Short Communication

Wolfram Weckwerth
Kathrin Wenzel
Oliver Fiehn

Max-Planck-Institut
für Molekulare
Pflanzenphysiologie,
Golm, Germany

Process for the integrated extraction, identification and quantification of metabolites, proteins and RNA to reveal their co-regulation in biochemical networks

A novel extraction protocol is described with which metabolites, proteins and RNA are sequentially extracted from the same sample, thereby providing a convenient procedure for the analysis of replicates as well as exploiting the inherent biological variation of independent samples for multivariate data analysis. A detection of 652 metabolites, 297 proteins and clear RNA bands in a single *Arabidopsis thaliana* leaf sample was validated by analysis with gas chromatography coupled to a time of flight mass spectrometer for metabolites, two-dimensional liquid chromatography coupled to mass spectrometry for proteins, and Northern blot analysis for RNA. A subset of the most abundant proteins and metabolites from replicate analysis of different *Arabidopsis* accessions was merged to form an integrative dataset allowing both classification of different genotypes and the unbiased analysis of the hierarchical organization of proteins and metabolites within a real biochemical network.

Keywords: Gas chromatography-time of flight mass spectrometer / Metabolomics / Multidimensional chromatography / Network topology / Plant systems biology / Transcript profiling
PRO 0500

Arabidopsis thaliana plants were cultivated in phytotrons under highly controlled light, gas and temperature conditions assuring approximately identical environmental conditions for each plant sample. Biological variation among independent samples of the same genotypes is attributed to the inherent fluctuation of the biochemical network due to slight environmental differences.

30–100 mg samples of *Arabidopsis* leaves at a developmental stage 1.1 [1] were harvested and immediately frozen in liquid nitrogen. Tissue was homogenized under liquid nitrogen using a Retsch mill. Two mL of a single phase solvent mixture of methanol/chloroform/water 2.5:1:1 v/v/v kept at −20°C was added to the tissue and thoroughly mixed at 4°C for 30 min to precipitate proteins and DNA/RNA and to disassociate metabolites from membrane and cell wall components. After centrifugation, the remaining pellet consisting of DNA/RNA, proteins, starch, membranes, and cell wall components was extracted in a second step with 1 mL methanol/chloroform 1:1 v/v at −20°C. The organic solvent extracts were combined and used for metabolite analysis *via* GC-TOF. For that purpose the chloroform phase was separated from the water/methanol phase by adding 500 μL water. The resulting water/methanol phase contained all hydrophilic metabolites such as sugars, amino acids and organic acids, and the chloroform phase all the lipophilic compounds, lipids, chlorophyll and waxes. The remaining white pellet was further partitioned according to the scheme in Fig. 1. The pellet was extracted with 1 mL extraction buffer (0.05 M Tris, pH 7.6; 0.5% SDS; 1% β-mercaptoethanol) and 1 mL water saturated phenol for 1 h at 37°C. After centrifugation at 14 000 *g* the remaining pellet was used for cell wall synthesis (data not shown). The phenol phase was separated from the buffer phase and the proteins were precipitated with ice-cold acetone at −20°C overnight, washed three times with ethanol and dried at room temperature. Remaining protein in the RNA-buffer phase was precipitated with 200 μL chloroform. After centrifugation and separation of the buffer phase, 40 μL of acetic acid and 1 mL ethanol were added

Correspondence: Dr. Wolfram Weckwerth, Max-Planck-Institut für Molekulare Pflanzenphysiologie, Am Mühlenberg 1, D-14476 Golm, Germany
E-mail: weckwerth@mpimp-golm.mpg.de
Fax: +49-331-567-8134

Abbreviation: FW, fresh weight

Supporting information for this article is available at www.proteomics-journal.de or from the author at www.mpimp-golm.mpg.de/fiehn/index-e.html.

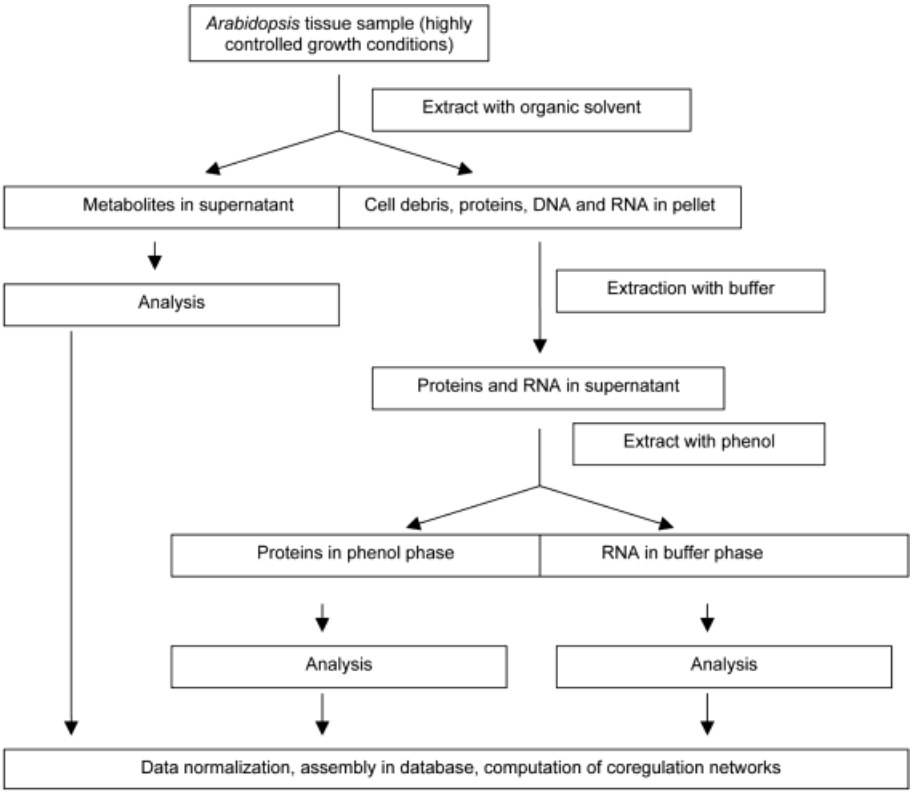

Figure 1. A separation scheme for the integrative extraction of metabolites, proteins and RNA from single samples enabling correlation analysis among these compound classes. Multiple system snapshots allow the analysis of dynamic biochemical interaction networks as a response of the observed genotype-environment-phenotype relationship [25].

to precipitate the RNA at 4°C for 30 min. The pellet was washed with one volume 3 M sodium acetate, and two times with one volume 70% ethanol. The remaining pellet was dissolved in 100 μL RNAse-free water. Amounts and purity of RNA were checked by absorbance at 260 nm and gel electrophoresis in agarose. Construction of *Arabidopsis* isopropyl-malate synthase (IPMS) probes for hybridization and Northern blots were performed using standard protocols.

For GC-TOF MS (Leco Pegasus II GC-TOF mass spectrometer; Leco, St. Joseph, MI, USA) analysis, the organic phase was dried and dissolved in 50 μL of methoxamine hydrochloride (20 mg/mL pyridine) and incubated at 30°C for 90 min with continuous shaking. Then 80 μL of *N*-methyl-*N*-trimethylsilyltrifluoroacetamid (MSTFA) was added to derivatize polar functional groups at 37°C for 30 min. The derivatized samples were stored at room temperature for 120 min before injection. GC-TOF analysis was performed on an HP 5890 gas chromatograph with tapered, deactivated split/splitless liners containing glasswool (Agilent, Böblingen, Germany) and 1 μL splitless injection at 230°C injector temperature. The GC was operated at constant flow of 1 mL/min helium and a 40 m 0.25 mm id 0.25 μm RTX-5 column with 10 m integrated precolumn. The temperature gradient started at 80°C, was held isocratic for 2 min, and subsequently ramped at 15°C/min to a final temperature of 330°C which was held for 6 min. Twenty spectra *per* second were recorded between *m/z* 85–500. Peak identification and quantification were performed using the Pegasus software package (Leco). Reference chromatograms were defined that had a maximum of detected peaks over a signal/noise threshold of 20 and used for automated peak identification based on mass spectral comparison to a standard NIST 98 library [2]. Automated assignments of unique fragment ions for each individual metabolite were taken as default as quantifiers, and manually corrected where necessary. All artifactual peaks caused by column bleeding or phtalates and polysiloxanes derived from MSTFA hydrolyzation were manually identified and removed from the results table. All data were normalized to plant mg fresh weight (FW) and internal references and log-transformed. *t*-test, correlation analysis, and variance analysis were performed in Microsoft Excel 5.0.

