



(19) 대한민국특허청(KR)
 (12) 공개특허공보(A)

(11) 공개번호 10-2017-0036801
 (43) 공개일자 2017년04월03일

(51) 국제특허분류(Int. Cl.)

C12Q 1/68 (2006.01)

(52) CPC특허분류

C12Q 1/6874 (2013.01)

C12Q 2521/301 (2013.01)

(21) 출원번호 10-2017-7007125

(22) 출원일자(국제) 2015년08월19일

심사청구일자 없음

(85) 번역문제출일자 2017년03월15일

(86) 국제출원번호 PCT/US2015/045805

(87) 국제공개번호 WO 2016/028843

국제공개일자 2016년02월25일

(30) 우선권주장

62/039,341 2014년08월19일 미국(US)

(71) 출원인

프레지던트 앤드 펠로우즈 오브 하바드 칼리지

미합중국, 메사추세츠 02138, 캠브리지, 퀸시스트리트17

(72) 별명자

처치, 조지 엠.

미국 02446 매사추세츠주 브루클린 센트 스트리트 218

비그놀트, 프레데릭

미국 01721 매사추세츠주 애쉬랜드 서드베리 로드 103

미르, 칼럼 유.

미국 02138 매사추세츠주 캠브리지 엘름우드 애비뉴 30

(74) 대리인

양영준, 김영

전체 청구항 수 : 총 76 항

(54) 발명의 명칭 **핵산의 프로빙 및 맵핑을 위한 RNA-가이드된 시스템**

(57) 요약

가이드 RNA 및 Cas9 단백질을 사용하여 표적 핵산을 검출, 프로빙, 맵핑 및 지정 서열분석하는 방법이 제공된다. 가이드 RNA/Cas9 복합체의 표적 핵산에의 결합을 검출하는 방법이며, 여기서 가이드 RNA는 프로브에 혼성화할 수 있는 3' 테일 서열을 포함하는 것인 방법이 제공된다. 가이드 RNA/Cas9 복합체의 표적 핵산에의 결합을 검출하는 방법이며, 여기서 복합체는 물리적으로 검출되는 것인 방법이 제공된다.

[대표도]

도 6A

대 표 도

도 6A



(52) CPC특허분류

C12Q 2543/10 (2013.01)

C12Q 2563/107 (2013.01)

C12Q 2565/601 (2013.01)

명세서

청구범위

청구항 1

표적 핵산 서열을, 표적 핵산 서열에 상보적인 부분을 갖는 가이드 RNA 서열 및 Cas9 단백질과 접촉시키는 단계를 포함하고,

여기서 가이드 RNA 및 Cas9 단백질은 표적 핵산 서열에 공동-국재화되어 복합체를 형성하고,

여기서 Cas9 단백질은 표적 핵산의 가닥을 낉킹하는 Cas9 낙카제이고, 여기서 프라이머 연장은 낙으로부터 개시되어 성장 쇄를 형성하고, 이 때 상보적 가닥은 주형으로서의 역할을 하며, 이에 의해 표적 핵산이 서열분석되는 것인, 표적 핵산 서열을 서열분석하는 방법.

청구항 2

제1항에 있어서, 서열분석이 개별 뉴클레오티드의 성장 쇄에의 부가를 검출하는 것을 포함하는 것인 방법.

청구항 3

제2항에 있어서, 뉴클레오티드가 형광 표지된 것인 방법.

청구항 4

제2항에 있어서, 뉴클레오티드 A, C, G, T가 차별적으로 표지되고, 매회 염기 부가 사이클 동안 용액 중에 제공되는 것인 방법.

청구항 5

제2항에 있어서, 뉴클레오티드가 가역성 종결인자이고, 합성에 의해 단계적 서열분석이 수행되는 것인 방법.

청구항 6

제2항에 있어서, 뉴클레오티드가 말단 포스페이트에서 표지되고, 실시간 서열분석이 수행되는 것인 방법.

청구항 7

제1항에 있어서, 표적 핵산 서열이 선형 스트링으로서 분석되는 것인 방법.

청구항 8

제1항에 있어서, 표적 핵산 서열이 실질적으로 신장되는 선형 스트링으로서 분석되는 것인 방법.

청구항 9

제1항에 있어서, 낙으로부터의 서열분석 개시가 표적 핵산 선형 스트링 상의 다중 부위로부터 개시되는 것인 방법.

청구항 10

제1항에 있어서, 각 핵산 및 다중 표적 핵산 상의 복수 개의 부위가 동시에 분석되는 것인 방법.

청구항 11

제1항에 있어서, 계놈의 하나 이상의 영역이 서열분석되는 것인 방법.

청구항 12

제1항에 있어서, gRNA/Cas9 복합체가 낉킹 이후에 제거되는 것인 방법.

청구항 13

제1항에 있어서, 핵산이 세포 계내에서 분석되는 것인 방법.

청구항 14

제1항에 있어서, 핵산이 세포 계내에서 분석되고, 세포가 분석 이전에 고정되는 것인 방법.

청구항 15

제1항에 있어서, RNA 문자가 분석 이전에 제거되는 것인 방법.

청구항 16

표적 핵산 서열을, 표적 핵산 서열에 상보적인 부분을 갖는 가이드 RNA 서열 및 Cas9 단백질과 접촉시키는 단계이며,

여기서 가이드 RNA 및 Cas9 단백질은 표적 핵산 서열에 공동-국재화되어 복합체를 형성하고,

여기서 가이드 RNA는 3' 테일 핵산 서열을 포함하는 것인 단계,

검출가능한 프로브 서열을 3' 테일 서열에 혼성화시키는 단계, 및

검출가능한 프로브 서열을 검출하여 표적 핵산 서열을 검출하는 단계

를 포함하는, 표적 핵산 서열을 검출하는 방법.

청구항 17

제16항에 있어서, 3' 테일 서열이 프라이머로서 작용할 수 있는 것인 방법.

청구항 18

제16항에 있어서, 3' 테일 서열이 gRNA 결합 위치에 인접해 있는 서열에 상보적인 서열을 포함하는 것인 방법.

청구항 19

제16항에 있어서, 3' 테일 서열이 DNA PAINT를 위한 도킹 부위 또는 핸들을 포함하는 것인 방법.

청구항 20

표적 핵산 서열을, 표적 핵산 서열에 상보적인 부분을 갖는 가이드 RNA 서열 및 Cas9 단백질과 접촉시키는 단계를 포함하고,

여기서 가이드 RNA 및 Cas9 단백질은 표적 핵산 서열에 공동-국재화되어 복합체를 형성하고,

여기서 표적 핵산은 선형 스트링으로서 분석되는 것인, 표적 핵산 서열을 검출하는 방법.

청구항 21

제20항에 있어서, 선형 스트링이 표면 상에서, 유동 스트림에서, 또는 마이크로 또는 나노채널에서 신장되는 것인 방법.

청구항 22

제20항에 있어서, 선형 스트링이 한쪽 단부는 표면에의 부착을 통해 신장되고, 나머지 다른 한쪽 단부는 유동 스트림에 매달려 있으면서, 광학 또는 자기 트위저를 포함하는 당김력에의 물리화학적 부착을 통해 신장되는 것인 방법.

청구항 23

제20항에 있어서, 핵산을 따라 하나 이상의 복합체의 결합 위치가 검출되는 것인 방법.

청구항 24

제20항에 있어서, 다중 복합체의 결합 위치를 통해 단일 문자 맵이 구축될 수 있는 것인 방법.

청구항 25

제20항에 있어서, gRNA가 반복 DNA를 표적화하고, 하나 이상의 반복 단위의 위치가 검출되는 것인 방법.

청구항 26

제20항에 있어서, gRNA가 DNA 상의 단일 카페 서열을 표적화하는 것인 방법.

청구항 27

제20항에 있어서, 다중 gRNA가 단일 게놈 유전자좌를 표적화하는 것인 방법.

청구항 28

없음

청구항 29

제20항에 있어서, 다중의 단일 카페 유전자좌가 각 유전자좌에 특이적인 gRNA에 의해 표적화되는 것인 방법.

청구항 30

제20항에 있어서, 한 유전자좌에의 결합이 또 다른 유전자좌에의 결합으로부터 구별되는 것인 방법.

청구항 31

제20항에 있어서, 다중 gRNA가 각 유전자좌에 결합하는 것인 방법.

청구항 32

제20항에 있어서, 선형 스트링이 염색질 섭유인 방법.

청구항 33

제20항에 있어서, 선형 스트링이 염색체 내에서 풀딩되는 것인 방법.

청구항 34

제20항에 있어서, 선형 스트링이 염색체 내에서 풀딩되고, 염색체가 중기 또는 유사분열 염색체인 방법.

청구항 35

표적 핵산 서열을, 표적 핵산 서열에 상보적인 부분을 갖는 가이드 RNA 서열 및 Cas9 단백질과 접촉시키는 단계를 포함하고,

여기서 가이드 RNA 및 Cas9 단백질은 표적 핵산 서열에 공동-국제화되어 복합체를 형성하고,

여기서 표적 핵산은 특성을 변경시키는 방식으로 나노포어 또는 나노캡을 통해 또는 그를 통과하여 획단하고,

여기서 복합체의 결합은 변경된 특성 검출에 의해 검출되며, 이에 의해 표적 핵산 서열이 검출되는 것인, 표적 핵산 서열을 검출하는 방법.

청구항 36

제35항에 있어서, 핵산을 따라 하나 이상의 복합체의 결합의 위치가 이온 전류, 전자 터널링 또는 광학 검출을 포함하는 방법에 의해 검출되는 것인 방법.

청구항 37

제35항에 있어서, 가이드 중의 인식 서열이 비교적 짧고, 가이드 중의 나머지 부분이 축중성 또는 범용 엔기를 포함하며, 이에 의해 표적 핵산을 따라 다수의 부위에 결합될 수 있는 것인 방법.

청구항 38

제35항에 있어서, 다중 복합체의 결합 위치를 통해 단일 문자 맵이 구축될 수 있는 것인 방법.

청구항 39

표적 핵산 서열을, 표적 핵산 서열에 상보적인 부분을 갖는 가이드 RNA 서열 및 Cas9 단백질과 접촉시키는 단계를 포함하고,

여기서 가이드 RNA 및 Cas9 단백질은 표적 핵산 서열에 공동-국재화되어 복합체를 형성하고,

여기서 복합체가 검출되며, 이에 의해 표적 핵산 서열이 검출되는 것인, 표적 핵산 서열을 검출하는 방법.

청구항 40

제39항에 있어서, 가이드 RNA가 검출가능한 표지를 포함하는 것인 방법.

청구항 41

제39항에 있어서, Cas9 단백질이 검출가능한 표지를 포함하는 것인 방법.

청구항 42

제39항에 있어서, 복합체가 검출가능한 표지를 포함하는 것인 방법.

청구항 43

제39항에 있어서, 복합체가 나노포어에 의해 검출되는 것인 방법.

청구항 44

제39항에 있어서, 복합체가 전자 현미경법에 의해 검출되는 것인 방법.

청구항 45

제39항에 있어서, 복합체가 스캐닝 프로브 현미경법에 의해 검출되는 것인 방법.

청구항 46

제39항에 있어서, 복합체가 캔틸레버에 의해 검출되는 것인 방법.

청구항 47

제39항에 있어서, 복합체가 수정 진동자 저울에 의해 검출되는 것인 방법.

청구항 48

제39항에 있어서, 복합체가 전계 효과 트랜지스터에 의해 검출되는 것인 방법.

청구항 49

제39항에 있어서, 가이드 RNA가 프로브 서열에 상보적인 3' 테일 서열을 포함하는 것인 방법.

청구항 50

제39항에 있어서, 가이드 RNA가 검출가능한 표지를 포함하는 프로브 서열에 상보적인 3' 테일 서열을 포함하고, 프로브 서열이 3' 테일 서열에 결합하는 것인 방법.

청구항 51

제39항에 있어서, 가이드 RNA가 복수 개의 검출가능한 표지를 포함하는 프로브 서열에 상보적인 3' 테일 서열을 포함하고, 프로브 서열이 3' 테일 서열에 결합하는 것인 방법.

청구항 52

제39항에 있어서, 가이드 RNA가 검출가능한 표지를 포함하는 프로브 서열에 상보적인 3' 테일 서열을 포함하고, 프로브 서열이 3' 테일 서열에 결합하고, 여기서 프로브 서열은 증폭되는 것인 방법.

청구항 53

제39항에 있어서, 가이드 RNA가 프로브 또는 검출가능한 표지에의 결합 쌍으로서 3' 테일 서열을 포함하는 것인 방법.

청구항 54

제39항에 있어서, 표적 핵산이 이중 가닥 게놈 DNA인 방법.

청구항 55

제39항에 있어서, 표적 핵산이 염색체 DNA인 방법.

청구항 56

제39항에 있어서, 표적 핵산이 기관상에서 연장되는 것인 방법.

청구항 57

제39항에 있어서, 표적 핵산이 평면 표면 상에서 연장되는 것인 방법.

청구항 58

제39항에 있어서, 표적 핵산이 포어 내에서 연장되는 것인 방법.

청구항 59

제39항에 있어서, 표적 핵산이 채널 내에서 연장되는 것인 방법.

청구항 60

제39항에 있어서, Cas9 단백질이 야생형 Cas9, cas9 낙카제 또는 뉴클레아제 널 Cas9인 방법.

청구항 61

제39항에 있어서, 검출가능한 표지가 직접 또는 간접적으로 Cas9 단백질에 결합되는 것인 방법.

청구항 62

제39항에 있어서, 검출가능한 표지가 직접 또는 간접적으로 가이드 RNA에 결합되는 것인 방법.

청구항 63

제39항에 있어서, 검출가능한 표지가 직접 또는 간접적으로 복합체에 결합되는 것인 방법.

청구항 64

제39항에 있어서, Cas9 단백질이 표적 핵산의 가닥을 낙킹하는 Cas9 낙카제이고, 여기서 프라이머 연장은 낙으로부터 개시되고, 이 때 상보적 가닥은 주형으로서의 역할을 하며, 이에 의해 표적 핵산이 서열분석되는 것인 방법.

청구항 65

제39항에 있어서, Cas9 단백질이 표적 핵산의 가닥을 낙킹하는 Cas9 낙카제이고, 여기서 프라이머 연장은 낙으로부터 개시되어 검출가능한 표지를 포함하고, 이 때 상보적 가닥은 주형으로서의 역할을 하며, 이에 의해 표적 핵산이 검출되는 것인 방법.

청구항 66

제39항에 있어서, 가이드 RNA 및 Cas9 단백질을 조합하고, 이어서 표적 핵산과 접촉시키는 것인 방법.

청구항 67

제39항에 있어서, 가이드 RNA 및 Cas9 단백질을 조합하고, 이어서 샘플 내의 표적 핵산과 접촉시키는 것인 방법.

청구항 68

제39항에 있어서, 가이드 RNA가 시드 영역 서열을 포함하는 것인 방법.

청구항 69

제39항에 있어서, 가이드 RNA가 가이드 RNA의 비-시드 영역에 축중성 위치 또는 서열 또는 범용 엔기를 포함하는 것인 방법.

청구항 70

제39항에 있어서, 표적 핵산 서열을, 각각이 표적 핵산 서열에 상보적인 부분을 갖는 것인 복수 개의 가이드 RNA 서열과 접촉시키는 단계를 포함하는 방법.

청구항 71

제39항에 있어서, 표적 핵산이 용액 샘플 내에 존재하는 것인 방법.

청구항 72

제39항에 있어서, 표적 핵산이 기관상에 존재하는 것인 방법.

청구항 73

제39항에 있어서, 표적 핵산이 기관에 결합되어 있는 것인 방법.

청구항 74

제39항에 있어서, 표적 핵산이 세포 내에 존재하는 것인 방법.

청구항 75

표적 핵산 서열을, 표적 핵산 서열에 상보적인 부분을 갖고 3' 프라이머 연장부를 갖는 가이드 RNA 서열 및 Cas9 단백질과 접촉시키는 단계이며,

여기서 가이드 RNA 및 Cas9 단백질은 표적 핵산 서열에 공동-국제화되어 복합체를 형성하는 것인 단계,

주형을 따라 프라이머를 연장시켜 연장 생성물 내로 하나 이상의 검출가능한 표지를 도입하는 단계, 및

하나 이상의 검출가능한 표지를 검출함으로써 표적 핵산 서열을 검출하는 단계

를 포함하는, 표적 핵산 서열을 검출하는 방법.

청구항 76

표적 핵산 서열을, 표적 핵산 서열에 상보적인 부분을 갖고, 3' 프라이머 연장부를 갖는 가이드 RNA 서열 및 Cas9 단백질과 접촉시키는 단계이며,

여기서 가이드 RNA 및 Cas9 단백질은 표적 핵산 서열에 공동-국제화되어 복합체를 형성하는 것인 단계,

롤링 서클 중복 주형을 따라 프라이머를 연장시켜 롤링 서클 콘카테머 내로 복수 개의 검출가능한 표지를 도입하는 단계, 및

복수 개의 검출가능한 표지를 검출함으로써 표적 핵산 서열을 검출하는 단계

를 포함하는, 표적 핵산 서열을 검출하는 방법.

발명의 설명

기술 분야

[0001] 관련 출원 데이터

[0002] 본 출원은 2014년 8월 19일에 출원된 미국 특허 가출원 번호 62/039,341을 우선권 주장하며, 이로써 상기 문헌은 모든 목적을 위해 그 전문이 본원에 참조로 포함된다.

배경 기술

[0003] 박테리아성 및 고세균성 CRISPR-Cas 시스템은 침입 외래 핵산 내에 존재하는 상보적 서열의 분해를 지시하는 Cas 단백질과 복합체를 형성하는 짧은 가이드 RNA에 의존한다. 문헌 [Deltcheva, E. et al. CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III. *Nature* 471, 602-607 (2011)]; [Gasiusas, G., Barrangou, R., Horvath, P. & Siksnys, V. Cas9-crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria. *Proceedings of the National Academy of Sciences of the United States of America* 109, E2579-2586 (2012)]; [Jinek, M. et al. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* 337, 816-821 (2012)]; [Sapranauskas, R. et al. The Streptococcus thermophilus CRISPR/Cas system provides immunity in *Escherichia coli*. *Nucleic acids research* 39, 9275-9282 (2011)]; 및 [Bhaya, D., Davison, M. & Barrangou, R. CRISPR-Cas systems in bacteria and archaea: versatile small RNAs for adaptive defense and regulation. *Annual review of genetics* 45, 273-297 (2011)]을 참조할 수 있다. 최근, S. 피오게네스(*S. pyogenes*) 유형 II CRISPR 시스템을 시험관내에서 재구성한 결과, 정상적으로 트랜스-코딩된 tracrRNA ("트랜스-활성화 CRISPR RNA")에 융합된 crRNA ("CRISPR RNA")가, crRNA와 매칭되는 표적 DNA 서열을 서열-특이적으로 절단하도록 Cas9 단백질을 지시하는데 충분한 것으로 입증되었다. 표적 부위에 상동성인 gRNA의 발현은 Cas9 동원 및 표적 DNA의 분해를 유도한다. 문헌 [H. Deveau et al., Phage response to CRISPR-encoded resistance in *Streptococcus thermophilus*. *Journal of Bacteriology* 190, 1390 (Feb, 2008)]을 참조할 수 있다. CRISPR/Cas9 시스템의 다양한 용도가 공지되어 있다. WO2014/099744, WO2013176772, US 8,697,359 및 문헌 [Sternberg et al., *Nature*, Vol. 507, pp. 62-67 (2014)]를 참조할 수 있다.

발명의 내용

[0004] 본 개시내용의 측면은 표적 핵산 서열을, 표적 핵산 서열에 상보적인 부분을 갖는 가이드 RNA 서열 및 Cas9 단백질과 접촉시키는 단계를 포함하고, 여기서 가이드 RNA 및 Cas9 단백질은 표적 핵산 서열에 국재화되어 복합체를 형성하고, 여기서 복합체가 검출되며, 이에 의해 표적 핵산 서열이 검출되는 것인, 표적 핵산 서열을 검출하는 방법에 관한 것이다. 한 측면에 따라, 본 방법은 생체외에서, 즉 시험관내에서, 예컨대 배울 또는 기관 상에서 수행된다. 한 측면에 따라, 가이드 RNA 및 Cas9 단백질은 본 개시내용의 시험관내 방법에서 시약으로서 사용되도록 제조 및 단리된다. 본 방법의 측면은 핵산, 예컨대 DNA의 프로빙, 예컨대 분석적 프로빙 또는 분취용 프로빙, 검출, 표지, 맵핑 및 서열분석을 포함한다. 예를 들어, 본 개시내용은 DNA의 존재를 확인하기 위한 목적으로 예컨대, 단일 분자 수준에서 DNA를 프로빙하거나, DNA를 친화도 정제하기 위한 목적으로 DNA를 프로빙하거나, DNA를 따라 중요한 특정 영역을 마킹하거나, 또는 서열분석 출발 부위를 생성하기 위해 DNA를 맵핑하는 방법에 관한 것이다.

[0005] 본원에 기술된 방법에 따라, 가이드 RNA, DNA 결합 단백질, 예컨대 Cas9 단백질, 및 이중 가닥 DNA 표적 서열을 포함하는 복합체가 형성된다. 특정 측면에 따라, 본 개시내용의 범주 내의 DNA 결합 단백질은 가이드 RNA와 복합체를 형성하는 단백질을 포함하며, 여기서 가이드 RNA는 상기 복합체를 이중 가닥 DNA 서열로 가이드하고, 여기서 상기 복합체는 DNA 서열에 결합한다. 본 개시내용의 이러한 측면은 이중 가닥 DNA로의 또는 그와의 RNA 및 DNA 결합 단백질의 공동-국재화로서 지칭될 수 있다. 이러한 방식으로, DNA 결합 단백질-가이드 RNA 복합체는 특정의 표적 DNA 서열에서 검출가능한 복합체를 형성함으로써 표적 DNA 서열의 존재를 검출하는데 사용될 수 있다. 특정 측면에 따라, 복합체는 검출가능한 표지의 존재에 기인하여 검출될 수 있다. 특정 측면에 따라, 복합체는 직접적으로 표지될 수 있거나, 또는 간접적으로 표지될 수 있다. 특정 측면에 따라, 검출가능한 표지는 가이드 RNA, Cas9 단백질 또는 복합체 상에 존재할 수 있다.

[0006] 특정 측면에 따라, 가이드 RNA에 대한 공동-국재화 인자는 DNA-결합 단백질이 아닐 수 있다. 표적 핵산 서열에서 가이드 RNA와 함께 공동-국재화하는데 시약이 사용될 수 있다. 특정 측면에 따라, 본 개시내용의 특정 측면에서 유용한 DNA 결합 단백질의 존재를 가이드 RNA가 반드시 필요로 하는 것은 아니다. DNA 결합 단백질은 존

재하지 않을 수도 있다. 예를 들어, 가이드 RNA 그 자체가 표적 핵산 서열에 결합할 수 있고, 가이드 RNA는 표적 핵산 서열에 또는 그 부근에 표지 또는 다른 기능성 모이어티를 국제화하기 위해 표지 또는 다른 기능성 모이어티를 그에 부착시킬 수 있다.

[0007] 특정 측면에 따라, 복합체는 검출가능한 표지를 갖지 않고, 복합체의 구조를 검출함으로써 검출될 수 있다. 형광 또는 다른 시각적으로 또는 분광학적으로 검출가능한 모이어티를 시각화하는 것과 달리, 복합체의 물리적 구조는 프로빙된다. 특정 측면에 따라, 복합체는 검출가능한 표지를 갖지 않고, 복합체의 물리 화학적 특성, 예컨대 정전하를 검출함으로써 검출될 수 있다. 상기 방법은, 모두가 관련 기술분야의 통상의 기술자에게 공지되어 있는 것인, 나노포어 검출 방법, 전자 현미경법, 광학 현미경법, 스캐닝 프로브 현미경법, 원자간력 현미경법, 캔틸레버 검출 방법, 수정 검출 방법, 전계 효과 트랜ジ스터 검출 방법을 사용하여 복합체를 검출하는 것을 포함한다. 통상의 기술자는 본 개시내용에 기반한 복합체의 구조를 검출할 수 있는 다른 방법도 쉽게 구상할 수 있을 것이다.

[0008] 특정 측면에 따라, CRISPR Cas9 시스템과 관련하여 "가이드 RNA"라는 용어는 관련 기술분야의 통상의 기술자에게 공지되어 있으며, 표적 핵산에 상보적인 부분, 예컨대 20개의 뉴클레오티드로 이루어진 부분을 포함한다. 가이드 RNA를 디자인하는 방법은 관련 기술분야의 통상의 기술자에게 널리 공지되어 있다. 본원에 기술된 방법은 각각이 표적 핵산 서열에 상보적인 부분을 갖는 것인 복수 개의 가이드 RNA 서열과 표적 핵산 서열을 접촉시키는 단계를 포함한다. 본원에 기술된 방법은 각각이 상응하는 표적 핵산 서열에 상보적인 부분을 갖는 것인 복수 개의 상응하는 가이드 RNA 서열과 복수 개의 표적 핵산 서열을 접촉시키는 단계를 포함한다.

[0009] 특정 측면에 따라, 본 개시내용에 따른 가이드 RNA는 표적 핵산에 상보적인 부분, 및 프로브 서열 또는 검출가능한 표지에 상보적이거나, 또는 그려할 수 있거나, 또는 다르게는 그에 결합하는 3' 테일 부분 또는 서열을 포함한다. 한 측면에 따라, 3' 테일 부분이 특이적인 관능기를 제공한다. 3' 테일 부분은 동일 및 다중의 관능기를 위해 모듈식일 수 있고, 다중의 요소를 포함할 수 있다. 한 측면에 따라, 3' 테일 부분은 하나 이상의 또는 다중의 프로브 서열(들) 또는 검출가능한 표지(들)에 상보적일 수 있거나, 또는 다르게는 (예컨대, 압타미메커니즘을 통해) 그에 결합할 수 있다. 각 프로브 서열 또는 검출가능한 표지는 상이한 역할을 할 수 있다 (예를 들어, 한 서열의 역할은 CY3 표지된 올리고뉴클레오티드에 결합하는 것일 수 있고, 제2 서열의 역할은 Cy5 표지된 올리고뉴클레오티드에 결합하는 것일 수 있다). 예를 들어, 테일 서열은 기능성 단백질을 표적 핵산 서열에 국제화하는데 사용될 수 있다. 예를 들어, 테일 서열은 가이드 RNA에 의해 치환된 표적 이중체의 일부에 결합할 수 있다.

[0010] 한 측면에 따라, 프로브 서열은 검출가능한 표지를 포함하고, 프로브 서열은 3' 테일 서열에 결합한다. 한 측면에 따라, 프로브 서열은 복수 개의 검출가능한 표지를 포함하고, 프로브 서열은 3' 테일 서열에 결합한다. 한 측면에 따라, 프로브 서열은 검출가능한 표지를 포함하고, 프로브 서열은 3' 테일 서열에 결합하고, 여기서 프로브 서열은 증폭된다. 한 측면에 따라, 프로브 서열은 3' 테일 서열을 포함/그에 결합하고, 여기서 프로브 서열은 증폭된다. 한 측면에 따라, 가이드 RNA는 프로브 또는 검출가능한 표지와의 결합 결합 쌍으로서 3' 테일 서열을 포함한다. 한 측면에 따라, 테일 서열은 주형 서열에 결합하였을 때, 프라이머로서의 역할을 할 수 있다. 이어서, 테일 서열을 연장시켜 하나 이상의 검출가능한 표지, 예컨대 형광 뉴클레오티드, 또는 하나 이상의 표지가 직접 또는 간접적으로 결합할 수 있는 하나 이상의 결합 모이어티, 예컨대 비오틴 또는 dig 표지된 뉴클레오티드를 도입한다. 한 측면에 따라, 테일 프라이머 서열 및 롤링 서클 증폭 주형을 이용하여 롤링 서클 증폭을 사용함으로써 복수 개의 검출가능한 모이어티, 또는 검출가능한 모이어티가 부착될 수 있는 결합 모이어티를 갖는 롤링 서클 콘카테머 생성물을 생성할 수 있다. 이러한 방식으로, 롤링 서클 증폭을 사용하여 신호 강도를 증폭시킬 수 있다. 롤링 서클 증폭 방법은 관련 기술분야의 통상의 기술자에게 공지되어 있고, 문헌 [Drmanac et al., Human genome sequencing using unchained base reads on self-assembling DNA nanoarrays, Science, vol. 327, p. 78-81 (2009)]를 포함한다.

[0011] 한 측면에 따라, 표적 핵산은 이중 가닥 핵산이다. 한 측면에 따라, 표적 핵산은 이중 가닥 게놈 DNA이다. 한 측면에 따라, 표적 핵산은 염색체 DNA이다.

[0012] 특정 측면에 따라, 가이드 RNA는 시드 영역 서열을 포함한다. 특정 측면에 따라, 가이드 RNA는 가이드 RNA의 비-시드 영역에 축중성 위치 또는 서열 또는 범용 염기를 포함한다.

[0013] 한 측면에 따라, 표적 핵산은 기판, 예컨대 평면 기판 또는 포어 또는 채널 상에서 신장되고, 즉 표적 핵산은 포어 또는 채널 내에서 신장된다.

- [0014] 한 측면에 따라, Cas9 단백질은 관련 기술분야의 통상의 기술자에게 공지된 바와 같이, 야생형 Cas9, Cas9 낙카제 또는 뉴클레아제 널 Cas9이다. 야생형 Cas9를 단리시키는 방법은 관련 기술분야의 통상의 기술자에게 공지되어 있다. Cas9 낙카제를 제조하는 방법은 관련 기술분야의 통상의 기술자에게 공지되어 있다. 뉴클레아제 널 Cas9를 제조하는 방법은 관련 기술분야의 통상의 기술자에게 공지되어 있다.
- [0015] 한 측면에 따라, 검출가능한 표지는 Cas9 단백질에 직접 또는 간접적으로 결합한다. 한 측면에 따라, 검출가능한 표지는 가이드 RNA에 직접 또는 간접적으로 결합한다. 한 측면에 따라, 검출가능한 표지는 가이드 RNA의 일부이다 (예컨대, (예컨대, 시험관내 전사에 의해) 가이드 RNA를 제조하는 동안 형광 표지된 뉴클레오티드가 도입된 경우). 한 측면에 따라, 검출가능한 표지는 복합체에 직접 또는 간접적으로 결합한다.
- [0016] 한 측면에 따라, 서열분석 방법에 의해 표적 핵산의 서열을 측정하는 것인 방법을 제공한다. 한 측면에 따라, Cas9 단백질은 표적 핵산 가닥을 낙킹하는 Cas9 낙카제이고, 여기서 프라이머 연장 또는 쇄 연장은 상기 낙으로부터 개시되고, 이 때 상보적 가닥은 주형으로서의 역할을 하며, 이에 의해 표적 핵산, 예컨대 표적 핵산 가닥 중 하나가 서열분석된다. 다른 낙킹 접근법, 예컨대 낙킹 엔도뉴클레아제 및 DNase 1 사용에 비하여 우수한 이점은 낙이 생성되는 위치가 가이드 RNA 사용을 통해 프로그램 가능하고, 다중의 특이적 위치가 표적화될 수 있다는 점이다. 컴퓨터 실행 방법 및 소프트웨어를 사용하여 합성을 위해 관심 게놈의 일부를 확인할 수 있고, gRNA 제조에 요구되는 DNA 주형을 주문할 수 있고, 가이드 RNA를 시험관내에서 전사시킬 수 있고, 이어서 원하는 위치에서 낙킹을 실행함으로써 표적화된 서열분석을 수행할 수 있다. 주형을 따라 진행되는 프라이머 연장에 의한 서열분석 방법은 관련 기술분야의 통상의 기술자에게 공지되어 있다. 프라이머로서 낙을 사용하는 것은 관련 기술분야의 통상의 기술자에게 공지되어 있다. 본 개시내용의 상기 특징은 예컨대, 표적 핵산으로부터 서열 정보를 수득하는 것 대신, 또는 그 이외에도 또한 연장 생성물에 검출가능한 표지를 포함시키며, 이에 의해 표적 핵산을 검출하는데 사용될 수 있다. 따라서, Cas9 단백질은 표적 핵산의 가닥을 낙킹하는 Cas9 낙카제이고, 여기서 프라이머 연장은 낙으로부터 개시되어 검출가능한 표지를 포함하고, 이 때 상보적 가닥은 주형으로서의 역할을 하며, 이에 의해 표적 핵산이 검출된다. 본 측면에 따라, 일단 표지가 연장 생성물 내로 도입되고 나면, 연장 생성물 내로 도입된 표지가 검출되기 때문에, gRNA/Cas9 복합체는 표적 DNA와 함께 그대로 유지될 필요는 없다.
- [0017] 한 측면에 따라, 단리된 가이드 RNA 및 단리된 Cas9 단백질을 적합한 조건하에서 조합한 후, 시험관내 반응 또는 복합체 형성 매질 중에서 표적 핵산과 접촉시킨다. 한 측면에 따라, 표적 핵산은 샘플, 예컨대 핵산 샘플 내에 존재한다. 핵산 샘플은 복수 개의 핵산을 포함할 수 있고, 이는 핵산의 복합 혼합물로서 지칭될 수 있다. 특정 측면에 따라, 표적 핵산에 특이적인 가이드 RNA를 이용하여 핵산의 복합 혼합물 내의 표적 핵산을 확인하는 방법을 제공한다. 한 측면에 따라, 하나 이상의 또는 복수 개의 표적 핵산에 특이적인 가이드 RNA를 이용하여 핵산의 복합 혼합물 내의 하나 이상의 또는 복수 개의 표적 핵산을 확인하는 방법을 제공한다. 이러한 측면에서, 복수 개의 표적 핵산을 검출하는 다중 방법을 제공한다. 복수 개의 표적 핵산은 각각 상응하는 가이드 RNA/Cas 9 단백질 복합체에 의해 결합될 수 있으며, 이에 의해 본원에 기술된 바와 같이, 또는 관련 기술분야에 공지된 바와 같이 검출 또는 서열분석될 수 있다.
- [0018] 특정 측면에 따라, 표적 핵산에 특이적인 가이드 RNA를 이용하여 핵산의 복합 혼합물 내의 표적 핵산을 친화도 정제하는 방법을 제공한다. 한 측면에 따라, 하나 이상의 또는 복수 개의 표적 핵산에 특이적인 가이드 RNA를 이용하여 핵산의 복합 혼합물 내의 하나 이상의 또는 복수 개의 표적 핵산을 친화도 정제하는 방법을 제공한다. 이러한 측면에서, 복수 개의 표적 핵산을 친화도 정제하는 다중 방법을 제공한다. 복수 개의 표적 핵산은 각각 상응하는 가이드 RNA/Cas 9 단백질 복합체에 의해 결합될 수 있다. 상기 측면에 따라, 관련 기술분야의 통상의 기술자에게 공지된 결합 쌍을 사용하는 친화도 시스템이 사용될 수 있다. 표적 핵산은 복합 혼합물 내의 다른 핵산을 고갈시킴으로써 정제된다. 표적 핵산은 복합 혼합물로부터 다른 핵산을 고갈시킴으로써 정제된다. 고갈은 고갈시키고자 하는 표적의 친화도 포획에 의해 이루어질 수 있다. 대안적으로, 고갈은 고갈시키고자 하는 표적의 절단에 의해 이루어질 수 있다. 일부 경우에서, 복합 혼합물 내에 높은 존재비로 존재하는 핵산은 존재비가 더 낮은 표적을 분석하기 위해서는 혼합물로부터 고갈될 필요가 있을 수 있다. 예를 들어, 반복 DNA 또는 다른 고농도의 핵산을 샘플로부터 고갈시키는 것이 바람직할 수 있다.
- [0019] 특정 측면에 따라, 고갈을 위해 표적화된 핵산에 특이적인 가이드 RNA를 이용하여 핵산의 복합 혼합물 내의 표적 핵산을 고갈시키는 방법을 제공한다. 한 측면에 따라, 하나 이상의 또는 복수 개의 표적 핵산에 특이적인 가이드 RNA를 이용하는 핵산의 복합 혼합물 내의 하나 이상의 또는 복수 개의 표적 핵산에 대한 방법을 제공한다. 이러한 측면에서, 복수 개의 표적 핵산을 고갈시키는 다중 방법을 제공한다. 고갈을 위해 표적화된 복수

개의 핵산은 각각 상응하는 가이드 RNA/Cas 9 단백질 복합체에 의해 결합될 수 있다.

[0020] 특정 측면에 따라, 표적 핵산은 샘플 용액 내에 존재한다. 특정 측면에 따라, 표적 핵산은 기관 상에 존재한다. 특정 측면에 따라, 표적 핵산은 기관에 결합되어 있다. 특정 측면에 따라, 표적 핵산은 세포 내에 있다. 한 측면에 따라, 세포는 진핵 세포이다. 한 측면에 따라, 세포는 효모 세포, 식물 세포 또는 동물 세포이다. 한 측면에 따라, 세포는 포유동물 세포이다. 특정 실시양태에서, 포유동물 세포는 살아있는 세포이고, 가이드 RNA 또는 DNA 결합 단백질, 예컨대 Cas9 또는 다른 DNA 결합 단백질은 전기천공, 운반체 매개 전달 (예컨대, 리포펙틴), 미세주입 및 관련 기술분야의 통상의 기술자에게 공지된 다른 방법에 의해 전달된다. 특정 실시양태에서, 포유동물 세포는 가이드 RNA 및 DNA 결합 단백질, 예컨대 전달하고자 하는 Cas9 또는 다른 DNA 결합 단백질을 함유하는 용액 중에 침지된 고정 세포이다. 유사한 방식으로, 표적 핵산 서열에 공동-국제화되는 gRNA 및 DNA 결합 단백질을 사용하는 본원에 기술된 방법은 중기 염색체 스프레드 상에서 수행될 수 있다.

[0021] 한 측면에 따라, 가이드 RNA는 약 10 내지 약 500개의 뉴클레오티드이다. 한 측면에 따라, 가이드 RNA는 약 20 내지 약 100개의 뉴클레오티드이다. 한 측면에 따라, 가이드 RNA는 tracrRNA-crRNA 융합물이다. 한 측면에 따라, tracrRNA 및 crRNA는 별개 종이고, 융합된다.

[0022] 한 측면에 따라, DNA는 게놈 DNA, 미토콘드리아 DNA, 바이러스 DNA, 또는 외인성 DNA이다.

[0023] 한 측면에 따라, 그의 DNA 함량을 특징으로 하는 것인, 샘플 중 하나 이상의 폴리뉴클레오티드 종에 상보적인 하나 이상의 서열을 선택하고, 상보적 서열을 포함하는 하나 이상의 gRNA를 제공하고, 하나 이상의 gRNA를 Cas9 와 조합하고, 샘플을 gRNA 및 Cas9에 노출시키고, 샘플 중 하나 이상의 폴리뉴클레오티드 종에 결합하는 gRNA/Cas9를 검출함으로써 2개 이상의 상이한 폴리뉴클레오티드 종 또는 세포의 혼합물을 포함하는 샘플을 프로빙하는 방법; 및 상기 검출에 기초하여 샘플의 세포 또는 폴리뉴클레오티드 구성 성분의 정체성을 측정하는 방법을 제공한다. 한 측면에 따라, gRNA 및 Cas9는 시험관내에서 제조되거나, 또는 시험관내 존재한다. 한 측면에 따라, gRNA 및 Cas9는 시험관내에서 조합된다. 한 측면에 따라, gRNA는 시험관내에서 제조되거나, 또는 존재하는 반면, Cas9 단백질은 생체내에서 제조된다. 추가 측면에 따라, 샘플은 복수 개의 상이한 폴리뉴클레오티드 종을 포함하고, 예컨대 10 또는 100 또는 1,000 또는 10,000개의 상이한 폴리뉴클레오티드 종으로 이루어진 복합 혼합물 중에 존재할 수 있다.

[0024] 다른 출원은 가이드 RNA 및 Cas9를 이용하여 표적 유기체의 정체성을 평가하는 방법; 가이드 RNA 및 Cas9를 이용하는 단계를 포함하는, 표적 유기체의 상태를 평가하는 방법; DNA 문자 상에 결합된 복수 개의 Cas9 및 가이드 RNA 복합체를 분석하는 단계를 포함하는, DNA 문자를 맵핑하는 방법; 또는 복수 개의 Cas9 및 가이드 RNA 복합체 및 복수 개의 프로브를 이용하는 단계를 포함하는, DNA 문자 중 대립유전자 변이체를 분석하는 방법을 포함한다. 이들 구체적인 출원들은 각각 본원에 기술된 gRNA/Cas9 시스템을 이용하여 표적 DNA 부위에서 복합체를 형성하는 DNA 프로빙 방법 및 gRNA/Cas9 복합체 검출 방법에 기초한다.

[0025] DNA 문자는 염색체 또는 염색체외 DNA 문자일 수 있다. Cas9 엔도뉴클레아제는 활성일 수 있거나, 또는 비활성일 수 있거나, 또는 부분적으로 비활성일 수 있다. Cas9는 형광 단백질 (예컨대, GFP, 루시페라제 등) 및/또는 하나 또는 다중의 친화성 태그와 함께 융합되어 있을 수 있다. 친화성 태그는 하나 또는 다중의 형광 프로브에 의해 인식될 수 있다. 친화성 태그는 Cas9에 측정가능한 속성 (예컨대, 전하 또는 형상)을 부가하는 하나 또는 다중의 태그에 의해 인식될 수 있다. Cas9는 하나 또는 다중의 오르토고날 아미노산을 함유할 수 있다. 오르토고날 아미노산은 다른 문자, 예컨대 프로브, 태그, 링커에 친화성을 제공할 수 있다.

[0026] 가이드 RNA는 하나 또는 다중의 형광 프로브를 사용하여 직접적으로 프로빙될 수 있다. 가이드 RNA는 Cas9에 측정가능한 속성 (예컨대, 전하 또는 형상)을 부가하는 하나 또는 다중의 태그에 의해 직접적으로 프로빙될 수 있다. 가이드 RNA는 하나 또는 다중의 변형된 엔기를 함유할 수 있다. 변형된 엔기는 다른 문자, 예컨대 프로브, 태그, 링커에 친화성을 제공할 수 있다.

[0027] 유기체는 원핵생물 또는 진핵생물, 단세포 또는 다세포일 수 있다. DNA는 유기체로부터 추출된다. DNA는 그의 천연 형태일 수 있거나, 또는 DNA는 표면 상에서 또는 장치 내에서 신장될 수 있다. DNA는 채널 또는 나노포어를 통해 전위될 수 있다. 유기체를 고정시키고, 시험관내에서 합성된 Cas9 및 가이드 RNA 복합체로 투과가능하게 만든다.

[0028] 특정 실시양태에서, Cas9 및 가이드 RNA는 DNA 표적화를 위해 사용되기 이전에 복합체를 형성하게 된다. 가이드 RNA는 표적 DNA에 상보적이다. DNA에 결합된 복합체는 형광 신호(들)를 측정함으로써 검출된다. DNA에 결합된 복합체는 나노포어 센서를 통해 또는 나노포어 또는 나노캡 센서에 근접하게 전위되는 동안 전류 신호를

측정함으로써 검출된다.

- [0029] DNA 상의 하나 또는 다중 복합체는 한번에 검출될 수 있다. DNA 상의 임의의 두 복합체 사이의 해상도는 1 나노미터 또는 5 나노미터 또는 10 나노미터만큼 낮을 수 있고, 1,000 밀리미터만큼 높을 수 있다 (그리고, 그 사이의 임의의 수치일 수 있다). 특이적인 복합체 검출은 특이적인 대립유전자의 존재를 나타낸다. DNA에 결합된 복합체 패턴이 생성될 수 있고, 이를 사용하여 DNA 분자의 맵을 제공할 수 있다. DNA에 결합된 복합체 패턴이 생성될 수 있고, 이를 사용하여 유기체의 정체성 및/또는 상태를 제공할 수 있다. 한 측면에 따라, 가이드 RNA 또는 Cas9 또는 가이드 RNA/Cas9 복합체는 살아있는 세포 또는 비생존 가능한 (즉, 사멸) 세포에 제공될 수 있다.
- [0030] DNA 분자 상에 서열 특이적 출발 부위를 생성하기 위해 Cas9 및 가이드 RNA 복합체를 사용하는 방법을 제공한다. 폴리머라제 또는 리가제를 사용하여 단일 분자 서열분석을 수행할 수 있다. 출발 부위는 게놈 변이체, 반복 서열, 고도의 가변 영역에 인접해 있을 수 있다.
- [0031] DNA 분자를 풀다운시키기 위해, 즉 친화도 정제를 위해 Cas9 및 가이드 RNA 복합체를 사용하는 방법을 제공한다. 풀다운을 허용하는 친화성 태그는 Cas9 및 가이드 RNA 복합체에 결합된다. 복합체가 DNA에 결합하기 전 또는 그 이후에 친화성 태그는 복합체에 결합된다. 하나 또는 다중의 Cas9 및 가이드 RNA 복합체에 결합된 하나의 특이적인 또는 다중의 특이적인 표적 DNA 분자(들)는 풀로부터 추출될 수 있다. 추출된 표적 DNA 분자는 서열분석, 예컨대 심층 서열분석될 수 있다.
- [0032] 특정 실시양태에서, 가이드 RNA는 특정 조건하에서 Cas9 (또는 다른 DNA 결합 단백질) 없이 사용되고, 특정 유형의 서열에 대해 표적화될 때, 가이드 RNA는 DNA 표적의 안정적 또는 일시적 부착을 형성하는데 충분하다.
- [0033] 특정 실시양태에서, gRNA/Cas9 공동-국재화 복합체는 이중 가닥 절단을 제공하지만, gRNA/Cas9 공동-국재화 복합체는 표적 핵산에 결합된 상태 그대로 유지된다. 상기 gRNA/Cas9 공동-국재화는 예컨대, 복합체 파괴를 시작하기 위해 7 M 우레이를 첨가하는 것과 같은 조건을 사용하여 표적 핵산으로부터 제거될 수 있다 (문헌 [Sternberg et al. Nature 507:62 (2014)] (본원에서 그 전문이 포함됨) 참조).
- [0034] 일부 실시양태에서, gRNA 및 Cas9의 복합체는 표적 핵산 서열, 예컨대 DNA에의 결합 이전에 형성된다. 일부 실시양태에서, 복합체는 Cas9가 먼저 표적 핵산과 상호작용한 후, 형성되고, 즉 Cas9가 표적 핵산과 상호작용한 후, 이어서 가이드 RNA와의 공동-국재화 복합체가 형성된다.
- [0035] 본 발명의 특정 실시양태의 추가의 특징 및 이점은 실시양태에 관한 하기의 상세한 설명 및 그의 도면, 및 특히 청구범위로부터 더욱 충분하게 자명해질 것이다.

도면의 간단한 설명

- [0036] 본 실시양태의 상기 및 다른 특징 및 이점은 첨부된 도면과 함께 포함된 예시적인 실시양태에 관한 하기의 상세한 설명으로부터 더욱 상세하게 이해될 것이다:
- 도 1은 표지된 gRNA/Cas9에 의해 프로빙된 고정 마우스 세포의 영상을 보여주는 것이다.
- 도 2는 도 1에서와 같은 Cas9 프로빙 프로토콜에 따라 표지된 올리고뉴클레오티드로 프로빙된 고정 마우스 세포의 영상을 보여주는 것이다.
- 도 3은 gRNA/Cas9 절단의 아가로스 겔 및 겔 이동 검정법을 보여주는 것이다.
- 도 4는 gRNA 테일을 프로빙하는 것에 관한 다양한 개략도를 보여주는 것이다.
- 도 5는 측면 유동 검정법의 다이어그램을 보여주는 것이다.
- 도 6은 표지된 gRNA/Cas9에 의해 프로빙된 신장된 DNA의 영상을 보여주는 것이다.
- 도 7은 gRNA/Cas9 프로빙을 이용하여 게놈 재배열을 확인하는 것에 관한 다이어그램을 보여주는 것이다.
- 도 8은 gRNA/Cas9에 부착된 오리가미 바코드를 이용하여 게놈 영역을 확인하는 것에 관한 다이어그램을 보여주는 것이다.
- 도 9는 낙 부위로부터 서열분석을 개시하는 것에 관한 다이어그램을 보여주는 것이다.
- 도 10은 Her2를 확인하기 위해, CHOPCHOP를 이용하여 gRNA/Cas9 표적 부위를 확인하는 결과 다이어그램을 보여

주는 것이다.

