



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2020-0109470
(43) 공개일자 2020년09월23일

(51) 국제특허분류(Int. Cl.)
G16B 25/10 (2019.01) C12Q 1/6809 (2018.01)
C12Q 1/6886 (2018.01)
(52) CPC특허분류
G16B 25/10 (2019.02)
C12Q 1/6809 (2018.05)
(21) 출원번호 10-2019-0028539
(22) 출원일자 2019년03월13일
심사청구일자 2019년03월13일

(71) 출원인
울산대학교 산학협력단
울산광역시 남구 대학로 93(무거동)
송실대학교산학협력단
서울특별시 동작구 상도로 369 (상도동)
재단법인 아산사회복지재단
서울특별시 송파구 올림픽로43길 88 (풍납동)
(72) 발명자
공경엽
경기도 성남시 수정구 대왕판교로 801, 101동 101호
이희진
서울특별시 강남구 압구정로39길 58, 63동 408호
(뒷면에 계속)
(74) 대리인
특허법인태백

전체 청구항 수 : 총 8 항

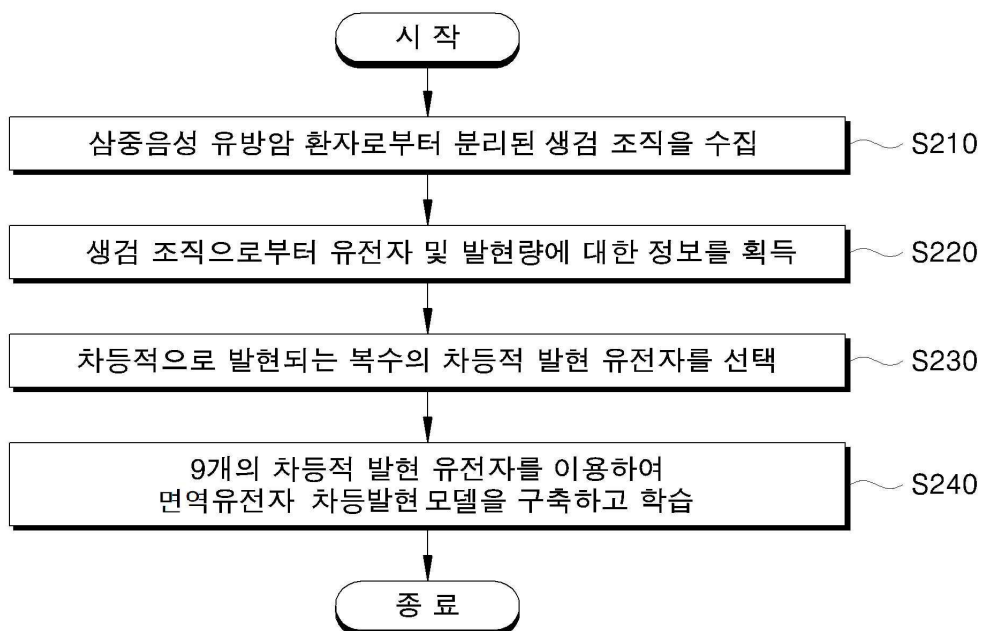
(54) 발명의 명칭 면역유전자 차등발현 예측 모델을 이용한 유방암 환자에 대한 완전 관해 예측 방법

(57) 요약

본 발명은 면역유전자 차등발현 모델을 이용한 유방암 환자에 대한 완전 관해 예측 방법에 대한 것이다.

본 발명에 따르면, 피검자인 유방암 환자로부터 생검 조직을 수집하는 단계, RNA 추출 키트를 사용하여 상기 생검 조직으로부터 복수의 차등적 발현 유전자 및 발현량 정보를 획득하는 단계, 그리고 상기 복수의 차등적 발현 (뒷면에 계속)

대표도 - 도2



유전자에 대한 발현량 정보를 기 학습된 면역유전자 차등발현 모델에 적용하여 상기 피검자의 완전 관해 가능성을 예측하는 단계를 포함한다.

이와 같이 본 발명에 따르면, 유방암 환자에 대한 완전 관해 예측 방법은 유방암 내에 면역 관련 유전자 중에서 차등적으로 발현하는 유전자를 추출하고, 추출된 차등적 발현 유전자를 통해 샘플의 수가 적은 데이터로부터 완전 관해를 예측할 수 있는 면역유전자 차등발현 모델을 구축할 수 있다. 따라서, 적은 비용으로 높은 정확성을 통해 유방암 환자의 완전 관해를 예측할 수 있다.

