

メタボローム解析のための計測技術開発とそれを用いた代謝経路推定

●富田 勝^{1),2)} ◆西岡 孝明³⁾ ◆曾我 朋義^{1),2)} ◆金井 昭夫^{1),2),4)}

1) 慶應義塾大学 先端生命科学研究 2) 慶應義塾大学 環境情報学部 3) 慶應義塾大学 先端生命科学研究 (2008～2009年度)、京都大学 農学研究科 (2005～2007年度) 4) 2005～2007年度参加

<研究の目的と進め方>

DNA シーケンシング技術の各段の進歩によって、ヒトをはじめとする各種生物のゲノム DNA 塩基配列情報を短時間で安価に決定することができる。このようなポストゲノム科学の時代においては、ゲノム解析された生物種について“omics”という言葉で表されるゲノミクス以下の階層（トランスクリプトミクス、プロテオミクス、メタボロミクス）ごとに定量的な測定を行うことにより、生命活動を理解しモデル化することが生命科学の大きな目標となっている。なかでも生体内に含まれる代謝物質の包括的な理解を追究するメタボロミクスは、代謝物質の生合成ネットワークや代謝制御などのキーワードで代表されるように、omics 測定データを互に関連づけて総合し、生物学的意味を与えることができる要となるデータを与えるものと期待されている。また、これまで長年にわたって研究されてきた生化学知識の全てがメタボロームとして蓄積しているともいえる。とりわけ3つの omics 研究を互に関連付け、統合している代謝反応ネットワーク（代謝経路のネットワーク）の解明が重要課題である。

これまで我々は、主にキャピラリー電気泳動質量分析装置（CE-MS）を用いたイオン性低分子化合物の網羅的測定法の開発に力を注いできた。また、同装置によるペプチド類の測定や液体クロマトグラフィー質量分析装置（LC-MS）およびガスクロマトグラフィー質量分析装置（GC-MS）による中性物質の測定法の開発・改良も行い、解析対象となる物質の網羅性を各段に向上させてきた。さらに未知物質の同定に欠かせない MS/MS 解析や有機合成技術、インフォマティクスなどを相補的に駆使した物質同定手法を強化し、メタボローム研究における基盤技術を次々と確立した。ところが、これまでに蓄積した微生物や動植物のメタボローム解析データから、現時点で最も大規模な代謝物質データベースである KEGG や METACYC にさえその詳細が述べられていない物質が生体内に数多く存在することが分かり、依然として多くの代謝物質に関する未知の異化・同化経路が存在することが示唆された。

CE-, LC-MS で検出される代謝物質のうち MS/MS で同定できる代謝物質の割合（＝同定率）はそれほど高くはない。最近の研究例でもその傾向は変わらない。例えば、ヒトの前立腺組織の例では LC-MS で検出された生体内物質 626 のうち、LC-MS/MS 解析して化合物を同定できたものは 175 代謝物質であった（同定率 28 %）と報告されている（Sreekumar ら 2009）。ヒトは大部分の代謝物質を食物や栄養素として摂取しているので代謝物質の種類は少ない。それに対して2次代謝物質が多く、全ての有機化合物を生合成しているシロイヌナズナでは、検出した1,233 代謝物質のうち同定できたものはわずか 48 代謝物質と同定率（3.9%）は極めて低い（Matsuda ら 2009）。このように質量分析の性能向上にもかかわらず同定率は低い。仮にバイオマーカー候補として未同定化合物が見つかったとしても、それを同定するためにはマスマススペクトル解析の豊富な経験を有する熟練者であっても数週間を要しているのが実情である。

従って本研究課題の目的は、メタボローム解析技術を駆使した網羅的な代謝物質の測定および代謝ネットワークの構築、の二点である。この目的に向かって次のように研究を進める計画である。すなわち、（1）未知代謝物質を同定し、（2）同定した代謝物質から生合成代謝経路を推定する、ために必要な測定技術の開発とバイオインフォマティクスとケモインフォマティクスを融合した新しい代謝経路推定法を開発する。質量分析はメタボローム解析のみならず生命科学研究の基盤技術である。質量分析から未同定物質の化学構造を推定できると、生命科学研究全体に与えるインパクトは大きい。

これまでに確立したメタボローム解析に関する測定基盤技術および未知物質同定手法のさらなる改良に加え、トランスクリプトームやプロテオーム解析技術をも統合し、細胞によるエネルギー生成、様々な細胞ストレスに対する応答反応や恒常性の維持、外来性物質の無毒化などに関する新規の細胞代謝経路の解明を試み、微生物や動植物細胞におけるより完全な代謝ネットワークの構築を目指す。

<研究開始時の研究計画>

メタボローム解析の高感度化、ハイスループット化、代謝物質の同定率向上、代謝経路推定法の開発の4項目について研究開発をおこなう。

高感度化とハイスループット化は分析機器のハードとしての性能に依存するだけでなく、測定条件の改良と、得られる大量データの自動処理技術の開発に重点を置く。

代謝物質の同定率の向上については以下の方法をとる。メタボローム解析では、クロマトグラフィー（LC）やキャピラリー電気泳動（CE）による分離と質量分析の2つを組み合わせ（LC-MS, CE-MS）、代謝物質の同定、定量をおこなっている。代謝物質の標準試薬が入手できる場合には、（1）分離における保持時間あるいは泳動時間 t 、（2）質量分析における分子イオンの m/z 、この2つのパラメータ（ $t, m/z$ ）値をあらかじめ測定しておく。生体試料を分析して2つのパラメータが標準試薬のそれらと一致すると、同じ代謝物質と同定することができる。

