



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2024-0015624
(43) 공개일자 2024년02월05일

- (51) 국제특허분류(Int. Cl.)
C12Q 1/6886 (2018.01) G16B 20/00 (2019.01)
G16B 30/10 (2019.01) G16H 50/30 (2018.01)
- (52) CPC특허분류
C12Q 1/6886 (2022.01)
G16B 20/00 (2019.02)
- (21) 출원번호 10-2023-7035747
- (22) 출원일자(국제) 2022년04월07일
심사청구일자 없음
- (85) 번역문제출일자 2023년10월18일
- (86) 국제출원번호 PCT/US2022/023907
- (87) 국제공개번호 WO 2022/216981
국제공개일자 2022년10월13일
- (30) 우선권주장
63/172,493 2021년04월08일 미국(US)

- (71) 출원인
델피 다이아그노스틱스, 인코포레이티드
미국, 메릴랜드 21211, 볼티모어, 301 웨스트 29 스트리트
- (72) 발명자
드레이코폴리 니콜라스 씨.
미국, 메릴랜드 21211, 볼티모어, 301 웨스트 29 스트리트, 씨/오 델피 다이아그노스틱스, 인코포레이티드
릴 알레산드로
미국, 메릴랜드 21211, 볼티모어, 301 웨스트 29 스트리트, 씨/오 델피 다이아그노스틱스, 인코포레이티드
캐리 제이콥
미국, 메릴랜드 21211, 볼티모어, 301 웨스트 29 스트리트, 씨/오 델피 다이아그노스틱스, 인코포레이티드
- (74) 대리인
특허법인한얼

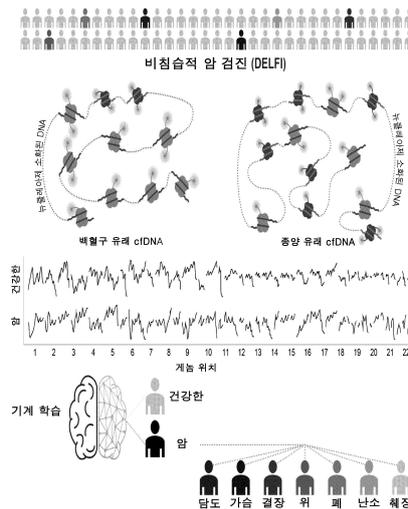
전체 청구항 수 : 총 59 항

(54) 발명의 명칭 **계놈 전체 cfDNA 단편화 프로파일을 이용한 암을 검출하는 방법**

(57) 요약

본 개시는 암 상태를 진단하고 예측하기 위해 환자로부터 얻은 샘플에서 무세포 DNA(cfDNA) 단편의 분석을 활용하는 방법 및 시스템을 제공한다. 본 개시는 대상체에서 암을 검출하는 방법을 제공한다. 본 개시는 또한 암이 있는 대상체의 전체 생존을 결정하는 방법을 제공한다. 본 개시는 대상체에서 암을 모니터링하는 방법을 추가로 제공한다. 또한 유전적 분석을 위한 시스템이 제공된다.

대표도 - 도1



(52) CPC특허분류

G16B 30/10 (2019.02)

G16H 50/30 (2018.01)

C12Q 2537/165 (2013.01)

C12Q 2600/118 (2013.01)

명세서

청구범위

청구항 1

대상체에서 암을 검출하는 방법으로서, 다음을 포함하는, 방법:

a) 대상체로부터의 샘플의 무세포 DNA(cfDNA) 단편화 프로파일을 결정하는 단계로서, cfDNA 단편화 프로파일은 다음에 의해 결정된다:

대상체로부터의 cfDNA 단편을 얻어 분리하는 단계,

cfDNA 단편을 서열 분석하여 서열 분석된 단편을 얻는 단계,

서열 분석된 단편을 게놈에 매핑하여 매핑된 서열의 윈도우를 얻는 단계, 및

매핑된 서열의 윈도우를 분석하여 cfDNA 단편 길이를 결정하고 cfDNA 단편화 프로파일을 생성하는 단계; 및

b) cfDNA 단편화 프로파일을 기반으로 대상체에서 암의 존재 가능성을 나타내는 점수를 계산하여 대상체를 암이 있거나 암이 없는 것으로 분류함으로써 대상체에서 암을 검출하는 단계.

청구항 2

제1항에 있어서, 점수를 계산하는 단계는 다음을 포함하는, 방법: i) 짧은 cfDNA 단편 대 긴 cfDNA 단편의 비율을 결정하는 단계, ii) 염색체 팔(arm)에 의한 cfDNA 단편에 대한 Z-점수를 결정하는 단계, iii) 전산 혼합 모델 분석을 사용하여 cfDNA 단편 밀도를 정량화하는 단계 및 iv) 기계 학습 모델을 사용하여 i)-iii)의 출력을 처리하여 점수를 정의하는 단계.

청구항 3

제2항에 있어서, 점수는 0 내지 1의 범위를 갖는, 방법.

청구항 4

제3항에 있어서, 대상체에서 암의 존재 가능성은 점수 값의 증가에 따라 증가하는, 방법.

청구항 5

제4항에 있어서, 암이 있는 것으로 분류된 대상체에 대하여, 점수를 기반으로 대상체의 전체 생존 가능성을 결정하는 단계를 추가로 포함하는, 방법.

청구항 6

제5항에 있어서, 대상체의 전체 생존 가능성은 점수 값의 증가에 따라 감소하는, 방법.

청구항 7

제6항에 있어서, 점수를 높은 점수 또는 낮은 점수로 분류하는 단계를 추가로 포함하고, 여기서 높은 점수는 0.5 초과 값을 갖고, 낮은 점수는 0.5 미만의 값을 가지며, 높은 점수는 대상체의 감소된 전체 생존을 나타내는, 방법.

청구항 8

제1항에 있어서, 서열 분석하는 단계는 cfDNA 단편을 낮은 적용 범위의 전체 게놈 서열분석에 적용하여 서열 분석된 단편을 얻는 단계를 포함하는, 방법.

청구항 9

제1항에 있어서, cfDNA 단편을 분리하는 단계는 105bp 미만 및 170bp 초과 단편 크기를 제외하는 단계를 포함

하는, 방법.

청구항 10

제1항에 있어서, 매핑된 서열의 윈도우는 수십 내지 수천 개의 윈도우를 포함하는, 방법.

청구항 11

제10항에 있어서, 윈도우는 비중첩 윈도우인, 방법.

청구항 12

제11항에 있어서, 윈도우 각각은 약 5백만 개의 염기쌍을 포함하는, 방법.

청구항 13

제12항에 있어서, cfDNA 단편화 프로파일은 각 윈도우 내에서 결정되는, 방법.

청구항 14

제1항에 있어서, cfDNA 단편화 프로파일은 매핑된 서열의 윈도우에서 작은 cfDNA 단편 대 큰 cfDNA 단편의 비율을 포함하는, 방법.

청구항 15

제1항에 있어서, cfDNA 단편화 프로파일은 게놈 전체의 윈도우에서 작은 cfDNA 단편 및 큰 cfDNA 단편의 서열 적용 범위를 포함하는, 방법.

청구항 16

제1항에 있어서, cfDNA 단편화 프로파일은 전체 게놈에 걸쳐 있는, 방법.

청구항 17

제1항에 있어서, cfDNA 단편화 프로파일은 서브게놈 간격에 걸쳐 있는, 방법.

청구항 18

제1항에 있어서, 분류하는 단계는 cfDNA 단편화 프로파일을 기준 cfDNA 단편화와 비교하는 단계를 포함하는, 방법.

청구항 19

제18항에 있어서, 기준 cfDNA 단편화 프로파일은 건강한 대상체의 cfDNA 단편화 프로파일인, 방법.

청구항 20

제1항에 있어서, 암은 고형 종양인, 방법.

청구항 21

제20항에 있어서, 암은 육종, 암종 또는 림프종인, 방법.

청구항 22

제20항에 있어서, 암은 다음으로 구성되는 군으로부터 선택되는 것인, 방법: 대장암, 전립선암, 유방암, 췌장암, 담관암, 간암, CNS, 위암, 식도암, 위장관 간질 종양(GIST), 자궁암 또는 난소암.

청구항 23

제1항에 있어서, 암은 혈액암인, 방법.

청구항 24

제23항에 있어서, 암은 다음으로 구성되는 군으로부터 선택되는 것인, 방법: 골수종, 다발성 골수종, B 세포 림프종, 여포성 림프종, 림프구성 백혈병, 백혈병 또는 골수성 백혈병.

청구항 25

제1항에 있어서, 대상체에게 암 치료를 투여하는 단계를 추가로 포함하는, 방법.

청구항 26

제25항에 있어서, 암 치료는 수술, 보조 화학치료요법, 신보조 화학치료요법, 방사선 치료요법, 호르몬 치료요법, 세포독성 치료요법, 면역치료요법, 입양 T 세포 치료요법, 표적화 치료요법 또는 이들의 임의의 조합으로 이루어진 군으로부터 선택되는, 방법.

청구항 27

암이 있는 대상체의 전체 생존을 결정하는 방법으로서, 다음을 포함하는, 방법:

- a) 대상체로부터의 샘플의 무세포 DNA(cfDNA) 단편화 프로파일을 결정하는 단계;
- b) cfDNA 단편화 프로파일을 기반으로 점수를 계산하는 단계로서, 여기서 점수를 계산하는 단계는 다음을 포함한다: i) 샘플의 짧은 cfDNA 단편 대 긴 cfDNA 단편의 비율을 결정하는 단계, ii) 염색체 팔에 의해 샘플의 cfDNA 단편에 대한 Z-점수를 결정하는 단계, iii) 전산 혼합 모델 분석을 사용하여 cfDNA 단편 밀도를 정량화하는 단계 및 iv) 기계 학습 모델을 사용하여 i)-iii)의 출력을 처리하여 점수를 정의하는 단계; 및
- c) 점수를 기반으로 대상체의 전체 생존 가능성을 결정함으로써 대상체의 전체 생존을 결정하는 단계.

청구항 28

제27항에 있어서, 점수는 0 내지 1의 범위를 갖는, 방법.

청구항 29

제28항에 있어서, 대상체의 전체 생존 가능성은 점수 값의 증가에 따라 감소하는, 방법.

청구항 30

제29항에 있어서, 점수를 높은 점수 또는 낮은 점수로 분류하는 단계를 추가로 포함하고, 여기서 높은 점수는 0.5 초과의 값을 갖고, 낮은 점수는 0.5 미만의 값을 가지며, 높은 점수는 대상체의 감소된 전체 생존을 나타내는, 방법.