도 11은 올리고뉴클레오티드 양상별로부터 gRNA 주형을 제조하는 PCR 어셈블리 전략법을 보여주는 것이다.

도 12는 gRNA/Cas9 프로빙을 이용하여 계놈 융합물을 확인하는 것에 관한 다이어그램을 보여주는 것이다.

발명을 실시하기 위한 구체적인 내용

[0037]

본 개시내용의 실시양태는 DNA 결합 단백질 및 가이드 RNA를 사용하여 표적 핵산에 공동-국재화하거나, 표적 핵산에서 복합체를 형성한 후, 복합체와 회합된 또는 복합체에 부착된 검출가능한 모이어티의 검출에 의해, 또는 복합체 그 자체를 물리적으로 프로빙함으로써 표적 핵산을 검출하는 것에 기초한다. 상기 DNA 결합 단백질은 관련 기술분야의 통상의 기술자에게 쉽게 공지되어 있는 것으로서, 다양한 목적으로 DNA에 결합하는 RNA-가이드된 DNA 결합 단백질을 포함한다. 상기 DNA 결합 단백질은 자연적으로 발생된 것일 수 있다. 본 개시내용의 범주 내에 포함된 DNA 결합 단백질은 본원에서 가이드 RNA로서 지칭되는, RNA에 의해 가이드될 수 있는 것을 포함한다. 본 측면에 따라, 가이드 RNA 및 RNA 가이드된 DNA 결합 단백질은 DNA에서 공동-국재화 복합체를 형성한다. 특정 측면에 따라, DNA 결합 단백질은 뉴클레아제-널 DNA 결합 단백질일 수 있다. 본 측면에 따라, 뉴클레아제-널 DNA 결합 단백질은 뉴클레아제 활성을 가진 DNA 결합 단백질의 변경 또는 변형으로부터 생성될 수 있다. 상기의, 뉴클레아제 활성을 가진 DNA 결합 단백질은 관련 기술분야의 통상의 기술자에게 공지되어 있고, 예컨대 예를 들어, 유형 II CRISPR 시스템에 존재하는 Cas9 단백질과 같은, 자연적으로 발생된, 뉴클레아제 활성을 가진 DNA 결합 단백질을 포함한다. 상기 Cas9 단백질 및 유형 II CRISPR 시스템은 관련 기술분야에 문서로 잘 입증되어 있다. 모든 보충 정보를 비롯한, 문헌 [Makarova et al., *Nature Reviews Microbiology*, Vol. 9, June 2011, pp. 467-477] (상기 문헌은 그 전문이 본원에서 참조로 포함된다)을 참조할 수 있다.

[0038]

예시적인, 뉴클레아제 활성을 가진 DNA 결합 단백질은 이중 가닥 DNA를 낉king 또는 컷팅하는 기능을 한다. 상기 뉴클레아제 활성을 뉴클레아제 활성을 보이는 하나 이상의 폴리펩티드 서열을 갖는 DNA 결합 단백질로부터 비롯될 수 있다. 상기 예시적인 DNA 결합 단백질은 별개의 두 뉴클레아제 도메인을 가질 수 있으며, 여기서 각 도메인은 이중 가닥 DNA의 특정 가닥을 컷팅하거나, 낉king하는 것을 담당한다. 관련 기술분야의 통상의 기술자에게 공지된, 예시적인, 뉴클레아제 활성을 갖는 폴리펩티드 서열은 McrA-HNH 뉴클레아제 관련 도메인 및 RuvC-유사 뉴클레아제 도메인을 포함한다. 따라서, 예시적인 DNA 결합 단백질은 사실상 McrA-HNH 뉴클레아제 관련 도메인 및 RuvC-유사 뉴클레아제 도메인 중 하나 이상의 것을 포함하는 것이다. 특정 측면에 따라, DNA 결합 단백질은 뉴클레아제 활성을 불활성화시키도록 변경되거나, 또는 다르게는 변형된다. 상기 변경 또는 변형은 뉴클레아제 활성을 또는 뉴클레아제 도메인을 불활성화시키도록 하나 이상의 아미노산을 변경시키는 것을 포함한다. 상기 변형은 뉴클레아제 활성을 보이는 폴리펩티드 서열 또는 폴리펩티드 서열들, 즉 뉴클레아제 도메인을 제거함으로써, 뉴클레아제 활성을 보이는 폴리펩티드 서열 또는 폴리펩티드 서열들, 즉 뉴클레아제 도메인이 DNA 결합 단백질에 존재하지 않도록 하는 것을 포함한다. 뉴클레아제 활성을 불활성화시키기 위한 다른 변형은 본 개시내용에 기초하여 관련 기술분야의 통상의 기술자에게 쉽게 자명해질 것이다. 따라서, 뉴클레아제-널 DNA 결합 단백질은 뉴클레아제 활성을 불활성화시키기 위해 변형된 폴리펩티드 서열, 또는 뉴클레아제 활성을 불활성화시키기 위한 폴리펩티드 서열 또는 서열들의 제거를 포함한다. 뉴클레아제-널 DNA 결합 단백질은 비록 뉴클레아제 활성이 불활성화되기는 하였지만, DNA에 결합할 수 있는 능력은 유지한다. 따라서, DNA 결합 단백질은 DNA 결합에 필요한 폴리펩티드 서열 또는 서열들은 포함하지만, 뉴클레아제 활성을 보이는 뉴클레아제 서열 중 하나 이상 또는 그들 모두는 포함하지 않을 수 있다. 따라서, DNA 결합 단백질은 DNA 결합에 필요한 폴리펩티드 서열 또는 서열들은 포함하지만, 불활성화된 뉴클레아제 활성을 보이는 뉴클레아제 서열 중 하나 이상 또는 그들 모두를 가질 수 있다.

[0039]

한 측면에 따라, 2개 이상의 뉴클레아제 도메인을 갖는 DNA 결합 단백질은 1개를 제외한 모든 뉴클레아제 도메인을 불활성화시키기 위해 변형 또는 변경될 수 있다. 상기 변형 또는 변경된 DNA 결합 단백질은, DNA 결합 단백질이 이중 가닥 DNA 중 단 하나의 가닥만을 컷팅 또는 낉king할 수 있는 한, DNA 결합 단백질 낙카제로서 지칭된다. RNA에 의해 DNA로 가이드된 경우, DNA 결합 단백질 낙카제는 RNA 가이드된 DNA 결합 단백질 낙카제로서 지칭된다. 따라서, 유용한 Cas9 단백질은 야생형 Cas9, Cas9 낙카제 또는 뉴클레아제 널 Cas9 및 그의 호모로그 및 오솔로그일 수 있다. 문헌 [Jinek et al., *Science* 337, 816-821 (2012)] (상기 문헌은 그 전문이 본원에서 참조로 포함된다)를 참조할 수 있다.

[0040]

S. 피오제네스에서, Cas9는 단백질에서 2개의 촉매 도메인: DNA의 상보적 가닥을 절단하는 HNH 도메인 및 비-상보적 가닥을 절단하는 RuvC-유사 도메인에 의해 매개되는 프로세스를 통해 프로토스페이서-인접 모티프 (PAM)로부터 상류 쪽으로 3 bp 만큼 떨어져 있는 곳에서 평활-말단 이중-가닥 파괴를 생성한다. 문헌 [Jinek et al.,

Science 337, 816–821 (2012)] (상기 문헌은 그 전문이 본원에서 참조로 포함된다)를 참조할 수 있다. Cas9 단백질은 문헌 [Makarova et al., *Nature Reviews, Microbiology*, Vol. 9, June 2011, pp. 467–477]에 대한 보충 정보에서 확인되는 하기 것들을 비롯하여 다수의 유형 II CRISPR 시스템으로 존재하는 것으로 공지되어 있다: 메타노코쿠스 마리팔루디스(*Methanococcus maripaludis*) C7; 코리네박테리움 디프테라이에 (*Corynebacterium diphtheriae*); 코리네박테리움 에피시엔스(*Corynebacterium efficiens*) YS-314; 코리네박테리움 글루타미쿰(*Corynebacterium glutamicum*) ATCC 13032 키타사토(Kitasato); 코리네박테리움 글루타미쿰 ATCC 13032 빌레펠트(Bielefeld); 코리네박테리움 글루타미쿰 R; 코리네박테리움 크로펜스테티이 (*Corynebacterium kroppenstedtii*) DSM 44385; 미코박테리움 압세수스(*Mycobacterium abscessus*) ATCC 19977; 노카르디아 파르시니카(*Nocardia farcinica*) IFM10152; 로도코쿠스 에리트로폴리스(*Rhodococcus erythropolis*) PR4; 로도코쿠스 조스티이(*Rhodococcus jostii*) RHA1; 로도코쿠스 오파쿠스(*Rhodococcus opacus*) B4 uid36573; 아시도테르무스 셀룰롤리티쿠스(*Acidothermus cellulolyticus*) 11B; 아르트로박터 클로로페놀리쿠스 (*Arthrobacter chlorophenolicus*) A6; 크리밸라 플라비다(*Kribbella flava*) DSM 17836 uid43465; 써모모노스 포라 쿠르바타(*Thermomonospora curvata*) DSM 43183; 비피도박테리움 덴티움(*Bifidobacterium dentium*) Bd1; 비피도박테리움 롱굼(*Bifidobacterium longum*) DJ010A; 슬라키아 헬리오토리니레두센스(*Slackia heliotrinireducens*) DSM 20476; 페르세포넬라 마리나(*Persephonella marina*) EX H1; 박테리오데스 프라길리스 (*Bacteroides fragilis*) NCTC 9434; 카프노시토파가 오크라세아(*Capnocytophaga ochracea*) DSM 7271; 플라보박테리움 사이크로필룸(*Flavobacterium psychrophilum*) JIP02 86; 악케르만시아 무시니필라(*Akkermansia muciniphila*) ATCC BAA 835; 로세이플렉수스 카스텐홀치이(*Roseiflexus castenholzii*) DSM 13941; 로세이플렉수스(*Roseiflexus*) RS1; 시네코시스티스(*Synechocystis*) PCC6803; 엘루시미크로비움 미누툼(*Elusimicrobium minutum*) Pei191; 비배양된 흰개미 군 1 박테리아 계통형 Rs D17; 피브로박터 숙시노게네스(*Fibrobacter succinogenes*) S85; 바실러스 세레우스(*Bacillus cereus*) ATCC 10987; 리스테리아 이노쿠아(*Listeria innocua*); 락토바실러스 카세이(*Lactobacillus casei*); 락토바실러스 람노수스(*Lactobacillus rhamnosus*) GG; 락토바실러스 살리바리우스(*Lactobacillus salivarius*) UCC118; 스트렙토코쿠스 아갈락티아에(*Streptococcus agalactiae*) A909; 스트렙토코쿠스 아갈락티아에 NEM316; 스트렙토코쿠스 아갈락티아에 2603; 스트렙토코쿠스 디스갈락티아에 에퀴시밀리스(*Streptococcus dysgalactiae equisimilis*) GGS 124; 스트렙토코쿠스 에퀴 주에피데미쿠스(*Streptococcus equi zooepidemicus*) MGCS10565; 스트렙토코쿠스 갈룰리티쿠스(*Streptococcus gallolyticus*) UCN34 uid46061; 스트렙토코쿠스 고르도니이 칼리스(*Streptococcus gordoni Chassis*) subst CH1; 스트렙토코쿠스 뮤탄스(*Streptococcus mutans*) NN2025 uid46353; 스트렙토코쿠스 뮤탄스; 스트렙토코쿠스 피오계네스(*Streptococcus pyogenes*) M1 GAS; 스트렙토코쿠스 피오계네스 MGAS5005; 스트렙토코쿠스 피오계네스 MGAS2096; 스트렙토코쿠스 피오계네스 MGAS9429; 스트렙토코쿠스 피오계네스 MGAS10270; 스트렙토코쿠스 피오계네스 MGAS6180; 스트렙토코쿠스 피오계네스 MGAS315; 스트렙토코쿠스 피오계네스 SSI-1; 스트렙토코쿠스 피오계네스 MGAS10750; 스트렙토코쿠스 피오계네스 NZ131; 스트렙토코쿠스 써모필레스(*Streptococcus thermophiles*) CNRZ1066; 스트렙토코쿠스 써모필레스 LMD-9; 스트렙토코쿠스 써모필레스 LMG 18311; 클로스트리디움 보툴리눔(*Clostridium botulinum*) A3 로크 마리(Loch Maree); 클로스트리디움 보툴리눔 B 에클룬드(Eklund) 17B; 클로스트리디움 보툴리눔 Ba4 657; 클로스트리디움 보툴리눔 F 랑겔란드(Langeland); 클로스트리디움 셀룰롤리티쿰(*Clostridium cellulolyticum*) H10; 피네골디아 마그나(*Finegoldia magna*) ATCC 29328; 유박테리움 텍탈레(*Eubacterium rectale*) ATCC 33656; 미코플라스마 갈리셉티쿰(*Mycoplasma gallisepticum*); 미코플라스마 모빌레(*Mycoplasma mobile*) 163K; 미코플라스마 페네트란스(*Mycoplasma penetrans*); 미코플라스마 시노비아에(*Mycoplasma synoviae*) 53; 스트렙토바실러스 모닐리포르미스(*Streptobacillus moniliformis*) DSM 12112; 브라디리조비움(*Bradyrhizobium*) BTai1; 니트로박터 함부르겐시스(*Nitrobacter hamburgensis*) X14; 로도슈도모나스 팔루스트리스(*Rhodopseudomonas palustris*) BisB18; 로도슈도모나스 팔루스트리스 BisB5; 파르비바콜룸 라바멘티보란스(*Parvibaculum lavamentivorans*) DS-1; 디노로세오박터 쉬바에(*Dinoroseobacter shibae*) DFL 12; 글루코나세토박터 디아조트로피쿠스(*Gluconacetobacter diazotrophicus*) Pal 5 FAPERJ; 글루코나세토박터 디아조트로피쿠스(*Gluconacetobacter diazotrophicus*) Pal 5 JGI; 아조스파릴룸(*Azospirillum*) B510 uid46085; 로도스파릴룸 루브룸(*Rhodospirillum rubrum*) ATCC 11170; 디아포로박터(*Diaphorobacter*) TPSY uid29975; 베르미네프로박터 에이세니아에(*Verminephrobacter eiseniae*) EF01-2; 네이세리아 메닌기티데스 (*Neisseria meningitidis*) 053442; 네이세리아 메닌기티데스(*Neisseria meningitidis*) 알파14; 네이세리아 메닌기티데스 Z2491; 데술포비브리오 살렉시겐스(*Desulfovibrio salexigens*) DSM 2638; 캄필로박터 제주니 도일레이(*Campylobacter jejuni doylei*) 269 97; 캄필로박터 제주니 81116; 캄필로박터 제주니; 캄필로박터 라리(*Campylobacter lari*) RM2100; 헬리코박터 헤파티쿠스(*Helicobacter hepaticus*); 월리넬라 숙시노게네스 (*Wolinella succinogenes*); 톨루모나스 아우엔시스(*Tolumonas auensis*) DSM 9187; 슈도알테로모나스 아틀란티

카(*Pseudoalteromonas atlantica*) T6c; 슈와넬라 페알레아나(*Shewanella pealeana*) ATCC 700345; 레지오넬라 뉴모필라 파리(*Legionella pneumophila Paris*); 악티노바실러스 숙시노게네스(*Actinobacillus succinogenes*) 130Z; 파스테우렐라 물토시다(*Pasteurella multocida*); 프란시셀라 툴라렌시스 노비시다(*Francisella tularensis novicida*) U112; 프란시셀라 툴라렌시스 홀라르크티카(*Francisella tularensis holarctica*); 프란시셀라 툴라렌시스 FSC 198; 프란시셀라 툴라렌시스 툴라렌시스; 프란시셀라 툴라렌시스 WY96-3418; 및 트레포네마 덴티콜라(*Treponema denticola*) ATCC 35405. 따라서, 본 개시내용의 측면은 뉴클레아제 기능부재된, 또는 본원에 기술된 바와 같은 닉카제가 된 유형 II CRISPR 시스템에 존재하는 Cas9 단백질에 관한 것이다.

[0041]

Cas9 단백질은 관련 기술분야의 통상의 기술자에 의해 문헌에서 Csn1로서 지칭될 수 있다. S. 피오게네스 Cas9 단백질 서열은 하기 제시된다. 문헌 [Deltcheva et al., *Nature* 471, 602-607 (2011)] (상기 문헌은 그 전문이 본원에서 참조로 포함된다)을 참조할 수 있다.

```
MDKKYSIGLDIGTNSVGWAVITDEYKVPSSKKFKVLGNTDRHSIKKNLIGALLFDGETAE
ATRLKRTARRRYTRRKNRICYLQEIFSNEAKVDDDSFFHRLEESFLVEEDKKHERHPIFG
NIVDEVAYHEKYPTIYHLRKKLVDSTDKAIDLRLIYLALAHMIKFRGHFLIEGDLNPDNSD
VDKLFQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGN
LIALSGLTPNFKNFDSLAEAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAI
LLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSCKNGYA
GYIDGGASQEEFYKFKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELH
AILRRQEDFYPFLKDNRREKIEKILTFRIPYYVGPLARGNSRFAMTRKSEETITPWNFEE
VVDKGASAQSFIERNMTNFDKNLPNEVKLPKHSLLYEFVYNEALKVKYVTEGMRKPAFL
SGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKI
IKDKDFLDNEENEDILEDIVLTTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRTGWG
RLSRKLINGIRDQSGKTIIDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSDL
HEHIANLAGSPAIIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRER
MKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYLQNGRDMYVDQELDINRLSDYDVDH
IVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLNAKLITQRKFNDL
TKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDLSRMNTKYDENDKLIREVKVITLKS
KLVSDFRKDFQFYKVREINNYHHAAHDAYLNAVVTALIKKYPKLESEFVYGDYKVYDVR
K
MIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDF
ATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWPKKYGGFDSPPTVA
YSVLVVAKVEKGKSKKLKSVEKLLGITMERSFEKNPIDFLEAKGYKEVKKDLIILPK
YSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLAHYEKLKGSPEDNEQKQLFVE
QHKHYLDEIIEQISEFSKRVILADANLDKVL SAYNKHRDKPIREQAENIHLFTLNLGA
PAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGD-
```

[0042]

표적 핵산은 그를 검출하는데 본원에 기술된 바와 같은 공동-국제화 복합체가 유용할 수 있는 것인 임의의 핵산 서열을 포함한다. 표적 핵산은 유전자를 포함한다. 표적 핵산은 단일 세포로부터 추출된 DNA 내에 존재할 수 있다. 표적 핵산은 단일 염색체로부터 추출된 DNA일 수 있다. 본 개시내용의 목적을 위해, DNA, 예컨대 이중 가닥 DNA는 표적 핵산을 포함할 수 있고, 공동-국제화 복합체는 표적 핵산을 검출하는 방식으로 표적 핵산의 또는 표적 핵산에 인접해 있는 DNA, 또는 표적 핵산에 근접한 위치의 DNA에 결합하거나, 또는 다르게는 그와 공동-국제화할 수 있다. 상기 표적 핵산은 내인성 (또는 자연적으로 발생된) 핵산 및 외인성 (또는 외래) 핵산을 포함할 수 있다. 상기 표적 핵산은 내인성 (또는 자연적으로 발생된) 핵산 및 외인성 (또는 외래) 핵산을 포함할 수 있다. 상기 표적 핵산은 관련 기술분야의 통상의 기술자에 공지된 방법을 사용하여 연장 또는 신장시킬 수 있다. DNA를 신장시키는 방법은 문헌 [KH Rasmussen, R Marie, JM Lange, WE Svendsen, A Kristensen, and KU Mir, *Lab Chip*, 2011, 11:1431-44 and A device for extraction, manipulation and stretching of DNA from single human chromosomes]; [DLV Bauer, R Marie, KH Rasmussen, A Kristensen, KU Mir, 2012 *Nuc1 Acids Res*, 2012, 1-7, DNA catenation maintains structure of human metaphase chromosomes]에 기술되어 있다.

[0044]

검출가능한 표지 또는 모이어티는 관련 기술분야의 통상의 기술자에게 공지되어 있다. 본원에서 사용되는 바,

"검출가능한 표지"라는 용어는 표적 핵산을 확인하는데 사용될 수 있는 표지를 지칭한다. 검출가능한 표지는 관련 기술분야의 통상의 기술자에게 공지된 방법을 사용하여 gRNA 또는 Cas9 단백질에 부착된다. 대안적으로, gRNA 또는 Cas9 단백질은 결합 쌍의 절반부를 포함할 수 있고, 상응하는 결합 쌍의 나머지 절반부는 검출가능한 표지에 결합될 수 있다. 이러한 방식으로, 표지는 결합 쌍의 결합에 기인하여 gRNA 또는 Cas9 단백질에 간접적으로 결합될 수 있다. 적합한 결합 쌍 또는 결합력은 관련 기술분야의 통상의 기술자에게 공지되어 있고, 상보적인 핵산 서열, 비오틴-아비딘, 비오틴-스트렙트아비딘, NHS-에스테르 등, 티오에테르 결합, 정전하 상호작용, 반 데르 발스의 힘 등을 포함한다 (예컨대, 홀트케(Holtke) 등의 미국 특허 번호 5,344,757; 5,702,888; 및 5,354,657; 후버(Huber) 등의 미국 특허 번호 5,198,537; 미요시(Miyoshi)의 미국 특허 번호 4,849,336; 미시우라(Misiura) 및 가이트(Gait)의 PCT 공보 WO 91/17160 참조). 비오틴, 또는 그의 유도체는 올리고뉴클레오티드 표지로서 (예컨대, 표적화 모이어티, 검색가능한 모이어티 및/또는 검출가능한 표지로서) 사용될 수 있고, 이어서 아비딘/스트렙트아비딘 유도체 (예컨대, 검출가능하게 표지된, 예컨대 피코에리트린-접합된 스트렙트아비딘), 또는 항-비오틴 항체 (예컨대, 검출가능하게 표지된 항체)에 의해 결합될 수 있다. 디곡시게닌은 표지로서 도입될 수 있고, 이어서 검출가능하게 표지된 항-디곡시게닌 항체 (예컨대, 검출가능하게 표지된 항체, 예컨대 플루오레세인화된 항-디곡시게닌)에 의해 결합될 수 있다. 아미노알릴-dUTP 잔기는 올리고뉴클레오티드 내로 도입될 수 있고, 이어서 N-히드록시 숙신이미드 (NHS) 유도체화된 형광 염료에 커플링될 수 있다. 일반적으로, 검출가능하게 표지된 접합체 파트너가 결합될 수 있고, 이에 의해 검출될 수 있다면, 접합체 쌍의 임의의 구성원이 검색가능한 모이어티 및/또는 검출가능한 표지 내로 도입될 수 있다. 본원에서 사용되는 바, 항체라는 용어는 임의 부류의 항체 분자, 또는 그의 임의의 하위 단편, 예컨대 Fab를 의미한다.

[0045] 검출가능한 표지는 크기 및 조성에 있어서 광범위하게 달라질 수 있고; 하기 참고 문헌이 특정 실시양태에 적절한 올리고뉴클레오티드 태그를 선택하는 것에 관한 가이던스를 제공한다: 브렌너(Brenner)의 미국 특허 번호 5,635,400; [Brenner et al., Proc. Natl. Acad. Sci., 97: 1665]; [Shoemaker et al. (1996) Nature Genetics, 14:450]; 모리스(Morris) 등의 EP 특허 공개 0799897A1; 월리스(Wallace)의 미국 특허 번호 5,981,179 등.

[0046] 검출가능한 표지를 핵산 프로브 내로 도입하는 방법은 널리 공지되어 있다. 전형적으로, 검출가능한 표지 (예컨대, 합텐- 또는 플루오로크롬-접합된 데옥시리보뉴클레오티드)는 중합화 또는 증폭 단계 동안, 예컨대 PCR, 낙 번역, 무작위 프라이머 표지, 말단 트랜스퍼라제 테일링 (예컨대, 하나 이상의 표지가 프라이머 서열의 절단 이후에 부가될 수 있다) 등에 의해 핵산, 예컨대 핵산 프로브 내로 도입된다 (문헌 [Ausubel et al., 1997, Current Protocols In Molecular Biology, Greene Publishing and Wiley-Interscience, New York] 참조).

[0047] 검출가능한 모이어티, 표지 또는 리포터는 본원에 기술된 바와 같이 표적 핵산을 검출하는데 사용될 수 있다. 가이드 RNA 또는 Cas9 단백질은 검출가능한 모이어티, 예컨대 형광 모이어티, 합텐, 비색 모이어티 등의 직접 또는 간접적인 부착을 비롯한, 다양한 방식으로 표지될 수 있다. 표지가 부착될 수 있는 위치는 본원에서 표지 부가 부위 또는 검출가능한 모이어티 부가 부위로 지정되고, 표지가 부착될 수 있는 뉴클레오티드를 포함할 수 있다. 관련 기술분야의 통상의 기술자는 핵산 또는 단백질의 표지에 관하여 참고문헌을 참조할 수 있다. 검출가능한 모이어티의 예는 각종 방사성 모이어티, 효소, 보결분자단, 형광 마커, 발광성 마커, 생체발광성 마커, 금속 입자, 단백질-단백질 결합 쌍, 단백질-항체 결합 쌍 등을 포함한다. 형광 모이어티의 예는 황색 형광 단백질 (YFP), 녹색 형광 단백질 (GFP), 청록색 형광 단백질 (CFP), 웜밸리페론, 플루오레세인, 플루오레세인 이소티오시아네이트, 로다민, 디클로로트리아지닐아민 플루오레세인, 시아닌, 단실 클로라이드, 피코시아닌, 피코에리트린 등을 포함하나, 이에 제한되지 않는다. 생체발광성 마커의 예는 루시페라제 (예컨대, 박테리아성, 반딧불이, 방아벌레 등), 루시페린, 에퀴린 등을 포함하나, 이에 제한되지 않는다. 시작적으로 검출가능한 신호를 갖는 효소 시스템의 예는 갈락토시다제, 글루코리니다제, 포스파타제, 페옥시다제, 콜린에스테라제 등을 포함하나, 이에 제한되지 않는다. 확인가능한 마커는 또는 방사성 화합물, 예컨대 ¹²⁵I, ³⁵S, ¹⁴C, 또는 ³H를 포함한다. 확인가능한 마커는 다양한 공급처로부터 상업적으로 이용가능하다.

[0048] 형광 표지 및 뉴클레오티드 및 올리고뉴클레오티드에의 그의 부착점은 문헌 [Haugland, Handbook of Fluorescent Probes and Research Chemicals, Ninth Edition (Molecular Probes, Inc., Eugene, 2002)]; [Keller and Manak, DNA Probes, 2nd Edition (Stockton Press, New York, 1993)]; [Eckstein, editor, Oligonucleotides and Analogues: A Practical Approach (IRL Press, Oxford, 1991)]; 및 [Wetmur, Critical Reviews in Biochemistry and Molecular Biology, 26:227-259 (1991)]을 비롯한, 다수의 리뷰에 기술되어 있다. 본 발명에 적용가능한 특정 방법은 하기 참고문헌 샘플에 개시되어 있다: 미국 특허 번호 4,757,141, 5,151,507 및 5,091,519. 한 측면에서, 하나 이상의 형광 염료는 예컨대, 미국 특허 번호 5,188,934 (4,7-디

클로로플루오레세인 염료); 5,366,860 (스펙트럼으로 분석가능한 로다민 염료); 5,847,162 (4,7-디클로로로다민 염료); 4,318,846 (에테르-치환된 플루오레세인 염료); 5,800,996 (에너지 전달 염료); 리(Lee) 등의 5,066,580 (크산틴 염료); 5,688,648 (에너지 전달 염료) 등에 개시된 바와 같이, 표지된 표적 서열에 대한 표지로서 사용된다. 표지는 또한 하기 특허 및 특허 공개: 미국 특허 번호 6,322,901, 6,576,291, 6,423,551, 6,251,303, 6,319,426, 6,426,513, 6,444,143, 5,990,479, 6,207,392, 2002/0045045 및 2003/0017264에 개시된 바와 같이, 양자점으로 수행될 수 있다. 본원에서 사용되는 바, "형광 표지"라는 용어는 하나 이상의 분자의 형광 흡수 및/또는 방출 특성을 통해 정보를 전달하는 신호전달 모이어티를 포함한다. 상기 형광 특성은 형광 강도, 형광 수명, 방출 스펙트럼 특징, 에너지 전달 등을 포함한다.

[0049] 뉴클레오티드 및/또는 올리고뉴클레오티드 서열 내로 쉽게 도입되는 상업적으로 이용가능한 형광 뉴클레오티드 유사체는 Cy3-dCTP, Cy3-dUTP, Cy5-dCTP, Cy5-dUTP (아머샴 바이오사이언시스(Amersham Biosciences: 미국 뉴저지주 피스카타웨이)), 플루오레세인-12-dUTP, 테트라메틸로다민-6-dUTP, 텍사스 레드(TEXAS RED)TM-5-dUTP, 캐스캐이드 블루(CASCADE BLUE)TM-7-dUTP, 바디파이(BODIPY) TMFL-14-dUTP, 바디파이 TMR-14-dUTP, 바디파이 TMTR-14-dUTP, 로다민 그린TM-5-dUTP, 오레곤(OREGON) 그린RTM 488-5-dUTP, 텍사스 레드TM-12-dUTP, 바디파이 TM 630/650-14-dUTP, 바디파이 TM 650/665-14-dUTP, 알렉사 플루오르(ALEXA FLUOR)TM 488-5-dUTP, 알렉사 플루오르TM 532-5-dUTP, 알렉사 플루오르TM 568-5-dUTP, 알렉사 플루오르TM 594-5-dUTP, 알렉사 플루오르TM 546-14-dUTP, 플루오레세인-12-UTP, 테트라메틸로다민-6-UTP, 텍사스 레드TM-5-UTP, m체리(mCherry), 캐스캐이드 블루TM-7-UTP, 바디파이 TM FL-14-UTP, 바디파이 TMR-14-UTP, 바디파이 TM TR-14-UTP, 로다민 그린TM-5-UTP, 알렉사 플루오르TM 488-5-UTP, 렉사 플루오르(LEXA FLUOR)TM 546-14-UTP (몰레큘라 프로브즈, 인크.(Molecular Probes Inc.: 미국 오레곤주 유진)) 등을 포함하나, 이에 제한되지 않는다. 대안적으로, 상기 형광단 및 본원에서 언급된 것은 예를 들어, 포스포로아미다이트 및 NHS 화학법을 사용하여 올리고뉴클레오티드 합성 동안에 부가될 수 있다. 프로토콜은 다른 형광단을 갖는 뉴클레오티드의 통상의 합성법에 대해 관련 기술분야에 공지되어 있다 (문헌 [Henegariu et al. (2000) Nature Biotechnol. 18:345] 참조). 2-아미노퓨린은 올리고뉴클레오티드 서열 합성 동안 그에 직접 도입될 수 있는 형광 염기이다. 핵산은 또한 선형적으로 인터칼레이팅 염료, 예컨대 DAPI, YOYO-1, 에티디움 브로마이드, 시아닌 염료 (예컨대, SYBR 그린) 등으로 염색될 수 있다.

[0050] 합성 후 부착에 이용가능한 다른 형광단은 알렉사 플루오르TM 350, 알렉사 플루오르TM 405, 알렉사 플루오르TM 430, 알렉사 플루오르TM 532, 알렉사 플루오르TM 546, 알렉사 플루오르TM 568, 알렉사 플루오르TM 594, 알렉사 플루오르TM 647, 바디파이 493/503, 바디파이 FL, 바디파이 R6G, 바디파이 530/550, 바디파이 TMR, 바디파이 558/568, 바디파이 558/568, 바디파이 564/570, 바디파이 576/589, 바디파이 581/591, 바디파이 TR, 바디파이 630/650, 바디파이 650/665, 캐스캐이드 Blue, 캐스캐이드 엘로우(Cascade Yellow), 단실, 리사민 로다민 B, 마리나 블루(Marina Blue), 오레곤 그린 488, 오레곤 그린 514, 퍼시픽 블루(Pacific Blue), 퍼시픽 오렌지(Pacific Orange), 로다민 6G, 로다민 그린, 로다민 레드, 테트라메틸 로다민, 텍사스 레드 (몰레큘라 프로브즈, 인크. (미국 오레곤주 유진)로부터 이용가능), Cy2, Cy3, Cy3.5, Cy5, Cy5.5, Cy7 (아머샴 바이오사이언시스: 미국 뉴저지주 피스카타웨이) 등을 포함하나, 이에 제한되지 않는다. PerCP-Cy5.5, PE-Cy5, PE-Cy5.5, PE-Cy7, PE-텍사스 레드, APC-Cy7, PE-알렉사 염료 (610, 647, 680), APC-알렉사 염료 등을 포함하나, 이에 제한되지 않는, FRET 텐덤 형광단 또한 사용될 수 있다.

[0051] FRET 텐덤 형광단, 예컨대 PerCP-Cy5.5, PE-Cy5, PE-Cy5.5, PE-Cy7, PE-텍사스 레드, 및 APC-Cy7; 또한 PE-알렉사 염료 (610, 647, 680) 및 APC-알렉사 염료 또한 사용될 수 있다.

[0052] 형광 표지된 뉴클레오티드 및/또는 올리고뉴클레오티드 서열로부터의 신호를 증강시키는데 금속 은 또는 금 입자가 사용될 수 있다 (Lakowicz et al. (2003) BioTechniques 34:62).

[0053] 비오틴, 또는 그의 유도체 또한 뉴클레오티드 및/또는 올리고뉴클레오티드 서열 상의 표지로서 사용될 수 있고, 이어서 검출가능하게 표지된 아비딘/스트렙트아비딘 유도체 (예컨대, 피코에리트린-접합된 스트렙트아비딘), 또는 검출가능하게 표지된 항-비오틴 항체에 의해 결합될 수 있다. 비오틴/아비딘이 리간드-리간드 결합 쌍의 한 예이다. 항체/항원 결합 쌍 또한 본원에 기술된 방법과 함께 사용될 수 있다. 다른 리간드-리간드 결합 쌍 또는 접합체 결합 쌍은 관련 기술분야의 통상의 기술자에게 널리 공지되어 있다. 디곡시게닌은 표지로서 도입될 수 있고, 이후, 검출가능하게 표지된 항-디곡시게닌 항체 (예컨대, 플루오레세인화된 항-디곡시게닌)에 의해 결합될 수 있다. 아미노알릴-dUTP 또는 아미노헥실아크릴아미드-dCTP 잔기가 올리고뉴클레오티드 서열 내로 도입될 수 있고, 이어서 이는 N-히드록시 숙신이미드 (NHS) 유도체화된 형광 염료에 커플링될 수 있다. 일반적으로, 검출가능하게 표지된 접합체 파트너가 결합될 수 있고, 이에 의해 검출될 수 있다면, 임의 개수의 접합체 쌍이 검출 올리고뉴클레오티드 내로 도입될 수 있다. 본원에서 사용되는 바, 항체라는 용어는 임의 부

류의 항체 분자, 또는 그의 임의의 하위 단편, 예컨대 Fab를 의미한다.

[0054] 다른 적합한 표지는 플루오레세인 (FAM, FITC), 디옥시게닌, 디니트로페놀 (DNP), 단실, 비오틴, 브로모데옥시우리딘 (BrdU), 헥사히스티딘 (6xHis), 포스포르-아미노산 (예컨대, P-tyr, P-ser, P-thr) 등을 포함할 수 있다. 한 실시양태에서, 하기 합텐/항체 쌍은 검출을 위해 사용되며, 여기서 각각의 항체는 검출가능한 표지로 유도체화된다: 비오틴/α-비오틴, 디옥시게닌/α-디옥시게닌, 디니트로페놀 (DNP)/α-DNP, 5-카르복시플루오레세인 (FAM)/α-FAM.

[0055] 특정 예시적인 실시양태에서, 뉴클레오티드 및/또는 올리고뉴클레오티드 서열은 예컨대, 미국 특허 번호 5,344,757, 5,702,888, 5,354,657, 5,198,537 및 4,849,336, PCT 공보 WO 91/17160 등에 개시된 바와 같이, 특히, 후속하여 포획제에 의해 결합되는 합텐으로 간접적으로 표지될 수 있다. 다수의 상이한 합텐-포획제 쌍이 이용가능하다. 예시적인 합텐은 비오틴, 테스-비오틴 및 다른 유도체, 디니트로페놀, 단실, 플루오레세인, CY5, 디옥시게닌 등을 포함하나, 이에 제한되지 않는다. 비오틴의 경우, 포획제는 아비딘, 스트렙트아비딘, 또는 항체일 수 있다. 항체는 다른 합텐에 대한 포획제로서 사용될 수 있다 (다수의 염료-항체 쌍이 예컨대, 몰레큘라 프로브즈 (미국 오래온주 유진)로부터의 것과 같이 상업적으로 이용가능하다).

[0056] 특정 측면에 따라, 본원에 기술된 검출가능한 모이어티는 스펙트럼으로 분석가능하다. 복수 개의 형광 표지와 관련하여 "스펙트럼으로 분석가능한"이라는 것은 표지의 형광 방출 대역이 충분히 상이하다는 것, 즉 충분히 비-중첩성이라는 것, 각 표지가 부착되는 분자 태그가 미국 특허 번호 4,230,558; 4,811,218 등, 또는 문헌 [Wheeler et al., pgs. 21-76, in Flow Cytometry: Instrumentation and Data Analysis (Academic Press, New York, 1985)]에 기술된 시스템에 의해 예시되는 바와 같이, 표준 광검출 시스템에 의해, 예컨대 대역 통과 필터 시스템 및 광전자 증배관 등을 사용함으로써 각 표지에 의해 발생되는 형광 신호에 기초하여 구별될 수 있다는 것을 의미한다. 한 측면에서, 스펙트럼으로 분석가능한 유기 염료, 예컨대 플루오레세인, 로다민 등은 파장 방출 최대가 적어도 20 nm 만큼 이격되어 있고, 또 다른 측면에서는 적어도 40 nm 만큼 이격되어 있는 것을 의미한다. 또 다른 측면에서, 스펙트럼으로 분석가능한 킬레이팅된 란타나이드 화합물, 양자점 등은 적어도 10 nm 만큼 이격되어 있고, 추가 측면에서는 적어도 15 nm 만큼 이격되어 있는 것을 의미한다.

[0057] 특정 실시양태에서, 검출가능한 모이어티는 일반 핵산과 비교하였을 때, 전자 현미경에서 사용될 경우, 더 높은 검출가능성을 제공할 수 있다. 검출가능성이 더 높은 모이어티는 대개 금속 및 유기금속 죽, 예컨대 머큐리 아세테이트, 플래티넘 디메틸су 폭시드, 수개의 금속-비피리딜 복합체 (예컨대, 오스뮴-비피, 루테늄-비피, 플래티넘-비피)에 있다. 이들 모이어티 중 일부는 핵산을 특이적으로 쉽게 염색시킬 수 있고, 링커 또한 상기 모이어티를 핵산에 부착시키는데 사용될 수 있다. 합성 동안 뉴클레오티드에 첨가되는 상기 링커는 아크리다이트- 및 티올-변형된 엔티티, 아민 반응성 기, 및 클릭 화학법 수행을 위한 아지드 및 알킨 기이다. 일부 핵산 유사체, 예컨대 일반적으로, 감마-아데노신-티오트리포스페이트, 아이오도데옥시시티딘-트리포스페이트, 및 메탈로뉴클레오시드 또한 더 큰 검출가능성을 갖는다 (문헌 [Dale et al., Proc. Nat. Acad. Sci. USA, Vol. 70, No. 8, pp. 2238-2242 (1973)] 참조). 변형된 뉴클레오티드는 합성 동안 첨가된다. 합성은 예로서, 올리고뉴클레오티드의 고체 지지체 합성을 지칭할 수 있다. 이 경우, 핵산 유사체, 또는 검출가능한 모이어티로, 또는 부착 화학 링커로 변형된 핵산일 수 있는, 변형된 핵산은 고체 지지체 상에 형성된 핵산 단편에 차례로 부가되며, 포스포르아미다이트에 의한 합성이 가장 일반적인 방법이다. 합성은 또한 핵산 주형의 상보적인 가닥을 합성하면서, 폴리머라제에 의해 수행되는 과정을 지칭할 수 있다. 특정 DNA 폴리머라제는 핵산 유사체, 또는 검출가능한 모이어티로 변형되거나, 또는 상보적인 핵산 주형에의 부착 화학 링커에 의해 변형된 것인, 변형된 핵산을 사용 및 도입할 수 있다.

[0058] 사용되는 검출 방법(들)은 반응성 표지, 검색가능한 표지 및/또는 검출가능한 표지에서 사용되는 특정의 검출가능한 표지에 의존할 것이다. 특정 예시적인 실시양태에서, 간기, 전기전, 전기, 중기전, 중기, 후기, 말기 및 세포질분열을 포함하나, 이에 제한되지 않는 세포 주기의 다양한 단계 동안, 본원에 기술된 프로브에 의해 하나 이상의 반응성 표지, 검색가능한 표지, 또는 검출가능한 표지가 그에 결합된, 표적 핵산, 예컨대 염색체 및 염색체의 하위염색체 영역은 현미경, 분광광도계, 튜브 광도계 또는 플레이트 광도계, x선 필름, 신틸레이터, 형광 활성화 세포 분류 (FACS) 장치, 마이크로플루이딕스 장치 등을 위해 선택 및/또는 그의 사용을 위해 스크리닝될 수 있다.

[0059] 본원에서 사용되는 바, "염색체"라는 용어는 DNA, 단백질, RNA 및 다른 관련 인자를 비롯한, 살아있는 세포에서 유전적 특징을 전달하는 유전자에 대한 지지체를 의미한다. 인간 계놈의 염색체를 확인하고, 넘버링하는 통상의 국제 체계가 본원에서 사용된다. 개별 염색체의 크기는 다중 염색체 계놈 내에서 및 계놈마다 달라질 수 있

다. 염색체는 임의 종으로부터 수득될 수 있다. 염색체는 성인 대상체, 청소년 대상체, 유아 대상체로부터, 태동 대상체로부터 (예컨대, 태아기 검사, 예컨대 양수 검사, 용모막 용모 샘플링 등을 통해 태아로부터, 또는 예컨대, 태아 수술 동안 태아로부터 직접적으로), 생물학적 샘플로부터 (예컨대, 생물학적 조직, 체액 세포, 예컨대 객담, 혈액, 혈액 세포, 조직 또는 미세 바늘 생검 샘플, 소변, 뇌척수액, 복막액, 및 흉수, 또는 그로 부터의 세포)로부터, 또는 세포 배양물 샘플 (예컨대, 1차 세포, 무한증식 세포, 부분 무한증식 세포 등)로부터 수득될 수 있다. 특정 예시적인 실시양태에서, 하나 이상의 염색체는 호모(*Homo*), 드로소필라(*Drosophila*), 카에노르하비디티스(*Caenorhabditis*), 다니오(*Danio*), 시프리누스(*Cyprinus*), 에쿠우스(*Equus*), 카니스(*Canis*), 오비스(*Ovis*), 오코린кус(*Ocorynchus*), 살모(*Salmo*), 보스(*Bos*), 수스(*Sus*), 갈루스(*Gallus*), 솔라눔(*Solanum*), 트리티쿰(*Triticum*), 오리자(*Oryza*), 제아(*Zea*), 호르데움(*Hordeum*), 무사(*Musa*), 아베나(*Avena*), 포풀루스(*Populus*), 브라씨카(*Brassica*), 사카룸(*Saccharum*) 등을 포함하나, 이에 제한되지 않는, 하나 이상의 속으로부터 수득될 수 있다.

[0060] 형광 표지된 표적화 모이어티, 또는 검출가능한 표지가 사용될 때, 관련 기술분야에 공지된 통상의 방법을 이용하여 계내 혼성화 결과를 검출 및 기록하는데 형광 광현미경법이 사용될 수 있다. 대안적으로, 영상 처리 능력이 있는 (컴퓨터로 실행되는) 디지털 형광 현미경법이 사용될 수 있다. 다중 컬러 표지가 그에 결합되어 있는 염색체의 FISH를 영상화하기 위한 시스템으로서 널리 알려져 있는 2가지 시스템은 다중-FISH (M-FISH) 및 스펙트럼 핵형분석 (SKY)을 포함한다. 염색체 페인트 방법 및 페이튼된 염색체 검출 방법에 관한 리뷰를 위해 문헌 [Schrock et al. (1996) *Science* 273:494]; [Roberts et al. (1999) *Genes Chrom. Cancer* 25:241]; [Fransz et al. (2002) *Proc. Natl. Acad. Sci. USA* 99:14584]; [Bayani et al. (2004) *Curr. Protocol. Cell Biol.* 22.5.1-22.5.25]; [Danilova et al. (2008) *Chromosoma* 117:345]; 미국 특허 번호 6,066,459; 및 FISH TAGTM DNA 멀티컬러 키트(FISH TAGTM DNA Multicolor Kit) 설명서 (몰레큘라 프로브즈)를 참조할 수 있다.

[0061] 특정 예시적인 실시양태에서, 형광 표지된 염색체의 영상은 컴퓨터화된 영상화 시스템, 예컨대 수정된 (예컨대, 예컨대 소프트웨어, 크로마(Chroma) 84000 필터 세트, 및 증강된 필터 훈), 어플라이드 이미징 코포레이션 시토비전 시스템(Applied Imaging Corporation CytoVision System) (어플라이드 이미징 코포레이션(Applied Imaging Corporation: 미국 캘리포니아주 산타클라라))를 사용하여 검출 및 기록된다. 다른 적합한 시스템은 자이스 악시오포트(Zeiss Axiophot) 현미경에 커플링된 냉각 CCD 카메라 (포토메트릭스(Photometrics), 코닥 KAF 1400 CCD(Kodak KAF 1400 CCD)가 장착된 NU200 시리즈)를 사용하는 컴퓨터화된 영상화 시스템을 포함하며, 영상은 문헌 [Ried et al. (1992) *Proc. Natl. Acad. Sci. USA* 89:1388]에 기술되어 있는 바와 같이 처리된다. 다른 적합한 영상화 및 분석 시스템은 문헌 [Schrock et al., 상기 문헌 동일]; 및 [Speicher et al., 상기 문헌 동일]에 기술되어 있다.

[0062] 본원에 기술된 방법에 의해 생성된 프로보를 사용하는 계내 혼성화 방법은 다양한 생물학적 또는 임상 샘플에서, 임의의 (또는 모든) 세포 주기 단계(들) (예컨대, 유사분열, 감수분열, 간기, G0, G1, S 및/또는 G2)에 있는 세포에서 수행될 수 있다. 예는 모든 유형의 세포 배양물, 동물 또는 식물 조직, 말초 혈액 립프구, 협축 도말, 배양되지 않은 원발성 종양으로부터 제조된 터치식 표본, 암 세포, 골수, 생검으로부터 수득된 세포, 또는 체액 (예컨대, 혈액, 소변, 객담) 중 세포, 양수로부터의 세포, 모체 혈액으로부터의 세포 (예컨대, 태아 세포), 고환 및 난소로부터의 세포 등을 포함한다. 샘플은 종래 기술을 사용하여 본 방법의 검정법을 위해 제조되며, 이는 전형적으로 샘플 또는 표본을 채취한 공급원에 의존한다. 이를 예는 본원에 기술된 방법 및 /또는 조성물에 적용가능한 샘플 유형을 제한하는 것으로 해석되지 않아야 한다.

