

(19) 日本国特許庁(JP)

(12) 公表特許公報(A)

(11) 特許出願公表番号

特表2006-509186

(P2006-509186A)

(43) 公表日 平成18年3月16日(2006.3.16)

(51) Int. Cl.	F I	テーマコード (参考)
GO 1 N 33/574 (2006.01)	GO 1 N 33/574 A	2GO41
GO 1 N 27/62 (2006.01)	GO 1 N 27/62 F	
GO 1 N 27/64 (2006.01)	GO 1 N 27/62 V	
GO 1 N 30/00 (2006.01)	GO 1 N 27/64 B	
GO 1 N 30/26 (2006.01)	GO 1 N 30/00 E	

審査請求 未請求 予備審査請求 未請求 (全 55 頁) 最終頁に続く

(21) 出願番号	特願2004-541437 (P2004-541437)	(71) 出願人	504414086
(86) (22) 出願日	平成15年5月9日 (2003.5.9)		イースタン バージニア メディカル ス
(85) 翻訳文提出日	平成17年1月5日 (2005.1.5)		クル
(86) 国際出願番号	PCT/US2003/014432		アメリカ合衆国 バージニア 23501
(87) 国際公開番号	W02004/030511		-6006, ノーフォーク, フェアフ
(87) 国際公開日	平成16年4月15日 (2004.4.15)		ァックス アベニュー 825, ピー.
(31) 優先権主張番号	60/379,018		オー. ボックス 1980
(32) 優先日	平成14年5月10日 (2002.5.10)	(71) 出願人	398048028
(33) 優先権主張国	米国 (US)		フレッド ハッチンソン キャンサー リ
			サーチ センター
			アメリカ合衆国, 98109 ワシントン
			, シアトル, シー2エム-027, フェア
			ビュー アベニュー ノース 1100
		(74) 代理人	100078282
			弁理士 山本 秀策

最終頁に続く

(54) 【発明の名称】 前立腺癌バイオマーカー

(57) 【要約】

前立腺癌、良性前立腺肥大を診断する際に、または陰性診断を下す際に、有利に利用される可能性があるタンパク質バイオマーカーについて説明している。したがって、本発明の一面では、前立腺癌または良性前立腺肥大の診断を助ける、さもなければ該診断を下す方法を提供する。本発明の一形態では、所定の分子量からなる種々のタンパク質バイオマーカーを検出し、該検出を前立腺癌、良性前立腺肥大の診断と、または陰性診断と関連させることを含む。本発明の別の局面では、本明細書に記載したバイオマーカーの検出に利用可能なキットを提供する。本発明のさらなるものでは、複数の分類子を用いて、良性前立腺肥大の前立腺癌の可能性があると診断を下す方法を提供する。本発明の特定の形態では、方法は促進決定木分析の使用を含む。種々のコンピュータ可読媒体も提供する。

【特許請求の範囲】

【請求項 1】

良性前立腺肥大または前立腺癌の診断を助ける方法であって、

(a) 被験体から得た試験試料中の少なくとも 2 種類のタンパク質バイオマーカーを検出することと、

(b) 良性前立腺肥大、前立腺癌の疑診もしくは陰性診断と前記検出を相関させることとを含み、

前記タンパク質バイオマーカーが約 4475 ± 81 、約 5074 ± 91 、約 5382 ± 97 、約 7024 ± 13 、約 7820 ± 14 、約 8141 ± 15 、約 9149 ± 16 、約 9508 ± 17 、および約 9656 ± 17 ダルトンからなる群から選択される分子量を有することを特徴とする、方法。

10

【請求項 2】

前記検出ステップが前記バイオマーカーの差次的発現を同定することをさらに含むことを特徴とする、請求項 1 に記載の方法。

【請求項 3】

前記少なくとも 2 種類のタンパク質バイオマーカーが前記約 7820 ± 14 ダルトンおよび 7024 ± 13 ダルトン・タンパク質バイオマーカーであることを特徴とする、請求項 1 に記載の方法。

【請求項 4】

前記少なくとも 2 種類のタンパク質バイオマーカーが前記約 7820 ± 14 ダルトン、約 7024 ± 13 ダルトン、約 5382 ± 97 ダルトン、および約 4475 ± 81 ダルトン・バイオマーカーであることを特徴とする、請求項 1 に記載の方法。

20

【請求項 5】

前記少なくとも 2 種類のタンパク質バイオマーカーが前記約 8141 ± 15 ダルトン、約 9149 ± 16 ダルトン、および約 9656 ± 17 ダルトン・バイオマーカーであることを特徴とする、請求項 1 に記載の方法。

【請求項 6】

前記少なくとも 2 種類のタンパク質バイオマーカーが前記約 5074 ± 91 ダルトン、約 9149 ± 16 ダルトン、および約 9656 ± 17 ダルトン・バイオマーカーであることを特徴とする、請求項 1 に記載の方法。

30

【請求項 7】

前記少なくとも 2 種類のタンパク質バイオマーカーが前記約 9149 ± 16 ダルトンおよび約 9508 ± 17 ダルトン・バイオマーカーであることを特徴とする、請求項 1 に記載の方法。

【請求項 8】

前記少なくとも 2 種類のタンパク質バイオマーカーが前記約 5382 ± 97 ダルトン、約 7024 ± 13 ダルトン、および約 7820 ± 14 ダルトン・バイオマーカーであることを特徴とする、請求項 1 に記載の方法。

【請求項 9】

前記少なくとも 2 種類のタンパク質バイオマーカーが前記約 7024 ± 13 ダルトン、約 7820 ± 14 ダルトン、約 5382 ± 97 ダルトン、および約 4475 ± 81 ダルトン・バイオマーカーであることを特徴とする、請求項 1 に記載の方法。

40

【請求項 10】

前記方法が前記約 7820 ± 14 ダルトン・タンパク質バイオマーカーの非存在を測定することと、前記約 7024 ± 13 ダルトン・タンパク質バイオマーカーの存在を測定することを含むことを特徴とする、請求項 3 に記載の方法。

【請求項 11】

前記方法が前記約 7820 ± 14 ダルトンおよび約 7024 ± 13 ダルトン・バイオマーカーの非存在、前記約 5382 ± 97 ダルトン・バイオマーカーの存在、ならびに前記約 4475 ± 81 ダルトン・バイオマーカーの量を測定することを含むことを特徴とする、

50

請求項 4 に記載の方法。

【請求項 1 2】

前記方法が前記約 8 1 4 1 ± 1 5 ダルトン、約 9 1 4 9 ± 1 6 ダルトン、および 9 6 5 6 ± 1 7 ダルトン・バイオマーカの量を測定することを含むことを特徴とする、請求項 5 に記載の方法。

【請求項 1 3】

前記方法が前記約 5 0 7 4 ± 9 1 ダルトン、約 9 1 4 9 ± 1 6 ダルトン、および約 9 6 5 6 ± 1 7 ダルトン・バイオマーカの量を測定することを含むことを特徴とする、請求項 6 に記載の方法。

【請求項 1 4】

前記方法が前記約 9 1 4 9 ± 1 6 ダルトンおよび約 9 5 0 8 ± 1 7 ダルトン・バイオマーカの量を測定することを含むことを特徴とする、請求項 7 に記載の方法。

【請求項 1 5】

前記方法が前記約 5 3 8 2 ± 9 7 ダルトン、約 7 0 2 4 ± 1 3 ダルトン、および約 7 8 2 0 ± 1 4 ダルトン・バイオマーカの非存在を測定することを特徴とする、請求項 8 に記載の方法。

【請求項 1 6】

前記方法が前記約 7 0 2 4 ± 1 3 ダルトンおよび約 7 8 2 0 ± 1 4 ダルトン・バイオマーカの非存在、前記約 5 3 8 2 ± 9 7 ダルトン・バイオマーカの存在、ならびに前記約 4 4 7 5 ± 8 1 ダルトン・バイオマーカの量を測定することを含むことを特徴とする、請求項 9 に記載の方法。

【請求項 1 7】

前記被験体から得た試験試料中の少なくとも 2 種類のタンパク質バイオマーカを検出することが質量分析によって実行されることを特徴とする、請求項 1 に記載の方法。

【請求項 1 8】

前記質量分析がレーザー脱着質量分析であることを特徴とする、請求項 1 7 に記載の方法。

【請求項 1 9】

前記質量分析が表面活性化レーザー脱着イオン化質量分析であることを特徴とする、請求項 1 8 に記載の方法。

【請求項 2 0】

前記表面活性化レーザー脱着イオン化質量分析が

(a) 基質に結合する吸着剤を含む前記基質を提供することと、

(b) 前記吸着剤と前記試験試料を接触させることと、

(c) 前記基質から前記バイオマーカを脱離およびイオン化させることと、

(d) 質量分析計を用いて前記脱着 / イオン化バイオマーカを検出することとを含むことを特徴とする、請求項 1 9 に記載の方法。

【請求項 2 1】

前記吸着剤と前記試験試料を接触させる前に、前記試験試料を精製することをさらに含むことを特徴とする、請求項 2 0 に記載の方法。

【請求項 2 2】

前記被験体から得た試験試料中の少なくとも 2 種類のタンパク質バイオマーカを検出することが免疫アッセイによって実行されることを特徴とする、請求項 1 に記載の方法。

【請求項 2 3】

前記免疫アッセイが酵素免疫アッセイであることを特徴とする、請求項 2 2 に記載の方法。

【請求項 2 4】

前記試験試料が前記被験体から得た生物学的流体であることを特徴とする、請求項 1 に記載の方法。

【請求項 2 5】

10

20

30

40

50

前記生物学的流体が血清であることを特徴とする、請求項 24 に記載の方法。

【請求項 26】

前記試験試料が精液、精漿、唾液、血液、リンパ液、肺ノ気管支洗浄液、粘液、大便、乳頭分泌、痰、涙、または尿からなる群から選択されることを特徴とする、請求項 1 に記載の方法。

【請求項 27】

2 から 9 種類のマーカーが検出されることを特徴とする、請求項 1 に記載の方法。

【請求項 28】

前立腺癌または良性前立腺肥大を診断する方法であって、

(a) 被験体から得た試験試料中でタンパク質バイオマーカーを検出することと、

10

(b) 良性前立腺肥大、前立腺癌の診断、または陰性診断と測定を相関させることとを含み、

前記タンパク質バイオマーカーが

(i) 約 7024 ± 13 ダルトンおよび約 7820 ± 14 ダルトン、

(ii) 約 7820 ± 14 ダルトン、約 7024 ± 13 ダルトン、約 5382 ± 97 ダルトン、および約 4475 ± 81 ダルトン、

(iii) 約 8141 ± 15 ダルトン、約 9149 ± 16 ダルトン、および約 9656 ± 17 ダルトン、

(iv) 約 9149 ± 16 ダルトンおよび約 9508 ± 17 ダルトン、

(v) 約 5074 ± 91 ダルトン、約 9149 ± 16 ダルトン、および約 9656 ± 17 ダルトン、または

20

(vi) 約 5382 ± 97 ダルトン、約 7024 ± 13 ダルトン、および約 7820 ± 14 ダルトンの群にあり、かつ上記の分子量を有することを特徴とする、方法。

【請求項 29】

前記検出が質量分析によって実行されることを特徴とする、請求項 28 に記載の方法。

【請求項 30】

前記質量分析がレーザー脱着質量分析であることを特徴とする、請求項 28 に記載の方法。

【請求項 31】

前記質量分析が表面活性化レーザー脱着イオン化質量分析であることを特徴とする、請求項 29 に記載の方法。

30

【請求項 32】

前記表面活性化レーザー脱着イオン化質量分析が

(a) 基質に結合する吸着剤を含む前記基質を提供することと、

(b) 前記吸着剤と前記試験試料を接触させることと、

(c) 前記基質から前記バイオマーカーを脱着およびイオン化させることと、

(d) 質量分析計を用いて前記脱着ノイオン化バイオマーカーを検出することとを含むことを特徴とする、請求項 30 に記載の方法。

【請求項 33】

前記吸着剤と前記試料を接触させる前に、前記試料を精製することをさらに含むことを特徴とする、請求項 28 に記載の方法。

40

【請求項 34】

前記検出が免疫アッセイによって実行されることを特徴とする、請求項 28 に記載の方法。

【請求項 35】

前記免疫アッセイが酵素免疫アッセイであることを特徴とする、請求項 34 に記載の方法。

【請求項 36】

前記方法が前記分子量約 7024 ± 13 を有するタンパク質バイオマーカーの存在および前記分子量 7820 ± 14 ダルトンを有するタンパク質バイオマーカーの非存在を測定す

50

ることを含むことを特徴とする、請求項 28 に記載の方法。

【請求項 37】

前記方法が約 7820 ± 14 ダルトンおよび約 7024 ± 13 ダルトン・バイオマーカーの非存在、前記約 5382 ± 97 ダルトン・バイオマーカーの存在、ならびに前記約 4475 ± 81 ダルトン・バイオマーカーの量を測定することを含むことを特徴とする、請求項 28 に記載の方法。

【請求項 38】

前記方法が前記約 8141 ± 15 ダルトン、約 9149 ± 16 ダルトン、および約 9656 ± 17 ダルトン・バイオマーカーの量を測定することを含むことを特徴とする、請求項 28 に記載の方法。

10

【請求項 39】

前記方法が前記約 9149 ± 16 ダルトンおよび約 9508 ± 17 ダルトン・バイオマーカーの量を測定することを含む、請求項 28 に記載の方法。

【請求項 40】

前記方法が前記約 5074 ± 91 ダルトン、約 9149 ± 16 ダルトン、および約 9656 ± 17 ダルトン・バイオマーカーの量を測定することを含むことを特徴とする、請求項 28 に記載の方法。

【請求項 41】

前記方法が前記約 5382 ± 97 ダルトン、約 7024 ± 13 ダルトン、および約 7820 ± 14 ダルトン・バイオマーカーの非存在を測定することを特徴とする、請求項 28 に記載の方法。

20

【請求項 42】

前記群の 1 から 6 つまでからバイオマーカーが検出されることを特徴とする、請求項 28 に記載の方法。

【請求項 43】

良性前立腺肥大または前立腺癌を診断する方法であって、

(a) 被験体から得た試験試料中の少なくとも 2 種類のタンパク質バイオマーカーを検出することと、

(b) 良性前立腺肥大、前立腺癌の診断、または陰性診断を前記検出と関連させることとを含み、

30

前記タンパク質バイオマーカーが約 3486 ± 6、約 3963 ± 7、約 4071 ± 7、約 4079 ± 7、約 4580 ± 8、約 5298 ± 10、約 6099 ± 11、約 6542 ± 12、約 6797 ± 12、約 6949 ± 13；約 6990 ± 13、約 7024 ± 13、約 7054 ± 13、約 7820 ± 14、約 7844 ± 14、約 7885 ± 14、約 8067 ± 15、約 8356 ± 15、約 8943 ± 16、約 9656 ± 17、および約 9720 ± 18 ダルトンからなる群から選択される分子量を有することを特徴とする、方法。

【請求項 44】

前記方法が、

(a) 約 3963 ± 7、約 4079 ± 7、約 6542 ± 12、約 6797 ± 12、約 6949 ± 13、約 6990 ± 13、約 7024 ± 13、約 7885 ± 14、約 8067 ± 15、約 8356 ± 15、約 9656 ± 17、約 9720 ± 18 ダルトン、およびその組み合わせからなる群から選択される分子量を有するタンパク質バイオマーカーを検出することと、

40

(b) 前立腺癌の診断と前記検出を関連させることとを含むことを特徴とする、請求項 43 に記載の方法。

【請求項 45】

前記方法が、

約 3486 ± 6、約 4071 ± 7、約 4580 ± 8、約 5298 ± 10、約 6099 ± 11、約 7054 ± 13、約 7820 ± 14、約 7844 ± 14、約 8943 ± 16、およびその組み合わせからなる群から選択される分子量を有するタンパク質バイオマーカー

50

を検出することと、

(b) 良性前立腺肥大の診断と前記検出を相関させることとを含むことを特徴とする、請求項43に記載の方法。

【請求項46】

前記方法が2から12種類の前記タンパク質バイオマーカーから検出することを含むことを特徴とする、請求項43に記載の方法。

【請求項47】

前記方法が2から9種類の前記タンパク質バイオマーカーから検出することを含むことを特徴とする、請求項45に記載の方法。

【請求項48】

前記方法が2から21種類の前記タンパク質バイオマーカーから検出することを含むことを特徴とする、請求項43に記載の方法。

【請求項49】

キットであって、以下：

(a) 基質に結合する吸着剤を含む前記基質と、

(b) 試験試料を吸着剤と接触させて該吸着剤に保持されたタンパク質バイオマーカーを検出することでタンパク質バイオマーカーを検出するためのインストラクションとを含み、

前記吸着剤が約4475 ± 81、約5074 ± 91、約5382 ± 97、約7024 ± 13、約7820 ± 14、約8141 ± 15、約9149 ± 16、約9508 ± 17、および約9656 ± 17ダルトンからなる群から選択される少なくとも1種類のタンパク質バイオマーカーを保持することができることを特徴とする、キット。

【請求項50】

前記基質が気相イオン質量分析計での使用に適合させたプローブであり、前記プローブが、前記吸着剤が結合する表面を有することを特徴とする、請求項49に記載のキット。

【請求項51】

前記吸着剤が金属器レート吸着剤であることを特徴とする、請求項49に記載のキット。

【請求項52】

前記吸着剤が陽イオン基を含むことを特徴とする、請求項49に記載のキット。

【請求項53】

前記基質が複数の異なる種類の吸着剤を含むことを特徴とする、請求項49に記載のキット。

【請求項54】

前記吸着剤が前記バイオマーカーに対して特異的に結合する抗体であることを特徴とする、請求項49に記載のキット。

【請求項55】

前記キットが(1)溶離剤をさらに含み、前記バイオマーカーが前記溶離剤で洗浄されたとき吸着剤上で保持されることを特徴とする、請求項49に記載のキット。

【請求項56】

キットであって、以下：

(a) 基質に結合する吸着剤を含む前記基質と、

(b) 試験試料を吸着剤と接触させて該吸着剤に保持されたタンパク質バイオマーカーを検出することでタンパク質バイオマーカーを検出するためのインストラクションとを含み、

前記吸着剤が約3486 ± 6、約3963 ± 7、約4071 ± 7、約4079 ± 7、約4580 ± 8、約5298 ± 10、約6099 ± 11、約6542 ± 12、約6797 ± 12、約6949 ± 13；約6990 ± 13、約7024 ± 13、約7054 ± 13、約7820 ± 14、約7844 ± 14、約7885 ± 14、約8067 ± 15、約8356 ± 15、約8943 ± 16、約9656 ± 17、および約9720 ± 18ダルトンからなる群から選択される少なくとも1種類のタンパク質バイオマーカーを保持すること

10

20

30

40

50

ができることを特徴とする、キット。

【請求項 57】

前記基質が気相イオン質量分析計での使用に適合させたプローブであり、前記プローブが、前記吸着剤が結合する表面を有することを特徴とする、請求項 49 に記載のキット。

【請求項 58】

前記吸着剤が金属器レート吸着剤であることを特徴とする、請求項 49 に記載のキット。

【請求項 59】

前記吸着剤が陽イオン基を含むことを特徴とする、請求項 49 に記載のキット。

【請求項 60】

前記基質が複数の異なる種類の吸着剤を含むことを特徴とする、請求項 49 に記載のキット。 10

【請求項 61】

前記吸着剤が前記バイオマーカーに対して特異的に結合する抗体であることを特徴とする、請求項 49 に記載のキット。

【請求項 62】

前記キットが(1)溶離剤をさらに含み、前記バイオマーカーが前記溶離剤で洗浄されたとき吸着剤上で保持されることを特徴とする、請求項 49 に記載のキット。

【請求項 63】

前立腺癌もしくは良性前立腺肥大の疑診断または陰性診断を下すための複数の分類子を用いる方法であって、 20

a) 正常な被験体、前立腺癌と診断された被験体、および良性前立腺肥大と診断された被験体から得た複数の試料から質量スペクトルを得るステップと、

b) ピーク強度値および関連閾値を含む複数の重みつき基準分類子を得るために前記質量スペクトルの少なくとも一部分に対して促進決定木分析を適用するステップと、

c) 前記複数の重みつき基準分類子の線形結合にもとづいて前立腺癌と良性前立腺肥大と陰性診断との少なくとも一つの疑診を下すステップとを含むことを特徴とする、方法。

【請求項 64】

前立腺癌もしくは良性前立腺肥大の疑診断または陰性診断を下す際に使用するための複数の分類子を開発する方法であって、 30

a) 正常な被験体、前立腺癌と診断された被験体、および良性前立腺肥大と診断された被験体から得た複数の試料から質量スペクトルを得るステップと、

b) ピーク強度値および関連閾値を含む複数の重みつき基準分類子を得るために前記質量スペクトルの少なくとも一部分に対して促進決定木分析を適用するステップとを含み、

前記値が線形結合の状態で用いられて、前立腺癌と陰性診断との少なくとも一つの疑診を下すことを特徴する、方法。

【請求項 65】

前立腺癌もしくは良性前立腺肥大の疑診断または陰性診断を下すための複数の分類子を用いて、コンピュータに命令してコンピュータ実行プロセスを実行するためにその中にコンピュータ命令を記憶するコンピュータ・プログラムであって、 40

a) 正常な被験体、前立腺癌と診断された被験体、および良性前立腺肥大と診断された被験体から得た複数の試料から質量スペクトルを得るための第1のコンピュータ・プログラム・コード手段と、

b) ピーク強度値および関連閾値を含む複数の重みつき基準分類子を得るために、前記質量スペクトルの少なくとも一部分に対して促進決定木を適用するための第2のコンピュータ・プログラム・コード手段と、

c) 前記複数の重みつき基準分類子の線形結合にもとづいて前立腺癌および良性前立腺肥大の少なくとも一つの疑診または陰性診断を下すための第3のコンピュータ・プログラム・コード手段とを含むことを特徴とする、コンピュータ・プログラム媒体。 50

【請求項 66】

前立腺癌もしくは良性前立腺肥大の疑診断または陰性診断を下す際に試料するための複数の分類子を開発するために、コンピュータに命令してコンピュータ実行プロセスを実行するためにその中にコンピュータ命令を記憶するコンピュータ・プログラム媒体であって、

a) 正常な被験体、前立腺癌と診断された被験体、および良性前立腺肥大と診断された被験体から得た複数の試料から質量スペクトルを得るための第1のコンピュータ・プログラム・コード手段と、

b) ピーク強度値および関連閾値を含む複数の重みつき基準分類子を得るために前記質量スペクトルの少なくとも一部分に対して促進決定木を適用する第2のコンピュータ・プログラム・コード手段とを含み、

前記値が線形結合の状態 で用いられて、前立腺癌と陰性診断との少なくとも1つの疑診を下すことを特徴とする、コンピュータ・プログラム媒体。

10

【請求項 67】

コンピュータに命令して、良性前立腺肥大または前立腺癌の診断を助けるコンピュータプログラムを実行するためにその中にコンピュータ命令を記憶するコンピュータ・プログラム媒体であって、

