Differential metabolic networks unravel the effects of silent plant phenotypes

Wolfram Weckwerth*, Marcelo Ehlers Loureiro†, Kathrin Wenzel*, and Oliver Fiehn*‡

*Department of Molecular Physiology, Max Planck Institute of Molecular Plant Physiology, 14424 Potsdam, Germany; and †Departamento de Biologia Vegetal, Universidade Federal de Viçosa, 35670001, Viçosa, Brazil

Edited by Maarten Koornneef, Wageningen University and Research Centre, Wageningen, The Netherlands, and approved March 5, 2004 (received for review June 5, 2003)

Current efforts aim to functionally characterize each gene in model plants. Frequently, however, no morphological or biochemical phenotype can be ascribed for antisense or knock-out plant genotypes. This is especially the case when gene suppression or knock-out is targeted to isoenzymes or gene families. Consequently, pleiotropic effects and gene redundancy are responsible for phenotype resistance. Here, techniques are presented to detect unexpected pleiotropic changes in such instances despite very subtle changes in overall metabolism. The method consists of the relative quantitation of >1,000 compounds by GC/time-of-flight MS, followed by classical statistics and multivariate clustering. Complementary to these tools, metabolic networks are constructed from pair-wise analysis of linear metabolic correlations. The topology of such networks reflects the underlying regulatory pathway structure. A differential analysis of network connectivity was applied for a silent potato plant line suppressed in expression of sucrose synthase isoform II. Metabolic alterations could be assigned to carbohydrate and amino acid metabolism even if no difference in average metabolite levels was found.

metabolomics | metabonomics | data mining | regulatory networks | functional genomics

Silent phenotypes are genetically modified organisms that do not show obvious changes in morphology, yield, growth rates, or related parameters when compared with parental lines under given physiological conditions (1). This phenomenon is especially astonishing when genes are altered that are known to play pivotal roles in overall plant fitness (2). It is thought that such organisms might have found ways to circumvent the deleterious effects of the mutated genes or that redundancy in gene families (3) could prevent injurious outcomes. The most apparent form of gene redundancy is the frequent coexpression of enzyme isoforms involved in various metabolic pathways (4–6). In most cases, enzyme isoforms cannot be distinguished on the basis of enzyme activities. Here, techniques like metabolomics might aid functional characterization (7), assuming that the primary alteration of enzyme-encoding genes pleiotropically affects biochemical pathways. The working hypothesis is that a network of metabolic associations represents a snapshot response of the underlying biochemical network at a given biological situation, which then can be used to observe changes between different genotypes (8, 9). This view is theoretically supported by the concept of "metabolic control analysis" (10) and the concept of maximal connectivity in a biochemical network (11). Specifically, the effects on metabolite pool concentrations may be higher than the alteration in enzymatic flux control or enzyme activities. Recently, silent yeast phenotypes were discriminated from WT strains by using multivariate statistics applied to NMR (12) or MS (13) based metabolic fingerprints, with the objective to cluster genotypes together that were defective in genes with similar functions. However, resolution of the NMR and MS data prohibited the actual determination of individual metabolites to underpin the biochemical basis of the clustering results.

Here, we present an approach to the study of silent plant phenotypes. We propose to use unbiased metabolite detection concomitant with differential topology analysis of metabolic correlation networks as complementary tools to classical univariate and multivariate statistical tests. As a test case for a silent plant phenotype, an antisense potato plant line was used that was constitutively reduced in gene expression encoding for the sucrose synthase isoform II (SS2) under the control of the 35S promoter. The primary biochemical action of sucrose synthase is sucrose cleavage to UDP-glucose and fructose, which may also work bidirectionally *in vivo* in tubers (14). In general, several isoforms of sucrose synthase are known to play pivotal roles in plant development, carbon partitioning, phloem unloading, and sink strength (15–19). However, it is difficult to allocate specific functions of particular isoforms. When two genes encoding sucrose synthase are expressed in the same cell, the proteins form homo- or heterotetramers (20–22), suggesting that the isozymes are interchangeable in at least some cellular roles. This gene redundancy might lead to silent phenotypes if specific isoforms have decreased activities using the antisense approach. Specifically, no change in total enzyme activity can be anticipated. The objective of this study was now to investigate with metabolomic and statistical tools whether (*i*) the silent SS2 antisense genotype can be discriminated from the parental line and whether (*ii*) biochemical effects on primary metabolism can be found as an important classifier in discrimination.

