



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2025-0092577
(43) 공개일자 2025년06월24일

(51) 국제특허분류(Int. Cl.)
G01N 33/574 (2006.01) C12Q 1/6886 (2018.01)
(52) CPC특허분류
G01N 33/574 (2013.01)
C12Q 1/6886 (2022.01)
(21) 출원번호 10-2023-0182898
(22) 출원일자 2023년12월15일
심사청구일자 2023년12월15일

(71) 출원인
재단법인 아산사회복지재단
서울특별시 송파구 올림픽로43길 88 (풍납동)
울산대학교 산학협력단
울산광역시 남구 대학로 93(무거동)
(72) 발명자
김원구
서울특별시 송파구 잠실로 62, 316동 1702호
김경곤
서울특별시 송파구 올림픽로 435, 108동 2302호
(74) 대리인
이명진

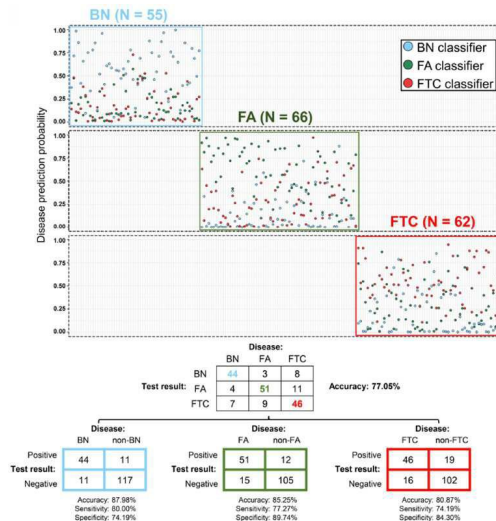
전체 청구항 수 : 총 22 항

(54) 발명의 명칭 갑상선 종양의 감별진단용 바이오마커 및 이의 용도

(57) 요약

본 발명은 갑상선 종양의 감별진단용 바이오마커 및 이의 용도에 관한 것으로서, 본 발명에서는 갑상선 여포암(follicular thyroid cancer, FTC)과 갑상선 여포선종(follicular adenoma, FA) 및 양성 결절(benign nodule, BN) 조직을 이용하여 광범위한 정량, 정성적 단백질체 분석을 수행하여 상기와 같은 갑상선 종양의 3가지 아형 조직에서 특이적으로 발현되는 단백질을 확인하였다. 이에, 본 발명에 따른 상기 단백질은 수술 없이 갑상선 여포암, 갑상선 여포선종, 및 양성 결절을 감별진단하기 위한 바이오마커로서 유용하게 이용될 수 있으며, 이에 따라 불필요한 수술 및 환자의 고통을 줄일 수 있을 것으로 기대된다.

대표도 - 도5i



(52) CPC특허분류

C12Q 2600/158 (2013.01)

G01N 2800/046 (2013.01)

이 발명을 지원한 국가연구개발사업

과제고유번호	1711182469
과제번호	2022R1A2C1004016
부처명	과학기술정보통신부
과제관리(전문)기관명	한국연구재단
연구사업명	개인기초연구(과기정통부)
연구과제명	단백질체 분석을 통한 미결정형 갑상선 결절의 분자진단법 개발
기 여 율	1/1
과제수행기관명	울산대학교
연구기간	2023.03.01 ~ 2024.02.29

명세서

청구범위

청구항 1

CNDP2, DNAAF5, DYNC1H1, FARSF, PDCD4, FAM149B1, RPS9, TAGLN2, UPF1, UQCRC, ACTN4, DSTN, MACROH2A1, NUCB1, SPTAN1, TAGLN, 및 XRCC5로 이루어진 군으로부터 선택된 하나 이상의 단백질을 포함하는, 갑상선 종양의 감별진단용 바이오마커.

청구항 2

CNDP2, DNAAF5, DYNC1H1, FARSF, PDCD4, FAM149B1, RPS9, TAGLN2, UPF1, UQCRC, ACTN4, DSTN, MACROH2A1, NUCB1, SPTAN1, TAGLN, 및 XRCC5로 이루어진 군으로부터 선택된 하나 이상의 단백질 또는 이의 mRNA 발현 수준을 측정하는 제제를 유효성분으로 포함하는, 갑상선 종양의 감별진단용 조성물.

청구항 3

제2항에 있어서,

상기 갑상선 종양은 갑상선 여포암(thyroid carcinoma, FTC), 갑상선 여포선종(follicular adenoma, FA), 및 양성 결절(benign nodule, BN)로 이루어진 군으로부터 선택된 하나 이상인 것을 특징으로 하는, 갑상선 종양의 감별진단용 조성물.

청구항 4

제3항에 있어서,

상기 갑상선 종양의 감별진단은 갑상선 여포암 또는 갑상선 여포선종을 양성 결절과 각각 구별하여 진단하는 것을 특징으로 하는, 갑상선 종양의 감별진단용 조성물.

청구항 5

제2항에 있어서,

상기 제제는 CNDP2, DNAAF5, DYNC1H1, FARSF, PDCD4, FAM149B1, RPS9, TAGLN2, UPF1, UQCRC, ACTN4, DSTN, MACROH2A1, NUCB1, SPTAN1, TAGLN, 및 XRCC5로 이루어진 군으로부터 선택된 하나 이상의 단백질에 특이적인 항체 또는 앵타머인 것을 특징으로 하는, 갑상선 종양의 감별진단용 조성물.

청구항 6

제2항에 있어서,

상기 제제는 CNDP2, DNAAF5, DYNC1H1, FARSF, PDCD4, FAM149B1, RPS9, TAGLN2, UPF1, UQCRC, ACTN4, DSTN, MACROH2A1, NUCB1, SPTAN1, TAGLN, 및 XRCC5로 이루어진 군으로부터 선택된 하나 이상의 단백질의 mRNA에 특이적으로 결합하는 프라이머 또는 프로브인 것을 특징으로 하는, 갑상선 종양의 감별진단용 조성물.

청구항 7

제4항에 있어서,

상기 조성물은 ACTN4, DSTN, MACROH2A1, NUCB1, SPTAN1, TAGLN, 및 XRCC5 단백질 또는 이의 mRNA 발현을 측정하여 갑상선 여포암을 진단하기 위한 것을 특징으로 하는, 갑상선 종양의 감별진단용 조성물.

청구항 8

제4항에 있어서,

상기 조성물은 DNAAF5, FAM149B1, RPS9, TAGLN2, UPF1, 및 UQCRC1 단백질 또는 이의 mRNA 발현을 측정하여 갑상선 여포선종을 진단하기 위한 것을 특징으로 하는, 갑상선 종양의 감별진단용 조성물.

청구항 9

제4항에 있어서,

상기 조성물은 CNDP2, DNAAF5, DYNC1H1, FARSB, 및 PDCD4 단백질 또는 이의 mRNA 발현을 측정하여 양성 결절을 진단하기 위한 것을 특징으로 하는, 갑상선 종양의 감별진단용 조성물.

청구항 10

제2항 내지 제9항 중 어느 한 항의 조성물 및 설명서를 포함하는, 갑상선 종양의 감별진단용 키트.

청구항 11

피검체로부터 분리된 생물학적 시료에서 CNDP2, DNAAF5, DYNC1H1, FARSB, PDCD4, FAM149B1, RPS9, TAGLN2, UPF1, UQCRC, ACTN4, DSTN, MACROH2A1, NUCB1, SPTAN1, TAGLN, 및 XRCC5로 이루어진 군으로부터 선택된 하나 이상의 단백질 또는 이의 mRNA 발현 수준을 측정하는 단계를 포함하는, 갑상선 종양의 감별진단을 위한 정보제공방법.

청구항 12

제11항에 있어서,

상기 갑상선 종양은 갑상선 여포암(thyroid carcinoma, FTC), 갑상선 여포선종(follicular adenoma, FA), 및 양성 결절(benign nodule, BN)로 이루어진 군으로부터 선택된 하나 이상인 것을 특징으로 하는, 갑상선 종양의 감별진단을 위한 정보제공방법.

청구항 13

제11항에 있어서,

상기 생물학적 시료는 혈액, 전혈, 혈장, 소변, 조직, 세포, 기관, 골수, 미세침흡인검체, 미세침세척액, 중심 부바늘생검(core needle biopsy) 검체 및 타액으로 이루어진 군으로부터 선택된 하나 이상인 것을 특징으로 하는, 갑상선 종양의 감별진단을 위한 정보제공방법.

청구항 14

제11항에 있어서,

상기 피검체는 갑상선 여포암 또는 여포선종이 의심되는 환자인 것을 특징으로 하는, 갑상선 종양의 감별진단을 위한 정보제공방법.

청구항 15

제11항에 있어서,

상기 방법은 CNDP2, DNAAF5, DYNC1H1, FARSB, PDCD4, FAM149B1, RPS9, TAGLN2, UPF1, UQCRC, ACTN4, DSTN, MACROH2A1, NUCB1, SPTAN1, TAGLN, 및 XRCC5 단백질 또는 이의 mRNA 발현 수준을 측정하는 단계; 및

상기 ACTN4, DSTN, MACROH2A1, NUCB1, SPTAN1, TAGLN, 및 XRCC5 단백질 또는 이의 mRNA 발현이 검출되거나 증가된 경우, 갑상선 여포암으로 판정하는 단계를 포함하는 것을 특징으로 하는, 갑상선 종양의 감별진단을 위한 정보제공방법.

청구항 16

제11항에 있어서,

상기 방법은 CNDP2, DNAAF5, DYNC1H1, FARSB, PDCD4, FAM149B1, RPS9, TAGLN2, UPF1, UQCRC, ACTN4, DSTN, MACROH2A1, NUCB1, SPTAN1, TAGLN, 및 XRCC5 단백질 또는 이의 mRNA 발현 수준을 측정하는 단계; 및

상기 FAM149B1, RPS9, TAGLN2, UPF1, 및 UQCRC1 단백질 또는 이의 mRNA 발현이 검출되거나 증가된 경우, 또는 DNAAF5 단백질 또는 이의 mRNA 발현이 검출되지 않거나 감소된 경우 갑상선 여포선종으로 판정하는 단계를 포함하는 것을 특징으로 하는, 갑상선 종양의 감별진단을 위한 정보제공방법.

청구항 17

제11항에 있어서,

상기 방법은 CNDP2, DNAAF5, DYNC1H1, FARSB, PDCD4, FAM149B1, RPS9, TAGLN2, UPF1, UQCRC, ACTN4, DSTN, MACROH2A1, NUCB1, SPTAN1, TAGLN, 및 XRCC5 단백질 또는 이의 mRNA 발현 수준을 측정하는 단계; 및

상기 DNAAF5 단백질 또는 이의 mRNA 발현이 검출되거나 증가된 경우, 또는 CNDP2, DYNC1H1, FARSB, 및 PDCD4 단백질 또는 이의 mRNA 발현이 검출되지 않거나 감소된 경우 양성 결절로 판정하는 단계를 포함하는 것을 특징으로 하는, 갑상선 종양의 감별진단을 위한 정보제공방법.

청구항 18

후보물질을 투여한 피검체로부터 분리된 생물학적 시료에서 CNDP2, DNAAF5, DYNC1H1, FARSB, PDCD4, FAM149B1, RPS9, TAGLN2, UPF1, UQCRC, ACTN4, DSTN, MACROH2A1, NUCB1, SPTAN1, TAGLN, 및 XRCC5로 이루어진 군으로부터 선택된 하나 이상의 단백질 또는 이의 mRNA 발현 수준을 측정하는 단계를 포함하는, 갑상선 종양의 치료용 후보물질의 스크리닝 방법.

청구항 19

제18항에 있어서,

상기 갑상선 여포 종양은 갑상선 여포암(thyroid carcinoma, FTC), 갑상선 여포선종(follicular adenoma, FA), 및 양성 결절(benign nodule, BN)로 이루어진 군으로부터 선택된 하나 이상인 것을 특징으로 하는, 갑상선 종양의 치료용 후보물질의 스크리닝 방법.

청구항 20

제18항에 있어서,

상기 방법은 CNDP2, DNAAF5, DYNC1H1, FARSB, PDCD4, FAM149B1, RPS9, TAGLN2, UPF1, UQCRC, ACTN4, DSTN, MACROH2A1, NUCB1, SPTAN1, TAGLN, 및 XRCC5 단백질 또는 이의 mRNA 발현 수준을 측정하는 단계; 및

상기 ACTN4, DSTN, MACROH2A1, NUCB1, SPTAN1, TAGLN, 및 XRCC5 단백질 또는 이의 mRNA 발현이 후보물질 투여 전에 비해 감소된 경우, 상기 후보물질을 갑상선 여포암 치료용 후보물질로 선택하는 단계를 포함하는 것을 특징으로 하는, 갑상선 종양의 치료용 후보물질의 스크리닝 방법.

청구항 21

제18항에 있어서,

상기 방법은 CNDP2, DNAAF5, DYNC1H1, FARSB, PDCD4, FAM149B1, RPS9, TAGLN2, UPF1, UQCRC, ACTN4, DSTN, MACROH2A1, NUCB1, SPTAN1, TAGLN, 및 XRCC5 단백질 또는 이의 mRNA 발현 수준을 측정하는 단계; 및

상기 FAM149B1, RPS9, TAGLN2, UPF1, 및 UQCRC1 단백질 또는 이의 mRNA 발현이 후보물질 투여 전에 비해 감소된 경우, 또는 DNAAF5 단백질 또는 이의 mRNA 발현이 후보물질 투여 전에 비해 증가된 경우 상기 후보물질을 갑상선 여포선종 치료용 후보물질로 선택하는 단계를 포함하는 것을 특징으로 하는, 갑상선 종양의 치료용 후보물질의 스크리닝 방법.

청구항 22

제18항에 있어서,

상기 방법은 CNDP2, DNAAF5, DYNC1H1, FARSB, PDCD4, FAM149B1, RPS9, TAGLN2, UPF1, UQCRC, ACTN4, DSTN, MACROH2A1, NUCB1, SPTAN1, TAGLN, 및 XRCC5 단백질 또는 이의 mRNA 발현 수준을 측정하는 단계; 및

상기 DNAAF5 단백질 또는 이의 mRNA 발현이 후보물질 투여 전에 비해 감소된 경우, 또는 CNDP2, DYNC1H1, FARSB, 및 PDCD4 단백질 또는 이의 mRNA 발현이 후보물질 투여 전에 비해 증가된 경우 상기 후보물질을 양성 결절 치료용 후보물질로 선택하는 단계를 포함하는 것을 특징으로 하는, 갑상선 종양의 치료용 후보물질의 스크리닝 방법.

발명의 설명

기술 분야

[0001] 본 발명은 갑상선 종양의 감별진단용 바이오마커 및 이의 용도에 관한 것이다.

배경 기술

[0003] 갑상선 여포암(follicular thyroid carcinoma, FTC)은 두 번째로 흔한 갑상선암으로 전체 갑상선 악성 종양의 10~15 %를 차지한다. FTC의 진단 특징은 세포학 검사만으로는 식별할 수 없는 종양 세포의 피막 또는 혈관 침윤으로서, 갑상선 결절의 가는바늘흡인(Fine needle aspiration, FNA)은 FTC를 갑상선 여포선종(follicular adenoma, FA) 또는 갑상선 양성 결절(benign nodule, BN)과 구별할 수 없다. 이는 FA와 BN은 FTC와 중복되는 세포학적 특징을 공유할 수 있기 때문이다.

[0004] 현재 갑상선 여포암을 확진하기 위하여서는 수술을 통하여 갑상선 절제술을 시행한 후 조직을 채취하여 조직 내 병리를 확인하는 방법이 유일하다. 이처럼 수술을 포함하는 병리 검사는 우선 결절이 발견된 갑상선의 절반을 수술로 절제하여 조직 검사를 실시하고, 이후 갑상선 여포암으로 진단되면 예후가 좋은 초기암을 제외하고는 다시 수술을 통해 나머지 절반을 다시 제거해내야 하기 때문에 매우 번거로운 방법이다. 또한, 암이 아닌 경우에

도 어쩔 수 없이 진단적인 수술을 받아야 하는 환자들이 생기고 일부 환자들은 평생 갑상선 호르몬제를 복용해야 할 수도 있다. 갑상선 세포병리학 보고를 위한 2017 Bethesda 시스템에 기초하여, FTC를 진단받은 환자들의 수술 전 세포병리 검사 결과는 일반적으로 카테고리 III(비정형(atypia of undetermined significance, AUS)/미결정형 여포 병변(follicular lesion of undetermined significance, FLUS)) 또는 IV(여포 종양(follicular neoplasm, FN)/여포 종양 의심(suspicious for a follicular neoplasm, SFN))인 경우가 대부분이다. 이러한 세포학적 범주는 실제 악성 종양 위험의 6~18 % 및 10~40 %와 각각 연관되어 있고, 이는 대다수의 환자가 불필요한 수술을 받고 있음을 나타낸다.

[0005] 따라서, 수술 전 FNA의 제한된 진단 정확도를 고려하여 FTC를 FA 또는 BN과 구별하기 위한 새로운 바이오마커에 대한 개발이 필요하다. 특히, 단백질 기반의 바이오마커를 개발하게 되면, 항체 등을 이용하여 해당 마커의 발현 수준을 측정하여 신속하게 진단에 필요한 정보를 도출할 수 있으며, 병리 검사를 위한 수술을 할 필요가 없다.

[0006] 갑상선 포르말린 고정 파라핀 포매(FFPE) 조직 샘플에서 발현된 모든 단백질의 정량 분석인 프로테오믹스는 새로운 분자 바이오마커를 식별하기 위한 강력한 도구이며 갑상선 결절 환자의 치료에 대한 의사결정에 관련된 여러 생물학적 경로를 반영하는 분자 시그니처의 검출을 가능하게 한다. 또한 요오드화, 인산화 또는 아세틸화와 같은 번역 후 변형 분석을 포함하는 질량 분석 기반 프로테오믹스는 갑상선 종양에 대한 보다 포괄적인 설명을 제공할 수 있다. 또한, 조직 프로테옴 분석은 더 빠른 시간(<2일)으로 작은 샘플(100 µg 미만 조직)을 사용하므로 기존의 바이오마커 및 진단 방법의 한계를 극복할 수 있다. 그러나, FTC에 대하여 신뢰할 수 있는 바이오마커에 대해서는 잘 알려진 바가 없다.

[0007] 이에, 본 발명자들은 갑상선 여포암, 갑상선 여포선종, 및 양성 결절의 FFPE 조직에 대한 전반적인 프로테옴 분석을 수행하여 FTC, FA, 및 BN에서 차등 발현되는 단백질을 확인함으로써 수술 없이도 FTC, FA, 및 BN을 구별하여 진단할 수 있는 바이오마커로 활용하고자 하였다.

선행기술문헌

비특허문헌

[0009] (비특허문헌 0001) Aschebrook-Kilfoy B, Grogan RH, Ward MH, Kaplan E, Devesa SS. Follicular thyroid cancer incidence patterns in the United States, 1980-2009. *Thyroid* 2013;23:1015-21.

발명의 내용

해결하려는 과제

[0010] 본 발명자들은 갑상선 여포암(follicular thyroid cancer, FTC)과 갑상선 여포선종(follicular adenoma, FA) 및 양성 결절(benign nodule, BN) 조직을 이용하여 광범위한 정량, 정성적 단백질체 분석을 수행한 결과, 상기와 같은 갑상선 종양의 3가지 아형 조직에서 특이적으로 발현되는 단백질을 발견하였는바, 이에 기초하여 본 발명을 완성하였다.

[0011] 이에, 본 발명의 목적은 CNDP2, DNAAF5, DYNC1H1, FARSB, PDCD4, FAM149B1, RPS9, TAGLN2, UPF1, UQCRC, ACTN4, DSTN, MACROH2A1, NUCB1, SPTAN1, TAGLN, 및 XRCC5로 이루어진 군으로부터 선택된 하나 이상의 단백질을 포함하는, 갑상선 종양의 감별진단용 바이오마커를 제공하는 것이다.

[0012] 본 발명의 다른 목적은 CNDP2, DNAAF5, DYNC1H1, FARSB, PDCD4, FAM149B1, RPS9, TAGLN2, UPF1, UQCRC, ACTN4, DSTN, MACROH2A1, NUCB1, SPTAN1, TAGLN, 및 XRCC5로 이루어진 군으로부터 선택된 하나 이상의 단백질 또는 이의 mRNA 발현 수준을 측정하는 제제를 유효성분으로 포함하는, 갑상선 종양의 감별진단용 조성물 또는 키트를 제공하는 것이다.

