

명세서

발명의 명칭: 발암성 돌연변이 유전자 검출용 가이드 RNA 내지 이의 용도

기술분야

- [1] 본 출원은 2021년 05월 11일 출원된 대한민국 특허출원 제10-2021-0061037호를 우선권으로 주장하고, 상기 명세서 전체는 본 출원의 참고문헌이다.
- [2] 본 발명은 CRISPR/Cas9 시스템을 이용하여 야생형 암세포 유전자를 절단함으로써 발암성 돌연변이 유전자를 검출하는데 사용되는 가이드 RNA 내지 이의 용도에 대한 것이다.
- [3] 이 출원은 보건복지부로부터 지원받았다(No. HR14C0005020021(1465033586)).
- [4]

배경기술

- [5] 시퀀싱 기술의 발전에 기초하여 암은 일반적으로 발암성 돌연변이에 의해 유발된다는 것으로 알려져 있다. 여러 연구에서 종양 형성이 특정 유형의 암에서 체세포 돌연변이의 유병률과 밀접한 관련이 있다는 증거를 제공했다. COSMIC (Catalogue of Somatic Mutations in Cancer) 데이터베이스에는 종양 유형 및 질병별로 분류된 수십만 개의 인간 암 관련 체세포 돌연변이가 포함되어 있다.
- [6] 이와 함께, 최근 액체 생검은 기존의 조직 생검에 대한 유용한 대안으로 부상하고 있으며, 혈류에서 순환하는 종양 유래 요소를 탐지하고 모니터링하기 위한 비 침습적 접근 방식을 제공하며 암 바이오 마커의 원천으로 직접 또는 간접적으로 사용할 수 있다. 혈액에는 상세한 분자 분석에 취약한 두 가지 유형의 암 유래 성분이 포함되어 있다. 하나는 온전한 순환 종양 세포 (CTC, circulating tumor cells)와 무 세포 순환 종양 DNA (cfDNA 또는 순환 종양 DNA; ctDNA)이다. 종양의 부피가 증가함에 따라 세포 사멸 및 괴사성 단편을 제거하고 제거하는 식세포의 능력이 초과되어 cfDNA가 혈류로 수동 방출될 수 있다. 이 액체 생검은 종양이 일반적으로 혈류로 방출하는 DNA를 분석하는데 도움이 된다. 종양의 크기에 따라 혈류로 방출되는 cfDNA의 양은 혈장에 존재하는 모든 DNA의 0.01 %에서 90 %까지 다양하다. 따라서 액체 생검은 종양 조직을 확보하지 않고도 종양 분자 프로파일링에 대한 대규모 비 침습적 접근 방식을 제공한다.
- [7] 그러나 혈장의 cfDNA에는 일반적으로 매우 적은 양의 종양 DNA가 포함되어 있으며 이는 종양 부담이나 암 단계에 따라 달라지는 것으로 알려졌다. 따라서 특히 암 초기 단계에서는 ctDNA를 검출하여 종양을 진단하기 위해 매우 민감하고 구체적인 방법이 필요하다. 원핵 생물의 적응성 면역 반응인 CRISPR (Clustered regular interspaced short palindromic repeats) 및 Cas(CRISPR-associated) 단백질 시스템은 특정 DNA 표적 인식 및 절단으로 잘 알려져 있다. CRISPR

시스템의 다양성은 우리와 다른 그룹은 다양한 유기체의 게놈 편집뿐만 아니라 시험관 내에서 DNA를 절단하는 데 광범위하게 활용한다. CRISPR 엔도뉴클레아제(예 : II 형 Cas9 또는 V 형 Cpf1)는 가이드 RNA(gRNA)에 해당하는 표적 DNA 서열의 각각 프로토 스페이스-인접 모티프 (PAM) 다운 스트림 또는 업스트림을 각각 인식하여 표적 DNA에서 이중 가닥 파손을 유도한다. 암 진단에 ctDNA를 사용하려면 비교적 많은 양의 야생형 cfDNA 중에서 미스 센스 돌연변이(missense mutation)가 있는 소량의 ctDNA를 매우 정확한 방식으로 검출해야 한다. 야생형 DNA를 특이적으로 절단함으로써 중앙 특이적 대 야생형 대립 유전자를 풍부하게 할 수 있는데, Cas9-sgRNA 를 이용하여 야생형 표적 DNA를 절단하려면 단일 가이드 RNA가 PAM 부위의 서열을 표적화하도록 설계되어야 한다. 그러나 cfDNA에서 표적화된 DNA PAM 부위는 발암성 돌연변이에 의해 파괴되므로 CRISPR 엔도뉴클레아제(endonuclease)에 의해 인식되지 않아 검출하기 어렵다는 문제점이 존재한다.

[8]

발명의 상세한 설명

기술적 과제

[9] 이에 본 발명자들은 PAM 서열이 존재하지 않는 암 관련 돌연변이에도 활용할 수 있는 암 진단 등을 위한 CRISPR 엔도뉴클레아제 및 Cas9(CRISPR/Cas9) 을 이용하여 검출하는 시스템을 제공하고자 예의 노력한 결과, 암세포의 돌연변이 DNA 가 아닌 야생형 DNA 만을 특이적으로 절단할 수 있는 가이드 RNA 를 제작하여 본 발명을 완성하였다.

[10] 따라서, 본 발명의 목적은 CRISPR/Cas9 시스템의 가이드 RNA(guide RNA) 내지 이를 포함하는 암 진단 또는 예후 예측용 조성물 내지 키트를 제공하는 것이다.

[11] 본 발명의 또 다른 목적은, 환자의 시료에 CRISPR/Cas9 시스템의 가이드 RNA(guide RNA) 를 처리하는 단계를 포함하는 포함하는 암 진단 또는 예후 예측을 위한 정보제공 방법을 제공하는 것이다.

[12]

과제 해결 수단

[13] 본 발명은, CRISPR/Cas9 시스템에 이용되는 가이드 RNA(guide RNA) 로서,

[14] 상기 가이드 RNA 는 야생형 무 세포 DNA 를 절단하는 것을 특징으로 하는 가이드 RNA를 제공한다.

[15] 본 발명의 바람직한 일실시예에 따르면, 상기 가이드 RNA는 서열번호 1 내지 17 로 이루어진 군에서 선택되는 어느 하나 이상의 뉴클레오타이드 서열을 포함하는 것일 수 있다.

[16] 본 발명의 바람직한 일실시예에 따르면, 상기 가이드 RNA 는 가이드 RNA 의

트렁크(trunk) 영역, 세미 시드(semi-seed) 영역 및 시드(seed) 영역으로 이루어진 군에서 선택되는 어느 하나 이상의 영역에서 전환 돌연변이(transversion mutation)가 발생한 것일 수 있다.

- [17] 본 발명의 바람직한 일실시예에 따르면, 상기 가이드 RNA 는 서열번호 2, 서열번호 4, 서열번호 9 및 서열번호 13 으로 이루어진 군에서 선택되는 어느 하나 이상의 뉴클레오타이드 서열을 포함하는 것일 수 있다.
- [18] 본 발명의 바람직한 일실시예에 따르면, 상기 가이드 RNA 는 EGFR, MET, KRAS, PIK3CA 및 BRAF 로 이루어진 군에서 선택되는 어느 하나 이상의 암 관련 돌연변이를 검출할 수 있는 것일 수 있다.
- [19] 본 발명의 바람직한 일실시예에 따르면, 상기 EGFR 돌연변이를 검출하는 경우, 상기 가이드 RNA 는 서열번호 1 내지 4, 8, 9 및 11 내지 13 으로 이루어진 군에서 선택되는 어느 하나 이상의 뉴클레오타이드 서열을 포함하는 것일 수 있다.
- [20] 본 발명의 바람직한 일실시예에 따르면, 상기 MET 돌연변이를 검출하는 경우, 상기 가이드 RNA 는 서열번호 5 내지 7 로 이루어진 군에서 선택되는 어느 하나 이상의 뉴클레오타이드 서열을 포함하는 것일 수 있다.
- [21] 본 발명의 바람직한 일실시예에 따르면, 상기 KRAS 돌연변이를 검출하는 경우, 상기 가이드 RNA 는 서열번호 10 의 뉴클레오타이드 서열을 포함하는 것일 수 있다.
- [22] 본 발명의 바람직한 일실시예에 따르면, 상기 PIK3CA 돌연변이를 검출하는 경우, 상기 가이드 RNA 는 서열번호 14 내지 16 으로 이루어진 군에서 선택되는 어느 하나 이상의 뉴클레오타이드 서열을 포함하는 것일 수 있다.
- [23] 본 발명의 바람직한 일실시예에 따르면, 상기 BRAF 돌연변이를 검출하는 경우, 상기 가이드 RNA 는 서열번호 17 의 뉴클레오타이드 서열을 포함하는 것일 수 있다.
- [24] 본 발명은, 서열번호 1 내지 17 로 이루어진 군에서 선택되는 어느 하나 이상의 뉴클레오타이드 서열을 포함하는 CRISPR/Cas9 의 가이드 RNA 를 포함하는 암 진단 또는 예후 예측용 조성물을 제공한다.
- [25] 본 발명의 바람직한 일실시예에 따르면, 상기 가이드 RNA 는 가이드 RNA 의 트렁크(trunk) 영역, 세미 시드(semi-seed) 영역 및 시드(seed) 영역으로 이루어진 군에서 선택되는 어느 하나 이상의 영역에서 전환 돌연변이(transversion mutation)가 발생한 것일 수 있다.
- [26] 본 발명의 바람직한 일실시예에 따르면, 상기 가이드 RNA 는 EGFR, MET, KRAS, PIK3CA 및 BRAF 로 이루어진 군에서 선택되는 어느 하나 이상의 암 관련 돌연변이를 검출할 수 있는 것일 수 있다.
- [27] 본 발명의 바람직한 일실시예에 따르면, 상기 가이드 RNA 는 야생형 무 세포 DNA 를 절단하는 것일 수 있다.
- [28] 본 발명은 또한, 상기 본 발명의 바람직한 일실시예에 따르면, 상기 암은 폐암,

