論文要旨

総合論文 ソフトイオン化質量分析による脂質の網羅的、特異的検出

北條俊章^a・田口 良^b

a 名古屋市立大学大学院薬学研究科分子生物薬学分野

(〒467-8603 愛知県名古屋市瑞穂区田辺通 3-1)

b 東京大学大学院医学系研究科メタボローム寄付講座 (〒113-0033 東京都文京区本郷 7-3-1)

エレクトロスプレーイオン化 (ESI) やマトリックス支援型レーザー脱離イオン化 (MALDI) などに代表されるソフトなイオン化法を用いた質量分析法は、微量成分を混合物中においても検出できるという点で、ある特定の分子群にフォーカスをおいた網羅的解析には非常に有効な手段となり得る。その対象分子として、我々は特にリン脂質に注目して研究を進めている。リン脂質は、生体膜の構成成分としてだけではなく、シグナリングのメディエーターや受容体のリガンドとしてなど、生体内で非常に重要な役割を担っているが、分子種レベルでの特異性など未だに不明な部分は多く、現在これらの解明が急がれている。このような分野において、質量分析 (MS) を適用していくには、多種多様に存在する機種の中から一番適したものを選択する必要がある。それだけではなく、液体クロマトグラフィーと組み合わせる方法 (LC-MS) やフラグメント情報を得るモード (MS/MS) などを用いていく必要がある。そこで本総説では、現在主に使用されている代表的な MS の特徴とともに、脂質研究の分野における MS の適用法を紹介する。

(連絡者：北條俊章) Vol.4, No.8, 319 (2004)

総合論文 脂質データベースからバイオインフォマティクスへ

大島美恵子

東北公益文科大学公益学部

(〒998-8580 酒田市飯森山 3-5-1)

生理活性脂質データベース “LIPID_{BANK} for Web” は、日本が世界にさきがけて作製した、脂質を化合物ごとに整理してデータベース化したもので、インターネットで公開されている (<http://lipidbank.jp/>)。扱うデータは、文字や数値情報以外に、化学構造式の挿画データやクロマトグラムなどのグラフィックデータおよび文献データが含まれる。本データベース作製の歴史とその使い方について解説し、今後この脂質データベースがどのような方向に進化して、バイオインフォマティクスの発展にどのように寄与してゆくかについて、生理活性脂質のメタボロームについての最近の話題とともにまとめる。

(連絡者：大島美恵子) Vol.4, No.8, 329 (2004)

脂質データベースからバイオインフォマティクスへ The Newly Developed Lipid Database “LIPID_{BANK} for Web”, its' Contribution to the Development of Bioinformatics

大島美恵子

東北公益文科大学公益学部
〒998-8580
酒田市飯森山 3-5-1

Mieko OSHIMA

Tohoku University of Community Service and Science
3-5-1, Iimoriyama, Sakata-shi, Yamagata-ken,
998-8580, JAPAN

Abstract: “LIPID_{BANK} for Web” is a database system offering information on lipids over the Internet (<http://lipidbank.jp/>). This database is developed by the members of the Japanese Conference of the Biochemistry of Lipids, and the first lipid database open to the world from Japan. The database is composed of three parts: 1) The factual data including name of the lipid, chemical and physical properties, information of biological activities, genetic information, etc., 2) The graphic data including chemical structure, chromatogram, IR or NMR spectra and other graphic data, and 3) The related reference data. This article describes the brief history of the “LIPID_{BANK} for Web” and how to use the database and its' future scope. In near future, the database will be linked to the other databases, such as, biological active lipid database, genomic-database, metabolome database, sugar-database, enzyme-database. etc. And it will contribute to the future development of bioinformatics.

