



(19) 대한민국특허청(KR)

(12) 등록특허공보(B1)

(45) 공고일자 2015년01월22일

(11) 등록번호 10-1484969

(24) 등록일자 2015년01월15일

(51) 국제특허분류(Int. Cl.)

C07K 1/14 (2006.01) G01N 33/574 (2006.01)

G01N 33/68 (2006.01) G01N 33/58 (2006.01)

(21) 출원번호 10-2014-0035062

(22) 출원일자 2014년03월26일

심사청구일자 2014년03월26일

(56) 선행기술조사문헌

JP2009168470 A

WO2013106886 A1

KR1020100120788 A

JP2009270996 A

(73) 특허권자

충남대학교산학협력단

대전광역시 유성구 대학로 99 (궁동, 충남대학교)

한국과학기술원

대전광역시 유성구 대학로 291(구성동)

(72) 발명자

안현주

대전광역시 유성구 지족로 343, 211동 2002호 (지족동, 반석마을아파트)

김제한

대전광역시 유성구 지족로 343, 211동 2002호 (지족동, 반석마을아파트)

(뒷면에 계속)

(74) 대리인

강성혜

전체 청구항 수 : 총 9 항

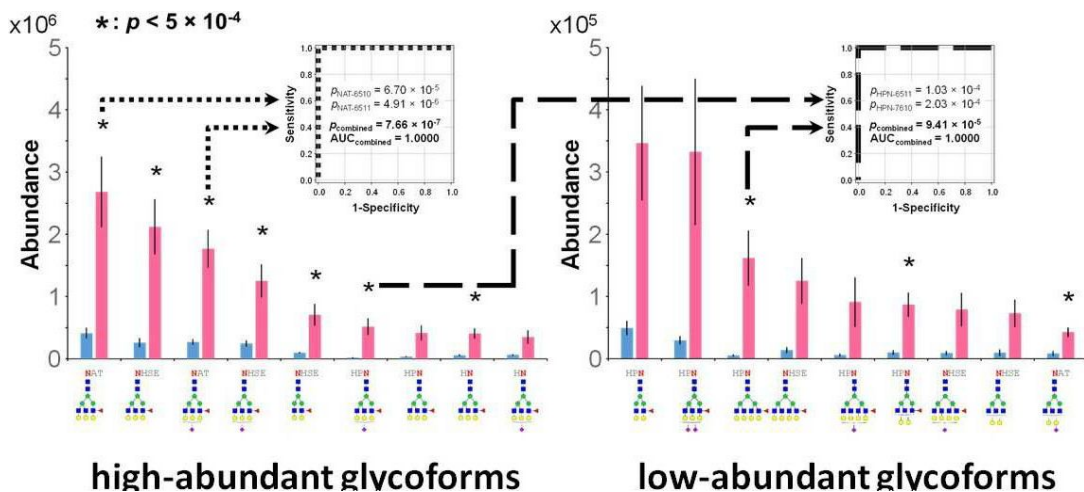
심사관 : 문동현

(54) 발명의 명칭 N-당쇄화된 당펩타이드를 이용한 암 진단방법

(57) 요약

본 발명은 바이오마커로서 부위 특이적으로 N-당쇄화된 당펩타이드를 이용하여 바이오마커의 N-당쇄변화를 탐지하는 방법에 관한 것이다. 좀 더 구체적으로, 본 발명에서는 비특이적 단백질 분해효소를 이용하여 혈액의 헵토글로빈을 펩타이드와 당펩타이드로 분해하였으며, 207번과 211번 등 총 4개의 N-당쇄화 부위를 가진 당펩타이드만을 분리하였다. 본 발명자들은 이렇게 얻은 4개의 N-당쇄화 부위에 결합된 당펩타이드를 이용하여 종양 특이적 바이오마커를 발굴함으로써 고특이적이고 고감도인 종양의 N-당쇄변화를 탐지하는 방법을 발명하기에 이르렀다.

대표도



(72) 발명자

씨리너스 화

대전광역시 유성구 대학로 179 그린빌 302호

오명진

대전광역시 유성구 봉명동 1105-1번지 301호

김정희

대전 유성구 대학로 291, 한국과학기술원 생명과학
과 (구성동)

이성현

서울특별시 서초구 방배로18길 67, 102-601 (방배
동, 방배자이아파트)

박승열

경기도 화성시 10용사로 221, 102동 403호 (병점
동, 성호아파트)

이 발명을 지원한 국가연구개발사업

과제고유번호 2013K000430

부처명 미래창조과학부

연구관리전문기관 한국기초과학지원연구원

연구사업명 신기술융합형성장동력사업

연구과제명 MALDI 질량분석기반 암세포 표지자 발굴 및 암진단 시스템 개발

기 여 율 1/3

주관기관 충남대학교 산학협력단

연구기간 2013.07.01 ~ 2014.06.03

이 발명을 지원한 국가연구개발사업

과제고유번호 NRF-2013M3A9B6075933

부처명 미래창조과학부

연구관리전문기관 한국연구재단

연구사업명 차세대응용오믹스사업

연구과제명 당단백질 의약품의 당 변이체 특성분석 및 당분석과학기술법 개발

기 여 율 1/3

주관기관 충남대학교

연구기간 2013.11.20 ~ 2014.11.19

이 발명을 지원한 국가연구개발사업

과제고유번호 2013075931

부처명 미래창조과학부

연구관리전문기관 한국연구재단

연구사업명 차세대응용오믹스사업

연구과제명 글리칸 변이체 제작을 위한 글리칸 조절.제어 원천기술 개발

기 여 율 1/3

주관기관 KAIST

연구기간 2013.11.20 ~ 2014.11.19

특허청구의 범위

청구항 1

- a) 시료에서 타겟 단백질을 분리 및 정제하는 단계;
- b) 분리 및 정제한 타겟 단백질에 비특이적 단백질 분해효소를 처리하는 단계;
- c) 효소 처리 후 N-당쇄화된 당펩타이드를 분리하는 단계;
- d) 분리한 N-당쇄화된 당펩타이드를 질량분석하여 이성질체를 분리하는 단계;
- e) 분리한 이성질체의 구조 분석 및 당쇄와 펩타이드 구성을 확인하는 단계; 및
- f) 분리한 이성질체의 정량적 프로파일링을 수행하는 단계;를 포함하는 바이오마커 단백질의 N-당쇄변화 탐지방법.

청구항 2

청구항 1에 있어서,

상기 타겟 단백질은 헵토글로빈임을 특징으로 하는, 바이오마커 단백질의 N-당쇄변화 탐지방법.

청구항 3

청구항 1에 있어서,

상기 N-당쇄화된 당펩타이드는 혈청 헵토글로빈의 184번, 207번, 211번 및 241번 아스파라긴 부위가 당쇄화된 것임을 특징으로 하는, 바이오마커 단백질의 N-당쇄변화 탐지방법.

청구항 4

청구항 1에 있어서,

상기 d) 단계의 분리한 N-당쇄화된 당펩타이드는 혈청 헵토글로빈 유래 당펩타이드로서

히스티딘-184번 아스파라긴-Hexose5_HexNAc4_Fucose1,

히스티딘-184번 아스파라긴-Hexose6_HexNAc5_Fucose1,

히스티딘-184번 아스파라긴-Hexose4_HexNAc3,

히스티딘-184번 아스파라긴-Hexose5_HexNAc4,

히스티딘-184번 아스파라긴-Hexose5_HexNAc4_NeuAc1,

Hexose4_HexNAc3-207번 아스파라긴-히스티딘-세린-글루탐산,

Hexose5_HexNAc4-207번 아스파라긴-히스티딘-세린-글루탐산,

Hexose4_HexNAc4-207번 아스파라긴-히스티딘-세린-글루탐산,

Hexose6_HexNAc5_Fucose1-207번 아스파라긴-히스티딘-세린-글루탐산,

Hexose6_HexNAc5-207번 아스파라긴-히스티딘-세린-글루탐산,

Hexose8_HexNAc2-207번 아스파라긴-히스티딘-세린-글루탐산,

Hexose6_HexNAc5_Fucose1_NeuAc1-207번 아스파라긴-히스티딘-세린-글루탐산,

Hexose5_HexNAc4_Fucose1-207번 아스파라긴-히스티딘-세린-글루탐산,
 Hexose7_HexNAc6-207번 아스파라긴-히스티딘-세린-글루탐산,
 Hexose6_HexNAc5_Fucose1_NeuAc1-211번 아스파라긴-알라닌-트레오닌,
 Hexose6_HexNAc5_Fucose1-211번 아스파라긴-알라닌-트레오닌,
 Hexose7_HexNAc6_Fucose1-211번 아스파라긴-알라닌-트레오닌,
 Hexose5_HexNAc5_NeuAc1-211번 아스파라긴-알라닌-트레오닌,
 Hexose4_HexNAc4_Fucose1-211번 아스파라긴-알라닌-트레오닌,
 Hexose7_HexNAc6_Fucose1_NeuAc1-211번 아스파라긴-알라닌-트레오닌,
 Hexose7_HexNAc6_NeuAc1-211번 아스파라긴-알라닌-트레오닌,
 히스티딘-프롤린-241번 아스파라긴-Hexose5_HexNAc4,
 히스티딘-프롤린-241번 아스파라긴-Hexose5_HexNAc5_Fucose1,
 히스티딘-프롤린-241번 아스파라긴-Hexose6_HexNAc5_Fucose1_NeuAc1,
 히스티딘-프롤린-241번 아스파라긴-Hexose6_HexNAc5,
 히스티딘-프롤린-241번 아스파라긴-Hexose7_HexNAc6_Fucose1 및
 히스티딘-프롤린-241번 아스파라긴-Hexose7_HexNAc6 중 선택된 1종 이상임을 특징으로 하는, 바이오마커 단백질의 N-당쇄변화 탐지방법.

청구항 5

청구항 1에 있어서,

상기 d) 단계의 이성질체 분리는 nanoLC chip/Q-TOF MS로 수행함을 특징으로 하는, 바이오마커 단백질의 N-당쇄변화 탐지방법.

청구항 6

청구항 1에 있어서,

상기 f) 단계의 정량적 프로파일링은 T 검정 p-값 분석, ROC 커브 (Receiver-Operating Characteristic curve) 및 AUC (Area under the ROC curve) 분석 중 한 가지 이상을 선택하여 수행함을 특징으로 하는, 바이오마커 단백질의 N-당쇄변화 탐지방법.

