ELSEVIER

Integrating genomics and metabolomics for engineering plant metabolic pathways

Kirsi-Marja Oksman-Caldentey[1] and Kazuki Saito[2,3,4]

Plant metabolites are characterized by an enormous chemical diversity, every plant having its own complex set of metabolites. This variety poses analytical challenges, both for profiling multiple metabolites in parallel and for the quantitative analysis of selected metabolites. We are only just starting to understand the roles of these metabolites, many of them being involved in adaptations to specific ecological niches and some finding beneficial use (e.g. as pharmaceuticals). Spectacular advances in plant metabolomics offer new possibilities, together with the aid of systems biology, to explore the extraordinary complexity of the plant biochemical capacity. State-of-the art genomics tools can be combined with metabolic profiling to identify key genes that could be engineered for the production of improved crop plants.

Addresses
[1] VTT Biotechnology, PO Box 1500, 02044 VTT (Espoo), Finland
[2] Graduate School of Pharmaceutical Sciences, Chiba University, Inage-ku, Chiba 263-8522, Japan
[3] Kazusa DNA Research Institute, Kisarazu, Chiba 292-0818, Japan
[4] RIKEN Plant Science Center, Tsurumi-ku, Yokohama 230-0045, Japan

Corresponding author: Oksman-Caldentey, Kirsi-Marja (kirsi-marja.oksman@vtt.fi)

Current Opinion in Biotechnology 2005, **16**:174–179

This review comes from a themed issue on
Plant biotechnology
Edited by Dirk Inzé

Available online 3rd March 2005

0958-1669/$ – see front matter

DOI 10.1016/j.copbio.2005.02.007

Introduction

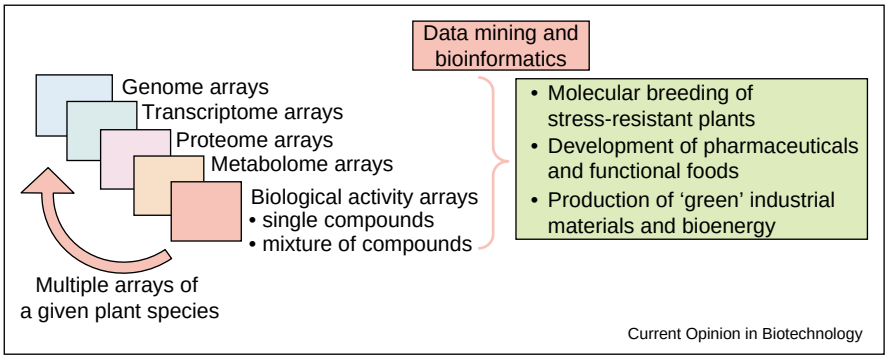
Plants are the most excellent designers and producers of a variety of small compounds that are beneficial to mankind as foods, medicines and industrial raw materials [1••]. Plants produce materials independently of fossil energies and resources and are thus regarded as 'ultimate factories'. However, to rely solely on naturally growing plants as a production system (e.g. for pharmaceuticals) can be extremely dangerous, as shown recently by severe shortage problems of artemisinin [2]. Therefore, an alternative biotechnological production system based on genetically engineered plant cells is an attractive approach. Engineering plant metabolic pathways is, however, a difficult endeavor, because our fundamental knowledge of metabolic control in plants is not yet sufficient for rational engineering of complicated metabolic networks. Of course the relatively simple approaches, such as the introduction of single genes for rate-limiting enzymes or new branch pathway enzymes (often from heterologous organisms), have led to spectacular successes [3,4]. However, there are also 'failures' that have not been published, and thus it is difficult to estimate the success rate of all trials of metabolic engineering. In most cases, difficulties arise from the fundamental lack of precise understanding of the entire network between genes, transcripts, proteins and metabolites in biological systems [5].

In contrast to traditional analysis that deals with a limited number of genes, proteins and metabolites, a genome-wide large-scale approach is now possible for the holistic systems analysis of some model plants such as *Arabidopsis thaliana*. Non-targeted analysis of cell components, integrated with genomics, analytical chemistry and bioinformatics, allows a better understanding of plant systems and will increase the practical potential of metabolic engineering. In particular, recent technological advances in metabolomics could open new avenues for precise metabolic engineering in plants. This article describes the cutting edge of the integrated analysis of genomics and metabolomics for metabolic engineering in plants. We also discuss some limitations of the technology and future challenges that need to be faced.

