

ゲノミクス・プロテオミクスの新展開 生物情報の解析と応用

第2編 第1章 第7節

発現プロファイル解析の基礎技術と方法

二次元電気泳動によるディファレンシャル
プロテオミクスとデータベースの構築

戸田 年総

東京都老人総合研究所

発現プロファイル解析の 基礎技術と方法

第7節

1. 二次元電気泳動によるディファレンシャルプロテオミクスとデータベースの構築

1.1 はじめに

プロテオミクスがゲノミクスやトランスクリプトミクスと大きく異なる点は、遺伝子の発現動態を、機能と直結したタンパク質のレベルで定量的に解析できるということにある。プロテオミクスとよく対比されるトランスクリプトーム解析は、遺伝子発現の有無や増減を簡単に知ろうとする場合には有用であるが、すでにGriffinら¹⁾の研究によって明らかにされているようにmRNAのレベルはタンパク質のレベルには直接反映せず、やはり翻訳後修飾や分解除去の速度も含めて包括的にタンパク質の動態を定量的に観察することのできるプロテオーム解析を行うことが重要である。本章では、二次元電気泳動に基づいたディファレンシャルプロテオミクスの基礎技術と、解析結果のデータベース化について述べる。

1.2 二次元電気泳動の再現性

二次元電気泳動は、細胞や組織の粗抽出液中に含まれる数千種のタンパク質を、1枚のゲル板上で分離展開できる高い分解能を有しており、これが、1995年にプロテオーム解析が誕生したときにタンパク質の分離方法として採用された最大の理由である。近年、キャピラリー電気泳動や二次元LCを質量分析計に連結させた方法なども行われるようになっているが、数千種のタンパク質を一度に分離できるほどの高分解能は達成されていない。二次元電気泳動の分解能は、1975年にO'Farrell²⁾によって発表された当初から多くの研究者によって高く評価されていたが、いざ実際に行ってみると、とくに一次元目の等電点電気泳動において再現性の問題に直面す

ることが多く、これが二次元電気泳動の普及の大きな妨げとなっていた。

ガラスの細管中で行うゲルディスク等電点電気泳動の最大の欠点は、可溶性の両性電解質(両性担体, carrier-ampholytes)によって形成されたpH勾配が、負に荷電したガラスの表面で発生する電気浸透流によって陰極方向に移動し続ける³⁾ことである。その結果、再現性の高い等電点集束(isoelectric focusing)を実現するに十分な時間、等電点電気泳動を続けることが困難であり、とくに塩基性のタンパク質をゲル内に留めておくことは不可能であった。このため苦肉の策として、一次元目の泳動を短時間で止め、塩基性タンパク質がゲルから押し出されてしまうのを防ぐ、いわゆる『非平衡の等電点電気泳動(NEpHGE: nonequilibrium pH gradient electrophoresis)⁴⁾』を組み合わせた二次元電気泳動法が開発された。その後、電気浸透流の問題は、RighettiとGörgらによって開発された『固定化pH勾配等電点電気泳動』を一次元目に行うIPG-DALT: two-dimensional gel electrophoresis with immobilized pH gradients in the first dimension⁵⁾によって根本的な解決を見、今日の『二次元電気泳動に基づく発現プロファイルのディファレンシャル解析』に道が拓かれた。

固定化pH勾配等電点電気泳動ではpH勾配のドリフト現象は起きないので、タンパク質が完全に等電点の位置に集束するまでの十分な時間、高電圧で電気泳動を続けることが可能である(図1参照)。これによって再現性と分解能が格段に改善されたばかりでなく、狭いpH領域ごとに二次元電気泳動を行い、得られたパターンをつなぎ合わせることによって、6,000個以上のタンパク質スポットを一度に解析することもできるようになった⁶⁾。初期の固定化pH勾配等電点電気泳動には、塩基性の等電点領域においてテーリングが若干大きくなるという問題点があったが、この原因がアルカリ条件下におけるシ

