

(19) 日本国特許庁(JP)

(12) 公表特許公報(A)

(11) 特許出願公表番号

特表2020-532732  
(P2020-532732A)

(43) 公表日 令和2年11月12日(2020.11.12)

(51) Int.Cl.	F I	テーマコード (参考)
GO 1 N 27/62 (2006.01)	GO 1 N 27/62 V	2 G O 4 1
GO 1 N 33/68 (2006.01)	GO 1 N 33/68	2 G O 4 5
GO 1 N 30/72 (2006.01)	GO 1 N 30/72 G	4 B O 6 3
C 1 2 Q 1/37 (2006.01)	GO 1 N 27/62 X	
	C 1 2 Q 1/37	

審査請求 未請求 予備審査請求 未請求 (全 19 頁)

(21) 出願番号 特願2020-512809 (P2020-512809)  
 (86) (22) 出願日 平成30年8月31日 (2018. 8. 31)  
 (85) 翻訳文提出日 令和2年4月17日 (2020. 4. 17)  
 (86) 国際出願番号 PCT/US2018/049256  
 (87) 国際公開番号 W02019/046814  
 (87) 国際公開日 平成31年3月7日 (2019. 3. 7)  
 (31) 優先権主張番号 62/553, 676  
 (32) 優先日 平成29年9月1日 (2017. 9. 1)  
 (33) 優先権主張国・地域又は機関  
 米国 (US)

(71) 出願人 520068618  
 ヴェン バイオサイエンス コーポレーション  
 アメリカ合衆国, カリフォルニア州 94080, サウス サンフランシスコ, ツー  
 タワー プレイス, フィフス フロア  
 (74) 代理人 100079108  
 弁理士 稲葉 良幸  
 (74) 代理人 100109346  
 弁理士 大貫 敏史  
 (74) 代理人 100117189  
 弁理士 江口 昭彦  
 (74) 代理人 100134120  
 弁理士 内藤 和彦

最終頁に続く

(54) 【発明の名称】 診断および治療モニタリングのためのバイオマーカーとしての糖ペプチドの同定および使用

(57) 【要約】

【課題】 診断および治療モニタリングのためのバイオマーカーとしての糖ペプチドの同定および使用に関する。

【解決手段】 プロテオミクス、ペプチドミクス、メタボリクス、プロテオグリコミクス、グリコミクス、質量分析および機械学習を用いて、様々な疾患のための新規バイオマーカーを同定する方法が提供される。本開示はまた、がんおよび自己免疫疾患などの様々な疾患のためのバイオマーカーとしての糖ペプチドを提供する。

【選択図】 図 1

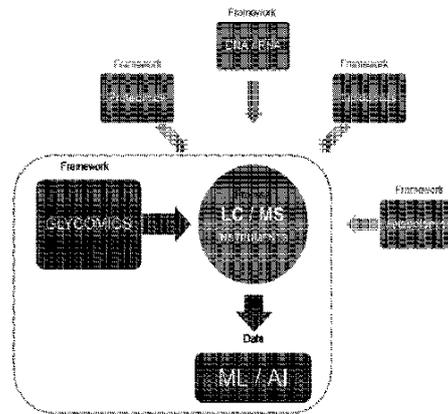


FIG. 1

## 【特許請求の範囲】

## 【請求項 1】

有望なバイオマーカーとしてグリコシル化ペプチド断片を同定する方法であって、対象から単離された複数の生体試料のそれぞれにおいて、1つ以上のプロテアーゼを用いてグリコシル化タンパク質を断片化することであって、グリコシル化ペプチド断片を生成する、ことと、

液体クロマトグラフィーおよび質量分析法(LC-MS)で前記グリコシル化ペプチド断片を定量することであって、定量結果を提供する、ことと、

機械学習法で前記定量結果を前記対象の分類と共に分析することであって、前記分類を予測するために有用なグリコシル化ペプチド断片を選択する、ことと、

グリコシル化ペプチド断片の同一性を決定することと、を含む、方法。

10

## 【請求項 2】

前記対象が、疾患または病態を有する対象と、前記疾患または病態を有しない対象とを含む、請求項 1 に記載の方法。

## 【請求項 3】

前記対象が、疾患の治療を受けている対象と、前記疾患を有するが治療を受けていない対象とを含む、請求項 1 または 2 に記載の方法。

## 【請求項 4】

前記疾患ががんまたは自己免疫疾患である、請求項 2 または 3 に記載の方法。

## 【請求項 5】

前記疾患が、乳がん、子宮頸がんまたは卵巣がんから選択されるがんである、請求項 4 に記載の方法。

20

## 【請求項 6】

前記疾患が、HIV、原発性硬化性胆管炎、原発性胆汁性肝硬変または乾癬から選択される自己免疫疾患である、請求項 5 に記載の方法。

## 【請求項 7】

前記グリコシル化ペプチド断片がN-グリコシル化されている、請求項 1 ~ 6 のいずれか 1 項に記載の方法。

## 【請求項 8】

前記グリコシル化ペプチド断片がO-グリコシル化されている、請求項 1 ~ 7 のいずれか 1 項に記載の方法。

30

## 【請求項 9】

前記グリコシル化タンパク質が、アルファ-1-酸性糖タンパク質、アルファ-1-アンチトリプシン、アルファ-1B-糖タンパク質、アルファ-2-HS-糖タンパク質、アルファ-2-マクログロブリン、アンチトロンビンIII、アポリポタンパク質B-100、アポリポタンパク質D、アポリポタンパク質F、ベータ-2-糖タンパク質1、セロプラスミン、フェチュイン、フィブリノゲン、免疫グロブリン(Ig)A、IgG、IgM、ハプトグロビン、ヘモペキシン、ヒスチジンリッチ糖タンパク質、キニノゲン-1、セロトランスフェリン、トランスフェリン、ビトロネクチン、亜鉛-アルファ-2-糖タンパク質のうちの一つ以上である、請求項 1 ~ 8 のいずれか 1 項に記載の方法。

40

## 【請求項 10】

前記グリコシル化タンパク質が、アルファ-1-酸糖タンパク質、免疫グロブリン(Ig)A、IgGまたはIgMのうちの一つ以上である、請求項 9 に記載の方法。

## 【請求項 11】

前記グリコシル化ペプチド断片が5~50アミノ酸残基の平均長を有する、請求項 1 ~ 10 のいずれか 1 項に記載の方法。

## 【請求項 12】

前記1つ以上のプロテアーゼが少なくとも2つのプロテアーゼを含む、請求項 1 ~ 11 のいずれか 1 項に記載の方法。

## 【請求項 13】

50

質量分析法が多重反応モニタリング質量分析法 ( M R M - M S ) を含む、請求項 1 ~ 1 2 のいずれか 1 項に記載の方法。

【請求項 1 4】

前記生体試料が、体組織、唾液、涙、喀痰、脊髄液、尿、滑液、全血、血清または血漿である、請求項 1 ~ 1 3 のいずれか 1 項に記載の方法。

【請求項 1 5】

前記生体試料が全血、血清または血漿である、請求項 1 4 に記載の方法。

【請求項 1 6】

前記対象が哺乳動物である、請求項 1 ~ 1 5 のいずれか 1 項に記載の方法。

【請求項 1 7】

前記対象がヒトである、請求項 1 6 に記載の方法。

【請求項 1 8】

前記機械学習アプローチが、深層学習、ニューラルネットワーク、線形判別分析、二次判別分析、サポートベクターマシン、ランダムフォレスト、最近傍またはそれらの組み合わせである、請求項 1 ~ 1 7 のいずれか 1 項に記載の方法。

【請求項 1 9】

前記機械学習アプローチが、深層学習、ニューラルネットワーク、またはそれらの組み合わせである、請求項 1 ~ 1 8 のいずれか 1 項に記載の方法。

【請求項 2 0】

前記分析が、ゲノムデータ、プロテオミクス、メタボリクス、リポドミクスデータ、またはそれらの組み合わせをさらに含む、請求項 1 ~ 1 9 のいずれか 1 項に記載の方法。

【発明の詳細な説明】

【技術分野】

【0001】

本開示は、一般に、疾患診断および治療モニタリング用のバイオマーカーを同定するための、マルチオミクス、特に、グライコミクスおよびグリコプロテオミクス、高度計装ビッグデータ、機械学習、および人工知能の分野に関する。