청구항 31

제27항에 있어서, cfDNA 단편화 프로파일은 다음에 의해 결정되는, 방법:

- 대상체로부터의 cfDNA 단편을 얻어 분리하는 단계,
- cfDNA 단편을 서열 분석하여 서열 분석된 단편을 얻는 단계,
- 서열 분석된 단편을 게놈에 매핑하여 매핑된 서열의 윈도우를 얻는 단계, 및
- 매핑된 서열의 윈도우를 분석하여 cfDNA 단편 길이를 결정하고 cfDNA 단편화 프로파일을 생성하는 단계.

청구항 32

제31항에 있어서, 서열 분석하는 단계는 cfDNA 단편을 낮은 적용 범위의 전체 게놈 서열분석에 적용하여 서열 분석된 단편을 얻는 단계를 포함하는, 방법.

청구항 33

제31항에 있어서, cfDNA 단편을 분리하는 단계는 105bp 미만 및 170bp 초과의 단편 크기를 제외하는 단계를 포함하는, 방법.

청구항 34

제31항에 있어서, 매핑된 서열의 윈도우는 수십 내지 수천 개의 윈도우를 포함하는, 방법.

청구항 35

제34항에 있어서, 윈도우는 비중첩 윈도우인, 방법.

청구항 36

제35항에 있어서, 윈도우 각각은 약 5백만 개의 염기쌍을 포함하는, 방법.

청구항 37

제36항에 있어서, cfDNA 단편화 프로파일은 각 윈도우 내에서 결정되는, 방법.

청구항 38

제31항에 있어서, cfDNA 단편화 프로파일은 매핑된 서열의 윈도우에서 작은 cfDNA 단편 대 큰 cfDNA 단편의 비율을 포함하는, 방법.

청구항 39

제31항에 있어서, cfDNA 단편화 프로파일은 게놈 전체의 윈도우에서 작은 cfDNA 단편 및 큰 cfDNA 단편의 서열 적용 범위를 포함하는, 방법.

청구항 40

제31항에 있어서, cfDNA 단편화 프로파일은 전체 게놈에 걸쳐 있는, 방법.

청구항 41

제31항에 있어서, cfDNA 단편화 프로파일은 서브게놈 간격에 걸쳐 있는, 방법.

청구항 42

제27항에 있어서, 암은 고형 종양인, 방법.

청구항 43

제42항에 있어서, 암은 육종, 암종 또는 림프종인, 방법.

청구항 44

제42항에 있어서, 암은 다음으로 구성되는 군으로부터 선택되는 것인, 방법: 대장암, 전립선암, 유방암, 췌장암, 담관암, 간암, CNS, 위암, 식도암, 위장관 간질 종양(GIST), 자궁암 또는 난소암.

청구항 45

제27항에 있어서, 암은 혈액암인, 방법.

청구항 46

제45항에 있어서, 암은 다음으로 구성되는 군으로부터 선택되는 것인, 방법: 골수종, 다발성 골수종, B 세포 림프종, 여포성 림프종, 림프구성 백혈병, 백혈병 또는 골수성 백혈병.

청구항 47

제27항에 있어서, 대상체에게 암 치료를 투여하는 단계를 추가로 포함하는, 방법.

청구항 48

제47항에 있어서, 암 치료는 수술, 보조 화학치료요법, 신보조 화학치료요법, 방사선 치료요법, 호르몬 치료요법, 세포독성 치료요법, 면역치료요법, 입양 T 세포 치료요법, 표적화 치료요법 또는 이들의 임의의 조합으로 이루어진 군으로부터 선택되는, 방법.

청구항 49

암이 있는 대상체를 치료하는 방법으로서, 다음을 포함하는, 방법:

- a) 제1항 내지 제19항 중 어느 한 항의 방법을 사용하여 대상체에서 암을 검출하거나, 제27항 내지 제41항 중 어느 한 항의 방법을 사용하여 대상체의 전체 생존을 결정하는 단계; 및
- b) 대상체에게 암 치료를 투여하여 대상체를 치료하는 단계.

청구항 50

제49항에 있어서, 암은 고형 종양인, 방법.

청구항 51

제50항에 있어서, 암은 육종, 암종 또는 림프종인, 방법.

청구항 52

제50항에 있어서, 암은 다음으로 구성되는 군으로부터 선택되는 것인, 방법: 폐암, 대장암, 전립선암, 유방암, 췌장암, 담관암, 간암, CNS, 위암, 식도암, 위장관 간질 종양(GIST), 자궁암 또는 난소암.

청구항 53

제49항에 있어서, 암은 혈액암인, 방법.

청구항 54

제53항에 있어서, 암은 다음으로 구성되는 군으로부터 선택되는 것인, 방법: 골수종, 다발성 골수종, B 세포 림프종, 여포성 림프종, 림프구성 백혈병, 백혈병 또는 골수성 백혈병.

청구항 55

제49항에 있어서, 암 치료는 수술, 보조 화학치료요법, 신보조 화학치료요법, 방사선 치료요법, 호르몬 치료요법, 세포독성 치료요법, 면역치료요법, 입양 T 세포 치료요법, 표적화 치료요법 또는 이들의 임의의 조합으로 이루어진 군으로부터 선택되는, 방법.

청구항 56

제47항에 있어서, 대상체는 인간인, 방법.

청구항 57

대상체에서 암을 모니터링하는 방법으로서, 다음을 포함하는, 방법:

- a) 제1항 내지 제19항 중 어느 한 항의 방법을 사용하여 대상체에서 암을 검출하거나 제27항 내지 제41항 중 어느 한 항의 방법을 사용하여 대상체의 전체 생존을 결정하는 단계;
- b) 대상체에게 암 치료를 투여하는 단계; 및
- c) 암 치료가 투여된 후 제27항 내지 제41항 중 어느 한 항의 방법을 사용하여 대상체의 전체 생존을 결정함으로써 대상체에서 암을 모니터링하는 단계.

청구항 58

컴퓨터 프로그램으로 암호화된 비밀시적 컴퓨터 판독 가능 저장 매체로서, 프로그램은 하나 이상의 프로세서에 의해 실행될 때 하나 이상의 프로세서가 제1항 내지 제24항 또는 제27항 내지 제46항 중 어느 한 항의 방법을 수행하기 위한 동작을 수행하게 하는 명령을 포함하는, 비밀시적 컴퓨터 판독 가능 저장 매체.

청구항 59

컴퓨팅 시스템으로서, 메모리; 및 메모리에 연결된 하나 이상의 프로세서를 포함하고, 하나 이상의 프로세서는 제1항 내지 제24항 또는 제27항 내지 제46항 중 어느 한 항의 방법을 수행하는 동작을 수행하도록 구성되는, 컴

퓨팅 시스템.

발명의 설명

기술 분야

[0001] **관련 출원(들)에 대한 상호 참조**

[0002] 본 출원은 2021년 4월 8일에 출원된 미국 특허 가출원 일련 번호 63/172,493의 35 U.S.C. § 119(e)에 따른 우선권의 이익을 주장한다. 선행 출원의 개시는 본 출원의 개시의 일부로 간주되며 참조로 포함된다.

[0003] **발명의 기술분야**

[0004] 본 발명은 일반적으로 유전적 분석에 관한 것이며, 더욱 구체적으로는 대상체에서 암을 검출하고/검출하거나 대상체의 전체 생존을 평가하기 위한 무세포 DNA(cfDNA) 단편의 분석을 위한 방법 및 시스템에 관한 것이다.

배경 기술

[0005] 전 세계적으로 인간 암의 이환율과 사망률의 대부분은 치료가 덜 효과적인 이들 질병의 늦은 진단의 결과이다. 불행하게도, 조기 암 환자를 광범위하게 진단하고 치료하는 데 사용될 수 있는 임상적으로 입증된 바이오마커는 널리 이용 가능하지 않다.

[0006] 무세포 DNA(cfDNA) 분석은 이러한 접근법이 조기 진단 및 치료를 위한 새로운 길을 제공할 수 있음을 시사한다. 순환 종양 DNA(ctDNA) 단편은 비종양 세포로부터의 다른 cfDNA보다 평균적으로 더 짧은 것으로 나타났다. 기존의 작업에서는 단편을 히스톤 코어 또는 링커 단백질(예를 들어, 짧고 긴 또는 상호 배타적인 크기 세트)에 결합하여 발생하는 상이한 크기의 군으로 분리하고 이들 단편의 수를 사용하여 ctDNA를 정량화 하고/하거나 개별 샘플을 종양의 존재/부재로 분류하여 탐구하였다. 그러나 기존의 연구들은 암 진단을 받은 환자의 전체 생존을 결정하는 능력뿐만 아니라 암 검출에 있어 강력한 민감도 및 특이성을 제공하는 능력이 부족하였다.

발명의 내용

[0007] 본 개시는 대상체로부터 얻은 샘플에서 cfDNA 단편의 분석을 통해 얻은 cfDNA 단편화 프로파일을 점수화함으로써 대상체의 전체 생존을 검출하고 예측하기 위한 cfDNA 분석을 활용하는 방법 및 시스템을 제공한다. 점수화 방법론은 대상체의 전체 생존 가능성의 척도를 제공한다.

[0008] 따라서, 일 구현예에서, 본 발명은 대상체에서 암을 검출하는 방법을 제공한다. 방법은 다음을 포함한다:

[0009] a) 대상체로부터의 샘플의 무세포 DNA(cfDNA) 단편화 프로파일을 결정하는 단계로서, cfDNA 단편화 프로파일은 다음에 의해 결정된다:

[0010] 대상체로부터의 cfDNA 단편을 얻어 분리하는 단계,

[0011] cfDNA 단편을 서열 분석하여 서열 분석된 단편을 얻는 단계,

[0012] 서열 분석된 단편을 게놈에 매핑하여 매핑된 서열의 윈도우를 얻는 단계, 및

[0013] 매핑된 서열의 윈도우를 분석하여 cfDNA 단편 길이를 결정하고 cfDNA 단편화 프로파일을 생성하는 단계; 및

[0014] b) cfDNA 단편화 프로파일을 기반으로 점수를 계산하여 대상체에서 암의 존재 가능성을 나타내는 점수를 계산하여 대상체를 암이 있거나 암이 없는 것으로 분류함으로써 대상체에서 암을 검출하는 단계. 일부 양태에서, 암은 폐암을 제외한다. 일부 양태에서, 화학치료요법제, 방사선, 면역치료요법 또는 기타 치료적 요법이 대상체에게 투여된다.

