

(19) 世界知的所有権機関  
国際事務局



(43) 国際公開日  
2006年12月7日 (07.12.2006)

PCT

(10) 国際公開番号  
WO 2006/129401 A1

- (51) 国際特許分類:  
G01N 27/62 (2006.01)
- (21) 国際出願番号: PCT/JP2006/302594
- (22) 国際出願日: 2006年2月8日 (08.02.2006)
- (25) 国際出願の言語: 日本語
- (26) 国際公開の言語: 日本語
- (30) 優先権データ:  
特願2005-159562 2005年5月31日 (31.05.2005) JP
- (71) 出願人 (米国を除く全ての指定国について): 株式会社 JCLバイオアッセイ (JCL BIOASSAY CORPORATION) [JP/JP]; 〒5600082 大阪府豊中市新千里東町1-4-2 千里ライフサイエンスセンタービル16F Osaka (JP).
- (72) 発明者; および
- (75) 発明者/出願人 (米国についてのみ): 後藤 理恵子

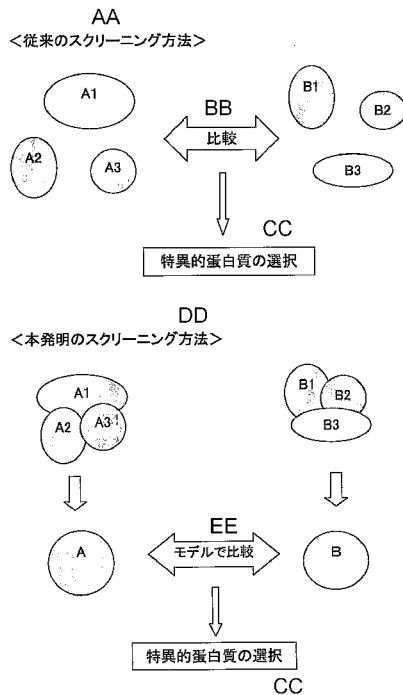
- (GOTO, Rieko) [JP/JP]; 〒6200841 京都府福知山市駒場新町2丁目105 Kyoto (JP). 塩山 昇平 (SHIOYAMA, Shohei) [JP/JP]; 〒6710255 兵庫県姫路市花田町小川746-3 フレグランス小川A101 Hyogo (JP). 戸塚 善三郎 (TOZUKA, Zenzaburo) [JP/JP]; 〒5620035 大阪府箕面市船場東3-2-35-2203 Osaka (JP). 靱山 邦男 (MOMIYAMA, Kunio) [JP/JP]; 〒5760053 大阪府交野市郡津1丁目25-20 Osaka (JP). 中村 靖司 (NAKA-MURA, Yasushi) [JP/JP]; 〒6408136 和歌山県和歌山市堀止東1丁目5-3 Wakayama (JP). 覚道 健一 (KAKUDO, Kennichi) [JP/JP]; 〒6310006 奈良県奈良市西登美ヶ丘3丁目11-2 Nara (JP).
- (74) 代理人: 南條 博道 (NANJO, Hiromichi); 〒5300047 大阪府大阪市北区西天満3丁目2番9号 翁ビル5階 Osaka (JP).

- (81) 指定国 (表示のない限り、全ての種類の国内保護が可能): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR,

[ 続葉有 ]

(54) Title: SCREENING METHOD FOR SPECIFIC PROTEIN IN PROTEOME COMPREHENSIVE ANALYSIS

(54) 発明の名称: プロテオーム網羅的解析における特異的蛋白質のスクリーニング方法



AA CONVENTIONAL SCREENING METHOD  
 BB COMPARISON  
 CC SELECTION OF SPECIFIC PROTEIN  
 DD SCREENING METHOD OF THE INVENTION  
 EE COMPARISON USING MODEL

(57) Abstract: A screening method for a specific protein in a proteome comprehensive analysis of the invention comprises the steps of (a) obtaining a sample containing a protein or a digested substance of a protein from a cell or a tissue of a specific group or a control group; (b) analyzing the sample obtained in the step (a) with a mass spectrometer and obtaining mass spectrometry data; (c) analyzing the mass spectrometry data obtained in the step (b) using an arbitrary database searching software and obtaining a protein list containing factors to specify a protein and indexes to identify a protein for each sample; (d) averaging the values of the indexes for each of the factors in all the protein lists obtained in the step (d) and obtaining protein list models of the specific group and the control group containing the average values of indexes; (e) calculating the difference of the average values for each of the factors between the protein list models of the specific group and the control group obtained in the step (d) and obtaining one protein list in which the factors are sorted in the order of the difference of the average values; and (f) selecting a protein whose difference in the average values is large from the protein list obtained in the step (e).

(57) 要約: 本発明のプロテオーム解析における特異的蛋白質のスクリーニング方法は、(a) 特異群および対照群の細胞または組織から、蛋白質または蛋白質消化物を含む試料を得る工程; (b) 該工程(a)で得られた試料を質量分析計で分析して、質量分析データを得る工程; (c) 該工程(b)により得られた質量分析データを、任意のデータベース検索ソフトウェアを用いて解析し、該試料ごとに、蛋白質を特定する要素および蛋白質を同定するための指標を含む蛋白質リストを取得する工程; (d) 該工程(c)で取得したすべての蛋白質リストにおいて、該要素ごとに該指標の値を平均化し、該指標の平均値を含む特異群および対照群の蛋白質リストモデルを取得する工程; (e) 該工程(d)で得られた特異群および対照群の蛋白質リストモデル間で該要素ごとに該平均値の差を算出して、該要素を該平均値の差の順に並べ替えた1つの蛋白質リストを取得する工程; および(f) 該工程

(e) で取得した蛋白質リストから、該平均値の差が大きい蛋白質を選択する工程を含む。

WO 2006/129401 A1



BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, LY, MA, MD, MG, MK, MN, MW, MX, MZ, NA, NG, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SM, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

- (84) 指定国 (表示のない限り、全ての種類の広域保護が可能): ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW); ユーラシア (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), ヨーロッパ (AT, BE, BG,

CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IS, IT, LT, LU, LV, MC, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

添付公開書類:

— 国際調査報告書

2文字コード及び他の略語については、定期発行される各PCTガゼットの巻頭に掲載されている「コードと略語のガイダンスノート」を参照。

## 明 細 書

## プロテオーム網羅的解析における特異的蛋白質のスクリーニング方法

## 5 技術分野

本発明は、プロテオーム網羅的解析における特異的蛋白質のハイスループットスクリーニング方法に関する。

## 背景技術

10 創薬や医療診断の基礎研究としてゲノミクスやプロテオミクスがある。ゲノミクスは、DNAマイクロアレイ、DNAチップなどの有効な分析ツールの開発およびこれらの実用化により、ヒト遺伝子が全て解明されるなどの成果を修めている。プロテオームの網羅的解析（プロテオミクス）も、蛋白質の構造および量の異常で起こる病気に対して、その蛋白質を特定し、診断方法、治療方法、および治療薬を開発するために盛んに行われている。しかし、  
15 プロテオミクスは1980年代にスタートしたにもかかわらず、サンプルの倫理的問題やゲノミクスにおけるDNAチップのような包括的に解析できるツールが開発されていないことなどのため、未だ大きな成果を挙げていない（ゲノミクス・プロテオミクスの新展開，今中忠行監修，2004年，株式会社エヌ・ティー・エス）。

20

一般に、蛋白質の研究において、分離精製には電気泳動や特異的吸着を利用するカラムクロマトグラフィーを用い、分析にはプロテインシーケンサー、NMR、X線解析を用いる（ゲノミクス・プロテオミクスの新展開，今中忠行監修，2004年，株式会社エヌ・ティー・エスおよびタンパク質実験ノート  
25 （上・下），改訂第3版，岡田雅人および宮崎香編，2004年，羊土社）。これらの手法は、コストがかかること、再現性のあるデータの取得が困難であ

ること、分析時間が長いことなどの問題があった。近年、質量分析計の著しい発展に伴い、質量分析計を用いたプロテオミクスが行われている。蛋白質分離後の測定方法としては、ESI、MALDIなどのイオン化方法を用いた質量分析計を用いる。