【背景技術】

【0002】

タンパク質のグリコシル化および他の翻訳後修飾は、ヒトの成長と発達のあらゆる面で重要な構造的および機能的役割を果たす。不完全なタンパク質のグリコシル化は、いくつかの疾患を伴う。疾患の初期段階でグリコシル化の変化を特定することにより、影響を受ける対象の早期発見の機会、介入、より高い生存の可能性が提供される。現在、早期がんを検出し、特定の種類のがんを他の疾患と区別できるバイオマーカーを同定する方法がある。これらの方法には、質量分析法 ( M S ) を使用したプロテオミクス、ペプチドミクス、メタボリクス ( m e t a b o l i c s ) 、プロテオグライコミクス ( p r o t e o g l y c o m i c s ) 、およびグライコミクスが含まれる。

【0003】

タンパク質のグリコシル化はがんおよびその他の疾患に関する有用な情報を提供するが、この方法の1つの欠点は、グリカン、由来するタンパク質の部位までさかのぼることができないことである。がんの生物学およびがんの早期発見に関するより多くの知識を得るためには、グリカンだけでなく、タンパク質での結合部位も同定することが重要である。糖タンパク質分析は、いくつかの理由により一般的に困難である。例えば、ペプチド中の単一のグリカン組成物は、異なるグリコシド結合、分岐、および同じ質量を有する多くの単糖のために、多数の異性体構造を含む場合がある。さらに、同じペプチド骨格を共有する複数のグリカンが存在すると、M S シグナルが様々なグリコフォームに分割され、グリコシル化されていないペプチドと比較して個々の存在量が低下する。したがって、タンデム M S データからグリカンとそのペプチドを同定できるアルゴリズムを開発することは困難であった。また、グリカンとペプチドの断片化効率が異なるため、グリカンとペプチドの両方について包括的な断片化を得ることも困難である。

10

20

30

40

50

## 【0004】

したがって、非疾患細胞と比較した、疾患細胞、組織または生体液のグリコシル化部位の不均一性に関する正確な定量的情報を提供するタンパク質グリコシル化パターンに関する重要かつ詳細な情報を得るための部位特異的糖タンパク質分析の方法を提供することの需要がある。そのような方法は、特にがんなどの疾患の疾患バイオマーカーの同定につながる。また、がんなどの疾患のためのグリカンベースの薬物標的などの新しい疾患標的を同定および検証するために、部位特異的な糖タンパク質分析データを深層学習および高度なLC/MS機器と組み合わせて、新しいバイオマーカーを同定する時間を短縮することへの需要もある。

## 【発明の概要】

10

## 【0005】

本開示は、様々な疾患のバイオマーカーを同定する方法に関する。これらのバイオマーカーは、生体試料からのグリコシル化タンパク質の断片化により得られたグリコシル化ペプチド断片である。バイオマーカーを同定する方法は、グリコシル化ペプチド断片の正確な質量測定と部位特異的グリコシル化分析を可能にする高度な質量分析技術の使用に依存する。本開示の質量分析法は、生体試料から一度に多数のグリコシル化タンパク質を分析するのに有利に役立つ。

## 【0006】

一実施形態では、本開示は、有望なバイオマーカーとしてグリコシル化ペプチド断片を同定する方法を提供し、本方法は、

20

対象から単離された複数の生物試料のそれぞれにおいて、1つ以上のプロテアーゼを用いてグリコシル化タンパク質を断片化することであって、グリコシル化ペプチド断片を生成する、ことと、

液体クロマトグラフィーおよび質量分析法(LC-MS)でグリコシル化ペプチド断片を定量することであって、定量結果を提供する、ことと、

分類の予測に役立つグリコシル化ペプチド断片を選択するために、機械学習法で対象の分類とともに定量結果を分析することと、

グリコシル化ペプチド断片の同一性を決定することと、を含む。

## 【0007】

別の実施形態では、本方法は、疾患または病態を有する対象と、疾患または病態を有しない対象とを含む。さらなる実施形態では、対象は、疾患または病態の治療を受けている対象と、疾患または病態を有するが治療を受けていない対象とを含む。

30

## 【0008】

別の実施形態では、本開示の方法は、対象の生体試料からのグリコシル化ペプチド断片を分析することにより検出され得るいずれかの疾患または病態に適用可能である。一実施形態では、疾患はがんである。別の実施形態では、疾患は自己免疫疾患である。別の実施形態では、本開示の方法は、O-グリコシル化またはN-グリコシル化されたグリコシル化ペプチド断片を提供する。別の実施形態では、本開示の方法は、5~50アミノ酸残基の平均長を有するグリコシル化ペプチド断片を提供する。

## 【0009】

40

別の実施形態では、本開示の方法は、アルファ-1-酸性糖タンパク質、アルファ-1-抗トリプシン、アルファ-1B-糖タンパク質、アルファ-2-HS-糖タンパク質、アルファ-2-マクログロブリン、アンチトロンビンIII、アポリポタンパク質B-100、アポリポタンパク質D、アポリポタンパク質F、ベータ2糖タンパク質1、セルロプラスミン、フェチュイン、フィブリノゲン、免疫グロブリン(Ig)A、IgG、IgM、ハプトグロビン、ヘモペキシン、ヒスチジンリッチ糖タンパク質、キニノゲン-1、セロトランスフェリン、トランスフェリン、ビトロネクチンおよび亜鉛-アルファ-2-糖タンパク質のうち1つ以上であるグリコシル化タンパク質を使用する。

## 【0010】

別の実施形態では、本開示の方法は、少なくとも2つのプロテアーゼを使用するグリコ

50

シル化タンパク質の断片化を含む。別の実施形態では、本開示の方法は、多重反応モニタリング質量分析法（MRM-MS）を使用するLC-MS技術を使用する。

【0011】

別の実施形態では、本開示は、本明細書に記載の様々な疾患の有望なバイオマーカーとしてグリコシル化ペプチド断片を同定する方法を提供し、生体試料は、対象から得られた体組織、唾液、涙、喀痰、脊髄液、尿、滑液、全血、血清または血漿である。一実施形態では、対象は哺乳動物である。別の実施形態では、対象はヒトである。

【0012】

別の実施形態では、本開示は、有望なバイオマーカーとしてグリコシル化ペプチド断片を同定する方法を提供し、本方法は、

対象から単離された複数の生物試料のそれぞれにおいて、1つ以上のプロテアーゼを用いてグリコシル化タンパク質を断片化することであって、グリコシル化ペプチド断片を生成する、ことと、

液体クロマトグラフィーおよび質量分析法（LC-MS）でグリコシル化ペプチド断片を定量することであって、定量結果を提供する、ことと、

分類の予測に役立つグリコシル化ペプチド断片を選択するために、機械学習法で対象の分類とともに定量結果を分析することと、

グリコシル化ペプチド断片の同一性を決定することと、を含み、機械学習アプローチは、深層学習、ニューラルネットワーク、線形判別分析、二次判別分析、サポートベクターマシン、ランダムフォレスト、最近傍、またはそれらの組み合わせである。別の実施形態では、機械学習アプローチは、深層学習、ニューラルネットワーク、またはそれらの組み合わせである。

【0013】

別の実施形態では、本開示は、本明細書に記載の様々な疾患の有望なバイオマーカーとしてグリコシル化ペプチド断片を同定する方法を提供し、分析は、ゲノムデータ、プロテオミクス、メタボリクス、リポミクスデータ、またはそれらの組み合わせをさらに含む。

【図面の簡単な説明】

【0014】

【図1】 グライコミクス、LC/MS、機械学習の統合を示す概略図であり、さらにプロテオミクス、ゲノム、リポミクス、メタボリクスと組み合わせることができる。

【図2】 乳がん患者と対照の血漿試料中の免疫グロブリンG（IgG）糖ペプチド比の変化を示す。

【図3】 原発性硬化性胆管炎（PSC）および原発性胆汁性肝硬変（PBC）試料と健常ドナーの血漿試料中のIgG糖ペプチド比の変化を示す。

【図4】 健常ドナーとPSCおよびPBC試料からの血漿試料中のIgG、IgAおよびIgM糖ペプチドの判別分析データを別々に示す。

【図5】 PSCおよびPBC患者と健常ドナーの血漿試料中のIgG、IgAおよびIgM糖ペプチドの複合判別分析データを示す。

【発明を実施するための形態】

【0015】

定義

本明細書で使用される場合、以下の単語および語句は、それらが使用される文脈がそうでないことを示す場合を除いて、一般に、以下に示される意味を有することを意図している。

