



등록특허 10-2640585



(19) 대한민국특허청(KR)

(12) 등록특허공보(B1)

(45) 공고일자 2024년02월23일

(11) 등록번호 10-2640585

(24) 등록일자 2024년02월21일

(51) 국제특허분류(Int. Cl.)
C12Q 1/6827 (2018.01) C12Q 1/6869 (2018.01)
G16B 30/00 (2019.01)
(52) CPC특허분류
C12Q 1/6827 (2018.05)
C12Q 1/6869 (2018.05)
(21) 출원번호 10-2022-7009793(분할)
(22) 출원일자(국제) 2014년12월11일
심사청구일자 2022년04월22일
(85) 번역문제출일자 2022년03월23일
(65) 공개번호 10-2022-0064973
(43) 공개일자 2022년05월19일
(62) 원출원 특허 10-2016-7018274
원출원일자(국제) 2014년12월11일
심사청구일자 2019년12월10일
(86) 국제출원번호 PCT/US2014/069848
(87) 국제공개번호 WO 2015/089333
국제공개일자 2015년06월18일
(30) 우선권주장
61/914,907 2013년12월11일 미국(US)
(뒷면에 계속)
(56) 선행기술조사문헌
W02001038580 A2

(73) 특허권자
아큐라젠 홀딩스 리미티드
케이먼 군도, 그랑 케이먼, 조지 타운, 사우스 처치 스트리트, 하버 플래스 2층, 피.오. 박스 472 (우편번호 케이와이1-1106)
(72) 발명자
런 생롱
미국 94555 캘리포니아주 프레몬트 윌브리지 테라스 34568
선 자오후이
미국 94706 캘리포니아주 알바니 가필드 애비뉴 1152
(뒷면에 계속)
(74) 대리인
김진희, 김태홍

전체 청구항 수 : 총 12 항

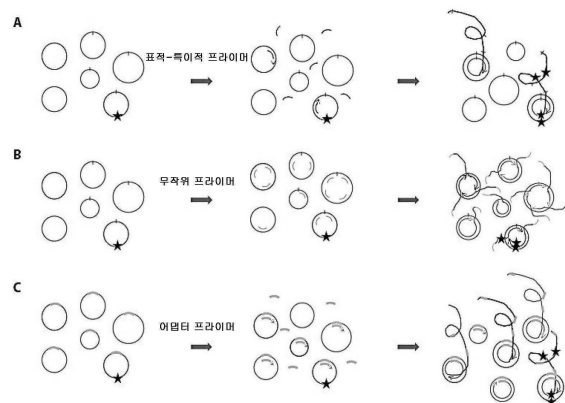
심사관 : 김현태

(54) 발명의 명칭 희귀 서열 변이를 검출하기 위한 조성물 및 방법

(57) 요약

일부 양태에서, 본 개시내용은 핵산 샘플에서 서열 변이를 확인하는 방법을 제공한다. 일부 구현예에서, 방법은 시퀀싱 리드와 참조 서열 간의 서열 차이를 확인하는 단계, 및 적어도 2개의 상이한 원형 폴리뉴클레오타이드, 예컨대 상이한 접합부를 갖는 2개의 원형 폴리뉴클레오타이드에서 발생하는 서열 차이를 서열 변이로서 판정하는 단계를 포함한다. 일부 양태에서, 본 개시내용은 상기 기재된 방법에서 유용한 조성물 및 시스템을 제공한다.

대표도 - 도7



(52) CPC특허분류

G16B 30/00 (2019.02)
C12Q 2521/501 (2013.01)
C12Q 2525/191 (2013.01)
C12Q 2525/307 (2013.01)
C12Q 2527/101 (2019.08)
C12Q 2531/113 (2013.01)
C12Q 2531/125 (2013.01)
C12Q 2531/131 (2013.01)
C12Q 2535/122 (2019.08)

(30) 우선권주장

61/987,414 2014년05월01일 미국(US)
62/010,975 2014년06월11일 미국(US)

(72) 발명자

자오 그레이스 쿼즈히

미국 94306 캘리포니아주 팔로 알토 던칸 플레이스
3969

탕 폴 링-퐁

미국 94132 캘리포니아주 샌프란시스코 19번 애비
뉴 2924

명세서

청구범위

청구항 1

서열 변이를 확인하는 방법으로서, 방법은 반응 혼합물에서 표적 서열을 포함하는 복수의 상이한 원형 폴리뉴클레오타이드를 증폭하는 단계를 포함하고, 표적 서열은 5'에서 3' 방향으로 배향된 서열 A 및 서열 B를 포함하고, 방법은 반응 혼합물에 대해 핵산 증폭 반응을 실시하는 것을 포함하고, 반응 혼합물은

(a) 복수의 원형 폴리뉴클레오타이드로서, 복수의 원형 뉴클레오타이드에서의 개개 원형 폴리뉴클레오타이드가 5' 말단 및 3' 말단을 갖는 개개 폴리뉴클레오타이드를 원형화시킴으로써 형성되는 상이한 접합부를 포함하는 것인 복수의 원형 폴리뉴클레오타이드;

(b) 서열 A'를 포함하는 제1 프라이머로서, 제1 프라이머는 서열 A와 서열 A' 간의 서열 상보성을 통해 표적 서열의 서열 A에 특이적으로 혼성화하는 것인 제1 프라이머;

(c) 서열 B를 포함하는 제2 프라이머로서, 제2 프라이머는 서열 B와 서열 B' 간의 서열 상보성을 통해 표적 서열의 보체를 포함하는 상보적 폴리뉴클레오타이드에 존재하는 서열 B'에 특이적으로 혼성화하는 것인 제2 프라이머; 및

(d) 제1 프라이머 및 제2 프라이머를 연장시켜 증폭된 폴리뉴클레오타이드를 생성하는 가닥 치환(strand displacement) 활성을 갖는 폴리머라아제

를 포함하고, 방법은

증폭된 폴리뉴클레오타이드를 시퀀싱하여 복수의 시퀀싱 리드(sequencing read)를 생성하는 단계;

시퀀싱 리드와 참조 서열 간의 서열 차이를 확인하는 단계; 및

서열 차이가 상이한 접합부를 갖는 2개 이상의 원형 폴리뉴클레오타이드에서 발생하거나, 서열 차이가 1개 초과 의 표적 서열의 카피에서 발생할 때에만 서열 차이를 서열 변이로서 판정(call)하는 단계

를 추가로 포함하고,

서열 A 및 서열 B는 내인성 서열이며, 표적 서열의 서열 A의 5' 말단과 서열 B의 3' 말단 사이의 거리가 200nt 이하인, 서열 변이를 확인하는 방법.

청구항 2

제1항에 있어서, 서열 A 및 서열 B는 동일한 유전자의 상이한 부분에 해당하는 것인 방법.

청구항 3

제1항에 있어서, 제1 프라이머 및 제2 프라이머는 무작위 프라이머인 방법.

청구항 4

제1항에 있어서, 반응 혼합물은 복수의 상이한 표적 서열에 관한 프라이머를 포함하여 복수의 상이한 표적 서열을 병렬로 증폭시키는 것인 방법.

청구항 5

제1항에 있어서, 서열 A의 5' 말단과 서열 B의 3' 말단 사이의 거리가 150nt 이하인 방법.

청구항 6

제1항에 있어서, 서열 A의 5' 말단과 서열 B의 3' 말단 사이의 거리가 75nt 이하인 방법.

청구항 7

제1항에 있어서, 복수의 상이한 원형 폴리뉴클레오타이드가 무세포 폴리뉴클레오타이드를 포함하는 것인 방법.

청구항 8

제1항에 있어서, 복수의 상이한 원형 폴리뉴클레오타이드가 순환 종양 DNA를 포함하는 것인 방법.

청구항 9

제1항에 있어서, 복수의 상이한 원형 폴리뉴클레오타이드가 단일 가닥인 방법.

청구항 10

제1항에 있어서, 개개 폴리뉴클레오타이드를 원형화시키는 것은 라이게이션 반응에 의해 수행되는 것인 방법.

청구항 11

제1항에 있어서, 원형화는 복수의 폴리뉴클레오타이드 중의 개개 폴리뉴클레오타이드의 5' 말단, 3' 말단, 또는 5' 말단 및 3' 말단 둘다에 어댑터 폴리뉴클레오타이드를 연결하는 단계를 포함하는 것인 방법.

청구항 12

제1항에 있어서, 복수의 상이한 원형 폴리뉴클레오타이드는 50 ng 미만의 폴리뉴클레오타이드를 포함하는 것인 방법.

청구항 13

삭제

청구항 14

삭제

청구항 15

삭제

청구항 16

삭제

청구항 17

삭제

청구항 18

삭제

청구항 19

삭제

청구항 20

삭제

청구항 21

삭제

청구항 22

삭제

청구항 23

삭제

청구항 24

삭제

청구항 25

삭제

청구항 26

삭제

청구항 27

삭제

청구항 28

삭제

청구항 29

삭제

청구항 30

삭제

청구항 31

삭제

청구항 32

삭제

청구항 33

삭제

청구항 34

삭제

청구항 35

삭제

청구항 36

삭제

청구항 37

삭제

청구항 38

삭제

청구항 39

삭제

청구항 40

삭제

청구항 41

삭제

청구항 42

삭제

청구항 43

삭제

청구항 44

삭제

청구항 45

삭제

청구항 46

삭제

청구항 47

삭제

청구항 48

삭제

청구항 49

삭제

청구항 50

삭제

청구항 51

삭제

청구항 52

삭제

발명의 설명

기술 분야

[0001]

상호참조

- [0002] 본원은 2013년 12월 11일에 출원된 미국 가출원 제61/914,907호; 2014년 5월 1일에 출원된 미국 가출원 제 61/987,414호; 및 2014년 6월 11일에 출원된 미국 가출원 제62/010,975호의 이익을 주장하며, 이들 모두는 본원에 참고로 통합되어 있다.
- [0003] 기술분야
- [0004] 본 발명은 회귀 서열 변이를 검출하기 위한 조성물 및 방법에 관한 것이다.

배경 기술

- [0005] 복잡한 집단 내에서 서열 변이를 확인하는 것은 특히 대규모 병렬 핵산 시퀀싱의 출현과 더불어 활발하게 성장하는 분야이다. 그러나 대규모 병렬 시퀀싱은 흔히 사용되는 기술에서 내재하는 오류 빈도가 집단에서의 많은 실제 서열 변이의 빈도보다 더 크다는 점에서 상당한 제약이 있다. 예를 들면, 표준 고속대량(high-throughput) 시퀀싱에서 0.1 - 1%의 오류율이 보고되었다. 회귀 서열 변이의 검출은 변이의 빈도가 예컨대 상기 오류율 이하로 낮을 때 높은 위양성 비율을 갖는다.
- [0006] 회귀 서열 변이를 검출하는 것에 관한 많은 이유가 있다. 예를 들면, 회귀한 특징적 서열의 검출은 박테리아 분류군과 같은 유해한 환경 오염물질의 존재를 확인하고 구별하는데 사용될 수 있다. 박테리아 분류군을 규명하는 일반적인 방법은 rRNA 서열과 같이 고도로 보존된 서열의 차이를 확인하는 것이다. 그러나, 이에 대한 전형적인 시퀀싱 기반의 접근법은, 이미 힘든 과정에 복잡한 문제를 야기하는, 주어진 샘플 내의 상이한 지놈들의 수 자체 및 멤버들 간의 상동성 정도와 관련된 도전에 직면하고 있다. 개선된 과정은 다양한 환경에서 오염 검출을 향상시킬 잠재성을 가질 것이다. 예를 들면, 인공위성 및 다른 우주선의 부품을 조립하는데 사용되는 무균실은 어떤 미생물 군집이 존재하는지 이해하기 위해 그리고 다른 행성 또는 이의 샘플에 지구상의 미생물이 도입되는 것을 막는 더 우수한 오염제거 및 정화 기술을 개발하기 위해 또는 추정되는 외계 미생물에 의해 생성된 데이터를 오염시키는 지구상의 미생물에 의해 생성된 데이터와 구별하는 방법을 개발하기 위해, 본 발명의 시스템 및 방법을 이용하여 조사될 수 있다. 식품 감시 응용은 식품 매개 병원균에 대해 식품 가공 공장에서 생산 라인의 정기적인 검사, 도축장을 조사하는 것, 식당, 병원, 학교, 교도소 및 다른 기관의 부엌 및 식품 보관 장소를 점검하는 것을 포함한다. 저수 및 수처리 공장 역시 유사하게 감시될 수 있다.
- [0007] 회귀 변이 검출은 또한 병리학적 돌연변이의 조기 검출에 중요할 수 있다. 예를 들면, 임상 샘플에서 암-연관 점 돌연변이의 검출은 화학요법 동안 최소 잔존 질병의 확인을 개선할 수 있고 재발 환자에서 종양 세포의 출현을 검출할 수 있다. 회귀 점 돌연변이의 검출은 또한 환경 변이원에 대한 노출을 평가하고, 내인성 DNA 복구를 감사하고, 노화하는 개체에서의 체세포 돌연변이의 축적을 연구하는데 중요하다. 또한, 회귀 변이를 검출하는 보다 민감한 방법은 산전 진단을 향상시켜 산모 혈액에 존재하는 태아 세포의 특징규명을 가능하게 할 수 있다.

발명의 내용

- [0008] 전술한 내용에 비추어, 회귀 서열 변이를 검출하는 개선된 방법이 필요하다. 본 개시내용의 조성물 및 방법은 이러한 필요를 다루며, 부가적인 이점도 제공한다. 특히, 본 개시내용의 다양한 양태들은 회귀 또는 저빈도 핵산 서열 변이(종종 돌연변이로 불림)의 고민감성 검출을 제공한다. 이것은 정상 서열의 배경에 소량의 변이체 서열을 함유할 수 있는 샘플에서 저빈도 핵산 변이(치환, 삽입 및 결실 포함)를 확인하고 설명하는 것뿐만 아니라 시퀀싱 오류의 배경에 저빈도 변이를 확인하는 것을 포함한다.
- [0009] 하나의 양태에서, 상기 개시내용은, 예컨대 핵산 샘플에서 서열 변이를 확인하는 방법을 제공한다. 일부 구현예에서, 상기 복수의 폴리뉴클레오타이드의 각각의 폴리뉴클레오타이드는 5' 말단 및 3' 말단을 가지고, 상기 방법은 (a) 상기 복수의 폴리뉴클레오타이드의 각각의 폴리뉴클레오타이드를 원형화시켜 복수의 원형 폴리뉴클레오타이드를 형성하는 단계로서, 이들 각각은 5' 말단과 3' 말단 사이에 접합부(junction)를 갖고; (b) 상기 (a)의 원형 폴리뉴클레오타이드를 증폭시키는 단계; (c) 상기 증폭된 폴리뉴클레오타이드를 시퀀싱하여 복수의 시퀀싱 리드(sequencing read)를 생성하는 단계; (d) 시퀀싱 리드와 참조 서열 간의 서열 차이를 확인하는 단계; 및 (e) 상이한 접합부를 갖는 적어도 2개의 원형 폴리뉴클레오타이드에서 발생하는 서열 차이를 서열 변이로서 판정(call)하는 단계를 포함한다. 일부 구현예에서, 상기 방법은 시퀀싱 리드와 참조 서열 간의 서열 차이를 확인하는 것, 및 상이한 접합부를 갖는 적어도 2개의 원형 폴리뉴클레오타이드에서 발생하는 서열 차이를 서열 변이로서 판정하는 것을 포함하며, (a) 상기 시퀀싱 리드는 상기 적어도 2개의 원형 폴리뉴클레오타이드의 증폭 산물(생성물)에 상응하며; (b) 상기 적어도 2개의 원형 폴리뉴클레오타이드 각각은 각각의 폴리뉴클레오타이드의 5' 말단 및 3' 말단을 라이게이션(ligation)함으로써 형성되는 상이한 접합부를 포함한다.

- [0010] 상기 복수의 폴리뉴클레오타이드는 단일- 또는 이중-가닥일 수 있다. 일부 구현예에서, 상기 폴리뉴클레오타이드는 단일 가닥이다. 일부 구현예에서, 원형화는 상기 복수의 폴리뉴클레오타이드에 대해 라이게이션 반응을 실시함으로써 달성된다. 일부 구현예에서, 각각의 원형 폴리뉴클레오타이드는 상기 원형화된 폴리뉴클레오타이드들 간에 공유한 접합부를 갖는다. 일부 구현예에서, 상기 서열 변이는 단일 뉴클레오타이드 다형(SNP)이다. 일부 구현예에서, 상기 참조 서열은 상기 서열 리드를 서로 정렬함으로써 형성된 콘센서스(consensus) 서열이다. 일부 구현예에서, 상기 참조 서열은 공지된 참조 서열, 예컨대 참조 지놈 또는 이의 일부이다. 일부 구현예에서, 원형화는 상기 복수의 폴리뉴클레오타이드 내의 하나의 폴리뉴클레오타이드의 5' 말단, 3' 말단, 또는 5' 말단 및 3' 말단 모두에 어댑터 폴리뉴클레오타이드를 연결하는 단계를 포함한다. 일부 구현예에서, 증폭은, 예컨대 롤링-서클 증폭(RCA)에서, 가닥 치환(strand-displacement) 활성을 갖는 폴리머라아제를 이용하여 달성된다. 일부 구현예에서, 증폭은 상기 원형 폴리뉴클레오타이드를 무작위 프라이머를 포함하는 증폭 반응 혼합물에 두는(subjecting) 것을 포함한다. 일부 구현예에서, 증폭은 상기 원형 폴리뉴클레오타이드를 하나 이상의 프라이머를 포함하는 증폭 반응 혼합물에 두는 것을 포함하고, 상기 프라이머 각각은 서열 상보성을 통해 상이한 표적 서열에 특이적으로 혼성화한다. 일부 구현예에서, 미생물 오염물질은 상기 판정 단계에 기초하여 확인된다.
- [0011] 증폭된 폴리뉴클레오타이드는, 시퀀싱 이전에 농축(enrichment) 단계를 수행함으로써 상기 증폭된 폴리뉴클레오타이드 중에 하나 이상의 표적 폴리뉴클레오타이드를 농축시키는 것과 같이, 농축을 수반하거나 농축 없이 시퀀싱될 수 있다. 일부 구현예에서, 상기 농축 단계는 증폭된 폴리뉴클레오타이드를 기질에 부착된 복수의 프로브에 혼성화시키는 것을 포함한다. 일부 구현예에서, 농축 단계는 하기를 포함하는 증폭 반응 혼합물에서 5'에서 3' 방향으로 배향된 서열 A 및 서열 B를 포함하는 표적 서열을 증폭시키는 것을 포함한다: (a) 상기 증폭된 폴리뉴클레오타이드; (b) 서열 A'를 포함하는 제1 프라이머, 여기서 상기 제1 프라이머는 서열 A와 서열 A' 간의 서열 상보성을 통해 표적 서열의 서열 A에 특이적으로 혼성화하며; (c) 서열 B를 포함하는 제2 프라이머, 여기서 상기 제2 프라이머는 B와 B' 간의 서열 상보성을 통해 표적 서열의 보체(complement)를 포함하는 상보적 폴리뉴클레오타이드에 존재하는 서열 B'에 특이적으로 혼성화하고; (d) 상기 제1 프라이머 및 상기 제2 프라이머를 연장시켜 증폭된 폴리뉴클레오타이드를 생성하는 폴리머라아제; 여기서 표적 서열의 서열 A의 5' 말단과 서열 B의 3' 말단 사이의 거리는 75nt 이하이다.
- [0012] 하나의 양태에서, 상기 개시내용은 50 ng 미만의 폴리뉴클레오타이드를 포함하는 핵산 샘플에서 서열 변이를 확인하는 방법을 제공하며, 각각의 폴리뉴클레오타이드는 5' 말단 및 3' 말단을 갖는다. 일부 구현예에서, 상기 방법은 (a) 상기 샘플 내의 각각의 폴리뉴클레오타이드를 리가아제를 이용하여 원형화시켜 복수의 원형 폴리뉴클레오타이드를 형성하는 단계; (b) 상기 리가아제를 상기 원형 폴리뉴클레오타이드로부터 분리시, 상기 원형 폴리뉴클레오타이드를 증폭시켜 연쇄체(concatemer)를 형성하는 단계; (c) 상기 연쇄체를 시퀀싱하여 복수의 시퀀싱 리드를 생성하는 단계; (d) 상기 복수의 시퀀싱 리드와 참조 서열 간의 서열 차이를 확인하는 단계; 및 (e) 50ng 미만의 폴리뉴클레오타이드의 상기 핵산 샘플로부터 상기 복수의 리드에서 0.05% 이상의 빈도로 발생하는 서열 차이를 서열 변이로서 판정하는 단계를 포함한다. 상기 폴리뉴클레오타이드는 단일- 또는 이중-가닥일 수 있다. 일부 구현예에서, 상기 폴리뉴클레오타이드는 단일 가닥이다. 일부 구현예에서, 각각의 원형 폴리뉴클레오타이드는 상기 원형화된 폴리뉴클레오타이드들 간에 공유한 접합부를 갖는다. 일부 구현예에서, 상기 서열 변이는 단일 뉴클레오타이드 다형이다. 일부 구현예에서, 상기 참조 서열은 상기 시퀀싱 리드를 서로 정렬함으로써 형성된 콘센서스 서열이다. 일부 구현예에서, 상기 참조 서열은 공지된 참조 서열, 예컨대 참조 지놈이다. 일부 구현예에서, 증폭은 가닥 치환 활성을 갖는 폴리머라아제를 이용함으로써 달성된다. 일부 구현예에서, 증폭은 상기 원형 폴리뉴클레오타이드를 무작위 프라이머를 포함하는 증폭 반응 혼합물에 두는 것을 포함한다. 일부 구현예에서, 증폭은 상기 원형 폴리뉴클레오타이드를 하나 이상의 프라이머를 포함하는 증폭 반응 혼합물에 두는 것을 포함하고, 상기 프라이머 각각은 서열 상보성을 통해 상이한 표적 서열에 특이적으로 혼성화한다.
- [0013] 하나의 양태에서, 상기 개시내용은 반응 혼합물에서 표적 서열의 2개 이상의 카피를 포함하는 복수의 상이한 연쇄체를 증폭하는 방법을 제공하며, 여기서 상기 표적 서열은 5'에서 3' 방향으로 배향된 서열 A 및 서열 B를 포함한다. 일부 구현예에서, 상기 방법은 상기 반응 혼합물을 핵산 증폭 반응시키는 것을 포함하고, 여기서 상기 반응 혼합물은 하기를 포함한다: (a) 복수의 연쇄체, 여기서 상기 복수 내에서 각각의 연쇄체는 5' 말단 및 3' 말단을 갖는 각각의 폴리뉴클레오타이드를 원형화시킴으로써 형성되는 상이한 접합부를 포함하고; (b) 서열 A'를 포함하는 제1 프라이머, 여기서 상기 제1 프라이머는 서열 A와 서열 A' 간의 서열 상보성을 통해 표적 서열의 서열 A에 특이적으로 혼성화하며; (c) 서열 B를 포함하는 제2 프라이머, 여기서 상기 제2 프라이머는 서열 B

와 B' 간의 서열 상보성을 통해 표적 서열의 보체를 포함하는 상보적 폴리뉴클레오타이드에 존재하는 서열 B'에 특이적으로 혼성화하고; 및 (d) 상기 제1 프라이머 및 상기 제2 프라이머를 연장시켜 증폭된 폴리뉴클레오타이드를 생성하는 폴리머라아제; 여기서 표적 서열의 서열 A의 5' 말단과 서열 B의 3' 말단 사이의 거리는 75nt 이하이다. 일부 구현예에서, 상기 제1 프라이머는 서열 A'에 대해 서열 C 5'를 포함하고, 상기 제2 프라이머는 서열 B에 대해 서열 D 5'를 포함하며, 서열 C 및 서열 D 어느 것도 제1 혼성화 온도에서 제1 증폭 과정(phase) 동안 상기 복수의 연쇄체에 혼성화하지 않는다. 일부 구현예에서, 증폭은 제1 과정 및 제2 과정을 포함하고; 상기 제1 과정은 프라이머 연장 전에 상기 제1 및 제2 프라이머가 연쇄체에 혼성화하는 제1 온도에서의 혼성화 단계를 포함하며; 상기 제2 과정은 상기 제1 및 제2 프라이머가 연장된 제1 또는 제2 프라이머 또는 이의 보체를 포함하는 증폭 산물에 혼성화하는, 상기 제1 온도보다 더 높은 제2 온도에서의 혼성화 단계를 포함한다. 일부 구현예에서, 제2 온도에서의 혼성화 및 프라이머 연장의 5 사이클 후, 상기 반응 혼합물 내의 증폭된 폴리뉴클레오타이드 중 적어도 5%가 표적 서열의 2개 이상의 카피를 포함한다.

[0014] 관련된 양태에서, 상기 개시내용은 반응 혼합물에서 표적 서열을 포함하는 복수의 상이한 원형 폴리뉴클레오타이드를 증폭하는 방법을 제공하고, 여기서 상기 표적 서열은 5'에서 3' 방향으로 배향된 서열 A 및 서열 B를 포함한다. 일부 구현예에서, 상기 방법은 상기 반응 혼합물을 핵산 증폭 반응시키는 것을 포함하고, 여기서 상기 반응 혼합물은 하기를 포함한다: (a) 복수의 원형 폴리뉴클레오타이드, 여기서 상기 복수의 원형 폴리뉴클레오타이드에서의 각각의 원형 폴리뉴클레오타이드는 5' 말단 및 3' 말단을 갖는 각각의 폴리뉴클레오타이드를 원형화시킴으로써 형성되는 상이한 접합부를 포함하고; (b) 서열 A'를 포함하는 제1 프라이머, 여기서 상기 제1 프라이머는 서열 A와 서열 A' 간의 서열 상보성을 통해 표적 서열의 서열 A에 특이적으로 혼성화하며; (c) 서열 B를 포함하는 제2 프라이머, 여기서 상기 제2 프라이머는 서열 B와 B' 간의 서열 상보성을 통해 표적 서열의 보체를 포함하는 상보적 폴리뉴클레오타이드에 존재하는 서열 B'에 특이적으로 혼성화하고; (d) 상기 제1 프라이머 및 상기 제2 프라이머를 연장시켜 증폭된 폴리뉴클레오타이드를 생성하는 폴리머라아제; 여기서 서열 A 및 서열 B는 내인성 서열이며, 표적 서열의 서열 A의 5' 말단과 서열 B의 3' 말단 사이의 거리는 75nt 이하이다. 일부 구현예에서, 상기 제1 프라이머는 서열 A'에 대해 서열 C 5'를 포함하고, 상기 제2 프라이머는 서열 B에 대해 서열 D 5'를 포함하며, 서열 C 및 서열 D 어느 것도 제1 혼성화 온도에서의 제1 증폭 과정 동안 상기 복수의 원형 폴리뉴클레오타이드에 혼성화하지 않는다. 일부 구현예에서, 증폭은 제1 과정 및 제2 과정을 포함하고; 상기 제1 과정은 프라이머 연장 전에 상기 제1 및 제2 프라이머가 상기 원형 폴리뉴클레오타이드 또는 이의 증폭 산물에 혼성화하는 제1 온도에서의 혼성화 단계를 포함하며; 상기 제2 과정은 상기 제1 및 제2 프라이머가 연장된 제1 또는 제2 프라이머 또는 이의 보체를 포함하는 증폭 산물에 혼성화하는, 상기 제1 온도보다 더 높은 제2 온도에서의 혼성화 단계를 포함한다.

[0015] 하나의 양태에서, 상기 개시내용은 본 개시내용의 방법에 따른 방법을 수행하기 위한 반응 혼합물을 제공한다. 상기 반응 혼합물은 임의의 다양한 방법에 대하여 본원에 기재된 바와 같은 다양한 구성요소 중 하나 이상을 포함할 수 있다. 일부 구현예에서, 상기 반응 혼합물은 표적 서열의 2개 이상의 카피를 포함하는 복수의 상이한 연쇄체를 증폭시키기 위한 혼합물이며, 여기서 상기 표적 서열은 5'에서 3' 방향으로 배향된 서열 A 및 서열 B를 포함하고, 상기 반응 혼합물은 하기를 포함한다: (a) 복수의 연쇄체, 여기서 상기 복수의 연쇄체 내의 각각의 연쇄체는 5' 말단 및 3' 말단을 갖는 각각의 폴리뉴클레오타이드를 원형화시킴으로써 형성되는 상이한 접합부를 포함하고; (b) 서열 A'를 포함하는 제1 프라이머, 여기서 상기 제1 프라이머는 서열 A와 서열 A' 간의 서열 상보성을 통해 표적 서열의 서열 A에 특이적으로 혼성화하고; (c) 서열 B를 포함하는 제2 프라이머, 여기서 상기 제2 프라이머는 B와 B' 간의 서열 상보성을 통해 표적 서열의 보체를 포함하는 상보적 폴리뉴클레오타이드에 존재하는 서열 B'에 특이적으로 혼성화하며; (d) 상기 제1 프라이머 및 상기 제2 프라이머를 연장시켜 증폭된 폴리뉴클레오타이드를 생성하는 폴리머라아제; 여기서 표적 서열의 서열 A의 5' 말단과 서열 B의 3' 말단 사이의 거리는 75nt 이하이다. 일부 구현예에서, 상기 제1 프라이머는 서열 A'에 대해 서열 C 5'를 포함하고, 상기 제2 프라이머는 서열 B에 대해 서열 D 5'를 포함하며, 서열 C 및 서열 D 어느 것도 증폭 반응에서 제1 증폭 단계 동안 상기 2개 이상의 연쇄체에 혼성화하지 않는다.

[0016] 하나의 양태에서, 상기 개시내용은, 예컨대 본 개시내용의 임의의 다양한 다른 양태에서, 본원에 기재된 방법에서 유용하거나 본원에 기재된 방법에 의해 생성된 조성물을 제공한다. 일부 구현예에서, 상기 조성물은 단일 가닥이며 실질적으로 리가아제가 없는 복수의 원형화된 폴리뉴클레오타이드를 포함한다. 일부 구현예에서, 상기 조성물은 복수의 연쇄체를 포함하고, 여기서 복수의 연쇄체는 10000개 이하의 표적 폴리뉴클레오타이드의 군에 상응하며, 추가로 상기 복수의 연쇄체 내의 각각의 연쇄체는: (a) 각각의 연쇄체가 서열 반복의 2개 이상의 카피를 포함하고, 여기서 상기 카피 모두는 동일한 표적 폴리뉴클레오타이드에 상응한다는 점; 및 (b) 각각의 연쇄체의 서열 반복의 2개 이상의 카피 사이의 접합부가 상기 조성물 내의 또 다른 각각의 연쇄체와 상이하다는

점을 특징으로 한다.

[0017] 하나의 양태에서, 상기 개시내용은 서열 변이를 검출하는 시스템을 제공한다. 일부 구현예에서, 상기 시스템은 하기를 포함한다 (a) 샘플에 대해 검출 반응을 수행해 달라는 사용자 요청을 수신(수용)하도록 구성된 컴퓨터; (b) 사용자 요청에 대한 반응으로 샘플 또는 이의 일부에 대해 핵산 증폭 반응을 수행하는 증폭 시스템, 여기서 상기 증폭 반응은 (i) 각각의 폴리뉴클레오타이드를 원형화시켜 복수의 원형 폴리뉴클레오타이드를 형성하는 단계로서, 이들 각각은 5' 말단과 3' 말단 사이에 접합부를 가지며; (ii) 상기 원형 폴리뉴클레오타이드를 증폭시키는 단계를 포함하며; (c) 상기 증폭 시스템에 의해 증폭된 폴리뉴클레오타이드에 대해 시퀀싱 리드를 생성하고 시퀀싱 리드와 참조 서열 간의 서열 차이를 확인하고, 상이한 접합부를 갖는 적어도 2개의 원형 폴리뉴클레오타이드에서 발생하는 서열 차이를 서열 변이로서 판정하는 시퀀싱 시스템; 및 (d) 수신자에게 보고서를 보내는 보고서 생성기, 여기서 상기 보고서는 서열 변이의 검출 결과를 함유한다. 일부 구현예에서, 상기 수신자는 사용자이다.

[0018] 하나의 양태에서, 상기 개시내용은 하나 이상의 프로세서에 의해 실행시, 서열 변이를 검출하는 방법을 구현하는 코드를 포함하는 컴퓨터 판독가능 매체를 제공한다. 일부 구현예에서, 상기 구현된 방법은 하기를 포함한다: (a) 샘플에 대해 검출 반응을 수행해 달라는 고객 요청을 수신하는 단계; (b) 상기 고객 요청에 대한 반응으로 샘플 또는 이의 일부에 대해 핵산 증폭 반응을 수행하는 단계, 여기서 상기 증폭 반응은 (i) 각각의 폴리뉴클레오타이드를 원형화시켜 복수의 원형 폴리뉴클레오타이드를 형성하는 단계로서, 이들 각각은 5' 말단과 3' 말단 사이에 접합부를 가지며; (ii) 상기 원형 폴리뉴클레오타이드를 증폭시키는 단계를 포함하고; (c) (i) 상기 증폭 반응에서 증폭된 폴리뉴클레오타이드에 대해 시퀀싱 리드를 생성하는 단계; (ii) 시퀀싱 리드와 참조 서열 간의 서열 차이를 확인하는 단계; 및 (iii) 상이한 접합부를 갖는 상기 적어도 2개의 원형 폴리뉴클레오타이드에 존재하는 서열 차이를 서열 변이로서 판정하는 단계; 및 (d) 상기 서열 변이의 검출 결과를 함유하는 보고서를 생성하는 단계를 포함하는 시퀀싱 분석을 수행하는 단계.

[0019] 본원에 개시된 임의의 다양한 양태의 일부 구현예에서, 상기 방법, 조성물, 및 시스템은, 예컨대 환자 샘플의 규명 및 선택적으로 대상(피험체)의 상태의 진단에서, 치료적 적용을 갖는다. 일부 구현예에서, 샘플은 대상으로부터의 샘플, 예컨대 소변, 대변, 혈액, 타액, 조직, 또는 체액이다. 일부 구현예에서, 샘플은, 예컨대 대상으로부터의 종양 조직의 샘플에서, 종양 세포를 포함한다. 일부 구현예에서, 샘플은 포르말린 고정된, 파라핀 포매된(FFPE) 샘플이다. 일부 구현예에서, 방법은 판정 단계에 기초하여 대상을 진단하는 단계를 추가로 포함한다. 일부 구현예에서, 서열 변이는 원인 유전 변이(causal genetic variant)이다. 일부 구현예에서, 서열 변이는 암의 유형 또는 병기(stage)와 연관된다. 일부 구현예에서, 복수의 폴리뉴클레오타이드는 무세포 폴리뉴클레오타이드, 예컨대 무세포 DNA 또는 순환 종양 DNA를 포함한다.

[0020] 참조에 의한 통합

[0021] 본 명세서에 언급된 모든 공개문헌, 특허, 및 특허출원은, 마치 각각의 개별적인 공개문헌, 특허, 및 특허출원이 명시적으로 그리고 개별적으로 참조에 의해 통합된 것으로 표시된 것처럼 동일한 정도로 참조에 의해 본원에 통합된다.

도면의 간단한 설명

[0022] 본 발명의 신규한 특징들은 첨부된 청구범위에 상세하게 제시되어 있다. 본 발명의 특징 및 이점은 본 발명의 원리가 이용되는 예시적인 구현예를 제시하는 하기 상세한 설명, 및 첨부 도면을 참조하여 이해될 것이다.

도 1은 본 개시내용에 따른 방법의 일 구현예의 도식을 묘사한다. DNA 가닥은 원형화되고, 조사 중인 유전자에 상응하는 표적 특이적 프라이머가 폴리머라아제, dNTP, 버퍼 등과 함께 부가되며, 이에 의해 롤링 서클 증폭(RCA)이 일어나 주형 DNA (예컨대, "모노머")의 연쇄체(예컨대, "멀티머")를 형성한다. 상기 연쇄체는 상응하는 상보적 가닥을 합성하도록 처리된 다음, 어댑터가 부가되어 시퀀싱 라이브러리를 만든다. 이후에 표준 기술을 이용하여 시퀀싱되는 이러한 수득된 라이브러리는 일반적으로 3개의 종을 함유할 것이다: 회귀 서열 변이(예컨대, 돌연변이)를 함유하지 않는 nDNA("정상" DNA); 효소 시퀀싱 오류를 함유하는 nDNA, 및 증폭 전 샘플 폴리뉴클레오타이드에 이미 존재하였던 "실제" 또는 서열 변이의 멀티머를 함유하는 DNA. 실질적으로 회귀한 돌연변이의 다중 카피의 존재는 서열 변이의 검출 및 확인을 가능하게 한다.

도 2는 폴리뉴클레오타이드 원형화를 용이하게 하기 위해 어댑터를 부가하는 것을 제외하고 도 1과 유사한 전략을 묘사한다. 도 2는 또한 표적 특이적 프라이머의 사용을 보여준다.

도 3은 어댑터 프라이머가 증폭에 사용되는 것을 제외하면 도 2와 유사하다.

도 4는 원형화된 ccDNA의 형성과 연관된 3가지 구현예를 묘사한다. 상부에서, 단일 가닥 DNA(ssDNA)가 어댑터의 부재하에 원형화되는 반면, 중앙 도식은 어댑터의 사용을 묘사하고, 하부 도식은 2개의 어댑터 올리고 (각 말단에 상이한 서열을 생성함)를 이용하며, 두 어댑터에 혼성화하여 두 말단을 근접시키는 스플린트 (splint) 올리고를 추가로 포함할 수 있다.

도 5는 라이게이션을 위해 단일 가닥 DNA의 2개의 말단을 공간적으로 근접시키는 "분자 클램프"의 사용을 통해 특정 표적을 원형화시키는 구현예를 묘사한다.

도 6A 및 6B는 핵산의 블로킹된 말단을 이용하여 어댑터를 부가하는 2가지 도식을 묘사한다.

도 7A, 7B, 및 7C는 롤링 서클 증폭(RCA) 반응을 프라이밍(priming)하는 3가지 상이한 방식을 묘사한다. 도 7A는 표적 특이적 프라이머, 예컨대, 특정한 표적 유전자 또는 관심있는 표적 서열의 사용을 보여준다. 이것은 일반적으로 표적 서열만이 증폭되게 한다. 도 7B는 전체 지놈 증폭(WGA)을 수행하기 위한 무작위 프라이머의 사용을 묘사하며, 이는 일반적으로 모든 샘플 서열을 증폭시킬 것이며, 이는 이후에 처리 동안 생물정보학적으로 분류된다. 도 7C는 어댑터가 사용될 때 어댑터 프라이머의 사용을 묘사하며, 이는 또한 일반적인 비-표적-특이적 증폭을 야기한다.

도 8은 이중 가닥 DNA 원형화 및 증폭의 예를 묘사하며, 이로써 일 구현예에 따라 두 가닥이 증폭된다.

도 9A, 9B, 9C, 및 9D는 차후 시퀀싱을 위해 상보적 가닥 합성을 달성하는 다양한 도식을 묘사한다. 도 9A는 표적 가닥의 무작위 프라이밍의 사용 후 라이게이션을 묘사한다. 도 9B는 표적 가닥의 어댑터 프라이밍의 사용 후 유사하게 라이게이션을 묘사한다. 도 9C는 "루프" 어댑터의 사용을 묘사하며, 상기 어댑터는 상보적인 서열들의 2개의 구획을 가지며, 이로써 이들이 서로 혼성화하여 루프 (예컨대, 스템-루프 구조)를 만든다. 연쇄체의 말단에 라이게이션시, 상기 루프의 자유 말단은 상보적 가닥에 대한 프라이머로서 작용한다. 도 9D는 제2 가닥 합성을 달성하는 과분지형(hyper-branching) 무작위 프라이머의 사용을 보여준다.

도 10은 표적 서열의 모노머 내에 정렬될 때 서로 떨어져 배열되는 한 쌍의 프라이머 (예컨대, 두 방향으로 배열되지만 증폭될 도메인의 말단 상에 있지 않은, "백투백(back to back)"으로도 지칭됨)를 이용하여, 적어도 2 카피의 표적 핵산 서열을 함유하는 원형 폴리뉴클레오타이드 또는 가닥의 시퀀싱을 촉진하는 일 구현예에 따른 PCR 방법을 보여준다. 일부 구현예에서, 이들 프라이머 세트는 연쇄체가 형성된 후에 사용되어 앰플리콘이 표적 서열의 더 큰 멀티머, 예컨대, 다이버, 트리머 등이 되도록 촉진한다. 선택적으로, 상기 방법은 다이버보다 더 작은 앰플리콘을 제거하기 위해 크기 선별을 추가로 포함할 수 있다.

도 11은 짧은 산물(예컨대, 모노머)의 증폭이 불리하도록, 백투백(B2B) 프라이머가 "터치업(touch up)" PCR 단계와 함께 사용되는 구현예를 묘사한다. 이 경우에, 상기 프라이머는 2개의 도메인; 표적 서열에 혼성화하는 제1 도메인 (회색 또는 검정색 화살표) 및 원래의 표적 서열에 혼성화하지 않는 "유니버설 프라이머" 결합 도메인인 제2 도메인 (굵은 사각형; 종종 어댑터로서도 지칭됨)을 갖는다. 일부 구현예에서, PCR의 제1 라운드는 유전자 특이적 서열이 결합하도록 낮은 온도 어닐링 단계로 수행된다. 상기 낮은 온도 진행은 짧은 산물을 포함하는 다양한 길이의 PCR 산물을 야기한다. 적은 수의 라운드 후, 전체 프라이머, 두 개의 도메인의 혼성화가 유리하도록 상기 어닐링 온도는 상승되며; 묘사한 바와 같이 이들은 주형의 말단에서 발견되는 반면, 내부 결합은 덜 안정적이다. 따라서 더 짧은 산물은 더 낮은 온도 또는 단지 단일 도메인에서보다 두 도메인으로 더 높은 온도에서 분리하다.

도 12A 및 12B는 2개의 상이한 시퀀싱 라이브러리 구축 방법을 묘사한다. 도 12A는 Illumina® Nextera 샘플 제조 시스템의 예를 예시하며, 상기 시스템에 의해 DNA는 단일 단계에서 동시에 단편화되고 시퀀싱 어댑터로 태그될 수 있다. 도 12B에서, 연쇄체는 초음파 처리에 의해 단편화된 후, 양 말단에 어댑터를 부가하고 (예컨대, KAPA Biosystem에 의해 키트를 이용하여), PCR 증폭한다. 다른 방법이 이용가능하다.

도 13A-C는 전통적인 PCR 프라이머 디자인 대비 백투백(B2B) 프라이머 디자인의 예시적인 이점의 예시를 제공한다. 전통적인 PCR 프라이머 디자인(좌측)은 돌연변이를 위한 핫스팟(검은색 별)일 수 있는 표적 서열의 측면에 위치한 영역에 프라이머(화살표, A 및 B)를 두며, 이들은 전형적으로 적어도 60 염기쌍 (bp) 떨어져 있어서 약 100bp의 전형적인 풋프린트(footprint)를 야기한다. 이 예시에서, B2B 프라이머 디자인(우측)은 프라이머를 표적 서열의 한 쪽에 둔다. 상기 2개의 B2B 프라이머는 반대 방향을 마주보고 있으며, 중첩될 수 있다 (예컨대, 약 12bp, 10bp, 5bp 이하, 또는 그 미만). B2B 프라이머 길이에 따라, 이 예시에서 총 풋프린트는 28-50bp 사이

일 수 있다. 더 큰 풋프린트로 인해, 단편화 이벤트가 전통적인 설계에서 프라이머 결합을 파괴할 가능성이 많으며, 이는 선형 단편(13A), 원형화된 DNA(13B), 또는 증폭 산물(13C)이건 간에, 서열 정보의 손실을 초래한다. 더욱이, 도 13C에서 예시된 바와 같이, B2B 프라이머 디자인은 상이한 폴리뉴클레오타이드를 구별하는데 사용될 수 있는 접합부 서열("천연 바코드"로도 지칭됨)을 포획한다.

도 14는 일 구현예 (예컨대, "네블라(Nebula)"로 지칭된, 원형화된 폴리뉴클레오타이드를 이용한 공정의 예시적인 구현)에 따라, 서열 변이를 검출하기 위한 주형을 생성하는 방법을 예시한다. 입력된 DNA는 ssDNA로 변성되고, 라이게이션에 의해 원형화되고, 비-원형화된 DNA는 엑소뉴클레아제 절단에 의해 분해된다. 라이게이션 효율은 입력 DNA 및 원형화된 DNA 양을 비교하는 정량적 PCR(qPCR)에 의해 정량되며, 이는 전형적으로 적어도 약 80%의 라이게이션 효율을 낳는다. 원형화된 DNA는 정제되어 버퍼를 교환한 후, 무작위 프라이머 및 Phi29 폴리머라아제를 이용하여 전체 지놈 증폭(WGA)을 수행한다. WGA 산물은 정제되고, 산물은 약 400bp 또는 그 미만의 짧은 단편으로 단편화된다 (예컨대, 초음파 처리에 의해). 증폭된 DNA의 온-표적 비율(on-target rate)이 동일한 양의 참조 지놈 DNA를 증폭된 DNA에 비교하는 qPCR에 의해 정량되며, 이는 전형적으로 약 95% 또는 초과와 평균 온-표적 비율을 나타낸다.

도 15는 테일을 가진 B2B 프라이머를 이용한 증폭의 추가 구현, 및 더 높은 온도에서 PCR의 "터치업" 제2 과정을 구현하는 것을 예시한다. B2B 프라이머는 서열-특이적 영역(진한 검정색 선) 및 어댑터 서열(개방형 박스)를 함유한다. 더 낮은, 제1 과정 어닐링 온도를 이용하여, 표적-특이적 서열은 주형에 어닐링하여 초기 모노머를 생성하며, PCR 산물은 직렬(tandem) 반복을 함유한다 (15A). 더 높은 온도에서의 제2 증폭 과정에서, 표적-특이적 및 어댑터 서열 혼성화 모두가 표적-특이적 서열 혼성화 단독보다 유리하며, 이는 짧은 산물이 우선적으로 생산되는 정도를 낮춘다 (15B). 전체 프라이머를 선호하지 않으면서, 표적-특이적 서열과의 내부 어닐링은 모노머의 비율을 빠르게 증가시킨다 (15C, 좌측).

도 16은 변이로서 계수될 2개의 상이한 폴리뉴클레오타이드(예컨대, 상이한 접합부에 의해 확인됨) 상에서 발생하는 것을 요구하는 (하부 선) 및 이를 요구하지 않는 (상부 선) Q30 필터를 이용한 표적 시퀀싱 방법에 의해 검출된 배경 노이즈 (변이의 빈도) 간의 비교를 예시한다. 이러한 검증 필터의 적용은 본원에서 "반딧불(Firefly)"로도 지칭된다. 인간의 지놈 DNA(12878, Coriell Institute)가 100-200bp로 단편화되었고, 이는 공지된 SNP(CYP2C19)를 함유하는 지놈 DNA(19240, Coriell Institute)의 2% 급등(spike-in)을 포함하였다. 상기 진정된 변이 신호(표시된 피크)는 배경을 유의미하게 초과하지 않았다 (상부, 열은 회색 플롯). 검증 필터를 적용함으로써 배경 노이즈는 약 0.1로 감소되었다 (하부, 검정색 플롯).

도 17은 폴리뉴클레오타이드의 집단에서 다양한 저빈도(2%, 0.2%, 및 0.02%)에서 급증한 서열 변이의 검출을 예시하며, 이는 그림에도 불구하고 본 개시내용의 방법을 적용할 때 배경을 유의미하게 초과한다.

도 18은 본 개시내용의 구현예의 라이게이션 효율 및 온-표적 비율의 분석 결과를 예시한다.

도 19는 본 개시내용의 일 구현예에 따른 방법에서, 대립유전자 빈도의 보존, 및 편향(bias)의 실질적인 부재를 예시한다.

도 20은 일 구현예에 따라, 작은 입력 샘플에서의 서열 변이의 검출 결과를 예시한다.

도 21은 표준 시퀀싱 방법에 따라, 서열 차이가 2개의 상이한 폴리뉴클레오타이드 상에서 발생하는 것을 요구하지 않으면서 수득된 서열 변이의 검출 결과에서 높은 배경의 예를 예시한다.

도 22는 본 개시내용의 일 구현예에 따른 방법에 따라 생산된 시퀀싱 결과 ("네블라-반딧불"; 좌측), 대안적인 시퀀싱 라이브러리 구축 키트(루비콘, Rubicon Genomics; 중간), 및 일반적으로 32ng에 대해 문헌에 보고된 바와 같은 무세포 DNA(cfDNA)(우측)를 이용한 시퀀싱 결과의 GC 함량 분포 및 상기 지놈의 GC 함량 분포 간의 비교를 예시하는 그래프를 제공한다.

도 23은 일 구현예에 따른 방법의 시퀀싱 리드로부터 수득된 입력 DNA의 크기 분포를 예시하는 그래프를 제공한다.

도 24는 일 구현예에 따른 무작위-프라이밍 방법에 의해 다수의 표적에 걸친 균일한 증폭을 예시하는 그래프를 제공한다.

도 25는 원형화의 부재시, 확인가능한 접합부를 갖는 폴리뉴클레오타이드 멀티머의 형성을 위한 구현예를 예시한다. 폴리뉴클레오타이드(예컨대, 폴리뉴클레오타이드 단편, 또는 무세포 DNA)는 연결되어 본 개시내용의 구현예에 따라 독립적인 폴리뉴클레오타이드를 구별하는데 유용한 비-천연 접합부를 갖는 멀티머를 형성한다 (본원

에서 "오토-태그"로도 지칭됨). 도 25A에서, 폴리뉴클레오타이드는 평활 말단(blunt-end) 라이게이션에 의해 서로 직접 연결된다. 도 25B에서, 폴리뉴클레오타이드는 바코드 서열을 추가로 포함할 수 있는, 하나 이상의 개입 어댑터 올리고뉴클레오타이드를 통해 연결된다. 그리고 나서, 멀티머는 임의의 다양한 방법에 의해, 예컨대 무작위 프라이머 (전체 지놈 증폭), 어댑터 프라이머, 또는 하나 이상의 표적 특이적 프라이머 또는 프라이머 쌍에 의해, 증폭된다. 복수의 별개의 폴리뉴클레오타이드로부터 확인가능한될 수 있는 접합부를 갖는 멀티머를 형성하는 공정은 본원에서 "이클립스(Eclipse)"로도 지칭된다.

도 26은 도 25의 공정에 대한 예시적인 변이를 예시한다. 폴리뉴클레오타이드 (예컨대, cfDNA, 또는 다른 폴리뉴클레오타이드 단편)은 말단-복구되고, A-테일을 가지며, 어댑터 라이게이션된다 (예컨대, KAPA Bio시스템의 키트와 같은 표준 키트를 이용하여). 총 DNA 입력을 원하는 수준으로 (예컨대, 약 20ng 또는 그 초과) 증가시키기 위해 내부 우라실(U)로 표지된 캐리어 DNA가 보충될 수 있다. 검출될 서열 변이는 "별"로 표시되어 있다. 라이게이션이 완료되는 경우, 캐리어 DNA는 우라실 DNA 글리코실라아제(UDG) 및 DNA 글리코실라아제-리아제 엔도뉴클레아제 VIII의 혼합물인 우라실-특이적 절제 시약 (사용자) 효소의 부가에 의해 분해될 수 있다. 산물은 캐리어 DNA의 단편을 제거하기 위해 정제된다. 정제된 산물은 증폭된다 (예컨대, 어댑터 서열에 대한 프라이머를 이용한 PCR에 의해). 어느 잔존 캐리어 DNA는 분해, 및 적어도 하나의 말단 상의 어댑터로부터의 분리로 인해 증폭되지 않을 가능성이 있다. 증폭된 산물은 짧은 DNA 단편을 제거하기 위해 정제될 수 있다.

도 27은 도 25의 공정에 대한 예시적인 변이를 예시한다. 표적 특이적 증폭 프라이머는 어댑터로서 기능하는 공통적인 5' "테일"을 포함한다 (회색 화살표). 초기 증폭 (예컨대, PCR에 의함)은 몇 개의 사이클 (예컨대, 적어도 약 5, 10, 또는 그 이상의 사이클) 동안 진행된다. PCR 산물은 또한 프라이머로서 작용할 수 있으며, 다른 PCR 산물에 어닐링(예컨대, 어닐링 온도가 제2 과정에서 감소될 때)하여 확인가능한 접합부를 갖는 연쇄체를 생성한다. 제2 과정은 많은 사이클 (예컨대, 5, 10, 15, 20, 또는 그 이상의 사이클)을 포함할 수 있고, 연쇄체 형성 및 증폭에 유리한 조건의 선택 또는 변형을 포함할 수 있다. 이 도식에 따른 방법은 "릴레이 앰프 세크(Relay Amp Seq)"로도 지칭되며, 이는 구분된 환경에서 (예컨대, 드롭렛에서) 유용할 수 있다.

도 28A-E는 폴리뉴클레오타이드를 원형화시키는 방법의 비제한적인 예를 예시한다. 도 28A에서, 이중 가닥 폴리뉴클레오타이드(예컨대, dsDNA)는 단일 가닥으로 변성된 후, 곧바로 원형화(예컨대, CircLigase에 의한 자가-연결 라이게이션)된다. 도 28B에서, 라이게이션 효율을 개선하기 위해 폴리뉴클레오타이드 (예컨대, DNA 단편)는 말단-복구되고 A-테일을 가진 (아데노신의 3' 말단에서의 단일-염기 연장) 후, 단일 가닥으로 변성되고, 및 원형화된다. 도 28C에서, 폴리뉴클레오타이드는 말단-복구되고 A-테일을 가지며 (이중 가닥인 경우), 티민 (T) 연장을 갖는 어댑터에 연결되고, 단일 가닥으로 변성되며, 원형화된다. 도 28D에서, 폴리뉴클레오타이드는 말단-복구되고 A-테일을 가지며 (이중 가닥인 경우), 양 말단이 3개의 요소 (라이게이션을 위한 T 연장, 어댑터 간의 상동성, 및 3' 테일)를 갖는 어댑터에 라이게이션되고, 가닥이 변성되며, 단일 가닥 폴리뉴클레오타이드가 원형화된다 (어댑터 서열 간의 상동성에 의해 촉진됨). 도 28E에서, 이중 가닥 폴리뉴클레오타이드는 단일 가닥 형태로 변성되며, 폴리뉴클레오타이드의 말단을 근접시켜 연결을 용이하게 하는 분자 클램프의 존재시에 원형화된다.

도 29는 특히 원형화된 폴리뉴클레오타이드에 대하여, 본 개시내용의 방법에 따라 서열 변이를 확인하기 위한 증폭 시스템의 예시적인 작업흐름 디자인을 예시한다.

도 30은 특히 원형화 단계 없는 선형 폴리뉴클레오타이드 입력에 대하여, 본 개시내용의 방법에 따라 서열 변이를 확인하기 위한 증폭 시스템의 예시적인 작업흐름 디자인을 예시한다.

도 31은 본 개시내용의 방법에 따라 서열 변이를 확인하는 예시적인 작업흐름의 요약 예시를 제공한다. "이클립스(Eclipse)" (선형 폴리뉴클레오타이드 분석) 분지(branch)를 따라, 분석은 디지털 PCR (예컨대, 디지털 드롭렛 PCR, ddPCR), 실시간 PCR, 접합부 서열(오토 태그)의 분석을 수반한 프로브 포획 (포획 서열)에 의한 농축, 삽입된 어댑터 서열 (바코드된 삽입)에 기초한 시퀀싱, 또는 릴레이 앰프 시퀀싱을 포함할 수 있다. "네블라(Nebula)" (원형화된 폴리뉴클레오타이드 분석)를 따라, 분석은 디지털 PCR (예컨대, 디지털 드롭렛 PCR, ddPCR), 실시간 PCR, 접합부 서열 (천연 바코드)의 분석을 수반한 프로브 포획 (포획 서열)에 의한 농축, 프로브 포획 또는 표적화된 증폭 (예컨대, B2B 증폭)에 의한 농축, 및 2개의 상이한 폴리뉴클레오타이드(예컨대, 상이한 접합부를 갖는 폴리뉴클레오타이드)에서 발생하는 차이로서 서열 변이를 확인하는 검증 단계를 갖는 서열 분석을 포함할 수 있다.

도 32는 일 구현예에 따른 시스템의 예시이다.

도 33은 일 예에 따른 표적 영역을 따라 포획 및 커버리지의 효율을 예시한다. 표적화 염기의 >90%가 20x 넘게 커버되며, 표적화된 염기의 >50%가 >50x 커버리지를 갖는다.

발명을 실시하기 위한 구체적인 내용

- [0023] 본원에 개시된 일부 구현예의 실행은, 달리 표시하지 않는한, 당해 기술에 속하는 면역학, 생화학, 화학, 분자 생물학, 미생물학, 생물학, 지놈 및 재조합 DNA의 종래의 기술을 이용한다. 예를 들면 문헌[Sambrook and Green, Molecular Cloning: A Laboratory Manual, 4th Edition (2012); the series Current Protocols in Molecular Biology (F. M. Ausubel, et al. eds.); the series Methods In Enzymology (Academic Press, Inc.), PCR 2: A Practical Approach (M.J. MacPherson, B.D. Hames and G.R. Taylor eds. (1995)), Harlow and Lane, eds. (1988) Antibodies, A Laboratory Manual, and Culture of Animal Cells: A Manual of Basic Technique and Specialized Applications, 6th Edition (R.I. Freshney, ed. (2010))]을 참고한다.
- [0024] 용어 "약" 또는 "대략"은 당해 분야에서 통상의 기술을 가진 자에 의해 결정되는 바와 같은 특정한 값에 대한 허용가능한 오차 범위 내를 의미하며, 이는 상기 값이 측정되거나 결정되는 방법, 즉 상기 측정 시스템의 한계에 일부 좌우될 것이다. 예를 들면, "약"은 본 기술분야에서의 관행에 따라, 1 또는 1 초과의 표준 편차를 의미할 수 있다. 대안적으로, "약"은 주어진 값의 최대 20%, 최대 10%, 최대 5%, 또는 최대 1%의 범위를 의미할 수 있다. 대안적으로, 특히 생물학적 시스템 또는 공정과 관련하여, 상기 용어는 하나의 값의 한자리수 이내, 바람직하게는 5배 이내, 및 더 바람직하게는 2배 이내를 의미할 수 있다. 본원 및 청구항에 특정한 값이 기재되는 경우, 달리 언급되지 않는 한, 특정한 값의 허용가능한 오차 범위 내를 의미하는 용어 "약"이 추정되어야 한다.
- [0025] 용어 "폴리뉴클레오타이드", "뉴클레오타이드", "뉴클레오타이드 서열", "핵산" 및 "올리고뉴클레오타이드"는 상호교환적으로 사용된다. 이들은 어느 길이의 뉴클레오타이드, 데옥시리보뉴클레오타이드 또는 리보뉴클레오타이드 중 하나, 또는 이의 유사체의 중합체 형태를 가리킨다. 폴리뉴클레오타이드는 어떤 3차원 구조를 가질 수 있고, 공지되거나 공지되지 않은 임의의 기능을 수행할 수 있다. 하기는 폴리뉴클레오타이드의 비제한적인 예이다: 유전자 또는 유전자 단편의 코딩 또는 비코딩 영역, 연결 분석으로부터 정의된 유전자좌(locus), 엑손, 인트론, 메신저 RNA(mRNA), 운반 RNA(tRNA), 리보솜 RNA(rRNA), 짧은 간섭 RNA(siRNA), 짧은-헤어핀 RNA(shRNA), 마이크로-RNA(miRNA), 리보자임, cDNA, 재조합 폴리뉴클레오타이드, 분지형 폴리뉴클레오타이드, 플라스미드, 벡터, 어떤 서열의 단리된 DNA, 어떤 서열의 단리된 RNA, 핵산 프로브, 및 프라이머. 폴리뉴클레오타이드는 하나 이상의 변형된 뉴클레오타이드, 예컨대 메틸화된 뉴클레오타이드 및 뉴클레오타이드 유사체를 포함할 수 있다. 존재하는 경우, 뉴클레오타이드 구조에 대한 변형은 폴리머의 조립 전 또는 후에 부여될 수 있다. 뉴클레오타이드의 서열은 비-뉴클레오타이드 구성요소에 의해 중단될 수 있다. 폴리뉴클레오타이드는, 예컨대 표지 구성요소와의 접합에 의해서, 중합 후에 추가 변형될 수 있다.
- [0026] 일반적으로, 용어 "표적 폴리뉴클레오타이드"는 존재, 양, 및/또는 뉴클레오타이드 서열, 또는 이들 중 하나 이상에서의 변화가 결정되기를 원하는 표적 서열을 갖는 핵산 분자의 출발 집단 내의 핵산 분자 또는 폴리뉴클레오타이드를 가리킨다. 일반적으로, 용어 "표적 서열"은 단일 가닥의 핵산 상의 핵산 서열을 가리킨다. 상기 표적 서열은 유전자, 조절 서열, 지놈 DNA, cDNA, mRNA, miRNA, rRNA를 포함하는 RNA 등의 일부일 수 있다. 표적 서열은 샘플로부터의 표적 서열 또는 증폭 반응의 산물과 같은 제2 표적일 수 있다.
- [0027] 일반적으로, "뉴클레오타이드 프로브", "프로브" 또는 "태그 올리고뉴클레오타이드"는 상응하는 표적 서열과의 혼성화에 의한 혼성화 반응에서 그의 상응하는 표적 폴리뉴클레오타이드를 검출하거나 확인하는데 사용되는 폴리뉴클레오타이드를 가리킨다. 따라서, 뉴클레오타이드 프로브는 하나 이상의 표적 폴리뉴클레오타이드에 혼성화가능하다. 태그 올리고뉴클레오타이드는 샘플 내의 하나 이상의 표적 폴리뉴클레오타이드에 완전히 상보적이거나, 또는 샘플 내의 하나 이상의 표적 폴리뉴클레오타이드에서 상응하는 뉴클레오타이드에 의해 상보되지 않는 하나 이상의 뉴클레오타이드를 함유할 수 있다.
- [0028] "혼성화"는 하나 이상의 폴리뉴클레오타이드가 반응하여 뉴클레오타이드 잔기의 염기 사이에 수소 결합을 통해 안정화되는 복합체를 형성하는 반응을 가리킨다. 수소 결합은 왓슨 크릭 염기 쌍, 후그스테인 결합, 또는 염기 상동성에 따른 어느 다른 서열 특이적 방식에 의해 일어날 수 있다. 상기 복합체는 이중체 구조를 형성하는 2개의 가닥, 다중 가닥 복합체를 형성하는 3개 이상의 가닥, 단일 자가-혼성화 가닥, 또는 이들의 어느 조합을 포함할 수 있다. 혼성화 반응은 PCR의 개시, 또는 엔도뉴클레아제에 의한 폴리뉴클레오타이드의 효소적 절단과 같은, 보다 광범위한 공정에서의 하나의 단계를 구성할 수 있다. 제1 서열에 상보적인 제2 서열은 제1 서열의 "보

체"로서 지칭된다. 폴리뉴클레오타이드에 적용되는 용어 "혼성화가능한"은 혼성화 반응에서 뉴클레오타이드 잔기의 염기 사이에 수소 결합을 통해 안정화되는 복합체를 형성하는 폴리뉴클레오타이드의 능력을 가리킨다.

[0029] "상보성(complementarity)"은 전통적인 왓슨-크릭 또는 다른 비-전통적인 유형에 의해 또 다른 핵산 서열과 수소 결합(들)을 형성하는 핵산의 능력을 가리킨다. 퍼센트 상동성은 제2 핵산 서열과 수소 결합 (예컨대, 왓슨-크릭 염기 쌍)을 형성하는 핵산 분자 내의 잔기의 백분율 (예컨대, 10개 중 5, 6, 7, 8, 9, 10개는 각각 50%, 60%, 70%, 80%, 90%, 및 100% 상보성임)을 가리킨다. "완전히 상보적"은 핵산 서열의 모든 인접한 잔기가 제2 핵산 서열 내의 동일한 수의 인접한 잔기와 수소 결합할 것임을 의미한다. 본원에 사용된 바와 같이 "실질적으로 상보적"은 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 30, 35, 40, 45, 50, 또는 그 이상의 뉴클레오타이드의 영역에 대해 적어도 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 97%, 98%, 99%, 또는 100%인 상동성의 정도를 가리키거나, 또는 엄격한 조건 하에 혼성화하는 2개의 핵산을 가리킨다. 예컨대 퍼센트 상동성을 평가하기 위한 서열 동일성은 비제한적으로 니들만-분쉬(Needleman-Wunsch) 알고리즘(예컨대, 선택적으로 디폴트 설정을 갖는, www.ebi.ac.uk/Tools/psa/emboss_needle/nucleotide.html에서 이용가능한, EMBOSS Needle 정렬기 참고), BLAST 알고리즘(예컨대, 선택적으로 디폴트 설정을 갖는, blast.ncbi.nlm.nih.gov/Blast.cgi에서 이용가능한 BLAST 정렬 도구), 또는 스미스-워터먼 알고리즘(예컨대, 선택적으로 디폴트 설정을 갖는, www.ebi.ac.uk/Tools/psa/emboss_water/nucleotide.html에서 이용가능한 EMBOSS Water 정렬기)를 포함하는, 어느 적합한 정렬 알고리즘에 의해 측정될 수 있다. 최적 정렬은 디폴트 파라미터를 포함하는, 선택된 알고리즘의 어느 적합한 파라미터를 이용하여 평가될 수 있다.