標準試薬が入手できない場合に代謝物質を推定する方法として、2つのパラメータそれぞれについて代謝物質の化学構造との関係を解析して関係を見出す。次ぎにその関係を利用して代謝物質を推定する、というアプローチが考えられる。さらに、標準試薬の MS/MS 測定は代謝物質の化学構造を反映した product ion を生成するので、化学構造と product ion との関係を明らかにすることができれば、MS/MS データから未知化合物の化学構造に関する手掛かりを得ることができると期待される。そこで、これら MS/MS 測定をできるだけ多様な代謝物質についておこなって、集積、解析をおこなう。

代謝経路推定には代謝物質量の時間変動や代謝フラックス分布の解析が有効であると考えられているが、実際に解析した例が少ない。そこでメタボローム解析の結果をこれらの解析に適用する

ことによって、それらの利点、欠点、有用性を調べる。さらに、代謝反応ネットワークを構成している要素は代謝物質と代謝経路である。代謝物質の生合成反応や制御には、酵素タンパク質遺伝子の発現制御や活性制御が関与している。代謝物質の化学分析だけでなく、その他の omics 解析データを組み合わせた代謝経路推定もおこなう。

<研究期間の成果>

1. 高感度化とハイスループット化

CE-MS に利用されていた 4 重極型質量分析計 (QMS) を飛行時間型質量分析計 (TOFMS) に換えることによってイオンの m/z の測定精度を 50-100 ppm まで向上させ、さらにイオン性代謝物質のスループットを約 30 倍改善した。同時に検出感度も数倍～数十倍向上した。また、LC よりも流量がはるかに少ないキャピラリー泳動に適したイオン化装置に改良した。QqTOFMS の分析条件の検討を重ね、1 μ M レベルでの CE-MS/MS 測定を実現させることに成功した。疎水性中性代謝物質に対しては C4 逆相カラムを用いた 0.02% 酢酸、イソプロピルアルコールによるグラジエント LC-TOFMS、親水性中性代謝物質に対しては両性イオン官能基 (ZIC-HILIC) カラムを用いた 100 mM 酢酸アンモニウム、アセトニトリルによるグラジエント LC-TOFMS 法による測定条件を確立した。

また、クロマトグラムのピーク測定データを高精度にアライメントする自動処理法および参照マスペクトルを用いた代謝物質の自動解析技術を開発した。

メタボローム測定で得られた千以上の代謝物質の中から試料間で変動した成分を探索することは難しく、データ解析に膨大な時間と労力を費やしていた。そこで、CE-TOFMS で計測した 2 群間のデータから、バイオマーカの候補になり得る顕著な差を検出するメタボローム・ディファレンシャルディスプレイを開発した。CE-TOFMS で得られた各代謝物の移動時間を X 軸に、質量数を Y 軸に表示し、二次元で可視化した。CE-TOFMS 測定では、同じ代謝物であっても測定毎に各代謝物質の移動時間と精密質量は少しずつ変動し、二次元表示の位置がずれる問題点があった。そこで検出されたすべての物質を用いて測定毎での測定時間のズレを測定し、補正曲線を求めた。この補正曲線を用いて各試料の測定時間の誤差を補正した。精密質量については、質量数を補正する物質を連続的に質量分析計に導入し、すべての MS スペクトルを自動補正した。この方法を用いることで、各試料中の同一物質を二次元上の同じ位置にプロットすることが可能になり、試料間での比較や大きな変動のある物質だけを表示することが可能になった。

さらに大量のデータを高速に処理できるソフトウェア JDAMP を新たに開発した。従来、Mathematica のスクリプトであったプログラムを、画面 (GUI) は Java 言語、計算エンジンは C++ で実装し、ユーザが対話的に処理を進められ、結果を検索・確認しながら解析できる仕組みとした。特にデータの比較に重要なバックグラウンドとノイズの除去のプロセスに、従来 6 時間程度 (1 計測データあたり) を要していたが、JDAMP では 30 分程度で処理することができるようになった。その後の時間補正、濃度の補正、結果の可視化のプロセスもそれぞれ数分程度で処理ができるよう高速化した。

また、CE-TOFMS の計測データに含まれる離散的な m/z と強度の値から、精密質量を高精度に計算するソフトウェアも開発した。計測時に質量既知の物質を連続的に質量分析装置に導入し、質量の補正を行なっているが、更にバックグラウンドの歪みを 3 次元的に捉えてマスペクトルの歪みをソフトウェア上で補正