청구항 7

청구항 5에 있어서,

상기 f) 단계의 프로파일링 수행시 T 검정 p-값 분석 및 AUC (Area under the ROC curve) 분석에 N-당쇄화된 당펩타이드의 가지 (antennae) 양 (abundance) 변수를 부가함을 특징으로 하는, 바이오마커 단백질의 N-당쇄변화 탐지방법.

청구항 8

청구항 1에 있어서,

상기 c)의 당펩타이드 분리는 다공성 흑연화 탄소 카트리지를 (Porous Graphitized Carbon Cartridge)를 이용하

여 고체상 추출 (Solid Phase Extraction)함을 특징으로 하는, 바이오마커 단백질의 N-당쇄변화 탐지방법.

청구항 9

청구항 8에 있어서,

상기 다공성 흡연화 탄소 카트리지에 흡착된 펩타이드는 아세트나이트릴 및 트리플루오로아세트산을 함유한 용액으로 용출하여 당펩타이드만 선택적으로 분리함을 특징으로 하는, 바이오마커 단백질의 N-당쇄변화 탐지방법.

청구항 10

삭제

청구항 11

삭제

명세서

기술분야

[0001]

본 발명은 바이오마커로서 부위 특이적으로 N-당쇄화된 당펩타이드를 이용하여 바이오마커의 N-당쇄변화를 탐지하는 방법에 관한 것이다.

배경기술

[0002]

임상에서의 암 진단은 전형적으로 병력 조사, 물리적 검사 및 임상적 평가후 조직 생검을 수행하여 확인하고 있다. 그러나, 임상 실험에 의한 암 진단은 암 세포의 수가 10억 개 이상이고 암의 직경이 1 cm 이상일 경우에만 가능하다. 이 경우, 암 세포는 이미 전이능력을 가지고 있으며, 적어도 이들 중 반은 이미 전이된 상태이다. 또한, 암으로부터 직접 또는 간접적으로 생산되는 물질을 모니터링하기 위한 종양 마커가 암 스크리닝에 사용되고 있지만, 암이 존재하는 경우에도 종양 마커 스크리닝 결과 반 이상이 정상으로 나타나고, 암이 없는 경우에도 자주 양성으로 나타나기 때문에, 정확성에 한계가 있어 혼란을 초래하고 있다. 뿐만 아니라, 암 치료에 주로 사용되는 항암물질은 종양의 크기가 작을 때만 효과를 나타낸다는 문제가 있다.

[0003]

위암환자의 조기진단 및 진단을 위한 방법 중 가장 많이 이용되는 방법은 위내시경 검사와 위장 조영술이 있다. 이러한 검사로 위암이 진단되면 암이 위에만 있는지 혹은 주위의 다른 장기에 퍼졌는지를 확인하기 위해, 즉 위암의 병기를 진단하기 위해 컴퓨터 단층 촬영(CT)을 시행하며 때로는 초음파 내시경, 복부 초음파 검사를 시행하게 된다. 그 외에 추가로 시행할 수 있는 검사로 혈액 내 종양 표지자 검사 등이 있다. 이는 생체조직검사나 고가의 의료장비를 필요로 하므로 피검자에게 육체적, 경제적인 부담을 줄 수 있다.

[0004]

글라이코믹스는 글라이칸 구조 및 특성에 관한 광범위한 연구를 포함한다. 글라이칸은 단당류 구조, 가지뻗기(branching)의 정도 및 연결구조가 다양하기 때문에 구조가 매우 복잡하여 분석이 용이하지 않다. 게놈의 단선적인 DNA 주형과 직접적으로 관련되어 있는 단백질 구조와 달리, 글라이칸 구성은 일련의 글라이코실트랜스퍼레이스 효소반응에 의해 결정된다.

[0005]

N-당쇄화 (N-Glycosylation)는 펩타이드 골격의 아스파라긴에 올리고당이 부가되는 것이다. 이것은 그 위치와 구조적 다양성을 결정하는 생물학적 원칙에 의해 지배된다. N-당쇄화의 변화는 종양과 같은 질환에서 정상적인 당쇄화 장치에 오류가 생긴 것이다. 이러한 이유로, N-당쇄를 질병 마커로 사용하려는 시도가 있어왔다.

[0006]

N-당쇄는 아주 다양하지만, O-당쇄, 글라이코사미노글라이칸, 및 글라이코리피드와 같은 다른 글라이코코분류제이트와 구별되는 특징이 있다. N-당쇄는 아스파라긴 다음 다양한 아미노산이 오고 그 뒤에 세린, 트레오닌 또는

덜 일반적인 시스테인 아미노산이 오는 구조에서만 발견된다. 프롤린을 제외하고, 다양한 아미노산들이 아스파라긴 다음에 올 수 있다. 펩타이드 N-글라이코시테이즈 F (PNGase F)와 같은 효소가 생리학적 조건 하에서 N-연결 당쇄를 자르는데 이용될 수 있다. N-연결 코어는 키토바이오스 코어 (chitobiose core; Man₅GlcNAc₂)에 부착된 세 개의 만노스 당으로 이루어져 있다. N-연결 당쇄는 N-연결 코어에 어떤 단당이 부가되느냐에 따라 종종 세 개의 주요 타입으로 분류된다. 고만노스 당쇄 (High Mannose Glycan) 형은 N-연결코어에 만노스 단당이 여섯 개 더해진 구조이다. 복합형 (Complex type) 당쇄는 N-연결 코어에서 돌리콜 전구체가 분해된 후 N-아세틸글루코사민 (GlcNAc), 갈락토스, 디옥시헥소스 (퓨코스) 및 사이알릭산 단당들이 효소적으로 결합된다. 하이브리드형 당쇄는 고만노스 구조와 복합형 구조를 모두 포함한다.

[0007]

발암 및 전이 과정에서 일어나는 세포 생물학적 변화를 보면, 세포막 표면에 존재하는 많은 종류의 당단백질이나 당지질이 암유전자(oncogene)와 같은 특정 신호의 명령을 받아 "잘못된 당질화(aberrant glycosylation)"가 일어나고 이로 인한 당쇄(sugar chain)의 변화가 세포간의 접착(adhesion), 인식(recognition)에 변화를 일으켜 세포의 암화 및 암전이를 일으키게 된다(Hakomori and Kannagi, 1983, J. Natl. Cancer Inst., 71:231-251; Feizi, 1985, Nature, 314:53-571). 세포 외부에서 자극이 오면 암유전자 ras, 전사인자 ets-1의 신호전달을 거쳐 당전이효소 GnT-V(Nacetylglucosaminyltransferase)의 발현이 강화된다. 상기 GnT-V는 당단백질의 기본적인 당쇄에 N-아세틸글루코사민을 β1,6의 위치로 부가하는 반응을 촉매하는 효소로 암의 공격(invasion)과 전이(metastasis)에 직접 연관성이 있는 것으로 알려져 있다(Dennis et al., 1987, Science, 236:582-5853).

[0008]

일반적으로 당단백질은 단백질이 합성된 후 ER(endoplasmic reticulum)에서 기본적인 당쇄가 만들어진 후 골지체로 이동하여 세포의 다양한 생명현상의 결과로 여러 당전이효소에 의해 당의 부가가 이루어지는데 1차적으로 만들어지는 N-아세틸글루코사민 당전이효소 여섯가지(I-VI)가 촉매하여 얻어지는 당쇄들을 보여주고 있으며, 이중 β1,6-N-아세틸글루코사민 당쇄를 만드는 GnT-V가 암의 발생과 전이에 가장 많이 관여한다고 알려져 있다. GnT-V 효소는 골지체에 위치하며, 목표 단백질들은 당쇄에 변화를 받아 세포 표면이나 밖으로 분비된다. 이때 당단백질은 대상세포의 표면 단백질과 인식(recognition), 접착(adhesion)하여 암을 유발하게 된다.

[0009]

이와 같이, 최근 유전자와 단백질의 기능에 관한 학문인 지노믹스, 프로테오믹스에 이어 단백질의 표면에 존재하면서 단백질의 기능, 인식 등에 매우 중요한 역할을 수행하는 당에 관한 연구인 글라이코믹스 (Glycomics)가 과학자들의 관심사로 떠오르고 있다. 글라이코믹스 중에서도 특히 각종 종양세포에서 특정 단백질의 당쇄에 의미 있는 변화가 일어난다는 것이 밝혀지면서 많은 연구자들은 당쇄의 변화를 탐지하여 종양을 초기에 진단하는 바이오마커 발굴에 관한 연구가 활발하게 진행되고 있다. N-당쇄화의 변화는 특정 종양에서만 발견되는 것이 아니라, 다양한 종양을 비롯한 여러 질병에서 발견되어 왔다.

[0010]

혈액 등의 생체 시료에서 추출한 당(glycan)들을 질량분석기를 이용하여 프로파일하는 분석법은 글라이코믹스 기반 바이오마커 발굴 연구에서 널리 사용되고 있다.

[0011]

한국특허 제475642호는 암의 발생과 전이에 관여하여 N-연결형 당쇄의 변화를 나타내는 단백질의 당쇄변화를 측정하여 암을 진단하는 방법 및 상기 단백질의 당쇄변화를 측정하는 방법을 이용한 진단키트에 관한 것이다. 이 방법은 펩타이드의 변화를 측정함으로써 당펩타이드가 변했을 것으로 추정하는 것이다.

[0012]

미국특허 제7,651,847호(2010.01.26.특허)는 종양에 특이적인 올리고당을 확인하는 방법, 개인의 종양 계통(strain)을 결정하는 방법 및 특정 종양 마커의 존재 또는 부재를 탐지함으로써 개인의 종양을 진단 또는 종양 진행단계를 진단하는 방법에 관한 것이다.

[0013]

그러나, 상기 종래 기술들은 민감성과 특이성이 그다지 높지 않으며 진단방법이 복잡하다는 문제점이 있다.