[0013] 본 발명의 또 다른 목적은 피검체로부터 분리된 생물학적 시료에서 CNDP2, DNAAF5, DYNC1H1, FARSB, PDCD4, FAM149B1, RPS9, TAGLN2, UPF1, UQCRC, ACTN4, DSTN, MACROH2A1, NUCB1, SPTAN1, TAGLN, 및 XRCC5로 이루어진 군으로부터 선택된 하나 이상의 단백질 또는 이의 mRNA 발현 수준을 측정하는 단계를 포함하는, 갑상선 종양의

감별진단을 위한 정보제공방법을 제공하는 것이다.

[0014] 본 발명의 또 다른 목적은 후보물질을 투여한 피검체로부터 분리된 생물학적 시료에서 CNBP2, DNAAF5, DYNC1H1, FARSB, PDCD4, FAM149B1, RPS9, TAGLN2, UPF1, UQCRC, ACTN4, DSTN, MACROH2A1, NUCB1, SPTAN1, TAGLN, 및 XRCC5로 이루어진 군으로부터 선택된 하나 이상의 단백질 또는 이의 mRNA 발현 수준을 측정하는 단계를 포함하는, 갑상선 종양의 치료용 후보물질의 스크리닝 방법을 제공하는 것이다.

[0015] 그러나, 본 발명이 이루고자 하는 기술적 과제는 이상에서 언급한 과제에 제한되지 않으며, 언급되지 않은 또 다른 과제들은 아래의 기재로부터 본 발명이 속하는 기술 분야의 통상의 지식을 가진 자에게 명확하게 이해될 수 있을 것이다.

과제의 해결 수단

[0017] 상기와 같은 목적을 달성하기 위해 본 발명은 CNBP2, DNAAF5, DYNC1H1, FARSB, PDCD4, FAM149B1, RPS9, TAGLN2, UPF1, UQCRC, ACTN4, DSTN, MACROH2A1, NUCB1, SPTAN1, TAGLN, 및 XRCC5로 이루어진 군으로부터 선택된 하나 이상의 단백질을 포함하는, 갑상선 종양의 감별진단용 바이오마커를 제공한다.

[0018] 또한, 본 발명은 CNBP2, DNAAF5, DYNC1H1, FARSB, PDCD4, FAM149B1, RPS9, TAGLN2, UPF1, UQCRC, ACTN4, DSTN, MACROH2A1, NUCB1, SPTAN1, TAGLN, 및 XRCC5로 이루어진 군으로부터 선택된 하나 이상의 단백질 또는 이의 mRNA 발현 수준을 측정하는 체제를 유효성분으로 포함하는, 갑상선 종양의 감별진단용 조성물을 제공한다.

[0019] 본 발명의 일 구현예로서, 상기 갑상선 종양은 갑상선 여포암(thyroid carcinoma, FTC), 갑상선 여포선종(follicular adenoma, FA), 및 양성 결절(benign nodule, BN)로 이루어진 군으로부터 선택된 하나 이상일 수 있으나, 이에 제한되지 않는다.

[0020] 본 발명의 다른 구현예로서, 상기 갑상선 종양의 감별진단은 갑상선 여포암 또는 갑상선 여포선종을 양성 결절과 각각 구별하여 진단하기 위한 것일 수 있으나, 이에 제한되지 않는다.

[0021] 본 발명의 또 다른 구현예로서, 상기 체제는 CNBP2, DNAAF5, DYNC1H1, FARSB, PDCD4, FAM149B1, RPS9, TAGLN2, UPF1, UQCRC, ACTN4, DSTN, MACROH2A1, NUCB1, SPTAN1, TAGLN, 및 XRCC5로 이루어진 군으로부터 선택된 하나 이상의 단백질에 특이적인 항체 또는 앵타머일 수 있으나, 이에 제한되지 않는다.

[0022] 본 발명의 또 다른 구현예로서, 상기 체제는 CNBP2, DNAAF5, DYNC1H1, FARSB, PDCD4, FAM149B1, RPS9, TAGLN2, UPF1, UQCRC, ACTN4, DSTN, MACROH2A1, NUCB1, SPTAN1, TAGLN, 및 XRCC5로 이루어진 군으로부터 선택된 하나 이상의 단백질의 mRNA에 특이적으로 결합하는 프라이머 또는 프로브일 수 있으나, 이에 제한되지 않는다.

[0023] 본 발명의 또 다른 구현예로서, 상기 조성물은 ACTN4, DSTN, MACROH2A1, NUCB1, SPTAN1, TAGLN, 및 XRCC5 단백질 또는 이의 mRNA 발현을 측정하여 갑상선 여포암을 진단하기 위한 것일 수 있으나, 이에 제한되지 않는다.

[0024] 본 발명의 또 다른 구현예로서, 상기 조성물은 DNAAF5, FAM149B1, RPS9, TAGLN2, UPF1, 및 UQCRC1 단백질 또는 이의 mRNA 발현을 측정하여 갑상선 여포선종을 진단하기 위한 것일 수 있으나, 이에 제한되지 않는다.

[0025] 본 발명의 또 다른 구현예로서, 상기 조성물은 CNBP2, DNAAF5, DYNC1H1, FARSB, 및 PDCD4 단백질 또는 이의 mRNA 발현을 측정하여 양성 결절을 진단하기 위한 것일 수 있으나, 이에 제한되지 않는다.

[0026] 또한, 본 발명은 상기 갑상선 종양의 감별진단용 조성물 및 설명서를 포함하는, 갑상선 종양의 감별진단용 키트를 제공한다.

[0027] 또한, 본 발명은 피검체로부터 분리된 생물학적 시료에서 CNBP2, DNAAF5, DYNC1H1, FARSB, PDCD4, FAM149B1, RPS9, TAGLN2, UPF1, UQCRC, ACTN4, DSTN, MACROH2A1, NUCB1, SPTAN1, TAGLN, 및 XRCC5로 이루어진 군으로부터 선택된 하나 이상의 단백질 또는 이의 mRNA 발현 수준을 측정하는 단계를 포함하는, 갑상선 종양의 감별진단을 위한 정보제공방법을 제공한다.

[0028] 본 발명의 일 구현예로서, 상기 생물학적 시료는 혈액, 전혈, 혈장, 소변, 조직, 세포, 기관, 골수, 미세침흡인검체, 미세침세척액, 중심부바늘생검(core needle biopsy) 검체 및 타액으로 이루어진 군으로부터 선택된 하나 이상일 수 있으나, 이에 제한되지 않는다.

[0029] 본 발명의 다른 구현예로서, 상기 피검체는 갑상선 여포암 또는 여포선종이 의심되는 환자일 수 있으나, 이에

제한되지 않는다.

- [0030] 본 발명의 또 다른 구현예로서, 상기 방법은 CNBP2, DNAAF5, DYNC1H1, FARSB, PDCD4, FAM149B1, RPS9, TAGLN2, UPF1, UQCRC, ACTN4, DSTN, MACROH2A1, NUCB1, SPTAN1, TAGLN, 및 XRCC5 단백질 또는 이의 mRNA 발현 수준을 측정하는 단계; 및
- [0031] 상기 ACTN4, DSTN, MACROH2A1, NUCB1, SPTAN1, TAGLN, 및 XRCC5 단백질 또는 이의 mRNA 발현이 검출되거나 증가된 경우, 갑상선 여포암으로 판정하는 단계를 포함할 수 있으나, 이에 제한되지 않는다.
- [0032] 본 발명의 또 다른 구현예로서, 상기 방법은 CNBP2, DNAAF5, DYNC1H1, FARSB, PDCD4, FAM149B1, RPS9, TAGLN2, UPF1, UQCRC, ACTN4, DSTN, MACROH2A1, NUCB1, SPTAN1, TAGLN, 및 XRCC5 단백질 또는 이의 mRNA 발현 수준을 측정하는 단계; 및
- [0033] 상기 FAM149B1, RPS9, TAGLN2, UPF1, 및 UQCRC1 단백질 또는 이의 mRNA 발현이 검출되거나 증가된 경우, 또는 DNAAF5 단백질 또는 이의 mRNA 발현이 검출되지 않거나 감소된 경우 갑상선 여포선종으로 판정하는 단계를 포함할 수 있으나, 이에 제한되지 않는다.
- [0034] 본 발명의 또 다른 구현예로서, 상기 방법은 CNBP2, DNAAF5, DYNC1H1, FARSB, PDCD4, FAM149B1, RPS9, TAGLN2, UPF1, UQCRC, ACTN4, DSTN, MACROH2A1, NUCB1, SPTAN1, TAGLN, 및 XRCC5 단백질 또는 이의 mRNA 발현 수준을 측정하는 단계; 및
- [0035] 상기 DNAAF5 단백질 또는 이의 mRNA 발현이 검출되거나 증가된 경우, 또는 CNBP2, DYNC1H1, FARSB, 및 PDCD4 단백질 또는 이의 mRNA 발현이 검출되지 않거나 감소된 경우 양성 결절로 판정하는 단계를 포함할 수 있으나, 이에 제한되지 않는다.
- [0036] 또한, 본 발명은 후보물질을 투여한 피검체로부터 분리된 생물학적 시료에서 CNBP2, DNAAF5, DYNC1H1, FARSB, PDCD4, FAM149B1, RPS9, TAGLN2, UPF1, UQCRC, ACTN4, DSTN, MACROH2A1, NUCB1, SPTAN1, TAGLN, 및 XRCC5로 이루어진 군으로부터 선택된 하나 이상의 단백질 또는 이의 mRNA 발현 수준을 측정하는 단계를 포함하는, 갑상선 종양의 치료용 후보물질의 스크리닝 방법을 제공한다.
- [0037] 본 발명의 또 다른 구현예로서, 상기 방법은 CNBP2, DNAAF5, DYNC1H1, FARSB, PDCD4, FAM149B1, RPS9, TAGLN2, UPF1, UQCRC, ACTN4, DSTN, MACROH2A1, NUCB1, SPTAN1, TAGLN, 및 XRCC5 단백질 또는 이의 mRNA 발현 수준을 측정하는 단계; 및
- [0038] 상기 ACTN4, DSTN, MACROH2A1, NUCB1, SPTAN1, TAGLN, 및 XRCC5 단백질 또는 이의 mRNA 발현이 후보물질 투여 전에 비해 감소된 경우, 상기 후보물질을 갑상선 여포암 치료용 후보물질로 선택하는 단계를 포함할 수 있으나, 이에 제한되지 않는다.
- [0039] 본 발명의 다른 구현예로서, 상기 방법은 CNBP2, DNAAF5, DYNC1H1, FARSB, PDCD4, FAM149B1, RPS9, TAGLN2, UPF1, UQCRC, ACTN4, DSTN, MACROH2A1, NUCB1, SPTAN1, TAGLN, 및 XRCC5 단백질 또는 이의 mRNA 발현 수준을 측정하는 단계; 및
- [0040] 상기 FAM149B1, RPS9, TAGLN2, UPF1, 및 UQCRC1 단백질 또는 이의 mRNA 발현이 후보물질 투여 전에 비해 감소된 경우, 또는 DNAAF5 단백질 또는 이의 mRNA 발현이 후보물질 투여 전에 비해 증가된 경우 상기 후보물질을 갑상선 여포선종 치료용 후보물질로 선택하는 단계를 포함할 수 있으나, 이에 제한되지 않는다.
- [0041] 본 발명의 일 구현예로서, 상기 방법은 CNBP2, DNAAF5, DYNC1H1, FARSB, PDCD4, FAM149B1, RPS9, TAGLN2, UPF1, UQCRC, ACTN4, DSTN, MACROH2A1, NUCB1, SPTAN1, TAGLN, 및 XRCC5 단백질 또는 이의 mRNA 발현 수준을 측정하는 단계; 및
- [0042] 상기 DNAAF5 단백질 또는 이의 mRNA 발현이 후보물질 투여 전에 비해 감소된 경우, 또는 CNBP2, DYNC1H1, FARSB, 및 PDCD4 단백질 또는 이의 mRNA 발현이 후보물질 투여 전에 비해 증가된 경우 상기 후보물질을 양성 결절 치료용 후보물질로 선택하는 단계를 포함할 수 있으나, 이에 제한되지 않는다.
- [0043] 또한, 본 발명은 CNBP2, DNAAF5, DYNC1H1, FARSB, PDCD4, FAM149B1, RPS9, TAGLN2, UPF1, UQCRC, ACTN4, DSTN, MACROH2A1, NUCB1, SPTAN1, TAGLN, 및 XRCC5로 이루어진 군으로부터 선택된 하나 이상의 단백질 또는 이의 mRNA 발현 수준을 측정하는 제제를 유효성분으로 포함하는 조성물의 갑상선 종양의 감별진단 용도를 제공한다.

[0044] 또한, 본 발명은 CNDP2, DNAAF5, DYNC1H1, FARSB, PDCD4, FAM149B1, RPS9, TAGLN2, UPF1, UQCRC, ACTN4, DSTN, MACROH2A1, NUCB1, SPTAN1, TAGLN, 및 XRCC5로 이루어진 군으로부터 선택된 하나 이상의 단백질 또는 이의 mRNA 발현 수준을 측정하는 제제를 유효성분으로 포함하는 조성물의, 갑상선 종양의 감별진단용 제제의 제조를 위한 용도를 제공한다.

발명의 효과

[0046] 본 발명에서는 갑상선 여포암(follicular thyroid cancer, FTC)과 갑상선 여포선종(follicular adenoma, FA), 및 양성 결절(benign nodule, BN) 조직을 이용하여 광범위한 정량, 정성적 단백질체 분석을 수행하여 상기와 같은 갑상선 여포 종양의 3가지 아형 조직에서 특이적으로 발현되는 단백질을 확인하였다. 이에, 본 발명에 따른 상기 단백질은 수술 없이 갑상선 여포암, 갑상선 여포선종, 및 양성 결절을 감별진단하기 위한 바이오마커로서 유용하게 이용될 수 있으며, 이에 따라 불필요한 수술 및 환자의 고통을 줄일 수 있을 것으로 기대된다.

도면의 간단한 설명

[0048] 도 1은 본 발명의 일 구현예에 따른 갑상선 FFPE 조직 단백질체 시료 전처리 및 분석 워크플로어를 나타낸 도면이다.

도 2a는 본 발명의 일 구현예에 따른 갑상선 FFPE 조직에서 확인된 MS 스펙트럼의 수를 나타낸 도면이다.

도 2b는 본 발명의 일 구현예에 따른 오픈 SSM 및 PSM 서치 및 클로즈 SSM 및 PSM 서치를 통해 확인된 단백질 수를 벤 다이어그램으로 나타낸 도면이다.

도 2c는 본 발명의 일 구현예에 따른 오픈 SSM 서치 결과에 의한 비표지 정량화를 기반으로 한 정규화된 단백질 존재비의 분포를 나타낸 도면이다.

도 2d는 본 발명의 일 구현예에 따른 티로글로불린(TG)에서 단백질 수와 TG 일치 스펙트럼의 백분율에 대한 산점도를 나타낸 도면이고, 내부 그래프는 PSM과 SSM으로 식별된 단백질의 수를 나타낸 도면이다.

도 2e는 본 발명의 일 구현예에 따른 단기 보관 샘플(1~5년) 및 장기 보관 샘플(6~11년)에서 확인된 단백질 수 및 단백질 번역 후 변형 5가지(산화 또는 수산화, 메틸화, 요오드화, 포름알데이드 부가물, 및 티로신 산화)의 백분율을 박스플롯으로 나타낸 도면이다.

도 2f는 본 발명의 일 구현예에 따른 샘플 보관 시간이 다른 두 샘플 세트의 단백질 정량화에 대한 산점도를 나타낸 도면이다.

도 3a는 본 발명의 일 구현예에 따른 갑상선 샘플의 PLS-DA 분석을 통한 양성 결절(BN), 갑상선 여포선종(FA), 및 갑상선 여포암(FTC)의 세 가지 갑상선 여포 종양의 아형에 대한 2차원 점수 플롯을 나타낸 도면이다.

도 3b는 본 발명의 일 구현예에 따른 갑상선 샘플의 PLS-DA 분석을 통해 분리된 각각의 구성 요소에서 중요도가 높은 상위 20개 단백질을 나타낸 도면이다.

도 3c는 본 발명의 일 구현예에 따른 세 가지 갑상선 여포 종양의 아형에 대한 KEGG 경로의 해당과정/포도당 신생합성의 분자맵으로 식별된 단백질의 상대적 정량값을 나타낸 도면이다.

도 3d는 본 발명의 일 구현예에 따른 MSigDB의 50개 특징 유전자 세트 중 ANOVA 테스트에서 p -값 <0.05 인 22개의 특징 유전자 세트에 대한 Hierarchical Spearman의 순위 상관 관계 클러스터링 결과를 나타낸 도면이다.

도 3e는 본 발명의 일 구현예에 따른 singscore 알고리즘에 의한 단일 샘플 점수를 사용하여 22가지 중 18가지 특징 신호 경로에 대한 세 가지 갑상선 여포 종양 아형의 상관 관계를 박스플롯으로 나타낸 도면이다.

도 3f는 본 발명의 일 구현예에 따른 singscore 알고리즘에 의한 단일 샘플 점수를 사용하여 대표적인 4가지 특징 신호 경로에 대한 세 가지 갑상선 여포 종양 아형의 상관 관계를 박스플롯으로 나타낸 도면이다.

도 4a 내지 4c는 본 발명의 일 구현예에 따른 BN 및 FA(도 4a), BN 및 FTC(도 4b), 및 FA 및 FTC(도 4c) 사이의 정량적 단백질체 분석 결과를 볼케이노 플롯으로 나타낸 도면이다.

도 4d는 본 발명의 일 구현예에 따른 BN, FTC, 및 FA의 차등 발현 단백질(DEP)에 대한 벤 다이어그램 결과를 나타낸 도면이다.

도 4e는 본 발명의 일 구현예에 따른 FA에 대한 FTC 상향 조절 DEP의 KEGG 및 유전자 온톨로지(GO) 용어를 나타낸 도면이다.

도 5a는 본 발명의 일 구현예에 따른 갑상선 샘플에서 정량화된 단백질에 대하여 BN, FTC, 및 FA 그룹에서 각각 ROC 분석을 수행한 결과를 나타낸 도면이다.

도 5b 내지 5d는 본 발명의 일 구현예에 따른 BN(도 5b), FA(도 5c), 및 FTC(도 5d) 그룹에서 AUC를 기반으로 한 예측 모델에서 선택된 단백질을 나타낸 도면이다.

도 5e내지 5g는 본 발명의 일 구현예에 따른 FTC(도 5e), FA(도 5f), 및 BN(도 5g) 그룹에서 진단 마커 각각의 발현 수준을 확인한 결과를 나타낸 도면이다.

도 5h는 본 발명의 일 구현예에 따른 BN, FA, 및 FTC 예측 모델에서 선택된 단백질의 조합에 따른 ROC 분석 결과를 나타낸 도면이다.

도 5i는 본 발명의 일 구현예에 따른 전체 샘플의 3가지 분류에 의한 질병 예측 점수와 이에 따라 예측된 질병 예측 결과를 나타낸 도면이다.

