

# メタボロミクス技術によるダイズ代謝制御関連マーカーのQTLマッピング

澤田有司\*

独立行政法人理化学研究所植物科学研究センター

## Metabolome Quantitative Locus Analysis in Soybean

Yuji SAWADA\*

RIKEN Plant Science Center, Yokohama 230-0045

### ABSTRACT

Metabolite quantitative trait locus (mQTL) analysis has great potential for discovering the associations involving diverse plant genes and metabolites. As a practical metabolic phenotyping, we have established widely targeted metabolomics using automated optimization of selected reaction monitoring (SRM) in 700 standard compounds. To detect the un-known metabolites, we also established a large-scale SRM assay system for thousands of un-targeted MS/MS tags (MS2Ts) obtained by plant extracts. As the populations for linkage based QTL mapping, soybean recombinant inbred lines (RILs) have been constructed from *Glycine max* (cultivated species) and *Glycine soja* (wild type accession). *G. soja* is expected to be a useful genetic resource for molecular breeding of soybean. Using our metabolomics platform, a total of 16 SRMs derived from 4 standard compounds and 12 MS2Ts were significantly associated with the four major mQTLs (LOD values > 10) in TK780 (*G. max*) x B01167 (*G. soja*) RILs, which were distributed from NBRP (Legume Base). The annotations of associated SRMs were flavonoids, phenolic compounds and unknown metabolites, which were highly accumulated in *G. soja*. Unfortunately, no significant mQTL was found in *G. max* x *G. max* RILs. These results suggested that the genetic resources of *G. soja* are useful for epoch-making metabolic engineering in soybean. *Soy Protein Research, Japan* **15**, 26-30, 2012.

Key words : Metabolomics, large-scale SRM assay system, QTL, *G. soja*

\* 〒230-0045 横浜市鶴見区末広町1-7-22

## 方 法

メタボロミクスは生物の定量的な情報を一挙に解析するオミックス手法であり、生物試料に含まれる代謝産物の一斉検出を目的としている<sup>1,2)</sup>。しかし、代謝産物の多くは未同定であるため、予め標的を絞らない非ターゲット解析が必要になる。これまでに、各種分離部分と質量分析装置を組み合わせた分析手法が提唱されている<sup>3~8)</sup>。これらの質量分析装置による代謝産物の同定には標準化合物による参照データが必要である<sup>9)</sup>。

多様な植物代謝産物は革新的な活性を示す医薬品の原料として注目されているが、利用可能な標準化合物は極めて限定的である。モデル植物シロイヌナズナの代謝産物の構造決定を非ターゲットで行った研究では37種の代謝産物が単離同定されたが、新規の代謝産物は2種であった<sup>10)</sup>。すなわち、一人の研究者が可能な代謝産物の単離に基づく新規の構造決定は年間数個程度である。

この植物代謝産物とそれを制御する遺伝的背景を解明する手法としてメタボローム情報を独立の表現型として量的形質座 (quantitative trait locus, QTL) 解析を行うゲノムワイドなメタボローム解析が注目されている<sup>11)</sup>。我々は、モデル植物シロイヌナズナを利用した大規模な遺伝学的解析<sup>12)</sup>、リンケージマッピングとアソシエーション解析によるmetabolite QTL (mQTL) 解析を行っている。これらの手法が確立できれば、構造決定されていない新規代謝産物とその合成遺伝子の同定が可能になる。

我々は、大規模な植物遺伝資源に適用可能なメタボローム解析を開発するために特定代謝産物の超高感度解析に利用されるselected reaction monitoring (SRM) 検出に着目し、約700種の化合物がSRM検出を行うワイドターゲット解析を確立した<sup>13)</sup>。さらに、未同定代謝産物の検出を行うために、非ターゲット検出したタンデムマススペクトラムタグ (MS2T) からSRM条件を同定する技術を確立した。

我々が独自開発したメタボロミクス技術を利用し、ダイズに含まれる豊富な機能性成分のmQTLを発見すれば交配育種による画期的な成分改変ダイズが作出できると期待される。本研究ではマメ科植物のナショナルバイオリソースプロジェクト (NBRP) で配布されている組み換え自殖系統 (recombination inbred line, RIL) のメタボローム情報を独立の表現型としたmQTL解析を行った。これらの結果はダイズの効率的な成分育種を可能にする基盤情報である。

### 植物種子材料

ダイズの親株 (ミスズダイズ, 秣食豆公, TK780, B01167) およびRILは、NBRP (<http://www.legumebase.brc.miyazaki-u.ac.jp/>) から取得した。

