

オミックス統合解析を用いたダイズ種子発生の代謝研究

澤田有司*

国立研究開発法人理化学研究所環境資源科学研究センター

Omics for Metabolism Study for Soybean Seed Development

Yuji SAWADA*

*RIKEN Center for Sustainable Resource Science, Yokohama 230-0045

ABSTRACT

Dynamic metabolic changes have been estimated in soybean seed development and fungal inoculation. In this study, soybean accession "Fukuyutaka", which is number one production in Japan, was used for our omics approaches. Our goal is integrated omics approaches for elucidation of metabolic change and its regulation. Generally, primary metabolites for plant physiology are known metabolites, which can purchase. In plant specific secondary metabolites, very limited standard metabolite can purchase. Thus, we need un-targeted metabolomics for detection of unknown structural metabolites. Using LC-QTOF-MS, molecular weight related ion (MS1) and its fragmented ions (MS2) were automatically detected for all detectable metabolites. The MS1 and MS2 information can search for MS and MS2 database. Using dry seed, seed development samples (day after seedling 1-8: DAS1-8) and fungal inoculated at DAS2, total 51 samples were applied for the un-targeted metabolomics. As the results, 472 metabolites candidate were detected as MS1 and MS2. The annotation rate based on the database search was more than 70%. The metabolic profiles were clearly shown the diverse and dynamic metabolic changes. In multivariate analysis, DAS3 was identified as a key point for the metabolic change. Key metabolic profiles in principal component 1 (PC1) of principal component analysis was significantly change based on the seed development. Key metabolic change in PC2 was significantly change based on the fungal inoculation. These metabolic changes also confirmed in widely targeted metabolomics based on the standard metabolites. To elucidate the co-accumulation of metabolites, we carried out correlation network analysis. We successfully found an isoflavone group, which was including genistin (genistein glycoside). In future study, gene expression

* 〒230-0045 横浜市鶴見区末広町1-7-22

profiles will be collected by using RNA sequence technology. The integrated omics approaches will give us overview of soybean metabolic changes in seed development and fungal inoculation. *Soy Protein Research, Japan* **18**, 28-34, 2015.

Key words : seed development, fungal inoculation, un-targeted metabolomics, multivariate analysis, RNA sequence.

ダイズ種子およびその加工品は多様な代謝産物を含む機能性食品として注目されている。特にダイズに特徴的なイソフラボン類の量や質の制御機構の解明と代謝育種への応用利用はダイズ代謝研究の中心課題の1つである。本研究ではダイズの種子発生と代謝の関連性に注目し、オミックス統合解析を試みた。

ゲノム、トランスクリプトーム、プロテオーム、メタボロームなどの広範囲な検出情報を利用するオミックス解析では複雑な生命現象の全体像をシステムとして解析可能である。さらに、これらのオミックス解析を統合利用すればより確度の高い現象解明が可能になると期待されている^{1,2)}。本研究では、メタボロミクスとトランスクリプトミクスを種子発生過程の解析に統合利用し、代謝変動の検出とその制御に関連する遺伝子発現の制御機構の解明を目指した。

ダイズ種子発生では可視の形態変化に伴い、代謝産物の量と質も大きく変動する。これまでに発芽ダイズの糖と糖代謝関連酵素に器官特異性(子葉, 胚軸)があることが報告されている³⁾。同様にマメ科植物に特徴的なイソフラボンも器官特異性(子葉, 胚軸, 上胚軸, 幼根)が調べられている⁴⁾。さらに、微生物処理したダイズでは、イソフラボン誘導体(グライセオリン)の増加することが報告されている⁵⁾。一方、ダイズ種子の抽出物をコウジカビで処理した発酵産物は、イソフラボンが顕著に減少する⁶⁾。これらの発芽ダイズおよび微生物発酵では、特定の代謝産物が短期間に数倍から数十倍以上変化する。これらの代謝変動をオミックス統合解析で検証すれば変動する代謝産物の生合成機構と制御機構が一挙に解明可能になる。

ダイズ種子発生では、成長生理に必須な一次代謝産物(アミノ酸, 糖, 有機酸など)とダイズに特徴的な二次代謝産物(イソフラボンなど)の変動が想定されている。一次代謝産物は標準化合物に基づく高感度検出が可能なターゲット解析で検出可能である。我々はこのターゲット解析を拡張したワイドターゲット解析を開発し、500種以上の代謝産物の検出を可能にしている^{7,8)}。一方、より多様な構造を示す二次代謝産物の多くは文献上の検出結果の報告はあるが、標準化合物

