(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2022-0129955
(43) 공개일자 2022년09월26일

(51) 국제특허분류(Int. Cl.)
G01N 33/68 (2006.01)
(52) CPC특허분류
G01N 33/6848 (2013.01)
G01N 2800/52 (2021.08)
(21) 출원번호 10-2021-0034926
(22) 출원일자 2021년03월17일
심사청구일자 2022년01월12일

(71) 출원인
재단법인 아산사회복지재단
서울특별시 송파구 올림픽로43길 88 (풍납동)
울산대학교 산학협력단
울산광역시 남구 대학로 93(무거동)
(72) 발명자
이새별
서울특별시 송파구 토성로 38-6, 한강극동아파트 106동 202호
안세현
서울특별시 강남구 도산대로87길 23, 11차 301호
(뒷면에 계속)
(74) 대리인
이명진

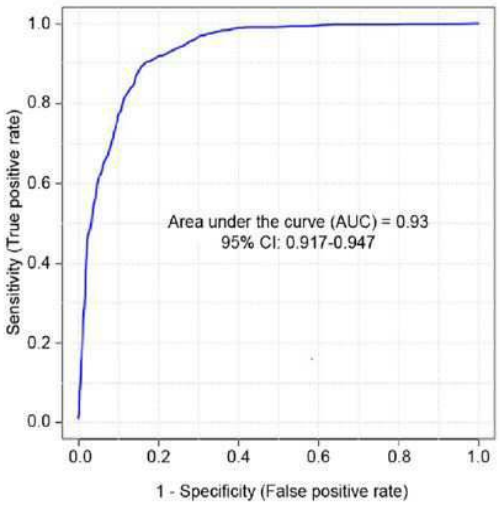
전체 청구항 수 : 총 19 항

(54) 발명의 명칭 **MALDI-TOF MS를 이용한 유방암 진단 또는 예후 예측을 위한 정보제공방법 및 이를 위한 키트**

(57) 요 약

본 발명자들은 다양한 병기 및 서브타입의 유방암 환자 및 정상인으로부터 혈청 *N*-글라이칸, 단백질, 및 펩티드를 추출하고 MALDI-TOF 질량분석법을 이용하여 분자량을 분석한 결과, 정상인과 유방암 환자간에 상이한 패턴을 가지는 *N*-글라이칸, 단백질, 및 펩티드의 *m/z* 값을 확인하였다. 따라서 본 발명자들이 발굴한 상기 *m/z* 값에 대하여 MALDI-TOF 질량분석법으로 피검체의 혈청 *N*-글라이칸, 단백질, 및 펩티드의 피크 강도를 측정하고 대조군과 비교하면 유방암 발생 여부는 물론 유방암 병기, 서브타입, 및 예후 등을 정확하고 신속하게 판단할 수 있다. 따라서, 본 발명에 따른 유방암 진단방법 등은 비침습적인 유방암 조기 진단방법으로 다양하게 활용될 수 있을 것으로 기대된다.

대 표 도 - 도2



(72) 발명자

이종원

서울특별시 강남구 도곡로14길 12, 101동 303호

이우창

서울특별시 송파구 잠실로 88, 101동 2301호

고명수

서울특별시 강남구 남부순환로365길 33, 101동 1203호

김지선

서울특별시 강남구 학동로 405, 102-601

김희정

서울특별시 강남구 선릉로 8, 223동 809호(개포동, 래미안블레스티지)

이예라

경기도 남양주시 다산중앙로81번길 25, 3511동 2303호

이 발명을 지원한 국가연구개발사업

과제고유번호	1345315338
과제번호	2018R1D1A1B07049969
부처명	교육부
과제관리(전문)기관명	한국연구재단
연구사업명	개인기초연구(교육부)(R&D)
연구과제명	혈중 글리칸을 이용한 유방암 진단 및 유방암 환자에서의 임상적 의의
기 여 율	50/100
과제수행기관명	서울아산병원
연구기간	2018.06.01 ~ 2021.05.31

이 발명을 지원한 국가연구개발사업

과제고유번호	1465027590
과제번호	HI18C1149000019
부처명	보건복지부
과제관리(전문)기관명	한국보건산업진흥원
연구사업명	연구자 주도 질병극복연구사업
연구과제명	혈중 글리칸을 이용한 유방암 진단 및 재발 모니터링
기 여 율	50/100
과제수행기관명	서울아산병원
연구기간	2018.04.30 ~ 2019.12.31

명 세 서

청구범위

청구항 1

1825, 1704, 1542, 1501, 1339, 1850, 1751, 1565, 1282, 1971, 1793, 1792, 1631, 2069, 2067, 1907, 1905, 1136, 1622, 1768, 1444, 1606, 1460, 1298, 1745, 1743, 1581, 1419, 1257, 1095, 및 1980 *m/z*로 이루어진 군에서 선택된 어느 하나 이상의 *m/z* 값에 대한 질량분석을 수행하는 단계를 포함하는, 유방암 진단을 위한 정보제공방법으로서,

(a) 피검체로부터 분리된 생물학적 시료에서 *N*-글라이칸을 분리하여 정제하는 단계;

(b) 단계 (a)에서 수득한 *N*-글라이칸을 대상으로 MALDI-TOF 질량분석을 수행하여 상기 *m/z* 값에 대한 피크 강도(intensity)를 측정하는 단계;

(c) 단계 (b)에서 측정한 피검체의 상기 *m/z* 값에 대한 피크 강도를 대조군의 상기 *m/z* 값에 대한 피크 강도와 비교하는 단계; 및

(d) 상기 *m/z* 값에 대한 피검체의 피크 강도가 대조군에 비해 높은 경우 유방암으로 판정하거나, 또는 유방암에 걸릴 위험이 있는 것으로 판정하는 단계를 포함하는 것인, 유방암 진단을 위한 정보제공방법.

청구항 2

제1항에 있어서,

(e) 1136 또는 1339 *m/z* 값에 대한 피크 강도가 대조군에 비해 피검체에서 높은 경우 유방암 1기로 판정하는 단계를 더 포함하는, 유방암 진단을 위한 정보제공방법.

청구항 3

제1항에 있어서,

(e) 1606, 1768, 1444, 1460, 및 1622로 이루어진 군에서 선택된 어느 하나 이상의 *m/z* 값에 대한 피크 강도가 대조군에 비해 피검체에서 높은 경우 유방암 2 내지 4기로 판정하는 단계를 더 포함하는, 유방암 진단을 위한 정보제공방법.

청구항 4

제1항에 있어서,

(e) 1095, 1136, 1257, 1282, 1298, 1419, 1444, 1460, 1565, 1581, 1631, 1743, 1745, 1751, 1768, 1793, 1905, 1907, 2067, 및 2069 *m/z*로 이루어진 군에서 선택된 어느 하나 이상의 *m/z* 값에 대한 피크 강도가 대조군에 비해 피검체에서 높은 경우 luminal 1 타입인 유방암으로 판정하거나, luminal 1 타입인 유방암에 걸릴 위험이 있는 것으로 판정하는 단계를 더 포함하는, 유방암 진단을 위한 정보제공방법.

청구항 5

제1항에 있어서,

(e) 1095, 1136, 1257, 1298, 1339, 1419, 1460, 1542, 1565, 1581, 1606, 1743, 1745, 1751, 1905, 1907, 2067, 및 2069 *m/z*로 이루어진 군에서 선택된 어느 하나 이상의 *m/z* 값에 대한 피크 강도가 대조군에 비해 피검

체에서 높은 경우 luminal 2 타입인 유방암으로 판정하거나, luminal 2 타입인 유방암에 걸릴 위험이 있는 것으로 판정하는 단계를 더 포함하는, 유방암 진단을 위한 정보제공방법.

청구항 6

제1항에 있어서,

(e) 1095, 1257, 1298, 1339, 1419, 1460, 1542, 1565, 1581, 1606, 1622, 1704, 1743, 1745, 1751, 1768, 1905, 1907, 2067, 및 2069 *m/z*로 이루어진 군에서 선택된 어느 하나 이상의 *m/z* 값에 대한 피크 강도가 대조군에 비해 피검체에서 높은 경우 non-luminal HER2+ 타입인 유방암으로 판정하거나, non-luminal HER2+ 타입인 유방암에 걸릴 위험이 있는 것으로 판정하는 단계를 더 포함하는, 유방암 진단을 위한 정보제공방법.

청구항 7

제1항에 있어서,

(e) 1095, 1136, 1282, 1298, 1339, 1444, 1542, 1565, 1631, 1751, 1980, 2067, 및 2069 *m/z*로 이루어진 군에서 선택된 어느 하나 이상의 *m/z* 값에 대한 피크 강도가 대조군에 비해 피검체에서 높은 경우 삼중음성(triple-negative) 타입인 유방암으로 판정하거나, 삼중음성 타입인 유방암에 걸릴 위험이 있는 것으로 판정하는 단계를 더 포함하는, 유방암 진단을 위한 정보제공방법.

청구항 8

제1항에 있어서,

(e) 1095, 1257, 1298, 1339, 1419, 1444, 1460, 1542, 1565, 1581, 1606, 1622, 1631, 1704, 1743, 1745, 1751, 1768, 1905, 1907, 2067, 및 2069 *m/z*로 이루어진 군에서 선택된 어느 하나 이상의 *m/z* 값에 대한 피크 강도가 대조군에 비해 피검체에서 높은 경우 림프절 침윤이 일어난 것으로 판정하는 단계를 더 포함하는, 유방암 진단을 위한 정보제공방법.

청구항 9

제1항에 있어서,

상기 대조군은 정상인으로부터 분리된 생물학적 시료로부터 분리한 *N*-글라이칸인 것을 특징으로 하는, 유방암 진단을 위한 정보제공방법.

청구항 10

제1항에 있어서,

상기 시료는 혈액, 혈장, 또는 혈청인 것을 특징으로 하는, 유방암 진단을 위한 정보제공방법.

청구항 11

제1항에 있어서,

상기 *m/z* 값은 양이온 모드의 *m/z* 값인 것을 특징으로 하는, 유방암 진단을 위한 정보제공방법.

청구항 12

제1항에 있어서,

상기 유방암은 침습성 유방암(invasive ductal carcinoma)인 것을 특징으로 하는, 유방암 진단을 위한 정보제공방법.

청구항 13

2138, 1663, 및 1419 *m/z*로 이루어진 1군; 및 1850 및 1688 *m/z*로 이루어진 2군에서 선택된 어느 하나 이상의 *m/z* 값에 대한 질량분석을 수행하는 단계를 포함하는, 유방암 진단 또는 예후 예측을 위한 정보제공방법으로서,

(a) 피검체로부터 분리된 생물학적 시료에서 *N*-글라이칸을 분리하여 정제하는 단계;

(b) 단계 (a)에서 수득한 *N*-글라이칸을 대상으로 MALDI-TOF 질량분석을 수행하여 상기 *m/z* 값에 대한 피크 강도(intensity)를 측정하는 단계;

(c) 단계 (b)에서 측정한 피검체의 상기 *m/z* 값에 대한 피크 강도를 대조군의 상기 *m/z* 값에 대한 피크 강도와 비교하는 단계; 및

(d) 1군에 속하는 *m/z* 값에 대한 피검체의 피크 강도가 대조군에 비해 높거나, 2군에 속하는 *m/z* 값에 대한 피검체의 피크 강도가 대조군에 비해 낮은 경우, 유방암으로 판정하거나 또는 유방암 재발 위험이 있는 것으로 판정하는 단계를 포함하는 것인, 유방암 진단 또는 예후 예측을 위한 정보제공방법.

청구항 14

제13항에 있어서,

상기 대조군은 정상인으로부터 분리된 생물학적 시료로부터 분리한 *N*-글라이칸인 것을 특징으로 하는, 유방암 진단 또는 예후 예측을 위한 정보제공방법.

청구항 15

1565, 1727, 1881, 및 2012 *m/z*로 이루어진 군에서 선택된 어느 하나 이상의 *m/z* 값에 대한 질량분석을 수행하는 단계를 포함하는, 유방암 예후 예측을 위한 정보제공방법으로서,

(a) 피검체로부터 분리된 생물학적 시료에서 *N*-글라이칸을 분리하여 정제하는 단계;

(b) 단계 (a)에서 수득한 *N*-글라이칸을 대상으로 MALDI-TOF 질량분석을 수행하여 상기 *m/z* 값에 대한 피크 강도(intensity)를 측정하는 단계;

(c) 단계 (b)에서 측정한 피검체의 상기 *m/z* 값에 대한 피크 강도를 대조군의 상기 *m/z* 값에 대한 피크 강도와 비교하는 단계; 및

(d) 상기 *m/z* 값에 대한 피검체의 피크 강도가 대조군에 비해 낮은 경우, 유방암 재발 위험이 있는 것으로 판정하는 단계를 포함하는 것인, 유방암 예후 예측을 위한 정보제공방법.

청구항 16

제15항에 있어서,

상기 대조군은 유방암 완치 환자 또는 유방암 치료 후 질병증거 없는(no evidence of disease, NED) 환자로부터 분리된 생물학적 시료로부터 분리한 *N*-글라이칸인 것을 특징으로 하는, 유방암 예후 예측을 위한 정보제공방법.

청구항 17

2896, 3444, 2550, 3237, 2656, 6888, 11099, 3879, 7757, 2101, 4206, 2373, 3260, 2167, 2057, 및 4207 *m/z* 로 이루어진 1군; 및 2447, 5001, 4150, 2724, 4299, 3167, 4280, 2352, 및 2111 *m/z*로 이루어진 2군에서 선택된 어느 하나 이상의 *m/z* 값에 대한 질량분석을 수행하는 단계를 포함하는, 유방암 진단을 위한 정보제공방법으로서,

(a) 피검체로부터 분리된 생물학적 시료에서 단백질 또는 펩티드를 분리하는 단계;

(b) 단계 (a)에서 수득한 단백질 또는 펩티드를 대상으로 MALDI-TOF 질량분석을 수행하여 상기 *m/z* 값에 대한 피크 강도(intensity)를 측정하는 단계;

(c) 단계 (b)에서 측정한 피검체의 상기 *m/z* 값에 대한 피크 강도를 대조군의 상기 *m/z* 값에 대한 피크 강도와 비교하는 단계; 및

(d) 1군에 속하는 *m/z* 값에 대한 피검체의 피크 강도가 대조군에 비해 높은 경우 또는 2군에 속하는 *m/z* 값에 대한 피검체의 피크 강도가 대조군에 비해 낮은 경우 유방암으로 판정하거나, 또는 유방암에 걸릴 위험이 있는 것으로 판정하는 단계를 포함하는 것인, 유방암 진단을 위한 정보제공방법.

청구항 18

제17항에 있어서,

상기 대조군은 정상인으로부터 분리된 생물학적 시료로부터 분리한 단백질 또는 펩티드인 것을 특징으로 하는, 유방암 진단을 위한 정보제공방법.

청구항 19

유방암 진단 또는 예후 예측을 위한 MALDI-TOF 질량분석 키트로서, MALDI 플레이트; 이온화제(ionizing agent); 및 제1항 내지 제18항 중 어느 한 항의 정보제공방법이 기재된 설명서를 포함하는, MALDI-TOF 질량분석 키트.

발명의 설명

기 술 분 야

[0001] 본 발명은 MALDI-TOF 질량분석법을 기반으로 한, 유방암 진단 또는 예후 예측을 위한 정보제공방법 등에 관한 것이다.

배 경 기 술

[0003] 유방암은 세계에서 가장 흔한 여성암으로서, 전체 여성암의 24.2%를 차지할 뿐만 아니라 사망률 또한 15%로 가장 높다. 우리나라의 경우 유방암은 젊은 여성(40대 및 50대)에게서 많이 발견되며, 국내 여성암 발생 빈도 중 1위를 기록하고 있다. 미국이나 유럽과 비교했을 때 유방암 사망률은 낮은 수준이나, 유방암 환자 비율이 점차 증가하고 있어 조기 점진의 필요성이 대두되고 있다. 유방암은 조직검사, 유방조영술, 초음파검사, 자기공명영상 등 다양한 방법으로 진단되고 있다. 그러나 위와 같은 진단방법들은 대개 침습적이고 많은 시간이 소요되며, 많은 수의 환자를 동시에 진단하는 데는 적합하지 않다는 단점이 있다. 때문에, 보다 간단하면서 비침습적인 진단방법의 발굴 필요성이 제기되어 왔으며, 실제로 이와 같은 방법들이 유방암 진단에 사용되고 있다. 대표적인 것이 환자의 체액 내에 존재하는 종양 관련 단백질 또는 핵산을 분석하는 것이다.

[0004] 당화(glycosylation)는 다양한 글리코실전달효소(glycosyltransferase)에 의해 촉매되는 기작으로, 특히 N-당화(N-linked glycosylation)는 진핵생물의 막 결합 단백질의 가장 흔한 번역 후 변형(post-translational modification) 중 하나이다. 글리코실전달효소는 당화를 촉매하여 당단백질(glycoproteins), 당지질(glycolipids), 및 프로테오글리칸(proteoglycan)과 같은 다양한 복합 탄수화물의 생성을 유도하며, 정상적인 당화는 단백질 생합성, 단백질의 안정, 단백질의 세포 내 국소화(localization) 등에 영향을 미친다. 글리코실전달효소의 활성과 발현은 암과 같은 다양한 병리학적 상황에서 변화하며, 따라서 암환자는 정상인과 비교하여

상이한 당화 패턴을 갖는 것으로 알려져 있다. 일부 당단백질 등은 종양에서 분비되거나 배출되기 때문에 혈청을 이용하여 암 관련 당단백질 또는 암 관련 당화 패턴의 변화를 검출할 수 있다. 실제로 최근에는 혈중 글라이칸(glycan), 특히 N-글라이칸을 암 특이적 바이오마커로 활용하여 암을 진단하고자 하는 연구가 활발히 진행되고 있다.

[0005] MALDI-TOF 질량분석법(Matrix-assisted laser desorption/ionization time-of-flight mass spectrometry, MALDI-TOF MS)은 분자량이 비교적 큰 단백질 등을 이온화시켜 분자량을 분석하는 질량분석 방법으로서, 단백질, 펩티드, 당 등을 포함한 유기분자의 분석에 널리 사용되고 있다. MALDI-TOF 질량분석법은 비침습적인 동시에 다량의 유기분자를 신속하게 검출할 수 있으므로 이를 이용하여 암을 진단하기 위한 연구가 지속적으로 이루어지고 있으나, 암을 조기에 효과적으로 진단할 수 있는 MALDI-TOF 기반 방법은 여전히 미비한 실정이다.

선행기술문헌

특허문헌

[0007] (특허문헌 0001) 대한민국 공개특허공보 제10-2014-0002149호

발명의 내용

해결하려는 과제

[0008] 본 발명은 상기와 같은 문제점을 해결하기 위해 안출된 것으로서, 본 발명자들은 MALDI-TOF 질량분석법을 바탕으로 유방암 환자에서 특이적인 *N*-글라이칸, 단백질, 및 펩티드 마커 등을 확인하여 본 발명을 완성하였다.

[0010] 따라서, 본 발명의 목적은 MALDI-TOF 질량분석을 수행하는 단계를 포함하는, 유방암 진단 또는 예후 예측을 위한 정보제공방법을 제공하는 것이다.

[0011] 본 발명의 다른 목적은 MALDI-TOF 질량분석을 수행하는 단계를 포함하는, 유방암 병기 판정을 위한 정보제공방법을 제공하는 것이다.

[0012] 본 발명의 또 다른 목적은 MALDI-TOF 질량분석을 수행하는 단계를 포함하는, 유방암 서브타입 판정을 위한 정보제공방법을 제공하는 것이다.

[0013] 본 발명의 또 다른 목적은 MALDI-TOF 질량분석을 수행하는 단계를 포함하는, 림프절 침윤 여부 예측을 위한 정보제공방법을 제공하는 것이다.

[0014] 본 발명의 또 다른 목적은 유방암 진단 또는 예후 예측을 위한 MALDI-TOF 질량분석 키트를 제공하는 것이다.

[0016] 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

[0018] 본 발명은 1825, 1704, 1542, 1501, 1339, 1850, 1751, 1565, 1282, 1971, 1793, 1792, 1631, 2069, 2067, 1907, 1905, 1136, 1622, 1768, 1444, 1606, 1460, 1298, 1745, 1743, 1581, 1419, 1257, 1095, 및 1980 *m/z*로 이루어진 군에서 선택된 어느 하나 이상의 *m/z* 값에 대한 질량분석을 수행하는 단계를 포함하는, 유방암 진단을 위한 정보제공방법으로서,

[0019] (a) 피검체로부터 분리된 생물학적 시료에서 *N*-글라이칸을 분리하여 정제하는 단계;

[0020] (b) 단계 (a)에서 수득한 *N*-글라이칸을 대상으로 MALDI-TOF 질량분석을 수행하여 상기 *m/z* 값에 대한 피크 강도

(intensity)를 측정하는 단계;

[0021] (c) 단계 (b)에서 측정한 피검체의 상기 *m/z* 값에 대한 피크 강도를 대조군의 상기 *m/z* 값에 대한 피크 강도와 비교하는 단계; 및

[0022] (d) 상기 *m/z* 값에 대한 피검체의 피크 강도가 대조군에 비해 높은 경우 유방암으로 판정하거나, 또는 유방암에 걸릴 위험이 있는 것으로 판정하는 단계를 포함하는 것인, 유방암 진단을 위한 정보제공방법을 제공한다.

[0023] 본 발명의 일 구현예에서, 상기 유방암 진단을 위한 정보제공방법은 (e) 1136 또는 1339 *m/z* 값에 대한 피크 강도가 대조군에 비해 피검체에서 높은 경우 유방암 1기로 판정하는 단계를 더 포함하는 것일 수 있다.

[0024] 본 발명의 다른 구현예에서, 상기 유방암 진단을 위한 정보제공방법은 (e) 1606, 1768, 1444, 1460, 및 1622로 이루어진 군에서 선택된 어느 하나 이상의 *m/z* 값에 대한 피크 강도가 대조군에 비해 피검체에서 높은 경우 유방암 2 내지 4기로 판정하는 단계를 더 포함하는 것일 수 있다.

[0025] 본 발명의 또 다른 구현예에서, 상기 유방암 진단을 위한 정보제공방법은 (e) 1095, 1136, 1257, 1282, 1298, 1419, 1444, 1460, 1565, 1581, 1631, 1743, 1745, 1751, 1768, 1793, 1905, 1907, 2067, 및 2069 *m/z*로 이루어진 군에서 선택된 어느 하나 이상의 *m/z* 값에 대한 피크 강도가 대조군에 비해 피검체에서 높은 경우 luminal 1 타입인 유방암으로 판정하거나, luminal 1 타입인 유방암에 걸릴 위험이 있는 것으로 판정하는 단계를 더 포함할 수 있다.

[0026] 본 발명의 또 다른 구현예에서, 상기 유방암 진단을 위한 정보제공방법은 (e) 1095, 1136, 1257, 1298, 1339, 1419, 1460, 1542, 1565, 1581, 1606, 1743, 1745, 1751, 1905, 1907, 2067, 및 2069 *m/z*로 이루어진 군에서 선택된 어느 하나 이상의 *m/z* 값에 대한 피크 강도가 대조군에 비해 피검체에서 높은 경우 luminal 2 타입인 유방암으로 판정하거나, luminal 2 타입인 유방암에 걸릴 위험이 있는 것으로 판정하는 단계를 더 포함할 수 있다.

[0027] 본 발명의 또 다른 구현예에서, 상기 유방암 진단을 위한 정보제공방법은 (e) 1095, 1257, 1298, 1339, 1419, 1460, 1542, 1565, 1581, 1606, 1622, 1704, 1743, 1745, 1751, 1768, 1905, 1907, 2067, 및 2069 *m/z*로 이루어진 군에서 선택된 어느 하나 이상의 *m/z* 값에 대한 피크 강도가 대조군에 비해 피검체에서 높은 경우 non-luminal HER2+ 타입인 유방암으로 판정하거나, non-luminal HER2+ 타입인 유방암에 걸릴 위험이 있는 것으로 판정하는 단계를 더 포함할 수 있다.

[0028] 본 발명의 또 다른 구현예에서, 상기 유방암 진단을 위한 정보제공방법은 (e) 1095, 1136, 1282, 1298, 1339, 1444, 1542, 1565, 1631, 1751, 1980, 2067, 및 2069 *m/z*로 이루어진 군에서 선택된 어느 하나 이상의 *m/z* 값에 대한 피크 강도가 대조군에 비해 피검체에서 높은 경우 삼중음성(triple-negative) 타입인 유방암으로 판정하거나, 삼중음성 타입인 유방암에 걸릴 위험이 있는 것으로 판정하는 단계를 더 포함할 수 있다.