[0063] 특정 예시적인 실시양태에서, 프로브는 차별적으로 표지된 다중 gRNA/Cas9 (즉, 적어도 2개의 gRNA/Cas9 복합체가 차별적으로 표지)를 포함한다. 각종의 멀티컬러 염색체 페인트 접근법이 관련 기술분야에 기술되어 있으며, 본원에서 제공하는 가이던스에 따라 본 발명에 맞게 적합화될 수 있다. 상기 차별적으로 표지된 표지의 예 ("멀티컬러 FISH")는 문헌 [Schrock et al. (1996) *Science* 273:494], 및 [Speicher et al. (1996) *Nature Genet.* 12:368]에 기술된 것을 포함한다. 슈록(Schrock) 등은 표면형광(epifluorescence) 필터 세트 및 컴퓨터 소프트웨어를 사용하여 표적 염색체 세트에 동시에 혼성화하는 다중의 차별적으로 표지된 DNA 프로브를 검출하고, 그를 구별하는, 스펙트럼 영상화 방법을 기술하였다. 스파이커(스파이커) 등은 27-컬럼 FISH로 명명되는 "조합 멀티플루오르 FISH"에서 인간 염색체 (염색체 아암) 각각을 표지하기 위해 5 플루오로크롬의 상이한 조합을 사용하는 것을 기술하였다. 다른 적합한 방법 또한 사용될 수 있다 (예컨대, 문헌 [Ried et al., 1992, *Proc. Natl. Acad. Sci. USA* 89:1388-92] 참조).

[0064] 특정 측면에 따라, Cas9-gRNA 복합체는 표적 DNA 단일 가닥을 제조할 필요 없이 천연 이중 가닥 DNA 상의 관심 영역을 프로빙하고, 그에 접근하는데 사용된다. 관심 표적 이중 가닥 핵산 서열에 특이적인 가이드 RNA는 관련

기술분야의 통상의 기술자에게 공지된 방법을 사용하여 디자인되고, Cas9와 함께 미리 인큐베이션되고, 표적 DNA를 함유하는 샘플에 첨가된다. 이어서, 가이드 RNA 및 Cas9는 표적 DNA에 공동-국재화되어, 그와 복합체를 형성하게 될 것이다.

[0065] 통상의 기술자는 본 개시내용에 기초하여 표적 핵산을 비롯한 DNA에 공동-국재화되는 가이드 RNA 및 Cas9 단백질을 쉽게 확인 또는 디자인할 수 있을 것이다. 통상의 기술자는 추가로 직접적이든 또는 간접적이든, 그에 상관없이, 가이드 RNA 또는 Cas9 단백질에의 결합에 대하여 검출가능한 모이어티를 확인할 수 있을 것이다. DNA는 게놈 DNA, 미토콘드리아 DNA, 바이러스 DNA 또는 외인성 DNA를 포함한다.

[0066] 한 측면에 따라, 관심 서열에 특이적인 가이드 RNA를 디자인하다. gRNA를 Cas9와 함께 미리 인큐베이션시킨 후, 그 조합을 표적 DNA를 함유하는 샘플에 첨가하거나, 또는 다르게는 표적 DNA에 접촉시킨다. gRNA 또는 Cas9는 검출가능한 표지를 포함할 수 있거나, 또는 검출가능한 표지는 복합체 형성 이후에 첨가될 수 있다. 성분의 혼합물은 모두 용액으로 제공될 수 있거나, 또는 표적 핵산은 표면 상에 고정화될 수 있거나, 또는 세포 또는 조직 내에 존재할 수 있다.

[0067] 본 개시내용의 측면에 따라, 본원에 기술된 CRISPR Cas9 시스템은 (적절히 디자인된 경우) 서열 특이성을 갖는다는 이점을 가지며, 표적 서열은 gRNA 상의 17-25개의 뉴클레오티드 스페이서 서열을 통해 "프로그래밍"된다.

[0068] 특정 서열이 gRNA의 "시드" 영역에서 사용될 때, Cas9 시스템은 매우 효율적인 결합을 보이며; 이는 이러한 짧은 서열의 게놈에서 빈번하게 존재하는 것에 기초하여 게놈 맵핑 도구로서 사용될 수 있다.

[0069] Cas9 시스템은 또한 gRNA의 비-시드 영역에서 축중성 위치 (및/또는 범용 염기)를 사용함으로써 맵핑 도구가 되도록 조작될 수 있다.

[0070] 표지 특이성을 증가시키기 위해, 가이드 클러스터는 관심 유전자 주변에 결합될 수 있다.

[0071] 가이드 RNA는 (업체, 예컨대 IDT로부터 이용가능한) RNA의 직접적인 고체상 합성에 의해, 또는 고체상에서 합성된 DNA 올리고의 시험관내 전사에 의해 제조될 수 있다.

[0072] gRNA는 어레이에서 합성된 올리고 (이는 필요한 ~100-200 nt보다 더욱 긴 길이로 이용가능하다)로부터 쉽게 합성될 수 있고, 증폭될 수 있으며, 이에 의해 각 가이드의 비용이 매우 저렴해지고, 다수의 gRNA 생성 규모가 쉽게 조정가능해질 수 있다. 예를 들어, 커스텀 어레이 인크.(Custom Array Inc.)는 그의 장치 1회 작동으로 gRNA 생성에 적합한 90,000개의 어레이에서 합성된 올리고를 제공할 수 있다.

[0073] 반응 동역학적 성질은 등온성이고 (37°C, 가능하게는 실온), 신속하며, 1분 미만이 소요되고, 생성된 복합체는 매우 안정적이며, 쉽게 프로빙될 수 있다.

[0074] 표적 DNA는 용액 중의 또는 표면 상에 고정화된 멀크 DNA, 세포 계내에 있는 DNA, 표면 상의 염색체 스프레드 상의 DNA, 표면 상에서 또는 나노채널 중에서 신장된 단일 DNA 분자일 수 있다.

[0075] 다른 조작된 뉴클레아제, 예컨대 호밍(Homing) 엔도뉴클레아제 (HE), 메가뉴클레아제, 전사 활성인자-유사 이펙터 뉴클레아제(TALEN), 아연 핑거 뉴클레아제 (ZFN), 원핵성 아르고노트(Argonaute) (pAgo), 또는 부르 H(BurrH)-기반 뉴클레아제 (BuDN)가 Cas9 대신, 또는 그와 동시에 사용될 수 있다. 예로서, TtAGO는 RNA에 대해서는 높은 친화도를 갖고, dsDNA에 대해서는 낮은 친화도를 갖는다.

[0076] 특정 측면에 따라, DNA 결합 Cas9-sgRNA는 직접적으로 (예컨대, 양자점 또는 유기 염료에 결합한 친화성 태그를 통해) Cas9 단백질을 표지함으로써 검출될 수 있다. 상업적으로 이용가능한 Cas9 단백질은 이미 친화성 태그를 함유하고 있다 (PNA바이오 인크.(PNABio Inc.)로부터의 Cas9는 인간 인플루엔자 헤마글루티닌 (HA)을 함유하고; 뉴잉글랜드 바이오랩스(New England Biolabs)로부터 이용가능한 Cas9는 히스티딘 (His) 태그를 함유한다). DNA 결합 Cas9-sgRNA는 예컨대, gRNA의 3' 단부의 테일 부분에서 gRNA를 표지함으로써 검출될 수 있고, 여기서 테일 부분은 프로빙될 수 있다. 예를 들어, 형광 모이어티는 테일에 결합될 수 있고, 상이한 컬러의 프로브는 결합되어 코딩 스킴을 생성할 수 있고, 프로브는 교환을 통해 코드 레퍼토리를 증가시킬 수 있다. DNA-PAINT와 조합하였을 때, 초해상도 영상화가 달성될 수 있다. 플루이딕 장치와 조합하였을 때, 초해상도로 다중화가 가능한, EXCHANGE-PAINT가 수행될 수 있다. 이러한 스ქ에서, 각 사이클에서 같은 컬러가 사용되지만, 상이한 상이한 DNA PAINT 이미저 서열에 연결되는 것인 시약 교환 사이클을 수행함으로써 다수의 유전자좌의 초해상도 영상화를 위해 제한된 개수의 코드를 사용할 수 있다. 예를 들어, Cy3B, Atto 655를 포함하는 단 2개의 표지가 5x 회에 사용되며, 매회 상이한 이미저 서열과 커플링되며, 이는 10개의 gRNA를 코딩할 수 있는 능력을 갖는다. 매 사이클에서, gRNA의 하위세트가 표지된다. 사이클 완료 후, 컬러, 및 특정 gRNA가 발광하게 된 사이클 수를

측정함으로써 gRNA의 정체성을 디코딩한다.

[0077] 신호 강도를 증가시키기 위해, 다중의 형광단으로 표지된 올리고를 gRNA의 테일 부분에 결합시킬 수 있다. 대안적으로, 테일에 결합하고, 패드락 프로브로서 환형화되는 올리고를 사용함으로써 테일 3' 단부에 의해 프라이밍된 롤링 서클 증폭이 수행될 수 있다. 혼성화 연쇄 반응 또는 관련 기술분야의 통상의 기술자에게 공지된 다른 신호 증폭 방법 또한 사용될 수 있다.

[0078] 검출 방법은 형광 검출 방법, 전계발광 검출 방법, 화학발광 검출 방법, 생체발광 검출 방법 및 비색 검출 방법을 포함한다.

[0079] 형광, 전계발광성, 화학발광성, 생체발광성 또는 비색 모이어티, 또는 복합체를 검출하는 것을 포함하는 것 이외의 다른 검출 방법, 예컨대 Cas9-sgRNA 결합 DNA 가닥을 나노포어 또는 나노캡 또는 나노채널을 통해 통과시켜 Cas9-sgRNA의 결합 위치를 측정하는 것, 표면 상에서 연장/신장된 DNA에의 Cas9-sgRNA의 결합 위치 검출을 위해 전자 현미경법 또는 스캐닝 프로브 현미경법을 사용하는 것, 캠틸레버, 수정 진동자 저울, 전계 효과 트랜지스터 등을 사용하여 Cas9-sgRNA의 표적 DNA에의 결합을 검출하는 것도 사용될 수 있다.

[0080] 특정 측면에 따라, 표적 핵산에서의 gRNA 및 Cas9의 복합체 존재는 관련 기술분야의 통상의 기술자에게 공지된 나노포어 또는 나노캡 검출 기술 또는 나노포어 또는 나노캡 서열분석 방법을 사용하여 측정된다. 요약하면, 전기 전도성 매질 중에서 그에 gRNA 및 Cas9가 결합되어 있는 표적 핵산을 전압차 영향하에서 나노포어를 통과시킨다. 계면 의존적 이온 전류 변화를 사용하여 개별 뉴클레오티드와 핵산에 결합된 gRNA/Cas9 복합체 사이의 구별짓는다. 이와 관련하여, gRNA/Cas9 복합체의 존재가 검출될 수 있다. 한 측면에 따라, 계면 의존적 이온 전류 변화는 표적 핵산의 나노포어 또는 나노캡 내로의 진입을 측정하고, gRNA/Cas9 복합체가 표적 핵산에 결합하였는지 여부, 및 결합 위치 또는 위치등을 측정한다. 핵산이 선형 중합체로서 나노포어로 진입하였을 때, 포어 중 중합체의 물리적 존재가 포어를 통과하는 이온의 유동을 교란시키기 때문에, 이온 전류는 하락한다. gRNA/Cas9 복합체가 DNA 상의 특정 위치에 결합한다면, 이때 상기 위치가 포어로 진입할 때, 이온은 추가로 감소하여, 이온 전류는 감소하게 된다. 이러한 전류 감소는 gRNA/Cas9 복합체가 표적 핵산에 결합하고 있음을 나타낸다. 그의 크기 및 물리화학적 특성에 의존하여, 표적 핵산 (즉, DNA 중합체)에 결합된 각 유형의 구조 또는 복합체는 특징적인 이온 전류 변화를 나타내게 될 것이다. 상이한 위치 또는 대립유전자를 표적화하는 gRNA-Cas9 복합체는 상이하게 표지될 수 있으며, 이에 의해 상기 복합체는 나노포어 판독에 의해 구별될 수 있다.

[0081] "나노포어"란, 너비가 나노미터 규모인 구멍 또는 통로, 예컨대 평면 표면 또는 막을 통과하는 구멍 또는 통로를 의미한다. 나노포어는 예컨대, 지질 이중층에서 다량체 단백질 고리에 의해 형성될 수 있다. 나노포어는 질화규소, 그래핀 또는 상기 비-생체 물질의 고체 상태의 평면 표면에 있는 물리적 구멍일 수 있다. 전형적으로, 통로 너비는 0.2-25 nm이다. 본원에서 사용되는 바, 나노포어는 문자가 막을 통과할 수 있도록 허용할 수 있는 막관통 구조를 포함할 수 있다. 나노포어의 예는 α -용혈소 (*Staphylococcus aureus*) 및 MspA (미코박테리움 스메그마티스 *Mycobacterium smegmatis*)를 포함한다. 나노포어의 다른 예는 나노포어 서열분석을 기술하는 관련 기술분야에서 찾아볼 수 있거나, 또는 포어-형성 독소, 예컨대 β -PFT 팬던-발렌타인 류코시딘 S(Panton-Valentine leukocidin S), 에어로리신, 및 클로스트리듐 엡실론(*Clostridial Epsilon*)-독소, α -PFT 시토리신 A, 이진 PFT 탄저균 독소, 또는 다른 것, 예컨대 뉴몰리신 또는 그라미시딘으로서 관련 기술분야에 기술될 수 있다. 나노포어는 나노포어 서열분석 기술 출현과 함께 기술적으로 및 경제적으로 중요해지고 있다. 나노포어 서열분석 방법은 예를 들어, U.S.P.N. 5,795,782 (이는 참조로 포함된다)에 기술되어 있는 바와 같이, 관련 기술분야에 공지되어 있다. 요약하면, 나노포어 검출은 예컨대, 예를 들어, KCl, NaCl, NiCl, LiCl을 비롯한 이온 용액 또는 관련 기술분야의 통상의 기술자에 공지된 다른 이온 형성 무기화합물과 같은, 전압-전도 유체에 침지된 나노포어-천공 막을 포함한다. 전압은 막을 통과하도록 인가되고, 전류는 나노포어를 통과하는 이온 전도로부터 생성된다. 나노포어가 중합체, 예컨대 DNA와 상호작용할 때, 나노포어를 통과하는 유동은 임의의 주어진 시간에 포어를 통해 전위되는 중합체 하위단편의 특징에 따라 예컨대, 단량체 특이적 방식으로 조절되며, 이에 의해 전류는 변하게 되고, 이를 통해 단량체(들) 또는 하위단편을 확인할 수 있게 된다. 본 개시내용의 범주에 포함되는 나노포어는 관련 기술분야의 통상의 기술자에 공지된 고체 상태의 비단백질 나노포어, 및 관련 기술분야의 통상의 기술자에 공지된 DNA 오리가미 나노포어를 포함한다. 상기 나노포어는, 복합체가 나노포어를 통과할 때, 이온 전류 변화를 검출하는데 있어서는 여전히 충분한 감도를 가지면서, 검출을 위한 더 큰 문자, 예컨대 이중 가닥 표적 핵산을 포함하는 Cas9/gRNA 복합체가 통과할 수 있도록 허용하는, 공지된 단백질 나노포어보다 너비가 더 넓은 나노포어를 제공한다.

- [0082] "나노포어 분석"은 나노포어와 중합체의 상호작용에 기초하여, 중합체, 예컨대 gRNA/Cas9 복합체를 포함하는 폴리뉴클레오티드의 성분을 측정하는 방법을 의미한다. 나노포어 분석은 중합체와의 상호작용에 의해 개구의 크기가 변경될 때 발생하는, 나노포어를 통과하는 이온의 전도도 변화를 측정함으로써 달성될 수 있다.
- [0083] 나노포어 이외에도, 본 개시내용은 두 전극 사이의 갭인 것으로서, 여기서 갭은 너비가 수 나노미터, 예컨대 약 0.2 nm 내지 약 25 nm 또는 약 2 내지 약 5 nm인 것으로, 관련 기술분야에 알려져 있는 나노갭의 사용도 구상한다. 갭은 나노포어에 있는 개구를 모방하고, DNA가 갭을 통해, 또는 그 위를 및 전극 사이를 통과하도록 허용한다. 본 개시내용의 측면은 또한 나노채널의 사용도 구상한다. 전극은 DNA가 통과하는 나노채널에 인접하게 위치한다. 추가로, 또는 대안적으로, 복합체가 광학적으로 표지될 때, 나노채널에서 신장된 DNA 중합체를 따라 결합하는 복합체의 위치가 측정될 수 있다. 통상의 기술자는 전기장을 통과하는 문자 또는 모이어티의 이동 및 전기장을 통과하는 구조를 나타내는 전기장의 왜곡 생성에 기초하여 문자 또는 모이어티 확인 및 서열분석의 상이한 실시양태를 쉽게 구상할 수 있을 것이라는 것을 이해하여야 한다.
- [0084] 추가 측면에 따라, Cas9 낙카제는 표적 이중 가닥 핵산을 낙킹하는데 사용될 수 있고, 낙은 폴리머라제- 또는 리가제 기반 서열분석을 위한 서열-정의된 프라이밍 부위로서 사용될 수 있으며, 이에 의해 Cas9-sgRNA 표적 부위 주변의 서열 정보를 밝혀낼 수 있다. 게놈의 선택된 부분 다수, 예컨대 액솜, GWAS 신호에 의해 확인되는 영역, 심장 질환, 암 등과 연관된 특이적인 유전자를 서열분석할 수 있는, 가이드의 라이브러리가 디자인될 수 있다. 프라이머 연장 또한 예컨대, 형광 뉴클레오티드 도입에 의해, 라이게이션에 의해, 또는 형광 뉴클레오티드 도입 후, 라이게이션에 의해 낙킹 부위를 표지하는데 사용될 수 있다. 프라이머 연장이 현존 가닥을 치환하는 경우, 치환된 플랩은 그에의 올리고의 혼성화에 의해 표지될 수 있다. 예를 들어, 다중의 낙킹 부위가 존재한다면, 예컨대 가이드의 부분이 축중성 서열을 함유할 경우, 이때 표지의 다중 부위가 맵핑 도구로서 사용될 수 있다. 프라이머 연장 방법은 관련 기술분야의 통상의 기술자에게 공지되어 있다.
- [0085] 본원에 기술된 방법의 특정 적용은 확인 또는 진단 또는 맵핑 방법을 포함한다. 예컨대, 동원체 반복부와 같은 인간 게놈의 "암흑 물질"에의 접근은 도 6A에 제시되어 있다. 동원체 DNA 서열은 그의 반복적 성질이 매우 높기 때문에, 대개는 현 참조 게놈에는 존재하지 않는다. CRISPR/Cas9를 사용하는 본원에 기술된 단일 문자 맵핑 또는 서열분석 방법을 통해 표적화된 방식으로 고해상도로 반복 부위를 맵핑함으로써 참조 게놈을 더욱 완전하게 어셈블리할 수 있고, 이로써 상기 부위로부터 서열분석할 수 있다. 상기 방법은 개인 게놈과 함께 사용될 수 있다.
- [0086] 본 개시내용에 따른 방법은 표지된 클론을 사용하여 수행되는 혈행 방법보다 더 높은 효율 및 해상도로 염색체 스프레드 상에서 FISH를 수행하는 것을 포함한다. 표지된 gRNA/Cas9 복합체는 더욱 깨끗한 신호를 획득할 수 있도록 하는 바, 이는 본원에 기술된 예시적인 방법에서 효과적인 FISH 프로브이다. 본 방법을 통해 다수의 암 및 불임 문제와 관련이 있는 염색체 파단점, 복합체 전위 또는 재배열을 더욱 신속하게, 및 더욱 잘 확인할 수 있다.
- [0087] 본 개시내용에 따른 방법을 통해 시험관내 진단을 수행할 수 있다.
- [0088] 본 개시내용에 따른 방법을 통해 신속한 진단 플랫폼과 관련하여 박테리아, 또는 바이러스에서 다중 약물 내성을 코딩하는 유전자를 프로빙할 수 있고; 이와 관련하여 Cas9-sgRNA를 사용하여 표적 dsDNA를 직접 및 안정적으로 표적화할 수 있다.
- [0089] 하기 예시적인 방법 또한 본 개시내용에 의해 구상된다.
- [0090] Mg의 존재를 포함하는 조건하의 야생형 Cas9
- [0091] 일부 실시양태에서, 본 발명은 (a) 복합체가 핵산에 결합하고, 핵산을 절단하지만, 복합체가 핵산으로부터 쉽게 해리되지 않는 (즉, 절단된 두 가닥을 함께 잡고 있으면서, 핵산에 부착된 상태로 그대로 유지되는) 조건하에서 핵산을, 뉴클레아제 활성을 갖거나, 또는 그를 보유하는 Cas9 단백질 및 가이드 RNA를 포함하는 복합체와 접촉시키는 단계, 및 (b) 단계 (a)의 생성물을 분석하는 단계를 포함하는, 핵산 중 한 부위를 검출, 표지, 풀 다운 또는 표적화하는 방법을 포함한다. 일부 실시양태에서, 상기 조건은 2가 양이온, 예컨대 Mg²⁺의 존재를 포함한다. 일부 실시양태에서, 단계 (a)와 단계 (b) 사이에 적어도 1회의 세척 단계가 존재한다. 일부 실시양태에서, 생성된 복합체 중 적어도 하나의 성분을 표지하거나, 태그부착한다. 일부 실시양태에서, 가이드 RNA는 천연 PAM 부위에 인접한 서열에 결합한다. 일부 실시양태에서, 가이드 RNA는 인공 PAM 부위에 인접한 서열에 결합한다. 일부 실시양태에서, Cas9는 Cas9의 변경된 버전이다. 일부 실시양태에서, 가이드 RNA는 말단절단된 가이드 RNA이다. 일부 실시양태에서, 말단절단된 가이드는 단지 가이드 RNA의 시드 영역만을, 즉 PAM 부위에

인접한 4-7개의 뉴클레오티드만을 포함한다.

[0092] Mg 부재하의 아생형 Cas9

일부 실시양태에서, 본 발명은 (a) 복합체가 핵산에 결합하고, 핵산으로부터 쉽게 해리되지 않는 조건하에서 핵산을, 뉴클레아제 활성을 갖거나, 또는 그를 보유하는 Cas9 단백질 및 가이드 RNA를 포함하는 복합체와 접촉시키는 단계, 및 (b) 단계 (a)의 생성물을 분석하는 단계를 포함하는, 핵산 중 한 부위를 검출, 표지, 풀 다운 또는 표적화하는 방법을 포함한다. 일부 실시양태에서, 상기 조건은 2가 양이온, 예컨대 Mg²⁺의 부재를 포함한다. 일부 실시양태에서, 단계 (a)와 단계 (b) 사이에 적어도 1회의 세척 단계가 존재한다. 일부 실시양태에서, 생성된 복합체 중 적어도 하나의 성분을 표지하거나, 태그부착한다. 일부 실시양태에서, 가이드 RNA는 천연 PAM 부위에 인접한 서열에 결합한다. 일부 실시양태에서, 가이드 RNA는 인공 PAM 부위에 인접한 서열에 결합한다. 일부 실시양태에서, Cas9는 Cas9의 변경된 버전이다. 일부 실시양태에서, 가이드 RNA는 말단절단된 가이드 RNA이다. 일부 실시양태에서, 말단절단된 가이드는 단지 가이드 RNA의 시드 영역만을, 즉 PAM 부위에 인접한 4-7개의 뉴클레오티드만을 포함한다.

[0094] 널/사멸 Cas9

일부 실시양태에서, 본 발명은 (a) 복합체가 핵산을 절단하지 않고, 핵산으로부터 쉽게 해리되지 않는 조건하에서 핵산을, 효소적으로 불활성인 Cas9 단백질, 또는 뉴클레아제 널 Cas9 단백질, 예컨대 (예컨대, D10A/H840A dCas9) 및 가이드 RNA를 포함하는 복합체와 접촉시키는 단계, 및 (b) 단계 (a)의 생성물을 분석하는 단계를 포함하는, 핵산 중 한 부위를 검출, 표지, 풀 다운 또는 표적화하는 방법을 포함한다. 일부 실시양태에서 단계 (a)와 단계 (b) 사이에 적어도 1회의 세척 단계가 존재한다. 일부 실시양태에서, 생성된 복합체 중 적어도 하나의 성분을 표지하거나, 태그부착한다. 일부 실시양태에서, 가이드 RNA는 천연 PAM 부위에 인접한 서열에 결합한다. 일부 실시양태에서, 가이드 RNA는 인공 PAM 부위에 인접한 서열에 결합한다. 일부 실시양태에서, Cas9는 Cas9의 변경된 버전이다. 일부 실시양태에서, 가이드 RNA는 말단절단된 가이드 RNA이다. 일부 실시양태에서, 말단절단된 가이드는 단지 가이드 RNA의 시드 영역만을, 즉 PAM 부위에 인접한 4-7개의 뉴클레오티드만을 포함한다.

[0096] 낙카제 Cas9

일부 실시양태에서, 본 발명은 (a) 복합체가 핵산으로부터 쉽게 해리되지 않도록 하는 조건하에서 핵산을, Cas9 낙카제 (즉, Cas9 단백질의 낙킹 돌연변이체, 예컨대 예컨대, Cas9 D10A 또는 H840A 돌연변이체) 및 가이드 RNA를 포함하는 복합체와 접촉시키는 단계, 및 (b) 단계 (a)의 생성물을 분석하는 단계를 포함하는, 핵산 중 한 부위를 검출, 표지, 풀 다운 또는 표적화하는 방법을 포함한다. 일부 실시양태에서, 단계 (a)와 단계 (b) 사이에 적어도 1회의 세척 단계가 존재한다. 일부 실시양태에서, 생성된 복합체 중 적어도 하나의 성분을 표지하거나, 태그부착한다. 일부 실시양태에서, 가이드 RNA는 천연 PAM 부위에 인접한 서열에 결합한다. 일부 실시양태에서, 가이드 RNA는 인공 PAM 부위에 인접한 서열에 결합한다. 일부 실시양태에서, Cas9는 Cas9의 변경된 버전이다. 일부 실시양태에서, 가이드 RNA는 말단절단된 가이드 RNA이다. 일부 실시양태에서, 말단절단된 가이드는 단지 가이드 RNA의 시드 영역만을, 즉 PAM 부위에 인접한 4-7개의 뉴클레오티드만을 포함한다.

[0098] DNA 결합 단백질 부재하에서 가이드 RNA를 사용하는 방법

일부 실시양태에서, 본 발명은 (a) 가이드 RNA가 핵산으로부터 쉽게 해리되지 않도록 하는 조건하에서 핵산을 가이드 RNA와 접촉시키는 단계, 및 (b) 단계 (a)의 생성물을 분석하는 단계를 포함하는, 핵산 중 한 부위를 검출, 표지, 풀 다운 또는 표적화하는 방법을 포함한다. 일부 실시양태에서, 단계 (a)와 단계 (b) 사이에 적어도 1회의 세척 단계가 존재한다. 일부 실시양태에서, 생성된 가이드 RNA-DNA 복합체는 표지하거나, 태그부착한다. 일부 실시양태에서, 표지 또는 태그는 테일 상에 존재한다. 일부 실시양태에서, 가이드 RNA는 천연 PAM 부위에 인접한 서열에 결합한다. 일부 실시양태에서, 가이드 RNA는 인공 PAM 부위에 인접한 서열에 결합한다. 일부 실시양태에서, 가이드 RNA는 PAM 부위에 인접한 서열에의 결합을 필요로 하지 않는다. 일부 실시양태에서, 가이드 RNA는 말단절단된 가이드 RNA이다. 일부 실시양태에서, 말단절단된 가이드 RNA는 단지 가이드 RNA의 시드 영역만을, 즉 PAM 부위에 인접한 4-7개의 뉴클레오티드만을 포함한다. 일부 실시양태에서, 표적 유전자좌는 반복 DNA로 이루어진 영역이고, 신호는 증폭된다. 일부 실시양태에서, 표적 핵산은 가이드 RNA 첨가 이전에 변성된다.

[0100] RNA 사용 방법

일부 실시양태에서, 본 발명은 (a) RNA가 핵산에 결합할 수 있도록 하는 조건하에서 핵산을 RNA와 접촉시키는

단계, 및 (b) 단계 (a)의 생성물을 분석하는 단계를 포함하는, 핵산 중 한 부위를 검출, 표지, 풀 다운 또는 표적화하는 방법을 포함한다. 일부 실시양태에서, 단계 (a)와 단계 (b) 사이에 적어도 1회의 세척 단계가 존재한다. 일부 실시양태에서, 생성된 RNA-DNA 복합체를 표지하거나, 태그부착한다. 일부 실시양태에서, 표지 또는 태그는 테일 상에 존재한다. 일부 실시양태에서, 표적 유전자좌는 반복 DNA로 이루어진 영역이고, 신호는 증폭된다.

[0102] 이중체 탈안정화/개방 시약 및 RNA

일부 실시양태에서, 본 발명은 (a) 복합체가 형성되어 핵산에 결합하고, 복합체가 핵산으로부터 쉽게 해리되지 않는 조건하에서 핵산을 이중체 탈안정화/개방 시약 및 가이드 RNA 또는 RNA와 접촉시키는 단계, 및 (b) 단계 (a)의 생성물을 분석하는 단계를 포함하는, 핵산 중 한 부위를 검출, 표지, 풀 다운 또는 표적화하는 방법을 포함한다. 일부 실시양태에서, 단계 (a)와 단계 (b) 사이에 적어도 1회의 세척 단계가 존재한다. 일부 실시양태에서, 생성된 복합체 중 적어도 하나의 성분을 표지하거나, 태그부착한다. 일부 실시양태에서, 표지 또는 태그는 테일 상에 존재한다. 일부 실시양태에서, Cas9는 Cas9의 변경된 버전이다. 일부 실시양태에서, 가이드 RNA는 말단절단된 가이드 RNA이다. 일부 실시양태에서, 말단절단된 가이드는 단지 가이드 RNA의 시드 영역만을, 즉 PAM 부위에 인접한 4-7개의 뉴클레오티드만을 포함한다. 일부 실시양태에서, 표적 유전자좌는 반복 DNA로 이루어진 영역이고, 신호는 증폭된다. (단일 가닥 DNA 안정화 시약을 포함하는) 탈안정화 시약은 헬리카제, 복제 단백질 A (RPA), E. 콜라이(*E. Coli*) 단일 가닥 결합 단백질 (SSB), 베타인, 베타인/글리신, 포름아미드, 우레아, DMSO 등을 포함한다. 이중체 개방 시약은 프리모솜 단백질 PriA, 삼중체 형성 비스-PNA, 감마 PNA 등을 포함한다.

[0104] 시험관내 RNA 합성 및 프로빙

일부 실시양태에서, 본 발명은 (a) 무세포 시스템에서 RNA를 합성하는 단계, (b) RNA가 핵산에 결합할 수 있도록 하는 조건하에서 핵산을 RNA와 접촉시키는 단계, 및 (c) 복합체가 형성되고, 상기 복합체가 핵산으로부터 쉽게 해리되지 않는 조건하에서 핵산을 RNA 및 다른 성분과 접촉시키는 단계, 및 (d) 단계 (b)의 생성물을 분석하는 단계를 포함하는, 핵산 중 한 부위를 검출, 표지, 풀 다운 또는 표적화하는 방법을 포함한다. 일부 실시양태에서, 단계 (b)와 단계 (c) 사이에 적어도 1회의 세척 단계가 존재한다. 일부 실시양태에서, 생성된 복합체 중 적어도 하나의 성분을 표지하거나, 태그부착한다. 일부 실시양태에서, 표지 또는 태그는 테일 상에 존재한다. 일부 실시양태에서, Cas9는 Cas9의 변경된 버전이다. 일부 실시양태에서, 가이드 RNA는 말단절단된 가이드 RNA이다. 일부 실시양태에서, 말단절단된 가이드는 단지 가이드 RNA의 시드 영역만을, 즉 PAM 부위에 인접한 4-7개의 뉴클레오티드만을 포함한다.

[0106] 낙킹 및 서열분석

일부 실시양태에서, 본 발명은 (a) 복합체가 핵산의 한 가닥에서 낙을 유도하는 조건하에서 Cas9 단백질의 낙킹 돌연변이체 및 특이적이 위치를 표적화하는 가이드 RNA를 형성하는 복합체와 핵산을 접촉시키는 단계, (b) 뉴클레오티드를 이용하여 낙의 3' 단부를 연장시키는 단계, (c) 단계 (b)의 생성물을 검출하고, DNA를 서열분석하는 방식으로 단계 (b)를 반복하는 단계를 포함하는, 표적화된 서열분석을 포함한다. 일부 실시양태에서, 뉴클레오티드를 표지한다. 일부 실시양태에서, 뉴클레오티드를 말단 포스페이트에서 표지한다. 일부 실시양태에서, 뉴클레오티드는 3개 초과의 포스페이트를 함유한다. 일부 실시양태에서, 뉴클레오티드를 절단 가능한 연결부를 통해 염기에서 표지한다. 일부 실시양태에서, 뉴클레오티드는 가역성 종결인자이다. 일부 실시양태에서, 염기에 있는 표지는 가역적 종결을 제공한다. 일부 실시양태에서, 당의 3' 또는 2'번 위치에서의 변형은 가역적 종결을 제공한다. 일부 실시양태에서, 4가지 뉴클레오티드 모두가 동시에 연장을 위해 이용될 수 있다. 일부 실시양태에서, 단일 문자 검출 방법이 사용된다. 일부 실시양태에서, 단일 문자는 선형 스트링으로서 분석된다. 일부 실시양태에서, 핵산은 세포 계내에서 분석된다. 일부 실시양태에서, 핵산이 세포 계내에서 분석되는 경우, 세포는 분석 이전에 고정된다. 일부 실시양태에서, 계내에서 분석되는 핵산은 DNA 문자이고, RNA 문자는 분석 이전에 제거된다. 일부 실시양태에서, 낙킹 이후 시약을 사용하여 핵산으로부터 복합체를 제거한다. 일부 실시양태에서, 단계 (b)와 단계 (c) 사이에 적어도 1회의 세척 단계가 존재한다. 일부 실시양태에서, Cas9는 Cas9의 변경된 버전이다. 일부 실시양태에서, 가이드 RNA는 말단절단된 가이드 RNA이다. 일부 실시양태에서, 말단절단된 가이드는 단지 가이드 RNA의 시드 영역만을, 즉 PAM 부위에 인접한 4-7개의 뉴클레오티드만을 포함한다.

[0108] 낙킹 및 포획

[0109]

일부 실시양태에서, 변형된 뉴클레오티드를 사용하여 닉의 3' 단부를 연장한다. 일부 실시양태에서, 변형된 뉴클레오티드는 비오틴 변형된 것이다. 일부 실시양태에서, 도입된 비오틴을 사용하여 예컨대, 그 자체가 표지된 (또는 표지되는) 스트랩트아비딘/뉴트라비딘 또는 항-비오틴 항체와의 상호작용을 통해 표적 DNA를 표지한다. 일부 실시양태에서, 도입된 비오틴을 사용하여 예컨대, 비오틴 코팅된 포획 물질, 예컨대 자기 또는 아가로스 비드에, 또는 표면에 그 자체가 결합된 (또는 결합되는) 스트랩트아비딘/뉴트라비딘 또는 항-비오틴 항체와의 상호작용을 통해 표적 DNA를 포획한다. 일부 실시양태에서, Cas9/gRNA 유도 낙킹 및 포획을 지원하도록 변형된 염기의 도입을 사용하여 용액으로 수행되는 반응물 중의 핵산 샘플의 특이적인 단일 또는 다중 부분을 단리시킨다. 본 실시양태에서, 예를 들어, 낙킹 및 비오틴화된 dUTP의 도입 후, 생성물을, 게놈의 비오틴화된 부분이 비드 상의 스트랩트아비딘에 결합할 수 있는 기간 (예컨대, 1시간) 동안 스트랩트아비딘 코팅된 자기 비드와 반응시킨다. 이어서, 자석을 가하고, 고체상 비드에 부착된 게놈의 표적 부분을 상청액으로부터 분리시켜 관심 표적화된 영역을 단리시킨다. 상청액을 폐기한다. 다양한 정도의 업격도로 세척하고, 상청액을 제거한 후, 관련 기술분야에 공지된 방법에 의해 (예컨대, 90°C 초과의 온도로 가열하여) 선택된 게놈 DNA를 비드로부터 분리한다. 이어서, 크기 선별, 폴리싱, 바코딩, 테일링, 라이브러리 제조, 클러스터 증폭, 로로니(rolony) 증폭 등으로부터 선택될 수 있는 단계를 포함할 수 있는, 관련 기술분야의 통상의 기술자에게 공지된 차세대 서열분석 방법 및 장치에 의해 포획된 문자를 서열분석할 수 있다. 핵산 샘플 중 일부 또는 그의 하위세트를 선별하거나, 또는 강화시키기 위한, 본원에 기술된 낙킹 및 도입 접근법은, 연장을 통해 다중의 비오틴이 도입될 수 있으며, 이에 의해 포획 효율이 개선될 수 있는 한, 가이드 RNA/Cas9의 결합보다 이롭다. 또 다른 이점은 오프-타겟 포획이 감소된다는 점인데, 이는 gRNA/cas9 결합, 낙킹 및 연장이라는 3 단계는 표적이 포획 가능하게 되기 이전에 필요하기 때문이다. 상기 선별 방법은 현준 접근법 (예컨대, 슈어셀렉트(Sureselect))보다 깨끗하고, 그 결과, 오프-타겟 서열은 더 적기 때문에, 서열분석은 더 적게 수행된다. 일부 실시양태에서, 반응, 예컨대 라이케이션 반응, 예컨대 예로서, 비오틴화된 올리고뉴클레오티드를 라이케이션시키는 것과 같은 반응은 닉의 5' 단부에서 수행된다. 일부 실시양태에서, DNA 이중체의 센스 및 안티센스 가닥은 분리될 수 있다. 본 측면에 따라, 한 가닥은 gRNA/Cas9 결합을 통해, 또는 비오틴화된 뉴클레오티드의 도입을 통해 포획되고, 나머지 다른 한 가닥은 상청액으로부터 수집된다. 대안적으로, gRNA 결합에 의해 치환된 이중체의 가닥은 단일 가닥 결합 단백질에, 히드록시아파타이트에 또는 서열 특이적인 올리고뉴클레오티드에의 결합에 의해 포획될 수 있다.

[0110]

낙킹 및 지정 서열분석

[0111]

일부 실시양태에서, gRNA 유도 낙킹은 표면에 부착된 표적 핵산, 즉 DNA를 사용하여 수행되거나, 또는 대안적으로, 낙킹 수행 후, DNA는 표면에 부착된다. 일부 실시양태에서, 이어서 닉의 3' 단부를 사용하여 DNA 서열분석을 개시할 수 있다. 닉의 3' 단부를 통해 폴리머라제 기반 서열분석, 예컨대 합성에 의한 일루미나 서열분석이 이루어질 수 있다. 닉의 3' 및 5' 단부, 둘 모두 라이케이션 기반 서열분석, 예컨대 SOLID (라이프 테크놀로지스(Life Technologies)) 및 라이케이션에 의한 서열분석 (컴플리트 게노믹스 인크.(Complete Genomics Inc.))를 지원한다. 일부 실시양태에서, 게놈 DNA는 장쇄로 유지되고, 낙킹은 게놈 DNA가 표면에 부착되기 이전 또는 그 이후에 수행된다. 일부 실시양태에서, DNA는 표면에 부착되고, 신장 또는 연장되며, 이에 의해 서열 또는 그의 길이에 따른 특징이 분석될 수 있다. 상기 분석은 광학, 전자, X선, 또는 스캐닝 프로브 현미경을 통해 수행될 수 있다. 일부 실시양태에서, 분석이 광학적 방법을 통해 수행되는 경우, 표면에 배치된 폴리뉴클레오티드에 대한 서열분석 반응에 대하여 내부 전반사 또는 소멸과/도파관 영상화가 수행된다. 일부 실시양태에서, 폴리뉴클레오티드는 표면에 부착되지 않고, 선형화될 수 있다. 일부 실시양태에서, 선형화는 한쪽 단부는 부착되어 있고, 유동 스트립에 매달려 있는 폴리뉴클레오티드에 의해 이루어진다. 일부 실시양태에서, 표적 핵산, 즉 DNA는 유체 역학적 항력을 통해 실질적으로 선형으로 제조된다. 일부 실시양태에서, DNA는 나노슬릿, 나노채널, 또는 나노그루브에 배치됨으로써 나노컨파인먼트에 의해 신장된다. 일부 실시양태에서, 선형 DNA는 실질적으로 직쇄이다. 일부 실시양태에서, gRNA 유도 낙은 합성에 의한 서열분석이 긴 선형 폴리뉴클레오티드 상의 다중의 선택된 서열 위치에서 수행되도록 유도한다. 본 실시양태에서, 폴리머라제 효소 (예컨대, 9 디그리 노쓰(9 Degree North) 또는 그의 돌연변이체, 또는 Phi29 또는 그의 돌연변이체)는 합성에 의한 서열분석에서 검출가능한 뉴클레오티드를 도입함으로써 닉으로부터 연장된다. 본 검출은 (이온 토렌트(Ion Torrent) 서열분석에서와 같이) pH를 통해 이루어질 수 있다. 일부 실시양태에서, 검출은 뉴클레오티드 상에서 형광 표지를 통해 이루어진다. 일부 실시양태에서, 뉴클레오티드 뿐만 아니라, 형광 표지된 것 또한 가역성 종결인자로서 작용하며, 이를 통해 관련 기술분야에 공지된 방법 (예컨대, 일루미나 또는 레이저젠(LaserGen) 서열분석)을 통해 4색 서열분석이 단계적으로 수행될 수 있다. 도입되는 뉴클레오티드는 라이트닝 종결인자 (레이저젠)일 수 있으며, 여기서 광절단가능한 모이어티의 절단은 UV 광에 의해 이루어진다. 일부 실시양태에서, gRNA/Cas9 복합체 형성

이후의 서열분석은 선택적 또는 표적화된 서열분석을 수행하는 것을 목적으로 한다. 다른 실시양태에서, gRNA는 축중성 위치 중 적어도 일부를 함유하고, 게놈 상에 분포되어 있는 다중의 출발 부위로부터 서열분석이 개시되고, 이는 특정 유전자좌에 대해 선택적인 것은 아니다. 일부 실시양태에서, DNA가 세포 내부에 있는 동안 낙킹이 일어난다. 일부 실시양태에서, 세포는 고정된다. 예를 들어, 낙킹은 형광 계내 서열분석 (FISSEQ) 반응의 일부를 형성할 수 있으며, 여기서 DNA는 서열분석된다. 일부 실시양태에서, 낙킹은 세포 내부의 게놈 DNA에서 유도 방식으로 낙킹한 후, 예컨대 가타 치환 폴리머라제, 예컨대 Phi29에 의한 분지형 또는 룰링 서클 증폭을 통해 이루어지는 낙에 인접한 영역의 증폭을 개시하는데 사용된다. 이어서, 증폭된 생성물에서의 서열분석이 수행된다. 대안적으로, 낙킹된 게놈 DNA는 단일 문자 서열분석 방법, 예컨대 TrueSeq (헬리코스 바이오 (Helicos Bio)/SeqLL)를 사용하여 낙으로부터 직접적으로 서열분석된다.

[0112] 결합 및 나노포어 분석

[0113] 일부 실시양태에서, 본 발명은 (a) 복합체가 핵산으로부터 쉽게 해리되지 않는 조건하에서 Cas9 단백질 및 가이드 RNA를 포함하는 복합체와 핵산을 접촉시키는 단계, (b) 핵산을 나노포어 또는 나노캡을 통과시키는 단계, 및 (c) 결합 위치를 분석하는 단계를 포함하는, 핵산 중 서열의 위치 또는 결합 부위를 검출하는 방법을 포함한다. 일부 실시양태에서, 단일 문자 검출 방법이 사용된다. 일부 실시양태에서, 단일 채널 기록이 사용된다. 일부 실시양태에서, 단일 문자는 선형 스트링으로서 분석된다.

[0114] 오프-타겟 결합

[0115] 일부 실시양태에서, 본 발명은 (a) 복합체가 핵산으로부터 쉽게 해리되지 않는 조건하에서 Cas9 단백질 및 가이드 RNA를 포함하는 복합체와 핵산을 접촉시키는 단계, (b) 결합 위치를 검출하는 단계, (c) 표적화된 결합 위치를 측정하는 단계, (d) 오프-타겟 결합 위치를 측정하는 단계, 및 (e) 오프-타겟 결합 위치를 통해 오프-타겟 결합의 서열의 정체성을 측정하는 단계를 포함하는, 가이드 RNA 오프-타겟 결합 부위를 검출하는 방법을 포함한다. 일부 실시양태에서, 표지를 사용하여 어떤 표적 또는 오프-타겟 결합이 측정될 수 있는지에 대해 참조에 의한 랜드마크를 제공한다. 일부 실시양태에서, 표지는 핵산 상에 물리적 맵을 생성하는 결합 시약을 포함한다. 일부 실시양태에서, 결합 시약은 비-무작위(promiscous) gRNA, 제한 효소, 낙카제 효소, 올리고뉴클레오티드 등, 그 중 하나 이상의 것일 수 있다. 일부 실시양태에서, 단일 문자 검출 방법이 사용된다. 일부 실시양태에서, 단일 문자는 선형 스트링으로서 분석된다.

[0116] 카피수 검출

[0117] 일부 실시양태에서, 본 발명은 (a) 표적 핵산 서열을, 그의 카피수를 측정하고자 하는 것인 염색체 또는 게놈 영역에 상보적인 부분을 갖는 가이드 RNA 서열 및 Cas9 단백질, 및 참조 염색체 및/또는 게놈 영역 및 Cas9 단백질과 접촉시키는 단계, (b) 그의 카피수를 측정하고자 하는 것인 염색체/게놈 영역으로부터의 신호 대 참조 염색체 또는 게놈 영역으로부터의 신호의 비율을 수득하는 단계를 포함하는, 염색체 또는 게놈 영역의 카피수를 측정하는 방법을 포함한다. 일부 실시양태에서, 본 방법은 이수성 검출에 적용된다. 일부 실시양태에서, 이수성은 삼염색체성 21이다. 일부 실시양태에서, 그의 카피수를 측정하고자 하는 것인 유전자좌는 LSI 21q22.13-q22.2이다.

[0118] Her2

[0119] 일부 실시양태에서, 본 발명은 (a) 표적 핵산 서열을, Her2 유전자좌에 상보적인 부분을 갖는 가이드 RNA 서열 및 Cas9 단백질, 및 참조 유전자좌에 상보적인 부분 및 Cas9 단백질과 접촉시키는 단계, 및 (b) 참조 유전자좌 대비 Her2 유전자좌로부터의 신호의 비율을 수득하는 단계를 포함하는, Her2 증폭 정도를 측정하는 방법을 포함한다.

[0120] 유전자 융합

[0121] 일부 실시양태에서, 본 발명은 (a) 표적 핵산 서열을, 제1 게놈 유전자좌에 상보적인 부분을 갖는 가이드 RNA 프로브 서열 및 Cas9 단백질, 및 제2 게놈 유전자좌에 상보적인 부분을 갖는 가이드 RNA 프로브 서열 및 Cas9 단백질을 접촉시키는 단계, (b) 제1 및 제2 유전자좌 사이의 공동-국재화 이벤트를 검출하는 단계이며, 여기서 유전자 융합은 게놈 영역 사이의 임의의 융합인 것인 단계를 포함하는, 유전자 융합의 존재를 측정하는 방법을 포함한다. 한 측면에 따라, 공동-국재화를 통해 서로 인접해 있는 프로브를 얻게 된다.