The dried protein pellet was dissolved in freshly prepared 1 M urea in 0.05 M Tris buffer pH 7.6. The complex protein mixture was digested with modified trypsin (Böhringer Mannheim, Mannheim, Germany) according to the manufacturer's instructions. The tryptic digest was dried down and dissolved in 300 μL water (1% formic acid). Insoluble

material was removed by centrifugation. An aliquot of the digest (~ 100 μg protein) was injected onto two-dimensional chromatography on a Thermofinnigan ProteomeX system (Thermofinnigan, San Jose, CA, USA) coupled to an LCQ DecaXp ion trap. The chromatographic separation was done according to manufacturer's instructions. After a 12 cycle run the MS/MS spectra were searched against an *A. thaliana* database (downloaded from the TAIR homepage www.arabidopsis.org) using TurboSequest implemented in Bioworks 3.0 (Thermofinnigan). Matches were filtered according to Wolters *et al.* [3] using the multiple scoring filter of Bioworks 3.0. For the quantification approach aliquots of the complex tryptic digest of *Arabidopsis* leaf protein (50 μg) were analyzed using 1-D reversed-phase chromatography. Quantification was achieved by integrating peak areas of target peptides representative for proteins. These peak areas were normalized to the sum of internal standard peptides that had been added to the mixture [4, 5].

All quantitative metabolite and protein data were normalized to internal standards and FW. Principal component analysis (PCA) and hierarchical cluster analysis (HCA) for pattern recognition was performed according to Fiehn *et al.* [6] using Pirouette software (Infometrix, Woodinville, WA, USA). The integrative data set of metabolites and proteins was $\log_{10}$ transformed. The HCA was performed using Euclidian distances and complete linkage grouping. Variance analyses were performed in MS Excel 5.0.

At the systems level, gene function is regarded as dependent on developmental stage, environmental conditions and expression levels of other genes, resulting in dynamic changes in transcript, protein and metabolite profiles. Thus, the next stage of understanding necessitates that biological tissues be described in depth on different levels, *i.e.* not only at the level of transcripts or protein expression, but also at the metabolite level and with consideration to the dynamic interaction of different gene products [7–11]. Here, we propose a novel extraction protocol for the integrative analysis of metabolites, proteins and RNA from the same sample. For each replicate, 30–100 mg FW leaf tissue of an individual *A. thaliana* plant was extracted at 4°C with chloroform/methanol/water (1:2.5:1 v/v/v) according to the scheme illustrated in Fig. 1. This single-phase mixture proved to have improved extraction strength for metabolites in comparison to the former extraction protocol utilizing a methanol/water mixture for 15 min at 70°C. However, when chloroform was left out of the cold extraction mix, *i.e.* if methanol/water extraction mixtures were used at −20°C, a strong decrease in sucrose content was detected in subsequent GC-TOF MS analysis, concomitant with a sharp increase in fructose and glucose contents. This indicates that chloroform may inhibit sucrose cleaving enzyme activity, *e.g.* invertase or sucrose synthase, by precipitating these enzymes.

Total metabolite analysis was performed with GC-TOF MS [12] (Fig. 2A) enabling the detection and quantification of 652 metabolites (see Supplementary Table 1). Replication of metabolite analysis revealed a high recovery and a mean coefficient of variance (CV) of 10%. In a subsequent step, proteins and mRNA were isolated from the remaining cell residue using buffer/phenol extraction and phase separation (see Fig. 1). In Fig. 2C, a comparison of this method with a conventional RNA extraction method is shown. The extraction procedure achieves a higher level of RNA recovery than does a typical RNA isolation kit extraction (see above) with 30% CV in 28 samples. To test the utility of the mRNA for hybridization we analyzed the expression of IPMS from *Arabidopsis* (Fig. 2D). The average amount of total protein extracted according to Fig. 1 was 1.3 mg *per* 100 mg FW with 17% CV. The overall extraction process resulted in good recovery of metabolites, proteins and transcripts. After complete extraction, the remaining cell pellet was used for cell wall analysis giving rise to clear and typical cell wall profiling (data not shown). The protein fraction was analyzed using shotgun proteomics [3, 13, 14].

The complex mixture of the tryptic *Arabidopsis* leaf protein digest was analyzed by 2-D capillary LC and MS/MS on an ion trap mass spectrometer (LCQ Deca Xp Plus) and a subsequent database search performed using TurboSequest implemented in ThermoFinnigan Bioworks 3.0. In a single *Arabidopsis* Col2 leaf sample extracted according to the scheme in Fig. 1, 586 peptides and 297 corresponding proteins were identified using very stringent criteria to avoid false positives (see above and Supplementary Table 2). A classification of detected proteins from one sample is shown in Fig. 2B. We applied the integrative extraction process to two *Arabidopsis* genotypes, C24 and Col2, to test if we were able to determine different biochemical phenotypes and general biochemical patterns using this strategy. C24 and Col2 showed an overlap of 153 proteins (see Supplementary Table 2). The data-dependent detection of peptides was strongly contingent on the estimated abundance of the corresponding proteins in the digest such as ribulose-1,5-biphosphate carboxylase/oxygenase (RUBISCO) [15]. Thus, the high number of nonoverlapping proteins also indicates differences in the protein-profiles of these different genotypes.

A set of 22 proteins appearing in both varieties was chosen for the quantification approach. These proteins were quantified by integrating their corresponding peptide areas in a 1-D LC-MS/MS analysis and normalizing these areas to internal standard peptides [4, 5]. Analytical

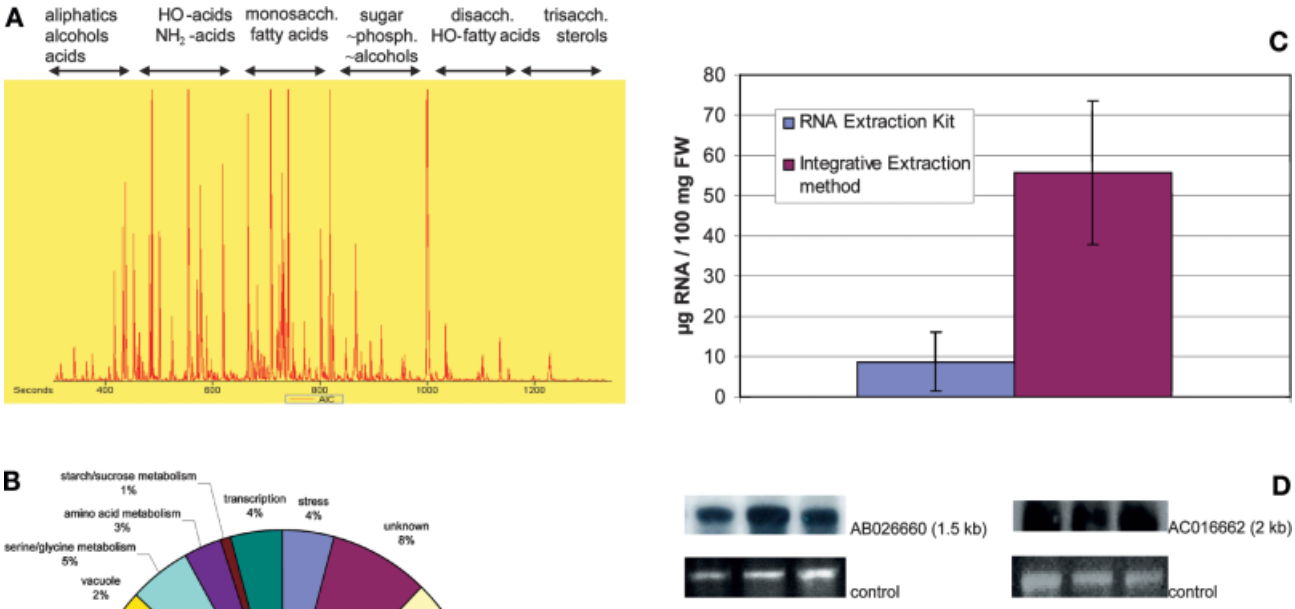

Figure 2. Analysis of metabolites, proteins, and transcripts extracted from a single *Arabidopsis* leaf sample. A, GC-TOF MS direct analysis of hydrophilic and lipophilic metabolites. B, Functional characterization of identified proteins from a single *Arabidopsis* Col2 leaf sample. Majority of the proteins are chloroplast-related. C, Comparison of the integrative extraction protocol with a conventional plant RNA extraction kit. D, RNA blot analysis of *Arabidopsis* isopropylmalate synthase (IPMS) transcripts (two isoforms, GI9758 and GI9759360) in three replicate samples of *Arabidopsis* Col2 leafs extracted according to the scheme in Fig. 1.

precision was tested by adding internal standard peptides to the sample. The deviation of the internal standards, in other words, the technical variation of the extraction process, stability of electrospray and matrix effects, was ~25% CV. Each genotype was represented by ten independent samples. The relative integrals of the peptides in each sample were normalized to the FW of the corresponding sample.