することによって、より高精度に精密質量を計算する機能を実装した。標準物質でベンチマーク試験を行ったところ、ほぼ全てのピークの m/z 値を理論値から 5 ppm 程度の誤差に収めることができた。更に CE-TOFMS に特異的に検出されるフラグメント、アダクト、リングング、スパイクノイズなどの特徴を調査し、検出したピークの全てにピーク間の関連付けと意味付けを自動的に施す機能を実装した。複数の計測データを同時に解析するため、JDAMP と同様、計測データ間で高精度な時間補正を行い、ピークのアライメントを行うこともできる。また、標準物質の泳動時間や分子構造から neural network を用いて予測した移動時間 t (2.2 参照) も同時にアライメントすることで、各ピークに物質名の候補をアノテーションすることができることもできる。これらの機能の実装により、CE-TOFMS で計測したデータにどのような代謝物質が入っているのかを解析する時間を、大幅に短縮させることができた。

未知物質の同定および定量には、組成式の決定に加え候補物質の標準物質によるピークの検証が必要であるが、生体サンプル中に検出される未知ピークに対する候補物質の多くは標準物質として市販されていない。従って、化学合成や生合成によって膨大な数の候補物質の標準物質を効率的に合成する必要があるが、新規の合成戦略として本年度新たに「pH 制御合成法」を開発した。これは、従来酸性・塩基性と二極的に考えていた反応条件をさらに細かい pH 条件に割り振ることで、より精密な合成を可能とした合成法であり、これまでに N-アセチルオクトパミン、N-アセチルドーパミン、N-アセチルチラミンなどの高収率な化学合成に成功した。これらは脳内神経伝達システムの解明、うつ病及びパーキンソン病治療への応用研究などに用いられる予定である。

2. 代謝物質同定の精度向上および推定に関する技術開発

質量分析で検出される分子イオンの m/z を精密に測定することによって代謝物質の化学式を推定できるものと期待される。しかし、仮に近い将来に m/z を超高精度 (< 0.5 ppm) で測定する質量分析機器が開発されたとしても、分子イオンの m/z だけから分子式を決めることが難しいことが指摘されている (Breitling *et al.* *TREND in Biotech*, 2006)。未同定物質の化学構造推定をおこなううえで、高精度な分子イオンの m/z 測定と同位体イオンの精度よい強度比測定のための技術開発、分子式から CE の泳動時間 t の推定、MS/MS データを参照スペクトルとした物質同定、これら 3 つを組み合わせることが最も有効であると考えた。

2. 1. 炭素同位体ピークを考慮した分子式の絞り込み

高分解能な質量分析計で検出した代謝物質の分子イオンの m/z 値にあてはまる元素の組合せ (分子式) にはいくつもの組合せが可能である。これらのうち、KEGG や KNApSAcK などの代謝物質 (天然物) データベースに収録されている既知代謝物質の分子式に当てはまるものだけを選んで、複数の候補分子式が残る。質量分析では同位体ピーク (A, A+1, A+2) 比を精度よく測定することができるので、分子式に含まれる炭素数やハロゲン原子数を同位体ピーク比にあてはめることによって精度よく絞り込むことができる。

当初、質量精度が 2ppm 以内のスペックを持つ、LTQ-Orbitrap 質量分析計を用いてモノアイソトピック質量 (分子を構成する各元素の代表同位体の精密質量) および同位体イオンから未知物質の組成式を決定する計画をした。多様な測定条件を検討したが、分子式の決定に必要な感度を得ることができなかった。むしろ、これまでに開発した精密質量を高精度に計算するソフトウェアを改善して用いると、CE-TOFMS の計測データから、同位体イオ

ンを感度よく測定できることがわかった。

2. 2. 泳動時間tから代謝物質を推定する

標準試薬を用いて約700化合物についてCEとLCの泳動時間tを測定した。これらtの値と化合物の化学組成(原子の種類と数)、結合次数とその数や芳香族性などの構造要素、分子電気陰性度、双極子モーメント、水/オクタノール系分配係数(LogP)など物理化学的性質をパラメータとして解析した。人工知能技術(Artificial Neural Network)を用いて陽イオン性代謝物質について、化合物の構造式からCEの泳動時間tを推定する方法を開発した。これを用いると、実測した泳動時間tが同定した代謝物質の化学構造と矛盾しないかどうかを検証することができる。泳動時間tと化学構造とを関係づける人工知能技術を我々がこれまでに開発してきたアノテーションツールに組み込むことによって、CE-MSによる代謝物質の同定精度を飛躍的に向上させることができた。

2. 3. MS/MS測定を利用した物質同定

MS/MSはHe, N₂, Arなどの気体(collision gas)を高速で分子イオンと衝突させることによって分子イオンを解離させて(collision induced dissociation, CID)、生成したイオン(product ion)を質量分析する分析法である。この解離は、分子イオンの化学構造に依存して特徴的に生ずるのでMS/MS測定したマススペクトルには化学構造を特徴づけるproduct ionが観察される。

あらかじめ代謝物質の標準試薬を用いてCE-ESI-QqQMSを用いてMS/MSデータを測定、収集しておくことにした。これまでのCE-MSのtと分子イオンのm/zの2つを参照する同定法と比較して、泳動時間tとMS/MSデータの2つを参照した場合には、これまでよりもはるかに精度よく同定することができた。ここで「精度よく」ということを定量的に表現することは難しいが、同定した物質がいずれもある1つの同じ代謝経路の中間体であった、などによって実感できた。