[0014]

또한, M. Nakano et al.의 *Int. J. Cancer*:122, 2301-2309 (2008)는 췌장암 환자 혈청에서 헵토글로빈 N-당쇄의 부위 특이적 분석법에 대하여 개시한다. 본 발명이 비특이적 단백질 분해효소를 사용하여 단백질을 N-당쇄화된 당펩타이드로 만드는데 비하여 M. Nakano et al.의 이 논문은 트립신을 이용한다. 트립신을 이용하여 단백질을 분해하는 경우 207번과 211번 아스파라긴은 트립신 처리만으로는 서로 분리되지 않기 때문에 시알산을 제거한 후 GluC 엔도펩티데이즈로 한 번 더 처리해야만 207번 아스파라긴 포함 펩타이드와 211번 아스파라긴 포함 펩타이드가 분리된다. 그러므로, 당사들 정보 중 시알산에 대한 정보는 없어지므로 정확한 당쇄 분석이 불가능해진다. 또한, 이 논문에 개시된 바와 같이 트립신을 처리하면 최소 5개 이상의 아미노산으로 이루어진 펩타이

드가 얻어지지만, 본 발명의 경우 비특이적 단백질 분해효소를 사용하면 두 개 내지 네 개의 아미노산으로 이루어진 당펩타이드를 얻게 되어 좀 더 민감하게 당쇄 변이를 탐지할 수 있다. 뿐만 아니라, 이 논문은 췌장암 환자에서 퓨코스 증가 여부에 따른 진단 결과만을 제시하기 때문에 본 발명의 N-당사슬이 결합된 상태의 당펩타이드 분석방법과 같이 민감성과 특이성이 높지 않다.

발명의 내용

해결하려는 과제

[0015] 본 발명은 좀 더 민감하고 특이적으로 N-당쇄 변화를 간편하고 좀 더 빠르게 탐지하는 방법을 제공하려는 것을 목적으로 한다.

[0016] 또한, 본 발명은 좀 더 민감하고 특이적으로 이용 가능한 종양 바이오마커를 제공하는 것을 목적으로 한다.

과제의 해결 수단

[0017] 단백질 분해효소로 당단백질을 처리하여 당펩타이드로 만들면 하나의 당 부위에 하나의 당사슬이 결합된 형태의 당펩타이드를 얻게 된다. 이를 통하여 어떠한 아미노산 부위에 어떠한 당사슬이 결합되어 있는지 정성 및 정량 정보를 얻을 수 있다. 이 방법은 당사슬만 분리하여 실시하는 당사슬 바이오마커보다 특이성을 높임으로써 고특이성 고민감성 바이오마커 발굴에 유용할 것이다.

[0018] 따라서, 본 발명에서는 비특이적 단백질 분해효소를 이용하여 혈액의 헵토글로빈을 펩타이드와 당펩타이드로 분해하였으며, 207번과 211번 등 총 4개의 N-당쇄화 부위만을 순수하게 분리하였다. 본 발명자들은 이렇게 얻은 4개의 N-당쇄화 부위에 결합된 당펩타이드를 이용하여 위암 특이적 바이오마커를 발굴함으로써 고특이적이고 고감도 위암 진단 방법을 발명하기에 이르렀다.

[0019] 본 발명에서 사용한 비특이적 단백질 분해효소를 사용하지 않으면 헵토글로빈의 207번 아미노산과 211번 아미노산을 구별할 수 없으므로 본 발명자들이 발명한 방법은 암환자의 바이오마커 개발에서 강력한 분석도구로 활용될 수 있다.

[0020] 헵토글로빈의 4개의 N-당쇄화 위치를 구별하여 각 당쇄화 위치에 대한 당사슬의 정성 및 정량정보를 얻는 방법은 지금까지 없었던 새롭고 진보적인 기술이다.

[0021] 본 발명자들은 단백질 수준이나 당사슬만 분리하여 바이오마커를 찾는 경우보다 당펩타이드를 이용하여 특정 4개의 부위와 관련된 바이오마커를 발굴함으로써 고감도, 고특이성 바이오마커를 발굴하였으며, 당펩타이드를 이용한 바이오마커의 발굴은 본 발명이 세계 최초이다. 본 발명은 N-당쇄화 위치에 특이적으로 결합한 당쇄 구조 중 가지 (antennae)의 증가와 당쇄화 위치, 당쇄 합성과정에 관련된 당쇄의 조합을 이용한 고감도 진단 바이오마커에 관한 것이다.

[0022] 현재까지 FDA에서 승인된 바이오마커 중 위암관련 바이오마커는 없는 실정이며, 특히 본 발명을 통하여 발굴한 위암 당펩타이드 바이오마커는 AUC값 1을 보이는 고특이성 고민감성을 가진 강력한 바이오마커이다. 본 발명자들이 발굴한 바이오마커 리스트는 표 1에 나타내었으며, 이들은 p-값 0.005 이하의 값을 보이는 유의성이 높은 바이오마커들이다. 본 발명은 위암 환자에서 정상인에 비해 당쇄의 가지 (antennae)가 증가하는 것을 이용하였고, 또한 말단이 잘린 복합 타입 당쇄와 퓨코스가 결합된 복합 타입 당쇄가 증가하였으며, 생합성적으로 관련되는 당펩타이드의 조합이 증가하는 것을 이용하여 위암을 진단하는 방법을 제공한다.

[0023] 본 발명은 바이오마커로서 부위 특이적으로 N-당쇄화된 당펩타이드를 이용하여 바이오마커의 N-당쇄변화를 탐지하는 방법에 관한 것이다.

[0024] 또한, 본 발명은 상기 부위 특이적으로 N-당쇄화된 당펩타이드가 혈청 헵토글로빈에서 유래한 것임을 특징으로 한다.

- [0025] 또한, 본 발명은 상기 부위 특이적으로 N-당쇄화된 당펩타이드가 혈청 헤파토글로빈의 184번, 207번, 211번 및 241번 아스파라긴 부위가 당쇄화된 것임을 특징으로 한다.
- [0026] 또한, 본 발명은 상기 부위 특이적으로 N-당쇄화된 당펩타이드가 혈청 헤파토글로빈 유래 당펩타이드이며,
- [0027] 히스티딘-184번 아스파라긴-Hexose5_HexNac4_Fucose1,
- [0028] 히스티딘-184번 아스파라긴-Hexose6_HexNac5_Fucose1,
- [0029] 히스티딘-184번 아스파라긴-Hexose4_HexNac3,
- [0030] 히스티딘-184번 아스파라긴-Hexose5_HexNac4,
- [0031] 히스티딘-184번 아스파라긴-Hexose5_HexNac4_NeuAc1,
- [0032] Hexose4_HexNac3-207번 아스파라긴-히스티딘-세린-글루탐산,
- [0033] Hexose5_HexNac4-207번 아스파라긴-히스티딘-세린-글루탐산,
- [0034] Hexose4_HexNac4-207번 아스파라긴-히스티딘-세린-글루탐산,
- [0035] Hexose6_HexNac5_Fucose1-207번 아스파라긴-히스티딘-세린-글루탐산,
- [0036] Hexose6_HexNac5-207번 아스파라긴-히스티딘-세린-글루탐산,
- [0037] Hexose8_HexNac2-207번 아스파라긴-히스티딘-세린-글루탐산,
- [0038] Hexose6_HexNac5_Fucose1_NeuAc1-207번 아스파라긴-히스티딘-세린-글루탐산,
- [0039] Hexose5_HexNac4_Fucose1-207번 아스파라긴-히스티딘-세린-글루탐산,
- [0040] Hexose7_HexNac6-207번 아스파라긴-히스티딘-세린-글루탐산,
- [0041] Hexose6_HexNac5_Fucose1_NeuAc1-211번 아스파라긴-알라닌-트레오닌,
- [0042] Hexose6_HexNac5_Fucose1-211번 아스파라긴-알라닌-트레오닌,
- [0043] Hexose7_HexNac6_Fucose1-211번 아스파라긴-알라닌-트레오닌,
- [0044] Hexose5_HexNac5_NeuAc1-211번 아스파라긴-알라닌-트레오닌,
- [0045] Hexose4_HexNac4_Fucose1-211번 아스파라긴-알라닌-트레오닌,
- [0046] Hexose7_HexNac6_Fucose1_NeuAc1-211번 아스파라긴-알라닌-트레오닌,
- [0047] Hexose7_HexNac6_NeuAc1-211번 아스파라긴-알라닌-트레오닌,
- [0048] 히스티딘-프롤린-241번 아스파라긴-Hexose5_HexNac4,
- [0049] 히스티딘-프롤린-241번 아스파라긴-Hexose5_HexNac5_Fucose1,
- [0050] 히스티딘-프롤린-241번 아스파라긴-Hexose6_HexNac5_Fucose1_NeuAc1,
- [0051] 히스티딘-프롤린-241번 아스파라긴-Hexose6_HexNac5,
- [0052] 히스티딘-프롤린-241번 아스파라긴-Hexose7_HexNac6_Fucose1 및
- [0053] 히스티딘-프롤린-241번 아스파라긴-Hexose7_HexNac6 중 선택된 1종 이상임을 특징으로 한다.
- [0054] 또한, 본 발명은
- [0055] a) 시료에서 타겟 단백질을 분리 및 정제하는 단계;
- [0056] b) 분리한 타겟 단백질에 비특이적 단백질 분해효소를 처리하는 단계;
- [0057] c) 효소 처리 후 N-당쇄화된 당펩타이드를 분리하는 단계;
- [0058] d) 분리한 N-당쇄화된 당펩타이드를 질량분석하여 이성질체를 분리하는 단계;
- [0059] e) 분리한 이성질체의 구조 분석 및 당쇄와 펩타이드 구성을 확인하는 단계; 및

