

(19) 日本国特許庁(JP)

## (12) 特許公報(B2)

(11) 特許番号

特許第5160447号  
(P5160447)

(45) 発行日 平成25年3月13日(2013.3.13)

(24) 登録日 平成24年12月21日(2012.12.21)

(51) Int.Cl.

F 1

GO 1 N	33/574	(2006.01)	GO 1 N	33/574	Z N A A
CO 7 K	16/18	(2006.01)	CO 7 K	16/18	
CO 7 K	14/47	(2006.01)	CO 7 K	14/47	
GO 1 N	33/53	(2006.01)	GO 1 N	33/53	N
			GO 1 N	33/574	E

請求項の数 12 (全 102 頁) 最終頁に続く

(21) 出願番号	特願2008-547772 (P2008-547772)
(86) (22) 出願日	平成18年12月21日 (2006.12.21)
(65) 公表番号	特表2009-521692 (P2009-521692A)
(43) 公表日	平成21年6月4日 (2009.6.4)
(86) 国際出願番号	PCT/US2006/062488
(87) 国際公開番号	W02007/076439
(87) 国際公開日	平成19年7月5日 (2007.7.5)
審査請求日	平成21年12月16日 (2009.12.16)
(31) 優先権主張番号	60/753,331
(32) 優先日	平成17年12月22日 (2005.12.22)
(33) 優先権主張国	米国 (US)

(73) 特許権者	391008788 アボット・ラボラトリーズ A B B O T T   L A B O R A T O R I E S アメリカ合衆国 イリノイ州 アボット パーク アボット パーク ロード 10 O
(74) 代理人	110001173 特許業務法人川口國際特許事務所
(74) 代理人	100114188 弁理士 小野 誠
(74) 代理人	100140523 弁理士 渡邊 千尋
(74) 代理人	100119253 弁理士 金山 賢教

最終頁に続く

(54) 【発明の名称】肺がんへの傾向についてのスクリーニングのための方法およびマーカー組合せ

## (57) 【特許請求の範囲】

## 【請求項 1】

- 肺がんの疑いのある被験者の診断を補助する方法であり、  
 a . 被験者から少なくとも 1 つのバイオメトリックパラメーターを得る段階、  
 b . 少なくとも 1 つのバイオメトリックパラメーターを、前記各バイオメトリックパラメーターに対して先に決定したカットオフと比較し、および前記比較に基づいて各バイオメトリックパラメーターに対してスコアを割り当てる段階、  
 c . 被験者から試験試料を得る段階、  
 d . 試験試料において、パネル内の 2 つ以上のバイオマーカーの量を定量する段階であつて、該パネルは、少なくとも 1 つの抗体、少なくとも 1 つの抗原および A c n 6 3 9 9 、  
 A c n 9 4 5 9 、 P u b 1 1 5 9 7 、 P u b 4 7 8 9 、 T F A 2 7 5 9 、 T F A 9 1 3 3  
 、 P u b 3 7 4 3 、 P u b 8 6 0 6 、 P u b 4 4 8 7 、 P u b 4 8 6 1 、 P u b 6 7 9 8  
 、 P u b 6 4 5 3 、 P u b 2 9 5 1 、 P u b 2 4 3 3 、 P u b 1 7 3 3 8 、 T F A 6 4 5  
 3 および H I C 3 9 5 9 からなる群より選択される、少なくとも 1 つの関心領域を含む段階、  
 e . パネル中の定量された各バイオマーカーの量を、前記バイオマーカーに対する所定のカットオフに対して比較し、および前記比較に基づいて各バイオマーカーに対してスコアを割り当てる段階、  
 f . 段階 b にて決定した各バイオメトリックパラメーターに対して割り当てられたスコアを、段階 e にて定量した各バイオマーカーに対して割り当てられたスコアと統合して、前 10  
 20

記被験者の総スコアを見つけ出す段階、

g . 段階 f で決定された総スコアを、所定の総スコアと比較する段階、ならびに

h . 段階 f で決定した総スコアに基づいて、前記被験者が肺がんの危険性を有するかどうかを決定する段階、

を含む、方法。

【請求項 2】

パネルが、抗 - p 5 3 、抗 - T M P 2 1 、抗 - N P C 1 L 1 C - ドメイン、抗 - T M O D 1 、抗 - C A M K 1 、抗 - R G S 1 、抗 - P A C S I N 1 、抗 - R C V 1 、抗 - M A P K A P K 3 および免疫反応性サイクリン E 2 に対する少なくとも 1 つの抗体からなる群より選択される少なくとも 1 つの抗体を含む、請求項 1 の方法。 10

【請求項 3】

パネルが、サイトケラチン 8 、サイトケラチン 19 、サイトケラチン 18 、 C E A 、 C A 1 2 5 、 C A 1 5 - 3 、 S C C 、 p r o G R P 、 C A 1 9 - 9 、血清アミロイド A 、 - 1 - 抗 - トリプシンおよびアポリポタンパク質 C I I I からなる群から選択される少なくとも 1 つの抗原を含む、請求項 1 の方法。

【請求項 4】

D F I が、  $[ ( 1 - \text{感度} )^2 + ( 1 - \text{特異性} )^2 ]^{1/2}$  として定義され、肺がんに関連したバイオマーカーの D F I が、 0 . 4 未満である、請求項 1 の方法。

【請求項 5】

バイオメトリックパラメーターが、被験者の喫煙歴、年齢、発がん物質への暴露および性別からなる群より選択される、請求項 1 の方法。 20

【請求項 6】

バイオメトリックパラメーターが、喫煙のパッケイヤーである、請求項 5 の方法。

【請求項 7】

段階 b が、各バイオメトリックパラメーターの量を、前記バイオメトリックパラメーターに対する所定のカットオフの数に対して比較し、および前記比較に基づいて、各前記バイオメトリックパラメーターに対して可能性あるいくつかのスコアの 1 つを割り当てることを含み、段階 e が、パネル中の各バイオマーカーの量を、前記バイオマーカーに対する所定のカットオフの数に対して比較し、および前記比較に基づいて、各バイオマーカーに対してスコアを割り当てることを含み、段階 f が、段階 b でのバイオメトリックパラメーターに対する割り当てられたスコアを、段階 e で定量された各バイオマーカーに対する割り当てられたスコアと統合して、前記被験者に対する総スコアを見つけ出すことを含み、段階 g が、段階 f で決定した総スコアを、先に決定したいくつかの総スコアと比較することを含み、段階 g が、段階 g で決定した総スコアに基づいて、前記被験者が肺がんを有するかどうかを決定することを含む、請求項 1 の方法。 30

【請求項 8】

肺がんの疑いのある被験者の診断を補助する方法であり、

a . 被験者から試験試料を得る段階、

b . 試験試料において、パネル中の 2 つ以上のバイオマーカーの量を定量する段階であつて、該パネルは、少なくとも 1 つの抗体、少なくとも 1 つの抗原および A c n 6 3 9 9 、 A c n 9 4 5 9 、 P u b 1 1 5 9 7 、 P u b 4 7 8 9 、 T F A 2 7 5 9 、 T F A 9 1 3 3 、 P u b 3 7 4 3 、 P u b 8 6 0 6 、 P u b 4 4 8 7 、 P u b 4 8 6 1 、 P u b 6 7 9 8 、 P u b 6 4 5 3 、 P u b 2 9 5 1 、 P u b 2 4 3 3 、 P u b 1 7 3 3 8 、 T F A 6 4 5 3 および H I C 3 9 5 9 からなる群より選択される、少なくとも 1 つの関心領域を含む段階、 40

c . パネル中で定量された各バイオマーカーの量を、前記バイオマーカーに対する所定のカットオフと比較し、および前記比較に基づいて各バイオマーカーに対してスコアを割り当てる段階、

d . 段階 c において定量された各バイオマーカーに対して割り当てられたスコアを統合して、前記被験者に対する総スコアを見つけ出す段階、 50

e. 段階dにて決定された総スコアを、所定の総スコアと比較する段階、ならびに  
f. 段階eにて決定された総スコアに基づいて、前記被験者が肺がんの危険性を有するかどうかを決定する段階、  
を含む、方法。

【請求項9】

パネルが、抗-p53、抗-TMP21、抗-NPC1L1C-ドメイン、抗-TMO  
D1、抗-CAMK1、抗-RGS1、抗-PACSIN1、抗-RCV1、抗-MAP  
KAPK3および免疫反応性サイクリンE2に対する少なくとも1つの抗体からなる群より選択される少なくとも1つの抗体を含む、請求項8の方法。

【請求項10】

パネルが、サイトケラチン8、サイトケラチン19、サイトケラチン18、CEA、CA  
125、CA19-9、CA15-3、SCC、proGRP、血清アミロイドA、  
-1-抗-トリプシンおよびアポリポタンパク質CIIIからなる群から選択される少なくとも1つの抗原を含む、請求項8の方法。

【請求項11】

DFIが、 $[(1 - \text{感度})^2 + (1 - \text{特異性})^2]^{1/2}$ として定義され、肺がんに関連したバイオマーカーのDFIが、0.4未満である、請求項8の方法。

【請求項12】

段階cが、パネル中の各バイオマーカーの量を、前記バイオマーカーに対する所定のカットオフの数と比較し、および前記比較に基づいて、各バイオマーカーに対してスコアを割り当てるなどを含み、段階dが、段階cにて定量した各バイオマーカーに対して割り当てられたスコアを統合して、前記被験者に対する総スコアを見つけ出すなどを含み、段階eが、段階dで決定した総スコアを、所定の総スコアの数と比較することを含み、ならびに段階fが、段階eで決定した総スコアに基づいて、前記被験者が肺がんを有するかどうかを決定することを含む、請求項8の方法。

【発明の詳細な説明】

【背景技術】

【0001】

肺がんは米国において男性および女性両方にて、第二のもっとも一般的ながんであり、2005年には、推定172,500新規症例が診断されると推定される(American Cancer Society統計)。両性にて、がんの死のもっとも一般的な原因であり、2005年に163,000以上の肺がんが死に関連すると推測される。肺がんはまた、世界の他の地域でも主要な健康問題である。欧州連合において、およそ135,000の新規症例が毎年起こっている。Genesis Report, 1995年2月。また、男性のタバコ消費率が世界でもっとも高い中央および東ヨーロッパにて、罹患率が急速に増加している。T. Reynolds, J. Natl. Cancer Inst. 87: 1348-1349 (1995)。タバコのみで、肺、気管および気管支のがんのすべての症例の、90%以上の原因である。CPMCnet, Guide to Clinical Preventive Services。世界保健機関(World Health Organization)のがん研究の国際機関(International Agency for Research on Cancer)は、2002年に、全世界で、1,352,000症例の肺がんが存在し、1,179,000の死がこの疾患によるものと推定した。

【0002】

早期段階の肺がんは、胸部レントゲン写真および喀痰細胞診によって検出可能であるが、これらの手順は、無症状の個体に対するスクリーニング試験として日常的に使用するほど十分な精度を有していない。胸部レントゲン写真の感度を制限しうる可能性のある技術問題には、準最適技術、不十分な暴露、および患者の位置決めと協力が含まれる。Ann. Intern. Med. 104: 663-670 (1986)。放射線科医師はしばしば、胸部レントゲン写真の解釈に同意せず、これらの40%以上が有意であるか、または

10

20

30

40

50

潜在的に有意である。P. G. Herman, et al., *Chest* 68: 278 - 282 (1975)。偽陰性解釈が、ほとんどのエラーの原因であり、決定的でない結果が、明確化のための追加試験を必要とする。T. G. Tape et al., 上記。

【0003】

喀痰細胞診は、早期肺がんを検出することにおいて、胸部レントゲン写真よりも感度が低い。国立がんセンター (The National Cancer Institute) 共同早期肺がん検出プログラム (Cooperative Early Lung Cancer Detection Program)、Am. Rev. Resp. Dis. 130: 565 - 567 (1984)。肺がんを診断する喀痰細胞診の能力に影響を与える因子には、十分な唾液を産出する患者の能力、腫瘍の大きさ、主要な気道への腫瘍の近さ、腫瘍の組織学的型、および細胞病理学者の経験およびトレーニングが含まれる。R. J. Ginsberg et al., *Cancer: Principles and Practice of Oncology*, Fourth Edition, pp. 673 - 723, Philadelphia, PA: J. B. Lippincott Co. (1993)。

10

【0004】

もっとも新しい肺がんは、この疾患が肺の下に広がっている場合に検出されうる。米国において、新規非小細胞肺がんのたった 16% が、局在化ステージで検出され、5 年生存率は、49.5% もっとも高い。反対に、新規症例の 68% が、それぞれ 18.5% より 20 および 1.8% の 5 年生存率である、疾患が局所的に広く広がっている場合、または遠位部に転移した場合に検出される。同様に、新規検出小細胞肺がんの 80% が、局所浸潤または遠位転移にて発見され、それぞれ 9.5% より 1.7% の 5 年生存率である。Statute, J. Natl. Cancer Inst. 87: 1662 (1995)。これらの統計は、現在の手段が、疾患の早期、治療可能ステージでの肺がんを検出できていないこと、および死亡率を減少させるために、検出および治療の方法の改善を必要とすることを示している。

【0005】

最初の治療後に肺がん患者をモニタリングするためのもっとも頻繁に使用される方法は、病院訪問、胸部 X 線、全血球計数、肝臓機能検査および胸部コンピュータ断層撮影法 (CT) である。しかしながら、定期的モニタリングによる再発の検出は、治療の様式および総合生存時間に大きな影響を与える、これによって、現在のモニタリング法が費用効率が高くないという結論を導く。K. S. Naunheim et al., Ann. Thorac. Surg. 60: 1612 - 1616 (1995)。G. L. Walsh et al., Ann. Thorac. Surg. 60: 1563 - 1572 (1995)。

30

【0006】

最近になり、肺がんに対して高い危険性のある無症状の患者を検査するために、CT の利用が再検証されてきた。C. I. Henschke et al., Clin. Imaging 28: 317 - 321 (2004) は、CTスキャニングが、多くの陽性を産出することなしに、無症状性肺がんを検出可能であることを示唆している 2 つの研究を報告している。J. Gohagen et al., Chest 126: 114 - 121 (2004) は、胸部 X 線を低用量スパイラルコンピュータ断層撮影法 (CT) と比較している無作為研究のための試験プロトコールを評価し、肺がんに対する検査への大規模無作為臨床試験が適していたと結論づけた。しかしながら、臨床診療にて実行する場合でさえ、CTスクリーニングのコストは高く、さらなる検査を導く偽陽性の数は高い。良好な特異性を有する低コスト血液試験が、がんの早期検出に対して CT を補完する。CT の利用を改善するための他の戦略には、早期ステージ肺がんに対する高感度血液試験の利用が関与する。そのような試験が、CT または X 線に対する代替案として患者に提案され、試験が陽性の場合、患者は撮像され、試験が陰性の場合、患者はスキャンされず、将来再試験される。血液試験が高感度または高特異性、または理想的には両方を提供するかどう

40

50

か、そのような試験が、早期ステージ肺がんを検出するために使用される現在のプロトコールに有用性を発見する。

#### 【0007】

さらに、肺がんの危険性のある個体を同定するために、パネル内に統合した場合に、腫瘍マーカーおよびその有用性の最近の再評価が存在する。しかしながら、個々のマーカーの特性である感度の欠如によって、腫瘍マーカーのパネルが、肺がんの早期検出のために有用であることを妨げる。反対に、公知の免疫アッセイマーカー、すなわちCEA、NSEおよびProgRPのパネルが、生検試料を得ることが難しい場合に、肺がんの組織学的を行うのに有用であると知られている。(C. Gruber et al., *Tumor Biology* 27 (Supplement 1) : 71 (2006) およびP. Stieber et al., *Tumor Biology*, 27 (Supplement 2) : S5-4 (2006)).

10

#### 【0008】

正常肺組織と比較して、肺腫瘍組織中で異なった形で発現した細胞成分をまず同定することによって、肺がんに対する改善された腫瘍マーカーを発見する試みがなされてきた。二次元ポリアクリルアミドゲル電気泳動が、ポリペプチド組成中の定量的および定性的差違を特性化するために使用されてきた。T. Hirano et al., *Br. J. Cancer* 72 : 840-848 (1995)。A. T. Endler et al., *J. Clin. Chem. Clin. Biochem.* 24 : 981-992 (1986)。しかしながら、本技術の感度は、2つの電気泳動段階のタンパク質分解の程度によって、そしてゲル中のタンパク質の染色に依存する検出段階によって制限される。また、ポリペプチドの不安定性が、二次元パターンにて不自然な結果を産出しうる。

20

#### 【0009】

イースタン バージニア メディカル スクール (Eastern Virginia Medical School) による、国際公開番号WO2005/098445 A-2 にて記述されたもののような、バイオマーカーおよび肺がんの診断での補助におけるこれら利用を同定する試みも行われてきた。国際公開番号WO2005/098445 にて議論されたバイオマーカーは、表面増強レーザー脱着 / イオン化質量分析 (SELDI) を用いて同定された。種々のマーカー、キット、方法およびディシジョンツリー解析法が開示されている。しかしながら、これらのマーカー、キットおよび方法は、これらのマーカーおよび方法が、任意の研究室にて再現できなかつたので、日常の診療での使用に適合できなかつた。

30

#### 【0010】

疾患を患う、または患っていない個体からの血清で、酵母または細菌中で発現されたペプチドライブラーを調査することによって、肺がんに特異的な免疫応答を発見する試みも行われてきた。Hirschowitz (L. Zhong et al., *Chest* 125 : 105-106 (2004)、L. Zhong et al., *Am. J. Respir. Crit. Care Med.* 15 : 1308-1314 (2005)) の研究所からの発行物が、肺がんの患者に対する自己抗原であるタンパク質を発見するための、ファージドライブラーの利用を記述した。著者は、対照試験にて、症候性および無症候性両方の肺がん患者の良好な同定について報告した。しかしながら、症例数および対照が制限されており (< 200 総被験者)、本方法は、非常の多くの集団において正当性を立証する必要がある。

40

#### 【0011】

現在、肺がんの危険性がある個体の同定は、個体の喫煙歴に大きく基づいている。アスベスト、微粒子などのような他の環境暴露がまた、肺がんを発現させる危険性を増加させる。これらの公知の危険性因子が、1つ以上のアルゴリズム内に統合され、臨床医が利用可能であり、肺がんの個体の危険性を査定するために公的である (P. B. Bach et al., *J. Natl. Cancer Inst.* 95 : 470-478 (2003))。残念なことに、本アルゴリズムは、早期ステージ肺がんの検出のために有用である

50

ほど十分感度がよくなく、特異的でない。実際、引用されたアルゴリズムに基づいて、明らかな喫煙歴を有する個体が、肺がんを発現される、1/500から1/100の相対危険性を有する。これは、Bach等の方法を用いても、500のCTスキャンのうち499が、肺がんの症例の発見にいたらないことを意味する。

【0012】

その結果、実施するために迅速で、簡便で、費用効率が高い肺がんを検出するために有用な方法およびマーカーに対して、当技術分野で必要性が残っている。肺がんを発現させる、患者の可能性ある素因または危険性を示唆するために使用可能な、特定の方法およびマーカーを提供する。そのような方法には、肺がんを示唆するバイオマーカーについて試料を検査し、そのようなマーカーを検出するための方法が含まれうる。そのような方法には、マーカーに対して、生体試料のマススペクトルを解析する、または試料をアッセイし、ついでバイオマーカーを、肺がんの指標として、または肺がんを発現させる危険性として検出するための改善方法が含まれうる。

10

【発明の開示】

【0013】

(発明の要約)

本発明は部分的に、肺がんを患っていることが疑われる被験者中の肺がんの検出において補助するための、迅速で感度のよい方法が、バイオマーカーおよびバイオマーカーとバイオメトリックパラメーターの特定の組合せに基づきうるという発見に基づく。

【0014】

20

1つの態様において、本方法には、

- a . 被験者より試験試料を得る段階
- b . 試験試料中、パネル中の1つ以上のバイオマーカーの量を定量する段階、
- c . 前記バイオマーカーに対して、所定のカットオフと、パネル中の各バイオマーカーの量を比較し、前記比較に基づいた各バイオマーカーに対するスコアを割り当てる段階、
- d . 段階cにおいて決定された各バイオマーカーに対する割り当てられたスコアを比較して、前記被験者に関する総スコアを見つけ出す段階、
- e . 段階dにて決定された総スコアを、所定の総スコアと比較する段階、および
- f . 前記被験者が総スコアに基づく肺がんの危険性を有するかどうか決定する段階、

が含まれうる。

30

【0015】

上記方法において、肺がんに関連するバイオマーカーのDFIは、好ましくは約0.4未満である。

【0016】

場合により、上記方法はさらに、被験者より少なくとも1つのバイオメトリックパラメーターを得る段階を含み得る。得ることが可能なバイオメトリックパラメーターの例は、被験者の喫煙歴である。上記方法がさらに、被験者から少なくとも1つのバイオメトリックパラメーターを得る段階を含む場合、本方法にはさらに、少なくとも1つのバイオメトリックパラメーターを、各前記バイオメトリックパラメーターに対する所定のカットオフに対して比較する段階、および前記比較に基づいた各バイオメトリックパラメーターに対してスコアを割り当てる段階、各バイオメトリックパラメーターに対する割り当てられたスコアを、段階cにて定量した各バイオマーカーに対する割り当てられたスコアと統合して、段階dでの前記被験者に対する総スコアを見つけ出す段階、総スコアを、段階eでの所定の総スコアと比較する段階、および前記被験者が、段階f中の総スコアに基づいて、肺がんの危険性を有するかどうかを決定する段階、が含まれうる。

40

【0017】

上記方法にて定量可能なバイオマーカーの例には、抗体、抗原および関心領域の群から選択される1つ以上のバイオマーカーがあげられる。より具体的には、定量可能なバイオマーカーには、限定はしないが、1つ以上の、抗-p53、抗-TMP21、抗-ニーマン-ピックC1-様タンパク質1、C末端ペプチド) - ドメイン(抗-NPC1L1C -

50

ドメイン)、抗-TMOD1、抗-CAMK1、抗-RGS1、抗-PACsin1、抗-RCV1、抗-MAPKAPK3、免疫反応性サイクリンE2に対する少なくとも1つの抗体、サイトケラチン8、サイトケラチン19、サイトケラチン18、CEA、CA125、CA15-3、SCC、CA19-9、proGRP、血清アミロイドA、-1-抗-トリプシン、アポリポタンパク質CII、Acn6399、Acn9459、Pub11597、Pub4789、TFA2759、TFA9133、Pub3743、Pub8606、Pub4487、Pub4861、Pub6798、Pub6453、Pub2951、Pub2433、Pub17338、TFA6453およびHIC3959が含まれる。

## 【0018】

10

他の態様において、方法には、

- a. 被験者から少なくとも1つのバイオメトリックパラメーターを得る段階、
- b. 少なくとも1つのバイオメトリックパラメーターを、前記各バイオメトリックパラメーターに関して、所定のカットオフと比較し、前記比較に基づいた各バイオメトリックパラメーターに対してスコアを割り当てる段階、
- c. 被験者から試験試料を得る段階、
- d. 試験試料中、パネル内の2つ以上のバイオマーカーの量を定量し、パネルは、少なくとも1つの抗体および少なくとも1つの抗原を含む段階、
- e. パネル中で定量された各バイオマーカーの量を、前記バイオマーカーに関する所定のカットオフに対して比較する段階と、前記比較に基づいて各バイオマーカーに対してスコアを割り当てる段階、
- f. 段階bにて決定した各バイオメトリックパラメーターに対する割り当てられたスコアを、段階eにて定量した各バイオマーカーに対する割り当てられたスコアと統合して、前記被験者の総スコアを見つけ出す段階、
- g. 段階fで決定された総スコアを、所定の総スコアと比較する段階、
- h. 前記被験者が、段階fで決定した総スコアに基づいて、肺がんの危険性を有するかどうかを決定する段階、

が含まれる。

## 【0019】

30

上記方法において、肺がんに関連するバイオマーカーのDFIは、好ましくは約0.4未満である。

## 【0020】

上記方法において、パネルは、抗-p53、抗-TMP21、抗-NPC1L1C-ドメイン、抗-TMOD1、抗-CAMK1、抗-RGS1、抗-PACsin1、抗-RCV1、抗-MAPKAPK3および免疫反応性サイクリンE2に対する少なくとも1つの抗体からなる群より選択される少なくとも1つの抗体、およびサイトケラチン8、サイトケラチン19、サイトケラチン18、CEA、CA125、CA15-3、SCC、CA19-9、proGRP、血清アミロイドA、-1-抗-トリプシンおよびアポリポタンパク質CIIからなる群から選択される少なくとも1つの抗原を含み得る。

## 【0021】

40

上記方法において、被験者から得たバイオメトリックパラメーターは、被験者の喫煙歴、年齢、発がん物質への暴露および性別からなる群より選択される。好ましくは、バイオメトリックパラメーターは、経験者の喫煙のパッケイヤーである。

## 【0022】

場合により、方法には、試料試験中の少なくとも1つの関心領域を定量することが含まれる。関心領域が、試験試料中で定量される場合、パネルにはさらに、Acn6399、Acn9459、Pub11597、Pub4789、TFA2759、TFA9133、Pub3743、Pub8606、Pub4487、Pub4861、Pub6798、Pub6453、Pub2951、Pub2433、Pub17338、TFA6453およびHIC3959からなる群から選択される、少なくとも1つの関心領域を含み

50

得る。

【0023】

場合により、上記方法はまた、患者が肺がんを発現させる危険性があるかどうかを決定するため、分割加重スコア法 (S p l i t a n d W e i g h t e d S c o r i n g M e t h o d) を利用可能である。上記方法が、そのような分割加重スコア法を利用する場合、前記方法にて、段階 b が、少なくとも 1 つのバイオメトリックパラメーターを、前記バイオメトリックパラメーターに対する所定のカットオフの数に対して比較する段階、および前記比較に基づいて、各前記バイオメトリックパラメーターに対してスコアを割り当てる段階、を含み、段階 e が、パネル中の各バイオマーカーの量を、前記バイオマーカーに対する所定のカットオフの数に対して比較する段階、および前記比較に基づいて、各バイオマーカーに対してスコアを割り当てる段階を含み、段階 f が、段階 b での各バイオメトリックパラメーターに対する割り当てられたスコアを、段階 e で定量された各バイオマーカーに対する割り当てられたスコアと統合して、前記被験者に対する総スコアを見つけ出す段階を含み、段階 g が、段階 f で決定した総スコアを、所定の総スコアの数と比較する段階を含み、段階 h が、前記被験者が、段階 g で決定した総スコアに基づいて、肺がんを有するかどうかを決定する段階を含む。

10

【0024】

他の態様において、方法には、

- a . 被験者から試験試料を得る段階、
- b . パネル中の 2 つ以上のバイオマーカーの量を、試験試料中で定量し、パネルは、少なくとも 1 つの抗体と少なくとも 1 つの抗原を含む段階、
- c . パネル中で定量された各バイオマーカーの量を、前記バイオマーカーに対して所定のカットオフと比較し、前記比較に基づいて各バイオマーカーに対してスコアを割り当てる段階、
- d . 段階 c において定量された各バイオマーカーに対して割り当てたスコアを統合して、前記被験者に対する総スコアを見つけ出す段階、
- e . 段階 d にて決定された総スコアを、所定の総スコアと比較する段階、
- f . 前記被験者が、段階 e にて決定された総スコアに基づいて、肺がんの危険性を有するかどうかを決定する段階、

が含まれうる。

30

【0025】

上記方法において、肺がんに関連するバイオマーカーの D F I は、好ましくは約 0 . 4 未満である。

【0026】

上記方法において、パネルには、抗 - p 5 3 、抗 - T M P 2 1 、抗 - N P C 1 L 1 C - ドメイン、抗 - T M O D 1 、抗 - C A M K 1 、抗 - R G S 1 、抗 - P A C S I N 1 、抗 - R C V 1 、抗 - M A P K A P K 3 および免疫反応性サイクリン E 2 に対する少なくとも 1 つの抗体からなる群より選択される少なくとも 1 つの抗体が含まれうる。パネルには、サイトケラチン 8 、サイトケラチン 1 9 、サイトケラチン 1 8 、C E A 、C A 1 2 5 、C A 1 5 - 3 、S C C 、C A 1 9 - 9 、p r o G R P 、血清アミロイド A 、- 1 - 抗 - トリプシンおよびアポリポタンパク質 C I I I からなる群から選択される少なくとも 1 つの抗原が含まれうる。

40

【0027】

場合により、方法にはさらに、試験試料中、少なくとも 1 つの関心領域を定量化することが含まれうる。関心領域が定量化される場合、パネルにはさらに、A c n 6 3 9 9 、A c n 9 4 5 9 、P u b 1 1 5 9 7 、P u b 4 7 8 9 、T F A 2 7 5 9 、T F A 9 1 3 3 、P u b 3 7 4 3 、P u b 8 6 0 6 、P u b 4 4 8 7 、P u b 4 8 6 1 、P u b 6 7 9 8 、P u b 6 4 5 3 、P u b 2 9 5 1 、P u b 2 4 3 3 、P u b 1 7 3 3 8 、T F A 6 4 5 3 およびH I C 3 9 5 9 からなる群より選択される、少なくとも 1 つの関心領域が含まれうる。

50

## 【0028】

場合により、上記方法はまた、患者が肺がんを発現させる危険性があるかどうかを決定するため、分割加重スコア法 (S p l i t a n d W e i g h t e d S c o r i n g M e t h o d) を利用可能である。上記方法が、そのような分割加重スコア法を利用する場合、前記方法にて、段階 c が、パネル中の各バイオマーカーの量を、前記バイオマーカーに関して所定のカットオフの数と比較する段階、および前記比較に基づいて、各バイオマーカーに対してスコアを割り当てる段階、を含み、段階 d が、段階 c にて定量した各バイオマーカーに対して割り当てられたスコアを統合して、前記被験者に対する総スコアを見つけ出す段階を含み、段階 e が、段階 d で決定した総スコアを、所定の総スコアの数と比較する段階を含み、段階 f が、段階 e で決定した総スコアに基づいて、前記被験者が肺がんを有するかどうかを決定する段階を含む。10

## 【0029】

他の態様において、方法には、

- a . 被験者から試験試料を得る段階、
- b . パネル中、少なくとも 1 つのバイオマーカーの量を、試験試料中で定量し、パネルが、免疫反応性サイクリン E 2 に対する少なくとも 1 つの抗体を含む段階、
- c . パネル中で定量した各バイオマーカーの量を、前記バイオマーカーに対する所定のカットオフに対して比較する段階、および前記比較に基づいて各バイオマーカーに対してスコアを割り当てる段階、
- d . 段階 c にて定量した各バイオマーカーに対する割り当てられたスコアを統合して、前記被験者に対する総スコアを見つけ出す段階、20
- e . 段階 d にて決定した総スコアを、所定の総スコアと比較する段階、
- f . 段階 e で決定した総スコアに基づいて、前記被験者が肺がんを有するかどうか決定する段階、

を含み得る。

## 【0030】

上記方法において、肺がんに関連するバイオマーカーの D F I は、好ましくは約 0 . 4 未満である。

## 【0031】

場合により、上記方法にはさらに、試験試料中の少なくとも 1 つの抗原を定量する段階、試験試料中の少なくとも 1 つの抗体を定量する段階、または試験試料中の少なくとも 1 つの抗原と少なくとも 1 つの抗体の組合せを定量する段階、が含まれうる。その結果、少なくとも 1 つの抗原、少なくとも 1 つの抗体、または少なくとも 1 つの抗原と少なくとも 1 つの抗体の組合せが、試験試料中で定量されるべき場合、パネルにはさらに、サイトケラチン 8、サイトケラチン 19、サイトケラチン 18、C E A、C A 1 2 5、C A 1 5 - 3、S C C、C A 1 9 - 9、p r o G R P、血清アミロイド A、- 1 - 抗 - トリプシンおよびアボリポタンパク質 C I I I からなる群から選択される少なくとも 1 つの抗原、抗 - p 5 3、抗 - T M P 2 1、抗 - N P C 1 L 1 C - ドメイン、抗 - T M O D 1、抗 - C A M K 1、抗 - R G S 1、抗 - P A C S I N 1、抗 - R C V 1、抗 - M A P K A P K 3 および免疫反応性サイクリン E 2 に対する少なくとも 1 つの抗体からなる群より選択される少なくとも 1 つの抗体、またはそれらの任意の組合せが含まれる。3040

## 【0032】

場合により、方法にはさらに、試験試料中、少なくとも 1 つの関心領域を定量化することが含まれうる。関心領域が定量化される場合、パネルにはさらに、A c n 6 3 9 9、A c n 9 4 5 9、P u b 1 1 5 9 7、P u b 4 7 8 9、T F A 2 7 5 9、T F A 9 1 3 3、P u b 3 7 4 3、P u b 8 6 0 6、P u b 4 4 8 7、P u b 4 8 6 1、P u b 6 7 9 8、P u b 6 4 5 3、P u b 2 9 5 1、P u b 2 4 3 3、P u b 1 7 3 3 8、T F A 6 4 5 3 およびH I C 3 9 5 9 からなる群より選択される、少なくとも 1 つの関心領域が含まれうる。

## 【0033】

場合により、上記方法はまた、患者が肺がんを発現させる危険性があるかどうかを決定するため、分割加重スコア法 (S p l i t a n d W e i g h t e d S c o r i n g M e t h o d) を利用可能である。上記方法が、そのような分割加重スコア法を利用する場合、前記方法にて、段階 c が、パネル中の各バイオマーカーの量を、前記バイオマーカーに対する所定のカットオフの数に対して比較して、前記比較に基づいて、各バイオマーカーに対してスコアを割り当てる段階を含み、段階 d が、段階 c にて定量した各バイオマーカーに対する割り当てられたスコアを統合して、前記被験者に対する総スコアを見つけ出す段階、を含み、段階 e が、段階 d で決定した総スコアを、所定の総スコアの数と比較する段階を含み、段階 f が、段階 e で決定された総スコアに基づいて、前記被験者が肺がんを有するかどうかを決定する段階を含む。

10

#### 【 0 0 3 4 】

場合により、上記方法はさらに、被験者から少なくとも 1 つのバイオメトリックパラメーターを得る段階を含み得る。被験者から得られるバイオメトリックパラメーターは、被験者の喫煙歴、年齢、発がん物質への暴露および性別からなる群より選択される。好ましくは、バイオメトリックパラメーターは、被験者の喫煙のパックイヤーである。上記方法がさらに、被験者から少なくとも 1 つのバイオメトリックパラメーターを得る段階を含む場合、方法にはさらに、少なくとも 1 つのバイオメトリックパラメーターを、各前記バイオメトリックパラメーターに対して所定のカットオフに対して比較する段階、および前記比較に基づいて、各バイオメトリックパラメーターに対してスコアを割り当てる段階、各バイオメトリックパラメーターに対して割り当てられたスコアを、段階 c にて定量された各バイオマーカーに対して割り当てられたスコアと統合して、前記被験者中の総スコアを見つけ出す段階と、総スコアを、段階 e での所定の総スコアと比較する段階、および段階 f での総スコアに基づいて、前記被験者が肺がんの危険性を有するかどうかを決定する段階、が含まれうる。

20

#### 【 0 0 3 5 】

他の態様において、方法には、

- a . 被験者から試験試料を得る段階、
- b . パネル中、少なくとも 1 つのバイオマーカーを試験試料中で定量し、パネルには、サイトケラチン 8、サイトケラチン 19、サイトケラチン 18、C E A、C A 1 2 5、C A 1 5 - 3、S C C、C A 1 9 - 9、p r o G R P、血清アミロイド A、- 1 - 抗 - トリプシンおよびアポリポタンパク質 C I I I からなる群より選択される少なくとも 1 つのバイオマーカーを含む段階、
- c . パネル中で定量した各バイオマーカーの量を、前記バイオマーカーに対する所定のカットオフと比較する段階、および前記比較に基づいて、各バイオマーカーに対してスコアを割り当てる段階、
- d . 段階 c にて定量した各バイオマーカーに対する割り当てたスコアを統合し、前記被験者に対する総スコアを見つけ出す段階、
- e . 段階 d にて定量した総スコアを、所定の総スコアと比較する段階、および
- f . 総スコアに基づいて、前記被験者が肺がんを有するかどうか決定する段階、が含まれうる。

30

#### 【 0 0 3 6 】

上記方法において、肺がんに関連するバイオマーカーの D F I は、好ましくは約 0 . 4 未満である。

#### 【 0 0 3 7 】

場合により、上記方法はさらに、試験試料中、少なくとも 1 つの抗体を定量する段階が含まれうる。その結果、パネルにはさらに、抗 - p 5 3、抗 - T M P 2 1、抗 - N P C 1 L 1 C - ドメイン、抗 - T M O D 1、抗 - C A M K 1、抗 - R G S 1、抗 - P A C S I N 1、抗 - R C V 1、抗 - M A P K A P K 3、および免疫反応性サイクリン E 2 に対する少なくとも 1 つの抗体、または任意のこれらの組合せからなる群より選択される少なくとも 1 つの抗体が含まれうる。

40

50

## 【0038】

場合により、方法にはさらに、試験試料中、少なくとも1つの関心領域を定量する段階が含まれる。関心領域が定量される場合、パネルにはさらに、A c n 6 3 9 9、A c n 9 4 5 9、P u b 1 1 5 9 7、P u b 4 7 8 9、T F A 2 7 5 9、T F A 9 1 3 3、P u b 3 7 4 3、P u b 8 6 0 6、P u b 4 4 8 7、P u b 4 8 6 1、P u b 6 7 9 8、P u b 6 4 5 3、P u b 2 9 5 1、P u b 2 4 3 3、P u b 1 7 3 3 8、T F A 6 4 5 3およびH I C 3 9 5 9からなる群より選択される少なくとも1つの関心領域が含まれる。

## 【0039】

場合により、上記方法はまた、患者が肺がんを発現させる危険性があるかどうかを決定するに、分割加重スコア法 (S p l i t a n d W e i g h t e d S c o r i n g M e t h o d) を利用可能である。上記方法が、そのような分割加重スコア法を利用する場合、前記方法にて、段階 c が、パネル中の各バイオマーカーの量を、前記バイオマーカーに対する所定のカットオフの数を比較する段階、および前記比較にもとづいて、各バイオマーカーに対してスコアを割り当てる段階を含み、段階 d が、段階 c にて定量された各バイオマーカーに対して割り当てられたスコアを統合して、前記被験者に対する総スコアを見つける段階を含み、段階 e は、段階 d で決定された総スコアを、所定の総スコアの数を比較する段階を含み、段階 f は、段階 e にて決定された総スコアに基づいて、前記被験者が肺がんを有するかどうかを決定する段階を含む。

## 【0040】

場合により、上記方法はさらに、被験者から少なくとも1つのバイオメトリックパラメーターを得る段階を含み得る。被験者から選られるバイオメトリックパラメーターは、被験者の喫煙歴、年齢、発がん物質への暴露および性別からなる群より選択される。好ましくは、バイオメトリックパラメーターは、喫煙のパックイヤーである。上記方法がさらに、被験者から少なくとも1つのバイオメトリックパラメーターを得る段階を含む場合、方法にはさらに、少なくとも1つのバイオメトリックパラメーターを、各前記バイオメトリックパラメーターに対して所定のカットオフに対して比較する段階、および前記比較に基づいて、各バイオメトリックパラメーターに対してスコアを割り当てる段階、各バイオメトリックパラメーターに対して割り当てられたスコアと統合して、前記被験者中の総スコアを見つける段階と、総スコアを、段階 e での所定の総スコアと比較する段階、および段階 f での総スコアに基づいて、前記被験者が肺がんの危険性を有するかどうかを決定する段階、が含まれる。

## 【0041】

他の態様において、方法には、

- a . 被験者から試験試料を得る段階、
  - b . パネル中、少なくとも1つのバイオマーカーを、試験試料中で定量し、パネルは、少なくとも1つのバイオマーカーを含み段階で、バイオマーカーは、A c n 6 3 9 9、A c n 9 4 5 9、P u b 1 1 5 9 7、P u b 4 7 8 9、T F A 2 7 5 9、T F A 9 1 3 3、P u b 3 7 4 3、P u b 8 6 0 6、P u b 4 4 8 7、P u b 4 8 6 1、P u b 6 7 9 8、P u b 6 4 5 3、P u b 2 9 5 1、P u b 2 4 3 3、P u b 1 7 3 3 8、T F A 6 4 5 3およびH I C 3 9 5 9からなる群より選択される、関心領域である、
  - c . パネル中で定量した各バイオマーカーの量を、前記バイオマーカーに対して所定のカットオフと比較する段階、および前記比較に基づいて、各バイオマーカーに対してスコアを割り当てる段階、
  - d . 段階 c で定量した各バイオマーカーに対して割り当てられたスコアを統合して、前記被験者に対する総スコアを見つける段階、
  - e . 段階 d にて定量した総スコアを、所定の総スコアと比較する段階、
  - f . 段階 e にて決定した総スコアに基づいて、前記被験者が肺がんを有するかどうかを決定する段階、
- が含まれる。

