

仮説なしの多変量解析によるNMRメタボロームデータの要約

関連製品: 核磁気共鳴装置(NMR)

近年、NMRによって代謝産物の総体(メタボローム)を混合物のまま解析するNMR代謝プロファイリング(NMR metabolic profiling; NMR-MP) がポストゲノム研究や品質管理などで盛んに用いられています。NMR-MPIは、ノンターゲットメタボロミクスの一手法です。ノンターゲットの解析では、解析の対象が広範にわたるため、NMRスペクトルから直接データを解釈するのは困難です。多変量解析を適用することで、多変量データから有用な知見を発掘できます(データマイニング)。

データマイニングでは、実験系および目的に合わせてデータの要約、分類、モデリング、回帰(予測)などの適切な多変量解析を適用することが重要です。ここでは、仮説なし*(試料の群情報や外部変数を使わない)でデータを要約する手法である“主成分分析”と“階層的クラスター解析”をNMRメタボロームデータに適用した例を紹介します。

試料と方法

6種類の市販の野菜のsprout(ダイコン、ブロッコリ、マスタード、アルファルファ、エンドウ、オクラ)から、Bligh-Dyer法に準ずる方法で代謝産物を抽出し、極性画分を遠心エバポレータで乾固しました。乾固した極性代謝産物を、重水で調製したpH = 7.0の100 mM リン酸カリウム緩衝液で再溶解させ、NMR試料としました。¹H-NMR計測には、2次元J分解分光法を用いました。J分解スペクトルのf2プロジェクションからは、 J_{HH} による分裂のない(Jカップリングの消去された)化学シフトのみのスペクトルが得られるため[1]、スペクトルを単純化させることができます(Figure 1)。それぞれのsproutにつき3試料を調製しました。

それぞれのスペクトルの0.5-9.0 ppmの範囲を、等幅の積分範囲(0.02 ppm幅)で積分しました(バケット積分)。ここでノイズレベルの変数は、多変量解析に供する前に削除しました。結果として101変数* 18試料からなる行列(多変量データ)が得られました。主成分分析および階層的クラスター分析は、R言語の統計解析関数(無料)を用いて計算しました。

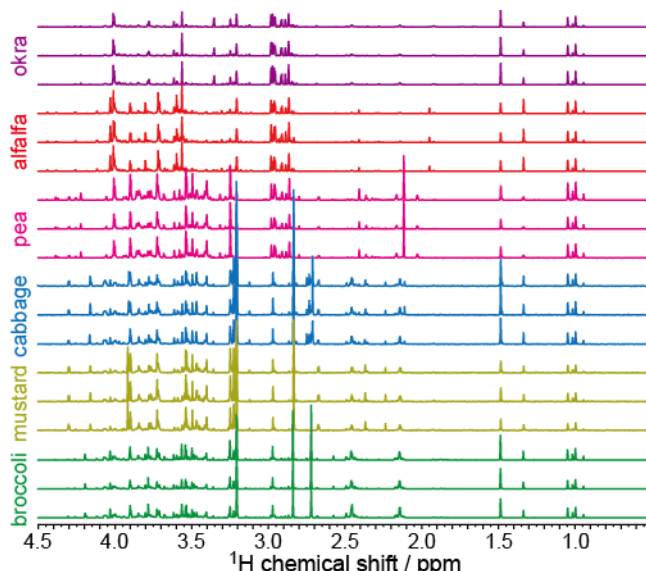


Figure 1. sprout抽出極性代謝産物¹H-NMRスペクトル。J分解分光法のf2プロジェクションを用いた。NMR分光計はJNM-ECZ400Sを、検出器はROYALプローブを用い、ASC30オートサンプルチェンジャーを用いて連続測定をおこなった。

主成分分析の適用

主成分分析では、与えられた多変量から分散(≒情報量)が最大となるように各変数の線形結合として主成分を作成します。その結果、多次元のデータ(多変量)をより少ない次元で説明できるようになります。式 1 の線形結合式から主成分分析におけるスコアと因子負荷量の関係を示します。通常、中央化(各変数から平均値を差し引く処理)と、必要に応じてスケールリングをおこないます。

元の変量(中央化およびスケールリングされた)

$$t_{ij} = \bar{x}_{i1} w_{j1} + \bar{x}_{i2} w_{j2} + \dots + \bar{x}_{in} w_{jn} \quad (1)$$

スコア

因子負荷量

式 1 から、スコアと因子負荷量の対応から試料間の特徴を解釈できます。 i は、試料番号を、 j は変数番号または主成分数を表します。ここでは、固有ベクトルを因子負荷量としています。