- [0060] f) 분리한 이성질체의 정량적 프로파일링을 수행하는 단계;를 포함하는 바이오마커 단백질의 N-당쇄변화 탐지방법에 관한 것이다.
- [0061] 상기 e)단계의 분리한 이성질체의 구조 분석 및 당쇄와 펩타이드 구성을 확인하는 방법 중 대표적인 방법은 PGC 컬럼을 이용하여 질량값은 같으나 구조가 다른 당펩타이드 이성질체들을 컬럼 내 머무름 시간이 다를을 이용하여 분리한 후 각각에 대하여 텐덤 질량분석법을 이용하여 구조분석을 수행하였다. 이때 펩타이드와 당쇄의 조성은 같으나 구조가 다른 당펩타이드의 구조를 확인하였다. 텐덤 질량분석시 얻어지는 산물 이온 (product ion) 중 당펩타이드에서 떨어진 당쇄 이온과 펩타이드에 당쇄 파편이 결합된 당펩타이드 이온을 확인하여 당쇄와 펩타이드의 구성을 확인하였다.
- [0062] 또한, 본 발명은 상기 f) 단계의 정량적 프로파일링이 T 검정 p-값 분석, ROC 커브 (Receiver-Operating Characteristic curve) 및 AUC (Area under the ROC curve) 분석임을 특징으로 한다.
- [0063] 또한, 본 발명은 상기 타겟 단백질이 헵토글로빈임을 특징으로 한다.
- [0064] 또한, 본 발명은 상기 d) 단계의 분리한 N-당쇄화된 당펩타이드가
- [0065] 히스티딘-184번 아스파라긴-Hexose5_HexNAc4_Fucose1,
- [0066] 히스티딘-184번 아스파라긴-Hexose6_HexNAc5_Fucose1,
- [0067] 히스티딘-184번 아스파라긴-Hexose4_HexNAc3,
- [0068] 히스티딘-184번 아스파라긴-Hexose5_HexNAc4,
- [0069] 히스티딘-184번 아스파라긴-Hexose5_HexNAc4_NeuAc1,
- [0070] Hexose4_HexNAc3-207번 아스파라긴-히스티딘-세린-글루탐산,
- [0071] Hexose5_HexNAc4-207번 아스파라긴-히스티딘-세린-글루탐산,
- [0072] Hexose4_HexNAc4-207번 아스파라긴-히스티딘-세린-글루탐산,
- [0073] Hexose6_HexNAc5_Fucose1-207번 아스파라긴-히스티딘-세린-글루탐산,
- [0074] Hexose6_HexNAc5-207번 아스파라긴-히스티딘-세린-글루탐산,
- [0075] Hexose8_HexNAc2-207번 아스파라긴-히스티딘-세린-글루탐산,
- [0076] Hexose6_HexNAc5_Fucose1_NeuAc1-207번 아스파라긴-히스티딘-세린-글루탐산,
- [0077] Hexose5_HexNAc4_Fucose1-207번 아스파라긴-히스티딘-세린-글루탐산,
- [0078] Hexose7_HexNAc6-207번 아스파라긴-히스티딘-세린-글루탐산,
- [0079] Hexose6_HexNAc5_Fucose1_NeuAc1-211번 아스파라긴-알라닌-트레오닌,
- [0080] Hexose6_HexNAc5_Fucose1-211번 아스파라긴-알라닌-트레오닌,
- [0081] Hexose7_HexNAc6_Fucose1-211번 아스파라긴-알라닌-트레오닌,
- [0082] Hexose5_HexNAc5_NeuAc1-211번 아스파라긴-알라닌-트레오닌,
- [0083] Hexose4_HexNAc4_Fucose1-211번 아스파라긴-알라닌-트레오닌,
- [0084] Hexose7_HexNAc6_Fucose1_NeuAc1-211번 아스파라긴-알라닌-트레오닌,
- [0085] Hexose7_HexNAc6_NeuAc1-211번 아스파라긴-알라닌-트레오닌,
- [0086] 히스티딘-프롤린-241번 아스파라긴-Hexose5_HexNAc4,
- [0087] 히스티딘-프롤린-241번 아스파라긴-Hexose5_HexNAc5_Fucose1,
- [0088] 히스티딘-프롤린-241번 아스파라긴-Hexose6_HexNAc5_Fucose1_NeuAc1,
- [0089] 히스티딘-프롤린-241번 아스파라긴-Hexose6_HexNAc5,
- [0090] 히스티딘-프롤린-241번 아스파라긴-Hexose7_HexNAc6_Fucose1 및

- [0091] 히스티딘-프롤린-241번 아스파라긴-Hexose7_HexNac6 중 선택된 1종 이상임을 특징으로 한다.
- [0092] 또한, 본 발명은 상기 f) 단계의 프로파일링 수행시 T 검정 p-값 분석 및 AUC (Area under the ROC curve) 분석에 N-당쇄화된 당펩타이드의 가지 (antennae) 양 (abundance) 변수를 부가함을 특징으로 하는, 바이오마커 단백질의 N-당쇄변화 탐지방법에 관한 것이다.
- [0093] 또한, 본 발명은 상기 c) 단계의 당펩타이드 분리를 다공성 흑연화 탄소 카트리지 (Porous Graphitized Carbon Cartridge)를 이용하여 고체상 추출 (Solid Phase Extraction)함을 특징으로 하는, 바이오마커 단백질의 N-당쇄변화 탐지방법에 관한 것이다.
- [0094] 또한, 본 발명은 상기 다공성 흑연화 탄소 카트리지에 흡착된 펩타이드를 아세트로나이트릴 및 트리플루오로아세트산을 함유한 용액으로 용출하여 당펩타이드만 선택적으로 분리함을 특징으로 하는, 바이오마커 단백질의 N-당쇄변화 탐지방법에 관한 것이다.
- [0095] 또한, 본 발명은 헵타글로빈의 184번, 207번, 211번 또는 241번 아스파라긴 부위가 N-당쇄화된 위암 진단용 당펩타이드 바이오마커에 관한 것이다.
- [0096] 또한, 본 발명은 상기 당펩타이드 바이오마커가
- [0097] 히스티딘-184번 아스파라긴-Hexose5_HexNac4_Fucose1,
- [0098] 히스티딘-184번 아스파라긴-Hexose6_HexNac5_Fucose1,
- [0099] 히스티딘-184번 아스파라긴-Hexose4_HexNac3,
- [0100] 히스티딘-184번 아스파라긴-Hexose5_HexNac4,
- [0101] 히스티딘-184번 아스파라긴-Hexose5_HexNac4_NeuAc1,
- [0102] Hexose4_HexNac3-207번 아스파라긴-히스티딘-세린-글루탐산,
- [0103] Hexose5_HexNac4-207번 아스파라긴-히스티딘-세린-글루탐산,
- [0104] Hexose4_HexNac4-207번 아스파라긴-히스티딘-세린-글루탐산,
- [0105] Hexose6_HexNac5_Fucose1-207번 아스파라긴-히스티딘-세린-글루탐산,
- [0106] Hexose6_HexNac5-207번 아스파라긴-히스티딘-세린-글루탐산,
- [0107] Hexose8_HexNac2-207번 아스파라긴-히스티딘-세린-글루탐산,
- [0108] Hexose6_HexNac5_Fucose1_NeuAc1-207번 아스파라긴-히스티딘-세린-글루탐산,
- [0109] Hexose5_HexNac4_Fucose1-207번 아스파라긴-히스티딘-세린-글루탐산,
- [0110] Hexose7_HexNac6-207번 아스파라긴-히스티딘-세린-글루탐산,
- [0111] Hexose6_HexNac5_Fucose1_NeuAc1-211번 아스파라긴-알라닌-트레오닌,
- [0112] Hexose6_HexNac5_Fucose1-211번 아스파라긴-알라닌-트레오닌,
- [0113] Hexose7_HexNac6_Fucose1-211번 아스파라긴-알라닌-트레오닌,
- [0114] Hexose5_HexNac5_NeuAc1-211번 아스파라긴-알라닌-트레오닌,
- [0115] Hexose4_HexNac4_Fucose1-211번 아스파라긴-알라닌-트레오닌,
- [0116] Hexose7_HexNac6_Fucose1_NeuAc1-211번 아스파라긴-알라닌-트레오닌,
- [0117] Hexose7_HexNac6_NeuAc1-211번 아스파라긴-알라닌-트레오닌,
- [0118] 히스티딘-프롤린-241번 아스파라긴-Hexose5_HexNac4,
- [0119] 히스티딘-프롤린-241번 아스파라긴-Hexose5_HexNac5_Fucose1,
- [0120] 히스티딘-프롤린-241번 아스파라긴-Hexose6_HexNac5_Fucose1_NeuAc1,
- [0121] 히스티딘-프롤린-241번 아스파라긴-Hexose6_HexNac5,

- [0122] 히스티딘-프롤린-241번 아스파라긴-Hexose7_HexNAc6_Fucose1 및
- [0123] 히스티딘-프롤린-241번 아스파라긴-Hexose7_HexNAc6 중 선택된 1종 이상임을 특징으로 한다.
- [0124] 상기 방법에서, 시료로는 혈액뿐만 아니라, 혈장, 혈청, 타액, 소변 등의 체액을 이용하는 것이 바람직하다.
- [0125] 또한, 상기 방법에서, a) 단계의 단백질은 크기에 특별한 제한이 없으며, 올리고펩타이드, 폴리펩타이드 또는 단백질 모두 가능하다.
- [0126] 또한, 상기 a) 단계의 타겟 단백질 분리 및 정제는 1차원 겔 단백질 분리 (1D-gel protein separation), 2D-PAGE, 크기 배제 크로마토그래피 (Size Exclusion Chromatography), 자유흐름 전기영동시스템 (Free Flow Electrophoresis System) 등을 이용하여 수행하는 것이 바람직하나 반드시 이에 한정되는 것은 아니다.
- [0127] 또한, 본 발명의 방법에서 상기 b) 단계의 비특이적 단백질 분해효소는 특정 펩타이드 결합에만 작용하는 특이적 단백질 분해효소가 아닌 단백질 분해효소를 말하며, 예컨대 돼지 엘라스테이스 (porcine elastase), 파파인 (papain), 돼지 펩신 (porcine pepsin), 프로네이즈 (pronase), 프로테이네이즈 K (proteinase K), 섭틸리신 (subtilisin), 썬모라이신 (thermolysin) 등이 있으나, 반드시 이에 한정되는 것은 아니다.

발명의 효과

- [0128] 본 발명의 방법에 따르면 종래의 방법과 같이 당쇄의 당만을 분석하는 것보다 당펩타이드를 분석하는 것이 더 특이적이며, 민감도가 높은 결과를 얻을 수 있다.
- [0129] 또한, 본 발명의 방법에 따르면, 특정 단백질에서 부위 특이적으로 N-당쇄화 변화가 일어나기 때문에 AUC 값을 높여주어 중앙 환자와 정상인을 확연하게 구분할 수 있다.