- 5 現在使用されている蛋白質混合物の分離方法として、蛋白質の等電点および大きさの違いに基づいて分離する2次元電気泳動がある。また、酵素消化後のペプチドを分離する方法として、イオン交換カラムと逆相カラムとを組み合わせた2次元HPLCもある (S.P. Gygiら, J. Proteome Research, 2003年, 43巻, pp. 43-50)。このような2次元電気泳動あるいは2次元HPLC (2DLC) と質量分析計とを組み合わせることによって、蛋白質の分離精製の必要がないプロテオーム解析方法が開発されている (S.P. Gygiら, J. Proteome Research, 2003年, 43巻, pp. 43-50およびS.P. Gygiら, J. Mass Spectrom., 2001年, 36巻, pp. 1083-1091)。最近の測定方法では、蛋白質をそのまま質量分析計に注入する、ECD-FTICRMS<sup>n</sup>やETD/LTQMS<sup>n</sup>のトップダウンシーケンスの手法もある (R.A. Zubarevら, J. Am. Chem. Soc., 1998年, 120巻, pp. 3265-3266 ; R.A. Zubarevら, Curr. Opin. Biotechnol., 2004年, 15巻, pp. 12-16 ; J.E. Sykaら, Proc. Natl. Acad. Sci. U.S.A., 2004年, 101巻, pp. 9528-9533 ; およびJ.J. Coonら, Int. J. Mass Spectrom., 2004年, 236巻, pp. 33-42)。

- 20 一般的に、特異的蛋白質をスクリーニングする場合、目的蛋白質を含む細胞または組織、および目的蛋白質を含まない細胞または組織の2種類を用意し、この2種類の細胞または組織から抽出された各試料について蛋白質の同定を行い、その同定結果を比較する。プロテオーム解析の場合は、各細胞または組織から蛋白質を分画・精製し、得られた蛋白質混合物を、蛋白質分解酵素によりペプチドフラグメントに分解し、ペプチドフラグメントを測定する。測定結果と蛋白質分解酵素の情報とを組み合わせ、ゲノムのデータベ

ースと照合し、蛋白質を同定する。これらの質量分析によって得られたデータについてのデータベース検索ソフトが市販されている。

上述のように、プロテオーム解析方法については、様々な方法がある。しかし、いずれの方法においても、例えば以下の理由から、種類の異なる蛋白質検索結果を比較して、効率的に特異的蛋白質をスクリーニングする方法は  
5 確立されていない。

(1) 検索結果から得られる蛋白質の種類が非常に多く、データが膨大であること；

(2) 蛋白質の多くは発現量の多い蛋白質であり (S.P. Gygiら, Mol. Cell Biol., 1999年, 19巻, p.1720およびS.P. Gygiら, Proc. Natl. Acad. Sci. U.S.A., 2000年, 97巻, pp.9390-9395)、発現量の少ない蛋白質は、その発現量の変化を見つけることが非常に困難であること；  
10

(3) 溶解性の悪い蛋白質では細胞からの抽出による再現性が要求されること；

(4) 細胞分画や蛋白質の粗精製・濃縮操作による再現性が要求されること；  
15

(5) 酵素消化処理での再現性が要求されること；

(6) 上記(3)～(5)の問題を解消するために、内部標準物質を加える方法があるが、適切な内部標準物質が必要であり、多量の内部標準物質と発現量の少ない蛋白質とが重なると、その蛋白質を検出することが困難となること；および  
20

(7) 上記(3)～(5)の問題を解消するために、ICAT (isotope-coded affinity tag; 同位体標識アフィニティータグ) 試薬を蛋白質のシステイン残基に結合させる方法があり、これは微量な蛋白質の発現を比較するために有効な手段であるが、ICAT試薬を必要とすること (S.P. Gygiら, Nat. Biotechnol., 1999年, 17巻, pp.994-999)。  
25

上記のような膨大なデータを解析するためのデータ処理方法も検討されている（特開2005-031021号公報）。しかし、処理して得られたデータが、実際に特異的蛋白質のスクリーニングに有効であるかどうかについては十分に検討されていない。

5

#### 発明の開示

上述のプロテオミクスの手法は、コスト面や分析時間、データの再現性の問題をある程度解消し、多量の未知の蛋白質混合物を包括的に分析できる手法として将来的に医療診断に応用が期待されている。しかし、包括的な分析を行うには非常に膨大なデータ処理が必要であること、さらに質量分析計を用いたプロテオミクス特有の擬陽性データを排除しきれないこと、定量的な考察が困難であることなどの問題点があり、実用化は非常に困難な状況にある。

10

本発明は、膨大な蛋白質のハイスループットな機能解析が求められるプロテオーム解析において、特異的蛋白質を効率よくハイスループットにスクリーニングするための新規な方法を提供することを目的とする。

15

本発明は、プロテオーム解析における特異的蛋白質のスクリーニング方法を提供し、該方法は、

(a 1) 特異群の細胞または組織から、蛋白質または蛋白質消化物を含む試料を得る工程；

20

(a 2) 対照群の細胞または組織から、蛋白質または蛋白質消化物を含む試料を得る工程；

(b 1) 該工程 (a 1) で得られた試料を質量分析計で分析して、質量分析データを取得する工程；

25

(b 2) 該工程 (a 2) で得られた試料を質量分析計で分析して、質量分析データを取得する工程；

(c 1) 該工程 (b 1) により得られた質量分析データを、任意のデータベース検索ソフトウェアを用いて解析し、該試料ごとに、蛋白質を特定する要素および蛋白質を同定するための指標を含む蛋白質リストを取得する工程；

5 (c 2) 該工程 (b 2) により得られた質量分析データを、任意のデータベース検索ソフトウェアを用いて解析し、該試料ごとに、蛋白質を特定する要素および蛋白質を同定するための指標を含む蛋白質リストを取得する工程；

10 (d 1) 該工程 (c 1) で取得したすべての蛋白質リストにおいて、該要素ごとに該指標の値を平均化し、該指標の平均値を含む特異群の蛋白質リストモデルを取得する工程；

(d 2) 該工程 (c 2) で取得したすべての蛋白質リストにおいて、該要素ごとに該指標の値を平均化し、該指標の平均値を含む対照群の蛋白質リストモデルを取得する工程；

15 (e) 該工程 (d 1) で得られた特異群の蛋白質リストモデルと該工程 (d 2) で得られた対照群の蛋白質リストモデルとの間で該要素ごとに該平均値の差を算出して、該要素を該平均値の差の順に並べ替えた1つの蛋白質リストを取得する工程；および

20 (f) 該工程 (e) で取得した蛋白質リストから、該平均値の差が大きい蛋白質を選択する工程；  
を含む。

好適な実施態様では、上記蛋白質を同定するための指標は、スコア、Coverage、またはランキングである。

25 より好適な実施態様では、上記蛋白質を同定するための指標は、スコアである。

好適な実施態様では、上記蛋白質を特定する要素は、Accession numberま

たは蛋白質名である。

好適な実施態様では、上記工程（d 1）、（d 2）、および（e）は、任意のコンピュータプログラムによって実行される。

5 本発明の方法によれば、多量の未知の蛋白質混合物を包括的に分析した際の膨大なデータを解析するための手法が提供され、実験誤差および擬陽性データを  
取り除いて、特異的蛋白質の候補を効率よく絞り込むことができる。本発明の方法では、従来のプロテオーム解析と比較して、スクリーニングの結果の再現性および正確性が向上している。また、本発明の方法は、比較的  
10 低コストかつハイスループットでのスクリーニングが可能である。さらに、本発明のスクリーニング方法によって選択された特異的蛋白質についての半定量的な判定も可能である。

#### 図面の簡単な説明

15 図1は、従来のスクリーニング方法および本発明のスクリーニング方法の原理を説明するための概念図である。

図2は、ヒト由来肝細胞のエストロゲンレセプター（A）およびグルタミン酸レセプター（B）についてのスコア値を示すグラフである。

図3は、各症例別のスコアの分布を示すグラフである。

20 図4は、特異的な3種類の蛋白質（A～C）についての、各試料のスコア値を示すグラフである。

図5は、特異的な3種類の蛋白質（D～F）についての、各試料のスコア値を示すグラフである。

図6は、種々の範囲のモデルスコア値に対応する蛋白質名およびAccession numberの数を示すグラフである。

25 図7は、各試料のスコア35以上のAccession numberの数および蛋白質の濃度を示すグラフである。

### 発明を実施するための最良の形態

本発明のプロテオーム解析における特異的蛋白質のスクリーニング方法は、種々の要因（例えば、病態、薬物への曝露など）によって発現量が特異的に変動する蛋白質を特定するために特に有用な方法である。

