

メタボロミクス（代謝物総体解析）の食品工学への応用

福崎英一郎

近年、代謝物総体（メタボローム）に基づくオーム科学であるメタボロミクス（メタボローム解析）が、ゲノム情報にもっとも近接した高解像度表現型解析手段として注目されているが、その最大のアドバンテージは、適応範囲の広さである。基幹代謝物は、当然のことながら生物間で互換性を有するので、ゲノム情報が利用できない実用植物や実用微生物にも適用可能な唯一のオーム科学といえる。また、観測感度や解像度にこだわらなければきわめて安価に実施できるコストパフォーマンスの高いツールでもある。その応用範囲はポストゲノム科学にとどまらず、医療診断、病因解析、品種判別、品質予測などの多岐に及ぶ。さらに、メタボロミクスの技術は生物以外の対象にも適応可能である。たとえば、食品や生薬原料のプロファイリングによる品質予測、あるいは酵素反応混合物のプロファイリングによる反応機構推定などの応用分野も想定できる。

さて、上記のように有望技術として期待されているメタボロミクスであるが、観測対象の代謝物が多岐に渡るゆえに手法の標準化が困難であり自動化も進んでいない。また、得られた膨大なデータから有用な結論を導く作業、すなわち「データマイニング」についても標準技術は確立されていない。結果として研究対象ごとに各論が展開されている。当該状況がメタボロミクスの正しい理解を困難とし、一般の研究者に普及しない一因となっている（図1）。本稿では、メタボロミクスの戦術を解説するとともに、食品工学を含む種々の用途への適用可能性を提示することを目的とした。

メタボロミクスに用いる質量分析

質量分析の利点は、大量の定性情報が得られる上に高感度で定量可能であることである。観測標的が決まっている場合、しかるべき内部標準化合物を用いた厳密な定量分析手法が適用可能である。しかしながら、特定の標的を決めずに網羅的に解析することを主眼とするメタボロミクスにおいては、従来の内部標準法による標準化は困難である。いかにして定量性を検証するかが重要となる。また、大量に存在する代謝物に混在する微量代謝物を網羅するためのノウハウも必要である¹⁾。

メタボロミクスに用いる質量分析は特に限定されない

が、種々の分離手段と組み合わせて運用し、「保持時間」と「質量分析データ」の両方のデータを代謝物情報とする場合が多い。「解像度」と「再現性」に最もすぐれた手法としてガスクロマトグラフィー質量分析（GC/MS）が挙げられる。GC/MSは、ほぼ完成した分析システムであり、質量分析の経験のないバイオサイエンス分野の研究者にも容易に扱えることが特長である。キャピラリー電気泳動質量分析（CE/MS）は、イオン性代謝物を観測するのに適した手法でありメタボロミクスにおける重要手法のひとつであるが、残念ながらGC/MSほど一般化した観測手法とは言えず、実用運用には若干のノウハウが必要である^{2,3)}。筆者らのグループも、極性を反転させることにより、未修飾のキャピラリー管でも分析可能なアニオン分析システムを開発し実用化している⁴⁾。

上記の他にも、対象代謝物の性質に応じて種々の分離手法が用いられる。HPLC質量分析（LC/MS）は、低分子から高分子まであらゆる化合物に対応したすぐれた分離分析系だが、ピークキャパシティがGCやCEに劣ることから、定量メタボロミクスでは十分に活用されていなかった。近年、高性能微細充填剤カラムに100 MPa以上の高压で安定して送液できるシステムが各社から開発された。さらに、高速スキャンが可能なESI飛行時間型質量分析計も続々と実用化されており、特に特殊技術を用いなくてもLC/MSを用いたメタボロミクスが可能となってきた。近年開発されたフーリエ変換イオンサイクロトロン質量分析計（FT-ICR-MS）などの高解像度質量分析計を用いて精密質量数を観測し、クロマトグラフィーによる分離を行うことなく、多数の代謝物を同時一斉分析する試みがなされている^{5,6)}。これらの方法論は、目的によってはきわめて有用である。

メタボロミクスにおけるデータ解析

膨大な変数を同時に扱うメタボロミクスにおけるデータの解析には多変量解析が必須である。多変量解析を行うためには、種々の分析手法により得られた生データ（主としてクロマトグラム）を数値データに変換する必要がある。メタボロミクスでは、当該第一ステップがきわめて重要となる。クロマトグラフィーなどにより観測したデータはピークを同定し、ピーク面積を積分することに