The development of plant metabolomics

Metabolomics, which deals with all cellular metabolites, has been recently recognized as an important sector of post-genome science. The general idea of 'metabolomics' or the 'metabolome' was first defined several years ago in the field of microbiology [6], and its importance in plant science was pointed out soon after [7]. Today, metabolomics is also a powerful tool in drug discovery and development; for instance, in the identification of drug metabolites or biomarkers for organ-specific toxicities [8]. Even in the absence of any visible change in a cell or individual plant, metabolomics, which allows phenotyping by exhaustive metabolic profiling, can show how cells respond as a system. Plant metabolomics is of particular importance because of the huge chemical diversity of plants compared with microorganisms and animals. The number of metabolites from the plant kingdom has been estimated at ~200 000 [9] or even more [10]. Even a single plant species such as *A. thaliana* might produce ~5000 metabolites [11•,12•,13]. These numbers are significantly greater than those of microorganisms (~1500)

Figure 1



The 'phytochemical array' — a concept for using various functional genomics-based arrays in plants. Multiple arrays to analyse the genome, transcriptome, proteome, metabolome and bioactivity of metabolites of a given plant species are the modern tools for understanding plants at various biological levels. These tools can lead to applications for crop improvement (e.g. molecular breeding of biotic and abiotic stress-resistant plants), to the discovery and development of plant-based pharmaceuticals, to the development of functional foods, and to the production of plant-derived industrial materials and energy. All of these advances will be valuable in the future and can be developed through genome-based plant biotechnology. The phytochemical array exhibits all of the links from the genome to the activity of metabolites for desired traits, and will lead to accelerated development.

and animals (~2500). Although the term 'metabolite (or metabolic) profiling' has been understood in the past to mean comprehensive profiling of metabolic change, metabolomics has an intrinsically distinct connotation because it implies integration with the other 'omic' sciences (Figure 1).

The comprehensive chemical analysis of metabolites and the computation of huge datasets are the key components of metabolomics [11•,14•,15,16]. Metabolomics is principally required to determine all metabolites in a plant extract; however, no single technology for metabolomics, such as a DNA sequencer for genomics or DNA arrays for transcriptomics, is available, and such a method may never be possible. This is because the analysis of metabolites of divergent physicochemical properties needs a wide range of chemistries: a single chemistry cannot deal with metabolomics as it can with nucleic acids and proteins. Metaphorically speaking, metabolomics is like attempting whole-genome sequencing without either the Sanger method or the Maxam–Gilbert method. At present, combinations of different analytical methods of high sensitivity are generally used for comprehensive non-targeted chemical analysis [17,18].

One of the key challenges of metabolite profiling and analysis is to find an optimal balance between accuracy and coverage of metabolite measurements. This has become particularly apparent when analysing plant secondary metabolites, which have very different chemical natures compared with primary metabolites. If one general extraction and analytical system is used it is likely that many metabolites will remain in the plant matrix and will not be profiled. One possibility is to split the metabolomics platform into subgroups of compounds sharing similar chemical extraction conditions, chromatographic separation and subsequent instrumental analysis [5].

Advances in instrument development have been tremendous. Traditionally, mass spectrometry (MS) combined with a number of chromatographic methods is used for metabolome analysis. These methods include gas chromatograpy mass spectrometry (GC-MS) [19•,20•], high-performance liquid chromatography mass spectrometry (HPLC-MS) [12•,21,22] and, more recently, capillary electrophoresis mass spectrometry (CE-MS) [23]. Analytical methods without pre-separation by chromatography, for example, Fourier-transform ion cyclotron resonance (FT-ICR) mass spectrometry and time-of-flight (TOF) mass spectrometry, are also used, although largely for fingerprinting purposes [24,25]. Nuclear magnetic resonance (NMR) spectroscopy analysis of the whole-cell extract and on-line NMR analysis coupled with a liquid chromatography stop flow method are also available [26–29].

Data analysis forms the second major component of metabolomics. Chemometrics and multivariate analysis, such as a principal component analysis, hierarchical cluster analysis, and self-organization mapping, are often used for data mining [14•,30–32,33••,34]. Integration of metabolome data with other omics data, such as transcriptomic data, can be performed *in silico* using a software tool to depict gene-protein-metabolite networks [35]. Cell simulation can also be carried out to predict the metabolic phenotype of a particular gene knockout [36].

Metabolomics-based functional genomics in model plants

An advantage of working with *A. thaliana*, a model plant for modern plant science, is that genome-related

resources including the whole-genome sequence with functional annotations, DNA microarrays, DNA-tagged insertion mutants and metabolic maps (AraCyc; http://www.arabidopsis.org/tools/aracyc) are readily accessible [37].