- 췌장암, 대장암, 위암, 유방암, 신장암, 간암, 혈액암, 방광암 및 자궁암 으로 이루어진 군에서 선택되는 어느 하나 이상의 암인 것일 수 있다.
- [29] 본 발명은 또한, 상기 조성물을 포함하는 암 진단 또는 예후 예측용 키트를 제공한다.
- [30] 본 발명은 또한, 환자의 치료에 상기 조성물을 처리하는 단계를 포함하는 암 진단 또는 예후 예측을 위한 정보제공 방법을 제공한다.
- [31]
- [32] 상기 기재한 바와 같이, 기존 CRISPR/Cas9 시스템 으로 무 세포 DNA 를 이용한 발암성 돌연변이 유전자 검출은 3 개의 PAM 서열 (NGG) 뉴클레오타이드의 돌연변이를 타겟으로 하여 가이드 RNA 로 검출할 확률이 낮다는 점 등 발암성 돌연변이 유전자를 검출하는데에 한계가 존재하였다.
- [33] 상기와 같은 한계점을 극복하기 위하여, 본 발명의 가이드 RNA 는 CRISPR/Cas9 시스템을 통하여 발암성 돌연변이 유전자가 아닌 야생형 무 세포 DNA 를 타겟 및 절단하며, 특히 가이드 RNA의 트렁크(trunk) 영역에서 전환 돌연변이(transversion mutation)를 일으킨 가이드 RNA를 이용함으로써 발암성 돌연변이 유전자만을 선택적으로 타겟 및 증폭할 수 있다.
- [34]
- [35] 따라서, 본 발명은 CRISPR/Cas9 시스템에 이용되는 가이드 RNA(guide RNA)로서, 상기 가이드 RNA 는 야생형 무 세포 DNA 를 절단하는 것일 수 있다.
- [36] 상기 'CRISPR/Cas9 시스템'이란, Cas9 이라는 특정한 염기서열을 인식하여 유전자를 자를 수 있는 효소가 그 염기서열을 포함하고 있는 단일가닥 가이드 RNA(single guide RNA; sgRNA)를 인식하여 특정한 DNA를 자를 수 있는 시스템을 의미한다. CRISPR/Cas9 시스템은 타겟 DNA를 정확히 인식할 수 있게 디자인된 sgRNA와 이를 토대로 타겟 DNA를 자를 수 있는 Cas9으로 구성되며, sgRNA를 인식한 Cas9은 DNA 이중나선을 절단한다.
- [37] 본 발명의 바람직한 일 실시예에 따르면, 상기 가이드 RNA는 서열번호 1 내지 17 로 이루어진 군에서 선택되는 어느 하나 이상의 뉴클레오타이드 서열을 포함하는 것일 수 있다.
- [38] 서열번호 1 의 뉴클레오타이드 서열을 포함하는 가이드 RNA 는 EGFR (p.T790M) 돌연변이를 검출할 수 있는 야생형 가이드 RNA(wild sgRNA) 로서, in vitro 절단 분석에 사용되는 경우 서열번호 18 을 타겟 서열로 할 수 있다.
- [39] 서열번호 2 의 뉴클레오타이드 서열을 포함하는 가이드 RNA 는 EGFR (p.T790M) 돌연변이를 검출할 수 있는 전환 돌연변이 가이드 RNA(transversion sgRNA) 로서, in vitro 절단 분석에 사용되는 경우 서열번호 19 를 타겟 서열로 할 수 있다. 서열번호 2 의 뉴클레오타이드 서열을 포함하는 가이드 RNA 를 이용하는 경우 1% 내지 32% 로 돌연변이 서열을 증폭시킬 수 있다.
- [40] 서열번호 3 의 뉴클레오타이드 서열을 포함하는 가이드 RNA 는 EGFR (p.C797S) 돌연변이를 검출할 수 있는 야생형 가이드 RNA(wild sgRNA) 로서, in

- vitro 절단 분석에 사용되는 경우 서열번호 20 을 타겟 서열로 할 수 있다.
- [41] 서열번호 4 의 뉴클레오타이드 서열을 포함하는 가이드 RNA 는 EGFR (p.C797S) 돌연변이를 검출할 수 있는 전환 돌연변이 가이드 RNA(transversion sgRNA) 로서, in vitro 절단 분석에 사용되는 경우 서열번호 21 을 타겟 서열로 할 수 있다. 서열번호 4 의 뉴클레오타이드 서열을 포함하는 가이드 RNA 를 이용하는 경우 1% 내지 13% 로 돌연변이 서열을 증폭시킬 수 있다.
- [42] 서열번호 5 의 뉴클레오타이드 서열을 포함하는 가이드 RNA 는 MET c.2888 Int13-Exon14 Acceptor (Exon-14 Skipping) 돌연변이를 검출할 수 있는 야생형 가이드 RNA(wild sgRNA) 로서, in vitro 절단 분석에 사용되는 경우 서열번호 22 또는 23 을 타겟 서열로 할 수 있다.
- [43] 서열번호 6 의 뉴클레오타이드 서열을 포함하는 가이드 RNA 는 MET c.3028 Exon14-Int14 Donor (Exon-14 Skipping) 돌연변이를 검출할 수 있는 야생형 가이드 RNA(wild sgRNA) 로서, in vitro 절단 분석에 사용되는 경우 서열번호 24 또는 25 를 타겟 서열로 할 수 있다.
- [44] 서열번호 7 의 뉴클레오타이드 서열을 포함하는 가이드 RNA 는 MET 'TAC' Y1003X (Exon-14 Skipping) 돌연변이를 검출할 수 있는 야생형 가이드 RNA(wild sgRNA) 로서, in vitro 절단 분석에 사용되는 경우 서열번호 26 또는 27 을 타겟 서열로 할 수 있다.
- [45] 서열번호 8 의 뉴클레오타이드 서열을 포함하는 가이드 RNA 는 EGFR (L718Q) 돌연변이를 검출할 수 있는 야생형 가이드 RNA(wild sgRNA) 로서, in vitro 절단 분석에 사용되는 경우 서열번호 28 을 타겟 서열로 할 수 있다.
- [46] 서열번호 9 의 뉴클레오타이드 서열을 포함하는 가이드 RNA 는 EGFR (L718Q) 돌연변이를 검출할 수 있는 전환 돌연변이 가이드 RNA(transversion sgRNA) 로서, in vitro 절단 분석에 사용되는 경우 서열번호 29 를 타겟 서열로 할 수 있다. 서열번호 9 의 뉴클레오타이드 서열을 포함하는 가이드 RNA 를 이용하는 경우 0.8% 내지 23.7% 로 돌연변이 서열을 증폭시킬 수 있다.
- [47] 서열번호 10 의 뉴클레오타이드 서열을 포함하는 가이드 RNA 는 KRAS (c.35G) 돌연변이를 검출할 수 있는 야생형 가이드 RNA(wild sgRNA) 로서, in vitro 절단 분석에 사용되는 경우 서열번호 30 또는 31 을 타겟 서열로 할 수 있다. 서열번호 10 의 뉴클레오타이드 서열을 포함하는 가이드 RNA 를 이용하는 경우 1.5% 내지 50% 로 돌연변이 서열을 증폭시킬 수 있다.
- [48] 서열번호 11 의 뉴클레오타이드 서열을 포함하는 가이드 RNA 는 EGFR, 15bp Del p.(729_761) 돌연변이를 검출할 수 있는 야생형 가이드 RNA(wild sgRNA) 로서, in vitro 절단 분석에 사용되는 경우 서열번호 32 또는 33 을 타겟 서열로 할 수 있다. 서열번호 11 의 뉴클레오타이드 서열을 포함하는 가이드 RNA 를 이용하는 경우 0.5% 내지 23% 로 돌연변이 서열을 증폭시킬 수 있다.
- [49] 서열번호 12 의 뉴클레오타이드 서열을 포함하는 가이드 RNA 는 EGFR (p.L858R) 돌연변이를 검출할 수 있는 야생형 가이드 RNA(wild sgRNA) 로서, in

in vitro 절단 분석에 사용되는 경우 서열번호 34 를 타겟 서열로 할 수 있다.

