Metabolic Fingerprinting and Profiling of *Arabidopsis thaliana* Leaf and its Cultured Cells T87 by GC/MS[§]

Eiichiro Fukusaki[a,*], Kanokwan Jumtee[a], Takeshi Bamba[a], Takehiro Yamaji[b], and Akio Kobayashi[a]

[a] Department of Biotechnology, Graduate School of Engineering, Osaka University, 2-1 Yamadaoka, Suita, Osaka, 565-0871, Japan. Fax: +81-6-6879-7424. E-mail: fukusaki@bio.eng.osaka-u.ac.jp

[b] Production Department, GL Science Inc., 237-2 Sayamagahara, Iruma, Saitama, 358-0032, Japan

* Author for correspondence and reprint requests

Z. Naturforsch. **61c**, 267–272 (2006); received August 17/September 16, 2005

Cell suspension cultures are now recognized as important model materials for plant bioscience and biotechnology. Very few studies of metabolic comparisons between cell cultures and original plants have been reported, even though the biological identity of cultured cells with the normally grown plant is of great importance. In this study, a comparison of the metabolome for primary metabolites extracted from the leaves of *Arabidopsis thaliana* and cultured cells from an *Arabidopsis* suspension culture (cell line T87) was performed. The results suggest that although cell suspension cultures and *Arabidopsis* leaves showed similarities in the common primary metabolite profile, nonetheless, moderate differences in quantitative profile were revealed.

Key words: Arabidopsis thaliana, Metabolomics, T87 Cultured Cells

Introduction

Arabidopsis thaliana, which can be grown under a wide variety of environment conditions and systems, and whose genome sequence has been completely revealed, is widely used as a model plant in molecular biology. Several factors affect the general and specific pathways in plant cells resulting in a variety of gene expression profiles and their downstream products, for example primary and secondary metabolites. Generally, *Arabidopsis* plants cultivated in soil or hydroponic systems are utilized for studies of changes in gene expression (Richards *et al.*, 1998; Maathuis *et al.*, 2003), and morphological and physiological characteristics (Sun *et al.*, 2004) of systems responding to environmental stress. Likewise, cell suspension cultures are also appropriate model systems due to their homogenous and rapid rate of growth and productive amounts of cells. *Arabidopsis* suspension cultured cells have been used for preliminary observation of plant response to exogenous stimuli (Yamada *et al.*, 2004), for investigating cell cycle regulation (Callard *et al.*, 1996; Menges *et al.*, 2003), and as ideal materials for isolating protoplasts used for transient gene expression assays (An *et al.*, 2003). However, it has been reported that some secondary metabolites of tissue cultures and cultured cells are different from those of whole plants (Jalal and Collin, 1977; Koulman *et al.*, 2003). Despite cell suspension cultures being recognized as important model materials to plant biologists, no comparison study between the metabolome of *Arabidopsis* plants and suspension cultured cells has been reported. In this paper, we aimed to investigate the similarity and dissimilarity of the metabolome focusing on primary metabolites from the leaves of *Arabidopsis thaliana* plants and a cell suspension culture of *Arabidopsis thaliana* cell line T87.

Materials and Methods

Plant material

Seeds of *Arabidopsis thaliana* (L.) Heynh. ecotype Columbia were stratified for 2 d at 4 °C on a

[§] The study represents a portion of the dissertation submitted by Kanokwan Jumtee to Osaka University in partial fulfilment of the requirements for her PhD. degree.

0939–5075/2006/0300–0267 $ 06.00 © 2006 Verlag der Zeitschrift für Naturforschung, Tübingen · http://www.znaturforsch.com · D

moist filter paper and then transferred to Metro-Mix-350 soil in 6.35-cm pots. The pots were kept in a growth chamber at 23 °C with 16 h/8 h daylight/dark. Plants were watered by sub-irrigation every 2 or 3 d. Relative humidity was maintained at 55 to 60%. The average light intensity at the top of the pots was approx. 40 μmol m^{-2} s^{-1}.

Cell suspension culture

A suspension culture of *Arabidopsis thaliana* cell line T87, which was established from the ecotype Columbia, was maintained as previously described (Axelos *et al.*, 1992). The cell suspension was mixotrophically cultivated in 100 mL of modified LS medium supplied with 3% sucrose at 23 °C under continuous illumination with the average light intensity 40 μmol m^{-2} s^{-1} and was subcultured into a new fresh medium every week.