(a) 被験体から得た試験試料中の少なくとも2種類のタンパク質バイオマーカーを検出するための第1のコンピュータ・プログラム・コード手段と、

(b) 上記検出を良性前立腺肥大、前立腺癌の疑診もしくは陰性診断に相関させるための第2のコンピュータ・プログラム・コード手段とを含み、

前記タンパク質バイオマーカーが約 4475 ± 81 、約 5074 ± 91 、約 5382 ± 97 、約 7024 ± 13 、約 7820 ± 14 、約 8141 ± 15 、約 9149 ± 16 、約 9508 ± 17 、および約 9656 ± 17 ダルトンからなる群から選択される分子量を有することを特徴とする、コンピュータ・プログラム媒体。

20

【請求項 68】

前記少なくとも2種類のタンパク質バイオマーカーが前記約 7820 ± 14 ダルトンおよび約 7024 ± 13 ダルトン・タンパク質バイオマーカーであることを特徴とする、請求項 67 に記載の方法。

【請求項 69】

前記少なくとも2つのタンパク質バイオマーカーが前記約 7820 ± 14 ダルトン、約 7024 ± 13 ダルトン、約 5382 ± 97 ダルトン、および約 4475 ± 81 ダルトン・バイオマーカーであることを特徴とする、請求項 67 に記載の方法。

30

【請求項 70】

前記少なくとも2つのタンパク質バイオマーカーが前記約 8141 ± 15 ダルトン、約 9149 ± 16 ダルトン、および約 9656 ± 17 ダルトン・バイオマーカーであることを特徴とする、請求項 67 に記載の方法。

【請求項 71】

前記少なくとも2つのタンパク質バイオマーカーが前記約 5074 ± 91 ダルトン、約 9149 ± 16 ダルトン、および約 9656 ± 17 ダルトン・バイオマーカーであることを特徴とする、請求項 67 に記載の方法。

40

【請求項 72】

前記少なくとも2つのタンパク質バイオマーカーが前記約 9149 ± 16 ダルトンおよび約 9508 ± 17 ダルトン・バイオマーカーであることを特徴とする、請求項 67 に記載の方法。

【請求項 73】

前記少なくとも2つのタンパク質バイオマーカーが前記約 5382 ± 97 ダルトン、約 7024 ± 13 ダルトン、および約 7820 ± 14 ダルトン・バイオマーカーであることを特徴とする、請求項 67 に記載の方法。

【請求項 74】

前記少なくとも2つのタンパク質バイオマーカーが前記約 7024 ± 13 ダルトン、約 7

50

820 ± 14 ダルトン、約 5382 ± 97 ダルトン、および約 4475 ± 81 ダルトン・バイオマーカーであることを特徴とする、請求項 67 に記載の方法。

【請求項 75】

前記約 7820 ± 14 ダルトン・タンパク質バイオマーカーの非存在を測定するため、および前記約 7024 ± 13 ダルトン・タンパク質バイオマーカーの存在を測定するための第 3 のコンピュータ・プログラム・コード手段をさらに含むことを特徴とする、請求項 68 に記載の媒体。

【請求項 76】

前記媒体が前記約 7820 ± 14 ダルトンおよび約 7024 ± 13 ダルトン・バイオマーカーの非存在、前記約 5382 ± 97 ダルトン・バイオマーカーの存在、および前記約 4475 ± 81 ダルトン・バイオマーカーの量を測定するための第 4 のコンピュータ・プログラム・コード手段を含むことを特徴とする、請求項 69 に記載の方法。

【請求項 77】

前記媒体が約 8141 ± 15 ダルトン、約 9149 ± 16 ダルトン、および約 9656 ± 17 ダルトン・バイオマーカーの量を測定するための第 5 のコンピュータ・プログラム・コード手段を含むことを特徴とする、請求項 70 に記載の媒体。

【請求項 78】

前記約 5074 ± 91 ダルトン、約 9149 ± 16 ダルトン、および約 9656 ± 17 ダルトン・バイオマーカーの量を測定するための第 6 のコンピュータ・プログラム・コード手段をさらに含むことを特徴とする、請求項 71 に記載の媒体。

【請求項 79】

前記約 9149 ± 16 ダルトンおよび約 9508 ± 17 ダルトン・バイオマーカーの量を測定するための第 7 のコンピュータ・プログラム・コード手段をさらに含むことを特徴とする、請求項 72 に記載の媒体。

【請求項 80】

前記約 5382 ± 97 ダルトン、約 7024 ± 13 ダルトン、および約 7820 ± 14 ダルトン・バイオマーカーの非存在を測定するための第 8 のコンピュータ・プログラム・コード手段をさらに含むことを特徴とする、請求項 73 に記載の媒体。

【請求項 81】

前記約 7024 ± 13 ダルトンおよび約 7820 ± 14 ダルトン・バイオマーカーの非存在、前記約 5382 ± 97 ダルトン・バイオマーカーの存在、前記約 4475 ± 81 ダルトン・バイオマーカーの量を測定するための第 9 のコンピュータ・プログラム・コード手段をさらに含むことを特徴とする、請求項 74 に記載の媒体。

【請求項 82】

前立腺癌または良性前立腺肥大を診断するために、コンピュータに命令して、コンピュータ実行プロセスを実行するためにその中にコンピュータ命令を記憶するコンピュータ・プログラム媒体であって、

(a) 被験体から得た試験試料中でタンパク質バイオマーカーを検出するための第 1 のコンピュータ・プログラム・コード手段と、

(b) 良性前立腺肥大、前立腺癌、または陰性診断の診断と測定を相関させるための第 2 のコンピュータ・プログラム・コード手段とを含み、

前記タンパク質バイオマーカーが

(i) 約 7024 ± 13 ダルトン、および約 7820 ± 14 ダルトン、

(ii) 約 7820 ± 14 ダルトン、約 7024 ± 13 ダルトン、約 5382 ± 97 ダルトン、および約 4475 ± 81 ダルトン、

(iii) 約 8141 ± 15 ダルトン、約 9149 ± 16 ダルトン、および約 9656 ± 17 ダルトン、

(iv) 約 9149 ± 16 ダルトンおよび約 9508 ± 17 ダルトン、

(v) 約 5074 ± 91 ダルトン、約 9149 ± 16 ダルトン、および約 9656 ± 17 ダルトン、または

(v i) 約 5382 ± 97 ダルトン、約 7024 ± 13 ダルトン、および約 7820 ± 14 ダルトンの群にあり、かつ上記の分子量を有することを特徴とする、媒体。

【請求項 83】

前記分子量約 7024 ± 13 を有するタンパク質バイオマーカーの存在および前記分子量約 7820 ± 14 ダルトンを有するタンパク質バイオマーカーの非存在を検出するための第 3 のコンピュータ・プログラム・コード手段をさらに含むことを特徴とする、請求項 82 に記載の媒体。

【請求項 84】

前記約 7820 ± 14 ダルトンおよび約 7024 ± 13 ダルトン・バイオマーカーの非存在、前記約 5382 ± 97 ダルトン・バイオマーカーの存在、ならびに前記約 4475 ± 81 ダルトン・バイオマーカーの量を測定するための第 4 のコンピュータ・プログラム・コード手段をさらに含むことを特徴とする、請求項 82 に記載の媒体。

【請求項 85】

前記約 8141 ± 15 ダルトン、約 9149 ± 16 ダルトン、および約 9656 ± 17 ダルトン・バイオマーカーの量を測定するための第 5 のコンピュータ・プログラム・コード手段をさらに含むことを特徴とする、請求項 82 に記載の媒体。

【請求項 86】

前記約 9149 ± 16 ダルトンおよび約 9508 ± 17 ダルトン・バイオマーカーの量を検出するための第 6 のコンピュータ・プログラム・コード手段をさらに含むことを特徴とする、請求項 82 に記載の方法。

【請求項 87】

前記約 5074 ± 91 ダルトン、約 9149 ± 16 ダルトン、および約 9508 ± 17 ダルトン・バイオマーカーの量を測定するための第 7 のコンピュータ・プログラム・コード手段をさらに含むことを特徴とする、請求項 82 に記載の媒体。

【請求項 88】

約 5382 ± 97 ダルトン、約 7024 ± 13 ダルトン、および約 7820 ± 14 ダルトン・バイオマーカーの非存在を測定するための第 8 のコンピュータ・プログラム・コード手段をさらに含むことを特徴とする、請求項 82 に記載の媒体。

【請求項 89】

良性前立腺肥大または前立腺癌を診断するために、コンピュータに命令してコンピュータ実行プロセスを実行するためにその中にコンピュータ命令を記憶するコンピュータ・プログラム媒体であって、

(a) 被験体から得た試験試料中の少なくとも 2 種類のタンパク質バイオマーカーを検出するための第 1 のコンピュータ・プログラム・コード手段と、

(b) 良性前立腺肥大、前立腺癌、または陰性診断の診断と測定を相関させるための第 2 のコンピュータ・プログラム・コード手段とを含み、

前記タンパク質バイオマーカーが約 3486 ± 6、約 3963 ± 7、約 4071 ± 7、約 4079 ± 7、約 4580 ± 8、約 5298 ± 10、約 6099 ± 11、約 6542 ± 12、約 6797 ± 12、約 6949 ± 13；約 6990 ± 13、約 7024 ± 13、約 7054 ± 13、約 7820 ± 14、約 7844 ± 14、約 7885 ± 14、約 8067 ± 15、約 8356 ± 15、約 8943 ± 16、約 9656 ± 17、および約 9720 ± 18 ダルトンからなる群から選択される分子量を有することを特徴とする、媒体。

【請求項 90】

(a) 約 3963 ± 7、約 4079 ± 7、約 6542 ± 12、約 6797 ± 12、約 6949 ± 13、約 6990 ± 13、約 7024 ± 13、約 7885 ± 14、約 8067 ± 15、約 8356 ± 15、約 9656 ± 17、約 9720 ± 18 ダルトン、およびその組み合わせからなる群から選択される分子量を有するタンパク質バイオマーカーを検出するための第 3 のコンピュータ・プログラム・コード手段と、

(b) 前立腺の診断と前記検出を相関するための第 4 のコンピュータ・プログラム・コード手段とをさらに含むことを特徴とする、請求項 89 に記載の方法。

10

20

30

40

50

【請求項 9 1】

(a) 約 3 4 8 6 ± 6、約 4 0 7 1 ± 7、約 4 5 8 0 ± 8、約 5 2 9 8 ± 1 0、約 6 0 9 9 ± 1 1、約 7 0 5 4 ± 1 3、約 7 8 2 0 ± 1 4、約 7 8 4 4 ± 1 4、約 8 9 4 3 ± 1 6 ダルトン、およびその組み合わせからなる群から選択される分子量を有するタンパク質バイオマーカーを検出するための第 3 のコンピュータ・プログラム・コード手段と、

(b) 前立腺の診断と前記検出を相関するための第 4 のコンピュータ・プログラム・コード手段とをさらに含むことを特徴とする、請求項 8 9 に記載の方法。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、国立癌研究所早期発見研究ネットワーク (National Cancer Institutes Early Detection Research Network) から得た研究助成番号 CA 8 5 0 6 7、国防総省 (Department of Defense) から得られた研究助成番号 DAMD 1 7 - 0 2 - 1 - 0 0 5 4、さらににバージニア州前立腺センター (Virginia Prostate Center) から得た研究助成による政府の支援によりなされた。

【0002】

(関連出願)

この出願は、仮米国特許出願第 6 0 / 3 7 9 , 0 1 8 号 (2 0 0 2 年 5 月 1 0 日出願) の優先権を主張するもので、該米国特許出願の全体を本明細書では援用する。

【背景技術】

【0003】

(発明の背景)

前立腺癌の症例数は、血清前立腺特異的抗原 (P S A) 検査および直腸指診の普及が原因で、過去 1 0 年間のあいだに 3 倍になった (Howe et al., J. Natl. Cancer Inst. 93: 824 - 842 (2001))。これらの効果として癌であると確認される人の数が増えたが、全体的な「早期」発見および侵襲性の強い癌の判断が求められている P S A は、実際の臨床に用いる場合、現在のところ前立腺癌 (P C A) にとって、最も全体的な血清マーカーである。それにもかかわらず、P S A 検査は特異性が欠けている (Djava (登録商標) n, et al. Urology 54: 517 - 5222, (1999) ; Pannek and Partin, Semin. Urol. Oncol. 16: 100 - 105 (1998))。早期発見バイオマーカーとしてその使用を制限すること、および生物学的活性とのその関係については、疑問視されていた (Stamey et al., J. Urol. 167: 103 - 111 (2002))。P C A 死亡率を低下させるために新たな診断用バイオマーカーを同定することが重要である。しかし、P C A が強力な分子および細胞異質性を示すことから、前立腺癌の早期発見を改善するために、複数のバイオマーカーの組み合わせまたはそれらが並んだパネルが必要となるかもしれない。

【0004】

細胞プロテオームの研究によって、バイオマーカー発見のための新たな地平線が示される。二次元ポリアクリルアミドゲル電気泳動 (2 D - E P) は、タンパク質発現の違いの検出および分離に関してプロテオームを調べる古典的なアプローチであった (Srinivas et al., Clin. Chem. 47: 1901 - 1911 (2001) ; Adam et al., Proteomics 1: 1264 - 1270 (2001))。潜在的なタンパク質変異を同定するためのソフトウェア・プログラムおよびロボット工学とが一体となった 2 D - E P 技術の進歩によって、このようなプロテオミクス系の改善がなされている。それにもかかわらず、2 D - E P はいまだに扱いにくく、労働集約的であり、再現性の問題があり、さらに臨床分析への応用が容易ではない。質量分析でも進歩がみられ、タンパク質の分離および分析が高い処理能力 (ハイ・スループット) でおこなわれる (Chong et al., Anal. Chem. 73

10

20

30

40

50

: 1219 - 1227 (2001); Ferrari et al., Rapid Comm. Mass Spec. 14:1149 - 1154 (2000); Keough et al., Electrophoresis 21:2252 - 2265 (2000)。最近の進歩の1つは、Ciphergen Biosystems, Inc. (Fremont, CA) が製造した ProteinChip (登録商標) システムである。このシステムは、表面活性化レーザー脱着イオン化飛行時間計測式 (SELDI - TOF) 質量分析法を用いて、タンパク質チップ配列に結合したタンパク質を検出する (Merchant and Weinberger, Electrophoresis 21:1164 - 1177 (2000); Kuwata et al., Biochem. Biophys. Res. Comm. 245:764 - 773 (1998))。 10

【0005】

このシステムは、タンパク質およびペプチドの複合混合物を分析するための高感度かつ迅速な方法である。当研究所の初期の研究では、前立腺および膀胱癌バイオマーカーを体液および細胞溶解物中に発見し、かつプロファイリングするための SELDI の潜在的能力を立証した (Wright et al., Prostate Cancer and Prostate Diseases 2:264 - 267 (1999); Vlahous et al., Am. J. Pathol., 158:1491 - 1502 (2001))。しかし、良性前立腺肥大、前立腺癌、および陰性診断を見分ける方法の改善がいまだに求められている。 20

【発明の開示】

【課題を解決するための手段】

【0006】

(発明の要旨)

前立腺癌もしくは良性前立腺肥大の診断または該診断の補助、さもなければ陰性診断を下すことに用いることが可能なタンパク質バイオマーカーを発見した。したがって、前立腺癌もしくは良性前立腺肥大の診断を助け、さもなければ診断を下す方法を提供する。

【0007】

本発明の一形態では、診断を助け、さもなければ該診断を下す方法は、被験体から得た試験試料中の少なくとも2種類のタンパク質バイオマーカーを検出することを含む。タンパク質バイオマーカーは、約 4475 ± 81 、約 5074 ± 91 、約 5382 ± 97 、約 7024 ± 13 、約 7820 ± 14 、約 8141 ± 15 、約 9149 ± 16 、約 9508 ± 17 、および約 9656 ± 17 ダルトンからなる群から選択される分子量を持つ。上記方法は、さらに、上記検出と良性前立腺肥大、前立腺癌の疑診 (probable diagnosis) または陰性診断とを相関させることを含む。 30

【0008】

本発明の特定の形態では、被験体から得た試験試料中のタンパク質バイオマーカーを以下の群から選択し、かつ該タンパク質バイオマーカーが以下の分子量を持つものであってもよい。

【0009】

すなわち、(i) 約 7024 ± 13 ダルトンおよび約 7820 ± 14 ダルトン、(ii) 約 7820 ± 14 ダルトン、約 7024 ± 13 ダルトン、約 5382 ± 97 ダルトン、および約 4475 ± 81 ダルトン、(iii) 約 8141 ± 15 ダルトン、約 9149 ± 16 ダルトン、および約 9656 ± 17 ダルトン、(iv) 約 9149 ± 16 ダルトン、および約 9508 ± 17 ダルトン、(v) 約 5074 ± 91 ダルトン、約 9149 ± 16 ダルトン、および約 9656 ± 17 ダルトン、または (vi) 約 5382 ± 97 ダルトン、約 7024 ± 13 ダルトンおよび約 7820 ± 14 ダルトンである。つぎに、検出と良性前立腺肥大、前立腺癌の診断または陰性診断との相関を、上記タンパク質バイオマーカー群の一種以上を分析することによって、形成することができる。 40

【0010】

本発明のさらに別の形態では、検出可能なタンパク質バイオマーカは、約3486 ± 6、約3963 ± 7、約4071 ± 7、4079 ± 7、約4580 ± 8、約5298 ± 10、約6099 ± 11、約6542 ± 12、約6797 ± 12、約6949 ± 13；約6990 ± 13、約7024 ± 13、約7054 ± 13、約7820 ± 14、約7844 ± 14、約7885 ± 14、約8067 ± 15、約8356 ± 15、約8943 ± 16、約9656 ± 17、および約9720 ± 18ダルトンからなる群から選択される分子量を有する。

【0011】

本明細書に記載したタンパク質バイオマーカの少なくとも2種類のものが概ね検出される。本明細書では、一種類以上のバイオマーカの検出とその後の分析が可能であることが理解され、かつ記載されており、上記バイオマーカの全てが包含される。

10

【0012】

本発明のさらに別の局面では、本明細書に記載するバイオマーカを検出するために利用可能で、またさもなければ診断、もしくはさもなければ前立腺癌または前立腺肥大の診断を助けるキットを提供する。本発明の一形態では、キットは、該キットに取り付けられ、かつ約4475 ± 81、約5074 ± 91、約5382 ± 97、約7024 ± 13、約7820 ± 14、約8141 ± 15、約9149 ± 16、約9508 ± 17、および約9656 ± 17ダルトンからなる群から選択される少なくとも一種類のタンパク質バイオマーカを保持することができる吸着剤と、試験試料を吸着剤と接触させて該吸着剤に保持されたタンパク質バイオマーカを検出することでタンパク質バイオマーカを検出するためのインストラクションとを含む。

20

【0013】

本発明のさらに別の実施形態では、キットは、該キットに取り付けられ、かつ約3486 ± 6、約3963 ± 7、約4071 ± 7、4079 ± 7、約4580 ± 8、約5298 ± 10、約6099 ± 11、約6542 ± 12、約6797 ± 12、約6949 ± 13；約6990 ± 13、約7024 ± 13、約7054 ± 13、約7820 ± 14、約7844 ± 14、約7885 ± 14、約8067 ± 15、約8356 ± 15、約8943 ± 16、約9656 ± 17、および約9720 ± 18ダルトンからなる群から選択される少なくとも一種類のタンパク質バイオマーカを保持することができる吸着剤と、試験試料を吸着剤と接触させて該吸着剤に保持されたタンパク質バイオマーカを検出することでタンパク質バイオマーカを検出するためのインストラクションとを含む。

30

【0014】

本発明のさらに別の局面では、前立腺癌もしくは良性前立腺肥大の疑診断または陰性診断を下すための複数の分類子を用いる方法を提供する。本発明の一形態では、方法は、(a) 正常な被験体、前立腺癌と診断された被験体、および良性前立腺肥大と診断された被験体から得た複数の試料から質量スペクトルを得ること、(b) ピーク強度値および関連閾値を含む複数の重みつき基準分類子 (weighted base classifiers) を得るために、質量スペクトルの少なくとも一部分に対して促進決定木 (促進決定木分析) を適用すること、(c) 上記複数の重みつき基準分類子の線形結合 (linear combination) にもとづいて前立腺癌および良性前立腺肥大の少なくとも一つの疑診または陰性診断を下すこと、を含む。本発明の特定の形態では、上記方法は、前立腺癌、良性前立腺癌の疑診を下すか、もしくは陰性診断を下すために、線形結合の状態ピーク強度値と関連閾値とを用いることを含む。

40

【0015】

本発明のさらに別の局面では、コンピュータ命令を格納するコンピュータ・プログラム媒体が提供され、該プログラム媒体は複数の分類子用いてコンピュータ実行プロセスを実行するようにコンピュータに命令して前立腺癌もしくは良性前立腺肥大の疑診断または陰性診断を下す。本発明の一形態では、コンピュータ・プログラム媒体は、(a) 正常な被験体、前立腺癌と診断された被験体、および良性前立腺肥大と診断された被験体から得た複数の試料から質量スペクトルを得るための第1のコンピュータ・プログラム・コード手

50

段と、(b)ピーク強度値および関連閾値を含む複数の重みつき基準分類子を得るために、質量スペクトルの少なくとも一部分に対して促進決定木(促進決定木分析)を適用するための第2のコンピュータ・プログラム・コード手段と、(c)上記複数の重みつき基準分類子の線形結合(linear combination)にもとづいて前立腺癌および良性前立腺肥大の少なくとも1つの疑診または陰性診断を下すための第3のコンピュータ・プログラム・コード手段と、を含む。前立腺癌、良性前立腺肥大の少なくとも1つの疑診を下すか、もしくは陰性診断を下すために、線形結合の状態でピーク強度値と関連閾値とを用いてもよい。

【0016】

本発明の別の形態では、コンピュータ・プログラム媒体は(a)被験体から得た試験試料中にあり、かつ約4475±81、約5074±91、約5382±97、約7024±13、約7820±14、約8141±15、約9149±16、約9508±17、および約9656±17ダルトンからなる群から選択される分子量を持つ少なくとも2種類のタンパク質バイオマーカーを検出するための第1のコンピュータ・プログラム・コード手段と、(b)上記検出を良性前立腺肥大、前立腺癌の疑診もしくは陰性診断に関連させるための第2のコンピュータ・プログラム・コード手段とを含む。

10

【0017】

本発明のさらに別の形態では、(i)約7024±13ダルトンおよび約7820±14ダルトン；(ii)約7820±14ダルトン、約7024±13ダルトン、約5382±97ダルトン、および約4475±81ダルトン、(iii)約8141±15ダルトン、約9149±16ダルトン、および約9656±17ダルトン、(iv)約9149±16ダルトン、および約9508±17ダルトン、(v)約5074±91ダルトン、約9149±16ダルトン、および約9656±17ダルトン、または(vi)約5382±97ダルトン、約7024±13ダルトン、および約7820±14ダルトンの群にあり、かつそれらの分子量を持つ被験体タンパク質バイオマーカーを、被験体から得た試験試料中で検出するための、第1のコンピュータ・コード手段と、(b)上記検出を良性前立腺肥大、前立腺癌の疑診もしくは陰性診断に関連させるための第2のコンピュータ・プログラム・コード手段とを含む。