[0030] 일반적으로, 혼성화를 위한 "엄격한 조건"은 표적 서열에 상동성을 갖는 핵산이 대부분 표적 서열과 혼성화하고, 비-표적 서열에는 실질적으로 혼성화하지 않는 조건을 가리킨다. 엄격한 조건은 일반적으로 서열-의존적이며, 많은 인자에 따라 달라진다. 일반적으로, 서열이 더 길수록, 서열이 그의 표적 서열에 특이적으로 혼성화하는 온도는 더 높다. 엄격한 조건의 비제한적인 예는 문헌[Tijssen (1993), Laboratory Techniques In Biochemistry and Molecular Biology-Hybridization With Nucleic Probes Part I, Second Chapter "Overview of principles of hybridization and the strategy of nucleic acid probe assay", Elsevier, N.Y.]에 상세히 기재되어 있다.

[0031] 하나의 양태에서, 상기 개시내용은, 예컨대 핵산 샘플에서 서열 변이를 확인하는 방법을 제공한다. 일부 구현예에서, 상기 복수의 폴리뉴클레오타이드 중 각각의 폴리뉴클레오타이드는 5' 말단 및 3' 말단을 가지고, 상기 방법은 (a) 상기 복수의 폴리뉴클레오타이드 중 각각의 폴리뉴클레오타이드를 원형화시켜 복수의 원형 폴리뉴클레오타이드를 형성하는 단계로서, 이들 각각은 5' 말단과 3' 말단 사이에 접합부를 갖고; (b) 상기 (a)의 원형 폴리뉴클레오타이드를 증폭시키는 단계; (c) 상기 증폭된 폴리뉴클레오타이드를 시퀀싱하여 복수의 시퀀싱 리드를 생성하는 단계; (d) 시퀀싱 리드와 참조 서열 간의 서열 차이를 확인하는 단계; 및 (e) 상이한 접합부를 갖는 적어도 2개의 원형 폴리뉴클레오타이드에서 발생하는 서열 차이를 서열 변이로서 판정하는 단계를 포함한다. 일부 구현예에서, 상기 방법은 시퀀싱 리드와 참조 서열 간의 서열 차이를 확인하는 것, 및 상이한 접합부를 갖는 적어도 2개의 원형 폴리뉴클레오타이드에서 발생하는 서열 차이를 서열 변이로서 판정하는 것을 포함하며, (a) 상기 시퀀싱 리드는 상기 적어도 2개의 원형 폴리뉴클레오타이드의 증폭 산물에 상응하며; (b) 상기 적어도 2개의 원형 폴리뉴클레오타이드 각각은 각각의 폴리뉴클레오타이드의 5' 말단 및 3' 말단을 라이게이션함으로써 형성된 상이한 접합부를 포함한다.

[0032] 일반적으로, 용어 "서열 변이"는 하나 이상의 참조 서열에 대비 서열에서의 임의의 변이를 가리킨다. 전형적으로, 상기 서열 변이는 참조 서열이 공지된 개체들의 주어진 집단에 대한 참조 서열보다 더 낮은 빈도로 일어난다. 예를 들면, 특정한 박테리아 속은 16S rRNA 유전자에 대해 공통 참조 서열을 가질 수 있지만, 상기 속에 속하는 각각의 종은 상기 유전자 (또는 이의 일부) 내에 하나 이상의 서열 변이를 가질 수 있으며, 이는 박테리아의 집단에서 상기 종을 확인하는데 유용하다. 추가 예로서, 동일한 종의 다수의 개체에 대한 서열 (또는 동일한 개체에 대한 다수의 시퀀싱 리드)은 최적으로 정렬될 때 컨센서스 서열을 생성할 수 있고, 상기 컨센서스 서열에 대한 서열 변이는 위험한 오염을 나타내는 집단 내 돌연변이를 확인하는데 사용될 수 있다. 일반적으로, "컨센서스 서열"은 일련의 관련된 핵산들이 집약적인 수학적 및/또는 서열 분석, 예컨대 임의의 다양한 서열 정렬 알고리즘에 따른 최적 서열 정렬을 거친 경우, 서열 내의 각 위치에 있는 염기의 가장 공통된 선택을 반영하는 뉴클레오타이드 서열을 가리킨다. 다양한 정렬 알고리즘이 이용가능하며, 이들 중 일부는 본원에 기재되어 있다. 일부 구현예에서, 참조 서열은 하나의 공지된 참조 서열, 예컨대 단일 개체의 지놈 서열이다. 일부 구현예에서, 참조 서열은 다수의 공지된 서열들, 예컨대 참조 집단으로서 역할을 하는 다수의 개체의 지놈 서열, 또는 동일한 개체로부터의 폴리뉴클레오타이드의 다수의 시퀀싱 리드를 정렬함으로써 형성된 컨센서스 서열이다.

일부 구현예에서, 참조 서열은 분석 중인 샘플로부터의 서열들을 최적 정렬함으로써 형성된 콘센서스 서열이며, 이로써 서열 변이는 동일한 샘플에서 상응하는 서열 대비 변이에 해당한다. 일부 구현예에서, 서열 변이는 상기 집단에서 낮은 빈도로 발생한다 ("희귀" 서열 변이로도 지칭됨). 예를 들면, 서열 변이는 약 5%, 4%, 3%, 2%, 1.5%, 1%, 0.75%, 0.5%, 0.25%, 0.1%, 0.075%, 0.05%, 0.04%, 0.03%, 0.02%, 0.01%, 0.005%, 0.001% 이하, 또는 그 미만의 빈도로 일어날 수 있다. 일부 구현예에서, 서열 변이는 약 0.1% 또는 그 미만의 빈도로 발생한다.

[0033] 서열 변이는 참조 서열에 대해 어떠한 변이일 수 있다. 서열 변이는 단일 뉴클레오타이드 또는 복수의 뉴클레오타이드(예컨대, 2, 3, 4, 5, 6, 7, 8, 9, 10, 또는 그 이상의 뉴클레오타이드)에서의 변화, 이의 삽입, 또는 결실로 이루어질 수 있다. 서열 변이가 2개 이상의 뉴클레오타이드 차이를 포함하는 경우, 상이한 뉴클레오타이드들은 서로 인접하거나, 또는 비연속적일 수 있다. 서열 변이의 유형의 비제한적인 예는 단일 뉴클레오타이드 다형(SNP), 결실/삽입 다형(DIP), 카피수 변이(CNV), 짧은 직렬 반복(STR), 단순 서열 반복(SSR), 다양한 수의 직렬 반복(VNTR), 증폭된 단편 길이 다형(AFLP), 레트로트랜스포존-기반의 삽입 다형, 서열 특이적 증폭된 다형, 및 서열 변이로서 검출될 수 있는 후성유전학적 표지에서의 차이(예컨대, 메틸화 차이)를 포함한다.

[0034] 본원에 기재된 방법을 거칠 수 있는 핵산 샘플은 임의의 적합한 공급원으로부터 유래될 수 있다. 일부 구현예에서, 사용되는 샘플은 환경 샘플이다. 환경 샘플은 어느 환경 공급원, 예를 들면, 천연발생 또는 인공 공기, 수계, 토양, 또는 관심있는 어느 다른 샘플로부터 유래될 수 있다. 일부 구현예에서, 환경적 샘플은, 예를 들면, 공기 병원균 수집 시스템, 표면하 침전물, 지하수, 땅속 깊숙한 곳의 오래된 물, 초지의 식물 뿌리-토양 경계, 연안수 및 하수 처리장으로부터 수득될 수 있다.

[0035] 샘플로부터의 폴리뉴클레오타이드는 비제한적으로, DNA, RNA, 리보솜 RNA(rRNA), 운반 RNA(tRNA), 마이크로 RNA(miRNA), 메신저 RNA(mRNA), 이들 중 어느 것의 단편, 또는 이들 중 어느 2개 이상의 조합을 포함하는, 임의의 다양한 폴리뉴클레오타이드일 수 있다. 일부 구현예에서, 샘플은 DNA를 포함한다. 일부 구현예에서, 샘플은 지놈 DNA를 포함한다. 일부 구현예에서, 샘플은 미토콘드리아 DNA, 엽록체 DNA, 플라스미드 DNA, 박테리아성 인공 염색체, 효모 인공 염색체, 올리고뉴클레오타이드 태그, 또는 이의 조합을 포함한다. 일부 구현예에서, 샘플은 증폭에 의해, 예컨대 비제한적으로 폴리머라아제 연쇄 반응(PCR), 역전사, 및 이의 조합을 포함하는, 프라이머들의 임의의 적합한 조합 및 DNA 폴리머라아제를 이용한 프라이머 연장 반응에 의해 생성된 DNA를 포함한다. 프라이머 연장 반응을 위한 주형이 RNA인 경우, 역전사의 산물은 상보적 DNA(cDNA)로 지칭된다. 프라이머 연장 반응에서 유용한 프라이머는 하나 이상의 표적에 특이적인 서열, 무작위 서열, 부분적 무작위 서열, 및 이의 조합을 포함할 수 있다. 일반적으로, 샘플 폴리뉴클레오타이드는 샘플 내에 존재하는 어느 폴리뉴클레오타이드를 포함하며, 이는 표적 폴리뉴클레오타이드를 포함하거나 포함하지 않을 수 있다. 상기 폴리뉴클레오타이드는 단일 가닥, 이중 가닥, 또는 이들의 조합일 수 있다. 일부 구현예에서, 본 개시내용에 거친 폴리뉴클레오타이드는 단일 가닥 폴리뉴클레오타이드이며, 이는 이중 가닥 폴리뉴클레오타이드가 존재하거나 존재하지 않을 수 있다. 일부 구현예에서, 폴리뉴클레오타이드는 단일 가닥 DNA이다. 단일 가닥 DNA(ssDNA)는 단일 가닥 형태로 단리된 ssDNA, 또는 이중 가닥 형태로 단리된 후 본 개시내용의 방법에서 하나 이상의 단계의 목적을 위해 단일 가닥으로 만들어진 DNA일 수 있다.

[0036] 일부 구현예에서, 폴리뉴클레오타이드는 추출 단계 없이, 및/또는 정제 단계 없이 차후의 단계(예컨대, 원형화 및 증폭)를 거친다. 예를 들면, 유체 샘플은 추출 단계 없이 세포를 제거하도록 처리되어 정제된 액체 샘플 및 세포 샘플을 생성한 후, 상기 정제된 유체 샘플로부터 DNA가 단리될 수 있다. 폴리뉴클레오타이드의 단리를 위한 다양한 과정, 예컨대 침전 또는 기질의 비특이적 결합 후 상기 기질을 세척하여 결합된 폴리뉴클레오타이드를 방출하는 것에 의한 것이 이용가능하다. 폴리뉴클레오타이드가 세포 추출 단계 없이 샘플로부터 단리되는 경우, 폴리뉴클레오타이드는 주로 세포외 또는 "무세포" 폴리뉴클레오타이드일 것이며, 이는 사멸하거나 손상된 세포에 상응할 수 있다. 그러한 세포의 동일성은 종양 세포 (예컨대, 암 검출에서), 태아 세포 (예컨대, 산전 진단에서), 이식된 조직으로부터의 세포 (예컨대, 이식 실패의 조기 검출에서), 또는 미생물의 군집의 구성원과 같이, 이들이 유래된 세포 또는 세포의 집단을 규명하는데 사용될 수 있다.

[0037] 샘플이, 예컨대 샘플 내의 세포로부터, 폴리뉴클레오타이드를 추출하기 위해 처리되는 경우, 다양한 추출 방법이 이용가능하다. 예를 들면, 핵산은 페놀, 페놀/클로로포름/이소아밀 알콜, 또는 TRIzol 및 TriReagent를 포함하는 유사한 제형을 이용한 유기 추출에 의해 정제될 수 있다. 추출 기술의 다른 비제한적인 예는: (1) 자동화 핵산 추출기, 예컨대, Applied Biosystems(캘리포니아주 포스터 시티)로부터 이용가능한 Model 341 DNA 추출기를 이용하거나, 이용하지 않고, 예컨대 페놀/클로로포름 유기 시약(Ausubel et al., 1993)을 이용하여, 유기 추출 후 에탄올 침전; (2) 고정상 흡착 방법(미국 특허 제5,234,809; Walsh et al., 1991); 및 (3) "염석" 방법으로 지칭되는, 염-유도 핵산 침전 방법(Miller et al., (1988))을 포함한다. 핵산 단리 및/또는 정제의 또 다

른 예는 핵산이 특이적으로 또는 비-특이적으로 결합할 수 있는 자기 입자를 사용한 후, 자석을 이용하여 상기 비드를 분리하고, 세척하고, 상기 비드로부터 핵산을 용리시키는 것을 포함한다 (예컨대, 미국 특허 제 5,705,628호 참조). 일부 구현예에서, 상기 분리 방법은 샘플로부터 원하지 않는 단백질의 제거를 돕기 위한 효소 절단 단계, 예컨대, 프로테이나아제 K, 또는 다른 유사 프로테아제를 이용한 절단 후에 수행될 수 있다. 예컨대, 미국 특허 제 7,001,724호를 참조한다. 원하는 경우, RNase 억제제가 용해 버퍼에 부가될 수 있다. 어떤 세포 또는 샘플 유형의 경우, 단백질 변성/분해 단계를 프로토콜에 부가하는 것이 바람직할 수 있다. 정제 방법은 DNA, RNA, 또는 이들 모두를 분리하는 것에 관한 것일 수 있다. 추출 과정 동안 또는 추출 과정 이후에 DNA 및 RNA가 함께 분리되는 경우, 다른 것들로부터 하나 또는 둘 모두를 분리하여 정제하기 위해 추가 단계가 이용될 수 있다. 추출된 핵산의 하위-분획이 또한, 예를 들면, 크기, 서열, 또는 다른 물리적 또는 화학적 특징에 의한 정제에 의해 생성될 수 있다. 초기 핵산 분리 단계 외에, 핵산의 정제는, 예컨대 과잉의 또는 원하지 않는 시약, 반응물, 또는 산물을 제거하기 위해, 상기 개시된 방법의 임의의 단계 이후에 수행될 수 있다. 샘플 내의 핵산의 양 및/또는 순도를 결정하는 다양한 방법, 예컨대 흡광도 (예컨대, 260 nm, 280 nm에서의 빛의 흡광도, 및 이들의 비율) 및 표지의 검출 (예컨대, 형광 염료 및 인터칼레이팅제, 예컨대 SYBR 그린, SYBR 블루, DAPI, 프로피듐 요오드, Hoechst 염색, SYBR 골드, 에티디움 브로마이드)에 의한 방법이 이용가능하다.

[0038] 원하는 경우, 샘플로부터의 폴리뉴클레오타이드는 추가 가공 전에 단편화될 수 있다. 단편화는 화학적, 효소적, 및 기계적 단편화를 포함하는, 임의의 다양한 방법에 의해 달성될 수 있다. 일부 구현예에서, 상기 단편은 약 10 내지 약 1,000 뉴클레오타이드 길이, 예컨대 10-800, 10-500, 50-500, 90-200, 또는 50-150 뉴클레오타이드의 평균 또는 중앙 길이를 갖는다. 일부 구현예에서, 상기 단편은 약 100, 200, 300, 500, 600, 800, 1000, 또는 1500 뉴클레오타이드 또는 그 미만의 평균 또는 중앙 길이를 갖는다. 일부 구현예에서, 상기 단편은 약 90-200 뉴클레오타이드의 범위이고/거나, 약 150 뉴클레오타이드의 평균 길이를 갖는다. 일부 구현예에서, 상기 단편화는 샘플 폴리뉴클레오타이드를 음향 초음파처리하는 것을 포함하여 기계적으로 달성된다. 일부 구현예에서, 상기 단편화는 하나 이상의 효소가 이중 가닥 핵산 파괴를 생성하는데 적합한 조건 하에 상기 샘플 폴리뉴클레오타이드를 하나 이상의 효소로 처리하는 것을 포함한다. 폴리뉴클레오타이드 단편의 생성에 유용한 효소의 예는 서열 특이적 및 비-서열 특이적 뉴클레아제를 포함한다. 뉴클레아제의 비제한적인 예는 DNase I, 단편화효소, 제한 엔도뉴클레아제, 이의 변이체, 및 이의 조합을 포함한다. 예를 들면, DNase I을 이용한 절단은 Mg^{++} 의 부재 및 Mn^{++} 의 존재하에 무작위 이중 가닥 파괴를 유도할 수 있다. 일부 구현예에서, 단편화는 샘플 폴리뉴클레오타이드를 하나 이상의 제한 엔도뉴클레아제로 처리하는 것을 포함한다. 단편화는 5' 오버행, 3' 오버행, 평활 말단, 또는 이의 조합을 갖는 단편을 생성할 수 있다. 일부 구현예에서, 예컨대 하나 이상의 제한 엔도뉴클레아제의 사용을 포함하는 경우, 샘플 폴리뉴클레오타이드의 절단은 예측가능한 서열을 갖는 오버행을 남긴다. 단편화된 폴리뉴클레오타이드는 컬럼 정제 또는 아가로스 겔로부터의 분리와 같은 표준 방법을 통해 단편을 크기 선별하는 단계를 거칠 수 있다.

[0039] 일부 구현예에 따르면, 샘플로부터의 상기 복수의 폴리뉴클레오타이드 중의 폴리뉴클레오타이드는 원형화된다. 원형화는 폴리뉴클레오타이드의 5' 말단을 동일한 폴리뉴클레오타이드의 3' 말단에, 샘플 내의 또 다른 폴리뉴클레오타이드의 3' 말단에, 또는 상이한 공급원으로부터의 폴리뉴클레오타이드(예컨대, 올리고뉴클레오타이드 어댑터와 같은, 인공 폴리뉴클레오타이드)의 3' 말단에 연결하는 것을 포함할 수 있다. 일부 구현예에서, 폴리뉴클레오타이드의 5' 말단은 동일한 폴리뉴클레오타이드의 3' 말단에 연결된다 ("자가-연결"으로도 지칭됨). 일부 구현예에서, 원형화 반응의 조건은 특정한 평균 길이의 원형화된 폴리뉴클레오타이드의 집단을 생성하기 위해, 특정한 범위의 길이 내의 폴리뉴클레오타이드의 자가-연결이 유리하도록 선택된다. 예를 들면, 원형화 반응 조건은 약 5000, 2500, 1000, 750, 500, 400, 300, 200, 150, 100, 50 이하의 뉴클레오타이드 길이보다 더 짧은 폴리뉴클레오타이드의 자가-연결이 유리하도록 선택된다. 일부 구현예에서, 50-5000 뉴클레오타이드, 100-2500 뉴클레오타이드, 또는 150-500 뉴클레오타이드의 길이를 갖는 단편이 유리하며, 이로써 원형화된 폴리뉴클레오타이드의 평균 길이가 상기 각각의 범위에 속한다. 일부 구현예에서, 상기 원형화된 단편의 80% 이상이 50-500 뉴클레오타이드 길이, 예컨대 50-200 뉴클레오타이드 길이이다. 최적화될 수 있는 반응 조건은 연결 반응을 위해 할당된 시간의 길이, 다양한 시약의 농도, 및 연결될 폴리뉴클레오타이드의 농도를 포함한다. 일부 구현예에서, 원형화 반응은 원형화 이전에 샘플 내에 존재하는 단편 길이의 분포를 보존시킨다. 예를 들면, 원형화 전 샘플에서 그리고 원형화된 폴리뉴클레오타이드에서 단편 길이의 평균, 중앙, 모드, 및 표준 편차 중 하나 이상은 또 다른 것의 75%, 80%, 85%, 90%, 95%, 또는 그 이상이다.

[0040] 우선적으로 자가-연결 원형화 산물을 형성하기보다는, 샘플 내의 폴리뉴클레오타이드의 5' 말단 및 3' 말단이 하나 이상의 개입 어댑터 올리고뉴클레오타이드에 의해 연결되어 원형 폴리뉴클레오타이드를 형성하도록, 하나 이상의 어댑터 올리고뉴클레오타이드가 사용된다. 예를 들면, 폴리뉴클레오타이드의 5' 말단은 어댑터의 3' 말

단에 연결될 수 있고, 동일한 어댑터의 5' 말단은 동일한 폴리뉴클레오타이드의 3' 말단에 연결될 수 있다. 어댑터 올리고뉴클레오타이드는 샘플 폴리뉴클레오타이드에 연결될 수 있는, 적어도 일부가 공지된 서열을 갖는 임의의 올리고뉴클레오타이드를 포함한다. 어댑터 올리고뉴클레오타이드는 DNA, RNA, 뉴클레오타이드 유사체, 비정규(non-canonical) 뉴클레오타이드, 표지된 뉴클레오타이드, 변형된 뉴클레오타이드, 또는 이의 조합을 포함할 수 있다. 어댑터 올리고뉴클레오타이드는 단일 가닥, 이중 가닥, 또는 부분적 이중체일 수 있다. 일반적으로, 부분적 이중체 어댑터는 하나 이상의 단일 가닥 영역 및 하나 이상의 이중 가닥 영역을 포함한다. 이중 가닥 어댑터는 서로 혼성화된 2개의 별개의 올리고뉴클레오타이드를 포함할 수 있고 ("올리고뉴클레오타이드 이중체"로도 지칭됨), 혼성화는 하나 이상의 평활 말단, 하나 이상의 3' 오버행, 하나 이상의 5' 오버행, 미스매치된 및/또는 쌍을 이루지 않은 뉴클레오타이드로부터 비롯된 하나 이상의 불룩한 것(bulge), 또는 이들의 임의의 조합을 포함할 수 있다. 어댑터의 2개의 혼성화된 영역이 비-혼성화된 영역에 의해 서로 분리되는 경우, "버블" 구조가 야기된다. 상이한 종류의 어댑터들, 예컨대 상이한 서열의 어댑터들이 조합되어 사용될 수 있다. 상이한 어댑터들은 순차적 반응에서 또는 동시에 샘플 폴리뉴클레오타이드에 연결될 수 있다. 일부 구현예에서, 동일한 어댑터들이 표적 폴리뉴클레오타이드의 양 말단에 부가된다. 예를 들면, 제1 및 제2 어댑터가 동일한 반응에 부가될 수 있다. 어댑터는 샘플 폴리뉴클레오타이드와 조합하기 전에 조작될 수 있다. 예를 들면, 말단 포스페이트가 부가되거나 제거될 수 있다.

[0041] 어댑터 올리고뉴클레오타이드가 사용되는 경우, 상기 어댑터 올리고뉴클레오타이드는 하나 이상의 다양한 서열 요소를 포함할 수 있으며, 이는 비제한적으로, 하나 이상의 증폭 프라이머 어닐링 서열 또는 이의 보체, 하나 이상의 시퀀싱 프라이머 어닐링 서열 또는 이의 보체, 하나 이상의 바코드 서열, 다수의 상이한 어댑터 또는 상이한 어댑터의 서브셋 사이에서 공유된 하나 이상의 공통 서열, 하나 이상의 제한 효소 인식 부위, 하나 이상의 표적 폴리뉴클레오타이드 오버행에 상보적인 하나 이상의 오버행, 하나 이상의 프로브 결합 부위 (예컨대, Illumina사에 의해 개발된 바와 같은 유세포와 같은, 대규모 병렬 시퀀싱을 위한 유세포와 같은, 시퀀싱 플랫폼에의 부착을 위함), 하나 이상의 무작위 또는 거의-무작위(near-random) 서열 (예컨대, 하나 이상의 위치에 있는 2개 이상의 상이한 뉴클레오타이드의 세트로부터 무작위로 선택된 하나 이상의 뉴클레오타이드, 상기 하나 이상의 위치에서 선택된 상이한 뉴클레오타이드 각각은 무작위 서열을 포함하는 어댑터의 집합에 나타남), 및 이의 조합을 포함한다. 일부 경우에, 상기 어댑터는 예를 들면 혼성화에 의해 정확한 어댑터를 갖는 폐쇄된 원형을 "포획"할 수 있는, 상기 어댑터에 대해 상보적인 서열을 포함하는 올리고뉴클레오타이드로 코팅된 비드(특히 취급하기 쉬운 자기 비드)를 이용하여 상기 어댑터를 함유하는 상기 원형을 정제하고, 상기 어댑터 및 어떤 라이게이션되지 않은 구성성분을 함유하지 않은 상기 원형을 세척해낸 다음, 상기 비드로부터 상기 포획된 원형을 방출하는데 사용될 수 있다. 또한, 일부 경우에, 혼성화된 포획 프로브 및 표적 원형의 복합체가, 예컨대 직접적인 롤링 서클 증폭(RCA)에 의해서, 연쇄체를 생성하는데 직접 사용될 수 있다. 일부 구현예에서, 상기 원형 내의 어댑터는 시퀀싱 프라이머로서 사용될 수 있다. 2개 이상의 서열 요소들은 서로-비인접하거나 (예컨대, 하나 이상의 뉴클레오타이드에 의해 분리되거나), 서로 인접하거나, 부분적으로 중첩되거나, 완전히 중첩될 수 있다. 예를 들면, 증폭 프라이머 어닐링 서열은 또한 시퀀싱 프라이머 어닐링 서열로서 작용할 수 있다. 서열 요소들은 3' 말단에 또는 그 부근에, 5' 말단에 또는 그 부근에, 또는 어댑터 올리고뉴클레오타이드의 내부에 위치할 수 있다. 서열 요소는 임의의 적합한 길이, 예컨대 약 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25, 30, 35, 40, 45, 50 이상, 또는 그 이하의 뉴클레오타이드 길이일 수 있다. 어댑터 올리고뉴클레오타이드는 이들이 구성하는 하나 이상의 서열 요소를 수용하는데 적어도 충분한 임의의 적합한 길이를 가질 수 있다. 일부 구현예에서, 어댑터는 약 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 90, 100, 200 이상, 또는 그 이하의 뉴클레오타이드 길이이다. 일부 구현예에서, 어댑터 올리고뉴클레오타이드는 약 12 내지 40 뉴클레오타이드 길이의 범위, 예컨대 약 15 내지 35 뉴클레오타이드 길이이다.

[0042] 일부 구현예에서, 하나의 샘플로부터의 단편화된 폴리뉴클레오타이드에 연결된 어댑터 올리고뉴클레오타이드는 모든 어댑터 올리고뉴클레오타이드에 공통적인 하나 이상의 서열 및 상기 특정 샘플의 폴리뉴클레오타이드에 연결된 어댑터에 독특한 바코드를 포함하며, 이로써 상기 바코드 서열은 하나의 샘플 또는 어댑터 연결 반응으로부터 비롯되는 폴리뉴클레오타이드를 또 다른 샘플 또는 어댑터 연결 반응으로부터 비롯되는 폴리뉴클레오타이드와 구별하는데 사용될 수 있다. 일부 구현예에서, 어댑터 올리고뉴클레오타이드는 하나 이상의 표적 폴리뉴클레오타이드 오버행에 상보적인 5' 오버행, 3' 오버행, 또는 이 둘을 포함한다. 상보적 오버행은 비제한적으로 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15 이상의 뉴클레오타이드 길이를 포함하는, 하나 이상의 뉴클레오타이드 길이일 수 있다. 상보적 오버행은 고정된 서열을 포함할 수 있다. 어댑터 올리고뉴클레오타이드의 상보적 오버행은 하나 이상의 뉴클레오타이드의 무작위 서열을 포함할 수 있고, 이로써 하나 이상의 뉴클레오타이드 하나 이상의 위치에 있는 2개 이상의 상이한 뉴클레오타이드의 세트로부터 무작위로 선택되고, 하나 이상

의 위치에서 선택된 상이한 뉴클레오타이드 각각은 상기 무작위 서열을 서열을 포함하는 상보적 오버행을 갖는 어댑터의 집합을 나타낸다. 일부 구현예에서, 어댑터 오버행은 제한 엔도뉴클레아제 절단에 의해 생성된 표적 폴리뉴클레오타이드 오버행에 상보적이다. 일부 구현예에서, 어댑터 오버행은 아데닌 또는 티민으로 이루어진다.

[0043] 폴리뉴클레오타이드를 원형화시키는 다양한 방법이 이용가능하다. 일부 구현예에서, 원형화는 효소적 반응, 예컨대 리가아제(예컨대, RNA 또는 DNA 리가아제)의 사용을 포함한다. 다양한 리가아제가 이용가능하며, 이는 비제한적으로, CircLigase™ (Epicentre; Madison, WI), RNA 리가아제, T4 RNA 리가아제 1 (DNA 및 RNA 모두에 작용하는 ssRNA 리가아제)를 포함한다. 또한, T4 DNA 리가아제 역시 dsDNA 주형이 존재하지 않는 경우 ssDNA를 라이게이션할 수 있는데, 이 반응은 일반적으로 느린 반응이다. 리가아제의 다른 비제한적인 예는 Taq DNA 리가아제, 썬머스 필리포미스 DNA 리가아제, 에스케리키아 콜라이 DNA 리가아제, Tth DNA 리가아제, 썬머스 스코토덕터스 DNA 리가아제 (I 및 II), 내열성 리가아제, Amp리가아제 내열성 DNA 리가아제, VanC-유형 리가아제, 9° N DNA 리가아제, Tsp DNA 리가아제, 및 생물탐사(bioprospecting)에 의해 발견된 신규 리가아제를 포함하는 NAD-의존적 리가아제; T4 RNA 리가아제, T4 DNA 리가아제, T3 DNA 리가아제, T7 DNA 리가아제, Pfu DNA 리가아제, DNA 리가아제 1, DNA 리가아제 III, DNA 리가아제 IV, 및 생물탐사에 의해 발견된 신규 리가아제를 포함하는 ATP-의존적 리가아제; 및 야생형, 돌연변이 아형, 및 이의 유전자 조작된 변이체를 포함한다. 자가-연결이 요구되는 경우, 분자간 구조보다 분자내 원형의 형성을 용이하게 하기 위해 폴리뉴클레오타이드 및 효소의 농도는 조절될 수 있다. 반응 온도 및 시간 역시 조절될 수 있다. 일부 구현예에서, 분자내 원형을 용이하게 하기 위해 60°C가 사용된다. 일부 구현예에서, 반응 시간은 12-16시간이다. 반응 조건은 선택된 효소의 제조사에 의해 명시된 것일 수 있다. 일부 구현예에서, 원형화 반응 후 임의의 라이게이션되지 않은 핵산을 절단하기 위해 엑소뉴클레아제 단계가 포함될 수 있다. 즉, 폐쇄된 원형은 자유 5' 또는 3' 말단을 함유하지 않으며, 따라서, 5' 또는 3' 엑소뉴클레아제의 도입은 상기 폐쇄된 원형을 절단하지 않지만 라이게이션되지 않은 구성요소는 절단할 것이다. 이것은 멀티플렉스 시스템에서 특히 유용할 수 있다.

[0044] 일반적으로, 폴리뉴클레오타이드의 말단을 서로 연결하여 원형 폴리뉴클레오타이드를 형성하는 것 (직접, 또는 하나 이상의 중간 어댑터 올리고뉴클레오타이드를 이용하여)은 접합부 서열을 갖는 접합부를 생성한다. 폴리뉴클레오타이드의 5' 말단 및 3' 말단이 어댑터 폴리뉴클레오타이드를 통해 연결되는 경우, 용어 "접합부"는 폴리뉴클레오타이드 및 어댑터 사이에 접합부를 가리킬 수 있거나 (예컨대, 5' 말단 접합부 또는 3' 말단 접합부 중 하나), 또는 어댑터 폴리뉴클레오타이드에 의해 형성되고 이를 포함하는 것과 같은 폴리뉴클레오타이드의 5' 말단 및 3' 말단 사이에 접합부를 가리킬 수 있다. 폴리뉴클레오타이드의 5' 말단 및 3' 말단이 개입 어댑터 없이 연결되는 경우 (예컨대, 단일 가닥 DNA의 5' 말단 및 3' 말단), 용어 "접합부"는 이들 두 말단이 연결되는 점을 가리킨다. 접합부는 상기 접합부를 포함하는 뉴클레오타이드의 서열("접합부 서열"로도 지칭됨)에 의해 확인될 수 있다. 일부 구현예에서, 샘플은 자연 분해 공정 (예컨대 세포 용해, 세포 사멸, 및 DNA가 세포로부터 그 주변 환경으로 방출되고, 예컨대 무세포 폴리뉴클레오타이드에서 추가로 분해될 수 있는 다른 공정), 샘플 가공 (예컨대 고정, 염색, 및/또는 보관 과정)의 부산물인 단편화, 및 특이적 표적 서열에 대한 제한 없이 DNA를 절단하는 방법에 의한 단편화(예컨대 초음파처리)에 의한, 예컨대 기계적 단편화; 비-서열 특이적 뉴클레아제 치료, 예컨대 DNase I, 단편화효소)에 의해 형성된 말단의 혼합물을 갖는 폴리뉴클레오타이드를 포함한다. 샘플이 말단의 혼합물을 갖는 폴리뉴클레오타이드를 포함하는 경우, 2개의 폴리뉴클레오타이드가 동일한 5' 말단 또는 3' 말단을 가질 가능성은 낮으며, 2개의 폴리뉴클레오타이드가 독립적으로 동일한 5' 말단 및 3' 말단 모두를 가질 가능성은 극히 낮다. 따라서, 일부 구현예에서, 상기 2개의 폴리뉴클레오타이드가 동일한 표적 서열을 갖는 부분을 포함하는 경우에도, 접합부는 상이한 폴리뉴클레오타이드를 구별하는데 사용될 수 있다. 폴리뉴클레오타이드 말단이 개입 어댑터 없이 연결되는 경우, 접합부 서열은 참조 서열에 대한 정렬에 의해 확인될 수 있다. 예를 들면, 2개의 구성요소 서열들의 순서가 참조 서열에 대해 역전된 것으로 보이는 경우, 상기 역전이 발생하는 것으로 보이는 지점이 상기 지점에서 접합부를 표시하는 것일 수 있다. 폴리뉴클레오타이드 말단이 하나 이상의 어댑터 서열을 통해 연결되는 경우, 접합부는 공지된 어댑터 서열에 대한 근접에 의해, 또는 시퀀싱 리드가 상기 원형화된 폴리뉴클레오타이드의 5' 및 3' 말단 모두로부터 서열을 얻는데 충분한 길이인 경우 정렬에 의해 확인될 수 있다. 일부 구현예에서, 특정한 접합부의 형성은 충분히 희귀한 사건이어서 그것은 샘플의 원형화된 폴리뉴클레오타이드 중에서 독특하다.

[0045] 도 4는 원형화된 폴리뉴클레오타이드의 방법의 3개의 비제한적인 예를 예시한다. 상부에서, 폴리뉴클레오타이드는 어댑터의 부재하에 원형화되는 반면, 중앙 도식은 어댑터의 사용을 묘사하고, 하부 도식은 2개의 어댑터를 이용한다. 2개의 어댑터가 사용되는 경우, 하나는 폴리뉴클레오타이드의 5' 말단에 연결될 수 있는 반면, 제2 어댑터는 동일한 폴리뉴클레오타이드의 3' 말단에 연결될 수 있다. 일부 구현예에서, 어댑터 라이게이션은 라이

게이션을 용이하게 하는 2개의 어댑터에 상보적인 "스플린트" 핵산과 함께 2개의 상이한 어댑터의 사용을 포함할 수 있다. 갈라진(Forked) 또는 "Y" 어댑터가 또한 사용될 수 있다. 2개의 어댑터가 사용되는 경우, 양 말단에 동일한 어댑터를 갖는 폴리뉴클레오타이드가 자가-어닐링으로 인해 차후 단계에서 제거될 수 있다.

[0046] 도 6은 단일 가닥 DNA와 같은 폴리뉴클레오타이드를 원형화시키는 추가적인 비제한적인 예시적 방법을 예시한다. 상기 어댑터는 폴리뉴클레오타이드의 5' 또는 3' 말단 중 하나에 비대칭적으로 부가될 수 있다. 도 6A에 나타난 바와 같이, 단일 가닥 DNA(ssDNA)는 3' 말단에 자유 하이드록실 그룹을 가지고, 어댑터는 블로킹된 3' 말단을 가지며, 이로써 리가아제의 존재시, 바람직한 반응이 ssDNA의 3' 말단을 어댑터의 5' 말단에 연결한다. 이러한 구현예에서, 원형을 형성하는 분자내 라이게이션 이전에, 단일 ssDNA 단편과 단일 어댑터의 분자간 라이게이션을 유도하는 폴리프로필렌 글리콜(PEG)과 같은 제제를 사용하는 것이 유용할 수 있다. 말단의 역 순서가 또한 이뤄질 수 있다 (블로킹된 3', 자유 5', 등). 선형 라이게이션이 달성되면, 블로킹 모이어티를 제거하기 위해, 예컨대 키나아제 또는 다른 적합한 효소 또는 화학의 사용을 통해, 상기 라이게이션된 조각들은 효소로 처리될 수 있다. 상기 블로킹 모이어티가 제거되면, 원형화 효소, 예컨대 CircLigase와 같은 원형화 효소의 부가가 원형화된 폴리뉴클레오타이드를 형성하는 분자내 반응을 허용한다. 도 6B에 나타난 바와 같이, 블로킹된 5' 또는 3' 말단을 갖는 하나의 가닥을 갖는 이중 가닥 어댑터를 이용함으로써, 이중 가닥 구조가 형성될 수 있고, 이는 상기 라이게이션시 닉(nick)을 갖는 이중 가닥 단편을 생성한다. 그리고 나서, 상기 2개의 가닥은 분리될 수 있고, 클로킹 모이어티가 제거되고, 단일 가닥 단편이 원형화되어 원형화된 폴리뉴클레오타이드를 형성한다.

[0047] 일부 구현예에서, 분자내 원형화 속도를 향상시키기 위하여 폴리뉴클레오타이드(예컨대, 단일 가닥 DNA)의 2개의 말단을 합치기 위해 분자 클램프가 사용된다. 그러한 한 가지 공정의 예시적 예가 도 5에 예시되어 있다. 이것은 어댑터를 이용하거나 이용하지 않고 수행될 수 있다. 분자 클램프의 사용은 평균 폴리뉴클레오타이드 단편이 약 100 뉴클레오타이드 길이보다 더 큰 경우에 특히 유용할 수 있다. 일부 구현예에서, 분자 클램프 프로브는 3개의 도메인: 제1 도메인, 개입 도메인, 및 제2 도메인을 포함한다. 상기 제1 및 제2 도메인은 서열 상보성을 통해 표적 폴리뉴클레오타이드 내의 상응하는 서열에 먼저 혼성화할 것이다. 분자 클램프 프로브의 개입 도메인은 표적 서열과 유의미하게 혼성화하지 않는다. 따라서, 상기 클램프와 표적 폴리뉴클레오타이드와의 혼성화는 상기 표적 서열의 2개의 말단을 더 근접시켜, 원형화 효소의 존재시에 표적 서열의 분자내 원형화를 용이하게 한다. 일부 구현예에서, 이것은 분자 클램프가 또한 증폭 프라이머로서 작용할 수 있으므로 추가적으로 유용하다.

[0048] 원형화 후, 반응 산물은 서브시퀀스 단계에 참여하는데 이용가능한 원형화된 폴리뉴클레오타이드의 상대적 농도 또는 순도를 증가시키기 위해 증폭 또는 시퀀싱 전에 정제될 수 있다 (예컨대, 반응에서 원형 폴리뉴클레오타이드의 단리 또는 하나 이상의 다른 분자의 제거에 의해). 예를 들면, 원형화 반응 또는 이의 구성성분들은, 예컨대 엑소뉴클레아제 처리에 의해, 단일 가닥 (비-원형화된) 폴리뉴클레오타이드를 제거하기 위해 처리될 수 있다. 추가 예로써, 원형화 반응 또는 이의 일부는 크기 배제 크로마토그래피를 거칠 수 있고, 이에 의해 작은 시약들이 유지되고 버려지거나 (예컨대, 반응하지 않은 어댑터), 또는 원형화 산물이 별개의 부피에 유지되고 방출된다. 라이게이션 반응을 정화하기 위한 다양한 키트, 예컨대 Zymo Research에 의해 제조된 자이모 올리고 정제 키트가 이용가능하다. 일부 구현예에서, 정제는 원형화 반응에 사용된 리가아제를 제거하거나 분해하기 위한 처리, 및/또는 상기 리가아제로부터 원형화된 폴리뉴클레오타이드를 정제해내기 위한 처리를 포함한다. 일부 구현예에서, 리가아제를 분해하는 처리는 프로테아제, 예컨대 프로테이나아제 K 처리를 포함한다. 프로테이나아제 K 처리는 제조사 프로토콜, 또는 표준 프로토콜(예컨대, Sambrook and Green, Molecular Cloning: A Laboratory Manual, 4th Edition (2012)에서 제공되는 바와 같음)을 따를 수 있다. 프로테아제 처리 후에 또한 추출 및 침전이 뒤따를 수 있다. 한 가지 예에서, 원형화된 폴리뉴클레오타이드는 0.1% SDS 및 20 mM EDTA의 존재하에 프로테이나아제 K(Qiagen)에 의해 정제되고, 1:1 페놀/클로로포름 및 클로로포름으로 추출되며, 에탄올 또는 이소프로판올로 침전된다. 일부 구현예에서, 침전은 에탄올 중에서 수행된다.

[0049] 원형화 이후에 곧바로 상기 원형화된 폴리뉴클레오타이드를 시퀀싱하는 것이 뒤따를 수 있다. 대안적으로, 시퀀싱은 하나 이상의 증폭 반응 후에 수행될 수 있다. 일반적으로, "증폭"은 하나 이상의 카피가 표적 폴리뉴클레오타이드 또는 이의 일부로 만들어지는 공정을 가리킨다. 폴리뉴클레오타이드(예컨대, DNA 및/또는 RNA)를 증폭시키는 다양한 방법이 이용가능하다. 증폭은 선형이거나, 지수이거나, 또는 다단계 증폭 공정에서 선형기 및 지수기 모두를 포함할 수 있다. 증폭 방법은 온도 변화, 예컨대 열 변성 단계를 포함할 수 있거나, 또는 열 변성을 필요로 하지 않는 등온 공정일 수 있다. 폴리머라아제 연쇄 반응(PCR)은 변성, 프라이머 쌍의 반대 가닥에의 어닐링, 및 표적 서열의 카피수를 지수적으로 증가시키는 프라이머 연장의 다수의 사이클을 이용한다. 어닐링된

핵산 가닥의 변성은 국소적인 금속 이온 농도를 증가시키는 열의 적용(예컨대, 미국 특허 제6,277,605호), 초음파 방사선(예컨대, 제WO/2000/049176호), 전압의 적용(예컨대, 미국 특허 제5,527,670호, 미국 특허 제6,033,850호, 미국 특허 제5,939,291호, 및 미국 특허 제6,333,157호), 및 자기 반응 물질에 결합된 프라이머와 조합된 전자기장의 적용(예컨대, 미국 특허 제5,545,540호)에 의해 달성될 수 있다. 변이 판정된 RT-PCR에서, RNA로부터 상보적 DNA(cDNA)를 만들기 위해 역전사효소(RT)가 사용되며, 그리고 나서 상기 cDNA는 PCR에 의해 증폭되어 DNA의 다수의 카피를 생성한다 (예컨대, 미국 특허 제5,322,770호 및 미국 특허 제5,310,652호). 등온 증폭 방법의 한 가지 예는 통상 SDA로 불리는 가닥 치환 증폭이며, 이는 표적 서열의 반대 가닥에 프라이머 서열의 쌍의 어닐링, 이중체 헤미포스포리토에이트된 프라이머 연장 산물을 생성하는 dNTP의 존재하에 프라이머 연장, 헤미변형된 제한 엔도뉴클레아제 인식 부위의 엔도뉴클레아제-매개된 Nicking(nicking), 및 존재하는 가닥을 치환하고 프라이머 어닐링, Nicking 및 가닥 치환의 다음 라운드를 위한 가닥을 생성하기 위해 상기 Nick의 3' 말단으로부터의 폴리머라아제-매개된 프라이머 연장의 사이클을 이용하여, 산물의 기하학적 증폭을 야기한다 (예컨대, 미국 특허 제5,270,184호 및 미국 특허 제5,455,166호). 고온성 SDA(tSDA)는 본질적으로 동일한 방법에서 더 높은 온도에서 고온성 엔도뉴클레아제 및 폴리머라아제를 사용한다 (유럽 특허 제0 684 315호). 다른 증폭 방법은 롤링 서클 증폭(RCA) (예컨대, Lizardi, "Rolling Circle Replication Reporter Systems," 미국 특허 제5,854,033호); 헬리카아제 의존적 증폭(HDA) (예컨대, Kong et al., "Helicase Dependent Amplifications," 미국 특허출원 공개 제US 2004-0058378호 A1); 및 루프-매개된 등온 증폭(LAMP) (예컨대, Notomi et al., "Process for Synthesizing Nucleic Acids," 미국 특허 제6,410,278호)을 포함한다. 일부 경우에, 등온 증폭은 올리고뉴클레오타이드 프라이머 내로 혼입될 수 있는, 프로모터 서열로부터 RNA 폴리머라아제에 의한 전사를 이용한다. 전사-기반의 증폭 방법은 NASBA로도 불리는 핵산 서열 기반 증폭 (예컨대, 미국 특허 제5,130,238호); 통상 Q β 레플리카아제로도 불리는, 프로브 분자 자체를 증폭시키는 RNA 레플리카아제의 사용에 의존하는 방법 (예컨대, Lizardi, P. et al. (1988) *BioTechnol.* 6, 1197-1202); 자립(self-sustained) 서열 복제 (예컨대, Guatelli, J. et al. (1990) *Proc. Natl. Acad. Sci. USA* 87, 1874-1878; Landgren (1993) *Trends in Genetics* 9, 199-202; 및 HELEN H. LEE et al., NUCLEIC ACID AMPLIFICATION TECHNOLOGIES (1997)); 및 부가적인 전사 주형을 생성하는 방법 (예컨대, 미국 특허 제5,480,784호 및 미국 특허 제5,399,491호)을 포함한다. 추가적인 등온 핵산 증폭 방법은 부가적인 프라이머를 위한 결합 부위를 노출시키는 비-기본형 뉴클레오타이드에 있는 핵산을 절단하는 효소(예컨대, DNA 글리코실라아제 또는 RNaseH)와 조합하여 비-기본형 뉴클레오타이드 (예컨대, 우라실 또는 RNA 뉴클레오타이드)를 함유하는 프라이머의 사용을 포함한다(예컨대, 미국 특허 제6,251,639호, 미국 특허 제6,946,251호, 및 미국 특허 제7,824,890호). 등온 증폭 공정은 선형이거나 지수일 수 있다.

[0050]

일부 구현예에서, 증폭은 롤링 서클 증폭(RCA)을 포함한다. 전형적인 RCA 반응 혼합물은 하나 이상의 프라이머, 폴리머라아제, 및 dNTP를 포함하고, 연쇄체를 생성한다. 전형적으로, RCA 반응에서의 폴리머라아제는 가닥 치환 활성을 갖는 폴리머라아제이다. 그러한 다양한 폴리머라아제가 이용가능하며, 이의 비제한적인 예는 엑소뉴클레아제가 결합된 DNA 폴리머라아제 I 거대 (Klenow) 단편, Phi29 DNA 폴리머라아제, Taq DNA 폴리머라아제 등을 포함한다. 일반적으로, 연쇄체는 주형 폴리뉴클레오타이드로부터 표적 서열의 2개 이상의 카피를 포함하는 폴리뉴클레오타이드 증폭 산물 (예컨대, 표적 서열의 약 2, 3, 4, 5, 6, 7, 8, 9, 10 이상 또는 그 이상의 카피; 일부 구현예에서, 약 2 카피 또는 그 이상의 카피)이다. 증폭 프라이머는 임의의 적합한 길이, 예컨대 약 또는 적어도 약 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 90, 100 이상의 뉴클레오타이드일 수 있으며, 이의 어느 일부 또는 모두는 프라이머가 혼성화하는 상응하는 표적 서열에 상보적일 수 있다 (예컨대, 약, 또는 적어도 약 5, 10, 15, 20, 25, 30, 35, 40, 45, 50 이상의 뉴클레오타이드). 도 7은 적합한 프라이머의 3가지 비제한적인 예를 묘사한다. 도 7A는 어댑터를 사용하지 않고 특이적 표적 서열 내의 서열 변이의 존재 또는 부재의 검출을 위해 사용될 수 있는 표적 특이적 프라이머를 사용하는 것을 보여준다. 일부 구현예에서, 복수의 표적을 위한 다중 표적-특이적 프라이머가 동일한 반응에서 사용된다. 예를 들면, 약 또는 적어도 약 10, 50, 100, 150, 200, 250, 300, 400, 500, 1000, 2500, 5000, 10000, 15000 이상의 상이한 표적 서열에 대한 표적-특이적 프라이머가 상응하는 수의 표적 서열 (존재하는 경우)을 병렬로 증폭시키기 위해 단일 증폭 반응에서 사용될 수 있다. 다수의 표적 서열은 동일한 유전자, 상이한 유전자, 또는 비-유전자 서열의 상이한 부분에 상응할 수 있다. 다수의 프라이머가 단일 유전자 내의 다수의 표적 서열을 표적화하는 경우, 프라이머는 표적 유전자의 모두 또는 특정 부분을 커버하기 위해 유전자 서열을 따라 간격을 둘 수 있다 (예컨대, 약 또는 적어도 약 50 뉴클레오타이드, 50-150 뉴클레오타이드마다, 또는 50-100 뉴클레오타이드마다 간격을 두고 떨어져 있음). 도 7C는 어댑터 서열에 혼성화하는 프라이머 (이는 일부 경우에 어댑터 올리고뉴클레오타이드 자체일 수 있음)의 사용을 예시한다.

- [0051] 도 7B는 무작위 프라이머에 의한 증폭의 예를 예시한다. 일반적으로, 무작위 프라이머는 하나 이상의 무작위 또는 거의-무작위 서열(예컨대, 하나 이상의 위치에 있는 2개 이상의 상이한 뉴클레오타이드의 세트로부터 무작위로 선택된 하나 이상의 뉴클레오타이드, 하나 이상의 위치에서 선택된 상이한 뉴클레오타이드 각각은 무작위 서열을 포함하는 어댑터의 집합에 나타남)을 포함한다. 이러한 방식으로, 폴리뉴클레오타이드(예컨대, 모든 또는 실질적으로 모든 원형화된 폴리뉴클레오타이드)는 서열 비특이적 방식으로 증폭될 수 있다. 그러한 과정은 "전체 지놈 증폭" (WGA)으로 지칭될 수 있으나, 전형적인 WGA 프로토콜 (원형화 단계를 포함하지 않음)은 짧은 폴리뉴클레오타이드, 예컨대 본 개시내용에 의해 고려되는 폴리뉴클레오타이드 단편을 효율적으로 증폭하지 못한다. WGA 과정의 추가의 예시적인 논의를 위해, 예를 들면 문헌[Li et al (2006) *J Mol. Diagn.* 8(1):22-30]을 참고한다.
- [0052] 원형화된 폴리뉴클레오타이드가 시퀀싱 전에 증폭되는 경우, 증폭된 산물은 농축 없이 곧바로 시퀀싱될 수 있거나, 하나 이상의 농축 단계 이후 시퀀싱될 수 있다. 농축은, 예컨대 증폭 산물의 보유 또는 하나 이상의 시약의 제거에 의해, 하나 이상의 구성성분을 정제하는 것을 포함할 수 있다. 예를 들면, 증폭 산물은 기질에 부착된 복수의 프로브에의 혼성화 후, 예컨대 세척 단계에 의해, 포획된 폴리뉴클레오타이드를 방출함으로써 정제될 수 있다. 대안적으로, 증폭 산물은 결합 쌍 중 하나의 멤버로 표지될 수 있고, 이후, 기질에 부착된 결합 쌍 중 다른 멤버에 결합되고, 세척되어 증폭 산물을 방출한다. 가능한 기질은, 비제한적으로 유리 및 변형된 또는 관능화된 유리, 플라스틱 (아크릴계, 폴리스티렌 및 스티렌 및 다른 물질의 공중합체, 폴리프로필렌, 폴리에틸렌, 폴리부틸렌, 폴리우레탄, 테플론™, 등 포함), 다당류, 나일론 또는 니트로셀룰로오스, 세라믹, 레진, 실리카 또는 실리콘 및 변형된 실리콘을 포함하는 실리카계 물질, 탄소, 금속, 무기 유리, 플라스틱, 광섬유 다발, 및 다양한 다른 폴리머를 포함한다. 일부 구현예에서, 기질은 비드 또는 다른 작은 별개의 입자의 형태이며, 이는 자기장의 적용을 통한 단리를 용이하게 하기 위해 자기 또는 상자성 비드일 수 있다. 일반적으로, "결합 쌍"은 제1 및 제2 모이어티 중 하나를 가리키며, 여기서 상기 제1 및 제2 모이어티는 서로에 대해 특이적 결합 친화도를 갖는다. 적합한 결합 쌍은, 비제한적으로, 항원/항체 (예를 들면, 디곡시게닌/항-디곡시게닌, 디니트로페닐 (DNP)/항-DNP, 단실-X-항-단실, 플루오레세인/항-플루오레세인, 루시퍼 엘로우/항-루시퍼 엘로우, 및 로다민 항-로다민); 바이오틴/아비딘 (또는 바이오틴/스트렙타비딘); 칼모둘린 결합 단백질 (CBP)/칼모둘린; 호르몬/호르몬 수용체; 렉틴/탄수화물; 펩타이드/세포 막 수용체; 단백질 A/항체; 합텐/안티합텐; 효소/보조인자; 및 효소/기질을 포함한다.
- [0053] 일부 구현예에서, 원형화된 폴리뉴클레오타이드의 증폭 후 농축은 하나 이상의 추가적인 증폭 반응을 포함한다. 일부 구현예에서, 농축은 하기를 포함하는 증폭 반응 혼합물에서 서열 A 및 서열 B (5'에서 3' 방향으로 배향됨)를 포함하는 표적 서열을 증폭하는 것을 포함한다: (a) 증폭된 폴리뉴클레오타이드; (b) 서열 A'를 포함하는 제1 프라이머, 여기서 상기 제1 프라이머는 서열 A와 서열 A' 간의 서열 상보성을 통해 표적 서열의 서열 A에 특이적으로 혼성화하고; (c) 서열 B를 포함하는 제2 프라이머, 여기서 상기 제2 프라이머는 B와 B' 간의 서열 상보성을 통해 표적 서열의 보체를 포함하는 상보적 폴리뉴클레오타이드에 존재하는 서열 B'에 특이적으로 혼성화하며; 및 (d) 상기 제1 프라이머 및 상기 제2 프라이머를 연장시켜 증폭된 폴리뉴클레오타이드를 생성하는 폴리머라아제; 여기서 표적 서열의 서열 A의 5' 말단과 서열 B의 3' 말단 사이의 거리는 75nt 이하이다. 도 10은 단일 반복 (이는 원형이 아니면 전형적으로 증폭되지 않을 것임) 및 표적 서열의 다수의 카피를 포함하는 연쇄체의 문맥에서 표적 서열에 대한 제1 및 제2 프라이머의 예시적 배열을 예시한다. 표적 서열의 모노머에 대한 프라이머의 방향을 고려할 때, 이 배열은 "백투백"(B2B) 또는 "역전된(inverted)" 프라이머로서 지칭될 수 있다. B2B 프라이머를 이용한 농축은 원형 및/또는 연쇄체 증폭 산물의 농축을 용이하게 한다. 더욱이, 상대적으로 작은 풋프린트 (한 쌍의 프라이머가 포괄하는 총 거리)와 조합된 이러한 방향은 표적 서열 주위의 더 다양한 단편화 사건의 증폭을 가능하게 하는데, 접합부가 전형적인 증폭 반응에서 발견되는 프라이머의 배열 (서로 마주보아 표적 서열을 포괄함)에서보다 프라이머 사이에서 존재할 가능성이 낮기 때문이다. 일부 구현예에서, 서열 A의 5' 말단과 서열 B의 3' 말단 사이의 거리는 약 200, 150, 100, 75, 50, 40, 30, 25, 20, 15 이하, 또는 그 미만의 뉴클레오타이드이다. 일부 구현예에서, 서열 A는 서열 B의 보체이다. 일부 구현예에서, 복수의 상이한 표적 서열에 관한 B2B 프라이머의 다수의 쌍이 복수의 상이한 표적 서열을 병렬로 (예컨대, 약 또는 적어도 약 10, 50, 100, 150, 200, 250, 300, 400, 500, 1000, 2500, 5000, 10000, 15000 이상의 상이한 표적 서열) 증폭시키기 위해 동일한 반응에서 사용된다. 프라이머는, 예컨대 본원의 다른 곳에 기재된, 임의의 적합한 길이일 수 있다. 증폭은, 예컨대 증폭 반응 본원에 기재된, 적절한 반응 조건 하에 임의의 적합한 증폭 반응을 포함할 수 있다. 일부 구현예에서, 증폭은 폴리머라아제 연쇄 반응이다.
- [0054] 일부 구현예에서, B2B 프라이머는 적어도 2개의 서열 요소, 즉 서열 상보성을 통해 표적 서열에 혼성화하는 제1

요소, 및 상기 제1 요소가 혼성화하는 제1 혼성화 온도에서의 제1 증폭 과정 동안 표적 서열에 혼성화하지 않는 5' "테일" (예컨대, 제1 요소가 결합하는 곳에 대하여 바로 3'에 있는 표적 서열의 일부와 상기 테일 사이의 서열 상보성의 결여로 인해)을 포함한다. 예를 들면, 상기 제1 프라이머는 서열 A'에 대해 서열 C 5'를 포함하고, 상기 제2 프라이머는 서열 B에 대해 서열 D 5'를 포함하며, 서열 C 및 서열 D 어느 것도 제1 혼성화 온도에서의 제1 증폭 과정 동안 복수의 연쇄체에 혼성화하지 않는다. 상기 테일을 가진 프라이머가 사용되는 일부 구현예에서, 증폭은 제1 과정 및 제2 과정을 포함할 수 있고; 상기 제1 과정은 상기 제1 및 제2 프라이머가 연쇄체 (또는 원형화된 폴리뉴클레오타이드)에 혼성화하는 제1 온도에서의 혼성화 단계 및 프라이머 연장을 포함하며; 상기 제2 과정은 제1 및 제2 프라이머가 연장된 제1 또는 제2 프라이머, 또는 이의 보체를 포함하는 증폭 산물에 혼성화하는, 제1 온도보다 더 높은 제2 온도에서의 혼성화 단계, 및 프라이머 연장을 포함한다. 상기 더 높은 온도는 단지 프라이머 내의 제1 요소 및 연쇄체 내의 내부 표적 서열 사이의 혼성화에 의해 형성된 더 짧은 단편보다 프라이머 연장 산물 내의 제1 요소 및 테일 요소 사이의 혼성화에 유리하다. 따라서, 상기 2-과정 증폭은 짧은 증폭 산물이 유리할 수 있는 정도를 감소시키는데 사용될 수 있고, 이에 의해 표적 서열의 2개 이상의 카피를 갖는 증폭 산물의 상대적으로 더 높은 비율을 유지한다. 예를 들면, 제2 온도에서의 혼성화 및 프라이머 연장의 5 사이클 (예컨대, 적어도 5, 6, 7, 8, 9, 10, 15, 20 이상의 사이클) 후, 상기 반응 혼합물 내의 증폭된 폴리뉴클레오타이드 중 적어도 5% (예컨대, 적어도 5%, 6%, 7%, 8%, 9%, 10%, 15%, 20%, 25%, 30% 이상)가 표적 서열의 2개 이상의 카피를 포함한다. 이러한 2-과정의 테일을 가진 B2B 프라이머 증폭 공정에 따른 구현예의 예시가 도 11에 예시되어 있다.

[0055] 일부 구현예에서, 농축은 연쇄체로부터 앰플리콘의 길이를 증가시키기 위해 편향된 조건하에서의 증폭을 포함한다. 예를 들면, 프라이머 농도는 낮아질 수 있으며, 이로써 모든 프라이밍 부위가 프라이머와 혼성화하지는 않을 것이며, 따라서 이는 PCR 산물을 더 길게 만든다. 유사하게, 사이클 동안 프라이머 혼성화 시간을 감소시키는 것은 유사하게 더 적은 프라이머가 혼성화하게 하여, 또한 평균 PCR 앰플리콘 크기를 증가시킬 것이다. 또한, 사이클의 온도 및/또는 연장 시간을 증가시키는 것은 유사하게 PCR 앰플리콘의 평균 길이를 증가시킬 것이다. 이들 기술의 어느 조합이 이용될 수 있다.

[0056] 일부 구현예에서, 특히 B2B 프라이머를 이용한 증폭이 수행된 경우, 증폭 산물은 연쇄체를 포함하는 혼합물로부터 모노머의 수를 감소시키고/거나 제거하기 위해 크기를 기준으로 수득된 앰플리콘을 여과하도록 처리된다. 이는 다양한 이용가능한 기술을 이용하여 수행될 수 있으며, 이는 비제한적으로, 겔로부터 단편 절제 및 겔 여과 (예컨대, 약 300, 400, 500 이상의 뉴클레오타이드 길이보다 더 큰 단편을 농축하기 위해) 뿐만 아니라 결합 버퍼 농도를 미세 조정하여 크기 선별하기 위한 SPRI 비드(Agencourt AMPure XP)를 포함한다. 예를 들면, DNA 단편과 혼합하는 동안 0.6x 결합 버퍼의 사용은 약 500 염기쌍(bp)보다 큰 DNA 단편에 우선적으로 결합시키는데 사용될 수 있다.

[0057] 일부 구현예에서, 증폭이 단일 가닥 연쇄체를 야기하는 경우, 상기 단일 가닥은 시퀀싱 반응을 위해 생성되는 시퀀싱 라이브러리의 형성 전에 또는 형성의 일부로서 이중 가닥 작제물로 전환된다. 단일 가닥 핵산으로부터 이중 가닥 작제물을 생성하는 다양한 적합한 방법들이 이용가능하다. 많은 가능한 방법들이 도 9에 묘사되어 있으나, 많은 다른 방법들 역시 사용될 수 있다. 도 9A에 나타난 바와 같이, 예를 들면, 무작위 프라이머, 폴리머라아제, dNTP 및 리가아제의 사용은 이중 가닥을 야기할 것이다. 도 9B는 연쇄체가 상기 반응에서 프라이머로서 사용될 수 있는 어댑터 서열을 함유할 때의 제2 가닥 합성을 묘사한다. 도 9C는 "루프"의 사용을 묘사하며, 여기서 상기 루프 어댑터의 하나의 말단은 상기 연쇄체의 말단에 부가되며, 상기 루프 어댑터는 자가-혼성화 핵산의 작은 구획을 갖는다. 이 경우에, 상기 루프 어댑터의 라이제이션은 자가 혼성화되어 폴리머라아제 프라이머 주형으로서 작용하는 루프를 야기한다. 도 9D는 일반적으로 표적 서열이 공지된 경우에 가장 많이 이용되는 과분지형 프라이머의 사용을 보여주며, 여기서, 특히 강한 가닥 치환 기능을 갖는 폴리머라아제가 사용될 때, 다수의 가닥이 형성된다.

[0058] 일부 구현예에 따르면, 원형화된 폴리뉴클레오타이드 (또는 선택적으로 농축되었을 수 있는, 이의 증폭 산물)는 시퀀싱 리드를 생성하기 위해 시퀀싱 반응을 거친다. 상기 방법에 의해 생성된 시퀀싱 리드는 본원에 개시된 다른 방법에 따라 사용될 수 있다. 다양한 시퀀싱 방법, 특히 고처리 시퀀싱 방법이 이용가능하다. 예는, 비제한적으로, 일루미나에 의해 제작된 시퀀싱 시스템(HiSeq® 및 MiSeq®과 같은 시퀀싱 시스템), 라이프 테크놀로지스(Ion Torrent®, SOLiD®, 등), 로슈의 454 Life Sciences 시스템, 퍼시픽 바이오사이언시스 시스템 등을 포함한다. 일부 구현예에서, 시퀀싱은 약 50, 75, 100, 125, 150, 175, 200, 250, 300 이상, 또는 그 이상의 뉴클레오타이드 길이를 초과하는 리드를 생성하기 위해 HiSeq® 및 MiSeq® 시스템의 사용을 포함한다. 일부 구현예에서, 시퀀싱은 합성 공정에 의한 시퀀싱을 포함하고, 여기서 각각의 뉴클레오타이드는, 이들이 성장하는 프

라이머 연장 산물에 부가됨에 따라, 반복적으로 확인된다. 피로시퀀싱은 시퀀싱 반응의 부산물, 즉 피로포스페이트의 존재를 위해 수득된 합성 혼합물을 분석함으로써 뉴클레오타이드의 혼입을 확인하는 합성 공정에 의한 시퀀싱의 예이다. 특히, 프라이머/주형/폴리머라아제 복합체는 단일 유형의 뉴클레오타이드와 접촉된다. 만약 상기 뉴클레오타이드가 혼입되면, 중합 반응은 삼인산염 사슬의 α 및 β 인산염 사이에서 뉴클레오사이드 삼인산염을 절단하여, 피로포스페이트를 방출한다. 그리고 나서, 방출된 피로포스페이트의 존재를 AMP를 이용하여 피로포스페이트를 ATP로 전환시키는 화학발광성 효소 리포터 시스템을 이용하여 확인한 후, 측정가능한 광 신호를 생성하는 루시페라아제 효소를 이용하여 ATP를 측정한다. 광이 검출되는 경우, 염기가 혼입된 것이며, 광이 검출되지 않는 경우, 염기가 혼입되지 않은 것이다. 적절한 세척 단계 후, 다양한 염기들이 사이클적으로 상기 복합체와 접촉되어 순차적으로 주형 서열 내의 그 다음 염기를 확인한다. 예컨대, 미국 특허 제6,210,891호를 참고한다.