Key words: Lipid-database, LIPID_{BANK} for Web, Bioinformatics, Bio-active lipids

1 はじめに

日本では近年、ようやくバイオインフォマティクスということばが社会のなかに定着しつつあるように思われる。生理活性脂質データベースを公開してから五年を経過したのを機に、脂質のデータベースをこれから新しいバイオインフォマティクスの世界へどのように展開してゆくかについて、ひとつの時代を終えた脂質生化学研究者の感慨を込めてまとめてみたい。

2 ポストゲノム時代とバイオインフォマティクス

ワトソンとクリックのDNA二重らせん構造についての論文から丁度50年目にあたる2003年4月は、ヒトゲノムの全構造が発表された記念すべき時であった。分析技術の革新と各国間の科学者達の競争とが相まって、まだ数年はかかると思われていたゲノム構造が一挙に解明に至ったからである。そして今、ポストゲノム時代とい

われて、ゲノムDNA解析の次に進めるべき研究がいくつも浮上している。RNA転写産物としてのtranscriptome、それ以後の代謝産物としてのmetabolomeの分析と、それらの情報の集積をもとに、将来は計算機上で生命を推測しようというバイオインフォマティクスが世界の注目を浴びるようになってきた。Metabolomeとは、生きた細胞の代謝産物のすべてをさす言葉であり、たんぱく質(proteome)のみならず、糖(glycome)、脂質(lipidome)その他の低分子物質などの代謝産物をも含んでいる。

さて、ここで少し歴史を振り返ると、DNAの複製とRNAへの転写機構とたんぱく合成機構がほぼ解明された1960年代の末以降は、若い研究者達が大量に分子生物学に流出していった時代であった。20世紀初頭から始まり急速に発展した生化学、すなわち代謝産物を化学的に追求してゆくという、いわゆるオールドファッションの生化学から、DNA複製とRNAへの転写とたんぱく合成というドグマに沿って研究を進める生物科学の新しい分野、分子生物学の擡頭である。

連絡者：大島 美恵子
E-mail : oshima@koeki-u.ac.jp

一方、いわゆる細胞生物学とよばれる分野では、1970年代に入ってから、がん抑制の研究、細胞の増殖と増殖抑制、細胞の周期、細胞膜の受容体分子など、培養細胞を使った研究が進み、がん遺伝子、がん抑制遺伝子、細胞増殖因子、細胞増殖抑制因子、細胞周期を司る物質の遺伝子などの発見が続いていた。そして、活性をもつたんぱく質をまず精製してアミノ酸配列を比較するという過去の手法から、遺伝子をクローニングし、すでに報告されたDNA配列のデータベースから相同性検索によってアミノ酸配列を整列 (alignment) させ比較検討することがはやりとなった。同時にDNA配列データの網羅的な集積がはじまり、日¹⁾、米²⁾、欧³⁾、などで現在も着実にその努力は継続されている。

そして1980年初頭には、例えば、ウイルス由来のがん遺伝子のDNA配列からアミノ酸配列を決定した研究が、全く関係のない他の研究グループが行ってきた血小板由来細胞増殖因子とのアミノ酸配列の類似からその機能を推察するという論文が出され⁴⁾、DNA配列からその機能を推察する時代に突入した。相同性検索-アミノ酸alignmentの手法は、今日研究室において日常的に利用されている。たとえば、小児の眼の腫瘍、網膜芽細胞腫 (retinoblastoma) では13番染色体長腕 (13q14) の短縮が知られていたが、がん抑制遺伝子としてRbがクローニングされると⁵⁾、がん細胞増殖抑制に働く分子量11万の核たんぱく (retinoblastoma protein, RB) の異常による網膜芽細胞腫の発症機構が明かになった⁶⁾。いまではRBは、核の中で他のがん抑制遺伝子 (p53) が産生するたんぱく質とクロストークして細胞のアポトーシスを誘導することや、転写因子E2Fとの相互作用によって活性が抑制される機構なども判明し、細胞周期や細胞分化に大きな役割をもつことが知られるようになった⁷⁾。このように、お互いに全く関係がないままに進んできた研究が、DNA配列の類似性から、ある日突然仲間としてつながってしまい、その機能が判明するという現象があちこちでしばしば起こるようになった。その時の驚きを我々のようなオールドファッションの生化学者は決して忘れない。ここに至って初めて、DNA配列がその機能にまで繋がるというバイオインフォマティクスの時代が始まったのであった。そして今は、DNA配列やアミノ酸配列のみならず、たんぱく質の立体構造や分子間相互作用など、多くの基本情報が世界各地のデータベースから豊富に供給されるような時代になってきた。

さて、細胞が生きているということは、たんぱく質を主体とする多くの酵素が各々の代謝活性を十分に発揮しているということである。そのためには、酵素が基質とうまく結合するような立体構造をとる必要があり、酵素