10

20

30

40

50

## 【0042】

上記方法において、肺がんに関連するバイオマーカーのDFIは、好ましくは約0.4未満である。

## 【0043】

場合により、上記方法にはさらに、試験試料中の少なくとも1つの抗原を定量する段階、試験試料中の少なくとも1つの抗体を定量する段階、または試験試料中の少なくとも1つの抗原と少なくとも1つの抗体の組合せを定量する段階、が含まれうる。その結果、少なくとも1つの抗原、少なくとも1つの抗体、または少なくとも1つの抗原または抗体の組合せが、試験試料中で定量されるべき場合、パネルにはさらに、サイトケラチン8、サイトケラチン19、サイトケラチン18、CEA、CA125、CA15-3、SCC、CA19-9、proGRP、血清アミロイドA、-1-抗-トリプシンおよびアボリポタンパク質C111からなる群から選択される少なくとも1つの抗原、抗-p53、抗-TMP21、抗-NPC1L1C-ドメイン、抗-TMOD1、抗-CAMK1、抗-RGS1、抗-PACSIN1、抗-RCV1、抗-MAPKAPK3および免疫反応性サイクリンE2に対する少なくとも1つの抗体からなる群より選択される少なくとも1つの抗体、またはそれらの任意の組合せが含まれる。

10

## 【0044】

場合により、上記方法はまた、患者が肺がんを発現させる危険性があるかどうかを決定するために、分割加重スコア法(Split and Weighted Scoring Method)を利用可能である。上記方法が、そのような分割加重スコア法を利用する場合、前記方法にて、段階cが、パネル中の各バイオマーカーの量を、前記バイオマーカーに対する所定のカットオフの数を比較する段階、および前記比較にもとづいて、各バイオマーカーに対してスコアを割り当てる段階を含み、段階dが、段階cにて定量された各バイオマーカーに対して割り当てられたスコアを統合して、前記被験者に対する総スコアを見つけ出す段階を含み、段階eは、段階dで決定された総スコアを、所定の総スコアの数を比較する段階を含み、段階fは、段階eにて決定された総スコアに基づいて、前記被験者が肺がんを有するかどうかを決定する段階を含む。

20

## 【0045】

場合により、上記方法はさらに、被験者から少なくとも1つのバイオメトリックパラメーターを得る段階を含み得る。被験者から選られうるバイオメトリックパラメーターは、被験者の喫煙歴、年齢、発がん物質への暴露および性別からなる群より選択される。好ましくは、バイオメトリックパラメーターは、喫煙のパックイヤーである。上記方法はさらに、被験者から少なくとも1つのバイオメトリックパラメーターを得る段階を含む場合、方法にはさらに、少なくとも1つのバイオメトリックパラメーターを、各前記バイオメトリックパラメーターに対して所定のカットオフに対して比較する段階、および前記比較に基づいて、各バイオメトリックパラメーターに対して割り当たされたスコアを、段階cにて定量された各バイオマーカーに対して割り当たされたスコアと統合して、前記被験者中の総スコアを見つけ出す段階と、総スコアを、段階eでの所定の総スコアと比較する段階、および段階fでの総スコアに基づいて、前記被験者が肺がんの危険性を有するかどうかを決定する段階、が含まれうる。

30

## 【0046】

他の態様において、発明は、

- a. 被験者から試験試料を得る段階、
- b. パネル中、2つ又はそれ以上のバイオマーカーの量を試験試料中で定量し、パネルは、サイトケラチン19、サイトケラチン18、CA19-9、CEA、CA15-3、CA125、SCC、proGRP、ACN9459、Pub11597、Pub4789、TFA2759、TFA9133、Pub3743、Pub8606、Pub4487、Pub4861、Pub6798、Tfa6453およびHic3959の2つ以上を含む段階、

40

50

c. パネル中の各バイオマーカーの量を、前記バイオマーカーに対する所定のカットオフに対して比較し、前記比較に基づいて、各バイオマーカーに対してスコアを割り当てる段階、

d. 段階cにて決定された各バイオマーカーに対する割り当てられたスコアを統合し、前記被験者に対する総スコアを見つける段階、

e. 段階dで決定された総スコアを、所定の総スコアと比較する段階、および

f. 段階eにて決定された総スコアに基づいて、前記被験者が肺がんを有するかどうかを決定する段階、

を含み得る。

【0047】

10

上記方法において、肺がんに関連するバイオマーカーのDFIは、好ましくは約0.4未満である。

場合により、上記方法でのパネルは、サイトケラチン19、CEA、ACN9459、Pub11597、Pub4789およびTFA2759、サイトケラチン19、CEA、ACN9459、Pub11597、Pub4789、TFA2759およびTFA9133、サイトケラチン19、CA19-9、CEA、CA15-3、CA125、SCC、サイトケラチン18およびProGRP、Pub11597、Pub3743、Pub8606、Pub4487、Pub4861、Pub6798、Tfa6453およびHic3959またはサイトケラチン19、CEA、CA125、SCC、サイトケラチン18、ProGRP、ACN9459、Pub11597、Pub4789、TFA2759、TFA9133を含み得る。

20

【0048】

30

場合により、上記方法はまた、患者が肺がんを発現させる危険性があるかどうかを決定するために、分割加重スコア法(Split and Weighted Scoring Method)を利用可能である。上記方法が、そのような分割加重スコア法を利用する場合、前記方法にて、段階cが、パネル中の各バイオマーカーの量を、前記バイオマーカーに対する所定のカットオフの数と比較する段階、および前記比較に基づいて各バイオマーカーに対してスコアを割り当てる段階、を含み、段階dが、段階cにて定量した各バイオマーカーに対する割り当てられたスコアを統合して、前記被験者に対する総スコアを見つける段階を含み、段階eが、段階dにて決定した総スコアを、所定の総スコアの数と比較することを含む段階、および段階fが、段階eにて決定された総スコアに基づいて、全被験者が肺がんを有するかどうかを決定する段階を含む。

【0049】

40

本発明は、以上で記述した方法にて使用される種々の異なるキットにも関する。1つの態様において、キットは、配列番号1、配列番号3、配列番号4、配列番号5またはこれらの任意の組合せからなる群より選択されるペプチドを含み得る。他の態様において、キットは、免疫反応性サイクリンE2に対する少なくとも1つの抗体、またはそれらの任意の組合せを含み得る。さらなる態様において、キットは、(a)試験試料中1つ以上の抗原を定量するための、少なくとも1つの抗体を含む試薬で、前記抗原は、サイトケラチン8、サイトケラチン19、サイトケラチン18、CEA、CA125、CA15-3、SCC、CA19-9、ProGRP、血清アミロイドA、-1-抗-トリプシンおよびアポリポタンパク質CIIである、(b)試験試料中の少なくとも1つの抗体を定量するための、1つ以上の抗体を含む試薬で、前記抗体は、抗-p53、抗-TMP21、抗-NPC1L1C-ドメイン、抗-TMOD1、抗-CAMK1、抗-RGS1、抗-PAC SIN1、抗-RCV1、抗-MAPKAPK3および免疫反応性サイクリンE2に対する少なくとも1つの抗体である、(c)ACN9459、Pub11597、Pub4789、TFA2759、TFA9133、Pub3743、Pub8606、Pub4487、Pub4861、Pub6798、Tfa6453およびHIC3959からなる群から選択される、1つ以上の関心領域を定量するための試薬、および(d)試験試料中で定量した各抗原、抗体および関心領域の量を、所定のカットオフに対して比較す

50

ること、および前記比較に基づいて、定量された各抗原、抗体および関心領域に対してスコアを割り当てること、定量した各抗原、抗体および関心領域に対する割り当てられたスコアを統合して、総スコアを得ること、総スコアを、所定の総スコアと比較すること、および前記比較を被験者が肺がんを有するかどうかを決定することにおける補助として使用することのための1つ以上のアルゴリズム、を含み得る。またさらに他の態様において、キットは、(a)試験試料中、1つ以上の抗原を定量するための、少なくとも1つの抗体を含む試薬で、前記抗原は、サイトケラチン19、サイトケラチン18、CA19-9、CEA、CA15-3、CA125、SCCおよびproGRPである、(b)Acn9459、Pub11597、Pub4789、TFA2759、TFA9133、Pub3743、Pub8606、Pub4487、Pub4861、Pub6798、Tfa6453およびHic3959からなる群から選択される、1つ以上の関心領域を定量するための試薬、および

(c)試験試料にて定量された、各抗原、抗体および関心の域の量を統合し、所定のカットオフに対して比較するため、前記比較に基づいて定量した、各抗原およびバイオマーカーに対してスコアを割り当てるため、総スコアを得るために定量した各抗原および関心領域に対して割り当てられたスコアを統合すること、総スコアを所定の総スコアと比較すること、および被験者が肺がんを有するかどうか決定することにおける補助として、前記比較を用いるための、1つ以上のアルゴリズム、を含み得る。定量可能な抗原および関心領域の例は、(a)サイトケラチン19、CEAおよびACN9459、Pub11597、Pub4789およびTfa2759、(b)サイトケラチン19およびCEAおよびACN9459、Pub11597、Pub4789、Tfa2759およびTfa9133、および(c)サイトケラチン19、CEA、CA125、SCC、サイトケラチン18、およびProGRPおよびACN9459、Pub11597、Pub4789およびTfa2759である。他の態様において、キットは、(a)試験試料中、1つ以上の抗原を定量するための、少なくとも1つの抗体を含む試薬で、前記抗原は、サイトケラチン19、サイトケラチン18、CA19-9、CEA、CA15-3、CA125、SCCおよびproGRPである、および(b)試験試料にて定量された、各抗原の量を統合し、所定のカットオフに対して比較するため、および前記比較に基づいて定量した、各抗原に対してスコアを割り当てるため、総スコアを得るために定量した各抗原に対して割り当てられたスコアを統合すること、総スコアを所定の総スコアと比較すること、および被験者が肺がんを有するかどうか決定することにおける補助として、前記比較を用いるための、1つ以上のアルゴリズム、を含み得る。キットを用いて定量可能な抗体の例は、サイトケラチン19、サイトケラチン18、CA19-9、CEA、CA15-3、CA125、SCCおよびProGRPである。他の態様において、キットは、(a)1つ以上のバイオマーカーを定量するための試薬で、前記バイオマーカーが、Acn9459、Pub11597、Pub4789、TFA2759、TFA9133、Pub3743、Pub8606、Pub4487、Pub4861、Pub6798、Tfa6453およびHic3959からなる群から選択される、関心領域である、および(b)試験試料にて定量された、各バイオマーカーの量を統合し、所定のカットオフに対して比較するため、および前記比較に基づいて定量した、各バイオマーカーに対してスコアを割り当てるため、総スコアを得るために定量した各バイオマーカーに対して割り当てられたスコアを統合すること、総スコアを所定の総スコアと比較すること、および被験者が肺がんを有するかどうか決定することにおける補助として、前記比較を用いるための、1つ以上のアルゴリズム、を含む。キットを使用して定量されうる関心領域の例は、Pub11597、Pub3743、Pub8606、Pub4487、Pub4861、Pub6798、Tfa6453およびHic3959からなる群から選択可能である。

#### 【0050】

本発明は、単離または精製ポリペプチドにも関する。本発明によって企図される単離または精製ポリペプチドは、(a)配列番号3および配列番号3のアミノ酸配列に対して60%相同性を有するポリペプチドからなる群より選択されるアミノ酸配列を有する(含む

10

20

30

40

50

) 単離または精製ポリペプチド、(b) 配列番号3および配列番号3のアミノ酸配列に対して60%相同性を有するポリペプチドからなる群より選択されたアミノ酸配列から本質的になる単離または精製ポリペプチド、(c) 配列番号3のアミノ酸配列からなる単離または精製ポリペプチド、(d) 配列番号4および配列番号4のアミノ酸配列に対して60%相同性を有するポリペプチドからなる群より選択されるアミノ酸配列を有する単離または精製ポリペプチド、(e) 配列番号4および配列番号4のアミノ酸配列に対して60%相同性を有するポリペプチドからなる群より選択されたアミノ酸配列から本質的になる単離または精製ポリペプチド、(f) 配列番号4のアミノ酸配列からなる単離または精製ポリペプチド、(g) 配列番号5および配列番号5のアミノ酸配列に対して60%相同性を有するポリペプチドからなる群より選択されるアミノ酸配列を有する単離または精製ポリペプチド、(h) 配列番号5および配列番号5のアミノ酸配列に対して60%相同性を有するポリペプチドからなる群より選択されたアミノ酸配列から本質的になる単離または精製ポリペプチド、(i) 配列番号5のアミノ酸配列からなる単離または精製ポリペプチド、である。

#### 【0051】

本発明はまた、固有の分割加重スコア法 (Split and Weighted Scoring Method) に関する。本方法は、被験者から得た1つ以上のマーカーをスコア化するために使用可能である。本方法は、

a. 被験者から少なくとも1つのマーカーを得る段階、

b. 前記被験者からのマーカーの量を定量する段階、

c. 定量した各マーカーの量を、前記マーカーに対して所定のカットオフの数に対して比較する段階、および前記比較に基づいて、各マーカーに対してスコアを割り当てる段階、および、

d. 段階cにて定量した各マーカーに対する割り当てられたスコアを統合して、前記被験者に対する総スコアを見つけ出す段階、

を含み得る。

#### 【0052】

上記方法において、所定のカットオフは、ROC曲線に基づき、各マーカーに対するスコアは、マーカーの特異性に基づいて計算する。さらに、上記方法におけるマーカーは、バイオマーカー、バイオメトリックパラメーター、またはバイオマーカーとバイオメトリックパラメーターの組合せでありうる。

#### 【0053】

さらに、本発明は、分割加重スコア法 (Split and Weighted Scoring Method) を用いる、病状を発現させる患者の危険性を決定するための方法を提供する。本方法には、

a. 被験者から少なくとも1つのマーカーを得る段階、

b. 前記被験者からのマーカーの量を定量する段階、

c. 定量した各マーカーの量を、前記マーカーに対して所定のカットオフの数に対して比較する段階、および前記比較に基づいて、各マーカーに対してスコアを割り当てる段階、

d. 段階cにて定量した各マーカーに対する割り当てられたスコアを統合して、前記被験者に対する総スコアを見つけ出す段階、

e. 段階dにて決定された総スコアを、所定の総スコアと比較する段階、および、

f. 段階eにて決定した総スコアに基づいて、病状を発現させる危険性を前記被験者が有するかどうかを決定する段階、

が含まれうる。

#### 【0054】

上記方法において、所定のカットオフは、ROC曲線に基づき、各マーカーに対するスコアは、マーカーの特異性に基づいて計算する。さらに、上記方法におけるマーカーは、バイオマーカー、バイオメトリックパラメーター、またはバイオマーカーとバイオメトリックパラメーターの組合せでありうる。

10

20

30

40

50

## 【0055】

(発明の詳細な記述)

## 定義

本明細書で使用するように、以下の語句は以下の意味を有する。すべての他の技術的および科学的語句が、当業者に一般的に理解される意味を有する。

## 【0056】

語句「吸着剤 (adsorbent)」は、生体分子を集積 (結合) 可能な任意の物質を意味する。吸着剤は典型的に、生物学的に活性な表面を被覆し、それらの物理的特性に基づいて、生体分子または種々の生体分子に結合可能である、単一物質または多数の異なる物質からなる。そのような物質には、限定はしないが、アニオン交換物質、カチオン交換物質、金属キレーター、ポリヌクレオチド、オリゴヌクレオチド、ペプチド、抗体、ポリマー (合成または天然)、紙などが含まれる。

10

## 【0057】

本明細書で使用するところの語句「抗体 (antibody)」は、免疫グロブリン分子またはその免疫学的に活性な部分、すなわち抗原結合部分を意味する。免疫グロブリン分子の免疫学的に活性な部分の例には、ペプシンのような酵素で抗体を処理することによって産出可能な、 $F(ab)$  および  $F(ab')_2$  断片が含まれる。抗体の例には、限定はしないが、ポリクローナル抗体、モノクローナル抗体、キメラ抗体、ヒト抗体、ヒト化抗体、組換え抗体、一本鎖  $Fvs$  (「scFv」)、アフィニティー飽和抗体、一本鎖抗体、單一ドメイン抗体、 $F(ab)$  断片、 $F(ab')_2$  断片、ジスルフィド - 結合  $Fvs$  (「sdFv」) および抗イデオタイプ (「抗 - Id」) 抗体および以上の任意の機能的に活性なエピトープ - 結合断片が含まれる。本明細書で使用するところの語句「抗体」にはまた、自己抗体 (自己抗体は、被験者または患者が合成する、たとえば、限定はしないが、p53、カルレチクリン、- エノラーゼおよびHOXB7のような、正常の自己タンパク質 (または自己抗原) に対して指向する抗体である) が含まれる。広範囲の自己抗原に対する自己抗体が当業者に周知であり、特に肺がん、乳がん、前立腺がん、および膵臓がんなどを含む、多くの悪性疾患中で記述されてきている。抗体はバイオマーカーの1つの型である。

20

## 【0058】

本明細書で使用するところの語句「抗原 (antigen)」は、抗体によって結合されるうる分子を意味し、さらに動物が、その抗原の少なくとも1つのエピトープに結合可能な抗体を産出するように誘導可能である。さらに、関心領域が、抗原であり得る (言い換えれば、最終的に抗原であると決定されうる)。抗原はバイオマーカーの1つの型である。

30

## 【0059】

語句「AUC」は、ROC曲線の、曲線下面積 (Area Under the Curve) を意味する。これは、所与の試料集団における試験のためのメリットの図として使用され、完全な試験に対して1から、試験が試験被験者を分類することにおける完全なランダム応答を与える0.5までの範囲の値を与える。AUCの範囲が、0.5から1.0だけであるため、AUC中の小さな変化が、0から1または0から100%の範囲である計量中の同様の変化よりもより有意である。AUCにて%変化が与えられる場合、計量の全範囲が、0.5から1.0である事実に基づいて計算される。JMP (商標) 統計パッケージが、産出された各ROC曲線に対するAUCを報告する。AUC測定は、完全データ範囲にわたる分類アルゴリズムの精度を比較するための価値のある手段である。より大きなAUCを有する分類アルゴリズムは定義によって、2つの関心の群 (疾患および非疾患) 間で正しく未知を分類するより大きな能力を有する。分類アルゴリズムは、単一分子の測定と同様に単純であってよく、または多数の分子の測定および統合と同様に複雑であってよい。

40

## 【0060】

語句「良性肺疾患 (benign lung disease)」または「良性 (be

50

n i g n )」は、任意の所与の被験者の肺系と関連した疾患状態を意味する。本発明の文脈において、良性肺疾患には、限定はしないが、慢性閉塞性肺疾患（COPD）、急性または慢性炎症、良性小結節、良性新生組織形成、異形成症、過形成、異型性、気管支拡張症、ヒストラスマ症、サルコイドーシス、線維症、肉芽腫、血腫、肺気腫、肺拡張不全、組織球増殖症および他の非がん性疾患が含まれる。

【0061】

語句「生物学的に活性な表面（biologicalidcally active surface）」は、物質および生体分子の特異的な生化学的特性のために、生体分子が結合、または相互作用可能な物質の任意に二次元または三次元伸展を意味する。そのような生化学的特性には、限定はしないが、イオン性特性（荷電）、疎水性または親水性が含まれる。  
10

【0062】

語句「生物学的試料（biological sample）」および「試験試料（test sample）」は、任意の所与の被験者から単離されたすべての生物学的液体および排せつ物を意味する。本発明の文脈において、そのような試料には、限定はしないが、血液、血清、血漿、乳頭吸引液、尿、精液（semen）、精液（seminal fluid）、精漿、前立腺液、排泄物、涙、唾液、汗、生検、腹水、脳脊髄液、乳、リンパ液、気管支および他の洗浄液試料、または組織抽出試料が含まれる。典型的に、血液、血清、血漿および気管支洗浄液が、本発明の文脈における使用のために好ましい試験試料である。  
20

【0063】

語句「バイオマーカー（biomarker）」は、生理的状態と相関する生体分子（または生体分子の断片）を意味する。たとえば、本発明のバイオマーカーはがん、好ましくは肺がんに相関し、肺がんの存在または不在の検出における補助として使用可能である。そのようなバイオマーカーには、限定はしないが、ヌクレオチド、アミノ酸、糖、脂肪酸、ステロイド、代謝物、ポリペプチド、（限定はしないが、抗原および抗体のような）タンパク質、糖質、脂質、ホルモン、抗体、生体分子に対する代用として役目を果たす関心領域、これらの組合せ（たとえば糖タンパク質、リボヌクレオタンパク質、リポタンパク質）、および限定はしないが、抗原と、前記抗原上の利用可能なエピトープに結合する自己抗体間で形成される複合体のような、任意のそのような生体分子を含む任意の複合体が含まれる。語句「バイオマーカー」はまた、少なくとも5個の保存アミノ酸残基、好ましくは少なくとも10個の保存アミノ酸残基、より好ましくは15個の保存アミノ酸残基を含み、親ポリペプチドの生物学的活性および/またはいくつかの機能特性、たとえば抗原性または構造ドメイン特性を維持する、ポリペプチド（親）配列の一部分も意味する。  
30

【0064】

語句「バイオメトリックパラメーター（biometric parameter）」は、よく定義された群または集団に属しているような患者を一意的に同定するために使用する、1つ以上の固有の生理学的または行動特性を意味する。本発明の文脈において、「バイオメトリックパラメーター」には、限定はしないが、患者の理学的記述子が含まれる。バイオメトリックパラメーターの例には、限定はしないが、患者の身長、患者の体重、患者の性別、喫煙歴、職業歴、発がん性物質に対する暴露、受動喫煙への暴露、肺がんの家族歴などが含まれる。喫煙歴は通常、パックイヤー（pack years）において定量される。本明細書で使用するところの語句「パックイヤー（pack Years）」は、一日あたりで喫煙したパックの平均数でわった、個人の喫煙した年数を意味する。平均で、35年間、一日あたり1パックのたばこを吸った個人は、35パックイヤーの喫煙歴を有するとされる。バイオメトリックパラメーター情報は、患者の日常質問書または健康歴質問書などの利用によって、被験者自身からのような、当技術分野で公知の日常の技術を用いて、被験者より得られる。あるいは、バイオメトリックパラメーターは、看護士、ナース・プラクティショナー、医師のアシスタントまたは医師より、被験者から得ることができる。  
40  
50

## 【0065】

「保存アミノ酸置換 (conservative amino acid substitution)」は、アミノ酸残基が、類似側鎖を有するアミノ酸残基で置換されるものである。類似の側鎖を有するアミノ酸残基のファミリーが、当技術分野で定義されている。これらのファミリーには、塩基性側鎖（たとえばリシン、アルギニン、ヒスチジン）、酸性側鎖（たとえばアスパラギン酸、グルタミン酸）、未荷電極性側鎖（たとえばグリシン、アスパラギン、グルタミン、セリン、スレオニン、チロシン、システィン）、非極性側鎖（たとえばアラニン、バリン、ロイシン、イソロイシン、プロリン、フェニルアラニン、メチオニン、トリプトファン）、- 分岐側鎖（たとえばスレオニン、バリン、イソロイシン）および芳香族側鎖（たとえばチロシン、フェニルアラニン、トリプトファン、ヒスチジン）が含まれる。したがって、タンパク質内の予想非必須アミノ酸残基が好ましくは、同一の側鎖ファミリーからの他のアミノ酸残基と置換される。

## 【0066】

語句「決定木解析 (Decision Tree Analysis)」は、連続した單一二股ルール（または症状）が、決定木を介して、最終分類出力または木の末端節へのガイドを提供する、古典的アプローチを意味する。その本質的に単純で、直感的な性質が、帰納的な分割を診断工程に非常に従順にする。

## 【0067】

この方法は、2つの型の変数、因子変数 (X's) および応答変数 (Y's) を必要とする。実行されるように、X変数が連続的であり、Y変数が絶対的である（名目上）。そのような場合、JMP統計パッケージは、多数の節を最大化する、カットオフ値を産出するアルゴリズムを利用する。試料を、このカットオフ値より上および下である値に基づいて、枝または節に分割する。

## 【0068】

絶対的応答変数に関して、この場合でのように、フィットした値が、各応答レベルに対する推定可能性となる。この場合、スプリットを、もっとも大きな尤度比カイ二乗統計 ( $G^2$ ) によって決定される。これは、2つのスプリットの枝間の応答における差を最大化する効果を有する。本方法のより詳細な議論およびその実施が、JMP統計およびグラフィックスガイドで見ることができる。

## 【0069】

しかしながら、木を構築することは、それに関連したその固有の懸案事項を有する。共通の懸案事項は、データを過剰フィッティングすることなしに、もっともよい予測モデルを提供する木の最適な大きさを決定することである。このことを考慮に入れて、ROC曲線を構築するために、種々の木の節にて抽出可能な情報の利用を作成するよう、方法が開発された。実施されたように、本方法は、モデルされているデータ組を確実に過剰フィットされる十分な節で、参照木を構築することを含む。結果として、木が切り戻され、引き続いて、節の最小数が到達するまで（2末端節）、各段階でもっとも悪い節を除去する。これによって、複雑性が減少した（節が少ない）一連の木、または木のファミリーが作製される。

## 【0070】

帰納的分割プログラムは、純粋な末端節を作製することを試み、すなわち1つの分類型の標本のみが含まれる。しかしながら、常に可能ではない。しばしば、末端節は、混合集団を有する。したがって、各末端節は、がんに対する異なる可能性を持ちうる。がんに対する純粋な末端節において、がん標本である可能性は100%ありえ、反対に、非がんに対する純粋な末端節に対して、がん標本である可能性は、0%でありうる。各末端節でのがんの可能性は、各節にて、（1 - 非がんの可能性）に対してプロットされる。

## 【0071】

これらの値を、特定の木の代表である、ROC曲線を発生するためにプロットする。各木に対して計算したAUCは、木またはモデルの「有益な部分 (goodness)」を表している。任意の診断適用と同様に、AUCが高ければ高いほど、アッセイが、または

10

20

30

40

50

この場合モデルがよりよい。木の大きさ（節の数）に対する A U C のプロットが、その最大として、トレーニング組に対してもっともよいモデルを持ちうる。同様の手順を第二の、しかしそれより小さなデータのサブセットで実施して、結果を認証する。トレーニングおよび認証組両方での同様の実施のモデルが、さらなる解析および／または認証のために最適であり、したがって選択されると考えられる。

【 0 0 7 2 】

語句「開発データ組（d e v e l o p m e n t a l d a t a s e t ）」または「データ組（d a t a s e t ）」は、生体試料の組に対して回収した、完全バイオマーカーまたはバイオマーカーおよびバイオメトリックパラメーターを含む特徴を意味する。これらの試料自体が、公知の診断結果で患者から引き出される。特徴または特徴の組を統計解析にかけ、公知の患者予後に相関する、2つ以上の異なる試料群（たとえばがんおよび非がん）への試料の分類を意図する。質量スペクトルを使用する場合、組内の質量スペクトルが、それらの強度において異なりえるが、器具類の精度内の、それらの見かけ分子質量内ではない。

10

【 0 0 7 3 】

語句「分類子（c l a s s i f i e r ）」は、試料に関連した疾患を決定するために、試料の組に対して由来した特徴を用いる任意のアルゴリズムを意味する。分類子の1つの型が、トレーニング組からのデータで、アルゴリズムを「トレーニングすること（t r a i n i n g ）」によって作製し、その性能は、試験組データで評価する。本発明に関して使用する分類子の例は、判別分析、決定木解析、受信者動作曲線または分割スコア解析である。

20

【 0 0 7 4 】

語句「決定木（d e c i s i o n t r e e ）」は、分類のために使用される、フローチャート様木構造を有する分類子を意味する。決定木は、サブセット内へのデータ組の繰り返しスプリットからなる。各スプリットは、1つの変数に適用した単一のルールからなり、たとえば「変数1」の値が「閾値1」よりも大きい場合、「左にいき、それ以外の場合右に行く」である。したがって、所与の特徴スペースを、1つのクラスに割り当てた各長方形を有する、長方形の組に分割する。

【 0 0 7 5 】

語句「診断アッセイ（d i a g n o s t i c a s s a y ）」および「診断法（d i a g n o s t i c m e t h o d ）」は、病的状態の存在または性質の決定を意味する。診断アッセイは、これらの感度および特異性で異なる。肺がんに対して試験陽性であり、実際に疾患である患者が、「真の陽性（t r u e p o s i t i v e s ）」と考えられる。本発明の文脈内で、診断アッセイの感度は、疾患集団内の真の陽性の割合として定義される。肺がんを患っているが、診断アッセイによって検出されていない被験者が、「偽陰性（f a l s e n e g a t i v e s ）」と考えられる。疾患をもっておらず、診断アッセイで試験陰性である被験者は、「真の陰性（t r u e n e g a t i v e s ）」と考えられる。本明細書で使用するところの語句診断アッセイの特異性は、非疾患集団における真の陰性の割合として定義される。

30

【 0 0 7 6 】

語句「判別分析（d i s c r i m i n a t o r a n a l y s i s ）」は、2つ以上の天然に存在する群間を最適に判別する特徴を選別するために使用する統計法の組を意味する。判別分析のデータ組への適用によって、利用者は、さらなる解析のために、もっとも判別する特徴に焦点をあてることができる。

40

【 0 0 7 7 】

語句「理想からの距離（D i s t a n c e F r o m I d e a l ）」または「D F I」は、感度および特異性両方を組み込み、 $[ ( 1 - \text{感度} )^2 + ( 1 - \text{特異性} )^2 ]^{1/2}$ として定義される、理想からの距離である、R O C 曲線からとるパラメーターを意味する。D F I は、100% 感度および100% 特異性の性能のアッセイに対して0であり、0% 感度および0% 特異性を有するアッセイに対して1.414まで増加する。その決定に

50

対して完全データ範囲を使用する A U C とは異なって、 D F I が、 R O C 曲線上の特定の点における試験の性能を測定する。より低い D F I 値を有する試験は、より高い D F I 値を有するものより良好に機能する。 D F I は、米国特許明細書番号第 2 0 0 6 / 0 2 1 1 0 1 9 A 1 号にて詳細に議論されている。

## 【 0 0 7 8 】

語句「アンサンブル ( e n s e m b l e ) 」、「木アンサンブル ( t r e e e n s e m b l e ) 」または「アンサンブル分類子 ( e n s e m b l e c l a s s i f i e r ) 」が、相互互換的に使用可能であり、多くの単一の要素分類子からなる分類子を意味し、たとえば、決定木のアンサンブルは、決定木からなる分類子である。アンサンブル分類子の結果は、たとえば、すべての構成要素分類子を等しく重みを加える過半数投票によって、その構成要素分類子の結果のすべてを統合することによって得られる。過半数投票は、構成要素分類子を、これらが発生される頻度によって天然に重みを加える場合に、とりわけ道理にかなっている。 10

## 【 0 0 7 9 】

語句「エプトープ ( e p i t o p e ) 」は、また抗体によって認識されうる抗体によって結合されうる抗原の部分を引用することを意味する。エプトープ決定因子は通常、アミノ酸または糖側鎖のような分子の化学的に活性な表面グルーピングからなり、特定の三次元構造特性ならびに特定の荷電特性を有する。

## 【 0 0 8 0 】

語句「特徴 ( f e a t u r e ) 」および「変数 ( v a r i a b l e ) 」は、相互互換的に使用してよく、試料の特性の測定の値を意味する。これらの測定は、試料の物理的、化学的または生物学的特徴から由来してよい。測定の例には、限定はしないが、質量スペクトルピーク、質量スペクトルシグナル、 R O I の強度の関数が含まれる。 20

## 【 0 0 8 1 】

配列間の相同性または配列同一性の計算（語句は本明細書で相互互換的に使用される）は以下のように実施する。

## 【 0 0 8 2 】

2 つのアミノ酸配列の、または 2 つの核酸配列のパーセント同一性を決定するために、配列を、最適な比較目的のために並べる（たとえば、ギャップを、最適なアライメントのために、第一および第二アミノ酸または核酸配列の 1 つまたは両方に導入可能であり、非相同配列を比較の目的のために無視可能である）。好ましい実施形態において、比較目的のために並べた参照配列の長さは、並べる参照アミノ酸残基配列の長さの、少なくとも 3 0 %、好ましくは少なくとも 4 0 %、より好ましくは少なくとも 5 0 %、より好ましくは少なくとも 6 0 %、およびより好ましくは少なくとも 7 0 %、8 0 %、9 0 %、9 5 %、9 9 % または 1 0 0 % である。ついで、相当するアミノ酸位置またはヌクレオチド位置における、アミノ酸残基またはヌクレオチドを比較する。第一配列中の位置が、第二配列中の相当する位置と同一のアミノ酸残基またはヌクレオチドによって占有されている場合、分子がその位置で同一である（本明細書で使用するように、アミノ酸または核酸「同一性 ( i d e n t i t y ) 」は、アミノ酸または核酸「相同性 ( h o m o l o g y ) 」と等価である）。2 つの配列間のパーセント同一性は、2 つのアライメントの最適な配列のために導入される必要のある、ギャップの数、および各ギャップの長さを考慮にいれて、配列によって共有される同一の位置の数の関数である。 30

## 【 0 0 8 3 】

配列の比較および 2 つの配列間のパーセント同一性の決定は、数学的アルゴリズムを用いて達成可能である。好ましい実施形態において、2 つのアミノ酸配列間のパーセント同一性は、 B l o s s u m 6 2 マトリックスまたは P A M 2 5 0 マトリックスいずれか、および 1 6、1 4、1 2、1 0、8、6、または 4 のギャップウェイトおよび 1、2、3、4、5、または 6 の長さウェイトを用いて、 G C G ソフトウェアパッケージ中の G A P プログラム内に組み込まれた、 N e e d l e m a n a n d W u n s c h ( J . M o l . B i o l . 4 8 : 4 4 4 - 4 5 3 ( 1 9 7 0 ) ) アルゴリズムを用いて決定する。また 40

他の好ましい実施形態において、2つのヌクレオチド配列間のパーセント同一性は、N W S g a p d n a . C M Pマトリックスおよび40、50、60、70、または80のギャップウェイトと1、2、3、4、5、または6の長さウェイトを用いて、G C Gソフトウェアパッケージ中のG A Pプログラムを用いて決定する。パラメーターの特に好ましい組（および分子が本発明の配列同一性または相同性限界内であるかどうかを決定するために何のパラメーターを適用すべきかについて、実施者が不確かである場合に使用すべきもの）は、12のギャップオーブンペナルティー、4のギャップエクステンドペナルティー、および5のフレームシフトギャップペナルティーでの、B l o s s u m 6 2スコアリングマトリックスを用いることである。

【0084】

10

2つのアミノ酸またはヌクレオチド配列間のパーセント同一性は、P A M 1 2 0 ウェイト残余表、12のギャップ長さペナルティーおよび4のギャップペナルティーを用いて、A L I G Nプログラム（バージョン2.0）内に組み込まれた、E. Meyers and W. Miller (C A B I O S , 4 : 1 1 - 1 7 ( 1 9 8 9 ) )のアルゴリズムを用いて決定可能である。

【0085】

本明細書で記述された核酸およびタンパク質配列を、たとえば他のファミリーメンバーまたは関連配列を同定するために、公的なデータベースに対する検索を実施するために、「クエリー配列（q u e r y s e q u e n c e ）」として使用可能である。そのような検索は、A l t s c h u l , e t a l . , J . M o l . B i o l . 2 1 5 : 4 0 3 - 1 0 ( 1 9 9 0 ) のN B L A S T およびX B L A S T プログラム（バージョン2.0）を用いて実施可能である。B L A S T タンパク質検索は、本発明の免疫反応性サイクリンE 2 に相同性のアミノ酸配列を得るために、X B L A S T プログラム、スコア = 5 0 、ワード長 = 3 で実施可能である。比較の目的のためのギャップアライメントを得るために、G a p p e d B L A S T を、A l t s c h u l e t a l . , N u c l e i c A c i d s R e s . 2 5 ( 1 7 ) : 3 3 8 9 - 3 4 0 2 ( 1 9 9 7 ) にて記述されたように利用可能である。B L A S T およびG a p p e d B L A S T プログラムを使用する場合、それぞれのプログラム（X B L A S T およびN B L A S T ）のデフォルトパラメーターを利用可能である。

20

【0086】

30

本明細書で使用するところの語句「免疫反応性サイクリンE 2 ( i m m u n o r e a c t i v e サイクリンE 2 ）」は、（1）配列番号1、配列番号3、配列番号4、または配列番号5の任意のアミノ酸を有するポリペプチド、（2）配列番号1、配列番号3、配列番号4、または配列番号5の任意の組合せ、（3）配列番号1に対して、少なくとも60%、好ましくは少なくとも70%、より好ましくは少なくとも75、80、81、82、83、84、85、86、87、88、89、90、91、92、93、94、95、96、97、98、99%の相同性を有するアミノ酸配列を有するポリペプチド、配列番号3に対して、少なくとも60%、好ましくは少なくとも70%、より好ましくは少なくとも75、80、81、82、83、84、85、86、87、88、89、90、91、92、93、94、95、96、97、98、99%の相同性を有するアミノ酸配列を有するポリペプチド、配列番号4に対して、少なくとも60%、好ましくは少なくとも70%、より好ましくは少なくとも75、80、81、82、83、84、85、86、87、88、89、90、91、92、93、94、95、96、97、98、99%の相同性を有するアミノ酸配列を有するポリペプチド、配列番号5に対して、少なくとも60%、好ましくは少なくとも70%、より好ましくは少なくとも75、80、81、82、83、84、85、86、87、88、89、90、91、92、93、94、95、96、97、98、99%の相同性を有するアミノ酸配列を有するポリペプチド、およびこれらの任意の組合せ、（4）配列番号1、配列番号3、配列番号4または配列番号5と同様の免疫反応性を示すサイクリンE 2 ポリペプチド、および（5）配列番号1、配列番号3、配列番号4、または配列番号5と同様の免疫反応性を示すポリペプチド、を意味する

40

50

。

## 【0087】

「単離 (isolated)」または「精製 (purified)」ポリペプチドまたはタンパク質は、このタンパク質が由来する細胞または細胞起源からの細胞物質または他の汚染細胞を本質的に含まず、または化学的に合成された場合、化学的前駆体または他の化学物質を本質的に含まない。タンパク質またはその生物学的活性な部分を組換え的に産出する場合、培養培地を本質的に含まないことも好ましく、すなわち、培養培地が、タンパク質調製物の容量の、約20%未満、より好ましくは、約10%未満、もっとも好ましくは約5%未満であることを表す。

## 【0088】

10

本明細書で使用するところの語句「直線区別解析 (Linear Discriminate Analysis)」は、試料を正しくカテゴリー化する際にもっともよいそれらの変数または特徴を同定するためのツールを提供し、たとえばJMP (商標) 統計パッケージによって実施可能である、解析の1つの型を意味する。ソフトウェアの段階的特徴を用いて、変数を、すべての試料を正しく分類するまで、モデルに加えてよい。一般的に、この様式で選別された変数の組は、データ組中の変数の本来の数よりも、本質的に小さい。特徴の数のこの減少が、任意の次の解析、たとえば、決定木、人工神経ネットワークなどを用いる、より一般的な分類エンジンの開発を単純化する。

## 【0089】

20

語句「肺がん (lung cancer)」は、任意の所与の被験者の肺系に関連したがん状態を意味する。本発明の文脈において、肺がんには、限定はしないが、腺がん、扁平上皮がん、扁平上皮細胞がん、大細胞がん、小細胞がん、非小細胞がんおよび細気管支肺胞上皮がんが含まれる。本発明の文脈内で、肺がんは、異なるステージであってよく、グレードの程度は様々であってよい。肺がんのステージ、またはそのグレード程度を決定するための方法が、当業者に周知である。