발명을 실시하기 위한 구체적인 내용

- [0049] 본 발명자들은 갑상선 포르말린 고정 파라핀 포매(FFPE) 조직 샘플에 대하여 단백질체 분석을 수행함으로써, 갑상선 종양의 3가지 아형인 갑상선 여포암(follicular thyroid cancer, FTC), 갑상선 여포선종(follicular adenoma, FA), 및 양성 결절(benign nodule, BN)을 구별하여 진단할 수 있는 단백질 바이오마커를 확인하였는바, 이에 기초하여 본 발명을 완성하였다.
- [0051] 이하, 본 발명을 상세히 설명한다.
- [0053] 본 발명은 CNDP2, DNAAF5, DYNC1H1, FARSB, PDCD4, FAM149B1, RPS9, TAGLN2, UPF1, UQCRC, ACTN4, DSTN, MACROH2A1, NUCB1, SPTAN1, TAGLN, 및 XRCC5로 이루어진 군으로부터 선택된 하나 이상의 단백질을 포함하는, 갑상선 종양의 감별진단용 바이오마커를 제공한다.
- [0054] 또한, 본 발명은 CNDP2, DNAAF5, DYNC1H1, FARSB, PDCD4, FAM149B1, RPS9, TAGLN2, UPF1, UQCRC, ACTN4, DSTN, MACROH2A1, NUCB1, SPTAN1, TAGLN, 및 XRCC5로 이루어진 군으로부터 선택된 하나 이상의 단백질 또는 이의 mRNA 발현 수준을 측정하는 제제를 유효성분으로 포함하는, 갑상선 종양의 감별진단용 조성물을 제공한다.
- [0055] 또한, 본 발명은 CNDP2, DNAAF5, DYNC1H1, FARSB, PDCD4, FAM149B1, RPS9, TAGLN2, UPF1, UQCRC, ACTN4, DSTN, MACROH2A1, NUCB1, SPTAN1, TAGLN, 및 XRCC5 단백질을 또는 이의 mRNA 발현 수준을 측정하는 제제를 유효성분으로 포함하는, 갑상선 종양의 감별진단용 조성물을 제공한다.
- [0056] 본 발명에 있어서, “바이오마커”는 정상이나 병적인 상태를 구분할 수 있거나 치료반응을 예측할 수 있고 객관적으로 측정이 가능한 표지자를 의미한다.
- [0057] 본 발명에 있어서, “갑상선 종양(thyroid neoplasm)”은 갑상선에 발생할 수 있는 모든 종양으로서 갑상선 세포가 비정상적으로 자라서 덩어리를 이루는 것을 의미하며, “갑상선 결절(thyroid nodule)”과 동일한 의미로 사용된다. 본 발명에 있어서 상기 갑상선 종양은 악성인 갑상선 여포암(thyroid carcinoma, FTC), 양성인 갑상선 여포선종(follicular adenoma, FA), 및 양성 결절(benign nodule, BN)을 모두 포함하는 개념으로 사용될 수 있다.
- [0058] 본 발명에 있어서, “단백질(protein)”은 “폴리펩타이드(polypeptide)” 또는 “펩타이드(peptide)”와 호환성 있게 사용되며, 예컨대, 자연 상태의 단백질에서 일반적으로 발견되는 바와 같이 아미노산 잔기의 중합체를 말한다. “폴리뉴클레오티드(polynucleotide)” 또는 “핵산”은 단일-또는 이중-가닥의 형태로 된 데옥시리보뉴클레오티드(DNA) 또는 리보뉴클레오티드(RNA)를 말한다. 다른 제한이 없는 한, 자연적으로 생성되는 뉴클레오티드와 비슷한 방법으로 핵산에 혼성화되는 자연적 뉴클레오티드의 공지된 아날로그도 포함된다. “mRNA”는 단

백질 합성 과정에서 특정 유전자로부터 아미노산 서열을 특정하게 되는 리보솜으로 유전 정보(유전자 특이적 염기 서열)를 전달하는 RNA를 의미한다.

- [0059] 본 발명에 있어서, 상기 단백질은 하기와 같은 Uniprot Accession의 아미노산 서열로 이루어지거나 이를 포함하는 것일 수 있으나, 이에 제한되지 않는다:
- [0060] CNDP2(Cytosolic non-specific dipeptidase): Q96KP4,
- [0061] DNAAF5(Dynein axonemal assembly factor 5): Q86Y56,
- [0062] DYNC1H1(Cytoplasmic dynein 1 heavy chain 1): Q14204,
- [0063] FARSB(Phenylalanine--tRNA ligase beta subunit): Q9NSD9,
- [0064] PDCD49(Programmed cell death protein 4): Q53EL6,
- [0065] FAM149B1(Primary cilium assembly protein FAM149B1): Q96BN6,
- [0066] RPS9(Small ribosomal subunit protein uS4): P46781,
- [0067] TAGLN2(Transgelin-2): P37802,
- [0068] UPF1(Regulator of nonsense transcripts 1): Q92900,
- [0069] UQCRC(Cytochrome b-c1 complex subunit 1, mitochondrial): P31930,
- [0070] ACTN4(Alpha-actinin-4): O43707,
- [0071] DSTN(Destrin): P60981,
- [0072] MACROH2A1(Core histone macro-H2A.1): O75367,
- [0073] NUCB1(Nucleobindin-1): Q02818,
- [0074] SPTAN1(Spectrin alpha chain, non-erythrocytic 1): Q13813
- [0075] TAGLN(Transgelin): Q01995, 및
- [0076] XRCC5(X-ray repair cross-complementing protein 5): P13010.
- [0077] 본 발명에 있어서, “발현(expression)”은 세포에서 단백질 또는 핵산이 생성되는 것을 의미한다.
- [0078] 본 발명에 있어서, “감별진단(differential diagnosis)”은 특정한 질환 또는 상태를 유사한 증상을 나타내는 다른 것으로부터 구별하는 것을 지칭한다. 감별진단 방법은 다중 대안이 가능한 상태의 존재를 확인하는데 사용되는 체계적 진단 방법이다. 이러한 방법은 본질적으로 후보 상태의 확률을 무시할만한 수준으로 축소시키는 제거 프로세스 또는 정보 수득 프로세스이다. 해당 용어는 또한 갑상선 종양이 갑상선 여포암인지, 갑상선 여포선종인지, 또는 양성 결절인지 여부를 결정하는 것을 언급하기 위해 사용된다. 본 발명에 있어서, 갑상선 종양의 감별진단은 갑상선 여포암 또는 갑상선 여포선종을 양성 결절과 각각 구별하여 진단하거나, 갑상선 여포암 또는 갑상선 여포선종을 서로 구별하여 진단하기 위한 것일 수 있으나, 이에 제한되지 않는다.
- [0079] 본 발명에 있어서, “갑상선 여포암(follicular thyroid carcinoma, FTC)”은 갑상선 유두암(papillary thyroid carcinoma, PTC) 다음으로 흔한 갑상선암으로 약 10 % 정도에서 발생하고 있고, 요오드 결핍지역에서 발생률이 높은 경향이 있다. 여포암은 유두암과 같은 고분화 갑상선암(well-differentiated thyroid cancer)으로 분류되고 있지만 유두암보다 고령에서 발생하고 좀 더 공격적인 성향을 보이는 등 다른 임상병리학적 특징을 보이고 있다. 여포암은 갑상선 상피에서 발생하는 악성 종양으로 피막, 조직 내 림프관, 혈관 또는 주위 조직으로의 침윤 및 원격전이 여부 등으로 조직학적 진단이 가능하므로 유두암과 달리 수술 전 가는바늘흡인검사(fine needle aspiration cytology, FNA)나 수술 중 동결절편 검사만으로 정확한 진단이 거의 불가능하다. 여포암, 여포선종 등이 수술 전 세포흡인검사상 여포성 종양(follicular neoplasm)으로 진단될 수 있다. 여포암은 보통 종양 내 국한된 피막 또는 혈관 침범 또는 두 가지 모두를 침범하는 미세 침윤형(minimally invasive type)과 육안적으로 피막침범 소견이 보이거나 갑상선 주변 조직으로의 침범을 특징으로 하는 광범위 침윤형(widely invasive type)으로 분류한다. 이 중 미세 침윤형 여포암 중에서도 피막이나 혈관 침윤에 따라 서로 다른 예후를 갖는다는 보고들도 있고, 피막침범에 관계없이 혈관침범이 있는 미세 침윤형 여포암을 중등도 침윤형으로 추가로 분류하는 경우도 있다. 그러나 갑상선 여포암은 현재까지 수술 전 진단 할 수 있는 방법이 없어 임상적으

로 의심되는 경우 수술을 통해 진단하게 되나, 진단적 수술에서 여포암으로 확인되는 경우는 10~40 %에 불과하다. 더욱이, 수술 후 여포암이 진단되지 않은 환자들 중 상당수는 불필요한 수술 이후 평생 동안 갑상선 호르몬제 보충이 필요하며, 수술 후 여포암이 진단된 환자들은 병기/재발 위험도에 따라 추가적인 잔존 갑상선 절제술 및 동위원소 치료가 필요하게 된다는 단점이 있다. 다양한 면역염색 및 유전자 분석이 수술 전 진단에 적용되고 있으나, 현재까지도 명확하게 입증된 방법은 없으며, 최근 연구에 따르면, 크기의 변화로도 악성과 양성의 구별이 불가능한 실정이다.

[0080] 본 발명에 있어서, “갑상선 여포선종(follicular adenoma, FA)”은 양성 갑상선 종양으로서, 종양세포가 피막이나 혈관을 침범한 소견이 없을 경우 여포선종으로 진단될 수 있다.

[0081] 본 발명에 있어서, “측정”이란 목적하는 물질의 존재(발현) 여부를 검출 및 확인하는 것, 또는 목적하는 물질의 존재 수준(발현 수준)의 변화를 검출 및 확인하는 것을 모두 포함하는 의미이다. 상기 측정은 정성적인 방법(분석)과 정량적인 방법을 모두 포함하여 제한 없이 수행될 수 있다. 상기 단백질 또는 이의 mRNA로 이루어진 군으로부터 선택된 하나 이상의 존재 여부 측정에 있어서 정성적 방법과 정량적 방법의 종류는 당업계에 잘 알려져 있으며, 본 명세서에서 기술한 실험법들이 이에 포함된다.

[0082] 본 발명에 있어서, 상기 제제는 CNDP2, DNAAF5, DYNC1H1, FARSF, PDCD4, FAM149B1, RPS9, TAGLN2, UPF1, UQCRC, ACTN4, DSTN, MACROH2A1, NUCB1, SPTAN1, TAGLN, 및 XRCC5로 이루어진 군으로부터 선택된 하나 이상의 단백질에 특이적인 항체 또는 앵타머일 수 있으나, 이에 제한되지 않는다.

[0083] 본 발명에 있어서, “항체(antibody)”는 항원성 부위에 대해서 지시되어 특이적으로 결합하는 단백질 분자를 의미한다. 항체는 당업계에서 통상적으로 실시되는 방법들, 예를 들어, 융합 방법, 재조합 DNA 방법, 또는 파아지 항체 라이브러리 방법 등에 의해 제조될 수 있다. 일부 실시형태에서, 항체 또는 항체의 단편은 인간, 마우스, 래트, 햄스터, 토끼, 또는 낙타 등을 포함한, 상이한 유기체에서 유래될 수 있으며, 예컨대 단클론 또는 다클론 항체, 면역학적으로 활성인 단편, 항체 중쇄, 인간화 항체, 항체 경쇄, 유전자 조작된 단일쇄 Fv 분자, 또는 키메라 항체 등일 수 있다. 본 발명에 있어서, 상기 항체는 본 발명의 단백질을 검출할 수 있는 것이라면 특정 항체의 종류로 한정되지 않는다.

[0084] 본 발명에 있어서, “앵타머(aptamer)”란 그 자체로 안정된 삼차구조를 가지면서 표적분자에 높은 친화성과 특이성으로 결합할 수 있는 특징을 가진 단일가닥 핵산(DNA, RNA 또는 변형핵산)을 의미하며, SELEX(Systematic Evolution of Ligands of Exponential enrichment)라는 방법으로 원하는 다양한 목적 물질(단백질, 당, 염색물질, DNA, 금속이온, 세포 등)에 대한 앵타머를 개발할 수 있다.

[0085] 본 발명에 있어서, 상기 제제는 CNDP2, DNAAF5, DYNC1H1, FARSF, PDCD4, FAM149B1, RPS9, TAGLN2, UPF1, UQCRC, ACTN4, DSTN, MACROH2A1, NUCB1, SPTAN1, TAGLN, 및 XRCC5로 이루어진 군으로부터 선택된 하나 이상의 단백질의 mRNA에 특이적으로 결합하는 프라이머 또는 프로브일 수 있으나, 이에 제한되지 않는다.

[0086] 본 발명에 있어서, “프라이머(primer)”는 DNA 합성의 개시점(starting point)으로 작용하는 짧은 단일가닥 올리고뉴클레오타이드(single strand oligonucleotide)이다. 프라이머는 적합한 완충액(buffer)와 온도 조건에서 주형(template)인 폴리뉴클레오타이드에 특이적으로 결합하고, DNA 중합효소가 프라이머에 주형 DNA에 상보적인 염기를 갖는 뉴클레오타이드 트리포스페이트를 추가하여 연결함으로써 DNA가 합성된다. 프라이머는 일반적으로 15 내지 30개의 염기서열로 이루어져 있으며, 염기 구성과 길이에 따라 주형 가닥에 결합하는 온도(melting temperature, Tm)가 달라진다. 프라이머의 서열은 주형의 일부 염기 서열과 완전하게 상보적인 서열을 가질 필요는 없으며, DNA 합성을 통해 mRNA 또는 cDNA의 특정 구간을 증폭하여 유전자 또는 mRNA의 양을 측정하려는 목적에 맞는 길이와 상보성을 갖는 것이면 충분하다. 따라서 본 발명에서 상기 유전자, mRNA 또는 이의 cDNA의 염기서열을 참조하여 프라이머 쌍을 용이하게 디자인할 수 있다. 상기 증폭 반응을 위한 프라이머는 증폭하고자 하는 유전자 또는 mRNA의 특정 구간의 양쪽 끝부분의 주형(또는 센스, sense)과 반대편(안티센스, antisense)에 각각 상보적으로 결합하는 한 세트(쌍)으로 구성된다.

[0087] 본 발명에 있어서, “프로브(probe)”는 특정 유전자의 mRNA나 cDNA(complementary DNA), DNA, miRNA 등에 특이적으로 결합할 수 있는 짧게는 수개 내지 길게는 수백 개의 염기(base pair) 길이의 RNA 또는 DNA 등 폴리뉴클레오타이드의 단편을 의미하며, 표지(labeling)되어 있어서 결합하는 대상 유전자, mRNA나 cDNA의 존재 유무, 발현양 등을 확인할 수 있다. 프로브의 선택 및 혼성화 조건은 당업계에 공지된 기술에 따라 적절하게 선택할 수 있다. 상기 프로브는 대립형질(또는 대립유전자, allele)을 검출하기 위한 진단 방법 등에 사용될 수 있다. 상기 진단 방법에는 서던 블롯 등과 같은 핵산의 혼성화에 근거한 검출 방법들이 포함되며, DNA 칩을 이용한 방

법에서 DNA 칩의 기관에 미리 결합된 형태로 제공될 수도 있다.

- [0088] 본 발명에서, 프라이머 또는 프로브는 포스포아미다이트(phosphoramidite) 고체지지체 합성법이나 기타 널리 공지된 방법을 이용하여 화학적으로 합성할 수 있다. 또한 프라이머 또는 프로브는 검출하고자 하는 표적이 되는 폴리뉴클레오티드와의 혼성화를 방해하지 않는 범위에서 당해 기술분야에 공지된 방법에 따라 다양하게 변형시킬 수 있다. 이러한 변형의 예로는 메틸화, 캡화, 천연 뉴클레오티드 하나 이상의 동족체로의 치환 및 뉴클레오티드 간의 변형, 예를 들면 하전되지 않은 연결체(예: 메틸 포스포네이트, 포스포트리에스테르, 포스포로아미데이트, 카바메이트 등) 또는 하전된 연결체(예: 포스포로티오에이트, 포스포로디티오에이트 등), 그리고 형광 또는 효소를 이용한 표지물질(labeling material)의 결합 등이 있다.
- [0089] 본 발명에 있어서, 상기 프라이머 또는 프로브는 상기 단백질의 mRNA 발현을 측정할 수 있는 것이라면 특정 서열로 한정되지 않는다.
- [0090] 본 발명에 있어서, 상기 조성물은 갑상선 여포 종양의 수술 여부 및 수술 범위를 결정하기 위한 것일 수 있으나, 이에 제한되지 않는다.
- [0091] 본 발명에 있어서, 상기 제제 또는 조성물은 ACTN4, DSTN, MACROH2A1, NUCB1, SPTAN1, TAGLN, 및 XRCC5로 이루어진 군으로부터 선택된 하나 이상의 단백질 또는 이의 mRNA 발현; 또는 ACTN4, DSTN, MACROH2A1, NUCB1, SPTAN1, TAGLN, 및 XRCC5 단백질 또는 이의 mRNA 발현을 측정하여 갑상선 여포암을 진단하기 위한 것일 수 있으나, 이에 제한되지 않는다.
- [0092] 본 발명에 있어서, 상기 제제 또는 조성물은 DNAAF5, FAM149B1, RPS9, TAGLN2, UPF1, 및 UQCRC1으로 이루어진 군으로부터 선택된 하나 이상의 단백질 또는 이의 mRNA 발현; 또는 DNAAF5, FAM149B1, RPS9, TAGLN2, UPF1, 및 UQCRC1 단백질 또는 이의 mRNA 발현을 측정하여 갑상선 여포선종을 진단하기 위한 것일 수 있으나, 이에 제한되지 않는다.
- [0093] 본 발명에 있어서, 상기 제제 또는 조성물은 CNBP2, DNAAF5, DYNC1H1, FARSB, 및 PDCD4로 이루어진 군으로부터 선택된 하나 이상의 단백질 또는 이의 mRNA 발현; 또는 CNBP2, DNAAF5, DYNC1H1, FARSB, 및 PDCD4 단백질 또는 이의 mRNA 발현을 측정하여 양성 결절을 진단하기 위한 것일 수 있으나, 이에 제한되지 않는다.
- [0095] 또한, 본 발명은 상기 갑상선 종양의 감별진단용 조성물을 포함하는, 갑상선 종양의 감별진단용 키트를 제공하며, 상기 키트는 본 발명에 따른 갑상선 종양의 감별진단을 위한 정보제공방법을 교시하는 설명서를 포함할 수 있으나, 이에 제한되지 않는다.
- [0096] 본 발명에 있어서, “키트”는 상기 단백질 또는 이의 mRNA의 발현을 측정하는 제제를 포함함으로써, 갑상선 종양을 감별진단할 수 있도록 하는 도구를 의미한다. 본 발명의 키트에는 상기 제제 외에도 이들의 측정 또는 검출 방법에 통상적으로 필요한 다른 구성성분, 조성물, 용액, 장치 등이 포함될 수 있다. 이때, 상기 유전자, mRNA, 또는 단백질의 발현을 검출하는 물질은 1회 이상 횟수에 제한 없이 작용시킬 수 있으며, 각 물질을 적용하는 선후에는 제한이 없고, 각 물질의 적용은 동시에 진행될 수도 있고 미시에 진행될 수도 있다.
- [0097] 본 발명에 있어서, 상기 키트는 컨테이너; 지시서(설명서); 및 상기 단백질 또는 이의 mRNA의 발현을 측정하는 제제를 포함할 수 있다. 상기 컨테이너는 상기 제제를 포장하는 역할을 할 수 있고, 보관 및 고정하는 역할을 할 수도 있다. 상기 컨테이너의 재질은 예컨대, 병, 통(tub), 작은 봉지(sachet), 봉투(envelope), 튜브, 앰플(ampoule) 등과 같은 형태를 취할 수 있고, 이들은 부분적 또는 전체적으로 플라스틱, 유리, 종이, 호일, 왁스 등으로부터 형성될 수 있다. 상기 용기는 처음에는 용기의 일부이거나 또는 기계적, 접착성, 또는 기타 수단에 의해 용기에 부착될 수 있는, 완전히 또는 부분적으로 분리가 가능한 마개를 장착할 수 있으며, 또한 주사바늘에 의해 내용물에 접근할 수 있는 스톱퍼가 장착될 수 있다. 상기 키트는 외부 패키지를 포함할 수 있으며, 외부 패키지는 구성 요소들의 사용에 관한 지시서를 포함할 수 있다.
- [0099] 또한, 본 발명은 피검체로부터 분리된 생물학적 시료에서 CNBP2, DNAAF5, DYNC1H1, FARSB, PDCD4, FAM149B1, RPS9, TAGLN2, UPF1, UQCRC, ACTN4, DSTN, MACROH2A1, NUCB1, SPTAN1, TAGLN, 및 XRCC5로 이루어진 군으로부터 선택된 하나 이상의 단백질 또는 이의 mRNA 발현 수준을 측정하는 단계를 포함하는, 갑상선 종양의 감별진단을 위한 정보제공방법 또는 갑상선 종양의 감별진단 방법을 제공한다.