도면의 간단한 설명

- [0130] 도 1은 혈청 헵토글로빈의 부위-특이적 N-당쇄화 맵을 나타낸다.
- 녹색 원은 만노스, 노란색 원은 갈락토스, 청색 네모는 N-아세틸글루코사민, 적색 마름모는 시알산, 적색 세모는 퓨코스를 나타낸다.
- 도 2는 같은 글라이칸 모이어터를 갖는 세 개의 당펩타이드의 크로마토그램이다. 초기암은 핑크색, 말기암은 적색, 정상은 청색으로 표시하였다.
- 도 3은 암환자에서 특이적으로 글라이칸의 가지 (antennae) 수가 증가하는 경향성을 보여주는 결과이다.
- 도 4는 글라이칸 생합성 경로에 따라 구분한 경우 AUC가 증가하는 결과를 얻을 수 있음을 보여준다.
- 도 5는 위암 바이오마커로 이용 가능한 당펩타이드를 보여준다.

발명을 실시하기 위한 구체적인 내용

- [0131] 아래에서는 실시예를 들어 본 발명의 구성을 좀 더 자세히 설명한다. 그러나, 본 발명의 범위가 실시예의 기재에 한정되는 것이 아님은 본 발명이 속하는 기술분야에서 통상의 지식을 가진 자에게 자명하다.

[0132] 실시예

- [0133] 1. 정상인 15명, 초기 위암환자 5명, 말기 위암환자 5명의 혈청에서 헵토글로빈을 선택적으로 분리하였다.
- [0134] 2. 분리된 헵토글로빈 50 μ g은 50 μ l PBS (pH 7.5)에 충분히 녹인 후 비특이적 단백질 분해효소 50 μ g (1:1)를 첨가하여 37℃에서 6시간 동안 반응시켜 당펩타이드를 얻었다.

[0135] 3. 얻어진 당펩타이드 (glycopeptide)는 다공성 흑연화 탄소 카트리지 (Porous Graphitized Carbon Cartridge)를 이용하여 고체상 추출 (Solid Phase Extraction)하였다. 좀 더 자세히는, 흑연화 탄소 카트리지를 80% 아세트나이트릴 및 0.1% TFA (trifluoroacetic acid) 수용액으로 세척한 후 순수로 행구었다. 당펩타이드 수용액을 카트리지에 로딩한 후 염과 완충액 제거를 위해 약 1 mL/min의 속도로 순수로 워싱하였다. 당펩타이드는 40%(v/v) ACN (acrylonitrile) 및 0.1%(v/v) TFA (trifluoroacetic acid)를 함유한 용액으로 용출하여 당펩타이드만 선택적으로 분리하였다.

[0136] 4. nanoLC chip/Q-TOF MS 분석을 통하여 이성질체를 분리하고, 구조를 분석하였고, 당쇄 및 펩타이드 구성을 확인하고, 정량적인 프로파일링을 수행하였다.

[0137] 분리된 당펩타이드를 물에 재용해하여 오토샘플러 (4 °C로 유지됨), 모세관 펌프, 나노펌프, HPLC-Chip/MS 인터페이스 및 6530 Q-TOF MS 탐지기를 갖춘 HPLC-Chip/Q-TOF(Chip Quadrupole Time-of-Flight) MS system (Agilent Technologies, Santa Clara, CA)을 이용하여 분석하였다. 9 × 0.75 mm 농축용 컬럼 및 43 × 0.075 mm 분석용 컬럼의 정지상 (stationary phase)으로 통합된 나노-ESI 스프레이 팁과 함께 5 μm 다공성 흑연화 탄소 칩을 패키징하여 사용하였다. 크로마토그래피 방법에 의한 분리는 종래 최적화된 방법에 따라 수행하였다. 간단히 설명하면, 시료를 로딩한 후, 빠른 당펩타이드 용출 농도구배 용액을 분당 0.3 μL씩 흘려주었다. 농도구배 용액은 (A) 3.0% 아세트나이트릴 및 0.5% 포름산(v/v) 수용액, 및 (B) 90.0% 아세트나이트릴 및 0.5% 포름산(v/v) 수용액이고, 각각 아래 시각에 아래 농도로 처리하였다: 6% B, 0~2.5min; 6~16% B, 2.5~5min; 16~44% B, 5~15min; 및 44~99.9% B, 15~20min. 남아있는 당펩타이드는 99.9% B 용액으로 10분간 씻어내었다. 최종적으로 분석용 컬럼은 6% B로 20분간 재평형화하고, 농축용 컬럼은 0% B로 20분간 재평형화하였다.

[0138] 건조가스 온도는 325 °C로 조정하고 흐름 속도는 4 L/min(여과한 질소가스 2 L와 여과한 건조 압축공기 2 L)로 하였다.

[0139] MS 스펙트럼은 스펙트럼당 1.5초의 획득시간 (acquisition time)으로 m/z 500-2000의 질량 범위로 양성 이온화 모드에서 얻었다. MS/MS 스펙트럼은 스펙트럼당 1.5초의 획득시간으로 m/z 100-3000 질량 범위로 양성 이온화 모드에서 얻었다. MS 스캔 후 전구체 화합물들은 이온의 양과 대전 상태 (z = 2, 3 또는 4)를 기초로 획득 소프트웨어에 의해 자동으로 MS/MS 분석을 위해 선택되었고 질량 밴드패스 fwhm (full width at half-maximum) 1.3m/z로 4극자 (quadrupole)에서 분리되었다. 통상 CID 단편화의 충돌 에너지는 다음의 식을 기초로 전구체 화합물에 대해 계산하였다:

수학식 1

$$V_{\text{collision}} = 3.6 V \left(\frac{m/z}{100 \text{ Da}} \right) - 4.8 V$$

[0140]

[0141] 위 식에서, $V_{\text{collision}}$ 은 전 충돌 세포에서의 전위차이다. 기울기와 에너지 m/z 램프의 오프셋 값은 필요에 따라 다소의 단편화 (fragmentation)를 제조하기 위하여 변화시킬 수 있다.

[0142] 데이터를 얻은 후, LC/MS의 원래 데이터는 Mass Hunter Qualitative Analysis software (version B.04.00 SP2, Agilent Technologies)에 포함된 분자 특성 추출 알고리즘 (Molecular Feature Extractor algorithm)을 이용하여 처리되었다. MS 피크는 5.0의 신호 대 노이즈 비율로 여과되었고, 개별 이온 종으로 분석되었다.

[0143] 예상 동위원소 분포 (expected isotopic distribution)를 이용하여 전위상태정보 및 머무름 시간, 단일 화합물과 연관된 모든 이온종 {예컨대, 이중 양성자화된 이온, 삼중 양성자화된 이온 및 모든 관련 동위 이성질체 (isotopologue)}을 함께 합산하여 화합물의 중성질량 (neutral mass)을 계산하였다.

[0144] 이 정보를 이용하여, 크로마토그래피 피크 면적에 의해 나타나는 양으로 시료 내 모든 화합물 피크의 리스트를 생성하였다. 당펩타이드들은 GlycoX 알고리즘에 기반한 본 발명자들의 소프트웨어 툴로 확인하였다. 가능한 당펩타이드의 질량을 고려하여 당쇄화 타입 (N 또는 O) 및 질량 내성 (mass tolerance) (여기서는 5 ppm), 주어진 화합물과 연관된 당쇄 구조와 가능한 아미노산 리스트를 생성하는 알고리즘이 단백질 또는 단백질들의 아미노산

서열들로부터 유래하였다. 결과는 알려진 당쇄화 패턴 및/또는 다른 생물학적 법칙들에 기반하여 좀 더 여과되었다.

[0145] 5. 당쇄타이드 분석 결과를 이용하여 T 검정 p-값 분석, ROC 커브 (Receiver-Operating Characteristic curve), AUC (Area under the ROC curve) 분석을 통하여 위암환자의 진단을 위한 고감도, 고특이적 바이오마커를 발굴하였다.

[0146] 진단 방법의 효율성을 판단하는 방법 중 널리 사용되는 것이 ROC 커브이다. 민감도 (sensitivity)와 특이도 (specificity)가 어떤 관계를 갖고 변하는지를 이차원 평면상에 표현한 것이 ROC 커브인데, ROC 커브 아래의 면적 즉, AUC (area under ROC curve)가 넓을수록 좋은 진단 방법이라 할 수 있다.

[0147] ROC 커브 (Receiver-Operating Characteristic curve)는 특정 진단 방법의 민감도와 특이도가 어떤 관계를 갖고 있는지를 표현한 그래프이다. 진단 방법에 의한 환자의 분류는, 점수에 따라 사람을 환자/정상인으로 분류하는 것으로 결국 분류 (classification)의 한 예이다. 민감도는 진짜 환자 중 검사 방법이 환자를 얼마나 잘 골라내는가, 하는 것이고, 특이도는 진짜 정상 중 검사 방법이 정상을 얼마나 잘 골라내는가를 의미한다. 예를 들어 혈압으로 심근경색을 진단할 수 있다고 가정할 때, '혈압이 얼마 이상이면 심근경색이다'라는 진단은 민감도 100%, 특이도 100%를 갖게 혈압의 기준을 설정할 수 있다. 가령, '혈압이 5 이상이면 심근경색' 이라면 전체 인구가 모두 환자로 판별되어, 누구나 환자로 판정이 나고, 따라서 환자를 100% 환자라고 진단할 수 있다. 문제는, 환자가 아닌 사람도 몽땅 환자라고 하기 때문에 민감도가 거의 0에 가깝게 떨어진다는 것. 또는 반대로 혈압이 '1000 이상이면 심근경색' 이라면 모두 정상이라 판별되기 때문에 정상을 100% 골라낼 수 있기 때문에 특이도가 1 이 되지만 환자는 하나도 못 골라내기 때문에 민감도가 0 이 된다. 이렇게, 모든 분류 방법은 민감도와 특이도를 동시에 보아야 좋은 진단 방법인지 확인할 수 있다. 현재 적당한 기준은 민감도/특이도 모두 0.8에서 0.9 정도는 되어야 임상에서 사용할 수 있다.

[0148] ROC 커브를 이용하여 만든 영역 즉, AUC (area under the ROC curve)는 측정하고자 하는 이분형 변수를 판별하는데 얼마만큼의 정확도가 있는지를 평가하는데 쓰인다. 일반적으로 AUC 의 수치를 해석하는 방법은 다음과 같다.