[0059]

관련된 시퀀싱 공정에서, 프라이머/주형/폴리머라아제 복합체는 기질 상에 고정화되고, 상기 복합체는 표지된 뉴클레오타이드와 접촉된다. 상기 복합체의 고정화는 프라이머 서열, 주형 서열 및/또는 폴리머라아제 효소를 통해 이뤄지며, 공유성이거나 비공유성일 수 있다. 예를 들면, 복합체의 고정화는 폴리머라아제 또는 프라이머 및 기질 표면 사이의 연결을 통해서 이뤄질 수 있다. 대안적인 구성에서, 상기 뉴클레오타이드는 제거가능한 종결 그룹과 함께 또는 상기 그룹 없이 제공된다. 혼입시, 표지는 상기 복합체와 결합되어 검출될 수 있다. 종결자를 갖는 뉴클레오타이드의 경우, 개별적으로 확인가능한 표지를 갖는 4개의 상이한 뉴클레오타이드 모두가 상기 복합체와 접촉된다. 상기 표지된 뉴클레오타이드의 혼입은 상기 종결자의 존재에 의해 연장을 정지시키고, 상기 복합체에 표지를 부가하며, 이는 상기 혼입된 뉴클레오타이드의 확인을 가능하게 한다. 그리고 나서, 상기 표지 및 종결자가 상기 혼입된 뉴클레오타이드로부터 제거되고, 적절한 세척 단계 이후, 상기 공정은 반복된다. 비-종결된 뉴클레오타이드의 경우, 상기 뉴클레오타이드가 혼입된 여부를 결정하기 위해 단일 유형의 표지된 뉴클레오타이드가 상기 복합체에 부가된다. 뉴클레오타이드 상의 표지 그룹의 제거 및 적절한 세척 단계 후, 다양한 상이한 뉴클레오타이드가 상기 동일한 공정에서 상기 반응 혼합물을 통해 순환한다. 예컨대, 미국 특허 제6,833,246호를 참고하며, 이는 모든 목적을 위해 그 전체가 참고로 본원에 통합되어 있다. 예를 들면, 일루미나 지놈 분석기 시스템은 제WO 98/44151호에 기술된 기술에 기반을 두고 있으며, 여기서 DNA 분자는 앵커 프로브 결합 부위 (다르게는 유세포 결합 부위로도 불림)를 통해 시퀀싱 플랫폼 (유세포)에 결합되고 유리 슬라이드 상에서 그 자리에서 증폭된다. DNA 분자가 증폭되는 고체 표면은 전형적으로 복수의 제1 및 제2 결합된 올리고뉴클레오타이드, 즉 표적 폴리뉴클레오타이드의 하나의 말단 부근 또는 말단에 있는 서열에 상보적인 제1 결합된 올리고뉴클레오타이드 및 표적 폴리뉴클레오타이드의 다른 말단 부근 또는 말단에 있는 서열에 상보적인 제2 결합된 올리고뉴클레오타이드를 포함한다. 이러한 배열은, 예컨대 제US20140121116호에 기술된, 브릿지 증폭을 가능하게 한다. 그리고 나서, 상기 DNA 분자는 시퀀싱 프라이머에 어닐링되고 가역적 종결자 접근법을 이용하여 염기별로 병렬 시퀀싱된다. 시퀀싱 프라이머의 혼성화는 브릿지를 앵커링하는 상기 결합된 올리고뉴클레오타이드 중 하나에서의 절단 부위에서 이중 가닥 브릿지 폴리뉴클레오타이드 중 하나의 가닥의 절단보다 뒤에 일어나며, 따라서 변성에 의해 제거될 수 있는 고체 기질에 결합되지 않은 하나의 단일 가닥 및 결합되고 시퀀싱 프라이머에의 혼성화에 이용가능한 다른 가닥을 남긴다. 전형적으로, 일루미나 지놈 분석기 시스템은 8 채널을 갖는 유세포를 이용하여 18 내지 36 염기 길이의 시퀀싱 리드를 생성하고, 런당 >1.3 Gbp의 고품질 데이터를 생성한다 (www.illumina.com 참조).

[0060]

합성 공정에 의한 추가 시퀀싱에서, 주형 의존적 합성이 수행됨에 따라 상이하게 표지된 뉴클레오타이드의 혼입이 실시간으로 관찰된다. 특히, 형광 표지된 뉴클레오타이드가 혼입됨에 따라 각각의 고정화된 프라이머/주형/폴리머라아제 복합체가 관찰되며, 이는 그것이 부가될 때 각각의 부가된 염기의 실시간 확인을 가능하게 한다. 이러한 공정에서, 표지 그룹은 뉴클레오타이드의 일부에 부착되어 혼입 동안 절단된다. 예를 들면, 표지 그룹을 혼입 동안 제거된 인산염 사슬의 일부, 즉, 뉴클레오사이드 다인산염 상의 β , γ , 또는 다른 말단 인산염기에 부착시킴으로써, 상기 표지는 상기 생성되려고 하는 가닥 내로 혼입되지 않으며, 대신 천연 DNA가 생성된다. 각각의 분자의 관찰은 전형적으로 매우 작은 조명 부피 내에 상기 복합체의 광 집중(optical confinement)을 포함한다. 상기 복합체를 광집중시킴으로써, 관찰 영역을 만들며, 여기에서 무작위로 확산하는 뉴클레오타이드는 매우 짧은 시간 동안 존재하는 대신 혼입된 뉴클레오타이드는 이들이 혼입될 때 더 긴 시간 동안 관찰 부피 내에서 유지된다. 이는 혼입 사건과 연관된 특징적인 신호를 야기하며, 이는 또한 부가되는 염기의 특징인 신호 프로파일을 특징으로 한다. 관련된 양태에서, 상호작용하는 표지 구성성분, 예컨대 형광 공명 에너지 전달(FRET) 염료 쌍이 폴리머라아제 또는 상기 복합체의 다른 부분 및 혼입하는 뉴클레오타이드에 제공되며, 이로써 혼입 사건이 표지 구성성분들을 상호적으로 근접시키고, 혼입되는 염기의 특징인 특징적인 신호를 야기한다 (예컨대, 미국 특허 제6,917,726호, 제7,033,764호, 제7,052,847호, 제7,056,676호, 제7,170,050호, 제7,361,466호, 및

제7,416,844호; 및 US 제20070134128호 참고).

- [0061] 일부 구현예에서, 샘플 내의 핵산은 라이게이션에 의해 시퀀싱될 수 있다. 이러한 방법은, 예를 들면, 폴로니(polony) 방법에서 그리고 SOLiD 기술(Applied Biosystems, 현재는 Invitrogen)에서와 같이, 표적 서열을 확인하기 위해 전형적으로 DNA 리가아제 효소를 이용한다. 일반적으로, 고정된 길이의 모든 가능한 올리고뉴클레오타이드의 집단이 제공되며, 시퀀싱된 위치에 따라 표지된다. 올리고뉴클레오타이드들은 어닐링되고 라이게이션되며; 매칭 서열에 대한 DNA 리가아제에 의한 우선적인 라이게이션이 상기 위치에서 상보적인 서열에 상응하는 신호를 야기한다.
- [0062] 일부 구현예에 따르면, 시퀀싱 리드와 참조 서열 간의 서열 차이는, 만약 상기 차이가 적어도 2개의 상이한 폴리뉴클레오타이드(예컨대, 상이한 접합부를 갖는 결과로서 구별될 수 있는, 2개의 상이한 원형 폴리뉴클레오타이드)에서 발생하면, 진정한 서열 변이(예컨대, 증폭 또는 시퀀싱 전에 샘플 내에 존재하며, 이들 공정의 결과가 아님)로 불린다. 증폭 또는 시퀀싱 오류의 결과인 서열 변이는 동일한 표적 서열을 포함하는 2개의 상이한 폴리뉴클레오타이드 상에서 정확하게 (예컨대, 위치 및 유형) 복제되지 않을 가능성이 있기 때문에, 이러한 검증 파라미터를 추가하는 것은 오류 서열 변이의 배경을 크게 감소시키며, 이와 동시에 샘플 내의 실제 서열 변이를 검출하는 민감도 및 정확도를 증가시킨다. 일부 구현예에서, 5%, 4%, 3%, 2%, 1.5%, 1%, 0.75%, 0.5%, 0.25%, 0.1%, 0.075%, 0.05%, 0.04%, 0.03%, 0.02%, 0.01%, 0.005%, 0.001% 이하, 또는 그 미만의 낮은 빈도를 갖는 서열 변이는 유의미하게 배경을 초과하여 정확한 판정을 가능하게 한다. 일부 구현예에서, 상기 서열 변이는 약 0.1% 또는 그 미만의 빈도로 일어난다. 일부 구현예에서, 서열 변이의 빈도는 상기 빈도가 통계학적으로 유의미하게 배경 오류율을 초과할 때 (예컨대, 약 0.05, 0.01, 0.001, 0.0001 이하, 또는 그 미만의 p-값을 가짐) 유의미하게 배경을 초과한다. 일부 구현예에서, 서열 변이의 빈도는 상기 빈도가 배경 오류율을 약 또는 적어도 약 2배, 3배, 4배, 5배, 6배, 7배, 8배, 9배, 10배, 25배, 50배, 100배 이상 초과할 때 (예컨대, 적어도 5배 더 높음), 유의미하게 배경을 초과한다. 일부 구현예에서, 주어진 위치에서 서열을 정확히 결정하는데 있어서 배경 오류율은 약 1%, 0.5%, 0.1%, 0.05%, 0.01%, 0.005%, 0.001%, 0.0005% 이하, 또는 그 미만이다. 일부 구현예에서, 오류율은 0.001%보다 낮다.
- [0063] 일부 구현예에서, 진정한 서열 변이를 확인하는 것 ("판정하는 것" 또는 "판정을 행하는 것"으로도 지칭됨)은 하나 이상의 시퀀싱 리드를 참조 서열과 최적으로 정렬하여 상기 둘 사이의 차이를 확인하는 것뿐만 아니라 접합부를 확인하는 것을 포함한다. 일반적으로, 정렬은 하나의 서열을 또 다른 서열과 함께 두고, 각각의 서열을 따라 갭을 반복적으로 도입하고, 두 서열이 얼마나 잘 매치하는지 점수를 매기고, 바람직하게는 참조를 따라 다양한 위치에 대해 반복하는 것을 포함한다. 최적-스코어링 매치는 정렬인 것으로 여겨지며, 서열들 간의 관련성 정도에 관한 추론에 해당한다. 일부 구현예에서, 시퀀싱 리드가 비교되는 참조 서열은 참조 지놈, 예컨대 대상과 동일한 종의 멤버의 지놈이다. 참조 지놈은 완전하거나 불완전할 수 있다. 일부 구현예에서, 참조 지놈은, 예컨대 참조 지놈으로부터 또는 분석 중인 시퀀싱 리드로부터 생성된 컨센서스로부터의 표적 폴리뉴클레오타이드를 함유하는 영역만으로 이루어진다. 일부 구현예에서, 참조 서열은 하나 이상의 유기체의 폴리뉴클레오타이드의 서열, 예컨대 하나 이상의 박테리아, 고세균, 바이러스, 원생생물, 진균, 또는 다른 유기체로부터의 서열을 포함하거나 이들로 이루어진다. 일부 구현예에서, 참조 서열은 참조 지놈의 일부, 예컨대 분석 중인 하나 이상의 표적 서열에 상응하는 영역(예컨대, 하나 이상의 유전자, 또는 이의 일부)만으로 이루어진다. 예를 들면, 병원균의 검출을 위해 (예컨대 오염 검출의 경우), 참조 지놈은 병원균(예컨대, HIV, HPV, 또는 유해한 박테리아 균주, 예컨대, E. 콜라이)의 전체 지놈, 또는 예컨대 특정한 균주 또는 혈청형의 확인에 유용한 이의 일부이다. 일부 구현예에서, 시퀀싱 리드는, 예컨대 다수의 상이한 유기체 또는 균주를 스크리닝하기 위해, 다수의 상이한 참조 서열에 정렬된다.
- [0064] 전형적인 정렬에서, 참조 내의 비-매칭 염기 옆의 시퀀싱 리드 내의 염기는 상기 지점에서 치환 돌연변이가 일어났음을 가리킨다. 유사하게, 하나의 서열이 다른 서열 내의 염기 옆에 갭을 포함하는 경우, 삽입 또는 결실 돌연변이("인델")가 일어난 것으로 추론된다. 하나의 서열이 다른 것과 정렬되고 있음을 명시하고자 하는 경우, 상기 정렬은 종종 쌍별 정렬(pairwise alignment)로 불린다. 다중 서열 정렬은 일반적으로 둘 이상의 서열의 정렬을 가리키며, 이는 예를 들면, 일련의 쌍별 정렬에 의한 것을 포함한다. 일부 구현예에서, 정렬을 스코어링하는 것은 치환 및 인델의 확률에 대해 값을 설정하는 것을 포함한다. 각각의 염기가 정렬되는 경우, 매치 또는 미스매치는 치환 확률에 의한 정렬 스코어에 기여하며, 이는, 예를 들면, 매치의 경우 1 및 미스매치의 경우 0.33일 수 있다. 인델은 갭 패널티에 의해 정렬 스코어로부터 공제되며, 이는 예를 들면, -1일 수 있다. 갭 패널티 및 치환 확률은 서열들이 얼마나 돌연변이되는지에 관한 실험적 지식 또는 선형적 가정에 기초할 수 있다. 이러한 값들은 수득되는 정렬에 영향을 미친다. 정렬을 수행하는 알고리즘의 예는, 비제한적으로, 스미스-워터

맨(SW) 알고리즘, 니들맨-분쉬(NW) 알고리즘, 버로우-휠러 트랜스폼(BWT)에 기반한 알고리즘, 및 Novoalign(Novocraft Technologies; www.novocraft.com에서 이용가능함), ELAND(Illumina, San Diego, Calif.), SOAP(soap.genomics.org.cn에서 이용가능함), 및 Maq(maq.sourceforge.net에서 이용가능함)와 같은 해시 함수 정렬기(hash function aligner)를 포함한다. BWT 접근법을 실행하는 한 가지 예시적인 정렬 프로그램은 Geeknet (Fairfax, Va.)에 의해 유지된 SourceForge 웹사이트로부터 이용가능한 버로우-휠러 정렬기(BWA)이다. BWT는 전형적으로 뉴클레오타이드 당 2 비트의 메모리를 차지하므로, 전형적인 데스크탑 또는 노트북 컴퓨터로 4G 염기 정도의 큰 뉴클레오타이드 서열을 인덱스할 수 있다. 상기 전처리 BWT의 구축 (즉, 참조를 인덱싱) 및 지지하는 보조 데이터 구조의 구축을 포함한다. BWA는 모두 BWT에 기초한 2개의 상이한 알고리즘을 포함한다. BWA에 의한 정렬은 낮은 오류율 (<3%)로 최대 200 bp의 짧은 질의를 위해 설계된 알고리즘 bwa-short를 이용하여 진행될 수 있다 (Li H. 및 Durbin R. *Bioinformatics*, 25:1754-60 (2009)). 제2 알고리즘인 BWA-SW는 더 많은 오류로 긴 리드를 위해 설계되어 있다 (Li H. 및 Durbin R. (2010). *Fast and accurate long-read alignment with Burrows-Wheeler Transform*. *Bioinformatics*, Epub.). bwa-sw 정렬기는 종종 "bwa-long", "bwa long 알고리즘", 또는 유사한 용어로 지칭된다. 스미스-워터맨의 한 형태를 실행하는 정렬 프로그램은 Geeknet(Fairfax, Va.)에 의해 유지된 SourceForge 웹사이트로부터 이용가능한 MUMmer이다. MUMmer는 완전한 형태 또는 초안 형태이든 간에 전체 지놈을 빠르게 정렬하는 시스템이다 (Kurtz, S., et al., *Genome Biology*, 5:R12 (2004); Delcher, A. L., et al., *Nucl. Acids Res.*, 27:11 (1999)). 예를 들면, MUMmer 3.0은 2.4 GHz 리눅스 데스크탑 컴퓨터로, 78 MB의 메모리를 이용하여 13.7초만에 5-메가베이스 지놈의 쌍 간의 모든 20-염기 쌍 이상의 정확한 매치를 찾아낼 수 있다. MUMmer는 또한 불완전한 지놈을 정렬할 수 있고, 그것은 샷건 시퀀싱 프로젝트로부터 100s 또는 1000s의 컨티그(contig)를 쉽게 처리할 수 있으며, 이들을 상기 시스템에 포함된 NUCmer 프로그램을 이용하여 또 다른 세트의 컨티그 또는 지놈에 정렬할 것이다. 다른 정렬 프로그램의 비제한적인 예는 하기를 포함한다: Kent Informatics (Santa Cruz, Calif.)로부터의 BLAT (Kent, W. J., *Genome Research* 4: 656-664 (2002)); Beijing Genomics Institute (Beijing, Conn.) 또는 BGI Americas Corporation (Cambridge, Mass.)로부터의 SOAP2; Bowtie (Langmead, et al., *Genome Biology*, 10:R25 (2009)); Efficient Large-Scale Alignment of Nucleotide Databases (ELAND) 또는 ELANDv2 component of the Consensus Assessment of Sequence and Variation (CASAVA) 소프트웨어 (Illumina, San Diego, Calif.); Real Time Genomics, Inc. (San Francisco, Calif.)로부터의 RTG Investigator; Novocraft (Selangor, Malaysia)로부터의 Novoalign; Exonerate, European Bioinformatics Institute (Hinxton, UK) (Slater, G., 및 Birney, E., *BMC Bioinformatics* 6:31(2005)), University College Dublin (Dublin, Ireland)로부터의 Clustal Omega (Sievers F., et al., *Mol Syst Biol* 7, article 539 (2011)); University College Dublin (Dublin, Ireland)으로부터의 ClustalW 또는 ClustalX (Larkin M. A., et al., *Bioinformatics*, 23, 2947-2948 (2007)); 및 FASTA, European Bioinformatics Institute (Hinxton, UK) (Pearson W. R., et al., *PNAS* 85(8):2444-8 (1988); Lipman, D. J., *Science* 227(4693):1435-41 (1985)).

[0065] 전형적으로, 시퀀싱 데이터는 대규모 병렬 시퀀싱 반응으로부터 획득된다. 많은 차세대 고속대량 시퀀싱 시스템은 데이터를 FASTQ 파일로서 내보내나, 다른 포맷이 이용될 수도 있다. 일부 구현예에서, 반복 단위 길이 (예컨대, 모노머 길이), 원형화에 의해 형성된 접합부, 및 참조 서열에 대한 임의의 정확한 변이를 확인하기 위하여 서열이 전형적으로 서열 정렬을 통해 분석된다. 반복 단위 길이를 확인하는 것은 반복된 단위의 영역을 컴퓨팅하는 것, 상기 서열의 참조 유전자좌 (예컨대, 증폭, 농축, 및/또는 시퀀싱을 위해 하나 이상의 서열이 특히 표적화된 경우), 각각의 반복된 영역의 경계, 및/또는 각각의 시퀀싱 런 내의 반복의 수를 찾는 것을 포함할 수 있다. 서열 분석은 이중체의 두 가닥에 대한 서열 데이터를 분석하는 것을 포함할 수 있다. 상기 언급한 바와 같이, 일부 구현예에서, 샘플로부터의 상이한 폴리뉴클레오타이드(예컨대, 상이한 접합부를 갖는 원형화된 폴리뉴클레오타이드)로부터 리드의 서열들을 나타내는 동일한 변이는 확인된 변이로 간주된다. 일부 구현예에서, 만약 서열 변이가 동일한 폴리뉴클레오타이드의 하나를 초과하는 반복된 단위 내에 발생하면, 상기 서열 변이는 확인되거나 진정한 변이로 간주될 수 있는데, 상기 동일한 서열 변이는 동일한 연쇄체 내의 반복된 표적 서열에서 동일한 위치에 발생하지 않을 가능성이 높기 때문이다. 변이 및 확인된 변이를 확인하는데 있어 서열의 품질 스코어가 고려될 수 있으며, 예를 들면, 역치보다 낮은 품질 스코어를 갖는 서열 및 염기는 걸러내질 수 있다. 변이 판정의 민감도 및 특이도를 더 증가시키기 위해 다른 생물정보학적 방법이 사용될 수 있다.

[0066] 일부 구현예에서, 변이(돌연변이)의 결정 및 총 DNA 샘플에서 변이의 비율을 정량화하기 위해 통계적 분석이 적용될 수 있다. 특정한 염기의 총 측정은 시퀀싱 데이터를 이용하여 계산될 수 있다. 예를 들면, 이전 단계에서 계산된 정렬 결과로부터, "유효 리드"의 수, 즉 각각의 유전자좌에 대해 확인된 리드의 수를 계산할 수 있다. 변이의 대립유전자 빈도는 유전자좌에 대한 유효 리드 수에 의해 정규화될 수 있다. 종합 노이즈 수준, 즉 모든

유전자좌에 대해 관찰된 변이의 평균 비율이 계산될 수 있다. 다른 인자와 조합된, 변이의 빈도 및 종합 노이즈 수준이 변이 판정의 신뢰 구간을 결정하는데 사용될 수 있다. 푸아송 분포(Poisson distribution)와 같은 통계적 모델이 변이 판정의 신뢰 구간을 평가하는데 사용될 수 있다. 변이의 대립유전자 빈도 역시 총 샘플에서 변이의 상대적 양의 지시자로서 사용될 수 있다.

[0067]

일부 구현예에서, 미생물 오염물질은 판정 단계에 기초하여 확인된다. 예를 들면, 특정한 서열 변이가 잠재적 감염성 미생물에 의한 오염을 나타낼 수 있다. 서열 변이는 미생물을 확인하기 위해 고도로 보존된 폴리뉴클레오타이드 내에서 확인될 수 있다. 미생물의 계통발생적 특성규명 및 확인에 유용한 예시적인 고도로 보존된 폴리뉴클레오타이드는 16S rRNA 유전자, 23S rRNA 유전자, 5S rRNA 유전자, 5.8S rRNA 유전자, 12S rRNA 유전자, 18S rRNA 유전자, 28S rRNA 유전자, gyrB 유전자, rpoB 유전자, fusA 유전자, recA 유전자, coxI 유전자 및 nifD 유전자에서 발견되는 뉴클레오타이드 서열을 포함한다. 진핵생물의 경우, rRNA 유전자는 핵, 미토콘드리아, 또는 둘 모두일 수 있다. 일부 구현예에서, 16S-23S rRNA 유전자 내부 전사된 스페이서(ITS)에서의 서열 변이가 다른 rRNA 유전자를 사용하여 또는 사용하지 않고 밀접히 관련된 분류군(taxon)의 구분 및 확인하는데 사용될 수 있다. 16S rRNA의 구조적 제한으로 인해, 비-특이적 구획들이 높은 정도의 가변성을 가질 수 있음에도 불구하고, 상기 유전자 전역의 특정 영역들은 고도로 보존된 폴리뉴클레오타이드 서열을 갖는다. 서열 변이를 확인하는 것은 아속, 속, 아과, 과, 아목, 목, 아강, 강, 아문, 문, 아계, 또는 계에 해당하는 운영 분류 단위(OTU)를 확인하는데 사용될 수 있고, 선택적으로 집단에서의 이들의 빈도를 결정하는데 사용될 수 있다. 특정한 서열 변이의 검출은 오염을 나타내는 미생물의 존재, 및 선택적으로 양 (상대적 또는 절대적)을 검출하는데 사용될 수 있다. 예시적인 적용은 분변 또는 다른 오염에 대한 수질을 검사하는 것, 동물 또는 인간의 병원균에 대해 검사하는 것, 물 오염의 출처를 찾아내는 것, 재생수 또는 재활용수를 검사하는 것, 해양 방류 플룸(plume)을 포함하는 하수 방류를 검사하는 것, 병원균에 대한 양식 시설을 점검하는 것, 해변, 수영 장소 또는 다른 물과 관련된 레크레이션 시설을 점검하는 것, 및 독성 조류 최고점을 예측하는 것을 포함한다. 식품 감시 적용은 E. 콜라이 균주 O157:H7 또는 O111:B4, 리스테리아 모노시토제네스, 또는 살모넬라 엔테리카 아종 엔테리카 세로바르 엔테리티디스와 같은 식품 매개 병원균에 대해 식품 가공 공장에서의 생산 라인의 정기적 검사, 도축장을 점검하는 것, 음식점, 병원, 학교, 교도소 및 다른 기관의 부엌 및 음식 저장소를 조사하는 것을 포함한다. 조개류 및 조개류를 생산하는 물은 마비성 패독, 신경독성 패독, 설사성 패독 및 건망증성 패독의 원인이 되는 조류에 대해 조사될 수 있다. 또한, 수입 식품은 식품안전을 보장하기 위해 방출하기 전에 세관에 있는 동안 확인될 수 있다. 식물 병원균 점검 적용은 원예학 및 보육시설 점검, 예를 들면 참나무 급사병(Sudden Oak Death), 농작물 병원균 감시 및 질환 관리 및 산림 병원균 감시 및 질환 관리의 원인이 되는 미생물인 파이토포thora 라모룸(Phytophthora ramorum)에 대한 점검을 포함한다. 미생물 오염이 주된 안전 문제인 약제의 제조 환경, 의료 장비, 및 다른 소비재 또는 중요한 구성성분이 슈모노나스 애틀지노사, 또는 스타펠로코커스 아우레우스와 같은 특정 병원균의 존재, 인간과 연관된 보다 흔한 미생물, 물의 존재와 연관된 미생물 또는 특정한 환경에서 또는 비슷한 환경에서 이전에 확인된 생균수(bioburden)를 나타내는 다른 것의 존재에 대해 조사될 수 있다. 유사하게, 우주선을 포함하는 민감한 장비에 대한 제조 및 조립 지역이 상기 환경에서 서식하는 것으로 알려져 있거나 상기 환경 내로 가장 흔히 도입되는 이전에 확인된 미생물에 대해 점검될 수 있다.

[0068]

하나의 양태에서, 본 개시내용은 50 ng 미만의 폴리뉴클레오타이드들을 포함하는 핵산 샘플에서 서열 변이를 확인하는 방법을 제공하며, 각각의 폴리뉴클레오타이드는 5' 말단 및 3' 말단을 갖는다. 일부 구현예에서, 상기 방법은: (a) 상기 샘플 내의 각각의 폴리뉴클레오타이드를 리아가제를 이용하여 원형화시켜 복수의 원형 폴리뉴클레오타이드를 형성하는 단계; (b) 상기 리아가제를 상기 원형 폴리뉴클레오타이드로부터 분리시, 상기 원형 폴리뉴클레오타이드를 증폭시켜 연쇄체를 형성하는 단계; (c) 상기 연쇄체를 시퀀싱하여 복수의 시퀀싱 리드를 생성하는 단계; (d) 상기 복수의 시퀀싱 리드와 참조 서열 간의 서열 차이를 확인하는 단계; 및 (e) 50ng 미만의 폴리뉴클레오타이드의 상기 핵산 샘플로부터 상기 복수의 리드에서 0.05% 이상의 빈도로 발생하는 서열 차이를 서열 변이로서 판정하는 단계를 포함한다.

[0069]

샘플 내의 폴리뉴클레오타이드들의 출발량은 작을 수 있다. 일부 구현예에서, 출발하는 폴리뉴클레오타이드의 양은 50 ng 미만, 예컨대 45 ng, 40 ng, 35 ng, 30 ng, 25 ng, 20 ng, 15 ng, 10 ng, 5 ng, 4 ng, 3 ng, 2 ng, 1 ng, 0.5 ng, 0.1 ng 미만, 또는 그 미만이다. 일부 구현예에서, 출발하는 폴리뉴클레오타이드의 양은 0.1-100 ng, 예컨대 1-75 ng, 5 - 50 ng, 또는 10 - 20 ng의 범위이다. 일반적으로, 더 낮은 출발 물질은 다양한 처리 단계로부터 회수율을 증가시킬 중요성을 증가시킨다. 차후 반응에 참여하기 위한 샘플 내의 폴리뉴클레오타이드의 양을 감소시키는 공정은 회귀 돌연변이가 검출될 수 있는 민감도를 감소시킨다. 예를 들면, 문헌[Lou et al. (PNAS, 2013, 110 (49))]에 기재된 방법은 출발 물질의 10-20%만을 회수할 것으로 예상된다. 출발 물질의 양이 많은 경우 (예컨대, 실험실-배양된 박테리아로부터 정제된 경우), 이것은 실질적인 장애가 아닐 수

있다. 그러나, 출발 물질이 상당히 적은 샘플의 경우, 이러한 낮은 범위의 회수율은 충분한 희귀 변이를 검출하는데 실질적인 장애가 될 수 있다. 따라서, 일부 구현예에서, 본 개시내용의 방법에서 한 단계에서 다음 단계로의 샘플 회수율 (예컨대, 그 다음 증폭 단계 또는 시퀀싱 단계로의 입력을 위해 이용가능한 원형화 단계로의 입력의 질량비)은 약 또는 약 50%, 60%, 75%, 80%, 85%, 90%, 95% 이상, 또는 그 초과이다. 특정 단계로부터의 회수율은 100%에 근접할 수 있다. 회수율은 특정한 형태에 대한 것, 예컨대 비-원형 폴리뉴클레오타이드의 입력으로부터 원형 폴리뉴클레오타이드의 회수율일 수 있다.

[0070] 폴리뉴클레오타이드는 임의의 적합한 샘플, 예컨대 본 개시내용의 다양한 양태에 대해 본원에 기재된 샘플로부터 유래될 수 있다. 샘플로부터의 폴리뉴클레오타이드는 비제한적으로, DNA, RNA, 리보솜 RNA(rRNA), 운반 RNA(tRNA), 마이크로 RNA(miRNA), 메신저 RNA(mRNA), 이들의 임의의 단편, 또는 이들의 임의의 2개 이상의 조합을 포함하는, 다양한 폴리뉴클레오타이드 중 어느 것일 수 있다. 일부 구현예에서, 샘플은 DNA를 포함한다. 일부 구현예에서, 상기 폴리뉴클레오타이드는 수득된 또는 처리(예컨대, 변성)에 의한 단일 가닥이다. 적합한 폴리뉴클레오타이드의 추가적인 예가, 예컨대 본 개시내용의 임의의 다양한 양태에 대해 본원에 기재되어 있다. 일부 구현예에서, 폴리뉴클레오타이드는 추출 단계 없이, 및/또는 정제 단계 없이 차후 단계(예컨대, 원형화 및 증폭)를 거친다. 예를 들면, 유체 샘플은 추출 단계 없이 세포를 제거하도록 처리되어 정제된 액체 샘플 및 세포 샘플을 생성한 후, 상기 정제된 유체 샘플로부터 DNA를 단리할 수 있다. 예컨대 침전 또는 기질에의 비특이적 결합 후 상기 기질을 세척하여 결합된 폴리뉴클레오타이드를 방출시키는 것과 같은, 폴리뉴클레오타이드의 단리를 위한 다양한 과정이 이용가능하다. 폴리뉴클레오타이드가 세포 추출 단계 없이 샘플로부터 단리되는 경우, 폴리뉴클레오타이드는 주로 세포의 또는 "무세포" 폴리뉴클레오타이드일 것이며, 이는 사멸되거나 손상된 세포에 상응할 수 있다. 그러한 세포의 동일성은 미생물 군집에서와 같이 이들이 유래된 세포 또는 세포의 집단을 규명하는데 사용될 수 있다. 만약 샘플이, 예컨대 샘플 내의 세포로부터 폴리뉴클레오타이드를 추출하기 위해 처리되는 경우, 다양한 판정 방법이 이용가능하며, 이의 예가 본원에 제공되어 있다 (예컨대, 본 개시내용의 임의의 다양한 양태에 대해).

[0071] 핵산 샘플에서의 서열 변이는 임의의 다양한 서열 변이일 수 있다. 서열 변이의 다수의 비제한적인 예가, 예컨대 본 개시내용의 임의의 다양한 양태에 대해, 본원에 기재되어 있다. 일부 구현예에서 서열 변이는 단일 뉴클레오타이드 다형(SNP)이다. 일부 구현예에서, 서열 변이는 집단에서 낮은 빈도로 발생한다 ("희귀" 서열 변이로도 지칭됨). 예를 들면, 서열 변이는 약 5%, 4%, 3%, 2%, 1.5%, 1%, 0.75%, 0.5%, 0.25%, 0.1%, 0.075%, 0.05%, 0.04%, 0.03%, 0.02%, 0.01%, 0.005%, 0.001% 이하, 또는 그 미만의 빈도로 발생할 수 있다. 일부 구현예에서, 서열 변이는 약 0.1% 또는 그 미만의 빈도로 발생한다.

[0072] 일부 구현예에 따르면, 샘플의 폴리뉴클레오타이드는, 예컨대 리가아제의 사용에 의해, 원형화된다. 원형화는 폴리뉴클레오타이드의 5' 말단을 동일한 폴리뉴클레오타이드의 3' 말단에, 샘플 내의 또 다른 폴리뉴클레오타이드의 3' 말단에, 또는 상이한 공급원으로부터의 폴리뉴클레오타이드(예컨대, 인공 폴리뉴클레오타이드, 예컨대 올리고뉴클레오타이드 어댑터)의 3' 말단에 연결하는 것을 포함할 수 있다. 일부 구현예에서, 폴리뉴클레오타이드의 5' 말단은 동일한 폴리뉴클레오타이드의 3' 말단에 연결된다 ("자가-연결"로도 지칭됨). 원형화 공정 (예컨대, 어댑터 올리고뉴클레오타이드를 이용한 그리고 이용하지 않는), 시약 (예컨대, 어댑터의 유형, 리가아제의 사용), 반응 조건 (예컨대, 자가-연결에 유리한), 및 선택적인 부가적인 처리 (예컨대, 후-반응 정제)의 비제한적인 예가, 예컨대 본 개시내용의 임의의 다양한 양태에 대해 본원에 제공되어 있다.

[0073] 일반적으로, 폴리뉴클레오타이드의 말단을 서로 연결하여 원형 폴리뉴클레오타이드를 형성하는 것 (직접, 또는 하나 이상의 중간 어댑터 올리고뉴클레오타이드를 이용하여)은 접합부 서열을 갖는 접합부를 생성한다. 폴리뉴클레오타이드의 5' 말단 및 3' 말단이 어댑터 폴리뉴클레오타이드를 통해 연결되는 경우, 용어 "접합부"는 폴리뉴클레오타이드 및 어댑터 사이의 접합부 (예컨대, 5' 말단 접합부 또는 3' 말단 접합부 중 하나), 또는 어댑터 폴리뉴클레오타이드에 의해 형성되고 이를 포함하는 폴리뉴클레오타이드 5' 말단 및 3' 말단 사이의 접합부를 가리킬 수 있다. 폴리뉴클레오타이드의 5' 말단 및 3' 말단이 개입 어댑터 없이 연결되는 경우 (예컨대, 단일 가닥 DNA의 5' 말단 및 3' 말단), 용어 "접합부"는 이들 두 말단들이 연결되는 지점을 가리킨다. 접합부는 접합부를 포함하는 뉴클레오타이드의 서열 ("접합부 서열"로도 지칭됨)에 의해 확인될 수 있다. 일부 구현예에서, 샘플은 자연 분해 공정 (예컨대 세포 용해, 세포 사멸, 및 DNA가 세포로부터 그 주변 환경으로 방출되고, 그 곳에서 예컨대 무세포 폴리뉴클레오타이드로 추가 분해될 수 있는 다른 공정), 샘플 처리 (예컨대 고정화, 염색, 및/또는 보관 과정)의 부산물인 단편화, 및 특정 표적 서열에 대한 제한 없이 DNA를 절단하는 방법에 의한 단편화 (예컨대, 기계적 단편화, 예컨대 초음파처리; 비-서열 특이적 뉴클레아제 처리, 예컨대 DNase I, 단편화효소)에 의해 형성된 말단의 혼합물을 갖는 폴리뉴클레오타이드를 포함한다. 샘플이 말단의 혼합물을 갖

는 폴리뉴클레오타이드를 포함하는 경우, 2개의 폴리뉴클레오타이드가 동일한 5' 말단 또는 3' 말단을 가질 가능성은 낮으며, 2개의 폴리뉴클레오타이드가 독립적으로 동일한 5' 말단 및 3' 말단 모두를 가질 가능성은 극히 낮다. 따라서, 일부 구현예에서, 접합부는 상기 2개의 폴리뉴클레오타이드가 동일한 표적 서열을 갖는 부분을 포함하는 경우에도, 상이한 폴리뉴클레오타이드를 구별하는데 사용될 수 있다. 폴리뉴클레오타이드 말단이 개입 어댑터 없이 연결되는 경우, 접합부 서열은 참조 서열에 대한 정렬에 의해 확인될 수 있다. 예를 들면, 두 구성 성분 서열의 순서가 참조 서열에 대해 반전된 것으로 보이는 경우, 상기 반전이 발생한 것으로 보이는 지점이 상기 지점에서 접합부를 표시하는 것일 수 있다. 폴리뉴클레오타이드 말단이 하나 이상의 어댑터 서열을 통해 연결되는 경우, 접합부는 공지된 어댑터 서열에 대한 근접에 의해, 또는 만약 시퀀싱 리드가 원형화된 폴리뉴클레오타이드의 5' 및 3' 말단 모두로부터 서열을 획득하는데 충분한 길이이면 상기와 같이 정렬에 의해 확인될 수 있다. 일부 구현예에서, 특정한 접합부의 형성은 충분히 희귀한 사건이어서 그것은 샘플의 원형화된 폴리뉴클레오타이드 중에서 독특하다.

[0074] 원형화 후, 반응 산물은 서브시퀀스 단계에 참여하기 위해 이용가능한 원형화된 폴리뉴클레오타이드의 상대적 농도 또는 순도를 증가시키기 위해 (예컨대, 원형 폴리뉴클레오타이드의 단리 또는 반응에서 하나 이상의 다른 분자의 제거에 의해) 증폭 또는 시퀀싱 이전에 정제될 수 있다. 예를 들면, 원형화 반응 또는 이의 구성성분은, 예컨대 엑소뉴클레아제의 처리에 의해, 단일 가닥 (비-원형화된) 폴리뉴클레오타이드를 제거하도록 처리될 수 있다. 추가적인 예로서, 원형화 반응 또는 이의 부분은 크기 배제 크로마토그래피를 거칠 수 있고, 이에 의해 작은 시약이 유지되고 폐기되거나 (예컨대, 처리되지 않은 어댑터), 또는 원형화 산물은 보유되고 별개의 부피로 방출된다. 라이게이션 반응을 세척하기 위한 다양한 키트, 예컨대 Zymo Research에 의해 제조된 Zymo 올리고 정제 키트에 의해 제공된 키트가 이용가능하다. 일부 구현예에서, 정제는 원형화 반응에서 사용되는 리가아제를 제거하거나 분해하는, 및/또는 그러한 리가아제로부터 원형화된 폴리뉴클레오타이드를 정제하는 처리를 포함한다. 일부 구현예에서, 리가아제를 분해하는 처리는 프로테아제를 이용한 처리를 포함한다. 적합한 프로테아제는 원핵생물, 바이러스, 및 진핵생물로부터 이용가능하다. 프로테아제의 예는 프로테이나아제 K (트리티라키움 알부민 유래), 프로나제 E (스트렙토마이세스 그리세우스 유래), 바실러스 폴리믹사 프로테아제, 테로몰리신 (호열성 박테리아 유래), 트립신, 서브틸리신, 퓨린 등을 포함한다. 일부 구현예에서, 프로테아제는 프로테이나아제 K이다. 프로테아제 처리는 제조사 프로토콜을 따를 수 있거나, 또는 표준 조건 (예컨대, Sambrook and Green, Molecular Cloning: A Laboratory Manual, 4th Edition (2012)에서 제공됨)을 거친다. 또한 프로테아제 처리 후에 추출 및 침전이 뒤따를 수 있다. 한 가지 예에서, 원형화된 폴리뉴클레오타이드는 0.1% SDS 및 20 mM EDTA의 존재하에 프로테이나아제 K (Qiagen) 처리에 의해 정제되고, 1:1 페놀/클로로포름 및 클로로포름으로 추출되며, 에탄올 또는 이소프로판올로 침전된다. 일부 구현예에서, 침전은 에탄올에서 이뤄진다.

[0075] 본 개시내용의 다른 양태에 대해 기술된 바와 같이, 원형화 이후에 곧바로 상기 원형화된 폴리뉴클레오타이드를 시퀀싱할 수 있다. 대안적으로, 시퀀싱 전에 하나 이상의 증폭 반응이 수행될 수 있다. 폴리뉴클레오타이드(예컨대, DNA 및/또는 RNA)를 증폭하는 다양한 방법이 이용가능하다. 증폭은 선형, 지수일 수 있거나, 또는 다단계 증폭 공정에서 선형 및 지수기 모두를 포함할 수 있다. 증폭 방법은 온도 변화, 예컨대 열 변성 단계를 포함할 수 있거나, 또는 열 변성을 요구하지 않는 등온 공정일 수 있다. 적합한 증폭 공정의 비제한적인 예가, 예컨대 본 개시내용의 임의의 다양한 양태에 대해 본원에 기재되어 있다. 일부 구현예에서, 증폭은 롤링 서클 증폭(RCA)을 포함한다. 본원의 다른 곳에 기재된 바와 같이, 전형적인 RCA 반응 혼합물은 하나 이상의 프라이머, 폴리머라아제, 및 dNTP를 포함하고, 연쇄체를 생성한다. 전형적으로, RCA 반응에서의 폴리머라아제는 가닥 치환 활성을 갖는 폴리머라아제이다. 그러한 다양한 폴리머라아제가 이용가능하며, 이의 비제한적인 예는 엑소뉴클레아제가 결합된 DNA 폴리머라아제 I 거대(Klenow) 단편, Phi29 DNA 폴리머라아제, Taq DNA 폴리머라아제 등을 포함한다. 일반적으로, 연쇄체는 주형 폴리뉴클레오타이드로부터 표적 서열의 2개 이상의 카피 (예컨대, 표적 서열의 약 2, 3, 4, 5, 6, 7, 8, 9, 10 이상, 또는 이를 초과하는 카피; 일부 구현예에서, 약 2 카피, 또는 이의 초과)를 포함하는 폴리뉴클레오타이드 증폭 산물이다. 증폭 프라이머는 임의의 적합한 길이, 예컨대 약 또는 적어도 약 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 90, 100, 또는 그 이상의 뉴클레오타이드일 수 있으며, 이의 임의의 부분 또는 모두가 상기 프라이머가 혼성화하는 상응하는 표적 서열에 상보적일 수 있다 (예컨대, 약, 또는 적어도 약 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 또는 그 이상의 뉴클레오타이드). 무작위 프라이머, 표적-특이적 프라이머, 및 어댑터-표적화된 프라이머를 사용하는 것과 같은 다양한 RCA 공정의 예가 본원에 기재되어 있으며, 이들 중 일부가 도 7에 예시되어 있다.

[0076] 원형화된 폴리뉴클레오타이드는 시퀀싱 이전에 증폭되는 경우 (예컨대, 연쇄체를 생성하기 위해), 증폭된 산물은 농축 없이 곧바로, 또는 하나 이상의 농축 단계 이후에 시퀀싱을 거칠 수 있다. 적합한 농축 공정의 비제한적인 예는, 예컨대 본 개시내용의 임의의 다양한 양태에 대해 본원에 기재되어 있다 (예컨대, 제2 증폭 단계를

위한 B2B 프라이머의 사용). 일부 구현예에 따르면, 원형화된 폴리뉴클레오타이드 (또는 선택적으로 농축될 수 있는, 이의 증폭 산물)는 시퀀싱 반응을 거쳐 시퀀싱 리드를 생성한다. 상기 방법에 의해 생성된 시퀀싱 리드는 본원에 개시된 다른 방법에 따라 사용될 수 있다. 다양한 시퀀싱 방법, 특히 고속대량 시퀀싱 방법이 이용가능하다. 예는, 비제한적으로, 일루미나에 의해 제작된 시퀀싱 시스템 (HiSeq® 및 MiSeq®과 같은 시퀀싱 시스템), 라이프 테크놀로지스 (Ion Torrent®, SOLiD®, 등), 로슈의 454 라이프 사이언시스 시스템, 퍼시픽 바이오사이언시스 시스템 등을 포함한다. 일부 구현예에서, 시퀀싱은 약 50, 75, 100, 125, 150, 175, 200, 250, 300 이상, 또는 이를 초과하는 뉴클레오타이드 길이의 리드를 생성하기 위해 HiSeq® 및 MiSeq® 시스템을 사용하는 것을 포함한다. 증폭 플랫폼 및 방법의 추가적인 비제한적인 예가, 예컨대 본 개시내용의 임의의 다양한 양태에 대해 본원에 기재되어 있다.

[0077] 일부 구현예에 따르면, 만약 시퀀싱 리드와 참조 서열 간의 서열 차이가 적어도 2개의 상이한 폴리뉴클레오타이드(예컨대, 상이한 접합부를 갖는 결과로서 구별될 수 있는, 2개의 상이한 원형 폴리뉴클레오타이드)에서 발생하면, 그것은 진정한 서열 변이(예컨대, 증폭 또는 시퀀싱 이전에 샘플에 존재하고 이러한 어느 공정의 결과가 아님)로 불린다. 증폭 또는 시퀀싱 오류의 결과인 서열 변이는 동일한 표적 서열을 포함하는 2개의 상이한 폴리뉴클레오타이드 상에 정확하게 (예컨대, 위치 및 유형) 복제되지 않을 가능성이 있기 때문에, 이러한 검증 파라미터를 추가하는 것은 잘못된 서열 변이의 배경을 크게 감소시키며, 동시에 샘플 내의 실제 서열 변이를 검출하는 민감도 및 정확도를 증가시킨다. 일부 구현예에서, 약 5%, 4%, 3%, 2%, 1.5%, 1%, 0.75%, 0.5%, 0.25%, 0.1%, 0.075%, 0.05%, 0.04%, 0.03%, 0.02%, 0.01%, 0.005%, 0.001% 이하, 또는 그 미만의 빈도를 갖는 서열 변이는 배경을 충분히 초과하여 정확한 판정을 허용한다. 일부 구현예에서, 서열 변이는 약 0.1% 또는 그 미만의 빈도로 일어난다. 일부 구현예에서, 상기 방법은 약 0.0005% 내지 약 3%, 예컨대 0.001%-2%, 또는 0.01%-1%의 범위의 빈도를 갖는 상기 서열 차이를 진정한 서열 변이로 판정하는 것을 포함한다. 일부 구현예에서, 빈도가 배경 오류율 (예컨대, 약 0.05, 0.01, 0.001, 0.0001, 또는 그 이하 또는 그 미만의 p-값)을 통계학적으로 유의하게 초과할 때 서열 변이의 빈도는 충분히 배경을 초과한다. 일부 구현예에서, 서열 변이의 빈도가 배경 오류율을 약 또는 적어도 약 2배, 3배, 4배, 5배, 6배, 7배, 8배, 9배, 10배, 25배, 50배, 100배, 또는 그 이상 초과할 때 (예컨대, 적어도 5배 초과), 상기 빈도는 서열 변이의 빈도는 배경을 충분히 초과한다. 일부 구현예에서, 주어진 위치에서 서열을 정확히 결정하는데 있어서 배경 오류율은 약 1%, 0.5%, 0.1%, 0.05%, 0.01%, 0.005%, 0.001%, 0.0005% 이하, 또는 그 미만이다. 일부 구현예에서, 오류율은 0.001%보다 낮다. 빈도 및 오류율을 결정하는 방법은, 예컨대 본 개시내용의 임의의 다양한 양태에 대해 본원에 기재되어 있다.

[0078] 일부 구현예에서, 진정한 서열 변이를 확인하는 것("판정하는 것" 또는 "판정을 행하는 것"으로도 지칭됨)은 하나 이상의 시퀀싱 리드를 참조 서열과 최적으로 정렬하여 이 둘 간의 차이를 확인하는 것뿐만 아니라 접합부를 확인하는 것을 포함한다. 일반적으로, 정렬은 하나의 서열을 또 다른 서열을 따라 위치시키고, 각각의 서열을 따라 반복적으로 겹을 도입하고, 상기 2개의 서열이 얼마나 매치하는지 스코어를 매기고, 바람직하게는 참조를 따라 다양한 위치에 대해 반복하는 것을 포함한다. 최적-스코어링 매치는 정렬인 것으로 간주되며 서열들 간의 관련성 정도에 관한 추론을 나타낸다. 다양한 정렬 알고리즘 및 이들을 실행하는 정렬기가 이용가능하며, 이의 비제한적인 예는, 예컨대 본 개시내용의 임의의 다양한 양태에 대해 본원에 기재되어 있다. 일부 구현예에서, 시퀀싱 리드가 비교되는 참조 서열은 공지된 참조 서열, 예컨대 참조 지놈 (예컨대, 대상과 동일한 종의 멤버의 지놈)이다. 참조 지놈은 완전하거나 불완전할 수 있다. 일부 구현예에서, 참조 지놈은, 예컨대 참조 지놈으로부터의 또는 분석 중인 시퀀싱 리드로부터 생성된 컨센서스로부터의 표적 폴리뉴클레오타이드를 함유하는 영역만으로 이루어진다. 일부 구현예에서, 참조 서열은 하나 이상의 유기체의 폴리뉴클레오타이드의 서열, 예컨대 하나 이상의 박테리아, 조류, 바이러스, 원생동물, 진균류, 또는 다른 유기체로부터의 서열을 포함하거나 이들로 이루어진다. 일부 구현예에서, 참조 서열은 참조 지놈의 일부, 예컨대 분석 중인 하나 이상의 표적 서열에 상응하는 영역 (예컨대, 하나 이상의 유전자, 또는 이의 일부)으로만 이루어진다. 예를 들면, 병원균의 검출을 위해 (예컨대 오염 검출의 경우), 참조 지놈은 병원균 (예컨대, HIV, HPV, 또는 유해한 박테리아 균주, 예컨대, E. 콜라이)의 전체 지놈, 또는, 예컨대 특정한 균주 또는 혈청형의 확인에 유용한 이의 일부이다. 일부 구현예에서, 시퀀싱 리드는, 예컨대 다수의 상이한 유기체 또는 균주를 스크리닝하기 위해, 다수의 상이한 참조 서열에 정렬된다. 서열 차이가 확인될 수 있는 (그리고 서열 변이가 판정될 수 있는) 참조 서열의 추가적인 비제한적인 예가 본원에 기재된, 예컨대 본 개시내용의 임의의 다양한 양태에 대해 본원에 기재되어 있다.

[0079] 하나의 양태에서, 본 개시내용은 반응 혼합물에서 표적 서열의 2개 이상의 카피를 포함하는 복수의 상이한 연쇄체를 증폭하는 방법을 제공하며, 여기서 상기 표적 서열은 5'에서 3' 방향으로 배향된 서열 A 및 서열 B를 포함한다. 일부 구현예에서, 상기 방법은 상기 반응 혼합물을 핵산 증폭 반응시키는 것을 포함하고, 상기 반응 혼합물은 하기를 포함한다: (a) 복수의 연쇄체, 상기 복수의 연쇄체 내의 각각의 연쇄체는 5' 말단 및 3' 말단을 갖

는 각각의 폴리뉴클레오타이드를 원형화시킴으로써 형성된 상이한 접합부를 포함하고; (b) 서열 A'를 포함하는 제1 프라이머, 여기서 상기 제1 프라이머는 서열 A와 서열 A' 간의 서열 상보성을 통해 표적 서열의 서열 A에 특이적으로 결합하며; (c) 서열 B를 포함하는 제2 프라이머, 여기서 상기 제2 프라이머는 서열 B와 B' 간의 서열 상보성을 통해 표적 서열의 보체를 포함하는 상보적 폴리뉴클레오타이드에 존재하는 서열 '에 특이적으로 결합하고; 및 (d) 상기 제1 프라이머 및 상기 제2 프라이머를 연장시켜 증폭된 폴리뉴클레오타이드를 생성하는 폴리머라아제; 여기서 표적 서열의 서열 A의 5' 말단과 서열 B의 3' 말단 사이의 거리는 75nt 이하이다.

[0080] 관련된 양태에서, 본 개시내용은 반응 혼합물에서 표적 서열을 포함하는 복수의 상이한 원형 폴리뉴클레오타이드를 증폭하는 방법을 제공하며, 여기서 상기 표적 서열은 5'에서 3' 방향으로 배향된 서열 A 및 서열 B를 포함한다. 일부 구현예에서, 상기 방법은 상기 반응 혼합물을 핵산 증폭 반응시키는 것을 포함하고, 여기서 상기 반응 혼합물은 하기를 포함한다: (a) 복수의 원형 폴리뉴클레오타이드, 여기서 상기 복수의 원형 폴리뉴클레오타이드 내의 각각의 원형 폴리뉴클레오타이드는 5' 말단 및 3' 말단을 갖는 각각의 폴리뉴클레오타이드를 원형화시킴으로써 형성되는 상이한 접합부를 포함하고; (b) 서열 A'를 포함하는 제1 프라이머, 여기서 상기 제1 프라이머는 서열 A와 서열 A' 간의 서열 상보성을 통해 표적 서열의 서열 A에 특이적으로 혼성화하고; (c) 서열 B를 포함하는 제2 프라이머, 여기서 상기 제2 프라이머는 서열 B와 B' 간의 서열 상보성을 통해 표적 서열의 보체를 포함하는 상보적 폴리뉴클레오타이드에 존재하는 서열 B'에 특이적으로 결합하며; 및 (d) 상기 제1 프라이머 및 상기 제2 프라이머를 연장시켜 증폭된 폴리뉴클레오타이드를 생성하는 폴리머라아제; 여기서 서열 A 및 서열 B는 내인성 서열이며, 표적 서열의 서열 A의 5' 말단과 서열 B의 3' 말단 사이의 거리는 75nt 이하이다.

[0081] 원형 폴리뉴클레오타이드 또는 연쇄체를 증폭시키는 간에, 상기 폴리뉴클레오타이드는 임의의 적합한 샘플 공급원으로부터 유래될 수 있다(직접적, 또는 간접적, 예컨대 증폭에 의해). 다양한 적합한 샘플 공급원, 선택적인 추출 공정, 폴리뉴클레오타이드의 유형, 및 서열 변이의 유형이, 예컨대 본 개시내용의 임의의 다양한 양태에 대해 본원에 기재되어 있다. 원형 폴리뉴클레오타이드는 비-원형 폴리뉴클레오타이드를 원형화시키는 것으로부터 유래될 수 있다. 원형화 공정 (예컨대, 어댑터 올리고뉴클레오타이드를 이용하고 이용하지 않는), 시약 (예컨대, 어댑터의 유형, 리가아제의 사용), 반응 조건 (예컨대, 자가-연결이 유리한), 선택적인 부가적인 처리 (예컨대, 후-반응 정제), 및 이에 의해 형성된 접합부의 비제한적인 예가, 예컨대 본 개시내용의 임의의 다양한 양태에 대해 본원에 제공되어 있다. 연쇄체는 원형 폴리뉴클레오타이드의 증폭으로부터 유래될 수 있다. 폴리뉴클레오타이드 (예컨대, DNA 및/또는 RNA)를 증폭시키는 다양한 방법이 이용가능하며, 이의 비제한적인 예가 또한 본원에 기재되었다. 일부 구현예에서, 연쇄체는 원형 폴리뉴클레오타이드의 롤링 서클 증폭에 의해 생성된다.

[0082] 도 10은 단일 반복 (이는 원형이 아니면 전형적으로 증폭되지 않을 것임) 및 표적 서열의 다수의 카피를 포함하는 연쇄체의 문맥에서 표적 서열에 대한 제1 및 제2 프라이머의 예시적 배열을 예시한다. 본원에 기재된 다른 양태에 대해 언급된 바와 같이, 프라이머의 이러한 배열은 "백투백"(B2B) 또는 "역전된" 프라이머로 지칭될 수 있다. B2B 프라이머를 이용한 증폭은 원형 및/또는 연쇄체 주형의 농축을 용이하게 한다. 더욱이, 상대적으로 더 작은 풋프린트(한쌍의 프라이머가 포괄하는 총 거리)와 조합된 이러한 방향은 표적 서열 주위의 더 다양한 단편화 사건의 증폭을 허용하는데, 접합부가 전형적인 증폭 반응에서 발견되는 프라이머의 배열에서보다 프라이머 사이에서 발생하지 않을 가능성이 있기 때문이다 (서로 마주보며, 타겟 핵산 서열을 포괄함). 일부 구현예에서, 서열 A의 5' 말단 및 서열 B의 3' 말단 사이의 거리는 약 200, 150, 100, 75, 50, 40, 30, 25, 20, 15 이하, 또는 그 미만의 뉴클레오타이드이다. 일부 구현예에서, 서열 A는 서열 B의 보체이다. 일부 구현예에서, 복수의 상이한 표적 서열에 관한 다수의 쌍의 B2B 프라이머가 복수의 상이한 표적 서열을 병렬로 (예컨대, 약 또는 적어도 약 10, 50, 100, 150, 200, 250, 300, 400, 500, 1000, 2500, 5000, 10000, 15000 이상의 상이한 표적 서열) 증폭시키기 위해 동일한 반응에서 사용된다. 프라이머는, 예컨대 본원의 다른 곳에 기재된 바와 같은, 임의의 적합한 길이일 수 있다. 증폭은 적절한 조건 하에 임의의 적합한 증폭 반응, 예컨대 본원에 기재된 증폭 반응을 포함할 수 있다. 일부 구현예에서, 증폭은 폴리머라아제 연쇄 반응이다.

[0083] 일부 구현예에서, B2B 프라이머는 적어도 2개의 서열 요소, 즉 서열 상보성을 통해 표적 서열에 혼성화하는 제1 요소, 및 제1 요소가 혼성화하는 제1 혼성화 온도에서의 제1 증폭 과정 동안 표적 서열에 혼성화하지 않는 5' "테일" (예컨대, 테일 및 제1 요소가 결합하는 곳에 대해 바로 3'에 있는 표적 서열의 부분 사이의 서열 상보성의 부족으로 인해)을 포함한다. 예를 들면, 제1 프라이머는 서열 A'에 대해 서열 C 5'를 포함하고, 제2 프라이머는 서열 B에 대해 서열 D 5'를 포함하며, 서열 C 및 서열 D 어느 것도 제1 혼성화 온도에서의 제1 증폭 과정 동안 복수의 연쇄체 (또는 원형 폴리뉴클레오타이드)에 혼성화하지 않는다. 그러한 테일을 가진 프라이머가 사용되는 일부 구현예에서, 증폭은 제1 과정 및 제2 과정을 포함할 수 있고; 상기 제1 과정은 제1 및 제2 프라

이머가 연쇄체 (또는 원형 폴리뉴클레오타이드)에 혼성화하는 제1 온도에서의 혼성화 단계 및 프라이머 연장을 포함하며; 제2 과정은 제1 및 제2 프라이머가 연장된 제1 또는 제2 프라이머, 또는 이의 보체를 포함하는 증폭 산물에 혼성화하는, 제1 온도보다 더 높은 제2 온도에서의 혼성화 과정, 및 프라이머 연장을 포함한다. 상기 두 온도 각각에서의 증폭 사이클의 수는 원하는 산물에 기초하여 조정될 수 있다. 전형적으로, 제1 온도는 상대적으로 적은 수의 사이클, 예컨대 약 15, 10, 9, 8, 7, 6, 5 이하, 또는 그 미만의 사이클 동안 사용될 것이다. 더 높은 온도에서의 사이클의 수는 제1 온도에서의 사이클의 수와 독립적으로 선택될 수 있지만, 전형적으로 더 많은 사이클, 예컨대 약 또는 적어도 약 5, 6, 7, 8, 9, 10, 15, 20, 25, 또는 그 이상의 사이클일 것이다. 더 높은 온도는 단지 프라이머 내의 제1 요소 및 연쇄체 내의 내부 표적 서열 간의 혼성화에 의해 형성된 더 짧은 단편보다 프라이머 연장에서 제1 요소 및 프라이머의 테일 요소 간의 혼성화에 유리하다. 따라서, 상기 2-과정 증폭은 짧은 증폭 산물이 유리할 수 있는 정도를 낮추기 위해 사용될 수 있으며, 이에 의해 표적 서열의 2개 이상의 카피를 갖는 증폭 산물의 상대적으로 더 높은 비율을 유지한다. 예를 들면, 제2 온도에서의 혼성화 및 프라이머 연장의 5 사이클(예컨대, 적어도 5, 6, 7, 8, 9, 10, 15, 20 이상의 사이클) 후, 반응 혼합물 내의 증폭된 폴리뉴클레오타이드의 적어도 5% (예컨대, 적어도 5%, 6%, 7%, 8%, 9%, 10%, 15%, 20%, 25%, 30%, 또는 그 이상)는 표적 서열의 2개 이상의 카피를 포함한다. 이러한 2-과정의 테일을 가진 B2B 프라이머 증폭 공정에 따른 구현예의 예시가 도 11에 예시되어 있다.

[0084] 일부 구현예에서, 증폭은 연쇄체로부터 앰플리콘의 길이를 증가시키기 위해 치우친 조건 하에서 이뤄진다. 예를 들면, 프라이머 농도는 낮아질 수 있으며, 이로써 모든 프라이밍 부위가 프라이머에 혼성화하는 것은 아니어서 PCR 산물을 더 길게 만든다. 유사하게, 상기 사이클 동안 프라이머 혼성화 시간을 감소시키는 것은 유사하게 더 적은 프라이머가 혼성화하게 하여, 또한 평균 PCR 앰플리콘 크기를 증가시킬 것이다. 더욱이, 상기 사이클의 온도 및/또는 연장 시간을 증가시키는 것은 유사하게 PCR 앰플리콘의 평균 길이를 증가시킬 것이다. 이들 기술의 임의의 조합이 사용될 수 있다.

[0085] 일부 구현예에서, 특히 B2B 프라이머를 이용한 증폭이 수행되는 경우, 증폭 산물은 연쇄체를 포함하는 혼합물로부터 모노머의 수를 감소시키고/거나 제거하기 위해 크기에 기초하여 수득된 앰플리콘을 여과하도록 처리된다. 이것은 비제한적으로, 겔로부터의 단편 절제 및 겔 여과 (예컨대, 약 300, 400, 500 이상, 또는 그보다 큰 뉴클레오타이드 길이의 단편을 농축시키기 위해)를 포함하는 다양한 이용가능한 기술뿐만 아니라 결합 버퍼 농도를 미세조정함으로써 크기 선택하기 위한 SPRI 비드(Agencourt AMPure XP)를 이용하여 수행될 수 있다. 예를 들면, DNA 단편과 혼합하는 동안 0.6x 결합 버퍼의 사용이 약 500 염기쌍(bp)보다 큰 DNA 단편을 우선적으로 결합하기 위해 사용될 수 있다.

[0086] 일부 구현예에서, 제1 프라이머는 서열 A'에 대해 서열 C 5'를 포함하고, 상기 제2 프라이머는 서열 B에 대해 서열 D 5'를 포함하며, 서열 C 및 서열 D 어느 것도 제1 혼성화 온도에서의 제1 증폭 과정 동안 복수의 원형 폴리뉴클레오타이드에 혼성화하지 않는다. 증폭은 제1 과정 및 제2 과정을 포함할 수 있고; 상기 제1 과정은 프라이머 연장 전에 상기 제1 및 제2 프라이머가 원형 폴리뉴클레오타이드 또는 이의 증폭 산물에 혼성화하는, 제1 온도에서의 혼성화 과정을 포함하며; 상기 제2 단계는 상기 제1 및 제2 프라이머가 연장된 제1 또는 제2 프라이머 또는 이의 보체를 포함하는 증폭 산물에 혼성화하는, 제1 온도보다 더 높은 제2 온도에서의 혼성화 과정을 포함한다. 예를 들면, 상기 제1 온도는 서열 A', 서열 B의 약 T_m , 또는 이들의 평균, 또는 그 초과, 또는 이들 T_m 중 하나보다 1°C, 2°C, 3°C, 4°C, 5°C, 6°C, 7°C, 8°C, 9°C, 10°C 이상을 초과하는 온도로서 선택될 수 있다. 이 예에서, 상기 제2 온도는 조합된 서열 (A' + C), 조합된 서열 (B + D)의 약 T_m 또는 그 초과, 또는 이들의 평균, 또는 이들 T_m 중 하나보다 1°C, 2°C, 3°C, 4°C, 5°C, 6°C, 7°C, 8°C, 9°C, 10°C 이상을 초과하는 온도가 되도록 선택될 수 있다. 용어 " T_m "은 "용융 온도"로도 지칭되며 일반적으로 참조 서열로 이루어지는 올리고뉴클레오타이드 (이는 사실상 더 큰 폴리뉴클레오타이드 내의 하위-서열일 수 있음) 및 그의 상보적 서열의 50%가 혼성화되는 (또는 분리되는) 온도를 가리킨다. 일반적으로, T_m 은 길이가 증가함에 따라 증가하며, 그와 같이 서열 A'의 T_m 은 조합 서열 (A' + C)의 T_m 보다 더 낮은 것으로 예상된다.