の入手が困難である。この未同定代謝産物の検出が可能な非ターゲット解析が提唱されている。非ターゲット解析では、検出可能な代謝産物の分子量関連イオン(MS1)とそのフラグメント化イオン(MS2)を取得し、標準化合物のMS1とMS2情報を検索することで構造の推定を行う⁹⁾。本研究では既知および未知の代謝産物の検出に対応可能なこれらの二種の手法を利用し、広範囲な代謝産物プロファイルを発芽ダイズで取得した。さらに、微生物の処理で一過的に誘導される二次代謝産物を検出するために発芽後にコウジカビを接種した発芽発酵ダイズも同様に調べた。

これらの代謝プロファイルは、種子発生と微生物発酵の過程によるものであり、これらの制御関連遺伝子は代謝育種に利用可能な有用遺伝子候補である。そこで、代謝プロファイルを得たサンプルを利用し、広範囲な遺伝子発現を得られるRNAシーケンスを行った。

方 法

発芽ダイズおよび発芽発酵ダイズの調製

ダイズ種子は、国内生産量1位のフクユタカの乾燥種子を利用した。発芽ダイズは播種後1-8日目(Day after seedling 1-8, DAS1-8)を調整した。発芽発酵ダイズはDAS1およびDAS2にコウジカビを接種したサンプルを調整した。DAS1に接種した発酵ダイズは、著しく成長が阻害されたため、以降の実験には利用しなかった。本研究では、乾燥ダイズ、発芽ダイズ(DAS1-8)、発酵ダイズ(DAS1-8, DAS2で菌を接種)の計17種の実験区を設定した(Table 1)。これらの実験区は3回の生物学的な独立実験を行い、代謝変動を検証した。

分析試料の作成方法

乾燥種子、発芽ダイズ、発芽発酵ダイズは12時間凍結乾燥後、ミルミキサーで粉末化した(Tube Mill, IKA社)。このダイズ種子粉を4 mg秤量し、1 mLの抽出バッファー(80% MeOH, 0.1%ギ酸)を加えて直径

Table 1. Experimental design for soybean seed development and fermentation

Group ID	Development	Inoculation
G01	Seed	non
G02	DAS01	Aspergillus oryza
G03	DAS02	Aspergillus oryza
G04	DAS03	Aspergillus oryza
G05	DAS04	Aspergillus oryza
G06	DAS05	Aspergillus oryza
G07	DAS06	Aspergillus oryza
G08	DAS07	Aspergillus oryza
G09	DAS08	Aspergillus oryza
G10	DAS01	non
G11	DAS02	non
G12	DAS03	non
G13	DAS04	non
G14	DAS05	non
G15	DAS06	non
G16	DAS07	non
G17	DAS08	non

5 mmジルコニアビーズ（アズワン社）を入れたのち、1500 rpm, 5 minの条件で抽出した。抽出液を15,000 rpm, 10 min の条件で遠心分離し、上清を以降のサンプル処理に用いた。

抽出物の前処理

抽出液を含んだ2 mLチューブを自動分注システム（Microlabstar, ハミルトン社）にセットし、1/100に希釈した後に窒素ガスで抽出溶媒を蒸発乾固させた。さらに、LC-MS用の超純水で225 μ Lで再度溶解したのちに、384 well形式の0.45 μ mのフィルター（Whatman社）でろ過した。

サンプルの分析

抽出サンプルは、液体クロマトグラフィーにタンデム四重極飛行時間型質量分析装置をハイフネーションしたLC-QqQ-MS（UPLC-TQS, Waters社）、LC-QTOF-MS（LC, 島津製作所；QTOF-MS Bruker社）で分析した。LC-QqQ-MS解析では、分析結果をマスリンクス（Waters社）で解析し、検出対象ごとのMSクロマトグラムのピークエリアを算出し、Perl scriptでサンプル×代謝産物のデータ行列を作成した。非ターゲット解析の分析結果はMS-DIAL（<http://prime.psc.riken.jp/>）で解析した。MS DIALは未同定代謝産物を含むデータ行列からMS1とMS2の情報を含んだデータ行列を作成するアライメント処理を行った後に、MS/MSデータベースを参照し、フラグメント化イオンの

m/z 値と強度の類似度から構造推定を行った。

データ解析

各サンプルの分析結果から得られたメタボローム情報を独立の変数として統計ソフト（R, <http://www.r-project.org/>およびRstudio）で解析した。ノーマライズ（Log2変換した後、Z-score化）した各Area値を統計解析と多変量解析（heatmap, principal component analysis, Pearson correlation coefficient analysis）に利用した。