[0029] 본 발명의 또 다른 구현예에서, 상기 유방암 진단을 위한 정보제공방법은 (e) 1095, 1257, 1298, 1339, 1419, 1444, 1460, 1542, 1565, 1581, 1606, 1622, 1631, 1704, 1743, 1745, 1751, 1768, 1905, 1907, 2067, 및 2069 *m/z*로 이루어진 군에서 선택된 어느 하나 이상의 *m/z* 값에 대한 피크 강도가 대조군에 비해 피검체에서 높은 경우 림프절 침윤이 일어난 것으로 판정하는 단계를 더 포함할 수 있다.

[0030] 본 발명의 또 다른 구현예에서, 상기 대조군은 정상인으로부터 분리된 생물학적 시료로부터 분리한 *N*-글라이칸일 수 있으나, 이에 한정되지 않는다.

[0031] 본 발명은 또한 2138, 1663, 및 1419 *m/z*로 이루어진 1군; 및 1850 및 1688 *m/z*로 이루어진 2군에서 선택된 어느 하나 이상의 *m/z* 값에 대한 질량분석을 수행하는 단계를 포함하는, 유방암 진단 및 예후 예측을 위한 정보제공방법으로서,

[0032] (a) 피검체로부터 분리된 생물학적 시료에서 *N*-글라이칸을 분리하여 정제하는 단계;

[0033] (b) 단계 (a)에서 수득한 *N*-글라이칸을 대상으로 MALDI-TOF 질량분석을 수행하여 상기 *m/z* 값에 대한 피크 강도(intensity)를 측정하는 단계;

[0034] (c) 단계 (b)에서 측정한 피검체의 상기 *m/z* 값에 대한 피크 강도를 대조군의 상기 *m/z* 값에 대한 피크 강도와 비교하는 단계; 및

[0035] (d) 1군에 속하는 *m/z* 값에 대한 피검체의 피크 강도가 대조군에 비해 높거나, 2군에 속하는 *m/z* 값에 대한 피

검체의 피크 강도가 대조군에 비해 낮은 경우, 유방암으로 판정하거나 또는 유방암 재발 위험이 있는 것으로 판정하는 단계를 포함하는 것인, 유방암 진단 또는 예후 예측을 위한 정보제공방법을 제공한다.

[0036] 본 발명의 일 구현예에서, 상기 대조군은 정상인으로부터 분리된 생물학적 시료로부터 분리한 *N*-글라이칸일 수 있으나, 이에 한정되지 않는다.

[0037] 본 발명은 또한, 1565, 1727, 1881, 및 2012 *m/z*로 이루어진 군에서 선택된 어느 하나 이상의 *m/z* 값에 대한 질량분석을 수행하는 단계를 포함하는, 유방암 예후 예측을 위한 정보제공방법으로서,

[0038] (a) 피검체로부터 분리된 생물학적 시료에서 *N*-글라이칸을 분리하여 정제하는 단계;

[0039] (b) 단계 (a)에서 수득한 *N*-글라이칸을 대상으로 MALDI-TOF 질량분석을 수행하여 상기 *m/z* 값에 대한 피크 강도(intensity)를 측정하는 단계;

[0040] (c) 단계 (b)에서 측정한 피검체의 상기 *m/z* 값에 대한 피크 강도를 대조군의 상기 *m/z* 값에 대한 피크 강도와 비교하는 단계; 및

[0041] (d) 상기 *m/z* 값에 대한 피검체의 피크 강도가 대조군에 비해 낮은 경우, 유방암 재발 위험이 있는 것으로 판정하는 단계를 포함하는 것인, 유방암 예후 예측을 위한 정보제공방법을 제공한다.

[0042] 본 발명의 일 구현예에서, 상기 대조군은 유방암 완치 환자 또는 유방암 치료 후 질병증거 없는(no evidence of disease, NED) 환자로부터 분리된 생물학적 시료로부터 분리한 *N*-글라이칸일 수 있으나, 이에 한정되지 않는다.

[0043] 본 발명은 또한 2896, 3444, 2550, 3237, 2656, 6888, 11099, 3879, 7757, 2101, 4206, 2373, 3260, 2167, 2057, 및 4207 *m/z*로 이루어진 1군; 및 2447, 5001, 4150, 2724, 4299, 3167, 4280, 2352, 및 2111 *m/z*로 이루어진 2군에서 선택된 어느 하나 이상의 *m/z* 값에 대한 질량분석을 수행하는 단계를 포함하는, 유방암 진단을 위한 정보제공방법으로서,

[0044] (a) 피검체로부터 분리된 생물학적 시료에서 단백질 또는 펩티드를 분리하는 단계;

[0045] (b) 단계 (a)에서 수득한 단백질 또는 펩티드를 대상으로 MALDI-TOF 질량분석을 수행하여 상기 *m/z* 값에 대한 피크 강도(intensity)를 측정하는 단계;

[0046] (c) 단계 (b)에서 측정한 피검체의 상기 *m/z* 값에 대한 피크 강도를 대조군의 상기 *m/z* 값에 대한 피크 강도와 비교하는 단계; 및

[0047] (d) 1군에 속하는 *m/z* 값에 대한 피검체의 피크 강도가 대조군에 비해 높은 경우 또는 2군에 속하는 *m/z* 값에 대한 피검체의 피크 강도가 대조군에 비해 낮은 경우 유방암으로 판정하거나, 또는 유방암에 걸릴 위험이 있는 것으로 판정하는 단계를 포함하는 것인, 유방암 진단을 위한 정보제공방법을 제공한다.

[0048] 본 발명의 일 구현예에서, 상기 대조군은 정상인으로부터 분리된 생물학적 시료로부터 분리한 단백질 또는 펩티드일 수 있으나, 이에 한정되지 않는다.

[0049] 본 발명에 있어서, 상기 시료는 혈액, 혈장, 또는 혈청일 수 있으나, 이에 한정되지 않는다.

[0050] 본 발명에 있어서, 상기 *m/z* 값은 양이온 모드의 *m/z* 값일 수 있으나, 이에 한정되지 않는다.

[0051] 본 발명에 있어서, 상기 유방암은 침습성 유방암(invasive ductal carcinoma)일 수 있으나, 이에 한정되지 않는다.

[0052] 본 발명은 또한 유방암 진단 또는 예후 예측을 위한 MALDI-TOF 질량분석 키트로서, MALDI 플레이트; 이온화제(ionizing agent); 및 상술한 정보제공방법이 기재된 설명서를 포함하는, MALDI-TOF 질량분석 키트를 제공한다.

[0053] 본 발명은 또한 본 발명에서 제시된 *m/z* 값 중 어느 하나 이상의 *m/z* 값에 대한 질량분석을 수행하는 단계를 포함하는, 유방암 진단 또는 예후 예측방법을 제공한다.

[0054] 본 발명은 또한 본 발명에서 제시된 *m/z* 값 중 어느 하나 이상의 *m/z* 값에 대한 질량분석을 수행하는 단계를 포함하는, 유방암 병기 판정방법을 제공한다.

[0055] 본 발명은 또한 본 발명에서 제시된 *m/z* 값 중 어느 하나 이상의 *m/z* 값에 대한 질량분석을 수행하는 단계를 포함하는, 유방암 서브타입 판정방법을 제공한다.

[0056] 본 발명은 또한 본 발명에서 제시된 *m/z* 값 중 어느 하나 이상의 *m/z* 값에 대한 질량분석을 수행하는 단계를 포

함하는, 유방암의 림프절 침윤 여부 예측방법을 제공한다.

발명의 효과

[0058] 본 발명자들은 다양한 병기 및 서브타입의 유방암 환자 및 정상인으로부터 혈청 *N*-글라이칸, 단백질, 및 펩티드를 추출하고 MALDI-TOF 질량분석법을 이용하여 분자량을 분석한 결과, 정상인과 유방암 환자간에 상이한 패턴을 가지는 *N*-글라이칸, 단백질, 및 펩티드의 *m/z* 값을 확인하였다. 따라서 본 발명자들이 발굴한 상기 *m/z* 값에 대하여 MALDI-TOF 질량분석법으로 피검체의 혈청 *N*-글라이칸, 단백질, 및 펩티드의 피크 강도를 측정하고 대조군과 비교하면 유방암 발생 여부는 물론 유방암 병기, 서브타입, 및 예후 등을 정확하고 신속하게 판단할 수 있다. 따라서, 본 발명에 따른 유방암 진단방법 등은 비침습적인 유방암 조기 진단방법으로 다양하게 활용될 수 있을 것으로 기대된다.

도면의 간단한 설명

[0060] 도 1은 실시예 1에서 선별한 *N*-글라이칸 *m/z* 프로파일에 대해 정상인과 유방암 환자의 *N*-글라이칸을 MALDI-TOF로 질량분석한 결과를 히트맵으로 나타낸 것이다.

도 2는 실시예 1에서 선별한 *N*-글라이칸 *m/z* 프로파일에 대한 정상인 및 유방암 환자 시료의 receiver operating characteristic(ROC) 커브 및 곡선하 면적(area under the curve, AUC)을 나타낸 것이다.

도 3은 실시예 1에서 선별한 *N*-글라이칸 *m/z* 프로파일에 대한 정상인과 유방암 환자의 정규화된 강도(intensity)를 나타낸 것이다.

도 4a는 실시예 1에서 선별한 *N*-글라이칸 *m/z* 프로파일에 대한 정상인과 1기 유방암 환자의 정규화된 강도에 대해 PCA를 수행하여 클러스터링한 결과를 나타낸다(파란색: 정상인, 분홍색: 암환자).

도 4b는 정상인과 2 내지 4기 유방암 환자의 *N*-글라이칸 *m/z* 프로파일의 PCA 결과를 나타낸다(파란색: 정상인, 분홍색: 암환자).

도 5a는 실시예 1에서 선별한 *N*-글라이칸 *m/z* 프로파일에 대한 정상인 및 1기 유방암 환자 시료의 ROC 커브 및 AUC를 나타낸다.

도 5b는 실시예 1에서 선별한 *N*-글라이칸 *m/z* 프로파일에 대한 정상인 및 2 내지 4기 유방암 환자 시료의 ROC 커브 및 AUC를 나타낸다.

도 6a는 실시예 1에서 선별한 *N*-글라이칸 *m/z* 프로파일에 대한 정상인과 유방암 HR+/HER2- 타입 환자의 정규화된 강도를 나타낸 것이다.

도 6b는 실시예 1에서 선별한 *N*-글라이칸 *m/z* 프로파일에 대한 정상인과 유방암 HR+/HER2+ 타입 환자의 정규화된 강도를 나타낸 것이다.

도 6c는 실시예 1에서 선별한 *N*-글라이칸 *m/z* 프로파일에 대한 정상인과 유방암 HR-/HER2+ 타입 환자의 정규화된 강도를 나타낸 것이다.

도 6d는 실시예 1에서 선별한 *N*-글라이칸 *m/z* 프로파일에 대한 정상인과 유방암 HR-/HER2- 타입 환자의 정규화된 강도를 나타낸 것이다.

도 7은 실시예 1에서 선별한 *N*-글라이칸 *m/z* 프로파일에 대한 정상인과 림프절 침윤이 일어난 유방암 환자의 정규화된 강도를 나타낸 것이다.

도 8는 실시예 2에서 선별한 *N*-글라이칸 *m/z* 프로파일에 대해 정상인과 유방암 환자의 *N*-글라이칸을 MALDI-TOF로 질량분석한 결과를 히트맵으로 나타낸 것이다.

도 9은 실시예 2에서 선별한 *N*-글라이칸 *m/z* 프로파일에 대한 정상인과 유방암 환자의 정규화된 강도에 대해 PCA를 수행하여 클러스터링한 결과를 나타낸다(삼각형: 정상인, 원형: 유방암 환자).

도 10은 실시예 2에서 선별한 *N*-글라이칸 *m/z* 프로파일에 대한 정상인 및 유방암 환자 시료의 ROC 커브 및 곡선

하 면적을 나타낸 것이다.

도 11은 ROC 분석에서 선택된 분류기(classifier)를 사용하여 정상인(흰색 원) 및 유방암 환자(검정색 원)의 각 샘플에 대한 분류 예측 정확도를 나타낸 것이다.

도 12는 실시예 2에서 선별한 *N*-글라이칸 *m/z* 프로파일에 대한 정상인(파란색) 및 유방암 환자(붉은색)의 정규화된 강도를 나타낸 것이다.

도 13은 실시예 2에서 선별한 *N*-글라이칸 *m/z* 값에 대한 로우(raw) 스펙트럼 프로파일 및 이들의 *N*-글라이칸 구조를 나타낸 것이다.

도 14는 실시예 2에서 선별한 *N*-글라이칸 *m/z* 프로파일에 대해 정상인, 유방암 환자 및 유방암 재발 환자의 *N*-글라이칸을 MALDI-TOF로 질량분석한 결과를 히트맵으로 나타낸 것이다.

도 15a 내지 15e는 실시예 2에서 선별한 *N*-글라이칸 *m/z* 값(1419, 1663, 1688, 1850, 및 2138 *m/z*)에 대한 정상인, 유방암 환자, 및 유방암 재발 환자의 정규화된 강도를 비교하여 나타낸 것이다.

도 16은 실시예 2에서 선별한 *N*-글라이칸 *m/z* 프로파일에 대한 정상인, 유방암 환자, 및 유방암 재발 환자의 정규화된 강도에 대해 PCA를 수행하여 클러스터링한 결과를 나타낸다(삼각형: 정상인, 원형: 유방암 환자, 사각형: 유방암 재발 환자).

도 17은 실시예 2에서 선별한 *N*-글라이칸 *m/z* 프로파일에 대해 정상인 시료 및 유방암 환자의 치료 전, 치료 중, 치료 후 시료의 *N*-글라이칸을 MALDI-TOF로 질량분석한 결과를 히트맵으로 나타낸 것이다.

도 18은 실시예 2에서 선별한 *N*-글라이칸 *m/z* 프로파일에 대한 정상인 시료 및 유방암 환자의 치료 전, 치료 중, 치료 후 시료의 정규화된 강도에 대해 PCA를 수행하여 클러스터링한 결과를 나타낸다(원형: 정상인, 삼각형: 유방암-치료 전, 사각형: 유방암-치료 중, 별표: 유방암-치료 후).

도 19는 실시예 2에서 선별한 *N*-글라이칸 *m/z* 프로파일에 대해 정상인 시료 및 유방암 환자의 치료 전, 치료 중, 치료 후 시료의 *N*-글라이칸을 MALDI-TOF로 질량분석한 결과를 히트맵으로 나타낸 것이다.

도 20a 내지 20e는 실시예 2에서 선별한 *N*-글라이칸 *m/z* 값(1419, 1663, 1688, 1850, 및 2138 *m/z*)에 대한 유방암 완치 환자(NED)의 치료 전, 중, 후 시료의 정규화된 강도(흰색 원) 및 유방암 재발 환자의 치료 전, 중, 후 시료의 정규화된 강도(검정색 원)를 나타낸다.

도 21는 실시예 2에서 선별한 *N*-글라이칸 *m/z* 프로파일에 대해 유방암 완치 환자 및 유방암 재발 환자의 *N*-글라이칸을 MALDI-TOF로 질량분석한 결과를 히트맵으로 나타낸 것이다.

도 22는 실시예 3에서 선별한 단백질 및/또는 펩티드 *m/z* 프로파일에 대해 정상인 및 유방암 환자의 단백질 및/또는 펩티드를 MALDI-TOF로 질량분석하여 얻은 강도를 비교하여 나타낸 것이다.

도 23는 실시예 3에서 선별한 단백질 및/또는 펩티드 *m/z* 프로파일에 대해 정상인 및 유방암 환자의 단백질 및/또는 펩티드를 MALDI-TOF로 질량분석한 결과를 히트맵으로 나타낸 것이다.

도 24는 실시예 3에서 선별한 단백질 및/또는 펩티드 *m/z* 프로파일에 대한 정상인 및 유방암 환자의 정규화된 강도를 워터폴 플롯(waterfall plot)으로 나타낸 것이다.

도 25은 실시예 3에서 선별한 단백질 및/또는 펩티드 *m/z* 프로파일에 대한 정상인 및 유방암 환자의 정규화된 강도에 대해 PCA를 수행하여 클러스터링한 결과를 나타낸 것이다.

도 26은 실시예 3에서 선별한 단백질 및/또는 펩티드 *m/z* 프로파일에 대한 정상인 및 유방암 환자 시료의 ROC 커브 및 곡선하 면적을 나타낸 것이다.

도 27a 내지 27d는 실시예 3에서 선별한 단백질 및/또는 펩티드 *m/z* 프로파일에 대한 정상인 또는 유방암 환자의 MALDI-TOF MS 스펙트럼을 비교하여 나타낸 것이다(도 27a: 2,000 내지 3,000 *m/z* 범위 분석 결과; 도 27b: 3,000 내지 3,500 *m/z* 범위 분석 결과; 도 27c: 4,000 내지 4,250 *m/z* 범위 분석 결과; 도 27d: 4,900 내지 5,100 *m/z* 범위 분석 결과).

발명을 실시하기 위한 구체적인 내용

[0061] 본 발명자들은 다양한 병기 및 서브타입의 유방암 환자 및 정상인으로부터 혈청 *N*-글라이칸, 단백질, 및 펩티드

를 추출하여 MALDI-TOF 질량분석법을 이용하여 분자량을 분석한 결과, 정상인과 유방암 환자간에 상이한 패턴을 가지는 *N*-글라이칸, 단백질, 또는 펩티드의 *m/z* 값을 확인하였다.

[0063] 따라서, 본 발명은 1825, 1704, 1542, 1501, 1339, 1850, 1751, 1565, 1282, 1971, 1793, 1792, 1631, 2069, 2067, 1907, 1905, 1136, 1622, 1768, 1444, 1606, 1460, 1298, 1745, 1743, 1581, 1419, 1257, 1095, 및 1980 *m/z*로 이루어진 군에서 선택된 어느 하나 이상의 *m/z* 값에 대한 질량분석을 수행하는 단계를 포함하는, 유방암 진단을 위한 정보제공방법으로서,

[0064] (a) 피검체로부터 분리된 생물학적 시료에서 *N*-글라이칸을 분리하여 정제하는 단계;

[0065] (b) 단계 (a)에서 수득한 *N*-글라이칸을 대상으로 MALDI-TOF 질량분석을 수행하여 상기 *m/z* 값에 대한 피크 강도(intensity)를 측정하는 단계;

[0066] (c) 단계 (b)에서 측정한 피검체의 상기 *m/z* 값에 대한 피크 강도를 대조군의 상기 *m/z* 값에 대한 피크 강도와 비교하는 단계; 및

[0067] (d) 상기 *m/z* 값에 대한 피검체의 피크 강도가 대조군에 비해 높은 경우 유방암으로 판정하거나, 또는 유방암에 걸릴 위험이 있는 것으로 판정하는 단계를 포함하는 것인, 유방암 진단을 위한 정보제공방법을 제공한다.

[0068] 보다 구체적으로, 상기 *m/z* 값은 차례로 1825.62, 1704.59, 1542.54, 1501.51, 1339.47, 1850.66, 1751.61, 1565.51, 1282.45, 1971.70, 1793.64, 1792.63, 1631.58, 2069.71, 2067.70, 1907.64, 1905.63, 1136.40, 1622.55, 1768.61, 1444.50, 1606.56, 1460.49, 1298.44, 1745.58, 1743.57, 1581.52, 1419.47, 1257.42, 1095.38, 및 1980.670으로 나타낼 수 있다.

[0069] 본 발명에 있어서, “글라이칸(glycan)”은 탄수화물이나 당(예컨대, 올리고당 또는 다당)이 사슬 구조로 이어진 것으로서, 구체적으로는 글리코시드 결합에 의해 연결된 당 잔기의 단량체 또는 중합체이다. 특히, "N-글라이칸"은 글리코콘쥬게이트로부터 방출되었지만 이전에는 질소 결합을 통해 글리코콘쥬게이트에 결합된 당의 중합체를 나타낸다. 글라이칸은 천연 당 잔기(예컨대, 글루코스, N-아세틸글루코사민, N-아세틸 뉴라민산, 갈락토스, 만노스, 푸코스, 헥소스, 아라비노스, 리보스, 자일로스, 등) 및/또는 변형 당(예컨대, 2'-플루오로리보스, 2'-데옥시리보스, 포스포만노스, 6'-설포 N-아세틸글루코사민, 등)을 포함할 수 있다. 또한, 글라이칸은 당 잔기의 단일 및 이종중합체를 포함할 수 있다. 글라이칸은 또한 글리코콘쥬게이트(예컨대, 당단백질, 당지질, 프로테오글리칸, 등)의 글리칸 구성요소를 포함할 수 있다. 나아가, 글라이칸은 또한 글리코콘쥬게이트로부터 절단되거나 방출된 글라이칸을 포함한, 유리 글라이칸을 포함한다. 가장 바람직하게는, 본 발명에 있어서 글라이칸은 단백질 등 기타 물질이 제거된 순수한 글라이칸을 의미한다.

[0070] 본 발명에 있어서, “진단”은 특정 질병 또는 질환에 대한 한 객체의 감수성(susceptibility)을 판정하는 것, 한 객체가 특정 질병 또는 질환을 현재 가지고 있는지 여부를 판정하는 것, 특정 질병 또는 질환에 걸린 한 객체의 예후(prognosis)(예컨대, 유방암 상태의 동정, 유방암의 병기(단계) 결정 또는 치료에 대한 유방암의 반응성 결정)를 판정하는 것, 또는 테라메트릭스(therametrics)(예컨대, 치료 효능에 대한 정보를 제공하기 위하여 객체의 상태를 모니터링 하는 것)를 포함한다.

[0071] 특히, 본 명세서에 있어서, “예후”는 병의 발생, 진행, 회복, 재발, 및 약물 내성 등에 관한 예측을 의미하는 것으로, 전망 내지는 예비적 평가를 말한다. 본 발명에서 예후는 유방암의 발생, 치료 후 재발, 유방암으로 인한 사망, 유방암으로부터의 생존(overall survival), 또는 무병생존(disease-free survival)을 의미할 수 있으나, 이에 한정되지 않는다.

[0072] 본 명세서에 있어서, “생물학적 시료”는 유방암을 진단하거나 예후를 예측하고자 하는 피검체로부터 채취된 것이라면 제한 없이 사용할 수 있으며, 예를 들어 생검 등으로 얻어진 세포나 조직, 혈액, 전혈, 혈청, 혈장, 타액, 뇌척수액, 각종 분비물, 소변, 대변 등일 수 있다. 바람직하게는 혈액, 혈장, 혈청, 타액, 비액, 객담, 복수, 질 분비물 및 소변으로 이루어진 군에서 선택될 수 있으며, 바람직하게는 혈액, 혈장, 또는 혈청일 수 있다. 가장 바람직하게는 혈청일 수 있다. 혈액과 같은 액체 생검(liquid biopsy)을 이용하면, 조직검사와 같은 침습적인 방법과 달리 검사자에게 고통을 주지 않으면서 질병의 진단이 가능하다. 또한, 액체 생검을 이용하는 경우 암이 발병하지 않은 잠재적 환자를 대상으로 조기 진단을 실시하거나, 이미 발병한 환자의 치료 경과를 주기적으로 관찰하는데 있어 조직검사에 비하여 큰 이점을 가진다. 상기 시료는 검출 또는 진단에 사용하기 전에 전처리할 수 있다. 예를 들어, 균질화(homogenization), 여과, 증류, 추출, 농축, 방해 성분의 불활성화, 시약

의 첨가 등을 포함할 수 있다.

[0073] 본 명세서에 있어서, “*m/z*”는 질량 대 전하비(mass to charge ratio)를 의미하며, MALDI-TOF 질량분석기를 이용해 측정할 수 있다.

[0074] 본 발명에 있어서, “MALDI-TOF 질량분석”은 매트릭스-보조 레이저 탈착 이온화 질량분석기(Matrix-assisted laser desorption/ionization time-of-flight mass spectrometry)를 이용한 질량분석 방법을 의미한다. 본 명세서에서 “MALDI-TOF 질량분석”은 “MALDI 질량 분석”, “MALDI-TOF MS” 등의 용어와 상호교환적으로 사용될 수 있다.

[0075] 본 발명에 있어서 MALDI-TOF 질량 분석은 양이온 반사형 모드(positive-ion reflection mode)로 수행되는 것일 수 있다. 즉, 본 명세서 전반에 기재된 *m/z* 값은 MALDI-TOF 질량 분석을 양이온 모드로 수행하여 얻은 양이온 모드의 *m/z* 값일 수 있다. 또한, 본 명세서 전반에 기재된 *m/z* 값은 MALDI-TOF 질량 분석을 선형 모드(linear mode) 또는 반사형 모드(reflection mode)로 수행하여 얻은 *m/z* 값일 수 있다.