【0016】

本明細書および特許請求の範囲で使用される単数形「a」、「an」および「the」は、文脈からそうでないことが明確に示されない限り、複数の指示対象を含むことに留意されたい。

【0017】

10

20

30

40

50

「生体試料」という用語は、任意の生体液、細胞、組織、器官、またはその一部を意味する。また、生検によって得られた組織切片、または組織培養した細胞、または組織培養に適合した細胞も含まれるが、これらに限定されない。さらに、唾液、涙、喀痰、汗、粘液、糞便、胃液、腹水 (abdominal fluid)、羊水、嚢胞液、腹水 (peritoneal fluid)、脊髄液、尿、滑液、全血、血清、血漿、腓液、母乳、肺洗浄液、骨髓などが含まれるが、これらに限定されない。

【0018】

「バイオマーカー」という用語は、プロセス、イベント、または病態の、特徴的な生物学的指標または生物学的由来の指標を指す。バイオマーカーは、疾患または病態の存在、または疾患または病態のリスクなど、特定の生物学的状態も示す。それには、生物学的分子、または生物学的分子の断片が含まれ、その変化または検出は、特定の物理的状态または条件と相関させることができる。バイオマーカーの例には、ヌクレオチド、アミノ酸、脂肪酸、ステロイド、抗体、ホルモン、ステロイド、ペプチド、タンパク質、炭水化物などを含む生体分子が含まれるが、これらに限定されない。さらなる例には、グリコシル化ペプチド断片、リポタンパク質などが含まれる。

10

【0019】

「含む (comprising)」という用語は、組成物および方法が列挙された方法を含むが、他の方法を排除しないことを意味することを意図している。

【0020】

「グリカン」という用語は、糖ペプチド、糖タンパク質、糖脂質またはプロテオグリカンなどの複合糖質の炭水化物部分を指す。

20

【0021】

「グリコフォーム」という用語は、特定の構造のグリカンが結合したタンパク質のユニークな一次、二次、三次、および四次構造を指す。

【0022】

「グリコシル化ペプチド断片」という用語は、グリコシル化タンパク質のアミノ酸配列の一部と同じではあるがすべてではないアミノ酸配列を有するグリコシル化ペプチド (または糖ペプチド) を指し、断片化 (例えば1つ以上のプロテアーゼによる) によって、グリコシル化ペプチドが得られる。

【0023】

「多重反応モニタリング質量分析法 (MRM-MS)」という用語は、生体試料中のタンパク質/ペプチドの標的化された定量化のための高感度で選択的な方法を指す。従来の質量分析法とは異なり、MRM-MSは高度に選択的 (標的化) であるため、研究者は機器を微調整して、目的のペプチド/タンパク質断片を特に探すことができる。MRMにより、有望なバイオマーカーなど、目的のペプチド/タンパク質断片の感度、特異性、速度、および定量性が向上する。MRM-MSでは、三連四重極 (QQQ) 質量分析計または四重極飛行時間 (qTOF) 質量分析計を使用する。

30

【0024】

「プロテアーゼ」という用語は、タンパク質をより小さなポリペプチドまたはアミノ酸にタンパク質分解または分解する酵素を指す。プロテアーゼの例には、セリンプロテアーゼ、スレオニンプロテアーゼ、システインプロテアーゼ、アスパラギン酸プロテアーゼ、グルタミン酸プロテアーゼ、メタロプロテアーゼ、アスパラギンペプチドリアーゼおよびそれらの組み合わせが含まれる。

40

【0025】

「対象」という用語は、哺乳類を指す。哺乳動物の非限定的な例には、ヒト、非ヒト霊長類、マウス、ラット、イヌ、ネコ、ウマ、またはウシなどが含まれる。ヒト以外の哺乳類を、疾患、前疾患、または前疾患病態の動物モデルを表す対象として有利に使用できる。対象は雄でも雌でもよい。対象は、疾患または病態を有すると以前に同定された者であって、場合により、疾患または病態のための治療的介入をすでに受けているか、受けている者でよい。あるいは、対象は、疾患または病態を有すると以前に診断されていない対象

50

でもよい。例えば、対象は、疾患または病態の1つ以上の危険因子を示す対象、または疾患の危険因子を示さない対象、または疾患または病態の無症状の対象でもよい。対象はまた、疾患または病態に罹患しているかまたは発症するリスクがある対象でもよい。

【0026】

「治療」または「治療する」という用語は、哺乳動物などの対象における疾患または病態の任意の治療を意味し、1) 疾患または病態に対する予防または保護、すなわち、臨床症状を発症させないこと、2) 疾患または病態の抑制、すなわち臨床症状の発症の阻止または抑制、および/または3) 疾患または病態を緩和し、臨床症状の退行を引き起こすこと、を含む。

【0027】

方法

本開示は、いくつかの実施形態では、グリコプロテオミクス、高度なLC/MS機器を使用したバイオマーカー発見、標的発見、および検証のためのグリコプロテオミクスに関する。本開示は、機械学習法を利用して分子データを処理する。分析は、ゲノムデータ、プロテオミクス、メタボリクス、リポミクスデータ、またはそれらの組み合わせを様々な疾患の新しいバイオマーカーを発見する際に利用することをさらに含む。本開示の方法の一般的な概略図を図1に示す。

【0028】

本開示は、より高い感度および特異性を有する新しいタイプのバイオマーカーの同定につながる部位特異的グリコシル化分析のための方法を提供する。この方法は、グリコシル化ペプチドの定量化を含み、したがって、異なる部位で特定のタンパク質に関連する異なるグリコフォームの示差分析を容易にする。この方法は、タンパク質の量と部位特異的グリコシル化プロファイルに関する情報を提供するため、グリコシル化プロファイルの変化がタンパク質のグリコシル化の変化によるものか、タンパク質濃度の変化によるものかについての洞察を提供する。機械学習法と組み合わせた部位特異的グリコシル化分析は、様々な疾患または病態の新しいバイオマーカーの同定を提供する。

【0029】

本開示の定量的グリコプロテオミクス方法は、様々な疾患のバイオマーカーを発見するために使用される。これらの方法は、いくつかの疾患において特定のグリコフォームが上昇し、他のものがダウンレギュレートされるという事実に基づいており、本開示のLC/MS方法は、有意なグリコシル化変化を分析することにより、疾患対疾患なしを区別する。一実施形態では、部位特異的グリコシル化分析は、目的の糖タンパク質、修飾部位、修飾が何かを特定し、次に各修飾の相対存在量を測定することを含む。いくつかの実施形態では、疾患はがんである。他の実施形態では、疾患は自己免疫疾患である。

【0030】

本開示の方法を使用して、数千人の対象からの生物学的試料がデジタル化され、膨大な量のデータが生成され、深層機械学習分析を受けて様々な疾患の新しい標的が発見される。具体的には、深層学習を使用して、既知および未知のペプチドのクラスター化と、疾患状態と対照状態のLC/MSで見られるグリコシル化シグネチャを比較する。グリコシル化ペプチドのこのような判別分析は、疾患バイオマーカーの同定につながる。

【0031】

バイオマーカーの同定およびそれらの発現レベルなどの対応する機能は、次いで疾患または病態の診断試験方法を開発するために使用される。この方法は、少なくとも部分的には、1つ以上の選択されたバイオマーカーを測定し、疾患または病態との関連の結果を分析することに依拠する。この方法はさらに、1つ以上の治療の選択、治療計画の決定、あるいは特定の疾患または病態の治療に対する反応の監視に使用することができる。したがって、本開示は、疾患または病態の予防、診断、治療、監視および予後診断のための方法を提供する。いくつかの実施形態では、方法は、疾患または病態を有する対象と健常な対象とを区別するのに有用である。いくつかの実施形態では、この方法は、がんを有する対象と健常な対象を区別するのに有用である。いくつかの実施形態では、この方法は、がん

10

20

30

40

50

の診断を支援するか、がんを監視するのに有用である。

【0032】

標的化アプローチおよび非標的化アプローチ

本開示のバイオマーカー発見方法は、標的化アプローチおよび/または非標的化アプローチの両方を使用する。この方法は通常、3つの異なるフェーズ、つまり、発見フェーズ、事前検証フェーズ、および検証フェーズで構成される。

