

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/38026272>

Dual Labeling of Metabolites for Metabolome Analysis (DLEMMA): A New Approach for the Identification and Relative Quantification of Metabolites by Means of Dual Isotope Labeling an...

ARTICLE *in* ANALYTICAL CHEMISTRY · OCTOBER 2009

Impact Factor: 5.64 · DOI: 10.1021/ac901495a · Source: PubMed

CITATIONS

24

READS

166

5 AUTHORS, INCLUDING:



Sergey Malitsky

Weizmann Institute of Science

21 PUBLICATIONS 590 CITATIONS

[SEE PROFILE](#)



Ilana Rogachev

Weizmann Institute of Science

54 PUBLICATIONS 1,264 CITATIONS

[SEE PROFILE](#)



Asaph Aharoni

Weizmann Institute of Science

101 PUBLICATIONS 5,585 CITATIONS

[SEE PROFILE](#)

Dual Labeling of Metabolites for Metabolome Analysis (DLEMMA): A New Approach for the Identification and Relative Quantification of Metabolites by Means of Dual Isotope Labeling and Liquid Chromatography–Mass Spectrometry

Liron Feldberg,[†] Ilya Venger, Sergey Malitsky, Ilana Rogachev, and Asaph Aharoni*

Department of Plant Sciences, Weizmann Institute of Science, Rehovot, Israel

Advanced metabolomics technologies are anticipated to permit the identification and quantification of metabolites at the whole-metabolome scale. Yet, most of the metabolites either remain unknown or cannot be identified unambiguously. Moreover, the present approaches suffer from inaccuracies in relative quantification because of sample preparation and matrix effects. Here we present Dual Labeling of Metabolites for Metabolome Analysis (DLEMMA) as a valuable tool, which with analogy to DNA array assays enables the identification and relative quantification of differential metabolites in a single sample. DLEMMA was demonstrated as an efficient method for reducing the number of possible chemical structures assigned that exhibit the same elemental composition. Its strength was exemplified by the discovery of 10 novel Tryptophan derivatives. Furthermore, employing DLEMMA by feeding two Phenylalanine-labeled precursors, we could detect differential metabolites between transgenic and control plants. The accuracy of relative quantification is also enhanced since DLEMMA provides identical matrixes for both samples, thus avoiding the effects of different complex biological matrixes on electrospray ionization. Hence, DLEMMA will complement and contribute to the advancement of metabolomics technologies and boost metabolic pathway discovery in diverse organisms.

Metabolomics aims at the non-targeted detection, identification, and quantification of all metabolites in a given biological sample.^{1–3} At present, the two most common techniques for metabolite analysis employ NMR and chromatography coupled to mass spectrometry (e.g., GC-MS and LC-MS).^{4–7} The latter offers high

sensitivity and permits the analysis of numerous low-abundance metabolites. Despite the huge importance of identification, most of the metabolites in complex biological matrixes either remain unknown or cannot be identified unambiguously.

Accurate mass measurements provided by high resolution mass spectrometry in combination with the use of an isotopic ratio filter is at present the most promising technique for facing the identification challenge.^{8–10} However, identification is not unambiguous since for a single mass more than one elemental composition is possible, and a certain elemental composition may have many possible molecular structures. While standard compounds could greatly improve metabolite identification, they are frequently unavailable. Recently, isotopic labeling was employed as a tool for constraining the number of elemental composition assignable to exact mass measurements.^{11–13} Yet, an important issue of identification, that is, distinguishing between many chemical structures exhibiting the same elemental formula, remains problematic. Hence, new methods are required for substantially increasing the confidence of structural elucidation in metabolomics studies.

Relative quantification in metabolomics is a major compromise as compared to the classic targeted analyses methods. The most widespread approach for relative quantification uses multiple retention time standards coupled to sophisticated data analysis software for comparison between consecutively ana-

* To whom correspondence should be addressed. E-mail: asaph.aharoni@weizmann.ac.il. Phone: +972 8 934 3643. Fax: +972 8 934 4181.

[†] Current address: Department of Analytical Chemistry, Israel Institute for Biological Research, Ness Ziona, Israel.

- (1) Roessner, U.; Willmitzer, L.; Fernie, A. R. *Plant Physiol.* 2001, 127, 749–64.
- (2) Raamsdonk, L. M.; Teusink, B.; Broadhurst, D.; Zhang, N.; Hayes, A.; Walsh, M. C.; Berden, J. A.; Brindle, K. M.; Kell, D. B.; Rowland, J. J.; Westerhoff, H. V.; van Dam, K.; Oliver, S. G. *Nat. Biotechnol.* 2001, 19, 45–50.
- (3) Weckwerth, W.; Loureiro, M. E.; Wenzel, K.; Fiehn, O. *Proc. Natl. Acad. Sci. U.S.A.* 2004, 101, 7809–7814.

- (4) Birkemeyer, C.; Luedemann, A.; Wagner, C.; Erban, A.; Kopka, J. *Trends Biotechnol.* 2005, 23, 28–33.
- (5) Rochfort, S. *J. Nat. Prod.* 2005, 68, 1813–1820.
- (6) Sumner, L. W.; Mendes, P.; Dixon, R. A. *Phytochemistry* 2003, 62, 817–836.
- (7) Weckwerth, W.; Fiehn, O. *Curr. Opin. Biotechnol.* 2002, 13, 156–160.
- (8) Dettmer, K.; Aronov, P. A.; Hammock, B. D. *Mass Spectrom. Rev.* 2007, 26, 51–78.
- (9) Dunn, W. B.; Overy, S.; Quick, W. P. *Metabolomics* 2005, 1, 137–148.
- (10) von Roepenack-Lahaye, E.; Degenkolb, T.; Zerjeski, M.; Franz, M.; Roth, U.; Wessjohann, L.; Schmidt, J.; Scheel, D.; Clemens, S. *Plant Physiol.* 2004, 134, 548–559.
- (11) Giavalisco, P.; Hummel, J.; Lisec, J.; Inostroza, A. C.; Catchpole, G.; Willmitzer, L. *Anal. Chem.* 2008, 80, 9417–9425.
- (12) Giavalisco, P.; Kohl, K.; Hummel, J.; Seiwert, B.; Willmitzer, L. *Anal. Chem.* 2009, 81 (15), 6546–6551.
- (13) Hegeman, A. D.; Schulte, C. F.; Cui, Q.; Lewis, I. A.; Huttlin, E. L.; Eghbalnia, H.; Harms, A. C.; Ulrich, E. L.; Markley, J. L.; Sussman, M. R. *Anal. Chem.* 2007, 79, 6912–6921.

lyzed samples.^{10,14,15} This approach suffers from inaccuracies in relative quantification because of sample preparation and matrix effects in electrospray ionization (ESI).^{16–19} A different approach uses *in vivo* labeling of metabolites with stable isotopes enriched with heavy isotopes and comparison with a control sample with natural isotopic abundances.^{4,20–23} Recently, relative quantification of metabolites using labeling reagents that react with compounds containing a particular functional group was reported.^{24,25} This approach may suffer from inaccuracies as a result of different derivatization efficiencies in different matrixes^{26,27} and only contributes to identification by indicating the presence of a certain functional group.

Here we present Dual Labeling of Metabolites for Metabolome Analysis (DLEMMA) as a complementary component to the current metabolomics toolbox. It considerably enhances the effectiveness and confidence in metabolite identification and provides a valuable approach for relative quantification. Hence, DLEMMA will largely contribute to gene and pathway discovery in diverse organisms.

