

コラム

バイオインフォマティクスを用いた疾患 Omics による疾患マーカーの開発

近藤 格¹⁾

〔KEYWORDS〕 バイオインフォマティクス, 疾患 Omics, 疾患マーカー

1. はじめに

様々な生命現象の背景にある分子の有り様を DNA, mRNA そして蛋白質のレベルで網羅的に調べる研究が盛んに行われており, それぞれゲノミクス (genomics), トランスクリプトミクス (transcriptomics), プロテオミクス (proteomics) と呼ばれている. 本稿では疾患に関連した異常を分子レベルで網羅的に調べる研究を「疾患 Omics」としてまとめて扱い, 筆者の考える「疾患 Omics」の目ざすところ, バイオインフォマティクスの位置づけと期待される成果の臨床検査への応用の展望について述べる.

2. 「疾患 Omics」の現状

「疾患 Omics」はわが国では基礎医学の分野ではかなり普及している. ヒトゲノムプロジェクトによって網羅的なゲノム情報が利用できるようになったことに加え, 個々の実験にかかわる技術が著しく発展したことがその背景として挙げられる. 例えば DNA のコピー数の異常は従来 DNA プローブを用いてサザンブロッティングや FISH などで一つ一つ調べられていた. しかし現在ゲノミクスの分野では, DNA プローブが高密度にアレイされたスライドガラスに検体の DNA を反応させて一度に数千か所の DNA のコピー数を調べることが可能になっている. また, mRNA のコピー数の場合も同様で, 従来はノーザンブロッティングや RT-PCR で限られた数の遺伝子についてだけ調べていたのだが, トランスクリプトミク

スの分野では今では GeneChip などを用いて一度に数万種類の遺伝子の発現が調べられている. 蛋白質レベルの研究では, 今まで特異抗体を用いてウェスタンブロッティングや ELISA, あるいは比較的小さいゲルを用いた低感度の二次元電気泳動で少数の蛋白質の発現が調べられていたのが, プロテオミクスでは蛍光二次元電気泳動法, 質量分析装置, 抗体アレイなどを用いることで一度に数百から数千種類の蛋白質の発現を一度に定量的に調べることができるようになった. このような技術を利用して疾患に伴う変化が多くのレストランで調べられている.

しかし誤解のないように付け加えると, 「疾患 Omics」の本質は「調べる分子の数が多きこと」ではない. DNA, mRNA, 蛋白質の異常を広範囲に調べることは「疾患 Omics」の特徴ではあるが, それだけでは発想として従来の研究と何ら変わりはなく, わざわざ「疾患 Omics」という言葉を持ち出す必要はない. 例えば, サザンブロッティング, RT-PCR, ウェスタンブロッティングを並列に独立して 1,000 回したのでは「Omics」にはならない. それでは「疾患 Omics」の本質は何なのか? 癌の発生の過程においては DNA に生じた異常は mRNA, 蛋白質の異常へと波及し, 最終的に細胞の形質に大きな変化を与えている. その場合, 一つの遺伝子の変異は染色体の安定性, DNA 修復, 細胞周期など細胞を正常に複製しようとする様々な機能に影響を及ぼし, 結果的に多くの遺伝子・蛋白質の異常を引き起こす. 実際, 癌においては DNA, mRNA, 蛋白質それぞれのレベルで膨大な異常

1) KONDO Tadashi 国立がんセンター研究所 プロテオームバイオインフォマティクスプロジェクト

が発生しているのだが、その異常はお互いに関わりのかたちで関連していると考えられる。その関連性は網羅的に(Omics的に)異常を調べることで初めて発見されるものである。「疾患 Omics」の本質は、網羅的に分子の異常を調べ異常の全体像を包括的に理解することによって、網羅的に調べることでしかわからなかった「分子の異常の相互関係」あるいは「分子の機能的なつながり」を発見することにある。さらには「疾患 Omics」によって発見される分子の異常のパターンに基づき、疾患の新しい分類あるいは新しい疾患概念の構築も期待される。別の言い方をすると、調べる分子の種類を飛躍的に増やすことで「量から質への変換」を行うのが「疾患 Omics」の本質である。その意味ではいくら前述の最新の技術を使って広範囲に異常を調べていても、網羅的な理解を目指す視点がなければ「疾患 Omics」とは言い難い。逆に、比較的小規模な発現解析であっても「疾患 Omics」的なアプローチをとることは可能である。

「疾患 Omics」では技術的にかなりの発展が見られており、得られる成果をどのように臨床応用していくかということが次の重要な課題である。その場合よく挙げられるのが疾患マーカーの開発である。「疾患 Omics」で発見された異常な DNA, mRNA, 蛋白質の状態を、特定の病態の指標としようという試みである。「疾患 Omics」では従来は測定するのに数年かかっていたような量のデータが数日で発生する。したがって「疾患 Omics」で疾患マーカーを開発する際、膨大なデータを臨床情報と結びつけ、そこにある種の法則性を見いだしていくことが求められる。膨大な情報量の中から最も重要な変化を発見しようとする場合、経験や直感だけを頼りに研究を行うのは非科学的である。ゲノムミクス、トランスクリプトミクス、プロテオミクスの各分野では、腫瘍マーカーの開発を目指して機械学習法、多変量解析などいわゆるバイオインフォマティクスの手法を用いて分子の異常を臨床情報に関連づける研究が盛んに行われており、有望な結果が報告されている(文献 1~15)。

3. 「疾患 Omics」の成果の臨床検査への応用 首尾よく研究室レベルで疾患マーカーができた

として、それを実際に臨床検査で使用する際には克服しなければいけない問題が存在する。具体的には、臨床検査室で使えるような簡便で安定した安価な測定系が使えるかどうかである。DNA マイクロアレイ、質量分析、二次元電気泳動などは研究開発には優れたツールだが、検査センターに導入して誰でもどこでも安定した結果が出るようにすることは現状ではかなり難しい。使用される機械もまだ高額だし、実験質レベルでは1検体当たりのコストも安くない。しかしこれらの技術の臨床検査系への応用は技術的には不可能な話ではないし、生命予後や治療方針の決定を左右する診断が行えるのであれば、新たに機械を開発する価値は十分ある。また、仮に高額な検査になったとしても、そのことで治療成績が大きく向上するのであれば臨床の現場では受け入れられるだろう。研究室レベルではかなり確かな成果が得られているものもあり、実用化へ向けた試みが次の課題として挙げられている(文献 16)。

4. 文献

いずれも臨床応用を目指した網羅的発現解析の代表的あるいは最新の論文である。

mRNA の網羅的発現解析

1. Beer DG, Kardia SL, Huang CC, et al : Gene-expression profiles predict survival of patients with lung adenocarcinoma. *Nature Med* 8 : 816-824, 2002
2. van de Vijver MJ, He YD, van't Veer LJ et al : A gene-expression signature as a predictor of survival in breast cancer. *N Engl J Med* 347 : 1999-2009, 2002
3. Miller LD, Smeds J, George J, et al : An expression signature for p53 status in human breast cancer predicts mutation status, transcriptional effects, and patient survival. *Proc Natl Acad Sci USA* 102 : 13550-13555, 2005

血清ペプチドの網羅的発現解析

4. Petricoin EF, Ardekani AM, Hitt BA, et al : Use of proteomic patterns in serum to identify ovarian cancer. *Lancet* 359 : 572-577, 2002
5. Schwegler EE, Cazares L, Steel LF, et al : SELDI-TOF MS profiling of serum for detection of the progression of chronic hepatitis C to hepatocellular carcinoma. *Hepatology* 41 :

634-642, 2005

6. Poon TC, Hui AY, Chan HL, et al : Prediction of liver fibrosis and cirrhosis in chronic hepatitis B infection by serum proteomic fingerprinting : a pilot study. Clin Chem 51 : 328-335, 2005

血清蛋白質の網羅的発現解析

7. Block TM, Comunale MA, Lowman M, et al : Use of targeted glycoproteomics to identify serum glycoproteins that correlate with liver cancer in woodchucks and humans. Proc Natl Acad Sci USA 102 : 779-784, 2005
8. Wang X, Yu J, Sreekumar A, et al : Autoantibody signatures in prostate cancer. N Engl J Med 353 : 1224-1235, 2005

腫瘍組織内ペプチドの網羅的発現解析

9. Yanagisawa K, Shyr Y, Xu BJ, et al : Proteomic patterns of tumour subsets in non-small-cell lung cancer. Lancet 362 : 433-439, 2003
10. Schwartz SA, Weil RJ, Thompson RC, et al : Proteomic-based prognosis of brain tumor patients using direct-tissue matrix-assisted laser desorption ionization mass spectrometry. Cancer Res 65 : 7674-7681, 2005

腫瘍組織内蛋白質の網羅的発現解析

11. Chen G, Gharib TG, Wang H, et al : Protein profiles associated with survival in lung adenocarcinoma. Proc Natl Acad Sci USA 100 : 13537-13542, 2003
12. Iwadate Y, Sakaida T, Saegusa T, et al : Proteome-based identification of molecular markers predicting chemosensitivity to each category of anticancer agents in human gliomas. Int J Oncol 26 : 993-998, 2005
13. Jacquemier J, Ginestier C, Rougemont J, et al : Protein expression profiling identifies subclasses of breast cancer and predicts prognosis. Cancer Res 65 : 767-779, 2005
14. Chen R, Yi EC, Donohoe S, et al : Pancreatic cancer proteome : the proteins that underlie invasion, metastasis, and immunologic escape. Gastroenterology 129 : 1187-1197, 2005
15. Uhlen M, Bjorling E, Agaton C, et al : A human protein atlas for normal and cancer tissues based on antibody proteomics. Mol Cell Proteomics 2005