- [50] 서열번호 13 의 뉴클레오타이드 서열을 포함하는 가이드 RNA 는 EGFR (p.L858R) 돌연변이를 검출할 수 있는 전환 돌연변이 가이드 RNA(transversion sgRNA) 로서, in vitro 절단 분석에 사용되는 경우 서열번호 35 를 타겟 서열로 할 수 있다. 서열번호 13 의 뉴클레오타이드 서열을 포함하는 가이드 RNA 를 이용하는 경우 1.7% 내지 63% 로 돌연변이 서열을 증폭시킬 수 있다.
- [51] 서열번호 14 의 뉴클레오타이드 서열을 포함하는 가이드 RNA 는 PIK3CA (p.E542K 또는 c.1624G>A) 돌연변이를 검출할 수 있는 야생형 가이드 RNA(wild sgRNA) 로서, in vitro 절단 분석에 사용되는 경우 서열번호 36 또는 37 을 타겟 서열로 할 수 있다.
- [52] 서열번호 15 의 뉴클레오타이드 서열을 포함하는 가이드 RNA 는 PIK3CA (p.E545? 또는 c.1633G>?) 돌연변이를 검출할 수 있는 야생형 가이드 RNA(wild sgRNA) 로서, in vitro 절단 분석에 사용되는 경우 서열번호 38 또는 39 를 타겟 서열로 할 수 있다.
- [53] 서열번호 16 의 뉴클레오타이드 서열을 포함하는 가이드 RNA 는 PIK3CA (p.H1047? 또는 c.3140A>?) 돌연변이를 검출할 수 있는 야생형 가이드 RNA(wild sgRNA) 로서, in vitro 절단 분석에 사용되는 경우 서열번호 40 또는 41 을 타겟 서열로 할 수 있다.
- [54] 서열번호 17 의 뉴클레오타이드 서열을 포함하는 가이드 RNA 는 PIK3CA (p.V600E 또는 c.1799T>A) 돌연변이를 검출할 수 있는 야생형 가이드 RNA(wild sgRNA) 로서, in vitro 절단 분석에 사용되는 경우 서열번호 42 또는 43 을 타겟 서열로 할 수 있다.
- [55]
- [56] 상기 가이드 RNA 는 가이드 RNA 의 트렁크(trunk) 영역, 세미 시드(semi-seed) 영역 및 시드(seed) 영역으로 이루어진 군에서 선택되는 어느 하나 이상의 영역에서 전환 돌연변이(transversion mutation) 가 발생한 것을 특징으로 하는 가이드 RNA 일 수 있다.
- [57] SpCas9로 DNA 절단을 표적화하려면 주로 표적 서열의 3 '말단에 인접한 3 개의 PAM 서열 (NGG) 뉴클레오타이드와 함께 표적 프로토 스페이서 서열의 특정 20 개 뉴클레오타이드가 필요하다. PAM 서열 내 돌연변이 (mismatched PAM-NGN 또는 임의의 발암성 돌연변이에 의해 유발됨)는 Cas9가 무 세포 DNA에서 야생형 표적 대립 유전자 (PAM-NGG)로 절단하여 야생형 DNA에서 돌연변이 대립 유전자를 풍부하게하기 위한 쉬운 최상의 후보이다. 그러나 모든 암 관련 돌연변이에 대해 이와 같이 일치하지 않는 PAM 표적 서열을 갖는 것은 불가능하다. 본 발명자들은 상기 문제점을 해결하고 모든 종류의 돌연변이를 표적으로 하기 위해 가이드 RNA의 트렁크(trunk) 영역에서 전환 돌연변이(transversion mutation, 퓨린에서 피리미딘으로)를 제조하였다. sgRNA 활성화에 대한 내성에 따라 가이드 RNA 내의 1 염기 불일치(mismatch)는 세미

시드(semi-seed), 시드(seed), 트렁크(trunk) 및 다중(promiscuous) 영역으로 구분된다. 위치 4에서 7까지의 시드 영역 내 1 염기 불일치에 대한 가이드 RNA 내성은 해당 sgRNA의 가장 낮은 내성이다. 다음으로, 위치 1-3 및 8-11에서 시드 영역의 양쪽은 불일치에 대한 중간 허용 오차이며 세미 시드 영역으로 정의된다. 위치 12-19에서 1 염기 불일치는 절단 빈도를 중간 수준으로 줄였지만 sgRNA 활성을 유지하고 이 중간 영역을 트렁크 영역이라고 한다. 위치 20에서, 1 염기 불일치는 Cas9 뉴클레아제에 의해 유도된 절단 빈도를 약간 감소시키고 이 위치를 다중 영역으로 명명한다. 이 전략을 기반으로 Cas9 뉴클레아제를 전환 돌연변이(transversion mutation) 가이드 RNA로 처리하여 낮은 빈도의 돌연변이 대립 유전자를 강화한 다음 PCR 증폭을 통해 야생형 표적 대립 유전자를 제거할 수 있다. 돌연변이 표적 대립 유전자는 Cas9 뉴클레아제에 의해 인식되지 않거나 절단되지 않는다. 왜냐하면 2 개의 돌연변이에 대한 낮은 내성, 즉 하나는 sgRNA 내의 전환 돌연변이이고 두 번째는 발암성 돌연변이의 표적 돌연변이이기 때문이다(도 3a, 도 3b).

- [58] 상기 Cas9 뉴클레아제는 다양한 변이체를 포함할 수 있으며, Cas9, eCas9 또는 HFCas9 일 수 있다.
- [59] 본 발명의 바람직한 일실시예에 따르면, 상기 전환 돌연변이(transversion mutation)가 발생한 가이드 RNA는 서열번호 2, 서열번호 4, 서열번호 9 및 서열번호 13으로 이루어진 군에서 선택되는 어느 하나 이상의 뉴클레오타이드 서열을 포함하는 것일 수 있다.
- [60] 본 발명의 바람직한 일실시예에 따르면, 상기 가이드 RNA는 EGFR, MET, KRAS, PIK3CA 및 BRAF로 이루어진 군에서 선택되는 어느 하나 이상의 암 관련 돌연변이를 검출할 수 있는 것일 수 있다.
- [61] 본 발명의 바람직한 일실시예에 따르면, 상기 EGFR 돌연변이를 검출하는 경우, 상기 가이드 RNA는 서열번호 1 내지 4, 8, 9 및 11 내지 13으로 이루어진 군에서 선택되는 어느 하나 이상의 뉴클레오타이드 서열을 포함하는 것일 수 있다.
- [62] 본 발명의 바람직한 일실시예에 따르면, 상기 MET 돌연변이를 검출하는 경우, 상기 가이드 RNA는 서열번호 5 내지 7로 이루어진 군에서 선택되는 어느 하나 이상의 뉴클레오타이드 서열을 포함하는 것일 수 있다.
- [63] 본 발명의 바람직한 일실시예에 따르면, 상기 KRAS 돌연변이를 검출하는 경우, 상기 가이드 RNA는 서열번호 10의 뉴클레오타이드 서열을 포함하는 것일 수 있다.
- [64] 본 발명의 바람직한 일실시예에 따르면, 상기 PIK3CA 돌연변이를 검출하는 경우, 상기 가이드 RNA는 서열번호 14 내지 16으로 이루어진 군에서 선택되는 어느 하나 이상의 뉴클레오타이드 서열을 포함하는 것일 수 있다.
- [65] 본 발명의 바람직한 일실시예에 따르면, 상기 BRAF 돌연변이를 검출하는 경우, 상기 가이드 RNA는 서열번호 17의 뉴클레오타이드 서열을 포함하는

것일 수 있다.

[66]

[67] 또한, 본 발명은, 서열번호 1 내지 17 로 이루어진 군에서 선택되는 어느 하나 이상의 뉴클레오타이드 서열을 포함하는 CRISPR/Cas9 의 가이드 RNA를 제공할 수 있다. 상기 가이드 RNA 는 야생형 무 세포 DNA 를 절단하는 것 일 수 있다.

[68] 본 발명의 바람직한 일실시예에 따르면, 상기 가이드 RNA 는 가이드 RNA 의 트렁크(trunk) 영역, 세미 시드(semi-seed) 영역 및 시드(seed) 영역으로 이루어진 군에서 선택되는 어느 하나 이상의 영역에서 전환 돌연변이(transversion mutation) 가 발생한 것일 수 있다.

[69] 본 발명은 또한, 상기 가이드 RNA 를 포함하는 암 진단 또는 예후 예측용 조성물을 제공할 수 있다.

[70] 본 발명의 바람직한 일실시예에 따르면, 상기 암은 폐암, 췌장암, 대장암, 위암, 유방암, 신장암, 간암, 혈액암, 방광암 및 자궁암 으로 이루어진 군에서 선택되는 어느 하나 이상의 암인 것일 수 있으나, EGFR, MET, KRAS, PIK3CA 또는 BRAF 돌연변이로 인해 발생하는 암이라면 제한 없이 해당할 수 있다.

[71] 본 발명은 또한, 상기 조성물을 포함하는 암 진단 또는 예후 예측용 키트를 제공할 수 있다.

[72] 상기 키트는 통상적으로 사용되는 발현 수준 분석방법에 적합한 한 종류 또는 그 이상의 다른 구성성분 조성물, 용액 또는 장치로 구성될 수 있다. 예를 들어, 단백질 발현 수준을 측정하기 위한 키트는 항체의 면역학적 검출을 위하여 기질, 적당한 완충용액, 발색 효소 또는 형광물질로 표지된 2차 항체, 발색 기질 등을 포함할 수 있다.

[73] 본 발명의 키트는 평가 대상체로부터 샘플을 획득하기 위한 샘플 추출 수단을 포함할 수 있다. 샘플 추출 수단은 바늘 또는 시린지 등을 포함할 수 있다. 키트는 액체, 기체 또는 반고체일 수 있는, 추출된 샘플을 수용하기 위한 샘플 수집 용기를 포함할 수 있다. 키트는 사용 지침을 추가로 포함할 수 있다. 상기 샘플은 발암성 돌연변이 유전자가 존재하거나 분비될 수 있는 임의의 신체 샘플일 수 있다. 예를 들어, 샘플은 소변, 대변, 모발, 땀, 타액, 체액, 혈액, 세포 또는 조직일 수 있다. 신체 샘플에서 발암성 돌연변이 유전자의 측정은 전체 샘플 또는 가공된 샘플 상에서 수행될 수 있다.