たんぱくを修飾してその構造を変化させる糖質や、疎水的な環境作りに役立つ脂質や、補酵素となる低分子物質の存在などたんぱく質以外の多くの物質の関与が必要である。それらが細胞の中で各々の役割を担ってはじめて生きた細胞の代謝が成り立つのである。これらの代謝物質 (metabolome) 情報のデータベースへの集積は、核酸塩基配列データベースにはるかに及ばない。というよりも、多くの構造の異なる物質の中間代謝産物をまとめるほど解析が進んでいない分野もあることによる。しかしながら、糖の分野では、糖鎖構造データを集めたCarbBankが1997年から公開され⁸⁾、また、脂質分野では、生理活性脂質データベース「LIPID_{BANK} for Web」が1999年に公開されて⁹⁾、ようやく生体内の基本的な物質、核酸、たんぱく質、糖質、脂質についての情報のデータベースが揃ったのであった。

3 脂質データベース作製の歴史

脂質データベースの必要性が学会ではじめて提唱されたのは公開のさらに10年以上前、林陽近畿大学教授 (現名誉教授) と重井医学研究所の渡辺清博氏からであった。当時まだパソコンはNECの88型から98型への移行のころであり、脂質の学会でデータベース作製に賛成する人は殆どいなかったが、それでも文部省科学研究費の支援を受けて少しずつデータ集積がなされてきた。脂質データベースを日本で作りたいという夢は、日本が、歴史的にみて特に脂質の研究者が多い国であり、日本がその分野をリードしてきたところが多かったという理由にもよる。

そして、1996年になって、科学技術振興事業団による生理活性脂質データベースプロジェクトへの支援 (プロジェクトリーダー大島美恵子) が決定したことから、科学技術振興事業団と国立国際医療センター研究所との共同研究というかたちで本格的にデータベース構築が開始された。このデータのもとになったのは、すでに10年以上にわたり集積されてきた脂質データであった。

この支援をうけて、日本脂質生化学研究会 (現、日本脂質生化学会) の中に「生理活性脂質データベース構築検討会」が発足し、データを集積する脂質項目ごとに構築責任者を置き、この責任者の指示のもとで、データ提供者が文献を基にした脂質化合物情報を書き込み、これを集積するという方法をとった。実際には、市販のデータベースソフトを用いてデータ提供者が入力したものを事務局に送り、これを事務局で書式変換し、サーバへ収録した。データ提供者数は、今までに140名以上にのぼる。集積を始めた当時は、まだインターネット上でギリシャ文字さえ表示することができず、物理化学でよく使

う記号などの殆どを画像として配信せざるをえなかった事を懐かしく思います。

さて、研究費がついた3年間で約6000件を目標にデータ集積を開始し、1997年度1500件、1998年度2000件、1999年度800件を集積し、さらに9月の完成までに1,000件以上を追加した。現在6500件に達している。1999年7月末には、サーバを日本脂質生化学研究会事務局のある東京大学医学部に移転して10月から「LIPID_{BANK} for Web」として本格的に一般公開を開始した。その後東京大学医学部内でサーバを異動し、現在は、<http://lipidbank.jp> でサーバに接続可能となっている。完成後は、僅かな継続予算(2002年まで)と、文科省科学研究費(2003年から)、および経済産業省管轄の財団法人(イメージ情報科学研究所、2003年度のみ)などから少額の予算を獲得して、現在も継続して集積修正が行われている。

4 「LIPID_{BANK} for Web」とはどのようなデータベースか

生理活性脂質データベース「LIPID_{BANK} for Web」は、脂質データを化合物ごとに整理しデータベース化したもので、インターネット上で広く公開することを目的としてシステム開発がなされており、特別なソフトや機器がなくても誰でも利用できる。生理活性脂質データベースで扱うデータには、文字・数値情報以外に、化学構造式の描画データ(Chem drawによる画像)及びクロマトグラムや電気泳動などのグラフィックデータが含まれている。