(52) CPC특허분류

C12Q 1/6886 (2018.05)

C12Q 2545/114 (2019.08)

C12Q 2600/106 (2013.01)

C12Q 2600/158 (2013.01)

(72) 발명자

김영애

서울특별시 송파구 풍성로6가길 14-3, 101호

서정환

서울특별시 강동구 천호대로168나길 56-14, 302호
(성내동, 더존팰리스)

박성희

서울특별시 관악구 관악로30길 27, 105동 1302 (봉천동, 관악푸르지오아파트)

임정백

전라남도 나주시 다시면 랑동길 29-9

윤성민

강원도 춘천시 춘천로146번길 15, 5층

명세서

청구범위

청구항 1

면역유전자 차등발현 모델을 이용한 유방암 환자에 대한 완전 관해 예측 방법에 있어서,

피검자인 유방암 환자로부터 생검 조직을 수집하는 단계,

RNA 추출 키트를 사용하여 상기 생검 조직으로부터 복수의 차등적 발현 유전자 및 발현량 정보를 획득하는 단계, 그리고

상기 복수의 차등적 발현 유전자에 대한 발현량 정보를 기 학습된 면역유전자 차등발현 모델에 적용하여 상기 피검자의 완전 관해 가능성을 예측하는 단계를 포함하는 완전 관해 예측 방법.

청구항 2

제1항에 있어서,

상기 면역유전자 차등발현 모델을 구축하여 학습시키는 단계를 더 포함하며,

상기 면역유전자 차등발현 모델을 구축하여 학습시키는 단계는,

선행항암화학요법을 받은 복수의 유방암 환자로부터 생검 조직을 수집하는 단계,

RNA 추출키트를 이용하여 상기 수집된 생검 조직으로부터 유전자 정보를 획득하는 단계,

상기 복수의 유방암 환자를 완전 관해가 이루어진 환자 그룹과 관해에 도달하지 못한 환자 그룹으로 그룹핑하고, 각 그룹으로부터 획득한 유전자 정보를 이용하여, 상기 두 그룹 간의 차등적으로 발현되는 복수의 차등적 발현 유전자를 선택하는 단계,

상기 선택된 복수의 차등적 발현 유전자가 유방암 완전 관해에 영향력을 주는 유전자인지 검증을 하는 단계, 그리고

상기 검증된 복수의 차등적 발현 유전자 및 발현량 정보를 랜덤 포레스트를 기반으로 학습하여 면역유전자 차등 발현 모델을 구축하는 단계를 포함하는 완전 관해 예측 방법.

청구항 3

제2항에 있어서,

상기 복수의 차등적 발현 유전자를 선택하는 단계는,

음이항(Negative binominal) 모형을 적용한 edgeR 방법을 이용하여 완전 관해 그룹과 완전 관해에 도달하지 못한 그룹 사이에 차등적으로 발현되는 복수의 차등적 발현 유전자를 선택하는 완전 관해 예측 방법.

청구항 4

제2항에 있어서,

상기 복수의 차등적 발현 유전자를 선택하는 단계는,

상기 생검 조직으로부터 획득한 유전자 정보를 상기 edgeR 방법에 적용하여 FDR 값이 0.05보다 낮은 유전자를 차등적 발현 유전자로 선택하는 완전 관해 예측 방법.

청구항 5

제2항에 있어서,

상기 검증을 하는 단계는,

Cox모델에 상기 선택된 복수의 차등적 발현 유전자를 적용하여 생존에 대한 치료 효과를 추정하고, 추정된 결과

에 따라 차등적 발현유전자가 유방암 완전 관해에 미치는 영향력을 검증하는 완전 관해 예측 방법.

청구항 6

제2항에 있어서,

상기 분석모델을 구축하여 학습시키는 단계는,

상기 복수의 차등적 발현 유전자 중에서 무작위로 n 개 선택하여 테스트 셋을 생성하고, 생성된 테스트 셋을 이용하여 번역유전자 차등발현 모델을 구축하여 학습시키는 완전 관해 예측 방법.

청구항 7

제1 내지 제6항 중 어느 한 항에 있어서,

상기 복수의 차등적 발현 유전자는,

IL2RA, CCL5, SELE, CCL20, FCER1A, CD1A, GNLy, S100A9 및 C4A.B중에서 적어도 하나를 포함하는 완전 관해 예측 방법.

청구항 8

제1항에 있어서,

상기 유방암 환자는,

삼중음성 유방암 환자를 포함하는 완전 관해 예측 방법.

발명의 설명

기술 분야

- [0001] 본 발명은 번역유전자 차등발현 모델을 이용한 유방암 환자에 대한 완전 관해 예측 방법에 관한 것으로서, 더욱 상세하게는 유방암 내에 번역 관련 유전자 중에서 차등적으로 발현하는 유전자를 추출하여 번역유전자 차등발현 모델을 구축하고, 구축된 번역유전자 차등발현 모델을 이용하여 완전관해 여부를 예측하는 완전 관해 예측 방법에 관한 것이다.