## 【0090】

30

語句「質量分析法 (mass spectrometry)」は、表面上、試料からの気体相イオンを産出させるためのイオン源の利用、および質量分析器で気体相イオンを検出すること、を意味する。語句「レーザー脱着質量分析法 (laser desorption mass spectrometry)」は、表面上、試料からの気体相イオンを産出させるために、イオン源としてレーザーを利用し、および質量分析器で気体相イオンを検出することを意味する。生体分子のための質量分析法の好ましい方法は、マトリックス補助レーザー脱着 / イオン化質量分析法、またはMALDIである。MALDIにおいて、解析物を典型的に、乾燥に際して、解析物と共に結晶化する、マトリックス物質と混合する。マトリックス物質は、さもなければ不安定な生体分子または解析物を断片化する、エネルギー供給源からエネルギーを吸収する。他の好ましい方法は、表面増強レーザー脱着 / イオン化質量分析法またはSELDIである。SELDIにおいて、解析物を適用する表面が、解析物捕獲および / または脱着において活性な役割を果たす。本発明の文脈において、試料には、クロマトグラフィーまたは他の化学的処理を受けうる生物学的試料と、適切なマトリックス基質が含まれる。

40

## 【0091】

質量分析法において、「見かけの分子量 (apparent molecular mass)」は、検出されたイオンの、分子量 (ダルトン) 対電荷値、m/z を意味する。どのように見かけの分子量が由来するかは、使用する質量分析法の型に依存する。飛行時間質量分析法にて、見かけの分子量は、イオン化から検出までの時間の関数である。

## 【0092】

50

語句「マトリックス (matrix)」は、質量分析器中、適切な光源、たとえばUV / VIS またはIR レーザーからの光子としてエネルギーを吸着し、それによって、表面からの生体分子の脱着を可能にする分子を意味する。 - シアノ桂皮酸、シナピン酸およびジヒドロ安息香酸を含む桂皮酸誘導体が、生体分子のレーザー脱着において、エネル

ギー吸着分子としてよく使用される。エネルギー吸着分子は、参照によって本明細書に組み込まれている、米国特許第5,719,060号にて記述されている。

【0093】

語句「正規化 (normalization)」およびその誘導語は、質量スペクトルに関連して使用する場合、スペクトルの全体的な強度における、おもに機器パラメーターが原因の、差を除去または最小化するために、質量スペクトルの組に適用する、数学的方法を意味する。

【0094】

語句「関心領域 (region of interest)」または「ROI」は、質量スペクトルのサブセットの統計学的適合を意味する。ROIは、連続したシグナルの固定化された最小長を有する。連続シグナルは、どのようにROIが選ばれるかに依存して、固定最大長のギャップを含んでよい。関心領域は、バイオマーカーに関し、バイオマーカーに対する代用として働きうる。関心領域は、後で、タンパク質、ポリペプチド、抗原、抗体、脂質、ホルモン、糖質などであると決定されてよい。

【0095】

語句「受信者動作特性曲線 (Receiver Operator Characteristic Curve)」または「ROC曲線 (ROC curve)」は、そのもとも単純な適用において、2つの集団（たとえば、症例（すなわち、肺がんを患っている被験者）および対照（すなわち、肺がんに対して正常であるか、または良性である被験者））間を区別することにおける、特定の特徴（たとえば、バイオマーカーまたはバイオメトリックパラメーター）の性能のプロットを意味する。全集団（すなわち症例および対照）に渡る特徴データを、単一の特徴の値に基づいて、昇順で仕分ける。ついで、その特徴に対する各値に対して、各データに関して真の陽性および偽陽性率を計算する。真の陽性率は、検討中の特徴に対する値以上の症例の数を計数して、ついで、症例の総数によって割ることによって決定する。偽陽性率は、検討して、特徴に対する値以上の対照の数を計数して、対照の総数によってわることによって決定する。この定義が、対照と比較して、症例にて特徴が上昇するシナリオを記述している一方で、この定義はまた、対照と比較して、症例にて特徴が抑制されるシナリオも企図している。本シナリオにおいて、検討中の特徴の値以下の試料を計算しうる。

【0096】

ROC曲線は、単一特徴に対して、ならびに他の単一出力に対して産出可能であり、たとえば、2つ以上の特徴の組合せが、数学的に一緒に（加えられる、引かれる、かけられるなどのように）統合されて、単一の合計値が提供され、この単一合計値をROC曲線中でプロット可能である。さらに、それによって組合せが単一の出力値を誘導する、多数の特徴の任意の組合せが、ROC曲線中でプロット可能である。特徴のこれらの組合せが試験を含んでよい。ROC曲線は、試験の偽陽性率（1 - 特異性）に対する試験の真の陽性率（感度）のプロットである。ROC曲線下面積は、所与の試料集団に対する特徴に関するメリットの図であり得、試験が、試験被験者を分類することにおいて、完全に無作為な応答を与える完全な試験に対して1から0.5の範囲の値を与える。ROC曲線は、データ組を迅速に検査するための他の手段を提供する。診断的であると見られる特徴を、大きな特徴スペースの大きさを減少させるために好ましく使用可能である。

【0097】

語句「スクリーニング (screening)」は、肺がんへの患者の傾向に関する診断決定を意味する。患者は、陽性「スクリーニング試験」で高い肺がんの危険性であると決定される。結果として、患者は、所与の追加試験、たとえば撮像、唾液試験、肺機能試験、気管支鏡検査法および/または生検手順を実施して、および最終診断を下すことが可能である。

【0098】

語句「シグナル (signal)」は、調査している生体分子によって産出される任意の応答を意味する。たとえば、語句シグナルは、質量分析器の検出器にぶつかる生体分子

10

20

30

40

50

によって産出される応答を意味する。シグナル強度は、生体分子の量または濃度と比例する。シグナルは、2つの値、見かけ分子量値と、以上で記述したように産出された強度値によって定義される。質量値は、生体分子の要素特性であり、一方で強度値は、相当する見かけ分子量値を有する生体分子の特定の量または濃度に一致する。したがって「シグナル」はいつも、生体分子の特性を意味する。

#### 【0099】

語句「分割スコア法 (S p l i t a n d S c o r e M e t h o d )」は、M o r e t a l . , P N A S , 1 0 2 ( 2 1 ) : 7 6 7 7 - 7 6 8 2 ( 2 0 0 5 ) から適合した方法を意味する。本方法において、多数の測定をすべての試料で実施する。カットオフ値を各測定に対して決定する。このカットオフ値を、関心の群間（たとえば疾患および非疾患）の正確な分類の精度を最大化するために設定してよく、または1つの群の感度または特異性を最大化するために設定してよい。各測定のために、関心の群、たとえば疾患が、カットオフ値より上、またはカットオフ値より下であるかどうかを決定する。各測定に対して、その測定の値が、カットオフ値の疾患側において発見される時はいつでも、その試料にスコアを割り当てる。すべての測定を1つの試料で実施した後、スコアを合計して、測定のパネルに対する総スコアを産出する。10の測定のパネルが、最大スコア10をもち（各測定は1または0のいずれかのスコアを有する）、最小スコア0を有するように、すべての測定に等しく重みをかけることが一般的である。しかしながら、より有意な測定のために、より高い個々のスコアで測定に不均等に重みをかけることが価値がある。

10

20

#### 【0100】

総スコアを決定した後、測定のパネルに基づいて、非疾患試料から疾患を分類するために、もう一度カットオフを決定する。ここで再び、最大スコア10および最小スコア0の測定のパネルに対して、カットオフを、感度（カットオフとしてスコア0）を最大化するため、または特異性（カットオフとしてスコア10）を最大化するため、または分類の精度（カットオフとして0から10の間のスコア）を最大化するために選択してよい。

#### 【0101】

本明細書で使用するところの、語句「分割加重スコア法 (S p l i t a n d W e i g h t e d S c o r i n g M e t h o d )」は、試験試料中で同定および定量される1つのバイオマーカーまたはバイオメトリックパラメーター（本明細書にて、「マーカー（類）」（m a r k e r ( s ) ）と集合的によぶ）の測定を、多くの可能性あるスコアの1つに変換することを含む方法を意味する。スコアは、以下の等式を用いて得る。

$$\text{スコア} = A U C ^ * \text{因子} / (1 - \text{特異性})$$

式中「因子（f a c t o r ）」は（0、1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24、25などのような）整数であり、「特異性（s p e c i f i c i t y ）」は、1以下である値の選択値である。「因子」の程度は、限定はしないが、より高いA U C 値、相対的に小さな標準偏差、高い特異性または感度または低いD F I のような、改善された臨床性能を有するマーカーに対して増加する。その結果、1つのマーカーの測定を、所望の多くの、または少ないスコアに変換可能である。本方法は、関心の集団におけるマーカー/試験性能を反映する、受信者動作特性曲線に基づいている。R O C 曲線は、試験の真の陽性率（感度）の、試験の偽陽性率（1 - 特異性）に対するプロットである。曲線上の各点は、測定している特徴/試験（マーカー）の単一の値を表している。したがって、いくつかの値が、関心の集団（すなわち、肺がんを発現させる危険性である被験者）において、低い偽陽性率を持ち、一方、特徴の他の値が、その集団で高い偽陽性率を持ちうる。本方法は、関心の被験者の集団に対して、低い偽陽性率を有する（それによって高い特異性を有する）特徴値（すなわち、バイオマーカーまたはバイオメトリックパラメーター）に対して、より高いスコアを提供する。本方法は、それよりも低いと試験の結果がスコアの増加となる、偽陽性（1 - 特異性）の所望のレベルを選択することを含む。言い換えると、高い特異性のマーカーが、特異性の低いマーカーより大きなスコア、またはより大きな範

30

40

50

囲のスコアを与える。

【0102】

本明細書で使用するところの語句「被験者 (s u b j e c t )」は、ヒトまたは非ヒトを含む、動物、好ましくは哺乳動物を意味する。語句患者および被験者は、本明細書で相互互換的に使用されてよい。

【0103】

語句「DTモデルの10倍検証 (T e n - f o l d V a l i d a t i o n o f D T M o d e l s )」は、良好な解析実施が、モデルをそれらの予測値を査定するために、2つの集団に対して検証されることを必要とする。新規の集団のかわりに、データを独立したトレーニング組と検証組に分けることができる。10のランダム亜組を、検証組としての利用のために産出する。各検証組に対して、共通して試料を含まない、相当する独立トレーニング組が存在する。10のDTモデルを、上述のように、10トレーニング組から産出し、検証組と問い合わせを行う。

10

【0104】

語句「試験組 (t e s t s e t )」または「未知 (u n k n o w n )」または「検証組 (v a l i d a t i o n s e t )」は、トレーニング組中に含まれない登録からなる全利用可能データ組の亜組を意味する。試験データを、分類子性能を評価するために適用する。

【0105】

語句「トレーニング組 (t r a i n i n g s e t )」または「公知組 (k n o w n s e t )」または「参照組 (r e f e r e n c e s e t )」は、それぞれの全利用可能データ組の亜組を意味する。この亜組は典型的に、無作為に選択され、分類子構築のためにのみ使用される。

20

【0106】

語句「変換ロジスティック回帰 (T r a n s f o r m e d L o g i s t i c R e g r e s s i o n M o d e l )」は、J M P T M 統計パッケージに組み込まれてもいる、特徴の数を統合し、ROC曲線解析を可能にする方法を提供するモデルを意味する。このアプローチは、特徴の他の1つとの関係に関して、単純かモデルを仮定するので、減少した特徴の組にもっともよく適用される。陽性結果は、より洗練された分類方法が良好であるべきであることを示唆する。残念であるが、陰性結果は、他の方法に対して必ずしも失敗を暗示しない。

30

【0107】

サイクリンE2ポリペプチド

1つの実施形態において、本発明は、単離または精製免疫反応性サイクリンE2ポリペプチド、または本明細書で記述した方法にて使用可能である抗体を発生させる、または試験する（またはより一般的に、結合する）免疫原または抗原として使用可能な、それらの生物学的に活性な断片に関する。本発明の免疫反応性サイクリンE2ポリペプチドは、標準のタンパク質精製技術を用いて、細胞または組織起源より単離可能である。あるいは、単離または精製免疫反応性サイクリンE2ポリペプチおよびその生物学的に活性な断片は、組換えDNA技術によって産出可能であり、または化学的に合成可能である。本発明の単離または精製免疫反応性サイクリンE2ポリペプチドは、配列番号1、配列番号3、配列番号4および配列番号5にて示されたアミノ酸配列を有する。配列番号1は、ヒトサイクリンE2のcDNA発現形態のアミノ酸配列である (G e n b a n k 受入B C 0 0 7 0 1 5 . 1 )。配列番号3は、B C 0 0 7 0 1 5 . 1 のC - 末端 + 1アミノ酸（システイン）を含む38アミノ酸配列であり、本明細書ではまた「E 2 - 1」とも呼ばれる。配列番号4は、長さにして37アミノ酸であり、配列番号4がそのアミノ末端で、配列番号3の非常に初期のシステインを含まないことを除いて、配列番号3と同一である。配列番号5は、B C 0 0 7 0 1 5 . 1 のC - 末端を含む19アミノ酸配列であり、本明細書で「E 2 - 2」と呼ばれる。実施例にてより詳細に記述するが、免疫反応性配列番号1を、配列番号2の免疫反応性と比較した。配列番号2は、ヒトサイクリンE2 (G e n b a n k 受入

40

50

B C 0 2 0 7 2 9 . 1 ) の他の c D N A 発現形態である。配列番号 1 が、がん試料のいくつかのプールと強力な免疫反応性をしめすことがわかり、良性および正常（非がん）プールとは非常に低い反応性を示した。反対に、配列番号 2 は、がんまたは非がんプール試料とほとんど反応を示さなかった。配列番号 1 の免疫反応性は、配列番号 2 には存在しない、配列番号 1 の C - 末端に存在する、最初の 3 7 アミノ酸の結果であると決定された。両方とも配列番号 1 の C - 末端から由来する、配列番号 3 および 5 は、がんまたは非がんプール間で強力な免疫反応性を示すことがわかった。したがって、任意の配列番号 1 、配列番号 3 、配列番号 4 および配列番号 5 、または（配列番号 1 および配列番号 3 に対して産出された抗体、配列番号 1 および配列番号 4 に対して産出された抗体、配列番号 1 および配列番号 5 に対して産出された抗体、配列番号 1 、配列番号 3 および配列番号 4 に対して産出された抗体、配列番号 1 、配列番号 4 および配列番号 5 に対して産出された抗体、配列番号 1 、配列番号 3 、配列番号 4 および配列番号 5 に対して産出された抗体、配列番号 3 および配列番号 4 に対して産出された抗体、配列番号 3 、配列番号 4 および配列番号 5 に対して産出された抗体、配列番号 4 および配列番号 5 に対して産出された抗体のような）これらの配列の任意の組合せに 対して産出された抗体を、本明細書で記述した方法において利用可能である。たとえば、そのような抗体は、任意の配列番号 1 、配列番号 3 、配列番号 4 および配列番号 5 またはこれらの配列の任意の組合せに対して産出された被験者抗体であり得る。そのような抗体は、本明細書で記述された本発明の方法における利用のための、1 つ以上のキット内に含まれてよい。

10

20

#### 【 0 1 0 8 】

本発明はまた、1 つ以上の保存アミノ酸置換によって、本明細書で記述されたポリペプチド（すなわち配列番号 1 、配列番号 3 、配列番号 4 および配列番号 5 ）とは異なるポリペプチドも含む。さらに、本発明はまた、配列番号 1 、配列番号 3 、配列番号 4 および配列番号 5 のアミノ酸配列をもつポリペプチドと、少なくとも 6 0 % 、好ましくは少なくとも 7 0 % 、より好ましくは少なくとも 7 5 、 8 0 、 8 1 、 8 2 、 8 3 、 8 4 、 8 5 、 8 6 、 8 7 、 8 8 、 8 9 、 9 0 、 9 1 、 9 2 、 9 3 、 9 4 、 9 5 、 9 6 、 9 7 、 9 8 、 9 9 % 以上の総配列類似性（同一性）または相同性を有するポリペプチドも含む。

#### 【 0 1 0 9 】

肺がんの存在を検出することにおける、バイオマーカーおよびバイオメトリックパラメーターの利用

30

他の実施形態において、本発明は、正常被験者と、がんの被験者、または病状、好ましくはがん、より好ましくは肺がんを発現させる危険性にある被験者を区別することにおいて効果的に補助する方法に関する。正常被験者は、がんのような任意の病状で診断されないもの、より好ましくは肺がんと診断されないものと考えられる。

#### 【 0 1 1 0 】

本発明は、病状、好ましくはがん、より好ましくは肺がんの診断において補助するための、迅速で、感度がよく、利用が簡単な方法を有利に提供する。さらに、本発明は、病状を発現させる危険性にある個体を同定するため、病状の危険性にある被験者を検査するため、および病状と診断された、またはそのために治療されている患者をモニタするために利用可能である。本発明はまた、病状のために治療されている患者の治療の効果をモニタするために利用可能である。好ましくは、病状はがんであり、より好ましくは肺がんである。

40

#### 【 0 1 1 1 】

一般的に、本発明の方法には、被験者より試験試料を得ることが含まれる。典型的に、試験試料は、被験者より得、当業者に公知の標準の方法を用いて処理する。血液標本および血清またはそれらから由来する血漿に関して、試料は従来より、静脈穿刺によって肘正中静脈より得ることができ、またはより少量の容量が必要である場合は、手掌での採血によって得ることができる。両方の場合で、形成された要素および固まりを遠心によって除去する。尿または排泄物は、迅速に処理されるか、処理が即時にできない場合に保存剤で

50

安定化するという条件で、患者から直接回収できる。気管支洗浄液または胸膜液のようなより特定の試料は、気管支鏡検査法の間、または経皮またはオープン生検によって回収可能であり、微粒子物質が遠心によって除去されたならば、血清または血漿と同様に処理可能である。

【0112】

処理の後、被験者から得た試験試料を、肺がんの診断と相関可能である1つ以上のバイオマーカーの存在および量に関して問い合わせをおこなう。とくに、本出願者らは、1つ以上のバイオマーカー、または（少なくとも1つのバイオマーカー、少なくとも1つのバイオマーカーおよび少なくとも1つのバイオメトリックパラメーター、少なくとも2つのバイオマーカー、少なくとも2つのバイオマーカーおよび1つのバイオメトリックパラメーター、少なくとも1つのバイオマーカーおよび少なくとも2つのバイオメトリックパラメーター、少なくとも2つのバイオマーカーおよび少なくとも2つのバイオメトリックパラメーター、少なくとも3つのバイオマーカーなどのような）バイオマーカーとバイオメトリックパラメーターの組合せの検出および定量化が、患者において肺がんを診断することにおける補助として有用であることを発見した。本明細書で記述された方法にて同定および定量された1つ以上のバイオマーカーを、1つ以上のパネル内に含めることができる。パネルを含むバイオマーカーの数は重要ではなく、限定はしないが、1つのバイオマーカー、2つのバイオマーカー、3つのバイオマーカー、4つのバイオマーカー、5つのバイオマーカー、6つのバイオマーカー、7つのバイオマーカー、8つのバイオマーカー、9つのバイオマーカー、10つのバイオマーカー、11つのバイオマーカー、12つのバイオマーカー、13つのバイオマーカー、14つのバイオマーカー、15つのバイオマーカー、16つのバイオマーカー、17つのバイオマーカー、18つのバイオマーカー、19つのバイオマーカー、20つのバイオマーカー、などでありうる。

【0113】

以上で言及したように、試験試料を得た後、本発明の方法には、試験試料中の1つ以上のバイオマーカーの存在を同定し、ついで定量することが含まれる。肺がんの危険性にあると疑われる患者の診断において、有用であるか、または有用であると信じられている任意のバイオマーカーを、本明細書で記述している方法にて定量可能であり、1つ以上のパネル内に含めることができる。その結果、1つの態様において、パネルは、1つ以上のバイオマーカーを含んでよい。パネルに含めることができるバイオマーカーの例には、限定はしないが、抗-p53、抗-TMP21、抗-ニーマン-ピック C1-様タンパク質1、C末端ペプチド-ドメイン（抗-NPC1L1C-ドメイン）、抗-TMOD1、抗-CAMK1、抗-RGS1、抗-PAC SIN1、抗-RCV1、抗-MAPKAPK3、（配列番号1、配列番号3、配列番号4、配列番号5、またはそれらの任意の組合せに対する抗体のような）免疫反応性サイクリンE2に対する少なくとも1つの抗体、限定はしないが、がん胎児性抗原（CEA）、がん抗原125（CA125）、がん抗原15-3（CA15-3）、プロガストリン放出ペプチド（proGRP）、扁平上皮細胞抗原（SCC）、サイトケラチン8、サイトケラチン19ペプチドまたはタンパク質（本明細書で、単に「CK-19、CYFRA21-1、Cyfra」とも呼ばれる）およびサイトケラチン18ペプチドまたはタンパク質（CK18、TPS）、シアル酸残基が追加された、Lewis A血液集団であるがん抗原19-9（CA19-9）のような糖質抗原、血清アミロイドA、-1-抗-トリプシンおよびアポリポタンパク質CIII、および限定はしないが、Acn6399、Acn9459、Pub11597、Pub4789、TFA2759、TFA9133、Pub3743、Pub8606、Pub4487、Pub4861、Pub6798、Pub6453、Pub2951、Pub2433、Pub17338、TFA6453およびHIC3959のような関心領域のような抗原が含まれる。

【0114】

他の態様において、パネルは、少なくとも1つの抗体、少なくとも1つの抗原、少なくとも1つの関心領域、少なくとも1つの抗原および少なくとも1つの抗体、少なくとも1

10

20

30

40

50

つの抗原および少なくとも1つの関心領域、少なくとも1つの抗体および少なくとも1つの関心領域および少なくとも1つの抗原、少なくとも1つの抗体および少なくとも1つの関心領域を含んでよい。パネル中に含めることができる少なくとも1つの抗体の例には、限定はしないが、抗-p53、抗-TMP21、抗-NPC1L1C-ドメイン、抗-TMOD1、抗-CAMK1、抗-RGS1、抗-PACSI1、抗-RCV1、抗-MAPKAPK3、免疫反応性サイクリンE2に対する少なくとも1つの抗体が含まれる。パネル中に含めることができる少なくとも1つの抗原には、限定はしないが、サイトケラチン8、サイトケラチン19、サイトケラチン18、CEA、CA125、SCC、CA19-9、proGRP、血清アミロイドA、-1-抗-トリプシンおよびアポリポタンパク質CIIが含まれる。パネル内に含めることができる少なくとも1つの関心領域の例には、限定はしないが、Acn6399、Acn9459、Pub11597、Pub4789、TFA2759、TFA9133、Pub3743、Pub8606、Pub4487、Pub4861、Pub6798、Pub6453、Pub2951、Pub2433、Pub17338、TFA6453およびHIC3959が含まれる。さらに、特定の関心領域が、特定の他の関心領域と非常に相関し(これらの関心領域が、相互間で高い相関係数を有することを意味する)、したがって、本発明の文脈内で、相互に置換可能である。特に、これらの非常に相関した関心領域が、特定の相関しているファミリーまたは「群」へと組み立てられる。これらの「群」内に含まれる関心領域を、本発明の方法およびキット内で相互に置換可能である。これらの相関している関心領域のファミリーまたは「群」を以下に記述している。 10

## 【0115】

群A：関心領域：Pub3448およびPub3493。

## 【0116】

群B：関心領域：Pub4487およびPub4682。

## 【0117】

群C：関心領域：Pub8766、Pub8930、Pub9142、Pub9216、Pub9363、Pub9433、Pub9495、Pub9648およびPub9722。 20

## 【0118】

群D：関心領域：Pub5036、Pub5139、Pub5264、Pub5357、Pub5483、Pub5573、Pub5593、Pub5615、Pub6702、Pub6718、Pub10759、Pub11066、Pub12193、Pub13412、Acn10679およびAcn10877。 30

## 【0119】

群E：関心領域：Pub6391、Pub6533、Pub6587、Pub6798、Pub9317およびPub13571。

## 【0120】

群F：関心領域：Pub7218、Pub7255、Pub7317、Pub7413、Pub7499、Pub7711、Pub14430およびPub15599。 40

## 【0121】

群G：関心領域：Pub8496、Pub8546、Pub8606、Pub8662、Pub8734、Pub17121およびPub17338。

## 【0122】

群H：関心領域：Pub6249、Pub12501およびPub12717。

## 【0123】

群I：関心領域：Pub5662、Pub5777、Pub5898、Pub11597およびAcn11559。 50

## 【0124】

群J：関心領域：Pub7775、Pub7944、Pub7980、Pub8002およびPub15895。

## 【0125】

群K：関心領域：Pub17858、Pub18422、Pub18766およびPub18986。

## 【0126】

群L：関心領域：Pub3018、Pub3640、Pub3658、Pub3682、Pub3705、Pub3839、Hic2451、Hic2646、Hic3055、Tfa3016、Tfa3635およびTfa4321。

## 【0127】

群M：関心領域：Pub2331およびTfa2331。

## 【0128】

群N：関心領域：Pub4557およびPub4592。

## 【0129】

群O：関心領域：Acn4631、Acn5082、Acn5262、Acn5355、Acn5449およびAcn5455。

## 【0130】

群P：関心領域：Acn6399、Acn6592、Acn8871、Acn9080、Acn9371およびAcn9662。

## 【0131】

群Q：関心領域：Acn9459およびAcn9471。

## 【0132】

群R：関心領域：Hic2506、Hic2980、Hic3176およびTfa2984。

## 【0133】

群S：関心領域：Hic2728およびHic3276。

## 【0134】

群T：関心領域：Hic6381、Hic6387、Hic6450、Hic6649、Hic6816およびHic6823。

## 【0135】

群U：関心領域：Hic8791およびHic8897。

## 【0136】

群V：関心領域：Tfa6453およびTfa6652。

## 【0137】

群W：関心領域：Hic6005およびHic5376。

## 【0138】

群X：関心領域：Pub4713、Pub4750およびPub4861。

## 【0139】

本発明の方法にて使用可能な好ましいパネルには、限定はしないが、

1. 少なくとも1つのバイオマーカーを含むパネルで、前記バイオマーカーが、少なくとも1つの抗体と少なくとも1つの抗原である。本パネルはまたさらに、少なくとも1つの関心領域のようなさらなるバイオマーカーを含み得る。

## 【0140】

2. 少なくとも1つのバイオマーカーを含むパネルで、前記バイオマーカーが、免疫反応性サイクリンE2に対する少なくとも1つの抗体を含む。さらに、本パネルはまた場合によりさらに、試験試料中の、少なくとも1つの抗体、少なくとも1つの抗原、少なくとも1つの抗原および少なくとも1つの抗体、少なくとも1つの関心領域、少なくとも1つの抗原および少なくとも1つの関心領域および少なくとも1つの抗体および少なくとも1つの抗原、少なくとも1つの抗体および少なくとも1つの関心領域のような、さらなるバイオマーカーを含み得る。

## 【0141】

3. 少なくとも1つのバイオマーカーを含むパネルで、バイオマーカーが、サイトケラ

10

20

30

40

50

チソ 8、サイトケラチン 19、サイトケラチン 18、CEA、CA 125、SCC、proGRP、血清アミロイド A、-1-抗-トリプシンおよびアポリポタンパク質 CII からなる群より選択される。本パネルは場合によりさらに、試験試料中の、少なくとも 1 つの抗体、少なくとも 1 つの関心領域、少なくとも 1 つの関心領域と少なくとも 1 つの抗体のようなさらなるバイオマーカーを含み得る。

【0142】

4. 少なくとも 1 つのバイオマーカーを含むパネルで、バイオマーカーが、Acn 63 99、Acn 9459、Pub 11597、Pub 4789、TFA 2759、TFA 9133、Pub 3743、Pub 8606、Pub 4487、Pub 4861、Pub 6798、Pub 6453、Pub 2951、Pub 2433、Pub 17338、TFA 6453 および HIC 3959 からなる群より選択される、少なくとも 1 つの関心領域である。パネルはまた場合によりさらに、試験試料中の、少なくとも 1 つの抗原、少なくとも 1 つの抗体および少なくとも 1 つの抗原と少なくとも 1 つの抗体のような、さらなるバイオマーカーを含み得る。

【0143】

5. パネル中の少なくとも 1 つのバイオマーカーを含むパネルで、少なくとも 1 つのバイオマーカーが、サイトケラチン 8、サイトケラチン 19、サイトケラチン 18、CEA、CA 125、SCC、proGRP、血清アミロイド A、-1-抗-トリプシンおよびアポリポタンパク質 CII 、Acn 6399、Acn 9459、Pub 11597、Pub 4789、TFA 2759、TFA 9133、Pub 3743、Pub 8606、Pub 4487、Pub 4861、Pub 6798、Pub 6453、Pub 2951、Pub 2433、Pub 17338、TFA 6453 および HIC 3959 からなる群より選択される。パネルはまた場合によりさらに、少なくとも 1 つの抗体のようなさらなるバイオマーカーを含み得る。好ましいパネルは、サイトケラチン 19、CEA、ACN 9459、Pub 11597、Pub 4789 および TFA 2759；サイトケラチン 19、CEA、ACN 9459、Pub 11597、Pub 4789、TFA 2759 および TFA 9133；サイトケラチン 19、CA 19-9、CEA、CA 15-3、CA 125、SCC、サイトケラチン 18 および ProGRP；Pub 11597、Pub 3743、Pub 8606、Pub 4487、Pub 4861、Pub 6798、Tfa 6453 および Hic 3959；およびサイトケラチン 19、CEA、CA 125、SCC、サイトケラチン 18、ProGRP、ACN 9459、Pub 11597、Pub 4789、TFA 2759、TFA 9133 を含むパネルである。

【0144】

試験試料中の 1 つ以上のバイオマーカーの存在および量を、当業者に公知の日常の技術を用いて得ること、および定量が可能である。たとえば、試験試料中の抗原または抗体を定量するための方法が、当業者に周知である。たとえば、試験試料中の 1 つ以上の抗原または抗体の存在および定量化を、当技術分野で公知の 1 つ以上の免疫アッセイを用いて決定可能である。免疫アッセイには典型的に、(a) バイオマーカー(すなわち抗原または抗体)に特異的に結合する抗体(または抗原)を提供すること、(b) 試験試料を抗体または抗原に接触させること、および (c) 試験試料中の抗原に結合した抗体の複合体、または試験試料中の抗体に結合した抗原の複合体の存在を検出すること、が含まれる。

【0145】

抗原に特異的に結合する抗体を調製するために、精製抗原または精製抗原の核酸配列を使用可能である。抗原の核酸およびアミノ酸配列は、これらの抗原のさらなる特性化によって得ることができる。たとえば、抗原は、多数の酵素(たとえばトリプシン、V8 プロテアーゼなど)でマップしたペプチドでありうる。各抗原からの消化断片の分子量を、種々の酵素によって産出された消化断片の分子量と一致する配列に関して、SwissProt データベースのようなデータベースを検索するために使用可能である。本方法を用いて、他の抗原の核酸およびアミノ酸配列を、これらの抗原がデータベース中の公知のタンパク質であるかどうか同定可能である。

## 【0146】

あるいは、タンパク質を、タンパク質ラダーシークエンシングを用いて配列決定可能である。タンパク質ラダーは、たとえば分子を断片化し、断片を酵素消化にかけることによって、または断片の末端から単一アミノ酸を連続して除去する他の方法によって、産出可能である。タンパク質ラダーを調整する方法は、たとえば国際公開番号第WO93/24834号および米国特許第5,792,664号にて記述されている。ついでラダーを質量分析によって解析する。ラダー断片の質量における差違が、分子の末端から除去したアミノ酸を同定する。

## 【0147】

抗原がデータベース中で未知である場合、核酸およびアミノ酸配列は、抗原のアミノ酸配列の一部分の知識で決定可能である。たとえば、変性プローブを、抗原のN-末端アミノ酸配列に基づいて作製可能である。これらのプローブをついで、抗原が最初に検出された試料から作製されるゲノムまたはcDNAライブラリーを検査するために使用可能である。陽性クローニングを同定し、増幅し、それらの組換えDNA配列を、周知の技術を用いてサブクローニング可能である。たとえばCurrent Protocols for Molecular Biology (Ausubel et al., Green Publishing Assoc. and Wiley-Interscience 1989) およびMolecular Cloning: A Laboratory Manual, 2nd Ed. (Sambrook et al., Cold Spring Harbor Laboratory, NY 1989) を参照のこと。

10

20

## 【0148】

精製した抗原またはその核酸配列を用いて、抗原に特異的に結合した抗体を、当分野で公知の任意の適切な方法を用いて調製可能である(たとえば、Coligan, Current Protocols in Immunology (1991); Harlow & Lane, Antibodies: A Laboratory Manual (1988); Goding, Monoclonal Antibodies: Principles and Practice (2d ed. 1986) およびKohler & Milstein, Nature 256:495-497 (1975) を参照のこと)。そのような技術には、限定はしないが、ファージまたは同様のベクター中の組換え抗体のライブラリーからの抗体の選別による抗体の調製、ならびにウサギまたはマウスを免疫化することによるポリクローナルおよびモノクローナル抗体の調製が含まれる(たとえば、Huse et al., Science 246:1275-1281 (1989)、Ward et al., Nature 341:544-546 (1989) を参照のこと)。

30

## 【0149】

抗体が提供された後、抗原を、任意の多数のよく認識されている免疫学的結合アッセイを用いて検出および/または定量可能である(たとえば、米国特許番号第4,366,241号、第4,376,110号、第4,517,288号、および第4,837,168号を参照のこと)。本発明の使用可能なアッセイには、たとえば、「サンドイッチアッセイ」としても公知である酵素免疫吸着アッセイ(ELISA)、酵素免疫アッセイ(ELIA)、放射免疫アッセイ(RIA)、蛍光免疫アッセイ(FIA)、化学発光免疫アッセイ(CLIA)、カウンティング免疫アッセイ(CIA)、フィルターメディア酵素免疫アッセイ(MEIA)、蛍光連結免疫吸着アッセイ(FLISA)、凝集免疫アッセイおよび(Luminex(商標) LabMPAのような)多重蛍光免疫アッセイなどが含まれる。一般的な免疫アッセイの総説として、また、Methods in Cell Biology: Antibodies in Cell Biology, 37巻(Asai, ed. 1993)、Basic and Clinical Immunology (Stites & Terr, eds., 7th ed. 1991) を参照のこと。

40

## 【0150】

50

一般的に、被験者から得た試験試料は、抗原に特異的に結合する抗体を接觸させることができる。場合によって、抗体を、その抗体と試験試料を接觸させる前に、複合体の洗浄および続く単離を促進するために、固体支持体に固定化可能である。固体支持体の例には、たとえば、マイクロタイタープレート、ガラス顕微鏡スライドまたはカバースリップ、スティック、ビーズまたはマイクロビーズの形態での、ガラスまたはプラスチックが含まれる。抗体はまた、以上のように記述されたプローブ基質またはProtein Chip(商標)アレイに結合させることも可能である(たとえば、Xiao et al., Cancer Research 62: 6029-6033 (2001)を参照のこと)。

#### 【0151】

抗体との試料のインキュベートしたの後、混合液を洗浄し、形成した抗体-抗原複合体を検出可能である。このことは、洗浄した混合液を、検出試薬と一緒にインキュベートすることによって達成可能である。本検出試薬は、たとえば、検出可能な標識で標識化された二次抗体であってよい。検出可能な標識に関して、当分野で公知の任意の検出可能な標識を使用可能である。たとえば、検出可能な標識は、(たとえば、<sup>3</sup>H、<sup>125</sup>I、<sup>35</sup>S、<sup>14</sup>C、<sup>32</sup>Pおよび<sup>33</sup>Pのような)放射活性標識、(たとえばホースラディッシュペルオキシダーゼ、アルカリホスファターゼ、グルコース6-リン酸デヒドロゲナーゼなどのような)酵素標識、(たとえばアクリジニウムエステル類、アクリジニウムチオエステル類、アクリジニウムスルホニアミド類、フェナンスリジニウムエステル類、ルミナール、イソルミノールなどのような)化学発光標識、(たとえばフルオレセイン(たとえば5-フルオレセイン、6-カルボキシフルオレセイン、3'-6-カルボキシフルオレセイン、5(6)-カルボキシフルオレセイン、6-カルボキシフルオレセイン、6-テトラクロロフルオレセイン、フルオレセインイソチオシアネートなど)のような蛍光標識、ローダミン、ヒコビリタンパク質、R-フィコエリスリン、量子ドット(たとえば硫化亜鉛キャップカドミウムセレナイト)、温度標識、または免疫-ポリメラーゼ連鎖反応標識であってよい。標識への導入、標識化手順および標識の検出は、モレキュラー プローブス社(Molecular Probes Inc., Eugene, Oregon)によって発行された混合ハンドブックおよびカタログである、Polak and Van Noorden, Introduction to Immunocytochemistry, 2nd ed., Spring Verlag, N.Y. (1997)およびHaugland, Handbook of Fluorescent Probes and Research Chemicals (1996)で見つけられる。あるいは、試料中のマーカーを、間接的アッセイを用いて、たとえば二次、標識化抗体を用いて、結合したマーカー-特異的抗体を検出する、および/またはたとえば抗原の異なるエプトープに結合するモノクローナル抗体を混合液と同時にインキュベートする、競合または阻害アッセイにて検出可能である。

#### 【0152】

アッセイを通じて、インキュベーションおよび/または洗浄段階が、試薬の各混合の後に必要となりうる。インキュベーション段階は、約5病間~数時間で変化可能であり、好みしくは、約5分間~約24時間である。しかしながら、インキュベーション時間は、アッセイの形式、バイオマーカー(抗原)、溶液の容量、濃度などに依存しうる。10~40のような温度の範囲で実施可能であるけれども、通常、アッセイは、室温にて実施する。

#### 【0153】

免疫アッセイ技術は当分野で周知であり、適用可能な技術の一般的総説が、上記Harrow & Laneにて見ることができる。

#### 【0154】

免疫アッセイを、被験者からの試料中の抗原の試験量を決定するために使用可能である。まず、試料中の抗原の試験量を、以上で記述した免疫アッセイ法を用いて検出可能である。抗原が試料中に存在したならば、以上で記述した適切なインキュベーション条件下、抗原に特異的に結合する抗体との抗体-抗原複合体を形成する。抗体-抗原複合体の量を

10

20

30

40

50

、標準と比較して決定可能である。抗原に対するA U Cをついで、限定はしないが、R O C解析のような、公知の技術を用いて計算可能である。あるいは、D F Iを計算可能である。A U Cが約0.5より大きいか、またはD F Iが約0.5未満の場合、免疫アッセイを使用して、(がん、好ましくは肺がんのような)疾患の被験者を、正常(または良性)被験者から判別可能である。

#### 【0155】

多数の抗原に対する免疫アッセイキットが市販されている。たとえば、サイトケラチン19を定量化するためのキットが、H. Hoffmann-La Roche Ltd, (Basel, Switzerland) および Akitengesellschaft, Hennigsdorf, Germanyより入手可能であり、サイトケラチン18を定量化するためのキットが、IDL Biotech AD, Bromma, Swedenより、CA125k、CEA SCCおよびCA19-9を定量化するためのキットが、それぞれ Abbott Diagnostics, Abbott Park, IL、および H. Hoffmann-La Roche Ltd. から、血清アミロイドAおよびアポリポタンパク質CIIIを定量化するためのキットが、Linco Research, Inc. St. Charles, MOから、ProgRPを定量化するためのキットが、Advanced Life Science Institute, Inc. (和光、日本)、およびIBL イムノ-バイオロジカル ラボラトリーズ (IBL Immuno-Biological Laboratories, Hamburg, Germany) から、そして1抗トリプシンを定量化するためのキットが、オートイムノ ダイアグノスティカ GM BH (Autoimmune Diagnostics GMBH, Strassberg, Germany) およびゲンウェイ バイオテック社 (GenWay Biotech, Inc., San Diego, CA) から入手可能である。

