Plant Metabolomics and Strategies

Chapter · August 2016

DOI: 10.1007/978-3-319-31703-8_17

3 authors:

Halbay Turumtay
Joint BioEnergy Institute
27 PUBLICATIONS **239** CITATIONS
SEE PROFILE

Cemal Sandalli
Recep Tayyip Erdoğan Üniversitesi
57 PUBLICATIONS **912** CITATIONS
SEE PROFILE

Emine Akyuz Turumtay
Lawrence Berkeley National Laboratory
41 PUBLICATIONS **709** CITATIONS
SEE PROFILE

Plant Metabolomics and Strategies

Halbay Turumtay, Cemal Sandallı, and Emine Akyüz Turumtay

Contents

Abstract Plant metabolomics is the study of metabolic pathways and processes to understand how plants grow and carry out functions. Over the past years, there has been a switch to a more holistic view of metabolomes. Instead of looking at individual pathways and metabolites, scientists are looking at interactions, broad pictures, and regulations. Combination of spectrometry-based technology database with (un)supervised multivariate statistical methodologies reveals a deep insight into complex metabolite patterns of plant derived samples. The ambitious goal of identifying the structure of all metabolites and hence all metabolic pathways is almost becoming within reach as MS- and NMR-based technologies and data mining approaches are rapidly evolving.

Keywords LC-MS • Metabolomics • Processing of metabolic data • NMR • GC/ MS

H. Turumtay (✉)
Department of Energy System Engineering, Karadeniz Technical University, 61830 Trabzon, Turkey
e-mail: halbay.turumtay@gmail.com

C. Sandallı
Department of Biology, Recep Tayyip Erdogan University, 53100 Rize, Turkey

E.A. Turumtay
Department of Chemistry, Recep Tayyip Erdogan University, 53100 Rize, Turkey

K.R. Hakeem et al. (eds.), *Plant Omics: Trends and Applications*,
DOI 10.1007/978-3-319-31703-8_17

1 Introduction

Recent development in technologies on DNA sequencing (genomics), gene expression analysis (transcriptomics), and protein analysis (proteomics) were not enough to identify different metabolic pathway in plants. The missing link in functional genomics strategies might fill by metabolomics studies. Metabolomics is the term used for high-throughput analysis of complex metabolite mixtures in plant extracts. Metabolic data outputs reveal a completely different level of information from genomic source, which does not necessarily show collinearity with genome sequencing (Saito and Matsuda 2010). The total number of the metabolites of a single plant is very speculative and is expected to exceed the number of currently known metabolites multiple times. Roessner et al. (2001) and von Roepenack-Lahaye et al. (2004) estimated that Arabidopsis comprise approximately 5000 metabolites of which 1015 have been identified today. For the plant kingdom, almost 100,000 (mainly secondary) metabolites have been identified to date (Oksman-Caldenty et al. 2004). Many important crops have already been targeted for their metabolomics profiling such as rice (Hall et al. 2008), wheat (Graham et al. 2009), tomato (Moco et al. 2006), melon (Moing et al. 2011), coffee (Lindinger et al. 2009), and potato (Beckman et al. 2007).

The metabolome consist of a variety of metabolites with different physic-chemical properties (polarity, acidity, molecular weight, extractability, affinity.). It also extends over an estimated nine magnitudes of concentration (pM-mM) (Dunn and Ellis 2005). This complexity, in combination with the technical limits of analytic instruments, makes it currently impossible to get overview of the entire metabolome in one single or small number of analyses. All current extraction and detection techniques, irrespective of their level of sophistication, have an unavoidable intrinsic bias toward certain metabolite groups (Hall 2006).

Based on objectives of the study on the one hand and technical limits and capacities on the other hand, metabolite analysis is divided into four classes (Fiehn 2001).