배경 기술

- [0002] 삼중음성 유방암이란 전체 유방암의 15-25%를 차지하는 한 아형으로 에스트로겐 수용체, 프로게스테론 수용체, HER2 수용체 모두가 발현되지 않아 유방암의 대표적 표적치료제인 항호르몬제(tamoxifen)이나 트라스투주맙(trastuzumab)의 사용이 불가능하며 효과적인 표적치료제가 없어 다른 아형에 비해서 예후가 나쁘다.
- [0003] 그러나, 삼중음성 유방암 환자의 경우 선행항암화학요법을 시행하면 삼중음성 유방암이 아닌 환자들에 비해 유의하게 높은 완전 관해율(pathologic complete response, pCR)을 보인다.
- [0004] 여기서, 삼중음성 유방암 환자들의 항암제에 대한 반응성과 예후를 분석해보면, 수술 전에 선행항암화학요법을 받은 환자들 중에서 항암제에 대한 반응성이 좋으면 좋은 예후를 보이고, 항암제에 대해 반응성이 좋지 않으면 암의 진행이 급격히 이루어져 나쁜 예후를 보인다.
- [0005] 따라서 환자 개개인의 치료 반응을 최적화하고 항암치료로 인한 독성을 최소화하기 위한 개인 맞춤형치료전략(personalized therapy) 수립이 절실하며 이를 위해서는 항암치료에 대한 종양의 반응을 예측할 수 있는 생체 표지자의 개발이 필요하다.
- [0006] 본 발명의 배경이 되는 기술은 대한민국 공개특허공보 제10-2018-0094497호(2018.03.23. 공개)에 개시되어 있다.

발명의 내용

해결하려는 과제

- [0007] 본 발명이 이루고자 하는 기술적 과제는, 유방암 내에 번역 관련 유전자 중에서 차등적으로 발현하는 유전자를

추출하여 면역유전자 차등발현 모델을 구축하고, 구축된 면역유전자 차등발현 모델을 이용하여 완전 관해 여부 를 예측하는 완전 관해 예측 방법을 제공하는데 목적이 있다.

과제의 해결 수단

- [0008] 이러한 기술적 과제를 이루기 위한 본 발명의 실시예에 따르면, 면역유전자 차등발현 모델을 이용한 유방암 환 자에 대한 완전 관해 예측 방법에 있어서, 피검자인 유방암 환자로부터 생검 조직을 수집하는 단계, RNA 추출 키트를 사용하여 상기 수집된 생검 조직으로부터 복수의 차등적 발현 유전자 및 발현량 정보를 획득하는 단계, 그리고 상기 복수의 차등적 발현 유전자에 대한 발현량 정보를 기 학습된 면역유전자 차등발현 모델에 적용하여 상기 피검자의 완전 관해 가능성을 예측하는 단계를 포함한다.
- [0009] 상기 면역유전자 차등발현 모델을 구축하여 학습시키는 단계를 더 포함하며, 상기 면역유전자 차등발현 모델을 구축하여 학습시키는 단계는, 선행항암화학요법을 받은 복수의 유방암 환자로부터 생검 조직을 수집하는 단계, RNA 추출키트를 이용하여 상기 생검 조직으로부터 유전자 정보를 획득하는 단계, 상기 복수의 유방암 환자를 완 전 관해가 이루어진 환자 그룹과 관해에 도달하지 못한 환자 그룹으로 그룹핑하고, 각 그룹으로부터 획득한 유 전자 정보를 이용하여, 상기 두 그룹 간의 차등적으로 발현되는 복수의 차등적 발현 유전자를 선택하는 단계, 상기 선택된 복수의 차등적 발현 유전자가 유방암 완전 관해에 영향력을 주는 유전자인지 검증을 하는 단계, 그 리고 상기 검증된 복수의 차등적 발현 유전자 및 발현량 정보를 랜덤 포레스트를 기반으로 학습하여 면역유전자 차등발현 모델을 구축하는 단계를 포함할 수 있다.
- [0010] 상기 복수의 차등적 발현 유전자를 선택하는 단계는, 음이항(Negative binominal) 모형을 적용한 edgeR 방법을 이용하여 완전 관해 그룹과 완전 관해에 도달하지 못한 그룹 사이에서 차등적으로 발현되는 복수의 차등적 발현 유전자를 선택할 수 있다.
- [0011] 상기 복수의 차등적 발현 유전자를 선택하는 단계는, 상기 생검 조직으로부터 획득한 유전자 정보를 상기 edgeR 방법에 적용하여 FDR 값이 0.05보다 낮은 유전자를 차등적 발현 유전자로 선택할 수 있다.
- [0012] 상기 검증을 하는 단계는, Cox모델에 상기 선택된 복수의 차등적 발현 유전자를 적용하여 생존에 대한 치료 효 과를 추정하고, 추정된 결과에 따라 차등적 발현유전자가 유방암 완전 관해에 미치는 영향력을 검증할 수 있다.
- [0013] 상기 면역유전자 차등발현 모델을 구축하여 학습시키는 단계는, 상기 복수의 차등적 발현 유전자 중에서 무작위 로 n개 선택하여 테스트 셋을 생성하고, 생성된 테스트 셋을 이용하여 면역유전자 차등발현 모델을 구축하여 학 습시킬 수 있다.
- [0014] 상기 복수의 차등적 발현 유전자는, IL2RA, CCL5, SELE, CCL20, FCER1A, CD1A, GNLY, S100A9 및 C4A.B중에서 적어도 하나를 포함할 수 있다.
- [0015] 상기 유방암 환자는, 삼중음성 유방암 환자를 포함할 수 있다.

