Leu Thr Cys Leu His  
 770 775 780  
 Thr Arg Met Pro Phe Asp Pro Gln Asn Tyr Gln Pro Gly Glu Arg Gln  
  
 785 790 795 800  
 Ser Tyr Ala Arg Glu Trp Leu Pro Ala Ala Ser Gln Ala Gly Lys Ala  
 805 810 815  
 His Ser Glu Phe Val Gln Pro Leu Pro Phe Thr Leu Pro Glu Thr Val  
 820 825 830  
 Pro Leu Glu Thr Leu Gln Arg Phe Trp Ala His Pro Val Arg Ala Phe  
 835 840 845  
 Phe Gln Met Arg Leu Gln Val Asn Phe Arg Thr Glu Asp Ser Glu Ile  
 850 855 860  
  
 Pro Asp Thr Glu Pro Phe Ile Leu Glu Gly Leu Ser Arg Tyr Gln Ile  
 865 870 875 880  
 Asn Gln Gln Leu Leu Asn Ala Leu Val Glu Gln Asp Asp Ala Glu Arg  
 885 890 895  
 Leu Phe Arg Arg Phe Arg Ala Ala Gly Asp Leu Pro Tyr Gly Ala Phe  
 900 905 910  
 Gly Glu Ile Phe Trp Glu Thr Gln Cys Gln Glu Met Gln Gln Leu Ala  
 915 920 925

Asp Arg Val Ile Ala Cys Arg Gln Pro Gly Gln Ser Met Glu Ile Asp

930 935 940

Leu Ala Cys Asn Gly Val Gln Ile Thr Gly Trp Leu Pro Gln Val Gln

945 950 955 960

Pro Asp Gly Leu Leu Arg Trp Arg Pro Ser Leu Leu Ser Val Ala Gln

965 970 975

Gly Met Gln Leu Trp Leu Glu His Leu Val Tyr Cys Ala Ser Gly Gly

980 985 990

Asn Gly Glu Ser Arg Leu Phe Leu Arg Lys Asp Gly Glu Trp Arg Phe

995 1000 1005

Pro Pro Leu Ala Ala Glu Gln Ala Leu His Tyr Leu Ser Gln Leu Ile

1010 1015 1020

Glu Gly Tyr Arg Glu Gly Met Ser Ala Pro Leu Leu Val Leu Pro Glu

1025 1030 1035 1040

Ser Gly Gly Ala Trp Leu Lys Thr Cys Tyr Asp Ala Gln Asn Asp Ala

1045 1050 1055

Met Leu Asp Asp Asp Ser Thr Leu Gln Lys Ala Arg Thr Lys Phe Leu

1060 1065 1070

Gln Ala Tyr Glu Gly Asn Met Met Val Arg Gly Glu Gly Asp Asp Ile

1075 1080 1085

Trp Tyr Gln Arg Leu Trp Arg Gln Leu Thr Pro Glu Thr Met Glu Ala

1090 1095 1100

Ile Val Glu Gln Ser Gln Arg Phe Leu Leu Pro Leu Phe Arg Phe Asn

1105 1110 1115 1120

Gln Ser

<210> 9

<211> 608

<212> PRT

<213> Artificial Sequence

<220><223> RecBCD\_RecD

<400> 9

Met Lys Leu Gln Lys Gln Leu Leu Glu Ala Val Glu His Lys Gln Leu  
 1                      5                      10                      15

Arg Pro Leu Asp Val Gln Phe Ala Leu Thr Val Ala Gly Asp Glu His  
                     20                      25                      30

Pro Ala Val Thr Leu Ala Ala Ala Leu Leu Ser His Asp Ala Gly Glu  
                     35                      40                      45

Gly His Val Cys Leu Pro Leu Ser Arg Leu Glu Asn Asn Glu Ala Ser  
                     50                      55                      60

His Pro Leu Leu Ala Thr Cys Val Ser Glu Ile Gly Glu Leu Gln Asn  
                     65                      70                      75                      80

Trp Glu Glu Cys Leu Leu Ala Ser Gln Ala Val Ser Arg Gly Asp Glu  
                     85                      90                      95

Pro Thr Pro Met Ile Leu Cys Gly Asp Arg Leu Tyr Leu Asn Arg Met  
                     100                      105                      110

Trp Cys Asn Glu Arg Thr Val Ala Arg Phe Phe Asn Glu Val Asn His  
                     115                      120                      125

Ala Ile Glu Val Asp Glu Ala Leu Leu Ala Gln Thr Leu Asp Lys Leu  
                     130                      135                      140

Phe Pro Val Ser Asp Glu Ile Asn Trp Gln Lys Val Ala Ala Ala Val  
                     145                      150                      155                      160

Ala Leu Thr Arg Arg Ile Ser Val Ile Ser Gly Gly Pro Gly Thr Gly  
                     165                      170                      175

Lys Thr Thr Thr Val Ala Lys Leu Leu Ala Ala Leu Ile Gln Met Ala  
                     180                      185                      190

Asp Gly Glu Arg Cys Arg Ile Arg Leu Ala Ala Pro Thr Gly Lys Ala  
                     195                      200                      205

Ala Ala Arg Leu Thr Glu Ser Leu Gly Lys Ala Leu Arg Gln Leu Pro  
                     210                      215                      220

Leu Thr Asp Glu Gln Lys Lys Arg Ile Pro Glu Asp Ala Ser Thr Leu  
                     225                      230                      235                      240

His Arg Leu Leu Gly Ala Gln Pro Gly Ser Gln Arg Leu Arg His His

245	250	255
Ala Gly Asn Pro Leu His Leu Asp Val Leu Val Val Asp Glu Ala Ser		
260	265	270
Met Ile Asp Leu Pro Met Met Ser Arg Leu Ile Asp Ala Leu Pro Asp		
275	280	285
His Ala Arg Val Ile Phe Leu Gly Asp Arg Asp Gln Leu Ala Ser Val		
290	295	300
Glu Ala Gly Ala Val Leu Gly Asp Ile Cys Ala Tyr Ala Asn Ala Gly		
305	310	315
Phe Thr Ala Glu Arg Ala Arg Gln Leu Ser Arg Leu Thr Gly Thr His		
325	330	335
Val Pro Ala Gly Thr Gly Thr Glu Ala Ala Ser Leu Arg Asp Ser Leu		
340	345	350
Cys Leu Leu Gln Lys Ser Tyr Arg Phe Gly Ser Asp Ser Gly Ile Gly		
355	360	365
Gln Leu Ala Ala Ala Ile Asn Arg Gly Asp Lys Thr Ala Val Lys Thr		
370	375	380
Val Phe Gln Gln Asp Phe Thr Asp Ile Glu Lys Arg Leu Leu Gln Ser		
385	390	395
Gly Glu Asp Tyr Ile Ala Met Leu Glu Glu Ala Leu Ala Gly Tyr Gly		
405	410	415
Arg Tyr Leu Asp Leu Leu Gln Ala Arg Ala Glu Pro Asp Leu Ile Ile		
420	425	430
Gln Ala Phe Asn Glu Tyr Gln Leu Leu Cys Ala Leu Arg Glu Gly Pro		
435	440	445
Phe Gly Val Ala Gly Leu Asn Glu Arg Ile Glu Gln Phe Met Gln Gln		
450	455	460
Lys Arg Lys Ile His Arg His Pro His Ser Arg Trp Tyr Glu Gly Arg		
465	470	475
Pro Val Met Ile Ala Arg Asn Asp Ser Ala Leu Gly Leu Phe Asn Gly		
485	490	495
Asp Ile Gly Ile Ala Leu Asp Arg Gly Gln Gly Thr Arg Val Trp Phe		

500 505 510  
 Ala Met Pro Asp Gly Asn Ile Lys Ser Val Gln Pro Ser Arg Leu Pro

515 520 525  
 Glu His Glu Thr Thr Trp Ala Met Thr Val His Lys Ser Gln Gly Ser

530 535 540

Glu Phe Asp His Ala Ala Leu Ile Leu Pro Ser Gln Arg Thr Pro Val  
 545 550 555 560

Val Thr Arg Glu Leu Val Tyr Thr Ala Val Thr Arg Ala Arg Arg Arg  
 565 570 575

Leu Ser Leu Tyr Ala Asp Glu Arg Ile Leu Ser Ala Ala Ile Ala Thr  
 580 585 590

Arg Thr Glu Arg Arg Ser Gly Leu Ala Ala Leu Phe Ser Ser Arg Glu  
 595 600 605

<210> 10

<211> 475

<212> PRT

<213> Artificial Sequence

<220><223> Exodeoxyribonuclease I

<400> 10

Met Met Asn Asp Gly Lys Gln Gln Ser Thr Phe Leu Phe His Asp Tyr  
 1 5 10 15

Glu Thr Phe Gly Thr His Pro Ala Leu Asp Arg Pro Ala Gln Phe Ala  
 20 25 30

Ala Ile Arg Thr Asp Ser Glu Phe Asn Val Ile Gly Glu Pro Glu Val

35 40 45  
 Phe Tyr Cys Lys Pro Ala Asp Asp Tyr Leu Pro Gln Pro Gly Ala Val

50 55 60  
 Leu Ile Thr Gly Ile Thr Pro Gln Glu Ala Arg Ala Lys Gly Glu Asn

65 70 75 80  
 Glu Ala Ala Phe Ala Ala Arg Ile His Ser Leu Phe Thr Val Pro Lys

85 90 95  
 Thr Cys Ile Leu Gly Tyr Asn Asn Val Arg Phe Asp Asp Glu Val Thr



Glu Ala Glu Pro Phe Thr Pro Ser Asp Asn Val Asp Ala Gln Leu Tyr  
 355 360 365  
 Asn Gly Phe Phe Ser Asp Ala Asp Arg Ala Ala Met Lys Ile Val Leu  
 370 375 380  
 Glu Thr Glu Pro Arg Asn Leu Pro Ala Leu Asp Ile Thr Phe Val Asp  
 385 390 395 400  
  
 Lys Arg Ile Glu Lys Leu Leu Phe Asn Tyr Arg Ala Arg Asn Phe Pro  
 405 410 415  
 Gly Thr Leu Asp Tyr Ala Glu Gln Gln Arg Trp Leu Glu His Arg Arg  
 420 425 430  
 Gln Val Phe Thr Pro Glu Phe Leu Gln Gly Tyr Ala Asp Glu Leu Gln  
 435 440 445  
 Met Leu Val Gln Gln Tyr Ala Asp Asp Lys Glu Lys Val Ala Leu Leu  
 450 455 460  
 Lys Ala Leu Trp Gln Tyr Ala Glu Glu Ile Val  
  
 465 470 475  
 <210> 11  
 <211> 268  
 <212> PRT  
 <213> Artificial Sequence  
 <220><223> Exodeoxyribonuclease III  
 <400> 11  
 Met Lys Phe Val Ser Phe Asn Ile Asn Gly Leu Arg Ala Arg Pro His  
 1 5 10 15  
 Gln Leu Glu Ala Ile Val Glu Lys His Gln Pro Asp Val Ile Gly Leu  
 20 25 30  
 Gln Glu Thr Lys Val His Asp Asp Met Phe Pro Leu Glu Glu Val Ala  
 35 40 45  
  
 Lys Leu Gly Tyr Asn Val Phe Tyr His Gly Gln Lys Gly His Tyr Gly  
 50 55 60  
 Val Ala Leu Leu Thr Lys Glu Thr Pro Ile Ala Val Arg Arg Gly Phe  
 65 70 75 80

Pro Gly Asp Asp Glu Glu Ala Gln Arg Arg Ile Ile Met Ala Glu Ile  
85 90 95

Pro Ser Leu Leu Gly Asn Val Thr Val Ile Asn Gly Tyr Phe Pro Gln  
100 105 110

Gly Glu Ser Arg Asp His Pro Ile Lys Phe Pro Ala Lys Ala Gln Phe  
115 120 125

Tyr Gln Asn Leu Gln Asn Tyr Leu Glu Thr Glu Leu Lys Arg Asp Asn  
130 135 140

Pro Val Leu Ile Met Gly Asp Met Asn Ile Ser Pro Thr Asp Leu Asp  
145 150 155 160

Ile Gly Ile Gly Glu Glu Asn Arg Lys Arg Trp Leu Arg Thr Gly Lys  
165 170 175

Cys Ser Phe Leu Pro Glu Glu Arg Glu Trp Met Asp Arg Leu Met Ser  
180 185 190

Trp Gly Leu Val Asp Thr Phe Arg His Ala Asn Pro Gln Thr Ala Asp  
195 200 205

Arg Phe Ser Trp Phe Asp Tyr Arg Ser Lys Gly Phe Asp Asp Asn Arg  
210 215 220

Gly Leu Arg Ile Asp Leu Leu Leu Ala Ser Gln Pro Leu Ala Glu Cys  
225 230 235 240

Cys Val Glu Thr Gly Ile Asp Tyr Glu Ile Arg Ser Met Glu Lys Pro  
245 250 255

Ser Asp His Ala Pro Val Trp Ala Thr Phe Arg Arg  
260 265

<210> 12  
<211> 355  
<212> PRT  
<213> Artificial Sequence  
<220><223> Mungbean exonuclease  
<400> 12

Met Gln Thr Leu Gln Met Ser Leu Leu Thr Gln Pro Tyr Val Gln Pro  
1 5 10 15

Arg Phe Pro Cys Lys Arg Tyr Pro Thr Phe Ser Ala Ser Cys Arg Thr  
 20 25 30  
 Gln Lys Thr Ala Ile Thr Lys Thr Glu Lys Val Phe Phe Ser Glu Ser  
 35 40 45  
  
 Phe Asp Gln Thr Arg Cys Thr Gln Pro Leu Ser Glu Lys Lys Lys Arg  
 50 55 60  
 Val Phe Phe Leu Asp Val Asn Pro Leu Cys Tyr Glu Gly Ser Lys Pro  
 65 70 75 80  
 Ser Leu Arg Ser Phe Gly Arg Trp Leu Ser Leu Phe Leu His Gln Val  
 85 90 95  
 Ser Leu Thr Asp Pro Val Ile Ala Val Ile Asp Gly Glu Gly Gly Ser  
 100 105 110  
 Glu His Arg Arg Lys Leu Leu Pro Ser Tyr Lys Ala His Arg Lys Lys  
  
 115 120 125  
 Phe Met Arg His Met Ser Ser Gly His Val Gly Arg Ser His Gln Val  
 130 135 140  
 Ile Asn Asp Val Leu Gly Lys Cys Asn Val Pro Val Ile Lys Val Ala  
 145 150 155 160  
 Gly His Glu Ala Asp Asp Val Val Ala Thr Leu Ala Gly Gln Val Val  
 165 170 175  
 Asn Lys Gly Phe Arg Val Val Ile Gly Ser Pro Asp Lys Asp Phe Lys  
 180 185 190  
  
 Gln Leu Ile Ser Glu Asp Val Gln Ile Val Met Pro Leu Pro Glu Leu  
 195 200 205  
 Gln Arg Trp Ser Phe Tyr Thr Leu Arg His Tyr Arg Asp Gln Tyr Asn  
 210 215 220  
 Cys Asp Pro Glu Ser Asp Leu Ser Phe Arg Cys Ile Val Gly Asp Glu  
 225 230 235 240  
 Val Asp Gly Val Pro Gly Ile Gln His Leu Val Pro Ser Phe Gly Arg  
 245 250 255  
 Lys Thr Ala Met Lys Leu Ile Lys Lys His Gly Ser Leu Glu Thr Leu

260 265 270  
 Leu Asn Ala Ala Ala Ile Arg Thr Val Gly Arg Pro Tyr Ala Gln Asp  
 275 280 285  
 Ala Leu Lys Asn His Ala Asp Tyr Leu Arg Arg Asn Tyr Glu Val Leu  
 290 295 300  
 Ala Leu Lys Arg Asp Val Asn Ile Gln Leu Tyr Asp Glu Trp Leu Val  
 305 310 315 320  
 Lys Arg Asp Asn His Asn Asp Lys Thr Ala Leu Ser Ser Phe Phe Lys  
 325 330 335  
  
 Tyr Leu Gly Glu Ser Lys Glu Leu Ser Tyr Asn Gly Arg Pro Ile Ser  
 340 345 350  
 Tyr Asn Gly  
 355  
 <210> 13  
 <211> 577  
 <212> PRT  
 <213> Artificial Sequence  
 <220><223> RecJ  
 <400> 13  
 Val Lys Gln Gln Ile Gln Leu Arg Arg Arg Glu Val Asp Glu Thr Ala  
 1 5 10 15  
 Asp Leu Pro Ala Glu Leu Pro Pro Leu Leu Arg Arg Leu Tyr Ala Ser  
 20 25 30  
  
 Arg Gly Val Arg Ser Ala Gln Glu Leu Glu Arg Ser Val Lys Gly Met  
 35 40 45  
 Leu Pro Trp Gln Gln Leu Ser Gly Val Glu Lys Ala Val Glu Ile Leu  
 50 55 60  
 Tyr Asn Ala Phe Arg Glu Gly Thr Arg Ile Ile Val Val Gly Asp Phe  
 65 70 75 80  
 Asp Ala Asp Gly Ala Thr Ser Thr Ala Leu Ser Val Leu Ala Met Arg  
 85 90 95  
 Ser Leu Gly Cys Ser Asn Ile Asp Tyr Leu Val Pro Asn Arg Phe Glu

100 105 110  
 Asp Gly Tyr Gly Leu Ser Pro Glu Val Val Asp Gln Ala His Ala Arg  
 115 120 125  
 Gly Ala Gln Leu Ile Val Thr Val Asp Asn Gly Ile Ser Ser His Ala  
 130 135 140  
 Gly Val Glu His Ala Arg Ser Leu Gly Ile Pro Val Ile Val Thr Asp  
 145 150 155 160  
 His His Leu Pro Gly Asp Thr Leu Pro Ala Ala Glu Ala Ile Ile Asn  
 165 170 175  
  