20

【0018】

本発明のさらに別の形態では、コンピュータ可読媒体は、(a)被験体から得た試験試料中にあり、かつ約3486±6、約3963±7、約4071±7、4079±7、約4580±8、約5298±10、約6099±11、約6542±12、約6797±12、約6949±13、約6990±13、約7024±13、約7054±13、約7820±14、約7844±14、約7885±14、約8067±15、約8356±15、約8943±16、約9656±17、および約9720±18ダルトンからなる群から選択される分子量を持つ少なくとも2種類のタンパク質バイオマーカーを検出するための第1のコンピュータ・プログラム・コード手段と、(b)上記検出を良性前立腺肥大、前立腺癌の疑診もしくは陰性診断に関連させるための第2のコンピュータ・プログラム・コード手段とを含むものであってもよい。

30

【0019】

本発明の目的は、前立腺癌もしくは良性前立腺肥大を診断または該診断を助けるか、さもなければ陰性診断を下す方法を提供することである。

40

【0020】

本発明の別の目的は、本明細書に記載したバイオマーカーを検出するために利用可能であり、かつ前立腺癌または良性前立腺肥大を診断または該診断を助けるために利用可能であるキットを提供することである。

【0021】

本発明の別の目的は、前立腺癌または良性前立腺肥大の疑診を下すために、複数の分類しを用いる方法を提供することである。

【0022】

50

本発明の別の目的は、前立腺癌、良性前立腺肥大の疑診もしくは陰性診断を下すために複数の分類子を開発および/または使用するコンピュータ実行プロセスを実行するようにコンピュータに対して命令するためのコンピュータ命令を格納するコンピュータ・プログラム媒体を提供することである。

【0023】

本発明のこれらの目的および他の目的ならびに利点は、本明細書の記述から明瞭に理解できるだろう。

【0024】

(好ましい実施形態の説明)

本発明の原理を深めることを目的として、好ましい実施形態を以下参照するとともに、同一のものを説明するために特定の言葉を用いる。しかし、本発明の範囲に限定を加えることを意図したものではなく、本発明のそのような変更およびさらなる修飾、本明細書に例証したような本発明の原理のさらなる応用は、当業者にとっては通常思いつくものと考えられることがわかる。

【0025】

本発明は、良性前立腺肥大および前立腺癌の診断を助ける方法と良性前立腺肥大および前立腺癌を診断する方法とに関する。表面活性化レーザー脱着イオン化質量分析を種々のアルゴリズムと併用して、良性前立腺肥大、前立腺癌の診断を助けるために、および/または該診断をおこなうために、あるいは陰性診断を下すために、種々の決定木で利用可能なタンク質バイオマーカーを推測する。

【0026】

本発明の方法は、良性前立腺肥大、前立腺癌の個体を正常個体と効果的に区別する。この明細書に定義するように、正常個体は、良性前立腺肥大または前立腺癌に関して陰性診断される個体である。すなわち、正常個体は良性前立腺または前立腺癌を持たない。上記方法は、被験体から得た試験試料中のタンパク質バイオマーカーを検出することを含む。例えば、分子量が約4475 ± 81、約5074 ± 91、約5382 ± 97、約7024 ± 13、約7820 ± 14、約8141 ± 15、約9149 ± 16、約9508 ± 17、および約9656 ± 17ダルトンであるタンパク質バイオマーカーは、良性前立腺肥大、前立腺癌の疑診を助けるか、もしくは陰性診断を助けるものと同定されている。また、分子量が約3486 ± 6、約3963 ± 7、約4071 ± 7、4079 ± 7、約4580 ± 8、約5298 ± 10、約6099 ± 11、約6542 ± 12、約6797 ± 12、約6949 ± 13；約6990 ± 13、約7024 ± 13、約7054 ± 13、約7820 ± 14、約7844 ± 14、約7885 ± 14、約8067 ± 15、約8356 ± 15、約8943 ± 16、約9656 ± 17、および約9720 ± 18ダルトンであるタンパク質バイオマーカーもまた、診断を助けるものと同定されている。本発明によれば、少なくとも2種類のタンパク質バイオマーカーが検出される。本明細書で用いられるように、「検出 (d e t e c t i n g)」という用語は、タンパク質バイオマーカーの有無、量、または組み合わせを決定することを含む。タンパク質バイオマーカーの量は、質量分析法で同定されるようなピーク強度 (例えば、すなわちタンパク質バイオマーカーの濃度) によって表すことが可能である。

【0027】

本発明の特定の形態では、タンパク質バイオマーカーの選択された群は、前立腺癌または良性前立腺肥大の診断で有用性を見いだす。例えば、以下の群のマーカーは、特定の診断を下すか、さもなければ該診断を下すのを助けることに有用性が見いだされる。すなわち、(1) 前立腺癌を診断するための7820および7024ダルトン・バイオマーカー、(2) 良性前立腺肥大を診断するための7820、7024および5382ダルトン・バイオマーカー、(3) 前立腺癌と良性前立腺肥大とを区別するための7820、7024、5382、および4474ダルトン・バイオマーカー、(4) 前立腺癌と良性前立腺肥大とを区別する914および9508ダルトン・バイオマーカー、(5) 前立腺癌と正常個体とを区別する9149、9656、および8141ダルトン・バイオマーカー、な

10

20

30

40

50

らびに(6)前立腺癌の個体と正常個体とを区別する9149、9656、および5074ダルトン・バイオマーカーである。例えば、そのようなマーカーをどのように利用するかを示す決定木を図2に示す。

【0028】

上記したように、種々のバイオマーカーの有無および/または量を利用して、特定の診断を下したり、さもなければ該診断の手助けをおこなうことが可能である。例えば、また図2を参照すると、7820ピークがなく、7024ピークがある場合は、前立腺癌の診断と相関すると考えられる。ここで上記第2の群を参照すると、7024、および7820ダルトン・バイオマーカーが存在しないことは、良性前立腺肥大と相関している可能性がある。

10

【0029】

選択されたバイオマーカーの有無は、他のバイオマーカーの量とともに、特定の診断を下したり、さもなければ該診断を下すのを手助けしたりすることも可能である。例えば、上記した群3に関して、7820ダルトンおよび7024ダルトン・バイオマーカーがなく、図2に示した閾値を下回る4475ダルトン・バイオマーカーの存在および5382ダルトン・バイオマーカーの存在は、前立腺癌の診断に相関しており、一方この同一群に閾値を上回る4475ダルトン・バイオマーカーが存在するならば、陰性診断を下すことが可能である。図2の閾値は、バイオマーカーの正規化ピーク強度を示す。実施例1により完全に説明するように、それらの閾値は、特定のバイオマーカーの正規化ピーク強度またはバイオマーカーの濃度を表すことが可能である。正規化プロセスは、分析したタンパク質に関係がないイオン電流を差し引くことを伴うものであってもよい。あるいは、正規化プロセスは、内部または外部対照のピーク強度と相対的なピーク強度を報告することを含むこともあり得る。例えば、既知のタンパク質をシステムに加えてもよい。また、被験体によって生産された既知のタンパク質、例えばアルブミンは、内部標準または対照として機能することが可能である。

20

【0030】

ここで、つぎの群4について検討すると、図2で表示した閾値を上回る9149および9508ダルトン・ピークの存在は、前立腺癌の診断と相関しており、その一方で、表示した閾値を上回る9149ダルトン・バイオマーカーの存在と表示した閾値以下の9508ダルトン・バイオマーカーの存在は、良性前立腺肥大の診断と相関している可能性がある。

30

【0031】

ここで群5を検討すると、9149、9655、および8141ダルトン・バイオマーカーが図2で表示した閾値を下回る場合、前立腺癌に対する相関が形成される可能性がある。あるいは、9149および9655ダルトン・バイオマーカーが表示した閾値を下回り、かつ8141ダルトン・バイオマーカーが表示した閾値を上回る場合、陰性診断が下される可能性がある。

【0032】

ここで群6を検討すると、図2に表示した特定の閾値を上回る9656ダルトン・バイオマーカーが存在し、かつ表示した閾値を下回る9149および5074ダルトン・バイオマーカーが存在することは、前立腺癌の診断と相関する可能性がある。あるいは、図2に表示した閾値を上回る9656および5074ダルトン・バイオマーカーの存在ならびに表示した閾値を下回る9149ダルトン・バイオマーカーの存在は、陰性診断と相関する可能性がある。

40

【0033】

上記方法のさらに別の形態では、検出可能なタンパク質バイオマーカーとして、約3486±6、約3963±7、約4071±7、4079±7、約4580±8、約5298±10、約6099±11、約6542±12、約6797±12、約6949±13；約6990±13、約7024±13、約7054±13、約7820±14、約7844±14、約7885±14、約8067±15、約8356±15、約8943±

50

16、約9656 ± 17、および約9720 ± 18ダルトンからなる群から選択される分子量を持つものが挙げられる。これらのバイオマーカーの検出と前立腺癌、良性前立腺肥大、または陰性診断との相関は、好ましくは実施例2でより完全に記載されるように促進決定木分析を用いて達成される。

【0034】

上記方法は、少なくとも一種類のタンパク質バイオマーカーを検出することが含まれる。しかし、任意の数のバイオマーカーを検出することも可能である。好ましくは、少なくとも2種類のタンパク質バイオマーカーが検出される。しかし、3、4、またはそれ以上の本明細書に記載したバイオマーカーの全てを診断に利用することが可能であることが理解される。したがって、一種類以上のマーカーが検出しうるだけでなく、1ないし9、好ましくは2ないし9、2ないし12、および2ないし21種類のバイオマーカー、またはいくつかの他の組み合わせを、本明細書に記載したように検出および分析することが可能である。また、本明細書に記載していない他のタンパク質バイオマーカーを、現在開示しているタンパク質バイオマーカーのいずれかと組み合わせ、前立腺癌または良性前立腺肥大の診断の助けにすることも可能である。さらに、上記タンパク質バイオマーカーの任意の組み合わせを、本発明にもとづいて検出することができる。本発明は、タンパク質バイオマーカーの選択された群が前立腺癌または良性前立腺肥大の診断での有用性を見いだすことを発見した。例えば、約3963 ± 7、約4079 ± 7、約6542 ± 12、約6797 ± 12、約6949 ± 13；約6990 ± 13、約7024 ± 13、約7885 ± 14、約8067 ± 15、約8356 ± 15、約9656 ± 17、および約9720 ± 18ダルトンからなる群から選択された分子量を有するタンパク質バイオマーカーを、陰性診断から前立腺癌を診断するために有利に利用することが可能である。また、約3486 ± 6、約4071 ± 7、約4580 ± 8、約5298 ± 10、約6099 ± 11、約7054 ± 13、約7820 ± 14、約7844 ± 14、および約8943 ± 16からなる群から選択される分子量を持つタンパク質バイオマーカーは、陰性診断から良性前立腺肥大を区別するのに有利に利用することが可能である。

【0035】

試験試料中の本明細書に記載したタンパク質バイオマーカーの検出は、種々の方法でおこなうことが可能である。本発明の一形態では、バイオマーカー検出のための方法が、気相イオン分光計を用いる気相イオン分光分析によってバイオマーカーを検出することを含む。この方法は、本明細書に記載したタンパク質バイオマーカーとしてバイオマーカーを有する試験試料を、該バイオマーカーと吸着剤との結合が可能となる条件下で、吸着剤を含む基質と接触させ、吸着剤に結合したバイオマーカーを気相イオン分光分析によって検出することが含まれる。

【0036】

多種多様の吸着剤を使用することが可能である。吸着剤として、例えば疎水基、親水基、カチオン基、アニオン基、金属イオン・キレート基、抗原性バイオマーカーに特異的結合する抗体、もしくはそれらの何らかの組み合わせ（例えば、「混合モード(mixed mode)」吸着剤)が挙げることが可能である。疎水基を含む典型的な吸着剤として、脂肪族炭化水素（例えば、C₁ ~ C₁₈脂肪族炭化水素）を有するマトリックスおよびフェニル基が含まれる芳香族炭化水素官能基を有するマトリックスが挙げられる。親水基を含む典型的な吸着剤として、酸化ケイ素、またはポリアルキレングリコール等の親水性ポリマー（例えば、ポリエチレングリコール、デキストラン、アガロース、もしくはセルロース）が挙げられる。カチオン基を有する典型的な吸着剤として、二級、三級、または四級アミンのマトリックスが挙げられる。アニオン基を有する典型的な吸着剤として、硫酸アニオンのマトリックスおよびカルボン酸アニオンもしくはリン酸アニオンのマトリックスが挙げられる。金属キレート基を有する典型的な吸着剤として、金属イオン（例えば、銅、ニッケル、コバルト、亜鉛、鉄、アルミニウム、およびカルシウム）と配位共有結合を形成しうる1つ以上の電子供与体を持つ有機分子が挙げられる。抗体を含む典型的な吸着剤として、本明細書で提供されるバイオマーカーのいずれかに対する抗体が挙げられ

、当業者に周知の方法によって容易に作ることが可能である。

【0037】

あるいは、上記基質を、気相イオン分光計に着脱自在に挿入可能なプローブの形態とすることができる。例えば、基質は、その表面に吸着剤を有するヒモ状の形態であってもよい。本発明の別の形態では、基質が第2の基質上に配置されて、着脱自在に気相イオン分光計に挿入可能なプローブを形成する。例えば、基質は、第2の基質上に配置してプローブを形成することが可能な、マーカを結合するための官能基を有するポリマー・ビーズまたはガラス・ビーズ等の固相形状とすることができる。第2の基質をヒモ状の形態または所定の位置に形成された一連のウェルを有するプレートであってもよい。このようにして、バイオマーカを第1の基質に吸着させ、第2の基質に移すことができ、該第2の基質によって気相イオン分光分析による分析に上記バイオマーカがゆだねられる。

10

【0038】

プローブは、気相イオン分光計に着脱自在に挿入可能である限り、多種多様な所望の形状の形態にすることができ、該形状として円形、楕円形、正方形、矩形、もしくは他の多角形または所望の形状にすることができる。また、プローブは注入装置(inlet system)および検出器の用途に好ましくは適する、さもなければ構成されることができる。例えば、プローブは水平および/または垂直に平行移動可能なキャリッジへの搭載に適應することができ、該プローブの手作業による再配置等を必要とすることなく、水平および/または垂直にプローブを連続する位置に動かすことができる。

20

【0039】

プローブを形成する基質を、様々な吸着剤を指示することができる多種多様な材料から作ることができる。典型的な材料として、ガラスおよびセラミックス等の絶縁材料、シリコン・ウェハ等の半絶縁材料、導電材料(ニッケル、黄銅、銅、アルミニウム、金等の金属もしくは導電性ポリマーを含む)、有機ポリマー、バイオポリマー、またはそれらの組み合わせが挙げられる。

【0040】

本発明の別の実施形態では、基質の性質に応じて、基質表面に吸着剤を形成してもよい。別の例では、基質の表面を修飾して所望の吸着剤をその上に取り込ませることも可能である。プローブを形成する基質の表面を、基質それ自体がバイオマーカと結合することができない場合に該バイオマーカと結合しうる吸着剤に結合するように処置さもなければそのように調整することができる。あるいは、基質の表面を処置さもなければ調整して、所望のバイオマーカに結合するその自然な能力を高めることもできる。本発明に用いるのに適した他のプローブを、PCT国際特許出願番号WO01/25791(Tai-Tung他)およびWO01/71360(Wright他)に見いだすことが可能である。

30

【0041】

上記吸着剤を、多種多様なパターンでプローブ基質上に置くことが可能であり、該パターンとして連続または非連続的なパターンが挙げられる。単一種類の吸着剤または複数種類の吸着剤を基質表面に置いてよい。パターンは、線状、円等の曲線、さもなければ当業者に所望または公知の形状またはパターンであってもよい。

40

【0042】

プローブを生産するための方法は、当業者に公知のように基質材料および/または吸着剤の選択に依存するだろう。例えば、基質が金属である場合、その表面に塗布される吸着剤に依存して、その表面を調製することが可能である。例えば、基質表面を金属(例えば、酸化ケイ素、酸化チタニウム、または金)で被覆して金属表面の誘導体化により吸着剤を形成することも可能である。つぎに、基質表面を二官能性リンカー(bifunctional linker)により誘導体化し、そのうちの一つを該リンカーの表面および反対側の端部の官能基と結合(例えば共有結合)させる。さらなる例として、結晶シリコ

50

ンから生じた多孔性シリコン面を含む基質を化学的に修飾してマーカーを結合させるための吸着剤を含ませることができる。また、ヒドロゲル骨格を持つ吸着剤をモノマー溶液の *in situ* 重合により基質表面に直接形成することができ、該モノマー溶液として、例えば、置換アクリルアミドまたはアクリレート・モノマー、または吸着剤としてえり抜きの官能基を含むそれらの誘導体が挙げられる。

【0043】

本発明の好ましい形態では、プローブを、チップ、例えば *Ciphergen Biosystems, Inc.* (Palo Alto, CA) から入手可能なチップとすることが可能である。チップは、親水性、疎水性、アニオン交換性、カチオン交換性、固定化金属親和性、もしくは事前に活性化されたタンパク質チップ・アレイであってもよい。疎水性チップは、逆送相互作用によりタンパク質に結合する長鎖脂肪族表面を含む *ProteinChip H4* であってもよい。疎水性チップは、二酸化ケイ素基質表面を含む *ProteinChip NP1* および *ProteinChip NP2* であってもよい。カチオン交換 *ProteinChip* アレイを *ProteinChip WCX2* (カチオン性タンパク質に結合するカルボキシレート表面を持つ弱カチオン交換アレイ) としてもよい。あるいは、上記チップをアニオン交換タンパク質チップ・アレイとすることも可能であり、例えば二酸化ケイ素被覆アルミニウム基質から作られる *SAX1* (強アニオン交換) *ProteinChip*、またはアニオン性タンパク質に結合する能力の高い四級アンモニウム表面を持つ *ProteinChip SAX2* がある。さらに有用なチップとして、表面にニトリロ三酢酸を持つ固定化金属親和性補足チップ (*immobilized metal affinity capture chip*) (*IMAC3*) が考えられる。さらに、その代わりとして、*ProteinChip PS1* が利用可能であり、このチップはアミノ基と共有結合的に反応するエポキシ表面カルボニルジイミダゾール面を含む。あるいは、アミンおよびチオール基と共有結合的に反応するエポキシ面を含む *ProteinChip PS2* であってもよい。

【0044】

本発明によれば、プローブは試験試料に接する。この試験試料は多種多様な源から得ることが可能である。一般に、試料は前立腺癌、良性前立腺肥大に関する試験が行われている被験体または患者の体液から得られ、あるいは正常な個体から得られ、さもなければそのような疾患のリスクを持っていると思われる個体から得られる。好ましい体液は血液または血清である。バイオマーカーが見つかる可能性のある他の体液として、例えば、精液、精漿、唾液、リンパ液、肺/気管支洗浄液、粘液、乳頭分泌、痰、涙、唾液が挙げられる。他の試験試料源として、例えば、大便が挙げられる。必要に応じて、プローブと接触させる前に溶離剤に可溶化または該溶離剤と混合することができる。プローブと試験試料との接触は、多種多様な技術、例えば水浴、浸漬(ソーキング)、液浸(ディッピング)、噴霧、洗浄、ピペティング、または他の望ましい方法で、おこなうことが可能である。このような方法は、プローブの吸着剤が好ましくは試験試料溶液に接触するようにしておこなわれる。バイオマーカー濃度または試料中のバイオマーカー濃度が変動するにもかかわらず、吸着剤に結合させるために約 $1 \mu\text{l}$ ないし約 $500 \mu\text{l}$ の溶液中に約 1 アトモルないし約 100 ピコモルのマーカーを含む 1 容量の試験試料と接触させることが一般に望ましい。

【0045】

バイオマーカーと吸着剤とが結合するのに十分な期間にわたって、試料とプローブとをお互いに接触させる。この時間は試料の性質、バイオマーカーの性質、吸着剤の性質、およびバイオマーカーが溶かされる溶液の性質に応じて変化するが、試料と吸着剤との接触は概ね約 30 秒ないし約 12 時間、好ましくは約 30 秒ないし約 15 分間にわたっておこなう。

【0046】

プローブが試料と接触する際の温度は、試料の性質、バイオマーカーの性質、吸着剤の性質、およびバイオマーカーが溶かされる溶液の性質に依存する。一般に、試料は周囲温

度および周囲圧力ならびに周囲条件下でプローブと接触することが可能である。しかし、温度および圧力を必要に応じて変えてもよい。例えば、温度は約 4 °C ないし約 37 °C の範囲で変えることが可能である。

【0047】

マーカーが吸着剤または基質表面に結合するのに十分な時間にわたってプローブに試料が接触した後、吸着剤がいったい使用されていない場合、未結合材料を基質または吸着剤表面から洗浄して結合した材料のみそれぞれの表面に残るようにすることが可能である。洗浄は、例えば、水浴、浸漬（ソーキング）、液浸（ディッピング）、濯ぎ、噴霧、さもなければ溶離液または他の洗浄液で各々の表面を洗浄することで達成できる。プローブ上の吸着剤の小さな点に溶離液等の洗浄液が導入される場合、マイクロフルイディクス・プロセスが好ましくは用いられる。洗浄液の温度は変動してもよいが、概ね約 0 °C ないし約 100 °C、好ましくは約 4 °C ないし約 37 °C である。

10

【0048】

多種多様の洗浄液を利用してプローブの基質表面を洗うことが可能である。洗浄液は有機溶液でも水溶液であってもよい。典型的な水溶液は緩衝液であり、HEPES 緩衝液、Tris 緩衝液、リン酸緩衝液、または当業者に公知の他の類似の緩衝液が挙げられる。特定の洗浄液の選択は、バイオマーカーの性質および利用する吸着剤の性質に依存する。例えば、プローブが吸着剤として疎水基およびスルホン酸塩基を含む場合（例えば、SCXIP Protein Chip（登録商標））、HEPES 緩衝液等の水溶液を用いることができる。もしプローブが吸着剤として金属結合基を含む場合（例えば、Ni(II) Protein Chip（登録商標）アレイ）、リン酸緩衝液等の水溶液を用いることが好ましいと思われる。さらに別の例として、もしプローブが吸着剤として疎水基を含む場合（例えば、HF Protein Chip（登録商標）アレイによる）、水が好ましい洗浄液と考えられる。

20

【0049】

エネルギー吸収分子、例えば溶液中に含まれるものを、プローブの基質表面に結合したマーカーまたは他の物質に適用することが可能である。本明細書で用いられるように、「エネルギー吸収分子（energy absorbing molecule）」は、気相イオン分光計内のエネルギー源からエネルギーを吸収する分子のことをいい、プローブ表面からマーカー又は他の物質が脱着するのを助けることが可能である。典型的なエネルギー吸収分子として、桂皮酸誘導体、シナピン酸、ジヒドロキシ安息香酸、および当業者に公知の他の類似の分子が挙げられる。エネルギー吸収分子を、試料およびプローブ基質を接触させるために、例えば、好ましくは未結合材料をプローブ基質表面から洗い落とした後に、噴霧、ピペッティング、または液漬を含む本明細書で既に議論した多種多様な技術によって適用することが可能である。