疾患 Omics の臨床検査への応用の現状

16. Branca MA : 'Omic diagnostics trip up on way to clinic. Nat Biotechnol 23 : 769, 2005

MEDICAL BOOK INFORMATION

医学書院

学生のためのカレントメディカルイングリッシュ 第3版

飯田恭子

●A5 頁224 2005年
定価2,520円(本体2,400円+税5%)
[ISBN4-260-00117-5]

著者のユニークな解説によって、読者は医療にも英語にも興味を抱くはず。英文読解問題をはじめ、単語をはめこむ問題や図表を使った問題など、練習問題も豊富。自習学習者の利便性を考慮して練習問題の解答を巻末に収載。さらに教師用資料の内容を充実させ、教科書としての使い勝手も向上した。

THE LUNG

perspectives

VOL.13 NO.4 2005 別刷

株式会社 **メディカルレビュー社**

臨床応用を目指した肺がんの プロテオミクス

Lung cancer proteomics for clinical utility

国立がんセンター研究所プロテオーム・バイオインフォマティクス・
プロジェクト プロジェクトリーダー

国立がんセンター研究所所長

近藤 格 *Tadashi Kondo*

廣橋 説雄 *Setsuo Hirohashi*

Key words

オーダーメイド医療, 診断マーカー, 二次元電気泳動, バイオインフォマティクス, レーザーマイクロダイセクション

Summary

発現しているタンパク質の全体像を理解しようとする試み(プロテオミクス)が盛んに行われている。現在のプロテオミクスの背景には、タンパク質の全体像を捉えることの重要性に対する認識と、タンパク質・ペプチドの質量を精度よく測定して同定する技術がある。細胞の形質を直接コントロールするタンパク質を調べる試みは、がん研究において有効である。プロテオミクスは肺がん

の研究においてもメカニズムの解明から診断マーカーの開発まで幅広く応用されている。国立がんセンターでも抗がん剤の奏効性や転移に関わるタンパク質をプロテオミクスの手法で同定している。タンパク質の全体像を捉えるという新しい思想に基づくアプローチで、肺がんの研究はいつそう発展するであろう。

I プロテオミクスとは

プロテオーム (proteome) とは、発現しているタンパク質の全体像を表す概念であり、プロテオームを研究対象とする学問がプロテオミクス (proteomics) である。タンパク質の発現量、翻訳後修飾、細胞内局在、分解、そしてそれらの発生、老化、疾患などに伴う変化など、タンパク質に関わる事象をすべて網羅的に調べ、網羅的に調べ

ることでしかなかった知見を得ることがプロテオミクスの目標である。

2つの要因からプロテオミクスはここ数年注目を浴びるようになった。1つは、ヒトゲノムプロジェクトが一段落したという時代背景である。「全体をみることが重要である」という新しい概念から出発したビッグサイエンスが成功を収め、ゲノムの情報を利用して mRNA そしてタンパク質について

網羅的な解析を行うことが次の目標として設定されるようになった。mRNA の発現を網羅的に調べるトランスクリプトーム解析の分野ではマイクロアレイが普及し、サンプル間で発現差のある遺伝子の探索については、一昔前には想像もしなかったほど多くのデータが一度の実験で得られるようになった。今まで何年もかけて発現差のある遺伝子を探索していたことを考えると、驚異的な進歩である。当初は

特定の遺伝子を見つけることだけが目的であったが、ここ数年は発現レベルの全体像に基づくがんの新しい分類も試みられており、臨床医や病理医が経験的に知っていた事象の分子生物学的裏付けがとられつつある。ゲノム情報に基づきRNAレベルで全体を俯瞰することの重要性が提示されたので、同じことをタンパク質レベルで行おうという気運が高まってきたのは当然の流れである。さらに、後述するように、タンパク質レベルでの研究でしか得ることができない重要な現象ががん研究において存在することが従来から知られていたため、なおさらタンパク質の網羅的解析の可能性が注目されるようになった。

2つ目は、タンパク質の網羅的解析を可能にする技術革新があったことである。タンパク質研究のボトルネックは、実験に必要とされるタンパク質の量にあった。しかし、現在では発現解析の実験やタンパク質の同定は、従来の1,000倍の感度で行うことが可能になっている。鍵となったのは質量分析装置とゲノムデータベースである。現在では、①質量分析装置によってタンパク質・ペプチドおよびその分解産物の質量をきわめて正確に測定し、②測定された質量はどのようなアミノ酸配列に由来するかということゲノムデータベースで検索する、という方法が普及している。このような時代の流れと技術革新が現代のプロテオミクスの背景にある。

II がん研究とプロテオミクス

発がんやがんの進展の主要な原因は

ゲノムの異常の蓄積である。しかし、ゲノムの異常の最終的なアウトプットは、分子レベルでは多くの場合タンパク質の異常であって、タンパク質を直接調べなくてはわからないことが多い。たとえば、mRNAの発現解析を行う場合、mRNAの発現が高ければその翻訳産物であるタンパク質も多く存在しているだろうという考えが背景にある。確かに、mRNAとタンパク質の発現に相関がある遺伝子も多い。しかし、最近の網羅的な発現解析から、mRNAとタンパク質の発現の間には全体としては相関が乏しく、mRNAの発現量の測定でもってタンパク質の発現測定を代用することはできないことが示唆されている¹⁾²⁾。

タンパク質について、DNAやmRNAからわからないことは発現量だけではない。たとえば、タンパク質はmRNAから翻訳されたままの形では機能しないことが多い。タンパク質は翻訳後に適切な部位(膜、核、細胞質)に運ばれたり、リン酸化、糖鎖修飾、切断などを受けたりして初めて本来の働きをしている。しかしながら、タンパク質がどこに局在しどのように翻訳後修飾を受けているのかはDNAの配列やmRNAの量からは今のところほとんど予測することができない。また、タンパク質同士の相互作用やタンパク質の寿命については、配列情報からそれらを予測することは困難である。一方で、腫瘍細胞の形質を直接制御するのはいうまでもなくDNAやmRNAではなくタンパク質である。したがって、タンパク質を直接調べることで「がんの個性」を裏付ける分子機構を新しく提唱できる可能性が高い

と考えられる。どのようにすればタンパク質の様子を詳細かつ網羅的に調べて臨床情報や生物学実験の結果と結び付けることができるか、というのがこのアイデアを具体化する上での問題点である。

III 肺がんのプロテオミクスの現状

数ある悪性腫瘍の中で、肺がんは最もプロテオミクスが応用されている腫瘍であり、論文も多い。本稿では肺がんのプロテオミクスの分野での最近のトピックスを2つ紹介する。

1. オーダーメイド医療のための腫瘍マーカー開発

トピックスの1つ目は、個別化医療のための腫瘍マーカー開発である。遺伝子の発現をmRNAのレベルで網羅的に調べて、がん患者の予後や抗がん剤奏効性を推定しようとする試みが盛んに行われている。種々のデータマイニングの手法を用いることで、複数のmRNAの「発現パターン」を腫瘍マーカーとして使用しようとする試みである。タンパク質レベルで同様のことをしようとする試みを紹介する。

Hanashらのグループ(現・Fred Hutchinson Cancer Research Center)は二次元電気泳動法を用いて肺がんのプロテオーム解析を行っている。二次元電気泳動法は古くからあるタンパク質分離技術であり、タンパク質を定量性よく翻訳後修飾まで含めて解析できるというメリットがある³⁾⁴⁾。この方法は1975年に発表され、盛んに用いられていた時期もあったと聞くが、

1990年代にはすっかり廃れかけていた。その理由の1つは、二次元電気泳動で観察されるタンパク質が大変微量であり、その構造解析がきわめて困難であったためである。しかし、上述の質量分析装置とゲノムデータベースとによりその問題はほぼ完全に解決され、今ではゲル上で観察されるタンパク質は100%近く同定が可能になった。Hanashらは、肺がんの手術検体のタンパク質発現プロファイルを二次元電気泳動法で作成し、それらをバイオインフォマティクス手法で解析することにより、肺がん患者の予後の予測に最も役立つタンパク質を20種類同定した⁵⁾。さらに、そのうちPGK-1に関しては、予後の悪い症例において血中発現レベルも増加していることを報告している。

2. 診断や治療のためのプローブ開発

2つ目のトピックは、画像診断や治療のためのプローブ開発である。がん細胞にターゲットを絞った分子標的治療を行ったり、微小ながんを画像で検出したりするためには、がん細胞に特有の分子を見つけ、その分子に特異的に結合するプローブを作製することが重要である。プロテオミクスからのアプローチとしては、狙った細胞のみ特異的に発現するタンパク質を網羅的に同定して、それぞれに対応するプローブとして抗体などを作製することになる。腫瘍細胞特異的なタンパク質であっても細胞内タンパク質の場合は、それらに対する抗体を作製しても細胞内にまで到達しないことが十分考えられる。したがって、分子標的としてはプローブがアクセスしやすいタン