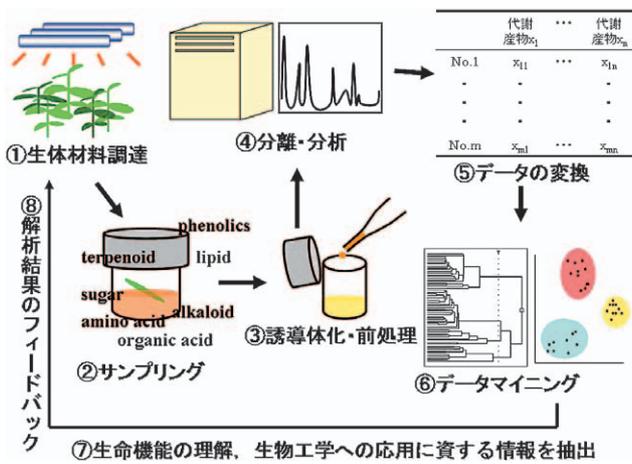


図1. メタボロミクスのスキーム

よりピークリストを作成すれば、それがそのまま多変量解析に適用可能なマトリクスになる。その場合、説明変数は代謝物名で、従属変数は各代謝物のピーク面積になる。分離が不十分な場合でもピークの重なりを多変量解析により分離することが原理的には可能である。GC/MSを用いた系においては、実際にコエリューションした代謝物ピークを質量分析情報からデコンボリューションする方法論が開発されており、抽出スペクトルを用いた高精度のピーク同定が可能になっている⁷⁾。

観測対象生物により、観測される代謝物の種類は異なる

るために、観測対象生物ごとにピークリストの整備が必要なのだが、ピークリスト作成は煩雑で時間を要するピーク同定作業を伴う。また、代謝改変などによりデッドエンドの代謝物に変化し、ピークリストにない代謝物が観測される事態が生じた場合、新たに生じたピークを見落とす可能性がある。危険回避のためにはマスクロマトグラムを目視によるチェックが安全確実であるが、技術的な問題が多く、一般的な方法ではない。

そこで、筆者らは、GC/MSクロマトグラムの結果をスペクトロメトリーと同様のデータ処理に供することによる解決策を開発した。具体的には、クロマトグラムの保持時間を独立変数とし、対応するピーク強度を従属変数としたクロマトデータ行列を作成し、サンプル間でデータセット（個々のサンプルのデータの数）標準化する作業を行い、ピークリストを作ることなく多変量解析を実施し、クラスター分離に寄与したピークを集中的に同定するシステムを開発した⁸⁾。当該操作は一般に、前処理の善し悪しとその後の解析の成否を左右することもしばしばある。前処理法は1) ノイズ除去法、2) ベースライン補正法、3) Resolution enhancement 法、4) 規格化法、などに分けることができる。一般的にスペクトルデータに用いられる代表的な前処理法として、スムージング、差スペクトル、微分処理、ベースライン補正、波形分離、中央化とスケールリングなどがある。当該方法は、代謝物