#### 【0156】

試験試料中の1つ以上の抗体の存在および定量化を、以上で記述したようなものと同様の免疫アッセイを用いて決定可能である。そのような免疫アッセイは、以上で記述した抗体および抗原の役割が逆であるという事実を除いて、上記免疫アッセイと同様の様式で実施する。たとえば、実施可能である免疫アッセイの1つの型は、自己抗体ビーズアッセイである。本アッセイにおいて、(バイオモル インターナショナル L. P. (Biomo1 International L. P., Plymouth Landing, Pennsylvania) から購入可能な)市販されているp53のような抗原を、固体支持体、たとえばビーズ、プラスチックマイクロプレート、ガラス顕微鏡スライドまたはカバースリップ、またはタンパク質抗原に結合するニトロセルロースのような物質からなる膜に、当分野で公知の日常の技術を用いて、または本明細書実施例3で記述した技術および方法を用いて、固定化可能である。あるいは、抗原が市販されていない場合、がん細胞株(好ましくは肺がん細胞株)、または被験者の自身のがん組織(好ましくは肺がん組織)から抗原を精製してよく(S-H Hong, et al., Cancer Research 64: 5504-5510 (2004)を参照のこと)、またはcDNAクローンより発現させてよい(Y-L Lee, et al., Clin. Chem. Acta 349: 87-96 (2004)を参照のこと)。抗原を含むビーズをついで、試験試料と接触させる。試験試料を、結合した抗原を含むビーズとともにインキュベートした後、ビーズを洗浄し、形成した任意の抗体-抗原複合体を検出する。この検出は、以上で記述したように実施可能であり、すなわち、洗浄したビーズを検出試薬とともにインキュベートすることによって実施可能である。本検出試薬は、たとえば検出可能標識によって標識化した、(限定はしないが、抗-ヒト免疫グロブリンG (IgG)、抗-ヒト免疫グロブリンA (IgA)、抗-ヒト免疫グロブリンM (IgM)のような)二次抗体であつてよい。検出の後、抗体-抗原複合体の量を、以上で記述したように、シグナルを、標準によって産出されたものと比較することによって決定可能である。あるいは、抗体-抗原複合体を、免疫グロブリンの多価特性の利点を利用することによって決定可能である。抗

体 - 抗原複合体を、抗 - ヒト抗体と反応させる代わりに、抗体 - 抗原複合体を、固体相に結合した抗原と同一のエピトープを含む、検出可能な標識で標識化した可溶性抗原に暴露可能である。任意の占有されていない抗体結合部位が、（検出可能な標識にて標識化された）可溶性抗原に結合する。洗浄の後、検出可能標識を、当業者に公知の日常の技術を用いて検出する。上記方法のいずれかによって、試験試料中の特異的抗体の、感度がよく、特異的な定量化が可能になる。抗体（およびしたがって、被験者の肺がんの検出のための、自己抗体のような抗体の利用）に対する A U C をついで、限定はしないが R O C 解析のような、当業者に公知の日常の方法を用いて計算可能である。あるいは、D F I を計算可能である。A U C が約 0.5 より大きいか、または D F I が約 0.5 未満の場合、免疫アッセイを使用して、（がん、好ましくは肺がんのような）疾患の被験者を、正常（または良性）被験者から判別可能である。

#### 【 0 1 5 7 】

関心領域の存在および量を、質量スペクトル技術を用いて決定可能である。質量分析を用いて、本発明者らは、試験試料中の肺がんの診断およびスクリーニングにおける補助として有用である、212 の関心領域を発見した。特に、質量スペクトル技術を、試験試料中の1つ以上のバイオマーカーを検出および定量するために使用する場合、試験試料はまず質量スペクトル解析のために調製されなければならない。試料調製を、種々の方法、しかしまっとも一般的に使用される、固相に結合した1つ以上の吸着剤と試料を接触させることを含む方法で実施可能である。吸着剤は、アニオン性またはカチオン性基、疎水性基、金属リガンドありまたはなしの金属キレーティング基、ポリクローナルまたはモノクローナルいずれかの抗体、または同族抗体に結合するのに好適な抗原であってよい。固相は、金属、ガラスまたはプラスチックからなる平面表面であってよい。固相はまた、マイクロ粒子性質を有するもの、表面積を増加させるための、マイクロビーズ、非結晶質性粒子、または不溶性ポリマーのいずれか、であってよい。さらに、マイクロ粒子物質は、操作の簡便性のために磁気的でありうる。関心のバイオマーカーを、固相に吸着させ、バルク分子を洗浄によって除去する。質量解析のために、関心のバイオマーカーを、吸着剤に対するバイオマーカーの親和力を減少させる溶媒で、固相より溶出する。ついでバイオマーカーを解析のために質量分析器に導入する。好ましくは、異常スペクトルが同定され、スペクトルを評価する場合に無視する。さらに、上記のような免疫アッセイも使用可能である。免疫アッセイの完了に際して、解析物を、免疫学的表面から溶出して、解析のために質量分析器に導入可能である。

#### 【 0 1 5 8 】

一旦試験試料が調製されたならば、質量解析器に導入する。レーザー脱着イオン化（たとえば M A L D I または S E L D I ）が、固体形態で存在する試料に対する一般的な技術である。本技術において、試料を、標的プレート上で、レーザーエネルギーを吸収し、試料に伝達することにおいて効果的なマトリックスとともに共結晶化する。作製されたイオンを分離し、計数し、公知の質量と電荷のイオンに対して較正する。任意の試料に対して回収した質量データは、特定の質量 / 電荷 (m / z) 比でのイオン計数である。異なる試料調製法および異なるイオン化技術が異なるスペクトルを産出することが予想される。

#### 【 0 1 5 9 】

質量スペクトルデータに対する認定試験は一般的に、本来のデータの最小限の前処理での外れ値解析の正確な処理を含む。外れ値の同定の工程は、生スペクトルの総イオン電流 (T I C) の計算で開始される。T I C 計算の前に、スムージングまたはベースライン補正アルゴリズムは、生スペクトルに対して適用しない。T I C は、検出した質量 (m / z) 範囲にわたる、各 m / z 値での強度を合計することによって計算する。これにより、器具不具合、試料スポットティング問題、および他の同様の欠陥が選別される。T I C に加えて、それぞれの試料の全スペクトルにわたる平均 % C V (パーセント変動係数) を計算する。それぞれの試料に対する反復測定の回数を用いて、検出された質量範囲にわたり、m / z 値ごとに % C V を計算する。次いでこれらの % C V を平均して、特定の試料を代表する平均 % C V を得る。平均 % C V は異常値を特定するための第一の選別工程として用いて

10

20

30

40

50

もよいし、そのように用いなくともよい。一般に、平均%CVが高い反復（30%以上またはその他の任意の許容範囲の値）は再現性が乏しい。

上記のように、それぞれのスペクトルの計算されたTICおよび平均%CVは再現性およびスペクトルの有益な部分を限定するための予測因子として用いることができる。しかし、これらの測定によって、スペクトルのバルク特性に関する良い記述子が得られる一方、それぞれのm/z値での個々の強度などのスペクトルの顕著な特性の再現性についての情報は少しも得られない。この障害は、Wan et al. (J. Am. Soc. Mass Spectrom. 2002, 13, 85-88) に報告されたSpectral Contrast Angle (SCA) 計算の適合によって克服された。SCA計算では、全スペクトルを、個々のm/z値から成るベクターとして扱う。この解釈では、標準的な数式から2つのベクター間のシータ(θ)角が得られる。

【0160】

【数1】

$$\cos(\theta) = \mathbf{v}_1 \bullet \mathbf{v}_2 / \left( \sqrt{\mathbf{v}_1 \bullet \mathbf{v}_1} * \sqrt{\mathbf{v}_2 \bullet \mathbf{v}_2} \right).$$

同じようなスペクトルでは、シータは小さく、0に近くなる。

【0161】

使用する際、試料複製についてかまたは特定の群（例えば、がん群）の中の全試料についての平均スペクトルを最初に計算することによって計算と比較の総数を減らす。次に、それぞれのスペクトルと計算された平均スペクトルとの間のSCAを計算する。平均スペクトルと大幅に異なるスペクトルは、もし下記の基準に合致すれば、異常値と見なす。

【0162】

1つの予測因子では質量スペクトルを完全に説明するのに不十分であるので、異常値を選択するためには1つ以上の予測因子を用いることが好ましい。多変量異常値分析は多数の予測因子を用いて行なうことができる。これらの予測因子は、TIC、平均%CV、およびSCAであってよいが、これらに限定されない。JMP（商標）統計パッケージ（SAS Institute Inc., Cary, NC）を用い、群（例えば、がん群）中のそれぞれの反復測定について、マハラノビスの距離を計算した。基準値（信頼限界ではない）は、観測値の約95%がその値を下回るように算出することができる。基準値を上回る残りの5%は異常値と見なし、さらなる分析からは除外する。

【0163】

質量スペクトルデータの限定後、通常スペクトルを正規化し、TICがデータセット中の全てのスペクトルに対して同じになるように強度を見積もるか、全スペクトルにおける1つのピークに応じて強度を見積もる。

【0164】

正規化した後、質量スペクトルを一連の強度特性に減少する。他の応用では、m/z値におけるスペクトル強度一覧へのこれらの限定は生体分子と関連する。別種の特性、関心領域またはROIを用いることが好ましい。

【0165】

関心領域は、被験者の2つ以上のデータセット間の比較の産物である。データセットは、関心の群（例えば、罹患群、非罹患群）を表す。全試料について、各々のm/z値における強度値に対してt検定を行なう。t検定のp値が技師指定の基準値未満であるm/z値を特定する。特定されたm/z値の、連続しているものをグループ化して被験者領域として定義する。ROIおよび連続した群の任意の許容間隙を形成するために必要な連続m/z値の最小数は、使用者が規定することができる。ROIを限定する他の因子は、2つの群の平均値の比率の対数の絶対値である。この値がある基準のカットオフ値（10進法の対数を用いた場合、約0.6）を上回る場合、質量対電荷の場所はROIに含める被験者になる。ROI法を用いる利点は、2種類のスペクトル間の高強度のパターンの違いの

10

20

30

40

50

フラグになるだけでなく、肩などのよりわずかな違いや、ピーク発見法では見逃してしまうであろう非常に低い強度も発見することができる。

#### 【0166】

被験者領域が決定されたら、R O I の範囲内の  $m/z$  値の平均値または中央値をその領域の特定因子として用いることが多い。それぞれの領域はデータセットを識別する潜在的マーカーである。さまざまなパラメーター（例えば、総強度、最大強度、中央強度、または平均強度）は、試料データから抽出することができ、R O I と関連する。このように、それぞれの試料スペクトルは何千もの  $m/z$ 、212 の R O I に対する強度対およびその特定因子、強度関数対から限定された。これらの記述子はデータ分析法の入力変数として用いる。

10

#### 【0167】

本発明の方法には、場合により、試験試料を得る前かまたは試験試料を得た後、そして試験試料中の1つ以上のバイオマーカーを同定し定量化する前または、試験試料中の1つ以上のバイオマーカーを同定して定量化した後に、被験者から少なくとも1つの生物測定パラメーターを得る段階を含むことができる。被験者から得る生物測定パラメーターの数は決定的に重要ではない。例えば、1、2、3、4、5、6、7、8、9、10などの生物測定パラメーターを被験者から得ることができる。あるいは、本発明の方法に被験者から任意の生物測定パラメーターを得る段階を含む必要はない。被験者から得る好ましい生物測定パラメーターは被験者の喫煙歴、具体的には被験者の喫煙パックイヤー（pack-years）である。被験者から得ることができる他の生物測定パラメーターとしては、年齢、発がん物質への暴露、性、家族の喫煙歴などが挙げられるが、それらに限定されない。

20

#### 【0168】

前記のように、本発明の方法においては、試験試料を分析してパネル中に含まれる1つ以上のバイオマーカーの存在を判定する。試験試料中にバイオマーカーが存在すると判定されたら、その検出されたバイオマーカーを定量化する（本明細書で前述した手法を用いて行なう）。試験試料中のそれぞれのバイオマーカーの量が定量化されたら、次に定量化されたそれぞれのバイオマーカーの量を特定のバイオマーカーに対する所定のカットオフ（通常は整数などの値または数字、あるいは本明細書では「分割点」（“split point”）と称する）と比較する。本発明の方法において採用される所定のカットオフは、当技術分野で公知の日常の手法（例えば、多変量解析（図1を参照されたい）、Transformedロジスティック回帰分析、分割スコア法（Split and Score Method）、またはそれらの任意の組合せなどが挙げられるが、それらに限定されない）を用いて決めることができる。例えば、分割スコア法（Split and Score Method）を用いた場合、規定のカットオフの値または数字は得られるべき所望の結果次第で決められる。得られるべき所望の結果が、関心の群のそれぞれのマーカーの正分類の正確さを最大にする、すなわち、肺がん発症の危険性がある被験者と肺がん発症の危険性がない被験者を正確に識別する目的のものであれば、こうした所望の結果に基づいたバイオマーカーに対する所定のカットオフのために、特定の値または数字が選択される。一方、所望の結果が、それぞれのマーカーの感度を最大にする目的のものであれば、所定のカットオフのための異なる値または数値が、その所望の結果に基づくバイオマーカーに対して選択される。同様に、所望の結果がそれぞれのマーカーの特異性を最大にする目的のものであれば、所定のカットオフのための異なる値または数値が、その所望の結果に基づくバイオマーカーに対して選択される。試験試料中に存在する任意のバイオマーカーの量が定量化されれば、この情報は、当業者が所望の結果次第で決まるそれぞれのバイオマーカーに対する適切な所定のカットオフを所定の手法を用いて決定するために使用するR O C 曲線、A U C およびの他の情報を作成するために用いることができる。

30

それぞれのバイオマーカーの量を所定のカットオフと比較した後、スコア（すなわち数字で、例えば0から100などの任意の整数であってよい）を、その比較に基づいてそれぞれのバイオマーカーに割り当てる。さらに、1つ以上のバイオマーカーに加え、被験者

40

50

に対して1つ以上の生物測定パラメーターを得た場合、それぞれの生物測定パラメーターを、その生物測定パラメーターに対する所定のカットオフと比較する。任意の生物測定パラメーターに対する所定のカットオフは、1つ以上のバイオマーカーに対する所定のカットオフの決定に関して本明細書に記載したものと同様の手法を用いて決定することができる。バイオマーカーの比較と同様に、スコア（すなわち数字で、例えば0から100などの任意の整数であってよい）を前述の比較に基づいて生物測定パラメーターに割り当てる。

#### 【0169】

あるいは、上記のスコア法を用いる代わりに、分割加重スコア法（S p l i t a n d W e i g h t e d S c o r i n g M e t h o d）を用いることができる。分割加重スコア法（S p l i t a n d W e i g h t e d S c o r i n g M e t h o d）を用いた場合、試験試料中のそれぞれのバイオマーカーの量を定量化したら、その、試験試料中に検出されたそれぞれのバイオマーカーの量をその特定のバイオマーカーに対する多数の所定のカットオフと比較する。得られる種々のカットオフ全てから、1つのスコア（すなわち数字で、例えば0から100などの任意の整数であってよい）をそのバイオマーカーに割り当てる。この分割加重スコア法（S p l i t a n d W e i g h t e d S c o r i n g M e t h o d）はまた、1つ以上の生物測定パラメーター（b i o m e t r i c p a r a m e t e r）にも同様に使用できる。

#### 【0170】

定量化したそれぞれのバイオマーカーに対して、または場合によっては被験者から得られた任意の生物測定パラメーターに対して、スコアが割り当てられたら、それぞれのバイオマーカー、またはそれぞれのバイオマーカーおよびそれぞれの生物測定パラメーターに対するスコアを総合して被験者の総スコアを出す。次にこの総スコアを既定の総スコアと比較する。この比較に基づいて、被験者に肺がんの危険性があるか否かを決定することができる。この、被験者に肺がんを発症する危険性があるか否かの決定は、総スコアが既定の総スコアよりも高いか、または低いか、否かに基づいてよい。例えば、既定の総スコアに割り当てられた値に基づいて、規定の総スコアよりも総スコアが高い被験者は、危険性が高いと考えてよく、よってさらなる試験または追加処置へと差し向けられうる。本方法で用いる既定の総スコア（あるいは、本明細書において「基準点（t h r e s h o l d）」と称する）は、バイオマーカーに対する既定のスコアに関して前記したものと同様の手法を用いて決定することができる。例えば、図5は3つのR O C曲線である。これらのR O C曲線はそれぞれ組み合わせたマーカーからの単一の出力を示すが、しかし、ただひとつのマーカーからも同様のR O C曲線ができると思われる。R O C曲線は、一方の端、低感度および低擬陽性率（1特異性）から他方の端、高感度および高偽陽性率までにわたる。これらの2端間の曲線の形状は大きく変化しうる。感度が少なくとも90%であることが求められる方法の場合、図5に示すR O C曲線に基づき、選択した曲線によって偽陽性率は60~70%になると思われる。偽陽性が多くとも10%であることが求められる方法の場合には、選択した曲線によって、感度は40~50%になると思われる。これらの方法はどちらもマーカーの同じパネルから得られるがしかし、異なる臨床成績特性を規定するために、パネルの基準点（または既定の総スコア）を変更する。計算上、R O C曲線上のそれぞれの点の下にあるのがデータ範囲の一方の端からデータ範囲の他方の端まで動く基準点（または既定の総スコア）である。基準点（または既定の総スコア）がデータ範囲の下端にある場合、全ての試料は陽性であり、R O C曲線上に高感度かつ高偽陽性率の点が生じる。基準点（または既定の総スコア）がデータ範囲の上端にある場合、全ての試料は陰性であり、R O C曲線上に低感度かつ低擬陽性の点が生じる。方法は所望の臨床特性、例えば感度の最小値（すなわち、90%）、特異性の最小値（すなわち90%）、またはその両方などを有することを求められることが多い。マーカーの基準点（または既定の総スコア）を変えることが、所望の臨床特性を得る助けになる。

#### 【0171】

上記の工程、（a）所定のカットオフとパネル中のそれぞれのバイオマーカーの量（ま

10

20

30

40

50

たは、分割加重スコア法 (S p l i t a n d W e i g h t e d S c o r i n g M e t h o d ) を用いる場合には多数の所定のカットオフとの比較、その比較に基づいてそれぞれのバイオマーカーにスコア (または、分割加重スコア法 (S p l i t a n d W e i g h t e d S c o r i n g M e t h o d ) を用いる場合には可能性のある多数のスコアの 1 つから得たスコア) の割り当てること、パネル中のそれぞれの生物測定パラメーターに対して割り当てられたスコアを総合して被験者の総スコアを出すこと、その総スコアを既定の総スコアと比較し、そして総スコアに基づいて被験者が肺がんの危険性を有するかどうかを決定すること、または (b) 少なくとも 1 つの生物測定パラメーターと所定のカットオフ (または、分割加重スコア法 (S p l i t a n d W e i g h t e d S c o r i n g M e t h o d ) を用いる場合には多数のカットオフ)との比較、およびその比較に基づいてそれぞれの生物測定パラメーターにスコア (または、分割加重スコア法 (S p l i t a n d W e i g h t e d S c o r i n g M e t h o d ) を用いる場合には可能性のある多数のスコアの 1 つから得たスコア) を割り当てること、パネル中のそれぞれのバイオマーカーの量と所定のカットオフとの比較、その比較に基づいてそれぞれのバイオマーカーにスコアを割り当てること、それぞれの生物測定パラメーターに割り当てられたスコアと定量化したそれぞれのバイオマーカーに割り当てられたスコアを総合して被験者の総スコアを出すこと、総スコアを既定の総スコアと比較し、そして総スコアに基づいて被験者が肺がんの危険性を有するかどうかを決定すること、は、例えばヒトによって手動で行なうことができ、またはコンピュータプログラムもしくはアルゴリズムを、入力装置、記憶装置、処理装置、表示装置、および出力装置などの必要なハードウェアと一緒に用いて完全にまたは部分的に行なうことができる。10

#### 【 0 1 7 2 】

例示のみの目的で、本発明の方法がどのように実行されるかをここで示す。この実施例においては、8 つのバイオマーカーを含むパネルと分割スコア法 (S p l i t a n d S c o r e M e t h o d ) を用いて、患者を試験し、その患者が肺がんを有する尤度を決定する。パネル中のバイオマーカーは、サイトケラチン 19、C E A、C A 1 2 5、C A 1 5 - 3、C A 1 9 - 9、S C C、p r o G R P、およびサイトケラチン 18 である。パネルに対する既定の総スコア (または基準点 (t h r e s h o l d)) は 3 である。患者から試験試料を得た後、その患者の試験試料中の 8 つのバイオマーカー (サイトケラチン 19、C E A、C A 1 2 5、C A 1 5 - 3、C A 1 9 - 9、S C C、p r o G R P、およびサイトケラチン 18) それぞれの量を定量化する。この実施例のために、試験試料中の 8 つのバイオマーカーそれぞれの量を以下の通り決定した。サイトケラチン 19 : 1 . 9 5、C E A : 2 . 7 5、C A 1 2 5 : 1 5 . 2 6、C A 1 5 - 3 : 1 1 . 9 2、C A 1 9 - 9 : 9 . 2 4、S C C : 1 . 0 6、p r o G R P : 2 5 . 2 9、サイトケラチン 18 : 6 1 . 1 3。次いでこれらのバイオマーカーそれぞれの量を対応する所定のカットオフ (または分割点) と比較する。バイオマーカーそれぞれに対する所定のカットオフは、サイトケラチン 19 : 1 . 8 9、C E A : 4 . 8 2、C A 1 2 5 : 1 3 . 6 5、C A 1 5 - 3 : 1 3 . 0 7、C A 1 9 - 9 : 1 0 . 8 1、S C C : 0 . 9 2、p r o G R P : 1 4 . 6 2、そしてサイトケラチン 18 : 5 7 . 3 7 である。対応する所定のカットオフ (分割点) よりも量が多いバイオマーカーそれぞれに対して、スコア 1 を与えることができる。対応する所定のカットオフよりも量が少ないかまたは同量であるバイオマーカーにたいしてはスコア 0 を与えることができる。その結果、前述の比較に基づいて、それぞれのバイオマーカーには以下の通りスコアが割り当てられることになる。サイトケラチン 19 : 1 、C E A : 0、C A 1 2 5 : 1、C A 1 5 - 3 : 0、C A 1 9 - 9 : 0、S C C : 1、p r o G R P : 1、そしてサイトケラチン 18 : 1。次いで 8 つのバイオマーカーそれぞれのスコアを数学的に (すなわち、バイオマーカーそれぞれのスコアを合計することによって) 総合して患者の総スコアにたどり着く。本患者の総スコアは 5 (この総スコアは以下のように計算される。 $1 + 0 + 1 + 0 + 0 + 1 + 1 + 1 = 5$ )。この患者の総スコアを既定の総スコア、3 と比較する。既定のスコア 3 よりも大きい総スコアは、その患者について陽性の結果であると示すことになる。3 よりも小さい、または同じ総スコアは、その患20304050

者について陰性の結果であると示すことになる。本実施例では、患者の総スコアは3よりも大きいので、この患者は陽性の結果であると考えられ、よって肺がんの兆候または疑いに対するさらなる試験に回されることになる。一方、患者の総スコアが2であったら、その患者は陰性の結果であると考えられ、それ以上のどんな試験にも回されないことになる。

【0173】

別の実施例においては、前述の8つのバイオマーカーパネルを再度用いるが、しかし本実施例では、分割加重スコア法 (S p l i t a n d W e i g h t e d S c o r i n g M e t h o d) を採用する。本実施例においては、パネルに対する既定の総スコア(または基準点 (t h r e s h o l d))は11.2であり、試験試料中の定量化されたバイオマーカーの量は前述したものと同じである。次いでそれぞれのバイオマーカーの量を、そのバイオマーカーそれぞれに対する3つの異なる所定のカットオフと比較する。例えば、それぞれのバイオマーカーに対する所定のカットオフを表Aに示す。

【0174】

【表1】

表A

	CEA	サイトケラチン18	ProGRP	CA15-3	CA125	SCC	サイトケラチン19	CA19-9
50%特異性に対する既定のカットオフ値	2.02	47.7	11.3	16.9	15.5	0.93	1.2	10.6
75%特異性に対する既定のカットオフ値	3.3	92.3	18.9	21.8	27	1.3	1.9	21.9
90%特異性に対する既定のカットオフ値	4.89	143.3	28.5	30.5	38.1	1.98	3.3	45.8
50%特異性以上のスコア	0	0	0	0	0	0	0	0
50%特異性以上のスコア	2.68	2.6	2.48	1.16	2.68	2.48	4.2	1.1
75%特異性以上のスコア	5.36	5.2	4.96	2.32	5.36	4.96	8.4	2.2
90%特異性以上のスコア	13.4	13	12.4	5.8	13.4	12.4	21	5.5

【0175】

したがって、それぞれのバイオマーカーに対して4つの可能性のあるスコアが得られる。定量化されたそれぞれのバイオマーカーの量を上記表Aに示される所定のカットオフ(分割点)と比較する。例えばCEAについてみると、試験試料中の定量化されたCEA量は2.75であるから、表A中の50%特異性に対する所定のカットオフ値2.02と75%特異性に対する所定のカットオフ値3.3との間に収まる。故に、CEAにはスコア2.68が割り当てられる。このことを残るバイオマーカーについても繰り返し、同様に評価してそれぞれに以下のスコアを割り当てる。サイトケラチン18:2.6、proGRP:4.96、CA15-3:0、CA125:0、SCC:2.48、サイトケラチ

10

20

30

40

50

ン 1 9 : 8 . 4 、 C A 1 9 - 9 : 0 。次いで 8 つのバイオマーカーそれぞれのスコアを数学的に（すなわち、バイオマーカーそれぞれのスコアを合計することにより）総合して患者の総スコアにたどり着く。患者の総スコアは 2 1 . 1 2 (この総スコアは次のように計算される。  $2 . 6 8 + 2 . 6 + 4 . 9 6 + 0 + 0 + 2 . 4 8 ? 8 . 4 + 0 = 2 1 . 1 2$  ) 。次に患者の総スコアを既定の総スコア、 1 1 . 2 と比較する。既定の総スコア 1 1 . 2 よりも大きい総スコアは、その患者については陽性の結果であると示すことになる。 1 1 . 2 よりも小さいかまたは等しい総スコアは、その患者については陰性の結果であると示すことになる。本実施例では、患者の総スコアは 1 1 . 2 よりも大きいので、この患者は陽性の結果を有すると考えられ、したがって、肺がんの兆候または疑いに対するさらなる試験に回されることになる。

10

#### 【 0 1 7 6 】

さらに、本明細書に記載の分割加重スコア法 ( S p l i t a n d W e i g h t e d S c o r i n g M e t h o d ) はまた、被験者から得られた 1 つ以上のマーカーにスコアをつけるためにも使用できる。好ましくは、そのようなマーカー ( 1 つ以上のバイオマーカー、 1 つ以上の生物測定パラメーター、またはバイオマーカーと生物測定パラメーターとの組合せのいずれにせよ ) は、被験者にがんまたは他のなんらかの疾病などの病状を発症する危険性があるかどうかを診断または評価するための補助として用いることができる。危険性を評価する為にマーカーを用いる、または用いることが可能な病状を本明細書の方法で用いることができる。そのような方法には、下記の段階が含まれる。

#### 【 0 1 7 7 】

- a . 被験者から少なくとも 1 つのマーカーを得ること、
- b . 該被験者から得たマーカーの量の定量化、
- c . 定量化されたそれぞれのマーカーの量を該マーカーに対するいくつかの所定のカットオフと比較し、その比較に基づいてそれぞれのマーカーのスコアを割り当てること、および
- d . 段階 C において、定量化されたそれぞれのマーカーに対して割り当てたスコアを総合して該被験者の総スコアを出すこと。

20

#### 【 0 1 7 8 】

好ましくは、この方法には下記の段階が含まれる。

#### 【 0 1 7 9 】

- a . 被験者から少なくとも 1 つのマーカーを得ること、
- b . 該被験者から得たマーカーの量の定量化、
- c . 定量化されたそれぞれのマーカーの量を該マーカーに対するいくつかの所定のカットオフと比較し、その比較に基づいてそれぞれのマーカーのスコアを割り当てること、
- d . 段階 C において、定量化されたそれぞれのマーカーに割り当てたスコアを総合して該被験者の総スコアを出すこと、
- e . 段階 d において決定した総スコアを既定の総スコアと比較すること、および、
- f . 段階 e において決定した総スコアに基づいて、該被験者が病状を発症する危険性を有するかどうかを決定すること。

30

#### 【 0 1 8 0 】

##### D F I

本明細書において前述のように、出願者らは、 1 つ以上のバイオマーカー、またはバイオマーカーと生物測定パラメーターの組合せを検出および定量化することが、患者の肺がんを診断する補助として有効であることを見いだした。さらに発明者らは、本明細書に記載の 1 つ以上のバイオマーカー、および 1 つ以上のバイオマーカーと 1 つ以上の生物測定パラメーターの組合せで、肺がんに関連する D F I を有するものが約 0 . 5 未満、好ましくは約 0 . 4 未満、さらに好ましくは約 0 . 3 未満、そしてさらに好ましくは約 0 . 2 未満であることを見いだした。表 2 5 ~ 2 9 は、 D F I が約 0 . 5 未満、約 0 . 4 未満、約 0 . 3 未満、および約 0 . 2 未満である、さまざまなバイオマーカーまたはバイオマーカーと生物測定パラメーターの組合せを含むパネルの例を示す。

40

50

## 【0181】

## キット

本発明の方法を実行する際に使用するために、1つ以上のバイオマーカー、1つ以上の免疫反応性サイクリンE2ポリペプチド、生物測定パラメーター、およびそれらの任意の組合せをキット(パネルなどの)形成に適用できる。一態様においては、キットには配列番号1、配列番号3、配列番号4、配列番号5、またはそれらの組合せからなる群から選択されたペプチドを含めることができる。

## 【0182】

別の態様においては、キットには免疫反応性サイクリンE2またはその任意の組合せに対する少なくとも1種の抗体を含めることができる。

10

## 【0183】

また別の態様においては、キットには下記を含めることができる。

## 【0184】

(a) 試験試料中の1種以上種の抗原を定量化するための少なくとも1種の抗体を含む試薬(ただし、上記抗原としては、サイトケラチン8、サイトケラチン19、サイトケラチン18、CEA、CA125, CA15-3、SCC、CA19-9、proGRP、血清アミロイドA、-1抗トリプシン、およびアポリポタンパク質CII)、

(b) 試験試料中の少なくとも1種の抗体を定量化するための1種以上種の抗原を含む試薬(ただし、上記抗体としては、抗p53、抗TMP21、抗NPCI1ICドメイン、抗TMOD1、抗CAMK1、抗RGS1、抗PACSI1、抗RCV1、抗MAPKAPK3、および免疫反応性サイクリンE2に対する少なくとも1つの抗体)、

20

(c) ACN9459、Pub11597、Pub4789、TFA2759、Pub3743、Pub8606、Pub4487、Pub4861、Pub6798、Tfa6453、およびHic3959からなる群から選択された1つ以上の関心領域を定量化するための試薬、および

(d) 試験試料中の、定量化されたそれぞれの抗原、抗体、および関心領域の量を総合して所定のカットオフ(または、多数の所定のカットオフ)に対する比較を行い、その比較に基づいて、定量化されたそれぞれの抗原、抗体、および関心領域に対するスコア(または、いくつか可能性のあるスコアの1つから得たスコア)を割り当て、定量化されたそれぞれの抗原、抗体、および関心領域に対して割り当てられたスコアを総合して総スコアを得て、その総スコアを既定の総スコアと比較し、患者が肺がんに罹患しているかどうかを決定する一補助としてその比較を用いる、という段階を行なうための1つ以上のアルゴリズムまたはコンピュータプログラム。

30

## 【0185】

あるいは、1つ以上のコンピュータプログラムの代わりに、人間が上記段階を手動で行なうための1つ以上の説明書を備えることもできる。

## 【0186】

1つ以上の関心領域を定量化するためにキットに含まれる試薬には、パネルに含まれる少なくとも1つの関心領域に結合して保持する吸着剤、その吸着剤に関連して用いる固体担体(ビーズなど)、1つ以上の検出可能な標識などを含めることができる。吸着剤は、分析化学および免疫化学において使用される多くの吸着剤のうち任意のものでよく、金属キレート、陽イオン基、陰イオン基、疎水基、抗原、および抗体が挙げられる。

40

## 【0187】

さらに別の態様においては、キットには下記を含めることができる。

## 【0188】

(a) 試験試料中の1種以上種の抗原を定量化するための少なくとも1種の抗体を含む試薬(ただし、上記抗原としては、サイトケラチン19、サイトケラチン18、CA19-9、CEA、CA15-3、CA125、SCC、およびproGRP)、

(b) ACN9459、Pub11597、Pub4789、TFA2759、TFA9133、Pub3743、Pub8606、Pub4487、Pub4861、Pub6

50

798、Tfa6453、およびHic3959からなる群から選択された1つ以上の関心領域を定量化するための試薬、および

(c) 試験試料中の、定量化されたそれぞれの抗原および被験者領域の量を総合して所定のカットオフ(または、いくつかの所定のカットオフ)に対する比較を行い、その比較に基づいて、定量化されたそれぞれの抗原および被験者領域に対するスコア(または、いくつかの可能性のあるスコアの1つから得たスコア)を割り当て、定量化されたそれぞれの抗原および被験者領域に対して割り当てられたスコアを総合して総スコアを得て、その総スコアを既定の総スコアと比較し、患者が肺がんに罹患しているかどうかを決定する一補助としてその比較を用いる、という段階を行なうための1つ以上のアルゴリズムまたはコンピュータプログラム。あるいは、1つ以上のコンピュータプログラムの代わりに、人間が上記段階を手動で行なうための1つ以上の説明書を備えることもできる。1つ以上の関心領域を定量化するためにキットに含まれる試薬には、パネルに含まれる少なくとも1つの関心領域に結合して保持する吸着剤、その吸着剤に関連して用いる固体担体(ビーズなど)、1つ以上の検出可能な標識などを含めることができる。10

#### 【0189】

キットには下記の抗原および関心領域を定量化するために必要な試薬が含まれることが好ましい。(a) サイトケラチン19およびCEAと、ACN9459、Pub11597、Pub4789およびTfa2759、(b) サイトケラチン19およびCAEと、ACN9459、Pub11597、Pub4789、Tfa2759、およびACN9459、および(c) サイトケラチン19、CEA、CA125、SCC、サイトケラチン18、およびProGRPと、ACN9459、Pub11597、Pub4789、およびTfa2759。20

#### 【0190】

別の態様においては、キットには下記を含めることができる。(a) 試験試料中の1種以上種の抗原を定量化するための少なくとも1種の抗体を含む試薬(ただし、上記抗原としては、サイトケラチン19、サイトケラチン18、CA19-9、CEA、CA15-3、CA125、SCC、およびProGRP)、(b) 試験試料中の、定量化されたそれぞれの抗原の量を総合して所定のカットオフ(または、いくつかの所定のカットオフ)に対する比較を行い、その比較に基づいて、定量化されたそれぞれの抗原に対するスコア(または、いくつかの可能性のあるスコアの1つから得たスコア)を割り当て、定量化されたそれぞれの抗原に対して割り当てられたスコアを総合して総スコアを得て、その総スコアを既定の総スコアと比較し、患者が肺がんに罹患しているかどうかを決定する一補助としてその比較を用いる、という段階を行なうための1つ以上のアルゴリズムまたはコンピュータプログラム。30

#### 【0191】

あるいは、1つ以上のコンピュータプログラムの代わりに、人間が上記段階を手動で行なうための1つ以上の説明書を備えることもできる。キットには1つ以上の検出可能な標識も含めることができる。キットには下記の抗原を定量化するために必要な試薬が含まれることが好ましい。サイトケラチン19、サイトケラチン18、CA19-9、CEA、CA15-3、CA125、SCC、およびProGRP。40

#### 【0192】

別の態様においては、キットには下記を含めることができる。(a) 1つ以上のバイオマーカーを定量化する試薬(ただし、上記バイオマーカーは、ACN9459、Pub11597、Pub4789、TFA2759、TFA9133、Pub3743、Pub8606、Pub4487、Pub4861、Pub6798、Tfa6453、およびHic3959からなる群から選択される被験者領域)、および、(b) 試験試料中の、定量化されたそれぞれのバイオマーカーの量を総合して所定のカットオフ(または、いくつかの所定のカットオフ)に対する比較を行い、その比較に基づいて、定量化されたそれぞれのバイオマーカーに対するスコア(または、いくつかの可能性のあるスコアの1つから得たスコア)を割り当て、定量化されたそれぞれのバイオマーカーに対して割り当てら50

れたスコアを総合して総スコアを得て、その総スコアを既定の総スコアと比較し、患者が肺がんに罹患しているかどうかを決定する一補助としてその比較を用いる、という段階を行なうための1つ以上のアルゴリズムまたはコンピュータプログラム。あるいは、1つ以上のコンピュータプログラムの代わりに、人間が上記工程を手動で行なうための1つ以上の説明書を備えることもできる。キット中の定量化されるべき関心領域は、Pub 11597、Pub 3743、Pub 8606、Pub 4487、Pub 4861、Pub 6798、Tfa 6453、およびHic 3959からなる群から選択されることが好ましい。1つ以上の関心領域を定量化するためにキットに含まれる試薬には、パネルに含まれる少なくとも1つの関心領域に結合して保持する吸着剤、その吸着剤と関連して用いる個体担体(ビーズなど)、1つ以上の検出可能な標識などを含めることができる。

10

### 【0193】

#### バイオマーカーの同定

本発明のバイオマーカーは、当技術分野の技術者に周知の技術によって単離、精製、および同定することができる。その技術としては、クロマトグラフ法、電気泳動技術、および遠心分離法が挙げられる。これらの技術はCurrent Protocols in Protein Science、J. Wiley and Sons, New York, NY, Coligan et al. (Eds) (2002)およびHarris, E. L. V., S. Angal in Protein Purification Applications: A Practical Approach, Oxford University Press, New York, NY (1990)などに記載されている。

20

### 【0194】

ここで本発明の実施例を、それに限定するものではなく、一例として示す。

### 【0195】

#### (実施例)

患者の血清の臨床試料を採取し(実施例1)、マーカー抗原免疫測定の分析をし(実施例2)、ビーズを用いてマーカー抗体免疫測定の分析(実施例3)またはスライドを用いてマーカー抗体免疫測定の分析をして(実施例4)、質量分析によりバイオマーカーを同定する(実施例5)。同定されたマーカーをさまざまなアルゴリズムを用いて分類し優先順位をつける(実施例6)。これらの優先順位をつけたマーカーを、スコア法を用いて総合し(実施例7)、予測モデルを同定して(実施例8)、臨床的有用性を評価する。肺がん罹患が疑われる患者における肺がん検出を補助する方法を用いた実施例を実施例9に例証している。対象領域の質量分析によって同定されたバイオマーカーは、分析しこの組成を決定して同定した(実施例10)。実施例11は、本発明に従って同定したバイオマーカーを、免疫測定法および免疫質量分析法を用いて検出、測定する方法が記載されている、机上実施例である。

30

### 【実施例1】

### 【0196】

#### 臨床標本

患者の血清の臨床試料を、施設内倫理委員会( Institutional Review Board)で承認されたプロトコールの下で採取した。標本を提供した全ての患者は、採取され本研究に使用される標本についてインフォームド・コンセントを受けた。血清試料は、血清分離管中に入れて室温で15分間、凝固させた。この血餅を遠沈させて試料を2mLアリコートとした。24時間以内にこの試料を-80°で凍結させ、さらに先の処理を行なうまでこの温度で維持した。受領次第、試料を解凍し、便宜上、より少ないう量に再等分して再凍結させた。この試料は最終的に分析の直前に解凍した。したがって、ひとそろいの試料すべては、分析の前に2度凍結および解凍された。

40

### 【0197】

全部で751標本を採取、分析した。群は、生検で肺がんが確認された250患者、生検で良性の肺疾患であると確認された274患者、および明らかに正常である227の被

50

験者から構成される。がんおよび良性の患者は全て最終的な生検による診断で確認された。正常な被験者には、このような最終的な診断方法は行なわず、顕性悪性疾患が見られないことにより正常と判断された。最終的な診断方法の後、50歳以上の患者のみを選択した。この選択の後には、がん231、良性182、および正常155が残った。このがんの大コホート、良性肺疾患、および明らかに正常な被験者をまとめて以下「大コホート」と証する。良性肺疾患と肺がんとの区別に焦点を当てるために大コホートのサブセットを用いる。このコホートを以下「小コホート」と称し、これは肺がん138、良性106、および明らかに正常な被験者13からなる。大コホートから小コホートを除いた後には、肺がん107、良性74、および明らかに正常な被験者142が残った。このコホートは、小コホートに無関係で、発生した予測モデルを検証するために用いたもので、以下「検証コホート」と称する。記載のように調製した臨床試料は実施例2から10で用いた。