[74] 본 발명은 또한, 환자의 시료에 제1항의 가이드 RNA 를 처리하는 단계를 포함하는 암 진단 또는 예후 예측을 위한 정보제공 방법을 제공할 수 있다.

[75] 상기 시료는 소변, 대변, 모발, 땀, 타액, 체액, 혈액, 세포 또는 조직일 수 있다.

[76]

발명의 효과

[77] 본 발명의 가이드 RNA 를 이용하여 CRISPR/Cas9 시스템으로 암 돌연변이를 검출하는 경우, 야생형 무 세포 DNA(cfDNA) 만을 절단하고 발암성 돌연변이

DNA 는 절단하지 않아 기존 CRISPR/Cas9 시스템의 가이드 RNA 를 이용하는 경우에 비하여 발암성 돌연변이 DNA 만을 선택적으로 증폭하는데 효과적으로 사용될 수 있다. 특히, 본 발명의 가이드 RNA(야생형 가이드 RNA)와 전환 돌연변이 가이드 RNA(transversion sgRNA)를 병용하는 경우 cf DNA 만을 절단하고 발암성 돌연변이 DNA 만을 선택적으로 증폭하는 효과가 유의적으로 증가할 수 있다.

[78]

도면의 간단한 설명

- [79] 도 1은 비편평성(non-squamous; 선암, adenocarcinoma; EGFR-TK, ALK 또는 ROS-1 양성) 암종 및 비소세포암종 환자를 위한 약물 치료 또는 전신 항암 요법을 위한 질병 진행 및 옵션을 폐암에 대한 NICE 가이드 라인과 함께 요약한 것을 나타낸다.
- [80] 도 2는 Cas9-sgRNA 를 이용하여 무세포 DNA(cfDNA) 내 돌연변이 대립 유전자를 증폭하기 위한 도식을 나타낸다.
- [81] 도 3a 및 도 3b 는 무세포 DNA(cfDNA) 에서 Cas9 처리로 가능한 모든 돌연변이 대립 유전자를 증폭시키기 위한 가이드 RNA 설계 전략을 나타낸다.
- [82] 도 4a 및 도 4b는 본 발명의 가이드 RNA를 이용하여 KRAS 돌연변이(c.35G>C) 서열을 검출하는 in vitro 절단 분석(in vitro cleavage assay) 및 그 결과를 나타낸다.
- [83] 도 5는 본 발명의 가이드 RNA를 이용하여 EGFR 돌연변이(c.2369C>T 또는 p.T790M) 서열을 검출하는 in vitro 절단 분석(in vitro cleavage assay) 및 그 결과를 나타낸다.
- [84] 도 6a 및 도 6b 는 본 발명의 가이드 RNA를 이용하여 EGFR 돌연변이(c.2573 T>G 또는 p.L858R) 서열을 검출하는 in vitro 절단 분석(in vitro cleavage assay) 및 그 결과를 나타낸다.
- [85] 도 7 는 본 발명의 가이드 RNA를 이용하여 환자(LA-29) 시료 내 EGFR 돌연변이(c.2573 T>G 또는 p.L858R) 서열을 검출하는 in vitro 절단 분석(in vitro cleavage assay) 및 그 결과를 나타낸다.
- [86] 도 8은 본 발명의 가이드 RNA를 이용하여 EGFR 돌연변이(p.C797S, L718Q 또는 p.(729_761)손실) 또는 MET 돌연변이(c.2888, c.3028 또는 TAC>Y1003X) 서열을 검출하는 in vitro 절단 분석(in vitro cleavage assay) 및 그 결과를 나타낸다.
- [87] 도 9 은 본 발명의 가이드 RNA를 이용하여 BRAF 돌연변이(p.V600E 또는 c.1799T>A) 또는 PIK3CA 돌연변이(c.1624G>A 또는 p.E542K; p.H1047? 또는 c.3140A>?) 서열을 검출하는 in vitro 절단 분석(in vitro cleavage assay) 및 그 결과를 나타낸다.
- [88] 도 10은 본 발명의 가이드 RNA를 이용하여 EGFR 돌연변이(T790M) 서열을 검출한 PCR 결과를 나타낸다.
- [89] 도 11은 본 발명의 가이드 RNA를 이용하여 EGFR 돌연변이(T790M) 서열을

- 검출한 PCR 결과를 정량적으로 나타낸다.
- [90] 도 12은 야생형 유전자 및 돌연변이 유전자(KRAS c.35G>C)를 다양한 비율로 혼합한 후 Cas9 을 처리한 결과를 나타낸다.
- [91] 도 13 은 야생형 유전자 및 돌연변이 유전자(EGFR-c.2369C>T 또는p.T790M)를 다양한 비율로 혼합한 후 Cas9 을 처리한 결과를 나타낸다.
- [92] 도 14 는 야생형 유전자 및 돌연변이 유전자(EGFR c.2390G>C 또는 p.C797S)를 다양한 비율로 혼합한 후 Cas9 을 처리한 결과를 나타낸다.
- [93] 도 15 은 야생형 유전자 및 돌연변이 유전자(EGFR c.2153T to A (L718Q))를 다양한 비율로 혼합한 후 Cas9 을 처리한 결과를 나타낸다.
- [94] 도 16는 cfDNA 참조 표준 세트로서, ddPCR(Droplet digital PCR)에 의해 수집된 대립 유전자 빈도 및 카피 수 데이터를 나타낸다. 다른 내인성 변이체가 미리 설정되어 있지만 제시된 데이터만 각 배치 내에서 테스트 되었다(분석 배경으로 인해 예상 AF에 대해 확인되지 않았지만 ddPCR에 의해 변이체가 검출되었다; 사용된 QC 분석으로 인해 NRAS Q61K와의 프로브 경쟁이 입증 되었기 때문에 예상 AF에 대해 확인되지 않았지만 ddPCR에 의해 변이체가 검출되었다).
- [95] 도 17은 cfDNA 참조 표준 세트를 기초로 본 발명의 가이드 RNA를 이용하여 돌연변이 유전자에 대한 PCR 결과를 나타낸다.
- [96] 도 18은 환자 시료 내 무 세포 DNA 검출을 위한 워크플로우를 나타낸다.
- [97] 도 19는 환자(LA-29, LA-209 또는 LA-292) 시료 내 EGFR 돌연변이 중 L858R 돌연변이를 본 발명의 가이드 RNA 및 Cas9 를 이용하여 검출한 결과를 나타낸다.
- [98] 도 20은 환자(LA-232) 시료 내 EGFR 돌연변이 중 T790M 또는 E19del 돌연변이를 본 발명의 가이드 RNA 및 Cas9 를 이용하여 단일 검출한 결과를 나타낸다.
- [99] 도 21는 환자(LA-232) 시료 내 EGFR 돌연변이 중 T790M 또는 E19del 돌연변이를 본 발명의 가이드 RNA 및 Cas9 를 이용하여 다중 검출한 결과를 나타낸다.
- [100] 도 22는 환자(LA-322) 시료 내 EGFR 돌연변이 중 T790M 또는 E19del 돌연변이를 본 발명의 가이드 RNA 및 Cas9 를 이용하여 단일 검출한 결과를 나타낸다.
- [101] 도 23은 환자(LA-322) 시료 내 EGFR 돌연변이 중 T790M 또는 E19del 돌연변이를 본 발명의 가이드 RNA 및 Cas9 를 이용하여 다중 검출한 결과를 나타낸다.
- [102] **발명의 실시를 위한 최선의 형태**
- [103] [실시예 1]
- [104] 약물 내성 돌연변이에 대한 가이드 RNA 제작

- [105] IVT(In-vitro transcription cleavage) 분석을 테스트하기 위한 선택된 유전자의 야생형 및 돌연변이 표적 서열과, 가이드 RNA 서열을 제작하였다.
- [106]

[107] [표1]

유전자/ Mut.AA	뉴클레오타이드 (WT/Mut)	In-vitro Cleavage Assay 타겟 서열	sgRNA + PAM 서열	sgRNA A 유형
EGFR (p.T790M or c.2369C>T)	WT - C	TCACCTCCACCGTGCAGCTCA TCAcGCAGCTCATGCCCTTC (서열번호 18)	CTGCgTGATGA GCTGCACGGT GG (서열번호 1)	Wild sgRNA
	Mut - T	TCACCTCCACCGTGCAGCTCA TCAtGCAGCTCATGCCCTTC (서열번호 19)	CTGCgTGcTGA GCTGCACGGT GG (서열번호 2)	Trans version sgRNA
EGFR (p.C797S or c.2390 G>C)	WT - G	ACGCAGCTCATGCCCTTCGGC TgCCTCCTGGACTATGTCCG (서열번호 20)	CATGCCCTTCG GCTgCCTCCT GG G (서열번호 3)	Wild sgRNA
	Mut - C	ACGCAGCTCATGCCCTTCGGC TcCCTCCTGGACTATGTCCG (서열번호 21)	CATGCaCTTCG GCTgCCTCCT GG G (서열번호 4)	Trans version sgRNA
MET c.2888 Int13-Exon14 Acceptor (Exon-14 Skipping)	WT - GA	TTAACAAGCTCTTTCTTTCTC TCTGTTTTAAgaTCTGGGCA (서열번호 22)	TTTCTCTCTGTT TTAAgaTCT GG (서열번호 5)	Wild sgRNA
	Mut - ??	TTAACAAGCTCTTTCTTTCTC TCTGTTTTAAgaTCTGGGCA (서열번호 23)	TTTCTCTCTGTT TTAAgaTCT GG	