5 一般的に、プロテオーム解析における特異的蛋白質のスクリーニングにおいては、特異的蛋白質が発現していると予想される場合であっても、蛋白質が存在するか否かは不明であり、特異的と判断するための基準も、蛋白質の種類、処理方法、試料濃度、個体差などのため一定ではない。これに対しては、n数を増やすことにより、スクリーニング結果の信頼性を高めることができる。しかし、n数の増加に伴って、リストアップされる蛋白質の数が増え、処理すべきデータが膨大になる。これらを個々について検討することは、非常に大きな労力および時間を必要とする（図1の従来のスクリーニング方法の概念図を参照のこと）。

15 そこで、本発明者らは、蛋白質を特異的と判断する基準が曖昧であること、試料間で再現性の問題があること、ならびに個体差が各検索結果に反映されていることに注目した。試料を、例えば、特異的蛋白質の変動があると予想される群とそれに対する対照群とに群分けし、各群内で個々の試料の質量分析データの解析によって得られる蛋白質を特定する要素（蛋白質名、Accession numberなど）ごとに、蛋白質の同定をするための指標（スコア、Coverage、ランキングなど）の平均値を算出することにより、各群の各蛋白質のモデル的な指標値を求めることができる。図1の本発明のスクリーニング方法の概念図に基づいて説明すると、A群の各試料A1～A3についての蛋白質リストモデルAおよびB群の各試料B1～B3についての蛋白質リストモデルBをそれぞれ作成し、このモデル間で比較を行う。比較とは、具体的には、  
20 各要素の指標の差を求めることである。次いで、その差に基づいて蛋白質リストを整理する。こうして得られた蛋白質リストを用いると、特異的蛋白質

を絞り込むことが容易である。例えば、整理された蛋白質リストの蛋白質Xという要素のモデル指標値を $X_A$ および $X_B$ とすると、 $X_A - X_B$ 値が大きい場合はXがA群に特異的な蛋白質であり、小さい場合はB群に特異的な蛋白質であると容易に判断できる。

5       したがって、本発明のプロテオーム解析における特異的蛋白質のスクリーニング方法は、以下の工程：

      (a 1) 特異群の細胞または組織から、蛋白質または蛋白質消化物を含む試料を得る工程；

10       (a 2) 対照群の細胞または組織から、蛋白質または蛋白質消化物を含む試料を得る工程；

      (b 1) 該工程 (a 1) で得られた試料を質量分析計で分析して、質量分析データを取得する工程；

      (b 2) 該工程 (a 2) で得られた試料を質量分析計で分析して、質量分析データを取得する工程；

15       (c 1) 該工程 (b 1) により得られた質量分析データを、任意のデータベース検索ソフトウェアを用いて解析し、該試料ごとに、蛋白質を特定する要素および蛋白質を同定するための指標を含む蛋白質リストを取得する工程；

20       (c 2) 該工程 (b 2) により得られた質量分析データを、任意のデータベース検索ソフトウェアを用いて解析し、該試料ごとに、蛋白質を特定する要素および蛋白質を同定するための指標を含む蛋白質リストを取得する工程；

25       (d 1) 該工程 (c 1) で取得したすべての蛋白質リストにおいて、該要素ごとに該指標の値を平均化し、該指標の平均値を含む特異群の蛋白質リストモデルを取得する工程；

      (d 2) 該工程 (c 2) で取得したすべての蛋白質リストにおいて、該要

素ごとに該指標の値を平均化し、該指標の平均値を含む対照群の蛋白質リストモデルを取得する工程；

5 (e) 該工程 (d 1) で得られた特異群の蛋白質リストモデルと (d 2) で得られた対照群の蛋白質リストモデルとの間で該要素ごとに該平均値の差を算出して、該要素を該平均値の差の順に並べ替えた1つの蛋白質リストを取得する工程；および

(f) 該工程 (e) で取得した蛋白質リストから、該平均値の差が大きい蛋白質を選択する工程；を含む。

10 以下、本発明を工程の順に詳細に説明する。

工程 (a 1) および (a 2) ；

本発明の方法では、まず、工程 (a 1) および (a 2) において、特異群および対照群の細胞または組織から、それぞれ蛋白質または蛋白質消化物を含む試料を得る。

15 「特異群」とは、スクリーニングの対象となる群をいい、発現量が特異的に変動している蛋白質の存在が予想される群をいう。例えば、特定の病態群；化学物質、光、温度などの特定の条件への曝露群が挙げられる。「対照群」とは、上記特異群と比較するための群であり、例えば、特定の病態ではない群（例えば、正常群）；種々の条件の曝露などを受けていない群が挙げられる。「細胞または組織」とは、上記の特異群および対照群に由来する、  
20 単離された細胞または組織をいう。例えば、培養細胞、血液細胞、バイオプシーにより体内から取り出された組織または細胞などが挙げられる。

まず、組織を用いる場合は、蛋白質分解酵素処理、例えば、コラゲナーゼ処理などの当業者が通常用いる手段によって細胞を分離する。細胞または組  
25 織から分離した細胞は、適切な緩衝液中で、例えば、ホモジナイザーなどの当業者が通常用いる手段を用いて破砕される。蛋白質を含む試料は、この破

砕によって得られた懸濁液自体であってもよく、あるいは必要に応じてさらに分画して得られた画分であってもよい。蛋白質を含む試料は、必要に応じて、さらにトリプシンなどの蛋白質消化酵素で消化してもよく、この消化処理によって、蛋白質消化物を含む試料を得ることができる。

- 5       これらの工程 (a 1) および (a 2) において、各群の試料の n 数は、特に限定されないが、試料の個体差の影響を除くことができる点で、n 数が多い方が好ましい。

      工程 (b 1) および (b 2) :

- 10       この工程 (b 1) および (b 2) では、上記工程 (a 1) および (a 2) で得られた各群の試料を質量分析計で分析して、試料ごとの質量分析データを  
      を得る。

- 15       「質量分析」 (MS) とは、分析する試料をイオン化させて導入し、電気力や磁気力により質量ごとの差をつくり、イオンの質量を分析することである。MSの測定原理としては、イオントラップ型MS法、フーリエ変換イオンサイクロトロン共鳴質量分析 (FT-ICR/MS) 法、イオンスキャン法、Q-TOF型MS法などが挙げられ、それぞれの原理に基づく質量分析計が存在する。本発明の方法においては、1方法だけ (すなわち、1つの質量分析計のみ) で分析してもよく、あるいは複数の質量分析計を連結させて分析 (以下MS/MS解析) してもよい。

- 20       工程 (c 1) および (c 2) :

      これらの工程 (c 1) および (c 2) では、上記工程 (b 1) および (b 2) で得られた各試料の質量分析データを、任意のデータベース検索ソフトウェアを用いて解析し、試料ごとに、蛋白質を特定する要素および蛋白質を同定するための指標を含む蛋白質リストを取得する。

- 25       「データベース検索ソフトウェア」は、MSデータから分子量が合致するペプチドフラグメントの候補の検出、さらにこのフラグメントから、任意の

データベースを検索して蛋白質全体を予測する解析ソフトウェアであればよい。市販されているソフトウェアとしては、Mascot (Matrix Science社)、Turbo Sequest (サーモエレクトロン社) などが挙げられる。利用可能なデータベースとしては、例えば、BLAST、Swiss-Protが挙げられる。このようなデータベース検索ソフトウェアは、MSデータを出力するために質量分析計とともに備えられているコンピュータ部に、予め組み込まれていることが好ましい。

上記データベース検索ソフトウェアによる解析の結果、蛋白質を特定する要素および該特定された蛋白質の同定をするための指標などを含む、試料ごとの蛋白質リストが得られる。蛋白質を特定する要素としては、Accession number、蛋白質名などが挙げられる。また、蛋白質の同定をするための指標としては、スコア (Score)、Coverage、ランキングなどが挙げられる。

工程 (d 1) および (d 2) :

これらの工程 (d 1) および (d 2) では、上記工程 (c 1) および (c 2) で取得した各群のすべての蛋白質リストにおいて、要素ごとに指標の値を平均化し、指標の平均値を含む特異群および対照群のそれぞれの蛋白質リストモデルを取得する。言い換えると、各群について、蛋白質リストに挙げられたすべての蛋白質を特定する要素と、該要素に対応する指標の平均値とを、1つのリストに統合することによって、各群のモデル的な蛋白質リストを得ることができる。