[0087] 하나의 양태에서, 본 개시내용은 본 개시내용의 방법에 따른 방법을 수행하기 위한 반응 혼합물을 제공한다. 상기 반응 혼합물은 본원에 기재된 방법에 기재된 반응 혼합물을 포함하는, 임의의 다양한 방법에 대해 본원에 기재된 바와 같은 다양한 구성성분 중 하나 이상을 포함할 수 있다. 일부 구현예에서, 상기 반응 혼합물은 표적 서열의 2개 이상의 카피를 포함하는 복수의 상이한 연쇄체, 또는 표적 서열의 하나 이상의 카피를 포함하는 원형 폴리뉴클레오타이드 (예컨대, 원형 모노머)를 증폭시키기 위한 혼합물이며, 상기 표적 서열은 5'에서 3' 방향으로 배향된 서열 A 및 서열 B를 포함하고, 상기 반응 혼합물은 하기를 포함한다: (a) 복수의 연쇄체 (또는 원형 폴리뉴클레오타이드), 여기서 상기 복수의 연쇄체 내의 각각의 연쇄체 (또는 원형 폴리뉴클레오타이드)는

5' 말단 및 3' 말단을 갖는 각각의 폴리뉴클레오타이드를 원형화시킴으로써 형성되는 상이한 접합부를 포함하고; (b) 서열 A'를 포함하는 제1 프라이머, 여기서 상기 제1 프라이머는 서열 A와 서열 A' 간의 서열 상보성을 통해 표적 서열의 서열 A에 특이적으로 혼성화하며; (c) 서열 B를 포함하는 제2 프라이머, 여기서 상기 제2 프라이머는 B와 B' 간의 서열 상보성을 통해 표적 서열의 보체를 포함하는 상보적 폴리뉴클레오타이드에 존재하는 서열B'에 특이적으로 혼성화하고; 및 (d) 상기 제1 프라이머 및 상기 제2 프라이머를 연장시켜 증폭된 폴리뉴클레오타이드를 생성하는 폴리머라아제; 여기서 표적 서열의 서열 A의 5' 말단 및 서열 B의 3' 말단 간의 거리는 75nt 이하이다. 샘플, 폴리뉴클레오타이드, 프라이머, 폴리머라아제, 다른 시약, 및 반응 조건은, 예컨대 임의의 다양한 양태에 대해 본원에 기재된 어느 것일 수 있으며, 이는 임의의 적합한 조합으로 반응 혼합물에 포함될 수 있다. 일부 구현예에서, 상기 제1 프라이머는 서열 A'에 대해 서열 C 5'를 포함하고, 상기 제2 프라이머는 서열 B에 대해 서열 D 5'를 포함하며, 서열 C 및 서열 D 어느 것도 증폭 반응에서의 제1 증폭 과정 동안 2개 이상의 연쇄체에 혼성화하지 않는다.

[0088] 하나의 양태에서, 본 개시내용은, 예컨대 본 개시내용의 임의의 다양한 다른 양태에서, 본원에 기재된 방법에 유용한, 또는 이에 의해 생성된 조성물을 제공한다. 일부 구현예에서, 상기 조성물은 단일 가닥이며 실질적으로 리가아제가 없는 복수의 원형화된 폴리뉴클레오타이드를 포함한다. 일부 구현예에서, 상기 조성물은 복수의 연쇄체를 포함하고, 여기서 복수의 연쇄체는 10000개 이하의 표적 폴리뉴클레오타이드의 군에 상응하며, 추가로 상기 복수의 연쇄체 내의 각각의 연쇄체는 하기를 특징으로 한다: (a) 이들이 서열 반복의 2개 이상의 카피를 포함하고, 여기서 상기 카피 모두가 동일한 표적 폴리뉴클레오타이드에 상응하는 점; 및 (b) 하나의 각각의 연쇄체의 서열 반복의 2개 이상의 카피 사이의 접합부가 상기 조성물 내의 또 다른 각각의 연쇄체와 상이한 점. 샘플, 폴리뉴클레오타이드, 프라이머, 폴리머라아제, 및 다른 시약은, 예컨대 임의의 다양한 양태에 대해, 본원에 기재된 것 중 어느 것일 수 있으며, 이는 임의의 적합한 조합으로 조성물 내에 포함될 수 있다. 상기 조성물은 하나 이상의 표적 서열을 증폭하기 위해, 설계된, 프라이머, 예컨대 본원에 기재된 B2B 프라이머의 하나 이상의 쌍을 포함할 수 있다. 조성물은 키트의 형태로 제공될 수 있다. 키트 내의 시약 및 다른 물질은 임의의 적합한 용기 내에 함유될 수 있고, 바로 사용가능한 형태이거나 키트 내의 다른 시약 또는 사용자에게 의해 공급된 시약과의 조합을 요구할 수 있다 (예컨대, 농축된 조성물의 희석 또는 동결건조된 조성물의 재구성). 키트는 버퍼를 제공할 수 있으며, 이의 비제한적인 예는 탄산나트륨, 중탄산나트륨 버퍼, 보레이트 버퍼, 트리스 버퍼, MOPS 버퍼, HEPES 버퍼, 및 이의 조합을 포함한다. 키트는 임의의 다양한 양태에 대해 본원에 기재된 하나 이상의 방법의 수행을 위한 지시를 추가로 포함할 수 있다. 지시는 하나 이상의 언어 (예컨대, 2, 3, 4, 5 이상의 언어)로 제공될 수 있다.

[0089] 하나의 양태에서, 본 개시내용은 서열 변이를 검출하기 위한 시스템을 제공한다. 일부 구현예에서, 상기 시스템은 (a) 샘플에 대한 검출 반응을 수행해 달라는 사용자 요청을 수신하도록 구성된 컴퓨터; (b) 사용자 요청에 대한 반응으로 샘플 또는 이의 일부 상에서 핵산 증폭 반응을 수행하는 증폭 시스템, 여기서 상기 증폭 반응은 (i) 각각의 폴리뉴클레오타이드를 원형화시켜 복수의 원형 폴리뉴클레오타이드를 형성하는 단계로서, 이들 각각은 5' 말단과 3' 말단 사이에 접합부를 가지며; 및 (ii) 상기 원형 폴리뉴클레오타이드를 증폭시키는 단계를 포함하며; (c) 증폭 시스템에 의해 증폭된 폴리뉴클레오타이드에 대한 시퀀싱 리드를 생성하고 시퀀싱 리드와 참조 서열 사이의 서열 차이를 확인하고, 상이한 접합부를 갖는 적어도 2개의 원형 폴리뉴클레오타이드에서 발생하는 서열 차이를 서열 변이로서 판정하는 시퀀싱 시스템; 및 (d) 수신자에게 보고서를 보내는 보고서 생성기, 여기서 상기 보고서는 서열 변이의 검출에 대한 결과를 포함한다. 일부 구현예에서, 상기 수신자는 사용자이다. 도 32는 본 개시내용의 방법에 유용한 시스템의 비제한적인 예를 예시한다.

[0090] 시스템에서 사용하기 위한 컴퓨터는 하나 이상의 프로세서를 포함할 수 있다. 프로세서는 하나 이상의 컨트롤러, 계산 유닛, 및/또는 컴퓨터 시스템의 다른 유닛과 결합될 수 있거나, 또는 원하는 대로 펌웨어에서 실행될 수 있다. 소프트웨어에서 실행되는 경우, 루틴이 어느 컴퓨터 판독가능 메모리에, 예컨대 RAM, ROM, 플래시 메모리, 자기 디스크, 레이저 디스크, 또는 다른 적합한 저장 매체에 저장될 수 있다. 마찬가지로, 이러한 소프트웨어는, 예를 들면, 통신 채널, 예컨대 전화선, 인터넷, 무선 연결 등으로, 또는 전송가능 매체, 예컨대 컴퓨터 판독가능 디스크, 플래시 드라이브 등을 통하는 것을 포함하는, 임의의 공지된 방법을 통해 계산 장치로 전달될 수 있다. 다양한 단계들이 다양한 블록, 작업, 도구, 모듈 및 기술로서 실행될 수 있고, 이는 결국 하드웨어, 펌웨어, 소프트웨어, 또는 하드웨어, 펌웨어, 및/또는 소프트웨어의 임의의 조합으로 실행될 수 있다. 하드웨어에서 실행되는 경우, 상기 블록, 작업, 기술 등의 일부 또는 전부가, 예를 들면, 사용자 집적 회로(IC), 주문형 집적 회로(application specific integrated circuit (ASIC)), 필드 프로그래머블 로직 어레이(field programmable logic array (FPGA)), 프로그래머블 로직 어레이(PLA), 등에서 실행될 수 있다. 클라이언트-서버, 관계형 데이터베이스 아키텍처가 상기 시스템의 구현예에서 사용될 수 있다. 클라이언트-서버

아키텍처는 네트워크 상의 각각의 컴퓨터 또는 프로세스가 클라이언트 또는 서버 중 하나인 네트워크 아키텍처이다. 서버 컴퓨터는 디스크 드라이브 (파일 서버), 프린터 (프린트 서버), 또는 네트워크 트래픽 (네트워크 서버)을 관리하기 위한 전형적으로 강력한 컴퓨터이다. 클라이언트 컴퓨터는 사용자가 애플리케이션뿐만 아니라 본원에 개시된 바와 같은 예시적 출력 장치를 작동시키는 PC (퍼스널 컴퓨터) 또는 워크스테이션이다. 클라이언트 컴퓨터는 리소스, 예컨대 파일, 장치, 및 심지어 프로세싱 파워를 위해 서버 컴퓨터에 의존한다. 일부 구현 예에서, 상기 서버 컴퓨터는 모든 데이터베이스 기능을 다룬다. 클라이언트 컴퓨터는 모든 프론트-엔드 데이터 관리를 다루는 소프트웨어를 가질 수 있고, 또한 사용자로부터 입력된 데이터를 받을 수 있다.

[0091] 상기 시스템은 샘플에 대해 검출 반응을 수행해 달라는 사용자 요청을 수신하도록 구성될 수 있다. 사용자 요청은 직접적이거나 간접적일 수 있다. 직접적인 요청의 예는 입력 장치, 예컨대 키보드, 마우스, 또는 터치 스크린과 같은 입력 장치에 의해 전송된 것을 포함한다. 직접적인 요청의 예는 통신 매체, 예컨대 인터넷 (유선 또는 무선)을 통한 전송을 포함한다.

[0092] 상기 시스템은 사용자 요청의 요청에 대응하여 샘플 또는 이의 일부에 대해 핵산 증폭 반응을 수행하는 증폭 시스템을 추가로 포함할 수 있다. 폴리뉴클레오타이드 (예컨대, DNA 및/또는 RNA)를 증폭시키는 다양한 방법이 이용가능하다. 증폭은 선형, 지수이거나, 또는 다-단계 증폭 공정에서 선형 및 지수기 모두를 포함할 수 있다. 증폭 방법은 온도 변화, 예컨대 열 변성 단계를 수반할 수 있거나, 또는 열 변성을 필요로하지 않는 등온 공정일 수 있다. 적합한 증폭 공정의 비제한적인 예는 본 개시내용의 임의의 다양한 양태에 대해서와 같이, 본원에 기재되어 있다. 일부 구현예에서, 증폭은 롤링 서클 증폭(RCA)을 포함한다. 폴리뉴클레오타이드를 증폭시키기 위한 다양한 시스템이 이용가능하며, 이는 수행되는 증폭 반응의 유형에 따라 달라질 수 있다. 예를 들면, 온도 변화의 사이클을 포함하는 증폭 반응의 경우, 상기 증폭 시스템은 열순환기(thermocycler)를 포함할 수 있다. 증폭 시스템은 실시간 증폭 및 검출 기기, 예컨대 어플라이드 바이오시스템스, 로슈, 및 스트라타진에 의해 제조된 시스템을 포함할 수 있다. 일부 구현예에서, 상기 증폭 반응은 (i) 각각의 폴리뉴클레오타이드를 원형화시켜 복수의 원형 폴리뉴클레오타이드를 형성하는 단계로서, 이들 각각은 5' 말단과 3' 말단 사이에 접합부를 가지며; 및 (ii) 상기 원형 폴리뉴클레오타이드를 증폭시키는 단계를 포함한다. 샘플, 폴리뉴클레오타이드, 프라이머, 폴리머라아제, 및 다른 시약은, 예컨대 임의의 다양한 양태에 대해, 본원에 기재된 것 중 어느 것일 수 있다. 원형화 공정 (예컨대, 어댑터 올리고뉴클레오타이드를 이용하고 이용하지 않는), 시약 (예컨대, 어댑터의 유형, 리가아제의 사용), 반응 조건 (예컨대, 자가-연결을 선호하는), 선택적인 부가적인 공정 (예컨대, 후-반응 정제), 및 이에 의해 형성된 접합부의 비제한적인 예가, 예컨대 본 개시내용의 임의의 다양한 양태에 관해, 본원에 제공되어 있다. 시스템은 임의의 상기 방법을 수행하도록 선택되고/거나 설계될 수 있다.

[0093] 시스템은 상기 증폭 시스템에 의해 증폭된 폴리뉴클레오타이드에 대한 시퀀싱 리드를 생성하고, 시퀀싱 리드와 참조 서열 간의 서열 차이를 확인하고, 상이한 접합부를 갖는 적어도 2개의 원형 폴리뉴클레오타이드에 존재하는 서열 차이를 서열 변이로서 판정하는 단계를 추가로 포함할 수 있다. 상기 시퀀싱 시스템 및 상기 증폭 시스템은 동일하거나, 중복되는 장비를 포함할 수 있다. 예를 들면, 증폭 시스템 및 시퀀싱 시스템 모두는 동일한 열순환기를 이용할 수 있다. 상기 시스템에서 사용하기 위한 다양한 시퀀싱 플랫폼이 이용가능하며, 이는 선택된 시퀀싱 방법에 기초하여 선택될 수 있다. 시퀀싱 방법의 예가 본원에 기재되어 있다. 증폭 및 시퀀싱은 액체 처리기(liquid handler)의 사용을 포함할 수 있다. 상업적으로 이용가능한 몇 가지 액체 처리 시스템이 이들 공정의 자동화를 실행하기 위해 이용될 수 있다 (예를 들면, Perkin-Elmer, Beckman Coulter, Caliper Life Sciences, Tecan, Eppendorf, Apricot Design, Velocity 11로부터의 액체 처리기). 다양한 자동화된 시퀀싱 기기가 상업적으로 이용가능하며, 라이프 테크놀로지스 (고체 플랫폼, 및 pH-기반 검출), 로슈 (454 플랫폼), 일루미나 (예컨대, 유세포 기반 시스템, 예컨대 지놈 분석기 장치)에 의해 제조된 시퀀서를 포함한다. 2, 3, 4, 5 개 이상의 자동화된 장치 사이에서의 이송(예컨대, 하나 이상의 액체 처리기 및 시퀀싱 장치 사이에서)은 수동이거나 자동일 수 있다.

[0094] 서열 차이를 확인하고 참조 서열에 대해 서열 변이를 판정하는 방법은, 예컨대 본 개시내용의 임의의 다양한 양태에 대해 본원에 기재되어 있다. 시퀀싱 시스템은 전형적으로 시퀀싱 데이터의 입력 및 원하는 파라미터의 입력(예컨대, 참조 지놈의 선택)에 대한 반응으로 이들 단계를 수행하는 소프트웨어를 포함할 것이다. 이들 알고리즘을 실행하는 정렬 알고리즘 및 정렬기의 예가 본원에 기재되어 있으며, 이는 시퀀싱 시스템의 일부를 형성할 수 있다.

[0095] 상기 시스템은 수신자에게 보고서를 보내는 보고서 생성기를 추가로 포함할 수 있으며, 상기 보고서는 서열 변이의 검출 결과를 포함한다. 보고서는, 공정이 진행됨에 따라 주기적 업데이트와 함께, 예컨대 시퀀싱 리드 동안 또는 시퀀싱 데이터가 분석되면서 실시간으로 생성될 수 있다. 또한, 또는 대안적으로, 보고서는 분석의 중

결시에 생성될 수 있다. 보고서는 상기 시퀀싱 시스템이 모든 서열 변이를 판정하는 단계를 완료할 때, 자동으로 생성될 수 있다. 일부 구현예에서, 보고서는 사용자로부터의 지시에 대한 반응으로 생성된다. 서열 변이의 검출 결과 외에도, 보고서는 하나 이상의 서열 변이에 기초한 분석을 포함할 수 있다. 예를 들면, 하나 이상의 서열 변이가 특정한 오염물질 또는 표현형과 연관된 경우, 보고서는 이러한 연관성과 관련된 정보, 예컨대 오염물질 또는 표현형이 어떤 수준으로 존재할 가능성, 및 선택적으로 이러한 정보에 기초한 제안(예컨대, 추가 시험, 관찰, 또는 규제 조치)을 포함할 수 있다. 보고서는 임의의 다양한 형태일 수 있다. 본 개시내용과 관련된 데이터는 수신 및/또는 수신자에 의한 검토를 위해 상기 네트워크 또는 연결(또는 비제한적으로 물리적 보고서, 예컨대 프린트 출력물을 메일로 보내는 것을 포함하는, 정보를 전송하기 위한 임의의 다른 적합한 수단)로 전송될 수 있음이 구상된다. 수신자는 비제한적으로 개인, 또는 전자 시스템(예컨대, 하나 이상의 컴퓨터, 및/또는 하나 이상의 서버)일 수 있다.

[0096] 하나의 양태에서, 본 개시내용은 하나 이상의 프로세서에 의해 실행시, 서열 변이를 검출하는 방법을 실행하는 코드를 포함하는 컴퓨터 판독가능 매체를 제공한다. 일부 구현예에서, 상기 실행된 방법은: (a) 샘플에 대해 검출 반응을 수행해 달라는 고객 요청을 수신하는 단계; (b) 고객 요청에 대한 반응으로 샘플 또는 이의 일부에 대해 핵산 증폭 반응을 수행하는 단계로서, 상기 증폭 반응은 (i) 각각의 폴리뉴클레오타이드를 원형화시켜 복수의 원형 폴리뉴클레오타이드를 형성하는 단계로서, 이들 각각은 5' 말단과 3' 말단 사이에 접합부를 갖고; (ii) 상기 원형 폴리뉴클레오타이드를 증폭시키는 단계를 포함하고; (c) (i) 상기 증폭 반응에서 증폭된 폴리뉴클레오타이드에 대한 시퀀싱 리드를 생성하는 단계; (ii) 시퀀싱 리드와 참조 서열 간의 서열 차이를 확인하는 단계; 및 (iii) 상이한 접합부를 갖는 적어도 2개의 원형 폴리뉴클레오타이드에서 발생하는 서열 차이를 서열 변이로서 판정하는 단계를 포함하는 시퀀싱 분석을 수행하는 단계; 및 (d) 서열 변이의 검출 결과를 포함하는 보고서를 생성하는 단계를 포함한다.

[0097] 컴퓨터-실행가능한 코드를 포함하는 기계 판독가능 매체는 비제한적으로, 유형의 저장 매체, 반송파(carrier wave) 매체 또는 물리적 전송 매체를 포함하는, 많은 형태를 취할 수 있다. 비-휘발성 저장 매체는, 예를 들면, 광학 또는 자기 디스크, 예컨대 어느 컴퓨터 내의 임의의 저장 장치 등을 포함하고, 예컨대 데이터베이스 등을 실행하는데 사용될 수 있다. 휘발성 저장 매체는 동적 메모리, 예컨대 그러한 컴퓨터 플랫폼의 주 메모리를 포함한다. 유형의 전송 매체는 동축 케이블; 컴퓨터 시스템 내의 버스를 포함하는 선을 포함하는, 구리선 및 광섬유를 포함한다. 반송파 전송 매체는 전기 또는 전자기 신호, 또는 음파 또는 광파, 예컨대 무선 주파수(RF) 및 적외선(IR) 데이터 통신 중에 생성되는 것의 형태를 취할 수 있다. 따라서, 컴퓨터 판독가능 매체의 일반적인 형태는 예를 들면: 플로피 디스크, 플렉서블 디스크, 하드 디스크, 자기 테이프, 임의의 다른 자기 매체, CD-ROM, DVD 또는 DVD-ROM, 임의의 다른 광학 매체, 펀치 카드 종이 테이프, 구멍 패턴을 갖는 임의의 다른 물리적 저장 매체, RAM, ROM, PROM 및 EPROM, FLASH-EPROM, 임의의 다른 메모리 칩 또는 카트리지, 데이터 또는 지시를 전송하는 반송파, 그러한 전송파를 전송하는 케이블 또는 링크, 또는 컴퓨터가 프로그래밍 코드 및/또는 데이터를 판독하는 임의의 다른 매체를 포함한다. 이러한 형태의 많은 컴퓨터 판독가능 매체는 하나 이상의 지시의 하나 이상의 시퀀스를 실행을 위한 프로세서에 전달하는데 관여할 수 있다.

[0098] 대상 컴퓨터-실행가능 코드는 서버, PC, 또는 모바일 장치, 예컨대 스마트폰 또는 태블릿을 포함하는, 프로세서를 포함하는 임의의 적합한 장치에서 실행될 수 있다. 임의의 컨트롤러 또는 컴퓨터는 선택적으로 모니터를 포함하며, 이는 음극선관("CRT") 디스플레이, 평면 패널 디스플레이(예컨대, 능동 매트릭스형 액정 디스플레이, 액정 디스플레이 등) 등일 수 있다. 컴퓨터 회로망은 종종 박스에 놓이며, 이는 많은 집적 회로 칩, 예컨대 마이크로프로세서, 메모리, 인터페이스 회로 등을 포함한다. 상기 박스는 또한 선택적으로 하드 디스크 드라이브, 플로피 디스크 드라이브, 고용량 제거가능 드라이브, 예컨대 기록가능 CD-ROM, 및 다른 일반적인 주변 요소를 포함한다. 입력 장치, 예컨대 키보드, 마우스, 또는 터치 민감 스크린은 선택적으로 사용자로부터의 입력을 돕는다. 컴퓨터는 파라미터 필드의 세트 내로의 사용자 입력의 형태, 예컨대, GUI, 또는 미리 프로그램된 지시의 형태, 예컨대, 다양한 상이한 특정 작업에 대해 미리 프로그램된 지시의 형태의 사용자 지시를 수신하기 위한 적절한 소프트웨어를 포함할 수 있다.

[0099] 본원에 개시된 임의의 다양한 양태의 일부 구현예에서, 상기 방법, 조성물, 및 시스템은, 예컨대 환자 샘플의 규명 및 선택적으로 대상의 상태의 진단에서 치료적 적용을 갖는다. 치료적 적용은 또한 본원에 기재된 방법의 결과에 기초하여, 환자가 가장 반응할 수 있는 요법의 선택을 통지하는 것("치료진단(theranostics)")으로도 지칭됨), 및 이를 필요로 하는 대상의 실제적인 치료를 포함할 수 있다. 특히, 본원에 개시된 방법 및 조성물은, 특히 분석된 폴리뉴클레오타이드가 cfDNA, ctDNA, 또는 단편화된 종양 DNA를 포함하거나 이들로 이루어진 경우, 종양 존재, 종양의 진행 및/또는 전이를 진단하는데 사용될 수 있다. 일부 구현예에서, 대상은 치료

효능에 대해 관찰된다. 예를 들면, 시간 경과에 따라 ctDNA를 관찰함으로써, ctDNA의 감소가 효과적인 치료의 표시로서 사용될 수 있는 반면, 증가는 상이한 치료 또는 상이한 투여의 선택을 촉진할 수 있다. 다른 용도는 이식 수용체에서의 장기 거부 평가 (이식 공여체 지능에 상응하는 순환 DNA의 양의 증가가 이식 거부의 조기 표지자로서 사용되는 경우), 및 병원균 감염, 예컨대 바이러스 또는 박테리아 감염의 유전자형 분석/아이소타입 분석을 포함한다. 순환 태아 DNA에서의 서열 변이의 검출은 태아의 상태를 진단하는데 사용될 수 있다.

[0100] 본원에 사용된 바와 같이, "치료" 또는 "치료하는" 또는 "완화시키는" 또는 "개선하는"은 상호교환적으로 사용된다. 이들 용어는 비제한적으로 치료적 이점 및/또는 예방적 이점을 포함하는 유익하거나 원하는 결과를 달성하기 위한 접근법을 가리킨다. 치료적 이점은 치료 중인 하나 이상의 질환, 병태, 또는 증상에서의 어떤 치료적으로 관련된 개선 또는 이에 대한 효과를 의미한다. 예방적 이점을 위해, 조성물은, 질환, 조건, 또는 증상이 아직 나타나지 않았을 수 있음에도, 특정한 질환, 병태, 또는 증상이 생길 위험성이 있는 대상에게, 또는 질환의 하나 이상의 생리학적 증상을 보고하는 대상에게 투여될 수 있다. 전형적으로, 예방적 이점은 치료 중인 하나 이상의 질환, 병태, 또는 증상의 발병을 및/또는 악화를 낮추는 것 (예컨대, 치료된 집단 및 치료되지 않은 집단 간에, 또는 대상의 치료된 상태 및 치료되지 않은 상태 간에)을 포함한다. 치료 결과를 개선하는 것은 대상을 하나 이상의 치료제, 또는 다른 치료적 개입 (예컨대 수술)을 이용한 치료로부터 이익을 얻거나 이익을 얻지 않을 대상으로서 확인하기 위해 대상의 병태를 진단하는 것을 포함할 수 있다. 상기 치료적 적용에 있어서, 하나 이상의 치료제를 이용한 성공적인 치료의 총 비율은 본 개시내용의 방법에 따른 진단 없이 그룹화된 환자 사이에서의 효과 대비 개선될 수 있다 (예컨대, 적어도 약 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 95% 이상의 치료 효능의 측정에서의 개선).

[0101] 용어 "대상" "개체" 및 "환자"는 척추동물, 바람직하게는 포유동물, 더 바람직하게는 인간을 가리키기 위해 본원에서 상호교환적으로 사용된다. 포유동물은 비제한적으로, 쥐과, 원숭이, 인간, 가축, 스포츠 동물, 및 애완동물을 포함한다. 생체내에서 수득되거나 시험관내에서 배양된 생물학적 개체의 조직, 세포, 및 이들의 자손 역시 포함된다.

[0102] 용어 "요법제", "치료가능제" 또는 "치료제"는 상호교환적으로 사용되며, 대상에게 투여시 상당한 유익한 효과를 제공하는 분자 또는 화합물을 가리킨다. 상기 유익한 효과는 진단 결정의 가능; 질환, 증상, 장애, 또는 병리학적 상태의 개선; 질환, 증상, 장애 또는 병태의 개시를 감소시키거나 예방하는 것; 및 일반적으로 질환, 증상, 장애 또는 병리학적 병태를 대항하는 것을 포함한다.

[0103] 본원에 기재된 다양한 방법의 일부 구현예에서, 샘플은 대상 유래이다. 대상은 임의의 유기체일 수 있으며, 임의의 비제한적인 예는 식물, 동물, 진균류, 원생동물, 모나라게, 바이러스, 미토콘드리아, 및 염색체를 포함한다. 샘플 폴리뉴클레오타이드는, 예를 들면, 배양된 세포주, 생검, 혈액 샘플, 볼 스왑, 또는 세포를 함유하는 유체 샘플 (예컨대, 타액)을 포함하는, 세포 샘플, 조직 샘플, 체액 샘플, 또는 장기 샘플 (또는 이들 중 어느 것으로부터 유래된 세포 배양물)과 같이, 대상으로부터 단리될 수 있다. 일부 경우에, 상기 샘플은 온전한 세포를 포함하지 않거나, 세포를 제거하기 위해 처리되거나, 또는 폴리뉴클레오타이드가 세포 추출 단계 없이 단리된다 (예컨대, 무세포 폴리뉴클레오타이드, 예컨대 무세포 DNA와 같은, 무세포 폴리뉴클레오타이드를 분리하기 위해). 샘플 공급원의 다른 예는 혈액, 소변, 대변, 콧구멍, 폐, 내장, 다른 체액 또는 배설물, 이들로부터 유래된 물질, 또는 이의 조합으로부터의 것들을 포함한다. 대상은 비제한적으로, 소, 돼지, 마우스, 랫트, 닭, 고양이, 개 등을 포함하는 동물일 수 있고, 보통 포유동물, 예컨대 인간일 수 있다. 일부 구현예에서, 샘플은 대상으로부터의 종양 조직의 샘플에서와 같이, 종양 세포를 포함한다. 일부 구현예에서, 샘플은 혈액 샘플 또는 임의의 일부 (예컨대, 혈액 혈장 또는 혈청)이다. 혈청 및 혈장은 상기 조직 중에서도 악성 세포 사멸의 높은 비율과 연관된 종양 DNA에 대한 상대적 농축으로 인해, 특히 관심 대상일 수 있다. 샘플은 신선한 샘플, 또는 하나 이상의 보관 공정을 거친 샘플 (예컨대, 파라핀-포매된 샘플, 특히포르말린-고정된 파라핀-포매된 (FFPE) 샘플)일 수 있다. 일부 구현예에서, 하나의 개체로부터의 샘플은 다수의 별개의 샘플 (예컨대, 2, 3, 4, 5, 6, 7, 8, 9, 10 이상의 별개의 샘플)로 분할되며, 이는 이반복, 삼반복, 사반복 이상의 분석과 같이, 독립적으로 본 개시내용의 방법을 거친다. 샘플이 대상으로부터 유래되는 경우, 참조 서열 역시 분석 중인 샘플로부터의 콘센서스 서열 또는 동일한 대상의 또 다른 샘플 또는 조직으로부터의 폴리뉴클레오타이드의 서열과 같이, 대상으로부터 유래될 수 있다. 예를 들면, 혈액 샘플은 ctDNA 돌연변이에 대해 분석될 수 있는 반면, 또 다른 샘플 (예컨대, 구강 또는 피부 샘플)로부터의 세포 DNA는 참조 서열을 결정하기 위해 분석된다.

[0104] 폴리뉴클레오타이드는 임의의 적합한 방법에 따라, 샘플 내의 세포로부터 추출을 이용하여 또는 추출 없이, 샘플로부터 추출될 수 있다. 다양한 키트가 폴리뉴클레오타이드의 추출을 위해 이용가능하며, 임의의 선택은 샘플의 유형, 또는 단리될 핵산의 유형에 좌우될 수 있다. 추출 방법의 예가 본원에 개시된 임의의 다양한 양태에 대해

기술된 것과 같이 본원에 제시되어 있다. 하나의 예에서, 샘플은 혈액 샘플, 예컨대 EDTA 튜브(예컨대, BD Vacutainer)에서 수집된 샘플일 수 있다. 혈장은 원심분리(예컨대, 4℃에서 1900xg에서 10분)에 의해 말초 혈액 세포로부터 분리될 수 있다. 6mL 혈액 샘플에 대해 이러한 방식으로 수행된 혈장 분리는 전형적으로 2.5 내지 3 mL의 혈장을 생성한다. 순환 무세포 DNA는 제조사의 프로토콜에 따라, 예컨대 QIAmp 순환 핵산 키트(Qiagen)를 이용하여 혈장 샘플로부터 추출될 수 있다. 그리고 나서, DNA는 정량될 수 있다(예컨대, 고민감도 DNA 키트(Agilent)를 이용하여 Agilent 2100 Bioanalyzer 상에서). 예로써, 건강한 사람으로부터의 이러한 혈장 샘플로부터의 순환 DNA의 수율은 혈장 mL 당 1ng 내지 10ng의 범위일 수 있으며, 암 환자 샘플에서는 유의미하게 더 많다.

[0105] 폴리뉴클레오타이드는 또한 보관된 샘플, 예컨대 냉동되거나 보관된 샘플로부터 유래될 수 있다. 샘플을 보관하는 한 가지 일반적인 방법은 이들을 포르말린-고정하고 파라핀-포매하는 것이다. 그러나, 이러한 공정은 또한 핵산의 분해와 연관되어 있다. FFPE 샘플로부터 처리되고 분석된 폴리뉴클레오타이드는 짧은 폴리뉴클레오타이드, 예컨대 50-200 염기쌍, 또는 더 짧은 염기쌍의 범위의 단편을 포함할 수 있다. 제W02007133703호에 기재된 것, 및 문헌[Foss, et al Diagnostic Molecular Pathology, (1994) 3:148-155 및 Paska, C., et al Diagnostic Molecular Pathology, (2004) 13:234-240]에 의해 기술된 방법과 같이, 고정된 파라핀-포매된 샘플로부터 핵산을 정제하기 위한 많은 기술들이 존재한다. Ambion의 Recoverall 총 핵산 단리 키트와 같이 상업적으로 이용가능한 키트가 FFPE 샘플로부터 폴리뉴클레오타이드를 정제하는데 사용될 수 있다. 전형적인 방법은 자일렌 또는 다른 유기 용매를 이용한 추출을 통해 조직으로부터 파라핀을 제거하는 단계로 시작한 후, 조직 및 단백질을 절단하고 조직으로부터 지능 물질을 방출하는 것을 돕는 열 및 프로테이나아제 K와 같은 프로테아제로 처리한다. 그리고 나서, 상기 방출된 핵산은 막 위에 포획될 수 있거나 또는 용액으로부터 침전되고, 불순물을 제거하기 위해 세척되고, mRNA 단리의 경우, 원치않는 DNA를 분해시키기 위해 DNase 처리 단계가 종종 추가된다. FFPE DNA를 추출하는 다른 방법이 이용가능하며, 본 개시내용의 방법에서 사용될 수 있다.

[0106] 일부 구현예에서, 상기 복수의 폴리뉴클레오타이드는 무세포 폴리뉴클레오타이드, 예컨대 무세포 DNA(cfDNA) 또는 순환 종양 DNA(ctDNA)를 포함한다. 무세포 DNA는 건강한 개체 및 병에 걸린 개체 모두에서 순환한다. 종양으로부터의 cfDNA(ctDNA)는 어느 특정 암 유형에 국한되지 않지만, 상이한 악성종양에 걸쳐 공통적으로 발견되는 것으로 보인다. 일부 측정에 따르면, 혈장 내에서 자유롭게 순환 DNA 농도는 대조군 대상에서 약 14-18 ng/mL이며, 종양형성을 갖는 환자에서 약 180-318 ng/mL이다. 아포토시스성 및 괴사성 세포 사멸은 체액 내의 무세포 순환 DNA에 기여한다. 예를 들면, 유의미하게 증가된 순환 DNA 수준은 전립선암 환자 및 다른 전립선 질환, 예컨대 양성 전립선 비대증 및 전립선염의 혈장에서 관찰되어 왔다. 또한, 순환 종양 DNA는 일차 종양이 발생하는 장기로부터 비롯된 유체에 존재한다. 따라서, 유방암 검출은 유관 세척물(ductal lavage)에서 달성될 수 있고; 결장암 검출은 대변에서; 폐암 검출은 객담에서, 그리고 전립선암 검출은 소변 또는 정액에서 달성될 수 있다. 무세포 DNA는 다양한 공급원으로부터 수득될 수 있다. 한 가지 흔한 공급원은 대상의 혈액 샘플이다. 그러나, cfDNA 또는 다른 단편화된 DNA는 다양한 다른 공급원으로부터 유래될 수 있다. 예를 들면, 소변 및 대변 샘플이 ctDNA를 포함하는 cfDNA의 공급원일 수 있다.

[0107] 일부 구현예에서, 폴리뉴클레오타이드는 추출 단계 없이, 및/또는 정제 단계 없이 그 다음 단계(예컨대, 원형화 및 증폭)를 거친다. 예를 들면, 유체 샘플은 정제된 액체 샘플 및 세포 샘플을 생성하기 위해 추출 단계 없이 세포를 제거하도록 처리된 후, 상기 정제된 유체 샘플로부터 DNA를 단리할 수 있다. 침전 또는 기질에의 비특이적 결합 후 기질을 세척하여 결합된 폴리뉴클레오타이드를 방출시키는 것과 같은, 폴리뉴클레오타이드의 단리를 위한 다양한 과정이 이용가능하다. 폴리뉴클레오타이드가 세포 추출 단계 없이 샘플로부터 단리되는 경우, 폴리뉴클레오타이드는 주로 세포의 또는 "무세포" 폴리뉴클레오타이드일 것이다. 예를 들면, 무세포 폴리뉴클레오타이드는 무세포 DNA("순환" DNA로도 불림)를 포함할 수 있다. 일부 구현예에서, 순환 DNA는 종양 세포, 예컨대 체액 또는 배설물(예컨대, 혈액 샘플)로부터의 순환 종양 DNA(ctDNA)이다. 종양은 자주 아포토시스 또는 괴사를 나타내며, 이로써 종양 핵산이 상이한 형태 및 상이한 수준으로, 다양한 기전을 통해, 대상의 혈류를 포함하는 체내로 방출된다. 전형적으로, ctDNA의 크기는 고농도의 작은 단편, 일반적으로 70 내지 200 뉴클레오타이드 길이 내지 저농도의 최대 1000kb의 큰 단편의 범위일 수 있다.

[0108] 본원에 기재된 임의의 다양한 양태의 일부 구현예에서, 서열 변이를 검출하는 것은 참조 서열에 대해 또는 돌연변이가 없는 배경에서 돌연변이(예컨대, 희귀 체세포 돌연변이)를 검출하는 것을 포함하고, 여기서 상기 서열 변이는 질환과 상호관련된다. 일반적으로, 질환 또는 특성과 연관된 통계학적, 생물학적, 및/또는 기능적 증거가 있는 서열 변이는 "원인 유전 변이"로 지칭된다. 하나의 원인 유전 변이는 하나를 초과하는 질환 또는 형질과 연관될 수 있다. 일부 구현예에서, 원인 유전 변이는 만델 형질, 비-만델 형질, 또는 이 둘과 연관될 수 있

다. 원인 유전 변이는 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 50 이상의 서열 차이(예컨대 원인 유전 변이를 포함하는 폴리뉴클레오타이드 및 동일한 상대적 지놈 위치에서 원인 유전 변이가 결여된 폴리뉴클레오타이드 사이에서)와 같이, 폴리뉴클레오타이드에서의 변이로서 나타날 수 있다. 원인 유전 변이의 유형의 비제한적인 예는 단일 뉴클레오타이드 다형(SNP), 결실/삽입 다형(DIP), 카피수 변이(CNV), 짧은 직렬 반복(STR), 제한 단편 길이 다형(RFLP), 단순 서열 반복(SSR), 가변적인 수의 직렬 반복(VNTR), 무작위로 증폭된 다형성 DNA(RAPD), 증폭된 단편 길이 다형(AFLP), 인터-레트로트랜스포손 증폭된 다형(IRAP), 길고 짧은 산재 요소(LINE/SINE), 긴 직렬 반복(LTR), 이동성 요소, 레트로트랜스포손 미세위성 증폭된 다형, 레트로트랜스포손-기반의 삽입 다형, 서열 특이적 증폭된 다형, 및 유전가능한 후생학적 변형 (예를 들면, DNA 메틸화)을 포함한다. 원인 유전 변이는 또한 밀접하게 관련된 원인 유전 변이의 셋트일 수 있다. 일부 원인 유전 변이는 RNA 폴리뉴클레오타이드에서의 서열 변이로서 영향을 미칠 수 있다. 이 수준에서, 일부 원인 유전 변이는 또한 RNA 폴리뉴클레오타이드의 종의 존재 또는 부재에 의해 표시된다. 또한, 일부 원인 유전 변이는 단백질 폴리펩타이드에서 서열 변이를 야기한다. 많은 원인 유전 변이들이 보고되어 왔다. 원인 유전 변이인 SNP의 예는 헤모글로빈의 Hb S 변이이며, 이는 겸상 세포 빈혈증을 유발한다. 원인 유전 변이인 DIP의 예는 CFTR 유전자의 delta508 돌연변이이며, 이는 낭포성 섬유증을 유발한다. 원인 유전 변이인 CNV의 예는 3염색체성 21이며, 이는 다운 증후군을 유발한다. 원인 유전 변이인 STR의 예는 직렬 반복이며, 이는 헌팅턴병을 유발한다. 이들이 연관된 원인 유전 변이 및 질환의 비제한적인 예가 표 4에 제공되어 있다. 원인 유전 변이의 부가적인 비제한적인 예가 W02014015084호에 기재되어 있다. 돌연변이가 질환과 연관된 유전자의 추가적인 예, 및 서열 변이가 본 개시내용의 방법에 따라 검출될 수 있는 유전자의 추가적인 예가 표 5에 제공되어 있다.

표 4

질환	유전자	변이체명
21-수산화효소 결핍증	CYP21A2	F306+t
21-수산화효소 결핍증	CYP21A2	F306+t
21-수산화효소 결핍증	CYP21A3	g.655A/C>G
21-수산화효소 결핍증	CYP21A4	g.655A/C>G
21-수산화효소 결핍증	CYP21A6	G110del8nt
21-수산화효소 결핍증	CYP21A5	G110del8nt
21-수산화효소 결핍증	CYP21A7	I172N, rs34607927
21-수산화효소 결핍증	CYP21A2	I236N
21-수산화효소 결핍증	CYP21A2	M239K, rs6476
21-수산화효소 결핍증	CYP21A2	P30L
21-수산화효소 결핍증	CYP21A2	P453S
21-수산화효소 결핍증	CYP21A2	Q318X
21-수산화효소 결핍증	CYP21A2	R356W
21-수산화효소 결핍증	CYP21A2	V237E, rs12530380
21-수산화효소 결핍증	CYP21A2	V281L, rs6471
ABCC8-관련 과인슐린증	ABCC8	3992-9G>A
ABCC8-관련 과인슐린증	ABCC8	delF1388
ABCC8-관련 과인슐린증	ABCC8	delF1388
ABCC8-관련 과인슐린증	ABCC8	V187D
연골무형성증	FGFR3	G375C
연골무형성증	FGFR3	G380R, rs28931614
색맹	CNGB3	c.1148delC
색맹	CNGB3	c.1148delC
색맹	CNGB3	c.819-826del8
색맹	CNGB3	c.819-826del8
색맹	CNGB3	c.886-896del11insT

[0109]

질환	유전자	변이체명
색맹	CNGB3	c.886-896del11insT
색맹	CNGB3	c.991-3T>G
색맹	CNGB3	p.Arg403Gln
색맹	CNGB3	p.Glu336X
아데노신 일인산 탈아미노효소 1	AMPD1	P48L
아데노신 일인산 탈아미노효소 1	AMPD1	Q12X, rs17602729
신경병증을 동반한 뇌량 무발생증	SLC12A6	c.2436delG
신경병증을 동반한 뇌량 무발생증	SLC12A6	c.2436delG
알칼톤뇨증	HGD	c.174delA
알칼톤뇨증	HGD	c.174delA
알칼톤뇨증	HGD	c.457_458insG
알칼톤뇨증	HGD	c.457_458insG
알칼톤뇨증	HGD	G161R
알칼톤뇨증	HGD	G270R
알칼톤뇨증	HGD	IVS1-1G>A
알칼톤뇨증	HGD	IVS5+1G>A
알칼톤뇨증	HGD	Met368Val
알칼톤뇨증	HGD	P230S
알칼톤뇨증	HGD	S47L
알칼톤뇨증	HGD	V300G
알파-1-항트립신 결핍증	SERPINA1	Arg101His, rs709932
알파-1-항트립신 결핍증	SERPINA1	Glu264Val
알파-1-항트립신 결핍증	SERPINA1	Glu342Lys, rs28929474
알파-1-항트립신 결핍증	SERPINA1	Glu376Asp, rs1303
알파-만노시도시스	MAN2B1	IVS14+1G>C
알파-만노시도시스	MAN2B1	p.L809P
알파-만노시도시스	MAN2B1	p.R750W
알파-사르코글리칸병증	SGCA	R77C, rs28933693
알파-탈라세미아	HBA2	H19D
알파-탈라세미아	HBA1	HbQ

[0110]

질환	유전자	변이체명
알파-탈라세미아	HBA1,HBA2	3.7 kb (유형 I) 결실 알파-2
알파-탈라세미아	HBA1,HBA2	3.7 kb (유형 I) 결실 알파-2
알파-탈라세미아	HBA1,HBA2	HbVar 데이터베이스 id # 1086, --(SEA); 알파-글로빈 유전자 알파-Thal-1 모두를 포함하는 ~20 kb의 결실
알파-탈라세미아	HBA1,HBA2	HbVar 데이터베이스 id # 1086, --(SEA); 알파-글로빈 유전자 알파-Thal-1 모두를 포함하는 ~20 kb의 결실
알파-탈라세미아	HBA1,HBA2	HbVar 데이터베이스 id # 1087, --(MED-I); 알파-글로빈 유전자 알파-Thal-1 모두를 포함하는 ~17.5 kb의 결실
알파-탈라세미아	HBA1,HBA2	HbVar 데이터베이스 id # 1087, --(MED-I); 알파-글로빈 유전자 알파-Thal-1 모두를 포함하는 ~17.5 kb의 결실
알파-탈라세미아	HBA1,HBA2	HbVar 데이터베이스 id # 1088, -(알파)20.5; 이러한 20.5 kb 결실은 알파 2 및 알파 1의 5' 말단을 포함함; 알파-Thal-1
알파-탈라세미아	HBA1,HBA2	HbVar 데이터베이스 id # 1088, -(알파)20.5; 이러한 20.5 kb 결실은 알파 2 및 알파 1의 5' 말단을 포함함; 알파-Thal-1
알파-탈라세미아	HBA1,HBA2	HbVar 데이터베이스 id # 1094, --(FIL); 알파 1, 알파 2, 및 제타 유전자 알파-Thal-1을 포함하는 30-34 kb의 결실
알파-탈라세미아	HBA1,HBA2	HbVar 데이터베이스 id # 1094, --(FIL); 알파 1, 알파 2, 및 제타 유전자 알파-Thal-1을 포함하는 30-34 kb의 결실
알파-탈라세미아	HBA1,HBA2	HbVar 데이터베이스 id # 1095, --(THAI); 알파 1, 알파 2, 및 제타 유전자 알파-Thal-1을 포함하는 34-38 kb의 결실
알파-탈라세미아	HBA1,HBA2	HbVar 데이터베이스 id # 1095, --(THAI); 알파 1, 알파 2, 및 제타 유전자 알파-Thal-1을 포함하는 34-38 kb의 결실
알파-탈라세미아	HBA1,HBA2	HbVar 데이터베이스 id # 1097, --(MED-II); 2개의 알파 및 제타 유전자 알파-Thal-1을 포함하는 26.5 kb의 결실
알파-탈라세미아	HBA1,HBA2	HbVar 데이터베이스 id # 1097, --(MED-II); 2개의 알파 및 제타 유전자 알파-Thal-1을 포함하는 26.5 kb의 결실
알파-탈라세미아	HBA2	HbVar 데이터베이스 id # 187
알파-탈라세미아	HBA2	HbVar 데이터베이스 id # 2598, IVS I-5 (G>A)
알파-탈라세미아	HBA1,HBA2	HbVar 데이터베이스 id # 703
알파-탈라세미아	HBA1,HBA2	HbVar 데이터베이스 id # 704
알파-탈라세미아	HBA1,HBA2	HbVar 데이터베이스 id # 705, Hb Koya Dora
알파-탈라세미아	HBA1,HBA2	HbVar 데이터베이스 id # 707, rs41412046
알파-탈라세미아	HBA1	HbVar 데이터베이스 id # 87
알파-탈라세미아	HBA1,HBA2	HbVar 데이터베이스 id # 969, Poly A (A->G); AATAAA->AATGAA 베타+

[0111]

질환	유전자	변이체명
알파-탈라세미아	HBA1,HBA2	HbVar 데이터베이스 id # 971, Poly A (A>G); AATAAA->AATAAG 베타+
알파-탈라세미아	HBA2	MIT
알파-탈라세미아	HBA1	W14X
안지오텐신 II 수용체, 유형 1	AGTR1	A1166C
아포지질단백질 E 유전자형 분석	APOE	p.C112R, rs429358
아포지질단백질 E 유전자형 분석	APOE	p.R158C, rs7412
아르기닌 숙신산뇨증	ASL	R385C
ARSACS	SACS	5254C>T
ARSACS	SACS	6594delT
ARSACS	SACS	6594delT
아스파틸글리코사민뇨증	AGA	c.199_200delGA
아스파틸글리코사민뇨증	AGA	c.199_200delGA
아스파틸글리코사민뇨증	AGA	C163S
비타민 E 결핍을 동반한 운동실조	TTPA	744delA
비타민 E 결핍을 동반한 운동실조	TTPA	744delA
모세혈관확장 운동실조	ATM	R35X
자가면역성 다선 증후군 유형 1	AIRE	c.1163_1164insA
자가면역성 다선 증후군 유형 1	AIRE	c.1163_1164insA
자가면역성 다선 증후군 유형 1	AIRE	c.769C>T
자가면역성 다선 증후군 유형 1	AIRE	c.967_979del
자가면역성 다선 증후군 유형 1	AIRE	c.967_979del

[0112]

질환	유전자	변이체명
자가면역성 다선 증후군 유형 1	AIRE	Y85C
바르데-비들 증후군	BBS1	M390R
바르데-비들 증후군	BBS10	p.C91LfsX4
바르데-비들 증후군	BBS10	p.C91LfsX4
베스트 난황형 황반 이영양증	BEST1	c.G383C
베타-사르코글리칸병증	SGCB	S114F
베타-탈라세미아	HBB	-28 (A->G) 베타+
베타-탈라세미아	HBB	-29 (A->G) 베타+
베타-탈라세미아	HBB	-29A>G
베타-탈라세미아	HBB	-30 (T->A) 베타+
베타-탈라세미아	HBB	-87 (C->G) 베타+
베타-탈라세미아	HBB	-88C>T
베타-탈라세미아	HBB	CAP+1 (A->C) 베타+
베타-탈라세미아	HBB	코돈 15 (G->A); TGG(Trp)->TAG(종결 코돈) 베타 0, rs34716011
베타-탈라세미아	HBB	코돈 15 (G->A); TGG(Trp)->TAG(종결 코돈) 베타 0, rs34716011
베타-탈라세미아	HBB	코돈 16 (-C); GGC(Gly)->GG-베타 0
베타-탈라세미아	HBB	코돈 16 (-C); GGC(Gly)->GG-베타 0
베타-탈라세미아	HBB	코돈 17 (A->T); AAG(Lys)->TAG(stop 코돈) 베타 0
베타-탈라세미아	HBB	코돈 24 (T->A); GGT(Gly)->GGA(Gly) 베타+
베타-탈라세미아	HBB	코돈 39 (C->T); CAG(Gln)->TAG(stop 코돈) 베타 0
베타-탈라세미아	HBB	코돈 5 (-CT); CCT(Pro)->C--베타 0
베타-탈라세미아	HBB	코돈 5 (-CT); CCT(Pro)->C--베타 0
베타-탈라세미아	HBB	코돈 6 (-A); GAG(Glu)->G-G 베타 0
베타-탈라세미아	HBB	코돈 6 (-A); GAG(Glu)->G-G 베타 0
베타-탈라세미아	HBB	코돈 8 (-AA); AAG(Lys)->--G 베타 0
베타-탈라세미아	HBB	코돈 8 (-AA); AAG(Lys)->--G 베타 0
베타-탈라세미아	HBB	코돈 41/42 (-TTCT); TTCCTT(Phe-Phe)->----TT 베타 0
베타-탈라세미아	HBB	코돈 41/42 (-TTCT); TTCCTT(Phe-Phe)->----TT 베타 0
베타-탈라세미아	HBB	코돈 71/72 (+A); TTT AGT(Phe Ser)->TTT A AGT; 베타 0
베타-탈라세미아	HBB	코돈 71/72 (+A); TTT AGT(Phe Ser)->TTT A AGT; 베타 0
베타-탈라세미아	HBB	코돈 8/9 (+G); AAG TCT(Lys;Ser)->AAG G TCT 베타 0

[0113]

질환	유전자	변이체명
베타-탈라세미아	HBB	코돈 8/9 (+G); AAG TCT(Lys;Ser)->AAG G TCT 베타 0
베타-탈라세미아	HBB	HbVar 데이터베이스 id # 889, IVS-II-654 (C->T); AAGGCAATA->AAG^GTAATA 베타+(중증)
베타-탈라세미아	HBB	HbVar 데이터베이스 id # 890, IVS-II-705 (T->G); GATGTAAGA->GAG^GTAAGA 베타+
베타-탈라세미아	HBB	HbVar 데이터베이스 id # 891, IVS-II-745 (C->G); CAGCTACCAT->CAG^GTACCAT 베타+
베타-탈라세미아	HBB	HbVar 데이터베이스 id # 979, 619 bp 결실 베타 0
베타-탈라세미아	HBB	619 bp 결실 베타 0
베타-탈라세미아	HBB	IVS-I-1 (G->A); AG^GTTGGT->AGATTGGT 베타 0
베타-탈라세미아	HBB	IVS-I-1 (G->T); AG^GTTGGT->AGTTTGGT 베타 0
베타-탈라세미아	HBB	IVS-I-110 (G->A) 베타+; 상기 돌연변이는 수용체 스플라이스 부위 AG^GC 에 대해 5'에 있는 21 뉴클레오타이드이다
베타-탈라세미아	HBB	IVS-I-5 (G->C) 베타+(중증)
베타-탈라세미아	HBB	IVS-II-1 (G->A); 베타 0
베타-탈라세미아	HBB	IVS-II-844 (C->G); 베타+
베타-탈라세미아	HBB	IVS1+6T>C
베타-탈라세미아	HBB	IVS11-849A>C
베타-탈라세미아	HBB	IVS11-849A>G
바이오티니다제 결핍증	BTBD	A171T, rs13073139
바이오티니다제 결핍증	BTBD	D252G, rs28934601
바이오티니다제 결핍증	BTBD	D444H, rs13078881
바이오티니다제 결핍증	BTBD	F403V
바이오티니다제 결핍증	BTBD	G98:d7i3
바이오티니다제 결핍증	BTBD	G98:d7i3
바이오티니다제 결핍증	BTBD	Q456H
바이오티니다제 결핍증	BTBD	R157H
바이오티니다제 결핍증	BTBD	R538C
블라우 증후군	NOD2	E383K
블라우 증후군	NOD2	L469F
블라우 증후군	NOD2	R334Q
블라우 증후군	NOD2	R334W
블룸 증후군	BLM	2407insT

[0114]

질환	유전자	변이체명
블룸 증후군	BLM	2407insT
블룸 증후군	BLM	736delATCTGAinsTAGATTC (2281del6/ins7)
블룸 증후군	BLM	736delATCTGAinsTAGATTC (2281del6/ins7)
BRCA1 유전적 유방/난소암	BRCA1	185delAG
BRCA1 유전적 유방/난소암	BRCA1	185delAG
BRCA1 유전적 유방/난소암	BRCA1	5382insC
BRCA1 유전적 유방/난소암	BRCA1	5382insC
BRCA1 유전적 유방/난소암	BRCA1	Tyr978X
BRCA2 유전적 유방/난소암	BRCA2	6174delT
BRCA2 유전적 유방/난소암	BRCA2	6174delT
BRCA2 유전적 유방/난소암	BRCA2	8765delAG
BRCA2 유전적 유방/난소암	BRCA2	8765delAG
카나반병	ASPA	A305E (914C>A), rs28940574
카나반병	ASPA	E285A (854A>C), rs28940279
카나반병	ASPA	IVS2-2A>G (433-2A>G)
카나반병	ASPA	Y231X (693C>A)
카르니틴 팔미토일전이효소 IA 결핍증	CPT1A	G710E
카르니틴 팔미토일전이효소 IA 결핍증	CPT1A	P479L
카르니틴 팔미토일전이효소 II 결핍증	CPT2	G549D

[0115]

질환	유전자	변이체명
카르니틴 팔미토일전효소 II 결핍증	CPT2	L178F 534 ins/25 bp del
카르니틴 팔미토일전효소 II 결핍증	CPT2	L178F 534 ins/25 bp del
카르니틴 팔미토일전효소 II 결핍증	CPT2	P227L
카르니틴 팔미토일전효소 II 결핍증	CPT2	P50H
카르니틴 팔미토일전효소 II 결핍증	CPT2	P604S
카르니틴 팔미토일전효소 II 결핍증	CPT2	Q413fs
카르니틴 팔미토일전효소 II 결핍증	CPT2	Q413fs
카르니틴 팔미토일전효소 II 결핍증	CPT2	Q550R
카르니틴 팔미토일전효소 II 결핍증	CPT2	R124X
카르니틴 팔미토일전효소 II 결핍증	CPT2	R503C
카르니틴 팔미토일전효소 II 결핍증	CPT2	R631C

[0116]

질환	유전자	변이체명
카르니틴 팔미토일전효소 II 결핍증	CPT2	S113L
카르니틴 팔미토일전효소 II 결핍증	CPT2	s38fs
카르니틴 팔미토일전효소 II 결핍증	CPT2	s38fs
카르니틴 팔미토일전효소 II 결핍증	CPT2	Y628S, rs28936673
연골모발 형성저하증	RMRP	g.262G>T
연골모발 형성저하증	RMPR	g.70A>G
CFTR-관련 장애	CFTR	1811+1.6kbA->G
CFTR-관련 장애	CFTR	2183AA>G
CFTR-관련 장애	CFTR	2183AA>G
CFTR-관련 장애	CFTR	3849+10kbC>T
CFTR-관련 장애	CFTR	A455E
CFTR-관련 장애	CFTR	A559T
CFTR-관련 장애	CFTR	C524X
CFTR-관련 장애	CFTR	574delA, 574delA
CFTR-관련 장애	CFTR	574delA, 574delA
CFTR-관련 장애	CFTR	2108delA, 2108delA
CFTR-관련 장애	CFTR	2108delA, 2108delA
CFTR-관련 장애	CFTR	3171delC, 3171delC
CFTR-관련 장애	CFTR	3171delC, 3171delC
CFTR-관련 장애	CFTR	621+1G->T
CFTR-관련 장애	CFTR	2105-2117del13insAGAAA
CFTR-관련 장애	CFTR	2105-2117del13insAGAAA
CFTR-관련 장애	CFTR	711+1G->T
CFTR-관련 장애	CFTR	711+5G->A
CFTR-관련 장애	CFTR	712-1G->T
CFTR-관련 장애	CFTR	1288insTA, 1288insTA

[0117]

질환	유전자	변이체명
CFTR-관련 장애	CFTR	1288insTA, 1288insTA
CFTR-관련 장애	CFTR	936delTA
CFTR-관련 장애	CFTR	936delTA
CFTR-관련 장애	CFTR	[delta]F311
CFTR-관련 장애	CFTR	[delta]F311
CFTR-관련 장애	CFTR	1078delT, 1078delT
CFTR-관련 장애	CFTR	1078delT, 1078delT
CFTR-관련 장애	CFTR	1161delC, 1161delC
CFTR-관련 장애	CFTR	1161delC, 1161delC
CFTR-관련 장애	CFTR	1609delCA, 1609delCA
CFTR-관련 장애	CFTR	1609delCA, 1609delCA
CFTR-관련 장애	CFTR	[delta]I507
CFTR-관련 장애	CFTR	[delta]I507
CFTR-관련 장애	CFTR	rs332, [delta]F508
CFTR-관련 장애	CFTR	rs332, [delta]F508
CFTR-관련 장애	CFTR	1677delTA, 1677delTA
CFTR-관련 장애	CFTR	1677delTA, 1677delTA
CFTR-관련 장애	CFTR	1717-1G->A
CFTR-관련 장애	CFTR	1812-1G->A
CFTR-관련 장애	CFTR	1898+1G->A
CFTR-관련 장애	CFTR	1898+1G->T
CFTR-관련 장애	CFTR	1898+5G->T
CFTR-관련 장애	CFTR	1949del84, 1949del84
CFTR-관련 장애	CFTR	1949del84, 1949del84
CFTR-관련 장애	CFTR	2043delG, 2043delG
CFTR-관련 장애	CFTR	2043delG, 2043delG
CFTR-관련 장애	CFTR	2055del9->A
CFTR-관련 장애	CFTR	2055del9->A
CFTR-관련 장애	CFTR	2143delT, 2143delT
CFTR-관련 장애	CFTR	2143delT, 2143delT
CFTR-관련 장애	CFTR	2184delA, 2184delA
CFTR-관련 장애	CFTR	2184delA, 2184delA
CFTR-관련 장애	CFTR	2184insA, 2184insA
CFTR-관련 장애	CFTR	2184insA, 2184insA

[0118]

질환	유전자	변이체명
CFTR-관련 장애	CFTR	2307insA, 2307insA
CFTR-관련 장애	CFTR	2307insA, 2307insA
CFTR-관련 장애	CFTR	296+12T->C
CFTR-관련 장애	CFTR	2789+5G->A
CFTR-관련 장애	CFTR	2869insG, 2869insG
CFTR-관련 장애	CFTR	2869insG, 2869insG
CFTR-관련 장애	CFTR	3120G->A
CFTR-관련 장애	CFTR	3120+1G->A
CFTR-관련 장애	CFTR	3272-26A->G
CFTR-관련 장애	CFTR	3659delC, 3659delC
CFTR-관련 장애	CFTR	3659delC, 3659delC
CFTR-관련 장애	CFTR	3667del4, 3667del4
CFTR-관련 장애	CFTR	3667del4, 3667del4
CFTR-관련 장애	CFTR	3791delC, 3791delC
CFTR-관련 장애	CFTR	3791delC, 3791delC
CFTR-관련 장애	CFTR	3821delT, 3821delT
CFTR-관련 장애	CFTR	3821delT, 3821delT
CFTR-관련 장애	CFTR	3905insT, 3905insT
CFTR-관련 장애	CFTR	3905insT, 3905insT
CFTR-관련 장애	CFTR	4016insT, 4016insT
CFTR-관련 장애	CFTR	4016insT, 4016insT
CFTR-관련 장애	CFTR	394delTT, 394delTT
CFTR-관련 장애	CFTR	394delTT, 394delTT
CFTR-관련 장애	CFTR	405+1G->A
CFTR-관련 장애	CFTR	405+3A->C
CFTR-관련 장애	CFTR	444delA
CFTR-관련 장애	CFTR	444delA
CFTR-관련 장애	CFTR	3876delA, 3876delA
CFTR-관련 장애	CFTR	3876delA, 3876delA
CFTR-관련 장애	CFTR	457TAT->G
CFTR-관련 장애	CFTR	457TAT->G
CFTR-관련 장애	CFTR	3199del6, 3199del6
CFTR-관련 장애	CFTR	3199del6, 3199del6
CFTR-관련 장애	CFTR	406-1G->A

[0119]

질환	유전자	변이체명
CFTR-관련 장애	CFTR	663delT, 663delT
CFTR-관련 장애	CFTR	663delT, 663delT
CFTR-관련 장애	CFTR	935delA, 935delA
CFTR-관련 장애	CFTR	935delA, 935delA
CFTR-관련 장애	CFTR	CFTR dele2,3 (21kb)
CFTR-관련 장애	CFTR	CFTR dele2,3 (21kb)
CFTR-관련 장애	CFTR	D1152H
CFTR-관련 장애	CFTR	E60X
CFTR-관련 장애	CFTR	E92X
CFTR-관련 장애	CFTR	F508C, rs1800093
CFTR-관련 장애	CFTR	G178R
CFTR-관련 장애	CFTR	G330X
CFTR-관련 장애	CFTR	G480C
CFTR-관련 장애	CFTR	G542X
CFTR-관련 장애	CFTR	G551D
CFTR-관련 장애	CFTR	G622D
CFTR-관련 장애	CFTR	G85E
CFTR-관련 장애	CFTR	G91R
CFTR-관련 장애	CFTR	I148T, rs35516286
CFTR-관련 장애	CFTR	I506V
CFTR-관련 장애	CFTR	IVS8-5T
CFTR-관련 장애	CFTR	IVS8-7T
CFTR-관련 장애	CFTR	IVS8-9T
CFTR-관련 장애	CFTR	K710X
CFTR-관련 장애	CFTR	L206W
CFTR-관련 장애	CFTR	M1101K, rs36210737
CFTR-관련 장애	CFTR	N1303K
CFTR-관련 장애	CFTR	P574H
CFTR-관련 장애	CFTR	Q1238X
CFTR-관련 장애	CFTR	Q359K/T360K_wt
CFTR-관련 장애	CFTR	Q493X
CFTR-관련 장애	CFTR	Q552X
CFTR-관련 장애	CFTR	Q890X
CFTR-관련 장애	CFTR	R1066C

[0120]

질환	유전자	변이체명
CFTR-관련 장애	CFTR	R1070Q
CFTR-관련 장애	CFTR	R1158X
CFTR-관련 장애	CFTR	R1162X
CFTR-관련 장애	CFTR	R117C
CFTR-관련 장애	CFTR	R117H
CFTR-관련 장애	CFTR	R1283M
CFTR-관련 장애	CFTR	R334W
CFTR-관련 장애	CFTR	R347H
CFTR-관련 장애	CFTR	R347P
CFTR-관련 장애	CFTR	R352Q
CFTR-관련 장애	CFTR	R553X
CFTR-관련 장애	CFTR	R560T
CFTR-관련 장애	CFTR	R709X
CFTR-관련 장애	CFTR	R75X
CFTR-관련 장애	CFTR	R764X
CFTR-관련 장애	CFTR	S1196X
CFTR-관련 장애	CFTR	S1235R, rs34911792
CFTR-관련 장애	CFTR	S1251N
CFTR-관련 장애	CFTR	S1255X
CFTR-관련 장애	CFTR	S364P
CFTR-관련 장애	CFTR	S549I
CFTR-관련 장애	CFTR	S549N
CFTR-관련 장애	CFTR	S549R
CFTR-관련 장애	CFTR	S549R
CFTR-관련 장애	CFTR	T338I
CFTR-관련 장애	CFTR	V520F
CFTR-관련 장애	CFTR	W1089X
CFTR-관련 장애	CFTR	W1204X
CFTR-관련 장애	CFTR	W1204X
CFTR-관련 장애	CFTR	W1282X
CFTR-관련 장애	CFTR	Y1092X
CFTR-관련 장애	CFTR	Y122X
맥락막결손	CHM	c.1609+2dupT
맥락막결손	CHM	c.1609+2dupT