MET c.3028 Exon14-Int14 Donor (Exon-14 Skipping)	WT - AC	ATTTCTCAGAACAAATAAACTG AAATAT ac CTTCTGGAAAAG (서열번호 24)	ATAAACTGAA ATAT ac CTTCT CG G (서열번호 6)	Wild sgRN A
	Mut - ??	ATTTCTCAGAACAAATAAACTG AAATAT ac CTTCTGGAAAAG (서열번호 25)	ATAAACTGAA ATAT ac CTTCT CG G	
MET 'TAC' Y1003X (Exon-14 Skipping)	WT - TAC	GAAATGGTTTCAAATGAATCT GTAGACT ac CGAGCTACTTT (서열번호 26)	AAATGAATCTG TAGACT ac CGA G (서열번호 7)	Wild sgRN A
	Mut - ???	GAAATGGTTTCAAATGAATCT GTAGACT ac CGAGCTACTTT (서열번호 27)	AAATGAATCTG TAGACT ac CGA G	
EGFR (L718Q or c.2153 T> A)	WT - T	GAATTCAAAAAGATCAAAGT GCtGGGCTCCGGTGC GTTCGG (서열번호 28)	AAGATCAAAG TGcGGGCTCC GG (서열번호 8)	Wild sgRN A
	Mut - A	GAATTCAAAAAGATCAAAGT GC a GGGCTCCGGTGC GTTCGG (서열번호 29)	AAGATCA a AGT GCtGGGCTCC GG G (서열번호 9)	
KRAS (c.35G)	WT - G	TGACTGAATATAAACTTGTGG TAGTTGGAGCTG g TGGCGTAG GCA (서열번호 30)	AAACTTGTGGT AGTTGGAGCT GG (서열번호 10)	Wild sgRN A
	Mut - C (A or T)	TGACTGAATATAAACTTGTGG TAGTTGGAGCTG c TGGCGTAG GCA (서열번호 31)	AAACTTGTGGT AGTTGGAGCT Gg	
EGFR, 15bp Del p.(729_761)	WT - GGAATT AAGAGA AGC	AAGTTAAAATTCCCGTCGCTA TCA Agga attaagagaagcAACATC (서열번호 32)	TAAAATTCCCG TCGCTATCA AG G (서열번호 11)	Wild sgRN A
	Mut - 15bp Deletion	AAGTTAAAATTCCCGTCGCTA TCA Agga attaagagaagcAACATC (서열번호 33)	TAAAATTCCCG TCGCTATCA AG G	

EGFR (p.L858R or c.2573 T>G)	WT - T	GATCACAGATTTTGGGCtGGC CAAAC T GCTGGGTGCGGAAG (서열번호 34)	TTTTGGGCtGG CCAAAC TGCTG G (서열번호 12)	Wild sgRN A
	Mut - G	GATCACAGATTTTGGGCgGGC CAAAC T GCTGGGTGCGGAAG (서열번호 35)	TTTTGtG Ct GGC CAAAC TGCTG G (서열번호 13)	Trans versio n sgRN A
PIK3CA p.E542K or c.1624G> A	WT - G	CACGAGATCCTCTCTCT g AAA TCACTGAGCAGGAGAAAG (서열번호 36)	TGCTCAGTGAT TTcAGAGAG AG G (서열번호 14)	Wild sgRN A
	Mut - A	CACGAGATCCTCTCTCT a AAA TCACTGAGCAGGAGAAAG (서열번호 37)	TGCTCAGTGAT TTcAGAGAG AG G	
PIK3CA p.E545? or c.1633G> ?	WT - G	CACGAGATCCTCTCTCTGAAA TCACT g AGCAGGAGAAAG (서열번호 38)	TCTCTCTGAAA TCACT g AGC AG G (서열번호 15)	Wild sgRN A
	Mut - ?? (A or C or T)	CACGAGATCCTCTCTCTGAAA TCACT a AGCAGGAGAAAG (서열번호 39)	TCTCTCTGAAA TCACT g AGC AG G	
PIK3CA p.H1047? or c.3140A> ?	WT - T	CATGAAACAAATGAATGATG CACaTCATGGTGGCTGG (서열번호 40)	CAAATGAATG ATGCACaTCAT GG (서열번호 16)	Wild sgRN A
	Mut - ?? (G or C or T)	CATGAAACAAATGAATGATG CAC g TCATGGTGGCTGG (서열번호 41)	CAAATGAATG ATGCACaTCAT GG	
BRAF p.V600E or c.1799T> A	WT - T	GATTTTGGTCTAGCTACAGtG AAATCTCGATGGAGTGGG (서열번호 42)	TAGCTACAGtG AAATCTCGAT G G (서열번호 17)	Wild sgRN A
	Mut - A	GATTTTGGTCTAGCTACAGaG AAATCTCGATGGAGTGGG (서열번호 43)	TAGCTACAGtG AAATCTCGAT G G	

[108]

[109] 상기 [표 1]은 약물 내성 돌연변이에 대한 가이드 RNA를 나타낸다. IVT(In-vitro

transcription cleavage) 분석을 테스트하기 위한 선택된 유전자의 야생형 및 돌연변이 표적 서열은 모두 IVT 테스트를 위해 T-블런트 클로닝 벡터로 서브 클로닝되고, 야생형 표적 서열에서 돌연변이 위치는 굵게 강조 표시되고 해당 뉴클레오타이드 돌연변이는 빨간색(소문자)으로 표시된다. 돌연변이의 '?' 표시는 야생형 뉴클레오타이드가 아닌 임의의 뉴클레오타이드를 의미한다. 각 표적에 대한 작은 가이드 RNA가 설계되었으며 PAM 서열은 파란색(기울임)으로 표시된다. Transversion 돌연변이는 녹색으로 표시된다. IVT 분석을 확인한 후 샘플을 PCR 증폭하여 NGS를 통해 분석하였다.

[110]

[111] [실시예 2]

[112] 야생형 및 KRAS 돌연변이 (c.35G>C) 서열 및 sgRNA 활성 결과를 포함하는 플라스미드를 사용한 *in vitro* 절단 분석

[113] 야생형 DNA를 특이적으로 절단하기 위해 sgRNA는 PAM 부위와 함께 20 개의 뉴클레오타이드를 표적으로 삼도록 설계되었지만, KRAS 돌연변이 (PAM)의 경우 발암성 돌연변이 (c.35G>C)에 의해 파괴되었다. *In vitro* 절단 분석(IVT)을 사용하여 이 접근법을 테스트하기 위해, 표적 서열을 T-블런트 클로닝 벡터 (야생형 및 돌연변이 KRAS 서열)에 서브 클로닝하고 벡터의 제한 효소로 플라스미드를 선형화했다. 선형화된 생성물은 Cas9 뉴클레아제 및 야생형 서열 특이적 sgRNA로 처리하여 추가로 절단되었다. 야생형 또는 돌연변이 DNA를 0.8 % 겔로 전기영동한 후, 선형화된 단편 (절단되지 않은) 및 절단된 단편 (절단된)을 크기로 구분하고 빨간색 화살표로 표시하였다(도 4a 및 도 4b). 야생형에서 표적화된 DNA는 Cas9/sgRNA에 의해 완전히 절단되지만 돌연변이 표적 서열은 Cas9 및 sgRNA에 의해 절단되지 않았다. 왜냐하면 PAM 부위의 돌연변이와 이 돌연변이 대립 유전자는 Cas9 뉴클레아제에 의해 인식되지 않기 때문이다.

[114]

[115] [실시예 3]

[116] 발암성 유전자의 야생형 또는 돌연변이 서열을 포함하는 플라스미드를 사용한 *in vitro* 절단 분석 및 Cas9 변이체를 사용한 sgRNA 활성 분석

[117] 야생형 표적 DNA를 특이적으로 절단하기 위해 sgRNA는 PAM 부위가 있는 20 개 뉴클레오타이드의 서열을 표적화하도록 설계되었지만 EGFR 돌연변이 (c.2369C> T 또는 p.T790M)의 경우 표적 돌연변이 위치는 sgRNA 결합 부위의 트렁크(trunk) 영역 내에서 발견된다. sgRNA 결합 부위의 스페이서의 트렁크 영역 (12 -19) 내에서 1 염기 불일치(mismatch), Cas9 절단 활성은 약간 또는 중간 수준으로 감소하지만 절단 활성은 유지한다. sgRNA를 표적으로 하는 야생형의 이러한 활성은 또한 무 세포 DNA에서 돌연변이 표적 대립 유전자를 절단할 수 있다. 이 문제를 극복하고 돌연변이 표적 대립 유전자의 절단을 피하기 위해 sgRNA (TmWT sgRNA)의 트렁크 영역에 전환 돌연변이(transversion mutation)을

생성했다. 이 TmWT-sgRNA는 두 돌연변이에 대한 낮은 내성 때문에 돌연변이 대립 유전자에서 Cas9 뉴클레아제에 의해 인식되거나 절단되지 않는다. In vitro 절단 분석을 사용하여 이 접근법을 테스트하기 위해 표적 서열을 T-블러스트 클로닝 벡터 (야생형 및 돌연변이 EGFR 서열)에 서브 클로닝하고 벡터의 제한 효소로 플라스미드를 선형화했다. 선형화된 생성물은 Cas9 뉴클레아제 변이체와 야생형 서열 특이적 sgRNA (WT sgRNA) 및 전환 돌연변이 sgRNA (TmWT sgRNA)의 상이한 조합으로 처리하여 추가로 절단되었다. 야생형 또는 돌연변이 DNA를 0.8 % 겔로 전기영동한 후, 선형화된 단편 (절단되지 않은) 및 절단된 단편 (절단된)을 크기로 구분하고 빨간색 화살표로 표시하였다. 야생형에서 표적화된 DNA는 WT-sgRNA로 Cas9 변이체에 의해 완전히 절단되고 심지어 돌연변이 표적 서열에서도 Cas9 및 WT-sgRNA에 의해 완전히 절단되었다. 그러나 전환 돌연변이에서 sgRNA (TmWT-sgRNA)는 Cas9 변이체에 대해 서로 다른 절단 패턴을 보여주고 있다. 이것은 TmWT-sgRNA가 두 돌연변이에 대해 낮은 내성이며 돌연변이 표적 대립 유전자를 완전히 절단할 수 없음을 나타낸다.