The metabolites in the corresponding samples were identified and quantified with GC-TOF. Fourteen of the most abundant metabolites were normalized to the internal standard and FW and combined with the protein data (normalized to internal peptide-standard and FW) to form an integrated dataset.

A homogeneous dataset was achieved by applying $\log_{10}$ transformation as described in [16]. Essential for the analysis is to test if we are able to discriminate between the two genotypes, to see if they have different biochemical phenotypes in identical environments. We applied principal component analysis (PCA) according to Fiehn *et al.* [6]. Both Col2 and C24 were completely separated into genotype-clusters (Fig. 3A). In contrast to NMR or other fingerprinting methods, the individual identification of compounds by our method enables the investigation of distinct metabolite-protein co-regulations in a multitude of samples. Examples are given in Figs. 3B and 3C. Based on the detection of these fundamental correlations in an integrative dataset (for instance a distance or a Pearsons matrix) and a comparison with hypothetical reaction pathway networks, it is possible to expose instantaneous connectivities in a regulatory network representing a snapshot of the actual state of the system [17–20].

To make use of such a refined analysis it is important to differentiate biological variability and technical measurement error. The quantified proteins showed an overall variability of ~39% whereas individual variation was up to 70%, exceeding clearly the overall analytical precision of ~ 25% CV. The same has been observed for meta-

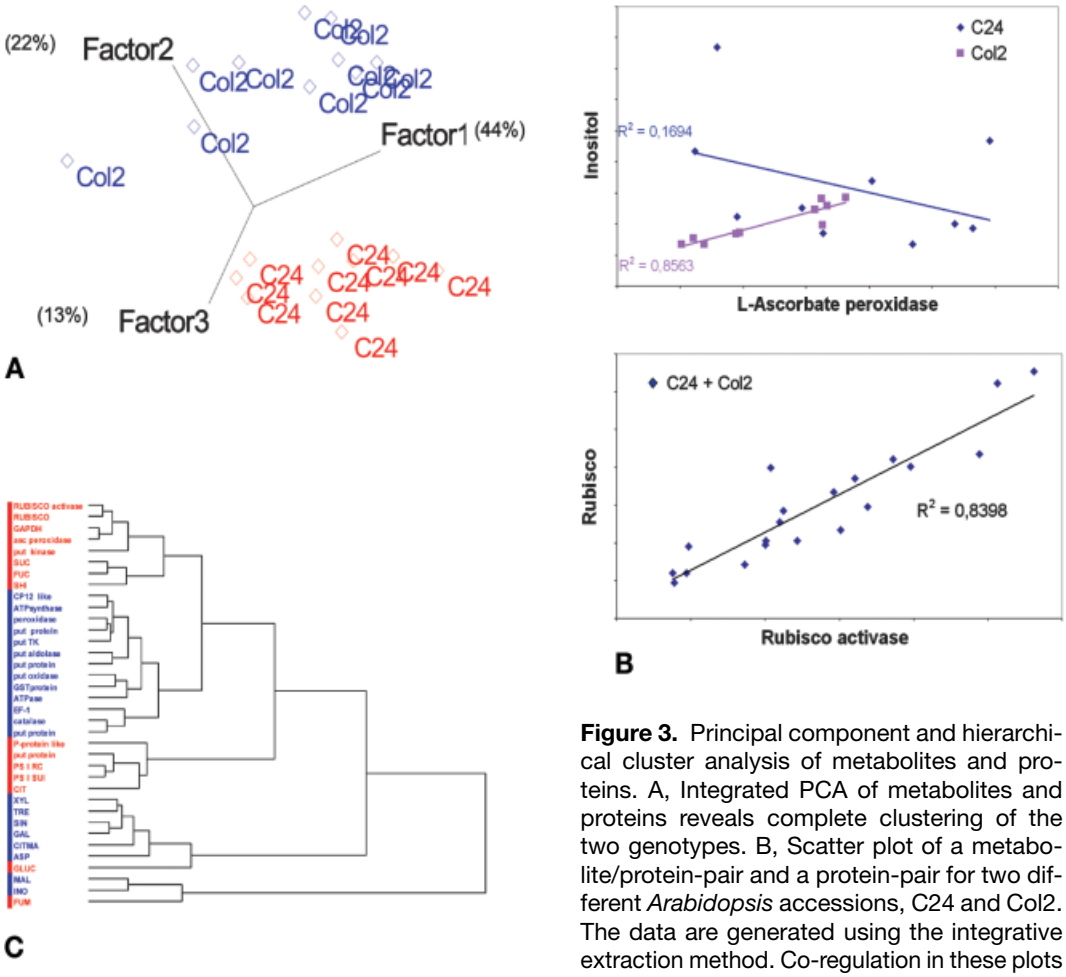

Figure 3. Principal component and hierarchical cluster analysis of metabolites and proteins. A, Integrated PCA of metabolites and proteins reveals complete clustering of the two genotypes. B, Scatter plot of a metabolite/protein-pair and a protein-pair for two different *Arabidopsis* accessions, C24 and Col2. The data are generated using the integrative extraction method. Co-regulation in these plots is defined as a linear correlation depending on the correlation coefficient (R^2). C, HCA of a merged metabolite-protein dataset (for details see text). Abbreviations of proteins and metabolites: RUBISCO activase: ribulose-1,5-bisphosphate carboxylase/oxygenase activase (S04048); RUBISCO: ribulose-1,5-bisphosphate carboxylase/oxygenase (NP_051067); GAPDH: glyceraldehyde-3-phosphate dehydrogenase (AAD10209); asc peroxidase: L-ascorbate peroxidase (S20866); put kinase: protein kinase, putative (At3g24550); SUC: sucrose; FUC: fucose; SHI: shikimate; CP12 like: CP12 protein precursor-like protein (At3g62410); ATPsynthase: ATP synthase CF1 beta chain (NP_051066); peroxidase: peroxidase, putative (At3g49120); put protein: protein, putative (At3g63190); put TK: transketolase-like protein (At3g60750), put aldolase: putative fructose-bisphosphate aldolase (AF428455_1); put oxidase: glycolate oxidase (At3g14420); GST protein: spindly (gibberellin signal transduction protein) (At3g11540); ATPase: ATPase alpha subunit (NP_051044); EF-1: translation elongation factor eEF-1 alpha chain (gene A4) (S08534); catalase: catalase (AAB07026); put protein: protein, putative (At3g47140); P-protein like: (At4g33010); put protein: protein, putative (At3g57190); PS I RC: putative photosystem I reaction center subunit II precursor (At1g03130); PS I SUI: photosystem I subunit III precursor (CAB52747); CIT: citrate; XYL: xylose; TRE: trehalose; SIN: sinapinate; GAL: galacturonate; CITMA: citramalate; ASP: aspartate; GLUC: gluconate; MAL: malate; INO: inositol; FUM: fumarate.

bolites according to Fiehn *et al.* [6]. We calculated the ratio of standard deviation to the mean for every variable, metabolite, and protein. These ratios appeared not to be correlated to the means ($r_{metabolites}$ = 0.38 and $r_{proteins}$ = 0.23), indicating that the relative variation of these compounds does not depend on their abundance. This is indicative of high biological variation among independent samples, even samples collected from tissues at seemingly identical developmental stage and grown under highly controlled environmental conditions.