【0033】

発見フェーズ

標的化アプローチは、対象の生物学的試料中の既知のグリコフォームで既知の糖タンパク質を特定および監視することを含む。様々な疾患についてFDAが承認した糖タンパク質バイオマーカーが知られており、それらは本開示の方法を使用して監視され、対象の分類を特定する。通常、バイオマーカーのグリコシル化の変化は腫瘍特異的であり、疾患または疾患のステージの有望なリスクを特定するのに役立つ。標的化アプローチでは、データ収集が実行される前の研究の開始時に確立された生物学的重要性が化学的に特徴付けられ、生物学的に注釈が付けられている、既知の糖タンパク質とそのグリコフォームに焦点を合わせる。定量化は、内部標準および化学的標準標品を使用して実行される。

10

【0034】

具体的には、標的化アプローチでは、部位特異的グリコシル化分析は、疾患または病態を有する多数の対象と、疾患または病態を有しない等しい数の一致した対照対象の、症例対照研究からの生体試料において行われる。疾患関連の糖タンパク質または生物学的活性を有する糖タンパク質などの目的の糖タンパク質が、最初に生体試料において特定される。次に、LC/MSを使用して、修飾部位、修飾の性質、修飾の同一性、および各修飾の相対存在量を分析し、ペプチド断片の同定と定量を行う。このアプローチでは、三連四重極(QQQ)質量分析計を使用して、グリコシル化ペプチド断片の定量を行い、対象の分類に関して分析する。

20

【0035】

非標的化アプローチは、既知および未知のペプチド断片のグリコシル化パターンを学習し、対象の分類を特定するのに役立つグリコシル化パターンの変化に関する詳細情報を提供することを含む。非標的化アプローチは、糖タンパク質の「上方または下方制御」を提供する相対定量技術に基づく。具体的には、糖タンパク質の上方制御または下方制御は、対象の分類に関連して監視される。例えば、糖タンパク質断片は、疾患または病態を有する対象に対して、疾患または病態を有しない対象に対して監視される。このアプローチでは、データが取得される前の各糖タンパク質断片の化学的同一性が不明である。一実施形態では、非標的化アプローチでは、グリコシル化ペプチド断片の分析のために四重極飛行時間(qTOF)質量分析計を使用する。このアプローチでは、機器を使用して、試料中の成分の質量を、これらの成分が何であるかについての何らの先入観もなく、正確に測定する。

30

【0036】

グループ間(疾患対疾患なし)で異なって表現された候補は、重要な臨床的特徴と特徴選択技術による分類の予測を可能にする機械学習法を使用して、さらなる評価のために選択される。パフォーマンスは、機能を選択し、トレーニングセットを使用してモデルを構築する内部相互検証を使用して評価される。得られたモデルは、モデルの構築に使用されなかった試験セットで評価される。偽陽性率は、BenjaminおよびHochbergによって導入された偽発見率(FDR)アプローチを使用して制御される。

40

【0037】

事前検証フェーズ

次いで、このようにして発見段階で特定された候補バイオマーカーは、疾患または病態を有する多数の対象および疾患または病態を有しないそれらの適合対照から得られた生体試料の独立した試験セットで試験され、候補バイオマーカーの性能が決定される。選択されたバイオマーカー、そのランキング、および発見フェーズで開発されたモデルのパラメ

50

ータ推定はすべてモデリングの一部であり、この独立した事前検証フェーズで試験される。候補バイオマーカーのシグナルによって、診断試験は生体試料を疾患のあるグループと疾患のないグループとの2つのグループに分類する。次いで、陽性適中率、陰性適中率、特異性および感度に基づいて、試験の有用性が評価される。また、診断性能は受信者動作特性（ROC）曲線を使用して評価され、どのバイオマーカーまたは複数のバイオマーカーの組み合わせが疾患または病態の統計的に優れた診断試験であるかを試験する。検証に成功した個々のバイオマーカーは、複合マーカーのパネルを形成するためにその後組み込むために検査される。複合マーカーは、重み付き多変数ロジスティック回帰または他の分類アルゴリズムによって構築される。

#### 【0038】

##### 検証フェーズ

事前検証フェーズで保持された候補バイオマーカーは、その後、多数の対象からの独立した盲検化された生体試料を使用した独立した検証を通じて検証される。このフェーズの目的は、選択したバイオマーカーの診断精度を評価することである。

#### 【0039】

一実施形態では、バイオマーカー発見方法は、がんを有する対象から得られた生体試料に適用される。いくつかの実施形態では、少なくとも20人、少なくとも40人、少なくとも60人、少なくとも80人または少なくとも100人の対象からの生体試料が各グループ（すなわち、がんを有するグループまたはがんを有しないグループ）で分析される。

#### 【0040】

標的化アプローチおよび/または非標的化アプローチの両方が、本明細書に記載の機械学習方法とともに、様々な疾患の可能性のあるリスクおよび/または初期段階の検出を特定するための新しい診断方法を提供する。一実施形態では、本開示は、機械学習法と組み合わせた標的化アプローチおよび非標的化アプローチの収束に基づくバイオマーカーの同定方法を提供する。本開示の方法により同定されたバイオマーカーは、診断方法、予後評価の方法、治療結果のモニタリング、特定の治療に应答する可能性のある対象の同定、薬物スクリーニングなどにおいて有用である。

#### 【0041】

一実施形態では、本開示は、有望なバイオマーカーとしてグリコシル化ペプチド断片を同定する方法を提供し、本方法は、

対象から単離された複数の生物試料のそれぞれにおいて、1つ以上のプロテアーゼを用いてグリコシル化タンパク質を断片化することであって、グリコシル化ペプチド断片を生成する、ことと、

液体クロマトグラフィーおよび質量分析法（LC-MS）でグリコシル化ペプチド断片を定量することであって、定量結果を提供する、ことと、

分類の予測に役立つグリコシル化ペプチド断片を選択するために、機械学習法で対象の分類とともに定量結果を分析することと、

グリコシル化ペプチド断片の同一性を決定することと、を含む。

#### 【0042】

別の実施形態では、本開示は、本明細書に記載の方法を提供し、対象は、疾患または病態を有する対象と、疾患または病態を有しない対象とを含む。さらなる実施形態では、対象は、疾患の治療を受けている対象と、疾患を有するが疾患の治療を受けていない対象とを含む。

#### 【0043】

本開示の方法は、対象の生体試料からのグリコシル化ペプチド断片を分析することにより検出され得るいずれかの疾患または病態に適用可能である。一実施形態では、疾患はがんである。別の実施形態では、がんは、乳がん、子宮頸がんまたは卵巣がんから選択される。別の実施形態では、疾患は自己免疫疾患である。別の実施形態では、自己免疫疾患は、HIV、原発性硬化性胆管炎、原発性胆汁性肝硬変または乾癬である。

#### 【0044】

10

20

30

40

50

別の実施形態では、本開示は、本明細書に記載の方法を提供し、グリコシル化タンパク質は、アルファ - 1 - 酸糖タンパク質、アルファ - 1 - アンチトリプシン、アルファ - 1 B - 糖タンパク質、アルファ - 2 - HS - 糖タンパク質、アルファ - 2 - マクログロブリン、アンチトロンピン - III、アポリポタンパク質 B - 100、アポリポタンパク質 D、アポリポタンパク質 F、ベータ - 2 - 糖タンパク質 1、セルロプラスミン、フェツイン、フィブリノゲン、免疫グロブリン (Ig) A、Ig G、Ig M、ハプトグロビン、ヘモペキシン、ヒスチジン - 豊富な糖タンパク質、キニノゲン - 1、セロトランスフェリン、トランスフェリン、ビトロネクチンおよび亜鉛 - アルファ - 2 - 糖タンパク質のうち 1 つ以上である。別の実施形態では、グリコシル化タンパク質は、アルファ - 1 - 酸糖タンパク質、免疫グロブリン (Ig) A、Ig G または Ig M のうち 1 つ以上である。