Experimental Protocol

Thirty to 40 individual SS2 antisense plants were compared with the Desirée cv. background line. Northern blot RNA analysis was performed from 1 g fresh weight (FW) of leaf tissue by using a *Kpn*I fragment of SuSy 2 clone as probe. Plants were grown in controlled greenhouse conditions in a randomized plot. One-hundred-milligram FW tuber slices (5 mm i.d.) were harvested 3 mm below the peels, perpendicular to the main tuber axis. For leaves, 300 mg FW disks were sampled. Samples were extracted and fractionated into a polar and a lipophilic fraction as given in ref. 23. GC-time-of-flight (TOF) analysis was performed on an HP 5890 gas chromatograph with standard liners containing glass wool in split mode (1:25) at 230°C injector temperature, with a liner exchange for every 50 samples. The GC was operated at constant flow of 1 ml/min helium on a 40-m, 0.25-mm i.d., 0.25-μm RTX-5 column with 10-m integrated precolumn, a start temperature of 80°C, 2 min isothermal, temperature ramping by 15°C/min to 330°C, 6 min isothermal. Data were acquired on a Pegasus II TOF mass spectrometer (LECO, St. Joseph, MI) and CHROMATOF software (Fig. 6, which is published as supporting information on the PNAS web site) at 20 s^{-1} from m/z 85–500, $R = 1$, 70 kV electron impact (EI), and autotuning with reference gas CF43. Samples were compared against reference chromatograms that had a maximum of detectable peaks at $S/n > 20$. For identification and alignment, peaks

This paper was submitted directly (Track II) to the PNAS office.

Abbreviations: SS2, sucrose synthase isoform II; GC-TOF, GC/time-of-flight; DFA, discriminant function analysis; SPS, sucrose-6-phosphate synthase; GABA, γ-aminobutyric acid.

‡To whom correspondence should be addressed. E-mail: fiehn@mpimp-golm.mpg.de.

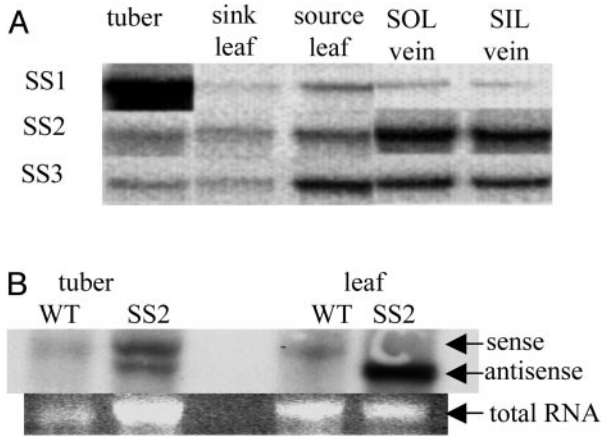

Fig. 1. RNA expression analysis of sucrose synthase isoforms in potato plants. (*A*) Tissue specificity of isoforms SS1, SS2, and SS3 using a 2.6-kb sense fragment. SOL, source leaf, SIL, sink leaf. (*B*) Hybridization of SS2-specific cDNA fragments in sense (2.6 kb) and antisense (2.1 kb) directions in a Northern blot from WT and SS2 antisense plants in comparison with the total amount of sample RNA applied.