In Fig. 3C, a hierarchical analysis of the set of quantified proteins and metabolites is shown. C24 and Col2 datasets are merged together to detect biochemical patterns conserved for both genotypes. A strongly conserved pattern for both varieties is detected for Calvin cycle enzymes such as RUBISCO and 3-glyceraldehyde dehydrogenase (GAPDH), which is in agreement with the literature [21]. Metabolites included in this cluster are sucrose and fucose suggesting the coordination of sucrose synthesis and degradation and photosynthetic activity. Surprisingly, ascorbate peroxidase is integrated into the Calvin cycle/sucrose cluster giving hints to the connectivity of the oxidative state and carbohydrate metabolism in plants. Inside the metabolite cluster, biochemically related structures such as malate and fumarate and carbohydrates form subclusters as expected.

These observed correlative patterns of metabolites and proteins suggest that connectivities with the underlying biochemical network can be adopted [17] but are not straightforwardly derivable from real biochemical networks at a systems level, pinpointing the need for extending these comprehensive studies. Additionally, the analysis of theoretical network topologies gives many hints about evolutionary relationships and regulation in biochemical networks [22–24] and delivers models which can be compared with the experimental network topologies discussed in this work. Many of the metabolites and proteins have unknown or putative structures and functions. Using a more comprehensive dataset and ultimately including quantitative transcript expression data, the integrative extraction protocol has the potential to unravel relationships among these compounds and to assign their linkage to already known functions in biochemical networks.

We thank Megan McKenzie for revising the manuscript.

Received December 23, 2002
Revised March 8, 2003
Accepted May 23, 2003

References

[1] Boyes, D. C., Zayed, A. M., Ascenzi, R., McCaskill, A. J. *et al.*, *Plant Cell* 2001, *13*, 1499–1510.

[2] McLafferty, F. W., Stauffer, D. A., Loh, S. Y., Wesdemiotis, C. *et al.*, *J. Am. Soc. Mass Spectrom.* 1999, *10*, 1229–1240.

[3] Wolters, D. A., Washburn, M. P., Yates, J. R., *Anal. Chem.* 2001, *73*, 5683–5690.

[4] Chelius, D., Bondarenko, P., *J. Proteome Res.* 2002, *1*, 317–323.

[5] Bondarenko, P., Chelius, D., Shaler, T., *Anal. Chem.* 2002, *74*, 4741–4749.

[6] Fiehn, O., Kopka, J., Dormann, P., Altmann, T. *et al.*, *Nat. Biotechnol.* 2000, *18*, 1157–1161.

[7] Hall, R., Beale, M., Fiehn, O., Hardy, N. *et al.*, *Plant Cell* 2002, *14*, 1437–1440.

[8] Beale, M., Dupree, P., Lilley, K., Beynon, J. *et al.*, *Trends Plant Sci.* 2002, *7*, 145–147.

[9] Marcotte, E. M., *Nat. Biotechnol.* 2001, *19*, 626–627.

[10] Oliver, S. G., Winson, M. K., Kell, D. B., Baganz, F. *et al.*, *Trends Biotechnol.* 1998, *16*, 373–378.

[11] Weckwerth, W., Fiehn, O., *Curr. Opin. Biotechnol.* 2002, *13*, 156–160.

[12] Weckwerth, W., Tolstikov, V., Fiehn, O., *Proc. 49th ASMS Conf. on Mass Spectrometry and Allied Topics*, ASMS, Chicago 2001, pp. 1–2.

[13] Washburn, M. P., Wolters, D., Yates, J. R., *Nat. Biotechnol.* 2001, *19*, 242–247.

[14] Koller, A., Washburn, M. P., Lange, B. M., Andon, N. L. *et al.*, *Proc. Natl. Acad. Sci. USA* 2002, *99*, 11969–11974.

[15] Pang, J. X., Ginanni, N., Dongre, A. R., Hefta, S. A. *et al.*, *J. Proteome Res.* 2002, *1*, 161–169.

[16] Yang, H., Haddad, H., Tomas, C., Alsaker, K. *et al.*, *Proc. Natl. Acad. Sci. USA* 2003, *100*, 1122–1127.

[17] Steuer, R., Kurths, J., Fiehn, O., Weckwerth, W. *et al.*, *Bioinformatics* 2003, 1019–1026.

[18] Kell, D., Mendes, P., in: Cornish-Bowden, A. J., Cardenas, M. L. (Eds.), *Technical and Medical Implications of Metabolic Control Analysis*, Kluwer, Netherlands 2000, pp. 2–25.

[19] Vance, W., Arkin, A., Ross, J., *Proc. Natl. Acad. Sci. USA* 2002, *99*, 5816–5821.

[20] Arkin, A., Shen, P. D., Ross, J., *Science* 1997, *277*, 1275–1279.

[21] Krapp, A., Hofmann, B., Schafer, C., Stitt, M. *et al.*, *Plant J.* 1993, *3*, 817–828.

[22] Ravasz, E., Somera, A., Mongru, D., Oltvai, Z. *et al.*, *Science* 2002, *297*, 1551–1555.

[23] Schuster, S., Hilgetag, C., Woods, J. H., Fell, D. A. *et al.*, *J. Math. Biol.* 2002, *45*, 153–181.

[24] Stelling, J., Klamt, S., Bettenbrock, K., Schuster, S. *et al.*, *Nature* 2002, *420*, 190–193.

[25] Cooper, M., Capman, S., Podlich, D., Hammer, G. *et al.*, *In Silico Biol.* 2002, *2*, 151–164.

PROTEOMICS

Supporting Information for Proteomics

DOI 10.1002/pmic.200200500

Wolfram Weckwerth, Kathrin Wenzel and Oliver Fiehn

Process for the integrated extraction, identification and quantification of metabolites, proteins and RNA to reveal their coregulation in biochemical networks

Supplementary Table 1. Identified metabolites of known structure

A list of all detected metabolites including classification of unknowns can be found on http://www.mpimp-golm.mpg.de/fiehn/index-e.html

Hydrophilic metabolites	Lipophilic metabolites
aconitate	2-Hydroxy-24:0 fatty acid
alanine	2-Hydroxy-C16:0 fatty acid
alpha-ketoglutarate	2-Hydroxy-C22:0 fatty acid
aminomalonate	2-Hydroxy-C24:1 fatty acid
arabinose	3-hydroxy-C16:0 fatty acid
arginine	3-hydroxy-C8:0 fatty acid
ascorbate	alpha-tocopherol
asparagine	benzylalcohol
aspartate	beta-sitosterol
benzoic acid	C10:0 fatty acid
benzoic acid, 3-hydroxy	C12:0 fatty acid
benzoic acid, 4-hydroxy	C12:0 fatty alcohol
beta-alanine	C14:0 fatty acid
butanoic acid trihydroxy, lactone	C14:0 fatty acid
butanoic acid, dihydroxy	C14:0 fatty alcohol
butyric acid, 4-amino	C15:0 fatty acid
citramalate	C16:0 fatty acid
citrate	C16:0 fatty alcohol
citrulline	C16:1 fatty acid
cycloserine	C17:0 fatty acid
erythritol	C18:0 fatty acid
erythrose	C18:0 fatty alcohol
ethanolamine	C18:1 cis6 fatty acid
fructose	C18:1 cis9 fatty acid
fructose-6-phosphate	C18:2 cis9,12 fatty acid
fumarate	C18:2 fatty acid
galactonate	C18:3 cis6,9,12 fatty acid
galactose	C18:3 cis9,12,15 fatty acid
galactose-6-phosphate	C20:0 fatty acid
galacturonic acid	C20:0 fatty alcohol
gluconic acid	C22:0 fatty acid
glucose	C24:0 fatty acid
glucose, 6-deoxy	C26:0 fatty acid
glucose-6-phosphate	C26:0 fatty alcohol
glutamate	C28:0 fatty acid
glutamine	C28:0 fatty alcohol
glyceric acid 2-phosphate	C30:0 fatty alcohol
glycerol	C6:0 fatty acid
glycerol 3-phosphate	C6:0 fatty alcohol
glycine	C8:0 fatty acid
glycolate	C8:0 fatty alcohol
homoserine	campesterol
indole-3-acetonitrile	cholesterol
inositol	cyclohexen-1-on
inositol-2-phosphate	decane
isoleucine	ergosterol
itaconate	sitosterol
kojic acid	stigmasterol
lactate	tetradecane
leucine	
lysine	
lyxitol	
lyxose	
malate	
maleate	
malonate	
maltose	
mannitol	
mannose	
melibiose	
methylergosterol	
nicotinate	
nicotinate	
nicoturic acid, 6-hydroxy	
nigerose	
norvaline	
ononitol	
ornithine	
oxoproline	
phenylalanine	
phosphate	
pipecolate	
proline	
psicose	
putrescine	
pyridine,2-Hydroxy	
pyridine,3-Hydroxy	
pyridine,4-Hydroxy	
pyruvate	
quinic acid	
raffinose	
rhamnose	
ribitol	
ribose	
salicylate	
serine	
serine O-acetyl	
serine, N-acetyl	
shikimate	
sinapinic acid	
sorbitol	
spermidine	
succinate	
sucrose	
tagatose	
tetronic acid	
threitol	
threonine	
trehalose	
tryptophane	
tyrosine	
urea	
uric acid	
valine	
xylitol	
xylose	