これまでに約700代謝物質についてMS/MS測定をおこなった。それらのうち、695代謝物質について測定した4,275 MS/MSデータはpublicマススペクトルデータベースであるMassBank(<http://www.massbank.jp>; 2006年12月にJST-BIRDプロジェクトとして開始)から公開している。化合物数に比べてデータが多い理由は、1化合物あたり平均5段階に異なるCID衝突エネルギーで測定していることによるものである。このように衝突エネルギーを変えて測定する理由は、このエネルギーによって解裂の程度が異なるからである。Collision gasのエネルギーが低いと分子イオンの解裂が少なく、m/zが大きなproduct ionしか生じない。それに対して、エネルギーが高いと解裂が進行してm/zが小さなproduct ionが生じやすくなる。すなわちcollision gasのエネルギーに依存して、同じ代謝物質であっても異なるマススペクトルを与える(=再現性が低い)ので、MS/MS測定で得られたマススペクトルから代謝物質を同定することが難しくなることを示している。

このようにMS/MSを利用すると代謝物質の構造情報が得られるという特徴がある一方で、再現性が低くなるという問題が生じた。この問題は次のような方法で解決した。

2. 4. 高精度なMS/MSマススペクトルの測定

先に(2.3)で695代謝物質について測定したMS/MSデータは、イオンの分析部が4重極型質量分析計(QMS)であったので、product ionのm/zの測定精度は整数(nominal mass)である。そこで同じ化合物についてESI-QTOF-MS/MSを用いて、より高精度にproduct ionのm/zの測定(平均誤差は約50 ppm)をおこなった。先ほどと同様に1化合物あたり平均5段階の異なるCID衝突エネルギーで測定した。695化合物についてMS/MS測

定した4,510件のデータを得た。

下の表は4-methylpyrazoleをESI-QTOF-MS/MSで測定したマススペクトルに現れたproduct ionのm/z値から化学組成を自動的に推定したものである。

表 4-methylpyrazole から生じた product ion

m/z 実測値	推定分子式	m/z 計算値	誤差 (ppm)
83.063	C ₄ H ₇ N ₂	83.06092	25
81.0471	C ₄ H ₅ N ₂	81.04527	23
67.0269	C ₃ H ₃ N ₂	67.02962	41
43.0276	CH ₃ N ₂	43.02962	47

いずれのproduct ionについても測定誤差範囲内で分子式を正しく推定できていることがわかる。いずれの分子式も4-methylpyrazoleの分子式の元素組成と原子数の範囲内であったことから、これらは4-methylpyrazoleから生じたproduct ionであることが確認できた。

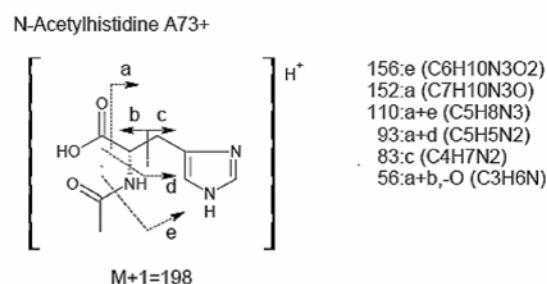
また、1化合物あたり平均5段階の異なるCID衝突エネルギーで測定したMS/MSデータを重ね合わせてみると、その化合物がCID解離で生じる全てのproduct ionを含んでいることがわかる。この人為的なマススペクトル(merged mass spectrum)を測定した化合物それぞれについて1つずつ作成した。695化合物について作成した914件の重ね合わせマススペクトルをMassBankで公開した。アミノ酸は陽イオンモードと陰イオンモードの2つの重ね合わせマススペクトルが公開されている。重ね合わせマススペクトルを参照マススペクトルとした場合、代謝物質の同定率が向上し、精度も改善されることがMassBankのスペクトル検索ツールを用いた検証実験から明らかにされた。

2. 5. MS/MSマススペクトルの化学情報化

これまでマススペクトルで検出されるイオンは全てm/z(数値)で表現されていた。高精度な質量分析で測定されたマススペクトルデータは、単に有効数字の桁数が多い、単なる数値のリストデータにすぎない。そこには測定した化合物を特徴づける化学構造や物理化学的性質などの化学情報は何も含まれていない。

未知代謝物質を同定するためには、product ionと化学構造に関する知識を獲得、集積することが必要である。そのとき、product ionをm/zで表すのではなく分子式で表すことができれば、このような知識を獲得することが容易になると予想した。

図1. 結合の切断位置とproduct ion.



ESI-QTOF-MS/MSで測定したマススペクトルのproduct ionはほとんど自動的に、しかも高い信頼性で分子式に置き換えることができた。また推定したproduct ionの分子式からprecursor ionにおいて解裂した共有結合も高い精度で推定することができた(図1参照)。

図1で示した化合物(N-acetylhistidine)のproduct ionのうち、この化合物を特徴づけるイオンはm/z 83, 152, 156であることがわかる。このような解析を他の化合物についてもおこなって、あ

る特定の部分化学構造を特徴づけるイオンの確かさを評価する必要がある。

MS/MS データを化学情報化すると、測定した質量分析計の分解能に依存しないし、測定誤差も無くなってしまう、というメリットがある。MassBank で公開された 695 化合物に関する MS/MS データの化学情報化と知識獲得は MassBank プロジェクトで作業をおこなっている。