[0015] 일부 양태에서, 점수를 계산하는 단계는 다음을 포함한다: i) 짧은 cfDNA 단편 대 긴 cfDNA 단편의 비율을 결정하는 단계, ii) 염색체 팔에 의한 cfDNA 단편에 대한 Z-점수를 결정하는 단계, iii) 전산 혼합 모델 분석을 사용하여 cfDNA 단편 밀도를 정량화하는 단계 및 iv) 기계 학습 모델을 사용하여 i)-iii)의 출력을 처리하여 점수를 정의하는 단계.

[0016] 또 다른 구현예에서, 본 발명은 암이 있는 대상체의 전체 생존을 결정하는 방법을 제공한다. 방법은 다음을 포함한다:

- [0017] a) 대상체로부터의 샘플의 무세포 DNA(cfDNA) 단편화 프로파일을 결정하는 단계;
- [0018] b) cfDNA 단편화 프로파일을 기반으로 점수를 계산하는 단계로서, 여기서 점수를 계산하는 단계는 다음을 포함한다: i) 샘플의 짧은 cfDNA 단편 대 긴 cfDNA 단편의 비율을 결정하는 단계, ii) 염색체 팔에 의해 샘플의 cfDNA 단편에 대한 Z-점수를 결정하는 단계, iii) 전산 혼합 모델 분석을 사용하여 cfDNA 단편 밀도를 정량화하는 단계 및 iv) 기계 학습 모델을 사용하여 i)-iii)의 출력을 처리하여 점수를 정의하는 단계; 및
- [0019] c) 점수를 기반으로 대상체의 전체 생존 가능성을 결정함으로써 대상체의 전체 생존을 결정하는 단계.
- [0020] 또 다른 양태에서, 본 발명은 암이 있는 대상체를 치료하는 방법을 제공한다. 방법은 다음을 포함한다:
- [0021] a) 본 발명의 방법론을 사용하여 대상체에서 암을 검출하거나, 본 발명의 방법론을 사용하여 대상체의 전체 생존을 결정하는 단계; 및
- [0022] b) 대상체에게 암 치료를 투여하여 대상체를 치료하는 단계. 일부 양태에서, 화학치료요법제, 방사선, 면역치료요법 또는 기타 치료적 요법이 대상체에게 투여된다.
- [0023] 또 다른 구현예에서, 본 발명은 대상체에서 암을 모니터링하는 방법을 제공한다. 방법은 다음을 포함한다:
- [0024] a) 본 발명의 방법론을 사용하여 대상체에서 암을 검출하고/검출하거나 본 발명의 방법론을 사용하여 대상체의 전체 생존을 결정하는 단계;
- [0025] b) 대상체에게 암 치료를 투여하는 단계; 및
- [0026] c) 암 치료가 투여된 후 본 발명의 방법론을 사용하여 대상체의 전체 생존을 결정함으로써 대상체에서 암을 모니터링하는 단계. 일부 양태에서, 화학치료요법제, 방사선, 면역치료요법 또는 기타 치료적 요법이 대상체에게 투여된다.
- [0027] 또 다른 구현예에서, 본 발명은 컴퓨터 프로그램으로 암호화된 비밀시적 컴퓨터 판독 가능 저장 매체를 제공한다. 컴퓨터 프로그램은 하나 이상의 프로세서에 의해 실행될 때 하나 이상의 프로세서가 본 발명의 방법을 수행하기 위한 동작을 수행하게 하는 명령을 포함한다.
- [0028] 또 다른 구현예에서, 본 발명은 컴퓨팅 시스템을 제공한다. 시스템은 메모리, 및 메모리에 연결된 하나 이상의 프로세서를 포함하며, 하나 이상의 프로세서는 본 발명의 방법을 구현하는 동작을 수행하도록 구성된다.
- [0029] 또 다른 구현예에서, 본 발명은 유전적 분석 및 암을 평가하기 위한 시스템을 제공하며 이는 다음을 포함한다: (a) 샘플에 대한 전체 게놈 서열 분석(WGS) 데이터 세트를 생성하도록 구성된 서열 분석기; 및 (b) 비밀시적 컴퓨터 판독 가능 저장 매체 및/또는 본 발명의 컴퓨터 시스템.

도면의 간단한 설명

- [0030] 도 1 은 본 발명의 일 구현예에서 본 개시의 방법론을 사용하는 예시적인 DELFI 접근법을 도시하는 개략도이다. 혈액은 건강한 개인과 암이 있는 환자 코호트로부터 수집된다. cfDNA는 혈장 분획에서 추출되어 서열 분석 라이브러리로 처리되고, 전체 게놈 서열 분석을 통해 검사되고, 게놈에 매핑되고, 분석되어 게놈 전체에 걸쳐 cfDNA 단편화 프로파일을 결정한다. 기계 학습 접근법을 사용하여 DELFI 점수를 생성하고 개인을 건강한 또는 암이 있는 것으로 분류한다.
- 도 2는 암의 비침습적 검출을 위한 cfDNA 단편화 검정의 성과를 나타내는 표이다. 포함 후 3개월 이내에 74명의 환자가 16개의 상이한 고탄암 중 1개로 진단받았고, 207명의 환자는 암이 없었다.
- 도 3은 본 발명의 일 구현예에서 본 개시의 방법론을 사용하여 생성된 데이터를 나타내는 그래픽 플롯이다. 그래프는 암 검출을 위한 cfDNA 단편화 검정의 전체 성과를 나타낸다.
- 도 4는 본 발명의 일 구현예에서 본 개시의 방법론을 사용하여 생성된 데이터를 나타내는 그래픽 플롯이다. 그래프는 DELFI 점수와 상관관계가 있는 대상체의 생존을 나타낸다. DELFI 점수가 높을수록 암 병기나 기타 임상적 특성과 관계없이 감소된 전체 생존과 연관이 있었다.
- 도 5는 본 발명의 일 구현예에서 개시된 방법론을 사용하여 생성된 데이터 곡선을 나타내는 일련의 그래픽 플롯이다. 계산된 DELFI 점수는 높은 점수(>0.5) 대 낮은 점수(<0.5)를 정의하는 데 사용된 컷오프 값에 관계없이 암(폐암 제외)이 있는 개인의 묘사된 카플란-마이어 곡선을 분리한다. 각 패널 상단의 숫자는 결정된 컷오프 값

을 가리킨다.

도 6은 본 발명의 일 구현예에서 본 개시의 방법론을 사용하여 생성된 데이터를 나타내는 그래픽 플롯이다. 도 6은 두 가지 설정에서 콕스 비례 위험 모델의 결과를 나타낸다. 첫 번째 설정(플롯의 왼쪽 패널)에서 DELFI 점수는 연속적으로 처리된다. 두 번째 설정(플롯의 오른쪽 패널)에서 DELFI 점수는 높음(>0.5) 또는 낮음(<0.5)으로 처리된다. 두 설정 모두에서 DELFI 점수는 재혈 및 병기에서 나이를 조정하더라도 생존에 대한 강력한 예측 변수이다. 병기는 1기에 관한 것임을 유념한다.

발명을 실시하기 위한 구체적인 내용

- [0031] 암의 조기 검출뿐만 아니라 암이 있는 대상체의 전체 생존의 예측을 위한 비침습적 방법이 본원에 설명되어 있다. 혈액 내 cfDNA는 암 환자에게 비침습적 진단 방안을 제공할 수 있다. 본원에서 입증된 바와 같이, 조기 차단을 위한 단편의 DNA 감정(DELFI)을 사용하여 다양한 유형의 암 환자뿐만 아니라 건강한 개인의 cfDNA의 게놈 전체 단편화 패턴을 감정하였다. cfDNA 감정에는 점수화 방법론이 포함되었다. 전체 생존과 상관관계가 있는 주어진 환자 샘플의 cfDNA 단편을 사용하여 얻은 cfDNA 단편화 프로파일에 대해 정의된 점수(본원에서는 'DELFI 점수'라고도 지칭됨)를 결정하였다. 본원에 설명된 방법론을 사용하여 cfDNA를 평가하면 암의 조기 발견 및 평가를 위한 검진 접근법을 제공할 수 있으며, 이는 암이 있는 환자의 성공적인 치료 기회를 증가시킬 수 있다. cfDNA 평가는 또한 암 모니터링을 위한 접근법을 제공할 수 있으며, 이는 성공적인 치료 및 암이 있는 환자의 개선된 결과를 위한 기회를 증가시킬 수 있다.
- [0032] 본 조성물 및 방법을 설명하기 전에, 본 발명은 설명된 특정 방법 및 시스템에 제한되지 않으며, 이러한 방법 및 시스템은 다양할 수 있다는 것이 이해되어야 한다. 본원에서 사용된 전문 용어는 단지 특정 구현예를 설명하기 위한 것이며, 제한하려는 의도가 아닌 것으로 또한 이해되어야 하는데, 이는 본 발명의 범위가 첨부된 청구항에만 제한될 것이기 때문이다.
- [0033] 본 명세서 및 첨부된 청구항에 사용된 바와 같이, 단수형 "a", "an" 및 "the"는 문맥상 명백하게 달리 지시하지 않는 한 복수 지칭을 포함한다. 따라서, 예를 들어, "방법"에 대한 지칭은 본 개시 등을 읽을 때 당업자에게 명백해질 본원에 설명된 유형의 하나 이상의 방법 및/또는 단계를 포함한다.
- [0034] 다르게 정의되지 않는 한, 본원에 사용된 모든 기술 및 과학 용어는 본 발명이 속하는 기술 분야의 숙련자가 일반적으로 이해하는 것과 동일한 의미를 갖는다. 본원에 설명된 것과 유사하거나 등가인 임의의 방법 및 재료가 본 발명의 실시 또는 테스트에 사용될 수 있지만, 이제 바람직한 방법 및 재료가 설명된다.
- [0035] 본 개시는 암을 검출하거나 평가하기 위한 cfDNA의 분석을 위한 혁신적인 방법 및 시스템을 제공한다. 선행 연구에서 가리킨 바와 같이, 평균적으로 암이 없는 개인은 더 긴 cfDNA 단편(평균 크기 167.09bp)을 갖는 반면, 암이 있는 개인은 더 짧은 cfDNA 단편(평균 크기 164.88bp)을 갖는다. 본원에 설명된 방법론은 cfDNA 단편화 패턴의 게놈 전체 분석을 통해 cfDNA의 수많은 이상의 동시 분석을 가능케한다.
- [0036] 따라서, 일 구현예에서, 본 발명은 대상체에서 암을 검출하는 방법을 제공한다. 방법은 다음을 포함한다:
- [0037] a) 대상체로부터의 샘플의 무세포 DNA(cfDNA) 단편화 프로파일을 결정하는 단계; 및
- [0038] b) cfDNA 단편화 프로파일을 기반으로 대상체에서 암의 존재 가능성을 나타내는 점수를 계산하여 대상체를 암이 있거나 암이 없는 것으로 분류함으로써, 단 암은 폐암을 포함하지 않는다는 조건으로, 대상체에서 암을 검출하는 단계.
- [0039] 또 다른 구현예에서, 본 발명은 암이 있는 대상체의 전체 생존을 결정하는 방법을 제공한다. 방법은 다음을 포함한다:
- [0040] a) 대상체로부터의 샘플의 무세포 DNA(cfDNA) 단편화 프로파일을 결정하는 단계;
- [0041] b) cfDNA 단편화 프로파일을 기반으로 점수를 계산하는 단계로서, 여기서 점수를 계산하는 단계는 다음을 포함한다: i) 샘플의 짧은 cfDNA 단편 대 긴 cfDNA 단편의 비율을 결정하는 단계, ii) 염색체 팔에 의해 샘플의 cfDNA 단편에 대한 Z-점수를 결정하는 단계, iii) 전산 혼합 모델 분석을 사용하여 cfDNA 단편 밀도를 정량화하는 단계 및 iv) 기계 학습 모델을 사용하여 i)-iii)의 출력을 처리하여 점수를 정의하는 단계; 및
- [0042] c) 점수를 기반으로 대상체의 전체 생존 가능성을 결정함으로써 대상체의 전체 생존을 결정하는 단계.
- [0043] 구현예에서, 본 발명은 암이 있는 대상체를 치료하는 방법을 제공한다. 방법은 다음을 포함한다:

- [0044] a) 본 발명의 방법론을 사용하여 대상체에서 암을 검출하거나, 본 발명의 방법론을 사용하여 대상체의 전체 생존을 결정하는 단계; 및
- [0045] b) 대상체에게 암 치료를 투여하여 대상체를 치료하는 단계. 일부 양태에서, 화학치료요법제, 방사선, 면역치료요법 또는 기타 치료적 요법이 대상체에게 투여된다.
- [0046] 또 다른 구현예에서, 본 발명은 대상체에서 암을 모니터링하는 방법을 제공한다. 방법은 다음을 포함한다:
- [0047] a) 본 발명의 방법론을 사용하여 대상체에서 암을 검출하거나 본 발명의 방법론을 사용하여 대상체의 전체 생존을 결정하는 단계;
- [0048] b) 대상체에게 암 치료를 투여하는 단계; 및
- [0049] c) 암 치료가 투여된 후 본 발명의 방법론을 사용하여 대상체의 전체 생존을 결정함으로써 대상체에서 암을 모니터링하는 단계.
- [0050] 본원에 설명된 방법론은 cfDNA 단편화 프로파일을 활용한다. 본원에 사용된 바와 같이, 용어 "단편화 프로파일"은 일부 양태에서, 포유동물에서 cfDNA 단편화 프로파일을 결정하는 것은 포유동물이 암이 있는 것으로 식별하는 데 사용될 수 있다. 예를 들어, 포유동물로부터 얻은 cfDNA 단편(예를 들어, 포유동물로부터 얻은 샘플로부터)은 낮은 적용 범위의 전체 게놈 서열 분석을 적용 받을 수 있으며, 서열 분석된 단편은 게놈에 매핑될 수 있고(예를 들어, 비중첩 윈도우에서) cfDNA 단편화 프로파일을 결정하기 위해 평가될 수 있다. 암이 있는 포유동물의 cfDNA 단편화 프로파일은 건강한 포유동물(예를 들어, 암이 없는 포유동물)의 cfDNA 단편화 프로파일보다 더 이질적이다(예를 들어, 단편 길이).
- [0051] cfDNA 단편화 프로파일에는 하나 이상의 cfDNA 단편화 패턴이 포함될 수 있다. cfDNA 단편화 패턴은 임의의 적절한 cfDNA 단편화 패턴을 포함할 수 있다. cfDNA 단편화 패턴의 예에는 단편 크기 밀도, 중간 단편 크기, 단편 크기 분포, 작은 cfDNA 단편 대 큰 cfDNA 단편의 비율 및 cfDNA 단편의 적용 범위가 비제한적으로 포함된다. 일부 양태에서, cfDNA 단편화 프로파일은 게놈 전체 cfDNA 프로파일(예를 들어, 게놈 전체 윈도우의 게놈 전체 cfDNA 프로파일)일 수 있다. 일부 양태에서, cfDNA 단편화 프로파일은 표적화 영역 프로파일일 수 있다. 표적화 영역은 게놈의 임의의 적절한 부분(예를 들어 염색체 영역)일 수 있다. cfDNA 단편화 프로파일이 본원에 설명된 바와 같이 결정될 수 있는 염색체 영역의 예에는 염색체의 부분(예를 들어, 2q, 4p, 5p, 6q, 7p, 8q, 9q, 10q, 11q, 12q 및/또는 14q의 부분) 및 염색체 팔(예를 들어, 8q, 13q, 11q 및/또는 3p의 염색체 팔)을 비제한적으로 포함한다. 어떤 경우에는, cfDNA 단편화 프로파일에 두 개 이상의 표적화 영역 프로파일이 포함될 수 있다.
- [0052] 다양한 양태에서, 샘플에서 얻은 cfDNA가 분리되고 특정 크기 범위의 단편이 분석에 활용된다. 일부 양태에서, 분석에서는 약 10, 50, 100 또는 105bp 미만 및 약 220, 250, 300, 350bp 이상의 단편 크기를 제외한다. 일부 양태에서, 분석에서는 105bp 미만 및 170bp 초과 단편 크기를 제외한다. 일부 양태에서, 분석에서는 약 230, 240, 250, 260bp 미만 및 약 420, 430, 440, 450bp 이상의 단편 크기를 제외한다. 일부 양태에서 분석에서는 260bp 미만 및 440bp 초과 단편 크기를 제외한다.
- [0053] 일부 양태에서, cfDNA 단편화 프로파일은 다음에 의해 결정될 수 있다: cfDNA 단편을 포함하는 대상체로부터의 샘플을 서열 분석 라이브러리로 처리하는 단계; 서열 분석 라이브러리를 낮은 적용 범위의 전체 게놈 서열 분석에 적용하여 서열 분석된 단편을 얻는 단계; 서열 분석된 단편을 게놈에 매핑하여 매핑된 서열의 윈도우를 얻는 단계; 및 매핑된 서열의 윈도우를 분석하여 cfDNA 단편 길이를 결정하는 단계를 포함한다.
- [0054] 일부 양태에서, cfDNA 단편화 프로파일은 다음에 의해 결정될 수 있다: 대상체로부터 cfDNA 단편을 얻고 분리하는 단계, cfDNA 단편을 서열 분석하여 서열 분석된 단편을 얻는 단계, 서열 분석된 단편을 게놈에 매핑하여 매핑된 서열의 윈도우를 얻는 단계, 매핑된 서열의 윈도우를 분석하여 cfDNA 단편 길이를 결정하고 cfDNA 단편화 프로파일을 생성하는 단계.
- [0055] 본 발명의 방법론은 낮은 적용 범위의 전체 게놈 서열 분석 및 분리된 cfDNA의 분석을 기반으로 한다. 일 양태에서, 본 발명의 방법론을 개발하는 데 사용되는 데이터는 얇은 전체 게놈 서열 데이터(1-2x 적용 범위)를 기반으로 한다.
- [0056] 일부 양태에서, 매핑된 서열은 게놈이 적용되는 비중첩 윈도우에서 분석된다. 개념적으로 윈도우는 크기가 수천에서 수백만 염기 범위에 이를 수 있어, 게놈에 수백에서 수천 개의 윈도우를 야기할 수 있다. 5Mb 윈도우는 cfDNA 단편화 패턴을 감정하는 데 사용되고, 이는 제한된 양의 1-2x 게놈 적용 범위에서도 윈도우 당 20,000개 초과 판독을 제공하기 때문이다. 각 윈도우 내에서, cfDNA 단편의 적용 범위와 크기 분포를 검사하였다. 일부

양태에서, 개인으로부터의 전체 게놈 패턴을 기준 집단과 비교하여 패턴이 건강한 것인지 또는 암에서 유래될 가능성이 있는 것인지를 결정할 수 있다.