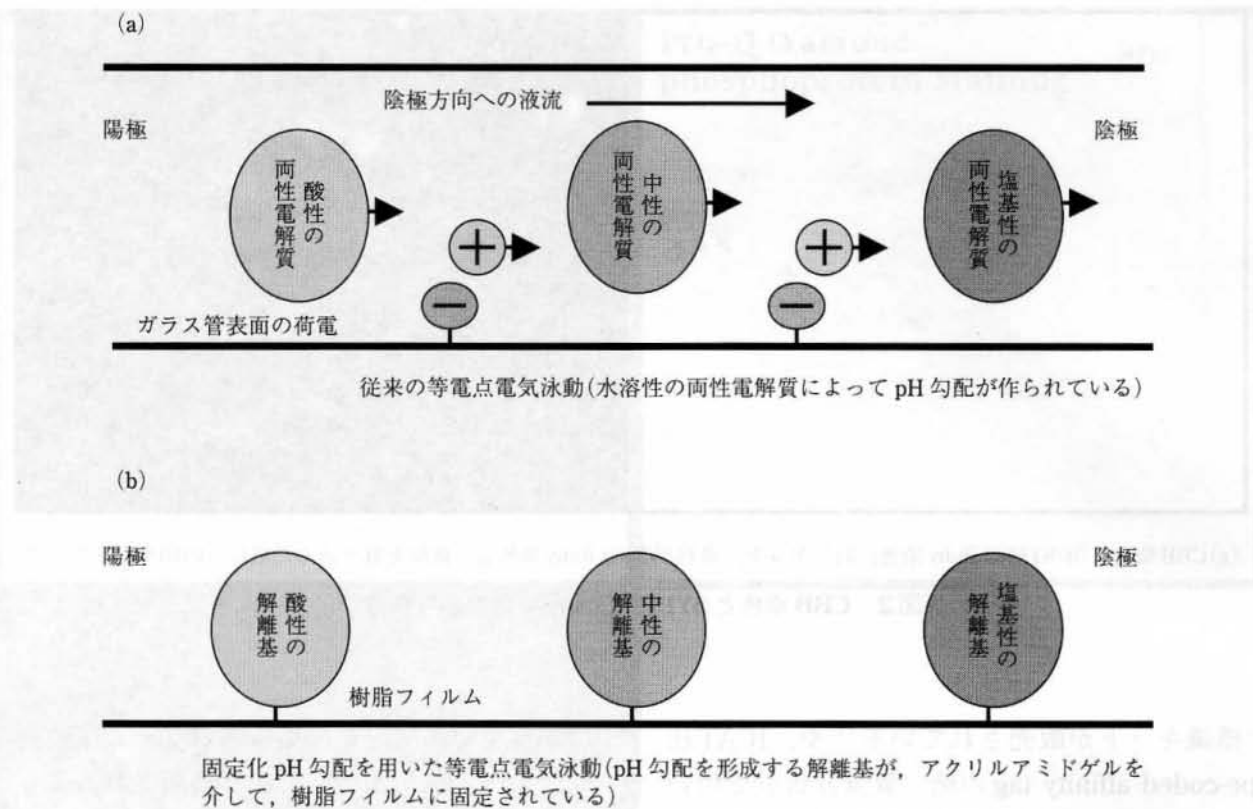


図1 水溶性の両性電解質(両性担体)を用いる従来の等電点電気泳動と固定化pH勾配を用いた等電点電気泳動の違い

従来法ではpH勾配が時間とともに陰極方向に移動するために長時間の通電ができなかったが、pH勾配を固定することによって再現性の高い等電点電気泳動を誰でも簡単にできるようになった

ステイン残基の再酸化⁶⁾にあることがわかり、等電点電気泳動の前に還元アルキル化を行うことや、等電点電気泳動中の再酸化を防ぐ工夫をすることによって、塩基性タンパク質の分解能も改善された。

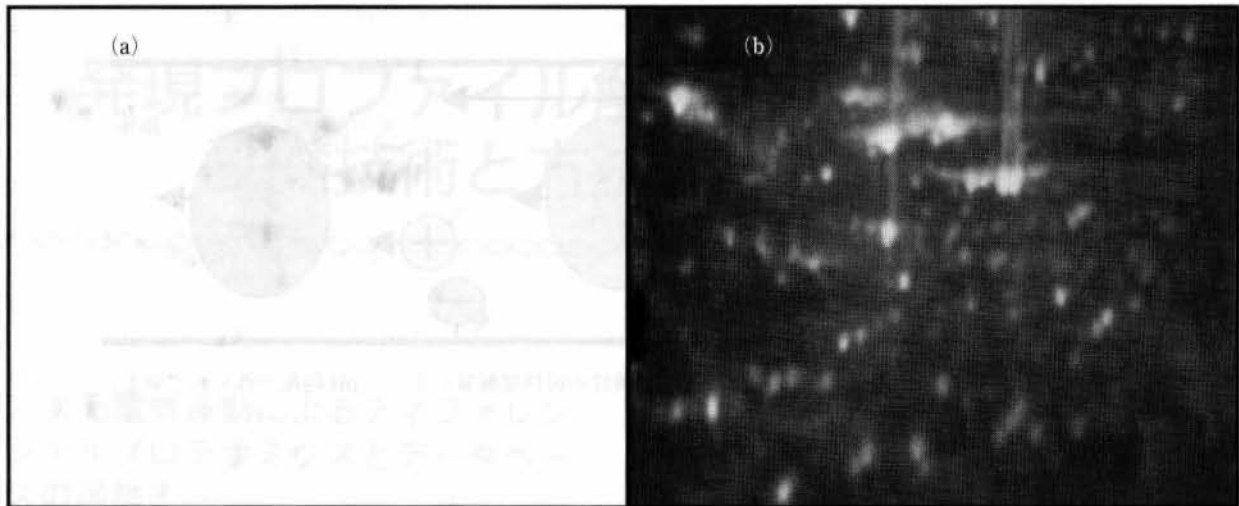
注1) IPGはimmobilized pH gradientの略。DALTはいわゆるSDS-PAGEによる分子量分画電気泳動法のことを指す。Daltonの意味)