30

【0050】

別の実施形態において、チップは、SENDチップであり得る。「表面活性化ニート脱着（Surface-Enhanced Neat Desorption）」または「SEND」は、プローブ面に化学的に結合したエネルギー吸収分子を含むプローブの使用を伴うSELDIの種類である（「SENDプローブ」）。「エネルギー吸収分子（energy absorbing molecules）」（「EAM」）は、レーザー脱着/イオン化源からのエネルギーを吸着することが可能な分子のことをいい、その後で該分子に接触している分析物分子の脱着およびイオン化に関与する。そのフレーズは、MALDIで用いられる分子を含み、該MALDIは「マトリックス（matrix）」ともしばしば呼ばれるもので、他のものと同様に、明確に桂皮酸誘導体、シナピン酸（「SPA」）、シアノヒドロキシ桂皮酸（「CHCA」）およびジヒドロキシ安息香酸、フェラ酸、ならびにヒドロキシアセトフェノン誘導体が挙げられる。それには、SELDIで用いられたEAMも含まれる。SENDは、さらに米国特許第5,719,060号および米国特許出願第60/408255号、2002年9月4日出願（Kitagawa, "Monomers And Polymers Having Energy Ab

40

50

sorbing Moieties Of Use I Desorption / Ionization Of Analytes”)に記載されている。SENDバイオチップは、レーザ脱着 / イオン化に先立ってチップに対して外部マトリックスを適用する必要性を排除する。

【0051】

バイオマーカを適切なプローブに結合させた後、該バイオマーカを検出および定量することが可能であり、さらに / またはその特徴をさもなければ気相イオン分光計を用いて決定することも可能である。当業者に公知のように、気相イオン分光計として、例えば、質量分析計、イオン移動度分光計、および全イオン電流測定装置が挙げられる。

【0052】

好ましい実施形態では、質量分析計を用いてプローブの基質表面に結合したバイオマーカを検出する。プローブはその表面に結合したマーカを有しており、該プローブを質量分析計の注入装置に導入することが可能である。つづいて、マーカを脱イオン化源（例えば、当業者に公知のレーザー、高速原子衝撃、プラズマ、または他の適当なイオン化源）によって脱イオン化してもよい。発生したイオンは、概ね、イオン光学装置によって回収された後、質量分析器によって、通過イオンを分散および分析する。検出器が検出されたイオンの情報を質量対電荷比 (mass-to-charge ratio) に翻訳する。マーカの検出および / または定量化は概ねシングナル強度の検出を伴う。

【0053】

本発明のさらに好ましい形態では、質量分析計はレーザー脱着飛行時間型質量分析計であり、より好ましくは表面活性化レーザー脱着飛行時間型試料分析法 (SELDI) が用いられる。SELDIは、生体分子に対する気相イオン分光分析法の改良型である。SELDIでは、分析物が塗布される表面が分析物補足および / または脱着に積極的役割を果たす。

【0054】

当業者に公知のように、レーザー脱着質量分析法では、結合マーカを有するプローブが注入装置に導入される。マーカは、気相イオン化源によって気相のなかに脱着およびイオン化される。生じたイオンをイオン光学装置によって回収する。イオンは、短い高電圧フィールドを介して飛行時間型質量分析器で加速化され、高真空チャンバーに流れ込むことが可能となる。加速化されたイオンは、異なる時間で高真空チャンバーの遠位端にある感度検出器表面にぶつかる。飛行時間はイオンの質量の関数であることから、イオン化と衝突との間の経過時間を用いて特定質量の分子の有無を識別することができる。バイオマーカの計量は、相対量または絶対量のいずれでもあっても、バイオマーカの表示シングナルの高度とバイオマーカの対照量もしくは当業者に知られているような標準と比較することによって達成可能である。レーザー脱着飛行時間型質量分析計の構成要素は、脱着、加速、検出、または測定時間の種々の手段を用いる本明細書に記載されている、および / または当業者に公知の他の構成要素と組み合わせることも可能である。

【0055】

さらに別の実施形態では、バイオマーカの検出および / または計量は、マトリックス支援レーザー脱着イオン化 (MALDI) によって達成することが可能である。MALDIもまた、固体相から生物試料を気化およびイオン化して直接的に気相にする。当業者に公知のように、好ましくはプロセスの過程での分解を好ましくは防ぐために、所望の分析物を含む分析物と共結晶化するマトリックスに試料を溶解し、さもなければ懸濁することが含まれる。

【0056】

本発明の別の形態では、イオン移動度分光計を用いて、本明細書に記載のバイオマーカを検出および特徴付けすることができる。イオン移動度分光分析の原理は、複数のイオンの異なる移動度に基づいている。特に、イオン化によって生じた試料のイオンは、質量、電荷、または形状等が異なることから、電場の影響下で管の中を異なる速度で移動する。イオン（概ね電流の形態で）が検出器に登録されると、つづいて該イオンを試料中のマ

10

20

30

40

50

ーカーまたは他の物質を同定するために使用することができる。イオン移動度分光分析の1つの利点は、それが大気圧で作用しうることである。

【0057】

別の実施形態では、全イオン電流測定装置を用いて、本発明に記載のバイオマーカーの検出および特徴付けをおこなうことができる。この装置は、例えば、プローブが、単一の種類のマーカーのみと結合することを許容する表面化学を有する。単一の種類のマーカーがプローブに結合した場合、イオン化したバイオマーカーから生じた全電流がマーカーの性質を反映する。つぎに、バイオマーカーによって生じた全イオン電流を、既知の化合物の格納された全イオン電流と比較することができる。これにより、バイオマーカーの特徴を決定することができる。

10

【0058】

TOFからM/Zへの変化は、飛行時間 (time s - o f - f l i g h t) から質量対電荷比 (M/Z) に変換させるアルゴリズムの適用を含む。このステップでは、シグナルが時間ドメインから質量ドメインに変換される。すなわち、各飛行時間が質量対電荷比 (すなわちM/Z) に変換される。キャリブレーションを内部または外部的におこなうことができる。内部キャリブレーションでは、分析した試料は、既知のM/Zからなる一種類以上の分析物を含む。これらの密集した (m a s s e d) 分析物を表す飛行時間での複数のシグナル・ピークが既知のM/Zに割り当てられる。これらの割り当てられたM/Z比にもとづいて、パラメータは、飛行時間をM/Zに変換する数学的関数に適している。外部キャリブレーションでは、飛行時間をM/Zに変換する関数 (例えば先行する内部キャリブレーションによって生成されたもの) が内部キャリブレーションを用いることなく飛行時間型スペクトルに適用される。

20

【0059】

ベースライン・サブトラクションは、スペクトルを混乱させる人工的かつ再現可能なインストルメント・オフセットを除去することによってデータ数量化を改善する。それは、ピーク幅等のパラメータを取り込むアルゴリズムを用いてスペクトルのベースラインを計算し、つづいて質量スペクトルからベースラインを差し引くことを含む。

【0060】

高周波ノイズ・シグナルを、スムージング機能の適用によって除去する。典型的スムージング機能は、各々の時間依存的瓶に対して移動平均関数を適用する。改良版では、移動平均フィルターが可変幅デジタル・フィルターであり、該フィルターのバンド幅がピーク・バンド幅等の関数として変化し、一般に増加する飛行時間によってより広くなる。例えば、WO 00/70648, November 23, 2000 (Gavin et al., "Variable Width Digital Filter for Time-of-flight Mass Spectrometry") を参照せよ。

30

【0061】

一般に分析は、分析物からのシグナルを表すスペクトル内のピークを同定することを含む。ピーク選択は、もちろん、目視によっておこなうことができる。しかし、ソフトウェアは、ピークの検出を自動化することができる Ciphergen's Protein Chip (登録商標) ソフトウェアの一部として入手可能である。一般に、このソフトウェアは、選択された閾値を上回る信号対雑音比を持つシグナルを同定し、該ピーク・シグナルの中心でのピークの質量を標識することで、機能する。1つの有用な応用では、多くのスペクトルを比較して、質量スペクトルの一部の選択されたパーセンテージで存在する同一のピークを同定する。このソフトウェアの一バージョンは、所定の質量範囲内の種々のスペクトルに現れるピーク全てを集め、質量 (M/Z) クラスターの中心点に近いピーク全てに対して質量 (M/Z) を割り当てる。

40

【0062】

1つ以上のスペクトルからのピーク・データは、被験体に対して、例えば各横列が特定の質量スペクトルを表し、各縦列が質量によって定義されるスペクトルでのピークを表し

50

、さらに各セルはその特定のスペクトルでのピーク強度を含む。種々の統計学的またはパターン認識的アプローチをデータに適用することができる。

【0063】

バイオマーカの脱着および検出によって生ずるデータを、プログラム可能なデジタル・コンピュータを用いて分析することができる。コンピュータプログラムは、一般にコードを格納する可読媒体を含む。特定のコードが、プローブ上の各特徴の位置、その特徴での吸着剤のアイデンティティ、および該吸着剤を洗浄するために使用される溶離条件を含むメモリに充てることができる。つぎに、この情報を用いてプログラムは特定の選択性特徴（例えば吸着剤の種類と使用される溶離液）を定義しているプローブ上の特徴の集合を同定することができる。コンピュータはまた、入力としてプローブ上の特定のアドレス指定可能位置から受け取った種々の分子量でのシグナル強度上でデータを受け取るコードを含む。このデータは、検出されたバイオマーカの数を示すことができ、任意に検出された各バイオマーカについてのシグナル強度および測定された分子量を含む。

10

【0064】

データ分析は、検出されたバイオマーカのシグナル強度（例えば、ピークの高さおよびピークの面積）を決定するステップと、「アウターライナー（outer liners）」（所定の統計学的分布から逸脱したデータ）を取り除くステップとを含む。例えば、既に言及したように、観察されたピークが正規化される（なんらかの基準に関連した各ピークの高さが計算されるプロセス）。例えば、基準は、装置および化学物質（例えば、エネルギー吸収分子）によって生ずるバックグラウンド・ノイズであり、スケールでゼロにセットすることができる。つぎに、シグナル強度を各々のバイオマーカについて検出することができ、あるいは他の物質が上記したスケール（例えば、100）での相対的強度のかたちで表示することができる。あるいは、標準は、標準からのピークが各バイオマーカで、あるいは既に述べたようにして検出された他のマーカに関して、観察されたシグナルの強度に関連した計算をおこなうための参照として用いられるようにして、標準を試料で含ませることが可能である。

20

【0065】

コンピュータは、得られたデータを表示するために該データを種々のフォーマットに変換することができる。「スペクトル・ビューまたは保持マップ（spectrum view or retentate map）」と呼ばれる1つのフォーマットでは、標準スペクトル・ビューを表示することができ、該ビューは各々の特定の分子量で検出器に達するバイオマーカの量を表現する。「ピーク・マップ（Peak map）」と呼ばれる別のフォーマットでは、ピークの高さおよび質量に関する情報のみがスペクトル・ビューから保持され、より鮮明な像が生成され、ほとんど同一の分子量でマーカがよりいっそう容易に見られるようになる。さらに別のフォーマットは、「ゲル・ビュー（gel view）」と呼ばれるもので、ピーク・ビューを各ピークの高さに基づいてグレースケール・イメージに変換することができ、それによって電気泳動ゲル上のバンドに類似した外観となる。さらに別のフォーマットは、「3-Dオーバーレイ（3-D overlays）」と呼ばれるもので、相対的なピークの高さでの微妙な変化を調べるために、いくつかのスペクトルをオーバーレイすることができる。さらに別のフォーマットでは、「軟差マップ・ビュー（difference map view）」と呼ばれるもので、2つ以上のスペクトルを比較することができ、試料間で上方制御または下方制御されるユニークなバイオマーカおよびバイオマーカが強調される。任意の2つの試料に由来するバイオマーカ・プロファイル（スペクトル）を視覚的に比較することも可能である。

30

40

【0066】

上記の表示フォーマットのいずれかを使用して、特定の分子量を持っているバイオマーカが試料から検出されたかどうかをシグナル表示から容易に決定されることができる。さらに、シグナル強度から、プローブ面で結合されるマーカの量を測定することができる。

【0067】

50

本発明の好ましい形態では、単一決定木分類アルゴリズムを用いて S E L D I から生じたデータを分析する。そのようなアルゴリズムは、本明細書の実施例 1 でより詳しく説明する。本発明の好ましい形態では、促進決定木アルゴリズムを利用して S E L D I から生じたデータを分析する。そのようなアルゴリズムは、例えば実施例 2 により具体的に記載されている。実施例 2 によりいっそう完全に記載したように、そのようなプロセスによって特異性および選択性が生ずる。

【 0 0 6 8 】

試験試料を、気相イオン分光分析にかける前に処理してもよい。例えば、試料を精製さもなければ予め分画化して、分析のための試料の複雑さを少なくすることができる。試験試料に含まれる生体分子を任意に精製するための手順は、生体分子の特性、大きさ、電荷、および機能に基づくものであってもよい。精製の方法として、遠心、電気泳動、クロマトグラフィー、透析、またはそれらの組み合わせが挙げられる。当業者に公知のように、大きさおよび電荷に基づいて試料内の生体分子を分離するために、電気泳動法を利用することが可能である。電気泳動法は、当業者に周知であり、等電点電気泳動、ドデシル硫酸ナトリウム・ポリアクリルアミドゲル電気泳動 (S D S - P A G E)、アガロース・ゲル電気泳動、および他の公知の電気泳動方法が挙げられる。

10

【 0 0 6 9 】

精製ステップは、クロマトグラフィーによる分画技術によって達成可能であり、サイズ分画、チャージによる分画、分離されている生体分子の他の特性による分画が挙げられる。当業者に公知のように、クロマトグラフィー・システムは固定相と移動相とが含まれ、さらに分離は、異なる相で分離される生体分子間の相互作用にもとづいている。本発明の好ましい形態では、カラム・クロマトグラフィー法を用いることが可能である。そのような方法として、分配クロマトグラフィー、吸着クロマトグラフィー、サイズ排除クロマトグラフィー、イオン交換クロマトグラフィー、およびアフィニティー・クロマトグラフィーが挙げられる。これらの方法は当業者に周知である。サイズ排除クロマトグラフィーでは、分子量が 10,000 Da を上回る分子をサイズ分画化カラムが排除する。

20

【 0 0 7 0 】

本発明の好ましい形態によれば、気相イオン分光分析に先立って、試料を精製し、さもなければ保持 (r e t e n t a t e) クロマトグラフィーによってバイオ・クロマトグラフィー・チップ上で分画化する。好ましいチップは、C i p h e r g e n B i o s y s t e m s , I n c . (P a l o A l t o , C A) から入手可能な P r o t e i n C h i p (登録商標) である。先に述べたように、チップまたはプローブは、質量分析計用に構成されている。チップは、該チップの表面に取り付けられた吸着剤を含む。この吸着剤は、特定の用途で i n s i t u クロマトグラフィー樹脂として機能する。操作の際、溶離液中の吸着剤に対して試料を作用させる。吸着剤が洗浄条件下で親和性を有する分子が、吸着剤に結合する。吸着剤に結合しない分子は、洗浄によって取り除かれる。つぎに、エネルギー吸収分子を吸着剤のスポットに加えて、脱着およびイオン化をさらに促進させることができる。分析物の検出は、吸着剤からの脱着、イオン化、および検出器による検出によっておこなわれる。したがって、保持 (r e t e n t a t e) クロマトグラフィーは、親和性材料によって保持された分析物を検出する点で従来のクロマトグラフィーとは異なるもので、一方従来のクロマトグラフィーでは、親和性材料から溶離した材料が検出される。

30

40

【 0 0 7 1 】

本発明のさらに別の形態では、本発明のバイオマーカーを免疫アッセイ法によって定性的または定量的に検出してもよい。免疫アッセイは、概して、バイオマーカーに特異的に結合さもなければ該バイオマーカーを認識する抗体と試験試料とを接触させ、試料中のバイオマーカーに結合した抗体からなる複合体の存在を検出することが含まれる。免疫アッセイ法は、抗体 / 抗原複合体の認識を伴う当業者に公知の多種多様な免疫アッセイ法から選択することができ、酵素結合免疫吸着アッセイ (E L I S A)、酵素免疫アッセイ (R I A)、ウェスタンブロット、およびマルチプレックス・アッセイの利用が挙げられ、該

50

マルチプレックス・アッセイの利用として、いくつかの所望の抗体がガラス・ビーズまたはプレート等の支持体に配置されて、試験試料と反応さなければ染色する抗体アレイの使用が挙げられる。そのようなアッセイは当業者に周知であり、例えばより完全に、定期的または周期的に更新されている *A Laboratory Manual (1988) by Harlow & Lane; Immunoassays: A Practical Approach, Oxford University Press, Gosling, J. P. (ed.) (2001)* および/または *Current Protocols in Molecular Biology (Ausubel et al.)* に記載されている。

【0072】

本明細書に記載されている免疫アッセイで使用される抗体は、ポリクローナル抗体であってもよく、また当業者二種打ちの方法で得ることが可能である。例えば、精製されたバイオマーカーを種々の動物に注射し、血清中に産生された抗体を単離することが挙げられる。抗体はモノクローナル抗体であってもよい。モノクローナル抗体の産生方法は当業者に周知であり、例えば、精製されたバイオマーカーをマウスに注射し、抗血清を産生する脾臓細胞を単離し、該細胞を腫瘍細胞と融合させることでハイブリドーマを作製し、該ハイブリドーマのスクリーニングをおこなう。バイオマーカーは、本明細書で既に記載したクロマトグラフィー、電気泳動、および遠心技術等、同様に当業者に周知の技術によって、まず最初に精製してもよい。そのような方法は、タンパク質バイオマーカーの大きさ、電荷、溶解性、選択された構成要素に対する結合親和性、それらの組み合わせ、または該タンパク質の他の特徴または特性からなる利点を利用してよい。そのような方法は、当業者に公知であり、例えば、*Current Protocols in Protein Science, J. Wiley and Sons, New York, NY, Coligan et al. (Eds.) (2002; Harris, E. L. V., and S. Angal in Protein purification applications: a practical approach, Oxford University Press, New York, NY (1990)* に見いだすことができる。ひとたび抗体が提供されると、本明細書で既に記載した免疫アッセイによって、バイオマーカーの検出/または計量をおこなうことができる。

【0073】

免疫アッセイのための特定の方法が当業者に周知であるけれども、一般に、免疫アッセイは本明細書に既に記載したように被験体から試料を最初に得ることによって実行することができる。抗体を試験試料に接触させる前に、該抗体を支持体に固定し、抗体/タンパク質バイオマーカー複合体の洗浄およびそれに続く単離を促進させてもよい。固相支持体の例は当業者に周知であり、例えば、マイクロタイター・プレート等の形状となったガラスまたはプラスチックが挙げられる。抗体を、プローブ基質、例えば本明細書に記載した *Protein Chip* (登録商標) アレイに結合させることもできる。

【0074】

試験試料を抗体とインキュベートさせた後に、この混合物を洗浄し、抗体 - マーカー複合体を検出することが可能である。検出は、洗浄混合物と検出試薬とをインキュベートし、例えば色または他の標識 (インジケーター) の発現を観察することで達成することができる。検出試薬は、例えば検出可能な標識でラベルされた二次抗体であってもよい。典型的な検出可能な標識として、磁気ビーズ (例えば *DYNABEADS* (登録商標))、蛍光色素、放射標識、エンザイム (例えばセイヨウワサビ・ペルオキシド、アルカリホスファターゼ、および酵素免疫アッセイ法で一般に使われる他のもの)、および比色ラベル (例えば金コロイド、色ガラス、またはプラスチック・ビーズ) が挙げられる。あるいは、試料中のマーカーは、そこで、間接測定法を使用して検出することができ、例えば、二次標識抗体を用いて結合マーカー特異的抗体の検出に用い、および/または競合または阻害アッセイで、例えばバイオマーカーの異なるエピトープに結合するモノクローナル抗体を混

10

20

30

40

50

合物と同時にインキュベートする。抗体・マーカー複合体の量は、標準と比較することで測定できる。

【0075】

アッセイの全体を通じて、インキュベーションおよび/または洗浄ステップは、試薬の各々の組み合わせの後、必要となる場合もある。インキュベーション・ステップは、約5秒ないし数時間、好ましくは約5分ないし約24時間の範囲内で変動可能である。しかし、インキュベーション時間は、特定の免疫アッセイ、バイオマーカー、およびアッセイ条件に依存する。一般に、アッセイは、一定の温度範囲内、例えば約0ないし約40でおこなわれるか、周囲温度でおこなう。

【0076】

本発明のさらに別の局面では、例えば、本明細書に記載したバイオマーカーを検出するために利用可能なキットが提供される。キットは、例えば、本発明に記載した任意の1種類以上のバイオマーカーを検出するために使うことができ、前立腺癌、両性前立腺肥大の診断または該診断を助けるために、有利に利用することが可能である。

【0077】

一実施形態では、キットは、好ましくは本明細書に記載した1種類以上のタンパク質バイオマーカーと結合するのに適している吸着剤を含む基質と、試験試料を吸着剤と接触させて該吸着剤に保持されたタンパク質バイオマーカーを検出することでタンパク質バイオマーカーを検出するためのインストラクションを含む。特定の実施形態では、キットは溶離液または溶離液を作るインストラクションを含むものであってもよく、溶離液と吸着剤との組み合わせが、気相イオン分光分析の使用等によって、タンパク質バイオマーカーの検出を可能とする。そのようなキットは、本明細書に記載されている材料から調製される。さらに別の実施形態では、キットは吸着剤を含む第1の基質（例えば、吸着剤によって機能が付与された粒子）と、第1の基質を置くことができる第2の基質とを含み、気相イオン分光器に着脱自在に挿入可能なプローブを形成する。別の実施形態では、キットは基質状の吸着剤を持つ着脱自在に挿入可能なタンパク質の形態の単一基質を含む単一の基質であってもよい。さらに、別の実施形態によれば、キットはさらにプレ分画スピン・カラム（例えば、K-30サイズ排除カラム）をさらに含むものであってもよい。

【0078】

キットは、さらに、ラベルまたは別体の挿入物の形で適当な操作パラメータのインストラクションを含むものであってもよい。例えば、キットは標準の指示をサンプルの特定の形状の後、プローブを洗う方法がプローブで接触することを消費者または他個人に知らせる標準的なインストラクションを持つものであってもよい。さらなる例として、キットは試料をプレ分画化して試料中に含まれるタンパク質の複雑さを減少させる。

【0079】

さらなる実施形態で、キットはマーカーおよび検出試薬に特異的に結合する抗体を含むものであってもよい。そのようなキットは本発明に記載した材料から調製することができる。さらにキットは、上記したように、プレ分画化スピン・カラムを含むものであってもよく、同様にラベルまたは別体の挿入物の形で適当な操作パラメータに適したインストラクションを含む。