### 植物材料の抽出

ダイズ種子は木槌で予備破砕し、シェイクマスターネオ (バイオメディカルサイエンス社) を利用し粉末化した (1,500 rpm, 10 min)。粉末化したダイズを $2 \pm 0.2$  mg秤量し、2 mLチューブに分注した。これに5 mmジルコニアビーズを1個と1 mLの抽出溶媒 (80% MeOH) を加え、破砕操作と同様にシェイクマスターネオで抽出した (1,000 rpm, 2 min)。

### 抽出物の前処理

自動分注システム (Microlabstar, ハミルトン社) を利用して抽出液50  $\mu$ Lを96 well plateに分注し、同システムのロボットアームで搬送し、乾燥窒素吹き付け型の溶媒乾燥機 (Ultra vap, Povair社) で溶媒をドライアップした。さらに同システムで120  $\mu$ LのH<sub>2</sub>Oで再溶解後、384 well filter plate (Whatman社) でフィルター濾過した。

### 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タンデム型質量分析装置 (LC-MS/MS) で分析した (Waters社UPLC-TQS/Q-ToF Premier)<sup>13)</sup>。

### データ解析

LC-MS/MS分析結果から得られた各代謝産物のピーク面積は、解析ソフトR (<http://www.r-project.org/>) を利用し、欠損値は0.1で置き換え、底2で対数変換し、サンプル間でZ-score化した。ノーマライズした各ピーク面積をRのパッケージRqt<sup>14)</sup> を利用し、mQTL解析した。

## 結 果

### タンデム型質量分析装置の統合解析による未同定代謝産物の検出

標準化合物に基づく検出に加え、未同定代謝産物の検出を行うために非ターゲット検出したMS2Tから超高感度検出が可能なSRM条件を最適化する新技术を確立した (Fig. 1)。本技術はUPLC-Q-ToF Premierで検出したMS2Tから分子量関連イオンと最も強度の高いフラグメント化イオンを計算科学的に選択する。次に同一試料をUPLC-TQSで再測定し、高感度なフラグ

メント化電圧を検討する (Fig. 1). この二種の質量分析装置の特性を統合した新技術によって、数千の代謝産物候補から安定的に検出可能な数百のSRMが決定できる。

本研究ではダイズRILの親株4種の抽出液を混合し、UPLC-Q-ToF Premierで分析した。この結果、合計46,717個のMS2Tが取得できた。これらのMS2T情報は公開情報として利用可能である (<http://prime.psc.riken.jp/lcms/>)。次に、ダイズのMS2T情報から安定的に検出可能なSRM条件を最適化するために、UPLC-TQSによるフラグメント化電圧 (collision energy) の検討を行った (10 Vごとに6段階 10-60 V)。MS2T取得に利用した混合試料を3回測定し、合計384個のSRM条件が最適化できた (LC-MS分析のピーク面積 $\geq 100$ , 相対標準偏差 $< 10\%$ )。

### ダイズおよびツルマメのmQTL解析

遺伝子組み換えが困難な作物品種では、有用形質を制御する遺伝子座を推定する手法としてQTL解析が行われている。本研究では、SRM検出した各代謝産物の蓄積様式を独立の表現形質としてmQTL解析を行った。ミスズダイズ $\times$ 秣食豆公のRIL 149種とTK780 $\times$ B01167のRIL 93種を3反復でメタボローム解析を行い、公開されている分子マーカー情報をQTL解析ソフトR/qtlで解析した。TK780 $\times$ B01167のRILでは4か所の主要mQTLが同定された (Fig. 2)。各mQTLと高い

LOD値 ( $< 10$ ) を示すSRMを調べた結果、標準化合物に基づく4 SRMとMS2Tに基づく12 SRMであることが分かった。さらに各SRMのアノテーションを調べた結果、フラボノイドとフェノール性化合物であることが解った (Table 1)。一方、ミスズダイズ $\times$ 秣食豆公のRILでは顕著なmQTLは発見できなかった。

## 考 察

本研究では標準化合物が入手できない未同定代謝産物に対応可能なメタボロームQTLプラットフォームを構築した。本技術をmQTLに適応した結果、ダイズとツルマメ間のRILで特徴的なmQTLが複数同定され

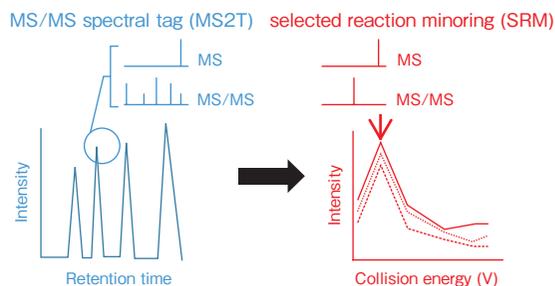


Fig. 1. Scheme of integrated metabolomics for large-scale SRM assay.