Sampling, extraction, and derivatization procedures

When the plants reached growth stage 6.00 (Boyes *et al.*, 2001) fully expanding rosette leaves were harvested during the day 8 h after the start of the photoperiod. Leaves were weighed (100 mg ± 1 mg) and immediately frozen in liquid nitrogen. After cultivation for 7 d, T87 suspension culture cells were transferred from the culture flask, washed thoroughly with MilliQ water, and filtered by vacuum filtration. The same amounts as above of the washed cells were weighed and immediately frozen in liquid nitrogen.

Leaves and cells were homogenized and disrupted with MM 301 mixer mills (Retsch GmbH & Co. KG, Haan, Germany) and then extracted with 1000 μL of a mixture of methanol/water/chloroform (2.5:1:1 v/v/v). A ribitol solution (60 μL, 0.2 mg mL^{-1}) was added as an internal standard. The extraction was performed at 37 °C with a mixing frequency of 1200 rpm using a thermomixer compact. The solutions were then centrifuged at 16,000 × *g* for 3 min. The polar phase (900 μL) was transferred to a new Eppendorf tube and mixed with 400 μL of MilliQ water. The polar phase was separated by centrifugation and transferred to a new Eppendorf tube for evaporation under vacuum to dryness. Derivatizations were carried out by adding 50 μL of methoxyamine hydrochloride (20 mg mL^{-1}) in pyridine and shaking at 30 °C for 90 min. The sample was then silylated for 30 min at 37 °C by adding 100 μL of *N*-methyl-*N*-(trimethylsilyl) trifluoroacetamide (MSTFA).

GC/MS analysis

1 μL of the TMS derivatized sample was injected by an AS 3000 autosampler into a TRACE GC gas chromatograph (Thermo Electron, Milan, Italy) equipped with a 30 m × 0.25 mm i.d. fused-silica capillary column coated with 0.25-μm CP-SIL 8 CB low bleed (Variance) and coupled to a TRACE DSQ mass spectrometer (Thermo Electron, Austin, TX). The injector temperature was 230 °C and the helium gas flow rate through the column was 1 mL/min. The column temperature was held at 70 °C for 5 min, then raised by 6 °C/min to 310 °C, and held there for 4 min. The transfer line and the ion source temperatures were 250 °C and 200 °C, respectively. Ions were generated by a 70-eV electron beam, and two scans per second were recorded in the mass range 50–600 *m/z*. The acceleration voltage was turned on after a solvent delay of 5 min.

Data analysis

The Automatic Mass Spectral Deconvolution and Identification System (AMDIS) (NIST, Gaithersburg, MD) was applied to the chromatographic/spectrometric data prior to the identification by comparing mass spectra with those in the NIST98 Mass Spectral Database and an in-house library (Halket *et al.*, 1999). Additionally, the library provided from Max-Planck-Institute of Molecular Plant Physiology, Germany was also used for this purpose (http://www.mpimp-golm.mpg.de/mms-library/index-e.html). For metabolite fingerprinting, the original data were converted to netCDF by the file converter tool in Xcalibur® software (Thermo Electron, Austin, TX). The converted files were extracted, normalized, and exported using the in-house program. The retention time variability was adjusted to a target chromatogram using LineUp® v2.0 software (Infometrix, Woodinville, WA). The resulting files were then further processed and visualized by hierarchical cluster analysis (HCA) and principal component analysis (PCA) using Pirouette® v3.11 software (Infometrix, Woodinville, WA). For HCA, the data was transformed by log 10 to allow better comparison of large and small variables. The Euclidean distance was used to calculate the matrix of all samples and the incremental linkage method was used in the assignment of clusters for HCA. For PCA, the data set was also transformed by log 10 and mean-centering scaling. The explained

factors, scores, loading, and outlier diagnostics, as well as other parameters, were observed for statistical classification.

Results and Discussion

Biological materials and method selection

We selected conventional cultivation systems for both *Arabidopsis* plants and *Arabidopsis* suspension cultured cells. However, it should be emphasized that the growth conditions were different. We decided to harvest fully expanded leaves when the plants reached growth stage 6.00 according to Boyes *et al.* (2001) and sampled the cultured cells after sub-cultivation for 7 d. Preliminary experiments showed that cultured cells were in the stationary phase during this period.