GC-MS was used to compare the metabolic profiles of four *Arabidopsis* genotypes (two ecotypes, Col-2 and C24, and a mutant of each ecotype) and indicated that the metabolic phenotypes of the two ecotypes diverged more from each other than did those of each mutant from its parent ecoptype, suggesting that the cell metabolome is influenced more predominantly by the difference in ecotype than by a single gene mutation [19•].

Pair-wise metabolite-to-metabolite [38••] and transcript-to-metabolite [39••] correlation analysis of transgenic potato plants expressing particular genes revealed novel correlations that had not been suggested by classical targeted approaches. Most of these transgenic plants exhibited no visible phenotypes. *Arabidopsis pal1* and *pal2* mutants lacking the function of two phenylalanine ammonia lyase genes also showed no clear phenotypic alterations. Phenotyping of the single mutants and the double-mutant by a combination of transcript profiling and detailed targeted metabolite profiling for sugars, amino acids and phenylpropanoids suggested the specific function of *PAL1* and *PAL2* [40•].

Integrated analyses of the transcriptome and a detailed chemical analysis employing LC-MS and FT-ICR MS were carried out for *Arabidopsis* lines overexpressing the *PAP1* gene, which codes for a Myb-like transcription factor [41••]. The changes in the metabolic profiles caused by *PAP1* gene expression were specific to the increased accumulation of anthocyanins. The expression of genes known to be involved in anthocyanin production was upregulated, and thus other upregulated genes could be tentatively assigned a role in the formation of anthocyanins, such as members of the glycosyltransferase, acyltransferase and glutathione *S*-transferase families. The function of some of these predicted genes was confirmed experimentally through classical analysis of T-DNA-inserted knockout mutant lines and *in vitro* enzymatic studies with recombinant proteins [42]. These approaches indicate the feasibility of integrating metabolomic and transcriptomic analyses in functional genomics studies of *Arabidopsis* [41••].

Nutritional and abiotic stresses modulate re-programming of the transcriptome and metabolome. Thus, an integrated analysis leads to the identification of gene functions that are modulated by these stresses. A good example of this was obtained from a study of sulfur starvation in *Arabidopsis* [33••]. Glucosinolate-related metabolites and genes involved in their metabolism were coordinately regulated by nutritional sulfur stress. Further detailed investigation resulted in the prediction and identification of particular gene functions involved in glucosinolate metabolism (MY Hirai and K Saito, unpublished). Through the investigation of stress responses to nitrogen [43] and low temperature [44], wide investigation of gene expression and metabolite levels, respectively, have revealed novel pathways responding to each stress.

Attempts to functionally identify genes involved in saponin biosynthesis by *Medicago truncatula*, a model Leguminosae plant, have also been made using a combination of genomics and metabolic profiling [45,46]. These studies demonstrated the differential transcriptional re-programming of multiple genes in the phenylpropanoid and triterpene pathways in response to methyl jasmonate and yeast extracts (see Update).

Gene and metabolite expression profiles in non-model plants

For decades, most beneficial compounds identified and isolated from plants were discovered through classical, complicated extraction systems combined with targeted chemical analysis. Secondary metabolites have been identified as a particularly rich source of useful compounds and have a crucial role in this regard. However, only a fragment of all higher plants has been investigated [47]. There is no doubt that the chemical diversity of plants is much greater than any chemical library made by man, and thus the plant kingdom represents an enormous reservoir of novel molecules waiting to be discovered. Efficient metabolome analysis is likely to be a key element for future success in this field.

Elicitation is a process used to mimic the natural reactions of plants towards various environmental stresses and has been successfully applied to plants and plant cell cultures to induce secondary metabolite production [48]. It has been shown that upon elicitation huge numbers of genes are activated, many of them involved in metabolite biosynthesis. Goossens and coworkers [49••] treated tobacco BY-2 cells with methyljasmonate and discovered by using cDNA-AFLP (amplified fragment length polymorphism) transcript profiling that close to 600 genes were differentially regulated by this elicitor. By linking this data to metabolite analysis in a time-course experiment, it was possible to build an ample inventory of genes that were already known or novel genes. Potentially, these genes could be involved — in either a structural or regulatory capacity — with tobacco secondary metabolism, and possibly with plant secondary metabolism in general. The great advantage of this technology is that it allows gene identification without prior sequence knowledge. This is crucial when working with non-model plants or rare (e.g. medicinal) plant species for which the genome is unknown. Furthermore, this approach provides quantitative information on gene expression. Applying transcript