[0122] 파단 검정법

[0123] 일부 실시양태에서, 본 발명은 (a) 표적 핵산 서열을, 게놈 유전자좌 (제1 유전자좌)에 상보적인 부분을 갖는

가이드 RNA 서열 및 Cas9 단백질, 및 (참조에 따라) 인접 게놈 유전자좌 (제2 유전자좌)에 상보적인 부분을 갖는 가이드 RNA 프로브 서열 및 Cas9 단백질을 접촉시키는 단계, 및 (b) 제1 및 제2 유전자좌 사이의 공동-국재화 이벤트가 검출되지는지 여부를 측정하는 단계를 포함하는, 유전자 융합의 존재를 측정하는 방법을 포함한다. 일부 실시양태에서, 본 방법은 역형성립프종 키나제, 예컨대 ALK에 적용된다 (도 12 참조). 일부 실시양태에서, 본 방법은 ROS1에 적용된다. ROS1은 인슐린 수용체 패밀리의 수용체 티로신 키나제이다. 일부 실시양태에서, 검정법은 비-소세포 폐암 진단에 적용된다.

[0124] 게놈 재배열

일부 실시양태에서, 본 발명은 (a) 게놈 DNA 샘플을, 각각이 게놈 영역의 특이적인 하위영역에 상보적인 부분을 갖는 것인 다중의 가이드 RNA 서열 및 Cas9 단백질과 접촉시키는 단계이며, 여기서 각 하위영역에 대한 gRNA는 그를 다른 하위영역에 대한 gRNA로부터 구별될 수 있게 하는 인코딩을 포함하는 것인 단계, 및 (b) 게놈 DNA를 영상화하고, 코드를 디코딩하고, 참조 기준으로 코드 순서를 비교하는 단계이며, 여기서 대략적인 길이의 관심 게놈 영역에 걸쳐 코드의 공동-국재화가 이루어져 있는 것인 단계를 포함하는, 게놈 영역 간의 재배열을 검출하는 방법을 포함한다. 일부 실시양태에서, gRNA는 테일에서 코딩된다. 일부 실시양태에서, 코드의 디코딩은 테일의 코딩된 부분을 디코더 문자, 예컨대 DNA 또는 단백질 프로브를 포함하는 디코더 문자와 접촉시킴으로써 수행된다. 일부 실시양태에서, 게놈 재배열 측정 이외에도, 특정 게놈 세그먼트에 대한 상이한 대립유전자 또한 구별된다. 일부 실시양태에서, 상이한 코드는 상이한 대립유전자에 대해서 뿐만 아니라, 상이한 게놈 세그먼트에 대해서도 사용될 수 있다. 일부 실시양태에서, 관심 영역은 게놈의 BRCA1 및/또는 BRCA2 영역(들)이다. 일부 실시양태에서, 관심 영역은 MHC 또는 HLA 영역이다. 일부 실시양태에서, 관심 영역은 MMR 유전자, MLH1-PMS2 및 MSH2-EPCAM-MSH6 주변의 영역이고, 유전성 비용종증 결장직장암 (HNPPCC) 진단 또는 분석에서 사용될 수 있다. 일부 실시양태에서, 다중 gRNA는 각각의 게놈 세그먼트에 대해 사용될 수 있고, 상기 다중의 가이드 RNA는 각각 같은 코드로 표지된다.

[0126] 반복부 개수 열거

일부 실시양태에서, 게놈 DNA의 가닥은 선형 스트링으로서 분석된다. 일부 실시양태에서, DNA는 신장된다. 일부 실시양태에서, 나노포어/나노캡 분석이 수행된다. 일부 실시양태에서, 본 방법은 인간 염색체 4 내지 10 위의 3.3 kb-D4Z4 반복부 함유 유전자좌 상의 반복 단위 횟수를 열거하는데 사용된다. 일부 실시양태에서, D4ZA 영역 평가는 안면 견갑 상완 근위축증 (FSHD) 환자를 진단 또는 분석하는데 사용된다. 일부 실시양태에서, 본 방법은 텔로미어 하위유닛 반복부 개수를 열거하는데 사용된다. 일부 실시양태에서, 본 방법은 동원체 반복부 개수를 열거하는데 사용된다. 일부 실시양태에서, 본 방법은 주요 위성 반복부 개수를 열거하는데 사용된다. 일부 실시양태에서, 본 방법은 부차 위성 반복부 개수를 열거하는데 사용된다.

[0128] Cas9/가이드 RNA 계내 혼성화

특정 측면에 따라, Cas9 매개 계내 혼성화를 수행하는 방법은 표적 핵산 서열을 관심 염색체 또는 게놈 영역에 상보적인 부분을 갖는 가이드 RNA 서열 및 Cas9 단백질과 접촉시키는 단계를 포함하고, 여기서 가이드 RNA 및 Cas9 단백질은 표적 핵산 서열에 공동-국재화되어 복합체를 형성하고, 여기서 관심 염색체 또는 게놈 영역은 유동 셀에 배치되고, 여기서 계내 혼성화용 시약 및 세척 시약은 관심 염색체 또는 게놈 영역 위로 유동한다. 특정 측면에 따라, 복합체 위치는 형광, 화학발광, 전계발광, 비색 검출 등을 포함하는 군으로부터의 방법을 사용하여 검출된다.

[0130] 대립유전자-특이적 검출

일부 실시양태에서, 본 발명은 (a) 표적 핵산 서열을, 검출하고자 하는 대립유전자에 상보적인, 그의 시드 부분 (PAM 부위에 인접해 있는 처음 4-7개의 뉴클레오티드)을 갖는 가이드 RNA 서열 및 Cas9 단백질과 접촉시키는 단계, 및 (b) 핵산과 공동-국재화된 gRNA/Cas9의 존재를 검출하는 단계를 포함하는, 특이적인 대립유전자의 존재를 측정하는 방법을 포함한다.

[0132] 결합 검정법

특정 측면에 따라, (a) 표적 핵산 서열을, 표적 핵산 서열에 상보적인 부분을 갖는 가이드 RNA 서열 및 Cas9 단백질과 접촉시키는 단계, (b) 표면 상의 한 위치에서 복합체를 포획하는 단계, (c) 표적 핵산 서열 상이 아닌, 오직 복합체에만 존재하는 표지를 통해 상기 위치에서 포획된 복합체를 검출하는 단계를 포함하는, 특이적인 서열을 검출하는 진단 방법을 제공한다. 특정 측면에 따라, (a) 표적 핵산 서열을, 표적 핵산 서열에 상보적인 부분을 갖는 가이드 RNA 서열 및 Cas9 단백질과 접촉시키는 단계, (b) 표면 상의 한 위치에서 복합체를 포획하

는 단계, (c) cas9/gRNA 상이 아닌, 오직 복합체에만 존재하는 표지를 통해 상기 위치에서 포획된 복합체를 검출하는 단계를 포함하는, 특이적인 서열을 검출하는 진단 방법을 제공한다. 일부 실시양태에서, 본 검정법은 측면 유동 검정법, 딥스티 검정법, 페이퍼 마이크로플루이딕스 검정법, 도트 블록 검정법, 마이크로어레이 검정법의 일부로서 수행된다. 일부 실시양태에서, 검정법은 진단 검정법이다.

[0134] 면역조직화학법 및 gRNA/Cas 9 계내 혼성화

일부 실시양태에서, 본 발명은 (a) 표적 염색질 내의 표적 핵산 서열을, 표적 핵산 서열에 상보적인 부분을 갖는 가이드 RNA 서열 및 Cas9 단백질과 접촉시키는 단계이며, 여기서 가이드 RNA 및 Cas9 단백질은 표적 핵산 서열에 공동-국재화되어 복합체를 형성하는 것이 단계, (b) 표적 염색질을 단백질 결합 시약과 접촉시키는 단계, 및 (c) 가이드 RNA/cas9 복합체 및 단백질 결합 시약의 비교 위치를 검출하는 단계를 포함하는, 동일한 샘플 상에서 면역조직화학법 (IHC) 및 gRNA/cas9 매개 계내 혼성화 (ISH)를 조합하는 방법을 포함한다. 일부 실시양태에서, 단백질 결합 시약은 항체이다. 일부 실시양태에서, 단백질 결합 시약은 암타мер이다. 일부 실시양태에서, 가이드 RNA/Cas9 복합체 및 단백질 결합 시약은 차별적으로 표지된다. 일부 실시양태에서, IHC 시약 및 가이드 RNA/Cas9 ISH 시약은 함께 첨가된다. 일부 실시양태에서, IHC 시약 및 가이드 RNA/Cas9 ISH 시약은 연속적으로, 즉 차례로 첨가된다. 일부 실시양태에서, 가이드 RNA/Cas9 ISH가 변성 단계를 필요로 하지 않기 때문에, 단백질 및 염색질 구조는 무손상 상태 그대로 유지되고, IHC는 수행될 수 있다. 일부 실시양태에서, 가이드 RNA/Cas9 복합체는 염색질의 특정 부분을 단리시키는데 사용되고, 단리된 염색질에 존재하는 단백질은 분석 방법을 사용하여 검출된다.

[0135] 가이드 RNA 테일 프로빙

일부 실시양태에서, 본 발명은 (a) 표적 핵산 서열을, 표적 핵산 서열에 상보적인 부분을 갖는 가이드 RNA 서열 및 Cas9 단백질과 접촉시키는 단계이며, 여기서 가이드 RNA 및 Cas9 단백질은 표적 핵산 서열에 공동-국재화되어 복합체를 형성하고, 여기서 가이드 RNA는 3' 테일 서열을 포함하고, 상기 테일은 프로브 서열에 상보적이거나, 또는 프라이머로서 작용할 수 있는 것인 단계, 및 (b) 복합체를 검출함으로써 표적 핵산 서열을 검출하는 단계를 포함하는, 표적 핵산 서열을 검출하는 방법을 포함한다. 특정 측면에 따라, 테일은 gRNA 결합 위치에 인접해 있는 서열에 상보적인 서열을 포함한다. 특정 측면에 따라, 테일은 이중체의 치환된 가닥에 상보적인 서열을 포함한다. 특정 측면에 따라, 이중체의 표적 가닥 중 하나는 가이드 RNA에 의해 격리되고, 나머지 다른 한 가닥은 개방된 상태로 다른 시약과 결합하게 된다. 예시적인 시약은 단일 가닥 결합 단백질, 표지될 수 있는 상보적인 올리고뉴클레오티드, 또는 상기 가닥에 상보적인 테일의 일부분을 포함할 수 있다. 한 측면에 따라, 테일은 DNA PAINT를 위한 도킹 부위 또는 핸들을 포함한다. 일부 실시양태에서, 나머지 다른 한 가닥과의 Cas9/가이드 RNA 복합체는 안정화된다. 일부 실시양태에서, 단일 가닥 결합 단백질, 예컨대 RPA 또는 결합 올리고뉴클레오티드 또는 그의 유사체/모방체의 치환된 가닥에의 결합이 나머지 다른 한 가닥과의 Cas9/가이드 RNA 복합체를 안정화시킬 수 있다. 일부 실시양태에서, 안정화 효과는 천연 이중체의 재-지핑으로부터의 경쟁이 감소되었기 때문이다.

[0136] PAM 특이성이 변경된 Cas9

일부 실시양태에서, 본 발명은 (a) 복합체가 핵산에 결합할 수 있도록 하는 조건하에서 핵산을, 비-정규 PAM 서열에 인접한 부분에 결합하도록 변경된 Cas9 단백질, 및 가이드 RNA를 포함하는, 표지된 복합체와 접촉시키는 단계, 및 (b) 단계 (a)의 생성물을 분석하는 단계를 포함하는, 핵산 중 부위를 표지 또는 표적화하는 방법을 포함한다. 한 측면에 따라, 임의적으로 단계 (a)와 단계 (b) 사이에 적어도 1회의 세척 단계가 존재하고, gRNA의 결합을 촉진시키기 위해 임의적으로 보조 시약이 제공된다.

일반적으로, 본원에 기술된 실시양태의 경우, 가이드 RNA 및 RNA는 관련 기술분야의 통상의 기술자에게 공지된 변형된 RNA 뉴클레오티드를 포함할 수 있다. 일반적으로, 본원에 기술된 실시양태의 경우, 가이드 RNA 및 RNA는 RNA/DNA 키메라 또는 RNA/PNA 키메라를 포함할 수 있다. 일부 실시양태에서, 가이드 RNA 또는 RNA는 무세포 시스템에서 제조된다. gRNA 또는 RNA를 제조하는 방법은 관련 기술분야의 통상의 기술자에게 공지된 시험관내 전사 및 자동 화학적 RNA 합성 방법을 포함한다. 일부 실시양태에서, gRNA 또는 Cas 단백질 또는 보조 단백질은 세포 시스템에서 발현되고, 무세포 시스템에서 정제된다. 일부 실시양태에서, gRNA 또는 RNA는 무세포 시스템에서 Cas 단백질 또는 보조 단백질과 복합체를 형성한다.

하기 실시예는 본 개시내용을 대표하는 것으로 기술된다. 상기 실시양태 및 다른 등가의 실시양태는 본 개시내용, 도면 및 첨부된 청구범위에 비추어 자명해지게 되는 바, 본 실시예는 본 개시내용의 범주를 제한하는 것으로

로 해석되지 않아야 한다.

[0142] 실시예 I

[0143] 프로토콜

[0144] gRNA 합성 프로토콜:

[0145] 하기와 같이 PCR 어셈블리를 수행한다.

[0146] · 12 $\mu\ell$ 의 Q5 DNA 폴리머라제 2X 마스터 믹스 (NEB)

[0147] · 3 $\mu\ell$ 의 10 μM T7 정방향 프라이머

[0148] · 3 $\mu\ell$ 의 10 μM 바코드 역방향 프라이머

[0149] · 3 $\mu\ell$ 의 10 μM Sp.gRNA.split60 (정방향)

[0150] · 3 $\mu\ell$ 의 10 μM gRNA.엔드 (역방향)를 포함하는 반응 믹스를 제조한다.

[0151] PCR 장치에서의 사이클 조건:

[0152] 1. 30초 동안 98°C

[0153] 2. 10 sec 동안 98°C

[0154] 3. 20 sec 동안 52°C

[0155] 4. 15 sec 동안 72°C

[0156] 5. 2 단계부터 29 사이클 반복

[0157] 6. 2 min 동안 72°C

[0158] 7. 4°C에서 유지.

[0159] 스판 칼럼 (자이모(Zymo)) 상에서 DNA를 정제한다. 전형적으로, ~1 μg 의 dsDNA 주형을 수득한다.

[0160] 시험관내 전사 (IVT)를 하기와 같이 수행한다:

[0161] · 5.8 $\mu\ell$ RNase-무함유 물

[0162] · 2.5 $\mu\ell$ 암플리스크라이브 T7-플래쉬(AmpliScribe T7-Flash) 10X 반응 완충제 (일루미나(Illumina))

[0163] · 1.8 $\mu\ell$ 100 mM ATP

[0164] · 1.8 $\mu\ell$ 100 mM CTP (+2 uL Cy3-dUTP)

[0165] · 1.8 $\mu\ell$ 100 mM GTP

[0166] · 1.8 $\mu\ell$ 100 mM UTP

[0167] · 2 $\mu\ell$ 100 mM DTT

[0168] · 0.5 $\mu\ell$ 리보가드(RiboGuard) RNase 억제제

[0169] · 5 $\mu\ell$ DNA 주형

[0170] · 2.0 $\mu\ell$ 암플리스크라이브 T7-플래쉬 효소 용액을 포함하는 반응 믹스를 제조한다.

[0171] 가이드 RNA 또한 변형된 NTP를 이용하여 합성할 수 있으며, 이 경우, 하기와 같은 변형이 수행된다:

[0172] · 변형시키고자 하는 NTP 및 변형된 NTP를 1:1 몰비로 (예컨대, 0.9 $\mu\ell$ UTP 및 0.9 $\mu\ell$ UTP-Cy3) 첨가한다.

[0173] PCR 장치에서 37°C에서 2h 내지 16h 동안 인큐베이션시킨 후, 4°C에서 유지시킨다. 스판 칼럼 (자이모) 상에서 RNA를 정제한다. 전형적으로, ~100 μg 의 RNA (즉, 바코드가 있는 gRNA)를 수득한다.

[0174] Cas9-gRNA 복합체 어셈블리 프로토콜:

[0175] 참조로, 테일 상에 단일 바코드가 있는 1 μg 의 gRNA는 25 pmol의 RNA이다. Cas9 단백질 및 gRNA를 전형적으로

1:1 몰비로 혼합하여 Cas9-gRNA 복합체를 형성한다. 야생형 Cas9, 닉카제 Cas9 및 뉴클레아제 널 또는 사멸 Cas9의 복합체 형성을 위해 동일한 반응 조건이 사용될 수 있다. 야생형 Cas9가 사용되는 경우, MgCl₂은 생략될 수 있는데, 이 목적은 절단을 막기 위함이다. 야생형 Cas9, 뉴클레아제 널 Cas9 및 Cas9 닉카제를 발현시키기 위한 플라스미드는 애드진(Addgene)으로부터 이용가능하다. 플라스미드를 적합한 숙주에서 발현시킬 수 있고, 관련 기술분야에 공지된 방법에 의해, 예컨대 발현된 단백질에서 His 태그를 사용함으로써 단백질을 정제할 수 있다. 더 높은 비율의 Cas9 대 gRNA (예컨대, 3:1)를 사용하여 확실하게는 더 많은 gRNA가 Cas9와 복합체를 형성하도록 할 수 있다. 전형적으로는 37°C에서 15분 동안 미리 인큐베이션시켜 활성 복합체를 형성한다. 반응 완충제는 달라질 수 있다. 예시적인 완충제는 20 mM HEPES, 100 mM NaCl, 5 mM MgCl₂, 0.1 mM EDTA, pH 6.5; 20 mM 트리스(Tris)-HCl, 100 mM KCl, 5 mM MgCl₂, 5% 글리세롤, 1 mM DTT, pH 7.5; 및 1X PBS, 5 mM MgCl₂, 0.5% 트윈(Tween)-20, pH 7.0을 포함한다. 활성 복합체는 즉시 사용될 수 있거나, 또는 수주 동안 4°C에서 보관될 수 있다.

[0176] 계내 검정법에서 Cas9-gRNA 형광:

본 검정법에 따라, 샘플을 현미경 슬라이드 (또는 커버슬립 슬라이드)에 고정시켰다. 슬라이드를 코플린(Coplin) 자에서 인큐베이션시키거나, 유동 챔버 또는 유동 셀에서 어셈블리하여 반응 부피를 한층 더 최소화시키고, 프로세스를 자동화시킬 수 있다. 달리 언급되지 않는 한, 하기 순서대로 반응을 수행하고, 슬라이드를 코플린 자에서 인큐베이션시킨다 (유동 챔버의 경우, 정보는 광호 안에 제공되어 있다). 0.5% 트윈-20을 포함하는 1X PBS와 함께 2분 동안 인큐베이션시킨다 (2 유동 챔버 부피를 이용하여 세척, 각 세척 사이에 30초 동안 인큐베이션). 0.5% 트리톤 X-100을 포함하는 1X PBS와 함께 5분 동안 인큐베이션시킨다 (2 유동 챔버 부피를 이용하여 세척, 각 세척 사이의 2분 동안 인큐베이션). 0.1 N HCl과 함께 5분 동안 인큐베이션시킨다 (2 유동 챔버 부피를 이용하여 세척, 각 세척 사이의 2분 동안 인큐베이션). 0.5% 트윈-20을 포함하는 1X PBS와 함께 2분 동안 인큐베이션시킨다 (2 유동 챔버 부피를 이용하여 세척, 각 세척 사이에 30초 동안 인큐베이션). Cas9 완충제 중에서 5분 동안 인큐베이션시킨다 (2 유동 챔버 부피를 이용하여 세척, 각 세척 사이에 30초 동안 인큐베이션). 이 시간을 사용하여 37°C에서 Cas9 완충제 중에서 Cas9-gRNA 복합체를 미리 형성할 수 있거나, 또는 냉장된 복합체를 37°C로 가온시킬 수 있다. 전형적으로, 샘플당 25 uL 부피 중에서 5 μM의 gRNA 및 5 μM의 Cas9가 함께 복합체를 형성한다. Cas9 완충제 중에 25 uL Cas9-gRNA 복합체를 첨가하고, 37°C에서 4h 동안 습식 챔버에서 인큐베이션시킨다. 대안적으로, 제거가능한 고무 시멘트로 실링하고, 37°C에서 인큐베이션시킨다. Cas9 완충제에서 37°C에서 5분 동안 2회에 걸쳐 인큐베이션시켜 세척한다 (4 유동 챔버 부피를 이용하여 세척, 각 세척 사이에 30초 동안 인큐베이션). 0.5% 트윈-20을 포함하는 1X PBS와 함께 2분 동안 인큐베이션시킨다 (2 유동 챔버 부피를 이용하여 세척, 각 세척 사이에 30초 동안 인큐베이션). 임의적: 필요한 경우, 각 샘플에 0.5% 트윈-20을 포함하는 20 μl 2X SCC에 1 μM 올리고 프로브를 첨가하여 프로빙하고, 15 min 동안 습식 챔버에서 인큐베이션시킨다. 대안적으로, 인큐베이션시키기 전에 제거가능한 고무 시멘트로 실링한다. 0.5% 트윈-20을 포함하는 1X PBS 중에서 2분 동안 2회에 걸쳐 인큐베이션시킴으로써 세척한다 (4 유동 챔버 부피를 이용하여 세척, 각 세척 사이에 30초 동안 인큐베이션). DAPI를 포함하는, 10 uL 안티페이드 현미경 매질을 탑재하고, 네일 폴리시 (예컨대, 안티페이드(AntiFade) 또는 벡타쉴드(VectaShield))로 실링한다. 슬라이드를 영상화를 위해 준비하거나, 또는 암실에서 1주 동안 보관할 수 있다. 일부 경우에, HCL 단계를 생략할 수 있다.

[0178] 유동 셀

관심 샘플을 함유하는 커버글래스 또는 슬라이드와, 제2 커버글래스 또는 슬라이 사이에 끼여 있는 장벽을 제작하기 위해 (어드헤시브 리서치(Adhesive Research) 또는 3M으로부터 입수한) 양면 테이프 또는 시트를 사용하여 유동 셀을 제작할 수 있다. 이비디(Ibidi)로부터 상업적으로 이용가능한 유동 셀 제작용 시스템 (첨착성-슬라이드 VI^{0.4} 또는 첨착성-슬라이드 I 루어(Luer))이 이용될 수 있다. 여기서 세포, 염색체, 또는 DNA가 그 위에 놓이게 되는 슬라이드 또는 커버글래스는 유동 셀의 첨착성 부분에 부착되며, 이에 의해 기판 상에 유동 셀이 생성된다. 유입구 영역으로 수동으로 피펫팅하고, 배출구 영역에서 (예컨대, 블록팅 페이퍼를 사용하여) 위킹(wicking)함으로써 시약을 유동 셀로 유동시킨다. 대안적으로, 유체를 이동시키는 자동 시약 유동 및 교환 시스템 및 다방향 밸브를 통해 시약을 유동시킨다. 이는 시린지 펌프, 압력 구동식 유동 및 흡입에 의해 달성될 수 있다. 자동 시스템은 현미경 또는 영상화 장치와 통합되고, 여기서 유동 셀이 로딩된다.

[0180] DNA의 분자적 코밍(combing)

- [0181] 예를 들어, 수컷 개놈 DNA (프로메가(Promega) 또는 노바젠(Novagen)) 또는 젤 플리그 방법을 사용하여 세포로부터 추출된 DNA를 비닐 실란 (7-옥테닐트리를로도실란)으로 코팅된 커버글래스 상에 코팅한다. 커버글래스 대부분을 덮는 0.5 M MES 완충제 용액 (예컨대, 22x22 커버글래스의 경우, 1-1.5 ml) 중 DNA를 함유하는 트로프내로 커버글래스 (예컨대, 22x22mm)를 침지시키고, DNA 단부가 표면 코팅에 결합될 수 있도록 한 후 (전형적으로, 1 min-10 min), 커버글래스를 트로프로부터 꺼냄으로써 수행된다. 트로프 중의 DNA 농도는 코팅되는 DNA의 원하는 밀도를 제공하도록 조정될 수 있다. 예를 들어, 개별 신장된 DNA 분자가 해동될 수 있는 경우, 최대 0.5 ng/ul의 농도가 적당한 밀도를 제공할 수 있다. 이어서, 커버글래스를 꺼낸 후에, DNA가 후진하는 메니스커스의 힘에 기인하여 신장될 수 있도록 허용하면서, DNA 용액으로부터 커버글래스를 일정한 속도 (예컨대, 300 μm/s)로 꺼낸다. 이어서, 임의적으로, DNA를 자외선 에너지 1 mJ당 대략 10-20 줄을 사용하여 커버글래스 상에 가교시킨다. 임의적으로, 유동 셀을 상기 기술된 바와 같이 커버글래스 상에서 형성시킨다. 코팅된 DNA는 코팅 프로세스 또는 그 이후에 하나 이상의 인터칼레이팅 염료, 예컨대 YOYO-1로 염색하여 시각화될 수 있다. 전형적으로, DNA 염기 쌍 대 YOYO-1 염색의 비는 5:1 내지 10:1인 것이 사용된다.
- [0182] 커버글래스 상에서 미리 신장된 DNA에의 Cas9/gRNA 결합
- [0183] 유동 셀 제작을 위해 그 사이에 끼여 있는, 코팅된 개놈 DNA를 포함하는 커버글래스를 먼저 PBS 트윈 및 PBS로 세척하여 수화시킨다. 임의적으로, 기판을 블록에이드 (라이프 테크놀로지스)로 차단시킨다. 임의적으로, 유동 셀을 Cas9 반응 완충제로 철저히 세척한다. 37°C에서 Cas9-gRNA 복합체를 미리 형성한 후, 유동 셀에 첨가하고, 30분 내지 1시간 동안 방치하여 인큐베이션시킨다. 완충제, 예컨대 반응 완충제 및 PBS 트윈20 및 PBS로 철저히 세척하여 반응을 정지시킨다.
- [0184] 도 6A 및 6B를 위해 사용된 가이드 RNA는 하기와 같다:
- [0185] 테일이 있는 동원체 16 gRNA:
- ```
GACGCCUUCGUUGGAAACGUUUUAGAGCUAGAAAAGCAAGGUAAAAAAGGCUA
GUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUUTCCTCTACCACCT
ACATCACTTATAACATCTA
```
- [0186]
- [0187] 형광 표지된 프로브의 테일에의 혼성화
- [0188] 유동 셀이 수화된 상태 그대로 유지될 수 있도록 하면서, Cas9/gRNA 복합체 형성 후, 하기 혼성화 완충제 중에서 테일에 DNA 프로브를 부가함으로써 복합체를 영상화한다: 200 ul 포름아미드, 20 ul sds, 120 ul 블록에이드, 40 ul의 20xssc, 및 20 ul 물.
- [0189] 테일을 프로빙하는데 사용된 서열은
- [0190] TAGATGTATAA GTGATGTAGGTGGTAGAGGA 이다.
- [0191] 이를 4°C에서 밤새도록 방치할 수 있다. 테일 하이브리드에 대한 프로브의 열 안정성에 따라 다른 혼성화 온도 및 혼성화 완충제 또한 사용될 수 있다.
- [0192] 형광 표지, 예컨대 Atto 647N, Atto 655, Alexa 647을 이용하여 한쪽 단부 또는 양쪽 모두에서 표지할 수 있다. 적절한 형광 표지 (예컨대, Atto 655)를 사용할 때, STORM 방법을 사용하여 초해상도 영상화를 수행할 수 있다.
- [0193] 프로브의 테일에의 혼성화 및 영상화
- [0194] 혼성화 완충제 중 영상화 용액에 DNA PAINT 이미저 프로브를 첨가하여 초해상에 필요한 추계적 온 및 오프 패턴을 수득한다. 테일 상의 서열에 상보적이거나, 또는 테일에 결합하고, 도 6A 및 도 6B, 및 유사 실험에서 사용된 프로브 상의 서열에 상보적인 PAINT 서열은
- ```
/5Alex647N/TAGATGTATAAAAAAATTTAATAAGGT/3AlexF647N  
또는  
/5ATTO647NN/TAGATGTATAAAAAAA/3ATTO647NN/
```
- [0195] 이었다.
- [0196] 내부 전반사 (TIRF) 모드로 니콘(Nikon) Ti-E 도립 현미경 상에서 633, 640 또는 647 nm 레이저 라인, 및 Cy3에 적절한 필터 블록을 사용하여 영상화를 달성하고, 후방 박막형 앤도르 익손(Andor Ixon) X3 EMCCD 카메라 상에서 포착한다.

- [0197] 초해상도 영상화
- [0198] 도 6B에 제시된 초해상도 영상화는 하기 이미지 가닥: P9 이미지: ACCTTATTA를 사용하여 DNA PAINT 방법을 실행함으로써 수행될 수 있다. 상기 이미지 가닥은 P9 헨들 (도킹 서열)을 함유하는 테일에 결합하거나, 또는 P9 헨들을 함유하는 테일에 이미 결합되어 있는 프로브: TAATAAGGT에 결합한다. DNA PAINT를 영상화하는데 적합한 완충제는 5 mM 트리스 pH 8, 10 mM MgCl₂, 1 mM EDTA pH 8 및 0.05% 트윈-20이다. 니콘 NIS 엘리먼츠(Nikon NIS Elements) 소프트웨어를 사용하여 15분 내지 30 min 동안 무비를 촬영할 수 있고, 랩뷰(LabVIEW)에 작성된 DNA PAINT 이미지 애널리시스(Analysis) 코드를 사용하여 초해상도 영상을 구성할 수 있다 (문헌 [Jungmann et al., Nano Lett., 2010, 10 (11), pp 4756-4761] 참조).
- [0199] 용액 중 게놈 DNA에의 Cas9/gRNA 결합
- [0200] (37°C에서 10 min 동안 미리 인큐베이션된) Cas9-gRNA 복합체 어셈블리를 완충제 용액의 존재하에 DNA에 첨가하고, 37°C에서 1시간 동안 인큐베이션시킨다. Cas9/gRNA가 결합된 DNA를 임의적으로 정제한다. 이어서, DNA 신장이 요구되는 경우, 반응물을 0.5 M MES 용액 중에 회석시키고, gRNA/cas9 복합체가 그 위에 데코레이션된 DNA를 비닐 실란 커버글래스 표면 상에 코팅한다. 도 6C에서는 이 접근법을 사용하였다.
- [0201] 복합체를 형성한 DNA의 정제
- [0202] 사용된 가이드 RNA의 크기에 의존하여, 표적 DNA/gRNA/Cas9 복합체를 단리시키는 데, 크기 배제, (예컨대, 스트렙트아비딘 비드에 결합하는데스티오비오틴 또는 절단 가능한 링커를 사용하는) 친화도 정제, 또는 투석 막을 비롯한, 다수의 상이한 정제 방법이 사용될 수 있다.
- [0203] Cas9 낉킹 및 표적화된 서열분석
- [0204] Cas9-gRNA는 표적 DNA 상의 특이적인 위치에 있는 dsDNA 주형에 결합한다. 표적 DNA는 천연 게놈 중에 존재할 수 있고, 반응은 생체내에서 수행된다. 표적 DNA는 세포 중에 존재할 수 있거나, 유기체는 표면에 고정되고, 반응은 계내에서 수행된다. 이러한 경우, 게놈의 공간상 위치는 그대로 보존하는 것이 바람직할 수 있다. 추가로, 표적 DNA를 추출할 수 있고, 반응은 생체외에서 수행된다. 추출된 표적 DNA 또한 표면 상에, 유체 시스템 (예컨대, 나노채널)에 고정화시킬 수 있고, 반응은 생체외에서 수행된다.
- [0205] 돌연변이체 낙카제 (예컨대, D10A 또는 H840A 돌연변이체를 보유하는 Cas9)를 이용하여 gRNA/Cas9 매개 낉킹 반응을 수행한다. 도 9를 참조할 수 있다. D10A 돌연변이에 의해 Cas9의 RuvC 도메인을 불활성화시키고, HNH 도메인은 H840A 돌연변이에 의해 불활성화시킬 수 있다.
- [0206] 두 낉킹 메커니즘 중 하나가 사용될 수 있으며, 상부 가닥에 있는 것은 gRNA와 염기쌍을 형성하고, 나머지 다른 하나는 하부 가닥인, 대체되는 가닥에 있다. D10A 위치의 돌연변이를 통해 하부 가닥이 아닌 상부 가닥의 낉킹이 이루어질 수 있다. H840A 위치의 돌연변이를 통해 상부 가닥이 아닌 하부 가닥의 낉킹이 이루어질 수 있다. 절단은 Cas9(D10A) 낙카제에 의해 프로토스페이서 인접 모티프 (PAM)로부터 상류 쪽으로 3 bp 만큼 떨어져 있는 곳에서 이루어진다. Cas9(H840A) 절단은 하부 가닥 상의 상보적인 위치에서 이루어진다. 게놈에서 다회에 걸쳐 존재하는 특이적인 짧은 서열을 인식하는 낉킹 엔도뉴클레아제와 비교하였을 때, gRNA는, 게놈 중에 단 한번만 존재할 수 있는 비교적 긴 인식 서열에서 나킹한다. 이는 게놈 중 특이적인 독특한 서열을 표적화하는데 사용될 수 있다. 표적화된 서열분석은 D10A 낙카제 또는 H840A 낙카제를 선택하는 단계; 서열분석하고자 하는 영역 부근의 가이드 RNA를 선택하는 단계; 선택된 gRNA 및 Cas9 돌연변이체와의 복합체 형성을 위해 gRNA/Cas9 반응을 수행하는 단계; 임의적으로, 표적 DNA로부터 Cas9/gRNA 복합체를 제거하는 단계; 적절한 완충제 중 폴리미리아제 및 형광 표지적 가역성 종결인자 뉴클레오티드를 첨가하여 도입, 세척, 및 절단으로 이루어진 서열분석 사이클을 수행하는 단계에 의해 수행될 수 있다. 이를 통해 표적화된 위치를 서열분석할 수 있다.
- [0207] 일부 실시양태에서, Cas9-gRNA 복합체를 표적 DNA로부터 및 샘플로부터 전체적으로 제거할 수 있다. 제거제는 계면활성제 (예컨대, 소듐 도데실 술페이트), 유기 화합물 (예컨대, 우레아, 구아니디늄 클로라이드/티오시아네이트), 아미드 (예컨대, 포름아미드), 단백질 분해 효소 (예컨대, 프로테아제), 물리적 특성 (예컨대, 온도) 또는 그의 조합일 수 있다.
- [0208] Cas9(D10A)가 사용될 때, PAM을 함유하는 가닥을 서열분석하면서, 서열분석은 PAM 유전자좌를 거쳐 진행되고, 상기 유전자좌의 하류로 진행된다. Cas9(H840A)가 사용될 때, 서열분석은 PAM 유전자좌의 상류 방향으로 진행되고, NGG PAM 서열에 상보적인 가닥을 서열분석한다. 가이드 RNA 서열을 거쳐 서열분석되는 것을 막기 위해, D10A 돌연변이체가 선택될 수 있다. 관심 영역 상에서 다수의 가이드가 선택될 수 있다.

- [0209] 합성에 의해 단계적 서열분석이 수행되는 경우, 도입된 뉴클레오티드는 각각 가역성 종결인자를 보유하며, 표적화된 각 위치의 단일 뉴클레오티드의 도입 결과는 (바람직하게는 TIRF 조명 및 CCD 또는 CMOS 검출기를 사용함으로써) 종결인자 및 표지 제거 이전에 및 표적화된 각 위치에서 다음 염기에 대한 사이클을 반복하기 이전에 검출된다.
- [0210] 시약 첨가 및 영상화와 함께 서열분석을 위해 다양한 일루미나 SBS 키트 (예컨대, SBS 키트 2)가 하기 순서로 사용될 수 있다: 유니버설 시퀀싱 버퍼(Universal Sequencing Buffer); 인코포레이션 마스터믹스 (Incorporation Mastermix); 유니버설 시퀀싱 버퍼; 이미징 타겟티드 로시(IMAGING TARGETED LOCI); 유니버설 스캔 믹스(Universal Scan Mix); 클리비지 리에이전트 마스터믹스(Cleavage Reagent Mastermix); 클리비지 워시 믹스(Cleavage Wash Mix). 일루미나 키트에 관한 상세한 설명은 월드 와이드 웹사이트 support.illumina.com/downloads/hiseq-rapid-sbs-kit-v2-reagent-prep-guide-15058772.html로부터 다운로드 받을 수 있다. 뉴클레오티드 상에서 4개의 염료 중 2개에 대해서는 532 nm 레이저를 사용하여, 염료 중 나머지 두 염료에 대해서는 660 nm 레이저를 사용하여 영상화를 수행한다. 각 레이저에 의해 여기된 두 염료는 각각 특이적인 방출 필터 및 각 염료의 시그니처를 측정하도록 디자인된 알고리즘을 사용함으로써 구별된다.
- [0211] 계놈 애널라이저 IIx(Genome Analyzer IIx)를 비롯한, 다수의 상이한 일루미나 서열분석 장치 중 하나가 사용될 수 있다. 일루미나 유동 셀 홀더 및 유입구 및 배출구 포트와 호환성인 유동 셀 풋프린트가 사용될 수 있다. 대안적으로, 사내 제작된(home-built) 시스템은 개구수가 높은 대물렌즈, 레이저, CCD 카메라, 형광단 선택적 필터 및 시린지 펌프 기반 또는 압력 구동식 시약 교환 시스템 및 가열식 스테이지와 함께 도립 현미경을 포함한다. 일루미나에 의해 제공되는 것 이외의 다른 다른 뉴클레오티드/염료 조합을 위해 사내 제작된 시스템이 적합화될 수 있다.
- [0212] 서열분석이 실시간 반응 (예컨대, 팩바이오(PacBio) 또는 스타라이트(Starlight) 서열분석)으로서 수행되는 경우, 뉴클레오티드는, 뉴클레오티드가 일단 도입되고 나면, 천연 이탈기인 말단 포스페이트에서 표지된다. 상기 반응은 CCD 또는 CMOS 카메라 상에서 연속적으로 모니터링된다.
- [0213] 신장된 DNA 이중체 상에서의 표적화된 서열분석
- [0214] 계놈 DNA를 (예컨대, 겔 플러그에서 추출을 수행함으로써) 세포로부터 추출할 수 있고, 장쇄로 유지시킬 수 있으며, 이에 의해 서열의 선형 위치는 보존된다. 이는 계놈 중 서열의 조직을 측정하는데 중요한 유용성을 갖는다. 본원에 기술된 ALK, BRCA, FSHM 및 HER2 예에 의해 입증된 바와 같이, 계놈 조직에서의 구조적 변이 (SV)가 질환을 일으킬 수 있다. 그러나, 예컨대 백혈병에서의 Bcr-ABL 전위와 같은 특정 SV는 약물, 예컨대 글리백(Gleevac)에 의해 표적화될 수 있으며, 이에 의해 백혈병 환자는 상기 약물에 의해 표적화될 수 있는 전위를 갖는지 여부를 측정하는데 중요하다.
- [0215] 바람직하게, 서열분석은 문자적 코밍을 사용하여 신장된 DNA에 대해 수행되며, 이에 의해 예를 들어, 표적화된 서열 중 계놈 중 선형 조직이 시각화될 수 있다. 표적화된 영역은 관심 위치에 결합하도록 디자인된 낙킹 돌연변이체 Cas9 및 gRNA를 사용하여, 낙킹 반응을 수행한 후, 낙으로부터 연장시킴으로써 선별될 수 있다. 낙킹 반응은 DNA 코밍 이전에 용액 중에서 수행될 수 있다. 대안적으로, 낙킹 반응은 DNA 코밍 이후에 코밍된 DNA 위에서 유동 셀을 사용하여 수행될 수 있다. 서열분석 반응은 먼저 도입 완충제로 DNA를 수화시키고, 미리 조절한 후, 서열분석 시약을 유동 셀 내로 유동시킴으로써 진행된다. 결과는 수득된 표적 영역의 서열 뿐만 아니라, 계놈 중의 그의 위치이다. 상기가 언제 중요한지에 관한 예는 서열분석의 gRNA 기반 표적화가 전위에 대한 핫스폿인 계놈 중의 서열에 대해 이루어지는 때이지만, 계놈 중 어느 곳에서 전위가 이루어지는지 및 어느 서열과 융합되는지에 대해서는 알려져 있지 않다.
- [0216] 계내 표적화된 서열분석
- [0217] 본원에 기술된 방법은 세포 또는 핵 내의 계놈 부분의 공간적 위치, 및 상기 공간적 위치가 유전자 조절 및 계놈 기능에 어떤 영향을 미치는지를 측정하는 것에 관한 것이다. 세포 중의 특이적인 계놈 서열의 공간적 위치는 상기 서열에 특이적인 gRNA를 사용하여 표적화될 수 있다. 고정된 세포 중 계놈 DNA 상에 gRNA/Cas9 매개된 낙이 생성되고, 형광 기반 서열분석이 계놈 DNA 상에서 계내에서 수행된다. 문헌 [Lee et al., 2015: (Nat Protoc. 2015, 10(3):442-58)]에 기술된 바와 같이, 세포를 유리 바닥 디ッシュ 상에서 성장시키거나, 또는 조직 절편을 관련 기술분야에 공지된 방법에 의해 봉입한다. 세포를 (예컨대, 25°C에서 15 min 동안 PBS 중 10% 포르말린, 또는 -20°C에서 20분 동안 100% 메탄을 사용) 고정시킨다. 트리톤 X-100을 포함하는 PBS 중에서 및 PBS 만을 단독으로 사용하여 세척을 수행한다. 임의적으로, 우레아를 첨가하여 낙king된 DNA로부터 gRNA/Cas9를 제거

한다. 임의적으로, RNase를 사용하여 RNA를 제거한다.

[0218] 서열분석은 세포 내의 계놈 DNA의 단일 문자 상에서 직접적으로 수행될 수 있다. 폴리머라제를 사용하여 뉘 부위에 뉴클레오티드를 도입한다. 바람직하게, 뉴클레오티드는 높은 양자 수율을 제공하는 형광단이 있는 형광 표지, 예컨대 Cy3B 또는 다중 표지된 뉴클레오티드를 보유한다. 관련 기술분야에 공지된 다양한 염격한 정도의 세척을 이용하여 도입되지 않는 뉴클레오티드로부터의 신호를 감소시킨다. 단일 문자 영상화 방법을 통해 영상화를 수행하지만, 세포 중 오직 일부분만이 표면과 접촉하고 있기 때문에, TIRF 영상화는 도입 신호의 소부분 (sub-fraction)에만 접근할 수 있다. 공초점 현미경법 또는 광 시트 현미경법을 사용하여 세포 및 조직 내에 3D 공간 상에 분포되어 있는 신호에 접근할 수 있다. 특히 신호가 심층부에 위치하는 경우, 다광자 또는 2-광자 레이저 현미경법을 사용하여 신호에 접근할 수 있다. 상기 방법은 또한 배경 신호를 감소시키는 데에도 효과적이다.

[0219] PBS 세척 후, 서열분석 반응에서 사용된 완충제를 사용하여 서열분석 반응 믹스를 첨가하기 이전에 샘플을 조절 한다. 사용된 프라이머, 영상화 설정 환경 및 온도 제어 이용가능 여부에 의존하여, 뉴클레오티드 도입, 및 표지 및 종결인자 절단을 위해 SOLID 화학법, CycLIC 화학법, SBL 화학법 또는 합성 화학법에 의한 서열분석, 예컨대 일루미나 서열분석 또는 페이저겐 서열분석이 적용된다. 각 도입 단계 이후에 영상을 촬영한다. 영상을 처리하여 강조하는 신호를 생성한다. 각각의 연속 사이클로부터 얻은 영상을 등록한다. 각 영상에 기본 호출을 하고, 기본 호출을 적재하여 각 포커스에 서열 리드를 제공한다. 리드 구축 프로세스에 다수의 등록된 영상에 걸쳐 일치하는 서열만을 포함시킴으로써 비-특이적인 신호는 필터링할 수 있다. 보우타이(Bowtie) 1.0 또는 다른 서열 정렬 알고리즘을 사용하여 참조에 대해 리드를 정렬하고, 관심 대상이 아닌 오프-표적 서열 또는 잡음을 추가로 필터링한다. 서열분석 리드의 뚜렷한 포커스를 시각화하여 세포 내의 표적화된 서열의 공간적 국재화 정보를 제공한다. 단지 표적화된 위치를 표지하는 것보다는 서열분석하는 것이 관심 영역 내의 서열 변이체 뿐만 아니라, 그의 공간적 위치를 밝혀낼 수 있는 잠재능을 갖는다.

[0220] 핵산 중합체를 따라 진행되는 gRNA/Cas9 복합체의 나노포어 분석

[0221] SiN 막을 제작하고, 투과 전자 현미경 (TEM)을 사용하여 직경이 20 nm인 나노포어를 천공시킨다 (Janssen, X. J. A.; Jonsson, M. P.; Plesa, C.; Soni, G. V.; Dekker, N. H. Nanotechnology 2012, 23 (47), 475302). 이어서, 폴리디메틸실록산 (PDMS) 층으로 막을 페인팅하여 정전 용량을 감소시키고, 신호 대 잡음비를 향상시킨다. 유동 셀에 막을 탑재하고, 이에 의해 유동 셀은 상기 막에 의해 분리된 상부 및 하부 저장소를 포함하게 되고, 이후, 두 저장소를 1 M KC1, 10 mM 트리스, 1 mM EDTA (pH 8)로 충전시킨다. 핵산 중합체를 상부 저장소에 첨가하고, 상부 (-ve) 및 하부 (+ve) 저장소 사이에 전압을 인가하여 실질적으로 선형 방식으로 포어를 통한 핵산 중합체의 전기영동 구동 수송이 이루어지도록 할 수 있다. 액소패치(Axopatch) 200B 증폭기 및 디기데이터(Digidata) 1322A DAQ 디지털화 장치를 포함하는 이온 채널 기록 시스템을 이용하여 전류를 기록 한다. 트랜스알라이저 매트랩(Transalyzer Matlab) 패키지를 사용하여 분석한다 (문헌 [Plesa, C.; Dekker, C. Nanotechnology 2015, 26 (8), 084003] 참조). 핵산 중합체가 나노포어를 진입함에 따른 전류 막힘 증가에 대하여 데이터를 분석한 후, 이어서 gRNA/Cas9 결합 영역이 포어의 좁은 수축 부분을 통과함에 따른 막힘의 추가적인 강조된 상승에 대하여 분석한다.

[0222] 실시예 II

[0223] Her2/ErbB2 진단을 위한 Cas9-gRNA 복합체 생성

[0224] (인간 계놈 기구 유전자 명명 위원회(Human Genome Organization Gene Nomenclature Committee)에 의해 Erbb2 로도 또한 알려져 있는) Her2 유전자 서열은 예컨대, NCBI 월드 와이드 웹사이트 ncbi.nlm.nih.gov/gene/2064 와 같은, 다른 온라인 저장소로부터 이용가능하다.

[0225] dsDNA를 따라 Her2 서열을 분석하여 PAM 모티프를 찾는다. 바람직한 모티프는 5'-GGNNNNNNNNNNNNNNNGG-3'이다. 5' GG는 T7 RNA 폴리머라제를 통한 최적의 RNA 합성을 제공한다. 3' NGG는 Cas9에 의해 인식되는 PAM 서열이다. 그 사이의 17개의 염기 N은 표적화된 서열에 대한 것이고, 높은 특이성을 제공한다. 본 방법은 유전자 변이체 (즉, 이소형)들을 구별할 수 있기 때문에 엑손이 바람직하다.