数値データを化学情報データに置き換えることによって、次の 3 つの新しい応用が開けると期待される。

- (1) 化学構造と product ion の関係に関する知識を獲得できる。
- (2) 代謝物質同定における参照マススペクトルとしての利用。誤差の無い、高精度な代謝物質同定が可能になる。
- (3) フラックス解析を利用した代謝経路推定への応用。

フラックス解析では、前駆代謝物質の特定の位置にある炭素原子を同位体で標識して、一定時間後に標識が下流の代謝物質のどの炭素原子に移動するのかを定量的に測定する。現在は ^{13}C 同位体で標識した前駆代謝物質を用いて、下流の代謝物質に取り込まれた炭素原子の位置を $\text{C}^{13}\text{-NMR}$ で同定している (Wiechert, *Metabo. Eng.*, 2001)。これは労力と時間、経費を要する。それに対して LC-MS/MS 分析の化学情報化を利用すると、下流の代謝物質の MS/MS 分析を超微量で一挙におこなって、標識された炭素原子を特定することができる。微量のサンプルを用いて短時間かつ安価に、代謝経路を精度よくトレースすることが可能になる。細胞内の代謝反応ネットワークには枝分かれ経路、バイパス経路など多くの重複 (redundancy) があるので、単に ^{14}C 原子がとりこまれているかどうか、だけで代謝経路を推定することは精度が悪く、時には致命的な誤りを犯す危険性がある。

3. 代謝経路の推定

これまでメタボローム解析を利用した代謝経路の推定は、流速解析 (フラックス解析) を用いてきた。この方法は上述した (2. 5) ように、代謝経路の重複の問題があるので、前駆代謝物質や分析対象とする下流の代謝物質を慎重に選ぶ必要がある。

本研究課題で開発した高感度でハイスループットなメタボローム解析を利用すると、多くの代謝物質を一定時間ごとに測定した時間分割データを容易にえることができる。このような時間変動データから代謝経路を推定する方法を開発した。

3. 1. 時間分割測定データに基づく代謝経路推定法

植物における代謝物質は光合成反応を利用した Rubisco による CO_2 固定反応を唯一の炭素源として生成される。すなわち代謝反応ネットワークの原点である。さらに、明暗周期に伴って代謝は CO_2 固定反応と呼吸反応の切り替えが行われ、代謝物質が変動すると推定される。これを代謝に与える摂動とみなすと、代謝物質の時間変動から代謝経路を推定できると考えた。

イネを 13 時間明 (25°C) -11 時間暗 (20°C) の環境下で育てた。この第 3 葉を 1 時間ごと、24 時間にわたって切り取って、代謝物質を抽出したのち、CE-MS と CE-DAD (diode array 検出器) を併用して代謝物質量を測定した。光合成と呼吸、窒素代謝、エネルギー代謝に関係する 1 次代謝物質 70 化合物のうち、57 化合物について定量することができた。残り 13 化合物は検出できなかった。

代謝物質ごとに時間変動を見ると次の 3 つのグループに分類された；(1) 明暗周期と同調して変動するグループ、(2) 変動するものの明暗周期とは同調しないグループ、(3) ほとんど変動しないグループ。グループ (1) は CO_2 固定反応に近い代謝中間体、解糖系やペントースリン酸サイクル、TCA サイクルなどのエネ

ルギー代謝系の代謝物質であった。グループ (3) は糖や糖鎖合成の反応中間体、ADP, NADP, NADH などであった。

酵素反応は可逆反応であるので、1 つの酵素反応の基質と生成物は平衡関係にある。基質と生成物の関係にある代謝物質は互いに似た時間変動をするものと予想される。互いに似た時間変動をする代謝物質群は、共通の代謝経路上にあるか、または代謝経路の一部を構成するモジュール内にある、と推定される。

Kohonen の自己組織化マップ (SOM) を利用して時間変動の類似性にしたがって代謝物質を分類することにした。各代謝物質の時間ごとの量を 24 次元ベクトルで表現して、ベクトルの類似性に従って 24×24 の SOM にマップした (図 2 A)。代謝物質は 2 つの大きなグループに別れ、左側は明期に濃度が増える代謝物質、右側は暗期に増える代謝物質に分けられている。

図 2 A. 代謝物質の明暗時間変動の SOM.

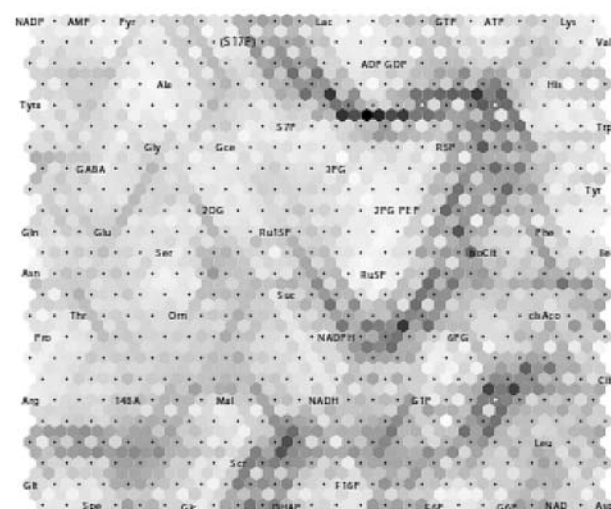
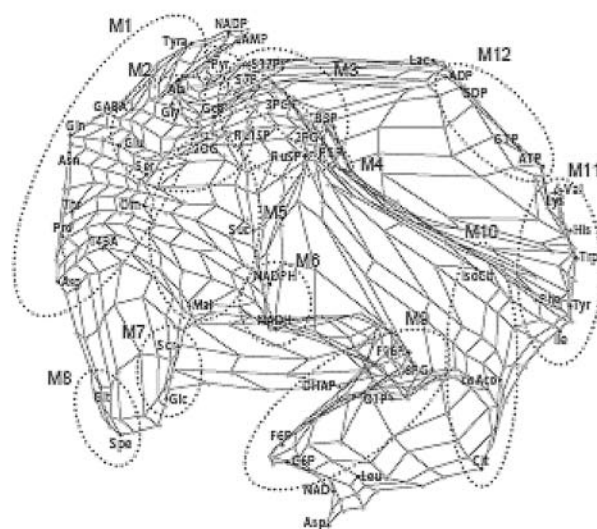