【0080】

本発明のさらに別の局面では、前立腺癌、良性前立腺肥大の疑診を下すために、あるいは陰性診断を下すために、複数の分類子を用いる方法が提供される。本発明の一形態では、方法は、(a)正常な被験体、前立腺癌と診断された被験体、および良性前立腺肥大と診断された被験体から得た複数の試料から質量スペクトルを得ること、(b)ピーク強度値と関連閾値とを含む複数の重みつき基準分類子を得るために、質量スペクトルの少なくとも一部分に対して促進決定木を適用すること、(c)上記複数の重みつき基準分類子の線形結合にもとづいて前立腺癌および良性前立腺肥大の少なくとも1つの疑診または陰性診断を下すこと、を含む。本発明の特定の形態で、方法は前立腺癌、良性前立腺肥大のうちの少なくとも1つの疑診、あるいは陰性診断を下すためにピーク強度値と関連閾値と

10

20

30

40

50

線形結合の状態を使用することを含む。好ましいアルゴリズムおよびデータ処理を、実施例 2 により完全に記載する。

【0081】

(コンピュータ・インプリメンテーション)

図 5 に示すもの等のコンピュータ処理システム 104 で、本発明の技術を実施することが可能である。この点では、図 5 は、少なくとも 1 つの本発明のコンピュータ実施形態にしたがってコンピュータ処理のいくつかまたは全てを実施することもできるコンピュータ・システム 104 の説明図である。

【0082】

外観から見て、図 5 では、参照番号 104 で示すコンピュータ・システムは、該コンピュータ・システムに収容可能な多数のディスク・ドライブを象徴しているにすぎないドライブ 502 および 504 を備えるコンピュータ部 112 を有する。通常、これらは、フロッピー（登録商標）・ディスク・ドライブ 502、ハード・ディスク・ドライブ（外観は図示せず）、および CDROM 504 を含むことが可能である。ドライブの数および種類は、通常各種のコンピュータ構成によって異なる。実際、ディスク・ドライブ 502 および 504 は任意であり、スペースを考慮して、コンピュータ・システムから省略することができる。

【0083】

コンピュータ・システム 104 は、任意のディスプレイ・モニタ 110 も有し、該ディスプレイ・モニタ上で、正常または異常な、正常と疑われる、あるいは異常と疑われる細胞等に関する視覚情報を表示することができる。いくつかの状況では、入力装置としてキーボード 116 およびマウス 114 が設けられており、それらを介して入力を実行ことが可能であるため、中央処理装置（CPU）604 とともにインタフェースへの入力が可能になる（図 6）。その反面、移植性を高めるために、キーボード 116 を限定機能キーボードにするか、または完全に省略することができる。さらに、マウス 114 は、任意で、タッチ・パッド制御装置またはトラック・ボール装置であり、あるいは完全に省略されることもあり、同様に入力装置として使用されることが可能である。さらに、コンピュータ・システム 104 は、任意で、赤外線信号を送信および/または受信するための少なくとも 1 つの赤外線（または無線）送信機および/あるいは赤外線（または無線）受信機を含むことも可能である。

【0084】

単一のプロセッサ、単一のハード・ディスク・ドライブ 614、および単一のローカル・メモリを有するコンピュータ・システム 104 を図示しているが、任意で、システム 104 は、任意の多数のまたは組み合わせたプロセッサまたは記憶装置を適当に備える。コンピュータ・システム 104 は、実際、本発明の原理にしたがって動作する任意の適当な処理システムと交換または組み合わせることができ、該処理システムとしては、携帯用ラップトップ/ノートブック、小型コンピュータ、大型コンピュータ、およびスーパー・コンピュータと共に、その処理システム・ネットワークの組み合わせが挙げられる。

【0085】

図 6 は、図 5 のコンピュータ・システム 104 の内部ハードウェアのブロック線図を示す。バス 602 は、コンピュータ・システム 104 の他の構成要素を関係させる主要情報ハイウェイとして機能する。CPU 604 は、該システムの中央処理装置であり、プログラムを実行するために必要な計算および論理演算を実行する。読み出し専用メモリ（ROM）606 およびランダム・アクセス・メモリ（RAM）608 は、コンピュータ・システム 104 の主記憶装置を構成する。ディスク・コントローラ 610 は、1 つ以上のディスク・ドライブをシステム・バス 602 に接続する。これらのディスク・ドライブは、例えば、502 等のフロッピー（登録商標）・ディスク・ドライブ、CDROM または DVD（デジタル・ビデオ・ディスク）ドライブ 504、あるいは内部または外部ハード・ディスク 614 である。先に示したように、これらの様々なディスク・ドライブおよびディスク・コントローラは、任意の装置である。

10

20

30

40

50

【0086】

ディスプレイ・インタフェース618は、ディスプレイ110のインタフェースとなり、バス602からの情報をディスプレイ110上で表示することを可能にする。すでに示したように、ディスプレイ110も任意の付属品である。例えば、ディスプレイ110を置換または省略することが可能である。外部装置、例えば本明細書に記載されるシステムの他の構成要素との通信は、通信ポート616を用いて生じる。例えば、光ファイバーおよび/または電気ケーブルおよび/または心線および/または光通信(例えば赤外線等)および/または無線通信(高周波(RF)等)を、外部装置と通信ポート606との間の伝送媒体として用いることができる。周辺インタフェース620は、キーボード116およびマウス114のインタフェースとなり、入力データをバス602に送信することを可能にする。 10

【0087】

他の実施形態では、上記に示したCPU604を任意の他の適当な処理回路と交換または組み合わせることができ、該処理回路としては、PAL(プログラマブル・アレイ・ロジック(programmable array logic))およびPLA(プログラマブル・ロジック・アレイ(programmable logic array))、DSP(デジタル・シグナル・プロセッサ(digital signal processor))、FPGA(フィールド・プログラマブル・ゲート・アレイ(field programmable gate arrays))、ASIC(特定用途向け集積回路(application specific integrated circuit))、VLSI(超大規模集積回路(very large scale integrated circuit))等のプログラム可能論理装置が挙げられる。 20

【0088】

任意の現在利用可能なまたは将来開発されるコンピュータ・ソフトウェア言語および/またはハードウェア構成要素を本発明のこのような実施形態で用いることができる。例えば、拡張可能マークアップ言語(Extensible Markup Language(XML))、HTML、ビジュアル・ベーシック(Visual Basic)、C、C++、または使用するプロセッサの観点から適当な任意のアセンブリ言語を用いて、上述の機能性の少なくともいくつかを実施することができる。それを、Java(登録商標)等のインタプリタ環境で書き込むことおよび様々なユーザへの複数の宛先に伝送することも可能である。 30

【0089】

本発明のインプリメンテーションのうちの1つは、命令セットとして、上記に記載したような一般的に構成される1つ以上のコンピュータ・システム104のランダム・アクセス・メモリ608に常駐している。コンピュータ・システム104が要求するまで、別のコンピュータ読み込み可能メモリ、例えばハード・ディスク・ドライブ614に、あるいはCD-ROM504での最終的な使用のための光ディスク等の着脱可能メモリにまたはフロッピー(登録商標)・ディスク・ドライブ502での最終的な使用のためのフロッピー(登録商標)・ディスク(例えば図7のフロッピー(登録商標)・ディスク702)に、該命令セットを保存することが可能である。さらに、命令セット(Java(登録商標)、HTML、XML、標準一般化マークアップ言語(Standard Generalized Markup Language(SGML))、および/または構造化問い合わせ言語(Structured Query Language(SQL))に書き込まれるもの等)を、ユーザが所望するとき、別のコンピュータのメモリに保存すること、およびローカル・エリア・ネットワークまたは広域ネットワーク(例えばインターネット)等の伝送媒体を介して送信することができる。当業者には、コンピュータ・プログラム媒体の保存または伝送は、媒体がコンピュータ読み込み可能情報を保持するように電氣的に、磁氣的に、または化学的に媒体を変化させることが既知である。 40

【0090】

任意のバイオマーカーはそれぞれに、前立腺癌状態の決定を助ける上で有用である。初 50

めに、本明細書に記載される方法、例えば S E L D I バイオチップ上での捕獲後の質量分析法による検出を用いて、選択したバイオマーカーを被験体試料で測定する。その後、前立腺癌状態と非癌状態を識別する診断量または対照と測定値を比較する。診断量は、本明細書で、特定のバイオマーカーが非癌状態に比べて癌状態で上方制御または下方制御されるという情報を反映する。当技術で十分に理解されているように、使用する特定の診断量を調整して、診断者の選好に応じて診断アッセイの感度または特異度を増加させることができる。したがって、診断量と比較した試験量は、前立腺癌状態を示す。

【 0 0 9 1 】

個々のバイオマーカーが有用な診断マーカーである一方で、バイオマーカーの組み合わせによって、単一マーカー単独よりも優れた予測値が得られることが判っている。特に、試料中の複数のマーカーの検出は、真陽性および真陰性診断の割合を増加させ、偽陽性または偽陰性診断の割合を減らす。したがって、本発明の好ましい方法は、1つより多いバイオマーカーの測定を含む。その後、考えられるヒト癌診断とマーカーまたはマーカー群の検出を相関させる。

10

【 0 0 9 2 】

他の実施形態では、マーカーの測定は、マーカーを定量して、考えられる癌診断とマーカーの検出を相関させることを含むことができる。したがって、試験される被験体で検出されたマーカー量が対照量と比べて異なる（すなわち、マーカーによって、対照より高いまたは低い）場合、試験される被験体は、癌を有する確率がより高い。

【 0 0 9 3 】

相関は、マーカーまたはマーカー群（マーカーまたはマーカー群の上方または下方制御）（例えば、ヒト癌が検出されない正常被験体中の）の対照量と比べた試料中のマーカーまたはマーカー群の量を考慮に入れることも可能でありうる。対照は、例えば、ヒト癌が検出されない正常被験体の比較可能試料に存在する平均または中位マーカー量である。試験量を測定する際と同じまたは実質的に類似の実験条件下で対照量を測定する。相関は、試験試料中のマーカーの有無および対照中の同マーカーの検出の頻度を考慮に入れることも可能である。相関は、そのような因子の両方を考慮に入れて、前立腺癌状態の決定を容易にすることが可能である。

20

【 0 0 9 4 】

癌状態を定量する方法の特定の実施形態では、該方法は、上記状態にもとづく被験体治療を管理することをさらに含む。前述のように、このような管理は、癌状態を決定した後の医師または臨床家の処置を含む。例えば、本発明の方法の結果が決定的ではないか、または状態の確認を必要とする理由がある場合、医師はより多くの試験を指示することが可能である。また、状態によって外科手術が適当であることが示される場合、医師は、患者に外科手術を計画することが可能である。他の例では、外科手術に代えてまたは加えて、患者が化学療法または放射線治療を受けることが可能である。同様に、結果が否定的である場合、例えば状態が末期の癌を示す場合、あるいは状態が急性である場合、それ以上処置の根拠とはなりえない。さらに、結果によって治療が成功したことが示される場合、それ以上の管理は必要ではないであろう。

30

【 0 0 9 5 】

本発明は、被験体管理後にバイオマーカー（またはバイオマーカーの特定の組み合わせ）を再び測定する方法も提供する。このような場合、該方法を用いて、癌の状態、例えばがん治療への応答、該疾患の一時的な小康、または該疾患の進行をモニタリングする。該方法の使用の容易さおよび該方法に侵襲性が無いことから、患者が受ける各治療の後に、該方法を繰り返すことができる。これによって、医師は、一連の治療の効果を追うことが可能になる。結果によって治療が効果的ではないことが示される場合、一連の治療をそれに応じて変更することができる。これによって、医師は、治療の選択肢で臨機応変に対応することが可能になる。

40

【 0 0 9 6 】

別の例では、マーカーを検出するための方法を用いて、これらのマーカーの発現を調節

50

する化合物を生体内 (*in vivo*) または生体外 (*in vitro*) でアッセイおよび同定することができる。

【 0 0 9 7 】

本発明の方法は他の用途も有する。例えば、マーカーを用いて、マーカーの発現を調節する化合物を生体外 (*in vitro*) または生体内 (*in vivo*) でスクリーニングすることができ、この化合物が今度は、患者の癌を治療または予防する上で有用となる。

【 0 0 9 8 】

別の例では、マーカーを用いて、癌に対する治療への応答をモニタリングすることができる。

10

【 0 0 9 9 】

さらに別の例では、マーカーを遺伝研究で用いて、被験体が癌を発症する危険があるかを決定することができる。例えば、特定のマーカーを遺伝的に関連付けることが可能である。例えば家族が前立腺癌の病歴を持つ前立腺癌患者の集団から得た試料を分析することによって、これを決定することができる。その後、例えば家族が前立腺癌の病歴を持たない癌患者から得たデータと結果を比較することができる。遺伝的に関連付けたマーカーをツールとして用いて、家族が前立腺癌の病歴を有する被験体が前立腺癌に罹患しやすいかを決定することが可能である。

【 0 1 0 0 】

ここで、上記のバイオマーカー、キット、コンピュータ・プログラム媒体、および方法を説明する特定の実施例に言及する。実施例は、好ましい実施形態を説明するために提供されるものであり、それによって、本発明の範囲を限定することを意図したものではないことが理解される。

20

【 実施例 】

【 0 1 0 1 】

(実施例 I)

(前立腺癌または良性前立腺肥大症を診断するための血清タンパク質指紋法およびパターン照合アルゴリズムの使用)

(材料および方法)

(血清試料)

30

血清試料は、バージニア・プロステート・センター・ティッシュ (Virginia Prostate Center Tissue) およびボディ・フルード・バンク (Body Fluid Bank) から得た。血清入手、データ管理、および血液採取プロトコールは、東バージニア医科大学倫理調査委員会 (Eastern Virginia Medical School Institutional Review Board) によって認証された。PCAまたはBPHであると診断された患者由来の血液試料は、東バージニア医科大学泌尿器科部 (Department of Urology , Eastern Virginia Medical School) から調達し、健康者 (HM) コホートは、一般公衆に開放されている無料スクリーニング診療所から得た。PCAまたはBPHの診断時に得た処理前の試料のみをこの研究に用いた。インフォームド・コンセント後に、10ccセラム・セパレータ・バキュテイナー・チューブ (Serum Separator Vacutainer Tube) 中に試料を採取し、30分後に3750 x 100 rpmで5分間遠心した。血清を500ulアリコートに分配して、-80 で凍結保存した。同齡HM群の各検体から等量の血清をプールして、100ulアリコートを-80 で保存することによって精度管理試料を調製した。再現性を決定するために、かつ各SELDI実験用の対照タンパク質プロファイルとして精度管理 (QC) 試料を用いた。

40

【 0 1 0 2 】

(患者およびドナー・コホート)

患者の4つの群から得た検体をこの研究で用いた。すなわち、同齡HM (対照) 96人

50

、良性前立腺肥大症 (BPH) 92人、臓器限局PCA (T1、T2)と診断された患者99人、および非臓器限局 (T3、T4) PCAと診断された98人である。ドナーは、正常な直腸指診、すなわちPSA < 4.0 ng/mlを有し、かつ前立腺疾患の証拠がなければHMとして選択された。HM群は、51歳から70歳 (平均60歳)の年齢範囲の白人男性48人およびアフリカ系アメリカ人男性48人からなっていた。BPH患者群には、48歳から86歳 (平均67歳)の年齢範囲の白人33人、アフリカ系アメリカ人2人、および人種不明57人であった。BPH患者は、4ないし10のPSA値、すなわち低PSA速度を有し、かつ複数の陰性の生検であった場合に選択された。臓器限局 (T1、T2) PCA群は、50歳から89歳 (平均71歳)の年齢範囲の白人76人、アフリカ系アメリカ人20人、アジア人1人、および人種不明2人からなっていた。非臓器限局PCA群 (T3、T4) に対しては、44歳から87歳 (平均69歳)の年齢範囲の白人80人、アフリカ系アメリカ人16人、および人種不明2人であった。HM群に関する範囲および平均PSA値は、0.15 ~ 3.83 ng/ml (1.32 ng/ml)であり、BPH群に関しては0.0 ~ 10.91 ng/ml (4.60 ng/ml)、PCA T1、T2群に関しては0.0 ~ 95.16 ng/ml (10.10 ng/ml)、PCA T3、T4群に関しては0.0 ~ 875.2 ng/ml (206.93 ng/ml)であった。

10

【0103】

(SELDIタンパク質プロファイリング)

様々なチップ化学 (疎水性、イオン、陽イオン、および金属結合) を初めに評価して、どの親和性化学がタンパク質数および分解能の点から最良の血清プロファイルを提供するかを決定した。IMAC-Cu金属結合チップが最良の結果を与えると観察された。IMAC-3チップ (Ciphergen Biosystems, Inc, Fremont CA) を、各アレイ上で100mM CuSO₄ (20 μl) で被覆し、トミー・マイクロ・チューブ・ミキサー (TOMY Micro Tube Mixer) (MT-360, Tomy Seiko Co., Ltd) に入れて、5分間攪拌した。該チップを脱イオン (DI) 水で10回すすぎ、100mM酢酸ナトリウム20 μlを各アレイに添加して、5分間振とうし、未結合の銅を除去した。チップを再びDI水ですすいで (x10)、バイオプロセッサ (Ciphergen Biosystems, Inc.) に入れた。該バイオプロセッサは、12個のチップを保有し、各チップ・アレイにより大量の血清を適用することを可能にする装置である。バイオプロセッサを洗浄し、プラットフォーム・シェーカー上で、各ウエルでPBS 200 μlとともに速度250 rpmで5分間振とうした。これをさらに2回繰り返し、ペパータオルの上でバイオプロセッサを反転させることによって1回ごとにPBS緩衝液を破棄した。1.5 ml微量遠心管内のPBS中で、8M尿素/1% CHAPS 30 μlとともに血清20 μlを4で10分間ボルテックスすることによって、SELDI分析用の血清試料を調製した。0.125% CHAPSと合わせて100 μlの1M尿素を血清/尿素混合物に添加して、短時間ボルテックスした。PBSを添加して、1:5で希釈し、タンパク質チップ・アレイに適用するまで氷上に置いた。希釈した血清/尿素混合物50 μlを各ウエルに適用し、バイオプロセッサを密閉して、プラットフォーム・シェーカー上で、速度250 rpmで30分間振とうした。血清/尿素混合物を破棄して、PBS洗浄ステップを3回繰り返した。チップをバイオプロセッサから取り出し、DI水で10回洗浄し、風乾させて、SELDI質量分子器にかけるまで室温で暗所に保存した。SELDI分析前に、50% (v/v) アセトニトリル、0.5% トリフルオロ酢酸中のエネルギー吸収分子 (EAM) シナピン酸の飽和溶液0.5 μlを各チップ・アレイ上に2回使用して、シナピン酸の各使用の間、アレイ表面を風乾させた。プロテイン・バイオリジカル・システムII (Protein Biological System II) 質量分析計読み取り装置 (Ciphergen Biosystems, Inc.) にチップを入れ、レーザー強度220、検出器感度7、および900 nsの集束遅れ時間でのポジティブ・モードで集められる192のレーザー照射 (ショット) を平均化することによって飛行時間スペクトルを生じさ

20

30

40

50

せた。オール・イン・ワン・ペプチドMWスタンダード(All-in-1 peptide MW standard)(Ciphergen)を用いて、質量精度を外部から校正した。

【0104】

(データ分析)

この研究で用いられるデータ分析プロセスは、3つの段階を含む。すなわち、(1)ピーク検出およびアラインメント、(2)最も高い識別能力を用いたピークの選択、および(3)決定木(Decision Tree)アルゴリズムを用いたデータ分析である。4つの層(PCA(T1/T2)、PCA(T3/M1)、BPH、HM)で層化無作為抽出を用いて、分析前に全データ・セットを訓練および試験データ・セットに分離した。訓練データ・セットは、167のPCA、77のBPH、および82の正常血清試料由来のSELDIスペクトルからなっていた。その後、分類アルゴリズムの有効性および精度を、30のPCA、15のBPH、および15の正常試料からなる盲検データ・セットを用いて調べた。

10

【0105】

(ピーク検出)

サイファージェン(Ciphergen)SELDIソフトウェア・バージョン3.0ベータおよび3.0(インターネット・アドレス:www.chiphergen.com)を用いて、ピーク検出を実行した。2000から4000Daの質量範囲は、分解されたタンパク質/ペプチドの大多数を含むため、この範囲を分析用に選択した。0から2000Daの分子質量は、EAMの付加物および人工物を含み、恐らく他の化学的不純物を含むので、この領域を分析から排除した。ピーク検出は、(1)基線減算、(2)質量精度校正、(3)自動ピーク検出、および(4)ピーク・アラインメント決定を含んだ。ソフトウェア・プログラムは、データ分析用の演算子によって選択される基準にもとづいてノイズ、ピーク領域、およびフィルタを算出する。この研究のために用いられる設定は、適応ウインドウ幅=100データ・ポイント、平均ノイズ=10ポイント、傾きにもとづく方法を用いてピーク領域を算出、低最小谷深度=ノイズの10倍、高最小谷深度=ノイズの0.5倍、ピーク高度の低感度および高感度=それぞれノイズの10倍および2倍、自動ピーク検出スライダ=質量範囲2から4KDaに対して8、質量範囲4から8KDaに対して11、および質量範囲8から40KDaに対して8であった。平均80のピークを各スペクトルで検出した。

20

30

【0106】

(ピーク・アラインメント)

772のスペクトルからのラベルつきピーク(総数63,157)全てをSELDIからエクセル表計算ソフトにエクスポートした。組織内で開発されたピークマイナー(Peak Miner)アルゴリズム(インターネット・アドレス:www.evms.edu/vpc/seld)を用いて、低質量から高質量までの質量値にもとづいて全てのピークを分類した。質量誤りスコア、すなわちピークXとピークX+1との間の質量差の測定を、 $(M_{pX} - M_{pX+1}) / M_{pX}$ を用いて各ピークに対して算出し、この際 M_{pX} は、ピークXの質量値である。例えば、質量誤りスコアが0.18%未満であった場合、ピーク値を平均化することによってピークXおよびピークX+1を1つのピークに位置合わせ(アライン)した。質量誤りスコアが0.18%より大きかった場合、ピークXおよびピークX+1を2つの異なるピークとみなした。これは、ピーク・アラインメントを決定するためにラベルつきピーク全てにわたって反復するプロセスであり、各ピーク質量に対応するピーク強度によって全ての試料を記録する。

40

【0107】

(特徴選択)

0.5(区別能力なし)から1.0(完全分離)の範囲であるAUCを評価することによって、正常対PCA、正常対BPH、およびBPH対PCAを区別する際の各ピーク的能力を決定した。

50

【0108】

(決定木分類)

326の試料(82の正常、77のBPH、および167のPCA)からなる訓練データ・セットを用いて、Breiman et al(14)によって記載されるように、かつ改変して(すなわち、基準として負の対数尤度を用い、データ整理のために0.62のAUCを用いて、779のピークから124のピークを選択した)決定木分類アルゴリズムの構築を実行した。分類木は、質問形態で1度に1つの規則を用いて、データ・セットを2つの瓶(bin)または節(node)に分割する。分割決定は、1つのピークの有無および強度レベルによって定義される。例えば、「質量Aは、X以下の強度を有するか」に対する回答は、データ・セットを2つの節、すなわち肯定の返事に対しては左の節かつ否定の返事に対しては右の節に分割する。この「分割(splitting)」プロセスは、末端の節または葉(leaf)が生成されるか、または分割がそれ以上増加しなくなるまで続く。その節中の試料の大多数を表す試料(すなわち、PCA, BPH, 正常)の群(「クラス」)によって末端の節の分類を決定する。各節の異質性を反映する「コスト(cost)」関数を算出する。すなわち $-\log L = -\sum_j n_j \log(p_j)$ であり、式中、Lは、多項分布の尤度であり、 n_j は、クラスj中の試料の数であり、 p_j は、クラスjの確率である。分割規則を形成するためにこのプロセスによって選択されるピークは、2つの派生する節中のコストの最大削減を達成するものである。