パク質が好ましいであろう。それでは、どのようにしてそのようなタンパク質をプロテオミクスの手法で同定できるのだろうか。

Ohらは、肺がん組織に含まれる血管の内皮細胞の膜タンパク質に注目した⁶⁾。血管内皮細胞は組織ごとに性格が異なるだけでなく、同じ臓器でも正常組織と腫瘍組織では含まれる血管内皮細胞の性格が異なる。彼らは、特定の組織の血管内皮細胞の細胞膜だけを回収する技術を開発し、回収された細胞膜のプロテオミクスを行った。このように特定の臓器や細胞内小器官を対象を絞ったプロテオーム解析は、「フォーカスド・プロテオミクス (focused proteomics)」と呼ばれている。発現しているタンパク質全部を一度の解析でカバーすることは、プロテオミクスのどのような技術をもってしても不可能である。研究対象を絞り、プロテオームの特定の分画だけを回収してから解析を行うことで、サンプルの複雑度を下げ、結果的に研究対象の事象に関わるタンパク質を多く観察できるようになる。彼らは正常肺組織と肺がん組織の血管の内皮細胞の膜タンパク質を比較し、後者に特有のタンパク質がいくつか存在することを発見した。同定されたタンパク質であるannexin Iに対する抗体を担がんマウスに静注してみると、腫瘍は縮小し生存時間は著しく改善された。本方法は、腫瘍細胞にターゲットを絞った治療法の開発や、通常の画像診断では発見が困難な微小な腫瘍の早期診断や転移の発見に役立つであろう。

IV

国立がんセンターにおける取り組み

1. がんのプロテオーム解析を支える技術基盤

国立がんセンター研究所プロテオーム・バイオインフォマティクス・プロジェクトでは、多くの臨床検体を用いた大規模な肺がんのプロテオーム解析を行っている。技術基盤になっているのは蛍光二次元電気泳動(2D-DIGE)法である。2D-DIGE法ではタンパク質を蛍光色素で標識してから泳動するが、実験サンプルとは異なる蛍光色素で標識した内部標準を泳動することで電気泳動によるばらつきを解消することが可能になっている。筆者らは、そこにさらにさまざまな工夫を凝らすことで、他では例をみないプロテオーム解析を行っている。その一端を紹介する。

質の高いプロテオーム解析への第一歩は、正確なサンプリングである。腫瘍組織の中には腫瘍細胞以外に実にさまざまな細胞が含まれている。プロテオーム解析をどんなに正確に行うことができたとしても、腫瘍組織をまとめてすり潰していたのでは何のプロテオームをみているのかわからなくなってしまう。腫瘍組織に含まれる間質もがんの生物学においては重要であり、ゆえに腫瘍組織をまるごと扱うべきだとの意見もある。腫瘍細胞の周囲の間質ががんの進展に重要な役割を果たしていることは確かであり、間質のプロテオミクスは大変興味深いテーマである。しかし、それならば間質細胞のみレーザーマイクロダイセクションで回収して比較するべきである。筆者は、

レーザーマイクロダイセクションで得られる腫瘍細胞のプロテオーム解析を、超高感度の蛍光色素を使って行っている⁷⁾。従来はレーザーマイクロダイセクションには膨大な時間がかかっていたが、通常の染色法の100倍もの感度をもつ蛍光色素を応用することで、レーザーマイクロダイセクションにかかる時間は大幅に短縮された。筆者のラボでは、手術検体を使う場合は、保存状態の許す限りレーザーマイクロダイセクションで凍結切片から細胞を回収している。肺がんに関しては、組織型⁸⁾や転移能に関係するタンパク質の同定を行っている。この方法は最近では普及し、胃がん⁹⁾や膵がん¹⁰⁾のプロテオミクスにも使用されている。

次に、二次元電気泳動の網羅性を高めるために、通常よりかなり大きなゲルを用いて二次元電気泳動を行っている。市販の泳動装置ではEttan Dalt II (アマシャムバイオサイエンス株式会社)が最大であるが、Ettan Dalt IIでは約2,000個以下のスポットしか観察できない。しかしゲルのサイズを大きくして解像度を上げると8,000個ものスポットを観察できることが報告されている¹¹⁾。大きなゲルで網羅性を高めるという方法の限界は、染色できるゲルのサイズが限られていることである。あまりに大きなゲルは、ゲルの強度の問題で染色が困難である。2D-DIGE法においては、タンパク質サンプルは蛍光色素で標識されてから二次元電気泳動で分離され、ゲルをレーザーキャナーでスキャンして画像を取得する。そのときゲルはガラス板に挟まったままの状態ですべて乗せられるので、ゲルの強

度は全く問題にならない。2D-DIGE法では、ゲルの強度についてだけいえば、理論的にはゲルはいくらでも大きくできる。この方法の限界はスキャナーの面積である。筆者のラボでは、市販されている中では最大サイズのスキャナーに合わせた大きさのゲルが12枚同時に泳動できる泳動装置を作成し、泳動を行っている。

プロテオームのデータ解析にはDNAマイクロアレイの解析と同様のバイオインフォマティクスの手法を用いている。臨床検体を用いたプロテオミクスでは、二次元電気泳動から得られる膨大なデータを臨床情報と結び付け重要な事象(転移、予後、抗がん剤耐性など)の背後にあるタンパク質のあり方を調べることが要求される。二次元電気泳動を用いたプロテオミクスといっても、スポットを1つ1つ見比べていたのではウエスタンブロットを違う抗体で何100回もしたのと同じことであり、その発想は30年前に二次元電気泳動が初めて世に出た頃と比べて本質的に進歩がない。タンパク質同士の発現の相互関係を調べ、発現の背後にある関係性をみつけることが全体像を理解することにつながるのであり、また近年の二次元電気泳動ではそのような解析を可能にするほどに性能が高くなっている。筆者のラボではDNAマイクロアレイ用に開発されたソフトや一般的なデータマイニングのソフトを使用して、二次元電気泳動から得られるデータを臨床情報と関連させて解析を行っている。発現レベルだけでなく翻訳後修飾や細胞内局在の情報などをどれだけ活用できるかが今後の課題であると考えている。

最後に、質量分析装置によるタンパク質の同定の工夫である。前述したように、二次元電気泳動の最大の弱点はスポットに対応するタンパク質の同定が著しく困難ということであった。昔は対応するタンパク質が何かかわからないので、スポットそのものを腫瘍マーカーとして使用するという論文も数多くあった。しかし、そのような考えは電気泳動を専門とする狭いコミュニティの中でしか通用せず、やがて二次元電気泳動は廃れていったわけである。今ではタンパク質の抽出効率を高めると同時に、よく整備された高性能の質量分析装置を使用することで、ほぼ100%の成功率でタンパク質の同定が可能である。

2. 肺がんの治療成績の向上を目指した取り組み

筆者のプロジェクトでは、肺がんのプロテオミクスとして、ゲフィチニブの奏効性に関連するタンパク質の発現を調べている。奏効性がある患者群をあらかじめ識別することで治療成績を向上させることができるし、同定されたタンパク質を調べることで耐性のメカニズムを調べたり、創薬のきっかけになるのではないかと考えている。ゲフィチニブの奏効性とEGFR(epidermal growth factor receptor)の変異が関係するとの報告がいくつも発表されているが、奏効性・耐性・副作用のメカニズムを理解するためにはさらなる研究が必要である。その他にも、肺腺がんの病理分類を裏付けるタンパク質発現パターンや、リンパ節転移に関連するタンパク質発現パターンなどを調べている。筆者のプロジェクトで

は、同じようなプロテオミクスの試みは他の悪性腫瘍(肝がん, 骨軟部腫瘍)でも行われており, それらでは既により結果が得られていることから, 肺がんでも成果を期待している。

V これからの展望

プロテオミクスは, タンパク質の全体像を研究対象とする今までになかった新しい学問である。ここ数年のプロテオミクス関連技術の発展は目覚ましく, 新しい技術をがん研究に応用することで今まで予想していなかった現象を発見できる時代にわれわれはいる。一方で, プロテオミクスの技術の応用の仕方は多くの場合は依然として従来の学問の延長線上にあることから, これからはプロテオミクスならではの発想に基づく研究がブレークスルーのきっかけになるのではないかと考えている。

文 献

- 1) Gygi SP, Rochon Y, Franz BR, et al : Correlation between protein and mRNA abundance in yeast. *Mol Cell Biol* 19 : 1720-1730, 1999
- 2) Chen G, Gharib TG, Huang CC, et al :

- Discordant protein and mRNA expression in lung adenocarcinomas. *Mol Cell Proteomics* 1 : 304-313, 2002
- 3) O'Farrell PH : High resolution two-dimensional electrophoresis of proteins. *J Biol Chem* 250 : 4007-4021, 1975
 - 4) Klose J : Protein mapping by combined isoelectric focusing and electrophoresis of mouse tissues. A novel approach to testing for induced point mutations in mammals. *Humangenetik* 26 : 231-243, 1975
 - 5) Chen G, Gharib TG, Wang H, et al : Protein profiles associated with survival in lung adenocarcinoma. *Proc Natl Acad Sci U S A* 100 : 13537-13542, 2003
 - 6) Oh P, Li Y, Yu J, et al : Subtractive proteomic mapping of the endothelial surface in lung and solid tumours for tissue-specific therapy. *Nature* 429 : 629-635, 2004
 - 7) Kondo T, Seike M, Mori Y, et al : Application of sensitive fluorescent dyes in linkage of laser microdissection and two-dimensional gel electrophoresis as a cancer proteomic study tool. *Proteomics* 3 : 1758-1766, 2003
 - 8) Seike M, Kondo T, Fujii K, et al : Proteomic signatures for histologi-

cal types of lung cancer. *Proteomics* 5 : 2939-2948, 2005

- 9) Greengauz-Roberts O, Stoppler H, Nomura S, et al : Saturation labeling with cysteine-reactive cyanine fluorescent dyes provides increased sensitivity for protein expression profiling of laser-microdissected clinical specimens. *Proteomics* 5 : 1746-1757, 2005
- 10) Sitek B, Luttes J, Marcus K, et al : Application of fluorescence difference gel electrophoresis saturation labelling for the analysis of microdissected precursor lesions of pancreatic ductal adenocarcinoma. *Proteomics* 5 : 2665-2679, 2005
- 11) Klose J, Nock C, Herrmann M, et al : Genetic analysis of the mouse brain proteome. *Nat Genet* 30 : 385-393, 2002