これらの工程において、平均化する指標は、スコア、Coverage、ランキングなどのいずれか1つであり、好ましくはスコアである。

工程 (e) :

工程 (e) では、上記工程 (d 1) および (d 2) で得られた特異群および対照群の2つの蛋白質リストモデル間で、要素ごとに指標の平均値の差を算出して、要素を平均値の差の順に並べ替えた1つの蛋白質リストを取得す

る。平均値の差は、(特異群の値) - (対照群の値) で表され得る。したがって、平均値の差は、正の値から負の値まで種々であり得る。差の順は、昇順であっても降順であってもよい。並べ替えることによって、特異的蛋白質の情報として有用なものの順に蛋白質を整理することができるため、以下の

5 工程 (f) において、差の大きな蛋白質の選択が容易になる。

ここで、工程 (d 1) および (d 2) ならびに (e) では、これらの工程を実行するようにプログラムされたコンピュータソフトウェアを用いてデータ処理され得る。例えば、上記のデータベース検索ソフトウェアとともに、質量分析計のコンピュータ部に組み込まれていてもよい。あるいは、上記の

10 工程 (c 1) および (c 2) でデータベース検索ソフトウェアによって得られた蛋白質リストは、サーバーやパーソナルコンピュータ (PC) などにエクスポートされ得る。例えば、Turbo Sequestで得られたデータを、PC用の表計算ソフトウェアであるマイクロソフト社のMicrosoft Excel®にエクスポートするためのソフトウェアがある。データがエクスポートされ得るソフト

15 ウェアにおいて、上記工程 (d 1) および (d 2) ならびに (e) を実行するためのマクロプログラムを組むことが可能である。したがって、PCなどにおいてこのプログラムを実行することによって、平均値の差の順に並べ替えた1つの蛋白質リストを取得することも可能である。

工程 (f) :

20 工程 (f) では、上記工程 (e) において得られた平均値の差の順に並べ替えた1つの蛋白質リストから、平均値の差が大きい蛋白質を選択する。ここで、平均値の差が大きいとは、その絶対値が大きいことをいう。これによって、蛋白質リストに挙げられた膨大な数の蛋白質から、特異的蛋白質の候補を効率よく絞り込むことができる。

25 この工程で選択した蛋白質は、必ずしも特異的蛋白質ではない。なぜなら、例えば、n数が小さい場合には、特異群および対照群のいずれにおいても非

常に発現量の多い蛋白質において、差の値が大きくなる場合があり、その差が発現量の変動の範囲内であることがある。したがって、候補として選択された蛋白質について、特異的蛋白質であるかどうかを個々に検証する必要がある。

5 検証の手段は、特に限定されない。例えば、上記のスクリーニング方法で使用した特異群および対照群に属する別の複数の試料について、質量分析データを解析し、候補蛋白質の各試料における指標値を蛋白質リストモデル中の候補蛋白質の指標値と比較することによって、変動の範囲内の可能性が高いかどうか、あるいは特異的蛋白質であると同定し得るかどうかを検証し得る。  
10 本発明の方法によれば、この検証の作業はやや煩雑なように見える。しかし、1つの試料から数万の蛋白質がリストアップされ、それらを個々に比較検討／検証しなければならない従来のスクリーニング作業を比較すると、本発明の方法では、検証すべき蛋白質が数種から数十種に絞り込まれ得るため、非常に効率よく特異的蛋白質を同定できる。

15 さらに、本発明の方法によって同定された特異的蛋白質については、未知試料の質量分析データの解析によって得られた蛋白質リスト中のスコアなどの要素の値から、蛋白質リストモデル中の平均値との比較によって、特異群であるか否かの半定量的判定を行うことも可能である。

## 20 実施例

以下の実施例において、蛋白質試料またはペプチド試料の質量分析には、  
n a n o 2 D L C - M S <sup>®</sup> L T Q MSシステム（サーモエレクトロン社）  
を用いた。このシステムは、質量分析計として2DLC/ESI/リニアイ  
オントラップ/MS/MS（サーモエレクトロン社）が採用され、得られた  
25 質量分析データはデータベース検索ソフトウェアであるTurbo Sequest（サーモエレクトロン社）によって解析される。

上記システムによる解析結果として、各蛋白質のスコア値を含む個々の試料についての蛋白質リストが得られる。試料群についての平均のスコア値を求める場合は、各試料群内において蛋白質ごとにスコア値の平均値を算出する。算出した平均値について、蛋白質ごとに群間で平均スコア値の差を算出し、その順に蛋白質リストを並べ替える。以下の実施例においては、上記データベース検索ソフトウェアから得られた解析結果を、マイクロソフト社の Microsoft Excel®にエクスポートし、各試料群において蛋白質ごとの平均スコア値を含む蛋白質リストモデルを取得し、群間で蛋白質の平均スコア値の差を求めて、その差が大きい順に並べ替えた蛋白質リストを作成するようにマクロプログラムを組み、これを実行することによって、整理された蛋白質リストを得た。

(実施例 1)

ウシ血清アルブミン (BSA) を、以下の表 1 に記載の種々の濃度の水溶液に調製し、トリプシン消化した後、質量分析計でそれぞれ 2 回分析し、質量分析データをデータベース検索ソフトウェアで解析して、蛋白質リストを得た。各濃度における BSA と同定された蛋白質のスコア値を表 1 に示す。

表 1

BSA濃度 (fmol)	スコア値		
	1回目	2回目	平均
3	558.3	370.2	464.3
6	700.3	570.3	635.3
30	1114.5	902.3	1008.4
60	1644.3	1468.3	1556.3
300	2140.3	2230.3	2185.3
600	3676.3	4090.3	3883.3
3000	4652.4	5366.3	5009.4
6000	4538.3	4236.3	4387.3

表1からわかるように、得られたスコア値と蛋白質濃度との間に相関関係が見られた。

(実施例2)

以下の表2に示すヒト由来の肝細胞を洗浄後、緩衝液を加えて、氷冷下に  
5 破砕した。得られた懸濁液をトリプシン消化し、質量分析計で測定を行い、  
質量分析データをデータベース検索ソフトウェアで解析して、蛋白質リスト  
を得た。

表2

10

試料番号	性別	年齢	人種
1	女性	44	白人
2	男性	59	白人
3	女性	64	白人
4	男性	52	白人
5	男性	43	白人

15

エストロゲンレセプターおよびグルタミン酸レセプターについてのスコア  
値を、それぞれ図2AおよびBに示す。

20

エストロゲンレセプター (A) については、女性のスコア値の平均値は約  
90であり、男性では約30であった。エストロゲンは女性ホルモンである  
ため、その受容体が女性群でスコア値が大きいのは妥当な結果である。グル  
タミン酸レセプター (B) では、試料番号3 (64歳女性) のスコア値が大  
きく、これが老化に関与する蛋白質である可能性が示唆された。なお、本実  
施例においては、各試料の蛋白質リストの蛋白質数は5~6万個であり、そ  
のうちの約3割に相当する2万個が全ての試料で見られた蛋白質であった。

(実施例3)

25

ヒトにおけるある疾患において異なる病態を示す症例から採取した組織を  
用いた。一方の病態を示す6症例を対照群 (試料番号1~6)、および他方  
の病態を示す13症例を特異群 (試料番号7~19) とした。得られた組織

をそれぞれコラゲナーゼ処理して、細胞を分離した。細胞を洗浄後、氷冷下にて破碎した。得られた懸濁液を1, 000×gで遠心分離し、上清を回収して細胞質画分を得た。上清をトリプシン消化し、質量分析計で測定を行い、質量分析データをデータベース検索ソフトウェアで解析して、それぞれの症

5 例に由来する試料についての蛋白質リストを得た。

各試料においてスコア>2.0を満たすAccession numberは、平均で56,050個であった。スコアの範囲は、2.0~2000強であった。各試料別のスコアの分布を図3に示す。1症例あたりのAccession number数は平均で、スコア2.0以上3.5未満が50677、スコア3.0以上100.0未満が4942、およびスコア100以上が431であった。

10

上記の解析結果、すなわち各試料の蛋白質リストを、Microsoft Excel®にエクスポートし、Accession numberでスコアの平均値を求めて整理するマクロプログラムを実行した。対照群については、試料番号1~6のすべての試料についてマクロプログラムを実行し、対照群モデルスコア値を得た。特異