[0076] 또한, 본 발명에 있어서 상기 유방암 진단을 위한 정보제공방법은 상기 *m/z* 값에 대한 피크 강도를 정규화하는 단계를 더 포함하는 것일 수 있다. 구체적으로는, 상기 정규화는 *m/z* 값에 대한 절대 피크 강도(absolute peak intensity)를 스펙트럼 내 전체 절대피크 강도의 총합으로 나눔으로써 이루어질 수 있다.

[0077] 달리 명시되지 않는 한, 본 명세서에서 분자량 또는 *m/z* 등을 표현하는 모든 숫자, 값, 및/또는 표현은 이러한 숫자들이 본질적으로 다른 것들 중에서 이러한 값을 얻는 데 발생하는 측정의 다양한 불확실성이 반영된 근사치들이므로, 모든 경우 "약"이라는 용어에 의해 수식되는 것으로 이해되어야 한다. “약”이라는 용어가 가치, 양 또는 범위로 명시적으로 나타나지는 않더라도 그러하다. 따라서, 달리 명시되지 않는 한, 본 명세서 전반에 걸쳐서 제시된 수치 또는 값은 정확하지 않고 정확할 필요가 없으며, 근사치 및/또는 원하는 대로 더 크거나 더 작을 수 있고, 이는 허용 오차, 변환 계수, 반올림, 측정 오차 및 기타 등등 및 본 발명이 목적을 달성하기 위해 요구되는 특성에 따라 통상의 기술자에게 알려진 인자들을 반영한다. 예를 들어, 본 명세서 전반에 걸쳐서 제시된 *m/z* 값은 특정된 값으로부터 ±5의 범위에 해당하는 값을, 바람직하게는 ±2의 범위에 해당하는 값을, 더욱 바람직하게는 ±1의 범위에 해당하는 값을 포함한다.

[0078] 본 명세서에 있어서 수치범위가 개시되는 경우, 이러한 범위는 연속적인 것이며, 달리 지적되지 않는 한 이러한 범 위의 최소값으로부터 최대값이 포함된 상기 최대값까지의 모든 값을 포함한다. 더 나아가, 이러한 범위가 정수를 지칭하는 경우, 달리 지적되지 않는 한 최소값으로부터 최대값이 포함된 상기 최대값까지를 포함하는 모든 정수가 포함된다.

[0079] 본 발명에 있어서 *m/z* 값에 대한 피크 강도가 대조군에 비해 높다는 것은, MALDI-TOF 질량분석으로 측정한 특정 *m/z*에 대한 피크 강도가 대조군의 그것에 비교하여 적어도 1%, 2%, 3%, 4%, 5%, 10% 또는 그 이상, 예를 들어, 5%, 10%, 20%, 30%, 40%, 또는 50%, 60%, 70%, 80%, 90% 또는 그 이상 높은, 및/또는 0.5배, 1.1배, 1.2배, 1.4배, 1.6배, 1.8배 또는 그 이상 높은 것을 의미한다.

[0080] 반대로 *m/z* 값에 대한 피크 강도가 대조군에 비해 낮다는 것은, MALDI-TOF 질량분석으로 측정한 특정 *m/z*에 대한 피크 강도가 대조군의 그것에 비교하여 적어도 1%, 2%, 3%, 4%, 5%, 10% 또는 그 이상, 예를 들어, 5%, 10%, 20%, 30%, 40%, 또는 50%, 60%, 70%, 80%, 90% 또는 그 이하로 낮은, 및/또는 0.9배, 0.8배, 0.6배, 0.4배, 0.2배, 0.1배 이하로 낮은 것을 의미한다.

[0081] 본 발명에 따른 유방암 진단을 위한 정보제공방법은, 유방암 병기를 판정하는 단계를 더 포함하는 것일 수 있다. 즉, 본 발명은 유방암 병기 판정을 위한 정보제공방법을 제공한다.

[0082] 유방암 병기는 종양 조직의 크기, 주변 조직으로의 성장 여부, 암의 림프절 침윤 여부, 및 신체 내 다른 장기로의 전이 여부 등을 고려하여 분류되며, 이와 같은 유방암 병기 분류 기준은 본 발명이 속하는 기술분야에 잘 알려져 있다. 유방암 병기는 일반적으로 0 내지 4기로 표현되며, 0기에 가까울수록 비침윤성 암에, 4기에 가까울수록 유방으로부터 신체의 다른 부위로 퍼진 침윤성 암에 해당한다. 유방암 병기에 따른 상태는 구체적으로 설명하면 다음과 같다: 0기-암세포가 상피 내에 국한되어 있는 상태인 비침윤성 유방암; 1기-종양 크기가 2cm 미만이며 암세포가 상피뿐만 아니라 주위 유방조직에 침범한 상태이나, 림프절 등으로는 전이가 없는 상태; 2기-종양 크기가 2cm 미만 또는 2 내지 5cm이면서 림프절 전이가 심하지 않은 상태, 또는 림프절 전이는 없으나 종양 크기가 5cm이상인 경우; 3기-종양 크기가 5cm 미만이지만 림프절 전이가 심하게 일어난 경우, 종양 크기가 5cm 이상이면서 림프절 전이가 일어난 경우, 또는 암이 흉벽이나 피부를 침범하여 국소적으로 진행되거나, 가슴

속 내유 림프절군으로 전이가 일어난 경우; 4기-목의 림프선, 뼈, 폐, 간 등 전신으로 전이가 일어난 경우.

[0083] 따라서, 본 발명은 유방암 병기 판정을 위한 정보를 제공하기 위해, 상기 유방암 진단을 위한 정보제공방법에 더하여 1136 또는 1339 *m/z* 값에 대한 피크 강도가 대조군에 비해 피검체에서 높은 경우 유방암 1기로 판정하는 단계; 또는 1606, 1768, 1444, 1460, 및 1622로 이루어진 군에서 선택된 어느 하나 이상의 *m/z* 값에 대한 피크 강도가 대조군에 비해 피검체에서 높은 경우 유방암 2 내지 4기로 판정하는 단계를 더 포함할 수 있다.

[0084] 보다 구체적으로, 상기 *m/z* 값은 차례로, 1136.401, 1339.467 (유방암 1기 마커); 1606.558, 1768.610, 1444.499, 1460.495, 및 1622.550 (유방암 2 내지 4기 마커)으로 나타낼 수 있다.

[0085] 본 발명에 따른 유방암 진단을 위한 정보제공방법은, 유방암 서브타입을 판정하는 단계를 더 포함하는 것일 수 있다. 즉, 본 발명은 유방암 서브타입 판정을 위한 정보제공방법을 제공한다.

[0086] 호르몬 수용체(hormone receptors, HR)는 유방암의 서브타입을 분류하는 주요 기준이자, 유방암 치료 방법의 결정 요인이다. 종양조직이 에스트로겐 수용체(estrogen receptor, ER) 또는 프로게스테론 수용체(progesterone receptor, PR) 양성이라면, 호르몬 요법(hormonal therapy)이 선호된다. HR 및 사람 표피성장인자 수용체-2(human epidermal growth factor receptor-2, HER2)의 발현 여부에 따라 유방암은 네 개의 주요 서브타입으로 나뉠 수 있다: luminal 1 (HR+/HER2-), luminal 2 (HR+/HER2+), non-luminal HER2+ (HR-/HER2+), 및 삼중음성 타입(triple-negative phenotype) (HR-/HER2-).

[0087] 따라서, 본 발명은 유방암 서브타입(예컨대, luminal 1 타입) 판정을 위한 정보를 제공하기 위해, 상기 유방암 진단을 위한 정보제공방법에 더하여 1095, 1136, 1257, 1282, 1298, 1419, 1444, 1460, 1565, 1581, 1631, 1743, 1745, 1751, 1768, 1793, 1905, 1907, 2067, 및 2069 *m/z*로 이루어진 군에서 선택된 어느 하나 이상의 *m/z* 값에 대한 피크 강도가 대조군에 비해 피검체에서 높은 경우 luminal 1 타입인 유방암으로 판정하거나, luminal 1 타입인 유방암에 걸릴 위험이 있는 것으로 판정하는 단계를 더 포함할 수 있다. 보다 구체적으로, 상기 *m/z* 값은 차례로 1095.379, 1136.401, 1257.422, 1282.446, 1298.441, 1419.470, 1444.499, 1460.495, 1565.514, 1581.520, 1631.580, 1743.573, 1745.580, 1751.608, 1768.610, 1793.636, 1905.631, 1907.639, 2067.698, 및 2069.711로 나타낼 수 있다.

[0088] 또한, 본 발명은 유방암 서브타입(예컨대, luminal 2 타입) 판정을 위한 정보를 제공하기 위해, 상기 유방암 진단을 위한 정보제공방법에 더하여 1095, 1136, 1257, 1298, 1339, 1419, 1460, 1542, 1565, 1581, 1606, 1743, 1745, 1751, 1905, 1907, 2067, 및 2069 *m/z*로 이루어진 군에서 선택된 어느 하나 이상의 *m/z* 값에 대한 피크 강도가 대조군에 비해 피검체에서 높은 경우 luminal 2 타입인 유방암으로 판정하거나, luminal 2 타입인 유방암에 걸릴 위험이 있는 것으로 판정하는 단계를 더 포함할 수 있다. 보다 구체적으로, 상기 *m/z* 값은 차례로 1095.379, 1136.401, 1257.422, 1298.441, 1339.467, 1419.470, 1460.495, 1542.538, 1565.514, 1581.520, 1606.558, 1743.573, 1745.580, 1751.608, 1905.631, 1907.639, 2067.698, 및 2069.711로 나타낼 수 있다.

[0089] 또한, 본 발명은 유방암 서브타입(예컨대, non-luminal HER2+ 타입) 판정을 위한 정보를 제공하기 위해, 상기 유방암 진단을 위한 정보제공방법에 더하여 1095, 1257, 1298, 1339, 1419, 1460, 1542, 1565, 1581, 1606, 1622, 1704, 1743, 1745, 1751, 1768, 1905, 1907, 2067, 및 2069 *m/z*로 이루어진 군에서 선택된 어느 하나 이상의 *m/z* 값에 대한 피크 강도가 대조군에 비해 피검체에서 높은 경우 non-luminal HER2+ 타입인 유방암으로 판정하거나, non-luminal HER2+ 타입인 유방암에 걸릴 위험이 있는 것으로 판정하는 단계를 더 포함할 수 있다. 보다 구체적으로, 상기 *m/z* 값은 차례로 1095.379, 1257.422, 1298.441, 1339.467, 1419.470, 1460.495, 1542.538, 1565.514, 1581.520, 1606.558, 1622.550, 1704.585, 1743.573, 1745.580, 1751.608, 1768.610, 1905.631, 1907.639, 2067.698, 및 2069.711로 나타낼 수 있다.

[0090] 또한, 본 발명은 유방암 서브타입(예컨대, 삼중음성(triple-negative) 타입) 판정을 위한 정보를 제공하기 위해, 상기 유방암 진단을 위한 정보제공방법에 더하여 1095, 1136, 1282, 1298, 1339, 1444, 1542, 1565, 1631, 1751, 1980, 2067, 및 2069 *m/z*로 이루어진 군에서 선택된 어느 하나 이상의 *m/z* 값에 대한 피크 강도가 대조군에 비해 피검체에서 높은 경우 삼중음성 타입인 유방암으로 판정하거나, 삼중음성 타입인 유방암에 걸릴 위험이 있는 것으로 판정하는 단계를 더 포함할 수 있다. 보다 구체적으로, 상기 *m/z* 값은 차례로 1095.379, 1136.401, 1282.446, 1298.441, 1339.467, 1444.499, 1542.538, 1565.514, 1631.580, 1751.608, 1980.670, 2067.698, 및 2069.711로 나타낼 수 있다.

[0091] 또한, 본 발명에 따른 유방암 진단을 위한 정보제공방법은, 유방암의 림프절 침윤 여부를 판정하는 단계를 더 포함하는 것일 수 있다. 즉, 본 발명은 유방암의 림프절 침윤 여부 예측을 위한 정보제공방법을 제공한다.

[0092] 구체적으로, 본 발명은 유방암의 림프절 침윤 여부 예측을 위한 정보를 제공하기 위해, 상기 유방암 진단을 위한 정보제공방법에 더하여 1095, 1257, 1298, 1339, 1419, 1444, 1460, 1542, 1565, 1581, 1606, 1622, 1631, 1704, 1743, 1745, 1751, 1768, 1905, 1907, 2067, 및 2069 *m/z*로 이루어진 군에서 선택된 어느 하나 이상의 *m/z* 값에 대한 피크 강도가 대조군에 비해 피검체에서 높은 경우 림프절 침윤이 일어난 것으로 판정하는 단계를 더 포함할 수 있다. 보다 구체적으로, 상기 *m/z* 값은 차례로 1095.379, 1257.422, 1298.441, 1339.467, 1419.470, 1444.499, 1460.495, 1542.538, 1565.514, 1581.520, 1606.558, 1622.550, 1631.580, 1704.585, 1743.573, 1745.580, 1751.608, 1768.610, 1905.631, 1907.639, 2067.698, 및 2069.711로 나타낼 수 있다.

[0093] 본 발명에 있어서, 유방암은 침습성 유방암(invasive ductal carcinoma)일 수 있다.

[0094] 본 발명은 또한 2138, 1663, 및 1419 *m/z*로 이루어진 1군; 및 1850 및 1688 *m/z*로 이루어진 2군에서 선택된 어느 하나 이상의 *m/z* 값에 대한 질량분석을 수행하는 단계를 포함하는, 유방암 진단 및 예후 예측을 위한 정보제공방법으로서,

[0095] (a) 피검체로부터 분리된 생물학적 시료에서 *N*-글라이칸을 분리하여 정제하는 단계;

[0096] (b) 단계 (a)에서 수득한 *N*-글라이칸을 대상으로 MALDI-TOF 질량분석을 수행하여 상기 *m/z* 값에 대한 피크 강도(intensity)를 측정하는 단계;

[0097] (c) 단계 (b)에서 측정한 피검체의 상기 *m/z* 값에 대한 피크 강도를 대조군의 상기 *m/z* 값에 대한 피크 강도와 비교하는 단계; 및

[0098] (d) 1군에 속하는 *m/z* 값에 대한 피검체의 피크 강도가 대조군에 비해 높거나, 2군에 속하는 *m/z* 값에 대한 피검체의 피크 강도가 대조군에 비해 낮은 경우, 유방암으로 판정하거나 또는 유방암 재발 위험이 있는 것으로 판정하는 단계를 포함하는 것인, 유방암 진단 또는 예후 예측을 위한 정보제공방법을 제공한다.

[0099] 보다 구체적으로, 상기 1군에 속하는 *m/z* 값은 차례로 2138.71, 1663.58, 1419.47로, 상기 2군에 속하는 *m/z* 값은 차례로 1850.66, 및 1688.61로 나타낼 수 있다.

[0100] 본 발명은 또한 1565, 1727, 1881, 및 2012 *m/z*로 이루어진 군에서 선택된 어느 하나 이상의 *m/z* 값에 대한 질량분석을 수행하는 단계를 포함하는, 유방암 예후 예측을 위한 정보제공방법으로서,

[0101] (a) 피검체로부터 분리된 생물학적 시료에서 *N*-글라이칸을 분리하여 정제하는 단계;

[0102] (b) 단계 (a)에서 수득한 *N*-글라이칸을 대상으로 MALDI-TOF 질량분석을 수행하여 상기 *m/z* 값에 대한 피크 강도(intensity)를 측정하는 단계;

[0103] (c) 단계 (b)에서 측정한 피검체의 상기 *m/z* 값에 대한 피크 강도를 대조군의 상기 *m/z* 값에 대한 피크 강도와 비교하는 단계; 및

[0104] (d) 상기 *m/z* 값에 대한 피검체의 피크 강도가 대조군에 비해 낮은 경우, 유방암 재발 위험이 있는 것으로 판정하는 단계를 포함하는 것인, 유방암 예후 예측을 위한 정보제공방법을 제공한다.

[0105] 보다 구체적으로, 상기 *m/z* 값은 차례로 1565.53, 1727.58, 1881.66, 및 2012.71로 나타낼 수 있다.

[0106] 본 발명은 또한 2896, 3444, 2550, 3237, 2656, 6888, 11099, 3879, 7757, 2101, 4206, 2373, 3260, 2167, 2057, 및 4207 *m/z*로 이루어진 1군; 및 2447, 5001, 4150, 2724, 4299, 3167, 4280, 2352, 및 2111 *m/z*로 이루어진 2군에서 선택된 어느 하나 이상의 *m/z* 값에 대한 질량분석을 수행하는 단계를 포함하는, 유방암 진단을 위한 정보제공방법으로서,

[0107] (a) 피검체로부터 분리된 생물학적 시료에서 단백질 또는 펩티드를 분리하는 단계;

[0108] (b) 단계 (a)에서 수득한 단백질 또는 펩티드를 대상으로 MALDI-TOF 질량분석을 수행하여 상기 *m/z* 값에 대한 피크 강도(intensity)를 측정하는 단계;

[0109] (c) 단계 (b)에서 측정한 피검체의 상기 *m/z* 값에 대한 피크 강도를 대조군의 상기 *m/z* 값에 대한 피크 강도와 비교하는 단계; 및

[0110] (d) 1군에 속하는 피검체의 *m/z* 값에 대한 피크 강도가 대조군에 비해 높은 경우 또는 2군에 속하는 *m/z* 값에 대한 피검체의 피크 강도가 대조군에 비해 낮은 경우 유방암으로 판정하거나, 또는 유방암에 걸릴 위험이 있는 것으로 판정하는 단계를 포함하는 것인, 유방암 진단을 위한 정보제공방법을 제공한다.

[0111] 보다 구체적으로, 상기 1군에 속하는 *m/z* 값은 차례로 2896.7, 3444.61, 2550.49, 3237.82, 2656.81, 6888.4775, 11099.6275, 3879.59, 7757.48, 2101.45, 4206.75, 2373.93, 3260.37, 2167.95, 2057.2, 및 4207로 나타낼 수 있고, 상기 2군에 속하는 *m/z* 값은 차례로 2447.97, 5001.02, 4150.49, 2724.24, 4299.86, 3167.65, 4280.45, 2352.486, 및 2111.38로 나타낼 수 있다.

[0112] 본 발명에 있어서, 상기 단백질 또는 펩티드는 당화된 것일 수 있다. 즉, 상기 단백질 또는 펩티드는 글라이칸을 포함하는 것일 수 있다.

[0113] 본 발명은 또한 유방암 진단 또는 예후 예측을 위한 MALDI-TOF 질량분석 키트로서, MALDI 플레이트; 이온화제(ionizing agent); 및 상술한 정보제공방법이 기재된 설명서를 포함하는, MALDI-TOF 질량분석 키트를 제공한다.

[0114] 상기 키트는 C18 칼럼(C18 columns)을 더 포함하는 것일 수 있다.

[0115] 상기 MALDI 플레이트는 MALDI-TOF 질량분석기에서 시료를 분석하기 위해 시료를 배열하는 장치를 지칭한다.

[0116] 상기 이온화제는 분석 대상인 시료를 이온화할 수 있는 용액 또는 물질이라면 제한 없이 포함될 수 있다. 구체적으로, 상기 이온화제는 시료를 결정화하여 레이저가 결정화된 시료에 가해졌을 때 시료가 이온화될 수 있게 하는 용액 또는 물질이라면 제한 없이 포함될 수 있다. 예를 들어, 상기 이온화제는 매트릭스일 수 있다. 또한, 매트릭스는 유기화합물일 수 있으나 이에 한정되지 않는다. 예를 들어, 매트릭스는 시나핀산(Sinapinic acid), α-시아노-4-하이드록시-시남산(-Cyano-4-hydroxy-cinnamic acid, HCA), 2,5-디하이드록시 벤조산(2,5-Dihydroxybenzoic acid, DHB), 슈퍼 DHB(Super DHB), 3-하이드록시피콜린산(3-Hydroxypicolinic acid), 또는 HABA일 수 있으나, 이에 한정되지 않는다. 바람직하게는, 매트릭스는 2,5-디하이드록시 벤조산 및 염화나트륨 용액의 혼합액, 또는 α-시아노-4-하이드록시-시남산 및 아세토나이트릴(acetonitrile, CAN)의 혼합액일 수 있다.

[0117] 상기 질량분석 키트는 본 발명에 따른 유방암 진단 또는 예후 예측을 위한 정보제공방법에 사용되는 것을 특징으로 하는 키트일 수 있다. 따라서, 상기 질량분석 키트는 본 발명에 따른 유방암 진단 또는 예후 예측을 위한 정보제공방법이 기재된 설명서를 포함할 수 있다. 이외에도 본 발명의 목적 달성을 가능하게 하는 것으로서, MALDI-TOF 질량분석기에 사용하기 위해 필요한, 당업계에 공지된 키트 구성이 본 발명에 따른 키트에 포함될 수 있다.

[0119] 이하, 본 발명의 이해를 돕기 위하여 바람직한 실시예를 제시한다. 그러나 하기의 실시예는 본 발명을 보다 쉽게 이해하기 위하여 제공되는 것일 뿐, 하기 실시예에 의해 본 발명의 내용이 한정되는 것은 아니다.

[0121] **<실시예 1> 유방암 병기 또는 유방암 서브타입 구별을 위한 *N*-글라이칸 마커의 발굴**

[0122] 유방암 환자의 진행단계(병기), 림프절 침윤 여부, 및 HR/HER2 발현 여부에 따른 바이오마커를 발굴하기 위해, 인간 혈청 *N*-글라이칸에 대한 MALDI-TOF MS를 수행하였다. 특히, 본 실시예에서는 NosQuest의 전매 소프트웨어인 "NosIDsys"로 이용하여, 정규화된 피크 강도를 기준으로 발현 및 풍부도를 분석해 글리칸 마커를 선별하였다.

[0124] **실험 방법**

[0125] **실험예 1-1. 정상인 및 환자의 혈액 채취**

[0126] 본 실험은 서울 아산병원 심사위원회의 승인을 받았다(IRB 승인번호: 2018-1234). 서울 아산병원의 유방암 환자 및 정상인으로부터 혈액 샘플을 채취하였다. 혈청을 수득하기 위해, 혈액을 25℃에서 짧게 배양한 후 25℃에서 10분 동안 1000×g로 원심분리 하였다. 분리된 혈청은 1.5mL 부피의 마이크로원심분리 튜브(microcentrifuge tubes)에 담아 -80℃에 보관하였다. 정량 평가시 표준 샘플로 사용하기 위해 Sigma-Aldrich사에서 인간 혈청을 구입하였다. 환자의 임상적 특징은 표 1에 나타낸 바와 같다.

표 1

[0127]

표시	변수	값	환자 수(%)
병기	병기	1기	113 (44.2)
		2기	102 (39.8)
		3기	33 (12.9)
		4기	8 (3.1)
N	림프절 침윤	음성	158 (61.7)
		양성	98 (38.3)
HR/HER2	Estrogen receptor (ER)	양성	187 (73.0)
		음성	69 (27.0)
	Progesterone receptor (PR)	양성	154 (60.2)
		음성	102 (39.8)
	Tissue HER2 (IHC)	양성	72 (28.3)
		음성	182 (71.7)
		알 수 없음	2

[0129] **실험예 1-2. *N*-글라이칸의 분리**

[0130] 마이크로튜브에서, 30㎕의 혈청을 1mM 다이티오트레이톨(dithiothreitol)을 함유한 동일한 부피의 200 mM NH_4HCO_3과 혼합하였다. 혈청 단백질은 히트블록(heat block)에서 65℃에서 1,500rpm으로 5분 동안 흔들어(shaking) 변성시켰다. 이어서 400 유닛의 peptide-*N*-glycosidase F(PNGase F; New England Biolabs)을 첨가하여 변성된 단백질로부터 *N*-연결된 글라이칸을 분리했다. 히트블록에서 45℃에서 20분 동안 1,500rpm에서 흔들어 PNGase F 효소 반응을 수행하였다. 이어서, 540 ㎕의 HPLC-급 물(HPLC-grade water)과 100㎕의 1% TFA를 연속해서 첨가하였다. 튜브 내벽에 붙어있는 액체를 스핀다운(spin down)하기 위해 3,000rpm에서 짧게 원심분리하였다.

