

特集②

バイオインフォマティクスの技術動向



情報・通信ユニット 客員研究官 乃木 篤*、香月祥太郎

1. はじめに

ヒトゲノム計画は国際協調のもと予想を上回るスピードで進展し、2000年6月にドラフトシーケンスが決定した。現在は完全解読に向けた最終作業の段階にある。日本の貢献は6~7%と予算規模に見合ったものであるが、戦略の欠如による出遅れが指摘されるなどの課題も残った。しかし、2002年8月の理化学研究所を中心とするマウス完全長cDNA^(注)の解読終了とイネゲノムの解読、12月の日米共同によるホヤゲノム解読等はゲノム科学の新しい進展に大きな足跡を残した。これらヒトゲノムやマウスcDNAの配列が明らかになったことで、今後は配列の解析から網羅的、体系的な機能の解析へと焦点が移っていくとみられる。日本はイネゲノムや完全長cDNAライブラリなど基盤技術

での強みがあり、学術的にも応用分野でも世界をリードするチャンスは十分にある。

近年、解析装置の高速化・自動化により、産出される実験データの量は膨大になっており、情報技術の役割はより一層大きなものとなりつつある。従来のデータ処理の効率化だけでなく、ゲノムを活用した網羅的な解析の複雑さに対応するための数学や情報科学の知識が必要となる。バイオインフォマティクスはこのような課題に体系的に取り組むために生物科学と情報科学の境界に生まれた新しい学問分野である。

本誌2001年12月の特集においてバイオインフォマティクスをとりあげ、その動向を生命科学の観点から概観した。本稿ではその後の動向を含めてゲノムや生命をシ

ステム的に理解するための方法論および今日的課題である人材育成策について述べる。

(注) 完全長cDNA：cDNAとは、ゲノムDNAの中から不要な配列を除き、タンパク質をコードする配列のみに整理された遺伝情報物質であるmRNA（メッセンジャーRNA）を鋳型にして作られたDNAのこと。完全長cDNAは、断片cDNAと異なり、タンパク質を合成するための設計情報をすべて有しているため、タンパク質を合成することができる。この完全長cDNAを効率的に合成するためには、非常に高い技術を必要とし、わが国が世界に先んじている。

2. バイオインフォマティクスにおける情報科学的手法

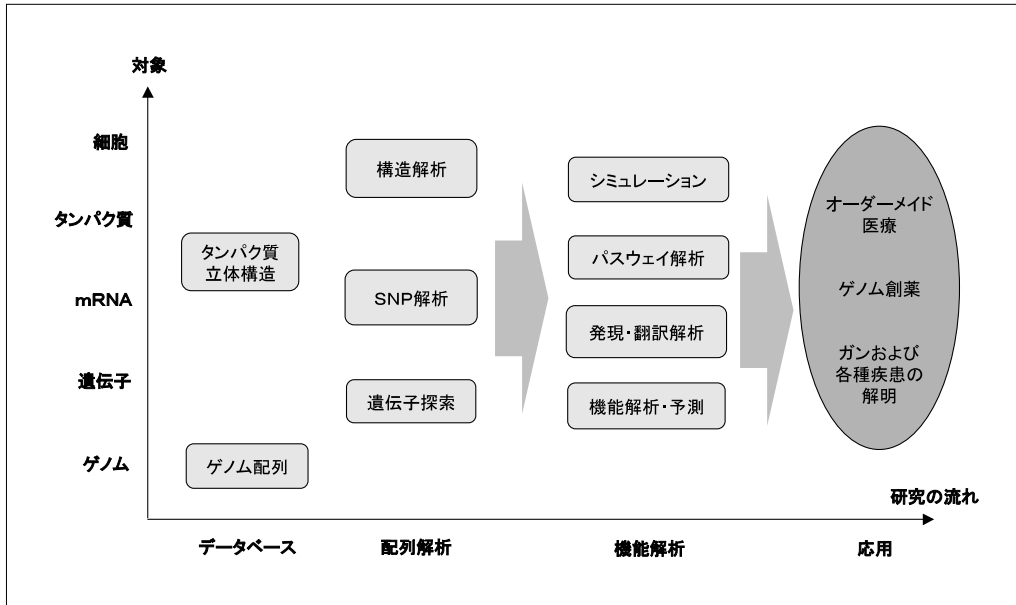
ゲノム研究における諸問題を情報科学的にとらえて定式化を行うと、その多くが現在の計算機の処理能力をはるかに超える非常に困難な問題となる。これはゲノム配列が非常に長いこと、および解析の対象とする問題の性質が組み合わせ的であることによる。したがって、近似解法や発見的手法が必要となる。そのための実用的アルゴリズムの開発はバイオインフォマティクスの主要な研究分野であ