were matched against a customized reference spectrum database, based on retention indices and mass spectral similarities. Relative quantification was performed on ion traces chosen by optimal selectivity from coeluting compounds. Tables on quantitative results were exported in ASCII format. Artifact peaks such as column bleeding, phtalates, and polysiloxanes were removed. All data were normalized to plant mg FW and to internal references (ribitol and nonadecanoic acid methyl ester). For all statistical tests, data were log-transformed, resulting in more Gaussian-type distributions (Fig. 7, which is published as supporting information on the PNAS web site). Significance levels for Pearson correlations r were computed depending on the number of metabolite pairs n found in the chromatograms by using $t = r\cdot(n-2)^{0,5}/(1-r^2)^{0,5}$ and controlled for a potential impact of outliers by robust fit assessments in MATLAB 6.5.0 (Mathworks, Natick, MA). Univariate ANOVA for production of box-whisker plots and F-statistics probabilities were also carried out in MATLAB. For multivariate statistics, empty cells were replaced by genotype means. The statistical package SAS/STAT 8.01 (SAS Institute, Cary, NC) was used to perform a normal kernel discriminant function analysis (DFA) method with equal bandwidth to estimate the group-specific probability density with $R = 0.4$. Principle component analysis was performed by using PIROUETTE 2.0 (Infometrix, Woodinville, WA) with mean-centering and unit variance scaling. Graphical visualization using the Fruchtermann-Reingold 2D layout algorithm was done by the PAJEK software package (http://vlado.fmf.uni-lj.si/pub/networks/pajek). For calculation of metabolite connectivity and changes in metabolite/metabolite ratios, lipophilic (secondary) metabolites were disregarded because biological regulatory mechanisms and turnover rates are known to be quite different compared with polar (primary) metabolites (24).

Results and Discussion

Sucrose Synthase II Antisense Potato Plants Have No Visible Phenotype. By Northern blot analysis using specific cDNAs for the sucrose synthase isoforms SS1, SS2, and SS3, it was shown that all sucrose synthase genes were expressed in all potato tissues, but that expression levels varied (Fig. 1*A*). The expression of SS2 was most specific for leaf veins, implying a role in carbon partitioning for long-distance transport. For this reason, we expected metabolic effects in mature leaves (as carbon source) and tubers (carbon sinks). Expression of the SS2 antisense construct was tissue specific. In leaf tissue, the sense SS2 band was completely suppressed concomitant with a strong antisense signal whereas in tuber the sense band was still expressed beside a strong antisense band (Fig. 1*B*). However, the total sucrose synthase activity was not found to be significantly decreased due to the coexpression of the other two enzyme isoforms. This phenomenon is often observed in plant molecular biology. In some cases, the total activity may even be found to have increased after suppression or knockout of a specific gene (25). In maize mutants, a loss of the sucrose synthase isoform *sus1* was reported to have no phenotypic effect but was associated with ectopic expression of the other gene isoform (*sh1* gene) complementing *sus1* (26).

Table 1. Biochemical and physiological comparison of WT and SS2 antisense plants

	WT	SS2
Leaf		
Morphology	Regular	Regular
Total sucrose synthase	92 ± 14	77 ± 13
SPS non-selective assay	382 ± 91	484 ± 66
SPS selective assay*	45 ± 4	152 ± 10
Fructobisphosphatase	4,177 ± 218	5,152 ± 518
Acid invertase	80 ± 30	110 ± 20
Alkaline invertase	65 ± 28	116 ± 11
Fructokinase	3,485 ± 218	3,883 ± 518
Tuber		
Morphology	Regular	Regular
Total sucrose synthase*	260 ± 8	303 ± 8
Yield: tuber number	12.8 ± 2.1	9.5 ± 0.8
Yield: tuber weight	22.6 ± 4.8	18.1 ± 1.8
Starch	309 ± 29	273 ± 14

Significant differences are highlighted by asterisks. Maximal catalytic activities for enzymes involved in carbohydrate metabolism are given as average values and SEs in nmol·min^{-1}·(g FW)$^{-1}$ for five individual plants per line. Morphology (see photographs in Fig. 8) and yield were assessed on eight plants per line, indicating the silence of the SS2 phenotype. Average tuber weight is given in g per tuber. Starch content is given in μM/g FW.