EXPERIMENTAL SECTION

Plant Material and Reagents. *Arabidopsis* plants (ecotype Columbia), were cultured in a short day (8 h light; 18 °C), 8 weeks before feeding experiments. For the Trp feeding experiments, leaves were treated by spraying with 500 µM Methyl Jasmonate (0.8% EtOH/H₂O in addition to 0.02% Silwet), three times in a 24 h period prior (on the intact plant) and during the feeding process (on single leaves in a tube). Tryptophan-2'4'5'6'7'-d₅ (98.6% atom D) was purchased from CDN Isotopes (Quebec, Canada). Tryptophan ¹³C₁₁¹⁵N₂ (¹⁵N > 97%, ¹³C > 95%), Phenyl-¹³C₆-alanine (¹³C > 99%) and Phenyl-¹³C₆,d₅-alanine (D > 99%, ¹³C > 98% D) were synthesized by ISOTEC (Sigma-Aldrich). All other reagents were purchased from Sigma-Aldrich Chemicals.

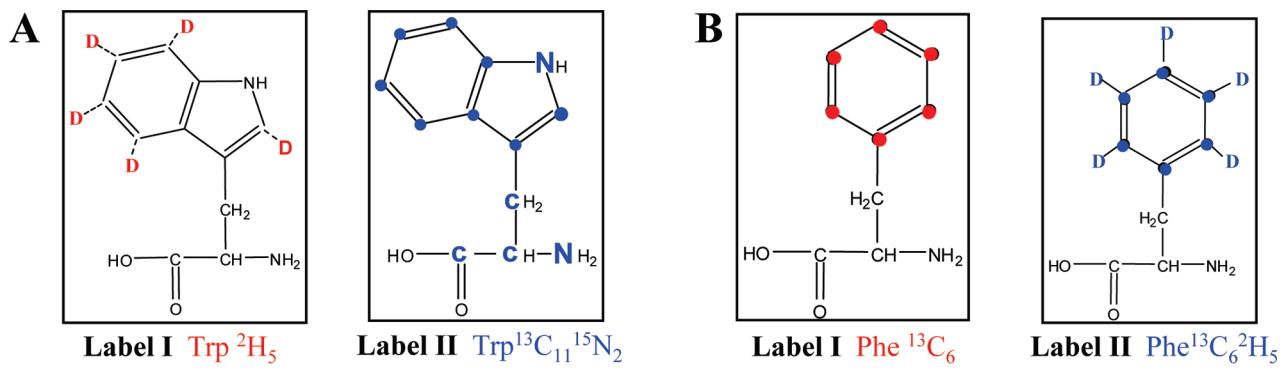
Feeding Experiments. Leaves (50–100 mg of fresh weight) of 8 weeks old plants were excised by cutting at the proximal side of the pedicel with scissors under water (to avoid air

penetration into the pedicel which may influence the feeding efficiency), leaving attached 1–2 cm of the pedicel. The leaves were immersed into a PCR tube with 0.5 mL aqua solution of 0.5 mg/mL of the required precursor [Trp (unlabeled), Trp ²H₅ (label I), Trp¹³C₁₁¹⁵N₂ (label II), Phe (unlabeled), Phe ¹³C₆ (label I), or Phe¹³C₆²H₅ (label II)]. All feeding experiments were performed in a controlled environment for 24 h under 25 °C and constant fluorescent illumination and humidity. Every feeding experiment was composed of six components (see Supporting Information, Methods S1), each conducted in three replicates (see Figure 1).

Sample Preparation. Fresh leaves, to which precursors have been fed, were rinsed with a small amount of water to remove all traces of externally fed precursor, dried gently, and weighed. For swapped label feeding experiments, leaves with similar weights from the WT-L1/ALT-L2 and WT-L2/ALT/L1 experimental components were combined into one sample. In cases of different leaves weighed in the combined sample, we used a correction factor to match concentrations. The leaves were frozen, and ground with 0.1% formic acid in 92% MeOH:8%H₂O (v/v) (400 mL per 100 mg leaf). The suspension was sonicated for 10 min, centrifuged at 14,000 g for 5 min, and the obtained supernatant was filtered through a 0.22 µm PTFE membrane filter before injection to the LC-MS instrument.

UPLC-QTOF-MS Analysis and Data Treatment. A UPLC-QTOF-MS instrument (Waters Premier QTOF, Milford, MA, U.S.A.) was used for metabolite analyses. Separation of metabolites was performed on a 100 × 2.1 mm i.d., 1.7 µm UPLC BEH C18 column. The mobile phase consisted of 0.1% formic acid in Acetonitrile:Water (5:95, v/v) (phase A), and 0.1% formic acid in Acetonitrile (phase B). The linear gradient program was as follows: 100–72% A over 22 min, 72–60% A over 0.5 min, 60 – 0% A over 0.5 min, held at 100% B for a further 1.5 min, then returned to the initial conditions (100% A) in 0.5 min, and conditioning at 100% A. The total analytical run was 26 min. The flow rate was 0.3 mL/min, the column temperature was kept at 35 °C. Masses of the eluted compounds were detected by a QTOF Premier MS, equipped with ESI source. Acquisition was performed separately in positive and negative ESI modes. The following settings were applied during the LC-MS runs: capillary voltage at 3.0 kV; cone voltage at 30 eV; collision energy at 3 eV. Full scan mass spectra were acquired from 50–1500 Da. The following settings were applied during the LC-MS/MS run: capillary spray at 3.0 kV; cone voltage at 30 eV; collision energies were 10–25 eV for positive mode and 15–40 for negative mode. Argon was used as the collision gas for collision-induced dissociation (CID) (MS-MS) experiments. The MS was calibrated using sodium formate, and leucine enkaphalin was used as the lock mass. A mixture of 15 standard compounds, injected after each 5 samples was used for quality control. The MassLynx software version 4.1 (Waters Inc.) was used to control the instrument and calculate accurate masses. Chromatograms signals were analyzed using the MarkerLynx software v. 4.1 (Waters Inc.) for peak picking, alignment and integration. To filter markers belonging to metabolites which are derived from the fed precursor, a custom script was developed in Matlab v. 7.3 (The Mathworks) which uses the MarkerLynx data obtained from the six components of the experimental setup (WT-UL, ALT-UL, WT-L1+L2, ALT-L1+L2, WT-L1/ALT-L2 and WT-

- (14) Want, E. J.; O'Maille, G.; Smith, C. A.; Brandon, T. R.; Uritboonthai, W.; Qin, C.; Trauger, S. A.; Siuzdak, G. *Anal. Chem.* **2006**, *78*, 743–752.
- (15) Wang, W.; Zhou, H.; Lin, H.; Roy, S.; Shaler, T. A.; Hill, L. R.; Norton, S.; Kumar, P.; Anderle, M.; Becker, C. H. *Anal. Chem.* **2003**, *75*, 4818–4826.
- (16) Bottcher, C.; Roepenack-Lahaye, E. V.; Willscher, E.; Scheel, D.; Clemens, S. *Anal. Chem.* **2007**, *79*, 1507–1513.
- (17) Matuszewski, B. K.; Constanzer, M. L.; Chavez-Eng, C. M. *Anal. Chem.* **2003**, *75*, 3019–3030.
- (18) Rogatsky, E.; Stein, D. *J. Am. Soc. Mass Spectrom.* **2005**, *16*, 1757–1759.
- (19) Choi, B. K.; Hercules, D. M.; Gusev, A. I. *J. Chromatogr. A* **2001**, *907*, 337–342.
- (20) Lafaye, A.; Labarre, J.; Tabet, J. C.; Ezan, E.; Junot, C. *Anal. Chem.* **2005**, *77*, 2026–2033.
- (21) Wu, L.; Mashego, M. R.; van Dam, J. C.; Proell, A. M.; Vinke, J. L.; Ras, C.; van Winden, W. A.; van Gulik, W. M.; Heijnen, J. J. *Anal. Biochem.* **2005**, *336*, 164–171.
- (22) Kruger, N. J.; Huddleston, J. E.; Le Lay, P.; Brown, N. D.; Ratcliffe, R. G. *Phytochemistry* **2007**, *68*, 2176–2188.
- (23) Libourel, I. G.; Shachar-Hill, Y. *Annu. Rev. Plant Biol.* **2008**, *59*, 625–650.
- (24) Shortreed, M. R.; Lamos, S. M.; Frey, B. L.; Phillips, M. F.; Patel, M.; Belshaw, P. J.; Smith, L. M. *Anal. Chem.* **2006**, *78*, 6398–6403.
- (25) Lamos, S. M.; Shortreed, M. R.; Frey, B. L.; Belshaw, P. J.; Smith, L. M. *Anal. Chem.* **2007**, *79*, 5143–5149.
- (26) Sternier, J. L.; Johnston, M. V.; Nicol, G. R.; Ridge, D. P. *J. Mass Spectrom.* **2000**, *35*, 385–391.
- (27) Souverain, S.; Rudaz, S.; Veuthey, J. L. *J. Chromatogr A* **2004**, *1058*, 61–66.