[0132] **실험예 1-3. 글라이칸의 정제**

[0133] PHGase F에 의해 방출된 글라이칸은 96웰 플레이트(96-well plate) 상에 장착된 HyperSep Hypercarb 고체상추출(solid-phase extraction, SPE) 카트리지(Thermo Fisher Scientific)를 사용하여 정제하였다. 카트리지는 30㎛의 구형인 100% 다공성 그래파이트 탄소(porous graphite carbon, PGC) 입자로 채웠고, 1mL의 증류수 및 1mL의 80% 아세토나이트릴(acetonitrile, ACN)/물(v/v)로 세척하였다. 준비된 *N*-글라이칸 용액을 카트리지 상에 로딩하고 1mL의 증류수로 세 번 세척하였다. 글라이칸은 20% ACN/물(v/v)로 용출하고, Genevac EZ-2 플러스 원심 진공 농축기(Genevac EZ-2 plus centrifugal vacuum evaporator)(Genevac)를 사용해 신속하게 건조시켰다. 이어서 MALDI-TOF 질량분광법(MALDI-TOF mass spectroscopy, MALDI-TOF MS)을 수행하기 위해 건조된 글라이칸을 15㎕의 HPLC-급 물에 재현탁시켰다.

[0135] **실험예 1-4. MALDI-TOF MS 분석**

[0136] 신선한 매트릭스 용액은 2,5-디히드록시 벤조산(2,5-dihydroxybenzoic acid, DHB)(ACN 중 20 mg/mL)을 40 mM의 염화나트륨 용액에 75:25(v/v)의 비율로 혼합하여 준비하였다. 글라이칸 용액은 매트릭스 용액과 1:2(v/v)의 비율로 혼합하였고, 2㎕의 혼합물을 STA μ Focus MALDI 타겟 플레이트(24 X 16 c 2,000㎛; ASTA) 상에 스폿팅하였다(spotted). MALDI 플레이트 상에 로딩된 샘플은 진공에서 신속하게 건조시켜 균일한 매트릭스-샘플 공동-결정화(co-crystallization)를 유도했다. 4800 Plus MALDI-TOF/TOF MS(AB SCIEX)를 양이온 반영 모드로 구동하여 질량 스펙트럼을 획득하였으며, 800 내지 3,000 *m/z* 범위를 모니터링하였다. 10 S/N(signal-to-noise) 이상의 글라이칸 질량 피크만을 유효한 것으로 간주했다.

[0138] **실험예 1-5. 데이터 처리 및 정규화**

[0139] 질량 스펙트럼 데이터는 전처리하였으며(pre-processed), 이온 피크 정보는 4000 Series Explorer™(Applied Biosystems); 및 스펙트럼으로부터 전체 피크 정보(질량 중심값(centroid mass values), 신호 대 잡음비(S/N), 높이)를 마이크로소프트 엑셀(Microsoft Excel)로 전달해주는 인하우스 소프트웨어(Nos-Quest)를 이용하여 추출했다. 각 *N*-글라이칸의 절대 피크 강도(absolute peak intensity, API_i)는 하기 공식에 따라 모든 API의 총합으로 정규화하여 상대적 강도(intensity)를 산출했다:

$$NAPI_i = \frac{API}{\sum_{j=1}^{n} API_j}$$

[0140]

[0141] 상기 식에서, $NAPI_i$는 질량 스펙트럼에서 정의된 *N*-글라이칸의 정규화된 절대 피크 강도를 나타내고, n은 질량 스펙트럼 중의 총 피크의 수(총 이온 크로마토그램)을 나타낸다. 각 API는 스펙트럼의 전체 API의 총합(총 이온 전류(total ion current, TIC))으로 나누었다. 계산 및 분석의 편의를 위해 NAPI 값은 1,000으로 곱했다.

[0143] **실험예 1-6. 구상화(visualization) 및 통계분석**

[0144] NosQuest 전매 바이오마커 패널로부터 얻은 *N*-글라이칸에 해당하는 *N*-글라이칸의 정규화된 강도 데이터를 모아 마이크로소프트 엑셀을 이용해 TSV 파일로 변환하였다. 이어서 Perseus™(Max Planck Institute of Biochemistry)를 사용해 분석하였다. ANOVA 분산분석을 수행하였으며, *P*-값을 0.05로 설정하여 *P*-값 절단법(P value truncation method)을 수행했다. Z 스코어 정규화에서, 각 데이터 포인트는 집계 데이터의 평균값에서 뺀 후 총 데이터의 표준편차로 나누었다. 계층적 클러스터링(hierarchical clustering), 주성분 분석(principal component analysis, PCA) 및 플로팅(plotting)을 위해 Perseus™ 1.5.2.6를 사용하였다. 획득한 피크는 먼저 900 내지 3,000 범위의 *m/z* 비율을 사용하여 필터링하였다. 이어서, *m/z*는 239가지의 공지된 인간 글라이칸 참조 리스트를 기반으로 필터링하였다(J Proteome Res. 2016; 15: 3358-3376). 각 *m/z*에 대해 네 개의 반복 데이터 포인트는 4개 값 중 2개 이상이 있는 경우에만 평균을 내었다. 그 외에는, Not-a-Number 또는 0으로 할당하였다. 이어서 *m/z*는 샘플 전체에 걸쳐 기존값의 90% 이상의 컷오프(cut-off) 빈도(frequency)를 사용하여 두 번째로 필터링하였다. *m/z*는 이어서 다중-샘플 ANOVA로, P값 컷오프는 0.05로 하여 분석하였다.

[0145] Nos-Quest 전매 바이오마커 패널의 *N*-글라이칸종에 해당하는 *N*-글라이칸의 정규화된 강도 데이터를 추출하고, CSV 파일로 변환하여 히트맵 분석에서 선별된 모든 클러스터 마커를 필터링, 저장, 및 사용하여 receiver operating characteristic(ROC) 커브를 구성할 수 있도록 하였다.

[0147] **결과**

[0148] **결과 1-1. 정상 대비 IDC 특이적 MALDI-TOF 마커의 발굴**

[0149] 하기 샘플군에 대해 MALDI-TOF MS를 이용하여 혈청 *N*-글라이칸을 프로파일링 하였다: 정상인(n=311) 및 침습성 유방암(invasive ductal carcinoma, IDC; n=256), 여기서 IDC군은 유방암(BrC) 1기(BrC1; n=113), 2기(BrC2; n=102), 3기(BrC3; n=33), 및 4기(BrC4; n=8) 환자를 포함함. 이들의 평균 나이는 각각 50.88세, 51.93세, 48.00세, 50.24세, 및 54.75세였으며, 모두 여성이었다. 질량 분석 후 *N*-글라이칸 피크 강도는 TIC로 정규화하여 상대적 강도를 산출했다.

[0150] 히트맵 분석 결과 NosQuest 전매 *N*-글라이칸 바이오마커 패널 중 30가지의 *N*-글라이칸종이 정상인과 비교하여 IDC 환자의 혈청에서 유의미한 발현 차이를 나타냈다(도 1). 상기 30 가지 *N*-글라이칸 *m/z* 값은 다음과 같다: 1825.62, 1704.59, 1542.54, 1501.51, 1339.47, 1850.66, 1751.61, 1565.51, 1282.45, 1971.70, 1793.64, 1792.63, 1631.58, 2069.71, 2067.70, 1907.64, 1905.63, 1136.40, 1622.55, 1768.61, 1444.50, 1606.56, 1460.49, 1298.44, 1745.58, 1743.57, 1581.52, 1419.47, 1257.42, 및 1095.38 *m/z*.

[0151] 또한, 곡선하 면적(area under the curve, AUC)는 0.93으로 나타난 바(95%의 신뢰구간(confidence interval, CI): 0.917 SOWL 0.947), IDC 환자 및 정상인의 구별을 위한 바이오마커의 진단 효능이 매우 높다는 것을 알 수

있었다(도 2).

[0152] 이어서, NosQuest 스크리닝에서 더 높은 정확도를 보이고, 정규화된 강도를기준으로 비교했을 때 정상인 및 IDC 환자 사이에서 10% 이상의 풍부도(abundance) 차이를 보이는 바이오마커 20개를 선별하였다. 이들 "NosID 바이오마커"는 정상인에 비해 IDC 환자에서 유의미하게 높은 발현을 보였다(도 3). 상기 20개의 *N*-글라이칸의 *m/z*값은 다음과 같다: 1095.38, 1136.40, 1257.42, 1282.45, 1298.44, 1419.47, 1444.50, 1542.54, 1565.51, 1581.52, 1622.55, 1631.58, 1743.57, 1745.58, 1751.61, 1768.61, 1905.63, 1907.64, 2067.70, 및 2069.71.

[0153] 유방암 환자 및 정상인에서 유의미한 MALDI-TOF 강도 차이를 보이는 *N*-글라이칸들 및 이에 관한 주석을 표 2에 정리하였으며, 각 유방암 병기 및 서브타입에 대해 특이적인 *N*-글라이칸 질량(*m/z*)값을 표시하였다.

표 2

Mass (M+Na)	1095.379	1136.401	1282.422	1298.441	1339.467	1542.538	1565.514	1631.580	1745.580	1751.608	1606.558	1768.610	1793.636
BrC Stage	Stage 1 Stage 2–4	Stage 1	Stage 1 Stage 2–4	Stage 1 Stage 2–4	Stage 1	Stage 1 Stage 2–4	Stage 1 Stage 2–4	Stage 1 Stage 2–4	Stage 1 Stage 2–4	Stage 1 Stage 2–4	Stage 2–4	Stage 2–4	
BrC Subtype	HR +/HER2+ HR +/HER2- HR-/ HER2+ HR-/ HER2- N(+), N (-)	HR +/HER2+ HR +/HER2- HR-/ HER2- N(-)	HR +/HER2- HR-/ HER2- N(-)	HR +/HER2+ HR +/HER2- HR-/ HER2+ HR-/ HER2- N(+), N (-)	HR +/HER2+ HR-/ HER2+ HR-/ HER2- N(+), N (-)	HR +/HER2+ HR-/ HER2+ HR-/ HER2+ HR-/ HER2- N(+), N (-)	HR +/HER2+ HR +/HER2- HR-/ HER2+ HR-/ HER2- N(+), N (-)	HR +/HER2+ HR-/ HER2- HR-/ HER2- N(+), N (-)	HR +/HER2+ HR +/HER2- HR-/ HER2+ N(+), N (-)	HR +/HER2+ HR +/HER2- HR-/ HER2+ HR-/ HER2- N(+), N (-)	HR +/HER2+ HR-/ HER2+ HR-/ HER2+ N(+)	HR +/HER2- HR-/ HER2+ N(+), N (-)	HR +/HER2-
Mass (M+Na)	1907.639	2067.698	2069.711	1257.422	1419.470	1581.520	1743.573	1905.631	1444.499	1460.495	1622.550	1704.585	
BrC Stage	Stage 1 Stage 2–4	Stage 1 Stage 2–4	Stage 1 Stage 2–4	Stage 1 Stage 2–4	Stage 1 Stage 2–4	Stage 1 Stage 2–4	Stage 1 Stage 2–4	Stage 1 Stage 2–4	Stage 2–4	Stage 2–4	Stage 2–4		
BrC Subtype	HR +/HER2+ HR +/HER2- HR-/ HER2+ HR-/ HER2+ N(+), N (-)	HR +/HER2+ HR +/HER2- HR-/ HER2+ HR-/ HER2- N(+), N (-)	HR +/HER2+ HR +/HER2- HR-/ HER2+ HR-/ HER2- N(+), N (-)	HR +/HER2+ HR +/HER2- HR-/ HER2+ N(+), N (-)	HR +/HER2+ HR +/HER2- HR-/ HER2+ N(+), N (-)	HR +/HER2+ HR +/HER2- HR-/ HER2+ N(+), N (-)	HR +/HER2+ HR +/HER2- HR-/ HER2+ N(+), N (-)	HR +/HER2+ HR +/HER2- HR-/ HER2+ N(+), N (-)	HR +/HER2- HR-/ HER2- N(+)	HR +/HER2+ HR +/HER2- HR-/ HER2+ N(+), N (-)	HR-/ HER2+ N(+)	HR-/ HER2+ N(+)	

[0154]

[0156] 구체적인 *N*-글라이칸 리스트는 다음과 같다: 1095.379, 1136.401, 1282.422, 1298.441, 1339.467, 1542.538, 1565.514, 1631.580, 1745.580, 1751.608, 1606.558, 1768.610, 1793.636, 1907.639, 2067.698, 2069.711, 1257.422, 1419.470, 1581.520, 1743.573, 1905.631, 1444.499, 1460.495, 1622.550, 및 1704.585. 여기서, *N*-글라이칸 1136.401 및 1339.467 *m/z* (M+Na)은 유방암 1기 마커에 해당하는 반면, 1606.558, 1768.610, 1444.499, 1460.495, 및 1622.550 *m/z* (M+Na)는 2 내지 4기에서만 우세하게 나타났다. 즉, 복합/하이브리드형 글라이칸인 *m/z* 1444.499 글라이칸과 네 가지 하이브리드형 글라이칸인 1460.495, 1606.558, 1622.550, 및 1768.610 *m/z* 가 유방암 2 내지 4기 특이적 마커임을 알 수 있었다.

[0157] 정상인 및 유방암 1기 환자 샘플을 비교하여 NosIDsys 분석을 수행한 결과, 특이도, 민감도, 및 정확도는 각각 82.3%, 84.1%, 및 82.8%로 나타났으며, 정상인 대 2 내지 4기 환자 샘플의 경우 각각 73.5%, 75.2%, 및 74.0%로 나타났다. 또한, PCA 결과 정상인 샘플 및 모든 병기의 유방암 환자 샘플이 서로 분리된 클러스터(cluster)를 형성한 바, 두 그룹의 *N*-글라이칸 패턴에 명확한 차이가 있는 것을 알 수 있었다. PCA는 정상인 대 1기 유방암 환자(도 4a); 및 정상인 대 2 내지 4기 환자(도 4b)에 대해 각각 수행하였다. 두 경우 모두, 정상인과 유방암 환자(1기 또는 2 내지 4기) 샘플군이 각각 분리된 클러스터를 형성하였다. 즉, 이로부터 유방암 병기를 불문하고 정상인과 유방암 샘플이 구별된 *N*-글라이칸 패턴을 갖는다는 것을 알 수 있었다. 이어서 ROC 곡선을 플롯팅하였으며(plotted), AUC를 계산하였다. 그 결과, 1기 유방암 분석의 AUC는 0.955로(도 5a), 2 내지 4기 분석의 AUC는 0.889(도 5b)로 나타난 바, 1기 유방암에 대한 분석 정확도가 더 높은 것을 알 수 있었다.

[0158]

[0159] **결과 1-2. 정상 대비 유방암 서브타입 특이적 MALDI-TOF 마커의 발굴**

[0160] 유방암 서브타입을 구별하기 위한 *N*-글라이칸 마커를 발굴하고자, HR, HER2, 및 N 병기(N stage)를 고려하였다. 정상인과 유방암의 네 가지 HR/HER2 서브타입을 구별할 수 있는 25가지의 NosID 글라이칸 바이오마커들이 확인되었다(표 2). 바이오마커들의 강도(풍부도)는 정상인 그룹에 비해 각각의 유방암 서브타입 그룹(HR+/HER2-, HR+/HER2+, HR-/HER2+, 및 HR-/HER2-)에서 유의미하게 높았다(도 6a 내지 6d). 특히, *m/z* 1793.636(M+Na)는 정상 그룹과 HR+/HER2- 그룹을 구별할 수 있는 것으로 확인되었다.

[0161] 또한, 표 2의 24가지 NosID 글라이칸 바이오마커가 정상 그룹과 림프절 침윤(lymph node invasion)이 일어난 유방암 그룹(N(+))을 구별할 수 있음을 확인하였다. 즉, N(+) 바이오마커가 정상인에 비해 림프절 침윤 유방암 그룹에서 유의미하게 높은 강도를 나타내는 것을 확인하였다(도 7).

[0162] 결과적으로, 유방암 병기 또는 서브타입에 특이적인 25가지의 *N*-글라이칸 바이오마커 후보가 선별되었다. 이 중 1622.550, 1704.585, 및 1793.636 *m/z*의 글라이칸은 표 2에 나타낸 바와 같이 MALDI-TOF 강도에 있어서 유방암 환자와 정상인간에 유의미한 차이를 나타냈다. 즉, 하이브리드형 글라이칸인 1622.550 *m/z* 및 복합/하이브리드형 글라이칸인 1704.585 *m/z* 가 HR-/HER2+ 서브타입 특이적 마커이며, 복합/하이브리드형 글라이칸인 1793.636 *m/z* 이 HR+/HER2- 서브타입 특이적 마커임을 알 수 있었다.

[0164] **<실시예 2> 유방암 진단 또는 유방암 예후 예측을 위한 *N*-글라이칸 마커의 발굴**

[0165] 유방암 진단 및 유방암의 재발 여부를 예측하기 위한 바이오마커를 발굴하기 위해, 인간 혈청 *N*-글라이칸의 MALDI-TOF MS를 수행하였다. 특히, 본 실시예에서는 IDsys RT를 이용하여 분석하였다.

[0167] **실험예 2-1. 정상인 및 환자의 혈액 채취**

[0168] 실험예 1-1과 동일한 방법으로 정상인 및 유방암 환자로부터 혈액을 채취한 후 혈청을 분리하였다. 임상적 특성은 하기 표 3에 정리했다. '모니터링' 그룹은 혈액을 세 번에 걸쳐 채취했다: 유방암 치료 전, 치료 중, 치료 후.

표 3

[0169]

분류	단계	재발 타입	샘플 수	평균 나이	총 수
정상 대조군	-	-	22	53.7	22

유방암	0	-	5	51.5	22
	1	-	5		
	2	-	4		
	3	-	6		
	4	-	2		
재발	-	국한(local)	6	47.5	22
		국소(regional)	5		
		원격(distant)	8		
		국한 및 국소	1		
		국소 및 원격	2		
모니터링					
치료 후 질병증거 없음(no evidence of disease, NED)	0	-	6	43.3	33
	1	-	6		
	2	-	12		
	3	-	9		
치료 후 재발	0	국한	3	44.9	33
	1	국한	3		
		국소	3		
		원격	3		
		국한 및 원격	3		
	2	국소	3		
		원격	6		
	3	원격	9		

[0171] **실험예 2-2. *N*-글라이칸의 분리**

[0172] 실험에 1-2 및 1-3에 나타낸 바와 같이 PNGase로 *N*-글라이칸을 분리하고 PGC로 채워진 칼럼 칼럼으로 글라이칸을 정제하였다. 글라이칸은 20% ACN/물(v/v)로 용출하였다.

[0174] **실험예 2-3. MALDI-TOF MS 분석**

[0175] 실험예 1-4와 동일한 방법으로 MALDI-TOF MS 분석을 위한 샘플 및 매트릭스를 준비하였다. 질량 스펙트럼(mass spectra)은 IDsys RT MALDI-TOF MS(ASTA)를 이용하여 획득했다. 양이온 반사형 모드(positive ion reflection mode)를 사용하여 900 내지 3000 달톤의 질량 범위를 분석하였으며, 2 S/N(signal 대 noise) 비율 이상의 *N*-글라이칸 질량 피크만이 유효한 피크로 간주하였다. IDsys RT 빌트-인 후처리 단계(IDsys RT built-in post-processing steps)를 사용하여 13가지의 '공지된-글라이칸' 질량(M+Na 1257, 1419, 1485, 1501, 1542, 1581, 1647, 1688, 1704, 1743, 1809, 1850, 및 1905 *m/z*, 소수는 나타내지 않은 것임)에 관하여 모든 피분석물에 대해 질량 교정(mass calibration)을 수행하였다.

[0177] **실험예 2-4. 데이터 처리**

[0178] 질량 스펙트럼으로부터의 피크 정보는 IDsys 3.0 (ASTA) 소프트웨어를 사용해 추출하였다. 획득한 모든 피크 중에서, 156가지의 인간 글라이칸으로 이루어진 참조 *N*-글라이칸 목록에 대응하는 *m/z*의 강도를 필터링하여 사용했다. 사용한 참조 *N*-글라이칸 목록은 표 4에 나타냈다. 실험예 1-5와 동일하게, 각 *N*-글라이칸의 절대 피크 강도(API_i)는 총 API_i의 총합으로 정규화하여, 정규화된 절대 피크 강도(normalized absolute peak intensity, $NAPI_i$)를 산출하였다.

표 4

[0179]

No.	*N*-글라이칸 질량 (M, Da)	단위 동일 원소 질량(Monoisotopic mass) (M+Na, Da)
1	910.328011	933.31778
2	1056.385919	1079.375688

3	1072.380834	1095.370603
4	1113.407383	1136.397152
5	1218.438743	1241.428512
6	1234.433657	1257.423426
7	1259.465292	1282.455061
8	1275.460207	1298.449976
9	1316.486756	1339.476525
10	1380.491566	1403.481335
11	1396.486481	1419.47625
12	1405.523201	1428.51297
13	1421.518115	1444.507884
14	1437.51303	1460.502799
15	1478.539579	1501.529348
16	1519.566128	1542.555897
17	1542.54439	1565.534159
18	1558.539304	1581.529073
19	1566.555623	1589.545392
20	1567.576024	1590.565793
21	1583.570939	1606.560708
22	1588.537568	1611.527337
23	1599.565853	1622.555622
24	1608.602573	1631.592342
25	1640.592402	1663.582171
26	1665.624037	1688.613806
27	1681.618952	1704.608721
28	1704.597213	1727.586982
29	1712.613532	1735.603301
30	1720.592128	1743.581897
31	1722.645501	1745.63527
32	1728.608446	1751.598215
33	1729.628848	1752.618617
34	1745.623762	1768.613531
35	1750.590392	1773.580161
36	1754.660482	1777.650251
37	1769.634996	1792.624765
38	1770.655397	1793.645166
39	1802.645226	1825.634995
40	1811.681946	1834.671715
41	1827.67686	1850.666629
42	1843.671775	1866.661544
43	1858.671441	1881.66121
44	1866.650037	1889.639806
45	1868.703409	1891.693178
46	1874.666355	1897.656124
47	1882.644951	1905.63472
48	1884.698324	1907.688093
49	1890.66127	1913.651039
50	1891.681671	1914.67144
51	1912.643215	1935.632984
52	1915.692904	1938.682673
53	1916.713305	1939.703074
54	1925.724873	1948.714642
55	1931.687819	1954.677588
56	1932.70822	1955.697989
57	1937.67485	1960.664619
58	1948.703135	1971.692904
59	1953.669764	1976.659533
60	1957.739855	1980.729624

61	1964.698049	1987.687818
62	1972.714368	1995.704137
63	1973.734769	1996.724538
64	1989.729684	2012.719453
65	2005.724598	2028.714367
66	2014.761318	2037.751087
67	2020.724264	2043.714033
68	2028.70286	2051.692629
69	2030.756233	2053.746002
70	2036.719179	2059.708948
71	2044.697775	2067.687544
72	2046.751148	2069.740917
73	2061.750813	2084.740582
74	2071.782782	2094.772551
75	2077.745728	2100.735497
76	2078.766129	2101.755898
77	2083.732758	2106.722527
78	2087.777697	2110.767466
79	2093.740642	2116.730411
80	2094.761044	2117.750813
81	2099.727673	2122.717442
82	2103.797763	2126.787532
83	2110.755958	2133.745727
84	2115.722588	2138.712357
85	2118.772277	2141.762046
86	2119.792678	2142.782447
87	2134.767192	2157.756961
88	2135.787593	2158.777362
89	2140.754222	2163.743991
90	2151.782507	2174.772276
91	2156.749137	2179.738906
92	2160.819227	2183.808996
93	2167.777422	2190.767191
94	2175.793741	2198.78351
95	2176.814142	2199.803911
96	2182.777088	2205.766857
97	2190.755683	2213.745452
98	2192.809056	2215.798825
99	2204.759033	2227.748802
100	2206.750598	2229.740367
101	2207.808722	2230.798491
102	2208.803971	2231.79374
103	2217.840691	2240.83046
104	2222.783236	2245.773005
105	2223.803637	2246.793406
106	2233.835605	2256.825374
107	2239.798551	2262.78832
108	2240.818952	2263.808721
109	2244.765181	2267.75495
110	2245.785582	2268.775351
111	2249.83052	2272.820289
112	2255.793466	2278.783235
113	2256.813867	2279.803636
114	2264.830186	2287.819955
115	2265.850587	2288.840356
116	2280.8251	2303.814869
117	2281.845501	2304.83527
118	2296.820015	2319.809784

119	2297.840416	2320.830185
120	2302.807045	2325.796814
121	2306.877136	2329.866905
122	2313.835331	2336.8251
123	2318.80196	2341.791729
124	2321.851649	2344.841418
125	2322.872051	2345.86182
126	2337.846564	2360.836333
127	2338.866965	2361.856734
128	2352.808507	2375.798276
129	2354.86188	2377.851649
130	2363.8986	2386.888369
131	2368.803421	2391.79319
132	2368.841144	2391.830913
133	2369.861545	2392.851314
134	2370.856794	2393.846563
135	2378.873113	2401.862882
136	2379.893514	2402.883283
137	2384.836059	2407.825828
138	2385.85646	2408.846229
139	2390.823089	2413.812858
140	2395.888429	2418.878198
141	2401.851375	2424.841144
142	2402.871776	2425.861545
143	2410.888095	2433.877864
144	2411.883344	2434.873113
145	2425.862608	2448.852377
146	2426.883009	2449.872778
147	2427.90341	2450.893179
148	2442.877924	2465.867693
149	2443.898325	2466.888094
150	2448.864954	2471.854723
151	2452.935045	2475.924814
152	2458.872838	2481.862607
153	2459.893239	2482.883008
154	2464.859869	2487.849638
155	2467.909558	2490.899327
156	2468.929959	2491.919728

[0181] **실험예 2-5. 구상화 및 통계분석**

[0182] 실시예 1-6과 동일하게 Perseus™를 이용해 데이터를 분석하였다. ANOVA 분산분석을 수행하였으며 *P*-값의 컷오프는 0.05로 하여 필터링하였다. 계층적 클러스터링 및 구상화를 위해 모든 데이터를 Z-스코어로 정규화하였다. 공공 웹-기반 생물학적 데이터 분석 패키지인 MetaboAnalystR의 바이오마커 분석툴을 사용해 ROC 분석 및 AUC 계산을 수행했다. 분류 방법으로는 선형 SVM을 사용했다.