1.3 タンパク質染色の定量性と特異性

ディファレンシャルな発現プロファイリングを行うためには、ゲル上のタンパク質スポットを定量的に染色できる方法が必要となる。電気泳動による分析において定量性が要求される場合、従来はもっぱらCBB(Coomassie Brilliant Blue)染色が利用されていたが、スポットの検出限界が0.2 μg前後と低いために、細胞内シグナル伝達などにかかわる発現量の少ないタンパク質の検出には不向きである。一方、高感度の銀染色法では数ng程度の微量のタンパク質も検出できるものの、タンパク質の種類により染

色性が大きく異なり、スポットの濃淡から直接タンパク量を推定することは困難である。これに対し、SYPRO Ruby(米国Molecular Probes社の製品で、販売はInvitrogen Life Technologies社が行っている)による蛍光染色は、感度が銀染色なみに高いうえに、定量性もCBB染色なみに高いという両者の長所を併せ持っていることから、ディファレンシャルプロテオミクスで広く利用されるようになっていく(図2参照)。

このほか、二次元電気泳動の再現性を気にせずに定量的な比較分析が行える方法として2-D DIGE法(two-dimensional differential gel electrophoresisの略。1枚の二次元電気泳動ゲル上で2種類のサンプル中の対応するスポットの定量的比較を行う方法。電気泳動を行う前にタンパク質をそれぞれ異なる波長特性の蛍光色素で標識しておき、それらを混合して二次元電気泳動を行うというもの。泳動後にそれぞれの励起波長・発光波長でパターンを観察し、画像を重ね合わせることによって定量的比較を行う。Amersham Biosciences社からEttan DIGEという商



(a)CBB染色, (b)SYPRO Ruby染色。同じゲルを, 最初SYPRO Ruby染色し, 画像を取り込んだ後に, CBB染色を行った

図2 CBB染色とSYPRO Ruby染色の感度の違い

品名で標識キットが販売されている)⁷⁾や, ICAT法 (isotope-coded affinity tagの略。質量分析計を用いてタンパク質の発現量を定量的に比較するための方法。それぞれのサンプル中のタンパク質を, あらかじめ安定同位体の混合比の異なる(質量の異なる)試薬で標識しておき, これを混合してトリプシン消化した後に, 質量分析計を用いて両者の存在比を測定するというもの)⁸⁾などの新しい方法も開発されているが, これらの方法の最大の問題点は, 常に1対1の比較において両者の違いが解析されるという点である。腫瘍細胞とその周辺の正常細胞の比較のように, コントロールとなるサンプルが得やすい場合には有効であるが, 肝障害のような慢性疾患や糖尿病などの生活習慣病のように, 同一個体の中でコントロールとなるサンプルを得にくい場合には健常者のサンプルを使わざるを得ず, 個体差の方が大きく出てしまうためにあまり有効ではない。