10

**【0045】**

別の実施形態では、本開示は、本明細書に記載の方法を提供し、グリコシル化ペプチド断片が N - グリコシル化または O - グリコシル化されている。

**【0046】**

別の実施形態では、本開示は、本明細書に記載の方法を提供し、グリコシル化ペプチド断片は、約 5 個 ~ 約 50 個のアミノ酸残基の平均長を有する。いくつかの実施形態では、グリコシル化ペプチド断片は、約 5 個 ~ 約 45 個、または約 5 個 ~ 約 40 個、または約 5 個 ~ 約 35 個、または約 5 個 ~ 約 30 個、または約 5 個 ~ 約 25 個、または約 5 個 ~ 約 20 個、または約 5 個 ~ 約 15 個、または約 5 個 ~ 約 10 個、または約 10 個 ~ 約 50 個、または約 10 個 ~ 約 45 個、または約 10 個 ~ 約 40 個、または約 10 個 ~ 約 35 個、または約 10 個 ~ 約 30 個、または約 10 個 ~ 約 25 個、または約 10 個 ~ 約 20 個、または約 10 個 ~ 約 15 個、または約 15 個 ~ 約 45 個、または約 15 個 ~ 約 40 個、または約 15 個 ~ 約 35 個、または約 15 個 ~ 約 30 個、または約 15 個 ~ 約 25 個または約 15 個 ~ 約 20 個のアミノ酸残基の平均長を有する。一実施形態では、グリコシル化ペプチド断片は、約 15 個のアミノ酸残基の平均長を有する。別の実施形態では、グリコシル化ペプチド断片は、約 10 個のアミノ酸残基の平均長を有する。別の実施形態では、グリコシル化ペプチド断片は、約 5 個のアミノ酸残基の平均長を有する。

20

**【0047】**

別の実施形態では、本開示は、本明細書に記載の方法を提供し、1 つ以上のプロテアーゼは、タンパク質の断片化に使用される任意のプロテアーゼを含む。一実施形態では、プロテアーゼは、セリンプロテアーゼ、スレオニンプロテアーゼ、システインプロテアーゼ、アスパラギン酸プロテアーゼ、グルタミン酸プロテアーゼ、メタロプロテアーゼ、アスパラギンペプチドリアーゼまたはそれらの組み合わせである。プロテアーゼのいくつかの代表的な例には、トリプシン、キモトリプシン、エンドプロテイナーゼ、Asp - N、Arg - C、Glu - C、Lys - C、ペプシン、サーモリシン、エラスターゼ、パバイン、プロテイナーゼ K、スプチリシン、クロストリパイン、カルボキシペプチダーゼが含まれるが、これらに限定されない。別の実施形態では、本開示は、本明細書に記載の方法を提供し、1 つ以上のプロテアーゼは、少なくとも 2 つのプロテアーゼを含む。

30

**【0048】**

本開示の方法には、いくつかのさらなる用途がある。例えば、1 つ以上のバイオマーカーは、疾患前の状態と疾患の状態、または疾患の状態と正常な状態を区別するのに役立つ。非疾患固有のその他の健康状態も判断できる。例えば、バイオマーカーの変化は、様々な時点で分析できる。疾患のある対象では、疾患の進行を監視し、治療を受けている対象では、治療の効果を監視し、治療後の対象では、再発の可能性を監視する。また、特定量のバイオマーカーのレベルにより、疾患の治療コースを選択することもできる。例えば、生物学的試料は、疾患の治療計画を受けている対象から提供され得る。そのような治療レジメンには、運動レジメン、食事サプリメント、減量、外科的介入、デバイス移植、および疾患または病態と診断または同定された対象に使用される治療薬または予防薬による治療が含まれるが、これらに限定されない。

40

**【0049】**

50

さらに、複数の糖タンパク質の糖ペプチド比の変化は、特定の疾患状態または疾患の欠如に関連している可能性がある。例えば、生体試料中の複数の特定の糖ペプチドの存在は、疾患の非存在を示す場合がある一方、生体試料中の複数の他の特定の糖ペプチドの存在は、疾患の存在を示す場合がある。したがって、様々な糖ペプチドプロファイルまたは糖ペプチドバイオマーカーのパネルは、疾患の様々な状態と関連する可能性がある。

#### 【0050】

実施例2は、対照に対する乳がん患者からの血漿試料中のIgG1、IgG0およびIgG2糖ペプチドの変化の定量結果を示す。図2は、糖ペプチドA1およびA2のレベルが対照と比較して上昇したのに対し、糖ペプチドA8、A9およびA10のレベルはこの実験で研究した乳がんのすべての段階で対照と比較して低下したことを示しており、糖ペ

10

#### 【0051】

実施例3は、PSCを有する患者およびPSCを有する患者からの血漿試料中のIgG、IgMおよびIgA糖ペプチドの変化の定量結果を示す。図3は、PBCおよびPSCを有する患者の血漿試料において、糖ペプチドAが健常ドナーと比較して上昇したのに対し、糖ペプチドH、IおよびJは、PBCおよびPSCを有する患者の血漿試料において健常ドナーと比較して減少したことを示す。したがって、糖ペプチドA、H、I、およびJは、PBCおよびPSCの有望なバイオマーカーである。さらに、図4および図5には、判別分析の個別分析結果と複合分析結果がそれぞれ示され、複合判別分析での、疾患状態を予測するための88%の精度が示唆される。

20

#### 【0052】

いくつかの実施形態では、本開示は、検出および分析されるバイオマーカーの数が1、または2以上、例えば3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24、25、26、27、28、30以上である方法を提供する。したがって、本開示はまた、疾患または病態の診断に有用なバイオマーカーのパネルを提供する。

#### 【0053】

##### 質量分析法

一実施形態では、本開示は、質量分析計を使用することによりグリコシル化ペプチド断片を定量化することを含む、本明細書に記載の方法を提供する。一実施形態では、本方法は、「多重反応モニタリング(MRM)」と呼ばれる技術を使用する。多くの場合、この手法は液体クロマトグラフィー(LC/MRM-MS)と組み合わせられ、1回のLC/MRM-MS分析で数百のグリコシル化ペプチド断片(およびその親タンパク質)を定量できる。本開示の高度な質量分析技術は、効果的なイオン源、より高い分解能、より速い分離、およびより高いダイナミックレンジを備えた検出器を提供し、標的化測定の特長を保持する幅広い非標的化測定を可能にする。

30

#### 【0054】

本開示の質量分析法は、一度にいくつかのグリコシル化タンパク質に適用可能である。例えば、少なくとも50超、または少なくとも60超、または少なくとも70超、または少なくとも80超、または少なくとも90超、または少なくとも100超、または少なくとも110超、または少なくとも120超のグリコシル化タンパク質を、質量分析計を使用して一度に分析できる。

40

#### 【0055】

一実施形態では、本開示の質量分析法は、QQQまたはqTOF質量分析計を使用する。別の実施形態では、本開示の質量分析法は、5,000以上、または10,000以上、または25,000以上、または50,000以上、または100,000以上の分解能で、10ppm以上、または5ppm以上、または2ppm以上、または1ppm以上、または0.5ppm以上、0.2ppm以上、または0.1ppm以上の高い質量精度を備えたデータを提供する。

50

## 【 0 0 5 6 】

## 生体試料

本開示は、生体試料からのグリコシル化ペプチド断片の定量に基づく方法を提供する。いくつかの実施形態では、生体試料は過去に収集された1つ以上の臨床試料であり、したがって、新しいバイオマーカーの同定に付さなければならないリソースおよび時間を削減する。いくつかの実施形態では、生物学的試料は、1～50年以上のスパンにわたって生じた1つ以上の過去の研究からのものである。いくつかの実施形態では、研究は、様々な他の臨床パラメータと、対象の年齢、身長、体重、民族性、病歴などの既知の情報を伴う。そのような追加情報は、対象を疾患または病態に関連付けるのに役立つ。いくつかの実施形態では、生体試料は、対象から前向きに収集された1つ以上の臨床試料である。

10

## 【 0 0 5 7 】

一実施形態では、本開示は、本明細書に記載の方法を提供し、対象から単離された生体試料は、唾液、涙、喀痰、汗、粘液、糞便、胃液、腹膜液、羊水、囊胞液、腹腔液、脊髄液、尿、滑液、全血、血清、血漿、臍液、母乳、肺洗浄液、骨髓の1つ以上である。別の実施形態では、対象から単離された生体試料は、体組織、唾液、涙、喀痰、脊髄液、尿、滑液、全血、血清または血漿である。別の実施形態では、対象から単離された生体試料は、全血、血清または血漿である。いくつかの実施形態では、対象は哺乳動物である。他の実施形態では、対象はヒトである。