近藤 格

平成 8 年 岡山大学医学部医学研究科

(博士課程)修了(細胞生物学)

現在, 国立がんセンター研究所プロテオーム・バイオインフォマティクス・プロジェクト プロジェクトリーダー

専門分野: がんのプロテオミクス

E-mail : takondo@gan2.res.ncc.go.jp

4. ペプチドーム解析の現状と展望

佐々木 一樹, 磯山正治, 南野直人

生体に内在するペプチドは前駆体タンパク質から固有の切断や修飾を受けて生成する。プロテオーム解析でこのようなペプチドも分析可能であると考えられがちであるが、ペプチドに特化した取扱いや操作を行わないとその実像はみえてこない。内在性ペプチドを対象とするペプチドーム解析は、生理活性ペプチドや疾患マーカーの探索、タンパク質の分解過程の解明に直結するもので、独自に実施しなければならない課題である。ペプチドーム解析への取り組みで少しずつ明らかになりつつある生体内ペプチドの実像と今後の展望について概説する。

はじめに

ヒトゲノム配列のうち、28.5億塩基対が99.999%の精度で決定され、遺伝子数は予想を下回る21,787程度と推定された。実在するタンパク質の総数はもとより、タンパク質から特異的切断・修飾を受けて生み出されるペプチドの総数は推測の域を超えている。その中で生理活性ペプチドの数は限定的と考えられる。哺乳類の生理活性ペプチドの3割以上がわが国の研究者の発見になることはほとんど知られていないが、生体に実在するペプチドの全貌、すなわちペプチドームの解析法確立とデータベース化はこのペプチド研究を大きく推進すると期待される。1999年度よりブタ脳ペプチドを対象にはじめたわれわれのペプチドーム解析は、マウスおよびヒトに枠を広げつつある。これまでに蓄積された情報から少しずつみえはじめた生体内ペプチドの実像と今後の可能性について概説する。

【キーワード】

ペプチドーム解析, ファクトデータベース, 内在性ペプチド, 質量分析, 二次元クロマトグラフィー

■ ペプチドとペプチドーム解析

ペプチドとタンパク質の両者を明確に区別する定義はないが、一般的にはアミノ酸が2個以上つながった分子量1万以下の物質をペプチドとよぶことが多い。ホルモン、循環調節因子、神経伝達・調節物質などとして作用する物質の多くはペプチドである。これらの生理活性ペプチドは、前駆体タンパク質へと翻訳された後にプロセッシングや修飾を受けてはじめて機能を発揮する。例えば、甲状腺ホルモン放出ホルモン (TRH: thyrotropin-releasing hormone) はわずか3つのアミノ酸で構成され、分子の両端に翻訳後修飾を受けている。このほかに、タンパク質の限定分解や代謝的分解の結果、生理活性ペプチドを圧倒する多種多量のペプチドが細胞の内外に存在する。限定分解で生じるペプチドとしてはフィブリノペプチド、プロテアーゼ活性化時のプロペプチドなどがよく知られている。生体内に存在するペプチドの全貌、すなわちペプチドームの実態はこれまで不明であった。実験結果や測定値などを構造化して収納し、効率的な検索を可能としたデータベースはファクトデータベースとよばれ、

Peptidomics—status quo and perspectives

Kazuki Sasaki¹⁾/Masaharu Isoyama²⁾/Naoto Minamino¹⁾: Department of Pharmacology, National Cardiovascular Center Research Institute¹⁾/Information Department, Protein Research Foundation²⁾ (国立循環器病センター研究所薬理部¹⁾/蛋白質研究奨励会情報管理室²⁾)

高精度なデータを入手できればその価値は将来も減少せず、情報基盤としての波及効果は非常に大きい。プロテオーム解析では、発現解析から相互作用、立体構造までを含めた研究が推進されているが、ペプチドーム解析では「実在するペプチド分子型のありさま」を記述し、生理活性ペプチドの発見の効率化や、タンパク質の代謝分解過程の情報収集が当面の課題である。

2 ペプチドーム解析は可能か？

ペプチドはタンパク質と同様に取扱・分析できると考えられやすいが、現在のプロテオーム解析法はペプチドーム解析には適していない。第1に、ペプチドはプロテオーム解析で繁用されるゲル電気泳動では分離が困難であること、第2に、多くの生理活性ペプチドは分子内に塩基性アミノ酸を含むため、通常のプロテオーム解析で実施される酵素消化によって同定が不可能になること、第3に、生理活性ペプチドの組織濃度や含量がタンパク質と比較して圧倒的に少ない（重量比でペプチドはタンパク質の0.1%以下）ことがあげられる。また、分析上の問題点として、試料の抽出操作時にタンパク質分解が起きる結果、多数の分解ペプチドが生じ、内在性のペプチドと判別不能になるからである。このため、ペプチドーム解析は非現実的とみなされ、限定的な試みしか行われていなかった¹⁾。われわれはこれらの問題点を克服しつつ、生体内に存在するペプチド情報を網羅的に収集するファクトデータベースの構築を1999年度より開始した^{2)~4)}。近年発見されたグレリンにおけるセリン残基のオクタノイル化修飾のように、生理活性ペプチドの活性発現には翻訳後修飾が必須である場合が多い⁵⁾。このような翻訳後修飾はゲノム構造からは決して予測しえず、生体内に存在するペプチドの構造情報を包括的に獲得するアプローチの価値は減弱しない。哺乳類では分解ペプチドが多量に存在するため、ペプチドーム解析による新規生理活性ペプチドの発見は困難が予想されたが、分析法の改良に伴って具体的成果が得られはじめている。

3 ペプチドームのプロファイリング法

われわれはタンパク質分解酵素を熱処理で失活させた後にペプチドの抽出を行うことにより、多くの活性ペプチドを内在する形で精製してきた^{6) 7)}。抽出操作の後、逆相および陽イオン交換樹脂を利用してタンパ

ク質、塩、糖、核酸、アミン類を効率的に除去し、ペプチド画分を短時間で回収している。この「粗ペプチド画分」に残存するタンパク質をゲル濾過で除去し、得られる分子量6,000以下の画分を現在の解析対象としている。

ペプチドのプロファイリングには、一次元目に陽イオン交換HPLC（70画分）、二次元目に逆相HPLC（75画分）を用いる二次元クロマトを行い、全体を約5,000画分に分割している（図1）。各画分の一部をマトリックス支援レーザー脱離イオン化飛行時間型（MALDI-TOF）質量分析計で分析する。ペプチドの構造解析は、エレクトロスプレーイオン化四重極飛行時間型（ESI-Q-TOF）^{*1}とMALDI-TOF-TOF^{*2, 3}のタンデム質量分析計で行う。ゲノム情報が蓄積されている生物種では、親イオンの質量と構造情報取得に十分なフラグメントイオンが得られれば同定はさほど困難ではない。しかし、糖ペプチドのように修飾構造を含む場合は構造決定のために分析試料を増やし、追

※1 Q-TOF (Quadrupole time-of-flight mass spectrometer : 四重極飛行時間型質量分析計)

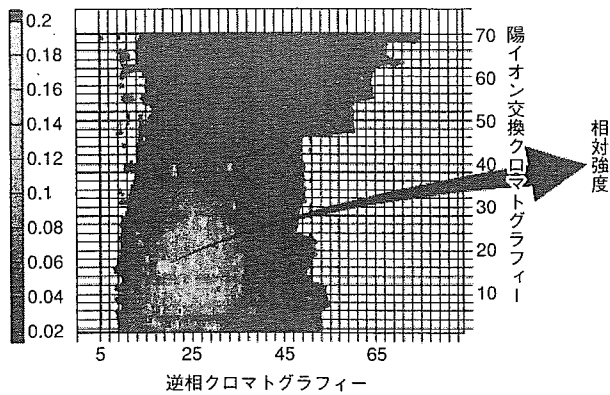
分析部に四重極と飛行管を垂直に連結した質量分析計。イオン化した試料を四重極質量分析計で分離した後、ペプチド鎖を開裂させ、飛行管型質量分析計で詳細な構造解析が実施できるタンデム質量分析装置である。後述のMALDI-TOF-TOFとならびプロテオミクス研究で用いられる代表的な質量分析計である。

※2 MALDI-TOF (Matrix-assisted laser desorption ionization time-of-flight mass spectrometer : マトリックス支援レーザー脱離イオン化飛行時間型質量分析計)