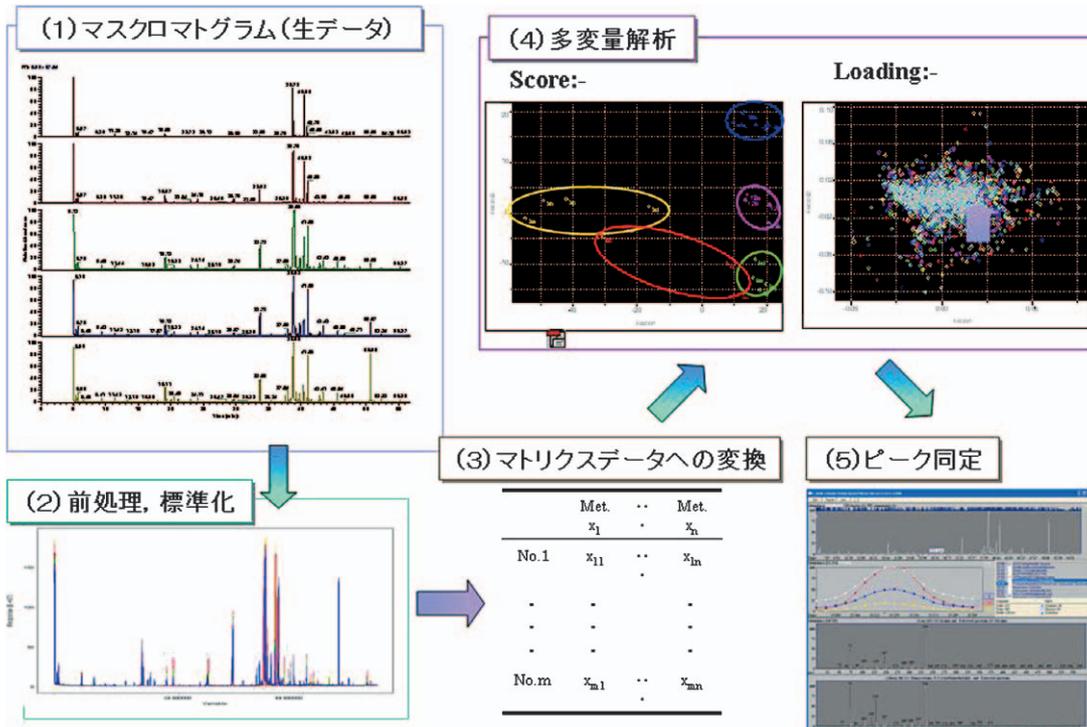


図2. データマイニングの基本的スキーム

フィンガープリンティングにも適応可能である(図2)。MoritzらのグループからもGC-MSクロマトグラムのデータ処理に関するすぐれた論文が発表されているので参照されたい⁹⁾。

多変量解析によるデータマイニング手法には重回帰分析、判別分析、主成分分析、クラスター分析、因子分析、正準相関分析などがあり、データ構造や解析の目的によって選択される。現在、メタボロミクスでもっともよく用いられている多変量解析手法は、探索的データ解析(exploratory analysis)であり、その目的は膨大な量のデータの特性を調査し、データが含んでいる情報の内容を判断することにある。また、探索的データ解析では、回帰分析や分類のモデルを構築する前にデータセットの可能性を確認できる。探索的データ解析の手法として主成分分析(principal component analysis; PCA)、および自己組織化マッピング(self organizing mapping; SOM)がもっとも頻繁に用いられる。図3は、シロイヌナズナ培養細胞T87に軽度の塩ストレス(100 mM食塩)を加えた後、経時的にサンプリングし、細胞内部メタボロームの経時変化を追跡し、親水性代謝物に注目してGC/MS分析を行い、結果を主成分分析(PCA)で解析したもので

である¹⁰⁾。主成分分析(PCA)は、なるべく少ない合成変数で、なるべく多くの情報を把握するという情報の縮約を目的とする。線形変換によりデータの次元数を削減する多変量解析の手法である。元の変数の線形関数の中で、元のデータの分散(ばらつき)をできるだけ保存する指標(主成分)を求める計算を行う。実際の計算は、前述のクロマトデータ行列の固有値問題の解答に他ならない。結果は固有値と固有ベクトルとして得られるが、固有ベクトルはクロマトグラフィーの各ピークにかかる係数ベクトルを意味し、固有値はその固有ベクトルの寄与率を意味する。固有ベクトルは、多くの場合、主成分あるいはファクターと呼ばれる。図3(A)は、経時ごとのサンプルの主成分得点を第1主成分(Factor1:横軸)と第2主成分(Factor2:縦軸)でプロットしたものである。この図からサンプルのばらつきの傾向ならびに類似性を知ることができる。本例の場合、サンプル(72時間)は、その他のサンプルと、第1主成分で明確にクラスター分離している。また、サンプル(24時間)は、他のサンプルと第2主成分でクラスター分離している。両者の分離にどのような代謝物が寄与しているかは、ローディング(主成分の係数)を見れば容易に類推することができる(図3)。Factor1からはサンプル(72時間)の特徴が、Factor2からはサンプル(24時間)の特徴がそれぞれ類推できる。主成分について、正に大きい係数と負に大きい係数に注目し、それらの係数に対応する代謝物を同定し、図中に記入した。たとえば、図3(B)のFactor2ではチロシン、シュクロース、トリプトファン、フェニルアラニンなどが正に大きい係数の代表であり、グリセロール、イノシトールなどが負に大きい係数の代表である。高得点クラスターを形成するサンプルは、正に大きい係数に対応する代謝物が多く、かつ負に大きい係数に対応する代謝物が少ないことにより、他サンプルとメタボロームレベルで分離したと考えられる。