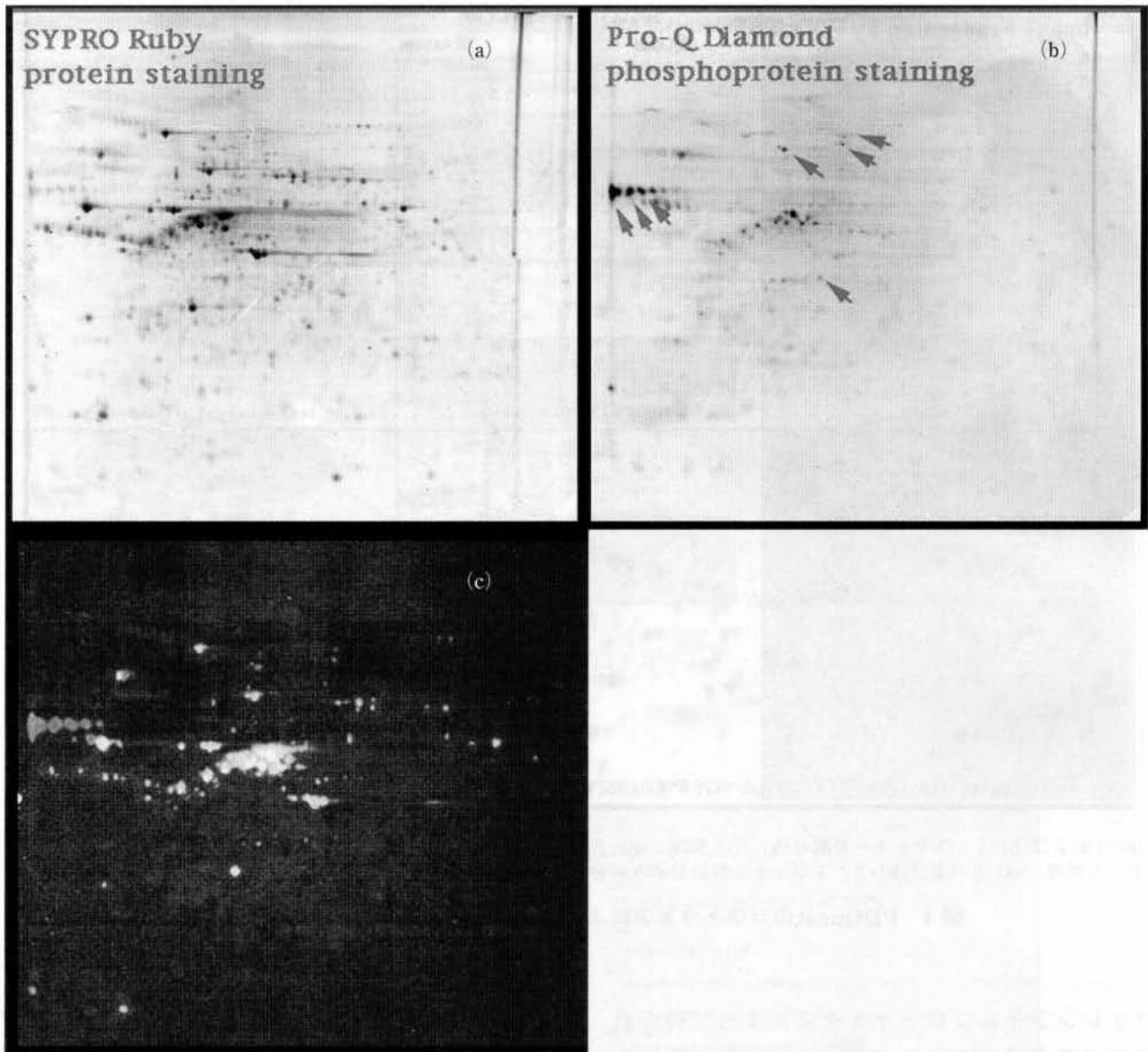
修飾タンパク質検出法においても, 最近新しい進歩が見られている。従来リン酸化タンパク質の特異的な検出には, もっぱら抗体によるウエスタンブロット法が利用されていたが, 定量的な解析には不向きであった。これに対し, Pro-Q Diamond(米国 Molecular Probes社の製品で, 販売はInvitrogen Life Technologies社が行っている。詳しい資料が, <http://www.probes.com/lit/bioprobes42/1.pdf>で公開されている)によるゲル内染色法⁹⁾は定量性も高く, 同一ゲル上でSYPRO Ruby¹⁰⁾による二重染色が行えるので, タンパク質発現プロファイルとの対応づ

けが容易であり, 今後細胞内シグナル伝達機能のプロテオーム解析法として広く利用されるようになるものと思われる(図3参照)。

1.4 画像解析による定量的なディファレンシャル解析

タンパク質発現プロファイルを定量的に比較分析するには, コンピュータ画像解析による数値化とスポットのマッチングが必要である。二次元電気泳動が開発された当初は, まだミニコンピュータが全盛の頃であり, 床置き型の大規模なコンピュータシステムが必要であった¹¹⁾。その後UNIXワークステーション上で動くPDQuestソフトウェアが市販されるようになり, さらにWindows PC上でも利用できるソフトウェアが市販されるに至って, 多くの大学や研究所で広く利用されるようになった¹²⁾ (Bio-Rad Laboratories社の製品。30日間無料の試用版がネット上で公開されている。http://www.proteome-works.bio-rad.com/B2B/vanity/proteomeworks/start.jsp?nextPage=product&childStr=imageAnalysis&specFile=pro_imaging.html)。

現在PDQuestのほかに, Image Master (Amersham Biosciences社の製品。http://www.jp.amershambiosciences.com/technologies/2d-electro/2d_11.asp)やMelanie 3 (Swiss Institute of Bioinformatics (SIB)が開発したソフトウェア。<http://www.ncbi.nlm.nih.gov/Class/NAWBIS/Modules/Protein/protein21.html>), GELLAB-II(もとも



(a)SYPRO Ruby 染色, (b)Pro-Q Diamond 染色, (c)擬似カラー表示による両染色法の比較画像。同じゲルを, 最初 Pro-Q Diamond 染色し, 画像を取り込んだ後に, SYPRO Ruby 染色を行った。擬似カラー表示による両染色法の比較画像は, PDQuest のソフトウェアを用いて作成した