り、その成果はゲノム配列の決定に大きく貢献してきた。米セラ・ジェノミクス社が用いた全ゲノムショットガン法という配列決定手法は、数千万個のランダムな配列断片をジグソーパズルのようにつなぎあわせるものであり、当初は手法そのものが疑問視されていた。しかし、同社は最高性能の計算機と独自のアルゴリズムを駆使して驚異的な速さで配列決定を成し遂げ、情報技術の威力と有効

性を強く印象付けた。

近似解法は必ずしも最適解を導くものではないため、解析結果の生物学的な評価とそれに基づくアルゴリズムやパラメータの改良が必要となる。バイオインフォマティクスの特徴は、実験で確かめべき膨大な探索空間を絞り、体系的かつ網羅的な解析を可能にし、個々の遺伝子の解析では見えてこなかった生命現象の全体像を明らかにできる点にあるといえる。

図表1 バイオインフォマティクス研究の流れ



(科学技術動向センターで作成)

これまでバイオインフォマティクスは、ゲノムデータベースの構築やデータ解析のためのツール開発とその利用技術が中心であった。今後、ゲノム研究において遺伝子発見や機能予測などで、より高度な情報処理能力が活用されるようになってくる。ゲノム解読からスタートした研究は、図表1に示すとおり、DNAから遺伝子、タンパク質、さらには個体へと対象の複雑性・多様性は増大し、解析レベルは配列から機能、作用へと、研究の焦点が移っていくことが明らかである。バイオインフォマティクスにおいてもこの流れに

沿った研究開発を行う必要がある。

2-1 データベース

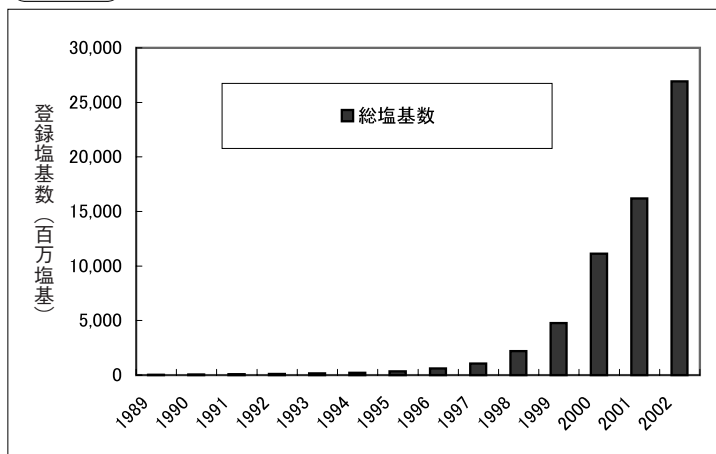
ゲノム研究によって生み出される大量のデータの多くは、公共のデータベースに登録されインターネット上で公開されている。これらのデータはバイオインフォマティクスのあらゆる研究の基点となる研究基盤として位置づけられる。公共のゲノムデータベース (GenBank (米)、EMBL (欧)、DDBJ (日)) に登録されるDNA配列データは指数的に成長してい

る (図表2)。これまでのデータベースはDNA配列やタンパク質構造データなどの一次データの集積であったが、ゲノムプロジェクトの進展とともに多様化が進み、一塩基多型 (SNP) や変異の情報、遺伝子やタンパク質の相互作用の情報など、従来とは異なる性質をもったデータがとりこまれるようになってきた。これらの多様なデータを駆使した解析を行う上で、データベースや情報技術の高度化がますます必要となっている。

(1) データベースの注釈づけ

多様なデータから新たな知見を発見するための一連のプロセスをデータマイニングと呼び、バイオインフォマティクスにおける重要な技術である。しかし、実験によって得られたデータが、データベースに登録されただけの状態で放置され、データマイニング技術を有効に適用できないことが問題となっている。そこで研究に使えるように解釈情報を付与する必要がある。これをアノテーション (注釈付け) と呼ぶ。現在アノテーションはコンピュータによる自動化が進められている。しかしコンピュータだけでは最適なアノテーシ