【実施例2】

【0198】

バイオマーカーの免疫測定検出

A. Abbott Laboratories (Abbott Park, IL, 以下「Abbott」) Architect (商標) 分析

以下の抗原、CEA、CA125、SCC、CA19-9、およびCA15-3のために Architect (商標) キットを入手した。分析は全てメーカーの説明書に従って行った。試料中の検体濃度は、Architect (商標) 計器によって得た。これらの濃度を用いて下記表1に示すAUCデータを作成した。

【0199】

【表2】

マーカー	大コホート		小コホート	
	観察番号	AUC	観察番号	AUC
Ca19-9	548	0.548	256	0.559
CEA	549	0.688	257	0.664
Ca15-3	549	0.604	257	0.569
Ca125	549	0.693	257	0.665
SCC	549	0.615	257	0.639

10

20

30

【0200】

B. Roche Elecsys (商標) 分析

Elecsys (商標) 2010 system (Roche Diagnostic GmbH, Mannheim, Germany) で、メーカーの説明書に従ってシフラー21-1 (Cyfra21-1) (サイトケラチン19、CK-19) 測定を行った。シフラー21-1 (Cyfra21-1) の濃度はElecsys (商標) 計器によって得た。このデータによってROC曲線を作成し、大コホートおよび小コホートについてのAUCを下記表2に記録した。

【0201】

【表3】

マーカー	大コホート		小コホート	
	観察番号	AUC	観察番号	AUC
CK-19	537	0.68	248	0.718

40

表2. サイトケラチン19の臨床性能(AUC)

【0202】

C. マイクロタイタープレート分析

以下のELISAキットを購入した。Advanced Life Science Institute, Inc. (Japan)のPROGRP、IDL Biotech AB (Bromma, Sweden)のTPS (サイトケラチン18、CK-18)、

50

およびIBL Immuno Biological Laboratories (Minneapolis, MN, USA) のパラインフルエンザ 1/2/3 IgG ELSA。メーカーの説明書に従って分析を行った。検体の濃度は、メーカーのプロトコールで指示、規定されている計算から導いた。個々の分析で得られたAUCを下記の表3に示す。

【0203】

【表4】

マーク	観察番号	大コホト		小コホト		10
		AUC	観察番号	AUC	観察番号	
CK-18	548	0.656	257	0.657		
ProGRP	548	0.698	257	0.533		
パラインフルエンザ 1/2/3	544	0.575	255	0.406		

表3. サイトケラチン18、proGRP、およびパラインフルエンザ1/2/3の臨床性能(AUC)

【実施例3】

【0204】

自己抗体ビーズアレイ

A. 市販のヒトタンパク質（下記の表4を参照されたい）を LumineX（商標）Seromap（商標）ビーズ（Austin, Texas）に付着させ、個々のビーズセットを合わせて試薬を調製した。試薬の一部を、存在する任意の抗体がタンパク質に結合できる条件下でヒト血清試料にさらした。結合していない物質を洗い落とし、次いでビーズを、ヒトIgGに特異的に結合する抗体にR-フィコエリトリンを結びつけた蛍光複合体にさらした。洗浄後、ビーズを、内部の色素に従ってそれぞれのビーズを同定する LumineX（商標）100計器に通して、ビーズに結合した蛍光発光（ビーズに結合した抗体の量に対応する）を測定した。このようにして、772の試料（肺がん251、正常244、良性277）について、21のヒトタンパク質、ならびに、対照としていくつかの非ヒトタンパク質（ウシ血清アルブミン（BSA）および破傷風毒素）に対する免疫応答を評価した。

【0205】

抗原MUC-1 (Fujiirebio Diagnostics INC, Malvern, PA)、サイトケラチン19 (Biodesign, Saco, ME)、およびCA-125 (Biodesign, Saco, ME)を、細胞培養物のイオン交換画分として入手した（下記の表4を参照されたい）。これらの比較的粗い調製物を、サイズ排除カラム (BioRad SEC-250, Hercules, CA)、移動相がPBS0.4mL/分のHPLCを用いた、分子量によるさらなる分別にかけた。15分から始めてそれぞれ1分ごとの画分を、それぞれの抗原に対して全部で23画分を集めた。MUC-1については250uLを注入し、サイトケラチン19およびCA-125については150uLを注入した。3つの試料全てが、15から24分に溶出する高分子量のタンパク質のさまざまな濃度を表す信号を示し、24分以降の時間においては測定できない程度高信号、低分子量の物質の濃度が高いことを示した。ビーズを被覆するために、以下の画分を合わせた。MUC-1-A画分6、7；MUC-1-B画分10、11；MUC-1-C画分12、13；サイトケラチン19-A画分4、5；サイトケラチン19-B画分8、9；サイトケラチン19-C画分16、17；CA125-A画分5、6；CA125-B画分12、13

【0206】

20

30

40

【表5】

ビーズID	抗原	供給源	
1	MUC-1-A	Fujirebio Diagnostics INC	
2	MUC-1-B	Fujirebio Diagnostics INC	
3	MUC-1-C	Fujirebio Diagnostics INC	
4	サイトケラチン19-A	Biodesign, Saco, ME	
5	サイトケラチン19-B	Biodesign, Saco, ME	
6	サイトケラチン19-C	Biodesign, Saco, ME	
7	CA125-A	Biodesign, Saco, ME	
8	CA125-B	Biodesign, Saco, ME	10
9	HSP27	US Biological, Swampscott, MA	
10	HSP70	Alexis, San Diego, CA	
11	HSP90	Alexis, San Diego, CA	
12	テタヌス	Sigma, St. Louis, MO	
13	HCG	Diosynth API, Des Plaines, IL	
14	VEGF	Biodesign, Saco, ME	
15	CEA	Biodesign, Saco, ME	
16	NY-ESO-1	NeoMarkers, Fremont, CA	
17	AFP	Cell Sciences, Canton, MA	
18	ERB-B2	Invitrogen, Grand Island, NY	20
19	PSA	Fitzgerald, Concord, MA	
20	P53	Lab Vision, Fremont, CA	
21	J0-1	Biodesign, Saco, ME	
22	ラクトフェリン	Sigma, St. Louis, MO	
23	HDJ1	Alexis, San Diego, CA	
24	ケラチン	Sigma, St. Louis, MO	
25	RECAF62	BioCurex, Vancouver, BC Canada	
26	RECAF50	BioCurex, Vancouver, BC Canada	
27	RECAFミク	BioCurex, Vancouver, BC Canada	
28	BSA	Sigma, St. Louis, MO	30

表4. タンパク質のリスト

【0207】

## B. 抗原の Lumines SeromAP (商標) ビーズへの被覆

Omega10K 限外ろ過プレート (Pall Corporation, Ann Arbor, Michigan) のウェルに水 50 uL を加えた。10 分後にプレートを真空状態に置いた。ウェルが空になった際には、水 10 uL を加えて水和状態を保った。それぞれのウェルに 5 mM のモルホリノエタンスルホン酸 (MES) (pH 5.6) を 50 から 100 uL、指示された Lumines (商標) SeromAP (商標) ビーズ、および 10 から 20 ug に相当する適切な量の表4に指示されたそれぞれの抗原を加えた。ビーズをピベットで懸濁した。ビーズに E D A C (2.0 mg / 5 mM の MES (pH 5.6) 1.0 mL) 10 uL を加えた。プレートに蓋をして暗所の振とう機に置いた。14 時間後にプレートを真空吸引し、水で洗浄し、最後にビーズを 20 mM のトリエタノールアミン (TEA) (pH 5.6) 50 uL 中に懸濁させた。プレートを暗所で振とう機で攪拌した。2 度目の E D A C (2.0 mg / 5 mM の MES (pH 5.6) 1.0 mL) をそれぞれのウェルに加え、プレートを暗所の振とう機に 1 時間置いた。洗浄後、1 % BSA および 0.08 % アジ化ナトリウム (PBN) を含有する P B S 緩衝液 200 uL をそれぞれのウェルに加え、続いてプローブで超音波処理して、暗所に置いた。

【0208】

## D. ビーズで覆われた血清試料の試験

血清試料を、以下のようにマイクロプレート中に調製した。PBN 中 1 : 20 希釀、各

マイクロプレートあたり 80 試料。上記のビーズセット 50  $\mu$ L にウサギ血清（本件で試験したものとは関係ない抗原の免疫を有するウサギ由来）5  $\mu$ L を加えた。ビーズセットをボルテックスして 37 下に置いた。35 分後に、5% ウサギ血清および 1% C H A P S (B R C) を含有する P B N 1 mL を加えた。ビーズセットをボルテックスし、遠沈させて、B R C 1.05 mL 中に再懸濁させた。Supor 1.2 u filter plate (Pall Corporation) のウェルを P B N 100  $\mu$ L で洗浄した。それぞれのウェルに B R C 50  $\mu$ L、それぞれの 1:20 の血清試料 10  $\mu$ L、および再懸濁させたビーズ 10  $\mu$ L を加えた。プレートを室温、暗所にて 1 時間振とうし、ろ過して、B R C 100  $\mu$ L で 10 分間、3 回洗浄を行った。検出複合体 (R P E 抗ヒト Ig G 200  $\mu$ L / B R C 5.0 mL) 50  $\mu$ L を加えてピペットでビーズを再懸濁させた後、プレートを暗所で 30 分間振とうした。次いで B R C 100  $\mu$ L を加え、ビーズをピペットで攪拌して、試料を Lumines (商標) 100 計器で分析した。

## 【0209】

結果（各試料および抗原に対するビーズの中間値強度）を、表 5 で以下に示した大および小コホートに対する以下の結果で、ROC 解析によって評価した。

## 【0210】

## 【表 6】

バイオマーカー	大コホート		小コホート	
	観察番号	AUC	観察番号	AUC
MUC-1-A	579	0.53	253	0.56
MUC-1-B	579	0.55	253	0.59
MUC-1-C	579	0.57	253	0.61
サイトケラチン 19-A	579	0.57	253	0.58
サイトケラチン 19-B	579	0.53	253	0.49
サイトケラチン 19-C	579	0.62	253	0.65
CA125-A	579	0.53	253	0.5
CA125-B	579	0.62	253	0.59
HSP27	579	0.56	253	0.56
HSP70	579	0.49	253	0.51
HSP90	579	0.54	253	0.53
テタヌス	579	0.57	253	0.56
HCG	579	0.54	253	0.5
VEGF	579	0.53	253	0.51
CEA	579	0.57	253	0.55
NY-ESO-1	579	0.58	253	0.58
AFP	579	0.51	253	0.55
ERB-B2	579	0.61	253	0.57
PSA	579	0.6	253	0.57
P53	579	0.6	253	0.54
J0-1	579	0.57	253	0.54
ラクトフェリン	579	0.49	253	0.49
HDJ1	579	0.62	253	0.63
ケラチン	579	0.58	253	0.55
RECAF62	579	0.54	253	0.53
RECAF50	579	0.53	253	0.53
RECAF ミク	579	0.54	253	0.62
BSA	579	0.57	253	0.59

表5. 大および小コホートでの、表4からのタック質を含む自己抗体ビーズアレイの臨床性能

## 【実施例 4】

## 【0211】

10

20

30

40

50

## 自己抗体スライドアレイ

## A. 抗原調製物

インビトロジェン (Invitrogen) の Ultimate ORF Collection (商標) (インビトロジェン、Grand Island, NY) から由来した、およそ 5000 タンパク質を、グルタチオン-S-トランスフェラーゼ (GST) 配列の全長ヒトタンパク質との組換え体融合物として調製した。GST タグによって、タンパク質の他の特徴に依存せずに、アレイに結合する各タンパク質の量の査定が可能になる。

## 【0212】

## B. スライドの抗原コーティング

ProtoArray は、以上で言及したおよそ 5000 タンパク質でスポットしたニトロセルロースでコートしたガラス表面 (スライド)、ならびに多数の対照特徴からなる。

## 【0213】

## C. コートスライドでの血清試料の試験

アレイをまず、PBS / 1% BSA / 0.1% Tween 20 で、4 にて 1 時間ブロックした。次いで、プロファイリング緩衝液 (Profiling Buffer) (本明細書で議論した「プロファイリング緩衝液」は、PBS、5 mM MgCl<sub>2</sub>、0.5 mM ジチオスレイトール、0.05% Triton X-100、5% グリセロール、1% BSA を含む) 中 1 : 120 で希釈した血清試料に、4 にて 90 分間暴露した。アレイを次いで、プロファイリング緩衝液にて 3 回、洗浄あたり 8 分間洗浄した。次いでアレイを、Alexa Fluor - 共役抗 - ヒト IgG、プロファイリング緩衝液中 0.5 ug / mL、44 にて 90 分間暴露した。次いでアレイを、プロファイリング緩衝液中で 3 回、1 洗浄あたり 8 分間洗浄した。遠心での乾燥後、Axon GenePix 4000B 蛍光マイクロアレイスキャナー (モレキュラー デバイゼス (Molecular Devices)、Sunnyvale, CA) を用いてスキャンした。

## 【0214】

## D. バイオマーカー選別

がん患者からの血清の陽性シグナルの分布を、正常患者からのものと比較することによって、がん患者の自己抗体特性を削除しているタンパク質の同定を決定した。制限された数のアレイでの、がん特異的自己抗体を発見する可能性を増加させるために、以下の試料のプールを使用した。それぞれ 4 から 5 人の肺がん患者の血清を含む 10 プール、それぞれ 4 または 5 人の正常患者の血清を含む 10 プール、それぞれ良性肺疾患有する 4 から 5 人の患者からの血清を含む 10 プール。これらのプールを、以上で記述したように、処理のためにインビトロジェンに送付した。各プローブに対する各タンパク質に相当する蛍光強度を、スプレッドシートにて表した。各タンパク質は 2 回表され、アレイ上の二重スポットに相当する。

## 【0215】

特定のタンパク質に対する免疫応答のがん特異性の査定に対する 1 つのアルゴリズムにおいて、カットオフ値が、がん試料のシグナル強度を非がん試料のものから最もよく区別した、製造業者 (インビトロジェン) によって供給された。本カットオフより高い強度を有する各群からの試料の数 (それぞれがんカウントおよび非がんカウント) を決定し、パラメーターとしてスプレッドシートに配置した。さらに、p - 値を計算し、これは、他と比較した 1 つの群において、シグナル増加がない可能性を表している。ついでデータを、非がん群において最も低い陽性、またがん群において最も陽性であるトップタンパク質を提供するようにソートし、さらに p - 値にて低から高でソートした。本式によるソートが、以下の表 7 に提供された以下の情報を提供した。

## 【0216】

10

20

30

40

【表7】

## 抗原検証

	がん カウト	非 んカウン ト	P-値
先体小胞タンパク質 1 (ACRV1)	6	0	0.0021
フォックヘッドボックス A3 (FOXA3)	6	0	0.0072
一般転写因子 IIA	6	0	0.5539
E3 キンシタンパク質リガーゼ 2 を含む WW ドメイン	5	0	0.0018
PDZ ドメイン含有 1 (PDZK1)	5	0	0.0018
サイクリン E2	5	0	0.0018
サイクリン E2	5	0	0.0018
2型ホスファチジル酸ホスファターゼドメイン含有 3 (PPAPDC3)	5	0	0.0088
アンキリン繰り返しおよび無菌アルファモチードメイン含有 3	5	0	0.0563
亜鉛フィンガー	5	0	0.0563
システィニル-tRNA シンセターゼ	4	0	0.0077
システィニル-tRNA シンセターゼ	4	0	0.0077
TGHMGH エンハンサー 3 に結合する転写因子 (TFE3)	4	0	0.0077
E3 キンシタンパク質リガーゼ 2 を含む WW ドメイン	4	0	0.0077
クロモソーム 21 オープンリーディングフレーム 7	4	0	0.0077
クロモソーム 21 オープンリーディングフレーム 7	4	0	0.0077
F1 を含む IQ モチード (IQCF1)	4	0	0.0077
リソバ球細胞質たんぱく質 1 (L-プラスチン) (LCP1)	4	0	0.0077
先体小胞タンパク質 1 (ACRV1)	4	0	0.0077
Dnaj (Hsp40) ホモログ	4	0	0.0077
Dnaj (Hsp40) ホモログ	4	0	0.0077
核レセプター結合因子 2	4	0	0.0077
核レセプター結合因子 2	4	0	0.0077
PDZ ドメイン含有 1 (PDZK1)	4	0	0.0077
ニューロン 2 中のタンパク質オーゼ C およびカゼインキナーゼ 基質	4	0	0.0077
LIM ドメインキナーゼ 2	4	0	0.0077
ポリメラーゼ (RNA) III (DNA 指向) ポリペプチド D	4	0	0.0077
RNA 結合モチーフタンパク質	4	0	0.0077
細胞分裂周期関連 4 (CDCA4)	4	0	0.0312
Rho G アンスクリオボンド 交換因子 (GEF) 1	4	0	0.076
Luc7 様 2 (S. セレビシエ (S. cerevisiae))	4	0	0.2302
RIKEN cDNA 2310008M10 に対する類似 (LOC202459)	4	0	0.2302
リブロース-5-リン酸-3-エピメラーゼ	3	0	0.0296
リブロース-5-リン酸-3-エピメラーゼ	3	0	0.0296
ヘム結合タンパク質 1 (HEBP1)	3	0	0.0296
ヘム結合タンパク質 1 (HEBP1)	3	0	0.0296
キラー細胞レクチン様レセプターアブファミリー C	3	0	0.0296
キラー細胞レクチン様レセプターアブファミリー C	3	0	0.0296
LATS	3	0	0.0296
N-アシルスフォンゴシンアミドヒドロラーゼ (酸クレアミダーゼ)	3	0	0.0296

セ) 1 (ASAHI)	3	0
N-アシルスフォンゴシンアミドヒドロラーゼ (酸クレアミダーゼ) 1 (ASAHI)	0	0.0296
パラメミン	3	0
パラメミン	3	0
PIN2-相互作用タンパク質 1	3	0
リボソームタンパク質 S6 キャセ	3	0
リボソームタンパク質 S6 キャセ	3	0
SH3 および PX ドメイン含有 3 (SH3PC3)	3	0
SHIGHM3 および PX ドメイン含有 3 (SH3PC3)	3	0
(小児白血病における) TCF3 (E2A) 融合パートナー (TFPT)	3	0
(小児白血病における) TCF3 (E2A) 融合パートナー (TFPT)	3	0
IGHM エンハンサー 3 に結合した転写因子 (TFE3)	3	0
クロモソーム 1 オープンリーティングフレーム 117	3	0
クロモソーム 1 オープンリーティングフレーム 117	3	0
シスプロテイン耐性関連過剰発現タンパク質	3	0
hsp70-相互作用タンパク質	3	0
仮説タンパク質 FLJ22795	3	0
仮説タンパク質 FLJ22795	3	0
インターフェロン誘導膜透過タンパク質 1 (9-27)	3	0
インターフェロン誘導膜透過タンパク質 1 (9-27)	3	0
F1 含有 IQ モチーフ (IQCF1)	3	0
ロイシン-リッチ繰り返しあおよび IQ モチーフ含有 2 (LRRIQ2)	3	0
ロイシン-リッチ繰り返しあおよび IQ モチーフ含有 2 (LRRIQ2)	3	0
パラメミン 2	3	0
パラメミン 2	3	0
RWD ドメイン含有 1	3	0
溶質担体ファミリー 7	3	0
溶質担体ファミリー 7	3	0
トキオミオシン 1 (アルファ)	3	0
トキオミオシン 1 (アルファ)	3	0
腫瘍抑制並伝達性候補 4	3	0
エビキチン様 4A	3	0
痕跡様 4 (ドロソフィラ (Drosophila)) (VGLL4)	3	0
WD 繰り返しドメイン 16	3	0
WD 繰り返しドメイン 16	3	0
マイトージェン-活性化タンパク質キャセ-活性化タンパク質キャセ 3	3	0
マイトージェン-活性化タンパク質キャセ-活性化タンパク質キャセ 3	3	0
死-関連タンパク質キャセ 1 (DAPK1)	3	0
ジメチルアルギニンジメチルアミノヒドロラーゼ 2 (DDAH2)	3	0
ジメチルアルギニンジメチルアミノヒドロラーゼ 2 (DDAH2)	3	0
熱ショック 70kDa タンパク質 2	3	0

黒色腫抗原ファミリー H  
 マトイジエン-活性化タンパク質オーバー-活性化タンパク質オーバー  
 ゼ 3 (MAPKAPK3)  
 ネイル様 2 (大腸菌 (E. coli)) (NEIL2)  
 ニューロン 2 中のタンパク質オーバー C およびオーバー インオーバー 基質  
 SMAD  
 SMAD  
 TIA1 細胞毒性顆粒関連 RNA 結合タンパク質  
 トレホイル因子 2 (鎮瘻性タンパク質 1) (TFF2)  
 ウロポルフィリノーゲン III シンターゼ (先天性造血性ポルフィリノーゲン  
 症) (UROS)  
 サイカイン誘導タンパク質 29kDa (CIP29)  
 膜透過性タンパク質 106C (TMEM106C)  
 クロモソーム 9 オープンリーディングフレーム 11  
 0-6-メチルグアニン-DNA メチルトランスフェラーゼ (MGMT)  
 PDGFA 関連タンパク質 1 (PDAP1)  
 PDGFA 関連タンパク質 1 (PDAP1)  
 ポリメラーゼ (RNA) III (DNA 指向) ポリペプチド D  
 Rho-関連  
 Rho-関連  
 RNA 結合モチーフタンパク質  
 テトラスパニン 17

3	0	0.0296
3	0	0.0296
3	0	0.0296
3	0	0.0296
3	0	0.0296
3	0	0.0296
3	0	0.0296
3	0	0.0296
3	0	0.0296
3	0	0.0296
3	0	0.0296
3	0	0.0296
3	0	0.0296
3	0	0.0296
3	0	0.0296
3	0	0.0296
3	0	0.0296
3	0	0.0296
3	0	0.0296
3	0	0.0296
3	0	0.0296
3	0	0.0296
3	0	0.0296
3	0	0.0296

10

20

30

表 7 抗原 ID リスト

## 【 0 2 1 7 】

第二のアルゴリズムが、非がん試料のシグナル強度の標準偏差で割った、がんに対する平均シグナルと、非がん試料に対する平均シグナル間の差として、タンパク質に対する免疫応答のがん特異性を計算した。これは、強い免疫応答が、弱いものよりも結果に影響を与えるという長所を有する。次いでデータを、最も高い値のトップタンパク質を与えるようにソートする。本ソートによって同定したトップ 100 リスティングを、以下表 8 にて示している。

## 【 0 2 1 8 】

【表8】

抗原検証	平均差/SD(非がん)
(小児白血病中) TCF3 (E2A) 融合パートナー (TFPT)	21.4
エピキシン特異的プロテアーゼ 45 (USP45)	16.1
エピキシン特異的プロテアーゼ 45 (USP45)	15.6
エピキシン-共役酵素 E20	15.1
(小児白血病中) TCF3 (E2A) 融合パートナー (TFPT)	13.9
エピキシン-共役酵素 E20	12.3
プロリン-リッヂコイルド-コイル 1 (PRRC1)	11.5
プロリン-リッヂコイルド-コイル 1 (PRRC1)	10
B-細胞 CLL/リンパ腫 10	9.8
溶質担体ファミリー 7	8.8
B-細胞 CLL/リンパ腫 10	8.7
DnaJ (Hsp40) ホモログ	8.2
DnaJ (Hsp40) ホモログ	8
溶質担体ファミリー 7	7.9
痕跡様 4 (ドロソフィラ (Drosophila)) (VGLL4)	6.5
SH3 および PX ドメイン含有 3 (SH3PX3)	6.3
サイクリン E2	6.1
SH3 および PX ドメイン含有 3 (SH3PC3)	6.1
サイクリン E2	6
cDNA クローン IMAGE:3941306	5.9
パラレムシン	5.8
インターフェロン誘導膜貫通タンパク質 1 (9~27)	5.6
パラレムシン	5.4
リブロース-5-リン酸-3-エピメラーゼ	5.4
ロイシン-リッヂ繰り返しおよび IQ モチーフ含有 2 (LRRIQ2)	5.3
リブロース-5-リン酸-3-エピメラーゼ	5.3
細胞分裂周期関連 4 (CDCA4)	5.2
インターフェロン誘導膜貫通タンパク質 1 (9~27)	4.8
ロイシン-リッヂ繰り返しおよび IQ モチーフ含有 2 (LRRIQ2)	4.7
マトイジエン-活性化タンパク質キナーゼ-活性化タンパク質キナーゼ 3	4.5
カルシウム/カルモジュリン-依存タンパク質キナーゼ I (CAMK1)	4.4
RAB3A 相互作用タンパク質 (テビン 3) 様 1 (RAB3IL1)	4.3
ジメチルアルギニンジメチルアミノヒドロラーゼ 2 (DDAH2)	4.2
hsp70-相互作用タンパク質	4.1
クロモソーム 9 オープンリード イング フレーム 11	4.1
マトイジエン-活性化タンパク質キナーゼ-活性化タンパク質キナーゼ 3	4.1
先体小胞タンパク質 1 (ACRV1)	4.1
トリオセホスフェート イソメラーゼ 1	4
トリオセホスフェート イソメラーゼ 1	3.8
ウロカルフィリノーケン III シンターゼ (先天性造血性カルフィリン症) (UROS)	3.7
キラー細胞レクチン様レセプターサブファミリー C	3.7

エストロゲン-関連レセプター-アルファ(ESRRA)	3.6	
先体小胞タンパク質 1(ACRV1)	3.6	
細胞分裂周期関連 4(CDCA4)	3.6	
RAB3A 相互作用タンパク質(ラビン 3)様 1(RAB3IL1)	3.5	
死-関連タンパク質キナーゼ 1(DAPK1)	3.5	
ニューロン 2 中のタンパク質キナーゼ C およびカゼイノキナーゼ基質	3.5	
トロボ モジュリン 1	3.4	
トロボ モジュリン 1	3.4	10
クロモーム 1 オープンリーディングフレーム 117	3.4	
ジメチルアキニンジメチルアミノヒドロラーゼ 2(DDAH2)	3.4	
エストロゲン-関連レセプター-アルファ(ESRRA)	3.2	
ブレクストリン相同ドメイン含有	3.1	
ウロポルフィノーゲン III シンターゼ(先天性造血性ポルフィrin症)(UROS)	3.1	
仮想タンパク質 FLJ22795	3.1	
SRC に関連した FYN オンコジー	3.1	
マイトージェン活性化タンパク質キナーゼ-活性化タンパク質キナーゼ 3(MAPKAPK3)	3.1	
CDC37 細胞分割周期 37 ホモログ(S. セルビッシュ(S. cerevisiae))様 1	3	20
腫瘍抑制亜伝達性候補 4	3	
RWD ドメイン含有 1	3	
仮想タンパク質 FLJ22795	3	
CDC37 細胞分割周期 37 ホモログ(S. セルビッシュ(S. cerevisiae))様 1	2.9	
E3 キンシタンパク質リガーゼ 2 を含む WW ドメイン	2.9	
PDZ ドメイン含有 1(PDZK1)	2.9	
マイトージェン活性化タンパク質キナーゼ-活性化タンパク質キナーゼ 3(MAPKAPK3)	2.9	
IGHM エンハンサー 3 に結合している転写因子(TFE3)	2.9	30
フォークヘッドボックス A3(FOXA3)	2.8	
クロモーム 1 オープンリーディングフレーム 117	2.8	
アンキリン繰り返しおよび無菌アルファモチードメイン含有 3	2.8	
OCIA ドメイン含有 1(OCIAD1)	2.8	
ポリメラーゼ(DNA 指向)	2.8	
SMAD	2.8	
KIAA0157(KIAA0157)	2.8	
B-細胞 CLL/リンパ腫 7C(BCL7C)	2.8	
リボソームタンパク質 S6 キナーゼ	2.8	40
クロモーム 9 オープンリーディングフレーム 11	2.7	
リボソームタンパク質 S6 キナーゼ	2.7	
サイトカイン誘導タンパク質 29kDa(CIP29)	2.7	
核レセプター結合因子 2	2.7	
宿主細胞因子 C1 調節物 1(XP01 依存)(HCFC1R1)	2.7	

STE20 様キーゼ (酵母) (SLK)	2.7	
OCIA ドメイン含有 1 (OCIAD1)	2.6	
ニューロン 2 中のタンパク質キーゼ C およびカゼインキーゼ基質	2.6	
クエイキング ホモログ	2.6	
ソーティング キシン 16 (SNX16)	2.6	
リンパ球細胞質タンパク質 1 (L-プロテイン) (LCP1)	2.6	
クロモソーム 21 オープンリーディングフレーム 7	2.5	
STE20 様キーゼ (酵母) (SLK)	2.5	10
宿主因子 C1 調節物 1 (XP01 依存) (HCFC1R1)	2.5	
hsp70-相互作用タンパク質	2.5	
クエイキング ホモログ	2.5	
IGHM エンハンサー 3 に結合している転写因子 (TFE3)	2.5	
SMAD	2.4	
E3 ユビキチンタンパク質リガーゼ 2 を含む WW ドメイン	2.4	
クロモソーム 21 オープンリーディングフレーム 7	2.4	
PDZ ドメイン含有 1 (PDZK1)	2.4	
アセチルセロトニン O-メチルトランスフェラーゼ 様	2.4	
B-細胞 CLL/リンパ腫 7C (BCL7C)	2.3	
リボソームタンパク質 S19 (RPS19)	2.3	20
O-6-メチルグアニン-DNA メチルトランスフェラーゼ (MGMT)	2.3	

表 8 最も高い S/N 比を持つトップタンパク質を紹介するためにソートされた抗原 ID リスト。S/N は、非がん群の標準偏差 (SD 非がん) によって、2 つの群の平均シグナル強度の差 (がん平均 - 非がん平均) を割ることによって計算した。

#### 【 0 2 1 9 】

表 7 および 8 のソート結果を比較し、各タンパク質に対して、がんおよび非がん試料によって産出されたシグナルを試験することによって、以下表 9 で示した以下の 25 個のタンパク質を、さらなる調査のために選別した。

#### 【 0 2 2 0 】

【表9】

クローン	抗原検証
BC007015.1	サイクリンE2
NM_002614.2	PDZドメイン含有1(PDZK1)
NM_001612.3	先体小胞タンパク質1(ACRV1)
NM_006145.1	DnaJ(Hsp49)ホモログ
BC011707.1	核レセプター結合因子2
BC008567.1	クロモソーム21オープンソリーティングフレーム7
BC000108.1	E3ユビキチンタンパク質カーセ2を含むWWドメイン
BC001662.1	マイトochondria活性化タンパク質カーセ3-活性化タンパク質カーセ3
BC008037.2	ニューロン2中のタンパク質カーセCおよびカーセイキカーセ基質
NM_005900.1	SMAD
NM_013974.1	ジメチルアルギニンジメチルアミノヒドロラーゼ2(DDAH2)
NM_000375.1	ウロポルフィリノゲンIIIシターゼ(先天性造血性ポルフィリン症)(UROS)
NM_145701.1	細胞分割周期関連4(CDCA4)
BC016848.1	クロモソーム1オープンソリーティングフレーム117
BC014307.1	クロモソーム9オープンソリーティングフレーム11
BC000897.1	インターフェロン誘導膜貫通タンパク質1(9~27)
NM_024548.2	ロイソーリッチ繰り返しおよびIQモチーフ含有2(LRRIQ2)
BC013778.1	溶質担体ファミリー7
BC032449.1	パラレムシン
NM_153271.1	SH3およびPXドメイン含有3(SH3PX3)
NM_013342.1	(小児白血病中)TCF3(E2A)融合パートナー(TFPT)
NM_006521.3	IGHMエンハンサー3に結合している転写因子(TFE3)
BC016764.1	リブロース-5-リン酸-3-エピメラーゼ
BC014133.1	CDC37細胞分割周期37ホモログ(S.セレビシエ(S. cerevisiae))様1
BC053545.1	トキミオシン1(アルファ)

表9 さらなる調査のために選別したトップ25タンパク質

## 【0221】

30

## E. サイクリンE2

サイクリンE2の2つの形態が、ProtoArray(商標)上で存在していることが分かった。Genebankアクセション番号BC007015.1(配列番号1)として同定された形態が、がん試料のプールのいくつかと、強力な免疫反応性を示し、良性および正常(非がん)プールと非常に低い応答性を示した。反対に、Genebankアクセション番号BC020729.1(配列番号2)として同定された形態は、任意のがんまたは非がんプール試料とほとんど反応性を示さなかった。以下に示したように、2つの形態の配列アライメントが、259アミノ酸にわたる同一性を示しており、N-末端およびC-末端領域両方で差がある。BC020729.1は、N-末端に110アミノ酸、C-末端に7アミノ酸を持ち、BC007015.1には存在しない。BC007015.1は、C-末端に、BC020729.1には存在しない37アミノ酸を有する。形態BC007015.1のみが免疫応答性を示しているので、この免疫応答性は、C-末端の37アミノ酸部分に起因する。

40

## 【0222】

BC007015.1のC-末端からの2つのペプチドを合成し、E2-1(配列番号3)は、BC007015.1のC-末端37アミノ酸を含む。E2-2(配列番号5)は、BC007015.1のC-末端18アミノ酸を含む。両方のペプチドを、N末端にてシステインを含むように合成し、担体タンパク質または表面への特異的共有結合のための反応性の部位を提供する。

## 【0223】

50

## 【化1】

BC007015.1 1 M  
 BC020729.1 1 MSRRSSRLQAKQQPQPSQTESPQEAQIIQAKKRKTTQDVKKRREEVTKKHQYEIRNCWPP  
 \*

BC007015.1  
 BC020729.1 61 VLSGGISPCLIIETPHKEIGTSDFSRFTNYRFKNLFINPSPLPDLSWGC

BC007015.1 2 SKEVWLNLKESRYVHDKHFVLSDEPQMRSLLLDWLLEVCEVYTLHRETFYLAQDF  
 BC020729.1 110 SKEVWLNLKESRYVHDKHFVLSDEPQMRSLLLDWLLEVCEVYTLHRETFYLAQDF  
 \*\*\*\*

BC007015.1 62 FDRFMLTQKDINKNMLQLIGITSLFIAKLEIYAPKLQEFAYVTDGACSEEDILRMELI  
 BC020729.1 170 FDRFMLTQKDINKNMLQLIGITSLFIAKLEIYAPKLQEFAYVTDGACSEEDILRMELI  
 \*\*\*\*

10

BC007015.1 122 ILKALKWELCPVTIISWLNLFLQVDALKDAPKVLLPQYSQETFIQIAQLLDLCILAIDSL  
 BC020729.1 230 ILKALKWELCPVTIISWLNLFLQVDALKDAPKVLLPQYSQETFIQIAQLLDLCILAIDSL  
 \*\*\*\*

BC007015.1 182 EFQYRILTAALCHFTSIEVVKKASGLEWDSISECVDWMVPFVNVKSTSPVKLKTFFKKI  
 BC020729.1 290 EFQYRILTAALCHFTSIEVVKKASGLEWDSISECVDWMVPFVNVKSTSPVKLKTFFKKI  
 \*\*\*\*

BC007015.1 242 PMEDRHNIQTHTNYLAMLEEVNYINTFRKGQLSPVCNGGIMTPPKSTEKPPGKH  
 BC020729.1 350 PMEDRHNIQTHTNYLAMLCMISSHV  
 \*\*\*\*

B C 0 0 7 0 1 5 . 1 (配列番号1) および B C 0 2 0 7 2 9 . 1 (配列番号2) の配  
 列アライメント

20

## 【0224】

## 【化2】

E2-1 : CEEVNYINTFRKGQLSPVCNGGIMTPPKSTEKPPGKH (配列番号3)

E2-2 : CNGGIMTPPKSTEKPPGKH (配列番号5)

## 【0225】

ペプチドE2-1およびE2-2は、BSAをマレイミドで活性化し、続いてペプチドの共役によって、BSAにそれぞれ連結した。活性化BSAを、以下のプロトコールに従って調製した。200uL PBS中のBSA8.0mgに、20uL DMRFおよび10uL 1M トリエタノールアミン pH 8.4中のGMB (N-(ガムマ-マレイミド-ブチリル-オキシ)スクシンイミド、ピアス (Pierce)、Rockford IL) 1mgを加えた。60分後、混合液を、PBS緩衝液でSephadex G50カラムを通して、400uL画分を回収した。活性化BSA-Mal (100uL) に、ペプチドE2-1、2.5mgまたはペプチドE2-2、3.2mgを加えた。両方の場合で、混合液をボルテックスして、氷上に15分間おき、その後混合液を、25分間室温まで動かした。結合産物、BSA-Mal-E2-1 (BM-E2-1) およびBSA-Mal-E2-2 (BM-E2-2)を、浄化のために、Sephadex G50カラムを通した。

30

## 【0226】

タンパク質およびペプチドを、2つの方法を用いて、Luminex (商標)マイクロスフィアに結合させた。第一の方法が、実施例10Cにて記述されており、「直接法 (direct methods)」と呼ばれる。第二の方法は、「前活性化法 (pre-activate method)」と呼ばれ、以下のプロトコールを使用する。オメガ (Omega) 10k超遠心プレートのウェル中に、水100uLを加え、10分後に真空状態に置いた。ウェルが空になったら、20uL MES (100mM) pH 5.6および50uL各Luminex (商標) Seromap (商標) ビーズセットを、以下表10にて示したように加えた。カラム1列A、B、CおよびD中のウェルに、およびカラム2列A、B、C、DおよびE中のウェルに、MES中のNHS (20mg/mL) 10uL、およびMES中EDAC (10mg/mL) 10uLを加えた。暗所での振盪45分

40

50

後に、プレートを真空状態に置き、緩衝液と未反応試薬を吸い上げた。プレートをバキュームから取り除き、20 uL MES および 50 uL 水を加えた。表10にて示唆したプレートに、各タンパク質またはペプチド(2 uLを加えた、DNAJB1以外)4 uLを加え、ピペットで攪拌して、ビーズを分散させた。プレートを30分間、シェーカー上で攪拌し、ついでMES中10 mg/mL EDAC 5 uLを(直接結合のために)カラム1、列EFGHに加え、プレートをシェーカー上で30分間攪拌し、ついで真空状態に置いて、緩衝液および未反応試薬を除去した。ウェルが空になったら、50 uL PBSを加え、混合液を攪拌し、プレートを真空状態に置いた。ウェルが空になったら、50 uL PBSを加え、ピペットで混合液を攪拌して、ビーズを分散させ、シェーカー上で60分間インキュベートした。反応を停止するために、200 uL PBNを加えて、反応液を超音波処理した。 10

### 【0227】

以下の表10は、異なるビーズセット上での、サイクリンE2ペプチドおよびタンパク質の異なる提示を要約している。ペプチドE2-1およびE2-2をBSAに結合し、ついで前活性化法(ビーズID 25および26)、または直接法(ビーズID 30および31)を用いてビーズに結合させた。ペプチド、E2-1およびE2-2をまた、前活性化法(ビーズID 28および29)または直接法(ビーズID 33および34)を用いて、BSAなしにビーズに結合させた。ビーズ35、37、38、39および40を、前活性化法を用いてタンパク質でコートした。 20

### 【0228】

#### 【表10】

カラム	列	ビーズID	抗原	供給源	結合方法
1	A	25	BM-E2-1	3.9mg/mL	前活性化
1	B	26	BM-E2-2	2.4mg/mL	前活性化
1	C	28	E2-1	21mg/mL	前活性化
1	D	29	E2-2	40mg/mL	前活性化
1	E	30	BM-E2-1	3.9mg/mL	直接
1	F	31	BM-E2-2	2.4mg/mL	直接
1	G	33	E2-1	21mg/mL	直接
1	H	34	E2-2	40mg/mL	直接
2	A	35	CCNE2	(ケンウェイ(GenWay)、San Diego、CA)0.6mg/mL	前活性化
2	B	37	MAPKAPK3	(ケンウェイ(GenWay)、San Diego、CA)0.5mg/mL	前活性化
2	C	38	p53	(バイオモル(Biomol)、Plymouth Meeting、PA)0.25mg/mL	前活性化
2	D	39	TMOD1	(ケンウェイ(GenWay)、San Diego、CA)0.8mg/mL	前活性化
2	E	40	DNAJB1	(アキソラ(Axxora)、San Diego、CA)1mg/mL	前活性化