- [0057] 특정 양태에서, 매핑된 서열은 수십 내지 수천 개의 게놈 윈도우, 예컨대 10, 50, 100 내지 1,000, 5,000, 10,000개 이상의 윈도우를 포함한다. 이러한 윈도우는 비중첩되거나 중첩될 수 있으며 약 1, 2, 3, 4, 5, 6, 7, 8, 9백만 또는 1천만 개의 염기쌍을 포함할 수 있다.
- [0058] 다양한 양태에서, cfDNA 단편화 프로파일은 각 윈도우 내에서 결정된다. 따라서, 본 발명은 대상체에서(예를 들어, 대상체로부터 얻은 샘플에서) cfDNA 단편화 프로파일을 결정하기 위한 방법을 제공한다.
- [0059] 일부 양태에서, cfDNA 단편화 프로파일은 cfDNA 단편 길이의 변화(예를 들어 변경)를 식별하는 데 사용될 수 있다. 변경은 게놈 전체 변경이거나 하나 이상의 표적화 영역/좌위의 변경일 수 있다. 표적 영역은 하나 이상의 암 특이적 변경을 함유하는 임의의 영역일 수 있다. 일부 양태에서, cfDNA 단편화 프로파일은 약 10개 변경 내지 약 500개 변경(예를 들어, 약 25개 내지 약 500개, 약 50개 내지 약 500개, 약 100개 내지 약 500개, 약 200개 내지 약 500개, 약 300개 내지 약 500개, 약 10개 내지 약 400개, 약 10개 내지 약 300개, 약 10개 내지 약 200개, 약 10개 내지 약 100개, 약 10개 내지 약 50개, 약 20개 내지 약 400개, 약 30개 내지 약 300개, 약 40개 내지 약 200개, 약 50개 내지 약 100개, 약 20개 내지 약 100개, 약 25개 내지 약 75개, 약 50개 내지 약 250개 또는 약 100개 내지 약 200개의 변경)을 식별(예를 들어, 동시에 식별)하는 데 사용될 수 있다.
- [0060] 다양한 양태에서, cfDNA 단편화 프로파일은 cfDNA 단편 크기 패턴을 포함할 수 있다. cfDNA 단편은 임의의 적절한 크기일 수 있다. 예를 들어, 일부 양태에서, cfDNA 단편은 길이로 약 50개 염기쌍(bp) 내지 약 400bp일 수 있다. 본원에 설명된 바와 같이, 암이 있는 대상체는 건강한 대상체의 중앙 cfDNA 단편 크기보다 더 짧은 중앙 cfDNA 단편 크기를 함유하는 cfDNA 단편 크기 패턴을 가질 수 있다. 건강한 대상체(예를 들어, 암이 없는 대상체)는 약 166.6bp 내지 약 167.2bp(예를 들어, 약 166.9bp)의 중앙 cfDNA 단편 크기를 갖는 cfDNA 단편 크기를 가질 수 있다. 일부 양태에서, 암이 있는 대상체는 건강한 대상체에서 cfDNA 단편 크기보다 평균적으로 약 1.28bp 내지 약 2.49bp(예를 들어, 약 1.88bp) 더 짧은 cfDNA 단편 크기를 가질 수 있다. 예를 들어, 암이 있는 대상체는 약 164.11bp 내지 약 165.92bp(예를 들어, 약 165.02bp)의 중앙 cfDNA 단편 크기를 갖는 cfDNA 단편 크기를 가질 수 있다.
- [0061] 일부 양태에서, 디뉴클레오솜 cfDNA 단편은 길이로 약 230 염기쌍(bp) 내지 약 450bp일 수 있다. 본원에 설명된 바와 같이, 암이 있는 대상체는 건강한 대상체의 중앙 디뉴클레오솜 cfDNA 단편 크기보다 더 짧은 중앙 디뉴클레오솜 cfDNA 단편 크기를 함유하는 디뉴클레오솜 cfDNA 단편 크기 패턴을 가질 수 있다. 일부 양태에서, 평균적으로, 암이 없는 대상체는 디뉴클레오솜 범위(평균 크기 334.75bp)에서 더 긴 cfDNA 단편을 갖는 반면, 암이 있는 대상체는 더 짧은 디뉴클레오솜 cfDNA 단편(평균 크기 329.6bp)을 갖는다. 따라서, 건강한 대상체(예를 들어, 암이 없는 대상체)는 약 334.75bp의 중앙 cfDNA 단편 크기를 갖는 디뉴클레오솜 cfDNA 단편 크기를 가질 수 있다. 일부 양태에서, 암이 있는 대상체는 건강한 대상체의 디뉴클레오솜 cfDNA 단편 크기보다 더 짧은 디뉴클레오솜 cfDNA 단편 크기를 가질 수 있다. 예를 들어, 암이 있는 대상체는 중앙 cfDNA 단편 크기가 약 329.6bp인 디뉴클레오솜 cfDNA 단편 크기를 가질 수 있다.
- [0062] cfDNA 단편화 프로파일에는 cfDNA 단편 크기 분포가 포함될 수 있다. 본원에 설명된 바와 같이, 암이 있는 대상체는 건강한 대상체의 cfDNA 단편 크기 분포보다 더 가변적인 cfDNA 크기 분포를 가질 수 있다. 일부 양태에서, 크기 분포는 표적화 영역 내에 있을 수 있다. 건강한 대상체(예를 들어, 암이 없는 대상체)는 약 1 또는 약 1 미만의 표적화 영역 cfDNA 단편 크기 분포를 가질 수 있다. 일부 양태에서, 암이 있는 대상체는 건강한 대상체의 표적화 영역 cfDNA 단편 크기 분포보다 더 긴(예를 들어, 10, 15, 20, 25, 30, 35, 40, 45, 50bp 이상 더 길거나 이들 숫자 사이의 임의의 염기쌍 수) 표적화 영역 cfDNA 단편 크기 분포를 가질 수 있다. 일부 양태에서, 암이 있는 대상체는 건강한 대상체의 표적화 영역 cfDNA 단편 크기 분포보다 더 짧은(예를 들어, 10, 15, 20, 25, 30, 35, 40, 45, 50bp 이상 더 짧거나 이들 숫자 사이의 임의의 염기쌍 수) 표적화 영역 cfDNA 단편 크기 분포를 가질 수 있다. 일부 양태에서, 암이 있는 대상체는 건강한 대상체의 표적화 영역 cfDNA 단편 크기 분포보다 약 47bp 더 작거나 약 30bp 더 긴 표적화 영역 cfDNA 단편 크기 분포를 가질 수 있다. 일부 양태에서, 암이 있는 대상체는 cfDNA 단편의 길이가 평균적으로 10, 11, 12, 13, 14, 15, 15, 17, 18, 19, 20bp 이상 차이인 표적화 영역 cfDNA 단편 크기 분포를 가질 수 있다. 예를 들어, 암이 있는 대상체는 평균적으로 cfDNA 단편의 길이가 약 13bp 차이인 표적화 영역 cfDNA 단편 크기 분포를 가질 수 있다. 일부 양태에서, 크기 분포는 게놈 전체 크기 분포일 수 있다.
- [0063] cfDNA 단편화 프로파일에는 작은 cfDNA 단편 대 큰 cfDNA 단편의 비율과 단편 비율 대 기준 단편 비율의 상관관

계가 포함될 수 있다. 본원에 사용된 바와 같이, 작은 cfDNA 단편 대 큰 cfDNA 단편의 비율과 관련하여, 작은 cfDNA 단편은 길이가 약 100bp 내지 약 150bp일 수 있다. 본원에 사용된 바와 같이, 작은 cfDNA 단편 대 큰 cfDNA 단편의 비율과 관련하여, 큰 cfDNA 단편은 길이가 약 151bp 내지 220bp일 수 있다. 본원에 설명된 바와 같이, 암이 있는 대상체는 건강한 대상체보다 더 낮은(예를 들어, 2배 낮음, 3배 낮음, 4배 낮음, 5배 낮음, 6배 낮음, 7배 낮음, 8배 낮음, 9배 낮음, 10배 낮음 또는 그 이상) 단편 비율의 상관관계(예를 들어, cfDNA 단편 비율 대 기준 DNA 단편 비율, 예컨대 한 명 이상의 건강한 대상체로부터의 DNA 단편 비율의 상관관계)를 가질 수 있다. 건강한 대상체(예를 들어, 암이 없는 대상체)는 약 1(예를 들어, 약 0.96)의 단편 비율의 상관관계(예를 들어, cfDNA 단편 비율 대 기준 DNA 단편 비율, 예컨대 한 명 이상의 건강한 대상체로부터의 DNA 단편 비율의 상관관계)를 가질 수 있다. 일부 양태에서, 암이 있는 대상체는 단편 비율의 상관관계(예를 들어, cfDNA 단편 비율 대 기준 DNA 단편 비율, 예컨대 한 명 이상의 건강한 대상체로부터의 DNA 단편 비율의 상관관계)를 가질 수 있으며, 즉 이는 평균적으로 건강한 대상체의 단편 비율의 상관관계(예를 들어, cfDNA 단편 비율 대 기준 DNA 단편 비율, 예컨대 한 명 이상의 건강한 대상체로부터의 DNA 단편 비율의 상관관계) 보다 약 0.19 내지 약 0.30(예를 들어, 약 0.25) 더 낮다.

[0064] 본 발명의 방법론은 cfDNA 단편화 프로파일을 기반으로 점수(예를 들어, DELFI 점수)를 계산하는 단계를 추가로 포함한다. 일부 양태에서, 점수를 계산하는 단계는 다음을 포함한다: i) 샘플의 짧은 cfDNA 단편 대 긴 cfDNA 단편의 비율을 결정하는 단계, ii) 염색체 팔에 의해 샘플의 cfDNA 단편에 대한 Z-점수를 결정하는 단계, iii) 전산 혼합 모델 분석을 사용하여 cfDNA 단편 밀도를 정량화하는 단계 및 iv) 기계 학습 모델을 사용하여 i)-iii)의 출력을 처리하여 점수를 정의하는 단계. 다양한 양태에서, 점수는 대상체의 전체 생존 가능성을 결정하는 데 활용된다.

[0065] 하나의 예시적인 예(실시예 1)에서, 다중 암 코호트에서, 발명자들은 각 개인의 5MB 빈에 의한 짧은 단편 대 긴 단편의 비율, 염색체 팔에 의한 Z-점수 및 cfDNA 단편 크기의 혼합 모델을 서열 분석하는 낮은 적용 범위의 전체 계층으로부터 계산하였다. 이러한 특징을 입력으로 사용하여, 발명자들은 교차 검증된 경사 증폭 기계를 각 사람의 암 상태(암/암 없음)에 맞춘다. 이 모델의 출력은 0에서 1의 범위의 점수이고, 숫자가 높을수록 암의 신호가 강하고 숫자가 낮을수록 비암의 신호와 유사하다는 것을 가리킨다. 완료되면 암으로 진단된 샘플만 유지된다.

[0066] 일부 양태에서, 출력된 점수를 다음과 같이 분석한다. 추적 시간, 추적 종료 시 환자의 생존 여부, 상기 기계 학습 모델로부터의 점수를 이용하여, cfDNA 단편화와 생존의 관계를 판단하였다. 도 5에서 나타나 있듯이, 카플란-마이어 곡선에서 암이 있는 개인에서 높은 점수 대 낮은 점수로 강력한 분리가 결정되었다. 추가적으로, 이 점수와 기타 임상 특징의 독립성은 콕스 비례 위험 모델을 맞추고 점수, 암 병기 및 환자 나이에 대한 회귀를 통해 평가되었다.

[0067] 도 5를 참조하면 상기에서 논의된 바와 같이, 계산된 DELFI 점수는 높은 점수(>0.5) 대 낮은 점수(<0.5)를 정의하는 데 사용된 컷오프 값에 관계없이 암(폐암 제외)이 있는 개인의 묘사된 카플란-마이어 곡선을 분리한다. 각 패널 상단의 숫자는 결정된 컷오프 값을 가리킨다.

[0068] 도 6은 두 가지 설정에서 콕스 비례 위험 모델의 결과를 나타낸다. 첫 번째 설정(플롯의 왼쪽 패널)에서, DELFI 점수는 연속적으로 처리된다. 두 번째 설정(플롯의 오른쪽 패널)에서 DELFI 점수는 높음(>0.5) 또는 낮음(<0.5)으로 처리된다. 두 설정 모두에서, DELFI 점수는 채혈 및 병기에서 나이를 조정하더라도 생존에 대한 강력한 예측 변수이다. 병기는 1기에 관한 것임을 유념한다.