발명의 효과

- [0016] 이와 같이 본 발명에 따르면, 유방암 내에 포함된 면역 관련 유전자 중에서 완전 관해를 예측하는 차등적 발현 유전자를 추출하고, 추출된 차등적 발현 유전자를 통해 샘플의 수가 적은 데이터로부터 완전 관해를 예측할 수 있는 면역유전자 차등발현 모델을 구축할 수 있다. 따라서, 적은 비용으로 높은 정확성을 통해 유방암 환자의 완전 관해를 예측할 수 있다.
- [0017] 또한, 본 발명에 따르면, 유방암 환자의 예후와 치료반응 예측에 중요한 것으로 대두되고 있는 면역관련 인자를 분석하여 치료제 개발 기반을 마련할 수 있고, 치료제 개발로 인해 유방암의 진료 지침이 개선될 수 있다.
- [0018] 또한, 본 발명에 따르면, 유방암 환자의 치료 반응성 예측 모델 및 예후 예측 모델을 임상에 적용하여 사용할 수 있으며, 모델 개발에 사용된 분석 방법을 다른 암종에도 적용하여 면역 프로파일링 모델의 확장이 가능하다.

도면의 간단한 설명

- [0019] 도 1은 본 발명의 실시예에 따른 완전 관해 예측 장치를 개략적으로 도시한 도면이다.
- 도 2는 본 발명의 실시예에 따른 완전 관해 예측을 위한 면역유전자 차등발현 모델을 구축하는 방법을 개략적으 로 도시한 순서도이다.

도 3은 도 2에 도시된 S230단계에서 선택된 차등적 발현 유전자를 나타내는 도면이다.

도 4는 도 2에 도시된 S230단계에서 선택된 9개의 차등적 발현 유전자에 대해 검증한 결과를 나타내는 도면이다.

도 5는 도 4에 도시된 검증 결과에 따른 생존곡선을 나타내는 그래프이다.

도 6은 도 2에 도시된 S240단계에서 구축된 면역유전자 차등발현 모델의 유효성을 평가한 결과를 나타내는 그래프이다.

도 7은 본 발명의 실시예에 따른 면역유전자 차등발현 모델을 통해 완전 관해를 예측하는 방법을 개략적으로 도시한 순서도이다.