[0184] **결과**

[0185] **결과 2-1. 유방암 환자 특이적 *N*-글라이칸 마커의 발굴**

[0186] 22명의 정상인 샘플과 동일한 수의 유방암 환자 샘플로부터 *N*-글라이칸을 추출 및 정제하여 MALDI-TOF 분석을 수행했다. 유방암 환자 샘플에는 유방암 각 병기 별로(0 내지 4) 2개 이상의 샘플이 포함되게 하였다. 또한, 환자 나이에 따른 편차를 최소화하기 위해, 정상인 및 유방암 환자의 평균 나이가 53.7세 및 51.5세가 되도록 조정하였다.

[0187] 도 8에 나타낸 바와 같이, 5개의 *N*-글라이칸 마커(1419, 1663, 1688, 1850, 및 2138 *m/z*)가 정상인과 유방암 환자 샘플간에 유의미한 차이를 보이는 것으로 나타났다. 히트맵 분석 결과, 이 중 세 가지 *N*-글라이칸 마커(*m/z* 1419, 1663, 및 2138)는 양성 마커로서 유방암 샘플에서 더 높은 강도를 가졌으며, 나머지 두 가지 *N*-글라이칸 마커(*m/z* 1688 및 1850)은 음성 마커로서 유방암 샘플에서 강도가 더 낮았다. PCA 분석 결과, 두 샘플 그룹은 분리된 클러스터로 분류되었고(도 9), 최대 AUC 값은 0.91이었다(도 10). 또한, 서포트 벡터 머신(support vector machine, SVM)으로 기계학습된 모델은 88.6%의 정확도로 정상인 샘플과 유방암 샘플을 구별하였다(도 11). 또한, 선별된 마커에 대한 강도 차이는 데이터 빈도(data frequency)의 90%를 초과하는 전체 *N*-글라이칸 강도의 평균을 비교할 때 더욱 명확하게 나타났다(도 12). 선별된 마커(1419, 1663, 1688, 1850, 및 2138 *m/z*)의 로우(raw) 스펙트럼 프로파일 및 이들의 *N*-글라이칸 구조는 도 13에 나타냈다.

[0189] **결과 2-2. 유방암 재발 환자 특이적 *N*-글라이칸 마커의 발굴**

[0190] 유방암 재발 환자 특이적 *N*-글라이칸을 발굴하기 위해, 22명의 유방암 재발 환자의 샘플을 정상인 및 유방암 환자 샘플과 비교분석하였다. 재발 유방암 샘플에는 다양한 재발 유형이 포함되었다: 국한(local), 국소(regional), 원격(distant) 재발 및 이들의 조합. 정상인, 유방암, 및 유방암 재발 환자 샘플의 *N*-글라이칸 프로파일을 도 14에 나타냈다. 총 41개의 *N*-글라이칸 *m/z* 마커가 확인되었으며, 구체적으로는 다음과 같다: 1866.66, 1825.63, 1663.58, 2028.71, 2138.71, 2163.74, 1996.72, 2174.77, 2158.77, 1913.65, 1938.68, 1987.68, 1971.69, 1834.67, 1960.66, 2126.78, 2122.71, 2179.73, 1955.69, 1980.72, 1622.55, 1606.56, 1460.50, 2012.71, 1793.64, 1792.62, 1631.59, 1444.50, 1768.61, 1727.58, 1751.59, 1428.51, 1419.47, 1907.68, 1745.63, 1743.58, 2069.74, 2067.68, 1850.66, 1688.61, 1881.66, 1257.42, 1136.39, 및 1013.31 *m/z*. 결과 2-1에서 확인된 모든 양성 및 음성 *N*-글라이칸 마커(1419, 1663, 1688, 1850, 및 2138 *m/z*) 또한 상기 41가지 마커 중에 포함되었고, 재발 유방암 샘플과 동일한 경향성을 보였다. 이와 같은 경향성은 상기 마커들을 세 그룹(정상, 유방암, 재발)에서 강도를 각각 비교하였을 때 더욱 두드러지게 나타났다(도 15a 내지 15e). 즉, 유방암 양성 마커들은 재발 환자 샘플에서 강도가 더욱 증가된 것으로 나타났으며, 유방암 음성 마커들은 재발 환자 샘플에서 강도가 더욱 감소된 것으로 나타났다. 이와 같은 결과는 상기 유방암 마커들이 유방암 재발 예측을 위한 마커로도 활용될 수 있음을 시사한다.

[0191] PCA 분석 결과, 유방암 재발 환자 샘플은 정상인 및 유방암 환자 샘플과 분리된 클러스터를 형성했다(도 16). 특히, 결과 2-1에서 확인한 바와 같이 정상인과 유방암 환자간에서도 유의미한 차이가 있음에도, 유방암 재발 환자 샘플에서 차이가 더욱 큰 것을 알 수 있었다.

[0193] **결과 2-3. 유방암 치료 반응 모니터링을 위한 *N*-글라이칸 마커 발굴**

[0194] 22명의 유방암 환자로부터 치료 전반에 걸쳐 혈청을 세 번 채취하였다(치료 전, 치료 중, 치료 후. 22명의 환자 중 11명의 환자는 유방암 재발을 진단받았고, 나머지 환자는 재발 증거를 보이지 않았다. 11명의 완치 환자의 경우, 샘플 그룹에 다양한 특성이 통합될 수 있도록, 초기에 특정 유방암 병기(0 내지 3)로 진단된 환자가 적어도 2명씩 포함되게 하였다. 상기 샘플들은 22명의 정상인 샘플과 비교분석하였다. 유방암 환자군을 모니터링하기 위한 중앙 팔로우업(median follow-up, f/u) 시간은 64.5개월이었다. 도 17은 정상인과 완치 환자(NED)(치료 전, 치료 중, 치료 후)의 *N*-글라이칸 강도를 비교한 히트맵 그래프이다. 그래프에 나타낸 바와 같이, 환자 그룹은 정상인과 비교하여 독특한 *N*-글라이칸 프로파일을 나타냈다. 비록 치료 후 유방암이 완치된 것으로 판정받은(NED) 환자라도, 치료 후 혈청 샘플의 *N*-글라이칸 프로파일은 정상인 샘플보다는 치료 전 및 치료 중 혈청 샘플의 프로파일과 유사한 것으로 나타났다. 이와 유사하게, PCA 분석 결과 정상인 그룹은 완전히 분리된 클러스터를 형성한 반면, 치료 전, 치료 중, 및 치료 후의 유방암 환자 샘플의 클러스터간에는 유의미한 차이가 없었다(도 18).

[0196] **결과 2-4. 유방암 재발 예측을 위한 *N*-글라이칸 마커 발굴**

[0197] 유방암 치료에 따른 *N*-글라이칸 패턴 변화를 분석하기 위해, 유방암 환자 샘플보다 정상인 샘플에서 두드러졌던 5가지 *N*-글라이칸 마커(1419, 1663, 1688, 1850, 및 2138 *m/z*)의 치료에 따른 평균 강도의 변화를 모니터링 및 플롯팅하였다(plotted). 또한, 치료 후 유방암이 재발한 11명의 환자 그룹을 완치 환자 그룹과 동일한 방식으로

모니터링 및 비교하였다. 재발 환자 내에는 다양한 유방암 병기 및 다양한 유형의 재발이 진단된 환자들이 포함되었다.

[0198] 유방암 재발 환자와 정상인을 비교하여 히트맵을 작성한 결과, 정상인 대비 완치 환자(NED) 분석 결과와 유사한 결과가 나타났다(도 19). 유방암 재발 예측을 위한 마커의 기준은 다음과 같다: 첫째, 치료 전 샘플은 정상인과 비교하여 상이한(증가 또는 감소된) 강도를 나타내야 한다; 둘째, NED 샘플에서는 정상인의 것과 유사한 경향을 보이나, 재발 환자 샘플에서는 정상인 샘플과 경향이 상이한 프로파일을 나타내야 한다. 5가지 *N*-글라이칸 마커(1419, 1663, 1688, 1850, 및 2138 *m/z*)에 대해 치료 과정에 따른 MALDI TOF 강도를 비교한 결과, 모든 치료 전 샘플이 정상인과 비교했을 때 변화된 *N*-글라이칸 강도를 보였다(도 20a 내지 20e). 완치 환자(NED) 및 유방암 재발 환자 샘플의 강도 패턴은 치료 중간 시점까지는 유사했으나, 치료 후 샘플의 경우 완치 환자만이 정상인과 유사한 수준으로 *N*-글라이칸 강도를 회복하거나 유지하는 것으로 나타났다.

[0199] 유방암 재발 환자와 완치 환자를 구별하기 위한 *N*-글라이칸 마커를 더 발굴하기 위해 추가 분석을 수행한 결과, 네 가지 글라이칸 마커가 추가로 발굴되었다: M+Na, 1565.53, 1727.58, 1881.66, 및 2012.71 *m/z* (도 21).

[0201] **<실시예 3> MALDI-TOF MS 분석을 이용한 유방암 진단용 혈청 단백질 마커 발굴**

[0202] 유방암 진단을 위한 단백질 마커를 추가로 발굴하기 위해, 혈청 단백질을 정제하고 C18 칼럼(C18 columns)으로 농축한 후 MALDI-TOF MS 분석을 수행하였다.

[0204] **실험 방법**

[0205] **실험예 3-1. 샘플의 준비**

[0206] 아산병원으로부터 30명의 정상인 및 30명의 유방암 환자의 혈청 샘플을 확보하였다. 환자 중 가장 어린 연령은 34세였으며, 가장 높은 연령은 64세였다. 유방암 샘플 중, 20명은 2기였고, 10명은 3기였다. 환자 분류표는 표 5에 나타냈다.

표 5

[0207]

특징	유방암 환자	대조군(정상)
수 (명)	30	30
성별		
남성	0	0
여성	30	30
나이		
평균 (SD)	45 (5.8)	52 (7.7)
범위	36-56	34-64
유방암 단계(병기)별 환자 수		
1기	0	
2기	20	
3기	10	
4기	0	

[0208] 단백질을 정제하기 전, 혈청 샘플은 먼저 0.5% TFA(v/v)를 함유한 물로 1:10의 비율로 희석하였다. 이어서 실험예 3-2에 따라 단백질 및 펩티드를 정제하기 위해, 혈청 샘플 중 5㎕을 45㎕의 탈이온수(DW)와 혼합하여 준비하였다.

[0209]

[0210] **실험예 3-2. 단백질/펩티드 정제 및 농축**

[0211] 혈청 단백질 또는 펩티드 분자는 Peptide Cleanup C18 피펫팁(pipette tips)(Agilent Technologies)을 이용하여 제조사의 설명에 따라 정제 및 농축하였다. Cleanup C18 피펫팁을 10-μL 피펫에 끼우고, 팁에 10μL의 ACN 용액을 흡인(aspirating) 및 분배(dispensing)한 후 다시 물에 희석된 0.1% TFA 10μL을 흡인 및 분배하였다. 이어서 희석된 샘플 10μL을 20회 흡인 및 배출(expelling)하여 샘플을 팁에 부착(bound)시켰다. 그 다음 팁으로 0.1% TFA를 함유한 60% ACN 용액 7.5μL을 10회 흡인 및 분배하여 부착된 단백질 및 펩티드를 원심분리 튜브에 용출시켰다.

[0213] **실험예 3-3. MALDI-TOF MS 분석**

[0214] 질량분석(mass spectrometry analysis)은 IDSys LT(ASTA)를 이용하여 수행했다. 용출된 단백질 및 펩티드 샘플 4μL을, 0.1% TFA가 첨가된 70% ACN 용액 중의 20 mg/ml α-시아노-4-하이드록시신남산(α-Cyano-4-hydroxycinnamic acid, CHCA)(Sigma) 매트릭스 4μL과 혼합하였다. 샘플/매트릭스 혼합용액 1.5μL을 96 타겟 MALDI 플레이트(ASTA)에 샘플별로 네 번씩 스폿팅하였다. 각 MALD 스폿(spot)마다, 30가지의 서로 상이한 자리(loci)에 40번씩 레이저를 조사하였으며(스폿당 총 1,200샷(shots)), 2,000 내지 50,000의 질량 범위에 대해 MALDI 분석을 수행했다. MALDI-TOF MS 분석은 선형 양성 모드(linear positive mode)로 수행했다.

[0216] **실험예 3-4. 데이터 처리 및 정규화**

[0217] 로우 질량 스펙트럼(raw mass spectra)는 모든 질량 스펙트럼 및 질량 분석 범위 전반에 걸쳐 나타나는 8가지 주요 *m/z* 피크로 정규화하였으며, 인하우스 프로그램인 NosIDSys를 사용하였다. 정규화된 질량 스펙트럼은 *m/z* 및 절대 강도 값을 추출한 인하우스 프로그램을 사용하여 피크의 수치 리스트로 변환하였으며, 상기 리스트는 XML 파일로 내보냈다. 임상 샘플에 대해 질량 수치 및 강도 값을 수집한 후 XML파일을 CSV 파일로 변환하였다.

[0218] 통계 분석을 수행하기 전, 모든 절대 강도 데이터는 총 이온 개수(total ion count, TIC)로 정규화 하였다(각각의 강도 값을 총 강도 값의 총합으로 나눔, TIC 정규화 강도).

[0220] **실험예 3-5. 통계적 분석**

[0221] TIC 정규화 데이터로 통계 분석을 수행했다. 먼저, 히트맵, PCA 그래프, 및 ROC 커브를 플롯팅하였다. 히트맵 및 PCA 그래프는 프리웨어 프로그램인 Perseus를 사용하여 생성했다. *P*-값은 0.05로 설정하여 ANOVA 분산분석을 수행하였으며, 유의미한 질량값만으로 히트맵 및 PCA 그래프를 생성했다. 또한, Origin이라는 써드파티 프로그램(3^{rd} party program)을 사용하여 정상인의 스펙트럼과 암환자 샘플의 스펙트럼을 비교하는 워터폴 플롯(waterfall plot)을 생성했다. ROC 곡선은 온라인 툴인 Metaboanalyst으로 생성하였다.

[0223] **실험예 3-6. 암 스크리닝 및 진단 분석**

[0224] *m/z* 값 및 대응 TIC 정규화 강도 값(TIC normalized intensity values)을 NosIDsys 소프트웨어에 입력하였다. 상기 프로그램은 TIC 정규화 강도 값과 두 임상 그룹간의 차이를 바탕으로 컷오프 값을 자동으로 계산하였는데, 이 값은 샘플이 정상 샘플로 선별되었는지 또는 유방암 샘플로 선별되었는지 여부를 결정하는 진단 기준 값으로 사용되었다. 두 개의 서로 다른 진단 알고리즘을 사용하였는데, 하나는 P-값 가중 방법(P-value weighted method)를 기반으로 하는 것이고, 다른 하나는 랜덤 포레스트(random forest) 모델링을 기반으로 하는 것이다. NosID 분석 후 특이성, 민감도, 및 정확도 값을 기록했다.

[0226] **결과**

[0227] **결과 3-1. 정상 대비 유방암 환자의 단백질 TIC 정규화 강도(Protein TIC nomrmalized intensity)**

[0228] 관측된 모든 *m/z* 피크의 절대 강도 값은 TIC로 정규화하였다. 도 22에 나타낸바와 같이, 정상인(n=30) 및 유방암 환자(n=30)간 TIC 정규화 강도 차이를 나란히 비교하였다.

[0229] 통계적으로 유의미한 질량 마커 후보를 결정하기 위해, TIC 정규화 강도 값을 히트맵으로 나타냈다. *P*-값 기준

을 0.05로 하여 두 그룹에 대해 ANOVA 테스트를 수행했으며, 통계학적으로 유의미한 질량에 대해 Z-스코어 분석을 수행했다. 총 24개의 통계적으로 유의미한 질량 마커가 검출되었으며, 이들의 히트맵 클러스터링 결과는 도 23에 나타냈다. 상기 24개의 질량 마커는 구체적으로 다음과 같다: 2447.97, 5001.02, 4150.49, 2724.24, 4299.86, 3167.65, 4280.45, 2352.486, 2111.38, 2896.7, 3444.61, 2550.49, 3237.82, 2656.81, 6888.4775, 11099.6275, 3879.59, 7757.48, 2101.45, 4206.75, 2373.93, 3260.37, 2167.95, 및 2057.2 *m/z*. 히트맵에 나타낸 바와 같이, 상기 질량 마커들은 정상인 샘플에 비해 유방암 환자 샘플에서 두드러지게 증가 또는 감소되어 있는 것을 확인할 수 있었다.

[0230] 또한, 워터폴 플롯에서도 도 23에 나타낸 *m/z* 마커들의 정규화된 강도 값이 정상인 샘플과 유방암 환자 샘플간에 가시적인 차이를 보이는 것을 확인했다(도 24). 특히, 4,150 *m/z* 가 정상인과 유방암 환자간에 분명한 차이를 나타냈다.

[0231] 이어서 두 그룹의 TIC 정규화 강도 값으로 PCA 그래프를 작성하여 샘플의 분포를 확인하고, 유사성에 기반하여 클러스터링하였다. 그 결과, 정상인 그룹과 유방암 환자 그룹이 서로 구별되는 클러스터를 갖는 것을 확인하였다(도 25).

[0232] 또한, 상기 마커들을 이용한 유방암 진단의 정확도를 파악하기 위해, 히트맵 분석에서 사용된 것과 동일한 데이터(정상인 및 유방암 환자 순으로 나열된 각 샘플의 *m/z* 값 및 TIC 정규화 강도)를 이용하여 ROC 곡선을 플롯팅하였다(도 26). 가장 높은 수의 마커의 AUC는 0.932로 나타난 바, 본 발명에 따른 바이오마커에 기반한 정상인 및 유방암 환자의 분류 성능이 뛰어나다는 것을 확인할 수 있었다.

[0234] **결과 3-2. NosID를 이용한 진단 결과의 신뢰도 파악**

[0235] 실제 암 스크리닝 또는 진단은 인하우스 프로그램인 NosID™을 이용하여 수행하였다. 상기 프로그램에 *m/z* 값 및 강도 값으로 이루어진 데이터를 입력하였으며, 강도 값은 추가적인 데이터 정규화 없이 TIC 정규화 강도 값으로 전환하여 이용했다. 전술한 바와 같이 두 가지 알고리즘을 기반으로 분석하였으며, 분석된 특이성, 민감도, 및 정확도를 표 6 및 표 7에 정리하였다.

표 6

[0236] *P*-값 가중 방법을 이용한 NosID 분석 결과

샘플	특이성	민감도	정확도
정상 vs 유방암	71.7%	87.5%	79.6%

표 7

[0238] 랜덤 포레스트-기반 알고리즘을 이용한 NosID 분석 결과

샘플	특이성	민감도	정확도
정상 vs 유방암	93.3%	94.9%	94.2%

[0240] 표 6 및 7에 나타낸 바와 같이, 본 실시예에 따른 바이오마커를 이용한 유방암 진단은 높은 특이성, 민감도, 및 정확도를 갖는 것으로 나타났다. 특이성(specifity)은 얼마나 많은 정상 대조군이 진단 분석에 의해 올바르게 식별되었는지를 나타내며, 민감도는 얼마나 많은 유방암 환자군이 진단 분석에 의해 올바르게 식별되었는지를 나타낸다. 정확도는 진단의 전반적인 정확성을 나타낸다. 위와 같은 결과는 본 실시예에 따른 바이오마커에 대해 두 그룹간에 유의미한 강도 차이가 있으며, 이들 차이로부터 도출한 임계 컷오프 값은 유방암 진단 또는 스크리닝에 사용될 수 있음을 보여준다.

[0242] **결과 3-4. TIC 정규화 강도간의 차이를 기반으로 한 *m/z* 값 분석**

[0243] *P*-값 가중 방법 또는 랜덤 포레스트-기반 방법을 이용해 각 *m/z*의 정상인 및 유방암 환자에서의 TIC 정규화 강도 값 간의 차이를 분석하는 것 외에, 간단하게 평균을 취한 후 비교하여 정상인과 유방암 환자 사이에서 차이를 보이는 *m/z*가 있는지 추가 분석하였다. 즉, 각 *m/z*에 대한 정상인 및 유방암 환자의 TIC 정규화 강도 값 사이의 *P*-값을 계산하여 두 그룹간 차이가 있는지 확인하였다. 특히 큰 차이를 보이는 바이오마커를 선별하기 위해, *P*-값이 0.001 이하인 *m/z*를 선별하였으며, 해당 *m/z*의 TIC 정규화 강도 (평균)값을 표 8에 나타냈다. 총 6개의 *m/z* 값이 정상인과 유방암 환자간에 유의미한 차이를 보이는 것으로 나타났다. 표 8에 나타낸 6가지 *m/z* 바이오마커에 대한 정상인 또는 유방암 환자의 MALDI-TOF MS 스펙트럼은 도 27a 내지 27d에 비교하여 나타냈다. 그래프로 확인되는 바와 같이, 6가지 바이오마커에 대해 절대 강도 및 TIC 정규화 강도 값 모두 정상인과 유방암 환자간에 두드러진 패턴 차이를 보이는 것을 알 수 있었다.