 Pro Asn Leu Arg Asp Cys Asn Phe Pro Ser Lys Ser Leu Ala Gly Val  
 180 185 190  
 Gly Val Ala Phe Tyr Leu Met Leu Ala Leu Arg Thr Phe Leu Arg Asp  
 195 200 205  
 Gln Gly Trp Phe Asp Glu Arg Asn Ile Ala Ile Pro Asn Leu Ala Glu  
 210 215 220  
 Leu Leu Asp Leu Val Ala Leu Gly Thr Val Ala Asp Val Val Pro Leu  
 225 230 235 240  
 Asp Ala Asn Asn Arg Ile Leu Thr Trp Gln Gly Met Ser Arg Ile Arg  
  
 245 250 255  
 Ala Gly Lys Cys Arg Pro Gly Ile Lys Ala Leu Leu Glu Val Ala Asn  
 260 265 270  
 Arg Asp Ala Gln Lys Leu Ala Ala Ser Asp Leu Gly Phe Ala Leu Gly  
 275 280 285  
 Pro Arg Leu Asn Ala Ala Gly Arg Leu Asp Asp Met Ser Val Gly Val  
 290 295 300  
 Ala Leu Leu Leu Cys Asp Asn Ile Gly Glu Ala Arg Val Leu Ala Asn  
 305 310 315 320  
  
 Glu Leu Asp Ala Leu Asn Gln Thr Arg Lys Glu Ile Glu Gln Gly Met  
 325 330 335  
 Gln Ile Glu Ala Leu Thr Leu Cys Glu Lys Leu Glu Arg Ser Arg Asp  
 340 345 350  
 Thr Leu Pro Gly Gly Leu Ala Met Tyr His Pro Glu Trp His Gln Gly



<213> Artificial Sequence

<220><223> RecE

<400> 14

Met Ser Thr Lys Pro Leu Phe Leu Leu Arg Lys Ala Lys Lys Ser Ser  
 1 5 10 15

Gly Glu Pro Asp Val Val Leu Trp Ala Ser Asn Asp Phe Glu Ser Thr  
 20 25 30

Cys Ala Thr Leu Asp Tyr Leu Ile Val Lys Ser Gly Lys Lys Leu Ser  
 35 40 45

Ser Tyr Phe Lys Ala Val Ala Thr Asn Phe Pro Val Val Asn Asp Leu  
 50 55 60

Pro Ala Glu Gly Glu Ile Asp Phe Thr Trp Ser Glu Arg Tyr Gln Leu  
 65 70 75 80

Ser Lys Asp Ser Met Thr Trp Glu Leu Lys Pro Gly Ala Ala Pro Asp  
 85 90 95

Asn Ala His Tyr Gln Gly Asn Thr Asn Val Asn Gly Glu Asp Met Thr  
 100 105 110

Glu Ile Glu Glu Asn Met Leu Leu Pro Ile Ser Gly Gln Glu Leu Pro  
 115 120 125

Ile Arg Trp Leu Ala Gln His Gly Ser Glu Lys Pro Val Thr His Val  
 130 135 140

Ser Arg Asp Gly Leu Gln Ala Leu His Ile Ala Arg Ala Glu Glu Leu  
 145 150 155 160

Pro Ala Val Thr Ala Leu Ala Val Ser His Lys Thr Ser Leu Leu Asp  
 165 170 175

Pro Leu Glu Ile Arg Glu Leu His Lys Leu Val Arg Asp Thr Asp Lys  
 180 185 190

Val Phe Pro Asn Pro Gly Asn Ser Asn Leu Gly Leu Ile Thr Ala Phe  
 195 200 205

Phe Glu Ala Tyr Leu Asn Ala Asp Tyr Thr Asp Arg Gly Leu Leu Thr  
 210 215 220

Lys Glu Trp Met Lys Gly Asn Arg Val Ser His Ile Thr Arg Thr Ala  
 225                      230                      235                      240  
 Ser Gly Ala Asn Ala Gly Gly Gly Asn Leu Thr Asp Arg Gly Glu Gly  
                                  245                      250                      255  
 Phe Val His Asp Leu Thr Ser Leu Ala Arg Asp Val Ala Thr Gly Val  
                                  260                      265                      270  
 Leu Ala Arg Ser Met Asp Leu Asp Ile Tyr Asn Leu His Pro Ala His  
                                  275                      280                      285  
 Ala Lys Arg Ile Glu Glu Ile Ile Ala Glu Asn Lys Pro Pro Phe Ser  
                                  290                      295                      300  
  
 Val Phe Arg Asp Lys Phe Ile Thr Met Pro Gly Gly Leu Asp Tyr Ser  
 305                      310                      315                      320  
 Arg Ala Ile Val Val Ala Ser Val Lys Glu Ala Pro Ile Gly Ile Glu  
                                  325                      330                      335  
 Val Ile Pro Ala His Val Thr Glu Tyr Leu Asn Lys Val Leu Thr Glu  
                                  340                      345                      350  
 Thr Asp His Ala Asn Pro Asp Pro Glu Ile Val Asp Ile Ala Cys Gly  
                                  355                      360                      365  
 Arg Ser Ser Ala Pro Met Pro Gln Arg Val Thr Glu Glu Gly Lys Gln  
  
                                  370                      375                      380  
 Asp Asp Glu Glu Lys Pro Gln Pro Ser Gly Thr Thr Ala Val Glu Gln  
 385                      390                      395                      400  
 Gly Glu Ala Glu Thr Met Glu Pro Asp Ala Thr Glu His His Gln Asp  
                                  405                      410                      415  
 Thr Gln Pro Leu Asp Ala Gln Ser Gln Val Asn Ser Val Asp Ala Lys  
                                  420                      425                      430  
 Tyr Gln Glu Leu Arg Ala Glu Leu His Glu Ala Arg Lys Asn Ile Pro  
                                  435                      440                      445  
  
 Ser Lys Asn Pro Val Asp Asp Asp Lys Leu Leu Ala Ala Ser Arg Gly  
                                  450                      455                      460  
 Glu Phe Val Asp Gly Ile Ser Asp Pro Asn Asp Pro Lys Trp Val Lys

465                      470                      475                      480  
 Gly Ile Gln Thr Arg Asp Cys Val Tyr Gln Asn Gln Pro Glu Thr Glu  
    485                      490                      495  
 Lys Thr Ser Pro Asp Met Asn Gln Pro Glu Pro Val Val Gln Gln Glu  
    500                      505                      510  
 Pro Glu Ile Ala Cys Asn Ala Cys Gly Gln Thr Gly Gly Asp Asn Cys  
  
    515                      520                      525  
 Pro Asp Cys Gly Ala Val Met Gly Asp Ala Thr Tyr Gln Glu Thr Phe  
    530                      535                      540  
 Asp Glu Glu Ser Gln Val Glu Ala Lys Glu Asn Asp Pro Glu Glu Met  
 545                      550                      555                      560  
 Glu Gly Ala Glu His Pro His Asn Glu Asn Ala Gly Ser Asp Pro His  
    565                      570                      575  
 Arg Asp Cys Ser Asp Glu Thr Gly Glu Val Ala Asp Pro Val Ile Val  
    580                      585                      590  
  
 Glu Asp Ile Glu Pro Gly Ile Tyr Tyr Gly Ile Ser Asn Glu Asn Tyr  
    595                      600                      605  
 His Ala Gly Pro Gly Ile Ser Lys Ser Gln Leu Asp Asp Ile Ala Asp  
    610                      615                      620  
 Thr Pro Ala Leu Tyr Leu Trp Arg Lys Asn Ala Pro Val Asp Thr Thr  
 625                      630                      635                      640  
 Lys Thr Lys Thr Leu Asp Leu Gly Thr Ala Phe His Cys Arg Val Leu  
    645                      650                      655  
 Glu Pro Glu Glu Phe Ser Asn Arg Phe Ile Val Ala Pro Glu Phe Asn  
  
    660                      665                      670  
 Arg Arg Thr Asn Ala Gly Lys Glu Glu Glu Lys Ala Phe Leu Met Glu  
    675                      680                      685  
 Cys Ala Ser Thr Gly Lys Thr Val Ile Thr Ala Glu Glu Gly Arg Lys  
    690                      695                      700  
 Ile Glu Leu Met Tyr Gln Ser Val Met Ala Leu Pro Leu Gly Gln Trp  
 705                      710                      715                      720  
 Leu Val Glu Ser Ala Gly His Ala Glu Ser Ser Ile Tyr Trp Glu Asp





<212> PRT

<213> Artificial Sequence

<220><223> Lambda exonuclease

<400> 16

Met Thr Pro Asp Ile Ile Leu Gln Arg Thr Gly Ile Asp Val Arg Ala

1 5 10 15

Val Glu Gln Gly Asp Asp Ala Trp His Lys Leu Arg Leu Gly Val Ile

20 25 30

Thr Ala Ser Glu Val His Asn Val Ile Ala Lys Pro Arg Ser Gly Lys

35 40 45

Lys Trp Pro Asp Met Lys Met Ser Tyr Phe His Thr Leu Leu Ala Glu

50 55 60

Val Cys Thr Gly Val Ala Pro Glu Val Asn Ala Lys Ala Leu Ala Trp

65 70 75 80

Gly Lys Gln Tyr Glu Asn Asp Ala Arg Thr Leu Phe Glu Phe Thr Ser

85 90 95

Gly Val Asn Val Thr Glu Ser Pro Ile Ile Tyr Arg Asp Glu Ser Met

100 105 110

Arg Thr Ala Cys Ser Pro Asp Gly Leu Cys Ser Asp Gly Asn Gly Leu

115 120 125

Glu Leu Lys Cys Pro Phe Thr Ser Arg Asp Phe Met Lys Phe Arg Leu

130 135 140

Gly Gly Phe Glu Ala Ile Lys Ser Ala Tyr Met Ala Gln Val Gln Tyr

145 150 155 160

Ser Met Trp Val Thr Arg Lys Asn Ala Trp Tyr Phe Ala Asn Tyr Asp

165 170 175

Pro Arg Met Lys Arg Glu Gly Leu His Tyr Val Val Ile Glu Arg Asp

180 185 190

Glu Lys Tyr Met Ala Ser Phe Asp Glu Ile Val Pro Glu Phe Ile Glu

195 200 205

Lys Met Asp Glu Ala Leu Ala Glu Ile Gly Phe Val Phe Gly Glu Gln

210 215 220

Trp Arg

225

<210> 17

<211> 80

<212> PRT

<213> Artificial Sequence

<220><223> Exonuclease VII small unit

<400> 17

Met Pro Lys Lys Asn Glu Ala Pro Ala Ser Phe Glu Lys Ala Leu Ser

1 5 10 15

Glu Leu Glu Gln Ile Val Thr Arg Leu Glu Ser Gly Asp Leu Pro Leu

20 25 30

Glu Glu Ala Leu Asn Glu Phe Glu Arg Gly Val Gln Leu Ala Arg Gln

35 40 45

Gly Gln Ala Lys Leu Gln Gln Ala Glu Gln Arg Val Gln Ile Leu Leu

50 55 60

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

65 70 75 80

<210> 18

<211> 456

<212> PRT

<213> Artificial Sequence

<220><223> Exonuclease VII large unit

<400> 18

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

1 5 10 15

Gln Thr Val Arg Leu Leu Leu Glu His Glu Met Gly Gln Val Trp Ile

20 25 30

Ser Gly Glu Ile Ser Asn Phe Thr Gln Pro Ala Ser Gly His Trp Tyr

35 40 45

Phe Thr Leu Lys Asp Asp Thr Ala Gln Val Arg Cys Ala Met Phe Arg

50 55 60



Arg Leu Gln Lys Arg Met Ser Phe Ala Leu Glu Asn Gln Leu Lys Arg  
 325 330 335

Thr Gly Gln Gln Gln Gln Arg Leu Thr Gln Arg Leu Asn Gln Gln Asn  
 340 345 350

Pro Gln Pro Lys Ile His Arg Ala Gln Thr Arg Ile Gln Gln Leu Glu  
 355 360 365

Tyr Arg Leu Ala Glu Thr Leu Arg Ala Gln Leu Ser Ala Thr Arg Glu  
 370 375 380

Arg Phe Gly Asn Ala Val Thr His Leu Glu Ala Val Ser Pro Leu Ser  
 385 390 395 400

Thr Leu Ala Arg Gly Tyr Ser Val Thr Thr Ala Thr Asp Gly Asn Val  
 405 410 415

Leu Lys Lys Val Lys Gln Val Lys Ala Gly Glu Met Leu Thr Thr Arg  
 420 425 430

Leu Glu Asp Gly Trp Ile Glu Ser Glu Val Lys Asn Ile Gln Pro Val  
 435 440 445

Lys Lys Ser Arg Lys Lys Val His  
 450 455

<210> 19

<211> 4101

<212> DNA

<213> Artificial Sequence

<220><223> Cas9 nucleotides sequence of Streptococcus pyogenes M1 GAS

<400> 19

gacaagaagt acagcatcgg cctggacatc ggcaccaact ctgtgggctg ggccgtgatc 60

accgacgagt acaaggtgcc cagcaagaaa ttcaaggtgc tgggcaacac cgaccggcac 120

agcatcaaga agaacctgat cggagccctg ctgttcgaca gcggcgaaac agccgaggcc 180

accggctga agagaaccgc cagaagaaga tacaccagac ggaagaaccg gatctgctat 240

ctgcaagaga tcttcagcaa cgagatggcc aaggtggacg acagcttctt ccacagactg 300

gaagatcct tcttgggtgga agaggataag aagcacgagc ggcaccccat ctteggcaac 360

atcgtggacg aggtggccta ccacgagaag taccaccca tctaccacct gagaaagaaa 420

ctggtggaca gcaccgacaa ggccgacctg cggctgatct atctggcctt ggcccacatg 480

atcaagtcc ggggccactt cctgatcgag ggcgacctga accccgacaa cagcgactg 540

gacaagctgt tcatccagct ggtgcagacc tacaaccagc tgttcgagga aaaccccatc 600

aacgccagcg gcgtggacgc caaggccatc ctgtctgcca gactgagcaa gacgagacgg 660

ctggaaaatc tgatcgccca gctgcccggc gagaagaaga atggcctgtt cggaaacctg 720

attgccctga gectgggctt gacccccaac ttcaagagca acttcgacct ggcccaggat 780

gccaaactgc agctgagcaa ggacacctac gacgacgacc tggacaacct gctggcccag 840

atcggcgacc agtacccga cctgtttctg gccccaaga acctgtccga cgccatcctg 900

ctgagcgaca tctgagagt gaacaccgag atcaccaagg cccccctgag cgcctctatg 960

atcaagagat acgacgagca ccaccaggac ctgacctgc tgaaagctct cgtgcggcag 1020

cagctgcctg agaagtacaa agagattttc ttcgaccaga gcaagaacgg ctacgccggc 1080

tacattgacg gcggagccag ccaggaagag ttctacaagt tcatcaagcc catcctggaa 1140

aagatggacg gcaccgagga actgctcgtg aagctgaaca gagaggacct gctgcggaag 1200

cagcggacct tcgacaacgg cagcatcccc caccagatcc acctgggaga getgcacgcc 1260

attctcggc ggaggaaga tttttacca ttctgaagg acaaccggga aaagatcgag 1320

aagatcctga cttccgcat ccctactac gtgggcctc tggccagggg aaacagcaga 1380

ttcgctgga tgaccagaaa gagcgaggaa accatcacc cctggaactt cgaggaagtg 1440

gtggacaagg gccttccgc ccagagcttc atcgagcgga tgaccaactt cgataagaac 1500

ctgccaacg agaaggtgct gccaagcac agcctgctgt acgagtaact caccgtgtat 1560

aacgagctga ccaaagtga atactgacc gagggaatga gaaagccgc cttcctgagc 1620

ggcgagcaga aaaaggccat cgtggacctg ctgttcaaga ccaaccgga agtgaccgtg 1680

aagcagctga aagaggacta cttcaagaaa atcgagtgct tcgactccgt ggaaatctec 1740

ggcgtggaag atcggttcaa cgcctccctg ggcacatacc acgatctgct gaaaattatc 1800

aaggacaagg acttctgga caatgaggaa aacgaggaca ttctggaaga tatcgtgctg 1860

accctgacac tgtttgagga cagagagatg atcgaggaac ggctgaaaac ctatgccac 1920

ctgttcgacg acaaagtgat gaagcagctg aagcggcgga gatacaccgg ctggggcagg 1980

ctgagccgga agctgatcaa ggcatccgg gacaagcagt ccggcaagac aatcctggat 2040

ttcctgaagt ccgacggctt cgccaacaga aacttcatgc agctgatcca cgacgacagc 2100

ctgacctta aagaggacat ccagaaagcc caggtgtccg gccagggcga tagcctgcac 2160

gagcacattg ccaatctggc cggcagcccc gccattaaga agggcatcct gcagacagtg 2220  
 aaggtggtgg acgagctcgt gaaagtgatg ggccggcaca agcccagaaa catcgtgatc 2280  
 gaaatggcca gagagaacca gaccacccag aaggacaga agaacagccg cgagagaatg 2340  
 aagcggatcg aagaggcat caaagagctg ggcagccaga tcctgaaaga acaccccggtg 2400  
 gaaaacacc agctgcagaa cgagaagctg tacctgtact acctgcagaa tggcgggat 2460  
 atgtacgtgg accaggaact ggacatcaac cggctgtccg actacgatgt ggacatc 2520  
 gtgcctcaga gctttctgaa ggacgactcc atcgacaaca aggtgctgac cagaagcgac 2580  
  