結 果

非ターゲットデータの取得とMS-DIAL によるアノテーション

本研究ではTable 1に示した17の実験区を3反復分析し、計51サンプルをLC-QTOF-MSを用いた30 Hzのデータ取得速度で3分間分析した（30 Hz×60 sec×3 min×51サンプル＝最大275,400データポイント）。この非ターゲット解析では分子量関連イオン（MS1）とこれをフラグメント化したイオン（MS2）を同時に取得できる。まずMS1の強度情報に基づく閾値を設定し、51サンプルを混合した品質チェックサンプルの代謝産物候補を472種選択した。さらに、MS2およびMS1に基づくアノテーションでは7割の検出イオンが標準化合物のデータベース情報のMS1またはMS2と一致した（Table 2）。これらの時系列をまとめたFig. 1では乾燥種子、発芽ダイズ、発芽発酵ダイズで顕著な代謝パターンを示すことが分かった。

Table 2. Summary for un-targeted metabolomics

	Annotation level					Total
	MS2			MS1		
	A	B	C	B	C	
Positive	111	123	68	17	13	332
Negative	2	87	40	8	3	140
Total	113	210	108	25	16	472

Annotation levels: A, MS2 based annotation; B, MS1 based annotation; C, No match.

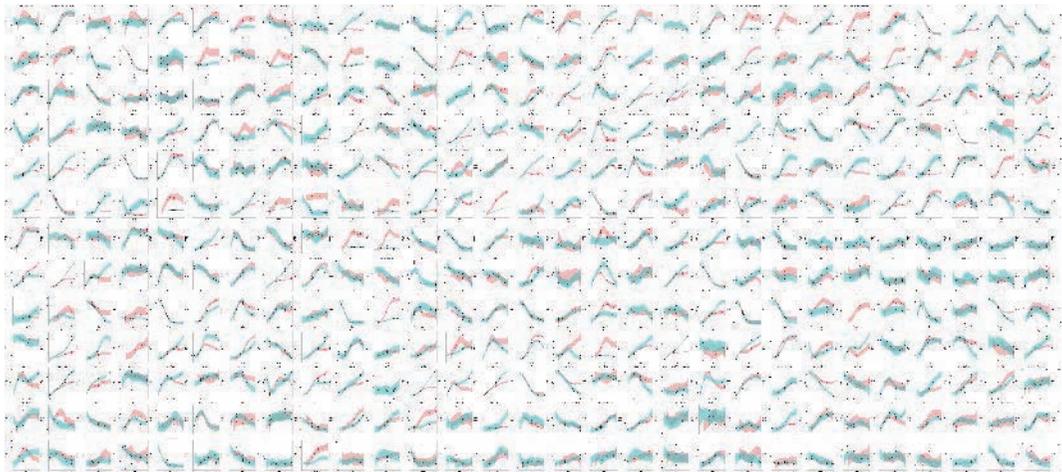


Fig. 1. Metabolic profiles in soybean development and fungal inoculation.