[0226] 표적 서열을 찾는데 있어서 다수의 프로그램이 인터넷 상에서 자유롭게 이용될 수 있다. 도 10은 CHOPCHOP를 사용하였을 때 제공되는 결과 다이어그램을 보여주는 것이다 (문헌 [Tessa G. Montague; Jose M. Cruz; James A. Gagnon; George M. Church; Eivind Valen. (2014). CHOPCHOP: a CRISPR/Cas9 and TALEN web tool for genome editing. Nucleic Acids Res. 42. W401-W407] 참조).

[0227] Her2에 대한 표적 서열의 일부 목록은 하기 표 1에 제시되어 있다. 이 표는 또한 인간 게놈 상의 표적의 위치, 엑손 번호, DNA 가닥 (즉, 센스 (+) 대 앤티센스 (-)), 및 게놈 중 다른 곳의 2차 표적 (즉, 미스매치, MM)의 잠재적인 개수에 관한 정보도 포함한다.

[0228] <표 1>

	표적 서열	게놈 위치	엑손	가닥	MM
1	GGGCGAGGAGGGAGCCCCCAGCGG	chr17:39700259	1	-	0
2	GGTGGCGGAGCATGTCCAGGTGG	chr17:39707039	2	-	0
3	GGTGGGTCTCGGGACTGGCAGGG	chr17:39707021	2	-	0
4	GGCAGCCCTGGTAGAGGTGGCGG	chr17:39707054	2	-	0
5	GGAGGCCCTGTGACAGGGTGG	chr17:39708472	3	-	0
6	GGGCCTCCCCAGGAGGCCTGCGG	chr17:39708487	3	+	0
7	GGCCTCCCCAGGAGGCCTGCGG	chr17:39708488	3	+	0
8	GGTGGCTGTGCCCGCTGCAAGGG	chr17:39710106	6	+	0

[0229]

9	GGGCAGTGGCCCCTTGCAGCGGG	chr17:39710116	6	-	0
10	GGGACAGGCAGTCACACAGCTGG	chr17:39710458	7	-	0
11	GGTTGTGCAGGGGGCAGACGAGG	chr17:39711962	8	-	0
12	GGGCATGGAGCACTTGCAGAGAGG	chr17:39712335	9	+	0
13	GGAGCACTTGCAGAGAGGTGAGGG	chr17:39712341	9	+	0
14	GGAGCTGCTCTGGCTGGAGCGGG	chr17:39715307	10	-	0
15	GGAAGACGCTGAGGTCAAGCAGG	chr17:39715477	11	-	0
16	GGTGGGTGTTATGGTGGATGAGG	chr17:39715824	12	-	0
17	GGCTGGGCTGCGCTCACTGAGG	chr17:39715781	12	+	0
18	GGTGC GGTTCCGAAAGAGCTGG	chr17:39715875	12	-	0
19	GGCTGGGCATCAGCTGGCTGGGG	chr17:39715766	12	+	0
20	GGGCTGGGCATCAGCTGGCTGGG	chr17:39715765	12	+	0
21	GGAGGAATGCCGAGTACTGCAGG	chr17:39716410	13	+	0
22	GGCATT CCTCCACGC ACTCCTGG	chr17:39716398	13	-	0
23	GGCTGACACTCAGGGTGGCACGG	chr17:39716552	14	-	0
24	GGTCAGGTTTACACCGCTGGGG	chr17:39717382	15	-	0
25	GGGCAGCGGGCACGCAGAAGGG	chr17:39717362	15	-	0
26	GGGGCAGCGGGCACGCAGAAGG	chr17:39717363	15	-	0
27	GGTTGGCATTCTGCTGGTCGTGG	chr17:39723346	17	+	0
28	GGAGAATGTGAAAATTCCAGTGG	chr17:39723932	19	+	0
29	GGGCATCTGCCTGACATCCACGG	chr17:39724776	20	+	0
30	GGATGTGCGGCTCGTACACAGGG	chr17:39725066	21	+	0
31	GGTGAACCGCCGGCGGAGAATGG	chr17:39725355	22	-	0
32	GGTCAGGGATCTCCCGGGCTGGG	chr17:39725759	23	-	0
33	GGATGACCACAAAGCGCTGGGG	chr17:39726635	24	-	0
34	GGATGATTGACTCTGAATGTCGG	chr17:39726565	24	+	0
35	GGTGTCTGAATTCTCCGCATGG	chr17:39726605	24	+	0
36	GGACAGAAGAAGCCCTGCTGGGG	chr17:39726920	25	-	0
37	GGGGGACCTGGTGGATGCTGAGG	chr17:39726886	25	+	0
38	GGACGATGACATGGGGACCTGG	chr17:39726874	25	+	0
39	GGTGGATGCTGAGGAGTATCTGG	chr17:39726895	25	+	0
40	GGCACCGCAGCTCATCTACCAGG	chr17:39726981	25	+	0
41	GGAGTATCTGGTACCCCAGCAGG	chr17:39726907	25	+	0
42	GGACCATGCCCCAGCGCCCGGG	chr17:39726952	25	-	0
43	GGGTGCCAGTGGAGACCTGGGGG	chr17:39727344	26	-	0
44	GGCGGTGGGGACCTGACACTAGG	chr17:39727298	26	+	0

[0230]

45	GGGGAGGGCTTGCAGCCCCTTGG	chr17:39727419	26	-	0
46	GGTCCTGGTCCCAGTAATAGAGG	chr17:39727933	27	-	0
47	GGGACCAAGGACCCACCAGAGCGG	chr17:39727944	27	+	0
48	GGTGTCCTTGAAGGTGCTGGG	chr17:39727976	27	-	0
49	GGGGGCTGGGCCAACATCTGG	chr17:39727703	27	-	0
50	GGCCCAAGACTCTCCCCAGGG	chr17:39727782	27	+	0

[0231]

[0232]

유전자 밖의 영역도 쉽게 표적화될 수 있다는 것에 주의하여야 한다. 특정 영역, 예컨대 엑손 및 인트론이 표적화되는데, 이는 1 kb 미만 내지 30 kb 전체 유전자, 수십 내지 수백 킬로베이스, 또는 수 메가베이스일 수 있다.

[0233]

서열 표적은 gRNA 제조를 위한 주형으로서 사용된다. 이는 각 표적 서열을 gRNA 스캐폴드에 부가하는 것을 포함한다. 어셈블리 모델의 다이어그램은 하기에 제시되어 있고, 어셈블리는 고충실 DNA 폴리머라제를 이용하여 PCR을 사용함으로써 수행된다(융점 (T_m) 제공). 요약하면, 올리고뉴클레오티드를 상업적 업체로부터 입수하거나, 또는 사내에서 합성한다. 2개의 범용 올리고뉴클레오티드: Fwd-T7-gRNA, T7 RNA 폴리머라제 인식 모티프의 일부분도 또한 포함하는 정방향 PCR 프라이머; 및 gRNA.split60인 범용 gRNA 스캐폴드가 존재한다. 추가로, 2개의 가변 올리고뉴클레오티드: Sp.gRNA.split60인 관심 표적에 특이적인 서열; 및 Rev-B1-gRNA.18인 다중 가닥 검출을 위한 바코드가 있는 핸들도 또한 포함하는 역방향 PCR 프라이머가 존재한다.

[0234]

본 디자인을 통해 포스포르아미다이트 화학법에 의해 합성되는 올리고뉴클레오티드를 최적의 서열 정확도로 제공하고, 표지된 엔티티가 도킹될 수 있는 코드인, 다양한 유형의 바코드 핸들을 부가하면서, 비용면에서 효율적으로 gRNA를 합성 및 어셈블리할 수 있다. DNA 주형은 올리고뉴클레오티드 어레이 합성기 상에서 이루어지는 것을 비롯하여, 상이한 규모로 합성될 수 있거나, 또는 상업적으로 입수할 수 있다. 주형을 PCR에 의해 증폭 및 재증폭시켜 주형을 생성하고, 영구화시킬 수 있는데, 이는 신생 합성보다 비용면에서 훨씬 더 효율적이다.

[0235]

도 11에 도시되어 있는 예시적인 PCR 어셈블리는 1시간 미만으로 소요될 수 있다. PCR 후, T7 RNA 폴리머라제 믹스를 주형 DNA에 첨가함으로써 시험판내 전사 (IVT)에 의해 gRNA를 합성한다. 이 경우, 더 긴 반응이 더 많은 gRNA를 제공할 것이다. 사용 횟수가 제한된 빠르게 제조되는 Cas9-gRNA 키트는 30분 이내에 제조될 수 있다. 다양은 16시간 동안 진행되는 반응으로 제조될 수 있다. IVT 이후, DNaseI 처리에 의해 DNA 주형을 분해하고, 시판용 칼럼 정제 키트를 사용하여, 또는 아세트산암모늄의 존재하에서 에탄올 침전에 의해 RNA를 정제 한다.

[0236]

이어서, 마그네슘 이온 존재하에서 바코드가 있는 gRNA를 Cas9 단백질과 1:1 비율로 복합체를 형성하여 Cas9-gRNA 복합체를 형성한다. 이어서, 시약을 사용할 수 있다. 바코드 디코딩을 촉진시키기 위해 바코드에 결합하는 표지된 올리고뉴클레오티드인 검출 프로브 또한 본 단계에서 첨가할 수 있으며, 이에 의해 하류에 있는 혼성화 검출 방법은 가속화될 수 있다. 이러한 결정은 예상되는 신호 대 잡음비에 기초하여 이루어진다. 예로서, 잡음 (예컨대, 현미경 슬라이드의 표면 상에서의 염색체 스프레드 상의 Her2의 검출)이 낮은 것으로 예상될 경우, 이때는 간단하게 Cas9-gRNA-표지 키트를 샘플에 첨가하고, 37°C에서 반응이 이루어질 수 있도록 할 수 있다. 비반응 복합체를 세척하고, 샘플을 영상화할 수 있다. 키트는 유전자좌를 프로빙하는데 필요한 모든 Cas9-gRNA 복합체의 양상을 봄이다. Cas9-gRNA-표지 키트는 미리 표지된 Cas9-gRNA 모두의 양상을 봄이다.

[0237]

다중 검출을 위해 Cas9-gRNA를 코딩하는 것에 관한 예

[0238]

하나 또는 다중의 코드 또는 바코드 핸들 (프로브에 대한 도킹 부위)을 gRNA에 첨가할 수 있다. 원리는 DNA 샘플에 연관된 Cas9-gRNA 복합체의 정체성을 측정하는 것이다. 이는 gRNA 핸들에 특이적인 검출가능한 "디코더" 프로브에 혼성화시킴으로써 이루어진다. gRNA 핸들은 하나 또는 다중의 바코드를 함유할 수 있고, 이는 다양한 방식으로 배열될 수 있으며, 예컨대 연속하여 또는 부분적으로 서로 중복되어 적재될 수 있다. 후자의 경우, 이를 통해 다중 바코드의 전체 길이는 단축될 수 있고, 이는 주어진 gRNA 핸들에 대한 다중 능력을 증가시킨다.

[0239]

다중 능력은 바코드 개수에 비례하고, 이는 바코드 조합 수, 예컨대 $nCr = r \cdot (n + r - 1)! / r!(n-1)!$ (여기서 n은 바코드 개수이고, r은 이벤트 수이고, 여기서 바코드는 리드이다)와 같다. 예로서, 바코드가 2개 B1 및 B2인 경우, 총 3개의 코드에 대하여 C = 1 및 P = 2를 얻게 된다 (B1B1, B1B2, B2B2). 바코드가 3개인 경우, 10개의 코드를 얻게 된다. 126개의 코드는 5개 만큼 적은 바코드로도 생성될 수 있는 반면, 10개의 바코드는 액

숨 검색에 충분한, 총 92378개의 코드를 제공한다.

[0240]

본원에서 제공하는 코딩 방법은 하기 내용을 따라 진행된다. 유전자좌 1개당 1개의 코드를 배정하여 단일 검정법에서 다중의 표적을 확인할 수 있도록 할 수 있다. 주어진 유전자좌에 대해 1개의 이소형당 1개의 코드를 배정하여 변이가 확인될 수 있도록 할 수 있다. 주어진 유전자좌에 대해 2개의 코드를 배정할 수 있으며, 이는 주어진 표적에 대한 신호 정체성을 공동-국제화하는데, 이를 통해 표적 정체성의 신뢰도는 증가할 수 있다. 이는 배경이 높은 샘플인 경우에 이로우며, 이때, 검정법은 두 신호 모두가 같은 유전자좌 상에서 검출되는 경우에만 오직 양성임을 나타낼 것이다. 주어진 유전자좌에 대해 2개 초과의 코드를 배정할 수 있으며, 이를 통해 서도 역시 공동-국제화할 수 있지만, 이는 확인을 향상시키고, 해상도를 향상시키는 시그니처를 제공할 것이다. 예로서, 첫번째 것은 바코드 B1B2 및 B1B3으로 코딩되고, 두번째 것은 바코드 B1B3 및 B1B4로 코딩된 것인, 2 세트의 Cas9-gRNA에 의해 2개의 유전자좌가 인식될 수 있다. 공통된 B1은 유전자좌 간에 쉽게 검출가능한 신호를 제공할 것이다. 신호 위치는 더욱 특이적으로 초점화될 수 있고, 유전자좌 1에 대한 나머지 바코드 B2B3, 및 유전자좌 2에 대한 나머지 바코드 B3B4가 검출될 수 있다.

[0241]

하기 표 2는 인간 게놈에는 1회 초과로 존재하지 않는, 길이가 12개의 뉴클레오티드 길이인 서열의 목록으로부터 생성된 것인, 5개의 바코드 목록을 제공한다:

[0242]

<표 2>

바코드	서열
B1	CGTCGATTACCA
B2	CCATACTCGTCG
B3	TCGATAGTACGT
B4	CGTACTGAACGA
B5	CGAAGTACTACG
B6	TCGTTACGACCA
B7	CCAGACGAATCG
B8	TCGAATAGTCGT
B9	CGTAACTAACGA
B10	CGAACAAATCGTA

[0243]

바코드 크기가 12개의 염기인 것이 다양한 길이 및 용점에 걸친 다양한 프로브를 디자인할 수 있도록 하는데 충분하다. 예로서, 10 내지 12 mer 프로브는, 여전히 변성체 (예컨대, 50% 포름아미드)를 사용하여 제거될 수 있으면서, 특정 이온 농도 (예컨대, 5 mM MgCl₂)하에서는 더욱 안정적인 상호작용이 이루어질 수 있도록 한다. 8 내지 9 mer 프로브는 문헌 [Jungmann R, Avendano MS, Woehrstein JB, Dai M, Shih WM, Yin P. Multiplexed 3D cellular super-resolution imaging with DNA-PAINT and Exchange-PAINT. Nat Methods. 2014 Mar;11(3):313-8]에 기술된 DNA-PAINT 기술을 사용하여 이루어지는 초해상도 영상화가 수행될 수 있도록 한다.

[0245]

바코드 특이적 프로브는 많은 유형이 있을 수 있고, 본 검정법에 맞게 적합화될 수 있다. 예로서, 프로브는 서열에 상보적인 올리고뉴클레오티드일 수 있다. 또 다른 예에서, 프로브는 예컨대, 패드락 프로브와 같이, 고리형 ssDNA 주형으로 이루어질 수 있으며, 여기서 프로브의 한 영역은 바코드와 혼성화하고, 하나 이상의 영역은 프로브의 2차 세트를 위한 보조 혼성화로서 작용한다. 이어서, 고리형 프로브는 롤링 서클 증폭을 사용하여 증폭될 수 있으며, 이에 의해 신호 강도는 증가될 것이다. 또 다른 예에서, 혼성화 프로브는 선형이고, 2차 또는 3 차 프로브 등의 것을 위한 부위도 또한 포함하는 2차 프로브에 대한 부위를 포함하고, 이에 의해 초분지형 프로브의 자기 어셈블리가 이루어질 수 있으며, 이에 의해 신호 강도는 증가될 것이다. 본원에서 논의되는 프로브는 대부분 형광 표지된 것이지만, 또 다른 예에서, 상기 프로브는 발색성 검출이 이루어질 수 있도록 허용하는 분자에 부착되며, 예컨대 고밀도의 금 나노입자라고 가정하면, 검출은 비색성일 수 있고, 이는 여전히 핵 내의 특정 영역으로 시각적으로 국제화될 수 있다.

[0246]

위치(Locational) RGB 코드의 예가 도 7에 제시되어 있고, 가이드 RNA 테일은 BRCA1 재배열을 형광으로 바코딩하는 gRNA를 열거하는 하기 표 3에 제시된 바와 같이 열거된 컬러 스킴에 따라 바코딩된다. R은 레드이고, G는 그린이며, B는 블루이다. 예상되는 순차적 컬러 패턴을 형성하는 방식으로 바코드를 분리시킨다. 상기 패턴에 따라, 각 바코드는 컬러와 연관성을 띠게 되며, 이는 3개의 바코드 서열이 생성되어야 하고 이어서, 컬러 코드

에 따라 조합된다는 것을 의미한다. 예로서, gRNA 코드가 R인 경우, 이때는 오직 바코드 R만이 gRNA 테일에 부가되어야 하지만, gRNA 코드가 RGB일 경우에는 3개의 바코드 모두가 gRNA 테일에서 조합되어야 한다. 바코드는 그의 각 형광 프로브 첨가를 통해 검출되며, 이에 의해 실제 순차적 컬러 패턴이 밝혀지게 된다. 예상되지 않는 임의의 패턴은 게놈 재배열로서 확인될 것이다.

[0247]

<표 3>

명칭	위치	코드	오리가미 코드		
			스폿1	스폿2	스폿3
gRNA.BRCA1.1	5	R	RG	RGB	RG
gRNA.BRCA1.2	177	R	RG	RGB	RB
gRNA.BRCA1.3	317	R	RG	RGB	BG
gRNA.BRCA1.4	460	R	RG	RGB	R
gRNA.BRCA1.5	585	R	RG	RGB	G
gRNA.BRCA1.6	610	R	RG	RGB	B
gRNA.BRCA1.7	730	R	RG	RB	RGB
gRNA.BRCA1.8	738	R	RG	RB	RG
gRNA.BRCA1.9	792	R	RG	RB	BG
gRNA.BRCA1.10	827	R	RG	RB	R
gRNA.BRCA1.11	881	R	RG	RB	G
gRNA.BRCA1.12	906	R	RG	RB	B
gRNA.BRCA1.13	960	R	RG	BG	RGB
gRNA.BRCA1.14	1,048	R	RG	BG	RG
gRNA.BRCA1.15	1,220	R	RG	BG	RB
gRNA.BRCA1.16	1,310	R	RG	BG	R
gRNA.BRCA1.17	1,337	R	RG	BG	G
gRNA.BRCA1.18	1,403	R	RG	BG	B
gRNA.BRCA1.19	1,410	R	RG	R	RGB
gRNA.BRCA1.20	1,446	R	RG	R	RG
gRNA.BRCA1.21	1,459	R	RG	R	RB
gRNA.BRCA1.22	1,509	R	RG	R	BG
gRNA.BRCA1.23	1,561	R	RG	R	G
gRNA.BRCA1.24	1,647	R	RG	R	B
gRNA.BRCA1.25	1,667	R	RG	G	RGB
gRNA.BRCA1.26	1,675	R	RG	G	RG
gRNA.BRCA1.27	1,813	R	RG	G	RB
gRNA.BRCA1.28	2,118	R	RG	G	BG
gRNA.BRCA1.29	2,119	R	RG	G	R

[0248]

gRNA.BRCA1.30	2,120	R	RG	B	RGB
gRNA.BRCA1.31	2,212	R	RG	B	RG
gRNA.BRCA1.32	2,304	R	RG	B	RB
gRNA.BRCA1.33	2,372	R	RG	B	BG
gRNA.BRCA1.34	2,533	R	RB	RGB	RG
gRNA.BRCA1.35	2,589	R	RB	RGB	RB
gRNA.BRCA1.36	2,609	R	RB	RGB	BG
gRNA.BRCA1.37	2,662	R	RB	RGB	R
gRNA.BRCA1.38	2,800	R	RB	RGB	G
gRNA.BRCA1.39	2,844	R	RB	RGB	B
gRNA.BRCA1.40	2,917	R	RB	RG	RGB
gRNA.BRCA1.41	2,951	R	RB	RG	RB
gRNA.BRCA1.42	3,012	R	RB	RG	BG
gRNA.BRCA1.43	3,110	R	RB	RG	R
gRNA.BRCA1.44	3,111	R	RB	RG	G
gRNA.BRCA1.45	3,112	R	RB	RG	B
gRNA.BRCA1.46	3,132	R	RB	BG	RGB
gRNA.BRCA1.47	3,333	R	RB	BG	RG
gRNA.BRCA1.48	3,362	R	RB	BG	RB
gRNA.BRCA1.49	3,363	R	RB	BG	R
gRNA.BRCA1.50	3,399	R	RB	BG	G
gRNA.BRCA1.51	3,877	R	RB	BG	B
gRNA.BRCA1.52	3,899	R	RB	R	RGB
gRNA.BRCA1.53	4,168	R	RB	R	RG
gRNA.BRCA1.54	4,174	R	RB	R	RB
gRNA.BRCA1.55	4,200	R	RB	R	BG
gRNA.BRCA1.56	4,401	R	RB	R	G
gRNA.BRCA1.57	4,408	R	RB	R	B
gRNA.BRCA1.58	4,422	R	RB	G	RGB
gRNA.BRCA1.59	4,453	R	RB	G	RG
gRNA.BRCA1.60	4,488	R	RB	G	RB
gRNA.BRCA1.61	4,489	R	RB	G	BG
gRNA.BRCA1.62	4,528	R	RB	G	R
gRNA.BRCA1.63	4,691	R	RB	G	B
gRNA.BRCA1.64	4,752	R	RB	B	RGB
gRNA.BRCA1.65	4,777	R	RB	B	RG

[0249]

gRNA.BRCA1.66	4,831	R	RB	B	RB
gRNA.BRCA1.67	4,919	R	RB	B	BG
gRNA.BRCA1.68	4,955	R	RB	B	R
gRNA.BRCA1.69	5,016	R	RB	B	G
gRNA.BRCA1.70	5,123	R	BG	RGB	RG
gRNA.BRCA1.71	5,146	R	BG	RGB	RB
gRNA.BRCA1.72	5,191	R	BG	RGB	BG
gRNA.BRCA1.73	5,282	R	BG	RGB	R
gRNA.BRCA1.74	5,448	R	BG	RGB	G
gRNA.BRCA1.75	5,473	R	BG	RGB	B
gRNA.BRCA1.76	5,675	R	BG	RG	RGB
gRNA.BRCA1.77	5,774	R	BG	RG	RB
gRNA.BRCA1.78	5,808	R	BG	RG	BG
gRNA.BRCA1.79	5,843	R	BG	RG	R
gRNA.BRCA1.80	5,958	R	BG	RG	G
gRNA.BRCA1.81	6,492	R	BG	RG	B
gRNA.BRCA1.82	6,493	R	BG	RB	RGB
gRNA.BRCA1.83	6,642	R	BG	RB	RG
gRNA.BRCA1.84	6,903	R	BG	RB	BG
gRNA.BRCA1.85	6,996	R	BG	RB	R
gRNA.BRCA1.86	7,027	R	BG	RB	G
gRNA.BRCA1.87	7,347	R	BG	RB	B
gRNA.BRCA1.88	7,431	R	BG	R	RGB
gRNA.BRCA1.89	7,444	R	BG	R	RG
gRNA.BRCA1.90	7,453	R	BG	R	RB
gRNA.BRCA1.91	7,474	R	BG	R	BG
gRNA.BRCA1.92	7,479	R	BG	R	G
gRNA.BRCA1.93	7,489	R	BG	R	B
gRNA.BRCA1.94	7,490	R	BG	G	RGB
gRNA.BRCA1.95	7,510	R	BG	G	RG
gRNA.BRCA1.96	7,511	R	BG	G	RB
gRNA.BRCA1.97	7,720	R	BG	G	BG
gRNA.BRCA1.98	7,721	R	BG	G	R
gRNA.BRCA1.99	7,997	R	BG	B	RGB
gRNA.BRCA1.100	8,022	R	BG	B	RG
gRNA.BRCA1.101	8,200	R	BG	B	RB

[0250]

gRNA.BRCA1.102	8,258	R	BG	B	BG
gRNA.BRCA1.103	8,279	R	BG	B	R
gRNA.BRCA1.104	8,380	R	R	RGB	RG
gRNA.BRCA1.105	8,544	R	R	RGB	RB
gRNA.BRCA1.106	8,556	R	R	RGB	BG
gRNA.BRCA1.107	8,702	G	R	RGB	R
gRNA.BRCA1.108	8,859	G	R	RGB	G
gRNA.BRCA1.109	8,990	G	R	RG	RGB
gRNA.BRCA1.110	9,079	G	R	RG	RB
gRNA.BRCA1.111	9,080	G	R	RG	BG
gRNA.BRCA1.112	9,315	G	R	RG	R
gRNA.BRCA1.113	9,339	G	R	RG	G
gRNA.BRCA1.114	9,407	G	R	RG	B
gRNA.BRCA1.115	9,539	G	R	RB	RGB
gRNA.BRCA1.116	9,814	G	R	RB	RG
gRNA.BRCA1.117	9,835	G	R	RB	BG
gRNA.BRCA1.118	10,035	G	R	RB	R
gRNA.BRCA1.119	10,047	G	R	RB	G
gRNA.BRCA1.120	10,107	G	R	RB	B
gRNA.BRCA1.121	10,152	G	R	BG	RGB
gRNA.BRCA1.122	10,201	G	R	BG	RG
gRNA.BRCA1.123	10,303	G	R	BG	RB
gRNA.BRCA1.124	10,478	G	R	BG	R
gRNA.BRCA1.125	10,545	G	R	BG	G
gRNA.BRCA1.126	10,548	G	R	BG	B
gRNA.BRCA1.127	10,549	G	R	G	RGB
gRNA.BRCA1.128	10,551	G	R	G	RG
gRNA.BRCA1.129	10,760	G	R	G	RB
gRNA.BRCA1.130	10,866	G	R	G	BG
gRNA.BRCA1.131	10,874	G	R	B	RG
gRNA.BRCA1.132	10,894	G	R	B	RB
gRNA.BRCA1.133	10,990	G	R	B	BG
gRNA.BRCA1.134	10,997	G	G	RGB	RG
gRNA.BRCA1.135	11,082	G	G	RGB	RB
gRNA.BRCA1.136	11,117	G	G	RGB	BG
gRNA.BRCA1.137	11,137	G	G	RGB	R

[0251]

gRNA.BRCA1.138	11,151	G	G	RGB	G
gRNA.BRCA1.139	11,172	G	G	RGB	B
gRNA.BRCA1.140	11,602	G	G	RG	RGB
gRNA.BRCA1.141	11,796	G	G	RG	RB
gRNA.BRCA1.142	11,804	G	G	RG	BG
gRNA.BRCA1.143	12,424	G	G	RG	R
gRNA.BRCA1.144	12,737	G	G	RG	G
gRNA.BRCA1.145	12,878	G	G	RB	RGB
gRNA.BRCA1.146	12,922	G	G	RB	RG
gRNA.BRCA1.147	12,977	G	G	RB	BG
gRNA.BRCA1.148	12,998	G	G	RB	R
gRNA.BRCA1.149	13,156	G	G	RB	G
gRNA.BRCA1.150	13,178	G	G	RB	B
gRNA.BRCA1.151	13,219	G	G	BG	RGB
gRNA.BRCA1.152	13,483	G	G	BG	RG
gRNA.BRCA1.153	13,578	G	G	BG	RB
gRNA.BRCA1.154	13,637	G	G	BG	R
gRNA.BRCA1.155	13,741	G	G	BG	G
gRNA.BRCA1.156	13,831	G	G	R	RGB
gRNA.BRCA1.157	13,997	G	G	R	RG
gRNA.BRCA1.158	14,306	G	G	R	RB
gRNA.BRCA1.159	14,476	G	G	R	BG
gRNA.BRCA1.160	14,501	G	G	B	RB
gRNA.BRCA1.161	14,522	G	G	B	BG
gRNA.BRCA1.162	14,575	G	B	RGB	RG
gRNA.BRCA1.163	14,587	G	B	RGB	RB
gRNA.BRCA1.164	15,033	G	B	RGB	BG
gRNA.BRCA1.165	15,219	G	B	RG	RGB
gRNA.BRCA1.166	15,222	G	B	RG	RB
gRNA.BRCA1.167	15,475	G	B	RG	BG
gRNA.BRCA1.168	15,522	G	B	RG	R
gRNA.BRCA1.169	15,579	G	B	RB	RGB
gRNA.BRCA1.170	15,818	G	B	RB	RG
gRNA.BRCA1.171	15,901	G	B	RB	BG
gRNA.BRCA1.172	16,111	G	B	RB	R
gRNA.BRCA1.173	16,136	G	B	BG	RGB

[0252]

gRNA.BRCA1.174	16,448	G	B	BG	RG
gRNA.BRCA1.175	16,595	G	B	BG	RB
gRNA.BRCA1.176	16,688	G	B	BG	R
gRNA.BRCA1.177	16,747	G	B	R	RGB
gRNA.BRCA1.178	16,748	G	B	R	BG
gRNA.BRCA1.179	16,918	G	B	G	RG
gRNA.BRCA1.180	16,974	G	B	G	RB
gRNA.BRCA1.181	17,418	G	B	G	BG
gRNA.BRCA1.182	17,422	G	RGB	RG	RGB
gRNA.BRCA1.183	17,589	G	RGB	RG	RB
gRNA.BRCA1.184	17,669	G	RGB	RG	BG
gRNA.BRCA1.185	18,262	G	RGB	RG	R
gRNA.BRCA1.186	18,269	G	RGB	RG	G
gRNA.BRCA1.187	18,970	G	RGB	RG	B
gRNA.BRCA1.188	18,971	G	RGB	RB	RGB
gRNA.BRCA1.189	19,005	G	RGB	RB	RG
gRNA.BRCA1.190	19,026	G	RGB	RB	BG
gRNA.BRCA1.191	19,289	G	RGB	RB	R
gRNA.BRCA1.192	19,392	G	RGB	RB	G
gRNA.BRCA1.193	19,445	G	RGB	RB	B
gRNA.BRCA1.194	19,545	G	RGB	BG	RGB
gRNA.BRCA1.195	19,570	G	RGB	BG	RG
gRNA.BRCA1.196	19,818	G	RGB	BG	RB
gRNA.BRCA1.197	19,819	G	RGB	BG	R
gRNA.BRCA1.198	19,839	G	RGB	BG	G
gRNA.BRCA1.199	20,595	G	RGB	BG	B
gRNA.BRCA1.200	20,810	G	RGB	R	RGB
gRNA.BRCA1.201	20,952	G	RGB	R	RG
gRNA.BRCA1.202	20,988	G	RGB	R	RB
gRNA.BRCA1.203	21,054	G	RGB	R	BG
gRNA.BRCA1.204	21,075	G	RGB	R	G
gRNA.BRCA1.205	21,352	G	RGB	G	RGB
gRNA.BRCA1.206	21,440	G	RGB	G	RG
gRNA.BRCA1.207	21,499	G	RGB	G	RB
gRNA.BRCA1.208	21,506	G	RGB	G	BG
gRNA.BRCA1.209	21,581	G	RGB	G	R

[0253]

gRNA.BRCA1.210	21,592	G	RGB	B	RGB
gRNA.BRCA1.211	22,366	G	RGB	B	RG
gRNA.BRCA1.212	22,373	G	RGB	B	RB
gRNA.BRCA1.213	22,729	B	RGB	B	BG
gRNA.BRCA1.214	22,760	B	RGB	B	G
gRNA.BRCA1.215	22,824	B	RB	RGB	RG
gRNA.BRCA1.216	22,849	B	RB	RGB	RB
gRNA.BRCA1.217	23,339	B	RB	RGB	BG
gRNA.BRCA1.218	23,739	B	RB	RGB	R
gRNA.BRCA1.219	24,253	B	RB	RGB	G
gRNA.BRCA1.220	24,443	B	RB	RGB	B
gRNA.BRCA1.221	24,530	B	RB	RG	RGB
gRNA.BRCA1.222	24,877	B	RB	RG	RB
gRNA.BRCA1.223	24,878	B	RB	RG	BG
gRNA.BRCA1.224	24,927	B	RB	RG	R
gRNA.BRCA1.225	25,105	B	RB	RG	G
gRNA.BRCA1.226	25,350	B	RB	RG	B
gRNA.BRCA1.227	25,368	B	RB	BG	RGB
gRNA.BRCA1.228	25,472	B	RB	BG	RG
gRNA.BRCA1.229	25,525	B	RB	BG	RB
gRNA.BRCA1.230	25,614	B	RB	BG	R
gRNA.BRCA1.231	25,693	B	RB	BG	G
gRNA.BRCA1.232	25,852	B	RB	BG	B
gRNA.BRCA1.233	25,936	B	RB	R	RGB
gRNA.BRCA1.234	25,972	B	RB	R	RG
gRNA.BRCA1.235	26,984	B	RB	R	RB
gRNA.BRCA1.236	27,130	B	RB	R	BG
gRNA.BRCA1.237	27,155	B	RB	R	G
gRNA.BRCA1.238	27,252	B	RB	G	RGB
gRNA.BRCA1.239	27,256	B	RB	G	RG
gRNA.BRCA1.240	27,431	B	RB	G	RB
gRNA.BRCA1.241	27,447	B	RB	G	BG
gRNA.BRCA1.242	27,452	B	RB	G	R
gRNA.BRCA1.243	27,610	B	RB	B	RGB
gRNA.BRCA1.244	27,631	B	RB	B	RG
gRNA.BRCA1.245	27,682	B	RB	B	RB

[0254]

gRNA.BRCA1.246	27,722	B	RB	B	BG
gRNA.BRCA1.247	28,262	B	RB	B	R
gRNA.BRCA1.248	28,287	B	RB	B	G
gRNA.BRCA1.249	28,400	B	BG	RGB	RG
gRNA.BRCA1.250	28,425	B	BG	RGB	RB
gRNA.BRCA1.251	28,464	B	BG	RGB	BG
gRNA.BRCA1.252	28,581	B	BG	RGB	R
gRNA.BRCA1.253	28,817	B	BG	RGB	G
gRNA.BRCA1.254	28,958	B	BG	RGB	B
gRNA.BRCA1.255	28,983	B	BG	RG	RGB
gRNA.BRCA1.256	29,080	B	BG	RG	RB
gRNA.BRCA1.257	29,202	B	BG	RG	BG
gRNA.BRCA1.258	29,516	B	BG	RG	R
gRNA.BRCA1.259	30,531	B	BG	RG	G
gRNA.BRCA1.260	30,964	B	BG	RG	B
gRNA.BRCA1.261	30,989	B	BG	RB	RGB
gRNA.BRCA1.262	31,104	B	BG	RB	RG
gRNA.BRCA1.263	31,374	B	BG	RB	BG
gRNA.BRCA1.264	31,469	B	BG	RB	R
gRNA.BRCA1.265	31,486	B	BG	RB	G
gRNA.BRCA1.266	31,547	B	BG	RB	B
gRNA.BRCA1.267	31,698	B	BG	R	RGB
gRNA.BRCA1.268	32,906	B	BG	R	RG
gRNA.BRCA1.269	32,942	B	BG	R	RB
gRNA.BRCA1.270	32,973	B	BG	R	BG
gRNA.BRCA1.271	33,121	B	BG	R	G
gRNA.BRCA1.272	33,128	B	BG	R	B
gRNA.BRCA1.273	33,335	B	BG	G	RGB
gRNA.BRCA1.274	33,496	B	BG	G	RG
gRNA.BRCA1.275	33,994	B	BG	G	RB
gRNA.BRCA1.276	34,521	B	BG	G	BG
gRNA.BRCA1.277	34,751	B	BG	G	R
gRNA.BRCA1.278	34,892	B	BG	B	RGB
gRNA.BRCA1.279	34,917	B	BG	B	RG
gRNA.BRCA1.280	34,988	B	BG	B	RB
gRNA.BRCA1.281	35,188	B	BG	B	BG

[0255]

gRNA.BRCA1.282	35,232	B	BG	B	R
gRNA.BRCA1.283	35,330	B	R	RGB	RG
gRNA.BRCA1.284	35,366	B	R	RGB	RB
gRNA.BRCA1.285	35,445	B	R	RGB	BG
gRNA.BRCA1.286	35,931	B	R	RGB	R
gRNA.BRCA1.287	36,116	B	R	RGB	G
gRNA.BRCA1.288	36,134	B	R	RGB	B
gRNA.BRCA1.289	36,764	B	R	RG	RGB
gRNA.BRCA1.290	36,884	B	R	RG	RB
gRNA.BRCA1.291	36,905	B	R	RG	BG
gRNA.BRCA1.292	37,050	B	R	RG	R
gRNA.BRCA1.293	37,334	B	R	RG	G
gRNA.BRCA1.294	37,867	B	R	RB	RGB
gRNA.BRCA1.295	38,007	B	R	RB	RG
gRNA.BRCA1.296	38,032	B	R	RB	BG
gRNA.BRCA1.297	38,275	B	R	RB	R
gRNA.BRCA1.298	38,325	B	R	RB	G
gRNA.BRCA1.299	38,470	B	R	RB	B
gRNA.BRCA1.300	38,471	B	R	BG	RGB
gRNA.BRCA1.301	38,472	B	R	BG	RG
gRNA.BRCA1.302	38,825	B	R	BG	RB
gRNA.BRCA1.303	38,842	B	R	BG	R
gRNA.BRCA1.304	39,099	B	R	BG	G
gRNA.BRCA1.305	39,109	B	R	BG	B
gRNA.BRCA1.306	39,114	B	R	G	RGB
gRNA.BRCA1.307	39,115	B	R	G	RG
gRNA.BRCA1.308	39,786	B	R	G	RB
gRNA.BRCA1.309	39,801	B	R	G	BG
gRNA.BRCA1.310	39,865	B	R	B	RGB
gRNA.BRCA1.311	40,010	B	R	B	RG
gRNA.BRCA1.312	40,054	B	R	B	RB
gRNA.BRCA1.313	40,221	B	G	RGB	RG
gRNA.BRCA1.314	40,663	B	G	RGB	RB
gRNA.BRCA1.315	40,927	B	G	RGB	BG
gRNA.BRCA1.316	41,058	B	G	RGB	R
gRNA.BRCA1.317	41,059	B	G	RGB	G

[0256]

gRNA.BRCA1.318	41,341	B	G	RG	RGB
gRNA.BRCA1.319	41,688	B	G	RG	RB
gRNA.BRCA1.320	42,346	RG	G	RG	BG
gRNA.BRCA1.321	42,400	RG	G	RG	R
gRNA.BRCA1.322	42,753	RG	G	RG	G
gRNA.BRCA1.323	42,862	RG	G	RG	B
gRNA.BRCA1.324	42,870	RG	G	RB	RGB
gRNA.BRCA1.325	42,891	RG	G	RB	RG
gRNA.BRCA1.326	42,912	RG	G	RB	BG
gRNA.BRCA1.327	42,943	RG	G	RB	R
gRNA.BRCA1.328	42,952	RG	G	RB	G
gRNA.BRCA1.329	42,964	RG	G	RB	B
gRNA.BRCA1.330	43,471	RG	G	BG	RGB
gRNA.BRCA1.331	43,514	RG	G	BG	RG
gRNA.BRCA1.332	43,594	RG	G	BG	RB
gRNA.BRCA1.333	43,783	RG	G	BG	R
gRNA.BRCA1.334	44,610	RG	G	BG	G
gRNA.BRCA1.335	44,956	RG	G	R	RGB
gRNA.BRCA1.336	45,085	RG	G	R	RG
gRNA.BRCA1.337	45,408	RG	G	R	RB
gRNA.BRCA1.338	45,642	RG	G	R	BG
gRNA.BRCA1.339	45,834	RG	G	B	RGB
gRNA.BRCA1.340	45,939	RG	G	B	BG
gRNA.BRCA1.341	46,018	RG	B	RGB	RG
gRNA.BRCA1.342	46,057	RG	B	RGB	RB
gRNA.BRCA1.343	46,159	RG	B	RGB	BG
gRNA.BRCA1.344	46,184	RG	B	RGB	G
gRNA.BRCA1.345	46,866	RG	B	RG	RGB
gRNA.BRCA1.346	47,006	RG	B	RG	RB
gRNA.BRCA1.347	47,161	RG	B	RG	BG
gRNA.BRCA1.348	47,295	RG	B	RG	G
gRNA.BRCA1.349	47,316	RG	B	RB	RGB
gRNA.BRCA1.350	47,624	RG	B	RB	RG
gRNA.BRCA1.351	48,296	RG	B	RB	BG
gRNA.BRCA1.352	48,417	RG	B	RB	R
gRNA.BRCA1.353	48,597	RG	B	BG	RGB

[0257]

gRNA.BRCA1.354	48,709	RG	B	BG	RG
gRNA.BRCA1.355	48,741	RG	B	BG	RB
gRNA.BRCA1.356	49,338	RG	B	BG	G
gRNA.BRCA1.357	49,509	RG	B	R	RG
gRNA.BRCA1.358	49,850	RG	B	R	BG
gRNA.BRCA1.359	50,036	RG	B	G	RG
gRNA.BRCA1.360	50,332	RG	B	G	RB
gRNA.BRCA1.361	50,553	RG	RGB	RG	RGB
gRNA.BRCA1.362	50,611	RG	RGB	RG	RB
gRNA.BRCA1.363	50,725	RG	RGB	RG	BG
gRNA.BRCA1.364	51,015	RG	RGB	RG	R
gRNA.BRCA1.365	51,082	RG	RGB	RG	G
gRNA.BRCA1.366	51,519	RG	RGB	RG	B
gRNA.BRCA1.367	51,797	RG	RGB	RB	RGB
gRNA.BRCA1.368	51,939	RG	RGB	RB	RG
gRNA.BRCA1.369	51,964	RG	RGB	RB	BG
gRNA.BRCA1.370	52,106	RG	RGB	RB	R
gRNA.BRCA1.371	52,142	RG	RGB	RB	G
gRNA.BRCA1.372	52,240	RG	RGB	RB	B
gRNA.BRCA1.373	52,265	RG	RGB	BG	RGB
gRNA.BRCA1.374	52,396	RG	RGB	BG	RG
gRNA.BRCA1.375	52,687	RG	RGB	BG	RB
gRNA.BRCA1.376	52,906	RG	RGB	BG	R
gRNA.BRCA1.377	53,519	RG	RGB	BG	G
gRNA.BRCA1.378	54,362	RG	RGB	BG	B
gRNA.BRCA1.379	54,794	RG	RGB	R	RGB
gRNA.BRCA1.380	54,956	RG	RGB	R	RG
gRNA.BRCA1.381	54,957	RG	RGB	R	RB
gRNA.BRCA1.382	54,976	RG	RGB	R	BG
gRNA.BRCA1.383	55,494	RG	RGB	R	G
gRNA.BRCA1.384	55,623	RG	RGB	G	RGB
gRNA.BRCA1.385	55,661	RG	RGB	G	RG
gRNA.BRCA1.386	55,668	RG	RGB	G	RB
gRNA.BRCA1.387	55,678	RG	RGB	G	BG
gRNA.BRCA1.388	56,191	RG	RGB	G	R
gRNA.BRCA1.389	56,192	RG	RGB	B	RGB

[0258]

gRNA.BRCA1.390	56,200	RG	RGB	B	RG
gRNA.BRCA1.391	57,155	RG	RGB	B	RB
gRNA.BRCA1.392	57,180	RG	RGB	B	BG
gRNA.BRCA1.393	57,839	RG	RGB	B	R
gRNA.BRCA1.394	57,864	RG	RGB	B	G
gRNA.BRCA1.395	58,010	RG	RG	RGB	RG
gRNA.BRCA1.396	58,204	RG	RG	RGB	RB
gRNA.BRCA1.397	58,282	RG	RG	RGB	BG
gRNA.BRCA1.398	58,325	RG	RG	RGB	R
gRNA.BRCA1.399	58,449	RG	RG	RGB	G
gRNA.BRCA1.400	58,474	RG	RG	RGB	B
gRNA.BRCA1.401	58,672	RG	RG	RB	RGB
gRNA.BRCA1.402	58,835	RG	RG	RB	RG
gRNA.BRCA1.403	59,401	RG	RG	RB	BG
gRNA.BRCA1.404	59,800	RG	RG	RB	R
gRNA.BRCA1.405	59,983	RG	RG	RB	G
gRNA.BRCA1.406	60,099	RG	RG	RB	B
gRNA.BRCA1.407	60,227	RG	RG	BG	RGB
gRNA.BRCA1.408	60,231	RG	RG	BG	RG
gRNA.BRCA1.409	60,248	RG	RG	BG	RB
gRNA.BRCA1.410	60,252	RG	RG	BG	R
gRNA.BRCA1.411	61,036	RG	RG	BG	G
gRNA.BRCA1.412	61,740	RG	RG	BG	B
gRNA.BRCA1.413	61,926	RG	RG	R	RGB
gRNA.BRCA1.414	62,214	RG	RG	R	RG
gRNA.BRCA1.415	62,304	RG	RG	R	RB
gRNA.BRCA1.416	62,672	RG	RG	R	BG
gRNA.BRCA1.417	62,820	RG	RG	R	G
gRNA.BRCA1.418	63,273	RG	RG	R	B
gRNA.BRCA1.419	63,414	RG	RG	G	RGB
gRNA.BRCA1.420	63,439	RG	RG	G	RG
gRNA.BRCA1.421	63,505	RG	RG	G	RB
gRNA.BRCA1.422	64,068	RG	RG	G	BG
gRNA.BRCA1.423	64,187	RG	RG	G	R
gRNA.BRCA1.424	64,302	RG	RG	G	B
gRNA.BRCA1.425	64,404	RG	RG	B	RGB

[0259]

gRNA.BRCA1.426	64,411	RG	RG	B	RG
gRNA.BRCA1.427	64,455	RB	RG	B	RB
gRNA.BRCA1.428	64,479	RB	RG	B	BG
gRNA.BRCA1.429	64,500	RB	RG	B	R
gRNA.BRCA1.430	65,444	RB	RG	B	G
gRNA.BRCA1.431	65,692	RB	BG	RGB	RG
gRNA.BRCA1.432	65,759	RB	BG	RGB	RB
gRNA.BRCA1.433	66,292	RB	BG	RGB	BG
gRNA.BRCA1.434	66,442	RB	BG	RGB	R
gRNA.BRCA1.435	66,751	RB	BG	RGB	G
gRNA.BRCA1.436	66,897	RB	BG	RGB	B
gRNA.BRCA1.437	66,922	RB	BG	RG	RGB
gRNA.BRCA1.438	67,204	RB	BG	RG	RB
gRNA.BRCA1.439	67,370	RB	BG	RG	BG
gRNA.BRCA1.440	67,444	RB	BG	RG	R
gRNA.BRCA1.441	67,977	RB	BG	RG	G
gRNA.BRCA1.442	68,830	RB	BG	RG	B
gRNA.BRCA1.443	68,873	RB	BG	RB	RGB
gRNA.BRCA1.444	68,912	RB	BG	RB	RG
gRNA.BRCA1.445	69,130	RB	BG	RB	BG
gRNA.BRCA1.446	69,271	RB	BG	RB	R
gRNA.BRCA1.447	69,436	RB	BG	RB	G
gRNA.BRCA1.448	69,925	RB	BG	RB	B
gRNA.BRCA1.449	70,004	RB	BG	R	RGB
gRNA.BRCA1.450	70,147	RB	BG	R	RG
gRNA.BRCA1.451	70,290	RB	BG	R	RB
gRNA.BRCA1.452	70,432	RB	BG	R	BG
gRNA.BRCA1.453	70,467	RB	BG	R	G
gRNA.BRCA1.454	70,754	RB	BG	R	B
gRNA.BRCA1.455	70,800	RB	BG	G	RGB
gRNA.BRCA1.456	70,924	RB	BG	G	RG
gRNA.BRCA1.457	71,133	RB	BG	G	RB
gRNA.BRCA1.458	71,158	RB	BG	G	BG
gRNA.BRCA1.459	71,397	RB	BG	G	R
gRNA.BRCA1.460	71,398	RB	BG	B	RGB
gRNA.BRCA1.461	71,773	RB	BG	B	RG