 aagaaccggg gcaagagcga caactgccc tccgaagagg tcgtgaagaa gatgaagaac 2640  
 tactggcggc agctgctgaa cgccaagctg attaccaga gaaagtcca caatctgacc 2700  
 aaggccgaga gaggcggcct gagcgaactg gataaggccg gcttcatcaa gagacagctg 2760  
 gtggaaccc ggagatcac aaagcagctg gcacagatcc tggactccg gatgaacact 2820  
 aagtacgacg agaatgaaa gctgatccgg gaagtgaag tgatcacct gaagtccaag 2880  
 ctggtgtccg atttccgaa ggatttccag ttttcaaag tgcgcgagat caacaactac 2940  
 caccacgcc acgagccta cctgaacgcc gtctgggaa ccgcctgat caaaaagtac 3000  
  
 cctaagctgg aaagcgagt cgtgtacggc gactacaagg tgtacgacgt gcggaagatg 3060  
 atcgccaaga gcgagcagga aatcggaag gctaccgcca agtacttctt ctacagcaac 3120  
 atcatgaact tttcaagc cgagattacc ctggccaac gcgagatccg gaagcggcct 3180  
 ctgatcgaga caaacggca aaccggggag atcgtgtggg ataaggccg ggattttgcc 3240  
 accgtgcgga aagtgtgag catgcccga gtgaatatcg tgaaaaagac cgaggtgcag 3300  
 acaggcggct tcagcaaga gtctatcctg cccaagagga acagcgataa gctgatgcc 3360  
 agaaagaagg actgggacc taagaagtac ggcggcttcg acagcccac cgtggcctat 3420  
  
 tctgtgctgg tgggtggcaa agtgaaaaag ggcaagtcca agaaactgaa gagtgtgaaa 3480  
 gagctgctgg ggatcacat catgaaaga agcagcttcg agaagaatcc catcgacttt 3540  
 ctggaagcca agggctaaa agaagtgaaa aaggacctga tcatcaagct gcctaagtac 3600  
 tccctgttcg agctgaaaa cggccggaag agaatgctgg cctctgccg gaaactgcag 3660  
 aaggaaacg aactggcct gccctcaaa tatgtgaact tcctgtacct ggccagccac 3720  
 tatgagaagc tgaaggctc ccccgaggat aatgagcaga aacagctgtt tgtggaacag 3780  
 cacaagcact acctggacga gatcatcgag cagatcagcg agttctcaa gagagtgatc 3840  
  
 ctggccgacg ctaatctgga caaagtgtg tccgctaca acaagcaccg ggataagccc 3900  
 atcagagagc agccgagaa tatcatccac ctgtttacc tgaccaatct gggagcccct 3960  
 gccgcttca agtacttga caccacatc gaccggaaga ggtacaccag caccaagag 4020

gtgctggacg ccaccctgat ccaccagagc atcaccggcc tgtacgagac acggatcgac 4080  
ctgtctcagc tgggaggcga c 4101

<210> 20  
<211> 1450  
<212> PRT  
<213> Artificial Sequence  
<220><223> Cas9 amino acids sequence of Streptococcus pyogenes M1 GAS

<400> 20  
Asp Tyr Lys Asp His Asp Gly Asp Tyr Lys Asp His Asp Ile Asp Tyr  
1 5 10 15  
Lys Asp Asp Asp Asp Lys Met Glu Thr Ala Pro Lys Lys Lys Arg Lys  
20 25 30  
Val Gly Ile His Gly Val Pro Ala Ala Asp Lys Lys Tyr Ser Ile Gly  
35 40 45  
Leu Asp Ile Gly Thr Asn Ser Val Gly Trp Ala Val Ile Thr Asp Glu  
50 55 60  
Tyr Lys Val Pro Ser Lys Lys Phe Lys Val Leu Gly Asn Thr Asp Arg  
65 70 75 80  
His Ser Ile Lys Lys Asn Leu Ile Gly Ala Leu Leu Phe Asp Ser Gly  
85 90 95  
Glu Thr Ala Glu Ala Thr Arg Leu Lys Arg Thr Ala Arg Arg Arg Tyr  
100 105 110  
Thr Arg Arg Lys Asn Arg Ile Cys Tyr Leu Gln Glu Ile Phe Ser Asn  
115 120 125  
Glu Met Glu Thr Ala Lys Val Asp Asp Ser Phe Phe His Arg Leu Glu  
130 135 140  
Glu Ser Phe Leu Val Glu Glu Asp Lys Lys His Glu Arg His Pro Ile  
145 150 155 160  
Phe Gly Asn Ile Val Asp Glu Val Ala Tyr His Glu Lys Tyr Pro Thr  
165 170 175  
Ile Tyr His Leu Arg Lys Lys Leu Val Asp Ser Thr Asp Lys Ala Asp  
180 185 190

Leu Arg Leu Ile Tyr Leu Ala Leu Ala His Met Glu Thr Ile Lys Phe  
 195 200 205  
 Arg Gly His Phe Leu Ile Glu Gly Asp Leu Asn Pro Asp Asn Ser Asp  
 210 215 220  
 Val Asp Lys Leu Phe Ile Gln Leu Val Gln Thr Tyr Asn Gln Leu Phe  
 225 230 235 240  
 Glu Glu Asn Pro Ile Asn Ala Ser Gly Val Asp Ala Lys Ala Ile Leu  
 245 250 255  
 Ser Ala Arg Leu Ser Lys Ser Arg Arg Leu Glu Asn Leu Ile Ala Gln  
 260 265 270  
 Leu Pro Gly Glu Lys Lys Asn Gly Leu Phe Gly Asn Leu Ile Ala Leu  
 275 280 285  
  
 Ser Leu Gly Leu Thr Pro Asn Phe Lys Ser Asn Phe Asp Leu Ala Glu  
 290 295 300  
 Asp Ala Lys Leu Gln Leu Ser Lys Asp Thr Tyr Asp Asp Asp Leu Asp  
 305 310 315 320  
 Asn Leu Leu Ala Gln Ile Gly Asp Gln Tyr Ala Asp Leu Phe Leu Ala  
 325 330 335  
 Ala Lys Asn Leu Ser Asp Ala Ile Leu Leu Ser Asp Ile Leu Arg Val  
 340 345 350  
 Asn Thr Glu Ile Thr Lys Ala Pro Leu Ser Ala Ser Met Glu Thr Ile  
 355 360 365  
 Lys Arg Tyr Asp Glu His His Gln Asp Leu Thr Leu Leu Lys Ala Leu  
 370 375 380  
 Val Arg Gln Gln Leu Pro Glu Lys Tyr Lys Glu Ile Phe Phe Asp Gln  
 385 390 395 400  
 Ser Lys Asn Gly Tyr Ala Gly Tyr Ile Asp Gly Gly Ala Ser Gln Glu  
 405 410 415  
 Glu Phe Tyr Lys Phe Ile Lys Pro Ile Leu Glu Lys Met Glu Thr Asp  
 420 425 430  
  
 Gly Thr Glu Glu Leu Leu Val Lys Leu Asn Arg Glu Asp Leu Leu Arg

435                      440                      445  
 Lys Gln Arg Thr Phe Asp Asn Gly Ser Ile Pro His Gln Ile His Leu  
 450                      455                      460  
 Gly Glu Leu His Ala Ile Leu Arg Arg Gln Glu Asp Phe Tyr Pro Phe  
 465                      470                      475                      480  
 Leu Lys Asp Asn Arg Glu Lys Ile Glu Lys Ile Leu Thr Phe Arg Ile  
 485                      490                      495  
 Pro Tyr Tyr Val Gly Pro Leu Ala Arg Gly Asn Ser Arg Phe Ala Trp  
  
 500                      505                      510  
 Met Glu Thr Thr Arg Lys Ser Glu Glu Thr Ile Thr Pro Trp Asn Phe  
 515                      520                      525  
 Glu Glu Val Val Asp Lys Gly Ala Ser Ala Gln Ser Phe Ile Glu Arg  
 530                      535                      540  
 Met Glu Thr Thr Asn Phe Asp Lys Asn Leu Pro Asn Glu Lys Val Leu  
 545                      550                      555                      560  
 Pro Lys His Ser Leu Leu Tyr Glu Tyr Phe Thr Val Tyr Asn Glu Leu  
 565                      570                      575  
  
 Thr Lys Val Lys Tyr Val Thr Glu Gly Met Glu Thr Arg Lys Pro Ala  
 580                      585                      590  
 Phe Leu Ser Gly Glu Gln Lys Lys Ala Ile Val Asp Leu Leu Phe Lys  
 595                      600                      605  
 Thr Asn Arg Lys Val Thr Val Lys Gln Leu Lys Glu Asp Tyr Phe Lys  
 610                      615                      620  
 Lys Ile Glu Cys Phe Asp Ser Val Glu Ile Ser Gly Val Glu Asp Arg  
 625                      630                      635                      640  
 Phe Asn Ala Ser Leu Gly Thr Tyr His Asp Leu Leu Lys Ile Ile Lys  
  
 645                      650                      655  
 Asp Lys Asp Phe Leu Asp Asn Glu Glu Asn Glu Asp Ile Leu Glu Asp  
 660                      665                      670  
 Ile Val Leu Thr Leu Thr Leu Phe Glu Asp Arg Glu Met Glu Thr Ile  
 675                      680                      685  
 Glu Glu Arg Leu Lys Thr Tyr Ala His Leu Phe Asp Asp Lys Val Met



Lys Lys Met Glu Thr Lys Asn Tyr Trp Arg Gln Leu Leu Asn Ala Lys  
 945                      950                      955                      960  
 Leu Ile Thr Gln Arg Lys Phe Asp Asn Leu Thr Lys Ala Glu Arg Gly  
                                  965                      970                      975  
 Gly Leu Ser Glu Leu Asp Lys Ala Gly Phe Ile Lys Arg Gln Leu Val  
                                  980                      985                      990  
 Glu Thr Arg Gln Ile Thr Lys His Val Ala Gln Ile Leu Asp Ser Arg  
                                  995                      1000                      1005  
  
 Met Glu Thr Asn Thr Lys Tyr Asp Glu Asn Asp Lys Leu Ile Arg Glu  
                                  1010                      1015                      1020  
 Val Lys Val Ile Thr Leu Lys Ser Lys Leu Val Ser Asp Phe Arg Lys  
 1025                      1030                      1035                      1040  
 Asp Phe Gln Phe Tyr Lys Val Arg Glu Ile Asn Asn Tyr His His Ala  
                                  1045                      1050                      1055  
 His Asp Ala Tyr Leu Asn Ala Val Val Gly Thr Ala Leu Ile Lys Lys  
                                  1060                      1065                      1070  
 Tyr Pro Lys Leu Glu Ser Glu Phe Val Tyr Gly Asp Tyr Lys Val Tyr  
  
                                  1075                      1080                      1085  
 Asp Val Arg Lys Met Glu Thr Ile Ala Lys Ser Glu Gln Glu Ile Gly  
                                  1090                      1095                      1100  
 Lys Ala Thr Ala Lys Tyr Phe Phe Tyr Ser Asn Ile Met Glu Thr Asn  
 1105                      1110                      1115                      1120  
 Phe Phe Lys Thr Glu Ile Thr Leu Ala Asn Gly Glu Ile Arg Lys Arg  
                                  1125                      1130                      1135  
 Pro Leu Ile Glu Thr Asn Gly Glu Thr Gly Glu Ile Val Trp Asp Lys  
                                  1140                      1145                      1150  
  
 Gly Arg Asp Phe Ala Thr Val Arg Lys Val Leu Ser Met Glu Thr Pro  
                                  1155                      1160                      1165  
 Gln Val Asn Ile Val Lys Lys Thr Glu Val Gln Thr Gly Gly Phe Ser  
                                  1170                      1175                      1180  
 Lys Glu Ser Ile Leu Pro Lys Arg Asn Ser Asp Lys Leu Ile Ala Arg  
 1185                      1190                      1195                      1200

Lys Lys Asp Trp Asp Pro Lys Lys Tyr Gly Gly Phe Asp Ser Pro Thr  
 1205 1210 1215  
 Val Ala Tyr Ser Val Leu Val Val Ala Lys Val Glu Lys Gly Lys Ser  
 1220 1225 1230  
 Lys Lys Leu Lys Ser Val Lys Glu Leu Leu Gly Ile Thr Ile Met Glu  
 1235 1240 1245  
 Thr Glu Arg Ser Ser Phe Glu Lys Asn Pro Ile Asp Phe Leu Glu Ala  
 1250 1255 1260  
 Lys Gly Tyr Lys Glu Val Lys Lys Asp Leu Ile Ile Lys Leu Pro Lys  
 1265 1270 1275 1280  
 Tyr Ser Leu Phe Glu Leu Glu Asn Gly Arg Lys Arg Met Glu Thr Leu  
 1285 1290 1295  
  
 Ala Ser Ala Gly Glu Leu Gln Lys Gly Asn Glu Leu Ala Leu Pro Ser  
 1300 1305 1310  
 Lys Tyr Val Asn Phe Leu Tyr Leu Ala Ser His Tyr Glu Lys Leu Lys  
 1315 1320 1325  
 Gly Ser Pro Glu Asp Asn Glu Gln Lys Gln Leu Phe Val Glu Gln His  
 1330 1335 1340  
 Lys His Tyr Leu Asp Glu Ile Ile Glu Gln Ile Ser Glu Phe Ser Lys  
 1345 1350 1355 1360  
 Arg Val Ile Leu Ala Asp Ala Asn Leu Asp Lys Val Leu Ser Ala Tyr  
 1365 1370 1375  
 Asn Lys His Arg Asp Lys Pro Ile Arg Glu Gln Ala Glu Asn Ile Ile  
 1380 1385 1390  
 His Leu Phe Thr Leu Thr Asn Leu Gly Ala Pro Ala Ala Phe Lys Tyr  
 1395 1400 1405  
 Phe Asp Thr Thr Ile Asp Arg Lys Arg Tyr Thr Ser Thr Lys Glu Val  
 1410 1415 1420  
 Leu Asp Ala Thr Leu Ile His Gln Ser Ile Thr Gly Leu Tyr Glu Thr  
 1425 1430 1435 1440  
  
 Arg Ile Asp Leu Ser Gln Leu Gly Gly Asp

	1445	1450	
<210>	21		
<211>	3903		
<212>	DNA		
<213>	Artificial Sequence		
<220><223>	Cpf1 nucleotides sequences, FnCpf1, NCBI accession no. CP009633.1		
<400>	21		
atgtcaattt atcaagaatt tgtaataaa tatagtttaa gtaaaactct aagatttgag			60
ttaatccac agggtaaaac acttgaaaac ataaaagcaa gaggtttgat tttagatgat			120
gagaaaagag ctaaagacta caaaaaggct aaacaaataa ttgataaata tcatcagttt			180
tttatagagg agatattaag ttcggtttgt attagcgaag atttattaca aaactattct			240
gatgtttatt ttaaacttaa aaagagtgat gatgataatc tacaaaaaga ttttaaagt			300
gcaaaagata cgataaagaa acaaatatct gaatatataa aggactcaga gaaatttaag			360
aatttgttta atcaaaacct tatc gatgct aaaaaaggc aagagtcaga ttttaattcta			420
tggctaaagc aatctaagga taatggtata gaactattta aagccaatag tgatatcaca			480
gatatagatg aggcgttaga aataatcaaa tcttttaaag gttggacaac ttattttaag			540
ggttttcatg aaaaatagaaa aaatgtttat agtagcaatg atattcctac atctattatt			600
tataggatag tagatgataa tttgcctaaa tttctagaaa ataaagctaa gtatgagagt			660
ttaaaagaca aagctccaga agctataaac tatgaacaaa ttaaaaaaga tttggcagaa			720
gagctaacct ttgatattga ctacaaaaca tctgaagtta atcaaagagt tttttcactt			780
gatgaagttt ttgatagtc aaactttaat aattatctaa atcaaagtg tattactaaa			840
tttaatacta ttattggtgg taaatttgta aatggtgaaa atacaaagag aaaaggtata			900
aatgaatata taaatctata ctacacagca ataaatgata aaactctca aaaatataaa			960
atgagtgttt tatttaagca aattttaagt gatacagaat ctaaactttt tgtaattgat			1020
aagttagaag atgatagtg tgtagttaca acgatgcaaa gtttttatga gcaaatagca			1080
gcttttaaaa cagtagaaga aaaatctatt aaagaaacac tatctttatt atttgatgat			1140
ttaaaagctc aaaaacttga tttgagtaaa atttatttta aaaatgataa atctcttact			1200
gatctatcac aacaagtttt tgatgattat agtgttattg gtacagcggg actagaatat			1260
ataactcaac aaatagcacc taaaaatctt gataacccta gtaagaaaga gcaagaatta			1320
atagccaaaa aaactgaaaa agcaaaatac ttatctctag aaactataaa gcttgccctta			1380
gaagaattta ataagcatag agatatagat aaacagtgta ggtttgaaga aatacttgca			1440