## 【 0 0 5 8 】

## 疾患

本開示の方法は、対象の生体試料からのグリコシル化ペプチド断片を分析することにより検出され得るいずれかの疾患または病態に適用可能である。いくつかの実施形態では、疾患または病態はがんである。他の実施形態では、がんは、急性リンパ性白血病（ALL）、急性骨髄性白血病（AML）、副腎皮質がん、肛門がん、膀胱がん、血液がん、骨がん、脳腫瘍、乳がん、女性生殖器系のがん、がん男性生殖器系、中枢神経系リンパ腫、子宮頸がん、小児横紋筋肉腫、小児肉腫、慢性リンパ性白血病（CLL）、慢性骨髄性白血病（CML）、結腸および直腸がん、結腸がん、子宮内膜がん、子宮内膜肉腫、食道がん、眼がん、胆嚢がん、胃がん、消化管がん、有毛細胞白血病、頭頸部がん、肝細胞がん、ホジキン病、下咽頭がん、カポジ肉腫、腎臓がん、喉頭がん、白血病、肝臓がん、肺がん、悪性腫瘍線維性組織球腫、悪性胸腺腫、黒色腫、中皮腫、多発性骨髄腫、骨髄腫、鼻腔および副鼻腔がん、鼻咽頭がん、神経系がん、神経芽細胞腫、非ホジキンリンパ腫、口腔がん、口腔咽頭がん、骨肉腫、卵巣がん、膵臓がん、副甲状腺がん、陰茎がん、咽頭がん、下垂体腫瘍、形質細胞新生物、原発性CNSリンパ腫、前立腺がん、直腸がん、呼吸器系、網膜芽細胞腫、唾液腺がん、皮膚がん、小腸がん、軟部組織肉腫、胃がん、精巣がん、甲状腺がん、尿路がん、子宮肉腫、膣がん、血管系、ウォルデンストロムのマクログロブリン血症、ウィルムス腫瘍などである。別の実施形態では、がんは乳がん、子宮頸がんまたは卵巣がんである。

20

30

## 【 0 0 5 9 】

別の実施形態では、疾患は自己免疫疾患である。別の実施形態では、自己免疫疾患は、急性播種性脳脊髄炎、アディソン病、無ガンマグロブリン血症、加齢黄斑変性、円形脱毛症、筋萎縮性側索硬化症、強直性脊椎炎、抗リン脂質症候群、抗シンセターゼ症候群、アトピー性アレルギー、アトピー性皮膚炎、自己免疫性貧血、自己免疫性心筋症、自己免疫性腸疾患、自己免疫性溶血性貧血、自己免疫性肝炎、自己免疫性内耳疾患、自己免疫性リンパ球増殖性症候群、自己免疫性末梢神経障害、自己免疫性膵炎、自己免疫性ポリ内分泌症候群、自己免疫性プロゲステロン皮膚炎、自己免疫性血小板減少性紫斑病、自己免疫性湿疹、自己免疫性ぶどう膜炎、パロー病/パロー同心円性硬化症、ベーチェット病、ベルガー病、ピッカースタッフ脳炎、ブラウ症候群、水疱性類天疱瘡、がん、キャスルマン病、セリアック病、シャーガス病、慢性炎症性脱髄性多発神経障害、慢性再発性イフォカル骨髄炎、慢性閉塞性肺疾患、チャグ・ストラウス症候群、癩痕性類天疱瘡、コーガン症候群、寒冷凝集素病、補体2成分欠乏症、接触皮膚炎、頭蓋動脈炎、CREST症候群

40

50

、クローン病、クッシング症候群、皮膚白血球破壊血管炎、デゴ病皮膚病、疱疹状皮膚炎、皮膚筋炎、1型糖尿病、びまん性皮膚全身性硬化症、ドレスラー症候群、薬物誘発性ループス、円板状エリテマトーデス、湿疹、子宮内膜症、腱炎関連関節炎、好酸球性筋膜炎、好酸球性胃腸炎、後天性表皮水疱症、結節性紅斑、胎児赤芽球症、本態性混合クリオグロブリン血症、エバン症候群、進行性線維性異形成症、線維化性肺胞炎、胃炎、胃腸類天疱瘡、糸球体腎炎、グッドパスチャー症候群、グレーブス病、ギランバレー症候群、橋本筋症、橋本病エノック・シェーンライン紫斑病、H I V、妊娠性類天疱瘡、化膿性汗腺炎、ヒューズ・ストピン症候群、低ガンマグロブリン血症、特発性炎症性脱髄疾患、特発性肺線維症、特発性血小板減少性紫斑病、I g A腎症、封入体筋炎、慢性炎症性脱髄性多発神経炎、間質性膀胱炎、若年性特発性関節炎、川崎病、ランバート・イートン筋無力症候群、白血球破碎性血管炎、扁平苔癬、硬化性苔癬、線状I g A疾患、紅斑性狼瘡、マジエード症候群、メニエール病、顕微鏡的多発血管炎、混合性結合組織病、斑状強皮症、ムッカハーベルマン病、多発性硬化症、重症筋無力症、筋炎、ナルコレプシー、視神経脊髄炎、神経ミオトニー、眼性癬痕性類天疱瘡、眼球クロウヌスミオクロウヌス症候群、オード甲状腺炎、パリンドローム性リウマチ、連鎖球菌に関連する小児自己免疫神経精神疾患、傍腫瘍性小脳変性症、腫瘍随伴性発作性夜間血色素尿症、パリーロンベルグ症候群、パーソナージターナー症候群、パーズブラニチス、尋常性天疱瘡、悪性貧血、静脈性脳脊髄炎、P O E M S症候群、結節性多発動脈炎、多発性筋痛リウマチ、多発性筋炎、原発性胆汁性肝硬変、原発性硬化性胆管炎、進行性炎症性神経障害、乾癬、乾癬性関節炎、壊疽性膿皮症、純粋な赤血球無形成、ラスムッセン脳炎、レイノー現象、再発性多発性軟骨炎、ライター症候群、下肢静止不能症候群、後腹膜線維症、関節リウマチ、リウマチ熱、サルコイドーシス、統合失調症、シュミット症候群、シュニッツラー症候群、強膜炎、皮膚硬化症、血清病、シェーグレン症候群、脊椎関節症、スティッフパーソン症候群、亜急性細菌性心内膜炎、ササック症候群、スウィート症候群、交感神経性眼炎、高安動脈炎、側頭動脈炎、血小板減少症、トロサハント症候群、横断性脊髄炎、潰瘍性大腸炎、無差別結合組織病、じん麻疹様血管炎、血管炎、白斑およびウェゲナー肉芽腫症などである。別の実施形態では、自己免疫疾患は、H I V、原発性硬化性胆管炎、原発性胆汁性肝硬変または乾癬である。

#### 【0060】

##### 機械学習

生体試料は、数千人の対象から取得される。これらの対象は、これまでに発見されていないマーカーのディープマイニングと検証を目的としたデジタル化に使用される。いくつかの実施形態では、生体試料は腫瘍試料または血液試料である。これらはL C / M S機器を使用してデジタル化され、膨大な量のデータが生成され、深層機械学習分析を受けて様々な疾患の新しい標的を発見する。いくつかの実施形態では、疾患はがんまたは自己免疫疾患である。

#### 【0061】

一実施形態では、本開示は、有望なバイオマーカーとしてグリコシル化ペプチド断片を同定する方法を提供し、本方法は、

対象から単離された複数の生物試料のそれぞれにおいて、1つ以上のプロテアーゼを用いてグリコシル化タンパク質を断片化することであって、グリコシル化ペプチド断片を生成する、ことと、

液体クロマトグラフィーおよび質量分析法(L C - M S)でグリコシル化ペプチド断片を定量することであって、定量結果を提供する、ことと、

分類の予測に役立つグリコシル化ペプチド断片を選択するために、機械学習法で対象の分類とともに定量結果を分析することと、

グリコシル化ペプチド断片の同一性を決定することと、を含み、ここで、機械学習アプローチは、深層学習、ニューラルネットワーク、線形判別分析、二次判別分析、サポートベクターマシン、ランダムフォレスト、最近傍、またはそれらの組み合わせである。いくつかの実施形態では、機械学習アプローチは、深層学習、ニューラルネットワーク、また

はそれらの組み合わせである。分析は、ゲノムデータ、プロテオミクス、メタボリクス、リポミクスデータ、またはそれらの組み合わせをさらに含む。図1は、グライコミクス、LC/MS、および機械学習の統合を示す概略図であり、さらにプロテオミクス、ゲノム、リポミクス、およびメタボリクスと組み合わせ、様々な疾患のバイオマーカーを同定する。

#### 【実施例】

#### 【0062】

##### 実施例1

##### バイオマーカー発見の一般的な方法

標的化アプローチでは、まず目的の糖タンパク質を生体試料で特定し、次にLC/MSを使用して修飾部位、修飾の性質、修飾のアイデンティティ、各修飾の相対存在量を分析して、ペプチド断片の同一性とペプチド断片の定量を行った。このアプローチでは、三連四重極(QQQ)質量分析計を使用して、グリコシル化されたペプチド断片を定量し、対象の分類との関係について分析する。