WT-UL	ALT-UL	WT-L1+L2	ALT-L1+L2	WT-L1/ALT-L2		WT-L2/ALT/L1	
WT Sample + unlabeled precursor	ALT sample + unlabeled precursor	WT sample + Label I + Label II	ALT sample + Label I + Label II	WT sample + Label I + Label II	ALT sample + Label II	WT sample + Label II	ALT sample + Label I

Figure 1. Experimental set up for the DLEMMA approach. **(A)** The two labeling of the Trp precursor used in this study. **(B)** The two labeling of the Phe precursor used in this study. **(C)** The six components of the DLEMMA experimental set up: WT-UL (Wild Type leaves fed with Unlabeled precursor), ALT-UL (Altered leaves fed with Unlabeled precursor), WT-L1+L2 (WT leaves fed with a mixture of two different labeling), ALT-L1+L2 (Altered leaves fed with a mixture of two different labeling), WT-L1/ALT-L2 (Wild Type leaves fed with labeled 1 precursor and Altered leaves fed with labeled 2 precursor), WT-L2/ALT-L1 (Wild Type leaves fed with labeled 2 precursor and Altered leaves fed with labeled 1 precursor). Altered (ALT) leaves were either genetically modified or altered by treatment).

L2/ALT-L1). A detailed description is available in Supporting Information, Methods S1.

Putative Identification and Relative Quantification of the Detected Metabolites. Elemental composition (in most cases mass accuracy between 1 to 5 ppm, with relative low i-Fit values, Supporting Information, Tables S1) of the unlabeled metabolite was predicted using the Masslynx software (Waters, Inc.). The SciFinder Scholar tool which provides access to the CAS and MEDLINE databases (SciFinder Scholar version 2007), was used for elemental composition search. Taking into consideration the labeling pattern obtained for the fed precursor detected metabolite, the number of possible structures per a given elemental composition was reduced using the following steps: (1) narrowing the number of molecular structure to fed precursor derivatives only (for example: structures containing the indolic ring in the case of Trp derivatives), by using the ability of SciFinder Scholar tool for substructure search. (2) Remaining only with structures which can match the combined labeling pattern of label I + label II by the combination of SciFinder sub structural search and manual inspection of the optional structures. (3) Remaining with molecular structures that were reported previously in biological literature, information generated by the SciFinder Scholar tool for specific literature search. (4) Verifying one of the few left molecular structures with MS-MS observed fragmentation (of unlabeled and labeled precursors). (5) Confirming the putative structure with standard compound when available.

Swapped-label feeding experiments (WT-L1/ALT-L2 and WT-L2/ALT/L1), providing chromatograms peaks intensities of metabolites with different labeling (L1 and L2) which are derived

from different samples, were used for relative quantification according to the following equation:

$$\text{Differential ratio} = \sqrt{\frac{\text{ALT-L2}}{\text{WT-L1}}} \times \sqrt{\frac{\text{ALT-L1}}{\text{WT-L2}}}$$

when for example ALT-L2/WT-L1 is the mean ratio (calculated from 3 replicates) between the intensities of labeled II metabolite belonging to altered metabolic leaves and the intensities of labeled I metabolite belonging to wild type leaves (derived from the same chromatogram).

RESULTS

DLEMMA Concept. To advance the capabilities of metabolomics in relative quantification and identification we developed a new system (DLEMMA) in which two samples are fed with the same precursor that is labeled differently with stable isotopes (Figure 1; Figure 2; Supporting Information, Results S1). The principle of DLEMMA is analogous to the one often used for gene expression assays using DNA microarrays. Label-fed samples are combined to a single sample prior to extraction, and label bias is corrected through swapped-label feeding. Metabolites derived from the fed precursor appear in the UPLC-QTOF-MS chromatogram as three co-eluting peaks (unlabeled and two differently labeled metabolites). Subsequently, peak picking, integration, and alignment, data filtering for precursor-derived masses (i.e., markers), and clustering to metabolites (manually according to retention time (r.t.) and peak shape), enables the detection of compounds derived from the fed precursor (see Supporting Information, Methods S1). Public metabolite databases search in

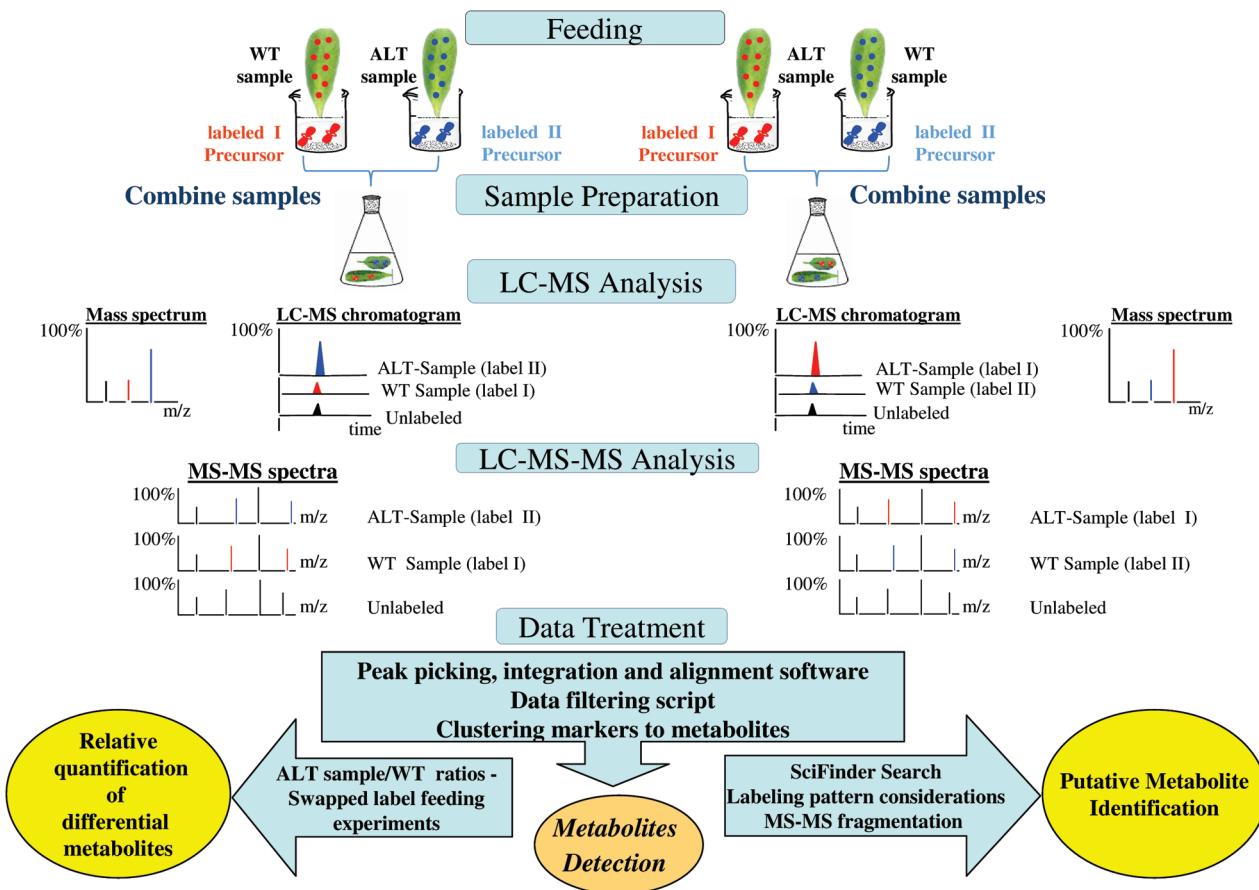


Figure 2. Schematic representation of the DLEMMA approach. The procedure begins by feeding leaves with different labeled precursors in swapped label experiments, followed by sample preparation in which both leaves are combined into one sample, prior to LC-MS and LC-MS-MS analysis and data treatment. The raw data is further processed using a peak picking, integration, and alignment software and subsequent data filtering and a clustering process to detect metabolites derived from the fed precursor. Using the SciFinder tool for comprehensive database search, labeling pattern considerations, and MS-MS fragmentation enables putative metabolite identification. Using metabolites ratios derived from metabolically different samples obtained from swapped label experiments, allows metabolite relative quantification. The scheme describes the fifth and the sixth components of the experimental setup presented in 1C (WT-L1/ALT-L2 and WT-L2/ALT-L1). The other four components, control experiments, were performed in a similar way as described in the Experimental Section.