Supplementary Table 2. Identified proteins

ProteinID	Function	Peptides identified	Found in
NP_051067	riblose 1,5-bisphosphate carboxylase/oxygenase large chain	23	Col2/C24
NP_051066	ATP synthase CF1 beta chain	10	Col2/C24
S04048	ribulose-bisphosphate carboxylase activase (EC 6.3.4.-) precursor	8	Col2/C24
NP_051044	ATPase alpha subunit	13	Col2/C24
AAL38341	chlorophyll a/b-binding protein	5	Col2/C24
CAB40384	16 kDa polypeptide of oxygen-evolving complex	9	Col2/C24
At3g01500	carbonic anhydrase, chloroplast precursor	8	Col2/C24
At3g60750	transketolase - like protein	6	Col2/C24
At4g10340	light-harvesting chlorophyll a/b binding protein	7	Col2/C24
GCST_MESCR	Aminomethyltransferase, mitochondrial precursor (Glycine cleavage system T protein) (GCVT)	7	Col2/C24
AAD10209	glyceraldehyde 3-phosphate dehydrogenase A subunit	4	Col2/C24
At2g30860	glutathione transferase, putative	4	Col2/C24
At1g06680	photosystem II oxygen-evolving complex 23 (OEC23)	5	Col2/C24
AT5g25980	myrosinase TGG2	5	Col2/C24
At5g35630	glutamate-ammonia ligase (EC 6.3.1.2) precursor, chloroplast	4	Col2/C24
S08534	translation elongation factor eEF-1 alpha chain (gene A4)	3	Col2/C24
At3g55800	sedoheptulose-bisphosphatase precursor	3	Col2/C24
At1g03130	putative photosystem I reaction center subunit II precursor	4	Col2/C24
AAM12979	chlorophyll a/b-binding protein CP29	4	Col2/C24
At3g26060	putative peroxiredoxin	4	Col2/C24
NP_051054	photosystem II protein D2	3	Col2/C24
At2g40840	glycosyl hydrolase family 77 (4-alpha-glucanotransferase)	1	Col2/C24
BAB08951	2-cys peroxiredoxin-like protein	3	Col2/C24
AAM62639	unknown	2	Col2/C24
At1g12900	putative calcium-binding protein, calreticulin	1	Col2/C24
At1g44575	photosystem II 22kDa protein, putative	4	Col2/C24
At4g05180	oxygen-evolving complex protein 16, chloroplast precursor (OEC16)	4	Col2/C24
NP_051072	cytochrome f	3	Col2/C24
AF217459_1	heat shock protein 70	2	Col2/C24
At1g64290	hypothetical protein	1	Col2/C24
At5g26000	glycosyl hydrolase family 1, myrosinase precursor	3	Col2/C24
At3g57190	putative protein	1	Col2/C24
AAK64040	unknown protein	4	Col2/C24
AF428455_1	putative fructose-bisphosphate aldolase	1	Col2/C24
At1g32060	phosphoribulokinase precursor	2	Col2/C24
BAA20945	beta subunit of coupling factor one	2	Col2/C24
At2g39730	auxin-regulated protein	1	Col2/C24
At1g20020	ferredoxin--NADP reductase precursor, putative	3	Col2/C24
At3g08940	putative chlorophyll a/b-binding protein	3	Col2/C24
NP_051055	photosystem II 44 kDa protein	3	Col2/C24
At3g03530	expressed protein, supported by cDNA: gi_14335155	1	Col2/C24
At2g14380	putative retroelement pol polyprotein	3	Col2/C24
At3g26490	non-phototropic hypocotyl, putative	1	Col2/C24
At5g66520	selenium-binding protein-like	1	Col2/C24
CAB52747	photosystem I subunit III precursor	2	Col2/C24
AAK68813	H+-transporting ATP synthase-like protein	2	Col2/C24
At3g22520	unknown protein	1	Col2/C24
At3g63140	mRNA binding protein precursor - like	1	Col2/C24
At1g79040	photosystem II polypeptide, putative	2	Col2/C24
At5g20290	putative protein	1	Col2/C24
At1g49290	hypothetical protein	1	Col2/C24
At5g13160	protein kinase-like	1	Col2/C24
At5g24630	unknown protein	1	Col2/C24
At1g36990	hypothetical protein	1	Col2/C24
A96754	Similar to part of disease resistance protein [imported]	1	Col2/C24
T05822	hypothetical protein T5K18.170	1	Col2/C24
At3g57330	potential calcium-transporting ATPase 11, plasma membrane-type (Ca2+-ATPase, isoform 11)	1	Col2/C24
At4g25430	hypothetical protein	1	Col2/C24
At1g16240	expressed protein	1	Col2/C24
At1g48280	expressed protein	1	Col2/C24
AAM62447	glycine-rich RNA binding protein 7	1	Col2/C24
BAB08888	gene_id:MIJ24.6~ref|NP_013897.1~similar to unknown protein	1	Col2/C24

Supplementary Table 2. Continued

AAM62639	unknown	2	Col2/C24
At1g12900	putative calcium-binding protein, calreticulin	1	Col2/C24
At1g44575	photosystem II 22kDa protein, putative	4	Col2/C24
At4g05180	oxygen-evolving complex protein 16, chloroplast precursor (OEC16)	4	Col2/C24
NP_051072	cytochrome f	3	Col2/C24
AF217459_1	heat shock protein 70	2	Col2/C24
At1g64290	hypothetical protein	1	Col2/C24
At5g26000	glycosyl hydrolase family 1, myrosinase precursor	3	Col2/C24
At3g57190	putative protein	1	Col2/C24
AAK64040	unknown protein	4	Col2/C24
AF428455_1	putative fructose-bisphosphate aldolase	1	Col2/C24
At1g32060	phosphoribulokinase precursor	2	Col2/C24
BAA20945	beta subunit of coupling factor one	2	Col2/C24
At2g39730	auxin-regulated protein	1	Col2/C24
At1g20020	ferredoxin--NADP reductase precursor, putative	3	Col2/C24
At3g08940	putative chlorophyll a/b-binding protein	3	Col2/C24
NP_051055	photosystem II 44 kDa protein	3	Col2/C24
At3g03530	expressed protein, supported by cDNA: gi_14335155	1	Col2/C24
At2g14380	putative retroelement pol polyprotein	3	Col2/C24
At3g26490	non-phototropic hypocotyl, putative	1	Col2/C24
At5g66520	selenium-binding protein-like	1	Col2/C24
CAB52747	photosystem I subunit III precursor	2	Col2/C24
AAK68813	H+-transporting ATP synthase-like protein	2	Col2/C24
At3g22520	unknown protein	1	Col2/C24
At3g63140	mRNA binding protein precursor - like	1	Col2/C24
At1g79040	photosystem II polypeptide, putative	2	Col2/C24
At5g20290	putative protein	1	Col2/C24
At1g49290	hypothetical protein	1	Col2/C24
At5g13160	protein kinase-like	1	Col2/C24
At5g24630	unknown protein	1	Col2/C24
At1g36990	hypothetical protein	1	Col2/C24
A96754	Similar to part of disease resistance protein [imported]	1	Col2/C24
T05822	hypothetical protein T5K18.170	1	Col2/C24
At3g57330	potential calcium-transporting ATPase 11, plasma membrane-type (Ca2+-ATPase, isoform 11)	1	Col2/C24
At4g25430	hypothetical protein	1	Col2/C24
At1g16240	expressed protein	1	Col2/C24
At1g48280	expressed protein	1	Col2/C24
AAM62447	glycine-rich RNA binding protein 7	1	Col2/C24
BAB08888	gene_id:MIJ24.6~ref|NP_013897.1~similar to unknown protein	1	Col2/C24
At3g14420	glycolate oxidase, putative	1	Col2/C24
AAM65044	60S acidic ribosomal protein P2	1	Col2/C24
At2g29450	glutathione transferase (103-1A)	1	Col2/C24
AAM98072	unknown protein	1	Col2/C24
AF428455_1	putative fructose-bisphosphate aldolase	1	Col2/C24
At3g50820	photosystem II oxygen-evolving complex 33 (OEC33)	1	Col2/C24
At3g14415	glycolate oxidase	1	Col2/C24
At2g20230	expressed protein	1	Col2/C24
At2g13360	alanine-glyoxylate aminotransferase	1	Col2/C24
At5g42650	allene oxide synthase	1	Col2/C24
At2g05100	light-harvesting chlorophyll a/b binding protein	1	Col2/C24
At5g38750	putative protein	1	Col2/C24
At3g44890	RP19 gene for chloroplast ribosomal protein CL9	1	Col2/C24
At3g45590	putative protein	1	Col2/C24
At5g56810	F-box protein	1	Col2/C24
At1g69070	hypothetical protein	1	Col2/C24
At2g15325	hypothetical protein	1	Col2/C24
At3g63190	putative protein	1	Col2/C24
At1g13800	hypothetical protein	1	Col2/C24
At3g30843	hypothetical protein	1	Col2/C24
S49030	RNA-binding protein RNP-D precursor	1	Col2/C24
AAM66135	unknown	1	Col2/C24
At4g37460	putative protein	1	Col2/C24
At5g44870	disease resistance protein (TIR-NBS-LRR class), putative	1	Col2/C24