15 群については、試料番号7~19のすべてについて蛋白質リストをAccession numberごとに整理させたが、蛋白質リストモデルの作成については、特に病態の著しい試料番号7、10、11、および12のみを用いて実行し、特異群モデルスコア値を得た。

特異群モデルスコア値と対照群モデルスコア値との差 (Score of difference number) の大きい順に並べ替えたところ、19症例由来の試料の163780の総Accession number数のうち、Score of difference numberの上位20番以内で、6種類の特異的蛋白質A~Fを同定した。これらの蛋白質についての各試料のスコアを、図4および5に示す。これらの蛋白質は、対照群よりも特異群でスコア値が高い傾向にあるため、この病態の指標になり得

20

25 ることがわかる。

例えば、特異的な蛋白質の一例である蛋白質D (図5) を検証してみると、

特異群では、115～5587までのランキング値であり、対照群では、該当なしならびに6354～25515であった（データは示さず）。このように、蛋白質Dは、蛋白質の発現量が非常に少ないため、本発明の方法では特異的蛋白質であると同定可能であったが、従来のスクリーニング方法では見出すことができなかつたと考えられる。

スクリーニング方法の信頼性を確保するために、蛋白質を特定する要素として、Accession numberの代わりに蛋白質名を用いて上記と同様に蛋白質リストモデルを作成し、モデルスコア値の差（特異群モデルスコア値－対照群モデルスコア値）を算出した。種々の範囲のモデルスコア値の差に対応する蛋白質名およびAccession numberの数を、図6に示す。

19試料の総蛋白質数は、蛋白質名で検索した場合は75195個であり、Accession numberで検索した場合は163780個であった。Accession numberの方が88585個多いが、これは総蛋白質数の集計ではUnnamed proteinが含まれていないこと、ならびに異なるAccession numberでも同じ蛋白質名であるものは、蛋白質数の集計に含まれていないためである。

特異群モデルスコア－対照群モデルスコア値は、ほとんどが±5以内であり、特異的な蛋白質の検索にはどちらを用いても、大差はなかつた。また、スコア値の差が10以上の蛋白質は、いずれの場合も総蛋白質の量と比較するとごく僅かとなっていることが確認できる。特異的蛋白質は、スコア値の差が10以上の枠にあると考えられ、蛋白質名およびAccession numberのいずれで検索した場合も、選択された特異的蛋白質は同じであった（データは示さず）。

さらに、蛋白質の濃度の違いによるスコア値への影響についても確認した。ここでは、スクリーニング結果より、スコア35以上のAccession numberの数と蛋白質の濃度とを比較した結果を図7に示す。蛋白質の濃度に依存して、スコア値が変動していることがわかる。また試料番号3および9のように、

濃度が高いにもかかわらず、スコア値35以上のAccession numberの数が少ないものもあった。この原因として、質量分析測定での噴霧状態が悪くイオン化効率が悪かったこと、酵素消化での消化効率が悪かったことなどが考えられる。この結果からも、各試料を全く同条件で測定することが困難であることがわかる。したがって、このようばらつきのある試料の場合でも、本発明のスクリーニング方法が有効に利用できることがわかる。

#### 産業上の利用可能性

本発明の方法によれば、多量の未知の蛋白質混合物を包括的に分析した際の膨大なデータを解析するための手法が提供され、実験誤差および擬陽性データを統計的に取り除いて、特異的蛋白質の候補を効率よく絞り込むことができる。本発明の方法では、従来のプロテオーム解析と比較して、スクリーニングの結果の再現性および正確性が向上している。また、本発明の方法は、比較的 low コストかつ high スループットでのスクリーニングが可能である。さらに、本発明のスクリーニング方法によって選択された特異的蛋白質についての半定量的な判定も可能である。

したがって、本発明のスクリーニング方法によって種々の病態時や薬物などの曝露によって発現する特異的蛋白質を同定できる。そのため、この蛋白質が関連する疾患などの診断、治療、または予防、あるいはそのための薬剤の開発に非常に有用である。

## 請求の範囲

1. プロテオーム解析における特異的蛋白質のスクリーニング方法であつて：

5 (a 1) 特異群の細胞または組織から、蛋白質または蛋白質消化物を含む試料を得る工程；

(a 2) 対照群の細胞または組織から、蛋白質または蛋白質消化物を含む試料を得る工程；

10 (b 1) 該工程 (a 1) で得られた試料を質量分析計で分析して、質量分析データをj得る工程；

(b 2) 該工程 (a 2) で得られた試料を質量分析計で分析して、質量分析データをj得る工程；

15 (c 1) 該工程 (b 1) により得られた質量分析データを、任意のデータベース検索ソフトウェアを用いて解析し、該試料ごとに、蛋白質を特定する要素および蛋白質を同定するための指標を含む蛋白質リストを取得する工程；

20 (c 2) 該工程 (b 2) により得られた質量分析データを、任意のデータベース検索ソフトウェアを用いて解析し、該試料ごとに、蛋白質を特定する要素および蛋白質を同定するための指標を含む蛋白質リストを取得する工程；

(d 1) 該工程 (c 1) で取得したすべての蛋白質リストにおいて、該要素ごとに該指標の値を平均化し、該指標の平均値を含む特異群の蛋白質リストモデルを取得する工程；

25 (d 2) 該工程 (c 2) で取得したすべての蛋白質リストにおいて、該要素ごとに該指標の値を平均化し、該指標の平均値を含む対照群の蛋白質リストモデルを取得する工程；

(e) 該工程 (d 1) で得られた特異群の蛋白質リストモデルと該工程 (d 2) で得られた対照群の蛋白質リストモデルとの間で該要素ごとに該平均値の差を算出して、該要素を該平均値の差の順に並べ替えた1つの蛋白質リストを取得する工程；および

- 5 (f) 該工程 (e) で取得した蛋白質リストから、該平均値の差が大きい蛋白質を選択する工程；  
を含む、方法。

10 2. 前記蛋白質を同定するための指標が、スコア、Coverage、またはランキングである、請求項1に記載の方法。

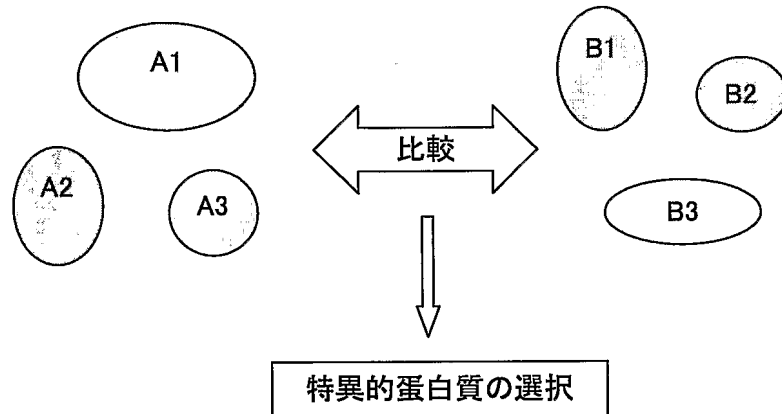
3. 前記蛋白質を同定するための指標が、スコアである、請求項2に記載の方法。

15 4. 前記蛋白質を特定する要素が、Accession numberまたは蛋白質名である、請求項1から3のいずれかに記載の方法。

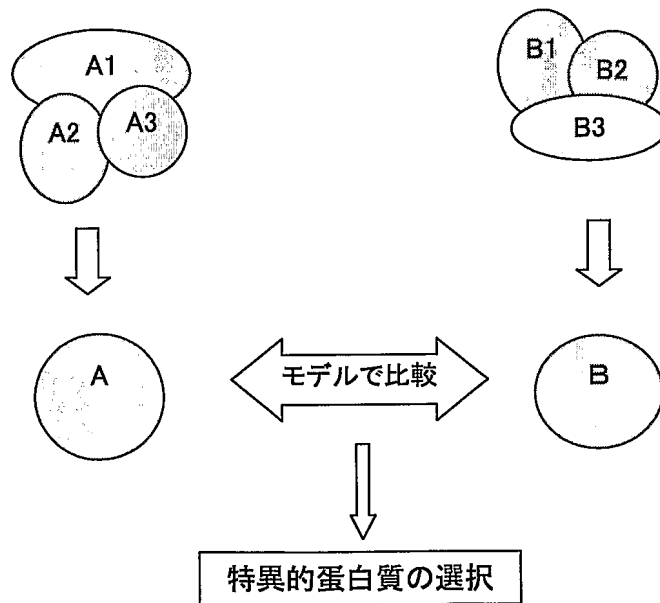
5. 前記工程 (d 1)、(d 2)、および (e) が、任意のコンピュータプログラムによって実行される、請求項1から4のいずれかに記載の方法。