combination with labeling pattern consideration and MS-MS fragmentation, allows putative metabolites identification. In the MS-MS spectra, common fragments between the unlabeled and the labeled metabolites indicate the unlabeled parts of the molecule while different fragments point toward the labeled parts of the molecule. Differential metabolites between samples are determined by the use of intensity ratios obtained in the swapped-label experiments.

Experimental Setup for DLEMMA-Based Analysis. The DLEMMA approach was employed for the study of metabolic pathways derived from the aromatic amino acids phenylalanine (Phe) and tryptophan (Trp) precursors in *Arabidopsis* leaves. Since the DLEMMA approach is based on MS analysis and has two applications, relative quantification and identification, rational selection of labeling the two precursors is significant. We specified three criteria for selecting labeled precursors: (i) The mass difference between both labeled molecules and the unlabeled ones should be higher than 3 Da, to avoid isotopes overlapping which can interrupt with quantification. (ii) The labeled region in the precursor molecule should be common inasmuch as possible related metabolites to enable following the downstream metabolites. (iii) Each precursor molecule should contain different labeled

atoms (expect of labile hydrogens as carboxylic or hydroxylic ones) for full coverage of the molecule, to have as much information as possible for structure elucidation. The labeled precursor metabolites Trp²H₅ and Trp¹³C₁₁¹⁵N₂ meet the specified criteria. On the other hand, Phe¹³C₆ and Phe¹³C₆²H₅, do not meet the third criteria, and hence, have less power in structure elucidation, although still appropriate for the purpose of relative quantification.

One experimental system consisted of wild-type (WT) leaves and those treated with the signal molecule methyl jasmonate (MJ) that induces the production of Trp-derived metabolites²⁸ (leaves were fed with Trp²H₅ and Trp¹³C₁₁¹⁵N₂; Figure 1A). The second experimental system consisted of WT leaves and leaves of transgenic plants (PAP1) altered in the production of the Phe-derived metabolites as a result of overexpressing the *PAPI* regulatory gene²⁹ (leaves were fed with Phe¹³C₆ and Phe¹³C₆²H₅;

- (28) Skirycz, A.; Reichelt, M.; Burow, M.; Birkemeyer, C.; Rolcik, J.; Kopka, J.; Zanor, M. I.; Gershenson, J.; Strnad, M.; Szopa, J.; Mueller-Roeber, B.; Witt, I. *Plant J.* **2006**, *47*, 10–24.
- (29) Tohge, T.; Nishiyama, Y.; Hirai, M. Y.; Yano, M.; Nakajima, J.; Awazuhara, M.; Inoue, E.; Takahashi, H.; Goodenowe, D. B.; Kitayama, M.; Noji, M.; Yamazaki, M.; Saito, K. *Plant J.* **2005**, *42*, 218–235.

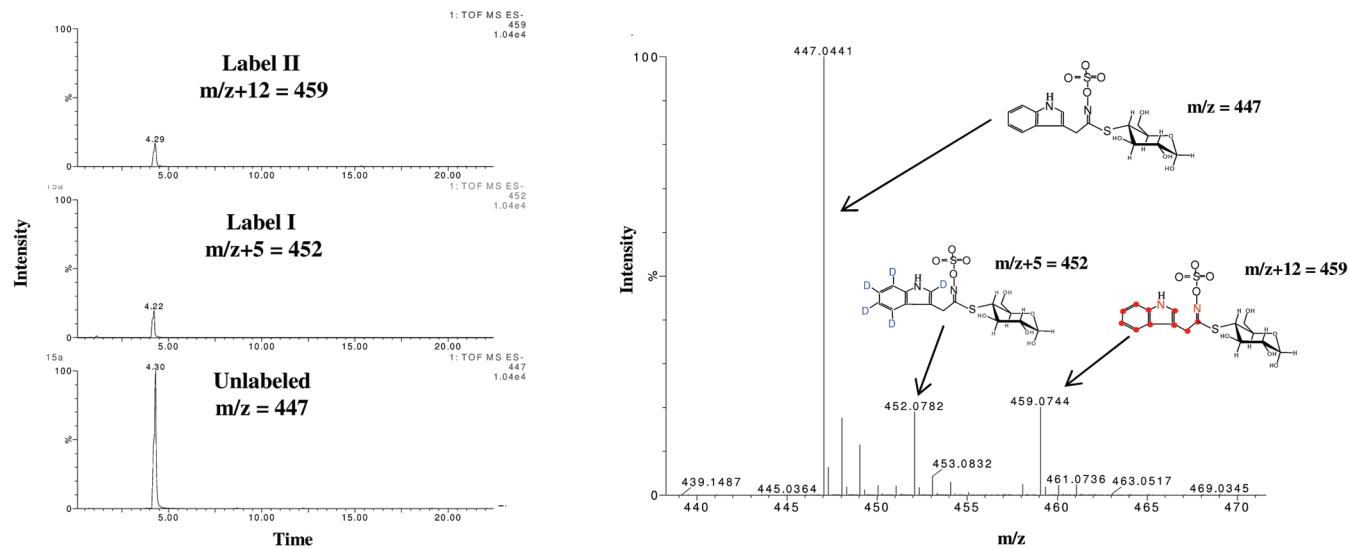


Figure 3. Metabolite detection using the DLEMMA approach. LC/ESI($-$)/MS/RSIM chromatograms retrieved from the chromatogram of dual labeled Trp fed leaves extract (left picture) and spectrum of 3-indolylmethyl glucosinolate (right picture). The chromatogram contains three co-eluting peaks, m/z , the unlabeled metabolite, $m/z+5$, the deuterated metabolite and $m/z+12$, metabolite labeled with ^{13}C and ^{15}N . The mass spectrum obtained from this chromatogram provides an accurate mass of the three signals.

Figure 1B). In preliminary experiments for tuning and examining the system, the following aspects were considered: (i) feeding time, solution volume, and concentration for optimized labeling efficiency while minimizing the amount of fed precursor, (ii) reproducibility, (iii) reliability of relative quantification using the dual labeling approach, (iv) labeling efficiency and the bias of different labels for incorporation. The exact details of these experiments are described in Supporting Information, Results S1, and Tables S2 and S3.

Next, we designed a complete, six components experimental setup that integrates all control and sample analyses required when carrying out a DLEMMA identification and quantification experiment (Figure 1C). These include: (I) WT leaves fed with unlabeled precursor (WT-UL), (II) Altered (ALT; either genetically or by a certain treatment) leaves fed with unlabeled precursor (ALT-UL), (III and IV) feeding with a mixture of both labels (L1, L2), to evaluate the label bias (samples WT-L1+L2, ALT-L1+L2), (V and VI) swapped label experiments (WT-L1/ALT-L2 and WT-L2/ALT-L1). In cases when the purpose of the experiment is only for identification, only two out of the six experimental setup components could be performed, namely, feeding leaves with unlabeled precursor (e.g., WT-UL) and feeding with a mixture of the two labels (e.g., WT-L1+L2).