表10 異なるビーズ上のサイクリンE2ペプチドおよびタンパク質の異なる提示の要約

### 【0229】

ビーズを、以下の様式にて、患者血清で処理した。1mL PBNに、各ビーズ調製物5 uLを加えた。混合液を超音波処理して遠心し、ペレットビーズを、PBS中のBSA 1% 1mLで洗浄し、同一の緩衝液1mL中に再懸濁させた。1.2u Suporフィルタープレート(パル コーポレーション(Pal1 Corporation)、East Hills, NY)に、100 uL PBS/Tween(0.2% Tween 20を含むPBS中1% BSA)を加えた。10分後、プレートをろ過し、50 uL PBS 0.2% Tween(0.2% Tween 20を含むPBS中1% BSA)を加えた。ろ過液を、100 uL PBS 0.2% Tween(0.2% Tween 20を含むPBS中1% BSA)に加え、100 uL PBS 0.2% Tween(0.2% Tween 20を含むPBS中1% BSA)を加えた。 30

S A )を加えた。各ウェルに、表 1 1 で示したように、20 u L ビーズ混合物および20 u L 血清(1:50)を加えた。血清は、ヒト患者血清またはウサギ抗-GST血清のいずれかであった。プレートを暗所にてシェーカー上に配置した。1時間後、プレートをろ過し、100 u L P B N / T w e e n にて三回洗浄した。R P E - 抗ヒト - I g G (1:400)(シグマ(Sigma)、St. Louis, MO) 50 u L を加えて、ヒト抗体を検出し、R P E - 抗ウサギ - I g G (1:200)を加えて、ウサギ抗-GST抗体を検出した。プレートを暗所にて30分間、シェーカー上に配置して、その後ビーズをろ過し、洗浄して、Luminex(商標)上で走らせた。

#### 【0230】

6つの血清試料とウサギ抗-GSTの結果を、以下表 1 1 で示している。

10

#### 【0231】

#### 【表 1 1】

		ビーズ ID								
		前活性化				直接				
血清 ID		BM-E2-1	BM-E2-2	E2-1	E2-2	CCNE2	BM-E2-1	BM-E2-2	E2-1	E2-2
A2	18	12	7	4	17	16	13	9	5	
A4	4	4	3	3	4	2	5	4	3	
B2	9	16	5	4	12	8	10	9	5	
B4	4380	172	1985	11	358	4833	132	2298	18	
C4	227	44	66	9	50	243	40	87	7	
D4	406	15	64	7	19	440	13	107	8	
F4	3721	156	1592	8	299	4034	140	1997	19	
ウサギ抗GST	13	14	40	21	1358	10	13	56	22	

表 11 患者血清に暴露した、サイクリン E2 ペプチドおよびタンパク質でコートしたビーズに対するルミネクス結果

20

#### 【0232】

B S A に連結し、それぞれビーズに(直接法を用いて)直接、または前活性化(または前活性化法)を介して結合したペプチド E 2 - 1 を含む、ビーズ 25 および 30 が、最も強力なシグナルを与えたことが、以上表 1 1 より明らかである。約半分だけが B S A 担体を介したが、B S A 担体なしで結合したペプチド E 2 - 1 がまた、強力なシグナルを与えた。ペプチド E 2 - 2 は、B S A 担体を介して結合した場合に、非常に低いシグナルを与え、B S A 担体なしではほぼ検出不能なシグナルであった。(N-末端GST融合タグを含む)全長タンパク質 C C N E 2 は、ペプチド E 2 - 2 の任意の形態のもの以上の、しかしペプチド E 2 - 1 のものより非常に低いシグナルを示し、これは、配列の免疫応答性部分を含むが、ビーズ上より低い密度であることを示唆している。ウサギ抗-GSTでのこのシグナルは、本GST融合タンパク質が、ミクロスフィアに対して成功裏に結合したことを見ている。

30

#### 【0233】

以下表 1 2 で示されたタンパク質を、以上で記述したような前活性化および直接法によって、および受動的コーティングによって、Luminex Seromap(商標)上にコートした。受動的コーティングのために、供給メーカーから受け取ったままの溶液中、タンパク質 5 u g を、Seromap(商標)ビーズ 200 u L に加え、混合液をボルテックスし、室温にて 5 時間、次いで 4 ℃ にて 18 時間インキュベートし、次いで沈殿するまで遠心し、ペレットを洗浄して、P B N 中に再懸濁させた。

40

#### 【0234】

50

【表12】

コード	タンパク質	ビーズ	供給源	
前活性化	TMP21-ECD	1	アボット(Abbott, North Chicago, IL)	
前活性化	NPC1L1C-ドメイン	5	アボット(Abbott, North Chicago, IL)	
前活性化	PSEN2(1-86aa)	14	アボット(Abbott, North Chicago, IL)	
前活性化	IgG ヒト	22	アボット(Abbott, North Chicago, IL)	
前活性化	BM-E2-2	26	アボット(Abbott, North Chicago, IL)	
直接	BM-E2-1	30	アボット(Abbott, North Chicago, IL)	
前活性化	TMOD1	39	ゲンウェイ(GenWay, San Diego, CA)	10
前活性化	DNAJB1	40	アキソラ(Axxora, San Diego, CA)	
前活性化	PSMA4	41	アブノバ(Abnova, Taipei City, Taiwan)	
前活性化	RPE	42	アブノバ(Abnova, Taipei City, Taiwan)	
前活性化	CCNE2	43	アブノバ(Abnova, Taipei City, Taiwan)	
前活性化	PDZK1	46	アブノバ(Abnova, Taipei City, Taiwan)	
直接	CCNE2	49	ゲンウェイ(GenWay, San Diego, CA)	
前活性化	パキシリン	53	ハイオレジエント(BioLegend, San Diego, CA)	20
	AMPHIPHYSIN	54	ラボビジョン(LabVision, Fremont, CA)	
	CAMK1	55	アップステート(Upstate, Charlottesville, VA)	
	DNAJB11	67	アブノバ(Abnova, Taipei City, Taiwan)	
	RGS1	68	アブノバ(Abnova, Taipei City, Taiwan)	
	PAC SIN1	70	アブノバ(Abnova, Taipei City, Taiwan)	
	SMAD1	71	アブノバ(Abnova, Taipei City, Taiwan)	30
	p53	72	ハイオモル(Biomol, Plymouth Meeting, PA)	
	RCV1	75	ゲンウェイ(GenWay, San Diego, CA)	
	MAPKAPK3	79	ゲンウェイ(GenWay, San Diego, CA)	

表12 前活性化および直接方法による、ルミネクス(Luminex)SeroMap(商標)上へコートしたタンパク質

【0235】

234人の患者(87人がん、70人良性、および77人正常)からの血清試料を試験した。本試験からの結果を、ROC曲線によって解析した。各抗原に対する計算されたAUCを、以下表13で示している。

【0236】

【表13】

タンパク質	AUC
サイクリンE2ペプチド1	0.81
サイクリンE2ペプチド(ゲンウェイ(Genway))	0.74
サイクリンE2ペプチド2	0.71
TMP21-ECD	0.66
NPC1L1C-トドメイン	0.65
PAC5IN1	0.65
p53	0.63
マイトージエン活性化タンパク質キーゼ 活性化タンパク質キーゼ(MAPKAPK3)	0.62
トロボモジュリン1(TM0D1)	0.61
PSEN2(1-86aa)	0.60
DNA J結合タンパク質1(DNAJB1)	0.60
DNA J結合タンパク質11(DNAJB11)	0.58
RCV1	0.58
(カルシウム/カルモジュリン-依存タンパク質キーゼ1.CAMK1)	0.57
SMAD1	0.57
AMPHIPHYSIN ラボ ビジョン(Lab Vision)	0.55
RGS1	0.55
PSMA4	0.51
リブロース-5-リン酸-3-エピメラーゼ(RPE)	0.51
パキシリン	0.51
サイクリンE2タンパク質(アブノバ(Abnova))	0.49
PDZドメイン含有タンパク質1(PDZK1)	0.47

表13 血清試料から由来した抗原に対する計算AUC

## 【実施例5】

## 【0237】

## 質量分析法

## A. 混合磁気ビーズ( M M B )の連続希釈による試料調製

血清試料を融解し、当量のインビトロジエンゾルB緩衝液と混合した。この混合液をボルテクスし、0.8  $\mu$ mフィルター(サルトリウス(Sartorius)、Goettingen, Germany)を通してろ過して、さらなる処理の前に残余物を精製し取り除いた。自動化試料調製を、Dynal(登録商標)(インビトロジエン)強アニオン交換およびアボットラボラトリーズ(ABBOTT Laboratories(アボット、ABBOTT Park, IL)弱力チオノン交換磁気ビーズの混合物を用いて、96-ウェルプレートKing Fisher(登録商標)(サーモフィッシャー サイエンティフィック(Thermo Fisher Scientific Inc.)、Waltham, MA)上で実施した。典型的に、アニオノン交換ビーズは、アミンに基づく炭化水素、四級アミンまたはジエチルアミン基を、官能末端基として持ち、弱力チオノン交換ビーズは典型的に、スルホン酸(カルボキシル酸)に基づく官能基を有する。アボットのカチオノン交換ビーズ(CXビーズ)は、濃度2.5%(質量/容量)であり、Dynal(登録商標)強アニオノン交換ビーズ(AXビーズ)は、10mg/ml濃度であった。試料調製の直前に、カチオノン交換ビーズを、20mM Tris-HCl、pH7.5、0.1%還元 Triton X100(Tris-Triton緩衝液)で一回洗浄した。本試料調製で使用する他の試薬、20mM Tris-HCl、pH7.5(Tris緩衝液)、0.5%トリフルオロ酢酸(本明細書以下「TFA溶液(TFA solution)」および50%アセトニトリル(本明細書以下「アセトニトリル溶液(Acetonitrile solution)」は自ら調製した。試薬および試料は以下のように96-ウェルプレートで準備した。

## 【0238】

10

20

40

50

列Aは、AXビーズ30uL、CXビーズ20uLおよびTris緩衝液50uLの混合液を含んだ。

## 【0239】

列Bは、Tris緩衝液100uLを含んだ。

## 【0240】

列Cは、Tris緩衝液120uLと試料30uLを含んだ。

## 【0241】

列Dは、Tris緩衝液100uLを含んだ。

## 【0242】

列Eは、脱イオン水100uLを含んだ。

10

## 【0243】

列Fは、TFA溶液50uLを含んだ。

## 【0244】

列Gは、アセトニトリル溶液50uLを含んだ。

## 【0245】

列Hは空だった。

## 【0246】

列Aのビーズおよび緩衝液を先に混合し、ビーズをCollectカウント3（何回磁気プロープが溶液内に入り、磁気ビーズを回収したかを示唆する器具パラメーター）で回収し、Tris緩衝液中、放出設定「fast」、洗浄設定 - medium、20秒間の洗浄時間で、洗浄のために列B上へ移した。ビーズ洗浄段階の最後に、ビーズをCollectカウント3で回収し、列C上へ移して試料と合わせた。ビーズ放出設定はfastである。試料結合を、5分間の結合時間で、「slow」設定にて実施した。結合段階の最後に、ビーズをCollectカウント3で回収する。回収したビーズを、放出設定「fast」、洗浄設定 - medium、20秒間の洗浄時間で、第一洗浄段階のために列D上に移した。第一洗浄段階の最後に、ビーズをCollectカウント3で回収する。回収したビーズを、放出設定「fast」、洗浄設定 - medium、20秒間の洗浄時間で、第二洗浄段階のために、列E上に移した。第二洗浄段階の最後に、ビーズをCollectカウント3で回収した。回収したビーズを、放出設定「fast」、溶出設定 - fastおよび2分間の溶出時間で、TFA溶液中の溶出のために列Fに移した。TFA溶出段階の最後に、ビーズをCollectカウント3で回収した。このTFA溶出液を回収し、質量分析によって解析した。回収したビーズを、放出設定「fast」、溶出設定 - fastおよび2分間の溶出時間で、アセトニトリル溶液中の溶出のために列Gに移した。溶出後、ビーズをCollectカウント3で除去し、列Aに捨てた。アセトニトリル(ACN)溶出液を回収して、質量分析によって解析した。

20

## 【0247】

すべての試料を二回ずつ、しかしシステムエラーを避けるために異なるプレート上で行った。溶出した試料を手動で吸引し、自動化MALDI標的化試料調製のために96-ウェルプレート中に配置した。したがって、各試料は、質量分析解析のために2つの溶出液を提供した。

30

## 【0248】

CLINPROTロボット(ブルッカー ダルトニックス社(Bruker Daltonics Inc.)、Billerica, MA)を、MS照合の前に、MALDI標的を調製するために使用した。簡単に記すと、処理には、溶出した血清試料を含む試料プレート、およびMALDIマトリックス溶液(70%アセトニトリル中の10mg/mlシナピン酸)を含むバイアルをロボット上の設計された位置にロードすることが含まれる。ロボットを制御するコンピュータからスポットティング手順を含むファイルをロードして、開始した。この場合、スポットティング手順には、5μLのマトリックス溶液を吸引すること、およびマトリックスプレート中に分配し、続いて5μL試料を分配することが含まれる。試料およびマトリックスの前混合を、混合液5μLを吸引し、マトリックスプレ

40

50

ート中で数回分配することによって実施した。前混合前に、混合液 5  $\mu$  L を吸引し、0.5  $\mu$  L を、アンカーチップ標的（ブルッカー ダルトニックス社、Billerica, MA）上、4 連続スポットに沈着させた。溶液の残り 3  $\mu$  L をゴミ箱容器中に捨てた。必要以上の試料を吸引することで、アンカーチップ標的上への試料沈着の間に誤ったスポットを導き得るディスポーザブルチップ中の空気バブルの形成を最小化した。

【0249】

B. C 8 磁気ビーズ水性相互作用クロマトグラフィー（C 8 M B - H I C）による試料調製

血清試料を、SOLB 緩衝液と混合し、実施例 5 A にて記述したようなフィルターで精製した。自己試料調製を、100MB - H I C 8（ブルッカー ダルトニックス社、Billerica, MA）として公知の CLINPROT Purification Kits を用いて、96 - ウェルプレート King Fisher（登録商標）上で実施した。キットには、C 8 磁気ビーズ、結合溶液および洗浄溶液が含まれる。すべての他の試薬は、他に言及がない限り、シグマ ケミカル社（Sigma Chem. Co.）から購入した。試薬および試料を、以下のように 96 - ウェルプレート中で設定した。

【0250】

列 A は、ブルカーノの C 8 磁気ビーズ 20  $\mu$  L と DI 水 80  $\mu$  L の混合液を含んだ。

【0251】

列 B は、血清試料 10  $\mu$  L と結合溶液 40  $\mu$  L の混合液を含んだ。

【0252】

列 C から E は、洗浄溶液 100  $\mu$  L を含んだ。

【0253】

列 F は、（有機溶媒の蒸発を最小にするために溶出段階の直前に加えた）70% アセトニトリル 50  $\mu$  L を含んだ。

【0254】

列 G は、DI 水 100  $\mu$  L を含んだ。

【0255】

列 H は空であった。

【0256】

列 A 中のビーズを前混合して、「Collect Count」3 で回収し、列 B 上に移して試料に結合させた。ビーズ放出設定を、10 秒間の放出時間で「fast」に設定した。試料結合を、5 分間の「slow」設定で実施した。結合段階の最後に、ビーズを「Collect Count」3 で回収して、第一洗浄段階（放出設定 = fast、時間 = 10 秒間、洗浄設定 = medium、時間 = 20 秒間）のために、列 C 上に移した。第一洗浄段階の終わりに、ビーズを「Collect Count」3 で回収し、第一洗浄段階でと同一のパラメーターでの第二洗浄段階のために、列 D 上に移した。第二洗浄段階の終わりに、ビーズをもう一度回収して、先に記述したように第三および最終洗浄段階のために、列 E 上に移した。第三洗浄段階の終わりに、King Fisher（商標）を、列 E から列 F の移行段階の間に中断し、70% アセトニトリル 50  $\mu$  L を列 F に加えた。アセトニトリルの添加後、処理を再開した。列 E からの回収したビーズを、溶出段階（放出設定 = fast、時間 = 10 秒間、溶出設定 = fast、時間 = 2 分間）のために、列 F に移した。溶出段階の後、ビーズを除去し、列 G 中に捨てた。すべての試料を、実施例 5 a にて以上で記述したように、二回ずつ行った。

【0257】

CLINPROT ロボット（ブルッカー ダルトニックス社、Billerica, MA）を、使用した MALDI マトリックス中で多少の変更をしただけの、先の項目で記述したような MS 照合の前に、MALDI 標的を調製するために使用した。この場合、SA の代わりに、HCCA を使用した（40% ACN / 50% MeOH / 10% 水、v/v/v 中、1 mg / mL HCCA）。すべての他のパラメーターは同じままであった。

10

20

30

40

50

## 【0258】

C. S E L D I チップを用いた試料調製

以下の試薬を使用した。

## 【0259】

1. 100 mM リン酸緩衝液、pH 7.0、250 mL 脱イオン水を、200 mM リン酸二ナトリウム溶液 152.5 mL と、200 mM リン酸一ナトリウム溶液 97.5 mM と混合することによって調製した。

## 【0260】

2. 10 mg / mL シナピン酸溶液、等容量のアセトニトリルおよび 0.4% 水性トリフル酢酸 (v/v) を混合して、最終濃度 10 mg シナピン酸 / mL 溶液を作製することによって調製した、十分な量の溶液中に、計量した量のシナピン酸を溶解することによって調製した。

10

## 【0261】

3. 脱イオン水、シナピン酸およびトリフルオロ酢酸は、フルカ ケミカルズ (Fulleka Chemicals) からであった。アセトニトリルは、ブルディック (Burdick) およびジャクソン (Jackson) からであった。

## 【0262】

8 スポット配座中の Q10 Protein Chip アレイと、標準のマイクロプレートと同一のフットプリントを有する、12 × 8 アレイ中でアレイを維持するために使用した Bioprocessors を、シフェルゲン (Ciphergen) より得た。Q10 活性表面は、四級アミン強アニオン交換器である。Ciphergen Protein Chip System, Series 4000 Matrix Assisted Laser Desorption Ionization (MALDI) 飛行時間質量分析器を使用して、チップ表面に結合したペプチドを解析した。すべての Ciphergen 産物を、Ciphergen Biosystems, Dumbarton, California より得た。

20

## 【0263】

すべての液体輸送、希釈液および洗浄液を、ハミルトン社 (Hamilton Company)、Reno, Nevada からの、Hamilton Microlab STA R 口ボットピペッターによって実施した。

30

## 【0264】

血清試料を室温にて融解し、穏やかにボルテックスすることで混合した。試料を含むバイアルを、ハミルトンピペッター上の 24 位試料ホルダー内にロードし、合計 96 試料を含む 4 試料ホルダーをロードした。2 つの Bioprocessors ホールディング Q10 チップ (192 総スポット) を、ハミルトンピペッターのデッキ上に配置した。100 mM リン酸緩衝液と脱イオン水を含む容器を、ハミルトンピペッター上にロードした。使い捨てピペットチップをまた、器具のデッキ上に配置した。

## 【0265】

すべての試料処理は、総合的に自動化した。各試料を、ハミルトンピペッターのデッキ上のマイクロプレートの 2 つの別々のウェル中で、血清 5 マイクロリットルを、リン酸緩衝液 4.5 マイクロリットルと混合することによって、2 つの別々の分液内に、1 から 10 希釈した。Q10 チップを、各スポットを、2 つのリン酸緩衝液の 150 マイクロリットル分液に暴露することによって活性化した。緩衝液が、各添加後 5 分間、表面を活性化させた。第二分液を各スポットより吸引した後、希釈血清 2.5 マイクロリットルを各スポットに加え、室温にて 30 分間インキュベートした。各試料を、Q10 チップのスポット上に配置した各希釈からの單一分液で 2 回希釈した。希釈血清の吸引に続いて、各スポットを、リン酸緩衝液 150 マイクロリットルで、および最後に脱イオン水 150 μリットルで 4 回洗浄した。処理したチップを風乾し、シナピン酸で処理し、Ciphergen 4000 中での MALDI 処理を可能にするようにマトリックスを使用した。シナピン酸マトリックス溶液を、それぞれシナピン酸溶液で十分満たされた、96 ウェルマイ

40

50

クロプレートを器具のデッキ上に配置することによって、ハミルトンピペット上にロードした。96ヘッドピペットを使用して、シナピン酸マトリックス 1マイクロリットルを、Bioprocessor上の各スポットに同時に加えた。15分間の乾燥期間の後、第二の1マイクロリットル分液を各スポットに加えて、乾燥させた。

#### 【0266】

D. 混合ビーズ試料プレップのAutoFlex MALDI-TOFデータ獲得  
 器具の捕獲範囲を、m/z 400から100,000に設定した。器具を、2から17kDaの範囲の質量をカバーしているブルッカーの較正標準を用いて、直線モードにて外部にて較正した。高品質スペクトルを回収するために、捕獲を、レーザーを除いて、ファジー制御にて完全に自動化した。レーザーのファジー制御を、レーザー出力が実験の間、一定のままになるようにとめた。器具は一般的に、固定レーザー出力にて較正するので、精度が、一定のレーザー出力を維持することより利益を受ける。他のファジー制御設定は、2から10kDaの質量範囲にて、ピークの分解能およびS/Nを制御した。これらの値は、各捕獲の前に最適化し、試料から試料への、またはスポットからスポットへの失敗した捕獲の数を最小化する一方で、スペクトルの品質を最大化するために選択した。偏向板をまた、低分子量イオン(<400m/z)を偏向するためにオンにし、マトリックスイオンでの検出器の飽和と、試料から発生するシグナルの最大化を防止する。さらに、レーザービームが試料表面を介してラスターしたので、各捕獲の前に、5つのワーニングショット(LP閾値の約5から10%上)を発射して、任意の過剰なマトリックスを除去した。各分子量スペクトルに対して、600レーザーショットを、以上で設定した分解およびS/N基準を満たす場合にのみ、一緒に加えた。すべての他の粗悪な品質のスペクトルを無視し、廃棄し、ベースライン補正または平滑化アルゴリズムは、生スペクトルの捕獲の間使用しなかった。

#### 【0267】

データを保存し、共通のm/z軸に変換して、比較を促進し、種々の統計学的ソフトウェアパッケージによって解析可能なポータブルASCIIフォーマットにて出力した。共通のm/z軸への変換は、組織内で開発した補間アルゴリズムを用いることによって達成した。

#### 【0268】

E. C8 MB-HICのAutoFlex MALDI-TOFデータ捕獲  
 器具の捕獲範囲を、m/z 1000から20,000に設定し、感度および分解能に対して最適化した。すべての他の捕獲パラメーターおよび較正法を、400レーザーショットを各質量スペクトルに対して一緒に加えたことを除いて、実施例5dにて以上で記述したように設定した。F.Q-10チップのCiphergen 4000 SELDI-TOFデータ捕獲

Bioprocessorsを、質量範囲0から50,000Daに対する最適化パラメーターを用いて、Ciphergen 4000 MALDI飛行時間質量分析器上にロードした。データをデジタル化し、530捕獲/スポット上で平均化して、イオン電流対質量/電荷(m/z)の単一スペクトルを得た。各スペクトルをサーバーに出力して、続いて捕獲後解析のために、ASCIIファイルとして読み出した。

#### 【0269】

##### G. 質量分析データの関心領域解析

質量分析データは、0から50,000の質量/電荷値と、これらの相当する強度値からなる。がんおよび非がんデータ組を構築した。がんデータ組は、すべてのがん試料からの質量スペクトルからなり、一方で、非がんデータ組は、正常被験者および良性肺疾患の患者を含む、各非がん試料からの質量スペクトルからなる。がんおよび非がんデータ組を、以下を実施するソフトウェアプログラム内で別々にアップロードした。

#### 【0270】

a) スチュードントt-検定を、各記録質量/電荷値にて決定し、p-値を与える。

#### 【0271】

10

20

30

40

50

b) がんおよび非がんスペクトルを、各群に対して1つの代表に平均化する。

【0272】

c) 平均がんスペクトルおよび平均非がんスペクトルの強度の対数比 (Log Ratio) を決定する。

【0273】

ROIsを、0.01より小さなp-値、および0.1より大きな絶対Log Ratioを10以上の連続質量値を有するように特定化し、18、36および26 ROIsが、それぞれMMB-TFA、MMB-ACNおよびMB-HICデータ組にて発見された(表14a-14c)。さらに、124 ROIs (< 20 kDa)が、表14dにて示した SELDIデータにて見られた。表14aから14dは、平均質量値を昇順に並べた、本発明のROIsを列記している。本表中で提供されたROIは、計算された間隔に対する平均質量値である(所与の間隔に対する開始および終了質量値の平均)。平均ROI質量は今後、単純化のためにROIと呼ぶ。各試料に対する各ROIの強度を、ROC解析にかけた。各マーカーに対するAUCをまた、以下の表14aから14d中で報告している。以下表14aから14cにおいて、計算ROIが、疾患および非疾患群のMSプロファイルの解析から得られた。個々の試料を、3つの異なる方法を用いて処理した。a) TFA(tfa)での溶出、およびb)アセトニトリル(acn)で連続溶出した、c)疎水性相互作用クロマトグラフィー(hic)を用いる混合磁気ビーズアニオン/カチオン交換クロマトグラフィー。各試料調製法を、ROIを得る目的のために独立して解析した。すべてのスペクトルを、ブルッカー Autoflex MALDO-TOF質量分析器で回収した。以下表14dにて、計算ROIを、疾患および非疾患群のMSプロファイルの解析から得た。すべての試料を、Q-10チップを用いて処理した。すべてのスペクトルを、Ciphergen 4000 SELDI-TOF質量分析器を用いて回収した。

【0274】

【表14】

ROI	ROI	平均	ROI	大コホート		小コホート	
開始m/z	終了m/z	ROI	名前	観察番号	AUC	観察番号	AUC
2322.911	2339.104	2331	tfa2331	538	0.66	236	0.52
2394.584	2401.701	2398	tfa2398	538	0.68	236	0.55
2756.748	2761.25	2759	tfa2759	538	0.65	236	0.60
2977.207	2990.847	2984	tfa2984	538	0.69	236	0.52
3010.649	3021.701	3016	tfa3016	538	0.63	236	0.48
3631.513	3639.602	3636	tfa3635	538	0.61	236	0.54
4188.583	4198.961	4194	tfa4193	538	0.60	236	0.56
4317.636	4324.986	4321	tfa4321	538	0.61	236	0.51
5000.703	5015.736	5008	tfa5008	538	0.70	236	0.57
5984.935	5990.126	5988	tfa5987	538	0.70	236	0.49
6446.144	6459.616	6453	tfa6453	538	0.74	236	0.65
6646.05	6658.513	6652	tfa6652	538	0.72	236	0.71
6787.156	6837.294	6812	tfa6815	538	0.71	236	0.53
8141.621	8155.751	8149	tfa8148	538	0.62	236	0.64
8533.613	8626.127	8580	tfa8579	538	0.71	236	0.58
8797.964	8953.501	8876	tfa8872	538	0.68	236	0.52
9129.621	9143.87	9137	tfa9133	538	0.63	236	0.60
12066.33	12093.36	12080	tfa12079	538	0.66	236	0.63

表14a.

【0275】

【表 15】

ROI	ROI	平均	ROI	大コホート		小コホート	
開始 m/z	終了 m/z	ROI	名前	観察番号	AUC	観察番号	AUC
3022.726	3026.825	3025	acn3024	519	0.63	244	0.51
3144.614	3182.554	3164	acn3163	519	0.70	244	0.60
3183.395	3188.023	3186	acn3186	519	0.63	244	0.54
4128.262	4135.209	4132	acn4132	519	0.61	244	0.59
4152.962	4161.372	4157	acn4157	519	0.65	244	0.65
4183.519	4194.373	4189	acn4189	519	0.52	244	0.55
4627.389	4635.759	4632	acn4631	519	0.74	244	0.68
5049.048	5114.402	5082	acn5082	519	0.68	244	0.62
5229.648	5296.428	5263	acn5262	519	0.68	244	0.61
5338.006	5374.554	5356	acn5355	519	0.64	244	0.52
5375.101	5383.848	5379	acn5378	519	0.67	244	0.62
5446.925	5457.382	5452	acn5455	519	0.68	244	0.54
5971.68	5981.476	5977	acn5976	519	0.64	244	0.58
6150.986	6166.194	6159	acn6158	519	0.63	244	0.54
6314.273	6338.877	6327	acn6326	519	0.62	244	0.58
6391.206	6406.112	6399	acn6399	519	0.67	244	0.60
6455.723	6461.713	6459	acn6458	519	0.56	244	0.65
6574.845	6607.218	6591	acn6592	519	0.68	244	0.58
6672.509	6689.568	6681	acn6681	519	0.53	244	0.70
8759.205	8791.323	8775	acn8775	519	0.64	244	0.58
8850.827	8888.382	8870	acn8871	519	0.69	244	0.55
9067.056	9095.468	9081	acn9080	519	0.65	244	0.57
9224.586	9277.996	9251	acn9251	519	0.64	244	0.59
9358.22	9384.195	9371	acn9371	519	0.65	244	0.55
9453.639	9467.414	9461	acn9459	519	0.66	244	0.76
9470.315	9473.579	9472	acn9471	519	0.70	244	0.71
9651.055	9674.867	9663	acn9662	519	0.66	244	0.52
10008.34	10022.51	10015	acn10015	519	0.63	244	0.56
10217.84	10221.98	10220	acn10216	519	0.64	244	0.55
10669.51	10689.53	10680	acn10679	519	0.61	244	0.52
10866.73	10886.56	10877	acn10877	519	0.63	244	0.50
11371.68	11745.49	11559	acn11559	519	0.63	244	0.68
14293.87	14346.94	14320	acn14319	519	0.62	244	0.58
22764.38	22771.69	22768	acn22768	519	0.68	244	0.62
22778.44	22788	22783	acn22783	519	0.68	244	0.63
22791.38	23147.21	22969	acn22969	519	0.70	244	0.63

表 14b.

【0276】

10

20

30

40

【表16】

ROI	ROI	平均	ROI	大コホート		小コホート	
開始 m/z	終了 m/z	ROI	名前	観察番号	AUC	観察番号	AUC
2016.283	2033.22	2025	hic2025	529	0.65	245	0.53
2304.447	2308.026	2306	hic2306	529	0.64	245	0.66
2444.629	2457.914	2451	hic2451	529	0.60	245	0.50
2504.042	2507.867	2506	hic2506	529	0.65	245	0.53
2642.509	2650.082	2646	hic2646	529	0.54	245	0.45
2722.417	2733.317	2728	hic2728	529	0.61	245	0.56
2971.414	2989.522	2980	hic2980	529	0.64	245	0.53
3031.235	3037.804	3035	hic3035	529	0.54	245	0.45
3161.146	3191.075	3176	hic3176	529	0.70	245	0.61
3270.723	3280.641	3276	hic3276	529	0.64	245	0.57
3789.504	3797.883	3794	hic3794	529	0.64	245	0.57
3942.315	3975.73	3959	hic3959	529	0.74	245	0.59
4999.913	5006.107	5003	hic5003	529	0.66	245	0.56
5367.59	5384.395	5376	hic5376	529	0.68	245	0.48
6002.824	6006.289	6005	hic6005	529	0.69	245	0.51
6181.86	6195.934	6189	hic6189	529	0.72	245	0.51
6380.634	6382.272	6381	hic6381	529	0.70	245	0.55
6382.569	6392.1	6387	hic6387	529	0.71	245	0.54
6438.218	6461.563	6450	hic6450	529	0.66	245	0.57
6640.279	6658.057	6649	hic6649	529	0.62	245	0.59
6815.125	6816.816	6816	hic6816	529	0.72	245	0.56
6821.279	6823.896	6823	hic6823	529	0.71	245	0.58
8788.878	8793.595	8791	hic8791	529	0.58	245	0.47
8892.247	8901.211	8897	hic8897	529	0.61	245	0.52
8908.948	8921.088	8915	hic8915	529	0.64	245	0.55
9298.469	9318.065	9308	hic9308	529	0.68	245	0.59

10

20

30

表14c.

【0277】

【表 1 7】

ROI	ROI	平均	ROI	大コホート		小コホート	
開始 m/z	終了 m/z	ROI	名前	観察番号	AUC	観察番号	AUC
2327	2336	2331	Pub2331	513	0.65	250	0.62
2368	2371	2369	Pub2369	513	0.64	250	0.60
2384	2389	2387	Pub2386	513	0.67	250	0.62
2410	2415	2413	Pub2412	513	0.67	250	0.63
2431	2435	2433	Pub2433	513	0.72	250	0.72
2453	2464	2459	Pub2458	513	0.70	250	0.62
2672	2682	2677	Pub2676	513	0.73	250	0.68
2947	2955	2951	Pub2951	513	0.72	250	0.64
2973	2979	2976	Pub2976	513	0.63	250	0.58
3016	3020	3018	Pub3018	513	0.50	250	0.51
3168	3209	3189	Pub3188	513	0.69	250	0.59
3347	3355	3351	Pub3351	513	0.70	250	0.67
3409	3414	3412	Pub3411	513	0.60	250	0.57
3441	3456	3449	Pub3448	513	0.72	250	0.58
3484	3503	3494	Pub3493	513	0.72	250	0.67
3525	3531	3528	Pub3527	513	0.62	250	0.55
3548	3552	3550	Pub3550	513	0.62	250	0.62
3632	3650	3641	Pub3640	513	0.63	250	0.57
3656	3662	3659	Pub3658	513	0.51	250	0.49
3678	3688	3683	Pub3682	513	0.72	250	0.69
3702	3709	3706	Pub3705	513	0.57	250	0.55

10

20

3737	3750	3744	Pub3743	513	0.69	250	0.67
3833	3845	3839	Pub3839	513	0.62	250	0.59
3934	3955	3944	Pub3944	513	0.65	250	0.57
4210	4217	4214	Pub4213	513	0.62	250	0.56
4299	4353	4326	Pub4326	513	0.69	250	0.59
4442	4448	4445	Pub4444	513	0.61	250	0.52
4458	4518	4488	Pub4487	513	0.75	250	0.69
4535	4579	4557	Pub4557	513	0.73	250	0.68
4590	4595	4592	Pub4592	513	0.70	250	0.66
4611	4647	4629	Pub4628	513	0.77	250	0.66
4677	4687	4682	Pub4682	513	0.72	250	0.69
4698	4730	4714	Pub4713	513	0.73	250	0.70
4742	4759	4751	Pub4750	513	0.76	250	0.73
4779	4801	4790	Pub4789	513	0.70	250	0.72
4857	4865	4861	Pub4861	513	0.72	250	0.75
4987	4996	4992	Pub4991	513	0.67	250	0.57
5016	5056	5036	Pub5036	513	0.65	250	0.54
5084	5194	5139	Pub5139	513	0.61	250	0.51
5208	5220	5214	Pub5213	513	0.57	250	0.52
5246	5283	5265	Pub5264	513	0.59	250	0.56
5295	5420	5357	Pub5357	513	0.64	250	0.54
5430	5537	5484	Pub5483	513	0.62	250	0.54
5570	5576	5573	Pub5573	513	0.59	250	0.57
5590	5595	5593	Pub5592	513	0.60	250	0.54
5612	5619	5615	Pub5615	513	0.55	250	0.53
5639	5648	5644	Pub5643	513	0.68	250	0.63
5679	5690	5685	Pub5684	513	0.66	250	0.59
5752	5804	5778	Pub5777	513	0.71	250	0.63
5839	5886	5862	Pub5862	513	0.73	250	0.67
5888	5909	5898	Pub5898	513	0.63	250	0.56
6008	6018	6013	Pub6013	513	0.61	250	0.57
6047	6058	6053	Pub6052	513	0.64	250	0.63
6087	6103	6095	Pub6094	513	0.59	250	0.54
6111	6124	6118	Pub6117	513	0.70	250	0.67
6153	6160	6156	Pub6156	513	0.57	250	0.51
6179	6188	6183	Pub6183	513	0.65	250	0.60
6192	6198	6195	Pub6194	513	0.57	250	0.49
6226	6272	6249	Pub6249	513	0.66	250	0.63
6277	6286	6281	Pub6281	513	0.62	250	0.65
6297	6307	6302	Pub6302	513	0.71	250	0.67
6352	6432	6392	Pub6391	513	0.65	250	0.56
6497	6570	6534	Pub6533	513	0.63	250	0.59
6572	6603	6587	Pub6587	513	0.60	250	0.55
6698	6707	6702	Pub6702	513	0.57	250	0.52
6715	6723	6719	Pub6718	513	0.64	250	0.57
6748	6849	6799	Pub6798	513	0.77	250	0.69
7197	7240	7219	Pub7218	513	0.73	250	0.65

10

20

30

40

7250	7262	7256	Pub7255	513	0.72	250	0.65
7310	7326	7318	Pub7317	513	0.71	250	0.65
7401	7427	7414	Pub7413	513	0.73	250	0.69
7435	7564	7499	Pub7499	513	0.76	250	0.73
7611	7616	7614	Pub7613	513	0.67	250	0.60
7634	7668	7651	Pub7651	513	0.70	250	0.63
7699	7723	7711	Pub7711	513	0.72	250	0.66
7736	7748	7742	Pub7742	513	0.69	250	0.65
7768	7782	7775	Pub7775	513	0.63	250	0.57
7935	7954	7945	Pub7944	513	0.64	250	0.61
7976	7985	7981	Pub7980	513	0.62	250	0.59
7999	8006	8003	Pub8002	513	0.58	250	0.60
8134	8239	8186	Pub8186	513	0.73	250	0.62
8286	8308	8297	Pub8297	513	0.69	250	0.62
8448	8461	8455	Pub8454	513	0.61	250	0.59
8476	8516	8496	Pub8496	513	0.69	250	0.64
8526	8567	8547	Pub8546	513	0.73	250	0.66
8579	8634	8606	Pub8606	513	0.80	250	0.70
8640	8684	8662	Pub8662	513	0.80	250	0.71
8710	8758	8734	Pub8734	513	0.74	250	0.67
8771	8781	8776	Pub8776	513	0.56	250	0.59
8913	8947	8930	Pub8930	513	0.68	250	0.64
8961	8977	8969	Pub8969	513	0.65	250	0.57
9122	9162	9142	Pub9142	513	0.66	250	0.66
9199	9233	9216	Pub9216	513	0.59	250	0.62
9311	9323	9317	Pub9317	513	0.57	250	0.60
9357	9370	9364	Pub9363	513	0.58	250	0.63
9409	9458	9434	Pub9433	513	0.67	250	0.65
9478	9512	9495	Pub9495	513	0.61	250	0.63
9629	9667	9648	Pub9648	513	0.62	250	0.64
9696	9749	9722	Pub9722	513	0.70	250	0.67
9977	10281	10129	pub10128	513	0.66	236	0.48
10291	10346	10318	pub10318	513	0.66	236	0.56
10692	10826	10759	pub10759	513	0.62	236	0.51
10867	11265	11066	pub11066	513	0.61	236	0.55
11339	11856	11597	pub11597	513	0.75	236	0.77
12080	12121	12100	pub12100	513	0.63	236	0.54
12159	12228	12194	pub12193	513	0.59	236	0.49
12422	12582	12502	pub12501	513	0.66	236	0.64
12620	12814	12717	pub12717	513	0.73	236	0.60
12839	12854	12846	pub12846	513	0.72	236	0.56
13135	13230	13182	pub13182	513	0.69	250	0.53
13386	13438	13412	pub13412	513	0.54	250	0.56
13539	13604	13572	pub13571	513	0.71	250	0.64
14402	14459	14430	pub14430	513	0.74	250	0.67
15247	15321	15284	pub15284	513	0.69	250	0.60
15414	15785	15600	pub15599	513	0.76	250	0.71

10

20

30

40

15872	15919	15896	pub15895	513	0.58	250	0.57
16366	16487	16427	pub16426	513	0.66	250	0.60
16682	16862	16772	pub16771	513	0.69	250	0.61
16984	17260	17122	pub17121	513	0.68	250	0.60
17288	17389	17339	pub17338	513	0.81	250	0.72
17431	18285	17858	pub17858	513	0.81	250	0.68
18321	18523	18422	pub18422	513	0.73	250	0.59
18728	18804	18766	pub18766	513	0.65	250	0.52
18921	19052	18987	pub18986	513	0.69	250	0.55

10

表14d.

