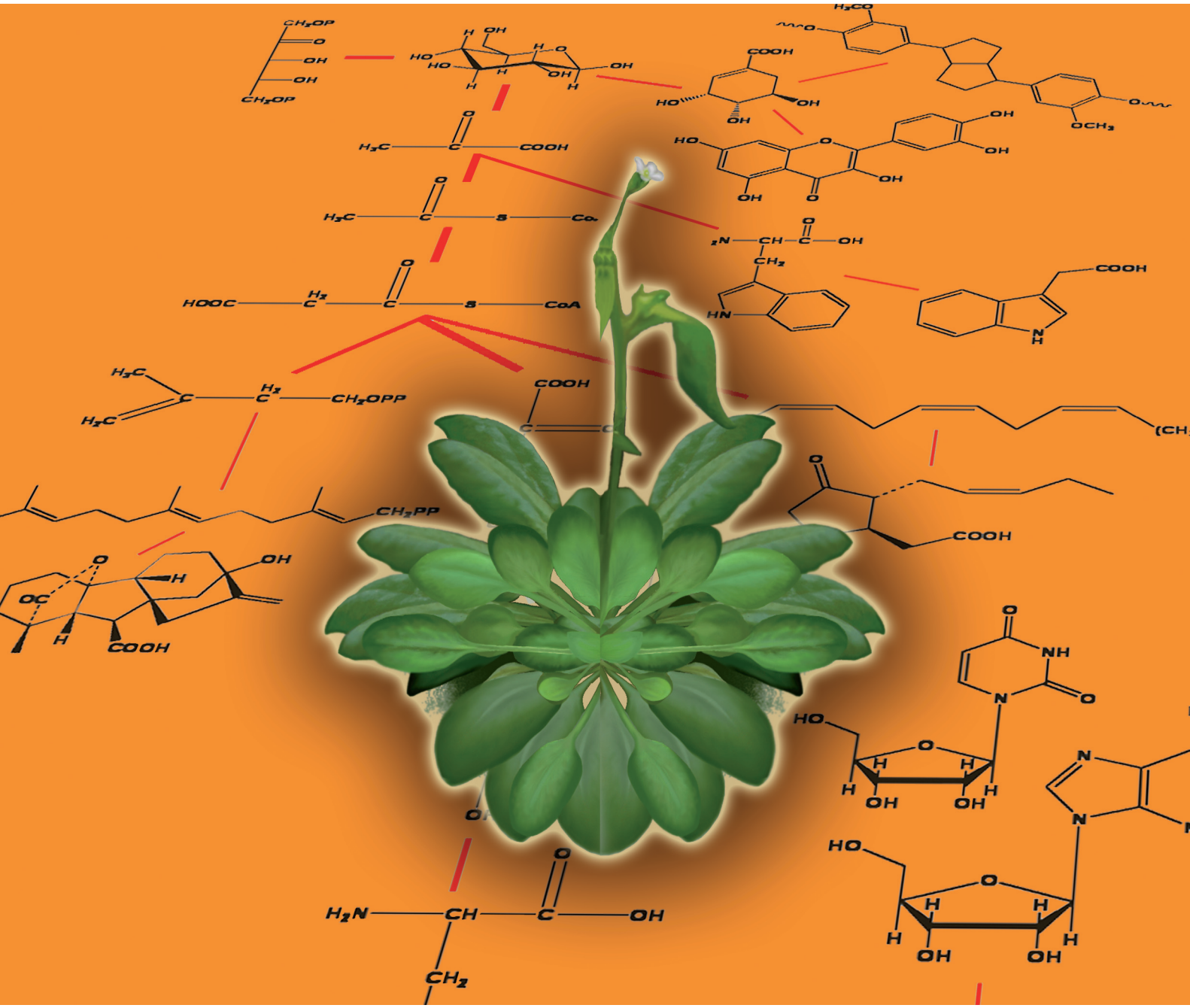
Molecular BioSystems

Indexed in MEDLINE!

www.molecularbiosystems.org

Volume 4 | Number 10 | October 2008 | Pages 957–1032



ISSN 1742-206X

HIGHLIGHT
Hirai and Saito
Analysis of systemic sulfur metabolism in plants using integrated '-omics' strategies

PAPER
Gipson *et al.*
Multi-platform investigation of the metabolome in a leptin receptor defective murine model of type 2 diabetes

RSC Publishing

1742-206X(2008)4:10;1-T

Analysis of systemic sulfur metabolism in plants using integrated '-omics' strategies

Masami Yokota Hirai[ab*] and Kazuki Saito[ac*]

DOI: 10.1039/b802911n

Sulfur (S) assimilation by plants plays a key role in the S cycle in nature, and metabolism of the assimilated S provides various compounds that are useful for animals, including humans. It is important to understand the mechanisms involved in systemic S metabolism in order to improve crop agronomy and the production of dietary and nutraceutical plants. Recent advances in '-omic' studies have provided comprehensive insights into S metabolism. These studies may be regarded as important case studies that provide information on the complicated regulatory mechanisms involved in plant metabolism.