For screening hydrophilic compounds, we chose an analytical method using GC/MS as described previously (Fiehn *et al.*, 2000). This method, which covered various classes of compounds, *e.g.*, amino acids, amines, fatty acids, mono- and disaccharides, and organic acids, was validated as an accurate and highly reproducible technique. Before identification, deconvolution was done using the Automatic Mass Spectral Deconvolution and Identification System (AMDIS) program, which is available with the NIST98 Mass Spectral Database (Halket *et al.*, 1999). The identification of compounds was based on retention times and comparison of mass spectra with those of standards in libraries.

Identification of common metabolites

Thirty leaf samples and 21 cell culture samples obtained from 30 identical plants and two cultivation flasks, respectively, were investigated. From GC/MS chromatograms obtained from leaf and T87 cell extracts, it was clear that, given the same weight of fresh tissue, the T87 cells contained higher metabolite concentrations than the leaves of *Arabidopsis.* Analysis of the complex chromatograms by the AMDIS program and NIST library revealed both known and unknown compounds. The identified compounds which were present in both *Arabidopsis* leaves and *Arabidopsis* T87 cells are shown in Table I and those which were presented in leaves or T87 cells, but not in both, are shown in Table II. The list of compounds in *Arabidopsis* leaves presented here is slightly different from the list reported previously (Fiehn *et al.*, 2000) because the samples in this study were merely derivatized with the trimethylsilyl (TMS) group.

Table I. Identified compounds present in both *Arabidopsis* leaf and *Arabidopsis* T87 cell extracts.

Amino acids	*Nitrogen compounds*
Alanine	1,4-Butanediamine
Asparagine	*n*-Butylamine
Aspartate	Cadaverine
Glutamine	Ethanolamine
Glutamate	Ethylamine
Glycine	Histamine
Isoleucine	3-Hydroxypyridine
Leucine	Urea
Phenylalanine	
Proline	*Organic acids derivatives*
Serine	4-Aminobutyrate (GABA)
Threonine	Citrate
Valine	Fumarate
β-Alanine	Glycerate
Pyroglutamate	Glyoxylate
Lysine	4-Hydroxybutanoate (GHB)
Sugars/Sugar alcohols	2-Keto-l-gluconate
Arabinose	Lactate
Fructose	Maleate
Galactose	Malate
Glucose	Oxalate
Gulose	Phosphoric acid
Mannose	Pipecolate
Sucrose	Pyruvate
Glycerol	Succinate
Myo-inositol	

Out of a total of 68 metabolites of *Arabidopsis* leaves and T87 cultured cells identified here, 48 metabolites were identical. These common compounds consist mainly of amino acids, mono- and disaccharides, and organic acids derivatives from TCA cycles. T87 cultured cells shared these com-

Table II. Identified compounds presented in either *Arabidopsis* leaf or *Arabidopsis* T87 cell extracts.

Detected only in leaf	Detected only in T87 cell culture
Ascorbate	Allantoin
Cellobiose	Allothreonine
Galacitol	Galactonate
Raffinose	Glucarate
Shikimate	Glucose-6-phosphate
Spermidine	Histidine
Trehalose	Homoserine
	Methionine
	S-Methylcysteine
	Nicotinate
	Ribose-5-phosphate
	Tryptophan
	Tyrosine