Supplementary Table 2. Continued

At2g47610	60S ribosomal protein L7A	1	Col2/C24
At5g25590	putative protein	1	Col2/C24
AAM63618	putative rubisco subunit binding-protein alpha subunit	1	Col2/C24
At5g23700	putative protein	1	Col2/C24
NP_051045	ATP synthase CF0 B chain	1	Col2/C24
At3g23400	expressed protein	1	Col2/C24
At2g40630	expressed protein	1	Col2/C24
At4g22780	Translation factor EF-1 alpha - like protein	1	Col2/C24
At1g04800	unknown protein	1	Col2/C24
At1g20060	kinesin-related protein	1	Col2/C24
At5g60120	APETALA2 protein - like	1	Col2/C24
At2g01620	expressed protein	1	Col2/C24
At2g27000	cytochrome p450 family	1	Col2/C24
At5g17370	hypothetical protein	1	Col2/C24
At5g09660	microbody NAD-dependent malate dehydrogenase	1	Col2/C24
At4g31050	putative protein	1	Col2/C24
At2g37310	hypothetical protein	1	Col2/C24
At5g49120	putative protein	1	Col2/C24
At3g24550	protein kinase, putative	1	Col2/C24
At2g26940	putative C2H2-type zinc finger protein	1	Col2/C24
T51531	biotin carboxyl carrier protein homolog T20K14.140 [imported]	1	Col2/C24
At5g16860	putative protein	1	Col2/C24
BAB10393	contains similarity to En/Spm-like transposon	1	Col2/C24
At4g18820	putative protein	1	Col2/C24
At5g01730	putative protein	1	Col2/C24
At1g80910	myrosinase precursor, putative	1	Col2/C24
At2g05170	expressed protein	1	Col2/C24
At5g42920	putative protein	1	Col2/C24
At5g51200	putative protein	1	Col2/C24
At3g53720	putative protein	1	Col2/C24
At5g22450	putative protein	1	Col2/C24
At1g22410	3-deoxy-D-arabino-heptulosonate 7-phosphate, putative	1	Col2/C24
At4g30990	putative protein	1	Col2/C24
At3g20860	putative serine/threonine protein kinase	1	Col2/C24
At3g05470	unknown protein	1	Col2/C24
At1g06380	hypothetical protein	1	Col2/C24
At3g47140	putative protein	1	Col2/C24
At5g01630	putative protein	1	Col2/C24
At5g39960	putative protein	1	Col2/C24
At2g35300	similar to late embryogenesis abundant proteins	1	Col2/C24
At4g30830	putative protein	1	Col2/C24
At1g68940	hypothetical protein	1	Col2/C24
At1g79680	WAK-like kinase (WLK)	1	Col2/C24
At3g66658	betaine aldehyde dehydrogenase, putative	1	Col2/C24
At4g19320	hypothetical protein	1	Col2/C24
At3g49350	GTPase activating -like protein	1	Col2/C24
At1g16140	WAK-like kinase (WLK)	1	Col2/C24
At3g04740	hypothetical protein	1	Col2/C24
At3g60890	putative protein	1	Col2/C24
At2g07010	putative retroelement pol polyprotein	1	Col2/C24
At5g02060	putative protein	1	Col2/C24
At3g60310	putative protein	1	Col2/C24
AAA32797	geranylgeranyl pyrophosphate synthase	1	Col2/C24
BAB09274	histidine kinase-like protein	1	Col2/C24
At1g72500	hypothetical protein	1	Col2/C24
At3g42320	putative protein	1	Col2/C24
H86321	hypothetical protein F6A14.10 [imported]	1	Col2/C24
At5g58980	random slug protein - like	1	Col2/C24
At5g14350	putative protein	1	Col2/C24
At4g27720	putative protein	1	Col2/C24
S20866	L-ascorbate peroxidase (EC 1.11.1.11) precursor	3	Col2/C24
At3g62410	CP12 protein precursor-like protein	1	Col2/C24
At3g49120	peroxidase, putative	1	Col2/C24

Supplementary Table 2. Continued

At3g11540	spindly (gibberellin signal transduction protein)	1	Col2/C24
AAB07026	catalase	3	Col2/C24
At4g33010	P-Protein - like protein	3	Col2/C24
RBS1_ARATH	Ribulose bisphosphate carboxylase small chain 1A, chloroplast precursor (RuBisCO small subunit 1A)	6	Col2/C24
At3g45140	lipoxygenase AtLOX2	7	Col2
S11852	photosystem II oxygen-evolving complex protein 1 precursor	8	Col2
At5g54270	light-harvesting chlorophyll a/b binding protein, putative	3	Col2
JQ1286	glyceraldehyde-3-phosphate dehydrogenase (NADP) (phosphorylating) (EC 1.2.1.13) B precursor, chloropla	8	Col2
At1g56190	phosphoglycerate kinase, putative	7	Col2
At3g04120	glyceraldehyde-3-phosphate dehydrogenase C subunit (GapC)	3	Col2
At2g02930	glutathione transferase, putative	5	Col2
AAA50156	carbonic anhydrase	5	Col2
AT4g37930	glycine hydroxymethyltransferase-like protein	5	Col2
At4g04640	coded for by A. thaliana cDNA AA041141	5	Col2
T52072	hypothetical protein g5bf [imported]	5	Col2
NP_051058	photosystem I P700 apoprotein A2	5	Col2
AT3g48870	AtClpC endopeptidase Clp ATP-binding chain C	5	Col2
T12970	hypothetical protein T6H20.190	3	Col2
A96602	elongation factor EF-2 [imported]	5	Col2
At2g39730	Rubisco activase	3	Col2
CAA70862	ferredoxin-dependent glutamate synthase	4	Col2
AF326861_1	putative photosystem I subunit PSI-E	3	Col2
AAN31836	putative 5-methyltetrahydropteroyltriglutamate-homocysteine S-methyltransferase	4	Col2
AT4g38970	putative fructose-bisphosphate aldolase	2	Col2
T52314	chlorophyll a/b-binding protein Lhcb6 [imported]	4	Col2
At2g35370	glycine decarboxylase complex H-protein	2	Col2
NP_051084	photosystem II 47 kDa protein	2	Col2
At4g13400	putative protein	1	Col2
AAN31832	putative chloroplast translation elongation factor EF-Tu precursor	3	Col2
At5g38410	ribulose bisphosphate carboxylase small chain 3b precursor	3	Col2
AAA32813	plasma membrane proton pump H+ ATPase	3	Col2
NP_051039	photosystem II protein D1	3	Col2
At4g18480	protein ch-42 precursor, chloroplast	3	Col2
S16582	fructose-bisphosphatase (EC 3.1.3.11) precursor, chloroplast	4	Col2
At5g47210	putative protein	2	Col2
JT0901	chaperonin 60 beta precursor	3	Col2
CAA74895	ribosomal protein L4	1	Col2
At1g76030	vacuolar ATP synthase subunit B	2	Col2
S33707	DNA-damage repair protein DRT112 precursor	1	Col2
AAN31859	putative heat shock protein 81-2 (HSP81-2)	3	Col2
AAM63250	cyanate lyase	1	Col2
AF360195_1	putative alanine aminotransferase	2	Col2
At1g53310	phosphoenolpyruvate carboxylase 1, putative	2	Col2
At2g26080	putative glycine dehydrogenase	2	Col2
At1g04410	putative malate dehydrogenase	2	Col2
At1g50250	chloroplast FtsH protease	2	Col2
G84888	probable transketolase precursor [imported]	2	Col2
AAK73957	putative ftsH chloroplast protease	3	Col2
AF083913_1	annexin	2	Col2
At1g77160	hypothetical protein	1	Col2
At3g17820	glutamine synthetase, putative	2	Col2
C84582	hypothetical protein At2g19880 [imported]	1	Col2
At2g41560	potential calcium-transporting ATPase 4, plasma membrane-type (Ca2+-ATPase, isoform 4)	2	Col2
AAM62764	glutamine synthetase, putative	2	Col2
At1g24706	F5A9.21, unknown	1	Col2
At2g22010	unknown protein	1	Col2
At5g14740	carbonic anhydrase 2	1	Col2
At1g76180	dehydrin, putative	2	Col2
AF428455_1	putative fructose-bisphosphate aldolase	1	Col2
At4g31700	ribosomal protein S6 - like	1	Col2
At1g74060	putative 60S ribosomal protein L6	1	Col2
At3g47070	putative protein	2	Col2
At3g58730	v-ATPase subunit D (vATPD)	2	Col2