食品工学への応用

メタボロミクスは、代謝物総体に基づく解析であるが、その方法論は生物以外をも対象として捕らえている。たとえば食品、食品原料なども解析対象として重要であり、それらを対象とした研究をフードメタボロミクスと称することがある。目的として、食品の品質鑑定、品質予測、品質を決定する要因解析、食品製造・保管工程の改善、品質保証システムの考案など多岐にわたる。その場合、探索的データ解析の他に、モデル化、予測のための多変量解析手法が時として用いられる。その中で、PLS回帰分析(partial least squares projection to latent structure:

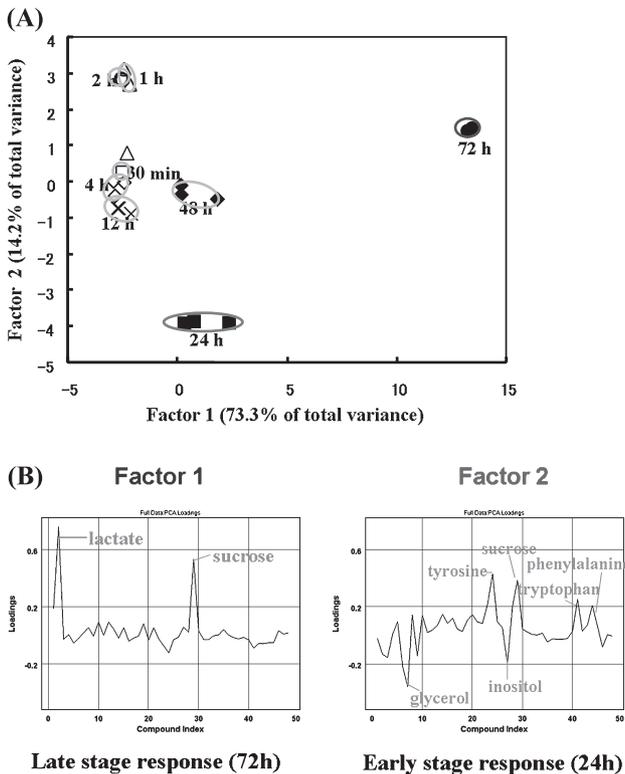


図3. シロイヌナズナ培養細胞(T87)に軽微な塩ストレスを加えたときのメタボロームの変化。(A) GC/MS分析結果を主成分分析(PCA)に供した際の主成分得点プロット。(B) 第1主成分(Factor1)と第2主成分(Factor2)のローディングプロット。(文献22の図を改変)

PLS)などの回帰分析手法が特に有用と思われる。

実例として、メタボロミクス技術を用いた緑茶の品質予測を紹介したい。製品緑茶の品質決定は熟練者による官能試験に委ねられる。鑑定は緑茶の熱湯抽出液の「味」、「香り」、「色」および緑茶葉の「外観」の総合得点(200点満点)によって行われる。4項目の中で最も重要なのは「味」である。そこで、味を形成する重要成分を網羅的に解析するために、茶葉から親水性低分子代謝物を抽出し、シリル化、メトキシ化の誘導体化を経て、GC/MS分析に供した。図4(A)は、高級茶と低級茶を比較した例である。GC/MS分析結果を行列データに変換し、主成分分析に供した結果を示した。第1主成分で明確に分離した。第1主成分ベクトルの係数を見ると、高級茶にはテアニン、キナ酸、リボース、アラビノースが多く含まれ、低級茶にはフルクトース、マンノース、スクロースなどの糖類が多く含まれていることが示唆された。同様の分析を奈良県緑茶品評会の出品サンプルに対して行い、GC/MS分析結果を行列データに変換後、PLS (projection of latent structure by means of partial least square) 回帰分析を行い、ランキング予測モデルを作製

した(図4B)。グラフ横軸は予測ランキング、縦軸は実測ランキングを示す。GC/MSデータだけでもある程度のランキング予測が可能であることが示唆された¹¹⁾。加えてLC/MS, FT-NIR, FT-NMRなどによるプロファイリングも研究されており、さらなる発展が期待される。