Detection of Metabolites Using DLEMMA. Examples of reconstructed selected ion monitoring (RSIM) chromatograms and spectra containing unlabeled Trp metabolites together with stable isotopes co-eluting metabolites are presented in Figure 3. In the Trp fed precursor data, a total of 54,204 and 22,470 masses were detected by the peak picking program in ESI($+$) and in ESI($-$), respectively. In the Phe fed precursor data, 55,338 and 32,790 masses were detected in ESI($+$) and in ESI($-$), respectively. To filter masses belonging to metabolites which are derived from Phe and Trp we developed a program that first automatically identifies the labeled masses and subsequently detects the corresponding unlabeled masses (see Experimental Section). The program output contains the average values for retention time (r.t.), m/z , and intensities of unlabeled metabolites and their

labeled counterparts in the WT-L1/ALT-L2 and WT-L2/ ALT-L1 components of the experimental setup.

Applying the program to the Phe feeding data resulted in the detection of 81 and 61 mass signals in ESI($+$) and ESI($-$), respectively, while in the Trp feeding 93 and 40 mass signals were detected in ESI($+$) and ESI($-$), respectively. Clustering the 81 and 61 Phe-derived signals resulted in 32 and 34 metabolites, respectively. The 93 and 40 Trp-derived mass signals were clustered to 34 and 20 metabolites, respectively.

Identification of Metabolites Using DLEMMA. Metabolite identification using DLEMMA is based on the observed pattern of labeled atoms of either one of the labeled metabolites detected in their mass spectrum. The following example from the MJ treatment experiment (i.e., Trp feeding) illustrates the power of DLEMMA for metabolite identification. The metabolite detected in ESI($-$) at m/z 332.1410 displays the labeling pattern of $m/z+4$ and $m/z+13$ (Figure 4A), indicating that one labeled metabolite is labeled with 4 deuterium [hence one of the deuteriums in 5 deuterated Trp, label I (Figure 1A), was substituted during the biosynthesis of the metabolite] and the second labeled metabolite was labeled with ^{13}C and ^{15}N [this metabolite still contains all the labeled atoms derived from the Trp labeled II precursor (Figure 1A)]. Elemental composition search using the observed mass (m/z 332.1410) of the unlabeled metabolite provided three possible elemental compositions with high mass accuracy, two of them ($\text{C}_{18}\text{H}_{23}\text{NO}_3\text{P}$ and $\text{C}_{20}\text{H}_{18}\text{N}_3\text{O}_2$) with a low enough i-Fit [the fitness between the observed and calculated isotopic ratio; better match with a lower i-Fit (Figure 4B)]. Both elemental compositions ($\text{C}_{18}\text{H}_{24}\text{NO}_3\text{P}$ and $\text{C}_{20}\text{H}_{19}\text{N}_3\text{O}_2$ for neutral molecules) were used for a SciFinder search in comprehensive databases that together contain more than 33 million organic and inorganic substances. The formula $\text{C}_{18}\text{H}_{24}\text{NO}_3\text{P}$ exhibited 82 possible hits but none of these was a Trp derivative (Figure 4C). The formula $\text{C}_{20}\text{H}_{19}\text{N}_3\text{O}_2$, exhibited 2074 possible hits, and 227 of them were Trp derivatives. Out of the 227, 153 compounds could match the labeling pattern of $m/z+4$ (label I), meaning that 153 molecular structures contained four

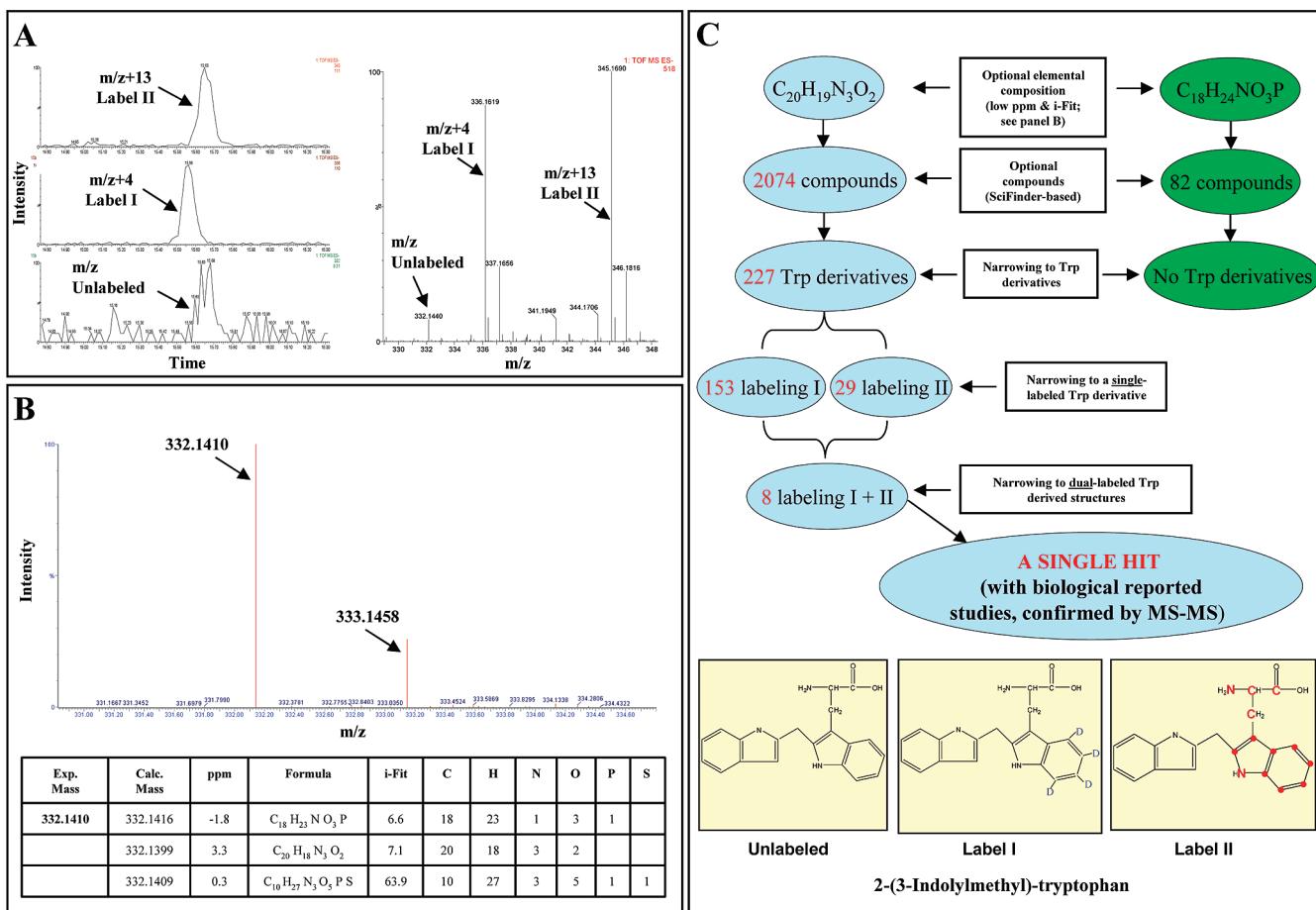


Figure 4. Various stages in the metabolite identification procedure when using the DLEMMA approach. (A) LC-ESI(-)MS/RSIM chromatograms (left picture) and spectrum (right picture) of a Trp unlabeled metabolite (m/z) and two labeled metabolites ($m/z+4$ and $m/z+13$). (B) Elemental composition search of the Trp unlabeled metabolite detected at m/z 332.1410 Da. (C) Example of a structure elucidation procedure obtained by the use of SciFinder search of the optional formulas (selected based on low ppm and iFit). Narrowing down the amount of possible structures (up to a single hit) was carried out by using labeling pattern consideration (introduction of parts of the molecule that did not change), manual inspection, and MS-MS fragmentation (see Experimental Section).