Consistent with this background knowledge, it was not expected to find a strong phenotype for the antisense expression of the SS2 isoform in potato. No significant change in plant morphology or development, or target metabolites such as starch, sucrose, glucose, or fructose contents was observed in comparison with the parental line (Table 1; see also Fig. 8, which is published as supporting information on the PNAS web site). Such a finding is called a "silent phenotype" in plant biotechnology. However, we assumed that the antisense expression of sucrose synthase should still have measurable effects on related key enzymes in carbohydrate metabolism that counteracted the primary effect of the altered biochemical network in a pleiotropic way. Based on the strength of the SS2 suppression in leaves, enzyme activity measurements focused on this organ. No significant difference was found for the acid and alkaline invertases, fructobisphosphatase, fructokinase, and the maximum catalytic activity V_{max} of sucrose-6-phosphate synthase (SPS). In contrast, SPS was found to be more activated in SS2 plants by using the selective SPS assay (V_{sel}) (27). This finding means that the SPS enzyme was less sensitive to inhibition by inorganic phosphate (the SPS allosteric inhibitor), which suggests a higher sucrose mobilization in SS2 plants. Intriguingly, the maximal catalytic activity of sucrose synthase was increased in the antisense tubers (Table 1), pointing to an increased sucrose influx. In the next sections, we describe how metabolic networks respond to differences in the SPS activation state.

Comprehensive GC-TOF Analysis Detects >1,000 Metabolites. In previous studies, metabolite profiling using GC/quadrupole MS resulted in the detection of >80 metabolites in potato tubers (23) and over 300 predefined compounds in leaf extracts (7). Although these data sets were sufficient to distinguish metabolic states of clearly

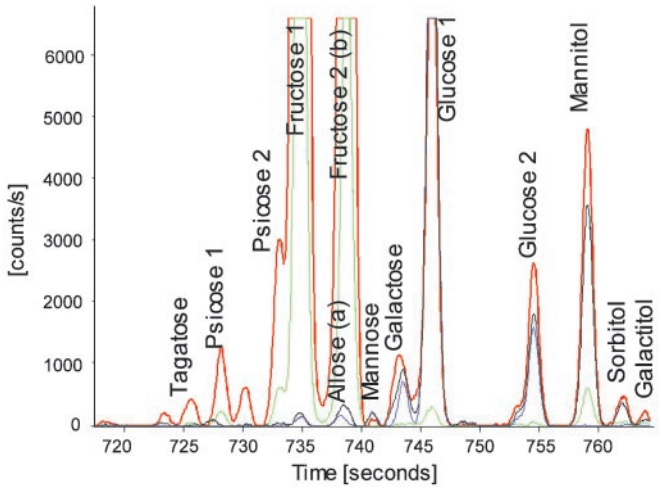

Fig. 2. GC-TOF analysis deconvoluting uncommon monosaccharides such as tagatose, psicose, and allose in polar extracts of potato leaves. Unique ion traces such as m/z 103 (red), 307 (green), 160 (blue), or 319 (black) are automatically selected for relative quantitation. Mass spectra for coeluting peaks a and b (allose and fructose) are given in Fig. 6.

different plant genotypes, they did not extract the full information contained in these chromatograms. To increase the efficiency of information extraction, GC/TOF MS (28) was applied concomitant with automated mass spectral deconvolution, identification, and export of metabolite peaks. Compared with quadrupole MS, GC-TOF data acquisition has the great advantage that far more spectra can be acquired across a chromatographic peak. Additionally, relative ion intensities in mass spectra remain constant over the chromatographic elution profile. Both properties largely enhance the suitability of (*i*) finding peaks in a completely unbiased manner, even for low abundant trace compounds, and (*ii*) deconvoluting coeluting mass spectra in complex chromatograms, in which very often more than two side peaks overlap with any specific metabolite. This MS deconvolution facilitates correct peak annotation by spectra purification.