イオン源にはマトリックス支援レーザー脱離イオン化法を採用し、分析部に飛行時間型質量分析計を搭載した装置。本イオン化法は飛行時間型質量分析計との適合性がよく感度が高いため初期から実用化され、操作も比較的容易であることから質量測定に繁用される。反射電場を備えた装置では構造解析も可能であるが、詳細な構造解析は困難である。

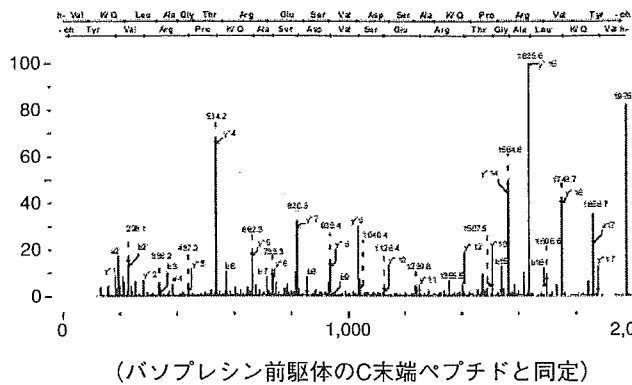
※3 MALDI-TOF-TOF (Matrix-assisted laser desorption ionization time-of-flight/time-of-flight mass spectrometer : マトリックス支援レーザー脱離イオン化二重連結型飛行時間型質量分析計)

イオン化にはMALDI法（マトリックス支援レーザー脱離イオン化法）を用い、分析部には飛行時間型質量分析計をタンデムに2つ連結した装置。1つの質量分析計で分離後ペプチド鎖を開裂させ、2段階目の質量分析計で構造を詳細に解析する。MALDI法による従来の構造情報解析法では困難であった分解能と質量精度の達成が可能になり、ハイスループットな構造解析に適している。

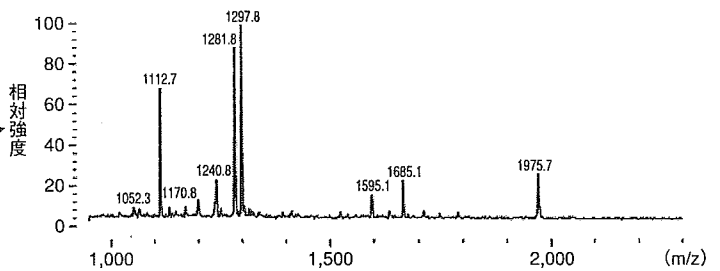


逆相クロマトグラフィー

タンデム質量分析 (ESI-Q-TOF)



マウス脳ペプチドの1画分の分析結果



ペプチドピークリストの作成

F19	m/z	強度
	1052.28	1,826
	1112.66	13,000
	1170.76	1,927
	1197.69	2,277
	1200.73	2,577
	1240.78	4,400
	1281.78	17,000
	1297.77	19,000
	1312.78	1,695
	1319.74	1,572
	1339.79	1,343
	1522.87	1,372
	1595.08	3,011
	1633.99	1,467
	1665.08	4,392
	1712.07	1,452
	1748.05	1,215
	1789.07	1,328
	1975.68	4,876

図1 ペプチドームのプロファイリング方法

二次元HPLCにより約5,000に分けられた画分(左上)は、質量分析によってペプチドの数と分子量が測定され(右)、次にタンデム質量分析により構造が決定され(左下)、データベースに登録される(文献18より引用改変)

加の生化学的操作が必要である。ただし、いったんデータベースが構築されると、物性と分子量の情報から標的ペプチドの大部分は容易に同定できる。

この方法は二次元HPLCを用いるため、溶媒、カラム、流速などについて条件を決め、機器や溶媒ごとに得られたデータの補正、標準化が必要である。われわれはイオン交換HPLC、逆相HPLCについて標準ペプチドが溶出される塩濃度、アセトニトリル濃度を基準として補正、標準化を行っている。分析条件が標準化されると、基本的には研究室間でもデータの比較検討が可能となる。質量分析法による構造解析アルゴリズムの改良、実験情報管理システム(LIMS: laboratory information management system)の開発も共同研究によって実施している^{8) 9)}。

4 ペプチドーム解析の現状と可能性

われわれはブタおよびマウス脳組織についてペプチドーム解析を開始したが、これはこれまでに多数の生

理活性ペプチドをブタ脳組織より発見したため、またマウスについてはゲノム情報が蓄積されていたためである³⁾。二次元HPLCは分子量6,000までのペプチド分離に適するよう条件を設定しており、既知の生理活性ペプチドは全体に広く溶出する(図2)。図1は、マウス脳の分子量3,000以下のペプチドを示したもので、二次元HPLC上の電荷や疎水性の低い部分に多くのペプチドが溶出している。分離した5,000画分の1つのMALDI-TOF質量分析計による分析結果を図1右上に示すが、ここで観測されるピークについて、そのm/z値と強度(イオンカウント)を整理するとピークリストが作成できる。順次、タンデム質量分析を行い、データベース検索より構造を推定する。このような結果を5,000画分についてまとめ、分子量、疎水性、電荷を座標軸とする仮想的な三次元空間に表示すると、図3のようにペプチドを視覚的に把握できる。各点はペプチドを、大きさは質量分析でのイオンカウントを示し、後者はペプチド量と大まかに相関する。こ

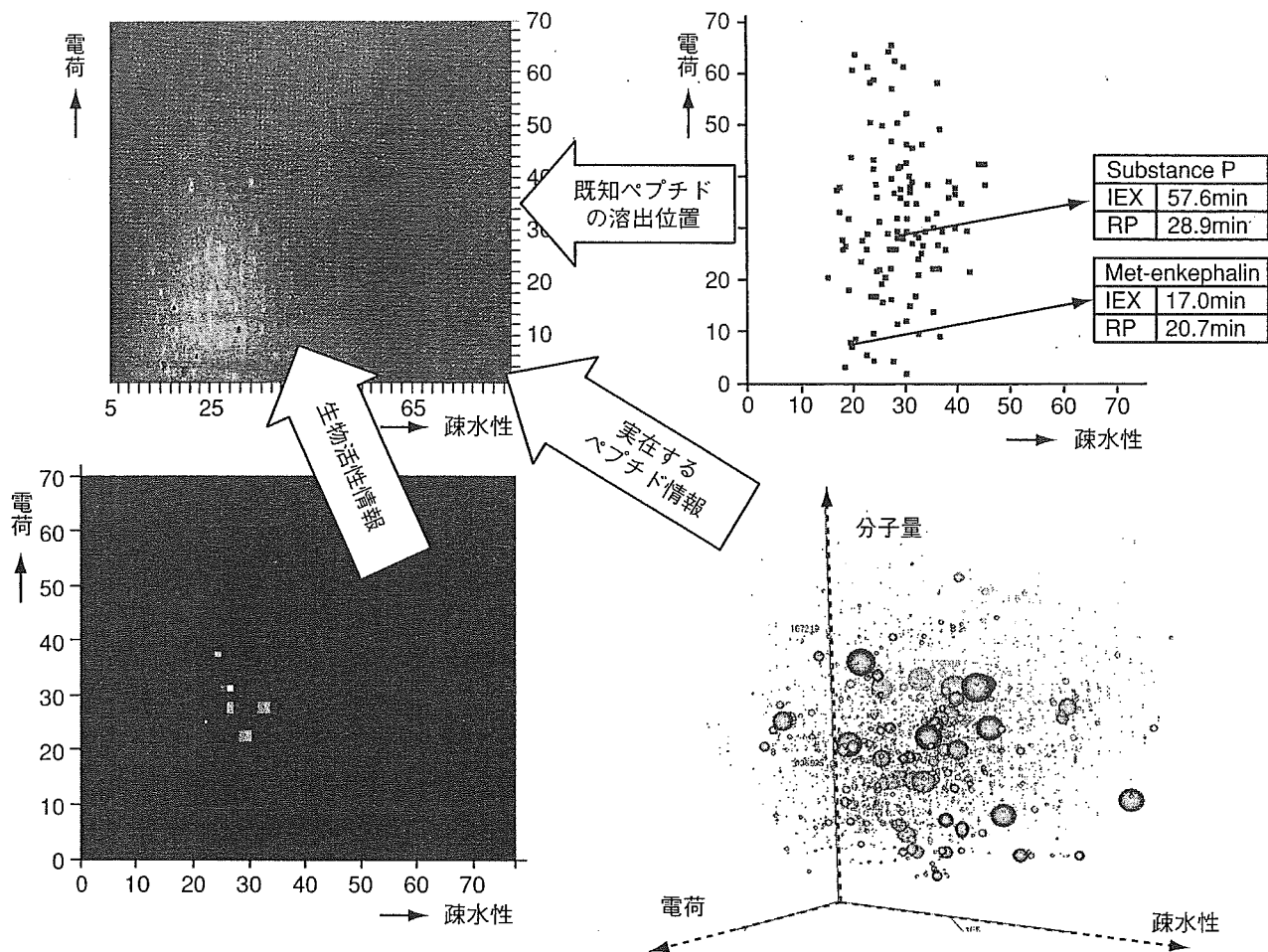


図2 ペプチドーム・データベースにおける各種情報の集積

ペプチドーム・データベースでは規格化された二次元HPLCで分離されたペプチド試料を以後の分析に使用するため、この二次元HPLCを共通のプラットフォームとしてすべての情報を集積でき、比較、検索だけでなく、研究グループ間での情報の交換も可能である。この図は、生物活性情報、実在するペプチドの情報、既知生理活性ペプチドの溶出位置情報が集積され、新規活性ペプチド探索に利用可能なことを示す（文献18より引用改変）