図表2 登録塩基数の推移



(日本DNAデータバンクホームページをもとに科学技術動向センターで作成)

オンが得られないので、人による確認作業が必要となる。理化学研究所で決定されたマウス完全長cDNA配列のアノテーションに関する会議が2000年8月に理化学研究所（筑波）において開かれ、各国の研究者による2週間にわたる白熱した議論の中でアノテーション作業が行われた。アノテーションには個人の解釈を超えた、各機関の研究者による確認の協調作業が不可欠となる。

(2) 概念と語彙の統合

遺伝子やタンパク質は、それを対象とする研究の過程で慣習的に付けられた名前をもっている。したがって異なる研究分野において同じタンパク質がいくつもの名前を持つこともあり、統合が求められている。またヒトをはじめとするモデル生物の相次ぐ解読で、生物界全体を対象とする研究が増えている。これまでモデル生物（大腸菌、酵母、マウスなど）ごとにデータベースが作られてきたが、使われる用語や分子の呼び名に統一の基準がなく利用上の非効率が問題とされてきた。

このような問題を解決するため、語彙の不統一を整理し、体系的に記述する試みがはじまっており、オントロジーと呼ばれている。オントロジーとは、概念に首尾一貫した用語と定義をあたえることである。遺伝子を対象とした遺伝子オントロジー、分子や細胞の相互作用を対象としたインタラクションオントロジー、シグナル伝達系を対象としたシグナルオントロジーなど、研究分野ごとにオントロジーの構築が進められている。このような語彙の統合と体系的分類が進めば、データベース間の相互運用性が高まり、異種のデータベースが互いに参照しあえるようになるというメリットがある。

2 - 2

ホモロジー解析

遺伝子やタンパク質の機能を、生物種間のアミノ酸配列の類似性に基づいて調べる手法をホモロジー解析と呼ぶ。ホモロジー解析は遺伝子およびタンパク質の機能を調べる有力な手段となる。タンパク質は、近縁の生物種ほど構造的、機能的に類似している。したがって、類似した配列の遺伝子が別の生物種に見つかり、その機能が同定されていれば、当該の遺伝子の機能的性質を予測する手がかりとなる。

ホモロジー解析において、対象がゲノム配列のように非常に長い配列の場合、配列が一致する領域を高速に見つけることが重要となる。このための代表的な解析ソフトウェアとしてFASTA (Fast Alignment) およびBLAST (Basic Local Alignment Search Tool) がある。FASTAとBLASTを比べるとBLASTのほうが処理速度は速いため、より一般的に使われている。FASTAは検索感度に優れているので、より詳しい解析に用いられる。

BLASTやFASTAを用いても大規模なゲノムデータベースの検索には時間がかかる。この対策としては検索処理の並列化が有効である。並列化には、SMP (Symmetric MultiProcessor) と呼ばれるマルチプロセッサ方式、PCクラスタ、グリッドなどの方式が考えられる。マルチプロセッサ方式は一つのコンピュータシステム上に複数のCPUを搭載して処理能力を高める方式である。またPCクラスタは通常のパソコンを複数台ネットワーク接続することにより低コストで並列計算機を実現するものである。広く使われているNCBI (National Center for Biotechnology Information: 米国の統合生物

情報データベース) 提供のBLASTプログラムではマルチプロセッサ型の並列計算に対応しているが、マルチプロセッサ方式はハードウェアが高価であるという難点がある。そこで、低コストなPCクラスタに対応したBLASTが実用化されている。

また、グリッドに代表されるような共用計算環境の利用も考えられる。グリッドは近年、注目を集めている情報技術の一つである。国内でのバイオ関連グリッドの動きとして、大阪大学のバイオグリッドプロジェクトや並列生物情報処理イニシアティブ (IPAB) のOBIGrid (後述) などがある。OBIGridでは、バイオインフォマティクスで必要となる最新データベースやアプリケーションが利用できる環境が提供される。これらを利用すれば、個々の研究室では高価な計算設備を持つ必要はなくなる。こうした国内での取り組みについてはより一層の充実を図っていく必要がある。