## 【0278】

H. ROIs ファミリーの同定： JMP (商標) 統計パッケージ (SAS インスティテュート社 (SAS Institute Inc.)、Cary, NC) プログラムの他変数解析機能を使用して、非常に相關した ROIs を同定した。二次元相關係数マトリックスを、JMP プログラムから抽出し、マイクロソフト (Microsoft) Excel によってさらに解析した。各 ROIs に対して、相關係数が 0.8 を超える ROIs の組を同定した。これらの ROIs が一緒に、相關 ROIs のファミリーになる。表15は、相關ファミリー、これらの相關メンバー ROIs 、大コホートでの、メンバー ROIs に対する AUC 値、およびファミリーの他のメンバーに対する相關係数の平均を示している。したがって、3449と3494の質量を有する ROIs が非常に相關しており、本発明の文脈内で、互いに対して代用可能であることが分かる。

20

## 【0279】

【表18】

グループ A (n=2)			
ROI名	メンバー	AUCs	相関係数
Pub3448	3449	0.72	0.81
Pub3493	3494	0.72	0.81
グループ B (n=2)			
ROI名	メンバー	AUCs	相関係数
Pub4487	4488	0.75	0.8
Pub4682	4682	0.72	0.8
グループ C (n=9)			
ROI名	メンバー	AUCs	相関係数
Pub8776	8776	0.56	0.8
Pub8930	8930	0.68	0.83
Pub9142	9142	0.66	0.92
Pub9216	9216	0.59	0.91
Pub9363	9363	0.58	0.88
Pub9433	9434	0.67	0.94
Pub9495	9495	0.61	0.94
Pub9648	9648	0.62	0.93
Pub9722	9722	0.7	0.89
グループ D (n=15)			
ROI名	メンバー	AUCs	相関係数
Pub5036	5036	0.65	0.71
Pub5139	5139	0.61	0.81
Pub5264	5265	0.59	0.79
Pub5357	5357	0.64	0.85
Pub5483	5484	0.62	0.87
Pub5573	5573	0.59	0.8
Pub5593	5593	0.6	0.78
Pub5615	5615	0.55	0.77
Pub6702	6702	0.57	0.79
Pub6718	6718	0.64	0.73
Pub10759	10759	0.62	0.77
Pub11066	11066	0.61	0.84
Pub12193	12194	0.59	0.79
Pub13412	13412	0.54	0.78
acn10679	acn10679	0.61	0.73
acn10877	acn10877	0.62	0.77
グループ E (n=6)			
ROI名	メンバー	AUCs	相関係数
Pub6391	6392	0.65	0.9
Pub6533	6534	0.63	0.9
Pub6587	6587	0.6	0.87
Pub6798	6799	0.76	0.85
Pub9317	9317	0.57	0.7
Pub13571	13571	0.71	0.67
グループ F (n=8)			

10

20

30

40

ROI名	メンバー	AUCs	相関係数
Pub7218	7219	0.73	0.82
Pub7255	7255	0.72	0.73
Pub7317	7318	0.71	0.88
Pub7413	7414	0.73	0.81
Pub7499	7499	0.76	0.84
Pub7711	7711	0.72	0.76
Pub14430	14430	0.74	0.77
Pub15599	15600	0.76	0.82
<b>グループ G (n=7)</b>			
ROI名	メンバー	AUCs	相関係数
Pub8496	8496	0.69	0.78
Pub8546	8547	0.73	0.88
Pub8606	8606	0.8	0.84
Pub8662	8662	0.79	0.77
Pub8734	8734	0.74	0.45
Pub17121	17122	0.68	0.78
Pub17338	17339	0.81	0.54
<b>グループ H (n=3)</b>			
ROI名	メンバー	AUCs	相関係数
Pub6249	6249	0.66	0.82
Pub12501	12502	0.66	0.87
Pub12717	12717	0.73	0.87
<b>グループ I (n=5)</b>			
ROI名	メンバー	AUCs	相関係数
Pub5662	5662	0.73	0.93
Pub5777	5777	0.71	0.92
Pub5898	5898	0.63	0.89
Pub11597	11597	0.75	0.93
acn11559	acn11559	0.63	0.84
<b>グループ J (n=5)</b>			
ROI名	メンバー	AUCs	相関係数
Pub7775	7775	0.63	0.39
Pub7944	7944	0.64	0.83
Pub7980	7980	0.62	0.72
Pub8002	8002	0.58	0.77
Pub15895	15895	0.58	0.75
<b>グループ K (n=4)</b>			
ROI名	メンバー	AUCs	相関係数
Pub17858	17858	0.81	0.84
Pub18422	18422	0.73	0.92
Pub18766	18766	0.69	0.89
Pub18986	18986	0.65	0.91
<b>グループ L (n=12)</b>			
ROI名	メンバー	AUCs	相関係数
Pub3018	3018	0.5	0.78
Pub3640	3640	0.62	0.82

10

20

30

40

Pub3658	3658	0.51	0.81
Pub3682	3682	0.72	0.77
Pub3705	3705	0.57	0.79
Pub3839	3839	0.62	0.75
hic2451	hic2451	0.6	0.78
hic2646	hic2646	0.54	0.7
hic3035	hic3035	0.54	0.72
tfa3016	tfa3016	0.63	0.78
tfa3635	tfa3635	0.61	0.78
tfa4321	tfa4321	0.61	0.74
<b>グループ M (n=2)</b>			
ROI 名	メンバー	AUCs	相関係数
Pub2331	2331	0.65	0.9
tfa2331	tfa2331	0.66	0.9
<b>グループ N (n=2)</b>			
ROI 名	メンバー	AUCs	Corr Coeff
Pub4557	4557	0.73	0.81
Pub4592	4592	0.71	0.81
<b>グループ O (n=6)</b>			
ROI 名	メンバー	AUCs	Corr Coeff
acn4631	acn4631	0.74	0.81
acn5082	acn5082	0.68	0.85
acn5262	acn5262	0.68	0.9
acn5355	acn5355	0.64	0.87
acn5449	acn5449	0.7	0.88
acn5455	acn5455	0.68	0.88
<b>グループ P (n=6)</b>			
ROI 名	メンバー	AUCs	相関係数
acn6399	acn6399	0.67	0.78
acn6592	acn6592	0.68	0.8
acn8871	acn8871	0.69	0.79
acn9080	acn9080	0.65	0.84
acn9371	acn9371	0.65	0.83
acn9662	acn9662	0.66	0.79
<b>グループ Q (n=2)</b>			
ROI 名	メンバー	AUCs	相関係数
acn9459	acn9459	0.66	0.91
acn9471	acn9471	0.7	0.91
<b>グループ R (n=4)</b>			
ROI 名	メンバー	AUCs	相関係数
hic2506	hic2506	0.65	0.82
hic2980	hic2980	0.64	0.87
hic3176	hic3176	0.69	0.8
tfa2984	tfa2984	0.69	0.78
<b>グループ S (n=2)</b>			
ROI 名	メンバー	AUCs	相関係数
hic2728	hic2728	0.61	0.81

10

20

30

40

hic3276	hic3276	0.64	0.81
グループ T (n=6)			
ROI名	メンバー	AUCs	相関係数
hic6381	hic6381	0.7	0.83
hic6387	hic6387	0.71	0.84
hic6450	hic6450	0.66	0.81
hic6649	hic6649	0.62	0.73
hic6816	hic6816	0.72	0.81
hic6823	hic6823	0.71	0.79
グループ U (n=2)			
ROI名	メンバー	AUCs	相関係数
hic8791	hic8791	0.58	0.8
hic8897	hic8897	0.61	0.8
グループ V (n=2)			
ROI名	メンバー	AUCs	相関係数
tfa6453	tfa6453	0.74	0.84
tfa6652	tfa6652	0.72	0.84
グループ W (n=2)			
ROI名	メンバー	AUCs	相関係数
hic6005	hic6005	0.69	0.74
hic5376	hic5376	0.68	0.74
グループ X (n=3)			
ROI名	メンバー	AUCs	相関係数
Pub4713	4714	0.73	0.83
Pub4750	4751	0.76	0.66
Pub4861	4861	0.72	0.65

10

20

30

40

50

表 15. 相関する関心領域のファミリー

## 【実施例 6】

## 【0280】

判別解析、決定木解析および主成分分析を用いた、バイオマーカーの多変数解析

多変数解析を、免疫アッセイバイオマーカーおよび関心領域上で実施した。すべての異なる解析を、JMP統計学的パッケージを用いて実施した。目的を単純化するために、判別解析 (DA)、主成分分析 (PCA) および決定木 (DT) を一般的に、本明細書で多変数法 (MVM) と呼ぶ。PCAにおいて、データ中の総変数の 90% 超に相当する、たった最初の 15 の主成分が抽出されたということは注目に値する。因子ローディングおよび / または平等分配を使用して、各主成分に最も寄与した 1 つの因子のみ (バイオマーカー) を抽出した。因子ローディングの平方が、各主成分における各因子の相対的寄与を反映するので、これらの値を、各主成分に最も寄与したマーカーの選別のための基礎として使用した。したがって、最初の 15 主成分に最も寄与している 15 因子 (バイオマーカー) を抽出した。DAにおいて、マーカーを選別する工程を、より多くのマーカーの添加が、分類出力に影響を与えるくなるまで実施した。一般的に、DA は 5 から 8 バイオマーカーを使用した。DTs の場合に、約 5 つのバイオマーカーを有する 6 - 結節木を構築して、評価した。

## 【0281】

バイオマーカーを、既知のブートストラッピングおよびリープ - ワン - アウトバリデーション法 (Richard O. Duda et al. In Pattern Classification, 2<sup>nd</sup> Edition, pp. 485, Wiley - Interscience (2000)) を使用することにより評価した。10 倍トレーニング工程を使用して、定期的に現れる強固なバイオマーカーを同定した。強固なバイオマーカー

ーは、トレーニング組の少なくとも50%にて出現するマーカーとして定義した。したがって、本発明者らの10倍トレーニング工程において、5以上の頻度でのバイオマーカーを、さらなる評価のために選別した。以下表16は、各コホートでの各方法において、定期的に現れるバイオマーカーを要約している。

【0282】

種々の統計方法を用いたバイオマーカー探索のためのアプローチによって、より広い候補バイオマーカーのレパートリーを提供することによる、異なった利点が提供される(図1)。DAおよびPCAのようないくつかの方法が、正常に分布したデータでよく機能する一方で、対数回帰および決定木のような他の非パラメーター法は、分離した、均一に分散していない、または著しい変動を有するデータでよりよく機能する。マーカーが、集団内で正常に分散していてよく、またはしていなくてもよいため、このようなアプローチは、異なる供給源(質量分析器、免疫アッセイ、臨床歴など)からの(バイオマーカーおよびバイオメトリックパラメーターのような)マーカーが、単一のパネルに統合されることが理想的である。

【0283】

【表19】

小コホート					大コホート				
AUC	トップマーカー	DA	PCA	DT	AUC	トップマーカー	DA	PCA	DT
1 0.76	acn9459	x			1 0.81	pub17858	x		x
2 0.75	pub4861	x		x	2 0.81	pub17338		x	
3 0.66	CEA			x	3 0.8	pub8606	x		
4 0.65	pub9433		x		4 0.72	pub4861	x		x
5 0.64	pub9648			x	5 0.69	pub3743	x		x
6 0.64	pub2951	x			6 0.67	acn6399		x	
7 0.63	pub6052	x			7 0.66	tfa2331		x	
8 0.6	tfa2759	x			8 0.65	pub9433		x	
9 0.6	tfa9133	x			9 0.58	acn6592		x	
10 0.59	acn4132		x		10 0.56	pub4213		x	
11 0.58	acn6592		x		11 0.55	acn9371		x	
12 0.57	pub7775		x						
13 0.56	pub4213		x						
14 0.55	acn9371		x						
	合計	6	6	3		合計	4	6	4

表16. 多変数解析を用いて同定されたマーカー。少くとも50%時間を示したマーカーのみ選択して、別の考察に付した。上表における“x”と“X”に差はない。

【実施例7】

【0284】

分割スコア法(Split and Score Method)(本明細書以下「SSM」)

A. 改善分割スコア法(SSM)

Moret al. (PNAS, 102(21):7677(2005)を参照のこと)によって記述されたスプリットポイントスコアリング法を実行するためのインタラクティブソフトウェアが、マイクロソフト(Microsoft)(登録商標)ウィンドウズ(登録商標)下で実行すると記述された。本ソフトウェアは、試料の組に対するマーカー

10

20

30

40

50

-(バイオマーカーおよびバイオメトリックパラメーター) 解析の結果を保存するための普通の媒体である、マイクロソフト(登録商標)Excelスプレッドシートを読む。データを、試料の疾患を指定するための領域を有する単一ワークシート上に保存可能であり、1つが疾患試料に対して、他が非疾患試料に対しての、2つのワークシート上に、および1対がトレーニング試料、疾患および非疾患で、他の対がトレーニング試料、疾患および非疾患に対してである、4つのワークシート上に保存可能である。最初の2つの場合、ユーザーは、入力からの無作為に選別したトレーニングおよび試験対を自動的に産出するために、本ソフトウェアを使用してよい。最後の場合、多数のExcelファイルを、一度に読み出して、単一の履行で解析してよい。

## 【0285】

10

ソフトウェアは、データ上で回収したマーカーすべてのリストを表している。ユーザーは、このリストから解析にて使用すべきマーカーの組を選別する。ソフトウェアが自動的に、疾患および非疾患トレーニングデータ組からの各マーカーに対するスプリット点を計算し、疾患群が非疾患に対して上昇したか、または減少したかどうかを決定する。スプリット点は、各単一マーカーの精度を最大化するために選択する。スプリット点はまた、手動で設定し、調節してよい。

## 【0286】

すべての解析において、選択したマーカーの組を用いて、各可能性のある閾値での精度、特異性および感度を、トレーニングおよび試験組両方に対して計算する。多数の結果を産出する解析において、これらの結果を、トレーニング組精度によって順序づける。

20

## 【0287】

3つの様式の解析が利用可能である。最も単純な様式は、選択したマーカーのみを用いて標準の結果を計算する。第二の様式は、選択したリスト中の、価値が最小のマーカーを決定する。多重計算を実施して、それぞれの可能性のあるマーカーのサブセットに対するものを、単一マーカーを除去することによって形成した。最もよい精度を有するサブセットによって、サブセットを作製するために除去したマーカーが、全組中、最小の寄与を作ることが示唆されている。これらの最初の2つのモードに対する結果は、本質的に即時である。最も関与した計算は、選択したマーカーのすべての可能性のある組合せを探索する。20の最もよい出力が報告されている。この最終選択肢は、多数の候補に関与し得る。したがって、これは計算的に非常に強く、しばしば完了してよい。使用した各追加マーカーは操作時間を倍にした。

30

## 【0288】

およそ20マーカーに対して、20の最もよい結果のすべてにおいて見られる、通常6から10マーカーが存在することがしばしば分かった。次いでこれらは、組からの2から4の他のマーカーと適合する。このことは、診断パネルのためにマーカーを選択することにおいて、自由度が存在し得ることを指摘している。トップ20の最もよい出力が、精度において、一般的に同様であるが、感度および特異性において、有意に異なってよい。この方法でのマーカーのすべての可能性のある組合せを見ることによって、臨床的に最も有用であり得る組合せの洞察を提供する。

## 【0289】

40

B. 分割加重スコア法(Split and Weighted Scoring Method)(本明細書以下「SWSM」)

本明細書で先に議論したように、本方法は、1つのマーカーの測定を、1つの多くの可能性のあるスコアに変換することに關与する、加重スコアリング法である。これらのスコアは、等式：

スコア = AUC \* 因子 / (1 - 特異性)

を用いて誘導される。

## 【0290】

マーカーサイトケラチン19を、具体例として使用可能である。サイトケラチン19レベルは、小コホート中で、0.4から89.2ng/mLの範囲である。解析ソフトウェ

50

アを用いて、R O C 曲線を、がんが陽性であるように、サイトケラチン 19 データで産出した。擬陽性率 (1 - 特異性) を x - 軸にプロットし、真の陽性率 (感度) を y - 軸にプロットし、曲線上の各点に相当するサイトケラチン 19 値を含むスプレッドシートを作成した。カットオフ 3 . 3 n g / m L にて、特異性は 90 % であり、擬陽性率は 10 % であった。この A U C が 0 . 7 より大きく、0 . 8 より小さかったので (表 2 を参照のこと) 、このマーカーに対して、3 つの因子が任意に与えられた。しかしながら、任意の整数を因子として使用可能である。この場合、数を増やすことで、よりよい臨床性能を示唆している、より高い A U C を有するバイオマーカーとともに使用する。したがって、3 . 3 n g / m L 以上のサイトケラチン 19 を有する個体に対するスコアを計算した。

スコア = A U C \* 因子 / (1 - 特異性)

10

スコア = 0 . 70 \* 3 / (1 - 0 . 90)

スコア = 21

3 . 3 n g / m L 超のサイトケラチン 19 の任意の値に対して、このようにスコア 21 を与えた。1 . 9 より大きいが、3 . 3 より小さいサイトケラチン 19 の任意の値に対して、スコア 8 . 4 を与え、以下同様である (以下の表 17 a を参照のこと)。

【0291】

【表 20】

カットオフ	サイトケラチン 19 AUC 特異性	スコア
3.3	0.90	21
1.9	0.75	8.4
1.2	0.50	4.2
0	0	0.0

表 17a サイトケラチン 19 に対して与えられた 4 つの可能性のあるスコア

【0292】

特異性レベルが増加するにつれて、スコアの値が増加する。選択した特異性の値を、任意の 1 つのマーカーに対して調整可能である。任意の 1 つのマーカーに対して選択した特異性レベルの数を調整可能である。本方法によって、特異性はバイオマーカーのパネルへの寄与を改善することができる。

30

【0293】

加重スコアリング法の比較を、上の実施例 7 A で記述した二成分スコアリング法に対して実施した。本実施例において、パネルは、8 つの免疫アッセイバイオマーカー、C E A 、サイトケラチン 19 、サイトケラチン 18 、C A 125 、C A 15 - 3 、C A 19 - 9 、p r o G R P および S C C からなった。A U C s 、因子、選択した特異性レベル、これらの特異性レベルのそれぞれにおけるスコアを、以下表 17 b 中、各マーカーに対して表で示している。これらの個々のカットオフおよびスコアを用いて、各試料を 8 つのバイオマーカーに対して表で示した。各試料の総スコアを合計し、R O C 曲線中にプロットした。本 R O C 曲線を、表 18 にて提供された小コホートスプリット点または大コホートスプリット点いずれかでの二成分スコアリング方を用いて産出された R O C 曲線と比較した (実施例 8 A を参照のこと)。加重スコアリング法、二成分スコアリング法大コホートスプリット点、および二成分スコアリング法小コホートスプリット点に対する A U C 値はそれぞれ、0 . 78 、0 . 76 および 0 . 73 であった。A U C 値によって示唆されたようなパネルの総合性能の改善は別として、加重スコアリング法は、パネルに対して、多数の可能性のあるスコア値を提供する。より多数の可能性のあるパネルスコアの 1 つの利点は、陽性試験に対して、カットオフを設定するために、より多くの選択肢が存在することである (図 5 を参照のこと)。8 バイオマーカーパネルに適用した二成分スコアリング法は、増分 1 をともなう、0 から 8 の範囲のパネル出力値を持ち得る (図 5 を参照のこと)。

40

【0294】

50

【表21】

	CEA	CK-18	proGRP	CA15-3	CA125	SCC	CK-19	CA19-9
AUC	0.67	0.65	0.62	0.58	0.67	0.62	0.7	0.55
因子	2	2	2	1	2	2	3	1
値@50%特異性*	2.02	47.7	11.3	16.9	15.5	0.93	1.2	10.6
値@75%特異性*	3.3	92.3	18.9	21.8	27	1.3	1.9	21.9
値@90%特異性*	4.89	143.3	28.5	30.5	38.1	1.98	3.3	45.8
50%特異性以下のスコア	0	0	0	0	0	0	0	0
50%特異性以上のスコア	2.68	2.6	2.48	1.16	2.68	2.48	4.2	1.1
75%特異性以上のスコア	5.36	5.2	4.96	2.32	5.36	4.96	8.4	2.2
90%特異性以上のスコア	13.4	13	12.4	5.8	13.4	12.4	21	5.5

表17b

\*これらの値のそれぞれがスプリット点を表している

## 【実施例8】

## 【0295】

スプリット&amp;スコア法( SSM )を用いる、肺がんに対する予測法

## A. 免疫アッセイバイオマーカーのSSM

実施例2にて議論したように、いくつかのバイオマーカーを、免疫学的アッセイによって検出した。これらには、サイトケラチン19、CEA、CA125、SCC、proGRP、サイトケラチン18、CA19-9およびCA15-3が含まれた。これらのデータを、SSMを用いて評価した。一緒に示されたこれらのバイオマーカーが臨床利用を制限した。良性肺疾患および肺がんを表している小コホートにおいて、陽性結果として4以上の閾値を有する、8つのバイオマーカーパネルの精度が、10小コホート試験組にわたり、平均64.8%精度(AUC 0.69)を達成した。正常ならびに良性肺疾患および肺がんを表している、大コホートにおいて、陽性結果として4以上の閾値を有する、8つのバイオマーカーパネルの精度は、10大コホート試験組にわたり、平均77.4%精度(AUC 0.79)を達成した。

## 【0296】

パックイヤーのバイオメトリックパラメーターを含むことが、大体5%分

のこれらのバイオマーカーの予想精度を改善した。したがって、4以上の閾値を陽性結果

10

20

30

40

50

として有する、8つのバイオマーカーおよび1つのバイオメトリックパラメータパネルの精度が、10小コホート試験組にわたり、平均69.6%精度(AUC 0.75)を達成した。

【0297】

【表22】

小コホート ト点(所定 のカットオフ)	大コホート ト点(所定 のカットオフ)	対照群
---------------------------	---------------------------	-----

10

CEA	4.82	0	9.2	0	正常<=スブ リット点
CK 19	1.89	0.45	2.9	0.3	正常<=スブ リット点
CA 125	13.65	8.96	26	2.6	正常<=スブ リット点
CA 15-3	13.07	3.36	20.1	2.6	正常<=スブ リット点
CA 19-9	10.81	11.25	41.1	18.5	正常<=スブ リット点
SCC	0.92	0.11	1.1	0.1	正常<=スブ リット点
proGRP	14.62	8.53	17.6	0	正常<=スブ リット点
CK-18	57.37	2.24	67.2	9.5	正常<=スブ リット点
パラインフルエン ザ	103.53	32.64	79.2	9.8	正常>=スブ リット点
PACK-yr	30		30		正常<=スブ リット点

20

30

表 18. SSM アルゴリズムを用いて各個々の免疫マーカーに対して計算したスブリット点

【0298】

B. ROC/AUC によって選別されたバイオマーカーおよびバイオメトリックパラメーターのSSM

推定バイオマーカーを、多変数統計学的方法を用いて同定した、実施例6とは反対に、ROC/AUC 解析が関与する、単純な、非パラメトリ一法をこの場合に使用して、推定バイオマーカーを同定した。本方法を適用することによって、許容可能な臨床性能を有する(AUC > 0.6)個々のマーカーをさらなる解析のために選択した。トップ15のバイオマーカーとバイオメトリックパラメーター(ピークイヤー)のみが、選択され、群は、本明細書以下で、16 AUC 群(小および大)と意味される。

40

【0299】

【表 2 3】

大コホート			小コホート		
マーカー	観察番号	AUC	マーカー	観察番号	AUC
pub17338	513	0.813	pub11597	236	0.766
pub17858	513	0.812	acn9459	244	0.761
pub8606	513	0.798	pub4861	250	0.75
pub8662	513	0.796	pack-yr	257	0.739
pub4628	513	0.773	pub4750	250	0.729
pub6798	513	0.765	pub7499	250	0.725
pub7499	513	0.762	pub2433	250	0.719
pub4750	513	0.76	CK 19	248	0.718
pub15599	513	0.757	pub4789	250	0.718
pub11597	513	0.751	pub17338	250	0.718
pub4487	513	0.747	pub8662	250	0.713
tfa6453	538	0.744	acn9471	244	0.712
pack years	249	0.741	pub15599	250	0.711
pub8734	513	0.741	tfa6652	236	0.71
pub14430	513	0.741	pub8606	250	0.703
hic3959	529	0.741	acn6681	244	0.703

表19. バイオマーカーおよびバイオメトリックパラメーター、Top 15 (パックイヤー)

10

20

## 【0300】

16 AUC 小コホートマーカーの最適化組合せ (パネル) を、10 のトレーニングサブセットのそれれにおいて、SSM を用いて決定した。この工程は、SSM を用いて、バイオメトリックパラメーター 喫煙歴 (パックイヤー) のない状態 (表 20 a)、およびある状態 (表 20 b) 両方で実施した。したがって、(バイオメトリックパラメーター パック - 年を除く) 15 バイオマーカーまたは 15 バイオマーカーと 1 バイオメトリックパラメーター (パックイヤー) (16 AUC) が、分割スコア法に対する入力変数であった。10 トレーニング組のそれれに対数する最適パネルを、総精度に基づいて決定した。各パネルを残っている、処理していない試料に対して試験し、性能統計解析を記録した。10 のパネルを次いで比較して、各バイオマーカーの頻度を記録した。処理を、バイオメトリックパックイヤーを含む、および含まない、で 2 回実施した。これら 2 つの処理の結果を、以下表 20 a および 20 b にて表している。今一度、5 以上の頻度での強固なマーカーを、さらなる考慮のために選択した。処理を、大コホートのために繰り返し、結果を表 20 c にて示している。表 20 a および 20 b には、a) 15 AUC バイオマーカーのみ、および b) 15 AUC バイオマーカーおよびバイオメトリックパラメーター パックイヤーに対するマーカーの頻度を示している小コホートの SSM の結果の部分リストが含まれる。第一の表 (20 a) において、5 つのマーカーのみが、5 以上の頻度であることに注意すること。表 20 b において、7 マーカーがこの基準に適合する。表 20 c は、大コホートの SSM 結果の部分リストを含み、15 AUC マーカーに対するマーカーの頻度を示している。11 マーカーが、5 以上の頻度を有することに注意すること。

30

## 【0301】

40

【表24】

トレーニング 組番号	CK 19	pub4789	acn9459	Pub11597	tfa6652	pub2433	pub4713
1	x	x	x	x			
2	x	x	x	x			x
3	x	x	x			x	
4	x	x		x			x
5	x	x	x			x	
6	x	x	x			x	
7	x	x	x	x			x
8	x	x	x	x	x	x	
9	x	x	x			x	
10	x	x	x	x		x	x
頻度	10	10	9	6	5	3	3

表20a.

【0302】

【表25】

トレーニング 組番号	acn 9459	CK 19	pkrys	Pub 11597	pub 4789	pub 2433	pub 4861	tfa 6652	acn 9471
1	x	x	x			x	x		
2	x	x	x	x	x	x			
3	x	x	x	x	x			x	
4	x	x	x	x	x	x		x	x
5	x	x	x	x			x	x	
6	x		x	x	x				x
7	x	x	x	x	x	x	x		
8	x	x	x		x		x		x
9	x	x	x	x			x	x	
10	x	x		x	x	x			x
頻度	10	9	9	8	7	5	5	4	4

表20b.

【0303】

【表26】

トレーニング 組番号	pub 11597	pub 4487	pub 17338	pub 8606	pub 6798	tfa 6453	pub 4750	hic 3959	pub 8662	pub 4628	pub 17858
1	x	x	x	x	x	x		x			
2	x	x	x	x	x	x	x	x	x		
3	x	x	x	x			x	x	x		x
4	x		x		x	x		x			
5	x	x	x	x		x	x		x	x	x
6	x	x		x	x	x	x	x	x	x	x
7	x	x			x	x	x	x	x	x	x
8	x	x	x		x			x	x	x	x
9	x	x		x		x			x	x	x
10	x	x	x	x	x	x	x	x	x	x	x
頻度	10	9	7	7	7	7	7	7	6	6	5

表20c.

【0304】

## C. MVM によって選択されたバイオマーカーの SSM

1つの多変数法の例は、決定木解析である。決定木解析のみを用いて同定されたバイオマーカーを、ひとまとめにして考え、SSMにて使用した。この群のバイオマーカーは、16AUCとして指定されたバイオマーカーの群と同様の臨床利用性を示した。例として、試験組1(10のうち)は、バイオメトリックパラメーターパックイヤーなしで、0.90、バイオメトリックパラメーターパックイヤーありで、0.91(試験)のAUC(試験)を有する。

## 【0305】

DTバイオマーカーを、PCAおよびDAを用いて同定したバイオマーカーと組み合わせて、MVM群を産出した。14MVM群を、SSMを用いて、バイオメトリックパラメーター喫煙歴(パックピーク)あり、またはなしで評価した。もう一度、5以上の頻度を有する強固なマーカーを、さらなる考慮のために選択した(結果は示していない)。以上の表中で分かるように、パックイヤー(喫煙歴)は、強固なマーカーとして出現するバイオマーカーの数および型において効果を有する。このことは、バイオマーカーが他のバイオマーカーにおいて、相乗的または欠損効果を持ち得るので、全く予想されなかつことではない。本発明の1つの態様は、モデルの予想能力を改善することにおいて、パネルとして一緒に機能するマーカーを発見することに関与する。同様の筋にそって、両方の方法(AUCおよびMVM)において、バイオメトリックパラメーターパックイヤーと相乗的に機能すると同定されたバイオマーカーを、なんとか統合して、マーカーのよりよいパネルを同定した(実施例8Dを参照のこと)。

10

20

## 【0306】

大コホートに対して同定した多変数マーカーを、SSMで評価した。再度、5以上の頻度を有するマーカーのみを、さらなる考慮のために選択した。以下の表21は、大コホートに対するSSM結果を要約している。

## 【0307】

## 【表27】

トレーニング組番号 pub 3743 pub 4861 pub 8606 Pub 17338 pub 17858 acn 6399 tfa 2331

1	x	x	x	x	x	x	x	
2	x	x	x	x	x	x		x
3	x	x	x	x			x	
4	x	x		x	x	x	x	
5	x	x	x	x				x
6	x	x	x		x	x	x	
7	x	x	x	x	x	x	x	
8	x	x	x	x				
9	x	x	x	x		x	x	
10	x		x	x	x	x	x	
頻度	10	9	9	8	6	6	5	

30

表21. 11MVMマーカーに対するマーカー頻度を示す大コホートのSSM結果の部分表。

40

注: 7つのマーカーは5以上の頻度を有する。

## 【0308】

## D. 統合マーカーのSSM(AUC + MVM + パックイヤー)

続く段階において、(10トレーニング組中)5以上の頻度を有するすべてのマーカー(バイオマーカーおよびバイオメトリックパラメーター)を統合して、両コホートに対する、AUCおよびMVM群両方からのマーカーを含むマーカーの第二リストを産出した。SSM結果より、5以上の頻度を有する、小コホートから16の固有マーカー、大コホートから15の固有マーカーを選択した。以下の表22は、選択したマーカーを要約してい

50

る。

【0309】

【表28】

小コホート				大コホート			
AUC	マーカー	16AUC	14MVM	AUC	マーカー	15AUC	11MVM
1	0.77	Pub11597	x	1	0.813	Pub17338	x x
2	0.76	Acn9459	x x	2	0.812	pub17858	x x
3	0.75	Pub4861	x x	3	0.798	pub8606	x x
4	0.74	pkyrs	x x	4	0.796	pub8662	x
5	0.72	Pub2433	x	5	0.773	pub4628	x
6	0.72	CK 19	x	6	0.765	pub6798	x
7	0.72	Pub4789	x	7	0.76	pub4750	x
8	0.71	Tfa6652	x	8	0.751	pub11597	x
9	0.66	cea	x	9	0.747	pub4487	x
10	0.64	Pub2951	x	10	0.744	tfa6453	x
11	0.63	Pub6052	x	11	0.741	hic3959	x
12	0.6	Tfa2759	x	12	0.72	pub4861	x
13	0.6	Tfa9133	x	13	0.69	pub3743	x
14	0.59	Acn4132	x	14	0.67	acn6399	x
15	0.58	Acn6592	x	15	0.66	tfa2331	x
16	0.57	Pub7775	x	合計		11	7
合計							
		8	11				

表22. AUCおよびMVM群からのマーカー組合せ

【0310】

30

マーカーの以上のリストを、SSMでの最終評価周期を通して得た。先に言及したように、マーカーの組合せを、10トレーニングサブセットに対して最適化し、各バイオマーカーおよびバイオメトリックパラメーターの頻度を決定した。マーカーがトレーニング組の少なくとも50%で存在するという選択基準を適用することによって、小コホートに対して16マーカーのうち13個を選択し、大コホートに対して15マーカーのうち9個を選択した。

【0311】

【表29】

小コホート			大コホート		
	AUC	マーカー		AUC	マーカー
1	0.718	CK 19	9	1	0.67
2	0.761	acn9459	8	2	0.69
3	0.74	pkrys	8	3	0.798
4	0.664	cea	8	4	0.751
5	0.603	tfa2759	8	5	0.744
6	0.766	pub11597	7	6	0.747
7	0.718	pub4789	7	7	0.72
8	0.6	tfa9133	7	8	0.765
9	0.75	pub4861	6	9	0.741
11	0.719	pub2433	6		
10	0.589	acn4132	6		
12	0.57	Pub7775	6		
13	0.635	pub2951	5		

10

20

表23a. 5以上の頻度を有するマーカー

## 【0312】

各マーカーに対して、マーカーのレベルを最適化した時、スプリット点を、各トレーニングデータ組を、分類における最も高い精度に関して評価することによって決定した。小コホートにて使用した8つの最も頻度の高いマーカーに対するスプリット点を以下に列記している。

## 【0313】

## 【表30】

30

	マーカー	対照群	平均値	標準偏差
1	CK 19	Norm <= SP	1.89	0.45
2	acn9459	Norm >= SP	287.3	23.67
3	pkrys	Norm <= SP	30.64	4.21
4	cea	Norm <= SP	4.82	0
5	tfa2759	Norm >= SP	575.6	109.7
6	pub11597	Norm <= SP	34.4	2.52
7	pub4789	Norm <= SP	193.5	18.43
8	tfa9133	Norm >= SP	203.6	46.38

表23b.

40

## 【0314】

表23bは、これらの平均（Ave）スプリット点（それぞれ所定のカットオフ）を有する8つの最も頻度の高いマーカーのリストを示している。各スプリット点に対する標準偏差もまた含む（Std dev）。スプリット点に関する対照群の位置は、左から第二カラムにて与えられる。例として、サイトケラチン19において、正常群または対照群（非がん）は、1.89以下のスプリット点値を有する。

## 【実施例9】

## 【0315】

## 予測モデルの検証

50

小コホートに対する 13 バイオマーカーおよびバイオメトリックパラメーターのリストのサブセット（以上表 23 a を参照のこと）が、良好な臨床利用性を提供する。たとえば、分割スコア法におけるパネルとして一緒に利用する 8 つの最も頻度の高いバイオマーカーおよびバイオメトリックパラメーターは、試験サブセット 1 に対して、AUC 0.90 を有する（以上表 23 b を参照のこと）。

【0316】

7 - マーカーパネル（マーカー 1 から 7、表 23 b）および 8 - マーカーパネル（マーカー 1 から 8、表 23 b）を含む予測モデルを、10 ランダム試験組を用いて検証した。以下表 24 a および 24 b は、2 つのモデルの結果を要約している。すべての条件および計算パラメーターは、各モデルにおけるマーカーの数を除いて、両方の場合で同一であった。

【0317】

【表 31】

試験組番号	AUC	精度 (%)	感度 (%)	特異性 (%)	マーカー数	閾値
1	0.91	85	80.7	90.7	7	3
2	0.92	85	78.2	93.3	7	3
3	0.89	80	78.8	82.4	7	3
4	0.89	82	78.0	86.0	7	3
5	0.90	85	78.7	90.6	7	3
6	0.89	83	76.9	89.6	7	3
7	0.92	86	78.4	93.9	7	3
8	0.89	83	79.6	87.0	7	3
9	0.91	84	79.6	89.1	7	3
10	0.92	86	81.8	91.1	7	3
平均値	<b>0.90</b>	<b>83.9</b>	<b>79.1</b>	<b>89.4</b>		
標準偏差	0.01	1.9	1.4	3.5		

表 24a.

10

20

30

【0318】

表 24 a は、10 タンデム試験組を含む 7 - マーカーパネルの臨床性能を示している。計算にて使用した 7 つのマーカーおよび平均スプリット点を表 16 b にて示した。3 の閾値を、疾患群を非疾患群から分けるために使用した。本モデルに対する平均 AUC は 0.90 であり、平均精度 83.9 %、およびそれぞれ 79.1 % および 89.4 % の感度および特異性に相当する。

【0319】

【表32】

試験組番号	AUC	精度 (%)	感度 (%)	特異性 (%)	マーカー数	閾値
1	0.90	81	91.2	67.4	8	3
2	0.91	86	92.7	77.8	8	3
3	0.89	83	90.9	67.6	8	3
4	0.89	83	90.0	76.0	8	3
5	0.91	83	91.5	75.5	8	3
6	0.90	83	88.5	77.1	8	3
7	0.92	88	92.2	83.7	8	3
8	0.90	85	92.6	76.1	8	3
9	0.93	84	92.6	73.9	8	3
10	0.92	85	92.7	75.6	8	3
平均値	<b>0.91</b>	<b>84.1</b>	<b>91.5</b>	<b>75.1</b>		
標準偏差	0.01	1.8	1.4	4.7		

表24b.