S-containing metabolites produced in plants

S metabolism is essential for both plants and animals because it produces various S-containing organic compounds that play pivotal roles in many aspects of life.[1,2]

Inorganic sulfate in soil is absorbed into plant cells *via* sulfate transporters and is metabolized first into cysteine and subsequently into methionine, which is an essential amino acid in animals. Cysteine residues are indispensable in proteins due to their presence in the active site of many enzymes, and due to the extensive functions they perform *via* the formation of disulfide bonds. Further, cysteine residues are incorporated into oligopeptides such as glutathione (GSH), which plays an important role along with Fe–S clusters in various redox reactions. GSH also plays a role in the detoxification of xenobiotics. *S*-Adenosylmethionine (SAM) is essential as a methyl-group donor for DNA methylation, 5′-capping of mRNA, and almost all other methylation reactions occurring in cells. It also functions as a precursor of bioactive polyamines. Many reactions occurring in living cells require S-containing cofactors such as biotin, thiamine pyrophosphate and coenzyme A.

In plants, sulfolipids are essential for the formation of thylakoid membranes in chloroplasts, the photosynthetic organelles. Plants produce species-specific secondary metabolites containing S, such as glucosinolates (GSLs), camalexin (produced by cruciferous plants) and alliins (produced by *Allium* species), which act as phytoalexins. Glucosinolates and alliins are broken down by specific degradation enzymes into volatile compounds such as isothiocyanates (ITCs) and allicin, respectively, which are defence compounds that possess antimicrobial activities. Interestingly, these volatile compounds are also beneficial to humans because they impart flavour to vegetables such as cabbage, broccoli, mustard (family, Cruciferae), garlic and onion (*Allium* species). These compounds have recently been attracting attention for their health-promoting effects such as their anticancer activities.

Regulation of S metabolism in plants

As mentioned above, inorganic sulfate is processed into various plant metabolites (Fig. 1). Since each S-containing metabolite is essential to the plant, the

[a] *RIKEN Plant Science Center, Suehiro-cho 1-7-22, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan*
[b] *Core Research for Evolutional Science and Technology, Japan Science and Technology Agency, 4-1-8 Hon-chou, Kawaguchi, Saitama 332-0012, Japan*
[c] *Graduate School of Pharmaceutical Sciences, Chiba University, Yayoi-cho 1-33, Inage-ku, Chiba, Chiba 263-8522, Japan*

Masami Yokota

Masami Yokota Hirai began her -omics studies in investigations of sulfur metabolism carried out at Chiba University. Since April, 2005, she has been a Principal Investigator at RIKEN Plant Science Center. Her research interests include discovering regulatory mechanisms of amino acid biosynthesis by omics-based strategies as well as identifying gene functions involved in secondary metabolism in plants.

Kazuki Saito

Kazuki Saito has been a full professor at Chiba University since 1995. Since 2005, he has been a group director at RIKEN Plant Science Center. His research interests are metabolome-based functional genomics, biochemistry, molecular biology and biotechnology of primary and secondary metabolism in plants. In particular, he is engaged in biosynthetic studies of sulfur compounds, flavonoids and alkaloids.

adequate distribution of S atoms to each metabolite is important for plant survival. For such distribution, the availability of S in the environment is monitored and S metabolism is regulated at various levels. Biochemical, molecular biological and genetic studies conducted over recent years have revealed the following independent mechanisms for the regulation of S metabolism: induction of gene expression during S-deficiency; S availability-dependent formation of an enzyme complex involving serine acetyltransferase and OAS(thiol)-lyase (OAS, *O*-acetylserine) leading to modulation of catalytic activity; and feedback regulation of enzymes involved in S assimilation at both the protein and transcript levels.[1–5] These studies investigated in detail the individual molecular mechanisms governing S homeostasis. However, their findings were generally restricted to the main pathways involved in S assimilation. Investigation from a broader perspective is expected to clarify the details of the complete regulatory mechanisms.

Transcriptome analyses in S-deficiency

In the model plants *Arabidopsis thaliana* and rice, genomic sequencing was completed in early 2000.[6–10] Since then, plant science has progressed beyond genomics. In *Arabidopsis* approximately 27 000 genes and their translation products have been targets of -omic studies.

S-deficiency is a serious agricultural problem that results in a decrease in crop yield and quality, and the symptoms and responses of plants to S-deficiency have been extensively reported in studies on plant physiology. Therefore, early -omic studies on S metabolism have focused on transcriptomic changes occurring in S-starved *Arabidopsis* plants to provide insights into the mechanisms of adaptation to S-deficiency.[11–14] In these studies,[11–14] wild-type *Arabidopsis* plants were grown under conditions of continuous S-deficiency or were transferred from the S-sufficient control condition to the S-deficiency condition and *vice versa*. Leaves, roots or whole seedlings were analyzed using DNA arrays in order to determine the transcript profiles of 8000–9000 genes. The status of S in these plants varied depending on the sulfate concentration in the culture media, the growth stage of the plant and the duration for which the plant was maintained under the S-deficiency condition (summarized in ref. 15). Therefore, the genes exhibiting the most notable changes in their expression levels differed among experiments. However, by considering transcriptomic data along with metabolic maps, metabolic pathways, which generally respond to S-deficiency, could be elucidated. As expected, the results confirmed the induction of several genes involved in the metabolism of sulfate into cysteine.

The most intriguing discovery is the involvement of jasmonate, which is a plant hormone, in metabolic pathways that are functional under S-deficiency conditions.[11,13,14] These studies were the first to suggest the involvement of jasmonates in S-deficiency responses. The biosynthetic pathway of another plant hormone, namely, auxin, is closely related to the pathway of indole GSL biosynthesis from tryptophan. In S-deficiency, the biosynthesis of tryptophan, indole GSL and auxin is affected.[12–14] On the other hand, by performing a transcriptomic analysis of jasmonate-regulated metabolic pathways in *Arabidopsis* plants, Sasaki-Sekimoto *et al*. reported that jasmonates activate the metabolic pathways of ascorbate and GSH and the biosynthesis of indole GSL.[16] All these studies suggest the involvement of complex metabolic and signalling networks in S assimilation and in the biosynthesis of GSH, indole GSL, auxin and jasmonate.