profiling in *Catharanthus roseus*, for example, the number of jasmonate-modulated genes was found to be of similar magnitude to that in tobacco cells. However, according to our preliminary metabolome experiments, more than 2000 metabolites were seen to respond to methyljasmonate treatment in *C. roseus* (H Rischer, M Oresic and KM Oksman-Caldentey, unpublished). Another huge challenge is to identify the function of the candidate genes obtained from this kind of transcript profiling [49••] and to find out their possible role in secondary metabolism. Again, high-throughput metabolite analysis will play a crucial role.

Some genes involved in the formation of volatile compounds in strawberry [50], rose [51•] and spider mint [52•] have been identified using DNA microarrays in combination with targeted analysis of volatile metabolites. The chemical analysis of fragrance-related metabolites by GC-MS is possible and leads to a sensitive chemical analysis which, together with gene expression profiles, is sufficient for gene identification. Activation-tagged lines, in which a gene is over-expressed by random insertion of an enhancer sequence in the genome, are good resources for gene hunting. By screening activation-tagged lines in tomato, a Myb transcription factor gene and co-regulated structural genes involved in anthocyanin formation were successfully identified [53].

The classical biochemical approach has resulted in considerable knowledge of the genes involved in the synthesis of flavonoids [54] and terpenoid indole alkaloids [55]. However, genetic maps of biosynthetic pathways, in general, are still far from complete and the regulation of these pathways is not fully understood. This point is well illustrated by a recent investigation of morphine biosynthesis in the opium poppy. Silencing codeinone reductase using a chimeric hairpin RNA construct led, through a feed-back mechanism, to the accumulation of (*S*)-reticuline, an intermediate of morphine biosynthesis; however, between (*S*)-reticuline and codeinone there are eight enzymatic steps and, surprisingly, all were inhibited by this RNA construct [56].

Conclusions

Tremendous advances in metabolomics and its integration into other omics (genomics, transcriptomics and proteomics) have brought us closer to understanding the links between different levels of biological systems, leading to the realization of systems biology [57]. This has been made possible by combining expertise from the areas of biology, chemistry, instrumentation and bioinformatics. In particular, *in silico* metabolic experiments could facilitate the network analysis of complicated metabolic pathways [58•,59]. Considering the huge chemical diversity of plants compared with those of animals and microorganisms, a major future challenge will be to explore the molecular genetic origins of chemical diversity in non-model exotic plants. A further prospect beyond conventional omics would be the comprehensive analysis of the function and activity of an array of plant metabolites, leading to the 'phytochemical array' concept (Figure 1). The phytochemical array consists of genomics, transcriptomics, proteomics, metabolomics and activity arrays of a given plant species. Such an array would allow visualization of all the connections between genes, transcripts, proteins, metabolites and their activities. Drawing links from the genome to the activity of metabolites will be necessary for the high-throughput discovery of plant-based pharmaceuticals and for the development of functional foods and stress-resistant plants. Although there are still several limitations in metabolomics, such as the urgent need for precise spacio-resolution (e.g. single cell and subcellular analysis) and temporal-resolution and the issue of metabolic channeling by protein complexes, we believe that many exciting developments are to be expected in the coming years.

Update

Recent work has demonstrated the differential transcriptional re-programming of multiple genes in the phenylpropanoid and triterpene pathways in response to methyl jasmonate and yeast extracts in *M. truncatula* [60].

Acknowledgements

Tekes, The National Technology Agency, Finland is kindly acknowledged for financial support from the NeoBio program to KMOC. KS thanks the Grants-in-Aid for Scientific Research from the Ministry of Education, Culture, Sports, Science and Technology, Japan and CREST of Japan Science and Technology Agency (JST) for financial support. Finally, the authors express their warm thanks to John Londesborough (VTT Biotechnology) for kindly revising the manuscript.

References and recommended reading

Papers of particular interest, published within the annual period of review, have been highlighted as:

- • of special interest
- •• of outstanding interest

1. •• Oksman-Caldentey KM, Inzé D: **Plant cell factories in the post-genomic era: new ways to produce designer secondary metabolites**. *Trends Plant Sci* 2004, **9**:433-440.
A concise and well-written up-to-date review on strategies and future trends for engineering plant secondary products, including a few attractive examples. New possibilities to use plant-cell cultures as a production host are presented.