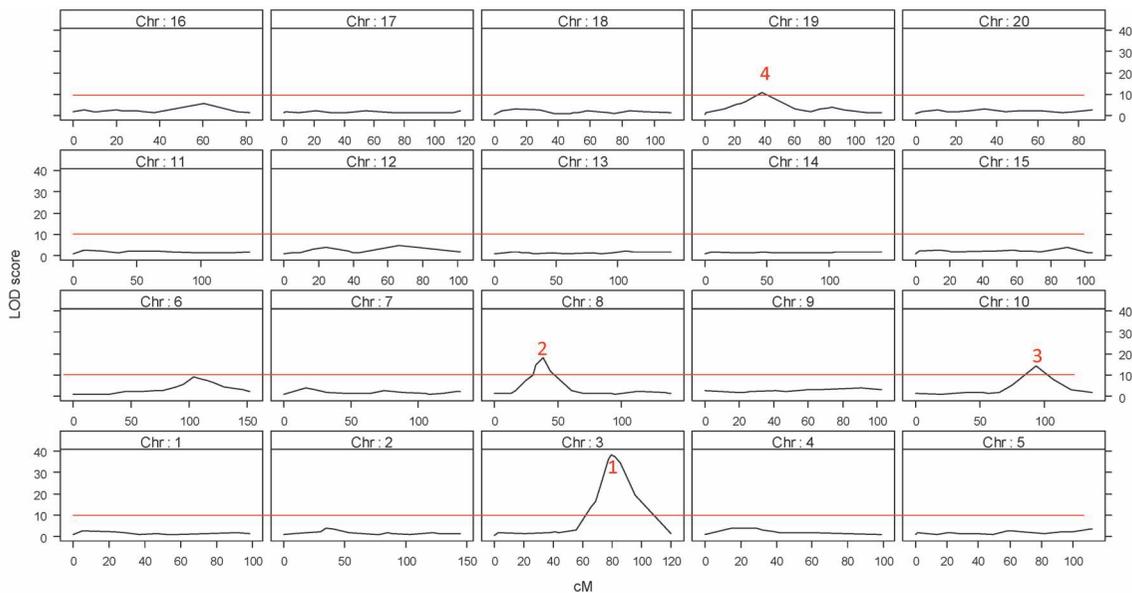


Fig. 2. mQTL of TK780 (*G. max*)  $\times$  B01167 (*G. soja*) RILs. The maximum logarithm of odds (LOD) score values obtained from triplicate experiments were plotted in each chromosome (Chr: 1-20). The major mQTLs (1-4) were estimated based on the LOD scores ( $< 10$ ) and are shown in red numbers.

Table 1. Summary of identified mQTLs in TK780 (*G. max*) x B01167 (*G. soja*) RILs.

mQTL <sup>1</sup>	Methods <sup>2</sup>	Annotations (number of associated SRMs)
1	UT	Unknown (2)
2	UT, WT	Flavonoid (3), Phenolic compound (1), Unknown (6)
3	WT	Flavonoid (1)
4	UT	Unknown (4)

1. The number of mQTL corresponds to Fig. 2

2. UT, un-targeted metabolomics; WT, widely targeted metabolomics

た。また、mQTLの多型と代謝産物蓄積量を調べた結果、いずれもツルマメの遺伝子型が代謝産物量を増加させることが解った。この結果、野生種ツルマメの遺伝資源がダイズの効率的な成分育種に適していることが示唆された。

今後、ツルマメ間の代謝産物量の多様性を解明すれば画期的な代謝産物の発見に加え、栽培種との交配育種への応用も期待できる。我々の先行研究では野生種モデル植物ミヤコグサでは、顕著なmQTL発見されている。また、本研究で特定のmQTLと顕著な関連性を示した未同定代謝産物 (MS2T) は、我々が構築したMS/MSデータベースReSpectで検索し、得られた

推定構造に基づく単離同定を行う予定である (<http://spectra.psc.riken.jp/>)。

一方、栽培種間に特徴的なmQTLは発見できなかった。従って、ダイズの栽培化の過程で可視の表現型の多様性は人為的に保たれてきたが、不可視の代謝産物蓄積量は味覚に基づく選抜を受けた結果、代謝産物の種類や量の差が見られない系統が選択されたことが示唆された。今後、高密度化が可能な一塩基多型 (single nucleotide polymorphism, SNP) タイピング情報と我々のメタボローム情報との関連解析が可能になれば、栽培種と野生種の代謝システムの比較が可能になると期待している。