aactttgcfg ctattccgat gatatttgat gaaatagctc aaaacaaaga caatttggca 1500

cagatatcta tcaaatatca aatcaaggt aaaaaagacc tacttcaagc tagtgcggaa 1560

gatgatgta aagctatcaa ggatctttta gatcaaaacta ataatctctt acataaaacta 1620

aaaatatttc atattagtca gtcagaagat aaggcaata ttttagacaa ggatgagcat 1680

ttttatctag tatttgagga gtgctacttt gagctagcga atagtgcc tctttataac 1740

aaaattagaa actatataac tcaaaagcca tatagtgatg agaaatftaa getcaatftt 1800

gagaactcga ctttggctaa tggttgggat aaaaataaag agcctgacaa tacggcaatt 1860

ttatttatca aagatgataa atattatctg ggtgtgatga ataagaaaa taacaaaata 1920

tttgatgata aagctatcaa agaaaataaa ggcgagggtt ataaaaaat tgtttataaa 1980

cttttacctg gcgcaaataa aatgttacct aaggttttct tttctgctaa atctataaaa 2040

ttttataate ctagtgaaga tatactttaga ataagaaatc attccacaca tacaaaaaat 2100

ggtagtcctc aaaaaggata tgaaaattt gagtttaata ttgaagattg ccgaaaattt 2160

atagatfttt ataacagtc tataagtaag catccggagt ggaaagattt tggattttaga 2220

ttttctgata ctcaaagata taattctata gatgaatftt atagagaagt tgaaaatcaa 2280

ggctacaaac taacttttga aatatatca gagagctata ttgatagcgt agttaatcag 2340

ggtaaattgt acctattcca aatctataat aaagatfttt cagcttatag caaagggcga 2400

ccaaatctac atactttata ttggaaagcg ctgtttgatg agagaaatct tcaagatgtg 2460

gtttataagc taaatggiga ggcagagctt ttttatcgta aacaatcaat acctaaaaa 2520

atcactcacc cagctaaaga ggcaatagct aataaaaaa aagataatcc taaaaaagag 2580

agtgtttttg aatgatftt aatcaaagat aaacgcttta ctgaagataa gtttttcttt 2640

cactgccta ttacaatcaa ttttaaatct agtggagcta ataagtttaa tgatgaaatc 2700

aatttattgc taaaagaaaa agcaaatgat gttcatatat taagtataga tagaggtgaa 2760

agacatttag ctactatac tttgtagat ggtaaaggca atatcatcaa acaagatact 2820

ttcaacatca ttgtaatga tagaatgaaa acaaaactacc atgataagct tgctgcaata 2880

gagaaagata gggattcagc taggaaagac tggaaaaaga taaataacat caaagagatg 2940

aaagagggtc atctatctca ggtagtftcat gaaatagcta agctagtftat agagtataat 3000

gctattgtgg tttttgagga tttaaatftt ggattftaaa gagggcgtft caaggttagag 3060

aagcaggtct atcaaaagtt agaaaaaatg ctaattgaga aactaaacta tctagtfttc 3120

aaagataatg agtttgataa aactggggga gtgcttagag cttatcagct aacagcacct 3180

tttgagactt ttaaaaagat gggtaaacaa acaggtatta tctactatgt accagctggt 3240  
 tttacttcaa aaatttgc tgaactggt tttgtaaac agttatatcc taagtatgaa 3300  
 agtgcagca aatctcaaga gttctttagt aagtttgaca agatttgta taacctgat 3360  
 aagggtatt ttgagtttag ttttgattat aaaaactttg gtgacaaggc tgccaaaggc 3420  
 aagtggacta tagctagctt tgggagtaga ttgattaact ttagaattc agataaaaat 3480  
 cataattggg ataactcgaga agtttatcca actaaagagt tggagaaatt gctaaaagat 3540  
 tattctatcg aatatgggca tggcgaatgt atcaaagcag ctatttgcgg tgagagcgac 3600

aaaaagtttt ttgctaagct aactagtgtc ctaaatacta tcttacaat gcgtaactca 3660  
 aaaacaggta ctgagttaga ttatctaatt tcaccagtag cagatgtaaa tggcaatttc 3720  
 tttgattcgc gacagggccc aaaaaatag cctcaagatg ctgatgcaa tggtgcttat 3780  
 catattgggc taaaaggtct gatgctacta ggtaggatca aaaataatca agagggcaaa 3840  
 aaactcaatt tggttatcaa aaatgaagag tattttgagt tcgtgcagaa taggaataac 3900  
 taa 3903

<210> 22  
 <211> 1300  
 <212> PRT

<213> Artificial Sequence

<220><223> Cpf1 amino acids sequences, FnCpf1, NCBI accession no. CP009633.1

<400> 22

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

1 5 10 15

Leu Arg Phe Glu Leu Ile Pro Gln Gly Lys Thr Leu Glu Asn Ile Lys

20 25 30

Ala Arg Gly Leu Ile Leu Asp Asp Glu Lys Arg Ala Lys Asp Tyr Lys

35 40 45

Lys Ala Lys Gln Ile Ile Asp Lys Tyr His Gln Phe Phe Ile Glu Glu

50 55 60

Ile Leu Ser Ser Val Cys Ile Ser Glu Asp Leu Leu Gln Asn Tyr Ser

65 70 75 80

Asp Val Tyr Phe Lys Leu Lys Lys Ser Asp Asp Asp Asn Leu Gln Lys

85 90 95

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



Gln Ser Phe Tyr Glu Gln Ile Ala Ala Phe Lys Thr Val Glu Glu Lys  
 355 360 365

Ser Ile Lys Glu Thr Leu Ser Leu Leu Phe Asp Asp Leu Lys Ala Gln  
 370 375 380

Lys Leu Asp Leu Ser Lys Ile Tyr Phe Lys Asn Asp Lys Ser Leu Thr  
 385 390 395 400

Asp Leu Ser Gln Gln Val Phe Asp Asp Tyr Ser Val Ile Gly Thr Ala  
 405 410 415

Val Leu Glu Tyr Ile Thr Gln Gln Ile Ala Pro Lys Asn Leu Asp Asn  
 420 425 430

Pro Ser Lys Lys Glu Gln Glu Leu Ile Ala Lys Lys Thr Glu Lys Ala  
 435 440 445

Lys Tyr Leu Ser Leu Glu Thr Ile Lys Leu Ala Leu Glu Glu Phe Asn  
 450 455 460

Lys His Arg Asp Ile Asp Lys Gln Cys Arg Phe Glu Glu Ile Leu Ala  
 465 470 475 480

Asn Phe Ala Ala Ile Pro Met Ile Phe Asp Glu Ile Ala Gln Asn Lys  
 485 490 495

Asp Asn Leu Ala Gln Ile Ser Ile Lys Tyr Gln Asn Gln Gly Lys Lys  
 500 505 510

Asp Leu Leu Gln Ala Ser Ala Glu Asp Asp Val Lys Ala Ile Lys Asp  
 515 520 525

Leu Leu Asp Gln Thr Asn Asn Leu Leu His Lys Leu Lys Ile Phe His  
 530 535 540

Ile Ser Gln Ser Glu Asp Lys Ala Asn Ile Leu Asp Lys Asp Glu His  
 545 550 555 560

Phe Tyr Leu Val Phe Glu Glu Cys Tyr Phe Glu Leu Ala Asn Ile Val  
 565 570 575

Pro Leu Tyr Asn Lys Ile Arg Asn Tyr Ile Thr Gln Lys Pro Tyr Ser  
 580 585 590

Asp Glu Lys Phe Lys Leu Asn Phe Glu Asn Ser Thr Leu Ala Asn Gly  
 595 600 605

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

850                      855                      860  
 Tyr Asp Leu Ile Lys Asp Lys Arg Phe Thr Glu Asp Lys Phe Phe Phe  
 865                      870                      875                      880  
 His Cys Pro Ile Thr Ile Asn Phe Lys Ser Ser Gly Ala Asn Lys Phe  
                                  885                      890                      895  
 Asn Asp Glu Ile Asn Leu Leu Leu Lys Glu Lys Ala Asn Asp Val His  
                                  900                      905                      910  
 Ile Leu Ser Ile Asp Arg Gly Glu Arg His Leu Ala Tyr Tyr Thr Leu  
  
                                  915                      920                      925  
 Val Asp Gly Lys Gly Asn Ile Ile Lys Gln Asp Thr Phe Asn Ile Ile  
                                  930                      935                      940  
 Gly Asn Asp Arg Met Lys Thr Asn Tyr His Asp Lys Leu Ala Ala Ile  
 945                      950                      955                      960  
 Glu Lys Asp Arg Asp Ser Ala Arg Lys Asp Trp Lys Lys Ile Asn Asn  
                                  965                      970                      975  
 Ile Lys Glu Met Lys Glu Gly Tyr Leu Ser Gln Val Val His Glu Ile  
                                  980                      985                      990  
  
 Ala Lys Leu Val Ile Glu Tyr Asn Ala Ile Val Val Phe Glu Asp Leu  
                                  995                      1000                      1005  
 Asn Phe Gly Phe Lys Arg Gly Arg Phe Lys Val Glu Lys Gln Val Tyr  
                                  1010                      1015                      1020  
 Gln Lys Leu Glu Lys Met Leu Ile Glu Lys Leu Asn Tyr Leu Val Phe  
 1025                      1030                      1035                      1040  
 Lys Asp Asn Glu Phe Asp Lys Thr Gly Gly Val Leu Arg Ala Tyr Gln  
                                  1045                      1050                      1055  
 Leu Thr Ala Pro Phe Glu Thr Phe Lys Lys Met Gly Lys Gln Thr Gly  
  
                                  1060                      1065                      1070  
 Ile Ile Tyr Tyr Val Pro Ala Gly Phe Thr Ser Lys Ile Cys Pro Val  
                                  1075                      1080                      1085  
 Thr Gly Phe Val Asn Gln Leu Tyr Pro Lys Tyr Glu Ser Val Ser Lys  
                                  1090                      1095                      1100  
 Ser Gln Glu Phe Phe Ser Lys Phe Asp Lys Ile Cys Tyr Asn Leu Asp

1105                    1110                    1115                    1120  
 Lys Gly Tyr Phe Glu Phe Ser Phe Asp Tyr Lys Asn Phe Gly Asp Lys  
                           1125                    1130                    1135

Ala Ala Lys Gly Lys Trp Thr Ile Ala Ser Phe Gly Ser Arg Leu Ile  
                           1140                    1145                    1150

Asn Phe Arg Asn Ser Asp Lys Asn His Asn Trp Asp Thr Arg Glu Val  
                           1155                    1160                    1165

Tyr Pro Thr Lys Glu Leu Glu Lys Leu Leu Lys Asp Tyr Ser Ile Glu  
                           1170                    1175                    1180

Tyr Gly His Gly Glu Cys Ile Lys Ala Ala Ile Cys Gly Glu Ser Asp  
 1185                    1190                    1195                    1200

Lys Lys Phe Phe Ala Lys Leu Thr Ser Val Leu Asn Thr Ile Leu Gln

1205                    1210                    1215

Met Arg Asn Ser Lys Thr Gly Thr Glu Leu Asp Tyr Leu Ile Ser Pro  
                           1220                    1225                    1230

Val Ala Asp Val Asn Gly Asn Phe Phe Asp Ser Arg Gln Ala Pro Lys  
                           1235                    1240                    1245

Asn Met Pro Gln Asp Ala Asp Ala Asn Gly Ala Tyr His Ile Gly Leu  
                           1250                    1255                    1260

Lys Gly Leu Met Leu Leu Gly Arg Ile Lys Asn Asn Gln Glu Gly Lys  
 1265                    1270                    1275                    1280

Lys Leu Asn Leu Val Ile Lys Asn Glu Glu Tyr Phe Glu Phe Val Gln  
                           1285                    1290                    1295

Asn Arg Asn Asn

1300

<210> 23

<211> 3459

<212> DNA

<213> Artificial Sequence

<220><223> C2C2 nucleotides sequences, Leptotrichia wadei's Cas13a

<400> 23

atgaaagtga ccaaggtcga cggcatcagc cacaagaagt acatcgaaga gggcaagctc

60

gtgaagtcca ccagcgagga aaaccggacc agcgagagac tgagcgagct gctgagcatc 120  
 cggctggaca tctacatcaa gaaccccgac aacgcctccg aggaagagaa cggatcaga 180

agagagaacc tgaagaagtt ctttagcaac aaggtgctgc acctgaagga cagcgtgctg 240  
 tatctgaaga accggaaaga aaagaacgcc gtgcaggaca agaactatag cgaagaggac 300  
 atcagcgagt acgacctgaa aaacaagaac agcttctccg tgctgaagaa gatcctgctg 360  
 aacgaggacg tgaactctga ggaactggaa atctttcgga aggacgtgga agccaagctg 420  
 aacaagatca acagcctgaa gtacagcttc gaagagaaca aggccaacta ccagaagatc 480  
 aacgagaaca acgtggaaaa agtgggaggc aagagcaagc ggaacatcat ctacgactac 540  
 tacagagaga gcgccaagcg caacgactac atcaacaacg tgcaggaagc cttcgacaag 600

ctgtataaga aagaggatat cgagaaactg tttttcctga tcgagaacag caagaagcac 660  
 gagaagtaca agatccgcga gtactatcac aagatcatcg gccggaagaa cgacaaagag 720  
 aacttcgcca agattatcta cgaagagatc cagaactgga acaacatcaa agagctgatt 780  
 gagaagatcc ccgacatgtc tgagctgaag aaaagccagg tgttctaca gtactacctg 840  
 gacaaagagg aactgaacga caagaatatt aagtacgcct tctgccactt cgtggaatc 900  
 gagatgtccc agctgctgaa aaactacgtg tacaagcggc tgagcaacat cagcaacgat 960  
 aagatcaagc ggatcttcga gtaccagaat ctgaaaaagc tgatcgaaaa caaactgctg 1020

aacaagctgg acacctacgt gcggaactgc ggcaagtaca actactatct gcaagtgggc 1080  
 gagatcgcca cctccgactt tatcgcccgg aaccggcaga acgaggcctt cctgagaaac 1140  
 atcatcgcg tgtccagcgt ggcctacttc agcctgagga acatcctgga aaccgagaac 1200  
 gagaacgata tcaccggccg gatgcggggc aagacctgga agaacaacaa gggcgaagag 1260  
 aaatacgtgt ccggcgaggt ggacaagatc tacaatgaga acaagcagaa cgaagtgaaa 1320  
 gaaaatctga agatgttcta cagctacgac ttcaacatgg acaacaagaa cgagatcgag 1380  
 gacttcttcg ccaacatcga cgaggccatc agcagcatca gacacggcat cgtgcacttc 1440

aacctggaac tggaaggcaa ggacatcttc gccttcaaga atatcgcccc cagcgagatc 1500  
 tccaagaaga tgtttcagaa cgaaatcaac gaaaagaagc tgaagctgaa aatcttcaag 1560  
 cagctgaaca gcgccaactg gttcaactac tacgagaagg atgtgatcat caagtacctg 1620  
 aagaatacca agttcaactt cgtgaacaaa aacatcccct tcgtgcccg cttcacciaag 1680  
 ctgtacaaca agattgagga cctgcggaat accctgaagt ttttttgag cgtgccaag 1740  
 gacaaagaag agaaggacgc ccagatctac ctgctgaaga atatctacta cggcgagttc 1800

ctgaacaagt tcgtgaaaa ctccaagggtg ttctttaaga tcaccaatga agtgatcaag 1860

attaacaagc agcggaacca gaaaaccggc cactacaagt atcagaagtt cgagaacatc 1920

gagaaaaaccg tgcccgtgga atacctggcc atcatccaga gcagagagat gatcaacaac 1980

caggacaaag aggaaaagaa tacctacatc gactttattc agcagatfff cctgaagggc 2040

ttcatcgact acctgaacaa gaacaatctg aagtatatcg agagcaacaa caacaatgac 2100

aacaacgaca tcttctcaa gatcaagatc aaaaaggata acaaagagaa gtacgacaag 2160

atcctgaaga actatgagaa gcacaatcgg acaaagaaa tccctcacga gatcaatgag 2220

ttcgtgcgag agatcaagct ggggaagatt ctgaagtaca ccgagaatct gaacatgttt 2280

tacctgatcc tgaagctgct gaaccacaaa gagctgacca acctgaaggg cagcctggaa 2340

aagtaccagt ccgccaacaa agaagaaacc ttcagcgacg agctggaact gatcaacctg 2400

ctgaacctgg acaacaacag agtgaccgag gacttcgagc tggaagccaa cgagatcggc 2460

aagttcctgg acttcaacga aaacaaaatc aaggaccgga aagagctgaa aaagttcgac 2520

accaacaaga tctatttca cggcgagaac atcatcaagc accgggcctt ctacaatate 2580

aagaaatagc gcatgctgaa tctgctggaa aagatcgccg ataaggccaa gtataagatc 2640

agcctgaaag aactgaaaga gtacagcaac aagaagaatg agattgaaaa gaactacacc 2700

atgcagcaga acctgcaccg gaagtacgcc agaccaaga aggacgaaaa gttcaacgac 2760

gaggactaca aagagtatga gaaggccatc ggcaacatcc agaagtacac ccacctgaag 2820

aacaagtggt aattcaatga gctgaacctg ctgcagggcc tgctgctgaa gatcctgcac 2880

cggctcgtgg gctacaccag catctgggag cgggacctga gattccgct gaaggcgag 2940

tttcccgaga accactacat cgaggaaatt ttcaatttcg acaactcaa gaatgtgaag 3000

tacaaaagcg gccagatcgt ggaaaagtat atcaacttct acaaagaact gtacaaggac 3060

aatgtggaaa agcggagcat ctactccgac aagaaagtga agaaactgaa gcaggaaaaa 3120

aaggacctgt acatccggaa ctacattgcc cacttcaact acatcccca cgccgagatt 3180

agcctgctgg aagtgtgga aaacctgcgg aagctgctgt cctacgaccg gaagctgaag 3240

aacgcatca tgaagtccat cgtggacatt ctgaaagaat acggcttctg ggccaccttc 3300

aagatcggcg ctgacaagaa gatcgaatc cagacctgg aatcagagaa gatcgtgcac 3360

ctgaagaatc tgaagaaaa gaaactgatg accgaccgga acagcgagga actgtgcgaa 3420

ctcgtgaaag tcatgttca gtacaaggcc ctggaataa 3459

<210> 24

<211> 1152

<212> PRT

<213> Artificial Sequence

<220><223> C2C2 amino acids sequences, *Leptotrichia wadei*'s Cas13a

<400> 24

Met Lys Val Thr Lys Val Asp Gly Ile Ser His Lys Lys Tyr Ile Glu

1 5 10 15

Glu Gly Lys Leu Val Lys Ser Thr Ser Glu Glu Asn Arg Thr Ser Glu

20 25 30

Arg Leu Ser Glu Leu Leu Ser Ile Arg Leu Asp Ile Tyr Ile Lys Asn

35 40 45

Pro Asp Asn Ala Ser Glu Glu Glu Asn Arg Ile Arg Arg Glu Asn Leu

50 55 60

Lys Lys Phe Phe Ser Asn Lys Val Leu His Leu Lys Asp Ser Val Leu

65 70 75 80

Tyr Leu Lys Asn Arg Lys Glu Lys Asn Ala Val Gln Asp Lys Asn Tyr

85 90 95

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

100 105 110

Ser Val Leu Lys Lys Ile Leu Leu Asn Glu Asp Val Asn Ser Glu Glu

115 120 125

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

130 135 140

Ser Leu Lys Tyr Ser Phe Glu Glu Asn Lys Ala Asn Tyr Gln Lys Ile

145 150 155 160

Asn Glu Asn Asn Val Glu Lys Val Gly Gly Lys Ser Lys Arg Asn Ile

165 170 175

Ile Tyr Asp Tyr Tyr Arg Glu Ser Ala Lys Arg Asn Asp Tyr Ile Asn

180 185 190

Asn Val Gln Glu Ala Phe Asp Lys Leu Tyr Lys Lys Glu Asp Ile Glu

195 200 205

Lys Leu Phe Phe Leu Ile Glu Asn Ser Lys Lys His Glu Lys Tyr Lys



465                      470                      475                      480  
 Asn Leu Glu Leu Glu Gly Lys Asp Ile Phe Ala Phe Lys Asn Ile Ala  
  