[0069] 현재 설명된 방법 및 시스템은 대상체의 암 상태를 검출하고, 예측하고, 치료하고/치료하거나 모니터링하는 데 유용하다. 임의의 적절한 대상체, 예컨대 포유동물은 본원에 설명된 바와 같이 평가되고, 모니터링되고/모니터링되거나 치료될 수 있다. 본원에 설명된 바와 같이 평가되고, 모니터링되고/모니터링되거나 치료될 수 있는 일부 포유동물의 예에는 인간, 영장류, 예컨대 원숭이, 개, 고양이, 말, 소, 돼지, 양, 마우스 및 래트가 비제한적으로 포함된다. 예를 들어, 암이 있거나 암이 의심되는 인간은 본원에 설명된 방법을 사용하여 평가될 수 있고, 선택적으로 본원에 설명된 바와 같은 하나 이상의 암 치료로 치료될 수 있다.

[0070] 임의의 적절한 유형의 암이 있거나 암이 있는 것으로 의심되는 대상체는 본원에 설명된 방법 및 시스템을 사용하여 평가되고/평가되거나 치료될 수 있다(예를 들어, 대상체에게 하나 이상의 암 치료 투여에 의한). 암은 임의의 병기의 암일 수 있다. 일부 양태에서, 암은 조기 암일 수 있다. 일부 양태에서, 암은 무증상 암일 수 있다. 일부 양태에서, 암은 잔여 질병 및/또는 재발(예를 들어, 외과적 절제 후 및/또는 암 치료요법 후)일 수 있다. 암은 임의의 유형의 암일 수 있다. 본원에 설명된 바와 같이 평가되고, 모니터링되고/모니터링되거나 치

료될 수 있는 암 유형의 예에는 폐암, 대장암, 전립선암, 유방암, 췌장암, 담관암, 간암, CNS, 위암, 식도암, 위장관 간질 종양(GIST), 자궁암 및 난소암이 비제한적으로 포함된다. 추가적인 유형의 암에는 골수종, 다발성 골수종, B세포 림프종, 여포성 림프종, 림프구성 백혈병, 백혈병 및 골수성 백혈병이 비제한적으로 포함된다. 일부 양태에서, 암은 고형 종양이다. 일부 양태에서, 암은 육종, 암종 또는 림프종이다. 일부 양태에서, 암은 폐암, 대장암, 전립선암, 유방암, 췌장암, 담관암, 간암, CNS, 위암, 식도암, 위장관 간질 종양(GIST), 자궁암 또는 난소암이다. 일부 양태에서, 암은 혈액암이다. 일부 양태에서, 암은 골수종, 다발성 골수종, B 세포 림프종, 여포성 림프종, 림프구성 백혈병, 백혈병 또는 골수성 백혈병이다.

[0071] 본원에 설명된 바와 같이 암이 있거나 암이 의심되는 대상체를 치료할 때, 대상체에게는 하나 이상의 암 치료가 투여될 수 있다. 암 치료는 임의의 적절한 암 치료일 수 있다. 본원에 설명된 하나 이상의 암 치료는 임의의 적절한 빈도(예를 들어, 며칠 내지 몇 주 범위의 기간에 걸쳐 1회 또는 여러 번)로 대상체에게 투여될 수 있다. 암 치료의 예는 외과적 개입, 보조 화학치료요법, 신보조 화학치료요법, 방사선 치료요법, 호르몬 치료요법, 세포독성 치료요법, 면역치료요법, 입양 T 세포 치료요법(예를 들어, 키메라 항원 수용체 및/또는 야생형 또는 변형된 T 세포 수용체를 갖는 T 세포), 표적화 치료요법, 예컨대 키나제 저해제(예를 들어, 특정 유전적 병변, 예컨대 전좌 또는 돌연변이를 표적으로 하는 키나제 저해제)의 투여(예를 들어, 키나제 저해제, 항체, 이중특이적 항체), 신호 전달 저해제, 이중특이적 항체 또는 항체 단편(예를 들어, BiTE), 단클론 항체, 면역 체크포인트 저해제, 수술(예를 들어, 수술적 절제) 또는 상기의 임의의 조합을 비제한적으로 포함한다. 일부 양태에서, 암 치료는 암의 증증도를 저감시키고, 암의 증상을 저감시키며/저감시키거나 대상체 내에 존재하는 암 세포의 수를 저감시킬 수 있다.

[0072] 일부 양태에서, 암 치료는 화학치료요법제일 수 있다. 화학치료요법제의 비제한적 예에는 다음이 포함된다: 암 사크린, 아자시티딘, 악사티오프린, 베바시주맙(또는 이의 항원 결합 단편), 블레오마이신, 부술판, 카르보플라틴, 카페시타빈, 클로람부실, 시스플라틴, 시클로포스파미드, 시타라빈, 다카르바진, 다우노루비신, 도세탁셀, 독시플루리딘, 독소루비신, 에피루비신, 에를로티닙 하이드로클로라이드, 에토포시드, 피우다라빈, 플록수리딘, 플루다라빈, 플루오로우라실, 켄시타빈, 하이드록시우레아, 이다루비신, 이포스파미드, 이리노테칸, 로무스틴, 메클로레타민, 멜팔란, 메르캅토포린, 메토티렉세이트, 미토마이신, 미톡산트론, 옥살리플라틴, 파클리탁셀, 페메트렉스드, 프로카바진, 전 트랜스 레티노산, 스트렙토조신, 타플루포시드, 테모졸로미드, 테니포시드, 티오구아닌, 토포테칸, 우라무스틴, 발루비신, 빈블라스틴, 빈크리스틴, 빈데신, 비노렐빈 및 이들의 조합. 항암 치료요법의 추가적인 예는 당업계에 알려져 있다; 예를 들어 미국임상종양학회(ASCO), 유럽종양학회(ESMO) 또는 국립종양암네트워크(NCCN)의 치료요법 지침을 참조한다.

[0073] 본원에 설명된 바와 같이 암이 있거나 암이 의심되는 대상체를 모니터링할 때, 모니터링은 암 치료 과정 전, 도중 및/또는 후일 수 있다. 본원에 제공된 모니터링 하는 방법은 하나 이상의 암 치료의 효능을 결정하고/결정하거나 증가된 모니터링을 위한 대상체를 선택하는 데 사용될 수 있다.

[0074] 일부 양태에서, 모니터링은 하나 이상의 암 치료(예를 들어, 하나 이상의 암 치료의 효능)를 모니터링할 수 있는 기존 기술을 포함할 수 있다. 일부 양태에서, 증가된 모니터링을 위해 선택된 대상체는 증가된 모니터링을 위해 선택되지 않은 대상체에 비해 증가된 빈도로 진단 테스트(예를 들어, 본원에 개시된 진단 테스트 중 임의의 것)를 투여받을 수 있다. 예를 들어, 증가된 모니터링을 위해 선택된 대상체는 매일 2회, 매일, 격주, 매주, 격월, 월간, 분기별, 반년마다, 매년 또는 그 중 임의의 빈도로 진단 테스트가 투여될 수 있다.

[0075] 다양한 양태에서, DNA는 대상체로부터 채취한 생물학적 샘플에 존재하며 본 발명의 방법론에 사용된다. 생물학적 샘플은 DNA를 포함하는 사실상 임의의 유형의 생물학적 샘플일 수 있다. 생물학적 샘플은 전형적으로 유체, 예컨대 전혈 또는 순환하는 cfDNA가 있는 그의 부분이다. 구현예에서, 샘플에는 종양 또는 액체 생검, 예컨대 비제한적으로 양수, 방수, 유리체액, 혈액, 전혈, 분획 혈액, 혈장, 혈청, 모유, 뇌척수액(CSF), 귀에지(귀지), 유미, 미즙, 내림프, 외림프, 대변, 호흡, 위산, 위액, 림프, 점액(비강 배액 및 가래 포함), 심낭액, 복막액, 흉막액, 고름, 점막 분비물, 타액, 호기 호흡 응축물, 피지, 정액, 가래, 땀, 윤활액, 눈물, 토사물, 전립선액, 유두 흡인액, 눈물액, 땀, 땀 면봉, 세포 용해물, 위장액, 생검 조직 및 소변 또는 기타 생물학적 체액으로부터의 DNA를 포함한다. 일 양태에서, 샘플은 순환 종양 세포로부터의 DNA를 포함한다.

[0076] 상기 개시된 바와 같이, 생물학적 샘플은 혈액 샘플일 수 있다. 혈액 샘플은 당업계에 알려진 방법, 예컨대 손가락 찌르기 또는 정맥절개술을 사용하여 얻을 수 있다. 적합하게는, 혈액 샘플은 대략 0.1 내지 20ml, 또는 대안적으로 대략 1 내지 15ml이고 혈액 부피는 대략 10ml이다. 혈액 내 순환하는 유리 DNA뿐만 아니라 소량도 사용될 수 있다. 바늘 생검, 카테터, 배설 또는 DNA를 함유하는 체액 생산에 의한 마이크로샘플링 및 샘플링도 감

재적인 생물학적 샘플원이다.