[0121]

질환	유전자	변이체명
CLN3-관련 신경 세로이드 리포푸신증	CLN3	c.461_677del
CLN3-관련 신경 세로이드 리포푸신증	CLN3	c.461_677del
CLN3-관련 신경 세로이드 리포푸신증	CLN3	c.791_1056del
CLN3-관련 신경 세로이드 리포푸신증	CLN3	c.791_1056del
CLN5-관련 신경 세로이드 리포푸신증	CLN5	c.1175_1176delAT
CLN5-관련 신경 세로이드 리포푸신증	CLN5	c.1175_1176delAT
CLN5-관련 신경 세로이드 리포푸신증	CLN5	c.225G>A
CLN8-관련 신경 세로이드 리포푸신증	CLN8	c.70C>G
코헨 증후군	VPS13B	c.3348_3349delCT
코헨 증후군	VPS13B	c.3348_3349delCT
선천성 백내장, 안면 이상형태증, 및 신경병증	CTDP1	IVS6+389C>T
당화 Ia의 선천성 장애	PMM2	p.F119L
당화 Ia의 선천성 장애	PMM2	p.R141H
당화 Ib의 선천성 장애	MPI	R295H, rs28928906
선천성 핀란드형 신장증	NPHS1	c.121_122del
선천성 핀란드형 신장증	NPHS1	c.121_122del
선천성 핀란드형 신장증	NPHS1	c.3325C>T
크론병	NOD2	3020 ins C
크론병	NOD2	3020 ins C
크론병	NOD2	G908R, rs2066845
크론병	NOD2	R702W, rs2066844
시스틴증	CTNS	1035insC

[0122]

질환	유전자	변이체명
시스틴증	CTNS	1035insC
시스틴증	CTNS	537del21
시스틴증	CTNS	537del21
시스틴증	CTNS	57kb 결실
시스틴증	CTNS	57kb 결실
시스틴증	CTNS	D205N
시스틴증	CTNS	L158P
시스틴증	CTNS	W138X
DFNA 9 (COCH)	COCH	P51S
당뇨병 및 청력 상실	mtDNA	3234A>G
당뇨병 및 청력 상실	mtDNA	3271T>C
당뇨병 및 청력 상실	mtDNA	G8363A
당뇨병 및 청력 상실	mtDNA	T14709C
조발형 원발성 근이긴장증 (DYT1)	TOR1A	904_906delGAG
조발형 원발성 근이긴장증 (DYT1)	TOR1A	904_906delGAG
연접형 수포성 표피박리증, 헤르리츠- 피어슨형	LAMB3	3024delT
연접형 수포성 표피박리증, 헤르리츠- 피어슨형	LAMB3	3024delT
연접형 수포성 표피박리증, 헤르리츠- 피어슨형	LAMB3	p.Q243X
연접형 수포성 표피박리증, 헤르리츠- 피어슨형	LAMB3	R144X
연접형 수포성 표피박리증, 헤르리츠- 피어슨형	LAMB3	R42X
연접형 수포성 표피박리증, 헤르리츠- 피어슨형	LAMB3	R635X

[0123]

질환	유전자	변이체명
연접형 수포성 표피박리증, 헤르리츠- 피어슨형	LAMA3	R650X
연접형 수포성 표피박리증, 헤르리츠- 피어슨형	LAMC2	R95X
인자 V 레이든 혈전성향증	F5	H1299R
인자 V 레이든 혈전성향증	F5	R506Q, rs6025
인자 V R2 돌연변이 혈전성향증	F5	rs6027
인자 XI 결핍증	F11	E117X (576G>T)
인자 XI 결핍증	F11	F283L (1074T>C)
인자 XI 결핍증	F11	IVS14 +1G>A
인자 XI 결핍증	F11	IVS14del14
인자 XI 결핍증	F11	IVS14del14
인자 XIII 결핍증	F13A1	V34L, rs5985
가족성 선종성 용종증	APC	I1307K, rs1801155
가족성 자율신경 실조증	IKBKAP	2507+6T>C
가족성 자율신경 실조증	IKBKAP	P914L
가족성 자율신경 실조증	IKBKAP	R696P
가족성 고콜레스테롤혈증유형 B	APOB	R3500Q, rs5742904
가족성 고콜레스테롤혈증유형 B	APOB	R3500W
가족성 고콜레스테롤혈증유형 B	APOB	R3531C, rs12713559
가족성 지중해성 열병	MEFV	A744S (2230G>T)
가족성 지중해성 열병	MEFV	delI692 (del2076_2078)

[0124]

질환	유전자	변이체명
가족성 지중해성 열병	MEFV	delI692 (del2076_2078)
가족성 지중해성 열병	MEFV	E148Q (442 G>C), rs3743930
가족성 지중해성 열병	MEFV	E167D (501 G>C)
가족성 지중해성 열병	MEFV	F479L (1437 C>G)
가족성 지중해성 열병	MEFV	K695R (2084A>G)
가족성 지중해성 열병	MEFV	M680I (2040G>C)
가족성 지중해성 열병	MEFV	M694I (2082G>A), rs28940578
가족성 지중해성 열병	MEFV	M694V (2080A>G)
가족성 지중해성 열병	MEFV	P369S (1105 C>T), rs11466023
가족성 지중해성 열병	MEFV	R408Q (1223G>A), rs11466024
가족성 지중해성 열병	MEFV	R653H (1958G>A)
가족성 지중해성 열병	MEFV	R761H (2282G>A)
가족성 지중해성 열병	MEFV	T267I (800 C>T)
가족성 지중해성 열병	MEFV	V726A (2177T>C), rs28940579
FANCC-관련 판코니 빈혈	FANCC	322delG
FANCC-관련 판코니 빈혈	FANCC	322delG
FANCC-관련 판코니 빈혈	FANCC	IVS4+4A>T (711 +4A>T)
FANCC-관련 판코니 빈혈	FANCC	Q13X (37C>T)
FANCC-관련 판코니 빈혈	FANCC	R547X
FGFR1-관련 두개골유합증	FGFR1	P252R
FGFR2-관련 두개골유합증	FGFR2	P253R
FGFR2-관련 두개골유합증	FGFR2	S252W
FGFR3-관련 두개골유합증	FGFR3	A391E, rs28931615
FGFR3-관련 두개골유합증	FGFR3	P250R, rs4647924

[0125]

질환	유전자	변이체명
자유 시알산 축적 장애	SLC17A5	c.1007_1008delTA
자유 시알산 축적 장애	SLC17A5	c.1007_1008delTA
자유 시알산 축적 장애	SLC17A5	c.115C>T
파킨슨병-17 동반 전두측 두엽성 치매	MAPT	IVS10+16
파킨슨병-17 동반 전두측 두엽성 치매	MAPT	P301L
파킨슨병-17 동반 전두측 두엽성 치매	MAPT	P301S
파킨슨병-17 동반 전두측 두엽성 치매	MAPT	R406W
푸마라아제 결핍증	FH	c.1431_1433dupAAA
푸마라아제 결핍증	FH	c.1431_1433dupAAA
갈락토오스혈증	GALT	5.0Kb 유전자 결실
갈락토오스혈증	GALT	5.0Kb 유전자 결실
갈락토오스혈증	GALT	5'UTR-119del
갈락토오스혈증	GALT	5'UTR-119del
갈락토오스혈증	GALT	IVS2-2 A>G
갈락토오스혈증	GALT	K285N
갈락토오스혈증	GALT	L195P T>C
갈락토오스혈증	GALT	L218L
갈락토오스혈증	GALT	N314D, rs2070074
갈락토오스혈증	GALT	Phe171Ser
갈락토오스혈증	GALT	Q169K
갈락토오스혈증	GALT	Q188R
갈락토오스혈증	GALT	S135L
갈락토오스혈증	GALT	T138MC>T
갈락토오스혈증	GALT	X380R
갈락토오스혈증	GALT	Y209C A>G
고셔병	GBA	1035insG
고셔병	GBA	1035insG
고셔병	GBA	84insG
고셔병	GBA	84insG
고셔병	GBA	D409H, rs1064651

[0126]

질환	유전자	변이체명
고셔병	GBA	D409V
고셔병	GBA	IVS2(+1)G>A
고셔병	GBA	L444P (1448T>C) _g , rs35095275
고셔병	GBA	N370S
고셔병	GBA	R463C
고셔병	GBA	R463H
고셔병	GBA	R496H (1604G>A)
고셔병	GBA	V394L
GJB2-관련 DFNA 3 비증후군성 청력 상실 및 난청	GJB2	167delT
GJB2-관련 DFNA 3 비증후군성 청력 상실 및 난청	GJB2	167delT
GJB2-관련 DFNA 3 비증후군성 청력 상실 및 난청	GJB2	235delC
GJB2-관련 DFNA 3 비증후군성 청력 상실 및 난청	GJB2	235delC
GJB2-관련 DFNA 3 비증후군성 청력 상실 및 난청	GJB2	35delG
GJB2-관련 DFNA 3 비증후군성 청력 상실 및 난청	GJB2	35delG
GJB2-관련 DFNA 3 비증후군성 청력 상실 및 난청	GJB2	IVS1+1G>A
GJB2-관련 DFNB 1 비증후군성 청력 상실 및 난청	GJB2	101delAG
GJB2-관련 DFNB 1 비증후군성 청력 상실 및 난청	GJB2	313del14

[0127]

질환	유전자	변이체명
GJB2-관련 DFNB 1 비증후군성 청력 상실 및 난청	GJB2	313del14
GJB2-관련 DFNB 1 비증후군성 청력 상실 및 난청	GJB2	delE120
GJB2-관련 DFNB 1 비증후군성 청력 상실 및 난청	GJB2	delE120
GJB2-관련 DFNB 1 비증후군성 청력 상실 및 난청	GJB2	M34T, rs35887622
GJB2-관련 DFNB 1 비증후군성 청력 상실 및 난청	GJB2	Q124X
GJB2-관련 DFNB 1 비증후군성 청력 상실 및 난청	GJB2	R184P
GJB2-관련 DFNB 1 비증후군성 청력 상실 및 난청	GJB2	V37I
GJB2-관련 DFNB 1 비증후군성 청력 상실 및 난청	GJB2	W24X
GJB2-관련 DFNB 1 비증후군성 청력 상실 및 난청	GJB2	W77R
GJB2-관련 DFNB 1 비증후군성 청력 상실 및 난청	GJB2	W77X
글루코스-6-인산 탈수소효소 결핍증	G6PD	A335V
글루코스-6-인산 탈수소효소 결핍증	G6PD	R459L

[0128]

질환	유전자	변이체명
글루코스-6-인산 탈수소효소 결핍증	G6PD	R459P
글루코스-6-인산 탈수소효소 결핍증	G6PD	rs1050828, rs1050828
글루코스-6-인산 탈수소효소 결핍증	G6PD	rs1050829, rs1050829
글루코스-6-인산 탈수소효소 결핍증	G6PD	rs5030868, rs5030868
글루타르산혈증 유형 1	GCDH	A421V
글루타르산혈증 유형 1	GCDH	R402W
글리코겐 축적 질환 유형 1a	G6PC	459insTA
글리코겐 축적 질환 유형 1a	G6PC	459insTA
글리코겐 축적 질환 유형 1a	G6PC	727G/T
글리코겐 축적 질환 유형 1a	G6PC	del F327
글리코겐 축적 질환 유형 1a	G6PC	del F327
글리코겐 축적 질환 유형 1a	G6PC	G188R
글리코겐 축적 질환 유형 1a	G6PC	G270V
글리코겐 축적 질환 유형 1a	G6PC	Q242X
글리코겐 축적 질환 유형 1a	G6PC	Q27fsdelC
글리코겐 축적 질환 유형 1a	G6PC	Q27fsdelC
글리코겐 축적 질환 유형 1a	G6PC	Q347X
글리코겐 축적 질환 유형 1a	G6PC	R83C

[0129]

질환	유전자	변이체명
글리코겐 축적 질환 유형 1a	G6PC	R83H
글리코겐 축적 질환 유형 1b	G6PT1	I211delCT
글리코겐 축적 질환 유형 1b	G6PT1	A367T
글리코겐 축적 질환 유형 1b	G6PT1	G339C
글리코겐 축적 질환 유형 1b	G6PT1	G339D
글리코겐 축적 질환 유형 1b	G6PT1	W118R
글리코겐 축적 질환 유형 II	GAA	Arg854X
글리코겐 축적 질환 유형 II	GAA	Asp645Glu, rs28940868
글리코겐 축적 질환 유형 II	GAA	IVS1(-13t>g)
글리코겐 축적 질환 유형 III	AGL	1484delT
글리코겐 축적 질환 유형 III	AGL	1484delT
글리코겐 축적 질환 유형 III	AGL	17delAG
글리코겐 축적 질환 유형 III	AGL	17delAG
글리코겐 축적 질환 유형 III	AGL	Q6X
글리코겐 축적 질환 유형 V	PYGM	G204S
글리코겐 축적 질환 유형 V	PYGM	K542T
글리코겐 축적 질환 유형 V	PYGM	K542X

[0130]

질환	유전자	변이체명
글리코겐 축적 질환 유형 V	PYGM	R49X
GNE-관련 근병증	GNE	M712T, rs28937594
Gracile 증후군	BCS1L	c.232A>G, rs28937590
헤모글로빈 S 베타- 탈라세미아	HBB	c.19G>A
헤모글로빈 S 베타- 탈라세미아	HBB	c.20A>T
헤모글로빈 S 베타- 탈라세미아	HBB	c.79G>A
헤모글로빈 S 베타- 탈라세미아	HBB	Hb CS
헤모글로빈 S 베타- 탈라세미아	HBB	Hb D
헤모글로빈 S 베타- 탈라세미아	HBB	Hb O
유전성 과당 불내증	ALDOB	A149P, rs1800546
유전성 과당 불내증	ALDOB	A174D
유전성 과당 불내증	ALDOB	Delta4E4
유전성 과당 불내증	ALDOB	Delta4E4
유전성 과당 불내증	ALDOB	N334K
유전성 과당 불내증	ALDOB	Y203X
유전성 췌장염	PRSS1	A16V
유전성 췌장염	SPINK1	M1T
유전성 췌장염	PRSS1	N29I
유전성 췌장염	SPINK1	N34S, rs17107315
유전성 췌장염	PRSS1	R122C
유전성 췌장염	PRSS1	R122H
유전적 티민- 우라실노증	DPYD	rs3918290
헥소사미니다아제 A 결핍증	HEXA	1278insTATC
헥소사미니다아제 A 결핍증	HEXA	1278insTATC

[0131]

질환	유전자	변이체명
헥소사미니다아제 A 결핍증	HEXA	G269S (805G>A)
헥소사미니다아제 A 결핍증	HEXA	IVS12 +1G>C
헥소사미니다아제 A 결핍증	HEXA	IVS7 +1G>A
헥소사미니다아제 A 결핍증	HEXA	IVS9 +1G>A
헥소사미니다아제 A 결핍증	HEXA	R178C
헥소사미니다아제 A 결핍증	HEXA	R178H
헥소사미니다아제 A 결핍증	HEXA	R247W (739C>T)
헥소사미니다아제 A 결핍증	HEXA	R249W (745C>T)
HFE-연관 유전성 혈색소침착증	HFE	E168Q
HFE-연관 유전성 혈색소침착증	HFE	E168X
HFE-연관 유전성 혈색소침착증	HFE	HM971246, H63H
HFE-연관 유전성 혈색소침착증	HFE	P160delC
HFE-연관 유전성 혈색소침착증	HFE	P160delC
HFE-연관 유전성 혈색소침착증	HFE	Q127H, rs28934595
HFE-연관 유전성 혈색소침착증	HFE	Q283P
HFE-연관 유전성 혈색소침착증	HFE	rs1799945, rs1799945
HFE-연관 유전성 혈색소침착증	HFE	rs1800562, rs1800562

[0132]

질환	유전자	변이체명
HFE-연관 유전성 혈색소침착증	HFE	rs1800730, rs1800730
HFE-연관 유전성 혈색소침착증	HFE	V53M
HFE-연관 유전성 혈색소침착증	HFE	V59M
HFE-연관 유전성 혈색소침착증	HFE	W169X
발한성 외배엽 이형성 2	GJB6	A88V, rs28937872
발한성 외배엽 이형성 2	GJB6	G11R
발한성 외배엽 이형성 2	GJB6	V37E
시스타티오닌 베타- 합성효소 결핍증에 의해 유발된 호모시스틴뇨증	CBS	G307S 919G->A
시스타티오닌 베타- 합성효소 결핍증에 의해 유발된 호모시스틴뇨증	CBS	I278T 833T->C, rs5742905
고칼륨혈증 주기성 마비 유형 1	SCN4A	I693T
고칼륨혈증 주기성 마비 유형 1	SCN4A	L689I
고칼륨혈증 주기성 마비 유형 1	SCN4A	L689V
고칼륨혈증 주기성 마비 유형 1	SCN4A	M1360V
고칼륨혈증 주기성 마비 유형 1	SCN4A	M1592V
고칼륨혈증 주기성 마비 유형 1	SCN4A	p.A1156T
고칼륨혈증 주기성 마비 유형 1	SCN4A	p.M1370V
고칼륨혈증 주기성 마비 유형 1	SCN4A	p.R1448C

[0133]

질환	유전자	변이체명
고칼륨혈증 주기성 마비 유형 1	SCN4A	p.T1313M
고칼륨혈증 주기성 마비 유형 1	SCN4A	R675G
고칼륨혈증 주기성 마비 유형 1	SCN4A	R675Q
고칼륨혈증 주기성 마비 유형 1	SCN4A	R675W
고칼륨혈증 주기성 마비 유형 1	SCN4A	T704M
고칼륨혈증 주기성 마비 유형 1	SCN4A	V781I
고오르니틴혈증- 고암모니아혈증- 호모시트룰린뇨증 증후군	SLC25A15	F188del
고오르니틴혈증- 고암모니아혈증- 호모시트룰린뇨증 증후군	SLC25A15	F188del
원발성 옥살산뇨증, 유형 1	AGXT	33insC
원발성 옥살산뇨증, 유형 1	AGXT	33insC
원발성 옥살산뇨증, 유형 1	AGXT	F152I
원발성 옥살산뇨증, 유형 1	AGXT	G170R
원발성 옥살산뇨증, 유형 1	AGXT	I244T
원발성 옥살산뇨증, 유형 2	GRHPR	I03delG
원발성 옥살산뇨증, 유형 2	GRHPR	I03delG
연골형성저하증	FGFR3	Asn328Ile

[0134]

질환	유전자	변이체명
연골형성저하증	FGFR3	I538V
연골형성저하증	FGFR3	K650M
연골형성저하증	FGFR3	K650N 1950G>T
연골형성저하증	FGFR3	K650Q
연골형성저하증	FGFR3	N540K 1620C>A
연골형성저하증	FGFR3	N540S
연골형성저하증	FGFR3	N540T
저칼륨혈증 주기성 마비 유형 1	CACNA1S	R528G
저칼륨혈증 주기성 마비 유형 1	CACNA1S	R528H
저칼륨혈증 주기성 마비 유형 1	CACNA1S	rs28930068, rs28930068
저칼륨혈증 주기성 마비 유형 1	CACNA1S	rs28930069, rs28930069
저칼륨혈증 주기성 마비 유형 2	SCN4A	R669H
저칼륨혈증 주기성 마비 유형 2	SCN4A	R672C
저칼륨혈증 주기성 마비 유형 2	SCN4A	R672G
저칼륨혈증 주기성 마비 유형 2	SCN4A	R672H
저칼륨혈증 주기성 마비 유형 2	SCN4A	R672S
저인산증	ALPL	Asp361Val
저인산증	ALPL	c.1559delT
저인산증	ALPL	c.1559delT
저인산증	ALPL	E174K
저인산증	ALPL	G317D
저인산증	ALPL	Phe310Leu
이소발레릭산혈증	IVD	A282V
이소발레릭산혈증	IVD	rs28940889
크라베병	GALC	EX11-17DEL

[0135]

질환	유전자	변이체명
크라베병	GALC	EX11-17DEL
크라베병	GALC	G270D
크라베병	GALC	rs1805078, rs1805078
크라베병	GALC	rs398607
레베르 유전성 시신경병증	mtDNA	14484T>C
레베르 유전성 시신경병증	mtDNA	15257G>A
레베르 유전성 시신경병증	mtDNA	G14459A
레베르 유전성 시신경병증	mtDNA	G3460A
레베르 유전성 시신경병증	mtDNA	m.11778G>A
레베르 유전성 시신경병증	mtDNA	m.13708G>A
레베르 유전성 시신경병증	mtDNA	m.15812G>A
레베르 유전성 시신경병증	mtDNA	m.3394T>C
레베르 유전성 시신경병증	mtDNA	m.4216T>C
레베르 유전성 시신경병증	mtDNA	m.4917A>G
리씨 증후군, 프랑스- 캐나다 유형	LRPPRC	A354V
LGMD2I	FKRP	L276I, rs28937900
장쇄 3-하이드록시아실- CoA 탈수소효소 결핍증	HADHA	E474Q c.1528G>C
장쇄 3-하이드록시아실- CoA 탈수소효소 결핍증	HADHA	Q342X 1132C>T
단풍당뇨증 유형 1A	BCKDHA	Y438N
단풍당뇨증 유형 1B	BCKDHB	E372X
단풍당뇨증 유형 1B	BCKDHB	G278S
단풍당뇨증 유형 1B	BCKDHB	R183P

[0136]

질환	유전자	변이체명
맥륜-알브라이트 증후군	GNAS	R201C
맥륜-알브라이트 증후군	GNAS	R201G
맥륜-알브라이트 증후군	GNAS	R201H
중쇄 아실-조효소 A 탈수소효소 결핍증	ACADM	244insT
중쇄 아실-조효소 A 탈수소효소 결핍증	ACADM	244insT
중쇄 아실-조효소 A 탈수소효소 결핍증	ACADM	250C>T
중쇄 아실-조효소 A 탈수소효소 결핍증	ACADM	583G>A
중쇄 아실-조효소 A 탈수소효소 결핍증	ACADM	616C>T
중쇄 아실-조효소 A 탈수소효소 결핍증	ACADM	617G>A
중쇄 아실-조효소 A 탈수소효소 결핍증	ACADM	799G>A
중쇄 아실-조효소 A 탈수소효소 결핍증	ACADM	K304E
중쇄 아실-조효소 A 탈수소효소 결핍증	ACADM	Y42H
피골하 골낭포를 동반한 거대뇌성 백질뇌병증	MLC1	135insC
피골하 골낭포를 동반한 거대뇌성 백질뇌병증	MLC1	135insC
MELAS	mtDNA	3243A>G
MELAS	mtDNA	3250T>C
MELAS	mtDNA	3252A>G
MELAS	mtDNA	A12770G
MELAS	mtDNA	C3256T
MELAS	mtDNA	G13513A
MELAS	mtDNA	T3291C

[0137]

질환	유전자	변이체명
MELAS	mtDNA	T8356C
MELAS	mtDNA	T9957C
MERRF	mtDNA	8361G>A
MERRF	mtDNA	A8296G
MERRF	mtDNA	m.8344A>G
이염성 백질이영양증	ARSA	c.459+1G>A
이염성 백질이영양증	ARSA	p.P426L, rs28940893
이염성 백질이영양증	ARSA	p.T274M
이염성 백질이영양증	ARSA	P377L
미토콘드리아 심근증	mtDNA	A3260T
미토콘드리아 심근증	mtDNA	A4300G
미토콘드리아 심근증	mtDNA	C3303T
미토콘드리아 심근증	mtDNA	T9997C
미토콘드리아 DNA- 연관 리씨 증후군 및 NARP	mtDNA	5537insT
미토콘드리아 DNA- 연관 리씨 증후군 및 NARP	mtDNA	5537insT
미토콘드리아 DNA- 연관 리씨 증후군 및 NARP	mtDNA	8993T>C
미토콘드리아 DNA- 연관 리씨 증후군 및 NARP	mtDNA	8993T>G
미토콘드리아 DNA- 연관 리씨 증후군 및 NARP	mtDNA	C11777A
미토콘드리아 DNA- 연관 리씨 증후군 및 NARP	mtDNA	T10158C
미토콘드리아 DNA- 연관 리씨 증후군 및 NARP	mtDNA	T10191C
미토콘드리아 DNA- 연관 리씨 증후군 및 NARP	mtDNA	T8851C

[0138]

질환	유전자	변이체명
미토콘드리아 DNA- 연관 리씨 증후군 및 NARP	mtDNA	T9176C
미토콘드리아 DNA- 연관 리씨 증후군 및 NARP	mtDNA	T9176G
MTHFR 결핍증	MTHFR	1298A>C
MTHFR 결핍증	MTHFR	rs1801133, rs1801133
MTRNR1-관련 청력 상실 및 난청	mtDNA	1095T>C
MTRNR1-관련 청력 상실 및 난청	mtDNA	1494C>T
MTRNR1-관련 청력 상실 및 난청	mtDNA	1555A>G
MTRNR1-관련 청력 상실 및 난청	mtDNA	961T>G
MTRNR1-관련 청력 상실 및 난청	mtDNA	A7445G
MTTS1-관련 청력 상실 및 난청	mtDNA	7443A>G
MTTS1-관련 청력 상실 및 난청	mtDNA	7444G>A
MTTS1-관련 청력 상실 및 난청	mtDNA	7472insC
MTTS1-관련 청력 상실 및 난청	mtDNA	7472insC
MTTS1-관련 청력 상실 및 난청	mtDNA	7510T>C
MTTS1-관련 청력 상실 및 난청	mtDNA	7511T>C
MTTS1-관련 청력 상실 및 난청	mtDNA	7512T>C
뮤코리피드증 IV	MCOLN1	delEx1 3 Ex7 (511>6944del)
뮤코리피드증 IV	MCOLN1	delEx1 3 Ex7 (511>6944del)
뮤코리피드증 IV	MCOLN1	IVS-2A>G

[0139]

질환	유전자	변이체명
뮤코다당류증 유형 I	IDUA	c.46_57del
뮤코다당류증 유형 I	IDUA	c.46_57del
뮤코다당류증 유형 I	IDUA	p.A327P
뮤코다당류증 유형 I	IDUA	p.P533R
뮤코다당류증 유형 I	IDUA	Q70X
뮤코다당류증 유형 I	IDUA	W402X
뮤코다당류증 유형 IIIA	SGSH	p.R245H
뮤코다당류증 유형 IIIA	SGSH	p.R74C
뮤코다당류증 유형 IIIA	SGSH	p.S66W
뮤코다당류증 유형 VII	GUSB	p.D152N
다발성 내분비 선종 유형 2	RET	2047T>A
다발성 내분비 선종 유형 2	RET	2047T>A
다발성 내분비 선종 유형 2	RET	2047T>C
다발성 내분비 선종 유형 2	RET	2047T>G
다발성 내분비 선종 유형 2	RET	2048G>A
다발성 내분비 선종 유형 2	RET	A883F 2647 G>T
다발성 내분비 선종 유형 2	RET	Glu768Asp G>C
다발성 내분비 선종 유형 2	RET	M918T
근육-눈-뇌병	POMGNT1	c.1539+1G>A
MYH-연관 폴립증	MUTYH	c.1376C>A
MYH-연관 폴립증	MUTYH	c.494A>G, rs34612342
MYH-연관 폴립증	유전자_SYMBOL TBD	rs36053993, rs36053993
스핑고마이엘리나제 결핍으로 인한 니만- 피크병	SMPD1	c.990delC

[0140]

질환	유전자	변이체명
스핑고마이엘리나제 결핍으로 인한 니만-피크병	SMPD1	c.990delC
스핑고마이엘리나제 결핍으로 인한 니만-피크병	SMPD1	fsP330
스핑고마이엘리나제 결핍으로 인한 니만-피크병	SMPD1	fsP330
스핑고마이엘리나제 결핍으로 인한 니만-피크병	SMPD1	L302P
스핑고마이엘리나제 결핍으로 인한 니만-피크병	SMPD1	R496L
스핑고마이엘리나제 결핍으로 인한 니만-피크병	SMPD1	R608del
니만-피크병 유형 C1	NPC1	I1061T
니즈메겐 파손증후군	NBN	657del5
니즈메겐 파손증후군	NBN	657del5
팔리스터-홀 증후군	GLI3	2012delG
팔리스터-홀 증후군	GLI3	2012delG
팔리스터-홀 증후군	GLI3	2023delG
팔리스터-홀 증후군	GLI3	2023delG
펜드레드 증후군	SLC26A4	1197delT
펜드레드 증후군	SLC26A4	1197delT
펜드레드 증후군	SLC26A4	E384G
펜드레드 증후군	SLC26A4	IV58+1(G->A)
펜드레드 증후군	SLC26A4	L236P
펜드레드 증후군	SLC26A4	T416P
퍼옥시좀 이작용성 효소 결핍증	HSD17B4	c.302+1G>C
퍼옥시좀 이작용성 효소 결핍증	HSD17B4	c.303-1G>A

[0141]

질환	유전자	변이체명
전반적 발달 장애	NLGN4X	D396X
전반적 발달 장애	NLGN4X	D396X
전반적 발달 장애	NLGN4X	NLGN4X:1253delAG
전반적 발달 장애	NLGN4X	NLGN4X:1253delAG
전반적 발달 장애	NLGN3	R451C
페닐알라닌 수산화효소 결핍증	PAH	G272X
페닐알라닌 수산화효소 결핍증	PAH	I65T
페닐알라닌 수산화효소 결핍증	PAH	IVS12+1G>T
페닐알라닌 수산화효소 결핍증	PAH	L48S, rs5030841
페닐알라닌 수산화효소 결핍증	PAH	R158Q, rs5030843
페닐알라닌 수산화효소 결핍증	PAH	R252W, rs5030847
페닐알라닌 수산화효소 결핍증	PAH	R261Q, rs5030849
페닐알라닌 수산화효소 결핍증	PAH	R408Q, rs5030859
페닐알라닌 수산화효소 결핍증	PAH	R408W, rs5030858
페닐알라닌 수산화효소 결핍증	PAH	rs5030855
페닐알라닌 수산화효소 결핍증	PAH	rs5030861
페닐알라닌 수산화효소 결핍증	PAH	Y414C, rs5030860
플라스미노겐 활성화제 억제제 I	SERPINE1	-844 G>A
플라스미노겐 활성화제 억제제 I	SERPINE1	4G/5G
상염색체 열성 다낭성 신장 질환	PKHD1	c.10412T>G

[0142]

질환	유전자	변이체명
상염색체 열성 다낭성 신장 질환	PKHD1	c.107C>T
상염색체 열성 다낭성 신장 질환	PKHD1	c.1486C>T
상염색체 열성 다낭성 신장 질환	PKHD1	c.5895dupA
상염색체 열성 다낭성 신장 질환	PKHD1	c.5895dupA
상염색체 열성 다낭성 신장 질환	PKHD1	c.9689delA
상염색체 열성 다낭성 신장 질환	PKHD1	c.9689delA
PPT1-관련 신경 세로이드 리포푸신증	PPT1	c.364A>T
PPT1-관련 신경 세로이드 리포푸신증	PPT1	p.L10X
PPT1-관련 신경 세로이드 리포푸신증	PPT1	p.R151X
PPT1-관련 신경 세로이드 리포푸신증	PPT1	T75P
PROP1-관련 뇌하수체 호르몬 결핍증	PROP1	301-302delAG
PROP1-관련 뇌하수체 호르몬 결핍증	PROP1	301-302delAG
프로트롬빈 혈전성향증	F2	rs1799963
프로트롬빈 혈전성향증	F2	rs6025, rs6025
거짓비타민 D 결핍 구루병	CYP27B1	엑손 8 에서 7bp 중복
거짓비타민 D 결핍 구루병	CYP27B1	엑손 8 에서 7bp 중복
거짓비타민 D 결핍 구루병	CYP27B1	958delG
거짓비타민 D 결핍 구루병	CYP27B1	958delG
레트 증후군	MECP2	806delG

[0143]

질환	유전자	변이체명
레트 증후군	MECP2	806delG
레트 증후군	MECP2	A140V, rs28934908
레트 증후군	MECP2	P152R
레트 증후군	MECP2	P225R
레트 증후군	MECP2	R106W, rs28934907
레트 증후군	MECP2	R133C
레트 증후군	MECP2	R168X
레트 증후군	MECP2	R255X
레트 증후군	MECP2	R270X
레트 증후군	MECP2	R294X
레트 증후군	MECP2	R306C, rs28935468
레트 증후군	MECP2	S134C
레트 증후군	MECP2	T158M, rs28934906
점상연골 이형성증 유형 1	PEX7	p.A218V
점상연골 이형성증 유형 1	PEX7	p.G217R
점상연골 이형성증 유형 1	PEX7	p.L292X, rs1805137
단쇄아실-CoA 탈수소효소 결핍증	ACADS	c.511C>T, rs1800556
단쇄아실-CoA 탈수소효소 결핍증	ACADS	c.625G>A
단쇄아실-CoA 탈수소효소 결핍증	ACADS	R107C
슈바크만-다이아몬드 증후군	SBDS	183_184TA>CT
슈바크만-다이아몬드 증후군	SBDS	183_184TA>CT
슈바크만-다이아몬드 증후군	SBDS	258+2T>C
쇼그렌-라손 증후군	ALDH3A2	c.943C>T
스미스-렘리-오피쯔 증후군	DHCR7	C380Y

[0144]

질환	유전자	변이체명
스미스-렘리-오피쯔 증후군	DHCR7	IVS8-1G>C
스미스-렘리-오피쯔 증후군	DHCR7	L109P
스미스-렘리-오피쯔 증후군	DHCR7	L157P
스미스-렘리-오피쯔 증후군	DHCR7	R352Q
스미스-렘리-오피쯔 증후군	DHCR7	R352W
스미스-렘리-오피쯔 증후군	DHCR7	R404C
스미스-렘리-오피쯔 증후군	DHCR7	R446Q
스미스-렘리-오피쯔 증후군	DHCR7	T93M
스미스-렘리-오피쯔 증후군	DHCR7	V326L
스미스-렘리-오피쯔 증후군	DHCR7	W151X
스미스-렘리-오피쯔 증후군	DHCR7	W151X
강직성 대마비 13	HSPD1	V72I
설페이트 수송체-관련 골연골이형성증	SLC26A2	340delV
설페이트 수송체-관련 골연골이형성증	SLC26A2	340delV
설페이트 수송체-관련 골연골이형성증	SLC26A2	c.837C>T
설페이트 수송체-관련 골연골이형성증	SLC26A2	C653S
설페이트 수송체-관련 골연골이형성증	SLC26A2	IVS1+2T>C
설페이트 수송체-관련 골연골이형성증	SLC26A2	R178X

[0145]

질환	유전자	변이체명
TFR2-관련 유전성 혈색소침착증	TFR2	AVAQ594-597del
TFR2-관련 유전성 혈색소침착증	TFR2	AVAQ594-597del
TFR2-관련 유전성 혈색소침착증	TFR2	E60X
TFR2-관련 유전성 혈색소침착증	TFR2	E60X
TFR2-관련 유전성 혈색소침착증	TFR2	M172K
TFR2-관련 유전성 혈색소침착증	TFR2	Y250X
치사성 이형성증	FGFR3	G370C
치사성 이형성증	FGFR3	K650E
치사성 이형성증	FGFR3	R248C
치사성 이형성증	FGFR3	S249C
치사성 이형성증	FGFR3	S371C
치사성 이형성증	FGFR3	X807C A>T
치사성 이형성증	FGFR3	X807G
치사성 이형성증	FGFR3	X807L
치사성 이형성증	FGFR3	X807R
치사성 이형성증	FGFR3	X807S
치사성 이형성증	FGFR3	X807W
치사성 이형성증	FGFR3	Y373C
TPP1-관련 신경 세로이드 리포푸신증	TPP1	c.509-1G>A
TPP1-관련 신경 세로이드 리포푸신증	TPP1	c.509-1G>C
TPP1-관련 신경 세로이드 리포푸신증	TPP1	G284V
TPP1-관련 신경 세로이드 리포푸신증	TPP1	p.R208X
트랜스티레틴 아밀로이드증	TTR	c.148G>A

[0146]

질환	유전자	변이체명
타이로신 수산화효소-결핍 DRD	TH	L205P
타이로신 수산화효소-결핍 DRD	TH	R202H
타이로신혈증 유형 I	FAH	E357X
타이로신혈증 유형 I	FAH	IVS12+5 G>A
타이로신혈증 유형 I	FAH	IVS7-6 T>G
타이로신혈증 유형 I	FAH	IVS8-1G>C
타이로신혈증 유형 I	FAH	p.W262X
타이로신혈증 유형 I	FAH	P261L
타이로신혈증 유형 I	FAH	Q64H
월슨병	ATP7B	I340del4
월슨병	ATP7B	3402delC
월슨병	ATP7B	3402delC
월슨병	ATP7B	H1069Q
월슨병	ATP7B	R778G
월슨병	ATP7B	W779X
월슨병	ATP7B	W779X
X-관련 연소망막층간분리	RS1	E72K
X-관련 연소망막층간분리	RS1	G109R
X-관련 연소망막층간분리	RS1	G74V
젤웨거 증후군 스펙트럼	PEX1	c.2097_2098insT
젤웨거 증후군 스펙트럼	PEX1	c.2097_2098insT
젤웨거 증후군 스펙트럼	PEX1	c.2916delA
젤웨거 증후군 스펙트럼	PEX1	c.2916delA
젤웨거 증후군 스펙트럼	PEX1	p.G843D

[0147]

표 5

질환/장애	유전자(들)
종양형성	PTEN; ATM; ATR; EGFR; ERBB2; ERBB3; ERBB4; Notch1; Notch2; Notch3; Notch4; AKT; AKT2; AKT3; HIF; HIF1a; HIF3a; Met; HRG; Bcl2; PPAR 알파; PPAR 감마; WT1 (윌름스 종양); FGF 수용체 패밀리 멤버 (5 멤버: 1, 2, 3, 4, 5); CDKN2a; APC; RB (망막아종); MEN1; VHL; BRCA1; BRCA2; AR (안드로겐 수용체); TSG101; IGF; IGF 수용체; Igf1 (4 변이체); Igf2 (3 변이체); Igf 1 수용체; Igf 2 수용체; Bax; Bcl2; 카스파제 패밀리 (9 멤버: 1, 2, 3, 4, 6, 7, 8, 9, 12); Kras; Apc
연령-관련 황반 변성	Aber; Ccl2; Cc2; cp (세룰로플라스민); Timp3; 카텝신 D; Vldlr; Ccr2
정신분열증	뉴레굴린 1 (Nrg1); Erb4 (뉴레굴린에 대한 수용체); 컴플렉스 1 (Cplx1); Tph1 트립토판 수산화효소; Tph2 트립토판 수산화효소 2; 뉴렉신 1; GSK3; GSK3a; GSK3b
삼뉴클레오타이드 반복 장애	HTT (헌팅틴병); SBMA/SMAX1/AR (케네디병); FXN/X25 (프레드리히 운동실조증); ATX3 (마카도-조셉병); ATXN1 및 ATXN2 (척수소뇌성 실조증); DMPK (근긴장성 이영양증); 아트로핀-1 및 Atm1 (DRPLA Dx); CBP (Creb-BP - 전체적 불안정성); VLDLR (알츠하이머); Atnx7; Atnx10
취약 X 증후군	FMR2; FXR1; FXR2; mGLUR5
세크레타아제 관련된 장애	APH-1 (알파 및 베타); 프레세닐린 (Psen1); 니카스트린 (Ncstn); PEN-2
프리온 - 관련된 장애	Prp
ALS	SOD1; ALS2; STX; FUS; TARDBP; VEGF (VEGF-a; VEGF-b; VEGF-c)
약물 중독	Prkce (알콜); Drd2; Drd4; ABAT (알콜); GRIA2; Grm5; Grin1; Htr1b; Grin2a; Drd3; Pdyn; Gria1 (알콜)

[0148]

질환/장애	유전자(들)
자폐증	Mecp2; BZRAP1; MDGA2; Sema5A; 뉴렉신 1; 취약 X (FMR2 (AFF2); FXR1; FXR2; Mglur5)
알츠하이머병	E1; CHIP; UCH; UBB; Tau; LRP; PICALM; Clusterin; PS1; SORL1; CR1; Vldlr; Uba1; Uba3; CHIP28 (Aqp1, Aquaporin 1); Uchl1; Uchl3; APP
염증	IL-10; IL-1 (IL-1a; IL-1b); IL-13; IL-17 (IL-17a (CTLA8); IL-17b; IL-17c; IL-17d; IL-17f); IL-23; Cx3cr1; ptpn22; TNFa; NOD2/CARD15 for IBD; IL-6; IL-12 (IL-12a; IL-12b); CTLA4; Cx3cl1
파킨슨병	α-시누클레인; DJ-1; LRRK2; 파킨; PINK1

[0149]

[0150]

일부 구현예에서, 방법은 판정 단계에 기초하여 대상을 진단하는 단계, 예컨대 검출된 원인 유전 변이와 연관된 질환을 갖는 대상을 진단하는 단계, 또는 상기 환자가 상기 질환을 갖거나 생길 가능성을 보고하는 단계를 추가로 포함한다. 질환, 연관된 유전자, 및 연관된 서열 변이의 예가 본원에 제공되어 있다. 일부 구현예에서, 결과는 본원에 기재된 바와 같은 보고서 생성기를 통해 보고된다.

- [0151] 일부 구현예에서, 하나 이상의 원인 유전 변이들은 특정한 유형 또는 단계의 암과 연관된 서열 변이, 또는 특정한 특징 (예컨대, 전이 잠재력, 약제 내성, 약물 민감성)을 갖는 암의 서열 변이이다. 일부 구현예에서, 본 개시내용은, 예컨대 어떤 돌연변이가 환자 결과와 연관된 것으로 공지된 경우, 예후를 결정하는 방법을 제공한다. 예를 들면, ctDNA는 순환 종양 세포의 전통적인 암 항원 53(CA-53) 및 목록보다 유방암 예후에 대한 우수한 바이오마커인 것으로 나타났다 (예컨대, Dawson, et al., N Engl J Med 368:1199 (20 13) 참고). 또한, 본 개시내용의 방법은 치료 결정, 지도 및 관찰뿐만 아니라 암 요법의 개발 및 임상 시험에서 사용될 수 있다. 예를 들면, 치료 효능은 분자 표적화된 요법 (단클론성 약물), 화학요법 약물, 방사선 프로토콜 등 또는 이들의 조합과 같은 특정한 요법을 이용한 치료 전, 동안, 및 후로부터 얻은 환자 ctDNA 샘플을 비교함으로써 관찰될 수 있다. 예를 들면, ctDNA는 치료 후 어떤 돌연변이가 증가하거나 감소하는지, 새로운 돌연변이가 나타나는지 등을 살펴보기 위하여 관찰될 수 있고, 이는 의사가 환자 증상을 추적하는 관찰 방법에 의해 제공되는 것보다 훨씬 더 짧은 시간 내에 치료를 변경하게 (예를 들면, 치료를 계속하거나, 멈추거나, 변화시키게) 할 수 있다. 일부 구현예에서, 방법은 판정 단계에 기초하여 대상을 진단하는 단계, 예컨대 검출된 서열 변이와 연관된 특정한 단계 또는 유형의 암을 갖는 대상을 진단하는 단계, 또는 상기 환자가 상기 질환을 갖거나 생길 가능성을 보고하는 단계를 추가로 포함한다.
- [0152] 예를 들면, 분자 마커(예컨대, Herceptin 및 her2/neu status)에 기초하여 환자에게 특이적으로 표적화되는 요법의 경우, 환자는 어떤 돌연변이가 이들 종양에 존재하는지 알아내기 위해 시험되고, 이들 돌연변이는 상기 요법에 대한 반응 및 내성을 예측하고 상기 요법을 사용할지 여부에 관한 결정을 안내하는데 사용될 수 있다. 따라서, 치료 과정 동안 ctDNA를 검출하고 관찰하는 것은 치료 선택을 안내하는데 매우 유용할 수 있다. 일부 1차 (치료 전) 또는 2차(치료 후) 암 돌연변이는 일부 요법에 대한 암의 내성의 원인인 것으로 확인된다 (Misale et al., Nature 486(7404):532 (2012)).
- [0153] 진단, 예후, 또는 치료 결정에서 유용할 수 있는 하나 이상의 종류의 암과 연관된 다양한 서열 변이가 알려져 있다. 본 개시내용의 방법에 유용한 종양학적으로 유의한 적합한 표적 서열은, 비제한적으로 TP53 유전자, ALK 유전자, KRAS 유전자, PIK3CA 유전자, BRAF 유전자, EGFR 유전자, 및 키트 유전자에서 변이를 포함한다. 특이적으로 증폭되고/거나 서열 변이에 대해 특이적으로 분석될 수 있는 표적 서열은 암-연관된 유전자의 전부 또는 일부일 수 있다. 일부 구현예에서, 하나 이상의 서열 변이가 TP53 유전자에서 확인된다. TP53은 인간의 암에서 가장 자주 돌연변이되는 유전자 중 하나이며, 예를 들면, TP53 돌연변이는 난소암의 45%, 대장암의 43%, 및 상부 소화도흡기암의 42%에서 발견된다 (예컨대, M. Olivier, et, al. TP53Mutations in Human Cancers: Origins, Consequences, and Clinical Use. Cold Spring Harb Perspect Biol. 2010 January; 2(1) 참고). TP53의 돌연변이 상태의 규명은 임상 진단을 돕고, 예후값을 제공하며, 암 환자의 치료에 영향을 미칠 수 있다. 예를 들면, TP53 돌연변이는 신경교 세포로부터 유래된 CNS 종양에서 환자에 대한 불량한 예후의 예측자 및 만성 림프성 백혈병을 가진 환자에서 빠른 질환 진행의 예측자로서 사용될 수 있다 (예컨대, McLendon RE, et al. Cancer. 2005 Oct 15; 1 04(8): 1693-9; Dicker F, et al. Leukemia. 2009 Jan;23(1):117-24 참고). 서열 변이는 상기 유전자 내의 어느 곳에서 일어날 수 있다. 따라서, TP53 유전자의 전부 또는 일부가 본원에서 평가될 수 있다. 즉, 본원의 다른 곳에 기재된 바와 같이, 표적 특이적 구성성분 (예컨대, 표적 특이적 프라이머)이 사용되는 경우, 예를 들면, 선택된 표적에 대해 사용될 수 있는 단지 하나 이상의 선택된 하위서열 (예컨대 돌연변이 "핫스팟")보다는, 상기 유전자를 포괄하는 단편을 증폭하고 검출하기 위해, 복수의 TP53 특이적 서열이 사용될 수 있다. 대안적으로, 표적-특이적 프라이머는 하나 이상의 선택된 하위서열 (또한 용어 "핫스팟"에 의해 포괄되는, 대상의 부류 중에서 증가된 돌연변이 비율과 연관된 뉴클레오타이드 또는 뉴클레오타이드 영역)의 상류 또는 하류에 혼성화하도록 설계될 수 있다. 그러한 하위서열을 포괄하는 표준 프라이머가 설계될 수 있고/거나, 그러한 하위서열의 상류 또는 하류에 혼성화하는 B2B 프라이머가 설계될 수 있다.
- [0154] 일부 구현예에서, 하나 이상의 서열 변이는 ALK 유전자의 모두 또는 일부에서 확인된다. ALK 융합이 폐 종양의 무려 7%에서 보고되어 왔으며, 이들 중 일부는 EGFR 타이로신 키나아제 억제제(TKI) 내성과 연관되어 있다 (예컨대, Shaw et al., J Clin Oncol. Sep 10, 2009; 27(26): 4247-4253 참고). 2013년까지, 전체 ALK 타이로신 키나아제 도메인에 걸쳐있는 몇 가지 상이한 점 돌연변이가 ALK 타이로신 키나아제 억제제(TKI)에 2차 내성을 갖는 환자에서 발견되었다 (Katayama R 2012 Sci Transl Med. 2012 Feb 8;4(120)). 따라서, ALK 유전자에서의 돌연변이 검출은 암 요법 결정을 돕는데 사용될 수 있다.
- [0155] 일부 구현예에서, 하나 이상의 서열 변이는 KRAS 유전자의 모두 또는 일부에서 확인된다. 폐 선암을 가진 환자의 약 15-25% 및 결장암을 가진 환자의 40%가 종양 연관된 KRAS 돌연변이를 갖는 것으로 보고되었다 (예컨대, Neuman 2009, Pathol Res Pract. 2009;205(12):858-62 참고). 상기 돌연변이 대부분은 KRAS 유전자의 코돈 12,

13, 및 61에 위치한다. 이들 돌연변이는 KRAS 신호전달 경로를 활성화시켜, 종양 세포의 성장 및 증식을 촉발시킨다. 일부 연구들은 KRAS 돌연변이를 갖는 종양 환자가 단독으로 또는 화학요법과 조합된 항-EGFR 항체 요법으로부터 혜택을 볼 것 같지 않다는 것을 보여준다 (예컨대, Amado et al. 2008 J Clin Oncol. 2008 Apr 1;26(10): 1626-34, Bokemeyer et al. 2009 J Clin Oncol. 2009 Feb 10;27(5):663-71 참고). 서열 변이를 확인하기 위해 표적화될 수 있는 서열 변이를 위한 한 가지 특정한 "핫스팟"은 상기 유전자의 위치 35에 있다. KRAS 서열 변이의 확인은 치료 선택에서, 예컨대 결장암을 가진 대상을 위한 치료 선택에서 사용될 수 있다.

[0156] 일부 구현예에서, 하나 이상의 서열 변이는 PIK3CA 유전자의 모두 또는 일부에서 확인된다. PIK3CA에서의 체세포 돌연변이는 다양한 유형의 암에서 빈번하게 발견되었으며, 예를 들면, 결장 암의 10-30%에서 발견되었다 (예컨대, Samuels et al. 2004 Science. 2004 Apr 23;304(5670):554. 참고). 이들 돌연변이는 엑손 9 (나선형 도메인) 및 엑손 20 (키나아제 도메인) 내의 2개의 "핫스팟" 영역 내에 가장 흔하게 위치하며, 이는 검출 서열 변이에 대한 증폭 및/또는 분석을 위해 특이적으로 표적화될 수 있다. 위치 3140 역시 특이적으로 표적화될 수 있다.

[0157] 일부 구현예에서, 하나 이상의 서열 변이는 BRAF 유전자의 모두 또는 일부에서 확인된다. 모든 악성 흑색종의 거의 50%가 BRAF에서 체세포 돌연변이를 갖는 것으로 보고되었다 (예컨대, Maldonado et al., J Natl Cancer Inst. 2003 Dec 17;95(24):1878-90 참고). BRAF 돌연변이는 모든 흑색종 아형에서 발견되지만, 만성적인 햇빛-유발 손상이 없는 피부로부터 유래된 흑색종에서 가장 빈번하다. 흑색종에서 가장 흔한 BRAF 돌연변이 중에는 미스센스 돌연변이 V600E가 있으며, 이는 위치 600의 발린을 글루타민으로 치환한다. BRAF V600E 돌연변이는 BRAF 억제제 요법의 임상적 이점과 연관된다. BRAF 돌연변이의 검출은 흑색종 치료 선택 및 표적화된 요법에 대한 내성의 연구에서 사용될 수 있다.

[0158] 일부 구현예에서, 하나 이상의 서열 변이는 EGFR 유전자의 모두 또는 일부에서 확인된다. EGFR 돌연변이는 비-소 세포 폐암과 흔히 연관된다 (미국에서 약 10% 및 동아시아에서 35%; 예컨대, Pao et al., Proc Natl Acad Sci US A. 2004 Sep 7;101(36):13306-11 참고). 이들 돌연변이는 전형적으로 EGFR 엑손 18-21 내에서 발생하며, 보통 이형접합성이다. 이들 돌연변이의 약 90%는 엑손 19 결실 또는 엑손 21 L858R 점 돌연변이이다.

[0159] 일부 구현예에서, 하나 이상의 서열 변이는 KIT 유전자의 모두 또는 일부에서 확인된다. 위장관 간질성 종양 (GIST)의 거의 85%가 KIT 돌연변이를 갖는 것으로 보고되었다 (예컨대, Heinrich et al. 2003 J Clin Oncol. 2003 Dec 1;21 (23):4342-9 참고). 다수의 KIT 돌연변이는 막근처 도메인 (엑손 11, 70%), 세포의 다이머화 모티프 (엑손 9, 10-15%), 타이로신 키나아제 I (TKI) 도메인 (엑손 13, 1-3%), 및 타이로신 키나아제 2 (TK2) 도메인 및 활성화 루프 (엑손 17, 1-3%)에서 발견된다. 2차 KIT 돌연변이는 표적 요법 이미타닙 후에 그리고 환자가 상기 요법에 내성이 생긴 후에 흔히 확인된다.

[0160] 본원에 기재된 방법에 따라 이의 모두 또는 일부가 서열 변이에 대해 분석될 수 있는 암과 연관된 유전자의 추가적인 비제한적인 예는 PTEN; ATM; ATR; EGFR; ERBB2; ERBB3; ERBB4; Notch1; Notch2; Notch3; Notch4; AKT; AKT2; AKT3; HIF; HIF1a; HIF3a; Met; HRG; Bcl2; PPAR 알파; PPAR 감마; WT1 (윌름스 종양); FGF 수용체 패밀 리 멤버 (5 멤버: 1, 2, 3, 4, 5); CDKN2a; APC; RB (망막아종); MEN1; VHL; BRCA1; BRCA2; AR; (안드로겐 수 용체); TSG101; IGF; IGF 수용체; Igf1 (4 변이체); Igf2 (3 변이체); Igf 1 수용체; Igf 2 수용체; Bax; Bcl2; 카스파제 패밀리 (9 멤버: 1, 2, 3, 4, 6, 7, 8, 9, 12); Kras; 및 Apc를 포함하나 이에 제한되지 않는 다. 추가적인 예가 본원의 다른 곳에 제공되어 있다. 본원에 개시된 방법에 따라 하나 이상의 서열 변이를 판정 하는 것에 기초하여 진단될 수 있는 암의 예는, 비제한적으로, 극세포종, 선방 세포 암종, 청신경종, 선단 흑자 성 흑색종, 선단한선종, 급성 호산구 백혈병, 급성 림프아구성 백혈병, 급성 거핵모구성 백혈병, 급성 단구성 백혈병, 성숙을 수반한 급성 골수아세포 백혈병, 급성 골수 수지상 세포 백혈병, 급성 골수 백혈병, 급성 전골 수구 백혈병, 범랑질종, 선암종, 선양 낭성 암종, 샘종, 샘모양 치원성 종양, 부신피질 암종, 성인 T-세포 백혈 병, 공격성 NK-세포 백혈병, AIDS-관련된 암, AIDS-관련된 림프종, 폐포 연질부 육종, 범랑아세포성 섬유종, 항 문암, 역형성 대세포 림프종, 역형성 갑상선암, 혈관면역모세포 T-세포 림프종, 혈관근육지방종, 맥관육종, 맹 장암, 별아교세포종, 비정형 기형 횡문근양 종양, 기저 세포 암종, 기저-유사 암종, B-세포 백혈병, B-세포 림 프종, 벨리니 관상 암종, 담관 암, 방광암, 모세포종, 골 암, 골 종양, 뇌간 신경아교종, 뇌종양, 유방암, 브레 너 종양, 기관지 종양, 세기관지폐포 암종, 브라운 종양, 버킷 림프종, 미공지 원발 부위 암, 유암종, 암종, 전 암상태, 음경의 암종, 미공지된 원발 부위의 암종, 암육종, 캐슬만병, 중추신경계 배아 종양, 소뇌 별아교세포 종, 뇌 별아교세포종, 자궁경부암, 담관암종, 척삭종, 연골육종, 척색종, 융모막암종, 맥락막종 유두종, 만성적 림프구성 백혈병, 만성적 단구성 백혈병, 만성적 골수성 백혈병, 만성적 골수증식성 장애, 만성적 중성구 백혈 병, 투명-세포 종양, 결장암, 결장직장암, 두개인두종, 피부 T-세포 림프종, 데고스 질환, 돌출성

피부섬유육종, 유피낭종, 결합조직형성 소원형 세포 종양, 미만성 큰 B 세포 림프종, 배엽부전성 신경상피 종양, 배아 암종, 내배엽종 종양, 자궁내막 암, 자궁내막 자궁암, 자궁내막모양 종양, 장 병증-관련된 T-세포 림프종, 상의모세포종, 뇌실막세포종, 상피모양 육종, 적백혈병, 식도암, 비강신경교세포종, 종양의 유잉 계열, 유잉 계열 육종, 유잉 육종, 두개의 생식세포 종양, 고환의 생식세포 종양, 간의 담도암, 유선의 파제트병, 나팔관 암, 태아내 태아, 섬유종, 섬유육종, 여포성 림프종, 여포성 갑상선암, 담낭암, 담낭암, 신경절교종, 신경절신경종, 위암, 위 림프종, 위장 암, 위장 유암종, 위장 기질 종양, 위장 기질 종양, 생식세포 종양, 종자세포 종, 임신성 용모막암종, 임신성 용모성 종양, 골의 거대세포 종양, 교모세포종 다형성, 신경아교종, 대뇌 교종 종, 사구 종양, 글루카곤종, 생식선아세포종, 과립막 세포 종양, 모발 세포 백혈병, 모발 세포 백혈병, 두경부 암, 두경부 암, 심장 암, 혈관모세포종, 혈관주위세포종, 혈관육종, 혈액 악성종양, 간세포 암종, 간비장 T-세포 림프종, 선천성 유방-난소암 증후군, 호지킨 림프종, 하인두 암, 시상하부 신경아교종, 염증성 유방암, 안구 내 흑색종, 소도세포 암종, 소도세포 종양, 유년성 골수단구 백혈병, 카포시 육종, 신장암, 클라스킨 종양, 크루켄베르크 종양, 후두암, 악성 흑자 흑색종, 백혈병, 구순암 및 구강암, 지방육종, 폐암, 황체종, 림프관종, 림프관육종, 림프상피종, 림프모양 백혈병, 림프종, 거대글로불린혈증, 악성 섬유질 조직구종, 골의 악성 섬유질 조직구종, 악성 신경아교종, 악성 중피종, 악성 말초 신경 덮개 종양, 악성 횡문근양 종양, 악성 트리톤 종양, 말트 림프종, 외투 세포 림프종, 미만 세포 백혈병, 종격 생식세포 종양, 종격 종양, 수질 갑상선암, 수모 세포종, 수질상피종, 흑색종, 수막종, 머켈 세포 암종, 중피종, 중피종, 잠복원발을 가진 전이성 편평상피 목암, 전이성 요상피 암종, 혼합된 플레리안 종양, 단구성 백혈병, 입 암, 점액성 종양, 다중 내분비 신조직형성 증후군, 다발성 골수종, 다발성 골수종, 균상식육종, 균상식육종, 골수이형성 질병, 골수이형성 증후군, 골수 백혈병, 골수 육종, 골수증식성 질병, 점액종, 비강 암, 비인두 암, 비인두 암종, 신생물, 신경초종, 신경교세포종, 신경섬유종, 신경종, 결절성 흑색종, 비-호지킨 림프종, 비흑색종 피부암, 비-소세포 폐암, 안구 종양학, 희소돌기별아교세포종, 희소돌기아교세포종, 종양세포종, 시신경 덮개 수막종, 구강암, 구강인두 암, 골육종, 골육종, 난소암, 난소 상피성 암, 난소 생식세포 종양, 난소 낮은 악성 잠재적 종양, 유방의 파제트병, 판코스트 종양, 췌장암, 유두상 갑상선암, 유두종증, 부신경절종, 부비동 암, 부갑상선암, 음경암, 말초혈관 상피모양 세포 종양, 인두 암, 크롬친화세포종, 중간체 분화의 송과체 실질 종양, 송과체아세포종, 과립세포종, 뇌하수체 샘종, 뇌하수체 종양, 형질 세포 신생물, 흉막폐 모세포종, 다배아종, 전구체 T-림프아구성 림프종, 원발성 중추신경계 림프종, 일차 삼출 림프종, 일차 간세포 암, 일차 간암, 일차 복막 암, 원시 신경외배엽성 종양, 전립선암, 복막 가점액종, 직장암, 신장 세포 암종, 염색체 15 상의 NUT 유전자와 관련된 기도 암종, 망막모세포종, 횡문근종, 횡문근육종, 리히터 형질전환, 천골미골 기형종, 타액샘 암, 육종, 신경초종증, 피지샘 암종, 2차 신생물, 정상피종, 장액 종양, 세르톨리-라이디히 세포 종양, 성기삭-기질 종양, 세자리(sezary) 증후군, 반지 세포 암종, 피부암, 청색 소원형 세포 종양, 소세포 암종, 소세포 폐암, 소세포 림프종, 작은 창자 암, 연조직 육종, 소마토스타티노마, 검맹 사마귀, 척수 종양, 척추 종양, 비장 변연부 림프종, 편평상피 세포 암종, 위암, 표재 확장성 흑색종, 천막상 원시 신경외배엽성 종양, 표면 상피성-기질 종양, 활막 육종, T-세포 급성 림프아구성 백혈병, T-세포 거대 과립 림프구 백혈병, T-세포 백혈병, T-세포 림프종, T-세포 전립프구 백혈병, 기형종, 말단 림프 암, 고환암, 난포막종, 인후두암, 흉선 암종, 흉선종, 갑상선암, 신장 골반 및 요관의 이행 세포 암, 이행 세포 암종, 요막관 암, 요도 암, 비뇨생식기 신생물, 자궁 육종, 포도막 흑색종, 질암, 베너 모리스 증후군, 사마귀모양 암종, 시각적 경로 신경아교종, 외음부암, 발덴스트롬 거대글로불린혈증, 와르틴 종양, 윌름스 종양, 및 이들의 조합을 포함한다. 암과 연관된 특정 서열 변이의 비제한적인 예가 표 6에 제공되어 있다.