- [0100] 또한, 본 발명은 피검체로부터 분리된 생물학적 시료에서 CNBP2, DNAAF5, DYNC1H1, FARSB, PDCD4, FAM149B1, RPS9, TAGLN2, UPF1, UQCRC, ACTN4, DSTN, MACROH2A1, NUCB1, SPTAN1, TAGLN, 및 XRCC5 단백질 또는 이의 mRNA 발현 수준을 측정하는 단계를 포함하는, 갑상선 종양의 감별진단을 위한 정보제공방법 또는 갑상선 종양의 감별진단 방법을 제공한다.
- [0101] 본 발명에 있어서, 상기 생물학적 시료는 갑상선 종양의 감별진단을 실시하고자 하는 피검체에서 채취된 것이라면 제한 없이 사용될 수 있으며, 예를 들어, 생검 등으로 얻어진 조직, 세포, 혈액, 혈장, 혈청, 타액, 비액, 객담, 관절낭액, 양수, 복수, 자궁경부 분비물, 질 분비물, 소변, 뇌척수액 등 생체 시료뿐만 아니라 미세침흡인검체 및 미세침세척액 또한 포함한다. 바람직하게는 생물학적 액체 시료, 예를 들어 혈액, 전혈, 혈장, 소변, 조직, 세포, 기관, 골수, 미세침흡인검체, 미세침세척액, 중심부바늘생검(core needle biopsy) 검체 및 타액으로 이루어진 군에서 선택된 하나 이상에서 측정될 수 있다. 본 발명은 갑상선 종양 또는 결절이 발생한 직접적 갑상선 병변조직 외의 체액시료(일례로, 혈장)를 이용할 수 있으나, 이에 제한되지 않는다.
- [0102] 상기 생물학적 시료는 검출 또는 진단에 사용하기 전에 전처리할 수 있다. 예를 들어, 균질화(homogenization), 여과, 증류, 추출, 농축, 방해 성분의 불활성화, 시약의 첨가 등을 포함할 수 있다. 상기 시료는 단백질 마커의 탐지 감도를 증가시키도록 준비될 수 있는데 예를 들어 환자로부터 수득한 혈청 시료는 음이온 교환 크로마토그래피, 친화도 크로마토그래피, 크기별 배제 크로마토그래피(size exclusion chromatography), 액체 크로마토그래피, 연속추출(sequential extraction) 또는 젤 전기영동 등의 방법을 이용하여 전처리될 수 있다.
- [0103] 본 발명에 있어서, “갑상선 가는바늘흡인검사(FNAC; fine needle aspiration cytology)”란 초음파를 보면서 의심되는 갑상선(경부) 종양 또는 결절을 미세한 주사바늘로 찔러 소량의 세포를 얻은 후 양성 혹은 악성의 세포 진단을 목적으로 시행되며, 가느다란 주사 바늘로 결절을 찔러 세포를 흡인함으로써 갑상선 수술의 진행 여부, 수술 범위 및 림프절 절제 범위 등을 결정하는 기준이 된다.
- [0104] 일반적으로 갑상선 여포암 또는 갑상선 여포선종이 의심되는 환자들의 수술 전 진단은 세침검사 또는 조직검사(중심부생검)를 통해서 이루어지는데, 이러한 검사의 결과는 갑상선 여포 종양(여포 종양 의심) 또는 비정형세포의 소견으로 나타난다.
- [0105] 본 발명에 있어서, 상기 피검체는 갑상선 결절에 대한 진단 검사에서 비정형세포 소견이나 갑상선 여포 종양이 의심되는 환자일 수 있으나, 이에 제한되지 않는다. 예를 들어, 상기 피검체는 갑상선 종양이 의심되는 환자일 수 있으며, 본 발명에 따른 방법에서의 피검체는 비정형세포나 여포성 종양으로 예비적 임상 진단 되었으나, 갑상선 여포암, 갑상선 여포선종, 또는 양성 결절이 모두 의심되는 환자일 수 있다. 즉, 갑상선에 종양이나 결절이 발견되어 종양이나 결절이 악성인 갑상선 여포암인지, 양성인 여포선종 또는 양성 결절인지를 구별해야 하는 환자가 피검체가 된다. 상기 환자는 처음 갑상선 종양이 발생한 환자 및 이전에 갑상선 종양으로 진단된 환자를 포함한다.
- [0106] 본 발명에 있어서, 상기 방법은 CNBP2, DNAAF5, DYNC1H1, FARSB, PDCD4, FAM149B1, RPS9, TAGLN2, UPF1, UQCRC, ACTN4, DSTN, MACROH2A1, NUCB1, SPTAN1, TAGLN, 및 XRCC5로 이루어진 군으로부터 선택된 하나 이상의 단백질 또는 이의 mRNA 발현 수준을 측정하는 단계; 및
- [0107] 상기 ACTN4, DSTN, MACROH2A1, NUCB1, SPTAN1, TAGLN, 및 XRCC5로 이루어진 군으로부터 선택된 하나 이상의 단백질 또는 이의 mRNA 발현이 검출되거나 증가된 경우, 갑상선 여포암으로 판정하는 단계를 포함할 수 있으나, 이에 제한되지 않는다.
- [0108] 본 발명에 있어서, 상기 방법은 CNBP2, DNAAF5, DYNC1H1, FARSB, PDCD4, FAM149B1, RPS9, TAGLN2, UPF1, UQCRC, ACTN4, DSTN, MACROH2A1, NUCB1, SPTAN1, TAGLN, 및 XRCC5 단백질 또는 이의 mRNA 발현 수준을 측정하는 단계; 및
- [0109] 상기 ACTN4, DSTN, MACROH2A1, NUCB1, SPTAN1, TAGLN, 및 XRCC5 단백질 또는 이의 mRNA 발현이 검출되거나 증가된 경우, 갑상선 여포암으로 판정하는 단계를 포함할 수 있으나, 이에 제한되지 않는다.
- [0110] 본 발명에 있어서, 상기 방법은 CNBP2, DNAAF5, DYNC1H1, FARSB, PDCD4, FAM149B1, RPS9, TAGLN2, UPF1, UQCRC, ACTN4, DSTN, MACROH2A1, NUCB1, SPTAN1, TAGLN, 및 XRCC5로 이루어진 군으로부터 선택된 하나 이상의 단백질 또는 이의 mRNA 발현 수준을 측정하는 단계; 및
- [0111] 상기 FAM149B1, RPS9, TAGLN2, UPF1, 및 UQCRC1으로 이루어진 군으로부터 선택된 하나 이상의 단백질 또는 이의 mRNA 발현이 검출되거나 증가된 경우, 또는 DNAAF5 단백질 또는 이의 mRNA 발현이 검출되지 않거나 감소된

경우 갑상선 여포선종으로 판정하는 단계를 포함할 수 있으나, 이에 제한되지 않는다.

- [0112] 본 발명에 있어서, 상기 방법은 CNBP2, DNAAF5, DYNC1H1, FARSB, PDCD4, FAM149B1, RPS9, TAGLN2, UPF1, UQCRC, ACTN4, DSTN, MACROH2A1, NUCB1, SPTAN1, TAGLN, 및 XRCC5 단백질 또는 이의 mRNA 발현 수준을 측정하는 단계; 및
- [0113] 상기 FAM149B1, RPS9, TAGLN2, UPF1, 및 UQCRC1 단백질 또는 이의 mRNA 발현이 검출되거나 증가된 경우, 또는 DNAAF5 단백질 또는 이의 mRNA 발현이 검출되지 않거나 감소된 경우 갑상선 여포선종으로 판정하는 단계를 포함할 수 있으나, 이에 제한되지 않는다.
- [0114] 본 발명에 있어서, 상기 방법은 CNBP2, DNAAF5, DYNC1H1, FARSB, PDCD4, FAM149B1, RPS9, TAGLN2, UPF1, UQCRC, ACTN4, DSTN, MACROH2A1, NUCB1, SPTAN1, TAGLN, 및 XRCC5로 이루어진 군으로부터 선택된 하나 이상의 단백질 또는 이의 mRNA 발현 수준을 측정하는 단계; 및
- [0115] 상기 DNAAF5 단백질 또는 이의 mRNA 발현이 검출되거나 증가된 경우, 또는 CNBP2, DYNC1H1, FARSB, 및 PDCD4로 이루어진 군으로부터 선택된 하나 이상의 단백질 또는 이의 mRNA 발현이 검출되지 않거나 감소된 경우 양성 결정으로 판정하는 단계를 포함할 수 있으나, 이에 제한되지 않는다.
- [0116] 본 발명에 있어서, 상기 방법은 CNBP2, DNAAF5, DYNC1H1, FARSB, PDCD4, FAM149B1, RPS9, TAGLN2, UPF1, UQCRC, ACTN4, DSTN, MACROH2A1, NUCB1, SPTAN1, TAGLN, 및 XRCC5 단백질 또는 이의 mRNA 발현 수준을 측정하는 단계; 및
- [0117] 상기 DNAAF5 단백질 또는 이의 mRNA 발현이 검출되거나 증가된 경우, 또는 CNBP2, DYNC1H1, FARSB, 및 PDCD4 단백질 또는 이의 mRNA 발현이 검출되지 않거나 감소된 경우 양성 결정으로 판정하는 단계를 포함할 수 있으나, 이에 제한되지 않는다.
- [0118] 본 발명에 있어서, 상기 특정 단백질 또는 이의 mRNA 발현이 증가 또는 감소된 경우, 이는 상기 특정 단백질 또는 이의 mRNA의 발현 수준이 대조군에서의 발현 수준에 비해 증가 또는 감소된 경우일 수 있고, 피검체에서 CNBP2, DNAAF5, DYNC1H1, FARSB, PDCD4, FAM149B1, RPS9, TAGLN2, UPF1, UQCRC, ACTN4, DSTN, MACROH2A1, NUCB1, SPTAN1, TAGLN, 및 XRCC5의 17개 단백질 중 특정 단백질의 조합이 다른 단백질들에 비해 특히 증가 또는 감소된 경우일 수도 있다.
- [0119] 본 발명에 있어서, 양성 결정으로 진단되는 경우 상기 대조군은 정상인, 갑상선 여포선종, 또는 갑상선 여포암 환자일 수 있고, 갑상선 여포선종으로 진단되는 경우 상기 대조군은 정상인, 양성 결정, 또는 갑상선 여포암 환자일 수 있고, 갑상선 여포암으로 진단되는 경우 상기 대조군은 정상인, 양성 결정, 또는 갑상선 여포선종 환자일 수 있으나, 이에 제한되지 않는다.
- [0120] 본 발명에 있어서, 상기 단백질 발현 수준은 질량 분석법, 웨스턴 블롯, ELISA(enzyme linked immunoassay), 캡처-ELISA, 억제 또는 경쟁 분석, 샌드위치 분석, 방사선면역분석법, 방사능면역분석, 방사면역확산법, 오우크테로니(Ouchterlony) 면역확산법, 로케트 면역전기영동, 면역조직화염색, 면역형광염색, 면역침화성정제, 방사능면역침전, 면역침전분석, 보체고정분석, 유세포 분석법(FACS) 및 단백질 칩으로 이루어진 군에서 선택된 어느 하나 이상의 방법을 이용하여 측정되는 것일 수 있으나, 이에 제한되지 않는다.
- [0121] 본 발명에 있어서, 상기 유전자 또는 mRNA 발현 수준은 중합효소연쇄반응(polymerase chain reaction, PCR), 역전사 중합효소연쇄반응(reverse transcription polymerase chain reaction, RT-PCR), 경쟁적 역전사 중합효소반응(Competitive RT-PCR), 정량적 실시간 역전사 중합효소 연쇄반응(qRT-PCR), 실시간 중합효소연쇄반응(Real-time PCR, quantitative PCR, quantitative real-time PCR), RNase 보호 분석법(RPA; RNase protection assay), 노던 블롯팅(Northern blotting), 및 DNA 칩으로 이루어진 군에서 선택된 어느 하나 이상의 방법을 이용하여 측정되는 것일 수 있으나, 이에 제한되지 않는다.
- [0123] 또한, 본 발명은 후보물질을 투여한 피검체로부터 분리된 생물학적 시료에서 CNBP2, DNAAF5, DYNC1H1, FARSB, PDCD4, FAM149B1, RPS9, TAGLN2, UPF1, UQCRC, ACTN4, DSTN, MACROH2A1, NUCB1, SPTAN1, TAGLN, 및 XRCC5로 이루어진 군으로부터 선택된 하나 이상의 단백질 또는 이의 mRNA 발현 수준을 측정하는 단계를 포함하는, 갑상선 종양의 치료용 후보물질의 스크리닝 방법을 제공한다.
- [0124] 본 발명의 또 다른 구현예로서, 상기 방법은 CNBP2, DNAAF5, DYNC1H1, FARSB, PDCD4, FAM149B1, RPS9, TAGLN2,

UPF1, UQCRC, ACTN4, DSTN, MACROH2A1, NUCB1, SPTAN1, TAGLN, 및 XRCC5 단백질 또는 이의 mRNA 발현 수준을 측정하는 단계; 및

- [0125] 상기 ACTN4, DSTN, MACROH2A1, NUCB1, SPTAN1, TAGLN, 및 XRCC5 단백질 또는 이의 mRNA 발현이 후보물질 투여 전에 비해 감소된 경우, 상기 후보물질을 갑상선 여포암 치료용 후보물질로 선택하는 단계를 포함할 수 있으나, 이에 제한되지 않는다.
- [0126] 본 발명의 다른 구현예로서, 상기 방법은 CNDP2, DNAAF5, DYNC1H1, FARSB, PDCD4, FAM149B1, RPS9, TAGLN2, UPF1, UQCRC, ACTN4, DSTN, MACROH2A1, NUCB1, SPTAN1, TAGLN, 및 XRCC5 단백질 또는 이의 mRNA 발현 수준을 측정하는 단계; 및
- [0127] 상기 FAM149B1, RPS9, TAGLN2, UPF1, 및 UQCRC1 단백질 또는 이의 mRNA 발현이 후보물질 투여 전에 비해 감소된 경우, 또는 DNAAF5 단백질 또는 이의 mRNA 발현이 후보물질 투여 전에 비해 증가된 경우 상기 후보물질을 갑상선 여포선종 치료용 후보물질로 선택하는 단계를 포함할 수 있으나, 이에 제한되지 않는다.
- [0128] 본 발명의 일 구현예로서, 상기 방법은 CNDP2, DNAAF5, DYNC1H1, FARSB, PDCD4, FAM149B1, RPS9, TAGLN2, UPF1, UQCRC, ACTN4, DSTN, MACROH2A1, NUCB1, SPTAN1, TAGLN, 및 XRCC5 단백질 또는 이의 mRNA 발현 수준을 측정하는 단계; 및
- [0129] 상기 DNAAF5 단백질 또는 이의 mRNA 발현이 후보물질 투여 전에 비해 감소된 경우, 또는 CNDP2, DYNC1H1, FARSB, 및 PDCD4 단백질 또는 이의 mRNA 발현이 후보물질 투여 전에 비해 증가된 경우 상기 후보물질을 양성 결절 치료용 후보물질로 선택하는 단계를 포함할 수 있으나, 이에 제한되지 않는다.
- [0130] 본 발명에 있어서, “후보물질”은 갑상선 종양 환자에서의 상기 단백질 또는 이의 mRNA 발현 여부 또는 발현 정도를 확인하기 위하여 스크리닝에 사용되는 미지의 물질을 의미하며, 예컨대 뉴클레오타이드, DNA, RNA, 아미노산, 펩타이드, 단백질, 화합물, 천연물, 및 천연추출물로 이루어진 군으로부터 선택되는 하나 이상일 수 있으나 이에 제한되지 않는다.
- [0132] 또한, 본 발명은 피검체로부터 분리된 생물학적 시료에서 CNDP2, DNAAF5, DYNC1H1, FARSB, PDCD4, FAM149B1, RPS9, TAGLN2, UPF1, UQCRC, ACTN4, DSTN, MACROH2A1, NUCB1, SPTAN1, TAGLN, 및 XRCC5 단백질 또는 이의 mRNA 발현 수준을 측정하는 단계;
- [0133] 상기 ACTN4, DSTN, MACROH2A1, NUCB1, SPTAN1, TAGLN, 및 XRCC5 단백질 또는 이의 mRNA 발현이 검출되거나 증가된 경우, 갑상선 여포암으로 판정하고,
- [0134] 상기 FAM149B1, RPS9, TAGLN2, UPF1, 및 UQCRC1 단백질 또는 이의 mRNA 발현이 검출되거나 증가된 경우, 또는 DNAAF5 단백질 또는 이의 mRNA 발현이 검출되지 않거나 감소된 경우 갑상선 여포선종으로 판정하고,
- [0135] 상기 DNAAF5 단백질 또는 이의 mRNA 발현이 검출되거나 증가된 경우, 또는 CNDP2, DYNC1H1, FARSB, 및 PDCD4 단백질 또는 이의 mRNA 발현이 검출되지 않거나 감소된 경우 양성 결절로 판정하는 단계; 및
- [0136] 양성 결절, 갑상선 여포선종, 또는 갑상선 여포암에 대한 치료제를 투여하는 단계를 포함하는, 갑상선 종양의 치료 방법을 제공한다.
- [0137] 이때, 상기 치료제는 갑상선 종양에 대하여 통상적으로 사용되는 치료제라면 종류에 제한이 없다.
- [0138] 또한, 본 발명은 갑상선 종양이 의심되는 피검체와 대조군에서 분리된 생물학적 시료에서 상기 단백질 또는 이의 mRNA의 발현 수준을 측정한 후, 이 수준 차이가 (2배 이상으로) 증가한 경우 양성 결절, 갑상선 여포선종, 또는 갑상선 여포암 중 하나를 가진 피검체임을 가리키는 것으로 보는, 갑상선 종양의 감별진단에 대한 감수성이 높은 피검체인지 결정/분석하기 위한 정보를 제공하는 방법을 제공한다.
- [0140] 또한, 본 발명은 CNDP2, DNAAF5, DYNC1H1, FARSB, PDCD4, FAM149B1, RPS9, TAGLN2, UPF1, UQCRC, ACTN4, DSTN, MACROH2A1, NUCB1, SPTAN1, TAGLN, 및 XRCC5로 이루어진 군으로부터 선택된 하나 이상의 단백질 또는 이의 mRNA 발현 수준을 측정하는 체제를 유효성분으로 포함하는 조성물의 갑상선 종양의 감별진단 용도를 제공한다.
- [0141] 또한, 본 발명은 CNDP2, DNAAF5, DYNC1H1, FARSB, PDCD4, FAM149B1, RPS9, TAGLN2, UPF1, UQCRC, ACTN4,

DSTN, MACROH2A1, NUCB1, SPTAN1, TAGLN, 및 XRCC5로 이루어진 군으로부터 선택된 하나 이상의 단백질 또는 이의 mRNA 발현 수준을 측정하는 제제를 유효성분으로 포함하는 조성물의, 갑상선 종양의 감별진단용 제제의 제조를 위한 용도를 제공한다.

[0143] 이하, 본 발명의 이해를 돕기 위하여 바람직한 실시예 또는 실험예를 제시한다. 그러나 하기의 실시예 또는 실험예는 본 발명을 보다 쉽게 이해하기 위하여 제공되는 것일 뿐, 하기 실시예 또는 실험예에 의해 본 발명의 내용이 한정되는 것은 아니다.

[0145] **[실시예]**

[0146] **실시예 1. 연구 대상**

[0147] 양성 결절(benign nodule, BN) 62개, 여포선종(follicular adenoma, FA) 72개, 갑상선 여포암(thyroid carcinoma, FTC) 68개를 포함하여 총 202개의 포르말린 고정 파라핀 포매(FFPE) 조직을 분석에 사용하였다. 이 조직은 모두 2008년부터 2018년 사이 서울아산병원에서 갑상선 절제술을 시행한 후 개별 환자의 수술 샘플에서 얻었다. 숙련된 내분비학 병리학자(D.E.S) 한 명이 모든 검체를 검토하였으며, 각 환자로부터 사전 동의를 얻었고 서울아산병원 임상연구심의위원회에서 모든 데이터 수집과 후속 분석을 승인받았다.