When *Solanum tuberosum* plants were analyzed by GC-TOF and mass spectral deconvolution, over 1,200 peaks were found from the two fractions of a single leaf extract, and ≈600 peaks were detected in a typical potato tuber. Considering method artifacts and double peaks known to be formed for some metabolites, ≈1,000 leaf metabolites and 500 tuber metabolites can be estimated to be detectable by this method. The deconvolution power is demonstrated by detection of low abundant hexose isomers like psicose, tagatose, and allose (Fig. 2). Limits of mass spectral deconvolution were found for trace compounds that coeluted with 1,000-fold more concentrated peaks as given for the example of allose and fructose (Fig. 6), where allose reached a match factor of only 67% identity to the pure reference compound. On average, peaks with signal/noise ratio >100 resulted in 89% match factor for identification. Low abundant peaks with signal/noise ratio <25 had an average match factor of 66%.

Difference in Metabolite Mean Levels and Variances Were Too Small for Deriving Functional Information from Genotype Clustering. To find subtle differences in SS2 antisense plants in comparison with their parental background, >30 individual plants per line were analyzed, according to the central limit theorem (29). Average levels of 262 of 1,216 leaf metabolites were found to be different at a *t* test significance level of $P < 0.05$ (Table 2; see also Fig. 9, which is published as supporting information on the PNAS web site), ranging from 1.1- to 4.6-fold differences between the silent transgenic and its parental genotype. Intriguingly, no metabolite with

Table 2. Differences in metabolite levels between SS2 and WT plants for tubers and leaves

Tuber			Leaf		
Metabolite	*P* value	x-fold SS2/WT	Metabolite	*P* value	x-fold SS2/WT
Lysine	0.000018	1.5	Allose	0.028	1.9
GABA	0.002	1.4	Arabinitol	0.023	1.1
Glutamine	0.002	1.4	GABA	0.007	0.7
Homoserine	0.002	1.4	Galactonic acid	0.000013	0.8
Citrulline	0.004	1.3	Glycerol	0.037	0.8
Valine	0.004	1.3	Glycine	0.050	1.2
Lactate	0.005	0.8	Isoleucine	0.021	1.6
Tyrosine	0.008	1.3	Mannitol	0.012	0.7
Isoleucine	0.009	1.3	Ribose	0.008	0.8
Oxoproline	0.010	1.4	Shikimate	0.001	0.8
Phenylalanine	0.010	1.4	Threonine	0.010	0.7
Tryptophane	0.015	1.4	Trehalose	0.003	0.6
Malonate	0.017	0.7	Tyramine	0.008	0.6
Methionine	0.021	1.2			
Glycine	0.046	1.2			
Fructose-6P	0.048	0.7			
β-alanine	0.050	1.3			

Tuber: 656 metabolites tested; 34 reached a significance level of $P < 0.05$; 18 metabolites had $P < 0.01$. Leaf: 1,216 metabolites tested; 262 reached a significance level of $P < 0.05$; 98 metabolites had $P < 0.01$. Only polar metabolites with known chemical structure are shown. fa, fatty acid; falc, fatty alcohol.

known chemical structure changed >2-fold in SS2/WT comparisons. Only 15 metabolites had a $P_{leaf} < 0.000041$, when significance thresholds were lowered for reduced risks of false positive findings to an overall probability of 95% by Bonferroni estimation. Among these compounds, the only metabolite that could be identified was galactonic acid. The other 14 were therefore primary candidates for *de novo* structural elucidation; 7 of these were classified as sugar-related compounds by their corresponding mass spectra. Only one