발명을 실시하기 위한 구체적인 내용

- [0020] 이하 첨부된 도면을 참조하여 본 발명에 따른 바람직한 실시예를 상세히 설명하기로 한다. 이 과정에서 도면에 도시된 선들의 두께나 구성요소의 크기 등은 설명의 명료성과 편의상 과장되게 도시되어 있을 수 있다.
- [0021] 또한 후술되는 용어들은 본 발명에서의 기능을 고려하여 정의된 용어들로서, 이는 사용자, 운용자의 의도 또는 관례에 따라 달라질 수 있다. 그러므로 이러한 용어들에 대한 정의는 본 명세서 전반에 걸친 내용을 토대로 내려져야 할 것이다.
- [0022] 이하에서는 도1을 이용하여 본 발명의 실시예에 따른 완전 관해 예측 장치를 더욱 상세하게 설명한다.
- [0023] 도 1은 본 발명의 실시예에 따른 완전 관해 예측 장치를 개략적으로 도시한 도면이다.
- [0024] 도 1에 도시된 바와 같이, 유방암 환자에 대한 완전 관해 예측 장치(100)는 수집부(110), 유전자 선택부(120), 모델생성부(130) 및 예측부(140)를 포함한다.
- [0025] 먼저, 수집부(110)는 선행항암화학요법을 받은 복수의 삼중음성 유방암 환자로부터 생검 조직을 수집한다. 여기서, 수집부(110)는 삼중음성 유방암 환자에 한정하지 않고 모든 유방암 환자로부터 생검 조직을 수집한다. 그리고, 생검 조직은 FFPE(Formalin-fixed, paraffin-embedded) 형태로 제공된다.
- [0026] 유전자 선택부(120)는 수집된 생검 조직에 FFPE RNA 추출 키트를 사용하여 RNA를 추출한다. 다만, 추출 키트는 FFPE RNA 추출 키트에 한정하지 않고, 당업자에게 알려져 있는 키트를 사용할 수도 있다.
- [0027] 그리고, 유전자 선택부(120)는 추출된 RNA를 nCounter®Analysis System에 적용하여 유전자 발현 데이터를 획득한다. 그 다음 유전자 발현 데이터를 완전 관해가 이루어진 환자 그룹과 관해에 도달하지 못한 그룹으로 나뉘어지면, 유전자 선택부(120)는 각 그룹 간 사이에 차등적으로 발현되는 복수의 차등적 발현 유전자를 선택한다.
- [0028] 모델생성부(130)는 선택된 복수의 차등적 발현 유전자를 이용하여 분석모델을 생성한다. 이때, 면역유전자 차등 발현 모델은 완전 관해 확률을 예측하는 모델이다.
- [0029] 부연하자면, 모델생성부(130)는 복수의 차등적 발현 유전자 중에서 무작위로 n개 선택하여 조합한 복수의 데이터 셋을 생성한다. 그리고, 모델생성부(130)는 생성된 복수의 데이터 셋을 7:3의 비율로 나뉘어, 7에 해당되는 데이터셋은 학습용으로 사용하고, 3에 해당하는 데이터 셋은 테스트용으로 사용한다. 즉, 모델생성부(130)는 데이터 셋을 이용하여 학습 및 테스트를 거침으로써 완전 관해를 예측하는 면역유전자 차등발현 모델을 생성한다.
- [0030] 그리고, 예측부(140)는 피검자로부터 획득한 차등적 발현 유전자 정보를 학습된 면역유전자 차등발현 모델에 적용하여 피검자의 완전 관해에 대한 확률을 예측한다.
- [0031] 이하에서는 도 2 내지 도 6을 이용하여 완전 관해 예측 장치를 이용하여 면역유전자 차등발현 모델을 구축하는 방법에 대해 더욱 상세하게 설명한다.
- [0032] 도 2는 본 발명의 실시예에 따른 완전 관해 예측을 위한 면역유전자 차등발현 모델을 구축하는 방법을 개략적으로 도시한 순서도이고, 도 3은 도2에 도시된 S230단계에서 선택된 차등적 발현 유전자를 나타내는 도면이다.
- [0033] 도 2에 도시된 바와 같이, 먼저 수집부(110)는 선행항암화학요법을 받은 유방암 환자로부터 분리된 생검 조직을 수집한다(S210).
- [0034] 부연하자면, 수집부(110)는 특정 기간 동안 선행항암화학요법을 받은 복수의 환자로부터 생검 조직을 획득하여 수집한다. 여기서, 복수의 환자는 완전 관해가 이루어진 환자와 완전 관해에 도달하지 못한 환자를 포함한다.

- [0035] 그 다음, 유전자 선택부(120)는 수집된 생검 조직으로부터 유전자 및 발현량에 대한 정보를 획득한다(S220).
- [0036] 상기 S220를 더욱 상세하게 설명하면, 유전자 선택부(120)는 완전 관해가 이루어진 환자 그룹과 완전 관해에 도달하지 못한 환자 그룹으로부터 각각 생검 조직을 획득하고, RNA 추출키트에 적용하여 생검 조직으로부터 RNA를 추출한다. 추출된 RNA는 그룹간의 상호 비교분석이 가능해지도록 정규화과정을 거친다.
- [0037] 즉, 유전자 선택부(120)는 추출된 RNA를 nCounter®Analysis System을 통해 분석한다. nCounter®Analysis System은 디지털 분석기가 RNA에 포함된 각 분자의 색을 포착 및 카운팅하여 유전자 및 발현량 정보를 획득한다.
- [0038] 그 다음, 유전자 선택부(120)는 획득한 유전자 및 발현량 정보를 음이항(Negative binominal) 모형을 적용한 edgeR 방법에 적용하여 완전 관해가 이루어진 환자 그룹과 완전 관해에 도달하지 않은 환자 그룹 사이에서 차등적으로 발현되는 복수의 차등적 발현 유전자를 선택한다(S230).
- [0039] 부연하자면, edgeR 방법을 통하여, 유방암에 대해 완전 관해가 이루어진 환자 그룹으로부터 분산을 추정된 후 분석하여 FDR(False Discovery Rate)값이 0.05보다 낮은 유전자를 추출한다.
- [0040] 즉, 도 3에 도시된 바와 같이, 유전자 선택부(120)는 완전 관해가 이루어진 환자 그룹과 완전 관해에 도달하지 않은 환자 그룹 사이에서 차등적으로 발현되는 9개의 차등적 발현 유전자를 선택한다.
- [0041] 9개의 차등적 발현 유전자는 하기의 표 1에 나타난 바와 같다.