positions of unsubstituted indolic ring. Twenty nine compounds could match the labeling pattern of $m/z+13$ (label II), meaning that 29 molecular structures contained 11 original Trp carbons and 2 original Trp nitrogen atoms. Only 8 compounds could fit the combination of both $m/z+4$ and $m/z+13$, meaning that 8 molecular structures contained four unsubstituted positions in the indolic ring together with 11 original Trp carbons and 2 original Trp nitrogen atoms. Only one compound (2-(3-Indolylmethyl)-tryptophan, metabolite no. 11, Supporting Information, Table S4) out of the eight structures was reported in any biological study, and its annotation was backed up by a MS-MS spectrum (Figure 4C). The comparison of the MS-MS spectra of labeled and unlabeled compounds provides additional information for structure elucidation. MS chromatograms and spectra of unlabeled and two-labeled 3-indolylmethyl glucosinolate (another Trp-derived metabolite) are presented in Figure 3. Their MS-MS spectra are shown in Figure 5. Common fragments indicate the unlabeled parts of the molecule while different fragments point toward the labeled parts of the molecule.

The significant reduction in the number of possibilities for metabolite identity obtained by the use of DLEMMA allowed the

putative identification of 17 Trp metabolites (Supporting Information, Table S4). The annotation of all structures was supported by MS-MS fragmentations, four of them were also examined and verified by commercially available standards (Supporting Information, Table S5). In the Phe labeling experiments we have putatively identified thirty two metabolites (will be published elsewhere).

Relative Quantification of Metabolites Using DLEMMA.

The LC-MS chromatograms displayed in Figure 6 provide an example for the detection of differential metabolites. A Phe-derived differential metabolite (Quercetin 3-O[6-O(ramnosyl)-glucoside] 7-O-rhamnoside) was detected using the swapped label feeding experiments ($m/z+9$ and $m/z+6$ metabolites derived from either PAP1 or WT fed leaves). A differential ratio between levels of the two samples was calculated using the ratios of intensities between the labeled metabolites obtained from both chromatograms (Figure 6 and equation in the Experimental Section). Twenty-six differential Phe-derived metabolites were detected by comparison of WT with PAP1 leaves. Twenty-five metabolites showed higher levels in PAP1 leaves and a single metabolite exhibited lower levels in PAP1 leaves. Five differential Trp metabolites were observed as a result of the MJ treatment (Supporting Information, Table S1), four of them exhibited higher levels in the MJ treated samples

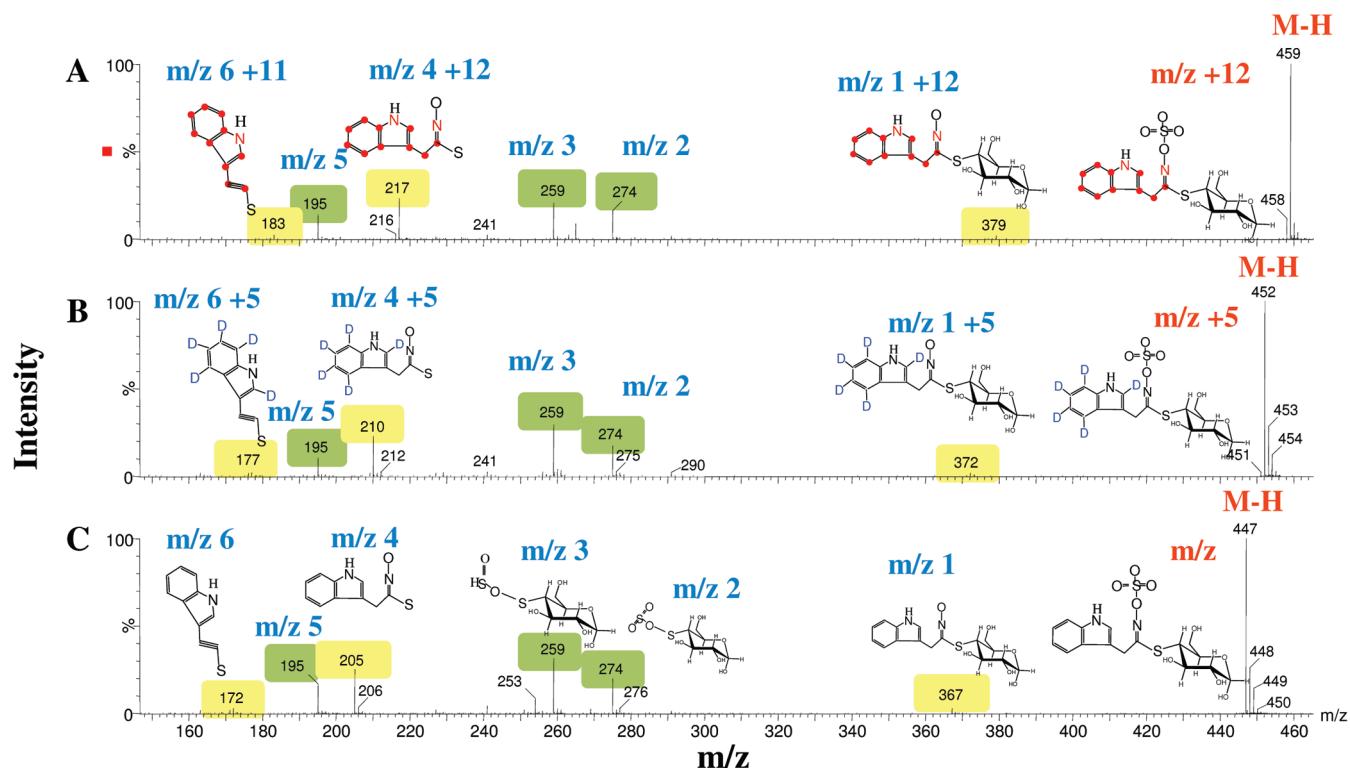


Figure 5. LC-ESI($-$)/MS-MS spectra of 3-indolylmethyl glucosinolate (unlabeled and labeled) obtained from dual labeled Trp fed leaves extract. **(A)** MS-MS of labeled II, 3-indolylmethy glucosinolate ($m/z+12$). **(B)** MS-MS of labeled I, 3-indolylmethy glucosinolate ($m/z+5$). **(C)** MS-MS of unlabeled 3-indolylmethy glucosinolate (m/z). From m/z 1 to m/z 6 are six fragments obtained in CID fragmentation. The green color indicates the unlabeled fragments, and the yellow one marks the labeled fragments.

(metabolites 5, 7, 10, and 11) and one showed lower levels (metabolite 16).

While the precision is maintained in DLEMMA experiments as compared to standard metabolomics assays, this approach has an advantage over the regular methods regarding accuracy. The precision obtained for biological replicates (3 replicates for each metabolite) is in the range of 4%–45% (with an average RSD of 25%). Dissimilar sample matrixes can influence differently on the ESI response of the same analytes and as a consequence decrease the accuracy of relative quantification. Figure 7 demonstrates how false results for relative quantification obtained because of problems of matrix effects can be prevented using DLEMMA. Coumaric acid can be determined as non-differential when performing the analysis separately (i.e., separated matrixes; Figure 7A), since in this example only the ALT sample matrix contains an interfering abundant peak which suppresses the peak of interest, thus leading to the unreliable relative quantification and as a result to an incorrect biological interpretation. DLEMMA (ESI analysis of the combined matrix; Figure 7B) provides the same matrix effect for both labeled peaks of coumaric acid and therefore leading to a more reliable relative quantification (Figure 7C).

DISCUSSION

In recent years, significant progress was made in developing technologies that could simultaneously gather information on the huge repertoire of small molecules that portray the organism's metabolic activities. Finding the optimal balance between the accuracy of identification and quantification on one hand and the full coverage of the metabolome on the other is a major challenge of today's metabolomics approaches. Here

we report on the development of a new tool for reliable detection, identification, and relative quantification of metabolites. DLEMMA is based on dual labeling of precursor metabolites and is therefore conceptually similar to strategies used in transcriptomics [array technology^{30,31} and proteomics (isotope-coded affinity tags techniques^{32,33}).