10

【0109】

(統計的分析)

AUCを計算して、疾患個体から得た試験結果が、非疾患個体から得たものよりも疾患の表示である確率にもとづいて、3つの群を区別する最も高い潜在能力を有するピークを同定した(15)。ベイズの手法(Bayesian approach)を用いて、各末端節中の各クラスの予想確率を算出し(16)、事後ディリクレ分布(posterior Dirichlet distribution)を用いてそれらの95%信頼区間を算出した(16)。事後ディリクレ分布(posterior Dirichlet distribution)に対して4000の試料を生成および分類することによって、95%信頼区間を算出し、100番目および3900番目の試料それぞれを95%信頼区間の下限および上限としてみなした。正確に分類された非疾患試料数と非疾患試料の総数との比率として特異度を算出した。正確に分類された疾患試料数と疾患試料の総数との比率で感度を算出した。真PCA陽性数と偽PCA陽性数との合計で真PCA陽性数を割ってPPVを算出した。偽陰性数と真陰性非疾患試料(BPH/HM)数との合計で真陰性非疾患試料(BPH/HM)数を割ってNPVを算出した。

20

30

【0110】

(結果)

(データ分析)

772のスペクトル(約81のピーク/スペクトルを有する386のスペクトルを二重)の分析に続いて、SELDIソフトウェア・プログラムを用いたピーク検出は、2~40KDa質量範囲で63,157のピークを検出した。これらのうち、クラスタリングおよびピーク・アラインメント・プロセス後に779のピークを同定した。779のピークのそれぞれに対してAUCを算出した。1.0のAUCを有する単一ピークは同定されず、単独で2つの群(すなわち、HM対PCA、またはHM対BPH、またはBPH対PCA)または3つの群(PCA対BPH対HM)を完全に分離することが可能なピークが検出されなかったことを示した。779のピークのうち、124は0.62以上のAUCを有していた。0.62より下のAUCを有するものは、分類用には不適切であるとみなした。その後、訓練セットで同定されたこれら124のピークを用いて、決定木分類アルゴリズムを構築した。図1は、ピーク検出から試料分類までのプロセスを概説するフロー・ダイアグラムである。分類アルゴリズムは、4から10KDa間の9つの質量を用いて、10の末端節(L1~L10)を生成した(図2A)。一旦アルゴリズムが最大識別ピークを同定したら、分類規則は極めて単純である。例えば、未知試料が質量7819.75

40

50

(「根 (root)」節) でピークを持たないが、質量 7024.02 でピークを有する場合、試料は終端節 L1 に位置し、PCA として分類される。試料が L2 に位置する場合は、BPH に割り当てられる。この分割プロセスの別の例を図 2 B に示しており、この際 5 から 10 KDa 間の 4 つの質量を用いて、167 の PCA 試料のうち 46 を終端節 L7 に割り当てる。現実の確率的性質にもとづいて、1 つの試料種類のみを含む純粋な節、例えば BPH 試料のみを含む L2 に対してでさえも、新規試料の誤分類を排除することはできない。未知試料が正確に分類または誤分類されるかどうかの考察を得るために、10 個の終端節中の各群に対して、予想確率および 95% 信頼水準を算出し、以下の表 1 に示す。

【0111】

10

(表 1 10 個の終端節に割り当てられたクラスのそれぞれに対する予想確率および 95% 信頼水準)

【0112】

【表 1】

節	クラス	観察	確率	5%信頼水準
L1	正常	1	0.0625	0.0081, 0.1693
	PCA	27	0.8750	0.7423, 0.9630
	BPH	1	0.0625	0.0087, 0.1698
L2	正常	0	0.0167	0.0005, 0.0584
	PCA	0	0.0167	0.0004, 0.0628
	BPH	57	0.9667	0.9072, 0.9952
L3	正常	1	0.2000	0.0247, 0.4793
	PCA	5	0.6000	0.3027, 0.8592
	BPH	1	0.2000	0.0248, 0.4753
L4	正常	0	0.0714	0.0018, 0.2509
	PCA	0	0.0714	0.0019, 0.2579
	BPH	11	0.8571	0.6311, 0.9823
L5	正常	0	0.1429	0.0040, 0.4725
	PCA	4	0.7143	0.3557, 0.9567
	BPH	0	0.1429	0.0040, 0.4504
L6	正常	74	0.9494	0.8950, 0.9858
	PCA	2	0.0380	0.0082, 0.0879
	BPH	0	0.0127	0.0003, 0.0459
L7	正常	0	0.0204	0.0005, 0.0738
	PCA	46	0.9592	0.8893, 0.9951
	BPH	0	0.0204	0.0005, 0.0726
L8	正常	4	0.5556	0.2458, 0.8416
	PCA	2	0.3333	0.0836, 0.6544
	BPH	0	0.1111	0.0032, 0.3732
L9	正常	0	0.1429	0.0034, 0.4560
	PCA	0	0.1429	0.0037, 0.4830
	BPH	4	0.7143	0.3354, 0.9566
L10	正常	2	0.0337	0.0068, 0.0784
	PCA	81	0.9213	0.8595, 0.9674
	BPH	3	0.0449	0.0123, 0.0964

20

30

【0113】

例えば、L2 で誤分類される HM および PCA 試料に対する予想確率は 1.67% である。ゼロではないが、この節に割り当てられる HM または PCA 試料の尤度は非常に低い。一方で、BPH は、L2 に正確に分類される 96.67% の確率を有する (90.72% ないし 99.52% の 95% 信頼区間を持って)。試料の不正確な割り当ての確率は、例えば終端節 L3、L5、および L9 のように、大多数の試料をほとんど含まない節で、または少数の試料のみが節に割り当てられるときに増加する (図 2 A)。

40

【0114】

10 個の終端節からの分類結果の合計を下記に示す表 2 で訓練および試験セットに対して提示する。

【0115】

(表 2 前立腺訓練および試験セットの決定木分類)

【0116】

【表 2】

A. 訓練セット

サンプル	正常		BPH		PCA		誤分類率	
正常 (N=82)	78	95.12%	0	0.00%	4	4.88%	4	4.88%
BPH (N=77)	0	0.00%	72	93.51%	5	6.49%	5	6.49%
PCA 病期 T1,T2 N=84	2	2.38%	0	0.00%	82	97.61%	2	2.38%
PCA 病期 T3,T4 N=83	2	2.40%	0	0.00%	81	97.59%	2	2.41%
全サンプル (N=326)							13	3.99%

10

B. 試験セット

サンプル	正常		BPH		PCA		誤分類率	
正常 (N=15)	15	100.00%	0	0.00%	0	0.00%	0	0.00%
BPH (N=15)	0	0.00%	14	93.33%	1	6.67%	1	6.67%
PCA 病期 T1,T2 (N=15)	3	6.67%	0	0.00%	12	80.00%	3	20.00%
PCA 病期 T3, T4 (N=15)	1	6.67%	1	6.67%	13	86.67%	2	13.33%
全サンプル (N=60)							6	10.00%

20

C. 盲検セットでの非疾患からの前立腺疾患の区別

疾患/非疾患	パーセント陽性(陽性数/試験数)					
	PCA/HM	PCA/BPH	PCA/(BPH/HM)	BPH/HC	BPH/T1,T2	T1,T2/T3,T4
感受性	83(25/30)	83(25/30)	83(25/30)	93(14/15)	93(14/15)	80(12/15)
特異性	100(15/15)	93(14/15)	97(29/30)	100(15/15)	80(12/15)	87(13/15)

30

【0117】

分類アルゴリズムは、96%の全体的な正確分類に対して、訓練セットの3つの群のそれぞれで試料の93.51%から97.59%を正確に予測した(表2A)。該アルゴリズムは、正確に分類されるHM由来の15の試料すべてを有する試験試料の90%(54/60)、BPH試料の93%(14/15)、およびPCA試料の83%(25/30)を正確に予測した(表2B)。非疾患群から疾患を区別するための分類システムの感度および特異度を表2Cに提示する。PCA対非癌(BPH/HM)を比較すると、感度は83%(25/30)であり、かつ特異度は97%(29/30)であった。PCA対HM(25/30)またはPCA対BPH(25/30)を比較した際、83%の感度も得た一方で、特異度は、PCA対HMに対して100%(15/15)、PCA対BPHに対して93%(14/15)であった。分類システムのPPVは、96.15%であり、NPVは96.67%であった。

40

【0118】

(再現性)

SELDISペクトルの再現性、すなわち、単一チップ上のアレイからアレイまでの(イントラアッセイ)および複数のチップ間(インターアッセイ)の質量位置および強度を、プールした正常血清QC試料を用いて測定した。研究過程にわたって無作為に選択した

50

スペクトル上で観察された3000から10000Da範囲の7つのタンパク質を用いて、変動係数を算出した。ピーク位置に対するイントラアッセイおよびインターアッセイCVは0.05%であり、正規化強度(ピーク高度または相対的濃度)に対しては、それぞれ15%および20%であった(データは図示せず)。スペクトル間の0.18%質量精度内にある質量は同じものであるとみなした。最も重要なことは、SELDIを実行する者には見せないで無作為に選択した試料は、数月の再実行または1年後でさえ、決定木分類アルゴリズムによって正確に分類されたという観察であった(図2C)。

【0119】

(考察)

前立腺癌のための現行の標準的スクリーニング手法は、前立腺特異的抗原(PSA)に関する血清試験であり、該試験が陽性の場合、前立腺の各葉から生検を入手する。PSA試験は感度が>90%であるが、その特異度は、25%のみである。この低特異度によって、人は生検を受ける上に、生検により検出可能なPCAを持たない際は相当の懸念が生じる。SELDIプロファイリング分類手法を用いて、BPHおよび同齡非罹患健康者から前立腺癌を区別する上で全体的な感度83%、特異度97%、および陽性予測値96%が得られた。より大きなかつより臨床的に多様な研究セットを用いてこのSELDIプロファイリング分類システムを検証することができれば、この手法は、不必要な生検数を減少する上で即座のかつ実質的な利益を有する。

【0120】

盲検セットに対して高陽性予測値(PPV)(96%)を達成する好結果の当診断システム開発は、無作為に選択した試料の大規模なかつ慎重に選択した訓練セットを用いることにもとづく。全ての検体は厳密に同齡であった。陰性DREおよびPSA<4.0ng/mlを有するとして同定された非罹患HM由来の血清試料を、無料の前立腺スクリーニング臨床の間に一般集団から入手した。BPH患者の大多数は、4~10ng/ml PSAおよび複数の陰性生検を有し、PCA患者は、少量の限局性疾患から局所および遠隔の転移性疾患までの範囲の癌と、0から8000ng/mlを超えるまでと多様なPSA値を有していた。好結果の分類システムを構築する際の別の重要な因子は、質量分析機器の特徴である「ノイズ(noise)」(アッセイに導入されるEAMおよび化学的不純物によって生成される偽のシグナル)、天然のランダムな日常の変動、および試料から試料への検証を選別除去することができるアルゴリズムを用いることであった。例えば真のペプチド/タンパク質ピークから人工物が原因のピークを識別する上で、この「正規化(normalization)」プロセスを用いることができる。上記の例での癌と非癌コホートとの間のタンパク質変化の大部分が過剰または減少発現するタンパク質に基づき、かつ該タンパク質が存在または被存在するかどうかのみには基づかないと考えられるとき、正規化プロセスはさらにより重要になる。発明者らは、正規化関数を有する精度のよいかつ再現可能な特徴選択またはピーク「ピッキング(picking)」アルゴリズムがSELDIプロファイリング・データに対する好結果の分類アルゴリズムを開発する上で重要な第1歩であると考えている。

【0121】

9つのペプチド/タンパク質質量の過剰または減少発現にもとづいて3つの研究コホートを分離可能であることが奨励される。当研究所からの先行研究(12)は、発明者らが認識する範囲では、潜在的診断手法としてSELDIタンパク質プロファイリングの概念を記載する最初の報告である。この研究は、SELDIにより分解される複数のタンパク質の組み合わせの選択が、単一マーカー(すなわち、尿細胞学)に比べて、早期の膀胱癌の検出率を劇的に改良することを記載している。全てのスペクトルのクラスタ分析および困難な手動の視診によってこの後者の研究での差分分析を実行したが、より良い癌バイオマーカーの発見を容易にするSELDIプロファイリングの能力を実証した。さらに、それによって、SELDIデータの高次元を効果的に取り扱うためのバイオインフォマティクス・アルゴリズムの必要性を明確に示された。この先行研究の結果にもとづいて、発明者らは、これらの臨床プロテオミクス研究から生じた大量のデータを調査および分析する

10

20

30

40

50

ために複数の異なるバイオインフォマティクス・モデルを考察した。該モデルは、純粋な生物統計学的アルゴリズム、遺伝子クラスター・アルゴリズム、サポート・ベクトル・マシン、および決定分類木を含んでいた。非癌（BPH/正常）試料からPCAを分離する上で、全てが83か90%間の精度で全てを得た（G. Wright, O. J. Semmes, P. Barlett, C. Harris、未発表観察）。

【0122】

分類木モデルは、解釈が容易であり、ニューラル・ネットワークおよび生物統計学的アルゴリズム等の「ブラック・ボックス（black box）」分類子と比べて、特に該分類子で用いられるタンパク質ピークを同定する際必要になるデコンヴォリューション・ステップに関連する問題の点で結果が明確に提示されることができるので、概モデルを選択した。決定木アルゴリズムを用いて、分類子で用いられるタンパク質ピークを規則の検査によって容易に達成でき、これらの規則は、SELDI処理スペクトルの検査によって容易に検証される。卵巣癌を含む他の研究（17）では、卵巣癌分類用の弁別パターンは、534Da、989Da、2111Da、2251Da、および2465Daの5つのタンパク質量からなっていた。それらは、複数の異なるタンパク質に結合しかつその後該タンパク質が本研究で用いられるIMAC-3Cuチップに結合すると考えられる疎水性チップ化学を用いたが、該質量は、前立腺分類システムで用いられたものとは明確に異なると示されたことが興味深い。これによって、SELDIタンパク質「指紋（fingerprint）」プロファイリング手法が癌の各種類に対して異なるタンパク質パターンを検出することが示唆される。当研究所で進行中の研究によって、これがその実例である可能性が示唆される。発明者らは、乳癌、卵巣癌、膀胱癌、および白血病のSELDIプロファイルが互いに異なり、かつこの報告に記載される前立腺分類プロファイルとは異なることを観察した（G. Wright, Jr., A. Vlahou, C. Laronga, J. Marks, O. J. Semmes、未発表観察）。当診断システムの頑健性を確認するために、非前立腺癌および非前立腺疾患を用いて前立腺分類アルゴリズムを調べて、タンパク質プロファイリング分類アルゴリズムが前立腺癌に特異的であることを決定した。類似のスキームが任意の疾患特異的分類システムに求められる。

【0123】

この研究の目標の1つは、癌の早期発見用に潜在的に用いられうる前立腺プロテオーム中のマーカーを同定することであった。5～10年にわたって経時的血清試料を評価する当研究所で進行中の研究によって、PSA試験によってPCAを5年以上早く疑うことが可能であることが示唆された（G. Wright, Jr., P. F. Schellhammer, B. L. Adam、未発表観察）。しかし、早期発見のためのこの分類システムを効果的に適用するためには、侵襲性の強い、すなわち臨床的に重要な癌を侵襲性の強くない癌から識別することができる他のバイオマーカーを同定することが重要である。現在の証拠によって、手術後の障害の存在、量、悪性度分類（grade）、または割合を予測するために10ng/mlを下回る手術前の血清PSAが有用なバイオマーカーではないことが示唆される（4、18）。したがって、PSAおよびその分子形態が提供することができるもの全てより良い生物マーカーに対する切迫した必要性がある。グリソン悪性度分類（Gleason grade）4/5（未分化癌）の量に比例したマーカーは、腫瘍生物学に適応させたより論理的に直接的な療法の重大な必要性を表す。侵襲性の強い癌を示すバイオマーカーまたは危険因子（リスク・ファクタ）を同定しようとして、治療後に再発性疾患に対する生化学的証拠を有する患者由来の手術前および手術後試料のSELDI血清スペクトルを評価する研究が当研究所で進行中である。

【0124】

本明細書に記載される前立腺分類システムの成功的な使用は、9つの質量のタンパク質「指紋（fingerprint）」パターンに完全に依拠する。これらの質量は、再現的に確実に検出されることが見出されているので、診断の正確な分類をするために質量値のみが必要となる。鑑別診断の目的のためにそれらのアイデンティティを知ることは必要ではない。しかし、それらの厳密なアイデンティティを知ることは、これらのペプチド/

10

20

30

40

50

プロテインがPCAの発癌で有する生物学的役割を理解するために不可欠であるので、新規の治療標的へと至る可能性がある。これらのタンパク質/ペプチド・バイオマーカーを精製、同定、および特徴づける研究が進行中である。さらに、それらのアイデンティティを知ることによって、PSAおよび前立腺特異的膜抗原(P SMA)の定量化に関して発明者らが先行して記載した単一および多重形式に類似した古典的なまたはSELDIイムノアッセイの発展用の抗体を生成することが可能になる(12、19)。このSELDIイムノアッセイ形式によって、多重バイオマーカーの定量化用の代替的なプラットフォームが提供される。

【0125】

この実施例で提示した血清タンパク質プロファイリング手法で得られる高感度、特異度、PPV、および陰性予測値(NPV)によって、人工知能分類アルゴリズムと組み合わせたSELDIタンパク質チップ質量分析が、前立腺疾患用のより良いバイオマーカーの発見を容易にすることができ(L. H. Cazares, B. L. Adam, M. D. Ward, S. Nasim, P. F. Schellhammer, O. J. Semmes, and G. L. Wright, Jr. Normal, benign, pre-neoplastic, and malignant prostate cells have distinct protein expression profiles resolved by SELDI mass spectrometry、発表済み)、かつ前立腺癌の早期発見および鑑別診断を改良する潜在能力を有する革新的な臨床診断プラットフォームを提供することができることが実証される。

【0126】

(実施例2)

(前立腺癌または良性前立腺肥大を診断するための血清タンパク質指紋法および促進決定木分析の使用)

(材料および方法)

(研究群および試料)

血清試料は、バージニア・プロステート・センター・ティッシュ(Virginia Prostate Center Tissue)およびボディ・フルード・バンク(Body Fluid Bank)から得た。全ての試料は、東バージニア医科大学倫理調査委員会(Eastern Virginia Medical School Institutional Review Board)によって認証されるプロトコールにしたがって同意した患者から入手し、-80で凍結保存された。その後2回より多く解凍された試料はなかった。治療前試料は、臓器限局癌と診断されたPCA患者(平均年齢71歳)99人、非臓器限局疾患と診断されたPCA患者(平均年齢69歳)98人、BPHと診断された患者(平均年齢67歳)92人から、検体は、陰性直腸指診(DRE)、すなわち4.0 ng/ml未満のPSAであり、かつ前立腺疾患の証拠がない健常者(平均年齢67歳)96人から得た。平均PSA値は、健常者が1.32 ng/ml、BPHが4.60 ng/ml、臓器限局PCAが10.10 ng/ml、非臓器限局が206.93 ng/mlであった。各正常ドナーから等量の血清をプールして、100ulアリコートで-80で保存することによって精度管理試料を調製した。

【0127】

(SELDIタンパク質プロファイリング)

1.5 ml微量遠心管内で、PBS中で8 M尿素/1% CHAPS 30 μlとともに、血清20 μlを4で10分間ボルテックスすることによって血清試料を調製した。その後、0.125% CHAPSとあわせて100 μlの1 M尿素を添加して、混合物を短時間ボルテックスした。PBSでの1:5希釈のうち50 μlを、CuSO₄で予め活性化させたIMAC-3チップを含有するバイオプロセッサ(Ciphergen Biosystems, Inc, Fremont, CA)の各ウエルに適用し、該バイオプロセッサを密閉して、プラットフォーム・シェーカー上で、速度250 rpmで30分間

攪拌した。同様に調製したブールしたQC血清試料を各チップ上のアレイに適用し、この研究で再現性照合として用いた。血清/尿素混合物を破棄して、PBSを用いてチップを3回洗浄し、バイオプロセッサからチップを取り出し、DI水で洗浄($\times 10$)し、風乾し、SELDI分析にかけるまで暗所で保存した。SELDI分析の直前に、50%(v/v)アセトニトリル、0.5%トリフルオロ酢酸中のシナピン酸の飽和溶液0.5 μ lを各チップ・アレイ上に2回使用して、シナピン酸の各使用の間、アレイ表面を風乾させた。PBS-I I質量分析計(Ciphergen Biosystems, Inc.)にチップを入れ、レーザー強度220、192のレーザー照射(ショット)(ポジティブ・モード、レーザー強度220、検出器感度7、および900nsの集束遅れ時間)を平均化することによって飛行時間スペクトルを生じさせた。オール・イン・ワン・ペプチドNWスタンダード(All-in-1 peptide MW standard)(Ciphergen)を用いて、質量精度を外部から校正した。わずかに改良したサイファージェン・プロテインチップ・ソフトウェア3.0(Ciphergen ProteinChip Software 3.0)を用いて、ピーク検出およびアラインメントを実行した。2000から40000Daの質量範囲は、分解されたタンパク質/ペプチドの大多数を含有しているため、この範囲を分析用に選択した。

10

【0128】

(データ分析)

(特徴選択)

AUCを評価することによって、正常からPCAを、正常からBPHを、およびPCAからBPHを区別する際の各ピークの能力を決定した。ROC曲線下面積は、0.5(区別なし)から1.0(絶対予測)の範囲である(6)。0.62以下のAUCを有するピークをその後のデータ分析から除外した。

20

【0129】

(促進決定スタンプ分類子)

167のPCA、77のBPH、および82の正常血清試料からなる訓練データ・セットを用いて分類子を開発した。その後、分類アルゴリズムの有効性および精度を、30のPCA、15のBPH、および15の正常試料からなる盲検データ・セットを用いて調べた。