표 1

No.	GeneSymbol	NCBI GeneID
1	IL2RA	3559
2	CCL5	6352
3	SELE	6401
4	CCL20	6364
5	FCER1A	2205
6	CD1A	909
7	GNLY	10578
8	S100A9	6280
9	C4A.B	720, 721

- [0042]
- [0043] 이하에서는 도 4 및 도 5를 이용하여 9개의 차등적 발현 유전자가 완전 관해에 영향력을 주는 유전자인지를 검증한 결과를 설명하기로 한다.
- [0044] 도 4는 도 2에 도시된 S230단계에서 선택한 9개의 차등적 발현 유전자에 대해 검증한 결과를 나타내는 도면이고, 도 5는 도 4에 도시된 검증 결과에 따른 생존곡선을 나타내는 그래프이다.
- [0045] 도 4에 도시된 바와 같이, Cox모형을 통하여 선택된 9개의 차등적 발현 유전자를 적용하여 생존에 대한 완전 관해 예측 결과를 추정할 수 있다.
- [0046] 도 5는 완전 관해가 이루어진 환자 그룹과 완전 관해에 도달하지 않은 환자 그룹 각각에 대하여 시간이 경과함에 따라서 변화하는 사망률을 생존곡선으로 나타낸 그래프이다. 여기서, 완전 관해가 이루어진 환자 그룹의 생존곡선과 완전 관해에 도달하지 않은 환자 그룹의 생존곡선이 서로 평행을 이루므로 9개의 차등적 발현 유전자의 선택이 합당함을 알 수 있다.
- [0047] 그 다음, 모델생성부(130)는 검증된 9개의 차등적 발현 유전자를 이용하여 면역유전자 차등발현 모델을 구축하고 학습시킨다(S240).
- [0048] 모델생성부(130)는 9개의 차등적 발현 유전자 및 발현량 정보를 랜덤 포레스트를 기반으로 학습하여 면역유전자 차등발현 모델을 구축한다. 즉, 모델생성부(130)는 9개의 차등적 발현 유전자 중에서 무작위로 n개의 차등적 발

현 유전자를 선택하여 조합된 복수개의 데이터 셋을 생성한다. 복수개의 데이터셋은 7:3의 비율로 나뉜 다음, 70%에 해당하는 데이터셋은 학습 데이터로 사용하여 면역유전자 차등발현 모델을 학습시킨다.