표 6

유전자	돌연변이 아미노산 뉴클레오타이드	흑색종에서 COSMIC 돌연변이 %	임상적 관련성
BRAF	V600E 1799T>A	50% (BRAF 돌연변이 중에서 80-90%)	BRAF 억제제에 대한 민감도를 증가시킴 MEK 억제제에 대한 민감도를 증가시킴
BRAF	V600E 1799_1800delTGinsAA	50% (BRAF 돌연변이 중에서 80-90%)	BRAF 억제제에 대한 민감도를 증가시킴 MEK 억제제에 대한 민감도를 증가시킴
BRAF	V600R 1798_1799delGTinsAG	50% (BRAF 돌연변이 중에서 <5%)	BRAF 억제제에 대한 민감도를 증가시킴 MEK 억제제에 대해 반응함
BRAF	V600M 1798G>A	50% (BRAF 돌연변이 중에서 <1%)	BRAF 억제제에 대한 민감도를 증가시킴 MEK 억제제에 대해 반응함
BRAF	V600K 1798_1799delGTinsAA	50% (BRAF 돌연변이 중에서 5%)	BRAF 억제제에 대한 민감도를 증가시킴 MEK 억제제에 대한 민감도를 증가시킴
BRAF	V600G 1799T>G	50% (BRAF 돌연변이 중에서 <1%)	BRAF 억제제에 대한 민감도를 증가시킴 MEK 억제제에 반응함
BRAF	V600D 1799-800delTGinsAT	50% (BRAF 돌연변이 중에서 <5%)	BRAF 억제제에 대한 민감도를 증가시킴 MEK 억제제에 반응함
BRAF	L597V 1789C>G	50% (BRAF 돌연변이 중에서 1%)	BRAF 억제제에 대해 반응함 MEK 억제제에 대해 반응함
BRAF	L597S	50% (BRAF 돌연변이	BRAF 억제제에 대해 반응함

[0161]

유전자	돌연변이 아미노산 뉴클레오타이드	흑색종에서 COSMIC 돌연변이 %	임상적 관련성
	1789_1790delCTinsTC	중에서 <1%)	MEK 억제제에 대해 반응함
BRAF	L597Q 1790T>A	50% (BRAF 돌연변이 중에서 <1%)	BRAF 억제제에 대해 반응함 MEK 억제제에 대해 반응함
BRAF	L597R 1790T>G	50% (BRAF 돌연변이 중에서 <1%)	BRAF 억제제에 대해 반응함 MEK 억제제에 대해 반응함
BRAF	D594N 1799_1780delTGinsGA	50% (BRAF 돌연변이 중에서 <1%)	BRAF 억제제에 대해 반응하지 않음 MEK 억제제에 대해 반응할 수 있음
BRAF	D594H 1780C>G	50% (BRAF 돌연변이 중에서 <1%)	BRAF 억제제에 대해 반응하지 않음 MEK 억제제에 대해 반응할 수 있음
BRAF	D594E 1782T>A	50% (BRAF 돌연변이 중에서 <1%)	BRAF 억제제에 대해 반응하지 않음 MEK 억제제에 대해 반응할 수 있음
BRAF	D594E 1782T>G	50% (BRAF 돌연변이 중에서 <1%)	BRAF 억제제에 대해 반응하지 않음 MEK 억제제에 대해 반응할 수 있음
BRAF	D594G 1781A>G	50% (BRAF 돌연변이 중에서 <1%)	BRAF 억제제에 대해 반응하지 않음 MEK 억제제에 대해 반응할 수 있음
BRAF	D594V 1781A>T	50% (BRAF 돌연변이 중에서 <1%)	BRAF 억제제에 대해 반응하지 않음 MEK 억제제에 대해 반응할 수 있음
NRAS	G12D 35G>A	13-25% (NRAS 돌연변이 중에서 4%)	세포독성 화학요법
NRAS	G13R 37G>C	13-25% (NRAS 돌연변이 중에서 2%)	세포독성 화학요법 BRAF 억제제에 대해 반응하지 않음 MEK 억제제에 대해 반응할 수 있음
NRAS	G13D 38G>A	13-25% (NRAS 돌연변이 중에서 2%)	세포독성 화학요법 BRAF 억제제에 대해 반응하지 않음 MEK 억제제에 대해 반응할 수 있음
NRAS	G13V 38G>T	13-25% (NRAS 돌연변이 중에서 2%)	세포독성 화학요법 BRAF 억제제에 대해 반응하지 않음 MEK 억제제에 대해 반응할

[0162]

유전자	돌연변이 아미노산 뉴클레오타이드	흑색종에서 COSMIC 돌연변이 %	임상적 관련성
			수 있음
NRAS	Q61K 181C>A	13-25% (NRAS 돌연변이 중에서 34%)	MEK 억제제에 대해 반응함
NRAS	Q61L 182A>T	13-25% (NRAS 돌연변이 중에서 8%)	MEK 억제제에 대해 반응함
NRAS	Q61R 182A>G	13-25% (NRAS 돌연변이 중에서 35%)	MEK 억제제에 대해 반응함
NRAS	Q61H 183A>C	13-25% (NRAS 돌연변이 중에서 2%)	MEK 억제제에 대해 반응함
NRAS	Q61H 183A>T	13-25% (NRAS 돌연변이 중에서 2%)	MEK 억제제에 대해 반응함
CTNNB1	S37F 110C>T	원발성 포도막흑색종에서 ~2-3% (CTNNB1 돌연변이 중에서 46%)	
CNA11	Q209L 626A>T	원발성 포도막흑색종에서 34% (CNA11 돌연변이 중에서 92%)	
CNA11	Q209P 626A>C	원발성 포도막흑색종에서 34% (CNA11 돌연변이 중에서 1%)	
GNAQ	Q209L 626A>T	원발성 포도막흑색종에서 ~50% (CNAQ 돌연변이 중에서 ~33%)	MEK 억제제에 대해 민감함
GNAQ	Q209P 626A>C	원발성 포도막흑색종에서 ~50% (CNAQ 돌연변이 중에서 ~64%)	MEK 억제제에 대해 민감함 MEK 억제제에 대해 민감함
GNAQ	Q209R 626A>G	원발성 포도막흑색종에서 ~50% (CNAQ 돌연변이 중에서 ~2%)	
KIT	K624E 1924A>G	KIT 돌연변이 악성 흑색종 중에서 ~20%	KIT 억제제에 대한 민감도를 증가시킴
KIT	D816H 2446G>C	KIT 돌연변이 악성 흑색종 중에서 ~5%	KIT 억제제에 대한 민감도를 증가시킴
KIT	L567P 1727T>C	KIT 돌연변이 악성 흑색종 중에서 ~25%	KIT 억제제에 대한 민감도를 증가시킴
KIT	V559A 1676T>C	KIT 돌연변이 악성 흑색종 중에서 ~20%	KIT 억제제에 대한 민감도를 증가시킴

[0163]

유전자	돌연변이 아미노산 뉴클레오타이드	흑색종에서 COSMIC 돌연변이 %	임상적 관련성
KIT	V559D 1676T>A	키트 돌연변이 악성 흑색종 중에서 ~5%	KIT 억제제에 대한 민감도를 증가시킴
KIT	W557R 1669T>C	키트 돌연변이 악성 흑색종 중에서 ~10%	KIT 억제제에 대한 민감도를 증가시킴
KIT	W557R 1669T>A	키트 돌연변이 악성 흑색종 중에서 ~10%	KIT 억제제에 대한 민감도를 증가시킴

[0164]

[0165]

또한, 본원에 개시된 방법 및 조성물은 하나 이상의 암 유형, 병기, 또는 암 특징과 연관된 새로운 희귀 돌연변이를 발견하는데 유용할 수 있다. 예를 들면, 분석하에 특징 (예컨대, 특정한 질환, 암 유형, 암의 병기 등)을 공유하는 개체의 집단은 서열 변이 또는 서열 변이의 유형(예컨대, 돌연변이, 특히 유전자 또는 유전자의 일부)을 확인하기 위해 본 개시내용에 따라 서열 변이를 검출하는 방법을 거칠 수 있다. 특징이 없는 개체에서보다 특징을 공유하는 개체의 그룹 사이에 통계학적으로 유의미하게 큰 빈도로 발생하는 것으로 확인된 서열 변이는 상기 특징과의 연관도를 할당받을 수 있다. 그리고 나서, 그렇게 확인된 서열 변이 또는 서열 변이의 유형은 이들을 갖는 것으로 발견된 개체를 진단하거나 치료하는데 사용될 수 있다.

[0166]

다른 치료적 적용은 비-침습적 태아 진단에서의 사용을 포함한다. 태아 DNA는 산모의 혈액에서 발견될 수 있다. 본원에 기재된 방법 및 조성물은 순환 태아 DNA에서 서열 변이를 확인하는데 사용될 수 있고, 따라서 태아에서

하나 이상의 유전적 질환, 예컨대 하나 이상의 원인 유전 변이와 연관된 질환을 질환을 진단하는데 사용될 수 있다. 원인 유전 변이의 비제한적인 예는 본원에 기재되어 있으며, 이는 삼염색체증(trisomy), 낭포성 섬유증, 겸상-세포 빈혈, 및 테이-삭스병을 포함한다. 이러한 구현예에서, 어머니는 비교를 위해 사용될 대조군 샘플 및 혈액 샘플을 제공할 수 있다. 상기 대조군 샘플은 임의의 적합한 조직일 수 있고, 전형적으로 세포 DNA를 추출하기 위해 처리될 것이며, 이후 시퀀싱되어 참조 서열을 제공할 수 있다. 그리고 나서, 태아 지놈 DNA에 상응하는 cfDNA의 서열이 어머니의 참조와 대비하여 서열 변이로서 확인될 수 있다. 아버지 역시 태아 서열, 및 서열 변이를 확인하는데 도움을 주는 참조 샘플을 제공할 수 있다.

[0167]

추가적인 치료적 적용은, 예컨대 병원균 (예컨대, 박테리아, 바이러스, 진균류, 및 미생물)으로부터 외인성 폴리뉴클레오타이드를 검출하는 것을 포함하며, 이 정보는 진단 및 치료 선택을 통지한다. 예를 들면, 일부 HIV 하위유형은 약제 내성과 상관관계가 있다 (예컨대, hivdb.stanford.edu/pages/genotype-rx 참조). 유사하게, HCV 유형분석, 하위유형분석 및 이소타입 돌연변이는 또한 본 개시내용의 방법 및 조성물을 이용하여 수행될 수 있다. 더욱이, HPV 하위유형이 자궁경부암의 위험성과 상관된 경우, 그와 같은 진단은 암 위험성의 평가를 추가로 통지할 수 있다. 검출될 수 있는 바이러스의 추가적인 비제한적인 예는 헤파드나바이러스 B형 간염 바이러스 (HBV), 우드치크 간염 바이러스, 얼룩다람쥐 (헤파드나비리다에) 간염 바이러스, 오리 B형 간염 바이러스, 왜가리 B형 간염 바이러스, 헤르페스바이러스 단순 포진 바이러스(HSV) 유형 1 및 2, 수두-대상포진 바이러스, 사이토메갈로바이러스(CMV), 인간 사이토메갈로바이러스(HCMV), 마우스 사이토메갈로바이러스(MCMV), 기니아 피그 사이토메갈로바이러스(GPCMV), 엡슈타인-바르 바이러스(EBV), 인간 헤르페스 바이러스 6 (HHV 변종 A 및 B), 인간 헤르페스 바이러스 7 (HHV-7), 인간 헤르페스 바이러스 8 (HHV-8), 카포시 육종-관련된 헤르페스 바이러스 (KSHV), B 바이러스 폭스바이러스 우두 바이러스, 두창 바이러스, 두창 바이러스, 원숭이두창 바이러스, 소두창 바이러스, 낙타두창 바이러스, 엑트로멜리아 바이러스, 마우스두창 바이러스, 토끼두창 바이러스, 라쿤두창 바이러스, 전염성 연속종 바이러스, 양두(orf) 바이러스, 젖짜는 사람의 결절 바이러스(milker's nodes virus), 소의 유행성 구내염 바이러스, 양두 바이러스, 산양두 바이러스, 림프 스킨 질병 바이러스, 계두 바이러스, 카나리두창 바이러스, 구두(pigeonpox) 바이러스, 참새두창 바이러스, 점액종 바이러스, 토끼(hare) 섬유종 바이러스, 토끼 섬유종 바이러스, 다람쥐 섬유종 바이러스, 돼지두창 바이러스, 타나폭스 바이러스, 야바폭스 바이러스, 플라비바이러스 탕기열 바이러스, C형 간염 바이러스(HCV), GB 간염 바이러스(GBV-A, GBV-B 및 GBV-C), 웨스트 나일 바이러스, 황열병 바이러스, 세인트 루이스 뇌염 바이러스, 일본 뇌염 바이러스, 포와산 바이러스, 진드기 매개 뇌염 바이러스, 키아사누르 포레스트 질병 바이러스, 토가바이러스, 베네주엘라 말 뇌염(VEE) 바이러스, 치쿤군야 바이러스, 로스 리버 바이러스, 마야로 바이러스, 신드비스 바이러스, 풍진 바이러스, 레트로바이러스 인간 면역결핍 바이러스(HIV) 유형 1 및 2, 인간 T 세포 백혈병 바이러스(HTLV) 유형 1, 2, 및 5, 마우스 유선 종양 바이러스(MMTV), 루(Rous) 육종 바이러스(RSV), 렌티바이러스, 코로나바이러스, 중증 급성 호흡기 증후군(SARS) 바이러스, 필로바이러스 에볼라 바이러스, 마르부르그 바이러스, 메타뉴모바이러스(MPV) 예컨대 인간 메타뉴모바이러스(HMPV), 램도바이러스 광견병 바이러스, 소포성 구내염 바이러스, 분야바이러스, 크립-콩고 출혈열 바이러스, 리프트밸리 열 바이러스, 라크로스 바이러스, 한탄 바이러스, 오르토믹소바이러스, 인플루엔자 바이러스 (유형 A, B, 및 C), 파라믹소바이러스, 파라인플루엔자 바이러스 (PIV 유형 1, 2 및 3), 호흡기 세포융합 바이러스 (유형 A 및 B), 홍역 바이러스, 볼거리 바이러스, 아레나바이러스, 림프구 맥락수막염 바이러스, 쥘린 바이러스, 마쉴포 바이러스, 파나리토 바이러스, 라사 바이러스, 암파리 바이러스, 플렉살 바이러스, 입파이 바이러스, 모발라 바이러스, 모페이아 바이러스, 란티게 바이러스, 파라나 바이러스, 피친데 바이러스, 폰타 토로 바이러스 (PTV), 타카리베 바이러스 및 타미아미 바이러스를 포함한다.

[0168]

본 개시내용의 방법에 의해 검출될 수 있는 박테리아 병원체의 예는 비제한적으로, 그 중에서도 아시네토박터 바우마니이, 악티노바실러스 종, 액티노마이세테스, 악티노마이세스 종(예컨대 악티노마이세스 이스라엘이 및 악티노마이세스 나에슬룬디이), 에어로모나스 종(예컨대 에어로모나스 하이드로필라, 에어로모나스 베로니이 바 이오바르 소브리아 (에어로모나스 소브리아), 및 에어로모나스 카비아에), 아나플라즈마 파고사이토펴, 알칼리게네스 크실로숙시단스, 아시네토박터 바우마니이, 악티노바실러스 악티노마이세테스코미탄스, 바실러스 종(예컨대 바실러스 안트라시스, 바실러스 세레우스, 바실러스 서브틸리스, 바실러스 투링기엔시스, 및 바실러스 스테아로테모필루스), 박테로이데스 종(예컨대 박테로이데스 프라길리스), 바르토넬라 종(예컨대 바르토넬라 바실리포르미스 및 바르토넬라 헨셀라에, 비피도박테리움 종, 보테텔라 종(예컨대 보테텔라 페르투스시스, 보테텔라 파라페르투스시스, 및 보테텔라 브론치세ptica), 보렐리아 종(예컨대 회귀열 보렐리아, 및 보렐리아 버그도르페리), 브루셀라 종(예컨대 브루셀라 아보르투스, 브루셀라 카니스, 브루셀라 멜리텐시스 및 브루셀라 수이스), 버크홀데리아 종(예컨대 버크홀데리아 슈포말레이 및 버크홀데리아 세파시아), 캄필로박터 종(예컨대 캄필로박터 제주니, 캄필로박터 콜라이, 캄필로박터 라리 및 캄필로박터 페투스), 카프노사이토파가 종, 카디오박테리움

호미니스, 클라미디아 트라코마티스, 클라마이도필라 뉴모니아에, 클라마이도필라 씨타시, 사이트로박터 중 콕시엘라 부르네티이, 코리네박테리움 중(예컨대, 코리네박테리움 디프테리애, 코리네박테리움 제이케이움 및 코리네박테리움), 클로스트리듐 중(예컨대 클로스트리듐 페르프린젠스, 클로스트리듐 디피실레, 클로스트리듐 보툴리눔 및 클로스트리듐 테타니), 에이케넬라 코로덴스, 엔테로박터 중(예컨대 엔테로박터 에어로게네스, 엔테로박터 아글로메란스, 엔테로박터 클로아케 및 기회감염 에스케리치아 콜라이를 포함하는 에스케리치아 콜라이, 예컨대 장독소생성 E. 콜라이, 장침입성 E. 콜라이, 장병원성 E. 콜라이, 장출혈성 E. 콜라이, 장응집성 E. 콜라이 및 요로병원성 E. 콜라이), 엔테로코쿠스 중(예컨대 엔테로코쿠스 파에칼리스 및 엔테로코쿠스 패슈), 에를리치아 중(예컨대 에를리치아 사판시스 및 에를리치아 카니스), 에리시펠로트릭스 루시오파티아에, 유박테륨 중, 프란시셀라 툴라렌시스, 푸소박테리움 뉴클레아툼, 가드넬라 바지날리스, 제멜라 보빌로움, 헤모필루스 중(예컨대 헤모필루스 인플루엔자, 헤모필루스 두크레이이, 헤모필루스 아에집티우스, 헤모필루스 파라인플루엔자, 헤모필루스 헤몰리티쿠스 및 헤모필루스 파라헤몰리티쿠스, 헬리코박터 중(예컨대 헬리코박터 파일로리 헬리코박터 시나에디 및 헬리코박터 펜넬리아에), 킨젤라 킨기, 클렙시엘라 중(예컨대 클렙시엘라 뉴모니아에, 클렙시엘라 그라눌라마티스 및 클렙시엘라 옥시토카), 락토바실러스 중, 리스테리아 모노사이토게네스, 랩토스피라 인테로간스, 레지오넬라 뉴모필라, 랩토스피라 인테로간스, 랩토스트렙토코쿠스 중, 모락셀라 카타르할리스, 모르가넬라 중, 모빌쿤쿠스 중, 마이크로코쿠스 중, 마이코박테리움 중(예컨대 마이코박테리움 레프라에, 마이코박테리움 투베르쿨로시스, 마이코박테리움 인트라셀룰라레, 마이코박테리움 아비움, 마이코박테리움 보비스, 및 마이코박테리움 마리넴), 마이코플라스마 중(예컨대 마이코플라스마 뉴모니아에, 마이코플라스마 호미니스, 및 마이코플라스마 제니탈리움), 노카르디아 중 (예컨대 노카르디아 아스테로이데스, 노카르디아 사이리아시게 오르기가 및 노카르디아 브라실리엔시스), 나이세리아 중 (예컨대 나이세리아 고노르호아에 및 나이세리아 메닌기티디스), 파스튜렐라 멀토시다, 플레시오모나스 쉬겔로이데스, 프레보텔라 중, 포르파이로모나스 중, 프레보텔라 펠라미노게니카, 프로테우스 중(예컨대 프로테우스 불가리스 및 프로테우스 미라빌리스), 프로비덴시아 중(예컨대 프로비덴시아 알칼리파엔시스, 프로비덴시아 레트게리 및 프로비덴시아 스투아르티이), 슈도모나스 에어루기노사, 프로피오니박테리움 아크네스, 로도코쿠스 에퀴, 리케차 중(예컨대 리케차 리케트시, 리케차 아카리 및 리케차 리케트시아, 오리엔티아 츠즈가무시 (예전에: 리케차 츠즈가무시) 및 리케차 타이파이), 로도코쿠스 중, 세라티아 마르세스센스, 스테노트로포모나스 말토필리아, 살모넬라 중(예컨대 살모넬라 엔테리카, 살모넬라 타이피, 살모넬라 파라타이피, 살모넬라 엔테리티디스, 살모넬라 콜레라uis 및 살모넬라 타이피유리움), 세라티아 중(예컨대 세라티아 마르세산스 및 세라티아 리쿠에파시엔스), 시겔라 중(예컨대 시겔라 다이센테리애, 시겔라 플렉스네리, 시겔라 보이디이 및 시겔라 손네이), 스타필로코쿠스 중(예컨대 스타필로코쿠스 아우레스, 스타필로코쿠스 에피더미디스, 스타필로코쿠스 헤몰리티쿠스, 스타필로코쿠스 사프로피티쿠스), 스트렙토코쿠스 중(예컨대 스트렙토코쿠스 뉴모니아에 (예를 들면 클로르암페니콜-내성 혈청형 4 스트렙토코쿠스 뉴모니아에, 스펙티노마이신-내성 혈청형 6B 스트렙토코쿠스 뉴모니아에, 스트렙토마이신-내성 혈청형 9V 스트렙토코쿠스 뉴모니아에, 에리트로마이신-내성 혈청형 14 스트렙토코쿠스 뉴모니아에, 옴토힌-내성 혈청형 14 스트렙토코쿠스 뉴모니아에, 리팜피신-내성 혈청형 18C 스트렙토코쿠스 뉴모니아에, 테트라사이클린-내성 혈청형 19F 스트렙토코쿠스 뉴모니아에, 페니실린-내성 혈청형 19F 스트렙토코쿠스 뉴모니아에, 및 트리메토프림-내성 혈청형 23F 스트렙토코쿠스 뉴모니아에, 클로르암페니콜-내성 혈청형 4 스트렙토코쿠스 뉴모니아에, 스펙티노마이신-내성 혈청형 6B 스트렙토코쿠스 뉴모니아에, 스트렙토마이신-내성 혈청형 9V 스트렙토코쿠스 뉴모니아에, 옴토힌-내성 혈청형 14 스트렙토코쿠스 뉴모니아에, 리팜피신-내성 혈청형 18C 스트렙토코쿠스 뉴모니아에, 페니실린-내성 혈청형 19F 스트렙토코쿠스 뉴모니아에, 또는 트리메토프림-내성 혈청형 23F 스트렙토코쿠스 뉴모니아에), 스트렙토코쿠스 아갈락티아에, 스트렙토코쿠스 류탄스, 스트렙토코쿠스 파이오제네스, 그룹 A 스트렙토코쿠스, 스트렙토코쿠스 파이오제네스, 그룹 B 스트렙토코쿠스, 스트렙토코쿠스 아갈락티아에, 그룹 C 스트렙토코쿠스, 스트렙토코쿠스 안기노서스, 스트렙토코쿠스 에퀴스밀리스, 그룹 D 스트렙토코쿠스, 스트렙토코쿠스 보비스, 그룹 F 스트렙토코쿠스, 및 스트렙토코쿠스 안기노서스 그룹 G 스트렙토코쿠스), 스피릴룸 마이너스, 스트렙토바실러스 모닐리포르미, 트레포네마 중 (예컨대 트레포네마 카라테움, 트레포네마 페르테누, 트레포네마 팔리둠 및 트레포네마 엔디미쿰, 트로페리마 위펠리, 우레아플라즈마 우레알리티쿰, 베일로넬라 중, 비브리오 중(예컨대 비브리오 콜레라에, 비브리오 파라헤몰리티쿠스, 비브리오 불니피쿠스, 비브리오 파라헤몰리티쿠스, 비브리오 불니피쿠스, 비브리오 알기놀리티쿠스, 비브리오 미미쿠스, 비브리오 홀리사에, 비브리오 플루비알리스, 비브리오 메치니코비, 비브리오 담셀라 및 비브리오 퍼니시이), 예르시니아 중 (예컨대 예르시니아 엔테로콜리티카, 예르시니아 페스티스, 및 예르시니아 슈도투베르쿨로시스) 및 크산토모나스 말토필리아 중 어느 하나 이상 (또는 어느 조합)을 포함한다.

[0169]

일부 구현예에서, 본 개시내용의 방법 및 조성물은 장기 이식 수용체를 관찰하는데 사용된다. 전형적으로, 공여

체 세포로부터의 폴리뉴클레오타이드는 수용체 세포로부터의 폴리뉴클레오타이드의 배경에서의 순환에서 발견될 것이다. 만약 장기가 잘 받아들여지면 상기 공여체 순환 DNA의 수준은 일반적으로 안정할 것이며, 공여체 DNA의 빠른 증가 (예컨대, 주어진 샘플에서의 빈도로서)는 이식 거부의 조기 징조로 사용될 수 있다. 이식 실패를 방지하기 위해 이 단계에서 치료가 제공될 수 있다. 공여체 장기의 거부는 혈액에서 공여체 DNA의 증가를 야기하는 것으로 나타났다 (Snyder et al., PNAS 108(15):6629 (2011) 참고). 본 개시내용은 이 분야에서 선행 기술 대비 유의한 민감도 개선을 제공한다. 이러한 구현예에서, 수용체 대조군 샘플(예컨대, 볼 스왑 등) 및 공여체 대조군 샘플이 비교를 위해 사용될 수 있다. 상기 수용체 샘플은 상기 참조 서열을 제공하는데 사용될 수 있는 반면, 공여체의 지놈에 상응하는 서열은 상기 참조 대비 서열 변이로서 확인될 수 있다. 관찰은 일정 기간 동안 수용체로부터 샘플(예컨대, 혈액 샘플)을 수득하는 것을 포함한다. 초기 샘플(예컨대, 처음 몇 주 이내)은 공여체 cfDNA의 비율에 대한 기준선 (baseline)을 확립하는데 사용될 수 있다. 이후의 샘플들은 상기 기준선과 비교될 수 있다. 일부 구현예에서, 약 또는 적어도 약 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 95%, 100%, 250%, 500%, 1000% 이상의 공여체 cfDNA의 비율 증가는, 수용체가 공여체 조직을 거부하는 과정에 있음을 가리키는 역할을 할 수 있다.

[0170] 실시예

[0171] 하기 실시예는 본 발명의 다양한 구현예를 예시하기 위한 목적으로 제공되며, 어떤 식으로든 본 발명을 제한하고자 하는 것이 아니다. 본 실시예는 본원에 기재된 방법과 더불어 바람직한 구현예를 나타내고, 예시적이며, 본 발명의 범위로 제한하고자 하는 것이 아니다. 이들에서의 변화 및 청구범위에 의해 정의된 본 발명의 취지에 속하는 다른 용도가 본 기술분야의 숙련가에게 떠오를 것이다.

[0172] 실시예 1: 돌연변이 검출을 위한 직렬 반복 시퀀싱 라이브러리 제조

[0173] 12 μ L 물 또는 10mM Tris-HCl pH 8.0 중의 >10ng의 ~150bp DNA 단편으로 시작하여, 2 μ L 10X CircLigase 버퍼 믹스를 부가하고 혼합물을 2분간 95°C로 가열하고 얼음 위에서 5분간 냉각시켰다. 여기에, 4 μ L 5M 베타인, 1 μ L 50mM MnCl₂, 및 1 μ L CircLigase II를 부가하였다. 상기 반응을 60°C에서 적어도 12시간 동안 인큐베이션 하였다. 그 다음, 2 μ L RCA 프라이머 믹스 (각각 50nM, 5nM 최종 농도로)를 부가하고 혼합하였다. 상기 혼합물을 2분간 95°C로 가열하고 42°C로 2시간 동안 냉각하였다. 상기 CirLigation 산물을 Zymo 올리고 뉴클레오타이드 정제 키트를 이용하여 정제하였다. 제조사의 지시에 따라, 50 μ L의 총 부피를 위해 28 μ L 물을 22 μ L CirLigation 산물에 부가하였다. 이를 100 μ L 올리고 결합 버퍼 및 400 μ L 에탄올과 혼합하였다. 이를 >10,000xg에서 30초간 회전시키고, 통과액을 버렸다. 750 μ L DNA세척 버퍼를 부가한 다음, >10,000xg에서 30초간 회전시키고, 통과액을 버리고, 최고 속도로 1분간 더 회전시켰다. 상기 컬럼을 새로운 에펜드르프 튜브로 이동시키고 17 μ L 물로 용리시켰다 (최종 용리된 부피는 약 15 μ L였음).

[0174] 약 50 μ L의 부피에서 롤링 서클 증폭을 수행하였다. 15 μ L 용리 샘플에, 5 μ L 10X RepliPhi 버퍼 (Epicentre), 1 μ L 25mM dNTP, 2 μ L 100mM DTT, 1 μ L 100U/ μ L RepliPhi Phi29, 및 26 μ L 물을 부가하였다. 상기 반응 믹스를 30°C에서 1시간 동안 인큐베이션하였다. 남아있는 세척 단계를 위해 제조사의 프로토콜에 따라 80 μ L의 Ampure 비드를 부가하여 RCA 산물을 정제하였다. 용리를 위해, 22.5 μ L 용리 버퍼를 부가하고, 상기 비드를 65°C에서 5분간 인큐베이션하였다. 가볍게 회전시킨 후, 튜브를 자석에 되돌렸다.

[0175] RCA 반응으로부터의 약 20 μ L의 용리된 산물을 25 μ L 2X Phusion 마스터 믹스, 2.5 μ L DMSO, 및 0.5 μ L의 10 μ M의 각각의 B2B 프라이머 믹스와 혼합하였다. 증폭은 하기 PCR 프로그램을 사용하였다: 95°C에서 1분, 5 사이클의 연장 (95°C에서 15초, 55°C에서 15초, 72°C에서 1분), 13-18 사이클의 복제 (95°C에서 15초, 68°C에서 15초, 72°C에서 1분), 및 72°C에서 7분간의 최종 연장. E-gel을 수행하여 PCR 산물 크기를 확인하였다. 만약 상기 범위가 100-500bp이면, 0.6X Ampure 비드 정제를 수행하여 300-500bp를 농축시키고 작은 RNA 라이브러리 어댑터 프라이머를 이용한 PCR의 또 다른 라운드를 위해 1-2ng을 취하였다. 만약 산물 크기 범위가 >1000bp이면, 산물을 1.6X Ampure 비드를 이용하여 정제하고, Nextera XT 앰플리콘 라이브러리 준비를 위해 2-3ng을 취하여 0.6X Ampure 비드 정제에 의해 400-1000bp 범위의 크기를 농축하였다.

[0176] 시퀀싱 데이터에 대한 생물정보처리를 수행하기 위해, MiSeq 실행으로부터 FASTQ 파일을 수득하였다. 상기 서열을 BWA를 이용하여 표적화된 서열 (예컨대, KRAS 및 EGFR)을 함유하는 참조 지놈 서열에 대해 FASTQ 파일로 정렬하였다. 정렬 결과를 이용하여 각각의 서열(2개의 리드)에 대해 반복 단위의 영역 및 길이 및 그의 참조 위치를 확인하였다. 상기 정렬 결과 및 각각의 서열의 반복 단위의 정보를 이용하여 모든 유전자좌에서 변이를 확인하였다. 상기 정렬 결과 및 각각의 서열의 반복 단위의 정보를 이용하여 모든 유전자좌에서의 변이를 확인하였다. 2개의 리드로부터의 결과를 조합하였다. 변이의 정규화된 빈도 및 노이즈 수준을 계산하였다. 확인된 변이

로부터의 변이 관정에서의 다수의 추가적인 기준을 적용하였고, 이는 $qscore > 30$ 및 $p-value < 0.0001$ 을 포함한다. 이러한 기준을 통과한 상기 확인된 변이를 진정한 변이(돌연변이)로서 보고하였다. 상기 과정은 컴퓨터 언어 (예컨대, 파이썬)에 의해 자동화될 수 있다.

[0177] **실시예 2: 서열 변이의 검출을 위한 직렬 반복 시퀀싱 라이브러리 제조**

[0178] 12 μ L 부피 중의 150bp 평균 길이를 갖는 10ng의 DNA 단편을 직렬 반복 시퀀싱 라이브러리 구축을 위해 사용하였다. 상기 DNA는 5' 말단에 인산염기를 부가하고 3' 말단에 수산화기를 남기기 위해 미리 T4 폴리뉴클레오타이드 키나아제 (New England Biolabs)로 처리되었다. DNase I 또는 효소적 단편화로부터 생성되거나 또는 혈청 또는 혈장으로부터 추출된 DNA 단편의 경우, 상기 말단 처리 단계를 생략하였다. 상기 DNA를 2 μ L 10X CircLigase 버퍼 (Epicentre CL9021K)와 혼합하였다. 상기 혼합물을 95°C에서 2분간 혼합하고 얼음 위에서 5분간 냉각한 다음, 4 μ L 베타인, 1 μ L 50mM $MnCl_2$, 및 1 μ L CircLigase II (Epicentre CL9021K)를 부가하였다. 상기 라이게이션 반응을 60°C에서 적어도 12시간 동안 수행하였다. 1 μ L의 200nM의 각각의 RCA 프라이머 믹스 (10nM 최종 농도)를 상기 라이게이션 산물에 부가하고 혼합하고, 96°C에서 1분 동안 가열하고, 42°C로 냉각하고, 42°C에서 2시간 동안 인큐베이션하였다.

[0179] 혼성화된 RCA 프라이머를 갖는 상기 CirLigation 산물을 Zymo 올리고 뉴클레오타이드 정제 키트(Zymo Research, D4060)를 이용하여 정제하였다. 이를 수행하기 위해, 21 μ L의 산물을 28 μ L 물 및 1 μ L의 캐리어 RNA (Sigma-Aldrich, R5636, 200 ng/ μ L with 1X TE 버퍼를 이용하여 200 ng/ μ L로 희석됨)를 이용하여 50 μ L로 희석하였다. 상기 희석된 샘플을 100 μ L 올리고 결합 버퍼 및 400 μ L의 100% 에탄올과 혼합하였다. 상기 혼합물을 컬럼에 로딩하고 >10,000xg에서 30초간 원심분리하였다. 통과액을 버렸다. 상기 컬럼을 >10,000xg에서 30초간 원심분리하여 750 μ L DNA 세척 버퍼로 세척하고, 통과액을 버리고 최고 속도에서 1분간 원심분리하였다. 상기 컬럼을 새로운 에펜도르프 튜브로 이동시키고, 상기 DNA를 17 μ L 용리 버퍼(10mM Tris-Cl pH 8.0, 최종 용리된 부피 약 15 μ L)를 이용하여 용리시켰다.

[0180] 50 μ L의 총 반응 부피를 위해, 5 μ L 10X RepliPhi 버퍼, 2 μ L 25mM dNTP, 2 μ L 100mM DTT, 1 μ L 100 U/ μ L RepliPhi Phi29, 및 25 μ L 물 (Epicentre, RH040210)을 상기 컬럼으로부터의 15 μ L 용리된 샘플에 부가하였다. 상기 반응 믹스를 30°C에서 2시간 동안 인큐베이션하였다. 80 μ L의 Ampure XP 비드 (Beckman Coulter, A63881)를 부가하여 상기 RCA 산물을 정제하였다. 세척 단계를 위해 제조사의 지침을 따랐다. 22.5 μ L 용리 버퍼에서 5분의 65°C 인큐베이션 후 RCA 산물을 용리시켰다. 상기 튜브를 잠시 원심분리한 후 자석에 되돌렸다.

[0181] 상기 RCA 반응으로부터의 용리된 산물 약 20 μ L를 25 μ L 2X Phusion 마스터 믹스(New England Biolabs M0531S), 2.5 μ L 물, 2.5 μ L DMSO, 및 0.5 μ L의 B2B 프라이머 믹스(각각 10 μ M)와 혼합하였다. 증폭을 하기 열순환 프로그램을 이용하여 수행하였다: 95°C에서 2분, 5 사이클의 연장(95°C에서 30초, 55°C에서 15초, 72°C에서 1분), 18 사이클의 복제(95°C에서 15초, 68°C에서 15초, 72°C에서 1분), 및 72°C에서 7분간의 최종 연장. 상기 PCR 산물 크기를 전기영동에 의해 확인하였다. 전기영동에 의해 긴 PCR 산물이 확인되면, 상기 PCR 산물을 정제를 위해 Ampure 비드(0.6X 부피)와 혼합하여 >500 bp PCR 산물을 농축시켰다. 상기 정제된 산물을 Qubit 2.0 정량 플랫폼(Invitrogen)을 이용하여 정량하였다. 약 1ng 정제된 DNA를 Nextera XT 앰플리콘 라이브러리 제조(Illumina FC-131-1024)를 위해 사용하였다. > 500bp의 삽입물 크기를 갖는 라이브러리 요소들을 0.6X Ampure 비드를 이용한 정제에 의해 농축시켰다.

[0182] 상기 증폭된 라이브러리의 농도 및 크기 분포를 2100 Bioanalyzer (Agilent Technologies Inc., Santa Clara, CA)용 Agilent DNA 고민감도 키트를 이용하여 분석하였다. 시퀀싱을 2-250 bp MiSeq 시퀀싱 키트를 갖는 일루미나 MiSeq를 이용하여 수행하였다. MiSeq 매뉴얼에 따라, 12pM 변성된 라이브러리를 시퀀싱 런에 로딩하였다.

[0183] 이 과정에 대한 변형에서, 일루미나 어댑터를 Nextera 제조 대신에 라이브러리 제조에서 사용하였다. 이를 수행하기 위해, 약 1ng의 유사하게 정제된 DNA를 B2B 프라이머의 유니버설 부분을 함유하는 프라이머의 쌍 및 일루미나 어댑터 서열 (P5 및 P7; 5'CAAGCAGAAGACGGCATACG3' 및 5'ACACTCTTCCCTACACGACGCTCTCCGATCT3')을 이용한 PCR 분석을 위해 사용하였다. Phusion 마스터 믹스를 이용하여, 12 사이클의 복제 단계(95°C에서 30초, 55°C에서 15초, 72°C에서 60초)를 수행하였다. 이러한 증폭 단계의 목적은 앰플리콘 시퀀싱을 위해 일루미나 어댑터를 부가하는 것이었다. >500bp 길이의 앰플리콘을 0.6X Ampure 비드를 이용하여 농축시켰다. 앰플리콘 라이브러리의 농도 및 크기 분포를 2100 Bioanalyzer를 위한 Agilent DNA 고민감도 키트(Agilent technologies Inc., Santa Clara, CA)를 이용하여 분석하였다. 시퀀싱을 2x250 bp MiSeq 시퀀싱 키트를 갖는 일루미나 MiSeq를 이용하여 수행하였다. B2B 프라이머의 유니버설 부분은 또한 시퀀싱 프라이머 서열로서 역할을 하였고, 만약 상기 프라이머가 일루미나 키트에 포함되어 있지 않으면 주문제작 시퀀싱 프라이머를 부가하였다. 12pM 변성된 라이

브러리를 시퀀싱 런에 로딩하였다.

[0184] 한 가지 예시적 분석에서의 표적 영역 커버리지가 도 33에 예시되어 있다. 하기 표 3은 상기 표적화된 영역의 분석 결과를 기재하고 있다.

[0185] 표 1은 본 개시내용의 방법에 유용한 RCA 프라이머의 예를 제공한다. 표 2는 본 개시내용의 방법에서 유용한 B2B 프라이머의 예를 제공한다.

표 1

유전자 명	RCA 프라이머 A 명	RCA 프라이머 A 서열	RCA 프라이머 B 명	RCA 프라이머 B 서열
PIK3CA	PIK3CA-E1a	GCTTTGAGCTGTTCTT TGTCATT	PIK3CA-E1b	AAAGCAATTTCTACACGA GATCC
PIK3CA	PIK3CA-E2a	TTTAATTGTGTGGAAG ATCCAATC	PIK3CA-E2b	ATTAAACAGCATGCATTG AACTG
EGFR	EGFR-E1a	CTTCTCACCTTCTGG GATCC	EGFR-E1b	AAATTCCCCTCGCTATCA AG
EGFR	EGFR-E2a	CCATCACGTAGGCTTC CTG	EGFR-E2b	ATGGCCAGCGTGGACAAC
EGFR	EGFR-E3a	GACATAGTCCAGGAG GCAGC	EGFR-E3b	TGTCCGGGAACACAAAG AC
EGFR	EGFR-E4a	AAGCGACGGTCCTCC AAG	EGFR-E4b	TGGCAGCCAGGAACGTAC
EGFR	EGFR-E5a	AGTACGTTCTGGCTG CC	EGFR-E5b	AACACCGCAGCATGTCAA G
EGFR	EGFR-E6a	ATCCACTTGATAGGC ACCTTG	EGFR-E6b	AAGTGGATGGCATTGGAA TC
EGFR	EGFR-E7a	TCTCGCTGGCAGGGA TTC	EGFR-E7b	CCTGGAGAAAGGAGAAC GC
EGFR	EGFR-E9a	AACTTTGGGCGACTAT CTGC	EGFR-E9b	AGTCCGTGAGTTGATCA TCG
EGFR	EGFR-E10a	TTGGAGTCTGTAGGA CTTGGC	EGFR-E10b	ACTTCTACCGTGCCCTGA TG
EGFR	EGFR-E11a	CTGCTGTGGGATGAG GTACTC	EGFR-E11b	CACAGCAGGGCTTCTTCA G
EGFR	EGFR-E12a	CATGGAATGCTTGATC CACATC	EGFR-E12b	CATGGGCAACTTCTCTGT TTC
BRAF	BRAF-E1a	CAGTTTGAACAGTTGT CTGGATC	BRAF-E1b	AAACTGATGGGACCCACT CC
PTEN	PTEN-E1a	TGTTTCTGCTAACGAT CTCTTG	PTEN-E1b	AGGAGATATCAAGAGGA TGGATT
PTEN	PTEN-E2a	CAGGAAATCCCATAG CAATAATG	PTEN-E2b	TCCTGCAGAAAGACTTGA AGG
PTEN	PTEN-E3a	GCTTTGAATCCAAAA ACCTTAAAC	PTEN-E3b	GGATTCAAAGCATAAAA ACCATTAC

[0186]

PTEN	PTEN-E4a	TACAGTACATTCATAC CTACCTCTGC	PTEN-E4b	TATGTTGTATAACTTAAA CCCGATAGAC
PTEN	PTEN-E5a	AAAGGATATTGTGCA ACTGTGG	PTEN-E5b	TGGAAGACCATAACCCAC CAC
PTEN	PTEN-E6a	CCATAGAAATCTAGG GCCTCTTG	PTEN-E6b	AAGTAAGGACCAGAGAC AAAAAGG
PTEN	PTEN-E7a	CCAGATGATTCTTTAA CAGGTAGC	PTEN-E7b	GGATTATAGACCAGTGGC ACTG
PTEN	PTEN-E9a	GAAC TTGTCTTCCCGT CGTG	PTEN-E9b	CATGTACTTTGAGTTCCC TCAGC
PTEN	PTEN-E12a	TCTGGTCCTGGTATGA AGAATG	PTEN-E12b	CAGGACCAGAGGAAACC TCAG
PTEN	PTEN-E13a	GCTCTATACTGCAAAT GCTATCG	PTEN-E13b	CGTGCAGATAATGACAAG GAATATC
PTEN	PTEN-E14a	TTGGAGAAAAGTATC GGTTGG	PTEN-E14b	GGTCAGTTAAATTAAACA TTTTGTGG
PTEN	PTEN-E15a	TGGTGTTACAGAAGTT GAAC TGC	PTEN-E15b	GATGTTAGTGACAATGAA CCTGATC
KRAS	KRAS-E1a	AAGAGTGCCTTGACG ATACAGC	KRAS-E1b	TCTTGCCTACGCCACCAG CAGGCCCTTCTGTCTTGA AC
TP53	TP53-E1a	CCTGACTCAGACTGA CATTCTCC	TP53-E1b	GAACATCTCGAAGCGCTC AC
TP53	TP53-E2a	ATGTTCCGAGAGCTG AATGAG	TP53-E2b	TTATGGTATAAGTTGGTG TTCTGAAG
TP53	TP53-E3a	TTAAAGGACCAGACC AGCTTTC	TP53-E3b	AGAGGTCCCAAGACTTAG TACCTG
TP53	TP53-E4a	CTTGGGACCTCTTATC AAGTGG	TP53-E4b	GCTTGCTTACCTCGCTTA GTG
TP53	TP53-E5a	AAGCAAGCAGGACAA GAAGC	TP53-E5b	TTCCGTCCCAGTAGATTA CCAC
TP53	TP53-E6a	GGGACGGAACAGCTT TGAG	TP53-E6b	CATGTAGTTGTAGTGGAT GGTGG
TP53	TP53-E7a	GTGGAGTATTTGGAT GACAGAAAC	TP53-E7b	ATACTCCACACGCAAATT TCC
TP53	TP53-E8a	TGCTCAGATAGCGAT GGTGAG	TP53-E8b	CTATCTGAGCAGCGCTCA TG
TP53	TP53-E9a	CTGTGCAGCTGTGGGT TGA	TP53-E9b	GCAGGTCTTGGCCAGTTG TGTCCCAGAATGCAAGAA GC
TP53	TP53-E10a	AAAGTCTGTGACTTGCA CGGTC	TP53-E10b	GATGACAGGGGCCAGGA G
TP53	TP53-E11a	CCTGTCATCTTCTGTC CCTTC	TP53-E11b	TGGGTCTTCAGTGAACCA TTG
TP53	TP53-E12a	AAGACCCAGGTCCAG ATGAAG	TP53-E12b	GAGCAGAAAGTCAGTCCC ATG
TP53	TP53-E13a	CTGCTCTTGTCTTTCA GACTTCC	TP53-E13b	GCTCGACGCTAGGATCTG AC
TP53	TP53-E14a	CTCTGAGTCAGGAAA CATTTTCAG	TP53-E14b	
TP53	TP53-E15a		TP53-E15b	
TP53	TP53-E16a		TP53-E16b	

[0187]

표 2

유전자 명	B2B 프라이 머 A	프라이머 A 서열	B2B 프라이 머 B	프라이머 B 서열
PIK3CA	PIK3CA- BX1a	GTTCAGAGTTCTACAGTCC GACGATCGCTTTGAGCTGT TCTTGTCAATT	PIK3CA- BX1b	CCTTGGCACCCGAGAATTCC AAAAGCAATTTCTACACGA GATCC
PIK3CA	PIK3CA- BX2a	GTTCAGAGTTCTACAGTCC GACGATCTTTAATTGTGTG GAAGATCCAATC	PIK3CA- BX2b	CCTTGGCACCCGAGAATTCC AATTAAACAGCATGCATTGA ACTG
EGFR	EGFR- BX1a	GTTCAGAGTTCTACAGTCC GACGATCCTTCTCACCTT CTGGGATCC	EGFR- BX1b	CCTTGGCACCCGAGAATTCC AAAATTCCCGTCGCTATCAA G
EGFR	EGFR- BX2a	GTTCAGAGTTCTACAGTCC GACGATCCCATCACGTAG GCTTCCTG	EGFR- BX2b	CCTTGGCACCCGAGAATTCC AATGGCCAGCGTGGACAAC
EGFR	EGFR- BX3a	GTTCAGAGTTCTACAGTCC GACGATCGACATAGTCCA GGAGGCAGC	EGFR- BX3b	CCTTGGCACCCGAGAATTCC ATGTCCGGGAACACAAAGA C
EGFR	EGFR- BX4a	GTTCAGAGTTCTACAGTCC GACGATCAAGCGACGGTC CTCCAAG	EGFR- BX4b	CCTTGGCACCCGAGAATTCC ATGGCAGCCAGGAACGTAC
EGFR	EGFR- BX5a	GTTCAGAGTTCTACAGTCC GACGATCAGTACGTTCTCTG GCTGCC	EGFR- BX5b	CCTTGGCACCCGAGAATTCC AAACACCGCAGCATGTCAA G
EGFR	EGFR- BX6a	GTTCAGAGTTCTACAGTCC GACGATCATCCACITGATA GGCACCTTG	EGFR- BX6b	CCTTGGCACCCGAGAATTCC AAAGTGGATGGCATTTGGAA TC
EGFR	EGFR- BX7a	GTTCAGAGTTCTACAGTCC GACGATCTCTCGCTGGCA GGGATTC	EGFR- BX7b	CCTTGGCACCCGAGAATTCC ACCTGGAGAAAGGAGAACG C
EGFR	EGFR- BX9a	GTTCAGAGTTCTACAGTCC GACGATCAACTTTGGGCG ACTATCTGC	EGFR- BX9b	CCTTGGCACCCGAGAATTCC AAGTTCCGTGAGTTGATCAT CG
EGFR	EGFR- BX10a	GTTCAGAGTTCTACAGTCC GACGATCTTGAGTCTGTA GGACTTGGC	EGFR- BX10b	CCTTGGCACCCGAGAATTCC AACTTCTACCGTGCCCTGAT G
EGFR	EGFR- BX11a	GTTCAGAGTTCTACAGTCC GACGATCCTGCTGTGGGA TGAGGTACTC	EGFR- BX11b	CCTTGGCACCCGAGAATTCC ACACAGCAGGGCTTCTTCAG
EGFR	EGFR- BX12a	GTTCAGAGTTCTACAGTCC GACGATCCATGGAATGCT TGTACCACATC	EGFR- BX12b	CCTTGGCACCCGAGAATTCC ACATGGGCAACTTCTCTGTT TC
BRAF	BRAF- BX1a	GTTCAGAGTTCTACAGTCC GACGATCCAGTTTGAACA GTTGTCTGGATC	BRAF- BX1b	CCTTGGCACCCGAGAATTCC AAAAGTATGGGACCCACT CC
PTEN	PTEN- BX1a	GTTCAGAGTTCTACAGTCC GACGATCTGTTTCTGCTAA CGATCTCTTG	PTEN- BX1b	CCTTGGCACCCGAGAATTCC AAGGAGATATCAAGAGGAT GGATTC
PTEN	PTEN- BX2a	GTTCAGAGTTCTACAGTCC GACGATCCAGGAAATCCC ATAGCAATAATG	PTEN- BX2b	CCTTGGCACCCGAGAATTCC ATCCTGCAGAAAGACTTGA AGG
PTEN	PTEN- BX3a	GTTCAGAGTTCTACAGTCC GACGATCGCTTTGAATCCA	PTEN- BX3b	CCTTGGCACCCGAGAATTCC AGGATTCAAAGCATAAAAA

[0188]

		AAAACCTTAAAC		CCATTAC
PTEN	PTEN-BX4a	GTTCAGAGTTCTACAGTCC GACGATCTACAGTACATTC ATACCTACCTCTGC	PTEN-BX4b	CCTTGGCACCCGAGAATTCC ATATGTTGTATAACTTAAAC CCGATAGAC
PTEN	PTEN-BX5a	GTTCAGAGTTCTACAGTCC GACGATCAAAGGATATTG TGCAACTGTGG	PTEN-BX5b	CCTTGGCACCCGAGAATTCC ATTGAAGACCATAACCCACC AC
PTEN	PTEN-BX6a	GTTCAGAGTTCTACAGTCC GACGATCCCATAGAAATC TAGGGCCTCTTG	PTEN-BX6b	CCTTGGCACCCGAGAATTCC AAAGTAAGGACCAGAGACA AAAAGG
PTEN	PTEN-BX7a	GTTCAGAGTTCTACAGTCC GACGATCCCAGATGATTCT TTAACAGGTAGC	PTEN-BX7b	CCTTGGCACCCGAGAATTCC AGGATTATAGACCAGTGGC ACTG
PTEN	PTEN-BX9a	GTTCAGAGTTCTACAGTCC GACGATCGAAGTGTCTTC CCGTCGTG	PTEN-BX9b	CCTTGGCACCCGAGAATTCC ACATGTACTTTGAGTTCCT CAGC
PTEN	PTEN-BX12a	GTTCAGAGTTCTACAGTCC GACGATCTCTGGTCTGGT ATGAAGAATG	PTEN-BX12b	CCTTGGCACCCGAGAATTCC ACAGGACCAGAGGAAACCT CAG
PTEN	PTEN-BX13a	GTTCAGAGTTCTACAGTCC GACGATCGCTCTATACTGC AAATGCTATCG	PTEN-BX13b	CCTTGGCACCCGAGAATTCC ACGTGCAGATAATGACAAG GAATATC
PTEN	PTEN-BX14a	GTTCAGAGTTCTACAGTCC GACGATCTTGGAGAAAAG TATCGGTTGG	PTEN-BX14b	CCTTGGCACCCGAGAATTCC AGGTCAGTTAAATTAACAT TTTGTGG
PTEN	PTEN-BX15a	GTTCAGAGTTCTACAGTCC GACGATCTGGTGTACAG AAGTTGAAGTGC	PTEN-BX15b	CCTTGGCACCCGAGAATTCC AGATGTTAGTGACAATGAA CCTGATC
KRAS	KRAS-BX1a	GTTCAGAGTTCTACAGTCC GACGATCAAGAGTGCCTT GACGATACAGC	KRAS-BX1b	CCTTGGCACCCGAGAATTCC ATCTTGCTACGCCACCAG
TP53	TP53-BX1a	GTTCAGAGTTCTACAGTCC GACGATCCCTGACTCAGA CTGACATTCTCC	TP53-BX1b	CCTTGGCACCCGAGAATTCC ACAGGCCCTTCTGTCTTGAA C
TP53	TP53-BX2a	GTTCAGAGTTCTACAGTCC GACGATCATGTTCCGAGA GCTGAATGAG	TP53-BX2b	CCTTGGCACCCGAGAATTCC AGAACATCTCGAAGCGCTC AC
TP53	TP53-BX3a	GTTCAGAGTTCTACAGTCC GACGATCTTAAAGGACCA GACCAGCTTTC	TP53-BX3b	CCTTGGCACCCGAGAATTCC ATTATGGTATAAGTTGGTGT TCTGAAG
TP53	TP53-BX4a	GTTCAGAGTTCTACAGTCC GACGATCCTTGGGACCTCT TATCAAGTGG	TP53-BX4b	CCTTGGCACCCGAGAATTCC AAGAGGTCCCAAGACTTAG TACCTG
TP53	TP53-BX5a	GTTCAGAGTTCTACAGTCC GACGATCAAGCAAGCAGG ACAAGAAGC	TP53-BX5b	CCTTGGCACCCGAGAATTCC AGCTTGCTTACCTCGCTTAG TG
TP53	TP53-BX6a	GTTCAGAGTTCTACAGTCC GACGATCGGGACGGAACA GCTTTGAG	TP53-BX6b	CCTTGGCACCCGAGAATTCC ATTCCGTCCCAGTAGATTAC CAC
TP53	TP53-BX7a	GTTCAGAGTTCTACAGTCC GACGATCCAACTACATGT GTAAACAGTCTCTGC	TP53-BX7b	CCTTGGCACCCGAGAATTCC ACATGTAGTTGTAGTGGATG GTGG
TP53	TP53-BX8a	GTTCAGAGTTCTACAGTCC GACGATCGTGGAGTATTT	TP53-BX8b	CCTTGGCACCCGAGAATTCC AATACTCCACACGCAAATTT

		GGATGACAGAAAC		CC
TP53	TP53-BX9a	GTTCAGAGTTCTACAGTCC GACGATCTGCTCAGATAG CGATGGTGAG	TP53-BX9b	CCTTGGCACCCGAGAATTCC ACTATCTGAGCAGCGCTCAT G
TP53	TP53-BX10a	GTTCAGAGTTCTACAGTCC GACGATCCTGTGCAGCTGT GGGTTGA	TP53-BX10b	CCTTGGCACCCGAGAATTCC AGCAGGTCTTGGCCAGTTG
TP53	TP53-BX12a	GTTCAGAGTTCTACAGTCC GACGATCAAGTCTGTGAC TTGCACGGTC	TP53-BX12b	CCTTGGCACCCGAGAATTCC ATGTCCCAGAATGCAAGAA GC
TP53	TP53-BX13a	GTTCAGAGTTCTACAGTCC GACGATCCCTGTCTCTTC TGTCCTTC	TP53-BX13b	CCTTGGCACCCGAGAATTCC AGATGACAGGGGCCAGGAG
TP53	TP53-BX14a	GTTCAGAGTTCTACAGTCC GACGATCAAGACCCAGGT CCAGATGAAG	TP53-BX14b	CCTTGGCACCCGAGAATTCC ATGGGTCTTCAGTGAACCAT TG
TP53	TP53-BX15a	GTTCAGAGTTCTACAGTCC GACGATCCTGTCTGTCT TTCAGACTTC	TP53-BX15b	CCTTGGCACCCGAGAATTCC AGAGCAGAAAGTCAGTCCC ATG
TP53	TP53-BX16a	GTTCAGAGTTCTACAGTCC GACGATCCTCTGAGTCAG GAAACATTTTCAG	TP53-BX16b	CCTTGGCACCCGAGAATTCC AGCTCGACGCTAGGATCTGA C

표 3

	결과
리드	1.5M
% 표적 염기 w/ 1x	97.8%
% 온 표적	63.4%
% 복제물	18.2%
평균 커버리지	74.5x
커버리지의 표준편차	0.21

[0191]

[0192] 실시예 3: 시퀀싱 라이브러리 구축을 위한 지놈 DNA의 단편화

[0193] 1 μ L의 지놈 DNA를 제조사의 프로토콜에 따라 NEBNext dsDNA 단편화효소 키트(New England Biolabs)를 이용하여 처리하였다. 배양 시간을 37℃에서 45분으로 연장시켰다. 5 μ L의 0.5M EDTA pH 8.0를 부가하여 단편화 반응을 정지시키고, 제조사의 프로토콜에 따라 2X 부피의 Ampure XP 비드(Beckman Coulter, A63881)를 부가하여 정제하였다. 단편화된 DNA를 고민감도 DNA 키트(Agilent)로 Bioanalyzer 상에서 분석하였다. 단편화된 DNA의 크기 범위는 약 150bp의 피크를 갖는 전형적으로 약 100bp 내지 약 200bp였다.

[0194] 실시예 4: 라이브러리 제조 과정

[0195] 본 실시예에서는, 예시 목적을 위해 KAPA 라이브러리 프렙 키트(KK8230)를 사용하였다.

[0196] 비드 정제를 포함하는 단계를 위해, AMPure XP 비드 (cat #A63881)를 실온으로 평형화하고 철저히 재현탁시킨 후 샘플과 혼합하였다. 볼텍스 믹서 상에서 샘플과 철저히 혼합한 후, 실온에서 15분 동안 인큐베이션하여 DNA를 비드에 결합시켰다. 그리고 나서, 상기 액체가 투명해질 때까지 비드를 자기 스탠드 위에 두었다. 그리고 나서, 비드를 200ul 80% 에탄올로 2회 세척하고 실온에서 15분 동안 건조시켰다.

[0197] 말단-복구 반응을 수행하기 위해, 최대 50 μ L (2-10ng)의 무세포 DNA를 20 μ L의 말단 복구 마스터 믹스 (8 μ L 물, 7 μ L 10X KAPA 말단 복구 버퍼, 및 5 μ L KAPA 말단 복구 효소 믹스)와 혼합하고, 20℃에서 30분간 인큐베이션하였다. 그리고 나서, 120 μ L의 AMPure XP 비드를 상기 70 μ L 말단 복구 반응에 부가하였다. 그리고 나서, 샘플을 상기와 같이 정제하였다.

[0198] A-테일링 반응을 수행하기 위해, 상기 말단-복구된 DNA 단편을 함유하는 건조된 비드를 A-테일링 마스터 믹스 (42 μ L 물, 5 μ L 10X KAPA A 테일링 버퍼, 및 KAPA A-테일링 효소)와 혼합하였다. 상기 반응을 30℃에서 30분간 인큐베이션하였다. 90 μ L의 PEG 용액(20% PEG 8000, 2.5M NaCl)을 부가한 후, 상기 혼합물을 상기 비드 정제 프로토콜에 따라 세척하였다. 평할 말단 라이게이션 반응의 경우 이 A-테일링 단계를 생략하였다.

[0199] 링커 라이게이션을 위해, 하기 서열 (5'에서 3')을 갖는 2개의 올리고를 사용하여 어댑터 폴리뉴클레오타이드 이중체를 형성하였다: /5Phos/CCATTTCATTACCTCTTCTCCGCACCCGACATAGAT*T 및 /5Phos/ATCTATGTCGGGTGCGGAGAAAGAGGTAATGAAATGG*T. 말단-복구 (평할 라이게이션의 경우) 또는 a-테일 (링커-기반의 라이게이션의 경우)을 함유하는 건조된 비드를 45 μ L의 라이게이션 마스터 믹스 (30 μ L 물, 10 μ L 5x KAPA 라이게이션 버퍼, 및 5 μ L KAPA T4 DNA 리가아제), 및 5 μ L 물 (평할 말단 라이게이션의 경우) 또는 5 μ L의 링커 올리고뉴클레오타이드의 동일한 물의 믹스 (링커-기반의 라이게이션의 경우)와 혼합하였다. 상기 비드를 철저히 재현탁시키고, 20℃에서 15분간 인큐베이션하였다. 50 μ L의 PEG 용액(상기 참조)을 부가한 후, 상기 혼합물을 상기 비드 정제 프로토콜에 따라 세척하였다.

[0200] Illustra Genomiphi V2 DNA 증폭 키트를 이용하여 다중 치환 증폭(MDA)을 수행하였다. 상기 라이게이션된 단편 사슬을 함유하는 상기 건조된 비드를 9 μ L의 무작위 헥사머-함유 버퍼에 재현탁시키고, 95℃에서 3분으로 가열한 후, 얼음 상에서 급속 냉각하였다. 1 μ L의 효소 믹스를 부가한 후, 상기 냉각된 샘플을 30℃에서 90분간 인

큐베이션하였다. 그리고 나서, 65℃에서 10분간 가열하여 상기 반응을 정지시켰다. 30 μ L의 PEG 용액(상기 참조)을 부가한 후, 상기 혼합물을 상기 기재된 정제 프로토콜에 따라 세척하고, 200 μ L TE에 재현탁시켰다 (65℃에서 5분간 인큐베이션)에 재현탁시켰다. 원하는 경우, 상기 정제된 산물은 정량적 PCR, 디지털 드롭렛 PCR(ddPCR)로 정량되거나, 또는 차세대 시퀀싱(NGS)될 수 있다.

[0201] MDA 후, 긴 라이게이션된 단편 사슬(예컨대, >2kb)을 130 μ L 총 부피로 Covaris S220를 이용하여 ~300bp로 초음파처리하였다. 제조사의 프로토콜은 140W 피크 파워, 10% 충격 계수(duty factor), 버스트(burst)당 200 사이클, 및 80초의 처리 시간을 지시하였다. 온전한 원래의 무세포 DNA 단편을 유지할 가능성을 증가시키기 위해 ~300bp의 단편 길이를 선택하였다. 원하는 경우 시퀀싱을 위해 어댑터를 초음파처리된 DNA 단편에 붙이기 위해 표준 라이브러리 제조 프로토콜이 사용될 수 있다. 다양한 리드 조성이 일루미나 시퀀서(HiSeq 또는 MiSeq 중 하나)에서의 쌍-말단 시퀀싱으로부터 나왔다. 접합부(자가-접합부, 또는 어댑터가 라이게이션 단계에서 포함된 경우 어댑터 접합부 중 어느 하나)가 리드의 내부에 있는 것 (비-어댑터 서열이 5' 및 3' 측면에 있음)을 관심 있는 서열을 바코드하기 위해 사용하였다.

[0202] 실시예 5: 원형화 및 증폭

[0203] 이는 원형화 및 증폭 과정("네볼라" 과정으로도 지칭됨)의 예시를 제공한다. 상기 과정은 하기 공급품을 이용하였다: PCR 기기 (예컨대, MJ research PTC-200 Peltier 열 순환기); CircLigase II, ssDNA 리가아제 Epicentre cat# CL9025K; 엑소뉴클레아제 (예컨대, ExoI, NEB Biolabs cat #M0293S; ExoIII, NEB Biolabs cat# M0206S); T4 폴리뉴클레오타이드 키나아제 (NEB Biolab cat #M0201S); 전체 지놈 증폭 키트 (예컨대, GE Healthcare, Illustra, ready-To-Go, Genomiphi, V3 DNA 증폭 키트); GlycoBlue (예컨대, Ambion cat# AM9515); 마이크로 원심분리 (예컨대, Eppendorf 5415D); DNA 정제 비드 (예컨대, Agencourt, AMPure XP, Beckman Coulter cat# A63881); 자기 스탠드 (예컨대, The MagnaRack™ Invitrogen cat# CS15000); Qubit® 2.0 형광계 (Invitrogen, cat#Q32866); 분자 프로브 ds DNA HS 분석 키트 (Life Technology cat #032854); 및 Bioanalyzer (Agilent 2100), 및 고민감도 DNA 시약 (cat #5067-4626).

[0204] 5' 말단 인산염이 결실된 DNA 단편 (예컨대, 무세포 DNA)의 증폭을 위해, 제1 단계는 말단-복구 및 단일 가닥의 형성이었다. DNA를 96℃에서 30초간 변성시켰다 (예컨대, PCR 기기 상에서). 40 μ L의 DNA 및 5 μ L 10X PNK 반응 버퍼를 조합하여 폴리뉴클레오타이드 키나아제 (PNK) 반응을 준비한 후, 37℃에서 30분간 인큐베이션하였다. 1mM ATP 및 PNK 효소를 상기 반응에 부가하고, 37℃에서 45분간 인큐베이션하였다. 상기 DNA를 침전시키고 재현탁시켜 버퍼 교환을 수행하였다. 상기 PNK 반응으로부터의 50 μ L DNA, 5 μ L 아세트산나트륨 0.5M pH 5.2, 1 μ L GlycoBlue, 1 μ L 올리고 (100ng/ μ L), 및 150 μ L 100% 에탄올을 조합하였다. 상기 혼합물을 -80℃에서 30분간 인큐베이션하고, 16k rpm에서 5분간 원심분리하여 DNA를 펠렛화하였다. 상기 DNA 펠렛을 500 μ L의 70% 에탄올로 세척하고, 실온에서 5분간 공기 건조시키고, DNA를 12 μ L 10mM Tris-Cl pH 8.0에 현탁시켰다.

[0205] 그리고 나서, 재현탁된 DNA를 라이게이션에 의해 원형화시켰다. 상기 DNA를 96℃에서 30초간 변성시키고, 상기 샘플을 얼음 위에서 2분간 냉각하고, 리가아제 믹스 (2 μ L 10X CircLigase 버퍼, 4 μ L 5M 베타인, 1 μ L 50mM MnCl₂, 1 μ L CircLigase II)를 부가하였다. 상기 라이게이션 반응을 PCR 기기 상에서 60℃에서 16시간 동안 인큐베이션하였다. 라이게이션되지 않은 폴리뉴클레오타이드를 엑소뉴클레아제 절단에 의해 분해하였다. 이를 위해, DNA를 80℃에서 45초 동안 변성시키고, 1 μ L 엑소 뉴클레아제 믹스 (엑소I 20U/ μ L : 엑소III 100U/ μ L = 1:2)를 각 튜브에 부가하였다. 5시간 동안 피펫팅하여 이를 혼합하고 가볍게 회전시켰다. 상기 절단 믹스를 37℃에서 45분간 인큐베이션하였다. 30 μ L의 물을 이용하여 부피를 50 μ L로 조정하고, 상기 기재된 바와 같이 침전 및 재현탁에 의해 추가적인 버퍼 교환을 수행하였다.

[0206] 전체 지놈 증폭(WGA)을 수행하기 위해, 정제된 DNA를 먼저 65℃에서 5분간 변성시켰다. GE WGA 키트로부터의 10 μ L의 변성 버퍼를 10 μ L의 정제된 DNA에 부가하였다. 상기 DNA를 얼음 블록 또는 얼음 위에서 2분간 냉각시켰다. 20 μ L의 DNA를 Ready-To-Go GenomPhi V3 cake (WGA)에 부가하였다. 상기 WGA 반응을 30℃에서 1.5시간 동안 인큐베이션한 후, 65℃에서 10분 동안 열 불활성화하였다.

[0207] 상기 샘플을 AmpureXP 자기 비드(1.6X)를 이용하여 정제하였다. 상기 비드를 불택성하고, 80 μ L를 1.5mL 튜브에 분주하였다. 그리고 나서, 30 μ L 물, 20 μ L 증폭된 DNA, 및 80 μ L의 비드를 조합하고, 실온에서 3분간 인큐베이션하였다. 상기 튜브를 2분간 자기 스탠드에 둔 다음, 투명한 용액을 피펫으로 제거하였다. 상기 비드를 80% 에탄올로 2회 세척하였다 200 μ L의 10mM Tris-Cl pH 8.0을 부가하여 DNA를 용리시켰다. 상기 DNA 비드 혼합물을 65℃에서 5분간 인큐베이션하였다. 상기 튜브를 2분간 자기 스탠드에 다시 두었다. 195 μ L의 DNA를 새로운 튜브로 옮겼다. 1 μ L를 Qubit에 의한 정량을 위해 사용하였다. 마지막으로, 약 400 bp의 크기에 도달하

록 130 μ L WGA 산물을 Covaris S220을 이용하여 초음파처리하였다.