主成分分析の結果をFigure 2に示します。アブラナ科のsproutであるダイコン、ブロッコリ、マスタードの第一主成分 (PC1) のスコアは負であり、それ以外のアルファルファ、エンドウ(マメ科)、オクラのスコアは正でした (Figure 2 (a))。負荷因子量では、アスパラギンなどがPC1の正に、グルタミンなどが負に寄与していることがわかります (Figure 2(b))。

sproutなどの若い植物では、アスパラギンやグルタミンなどの側鎖にアミノ基を有するアミノ酸が窒素代謝において重要な役割を担うことが知られています[2]。これらの組成は、植物の種や環境によって異なります。例えばアブラナ科の植物ではグルタミンが、マメ科の植物ではアスパラギンが側鎖にアミノ基を有するアミノ酸では主要であることが報告されています [3]。また、光合成ができない環境では、よりC/N比の小さいアスパラギンを選択的に合成することが知られています [4]。

階層的クラスタ分析の適用

階層的クラスタ分析は、各試料あるいは変数間の距離を定義し、距離が近いものを逐次的にクラスタとして連結することで、データ間の関係性をデンドログラム(樹形図)として要約する手法です。ヒートマップと併用することで、試料と代謝産物の関係性の解釈が容易になります。データ間の距離の定義として、ユークリッド距離(式 2)が一般的ですが、ここでは、ピアソンの積率相関係数に基づいた相関距離(式 3)を用いました。

$$dist_{euclidean}(A, B) = \left[\sum_{i=1}^n (a_i - b_i)^2 \right]^{1/2} \quad (2)$$

$$dist_{corr}(A, B) = (1 - corr(A, B))/2 \quad (3)$$

ここで a_i および b_i は、それぞれAおよびBの i 番目の変数、 $corr(A, B)$ はピアソンの積率相関係数を表します。

階層的クラスタ分析から得られたヒートマップをFigure 3に示します。ヒートマップの行および列方向は、それぞれ変数と試料に対応します。左および上部に示すデンドログラムは、それぞれ変数間および試料間の相関距離を階層構造で示しています。試料間のデンドログラムでは、主成分分析同様に、アブラナ科のsproutとそれ以外が分けられました。変数間のデンドログラムでは、同じ分子に由来する信号および異なる分子だが相関が高い信号がクラスタリングされています。緑～赤のグラデーションで示されたヒートマップは、デンドログラムに従って並べ替えされており、各試料および試料群に特異的に存在する代謝産物が可視化されています。例えば、イソチオシアネートのスルフォラファンが、ブロッコリに多く含まれており、グルコシノレート的一种であるシニグリンがマスタードに多く含まれていることがわかります [5]。

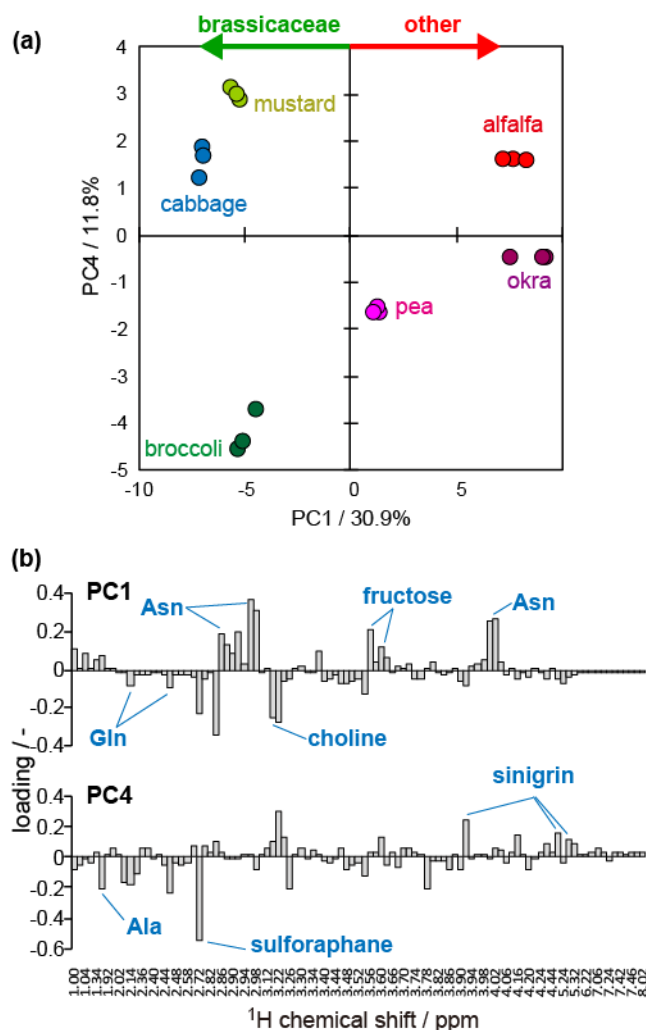


Figure 2. sprout抽出極性代謝産物¹H-NMRスペクトルの主成分分析。(a)スコアプロット(b)PC1およびPC4の因子負荷量。スペクトルの総和でノーマライズし、中央化したデータから主成分分析をおこなった。