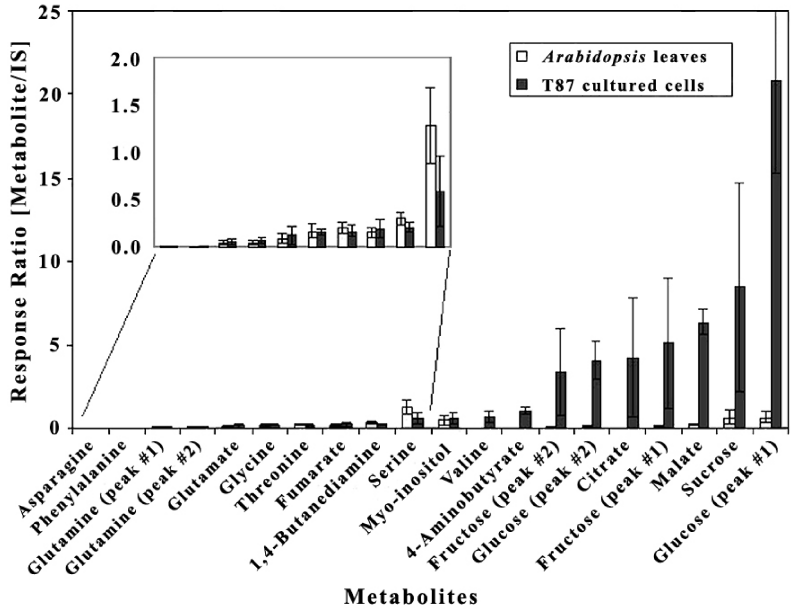

Fig. 1. Differences in metabolite levels of some hydrophilic metabolites in *Arabidopsis* rosette leaves and *Arabidopsis* T87 cells. Relative quantitative analyses were performed for individual metabolites obtained from GC/MS selected ion monitoring chromatogram by means of integration of peak area of each metabolite and normalization with peak area of internal standards. The insert represents the expansion of the bar graph. IS, internal standard; fructose (peak #1), peak at 27.0 min; fructose (peak #2), peak at 27.2 min; glucose (peak #1), peak at 27.3 min; glucose (peak #2), peak at 27.7 min; glutamine (peak #1), peak at 24.6 min; glutamine (peak #2), peak at 25.4 min.

mon metabolites with whole plants; however, there were some amino acids, organic acids and sugar phosphates that could not be detected in *Arabidopsis* leaf extracts analyzed by the present protocol. Moreover, beside mono-and disaccharides, trisaccharides were observed in the leaves from whole plants.

We selected certain metabolites and performed relative quantitative analyses by means of peak detection and deconvolution of co-eluting peaks with selected ion monitoring (SIM), followed by the integration of peak area of each metabolite and normalization with those of internal standards. The results are shown in Fig. 1. This revealed differences in metabolite levels of some hydrophilic metabolites present in both samples. Glucose, fructose, sucrose, malate, citrate, valine, and 4-aminobutyrate are significantly higher in the cells of T87 suspension cultures. Other primary metabolites such as glycine, phenylalanine, glutamine, asparagine, fumarate, myo-inositol were slightly higher in T87 cultured cells, whereas only 1,4-butanediamine, threonine and serine were higher in leaf tissue. In this study, T87 cultured cells were cultivated under mixotrophic conditions. It should be noted that exogenous sucrose that was uptaken by cells might drive the increasing of sugars and primary metabolism compounds.

We selected metabolite fingerprinting as a virtualization tool for these two sample sets. Although there is a freely available program for data processing of metabolic profiling (Duran *et al.*, 2003), we decided to perform data preprocessing using an in-house program as described in the experimental section. After processing the data, the processed files were subjected to the Pirouette® program for hierarchical cluster analysis (HCA) and principal component analysis (PCA).

HCA was applied to the processed data of T87 cell extracts and leaf extracts for preliminary observation of these data sets. The HCA described here was transformed by log 10 to allow better comparison of large and small variables. The Euclidean distance was used as the metric of data relatedness and the incremental linkage method was then used for data clustering. As a result, hierarchically ordered clusters were formed. The distances between clusters could then be depicted as a dendrogram. The dendrogram obtained from the HCA is illustrated in Fig. 2, indicating 2 clusters in these samples; of leaf and T87 cells. Even these sample sets showed similarities in the common primary metabolite profile, two separate clusters observed from the dendrogram suggest differences in metabolite levels between them.

PCA was then applied to the data as another approach. PCA is a powerful visualization tool and finds use in exploratory data analysis. It can represent graphically intersample and intervariable relationships, and finds linear combinations of

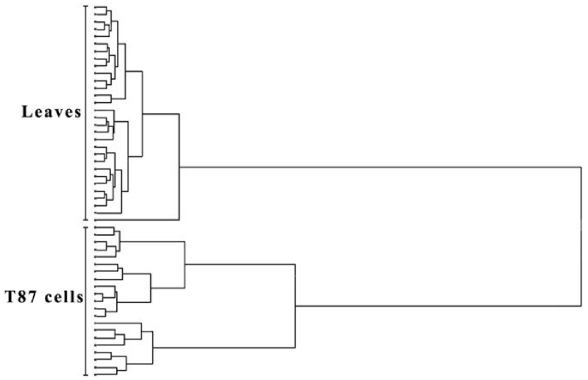

Fig. 2. Dendrogram obtained from HCA of *Arabidopsis* leaf (30 samples) and *Arabidopsis* T87 cell samples (21 samples). The distances between these samples were calculated as described in Materials and Methods. The data was transformed by log 10 to allow better comparison of large and small variables. The Euclidean distance was used to calculate the matrix of all samples and the incremental linkage method was used in the assignment of clusters. Hierarchical clustering of the samples using Euclidean distances results in the separation of leaves and T87 cells.

the original independent variables which account for maximal amounts of variation. PCA is based on the idea of expressing a matrix **X** as the product of two other matrices (score and loading matrices):

$$\mathbf{X} = \mathbf{T}_k\mathbf{L}_k{}^{\mathrm{T}} + \mathbf{E}_k,$$

where **X** is the original data matrix, **T** is the score matrix, **L** is the loading matrix, **E** is the residual matrix and k is the number of components. In other words, PCA can approximate the original matrix **X** to k dimensional subspace (Otto, 1999).