図 2 B. SOM 分類結果の Sammon マップ表現



SOM の分類結果 (図 2 A) を入力した 24 次元ベクトル間のユークリッド距離に従って 2 次元上で表現した Sammon マップを図 2 B に示す。このマップで距離が近い代謝物質どうしを、KEGG の代謝経路を考慮してグループ化すると、12 のグループができた。これらは代謝経路を構成しているモジュールとみなすことができる。このモジュールは光合成で固定された炭素原子が、解糖系やペントースリン酸回路、TCA サイクルを経て、アミノ酸や核酸

の生合成へと、上流から下流へ流れていく場合に、反応速度力学的に連続した酵素反応を表していると理解することができる。

これまで代謝経路を推定した研究は植物の環境変動に伴う生合成酵素遺伝子の発現量を microarray で測定して、対照実験区と比較して、SOM 解析した例が多い。本研究のように代謝物質量の時間変動データを1時間ごとに24時間測定して、代謝経路推定に応用したのは初めての例である。これだけ大量のデータを測定することが可能になったのは、メタボローム解析のハイスループット化によるところが大きい。

フラックス解析では代謝経路をあらかじめモデル化しておく必要がある。既知の代謝経路に依存するところが大きい。これに対して代謝変動の類似性を利用した SOM 解析では既知の代謝経路の知識だけに頼らないので新しい代謝経路を見つけるチャンスが高いと期待できる。しかし、そのためには代謝経路ネットワーク全体をある程度の密度でおおうことができるだけ多数の代謝物質の変動データを取得する必要がある。

3. 2. その他の開発

代謝経路の推定をおこなうために必要なその他の技術の開発をおこなった。既に開発した連続培養における代謝流束推定法に加え、応用範囲の広い回分培養に対応可能な手法を開発した。プロテオーム解析においては、既存の手法より広い濃度オーダーで各タンパク質の正確な定量値を得られる emPAI 法を確立し、生体サンプル中の数千のタンパク質の絶対量の推定を可能にした。

4. 応用例

1. から 3. までの開発研究の成果は統合されて、メタボローム解析をおこなう複数のツールに組み込まれている。これらに応用した研究成果のいくつかを示す。

4. 1. 大腸菌の1遺伝子欠損株のメタボローム解析

大腸菌の解糖系、ペントースリン酸回路の酵素遺伝子を1だけ欠損変異した24変異株を一定の増殖速度に保ちながら培養した。解糖系とペントースリン酸回路の代謝物質130の細胞内量を CE-TOFMS で定量した。欠失した酵素の近傍にある代謝物質量は変動した。しかし、この変動は増殖速度を 0.1 -0.7 回/h にしたときの変動と比べると、それほど大きいわけではなかった。増殖速度が遅いとき、酵素遺伝子の欠損が遺伝子転写や酵素タンパク質の量的変動に与える影響は、予想していたよりも小さかった。むしろ他のバイパス経路にある酵素遺伝子の発現が増加していることやフラックス比解析によって、炭素源がバイパス経路を迂回して (rerouting) 供給されるので代謝物質の量がほぼ一定に保たれていることが明らかになった。これに対して、生育速度を速くすると酵素タンパク質の量を増やすことによって代謝物質のプロファイルを一様に保っていることがわかった。このように大腸菌は2つの相補的な方法で恒常性を保っている。これは大腸菌の恒常性の機構を omics 解析で実証した最初の例である。これらの測定データは全てデータベース化し、Web 上で公開している (<http://ecoli.iab.keio.ac.jp/>)。

4. 2. 大腸がん、胃がんの代謝プロファイル

たいいていのがん細胞は十分な酸素が供給される環境においても、TCA サイクルを経由する酸化的リン酸化ではなくて解糖系でエネルギー代謝をおこなっていることが知られている。しかし、これまでがん細胞でメタボローム解析のデータはほとんどなかった。大腸がんおよび胃がんのがん組織における94代謝物質のプロファイルを健康な組織と比較した。乳酸および解糖系中間物質の濃度が高いのに対して、グルコース濃度は低いことが2つのがん組織に共通していた。組織の壊死によってアミノ酸含量が高く、グルタミン酸分解活性が活発であった。TCA サイクルの