図3 Pro-Q Diamond によるリン酸化タンパク質の特異的検出

とは米国 NIH で開発されたソフトウェア。現在は Scanalytics 社から販売されている。http://www-lecb.ncifcrf.gov/lemkin/gellab.html) といった様々なソフトウェアが販売されており, 各研究者が自分の好みと研究の目的によって使い分けができるようになってきている(図4参照)。

1.5 定量的な発現プロファイル解析結果のデータベース化

<プロテオームデータの階層的管理>

現在インターネット上には, SWISS-2DPAGE (http://us.expasy.org/ch2d/) に代表される, 数多

くのプロテオームデータベースが公開されているが, 本来プロテオーム解析の最大のメリットであるタンパク質の発現量に関する情報は, ほとんど見当たらない。われわれが行っている老化機構研究においては, 細胞の老化に伴い多くのタンパク質において発現量の増減が認められており, 変化の形も『単純に増加するもの』, 『減少傾向を示すもの』, 『いったん増加した後に減少に転じるもの』など, 様々なパターンを示している。老化形質とタンパク質の変化の間の関係を解明するためには, 老化以外の様々な生理的・病的プロセスにおいて生じる変化と対応させながら解析を進める必要がある。それには,

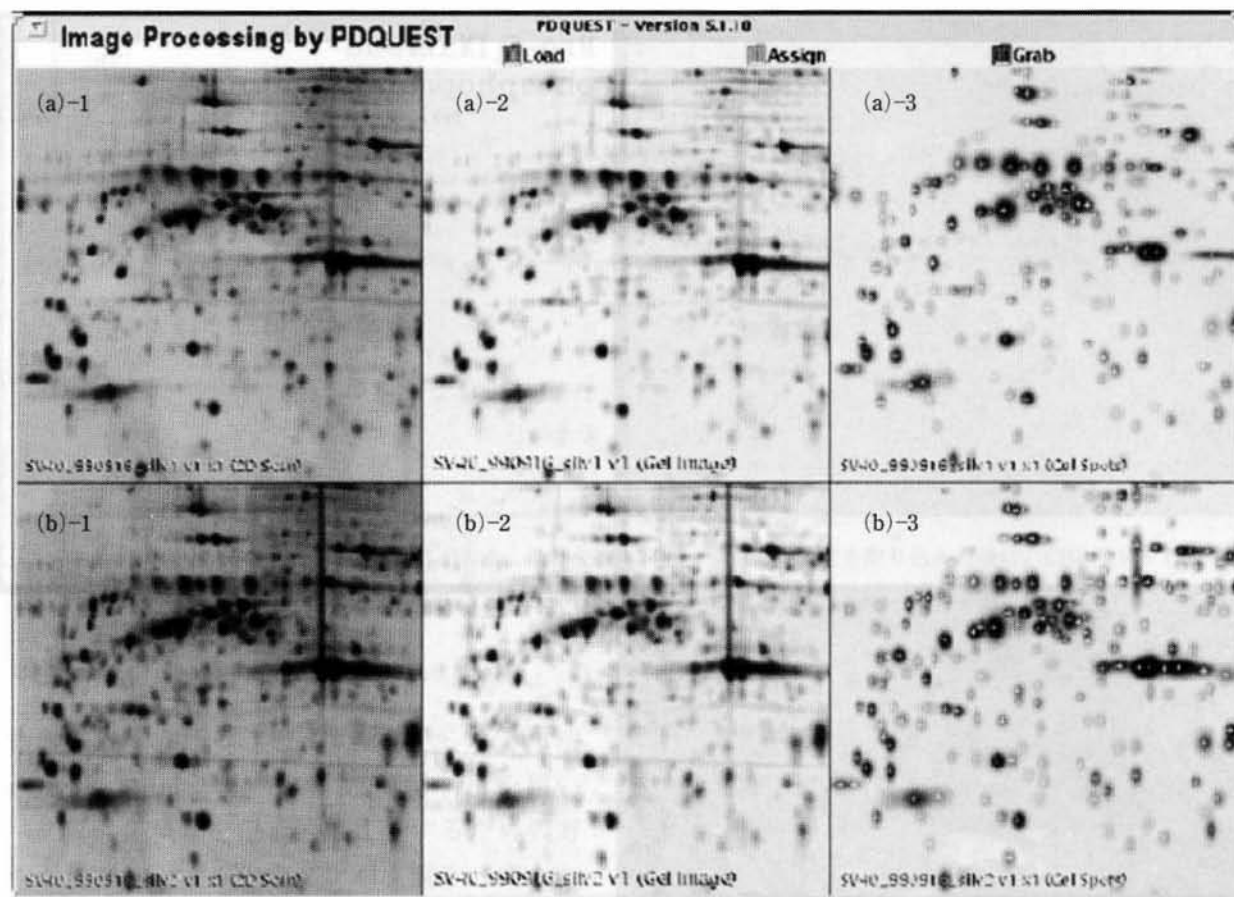


図4 PDQuestのソフトウェアによる二次元電気泳動パターン画像処理
 (a)-1および(b)-1: スキャナーで取り込んだ生画像。(a)-2および(b)-2: 細かい汚れによるノイズやバックグラウンドを消去した画像。(a)-3および(b)-3: スポットの検出とガウス分布近似の処理を行った後にスポットのマッチングを行った画像