- [118] 상기 전략을 기반으로 야생형 표적 대립 유전자를 절단하고 제거하여 cDNA에서 돌연변이 대립 유전자(EGFR 돌연변이 (c.2369C> T 또는 p.T790M; c.2573 T>G 또는 p.L858R; p.C797S, L718Q 또는 p.(729_761) 결실), MET 돌연변이(c.2888, c.3028 또는 "TAC"Y1003X), BRAF 돌연변이(p.V600E 또는 c.1799T>A) 또는 PIK3CA 돌연변이(c.1624G>A 또는 p.E542K; p.H1047? 또는 c.3140A>?)) 를 풍부하게 할 수 있었다(도 5 ~ 도 11). EGFR p.729-761 결실 및 MET 유전자 돌연변이의 경우 돌연변이 위치가 seed 영역에 있기 때문에 Transversion 돌연변이 sgRNA를 설계하지 않았다.

[119]

[120] [실시예 4]

[121] 다양한 비율의 야생형 및 돌연변이 플라스미드 혼합물을 사용한 EGFR 돌연변이의 CUT-PCR 기반 농축 분석

- [122] 상기 <실시예 3> 과 동일한 방법으로 다양한 타겟 서열 및 가이드 RNA 를 이용하여 분석하였다. 야생형 또는 돌연변이 EGFR (c.2369C> T 또는 p.T790M) 서열을 포함하는 다양한 비율의 플라스미드 혼합물에 대한 CUT-PCR 실험을 수행하였다. 야생형 및 돌연변이 서열을 포함하는 DNA 플라스미드를 다양한 비율로 혼합하고 시험관 내에서 Cas9 변이체 절단을 실시했다. 플라스미드 혼합물을 Cas9 변이체와 함께 Transversion 돌연변이 sgRNA (TmWT-sgRNA)로 처리하여 야생형 DNA를 절단했다. 처리 후 PCR cleanup kit를 사용하여 시료를 정제 하였다. 표적 특이적 절단을 정량화하기 위해 표적 서열의 배수 증가를 처리되지 않은 대조군 제품의 배수 증가와 비교했다. 각 레인에서 처리되지 않은 대조군 제품에 대한 EGFR 표적 영역의 양은 NGS 분석에 의해 계산되었다. EGFR (c.2369C> T 또는 p.T790M) 돌연변이의 경우, 돌연변이 플라스미드가

존재하는 플라스미드 혼합물에 대해 CUT-PCR을 처리한 후 (주황색 또는 분홍색 막대) 또는 하지 않은 (보라색 또는 파란색 막대) 표적화된 깊은 시퀀싱을 수행했다. 원래 0.5 ~ 5 %의 비율로 야생형 플라스미드와 혼합되었다. 생성된 PCR 앰플리콘은 MiSeq로 페어드 엔드 시퀀싱에 적용되었다. NGS 결과는 Cas9 및 eCas9 변이체 모두에서 돌연변이 DNA 단편의 모든 혼합 비율에 대한 야생형 및 돌연변이 서열 배수 변화를 TmWT-sgRNA와 비교하여 분석했다.

[123] 돌연변이 서열은 4.8 ~ 5.4 배로 풍부하였다(도 10, 도 11).

[124]

[125] [실시예 5]

[126] 다양한 비율의 야생형 및 돌연변이 플라스미드 혼합물을 사용한 다중 유전자 돌연변이의 CUT-PCR 기반 농축 분석

[127] 상기 <실시예 3> 과 동일한 방법으로 다양한 타겟 서열 및 가이드 RNA 를 이용하여 분석하였다. KRAS (c.35 G> C) 및 EGFR (c.2369 C> T, c.2390 G> C 및 c.2153 T> A)의 야생형 또는 돌연변이 서열을 포함하는 다양한 비율의 플라스미드 혼합물에 대한 CUT-PCR 실험을 수행하였다. 다중 유전자 돌연변이의 야생형 및 돌연변이 서열을 포함하는 DNA 플라스미드를 다양한 비율로 혼합하고 시험관 내에서 Cas9 변이체 절단을 수행했다. 플라스미드 혼합물은 야생형 DNA를 절단하기 위해 Cas9 변이체가 있는 각 유전자에 해당하는 혼합 sgRNA로 처리되었다. 처리 후 PCR cleanup kit를 사용하여 시료를 정제하였다. 다중 표적 특이적 절단을 정량화하기 위해 표적 서열의 배수 증가를 처리되지 않은 대조군 제품의 배수 증가와 비교했다. 각 레인에서 처리되지 않은 대조군 제품에 대한 각 표적 영역의 양은 NGS 분석에 의해 계산되었다. NGS 결과는 Cas9 및 eCas9 변이체 모두에서 돌연변이 DNA 단편의 모든 혼합 비율에 대한 야생형 및 돌연변이 서열 백분율을 다중 sgRNA와 비교하여 분석했다.

[128] 돌연변이 서열은 0.6 ~ 46.5 % 범위로 풍부했다(도 12 ~ 도 15).

[129]

[130] [실시예 6]

[131] cfDNA 참조 표준 세트를 사용한 다중 돌연변이의 CUT-PCR 기반 강화

[132] sgRNA 활성 및 cfDNA에서 돌연변이 대립 유전자를 강화하기 위한 실험 방법을 확인하거나 증명하기 위해 시판되는 다중 cfDNA 샘플을 참조 표준으로 활용했으며, 3 개의 sgRNA의 활성을 확인했다. 상기 다중 cfDNA 표준 샘플은 8 개의 다른 돌연변이를 가진 4 개의 유전자를 포함하며 각 유전자의 세 가지 비율을 포함하였다(도 16).

[133] sgRNA 활성 및 cfDNA에서 돌연변이 대립 유전자를 풍부하게하는 실험 방법을 확인하거나 증명하였다. 참조 표준 혼합물은 야생형 DNA를 절단하기 위해 SpCas9로 각 유전자에 상응하는 sgRNA로 처리되었다. 처리 후 PCR cleanup kit를 사용하여 시료를 정제하였다. 다중 표적 특이적 절단을 정량화하기 위해 각 표적 서열의 백분율을 처리되지 않은 대조군 제품의 백분율과 비교했다. 각 레인에서

처리되지 않은 대조군 제품에 대한 각 표적 영역의 양은 NGS 분석에 의해 계산되었다. NGS 결과는 SpCas9 및 multi-sgRNA와 다른 비율로 cfDNA의 모든 혼합 비율에 대한 미처리 및 처리된 샘플 비율을 비교하여 분석했으며, 돌연변이 서열은 1.44 ~ 9.39 % 범위로 풍부하였다(도 17).

[134] 또한, 혼합된 sgRNA 를 KRAS, EGFR-T790M 또는 EGFR 729-761bp deletion 샘플에 처리한 후 In vitro Cleavage assay 를 수행한 결과를 하기 [표 2]에 나타냈다.

[135]

[136] [표2]

Cas9 with Mixed-sgRNA's Treated for Multiplex 1	KRAS PCR				EGFR T790M PCR				EGFR Deletion PCR			
	0.1	1	5	100	0.1	1	5	100	0.1	1	5	100
Cas9 Not Treated	0.10	1.71	7.77	0.05	0.34	1.24	5.79	0.17	0.04	0.94	5.53	0.05
Cas9 Treated	2.33	9.91	4.12	0.09	0.05	0.02	7.92	0.05	4.59	0.02	25.96	0.27
Fold Change(mu t Seq Enreached)	22.28	5.81	0.53	1.87	0.14	0.02	1.37	0.32	102.34	0.02	4.69	5.86

[137]

[138] [실시예 7]

[139] 환자 샘플의 무 세포 DNA에서 돌연변이 표적의 초 민감도 검출

[140] 돌연변이를 보유한 폐암 환자의 혈액 내에서 본원발명의 가이드 RNA 를 이용하여 무세포 DNA 내 돌연변이 타겟을 검출하고자 하였다(도 18).

[141] 하기 [표 3] 은 L858R 돌연변이를 보유한 환자 샘플을, [표 4] 는 T790M 및 E19del 돌연변이를 보유한 환자 샘플을 나타낸다.