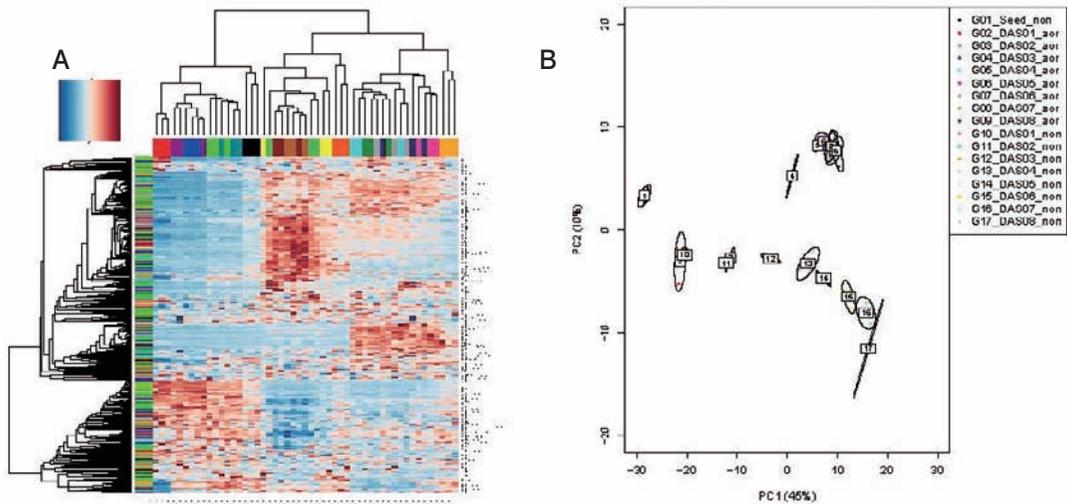


Fig. 2. Multivariate statistics for soybean metabolomics.
 A. Heatmap analysis: x axis was samples and y axis was metabolites.
 B. Principal component analysis.

多変量解析結果

メタボローム研究の結果は多変量情報であるため、単変量を対象とした統計解析では結果の解釈が困難である。そこで、多変量解析である階層クラスタリングと主成分解析を行った (Fig 2)。階層クラスタリングの結果から、DAS1-2とDAS3-4、DAS5-8で顕著な代謝産物の変化が見込まれた (Fig 2A)。同様に主成分解析でも顕著な代謝変化が示唆された (Fig 2B)。特に、DAS3は発芽サイズと発酵サイズの代謝パターンが顕著に変化する変曲点であった。

検出情報の集計

本研究では、多変量解析で顕著な変化が検出できる発芽サイズと発酵サイズを利用したため、各実験 (乾燥種子、DAS1-8、DAS1-8でDAS2に菌を接種したサンプル) で大きく検出代謝産物数が変化することが予測された。そこで、各実験区の検出数を比較した。抽出溶媒のみをコントロールとしてシグナルとノイズ比が5より大きく、各実験区のrelative standard deviationが30%以下の化合物数を比較すると乾燥種子に比べてDAS3は乾燥種子の1.5倍の検出数を示した。

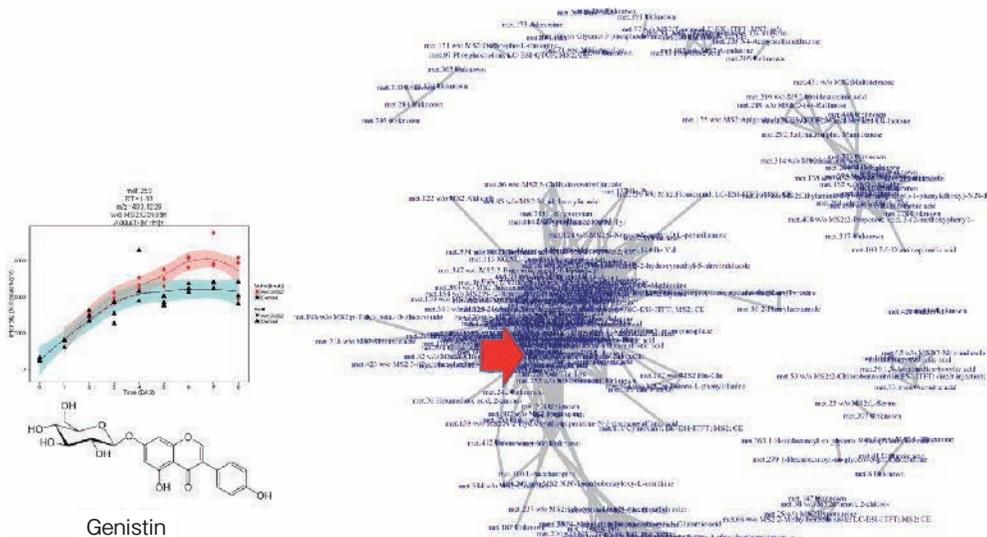


Fig. 3. Pearson correlation coefficient network analysis.

代謝産物の共蓄積ネットワーク解析

二次代謝では、合成関連遺伝子とその代謝産物が同調的に増減すると考えている。そこで、ダイズに特徴的なイソフラボンが含まれる共蓄積ネットワークの存在をピアソンの相関値を指標に検証した。検出した代謝産物間の相関値を総当たりで計算した後に各代謝産物をノードとし、各相関値をエッジとしてネットワーク図を作ると、イソフラボン配糖体を含むネットワークが示された (Fig. 3)。このゲニステイン配糖体は発芽の過程で増加すること、発芽発酵サンプルではより高蓄積することが示されている (Fig 3)。