2. Cyranoski D: **Campaign to fight malaria hit by surge in demand for medicine**. *Nature* 2004, **432**:259.

3. Ye X, Al-Babili S, Kloti A, Zhang J, Lucca P, Beyer P, Potrykus I: **Engineering the provitamin A (β-carotene) biosynthetic pathway into (carotenoid-free) rice endosperm**. *Science* 2000, **287**:303-305.

4. Niggeweg R, Michael AJ, Martin C: **Engineering plants with increased levels of the antioxidant chlorogenic acid**. *Nat Biotechnol* 2004, **22**:746-754.

5. Oksman-Caldentey K-M, Inzé D, Orešič M: **Connecting genes to metabolites by a systems biology approach**. *Proc Natl Acad Sci USA* 2004, **101**:9949-9950.

6. Tweeddale H, Notley-McRobb L, Ferenci T: **Effect of slow growth on metabolism of *Escherichia coli*, as revealed by global metabolite pool ('Metabolome') analysis**. *J Bacteriol* 1998, **180**:5109-5116.

7. Trethewey RN, Krotzky AJ, Willmitzer L: **Metabolic profiling: a Rosetta stone for genomics?** *Curr Opin Plant Biol* 1999, **2**:83-85.

8. Harrigan GG, Goodacre R (eds): *Metabolic Profiling: its Role in Biomarker Discovery and Gene Function Analysis*. Kluwer Academic Publishers: Boston; 2003.

9. Dixon RA, Strack D: **Phytochemistry meets genome analysis, and beyond**. . .. *Phytochemistry* 2003, **62**:815-816.

10. Trethewey R: **Metabolite profiling as an aid to metabolic engineering in plants**. *Curr Opin Plant Biol* 2004, **7**:196-201.

11. • Bino RJ, Hall RD, Fiehn O, Kopka J, Saito K, Draper J, Nikolau BJ, Mendes P, Roessner-Tunali U, Beale MH *et al.*: **Potential of metabolomics as a functional genomics tool**. *Trends Plant Sci* 2004, **9**:418-425.
This article proposes a framework for the study of plant metabolomics presented by the international plant metabolomics platform. Several community-based efforts are proposed for the development of plant metabolomics.

12. • Von Roepenack-Lahaye E, Degenkolb T, Zerjeski M, Franz M, Roth U, Wessjohann L, Schmidt J, Scheel D, Clemens S: **Profiling of *Arabidopsis* secondary metabolites by capillary liquid chromatography coupled to electrospray ionization quadrupole time-of-flight mass spectrometry**. *Plant Physiol* 2004, **134**:548-559.
This report describes the chemical analysis of *Arabidopsis* metabolites using HPLC coupled with a quadrupole time-of-flight mass spectrometer. Using this method, the authors detected 2000 metabolites in leaves and roots of *A. thaliana*.

13. Facchini PJ, Bird DA, St-Pierre B: **Can *Arabidopsis* make complex alkaloids?** *Trends Plant Sci* 2004, **9**:116-122.

14. • Fiehn O: **Metabolomics – the link between genotypes and phenotypes**. *Plant Mol Biol* 2002, **48**:155-171.
The first comprehensive review covering all aspects of plant metabolomics, definition, analytical methodology and data analysis.

15. Sumner LW, Mendes P, Dixon RA: **Plant metabolomics: large-scale phytochemistry in the functional genomics era**. *Phytochemistry* 2003, **62**:817-836.

16. Weckwerth W: **Metabolomics in systems biology**. *Annu Rev Plant Biol* 2003, **54**:669-689.

17. Kopka J, Fernie A, Weckwerth W, Gibon Y, Stitt M: **Metabolite profiling in plant biology: platforms and destinations**. *Genome Biol* 2004, **5**:109.

18. Stitt M, Fernie AR: **From measurements of metabolites to metabolomics: an 'on the fly' perspective illustrated by recent studies of carbon–nitrogen interactions**. *Curr Opin Biotechnol* 2003, **14**:136-144.

19. • Fiehn O, Kopka J, Doermann P, Altman T, Trethewey RN, Willmitzer L: **Metabolite profiling for plant functional genomics**. *Nat Biotechnol* 2000, **18**:1157-1161.
Together with Rossener *et al*. [20•], this is one of the first reports to show how comprehensive metabolic profiling using high-throughput GC-MS technology could help us to understand the systematic response of plants to genetic and environmental alterations.