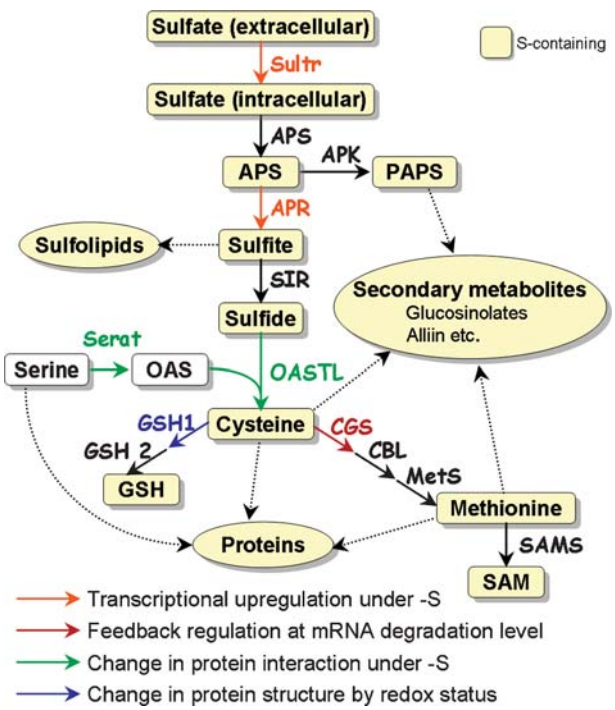


Fig. 1 S metabolism and regulatory mechanisms induced depending on S availability in *Arabidopsis*. Sultr, sulfate transporter; APS, ATP sulfurylase; APK, APS kinase; APR, adenosine 5′-phosphosulfate reductase; PAPS, 3′-phosphoadenosine 5′-phosphosulfate; SIR, sulfite reductase; Serat, serine acetyltransferase; OASTL, OAS(thiol)-lyase; CGS, cystathionine gamma-synthase; CBL, cystathionine beta-lyase; MetS, methionine synthase; SAMS, *S*-adenosylmethionine synthetase; GSH1, gamma-glutamylcysteine synthetase; GSH2, glutathione synthetase.

Integrated transcriptome and metabolome analysis in S-deficiency

The analytical technology of transcriptomics has developed considerably in recent years; for example, the 22-k *Arabidopsis* microarray has become commercially available. In contrast, the technology of metabolomics is still developing because the targets of this technology are chemically diverse metabolites, of which more than 200 000 exist in the plant

kingdom. Hence, no single method is sufficient for the simultaneous analysis of all existing metabolites.[17] In particular, the improvement of methods for identifying metabolites is an urgent requirement to enable the use of metabolomics as a tool for investigating the biology of systems. Nevertheless, metabolome analyses performed using current techniques can provide useful information on metabolic systems, particularly if these analyses are integrated with transcriptomics.[18,19] Nikiforova *et al.*[20] and Hirai *et al.*[12] analyzed the metabolome in the same plant samples that were used for the above-mentioned transcriptome analyses, and their studies were the first to integrate metabolomics and transcriptomics.

Hirai *et al.*[12] analyzed the metabolomes of S-starved, nitrogen (N)-starved, and OAS-treated *Arabidopsis* plants (see below) in a non-targeted manner by performing Fourier transform-ion cyclotron resonance (FT-ICR) mass spectrometry (MS) to detect approximately 2000 putative metabolites. The N-starved plants were expected to exhibit the reverse of the response exhibited by the S-starved plants because S-deficiency is considered to be a shortage of S in relation to N. OAS is considered to be a positive regulator of genes responsive to S-deficiency,[1] and its exogenous application induces responses that mimic several responses observed in S-deficiency. Therefore, the OAS-supplemented plants were expected to exhibit responses similar to those observed in S-deficiency. The treatments administered were divided into 2 categories, namely, short- and long-term treatments (summarized in Fig. 2A). The data obtained were subjected to a principal component analysis (PCA) to clarify the global changes occurring in the metabolome and transcriptome. As shown in Fig. 2B and C, the experiments (samples) were clustered according to the plant organ sampled (leaves or roots), method of plant culture (hydroponic culture or culture on an agar-solidified plate) and period of stress (long- or short-term stress). The results indicated the following features of global regulation: (1) long-term S-deficiency, N-deficiency, and S- and N-deficiency had similar effects on the metabolome and transcriptome; (2) the metabolite and transcript profiles differed considerably between long- and short-term S-deficiency; and (3) the effects of OAS treatment were similar to those of short-term S-deficiency, suggesting that OAS regulates the global metabolite and transcript profiles in short-term S-deficiency. The fact that similar clustering patterns were obtained by using the transcriptome and metabolome data indicates that the global transcript and metabolite profiles were strongly related to each other. Unexpectedly, similar changes were observed during S-deficiency and N-deficiency, suggesting the existence of general responses to nutritional deficiency in the global regulation of the metabolome and transcriptome. With regard to individual responses, however, general responses to both S-deficiency and N-deficiency as well as specific responses to either S-deficiency or N-deficiency were observed, for example, in the case of GSL metabolism.[12]