この図はマウス脳に存在する分子量3,000以下のペプチドを量の多い順に4,000個表示したものと見える。公開を開始したweb上では、各点をクリックすると分子量、疎水性、電荷のほか、アミノ酸配列、前駆体名やペプチドの位置、活性の有無、分類、作用、受容体、分布、局在などの情報が表示され、他のデータベースや文献情報へのリンクが可能となる（現在は一部のみ可能）。

物性と分子量情報の収集には、約50mgの組織より抽出したペプチドを使用している。使用する質量分析計の感度を考えると、組織濃度1 pmol/g程度のペプチドまで検出されると見積もられる。タンデム質量分析による構造情報の取得では数百fmol以上のペプチドを使用するため、組織濃度1 pmol/gとして1g程度の組織が必要である。脳ペプチドの解析より、実際

に1~2gの組織で10,000~20,000ペプチドの検出、3,000~5,000ペプチドの構造情報が入手できると推定している。

図4にこれまでに同定されたマウス脳ペプチドについて、その由来タンパク質の内訳を示した。同定されたペプチドの約8%は、ホルモンや生理活性ペプチドの前駆体由来であった。また、約7%は神経や内分泌細胞の分泌顆粒中に多く含まれるタンパク質由来で、約8%は分泌タンパク質由来であった。これら以外の約75%は細胞内や細胞膜に多量に含まれる構造タンパク質やエネルギー代謝系酵素類由来のペプチドである。これらを細胞内の局在に従って分類すると、細胞質由来タンパク質が約35%、細胞膜、ミトコンドリア、細胞内情報伝達、核のタンパク質がそれぞれ6~8%、神経細胞特異的なタンパク質が4%、

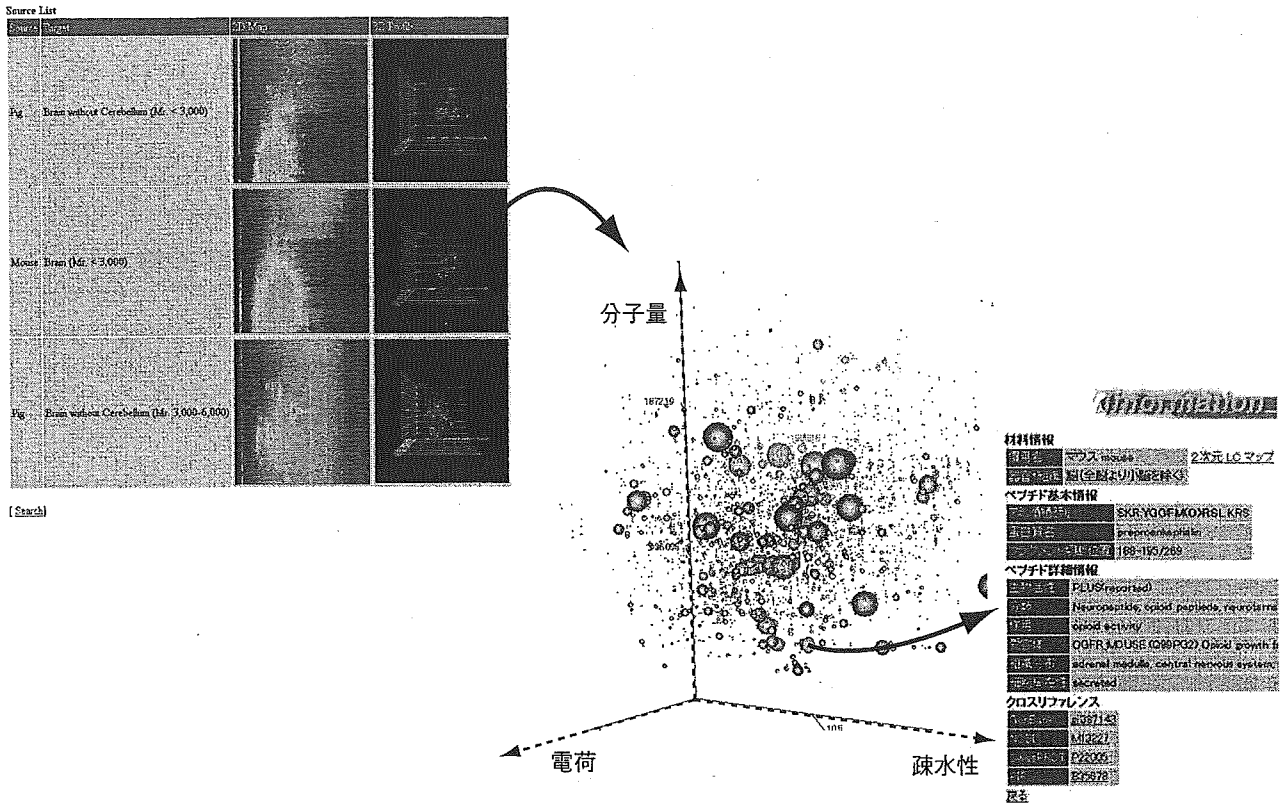


図3 仮想的三次元空間（疎水性，電荷，分子量）に表示した約4,000個のマウス脳ペプチド。各点がペプチドを示し，その大きさはイオンカウントを示し存在量とほぼ相関する。各点をクリックするとそのペプチドに関する情報が表示される。この図は分子量3,000以下のペプチドデータで，これに分子量3,000～6,000までのデータを加えると，マウス脳内ペプチド全体が概観できる（文献18より引用改変）（表紙参照）

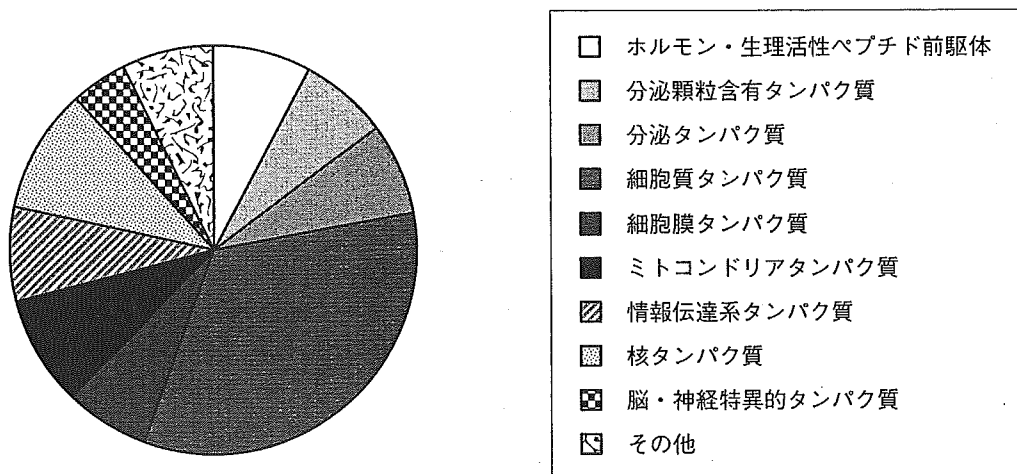
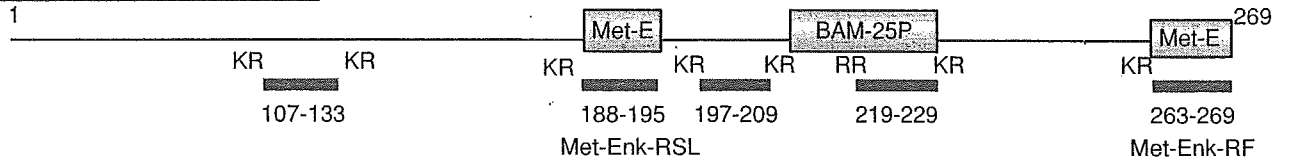


図4 マウス脳で同定されたペプチドの由来するタンパク質の内訳。分子量3,000以下のペプチド画分の解析結果に基づく。

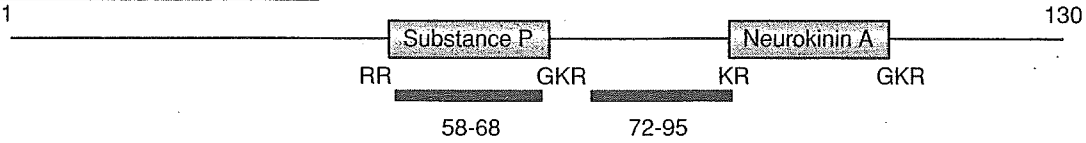
ヘモグロビンなどが残りの7%となっている。細胞内にはタンパク質の代謝的分解に由来するペプチドが多数含まれることから，これらの数値はほぼ妥当なものと考えている。図5に示すように，ペプチドホルモン

前駆体に由来するペプチドの大部分は，プロホルモン変換酵素（PC：prohormone convertase）やFurinにより塩基性アミノ酸対で切断を受け生成していた。しかし，単一アルギニン残基C末端側での切断もかな