## 要 約

量的形質座 (quantitative trait locus, QTL) 解析は有用形質に関わる遺伝的背景を解明する重要技術である。同手法でダイズに含まれる有用代謝産物の生合成機構に関わるQTLを同定すれば代謝育種の基盤情報となる。そこで本研究はダイズ代謝産物の一斉解析手法 (メタボロミクス) を利用した QTL解析を行った。QTL解析に利用した組み換え自殖系統は、Legume Baseから入手した。代謝産物測定は、予め標的を絞らない非ターゲット解析と標準化合物に基づくターゲット解析を行った。これらのメタボローム情報を解析し、検出強度と親株間の差から約100代謝産物候補を選択し、各代謝産物の蓄積情報を独立の量的形質としてQTL解析を行った。3回の独立実験を行い、再現性の高いQTLを調べた結果、栽培種と野生種ツルマメ間の系統 (TK780×B01167) で顕著なQTLが4か所発見された (LOD score>10)。対象化合物を我々が構築した MS/MS データベース ReSpect (<http://spectra.psc.riken.jp/>) で検索した結果、ツルマメに多く含まれるフラボノイドとフェノール性化合物であることが解った。これらは従来交配育種では未利用の有用代謝産物と考えられる。一方、栽培種間の系統 (ミスズダイズ×秣食豆公503) は顕著なQTLが発見できなかった。この結果から、栽培種では可視の表現型 (収量、形態など) に基づく選抜育種の過程で不可視の代謝産物の有用形質が脱落していることが推定された。今後、これらのツルマメに特徴的な代謝制御遺伝子を同定し、ダイズの代謝育種を実現する。

## 文 献

- 1) Fernie AR, Trethewey RN, Krotzky AJ and Willmitzer L (2004): Metabolite profiling: from diagnostics to systems biology. *Nat Rev Mol Cell Biol* **5**, 763-769.
- 2) Weckwerth W (2003): Metabolomics in systems biology. *Annu Rev Plant Biol* **54**, 669-689.
- 3) Fiehn O, Kopka J, Dormann P, Altmann T, Trethewey RN and Willmitzer L (2000): Metabolite profiling for plant functional genomics. *Nat Biotechnol* **18**, 1157-1161.
- 4) Guy C, Kopka J and Moritz T (2008): Plant metabolomics coming of age. *Physiol Plant* **132**, 113-116.
- 5) Hall RD, Brouwer ID and Fitzgerald MA (2008): Plant metabolomics and its potential application for human nutrition. *Physiol Plant* **132**, 162-175.
- 6) Saito K and Matsuda F (2010): Metabolomics for functional genomics, systems biology, and biotechnology. *Annu Rev Plant Biol* **61**, 463-489.
- 7) Schauer N and Fernie AR (2006): Plant metabolomics: towards biological function and mechanism. *Trends Plant Sci* **11**, 508-516.
- 8) Sumner LW, Mendes P and Dixon RA (2003): Plant metabolomics: large-scale phytochemistry in the functional genomics era. *Phytochemistry* **62**, 817-836.
- 9) Last RL, Jones AD and Shachar-Hill Y (2007): Towards the plant metabolome and beyond. *Nat Rev Mol Cell Biol* **8**, 167-174.
- 10) Nakabayashi R, Kusano M, Kobayashi M, Tohge T, Yonekura-Sakakibara K, Kogure N, Yamazaki M, Kitajima M, Saito K and Takayama H (2009): Metabolomics-oriented isolation and structure elucidation of 37 compounds including two anthocyanins from *Arabidopsis thaliana*. *Phytochemistry* **70**, 1017-1029.
- 11) Lisec J, Meyer RC, Steinfath M, Redestig H, Becher M, Witucka-Wall H, Fiehn O, Torjek O, Selbig J, Altmann T and Willmitzer L (2008): Identification of metabolic and biomass QTL in *Arabidopsis thaliana* in a parallel analysis of RIL and IL populations. *Plant J* **53**, 960-972.
- 12) Hirai MY, Sawada Y, Kanaya S, Kuromori T, Kobayashi M, Klausnitzer R, Hanada K, Akiyama K, Sakurai T, Saito K and Shinozaki K (2010): Toward genome-wide metabolotyping and elucidation of metabolic system: metabolic profiling of large-scale bioresources. *J Plant Res* **123**, 291-298.
- 13) 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.
- 14) Broman KW, Wu H, Sen S and Churchill GA (2003): R/qtl: QTL mapping in experimental crosses. *Bioinformatics* **19**, 889-890.