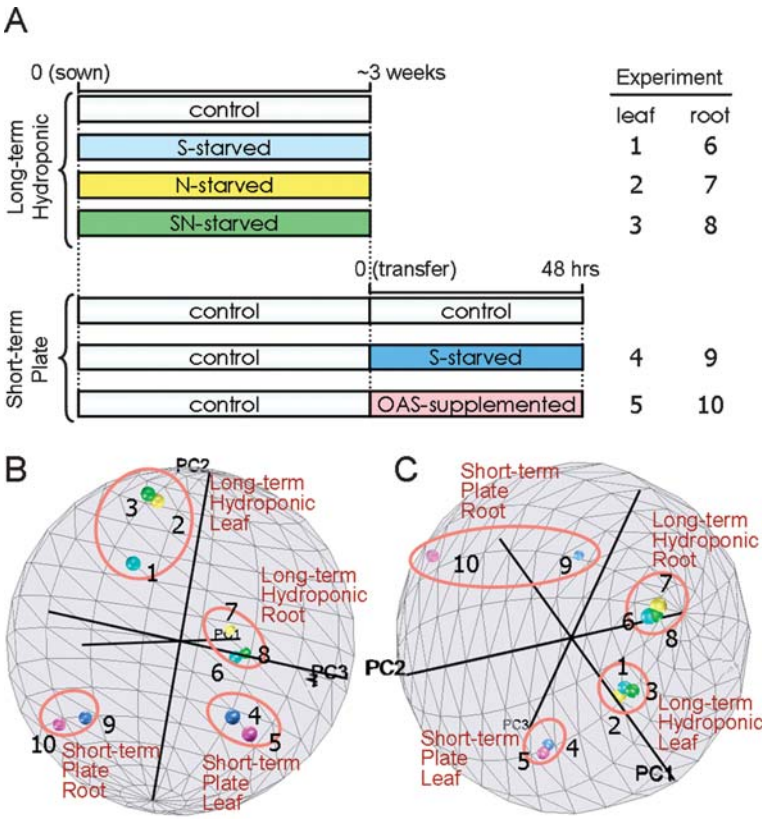


Fig. 2 Plant growth conditions and PCA performed to elucidate the changes induced in the metabolome and transcriptome by the culture conditions. (A) Plant growth conditions. (B) In order to classify the samples based on the global changes occurring in the metabolome, the log ratio of the accumulation level of each metabolite in S-deficiency to that in the corresponding control sample was calculated and was subjected to PCA. Each small globe represents a sample. The proportions of the first, second and third components are 27.6%, 18.8% and 15.8%, respectively. (C) In order to classify the samples based on the global changes occurring in the transcriptome, the log ratio of the expression level of each gene in S-deficiency to that in the corresponding control sample was calculated and was subjected to PCA. The proportions of the first, second and third components are 35.8%, 14.6% and 13.3%, respectively. The figures have been modified from ref. 12.

Nikiforova *et al.*[20] analyzed the metabolome of S-starved *Arabidopsis* plants by performing gas chromatography (GC)-MS and liquid chromatography (LC)-MS. These analyses revealed the response patterns of 6023 peaks of non-redundant ion traces and the relative concentration levels of 134 non-redundant compounds with known chemical structures. In order to understand the detailed response mechanisms, the metabolomic and transcriptomic data were considered in relation to known metabolic pathways. This revealed the interplay among S assimilation, N imbalance, lipid breakdown, purine metabolism and

enhanced photorespiration associated with S-deficiency stress.

Analysis of time-series data of the transcriptome and metabolome

Due to the complexity of the regulatory mechanisms functional at various levels, 'snapshot' analyses of the transcriptome and metabolome often merely describe the patterns of gene expression and metabolite accumulation. Time-series analyses of the transcriptome and metabolome provide deeper insights into gene–gene, metabolite–metabolite and gene–metabolite networks because time-series data are suitable for calculating the similarity distance between genes and metabolites. Time-series transcriptomic and metabolomic data obtained for S-starved *Arabidopsis* plants were combined into a single matrix and were analyzed using 2 different algorithms to elucidate the gene–metabolite networks.[21,22] Another study aimed to identify the transition time point during plant adaptation to S-deficiency.[23]

1 Batch-learning self-organizing mapping

Hirai *et al*.[21] applied batch-learning self-organizing mapping (BL-SOM)[24,25] to calculate the similarity distance between S-deficiency-responsive genes and metabolites. Based on the time-dependent changes in their accumulation/expression patterns in response to S-deficiency,[21] approximately 1000 putative metabolites and 10 000 genes were classified into clusters within single or adjacent cells on the 'feature map' obtained (Fig. 3A). For example, GSLs and their degradation products ITCs were clustered into their respective clusters on the feature map (Fig. 3A). The accumulation patterns of ITCs and GSLs were mirror images of each other (Fig. 3B), suggesting that GSL metabolism is coordinately regulated. In *Arabidopsis*, most of the genes involved in the synthesis of the GSL core structure have been identified. These genes known to be involved in GSL biosynthesis were clustered in the same region on the map (Fig. 3A), and this supports the notion that GSL metabolism is coordinately regulated (Fig. 3C). This analysis provides an effective method for predicting novel genes that may be involved in GSL biosynthesis (see below).