考 察

非ターゲットのメタボローム解析は、未同定の代謝産物を含む検出可能な全代謝産物のMS1とMS2の情報得られる。この情報を閾値 (強度とRSD) で選択した後に、MS-DIALの機能でMS1およびMS2のデータベースで検索した。この結果、73%の検出イオン (368/472) が既知の構造と類似していた。従来手法では検出情報の10%程度をマニュアル操作で注釈付けしていたため、大規模情報への対応が困難であった。予備的な検討では同手法を利用して100検体以上の非ターゲット解析に成功している。このアノテーション情報に基づいて各検出情報を精査すればダイズの種子発生の代謝変化の全容が解明できる。また、本報告書

ではフクユタカの乾燥種子、発芽ダイズ、発芽発酵ダイズの結果を示したが、すでに他のダイズ品種 (黄豆、黑豆、リポ欠品種など) を利用した分析に着手している。これらの結果から発芽ダイズと発芽発酵ダイズの多様性を利用した特定代謝産物の高効率な生産が実現できると期待している。

非ターゲットのメタボローム解析では数百以上の検出情報が得られるが、個々の現象の解明では全体像の把握は困難である。そこで各検出情報を独立の変数と仮定した多変量解析を行った。ヒートマップ解析では発芽の前半のサンプル (DAS1-2) は発芽ダイズと接種直後の発芽発酵ダイズに大きな差は見られないが、DAS3以降では顕著な変化が見られた。主成分解析でも同様にDAS3を分岐点として発芽ダイズと発芽発酵ダイズの代謝プロファイルが変化していることが示された。また、第一、第二主成分に寄与する代謝産物の時系列変化を調べた結果、第一主成分の正の方向に寄与する代謝プロファイルはDAS3以降に顕著な増加がみられること、負の方向に寄与する代謝プロファイルはDAS3以降で減少する傾向がみられることが分かった。さらに、第二主成分に寄与する代謝産物も正方向では発芽発酵ダイズで増加する代謝産物、負方向では発芽ダイズでは増加するが、発芽発酵ダイズでは減少する代謝産物であった。このように発芽ダイズと発芽発酵ダイズを利用したメタボローム解析ではダイズが本来示す代謝変動に加えて新たな多様性を発見できる

可能性が示された。すなわち、発芽と発酵を時系列解析し、この結果を利用すれば特定代謝産物の制御機構が解明できる。

本報告書では主に、非ターゲットメタボローム解析の結果を紹介したが、我々は同一サンプルでRNAシーケンスを行っている。これらの統合解析が実現できれば、各代謝産物の生合成遺伝子も一挙に解明できると期待している。さらに、発芽ダイズは出発材料の増加で比較的容易にスケールアップが可能である。今回検出したイソフラボンなどの主要成分を分画すれば従来は入手が困難であったglyceollinなどの高度に修飾されたイソフラボンの単離精製が可能になる。また、これらの既知構造を利用して各イソフラボンに特徴的に検出されるMS2情報を利用すれば、より高精度なアノテーションが実現できる。

最後に、本研究で挑戦した発芽および発芽発酵ダイ

ズのオミックス統合解析は再現性の高い時系列解析が可能であるため、ダイズ品種や発酵に利用する菌の組み合わせや発酵条件の検討でダイズ代謝の変動を最大限に利用できる可能性がある。今後は、他の植物種にも応用可能であるか検討する予定である。

謝 辞

メタボローム情報の取得では環境資源科学研究センターの技術員である坂田あかね（サンプル抽出）、佐藤心郎（LC-QqQ-MS分析）、明石寛道（LC-QTOF-MS分析）、山田豊（情報処理）の支援を得た。またMS-DIALの利用では同センターの津川祐司博士の支援を得た。生豆、発芽ダイズ、発芽発酵ダイズの調製では果実堂社の落合孝次氏、桑原杏奈氏、荒木涼子氏の支援を得た。