第1図

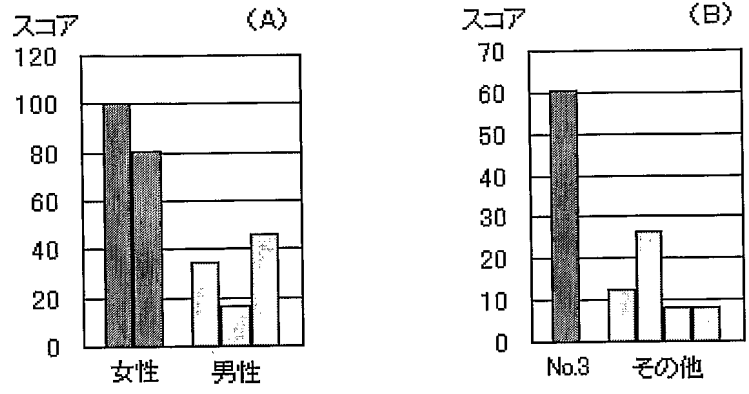
<従来のスクリーニング方法>



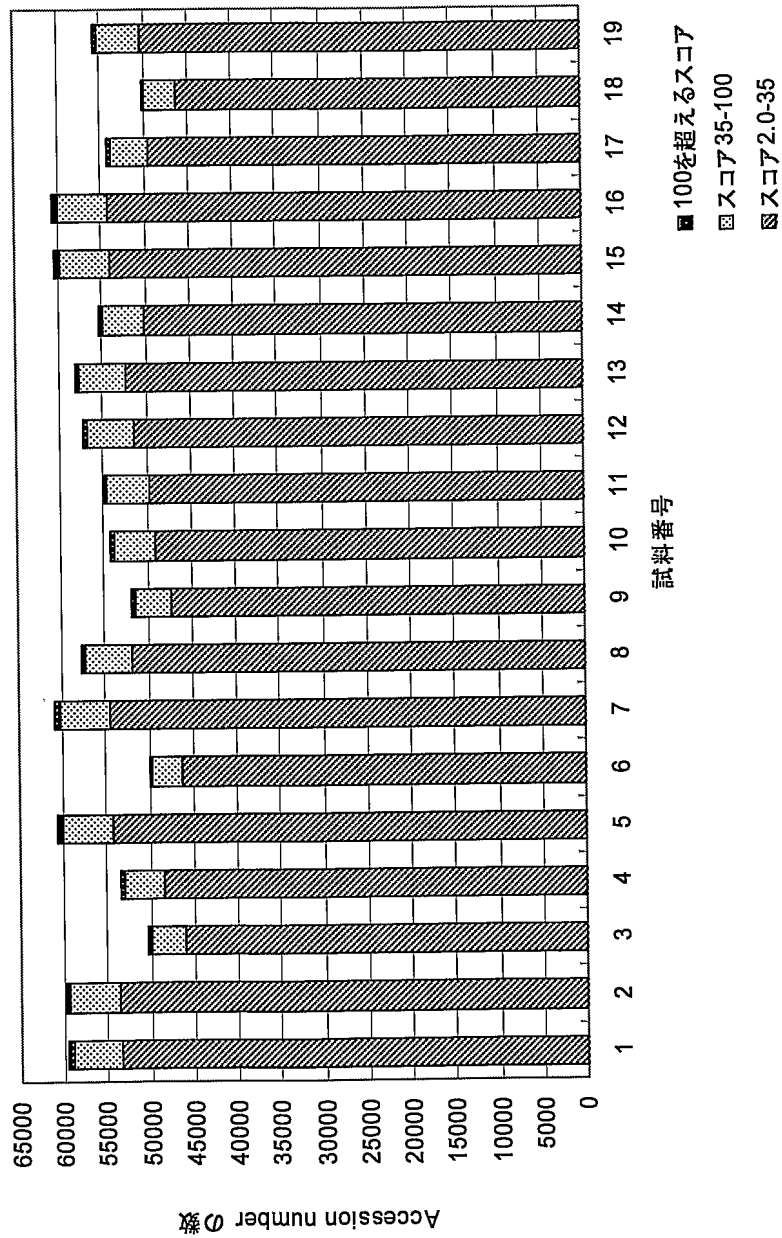
<本発明のスクリーニング方法>



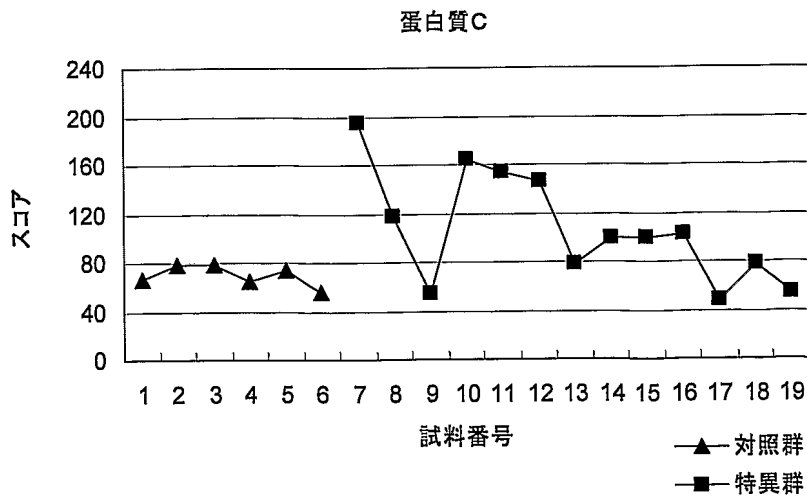
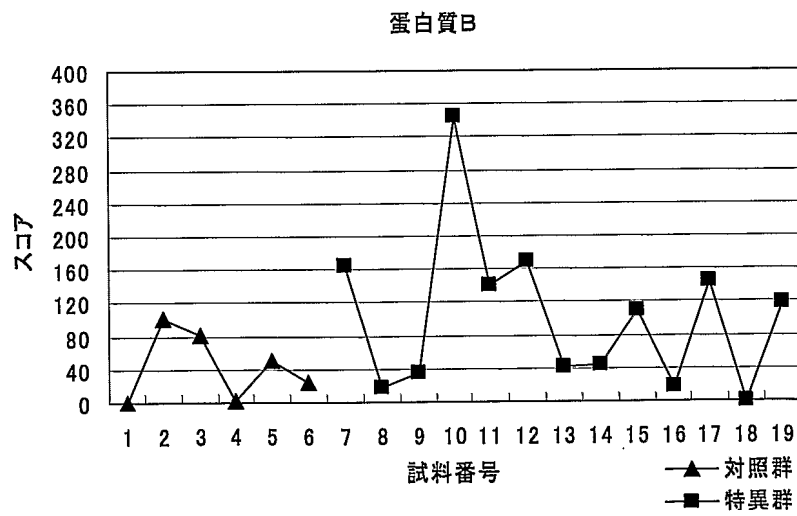
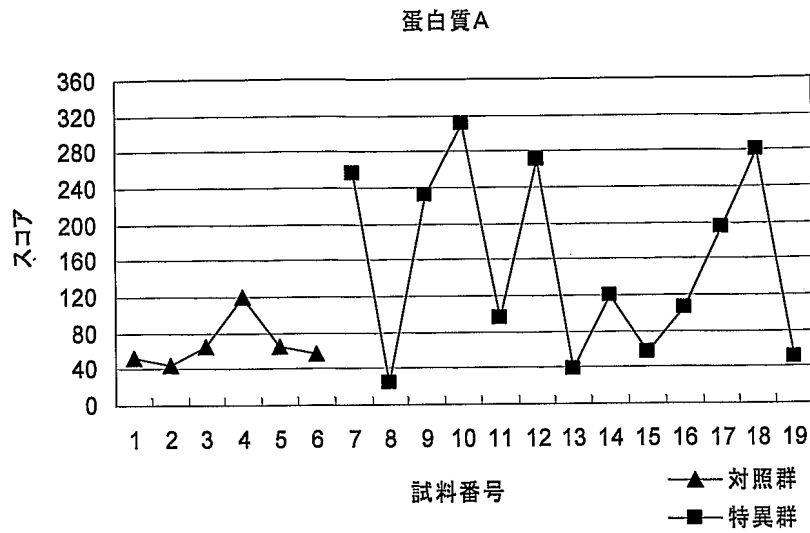
第2図