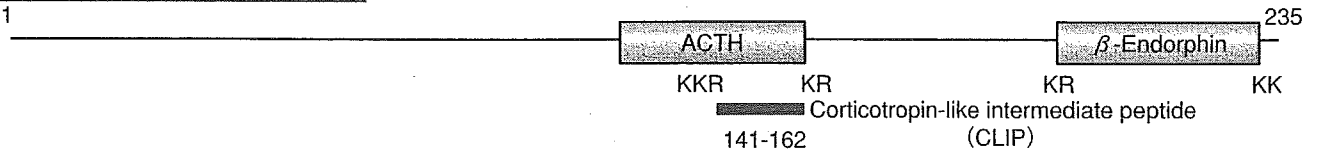
① Proenkephalin A 由来ペプチド



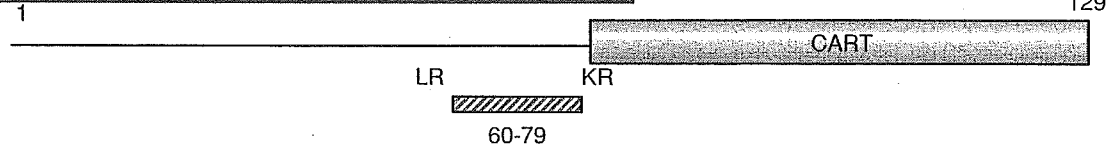
② Protachykinin-1 由来ペプチド



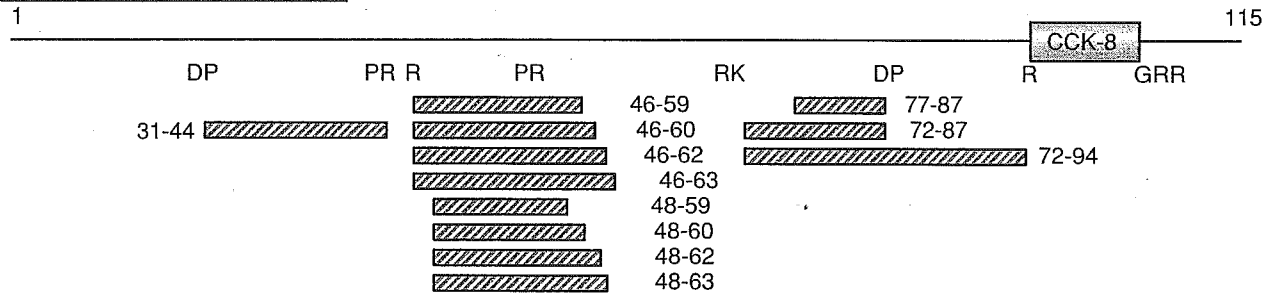
③ Proopiomelanocortin 由来ペプチド



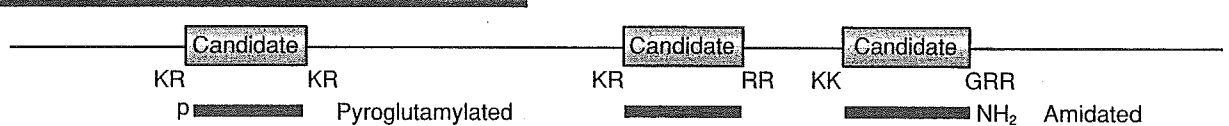
④ Cocaine- and amphetamine-regulated transcript 由来ペプチド



⑤ Procholecystokinin 由来ペプチド



⑥ 分泌タンパク質 (非ホルモン前駆体) 由来ペプチド



⑦ 非分泌タンパク質 (非ホルモン前駆体) 由来ペプチド

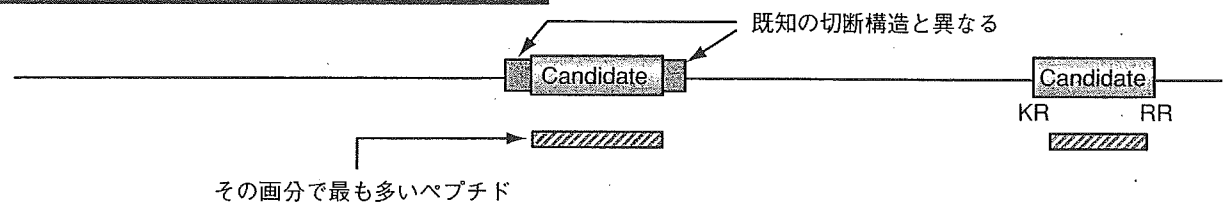


図5 マウス脳で同定されたペプチドと前駆体タンパク質との関係

①～⑤はホルモン前駆体由来ペプチドの実例、⑥と⑦は解析結果に基づき作成した非ホルモン前駆体の分泌タンパク質と非分泌タンパク質の仮想的プロセッシング例を示す。前駆体の下側の棒は同定したペプチドを示し、 は典型的なプロセッシング、 は非典型的なプロセッシングにより生成すると考えられる

りの頻度で観測されている。また、ホルモン前駆体とは想定されていないタンパク質でも、PCや他のプロセシング酵素により切断を受けたと考えられるペプチドが検出されている。これらの一部は能動的に生成され未知の機能を担う可能性があるため、抗体などを用いた選択的切断と内在性の確認、合成ペプチドを用いた生物活性の同定が次の関門である。ペプチドーム解析で見出された非ホルモン前駆体タンパク質由来ペプチドにある程度の頻度で生理活性が検出できれば、ホルモン前駆体に対する概念の再考も必要となりうる。将来的には本データベースにプロセシング酵素（タンパク質分解酵素）情報なども収録し、より利用価値を高めていきたい。

データベース構築の主要な目的は、ペプチド情報を総合的に収集し新規活性ペプチドの発見を促進することである。ブタ脳については、二次元HPLCで分離した各画分について生物活性情報も収集している。図2には、神経芽腫細胞株の細胞内Ca²⁺濃度上昇活性と、既知の活性ペプチドの溶出位置（座標）を示した。二次元HPLC上で観測された生物活性陽性の画分と既知の生理活性ペプチド溶出位置の比較より、生物活性と既知ペプチドの関連をかなり明確に推測できる。図2の細胞内Ca²⁺濃度上昇活性の場合、5つの活性はサブスタンスPとN末端2残基が欠如したペプチド、それらのメチオニン酸化物とニューロテンシンに由来していた。カルシトニン受容体刺激ペプチド（CRSP：calcitonin receptor-stimulating peptide）も、類似の二次元クロマト上で既知ペプチドに一致しない活性画分より精製に成功したものである¹⁰。オーファン受容体の検出系と組み合わせたリガンド探索も可能で、モデル実験としてノシセプチンを同定した¹¹。

厄介者の分解ペプチドも有効利用できる道がある。分泌ペプチドで安定な分子は、疾患マーカーとして利用できる可能性がある¹²。また、データベースから1つのタンパク質に由来するペプチドを整理すると、そのタンパク質の代謝分解経路がわかる可能性が高い。ホルモン前駆体の場合、塩基性アミノ酸対の両側で切断される事例が多いが、非典型的な切断部位はゲノム構造からは予測できない。例えば、ヒト試料の解析データより特異的変換酵素によるエンドセリンの切断部位も浮かび上がってくる。ペプチドホルモンは一般に血中半減期が短い、分解酵素が同定されれば阻害剤

の投与により活性ペプチドの濃度を維持することができる。例えば、ナトリウム利尿ペプチド、ブラジキニンなどは、細胞膜結合型タンパク質分解酵素のNEP（neutral endopeptidase）で主に分解される¹³。NEP阻害剤を投与すると、これらの降圧性ペプチドの血中半減期が延長し、血圧を低下させることが可能となる。残念ながらNEP阻害剤は治験段階で副作用が見出されたが、タンパク質分解経路の推定も本データベースの利用法の1つである。

5 世界のペプチド解析の流れ

プロテオーム解析は世界各国でいっせいに開始されているが、ペプチドーム解析を行っている施設は非常に少ない。データベース化を視野に入れて実際的なデータ収集を行っているのはドイツのバイオビジョン社である。キャピラリー逆相HPLCとMALDI-TOF質量分析計を活用し、再現性の高く高感度なシステムを構築しており、ヒト体液中の分子量2万までのペプチドと低分子量タンパク質の解析を進めている¹⁴。疾患マーカーの探索、薬剤によるペプチド変動、モデル動物のペプチド解析などを実施しており、先日開催された大阪大学蛋白質研究所国際シンポジウムでは糖負荷によって血中インスリンが検出可能となることを報告している。

ウプサラ大学のアンドレンらは、神経変性疾患モデルマウスにおける脳内ペプチドの解析を行っている。ラットやマウス脳を局所的にマイクロエーブで急速に温度上昇させてタンパク質分解酵素を瞬時に不活性化し、1mgの視床下部より550ペプチドを検出している¹⁵。ペプチドの表示は、アマシャム社とプロテオームの表示法を改良し、逆相HPLCの保持時間を横軸、質量を縦軸とする二次元表示法、イオンカウントを加えた三次元表示法などを作成し、視覚的な認識を可能としている。そのほかに非哺乳類のペプチドーム解析の事例として、国立遺伝学研究所のヒドラ¹⁶、欧州での昆虫の側心体の解析¹⁷などが知られている。2004年1月には、ペプチドームに関する国際シンポジウムがわが国ではじめて開催された⁴。今後は、ペプチドーム解析で得られた新規生理活性ペプチドの機能解析を進めるとともに、ペプチドーム用情報収集、解析システムを広く提供して、ペプチドームの概念を世界標準へと成長させたいと考えている。