2 - 3

タンパク質構造解析

タンパク質はゲノム情報に従って生体内で合成される最終の産物であり、生命活動の基本的な機能分子である。タンパク質のアミノ酸配列と立体構造の関係を理解することはバイオインフォマティクスの大きな目標の一つである。ここ数年の立体構造のデータ数は著しい伸びを示しており、10年前の10倍以上となっている。

アミノ酸の配列からタンパク質の立体構造は一意に決定される。すなわち、原理的には配列のみから立体構造の予測が可能である。タンパク質が細胞内で合成されると、数ミリ秒から数秒の時間をかけてエネルギー的に最も安定な立体構造に折りたたまれる。この過程をフォールディングという。し

かしたンパク質の立体構造を分子動力学から予測しようとする、膨大な計算量が必要となるため十分な精度が得られていない。

実用的な構造予測の手法にホモロジーモデリング法がある。これは、配列が類似しているタンパク質は構造も似ているという性質を利用するものである。構造が既知のタンパク質と配列レベルで30%以上一致しているとき、その構造を部分的に修正することで、求めるタンパク質立体構造を予測することができる。

タンパク質の構造予測が可能になった今日、タンパク質研究の焦点は機能解析に移りつつある。機能解析によるタンパク質の機能特性の予測は創薬開発につながる重要なものである。ここでのバイオインフォマティクスの役割は極めて大きい。

2 - 4

遺伝子ネットワーク

ゲノムの全配列やタンパク質の立体構造がわかったとしても生命を理解したことにはならない。次に重要となるのは、どの遺伝子がどのような相互作用を行うかという遺伝子ネットワークの解明である。そのためには、これまでに蓄積された生命科学の膨大な知識を相互作用という観点から体系化し、計算機で扱えるようにしていくことが必要となる。バイオインフォマティクスの新たな適用領域として注目される。

ゲノムの機能解析において、細胞中で発現している全遺伝子をトランススクリプトーム、また遺伝子から生成される全タンパク質をプロテオームという。機能未知の遺伝子が、いつ、どのように発現するかを知ることは、その機能を考える大きな手がかりとなる。

プロテオーム解析で得られたデ

ータは、他の研究者にとっても有用な情報となりうる。スイスバイオインフォマティクス研究所では実験試料に含まれるタンパク質の分析結果である電気泳動ゲル画像のデータベース化を推進している。しかしプロテオーム情報には特許性があり、創薬など産業利用にも直結するため、公共利用は制限される方向にある。したがって国内でのプロテオーム関連データベースの整備が急がれる。

さらに細胞をゲノムにコードされた遺伝子が作るダイナミックなシステムとして理解することが、バイオインフォマティクスの主要な研究課題となりつつある。京都大学のKEGG (Kyoto Encyclopedia of Genes and Genomes) システムでは、遺伝子ネットワークの研究成果をデータベースとして公開している。遺伝子ネットワークの解明は遺伝子機能の記述を中心とした従来の生物研究の方法論では不十分であり、情報科学との融合が必須となっている。

3. 情報技術からみたバイオインフォマティクスの課題

バイオインフォマティクスで使われる解析ツール(ソフトウェア)は大半が海外技術を導入したものに依存しており、ソースコードが公開されてないもの、インターネット上のサービスのみで提供されるものなど、機能がブラックボックスとなっている。また商用では、海外のソフトウェアを組み合わせ提供されるものがほとんどである。

図表3は米国の大学の標準的な教科書で紹介されている、評価の定まったゲノム配列解析ソフトウェアを示した。これを見るとわかるとおり、米国以外の国で開発されたソフトウェアは極端に少ない。タンパク質立体構造予測や遺伝子アノテーションなどアプリケーションソフトにおいて国産の優

れたソフトウェアがいくつか存在するものの、図表3に示したような基盤分野で広く使われるソフトウェアはなかなかでてこない。この理由として、複数の研究者へのヒアリングから以下のような意見が得られた。

- バイオインフォマティクス研究者の数は、我が国と米国とで大きな開きがある。米国では研究動向の変化に応じて研究者の転身、転職が迅速に行われるなど人材の流動性があり、バイオインフォマティクスの研究者が短期間に確保された。バイオインフォマティクス研究者の育成が喫緊の課題である。
- 情報系と生物系の研究領域の