표 8

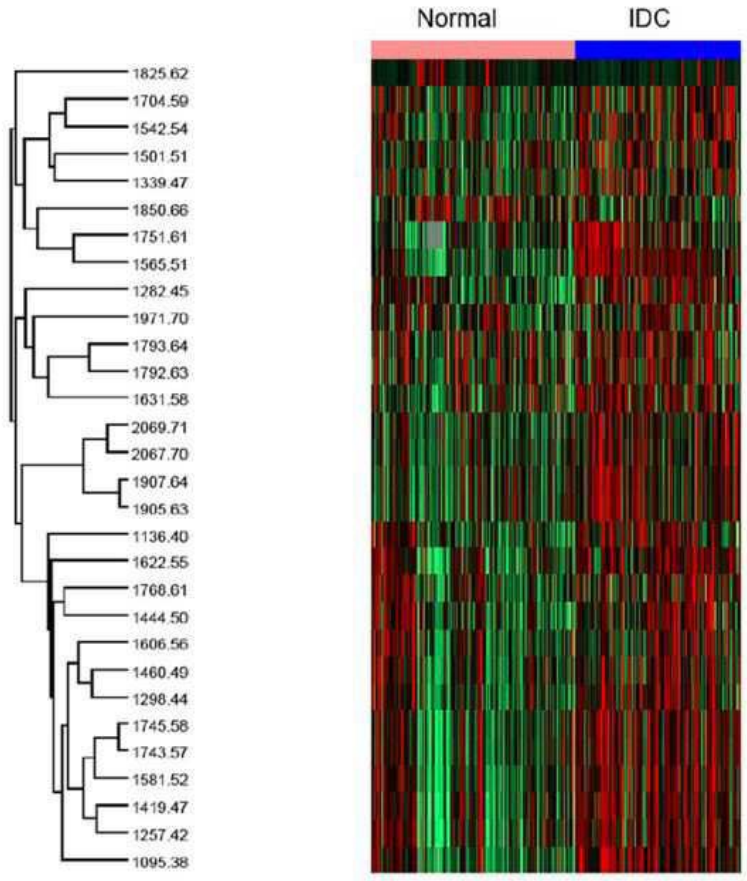
[0244] 정상인(n=30) 및 유방암 환자(n=30)간 유의미한 차이를 보이는 *m/z*값 및 TIC 정규화 강도 값

질량 (m/z)	정상	유방암	p-값
3260	2.34 ± 2.54	7.02 ± 4.14	0.000002
4207	8.95 ± 7.79	17.67 ± 6.61	0.000018
2101	1.19 ± 1.69	3.52 ± 2.66	0.000153
2550	3.92 ± 4.44	8.44 ± 4.79	0.000354
5001	9.53 ± 12.50	0.82 ± 2.03	0.000392
4150	4.88 ± 5.87	1.00 ± 1.60	0.000928

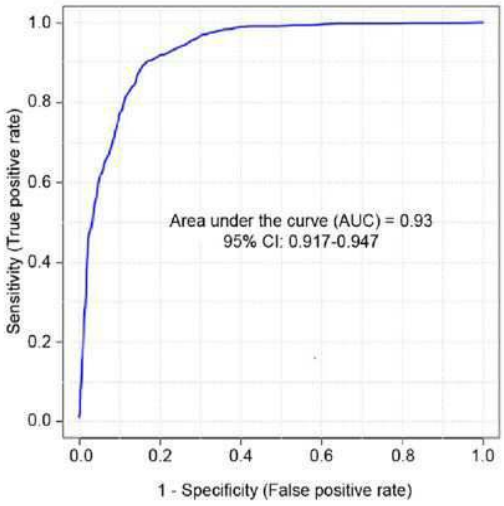
[0246] 전술한 본 발명의 설명은 예시를 위한 것이며, 본 발명이 속하는 기술분야의 통상의 지식을 가진 자는 본 발명의 기술적 사상이나 필수적인 특징을 변경하지 않고서 다른 구체적인 형태로 쉽게 변형이 가능하다는 것을 이해할 수 있을 것이다. 그러므로 이상에서 기술한 실시예들은 모든 면에서 예시적인 것이며 한정적이 아닌 것으로 이해해야 한다.

도면

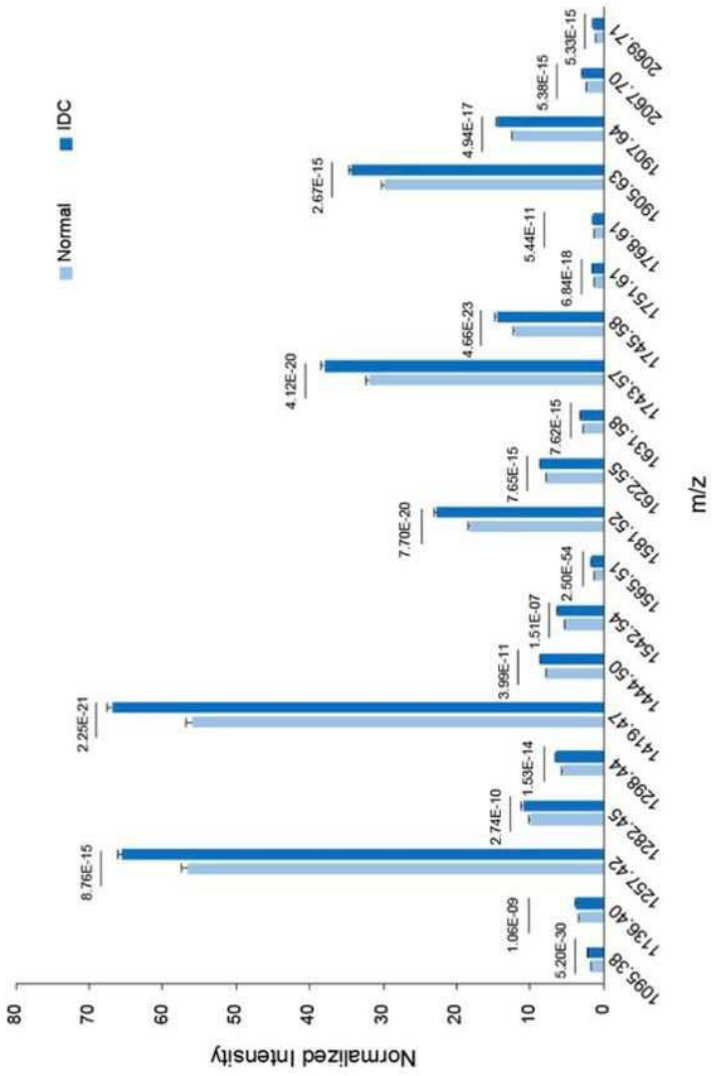
도면1


Normal
IDC
1825.62
1704.59
1542.54
1501.51
1339.47
1850.66
1751.61
1565.51
1282.45
1971.70
1793.64
1792.63
1631.58
2069.71
2067.70
1907.64
1905.63
1136.40
1622.55
1768.61
1444.50
1606.56
1460.49
1298.44
1745.58
1743.57
1581.52
1419.47
1257.42
1095.38

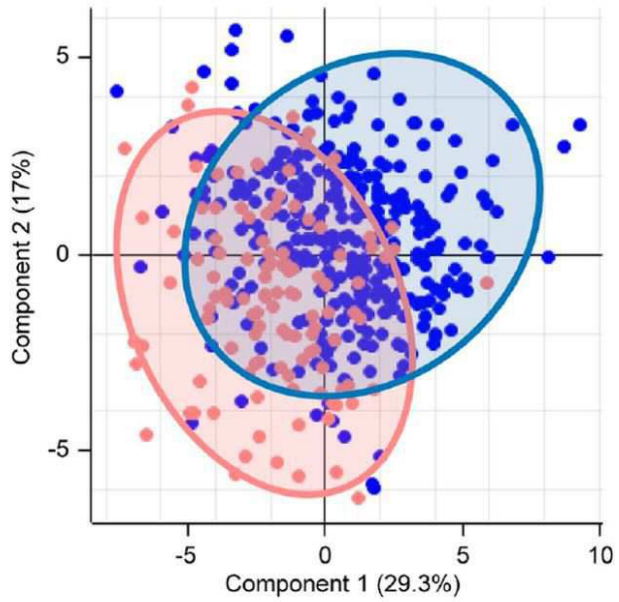
도면2


Sensitivity (True positive rate)
1 - Specificity (False positive rate)
Area under the curve (AUC) = 0.93
95% CI: 0.917-0.947
0.0
0.2
0.4
0.6
0.8
1.0

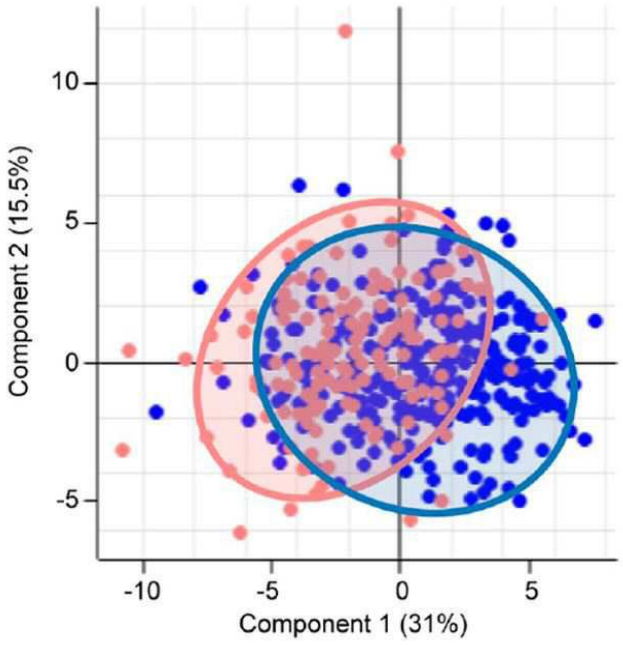
도면3


Normal
IDC
Normalized Intensity
m/z
0
10
20
30
40
50
60
70
80
1095.38
1136.40
1257.42
1282.45
1298.44
1419.47
1444.50
1542.54
1565.57
1581.52
1622.55
1631.58
1743.57
1745.58
1751.67
1766.67
1905.63
1907.64
2067.70
2069.71
5.20E-30
1.06E-09
8.76E-15
2.74E-10
1.53E-14
2.25E-21
3.99E-11
1.51E-07
2.50E-54
7.70E-20
7.65E-15
7.62E-15
4.12E-20
4.66E-23
6.84E-18
5.44E-11
2.67E-15
4.94E-17
5.38E-15
5.33E-15

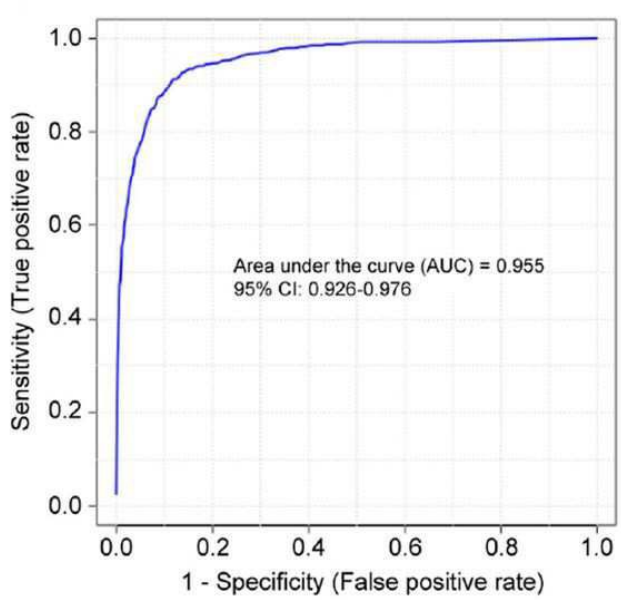
도면4a


5
0
-5
Component 2 (17%)
-5
0
5
10
Component 1 (29.3%)

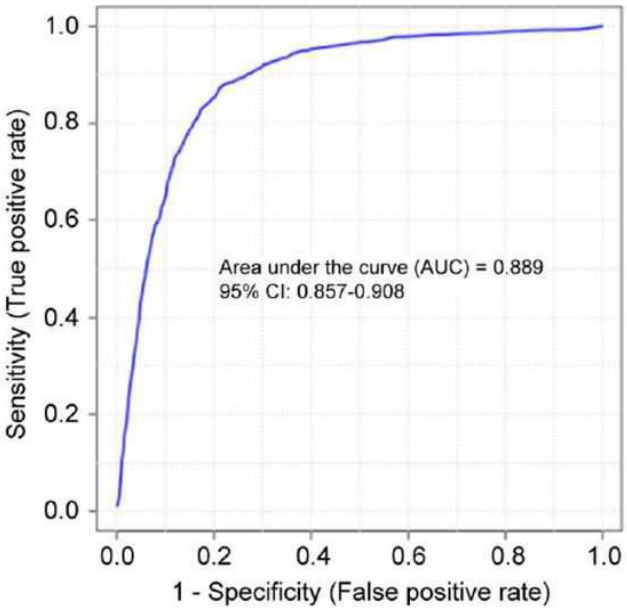
도면4b


10
5
0
-5
Component 2 (15.5%)
-10
-5
0
5
Component 1 (31%)

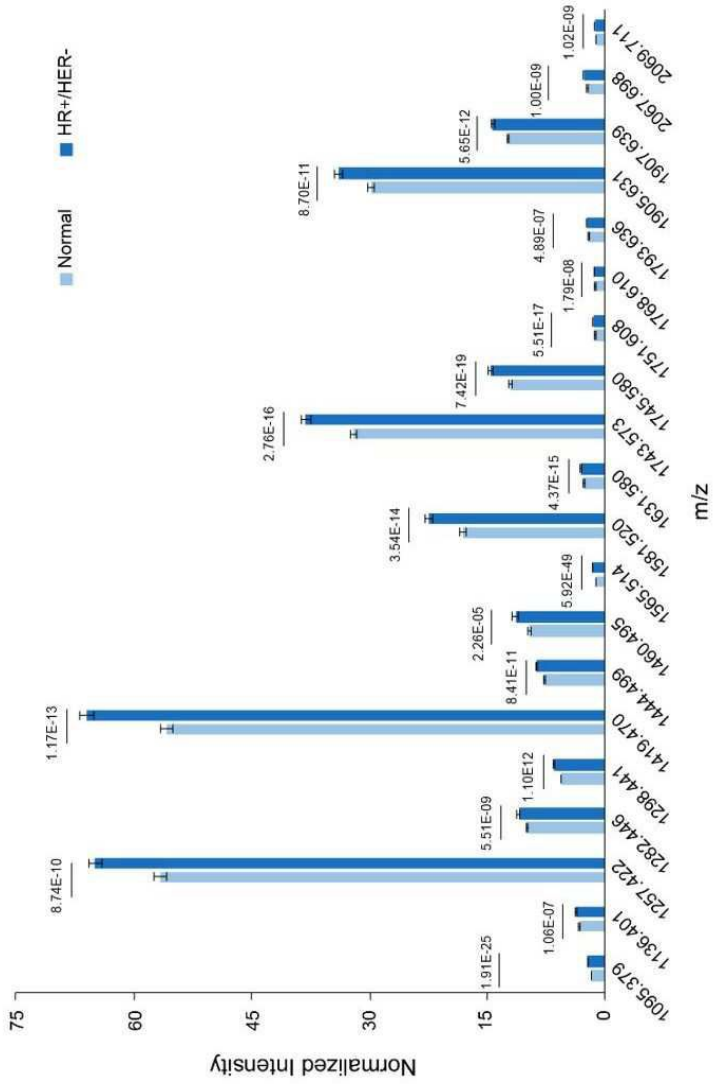
도면5a


Sensitivity (True positive rate)
1.0
0.8
0.6
0.4
0.2
0.0
Area under the curve (AUC) = 0.955
95% CI: 0.926-0.976
0.0
0.2
0.4
0.6
0.8
1.0
1 - Specificity (False positive rate)

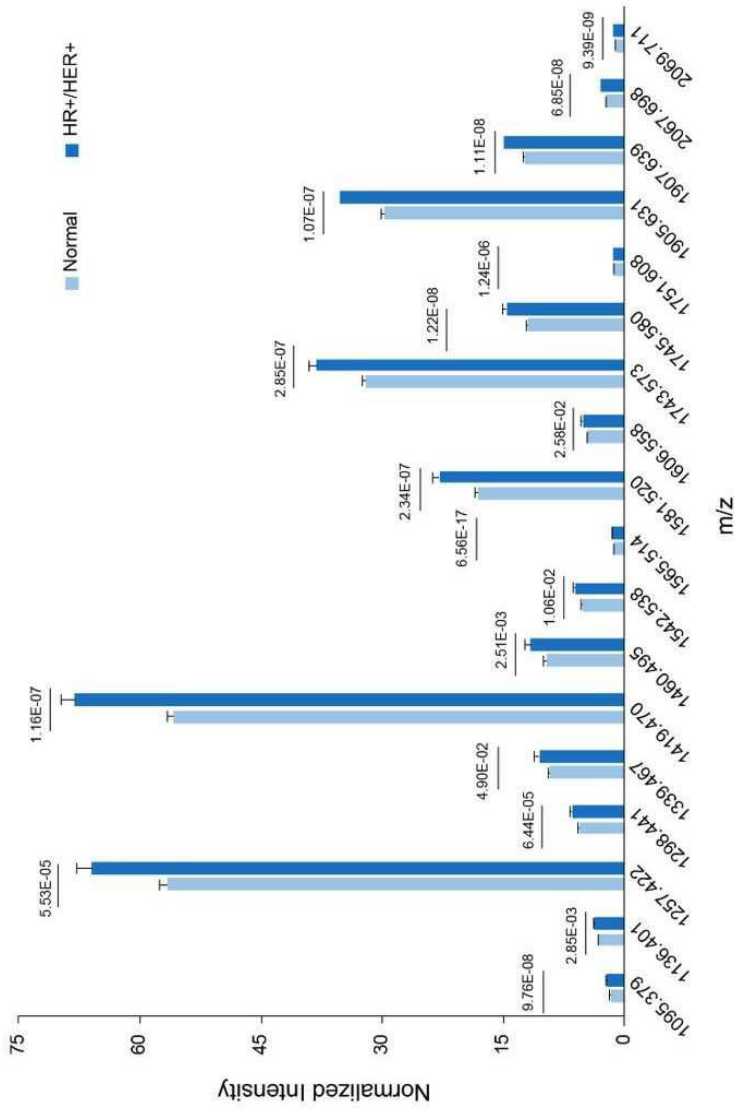
도면5b


Sensitivity (True positive rate)
1.0
0.8
0.6
0.4
0.2
0.0
Area under the curve (AUC) = 0.889
95% CI: 0.857-0.908
0.0
0.2
0.4
0.6
0.8
1.0
1 - Specificity (False positive rate)

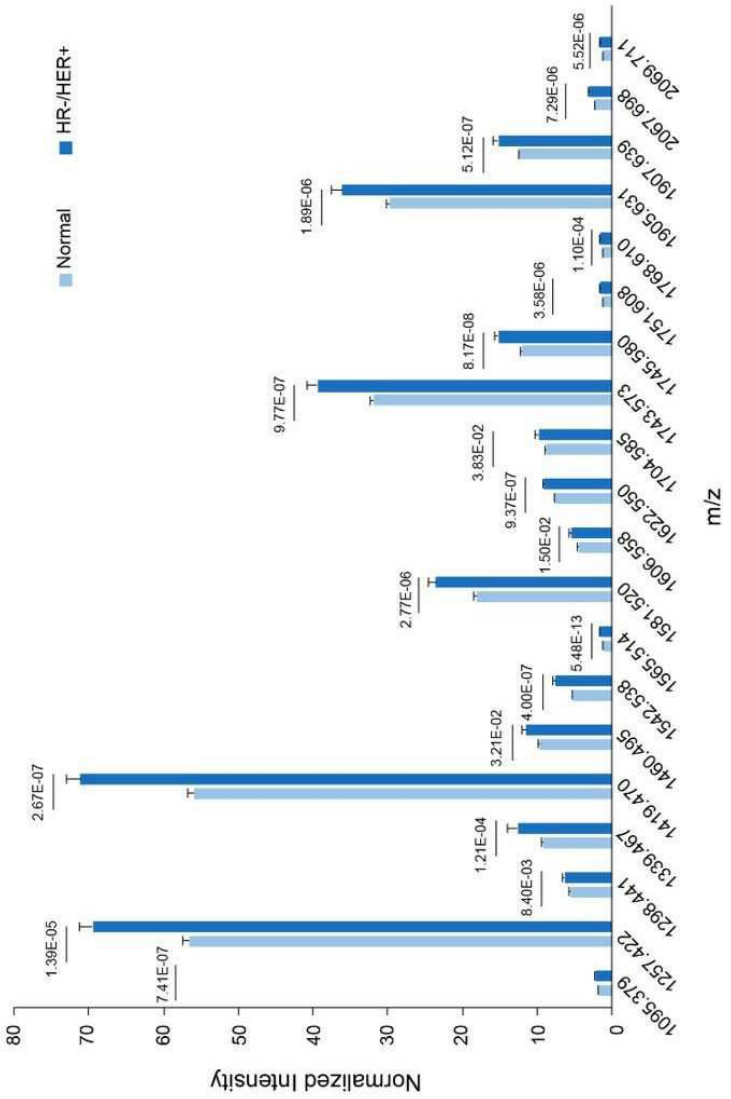
도면6a


HR+/HER-
Normal
Normalized Intensity
m/z
75
60
45
30
15
0
1095.379
1136.401
1257.422
1282.446
1298.441
1419.470
1444.499
1460.495
1565.514
1581.520
1631.580
1743.573
1745.580
1751.608
1766.610
1793.636
1905.631
1907.639
2067.698
2069.711
1.91E-25
1.06E-07
8.74E-10
1.17E-13
5.51E-09
1.10E12
8.41E-11
2.26E-05
5.92E-49
3.54E-14
4.37E-15
2.76E-16
7.42E-19
5.51E-17
1.79E-08
4.89E-07
8.70E-11
5.65E-12
1.00E-09
1.02E-09

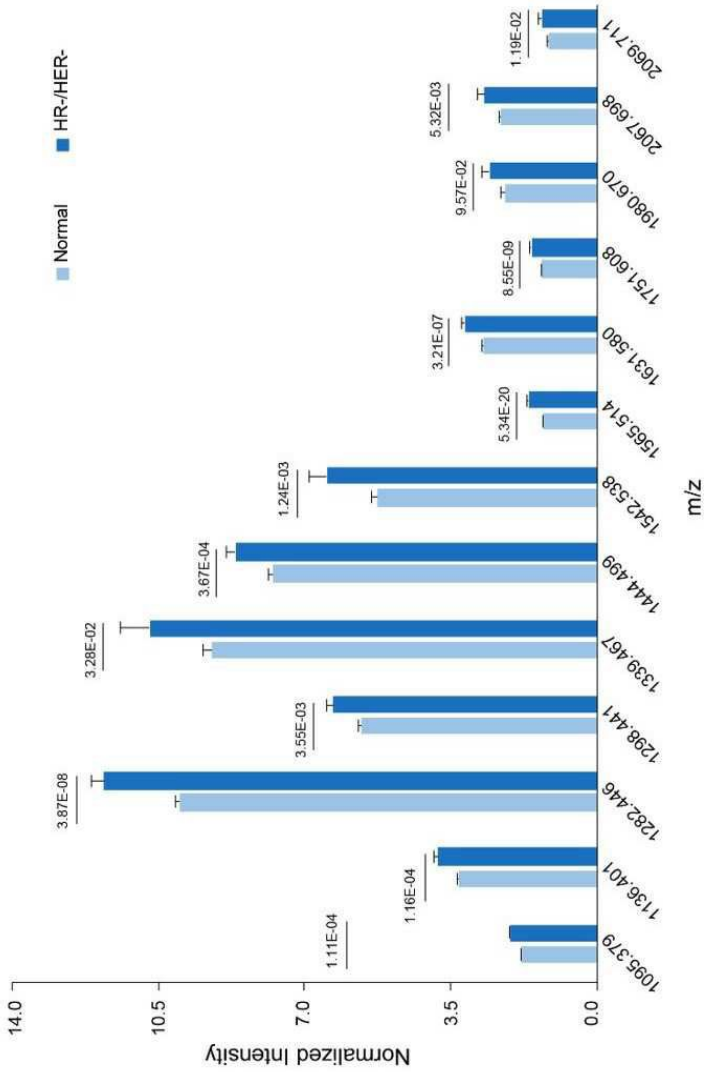
도면6b


HR+/HER+
Normal
Normalized Intensity
m/z
75
60
45
30
15
0
1095.379
1136.401
1257.422
1298.441
1339.467
1419.470
1460.495
1542.538
1565.514
1581.520
1606.558
1743.573
1745.580
1751.608
1905.631
1907.639
2067.698
2069.711
9.76E-08
2.85E-03
5.53E-05
6.44E-05
4.90E-02
1.16E-07
2.51E-03
1.06E-02
6.56E-17
2.34E-07
2.58E-02
2.85E-07
1.22E-08
1.24E-06
1.07E-07
1.11E-08
6.85E-08
9.39E-09