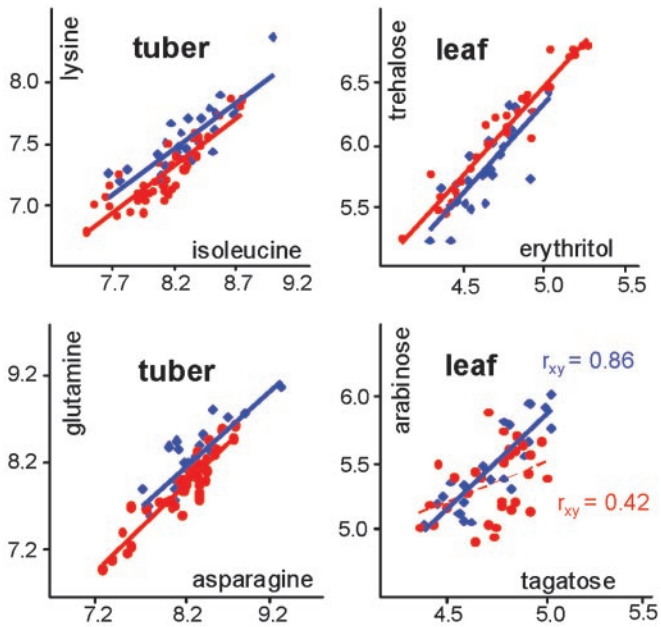

Fig. 3. Metabolite–metabolite correlations given in scatter plots. Blue, SS2; red, WT plants. Unitless relative metabolite peak areas (see *Experimental Protocol*) are given as log-scaled data. Differences in linear regression slopes are hence converted to differences in offsets ($P < 0.01$).

lipophilic compound was found among these hits, implying that leaf metabolism was more significantly affected for polar phase metabolites than for lipophilics.

For tubers, far fewer significant changes were found. Only 34 of 656 metabolites detected were found to be below the *t* test threshold level of $P < 0.05$ (among them 25 polar phase metabolites), and only lysine satisfied the Bonferroni criterion of $P'_{\text{tuber}} < 0.000076$. Mean values of metabolite levels directly involved in the enzymatic reaction of sucrose synthase, i.e., sucrose, glucose, or fructose, were not significantly changed in tubers or leaves.

Next, genotype discrimination was tested by clustering tools to find further evidence for biochemical alterations in the silent SS2 phenotype. According to its silent phenotype, SS2 tuber or leaf metabolic phenotypes were not distinguishable from WT plants by unsupervised principal component analysis (PCA) using the first two vectors, which explained 93–95% of the total metabolic variance. Apart from the fairly small differences in mean values between the two genotypes, this result is probably caused by the high amount of non-genotype related variance caused by all other metabolites. By using lower-order vectors, metabolic phenotypes could partly or fully be discriminated between SS2 and WT plants (Fig. 10, which is published as supporting information on the PNAS web site). On leaf metabolic phenotypes, glucose, fructose, and sucrose had almost no impact on this discrimination, but rather those compounds that were already observed to have significantly different mean values. For tubers, glucose and fructose were found among the most important metabolites for classification based on loading scores for vector 4 (4.0% of total variance, Fig. 10) when restricting the PCA analysis to polar metabolites. In accordance with the alteration in sucrose synthase and specific SPS activities, this finding indicated a subtle alteration in glucose and fructose metabolism between SS2 and WT tubers.

Alternatively, supervised learning methods such as DFA can be used to classify metabolite data sets (12). By using a nonparametric normal kernel DFA, genotypes could be discriminated by using internal cross-validation. Investigation of the discriminant functions, however, pointed to the same metabolites that had already been demonstrated to be significantly different in mean levels, e.g., lysine levels in tubers or galactonic acid in leaves. In this respect, DFA was not helpful for gaining further insights into metabolic alterations in the SS2 plants or pointing to the primary cause of the genetic defect.

