

生体メカニズムの物質変換構造を明らかにするための相互作用マップ

●有田 正規

東京大学大学院新領域創成科学研究科、理化学研究所 植物科学研究センター、慶應義塾大学 先端生命科学研究所

<研究の目的と進め方>

代謝ネットワークにおける任意の炭素や窒素原子をトレースでき、また部分ネットワーク全体の物質収支を自在に閲覧できるソフトウェアを作成する。同様に、インタラクティブなシグナル伝達のマップを作成する。ただし代謝マップと密接に関連する、生命維持機能に必要な部分を中心に扱う。マップはグラフとして取り扱い、これらをネットワーク越しに交換、検索、また重ね合わせる技術を開発する。以上のデータおよび技術を組み合わせ、生命に必要な物質の最小セットと生化学機能を構造的に定義する。

<2007年度の研究の当初計画>

インタラクティブな代謝マップを植物二次代謝に応用

生成マップをインタラクティブに閲覧できる機能をサポートする。代謝物の構造情報だけでなく、代謝物のマスペクトル情報や、同位体を用いた実験法のデータを可視化できるソフトウェアとして完成させる。

糖鎖構造レイアウトアルゴリズムの開発と、それに基づくデータの再整理

略記された糖構造から正確な構造を再現し、美しくレイアウトするアルゴリズムを開発する。特に糖脂質の構造表示を、省略表記とそうでない場合とインタラクティブに切り替えられる機能を実装する。また植物二次代謝物（フラボノイド）のデータベースでは、糖の構造表記が非常に曖昧（不正確）になっているため、そうした構造を全て正確なものに描きなおす。

脂質代謝、植物二次代謝等の要素をマップに統合

代謝の調節に大きな役割を果たすシグナル伝達系の要素をインタラクティブ代謝マップに統合する。また、代謝物とタンパク質の相互作用情報をデータベースに追加する。前年度に引き続き、脂質代謝、植物二次代謝に関与する物質のID番号付け、分類作業をおこなう。特にカロテノイドと糖脂質を中心に作業をおこなう。

<2007年度の成果>

1. フラボノイドの文献情報整理

前年度に引き続き作成するフラボノイドのデータベースにおける文献情報の整理を行った。フラボノイド7000分子について辞書等から收拾した文献情報は入力者や入力方法によってスタイルが異なったり、入力ミスがあったりする。当初5700文献（文献に記される物質数は重複を含め30078件）収集されていた文献情報のうち、①辞書などの孫引き情報を削除し、オリジナルの文献のみを記載、②重複を省く、処理を行った結果、19728件まで減らせることができた。ほぼすべての文献をPubMedやGoogle Scholar等で調べ、スペルミス等の間違いも修正した。

2. フラボノイドの名称入力、糖鎖構造の統一

多くのフラボノイドは慣用名を持ち、名称から構造を類推することが難しい。そこで構造を反映させた名称を入力する作業をおよそ1,100フラボノイドについて行った。また同時に、名称の出現頻度をもとにプログラムで検出したスペルミス300件を修正した。ただし188件については、知識不足のために適切な体系名を記述することができなかった。

また糖の立体構造表記が統一されていなかったため、1610フラボノイドについて糖鎖部分を修正し、統一した書式とした。

3. データベースシステムの移植

これまでJava appletとMySQLによりデータベースを構築してきたが、今回、全データをMediaWikiと呼ばれるWikipediaのために開発されたシステムに移植した。MediaWikiによるシステムは従来のデータベース構築に比較して修正やメンテナンスが格段に容易であり、外部ユーザの書き込みやディスカッションも可能である。ただ、基本はテキスト用のシステムであるため、分子構造や数値属性を扱うために様々な機能の追加を施した。

4. 糖認識アルゴリズムの開発

糖の立体構造表記法であるHaworth法で描かれたXY二次元座標情報から、グルコースやガラクトース等のタイプを判定するアルゴリズムを開発した。またそのアルゴリズムをフラボノイドのデータベースに適用し、ラムノースが付加される位置に特徴があることを明らかにした。

5. フラボノイドの仮想マスペクトル生成

マスペクトル情報をデータベースに付加するため、既知のフラグメンテーション規則に基づいてフラボノイド構造から特徴的なマスペクトルを生成するアルゴリズムを開発した。

6. 日本脂質生化学会公式データベースLipidBankのリリース

日本脂質生化学会にデータベース構築委員会を結成してその主メンバーとなり、新しいLipidBankの構築からメンテナンスまでを引き受けることになった。分子構造データの修正を続行し、糖脂質およびカロテノイドの構造とリファレンスを修正した。リファレンスの修正作業はフラボノイドにおける作業と同じ。