                                  485                      490                      495  
 Pro Ser Glu Ile Ser Lys Lys Met Phe Gln Asn Glu Ile Asn Glu Lys  
                                  500                      505                      510  
 Lys Leu Lys Leu Lys Ile Phe Lys Gln Leu Asn Ser Ala Asn Val Phe  
                                  515                      520                      525  
 Asn Tyr Tyr Glu Lys Asp Val Ile Ile Lys Tyr Leu Lys Asn Thr Lys  
                                  530                      535                      540  
 Phe Asn Phe Val Asn Lys Asn Ile Pro Phe Val Pro Ser Phe Thr Lys  
 545                      550                      555                      560  
  
 Leu Tyr Asn Lys Ile Glu Asp Leu Arg Asn Thr Leu Lys Phe Phe Trp  
                                  565                      570                      575  
 Ser Val Pro Lys Asp Lys Glu Glu Lys Asp Ala Gln Ile Tyr Leu Leu  
                                  580                      585                      590  
 Lys Asn Ile Tyr Tyr Gly Glu Phe Leu Asn Lys Phe Val Lys Asn Ser  
                                  595                      600                      605  
 Lys Val Phe Phe Lys Ile Thr Asn Glu Val Ile Lys Ile Asn Lys Gln  
                                  610                      615                      620  
 Arg Asn Gln Lys Thr Gly His Tyr Lys Tyr Gln Lys Phe Glu Asn Ile  
  
 625                      630                      635                      640  
 Glu Lys Thr Val Pro Val Glu Tyr Leu Ala Ile Ile Gln Ser Arg Glu  
                                  645                      650                      655  
 Met Ile Asn Asn Gln Asp Lys Glu Glu Lys Asn Thr Tyr Ile Asp Phe  
                                  660                      665                      670  
 Ile Gln Gln Ile Phe Leu Lys Gly Phe Ile Asp Tyr Leu Asn Lys Asn  
                                  675                      680                      685  
 Asn Leu Lys Tyr Ile Glu Ser Asn Asn Asn Asn Asp Asn Asn Asp Ile  
                                  690                      695                      700  
  
 Phe Ser Lys Ile Lys Ile Lys Lys Asp Asn Lys Glu Lys Tyr Asp Lys  
 705                      710                      715                      720

Ile Leu Lys Asn Tyr Glu Lys His Asn Arg Asn Lys Glu Ile Pro His  
 725 730 735  
 Glu Ile Asn Glu Phe Val Arg Glu Ile Lys Leu Gly Lys Ile Leu Lys  
 740 745 750  
 Tyr Thr Glu Asn Leu Asn Met Phe Tyr Leu Ile Leu Lys Leu Leu Asn  
 755 760 765  
 His Lys Glu Leu Thr Asn Leu Lys Gly Ser Leu Glu Lys Tyr Gln Ser  
 770 775 780  
 Ala Asn Lys Glu Glu Thr Phe Ser Asp Glu Leu Glu Leu Ile Asn Leu  
 785 790 795 800  
 Leu Asn Leu Asp Asn Asn Arg Val Thr Glu Asp Phe Glu Leu Glu Ala  
 805 810 815  
 Asn Glu Ile Gly Lys Phe Leu Asp Phe Asn Glu Asn Lys Ile Lys Asp  
 820 825 830  
 Arg Lys Glu Leu Lys Lys Phe Asp Thr Asn Lys Ile Tyr Phe Asp Gly  
 835 840 845  
 Glu Asn Ile Ile Lys His Arg Ala Phe Tyr Asn Ile Lys Lys Tyr Gly  
 850 855 860  
 Met Leu Asn Leu Leu Glu Lys Ile Ala Asp Lys Ala Lys Tyr Lys Ile  
 865 870 875 880  
 Ser Leu Lys Glu Leu Lys Glu Tyr Ser Asn Lys Lys Asn Glu Ile Glu  
 885 890 895  
 Lys Asn Tyr Thr Met Gln Gln Asn Leu His Arg Lys Tyr Ala Arg Pro  
 900 905 910  
 Lys Lys Asp Glu Lys Phe Asn Asp Glu Asp Tyr Lys Glu Tyr Glu Lys  
 915 920 925  
 Ala Ile Gly Asn Ile Gln Lys Tyr Thr His Leu Lys Asn Lys Val Glu  
 930 935 940  
 Phe Asn Glu Leu Asn Leu Leu Gln Gly Leu Leu Leu Lys Ile Leu His  
 945 950 955 960  
 Arg Leu Val Gly Tyr Thr Ser Ile Trp Glu Arg Asp Leu Arg Phe Arg  
 965 970 975

Leu Lys Gly Glu Phe Pro Glu Asn His Tyr Ile Glu Glu Ile Phe Asn  
 980 985 990

Phe Asp Asn Ser Lys Asn Val Lys Tyr Lys Ser Gly Gln Ile Val Glu  
 995 1000 1005

Lys Tyr Ile Asn Phe Tyr Lys Glu Leu Tyr Lys Asp Asn Val Glu Lys  
 1010 1015 1020

Arg Ser Ile Tyr Ser Asp Lys Lys Val Lys Lys Leu Lys Gln Glu Lys  
 1025 1030 1035 1040

Lys Asp Leu Tyr Ile Arg Asn Tyr Ile Ala His Phe Asn Tyr Ile Pro  
 1045 1050 1055

His Ala Glu Ile Ser Leu Leu Glu Val Leu Glu Asn Leu Arg Lys Leu  
 1060 1065 1070

Leu Ser Tyr Asp Arg Lys Leu Lys Asn Ala Ile Met Lys Ser Ile Val  
 1075 1080 1085

Asp Ile Leu Lys Glu Tyr Gly Phe Val Ala Thr Phe Lys Ile Gly Ala  
 1090 1095 1100

Asp Lys Lys Ile Glu Ile Gln Thr Leu Glu Ser Glu Lys Ile Val His  
 1105 1110 1115 1120

Leu Lys Asn Leu Lys Lys Lys Lys Leu Met Thr Asp Arg Asn Ser Glu  
 1125 1130 1135

Glu Leu Cys Glu Leu Val Lys Val Met Phe Glu Tyr Lys Ala Leu Glu  
 1140 1145 1150

<210> 25

<211> 43

<212> DNA

<213> Artificial Sequence

<220><223> T7 promoter and DNA sequence for producing crRNA

<400> 25

aattctaata cgactcacta taggaatttc tactgttgta gat

43

<210> 26

<211> 32

<212> DNA

<213> Artificial Sequence

<220><223> DNA sequence of BRCA1 Exon 7 in normal cell

<400> 26

caaagtatgg gctacagaaa ccgtgccaaa ag 32

<210> 27

<211> 32

<212> DNA

<213> Artificial Sequence

<220><223> DNA sequence of BRCA1 Exon 7 in cancer cell

<400> 27

caaagtatgg gcttcagaaa ccgtgccaaa ag 32

<210> 28

<211> 32

<212> DNA

<213> Artificial Sequence

<220><223> DNA sequence of BRCA1 Exon 10 in normal cell

<400> 28

tgggaaaacc tatcggaaga aggcaagcct cc 32

<210> 29

<211> 32

<212> DNA

<213> Artificial Sequence

<220><223> DNA sequence of BRCA1 Exon 10 in cancerl cell

<400> 29

tgggaaaacc tatcggtaga aggcaagcct cc 32

<210> 30

<211> 30

<212> DNA

<213> Artificial Sequence

<220><223> DNA sequence of BRCA1 Exon 11 in normal cell

<400> 30

ggggccaaga aattagagtc ctcagaagag 30

<210> 31

<211>	31	
<212>	DNA	
<213>	Artificial Sequence	
<220><223>	DNA sequence of BRCA1 Exon 11 in cancer cell	
<400>	31	
	ggggccaaga aaattagagt cctcagaaga g	31
<210>	32	
<211>	31	
<212>	DNA	
<213>	Artificial Sequence	
<220><223>	DNA sequence of BRCA1 Exon 11 in normal cell	
<400>	32	
	atgatgaaga aagaggaacg ggcttgaag a	31
<210>	33	
<211>	29	
<212>	DNA	
<213>	Artificial Sequence	
<220><223>	DNA sequence of BRCA1 Exon 11 in cancer cell	
<400>	33	
	atgatgaaga aaggaacggg cttggaaga	29
<210>	34	
<211>	33	
<212>	DNA	
<213>	Artificial Sequence	
<220><223>	DNA sequence of BRCA1 Exon 11 in normal cell	
<400>	34	
	catctcaggt ttgttctgag acacctgatg acc	33
<210>	35	
<211>	32	
<212>	DNA	
<213>	Artificial Sequence	
<220><223>	DNA sequence of BRCA1 Exon 11 in cancer cell	
<400>	35	

catctcaggt ttgttctaga cacctgatga cc 32

<210> 36

<211> 33

<212> DNA

<213> Artificial Sequence

<220><223> DNA sequence of BRCA1 Exon 15 in normal cell

<400> 36

atatacagga tatgcgaatt aagaagaaac aaa 33

<210> 37

<211> 33

<212> DNA

<213> Artificial Sequence

<220><223> DNA sequence of BRCA1 Exon 15 in cancer cell

<400> 37

atatacagga tatgtgaatt aagaagaaac aaa 33

<210> 38

<211> 32

<212> DNA

<213> Artificial Sequence

<220><223> DNA sequence of TP53 in normal cell

<400>

38

taggaggccg agctctgttg cttcgaactc ca 32

<210> 39

<211> 31

<212> DNA

<213> Artificial Sequence

<220><223> DNA sequence of TP53 in cancer cell

<400> 39

taggaggccg agctctttgc ttcgaactcc a 31

<210> 40

<211> 31

<212> DNA

<213> Artificial Sequence

<220><223> DNA sequence of MSH2 in normal cell  
 <400> 40  
 tgaggaggtt tcgacatggc ggtgcagccg a 31

<210> 41  
 <211> 31  
 <212> DNA  
 <213> Artificial Sequence

<220><223> DNA sequence of MSH2 in cancer cell  
 <400> 41  
 tgaggaggtt tcgacctggc ggtgcagccg a 31

<210> 42  
 <211> 32  
 <212> DNA  
 <213> Artificial Sequence

<220><223> DNA sequence of EGFR in normal cell  
 <400> 42  
 aaaaagatca aagtgctggg ctccggtgcg tt 32

<210> 43  
 <211> 32  
 <212> DNA  
 <213> Artificial Sequence

<  
 <220><223> DNA sequence of EGFR in cancer cell  
 <400> 43  
 aaaaagatca aagtgctgag ctccggtgcg tt 32

<210> 44  
 <211> 33  
 <212> DNA  
 <213> Artificial Sequence

<220><223> DNA sequence of FGFR3 in normal cell  
 <400> 44  
 atcctctctc tgaatcact gagcaggaga aag 33

<210> 45  
 <211> 33

<212> DNA

<213> Artificial Sequence

<220><223> DNA sequence of FGFR3 in cancer cell

<400> 45

atcctctctc tgaatcact gcgcaggaga aag

33

<210> 46

<211> 1442

<212> PRT

<213> Artificial Sequence

<220><223> SpyCas9

<400> 46

Met Gly Ser Ser His His His His His His Ser Ser Gly Leu Val Pro

1 5 10 15

Arg Gly Ser His Met Ala Ser Met Thr Gly Gly Gln Gln Met Gly Arg

20 25 30

Gly Ser Glu Phe Glu Leu Arg Arg Gln Ala Cys Gly Arg Met Asp Lys

35 40 45

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

50 55 60

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

65 70 75 80

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

85 90 95

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

100 105 110

Ala Arg Arg Arg Tyr Thr Arg Arg Lys Asn Arg Ile Cys Tyr Leu Gln

115 120 125

Glu Ile Phe Ser Asn Glu Met Ala Lys Val Asp Asp Ser Phe Phe His

130 135 140

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

145 150 155 160

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



Phe Tyr Lys Phe Ile Lys Pro Ile Leu Glu Lys Met Asp Gly Thr Glu  
 420 425 430  
 Glu Leu Leu Val Lys Leu Asn Arg Glu Asp Leu Leu Arg Lys Gln Arg  
 435 440 445  
 Thr Phe Asp Asn Gly Ser Ile Pro His Gln Ile His Leu Gly Glu Leu  
 450 455 460  
 His Ala Ile Leu Arg Arg Gln Glu Asp Phe Tyr Pro Phe Leu Lys Asp  
 465 470 475 480  
 Asn Arg Glu Lys Ile Glu Lys Ile Leu Thr Phe Arg Ile Pro Tyr Tyr  
 485 490 495  
 Val Gly Pro Leu Ala Arg Gly Asn Ser Arg Phe Ala Trp Met Thr Arg  
 500 505 510  
 Lys Ser Glu Glu Thr Ile Thr Pro Trp Asn Phe Glu Glu Val Val Asp  
 515 520 525  
 Lys Gly Ala Ser Ala Gln Ser Phe Ile Glu Arg Met Thr Asn Phe Asp  
 530 535 540  
 Lys Asn Leu Pro Asn Glu Lys Val Leu Pro Lys His Ser Leu Leu Tyr  
 545 550 555 560  
 Glu Tyr Phe Thr Val Tyr Asn Glu Leu Thr Lys Val Lys Tyr Val Thr  
 565 570 575  
 Glu Gly Met Arg Lys Pro Ala Phe Leu Ser Gly Glu Gln Lys Lys Ala  
 580 585 590  
 Ile Val Asp Leu Leu Phe Lys Thr Asn Arg Lys Val Thr Val Lys Gln  
 595 600 605  
 Leu Lys Glu Asp Tyr Phe Lys Lys Ile Glu Cys Phe Asp Ser Val Glu  
 610 615 620  
 Ile Ser Gly Val Glu Asp Arg Phe Asn Ala Ser Leu Gly Thr Tyr His  
 625 630 635 640  
 Asp Leu Leu Lys Ile Ile Lys Asp Lys Asp Phe Leu Asp Asn Glu Glu  
 645 650 655  
 Asn Glu Asp Ile Leu Glu Asp Ile Val Leu Thr Leu Thr Leu Phe Glu  
 660 665 670

Asp Arg Glu Met Ile Glu Glu Arg Leu Lys Thr Tyr Ala His Leu Phe  
 675 680 685  
 Asp Asp Lys Val Met Lys Gln Leu Lys Arg Arg Arg Tyr Thr Gly Trp  
 690 695 700  
  
 Gly Arg Leu Ser Arg Lys Leu Ile Asn Gly Ile Arg Asp Lys Gln Ser  
 705 710 715 720  
 Gly Lys Thr Ile Leu Asp Phe Leu Lys Ser Asp Gly Phe Ala Asn Arg  
 725 730 735  
 Asn Phe Met Gln Leu Ile His Asp Asp Ser Leu Thr Phe Lys Glu Asp  
 740 745 750  
 Ile Gln Lys Ala Gln Val Ser Gly Gln Gly Asp Ser Leu His Glu His  
 755 760 765  
 Ile Ala Asn Leu Ala Gly Ser Pro Ala Ile Lys Lys Gly Ile Leu Gln  
  
 770 775 780  
 Thr Val Lys Val Val Asp Glu Leu Val Lys Val Met Gly Arg His Lys  
 785 790 795 800  
 Pro Glu Asn Ile Val Ile Glu Met Ala Arg Glu Asn Gln Thr Thr Gln  
 805 810 815  
 Lys Gly Gln Lys Asn Ser Arg Glu Arg Met Lys Arg Ile Glu Glu Gly  
 820 825 830  
 Ile Lys Glu Leu Gly Ser Gln Ile Leu Lys Glu His Pro Val Glu Asn  
 835 840 845  
  