要 約

発芽ダイズに菌を接種した実験では、形態変化と発酵に伴う大きな代謝変動が期待できる。そこで、本研究では国内生産が最も多いフクユタカの代謝プロファイルと発現プロファイルを時系列解析し、特定代謝産物の増減とそれを制御する遺伝子の解明を試みた。植物の成長生理に共通する一次代謝産物は構造が既知であるが、ダイズに特異的かつ特徴的な二次代謝産物の多くは標準化合物の入手が困難である。本研究では標準化合物の高感度検出条件を利用するワイドターゲット解析に加え、あらかじめ検出対象を選択せずに検出範囲を指定して分析する、非ターゲット解析を行った。生豆、発芽ダイズ（播種後1-8日間培養：DAS1-8）、発芽発酵ダイズ（播種後2日目に菌を接種し、8日目まで培養）を生物学的な3反復の分析を行い、計51サンプルのメタボローム解析を行った。この結果、472種の代謝産物候補の検出情報が得られた。この非ターゲット解析では各代謝産物の分子量関連イオン（MS1）とフラグメント化イオン（MS2）が同時に得られる。このMS情報に基づくデータベース検索を行った結果、7割以上の検出イオンにアノテーション（推定構造情報）が付与できた。このメタボローム情報は、生豆から発芽の過程で増減する化合物と発芽と発酵により増減する多様な代謝プロファイルが得られた。各代謝産物の蓄積量を独立の変数として行った多変量解析ではDAS3で顕著な変曲点を示された。主成分解析の傾向では発芽の過程で増減する代謝産物が第一主成分に寄与し、発酵で増減する代謝産物が第二主成分に寄与することが分かった。この結果、教師なしの多変量解析でも各実験区を明確に区別できることが分かった。これらの顕著な代謝変動は、標準化合物に基づくワイドターゲット解析でも同様の傾向が見られた。各代謝産物の蓄積様式を相関解析した結果、イソフラボン配糖体と顕著な相関ネットワークを示すグループが発見できた。これらは、イソフラボン関連候補を含む重要なグループであり、各代謝産物候補の構造解析が進めば発芽ダイズと発芽発酵ダイズの代謝変動の詳細が解明できる。今後は、実験系をスケールアップし、glyceollinなどの修飾が進んだ二次代謝産物の単離精製を進める。さらに、同一サンプルでRNAシーケンス解析を行い、広範囲な遺伝子発現プロファイルを得る。これらのオミックスデータを統合解析し、ダイズ種子発生の全容解明を目指す予定である。

文 献

- 1) Prosser GA, Larrouy-Maumus G and de Carvalho LP (2014): Metabolomic strategies for the identification of new enzyme functions and metabolic pathways. *EMBO Rep*, **15**, 657-669.
- 2) Gomez-Cabrero D, Abugessaisa I, Maier D, Teschendorff A, Merckenschlager M, Gisel A, Ballestar E, Bongcam-Rudloff E, Conesa A and Tegnér J (2014): *BMC Syst Biol*, **8 Suppl 2**: I1.
- 3) Kuo TM1, Doehlert DC and Crawford CG (1990): Sugar metabolism in germinating soybean seeds: evidence for the sorbitol pathway in soybean axes. *Plant Physiol*, **93**, 1514-1520.
- 4) Quinhone Junior A and Ida EI (2014): Isoflavones of the soybean components and the effect of germination time in the cotyledons and embryonic axis. *J Agric Food Chem*, **62**, 8452-8459.
- 5) Burow ME1, Boue SM, Collins-Burow BM, Melnik LI, Duong BN, Carter-Wientjes CH, Li S, Wiese TE, Cleveland TE and McLachlan JA (2001): Phytochemical glyceollins, isolated from soy, mediate antihormonal effects through estrogen receptor alpha and beta. *J Clin Endocrinol Metab*, **86**, 1750-1758.
- 6) Lee S1, Seo MH, Oh DK and Lee CH (2014): Targeted metabolomics for *Aspergillus oryzae*-mediated biotransformation of soybean isoflavones, showing variations in primary metabolites. *Biosci Biotechnol Biochem*, **78**, 167-174.
- 7) Sawada Y, Akiyama K, Sakata A, Kuwahara A, Otsuki H, Sakurai T, Saito K and Hirai MY (2009): Widely targeted metabolomics based on large-scale MS/MS data for elucidating metabolite accumulation patterns in plants. *Plant Cell Physiol*, **50**, 37-47.
- 8) Sawada Y and Hirai MY (2013): Integrated LC-MS/MS system for plant metabolomics. *Comput Struct Biotechnol J*, **4**: e201301011.
- 9) Sawada Y, Nakabayashi R, Yamada Y, Suzuki M, Sato M, Sakata A, Akiyama K, Sakurai T, Matsuda F, Aoki T, Hirai MY and Saito K (2012): RIKEN tandem mass spectral database (ReSpect) for phytochemicals: a plant-specific MS/MS-based data resource and database. *Phytochemistry*, **82**, 38-45.