[0149] $AUC=0.5$; 전혀 구분에 도움이 되지 못함.

[0150] $0.7 \leq AUC < 0.8$; 구분을 받아들일 만함.

[0151] $0.8 \leq AUC < 0.9$; 구분하는 능력이 뛰어남.

[0152] $AUC \geq 0.9$; 거의 완벽한 구분 능력.

[0153] p-값 유의확률

[0154] 통계분석에서 가장 많이 나오는 중요한 개념 중의 하나로, 일반적으로 어느 집단 간의 통계적 유의성(A와 B가 서로 차이가 있는가? 아닌가?)을 검정할 때 F 검정이나, t 검정을 사용하며 그 값으로 p-값을 얻게 된다.

[0155] t 검정은 표본 수가 아주 작은(보통 30 이하) 두 실험구간의 평균을 비교할 때 사용하고, F 검정은 그 이상의 경우에 활용한다. 이때 t 검정, F 검정의 결과로 나타나는 p-값이 내가 제시한 위험율 (α)보다 낮으면 비교한 집단 간에 유의차이가 인정된다고 말할 수 있다. 즉 통계처리의 신뢰도를 의미한다. p-값은 값이 낮아질수록 신뢰도가 좋은 것으로 판단하며 $P < 0.01$ 에서의 값이 가장 좋은 것이겠지만, 이 신뢰구간에서 통계처리가 무의미한 경우엔 $p < 0.05$ 로 표기를 하기도 하며 $p < 0.05$ 에서도 통계처리의 의미가 없게 나타나면 이는 데이터로 사용할 수 없다는 것을 의미한다.

[0156] 도 1은 햅토글로빈의 N-당쇄화 부위-특이적 당쇄 지도를 나타낸다. 그림에 표기된 당쇄의 크기는 LC/MS 분석을 통하여 얻어진 강도를 표준화한 값을 이용하여 표시한다. 즉, 가장 높은 강도를 보이는 값을 100%로 환산하여 상대적인 값을 비교하였다. 50~100%는 큰 크기; 10~50%, 중간 크기; 1~10%, 작은 크기로 나타내고, 1% 이하의 값을 보이는 당쇄는 나타내지 않았다. 당쇄 구조 위에 표시된 별표는 정상인에서는 확인되지 않았지만 위암 환

자에서는 1% 이상 존재하는 당쇄로서, 위암환자 진단용 바이오마커로 사용할 수 있다.

[0157] 도 2는 위암 초기와 말기 환자, 정상인의 헵토글로빈에서 부위-특이적인 바이오마커를 나타낸다. 그림에서 보여주는 크로마토그램은 같은 당쇄 구조 (Hex7 HexNAc6 Fucose1)를 가지고, 세 개의 다른 N-당쇄화 부위에 결합되어 있는 당펩타이드의 정량분석 정보를 나타낸다. 184번에는 위와 같은 당쇄는 결합되어있지 않았다. 각 세 개의 다른 당펩타이드는 다른 정량정보를 가지고 있으며, T-검정 p-값, ROC 커브, AUCs (Area under the ROC curve)에서 각기 다른 값을 보여주고 있다. 세 개의 당펩타이드 모두 정상인과 암환자를 명확히 구별할 수 있으며, 211번 당펩타이드의 AUC=0.92, 207번 당펩타이드는 AUC=0.94로 높은 AUC 값을 보여주며, 특히 241번 당펩타이드는 AUC=1의 값을 보여 정상인과 암환자를 완벽히 구별할 수 있는 강력한 바이오마커로 보인다. 즉, 같은 당쇄 구조를 가지고 있더라도 어느 당쇄화 위치에 어떠한 당이 결합되어 있는지에 따라 위암환자와 정상인을 완벽하게 구별할 수 있는 바이오마커로 사용 가능하다. 부위-특이적인 당펩타이드를 이용한 바이오마커는 고감도 고특이성을 가진 바이오마커이다.

[0158] 도 3은 암환자에서 특이적으로 당펩타이드에 결합된 당쇄의 가지 (antennae) 수가 증가하는 경향성을 보여주는 결과이다. 헵토글로빈은 정상인에서는 주로 두 가지 (bi-antennae)의 구조를 가지지만, 암환자에서는 한 가지 (mono-antennae), 세 가지 (tri-antennae), 네 가지 (tetra-antennae)의 양이 증가하는 경향을 보여준다.

[0159] 당펩타이드의 정량정보는 LC/MS 분석시 얻은 피크 크로마토그램의 면적을 이용하여 정량값을 얻었다. N-당쇄의 기본 구조인 만노스 (Man)3, 엔-아세틸-글루코사민 (GlcNAc)2()에서 GlcNAc이 증가하면서 당쇄 구조의 가지가 증가하는 경향을 보인다. 한 가지 (mono-antennae) 구조는 Man3, GlcNAc3 구조()에 푸코스, 헥소스, 시알산이 결합된 형태인 당사슬의 구조를 의미한다. 두 가지 (bi-antennae) 구조는 Man3, GlcNAc4 구조()에 푸코스, 헥소스, 시알산이 결합된 형태인 당사슬의 구조를 의미한다. 세 가지 (tri-antennae) 구조는 Man3, GlcNAc5 구조()에 푸코스, 헥소스, 시알산이 결합된 형태인 당사슬의 구조를 의미한다. 또한, 네 가지 (tetra-antennae) 구조는 Man3, GlcNAc6 구조()에 푸코스, 헥소스, 시알산이 결합된 형태인 당사슬의 구조를 의미한다. 이와 같이, 각 특정 당쇄화 위치에 결합된 당쇄 구조에 따라 한 가지, 두 가지, 세 가지, 네 가지 구조로 구분하여 분석을 수행한다.

[0160] 찾아진 당펩타이드에서 모든 당쇄화 위치에 결합된 당쇄 가지 (antennae) 구조에 따라 당펩타이드를 구분하여 가지의 양 (abundance)을 합산한다. 즉, LC/MS 분석시 얻은 피크 크로마토그램의 면적을 이용하여 당펩타이드의 정량정보를 얻고, p-값과 AUC 값을 얻었다. 이와 같이, 각 당쇄화 위치별로 결합된 N-당쇄의 가지 (antennae)의 양을 계산에 넣은 경우 유의성이 높고 고특이성, 고민감도를 나타내어 당쇄 변화를 탐지하기에 효과적이다.

[0161] 각 특정 부위에 결합된 당사슬의 가지 구조를 분석함으로써 고감도 고특이성의 바이오마커를 발굴하였다. 이러한 경향성은 부위-특이적인 분석에서 유의성을 가진다. 예컨대, p-값은 모든 부위를 통합하여 분석하였을 때보다 하나의 부위만을 이용하여 분석하였을 경우 더 낮은 값을 나타낸다. 유사한 결과로, AUC 또한 하나의 부위를 분석한 경우 더 높은 값을 보였다. 즉 부위-특이적으로 분석한 경우 암환자 진단을 위한 유의성 및 민감도와 특이성이 증가되는 결과를 보였다.

[0162] 예컨대 모든 N-당쇄화 위치에서 한 가지 (mono-antennae)의 양이 증가되는 경향은 p-값=9.50x10⁻⁵ 와 AUC=0.9067의 값을 보였으나, 184 당쇄화 위치에서 한 가지 (mono-antennae)가 증가되는 경향은 p-값=1.81x10⁻⁵ 와 AUC=0.9533의 값을 보여 더 높은 유의성과 고감도 고특이성을 나타내었다.

[0163] 또한, 모든 N-당쇄화 위치에서 네 가지 (tetra-antennae)의 양이 증가되는 경향은 p-값= 3.64x10⁻⁵ 와 AUC=0.9400의 값을 보였으나 241 당쇄화 위치에서 네 가지 (tetra-antennae)의 증가는 p-값=2.32x10⁻⁶ 와 AUC=0.9667로 나타나 더 높은 유의성과 고감도 고특이성을 보여주었다.

[0164] 도 4는 당쇄의 생물학적 합성경로에 따라 구별하였을 경우 AUC 가 증가되는 결과를 보여준다. 고-만노스 (high-

mannose) 타입, 복합 (complex) 타입, 퓨코스가 결합된 복합 타입, 시알산이 결합된 복합 타입 당쇄로 나누어 바이오마커 검증분석을 수행하였다. 말단이 잘린 복합 타입 (truncated complex type)과 퓨코스가 결합된 복합 타입에서 유의성이 높은 고감도의 바이오마커를 발굴하였다. 두 당쇄 타입에서 보면 정상인에 비해 암환자에서 두 당쇄가 확연히 증가한 것을 확인할 수 있으며, 각 당쇄화 부위에서 모두 특이적으로 양이 증가하였다.

[0165] 또한, 이러한 경향성은 부위-특이적으로 분석할 경우 유의성이 커진다. p-값은 모든 부위를 통합하여 분석하였을 때보다 하나의 부위만을 이용하여 분석하였을 경우 더 낮은 값을 나타내었다. 유사한 결과로, AUC 또한 하나의 부위를 분석한 경우 더 높은 값을 보였다. 즉, 부위-특이적으로 분석한 경우 암환자 진단을 위한 유의성, 민감도와 특이성이 증가되는 결과를 보였다.

[0166] 말단이 잘린 복합 타입 당쇄가 결합된 당펩타이드의 증가를 이용하여 암환자와 정상인을 구별할 수 있으며, 이는 고감도 고평가적 바이오마커로 사용가능하다. 특히 부위-특이적으로 분석할 경우 감도와 특이성은 더욱 증가한다.

[0167] 모든 부위의 말단이 잘린 복합 타입 당쇄가 결합된 당펩타이드의 양적 증가를 이용한 경우 $p\text{-값}=5.52 \times 10^{-5}$ 와 $AUC=0.9533$ 값을 보였으나, 부위 207의 복합 타입 당사슬이 결합된 당펩타이드의 증가를 이용한 경우 $p\text{-값}=8.78 \times 10^{-6}$ 와 이의 AUC 값은 0.973으로 고감도, 고평가적인 바이오마커로 사용 가능하다.