[0149] **실시예 2. FFPE 조직 단백질 추출**

[0150] 수술로 제거된 갑상선 조직 샘플의 보관용 포르말린 고정 파라핀 포매(FFPE) 조직을 5 μ m 두께로 슬라이스 하였다. 잘게 슬라이스한 FFPE 조직 절편을 Thermomixer(Eppendorf, Hamburg, 독일)에 넣고 300 μ L의 헵탄에서 600 rpm으로 60분 동안 파라핀을 제거하였다. 그런 다음, 메탄올 15 L를 첨가한 후 15000 rpm에서 30분간 원심 분리하였다. Q-proteome FFPE Tissue Kit(Qproteome[®] FFPE Tissue, Hilden, 독일)는 제조업체의 지침에 따라 사용하였다. 즉, 6 % (v/v) β -머캅토에탄올(mercaptoethanol)(Sigma-Aldrich, St. Louis, MO, 미국)을 함유한 100 μ L의 EXB plus 추출 완충액(Qproteome[®] FFPE Tissue, Hilden, 독일)을 조직 펠릿에 첨가하였다. 그리고 나서 튜브를 100 $^{\circ}$ C에서 20분 동안 가열한 후, 튜브를 Thermomixer를 사용하여 750 rpm에서 웨이킹하면서 80 $^{\circ}$ C에서 2시간 동안 배양하고 얼음 위에서 5분간 배양하였다. 이후 4 $^{\circ}$ C에서 15,000 rpm으로 30분간 원심분리한 다음 상층액을 옮겼다. 이를 15,000 rpm, 4 $^{\circ}$ C에서 30분간 원심분리한 후 상층액을 옮기고 아세톤을 6회 첨가하여 16시간 동안 반응시켰다. 그런 다음, 15000 rpm, 4 $^{\circ}$ C에서 30분 동안 원심분리한 후 상층액을 버리고 100 μ L의 Strap 용해 완충액을 펠릿에 넣었다. 100 μ L의 5 % SDS의 분취량(aliquot)을 함유하는 50 mM TEAB(pH 7.55)를 재현탁에 사용하였으며, 제조업체의 지침에 따라 BCA 단백질 분석을 사용하여 단백질 농도를 측정하였다(Pierce, Rockford, IL, 미국).

[0151] 또한, 300 μ g의 단백질을 사용하여 환원 및 알킬화를 수행하였다. 디티오프레이톨(dithiothreitol)을 95 $^{\circ}$ C에서 10분 동안 최종 농도 20 mM로 첨가하여 이황화물을 감소시켰다. 알킬레이트 시스테인(alkylate cysteine)에 요오드아세트아마이드(iodoacetamide)를 첨가하여 최종 농도가 40 mM이 되도록 처리하고, 단백질 용액을 실온에서 냉각시킨 후 샘플을 암실에서 30분 동안 배양하였다. 현탁액 포획 샘플 준비를 이용하여 단백질 소화를 수행하였으며, 샘플을 10배로 희석하고, 12 % 인산을 샘플에 첨가하였다. 샘플을 S-Trap[™] 미니 컬럼(ProtiFi, NY, USA, Cat. No: CO2-mini-80)에 넣고 제조업체의 프로토콜에 따라 수행하였으며, 샘플 대 Lys-C/트립신의 비율(Promega Corp., Madison, WI, 미국, Cat. No: V5071)은 1:20이었다. 용출된 펩타이드 혼합물은 쿨드 트랩으로 동결건조 하였으며 사용 전까지 -80 $^{\circ}$ C에서 보관하였다.

[0153] **실시예 3. RAS 돌연변이 분석**

[0154] 신선 냉동 조직 및 FFPE 샘플의 게놈 DNA(genomic DNA)를 증합효소 연쇄 반응(PCR) 증폭 및 직접 시퀀싱에 사용하여 RAS 돌연변이의 존재를 평가하였다. NRAS, HRAS 및 KRAS 유전자의 엑손 2 및 3을 PCR로 증폭시키고, 이전 연구에서 보고된 바와 같이 RAS 돌연변이의 상태를 분석하였다(Thyroid. 2018 Apr;28(4):504-510.). 즉, 200 ng의 게놈 DNA를 주형으로 사용하여 최종 부피 20 μ l를 Accupower PCR premix(BIONEER, 대전, 한국)로 증폭시켰다. 증폭 프로토콜은 95 $^{\circ}$ C에서 5분 동안 초기 변성으로 구성되었으며, 이어서 95 $^{\circ}$ C에서 20초 동안 변성, 55

℃에서 20초 동안 프라이머를 주형에 어닐링, 72 ℃에서 30초 동안 프라이머 확장, 및 72 ℃에서 5분 동안 최종 확장을 35회 수행하였다. 그런 다음, PCR 산물을 4 ℃로 냉각하였고, 최종 PCR 산물은 전기영동을 이용하여 확인하였다. 증폭된 PCR은 젤/PCR DNA 미니 키트(Real Biotech Corporation, Delhi, 인도)를 사용하여 정제하고 시퀀싱하였다. DNA 분석기(ABI 3730xl System; Applied Biosystems, Waltham, MA)를 사용하여 DNA 서열을 판독하였으며, RAS 돌연변이의 상태를 확인하기 위해 각 DNA 샘플을 최소 두번 분석하였다.

[0156] **실시예 4. 나노-LC-ESI-MS/MS 분석**

[0157] 소화된 펩타이드는 Dionex UltiMate 3000 RSLCnano 시스템(Thermo Fisher Scientific)을 사용하여 분리하였다. 비드 컬럼의 트립신 펩타이드를 100 μ L의 0.1 % 포름산(formic acid)에서 재구성하고, C18 Pepmap 트랩 컬럼 (20 mm x 100 μ m i.d., 5 μ m, 100 Å; Thermo Scientific, 미국)이 장착된 AcclaimTM Pepmap 100 C18 컬럼 (500 mm x 75 μ m i.d., 3 μ m, 100 Å)에서 50 ℃에서 150분 동안 0.1 % 포름산 및 5 % DMSO의 0~48% 아세트오 니트릴 구배를 사용하여 200분에 걸쳐(250 nL/분) 분리하였다. LC(Liquid Chromatography)는 nano-ESI(Electrospray Ionization) 소스를 갖춘 Q ExactiveTM Plus Hybrid Quadrupole-OrbitrapTM 질량 분석기와 결합하였다. 질량 스캔은 전체 스캔과 10개의 데이터 의존 MS/MS 스캔 사이의 자동 전환을 통해 데이터 의존 모드에서 수집되었다. 최대 주입 시간 50 ms 및 m/z 400에서 분해능 70,000과 함께 350 m/z ~ 1800 m/z에서 선택된 전체 스캔 MS 스펙트럼에 대한 목표 값은 3,000,000이었다. 선택된 이온은 다음 매개변수의 고에너지 충돌 해리(HCD)에 의해 단편화되었다: 2 Da 전구체 이온 분리 창 및 27 % 정규화된 충돌 에너지. MS/MS에 대한 이온 목표 값은 100 ms의 최대 주입 시간과 m/z 400에서 분해능 17,500과 함께 1,000,000으로 설정하였다. 반복되는 펩타이드는 20초 동안 동적으로 제외되었다. 모든 MS 데이터는 PRIDE 아카이브(www.ebi.ac.uk/pride/archive/) PDX043154에 보관되었다.

[0159] **실시예 5. 단백질체 식별 및 정량화 전략**

[0160] raw 데이터 파일은 ProteoWizard msConvert에 의해 mzml 및 mgf 파일로 변환되었다. 본 발명에서 DDA-MS 데이터 분석을 위해 제안된 워크플로는 MSFragger(버전 2.3)의 오픈 서치 스펙트럼 결과를 기반으로 ANN-Solo 스펙트럼 라이브러리 검색 접근 방식, k-nearest hashing 접근 방식을 채택하였다. MS 데이터 분석은 두 가지 다른 방식으로 처리하였다. 스펙트럼 라이브러리 검색 결과를 기반으로 단백질 풍부도를 얻었고, FragPipe 플랫폼(버전 12.2)을 이용하여 단백질 번역 후 변형(post-translational modification, PTM) 스펙트럼 카운트 및 주석을 얻었다.

[0162] **실시예 6. MSFragger 오픈 서치 및 PTM 식별**

[0163] mzml 파일은 FragPipe(버전 12.2) 및 Philosopher(버전 2.1.1)와 결합된 MSFragger(버전 2.3)로 검색되어 오픈 서치를 수행하였다. 단백질 서열 데이터베이스는 검토된 H. sapiens(2020년 2월 21일에 UniProt으로부터 평가됨)와 미끼(decoy) 서열이 추가된 조합이었다. 전구체 질량 허용 오차(tolerance)는 -150 내지 +500 Da로 설정되었으며 전구체 실제 허용 오차 및 단편 질량 허용 오차는 20 ppm으로 설정되었다. 또한, 질량 눈금 및 매개변수 최적화가 허용되었다. 2개의 누락된 절단이 허용되었으며, 효소 말단의 수는 2개로 설정되었다. 아이소타입(isotype) 오류는 0으로 설정되었다. 모델링에 PSM을 포함하는 데 필요한 최소 단편 피크의 수는 2로 설정되었으며, 매치를 확인하는 데 필요한 최소 수는 4였다. 최소 15개의 단편 피크 및 상위 100개의 가장 강렬한 피크가 초기 설정으로 사용되었다. C-크리스탈 필터 MS 스펙트럼 및 PTM-shepherd는 MS 스펙트럼의 PTM 주석에 사용되었다.

[0165] **실시예 7. ANN-Solo 스펙트럼 검색 및 비표지(label-free) 단백질 정량화**

[0166] MSFragger 소프트웨어의 개별 검색 결과(interact.pep.xml)를 사용하여 SpectraST(버전 5.0.0)에 의해 스펙트럼 라이브러리를 구축하였고, 이는 202개의 스펙트럼을 펩타이드 이온에 대한 최고 점수를 가진 것으로 결합하였다. 이 스펙트럼 라이브러리의 경우, 202 mgf 파일로 변환된 raw를 ANN-solo를 통해 검색하였다. 전구체 질량 허용 오차는 $\pm$ 20 ppm으로 설정하였고 MS/MS 허용 오차는 0.02 Da로 설정하였다. 검색 매개변수는 기본 매개변

수로 설정하였다. 각 분석에 대해 오발견율(False discovery rate)은 1 %로 설정하였다. 스펙트럼-스펙트럼 매치(SSM)는 FlashLFQ 소프트웨어(버전 1.1.1)로 정량화 하였다. 이렇게 계산된 단백질 발현 값은 추가 분석에 사용하였다.

[0168] 실시예 8. 정규화 및 결측 데이터 채우기를 통한 차등 단백질 존재비 분석

[0169] 정규화 방법은 FlashLFQ로부터 얻은 정량적 단백질 값에 대해 perseus(버전1.6.15.0)별로 폭을 조정하여 수행하였다. 총 샘플과 그룹 샘플에서 측정된 단백질의 비율에 따라 4가지 클래스가 결정되었다. 클래스 1은 전체 샘플의 70 % 이상을 차지하는 정량적 값을 갖는 단백질이다. 클래스 2는 클래스 1의 단백질을 제외하고 각 군에서 정량적 값이 70 % 이상 측정된 단백질이다. 클래스 3은 클래스 1 및 2의 단백질을 제외하고 군별로 정량적 값이 20 % 이상 측정된 단백질이다. 클래스 4는 클래스 1, 2, 및 3의 단백질을 제외한 모든 샘플에서 하나 이상의 정량적 값이다. 클래스 2에 대하여 클래스 1,291, 클래스 3에 대하여 2,169, 클래스 4에 대하여 3,280에 속하는 592개의 단백질이 있다. 클래스 1의 결측값(missing value)은 MAR(missing at random)을 가정한 SVD(singular value decomposition) 방법으로 채워졌다. 클래스 3의 결측값은 MNAR(missing not at random)을 가정하는 왼쪽 검열(left-censored) 데이터(grilc)(시그마 = 0.3) 방법의 변위치 회귀 대체(quantile regression imputation)에 의해 채워졌다. 클래스 2의 결측값은 그룹별로 다르게 채워졌으며, 70 % 이상 정량화된 그룹은 MAR 방법으로 채워졌으나 MNAR 방법으로는 채워지지 않았다. 클래스 4 단백질은 통계 분석에서 제외되었다. 여기서 결측값 추정은 R의 imputeLCMD 패키지를 통해 수행되었다.

[0171] 실시예 9. 특징 선택을 기반으로 한 통계적 임상 모델 생성

[0172] 세 가지 질병 그룹을 예측하기 위해 세 가지 모델을 구축하였다. 예를 들어 FA를 예측하는 모델에서, FTC 및 BN은 대조군으로, FA는 질병 그룹으로 정의하였다. 나머지 두 질병 그룹도 같은 방식으로 진행하였다. 랜덤 포레스트(RF) 기반 역방향 제거 프로세스를 사용하여 대조군과 질병 그룹을 구별하기 위해, 클래스 1에 속하는 592개의 단백질로부터 최적의 하위 집합을 식별하기 위해 특징 선택을 수행하였다. 이 과정은 다음 두 단계로 구성된다: 첫째, 592개의 변수를 포함하는 5,000개의 의사결정 트리를 무작위로 생성하고, AUC(곡선 아래 면적) 값을 계산하였다. AUC 값은 Out-of-bag 오류 추정을 사용하여 최적의 단백질 수를 결정하는 데 사용되었으며, BN의 경우 8, FA의 경우 47, FTC의 경우 37의 값을 산출하였다. 둘째, 모델에 포함된 각 변수의 예측 중요도를 계산하기 위해 선택된 3개의 변수를 사용하여 20번의 반복과 5중 교차 검증을 수행하였다.

[0173] 단백질(BN=5(≥ 0.75 선택 확률), FA=6(≥ 0.9 선택 확률), FTC=7(≥ 0.75 선택 확률))을 선택하고(도 5b-d), 모델 구축 전에 센터링 및 스케일링을 포함한 데이터 전처리를 수행하였다. 트레이닝 및 검증 세트는 전체 데이터를 사용하여 3분의 1로 나누어졌다. RF 모델은 10,000개의 트리 및 nodeSize=3을 사용하여 3중 교차 검증을 3회 반복하여 mtry 매개변수로 최적화되었다. 기계 학습(ML) 모델 예측 값은 트레이닝 세트 샘플 없이 검증 세트에서 얻었으며, ROC 분석을 수행하였다. 이 과정은 세트를 100번 무작위로 변경하여 수행하였다.

[0175] [실험예]

[0176] 실험예 1. 기준선(baseline) 특징

[0177] 202명의 환자의 조직학적 아형별 특징을 하기 표 1에 정리하였다.

표 1

	Total (n = 202)	Benign (n=62)	FA (n=72)	FTC (n=68)	p-value
Age (years)	50 (40-58)	52 (44 - 59)	50 (39 - 56)	43 (34 - 57)	0.025
Sex (female)	167 (82.7%)	52 (83.9%)	61 (84.7%)	54 (79.4%)	0.454
Primary tumor size (cm)	3.0 (2.2-3.8)	2.8 [2.0; 4.3]	2.7 [2.0; 3.2]	3.3 [2.5; 4.0]	0.341
US characteristics					0.278
Benign	11 (5.4%)	5 (8.1%)	3 (4.2%)	3 (4.4%)	
Low suspicion	137 (67.8%)	47 (75.8%)	47 (65.3%)	43 (63.2%)	
Intermediate suspicion	52 (25.7%)	9 (14.5%)	22 (30.6%)	21 (30.9%)	

High suspicion	2 (1.0%)	1 (1.6%)	0 (0.0%)	1 (1.5%)	
Preoperative diagnosis					<0.001
Benign	7 (3.5%)	5 (8.1%)	1 (1.4%)	1 (1.5%)	
AUS	121 (59.9%)	53 (85.5%)	36 (50.0%)	32 (47.1%)	
Follicular neoplasm	69 (34.2%)	3 (4.8%)	35 (48.6%)	31 (45.6%)	
Suspicious for malignancy	2 (1.0%)	1 (1.6%)	0 (0.0%)	1 (1.5%)	
Malignancy	3 (1.5%)	0 (0.0%)	0 (0.0%)	3 (4.4%)	
RAS mutations	47 (23.3%)	2 (3.2%)	22 (30.6%)	23 (33.8%)	<0.001

평균연령은 50세였으며 아형간 유의한 연령차이가 관찰되었다. FTC 환자의 연령은 가장 낮았고(평균 43세) BN 결절 환자의 연령이 가장 높았다(평균 52세, $p=0.025$). 환자의 대부분은 여성(82.7 %)이었으며 평균 종양 크기는 3.0 cm(사분위간 범위 2.2~3.8)였다. 한국 갑상선 영상 보고 및 데이터 시스템(K-TIRADS)을 기반으로 대부분의 종양이 '낮은 의심'(67.8 %), 그 다음 '중간 의심'(25.7 %)으로 분류되었으며, 이러한 빈도 경향은 모든 세 가지 아형에서 동일하였다. 수술 전 가는바늘흡인(Fine needle aspiration, FNA)에 의한 진단에서는 조직학적 아형에 따라 유의한 차이가 있었다. FTC 그룹($n=68$)에서는 그 중 32명(47.1 %)이 비정형(atypia of undetermined significance, AUS)으로 진단되었고, 31명(45.6 %)이 FN으로 진단되었다. 수술 전 악성 종양으로 진단된 환자는 3명(4.4 %)에 불과했다. 마찬가지로, FA 환자($n=72$)는 FNA에 의해 거의 고르게 AUS($n=36$, 50 %)와 FN($n=35$, 48.6 %)으로 진단되었다. 반면, 양성 결절이 있는 환자의 대부분은 수술 전 AUS로 진단되었으며($n=53$, 85.5 %) FNA에 의해 양성으로 간주된 환자는 5명(8.1 %)이었다. RAS 돌연변이는 BN군, FA군, FTC군에서 각각 3.2 %, 30.6 %, 33.8 %로 나타났다.

실험예 2. 비교 단백질체학 결과

비교 FFPE 갑상선 조직 프로테옴 분석의 작업 흐름도는 도 1에 나타내었다. 202개 샘플의 전체 프로테옴은 DDA-LC-MS/MS로 분석되었다. MS 스펙트럼 분석 시 MSFragger를 이용하여 단백질 DB를 이용한 클로즈 서치(close search)와 오픈 서치(open search)를 분석하였고, 오픈 서치 결과를 바탕으로 스펙트럼 라이브러리를 생성한 후 스펙트럼-스펙트럼 비교 분석을 통해 ANN-SoLo 프로그램으로 단백질 분석을 수행하였다. 그 결과, 도 2a에 나타낸 바와 같이 전체 MS2 스펙트럼에서, 69.88 %(FA), 70.54 %(FTC), 66.60 %(BN)가 오픈 스펙트럼-스펙트럼 매치(spectrum-spectrum match, SSM)로 덮여 있었고, 45.54 %(FA), 49.45 %(FTC), 38.99 %(BN)가 오픈 펩타이드-스펙트럼 매치(peptide-spectrum match, PSM)로 덮여 있었다. 20.70 %(FA), 19.79 %(FTC), 15.35 %(BN)는 클로즈 SSM으로 덮여 있었고, 18.51 %(FA), 22.00 %(FTC), 14.30 %(BN)는 클로즈 PSM으로 덮여있었다.

도 2b에 나타낸 바와 같이 4가지 스펙트럼 매칭 방식을 통해 총 8,554개 중 공통적인 4,848개의 단백질을 정량화하였다. SSM의 경우, 오픈 서치에서는 PSM에서 22 % 더 많은 단백질이 정량화되었고, PSM에서보다 클로즈 서치에서는 42.7 % 더 많은 단백질이 정량화되었다. 특히, 오픈 SSM은 가장 많은 단백질을 매치하고 정량화 하였으며 다른 세 가지 방식과 비교하여 2.55 % 미만의 단백질을 제외하였다. 도 2c에 나타낸 바와 같이 Human Protein Atlas([https://www. Proteinatlas.org/](https://www.Proteinatlas.org/))와의 비교는 오픈 SSM에서 정량화된 8,341개의 단백질 중 85개가 갑상선 조직 샘플에서 증가했음을 나타내었다. 티로글로불린(thyroglobulin, TG)은 갑상선 조직에서 가장 높은 단백질로, 세 번째로 높은 알부민보다 5배 이상 높게 정량되었고, 액체 크로마토그래피-MS(LC-MS)로는 저농도 단백질을 측정하기 어렵다는 것을 알 수 있다. 도 2d에 나타낸 바와 같이 202개 샘플에서 오픈 SSM의 단백질 수는 TG의 오픈 SSM 비율과 음의 상관관계가 있었다($R^2=0.162$ 및 $p\text{-값}=1.77 \times 10^{-9}$). 또한, 오픈 PSM의 단백질 수는 TG의 오픈 PSM 비율과 강한 음의 상관관계가 있었다($R^2=0.525$ 및 $p\text{-값}< 2.0 \times 10^{-16}$). MSFragger의 펩타이드 스펙트럼 검색과 ANN-Solo의 스펙트럼 대 스펙트럼 검색이라는 두 가지 방법으로 식별된 단백질의 수를 비교해 보면, 스펙트럼 대 스펙트럼 검색에서 더 높은 수가 식별되었으며, 도 2d에 나타낸 바와 같이 두 방법 모두에서 단백질의 수가 양의 상관관계를 보인 것($R^2=0.589$ 및 $p\text{-값}<2.0 \times 10^{-16}$)은 MSFragger에서 확인된 대부분의 단백질이 ANN-Solo에서 확인된 단백질에 포함되어 있음을 나타낸다.