融合が進んでいない。BLASTの開発者はもともと数学を専門としていた。異なる分野の研究者同士が共同で研究できる環境が必要である。

- バイオインフォマティクスに必要なソフトウェアが研究レベルで動くプログラムにとどまっておらず流通しない。また、独創性のあるアルゴリズムが論文になった時点で、その仕事に区切りをつけてしまい、ソフトウェア開発にまでつながらないケースがある。また、ソフトウェアの流通のためには、マニュアルやインストールツール、配布メディアなどのパッケージ化が必要となるが個々の研究者がそこまでの労力をかける余裕がない。

図表3 主なゲノム配列解析ソフトウェア

| ソフトウェア | 考案者、作者 | 特徴 |
|------------------------|----------------------------|--|
| ホモロジー検索 | | |
| FASTA | Pearson 1988 (米・Virginia大) | 検出感度はBLASTより優れている。 |
| BLAST | Altschul 1990 (米・NCBI) | FASTAよりも高速。現在もっともよく使われている。 |
| PSI-BLAST | Altschul 1997 (米) | タンパク質ファミリーを見つけるための対話版BLAST。SSEARCHよりも検出感度は高い。 |
| SEG | Wootton, Federhen 1993 (米) | 低複雑度領域や反復配列を除いて類似性の比較精度を高める。 |
| SSEARCH | Pearson 1991 (米) | 動的計画法により最適なアラインメントを求める。非常に遅い。 |
| Bayes block aligner | Zhu 1998 (米) | ベイズ統計を利用するものでSSEARCHよりも遅いが、より遠縁の配列を見つけることができる。 |
| PROBE | Neuwald 1997 (米) | PSI-BLASTと同種の機能をもつ。ベイズ統計を用いて非対話的に最も有意な配列集合をみつける。 |
| 多重配列比較 (複数の配列の比較) | | |
| ClustalW | Higgins, Sharp 1988 (英) | 累進法により複数の配列の整列配置を求める。多重配列アラインメントでは最もよく使われている。 |
| PILEUP | Fen, Doolittle 1987 (米) | 累進法により複数の配列の整列配置を求める。配列の比較にはニードルマン・ヴァンシュ法を用いる。 |
| MSA | Lipman 1989 (米) | 多次元動的計画法により最適なアラインメントを求める。 |
| PRRP | 後藤 1996 (日・産総研CBRC) | 系統樹を求めて反復学習によりアラインメントを改善する。 |
| SAGA | Notredame, Higgins (仏) | 遺伝的アルゴリズムによりスコアの高いものを選ぶ。 |
| HMMER | Eddy 1998 (米) | 隠れマルコフモデルを用いる。 |
| プロファイル検索 (特徴的なパターンの検索) | | |
| ProfileSearch | Gribskov 1996 (米) | 配列パターン (モチーフ) の検索を行う。 |
| MAST | Bailey, Gribskov 1997 (米) | ギャップを含まない配列ブロックに合致する配列を検索する。 |
| 遺伝子発見 | | |
| RepeatMasker | Smit (米・Washington大) | 反復配列を検出して取り除き、遺伝子を発見しやすくする。 |
| TWINSKAN | Korf (米・Washington大) | 異なる生物種のゲノムを比較し、配列の保存領域から遺伝子を発見する。配列比較法と ab initio 法のハイブリッド法。 |

(参考文献¹⁾をもとに科学技術動向センターで作成)

これらの問題の解決には、バイオインフォマティクスの人材育成のための環境整備やソフトウェアの実用性を評価し流通させる体制の確立が有効と考えられる。これらを国として支援できる枠組みを検討すべきであろう。

4. 国内におけるバイオインフォマティクスの事例

これまで述べたように国内でのバイオインフォマティクス研究の遅れが指摘される一方で、一部には欧米に並ぶレベルの研究事例も出てきている。

4-1

タンパク質立体構造予測プログラム

現在、タンパク質の立体構造予測は創薬の観点から注目されている。北里大学の梅山教授らが開発したFAMSはタンパク質構造予測

の国際コンテストであるCASP (the Critical Assessment of Techniques for Protein Structure Prediction) の2000年の大会で優秀な成績を収めた。この分野では米国製をはじめとする海外のプログラムが広く使われているが、FAMSはそれらに勝る好成績を修めた。他のプログラムが部分構造から全体を構築するボトムアップアプローチであるのに対し、FAMSはまずおおまかな全体構造をとらえた後、部分構造を予測する。これは人間が立体構造を認識する際にまず全体