1. Metabolome targeted analysis. By means of this type of analysis, quantitative information on one or selected group of metabolites is congregate. Extensive extraction and separation are commonly obliged to avoid interfering metabolites. Mostly, a high sensitive detector is used (e.g., fluorescence and single ion monitoring (SIM-MS)). This techniques is needed when low detection limits are required such as detection of phytohormones (Prinsen et al. 2000)
2. Metabolic profiling. This approach aims at detecting several predefined targets that are typically metabolites of specific pathways or with related chemical structures such as amino acids. The techniques comprise both sensitive and specific instruments. Gas chromatography (GC), liquid chromatography (LC), and capillary electrophoresis (CE) are generally used in combination with common detector such as mass spectrometry (MS) and photodiode array (PDA) (Theodoridis et al. 2012).
3. Metabolomics. This class is rather theoretical since it refers to the identification and quantification of whole metabolome. Because this is practically impossible,

the term "metabolomics" is commonly used when the metabolome is studied in unbiased way, even if only a fraction of it is revealed. Generally same chromatographic tools are used as in metabolic profiling, commonly combined with MS-detection.

4. Metabolic fingerprinting. This is a high-throughput qualitative screening of metabolic composition with primary aim of sample comparison and classification. Although in general no attempt is made to identify the metabolites, these fingerprints can provide information on functional groups or compound classes. Typical tools are Fourier transformation cyclotron resonance mass spectrometry (FT-ICR/MS) (Johnson et al. 2003), NMR (Reo 2002; Ratcliffe and Shachar-Hill 2005), and direct injection mass spectroscopy (DIMS) (Dunn and Ellis 2005; Aharoni et al. 2004)

2 Recent Development on Plant Metabolomics

As stated before, a remarkably broad metabolic profile can be achieved in single analysis (Kikuchi et al. 2004; von Roepenack-Lahaye et al. 2004; Tohge et al. 2005), but not one single technique is able to measure the complete metabolome. Multiple parallel chromatographic techniques are often required to gain desired broad metabolic picture (Hirai et al. 2004). For instance, untargeted (GC-MS) and targeted (HPLC-MS) metabolic profiling analyses were performed on strawberry fruits (Zhang et al. 2011).

The techniques that come closest to the definition of metabolomics are generally hyphenated techniques. Among these are LC/MS, GC/MS, and CE/MS. LC/MS is the most used with reversed phase for detection of secondary metabolites (Murch et al. 2004). However, LC/MS with hydrophilic interaction chromatography (HILIC) was also used for detection of oligosaccharides and sugar nucleotides (Tolstikov and Fiehn 2002). GC/MS is the gold standard to detect sugars, sugar alcohols, amino acids, volatile compounds, small organic acids, and lignin monomers (G type, S type, and H type oligolignols (Roessner et al. 2001)).

In order to perform performing in vivo metabolomics in plants Laser Ablation Electrospray Ionization-Mass Spectrometry Imaging (LAESI-MSI) has been successfully applied (Etalo et al. 2015). These recent finding results indicate LAESI-based imaging approaches have a potential of as a reliable and fast way to perform metabolomics analyses on living tissues (Etalo et al. 2015).

3 LC-MS

LC/MS is the mostly used with reverse phase and gradient elution in metabolite profiling of plants. This makes it possible to separate rather less polar metabolites, usually secondary metabolites. Highly polar compounds (e.g., sugars) do not show

affinity for the hydrophobic stationary phase and will coelute in a cluttered and useless front. Change of the pH and the polarity of the mobile phase determine mostly which metabolites could be separated and thus detected. The most frequently used ionization sources in LC/MS are atmospheric pressure chemical ionization (APCI) and electrospray (ESI) to ionize molecules with acid and/or base functionalities.

A new generation of mass spectrometers such as FT-ICR/MS and Orbitrap show a dramatic improvement of scan rate, sensitivity, resolution, and accuracy (Hirai et al. 2004; Hu et al. 2005; Zubarev and Makarov 2013). These technical improvements might increase the number of the compounds that can be measured in one chromatographic separation and has become a powerful addition to mass spectrometric techniques for increasing selectivity and confidence of routine analyses.