[0260]

gRNA.BRCA1.462	71,798	RB	BG	B	RB
gRNA.BRCA1.463	71,915	RB	BG	B	BG
gRNA.BRCA1.464	72,051	RB	BG	B	R
gRNA.BRCA1.465	73,315	RB	BG	B	G
gRNA.BRCA1.466	73,457	RB	R	RGB	RG
gRNA.BRCA1.467	73,797	RB	R	RGB	RB
gRNA.BRCA1.468	73,909	RB	R	RGB	BG
gRNA.BRCA1.469	73,951	RB	R	RGB	R
gRNA.BRCA1.470	74,973	RB	R	RGB	G
gRNA.BRCA1.471	74,998	RB	R	RGB	B
gRNA.BRCA1.472	75,052	RB	R	RG	RGB
gRNA.BRCA1.473	75,139	RB	R	RG	RB
gRNA.BRCA1.474	75,148	RB	R	RG	BG
gRNA.BRCA1.475	75,175	RB	R	RG	R
gRNA.BRCA1.476	75,837	RB	R	RG	G
gRNA.BRCA1.477	75,851	RB	R	RB	RGB
gRNA.BRCA1.478	75,926	RB	R	RB	RG
gRNA.BRCA1.479	75,927	RB	R	RB	BG
gRNA.BRCA1.480	76,170	RB	R	RB	R
gRNA.BRCA1.481	76,487	RB	R	RB	G
gRNA.BRCA1.482	77,242	RB	R	RB	B
gRNA.BRCA1.483	77,401	RB	R	BG	RGB
gRNA.BRCA1.484	77,803	RB	R	BG	RG
gRNA.BRCA1.485	77,970	RB	R	BG	RB
gRNA.BRCA1.486	78,230	RB	R	BG	R
gRNA.BRCA1.487	78,231	RB	R	BG	G
gRNA.BRCA1.488	78,277	RB	R	G	RGB
gRNA.BRCA1.489	78,625	RB	R	G	RG
gRNA.BRCA1.490	78,933	RB	R	G	RB
gRNA.BRCA1.491	78,958	RB	R	G	BG
gRNA.BRCA1.492	79,186	RB	R	B	RG
gRNA.BRCA1.493	79,211	RB	R	B	RB
gRNA.BRCA1.494	79,475	RB	G	RGB	RG
gRNA.BRCA1.495	79,600	RB	G	RGB	RB
gRNA.BRCA1.496	79,654	RB	G	RGB	BG
gRNA.BRCA1.497	79,679	RB	G	RGB	R

[0261]

gRNA.BRCA1.498	79,998	RB	G	RGB	G
gRNA.BRCA1.499	80,198	RB	G	RGB	B
gRNA.BRCA1.500	80,737	RB	G	RG	RGB
gRNA.BRCA1.501	80,877	RB	G	RG	RB
gRNA.BRCA1.502	80,902	RB	G	RG	BG
gRNA.BRCA1.503	80,986	RB	G	RG	R
gRNA.BRCA1.504	81,300	RB	G	RG	G
gRNA.BRCA1.505	81,339	RB	G	RB	RGB
gRNA.BRCA1.506	81,479	RB	G	RB	RG
gRNA.BRCA1.507	81,765	RB	G	RB	BG
gRNA.BRCA1.508	81,907	RB	G	RB	R
gRNA.BRCA1.509	81,932	RB	G	RB	G
gRNA.BRCA1.510	82,467	RB	G	RB	B
gRNA.BRCA1.511	82,513	RB	G	BG	RGB
gRNA.BRCA1.512	82,521	RB	G	BG	RG
gRNA.BRCA1.513	82,522	RB	G	BG	RB
gRNA.BRCA1.514	82,537	RB	G	BG	R
gRNA.BRCA1.515	82,729	RB	G	BG	G
gRNA.BRCA1.516	82,919	RB	G	R	RGB
gRNA.BRCA1.517	83,157	RB	G	R	RG
gRNA.BRCA1.518	83,334	RB	G	R	RB
gRNA.BRCA1.519	83,681	RB	G	R	BG
gRNA.BRCA1.520	84,189	RB	G	B	RGB
gRNA.BRCA1.521	84,452	RB	G	B	BG
gRNA.BRCA1.522	84,618	RB	B	RGB	RG
gRNA.BRCA1.523	85,185	RB	B	RGB	RB
gRNA.BRCA1.524	85,276	RB	B	RGB	BG
gRNA.BRCA1.525	85,394	RB	B	RG	RGB
gRNA.BRCA1.526	85,527	RB	B	RG	RB
gRNA.BRCA1.527	85,571	RB	B	RG	BG
gRNA.BRCA1.528	85,596	RB	B	RG	R
gRNA.BRCA1.529	85,683	RB	B	RG	G
gRNA.BRCA1.530	85,936	RB	B	RB	RGB
gRNA.BRCA1.531	85,955	RB	B	RB	RG
gRNA.BRCA1.532	86,055	RB	B	RB	BG
gRNA.BRCA1.533	86,194	GB	B	BG	RGB

[0262]

gRNA.BRCA1.534	86,303	GB	B	BG	RG
gRNA.BRCA1.535	86,310	GB	B	BG	RB
gRNA.BRCA1.536	86,697	GB	B	BG	R
gRNA.BRCA1.537	86,810	GB	B	R	RGB
gRNA.BRCA1.538	86,840	GB	B	R	RG
gRNA.BRCA1.539	86,960	GB	B	R	BG
gRNA.BRCA1.540	87,126	GB	B	G	RG
gRNA.BRCA1.541	87,199	GB	B	G	BG
gRNA.BRCA1.542	87,247	GB	RGB	RG	RGB
gRNA.BRCA1.543	87,355	GB	RGB	RG	RB
gRNA.BRCA1.544	88,350	GB	RGB	RG	BG
gRNA.BRCA1.545	88,419	GB	RGB	RG	R
gRNA.BRCA1.546	88,446	GB	RGB	RG	G
gRNA.BRCA1.547	88,455	GB	RGB	RG	B
gRNA.BRCA1.548	88,543	GB	RGB	RB	RGB
gRNA.BRCA1.549	88,597	GB	RGB	RB	RG
gRNA.BRCA1.550	88,604	GB	RGB	RB	BG
gRNA.BRCA1.551	88,622	GB	RGB	RB	R
gRNA.BRCA1.552	88,720	GB	RGB	RB	G
gRNA.BRCA1.553	88,902	GB	RGB	RB	B
gRNA.BRCA1.554	88,909	GB	RGB	BG	RGB
gRNA.BRCA1.555	88,927	GB	RGB	BG	RG
gRNA.BRCA1.556	88,955	GB	RGB	BG	RB
gRNA.BRCA1.557	88,960	GB	RGB	BG	R
gRNA.BRCA1.558	88,998	GB	RGB	BG	G
gRNA.BRCA1.559	89,029	GB	RGB	BG	B
gRNA.BRCA1.560	89,153	GB	RGB	R	RGB
gRNA.BRCA1.561	89,207	GB	RGB	R	RG
gRNA.BRCA1.562	89,351	GB	RGB	R	RB
gRNA.BRCA1.563	89,515	GB	RGB	R	BG
gRNA.BRCA1.564	90,439	GB	RGB	R	G
gRNA.BRCA1.565	90,440	GB	RGB	G	RGB
gRNA.BRCA1.566	90,441	GB	RGB	G	RG
gRNA.BRCA1.567	90,462	GB	RGB	G	RB
gRNA.BRCA1.568	90,700	GB	RGB	G	BG
gRNA.BRCA1.569	90,736	GB	RGB	G	R

[0263]

gRNA.BRCA1.570	90,914	GB	RGB	G	B
gRNA.BRCA1.571	91,209	GB	RGB	B	RGB
gRNA.BRCA1.572	91,284	GB	RGB	B	RG
gRNA.BRCA1.573	91,323	GB	RGB	B	RB
gRNA.BRCA1.574	91,424	GB	RGB	B	BG
gRNA.BRCA1.575	91,460	GB	RGB	B	R
gRNA.BRCA1.576	91,603	GB	RG	RGB	RG
gRNA.BRCA1.577	91,678	GB	RG	RGB	RB
gRNA.BRCA1.578	91,710	GB	RG	RGB	BG
gRNA.BRCA1.579	92,123	GB	RG	RGB	R
gRNA.BRCA1.580	92,226	GB	RG	RGB	G
gRNA.BRCA1.581	92,269	GB	RG	RGB	B
gRNA.BRCA1.582	92,347	GB	RG	RB	RGB
gRNA.BRCA1.583	92,363	GB	RG	RB	RG
gRNA.BRCA1.584	92,368	GB	RG	RB	BG
gRNA.BRCA1.585	92,384	GB	RG	RB	R
gRNA.BRCA1.586	92,520	GB	RG	RB	G
gRNA.BRCA1.587	92,536	GB	RG	RB	B
gRNA.BRCA1.588	92,540	GB	RG	BG	RGB
gRNA.BRCA1.589	92,541	GB	RG	BG	RG
gRNA.BRCA1.590	92,557	GB	RG	BG	RB
gRNA.BRCA1.591	92,604	GB	RG	BG	R
gRNA.BRCA1.592	92,690	GB	RG	BG	G
gRNA.BRCA1.593	92,820	GB	RG	BG	B
gRNA.BRCA1.594	92,899	GB	RG	R	RGB
gRNA.BRCA1.595	92,914	GB	RG	R	RG
gRNA.BRCA1.596	92,935	GB	RG	R	RB
gRNA.BRCA1.597	92,936	GB	RG	R	BG
gRNA.BRCA1.598	92,957	GB	RG	R	G
gRNA.BRCA1.599	92,978	GB	RG	R	B
gRNA.BRCA1.600	92,979	GB	RG	G	RGB
gRNA.BRCA1.601	93,001	GB	RG	G	RG
gRNA.BRCA1.602	93,022	GB	RG	G	RB
gRNA.BRCA1.603	93,029	GB	RG	G	BG
gRNA.BRCA1.604	93,045	GB	RG	G	R
gRNA.BRCA1.605	93,067	GB	RG	B	RGB

[0264]

gRNA.BRCA1.606	93,088	GB	RG	B	RG
gRNA.BRCA1.607	93,107	GB	RG	B	RB
gRNA.BRCA1.608	93,108	GB	RG	B	BG
gRNA.BRCA1.609	93,123	GB	RG	B	G
gRNA.BRCA1.610	93,134	GB	RB	RGB	RG
gRNA.BRCA1.611	93,140	GB	RB	RGB	RB
gRNA.BRCA1.612	93,141	GB	RB	RGB	BG
gRNA.BRCA1.613	93,173	GB	RB	RGB	R
gRNA.BRCA1.614	93,364	GB	RB	RGB	G
gRNA.BRCA1.615	93,611	GB	RB	RGB	B
gRNA.BRCA1.616	93,612	GB	RB	RG	RGB
gRNA.BRCA1.617	93,774	GB	RB	RG	RB
gRNA.BRCA1.618	93,999	GB	RB	RG	BG
gRNA.BRCA1.619	94,032	GB	RB	RG	R
gRNA.BRCA1.620	94,130	GB	RB	RG	G
gRNA.BRCA1.621	94,341	GB	RB	RG	B
gRNA.BRCA1.622	94,367	GB	RB	BG	RGB
gRNA.BRCA1.623	94,388	GB	RB	BG	RG
gRNA.BRCA1.624	94,406	GB	RB	BG	RB
gRNA.BRCA1.625	94,429	GB	RB	BG	R
gRNA.BRCA1.626	94,511	GB	RB	BG	G
gRNA.BRCA1.627	94,577	GB	RB	BG	B
gRNA.BRCA1.628	94,608	GB	RB	R	RGB
gRNA.BRCA1.629	94,644	GB	RB	R	RG
gRNA.BRCA1.630	94,645	GB	RB	R	RB
gRNA.BRCA1.631	94,666	GB	RB	R	BG
gRNA.BRCA1.632	94,674	GB	RB	R	G
gRNA.BRCA1.633	94,712	GB	RB	R	B
gRNA.BRCA1.634	94,713	GB	RB	G	RGB
gRNA.BRCA1.635	94,725	GB	RB	G	RG
gRNA.BRCA1.636	94,803	GB	RB	G	RB
gRNA.BRCA1.637	94,804	GB	RB	G	BG
gRNA.BRCA1.638	94,816	GB	RB	G	R
gRNA.BRCA1.639	94,850	RGB	RB	G	B
gRNA.BRCA1.640	94,904	RGB	RB	B	RGB
gRNA.BRCA1.641	94,905	RGB	RB	B	RG

[0265]

gRNA.BRCA1.642	94,922	RGB	RB	B	RB
gRNA.BRCA1.643	94,923	RGB	RB	B	BG
gRNA.BRCA1.644	94,965	RGB	RB	B	R
gRNA.BRCA1.645	94,966	RGB	RB	B	G
gRNA.BRCA1.646	94,982	RGB	R	RGB	RG
gRNA.BRCA1.647	94,991	RGB	R	RGB	RB
gRNA.BRCA1.648	95,012	RGB	R	RGB	BG
gRNA.BRCA1.649	95,018	RGB	R	RGB	R
gRNA.BRCA1.650	95,033	RGB	R	RGB	G
gRNA.BRCA1.651	95,038	RGB	R	RG	RGB
gRNA.BRCA1.652	95,039	RGB	R	RG	RB
gRNA.BRCA1.653	95,636	RGB	R	RG	BG
gRNA.BRCA1.654	95,714	RGB	R	RG	R
gRNA.BRCA1.655	95,802	RGB	R	RG	G
gRNA.BRCA1.656	95,856	RGB	R	RB	RGB
gRNA.BRCA1.657	95,881	RGB	R	RB	RG
gRNA.BRCA1.658	96,183	RGB	R	RB	BG
gRNA.BRCA1.659	96,409	RGB	R	RB	R
gRNA.BRCA1.660	96,593	RGB	R	RB	G
gRNA.BRCA1.661	96,594	RGB	R	RB	B
gRNA.BRCA1.662	96,879	RGB	R	BG	RGB
gRNA.BRCA1.663	97,267	RGB	R	BG	RG
gRNA.BRCA1.664	97,461	RGB	R	BG	RB
gRNA.BRCA1.665	97,550	RGB	R	BG	R
gRNA.BRCA1.666	97,586	RGB	R	BG	G
gRNA.BRCA1.667	98,126	RGB	R	G	RGB
gRNA.BRCA1.668	98,248	RGB	R	G	RG
gRNA.BRCA1.669	98,443	RGB	R	G	RB
gRNA.BRCA1.670	99,499	RGB	R	G	BG
gRNA.BRCA1.671	99,578	RGB	R	B	RGB
gRNA.BRCA1.672	99,858	RGB	R	B	RB
gRNA.BRCA1.673	99,989	RGB	R	B	BG
gRNA.BRCA1.674	100,077	RGB	G	RGB	RG
gRNA.BRCA1.675	100,078	RGB	G	RGB	RB
gRNA.BRCA1.676	100,252	RGB	G	RGB	BG
gRNA.BRCA1.677	100,253	RGB	G	RGB	R

[0266]

gRNA.BRCA1.678	100,267	RGB	G	RGB	G
gRNA.BRCA1.679	100,333	RGB	G	RG	RGB
gRNA.BRCA1.680	100,512	RGB	G	RG	RB
gRNA.BRCA1.681	100,551	RGB	G	RG	BG
gRNA.BRCA1.682	100,672	RGB	G	RG	R
gRNA.BRCA1.683	100,697	RGB	G	RG	G
gRNA.BRCA1.684	100,848	RGB	G	RG	B
gRNA.BRCA1.685	100,884	RGB	G	RB	RGB
gRNA.BRCA1.686	100,942	RGB	G	RB	RG
gRNA.BRCA1.687	101,036	RGB	G	RB	BG
gRNA.BRCA1.688	101,094	RGB	G	RB	R
gRNA.BRCA1.689	101,186	RGB	G	RB	G
gRNA.BRCA1.690	101,204	RGB	G	RB	B
gRNA.BRCA1.691	101,219	RGB	G	BG	RGB
gRNA.BRCA1.692	101,286	RGB	G	BG	RG
gRNA.BRCA1.693	101,483	RGB	G	BG	RB
gRNA.BRCA1.694	101,488	RGB	G	BG	R
gRNA.BRCA1.695	101,709	RGB	G	BG	G
gRNA.BRCA1.696	101,841	RGB	G	BG	B
gRNA.BRCA1.697	101,844	RGB	G	R	RGB
gRNA.BRCA1.698	101,845	RGB	G	R	RG
gRNA.BRCA1.699	101,846	RGB	G	R	RB
gRNA.BRCA1.700	101,909	RGB	G	R	BG
gRNA.BRCA1.701	101,972	RGB	B	RGB	RG
gRNA.BRCA1.702	102,080	RGB	B	RGB	RB
gRNA.BRCA1.703	102,348	RGB	B	RGB	BG
gRNA.BRCA1.704	102,360	RGB	B	RG	RGB
gRNA.BRCA1.705	102,471	RGB	B	RG	RB
gRNA.BRCA1.706	102,654	RGB	B	RG	BG
gRNA.BRCA1.707	102,655	RGB	B	RB	RGB
gRNA.BRCA1.708	102,823	RGB	B	RB	RG
gRNA.BRCA1.709	102,882	RGB	B	RB	BG
gRNA.BRCA1.710	102,889	RGB	B	RB	G
gRNA.BRCA1.711	102,981	RGB	B	BG	RGB
gRNA.BRCA1.712	103,008	RGB	B	BG	RG
gRNA.BRCA1.713	103,018	RGB	B	BG	RB

[0267]

gRNA.BRCA1.714	103,039	RGB	B	BG	R
gRNA.BRCA1.715	103,059	RGB	B	R	RG
gRNA.BRCA1.716	103,143	RGB	B	R	BG
gRNA.BRCA1.717	103,168	RGB	B	G	RG
gRNA.BRCA1.718	103,438	RGB	RGB	RG	RGB
gRNA.BRCA1.719	103,446	RGB	RGB	RG	RB
gRNA.BRCA1.720	103,459	RGB	RGB	RG	BG
gRNA.BRCA1.721	103,472	RGB	RGB	RG	R
gRNA.BRCA1.722	103,555	RGB	RGB	RG	G
gRNA.BRCA1.723	103,559	RGB	RGB	RB	RGB
gRNA.BRCA1.724	103,606	RGB	RGB	RB	RG
gRNA.BRCA1.725	103,670	RGB	RGB	RB	BG
gRNA.BRCA1.726	103,699	RGB	RGB	RB	R
gRNA.BRCA1.727	103,700	RGB	RGB	RB	G
gRNA.BRCA1.728	103,738	RGB	RGB	RB	B
gRNA.BRCA1.729	103,755	RGB	RGB	BG	RGB
gRNA.BRCA1.730	103,756	RGB	RGB	BG	RG
gRNA.BRCA1.731	103,769	RGB	RGB	BG	RB
gRNA.BRCA1.732	103,790	RGB	RGB	BG	R
gRNA.BRCA1.733	103,811	RGB	RGB	BG	G
gRNA.BRCA1.734	103,824	RGB	RGB	R	RGB
gRNA.BRCA1.735	103,880	RGB	RGB	R	RG
gRNA.BRCA1.736	103,881	RGB	RGB	R	RB
gRNA.BRCA1.737	104,232	RGB	RGB	R	BG
gRNA.BRCA1.738	104,563	RGB	RGB	G	RGB
gRNA.BRCA1.739	104,659	RGB	RGB	G	RG
gRNA.BRCA1.740	105,115	RGB	RGB	G	RB
gRNA.BRCA1.741	105,270	RGB	RGB	G	BG

[0268]

[0269] 또 다른 예는 오리가미 코드를 이용한다 (문헌 [Nat Chem. 2012 Oct;4(10):832-9.] 참조); 도 8을 참조할 수 있다. 핵산 오리가미 방법은 관련 기술분야의 통상의 기술자에게 공지되어 있다. gRNA 테일을 표 3에 열거된 오리가미 멀티컬러 스Kim에 따라 바코딩한다. 예상되는 순차적 gRNA 패턴을 형성하는 방식으로 바코드를 분리시킨다. 이러한 스Kim에 따라, 오리가미는 gRNA 테일 프로브이다. 각 오리가미 멀티컬러 프로브는 gRNA에 특이적인 부착점을 가지며, 이는 741개의 오리가미 프로브에 대하여 본 발명자들은 741개의 독특한 gRNA 테일 서열을 갖는다는 것을 의미한다. 오리가미 멀티컬러 프로브는, 고해상도 또는 초해상도 영상화를 통해 물리적으로 분석될 수 있는, 형광 표지에 대한 3개의 부착 위치, 즉 스폿(Spot)1, 스폿2, 스폿3을 갖는다. 각 스폿은 그와 연관된 다양한 컬러 코드를 갖는다. 이를 코드의 조합이 오리가미 멀티컬러 프로브의 정체성을 제공한다.

[0270]

gRNA 바코드는 그의 각 형광 표지된 오리가미 멀티컬러 프로브의 부가를 통해 검출된다. 실제 순차적 gRNA 패턴이 밝혀진다. 예상되지 않는 임의의 패턴은 계놈 재배열로서 확인될 것이다. 유전자좌의 영역을 분석하는 위치 RGB 코드와 비교하였을 때, 오리가미 코드는 단일 gRNA 수준에서 분석할 수 있으며, 유전자좌 상의 계놈 영역 또는 배열을 훨씬 더 고도하게, 및 상세하게 확인할 수 있다.

[0271]

Her2 검정법

[0272]

모든 침습성 유방암에 대하여 계내 혼성화 기술을 사용하여 Her2 상태를 측정하는 것이 권고된다. 현재 승인받은 임상 진단에 대한 최근 리뷰는 상이한 임상 검정법의 현 도전 과제들을 강조하였다 (참고문헌: Franchet C, Filleron T, Cayre A, Mounie E, Penault-Llorca F, Jacquemier J, Macgrogan G, Arnould L, Lacroix-Triki M. Instant-quality fluorescence in-situ hybridization as a new tool for HER2 testing in breast cancer: a comparative study. Histopathology. 2014 Jan;64(2):274-83).

- [0273] 다수의 도전 과제들이 이용가능한 현 검정법과 함께 보고되어 있다. 우선 한 가지로서, 전통적인 FISH는 많은 시간이 소요되는 프로세스로서, 전형적으로 혼성화를 위해 샘플을 제조하는데 1일이 소요되지만 (파라핀 제거, 전처리, 펩신 분해, 변성), 많은 시간이 소요되는 혼성화는 12-24 h 동안 수행되고, 다음날 혼성화를 철저하게 세척하고, 시각화를 수행한다.
- [0274] 본원에 기술된 예시적인 측면에 따라, 본 방법은 샘플의 어떤 전처리, 펩신 분해, 변성도 필요로 하지 않는다. Cas9-gRNA 복합체는 고정된 샘플 중 천연 이중 가닥 DNA에 쉽게, 및 특이적으로 혼성화할 수 있다. 또한 혼성화되지 않는 복합체는 세척으로 쉽게 제거된다. 따라서, 전반적인 프로세스는 하루 이내에 수행될 수 있다. 이러한 타임라인은, FISH 검정법을 전통적으로 진행하는 Her2 면역조직화학 검정법과 유사하다. 또한, 코딩 방법은 IHC 및 Cas9-gRNA FISH, 둘 모두가 동시에 진행될 수 있도록 허용할 것이다.
- [0275] 정확성 및 검정법들 사이의 불일치 또한 문제점인 것으로 보고되고 있다. 유방암 치료 비용이 높고, 대개는 수년간 진행된다는 점을 고려해 볼 때, 잘못된 치료는 사회적으로 및 경제적으로 상당한 영향을 초래하게 된다. 본원에 기술된 특정 측면에 따라, Cas9-gRNA는 중복 및 공동-국제화를 제공하는 방식으로 코딩된다. 이상 현상과 비교하여 양성 샘플이 구별될 수 있다. 또한, 본원에 기술된 방법을 통해 샘플을 재프로빙할 수 있고, 즉 Cas9-gRNA는 표적 DNA에 결합된 상태로 그대로 유지되면서, 프로브는 제거될 수 있고, 새로운 회차의 프로빙이 더욱 엄격한 조건하에서 수행될 수 있다.
- [0276] 현행 FISH 검정법의 경우, 비용이 수백 달러가 든다 (참고: 월드 와이드 웹사이트 ncbi.nlm.nih.gov/pmc/articles/PMC2706184/). 이는 부분적으로는, 일부 경우에는 프로브가 인터칼레이팅 형광 분자 (페쓰비전(PathVysis))와 함께 큰 벡터 내로 삽입되는 매우 긴 DNA 카피이거나, 또는 값비싼 웹티드 핵산 (PNA) 프로브 (팜Dx(pharmDx))인 것인 프로브용으로 사용되는 물질, 또는 DNA 프로브 (INFORM) 검출을 위해 항체 및 2차 항체 사용에 기인하는 것이다. 본원에 기술된 방법은 FISH 검정법에 선행하는 전통적인 Her2 ICH 검사 (~ 100\$)와 유사한 비용이 들 것이다. Cas9는 박테리아 기원의 단백질이고, 효율적으로 발현되며, 저렴한 박테리아 시스템으로부터 정제된다. 이는 또한 상업적으로도 이용가능하다. 올리고뉴클레오티드 어레이 합성기에서의 gRNA 주형 합성이 DNA를 합성하는 현행 방법 중 가장 저렴한 방법이다 (0.0004\$/염기, 시판 가격). 어레이 합성 후, PCR을 사용하여 올리고뉴클레오티드를 증폭시킬 수 있으며, 이에 의해 무한대로 저렴하게 재증폭시킬 수 있는 gRNA 주형 폴을 생성할 수 있다. 시험관내 전사는 DNA 분자 1개당 500 내지 1,000개의 RNA 분자를 생성하는 것으로, 매우 효율적이다.
- [0277] Her2 검정법은 염색체 17의 비율로서 세포 1개당 Her2 유전자좌의 평균 카피수를 기록한다. 염색체 17에 대한 가장 일반적인 리포터는 동원체 영역의 알파 위성 반복부를 표적화한다. 이는 특징이 비교적 잘 규명되어 있고, 최대 1,000개의 단량체 반복부를 함유하는 영역이다 (문헌 [Waye JS, Willard HF. Structure, organization, and sequence of alpha satellite DNA from human chromosome 17: evidence for evolution by unequal crossing-over and an ancestral pentamer repeat shared with the human X chromosome. Mol Cell Biol. 1986 Sep;6(9):3156-65] 참조). 반복부는 본원에 기술된 Cas9-gRNA 방법에 맞게 특히 잘 적합화되어 있다. 이러한 경우, 오직 수개의 상이한 gRNA만이 요구된다. 하기 표 4에는 gRNA 표적 목록이 제시되어 있다. 이 목록은 특징이 잘 규명되어 있는 D17Z1 유전자좌로부터 추출되었고, 그로부터 다른 CEP17 프로브가 기반을 갖게 된다. 이 목록은 또한 오키프 CL(O'Keefe CL), 워버튼 PE(Warburton PE), 마테라 AG(Matera AG)에 의해 보고된 변이체도 포함할 것이다. 알파 위성 DNA 변이체에 대한 올리고뉴클레오티드 프로브는 FISH에 의해 상동성 염색체를 구별할 수 있다 (Hum Mol Genet. 1996 Nov;5(11):1793-9). 그러나, 본원에 기술된 방법의 경우, Cas9는 전형적으로 그의 5' 단부 부근에 1-2개의 미스매치를 허용하는 바, 이에 오직 11개의 gRNA만을 사용함으로써 상기 반복부 대부분을 표적화할 수 있다 (하기 표 5)

[0278]

<표 4>

표적 명칭	표적 서열
CEN17.1	GAGCGCTTCAGGCCTGTGGTGG
CEN17.2	GAGGGCTTGAGGCCTGTGGTGG
CEN17.3	GAGGGCTTGAGGTTGTGGTGG
CEN17.4	GGAATCTGCAAGTGGATATGTGG
CEN17.5	GTGTTGAAACTCTTTTGTTGTGG
CEN17.6	TTTCCAATCACTCTTGTTGTGG
CEN17.7	TTTGGAAACACTCTGTTGTGG

[0279]

CEN17.8	TTTTGAAACTCTTTCTGTGG
CEN17.9	ACACTGCTCTATCCATAGGAGG
CEN17.10	AGATATTTGACCGCTCTGAGG
CEN17.11	AGCGCTTCAGGCCTGTGGTGG
CEN17.12	AGGAATGTTCAACTCTGTGAGG
CEN17.13	AGGGCTTGAGGCCTGTGGTGG
CEN17.14	AGGGCTTGAGGTTGTGGTGG
CEN17.15	CATCACAGAGAAGCTCTGAGG
CEN17.16	CTGCATTCAACTCACAGTGTGG
CEN17.17	GAAAGGAAAGTTCAACTCGGGG
CEN17.18	GAATCTGCAAGTGGATATGTGG
CEN17.19	GAATGCAAACATCACGAAGAGG
CEN17.20	GCATATTTGACCTCTTGAGG
CEN17.21	GCTTCTGTTAGTTCTGTGCGG
CEN17.22	GCTTCTGTTAGTTCTGTGCGG
CEN17.23	GGACATTTGAGGGCTTGAGG
CEN17.24	GGACGTTGAGGGCTTGAGG
CEN17.25	GGAGATTGGAGCGCTTGAGG
CEN17.26	GGATATTTAGGCCTCTGTGAGG
CEN17.27	GGATATTTGACCACTCTGTGG
CEN17.28	GGATATTTGACCTCTGTGAGG
CEN17.29	GGGATCATTGCACTCTTGAGG
CEN17.30	TACTACCATAAGCCTAAAGCGG
CEN17.31	TATTTGAGAATGTGCAAGTGG
CEN17.32	TCCAAGACATCTCGGAGAGG
CEN17.33	TCCAACGAAATCCTCAGAGAGG
CEN17.34	TCCAACGAAATCCTCAGAGAGG
CEN17.35	TCCAACGAAATCCTCAGAGCGG
CEN17.36	TCCAACGAAATCTCAAAGAGG
CEN17.37	TCCAACGAAATGCTCAGAGAGG
CEN17.38	TCGAACGAAGGACACAGAGTGG
CEN17.39	TCGAACGAAGGCCACAGAGTGG
CEN17.40	TCTGCAAGTGGACATTGGAGG
CEN17.41	TCTGCAAGTGGACGTTGGAGG
CEN17.42	TGGAGCGCTTCAGGCCTGTGG
CEN17.43	TGGAGGGCTTGAGGCCTGTGG

[0280]

CEN17.44	TGGAGGGCTTGTGGTTGTGG
CEN17.45	TGTTGAAACTCTCTTTGTGG
CEN17.46	TTGTTGTGGAATGTCAAGTGG
CEN17.47	TTTCCAATCACTCTTGTGTGG
CEN17.48	TTTCTGTGGCATCTGCAAGGGG
CEN17.49	TTTGGAAACACTCTTGTGTGG
CEN17.50	TTTGTGTAGAATCTGCAAGTGG
CEN17.51	TTTGTGTGGAATCTGCAAGTGG
CEN17.52	TTTCGTAAGTGTCTACAAGTGG
CEN17.53	TTTGAAACTCTCTTGTGTGG
CEN17.54	TTTTCCAGAACATCTGCAAGTGG
CEN17.55	TTTTCTAGAACATCTGCAAGTGG
CEN17.56	TTTTGCAGGATCTACAAGTGG
CEN17.57	TTTTGTACAATCTACAAGTGG
CEN17.58	TTTTTAGAAACTGCAAGGGG
CEN17.59	TTTTTAGAAACTGCAAGTGG
CEN17.60	TTTTTAGGATCTGCAAGTGG
CEN17.61	TTTTGTGGAATCTGCAAGTGG

[0281]

[0282] <표 5>

gRNA명칭	gRNA 표적 서열
Cen17gRNA.1	GAGGGCTTGAGGCCTGTGG
Cen17gRNA.2	GTGTTGAAACTCTCTTTTG
Cen17gRNA.3	GACACTGCTCTATCCATAGG
Cen17gRNA.4	GAGGGCTTGAGGCCTGTGG
Cen17gRNA.5	GGAATCTGCAAGTGGATATG
Cen17gRNA.6	GGCATATTGGACCTCTTTG
Cen17gRNA.7	GTACTACCATAAGCCTAAAG
Cen17gRNA.8	GTCCAACGAAATCCTCAGAG
Cen17gRNA.9	GTCTGCAAGTGGACATTGG
Cen17gRNA.10	GTTGTTGTGGAATGTGCAAG
Cen17gRNA.11	GTTTGTGTAGAACATCTGCAAG

[0283]

[0284]

추가로, 본원에 기술된 방법을 통해 개인 맞춤형 의약에서 필수 정보가 되고, 치료에 영향을 줄 수 있는 대립유전자 변이체를 스크리닝할 수 있다. 예로서, Her2 변이체 I655V는 유방암 치료로서 타목시펜의 효능 감소와 관련이 있으며, 이는 고려되어야 한다 (문헌 [Chang NW, Chen DR, Chen FN, Lin C, Wu CT. HER2 codon 655 G-allele is associated with reductions in plasma high-density lipoprotein levels in breast cancer patients treated with tamoxifen. J Investig Med. 2011 Dec;59(8):1252-7. doi: 10.231/JIM.0b013e3182354923] 참조). 변이체를 확인하는데 확인된 두 gRNA 표적이 사용될 수 있다 (표적 1: TCTGACGTCCATCATCTCTGCGG, 및 표적 2: GCCAACACCAGCAGAGATGATGG). 대립유전자 특이적 바코드와 함께 4개의 gRNA (표적 1개당 2개)로 이루어진 한 세트를 이용하여, 표준 Her2 검정법의 일부로서, 이소류신 대립유전자 (ATC) 또는 발린 대립유전자 (GTC) 사이를 구별짓는 방법을 제공한다.

[0285]

염색체 17의 다른 영역도 표적화될 수 있다는 것도 이해한다. 예로서, Top2A는 Her2 유전자좌에 인접하게 위치하고, Her2와 함께 공동으로 증폭되는 것으로 관찰되었으며, 이는 또한 적절한 치료법 선택에 영향을 줄 수 있다. 문헌 [Smith K, Houlbrook S, Greenall M, Carmichael J, Harris AL. Topoisomerase II alpha co-amplification with erbB2 in human primary breast cancer and breast cancer cell lines: relationship to m-AMSA and mitoxantrone sensitivity. Oncogene. 1993 Apr;8(4):933-8]을 참조할 수 있다.

[0286]

Her2/CEN17을 스크리닝하는 것에 관한 FISH 프로토콜의 예는 하기와 같다. 면역조직화학법 및 FISH 검정법, 둘 모두 같은 샘플에 대해 수행하고자 하는 경우, 세포 또는 조직 샘플을 4°C에서 1h 내지 16h 동안 PBS 중 10% 파라포름알데히드 중의 현미경 적합 지지체 (예컨대, 깨끗한 보로실리케이트 현미경 슬라이드 또는 현미경 커버 슬립) 상에 고정시킨다. 대안적으로, 샘플을 -20°C에서 20분 내지 1h 동안 100% 메탄올 중에서 고정시킬 수 있다 (많은 단백질을 추출하면서, 샘플 중의 DNA를 유지시키는 데에는 메탄올이 더욱 효율적이다). 슬라이드에 부착시키기 전에 파라핀 포매된 조직 절편으로부터 파라핀을 제거하여야 하며, 이는 전형적으로는 크실렌 중에서 5분 동안 인큐베이션시킨 후, 에탄올로 2회, 및 물로 1회 세척함으로써 달성된다. 이어서, 고정 샘플을 PBST (0.5% 트윈-20을 포함하는 PBS 완충제)로 1회 세척한 후, 0.5% 트리톤 X-100을 함유하는 PBS와 함께 5분 동안 인큐ベ이션시킨다. 이어서, 샘플을 PBST 중에서 세척하고, Cas9-gRNA-표지 복합 혼합물을 5 mM MgCl₂를 포함하는 PBST 중 샘플에 첨가하고, 37°C에서 2h 내지 16h 동안 인큐ベ이션시킨다 (즉, 믹스는 Her2 및 CEN17에 대한 세트 및 그의 각각의 확인용 표지, 둘 모두를 함유한다). 인큐ベ이션 후, 37°C에서 PBST 중에서 3회 세척하여 비결합 Cas9를 제거한다. DAPI를 포함하는 안티페이드 탐재용 시약 (예컨대, DAPI를 함유하는 프로롱(ProLong) 또는 베타쉴드)을 샘플에 탐재하고, 실링한다. 유침용 63X 대물렌즈를 사용하여 샘플을 영상화한다. 대안적으로, 샘플을 4°C에서 빛 차단하에 유지시키면, 샘플은 수주 동안 안정적이다. 이어서, 종양 부위 중 적어도 20개의 핵을 계수하고, 그들 각각에 대하여 Her2 및 CEN17 국제화 포커스의 개수를 기록한다. 2013 ASCO/CAP 가이던스에 의해 규명된 바와 같은 현행 임상 실무에 따라 Her2 대 CEN17의 비율을 계산할 수 있고, 기록할 수 있다 (문헌 [Wolff AC, Hammond ME, Hicks DG, Dowsett M, McShane LM, Allison KH, Allred DC, Bartlett JM, Bilous M, Fitzgibbons P, Hanna W, Jenkins RB, Mangu PB, Paik S, Perez EA, Press MF, Spears PA, Vance GH, Viale G, Hayes DF; American Society of Clinical Oncology; College of American Pathologists. Recommendations for human epidermal growth factor receptor 2 testing in breast cancer: American Society of Clinical Oncology/College of American Pathologists clinical practice guideline update. J Clin Oncol. 2013 Nov 1;31(31):3997-4013] 참조).

[0287]

실시예 III

[0288]

도 1은 마우스 배아 섬유모세포 세포 핵의 주요 및 부차 위성, 및 텔로미어 영역의 계내 Cas9-gRNA 프로빙에 관한 것이다. 마우스 주요 위성 반복부 (Cy5), 부차 위성 반복부 (알렉사-488), 및 텔로미어 (Cy3)를 표적화하기 위해, 형광 표지된 UTP를 이용하여 3개의 gRNA를 합성하였다 (도 1 범례: 주요 위성 (A), 부차 위성 (B), 텔로미어 (C), 계놈 DNA의 DAPI 염색 (D), 및 중첩 사진 (E)). 본원에 기술된 프로토콜에 따라 PFA 고정 샘플에 첨가하기 이전에 gRNA가 Cas9와 복합체를 형성하게 하였다. 패턴은 상기 표적에 대한 예상 프로빙에 맞는다 (문헌 [Guenatri M, Bailly D, Maison C, Almouzni G. Mouse centric and pericentric satellite repeats form distinct functional heterochromatin. J Cell Biol. 2004 Aug 16;166(4):493-505] 참조). 63X/1.40 유침용 대물렌즈, 및 각 형광 채널에 적절한 LED 광원 및 필터가 장착된 자이스 악시오 옵저버 Z1(Zeiss Axio Observer) 상에서 사진 촬영하였다. 하기 범례는 도 1과 연관이 있다: A, 주요 위성; B, 부차 위성; C, 텔로미어; D, DNA의 DAPI 염색; E, 중첩.

[0289]

실시예 IV

[0290]

도 2는 도 1에 대한 대조군 실험에 대한 것이다. 마우스 주요 위성 반복부 (Cy5), 부차 위성 반복부 (알렉사-488), 및 텔로미어 (Cy3)를 표적화하는데 일반적으로 사용되는 FISH 올리고뉴클레오티드이며, 도 1에서 사용되는 것과 같은 Cas9 프로빙 프로토콜을 이용함으로써 사용되었다. DNA의 DAPI 염색 (D)을 제외하면, 다른 형광 신호는 도 1에서 포착된 사진보다 더 분산되어 있었고, 현미경 상에서 더 큰 대조를 이루어야 했다 (도 2 범례: 주요 위성 (A), 부차 위성 (B), 텔로미어 (C), 계놈 DNA의 DAPI 염색 (D), 및 중첩 사진 (E)). 변성되지 않은 경우, 이들 FISH 올리고는 그의 표적에의 혼성화를 막았고, 이는 또한 핵 밖에서 응집되었다. 63X/1.40 유침용 대물렌즈, 및 각 형광 채널에 적절한 LED 광원 및 필터가 장착된 자이스 악시오 옵저버 Z1 상에서 사진 촬영하였다. 하기 범례는 도 1과 연관이 있다: A, 주요 위성; B, 부차 위성; C, 텔로미어; D, DNA의 DAPI 염색; E, 중첩.

[0291]

실시예 V

[0292]

도 3은 천연 Cas9 결합 및 절단 활성이 비의존적이고, 마그네슘 이온의 존재 또는 부재에 의존한다는 것을 보여주는 Cas9-gRNA 겔 이동 및 절단 검정법에 관한 것이다. 하기와 같이 반응물을 제조하였다: 2 pmol의 뉴클레아제 활성 Cas9 (NEB), 2 pmol의 합성된 gRNA (L22116 위치의 람다 DNA 표적화), 및 0.2 pmol의 2 kb PCR 증폭된 DNA 단편 (람다 DNA L21333-L23332로부터의 것)을 혼합하고, 5 mM MgCl₂의 존재하에 (레인 1), 또는 마그네슘

부재하에 (레인 2) 37°C에서 1 h 동안 인큐베이션시켰다. 레인 1은 마그네슘 존재하의 2차 절단 생성물을 보여주는 반면, 레인 2는 절단 생성물은 존재하지 않지만, 여전히 그에 결합되어 있는 복합체에 의해 유발되는 작은 상향 이동을 보여준다. 레인 3 및 레인 4는 각각 레인 1 및 2와 유사한 조건이지만, 단, 예외적으로 반응물로부터 Cas9 단백질이 생략되어 있는 것인 조건하에서 수행되었다. 레인 3 및 4의 경우, 어떤 절단 및 어떤 이동도 관찰되지 않는다.

[0293] 실시예 VI

도 4A-4L은 gRNA 테일 바코드 및 프로빙의 예에 관한 것이다. 도 4A는 Cas9 단백질 (1), 및 돌출 테트라루프 (3) 및 바코드 (4)를 보유하는 돌출 gRNA 테일을 포함하는 복합체를 형성하는 gRNA (2)를 나타내는 것이다. 바코드는 검출가능한 모이어티, 또는 표지 (6)를 보유하는 혼성화 프로브 (5)에 의해 검출된다. 도 4B는 gRNA 테일이 1개 초과의 바코드에 의해 코딩될 수 있다는 것을 보여주는 것이다. 도 4B에서, 2개의 바코드 (4 및 7)가 도시되어 있는데, 이는 검출가능한 모이어티, 또는 표지 (6)를 보유하는, 그의 각 혼성화 프로브 (5 및 8)에 의해 검출될 수 있다. 도 4C에서, 혼성화 프로브는 신호를 증폭시키기 위해 다중의 검출가능한 모이어티, 또는 표지 (9)를 보유할 수 있다. 도 4D에서, 검출가능한 모이어티는 직접적으로 검출가능하지 않을 수도 있으며 (10), 검출가능한 모이어티에 대한 특이적인 친화성을 갖는 제2 검출가능한 작용제 (11)의 사용을 필요로 할 수 있다. 이러한 제2 작용제는 검출을 위해, 또는 검출되는 신호를 증폭시키기 위해 검출가능한 모이어티, 또는 표지 (12)를 보유한다. 도 4E 내지 4I는 롤링 서클 증폭을 통해 gRNA 테일을 프로빙하는 방법을 나타낸 것이다. 도 4E는 Cas9 단백질 (1), 및 돌출 테트라루프 (3) 및 바코드 (4)를 보유하는 gRNA 테일을 포함하는 복합체를 형성하는 gRNA (2)를 나타내는 것이다. 바코드는 프로브의 한쪽 영역에는 바코드에 대한 친화성을 갖고, 또 다른 영역에는 표지된-프로브 혼성화 표적 부위 (14)를 갖는, 고리형 혼성화 프로브 (13)에 의해 검출된다. 도 4F는 고리형 프로브 (13)가 gRNA 테일 (16)을 연장시키는데 있어서 롤링 서클 증폭 폴리머라제 (15)에 대한 주형으로서 작용할 수 있다는 것을 도시한 것이다. 도 4G는 롤링 서클 증폭이 다수의 가깝게 국제화된 표지된-프로브 혼성화 표적 부위 (18)를 갖는 국제화된 증폭된 혼성화 프로브 (17)를 생성한다는 것을 도시한 것이다. 도 4H는 표지된-프로브 혼성화 표적 부위가 검출가능한 프로브 (19) 첨가에 의해 검출가능하게 만들어질 수 있으며, 이에 의해 신호가 증폭된 gRNA 테일 프로브 (20)를 얻을 수 있다는 것을 도시한 것이다. 도 4I는 도 4H와 유사한 롤링 서클 증폭된 프로브 (20)를 도시한 것으로서, 이는 gRNA 바코드에 혼성화되는 고리형 프로브를 갖지 않고, 생성 및 표지되며 (21), 이에 의해 gRNA 그 자체에 대한 롤링 서클 증폭 및 표지 단계를 면할 수 있다. 이러한 오프-gRNA 생성된 프로브는 영역 (5)을 통해 도 4A에서와 같이 gRNA 테일의 바코드 영역 (4)에 혼성화할 수 있다. 도 4J 내지 4L은 핵산 자기-어셈블리를 통해 gRNA 테일을 프로빙하는 방법을 나타낸 것이다. 도 4J는 순차적 방식의 서로에 대한 핵산 프로브의 선형 어셈블리를 도시한 것이다. 바코드 혼성화 영역 (5), 어셈블리 영역 (21) 및 표지 (6)로 구성된 제1 단편을 먼저 gRNA 테일 바코드 (도 4A, 4)에 혼성화시킨다. 부분적으로 21 또는 22 또는 23에 상보적인 표지된 어셈블리 단편 (22 및 23)의 혼합물을 첨가하고, 이는, 반응물 중에 부분적으로 상보적인 단편이 존재하는 한, 그 기간 동안은 자기 어셈블리될 것이다. 그 구조 중 단지 일부분만이 본원에 도시되어 있다. 이는 앞서 기술된 혼성화 연쇄 반응 (HCR)과 유사하다. 문헌 [Dirks RM, Pierce NA. Triggered amplification by hybridization chain reaction. Proc Natl Acad Sci U S A. 2004 Oct 26;101(43):15275-8]을 참조할 수 있다. 도 4K는 초분지형 핵산 프로브 덴드리머 구조의 자기 어셈블리를 도시한 것이다. 바코드 혼성화 영역 (5), 어셈블리 영역 (24) 및 표지 (6)로 구성된 제1 단편을 먼저 gRNA 테일 바코드 (도 4A, 4)에 혼성화시킨다. 부분적으로 24 또는 25 또는 26에 상보적인 표지된 어셈블리 단편 (24, 25 및 26)의 혼합물을 첨가하고, 이는, 반응물 중에 부분적으로 상보적인 단편이 존재하는 한, 그 기간 동안은 초분지형 구조로 자기 어셈블리될 것이다. 그 구조 중 단지 일부분만이 본원에 도시되어 있다. 이러한 종류의 핵산 덴드리머는 앞서 기술된 바 있다. 문헌 [Li Y, Tseng YD, Kwon SY, D'Espaux L, Bunch JS, McEuen PL, Luo D. Controlled assembly of dendrimer-like DNA. Nat Mater. 2004 Jan;3(1):38-42]를 참조할 수 있다. 도 4L은 신호를 선형으로 증폭시키는 핵산 프로브의 분지의 어셈블리를 도시한 것이다. 바코드 혼성화 영역 (5), 어셈블리 영역 (27)으로 구성된 제1 단편을 먼저 gRNA 테일 바코드 (도 4A, 4)에 혼성화시킨다. 부분적으로 27 또는 29 또는 30에 상보적인 표지된 어셈블리 단편 (29 및 30)의 혼합물을 첨가하고, 이는 자기 어셈블리될 수 있다. 반응물 중에 부분적으로 상보적인 단편이 존재하는 한, 그 기간 동안은 검출가능한 프로브 (28)는 어셈블리 단편 (27) 및 (30)의 일부분에 혼성화할 수 있다. 그 구조 중 단지 일부분만이 본원에 도시되어 있다. 이는 앞서 기술된 분지형 DNA (bDNA)와 유사하다. 문헌 [Collins ML, Irvine B, Tyner D, Fine E, Zayati C, Chang C, Horn T, Ahle D, Detmer J, Shen LP, Kolberg J, Bushnell S, Urdea MS, Ho DD. A branched DNA signal amplification assay for quantification of nucleic acid targets below 100 molecules/ml. Nucleic Acids Res. 1997 Aug 1;25(15):2979-84]를 참조할 수 있다.