For PCA, the data sets were transformed by log 10 as those in HCA and represented as mean-centering scaling. After transformation, a large set of related variables was reduced into a new, smaller set of uncorrelated variables, or principal component (PC). Score plots of independent PC1 and PC2 for the processed data sets are shown in Fig. 3A. The explained variances of these data showed that 90% of the data variance can be explained by use of two principal components. Therefore, the first two principal components are sufficient to examine these samples. This also illustrated a clear clustering of two groups; leaf samples and T87 cell samples. However, it was clear from this figure that only PC1 played a significant role in discriminating leaf samples and T87 cell

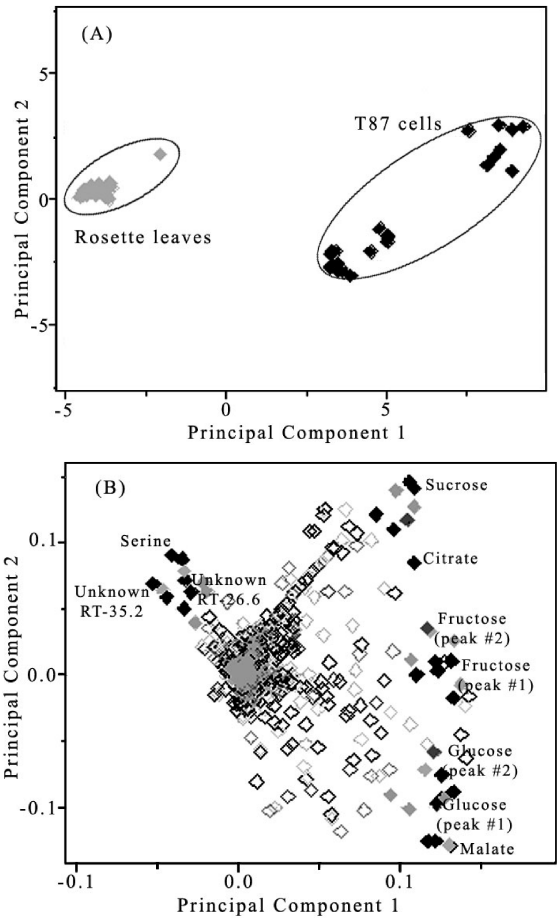

Fig. 3. Score (A) and loading (B) plots of principal component (PC) 1 and 2 generated from PCA of *Arabidopsis* leaf and *Arabidopsis* T87 cell samples. The distances between these samples were calculated as described in Materials and Methods. The data set was transformed by log 10 and mean-centering scaling. (A) The PCA score plot of PC1 and PC2 includes 90% of the information derived from data variances and distinguishes between series of 30 leaf samples (rosette leaves) and 21 T87 cell samples (T87 cells). (B) The PCA loading plot illustrates the contribution of individual metabolites to the PCA vector. The closer to the origin, the smaller the influence a given metabolite has on the discrimination. The most important metabolites playing a role in the separation of leaf and T87 cell extracts are labelled. Fructose (peak #1), peak at 27.0 min; fructose (peak #2), peak at 27.2 min; glucose (peak #1), peak at 27.3 min; glucose (peak #2), peak at 27.7 min

samples. These results also suggest that the leaf metabolome is not identical to that of the T87 cells even though both metabolomes are very similar. In addition, this illustrated a clear clustering of two groups; leaf samples and T87 cell samples,

strongly confirming that the metabolic profiles within these two samples were very similar.

The contribution of each variable to PC1 and PC2 was calculated giving each variable a weighing value or loading for PC. The corresponding PCA loading plot of PC1 and PC2 is illustrated in Fig. 3B. This shows the distribution of metabolites in the form of retention time to the principal component. Retention times closer to zero have less influence on the discrimination. As mentioned previously, PC1 predominantly separated T87 cell samples from leaf samples, with T87 cell samples having the higher loading value of PC1. For example, some metabolites such as sucrose (RT = 37.9), glucose (RT = 27.3 and 27.7), fructose (RT = 27.0 and 27.2), citrate (RT = 26.2), malate (RT = 19.9), and 4-aminobutyrate (RT = 20.8) had a comparatively strong impact on the PC1. This also confirms that *Arabidopsis* T87 cells contained higher amounts of these compounds than *Arabidopsis* rosette leaves given the cultivation conditions and sampling protocols mentioned in Materials and Methods.