- [0049] 면역유전자 차등발현 모델은 학습 데이터를 학습하면서 완전 관해를 예측할 수 있는 확률의 오차 범위를 감소시킬 수 있으며, 동시에 정확한 예측이 가능한 레이어 개수를 결정한다.
- [0050] 그리고 모델생성부(130)는 결정된 레이어 개수를 이용하여 구축된 면역유전자 차등발현 모델에 30%에 해당되는 데이터 셋을 입력 데이터로 사용하여 완전 관해 예측이 가능한지 테스트할 수 있다.
- [0051] 이하에서는 도 6을 통하여 본 발명의 실시예에 따른 면역유전자 차등발현 모델의 성능을 평가한 결과에 대하여 설명한다.
- [0052] 도 6은 도 2에 도시된 S240단계에서 구축된 분석모델의 유효성을 평가한 결과를 나타내는 그래프이다.
- [0053] 면역유전자 차등발현 모델에서 발생할 수 있는 분산문제를 해결하기 위하여, 모델생성부(130)는 GridSearch와 5-fold 교차검증을 통하여 분석모델의 유효성을 검증한다.
- [0054] 5-fold 교차검증에 의하면 데이터는 5개의 세트로 나뉘고, 나뉘어진 데이터 중에서 1개의 세트는 훈련용으로 사용되고 나머지 4개의 세트는 테스트용으로 사용된다. 그 다음, 모델생성부(130)는 첫번째 폴드에서 다섯번째 폴드가 진행되는 과정까지 사용하여 얻은 결과의 평균값을 표준편차값으로 사용하여 분산을 제거한다.
- [0055] 도 6에 도시된 바와 같이, 본 발명의 실시예에 따른 완전 관해 분석모델은 AUC(area under the curve)값이 0.71으로 높은 정확성을 보인다.
- [0056] 이하에서는 도 7을 이용하여 본 발명의 실시예에 따라 구축된 면역유전자 차등발현 모델을 이용하여 유방암 환자의 완전 관해 가능성을 예측하는 방법에 대하여 상세하게 설명한다.
- [0057] 도 7은 본 발명의 실시예에 따른 완전 면역유전자 차등발현 모델을 통해 예측하는 방법을 개략적으로 도시한 순서도이다.
- [0058] 도 7에 도시된 바와 같이, 먼저, 수집부(110)는 피검자인 유방암 환자로부터 생검 조직을 수집한다(S710).
- [0059] 그리고, 생검 조직에 대한 정보는 유전자 선택부(120)에 전달된다. 유전자 선택부(120)는 RNA 추출 키트를 사용하여 생검 조직으로부터 유전자 정보를 획득한다(S720).
- [0060] 그 다음 유전자 선택부(120)는 피검자의 유전자 정보로부터 9개의 차등적 발현 유전자의 발현량을 추출한다(S730).
- [0061] 부연하자면, 유전자 선택부(120)는 피검자의 유전자 정보 중에서 9개의 차등적 발현 유전자를 선택한다. 그리고 유전자 선택부(120)는 nCounter®Analysis System을 통해 선택된 9개의 차등적 발현 유전자에 대한 발현량 정보를 추출한다.
- [0062] 그리고, 유전자 선택부(120)는 선택된 9개의 차등적 발현 유전자에 대한 발현량 정보를 예측부(140)에 전달한다.
- [0063] 그러면, 예측부(140)는 완전 관해 분석모델에 9개의 차등적 발현 유전자의 발현량 정보를 입력하여 유방암이 걸린 피검자에 대하여 완전 관해 확률을 예측한다(S740).
- [0064] 이와 같이 본 발명의 실시예에 따른 유방암에 대한 완전 관해 예측 방법은 유방암 내에 면역 관련 유전자 중에서 차등적으로 발현하는 유전자를 추출하고, 추출된 차등적 발현 유전자를 통해 샘플의 수가 적은 데이터로부터 완전 관해를 예측할 수 있는 면역유전자 차등발현 모델을 구축할 수 있다. 따라서, 적은 비용으로 높은 정확성을 통해 유방암 환자의 완전 관해를 예측할 수 있다.
- [0065] 또한, 본 발명의 실시예에 따른 유방암에 대한 완전 관해 예측 방법은 유방암의 예후와 치료반응 예측에 중요한 것으로 대두되고 있는 면역관련 인자를 분석하여 치료제 개발 기반을 마련할 수 있고, 치료제 개발로 인해 유방암의 진료 지침이 개선될 수 있다.
- [0066] 또한, 본 발명의 실시예에 따른 유방암에 대한 완전 관해 예측 방법은 유방암의 치료 반응성 예측 모델 및 예후 예측 모델을 임상에 적용하여 사용할 수 있으며, 모델 개발에 사용된 분석 방법을 다른 암종에도 적용하여 면역 프로파일링 모델의 확장이 가능한 효과를 지닌다.

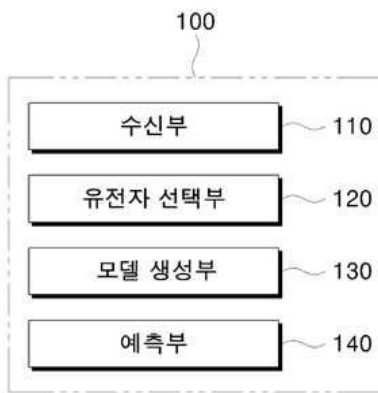
[0068] 본 발명은 도면에 도시된 실시예를 참고로 하여 설명되었으나 이는 예시적인 것에 불과하며, 당해 기술이 속하는 분야에서 통상의 지식을 가진 자라면 이로부터 다양한 변형 및 균등한 타 실시예가 가능하다는 점을 이해할 것이다. 따라서 본 발명의 진정한 기술적 보호범위는 아래의 특허청구범위의 기술적 사상에 의하여 정해져야 할 것이다.