- [0295] 검출가능한 모이어티 (예컨대, 6, 9, 10, 12)가 여러 검출 수단, 예컨대 형광, 화학발광성 또는 발색성, 또는 공명 수단을 제공할 수 있다는 것을 이해한다. 혼성화 프로브 (예컨대, 5, 8, 9, 14)는 다양한 핵산 성분, 예컨대 DNA, RNA, 변형된 DNA, 변형된 RNA로 제조될 수 있다는 것을 이해한다.
- [0296] 실시예 VII
- [0297] 도 5A는 측면 유동 시험 시스템의 표면에 부착된 Cas9-gRNA 복합체에 관한 것이다. 조사된 DNA 집단을 시스템에 로딩하고, 측면 유동은 로딩된 위치에서 시험 구역으로 이동하는데, 여기서 표적-특이적 Cas9-gRNA는 특이적인 DNA에 결합하게 된다. 나머지 DNA는 계속해서 검정 단부를 통과하고, 대조군 구역에 의해 포획되는 반면, 표적 DNA는 시험 구역에서 Cas9-gRNA 복합체에 결합된 상태 그대로 유지된다. DNA 검출은 핵산 염색을 통해, DNA에의 검출가능한 프로브의 혼성화를 통해, 또는 검출가능한 프로브의 DNA에의 공유 부착을 통해, 또는 검출가능한 모이어티를 보유하는 특이적인 또는 범용 올리고뉴클레오티드 프라이머의 존재하에서 DNA의 효소적 증폭 (예컨대, 등온 증폭)에 의해 수행될 수 있다. 측면 유동 검정법에서 자주 사용되는 검출가능한 모이어티는 금 또는 은 나노입자를 포함한다. 다른 모이어티, 예컨대 비오틴, 디옥시게닌, 디니트로페닐, 플루오레세인은 보편적으로는 2차 검출 작용제 (예컨대, 비오틴, 디옥시게닌, 디니트로페닐 또는 플루오레세인에 특이적인 금 또는 은 코팅된 항체)와 함께 조합하여 사용된다. 도 5B는 예컨대, 소수성 또는 전하와 같이, DNA 또는 단백질을 유지하기 위한 상이한 특성을 갖는 두 물질을 사용하는 측면 유동 시스템에 관한 것이다. 예로서, 니트로셀룰로스 막은 DNA보다는 단백질을 더 호의적으로 유지하는 반면, 나일론 막은 단백질보다는 DNA에 결합한다. 조사하고자 하는 DNA 집단을 로딩할 때, Cas9-gRNA 복합체는 표적 DNA에 결합한다. Cas9 결합 DNA는 시험 구역의 DNA 보유 막에 의해 유지되는 반면, 비결합 Cas9는 통과하여 대조군 구역의 단백질 보유 막에 결합하게 될 것이다. DNA와의 Cas9-gRNA 복합체를 검출하는데 검출을 수행한다. gRNA는 본원에 기술된 방식으로 프로빙된다. 대안적으로, Cas9 단백질, 그 자체가 모이어티 (예컨대, 금 나노입자, 은 나노입자, 비오틴, 디옥시게닌, 디니트로페닐, 플루오레세인)를 보유할 수 있다. 도 5C는 조사하고자 하는 DNA 집단이 먼저 분자 상호작용 (예컨대, 전하, 소수성, 공유 상호작용, 또는 친화성 상호작용, 예컨대 비오틴-스트렙트아비딘)을 통해 시험 구역의 표면 상에서 포획되는 것인 측면 유동 시스템에 관한 것이다. 이어서, Cas9-gRNA는 시험 구역의 표면-포획된 표적 DNA에 결합된다. 나머지 비반응 Cas9-gRNA는 통과하여 대조군 구역에서 포획된다. DNA와의 Cas9-gRNA 복합체를 검출하는데 검출을 수행한다. gRNA는 본원에 기술된 방식으로 프로빙된다. 대안적으로, Cas9 단백질, 그 자체가 모이어티 (예컨대, 금 나노입자, 은 나노입자, 비오틴, 디옥시게닌, 디니트로페닐, 플루오레세인)를 보유할 수 있다.
- [0298] 실시예 VIII
- [0299] 가이드 RNA/Cas9 복합체를 사용하여 이루어지는 표적 DNA 프로빙
- [0300] 시험관내에서 가이드 RNA를 디자인하고, 제조하고, 발현하고, 정제하였다. 단일 가닥 DNA (ssDNA) 올리고뉴클레오티드를 합성한다 (IDT 또는 커스텀어레이 (CustomArray) 칩). 상기 올리고는 시험관내 RNA 합성을 위한 T7 전사 출발 부위, 및 프로빙에 사용되는 3' 단부 테일을 형성하는 연장부를 포함한다. T7 RNA 합성을 수행하고, RNA를 정제한다.
- [0301] 한 측면에 따라, DNA-PAINT를 사용하여 초해상도 영상화를 달성하기 위해 연장된 서열 또는 테일은 입체 장해를 최소화하도록, 낮은 복잡도 및 낮은 자유 에너지를 갖도록 디자인된다 (이는 2.7 nm 해상도를 달성할 수 있고, 8개의 DNA 염기로 이루어져 있는 바, 이에 Cas9의 끗프린트는 20 bp 미만이다). 디자인된 가이드 RNA는 20개의 상이한 PAINT 도킹 부위를 갖는데, 그 길이는 9 mer이고, A, T 및 (G가 아닌) C 염기로 이루어진 상이한 정렬을 포함하는 반면, PAINT 프로브는 형광단을 포함하는 상보적 서열이다.
- [0302] 영상화 사이에 프로브를 세척하여 제거하고, 새 프로브를 로딩함으로써, 형광단 1개당 20개의 프로브를 달성할 수 있고, 즉 형광단 1개당 20개의 상이한 gRNA가 영상화될 수 있다. 따라서, 4개의 형광단이 80개의 gRNA를 구별해낼 수 있다.
- [0303] 가이드 RNA 구조의 한 실시양태는 (5' → 3') gRNA-UUUUU-PAINTdock이다. 가이드 RNA 서열은 일반적으로 SpCas9와 함께 사용되는 단일 가이드 RNA의 최소 길이를 갖는다.
- [0304] 고해상도 영상화를 위해, 연장된 서열 또는 테일은 입체 장해를 최소화하도록, 낮은 복잡도 및 낮은 자유 에너지를 갖도록 디자인된다. 그러나, 고해상도 영상화의 경우, 더욱 긴 어닐링 영역이 사용된다. 연장된 서열 또는 테일의 구조는 서열 특이적 올리고로 프로빙될 수 있는 바코드로서의 역할을 할 수 있거나, 또는 후속하여 롤링 서클 증폭이 이루어지는 패드락 프로브로서의 역할을 할 수 있다.

- [0305] SpCas9는 E. 콜라이에서 발현되었다. 김순(Gibson) 등은 어셈블리를 이용하여 E. 콜라이에서의 발현에 적합한 플라스미드에서 Cas9 코딩 유전자를 어셈블리하고 클로닝한다. 유도한 후, 세포를 용해시키고, 단백질을 정제 한다.
- [0306] 추가로 정제하지 않고, 세포를 용해시킴으로써 표적 DNA를 수득하였다. 따라서, 샘플은 염색체 및 염색체외 기원 (예컨대, 플라스미드)의 DNA의 혼합물을 포함한다.
- [0307] Cas9, gRNA 및 표적 핵산이 복합체를 형성하는데 충분한 시간 (약 15분) 동안 Cas9를 gRNA와 혼합하고, DNA 샘플에 첨가한 후, 이를 검출 프로브에 첨가한다. 이어서, 영상화 데이터를 획득한다 (수초 내지 수분 소요).
- [0308] 나노포어 검출을 위해, Cas9-gRNA 복합체는 디자인된 간격으로 dsDNA 단편에 결합한다. 이어서, DNA는 나노포어 (또는 나노캡 전극)을 통과하여 또는 그에 근접하게 전위된다. 전류 변화를 측정한다: dsDNA는 특정 전류에서 이동하는 반면, DNA에 결합된 복합체는 전류를 부분적으로 차단할 것이며, 본 발명자들은 이를 전류 스파이크로서 기록할 수 있다. 시간 경과에 따른 상기 스파이크 이벤트를 분석함으로써, 표적 DNA 상의 Cas9/gRNA의 위치를 추론해낼 수 있고, 가이드 RNA 디자인에 기초한 예상 위치와 비교할 수 있다.
- [0309] 상기 방법을 사용하여, 다중의 표적을 검출함으로써 DNA 표적의 성질, 예컨대 동원체의 반복부 영역, 상기 반복부에 연결된 염색체의 정체성, (박테리아 게놈 또는 플라스미드에 함유되어 있는) 약물 내성 유전자의 확인, 이동식 요소 (예컨대, 트랜스포존, 약물 내성 카세트)의 확인, 질환 관련 유전자에 대한 특이적인 대립유전자 (예컨대, 온코진, 자가면역, 신경퇴행성) 등에 관한 정보를 즉시 획득할 수 있다.
- [0310] 실시예 IX
- [0311] 가이드 RNA/Cas9 복합체를 사용하여 이루어지는 표적 DNA 프로빙
- [0312] 인간 게놈 DNA (노바겐(Novagen))를 0.5 M MES 완충제 (pH 5.5) 중에 약 0.2 ng/ul로 희석시키고, 문헌 [Michalet et al., (Dynamic Molecular Combing, Science 1999)]에 따라 그에 대하여 비닐 실란 코팅된 기판 (게노믹 비전(Genomic Vision: 파리)) 상에 분자적 코밍을 수행하였다. 요약하면, 이는 랭무어 블라젯 (Langmuir Blodgett) 설정 환경과 유사하게, DNA를 함유하는 용액 중에 비닐 실란 코팅된 커버글래스를 침지시킨 후, 그를 고정 속도로 꺼내는 것을 포함한다. 이는 "후진하는 메니스커스" 메커니즘에 의해 기판 상에서 DNA를 신장시키는 작용을 하였다. DNA가 용액 밖으로 완전히 들어 올려졌을 때, 1 cm²당 10,000 마이크로줄로 UV 가교 결합시켰다. 이에 의해 표면 상에 한 방향으로 정렬된 다량의 DNA를 얻었다. DNA 제조시 주의를 기울인다면, DNA 염색, 예컨대 YOYO-1을 사용하여 메가베이스 길이의 DNA를 시각화할 수 있다.
- [0313] 기판 상에서 신장된 게놈 DNA를 완충제로 습윤화시킨 후, (37°C에서 10 min 동안 미리 인큐베이션시켜) 미리 형성된 gRNA/cas9 (NEB) 복합체를 기판에 첨가하고, 1시간 동안 반응이 일어나도록 하였다. 이어서, 잉여량의 복합체는 세척하여 제거하였다. 가이드 RNA는 동원체 서열에 상보적인 부분을 갖도록 디자인하였고, 가이드의 3' 단부는 프로브 서열에 상보적인 테일 핵산 서열을 갖도록 디자인하였다. 가이드 RNA가 임의의 원하는 게놈 서열로 디자인될 수 있다는 것을 통상의 기술자는 쉽게 이해할 수 있을 것이다. 시험관내 전사 시스템에서 가이드 RNA 및 테일 서열, 프로모터 및 종결인자 신호 서열 (IDT)을 비롯한 서열을 사용하여 주형으로부터 프로브 서열에 상보적인 테일 서열을 갖는 가이드 RNA를 합성하였다. 자이모 클린/농축기를 사용하여 프로브 서열에 상보적인 테일 서열을 갖는 가이드 RNA를 정제하였다. 이어서, 신장된 DNA에 첨가하기 이전에 상기 기술된 바와 같이, 전사된 RNA를 Cas9와 함께 인큐베이션시켰다.
- [0314] 이어서, 슬라이드를 BLOCKAID (인비트로젠(Invitrogen))로 처리하였다. 이어서, gRNA 서열 상의 테일에 상보적인 16 nt DNA 프로브를 비-염격한 조건 (4XSSC, 50% 포름아미드, 블록에이드(Blockaid)) 하에서 복합체와 반응시켰다. 사용된 프로브 양단 모두 Atto 657N 염료로 표지되었다 (인비트로젠으로부터 맞춤식 합성 주문). 혼성화 반응물을 4°C에서 밤새도록 정치시켰다. 이어서, 슬라이드를 세척하여 잉여량의 염료를 제거하고, TIRF 현미경에서 영상화하였다.
- [0315] DNA 표적이 이중 가닥 상태 그대로 남아있기 때문에, Atto647N의 영상화 이전 또는 이후에 YOYO-1 염료로 염색 할 수 있었다. 적색 레이저 및 적절한 필터를 사용하여 가이드 RNA 상의 표지를 검출하고, 블루 레이저 및 적절한 필터를 사용하여 DNA 염색을 검출하였다. 광역 TIRF 현미경법에 의해 영상화를 수행하였다.
- [0316] 도 6A는 표면 상에서 연장 또는 미리 신장된 이중 가닥 인간 게놈 DNA (선)에 결합된 동원체 특이적 가이드 RNA (gRNA)/Cas9 복합체 (점)의 결과를 보여주는 영상이다. 도 6B는 gRNA/Cas9 (점)의 인간 동원체 DNA (회색)에의 결합에 대한 초해상도 영상이다.

- [0317] 도 6A에 제시된 바와 같이, 프로브는, 대략 추정컨대, 게놈 중 ~1%를 포함하는 동원체 DNA를 표적화하였다. 많은 시야를 조사하였고, 도 6A에 제시된 바와 같이, 선을 따라 상호관계가 있는 프로브 표지의 점이 있는 시야가 빈번하게 관찰되었다. DNA 염색 결과, 슬라이드는 선을 따라 상호관계가 있는 표지를 보이지 않은 다수의 신장된 DNA 문자를 갖는 것으로 나타났으며, 이에 의해 본원에 기술된 방법은 검출가능한 gRNA/Cas9 복합체를 사용하여 복수 개의 핵산을 포함하는 샘플 내에서 특정의 표적 핵산을 확인할 수 있다는 것이 입증된다.
- [0318] CRISPR/CAS9 완충제는 20 mM HEPES, 100 mM NaCl, 5 mM MgCl₂, 0.1 mM EDTA를 포함하였고, 25°C에서 pH 6.5였다. 혼성화 완충제는 200 uL "순(neat)" 포름아미드, 20 uL 10% SDS, 120 uL 블록에이드, 40 uL의 20xSSC 및 20 uL 물을 포함하였다. 사용된 Cas9는 최종 농도 25-30 nM의 Cas9 뉴클레아제, S. 피오케네스 (NEB)였다. 사용된, 표지된 올리고 (IDT) 농도는 ~ 800 nM이었다.
- [0319] 도 6C는 시험관내 Cas9 결합 검정법에 관한 것이다. 전장의 람다 DNA (~ 48 kb)를 Cas9 단독으로 프로빙하거나, 또는 CRISPR-gRNA와 복합체를 형성하게 하고, 비닐 실란 관능화된 유리 표면 상에서 신장시켰다. 이어서, Cas9를 피코에리트린-접합된 항체로 표지하고 (암으로 표시), DNA는 YOYO-1로 표지하였다 (명으로 표시). 100x/1.47 오일 대물렌즈가 장착된 레이카(Leica) DM1600 상에서 TIRF 모드로 영상을 획득하였다.
- [0320] 실시예 X
- [0321] 도 7은 BRCA1 유전자 유전자좌에 적용된 위치 레드-그린-블루 (RGB) 바코드 시스템에 대한 것이다. BRCA1 및 BRCA2 유전자 내의 큰 재배열은 여성 및 남성, 둘 모두에서 유방암 및 난소암 감수성에 대한 중요한 마커이다. 재배열이 암 발생 위험, 및 적절한 치료에 대한 정보를 제공할 수 있다 [Judkins T, Rosenthal E, Arnell C, Burbidge LA, Geary W, Barrus T, Schoenberger J, Trost J, Wenstrup RJ, Roa BB. Clinical significance of large rearrangements in BRCA1 and BRCA2. *Cancer*. 2012 Nov 1;118(21):5210-6.; Liede A, Karlan BY, Narod SA. Cancer risks for male carriers of germline mutations in BRCA1 or BRCA2: a review of the literature. *J Clin Oncol*. 2004 Feb 15;22(4):735-42]. 상기의 큰 재배열 스펙트럼은 매우 광범위하고, 전형적으로는 상기 재배열의 특징을 규명하는데에는 PCR 증폭된 DNA의 생어(Sanger) 서열분석이 사용된다. 본 발명자들의 gRNA/Cas9 프로빙 전략법은 유전자좌에 따른 특이적인 컬러 패턴 (표 3, 위치 코드)을 사용하여 고밀도로 (100 bp당 대략 1개의 부위가 프로빙됨) 전장의 각 BRCA 유전자 유전자좌를 표지한다. 그의 유전자좌 내에서 재배열된 임의 영역의 경우, 컬러 패턴은 바뀔 수 있고, 이에 의해 상기의 큰 재배열을 확인할 수 있다. 본 발명자들의 접근법은 서열분석에 대한 더욱 간단한 대안을 제공한다. 본 발명자들의 접근법은 또한 다른 암, 예컨대 급성 끌수성 백혈병, 발육 질환, 예컨대 부虺 증후군(Bushy syndrome), 및 신경 발달 질환, 예컨대 자폐증에서의 다른 유사 재배열을 검출하는데에도 사용된다. 본 접근법에 잘 맞는 재배열의 또 다른 카테고리는 면역 세포 수용체의 VDJ 재조합을 포함한다.
- [0322] 실시예 XI
- [0323] 도 8은 gRNA/cas9 복합체에 의해 표적화된 특이적인 위치에 적용된 DNA 오리가미 바코드에 대한 것이다. 각 바코드는 gRNA가 표적화하는 위치에 대하여 독특한 것이다. BRCA1 및 BRCA2 유전자는, 오직 생어 서열분석으로만 검출될 수 있는 것인, 좀 더 작은 삽입 및 결실을 비롯한, 다양한 규모로 쉽게 게놈 재배열이 이루어질 수 있다. 본 발명자들의 gRNA/Cas9 프로빙 전략법은 각 gRNA/Cas9 표적 부위에 독특한 바코드를 사용하여 고밀도로 (100 bp당 대략 1개의 부위가 프로빙됨) 전장의 각 BRCA 유전자 유전자좌를 표지한다. 상기 바코드는 이전에 기술된 바와 같이 형광 오리가미 바코드 형태를 취한다 [Lin C, Jungmann R, Leifer AM, Li C, Levner D, Church GM, Shih WM, Yin P. Submicrometre geometrically encoded fluorescent barcodes self-assembled from DNA. *Nat Chem*. 2012 Oct;4(10):832-9]. 각 바코드 (표 3, 오리가미 코드)를 디코딩함으로써, 본 발명자들은 각 gRNA/Cas9, 및 유전자좌 상의 그의 위치를 확인할 수 있고, 이는 삽입, 결실, 및 그의 유전자좌 내에서 재배열된 영역에 대한 정보를 제공한다. 이러한 분석 수준은 적절한 치료 결정시 의사에 의해 고려될 수 있다. 본 발명자들의 접근법은 서열분석에 대한 대안을 제공한다. 본 발명자들의 접근법은 또한 다른 암, 발육 질환, 및 신경 발달 질환에서의 다른 유사 재배열을 검출하는데에도 사용된다. 본 접근법에 잘 맞는 재배열의 또 다른 카테고리는 면역 세포 수용체의 VDJ 재조합을 포함한다.
- [0324] 실시예 XII
- [0325] 도 9A-9B는 gRNA/Cas9 네으로부터의 폴리미라제 연장에 관한 것이다. 도 9A에는 gRNA에 의해 절단된 표적 DNA의 상부 가닥 및 Cas9의 D10A 돌연변이체가 도시되어 있다. 도 9B에는 gRNA에 의해 절단된 표적 DNA의 하부 가닥 및 Cas9의 H840A 돌연변이체가 도시되어 있다. 곡선형 화살표 표시는 네으로부터의 프라이머 연장 방향을

나타낸다. 표적화된 영역을 표지하는데, 또는 표적화된 영역의 서열분석을 개시하는데 프라이머 연장이 사용될 수 있다. 본 발명자들의 gRNA/Cas9 낙 및 서열분석 접근법을 통해 계놈 상의 서열분석 출발 부위를 정확하게 표적화하여 국제화할 수 있다. 요약하면, gRNA/Cas9 낙카제는 DNA 가닥 중 표적화된 관심 위치에서 낙을 형성한다. 이어서, gRNA/Cas9는 수개의 방법들, 예컨대 계면활성제, 변성제, 또는 온도에 의해 치환된다. 이어서, 가닥 치환 활성, 및 표지된 dNTP를 포함하는 DNA 폴리머라제를 첨가함으로써 낙 부위로부터 DNA를 연장시킨다. 관심 계놈 영역 다수가 본 gRNA/Cas9 낙 개시 방법에 잘 맞는다. 반복부 영역, 예컨대 동원체 DNA는 근본적으로 그의 다수의 짧은 반복부로 인해 서열분석 및 정렬이 어렵다. 계놈 재배열은 대개 더욱 작은 돌연변이를 함유한다. 추가로, 다수의 계놈 재배열에 관한 특징 규명은 불충분하게 이루어져 있으며, 재배열 또는 융합 위치는 알려져 있지 않다.

[0326] 실시예 XIII

도 10은 CHOPCHOP를 사용하여 PAM 부위를 찾을 때 제공되는 결과에 관한 다이어그램이다. CHOPCHOP는 주어진 유전자좌 상의 gRNA/Cas9 표적 및 오프-타겟을 찾고, 점수화하는데 사용되는 다수의 온라인 도구 알고리즘 중 하나이다 [Tessa G. Montague; Jose M. Cruz; James A. Gagnon; George M. Church; Eivind Valen. (2014). CHOPCHOP: a CRISPR/Cas9 and TALEN web tool for genome editing. Nucleic Acids Res. 42. W401-W407]. 본 다이어그램은 Her2 엑손 상의 표적을 검색하는 것에 관한 그래픽 결과를 나타낸 것이다. 작은 삼각형 표시는 Her2 엑손 상의 gRNA/Cas9 표적 부위를 나타낸다. 표적 서열 목록을 추출하고, 큐레이팅하여 오프-타겟이 없거나, 또는 소수의 오프-타겟을 포함하는 표적만을 유지시킨 후, 이를 사용하여 gRNA를 디자인할 수 있다. 상기 도구를 사용하여 다른 유전자좌에 대한 gRNA/Cas9 표적을 찾는다. 인트론 및 엑손, 둘 모두, 또는 유전자좌 바깥쪽의 표적을 찾는데 다른 도구가 사용될 수 있다.

[0328] 실시예 XIV

도 11은 고충실도 폴리머라제 연쇄 반응 (PCR)을 사용하여 이루어지는 가이드 RNA의 시험관내 전사를 위한 DNA 주형의 어셈블리에 관한 것이다. 본 전략법은 2개의 범용 올리고뉴클레오티드: Fwd-T7-gRNA인 시험관내 전사를 위한 T7 RNA 폴리머라제 인식 모티프의 일부도 포함하는 정방향 PCR 프라이머; 및 gRNA.split60 gRNA 스캐폴드에 의존한다. 추가로, 2개의 특이적인 올리고뉴클레오티드: Sp.gRNA.split60인 관심 표적에 특이적인 서열; 및 Rev-B1-gRNA.18인, 다중 가닥 검출을 위한 바코드가 있는 핸들도 또한 포함하는 역방향 PCR 프라이머가 존재한다. 제안된 융점 (T_m) 또한 제공한다. 상기 디자인은 비용면에서 효율적이고, 증폭 오류를 최소화하며, 소규모 또는 대규모 올리고뉴클레오티드에도 잘 맞는다. 일단 증폭되고 나면, DNA 주형은 PCR에 의해 재증폭됨으로써 주형을 생성하고, 영구화할 수 있고, 이는 신생 합성보다 비용면에서 더 효율적이다. PCR 어셈블리는 1시간 미만으로 소요된다. PCR 후, gRNA는 T7 RNA 폴리머라제 믹스를 주형 DNA에 첨가함으로써 시험관내 전사 (IVT)에 의해 합성된다.

[0330] 실시예 XV

도 12는 ALK 전위의 Cas9-gRNA 표적화를 나타낸 개략도이다. 본 도면은 융합된 두 영역이 공지되어 있을 때, 융합물의 존재는 특이적인 바코드를 보유하는 gRNA/Cas9를 사용함으로써 검출될 수 있다는 것을 보여주는 것이다. ALK 유전자좌에서는 염색체내 및 염색체간 재배열 및 역위, 둘 모두가 쉽게 잘 일어날 수 있다. 7개의 상기 재배열은 유해한 임상 결과와 함께 공지되어 있다 [Solomon B, Varella-Garcia M, Camidge DR. ALK gene rearrangements: a new therapeutic target in a molecularly defined subset of non-small cell lung cancer. J Thorac Oncol. 2009 Dec;4(12):1450-4]. 본 발명자들의 gRNA/Cas9 프로빙 전략법은 각 유전자좌와 연관된 컬러를 사용하여 고밀도로 (100 bp당 대략 1개의 부위가 프로빙됨) 특이적인 유전자좌 영역을 표지한다. 상기 전략법은 유전자 융합 및 역위를 검출하는데 유용하다. 두 표지된 유전자좌가 인접해 있을 때, 그의 각 표지는 공동-국제화된 신호로서 검출된다. 재배열된 경우, 두 신호는 더 이상 공동-국제화되지 않고, 상당히 멀리 떨어져 있다. 역위된 경우, 신호는 더 이상 공동-국제화되지 않지만, 여전히 같은 인근 위치에 존재한다. 신뢰도를 부가하기 위해, 제3 유전자좌를 표지하는데, 이는 (ALK 염색체간 재배열에 기인한) 유전자 융합이 있는 경우, 또는 ALK 역위가 있는 경우, 상이한 공동-국제화 조합을 제공한다. 수개의 암, 예컨대 역형성 대세포 텁프종 (ALK-NPM1 융합), 폐 샘암종 (ALK-EML4 융합 및 역위), 및 특정 소아 신경모세포종을 확인하는데 있어 상기 ALK 결합을 검출하는 것이 중요하다. 특정 ALK 재배열에 대해 치료가 이루어진다. 수개의 다른 질환이 유전자 융합을 특징으로 하며, 이는 본 발명자들의 접근법에 잘 맞는다. 유전자 융합 표적의 일부 예로서 ABL1-BCR, AML1-RUNX1T1, AML1-ETV6, BCL-2-IGH, BCL-2-MLT, C-Myc-IGH, COL1A1-PDGFB, CycD1-IGH, ETV6-TRKC, ETV6-JAK, FLI1-EWS, PAX8-PPARG, PML-NR1B1, TCR-RBTN2, SS18-SSX를 포함한다.

도면

도면1

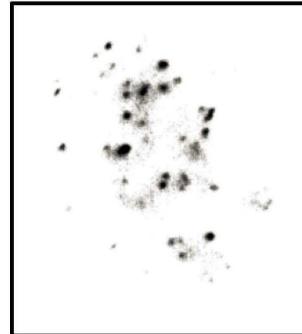
도 1A



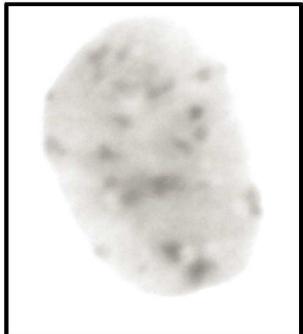
도 1B



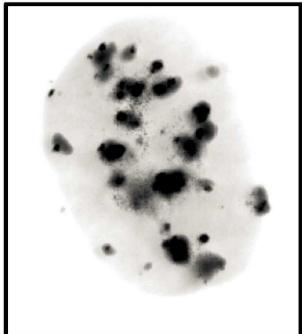
도 1C



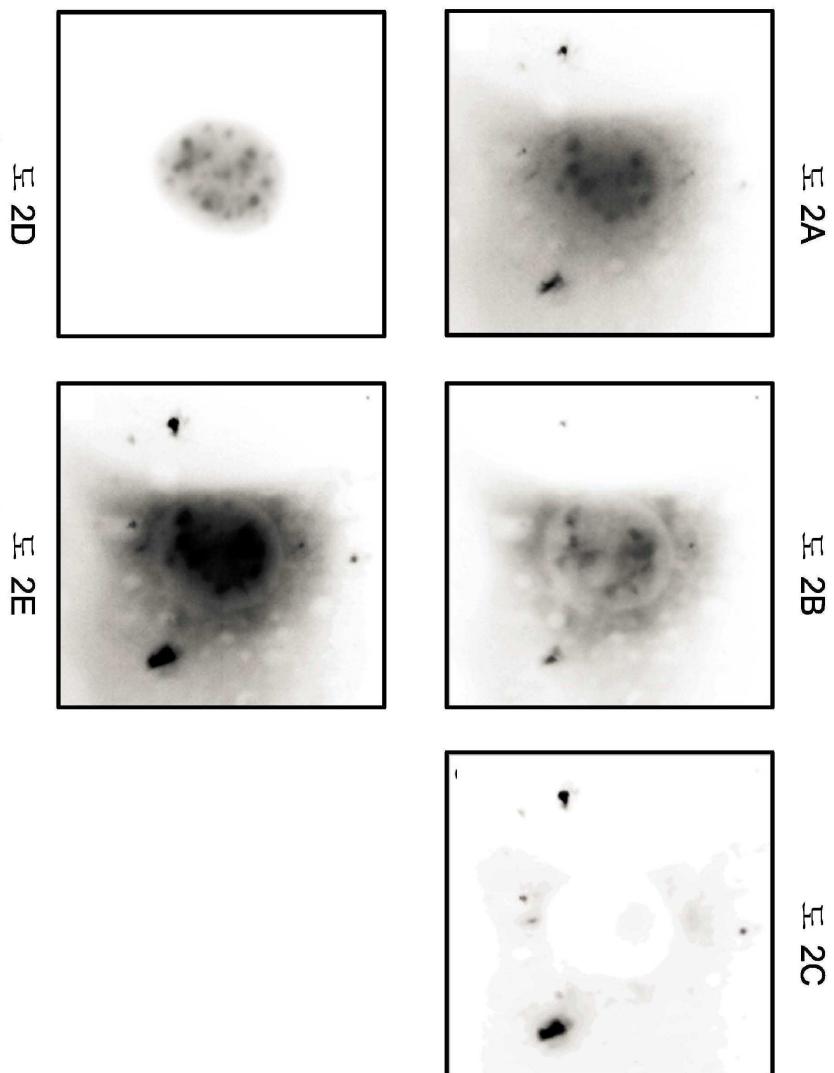
도 1D



도 1E



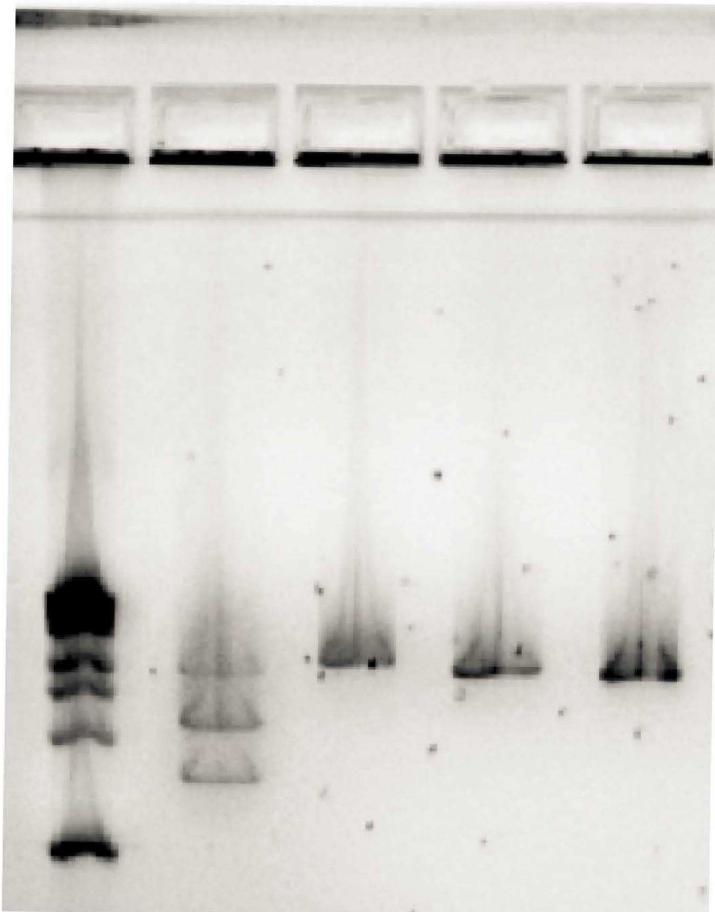
도면2



도면3

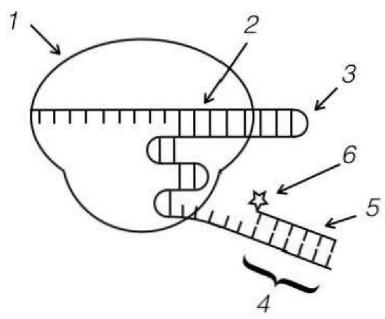
아가로스 2%

M 1 2 3 4

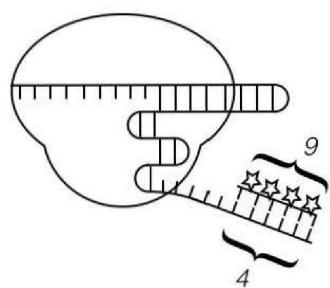
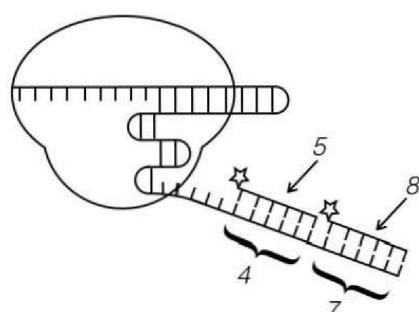


도면4*i*

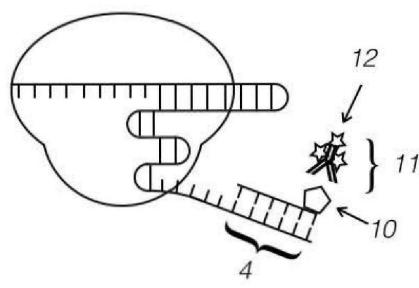
도 4A



도 4B



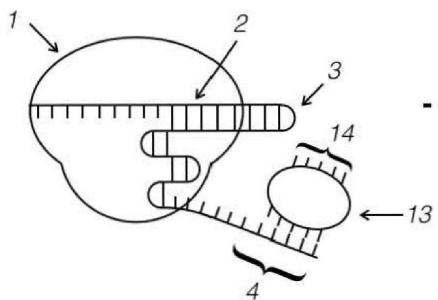
도 4C



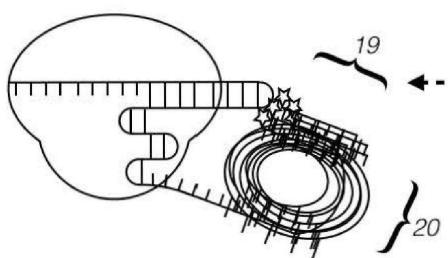
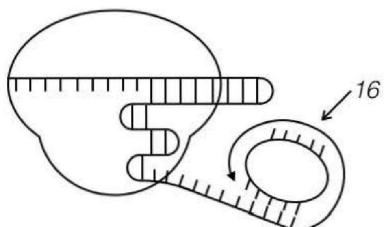
도 4D

도면4ii

도 4E



도 4F

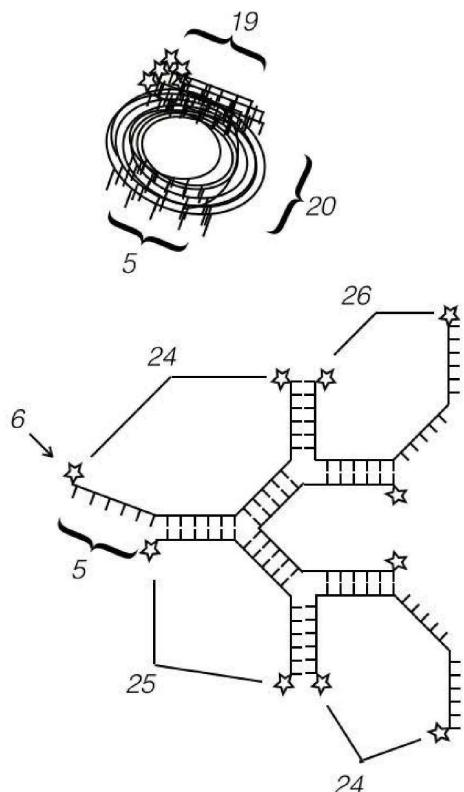


도 4G

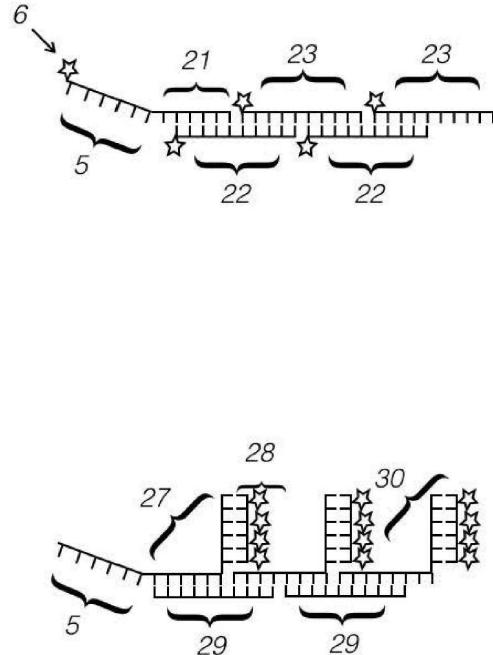
도 4H

도면4iii

도 4I



도 4J

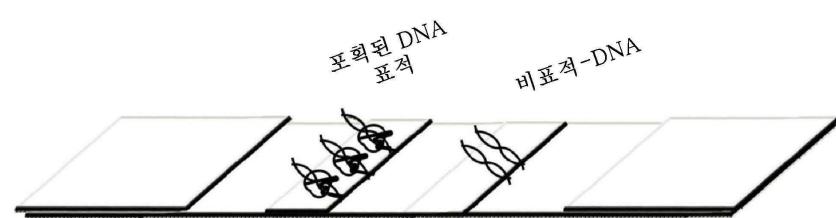
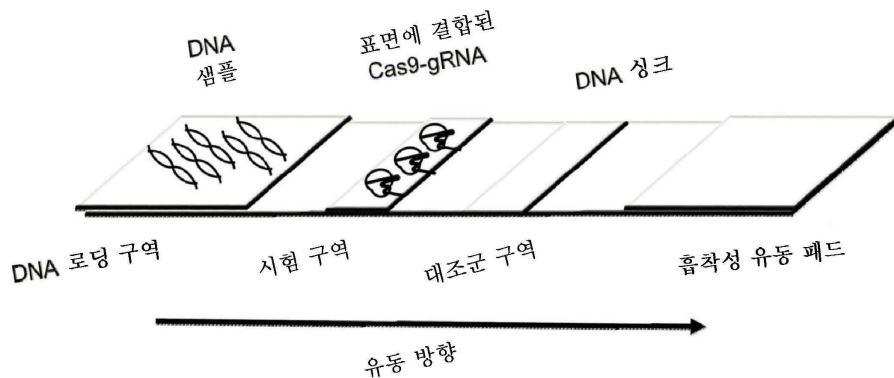


도 4K

도 4L

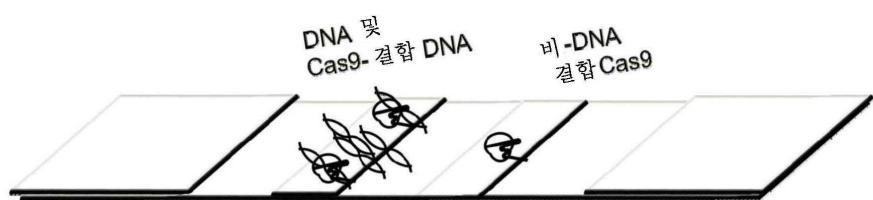
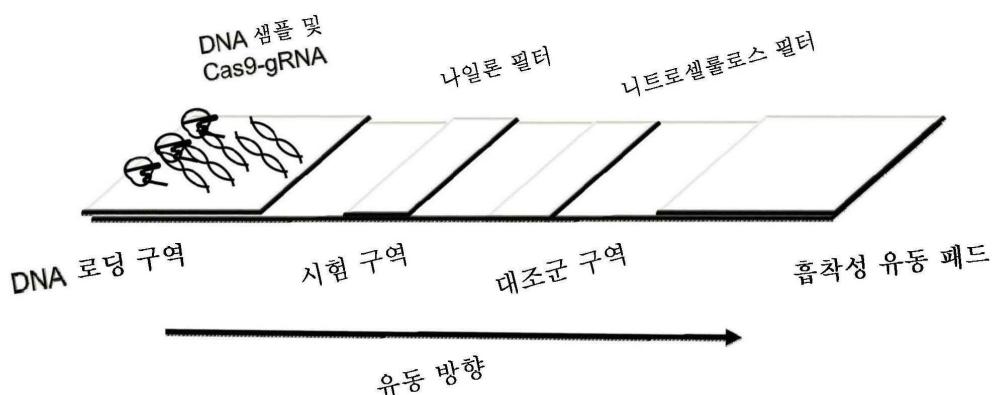
도면5i