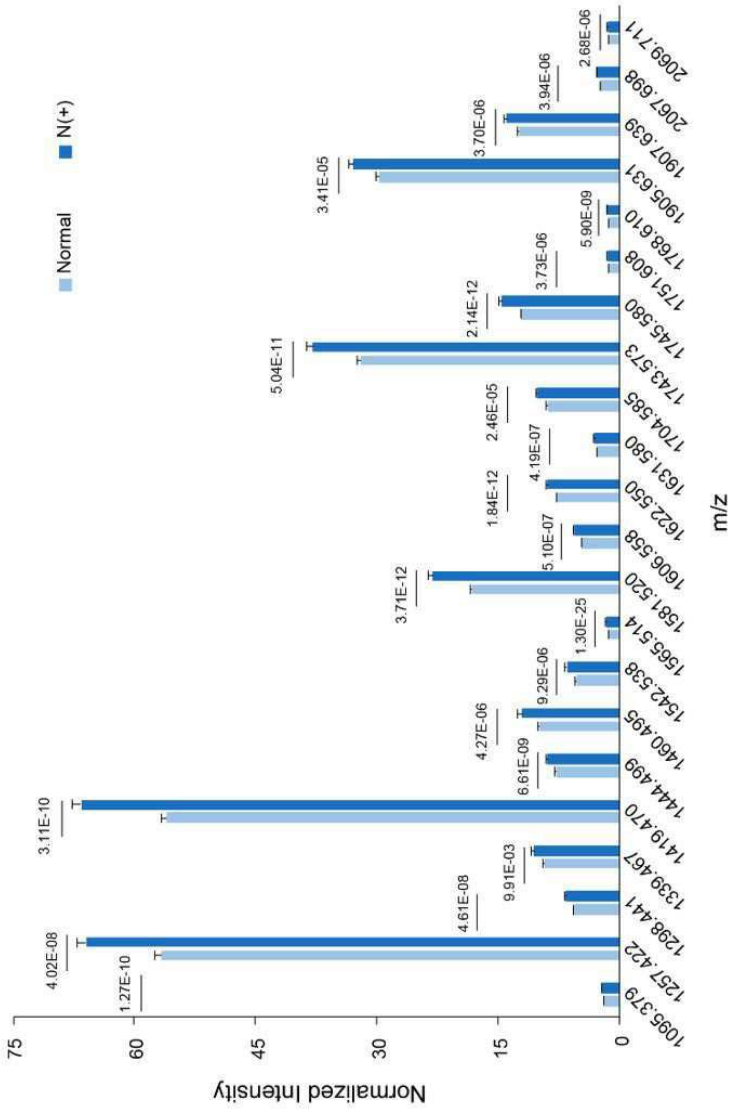
도면6c


HR-/HER+
Normal
Normalized Intensity
m/z
0
10
20
30
40
50
60
70
80
1095.379
1257.422
1298.441
1339.467
1419.470
1460.495
1542.538
1565.514
1581.520
1606.558
1622.550
1704.585
1743.573
1745.580
1751.608
1768.610
1905.631
1907.639
2067.698
2069.711
1.39E-05
7.41E-07
8.40E-03
1.21E-04
2.67E-07
3.21E-02
4.00E-07
5.48E-13
2.77E-06
1.50E-02
9.37E-07
3.83E-02
9.77E-07
8.17E-08
3.58E-06
1.10E-04
1.89E-06
5.12E-07
7.29E-06
5.52E-06

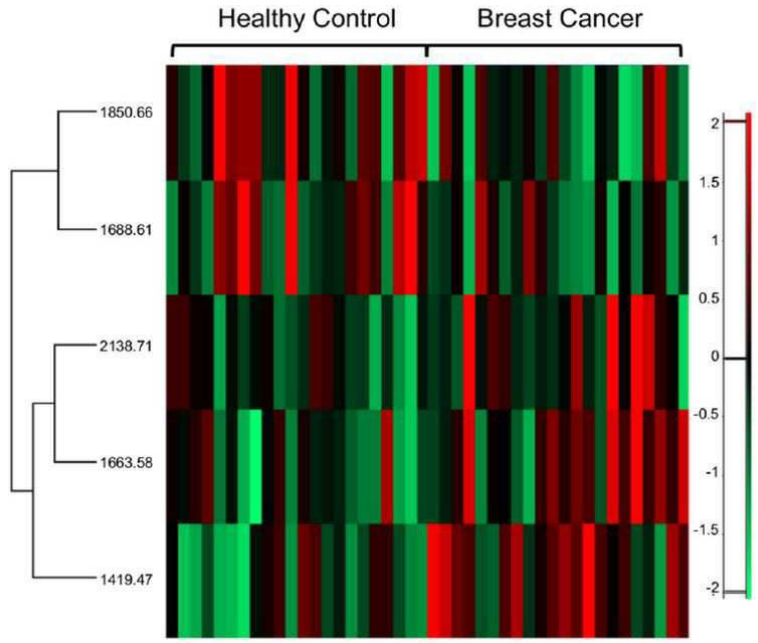
도면6d


HR-/HER-
Normal
Normalized Intensity
14.0
10.5
7.0
3.5
0.0
m/z
1096.379
1136.401
1282.446
1298.441
1339.467
1444.499
1542.538
1565.514
1631.580
1751.608
1980.670
2067.698
2069.711
1.11E-04
1.16E-04
3.87E-08
3.55E-03
3.28E-02
3.67E-04
1.24E-03
5.34E-20
3.21E-07
8.55E-09
9.57E-02
5.32E-03
1.19E-02

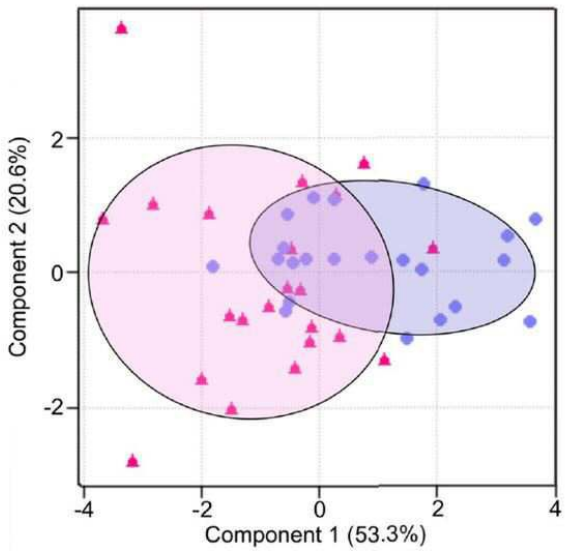
도면7


Normal
N(+)
Normalized Intensity
m/z
75
60
45
30
15
0
1095.379
1257.422
1298.441
1339.467
1419.470
1444.499
1460.495
1542.538
1565.514
1581.520
1606.558
1622.550
1631.580
1704.585
1743.573
1745.580
1751.608
1768.670
1905.631
1907.639
2067.698
2069.711
4.02E-08
1.27E-10
3.11E-10
4.61E-08
9.91E-03
6.61E-09
4.27E-06
9.29E-06
1.30E-25
3.71E-12
5.10E-07
1.84E-12
4.19E-07
2.46E-05
5.04E-11
2.14E-12
3.73E-06
5.90E-09
3.41E-05
3.70E-06
3.94E-06
2.68E-06

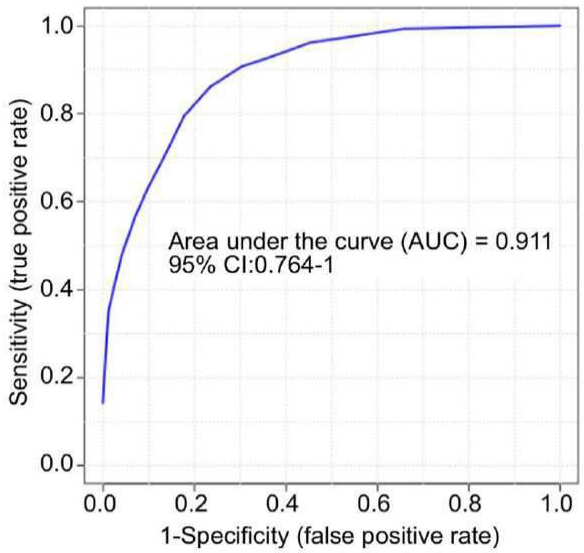
도면8


Healthy Control
Breast Cancer
1850.66
1688.61
2138.71
1663.58
1419.47
2
1.5
1
0.5
0
-0.5
-1
-1.5
-2

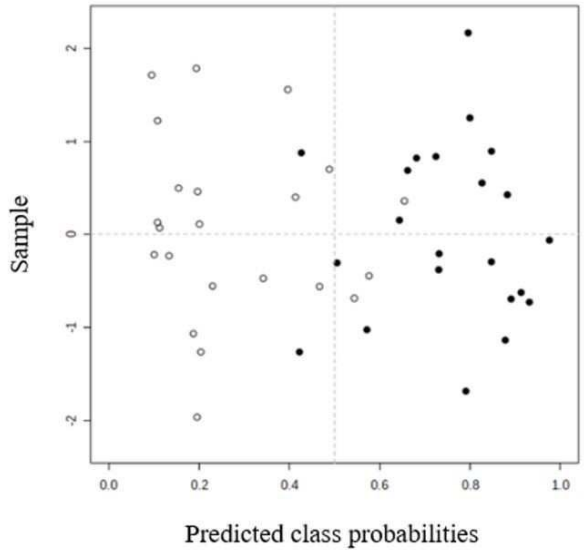
도면9


Component 2 (20.6%)
2
0
-2
-4
-2
0
2
4
Component 1 (53.3%)

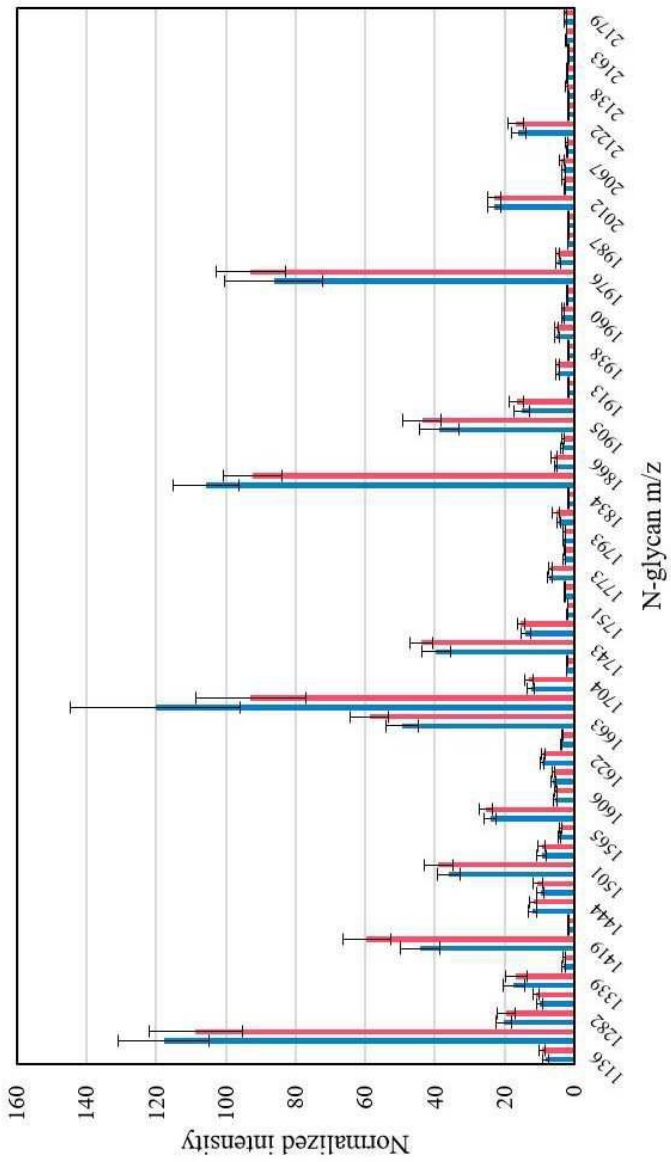
도면10


1.0
0.8
0.6
0.4
0.2
0.0
Sensitivity (true positive rate)
Area under the curve (AUC) = 0.911
95% CI:0.764-1
0.0
0.2
0.4
0.6
0.8
1.0
1-Specificity (false positive rate)

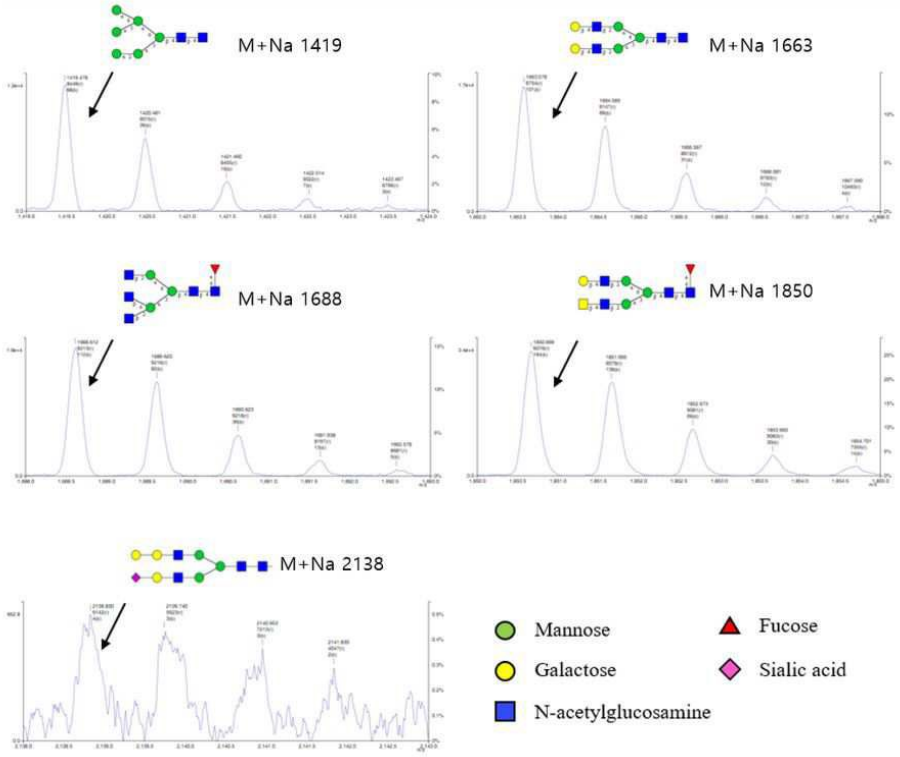
도면11


Sample
2
1
0
-1
-2
0.0
0.2
0.4
0.6
0.8
1.0
Predicted class probabilities

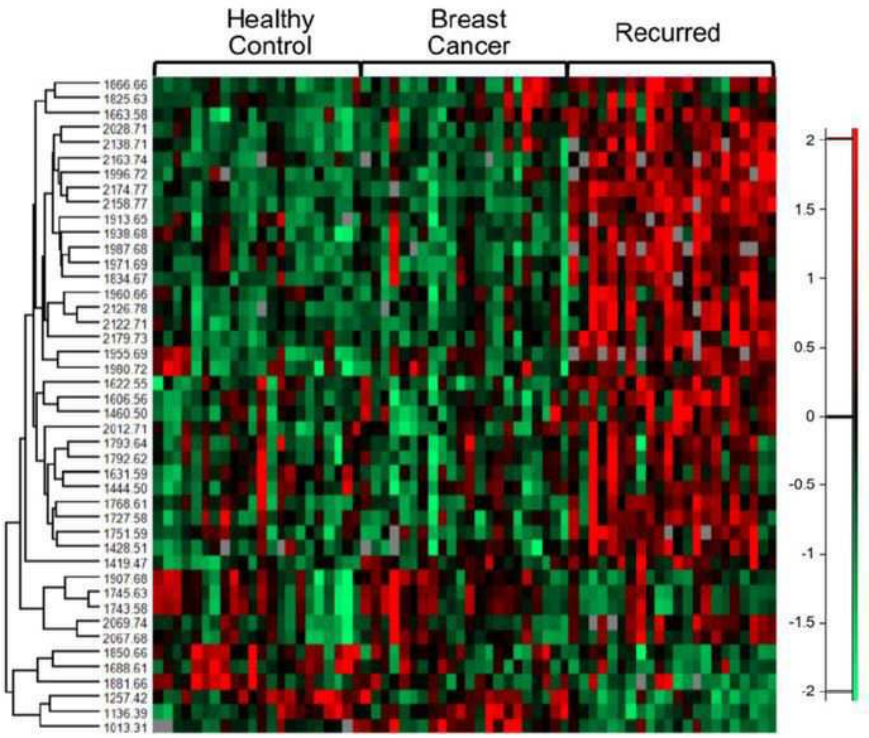
도면12


Normalized intensity
0
20
40
60
80
100
120
140
160
N-glycan m/z
1136
1282
1339
1419
1444
1501
1565
1606
1622
1663
1704
1743
1751
1773
1793
1834
1866
1905
1913
1938
1960
1976
1987
2012
2067
2122
2138
2163
2179

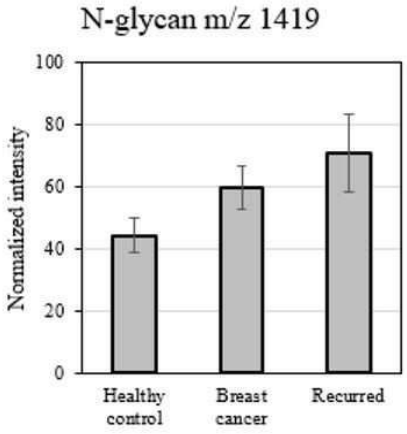
도면13


M+Na 1419
M+Na 1663
M+Na 1688
M+Na 1850
M+Na 2138
Mannose
Galactose
N-acetylglucosamine
Fucose
Sialic acid

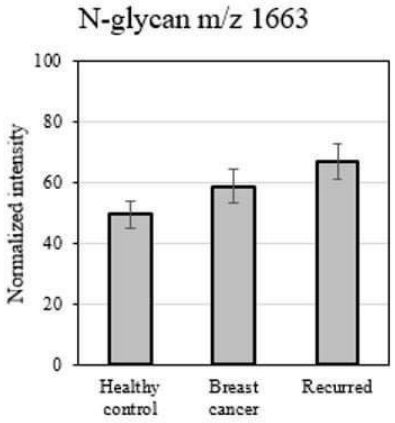
도면14


Healthy Control
Breast Cancer
Recurred
1666.66
1825.63
1663.58
2028.71
2138.71
2163.74
1996.72
2174.77
2158.77
1513.65
1938.68
1987.68
1971.69
1834.67
1960.66
2126.78
2122.71
2179.73
1955.69
1980.72
1622.55
1606.56
1460.50
2012.71
1793.64
1792.62
1631.59
1444.50
1768.61
1727.58
1751.59
1428.51
1419.47
1907.68
1745.63
1743.58
2069.74
2067.68
1650.66
1688.61
1881.66
1257.42
1136.39
1013.31
2
1.5
1
0.5
0
-0.5
-1
-1.5
-2