に着目することをアルゴリズムに取り入れたものであるという。こうした世界レベルのコンテストに参加することは研究開発の促進につながるものと期待される。

4-2

バイオグリッド

近年、注目を集めている情報技術の一つにグリッドがある。グリッドとは電気の送電網のことで、「電気のようにコンセントにつなぐだけで好きなだけの計算パワー

やディスクを使うことができるようなコンピュータシステム」をイメージしている（科学技術動向2002年9月号「グリッド技術の動向」参照）。

バイオグリッドは、バイオインフォマティクスに必要な設計パスをグリッド技術で共有しようというものである。

OBIGrid（Open Bioinformatics Grid）は、文科省科研特定領域研究ゲノム情報科学と並列情報処理イニシアティブ（IPAB）が母体となり、グリッドに接続するだけで最新データベースやアプリケーションが利用できる環境の構築を目指している。OBIGridは、ゲノム解析の拠点となりうる可能性を持っており、まずバイオインフォ

マティクスへの研究者の参入障壁を下げることができる。即ち、バイオインフォマティクスで必要となる様々なデータベースやアプリケーションの環境を自力では揃えられない研究者を多く引き寄せる効果が期待できる。またグリッドは実験データ開示の場として機能しうるものであり、LAN環境と同様の手軽さでファイルにアクセスできるため、そのままでは死蔵されてしまうような実験データから新たな発見をもたらすことが期待される。

4 - 3

産学連携による商用化

バイオインフォマティクス研究の成果として公的資金により開発されたソフトウェアの商用化による市場への流通の動きもある。理化学研究所のFANTOM（Functional Annotation of Mouse）プロジェクトで開発されたcDNA機能アノテーションシステムが2002年に製品化された。公的研究機関と企業との共同研究プロジェクトの成果が汎用的に利用可能なパッケージ化につながった例として注目される。

これらの事例は日本のバイオインフォマティクス研究の可能性を示すものとして注目される。

5. 人材育成の取り組み

5 - 1

バイオインフォマティクス研究に求められる人材

国内のバイオインフォマティクス研究が遅れた原因として、複数の研究者へのヒアリングによれば、人材の不足を指摘する声が圧倒的に多い。ゲノムやDNA解析からタンパク質の構造・機能解析までを強力に推進する技術としてバイオインフォマティクスは不可欠であるが、情報系技術と生物系技術の融合領域に属するだけに、わが国では人材の育成が大きな課題である。二つの言葉（生物・医科学系と情報系）を話せる人材が必要とされている。また、学部・学科を超えた横のコミュニケーションや交流の不足も原因として指摘されている。理化学研究所のFANTOMプロジェクトでは、生物系、医科学系、情報系の3分野の研究者がチームを組んで密接な協業の中で研究を進めていくこと

で大きな成果が得られたという。情報系と生物系の研究者がただデータのやりとりだけで役割を区切ってしまうと、実験研究と解析研究の双方の研究成果を互いに十分に生かすことが難しくなる。

求められる情報系人材としては、大量のデータを扱うことから、DB構造やプログラミングに長けている人が望ましい。また生物系人材としては、高度なIT知識までは必須ではないが、解析プログラムの仕組みを理解して使いこなすことができ、実験上の要求に応じてプログラムを改変したりできる能力を習得していることが望ましい。

人材育成の方針としては、豊富な人材を抱えるIT分野から生物系への参入を促すことが人材の量的な確保において有効であると考えられる。しかしながら、すぐれたバイオインフォマティクス研究には深い生物学的・医学的知識が欠かせないことから、生物・医学系の研究者が情報技術を学ぶこと

によってバイオインフォマティクスの研究に進むという流れを作っていくことも、研究人材の質的な確保には欠かせないと考えられる。

欧米、特に米国では情報科学を修得した研究者が新たな研究分野としてゲノム科学等に進出し、遺伝子研究者らと協業することによってすぐれたバイオインフォマティクスを確立してきた。そういう環境のもとでは、論理的・数学的に裏付けられた解析手法が遺伝子発現やタンパク質の構造・機能実験に適用され、試行錯誤を経て実用的解析技術として確立されてきている。

5 - 2

人材育成のための方策

学術的にもバイオインフォマティクスは新しい領域であり、大きな期待が寄せられている。一方でゲノム科学研究やタンパク質研究の進歩が急激なため、民間を含む多くの研究機関等で早急な人材の