- [0077] 본 개시의 방법 및 시스템은 핵산 서열 정보를 활용하므로, 핵산 증폭, 중합 효소 연쇄 반응(PCR), 나노포어 서열 분석, 454 서열 분석, 삽입 태그 서열 분석을 포함하는 핵산 서열 분석을 수행하기 위한 임의의 방법 또는 서열 분석 장치를 포함할 수 있다. 일부 양태에서, 본 개시의 방법론 또는 시스템은 시스템, 예컨대 주식회사 일루미나에 의해 제공되는 것들(HiSeq™ X10, HiSeq™ 1000, HiSeq™ 2000, HiSeq™ 2500, Genome Analyzers™, MiSeq™, NextSeq, NovaSeq 6000 시스템을 포함하나 이에 제한되지 않음), 어플라이드 바이오시스템스 라이프 테크놀로지(SOLiD™ System, Ion PGM™ Sequencer, ion Proton™ Sequencer) 또는 제넵시스 또는 BGI MGI 및 기타 시스템을 활용한다. 핵산 분석은 또한 옥스포드 나노포어 테크놀로지(GridiON™, MiniON™) 또는 퍼시픽 바이오사이언스(Pacbio™ RS II 또는 후속 I 또는 II)에 의해 제공되는 시스템에 의해 수행될 수 있다.
- [0078] 본 발명은 개시된 방법의 단계를 수행하기 위한 시스템을 포함하며 기능적 구성요소 및 다양한 처리 단계의 관점에서 부분적으로 설명된다. 이러한 기능적 구성요소 및 처리 단계는 특정된 기능을 수행하고 다양한 결과를 달성하도록 구성된 임의의 수의 구성요소, 동작 및 기술에 의해 실현될 수 있다. 예를 들어, 본 발명은 다양한 기능을 수행할 수 있는 다양한 생물학적 샘플, 바이오마커, 요소, 재료, 컴퓨터, 데이터원, 저장 시스템 및 매체, 정보 수집 기술 및 프로세스, 데이터 처리 기준, 통계 분석, 회귀 분석 등을 사용할 수 있다.
- [0079] 따라서, 본 발명은 암을 검출하고, 분석하고/분석하거나 평가하기 위한 시스템을 추가로 제공한다. 다양한 양태에서, 시스템은 다음을 포함한다: (a) 샘플에 대한 낮은 적용 범위의 전체 게놈 서열 분석 데이터 세트를 생성하도록 구성된 서열 분석기; 및 (b) 본 발명의 방법을 수행하는 기능성을 갖춘 컴퓨터 시스템 및/또는 프로세서.
- [0080] 일부 양태에서, 컴퓨터 시스템은 하나 이상의 추가적인 모듈을 추가로 포함한다. 예를 들어, 시스템은 적합한 유전적 구성성분 분석, 예를 들어 특정 크기의 cfDNA 단편을 선택하도록 작동 가능한 하나 이상의 추출 및/또는 분리 유닛을 포함할 수 있다.
- [0081] 일부 양태에서, 컴퓨터 시스템은 시각적 디스플레이 장치를 추가로 포함한다. 시각적 디스플레이 장치는 곡선 맞춤선, 기준 곡선 맞춤선 및/또는 두 가지의 비교를 표시하도록 작동할 수 있다.
- [0082] 본 발명의 다양한 양태에 따른 검출 및 분석을 위한 방법은 예를 들어 컴퓨터 시스템에서 작동하는 컴퓨터 프로그램 사용하여 임의의 적합한 방식으로 구현될 수 있다. 본원에 논의된 바와 같이, 본 발명의 다양한 양태에 따른 예시적인 시스템은 컴퓨터 시스템, 예를 들어 프로세서 및 랜덤 액세스 메모리, 예컨대 원격 액세스 가능한 애플리케이션 서버, 네트워크 서버, 개인용 컴퓨터 또는 워크스테이션을 포함하는 기존의 컴퓨터 시스템과 함께 구현될 수 있다. 컴퓨터 시스템은 또한 추가적인 메모리 장치 또는 정보 저장 시스템, 예컨대 대용량 저장 시스템 및 사용자 인터페이스, 예를 들어 기존의 모니터, 키보드 및 추적 장치를 적합하게 포함한다. 그러나 컴퓨터 시스템은 임의의 적합한 컴퓨터 시스템 및 연관 장비를 포함할 수 있으며 임의의 적합한 방식으로 구성될 수 있다. 일 구현예에서, 컴퓨터 시스템은 독립형 시스템을 포함한다. 또 다른 구현예에서, 컴퓨터 시스템은 서버와 데이터베이스를 포함하는 컴퓨터의 네트워크의 일부이다.
- [0083] 정보를 수신하고, 처리하고, 분석하는 데 요구되는 소프트웨어는 단일 장치에서 구현될 수도 있고, 복수 개의 장치에서 구현될 수도 있다. 소프트웨어는 정보의 저장 및 처리가 사용자에게 대해 원격으로 발생하도록 네트워크를 통해 액세스할 수 있다. 본 발명의 다양한 양태에 따른 시스템과 그 다양한 요소는 검출 및/또는 분석, 예컨대 데이터를 수집하고, 처리하고, 분석하고, 보고하고/보고하거나 진단을 용이하게 하는 기능 및 동작을 제공한다. 예를 들어, 본 양태에서, 컴퓨터 시스템은 인간 게놈 또는 그 영역과 관련된 정보를 수신하고, 저장하고, 검색하고, 분석하고 보고할 수 있는 컴퓨터 프로그램을 실행한다. 컴퓨터 프로그램은 다양한 기능 또는 작업, 예컨대 미가공 데이터를 처리하고 보충 데이터를 생성하기 위한 처리 모듈 및 질병 상태 모델 및/또는 진단 정보의 정량적 평가를 생성하기 위해 미가공 데이터 및 보충 데이터를 분석하기 위한 분석 모듈을 수행하는 다중 모듈을 포함할 수 있다.
- [0084] 시스템에 의해 수행되는 절차는 분석 및/또는 암 진단을 용이하게 하는 임의의 적합한 공정을 포함할 수 있다. 일 구현예에서, 시스템은 질병 상태 모델을 확립하고/확립하거나 환자의 질병 상태를 결정하도록 구성된다. 질병 상태를 결정하거나 식별하는 단계는 질병에 관한 환자의 상태에 관한 임의의 유용한 정보를 생성하는 단계, 예컨대 진단을 수행하고, 진단에 도움이 되는 정보를 제공하고, 질병의 병기 또는 진행을 평가하고, 질병에 대한 취약성을 가리킬 수 있는 상태를 식별하고, 추가 테스트가 권장될 수 있는지 여부를 식별하고, 하나 이상의

치료 프로그램의 효능을 예측하고/예측하거나 평가하고, 질병 상태, 질병 가능성 또는 환자의 기타 건강 양태를 평가하는 것을 포함할 수 있다.

[0085] 다음의 실시예는 본 발명의 장점과 특징을 더욱 도시하기 위해 제공되는 것이나, 이는 본 발명의 범위를 제한하려는 의도가 아니다. 이 실시예는 사용될 수 있는 것 중 전형적인 것이지만, 당업자에게 알려진 다른 절차, 방법론 또는 기술이 대안적으로 사용될 수 있다.

[0086] **실시예 1**

[0087] **전향적 진단 코호트에서 게놈 전체 cfDNA 단편화를 사용한 암 검출**

[0088] 게놈 전체 cfDNA 단편화 패턴은 암이 있는 개인과 암이 없는 개인의 혈장 샘플 사이의 높은 민감도와 특이성으로 구분하는 것으로 입증되었다.

[0089] 본 실시예에서, 본 개시의 방법론은 암을 검출하고 전체 환자 생존을 예측하기 위해 활용되었다.

[0090] 연구의 목적은 cfDNA 단편화 검정을 혈액 기반 검진 테스트로 감정하여 복수의 상이한 고형 종양을 검출하고 전산 점수 체계를 사용하여 전체 환자 생존을 예측하는 것이었다.

[0091] **방법**

[0092] 혈장 샘플: 암의 비기관 특이적 징후 및 증상으로 인해 헤르레브 및 겐토프테 병원(덴마크 코펜하겐의 코펜하겐 대학 병원)의 진단 외래 진료소로 의뢰된 281명의 환자로부터 샘플을 수집하였다.

[0093] cfDNA 단편화 접근법: cfDNA 단편화 접근법은 도 1에 요약되어 있다. cfDNA는 혈장에서 추출되어 서열 분석 라이브러리로 처리되고 낮은 적용 범위의 전체 게놈 서열 분석(WGS)으로 검사되고 게놈에 매핑되고 분석되어 게놈 전반에 걸쳐 cfDNA 단편화 프로파일을 결정하였다.

[0094] 기계 학습을 사용하여 DELFI 점수를 생성하고 개인을 건강하거나 암이 있는 것으로 분류하고 전체 환자 생존을 예측하였다.

[0095] **결과**

[0096] 암의 비침습적 검출을 위한 cfDNA 단편화 검정의 성과: 포함 후 3개월 이내에, 74명의 환자가 16개의 상이한 고형암 중 1개로 진단되었고 207명의 환자는 암이 없었다. 추가적인 결과는 도 2에 나타나 있다. 국소암과 전이암 및 모든 병기의 대장암, 폐암 및 기타 모든 암에 대한 곡선하면적(AUC)은 10회 반복, 10배 교차 검증을 사용하여 결정되었다.

[0097] 암 검출을 위한 cfDNA 단편화 검정의 전반적인 성과: 결과는 도 3에 요약되어 있다. 1-4기 암이 있는 개인 74명과 비암 대조군 207명의 분석에 대한 수신자 조작 특성(ROC)의 AUC이다.

[0098] DELFI 점수에 따른 생존: 도 4에 나타난 바와 같이 DELFI 점수가 높을수록 암 병기 또는 기타 임상적 특성과 관계없이 감소된 전체 생존과 연관이 있었다. 도 4는 DELFI 점수와 상관관계가 있는 대상체의 생존을 나타낸다. DELFI 점수가 높을수록 암 병기나 기타 임상적 특성과 관계없이 감소된 전체 생존과 연관이 있었다.

[0099] **결론**

[0100] 전향적으로 등록된 개인에 대한 이 연구는 암이 있는 개인과 암이 없는 개인을 구분하는 cfDNA 단편화 검정의 능력을 입증하였다. 본 발명의 검정은 낮은 적용 범위의 WGS로부터 얻은 단편화 관련 정보만을 사용하여 다중 암 설정에서 높은 성능을 나타냈다.

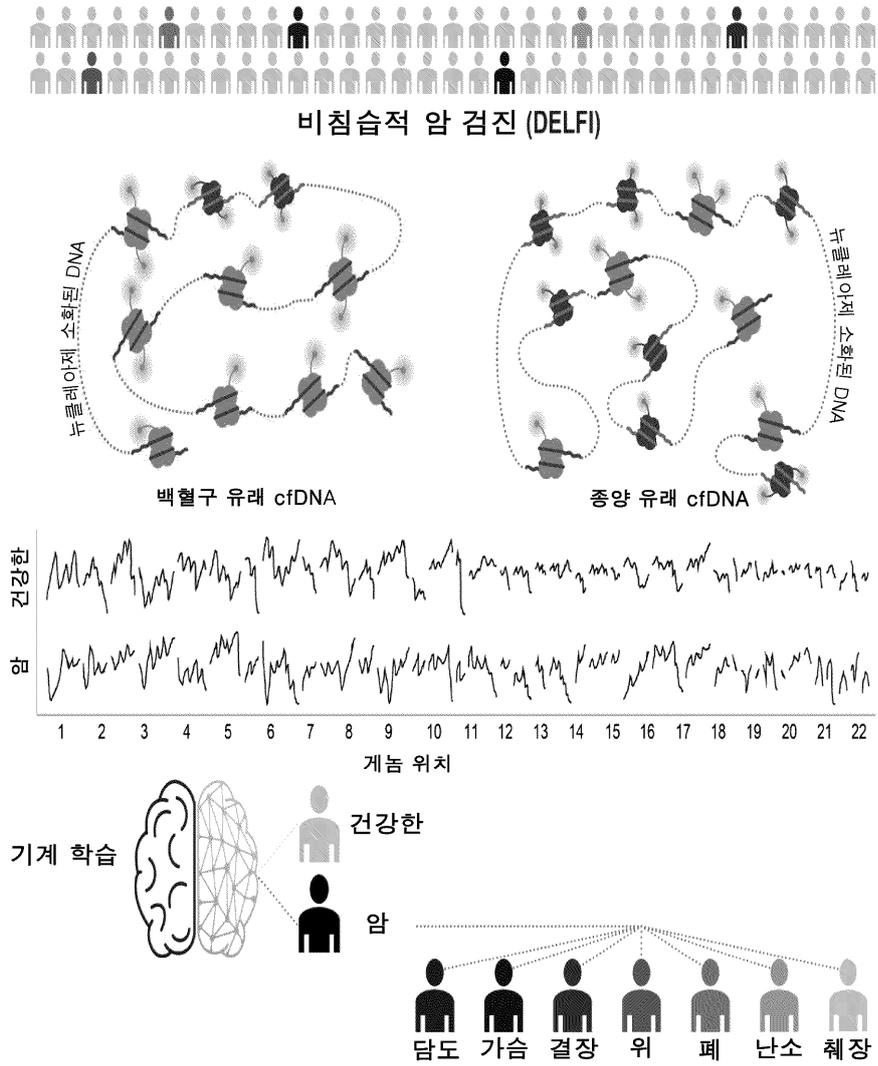
[0101] 결과는 기계 학습 모델이 cfDNA 단편화 프로파일을 사용하여 일반적인 비악성 상태(심혈관, 자가면역 또는 염증성 질병 포함)의 존재에도 불구하고 암과 비암 사이를 구분 지을 수 있음을 시사한다. 추가적으로, DELFI 점수가 높은 개인은 다른 특성과 관계없이 예후가 더 나빴다.