20. • Rossener U, Luedemann A, Brust D, Fiehn O, Linke T, Willmitzer L, Fernie AR: **Metabolic profiling allows comprehensive phenotyping of genetically or environmentally modified plant systems**. *Plant Cell* 2001, **13**:11-29.
See the annotation for [19•].

21. Yamazaki M, Nakajima J, Yamanashi M, Sugiyama M, Makita Y, Springob K, Awazuhara M, Saito K: **Metabolomics and differential gene expression in anthocyanin chemo-varietal froms of *Perilla frutescens***. *Phytochemistry* 2003, **62**:987-995.

22. Yamazaki Y, Urano A, Sudo H, Kitajima M, Takayama H, Yamazaki M, Aimi N, Saito K: **Metabolite profiling of alkaloids and strictosidine synthase activity in camptothecin producing plants**. *Phytochemistry* 2003, **62**:461-470.

23. Sato S, Soga T, Nishioka T, Tomita M: **Simultaneous determination of the main metabolites in rice leaves using capillary electrophoresis mass spectrometry and capillary electrophoresis diode array detection**. *Plant J* 2004, **40**:151-163.

24. Aharoni A, de Vos CH, Verhoeven HA, Mailiepaard CA, Kruppa G, Bino RJ, Goodenowe DB: **Nontargeted metabolome analysis by use of Fourier transform ion cyclotron mass spectrometry**. *OMICS* 2002, **6**:217-234.

25. Brown SC, Kruppa G, Dasseux JL: **Metabolomics applications of FT-ICR mass spectrometry**. *Mass Spectrom Rev* 2005, **24**:223-231.

26. Griffin JL: **Metabonomics: NMR spectroscopy and pattern recognition analysis of body fluids and tissues for characterisation of xenobiotic toxicity and disease diagnosis**. *Curr Opin Chem Biol* 2003, **7**:648-654.

27. Wolfender JL, Ndjoko K, Hostettmann K: **Liquid chromatography with ultraviolet absorbance-mass spectrometric detection and with nuclear magnetic resonance spectroscopy: a powerful combination for the on-line structural investigation of plant metabolites**. *J Chromatogr A* 2003, **1000**:437-455.

28. Ward JL, Harris C, Lewis J, Beale MH: **Assessment of ^{1}H NMR spectroscopy and multivariate analysis as a technique for metabolite fingerprinting of *Arabidopsis thaliana***. *Phytochemistry* 2003, **62**:949-957.

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

30. Kose F, Weckwerth W, Linke T, Fiehn O: **Visualizing plant metabolomics correlation networks using clique-metabolite matrices**. *Bioinformatics* 2001, **17**:1198-1208.

31. Duran AL, Yang J, Wang L, Sumner LW: **Metabolomics spectral formatting, alignment and conversion tools (MSFACTs)**. *Bioinformatics* 2003, **19**:2283-2293.

32. Jonsson P, Gullberg J, Nordström A, Kusano M, Kowalczyk M, Sjöström M, Moritz T: **A strategy for identifying differences in large series of metabolomics samples analyzed GC/MS**. *Anal Chem* 2004, **76**:1738-1745.

33. •• Hirai MY, Yano M, Goodenowe D, Kanaya S, Kimura T, Awazuhara M, Arita M, Fujiwara T, Saito K: **Integration of transcriptomics and metabolomics for understanding of global responses to nutritional stresses in *Arabidopsis thaliana***. *Proc Natl Acad Sci USA* 2004, **101**:10205-10210.
In this excellent article, the authors combined DNA array analysis with targeted and non-targeted chemical analysis for holistic investigation under conditions of nutritional stress in *Arabidopsis*. They found new networks of pathways, genes and metabolites in primary and secondary metabolism, enabled by the use of batch-learning self-organizing map analysis.

34. Jenkins H, Hardy N, Beckmann M, Draper J, Smith AR, Taylor J, Fiehn O, Goodacre R, Bino R, Hall R *et al.*: **A proposed framework for the description of plant metabolomics experiments and their results**. *Nat Biotechnol* 2004, **2**:1601-1606.

35. Thimm O, Bläsing O, Gibon Y, Nagel A, Meyer S, Krüger P, Selbig J, Müller LA, Rhee SY, Stitt M: **MAPMAN: a user-driven tool to display genomics data sets onto diagrams of metabolic pathways and other biological processes**. *Plant J* 2004, **37**:914-939.