[0186] **실험예 3. 프로테오믹스에 대한 조직 보관 시간의 영향**

[0187] 잠재적인 샘플링 편향을 피하기 위해 대규모 발견 기반 연구 설계와 특히 관련성이 있는 조직 보관 시간에 따른 단백질체학 결과를 조사하였다. 본 발명에 사용된 조직은 평균 보관 시간 7년, 최대 12년으로 2008년부터 2018년 사이에 수집 및 보관되었다. 샘플은 보관 시간에 따라 6~12년(N=164) 및 1~5년(N=38)으로 그룹화하였다. 먼저 보관 시간 그룹에 걸쳐 정량화된 단백질의 총 수를 비교한 결과, 도 2e에 나타난 바와 같이 단기 그룹에서 장기 그룹에 비해 정량화된 단백질이 다소 더 높게 나타났다(+9.48 %). 또한, 낮은 식별률이 시간 경과에 따른 메틸화, 산화 또는 수산화, 요오드화와 관련이 있는지 조사하였다. 이는 장기 보관 샘플에서 더 높은 펩타이드 변형률을 향한 일반적인 경향을 나타냈으며, 이는 시간이 지남에 따라 점진적인 단백질 변형을 나타낸다. 그러나 1 %의 차이는 가장 오래된 샘플에서 펩타이드 식별이 16 % 감소한 것을 설명하기에는 너무 작았다. 이전 연구에서는 20년 이상 저장된 조직 샘플의 식별률이 낮다고 보고되었으며, 이는 풍부하지 않은 단백질의 손상된 검색과 연결되었다(J Proteome Res. 2009 Feb;8(2):917-25.). 다양한 보관 시간 그룹 간의 높은 프로테오믹스 상관관계로 판단할 때 전체 단백질 정량화에 대한 영향도 도 2f에 나타난 바와 같이 미미했다(Pearson 상관 관계 0.93).

[0189] **실험예 4. 주요 경로에 따른 세 가지 하위 유형 구별**

[0190] 클로즈 SSM을 사용하여 전체 202개의 갑상선 FFPE 샘플에서 6,332개의 단백질을 정량화 하였으며, 도 3a에 나타난 바와 같이 부분 최소 제곱 판별 분석(partial least squares-discriminant analysis, PLS-DA)을 통해 FA, FTC, 및 BN 그룹은 구성 요소 1(9 %)과 2(2 %)의 두 구성 요소로 적당히 분리하였다. 각각의 구성 요소에서 중요도가 높은 상위 20개 단백질을 도 3b에 나타내었다. 프로테오믹스 결과에서는 도 3c에 나타난 바와 같이 FA와 FTC가 해당(glycolysis) 및 당신생합성(glyconeogenesis)에 관여하는 단백질로서 Wardburg 효과의 발현 시그니처를 나타냄을 확인하였다.

[0191] 또한, Molecular Signatures Database(MSigDB)의 50개 특징 유전자 세트에서 각 샘플의 유전자 세트에 대해 singscore 방법으로 농축 점수를 계산하였다. 그 결과, 도 3d 내지 3f에 나타난 바와 같이 세 그룹의 ANOVA 테스트를 통해 적어도 두 그룹에서 통계적으로 다른 22개 경로의 상관 관계 플롯을 나타내었다. 특히, 도 3f에 나타난 바와 같이 BN에 대하여 FA 및 FTC는 MYC_targets_V1 및 해당작용(glycolysis)을 상향 조절했고, BN 및 FA에 대하여 FTC는 염증 반응(inflammatory_response)을 하향 조절했으며, BN 및 FTC에 대하여 FA는 PI3K_AKT_MTOR_SIGNALING을 하향 조절하였다.

[0193] **실험예 5. 하위 유형 간의 정량 분석**

[0194] 세 그룹 사이에서 차등 발현된 단백질(DEP)을 확인하기 위해, 두 그룹의 스튜던트 t-테스트를 통해 배수 변화 및 p-값을 세 번 계산하였다. $-\log_{10}$ 조정된 p-값에 대한 $\log_2$ 배수 변화를 보여주는 첫 번째 볼케이노 플롯 결과, 도 4a에 나타난 바와 같이 BN에서 상향 조절되는 28개 단백질 및 FA에서 상향 조절되는 379개 단백질을 확인하였다($|\log_2$ 배수 변화|>1; 조정된 p-값<0.05). $-\log_{10}$ 조정 p-값에 대한 $\log_2$ 배수 변화를 보여주는 두 번째 볼케이노 플롯 결과, 도 4b에 나타난 바와 같이 BN에서 상향 조절되는 43개 단백질 및 FTC에서 상향 조절되는 482개 단백질을 확인하였다($|\log_2$ 배수 변화|>1; 조정된 p-값<0.05). BN, FTC, 및 FA의 DEP에 대한 벤 다이어그램을 도 4d에 나타내었으며, 그 결과 FA 및 FTC에 대해 상향 조절된 BN에서 2개의 DEP 세트는 공통적으로 12개였고, FA 특이적 단백질은 16개, 및 FTC 특이적 단백질은 31개였다. 또한, BN에 대한 FA-상향 조절 DEP와 BN에 대한 FTC 상향 조절 DEP는 공통적으로 186개였고, FA 특정 단백질은 193개, FTC 특정 단백질은 296개였으며, 이는 MYC_TARGETS_V1, OXIDATIVE_PHOSPHORYLATION, FATTY_ACID_METABOLISM, ADIPOGENESIS 등과 같은 암 신호 전달 경로와 관련이 있었다. 아울러, $-\log_{10}$ 조정된 p-값에 대한 $\log_2$ 배수 변화를 보여주는 세 번째 볼케이노 플롯 결과, 도 4c에 나타난 바와 같이 FA에서 상향 조절되는 59개 단백질 및 FTC에서 상향 조절되는 59개 단백질($|\log_2$ 배수 변화|>1; 조정된 p-값<0.05)을 확인하였다. FA에 대한 FTC 상향 조절 DEP는 도 4e에 나타난 바와 같이 여러 가지 유전자 온톨로지(GO) 용어 및 특징에 크게 관여하는 것을 확인하였다.

[0196] **실험예 6. 단백질체 기반 기계 학습(ML) 모델을 이용한 갑상선암 아형 예측**

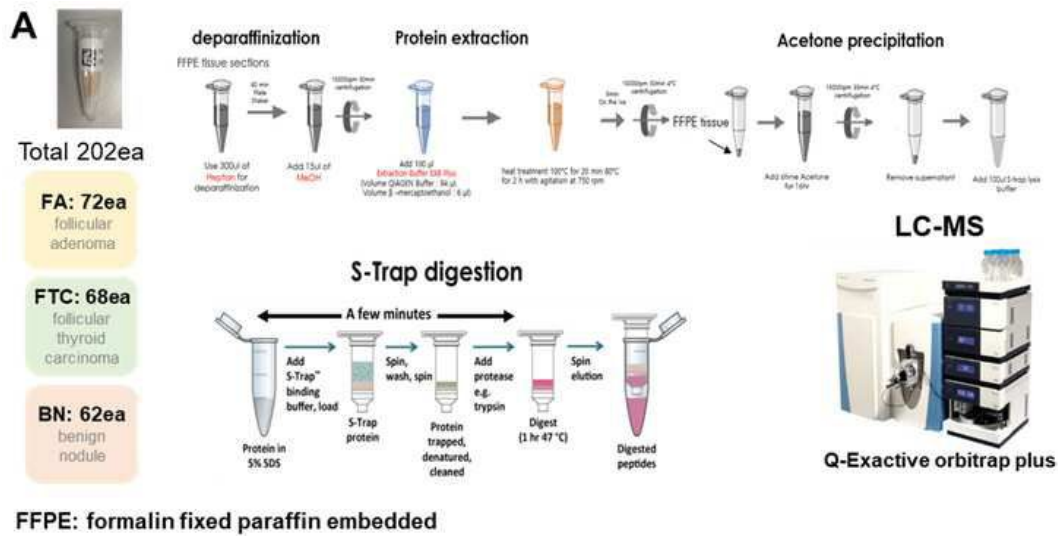
[0197] 본 발명에서 임상 목적은 FNA에 의한 2017년 Bethesda 분류 시스템을 기반으로 FTC가 일반적으로 카테고리

III(6~18 %) 또는 IV(10~40 %)로 진단되는 환자에 대하여 새로운 단백질체 기반 방법을 통해 FTC 스크리닝의 정확도를 향상시키는 것이다. 또한, 갑상선 절제술이 필요하지 않은 BN과 FA를 구분하여 불필요한 수술을 줄일 수 있다. 통계적 신뢰성을 보장하기 위해, 본 발명은 갑상선 조직 샘플의 70 % 이상에서 정량화된 592개의 단백질에 중점을 두었으며, 카테고리 III 또는 IV에 포함된 183개 샘플만 추가 분석을 거쳤다. 592개 단백질의 일변량(univariate) ROC(receiver operating characteristic) 분석은 세 그룹 각각에 대해 수행하였다. 그 결과, 도 5a에 나타난 바와 같이 FTC의 경우 MACROH2A1(AUROC=0.720), FA의 경우 UQCRC1(AUROC=0.736), BN의 경우 DNAAF5(AUROC=0.773)가 가장 높게 식별된 단백질이였다.

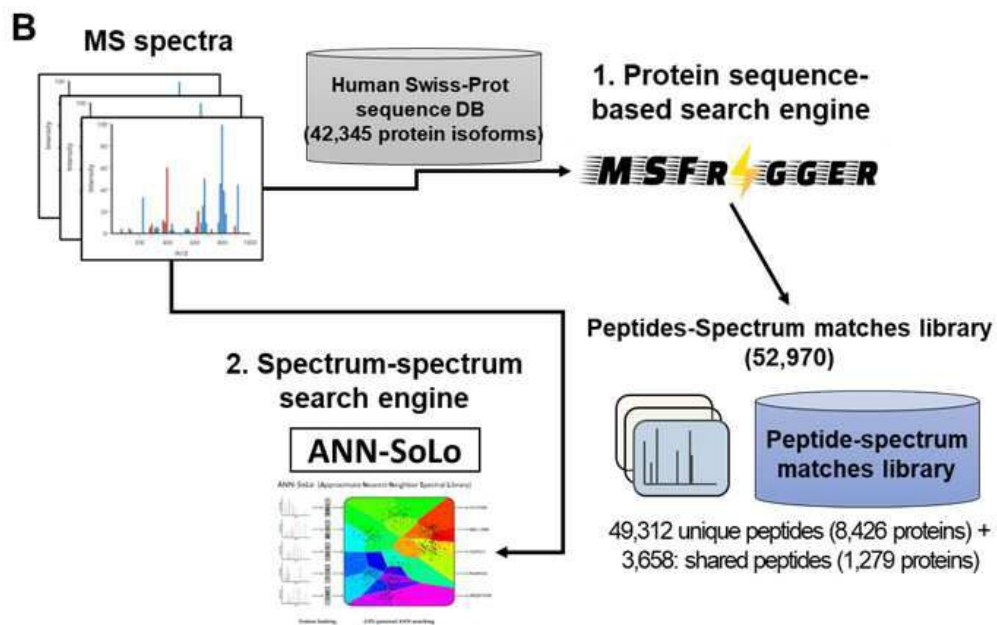
- [0198] 예측 성능을 향상시키고 세 가지 아형을 구별할 수 있는 의미 있는 단백질 조합을 찾기 위해 랜덤 포레스트(RF)를 기반으로 세 가지 임상 모델을 생성하였다. AUC 기반 RF 역방향 제거 프로세스에 의해, 도 5b 내지 5d에 나타난 바와 같이 BN 예측 모델에서 5개의 단백질(CNDP2, DNAAF5, DYNC1H1, FARSB 및 PDCD4; 확률 ≥ 0.65)을 선택하였고(도 5b), FA 예측 모델에서는 6개의 단백질(DNAAF5, FAM149B1, RPS9, TAGLN2, UPF1 및 UQCRC1; 확률 ≥ 0.9)을 선택하였으며(도 5c), FTC 예측 모델에서는 7개의 단백질(ACTN4, DSTN, MACROH2A1, NUCB1, SPTAN1, TAGLN 및 XRCC5; 확률 ≥ 0.75)을 선택하였다(도 5d).
- [0199] 상기 FTC, FA, 및 BN의 진단 마커에 대하여, 각각의 마커별 발현 수준을 비교하여 도 5e 내지 5g에 나타내었다.
- [0200] FTC 진단 마커의 발현 수준을 확인한 결과, 도 5e에 나타난 바와 같이, ACTN4, SPTAN1, XRCC5, DSTN, TAGLN, NUCB1, 및 MACROH2A1는 BN 및 FA 환자와 비교하여 FTC 환자에서 가장 발현 수준이 높은 것을 확인하였다.
- [0201] FA 진단 마커의 발현 수준을 확인한 결과, 도 5f에 나타난 바와 같이, FAM149B1, RPS9, TAGLN2, UPF1, 및 UQCRC1은 BN 및 FTC 환자에 비해 발현 수준이 높게 나타났으며, DNAAF5의 경우 BN 및 FTC 환자에 비해 발현 수준이 낮은 것을 확인하였다.
- [0202] BN 진단 마커의 발현 수준을 확인한 결과, 도 5g에 나타난 바와 같이, DNAAF5는 FA 및 FTC 환자에 비해 발현 수준이 높게 나타났으며, CNDP2, DYNC1H1, FARSB, 및 PDCD4의 경우 FA 및 FTC 환자에 비해 발현 수준이 낮은 것을 확인하였다.
- [0203] 또한, 각 모델의 특징을 기반으로 RF 모델을 구축하였으며, ML 모델의 견고성을 확인하기 위해 샘플을 무작위로 100회 삼등분한 후 샘플의 2/3로 모델을 구축하고 샘플의 1/3로 검증하였다. BN, FA 및 FTC 분류기의 성능을 평가한 결과, 도 5h에 나타난 바와 같이 RF에 대한 AUC 중앙값은 각각 0.832(95 % CI: 0.824-0.839), 0.826(95 % CI: 0.817-0.835) 및 0.870(95 % CI: 0.863-0.877)인 것으로 나타났다.
- [0204] 또한, 세 가지 예측 모델에 의해 파생된 가장 높은 세 가지 확률 점수를 가진 질병으로 샘플의 질병을 결정하였다. 그 결과, 도 5i에 나타난 바와 같이 피검체 183명 중 141명이 교정되었으며, 정확도는 77.05 %였다. BN 샘플(N=55)에서 정확도는 87.98 %, 민감도는 80 %, 특이도는 74.19 %였고, FA 샘플(N=66)의 경우 정확도는 85.25 %, 민감도는 77.27 %, 특이도는 89.74 %였으며, FTC 샘플(N=62)에서 정확도는 80.87 %, 민감도는 74.19 %, 특이도는 84.3 %였다.
- [0206] 진술한 본 발명의 설명은 예시를 위한 것이며, 본 발명이 속하는 기술분야의 통상의 지식을 가진 자는 본 발명의 기술적 사상이나 필수적인 특징을 변경하지 않고서 다른 구체적인 형태로 쉽게 변형이 가능하다는 것을 이해할 수 있을 것이다. 그러므로 이상에서 기술한 실시예들은 모든 면에서 예시적인 것이며 한정적이 아닌 것으로 이해해야 한다.