2 Network reconstruction based on the cause–effect relationship

Nikiforova *et al*.[22] used 2 methods to calculate the correlation distance (Pearson correlation coefficient and mutual information) for the combined time-series data of the transcriptome (6454 genes) and metabolome (81 chemical compounds) of S-starved *Arabidopsis* plants.[22] Assuming that an altered S level is the primary cause of systemic stimulation, they centralized S in the graph depicting the correlation network between S-deficiency-responsive genes and metabolites, and reconstructed the gene–metabolite network based on the cause–effect relationship.[22] By using this reconstructed network, the information flux from the primary cause (S-deficiency) to the response end-point (for example, anthocyanin accumulation) and the pathways involved in this flux (for example, the 'S-serine-tryptophan' pathway and hormone-related pathways) could be elucidated. Recently, Nikiforova and co-workers analyzed the topology of the metabolite–gene network in order to identify metabolites that regulate gene expression during adaptive responses to perturbation of the biological system.[26] By comparing 3 independent biological systems (S-starved *Arabidopsis*, herbicide-treated *Arabidopsis* and ripening tomato), they suggested that metabolites are less involved in gene expression in the S-deficiency stress response than in the ripening of tomatoes.

3 Detection of the transition time point

Morioka *et al*.[23] established a novel method based on a linear dynamical system model to predict the transition time point of the transcriptome and metabolome during the process of adaptation. The transition time point, at which the transcriptome and/or the metabolome change drastically in response to environmental changes, is determined based on the time-series data by using the logarithm of the 'likelihood value', which is the generative probability of current data based on the condition of previous data sets. If the likelihood value is low, then the current data cannot be adequately explained based on previous data sets; that is, a transition has occurred. Morioka *et al*.[23] applied this method to time-series transcriptomic and metabolomic data.[21] The results revealed that both the metabolome and transcriptome transitioned between 12 and 24 h after the plants were transferred to S-deficiency conditions.[23] Relevant factors such as genes and/or metabolites related to the transitions could then be identified using BL-SOM. For example, changes in lipid metabolism were observed around the transition time points.[23]

Functional identification of novel GSL biosynthesis genes

As mentioned above, known GSL biosynthesis genes were clustered on the feature map in the BL-SOM analysis (Fig. 3A). This suggested that unknown genes that clustered along with the known GSL biosynthesis genes may also be involved in GSL biosynthesis. Based on this assumption, Hirai *et al*.[21,27] comprehensively predicted the functions of these genes. The predicted functions of 3 sulfotransferase genes and an Myb transcription factor gene (*Myb28*) were subsequently confirmed by molecular biological and reverse genetic studies.[21,27] Further, the predicted functions of two other Myb genes (*Myb29* and *Myb76*) were recently confirmed;[28,29] this indicates the validity of this strategy for functional genomics.

Proteome analysis of S-starved seeds

Seed storage proteins (SSPs), which accumulate during seed maturation, are reservoirs of carbon (C), N and S for the growth of next-generation progenies during infancy in the absence of nutrient supply from the environment. SSPs comprise several proteins encoded by various gene families. The composition of SSPs varies depending on the availability of S in the environment; for maintenance of