도 5A



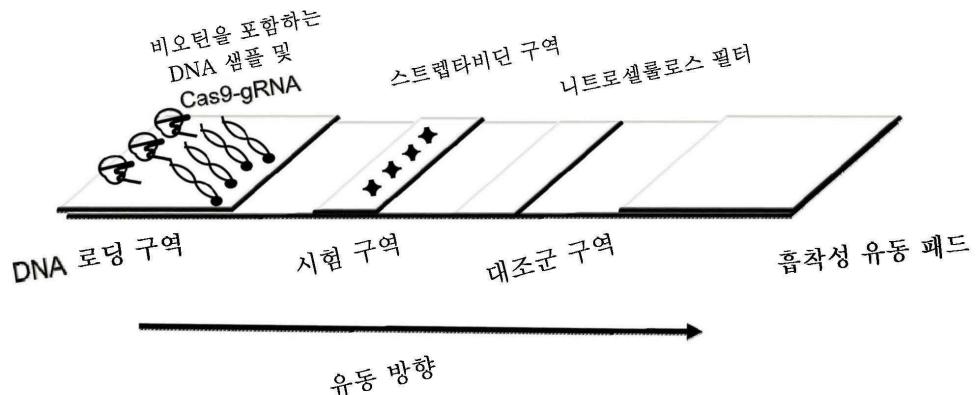
도면5ii

도 5B



도면5iii

도 5C

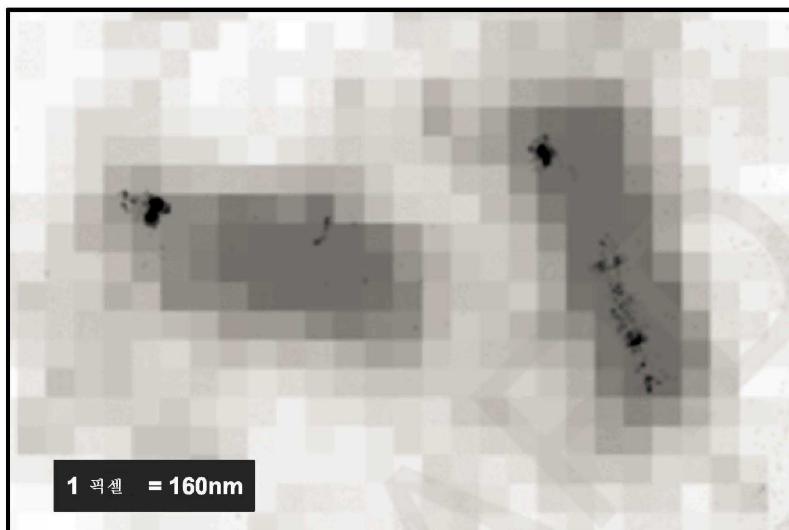


도면6

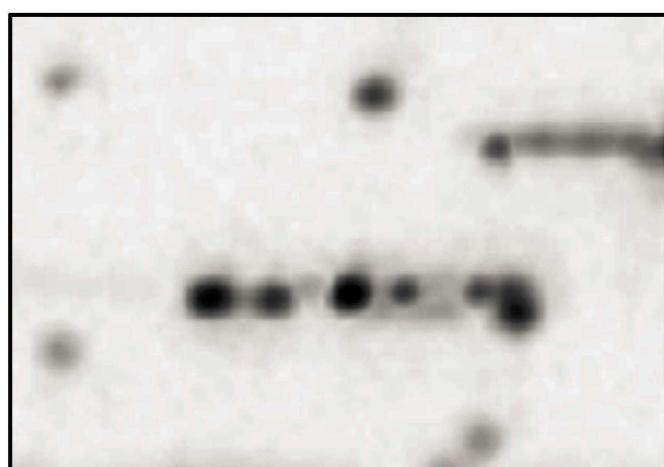
도 6A



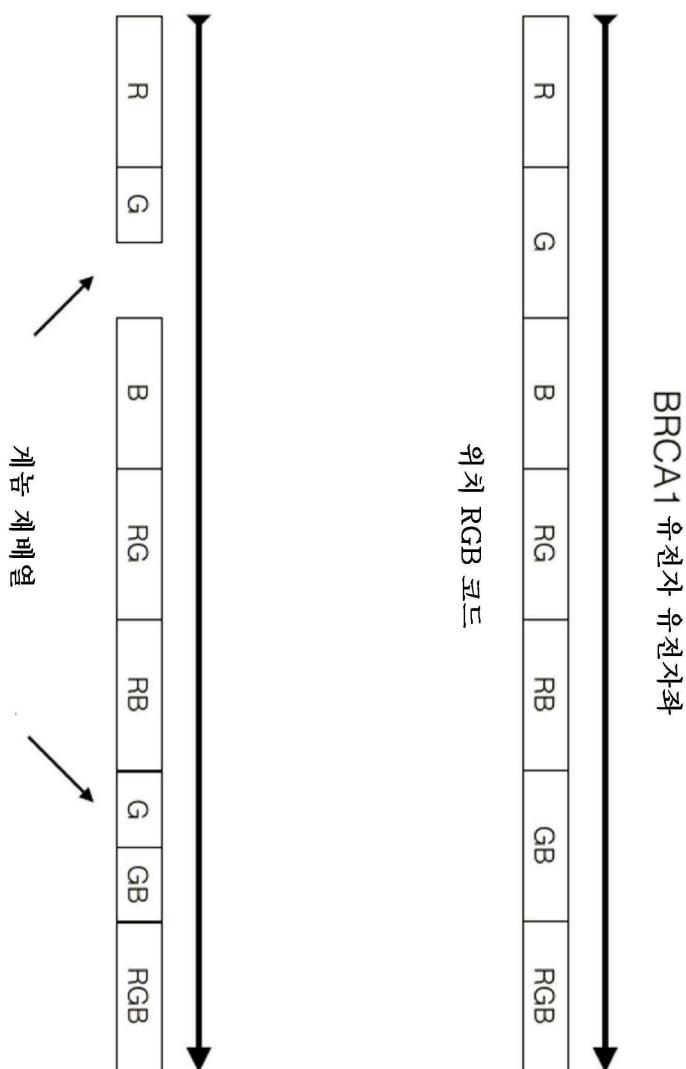
도 6B



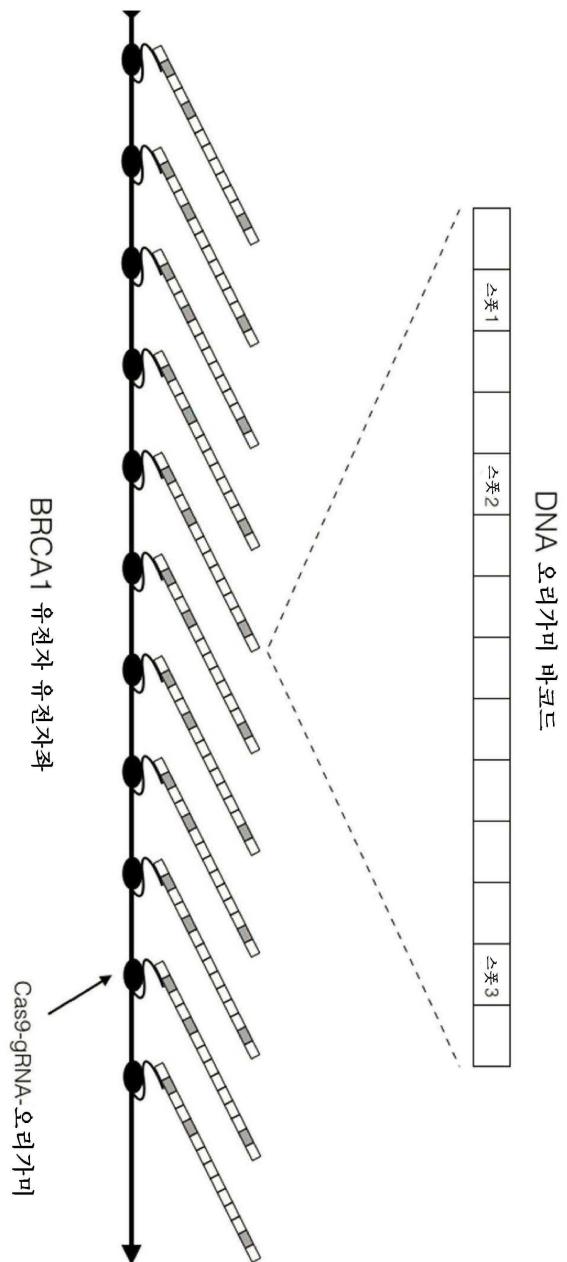
도 6C



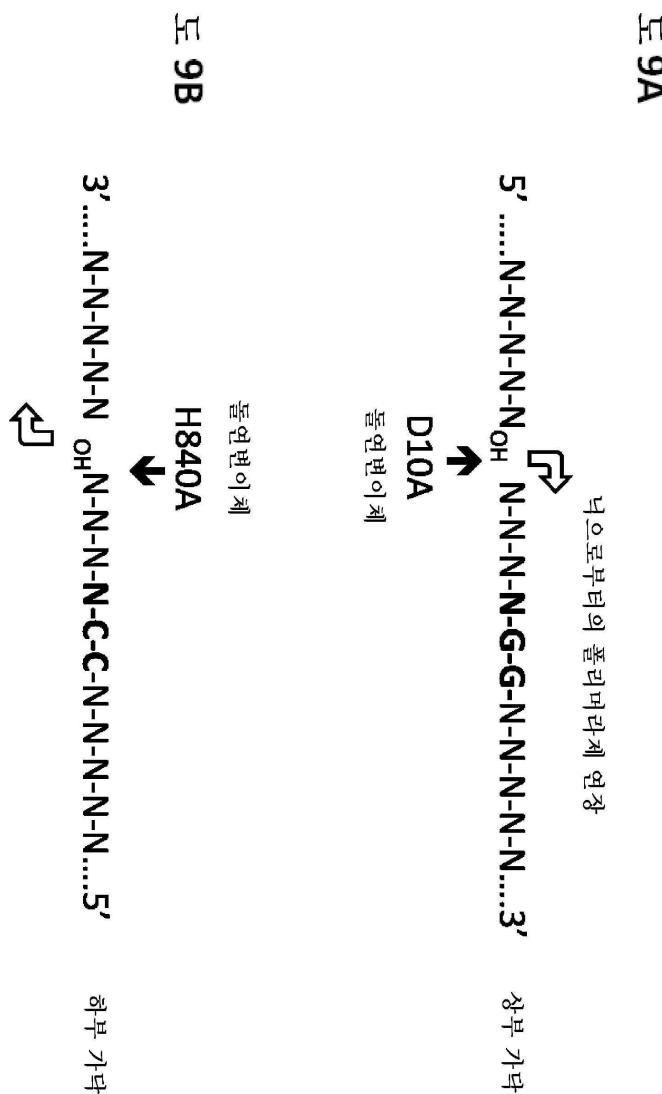
도면7



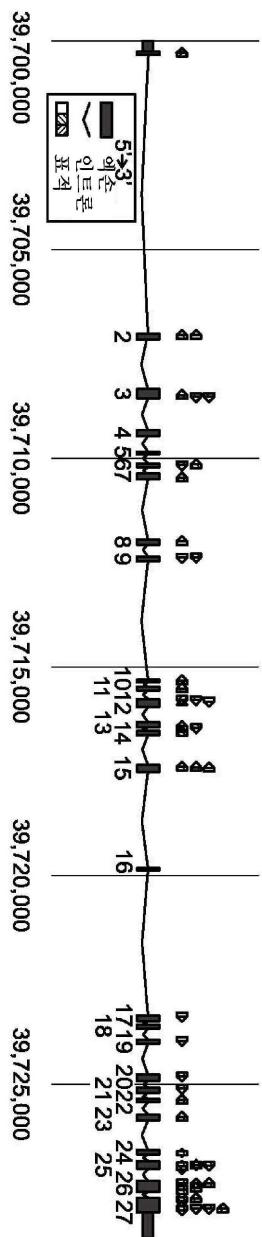
도면8



도면9



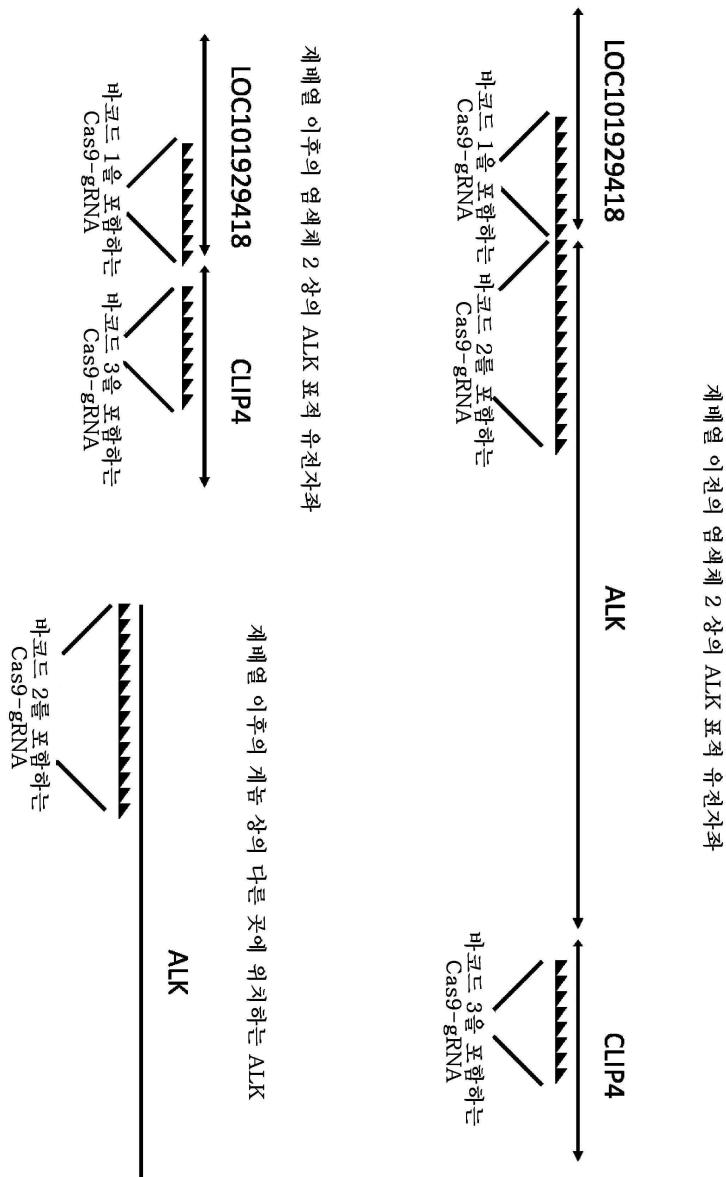
도면10



도면11

Fwd-t7-**gRNA** GAATTATACACTCACTAAG (Tm = 52°C)
Sp.**gRNA**.split60 TTAATCGACTCACTATGGNNNNNNNNNNNNNTTTAGACTAGAAATAGAA (Tm = 55°C)
gRNA.split60
Rev=B1-**gRNA**.18 GTTTAGACTAGAAATAGCAGTTAAATAGGCTAGTCGTTACCTGAAAAGTGCCACCAAGGTCTTTT (Tm = 55°C)
ACGGGGGGGGCTTTTTATACGCA (Tm = 62°C)

도면12



서 열 목록

SEQUENCE LISTING

<110> President and Fellows of Harvard College

<120> RNA-GUIDED SYSTEMS FOR PROBING AND MAPPING OF NUCLEIC ACIDS

<130> 010498_00633

<140> PCT/US2015/045805

<141> 2015-08-19

<150> US 62/039, 341

<151> 2014-08-19
<160> 146
<170> PatentIn version 3.5
<210> 1
<211> 16
<212> DNA
<213> Artificial
<220><223> Figure 9A Polymerase Extension from nick site
<220><221> misc_feature
<222> (1)..(9)
<223> n is a, c, g, or t

<220><221> misc_feature

<222> (12)..(16)
<223> n is a, c, g, or t
<400> 1

nnnnnnnnnng gnnnnn 16

<210> 2
<211> 16
<212> DNA
<213> Artificial
<220><223> Figure 9B Polymerase Extension from nick site

<220><221> misc_feature
<222> (1)..(9)

<223> n is a, c, g, or t

<220><221> misc_feature

<222> (12)..(16)

<223> n is a, c, g, or t

<400> 2

nnnnnnnnnnnc cnnnnn 16

<210> 3
<211> 24
<212> DNA
<213> Artificial

<220><223> Fwd-t7-gRNA

<400> 3

gaaatataa cgaactacta tagg 24

<210> 4

<211> 60

<212> DNA

<213> Artificial

<220><223> Figure 11 Sp.gRNA.split60

<220><221> misc_feature

<222> (21)..(38)

<223> n is a, c, g, or t

<400> 4

ttaatacgac tcactatagg nnnnnnnnnn nnnnnnnngt tttagagcta gaaatagcaa 60

<210> 5

<211> 81

<212> DNA

<213> Artificial

<220><223> gRNA.split60

<

400> 5

gttttagagc tagaaatagc aagttaaat aaggctagtc cgtttatcaac ttgaaaaagt 60

ggcacccgagt cggtgcttt t 81

<210> 6

<211> 28

<212> DNA

<213> Artificial

<220><223> Figure 11 Rev_B1_gRNA.18

<400> 6

accgagtcgg tgctttta tacatcta 28

<210> 7

<211> 1368

<212> PRT

<213> Streptococcus pyogenes

<400> 7

Met Asp Lys Lys Tyr Ser Ile Gly Leu Asp Ile Gly Thr Asn Ser Val

1	5	10	15
Gly Trp Ala Val Ile Thr Asp Glu Tyr Lys Val Pro Ser Lys Lys Phe			
20	25	30	
Lys Val Leu Gly Asn Thr Asp Arg His Ser Ile Lys Lys Asn Leu Ile			
35	40	45	
Gly Ala Leu Leu Phe Asp Ser Gly Glu Thr Ala Glu Ala Thr Arg Leu			
50	55	60	
Lys Arg Thr Ala Arg Arg Tyr Thr Arg Arg Lys Asn Arg Ile Cys			
65	70	75	80
Tyr Leu Gln Glu Ile Phe Ser Asn Glu Met Ala Lys Val Asp Asp Ser			
85	90	95	
Phe Phe His Arg Leu Glu Glu Ser Phe Leu Val Glu Glu Asp Lys Lys			
100	105	110	
His Glu Arg His Pro Ile Phe Gly Asn Ile Val Asp Glu Val Ala Tyr			
115	120	125	
His Glu Lys Tyr Pro Thr Ile Tyr His Leu Arg Lys Lys Leu Val Asp			
130	135	140	
Ser Thr Asp Lys Ala Asp Leu Arg Leu Ile Tyr Leu Ala Leu Ala His			
145	150	155	160
Met Ile Lys Phe Arg Gly His Phe Leu Ile Glu Gly Asp Leu Asn Pro			
165	170	175	
Asp Asn Ser Asp Val Asp Lys Leu Phe Ile Gln Leu Val Gln Thr Tyr			
180	185	190	
Asn Gln Leu Phe Glu Glu Asn Pro Ile Asn Ala Ser Gly Val Asp Ala			
195	200	205	
Lys Ala Ile Leu Ser Ala Arg Leu Ser Lys Ser Arg Arg Leu Glu Asn			
210	215	220	
Leu Ile Ala Gln Leu Pro Gly Glu Lys Lys Asn Gly Leu Phe Gly Asn			
225	230	235	240
Leu Ile Ala Leu Ser Leu Gly Leu Thr Pro Asn Phe Lys Ser Asn Phe			
245	250	255	

Asp Leu Ala Glu Asp Ala Lys Leu Gln Leu Ser Lys Asp Thr Tyr Asp

260 265 270

Asp Asp Leu Asp Asn Leu Leu Ala Gln Ile Gly Asp Gln Tyr Ala Asp

275 280 285

Leu Phe Leu Ala Ala Lys Asn Leu Ser Asp Ala Ile Leu Leu Ser Asp

290 295 300

Ile Leu Arg Val Asn Thr Glu Ile Thr Lys Ala Pro Leu Ser Ala Ser

305 310 315 320

Met Ile Lys Arg Tyr Asp Glu His His Gln Asp Leu Thr Leu Leu Lys

325 330 335

Ala Leu Val Arg Gln Gln Leu Pro Glu Lys Tyr Lys Glu Ile Phe Phe

340 345 350

Asp Gln Ser Lys Asn Gly Tyr Ala Gly Tyr Ile Asp Gly Gly Ala Ser

355 360 365

Gln Glu Glu Phe Tyr Lys Phe Ile Lys Pro Ile Leu Glu Lys Met Asp

370 375 380

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

385 390 395 400

Lys Gln Arg Thr Phe Asp Asn Gly Ser Ile Pro His Gln Ile His Leu

405 410 415

Gly Glu Leu His Ala Ile Leu Arg Arg Gln Glu Asp Phe Tyr Pro Phe

420 425 430

Leu Lys Asp Asn Arg Glu Lys Ile Glu Lys Ile Leu Thr Phe Arg Ile

435 440 445

Pro Tyr Tyr Val Gly Pro Leu Ala Arg Gly Asn Ser Arg Phe Ala Trp

450 455 460

Met Thr Arg Lys Ser Glu Glu Thr Ile Thr Pro Trp Asn Phe Glu Glu

465 470 475 480

Val Val Asp Lys Gly Ala Ser Ala Gln Ser Phe Ile Glu Arg Met Thr

485 490 495

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

500	505	510
Leu Leu Tyr Glu Tyr Phe Thr Val Tyr Asn Glu Leu Thr Lys Val Lys		
515	520	525
Tyr Val Thr Glu Gly Met Arg Lys Pro Ala Phe Leu Ser Gly Glu Gln		
530	535	540
Lys Lys Ala Ile Val Asp Leu Leu Phe Lys Thr Asn Arg Lys Val Thr		
545	550	555
Val Lys Gln Leu Lys Glu Asp Tyr Phe Lys Lys Ile Glu Cys Phe Asp		
565	570	575
Ser Val Glu Ile Ser Gly Val Glu Asp Arg Phe Asn Ala Ser Leu Gly		
580	585	590
Thr Tyr His Asp Leu Leu Lys Ile Ile Lys Asp Lys Asp Phe Leu Asp		
595	600	605
Asn Glu Glu Asn Glu Asp Ile Leu Glu Asp Ile Val Leu Thr Leu Thr		
610	615	620
Leu Phe Glu Asp Arg Glu Met Ile Glu Glu Arg Leu Lys Thr Tyr Ala		
625	630	635
His Leu Phe Asp Asp Lys Val Met Lys Gln Leu Lys Arg Arg Arg Tyr		
645	650	655
Thr Gly Trp Gly Arg Leu Ser Arg Lys Leu Ile Asn Gln Ile Arg Asp		
660	665	670
Lys Gln Ser Gly Lys Thr Ile Leu Asp Phe Leu Lys Ser Asp Gly Phe		
675	680	685
Ala Asn Arg Asn Phe Met Gln Leu Ile His Asp Asp Ser Leu Thr Phe		
690	695	700
Lys Glu Asp Ile Gln Lys Ala Gln Val Ser Gly Gln Gly Asp Ser Leu		
705	710	715
His Glu His Ile Ala Asn Leu Ala Gly Ser Pro Ala Ile Lys Lys Gly		
725	730	735
Ile Leu Gln Thr Val Lys Val Val Asp Glu Leu Val Lys Val Met Gly		
740	745	750

Arg His Lys Pro Glu Asn Ile Val Ile Glu Met Ala Arg Glu Asn Gln
 755 760 765
 Thr Thr Gln Lys Gly Gln Lys Asn Ser Arg Glu Arg Met Lys Arg Ile
 770 775 780

Glu Glu Gly Ile Lys Glu Leu Gly Ser Gln Ile Leu Lys Glu His Pro
 785 790 795 800
 Val Glu Asn Thr Gln Leu Gln Asn Glu Lys Leu Tyr Leu Tyr Tyr Leu
 805 810 815
 Gln Asn Gly Arg Asp Met Tyr Val Asp Gln Glu Leu Asp Ile Asn Arg
 820 825 830
 Leu Ser Asp Tyr Asp Val Asp His Ile Val Pro Gln Ser Phe Leu Lys
 835 840 845

Asp Asp Ser Ile Asp Asn Lys Val Leu Thr Arg Ser Asp Lys Asn Arg
 850 855 860
 Gly Lys Ser Asp Asn Val Pro Ser Glu Glu Val Val Lys Lys Met Lys
 865 870 875 880
 Asn Tyr Trp Arg Gln Leu Leu Asn Ala Lys Leu Ile Thr Gln Arg Lys
 885 890 895
 Phe Asp Asn Leu Thr Lys Ala Glu Arg Gly Gly Leu Ser Glu Leu Asp
 900 905 910

Lys Ala Gly Phe Ile Lys Arg Gln Leu Val Glu Thr Arg Gln Ile Thr
 915 920 925
 Lys His Val Ala Gln Ile Leu Asp Ser Arg Met Asn Thr Lys Tyr Asp
 930 935 940
 Glu Asn Asp Lys Leu Ile Arg Glu Val Lys Val Ile Thr Leu Lys Ser
 945 950 955 960
 Lys Leu Val Ser Asp Phe Arg Lys Asp Phe Gln Phe Tyr Lys Val Arg
 965 970 975

Glu Ile Asn Asn Tyr His His Ala His Asp Ala Tyr Leu Asn Ala Val
 980 985 990
 Val Gly Thr Ala Leu Ile Lys Lys Tyr Pro Lys Leu Glu Ser Glu Phe

995	1000	1005
Val Tyr Gly Asp Tyr Lys Val	Tyr Asp Val Arg Lys	Met Ile Ala
1010	1015	1020
Lys Ser Glu Gln Glu Ile Gly	Lys Ala Thr Ala Lys	Tyr Phe Phe
1025	1030	1035
Tyr Ser Asn Ile Met Asn Phe	Phe Lys Thr Glu Ile	Thr Leu Ala
1040	1045	1050
Asn Gly Glu Ile Arg Lys Arg	Pro Leu Ile Glu Thr	Asn Gly Glu
1055	1060	1065
Thr Gly Glu Ile Val Trp Asp	Lys Gly Arg Asp Phe	Ala Thr Val
1070	1075	1080
Arg Lys Val Leu Ser Met Pro	Gln Val Asn Ile Val	Lys Lys Thr
1085	1090	1095
Glu Val Gln Thr Gly Gly Phe	Ser Lys Glu Ser Ile	Leu Pro Lys
1100	1105	1110
Arg Asn Ser Asp Lys Leu Ile	Ala Arg Lys Lys Asp	Trp Asp Pro
1115	1120	1125
Lys Lys Tyr Gly Gly Phe Asp	Ser Pro Thr Val Ala	Tyr Ser Val
1130	1135	1140
Leu Val Val Ala Lys Val Glu	Lys Gly Lys Ser Lys	Lys Leu Lys
1145	1150	1155
Ser Val Lys Glu Leu Leu Gly	Ile Thr Ile Met Glu	Arg Ser Ser
1160	1165	1170
Phe Glu Lys Asn Pro Ile Asp	Phe Leu Glu Ala Lys	Gly Tyr Lys
1175	1180	1185
Glu Val Lys Lys Asp Leu Ile	Ile Lys Leu Pro Lys	Tyr Ser Leu
1190	1195	1200
Phe Glu Leu Glu Asn Gly Arg	Lys Arg Met Leu Ala	Ser Ala Gly
1205	1210	1215
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

1235 1240 1245

Pro Glu Asp Asn Glu Gln Lys Gln Leu Phe Val Glu Gln His Lys

1250 1255 1260

His Tyr Leu Asp Glu Ile Ile Glu Gln Ile Ser Glu Phe Ser Lys

1265 1270 1275

Arg Val Ile Leu Ala Asp Ala Asn Leu Asp Lys Val Leu Ser Ala

1280 1285 1290

Tyr Asn Lys His Arg Asp Lys Pro Ile Arg Glu Gln Ala Glu Asn

1295 1300 1305

Ile Ile His Leu Phe Thr Leu Thr Asn Leu Gly Ala Pro Ala Ala

1310 1315 1320

Phe Lys Tyr Phe Asp Thr Thr Ile Asp Arg Lys Arg Tyr Thr Ser

1325 1330 1335

Thr Lys Glu Val Leu Asp Ala Thr Leu Ile His Gln Ser Ile Thr

1340 1345 1350

Gly Leu Tyr Glu Thr Arg Ile Asp Leu Ser Gln Leu Gly Gly Asp

1355 1360 1365

<210> 8

<211> 130

<212> DNA

<213> Artificial

<220><223> Guide RNA used for Figure 6a and 6b_Centromere 16 gRNA with tail

<400> 8

gacgccuucg uuggaaacgu uuuagagcua gaaaugcaa guaaaaauaa ggcuaguccg 60

uuaucaacuu gaaaaagugg caccgagucg gugcuuuuut cctctaccac ctacatcact 120

tatacatcta 130

<210> 9

<211> 31

<212> DNA

<213> Artificial

<220><223> DNA probe to tail

<400> 9	
tagatgtata agtgatgtag gtggtagagg a	31
<210> 10	
<211> 27	
<212> DNA	
<213> Artificial	
<220><223> DNA Paint imager probe	
<400> 10	
tagatgtata aaaaaattta ataaggt	27
<210> 11	
<211> 16	
<212> DNA	
<213> Artificial	
<220><223> DNA PAINT imager probe	
<400> 11	
tagatgtata aaaaaaa	16
<210> 12	
<211> 22	
<212> DNA	
<213> Artificial	
<220><223> PAM motif	
<220><221> misc_feature	
<222> (3)..(20)	
<223> n is a, c, g, or t	
<400> 12	
ggnnnnnnnn nnnnnnnnnn gg	22
<210> 13	
<211> 23	
<212> DNA	
<213> Artificial	
<220><223> Target sequence for Her2	
<400> 13	
ggcgaggag gagccccag cg	23

<210> 14
<211> 23
<212> DNA
<213> Artificial
<220><223> Her2 target sequence
<400> 14
ggtgtccagg catgtccagg tgg 23
<210> 15
<211> 23
<212> DNA
<213> Artificial
<220><223> Her2 target sequence
<400> 15
ggtgggtctc gggactggca ggg 23
<210> 16
<211> 23
<212> DNA
<213> Artificial
<220><223> Her2 target sequence
<400> 16
ggcagccctg gttagaggtgg cg 23

<210> 17
<211> 23
<212> DNA
<213> Artificial
<220><223> Her2 target sequence
<400> 17
ggaggcccct gtgacagggg tgg 23
<210> 18
<211> 23
<212> DNA
<213> Artificial
<220><223> Her2 target sequence
<400> 18

gggcctcccc aggaggcctg cg	23
<210> 19	
<211> 23	
<212> DNA	
<213> Artificial	
<220><223> Her2 target sequence	
<400> 19	
ggcctccca ggaggcctgc ggg	23
<210> 20	
<211> 23	
<212> DNA	
<213> Artificial	
<220><223> Her2 target sequence	
<400> 20	
ggtggtgtg cccgctgcaa ggg	23
<210> 21	
<211> 23	
<212> DNA	
<213> Artificial	
<220><223> Her2 target sequence	
<400> 21	
ggcagtggc ccctgcagc ggg	23
<210> 22	
<211> 23	
<212> DNA	
<213> Artificial	
<220><223> Her2 target sequence	
<400> 22	
ggcacaggca gtcacacagc tgg	23
<210> 23	
<211> 23	
<212> DNA	
<213> Artificial	

<220><223> Her2 target sequence
<400> 23
ggtttgtcag gggcagacg agg 23
<210> 24
<211> 23
<212> DNA
<213> Artificial
<220><223> Her2 target sequence
<400> 24
ggcatggag cacttgcgag agg 23
<210> 25
<211> 23
<212> DNA
<213> Artificial
<220><223> Her2 target sequence
<400> 25
ggagcacttg cgagaggta ggg 23
<210> 26
<211> 23
<212> DNA
<213> Artificial
<220><223> Her2 target sequence
<400> 26
ggagctgctc tggctggagc ggg 23
<210> 27
<211> 23
<212> DNA
<213> Artificial
<220><223> Her2 target sequence
<400> 27
ggaagacgct gaggtcaggc agg 23
<210> 28
<211> 23
<212> DNA

<213> Artificial	
<220><223> Her2 target sequence	
<400> 28	
ggtgggttt atggatggatg agg	23
<210> 29	
<211> 23	
<212> DNA	
<213> Artificial	
<220><223> Her2 target sequence	
<400> 29	
ggctggggct gcgcgtcactg agg	23
<210> 30	
<211> 23	
<212> DNA	
<213> Artificial	
<220><223> Her2 target sequence	
<400> 30	
ggtgcggtt ccgaaagagc tgg	23
<210> 31	
<211> 23	
<212> DNA	
<213> Artificial	
<220><223> Her2 target sequence	
<400> 31	
ggctggcat cagctggctg ggg	23
<210> 32	
<211> 23	
<212> DNA	
<213> Artificial	
<220><223> Her2 target sequence	
<400> 32	
ggctggca tcagctggct ggg	23
<210> 33	

<211>	23	
<212>	DNA	
<213>	Artificial	
<220><223> Her2 target sequence		
<400>	33	
ggaggaatgc cgagtaactgc agg		23
<210>	34	
<211>	23	
<212>	DNA	
<213>	Artificial	
<220><223> Her2 target sequence		
<400>	34	
ggcatttcctc cacgcactcc tgg		23
<210>	35	
<211>	23	
<212>	DNA	
<213>	Artificial	
<220><223> Her2 target sequence		
<400>	35	
ggctgacact cagggtggca cgg		23
<210>	36	
<211>	23	
<212>	DNA	
<213>	Artificial	
<220><223> Her2 target sequence		
<400>	36	
ggtcaggttt cacaccgctg ggg		23
<210>	37	
<211>	23	
<212>	DNA	
<213>	Artificial	
<220><223> Her2 target sequence		
<400>	37	

gggcagcggg ccacgcagaa ggg

23

<210> 38

<211> 23

<212> DNA

<213> Artificial

<220><223> Her2 target sequence

<400> 38

ggggcagcgg gccacgcaga agg

23

<210> 39

<211> 23

<212> DNA

<213> Artificial

<220><223> Her2 target sequence

<400> 39

ggttggcatt ctgctggtcg tgg

23

<210> 40

<211> 23

<212> DNA

<213> Artificial

<220><223> Her2 target sequence

<400> 40

ggagaatgtg aaaattccag tgg

23

<210> 41

<211> 23

<212> DNA

<213> Artificial

<220><223> Her2 target sequence

<400> 41

ggccatctgc ctgacatcca cg

23

<210> 42

<211> 23

<212> DNA

<213> Artificial

<220><223> Her2 target sequence
<400> 42
ggatgtgcgg ctcgtacaca ggg 23

<210> 43
<211> 23
<212> DNA
<213> Artificial
<220><223> Her2 target sequence
<400> 43
ggtgaaccgc cggcggagaa tgg 23

<210> 44
<211> 23
<212> DNA
<213> Artificial
<220><223> Her2 target sequence
<400> 44
ggtcagggat ctcccggct ggg 23

<210> 45
<211> 23
<212> DNA
<213> Artificial
<220><223> Her2 target sequence
<400> 45
ggatgaccac aaagcgctgg ggg 23

<210> 46
<211> 23
<212> DNA
<213> Artificial
<220><223> Her2 target sequence
<400> 46
ggatgattga ctctgaatgt cg 23

<210> 47
<211> 23

<212> DNA	
<213> Artificial	
<220><223> Her2 target sequence	
<400> 47	
ggtgtctgaa ttctccgca tgg	23
<210> 48	
<211> 23	
<212> DNA	
<213> Artificial	
<220><223> Her2 target sequence	
<400> 48	
ggacagaaga agccctgctg ggg	23
<210> 49	
<211> 23	
<212> DNA	
<213> Artificial	
<220><223> Her2 target sequence	
<400> 49	
gggggacctg gtggatgctg agg	23
<210> 50	
<211> 23	
<212> DNA	
<213> Artificial	
<220><223> Her2 target sequence	
<400> 50	
ggacgatgac atggggacc tgg	23
<210> 51	
<211> 23	
<212> DNA	
<213> Artificial	
<220><223> Her2 target sequence	
<400> 51	
ggtggatgct gaggagtatc tgg	23
<210> 52	

<211> 23	
<212> DNA	
<213> Artificial	
<220><223> Her2 target sequence	
<400> 52	
ggcaccccgag ctcatctacc agg	23
<210> 53	
<211> 23	
<212> DNA	
<213> Artificial	
<220><223> Her2 target sequence	
<400> 53	
ggagtatctg gtaccccgagc agg	23
<210> 54	
<211> 23	
<212> DNA	
<213> Artificial	
<220><223> Her2 target sequence	
<400> 54	
ggaccatgcc cccagcgccc ggg	23
<210> 55	
<211> 23	
<212> DNA	
<213> Artificial	
<220><223> Her2 target sequence	
<400> 55	
gggtgccagt ggagacctgg ggg	23
<210> 56	
<211> 23	
<212> DNA	
<213> Artificial	
<220><223> Her2 target sequence	
<400> 56	

ggcggtgggg acctgacact agg	23
<210> 57	
<211> 23	
<212> DNA	
<213> Artificial	
<220><223> Her2 target sequence	
<400> 57	
ggggaggc ttgcagccct tgg	23
<210> 58	
<211> 23	
<212> DNA	
<213> Artificial	
<220><223> Her2 target sequence	
<400> 58	
ggtccctggc ccagtaatag agg	23
<210> 59	
<211> 23	
<212> DNA	
<213> Artificial	
<220><223> Her2 target sequence	
<400> 59	
gggaccagga cccaccagag cg	23
<210> 60	
<211> 23	
<212> DNA	
<213> Artificial	
<220><223> Her2 target sequence	
<400> 60	
ggtgtccctt tgaaggtgct ggg	23
<210> 61	
<211> 23	
<212> DNA	
<213> Artificial	
<220><223> Her2 target sequence	

<400> 61	
gggggctggg gccgaacatc tgg	23
<210> 62	
<211> 23	
<212> DNA	
<213> Artificial	
<220><223> Her2 target sequence	
<400> 62	
ggcccaagac tctctccca ggg	23
<210> 63	
<211> 12	
<212> DNA	
<213> Artificial	
<220><223> barcode B1	
<400> 63	
cgtcgattac ca	12
<210> 64	
<211> 12	
<212> DNA	
<213> Artificial	
<220><223> barcode B2	
<400> 64	
ccatactcgta cg	12
<210> 65	
<211> 12	
<212> DNA	
<213> Artificial	
<220><223> barcode B3	
<400> 65	
tcgatagtagtac gt	12
<210> 66	
<211> 12	
<212> DNA	

<213> Artificial	
<220><223> barcode B4	
<400> 66	
cgtactaac ga	12
<210> 67	
<211> 12	
<212> DNA	
<213> Artificial	
<220><223> barcode B5	
<400> 67	
cgaagtacta cg	12
<210> 68	
<211> 12	
<212> DNA	
<213> Artificial	
<220><223> barcode B6	
<400> 68	
tcgtaacgac ca	12
<210> 69	
<211> 12	
<212> DNA	
<213> Artificial	
<220><223> barcode B7	
<400> 69	
ccagacgaat cg	12
<210> 70	
<211> 12	
<212> DNA	
<213> Artificial	
<220><223> barcode B8	
<400> 70	
tcgaatagtc gt	12
<210> 71	

<211>	12		
<212>	DNA		
<213>	Artificial		
<220><223>	barcode B9		
<400>	71		
cgtaactaac	ga	12	
<210>	72		
<211>	12		
<212>	DNA		
<213>	Artificial		
<220><223>	barcode B10		
<400>	72		
cgaacaatcg	ta	12	
<210>	73		
<211>	23		
<212>	DNA		
<213>	Artificial		
<220><223>	CEN17.1 target		
<400>	73		
gagcgcttc	aggcctgtgg	tgg	23
<210>	74		
<211>	23		
<212>	DNA		
<213>	Artificial		
<220><223>	CEN17.2 target		
<400>	74		
gagggcatttgc	aggcctgtgg	tgg	23
<210>	75		
<211>	23		
<212>	DNA		
<213>	Artificial		
<220><223>	CEN17.3 target		
<400>	75		
gagggcatttgc	tggtttgtgg	tgg	23

<210> 76	
<211> 23	
<212> DNA	
<213> Artificial	
<220><223> CEN17.4 target	
<400> 76	
ggaatctgca agtggatatg tgg	23
<210> 77	
<211> 23	
<212> DNA	
<213> Artificial	
<220><223> CEN17.5 target	
<400> 77	
gtgtgaaac tctttttg tgg	23
<210> 78	
<211> 23	
<212> DNA	
<213> Artificial	
<220><223> CEN17.6 target	
<400> 78	
gtttccaaatc actcttgatg tgg	23
<210> 79	
<211> 23	
<212> DNA	
<213> Artificial	
<220><223> CEN17.7 target	
<400> 79	
gtttggaaac actcttgatg tgg	23
<210> 80	
<211> 23	
<212> DNA	
<213> Artificial	
<220><223> CENT17.8 target	

<400> 80	
gttttgaac tctttctg tgg	23
<210> 81	
<211> 22	
<212> DNA	
<213> Artificial	
<220><223> centromeric repeat target	
<400> 81	
acactgtct atccatagga gg	22
<210> 82	
<211> 22	
<212> DNA	
<213> Artificial	
<220><223> centromeric repeat target	
<400> 82	
agatatttg accgctctga gg	22
<210> 83	
<211> 22	
<212> DNA	
<213> Artificial	
<220><223> CENT17.11 TARGET	
<400> 83	
agcgcttca ggcctgtggc gg	22
<210> 84	
<211> 22	
<212> DNA	
<213> ARTIFICIAL	
<220><223> CEN17.12 TARGET	
<400> 84	
aggaatgttc aactctgtga gg	22
<210> 85	
<211> 22	
<212> DNA	
<213> ARTIFICIAL	

<220><223> CENT17.13 TARGET

<400> 85

agggcttga gccctgtgg gg

22

<210> 86

<211> 22

<212> DNA

<213> ARTIFICIAL

<220><223> CENT17.14 TARGET

<400> 86

agggcttgt gtttgtgg gg

22

<210> 87

<211> 22

<212> DNA

<213> ARTIFICIAL

<220><223> CEN17.15 TARGET

<400> 87

catcacagag aagttctga gg

22

<210> 88

<211> 22

<212> DNA

<213> ARTIFICIAL

<220><223> CEN17.16 TARGET

<400> 88

ctgcattcaa ctcacagtgt gg

22

<210> 89

<211> 22

<212> DNA

<213> ARTIFICIAL

<220><223> CEN17.17

<400> 89

gaaaggaaag ttcaactcg gg

22

<210> 90

<211> 22

<212> DNA

<213> ARTIFICIAL

<220><223> CEN17.18 TARGET

<400> 90

gaatctgcaa gtggatatgt gg

22

<210> 91

<211> 22

<212> DNA

<213> ARTIFICIAL

<220><223> CEN17.19 TARGET

<400> 91

gaatgcaaac atcacgaaga gg

22

<210> 92

<211> 22

<212> DNA

<213> ARTIFICIAL

<220><223> CEN17.20 TARGET

<400> 92

gcatattgg accttttga gg

22

<210> 93

<211> 22

<212> DNA

<213> ARTIFICIAL

<220><223> CEN17.21 TARGET

<400> 93

gcttctgttt agttctgtgc gg

22

<210> 94

<211> 22

<212> DNA

<213> ARTIFICIAL

<220><223> CEN17.22 TARGET

<400> 94

gcttctgttt agttctgtgc gg

22

<210> 95
<211> 22
<212> DNA
<213> ARTIFICIAL
<220><223> CEN17.23 TARGET
<400> 95
ggacat tgg agggcttga gg 22
<210> 96
<211> 22
<212> DNA
<213> ARTIFICIAL
<220><223> CENT17.24 TARGET
<400> 96
ggacgttgg agggcttgt gg 22
<210> 97
<211> 22
<212> DNA
<213> ARTIFICIAL
<220><223> CEN17.25 TARGET
<400> 97
ggagat tgg agcgcttga gg 22

<210> 98
<211> 22
<212> DNA
<213> ARTIFICIAL
<220><223> CENT17.26 TARGET
<400> 98
ggatatttag gcctctctga gg 22
<210> 99
<211> 22
<212> DNA
<213> ARTIFICIAL
<220><223> CEN17.27 TARGET
<400> 99

ggatattgg accactctgt gg	22
<210> 100	
<211> 22	
<212> DNA	
<213> ARTIFICIAL	
<220><223> CEN17.28 TARGET	
<400> 100	
ggatattgg acctctctga gg	22
<210> 101	
<211> 22	
<212> DNA	
<213> ARTIFICIAL	
<220><223> CEN17.29 TARGET	
<400> 101	
gggatcattg cacttttga gg	22
<210> 102	
<211> 22	
<212> DNA	
<213> ARTIFICIAL	
<220><223> CEN17.30 TARGET	
<400> 102	
tactaccata ggcctaaagc gg	22
<210> 103	
<211> 22	
<212> DNA	
<213> ARTIFICIAL	
<220><223> CEN17.31 TARGET	
<400> 103	
tatttgtaga atgtgcaagt gg	22
<210> 104	
<211> 22	
<212> DNA	
<213> ARTIFICIAL	

<220><223> CEN17.32 TARGET

<400> 104

tccaaagaca tcttcggaga gg	22
--------------------------	----

<210> 105

<211> 22

<212> DNA

<213> ARTIFICIAL

<220><223> CEN17.33 TARGET

<400> 105

tccaacgaaa tcctcagaga gg	22
--------------------------	----

<210> 106

<211> 22

<212> DNA

<213> ARTIFICIAL

<220><223> CEN17.34 TARGET

<400> 106

tccaacgaaa tcctcagaga gg	22
--------------------------	----

<210> 107

<211> 22

<212> DNA

<213> ARTIFICIAL

<220><223> CEN17.35 TARGET

<400> 107

tccaacgaaa tcctcagagc gg	22
--------------------------	----

<210> 108

<211> 22

<212> DNA

<213> ARTIFICIAL

<220><223> CENT17.36 TARGET

<400> 108

tccaacgaaa tcttcaaaga gg	22
--------------------------	----

<210> 109

<211> 22

<212> DNA

<213> ARTIFICIAL

<220><223> CEN17.37 TARGET

<400> 109

tccaacgaaa tgctcagaga gg

22

<210> 110

<211> 22

<212> DNA

<213> ARTIFICIAL

<220><223> CEN17.38 TARGET

<400> 110

tcgAACGAAG GACACAGAGT GG

22

<210> 111

<211> 22

<212> DNA

<213> ARTIFICIAL

<220><223> CEN17.39 TARGET

<400> 111

tcgAACGAAG GCCACAGAGT GG

22

<210> 112

<211> 22

<212> DNA

<213> ARTIFICIAL

<220><223> CEN17.40 TARGET

<400> 112

tctgcaagtg gacatTTGGA GG

22

<210> 113

<211> 22

<212> DNA

<213> ARTIFICIAL

<220><223> CEN17.41 TARGET

<400> 113

tctgcaagtg gacgtttGGA GG

22

<210> 114

<211> 22

<212> DNA

<213> ARTIFICIAL

<220><223> CEN17.42 TARGET

<400> 114

tggagcgctt tcaggcctgt gg

22

<210> 115

<211> 22

<212> DNA

<213> ARTIFICIAL

<220><223> CEN17.43 TARGET

<400> 115

tggaggcgtt tgaggcctgt gg

22

<210> 116

<211> 22

<212> DNA

<213> ARTIFICIAL

<220><223> CEN17.44 TARGET

<400> 116

tggaggcgtt tgggtttgt gg

22

<210> 117

<211> 22

<212> DNA

<213> ARTIFICIAL

<220><223> CEN17.45 TARGET

<400> 117

tgttgaaact ctcttttgt gg

22

<210> 118

<211> 22

<212> DNA

<213> ARTIFICIAL

<220><223> CEN17.46 TARGET

<400> 118

tttgtgtgga atgtgcaagt gg

22

<210> 119

<211> 22

<212> DNA

<213> ARTIFICIAL

<220><223> CEN17.47 TARGET

<400> 119

tttccaatca ctctttgtgt gg

22

<210> 120

<211> 22

<212> DNA

<213> ARTIFICIAL

<220><223> CEN17.48 TARGET

<400> 120

tttcgtggc atctgcaagg gg

22

<210> 121

<211> 22

<212> DNA

<213> ARTIFICIAL

<220><223> CEN17.49 TARGET

<400> 121

tttgaaaca ctcttgttgg gg

22

<210> 122

<211> 22

<212> DNA

<213> ARTIFICIAL

<220><223> CEN17.50 TARGET

<400> 122

tttgttaga atctgcaagt gg

22

<210> 123

<211> 22

<212> DNA

<213> ARTIFICIAL

<220><223> CEN17.51 TARGET

<400> 123

tttgtgtgga atctgcaagt gg

22

<210> 124

<211> 22

<212> DNA

<213> ARTIFICIAL

<220><223> CEN17.52 TARGET

<400> 124

tttcgttagt gtctacaagt gg

22

<210> 125

<211> 22

<212> DNA

<213> ARTIFICIAL

<220><223> CEN17.53 TARGET

<400> 125

tttgaaact ctcttctgt gg

22

<210> 126

<211> 22

<212> DNA

<213> ARTIFICIAL

<220><223> CEN17.54 TARGET

<400> 126

ttttccaga atctgcaagt gg

22

<210> 127

<211> 22

<212> DNA

<213> ARTIFICIAL

<220><223> CEN17.55 TARGET

<400> 127

ttttctaga atctgcaagt gg

22

<210> 128

<211> 22

<212> DNA

<213> ARTIFICIAL

<220><223> CEN17.56 TARGET

<400> 128

ttttgcagg atctacaagt gg

22

<210> 129

<211> 22

<212> DNA

<213> ARTIFICIAL

<220><223> CEN17.57 TARGET

<400> 129

ttttgtaca atctacaagt gg

22

<210> 130

<211> 22

<212> DNA

<213> ARTIFICIAL

<220><223> CEN17.58 TARGET

<400> 130

ttttgtaga aactgcaagg gg

22

<210> 131

<211> 22

<212> DNA

<213> ARTIFICIAL

<220><223> CEN17.59 TARGET

<400> 131

ttttgtaga aactgcaagt gg

22

<210> 132

<211> 22

<212> DNA

<213> ARTIFICIAL

<220><223> CEN17.60 TARGET

<400> 132

ttttgtagg atctgcaagt gg

22

<210> 133

<211> 22

<212> DNA

<213> ARTIFICIAL

<220><223> CEN17.61 TARGET

<400> 133

ttttgtgga atctgcaagt gg

22

<210> 134

<211> 20

<212> DNA

<213> ARTIFICIAL

<220><223> gRNA TARGET SEQUENCE

<400> 134

gagggctttg aggccctgtgg

20

<210> 135

<211> 20

<212> DNA

<213> ARTIFICIAL

<220><223> gRNA TARGET SEQUENCE

<400> 135

gtgttgaac tcttttttg

20

<210> 136

<211> 20

<212> DNA

<213> ARTIFICIAL

<220><223> gRNA TARGET SEQUENCE

<400> 136

gacactgctc tatccatagg

20

<210> 137

<211> 20

<212> DNA

<213> ARTIFICIAL

<220><223> gRNA TARGET SEQUENCE

<400> 137

gagggcatttgg	20
<210> 138	
<211> 20	
<212> DNA	
<213> ARTIFICIAL	
<220><223> gRNA TARGET SEQUENCE	
<400> 138	
ggaatctgca agtggatatg	20
<210> 139	
<211> 20	
<212> DNA	
<213> ARTIFICIAL	
<220><223> gRNA TARGET SEQUENCE	
<400> 139	
ggcatattttg gaccttttg	20
<210> 140	
<211> 20	
<212> DNA	
<213> ARTIFICIAL	
<220><223> gRNA TARGET SEQUENCE	
<400> 140	
gtactaccat aggctaaag	20
<210> 141	
<211> 20	
<212> DNA	
<213> ARTIFICIAL	
<220><223> gRNA TARGET SEQUENCE	
<400> 141	
gtccaaacgaa atcctcagag	20
<210> 142	
<211> 20	
<212> DNA	
<213> ARTIFICIAL	
<220><223> gRNA TARGET SEQUENCE	

<400> 142	
gtctgcaagt ggacatgg	20
<210> 143	
<211> 20	
<212> DNA	
<213> ARTIFICIAL	
<220><223> gRNA TARGET SEQUENCE	
<400> 143	
gttgttgtgg aatgtgcaag	20
<210> 144	
<211> 20	
<212> DNA	
<213> ARTIFICIAL	
<220><223> gRNA TARGET SEQUENCE	
<400> 144	
gtttgtgttag aatctgcaag	20
<210> 145	
<211> 23	
<212> DNA	
<213> ARTIFICIAL	
<220><223> gRNA TARGET	
<400> 145	
tctgacgtcc atcatctctg cggtt	23
<210> 146	
<211> 23	
<212> DNA	
<213> ARTIFICIAL	
<220><223> gRNA TARGET	
<400> 146	
gccaaaccacc gcagagatga tgg	23