10

#### 【0063】

非標的化アプローチでは、すべてのペプチド断片(既知および未知)のグリコシル化パターンが分析され、様々な対象のグリコシル化パターンの変化に関する情報が得られる。具体的には、糖タンパク質のアップレギュレーションまたはダウンレギュレーションは、対象の分類に関連して監視される。例えば、糖タンパク質断片は、疾患または病態を有する対象に対する、疾患または病態を有しない対象について、監視される。このアプローチでは、グリコシル化ペプチド断片の分析に四重極飛行時間(qTOF)質量分析計を使用する。

20

#### 【0064】

##### 実施例2

##### 乳がんの有望なバイオマーカーとしてのIgG糖ペプチドの定量化

がんの様々な段階にある乳がん患者と、年齢が一致する対照の血漿試料を、IgG1、IgG0、およびIgG2糖ペプチドについて分析し、それらの比率の変化を比較した。具体的には、Tisステージの20の試料、EC1ステージの50の試料、EC2ステージの試料、EC3ステージの25の試料、EC4ステージの9の試料、および73の年齢が一致した対照試料を、QQQ質量分析計でMRM定量分析にかけた。図2の定量的結果からわかるように、特定のIgG1糖ペプチドのレベルを対照と比較して評価したところ、特定のIgG1糖ペプチドのレベルはこの実験で研究された乳がんのすべての段階での対照と比較して減少した。例えば、A1~A11と名付けられたIgG1糖ペプチドをモニターしたところ、糖ペプチドA1およびA2のレベルは対照と比較して上昇し、一方、糖ペプチドA8、A9およびA10のレベルはこの実験で研究された乳がんのすべての段階での対照と比較して減少した。したがって、糖ペプチドA1、A2、A8、A9およびA10は、乳がんの有望なバイオマーカーである。

30

#### 【0065】

##### 実施例3

##### PSCおよびPBCの有望なバイオマーカーとしてのIgG糖ペプチドの定量化

原発性硬化性胆管炎(PSC)を有する患者、原発性胆汁性肝硬変(PBC)を有する患者からの血漿試料、および健常ドナーからの血漿試料をIgG1およびIgG2糖ペプチドについて分析し、グリコペイド比の変化を比較した。具体的には、100のPBC血漿試料、76のPSC血漿試料、49の健常ドナーからの血漿試料を、QQQ質量分析計でMRM定量分析にかけた。図3の定量結果からわかるように、特定のIgG1糖ペプチドは健常ドナーと比較して上昇したが、特定のIgG1糖ペプチドはPBCとPSCの患者の血漿試料の対照と比較して減少した。例えば、糖ペプチドAはPBCとPSCの患者の健常ドナーと比較して上昇したが、糖ペプチドH、IとJはPBCとPSCの患者の血漿試料の健常ドナーと比較して減少した。したがって、糖ペプチドA、H、I、およびJは、PBCおよびPSCの有望なバイオマーカーである。

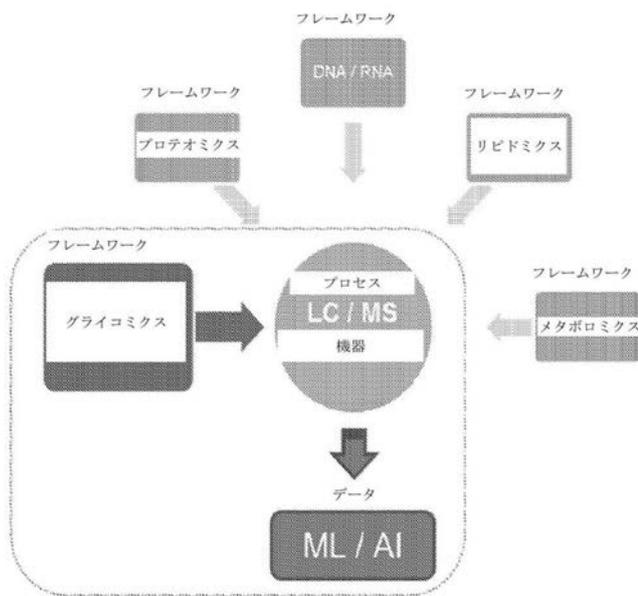
40

50

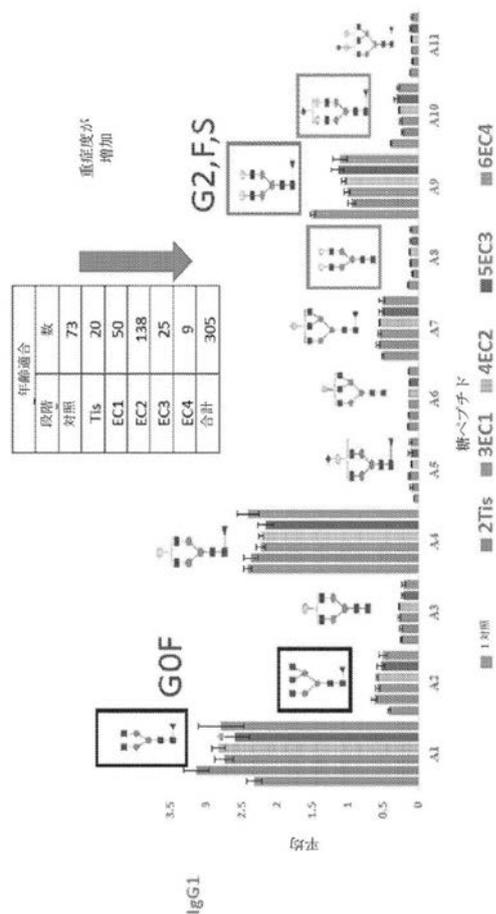
【 0 0 6 6 】

同様の分析が、PBCを有する患者の血漿試料およびPSCを有する患者の血漿試料中のIgAおよびIgM糖タンパク質に対して実施された。判別分析の結果を図4に示す。これは、IgG、IgM、およびIgAの個別のデータに基づいて予測できる精度がそれぞれ59%、69%、74%であることを示す。しかし、IgG、IgMおよびIgAのすべてについての結果を組み合わせると、図5に示すように判別分析により約88%の精度が得られる。