부호의 설명

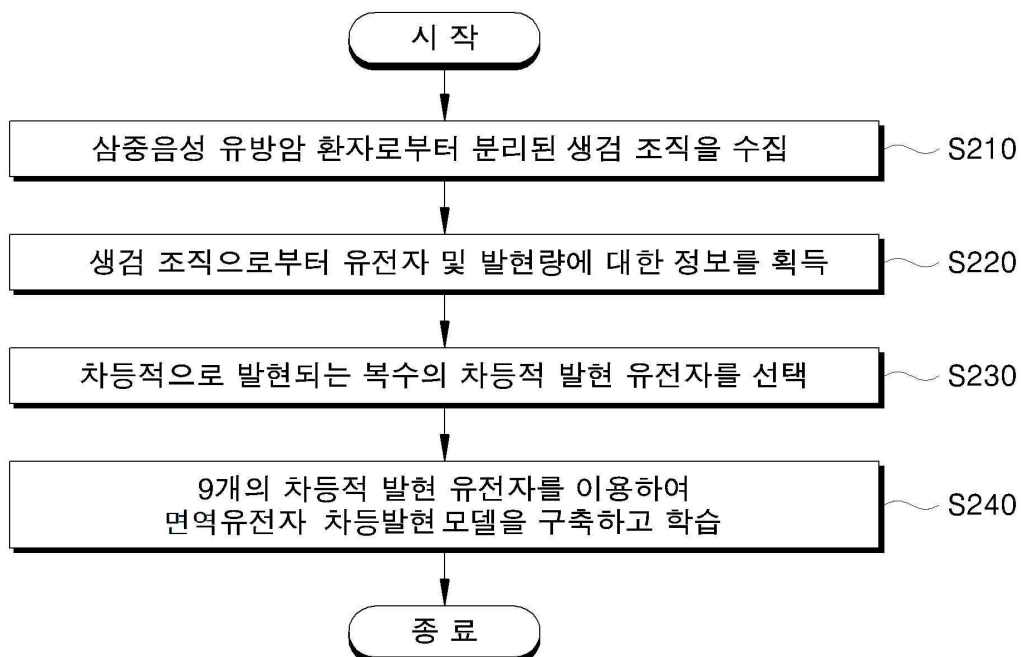
[0069] 100 : 완전 관해 예측 장치 110 : 수집부
120 : 유전자 선택부 130 : 모델생성부
140 : 예측부

도면

도면1



도면2



도면3

Gene	logFC	logCPM	LR	PValue	FDR
IL2RA	1.491432	7.954741	28.42154	9.76E-08	5.65E-05
CCL5	1.516217	11.20899	25.67787	4.03E-07	1.17E-04
SELE	1.571043	7.007868	20.21268	6.93E-06	1.34E-03
CCL20	1.776505	7.594408	18.79271	1.46E-05	2.11E-03
FCER1A	-1.294	6.128195	13.55537	2.32E-04	2.23E-02
CD1A	-1.55661	6.545621	13.28413	2.68E-04	2.23E-02
GNLY	1.159208	8.850026	13.26708	2.70E-04	2.23E-02
S100A9	1.55856	12.88563	12.18762	4.81E-04	3.48E-02
C4A.B	-1.27676	11.04598	11.86916	5.71E-04	3.67E-02

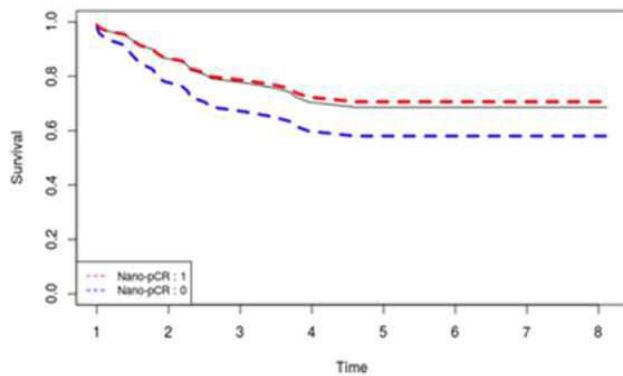
도면4

```
Call:
coxph(formula = Surv(time, DEATH) ~ ., data = j1)

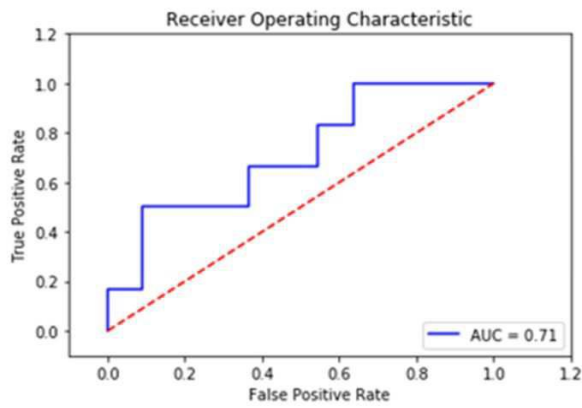
n = 55, number of events = 21

      coef exp(coef) se(coef)      z Pr(>|z|)
CD1A  -0.5757   0.5623   0.2651  -2.172 0.029877 *
S100A9 -0.2732   0.7609   0.1908  -1.432 0.152245
C4A.B   0.3560   1.4276   0.2162   1.647 0.099625 .
GNLY    1.0303   2.8020   0.4229   2.436 0.014831 *
CCL5   -1.6856   0.1853   0.5052  -3.336 0.000849 ***
FCER1A  1.7254   5.6149   0.4387   3.933 8.4e-05 ***
IL2RA  -0.2790   0.7565   0.4753  -0.587 0.557131
CCL20   0.7954   2.2153   0.3836   2.073 0.038135 *
SELE   -1.5349   0.2155   0.5126  -2.994 0.002750 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

도면5



도면6



도면7