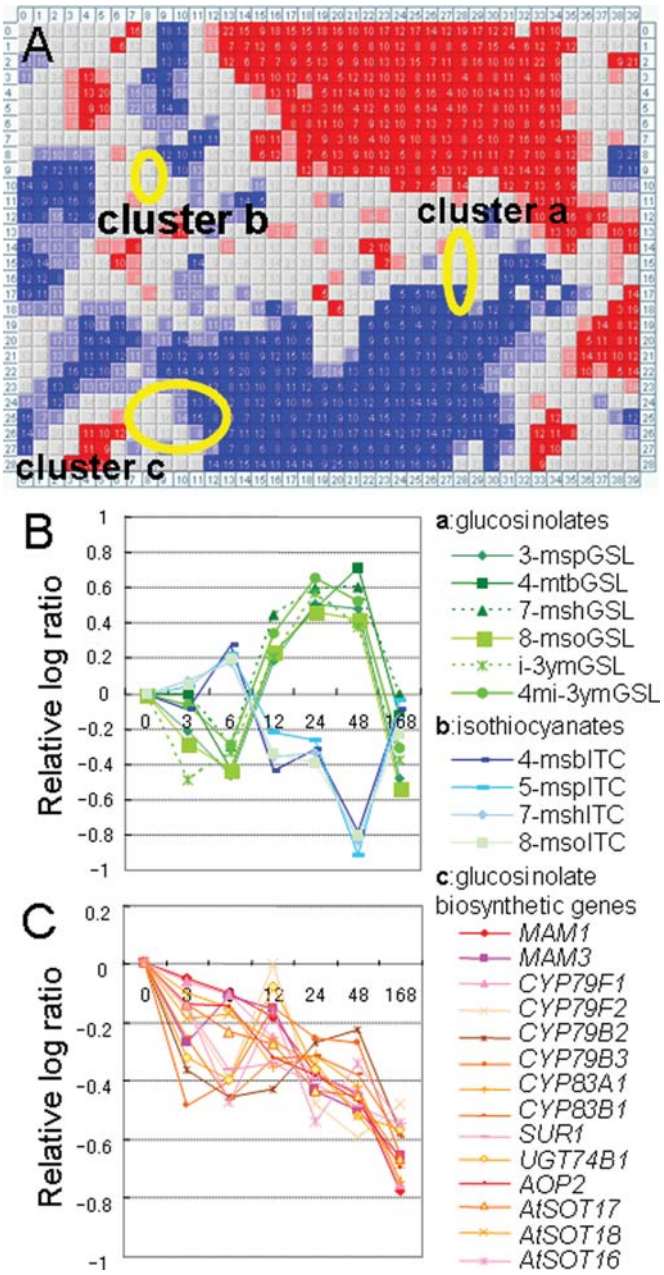


Fig. 3 BL-SOM of time-series metabolomic and transcriptomic data. The metabolome data obtained by performing FT-ICR-MS, high-pressure liquid chromatography and capillary electrophoresis were merged with the transcriptome data. Approximately 1000 metabolites and 10 000 genes responsive to S-deficiency were classified by BL-SOM based on the time-dependent changes in their patterns of accumulation and expression. (A) A feature map based on the data obtained for the leaf samples. Each cell was coloured based on the relative log ratio values of the metabolites and genes therein as follows: when all the relative log ratio values of the metabolites and genes in a particular cell were greater or smaller than the average, the cell was coloured in pink or pale blue, respectively. Red and blue cells indicated that at least one of the relative log ratio values was greater than the average plus standard deviation or smaller than the average minus standard deviation, respectively. Yellow circles indicate the clusters for GSLs (a), ITCs (b) and genes involved in GSL biosynthesis (c). (B) Changes in the content of GSLs and ITCs in the leaves. (C) Changes in the expressions of genes involved in GSL biosynthesis. Abbreviations indicate the side-chain groups of the GSLs and ITCs as follows: 3-msp, 3-methylsulfinylpropyl; 4-mtb, 4-methylthiolbutyl; 7-msh, 7-methylsulfinylheptyl; 8-mso, 8-methylsulfinyloctyl; i-3ym, indol-3-ylmethyl; 4-mi-3-ym, 4-methoxyindol-3-ylmethyl; 4-msb, 4-methylsulfinylbutyl; and 5-msp, 5-methysulfinylpentyl. The figures have been modified from ref. 21.

the total protein content, proteins having a low S content increase in abundance during S-deficiency, while those having a high S content decrease in abundance. In order to elucidate the changes occurring in the composition of SSPs during S-deficiency in detail, Higashi *et al.*[30] conducted proteome analysis on *Arabidopsis* seeds. In addition to the expected peptides that included all the amino acid residues encoded by the genes, many truncated versions were observed under both the S-sufficient control condition and the S-deficiency condition. The C- and N-terminal residues of all the observed SSPs were determined by performing extensive MS analysis, and the results indicated successive C-terminal degradation of 12S globulins from particular residues.[30] In addition to the expected hyperaccumulation of proteins having a low S content, the accumulation profiles of these truncated peptides were altered under the S-deficiency condition. Integration of this analysis with transcriptome analysis suggested that multi-level regulation of SSP gene expression, particularly post-translational modifications, enables the plant to respond efficiently to S-deficiency.

The content of S-containing amino acids in SSPs determines the quality of seeds used for feed and food. In order to improve the S content in seeds, a transgene encoding an S-rich protein under the control of a seed-specific promoter was introduced into legumes, canola and cereal plants (ref. 31–34 and the references therein). In the case of transgenic lupins and chickpeas, the methionine content in seeds increased concomitantly with the accumulation of foreign protein.[31,33] On the other hand, transgenic rice grains exhibited little change in the total S amino acid content regardless of foreign S-rich protein accumulation.[32,34] In all cases, a change in the composition of endogenous SSPs, but not in total S content, was observed. This change was assumed to be mediated by a signal transduction pathway that normally modulates SSP composition in response to environmental fluctuations in S availability, *via* both transcriptional and post-transcriptional control of gene expression.[32] A proteomics approach enabled detailed analysis of the change in SSP composition in rice.[34] These transgenic studies can be considered as good model cases of integrated -omics

designed to reveal the complicated mechanism of S homeostasis, involving sink/source relationships, in seeds.

Future perspectives and conclusion

In plant science, pioneering studies involving the integration of metabolomics and transcriptomics have been conducted in an attempt to understand the mechanisms of systemic S metabolism and regulation. Future improvements in metabolomics technology and metabolite databases will enable us to identify more metabolites detected using techniques such as mass spectroscopy and nuclear magnetic resonance, and hence to understand metabolism more clearly by considering metabolomic data along with metabolic maps. In addition, enhancement of public metabolome databases will contribute to a systematic and comprehensive understanding of metabolism as public transcriptome databases have already done (ref. 19 and the references therein).

The strategies introduced in this report may be applicable to other studies not only on sulfur metabolism in plants but also on other pathways in various organisms.[19] Since metabolism is a basic process occurring in life, plants have evolved a robust metabolic system to protect themselves against environmental and internal perturbations. Obtaining a global understanding of systemic metabolism is essential for effective manipulation of plant production.[18] Integrative studies on S metabolism will make it possible to elucidate the regulatory mechanisms involved in the conversion of sulfate to primary and secondary S-containing metabolites in response to environmental changes; this will enable the production of plants having valuable functions, such as cancer-preventing vegetables possessing a high GSL content and highly nutritious grains with higher S amino acid contents.

References

1 K. Saito, *Plant Physiol*., 2004, **136**, 2443–2450.

2 V. Nikiforova, B. Gakière, S. Kempa, M. Adamik, L. Willmitzer, H. Hesse and R. Hoefgen, *J. Exp. Bot*., 2004, **55**, 1861–1870.

3 L. M. Hicks, R. E. Cahoon, E. R. Bonner, R. S. Rivard, J. Sheffield and J. M. Jez, *Plant Cell*, 2007, **19**, 2653–2661.

4 T. Rausch, R. Gromes, V. Liedschulte, I. Müller, J. Bogs, V. Galovic and A. Wachter, *Plant Biol. (Stuttgart, Ger.)*, 2007, **9**, 565–572.