The use of Ionizable Isotopic Labeling Reagents for derivatization of amines and carboxylic acids and subsequent LC-MS analysis has been reported.^{24,25} However, this method is restricted to the detection of a unique functional group and thus metabolite identification and quantification is limited. Recently, Hegeman et al.¹³ proposed a powerful strategy for elemental composition elucidation by comparing observations of metabolites extracted from *Arabidopsis* seedlings grown on either normal or ^{13}C / ^{15}N enriched media. The same authors suggested that their approach was also promising for quantitative comparisons of metabolites in biological samples.¹² Giavalisco et al.,^{11,12} suggested a strategy for metabolite analysis based on high resolution mass spectrometry and ^{13}C -isotope labeling of the entire *Arabidopsis* metabolome (growing the *Arabidopsis* plants in a $^{13}\text{CO}_2$ atmosphere). It also enables relative quantification for differential analysis by spiking of ^{12}C grown sample with a fixed amount of ^{13}C -derived metabolite extract. Although these

- (30) Lockhart, D. J.; Dong, H.; Byrne, M. C.; Follettie, M. T.; Gallo, M. V.; Chee, M. S.; Mittmann, M.; Wang, C.; Kobayashi, M.; Horton, H.; Brown, E. L. *Nat. Biotechnol.* **1996**, *14*, 1675–1680.
- (31) Schena, M.; Shalon, D.; Davis, R. W.; Brown, P. O. *Science* **1995**, *270*, 467–470.
- (32) Aebersold, R.; Mann, M. *Nature* **2003**, *422*, 198–207.
- (33) Gygi, S. P.; Rist, B.; Gerber, S. A.; Turecek, F.; Gelb, M. H.; Aebersold, R. *Nat. Biotechnol.* **1999**, *17*, 994–999.

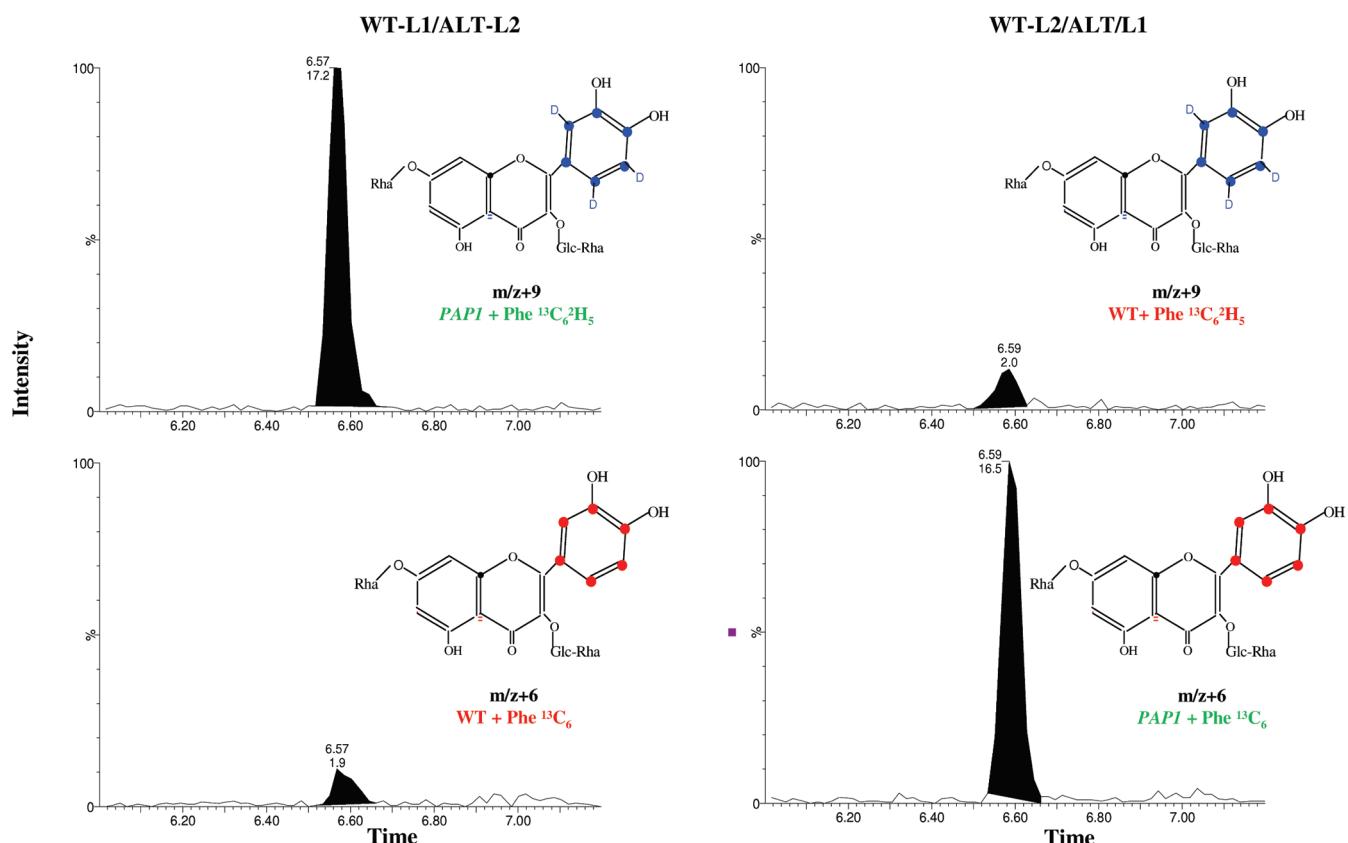


Figure 6. Relative quantification of metabolites using the DLEMMA approach. LC-MS chromatograms of Phe differential metabolite [Quercetin 3-O(6-O(ramnosyl)-glucoside) 7-O-rhamnoside] obtained from a swapped label feeding DLEMMA experiment. The left chromatogram contains the $m/z+6$ metabolite derived from WT and the $m/z+9$ metabolite derived from PAP1. The right chromatogram contains the $m/z+6$ metabolite derived from PAP1 and the $m/z+9$ metabolite derived from WT.

approaches are most valuable and lead to reduction of the ambiguity associated with assignment of elemental composition, they do not facilitate the reduction in the large number of structural isomer candidates for a given elemental composition. DLEMMA which is based on *in vivo* dual labeling of metabolites in pathways downstream of the fed precursor provides an efficient filter tool for isomer candidates as discussed below. Although we report on the use of DLEMMA for studies in plants, it might be possible to exploit this approach for metabolic investigations in diverse biological systems such as bacteria, yeast, animal and human cell lines.

Using DLEMMA for relative quantification minimizes experimental errors and increases accuracy because of the combined sample preparation and the subsequent single run. It provides identical matrixes for both samples and therefore avoids the effects of different complex biological matrixes on ESI that could alter compound measurements.^{4,18,19} In DLEMMA, the isotopic pair of a labeled compound (originating from different biological samples), co-elute within a single LC-MS run and hence exhibit a similar retention time ($\Delta r.t. < 0.1$ min) and are electrosprayed from identical solution conditions. Future developments might include the simultaneous analysis of more than two different samples (multiple labeling), for example, by triple labeling of Trp with Trp^2H_5 , $\text{Trp}^{13}\text{C}_{11}^{15}\text{N}_2$, and $\text{Trp}^{13}\text{C}_4^{2}\text{H}_3^{18}\text{O}_2$. As multiple fields of research utilize stable isotope labeled biomolecules, the commercial availability of such substances is constantly increasing.