【0130】

Freund and Schapire(7-9)によって記載されるアダブースト(AdaBoost)アルゴリズムを改変して用いて、分類子を構築した。このアルゴリズムでは、決定「スタンプ(stump)」を基準分類子として用い、該基準分類子のそれぞれが1つのピークを用いて1つの分割を有する。決定スタンプは通常、むしろ高誤り率を有する弱い分類子である。しかし、重みつき投票を用いて組み合わせたスタンプは、非常に精度のよい分類子であることが予想される。決定スタンプは、 (Z, c) によって表され、この際Zは、訓練セット中のピークから選択されたピークであり、cは閾値である。ith訓練試料は、 $(y_i; Z_{i1}, \dots, Z_{ip})$ によって表され、この際、 y_i は、観察i($i = 1; \dots, N$)のクラス(ラベル)を表し、 Z_{ij} は、ith観察のjthピークの強度である。2つのクラスがある場合は、クラス1(例えば非癌)は $y_i = 1$ によって表され、クラス2(例えば癌)は $y_i = -1$ によって表される。スタンプは、2つの葉を有し、左葉は、閾値c以下のピークZの強度を有する訓練試料を含み、右葉は他の試料の全てを含む。左葉中の大部分の試料が、例えば癌である場合、 $Z < c$ を有する試料は癌として分類される。分類子は、 $f(x)$ として表され、この際xは、ブール変数「 $Z < c$ 」であり、ith訓練試料がクラス1として分類される場合、 $f(x)$ は $\{-1, 1\}$: $f(x_i) = 1$ から値をとり、ith試料がクラス2として分類される場合は $f(x_i) = -1$ から値をとる。 $y_i f(x_i) = -1$ である場合、試料は誤分類される。左葉 $Z < c$ 、または $x = \text{「真」}$ に対しては、 n_{11} および n_{21} を、それぞれ $y_i = 1$ および $y_i = -1$ を有する観察の数とし、すなわち、

30

40

【0131】

50

【数 1】

$$n_{11} = \sum_{i=1}^N I\{(y_i=1) \& (Z \leq c)\}, \quad n_{21} = \sum_{i=1}^N I\{(y_i=-1) \& (Z \leq c)\}, \quad (1)$$

【0 1 3 2】

であり、式中、 $I\{\text{記載}\}$ は、特性関数であり、これは記載が真である場合は 1 に等しく、または記載が偽である場合は 0 に等しい。同様に、 n_{12} および n_{22} を、それぞれ $y_i = 1$ および $y_i = -1$ の観察の数とし、かつ $Z > c$ とし、すなわち、

【0 1 3 3】

【数 2】

$$n_{12} = \sum_{i=1}^N I\{(y_i=1) \& (Z > c)\}, \quad n_{22} = \sum_{i=1}^N I\{(y_i=-1) \& (Z > c)\}. \quad (2)$$

【0 1 3 4】

である。
その後、多項モデルに対するこの対数尤度は、

【0 1 3 5】

【数 3】

$$\log L = \sum_{u=1}^2 \sum_{v=1}^2 n_{uv} \log(p_{uv}), \quad (3)$$

【0 1 3 6】

であり、式中、 P_{uv} は、 $n_{uv} = (n_{1v} + n_{2v})$ によって評価される。ピーク Z およびその閾値 c は、対数尤度を最大化することによって得られる。前立腺癌バイオマーカー（括弧内はタンパク質バイオマーカーの分子量）に対するこの実施例では、以下の閾値を利用した。すなわち、0.1912 (9656)、1.0519 (9720)、0.0000 (6542)、0.0000 (6797)、2.2427 (6949)、0.0000 (7024)、0.1638 (8067)、1.7755 (8356)、13.8103 (3963)、0.8301 (4079)、0.3805 (7885)、および 0.0000 (6990) である。良性前立腺肥大バイオマーカー（括弧内はタンパク質バイオマーカーの分子量）に対するこの実施例では、以下の閾値を利用した。すなわち、0.0000 (7820、4580、7844、4071、6099 に対して)、0.2679 (7054)、0.1991 (5298)、3.3758 (3486)、および 20.1535 (8943) である。これらの閾値を介して、連続活性を二分（論理）値に変換する。

【0 1 3 7】

ブースティング分類子を開発する際、重みつき観察上の決定スタンプのアセンブルを構築する。例えば重み w_i を i th 観察 y_i に割り当てる場合、重みの合計は N である。その後、計数 n_{11} 、 n_{21} 、 n_{12} 、および n_{22} を変更して重みを包含すべき場合以外では、方程式 (3) をさらに用いて分割 (Z, c) を見出す。方程式 (1) は、

【0 1 3 8】

【数 4】

$$n_{11} = \sum_{i=1}^N w_i I\{(y_i=1) \& (Z \leq c)\}, \quad n_{21} = \sum_{i=1}^N w_i I\{(y_i=-1) \& (Z \leq c)\},$$

【0 1 3 9】

であり、方程式 (2) は、

【0 1 4 0】

10

20

30

40

50

【数 5】

$$n_{11} = \sum_{i=1}^N w_i I\{(y_i = 1) \& (Z > c)\}, n_{22} = \sum_{i=1}^N w_i I\{(y_i = -1) \& (Z > c)\}.$$

【0 1 4 1】

である。したがって、ブースティング・アルゴリズムの基本的概念は、重みつき観察上に基準分類子のアセンブルを構築することである。第 1 のラウンドでは、全ての観察に対して等しい重みを用い、すなわち、 $i = 1, \dots, N$ に対して $w_i = 1$ である。試料を分類して、次のラウンドでは、先行ラウンドで誤分類された観察に対して重みを増加させる一方で、正確に分類された観察に対しては重みを減少させる。したがって、次の決定スタンプは、先行スタンプにより誤分類された 4 つの試料に焦点をあてる。下記に概説するように、所定数のスタンプが構築されるまで、この手順を何度も繰り返す。

10

【0 1 4 2】

1. 第 1 のラウンドでは、等しい重み $w_i = 1$ ($i = 1, \dots, N$) を用いる。

【0 1 4 3】

2. M への $m = 1$ に関しては、重み w_i を有する訓練データに対して決定スタンプ $f_m(x)$ を構築し、誤り率を計算する。

【0 1 4 4】

【数 6】

$$err_m = \frac{1}{N} \sum_{i=1}^N w_i I(y_{if_m}(x_i) = -1)$$

20

【0 1 4 5】

3. 信頼度を計算する：

【0 1 4 6】

【数 7】

$$\alpha_m = \log\{(1 - err_m) / err_m\}.$$

30

【0 1 4 7】

4. 重みを更新する：

【0 1 4 8】

【数 8】

$$w_i \leftarrow w_i \cdot \exp\{\alpha_m I(y_{if_m}(x_i) = -1)\}, i = 1, \dots, N,$$

【0 1 4 9】

に設定し、重みを正規化する。すなわち、

【0 1 5 0】

【数 9】

$$\sum_i w_i = N.$$

40

【0 1 5 1】

5. 採集組み合わせ分類子として、線形結合を用いる：

【0 1 5 2】

【数 1 0】

$$f(x) = \sum_{m=1}^M \alpha_m f_m(x)$$

【0 1 5 3】

組み合わせ分類子 $f(x)$ は、 M 基準分類子の重みつき多数決票である。 i th 試料に対しては、クラス 1 として分類される場合は m th 基準分類子 $f_m(x_i) = 1$ であり、クラス 2 として分類される場合は $f_m(x_i) = -1$ である。最終投票への m th 決定スタンプの組み合わせは、クラス 1 に対する投票の場合は m であり、またはクラス 2 に対する投票の場合は $-m$ である。従って、総投票が正、すなわち、

【0 1 5 4】

【数 1 1】

$$f(x_i) = \sum_{m=1}^M \alpha_m f_m(x_i) > 0, i = 1, \dots, N, \quad (4)$$

【0 1 5 5】

であり、クラス 1 に対する多数決票である場合、試料はクラス 1 として分類され、 $(f(x_i) > 0)$ である場合はクラス 2 として分類される。これは、 α_m が等しくないので、重みつき多数決票である。

【0 1 5 6】

分類子に含まれる変数を非常に多くあるので、過度のあてはめ (over-fitting) の証拠なしに試験または一般化誤りはゼロである。なぜブースティング法が過度のあてはめ (over-fitting) を避けることができるかについての詳細な説明は、Schapire and Freund によって記載されている (10)。 $y_i f(x_i) > 0$ である場合またはある場合のみ、訓練セット中の試料は正確に分類された。後者は 2 つの部分からなる。すなわち、

【0 1 5 7】

【数 1 2】

$$y_i f(x_i) = \sum_{m=1}^M \alpha_m y_i f_m(x_i) = \sum_{m=1}^M \alpha_m I(y_i f_m(x_i) > 0) - \sum_{m=1}^M \alpha_m I(y_i f_m(x_i) \leq 0), i = 1, \dots, N,$$

【0 1 5 8】

であり、式中、重み α_m は正規化されている ($\sum_{m=1}^M \alpha_m = 1$)。方程式の右側にある第 1 の項目は正確な決定を下した多数決器 (voter) (M) の総投票であり ($y_i f_m(x_i) > 0$ なので)、第 2 の項目は、誤った決定を下した多数決器 (voter) の総投票である。 $y_i f(x_i)$ は、 i th 試料の誤差として言及される。組み合わせ分類子によって、負の誤差を有する試料は誤分類されている。負の誤差を有する訓練セット中の比率は、訓練誤り率である。誤差の最小値は、

【0 1 5 9】

【数 1 3】

$$\text{最小誤差}_M = \min_{1 \leq i \leq N} \left\{ \sum_{m=1}^M \alpha_m y_i f_m(x_i) \right\}$$

【0 1 6 0】

によって決定され、投票法にとって重要な量である。最小誤差がゼロを超える場合、誤分類される訓練試料がない。最小誤差が大きくなるほど、試料が誤分類される確率が少なく

なる（すなわち、一般化でのより高い信頼度）。したがって、訓練試料での最小誤差は、訓練および試験セット両方での学習アルゴリズムにより2つのクラスがどの程度良く分類されるかを測定する。最小誤差が増加し続けるにつれて、試験試料が組み合わせ分類子により正確に分類される余地がますます大きくなる。最小誤差のプロットは、いつ基準分類子をさらに追加することを停止するかを決定する上で重要である（11）。

【0161】

（統計学的分析）

受信者動作特性（ROC）のAUCを計算して、3つの群を区別する最高潜在能力を有するピークを同定した（6）。正確に分類された非疾患試料数と非疾患試料の総数との比率として特異度を算出した。正確に分類された疾患試料数と疾患試料の総数との比率として感度を算出した。

10

【0162】

（結果）

各SELDIスペクトルによって、2000～40000Da範囲中の平均80のピーク質量が明らかになった。ピーク強度に対して、0.05%のピーク位置と、それぞれ15%および20%のCVとに関するイントラアッセイおよびインターアッセイCVを用いて、QCスペクトルは非常に再現的であると判明した（データ図示せず）。図3は、SELDIスペクトルの代表例を示す。772にのスペクトル（336の試料で二重に実行）全ての分析は779のピークを同定し、そのうち124は、0.62以上のAUCを有していた。訓練セットで同定されたこれら124のピークを用いて、分類子を構築した。

20

【0163】

学習アルゴリズムの構築および使用の際での懸念の1つは、データの過度のあてはめ（over-fitting）の可能性である。しかし、ブースティング法は、最小誤差を算出することによって過度のあてはめ（over-fitting）を避けることができる。最小誤差が大きくなるほど、試験試料が誤分類される確率が少なくなる。図4Aは、M（癌から非がんを識別する促進決定木分類子に対する基準スタンプの数）に対する訓練誤り率、最小誤差、および一般化誤り率（試験誤り）を示す。訓練誤りがゼロに到達した（ラウンド47）後、最小誤差が増加し続けると同時に、一般化誤りが減少し続け、最終的にはラウンド265でゼロに到達し、その後ゼロで保たれる。同様に、図4Bは、BPHから正常を識別する促進決定木分類子のための基準スタンプの数に対する訓練誤り率および最小誤差を示す。ここでも、訓練誤りがゼロに到達した（ラウンド9で）後、最小誤差は増加し続ける。図示するように、ブースティング技術に対する過度のあてはめ（over-fitting）の証拠はない。訓練誤りがゼロになるとき、学習プロセスは停止しないが、それとは反対に、学習アルゴリズムは、2つのクラス間の最小誤差を拡大させ続ける。従って、最小誤差が増加し続ける限り、さらに基準分類子を付加することは、過度のあてはめ（over-fitting）を引き起こさない。

30

【0164】

PCAから非癌を識別するための第1のブースティング分類子（アダブースト分類子（AdaBoost Classifier））は、400の基準分類子からなり、該基準分類子には、326の訓練試料および60の試験試料両方で0誤り率を有する62のピークが含まれる。基準スタンプの数（すなわち、ラウンド数）が47より大きかったとき、訓練誤りはゼロであったが、試験誤り（一般化誤り）は0.0333であった。一般化誤りは、基準スタンプ数が増加するにつれて徐々に減少することが判明した。ラウンド265後に、一般化誤りは依然としてゼロであった。BPHから正常を識別するための12のピークを有する100の決定スタンプも、159の訓練試料および30の試験試料の両方に対して0誤り率を得た。この場合、訓練誤り率は、ラウンド9でゼロになり、30の試験試料に対する一般化誤りは、ラウンド1で始まって0であった。これら2つの促進決定スタンプを組み合わせると、3つのクラス、すなわち正常、BPH、およびPCAに対して100%分離を達成した（表1）。

40

【0165】

50

しかし、この分類子は、74のピークを用いた500の基準分類子を組み合わせた。解釈するために、どのピークが非癌から癌を識別する上で最も重要であるか、およびどのピークが正常からBPHを識別する上で最も重要であるかを知る必要がある。これは、マシン学習での特徴選択の問題である。決定木モデルでの本質的な構成要素である特徴選択のための多数の方法がある。実際、ROC曲線下面積を用いて、779のピークから124を選択することは、特徴選択の第1ステップであり、フィルタ法と呼ばれる。促進決定スタンプ分類子を構築する際、各ラウンドで、124のピークから1つのピークを選択する。この特徴選択手順はアルゴリズムに組み込まれているので、特徴(ピーク)を何度も選択することが可能である。同じピークを複数回選択することを避けるため、かつ各ラウンドで新規のピークのみを選択するために、先行して選択した特徴は無視する。このアルゴリズムは、促進決定スタンプ特徴選択(Boosted Decision Stump Feature Selection(BDSFS))として言及される(12)。この方法を用いて、BPHおよび正常者からPCAを識別する際に有用性を持つタンパク質質量のアイデンティティを容易に得ることができる。表3には、BDSFSアルゴリズムにより非癌から癌を識別するために選択された第1の12のピークと、正常からBPHを識別するために選択された第1の9のピークとに対する平均分子量を列記する。表4は、BDSFSアルゴリズムの結果を提示する。これらの表を以下に示す。

10

【0166】

(表3 ピーク質量選択のために促進決定スタンプ特徴選択分類子に用いられたピーク質量)

20

【0167】

【表3】

非癌 対 癌		正常 対 BPH	
No.	質量	No.	質量
1	9655.75	1	7819.75
2	9719.99	2	4579.73
3	6541.82	3	7844.00
4	6797.02	4	4071.18
5	6949.22	5	7054.17
6	7024.02	6	5297.55
7	8066.95	7	3486.21
8	8355.56	8	6099.08
9	3963.18	9	8943.08
10	4079.48		
11	7884.72		
12	6990.63		

30

【0168】

(表4 促進決定スタンプ特徴選択アルゴリズムによる訓練および試験セットの分類)

40

【0169】

【表 4】

予測		訓練セット			試験セット		
		正常	BPH	PCA	正常	BPH	PCA
BDSFS 分類	正常	82	0	7	14	0	0
	BPH	0	74	0	0	15	1
	PCA	0	3	160	1	0	29
	感受性	95.18%			96.67%		
	特異性	98.11%			96.67%		
	ピーク番号	1の基準分類子から21のピーク					
	最小誤差	-0.2555					

10

【0170】

この分類子（BDSFS分類子）は、BDSFSアルゴリズムによって選択された21のピークを用いており、該ピークは、非癌から癌を識別するための表3中の12のピークと、BPHから正常を識別するための第1の9のピークとからなっていた。この分類子は、96.67%の試験セットで感度および特異度を得た。この場合、解釈は、74のピークを含む（表1）アダブースト（AdaBoost）分類子よりもかなり容易である。しかし、BDSFS分類子に対する最小誤差は-0.2555である一方で、分類子1に対する最小誤差は、0.1143である。したがって、アダブースト（AdaBoost）分類子は、新規（未知）試料に対して、BDSFS分類子よりも精度がよい。

20

【0171】

(考察)

ヒスチジン、トリプトファン、システイン、またはリン酸化アミノ酸を介して、化学的活性化銅表面に選択的に結合するタンパク質能力にもとづいてタンパク質を捕獲するタンパク質チップを用いたSELDI質量分析法は、2,000~40,000ダルトン範囲の平均80の血清タンパク質/ペプチドを分解することができる。これは、二次元電気泳動によって選択されることができるタンパク質の100から数千未満であるが、2D-EFを越える利点は、20,000Daを下回るポリペプチドおよびペプチドを効果的に分解するSELDIの能力である。これによって、診断、予後用の潜在的バイオマーカーとして、かつ治療標的としてのペプチドを容易に分解および研究することに対して門戸が開かれた。興味深いことに、非癌群からPCAを分離するために最も有用であるところの研究で同定された24のタンパク質は、3,000から10,000Daの範囲であった。SELDIの導入以来、そのような「ペプチド(peptide)」は、より大きなタンパク質のランダムなフラグメントを表す可能性があるという関心があった。これが真実であった場合、発明者らは、タンパク質パターンでの高再現性を達成することができている（ピーク位置に対して0.05%CVおよびピーク強度に対して15~20%CV）。

30

40

【0172】

2から40kDa間で得られるスペクトルごとに平均80のピークのみを用いてでさえも、依然としてデータは高次元である。772の血清試料（386を二重に実行）の初期分析によって、63,157のピークが生じ、このピークは、クラスタ分析およびピーク・アラインメント後に779のピークに減少した。779のピークのうち、124のピークが、3つの群（正常対PCA、正常対BPH、およびBPH対PCA）を区別する最も高い潜在能力を有することが統計学的に見出された。772の試料のそれぞれでの124のピークのその後の分析によって、非癌群からPCAを分離する質量のパターンまたは組

50

み合わせを同定するための指示に応じて、95,000のデータ・ポイントにわたる処理が生じた。この高次元のために、人工知能アルゴリズムのみがこのような高データ量を分析して、効率的かつ再現的な分類子を開発することができる。発明者らは、生物統計学的アルゴリズム(13、14)、遺伝子クラスター・アルゴリズム、およびサポート・ベクトル・マシン・アルゴリズムを含む複数の異なるモデルを評価した。非癌群(BPH/非癌)からPCAを区別する上で、大部分が83~90%間の精度を得ることができたが、決定木モデル(15)は、ニューラル・ネットワークおよび生物統計学的アルゴリズム等の「ブラック・ボックス(black box)」分類子と比べて容易に解釈されるので、該決定木モデルを選択した。この研究に記載される同データ・セットを用いて、発明者らは、非癌群(すなわち、BPHおよび正常)からPCAを区別するために、83%の感度および97%の特異度を達成する2000から10000Da間の9つの質量を有する単一の決定木基準分類子を開発した(実施例1に論じている)。しかし、単一の決定木分類子の予測能力は、ニューラル・ネットワークおよび表面ベクトル・マシン等の他の学習アルゴリズムほど優れてはいない可能性がある。さらに、癌の早期発見のためのアッセイは、あまりに多くの偽陽性を生成することを避けるために高精度である必要がある。本研究を開始して、決定木分類子の予測能力を増加させることができるかを決定した。ブースティング(16)およびブートストラップ法等の投票法を用いた決定木分類子の予測能力での極めて大きい改善が近年報告されている。バギング法(17、18)と呼ばれる1つの投票法では、無作為に再抽出した観察(ブートストラップ準試料)に決定木モデルを何度も適合させ、その後、単一の投票を用いて決定木を組み合わせる。別の手法は、アダブースト(AdaBoost)アルゴリズムとして言及されるブースティング法(7)であり、これは、重みつき観察に決定木モデルを何度も適合させ、その後重みつき投票を用いて決定木を組み合わせる。バギングおよびブースティング両方では、組み合わせ分類子は、個々の基準決定木それぞれよりも優れた成果を有する。ブースティング手法は通常、試験試料で、バギング手法よりも精度がよい(10)ため、発明者らは、バギング・アルゴリズムではなくブースティング手法を選択した。アダブースト(AdaBoost)アルゴリズムを用いて、訓練セットおよび盲検セット療法に対して、試料がPCAまたはBPHと診断された患者由来のものであるか、または試料が健常ドナー由来のものであるかを予測する上で100%の精度である分類子を確立した。この分類子は、100%の感度および特異度を生成したが、該分類子は、74のタンパク質質量値(ピーク)を用い、かつ5000の基準決定木分類子を組み合わせることを必要とするので、高精度であるが解釈することが難しくなっていた。ウェーブレット(Wavelets)(13)および表面ベクトル・マシン(Wright、未発表データ)を用いた生物系等学的手法等の他のモデルは、類似の高精度に到達することができるが、特に分類子で用いられるタンパク質質量を同定する上で同様の困難さがある。同じ特徴(ピーク)を何度も選択するときこの困難さが生じる。同じピークを複数回選択することを避けるために、発明者らは、促進決定木スタンプ特徴選択(Boosted Decision Stump Feature Selection)(12)と呼ばれる改良型ブースティング・アルゴリズムを用いており、これは、各ラウンドで単一のピークのみを選択し、先行ラウンドから選択されたピークは除外する。この方法では、3つの群を識別する上で重要な24のピークのアイデンティティが容易に得られた。分類子は、15のBPHのうち1つをPCAとして、29のPCAのうち1つを正常として誤分類することによってアダブースト(AdaBoost)分類子よりもわずかに精度が落ちたが、一方で、正常非罹患男性から得た14の試料の全てを正確に同定した。この分類子はさらに、21のピークおよび21のみの基準決定木分類子を用いて、感度および特異度両方に対して、相当な96.67%を達成した。特異度は依然として、単一の基準分類子を用いて得られた97%と同じであったが、感度、すなわちPCA試料を正確に予測する能力は、このブースティング・アルゴリズムによって、83%から97%に増加した。

【0173】

PSA試験は、前立腺癌に対する現行のスクリーニング試験であり、陽性の場合、前

10

20

30

40

50

立腺の各葉から生検を得る。大勢が、この試験がいずれのヒト癌にも最善であるとみなしているが、PCAの早期発見用の完璧な試験とはかけ離れている。該試験は、>90%の高感度を有しているが、BPHからPCAを識別する上でその特異度は25%のみであり、さらに前立腺癌を有する者にはPSAの正常レベルを有する者もいる。その低特異度のために、人は、不必要な生検を受けなければならない、実際に癌ではないとき相当な懸念を引き起こすことになる。現在の証拠によって、手術後の障害の存在、量、悪性度分類(grade)、または割合を予測するために10ng/mlを下回る手術前の血清PSAが有用なバイオマーカーではないことも示唆される(1)。これらの事実にもとづいて、PSAおよびその分子形態が提供できる全てより良い生物マーカーに対する必要性がある。非癌患者から得た試料を用いてプロファイリング・アッセイの有効性を試験することを含む複数の部位で評価されるより大量の試料上でブースティング決定木分類子の精度が検証されうる場合、バイオインフォマティクス分類子と組み合わせたSELDIタンパク質プロファイリングは、前立腺癌の早期発見および診断用の「より良い(better)」試験を提供することが可能である。この潜在能力を裏付けるものは、その後この研究で用いた異なるバイオインフォマティクス分類子と組み合わせたSELDIを用いて、卵巣癌および乳癌に対して類似の結果を達成している他の研究者の報告である(19; Li, J, Zhang, Z, Rosenzweig, J, Wang, W, Chan, D. Potential serum biomarkers identified by SELDI mass spectrometry can discriminate breast cancer from non-cancer patients. Proc. Am. Assoc. Cancer Res. 2002; 43:136 [abstract 682])。全体的なこれらの初期研究によって、SELDIにより癌診断に対する革新的なプロテオミック手法を開発するための特有の機会が提供されることが示唆される。