おわりに

メタボロミクスが将来有望な技術であることは間違いないが、技術的に発展途上であり、標準的な運用方法確立されていない。当然のことながら、メタボロミクスをモデル生物を使った機能ゲノム学のツールと限定して考える必要はまったくない。本稿で示したとおり、メタボロミクスの手法を食品分析、生薬分析に応用し、品質予測、製造工程改善、保管工程最適化への応用が可能である。また、生物生産応用研究の基礎スクリーニングにもきわめて有用である。たとえば複数の基質を複数の酵素で反応させる複雑系酵素反応混合物などは好適なサンプルである。メタボロミクスの手法によれば、数百種類の酵素生成物混合物を同時一斉分析し、データマイニングすることが可能である。今までブラックボックスだった複雑な酵素反応の様式を解明する手がかりを掴めるのではないかと確信している。メタボロミクスという確立されていない方法論を研究ツールに用いることを躊躇される方が多いと思われるが、未確立の今だからこそ、独自の運用方法を開発し、競争者に先んじてご自分の研究を大きく推進できる可能性がある。数多くのバイオサイエンス、バイオテクノロジーの研究者がメタボロミクスに興味を示し、ツールとして使われることを期待したい。

本研究の一部は、文科省科学研究費特定領域「ライフサーベイヤ」、科学技術振興機構・地域結集型共同事業「古都奈良の新世紀植物機能開発」のサポートにより行われた。

文 献

- 1) Fukusaki, E. and Kobayashi, A.: *J. Biosci. Bioeng.*, **100**, 347 (2005).
- 2) Soga, T. *et al.*: *Anal. Chem.*, **74**, 6224 (2002).
- 3) Soga, T. *et al.*: *Anal. Chem.*, **74**, 2233 (2002).
- 4) Harada, K. *et al.*: *J. Biosci. Bioeng.*, **101**, 403 (2006).
- 5) Aharoni, A. *et al.*: *Omics*, **6**, 217 (2002).
- 6) 及川 彰ら: *生物工学*, **84**, 219 (2006).
- 7) Halket, J. M. *et al.*: *Rapid Commun. Mass Spectrom.*, **13**, 279 (1999).
- 8) Fukusaki, E. *et al.*: *Z. Naturforsch., C*, p. 267 (2006).
- 9) Jonsson, P. *et al.*: *Anal. Chem.*, **76**, 1738 (2004).
- 10) Kim, J. K. *et al.*: *J. Exp. Bot.*, **58**, 415 (2007).
- 11) Pongsuwan, W. *et al.*: *J. Agric. Food Chem.*, **55**, 231 (2007).

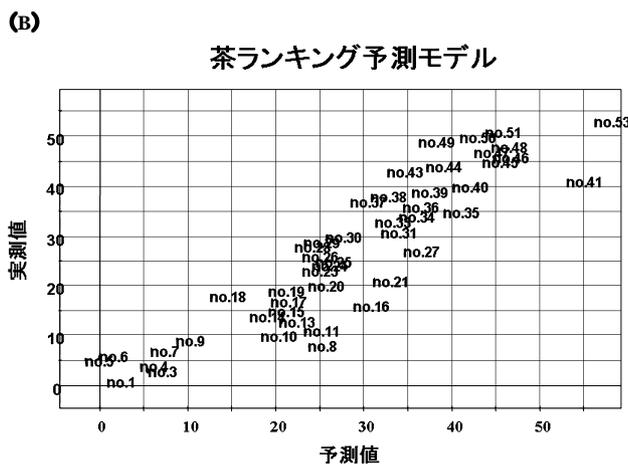
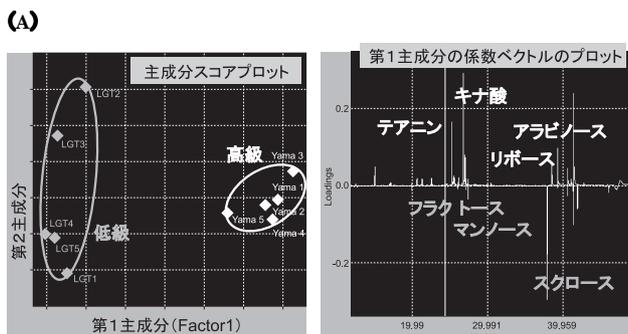


図4. GC/MS分析による緑茶品質予測。(A) 高級茶と低級茶の比較。(B) 製品緑茶品評会サンプルのランキング予測。(文献11の図を改変)