供給が求められている。そのためにはバイオインフォマティクス人材の育成が急務であり、以下のような環境を整備することが有効である。

- (1)情報科学系と生物科学・医学系の大学院生に対して両分野の学問を学ぶ選択肢を与え、融合領域の研究に取り組めるようにすること。
- (2)情報系の研究者・技術者と生物系医科学系の研究者が共同で研究できる環境をつくること。これによりDNAやタンパク質の実験研究の進捗に応じて必要な解析アルゴリズムを案出しコンピュータで解析するなどの、実験とバイオインフォマティクスが一体化した研究が進み、技術の開発と実地的な人材の育成が図られる。
- (3)バイオインフォマティクスで用いられる各種のコンピュータ解析アルゴリズムやソフトウェアツールを考案した人の

成果と技術貢献を評価する仕組みを作ること。

バイオインフォマティクスの人材育成については、上述の環境条件を整備するためには大学院のカリキュラムの弾力的運用や変更が有効であろう。また、既存の専門分野の研究者と同様のキャリアパスが得られるよう教官のポストを設置し、その枠組みの中で研究成果の評価や人材配置を行う必要がある。さらに民間も参画できるような研究体制の整備も望まれる。これらの施策によって関連学術分野からの研究者の参入も期待できる。特に、数学、統計学、数理工学などの研究分野はバイオインフォマティクスの理論的基盤の発展に大きく貢献するものである。上記条件を満たす体制に対して国の資金的な支援も必要であろう。

2001年に慶応大学は「IT主導型バイオサイエンス」を目指して山形県鶴岡市に先端生命科学研究所を設立した。ここでは情報系や

生物系の教授や若手研究者が学生とともに互いの学問分野を学びながら両分野の融合研究が行えるような環境が整備されている。また、2002年に大阪大学に新設された生命機能研究科は医学・生命科学系、生物工学、生物学、物理学など大阪大学の生命科学関連分野の研究室を結集して作られた学際的な研究科である。

また、国においては、平成13年度から科学技術振興調整費によりバイオインフォマティクス分野においてプロフェッショナルを早期に育成する人材養成ユニットを設置している。平成14年度までに東京大学、京都大学、産業技術研究所、慶應義塾大学、奈良先端技術大学院大学に計6ユニット設置され、各人材養成ユニットにおいて融合的な人材養成が行われている。このような動きが他の大学や研究機関にも広がっていくことによって、バイオインフォマティクスのような学際研究の活性化が促進されるものと期待される。

6. まとめ

バイオインフォマティクスの活躍の場となるゲノム研究はスピードが要求される分野であり、各国で研究資源の急速な投入が行われている。我が国においてもゲノム研究への投資は増大しているが、創薬やオーダーメイド医療に向けてますます発展していくであろうゲノム研究を支えるバイオインフォマティクス人材の育成はまだ十分ではない。

ゲノムやタンパク質研究は比較的単純な配列解析の段階から、機能解析、さらに応用へと新しい段階へ動きつつある。これに従いバイオインフォマティクスへの要求もより高度になり、また研究における重要性も高くなっていくと思われる。この要求を満たすために

は情報系・生物系双方の学術的・技術的知見を備えた人材の育成を急がなくてはならない。5章に述べたようにその動きはすでに始まっているが、日本全体で見た場合には十分とはいえないであろう。

人材育成は、短期的には情報系研究者・技術者と生物・医学系研究者との協業を促進し、情報系、生物・医学系双方が個々の分野の知識・ノウハウを共有しながら研究する環境を整備することが有効であると考えられる。現在進行中のバイオインフォマティクス関連のプロジェクトにおいてもこの観点を重視するとともに、情報系研究者・技術者と生物・医学系研究者の相互参入を促していくべきである。

謝 辞

本稿の執筆にあたり、理化学研究所ゲノム科学総合研究センターの林崎良英プロジェクトディレクター、岡崎康司チームリーダーおよび小長谷明彦プロジェクトディレクターの各氏に多くの情報と資料のご提供を頂いた。深く感謝申し上げます。

参考文献

- 1) David W. Mount 著、岡崎 康司／坊農 秀雄 監訳、「バイオインフォマティクス」、メディカル・サイエンス・インターナショナル (2002)