疎水基をもつ物質としての脂質は多様な化学構造をもっている。そこで28種類の主要な脂質化合物を選び、脂質項目ごとにデータが集積された。従って、28種類の個別の脂質データベースをひとつにまとめたものといってよい。28種類の化合物は16種類にグループ分けして収録されている。DERIVED LIPIDS(誘導脂質)のグループには、FATTY ACID(脂肪酸)、LONG CHAIN ALCOHOL(長鎖アルコール)、LONG CHAIN ALDEHYDE(長鎖アルデヒド)、LONG CHAIN BASE(長鎖塩基) and CERAMIDE(セラミド)が、FAT SOLUBLE VITAMIN(脂溶性ビタミン)には、CAROTENOID(カロテノイド)、COENZYME Q(補酵素Q)、VITAMIN A、VITAMIN D、VITAMIN E、VITAMIN K(ビタミンA、D、E、K)が、GLYCOLIPID(糖脂質)のグループには glycoSPHINGOLIPID(スフィンゴ糖脂質)、glycoGLYCEROLIPID(グリセロ糖脂質) and OTHERS(他の糖脂質)が、PHOSPHOLIPID(りん脂質)には、GLYCEROPHOSPHOLIPID(グリセロりん脂質)、PAF、SPHINGOPHOSPHO-

LIPID(スフィンゴりん脂質)がそれぞれグループ分けされて収録されている。

5 「LIPID_{BANK} for Web」の使い方

ここではデータを集積した28種類の脂質化合物のなかで、化学構造の簡単な脂肪酸と、疎水基以外に糖をもつ複雑な化学構造をもつスフィンゴ糖脂質のデータベースを中心にその操作法を説明する。

5.1 「LIPID_{BANK} for Web」へのアクセス

ブラウザアプリケーションを用いて<http://lipidbank.jp> にログインすると、「LIPID_{BANK} for Web」のホームページが表示される(Fig. 1)。

5.2 検索化合物の種類の選択

画面中央の[Next page →]のボタンを押すと脂質メニュー画面に進む(Fig. 2)。ブラウザの文字コードが日本語表示に設定されている場合は、ギリシャ文字等の特殊文字を正確に表示するため、「欧米(西ヨーロッパ言語)」に変更する(日本語表示のままではこれらの特殊文字は文字化けして表示される)。脂質メニュー画面(LIPID MENU)の28種類の脂質化合物の中から、検索したい化合物を選びクリックする。ここでは、脂肪酸(FATTY ACID)、あるいはスフィンゴ糖脂質(glycoSPHINGOLIPID)をクリックすると、検索画面へ移る。

5.3 検索画面

検索画面は、左右に2分割され、左画面に表示されている検索メニューにキーワードや数値を入力し「Search」ボタンを押すと検索が開始されて、しばらく待つと検索結果が右ウインドウに表示されるようになっている。

Fig. 1 “LIPID_{BANK} for Web”のホームページ
<http://lipidbank.jp/>

Fig. 2 脂質メニュー画面

この検索メニューウインドウで、検索条件をまったく入力せずに「Search」ボタンを押すと、収録されている全データの件数と化合物名のリストが表示される。2004年2月現在、755件の脂肪酸データ、580件のスフィンゴ糖脂質データが収録されている。

5.4 検索条件の設定方法

検索メニューには、キーワード検索 (KEY WORD)、分類検索 (CLASSIFICATION)、その他があるが、これは脂質項目によって異なっている。

脂肪酸の検索条件は、(KEY WORD) 検索、(CLASSIFICATION 1-3) 検索のほかに、側鎖構造についての (BRANCHED CHAIN) 検索と、不飽和結合の位置やヒドロキシル基の位置から検索する (POSITION) 検索、不飽和結合の数、炭素数、ヒドロキシル基の数、カルボキシル基の数などから検索する (NUMBER) 検索がある。

糖脂質の場合は、糖脂質の名称や含まれる臓器、組織名、生物活性などのキーワードを入力する (KEY WORD) 検索、中性、酸性、塩基性、両性などの糖脂質の種類から検索する (CLASSIFICATION) 検索、アシル基やステロールなどのアグリコンの種類から検索する (SPECIES OF AGLYCON) 検索、結合している単糖の数やシアル酸の数、フコースの数などから検索する (NUMBER) 検索の他に、IUPAC/IUBMBによる名称からの検索、糖鎖構造からの検索、脂肪酸の種類、長鎖塩基やグリセロールの構造などからも検索できるように