 Thr Gln Leu Gln Asn Glu Lys Leu Tyr Leu Tyr Tyr Leu Gln Asn Gly  
 850 855 860  
 Arg Asp Met Tyr Val Asp Gln Glu Leu Asp Ile Asn Arg Leu Ser Asp  
 865 870 875 880  
 Tyr Asp Val Asp His Ile Val Pro Gln Ser Phe Leu Lys Asp Asp Ser  
 885 890 895  
 Ile Asp Asn Lys Val Leu Thr Arg Ser Asp Lys Asn Arg Gly Lys Ser  
 900 905 910  
 Asp Asn Val Pro Ser Glu Glu Val Val Lys Lys Met Lys Asn Tyr Trp



1170                      1175                      1180  
 Ala Tyr Ser Val Leu Val Val Ala Lys Val Glu Lys Gly Lys Ser Lys  
 1185                      1190                      1195                      1200  
 Lys Leu Lys Ser Val Lys Glu Leu Leu Gly Ile Thr Ile Met Glu Arg  
  
                          1205                      1210                      1215  
 Ser Ser Phe Glu Lys Asn Pro Ile Asp Phe Leu Glu Ala Lys Gly Tyr  
                          1220                      1225                      1230  
 Lys Glu Val Lys Lys Asp Leu Ile Ile Lys Leu Pro Lys Tyr Ser Leu  
                          1235                      1240                      1245  
 Phe Glu Leu Glu Asn Gly Arg Lys Arg Met Leu Ala Ser Ala Gly Glu  
                          1250                      1255                      1260  
 Leu Gln Lys Gly Asn Glu Leu Ala Leu Pro Ser Lys Tyr Val Asn Phe  
 1265                      1270                      1275                      1280  
  
 Leu Tyr Leu Ala Ser His Tyr Glu Lys Leu Lys Gly Ser Pro Glu Asp  
                          1285                      1290                      1295  
 Asn Glu Gln Lys Gln Leu Phe Val Glu Gln His Lys His Tyr Leu Asp  
                          1300                      1305                      1310  
 Glu Ile Ile Glu Gln Ile Ser Glu Phe Ser Lys Arg Val Ile Leu Ala  
                          1315                      1320                      1325  
 Asp Ala Asn Leu Asp Lys Val Leu Ser Ala Tyr Asn Lys His Arg Asp  
                          1330                      1335                      1340  
 Lys Pro Ile Arg Glu Gln Ala Glu Asn Ile Ile His Leu Phe Thr Leu  
  
 1345                      1350                      1355                      1360  
 Thr Asn Leu Gly Ala Pro Ala Ala Phe Lys Tyr Phe Asp Thr Thr Ile  
                          1365                      1370                      1375  
 Asp Arg Lys Arg Tyr Thr Ser Thr Lys Glu Val Leu Asp Ala Thr Leu  
                          1380                      1385                      1390  
 Ile His Gln Ser Ile Thr Gly Leu Tyr Glu Thr Arg Ile Asp Leu Ser  
                          1395                      1400                      1405  
 Gln Leu Gly Gly Asp Ala Ala Ala Leu Asp Leu Glu Lys Arg Pro Ala  
 1410                      1415                      1420

Ala Thr Lys Lys Ala Gly Gln Ala Lys Lys Lys Glu His His His His  
 1425                      1430                      1435                      1440  
 His His

<210> 47  
 <211> 20  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> HCC827 EGFR Target Sequence  
 <400> 47  
 cggagatgtc ttgatagcga 20  
 <210> 48  
 <211> 20  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> HCC827 VSTM2A Target Sequence  
 <400> 48  
 agcttcctag caagtaacag 20  
 <210> 49  
 <211> 20  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> HCC827 KIF5A Target Sequence  
 <400> 49  
 gcgcatcttc cctttgttat 20  
 <210> 50  
 <211> 20  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> H1563 IRX1 Target Sequence  
 <400> 50  
 gtccggaaga ggaactagaa 20  
 <210> 51  
 <211> 20

<212> DNA  
 <213> Artificial Sequence  
 <220><223> H1563 ADAMTS16 Target Sequence  
  
 <400> 51  
 ctccgtgccg ctggtttatt 20  
 <210> 52  
 <211> 20  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> H1299 GNPDA2 Target Sequence  
 <400> 52  
 cagaagctct gcattcatcc 20  
 <210> 53  
 <211> 20  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> H1299 KCNE2 Target Sequence  
 <400> 53  
 ggtgatgtga gttctagtcc 20  
  
 <210> 54  
 <211> 20  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> H1299 SLC15A5 Target Sequence  
 <400> 54  
 ggaccgattg tgagaaatgc 20  
 <210> 55  
 <211> 20  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> H1299 SMIM11 Target Sequence  
 <400> 55  
 gtgcccagtg tgatgatatt 20

<210> 56  
 <211> 20  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> A549 DACH2 Target Sequence

<400> 56  
 gcatggcttt tggctgttcg 20  
 <210> 57  
 <211> 20  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> A549 HERC2P2 Target Sequence

<400> 57  
 gctgtgattt caacaggacg 20  
 <210> 58  
 <211> 20  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> A549 CD68 Target Sequence

<400> 58  
 agaccattgg agactacag 20  
 <210> 59  
 <211> 20  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> A549 SHBG Target Sequence

<400> 59  
 atagtactag gctgcctcac 20  
 <210> 60  
 <211> 20  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> SKBR3 ERBB2 Target Sequence (chr17:37,863,321(+))

<400> 60  
catgctccgc caccttacc 20

<210> 61  
<211> 20  
<212> DNA  
<213> Artificial Sequence

<220><223> SKBR3 ERBB2 Target Sequence (chr17:37,864,786(+))

<400> 61  
ctgcagcttc gaagcctcac 20

<210> 62  
<211> 20  
<212> DNA  
<213> Artificial Sequence

<220><223> SKBR3 ERBB2 Target Sequence (chr17:37,863,519(+))

<400> 62  
cttgttgtgg tttctcaacc 20

<210> 63  
<211> 20  
<212> DNA  
<213> Artificial Sequence

<220><223> SKBR3 ERBB2 Target Sequence (chr17:37,864,935(+))

<400> 63  
ggaagacgcc ctcagaagat 20

<210> 64  
<211> 20  
<212> DNA  
<213> Artificial Sequence

<220><223> SKBR3 ERBB2 Target Sequence (chr17:37,846,501(+))

<400> 64  
gcctgtaatc ccagctactc 20

<210> 65  
<211> 20  
<212> DNA

<213> Artificial Sequence  
 <220><223> SKBR3 ERBB2 Target Sequence (chr17:37,880,586(+))  
 <400> 65  
 caggctagag tgaatggtg 20  
  
 <210> 66  
 <211> 20  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> SKBR3 ERBB2 Target Sequence (locus N/A\_1)  
 <400> 66  
 cttccttgta ccaacacgta 20  
 <210> 67  
 <211> 20  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> SKBR3 ERBB2 Target Sequence (locus N/A\_2)  
 <400> 67  
 caggtgtgta ccaacacgta 20  
 <210> 68  
 <211> 20  
 <212> DNA  
 <213> Artificial Sequence  
  
 <220><223> SKBR3 KRT16 Target Sequence (chr17:39,766,652(-))  
 <400> 68  
 ccaggagtac cagctgcat 20  
 <210> 69  
 <211> 20  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> SKBR3 KRT16 Target Sequence (chr17:39,766,661(+))  
 <400> 69  
 tcttcacatc atgcagcagc 20  
 <210> 70

<211> 20  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> HT29 LINC00536 Target Sequence (chr8:117,336,549(+))

<400> 70  
 agtggccagg attgattcag 20

<210> 71  
 <211> 20  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> HT29 TRPS1 Target Sequence (chr8:116,630,763(-))

<400> 71  
 gacagcagaa tgaccttgg 20

<210> 72  
 <211> 20  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> HT29 CDK8 Target Sequence

<400> 72  
 cctgtcttgt tcccagtc 20

<210> 73  
 <211> 20  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> HT29 TRAPPC9 Target Sequence

<400> 73  
 gtaagcttac ctagagacc 20

<210> 74  
 <211> 20  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> HT29 HERC2P2 Target Sequence

<400> 74

gctgtgattt caacaggacg	20
<210> 75	
<211> 20	
<212> DNA	
<213> Artificial Sequence	
<220><223> HeLa PRDM9 Target Sequence	
<400> 75	
ggaccctatc tgaatgtgca	20
<210> 76	
<211> 20	
<212> DNA	
<213> Artificial Sequence	
<220><223> HeLa CDKN2B Target Sequence	
<400> 76	
gtgcattcca cgcgtaaaac	20
<210> 77	
<211> 20	
<212> DNA	
<213> Artificial Sequence	
<220><223> HeLa HPV Target Sequence (locus: N/A_1)	
<400> 77	
tgcttattgc caccacctgc	20
<210> 78	
<211> 20	
<212> DNA	
<213> Artificial Sequence	
<220><223> HeLa HPV Target Sequence (locus: N/A_2)	
<400> 78	
cctgcaggaa ccctaaaata	20
<210> 79	
<211> 20	
<212> DNA	
<213> Artificial Sequence	

<220><223> HeLa HPV Target Sequence (locus: N/A\_3)  
 <400> 79  
 ccatatTTTA gggTtCctgc 20  
 <210> 80  
 <211> 20  
 <212> DNA  
 <213> Artificial Sequence

<220><223> HeLa HPV Target Sequence (locus: N/A\_4)  
 <400> 80  
 tgcaggtggT ggcaataagc 20  
 <210> 81  
 <211> 20  
 <212> DNA  
 <213> Artificial Sequence

<220><223> LINE2(mt2) Target Sequence  
 <400> 81  
 aatctcccc acccttaaga 20  
 <210> 82  
 <211> 20  
 <212> DNA  
 <213> Artificial Sequence

<220><223> CCR5 Target Sequence  
 <400> 82  
 tgacatcaat tattatacat 20  
 <210> 83  
 <211> 20  
 <212> DNA  
 <213> Artificial Sequence

<220><223> HPRT1 Target Sequence  
 <400> 83  
 gcatttctca gtctaaaca 20  
 <210> 84  
 <211> 20

<212> DNA  
 <213> Artificial Sequence  
 <220><223> NT1 Target Sequence  
 <400> 84  
 ggtaaccgt gcggtcgtac 20  
 <210> 85  
 <211> 10  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> NT2 Target Sequence  
  
 <400> 85  
 ggtaaccgt 10  
 <210> 86  
 <211> 5  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> NT3 Target Sequence  
 <400> 86  
 gggt a 5  
 <210> 87  
 <211> 1951  
 <212> PRT  
 <213> Artificial Sequence  
 <220><223> Cas9-RecJ  
 <400> 87  
 Asp Lys Lys Tyr Ser Ile Gly Leu Asp Ile Gly Thr Asn Ser Val Gly  
 1 5 10 15  
  
 Trp Ala Val Ile Thr Asp Glu Tyr Lys Val Pro Ser Lys Lys Phe Lys  
 20 25 30  
 Val Leu Gly Asn Thr Asp Arg His Ser Ile Lys Lys Asn Leu Ile Gly  
 35 40 45  
 Ala Leu Leu Phe Asp Ser Gly Glu Thr Ala Glu Ala Thr Arg Leu Lys  
 50 55 60







Asn Gly Arg Asp Met Tyr Val Asp Gln Glu Leu Asp Ile Asn Arg Leu  
 820 825 830  
 Ser Asp Tyr Asp Val Asp His Ile Val Pro Gln Ser Phe Leu Lys Asp  
 835 840 845  
 Asp Ser Ile Asp Asn Lys Val Leu Thr Arg Ser Asp Lys Asn Arg Gly  
 850 855 860  
 Lys Ser Asp Asn Val Pro Ser Glu Glu Val Val Lys Lys Met Lys Asn  
 865 870 875 880  
  
 Tyr Trp Arg Gln Leu Leu Asn Ala Lys Leu Ile Thr Gln Arg Lys Phe  
 885 890 895  
 Asp Asn Leu Thr Lys Ala Glu Arg Gly Gly Leu Ser Glu Leu Asp Lys  
 900 905 910  
 Ala Gly Phe Ile Lys Arg Gln Leu Val Glu Thr Arg Gln Ile Thr Lys  
 915 920 925  
 His Val Ala Gln Ile Leu Asp Ser Arg Met Asn Thr Lys Tyr Asp Glu  
 930 935 940  
 Asn Asp Lys Leu Ile Arg Glu Val Lys Val Ile Thr Leu Lys Ser Lys  
  
 945 950 955 960  
 Leu Val Ser Asp Phe Arg Lys Asp Phe Gln Phe Tyr Lys Val Arg Glu  
 965 970 975  
 Ile Asn Asn Tyr His His Ala His Asp Ala Tyr Leu Asn Ala Val Val  
 980 985 990  
 Gly Thr Ala Leu Ile Lys Lys Tyr Pro Lys Leu Glu Ser Glu Phe Val  
 995 1000 1005  
 Tyr Gly Asp Tyr Lys Val Tyr Asp Val Arg Lys Met Ile Ala Lys Ser  
 1010 1015 1020  
  
 Glu Gln Glu Ile Gly Lys Ala Thr Ala Lys Tyr Phe Phe Tyr Ser Asn  
 1025 1030 1035 1040  
 Ile Met Asn Phe Phe Lys Thr Glu Ile Thr Leu Ala Asn Gly Glu Ile  
 1045 1050 1055  
 Arg Lys Arg Pro Leu Ile Glu Thr Asn Gly Glu Thr Gly Glu Ile Val  
 1060 1065 1070

Trp Asp Lys Gly Arg Asp Phe Ala Thr Val Arg Lys Val Leu Ser Met  
 1075 1080 1085  
 Pro Gln Val Asn Ile Val Lys Lys Thr Glu Val Gln Thr Gly Gly Phe  
 1090 1095 1100  
 Ser Lys Glu Ser Ile Leu Pro Lys Arg Asn Ser Asp Lys Leu Ile Ala  
 1105 1110 1115 1120  
 Arg Lys Lys Asp Trp Asp Pro Lys Lys Tyr Gly Gly Phe Asp Ser Pro  
 1125 1130 1135  
 Thr Val Ala Tyr Ser Val Leu Val Val Ala Lys Val Glu Lys Gly Lys  
 1140 1145 1150  
 Ser Lys Lys Leu Lys Ser Val Lys Glu Leu Leu Gly Ile Thr Ile Met  
 1155 1160 1165  
 Glu Arg Ser Ser Phe Glu Lys Asn Pro Ile Asp Phe Leu Glu Ala Lys  
 1170 1175 1180  
 Gly Tyr Lys Glu Val Lys Lys Asp Leu Ile Ile Lys Leu Pro Lys Tyr  
 1185 1190 1195 1200  
 Ser Leu Phe Glu Leu Glu Asn Gly Arg Lys Arg Met Leu Ala Ser Ala  
 1205 1210 1215  
 Gly Glu Leu Gln Lys Gly Asn Glu Leu Ala Leu Pro Ser Lys Tyr Val  
 1220 1225 1230  
 Asn Phe Leu Tyr Leu Ala Ser His Tyr Glu Lys Leu Lys Gly Ser Pro  
 1235 1240 1245  
 Glu Asp Asn Glu Gln Lys Gln Leu Phe Val Glu Gln His Lys His Tyr  
 1250 1255 1260  
 Leu Asp Glu Ile Ile Glu Gln Ile Ser Glu Phe Ser Lys Arg Val Ile  
 1265 1270 1275 1280  
 Leu Ala Asp Ala Asn Leu Asp Lys Val Leu Ser Ala Tyr Asn Lys His  
 1285 1290 1295  
 Arg Asp Lys Pro Ile Arg Glu Gln Ala Glu Asn Ile Ile His Leu Phe  
 1300 1305 1310  
 Thr Leu Thr Asn Leu Gly Ala Pro Ala Ala Phe Lys Tyr Phe Asp Thr

1315                      1320                      1325  
 Thr Ile Asp Arg Lys Arg Tyr Thr Ser Thr Lys Glu Val Leu Asp Ala  
 1330                      1335                      1340  
 Thr Leu Ile His Gln Ser Ile Thr Gly Leu Tyr Glu Thr Arg Ile Asp  
 1345                      1350                      1355                      1360  
 Leu Ser Gln Leu Gly Gly Asp Ala Ala Ala Leu Asp Leu Gln Val Lys  
                                  1365                      1370                      1375  
 Gln Gln Ile Gln Leu Arg Arg Arg Glu Val Asp Glu Thr Ala Asp Leu  
  
                                  1380                      1385                      1390  
 Pro Ala Glu Leu Pro Pro Leu Leu Arg Arg Leu Tyr Ala Ser Arg Gly  
                                  1395                      1400                      1405  
 Val Arg Ser Ala Gln Glu Leu Glu Arg Ser Val Lys Gly Met Leu Pro  
                                  1410                      1415                      1420  
 Trp Gln Gln Leu Ser Gly Val Glu Lys Ala Val Glu Ile Leu Tyr Asn  
 1425                      1430                      1435                      1440  
 Ala Phe Arg Glu Gly Thr Arg Ile Ile Val Val Gly Asp Phe Asp Ala  
                                  1445                      1450                      1455  
  
 Asp Gly Ala Thr Ser Thr Ala Leu Ser Val Leu Ala Met Arg Ser Leu  
                                  1460                      1465                      1470  
 Gly Cys Ser Asn Ile Asp Tyr Leu Val Pro Asn Arg Phe Glu Asp Gly  
                                  1475                      1480                      1485  
 Tyr Gly Leu Ser Pro Glu Val Val Asp Gln Ala His Ala Arg Gly Ala  
                                  1490                      1495                      1500  
 Gln Leu Ile Val Thr Val Asp Asn Gly Ile Ser Ser His Ala Gly Val  
 1505                      1510                      1515                      1520  
 Glu His Ala Arg Ser Leu Gly Ile Pro Val Ile Val Thr Asp His His  
  