[0208] 실시예 6: 부가적인 정제를 수반한 원형화 및 증폭

[0209] 이는 페놀 클로로포름 추출 단계를 수반한 원형화 및 증폭 과정 ("네블라" 과정으로도 지칭됨)의 예시를 제공한다.

[0210] 단계 1은 circligase 반응을 위한 경쟁적 RNA (추출에서 캐리어 RNA로부터) 및 천연 RNA (공-정제된)의 제거였다. 1 μ L의 RNase A (10mg/ml) (Qiagen 1007885)를 50 μ L의 cfDNA(2-10ng)에 부가하고 PCR 기기 (MJ research PTC-200 Peltier 열 순환기) 상에서 37°C에서 30분간 인큐베이션하여 RNA를 제거하였다.

[0211] 단계 2는 염 및 에탄올 침전을 이용한 버퍼 교환이었다. 이 단계는 거의 100% 회수율 (반면 컬럼은 일반적으로 30%만 회수함)로 라이게이션을 위한 입력물을 세척하고 농축하는데 유용하였다. 에탄올 공침전 믹스 (RNase 처리로부터의 50 μ L의 DNA, 5 μ L의 아세트산나트륨 0.5M pH5.2, 1 μ L의 GlycoBlue (Ambion AM9515), 1 μ L 캐리어 올리고 (100ng/ μ L) 150 μ L의 100% 에탄올)를 -80°C에서 30분간 인큐베이션하고, 16k rpm (Eppendorf 5415D)에서 5분간 원심분리하여 DNA를 침전시켰다. 20-머 비특이적 캐리어 올리고 (PCR 프라이머로 사용됨)의 사용은 수율 및 침전 회수율의 안정성을 약간 증가시켰다. 상기 DNA 펠렛을 500 μ L의 70% 에탄올로 세척하였다. 상기 DNA 펠렛을 실온에서 5분간 공기 건조시키고, 13 μ L 10mM Tris-Cl pH 8.0으로 재현탁시켰다.

[0212] 단계 3은 원형화였다. 12 μ L의 cfDNA를 96°C에서 30초간 변성시키고, 얼음 블록 위에서 2분간 냉각시켰다. 라이게이션 믹스(12 μ L cfDNA, 2 μ L의 10X CircLigase 버퍼, 4 μ L 5M 베타인, 1 μ L의 50 mM MnCl₂, 1 μ L의 CircLigase II (Epicentre # CL9025K))를 차가운 얼음 상에서 부가하고, 라이게이션을 60°C에서 16시간 동안 수행하였다.

[0213] 단계 4는 엑소뉴클레아제 절단이었다. 라이게이션 DNA 혼합물을 PCR 기기 상에서 80°C에서 45초 동안 인큐베이션한 후, 엑소뉴클레아제를 처리하였다. 1 μ L 엑소 뉴클레아제 믹스 (엑소I 20U/ μ L : 엑소III 100U/ μ L = 1:2)를 각 튜브에 부가하고, 반응을 37°C에서 30분간 인큐베이션하였다. 품질 관리 목적을 위해, 선형 주형의 제거는 필요하지 않았다.

[0214] 단계 5는 페놀 클로로포름 추출 및 염 및 에탄올 침전을 이용한 버퍼 교환이었다. 페놀/에탄올은 80%를 초과하는 라이게이션 효율을 달성하는데 유용하였다 (원형화된 산물의 양은 입력 폴리뉴클레오타이드의 양과 거의 동일하였음). 180 μ L의 10mM 트리스를 엑소뉴클레아제 처리로부터의 20 μ L의 DNA에 부가하여 200 μ L의 부피를 만들고, 200 μ L의 페놀을 이용하여 DNA를 추출하였다. 수층을 수집하고, 에탄올 침전에 의해 DNA를 회수하였다. 상기 에탄올 공침전 믹스(페놀 추출 후 200 μ L의 DNA 용액, 20 μ L 아세트산나트륨 0.5M pH 5.2, 1 μ L GlycoBlue, 1 μ L 캐리어 올리고 (100ng/ μ L), 600 μ L의 100% 에탄올)를 -80°C에서 30분간 인큐베이션하고 16k rpm에서 5분간 원심분리하여 DNA를 침전시켰다. 상기 DNA 펠렛을 500 μ L의 70% 에탄올로 세척하였다. 상기 DNA 펠렛을 실온에서 5분간 공기 건조시키고, 11 μ L 10mM Tris-Cl pH 8.0으로 재현탁시켰다.

[0215] 단계 6은 전체 지놈 증폭이었다. 10 μ L의 정제된 DNA를 가열 블록 위에서 65°C에서 5분간 인큐베이션하고, GE Healthcare로부터의 10 μ L의 변성 버퍼, Ready-To-Go, Genomiphi, V3 DNA 증폭 키트를 부가하였다. DNA를 실온에서 5분간 냉각한 후, 20 μ L의 DNA를 Ready-To-Go Genomiphi V3 cake에 두었다 (WGA). 상기 증폭 반응을 30°C에서 1.5시간 동안 인큐베이션하고, 상기 반응을 65°C에서 10분간 열 불활성화에 의해 종결시켰다.

[0216] 단계 7은 AmpureXP 자기 비드 (1.6X)를 이용한 비드 정제였다. 이는 이전 실시예와 마찬가지로 수행하였다.

[0217] 단계 8은 이전 실시예와 마찬가지로 초음파처리였다. 이후, 상기 DNA를 정량적 PCR, ddPCR, 또는 시퀀싱 라이브러리 구축을 위해 준비하였다.

[0218] 실시예 7: 라이게이션 효율 및 온-표적(on-target) 비율의 분석

[0219] 상기 실시예에서와 같이 원형화되고 전체 지놈 적용을 거친 cfDNA를 정량적 PCR(qPCR)에 의해 분석하였다. 샘플 표적에 대한 qPCR 증폭 곡선 결과 (KRAS 프라이머를 이용)가 도 18에 나타나 있다. 도 18A에 나타난 바와 같이, 입력 cfDNA의 1/10의 qPCR 증폭은 31.75의 평균 Ct (사이클 역치)를 나타내었고, 동일한 샘플의 라이게이션 산물의 1/10은 31.927의 평균 Ct를 나타내었는데, 이는 약 88%의 높은 라이게이션 효율을 가리킨다. 라이게이션 효율은 약 70%, 80%, 90%, 95% 이상, 또는 이를 초과하는 범위, 예컨대 약 100%였다. 원형화되지 않은 선형 DNA는 일부 실시예에서 제거되는데, 이로써 거의 모든 DNA가 원형 형태로부터 증폭될 수 있다. 각각의 샘플을 2반복으로 3회 수행하였다. 도 18B에 나타난 바와 같이, 10ng의 WGA 산물 및 참조 지놈 DNA (gDNA) (12878, 10n

g)의 증폭 곡선은 사실상 서로 중첩된다. WGA 샘플에 대한 평균 Ct는 26.655인 반면, gDNA 샘플의 Ct는 26.605이었는데, 이는 96%가 넘는 높은 온-표적 비율을 가리킨다. 주어진 양의 증폭된 DNA에서의 KRAS의 수는 증폭되지 않은 gDNA와 대등하였고, 이는 편향되지 않은 증폭 공정을 가리킨다. 각각의 샘플을 2반복으로 3회 시험하였다. 대조점으로서, Lou et al. (PNAS, 2013, 110 (49))에서 제공된 원형화 프로토콜을 또한 시험하였다. 상기 기재된 실시예 중 침전 및 정제 단계가 없는 Lou 방법을 이용하여, 단지 10-30%의 선형 입력 DNA만이 원형 DNA로 전환되었다. 이러한 낮은 회수율은 하류 시퀀싱 및 변이 검출에 대한 문제를 야기한다.

[0220] **실시예 8: ddPCR에 의한 증폭된 원형화된 DNA의 분석**

[0221] 드롭렛 디지털 PCR(ddPCR)을 사용하여 원형화된 폴리뉴클레오타이드로부터 생성된 전체 지놈 증폭 산물에서 대립유전자 빈도 보존 및 편향을 평가하였다. 일반적으로, ddPCR은 PCR 증폭을 지원하는 별개의, 부피가 정해진, 수중유 드롭렛 파티션에 캡슐화된 핵산 분자를 계수함으로써 절대량을 측정하는 디지털 PCR 분석을 가리킨다 (Hinson et al, 2011, Anal. Chem. 83:8604-8610; Pinheiro et al, 2012, Anal. Chem. 84: 1003-1011). 단일 ddPCR 반응은 웰당 적어도 20,000개의 파티션된 드롭렛으로 구성될 수 있다. 드롭렛 디지털 PCR은 PCR 증폭을 지원하는 별개의, 부피가 정해진, 수중유 드롭렛 파티션에 캡슐화된 핵산 분자를 계수함으로써 절대량을 측정하는 디지털 PCR 분석을 수행하는 임의의 플랫폼을 이용하여 수행될 수 있다. 드롭렛 디지털 PCR을 위한 전형적인 전략은 하기와 같이 요약될 수 있다: 샘플을 희석하고, 각각 관심있는 핵산 분자의 하나의 카피를 함유하거나 함유하지 않도록 이를 수천 내지 수백만개의 별개의 반응 챔버(수중유 드롭렛)로 분할한다. 표적 앵플리콘 (즉, 관심있는 핵산 분자)을 함유하는, 검출된 "양성" 드롭렛의 수 대 표적 앵플리콘 (관심있는 핵산 분자)을 함유하지 않는 "음성" 드롭렛의 수가 원래의 샘플에 존재하였던 관심있는 핵산 분자의 카피 수를 결정하는데 사용될 수 있다. 드롭렛 디지털 PCR 시스템의 예는 핵산 주형을 함유하는 샘플을 20,000 나노리터 크기의 드롭렛으로 분할하는 Bio-Rad사의 QX100™ 드롭렛 디지털 PCR 시스템; 및 핵산 주형을 함유하는 샘플을 1,000,000 내지 10,000,000 피코리터 크기의 드롭렛으로 분할하는 RainDance사의 RainDrop™ 디지털 PCR 시스템을 포함한다. ddPCR을 위한 방법의 부가적인 예가 제WO2013181276A1호에 제공되어 있다.

[0222] 본 실시예에서, 흑색종 세포주로부터의 BRAF V600E 지놈 DNA(gDNA)를 특정 비율 (0%, 0.67%, 2.0%, 6.67%, 20%, 또는 100%)로 참조 지놈 DNA 12878과 혼합하고, 단편화하여 cfDNA에서 발견되는 것과 유사한 크기의 단편 (이 경우, 약 150bp)을 생성하였다. 상기 혼합된 DNA 샘플 (10ng)을 실시예 2에 따라 원형화하고 증폭시켰다. 40ng의 증폭된 DNA를 BRAF V600E 및 야생형에 대해 ddPCR을 수행하였다. 관찰된 돌연변이 대립유전자 빈도가 도 19에 그래프로 예시되어 있고 표로 작성되어 있다. 나타난 바와 같이, 증폭 없이 100ng의 지놈 DNA로부터 수득한 ddPCR 결과 (하부 열) 뿐만 아니라, 증폭을 수반한 상기 관찰된 돌연변이 대립유전자 빈도 (도 19 표의 중간 열)는 입력 돌연변이 대립유전자 빈도 (상부 열)를 반영한다. ddPCR 출력에 의한 대립유전자 빈도는 돌연변이 및 야생형을 함유하는 드롭렛 모두의 합 대비 BRAF 돌연변이를 함유하는 드롭렛의 수로서 계산된다. 증폭된 DNA는 개방형 원으로 표시되어 있으며, 증폭이 없는 DNA는 작은 채워진 원으로 표시되어 있다. 0.67%의 작은 편차를 제외하면, 2개의 데이터 세트는 완전히 중복되었다. 이것은 실질적으로 편향없이, 돌연변이 대립유전자 빈도의 진정한 표현의 보존을 입증한다.

[0223] **실시예 9: 배경을 초과하는 서열 변이의 검출**

[0224] 10ng의 초음파처리된 gDNA (150bp, Multi-Gene Multiplex reference DNA, Horizon)를 실시예 2에 기재된 바와 같이 원형화시키고 증폭시킨 후, 초음파처리하였다. 그리고 나서, 단편화된 DNA를 루비콘 시퀀싱 라이브러리 구축을 수행하였다. 포획 시퀀싱 후, 참조 핫스팟으로부터 50bp이내의 변이를 플롯팅하였다. 변이 판정이 상이한 접합부에 의해 구별된 2개의 상이한 폴리뉴클레오타이드에서의 검출을 요구한 변이 검출에 대한 결과가 도 20에 나타나 있다. 7개의 예측된 참조 핫스팟 (KIT D816V, EGFR G719S, EGFR T790M, EGFR L858R, KRAS G13D, KRAS G12D, NRAS Q61K)이 위치 0에 플롯팅되어 있다. 2개의 다른 변이들 역시 확인되었고, 도 20에 개방된 삼각형 및 다이아몬드로 예시되어 있다.

[0225] 비교를 위해, gDNA를 상기와 같이 초음파처리하였으나, 10ng의 상기 초음파처리된 gDNA는 원형화 없이 그리고 2개의 상이한 폴리뉴클레오타이드 상에서 서열 변이의 확인을 요구하지 않으면서, 일반적인 관행에 따라 곧바로 루비콘 시퀀싱 라이브러리 구축을 수행하였다. 포획 시퀀싱 후, 참조 핫스팟으로부터 50bp 이내의 변이를 다시 플롯팅한 후, 결과를 도 21에 나타내었다. 7개의 예측된 참조 핫스팟 (KIT D816V, EGFR G719S, EGFR T790M, EGFR L858R, KRAS G13D, KRAS G12D, NRAS Q61K)이 위치 0에 플롯팅되어 있다. 다른 위치에 있는 변이들은 예상되지 않았고, 이는 대부분 시퀀싱 오류 때문일 가능성이 있다. 도 20을 생성하는데 사용된 방법의 결과와는 대조적으로, 도 21의 결과는 표준 시퀀싱 방법이 대립유전자 빈도가 낮을 때 (예컨대 5% 미만) 진정한 돌연변이

신호를 감출 수 있는 더 높은 무작위 오류율을 갖는다는 것을 가리킨다.

[0226] 2개의 상이한 폴리뉴클레오타이드에서의 검출을 수반하는 시퀀싱 방법 및 이를 필요로 하지 않는 시퀀싱 방법에 의해 검출된 민감도 및 배경 노이즈의 개별적인 분석 결과가 도 16-17에 예시되어 있다. 이들 도면이 예시하는 바와 같이, 검증 요구는 배경 노이즈를 크게 감소시키고 민감도를 증가시킨다.

[0227] **실시예 10: GC 조성 및 크기 분포의 분석**

[0228] 10ng의 초음파처리된 gDNA(150bp, Multi-Gene Multiplex reference DNA, Horizon)를 실시예 5에서와 같이 원형화시키고 증폭시키고, 시퀀싱하고, 변이-판정 2-폴리뉴클레오타이드 확인 필터(좌측)를 이용하여 분석하였다. 일정한 범위의 GC 백분율을 갖는 서열의 수를 표로 작성하고 도 22에 나타난 바와 같이 그래프로 플롯팅하였다. 가장 좌측 플롯에 나타난 바와 같이, 실시예 5에 따라 제조된 샘플에 대한 서열들은 중앙 피크(근본적인 지놈의 전체 GC 함량에 상응함)를 제외하고 이론적인 분포와 대부분 유사하다. 대조적으로, 동일한 양의 gDNA가 루비콘 시퀀싱 라이브러리 구축 키트를 이용하여 증폭 없이 시퀀싱 라이브러리를 구축하기 위해 곧바로 사용되었을 때, 시퀀싱 결과 및 이론적 분포 간의 차이는 매우 확연하다 (중간 플롯 참고). 이러한 직접적인 루비콘 시퀀싱의 중앙 피크는 이론적 분포보다 더 높다. Newman *et al.* (2014; Nature Medicine, (20):548-54)은 32ng의 cfDNA가 사용되었을 때 cfDNA 시퀀싱 GC 함량 분포가 이론적 분포와 유사하였음을 보고하였다. 이것이 가장 우측 플롯에 예시되어 있다.

[0229] 실시예 5에서와 같이 원형화되고, 증폭되고, 시퀀싱된 cfDNA에 대해 DNA 크기 분포를 평가하였다. 도 23에 나타난 바와 같이, 시퀀싱 결과에 의해 표시된 단편 길이의 분포의 피크는 약 150-180bp이며, 이는 cfDNA의 전형적인 분포 패턴과 유사하다.

[0230] **실시예 11: 증폭 균일성의 평가**

[0231] 실시예 5에 따라 원형화되고 증폭된 10개의 산물의 qPCR 결과를 증폭되지 않은 참조 DNA (12878 세포주로부터의 gDNA, Coriell Institute)와 비교하였다. 10ng의 지놈 참조 DNA 또는 증폭 산물을 각 실시간 qPCR 반응을 위해 사용하였고, 지놈 참조 대비 증폭 산물의 상대적 정량에 의해 비율을 생성하였다. 도 24에 나타난 바와 같이, 각각의 PCR의 비율은 2배 변화 이내이며, 이는 증폭된 DNA 집단 내의 이들 표적들의 카피수가 증폭되지 않은 참조 DNA와 매우 유사하다는 것을 제시한다. 6개 유전자(BRAF, cKIT, EGFR, KRAS, NRAS, PI3KCA)로부터 10쌍의 PCR 프라이머를 설계하고 미리 검증하였다.

[0232] **실시예 12: DNA 단편의 증폭 수율의 정량**

[0233] 4명의 환자 (환자1-4) 및 1명의 건강한 대조군으로부터 cfDNA를 단리하였다. 지놈 DNA (gDNA, Multi-Gene Multiplex Horizon)를 약 150bp 단편으로 초음파처리하였다. DNA를 원형화시키고 무작위 프라이머를 이용하여 증폭시켰다. 표 7은 증폭 반응 내로 입력된 DNA의 양, 및 증폭에 의해 생성된 DNA의 양을 보여준다. 가장 작은 샘플 (0.4ng)에 대해서도 유의미한 증폭이 얻어졌으며, 모든 샘플들은 적어도 600배 증폭되었다.

표 7

샘플 유형	입력 (ng)	수율 (ng)
gDNA	10	6100
gDNA	4	6880
gDNA	2	5700
gDNA	1	5760
gDNA	0.5	3280
cfDNA 건강한 대조군	0.4	6240
cfDNA 환자 1	4	13800
cfDNA 환자 2	3	9480
cfDNA 환자 3	3	7840
cfDNA 환자 4	1	3180

[0234]

[0235] **실시예 13: 암 환자의 cfDNA로부터 저-빈도 돌연변이의 검출**

[0236] 단계 1에서, cfDNA를 원형화시켰다. 원형 라이게이션 믹스를 실온에서 PCR 튜브에서 제조하였다. 4 ng-10ng cfDNA를 12 μ l의 부피로 PCR 튜브에 피펫팅하였다. DNA를 96°C에서 30초간 변성시킨 다음, PCR 튜브를 얼음 위에서 2분간 냉각시켰다. 라이게이션 믹스 (2 μ l의 10X CircLigase 버퍼, 4 μ l 5 M 베타인, 1 μ l 50 mM

MnCl₂, 1 μl CircLigase II)를 각 튜브에 부가하고, 반응을 PCR 기기 상에서 60℃에서 16시간 동안 진행하였다.

- [0237] 단계 2에서, 라이게이션 반응을 처리하여 라이게이션되지 않은 선형 DNA를 제거하였다. 1 μl 엑소뉴클레아제 믹스 (NEB M0206S, M0293S; ExoI 20u/μl : ExpIII 100u/μl = 1:2)를 각 튜브에 부가하고, 혼합하고, PCR 기기에서 37℃에서 30분간 인큐베이션하였다.
- [0238] 단계 3에서, 버퍼 교환을 위해 상기 라이게이션 반응을 정제하였다. 상기 라이게이션 산물을 올리고 Clean & Concentrator(Zymo Research)를 이용하여 정제하였다. 엑소뉴클레아제 처리 후 결합 믹스(30 μl의 10mM Tris, 100 μl의 올리고 결합 버퍼, 400 μl의 100% 에탄올)를 상기 라이게이션 반응에 부가하고, 혼합하고, 가볍게 회전시켰다. 자이모-스핀 컬럼을 로딩하고, 10,000xg 초과에서 30초간 회전시켰다. 컬럼을 750 μl의 DNA 세척 버퍼로 세척하고, 14,000xg에서 1분간 원심분리하였다. DNA를 10,000xg 초과에서 30초간 원심분리에 의해 15 μl의 10mM Tris로 용리시켰다.
- [0239] 단계 4에서, DNA를 무작위 프라이밍에 의해 증폭시켰다. 전체 지놈 증폭 (WGA)을 Ready-To-Go Genomiphi V3 DNA 증폭 키트 (GE Healthcare)를 이용하여 수행하였다. 10 μl의 정제된 라이게이션을 10 μl의 2x 변성 버퍼와 혼합하고, 95℃에서 3분간 인큐베이션한 다음, 얼음 위에서 4℃로 냉각시켰다. 20 μl의 변성된 DNA를 WGA 프리-믹스에 부가하고, 샘플을 30℃에서 1.5시간 동안 인큐베이션 후 65℃에서 10분간 불활성화시켰다.
- [0240] 단계 5에서, 상기 증폭 산물을 Agencourt AMPure XP 정제 (1.6X) (Beckman Coulter)를 이용하여 세척하였다. 30 μl의 10mM Tris 및 80 μl의 Ampure 비드를 20 μl의 WGA 반응에 부가하였다. 상기 혼합물을 실온에서 2분간 인큐베이션하였다. 튜브를 자석 스탠드 위에 두고, 2분간 인큐베이션하였다. 상등액을 제거하여 버렸다. 샘플을 200 μl의 에탄올(80%)로 2회 세척하고, 5분간 공기 건조하고, DNA를 200 μl의 10mM Tris pH 8.0으로 용리시켰다.
- [0241] 단계 6에서, WGA DNA를 단편화시켰다. 130 μl WGA 산물을 Covaris S220 초음파처리기를 이용하여 초음파처리하여 약 400 bp의 단편 크기를 획득하였다. Covaris S220 설정은 하기와 같았다: 피크 입사력=140W, 충격 계수=10%, 버스트 당 사이클= 200, 처리 시간=55초.
- [0242] 단계 7에서, 샘플을 qPCR에 의해 정량하였다. 라이게이션 입력의 1/10 및 라이게이션 산물을 라이게이션 효율을 측정하기 위해 3반복으로 qPCR 반응을 위해 사용하였다. 10ng의 참조 gDNA (12878 세포주)와 함께 10ng의 단편화된 WGA 산물을 qPCR에 사용하여 온 표적 비율을 측정하였다. 반응은 5 μl의 2x 마스터 믹스 (TaqMan Fast 유니버설 PCR 마스터 믹스 (2x), Applied Biosystems; Evagreen dye, Biotium), 0.5 μl의 프라이머 (5 μM), 1.2 μl의 H₂O, 10 μl의 DNA로 구성되었다. 증폭을 하기 프로그램에 따라 진행하였다: 95℃에서 2분; 및 [95℃, 10초; 60℃, 20초]의 40 사이클.
- [0243] 단계 8에서, 시퀀싱 라이브러리를 구축하였다. KAPA Hyper Prep 키트 (KK8500) 또는 표준 PCR 라이브러리를 갖는 KAPA 라이브러리 제조 키트 (KK8200)를 이용하여 500-1000ng의 초음파처리된 증폭된 DNA로부터 시퀀싱 라이브러리를 제조하였다. 어댑터 라이게이션 (1uM 어댑터 최종 농도를 이용하여)을 제조사의 프로토콜에 따라 제조하였다. 상기 라이게이션된 산물의 어댑터 라이게이션된 세척, 30 μl (0.3x)의 20% PEG 8000/ 2.5M NaCl 용액을 100 μl의 재현탁된 라이게이션된 산물에 부가하였다. 비드를 상기 라이게이션된 산물과 철저히 혼합하고 실온에서 15분간 인큐베이션하였다. 그리고 나서, 상기 액체가 투명할 때까지 비드를 자석 위에서 포획하였다. 그리고 나서, 130 μl의 상등액을 Ampure XP 비드를 이용하여 크기 선별을 수행하였다. 샘플을 새로운 플레이트로 옮긴 후 20 μl의 Ampure XP 비드 (0.5x)를 부가하였다. 라이게이션된 산물이 이제 비드에 포획되었고 200 μl 80% 에탄올로 2회 세척하였다. 그리고 나서, 라이게이션된 산물을 재현탁시키고 20 μl EB 버퍼에 용리시켰다. 크기 선택 및 정제 후, 20 μl 라이게이션된 산물을 25 μl 2x KAPA HiFi Hotstart ready 믹스 및 5 μl 10 μM P5+P7 프라이머 (5'CAAGCAGAAGACGGCATACGA3', 5'ACACTCTTTCCCTACACGACGCTCTTCCGATCT3')에 부가하여 하기 순환 프로그램을 이용하여 상기 라이브러리를 증폭시켰다: 98℃, 45초; (98℃, 15초; 60℃, 30초; 72℃, 30초)의 5 사이클; 72℃, 60초. 증폭된 라이브러리를 20x로 희석한 후 정량을 위해 단편 분석기 또는 생물분석기 (고민 감도 칩) 상에 로딩하였다. 겔 크기 선별기(Sage Science로부터의 Blue Pippin prep)를 통해 추가적인 크기 선별을 수행하였다.
- [0244] 단계 9에서, 시퀀싱 라이브러리를 xGEN Pan-Caner Panel v1.5, 127908597 (IDT)로부터의 프로브를 이용한 프로브 포획 농축에 의해 농축시켰다. 단계 10에서, 상기 라이브러리를 1000x의 평균 깊이로, HiSeq 2500에서 시퀀싱하였다.

[0245] 단계 11에서, 시퀀싱 데이터를 분석하여 변이를 판정하였다. 변이 판정은 변이로서 계수될 2개의 상이한 폴리뉴클레오타이드(예컨대, 상이한 접합부에 의해 확인됨) 상에서 서열 차이가 발생하는 것을 필요로 하는 단계를 포함하였다. 몇 가지 체세포 돌연변이가 검출되었고, 이들은 또한 공공 데이터베이스(COSMIC (Catalog of Somatic Mutations in Cancer))에서 보고되었다. 확인된 돌연변이 중에는 0.05% 대립유전자 빈도를 갖는 BRAF V600M이 있었고, 이는 입력이 낮을 때에도 이 시스템의 높은 민감도를 입증한다. 샘플에서의 이들의 빈도를 포함하는, 다양한 돌연변이의 검출 결과가 표 8에 나타나 있다.

표 8

암 유형	돌연변이된 유전자	돌연변이	샘플에서의 빈도
흑색종	FGFR2	V191I	0.20%
흑색종	NRAS	Q61K	0.20%
흑색종	CTNNB1	S37F	0.10%
흑색종	BRAF	V600E	0.10%
흑색종	BRAF	V600M	0.05%
유방암	PIK3CA	H1407L	53%
유방암	LIFR	S679L	0.60%
유방암	KRAS	G12D	0.60%
유방암	ATM	R2993*	0.30%
유방암	ATRX	S618F	0.10%
전립선	AR	T346A	6.40%
전립선	SPOP	F133L	3.30%
전립선	BRAF	V600M	1.10%
전립선	NSD1	R1557C	1.00%
전립선	SF3B1	K700E	1.60%
췌장	KRAS	G12V	1.70%
췌장	TP53	7578176C>T (스플라이스 공여체 변이체)	0.5%

[0246]

[0247] 실시예 14: FFPE 샘플로부터의 멀티플렉스 참조 DNA로부터 정확한 돌연변이 검출

[0248] 제조사의 프로토콜(Covaris truXTRAC™ FFPE DNA 키트)에 따라 샘플 Horizon FFPE-multiplex (HD200)로부터 DNA를 추출하였다. 130 µl FFPE gDNA를 Covaris S220 초음파처리기를 이용하여 초음파처리하여 약 150 bp의 단편 크기를 획득하였다 (Covaris S220 설정: 피크 입사력=175W, 충격 계수=10%, 버스트 당 사이클= 200, 처리 시간=430초). 11 µl 부피의 50 ng DNA를 95℃에서 30초 동안 변성시켰다. 10.5 µl H₂O, 2.5 µl 리가아제 버퍼(NEB B0202S), 및 1 µl T4 폴리뉴클레오타이드 키나아제(NEB M0201S)를 추가하였다. 인산화를 위해 반응을 37℃에서 30분간 인큐베이션하였다.

[0249] 샘플을 라이게이션한 다음, 올리고 Clean & Concentrator(Zymo Research)를 이용하여 정제하였다. 엑소뉴클레아제 처리 후 결합 믹스(30 µl의 10mM Tris, 100 µl의 올리고 결합 버퍼, 400 µl의 100% 에탄올)를 상기 라이게이션 반응에 추가하고, 볼텍스에 의해 혼합하고, 가볍게 회전시켰다. 샘플을 자이모-스핀 컬럼에 로딩하고, 10,000xg 초과에서 30초간 회전시켰다. 컬럼을 750 µl의 DNA 세척 버퍼로 세척하고, 14,000xg에서 1분간 원심 분리하였다. DNA를 10,000xg 초과에서 30초간 원심분리하여 15 µl의 10mM Tris를 이용하여 용리시켰다.

[0250] 샘플을 추가 처리하고 실시예 13에서의 단계 5-11에 따라 분석하였다. 결과가 표 9에 요약되어 있다. Horizon의 다중 돌연변이 표준 DNA에서의 9개 돌연변이의 표시는 이 과정에 의해 대략적으로 유지된 반면 DNA의 양은 적어도 600배 증가하였다.

표 9

유전자	돌연변이	대립유전자 빈도	측정된 대립유전자 빈도
BRAF	p.V600E	10.5%	8%
KIT	p.D816V	10%	9.4%
EGFR	p.L858R	3%	4%
EGFR	p.T790M	1%	1%
EGFR	p.G719S	24.5%	14%
KRAS	p.G13D	15%	9.1%
KRAS	p.G12D	6%	5.4%
NRAS	p.Q61K	12.5%	8.3%
PIK3CA	p.H1047R	17.5%	14%

[0251]

[0252] 실시예 15: 암 돌연변이 세포주 gDNA 멀티머로부터 저-빈도 돌연변이의 검출

[0253] 이 실시예에서, 초음파처리된 지놈 DNA를 라이게이션시켜 멀티머를 형성시킨 다음, 이를 증폭, 단편화, 및 분석하였다. 도 25가 이 과정의 예를 예시한다.

[0254] BRAF V600E 돌연변이를 함유하는 흑색종 세포주 SK--mel-28 (ATCC)로부터의 gDNA를 참조 gDNA (12878 Coriell Institute)와 혼합하여 1% BRAF V600E를 달성하였다. DNA를 실시예 14에서와 같이 초음파처리하여 약 150bp의 단편 크기를 수득하였다. 11 μ l 부피의 100 ng의 DNA를 95°C에서 30초간 변성시켰다. 10.5 μ l H₂O, 2.5 μ l 리가아제 버퍼(NEB B0202S), 및 1 μ l T4 폴리뉴클레오타이드 키나아제 (NEB M0201S)를 부가한 후, 37°C에서 30분간 인큐베이션하여 DNA를 인산화시켰다.

[0255] 샘플을 올리고 Clean & Concentrator (Zymo Research)를 이용하여 정제하였다. 이것은 엑소뉴클레아제 처리 후 결합 믹스 (25 μ l의 10mM Tris, 100 μ l의 올리고 결합 버퍼, 400 μ l의 100% 에탄올)를 상기 라이게이션 반응에 추가하는 것을 포함하였다. 이를 볼텍스에 의해 혼합하고 가볍게 회전시켰다. 자이모-스핀 컬럼을 로딩하고, 10,000xg 초과에서 30초간 회전시키고, 750 μ l의 DNA 세척 버퍼로 세척하고, 14,000xg에서 1분간 원심분리하였다. DNA를 10,000xg 초과에서 30초 동안 원심분리하여 15 μ l의 10mM Tris로 용리시켰다.

[0256] 라이게이션시키기 위해, 4 μ l 부피의 6ng DNA를 0.45 μ l 10x 말단 복구 버퍼 (Enzymatics), 0.05 μ l dNTP 25mM, 0.5 μ l ATP 10mM, 말단 복구 효소 믹스 (Enzymatics), 및 T4 리가아제 2000 unit/ μ l과 혼합하였다. 상기 반응을 25°C에서 30분간 인큐베이션 후, 75°C에서 20분간 인큐베이션하였다.

[0257] Ready-To-Go Genomiphi V3 DNA 증폭 키트 (GE Healthcare)를 이용하여 전체 지놈 증폭을 수행하였다. 8 μ l의 H₂O 및 10 μ l의 정제된 라이게이션을 10 μ l의 2x 변성 버퍼와 혼합하였다. DNA를 95°C에서 3분간 변성시킨 다음, 얼음 위에서 4°C로 냉각하였다. 20 μ l의 변성된 DNA를 WGA 프리-믹스에 부가하고, 30°C에서 1.5시간 동안 인큐베이션 후 65°C에서 10분간 불활성화시켰다.

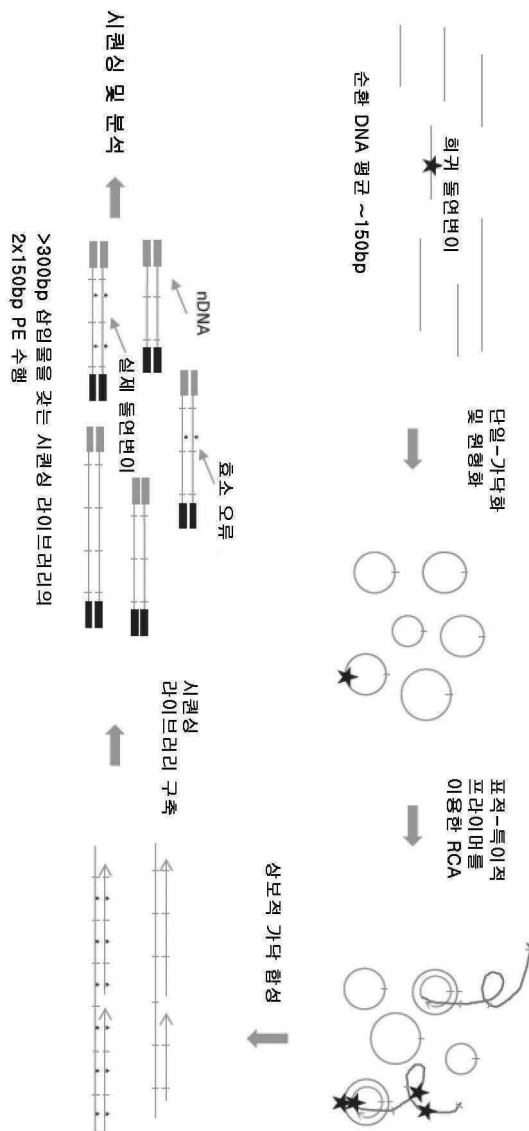
[0258] 그리고 나서, 상기 증폭 반응을 Agencourt AMPure XP 정제 (1.6X) (Beckman Coulter)를 이용하여 세척하였다. 30 μ l의 10mM Tris 및 80 μ l의 AMPure 비드를 20 μ l의 WGA 반응에 부가하였다. 이를 실온에서 2분간 인큐베이션하였다. 상기 튜브를 자석 스탠드 상에 두고, 2분간 인큐베이션하였다. 상등액을 제거하여 버렸다. 비드를 200 μ l의 에탄올 (80%)로 2회 세척한 다음, 5분간 공기건조하였다. DNA를 200 μ l의 10mM Tris pH 8.0으로 용리시켰다. 그리고 나서, 130 μ l WGA 산물을 Covaris S220 초음파처리기를 이용하여 단편화시켜 약 400 bp의 단편 크기를 수득하였다 (Covaris S220 설정: 피크 입사력=140W, 충격 계수=10%, 버스트 당 사이클= 200, 처리 시간=55초).

[0259] 돌연변이를 BioRad Prime PCR ddPCR 돌연변이 검출 분석을 이용하여 ddPCR에 의해 검출하였다. 돌연변이-검출 ddPCR 반응을 실온에서 PCR 튜브에서 모았다 (80ng의 증폭된 DNA, 프로브에 대한 10 μ l의 2x ddPCR 수퍼믹스, 1 μ l의 20x 표적 (BRAF V600E, BioRad) 프라이머 (9 μ M)/프로브 (FAM; 5 μ M), 1 μ l의 20x 야생형 프라이머 (9 μ M)/ 프로브 (HEX; 5 μ M), 8 μ l의 DNA 샘플(50 ng)). 상기 반응을 5시간 동안 피펫팅에 의해 혼합한 다음, 드롭렛 생성기 카트리지에 옮겼다. QX200 드롭렛 생성기를 이용하여 드롭렛을 생성시키고, 96-웰 PCR 플레이트로 옮기고, 하기 PCR 프로그램을 이용하여 증폭시켰다: 95℃, 10분; [94℃, 30초, 55℃ 1분]의 40 사이클; 98℃, 10분. PCR 반응 플레이트를 QX200 드롭렛 판독기로 옮겨 결과를 정량하였다. 입력 DNA에 기초하여, BRAF V600E 돌연변이의 예측된 빈도는 1%였다. 이 라이게이션 및 증폭 과정에 의해, 이 빈도는 개략적으로 유지된 반면 (ddPCR 분석에 따라 1.41%), DNA의 양은 약 200배 증가하였다.

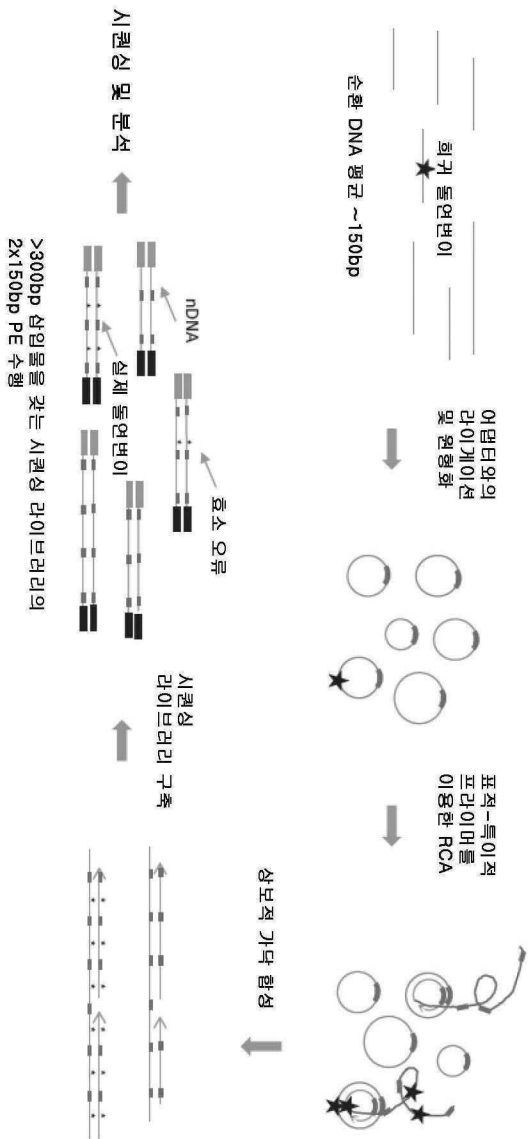
[0260] 본 발명의 바람직한 구현예가 본원에 나타나 있고 기재되어 있지만, 그러한 구현예는 단지 예로서 제공되는 것임이 당해 기술분야의 숙련자에게 자명할 것이다. 많은 변형, 변화, 및 치환이 본 발명을 벗어나지 않으면서 본 기술분야의 숙련자에게 일어날 것이다. 본원에 기재된 본 발명의 구현예에 대한 다양한 대안들이 본 발명을 실행하는데 이용될 수 있는 것으로 이해되어야 한다. 하기 청구범위는 본 발명의 범위를 한정하며, 이들 청구범위 및 이들의 등가물에 속하는 방법 및 구조가 이에 의해 보호되는 것으로 의도된다.