図4 PDQuestのソフトウェアによる二次元電気泳動パターン画像処理

ディファレンシャルプロテオミクスによって得られた発現量の数値を情報として盛り込んだ『定量的なプロテオームデータベース』を構築することが必要である。

われわれはこれを実現するために、外部に公開して差し支えない基本情報についてはホームページサーバー¹³⁾ (<http://proteome.tmig.or.jp/2D/>) 上に、医薬品等の開発につながるような未公開情報についてはセキュリティが守られるファイアウォール内のWorksBaseサーバー上にプロテオームデータベースを構築することを行っている(図5参照)。また別のwebサーバー上で、XMLフォーマットによるデータベースの構築も試みている(<http://www.proteome.jp/2D/>)。HTMLに比べてXMLフォーマットが有利な点はいろいろあるが、最も大きな違いは、HTMLではデータ本体とブラウザ上での表示形式が同一のファイルに書き込まれており、データベースの再構築が困難であるのに対し、XMLではデータの本体とブラウザ上での表示形式を指示する情報

(スタイルシート)が分割されているために、データベースとしての管理に向いている。例えば、東京都老人総合研究所で公開しているプロテオームデータベースのヒト二倍体線維芽細胞のssp5005タンパク質のデータの場合、HTML形式では図6のような書き方になっている。データの意味が定義づけられていないため、これを別の解析ソフトで利用することは困難である。これに対しXML形式で書かれた図7のデータでは、各数値が何を意味しているか<タグ>の中に定義づけられているので、再利用が容易である。現在、JHUPO (<http://www.jhupo.org/>)^{注2)}とAOHUPO (<http://www.hupo.org/aohupo/>)^{注3)}によって、HUP-XMLのタグの標準化とデータベース構築に使用するスキーマ(XML文書の取りうる構造と文法を記述したもの)の開発が行われている。