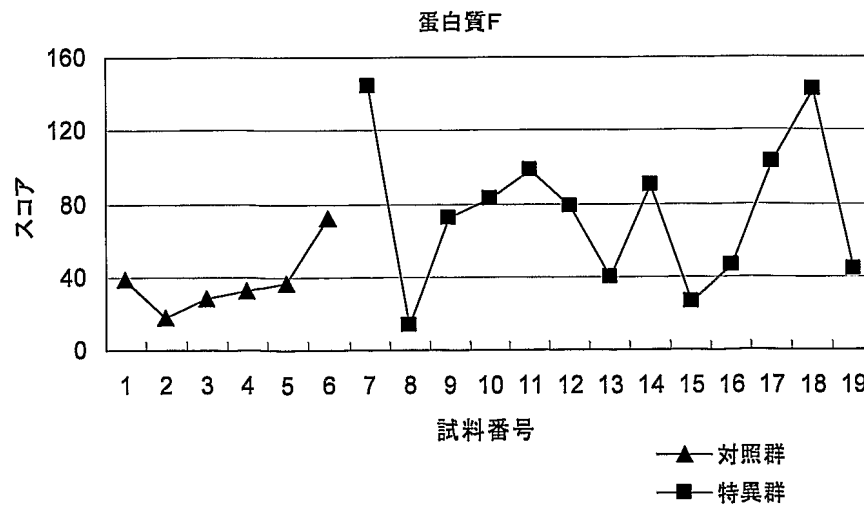
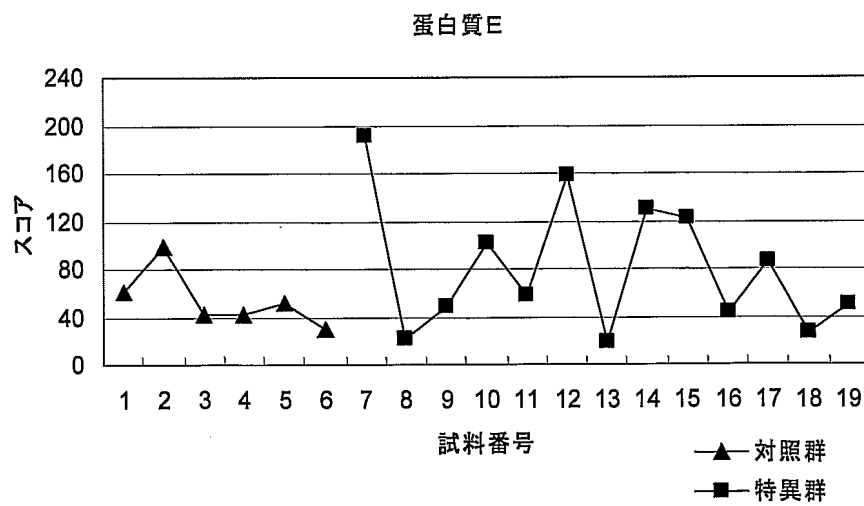
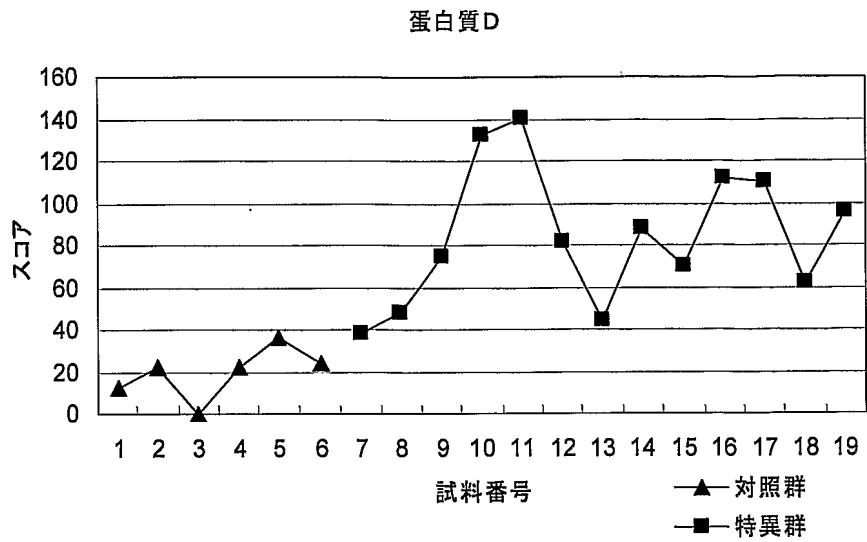
第3図