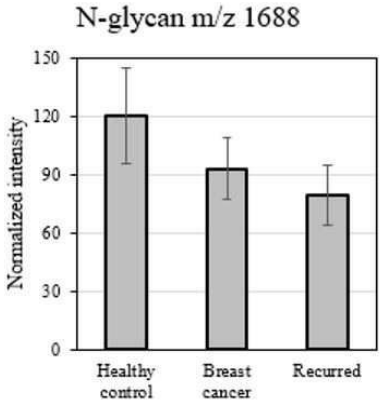
도면15a



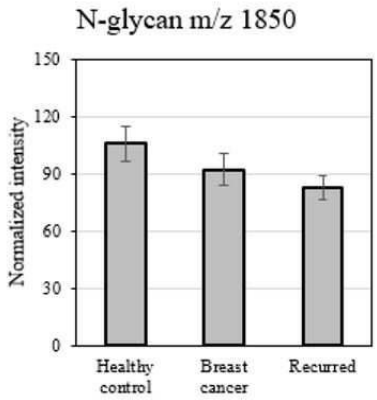
도면15b



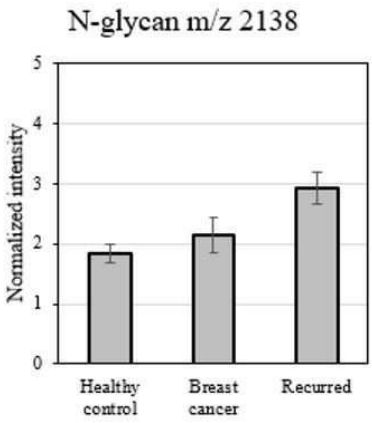
도면15c



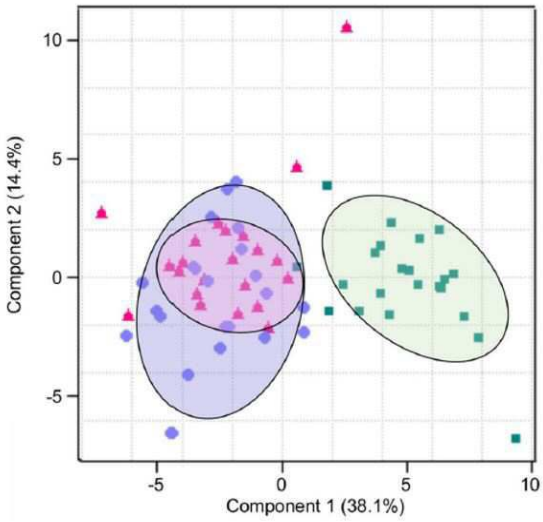
도면15d



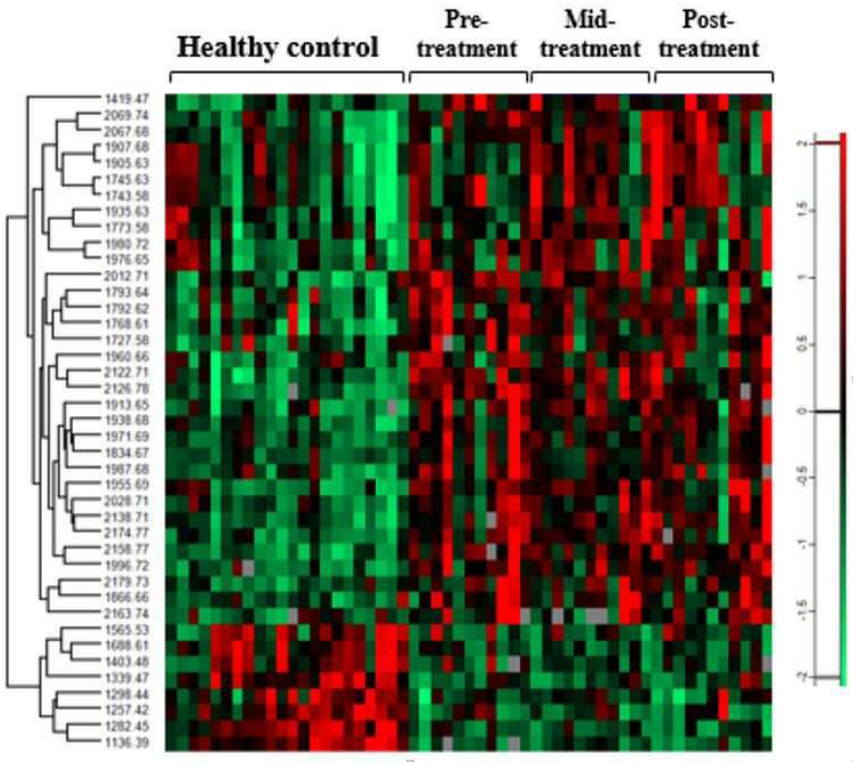
도면15e



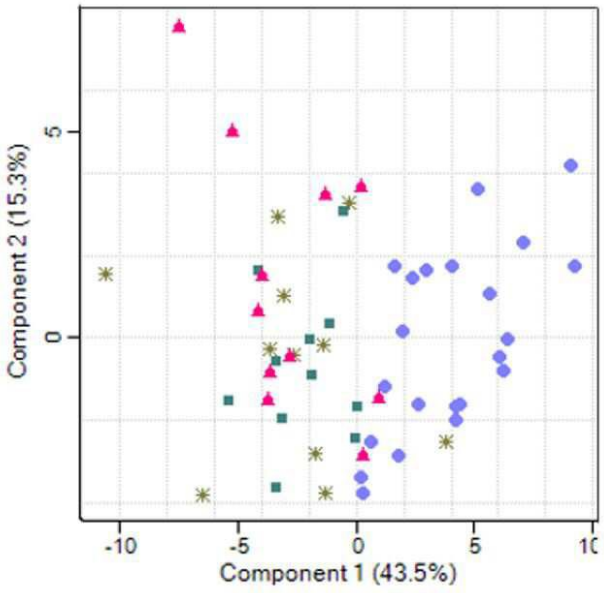
도면16



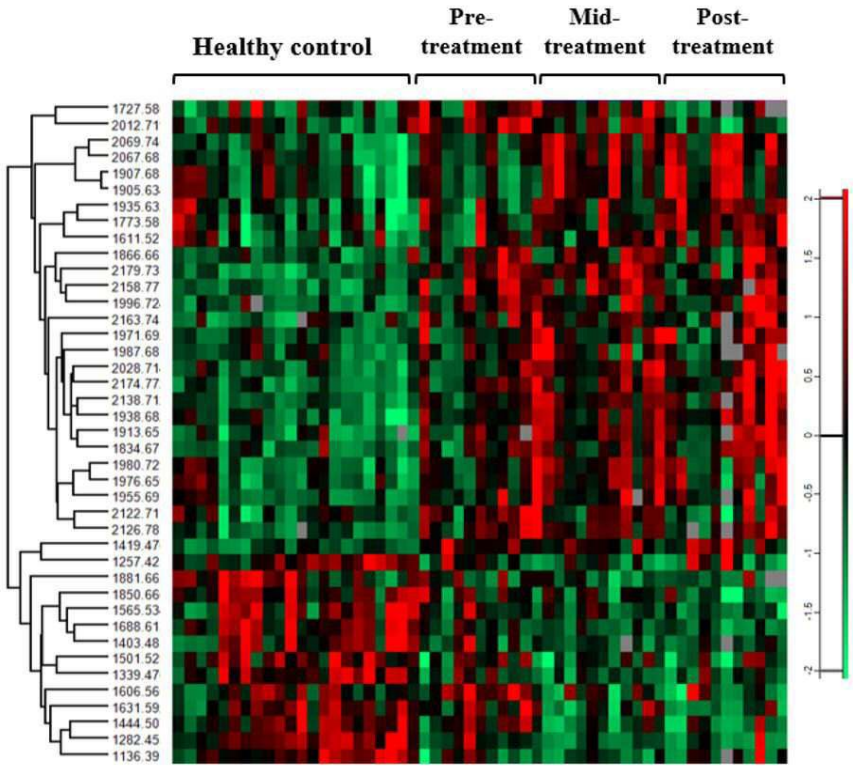
도면17


Healthy control
Pre-treatment
Mid-treatment
Post-treatment

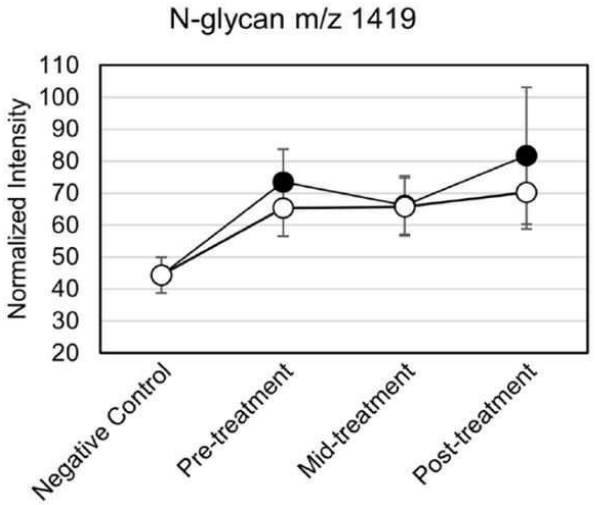
도면18


Component 2 (15.3%)
Component 1 (43.5%)
-10
-5
0
5
10

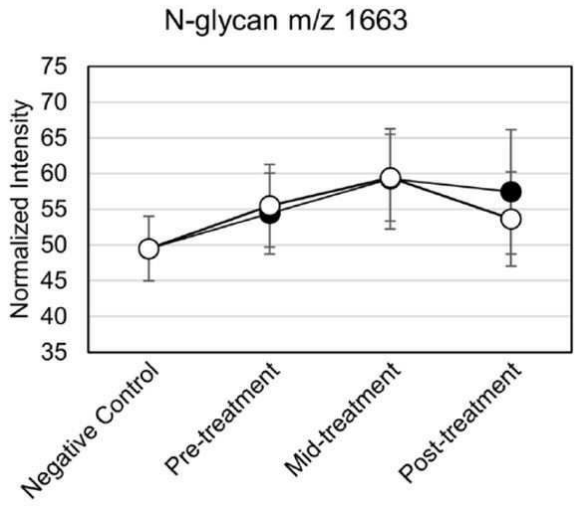
도면19


Healthy control
Pre-treatment
Mid-treatment
Post-treatment
1727.58
2012.71
2069.74
2067.68
1907.68
1905.63
1935.63
1773.58
1611.52
1866.66
2179.73
2158.77
1996.72
2163.74
1971.69
1987.68
2028.71
2174.77
2138.71
1938.68
1913.65
1834.67
1980.72
1976.65
1955.69
2122.71
2126.78
1419.47
1257.42
1881.66
1850.66
1565.53
1688.61
1403.48
1501.52
1339.47
1606.56
1631.59
1444.50
1282.45
1136.39

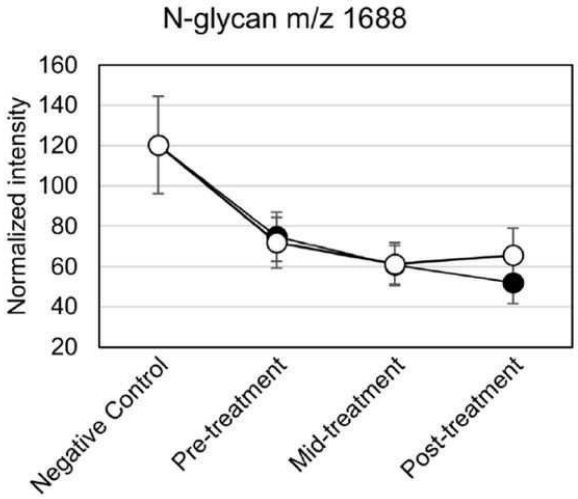
도면20a


N-glycan m/z 1419
Normalized Intensity
110
100
90
80
70
60
50
40
30
20
Negative Control
Pre-treatment
Mid-treatment
Post-treatment

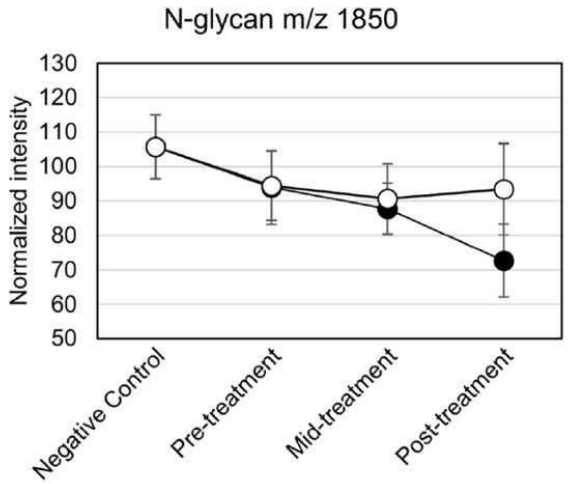
도면20b



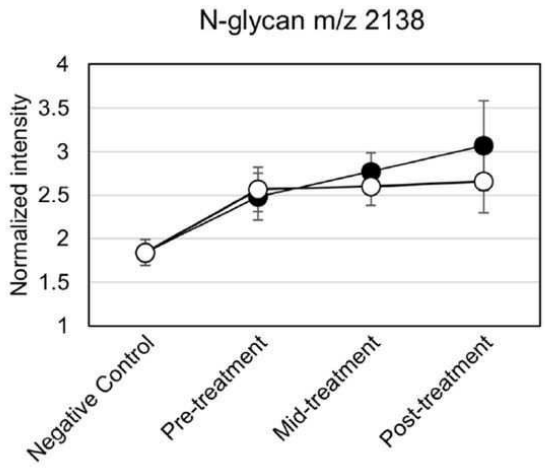
도면20c



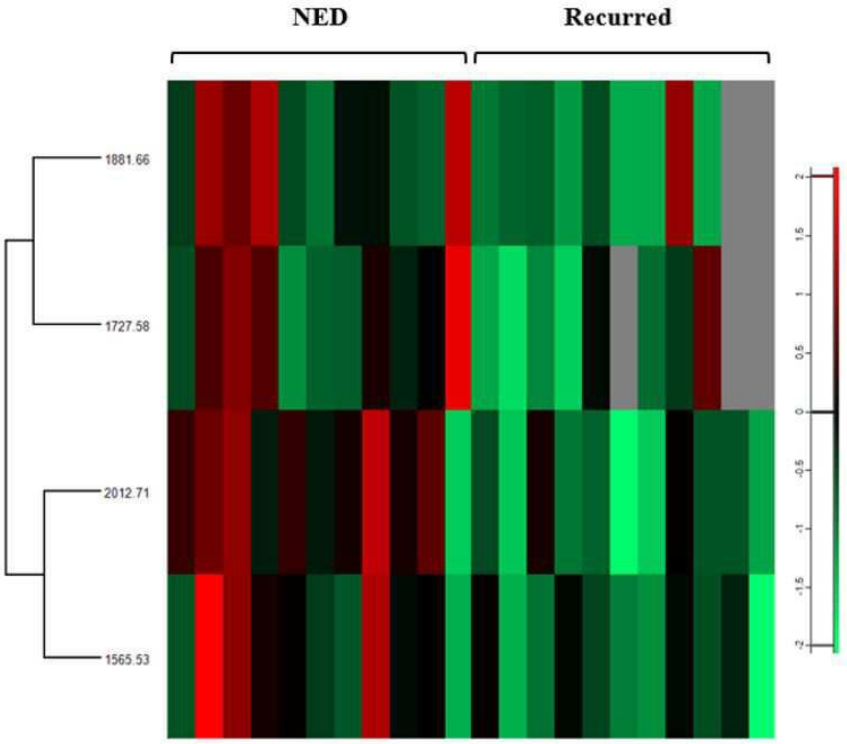
도면20d



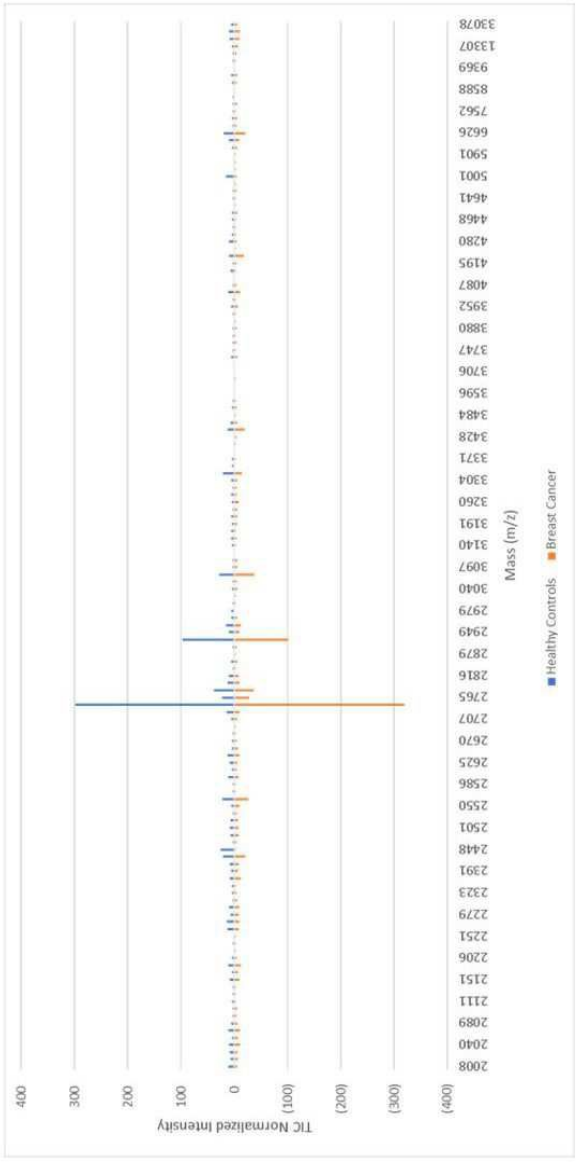
도면20e


N-glycan m/z 2138
Normalized intensity
4
3.5
3
2.5
2
1.5
1
Negative Control
Pre-treatment
Mid-treatment
Post-treatment

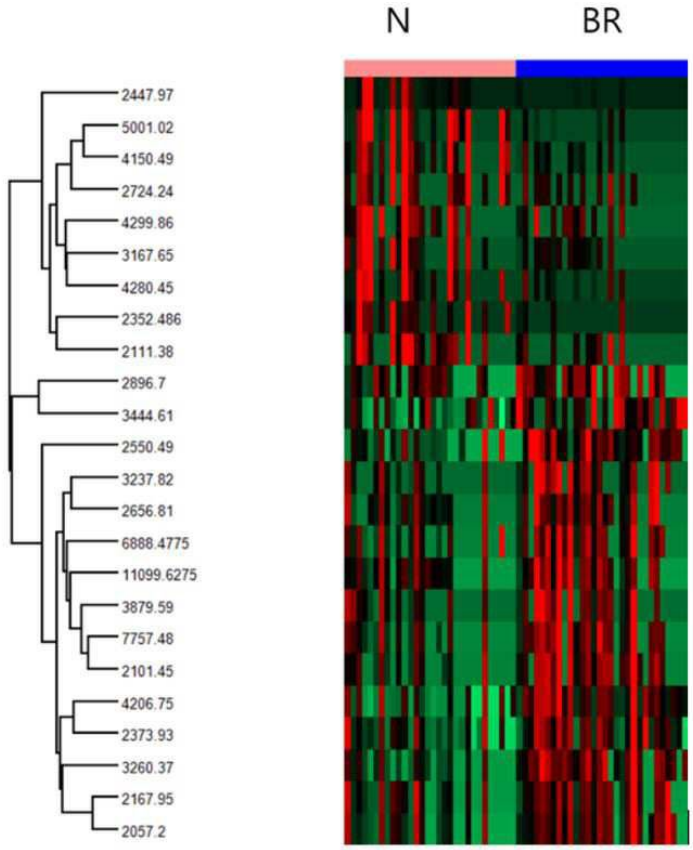
도면21


NED
Recurred
1881.66
1727.58
2012.71
1565.53

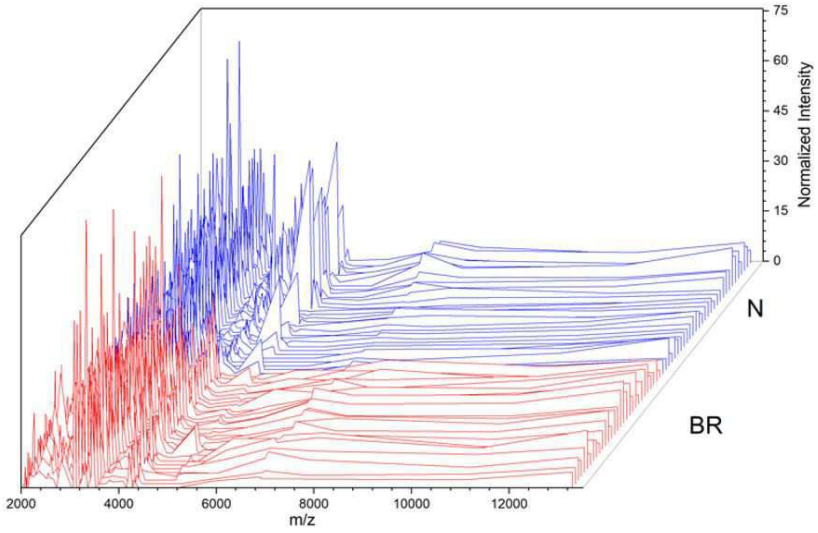
도면22


TIC Normalized Intensity
400
300
200
100
0
(100)
(200)
(300)
(400)
Mass (m/z)
Healthy Controls
Breast Cancer

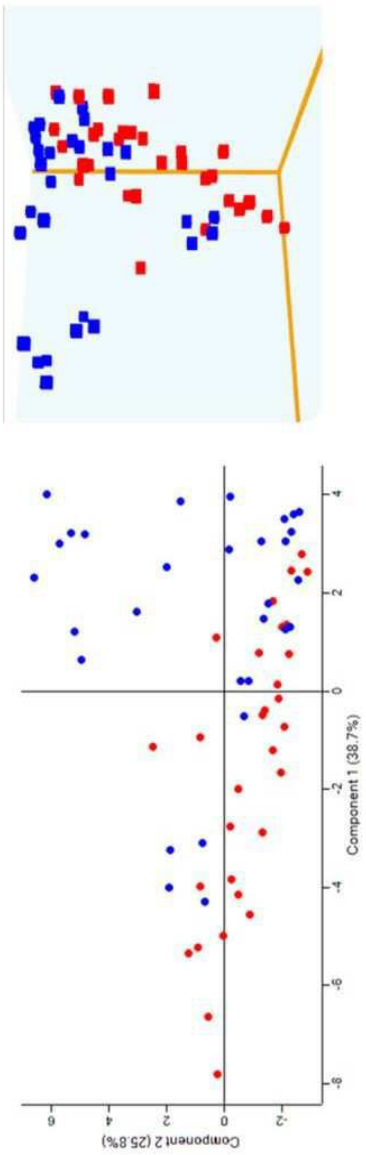
도면23


N
BR
2447.97
5001.02
4150.49
2724.24
4299.86
3167.65
4280.45
2352.486
2111.38
2896.7
3444.61
2550.49
3237.82
2656.81
6888.4775
11099.6275
3879.59
7757.48
2101.45
4206.75
2373.93
3260.37
2167.95
2057.2

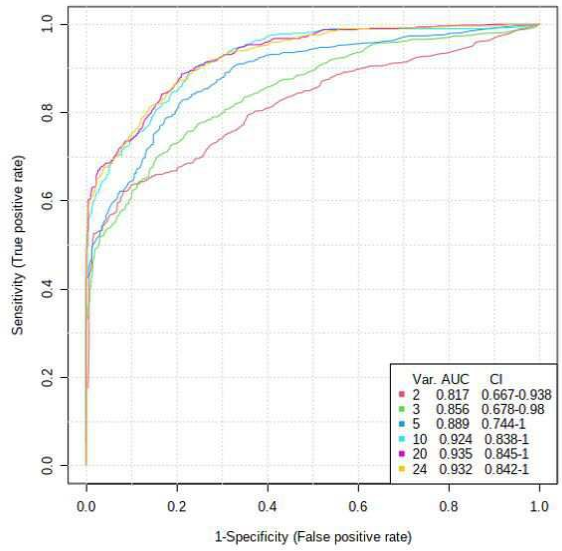
도면24


75
60
45
30
15
0
Normalized Intensity
N
BR
2000
4000
6000
8000
10000
12000
m/z

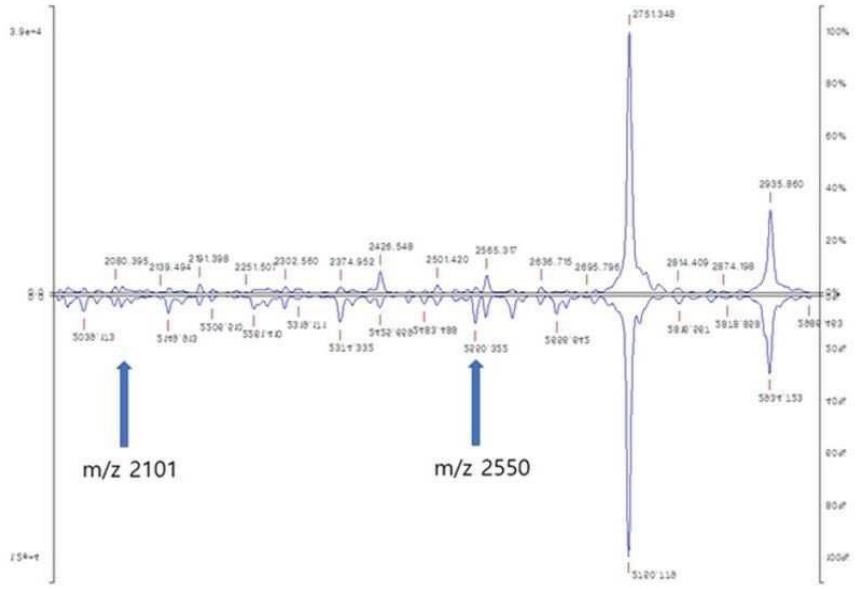
도면25



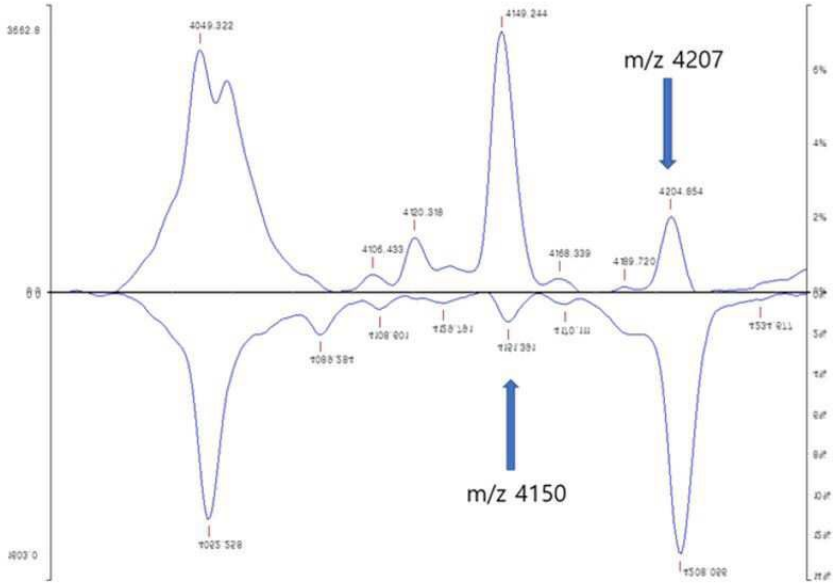
도면26


Sensitivity (True positive rate)
1-Specificity (False positive rate)
Var. AUC CI
2 0.817 0.667-0.938
3 0.856 0.678-0.98
5 0.889 0.744-1
10 0.924 0.838-1
20 0.935 0.845-1
24 0.932 0.842-1

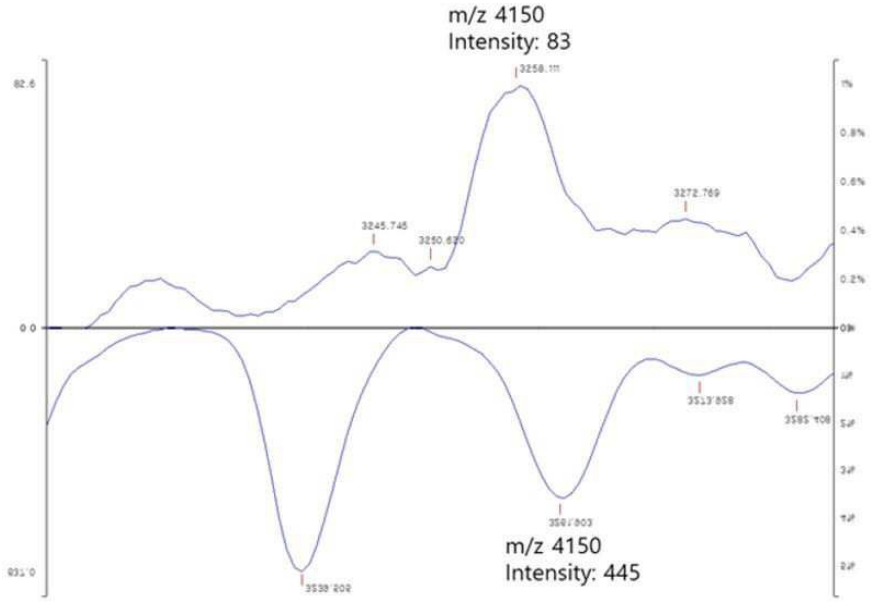
도면27a


m/z 2101
m/z 2550

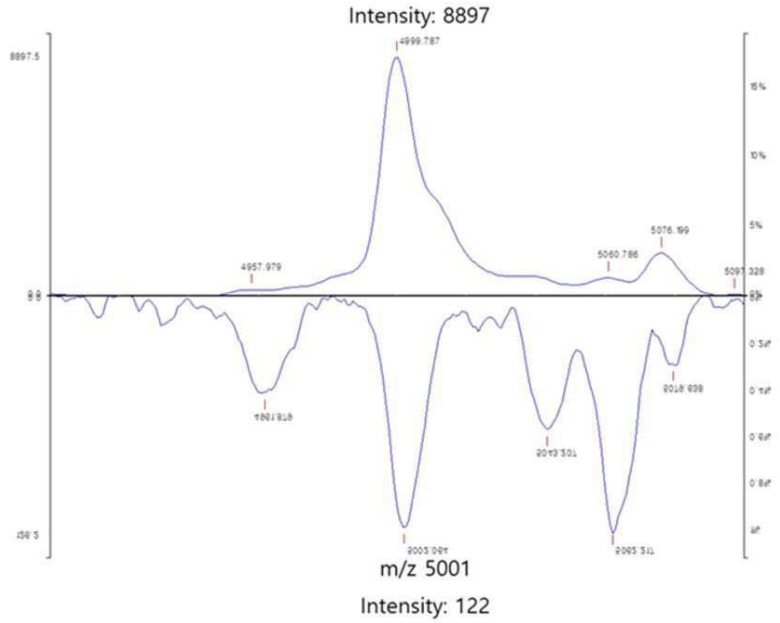
도면27b



도면27c



도면27d


Intensity: 8897
4999.787
5076.199
5060.786
4957.979
5097.328
m/z 5001
Intensity: 122