[0168] 퓨코실화 복합 타입 (fucosylated complex type) 당쇄가 결합된 당펩타이드의 양적 증가를 이용하여 정상인과 암환자를 구별할 수 있으며 이는 고감도 고평가적 바이오마커로 사용가능하다. 특히 부위-특이적으로 분석할 경우 유의성은 더욱 높아진다.

[0169] 모든 부위의 퓨코실화 복합 타입 당쇄가 결합된 당펩타이드의 증가를 이용하여 p-값과 AUC 값을 분석한 경우 각각 2.27×10^{-6} , 1의 값을 보여 고감도 고평가적인 바이오마커로 사용이 가능하며, 부위-특이적 당펩타이드를 분석한 경우 그 유의성은 더욱 증가되었다.

[0170] 부위 211의 퓨코실화 복합 타입 당쇄가 결합된 당펩타이드 증가는 암환자와 정상인을 완벽하게 구별할 수 있는 $AUC=1$, $p\text{-값}$ 2.21×10^{-7} 의 값을 보였으며, 따라서 이 당사슬은 고감도, 고평가적 바이오마커로 이용 가능하다.

[0171] 도 5는 바이오마커로 활용가능한 당펩타이드를 보여주고 있다. 도 5에서 보여주는 모든 당펩타이드는 정량적인 양이 암환자와 정상인에서 5배 이상 차이가 난다.

[0172] 별표로 표시된 모든 당펩타이드는 통계적인 유의성을 가지며 이들의 생합성적으로 관련되는 당펩타이드의 조합을 이용하여 분석한 결과 유의성은 더 증가했다.

[0173] 예를 들어, 211번 아미노산 N-당쇄화 위치에서 (Hex6 HexNAc5 Fucose1)을 가진 당펩타이드의 p-값은 6.70×10^{-5} , (Hex6 HexNAc5 Fucose1 NeuAc1)을 가진 당펩타이드의 p-값은 4.91×10^{-6} 인데, 이 두 당펩타이드의 조합을 이용하면 p-값은 7.66×10^{-7} , AUC는 1의 값을 보여 정상인과 암환자 구별을 완벽하게 할 수 있는 바이오마커가 된다.

[0174] 241번 아미노산 N-당쇄화 위치에 (Hex6 HexNAc5 Fucose1 NeuAc1)과 (Hex7 HexNAc6 Fucose1)을 가진 두 개의 당펩타이드는 p-값이 각각 1.03×10^{-4} , 2.03×10^{-4} 인데, 이 두 당펩타이드의 조합을 이용하면, p-값은 9.41×10^{-5} , $AUC=1$ 의 값을 가진 바이오마커가 된다. 이와 같이, 부위 특이적 당펩타이드의 조합을 이용하면 고감도, 고평가적의 바이오마커를 발굴할 수 있다.

[0175] 표 1은 p-값 0.005 이하의 고감도 고평가성을 가진 바이오마커의 리스트이다. 이와 같이, 본 발명자들은 특정 N-당쇄화 위치에 결합한 당사슬의 정성 및 정량정보를 이용한 바이오마커 발굴 및 진단법을 개발하였다.

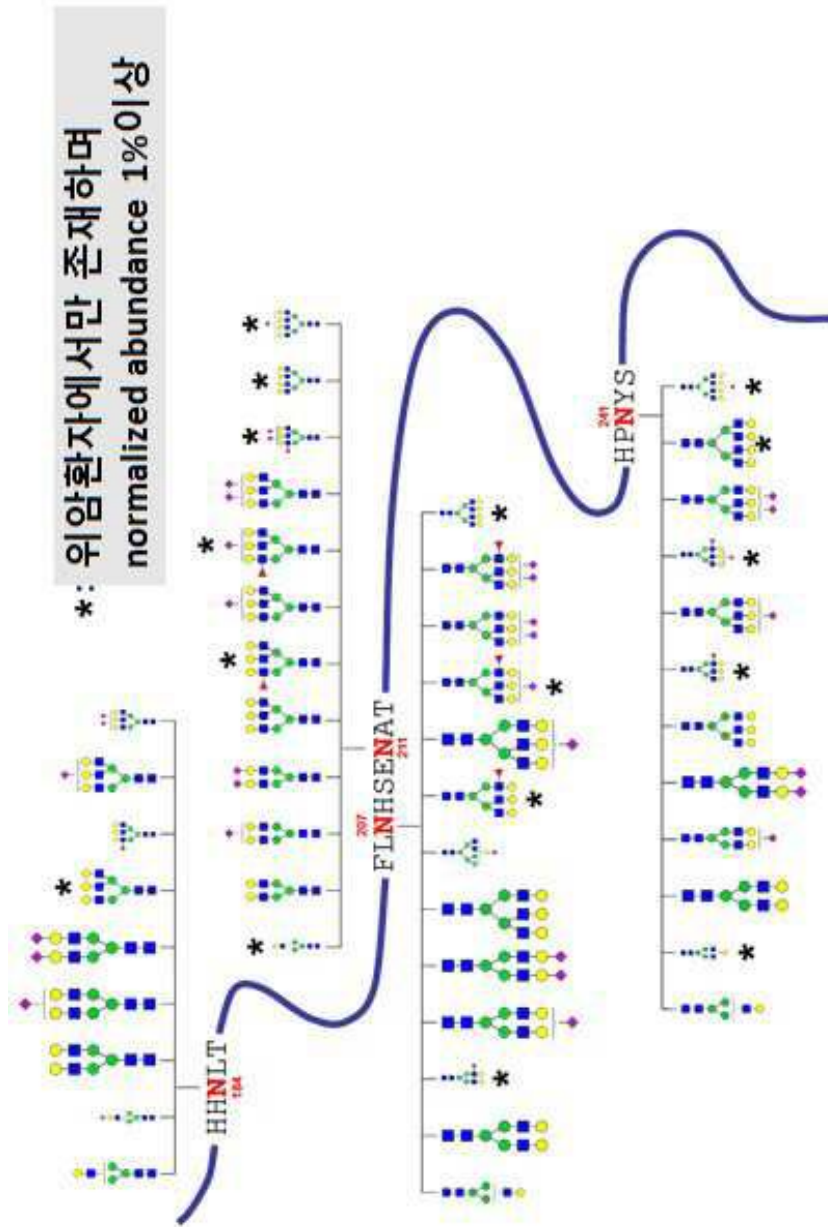
표 1

N-glycosylation site	Peptide	Glycan composition*	Glycopeptide (peptide-glycan composition)	p-value
184	HN	Hexose5_HexNAc4_Fucose1_NeuAc0	HN-5410	1.65E-05
	HN	Hexose6_HexNAc5_Fucose1_NeuAc0	HN-6510	1.92E-05
	HN	Hexose4_HexNAc3_Fucose0_NeuAc0	HN-4300	3.07E-05
	HN	Hexose5_HexNAc4_Fucose0_NeuAc0	HN-5400	5.7E-05
	HN	Hexose5_HexNAc4_Fucose0_NeuAc1	HN-5401	4.63E-04
207	NHSE	Hexose4_HexNAc3_Fucose0_NeuAc0	NHSE-4300	8.48E-06
	NHSE	Hexose5_HexNAc4_Fucose0_NeuAc0	NHSE-5400	9.35E-06
	NHSE	Hexose4_HexNAc4_Fucose0_NeuAc0	NHSE-4400	1.55E-05
	NHSE	Hexose6_HexNAc5_Fucose1_NeuAc0	NHSE-6510	3.67E-05
	NHSE	Hexose6_HexNAc5_Fucose0_NeuAc0	NHSE-6500	8.1E-05
	NHSE	Hexose8_HexNAc2_Fucose0_NeuAc0	NHSE-8200	1.23E-04
	NHSE	Hexose6_HexNAc5_Fucose1_NeuAc1	NHSE-6511	1.33E-04
	NHSE	Hexose5_HexNAc4_Fucose1_NeuAc0	NHSE-5410	2.24E-04
	NHSE	Hexose7_HexNAc6_Fucose0_NeuAc0	NHSE-7600	3.35E-04
211	NAT	Hexose6_HexNAc5_Fucose1_NeuAc1	NAT-6511	4.91E-06
	NAT	Hexose6_HexNAc5_Fucose1_NeuAc0	NAT-6510	6.7E-05
	NAT	Hexose7_HexNAc6_Fucose1_NeuAc0	NAT-7610	9.29E-05
	NAT	Hexose5_HexNAc5_Fucose0_NeuAc1	NAT-5501	2.14E-04
	NAT	Hexose4_HexNAc4_Fucose1_NeuAc0	NAT-4410	2.82E-04
	NAT	Hexose7_HexNAc6_Fucose1_NeuAc1	NAT-7611	3.64E-04
	NAT	Hexose7_HexNAc6_Fucose0_NeuAc1	NAT-7601	4.46E-04
241	HPN	Hexose5_HexNAc4_Fucose0_NeuAc0	HPN-5400	6.85E-05
	HPN	Hexose5_HexNAc5_Fucose1_NeuAc0	HPN-5510	7.32E-05
	HPN	Hexose6_HexNAc5_Fucose1_NeuAc1	HPN-6511	1.03E-04
	HPN	Hexose6_HexNAc5_Fucose0_NeuAc0	HPN-6500	1.72E-04
	HPN	Hexose7_HexNAc6_Fucose1_NeuAc0	HPN-7610	2.03E-04
	HPN	Hexose7_HexNAc6_Fucose0_NeuAc0	HPN-7600	2.15E-04