【0174】

分類子で用いられるピーク質量のアイデンティティは、診断を下すためには必要ではない。精度の良い診断を下すためのこの分類システムに対する唯一の要件は、バイオマーカーがSELDIによって再現的に検出されること、および分類子によって選択されることである。分類子で用いられる質量のそれぞれに対する名称を得ることは、分類システムをより良くまたはより高精度にはしない。しかし、タンパク質バイオマーカーのアイデンティティを知ることは、発見の観点から重要である。ペプチド/タンパク質バイオマーカーのアイデンティティは、これらのタンパク質がPCAの発癌で有する生物学的役割を理解するために必要となる。このような情報は、より良い治療上の介入へと至ることが可能である。プロファイリング手法が臨床アッセイへと発展されなかった場合に、アイデンティティを知ることは、古典的な多重イムノアッセイおよび抗体アレイの発展用の抗原および抗体試薬の生成を容易にする。これらの理由から、潜在的バイオマーカーのタンパク質同定が進行中である。

【0175】

結論として、多重血清バイオマーカーの組み合わせ使用によって達成される高感度および特異度によって、学習アルゴリズムと組み合わせたSELDIが、PCA用の新規かつより良いバイオマーカーの発見を容易にするばかりではなく、新規の臨床診断アッセイへと発展するための潜在能力を有する証拠の裏づけが提供される。

【0176】

(実施例1で引用した参考文献)

1. Howe, H. L., Wingo, P. A., Thun, M. J., Ries, L. A., Rosenberg, H. M., Feigal, E. G., and Edwards, B. K. Annual report to the nation on the status of cancer (1973 through 1998), featuring cancers with recent increasing trends. J. Natl. Can 50

- cer Inst., 93: 824-842, 2001.
2. Djava (登録商標)n, B., Zlotta, A., Kratzik, C., Remzi, M., Seitz, C., Schulman, C. C., and Marberger, M. PSA, PSA density, PSA density of transition zone, free/total PSA ratio, and PSA velocity for early detection of prostate cancer in men with serum PSA 2.5 to 4.0 ng/mL. *Urology*, 54: 517-522, 1999.
3. Pannek, J. and Partin, A. W. The role of PSA and percent free PSA for staging and prognosis prediction in clinically localized prostate cancer. *Semin. Urol. Oncol.*, 16: 100-105, 1998. 10
4. Stamey, T.A., Johnstone, I.M., McNeal, J.E., Lu, A.Y., and Yemoto, C.M. Preoperative serum prostate specific antigen levels between 2 and 22 ng/ml correlate poorly with post-radical prostatectomy cancer morphology: Prostate specific antigen cure rates appear constant between 2 and 9 ng/ml. *J. Urol.* 167:103-111, 2002. 20
5. Srinivas, P.R., Srivastava, S., Hanash, S., and Wright, Jr., G.L. Proteomics in early detection of cancer. *Clin. Chem.* 47:1901-1911, 2001.
6. Adam, B-L., Vlahou, A., Semmes, O.J., and Wright, Jr., G.L. Proteomic approaches to biomarker discovery in prostate and bladder cancers. *Proteomics* 1: 1264-1270, 2001. 30
7. Chong, B. E., Hamler, R. L., Lubman, D. M., Ethier, S. P., Rosenspire, A. J., and Miller, F. R. Differential screening and mass mapping of proteins from premalignant and cancer cell lines using nonporous reversed-phase HPLC coupled with mass spectrometric analysis. *Anal. Chem.*, 73: 1219-1227, 2001. 40
8. Ferrari, L., Seraglia, R., Rossi, C. R., Bertazzo, A., Lise, M., Allegri, G., and Traldi, P. Protein profiles in sera of patients with malignant cutaneous melanoma. *Rapid Commun. Mass Spectrom.*, 14: 1149-1154, 2000.
9. Keough, T., Lacey, M. P., Fieno, A. M., Grant, R. A., Sun, Y., Bauer, M. D., and Begley, K. B. Tandem mass spectrometry methods for definitive protein identi 50

- fication in proteomics research. Electrophoresis, 21: 2252-2265, 2000.
10. Merchant, M. and Weinberger, S. R. Recent advancements in surface-enhanced laser desorption/ionization-time of flight-mass spectrometry. Electrophoresis, 21: 1164-1177, 2000.
11. . Kuwata, H., Yip, T. T., Yip, C. L., Tomita, M., and Hutchens, T. W. Bactericidal domain of lactoferrin: detection, 10
quantitation, and characterization of lactoferricin in serum by SELDI affinity mass spectrometry. Biochem. Biophys. Res. Commun., 245: 764-773, 1998.
12. . Wright, G. L., Jr., Cazares, LH, Leung, Sau-Mei, Nasim, S, Adam, BL, Yip, TT, Schellhammer, PF, and Vlahou, A. Proteinchip Surface Enhanced Laser Desorption/Ionization (SELDI) Mass Spectrometry: 20
A Novel Proteomic Technology for Detection of Prostate Cancer Biomarkers in Complex Protein Mixtures. Prostate Cancer and Prostate Diseases 2, 264-267. 1-1-1999.
13. . Vlahou, A., Schellhammer, P. F., Mendrinos, S., Patel, K., Kondylis, F. I., Gong, L., Nasim, S., and Wright, J. G., Jr. Development of a novel proteomic approach for the detection of transitional cell carcinoma of the bladder in urine 30
. Am. J. Pathol., 158: 1491-1502, 2001.
14. . Breiman, L., Friedman, J.H., Olshen, R.A., and Stone, C.J. Classification and Regression Trees, Wadsworth International Group, Belmont, CA (1984), pp. 103-111 and 203-215.
15. . Pepe, M.S., Receiver operating characteristic methodology, Amer. Statistical Assoc., 95:308-311, 2000.
16. . Gelman, A., Carlin, J.B., Stern, H 40
.S., and Rubin, D.B. Bayesian Data Analysis. Chapman & Hall, London (1995).
17. . Stamey, T.A. Preoperative serum prostate specific antigen (PSA) below 10 ug/L predicts neither the presence of prostate cancer nor the rate of postoperative PSA failure. Clin. Chem. 47:631-634, 2001.
18. Petricoin, III, E. F., Ardkani, A.M., Hitt, B.A., Levine, P.J., Fusaro, V.A. 50

, Steinberg, S.M., Mills, G.B., Simone, C., Fishman, D.A., Kohn, E.C., Liotta, L.A. Use of proteomic patterns in serum to identify ovarian cancer. *The Lancet* 359:572-577, 2002.

19. Xiao, Z., Adam, B-L., Cazares, L.H., Clements, M.A., Davis, J.W., Schellhammer, P.F., Dalmaso, E.A., Wright, Jr., G.L. Quantitation of serum prostate specific antigen by a novel protein biochip immunoassay discriminates benign from malignant prostate disease. *Cancer Res.* 61:6029-6033, 2001. 10

(実施例2で引用した参考文献)

1. Stamey, TA, Johnstone, IM, McNeal, JE, Lu, AY, Yemoto, CM. Preoperative serum prostate specific antigen levels between 2 and 22 ng/ml correlate poorly with post-radical prostatectomy cancer morphology: Prostate specific antigen cure rates appear constant between 2 and 9 ng/ml. *J Urol* 2002; 167:103-111. 20

2. Merchant, M, Weinberger, SR. Recent advancements in surface-enhanced laser desorption/ionization-time of flight-mass spectrometry. *Electrophoresis* 2000; 21:1164-1177.

3. Kuwata, H, Yip, TT, Yip, CL, Tomita, M, Hutchens, TW. Bactericidal domain of lactoferrin: detection, quantitation and characterization of lactoferrin in serum by SELDI affinity mass spectrometry. *Biochem Biophys Res Commun* 1998; 245:764-773. 30

4. Wright, Jr, GL, Cazares, LH, Leung, S-M, Nasim, S, Adam, BL, Yip, TT, et al. Proteinchip surface enhanced laser desorption/ionization (SELDI mass spectrometry: A novel proteomic technology for detection of prostate cancer biomarkers in complex protein mixtures. *Prostate Cancer & Prostate Diseases* 1999; 2:264-267. 40

5. Vlahou, A, Schellhammer, PF, Mendrinos, S, Patel, K, Kondylis, Fl, Gong, L, et al. Development of a novel proteomic approach for the detection of transitional cell carcinoma of the bladder in urine. *Am J Pathol* 2001; 158:1491-1502.

6. Pepe, MS. Receiver operating characteristic methodology. *Amer Statistical Ass* 50

- oc. 2000; 95: 308-311.
7. Freund, Y, Schapire, R.. A decision-theoretical generalization of on-line learning and an application to boosting. J Computer & Syst Sci. 1997; 55:119-139.
8. Hastie, T, Tibshirani, R, Friedman, J. The Elements of Statistical Learning. New York: Springer-Verlag. 2001; 301 pp.
9. Friedman, J, Hastie, T, Tibshirani, R. Additive logistic regression: A statistical view of boosting. The Annals of Statistics. 2000; 28: 337-407. 10
10. Schapire, R, Freund, Y. Boosting the margin: A new explanation for the effectiveness of voting methods. The Annals of Statistics/ 1998; 26: 1651-1686.
11. Vapnik, VN. The Nature of Statistical Learning Theory. New York: Springer-Verlag. 20
12. Das, S. Filters, wrappers and a boosting-based hybrid for feature selection. In Proceeding of the Eighteenth International Conferences on Machine Learning. San Francisco, CA: Morgan Kaufmann Publishers, 2001; 74-81 pp.
13. Qu, Y, Adam, B-L, Thornquist, M, Potter, J, Yasui, Y, Davis, J, et al. Data reduction using discrete wavelet transform in discriminant analysis with very high dimension. 2002; Biometrics (In Press) 30
14. Yasui, Y, Pepe, M, Thompson, ML, Adam, BL, Wright, G, Jr, Qu, Y, et al. A data-analytic strategy for protein-biomarker discovery: Profiling of high-dimensional proteomic data for cancer detection. 2001; University of Washington, Department of Biostatistics, Technical Report No177. 40
15. Breiman, L, Friedman, JH, Olsen, RA, Stone, C.J. Classification and Regression Trees. Belmont, California: Wadsworth International Group, 1984.
16. Drucker, H, Cortes, C. Boosting decision trees. In Advances in Neural Information Processing Systems 1996; 8: pp 479-485.
17. Breiman, L. Bagging predictors. Machine Learning Breiman, L. 1996. 24:123-14 50

0 .

18 . Breiman , L . . Statistical modeling : The two cultures . Statistical Science . 2001 ; 16 : 199 - 231 .

19 . . Petricoin , III , EF , Ardkani , AM , Hitt , BA , Levine , PJ , Fusaro , VA , Steinberg , SM , et al . Use of proteomic patterns in serum to identify ovarian cancer . The Lancet 2002 ; 359 : 572 - 577 .

図面および先行の記載で本発明を詳細に例示および説明したが、これは説明として考慮され、特徴を制限するものではない。好ましい実施形態を図示および記載しただけであり、本発明の精神に逸脱しない全ての変更および改変が保護されることを望んでいることが理解される。さらに、本明細書で引用した全ての参考文献は、当技術の技術レベルを示すものであり、全体として参照により本明細書に組み込まれる。

【図面の簡単な説明】

【0177】

【図1】図1は、実施例1により完全に記載したように、ピーク検出から試料分類に至るプロセスがまとめられているフローチャートを示す。

【図2】図2Aは実施例1で用いた決定木分類系の模式図、図2Bは上記分類系の種々の特徴を示すSELDIタンパク質プロファイルを示す図、および図2Cは長時間にわたって試料を保存した後に得られたSELDIタンパク質プロファイルである。

【図3】図3は、実施例2により完全に記載したようなSELDIを用いて、2,000~40,000ダルトン間での分解ピークの典型的な低スペクトルを示す図であり、上側のスペクトルが約2,000から約10,000ダルトンの分子量範囲を持つ分解ピークのスペクトルを示し、下側のスペクトルは約10,000から約40,000ダルトンの分子量範囲を持つ分解ピークのスペクトルを示す。

【図4】図4は、実施例2に記載した促進決定木分類子に対する学習誤差率最小限界 (training error rate 最小限界) および / または試験誤差率 (test error rate) を示すグラフである。図4Aは、非癌と癌とを区別する促進決定木分類子の基部 (base stumps) の数であるMの学習誤差率、最小限界、および一般化誤差率 (試験誤差) を示す。学習誤差がゼロ (ラウンド47) に達した後、最小限界は増加し続け、一般化誤差は減少し続けて最終的にゼロに達する (ラウンド265 で) 。図4Bは、正常とBPHとを区別する促進決定木分類子の基部数に対する学習誤差率および最小限界を示す。学習誤差がゼロ (ラウンド9) に達した後、最小限界は増加し続ける。

【図5】図5は、本発明のコンピュータ実現実施形態にもとづくコンピュータ・プロセスを実行するための中央処理装置の一例を示す。

【図6】図5に示した中央処理装置内部ハードウェアの内部ハードウェアのブロック図の一例を示す。

【図7】コンピュータ命令が組み込まれるコンピュータ可読媒体を説明する図である。

【 図 1 】

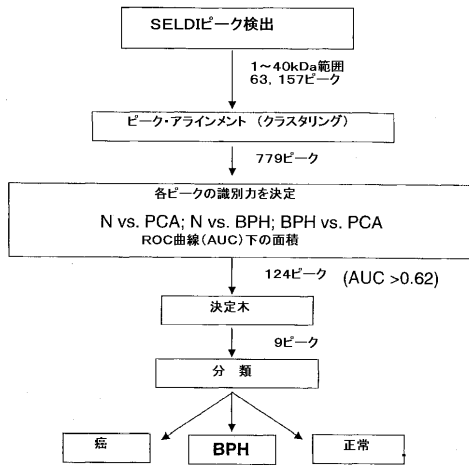


FIG. 1

【 図 2 】

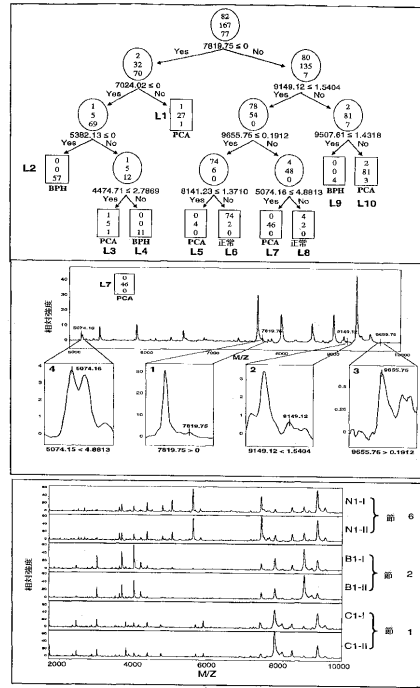


FIG. 2

【 図 3 】

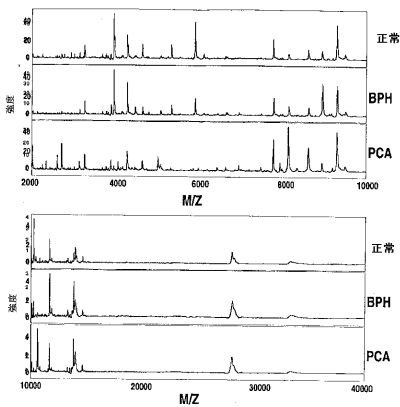


FIG. 3

【 図 4 】

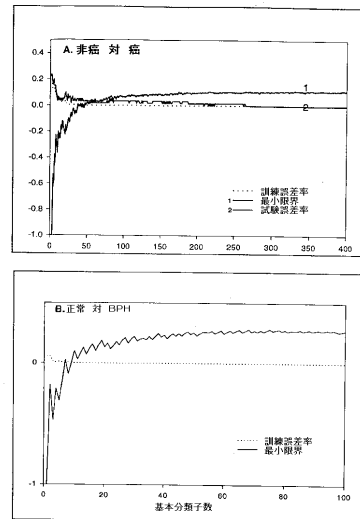


FIG. 4

【 図 5 】

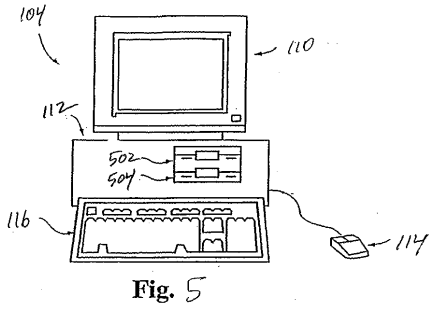


Fig. 5

【 図 7 】

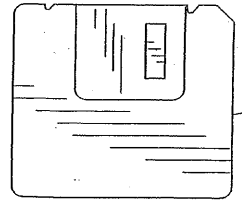


Fig. 7

【 図 6 】

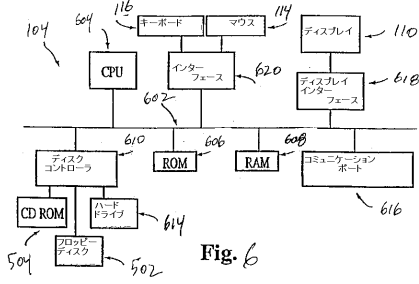


Fig. 6

フロントページの続き

(51) Int.Cl.	F I	テーマコード(参考)
B 0 1 J 20/281 (2006.01)	G 0 1 N 30/26	E
G 0 1 N 30/88 (2006.01)	G 0 1 N 30/48	R
G 0 1 N 33/53 (2006.01)	G 0 1 N 30/88	J
	G 0 1 N 33/53	D

(81) 指定国 AP(GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), EA(AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), EP(AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PT, RO, SE, SI, SK, TR), OA(BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG), AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW

(74) 代理人 100062409

弁理士 安村 高明

(74) 代理人 100113413

弁理士 森下 夏樹

(72) 発明者 ライト, ジョージ エル., ジュニア

アメリカ合衆国 バージニア 2 3 4 5 5, バージニア ビーチ, モルツリー シーティー.
8 2 9

(72) 発明者 アダム, パオ - リン

アメリカ合衆国 バージニア 2 3 5 0 8, ノーフォーク, ウェスト 5 2 エヌディー スト
リート 7 6 9

(72) 発明者 ク, インシェン

アメリカ合衆国 ワシントン 9 8 1 1 5, シアトル, ノースイースト, 4 0 ティーエイチ
アベニュー 6 5 5 6

F ターム(参考) 2G041 CA01 CA02 CA04 DA03 DA04 DA16 FA12 GA06

专利名称(译)	前列腺癌生物标志物		
公开(公告)号	JP2006509186A	公开(公告)日	2006-03-16
申请号	JP2004541437	申请日	2003-05-09
[标]申请(专利权)人(译)	东弗吉尼亚医学院 弗莱德哈钦森癌症研究中心		
申请(专利权)人(译)	东弗吉尼亚医学院 弗雷德·哈钦森癌症研究中心		
[标]发明人	ライトジョージエルジュニア アダムバオリン クインシエン		
发明人	ライト, ジョージ エル., ジュニア アダム, バオリン ク, インシエン		
IPC分类号	G01N33/574 G01N27/62 G01N27/64 G01N30/00 G01N30/26 B01J20/281 G01N30/88 G01N33/53 A61B A61B1/00 C12Q1/68 C12Q1/70 G01N33/68 G06F19/00 G06F19/18 G06F19/24		
CPC分类号	G01N33/57434 C12Q1/6886 G01N33/6851 G01N2800/52 G16B20/00 G16B40/00		
FI分类号	G01N33/574.A G01N27/62.F G01N27/62.V G01N27/64.B G01N30/00.E G01N30/26.E G01N30/48.R G01N30/88.J G01N33/53.D		
F-TERM分类号	2G041/CA01 2G041/CA02 2G041/CA04 2G041/DA03 2G041/DA04 2G041/DA16 2G041/FA12 2G041 /GA06		
代理人(译)	夏木森下		
优先权	60/379018 2002-05-10 US		
其他公开文献	JP2006509186A5		
外部链接	Espacenet		

摘要(译)

前列腺癌，在诊断良性前列腺肥大，或在做出否定的诊断中，描述了可以利用有利的蛋白生物标志物。因此，在本发明的一个方面，它有助于在前列腺癌或良性前列腺增生症的诊断，或以其他方式提供了用于使诊断的方法。在本发明的一种形式包括：检测具有预定分子量的多种蛋白质的生物标志物，关联所述检测前列腺癌的诊断和良性前列腺增生症，或者一个否定的诊断。在本发明的另一个方面，是提供用于本文所述的生物标记物的可检测的试剂盒。在本发明的进一步的，通过使用多个分类器，使得诊断有前列腺癌的前列腺增生一个可能性的方法。在本发明的一个具体实施方案中，该方法涉及使用加速决策树分析。还提供各种计算机可读介质。

節	クラス	観察	確率	5%信頼水準
L1	正常	1	0.0625	0.0081, 0.1693
	PCA	27	0.8750	0.7423, 0.9530
	BPH	1	0.0625	0.0087, 0.1698
L2	正常	0	0.0167	0.0005, 0.0584
	PCA	0	0.0167	0.0004, 0.0528
	BPH	57	0.9667	0.8072, 0.9952
L3	正常	1	0.2000	0.0247, 0.4793
	PCA	5	0.6000	0.3027, 0.8592
	BPH	1	0.2000	0.0248, 0.4753
L4	正常	0	0.0714	0.0018, 0.2509
	PCA	0	0.0714	0.0019, 0.2579
	BPH	11	0.8571	0.6311, 0.9823
L5	正常	0	0.1429	0.0040, 0.4725
	PCA	4	0.7143	0.3557, 0.9567
	BPH	0	0.1429	0.0040, 0.4504
L6	正常	74	0.9494	0.8950, 0.9858
	PCA	2	0.0380	0.0032, 0.0879
	BPH	0	0.0127	0.0003, 0.0458
L7	正常	0	0.0204	0.0005, 0.0738
	PCA	46	0.9582	0.8893, 0.9951
	BPH	0	0.0204	0.0005, 0.0726
L8	正常	4	0.5556	0.2458, 0.8416
	PCA	2	0.3333	0.0836, 0.6544
	BPH	0	0.1111	0.0032, 0.3732
L9	正常	0	0.1429	0.0034, 0.4560
	PCA	0	0.1429	0.0037, 0.4830
	BPH	4	0.7143	0.3354, 0.9566
L10	正常	2	0.0337	0.0068, 0.0784
	PCA	81	0.9213	0.8595, 0.9674
	BPH	3	0.0449	0.0123, 0.0964