なっている。

検索条件がわからないときは、画面の「？」ボタンをクリックすると、検索条件設定のためのヘルプが別ウインドウで表示されるので参考にしてほしい。

検索が終了すると、画面上部には検索条件を満たす化合物の数が表示され、20件ずつ化合物名が表示される。慣用名がある場合は、左から慣用名が太字で、次いで系統名が表示される。それぞれの化合物名の左側に「VIEW」ボタンが配置されているので、データの詳細を見たい化合物のボタンを押し、詳細画面に移る。詳細画面は別ウインドウとして表示される。

また、検索条件を設定した後、キーワード検索項目の下部にある「CHECK the box to search ALL LIPID CLASSES」の□にカーソルを合わせてここをチェックすると、入力された検索条件で28化合物の全てを検索対象として検索を開始する。脂質メニュー画面で選択した種類の化合物のみならず、他の種類の化合物を含めて設定した検索条件を満たす化合物を検索したい時に利用できる。但し収録されている全化合物が検索の対象のため検索に時間がかかる。

検索条件が設定されているときは、ヒットした化合物名のリストの上に、検索されたデータ数と共に、どのような条件を満たす化合物であるのかがわかるように、設定した検索条件が表示されている。検索条件を満たす化合物がない場合は、「0 RECORDS FOUND」と表示されるので、再度左ウインドウで検索条件を入力後「Search」ボタンを押す。前回の検索条件をクリアして新たな検索条件を設定するには、「Clear」ボタンを押す。複数の検索条件を設定した場合は、それぞれの検索条件が“AND”条件で検索される。また、(NUMBER) 検索では、ある特定の数値に合致するもののみならず、数値の範囲を指定して検索ができる。

5.5 化合物データ画面の詳細表示

検索が終了し、表示された化合物のリストから、各化合物の詳細データを見るには、化合物名の左にある「VIEW」ボタンを押す。別ウインドウが開き、データ番号、データ提供者名に続いて、化合物の系統名、慣用名、略名および化学構造式が表示される (Fig. 3)。構造式はChem drawで描かれており、「Download」ボタンを押すと構造データをダウンロードすることが出来る。このサービスは現在多くの研究者に利用されている。

画面の下方を見てゆくと、生物活性 (BIOLOGICAL ACTIVITY)、物理化学的性質 (PHYSICAL AND CHEMICAL PROPERTIES)、UV, IR, NMR, MASSなどのスペクトルデータ (SPECTRAL DATA)、TLCや液体クロマトグラム、ガスクロマトグラムなどのクロ

EICOSANOIDS
LIPID ?

DATA No : XPR1000 INFORMANT : Shouzo Yamamoto

NAME : 7-[2(R)-(3(S)-Hydroxy-1(E)-octenyl)-5-oxo-3-cyclopenten-1(R)-yl]h
acid / (8R,12S,13E,15S)-15-Hydroxy-9-oxo-10,13-prostadienoic acid

COMMON NAME : PROSTAGLANDIN A1
SYMBOL : PGA1
FORMULA : C₂₀H₃₂O₄ MOL.WT : 336.466

Download If you want to use ChemDraw structure data shown on this page, click on the "Download" at the left.

BIOLOGICAL ACTIVITY

Prostaglandin A1 relaxes uterine muscle and shows hypotensive effects on rat and dog <<0002>>.

PHYSICAL AND CHEMICAL PROPERTIES

MELTING POINT : 42-44° C <<1005>>

Fig. 3 エイコサノイドの PROSTAGLANDIN A1 の詳細表示画面 (上部分のみ)

マトデータ (CHROMATOGRAM DATA), 所在・分布 (SOURCE), 化学合成法 (CHEMICAL SYNTHESIS), 代謝データ (METABOLISM), 遺伝情報 (GENETIC INFORMATION), 備考 (NOTE) および文献リストが表示される。

各項目内に書かれたスペクトル番号やクロマト番号をクリックすると、該当するスペクトルやクロマトグラムが表示される。文献番号をクリックすると、該当する文献にジャンプする。文献のタイトルの末尾には、PubMed 番号が表記されており、クリックすると PubMed の抄録ページへリンクするようになっている。