도면

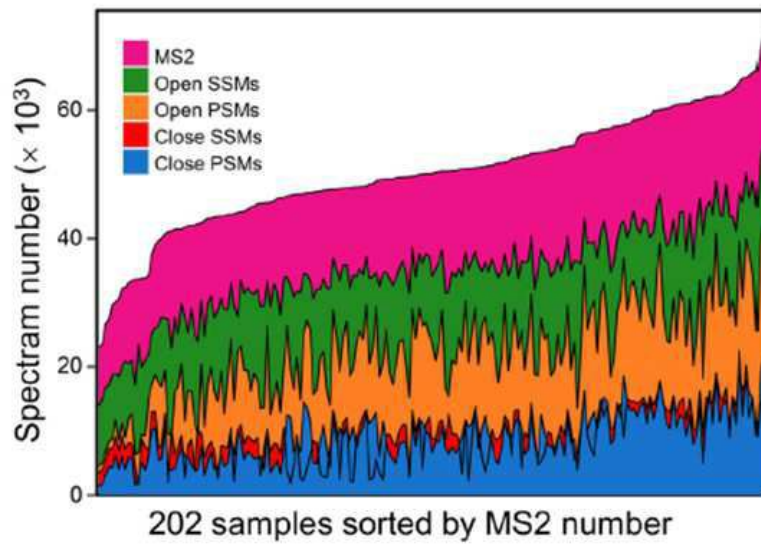
도면1



FFPE: formalin fixed paraffin embedded



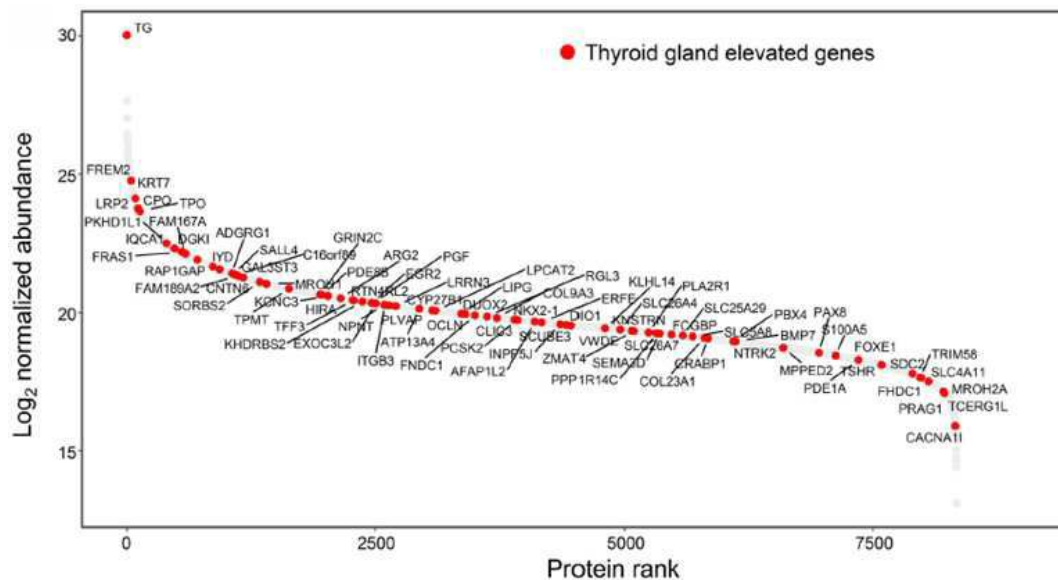
도면2a



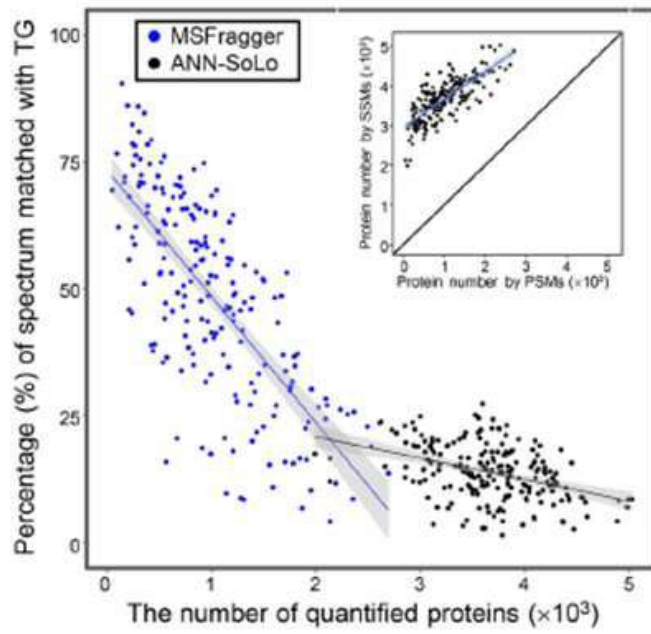
도면2b



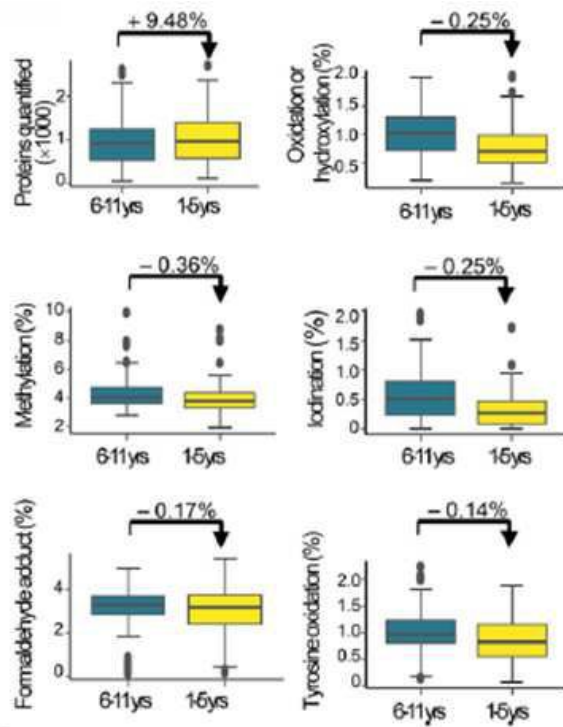
도면2c



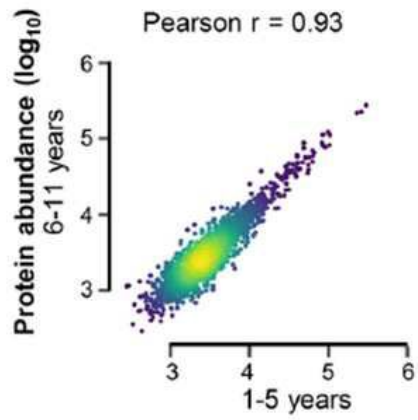
도면2d



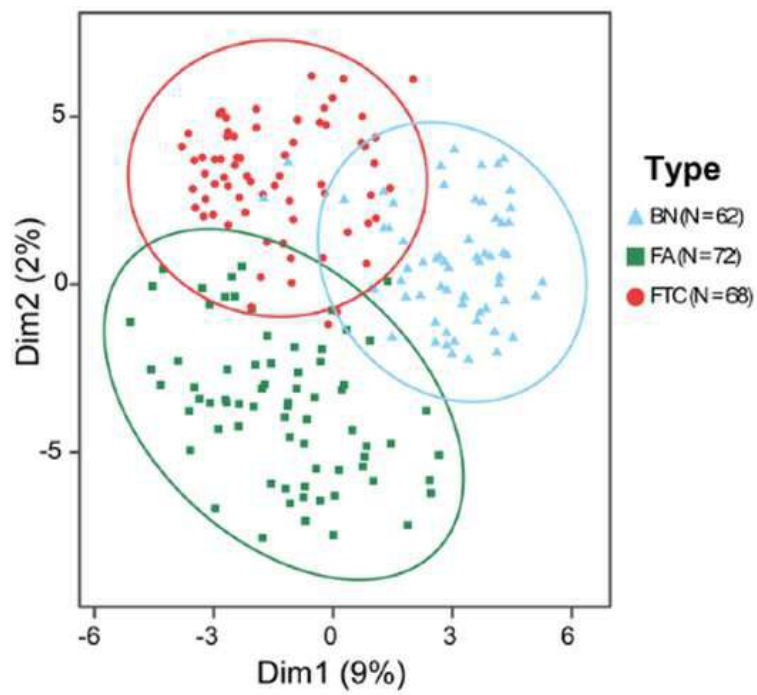
도면2e



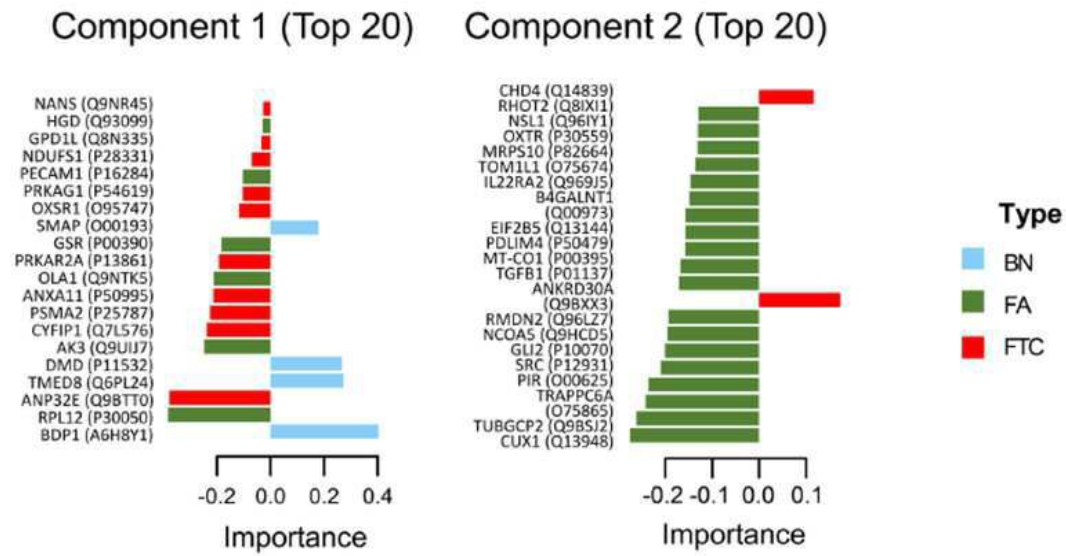
도면2f



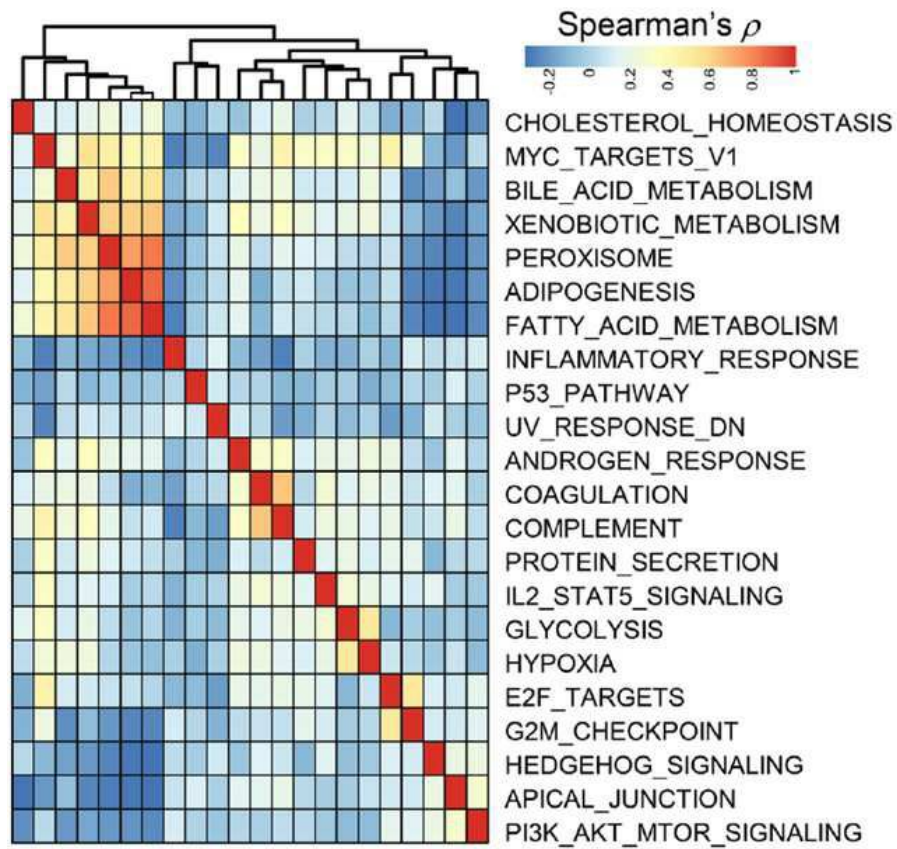
도면3a



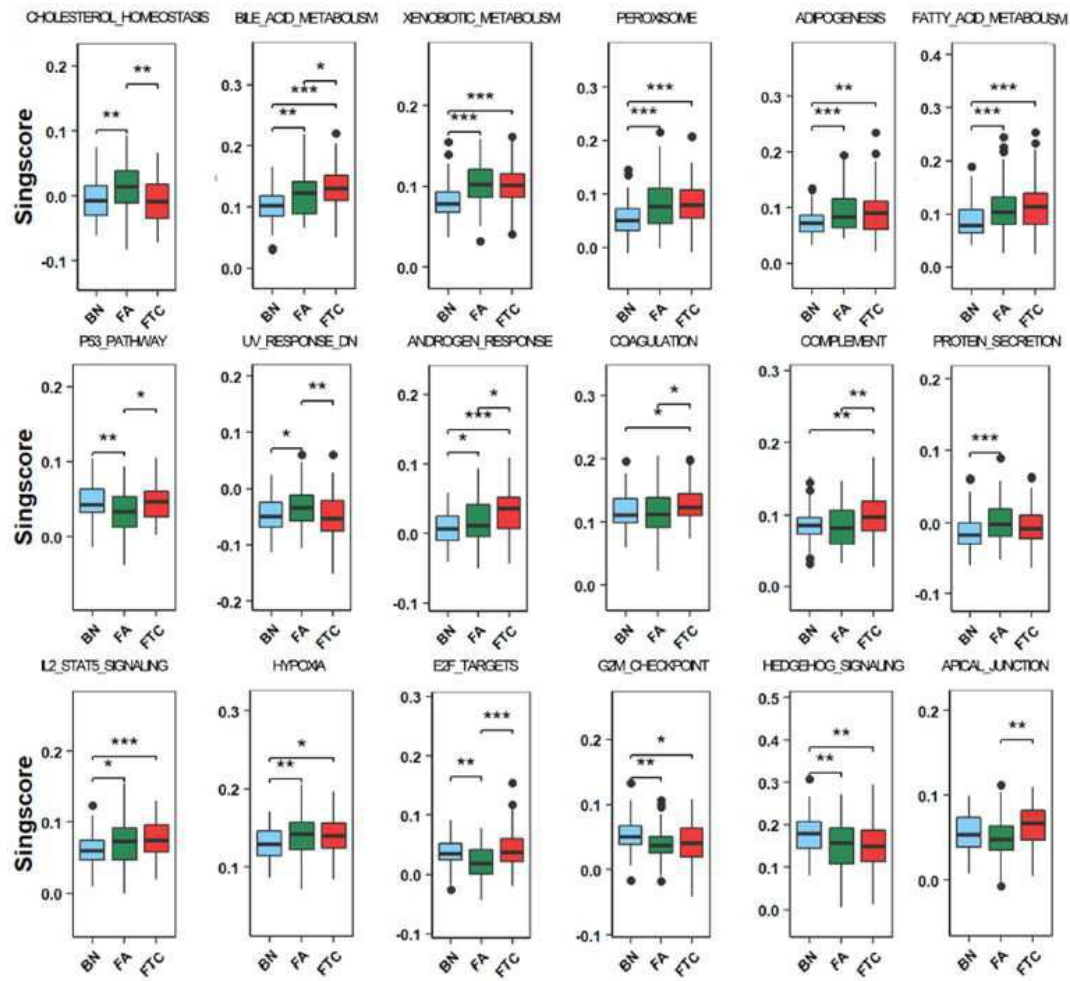
도면3b



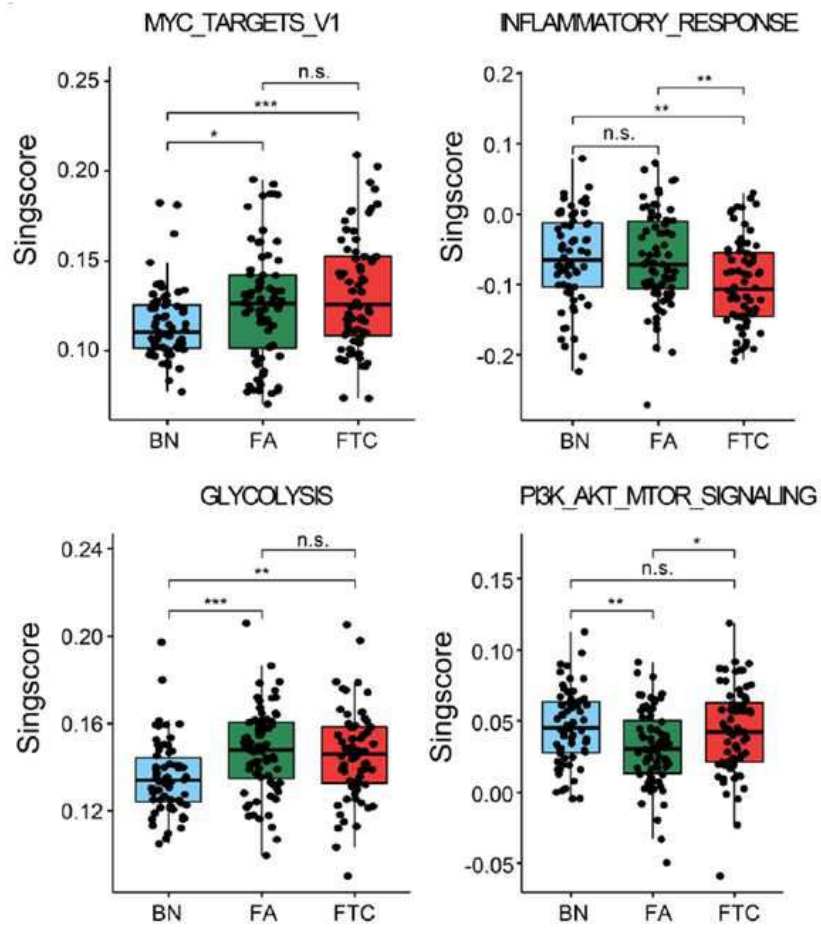
도면3d



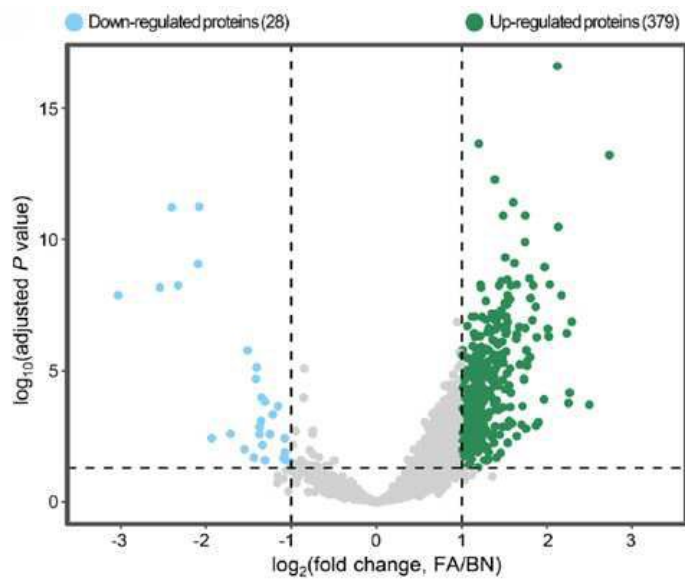
도면3e



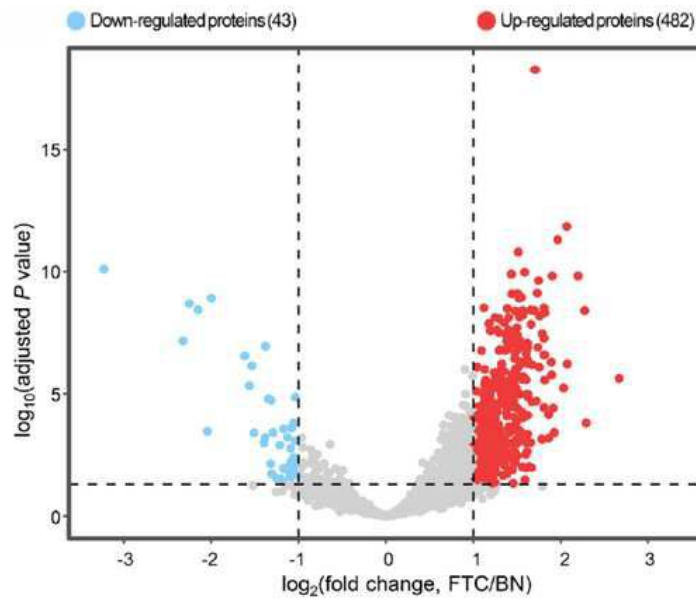
도면3f



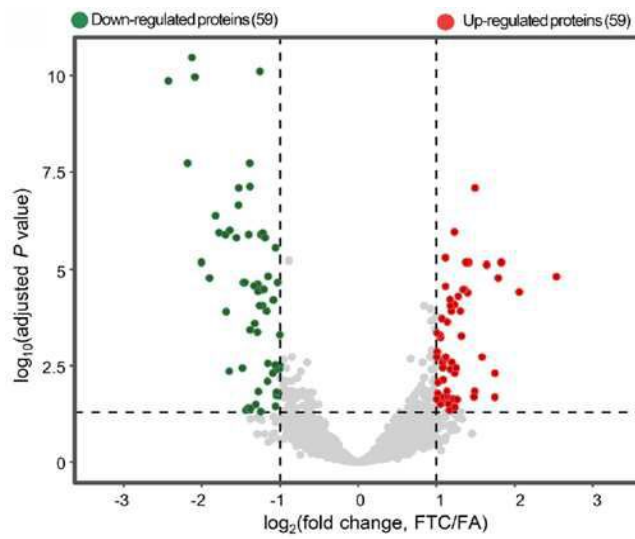
도면4a



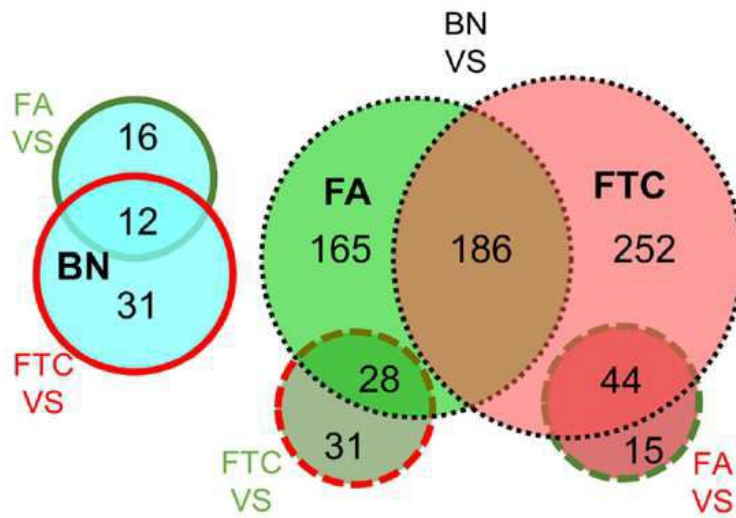
도면4b



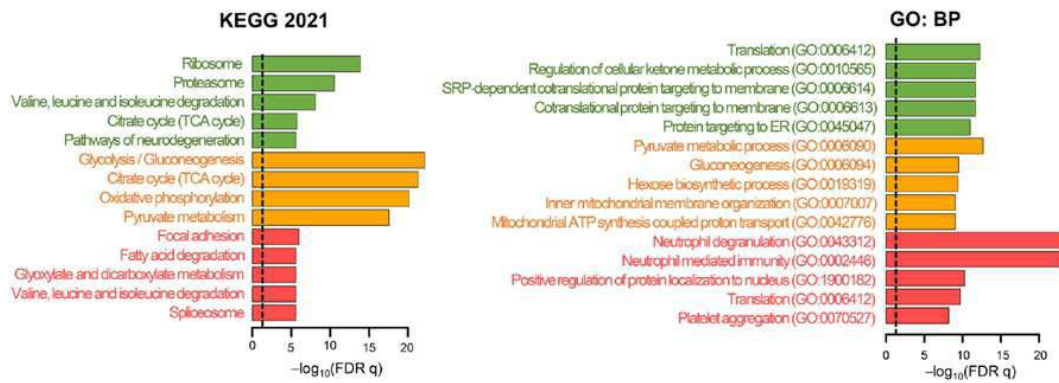
도면4c



도면4d

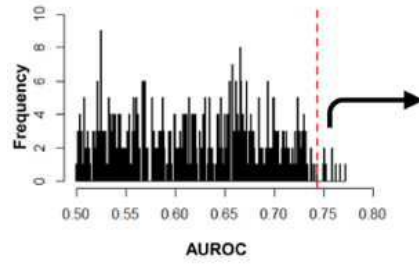


도면4e



도면5a

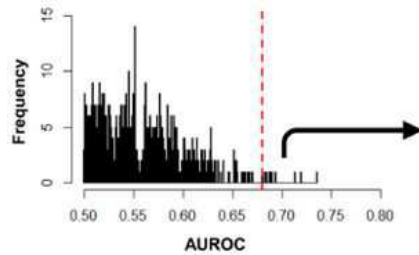
Histogram of AUROC of 592 proteins in BN