注2) 各国で独自に進められようとしていたプロテオーム研究を国際協調の下に進める事を目的として、2001年2月8日にヒトプロテオーム機構

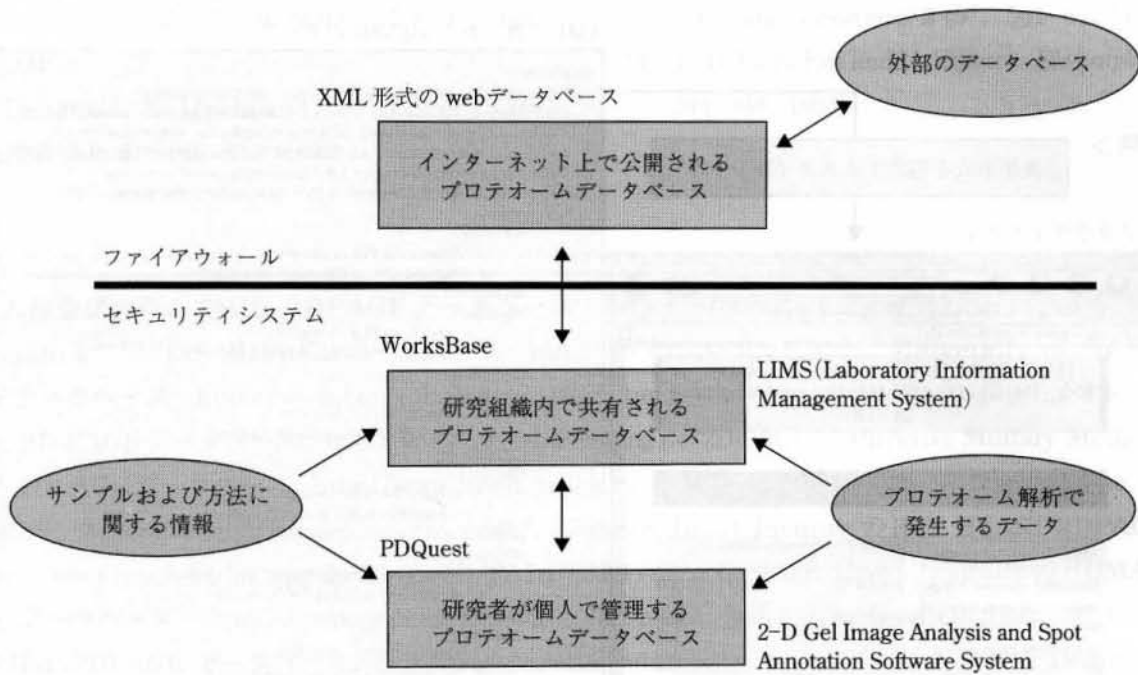


図5 東京都老人総合研究所におけるプロテオームデータベース構築システム

HTML で書かれた ssp5005 のデータ

```

<HTML>
<H2>ssp5005
</H2>
[Physical properties determined by 2-DE]<BR>
Molecular mass = 23.80 kDa
Isoelectric point = 5.78
<BR>
Tissue/cell specificity = ubiquitous
<BR>
Cellular localization = cytosolic
<BR>
Protein name = ubiquitin carboxyl-terminal
<BR><DD>hydrolase L1<BR>
Identified by <A HREF="/MS/ssp5005.jpg">
peptide mass fingerprinting</A><BR>
<DD>and microsequencing<BR>
<HR>
[Activities/functions]<BR>
involved in the hydrolysis of esters and amides
at the C-terminal glycine residue of ubiquitin.<BR>
<HR>
[Related databases]<BR>
ssp on keratinocyte database of Celis: <BR>
<DD>Theoretical Mr: <BR>
PIR: <A HREF="http://www.genome.ad.jp/dbget-bin/
www_bget?pir=A25856" TARGET="_top">A25856</A><BR>
PID: g35440<BR>
swiss-prot: <A HREF="http://www.expasy.ch/cgi-bin/
niceprot.pl?P09936" TARGET="top">P09936</A><BR>
genbank: <A HREF="http://www.genome.ad.jp/dbget-
bin/www_bget?genbank=today+X04741" TARGET="_top">
X04741</A><BR>
</HTML>
    
```

ブラウザ上の表示

図6 HTML で書かれた ssp5005 のデータとブラウザ上の表示

(Human Proteome Organisation : HUPO) が創設が宣言された。JHUPO は HUPO の日本国内組織として 2002 年の 5 月 14 日にスタートし、2003 年の 2 月に茨城県つくば市の産業技術総合研究所共用講堂に