6 「LIPID_{BANK} for Web」と他の脂質データベースサイト

生理活性脂質データベース「LIPID_{BANK} for Web」の概要と利用方法を説明してきたが、日本から情報を発信することを意識して構築された生理活性脂質データベースは、現在国内外からのアクセス数が加速度的に増加しつつある。大学研究室からのアクセスが多い国内に比して、海外からは、すでに 50 カ国からのアクセスがあり、企業その他からのアクセス数が大学よりはるかに多い。

今、インターネットで世界の脂質データベースを検索すると、「LIPID_{BANK} for Web」の他に「LIPIDAT」と

いうサイトが存在する。これは脂質の物理的性質をまとめたデータベースで、すでに約 2 万件の入力がなされているが、相転位温度などを調べるには便利であるが、脂質全体を網羅したものではない。その他のサイトの殆どは入力件数が少ないためデータベースとしての実用性は低い。従って現段階では、「LIPID_{BANK} for Web」が世界で唯一の脂質全体を網羅する化合物データベースである。

「LIPID_{BANK} for Web」では、すべての脂質化合物にデータ提供者名が記入されていて、後で修正される場合に、データ提供者のみならずデータ集積責任者にも連絡できるシステムができており、今後も責任あるデータの集積が可能である。内容を改良してゆくことが作製者としての責任であると思う一方で、誰でも無料でどこからでもアクセスできることから、悪意を持ってアクセスがなされる場合の防衛の方法や、維持管理など頭の痛い問題はまだ残っている。今後さまざまなサイトへのリンクが進むと、バイオ情報としてさらに充実してくるであろう。

7 脂質のダイナミズムとバイオインフォマティクス

ポストゲノム時代といわれる最近になって、強い生理活性をもつ脂質分子や環境づくりのための脂質の重要性が注目されるようになってきた¹⁰⁾。脂質分子は、その疎水性部分の分子多様性 (分子種) という特徴をもっており、脂質分子の生体における活性は分子種により多様に変化する。ありふれたレシチンの 1 分子種である dipalmitoyl phosphatidylcholine (DPPC) は、肺 surfactant の主成分として知られており、これが欠乏することが新生児呼吸窮迫症候群の病因であることからみても、分子種とその生理作用の関係を情報として集めることが重要である。最近田口らは、りん脂質の網羅的解析の新しい方法を報告している^{11, 12)}。このような方法を利用することにより、細胞内りん脂質の時間変動が追跡できるようになり、脂質 metabolome の解析も進むであろう。

脂質として最も単純な構造をもつ脂肪酸をとり上げてみても、分子種により生物体内で誘導される代謝産物 (脂質 metabolome) の生理活性は全く異なっている。たとえば、炭素数 20 (eicosa) のポリエン脂肪酸から生体内で誘導される物質はエイコサノイドと呼ばれ、化学構造と活性の全く異なる prostanoid, leukotriene, lipoxin のグループがある。一方、二重結合数の違いで、8, 11, 14 eicosa トリエン酸からできる第 1 系列、5, 8, 11, 14 eicosa テトラエン酸 (arachidonic acid) からの第 2 系列、5, 8, 11, 14, 17 eicosa ペンタエン酸からの第 3 系列エイコサノイドは、それぞれ多彩な生理活性作用を持

つ¹³⁾。また, prostanoid からは, さらに prostaglandin, prostacyclin, thromboxane のように異なる生理作用を持つ代謝産物ができる。現在「LIPID_{BANK} for Web」には, エイコサノイド (EICOSANOID) として 249 件, そのなかに prostaglandin として化学構造の異なる 60 件が登録されている。それらの生理作用はまだ解明されていないものも多いため, データベース中の「BIOLOGICAL ACTIVITY」欄への記載は充分ではない。また最近, arachidonic acid は, 上にのべたような第2系列のエイコサノイドを産生するほかに, アラキドン酸エタノールアミド (anandamide) や 2-アラキドニルグリセロールのような誘導体が強い生理活性をもつことが報告されるようになってきた^{14, 15)}。これらの脂質の構造と生理活性については, 今後の研究の進歩とともにデータベースへの集積が必要である。化学構造を主体としたデータベースとして開発した「LIPID_{BANK} for Web」から, 生きた細胞の脂質 metabolome とその代謝, 生体における生理活性の発現のデータを含んだ Systematic Bioactive Lipid Database (SBLD) へと今後発展させていくことが重要である^{16, 17)}。