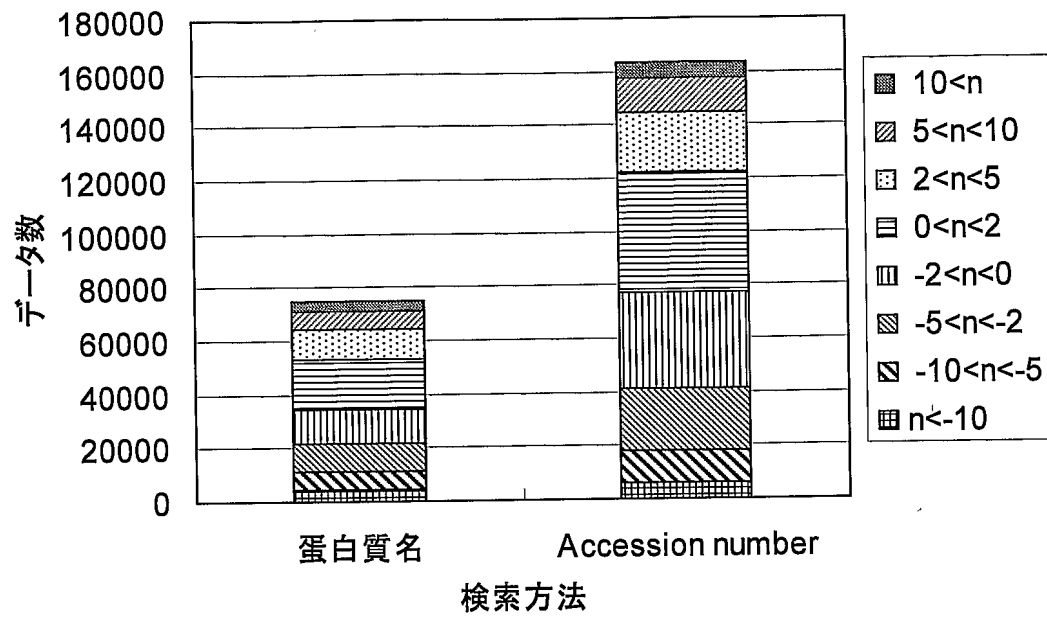
第4図



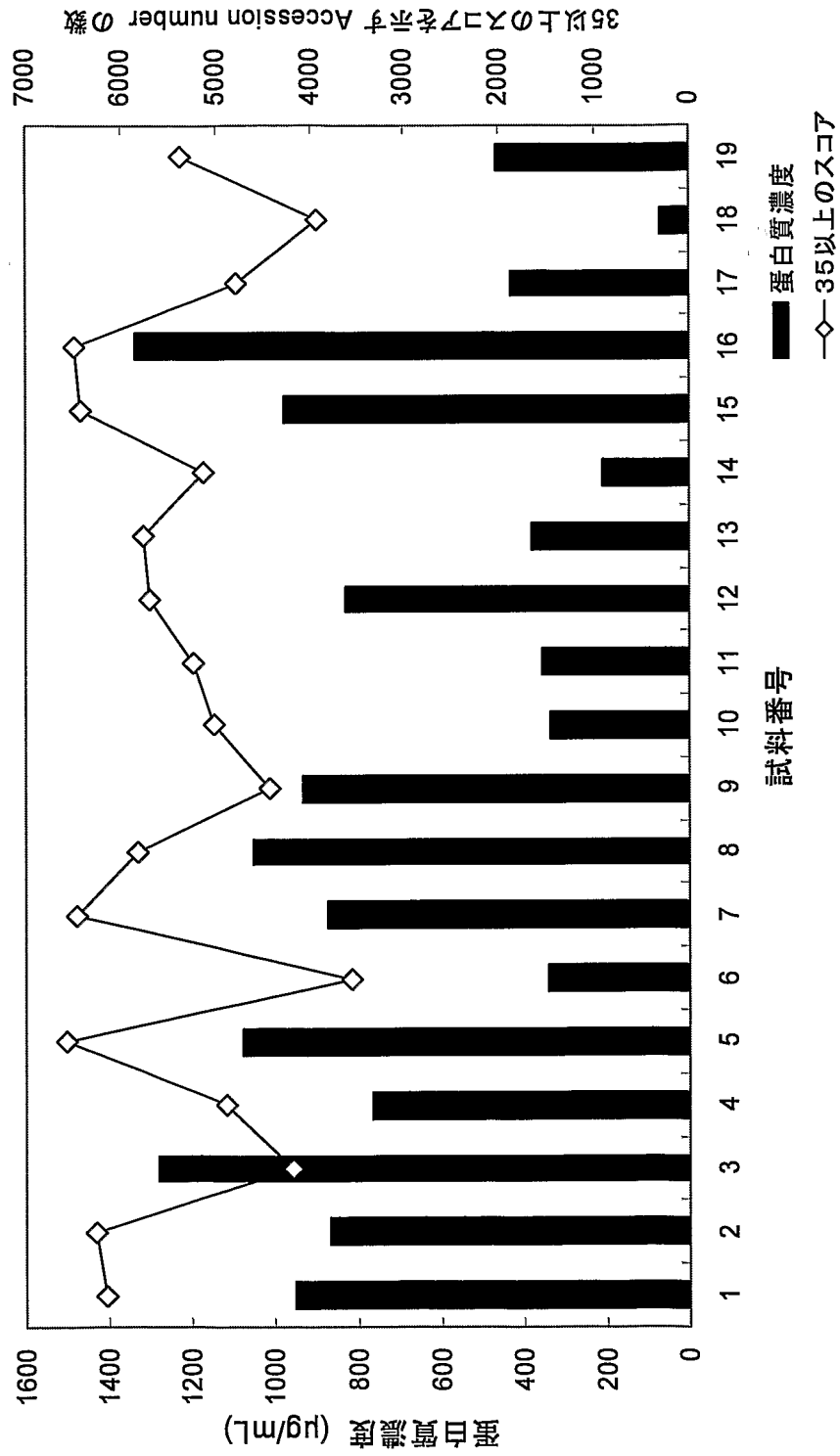
第5図



第6図



第7図



**INTERNATIONAL SEARCH REPORT**

International application No.

PCT/JP2006/302594

**A. CLASSIFICATION OF SUBJECT MATTER**

**G01N27/62** (2006.01)

According to International Patent Classification (IPC) or to both national classification and IPC

**B. FIELDS SEARCHED**

Minimum documentation searched (classification system followed by classification symbols)

G01N27/62-G01N27/70, H01J49/00-49/48

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Jitsuyo Shinan Koho	1922-1996	Jitsuyo Shinan Toroku Koho	1996-2006
Kokai Jitsuyo Shinan Koho	1971-2006	Toroku Jitsuyo Shinan Koho	1994-2006

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

PATENT FILE (PATOLIS), JSTPlus (JDream2)

**C. DOCUMENTS CONSIDERED TO BE RELEVANT**

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	JP 2003-279578 A (Director of Shikoku Cancer Center), 02 October, 2003 (02.10.03), Par. No. [0023] (Family: none)	1-5
A	JP 2004-533610 A (Perseptive Biosystems Inc.), 04 November, 2004 (04.11.04), Par. No. [0014] & WO 2002/090929 A2	1-5
A	JP 2005-506840 A (DIVERSA CORP.), 10 March, 2005 (10.03.05), Par. No. [0058] & WO 2003/029425 A2	1-5

Further documents are listed in the continuation of Box C.

See patent family annex.

\* Special categories of cited documents:

- "A" document defining the general state of the art which is not considered to be of particular relevance
- "E" earlier application or patent but published on or after the international filing date
- "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- "O" document referring to an oral disclosure, use, exhibition or other means
- "P" document published prior to the international filing date but later than the priority date claimed

- "T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
- "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
- "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
- "&" document member of the same patent family

Date of the actual completion of the international search  
07 April, 2006 (07.04.06)

Date of mailing of the international search report  
18 April, 2006 (18.04.06)

Name and mailing address of the ISA/  
Japanese Patent Office

Authorized officer

Facsimile No.

Telephone No.

**INTERNATIONAL SEARCH REPORT**

International application No.

PCT/JP2006/302594

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	WO 2004/019035 A2 (APPLERA Corp.), 04 March, 2004 (04.03.04), Page 29, lines 15 to 22 & JP 2005-536728 A	1-5
A	WO 2003/042774 A2 (CAPRION PHARMACEUTICALS INC.), 22 May, 2003 (22.05.03), Page 7, lines 5 to 11; page 16, lines 8 to 22; Page 19, lines 14 to 21 & JP 2005-536714 A	1-5

A. 発明の属する分野の分類 (国際特許分類 (IPC))  
 Int.Cl. G01N27/62(2006.01)

B. 調査を行った分野  
 調査を行った最小限資料 (国際特許分類 (IPC))  
 Int.Cl. G01N27/62-G01N27/70, H01J49/00-49/48

最小限資料以外の資料で調査を行った分野に含まれるもの  
 日本国実用新案公報 1922-1996年  
 日本国公開実用新案公報 1971-2006年  
 日本国実用新案登録公報 1996-2006年  
 日本国登録実用新案公報 1994-2006年

国際調査で使用した電子データベース (データベースの名称、調査に使用した用語)  
 特許ファイル (PATOLIS), JSTPlus (JDream2)

C. 関連すると認められる文献

引用文献の カテゴリー*	引用文献名 及び一部の箇所が関連するときは、その関連する箇所の表示	関連する 請求の範囲の番号
A	JP 2003-279578 A (国立病院四国がんセンター院長) 2003. 10. 02, 【0023】 (ファミリーなし)	1-5
A	JP 2004-533610 A (パーセプティブ バイオシステムズ インコー ポレーテッド) 2004. 11. 04, 【0014】 & WO 2002/090929 A2	1-5

C欄の続きにも文献が列挙されている。  パテントファミリーに関する別紙を参照。

* 引用文献のカテゴリー	の日の後に公表された文献
「A」特に関連のある文献ではなく、一般的技術水準を示すもの	「T」国際出願日又は優先日後に公表された文献であって出願と矛盾するものではなく、発明の原理又は理論の理解のために引用するもの
「E」国際出願日前の出願または特許であるが、国際出願日以後に公表されたもの	「X」特に関連のある文献であって、当該文献のみで発明の新規性又は進歩性がないと考えられるもの
「L」優先権主張に疑義を提起する文献又は他の文献の発行日若しくは他の特別な理由を確立するために引用する文献 (理由を付す)	「Y」特に関連のある文献であって、当該文献と他の1以上の文献との、当業者にとって自明である組合せによって進歩性がないと考えられるもの
「O」口頭による開示、使用、展示等に言及する文献	「&」同一パテントファミリー文献
「P」国際出願日前で、かつ優先権の主張の基礎となる出願	

国際調査を完了した日 07. 04. 2006	国際調査報告の発送日 18. 04. 2006
----------------------------	----------------------------

国際調査機関の名称及びあて先 日本国特許庁 (ISA/J P) 郵便番号100-8915 東京都千代田区霞が関三丁目4番3号	特許庁審査官 (権限のある職員) 横井 亜矢子 電話番号 03-3581-1101 内線 3292	2W	3311
---	---	----	------

C (続き) . 関連すると認められる文献		
引用文献の カテゴリー*	引用文献名 及び一部の箇所が関連するときは、その関連する箇所の表示	関連する 請求の範囲の番号
A	JP 2005-506840 A (ディヴァーサ コーポレイション) 2005.03.10, 【0058】 & WO 2003/029425 A2	1-5
A	WO 2004/019035 A2 (APPLERA CORPORATION) 2004.03.04, 第29頁第 15-22行 & JP 2005-536728 A	1-5
A	WO 2003/042774 A2 (CAPRION PHARMACEUTICALS INC.) 2003.05.22, 第 7頁第5-11行, 第16頁第8-22行, 第19頁第14-21行 & JP 2005-536714 A	1-5