[142]

[143] [표3]

ID	날짜	성별	구분
LA-29	2015-06-19	남	말기
LA-209	2016-10-04	여	후기
LA-292	2019-04-19	남	후기

[144] [표4]

ID	날짜	성별	구분
LA-232	2017-02-14	남	후기
LA-322	2019-10-07	남	후기
LA-323	2019-10-14	여	후기

[145]

[146] 상기 <실시에 3> 과 동일한 방법으로 다양한 타겟 서열 및 가이드 RNA 를 이용하여 분석하였다. NGS 결과는 환자 cfDNA 각각에 대한 처리되지 않은 샘플 비율과 SpCas9 및 해당 sgRNA를 비교하여 분석되었으며, 돌연변이 시퀀스는 풍부하게 검출되었다(도 19 ~ 도 25).

[147]

산업상 이용가능성

[148] 본 발명의 가이드 RNA 를 이용하여 CRISPR/Cas9 시스템으로 암 돌연변이를 검출하는 경우, 야생형 무 세포 DNA(cfDNA) 만을 절단하고 발암성 돌연변이 DNA 는 절단하지 않아 기존 CRISPR/Cas9 시스템의 가이드 RNA 를 이용하는 경우에 비하여 발암성 돌연변이 DNA 만을 선택적으로 증폭하는데 효과적으로 사용될 수 있다. 특히, 본 발명의 가이드 RNA(야생형 가이드 RNA)와 전환 돌연변이 가이드 RNA(transversion sgRNA)를 병용하는 경우 cfDNA 만을 절단하고 발암성 돌연변이 DNA 만을 선택적으로 증폭하는 효과가 유의적으로 증가할 수 있어 산업상 이용가능성이 있다.

[149]

서열목록 Free Text

[150] 서열번호 1 은 EGFR (p.T790M 또는 c.2369C>T) 돌연변이를 검출할 수 있는 야생형 가이드 RNA(wild sgRNA) 로서, in vitro 절단 분석에 사용되는 경우 서열번호 18 을 타겟 서열로 할 수 있다.

[151] 서열번호 2 는 EGFR (p.T790M 또는 c.2390 G>C) 돌연변이를 검출할 수 있는 전환 돌연변이 가이드 RNA(transversion sgRNA) 로서, in vitro 절단 분석에 사용되는 경우 서열번호 19 를 타겟 서열로 할 수 있다. 서열번호 2 의 뉴클레오타이드 서열을 포함하는 가이드 RNA 를 이용하는 경우 1% 내지 32%

로 돌연변이 서열을 증폭시킬 수 있다.

- [152] 서열번호 3 은 EGFR (p.C797S) 돌연변이를 검출할 수 있는 야생형 가이드 RNA(wild sgRNA) 로서, in vitro 절단 분석에 사용되는 경우 서열번호 20 을 타겟 서열로 할 수 있다.
- [153] 서열번호 4 는 EGFR (p.C797S) 돌연변이를 검출할 수 있는 전환 돌연변이 가이드 RNA(transversion sgRNA) 로서, in vitro 절단 분석에 사용되는 경우 서열번호 21 을 타겟 서열로 할 수 있다. 서열번호 4 의 뉴클레오타이드 서열을 포함하는 가이드 RNA 를 이용하는 경우 1% 내지 13% 로 돌연변이 서열을 증폭시킬 수 있다.
- [154] 서열번호 5 는 MET c.2888 Int13-Exon14 Acceptor (Exon-14 Skipping) 돌연변이를 검출할 수 있는 야생형 가이드 RNA(wild sgRNA) 로서, in vitro 절단 분석에 사용되는 경우 서열번호 22 또는 23 을 타겟 서열로 할 수 있다.
- [155] 서열번호 6 은 MET c.3028 Exon14-Int14 Donor (Exon-14 Skipping) 돌연변이를 검출할 수 있는 야생형 가이드 RNA(wild sgRNA) 로서, in vitro 절단 분석에 사용되는 경우 서열번호 24 또는 25 를 타겟 서열로 할 수 있다.
- [156] 서열번호 7 은 MET "TAC" Y1003X (Exon-14 Skipping) 돌연변이를 검출할 수 있는 야생형 가이드 RNA(wild sgRNA) 로서, in vitro 절단 분석에 사용되는 경우 서열번호 26 또는 27 을 타겟 서열로 할 수 있다.
- [157] 서열번호 8 은 EGFR (L718Q 또는 c.2153 T> A) 돌연변이를 검출할 수 있는 야생형 가이드 RNA(wild sgRNA) 로서, in vitro 절단 분석에 사용되는 경우 서열번호 28 을 타겟 서열로 할 수 있다.
- [158] 서열번호 9 는 EGFR (L718Q) 돌연변이를 검출할 수 있는 전환 돌연변이 가이드 RNA(transversion sgRNA) 로서, in vitro 절단 분석에 사용되는 경우 서열번호 29 를 타겟 서열로 할 수 있다. 서열번호 9 의 뉴클레오타이드 서열을 포함하는 가이드 RNA 를 이용하는 경우 0.8% 내지 23.7% 로 돌연변이 서열을 증폭시킬 수 있다.
- [159] 서열번호 10 은 KRAS (c.35G) 돌연변이를 검출할 수 있는 야생형 가이드 RNA(wild sgRNA) 로서, in vitro 절단 분석에 사용되는 경우 서열번호 30 또는 31 을 타겟 서열로 할 수 있다. 서열번호 10 의 뉴클레오타이드 서열을 포함하는 가이드 RNA 를 이용하는 경우 1.5% 내지 50% 로 돌연변이 서열을 증폭시킬 수 있다.
- [160] 서열번호 11 은 EGFR, 15bp Del p.(729_761) 돌연변이를 검출할 수 있는 야생형 가이드 RNA(wild sgRNA) 로서, in vitro 절단 분석에 사용되는 경우 서열번호 32 또는 33 을 타겟 서열로 할 수 있다. 서열번호 11 의 뉴클레오타이드 서열을 포함하는 가이드 RNA 를 이용하는 경우 0.5% 내지 23% 로 돌연변이 서열을 증폭시킬 수 있다.
- [161] 서열번호 12 는 EGFR (p.L858R 또는 c.2573 T>G) 돌연변이를 검출할 수 있는 야생형 가이드 RNA(wild sgRNA) 로서, in vitro 절단 분석에 사용되는 경우

서열번호 34 를 타겟 서열로 할 수 있다.

- [162] 서열번호 13 은 EGFR (p.L858R) 돌연변이를 검출할 수 있는 전환 돌연변이 가이드 RNA(transversion sgRNA) 로서, in vitro 절단 분석에 사용되는 경우 서열번호 35 를 타겟 서열로 할 수 있다. 서열번호 13 의 뉴클레오타이드 서열을 포함하는 가이드 RNA 를 이용하는 경우 1.7% 내지 63% 로 돌연변이 서열을 증폭시킬 수 있다.
- [163] 서열번호 14 는 PIK3CA (p.E542K 또는 c.1624G>A) 돌연변이를 검출할 수 있는 야생형 가이드 RNA(wild sgRNA) 로서, in vitro 절단 분석에 사용되는 경우 서열번호 36 또는 37 을 타겟 서열로 할 수 있다.
- [164] 서열번호 15 는 PIK3CA (p.E545? 또는 c.1633G>?) 돌연변이를 검출할 수 있는 야생형 가이드 RNA(wild sgRNA) 로서, in vitro 절단 분석에 사용되는 경우 서열번호 38 또는 39 를 타겟 서열로 할 수 있다.
- [165] 서열번호 16 은 PIK3CA (p.H1047? 또는 c.3140A>?) 돌연변이를 검출할 수 있는 야생형 가이드 RNA(wild sgRNA) 로서, in vitro 절단 분석에 사용되는 경우 서열번호 40 또는 41 을 타겟 서열로 할 수 있다.
- [166] 서열번호 17 은 PIK3CA (p.V600E 또는 c.1799T>A) 돌연변이를 검출할 수 있는 야생형 가이드 RNA(wild sgRNA) 로서, in vitro 절단 분석에 사용되는 경우 서열번호 42 또는 43 을 타겟 서열로 할 수 있다.
- [167] 서열번호 18 은 EGFR (p.T790M 또는 c.2369C>T) 돌연변이를 검출할 수 있는 서열번호 1 의 야생형 가이드 RNA(wild sgRNA)의 타겟 서열을 나타낸다.
- [168] 서열번호 19 는 EGFR (p.T790M 또는 c.2390 G>C) 돌연변이를 검출할 수 있는 서열번호 2 의 전환 돌연변이 가이드 RNA(transversion sgRNA)의 타겟 서열을 나타낸다.
- [169] 서열번호 20 은 EGFR (p.C797S) 돌연변이를 검출할 수 있는 서열번호 3 의 야생형 가이드 RNA(wild sgRNA)의 타겟 서열을 나타낸다.
- [170] 서열번호 21 은 EGFR (p.C797S) 돌연변이를 검출할 수 있는 서열번호 4 의 전환 돌연변이 가이드 RNA(transversion sgRNA)의 타겟 서열을 나타낸다.
- [171] 서열번호 22 는 MET c.2888 Int13-Exon14 Acceptor (Exon-14 Skipping) 돌연변이를 검출할 수 있는 서열번호 5 의 야생형 가이드 RNA(wild sgRNA)의 타겟 서열을 나타낸다.
- [172] 서열번호 23 은 MET c.2888 Int13-Exon14 Acceptor (Exon-14 Skipping) 돌연변이를 검출할 수 있는 서열번호 5 의 야생형 가이드 RNA(wild sgRNA)의 타겟 서열을 나타낸다.
- [173] 서열번호 24 는 MET c.3028 Exon14-Int14 Donor (Exon-14 Skipping) 돌연변이를 검출할 수 있는 서열번호 6 의 야생형 가이드 RNA(wild sgRNA)의 타겟 서열을 나타낸다.
- [174] 서열번호 25 는 MET c.3028 Exon14-Int14 Donor (Exon-14 Skipping) 돌연변이를 검출할 수 있는 서열번호 6 의 야생형 가이드 RNA(wild sgRNA)의 타겟 서열을

나타낸다.