5 H. Onouchi, Y. Nagami, Y. Haraguchi, M. Nakamoto, Y. Nishimura, R. Sakurai, N. Nagao, D. Kawasaki, Y. Kadokura and S. Naito, *Genes Dev*., 2005, **19**, 1799–1810.

6 The Arabidopsis Initiative, *Nature*, 2000, **408**, 796–815.

7 Q. Feng, Y. Zhang, P. Hao, S. Wang, G. Fu, Y. Huang, Y. Li, J. Zhu, Y. Liu, X. Hu, P. Jia, Y. Zhang, Q. Zhao, K. Ying, S. Yu, Y. Tang, Q. Weng, L. Zhang, Y. Lu, J. Mu, Y. Lu, L. S. Zhang, Z. Yu, D. Fan, X. Liu, T. Lu, C. Li, Y. Wu, T. Sun, H. Lei, T. Li, H. Hu, J. Guan, M. Wu, R. Zhang, B. Zhou, Z. Chen, L. Chen, Z. Jin, R. Wang, H. Yin, Z. Cai, S. Ren, G. Lv, W. Gu, G. Zhu, Y. Tu, J. Jia, Y. Zhang, J. Chen, H. Kang, X. Chen, C. Shao, Y. Sun, Q. Hu, X. Zhang, W. Zhang, L. Wang, C. Ding, H. Sheng, J. Gu, S. Chen, L. Ni, F. Zhu, W. Chen, L. Lan, Y. Lai, Z. Cheng, M. Gu, J. Jiang, J. Li, G. Hong, Y. Xue and B. Han, *Nature*, 2002, **420**, 316–320.

8 T. Sasaki, T. Matsumoto, K. Yamamoto, K. Sakata, T. Baba, Y. Katayose, J. Wu, Y. Niimura, Z. Cheng, Y. Nagamura, B. A. Antonio, H. Kanamori, S. Hosokawa, M. Masukawa, K. Arikawa, Y. Chiden, M. Hayashi, M. Okamoto, T. Ando, H. Aoki, K. Arita, M. Hamada, C. Harada, S. Hijishita, M. Honda, Y. Ichikawa, A. Idonuma, M. Iijima, M. Ikeda, M. Ikeno, S. Ito, T. Ito, Y. Ito, Y. Ito, A. Iwabuchi, K. Kamiya, W. Karasawa, S. Katagiri, A. Kikuta, N. Kobayashi, I. Kono, K. Machita, T. Maehara, H. Mizuno, T. Mizubayashi, Y. Mukai, H. Nagasaki, M. Nakashima, Y. Nakama, Y. Nakamichi, M. Nakamura, N. Namiki, M. Negishi, I. Ohta, N. Ono, S. Saji, K. Sakai, M. Shibata, T. Shimokawa, A. Shomura, J. Song, Y. Takazaki, K. Terasawa, K. Tsuji, K. Waki, H. Yamagata, H. Yamane, S. Yoshiki, R. Yoshihara, K. Yukawa, H. Zhong, H. Iwama, T. Endo, H. Ito, J. H. Hahn, H.-I. Kim, M.-Y. Eun, M. Yano, J. Jiang and T. Gojobori, *Nature*, 2002, **420**, 312–316.

9 S. A. Goff, D. Ricke, T.-H. Lan, G. Presting, R. Wang, M. Dunn, J. Glazebrook, A. Sessions, P. Oeller, H. Varma, D. Hadley, D. Hutchison, C. Martin, F. Katagiri, B. M. Lange, T. Moughamer, Y. Xia, P. Budworth, J. Zhong, T. Miguel, U. Paszkowski, S. Zhang, M. Colbert, W.-l. Sun, L. Chen, B. Cooper, S. Park, T. C. Wood, L. Mao, P. Quail, R. Wing, R. Dean, Y. Yu, A. Zharkikh, R. Shen, S. Sahasrabudhe, A. Thomas, R. Cannings, A. Gutin, D. Pruss, J. Reid, S. Tavtigian, J. Mitchell, G. Eldredge, T. Scholl, R. M. Miller, S. Bhatnagar, N. Adey, T. Rubano, N. Tusneem, R. Robinson, J. Feldhaus, T. Macalma, A. Oliphant and S. Briggs, *Science*, 2002, **296**, 92–100.

10 J. Yu, S. Hu, J. Wang, G. K.-S. Wong, S. Li, B. Liu, Y. Deng, L. Dai, Y. Zhou, X. Zhang, M. Cao, J. Liu, J. Sun, J. Tang, Y. Chen, X. Huang, W. Lin, C. Ye, W. Tong, L. Cong, J. Geng, Y. Han, L. Li, W. Li, G. Hu, X. Huang, W. Li, J. Li, Z. Liu, L. Li, J. Liu, Q. Qi, J. Liu, L. Li, T. Li, X. Wang, H. Lu, T. Wu, M. Zhu, P. Ni, H. Han, W. Dong, X. Ren, X. Feng, P. Cui, X. Li, H. Wang, X. Xu, W. Zhai, Z. Xu, J. Zhang, S. He, J. Zhang, J. Xu, K. Zhang, X. Zheng, J. Dong, W. Zeng, L. Tao, J. Ye, J. Tan, X. Ren, X. Chen, J. He, D. Liu, W. Tian, C. Tian, H. Xia, Q. Bao, G. Li, H. Gao, T. Cao, J. Wang, W. Zhao, P. Li, W. Chen, X. Wang, Y. Zhang, J. Hu, J. Wang, S. Liu, J. Yang, G. Zhang, Y. Xiong, Z. Li, L. Mao, C. Zhou, Z. Zhu, R. Chen, B. Hao, W. Zheng, S. Chen, W. Guo, G. Li, S. Liu, M. Tao, J. Wang, L. Zhu, L. Yuan and H. Yang, *Science*, 2002, **296**, 79–92.