36. Tomita M, Hashimoto K, Takahashi K, Shimizu TS, Matsuzaki Y, Miyoshi F, Saito K, Tanida S, Yugi K, Venter JC: **E-CELL: Software environment for whole-cell simulation**. *Bioinformatics* 1999, **15**:72-84.

37. Mueller LA, Zhang P, Rhee SY: **AraCyc. A biochemical pathway database for *Arabidopsis***. *Plant Physiol* 2003, **132**:453-460.

38. •• Weckwerth W, Loureiro ME, Wenzel K, Fiehn O: **Differential metabolic networks unravel the effects of silent plant phenotypes**. *Proc Natl Acad Sci USA* 2004, **101**:7809-7814.
Non-targeted metabolite analysis by GC-TOF MS was applied to the phenotyping of silent (i.e. no visible changes) transgenic potato plants suppressing a sucrose synthase gene. Metabolic networks were constructed to reveal alterations even if no apparent difference was seen in metabolite levels.

39. •• Urbanczyk-Wochniak E, Luedemann A, Kopka J, Selbig J, Roessner-Tunali U, Willmitzer L, Fernie AR: **Parallel analysis of transcript and metabolic profiles: a new approach in systems biology**. *EMBO Rep* 2003, **4**:989-993.

Correlations of gene expression and metabolite accumulation were made by the pair-wise analysis of transcript and metabolite levels. From 27 000 possible pairs, 571 showed significant correlation.

40. • Rohde A, Morreel K, Ralph J, Goeminne G, Hostyn V, De Rycke R, Kushnir S, Van Doorsselaere J, Joseleau JP, Vuylsteke M *et al.*: **Molecular phenotyping of the *pal1* and *pal2* mutants of *Arabidopsis thaliana* reveals far-reaching consequences on phenylpropanoid, amino acid, and carbohydrate metabolism**. *Plant Cell* 2004, **16**:2749-2771.

Extensive analysis of the metabolites and transcripts from *pal1* and *pal2* single- and double-mutants of *Arabidopsis* enabled the distinct functions of the *PAL1* and *PAL2* genes to be elucidated.

41. •• Tohge T, Nishiyama Y, Hirai MY, Yano M, Nakajima J, Awazuhara M, Inoue E, Takahashi H, Goodenowe DB, Kitayama M *et al*: **Functional genomics by integrated analysis of metabolome and transcriptome of *Arabidopsis* plants overexpressing a MYB transcription factor**. *Plant J* 2005, in press.

Detailed analysis of flavonoids together with non-targeted chemical analysis were used to characterize the metabolome of several *Arabidopsis* lines engineered to overexpress the transcription factor *PAP1*. Some of the genes upregulated in the *PAP1* overexpression lines were found to encode proteins involved in the biosynthesis of anthocyanins. This is a good example of how the combined efforts of transcriptome and metabolome analyses can lead to the identification of gene function.

42. Jones P, Messner B, Nakajima J, Schäffner AR, Saito K: **UGT73C6 and UGT78D1, glycosyltransferases involved in flavonol glycoside biosynthesis in *Arabidopsis thaliana***. *J Biol Chem* 2003, **278**:43910-43918.

43. Scheible WR, Morcuende R, Czechowski T, Fritz C, Osuna D, Palacios-Rojas N, Schindelasch D, Thimm O, Udvardi MK, Stitt M: **Genome-wide reprogramming of primary and secondary metabolism, protein synthesis, cellular growth processes, and the regulatory infrastructure of *Arabidopsis* in response to nitrogen**. *Plant Physiol* 2004, **136**:2483-2499.

44. Cook D, Fowler S, Fiehn O, Thomashow MF: **A prominent role for the CBF cold response pathway in configuring the low-temperature metabolome of *Arabidopsis***. *Proc Natl Acad Sci USA* 2004, **101**:15243-15248.

45. Suzuki H, Achnine L, Xu R, Matsuda SPT, Dixon RA: **A genomics approach to the early stages of triterpene saponin biosynthesis in *Medicago truncatula***. *Plant J* 2002, **32**:1033-1048.

46. Huhman DV, Sumner LW: **Metabolic profiling of saponins in *Medicago sativa* and *Medicago truncatula* using HPLC coupled to an electrospray ion-trap mass spectrometer**. *Phytochemistry* 2002, **59**:347-360.

47. Hostettmann K, Terreaux C: **Search for new lead compounds from higher plants**. *Chimia (Aarau)* 2000, **54**:652-657.