代謝物質のプロファイルが組織によって異なっていることなどが明らかになった。

4. 3. 唾液のメタボロームプロファイルによるがん診断

唾液は採取が簡単であるので様々な疾病の初期診断に利用できると期待されている。これまで唾液のメタボローム解析に基づいた診断法は未だ開発されていなかった。唾液のメタボローム解析法を開発し、それを口頭がん、膵臓がん、乳がんの診断に利用できることを実証した。がんの診断に利用できると予想された57主要代謝物質を CE-TOFMS で分析し、高感度で高精度にこれらの代謝物質を自動同定し、3つのがん患者を判別する自動化システムを提案した。

4. 4. がん細胞の新規エネルギー代謝機構の解明

がん細胞が好氣的環境下でも酸化的リン酸化よりむしろ解糖系依存的に代謝を行う現象は「ワールブルク効果」と呼ばれ、様々ながんにおいて確認されている。ところが膵臓癌など乏血性のがん種では、血流が乏しいために慢性的な虚血状態に曝された細胞が、酸素やグルコースの供給が不十分な栄養的に極めて劣悪な、微小環境で増殖していることが知られている。従って、このようながん細胞のエネルギー代謝を解糖系の亢進だけで説明することは難しい。血流不足により解糖も酸化的リン酸化も慢性的に制限された環境において、細々としかし着実に生き永らえるがん細胞は、増殖に必要なエネルギーをいったいどのように捻出しているのだろうか？そこで、Panc-1 (膵臓癌細胞株) および HDF (皮膚線維芽細胞) を、このようながんの微小環境を模倣したグルコース欠乏かつ低酸素状態で培養し、CE-TOFMS による時系列メタボローム解析をおこなって、栄養飢餓状態におけるがん細胞のエネルギー代謝に関するジレンマの解明を試みた。

エネルギー代謝に関連する120種以上の代謝物質の定量結果から、がん細胞特異的にコハク酸が細胞内外に顕著に蓄積することを見出した。グルコース欠乏かつ低酸素状態で培養した Panc-1 (膵臓癌細胞株) のメタボローム解析および同位体標識したピルビン酸の代謝流束解析を行った。これらからコハク酸の生成経路を計算的に予測したところ、これまである種の嫌気性微生物や寄生虫などにおいてしか確認されていなかった「フマル酸呼吸」 (= NADH- フマル酸還元系) に依存したエネルギー代謝 (= ATP 合成) をおこなっていることが示唆された。これはこの特殊な代謝機構が哺乳動物細胞をはじめヒト細胞で初めて確認された例である。酸素やグルコースに依存しない ATP 生成を可能にするため、栄養飢餓状態に曝されたがん細胞がエネルギー代謝の切り札として利用していると考えられる。さらに、駆虫薬として臨床的に使われているパモ酸ピルベニウムがフマル酸呼吸の阻害剤であることを突き止め、この薬剤の投与によりコハク酸および ATP の生成が顕著に抑制されることを確認した。がんの特殊なエネルギー代謝を標的とする全く新しい抗がん剤として、フマル酸呼吸阻害剤の開発が今後期待されると共に、本研究グループが開発してきたメタボローム解析技術が、新規代謝経路の推定に応用可能であることを実証した。

<国内外での成果の位置づけ>

本研究代表者を中心としたメタボローム解析技術開発とそれらに応用した研究のレベルは世界の最先端をいくものとして高く評価されている。これらの成果は近年世界中に普及しつつあり、測定技術の向上および応用研究が盛んに行われている。

CE-TOFMS で計測した大量のデータを高速に処理できるソフトウェアと m/z 値を高精度に計算するソフトウェアは国内外を問わず最初である。これらはメタボローム解析のための計測技術を格段に向上させ、広く利用されることが期待される。

約 700 代謝物質の標準試薬について異なる CID 条件で測定した 4500 件以上の MS/MS データを測定して公開したことは世界で他に例をみない。いずれのマススペクトルも極めて良質であり、本研究グループの高い測定技術を国内外にアピールした。

本研究課題によって開発されたこれらの基盤技術を活用した例としては、メタボローム解析をはじめとする omics データの完全なセットから大腸菌の恒常性の機構を解明したことや、がん細胞のメタボローム解析、代謝流束解析などを挙げることができる。いずれも世界初の快挙といえる。メタボローム解析における基盤技術の高さとその応用力を実証した。

<達成できなかったこと、予想外の困難、その理由>

当初の計画では質量精度が 2ppm 以内のスペックを持つ、LTQ-Orbitrap 質量分析計を CE や LC に接続し、得られた未知物質のモノアイソトピック質量（分子を構成する各元素の代表同位体の精密質量）および同位体イオンから未知物質の組成式を決定する予定であった。同位体イオンの感度がモノアイソトピック質量に比較し 2 桁以上低いことが判明した。その後、同位体イオンを感度よく検出できる実験条件を見つけることを主要課題の 1 つとして検討を続けたが、期待されるような感度の向上は得られなかった。CE で分離をした場合にはサンプル量が極微量であることや、高い分離のために鋭いピークとして代謝物質が分析されることなどの理由が考えられた。