11 M. Y. Hirai, T. Fujiwara, M. Awazuhara, T. Kimura, M. Noji and K. Saito, *Plant J*., 2003, **33**, 651–663.

12 M. Y. Hirai, M. Yano, D. B. Goodenowe, S. Kanaya, T. Kimura, M. Awazuhara, M. Arita, T. Fujiwara and K. Saito, *Proc. Natl. Acad. Sci. U. S. A*., 2004, **101**, 10205–10210.

13 A. Maruyama-Nakashita, E. Inoue, A. Watanabe-Takahashi, T. Yamaya and H. Takahashi, *Plant Physiol*., 2003, **132**, 597–605.

14 V. Nikiforova, J. Freitag, S. Kempa, M. Adamik, H. Hesse and R. Hoefgen, *Plant J*., 2003, **33**, 633–650.

15 M. Y. Hirai and K. Saito, *J. Exp. Bot*., 2004, **55**, 1871–1879.

16 Y. Sasaki-Sekimoto, N. Taki, T. Obayashi, M. Aono, F. Matsumoto, N. Sakurai, H. Suzuki, M. Y. Hirai, M. Noji, K. Saito, T. Masuda, K. Takamiya, D. Shibata and H. Ohta, *Plant J*., 2005, **44**, 653–668.

17 R. J. Bino, R. D. Hall, O. Fiehn, J. Kopka, K. Saito, J. Draper, B. J. Nikolau, P. Mendes, U. Roessner-Tunali, M. H. Beale, R. N. Trethewey, B. M. Lange, E. S. Wurtele and L. W. Sumner, *Trends Plant Sci*., 2004, **9**, 418–425.

18 K. M. Oksman-Caldentey and K. Saito, *Curr. Opin. Biotechnol*., 2005, **16**, 174–179.

19 K. Saito, M. Y. Hirai and K. Yonekura-Sakakibara, *Trends Plant Sci*., 2008, **13**, 36–43.

20 V. J. Nikiforova, J. Kopka, V. Tolstikov, O. Fiehn, L. Hopkins, M. J. Hawkesford, H. Hesse and R. Hoefgen, *Plant Physiol*., 2005, **138**, 304–318.

21 M. Y. Hirai, M. Klein, Y. Fujikawa, M. Yano, D. B. Goodenowe, Y. Yamazaki, S. Kanaya, Y. Nakamura, M. Kitayama, H. Suzuki, N. Sakurai, D. Shibata, J. Tokuhisa, M. Reichelt, J. Gershenzon, J. Papenbrock and K. Saito, *J. Biol. Chem*., 2005, **280**, 25590–25595.

22 V. J. Nikiforova, C. O. Daub, H. Hesse, L. Willmitzer and R. Hoefgen, *J. Exp. Bot*., 2005, **56**, 1887–1896.

23 R. Morioka, S. Kanaya, M. Y. Hirai, M. Yano, N. Ogasawara and K. Saito, *BMC Bioinf*., 2007, **8**, 343.

24 T. Abe, S. Kanaya, M. Kinouchi, Y. Ichiba, T. Kozuki and T. Ikemura, *Genome Res*., 2003, **13**, 693–702.

25 S. Kanaya, M. Kinouchi, T. Abe, Y. Kudo, Y. Yamada, T. Nishi, H. Mori and T. Ikemura, *Gene*, 2001, **276**, 89–99.
26 J. Szymanski, M. Bielecka, F. Carrari, A. R. Fernie, R. Hoefgen and V. J. Nikiforova, *Phytochemistry*, 2007, **68**, 2163–2175.
27 M. Y. Hirai, K. Sugiyama, Y. Sawada, T. Tohge, T. Obayashi, A. Suzuki, R. Araki, N. Sakurai, H. Suzuki, K. Aoki, H. Goda, O. I. Nishizawa, D. Shibata and K. Saito, *Proc. Natl. Acad. Sci. U. S. A.*, 2007, **104**, 6478–6483.
28 T. Gigolashvili, M. Engqvist, R. Yatusevich, C. Muller and U. I. Flugge, *New Phytol.*, 2007, **177**, 627–642.
29 I. E. Sonderby, B. G. Hansen, N. Bjarnholt, C. Ticconi, B. A. Halkier and D. J. Kliebenstein, *PLoS One*, 2007, **2**, e1322.
30 Y. Higashi, M. Y. Hirai, T. Fujiwara, S. Naito, M. Noji and K. Saito, *Plant J.*, 2006, **48**, 557–571.
31 L. M. Tabe and M. Droux, *Plant Physiol.*, 2002, **128**, 1137–1148.
32 N. D. Hagan, N. Upadhyaya, L. M. Tabe and T. J. V. Higgins, *Plant J.*, 2003, **34**, 1–11.
33 P. Chiaiese, N. Ohkama-Ohtsu, L. Molvig, R. Godfree, H. Dove, C. Hocart, T. Fujiwara, T. J. V. Higgins and L. M. Tabe, *J. Exp. Bot.*, 2004, **55**, 1889–1901.
34 N. Islam, N. M. Upadhyaya, P. M. Campbell, R. Akhurst, N. Hagan and T. J. V. Higgins, *Phytochemistry*, 2005, **66**, 2534–2539.