<国内外での成果の位置づけ>

前年度に引き続き、フラボノイドのデータベースは世界的にも群を抜いた規模と精度を有している。徐々に知名度も上がりつつあり、植物業界ではサントリーをはじめとした固定ユーザがついた。国内の研究者を優先して支援したいため、ソフトウェアや多くの情報は引き続き日本語でサービスしている。

脂質のデータベースも世界でトップの精度と分子数であり海外からの参照も多いが、論文発表をまだおこなっていない。日本脂質生化学会のデータベース構築委員会と強調して成果公開をおこなう。

<共同研究>**University of Louisville, NMR Center**

Louisville 大で開かれるメタボロミクスシンポジウムで講演したが、申請したグラントは採択されなかったため、一時的に中断している。

京都大学、奈良先端科学技術大学院大学

フラボノイドのデータベースは奈良先端大バイオサイエンス研究科の金谷研究室との共同開発。また、マススペクトルの解析やデータベース化に関して金谷研究室および京大農学系研究科の西岡研究室と共同研究。これらの研究室とは、JST-BIRD プロジェクト「メタボロームマススペクトルデータベースの開発」でも共同研究。

<達成できなかったこと、予想外の困難、その理由>**1. インタラクティブな代謝マップの構築をはじめ、代謝マップ関連の作業が未着手**

データベースシステムを MySQL から MediaWiki に移行する作業が難航し、今年度の当初計画がほぼ半年遅れることになった。また、データの整理作業をおこなう人材は確保できたものの、ポストドク不在の状況がおよそ1年続いてしまったため、研究面の遅れは否めない。Java Applet で作成した旧データベースはクリック代謝マップの表示まで完成していたため、今後早急に適切な人材を確保し、MediaWiki 版データベースでも代謝マップをサポートできるようにしたい。

2. マススペクトルデータベースとの連携

JST-BIRD で開発するマススペクトルデータベースと、パスウェイマップ、フラボノイド構造をリンクすると便利なデータベースになるが、単純な HTML リンク以上の連携を行うことが難しいことが判明した。システムをすべて MediaWiki に移行したため、Wiki ページへのリンクは作成できるものの、それ以上の処理を行うには Java applet を埋め込むしかない。しかし Applet がリンクをクリックするたびに起動するようでは使い勝手が悪い。今後、使いやすい連携手法を見出していく必要がある。

<今後の課題>**1. 論文として成果を発表**

フラボノイドの階層分類やデータベース作りは丸二年間おこなってきたので、論文として成果発表をする時期に来ていたが、まだ論文化できていない。2008 年内にはデータベースを公表したい。またこのデータベースを用いて行う解析に一刻も早く着手したい。

2. データベース開発の共同作業化

MediaWiki を基本にシステムを構築する以上、ユーザ参加型のデータベース開発を前提にするべきである。

<成果公表リスト>**1) 論文/プロシーディング****1. 0801230157**

Arita, M., Tokimatsu, T.: Detection of monosaccharide types from coordinates, Proceedings of the 18th International Conference on Genome Informatics (Genome Informatics Series 19), 3-14, 2007.

2. 0801230152

Kusano, M., Fukushima, A., Arita, M., Jonsson, P., Moritz, T., Kobayashi, M., Hayashi, N., Tohge, T., Saito, K.: Unbiased characterization of genotype-dependent metabolic regulations by metabolomic approach in *Arabidopsis thaliana*, BMC Systems Biology, 1:53, 2008.

2) データベース/ソフトウェア**1. 0701262355 フラボノイド WIKI**

Flavonoid Wiki <http://157.82.238.171/wiki/>

2. 0701262353 脂質データベース

LipidBank <http://lipidbank.jp/>

3) 解説記事、その他**1. 0801230201**

福島敦史, 草野都, 有田正規, 斉藤和季「植物メタボロミクス—代謝プロファイルからシステム解析へ」実験医学, 1月号, 羊土社, 2008

2. 招待講演 Arita, M. "Software Tools to Study Flavonoids", Systems Biology and Bioinformatics Symposium (SBBS07), March 27-29 Taipei Taiwan, 2007.**3. 招待講演 Arita, M. "Databases for Flavonoids and Measurement in *Arabidopsis*", 2nd International Symposium on Environmental Metabolomics, March 24-25, University of Louisville (KY), USA, 2007.**