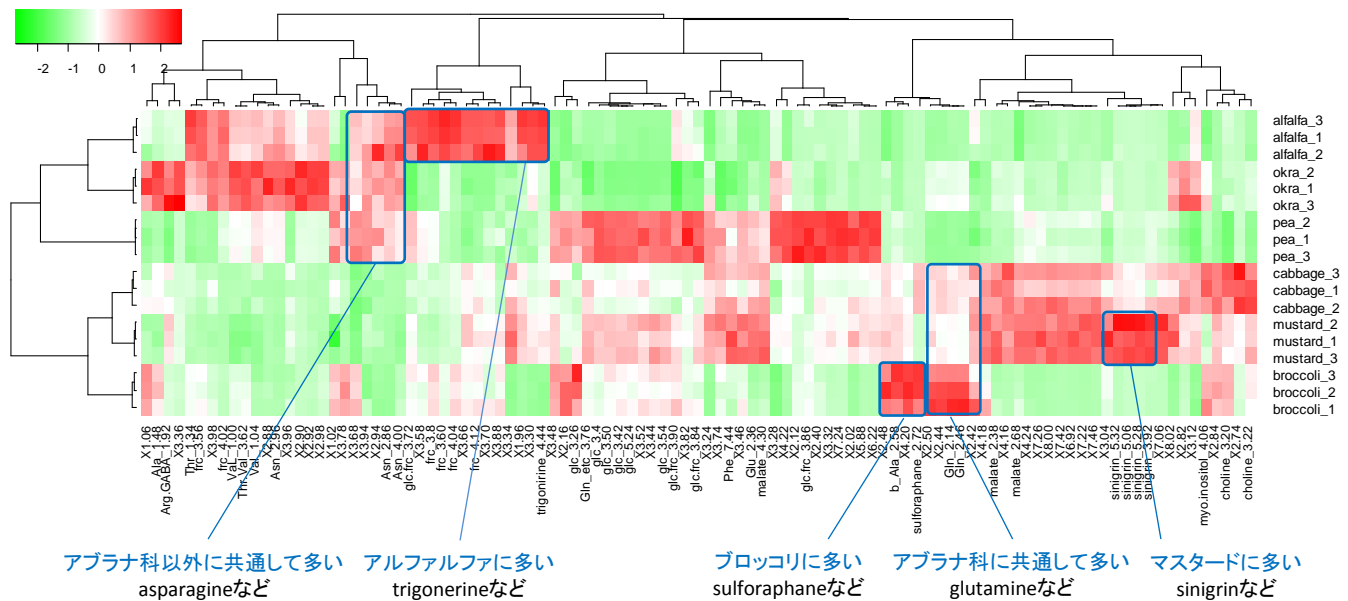


Figure 3. スプラウト抽出極性代謝産物の¹H-NMRスペクトルに基づくヒートマップおよび階層的クラスター分析。行および列は変数および試料に対応する。変数間および試料間では相関距離を用い完全連結法でクラスターを形成した。ヒートマップには、オートスケールされたデータを用いた。よって試料間で相対的に存在量が多い代謝産物と少ない代謝産物が赤と緑の濃淡で示される。ヒートマップからは、各試料に特有な代謝産物の情報が視覚的に得られる。glc: glucose、frc: fructose、アミノ酸は三文字略号で記す。

NMR-MPの特徴として、多くの質量分析法とは異なりクロマトグラフィなどによる分離操作をおこなわない点、再現性が非常に高い点が挙げられます。質量分析法によるメタボローム解析の多くは、対象とする分子種に合わせて固相、移動相、イオン化法、代謝産物抽出法の選択などの検討が重要です。一方、NMRでは分離操作がなく、溶液に溶けたほぼすべての分子が検出対象になるため、実験系の検討の時間と手間を省くことができます。また、分析試料のキャリアオーバーの心配もありません。選択性や感度が求められる詳細な代謝の解析には、質量分析法が第一選択肢となります。仮説などが無い状態から、データを計測して、教師なしの多変量解析を“とりあえず”おこなうfirst (or fast) screeningとしてNMR-MPは効果的です。