4 GC/MS

GC/MS is valuable tool to detect a variety compounds simultaneously, such as sugars, sugar alcohols, amino acids and small organic acid. However, it is limited in detection of organic diphosphates, cofactors, and metabolites larger than trisaccharides (Weckwerth and Fiehn 2002). The method is claimed to be sensitive, qualitative, reproducible, and relatively fast (one separation takes approximately 1 h). Besides, GC/MS is suitable for automation (Roessner et al. 2000).

Biological variation of plant metabolites measured by GC/MS has been shown to be substantial and to surpass the technical variation severalfold (Roessner et al. 2000; Hall 2006; Kose et al. 2001). These points to the remarkably high flexibility of the metabolome, without affecting the visual phenotype. Moreover, it also illustrates the need to analyze enough biological replicates to estimate the biological variation within the system whereas technical repeats are less important. In case of a typical metabolic profiling, it is advised to work with approximately 12 biological replicates. (Fuzfai et al. 2004; Kaiser and Benner 2012; Pitthard and Finch 2001; Ratsimba et al. 1999)

GC/MS can be applied to detect volatile metabolites (e.g., alcohols and monoterpenes). However, the technology is more broadly applicable to group of nonvolatile, polar (mainly primary) metabolites, such as amino acid, organic acids and sugars, by converting them into volatile and thermostable compounds through chemical derivatization. Derivatization reduces the polarity of the functional groups thereby facilitating their separation by GC. Carbonyl functions are generally converted to the methoxime and acidic protons, which are replaced by silyl-group (such as trimethylsilyl-TMS) (Gulberg et al. 2004; Halket et al. 2005). Despite the big advantages of derivatization, it has some drawback as well. The presence of sterically hindered groups can lead to partial derivatization, especially in case of sterical trimethylsilyl groups. Consequently multiple products are formed from a single metabolite, hereby complicating the chromatogram (Halket et al. 2005). Furthermore, derivatised products are relatively labile, which result in concentration changes over time. However, these drawbacks could be complemented by preparing samples in daily batches, with a common reference throughout the measurement period (Tikunov et al. 2005).

5 NMR

NMR based metabolomics studies is an important tool for biological system and have been applied in various organism including animals, plants, and bacteria. ^{1}H NMR has been predominant profiling method since it fast, simple and could be used for different purposes in plant metabolomics such as quality control (Rasmussen et al. 2006), chemotaxonomy (Roos et al. 2004; Le Gall et al. 2004), and analysis of transgenic plants (Le Gall et al. 2004; Colquhoun 2007). Up to now, a plethora of applications of NMR based metabolomics have been reported. Generally, some 30–150 metabolites are identified (Kim et al. 2011).

The major advantages of high-throughput analysis of NMR analysis are the ease of quantification and simple sample preparation (not at al in certain cases). NMR could also provide information on the absolute quality of metabolites, and thus the ratio and amount of components in a mixture can be determined (Kim and Ralph 2014). Furthermore, NMR also provides a data about molecules stereochemical details (Seger and Sturm 2007). Compared to MS, the weakness point of NMR is its low sensitivity, although recent findings have led to a considerable increase in sensitivity of NMR (Grivet and Delort 2009).

6 Metabolite Identification

One of the challenges in metabolomics is the identification of unknown metabolites. Because of the huge discrepancy between the amount of sample needed to detect unknown peaks with very low abundance and the amount of sample required for structural elucidation of unknown compounds, only a fraction of the measured peaks could be assigned to known compounds. In GC/MS typically could detect metabolites are known and numbers in LC/MS are comparable. Depending on used method, different strategies are possible to unveil the identity of unnamed peak. Interpretation of mass spectral data is often an intricate and time-consuming task. The use of MS-libraries in combination with the retention time or retention indices (RI) is a powerful tool to identify metabolites (Roessner et al. 2001). These MS libraries could be constructed based on chemically synthesized standards. If the molecule is not present in a library, classical analytic chemical tools can be considered (for instance purification and spectroscopic analysis, e.g., NMR, UV, IR).