おいて、第 1 回日本ヒトプロテオーム学会が開催された。

注 3) AOHUPO は HUPO のアジア・オセアニア地域の研究者によって組織されたもので、2002 年 3 月

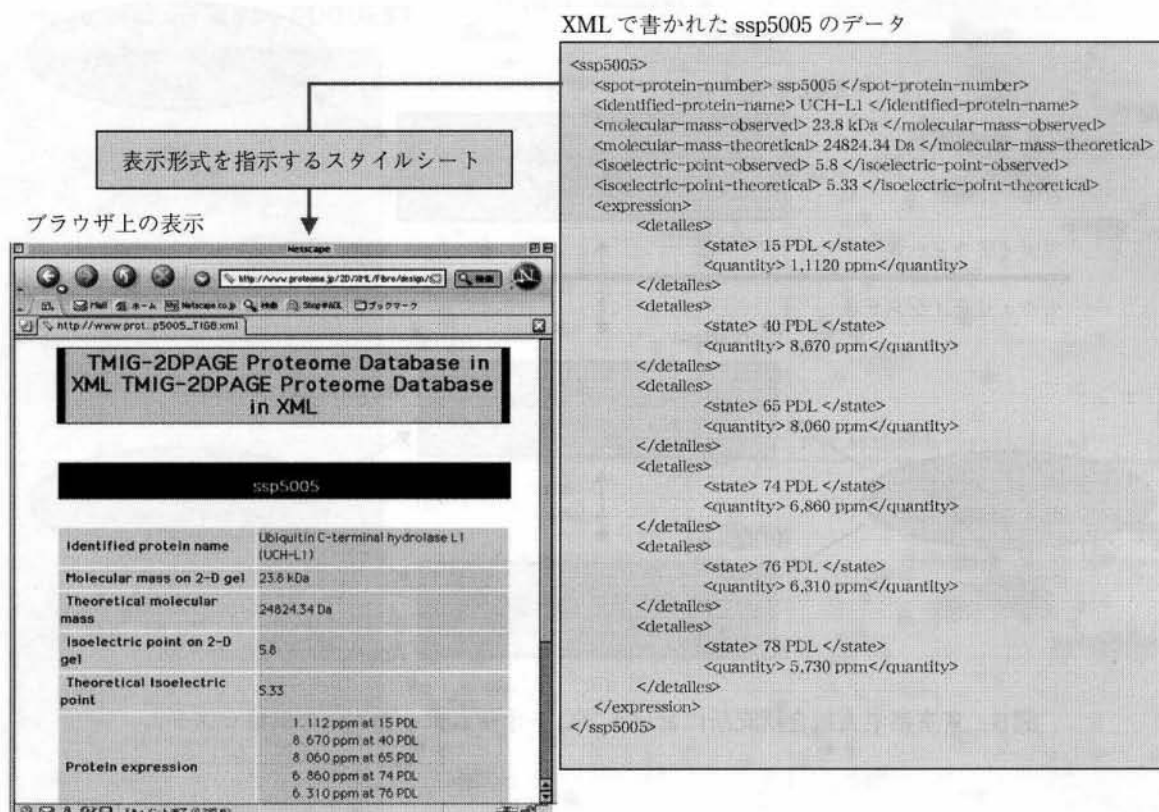


図7 XMLで書かれた ssp5005 のデータとブラウザ上の表示

にはソウルでAOHUPOの主催による第1回国際プロテオミクスシンポジウムが開催された。

1.6 おわりに

最近のプロテオーム解析は、自動化やハイスループトット化を急ぐあまり、定量性や再現性が犠牲にされるような方法に向かう傾向が見られる。単に、発現タンパク質を網羅的に検出するだけならそれでも十分と言えなくもないが、老化や分化、各種疾患の病態の解析などに利用しようとする、発現プロファイルのディファレンシャル解析における、定量性や再現性が十分に保証されていなければならない。現時点でこれに最適であるのは、やはり二次元電気泳動に基づくタンパク質ディファレンシャル解析である。また、得られた情報をデータベース化する場合、単なる同定結果だけでなく、定量的な分析結果も記録しておくことが重要である。現在HUPO(ヒトプロテオーム機構, <http://www.hupo.org/>)では、プロテオーム情報をデータベース化する際の基準作りを進めており、われわれはその中に定量的な情報を含めるように働きかけを行っていきたいと考えて

いる。

【参考・引用文献】

- 1) T. J. Griffin et al. : *Mol. Cell. Proteomics*, **1** (4), 323–333 (2002).
- 2) P. H. O'Farrell : *J. Biol. Chem.*, **250** (10), 4007–4021 (1975).
- 3) P. G. Righetti and C. Macelloni : *J. Biochem. Biophys. Methods*, **6** (1), 1–15 (1982).
- 4) P. Z. O'Farrell, H. M. Goodman and P. H. O'Farrell : *Cell*, **12** (4), 1133–1141 (1977).
- 5) R. Westermeier, W. Postel, J. Weser and A. Gorg : *J. Biochem. Biophys. Methods*, **8** (4), 321–330 (1983).
- 6) K. Altland et al. : *Electrophoresis*, **9** (9), 474–485 (1988).
- 7) M. Unlu, M. E. Morgan and J. S. Minden : *Electrophoresis*, **18** (11), 2071–2077 (1997).
- 8) S. P. Gygi et al. : *Nat. Biotechnol.*, **17** (10), 994–999 (1999).
- 9) K. N. Berggren et al. : *Proteomics*, **2** (5), 486–498 (2002).

- 10) T. H. Steinberg et al. : *Electrophoresis*, **21** (3), 486-496 (2000).
- 11) P. F. Lemkin, L. E. Lipkin and E. P. Lester : *Clin. Chem.*, **28** (4 Pt 2), 840-849 (1982).
- 12) J. E. Celis : *Leukemia*, **2** (9), 561-601 (1988).
- 13) T. Toda, K. Kaji and N. Kimura : *Electrophoresis*, **19**(2), 344-348 (1998).

<戸田 年総>

● URL一覧

東京都老人総合研究所のTMIG-2DPAGEデータベース <http://proteome.tmig.or.jp/2D/>, Swiss Institute of BioinformaticsのSWISS-2DPAGEデータベース <http://us.expasy.org/ch2d/>, University of SienaのSIENA-2DPAGEデータベース <http://www.bio-mol.unisi.it/2d/2d.html>, Max Planck Institute for Infection BiologyのProteome 2D-PAGEデータベース <http://www.mpiib-berlin.mpg.de/2D-PAGE/>, Purkyne Military Medical AcademyのPMMA-2DPAGEデータベース <http://www.pmma.pmfhk.cz/>, Danish Centre for Human Genome Researchの2D PAGEデータベース <http://proteomics.cancer.dk/>, German Heart InstituteのRAT HEART-2DPAGE データベース <http://www.mpiib-berlin.mpg.de/2D-PAGE/RAT-HEART/2d/>, German Heart InstituteのHUMAN HEART-2DPAGE データベース <http://userpage.chemie.fu-berlin.de/~pleiss/dhzb.html>, Harefield Hospital Heart Science CentreのHSC-2DPAGE データベース <http://www.harefield.nthames.nhs.uk/nhli/protein/>, Max Delbruck CenterのHP-2DPAGE データベース <http://www.mdc-berlin.de/~emu/heart/>, Oregon Health & Science UniversityのLENS-2DPAGE データベース <http://www.ohsu.edu/2d/2d.html>

ネット上で公開されているプロテオーム解析ツールのURL : ExPASy Proteomics tools サーバー <http://us.expasy.org/tools/>, Matrix ScienceのMascot Search サーバー http://www.matrixscience.com/search_form_select.html, UCSFのProteinProspector サーバー <http://prospector.ucsf.edu/>