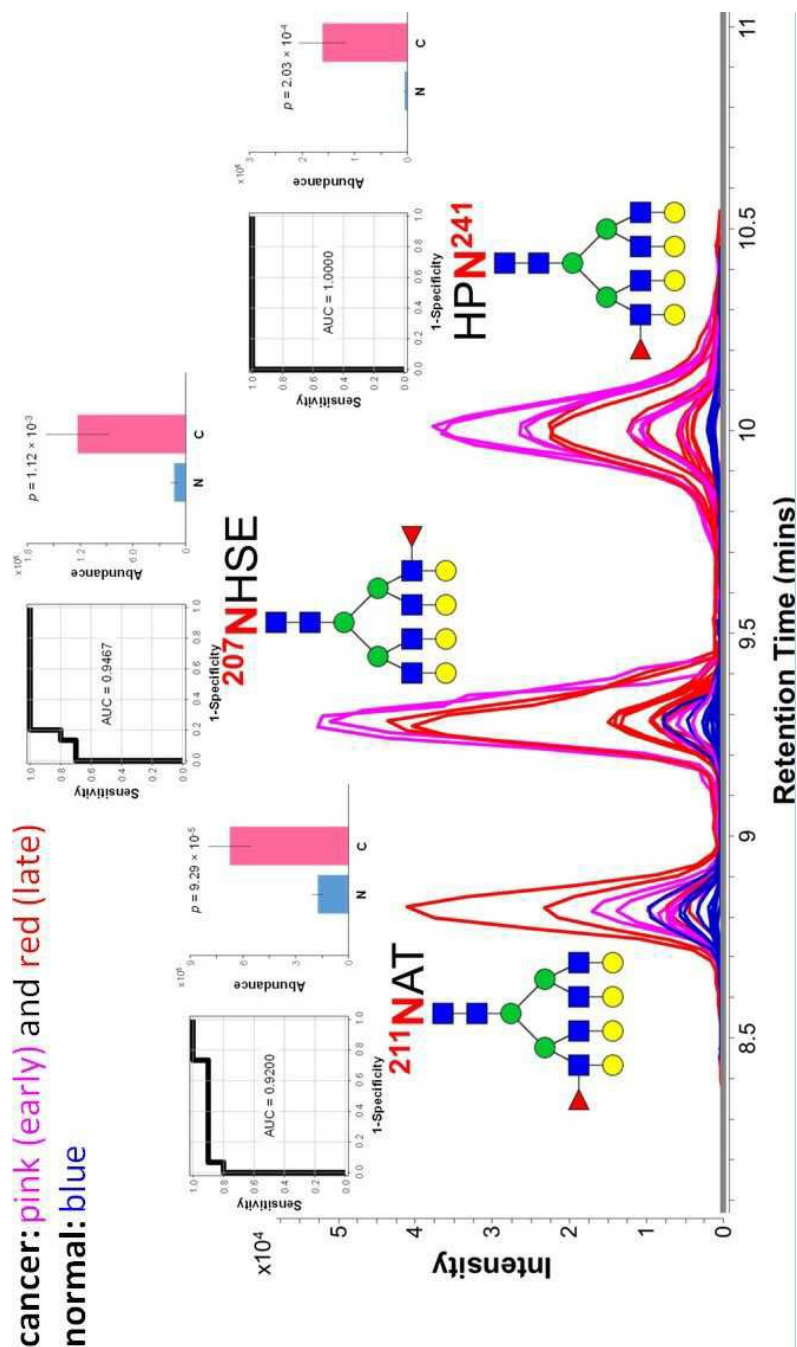
[0176]

도면

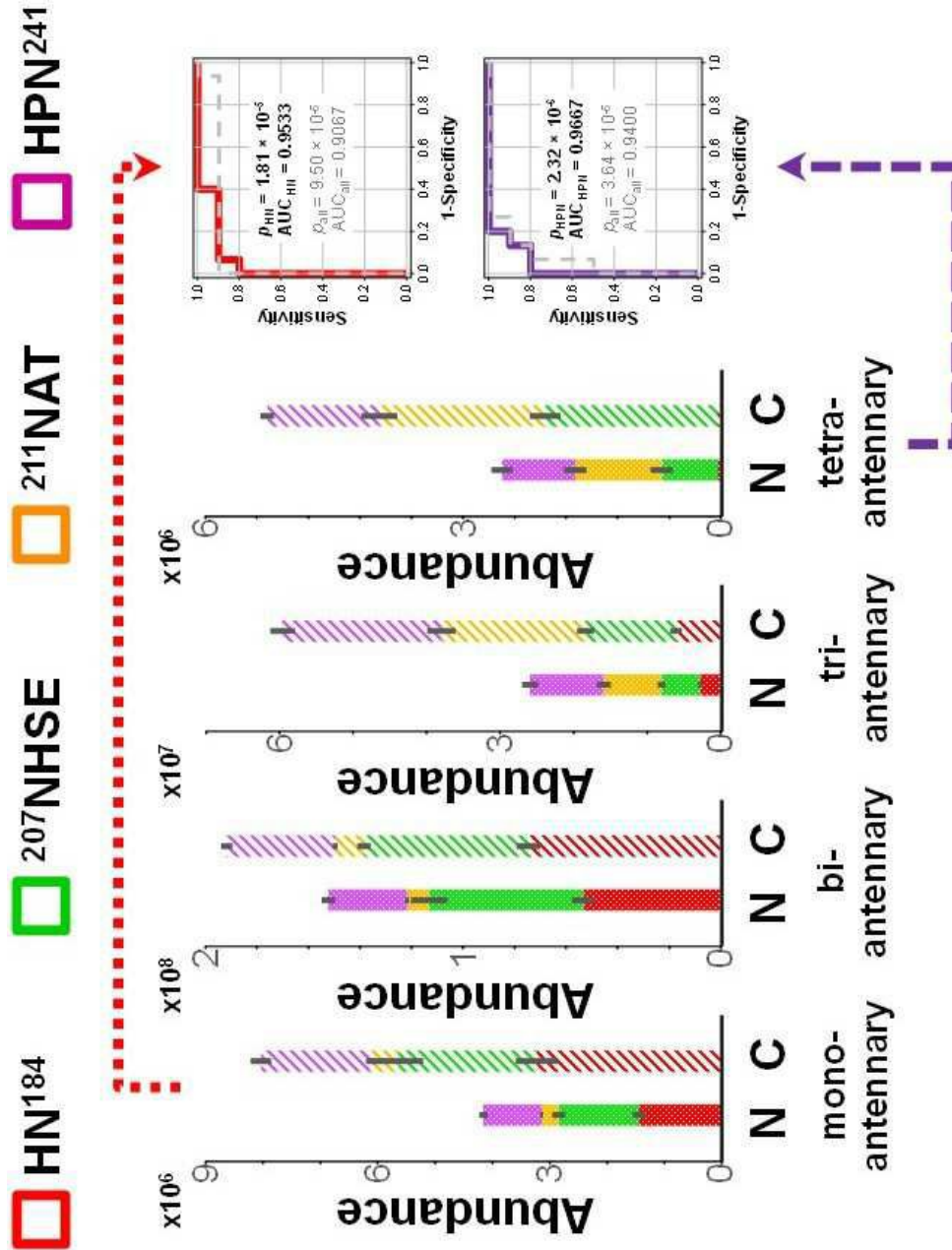
도면1



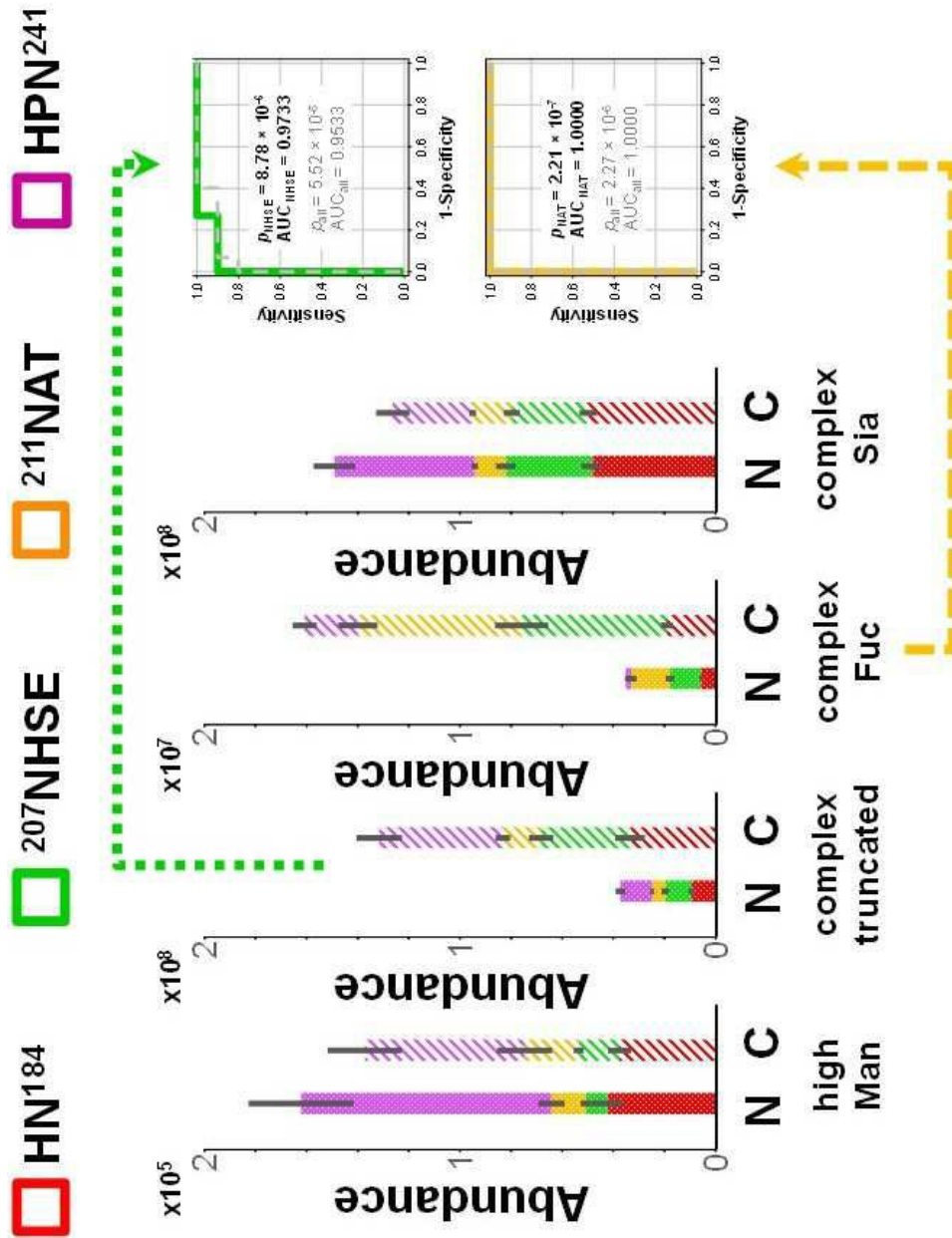
도면2



도면3



도면4



도면5