                                  1525                      1530                      1535  
 Leu Pro Gly Glu Thr Leu Pro Ala Ala Glu Ala Ile Ile Asn Pro Asn  
                                  1540                      1545                      1550  
 Leu Arg Asp Cys Asn Phe Pro Ser Lys Ser Leu Ala Gly Val Gly Val  
                                  1555                      1560                      1565  
 Ala Phe Tyr Leu Met Leu Ala Leu Arg Thr Phe Leu Arg Asp Gln Gly

1570                      1575                      1580  
 Trp Phe Asp Glu Arg Gly Ile Ala Ile Pro Asn Leu Ala Glu Leu Leu  
 1585                      1590                      1595                      1600  
  
 Asp Leu Val Ala Leu Gly Thr Val Ala Asp Val Val Pro Leu Asp Ala  
                                  1605                      1610                      1615  
 Asn Asn Arg Ile Leu Thr Trp Gln Gly Met Ser Arg Ile Arg Ala Gly  
                                  1620                      1625                      1630  
 Lys Cys Arg Pro Gly Ile Lys Ala Leu Leu Glu Val Ala Asn Arg Asp  
                                  1635                      1640                      1645  
 Ala Gln Lys Leu Ala Ala Ser Asp Leu Gly Phe Ala Leu Gly Pro Arg  
                                  1650                      1655                      1660  
 Leu Asn Ala Ala Gly Arg Leu Asp Asp Met Ser Val Gly Val Ala Leu  
  
 1665                      1670                      1675                      1680  
 Leu Leu Cys Asp Asn Ile Gly Glu Ala Arg Val Leu Ala Asn Glu Leu  
                                  1685                      1690                      1695  
 Asp Ala Leu Asn Gln Thr Arg Lys Glu Ile Glu Gln Gly Met Gln Val  
                                  1700                      1705                      1710  
 Glu Ala Leu Thr Leu Cys Glu Lys Leu Glu Arg Ser Arg Asp Thr Leu  
                                  1715                      1720                      1725  
 Pro Gly Gly Leu Ala Met Tyr His Pro Glu Trp His Gln Gly Val Val  
                                  1730                      1735                      1740  
  
 Gly Ile Leu Ala Ser Arg Ile Lys Glu Arg Phe His Arg Pro Val Ile  
 1745                      1750                      1755                      1760  
 Ala Phe Ala Pro Ala Gly Asp Gly Thr Leu Lys Gly Ser Gly Arg Ser  
                                  1765                      1770                      1775  
 Ile Gln Gly Leu His Met Arg Asp Ala Leu Glu Arg Leu Asp Thr Leu  
                                  1780                      1785                      1790  
 Tyr Pro Gly Met Ile Leu Lys Phe Gly Gly His Ala Met Ala Ala Gly  
                                  1795                      1800                      1805  
 Leu Ser Leu Glu Glu Asp Lys Phe Glu Leu Phe Gln Gln Arg Phe Gly  
  
 1810                      1815                      1820

Glu Leu Val Thr Glu Trp Leu Asp Pro Ser Leu Leu Gln Gly Glu Val  
 1825                      1830                      1835                      1840  
 Val Ser Asp Gly Pro Leu Ser Pro Ala Glu Met Thr Met Glu Val Ala  
                                  1845                      1850                      1855  
 Gln Leu Leu Arg Asp Ala Gly Pro Trp Gly Gln Met Phe Pro Glu Pro  
                                  1860                      1865                      1870  
 Leu Phe Asp Gly His Phe Arg Leu Leu Gln Gln Arg Leu Val Gly Glu  
                                  1875                      1880                      1885

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

<210> 88

<211

> 96

<212> DNA

<213> Artificial Sequence

<220><223> EGFR gRNA Sequence

<400> 88

cggagatgtc ttgatagcga gttttagagc tagaaatagc aagttaaaat aaggctagtc 60

cgttatcaac ttgaaaaagt ggcaccgagt cgggtgc 96

<210> 89

<211> 96

<212> DNA

<213> Artificial Sequence

<220><223> VSTM2A gRNA Sequence

<400> 89

agcttcctag caagtaacag gttttagagc tagaaatagc aagttaaaat aaggctagtc 60

cgttatcaac ttgaaaaagt ggcaccgagt cgggtgc 96

<210> 90  
 <211> 96  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> KIF5A gRNA Sequence  
 <400> 90  
 gcgcaccttc cctttgttat gttttagagc tagaaatagc aagttaaaat aaggctagtc 60  
 cgttatcaac ttgaaaaagt ggcaccgagt cggatgc 96  
 <210> 91  
 <211> 96  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> IRX1 gRNA Sequence  
 <400> 91  
 gtccggaaga ggaactagaa gttttagagc tagaaatagc aagttaaaat aaggctagtc 60  
 cgttatcaac ttgaaaaagt ggcaccgagt cggatgc 96  
 <210> 92  
 <211> 96  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> ADAMTS16 gRNA Sequence  
 <400> 92  
 ctccgtgccg ctggtttatt gttttagagc tagaaatagc aagttaaaat aaggctagtc 60  
 cgttatcaac ttgaaaaagt ggcaccgagt cggatgc 96  
 <210> 93  
 <211> 96  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> GNPDA2 gRNA Sequence  
 <400> 93  
 cagaagctct gcattcatcc gttttagagc tagaaatagc aagttaaaat aaggctagtc 60  
 cgttatcaac ttgaaaaagt ggcaccgagt cggatgc 96  
 <210> 94

<211>	96	
<212>	DNA	
<213>	Artificial Sequence	
<220><223>	KCNE2 Target Sequence	
<400>	94	
ggtgatgtga gttctagtc gttttagagc tagaaatagc aagttaaaat aaggctagtc		60
cgttatcaac ttgaaaaagt ggcaccgagt cgggtgc		96
<210>	95	
<211>	96	
<212>	DNA	
<213>	Artificial Sequence	
<220><223>	SLC15A5 gRNA Sequence	
<400>	95	
ggaccgattg tgagaaatgc gttttagagc tagaaatagc aagttaaaat aaggctagtc		60
cgttatcaac ttgaaaaagt ggcaccgagt cgggtgc		96
<210>	96	
<211>	96	
<212>	DNA	
<213>	Artificial Sequence	
<220><223>	SMIM11 gRNA Sequence	
<400>	96	
gtgccagtgt tgatgatatt gttttagagc tagaaatagc aagttaaaat aaggctagtc		60
cgttatcaac ttgaaaaagt ggcaccgagt cgggtgc		96
<210>	97	
<211>	96	
<212>	DNA	
<213>	Artificial Sequence	
<220><223>	DACH2 gRNA Sequence	
<400>	97	
gcatggcttt tggctgttcg gttttagagc tagaaatagc aagttaaaat aaggctagtc		60
cgttatcaac ttgaaaaagt ggcaccgagt cgggtgc		96
<210>	98	
<211>	96	

<212> DNA  
 <213> Artificial Sequence  
 <220><223> HERC2P2 gRNA Sequence  
 <400> 98  
 gctgtgattt caacaggacg gttttagagc tagaaatagc aagttaaaat aaggctagtc 60  
 cgttatcaac ttgaaaaagt ggcaccgagt cggtagc 96  
 <210> 99  
 <211> 96  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> CD68 gRNA Sequence  
 <400> 99  
 agaccattgg agactacacg gttttagagc tagaaatagc aagttaaaat aaggctagtc 60  
 cgttatcaac ttgaaaaagt ggcaccgagt cggtagc 96  
  
 <210> 100  
 <211> 96  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> SHBG gRNA Sequence  
 <400> 100  
 atagtactag gctgcctcac gttttagagc tagaaatagc aagttaaaat aaggctagtc 60  
 cgttatcaac ttgaaaaagt ggcaccgagt cggtagc 96  
 <210> 101  
 <211> 96  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> ERBB2 (1) gRNA Sequence  
 <400> 101  
 catgctccgc cacctctacc gttttagagc tagaaatagc aagttaaaat aaggctagtc 60  
 cgttatcaac ttgaaaaagt ggcaccgagt cggtagc 96  
  
 <210> 102  
 <211> 96  
 <212> DNA

<213> Artificial Sequence  
 <220><223> ERBB2 (2) gRNA Sequence  
 <400> 102  
 ctgcagcttc gaagcctcac gttttagagc tagaaatagc aagttaaaat aaggctagtc 60  
 cgttatcaac ttgaaaaagt ggcaccgagt cggatgc 96  
 <210> 103  
 <211> 96  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> ERBB2 (3) gRNA Sequence  
 <400> 103  
 ctgtgtgtgg tttctcaacc gttttagagc tagaaatagc aagttaaaat aaggctagtc 60  
 cgttatcaac ttgaaaaagt ggcaccgagt cggatgc 96  
  
 <210> 104  
 <211> 96  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> ERBB2 (4) gRNA Sequence  
 <400> 104  
 ggaagacgcc ctcaagaat gttttagagc tagaaatagc aagttaaaat aaggctagtc 60  
 cgttatcaac ttgaaaaagt ggcaccgagt cggatgc 96  
 <210> 105  
 <211> 96  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> ERBB2 (5) gRNA Sequence  
 <400> 105  
 gcctgtaatc ccagctactc gttttagagc tagaaatagc aagttaaaat aaggctagtc 60  
 cgttatcaac ttgaaaaagt ggcaccgagt cggatgc 96  
  
 <210> 106  
 <211> 96  
 <212> DNA  
 <213> Artificial Sequence

<220><223> ERBB2 (6) gRNA Sequence  
 <400> 106  
 caggctagag tgaatggtg gtttagagc tagaaatagc aagttaaaat aaggctagtc 60  
 cgttatcaac ttgaaaaagt ggcaccgagt cgggtgc 96  
 <210> 107  
 <211> 96  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> ERBB2 (7) gRNA Sequence  
 <400> 107  
 cttccttgta ccaacacgta gtttagagc tagaaatagc aagttaaaat aaggctagtc 60  
 cgttatcaac ttgaaaaagt ggcaccgagt cgggtgc 96  
 <210> 108  
 <211> 96  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> ERBB2 (8) gRNA Sequence  
 <400> 108  
 caggtgtgta ccaacacgta gtttagagc tagaaatagc aagttaaaat aaggctagtc 60  
 cgttatcaac ttgaaaaagt ggcaccgagt cgggtgc 96  
 <210> 109  
 <211> 96  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> KRT16 (1) gRNA Sequence  
 <400> 109  
 ccaggagtac cagctgcat gtttagagc tagaaatagc aagttaaaat aaggctagtc 60  
 cgttatcaac ttgaaaaagt ggcaccgagt cgggtgc 96  
 <210> 110  
 <211> 96  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> KRT16 (2) gRNA Sequence

<400> 110  
tcttcacatc atgcagcagc gttttagagc tagaaatagc aagttaaaat aaggctagtc 60  
cgttatcaac ttgaaaaagt ggcaccgagt cgggtgc 96

<210> 111  
<211> 96  
<212> DNA  
<213> Artificial Sequence  
<220><223> LINC00536 gRNA Sequence

<400> 111  
agtggccagg attgattcag gttttagagc tagaaatagc aagttaaaat aaggctagtc 60  
cgttatcaac ttgaaaaagt ggcaccgagt cgggtgc 96

<210> 112  
<211> 96  
<212> DNA  
<213> Artificial Sequence  
<220><223> TRPS1 gRNA Sequence

<400> 112  
gacagcagaa tgaccttggg gttttagagc tagaaatagc aagttaaaat aaggctagtc 60  
cgttatcaac ttgaaaaagt ggcaccgagt cgggtgc 96

<210> 113  
<211> 96  
<212> DNA  
<213> Artificial Sequence  
<220><223> CDK8 gRNA Sequence

<400> 113  
cctgtcttgt tcccagtcac gttttagagc tagaaatagc aagttaaaat aaggctagtc 60  
cgttatcaac ttgaaaaagt ggcaccgagt cgggtgc 96

<210> 114  
<211> 96  
<212> DNA  
<213> Artificial Sequence  
<220><223> TRAPPC9 gRNA Sequence

<400> 114

gtaagcttac cttagagacc gttttagagc tagaaatagc aagttaaaat aaggctagtc	60
cgttatcaac ttgaaaaagt ggcaccgagt cgggtgc	96
<210> 115	
<211> 96	
<212> DNA	
<213> Artificial Sequence	
<220><223> SIRPB1 gRNA Sequence	
<400> 115	
gacaccctaa cagggttatg gttttagagc tagaaatagc aagttaaaat aaggctagtc	60
cgttatcaac ttgaaaaagt ggcaccgagt cgggtgc	96
<210> 116	
<211> 96	
<212> DNA	
<213> Artificial Sequence	
<220><223> MRC1 gRNA Sequence	
<400> 116	
ctcactgcag ccttgacatc gttttagagc tagaaatagc aagttaaaat aaggctagtc	60
cgttatcaac ttgaaaaagt ggcaccgagt cgggtgc	96
<210> 117	
<211> 96	
<212> DNA	
<213> Artificial Sequence	
<220><223> ATP11A gRNA Sequence	
<400> 117	
cgccaagtg tgtgagtgc gttttagagc tagaaatagc aagttaaaat aaggctagtc	60
cgttatcaac ttgaaaaagt ggcaccgagt cgggtgc	96
<210> 118	
<211> 96	
<212> DNA	
<213> Artificial Sequence	
<220><223> POTE8 gRNA Sequence	
<400> 118	
gtggaaacct cagagatgac gttttagagc tagaaatagc aagttaaaat aaggctagtc	60

cgttatcaac ttgaaaaagt ggcaccgagt cgggtgc	96
<210> 119	
<211> 96	
<212> DNA	
<213> Artificial Sequence	
<220><223> PRDM9 gRNA Sequence	
<400> 119	
ggaccctatc tgaatgtgca gttttagagc tagaaatagc aagttaaaat aaggctagtc	60
cgttatcaac ttgaaaaagt ggcaccgagt cgggtgc	96
<210> 120	
<211> 96	
<212> DNA	
<213> Artificial Sequence	
<220><223> CDKN2B gRNA Sequence	
<400> 120	
gigcattcca cgcgtaaac gttttagagc tagaaatagc aagttaaaat aaggctagtc	60
cgttatcaac ttgaaaaagt ggcaccgagt cgggtgc	96
<210> 121	
<211> 96	
<212> DNA	
<213> Artificial Sequence	
<220><223> HPV (1) gRNA Sequence	
<400> 121	
tgcttattgc caccacctgc gttttagagc tagaaatagc aagttaaaat aaggctagtc	60
cgttatcaac ttgaaaaagt ggcaccgagt cgggtgc	96
<210> 122	
<211> 96	
<212> DNA	
<213> Artificial Sequence	
<220><223> HPV (2) gRNA Sequence	
<400> 122	
cctgcaggaa ccctaaaata gttttagagc tagaaatagc aagttaaaat aaggctagtc	60
cgttatcaac ttgaaaaagt ggcaccgagt cgggtgc	96

<210> 123  
 <211> 96  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> HPV (3) gRNA Sequence  
 <400> 123  
 ccatatttta gggttcctgc gttttagagc tagaaatagc aagttaaaat aaggctagtc 60  
 cgttatcaac ttgaaaaagt ggcaccgagt cgggtgc 96

<210> 124  
 <211> 96  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> HPV (4) gRNA Sequence  
 <400> 124  
 tgcagtggt ggcaataagc gttttagagc tagaaatagc aagttaaaat aaggctagtc 60  
 cgttatcaac ttgaaaaagt ggcaccgagt cgggtgc 96

<210> 125  
 <211> 96  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> LINE2 (mt2) gRNA Sequence  
 <400> 125  
 aatctcccc accettaaga gttttagagc tagaaatagc aagttaaaat aaggctagtc 60  
 cgttatcaac ttgaaaaagt ggcaccgagt cgggtgc 96

<210> 126  
 <211> 96  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> CCR5 gRNA Sequence  
 <400> 126  
 tgacatcaat tattatacat gttttagagc tagaaatagc aagttaaaat aaggctagtc 60  
 cgttatcaac ttgaaaaagt ggcaccgagt cgggtgc 96

<210> 127

<211>	96	
<212>	DNA	
<213>	Artificial Sequence	
<220><223>	HPRT1 gRNA Sequence	
<400>	127	
gcattttctca gtcctaaaca gtttttagagc tagaaatagc aagttaaaat aaggctagtc		60
cgttatcaac ttgaaaaagt ggcaccgagt cgggtgc		96
<210>	128	
<211>	96	
<212>	DNA	
<213>	Artificial Sequence	
<220><223>	NT (1) gRNA Sequence	
<400>	128	
gggtaaccgt gcggtcgtac gtttttagagc tagaaatagc aagttaaaat aaggctagtc		60
cgttatcaac ttgaaaaagt ggcaccgagt cgggtgc		96
<210>	129	
<211>	86	
<212>	DNA	
<213>	Artificial Sequence	
<220><223>	NT (2) gRNA Sequence	
<400>	129	
gggtaaccgt gtttttagagc tagaaatagc aagttaaaat aaggctagtc cgttatcaac		60
ttgaaaaagt ggcaccgagt cgggtgc		86
<210>	130	
<211>	81	
<212>	DNA	
<213>	Artificial Sequence	
<220><223>	NT (3) gRNA Sequence	
<400>	130	
gggtagtttt agagctagaa atagcaagtt aaataaggc tagtccgtta tcaactgaa		60
aaagtggcac cgagtcggtg c		81
<210>	131	
<211>	249	