* 仮説なしの多変量解析を教師なし多変量解析とよびます。一方、仮説(試料の群情報や外部変数)を利用する多変量解析を教師あり多変量解析とよびます。

参考文献

- [1]弊社アプリケーションノート“NMR代謝プロファイリングなどにおける混合物の¹Hスペクトルの単純化”(NM170014), [2] Komatsu; T. et al., *Metabolites* (2014) **4**, 1018. [3] Nord; F. F. and Street; H. E., “Nitrogen Metabolism of Higher Plants” in *Advances in Enzymology* (1949) **9**, John Wiley & Sons, Inc., 391. [4] McGrath; R.B. et al., *Plant J.* (1991) **1**, 275. [5] Rangkadilok; N. et al., *Sci. Hort.* (2002) **96**, 11.

本誌の記載内容は予告なしに変更することがあります。

Copyright © 2017 JEOL Ltd.

このカタログに掲載した商品は、外国為替及び外国貿易法の安全輸出管理の規制品に該当する場合がありますので、輸出するとき、または日本国外に持ち出すときは当社までお問い合わせ下さい。

JEOL 日本電子株式会社

本社・昭島製作所

〒196-8558 東京都昭島市武蔵野3-1-2 TEL: (042) 543-1111(大代表) FAX: (042) 546-3353

www.jeol.co.jp ISO 9001・ISO 14001 認証取得

東京事務所 〒100-0004 東京都千代田区大手町2丁目1番1号 大手町野村ビル13階

営業企画室 TEL: 03-6262-3560 FAX: 03-6262-3577

電子光学機器営業推進室 TEL: 03-6262-3567 分析機器営業推進室 TEL: 03-6262-3568

産業機器営業部 TEL: 03-6262-3570 医用機器ソリューション販売室 TEL: 03-6262-3571

東京支店 〒100-0004 東京都千代田区大手町2丁目1番1号 大手町野村ビル18階 TEL: 03-6262-3580 FAX: 03-6262-3588

電子光学機器営業グループ TEL: 03-6262-3581 分析機器営業グループ TEL: 03-6262-3582

医用機器営業グループ TEL: 03-6262-3583

東京第二事務所 〒190-0012 東京都立川市曙町2丁目8番3号 新鈴舎ビル9階

半導体機器営業室 TEL: 042-528-3491 ソリューションビジネス部 TEL: 042-526-5098

横浜事務所 〒222-0033 神奈川県横浜市港北区新横浜3丁目6番4号 新横浜千歳ビル6階 TEL: 045-474-2181 FAX: 045-474-2180

海外事業所・営業所 Boston, Paris, London, Amsterdam, Stockholm, Sydney, Milan, Singapore, Munich, Beijing, Moscow, Sao Paulo ほか

札幌支店 〒060-0809 北海道札幌市北区北9条西3丁目19番地 ノルテプラザ5階

仙台支店 〒980-0021 宮城県仙台市青葉区中央2丁目2番1号 仙台三菱ビル6階

筑波支店 〒305-0033 茨城県つくば市東新井18番1号

名古屋支店 〒450-0001 愛知県名古屋市中村区新緑1丁目47番1号 名古屋国際センタービル14階

大阪支店 〒532-0011 大阪府大阪市淀川区中島5丁目14番5号 ニッセイ新大阪南口ビル11階 TEL: 06-6304-3941 FAX: 06-6304-7377

西日本ソリューションセンター

〒532-0011 大阪府大阪市淀川区中島5丁目14番5号 ニッセイ新大阪南口ビル11階 TEL: 06-6305-0121 FAX: 06-6305-0105

広島支店 〒730-0015 広島県広島市中区橋本町10番6号 広島NSビル5階

高松支店 〒760-0023 香川県高松市寿町1-1-12 パシフィックシティ高松5階

福岡支店 〒812-0011 福岡市博多区博多駅前2丁目1番1号 福岡朝日ビル5階

TEL: 011-726-9680 FAX: 011-717-7305

TEL: 022-222-3324 FAX: 022-265-0202

TEL: 029-856-3220 FAX: 029-856-1639

TEL: 052-581-1406 FAX: 052-581-2887

TEL: 06-6304-3941 FAX: 06-6304-7377

TEL: 06-6305-0121 FAX: 06-6305-0105

TEL: 082-221-2500 FAX: 082-221-3611

TEL: 087-821-0053 FAX: 087-822-0709

TEL: 092-411-2381 FAX: 092-473-1649