[0102] 이들 데이터는 단일 암과 다중 암 모두의 비침습적 검출을 위한 게놈 전체 cfDNA 단편화 분석의 개발을 지원한다.

[0103] 본 발명이 상기 실시예를 참조하여 설명되었으나, 수정 및 변형이 본 발명의 사상 및 범위 내에 포괄된다는 것이 이해될 것이다. 따라서, 본 발명은 다음의 청구항에 의해서만 제한된다.

도면

도면1



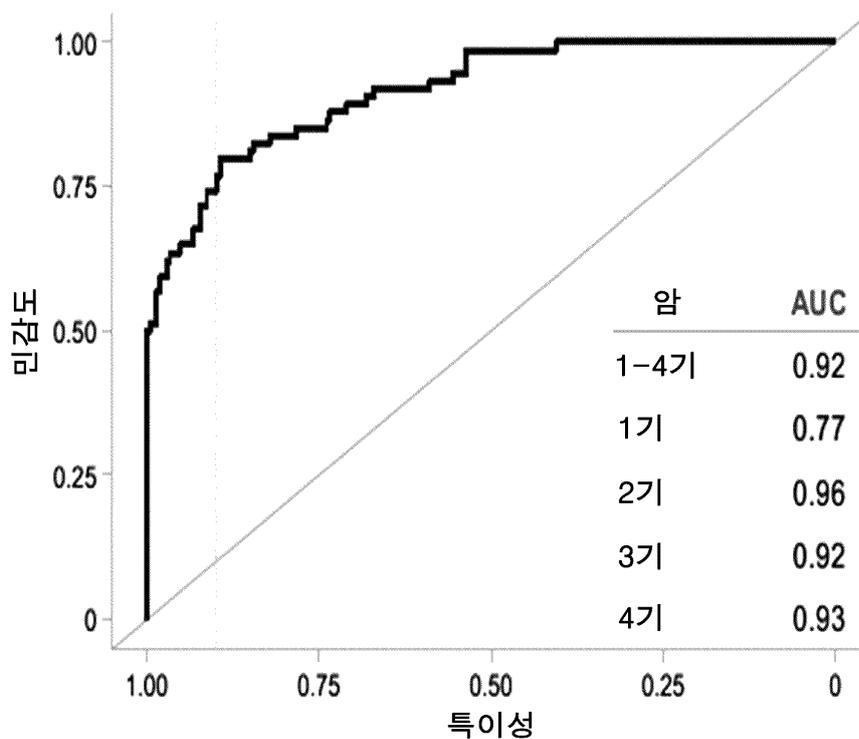
도면2

암	n	AUC	95% CI
 모든 암			
1-4기	72	0.92	0.88, 0.96
1-3기	24	0.90	0.84, 0.96
4기	48	0.93	0.89, 0.97
 대장암 (1-4기)	12	0.91	0.85, 0.97
 폐암 (1-4기)	12	0.93	0.86, 1.00
 기타 암	48	0.92	0.87, 0.96

10회 반복, 10배 교차 검증을 사용하여 결정된 국소 및 전이 암 및 모든 단계의 대장암, 폐암 및 모든 기타 암에 대한 곡선하면적(AUC).

도면3

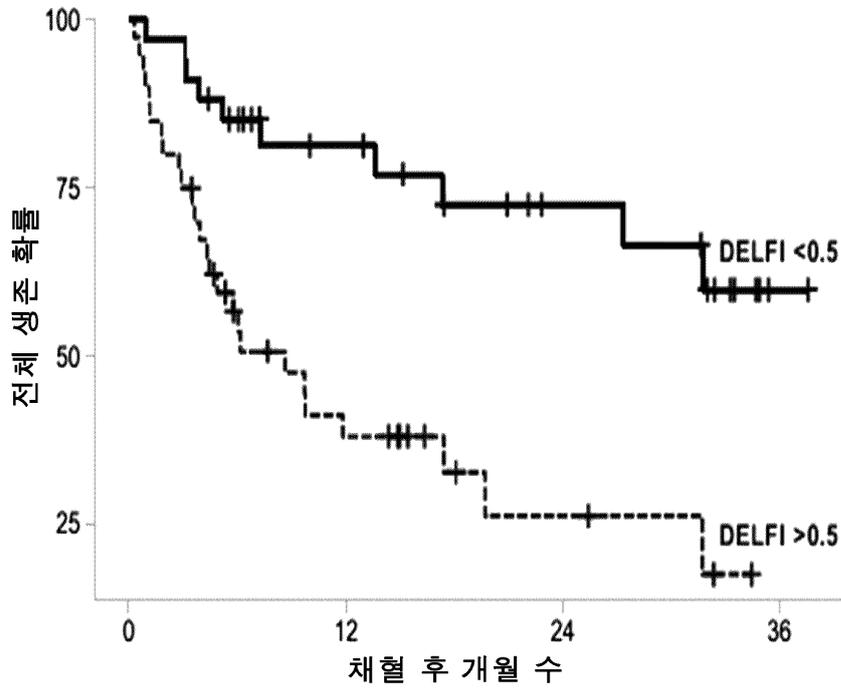
암 검출을 위한 cfDNA 단편화 검정의 전체 성과



1-4기 암을 가진 74명의 개인과 207명의 비암 대조군 분석을 위한 수신자 조작 특성(ROC)의 AUC.

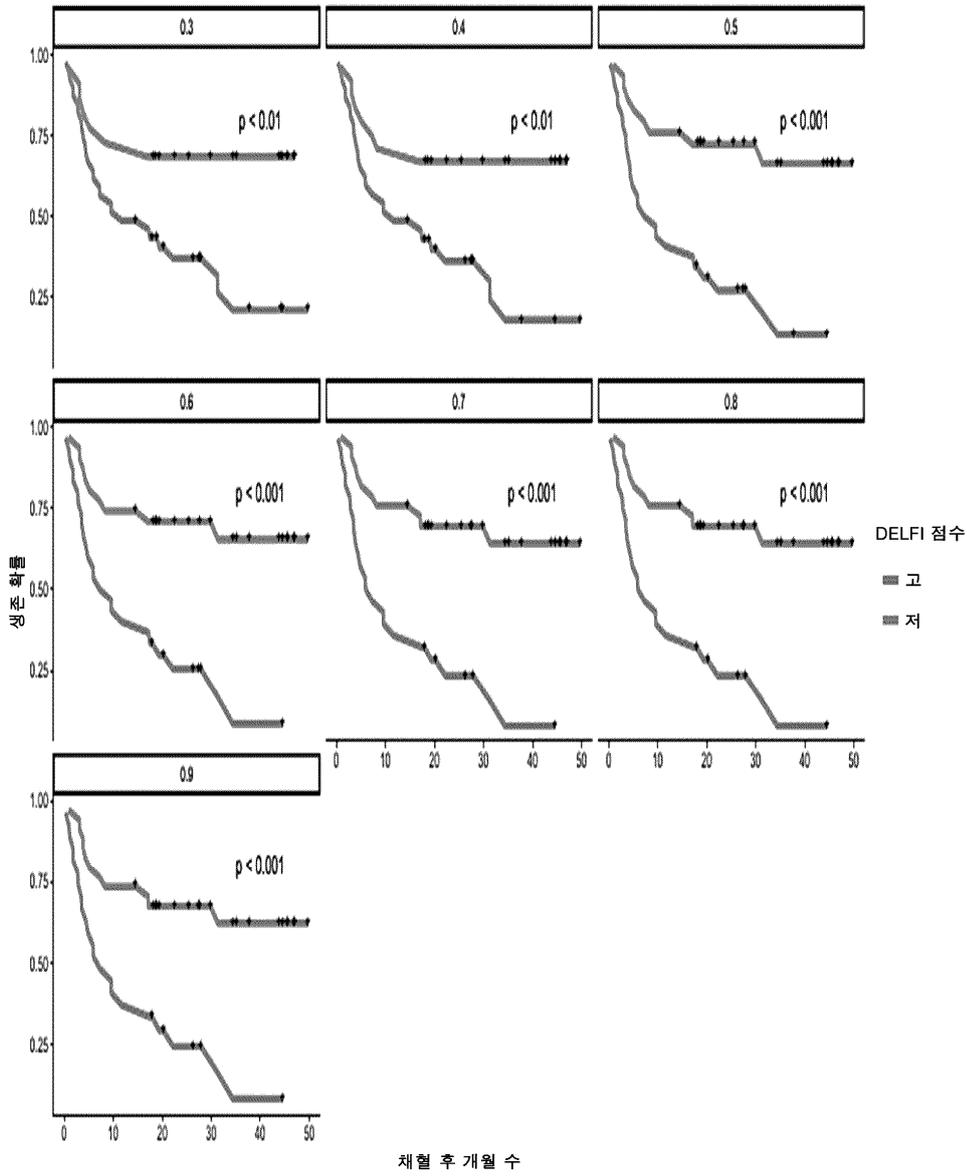
도면4

DELFI 점수에 따른 생존



* DELFI 점수가 높을수록 암 병기나 기타 임상적 특성과 관계없이 감소된 전체 생존과 연관이 있었다.

도면5



도면6