The current generation of high resolution mass spectrometers such as FT-ICR/MS, LC-PDA-SPE-NMR/MS, and Orbitrap might simplify the identification of unknown peaks based on the MS^{n} spectra. Different techniques were applied will lead to further development of plant metabolomics have been reviewed by Nakabayashi and Saito (2013). Moreover, the high sensitivity of this spectrometer allows us to detect lowly abundant metabolites.

7 Processing of Metabolic Data

Manual processing of small metabolite screening dataset is feasible but its time consuming and only dramatic changes could be observed. In practice, datasets are often too big to look at the metabolites one by one and in addition subtle changes are informative as well. Therefore, thigh-throughput data processing methods are required for metabolomics. One of the elegant ways of this high-throughput system is a candidate substrate–product pair (CSPP) system, which has been developed by Morreel et al. (2014). Morreel et al. (2014) developed an algorithm in which liquid chromatography-mass spectrometry profiles are searched for pairs of peaks that have mass and retention time differences corresponding with those of substrates and products from well-known enzymatic reactions. This method allows the annotation of low-abundance compounds that are otherwise not amenable to isolation and purification. This method will greatly advance the value of metabolomics in systems biology (Morreel et al. 2014).

8 Conclusion and Future Perspective

Metabolomics is a tool to improve our understanding of the metabolic pathways and biochemistry of organisms. It is certain that plant metabolomics is in its infancy and still in a dynamic phase of development. To get a deeper biological meaning of living organisms, metabolic studies need more data about known metabolites. In order to increase the number of known compounds, deeper technologies for identification of unknown compounds are certainly necessary. Recent technological improvements in both metabolite identification and data interpretation revealed that there are still lots of things to do.

References

Aharoni A, Ric de Vos CH, Verhoeven HA, Maliepaard CA, Kruppa G, Bino R, Goodenowe DB (2004) Nontargeted metabolome analysis by use of fourier transform ion cyclotron mass spectrometry. OMICS 6:217–234

Beckman M, Enot DP, Overy DP, Draper J (2007) Representation, comparison and interpretation of metabolome fingerprint data for total composition analysis and quality trait investigation in potato cultivars. J Agric Food Chem 55:3444–3451

Colquhoun IJ (2007) Use of NMR for metabolic profiling in plant systems. J Pestic Sci 32:200–212

Dunn WB, Ellis DI (2005) Metabolomics: current analytic platforms and methodologies. Trends Anal Chem 24:285–294

Etalo DW, De Vos CHR, Joosten MHAJ, Hall RD (2015) Spatially resolved plant metabolomics: some potentials and limitations of laser-ablation electrospray ionization mass spectrometry metabolite imaging. Plant Physiol 169:1424–1435

Fiehn O (2001) Combining genomics, metabolome analysis, and biochemical modelling to understand metabolic networks. Comp Funct Genomics 2:155–168

Fuzfai ZF, Katona E, Kovacs E, Molnar-Perl I (2004) Simultaneous identification and quantification of the sugar, sugar alcohol and carboxylic acid contents of sour cherry, apple, and ber fruits as their trimethylsilyl derivatives, by gas-chromatography-mass spectrometry. J Agric Food Chem 52:7444–7452

Graham SF, Amigues E, Migaud M, Browne RA (2009) Application of NMR based metabolomics for mapping metabolite variation in European wheat. Metabolomics 5:302–306

Grivet JP, Delort AM (2009) NMR for microbiology: in vivo and in situ applications. Prog Nucl Magn Reson Spectrosc 54:1–53

Gulberg J, Jonsson P, Nordstrom M, Moritz T (2004) Design of experiments: an efficient strategy to identify factors influencing extraction and derivatization of Arabidopsis thaliana samples in metabolomic studies with gas chromatography/mass spectrometry. Anal Biochem 331:283–295

Halket JM, Waterman D, Przyborowska AM, Patel RKP, Fraser PD, Bramley PM (2005) Chemical derivatization and mass spectral libraries in metabolic profiling by GC/MS and LC/MS/MS. J Exp Bot 56:219–243

Hall D (2006) Plant metabolomics: from holistic hope, to hype, to hot topic. New Phytol 169:453–468