<今後の課題、展望>

代謝物質の同定率を向上することは、質量分析に依存するメタボローム解析やプロテオーム解析など生命科学研究分野にとって共通の課題である。本研究課題の初期計画でも取上げたように、質量分析データであるマススペクトルから化学構造を自動的に推定する技術の開発が望まれる。そのためには、マススペクトルのデータを集積することと、その解釈、注釈をおこなった知識の集積や利用をする情報科学技術を開発する必要がある。これはチャレンジするに値する将来に残された大きな課題である。

<研究期間の全成果公表リスト>

1. Sugimoto, M., Wong, D. T., Hirayama, A., Soga, T. and Tomita, M. "Capillary electrophoresis mass spectrometry-based saliva metabolomics identified oral, breast and pancreatic cancer-specific profiles" *Metabolomics*, in press
2. Nishino, T., Yachie, A. K., Hirayama, A., Soga, T., Suematsu, T., and Tomita, M. "In silico modeling and metabolome analysis of long-stored erythrocytes to improve blood storage methods" *Journal of Biotechnology*, in press
3. Sugimoto, M., Hirayama, A., Ishikawa, T., Baran, R., Robert, M., Uehara, K., Soga, T. and Tomita, M. "Differential Metabolomics Software for Capillary Electrophoresis-Mass Spectrometry Data Analysis" *Metabolomics*, in press
4. 0911251310
Soga, T., Igarashi, K., Itoh, C., Mizobuchi, K., Zimmermann, H. P. and Tomita, M. "Metabolomic Profiling of Anionic Metabolites by Capillary Electrophoresis Mass Spectrometry" *Analytical Chemistry* 81:6165-6174 (2009)
5. 0911251311
Hirayama, A., Kami, K., Sugimoto, M., Sugawara, M., Toki, N.,

Onozuka, H., Kinoshita, T., Saito, N., Ochiai, A., Tomita, M., Esumi, H. and Soga, T. "Quantitative metabolome profiling of colon and stomach cancer microenvironment by capillary electrophoresis time-of-flight mass spectrometry" *Cancer Res* 69:4918-25 (2009)

6. 0901091419

Sato, S., Arita, M., Soga, T., Nishioka, T. and Tomita, M. "Time-resolved metabolomics reveals metabolic modulation in rice foliage" *BMC Systems Biology* 2:51(2008)

7. 801291233

Soga, T., Ishikawa, T., Igarashi, S., Sugawara, K., Kakazu, Y. and Tomita, M.: Analysis of nucleotides by pressure-assisted capillary electrophoresis mass spectrometry using silanol mask technique, *Journal of Chromatography A*, 1159, 125-133(2007)

8. 708071724

Toya, Y., Ishii, N., Hirasawa, T., Tomita, M. and Soga, T.: Direct measurement of isotopomer of intracellular metabolites using capillary electrophoresis time-of-flight mass spectrometry for efficient metabolic flux analysis, *Journal of Chromatography A*, 1159, 134-141 (2007)

9. 705011901

Ishii, N., Nakahigashi, K., Baba, T., Robert, M., Soga, T., Kanai, A., Hirasawa, T., Naba, M., Hirai, K., Hoque, A., Yee, Ho, Pei., Kakazu, Y., Sugawara, K., Igarashi, S., Harada, S., Masuda, T., Sugiyama, N., Togashi, T., Hasegawa, M., Takai, T., Yugi, K., Arakawa, K., Iwata, N., Toya, Y., Nakayama, Y., Nishioka, T., Shimizu, K., Mori, H. and Tomita, M. "Multiple high throughput analyses monitor the response of *E. coli* to perturbations" *Science*, 316, 593-597 (2007)

10. 0805201343

Baran, R., Robert, M., Suematsu, M., Soga, T. and Tomita, M. "Visualization of three-way comparisons of omics data" *BMC Bioinformatics*, 8, 72(2007)

11. 705021212

Ishii, N., Suga, Y., Hagiya, A., Watanabe, H., Mori, H., Yoshino, M. and Tomita, M.: Dynamic simulation of an in vitro multi-enzyme system, *FEBS Letters*, 581, 413-420(2007)

12. 0805201338

Baran, R., Kochi, H., Saito, N., Suematsu, M., Soga, T., Nishioka, T., Robert, M., Tomita, M. "MathDAMP: a package for differential analysis of metabolite profiles" *BMC Bioinformatics* 7, 530 (2006)

13. 0805201333

Soga, T., Baran, R., Suematsu, M., Ueno, Y., Ikeda, S., Sakurakawa, T., Kakazu, Y., Ishikawa, T., Robert, M., Nishioka, T. and Tomita, M. "Differential metabolomics reveals ophthalmic acid as an oxidative stress biomarker indicating hepatic glutathione consumption" *J. Biol. Chem.* 281: 16768-16776 (2006)

14. 0601261933

Ishii, N., Soga, T., Nishioka, T. and Tomita, M. "Metabolome Analysis and Metabolic Simulation" *Metabolomics* 1: 29-37(2005)

15. 0601261926

Sugimoto, M., Kikuchi, S., Arita, M., Soga, T., Nishioka, T. and Tomita, M. "Large-Scale Prediction of Cationic Metabolites Identity and Migration Time in Capillary Electrophoresis Mass Spectrometry using Artificial Neural Networks" *Analytical Chemistry* 77(1): 78-84(2005)