8 おわりに

生命科学の発展に伴って, 脂質のみならずあらゆる生体物質のデータ項目は日毎に複雑化し, それらの情報量は増大し続けている。既存のデータベースが継続され内容が充実していくと共に, 新しい観点からのさまざまなデータベースがうまれてくる。現在リンクというかたちでデータベース同士が繋がりがつつあるが, 望む情報を検索することは容易ではない。検索に便利な新しいソフトウェアが作られる事と, 既存のデータベースをあまり変更することなく, 何らかのインターフェースを付加して相互に運用できるような基盤技術の開発が望まれている。そのためには, 現在それぞれのデータベースによって異なる化合物の表示方法などを統一してゆく必要もあるだろう。

バイオインフォマティクスの最終的の目的は, 生命科学のデータベース群に解析ソフトウェアを組み合わせ, 計算機上で生きた生物のふるまいを予測して, モデル生物ならぬモデル・デジタル生物をつくることであるという¹⁸⁾。研究者が, *in vitro* 実験と *in vivo* 実験を組み合わせ出た情報を, データベースとして集積してゆくこ

とにより, それらのデータベース群から計算機上 (*in silico*) で生命を推測する時代が近づいてきている。しかし現在のところ, 進化したデータベースは, 核酸と一部のたんぱく質に限られており, バイオベンチャーなどが注目するバイオインフォマティクスも, 遺伝子発現からの疾患の診断や, SNP (Single Nucleotide polymorphism) をもとにした薬のデザインなどに限られている¹⁹⁾。バイオベンチャー企業の研究者などで, metabolome は, 核酸情報から産生されるたんぱく質 (proteome) と解釈している人もあり, 脂質や糖質などを含めたすべての metabolome が生体内で機能を発現し, 代謝調節されることが「生命」であるという視野がまだ欠けている。本当の生命体を知るには, まだまだ膨大な基礎データの地道な蓄積が必須である。何時の日にか, *in silico* で生命を考察できる日がくることを祈りつつ, 稿を終わりたい。

文 献

- 1) DDBJ, (Data Bank of Japan, <http://www.ddbj.nig.ac.jp>)
- 2) GenBank, (National Center for Biotechnology Information (NCBI))
- 3) EMBL, European Bioinformatics Institute (EMBL)
- 4) R.F. Doolittle, *et al.*, *Science*, 221 275 (1983).
- 5) S.H. Friend, *et al.*, *Nature*, 323, 643 (1986).
- 6) P. White, *et al.*, *Nature*, 334, 124 (1988).
- 7) S.P. Chellapan, *et al.*, *Cell*, 65, 1053 (1991).
- 8) K. Nagai & I. Ishizuka, *Trends Glycosci. Glycotechnol.*, 10, 257 (1998).
- 9) K. Watanabe, E. Yasugi & M. Oshima, *Trends Glycosci. Glycotechnol.*, 12, 175 (2000).
- 10) L. Chong & J. Marx, *Science*, 294, 1861 (2001).
- 11) R. Taguchi, J. Takeuchi & M. Ishida, *J. Mass Spectrom.*, 35, 953 (2000).
- 12) M. Ishida, M. Imagawa & R. Taguchi, *J. Mass Spectrom.*, 49, 150 (2001).
- 13) C.D. Funk, *Science*, 294, 1871 (2001).
- 14) W.A. Devane, *et al.*, *Science*, 258, 1946 (1992).
- 15) T. Sugiura, *et al.*, *Biochem. Biophys. Res. Commun.*, 215, 89 (1995).
- 16) 八杉悦子, 渡辺清博, あなたにも役立つバイオインフォマティクス (菅原秀明編, 共立出版) pp 100 (2002).
- 17) 大島美恵子, 東北公益文科大学総合研究論集, vol. 2, 19 (2001).
- 18) 菅原秀明, あなたにも役立つバイオインフォマティクス, 菅原秀明編, 共立出版, pp 4 (2002).
- 19) 浅井聡, バイオベンチャー, vol. 2, 33 (2002).