Supplementary Table 2. Continued

At1g23740	putative auxin-induced protein	1	Col2
CAA11554	2-oxoglutarate dehydrogenase, E3 subunit	2	Col2
At2g21170	putative triosephosphate isomerase	2	Col2
At3g46520	actin 12	1	Col2
At1g19570	dehydroascorbate reductase, putative	2	Col2
AAM97062	unknown protein	1	Col2
S71112	catalase (EC 1.11.1.6) 3, peroxisome/glyoxysome location signal (S-[RKH]-L) motif	1	Col2
At4g13940	adenosylhomocysteinase	2	Col2
NP_051087	photosystem II phosphoprotein	1	Col2
At2g15620	ferredoxin--nitrite reductase	1	Col2
At3g47470	light-harvesting chlorophyll a/b binding protein	1	Col2
At4g35090	catalase 2	1	Col2
S19226	cold-regulated protein cor47	2	Col2
CAC35872	H+-transporting ATP synthase beta chain (mitochondrial)-like protein	2	Col2
AAB80700	glycolate oxidase	1	Col2
At4g09000	14-3-3 protein GF14 chi (grf1)	1	Col2
At2g34430	photosystem II type I chlorophyll a /b binding protein	1	Col2
At4g34870	peptidylprolyl isomerase (cyclophilin)	2	Col2
At1g16880	expressed protein	2	Col2
At2g37660	expressed protein	2	Col2
AF360228_1	putative glutathione reductase	1	Col2
At3g49910	60S ribosomal protein - like	2	Col2
At5g15980	putative protein	2	Col2
T52122	chaperonin 10	2	Col2
At3g07570	unknown protein	1	Col2
AAM13161	ATP-dependent transmembrane transporter, putative	1	Col2
AAB09585	ADP glucose pyrophosphorylase small subunit	2	Col2
At3g02090	putative mitochondrial processing peptidase	2	Col2
At4g03430	putative pre-mRNA splicing factor	1	Col2
At1g71240	hypothetical protein	1	Col2
At4g31700	ribosomal protein S6 - like	1	Col2
AAK59424	putative DEF (CLA1) protein	1	Col2
At5g55180	glycosyl hydrolase family 17	1	Col2
At1g74770	hypothetical protein	1	Col2
AC012394_17	putative phytochrome A signaling protein	1	Col2
At2g24820	putative Rieske iron-sulfur protein	1	Col2
At2g36380	ABC transporter family protein	1	Col2
At2g35120	glycine decarboxylase complex H-protein	1	Col2
AAL32516	putative protein	1	Col2
At1g50730	hypothetical protein	1	Col2
At2g14470	putative helicase	2	Col2
At1g67560	putative lipoxygenase	1	Col2
At1g56190	phosphoglycerate kinase	1	Col2
At4g12180	putative reverse transcriptase	1	Col2
At3g51560	disease resistance protein (TIR-NBS-LRR class), putative	1	Col2
At1g50120	hypothetical protein	2	Col2
G86301	probable retroelement polyprotein [imported]	1	Col2
At5g16500	protein kinase-like protein	1	Col2
At1g60860	GCN4-complementing protein	1	Col2
CAB80674	putative protein transport factor	1	Col2
At2g34610	hypothetical protein	1	Col2
T01733	hypothetical protein A_IG002N01.31	2	Col2
At2g07698	hypothetical protein	1	Col2
At5g10790	ubiquitin-specific protease 22 (UBP22)	1	Col2
At1g62810	amine oxidase, putative	1	Col2
AC069473_9	unknown protein	1	Col2
At1g73980	unknown protein	1	Col2
T50928	calmodulin-binding protein [imported]	1	Col2
AAM62795	60S ribosomal protein L27A	1	Col2
At5g37670	low-molecular-weight heat shock protein - like	1	Col2
At5g48010	pentacyclic triterpene synthase (04C11) (ATPEN1), putative	1	Col2
At2g43560	FKBP-type peptidyl-prolyl cis-trans isomerase	1	Col2
AC007354_10	Strong similarity to gb|Y09533 involved in starch metabolism from Solanum tuberosum	1	Col2

Supplementary Table 2. Continued

At1g13790	hypothetical protein	1	Col2
At2g19380	RRM-containing RNA-binding protein	1	Col2
At5g06240	unknown protein	1	Col2
At1g67240	mutator-like transposase, putative	1	Col2
CAA69802	ATPase subunit 1	1	Col2
At4g20890	tubulin beta-9 chain	1	Col2
At2g40590	40S ribosomal protein S26	1	Col2
BAA97188	emb|CAB87273.1~gene_id:MMI9.7~similar to unknown protein	1	Col2
At5g09860	expressed protein	1	Col2
At1g60630	leucine-rich repeat transmembrane protein kinase	1	Col2
At1g02500	s-adenosylmethionine synthetase	1	Col2
AF462865_1	unknown protein	1	Col2
AF424618_1	membrane-associated salt-inducible protein	1	Col2
At1g05530	UDP-glycosyltransferase family	1	Col2
At1g74680	Exostosin family	1	Col2
At1g32470	glycine cleavage system H protein precursor, putative	1	Col2
At2g47470	putative protein disulfide-isomerase	1	Col2
T48997	epsin-like protein	1	Col2
AAK96795	acyl carrier protein (ACP) gene	1	Col2
At2g16890	putative glucosyltransferase	1	Col2
AAL91646	unknown protein	1	Col2
At5g59660	leucine-rich repeat transmembrane protein kinase, putative	2	Col2
At5g66190	ferredoxin-NADP+ reductase	1	C24
At3g22910	potential calcium-transporting ATPase 13, plasma membrane-type (Ca2+-ATPase, isoform 13)	1	C24
At4g28750	photosystem I subunit PSI-E - like protein	1	C24
At5g48310	putative protein	1	C24
At5g28300	GTL1 - like protein	1	C24
At1g29930	light-harvesting chlorophyll a/b binding protein	1	C24
At5g09660	microbody NAD-dependent malate dehydrogenase	1	C24
At3g11820	syntaxin SYP121	1	C24
At5g40480	nuclear pore protein -like	1	C24
At2g35920	putative ATP-dependent RNA helicase A	1	C24
At1g22490	expressed protein	1	C24
At5g14070	glutaredoxin-like protein	1	C24
NP_051048	ribosomal protein S2	1	C24
At4g19750	glycosyl hydrolase family 18	1	C24
At1g25340	myb-related transcription factor (cpm7), putative	1	C24
At2g25140	HSP100/ClpB, putative	1	C24
At2g20960	pEARLI 4 protein	1	C24
At2g27480	putative calcium binding protein	1	C24
AAD03443	contains similarity to human RNA polymerase II complex component SRB7 (GB:U52960)	1	C24
At5g03940	signal recognition particle 54CP protein precursor	1	C24
At1g07430	protein phosphatase 2C (PP2C), putative	1	C24
At4g07960	putative glucosyltransferase	1	C24
At3g11630	putative 2-cys peroxiredoxin BAS1 precursor (thiol-specific antioxidant protein)	1	C24
At4g01310	putative L5 ribosomal protein	1	C24
At5g61250	glycosyl hydrolase family 79 (endo-beta-glucuronidase/heparanase)	1	C24
At1g65010	hypothetical protein	1	C24
At1g67810	unknown protein	1	C24
At5g50260	cysteine proteinase	1	C24
At5g64040	photosystem I reaction center subunit PSI-N precursor (PSI-N)	1	C24
At5g24770	vegetative storage protein Vsp2	1	C24
At4g28630	ABC transporter family protein	1	C24
At5g09730	glycosyl hydrolase family 3	1	C24
At4g31300	20S proteasome beta subunit A (PBA1);	1	C24
T05498	hypothetical protein T19K4.190	1	C24
AC000103_3	unknown protein	1	C24
At5g09700	beta-glucosidase - like protein	1	C24
At5g15200	40S ribosomal protein - like	1	C24
At1g72300	leucine-rich repeat transmembrane protein kinase, putative	1	C24
At3g14350	leucine-rich repeat transmembrane protein kinase, putative	1	C24
At3g13160	expressed protein	1	C24
At5g59660	leucine-rich repeat transmembrane protein kinase, putative	1	C24