Acknowledgements

The authors would like to thank the Research Institute of Innovative Technology for the Earth (RITE), Japan and the New Energy and Industrial Technology Development Organization (NEDO), Japan for their supports to this project. We also thank the Max-Planck-Institute of Molecular Plant Physiology, Germany providing the library of TMS derivatized compounds.

An C., Sawada A., Fukusaki E., and Kobayashi A. (2003), Transient RNAi induction in *Arabidopsis* protoplasts. Biosci. Biotechnol. Biochem. **67**, 2674–2677.

Axelos M., Curie C., Mazzolini L., Bardet C., and Lescure B. (1992), A protocol for transient gene expression in *Arabidopsis thaliana* protoplasts isolated from cell suspension cultures. Plant Physiol. Biochem. **30**, 123–128.

Boyes D. C., Zayed A. M., Ascenzi R., McCaskill A. J., Hoffman N. E., Davis K. R., and Görlach, J. (2001), Growth stage-based phenotypic analysis of *Arabidopsis*: A model for high throughput functional genomics in plants. Plant Cell **13**, 1499–1510.

Callard D., Axelos M., and Mazzolini L. (1996), Novel molecular markers for late phase of the growth of *Arabidopsis thaliana* cell-suspension cultures are expressed during organ senescence. Plant Physiol. **112**, 705–715.

Duran A. L., Yang J., Wang L., and Sumner L. W. (2003), Metabolomics spectral formatting, alignment and conversion tools (MSFACTs). Bioinformatics **19**, 2283–2293.

Fiehn O., Kopka J., Trethewey R. N., and Willmitzer L. (2000), Identification of uncommon plant metabolites based on calculation of elemental compositions using gas chromatography and quadrupole mass spectrometry. Anal. Chem. **72**, 3573–3580.

Halket J. M., Przyborowska A., Stein S. E., Mallard W. G., Down S., and Chalmers R. A. (1999), Deconvolution gas chromatography/mass spectrometry of urinary organic acids- – potential for pattern recognition and automated identification of metabolic disorders. Rapid Commun. Mass Spectrom. **13**, 279–284.

Jalal M. A. F. and Collin H. A. (1977), Polyphenols of mature plant, seedling and tissue cultures of *Theobroma cacao*. Phytochemistry **16**, 1377–1380.

Koulman A., Kubbinga M. E., Batterman S., Woerdenbag H. J., Pras N., Woolley J. G., and Quax W. J. (2003), A phytochemical study of lignans in whole plants and cell suspension cultures of *Anthriscus sylvestris*. Planta Med. **69**, 733–738.

Maathuis F. J. M., Filatov V., Herzyk P., Krijger G. C., Axelsen K. B., Chen S., Green B. J., Li Y., Madagan K. L., Sánchez-Fernández R., Forde B. G., Palmgren M. G., Rea P. A., Williams L. E., Sanders D., and Amtmann A. (2003), Transcriptome analysis of root transporters reveals participation of multiple gene families in the response to cation stress. Plant J. **35**, 675–692.

Menges M., Hennig L., Gruissern W., and Murray J. A. H. (2003), Genome-wide gene expression in an *Arabidopsis* cell suspension. Plant Mol. Biol. **53**, 423–442.

Otto M. (1999), Chemometrics: Statistics and Computer Application in Analytical Chemistry. Wiley-VCH, Weinheim, p. 125.

Richards K. D., Schott E. J., Sharma Y. K., Davis K. R., and Gardner R. C. (1998), Aluminum induces oxidative stress genes in *Arabidopsis thaliana*. Plant Physiol. **116**, 409–418.

Sun K., Hunt K., and Hauser B. A. (2004), Ovule abortion in *Arabidopsis* triggered by stress. Plant Physiol. **135**, 2358–2367.

Yamada H., Koizumi N., Nakamichi N., Kiba T., Yamashino T., and Mizuno T. (2004), Rapid response of *Arabidopsis* T87 cultured cells to cytokinin through His-to-Asp phosphorelay signal transduction. Biosci. Biotechnol. Biochem. **68**, 1966–1976.