48. Dixon RA: **Natural products and plant disease resistance**. *Nature* 2001, **411**:843-847.

49. •• Goossens A, Häkkinen ST, Laakso I, Seppänen-Laakso T, Biondi S, De Sutter V, Lammertyn F, Nuutila AM, Soderlund H, Zabeau M *et al.*: **A functional genomics approach toward the understanding of secondary metabolism in plant cells**. *Proc Natl Acad Sci USA* 2003, **100**:8595-8600.

A cDNA-AFLP technology was applied to the high-throughput transcript profiling of methyljasmonate-treated tobacco BY-2 cells. The correlation of transcripts with the accumulation patterns of secondary metabolites allowed the identification of putative genes responsible for the inducible accumulation of secondary products. The method is generally applicable for any non-model or exotic plants for which genome information is unavailable. Furthermore, this approach provides quantitative data on gene expression.

50. Aharoni A, Keizer LCP, Bouwmeester HJ, Sun Z, Alvarez-Huerta M, Verhoven HA, Blaas J, van Houwelingen AMML, De Vos RCH, van der Voet H *et al.*: **Identification of the *SAAT* gene involved in strawberry flavor biogenesis by use of DNA microarrays**. *Plant Cell* 2000, **12**:647-661.

51. • Guterman I, Shalit M, Menda N, Piestun D, Dafy-Yelin M, Shalev G, Bar E, Davydov O, Ovadis M, Emanuel M *et al.*: **Rose scent: genomics approach to discovering novel floral fragrance-related genes**. *Plant Cell* 2002, **14**:2325-2338.

DNA chips were prepared from the expressed sequence tag information of rose and used for the expression profiling of transcripts in rose petal from different cultivars. Combined with a detailed chemical analysis of volatile compounds this work led to the discovery of novel flower scent-related genes.

52. • Mercke P, Kappers IF, Verstappen FWA, Vorst O, Dicke M, Bouwmeester HJ: **Combined transcript and metabolite analysis reveals genes involved in spider mint induced volatile formation in cucumber plants**. *Plant Physiol* 2004, **135**:2012-2024.

Using a similar approach to that described by Guterman *et al.* [51•], DNA microarray analysis was combined with the chemical analysis of volatile metabolites in cucumber. Insight was gained into the involvement of the studied genes in the production of volatile compounds.

53. Mathews H, Clendennen SK, Caldwell CG, Liu XL, Connors K, Matheis N, Schuster DK, Menasco DJ, Wagoner W, Lightner J, Wagner DR: **Activation tagging in tomato identifies a transcriptional regulator of anthocyanin biosynthesis, modification, and transport**. *Plant Cell* 2003, **15**:1689-1703.

54. Springob K, Nakajima J, Yamazaki M, Saito K: **Recent advances in the biosynthesis of and accumulation of anthocyanins**. *Nat Prod Rep* 2003, **20**:288-303.

55. Verpoorte R, Memelink J: **Engineering secondary metabolite production in plants**. *Curr Opin Biotechnol* 2002, **13**:181-187.

56. Allen RS, Millgate AG, Chitty JA, Thisleton J, Miller JAC, Fist AJ, Gerlach WL, Larkin PJ: **RNAi-mediated replacement of morphine with the nonnarcotic alkaloid reticuline in opium poppy**. *Nat Biotechnol* 2004, **22**:1559-1566.

57. Fernie AR, Trethewey RN, Krotzky AJ, Willmitzer L: **Metabolite profiling: from diagnostics to systems biology**. *Nat Rev Mol Cell Biol* 2004, **5**:763-769.

58. • Arita M: ***In silico* atomic tracing by substrate-product relationships in *Escherichia coli* intermediary metabolism**. *Genome Res* 2003, **13**:2455-2466.

A report on a specific software that allows *in silico* tracer experiments to be carried out at the atomic level for all possible pathways.

59. Yamazaki Y, Kitajima M, Arita M, Takayama H, Sudo H, Yamazaki M, Aimi N, Saito K: **Biosynthesis of camptothecin. *In silico* and *in vivo* tracer study from [1-^{13}C]glucose**. *Plant Physiol* 2004, **134**:161-170.

60. Suzuki H, Srinivasa Reddy MS, Naoumkina M, Aziz N, May GD, Huhman DV, Sumner LW, Blount JW, Mendes P, Dixon RA: **Methyl jasmonate and yeast elicitor induce differential transcriptional and metabolic re-programming in cell suspension cultures of the model legume *Medicago truncatula***. *Planta* 2005. DOI: 10.1007/s00425-004-1387-2.