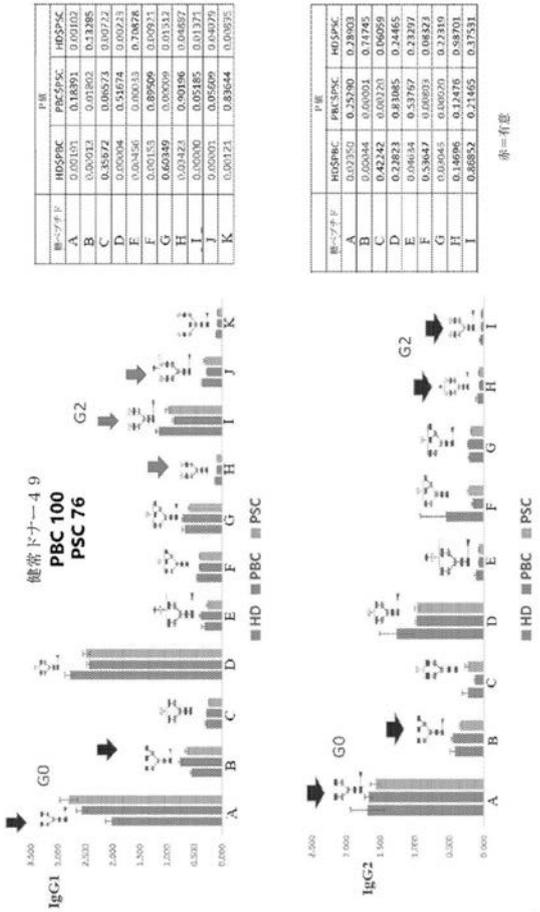
【 図 1 】



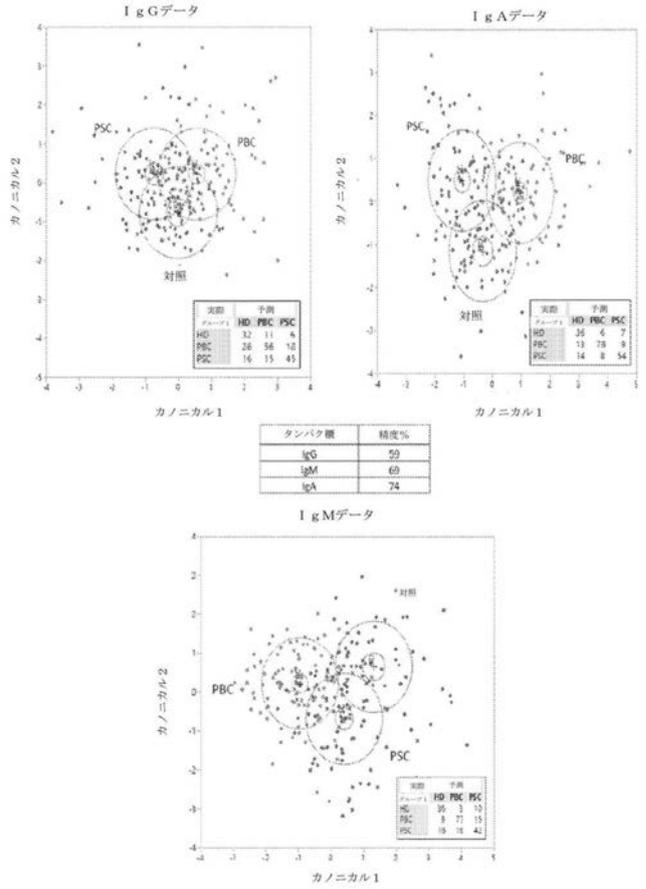
【 図 2 】



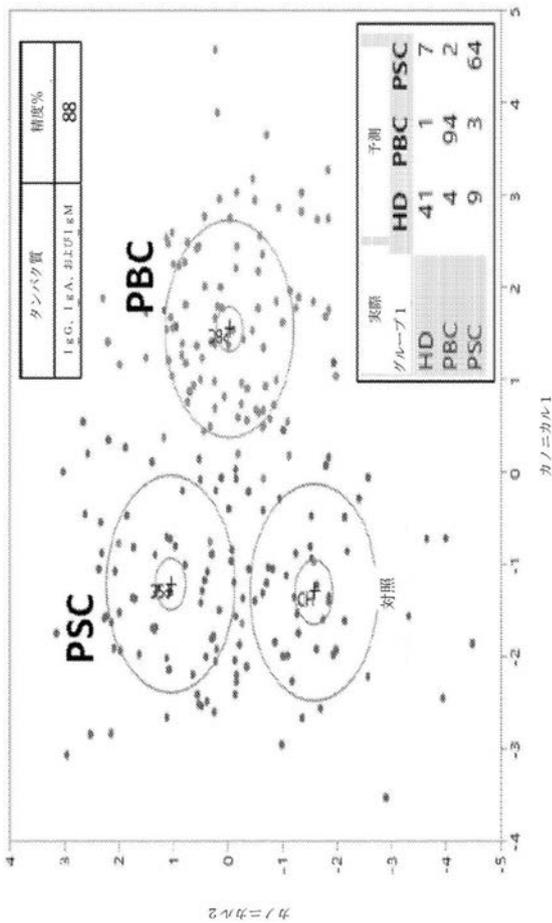
【 図 3 】



【 図 4 】



【 図 5 】



## 【 国際調査報告 】

## INTERNATIONAL SEARCH REPORT

International application No.

PCT/US2018/049256

<b>A. CLASSIFICATION OF SUBJECT MATTER</b> IPC(8) - C12Q 1/37; G01N 33/53; G01N 33/574; G01N 33/68 (2018.01) CPC - G01N 33/5091; G01N 33/57407; G01N 33/57419; G01N 33/57434; G01N 33/57484; G01N 33/6842; G01N 33/6848; G01N 33/6893; G01N 2333/47; G01N 2333/4722; G01N 2333/4728; G01N 2333/475; G01N 2333/495; G01N 2333/70525; G01N 2333/70596; G01N 2333/71; G01N 2333/8146; G01N 2458/15 (2018.08)		
According to International Patent Classification (IPC) or to both national classification and IPC		
<b>B. FIELDS SEARCHED</b>		
Minimum documentation searched (classification system followed by classification symbols) See Search History document		
Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched USPC - 424/178.1; 435/7.1; 435/7.21; 435/7.92 (keyword delimited)		
Electronic data base consulted during the international search (name of data base and, where practicable, search terms used) See Search History document		
<b>C. DOCUMENTS CONSIDERED TO BE RELEVANT</b>		
<b>Category*</b>	<b>Citation of document, with indication, where appropriate, of the relevant passages</b>	<b>Relevant to claim No.</b>
X	US 2006/0141528 A1 (AEBERSOLD et al) 29 June 2006 (29.06.2006) entire document	1-3
A	WO 2009/138392 A1 (ETH ZÜRICH et al) 19 November 2009 (19.11.2009) entire document	1-3
A	US 2009/0053828 A1 (REGNIER et al) 26 February 2009 (26.02.2009) entire document	1-3
A	WO 2015/023068 A1 (SEOUL NATIONAL UNIVERSITY R&DB FOUNDATION (SNU)) 19 February 2015 (19.02.2015) entire document	1-3
A	US 2013/0287801 A1 (UNIVERSITE DE LIEGE) 31 October 2013 (31.10.2013) entire document	1-3
<input type="checkbox"/> Further documents are listed in the continuation of Box C.		<input type="checkbox"/> See patent family annex.
* Special categories of cited documents: "A" document defining the general state of the art which is not considered to be of particular relevance "E" earlier application or patent but published on or after the international filing date "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified) "O" document referring to an oral disclosure, use, exhibition or other means "P" document published prior to the international filing date but later than the priority date claimed		"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art "&" document member of the same patent family
Date of the actual completion of the international search 08 October 2018		Date of mailing of the international search report <b>07 NOV 2018</b>
Name and mailing address of the ISA/US Mail Stop PCT, Attn: ISA/US, Commissioner for Patents P.O. Box 1450, Alexandria, VA 22313-1450 Facsimile No. 571-273-8300		Authorized officer Blaine R. Copenheaver PCT Helpdesk: 571-272-4300 PCT OSP: 571-272-7774

Form PCT/ISA/210 (second sheet) (January 2015)

## INTERNATIONAL SEARCH REPORT

International application No.

PCT/US2018/049256

**Box No. II Observations where certain claims were found unsearchable (Continuation of item 2 of first sheet)**

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1.  Claims Nos.:  
because they relate to subject matter not required to be searched by this Authority, namely:
  
2.  Claims Nos.:  
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:
  
3.  Claims Nos.: 4-20  
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

**Box No. III Observations where unity of invention is lacking (Continuation of item 3 of first sheet)**

This International Searching Authority found multiple inventions in this international application, as follows:

1.  As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2.  As all searchable claims could be searched without effort justifying additional fees, this Authority did not invite payment of additional fees.
3.  As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:
  
4.  No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

**Remark on Protest**

- The additional search fees were accompanied by the applicant's protest and, where applicable, the payment of a protest fee.
- The additional search fees were accompanied by the applicant's protest but the applicable protest fee was not paid within the time limit specified in the invitation.
- No protest accompanied the payment of additional search fees.

## フロントページの続き

(81) 指定国・地域 AP(BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), EA(AM, AZ, BY, KG, KZ, RU, TJ, TM), EP(AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OA(BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG), AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT

(72) 発明者 ダナン - レオン, リーザ マリエ アラウロ

アメリカ合衆国, カリフォルニア州 94080, サウス サンフランシスコ, デリー ウェイ  
2267

(72) 発明者 カッラスコーソ, アルド マリオ エデュアルド シルヴァ

アメリカ合衆国, カリフォルニア州 94015, デーリー シティー, ノースゲート アベニュー  
376

(72) 発明者 ベルトツツィ, キャロライン ルース

アメリカ合衆国, カリフォルニア州 94025, メンロー パーク, イェール ロード 220

(72) 発明者 レブリッラ, カルリト バンゲルス

アメリカ合衆国, カリフォルニア州 95618, デヴィス, マロルカ レーン 3035

F ターム(参考) 2G041 CA01 EA04 FA12 GA06 GA09 HA01 JA02 LA07 LA08 LA12

LA20

2G045 AA26 CA25 CA26 CB07 CB08 CB11

4B063 QA01 QA13 QA18 QA19 QQ02 QQ03 QQ67 QQ79 QR16 QS10

QS39