Hall D, Brouwer ID, Fitzgerald MA (2008) Plant metabolomics and its potential application for human nutrition. Physiol Plant 132:162–175

Hirai MY, Yano M, Goodenowe DB, Kanaya S, Kimura T, Awazuhara M, Arita M, Fujiwara T, Saito K (2004) Integration of transcriptomics and metabolomics for understanding of global responses to nutritional stresses in Arabidopsis thaliana. Proc Natl Acad Sci U S A 101:10205–10210

Hu Q, Noll RJ, Li H, Makarov A, Hardman M, Cooks RG (2005) The Orbitrap: a new mass spectrometer. J Mass Spectrom 40:430–443

Johnson HE, Broadhurst D, Goodacre R, Smith AR (2003) Metabolic fingerprinting of salt-stressed tomatoes. Phytochemistry 62:919–928

Kaiser K, Benner R (2012) Characterization of lignin by gas chromatography and mass spectrometry using a simplified CuO oxidation method. Anal Chem 84:459–464

Kikuchi J, Shinozaki K, Hirayama T (2004) Stable isotope labeling of Arabidopsis thaliana for an NMR-based metabolomics approach. Plant Cell Physiol 45:1099–1104

Kim H, Ralph J (2014) A gel-state 2D-NMR method for plant cell wall profiling and analysis: a model study with the amorphous cellulose and xylan from ball-milled cotton linters. RSC Adv 4:7549–7560

Kim HK, Choi YH, Verpoorte R (2011) NMR-based plant metabolomics: where do we stand, where do we go? Trends Biotechnol 29:267–275

Kose F, Weckwerth W, Linke T, Fiehn O (2001) Visualizing plant metabolomic correlation networks using clique-metabolite matrices. Bioinformatics 17:1198–1208

Le Gall G, Colquhoun IJ, Defernez M (2004) Metabolite profiling using 1H NMR spectroscopy for quality assessment of green tea, Camellia sinensis (L.). J Agric Food Chem 52:692–700

Lindinger C, Pollien P, de Vos RCH, Tikunov Y, Hageman JA, Lambot C, Fumeaux R, Voirol-Baliguet E, Blank I (2009) Identification of ethyl formate as a quality marker of the fermented off-note in coffee by a nontargeted chemometric approach. J Agric Food Chem 57:9972–9978

Moco S, Bino RJ, Vorst O, Vehoeven HA, De Groot J, Van Beek TA, Vervoort J, De Vos CHR (2006) A liquid chromatography-mass spectrometry-based metabolome database for tomato. Plant Physiol 141:1205–1218

Moing A, Aharoni A, Biais B, Rogachev I, Meir S, Brodsky L, Allwood JW, Erban A, Dunn WB, Kay L, de Koning S, de Vos RCH, Jonker H, Mumm R, Deborde C, Maucourt M, Bernillon S, Gibon Y, Hansen TH, Husted SGR, Kopka J, Schjoerring JK, Rolin D, Hall RD (2011) Extensive metabolic cross-talk in melon fruit revealed by spatial and developmental combinatorial metabolomics. New Phytol 190:683–696

Morreel K, Saeys Y, Dima O, Lu F, Van de Peer Y, Vanholme R, Ralph J, Vanholme B, Boerjan W (2014) Systematic structural characterization of metabolites in Arabidopsis via candidate substrate–product pair networks. Plant Cell 26:929–945

Murch SJ, Rupasinghe HP, Goodenowe D, Saxena PK (2004) A metabolomic analysis of medicinal diversity in Huang-qin (Scutellaria baicalensis Georgi) genotypes: discovery of novel compounds. Plant Cell Rep 23:419–425

Nakabayashi R, Saito K (2013) Metabolomics for unknown plant metabolites. Anal Bioanal Chem 405:5005–5011

Oksman-Caldenty KM, Inze D, Oresic M (2004) Connecting genes to metabolites by a systems biology approach. Proc Natl Acad Sci U S A 101:9949–9950