<212> DNA  
 <213> Artificial Sequence  
 <220><223> Promoter Sequence  
 <400> 131  
 gagggcctat ttcccatgat tccttcatat ttgcatatac gatacaaggc tgtagagag 60  
 ataattggaa ttaatttgac tgtaaacaca aagatattag tacaaaatac gtgacgtaga 120  
 aagtaataat ttcttgggta gtttgagtt ttaaaattat gttttaaaat ggactatcat 180  
 atgcttaccg taacttgaaa gtatttcgat ttcttggcctt tatatatctt gtggaaagga 240  
 cgaaacacc 249  
 <210> 132  
 <211> 24  
 <212> RNA  
 <213> Artificial Sequence  
 <220><223> DHCR7 crRNA  
 <400> 132  
 cactggcgag cgtcatcttc ctac 24  
 <210> 133  
 <211> 1431  
 <212> DNA  
 <213> Artificial Sequence  
 <220>  
 ><223> DHCR7 dsDNA #1 (1431bp)  
 <400> 133  
 gtaaaacgac ggccagtga ttgtaatacg actcactata gggcgaattg ggcccacgt 60  
 cgcattctcc cggccgccat ggcggccgc ggaattcgat tgtccctttg gggactgaga 120  
 tgctccttc cagacctggt gcaaggcgtg actcctgctg acaatggggc ttctcagga 180  
 cagcactgcc ctcccacggg gttttgctcc taccctcagc ttgtccctgc agagctgggc 240  
 gtgccagca actgcatgca ggcattccgt gaaggtgtat caaacgtga tgtgacaggt 300  
 gcctcctgcc ttacagetga ggatgaatct gaacatgtca gattccaggg aatgcctgc 360  
 tgggtcccgg gaaccagat gtcaacctga gccaggatcc atgtcccaga caaatggaag 420  
 gactaccca gcaggagggc acgtcccca cctgctgtgt cccaaccca gggcaggggc 480  
 tgctgacctg gaaggtgacc cacaaggtat agagctgggc ggctttcctc gttataggtg 540  
 gactcttggc ccagatgtcc gagagccgag catgtccggt gacgatgtcc accacagggc 600

cagtcagggc gcagctgtac tggtcacaag ccatgatgaa gtagtagacg atgaaggggg 660  
 cgaacagcag taggaagatg acgctcgcca gtgaaaacca gtccacctcc ctgcgaggac 720  
 ggatgcaggc agtcacactg gggcccatct gccctgggcc ccaccaggac cctaagaggc 780

tctgtgtggg agaactgttg ctcaaacca ccagtacccc atgacagaag gcattagctc 840  
 ctagcacggg cctccttgc ggccaggga gccactcaac atgcctgctc tactgtagtt 900  
 gattaactgg tggcactatc tgtggccccc atggaaggcc cagagctcag aatatgagcg 960  
 gaggtaggtc tttcacaacc accaaggcca gtggtttccc cagtccagg tcggagagga 1020  
 tactcacct gcacgaagtc cccatggttc tatggcgaaa tgggagctgt gacaataggt 1080  
 catacccca gatggtggga aggccccagc tgcctttta gcctttatct tttttctaa 1140  
 gtgttgact gaccatgatt tctagtttct tttataatt ccttacttta taaagtggaa 1200

aactgtttac ctttcattag tcttgacaaa aaaaaaaaa tcttttttaa aaagtcctt 1260  
 cattcctttg ggtccataca attccaaagg ctggaagaa tcactagtga attcgcggcc 1320  
 gcctgcaggt cgaccatag ggagagctcc caacgcgttg gatgcatagc ttgagtattc 1380  
 tatagtgtca cctaataagc ttggcgtaat catggtcata gctgtttcct g 1431

<210> 134

<211> 544

<212> DNA

<213> Artificial Sequence

<220><223> DHCR7 dsDNA #2 (544bp)

<400> 134

gagccaagct ctccatctag tggacagga agctagcagc aaaccttccc ttcactaaa 60

aacttcattg ctggccaaa aagagagtta attcaatgta gacatctatg taggcaatta 120  
 aaaacctatt gatgtataaa acagtttga ttcattggagg gcaactaat acattctagg 180  
 actttataaa agatcacttt ttatttatgc acagggtgga acaagatgga ttatcaagtg 240  
 tcaagtccaa tctatgacat caattattat acatcggagc cctgccaaaa aatcaatgtg 300  
 aagcaaatcg cagcccgct cctgcctccg ctctactcac tgggtttcat ctttggtttt 360  
 gtgggcaaca tctgtgtcat cctcatctg ataaactgca aaaggctgaa gagcatgact 420  
 gacatctacc tgetcaacct ggccatctct gacctgtttt tctttcttac tgtecccttc 480

tgggctcact atgctgccgc ccagtgggac tttggaaata caatgtgtca actcttgaca 540  
 gggc 544

<210> 135

<211> 3825

<212> DNA

<213> Artificial Sequence

<220><223> Cas12a

<400> 135

```

atgggcaaaa accaaaatTT ccaagaatTT atcggagtga gccccctgca gaaaaccctc      60
cggaacgagc ttattccgac tgagaccaca aagaaaaata taaccagct ggacttgctg      120
actgaagatg agatccgcgc ccagaaccgg gaaaagctca aagagatgat ggacgattat      180
taccgcaatg ttattgacag tacccttcac gtcgggatcg ctgtggattg gtcttatctg      240

ttcagctgca tgcggaacca tttgcgcgaa aattccaagg agtcaaacg ggaactggag      300
cgcacacagg acagcattcg gagtcagata cacaacaagt ttgccgaac gcgagatttc      360
aaagacatgt ttggcgctc tatcattacc aagctccttc ctacttacct caaacaaaat      420
agcgagtatt ccgaacggta cgatgagtca atggaaattc tgaagttgta tggtaaattc      480
accacaagcc tgaccgacta ctttgagact cgcaagaaca tattcagtaa agaaaagatc      540
tctagcgctg taggctatcg gatttgggag gaaaatgccg agatctttct ccagaaccag      600
aatgcatacg atgcatttg taaaatagcc ggacttgacc tgcatgggtt ggataacgaa      660

atcacgcgtt atgttgacgg caagacactg aaagaggtct gctccgatga aggtttcgcc      720
aaggcaatta cccaagaggg catcgatcgg tacaatgaag ccatgggagc tgtgaaccag      780
tatatgaatc tcctttgtca gaaaaacaag gccctgaaac cggggcaatt taagatgaaa      840
cgcttgacac agcagatact gtgcaaagc actacctcat tcgatatccc gaagaaatTT      900
gagaatgaca agcaggtata cgatgcagtg aacagcttca cagaaattgt taccaaaaat      960
aacgacctca agcggcttct gaatatcact caaacgccca atgattatga catgaacaaa      1020
atttacgtcg tggtgatgc ctatagtatg atatctcagt ttatcagcaa gaaatggaat      1080

ttgattgagg aatgtctgct cgactactat tccgataacc ttccaggtaa gggcaatgca      1140
aaagagaaca aggtaaaaaa ggccgtgaaa gaagagacct accgctcagt tagccagctg      1200
aatgaagtca tcgagaagta ttacgtggaa aaaacaggac aaagtgtatg gaaggtggag      1260
tcttatatta gtccttggc tgaaatgata aaactggagc tctgcatga aatcgacaac      1320
gatgagaagc acaatcttat tgaggacgat gagaaaatct cagaaattaa ggagctgttg      1380
gacatgtaca tggatgtttt ccatataatc aaagtctttc gggtgaacga agtactgaat      1440
ttcgacgaga ctttttatag cgaaatggat gagatttacc aggacatgca ggaaatcgtg      1500

cccctctata accacgttcg caattacgtc actcaaaagc cgtataaaca ggagaagtac      1560

```

cggttttatt tccatacccc tacactggcc aacgggtgga gtaaatctaa ggaatacgat 1620  
 aataacgcaa ttatattggt gcgcgaggac aaatattacc tgggcatcct caatgccaag 1680  
 aaaaagccca gcaagaanaa tatggctggt aaggaggatt gttccgaaca cgctatgca 1740  
 aaaatgaact aciatcttct gccggcgcc aataagatgt tgccaaaagt atttctgtca 1800  
 aagaaaggaa tccaggacta ccatcccagc agttatattg tggaggggta caacgaaaag 1860  
 aaacacataa agggctctaa aaatttcgat atccggtttt gccgcgacct cattgattat 1920  
  
 ttcaaggagt gtatcaaaaa gcatccggac tggacaanaa ttaatttcga atttagcgct 1980  
 accgagactt acgaagatat ttccgttttc tatcgggagg tcgaaaagca aggttaccgc 2040  
 gtggagtgga cctatataaa ctacaggagc atccagaaac ttgaggaaga tggccagctg 2100  
 tttttgttcc aaatttaca taaggacttt gccgtaggaa gcacaggga acctaacctg 2160  
 cacacctct atcttaagaa tctgttcagt gaggaaaact tgcgggatat cgtgctgaaa 2220  
 ctcaatggcg aggcagaaat tttttccgc aagtctagcg ttcagaaacc cgtcatacat 2280  
 aagtgcggtt ccatccttgt gaaccggact tacgagatta ccgaatcagg cacaaccgc 2340  
  
 gtacagagca tcccgagag tgaatatatg gagctgtatc ggtattttaa ttctgaaaa 2400  
 caaattgagt tgagcgacga agccaagaaa tacctggata aggtgcagtg taacaaagct 2460  
 aagactgaca tagttaaaga ttatcgctac accatggaca agttctttat ccacctcca 2520  
 attacaatca atttcaaagt cgataaggga aacaatgtga acgccattgc acageaatat 2580  
 atagccgggc gaaagacct tcatgtaatc ggcatgtatc gcggtgagcg gaatctgatc 2640  
 tacgtgtccg ttattgacat gtatggccgc atattggaac agaagtcatt taacctggtc 2700  
 gagcaggtga gcagtcaagg aaccaaaccg tactatgatt acaaggaaaa actccagaat 2760  
  
 cgcgaggaag agcgggacaa ggctcgccag tcttgaaaa ctatcggga gattaagaa 2820  
 cttaggagg gctatctgag ctccgtaatc cacgaaatg cccaatggt ggttaaac 2880  
 aacgcaataa tcgccatgga ggatttgaat tatggtttca agcggggccg ctttaaagtc 2940  
 gaacggcagg tgtaccagaa gttcgagacc atgctgattt caaaactca ctatcttct 3000  
 gacaagagcc aagccgtaga tgaacccgga gggattctgc gcgctacca gatgacatat 3060  
 gtgccggaca atattaanaa cgttggtcgg cagtgcggca taatcttta cgtccctgca 3120  
 gcctatacca gtaagattga tcccactacc ggattcatca atgctttta acgcgactg 3180  
  
 gtatctaca acgatgcca ggagaatttc ttgatgaat ttgacagcat tcaatacgat 3240  
 atagaaaagg gctgttcaa attttcttc gactataaga actttgcaac ccataaactc 3300  
 actcttgcca agaccaaatg ggatgtgtac acaaatggca cccgattca gaacatgaag 3360  
 gttgagggtc actggctgtc aatggaagtc gagttgacta ccaaatgaa ggaactgctc 3420

gacgatagcc atattccgta tgaggaagc cagaatatcc ttgacgatct gcgcgagatg 3480  
 aaagacatta caaccatagt gaacggaatc ttggaaatct tctggctgac tgtacaactc 3540  
 cggaatagtc gcatcgataa cccagactac gatcggatta tatctcccggt gcttaataag 3600

aacggggagt ttttcgacag cgatgaatat aattcctaca tcgacgctca gaaagccccc 3660  
 ctgcctattg atgcagacgc caacggcgct tttgtatcg ccttgaaggg tatgtatacc 3720  
 gcaaatcaga ttaaagagaa ctgggttgaa ggcgagaagc tgcccgccga ttgcctcaaa 3780  
 atagaacacg cttcatggct tgccttcatg caaggagagc gcggg 3825

<210> 136  
 <211> 1275  
 <212> PRT  
 <213> Artificial Sequence  
 <220><223> Cas12a

<400> 136  
 Met Gly Lys Asn Gln Asn Phe Gln Glu Phe Ile Gly Val Ser Pro Leu  
 1 5 10 15

Gln Lys Thr Leu Arg Asn Glu Leu Ile Pro Thr Glu Thr Thr Lys Lys  
 20 25 30

Asn Ile Thr Gln Leu Asp Leu Leu Thr Glu Asp Glu Ile Arg Ala Gln  
 35 40 45

Asn Arg Glu Lys Leu Lys Glu Met Met Asp Asp Tyr Tyr Arg Asn Val  
 50 55 60

Ile Asp Ser Thr Leu His Val Gly Ile Ala Val Asp Trp Ser Tyr Leu  
 65 70 75 80

Phe Ser Cys Met Arg Asn His Leu Arg Glu Asn Ser Lys Glu Ser Lys  
 85 90 95

Arg Glu Leu Glu Arg Thr Gln Asp Ser Ile Arg Ser Gln Ile His Asn  
 100 105 110

Lys Phe Ala Glu Arg Ala Asp Phe Lys Asp Met Phe Gly Ala Ser Ile  
 115 120 125

Ile Thr Lys Leu Leu Pro Thr Tyr Ile Lys Gln Asn Ser Glu Tyr Ser  
 130 135 140

Glu Arg Tyr Asp Glu Ser Met Glu Ile Leu Lys Leu Tyr Gly Lys Phe



Asn Glu Val Ile Glu Lys Tyr Tyr Val Glu Lys Thr Gly Gln Ser Val  
 405 410 415  
 Trp Lys Val Glu Ser Tyr Ile Ser Ser Leu Ala Glu Met Ile Lys Leu  
 420 425 430  
 Glu Leu Cys His Glu Ile Asp Asn Asp Glu Lys His Asn Leu Ile Glu  
 435 440 445  
  
 Asp Asp Glu Lys Ile Ser Glu Ile Lys Glu Leu Leu Asp Met Tyr Met  
 450 455 460  
 Asp Val Phe His Ile Ile Lys Val Phe Arg Val Asn Glu Val Leu Asn  
 465 470 475 480  
 Phe Asp Glu Thr Phe Tyr Ser Glu Met Asp Glu Ile Tyr Gln Asp Met  
 485 490 495  
 Gln Glu Ile Val Pro Leu Tyr Asn His Val Arg Asn Tyr Val Thr Gln  
 500 505 510  
 Lys Pro Tyr Lys Gln Glu Lys Tyr Arg Leu Tyr Phe His Thr Pro Thr  
  
 515 520 525  
 Leu Ala Asn Gly Trp Ser Lys Ser Lys Glu Tyr Asp Asn Asn Ala Ile  
 530 535 540  
 Ile Leu Val Arg Glu Asp Lys Tyr Tyr Leu Gly Ile Leu Asn Ala Lys  
 545 550 555 560  
 Lys Lys Pro Ser Lys Glu Ile Met Ala Gly Lys Glu Asp Cys Ser Glu  
 565 570 575  
 His Ala Tyr Ala Lys Met Asn Tyr Tyr Leu Leu Pro Gly Ala Asn Lys  
 580 585 590  
  
 Met Leu Pro Lys Val Phe Leu Ser Lys Lys Gly Ile Gln Asp Tyr His  
 595 600 605  
 Pro Ser Ser Tyr Ile Val Glu Gly Tyr Asn Glu Lys Lys His Ile Lys  
 610 615 620  
 Gly Ser Lys Asn Phe Asp Ile Arg Phe Cys Arg Asp Leu Ile Asp Tyr  
 625 630 635 640  
 Phe Lys Glu Cys Ile Lys Lys His Pro Asp Trp Asn Lys Phe Asn Phe  
 645 650 655

Glu Phe Ser Ala Thr Glu Thr Tyr Glu Asp Ile Ser Val Phe Tyr Arg  
 660 665 670  
 Glu Val Glu Lys Gln Gly Tyr Arg Val Glu Trp Thr Tyr Ile Asn Ser  
 675 680 685  
 Glu Asp Ile Gln Lys Leu Glu Glu Asp Gly Gln Leu Phe Leu Phe Gln  
 690 695 700  
 Ile Tyr Asn Lys Asp Phe Ala Val Gly Ser Thr Gly Lys Pro Asn Leu  
 705 710 715 720  
 His Thr Leu Tyr Leu Lys Asn Leu Phe Ser Glu Glu Asn Leu Arg Asp  
 725 730 735  
  
 Ile Val Leu Lys Leu Asn Gly Glu Ala Glu Ile Phe Phe Arg Lys Ser  
 740 745 750  
 Ser Val Gln Lys Pro Val Ile His Lys Cys Gly Ser Ile Leu Val Asn  
 755 760 765  
 Arg Thr Tyr Glu Ile Thr Glu Ser Gly Thr Thr Arg Val Gln Ser Ile  
 770 775 780  
 Pro Glu Ser Glu Tyr Met Glu Leu Tyr Arg Tyr Phe Asn Ser Glu Lys  
 785 790 795 800  
 Gln Ile Glu Leu Ser Asp Glu Ala Lys Lys Tyr Leu Asp Lys Val Gln  
  
 805 810 815  
 Cys Asn Lys Ala Lys Thr Asp Ile Val Lys Asp Tyr Arg Tyr Thr Met  
 820 825 830  
 Asp Lys Phe Phe Ile His Leu Pro Ile Thr Ile Asn Phe Lys Val Asp  
 835 840 845  
 Lys Gly Asn Asn Val Asn Ala Ile Ala Gln Gln Tyr Ile Ala Gly Arg  
 850 855 860  
 Lys Asp Leu His Val Ile Gly Ile Asp Arg Gly Glu Arg Asn Leu Ile  
 865 870 875 880  
  
 Tyr Val Ser Val Ile Asp Met Tyr Gly Arg Ile Leu Glu Gln Lys Ser  
 885 890 895  
 Phe Asn Leu Val Glu Gln Val Ser Ser Gln Gly Thr Lys Arg Tyr Tyr





<213> Artificial Sequence

<220><223> EGFR\_E2 crRNA

<400> 139

ggagatgtct tgatagcgac ggga

24