## 【0320】

表24bは、10ランダム試験組を含む8-マーカーパネルの臨床性能を示している。計算にて使用した8つのマーカーおよび平均スプリット点を表16bにて示した。3の閾値（所定の総スコア）を、疾患群を非疾患群から分けるために使用した。本モデルに対する平均AUCは0.91であり、平均精度84.1%、およびそれぞれ91.5%および71.5%の感度および特異性に相当する。

20

## 【0321】

表24aおよび24bの比較は、両モデルが、AUCおよび精度に関して同程度であり、感度および特異性においてのみ異なる。表24aにおいて見ることができるように、7-マーカーパネルは、より高い特性を示している（89.4%対75.1%）。反対に、8-マーカーパネルは、これらの平均値（Ave）から判断できるように、よりよい感度を示している（91.5%対79.1%）。分類の精度を最大化する閾値（または所定の総スコア）を選択し、ROC曲線のAUCを最大化することに類似することに注意すべきである。したがって、3の選択した閾値（所定の総スコア）は、精度を最大化するだけでなく、モデルの感度および特異性間の最もよい比較を提供した。実際、このことが何を意味するかは、正常の個体が、前記個体が本モデルにおける7つの可能性のあるマーカーのうち3つ以下（または第二モデルに関して、8のうち3つ以下）に対して陽性であると試験される場合、肺がんを発症させる低「危険性」であると考えられる。設定閾値（または所定の総スコア）より大きなスコア（総スコア）を有する個体が、より高い危険性であると考えられ、さらなる試験または追跡手順の候補となったことに注意されたい。モデルの閾値（すなわち、所定の総スコア）を、（精度を犠牲にして）このモデルの感度または特異性を最大化するために、増加または減少いずれかが可能である。この柔軟性は、異なる診断上の質問および/または危険性のあるある集団に対処するためにモデルを適合可能である、たとえば正常個体を、症候性および/または無症状個体から差別化可能であるので、都合がよい。

30

## 【0322】

種々の予測モデルを、以下表25aおよび25bにて要約している。各予測モデルに対して、10試験組を介して、これらの相当する標準偏差（角括弧中で開示された）を含む閾値（すなわち所定の総スコア）、平均AUC、精度、感度および特異性として、モデルを構成するバイオマーカーおよびバイオメトリックパラメーターを示唆する。以上で概説した8つのマーカーパネルは、混合モデル2であり、以上で概説した7つのマーカーパネルは、混合モデル3である。混合モデル1Aおよび混合モデル1Bは同一のマーカーを含

40

50

む。混合モデル 1 A と混合モデル 1 B 間の唯一の差は、閾値（すなわち、所定の総スコア）中である。同様に、混合モデル 10 A と混合モデル 10 B は同一のマーカーを含む。混合モデル 10 A と混合モデル 10 B 間の唯一の差は、閾値（すなわち、所定の総スコア）中である。

【0323】

【表33】

小コホト

マーカー	8 IA モデ ル	9 IA モデ ル	IA-p kyrs	MS モデ ル	MS pkyr s モ デル	混合 モデ ル 1A	混合 モデ ル 1B	混合 モデ ル 2	混合 モデ ル 3	混合 モデ ル 4	混合 モデ ル 5
CK 19	X	X	X			X	X	X	X		
CA 19-9	X	X	X								
CEA	X	X	X			X	X	X	X	X	X
CA 15-3	X	X	X								
CA 125	X	X	X								
SCC	X	X	X								
CK 18	X	X	X								
ProGRP	X	X	X								
パラインフルエン ザ			X	X							
Pkyrs				X	X			X	X	X	
Acn9459					X	X	X	X	X	X	X
Pub11597					X	X	X	X	X	X	X
Pub4789					X	X	X	X	X	X	X
TFA2759					X	X	X	X	X	X	X
TFA9133					X	X	X	X		X	X
pub3743											
pub8606											
pub4487											
pub4861											
pub6798											
tfa6453											
hic3959											
閾値*	1/8	4/9	4/1 0	3/5	3/6	2/7	3/7	3/8	3/7	3/7	3/6
AUC		0.73 (0.0 4)	0.80 (0.0 3)	0.83 (0.0 2)	0.86 (0.0 2)		0.87 (0.0 2)	0.91 (0.0 1)	0.90 (0.0 1)	0.89 (0.0 1)	0.86 (0.0 2)
精度		66. 0(4 .1)	70. 0(2 .4)	77. 0(3 .7)	80. 0(2 .1)		78. 8(2 .0)	84. 1(2 .0)	83. 9(1 .9)	83. 0(1 .9)	79. 4(3 .6)
感度		90. 2(3 .1)	69. 5(8 .5)	85. 0(5 .0)	63. 4(4 .6)	72. 0(3 .5)	91. 3(2 .0)	81. 6(2 .3)	91. 5(1 .4)	79. 1(1 .4)	81. 3(1 .8)
特異性		30(4. 7 .8)	62. 0(6 .9)	52. 3(3 .5)	93. 3(2 .6)	89. 0(2 .6)	42. 7(3 .6)	75. 5(3 .1)	75. 1(3 .1)	89. 4(3 .5)	84. 8(4 .7)
DFI		0.7 1	0.4 9	0.5 0	0.3 7	0.3 0	0.5 8	0.3 1	0.2 6	0.2 3	0.3 4

表 25a

\*所定の総スコア。以上の表において、「x」と「X」の間で差はない。

10

20

30

40

50

【0324】

【表34】

小コホート

マーカー	混合モデル6	混合モデル7	混合モデル8	混合モデル9	混合モデル10A	混合モデル10B
CK 19		X		X	X	X
CA 19-9						
CEA	X	X		X	X	X
CA15-3						
CA125				X	X	X
SCC				X	X	X
CK 18	X			X	X	X
ProGRP				X	X	X
パラインフルエンザ						
Pkyrs	X				X	X
Acn9459	X	X		X	X	X
Pub11597	X	X	X	X	X	X
Pub4789	X	X		X	X	X
TFA2759	X	X		X	X	X
TFA9133	X					
pub3743			X			
pub8606			X			
pub4487			X			
pub4861			X			
pub6798			X			
tfa6453			X			
hic3959			X			
閾値*	3/8	2/6	3/8	3/10	3/11	4/11
AUC	0.90 (0.01)					
精度	80.2 (1.7)					
感度	92.6 (2.0)	87.8 (2.3)	88.2 (3.3)	89.1 (3.4)	94.3 (1.2)	86.6 (4.40)
特異性	65.5 (2.7)	63.7 (4.9)	64.2 (3.7)	52.3 (3.9)	47.6 (4.9)	63.9 (4.0)
DFI	0.35	0.38	0.38	0.49	0.53	0.39

表25b

\*所定の総コア

表25aおよびb 種々の予測モデルの要約

10

20

30

40

【0325】

同様に、大コホートに対して、種々の予測モデルを、総精度、感度または特異性に対して最適化可能である。4つの可能性のあるモデルを以下表26にて要約している。

【0326】

【表35】

大コホート

マーク	MS モデル 1	MS モデル 2	MS モデル 3	MS モデル 4
acn6399	X	X	X	X
pub3743	X	X	X	X
pub8606	X	X	X	X
pub11597	X	X	X	X
tfa6453	X	X	X	X
pub4487	X	X	X	X
pub4861	X	X	X	
pub6798	X	X		
hic3959	X			
閾値*	3/9	3/8	3/7	2/6
AUC				
精度	75.7(2.6)	80.0(2.0)	84.2(1.7)	78.9(2.6)
感度	95.1(2.0)	89.7(2.6)	80.7(4.4)	88.5(4.0)
特異性	67.7(3.1)	76.0(2.2)	85.7(1.4)	74.9(2.7)
DFI	0.33	0.26	0.24	0.28

表26 4つの可能性あるモデル

\*所定の総数ア 以上の中において、「X」と「X」の間で差はない。

【0327】

同様に、サイクリンコホート（測定した抗サイクリンE 2タンパク質抗体および抗サイクリンE 2ペプチド抗体を有する個体のサブセット）に対する予測モデルを、以下表27aおよび27bにて要約している。

【0328】

【表36】

マーク	サイクリンコホート(234試料)											
	モデル A	モデル B	モデル C	モデル D	モデル E	モデル F	モデル G	モデル H	モデル I	モデル J	モデル K	
CK 19									X	X		
CA 19-9												
CEA												
CA15-3												
CA125						X	X	X	X			
SCC								X	X			
CK 18							X	X	X			
ProGRP					X	X	X	X	X			
パラインフルエ ンザ												
Pkyrs				X	X	X	X	X	X			
Acn9459												
Pub11597										X	X	
Pub4789												
TFA2759												
TFA9133												
Pub6453											X	
Pub2951											X	
Pub4861											X	
Pub2433											X	

10

20

30

40

50

Pub3743												
Pub17338												
TFA6652												
サイクリン E2-1 pep	x			x	X	x	x	x	x	x	x	
サイクリン E2 タンパク質		x										
サイクリン E2-2 pep			x									
閾値*	0/1	0/1	0/1	0/2	0/3	0/4	0/5	0/6	0/7	2/6	1/3	
精度	79.0	75.4	67.4	84.1	86.2	85.2	83.5	81.2	80.4	88.4	88.4	10
感度	61.2	44.7	31.8	93.2	87	91.8	95.3	95.3	95.5	80.0	74.1	
特異性	89.9	94.2	89.2	72.9	85.6	81.3	76.2	72.7	71.4	93.5	97.1	
DFI	0.40	0.56	0.69	0.28	0.19	0.20	0.24	0.28	0.29	0.21	0.26	

表 27a

\*所定の総スコア

【0329】

【表37】

マークー	モデル L	モデル M	モデル N	モデル O	モデル P	モデル Q	モデル R	モデル S	モデル T	モデル U	モデル V
CK 19		x	x	x	X						
CA 19-9											
CEA			x	x	X	x	x				
CA15-3											
CA125					X						
SCC					X						
CK 18					X	x					
ProGRP					X	x	x	x	x	x	
パライソルエ ンザ											
Pkyrs											
Acn9459											
Pub11597	x			x							
Pub4789											
TFA2759											
TFA9133											
Pub6453	x										
Pub2951											
Pub4861								x	x		
Pub2433								x			
Pub3743								x	x	x	
Pub17338								x	x	x	
TFA6652								x			
サイクリン E2-1 pep		x	x	x	X	x	x	x	x	x	

10

20

30

40

サイクリン E2 タンパク質	x										
サイクリン E2-2 pep											
閾値*	1/3	0/2	0/3	1/4	1/7	0/4	0/3	0/2	2/8	1/5	0/2
精度	84.4	80.3	80.8	82.6	63.8	82.1	83.0	82.1	93.8	92.9	85.2
感度	64.7	80.0	81.1	58.8	94.1	80	75.3	72.9	90.6	89.4	85.9
特異性	96.4	80.6	80.6	97.1	45.3	83.4	87.8	87.8	95.7	95	84.9
DFI	0.35	0.28	0.27	0.41	0.55	0.26	0.28	0.30	0.10	0.12	0.21

表 27b 表 27a および 27b は、サイクリンコホートに対する予測モデルを提供する。

\*所定の総スコア

【0330】

同様に、自己抗体アッセイを用いる予測モデルを、以下表 28 にて要約している。

【0331】

【表 38】

マーカー	モデル AAb1	モデル AAb2
TMP21	X	X
NPC1L1C-	X	X
CCNE2BM-E2-1	X	X
TMOD1	X	X
CAMK1	X	X
RGS1	X	X
PACSIN1	X	X
p53	X	X
RCV1	X	
MAPKAPK3	X	X
閾値*	1/10	1/9
精度	82	82.9
感度	74.7	73.5
特異性	86.4	88.4
DFI	0.29	0.29

表 28 自己 Ab アッセイを用いる予測モデル

\*所定の総スコア

【0332】

これらのモデルの 5 つを、検証コホートに対して使用した。以下表 29 は、独立コホート、小コホートおよび検証コホートに対するそれぞれの予測モデルの臨床性能を様似している。

【0333】

10

20

30

【表 3 9】

	混合モデル 7	混合モデル 1	8 1A モデル	MS モデル 5	混合モデル 9
CK 19	X	X	X		X
CEA	X	X	X		X
CA19-9			X		
CA15-3			X		
CA125			X		X
SCC			X		X
CK 18			X		X
proGRP			X		X
パラインフルエンザ					
acn9459	X	X			X
pub11597	X	X		X	X
pub4789	X	X			X
tfa2759	X	X			X
tfa9133		X			
pub3743				X	
pub8606				X	
pub4487				X	
pub4861				X	
pub6798				X	
pub6453				X	
hic3959				X	

10

閾値	2/6	2/7	1/8	3/8	3/10
小コポート					

	AUC	精度	感度	特異性	DFI
	87.8	91.3	90.2	88.2	89.1
	63.7	42.7	30.0	64.2	52.3
	0.38	0.58	0.71	0.38	0.49

20

	AUC	精度	感度	特異性	DFI
	75.6	87.2	94.2	82.5	88.4
	62.9	55.7	35.2	86.0	58.6
	0.44	0.46	0.65	0.22	0.43

30

表 29  
\*所定の総スコア 以上の中において、「x」と「X」の間で差はない。

40

【実施例 10】

【0334】

バイオマーカー同定

A. HPLC 画分化

表 22 中の MS バイオマーカー候補の同定を得るために、まずプールされた、および/または個々の血清試料を、標準のプロトコールを用いて、逆相 HPLC によって画分化する必要があった。ゲル電気泳動のため、および MS 解析のために十分な物質を得ることは、いくつかの画分化サイクルを必要とした。個々の画分を、MALDI - TOF MS に

50

よって特性化し、関心のピークを含む画分を一緒にプールし、スピードバック内で濃縮した。すべての他のバイオマーカー候補を以上で記述したように処理した。

### 【0335】

図2は、濃縮前後の推定バイオマーカー( p u b 1 1 5 9 7 )を示している。開始試料中の 11 kDa でのバイオマーカー候補が非常に低濃度であることに注意すること。濃縮の後、強度がより高いが、試料は解析のために十分な純度ではなく、関心のバイオマーカーを単離するために、 SDS - PAGE によるさらなる分離を必要とした。

### 【0336】

#### B. ゲル内消化およびLC - MS / MS 解析

濃縮後、候補バイオマーカーを含む画分を、 SDS - PAGE にかけて、候補バイオマーカーに相当する分子量を有する所望のタンパク質 / ペプチドを単離した。ゲル電気泳動( SDS - PAGE )を、製造業者( インビトロジエン社( Invitrogen, Inc. ) )によって提供された標準の方法を用いて実施した。簡単に記すと、手順には、候補バイオマーカーと公知の分子量の標準タンパク質を含む試料を、図3で示したような同一のゲル中、異なるウェル内にロードすることが含まれる。標準タンパク質の遊走距離を、「未知」試料のものと比較することによって、所望の分子量を有するバンドを同定し、ゲルより切り出した。

### 【0337】

切り出したゲルバンドを次いで、 Waters Mass Prep (商標)ステーションを用いて、自動化ゲル内トリプシン消化にかけた。続いて、消化した試料をゲルより抽出し、オンライン逆相 ESI - LC - MS / MC にかけた。産物イオンスペクトルを次いで、データベース検索のために使用した。可能である場合、同定したタンパク質は市販されており、先に記述したように、 SDS - PAGE およびゲル内消化にかけた。2つの試料間のゲル電気泳動、 MS / MS 結果およびデータベース検索における良好な一致が、バイオマーカーが正しく同定されたというさらなる証拠であった。図3にて見ることができるように、 11.5 kDa での画分化試料中の市販されているヒト血清アミロイドA ( HSA ) と推定バイオマーカー間で良好な一致が存在する。 MS / MS 解析およびデータ検索によって、両方の試料が同一のタンパク質であることが確認された。図4は、候補バイオマーカー P u b 1 1 5 9 7 の MS / MS スペクトルを示している。 b および y イオンから由來したアミノ酸配列を、各パネルのトップに注釈をつける。バイオマーカー候補を、ヒト血清アミロイドA ( HSA ) タンパク質の断片として同定した。

### 【0338】

消化の影響を受けにくい小候補バイオマーカーを、 ESI - q - TOF および / または MALDI - TOF - TOF 断片化、続いて de - novo シークエンシングおよびデータベース検索 ( BLAST ) にかけて、配列情報とタンパク質 ID を得た。

### 【0339】

#### C. データベース検索およびタンパク質 ID

バイオマーカー候補を完全に特性化するために、そこから由来するタンパク質を同定することが必須であった。未知タンパク質の同定は、ゲル内消化と、続くトリプシン断片のタンデム質量分析が関与した。 MS / MS 処理の結果である産物イオンを、 Swiss - Prot タンパク質データベースに対して検索して、供給源タンパク質を同定した。低分子量を有するバイオマーカー候補に対して、タンデム質量分析と、続く de - novo シークエンシング、およびデータベース検索が、供給源タンパク質を同定するための選択の方法であった。検索によって、ホモサピエンス( Homo sapiens )ゲノムのみを考慮し、前駆体イオンに対して  $\pm 1.2 \text{ Da}$  の質量精度、および産物イオンに対して  $\pm 0.8 \text{ Da}$  の質量精度が考慮された ( MS / MS )。1つの見逃された開裂のみがトリプシンに対して許可された。2つの可変変容のみが、データベース検索に対して許容され、カルバミドメチル化 ( C ) および酸化 ( M ) であった。 Masscot 検索エンジン結果および関連 MS および MS / MS スペクトルの手動解釈を調整した後、最終タンパク質 ID を帰した。結果の精度を、反復測定によって検証した。

10

20

30

40

50

【0340】

【表40】

候補マーク 号	受入番 号	タンパク質名 ンパク質 A	観察されたペプチド配列 IGSDKYFHAGNYDAAKRGPGGAWAAEVI SDARENQQRFFGHGAEDSLADQAANEWG RSGKDPNHERPAGLPEKY (配列番 号:7)	平均 MW (Da)	
Pub11597	Q6FG67	ヒトミロトータンパク質 A	SFFSFLGEAFDGDARMWRAYSMDMREANY IGSDKYFHAGNYDAAKRGPGGAWAAEVI SDARENQQRFFGHGAEDSLADQAANEWG RSGKDPNHERPAGLPEKY (配列番 号:7)	11526.51	
ACN9459	P02656	ApoCIII <sub>1</sub>	SEAEDASLLSFMQGYMKHATKTAKDALS SVQESQVAQQQARGWVTDGFSSLKDYWST VKDKFSEFWDLDPEVRP *(T) SAVAA (配列番号:8) *(グリコシル化部位)	9421.22	10
TFA9133	P02656	ApoCIII <sub>1</sub>	シアル酸の過失後 ApoCIII <sub>1</sub>	9129.95	
Pub4789	P01009	アルファ-1 抗トリプシン	LEAIPMSIPPEVKFN *(E)PFVFLMIDQ NTKSPLFMGKVNNPTQK (配列番 号:8) *(K から E へ置換可能)	4776.69	
TFA2759	Q56G89	ヒトアルブミンペプチド	DAHKSEVAHRFKDLGEENFKALVL (配 列番号:10)	2754.10	20

表 30

【0341】

上記表30は、これらのタンパク質IDを有する、種々の候補バイオマーカーのタンパク質源を与える。マーカーを、ゲル内消化およびLC-MS/MSおよび/またはde-novoシークエンシングによって同定した。観察された断片のアミノ酸配列のみが示されており、平均MWには、示唆した場所でPTMを含むことに注意すること。アクセシヨン番号を、Swiss-Protデータベースから得、参照のみとして与えた。ACN9459およびTFA9133は、後者が糖付加部分からシアル酸を欠く(-291.3Da)ことを除いて同一のタンパク質断片であることに注意することが興味深い。ACN9459およびTFA9133両方が、アボリポタンパク質CIIIのバリエントとして同定された。本発明者らの発見は、本タンパク質の発行されている公知の配列および分子量(Bondarenko et al., J. Lipid Research, 40: 543-555 (1999))と一致している。Pub4789は、アルファ-1-抗トリプシンタンパク質として同定された。産物イオンスペクトルの近接実験が、表30で示唆した部位で、KからEへの置換が存在し得ることを示唆している。質量精度における不確実性が課題を不可能にした。

【実施例11】

【0342】

肺がんの検出

A. ペプチドまたはタンパク質の免疫アッセイ：上記実施例9にて記述したバイオマーカーを、免疫アッセイ技術によって検出および測定可能である。たとえば、アボットダイアグノスティックス(ABBOTT Diagnostics)からのArchitect(商標)免疫アッセイ系を、本発明のバイオマーカーを含むことが予想される試料中、未知の自己アッセイのために使用する。当分野で公知のように、本システムは、抗体でコートした磁気マイクロ粒子を利用し、関心のバイオマーカーに結合可能である。器具制御下、試料の分液を、当量の抗体-コート磁気マイクロ粒子と混合し、緩衝液、塩、界面活性剤および可溶性タンパク質を含む、標本希釈液の容量の2倍と混合した。インキュベーション後、マイクロ粒子を、緩衝液、塩、界面活性剤および保存剤を含む洗浄緩衝液で洗浄した。アクリジニウム標識共役物の分液を、当量の標本希釈液にそって加え、粒子を

30

40

50

再分散させた。混合液をインキュベートし、洗浄緩衝液で洗浄した。洗浄した粒子を、硝酸および過酸化水素を含む酸性プレトリガー中に再分散させて、アクリジニウム共役物をマイクロ粒子から分離した。NaOHの溶液をついで加えて、化学発光反応をトリガーした。光を、光電子増倍管によって測定し、未知の結果は、標準曲線を構築するために使用した公知の量のバイオマーカーペプチドを含む一連の試料によって放射される光との比較によって定量する。次いで標準曲線を使用して、同一の様式で処理した臨床試料中のバイオマーカーの濃度を推測する。結果は、それ自身によって、または以下で記述したような他のマーカーとの組合せで、使用可能である。

#### 【0343】

B. ペプチドまたはタンパク質の多重免疫アッセイ：単一試料から本発明の多重バイオマーカーを検出する必要がある場合、多重アッセイを実施することがより経済的であり、簡便である。問題になっている各解析物に対して、一組の特異的抗体が必要であり、Luminex 100（商標）解析器上での利用のために、独自に染色したマイクロ粒子が必要である。対の各捕獲抗体は個々に固有マイクロ粒子上にコートする。対の他の抗体を、r Phycoerythrinのような蛍光団に共役させる。マイクロ粒子をプールし、約0.01%w/cに相当する、約1000固有粒子/マイクロタイターの濃度まで希釈する。希釈液には、緩衝液、塩および界面活性剤が含まれる。10マーカーがパネル中である場合、総固体は、約10,000粒子/マイクロタイター、または0.1%固体w/vである。共役物をプールし、マイクロ粒子希釈液中各最終濃度約1から10nMに調整した。アッセイを実施するために、1以上の解析物が含まれることが予想される試料の分液を、インキュベーションウェル中に配置し、続いて、プールしたマイクロ粒子の半量を配置した。懸濁液を30分間インキュベートし、プールした共役物溶液の半量を添加した。30分間のさらなるインキュベーションの後、反応液を、2容量の塩および界面活性剤を含む緩衝化溶液の添加によって希釈した。懸濁液を混合し、試料のおよそ二倍の容量をLuminex 100（商標）器具によって解析のために吸引した。場合により、マイクロ粒子を各インキュベーション後に洗浄し、次いで解析のために再懸濁させることができる。各個々の粒子の蛍光を、3波長にて測定し、2つを粒子とこの関連解析物を同定するために使用し、3つ目を、粒子に結合した解析物の量を定量するために使用した。各型の少なくとも100マイクロ粒子を測定し、各解析物に対する中央値蛍光を計算した。試料中の解析物の量を、公知の量のペプチドまたはタンパク質を含む一連の試料において同様の解析を実施し、公知の濃度に対して公知の試料の中央値蛍光をプロットすることによって産出した標準曲線に対して比較することによって計算する。未知試料を、実施例7でのように、分割スコア法（Split and Score Method）または分割加重スコア法（Split and Weighted Score Method）のようなモデルを用いる、公知がんまたは非がん標本に関する、解析物（上昇したかまたは弱まったか）の濃度に基づいて、がんであるかまたは非がんであるか分類する。

#### 【0344】

たとえば、患者を、表18の8免疫アッセイ（IA）パネルおよび分割スコア法を用いて、肺がんを有する患者の可能性を決定するために試験してよい。患者より試験試料を得た後、患者の試験試料（すなわち血清）中の8バイオマーカーのそれぞれの量を定量し、バイオマーカーそれぞれの量を次いで、表18にて列記されたもののような、バイオマーカーに対する相当する所定のスプリット点（所定のカットオフ）と比較する（すなわちサイトケラチン19に対して使用可能な所定のカットオフは1.89または2.9である）。この相当する所定のスプリット点（所定のカットオフ）より大きい量を有する各バイオマーカーに対して、スコア1が与えられ得る。この相当する所定のスプリット点（所定のカットオフ）以下である量を有する各バイオマーカーに対して、スコア0が与えられ得る。8バイオマーカーのそれぞれに対するスコアを次いで数学的に（すなわちバイオマーカーのスコアのそれぞれを一緒に加えることによって）統合し、患者に対する総スコアに行き着く。本総スコアがパネルスコアとなる。パネルスコアを、表25aの8IAモデルの所定の閾値（所定の総スコア）、すなわち1と比較する。1より大きなパネルスコアが患

10

20

30

40

50

者に対する陽性結果である。1以下のパネルスコアが、患者に対して陰性結果である。先の集団研究において、本パネルは、特異性30%、擬陽性率70%、および感度90%を示した。患者に対する陽性パネル結果は、擬陽性になる可能性が70%ある。さらに、90%の肺がん患者が、陽性パネル結果を持ち得る。したがって、陽性パネル結果を有する患者を、肺がんの示唆または疑念に対するさらなる試験に対して参考し得る。

#### 【0345】

さらなる例によって、再び8IAパネルおよび分割加重スコア法を用いて、患者より試験試料を得たのち、患者の試験試料（すなわち血清）中の8つのバイオマーカーそれぞれの量を定量し、次いでバイオマーカーのそれぞれの量を、表17bにて列記したスプリット点のような、所定のスプリット点（所定のカットオフ）と比較する（すなわち、サイトケラチン19のために使用可能な所定のカットオフは、1.2、1.9および3.3である）。本例において、各バイオマーカーは、3つの所定のスプリット点（所定のカットオフ）である。したがって、4つの可能性のあるスコアが、各バイオマーカーに対して与えられ得る。次いで8つのバイオマーカーのそれぞれに対するスコアを、数学的に（すなわちバイオマーカーのスコアのそれぞれを一緒に加えることによって）合わせ、患者に対する総スコアに行き着く。次いで、総スコアがパネルスコアになる。パネルスコアを、11.2であると計算された、8IAモデルに対する所定の閾値（または所定の総スコア）と比較可能である。11.2より大きな患者パネルスコアが陽性結果である。11.2以下の患者パネルスコアが、陰性結果である。先の集団研究において、本パネルが、特異性34%、擬陽性率66%および感度90%を示した。陽性パネル結果が、擬陽性である可能性66%を有する。さらに90%の肺がん患者が陽性パネル結果を有する。したがって、陽性パネル結果を有する患者が、肺がんの示唆または疑念のためのさらなる試験のために参考され得る。

10

#### 【0346】

C. 免疫質量分光分析：質量分析のための試料調製はまた、免疫学的方法ならびにクロマトグラフィーまたは電位測定法も使用可能である。ペプチドバイオマーカーに対して特異的な抗体でコートされた超常磁性マイクロ粒子を、塩を含む緩衝溶液中、および0.1% w/vの濃度まで調整する。患者血清試料の分液を当量の抗体コートマイクロ粒子と混合し、2倍容量の希釈液と混合する。インキュベーションの後、マイクロ粒子を、緩衝塩および場合により塩と界面活性剤を含む洗浄緩衝液で洗浄する。マイクロ粒子を次いで脱イオン水で洗浄する。免疫精製解析物を、トリフルオロ酢酸を含む、当量の水性アセトニトリルを加えることによって、マイクロ粒子から溶出する。ついで試料を当量のシナピン酸マトリックス溶液と混合し、少量（およそ1から3マイクロリットル）を、飛行時間質量解析のために、MALDI標的に適用する。所望のm/zでのイオン電流を、同一の様式で処理した公知の量のペプチドバイオマーカーを含む試料から得たイオン電流と比較する。

20

#### 【0347】

イオン電流が、濃度と直接関連すること、および特定のm/z値（またはROI）でのイオン電流（または強度）を望むのならば濃度に変換可能であることに注意すべきである。そのような濃度または強度を次いで、実施例7で記述した任意のモデル構築アルゴリズムへの入力として使用可能である。

30

#### 【0348】

D. ROI sに対する質量分析：血液試料を患者より得、凝血塊を許容して血清試料を形成する。試料をSELDI質量分光分析のために調製し、Bioprocessor中のProtein Chip上にロードし、実施例2にて提供されたように処理する。Protein Chipを、Ciphergen 4000 MALDI飛行時間質量分析上にロードし、実施例3でのように解析する。各スペクトルを、多変数解析を用いて受入に対して試験する。たとえば、（未知試料と公知の参照集団間の）総イオン電流およびスペクトルコントラスト角を計算する。次いでマハラノビスの距離を決定する。マハラノビスの距離が確立された臨界値より小さいスペクトルに対して、スペクトルを定量する。

40

50

マハラノビスの距離が、確立された臨界値より大きなスペクトルに対して、スペクトルをさらなる解析から排除し、試料を再度実行するべきである。適格後、質量スペクトルを正規化する。

【0349】

得られる質量スペクトルを、選択したデータ解析モデルに対して適切な関心領域内のイオン電流を測定することによって評価する。解析の結果に基づいて、患者が、肺がんを有する危険性があるか、または高い可能性を有するか判断し、さらなる診断手順を通して解釈すべきである。

【0350】

分割スコア法の利用のために、表5で与えられたm/z値でのR O I sにおける強度を患者に対して測定する。患者結果を、患者値が、表6で与えられた平均スプリット点値のがん側であるか、または非がん側であるかを注意することによってスコア化する。スコア1が、スプリット点のがん側であると分かった各R O I値に与えられる、スコア3およびそれ以上が、患者ががんの危険性が上昇していることを示唆しており、さらなる診断手順が参照されるべきである。

10

【0351】

当業者は、本発明が目的を実行し、言及された、末端および有利な点ならびにここで本来そなわっているものを得るために十分適合することを簡単に理解するであろう。本明細書で記述された組成物、処方、方法、手順、処置、分子、特異的化合物が、好ましい実施形態の現在代表であり、例示であり、本発明の目的において制限である意図はない。置換および改変を、本発明の範囲および精神から逸脱することなしに、本明細書で開示された本発明に対して実施してよい。

20

【0352】

本明細書で言及されたすべての特許および発行物が、本発明が関連する技術分野の当業者のレベルの示唆である。すべての特許および発行物が、各個々の発行物が参照により組み込まれていると特に、および個々に示唆された場合と同程度に、参照により本明細書で組み込まれる。

【図面の簡単な説明】

【0353】

【図1】バイオ-インフォマティクス・ワークフローのダイアグラムである。特に、MSデータおよびIAデータを、種々の統計学的方法にかけた。ロジスティック回帰を使用して、受信者動作特性(Receiver Operator Characteristic (ROC))曲線を産出し、各マーカーに関して、曲線化面積(Area Under the Curve (AUC))を得た。最も高いAUCを持つトップのマーカーを候補マーカーとして選択した。判別分析(Discriminant Analysis (DA))、主成分分析(Principal Component Analysis (PCA))、および決定木解析(Decision Trees (DT))のような多重変数解析が、モデルへの入力に関して、さらなるマーカーを同定した。バイオメトリックパラメーターをまた含めることができる。少なくとも50%のトレーニングセット中で発生する強固なマーカーを、分割スコア(Split and Score)法/アルゴリズム(SSM)によって同定し、推定バイオマーカーとして選別する。工程を、最終予測モデルのために公的な数のマーカーが得られるまで、n回繰り返す。

30

【図2】a) プールしたHPLC画分を濃縮した後、およびb) 濃縮工程前の、Pub 11597バイオマーカー候補を示しているMALDI-TOF MSプロファイルである。試料は、HPLC画分化後でさえ、複合体混合物である。

40

【図3】ゲルにのせた種々の試料の成分を示している染色ゲルである。レーンa、fおよびgは、較正目的のための、公知の分子量の標準タンパク質の混合物を示している。さらに、レーンbおよびeは、市販されている、ヒト血清アミロイドA(HSAA)として知られる推定されるタンパク質の非所運異性制された形態を示している。レーンcおよびdは、推定バイオマーカーを含む画分化試料を示している。HSAA標準と同一の距離を遊

50

走する混合液中の成分が存在する。H S A A と同一の誘導距離を持つバンドをゲルより切り出し、ゲル内消化およびM S / M S 解析にかけて、その同一性を確認した。

【図4 a】P u b 1 1 5 9 7 のトリプシン消化のL C - M S / M S である。パネルaからdは、4つの主要な前駆体イオンのM S / M S を示している。bおよびy 産物イオンは、注釈が施されており、誘導アミノ酸配列が、4つの前駆体イオンのそれぞれに与えられる。産出されたbおよびy イオンの分子量を用いたデータベース検索により、H S A A として供給タンパク質が同定された。観察された断片 (M W = 1 1 5 2 6 . 5 1 ) の完全配列が、配列番号6にて提供されている。

【図4 b】P u b 1 1 5 9 7 のトリプシン消化のL C - M S / M S である。パネルaからdは、4つの主要な前駆体イオンのM S / M S を示している。bおよびy 産物イオンは、注釈が施されており、誘導アミノ酸配列が、4つの前駆体イオンのそれぞれに与えられる。産出されたbおよびy イオンの分子量を用いたデータベース検索により、H S A A として供給タンパク質が同定された。観察された断片 (M W = 1 1 5 2 6 . 5 1 ) の完全配列が、配列番号6にて提供されている。10

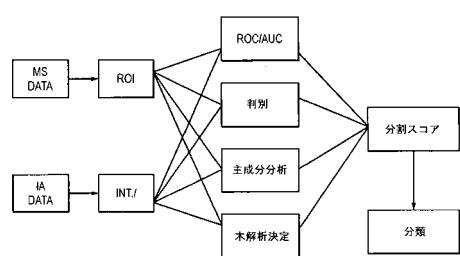
【図4 c】P u b 1 1 5 9 7 のトリプシン消化のL C - M S / M S である。パネルaからdは、4つの主要な前駆体イオンのM S / M S を示している。bおよびy 産物イオンは、注釈が施されており、誘導アミノ酸配列が、4つの前駆体イオンのそれぞれに与えられる。産出されたbおよびy イオンの分子量を用いたデータベース検索により、H S A A として供給タンパク質が同定された。観察された断片 (M W = 1 1 5 2 6 . 5 1 ) の完全配列が、配列番号6にて提供されている。20

【図4 d】P u b 1 1 5 9 7 のトリプシン消化のL C - M S / M S である。パネルaからdは、4つの主要な前駆体イオンのM S / M S を示している。bおよびy 産物イオンは、注釈が施されており、誘導アミノ酸配列が、4つの前駆体イオンのそれぞれに与えられる。産出されたbおよびy イオンの分子量を用いたデータベース検索により、H S A A として供給タンパク質が同定された。観察された断片 (M W = 1 1 5 2 6 . 5 1 ) の完全配列が、配列番号6にて提供されている。

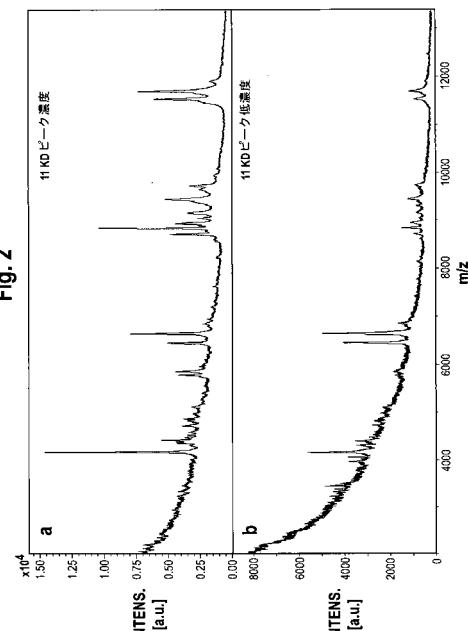
【図5】実施例1にて記述された751の患者試料にて実施された、8つの免疫アッセイバイオマーカーパネルから産出されたR O C 曲線を与える。黒ダイアモンドは、分割加重スコア法 (S p l i t a n d W e i g h t e d S c o r i n g M e t h o d ) を用いて総スコアより産出されたR O C 曲線を表している。四角形は、大コホートスプリット点を用いる二元スコアリング法を用いて総スコアより産出されたR O C 曲線を表している。三角形は、小コホートスプリット点を用いる二元スコアリング法を用いて総スコアより産出されたR O C 曲線を表している。30

【図1】

Fig. 1



【図2】



【図3】

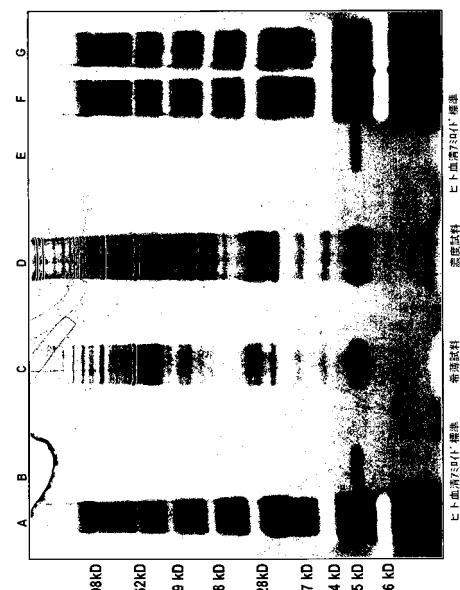


Fig. 3

【図4 a】

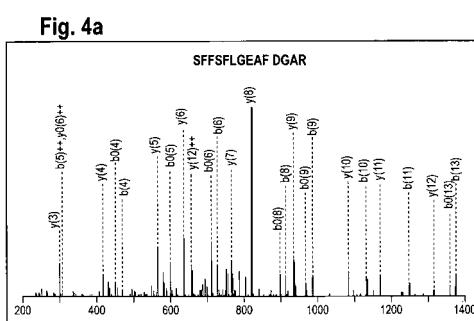


Fig. 4a

【図4 b】

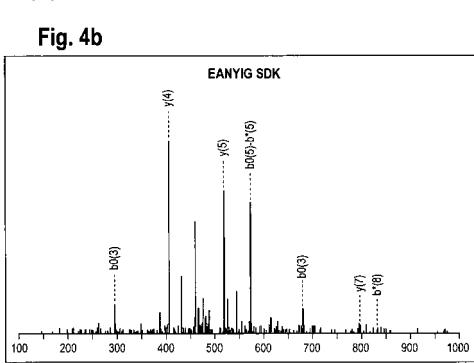
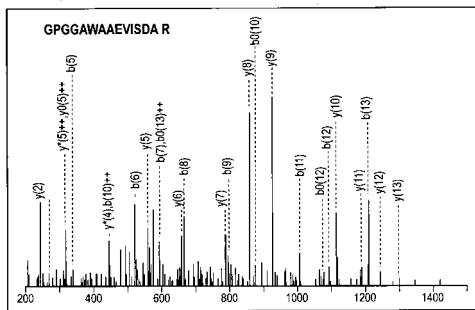


Fig. 4b

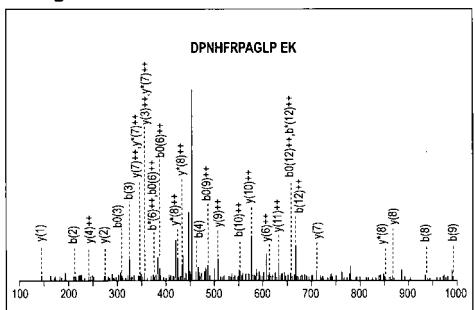
【図4c】

Fig. 4c



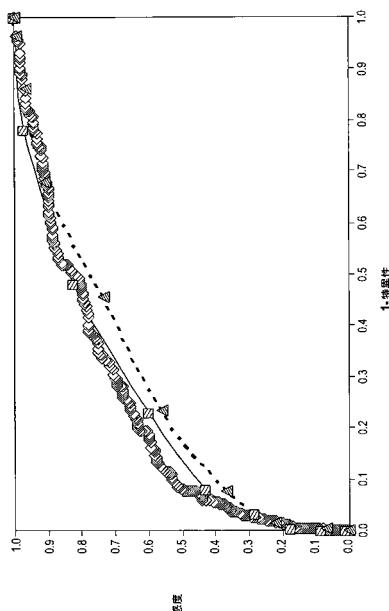
【図4d】

Fig. 4d



【図5】

Fig. 5



【配列表】

[0005160447000001.app](#)

---

フロントページの続き

(51)Int.Cl.

F I

G 0 1 N 33/574

B

(74)代理人 100103920

弁理士 大崎 勝真

(74)代理人 100124855

弁理士 坪倉 道明

(72)発明者 ラツセル, ジヨン

アメリカ合衆国、ウイスコンシン・53142、ケノーシヤ、シックステイス・アベニュー・80  
11

(72)発明者 コルピツツ, トレイシー

アメリカ合衆国、イリノイ・60073、ラウンド・レイク、ノース・サークル・ドライブ・34  
365

(72)発明者 ラツセル, エリツク

アメリカ合衆国、イリノイ・60031、ガーニー、アツブル・レイン・16730

(72)発明者 フロスト, ステイーブン

アメリカ合衆国、イリノイ・60031、ガーニー、ロング・ヒル・ロード・431

(72)発明者 ラミレス, ハビア

アメリカ合衆国、イリノイ・60031、ガーニー、グレイズテム・サークル・2155

(72)発明者 シン, パワーニ

アメリカ合衆国、イリノイ・60091、ウイルメット、リンデン・アベニュー・420

審査官 赤坂 祐樹

(56)参考文献 特表平07-505719 (JP, A)

特表2003-521669 (JP, A)

国際公開第00/058728 (WO, A1)

(58)調査した分野(Int.Cl., DB名)

G01N 33/53-33/574

专利名称(译)	用于筛选肺癌趋势的方法和标记物组合		
公开(公告)号	<a href="#">JP5160447B2</a>	公开(公告)日	2013-03-13
申请号	JP2008547772	申请日	2006-12-21
[标]申请(专利权)人(译)	雅培公司		
申请(专利权)人(译)	雅培制药		
当前申请(专利权)人(译)	雅培制药		
[标]发明人	ラツセルジョン コルピツツトレイシー ラツセルエリツク フロストステイブン ラミレスハビアー シンバワーニ		
发明人	ラツセル,ジョン コルピツツ,トレイシー ラツセル,エリツク フロスト,ステイブン ラミレス,ハビアー シン,バワーニ		
IPC分类号	G01N33/574 C07K16/18 C07K14/47 G01N33/53		
CPC分类号	G01N33/57423 Y02A90/26		
FI分类号	G01N33/574.ZNA.A C07K16/18 C07K14/47 G01N33/53.N G01N33/574.E G01N33/574.B		
代理人(译)	小野 诚 金山 贤教 Masarushin大崎		
优先权	60/753331 2005-12-22 US		
其他公开文献	JP2009521692A		
外部链接	<a href="#">Espacenet</a>		

## 摘要(译)

本发明涉及某些免疫反应性多肽，辅助受试者肺癌诊断的方法和进行所述方法的试剂盒。

$$\cos(\theta) = \frac{V_1 \cdot V_2}{\sqrt{V_1 \cdot V_1} \sqrt{V_2 \cdot V_2}}$$