Pitthard V, Finch P (2001) GC-MS analysis of monosaccharides mixtures as their diethyldithioacetal derivatives: application to plant gums used in art works. Chromatographia 53:317–321

Prinsen E, Van Laer S, Oden S, Van Onckelen H (2000) Auxin analysis. Humana Press Inc, Totowa, NJ

Rasmussen B, Cloarec O, Tang H, Staerk D, Jaroszewski JW (2006) Multivariate analysis of integrated and full-resolution 1H-NMR spectral data from complex pharmaceutical preparations: St John's Wort. Planta Med 72:556–563

Ratcliffe RG, Shachar-Hill Y (2005) Revealing metabolic phenotypes in plants: inputs from NMR analysis. Biol Rev 80:27–43

Ratsimba V, Garcia Fernandez JM, Defaye J, Nigay H, Voilley A (1999) Qualitative and quantitative evaluation of mono- and disaccharides in D-fructose, Dglucose and sucrose caramel by gas-liquid chromatography-mass spectrometry Di-D-fructose dianhydrides as tracers of caramel authenticity. J Chromatogr A 844:283–293

Reo N (2002) NMR-based metabolomics. Drug Chem Toxicol 25:375–382

Roessner U, Wagner C, Kopka J, Trethewey RN, Willmitzer L (2000) Simultaneous analysis of metabolites in potato tuber by gas chromatography-mass spectrometry. Plant J 23:131–142

Roessner U, Willmitzer L, Fernie A (2001) High-resolution metabolic phenotyping of genetically and environmentally diverse potato tuber systems Identification of phenocopies. Plant Physiol 127:749–764

Roos G, Roseler C, Buter KB, Simmen U (2004) Classification and correction of St John's wort extracts by nuclear magnetic resonance spectroscopy, multivariate data analysis and pharmacological activity. Planta Med 70:771–777

Saito K, Matsuda F (2010) Metabolomics for functional genomics, systems biology, and biotechnology. Annu Rev Plant Biol 61:463–489

Seger C, Sturm S (2007) Analytical aspects of plant metabolite profiling platforms: current standings and future aims. J Proteome Res 6:480–497

Theodoridis GA, Gika HG, Want EJ, Wilson ID (2012) Liquid chromatography-mass spectrometry based global metabolite profiling: a review. Anal Chim Acta 711:7–16

Tikunov Y, Lommen A, De Vos CHR, Verhoeven HA, Bino RJ, Hall RD, Bovy AG (2005) A novel approach for nontargeted data analysis for metabolomics large-scale profiling of tomato fruit volatiles. Plant Physiol 139:1125–1137

Tohge T, Nishiyama Y, Hirai MY, Yano M, Nakajima JI, Awazuhara M, Inoue E, Takahashi H, Goodenowe DB, Kitayama M, Noji M, Yamazaki M, Saito K (2005) Functional genomics by integrated analysis of metabolome and transcriptome of Arabidopsis plants over-expressing an MYB transcription factor. Plant J 42:218–235

Tolstikov VV, Fiehn O (2002) Analysis of highly polar compounds of plant origin: combination of hydrophilic interaction chromatography and electrospray ion trap mass spectrometry. Anal Biochem 301:298–307

von Roepenack-Lahaye E, Degenkolb T, Zerjeski M, Franz M, Roth U, Wessjohann L, Schmidt J, Scheel D, Clemens S (2004) Profiling of Arabidopsis secondary metabolites by capillary liquid chromatography coupled to electrospray ionization quadrupole time-of-flight mass spectrometry. Plant Physiol 134:548–559

Weckwerth W, Fiehn O (2002) Can we discover novel pathways using metabolomic analysis? Curr Opin Biotechnol 13:156–160

Zhang J, Wang X, Yu O, Tang J, Wan X, Fang C (2011) Metabolic profiling of strawberry (Fragaria × ananassa Duch) during fruit development and maturation. J Exp Bot 62:1103–1118

Zubarev RA, Makarov A (2013) Orbitrap mass spectrometry. Anal Chem 85:5288–5296