도면

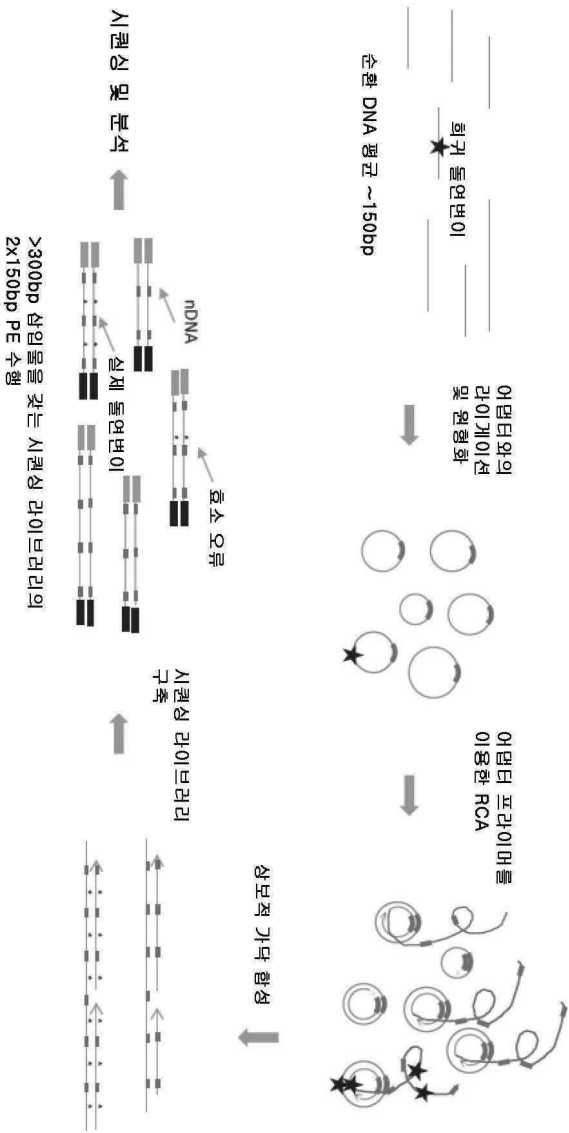
도면1



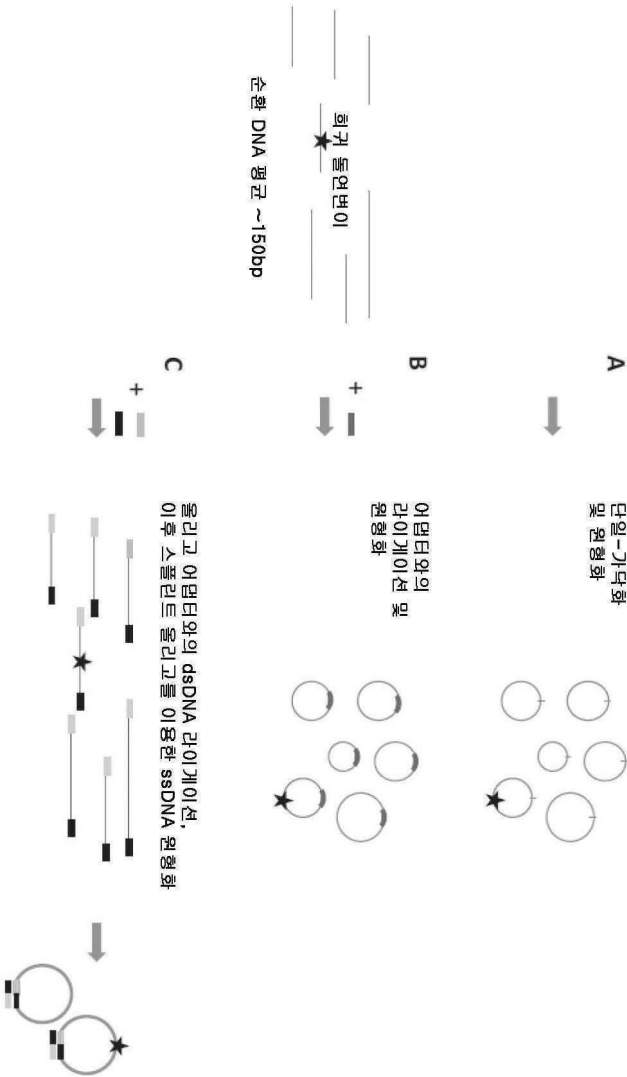
도면2



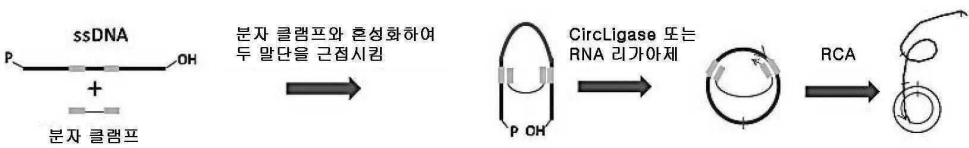
도면3



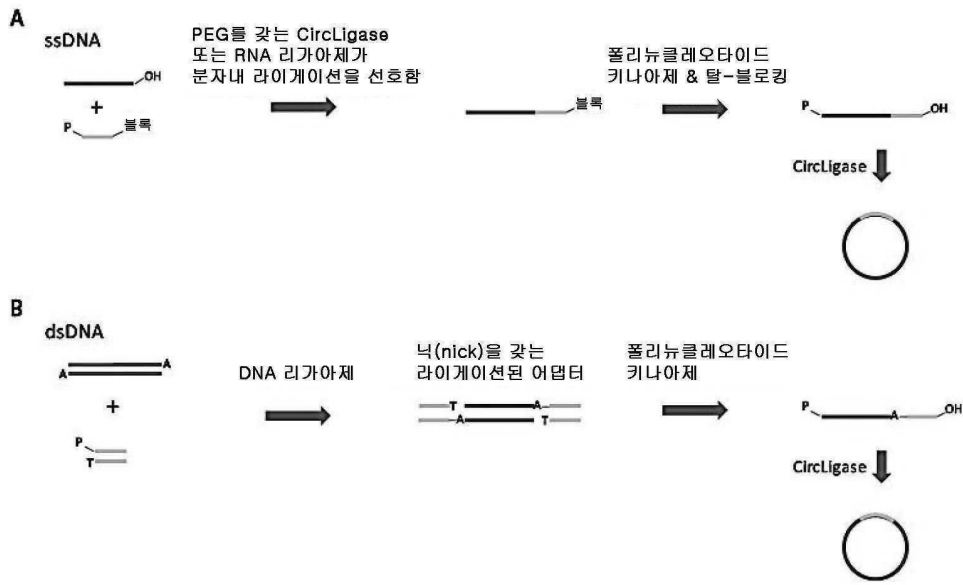
도면4



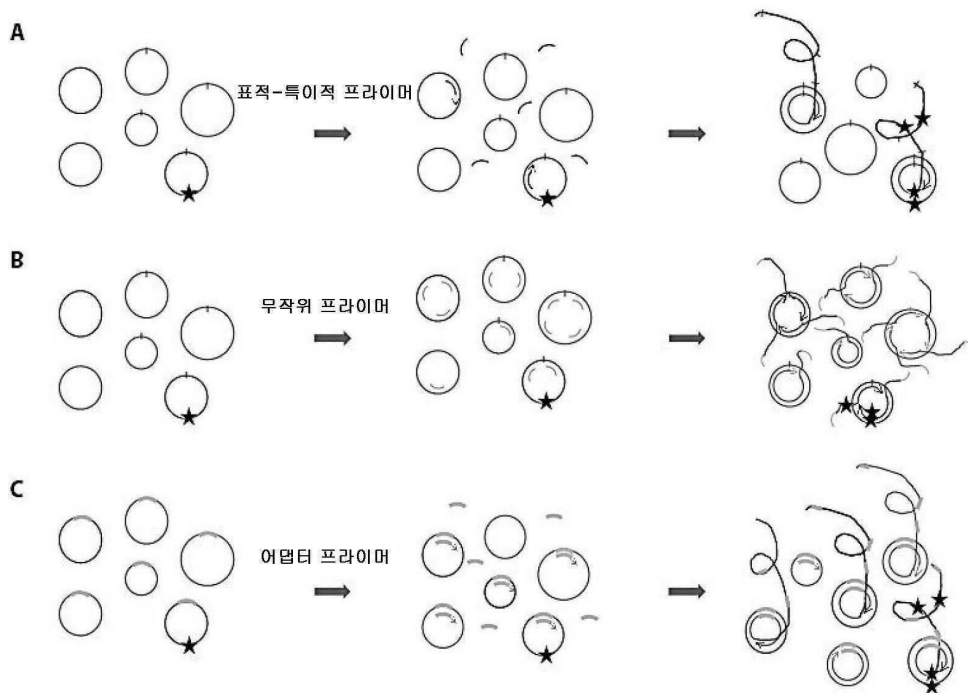
도면5



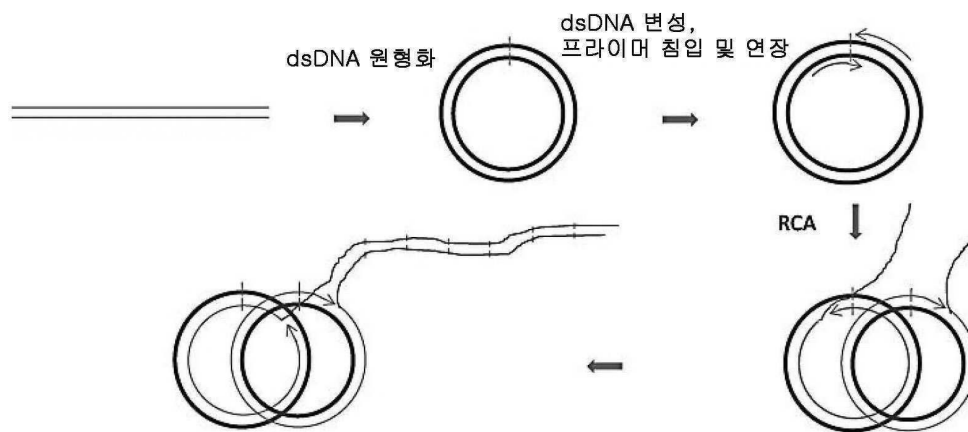
도면6



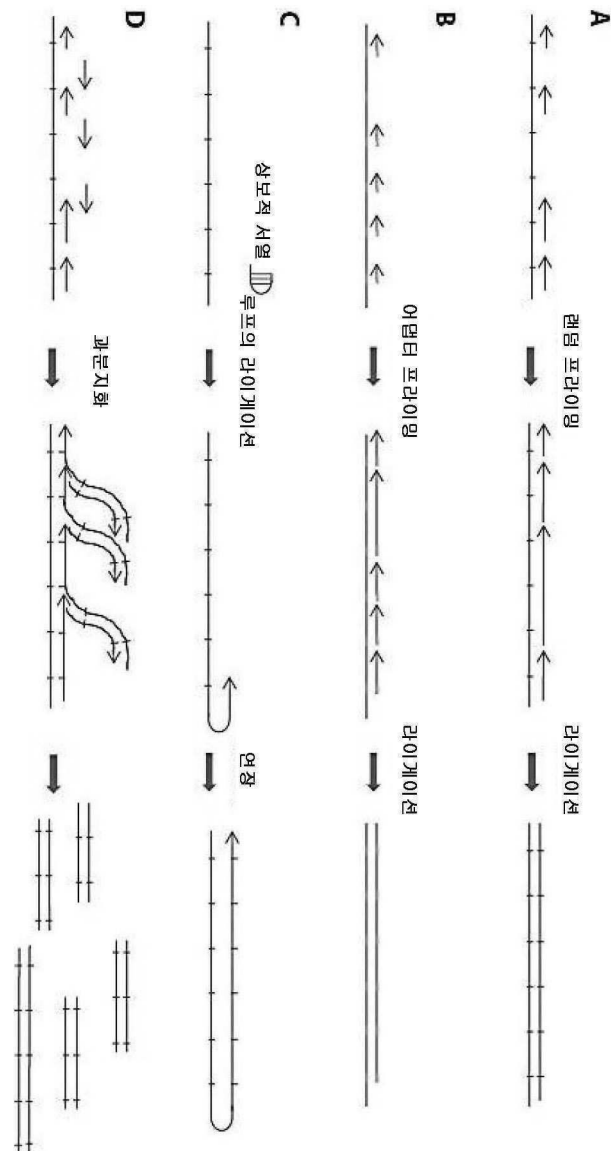
도면7



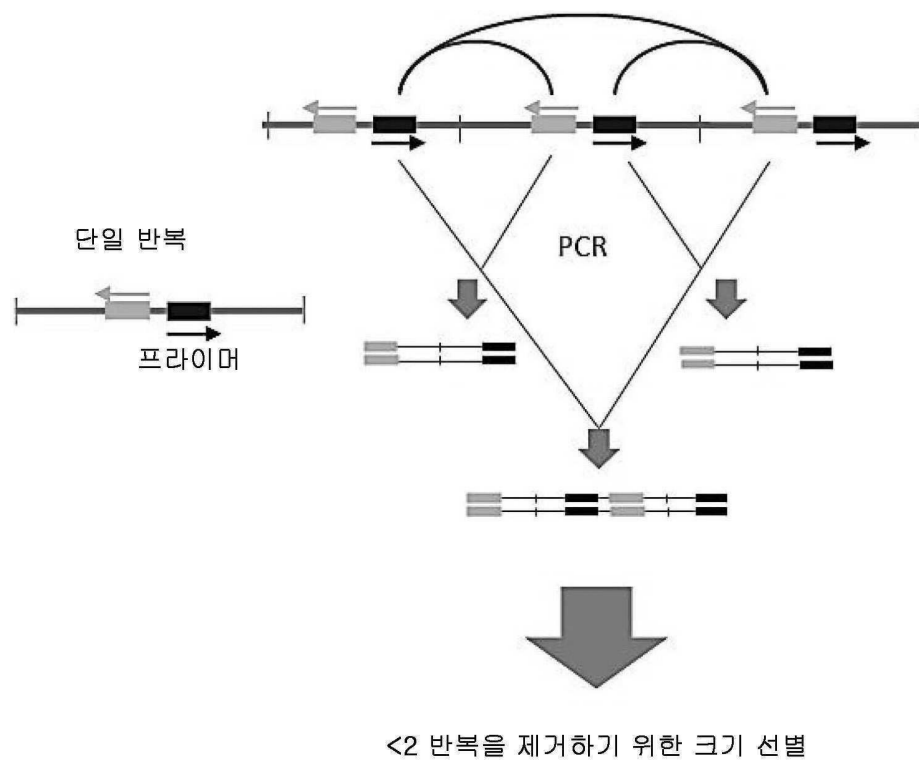
도면8



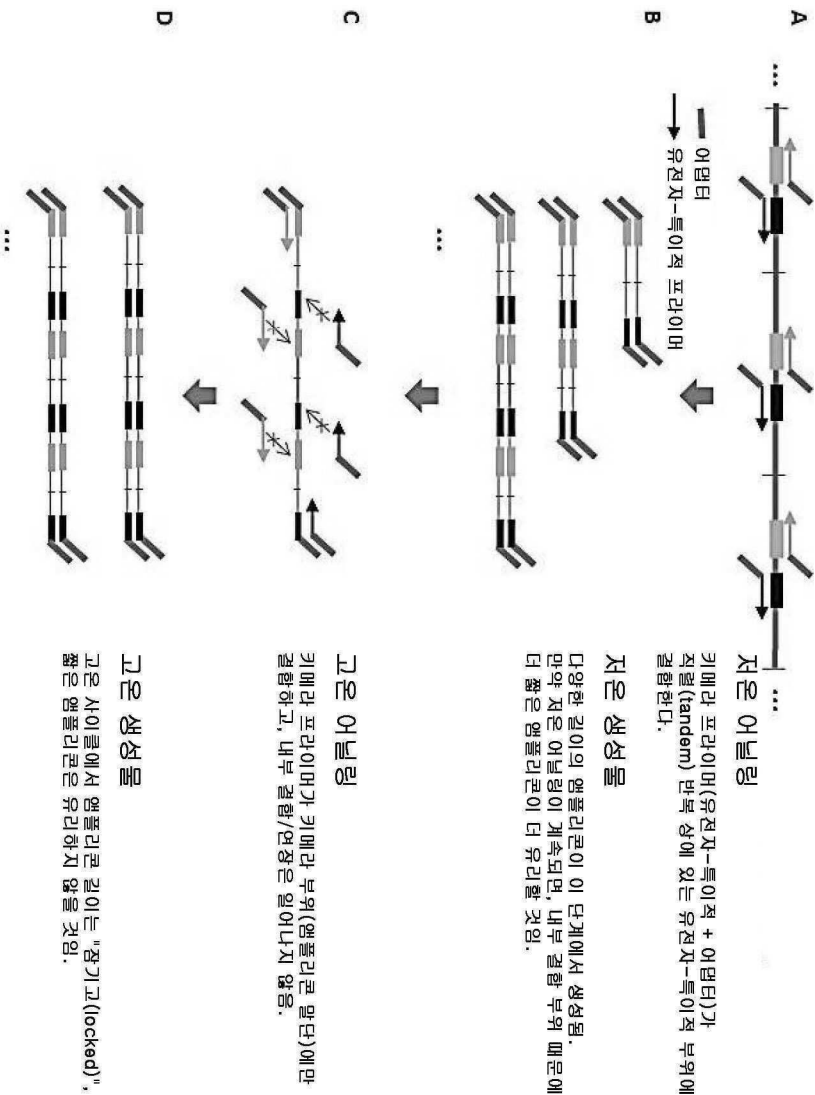
도면9



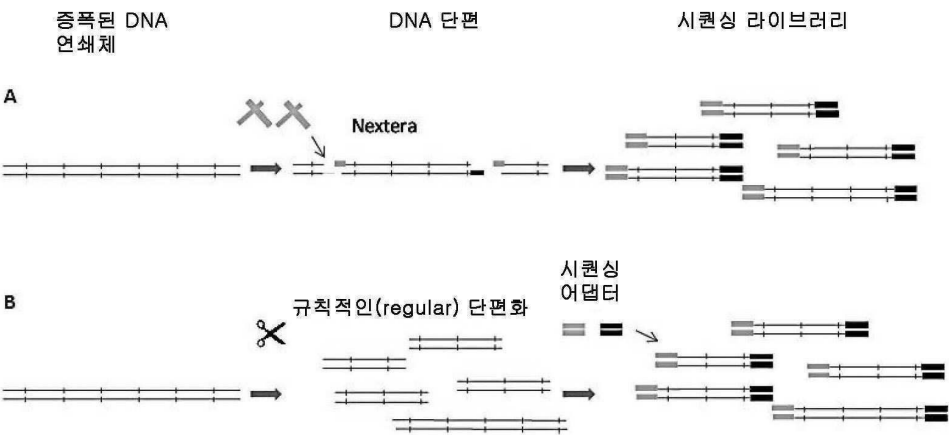
도면10



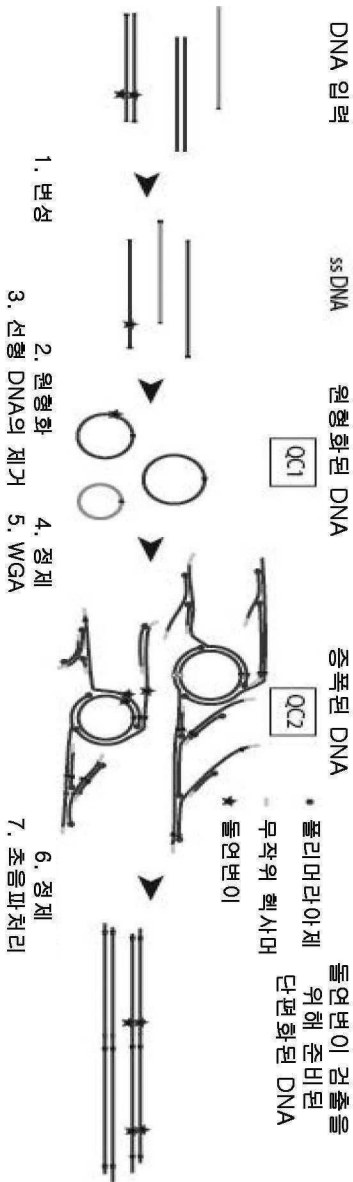
도면11



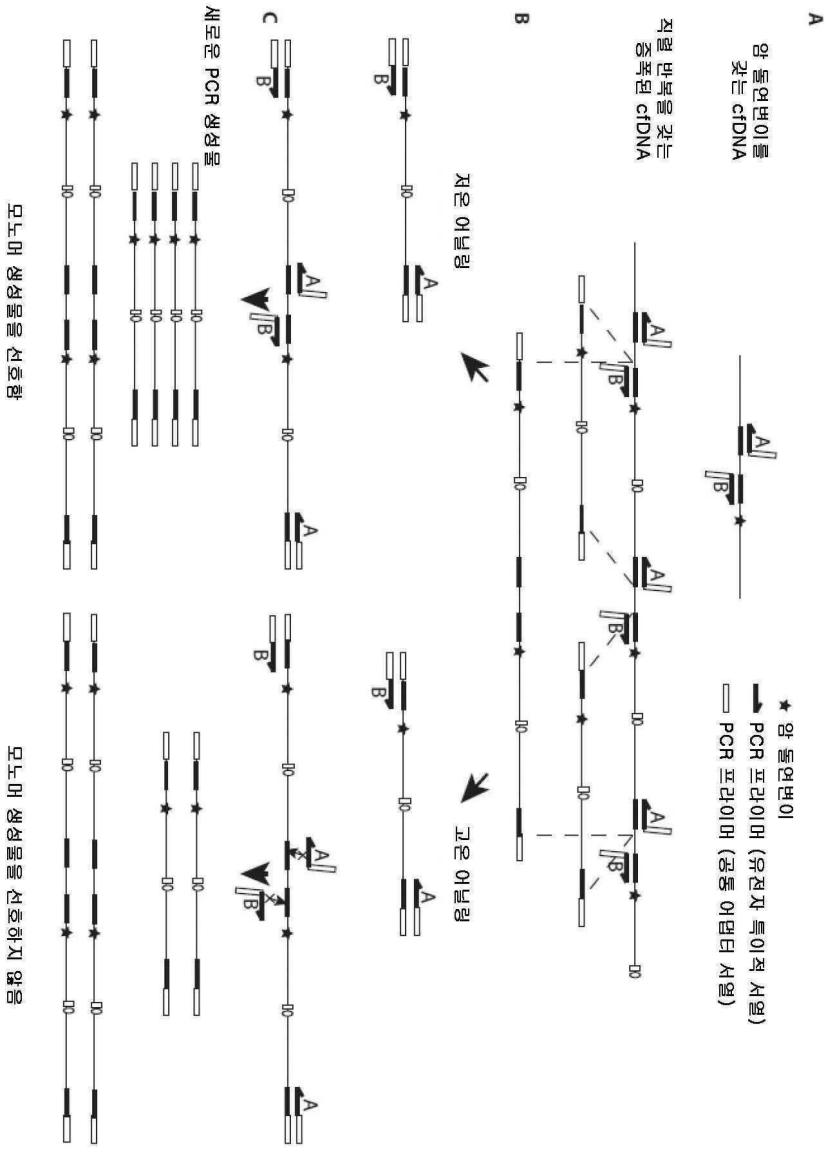
도면12



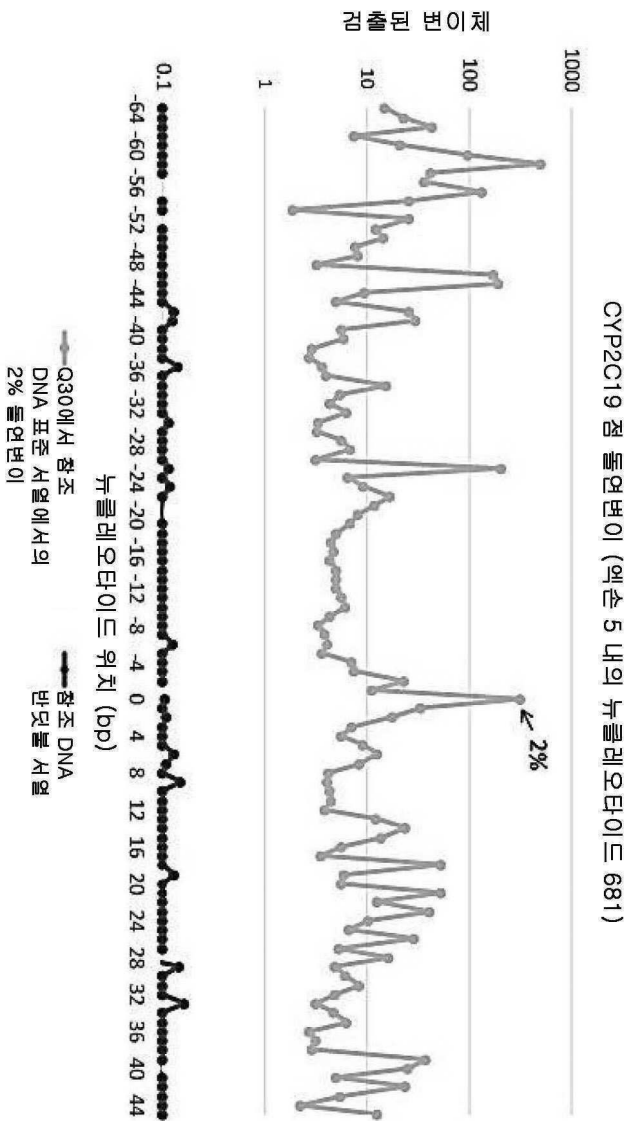
도면14



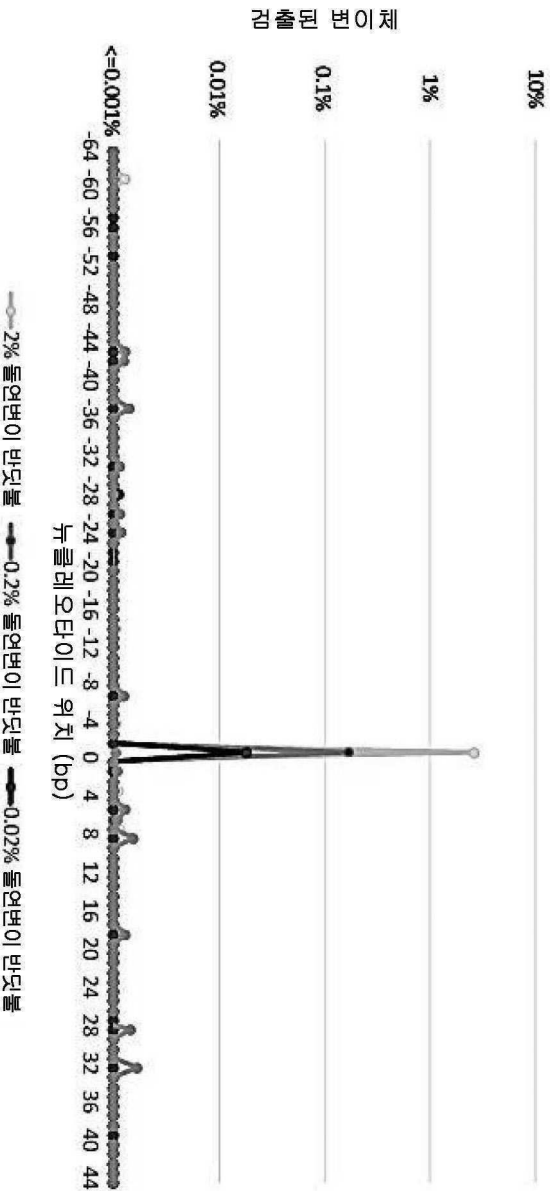
도면15



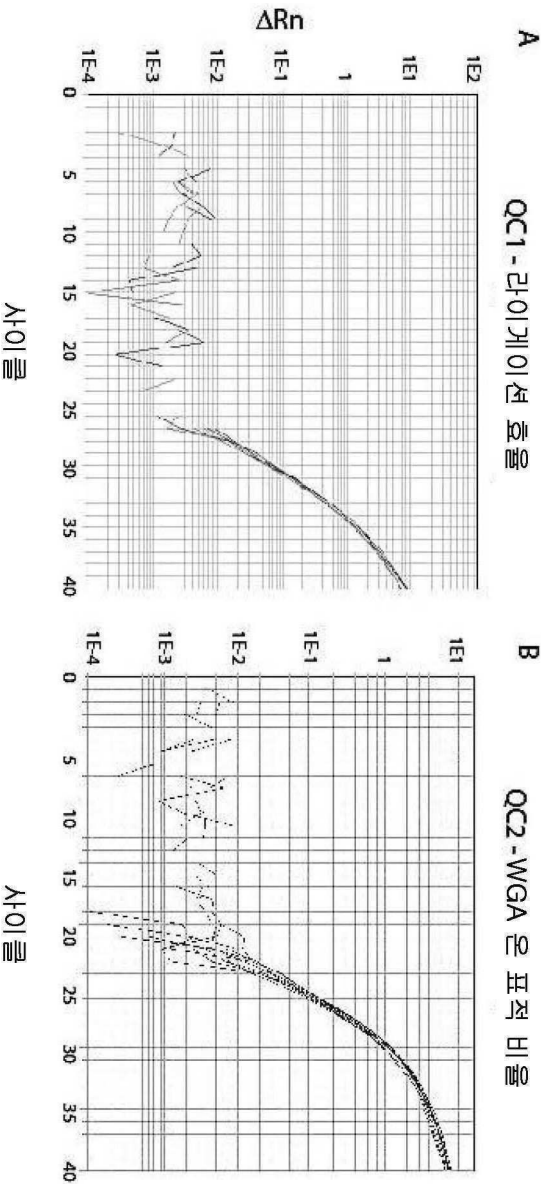
도면16



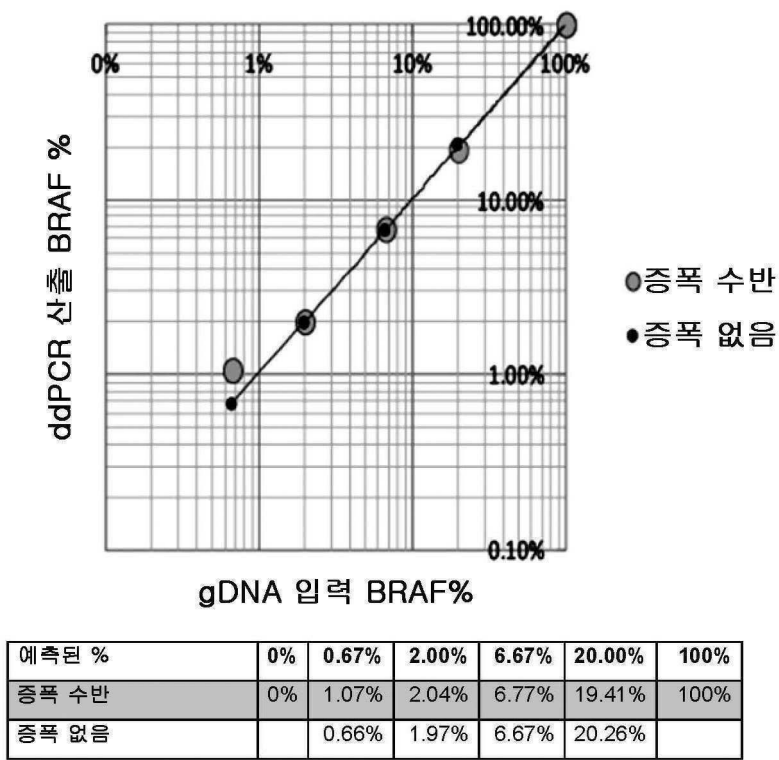
도면17



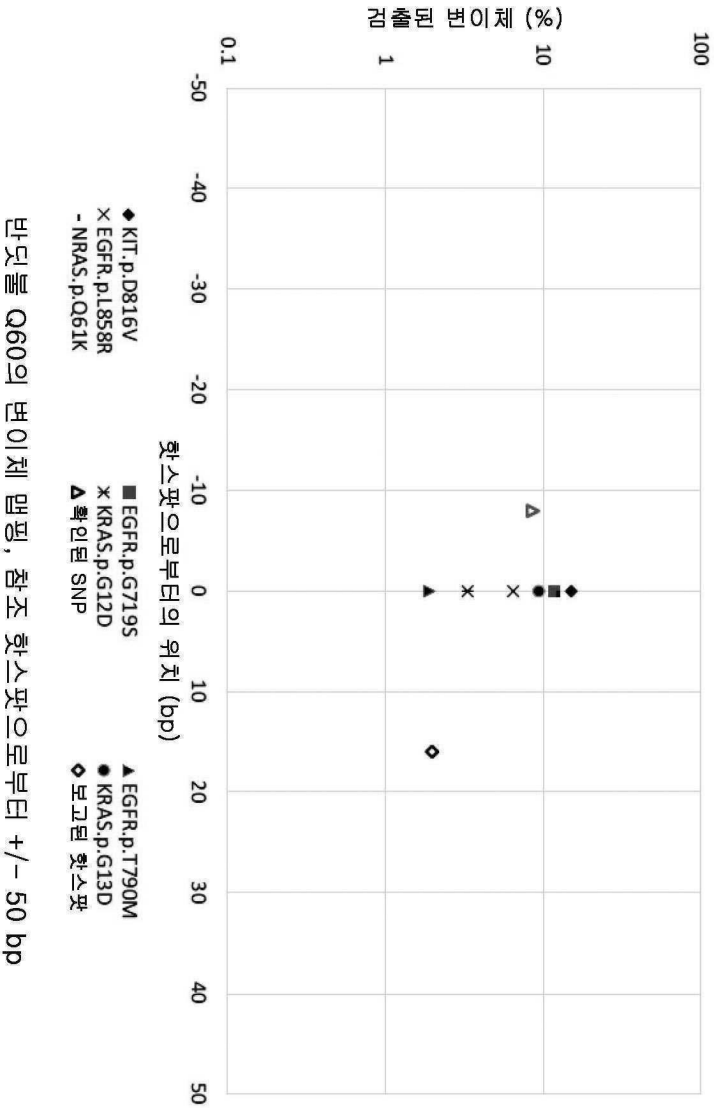
도면18



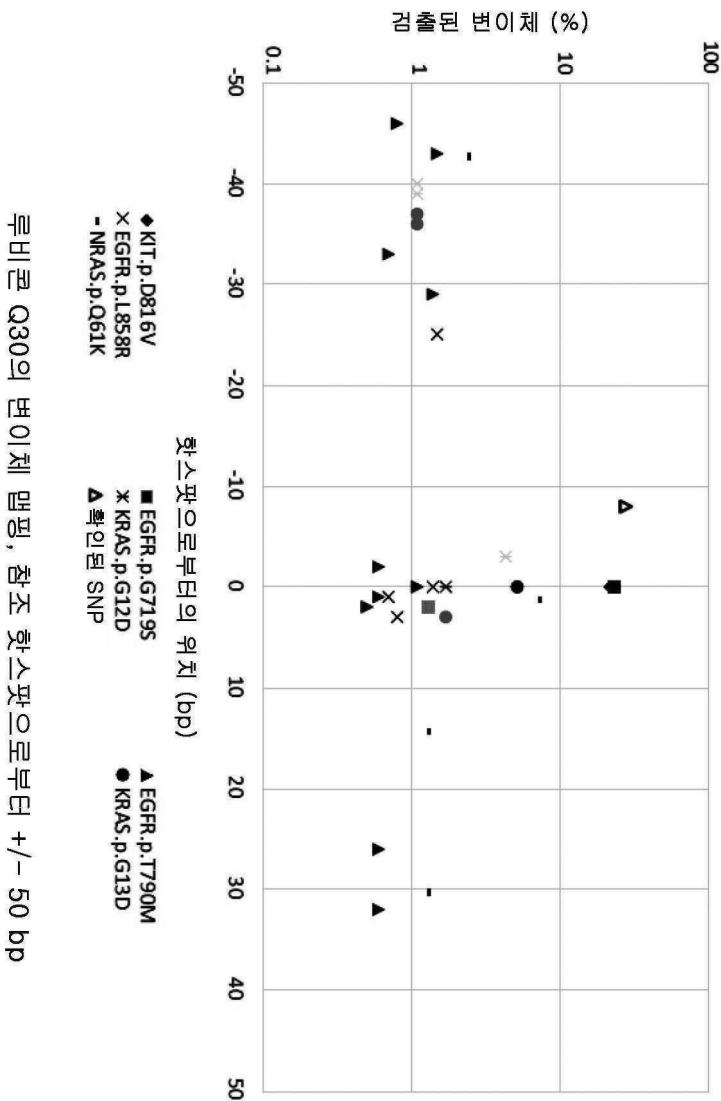
도면19



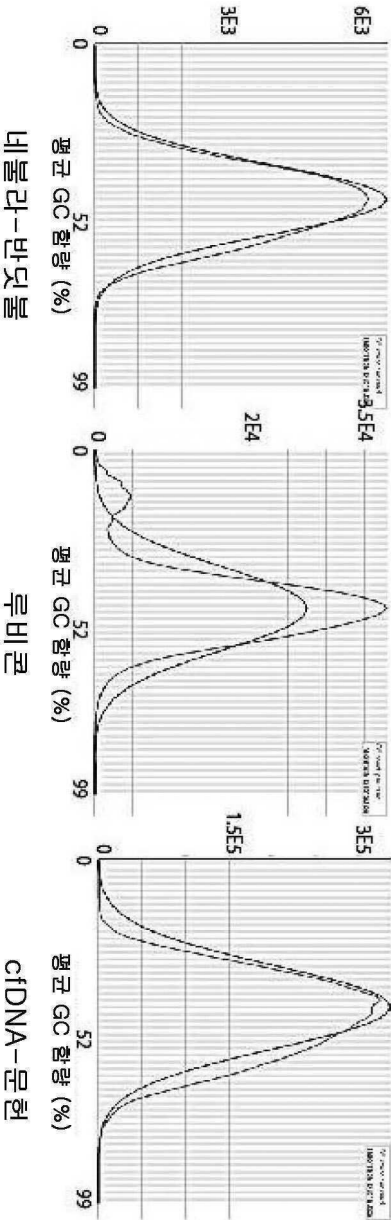
도면20



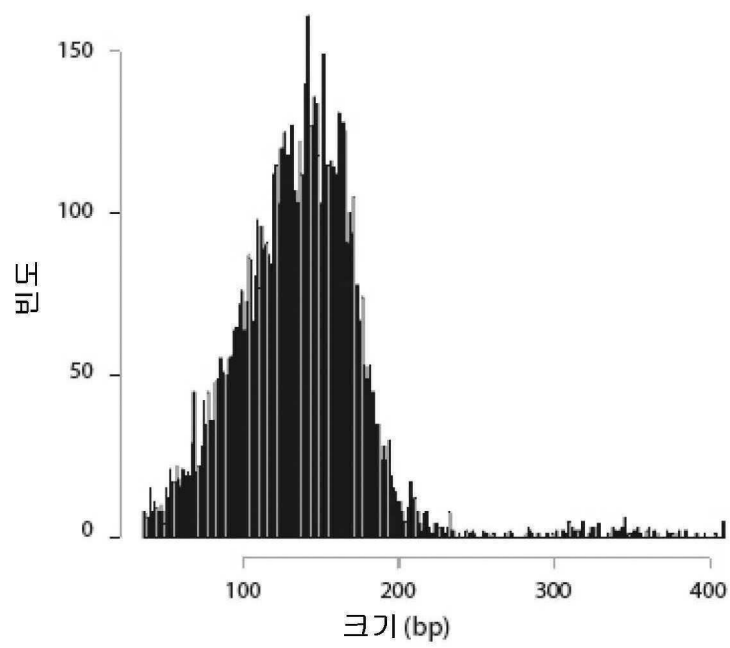
도면21



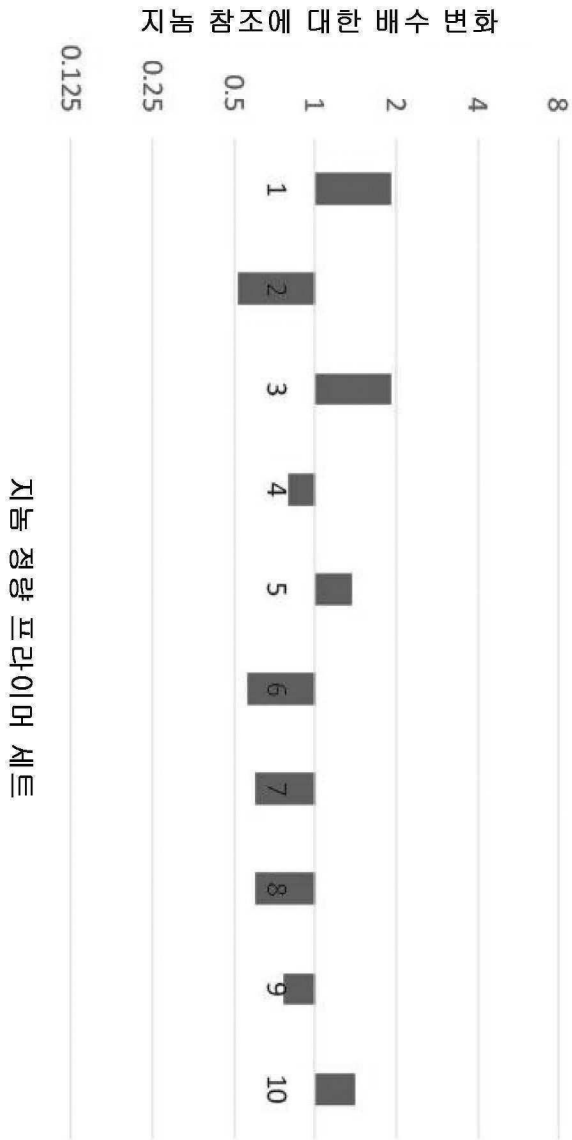
도면22



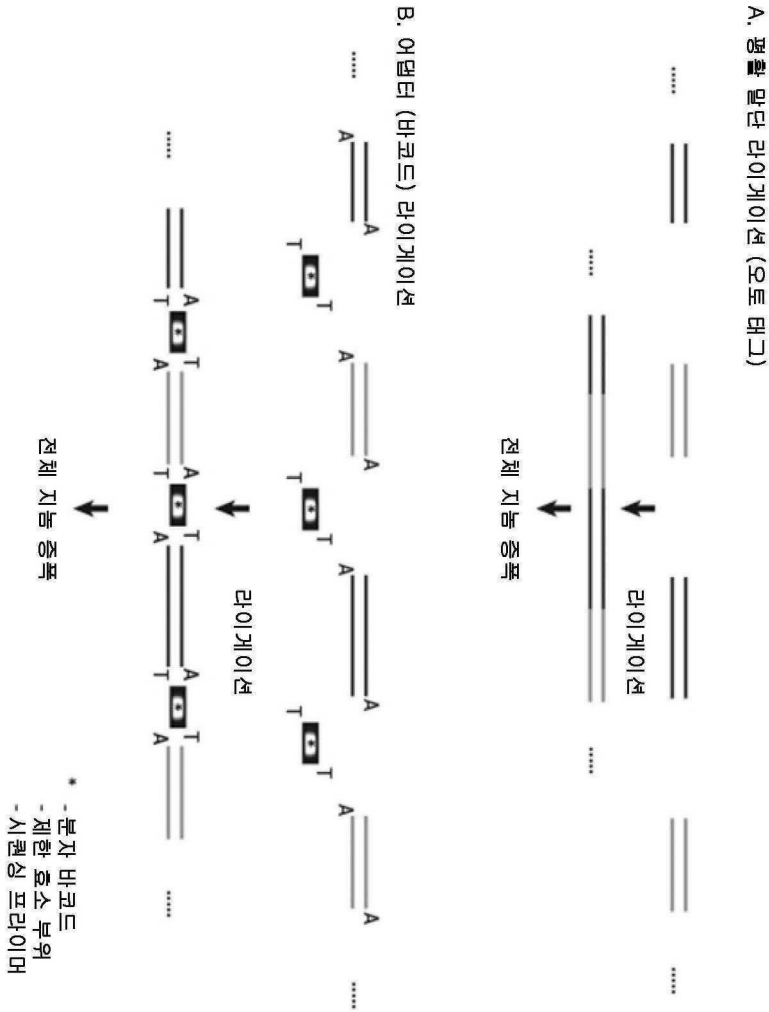
도면23



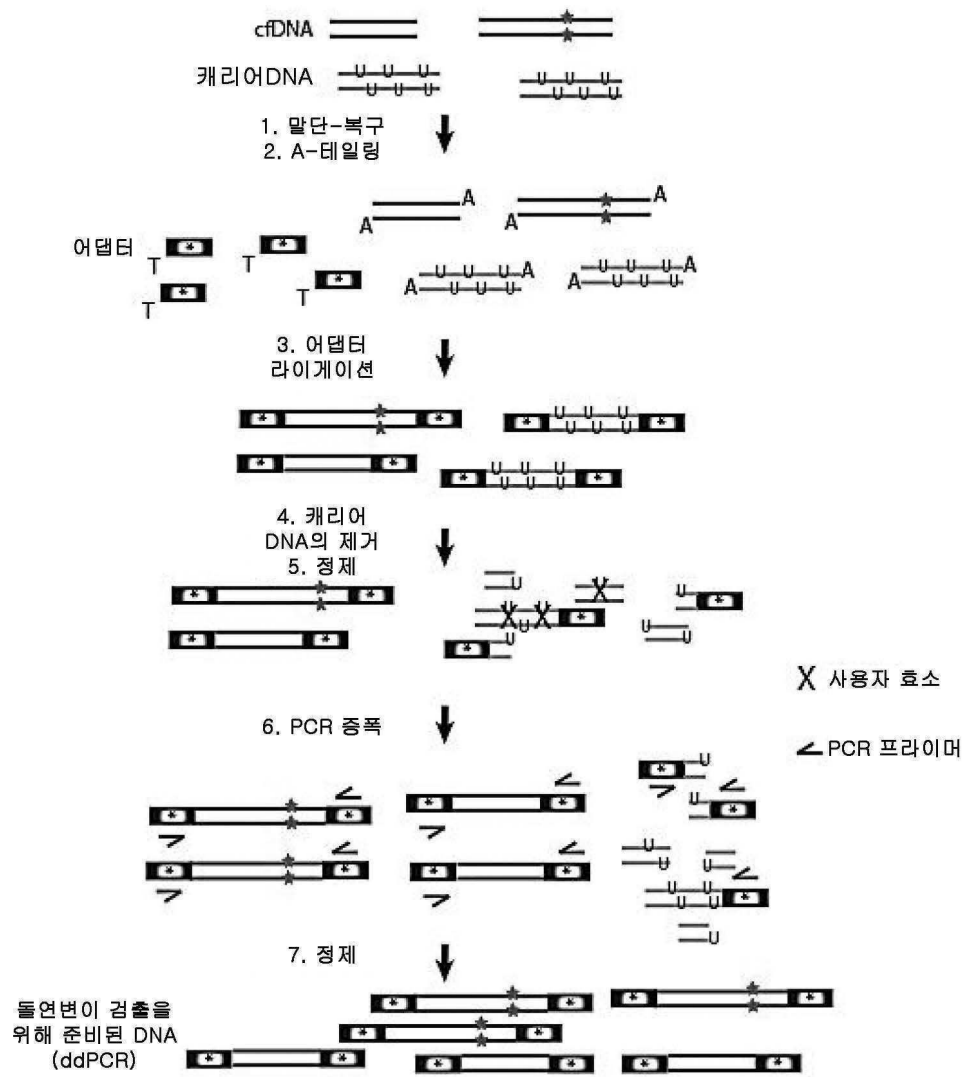
도면24



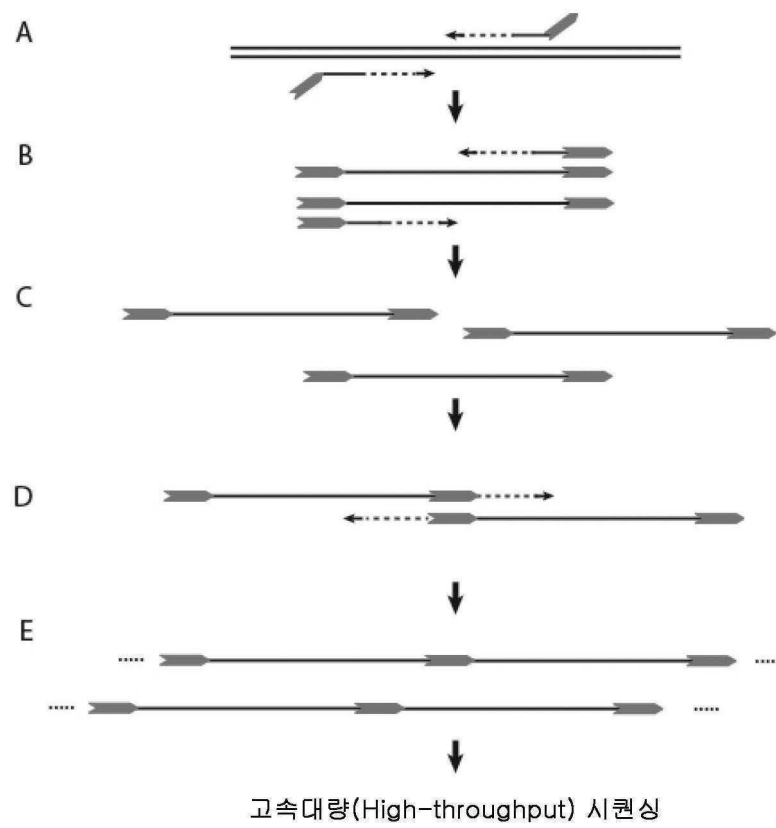
도면25



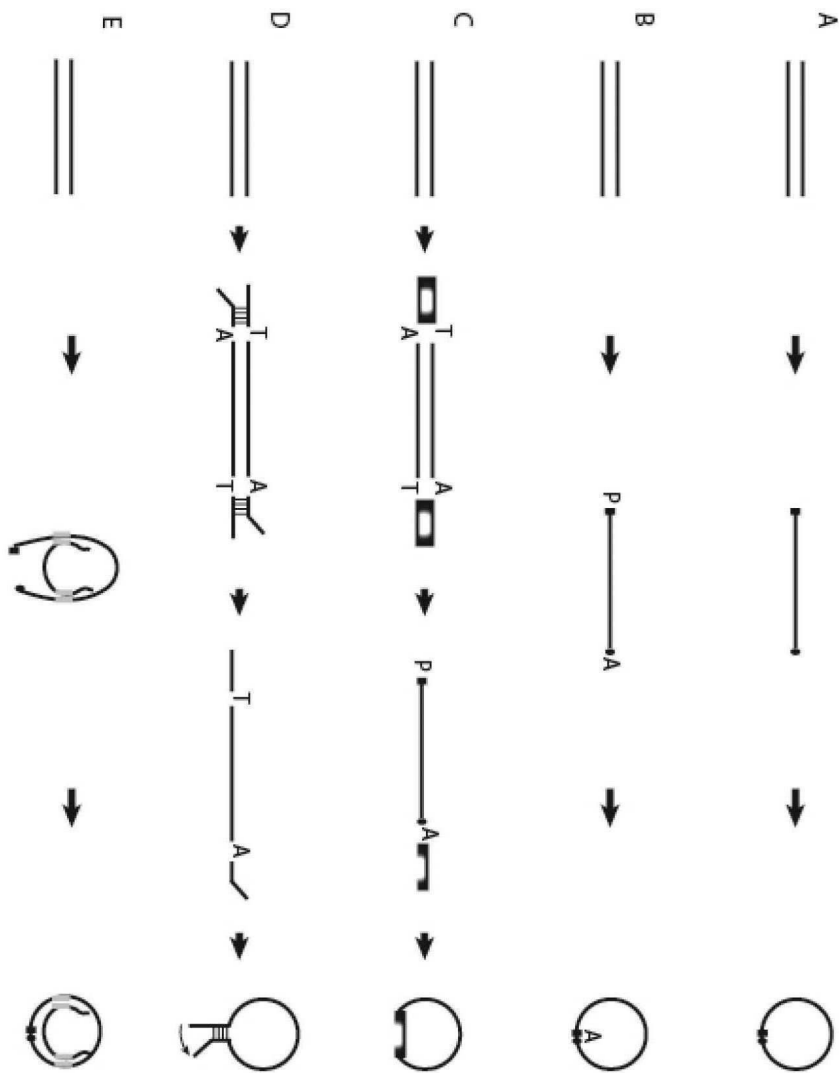
도면26



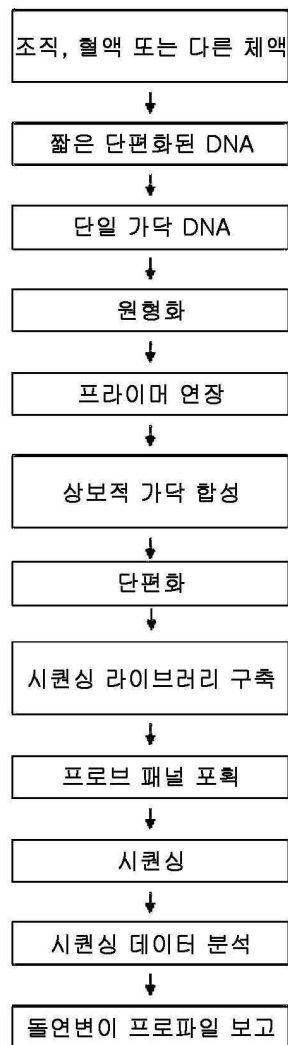
도면27



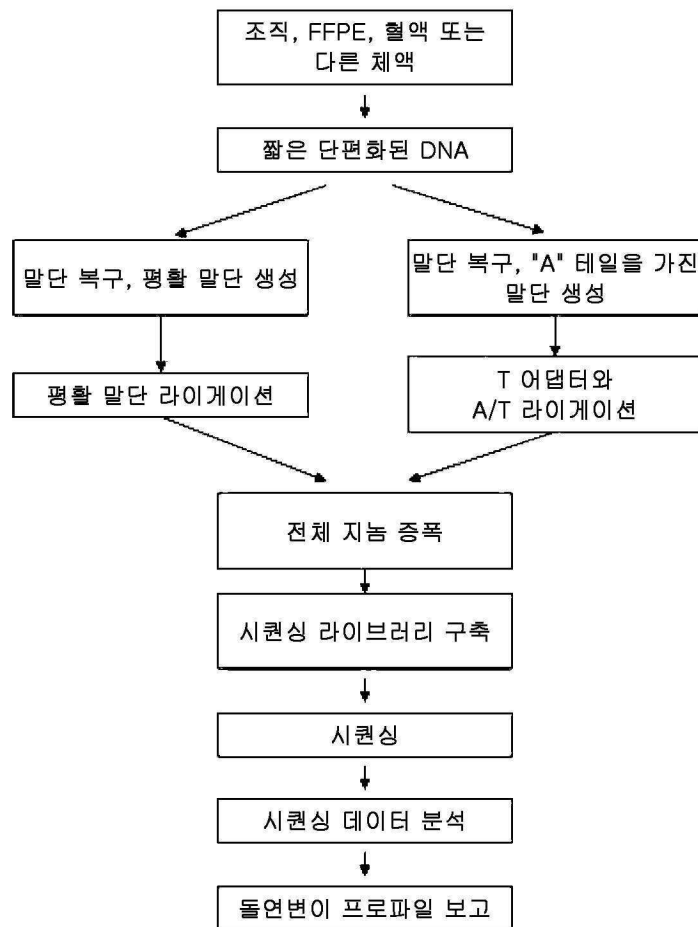
도면28



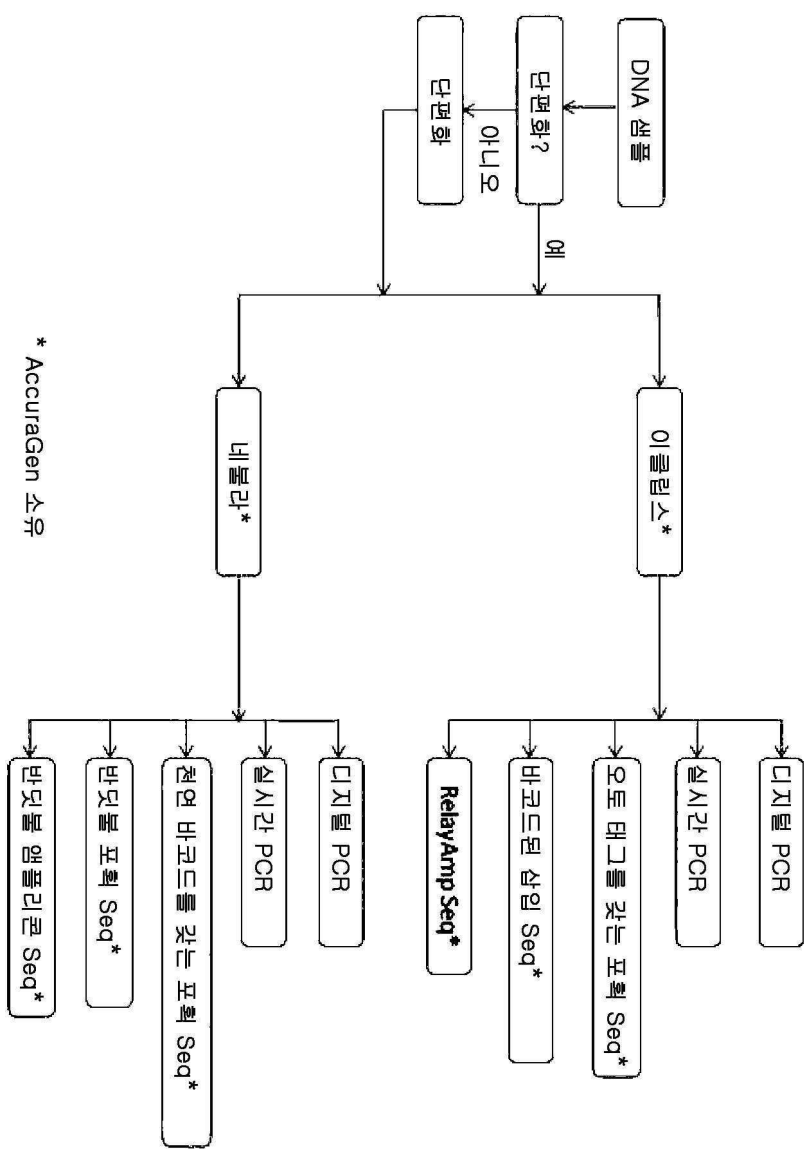
도면29



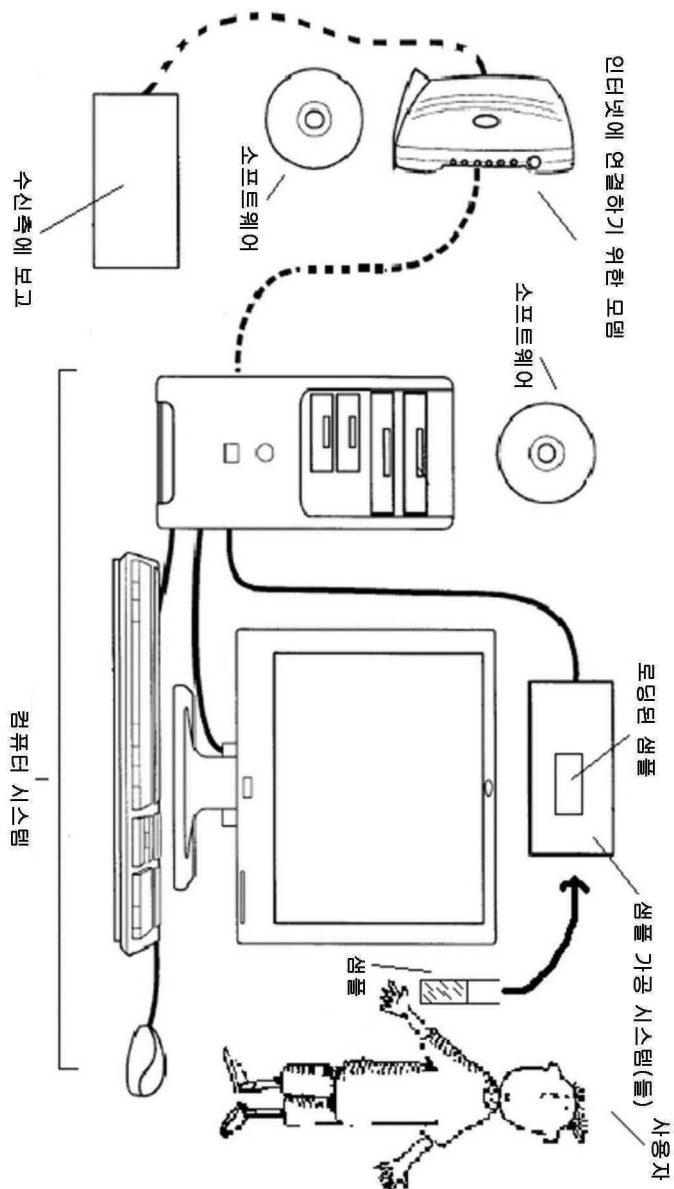
도면30



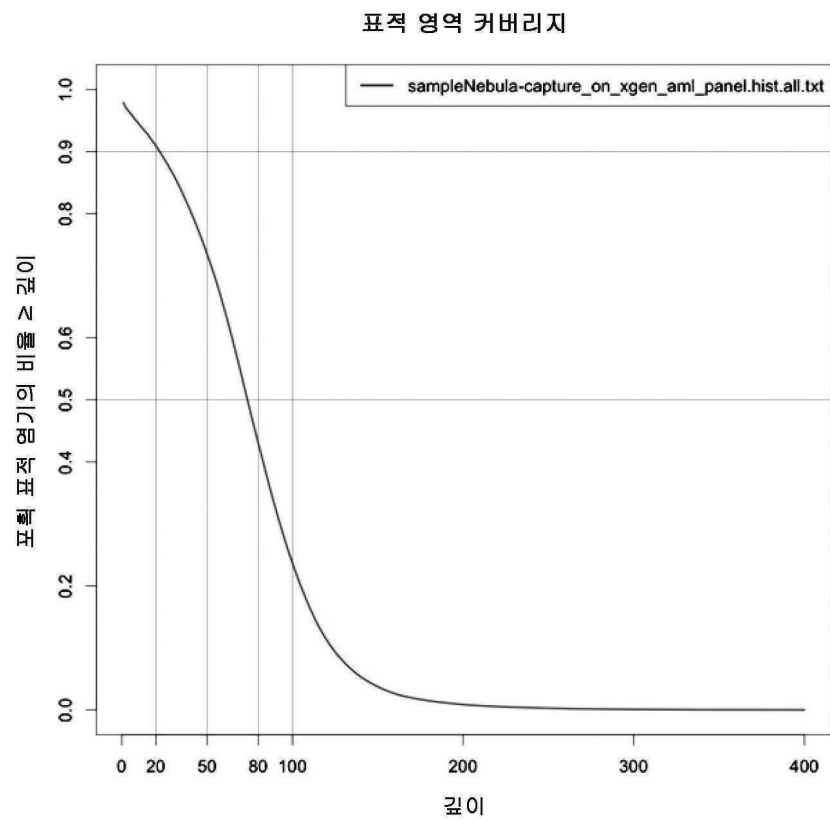
도면31



도면32



도면33



서열 목록

SEQUENCE LISTING

<110> ACCURAGEN, INC.

<120> COMPOSITIONS AND METHODS FOR DETECTING RARE SEQUENCE VARIANTS

<130> 47608-701.601

<140> PCT/US2014/069848

<141> 2014-12-11

<150> 62/010,975

<151> 2014-06-11

<150> 61/987,414

<151> 2014-05-01

<150> 61/914,907

<151> 2013-12-11

<160> 174

<170> PatentIn version 3.5

<210> 1

<211> 10

<212> DNA
 <213> Unknown
 <220><223> Description of Unknown: HBB oligonucleotide
 <400> 1
 cagctaccat 10

<210> 2
 <211> 10
 <212> DNA
 <213> Unknown
 <220><223> Description of Unknown: HBB oligonucleotide
 <400> 2
 caggtaccat 10

<210> 3
 <211> 21
 <212> DNA
 <213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic
 primer
 <400> 3
 caagcagaag acggcatagc a 21

<210> 4
 <211> 33
 <212> DNA
 <213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic
 primer
 <400> 4
 acactctttc cctacacgac gctcttccga tct 33

<210> 5
 <211> 23
 <212> DNA
 <213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic

primer
<400> 5
gctttgagct gttctttgtc att 23
<210> 6
<211> 24
<212> DNA
<213> Artificial Sequence
<220><223> Description of Artificial Sequence: Synthetic
primer
<400> 6
tttaattgtg tggaagatcc aatc 24
<210> 7
<211> 21
<212> DNA
<213> Artificial Sequence
<220><223> Description of Artificial Sequence: Synthetic
primer
<400> 7
ctttctcacc ttctgggatc c 21
<210> 8
<211> 19
<212> DNA
<213> Artificial Sequence
<220><223> Description of Artificial Sequence: Synthetic
primer
<400> 8
ccatcacgta ggcttcctg 19
<210> 9
<211> 20
<212> DNA
<213> Artificial Sequence
<220><223>
> Description of Artificial Sequence: Synthetic
primer

<400> 9
gacatagtcc aggaggcagc 20
<210> 10
<211> 18
<212> DNA
<213> Artificial Sequence
<220><223> Description of Artificial Sequence: Synthetic
primer
<400> 10
aagcgacggt cctccaag 18
<210> 11
<211> 18
<212> DNA
<213> Artificial Sequence
<220><223> Description of Artificial Sequence: Synthetic
primer
<
400> 11
agtacgttcc tggctgcc 18
<210> 12
<211> 21
<212> DNA
<213> Artificial Sequence
<220><223> Description of Artificial Sequence: Synthetic
primer
<400> 12
atccacttga taggcacctt g 21
<210> 13
<211> 18
<212> DNA
<213> Artificial Sequence
<220><223> Description of Artificial Sequence: Synthetic
primer
<400> 13

tctcgtggc agggattc	18
<210> 14	
<211> 20	
<212> DNA	
<213> Artificial Sequence	
<220><223> Description of Artificial Sequence: Synthetic primer	
<400> 14	
aactttgggc gactatctgc	20
<210> 15	
<211> 21	
<212> DNA	
<213> Artificial Sequence	
<220><223> Description of Artificial Sequence: Synthetic primer	
<400> 15	
ttggagtctg taggacttgg c	21
<210> 16	
<211> 21	
<212> DNA	
<213> Artificial Sequence	
<220>	
><223> Description of Artificial Sequence: Synthetic primer	
<400> 16	
ctgctgtggg atgaggtact c	21
<210> 17	
<211> 22	
<212> DNA	
<213> Artificial Sequence	
<220><223> Description of Artificial Sequence: Synthetic primer	
<400> 17	
catggaatgc ttgtaccaca tc	22

<210> 18
 <211> 23
 <212> DNA
 <213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic
 primer

<400> 18
 cagtttgaac agttgtctgg atc 23

<210> 19
 <211> 23
 <212> DNA
 <213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic
 primer

<400> 19
 tgtttctgct aacgatctct ttg 23

<210> 20
 <211> 23
 <212> DNA
 <213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic
 primer

<400> 20
 caggaaatcc catagcaata atg 23

<210> 21
 <211> 25
 <212> DNA
 <213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic
 primer

<400> 21
 gctttgaatc caaaacctt aaaac 25

<210> 22

<211> 26

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic
primer

<400> 22

tacagtacat tcatacctac ctctgc 26

<210> 23

<211> 22

<212> DNA

<213> Artificial Sequence

<220>

><223> Description of Artificial Sequence: Synthetic
primer

<400> 23

aaaggatatt gtgcaactgt gg 22

<210> 24

<211> 23

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic
primer

<400> 24

ccatagaaat ctagggcctc ttg 23

<210> 25

<211> 24

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic
primer

<400> 25

ccagatgatt cttaacagg tagc 24

<210> 26

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

primer

<400> 26

gaacttgtct tccgctcgtg 20

<210> 27

<211> 22

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

primer

<400> 27

tctggctctg gtatgaagaa tg 22

<210> 28

<211> 23

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

primer

<400> 28

gctctatact gcaaagcta tcg 23

<210> 29

<211> 21

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

primer

<400> 29

ttggagaaaa gtatcggttg g 21

<210> 30

<211> 23

<212> DNA

<213> Artificial Sequence

<220

><223> Description of Artificial Sequence: Synthetic

primer

<400> 30

tggtgttaca gaagttgaac tgc 23

<210> 31

<211> 22

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

primer

<400> 31

aagagtcct tgacgataca gc 22

<210> 32

<211> 23

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

primer

<400> 32

cctgactcag actgacattc tcc 23

<210> 33

<211> 21

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

primer

<400> 33

atgttccgag agctgaatga g 21

<210> 34

<211> 22

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

primer
<400> 34
ttaaaggacc agaccagctt tc 22

<210> 35
<211> 22
<212> DNA
<213> Artificial Sequence
<220><223> Description of Artificial Sequence: Synthetic

primer
<400> 35
cttgggacct cttatcaagt gg 22

<210> 36
<211> 20
<212> DNA
<213> Artificial Sequence
<220><223> Description of Artificial Sequence: Synthetic

primer
<400> 36
aagcaagcag gacaagaagc 20

<210> 37
<211> 19
<212> DNA
<213> Artificial Sequence
<220>
><223> Description of Artificial Sequence: Synthetic

primer
<400> 37
gggacggaac agctttgag 19

<210> 38
<211> 25
<212> DNA
<213> Artificial Sequence
<220><223> Description of Artificial Sequence: Synthetic

primer

<400> 38	
caactacatg tgtaacagtt cctgc	25
<210> 39	
<211> 24	
<212> DNA	
<213> Artificial Sequence	
<220><223> Description of Artificial Sequence: Synthetic primer	
<400> 39	
gtggagtatt tggatgacag aaac	24
<210> 40	
<211> 21	
<212> DNA	
<213> Artificial Sequence	
<220><223> Description of Artificial Sequence: Synthetic primer	
<400> 40	
tgctcagata gcgatggtga g	21
<210> 41	
<211> 19	
<212> DNA	
<213> Artificial Sequence	
<220><223> Description of Artificial Sequence: Synthetic primer	
<400> 41	
ctgtgcagct gtgggttga	19
<210> 42	
<211> 21	
<212> DNA	
<213> Artificial Sequence	
<220><223> Description of Artificial Sequence: Synthetic primer	
<400> 42	

aagtctgtga cttgcacggt c	21
<210> 43	
<211> 21	
<212> DNA	
<213> Artificial Sequence	
<220><223> Description of Artificial Sequence: Synthetic primer	
<400> 43	
cctgtcatct tctgtccctt c	21
<210> 44	
<211> 21	
<212> DNA	
<213> Artificial Sequence	
<220>	
><223> Description of Artificial Sequence: Synthetic primer	
<400> 44	
aagacccagg tccagatgaa g	21
<210> 45	
<211> 23	
<212> DNA	
<213> Artificial Sequence	
<220><223> Description of Artificial Sequence: Synthetic primer	
<400> 45	
ctgctcttgt ctttcagact tcc	23
<210> 46	
<211> 24	
<212> DNA	
<213> Artificial Sequence	
<220><223> Description of Artificial Sequence: Synthetic primer	
<400> 46	
ctctgagtca ggaaacattt tcag	24

<210> 47
 <211> 23
 <212> DNA
 <213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic
 primer
 <400> 47
 aaagcaattt ctacacgaga tcc 23
 <210> 48
 <211> 23
 <212> DNA
 <213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic
 primer
 <400> 48
 attaacagc atgcattgaa ctg 23
 <210> 49
 <211> 20
 <212> DNA
 <213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic
 primer
 <400> 49
 aaattcccgt cgctatcaag 20
 <210> 50
 <211> 18
 <212> DNA
 <213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic
 primer
 <400> 50
 atggccagcg tggacaac 18
 <210> 51
 <211> 19

<212> DNA

<213> Artificial Sequence

<220

><223> Description of Artificial Sequence: Synthetic

primer

<400> 51

tgtccgggaa cacaagac 19

<210> 52

<211> 18

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

primer

<400> 52

tggcagccag gaacgtac 18

<210> 53

<211> 19

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

primer

<400> 53

aacaccgcag catgtcaag 19

<210> 54

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

primer

<400> 54

aagtggatgg cattggaatc 20

<210> 55

<211> 19

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic
 primer

<400> 55

cctggagaaa ggagaacgc 19

<210> 56

<211> 21

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic
 primer

<400> 56

agttccgtga gttgatcatc g 21

<210> 57

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic
 primer

<400> 57

acttctaccg tgcctgatg 20

<210> 58

<211> 19

<212> DNA

<213> Artificial Sequence

<220>

><223> Description of Artificial Sequence: Synthetic
 primer

<400> 58

cacagcaggg cttcttcag 19

<210> 59

<211> 21

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic
 primer
 <400> 59
 catgggcaac ttctctgttt c 21
 <210> 60
 <211> 20
 <212> DNA
 <213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic
 primer
 <400> 60
 aaactgatgg gaccactcc 20
 <210> 61
 <211> 24
 <212> DNA
 <213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic
 primer
 <400> 61
 aggagatatc aagaggatgg attc 24
 <210> 62
 <211> 21
 <212> DNA
 <213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic
 primer
 <400> 62
 tcctgcagaa agacttgaag g 21
 <210> 63
 <211> 25
 <212> DNA
 <213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic

```

    primer
<400> 63
ggattcaaag cataaaaacc attac
25
<210> 64
<211> 28
<212> DNA
<213> Artificial Sequence
<220><223> Description of Artificial Sequence: Synthetic
    primer
<400> 64
tatgttgat aacttaaacc ctagagac
28
<210> 65
<211> 21
<212> DNA
<213> Artificial Sequence
<220>
><223> Description of Artificial Sequence: Synthetic
    primer
<400> 65
ttgaagacca taaccacac c
21
<210> 66
<211> 24
<212> DNA
<213> Artificial Sequence
<220><223> Description of Artificial Sequence: Synthetic
    primer
<400> 66
aagtaaggac cagagacaaa aagg
24
<210> 67
<211> 22
<212> DNA
<213> Artificial Sequence
<220><223> Description of Artificial Sequence: Synthetic
    primer

```

<400> 67	
ggattataga ccagtggcac tg	22
<210> 68	
<211> 23	
<212> DNA	
<213> Artificial Sequence	
<220><223> Description of Artificial Sequence: Synthetic primer	
<400> 68	
catgtacttt gaggccctc agc	23
<210> 69	
<211> 21	
<212> DNA	
<213> Artificial Sequence	
<220><223> Description of Artificial Sequence: Synthetic primer	
<400> 69	
caggaccaga ggaaacctca g	21
<210> 70	
<211> 25	
<212> DNA	
<213> Artificial Sequence	
<220><223> Description of Artificial Sequence: Synthetic primer	
<400> 70	
cgtgcagata atgacaagga atatc	25
<210> 71	
<211> 26	
<212> DNA	
<213> Artificial Sequence	
<220><223> Description of Artificial Sequence: Synthetic primer	
<400> 71	
ggtcagttaa attaaacatt ttgtgg	26

<210> 72
 <211> 25
 <212> DNA
 <213> Artificial Sequence
 <220>
 ><223> Description of Artificial Sequence: Synthetic
 primer
 <400> 72
 gatgttagtg acaatgaacc tgatc 25
 <210> 73
 <211> 18
 <212> DNA
 <213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic
 primer
 <400> 73
 tcttgctac gccaccag 18
 <210> 74
 <211> 20
 <212> DNA
 <213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic
 primer
 <400> 74
 caggcccttc tgtcttgaac 20
 <210> 75
 <211> 20
 <212> DNA
 <213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic
 primer
 <400> 75
 gaacatctcg aagcgctcac 20
 <210> 76

<211> 26

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic
 primer

<400> 76

ttatggtata agttggtgtt ctgaag 26

<210> 77

<211> 24

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic
 primer

<400> 77

agaggtccca agacttagta cctg 24

<210> 78

<211> 21

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic
 primer

<400> 78

gcttgcttac ctgccttagt g 21

<210> 79

<211> 22

<212> DNA

<213> Artificial Sequence

<220>

><223> Description of Artificial Sequence: Synthetic
 primer

<400> 79

ttccgtccca gtagattacc ac 22

<210> 80

<211> 23

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic
 primer

<400> 80

catgtagttg tagtgatgg tgg 23

<210> 81

<211> 21

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic
 primer

<400> 81

atactccaca cgcaaatttc c 21

<210> 82

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic
 primer

<400> 82

ctatctgagc agcgctcatg 20

<210> 83

<211> 18

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic
 primer

<400> 83

gcaggtcttg gccagttg 18

<210> 84

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic
 primer

<400> 84

tgtcccagaa tgcaagaagc 20

<210> 85

<211> 18

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic
 primer

<400> 85

gatgacaggg gccaggag 18

<210> 86

<211> 21

<212> DNA

<213> Artificial Sequence

<220>

><223> Description of Artificial Sequence: Synthetic
 primer

<400> 86

tgggtcttca gtgaaccatt g 21

<210> 87

<211> 21

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic
 primer

<400> 87

gagcagaaag tcagtcccat g 21

<210> 88

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

primer

<400> 88
gctcgacgct aggatctgac 20

<210> 89
<211> 49
<212> DNA
<213> Artificial Sequence
<220><223> Description of Artificial Sequence: Synthetic
primer

<400> 89
gttcagagtt ctacagtcg acgatcgctt tgagctgttc tttgtcatt 49

<210> 90
<211> 50
<212> DNA
<213> Artificial Sequence
<220><223> Description of Artificial Sequence: Synthetic
primer

<400> 90
gttcagagtt ctacagtcg acgatcctta attgtgtgga agatccaatc 50

<210> 91
<211> 47
<212> DNA
<213> Artificial Sequence
<220><223> Description of Artificial Sequence: Synthetic
primer

<400> 91
gttcagagtt ctacagtcg acgatccttt ctcaccttct gggatcc 47

<210> 92
<211> 45
<212> DNA
<213> Artificial Sequence
<220><223> Description of Artificial Sequence: Synthetic
primer

<400> 92
gttcagagtt ctacagtccg acgatcccat cacgtaggct tcctg 45
<210> 93
<211> 46
<212> DNA
<213> Artificial Sequence
<220>
><223> Description of Artificial Sequence: Synthetic
primer
<400> 93
gttcagagtt ctacagtccg acgatcgaca tagtccagga ggcagc 46
<210> 94
<211> 44
<212> DNA
<213> Artificial Sequence
<220><223> Description of Artificial Sequence: Synthetic
primer
<400> 94
gttcagagtt ctacagtccg acgatcaagc gacggctctc caag 44
<210> 95
<211> 44
<212> DNA
<213> Artificial Sequence
<220><223> Description of Artificial Sequence: Synthetic
primer
<400> 95
gttcagagtt ctacagtccg acgatcagta cgttcctggc tgcc 44
<210> 96
<211> 47
<212> DNA
<213> Artificial Sequence
<220><223> Description of Artificial Sequence: Synthetic
primer
<400> 96

gttcagagtt ctacagtccg acgatcatcc acttgatagg caccttg 47

<210> 97

<211> 44

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

primer

<400> 97

gttcagagtt ctacagtccg acgatctctc gctggcaggg attc 44

<210> 98

<211> 46

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

primer

<400> 98

gttcagagtt ctacagtccg acgatcaact ttgggcgact atctgc 46

<210> 99

<211> 47

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

primer

<400> 99

gttcagagtt ctacagtccg acgatcttgg agtctgtagg acttggc 47

<210> 100

<211> 47

<212> DNA

<213> Artificial Sequence

<220>

><223> Description of Artificial Sequence: Synthetic

primer

<400> 100

gttcagagtt ctacagtccg acgatcctgc tgtgggatga ggtactc 47

<210> 101

<211> 48

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic
primer

<400> 101

gttcagagtt ctacagtccg acgatccatg gaatgcttgt accacatc 48

<210> 102

<211> 49

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic
primer

<400> 102

gttcagagtt ctacagtccg acgatccagt ttgaacagtt gtctggatc 49

<210> 103

<211> 49

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic
primer

<400> 103

gttcagagtt ctacagtccg acgatctgtt tctgctaacg atctctttg 49

<210> 104

<211> 49

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic
primer

<400> 104

gttcagagtt ctacagtccg acgatccagg aaatcccata gcaataatg 49

<210> 105

<211> 51
 <212> DNA
 <213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic
 primer
 <400> 105
 gttcagagtt ctacagtccg acgatcgctt tgaatccaaa aaccttaaaa c 51
 <210> 106
 <211> 52
 <212> DNA
 <213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic
 primer
 <400> 106
 gttcagagtt ctacagtccg acgatctaca gtacattcat acctacctct gc 52
 <210> 107
 <211> 48
 <212> DNA
 <213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic
 primer
 <400> 107
 gttcagagtt ctacagtccg acgatcaaag gatattgtgc aactgtgg 48
 <210> 108
 <211> 49
 <212> DNA
 <213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic
 primer
 <400> 108
 gttcagagtt ctacagtccg acgatcccat agaaatctag ggcctcttg 49
 <210> 109
 <211> 50
 <212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic
primer

<400> 109

gttcagagtt ctacagtccg acgatcccag atgattcttt aacaggtagc 50

<210> 110

<211> 46

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic
primer

<400> 110

gttcagagtt ctacagtccg acgatcgaac ttgtcttccc gtcgtg 46

<210> 111

<211> 48

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic
primer

<400> 111

gttcagagtt ctacagtccg acgatctctg gtcttggtat gaagaatg 48

<210> 112

<211> 49

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic
primer

<400> 112

gttcagagtt ctacagtccg acgatcgctc tatactgcaa atgctatcg 49

<210> 113

<211> 47

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

primer

<400> 113

gttcagagtt ctacagtccg acgatcttgg agaaaagtat cggttgg

47

<210> 114

<211> 49

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

primer

<400> 114

gttcagagtt ctacagtccg acgatctggg gttacagaag ttgaactgc

49

<210> 115

<211> 48

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

primer

<400> 115

gttcagagtt ctacagtccg acgatcaaga gtgccttgac gatacagc

48

<210> 116

<211> 49

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

primer

<400> 116

gttcagagtt ctacagtccg acgatccctg actcagactg acattctcc

49

<210> 117

<211> 47

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

primer

<400> 117

gttcagagtt ctacagtccg acgatcatgt tccgagagct gaatgag 47

<210> 118

<211> 48

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

primer

<400> 118

gttcagagtt ctacagtccg acgatcttaa aggaccagac cagctttc 48

<210> 119

<211> 48

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

primer

<400> 119

gttcagagtt ctacagtccg acgatccttg ggacctctta tcaagtgg 48

<210> 120

<211> 46

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

primer

<400> 120

gttcagagtt ctacagtccg acgatcaagc aagcaggaca agaagc 46

<210> 121

<211> 45

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

primer

<400> 121
gttcagagtt ctacagtccg acgatcggga cggaacagct ttgag 45
<210> 122
<211> 51
<212> DNA
<213> Artificial Sequence
<220><223> Description of Artificial Sequence: Synthetic
primer
<400> 122
gttcagagtt ctacagtccg acgatccaac tacatgtgta acagttcctg c 51
<210> 123
<211> 50
<212> DNA
<213> Artificial Sequence
<220><223> Description of Artificial Sequence: Synthetic
primer
<400> 123
gttcagagtt ctacagtccg acgatcgtgg agtatttgga tgacagaaac 50
<210> 124
<211> 47
<212> DNA
<213> Artificial Sequence
<220><223> Description of Artificial Sequence: Synthetic
primer
<400> 124
gttcagagtt ctacagtccg acgatctgct cagatagcga tggtag 47
<210> 125
<211> 45
<212> DNA
<213> Artificial Sequence
<220><223> Description of Artificial Sequence: Synthetic
primer
<400> 125

gttcagagtt ctacagtccg acgatcctgt gcagctgtgg gttga 45

<210> 126

<211> 47

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic
primer

<400> 126

gttcagagtt ctacagtccg acgatcaagt ctgtgacttg cacggtc 47

<210> 127

<211> 47

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic
primer

<400> 127

gttcagagtt ctacagtccg acgatccctg tcatcttctg tcccttc 47

<210> 128

<211> 47

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic
primer

<400> 128

gttcagagtt ctacagtccg acgatcaaga cccaggtcca gatgaag 47

<210> 129

<211> 49

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic
primer

<400> 129

gttcagagtt ctacagtccg acgatcctgc tcttgtcttt cagacttcc 49

<210> 130

<211> 50

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic
primer

<400> 130

gttcagagtt ctacagtccg acgatacctct gaggcaggaa acattttcag 50

<210> 131

<211> 44

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic
primer

<400> 131

ccttggcacc cgagaattcc aaaagcaatt tctacacgag atcc 44

<210> 132

<211> 44

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic
primer

<400> 132

ccttggcacc cgagaattcc aattaacag catgcattga actg 44

<210> 133

<211> 41

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic
primer

<400> 133

ccttggcacc cgagaattcc aaaattcccg tcgctatcaa g 41

<210> 134

<211> 39

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic
primer

<400> 134

ccttggcacc cgagaattcc aatggccagc gtggacaac 39

<210> 135

<211> 40

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic
primer

<400> 135

ccttggcacc cgagaattcc atgtccggga acacaaagac 40

<210> 136

<211> 39

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic
primer

<400> 136

ccttggcacc cgagaattcc atggcagcca ggaacgtac 39

<210> 137

<211> 40

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic
primer

<400> 137

ccttggcacc cgagaattcc aaacaccgca gcatgtcaag 40

<210> 138

<211> 41

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

primer

<400> 138

ccttggcacc cgagaattcc aaagtggatg gcattggaat c 41

<210> 139

<211> 40

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

primer

<400> 139

ccttggcacc cgagaattcc acctggagaa aggagaacgc 40

<210> 140

<211> 42

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

primer

<400> 140

ccttggcacc cgagaattcc aagttccgtg agttgatcat cg 42

<210> 141

<211> 41

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

primer

<400> 141

ccttggcacc cgagaattcc aacttctacc gtgccctgat g 41

<210> 142

<211> 40

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

primer

<400> 142

ccttggcacc cgagaattcc acacagcagg gcttcttcag 40

<210> 143

<211> 42

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

primer

<400> 143

ccttggcacc cgagaattcc acatgggcaa cttctctgtt tc 42

<210> 144

<211> 41

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

primer

<400> 144

ccttggcacc cgagaattcc aaaactgatg ggaccacctc c 41

<210> 145

<211> 45

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

primer

<400> 145

ccttggcacc cgagaattcc aaggagatat caagaggatg gattc 45

<210> 146

<211> 42

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic
primer

<400> 146

ccttggcacc cgagaattcc atcctgcaga aagacttgaa gg

42

<210> 147

<211> 46

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic
primer

<400> 147

ccttggcacc cgagaattcc aggattcaaa gcataaaaac cattac

46

<210> 148

<211> 49

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic
primer

<400> 148

ccttggcacc cgagaattcc atatgttgta taacttaaac ccgatagac

49

<210> 149

<211> 42

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic
primer

<400> 149

ccttggcacc cgagaattcc attgaagacc ataaccacc ac

42

<210> 150

<211> 45

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

primer

<400> 150

ccttggcacc cgagaattcc aaagtaagga ccagagacaa aaagg 45

<210> 151

<211> 43

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

primer

<400> 151

ccttggcacc cgagaattcc aggattatag accagtggca ctg 43

<210> 152

<211> 44

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

primer

<400> 152

ccttggcacc cgagaattcc acatgtactt tgagttcct cagc 44

<210> 153

<211> 42

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

primer

<400> 153

ccttggcacc cgagaattcc acaggaccag aggaaacctc ag 42

<210> 154

<211> 46

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

primer

<400> 154
ccttggcacc cgagaattcc acgtgcagat aatgacaagg aatatc 46
<210> 155
<211> 47
<212> DNA
<213> Artificial Sequence
<220><223> Description of Artificial Sequence: Synthetic
primer
<400> 155
ccttggcacc cgagaattcc aggtcagtta aattaaacat tttgtgg 47
<210> 156
<211> 46
<212> DNA
<213> Artificial Sequence
<220><223> Description of Artificial Sequence: Synthetic
primer
<400> 156
ccttggcacc cgagaattcc agatgttagt gacaatgaac ctgac 46
<210> 157
<211> 39
<212> DNA
<213> Artificial Sequence
<220><223> Description of Artificial Sequence: Synthetic
primer
<400> 157
ccttggcacc cgagaattcc atcttgccta cgccaccag 39
<210> 158
<211> 41
<212> DNA
<213> Artificial Sequence
<220><223> Description of Artificial Sequence: Synthetic
primer
<400> 158

ccttggcacc cgagaattcc acaggccctt ctgtcttgaa c 41

<210> 159

<211> 41

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

primer

<400> 159

ccttggcacc cgagaattcc agaacatctc gaagcgtca c 41

<210> 160

<211> 47

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

primer

<400> 160

ccttggcacc cgagaattcc attatggat aagttggtgt tctgaag 47

<210> 161

<211> 45

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

primer

<400> 161

ccttggcacc cgagaattcc aagaggtccc aagacttagt acctg 45

<210> 162

<211> 42

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

primer

<400> 162

ccttggcacc cgagaattcc agcttgctta cctcgcttag tg 42

<210> 163

<211> 43

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic
primer

<400> 163

ccttggcacc cgagaattcc attccgtccc agtagattac cac 43

<210> 164

<211> 44

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic
primer

<400> 164

ccttggcacc cgagaattcc acatgtagtt gtagtggatg gtgg 44

<210> 165

<211> 42

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic
primer

<400> 165

ccttggcacc cgagaattcc aatactccac acgcaaattt cc 42

<210> 166

<211> 41

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic
primer

<400> 166

ccttggcacc cgagaattcc actatctgag cagcgctcat g 41

<210> 167

<211> 39

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic
primer

<400> 167

ccttggcacc cgagaattcc agcaggtctt ggccagttg 39

<210> 168

<211> 41

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic
primer

<400> 168

ccttggcacc cgagaattcc atgtcccaga atgcaagaag c 41

<210> 169

<211> 39

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic
primer

<400> 169

ccttggcacc cgagaattcc agatgacagg ggccaggag 39

<210> 170

<211> 42

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic
primer

<400> 170

ccttggcacc cgagaattcc atgggtcttc agtgaaccat tg 42

<210> 171

<211> 42

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic
primer

<400> 171

ccttggcacc cgagaattcc agagcagaaa gtcagtccca tg 42

<210> 172

<211> 41

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic
primer

<400> 172

ccttggcacc cgagaattcc agctcgacgc taggatctga c 41

<210> 173

<211> 38

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic
oligonucleotide

<400> 173

ccatttcatt acctctttct ccgcaccgga catagatt 38

<210> 174

<211> 38

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic
oligonucleotide

<400> 174

atctatgtcg ggtgcggaga aagaggtaat gaaatggt 38