- [175] 서열번호 26 은 MET 'TAC' Y1003X (Exon-14 Skipping) 돌연변이를 검출할 수 있는 서열번호 7 의 야생형 가이드 RNA(wild sgRNA)의 타겟 서열을 나타낸다.
- [176] 서열번호 27 은 MET 'TAC' Y1003X (Exon-14 Skipping) 돌연변이를 검출할 수 있는 서열번호 7 의 야생형 가이드 RNA(wild sgRNA)의 타겟 서열을 나타낸다.
- [177] 서열번호 28 은 EGFR (L718Q 또는 c.2153 T> A) 돌연변이를 검출할 수 있는 서열번호 8 의 야생형 가이드 RNA(wild sgRNA)의 타겟 서열을 나타낸다.
- [178] 서열번호 29 는 EGFR (L718Q) 돌연변이를 검출할 수 있는 서열번호 9 의 전환 돌연변이 가이드 RNA(transversion sgRNA)의 타겟 서열을 나타낸다.
- [179] 서열번호 30 은 KRAS (c.35G) 돌연변이를 검출할 수 있는 서열번호 10 의 야생형 가이드 RNA(wild sgRNA)의 타겟 서열을 나타낸다.
- [180] 서열번호 31 은 KRAS (c.35G) 돌연변이를 검출할 수 있는 서열번호 10 의 야생형 가이드 RNA(wild sgRNA)의 타겟 서열을 나타낸다.
- [181] 서열번호 32 는 EGFR, 15bp Del p.(729_761) 돌연변이를 검출할 수 있는 서열번호 11 의 야생형 가이드 RNA(wild sgRNA)의 타겟 서열을 나타낸다.
- [182] 서열번호 33 은 EGFR, 15bp Del p.(729_761) 돌연변이를 검출할 수 있는 서열번호 11 의 야생형 가이드 RNA(wild sgRNA)의 타겟 서열을 나타낸다.
- [183] 서열번호 34 는 EGFR (p.L858R 또는 c.2573 T>G) 돌연변이를 검출할 수 있는 서열번호 12 의 야생형 가이드 RNA(wild sgRNA)의 타겟 서열을 나타낸다.
- [184] 서열번호 35 는 EGFR (p.L858R) 돌연변이를 검출할 수 있는 서열번호 13 의 전환 돌연변이 가이드 RNA(transversion sgRNA)의 타겟 서열을 나타낸다.
- [185] 서열번호 36 은 PIK3CA (p.E542K 또는 c.1624G>A) 돌연변이를 검출할 수 있는 서열번호 14 의 야생형 가이드 RNA(wild sgRNA)의 타겟 서열을 나타낸다.
- [186] 서열번호 37 은 PIK3CA (p.E542K 또는 c.1624G>A) 돌연변이를 검출할 수 있는 서열번호 14 의 야생형 가이드 RNA(wild sgRNA)의 타겟 서열을 나타낸다.
- [187] 서열번호 38 은 PIK3CA (p.E545? 또는 c.1633G>?) 돌연변이를 검출할 수 있는 서열번호 15 의 야생형 가이드 RNA(wild sgRNA)의 타겟 서열을 나타낸다.
- [188] 서열번호 39 는 PIK3CA (p.E545? 또는 c.1633G>?) 돌연변이를 검출할 수 있는 서열번호 15 의 야생형 가이드 RNA(wild sgRNA)의 타겟 서열을 나타낸다.
- [189] 서열번호 40 은 PIK3CA (p.H1047? 또는 c.3140A>?) 돌연변이를 검출할 수 있는 서열번호 16 의 야생형 가이드 RNA(wild sgRNA)의 타겟 서열을 나타낸다.
- [190] 서열번호 41 은 PIK3CA (p.H1047? 또는 c.3140A>?) 돌연변이를 검출할 수 있는 서열번호 16 의 야생형 가이드 RNA(wild sgRNA)의 타겟 서열을 나타낸다.
- [191] 서열번호 42 는 PIK3CA (p.V600E 또는 c.1799T>A) 돌연변이를 검출할 수 있는 서열번호 17 의 야생형 가이드 RNA(wild sgRNA)의 타겟 서열을 나타낸다.
- [192] 서열번호 43 은 PIK3CA (p.V600E 또는 c.1799T>A) 돌연변이를 검출할 수 있는 서열번호 17 의 야생형 가이드 RNA(wild sgRNA)의 타겟 서열을 나타낸다.

청구범위

- [청구항 1] CRISPR/Cas9 시스템에 이용되는 가이드 RNA(guide RNA)로서, 상기 가이드 RNA는 야생형 무 세포 DNA를 절단하는 것을 특징으로 하는 가이드 RNA.
- [청구항 2] 제1항에 있어서, 상기 가이드 RNA는 서열번호 1 내지 17로 이루어진 군에서 선택되는 어느 하나 이상의 뉴클레오타이드 서열을 포함하는 것을 특징으로 하는 가이드 RNA.
- [청구항 3] 제1항에 있어서, 상기 가이드 RNA는 가이드 RNA의 트렁크(trunk) 영역, 세미 시드(semi-seed) 영역 및 시드(seed) 영역으로 이루어진 군에서 선택되는 어느 하나 이상의 영역에서 전환 돌연변이(transversion mutation)가 발생한 것을 특징으로 하는 가이드 RNA.
- [청구항 4] 제3항에 있어서, 상기 가이드 RNA는 서열번호 2, 서열번호 4, 서열번호 9 및 서열번호 13으로 이루어진 군에서 선택되는 어느 하나 이상의 뉴클레오타이드 서열을 포함하는 것을 특징으로 하는 가이드 RNA.
- [청구항 5] 제1항에 있어서, 상기 가이드 RNA는 EGFR, MET, KRAS, PIK3CA 및 BRAF로 이루어진 군에서 선택되는 어느 하나 이상의 암 관련 돌연변이를 검출할 수 있는 것을 특징으로 하는 가이드 RNA.
- [청구항 6] 제5항에 있어서, 상기 EGFR 돌연변이를 검출하는 경우, 상기 가이드 RNA는 서열번호 1 내지 4, 8, 9 및 11 내지 13으로 이루어진 군에서 선택되는 어느 하나 이상의 뉴클레오타이드 서열을 포함하는 것을 특징으로 하는 가이드 RNA.
- [청구항 7] 제5항에 있어서, 상기 MET 돌연변이를 검출하는 경우, 상기 가이드 RNA는 서열번호 5 내지 7로 이루어진 군에서 선택되는 어느 하나 이상의 뉴클레오타이드 서열을 포함하는 것을 특징으로 하는 가이드 RNA.
- [청구항 8] 제5항에 있어서, 상기 KRAS 돌연변이를 검출하는 경우, 상기 가이드 RNA는 서열번호 10의 뉴클레오타이드 서열을 포함하는 것을 특징으로 하는 가이드 RNA.
- [청구항 9] 제5항에 있어서, 상기 PIK3CA 돌연변이를 검출하는 경우, 상기 가이드 RNA는 서열번호 14 내지 16으로 이루어진 군에서 선택되는 어느 하나 이상의 뉴클레오타이드 서열을 포함하는 것을 특징으로 하는 가이드 RNA.
- [청구항 10] 제5항에 있어서, 상기 BRAF 돌연변이를 검출하는 경우, 상기 가이드 RNA는 서열번호 17의 뉴클레오타이드 서열을 포함하는 것을 특징으로 하는 가이드 RNA.
- [청구항 11] 서열번호 1 내지 17로 이루어진 군에서 선택되는 어느 하나 이상의 뉴클레오타이드 서열을 포함하는 CRISPR/Cas9의 가이드 RNA를 포함하는 암 진단 또는 예후 예측용 조성물.

- [청구항 12] 제11항에 있어서, 상기 가이드 RNA 는 가이드 RNA 의 트렁크(trunk) 영역, 세미 시드(semi-seed) 영역 및 시드(seed) 영역으로 이루어진 군에서 선택되는 어느 하나 이상의 영역에서 전환 돌연변이(transversion mutation)가 발생한 것을 특징으로 하는 조성물.
- [청구항 13] 제11항에 있어서, 상기 가이드 RNA 는 EGFR, MET, KRAS, PIK3CA 및 BRAF 로 이루어진 군에서 선택되는 어느 하나 이상의 암 관련 돌연변이를 검출할 수 있는 것을 특징으로 하는 조성물.
- [청구항 14] 제11항에 있어서, 상기 가이드 RNA 는 야생형 무 세포 DNA 를 절단하는 것을 특징으로 하는 조성물.
- [청구항 15] 제11항에 있어서, 상기 암은 폐암, 췌장암, 대장암, 위암, 유방암, 신장암, 간암, 혈액암, 방광암 및 자궁암 으로 이루어진 군에서 선택되는 어느 하나 이상의 암인 것을 특징으로 하는 조성물.
- [청구항 16] 제11항의 조성물을 포함하는 암 진단 또는 예후 예측용 키트.
- [청구항 17] 환자의 시료에 제11항의 조성물을 처리하는 단계를 포함하는 암 진단 또는 예후 예측을 위한 정보제공 방법.