Supplementary Table 2. Continued

At5g66190	ferredoxin-NADP+ reductase	1	C24
At3g22910	potential calcium-transporting ATPase 13, plasma membrane-type (Ca2+-ATPase, isoform 13)	1	C24
At4g28750	photosystem I subunit PSI-E - like protein	1	C24
At5g48310	putative protein	1	C24
At5g28300	GTL1 - like protein	1	C24
At1g29930	light-harvesting chlorophyll a/b binding protein	1	C24
At5g09660	microbody NAD-dependent malate dehydrogenase	1	C24
At3g11820	syntaxin SYP121	1	C24
At5g40480	nuclear pore protein -like	1	C24
At2g35920	putative ATP-dependent RNA helicase A	1	C24
At1g22490	expressed protein	1	C24
At5g14070	glutaredoxin-like protein	1	C24
NP_051048	ribosomal protein S2	1	C24
At4g19750	glycosyl hydrolase family 18	1	C24
At1g25340	myb-related transcription factor (cpm7), putative	1	C24
At2g25140	HSP100/ClpB, putative	1	C24
At2g20960	pEARLI 4 protein	1	C24
At2g27480	putative calcium binding protein	1	C24
AAD03443	contains similarity to human RNA polymerase II complex component SRB7 (GB:U52960)	1	C24
At5g03940	signal recognition particle 54CP protein precursor	1	C24
At1g07430	protein phosphatase 2C (PP2C), putative	1	C24
At4g07960	putative glucosyltransferase	1	C24
At3g11630	putative 2-cys peroxiredoxin BAS1 precursor (thiol-specific antioxidant protein)	1	C24
At4g01310	putative L5 ribosomal protein	1	C24
At5g61250	glycosyl hydrolase family 79 (endo-beta-glucuronidase/heparanase)	1	C24
At1g65010	hypothetical protein	1	C24
At1g67810	unknown protein	1	C24
At5g50260	cysteine proteinase	1	C24
At5g64040	photosystem I reaction center subunit PSI-N precursor (PSI-N)	1	C24
At5g24770	vegetative storage protein Vsp2	1	C24
At4g28630	ABC transporter family protein	1	C24
At5g09730	glycosyl hydrolase family 3	1	C24
At4g31300	20S proteasome beta subunit A (PBA1);	1	C24
T05498	hypothetical protein T19K4.190	1	C24
AC000103_3	unknown protein	1	C24
At5g09700	beta-glucosidase - like protein	1	C24
At5g15200	40S ribosomal protein - like	1	C24
At1g72300	leucine-rich repeat transmembrane protein kinase, putative	1	C24
At3g14350	leucine-rich repeat transmembrane protein kinase, putative	1	C24
At3g13160	expressed protein	1	C24
At5g59660	leucine-rich repeat transmembrane protein kinase, putative	1	C24
At3g05400	sugar transporter, putative	1	C24
At5g54290	cytochrome c biogenesis protein precursor (gb|AAF35369.1)	1	C24
At1g75350	chloroplast 50S ribosomal protein L31, putative	1	C24
At1g75350	chloroplast 50S ribosomal protein L31, putative	1	C24
At3g54890	light-harvesting chlorophyll a/b binding protein	1	C24
At5g47180	VAMP (vesicle-associated membrane protein)-associated protein-like	1	C24
At5g41610	Na+/H+ antiporter-like protein	1	C24
AC002423_15	unknown protein	1	C24
At5g58490	cinnamoyl-CoA reductase - like protein	1	C24
C86379	unknown protein	1	C24
At1g19640	S-adenosyl-L-methionine:jasmonic acid carboxyl methyltransferase (JMT)	1	C24
At5g04290	glycine-rich protein	1	C24
At1g35680	50S ribosomal protein L21 chloroplast precursor (CL21)	1	C24
NP_051097	ribosomal protein L22	1	C24
At5g48600	chromosome condensation protein	1	C24
At3g43190	sucrose synthase, putative	1	C24
At3g26790	transcriptional regulator (FUSCA3)	1	C24
At4g17300	asparagine--tRNA ligase	1	C24
At1g31000	hypothetical protein	1	C24
BAB02913	unknown protein	1	C24
AAK64154	unknown protein	1	C24
AAB61690	disease resistance protein homolog	1	C24

Supplementary Table 2. Continued

At2g24490	putative replication protein A1	1	C24
At5g66190	ferredoxin-NADP+ reductase	1	C24
At3g22910	potential calcium-transporting ATPase 13, plasma membrane-type (Ca2+-ATPase, isoform 13)	1	C24
At5g48310	putative protein	1	C24
At5g28300	GTL1 - like protein	1	C24
At1g29930	light-harvesting chlorophyll a/b binding protein	1	C24
At5g09660	microbody NAD-dependent malate dehydrogenase	1	C24
At3g11820	syntaxin SYP121	1	C24
At3g51570	disease resistance protein (TIR-NBS-LRR class), putative	1	C24
At3g46530	disease resistance protein, RPP13-like (CC-NBS class), putative	1	C24
At2g25710	biotin holocarboxylase synthetase	1	C24
At2g12150	Mutator-like transposase	1	C24
At5g32481	Athila retroelement ORF1, putative	1	C24
At3g24190	expressed protein	1	C24
At1g19390	WAK-like kinase (WLK)	1	C24
At4g22470	extensin - like protein	1	C24
At1g47560	hypothetical protein	1	C24
At2g16780	putative WD-40 repeat protein, MSI2	1	C24
At1g18030	protein phosphatase 2C (PP2C), putative	1	C24
At3g23790	AMP-binding protein, putative	1	C24
At1g66530	arginyl-tRNA synthetase	1	C24
At1g48150	MADS-box protein	1	C24
At4g14140	(cytosine-5-)-methyltransferase	1	C24
At1g62940	4-coumarate:coenzyme A ligase, putative	1	C24
At2g29500	putative small heat shock protein	1	C24
At3g07980	putative MAP3K epsilon protein kinase	1	C24
At4g23940	putative MAP3K epsilon protein kinase	1	C24
At1g21810	myosin-like protein	1	C24
At5g14950	glycosyl hydrolase family 38 (alpha-mannosidase)	1	C24
At3g53280	cytochrome P450 monooxygenase	1	C24
At2g41310	putative two-component response regulator 3 protein	1	C24
At1g50410	DNA-binding protein, putative	1	C24
At5g05340	peroxidase, putative	1	C24
A96721	probable peptide transporter	1	C24
At2g26790	putative salt-inducible protein	1	C24
At3g30570	putative reverse transcriptase	1	C24
At1g74080	putative transcription factor	1	C24
At3g21210	CHP-rich zinc finger protein, putative	1	C24
At2g42270	U5 small nuclear ribonucleoprotein helicase, putative	1	C24
At1g69320	CLE10, putative	1	C24
At3g42950	polygalacturonase, putative	1	C24