Top 10 proteins in BN

UniProt Accession ID	Gene	AUROC
Q86Y56	DNAAF5	0.773
Q53EL6	PDCD4	0.767
P07355	ANXA2	0.763
P63104	YWHAZ	0.759
Q9NSD9	FARSB	0.759
Q96KP4	CNDP2	0.753
P49591	SARS1	0.751
P27695	APEX1	0.751
P05455	SSB	0.750
Q9UL46	PSME2	0.744

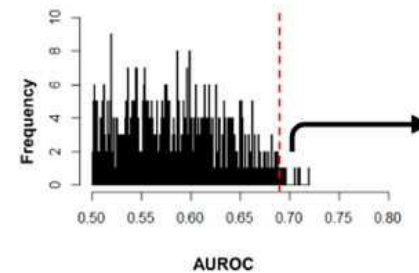
Histogram of AUROC of 592 proteins in FA



Top 10 proteins in FA

UniProt Accession ID	Gene	AUROC
P31930	UQCRC1	0.736
P46781	RPS9	0.720
Q86Y56	DNAAF5	0.714
Q02878	RPL6	0.694
P37802	TAGLN2	0.690
P17980	PSMC3	0.689
Q96BN6	FAM149B1	0.688
P35270	SPR	0.685
Q14011	CIRBP	0.683
Q92900	UPF1	0.681

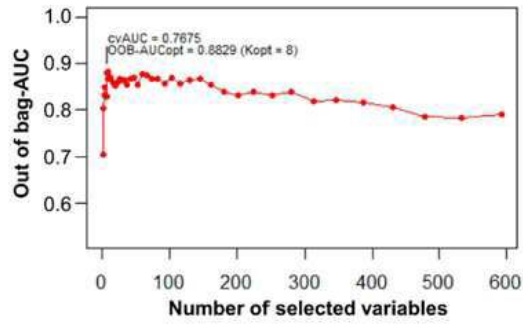
Histogram of AUROC of 592 proteins in FTC



Top 10 proteins in FTC

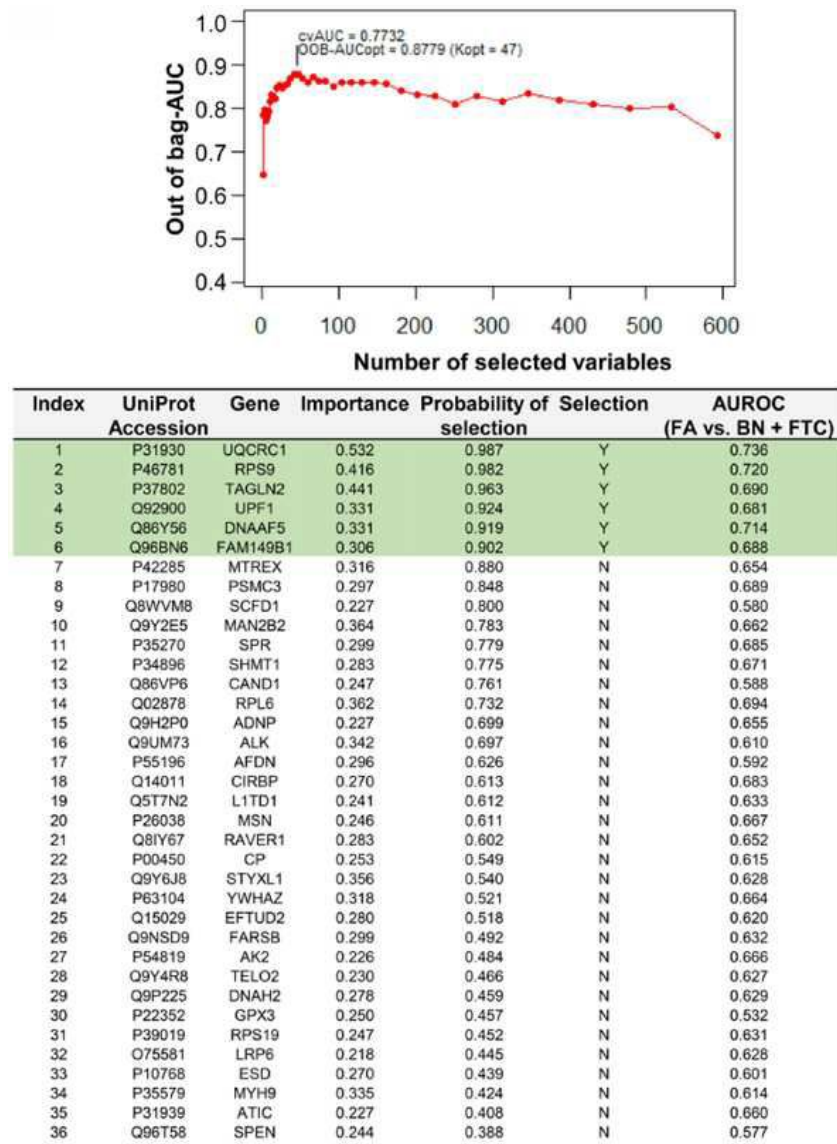
UniProt Accession ID	Gene	AUROC
O75367	MACROH2A1	0.720
Q53EL6	PDCD4	0.711
P60981	DSTN	0.709
P00450	CP	0.706
P02671	FGA	0.697
Q13813	SPTAN1	0.696
Q02818	NUCB1	0.694
P12956	XRCC6	0.693
P45974	USP5	0.692
Q01995	TAGLN	0.690

도면5b

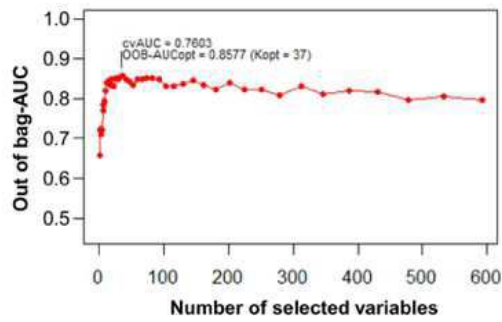


Index	UniProt Accession	Gene	Importance	Probability of selection	Selection	AUROC (BN vs. FA + FTC)
1	Q86Y56	DNAAF5	0.451	0.884	Y	0.773
2	Q14204	DYNC1H1	0.380	0.872	Y	0.599
3	Q96KP4	CNDP2	0.438	0.790	Y	0.753
4	Q9NSD9	FARSB	0.383	0.788	Y	0.759
5	Q53EL6	PDCD4	0.372	0.684	Y	0.767
6	P45974	USP5	0.403	0.584	N	0.731
7	P31930	UQCRC1	0.406	0.478	N	0.675
8	P49591	SARS1	0.435	0.436	N	0.751

도면5c

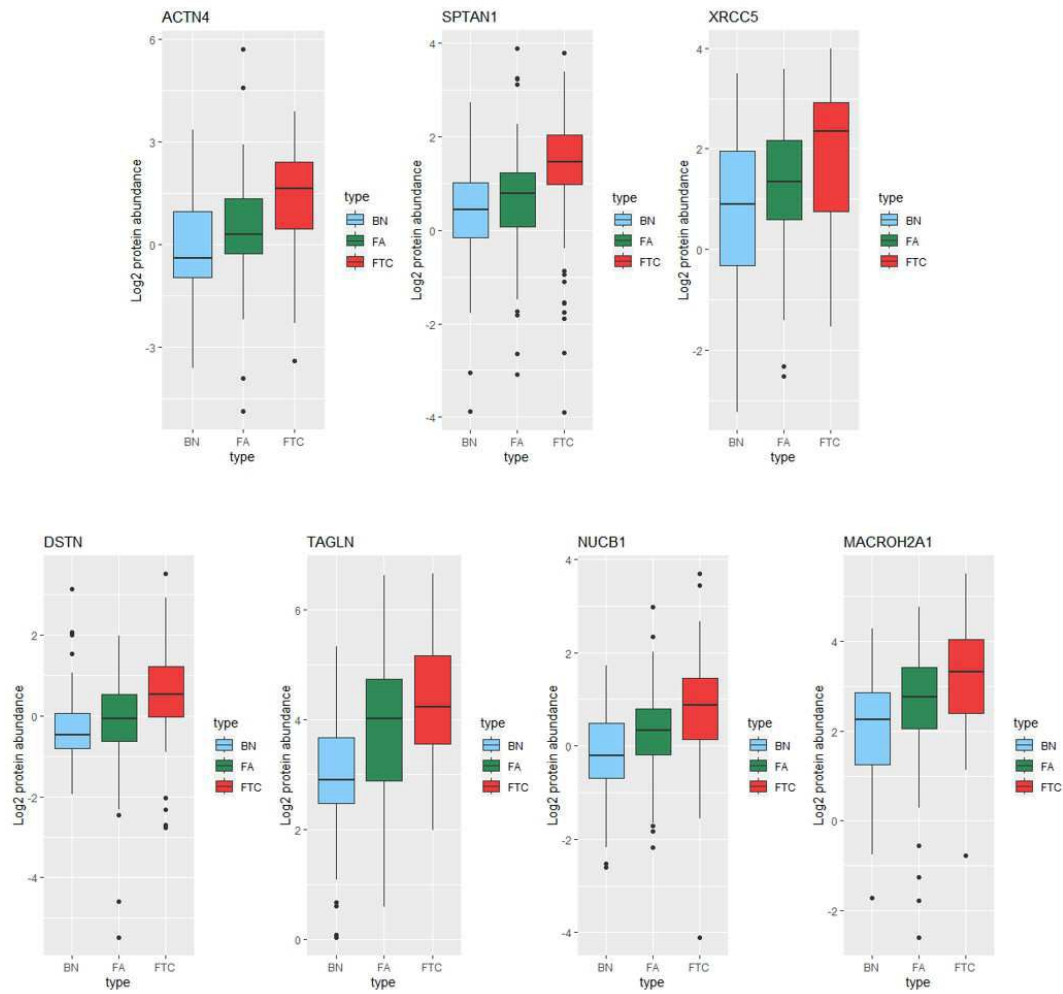


도면5d

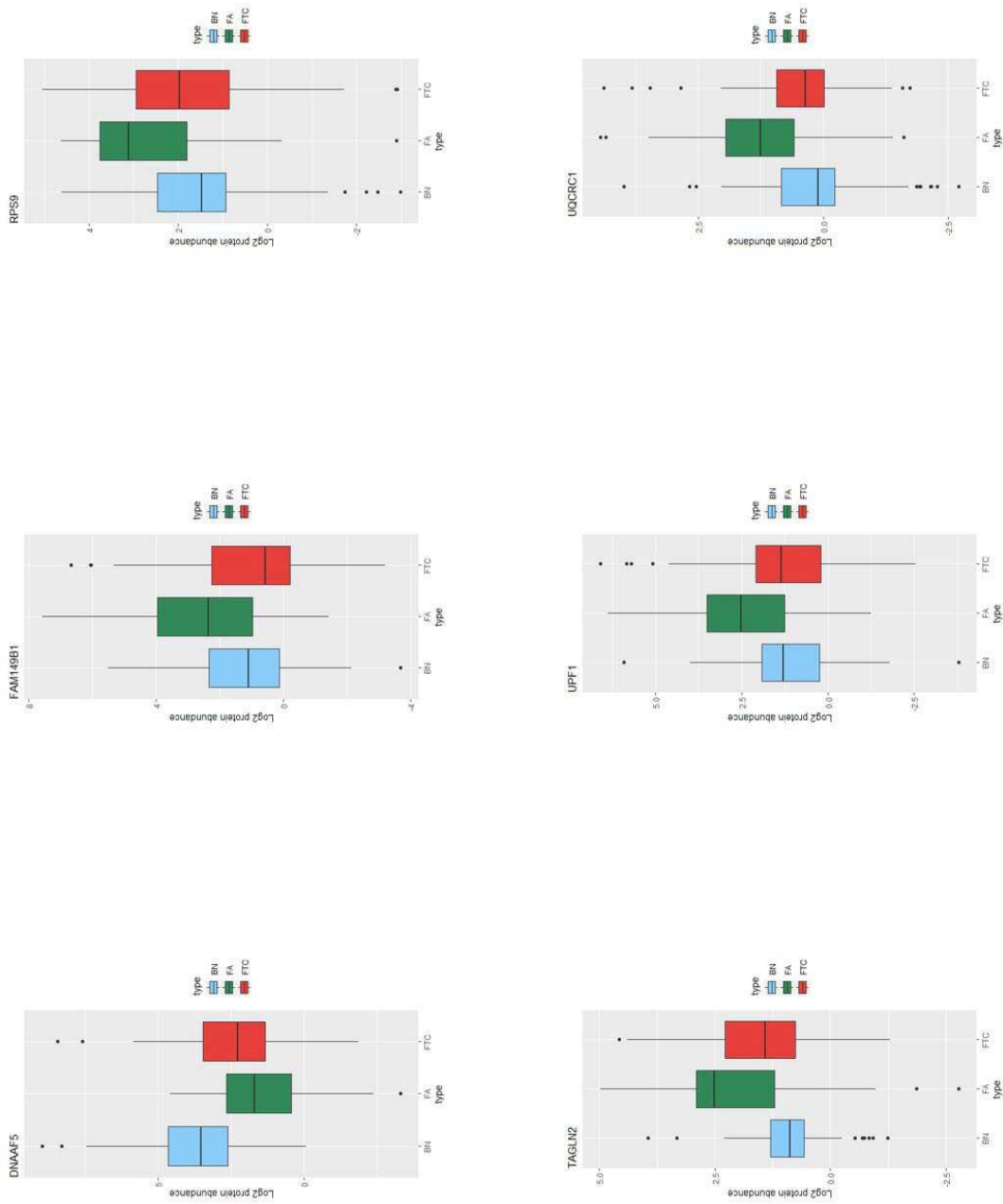


Index	UniProt Accession	Gene	Importance	Probability of selection	Selection	AUROC (FTC vs. BN + FA)
1	Q13813	SPTAN1	0.351	0.921	Y	0.696
2	Q01995	TAGLN	0.312	0.893	Y	0.690
3	P13010	XRCC5	0.523	0.809	Y	0.684
4	O75367	MACROH2A1	0.440	0.802	Y	0.720
5	P60981	DSTN	0.344	0.795	Y	0.709
6	Q02818	NUCB1	0.308	0.754	Y	0.694
7	Q43707	ACTN4	0.325	0.750	Y	0.687
8	P68871	HBB	0.370	0.716	N	0.649
9	P02511	CRYAB	0.262	0.704	N	0.663
10	P42858	HTT	0.328	0.682	N	0.669
11	A2PYH4	HFM1	0.308	0.674	N	0.645
12	Q53EL6	PDCD4	0.403	0.657	N	0.711
13	Q14103	HNRNPD	0.399	0.640	N	0.662
14	O75874	IDH1	0.253	0.636	N	0.686
15	P11142	HSPA8	0.265	0.614	N	0.677
16	P30040	ERP29	0.480	0.609	N	0.663
17	P78371	CCT2	0.265	0.528	N	0.688
18	P51888	PRELP	0.303	0.502	N	0.679
19	P17980	PSMC3	0.255	0.473	N	0.652
20	Q8N8W4	PNPLA1	0.284	0.459	N	0.650
21	P55327	TPD52	0.248	0.451	N	0.587
22	P13797	PLS3	0.304	0.450	N	0.680
23	P23528	CFL1	0.303	0.444	N	0.679
24	P12956	XRCC6	0.273	0.443	N	0.693
25	P29401	TKT	0.283	0.434	N	0.643
26	Q96J01	THOC3	0.305	0.434	N	0.671
27	Q92625	ANKS1A	0.289	0.403	N	0.675
28	P27695	APEX1	0.299	0.384	N	0.682
29	P26599	PTBP1	0.251	0.373	N	0.670
30	Q9Y4R8	TELO2	0.252	0.339	N	0.639
31	Q14697	GANAB	0.292	0.317	N	0.652
32	Q63ZY3	KANK2	0.246	0.288	N	0.644
33	Q8WVM8	SCFD1	0.249	0.277	N	0.551
34	P04264	KRT1	0.304	0.255	N	0.661
35	P35527	KRT9	0.249	0.199	N	0.623
36	Q9UM73	ALK	0.245	0.158	N	0.648
37	P05091	ALDH2	0.238	0.145	N	0.593

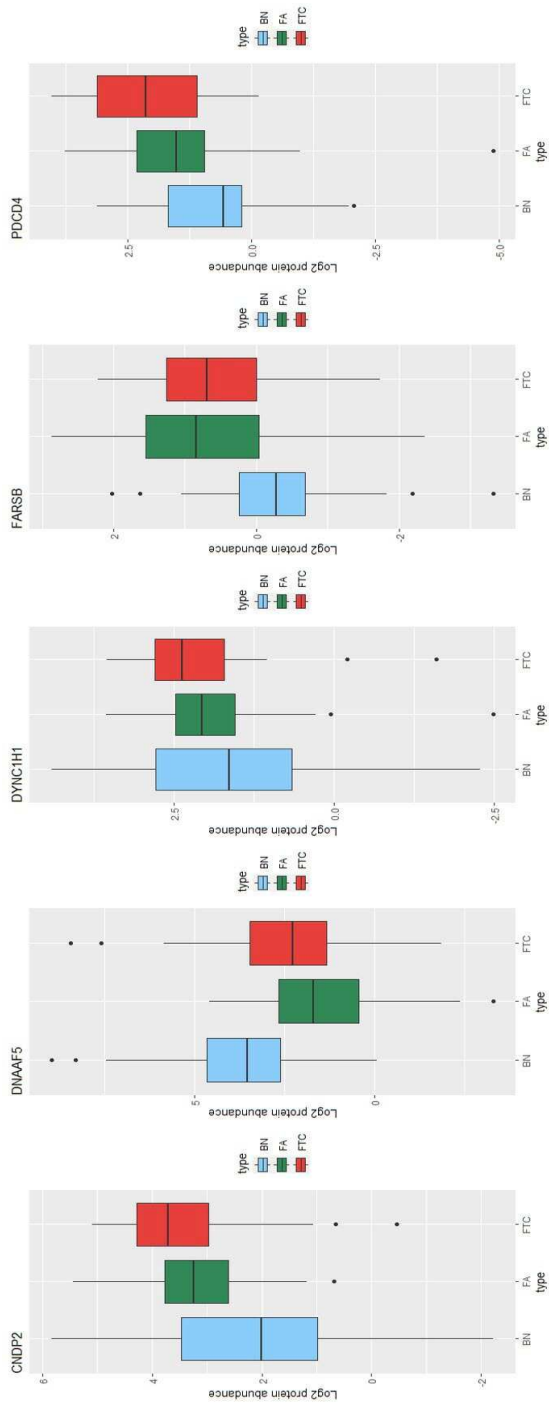
도면5e



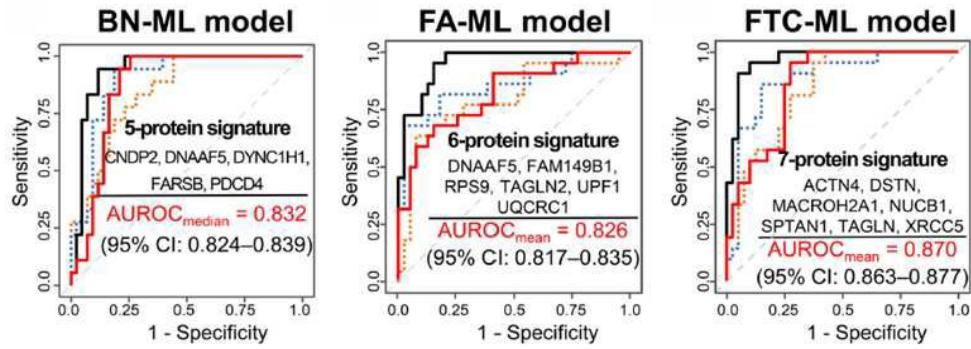
도면5f



도면5g



도면5h



도면5i