おわりに

本プロジェクトは科学技術振興調整費知的基盤研究により実施してきたもので、共同して提案・実施してきた大阪大学蛋白質研究所・高尾敏文教授に厚くお礼を申し上げます。データベース (<http://www.peptidome.org>) のホームページでは、マウス、ブタ脳の情報进行を掲示しているが、情報量が限られており、今後その質、量の向上に努めたい。幅広い研究に利用可能なデータベースとするため、ご意見、ご批判を頂ければ幸いである。

文献

- 1) Seiler, P. et al. : J. Chromatogr. A, 852 : 273-283, 1999
- 2) 南野直人 : 蛋白質核酸酵素, 46 : 1510-1517, 2001
- 3) Minamino, N. et al. : J. Chromatogr. B, 792 : 33-48, 2003
- 4) <http://www.peptidome.org>
- 5) Kojima, M. et al. : Nature, 402 : 656-660, 1999
- 6) Sudoh, T. et al. : Nature, 332 : 78-81, 1988
- 7) 南野直人 他 : 『シグナル伝達実験法』 (宇井理生／編), pp.23-34, 羊土社, 1996
- 8) Fernandez-de-Cossio, J. et al. : Nucleic Acids Res., 32 (Web Server issue) : W674-W678, 2004
- 9) Fernandez-de-Cossio, J. et al. : Rapid Commun. Mass Spectrom., 18 : 2465-2472, 2004
- 10) Katafuchi, T. et al. : J. Biol. Chem., 278 : 12046-12054, 2003
- 11) Takeda, S. et al. : J. Biochem. (Tokyo), 135 : 597-604, 2004
- 12) Sasaki, K. et al. : Cancer Res., 62 : 4894-4898, 2002
- 13) Kubota, E. et al. : Essays Biochem., 38 : 129-139, 2002
- 14) Schulz-Knappe, P. et al. : Comb. Chem. High Throu. Screen, 4 : 207-217, 2001
- 15) Svensson, M. et al. : J. Proteome Res., 2 : 213-219, 2003
- 16) <http://www.nig.ac.jp/labs/OntoGen/PEPTIDE.html>
- 17) Schoofs, L. et al. : Brief Funct. Genomic Proteomic., 2 : 114-120, 2003
- 18) 南野直人 : ファルマシア, 39 (12) : 1157-1162, 2003

<著者プロフィール>

佐々木 一樹 : 1990年東京大学医学部医学科卒業。'97年国立がんセンター研究所細胞増殖因子研究部室長。2004年より国立循環器病センター研究所薬理部室長。

E-mail : ksasaki@ri.ncvc.go.jp

磯山正治 : 1984年大阪大学理学部生物学科卒業。'89年より蛋白質研究奨励会情報室勤務。'94年より情報管理室長。

E-mail : isoyama@prf.or.jp

南野直人 : 1976年大阪大学理学部化学科卒業。長崎大学医学部、宮崎医科大学助手を経て'89年より国立循環器病センター研究所室長。2002年より現職。松尾壽之先生の指導の下、ペプチド探索研究を開始。最近はペプチドーム・データベースの構築研究を推進。

E-mail : minamino@ri.ncvc.go.jp

G.I. Research

別刷

発行：株式会社 先端医学社

〒103-0007 東京都中央区日本橋浜町 2-17-8 浜町花長ビル

膵癌マーカー

佐々木一樹*

血中低分子量蛋白質・ペプチドの発現解析にもとづく腫瘍マーカーの探索が、プロテインチップシステムとよばれる質量分析計の普及にともない現実味を帯びはじめている。本特集で紹介されているように、本法は、さまざまな化学表面をもつプレートを用いるのが特徴的であり、血液や臨床材料など、従来の質量分析法では不得手とされてきた試料の分析を可能とした。早期発見に有効なマーカーが存在しない膵癌の領域でも、実際の血液検体を用いてのマーカー探索が盛んに実施されている。本稿ではその現状と展望について言及する。

はじめに

膵腫瘍のおよそ9割を占め、最も予後不良な浸潤性膵管癌について本稿では記載する。膵管癌で広く使用されてきた腫瘍マーカーのCA 19-9は、糖鎖抗原の一種で、ヒト腫瘍を移植した担癌動物の免疫系で認識される物質として同定されている分子である。CA 19-9は、Stage II以上では80%以上の症例で陽性になるものの、良好な治療成績が得られる可能性の高いStage Iでは陰性であることが多い。また、ほかの消化管の腺癌でも陽性になり、肝・胆・膵の良性疾患でも陽性になる割合が無視できない点などから、現在のところ早期発見に有用な腫瘍マーカーとはいえない。また、CA 19-9の検査で検出されているのはシアリルLe^aとよばれる糖鎖であり、これはLewis 式血液

型物質の関連しているため、日本人の10%程度は膵癌になってもCA 19-9が陽性にならないといった問題も抱えている。現在、腫瘍マーカー開発の領域では、蛋白質・ペプチドの発現解析の流行に先立ち、SAGE(serial analysis of gene expression)法や、オリゴマイクロアレイ・cDNAマイクロアレイにより膵癌で高発現を示す遺伝子が同定され、その蛋白質レベルでの評価が徐々に報告されはじめているが^{1)~4)}、本稿では割愛する。

1. SELDI 法による癌診断の試み

1993年に表面改良型レーザー脱離イオン化質量分析法(Surface enhanced laser desorption ionization: SELDI)が発表された⁵⁾。当初はほとんど注目を受けなかったが、2001年の米国癌学会で、血清を用い、大腸癌、卵巣癌などについて早期を含め9割を超える感度と特異性を達成したという発表がされた。これは、血清をそのままSELDIプロテインチップに1 μ lほど塗布し、分子量10 kDa以下の領域のマスペクトルを取得し、その質量数とピーク強度をもとに、機械学習

【キーワード】

SELDI

膵癌

質量分析

血液検体

*Kazuki SASAKI/国立循環器病センター研究所薬理部

法でデータ解析をおこなう方法であった。手技の簡便さも相まって、類似手法が多数報告され、米国を中心に大きな反響をよんだ。翌年の2002年春に *Lancet* 誌に卵巣癌を対象として最初の論文発表をしたグループの解析データはウェブ上に公開されたこともあったが、データは改変されており再現不能であるという指摘もなされていた⁶⁾。また、癌と正常を識別しうるピークとして質量/電荷比 (m/z) 900 以下のピークが計上されていたが、SELDI のイオン化法は基本的に MALDI (matrix assisted laser desorption/ionization: マトリックス支援レーザー脱離イオン化) 法と同じであり、そのような低分子域に出るピークをマトリックスと識別するのは、量的な観点からも事実上不可能であることから、質量分析の分野からも疑問が投げかけられていた。

2. 昨今の解析

SELDI のプロテインチップは、試料添加表面がさまざまに加工されており、異なる種類のチップを用いることで、同一試料から異なるスペクトルパターンが得られる。膀胱癌のマーカー探索では、膀胱液、培養細胞の上清の分析が当初実施され、目的ピークを同定していく作業が一般的であった^{7,8)}。現在は、膀胱癌を含めほぼすべての癌で、100 μ l 以下の血清または血漿を用いて、試料をあらかじめイオン交換担体を用いて段階的に複数に分画し、異なる種類のプロテインチップを用いてスペクトルを取得し、さまざまな機械学習法で、癌と正常を明瞭に識別しうるピークセットを選択する手法が主流となっている⁹⁾。実際には2~5個のピークセットにもとづいて、癌症例と健常者との検討で、特異性と感度が9割以上であると報告している論文が多い。その質量値はほとんどの場合2~15 kDa 程度までに集中しているが、この領域の分子量のポリペプチドの構造決定は煩雑な操作を要することもあり、同定されている事例は少ない。しかしながら、同定されている事例では、ほ

とんど全部が血漿蛋白質の断片ペプチドである。これらが実際の病態生理上どのような意味合いをもっているかは現状では不明である。腫瘍塊の大きさと検出感度などから見積もると、癌細胞自身が産生分泌している物質であるとは考えにくい。一方で、癌細胞そのもの、あるいは宿主側との相互作用の結果として活性化されたプロテアーゼのカスケードの最終産物とみなすことを積極的に否定することは難しい。血液のようなきわめて複雑度の高い試料を解析する場合、段階的分画とプロテインチップの組み合わせで再現性よく検出できるピークの数に限定されており、質量分析計の感度を考慮すると、同定されたペプチドは存在量が多い(少なくとも10 μ g/ml 以上)蛋白質の断片に由来するのは妥当である。分画しているとはいえ、血液のような混合度の高い試料を質量分析している点にかわりはないため、健常群と疾患群で相対強度が2,3倍異なって呈示されてくるピークについて、実際に定量した場合の量的な差異がどの程度であるかについても不明な点が多い。

また、SELDI 法で用いられた当初の質量分析計は、感度を最優先する目的でリニア型の短い飛行管を搭載していたために、質量精度と分解能が犠牲になりやすい傾向があった。現在では SELDI プロテインチップは、通常四重極-飛行時間型質量分析計でも測定可能になっており、この装置によってマススペクトルを取得する試みもはじまっている¹⁰⁾。また、特定の操作で選択をかけた血液分画についてエレクトロスプレーイオン化法の質量分析計でその包括的なマススペクトルを取得し、SELDI と同様に機械学習法で診断マーカーを開発しようとする試みもあるが、実態は不明である。

3. 実際的な問題点

SELDI による分析では、チップ表面での試料調製が可能であり、いったん試料が準備できるとマトリックスを添加しての測定のみであるため、非常に簡単なものと考えられがちであるが、実際は