DLEMMA is a potent tool for metabolite annotation, a major drawback in the current metabolomics experiments. The strength of DLEMMA for the identification process is based on two main filtering factors; first, the structure of the fed precursor metabolite is known and can serve as a stringent parameter for filtering, and second, is the information obtained through dual labeling. Applying accurate mass with isotope pattern recognition reduces the search space for possible elemental composition.³⁴ Yet, for a given elemental composition, several hundred to billions of isomers can be constructed, depending on the number and nature of elements given by the chemical composition.³⁵ The filtering capability of DLEMMA for structure elucidation allows the query in very extensive metabolite databases such as CAS, which is the largest database available for small molecules containing more than 33,000,000 entries (~20,000,000 entries for them are organic substances).

The structural elucidation strategy performed in DLEMMA was demonstrated here by using Trp^2H_5 and $\text{Trp}^{13}\text{C}_{11}^{15}\text{N}_2$ for dual labeling of Trp metabolites. Three steps for reducing structural isomer candidates were typically performed. First, restriction to Trp derivatives (isomers containing the indole ring), which in most cases removed the majority of false candidates (yet, in most cases 100–200 potential isomers were remained). Second, reducing the number of the potential Trp derivatives based on the observed pattern of labeled atoms of either one of the

(34) Roussis, S. G.; Proulx, R. *Anal. Chem.* 2003, 75, 1470–1482.

(35) Faulon, J. L. *J. Chem. Inf. Comput. Sci.* 1994, 34, 1204–1218.

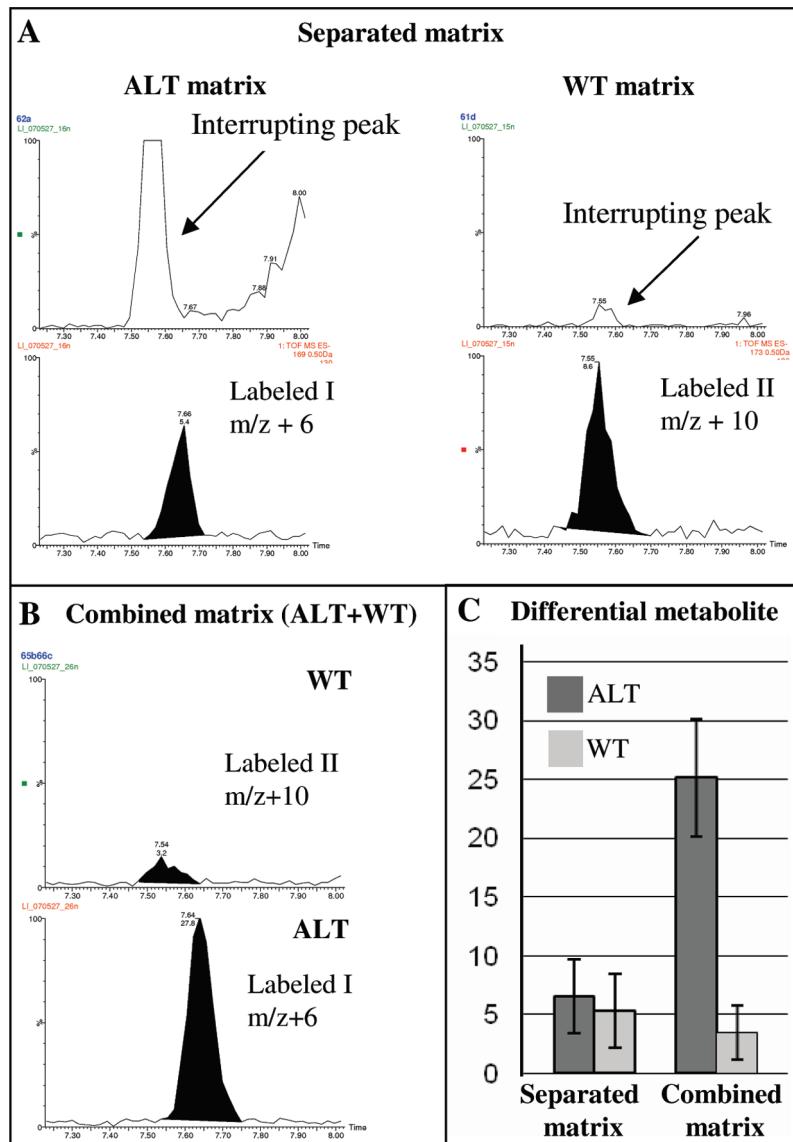


Figure 7. Reducing matrix effects using the DLEMMA approach. **(A)** Chromatograms obtained from feeding experiments followed by separate analysis of labeled coumaric acid in two different matrixes: The left chromatogram demonstrates the analysis of labeled I coumaric acid and an abundant interfering peak in the same r.t. obtained from the altered sample matrix (ALT, in this case leaves of PAP1). The right chromatograms were obtained from the WT matrix and display the analysis of labeled II coumaric acid without the interruption observed in the ALT sample matrix. **(B)** Chromatogram of labeled I and labeled II coumaric acid analyzed in the combined matrix (WT + ALT), as in the DLEMMA approach. **(C)** The pair of columns on the left represent the average response of coumaric acid derived from LC-MS analysis in separated matrixes while those on the right display the average response of coumaric acid in the combined matrix. A *t*-TEST ($n = 4$) was performed for separated matrixes (WT or ALT) and the combined matrix (WT + ALT). Only in the combined matrix experiment (as in DLEMMA), where labeled I and II samples were electrosprayed in the same conditions (therefore with the same matrix effects), coumaric acid was found to be statistically different ($p < 0.05$) between the two samples.

labeled metabolites, which removed 50%–97% of the false candidates from the potential Trp isomers. The combination of labeling patterns obtained from either one of the labeled metabolites was found to be significant for the reduction of Trp isomer candidates as compared to using the labeling pattern of a single labeled metabolite. A third, optional step, is the selection of Trp metabolites with biological reported studies (indicated in the CAS database output). After applying these selection criteria, we were typically left with one or two hits for each elemental composition. Still, confirmatory information from mass spectral fragmentation was required but using the DLEMMA approach much fewer structure elucidations had to be performed. When needed, MS-MS spectra of the labeled

metabolites were compared to those of the unlabeled ones to provide additional important supporting information for structure elucidation. Tryptophan derived metabolites putatively identified through DLEMMA included 11 that were not previously reported in the model plant *Arabidopsis* and 8 of them to the best of our knowledge were not described in any other plant (Supporting Information, Table S1). These new findings obtained using the DLEMMA approach, can be employed for the prediction of a Trp pathway map.

CONCLUSION

Here we report on a new method that complements information obtained by more global metabolomics analyses. This ap-

proach is based on feeding two differentially labeled precursors and following the downstream labeled metabolites using Liquid Chromatography coupled to high resolution mass spectrometry (UPLC-MS-QTOF). This approach is an efficient filtering tool, not only for eliminating the number of elemental compositions, but especially for reducing the number of structural isomer candidates for a given elemental composition. It also enables better relative quantification of differential metabolites, eliminating the effects of different matrixes on ESI. The comprehensiveness of this method is limited because of the use of a particular precursor molecule (e.g., aromatic amino acids) for feeding experiments. From the same reason, however, it is most valuable for certain experimental systems that examine a particular metabolic pathway. Thus, DLEMMA could have an important contribution for pathway discovery and the extension of the current attempts to reconstruct metabolic networks.

ACKNOWLEDGMENT

The first two authors (L.F. and I.V.) contributed equally to this work. A.A. is an incumbent of the Adolfo and Evelyn Blum Career Development Chair. The research in the A.A. laboratory was supported by a grant from The Israel Ministry of Science (IMOS project #3-2552), The EU project “META-PHOR”, contract number FOODCT-2006-036220, Mr. and Mrs. Mordechai Segal, and by the Benoziyo Institute. We thank Arie Tishbee for operating the LC-MS instrument.

SUPPORTING INFORMATION AVAILABLE

Additional information as noted in the text. This material is available free of charge via the Internet at <http://pubs.acs.org>.

Received for review July 6, 2009. Accepted October 9, 2009.

AC901495A