Metabolic Correlation Analysis as a Complementary Biochemical Tool. Analyzing a large number of snapshots of the same genotype permitted the search for metabolic correlations (Fig. 3) that *per se* contain inherent information on how metabolites are related within a complex network of reaction pathways and regulatory events. In a recent study, we demonstrated that metabolic fluctuations may cause linear associations between metabolite levels as a consequence of the underlying reaction pathway structure (8, 9). However, experimentally observed correlations between variables are not straightforwardly interpretable in a complex system (9, 30, 31). Studies on RNA levels have derived hypotheses from weak correlations of $r_{xy} \geq 0.40$ (32). In our study, several thousand metabolite pairs were found meeting such low thresholds. Therefore, we suggest restricting the analysis of metabolic correlations to higher thresholds that are more significant to display biological connectivity (moderate ($r_{xy} \geq 0.60$) to strong ($r_{xy} \geq 0.80$) linear relationships).

Some clear differences were observed between WT and SS2 antisense plants. For example, a significant change in correlation offsets was found for lysine and isoleucine levels in potato tubers after log transformation (Figs. 3 and 7), which was not apparent from the small differences in average levels. The same was found for the glutamine-asparagine pair. From this observation, it can be concluded that not only were average levels of these amino acids increased in SS2 antisense tubers, but also the ratio between these metabolite pairs was altered. This finding indicates a shift of relative control over carbon partitioning between the pools of soluble metabolites of both genotypes. Such observations might facilitate linking changes in metabolic levels back to the relative impact on different biochemical pathways.

The significance of changes in slopes was tested by transforming all data to metabolite/metabolite ratios, for which fixed correlations were found above a threshold of $r_{xy} = 0.80$ in one of both genotypes (*t* test, $P < 0.05$). From 17,205 tuber metabolite correlations tested, 1,181 (6.9%) showed a strong correlation in at least one of both genotypes, and, for 157 of these pairs, significant differences in metabolite/metabolite ratios were found (Table 3). The largest group (65 pairs) among these included at least one amino acid, and the most prominent effects were found for pairs that included lysine, glutamine, and γ-aminobutyric acid (GABA). For leaves, 1,733 of 75,078 (2.3%) searched polar metabolite pairs were found to have a correlation $r_{xy} > 0.80$ in at least one genotype. Among these, 317 were calculated to have significantly different metabolite/metabolite ratios. Unlike in tubers, the largest effect in leaves was found among sugar alcohols. As exemplified in Fig. 3, some metabolite pairs such as arabinose-tagatose showed clear correlation in SS2 leaves, which was not evident in WT leaves. For other pairs, like trehalose-erythritol, metabolite levels in both SS2 and WT leaves implied a strong linear correlation, and again a difference in metabolite ratios was found. Two hundred twenty-six of the metabolite pairs with altered correlation slopes were related to carbohydrate metabolism and only 15 to amino acids. The most apparent differences were found for leaf metabolite pairs that included erythritol, mannitol, trehalose, or tetronate. Comparing the results for leaf and tuber metabolite ratios, these differences in control of carbon partitioning suggest that the overall effect of the SS2 antisense construct was organ dependent.

Topology of Metabolite Correlation Networks Reveals Alteration in Carbohydrate Metabolism in Silent SS2 Plants. The working hypothesis was that primary carbohydrates like sucrose, fructose, and glucose are heavily interconnected with multiple other pathways and might thus have impact on overall reaction networks (33, 34). Therefore, we assumed changes of the structural properties of the metabolic network itself. In analogy to recent studies of protein interaction networks (35), transcriptional regulation frameworks (36), and reconstructed metabolic systems (33, 37, 38), we extended metabolic correlation analysis to the level of network topology, which may be used for further calculations or visualized by graphs. For WT and SS2 leaf total correlation networks, few direct correlations between polar and lipophilic metabolites were found. Due to the sheer number of pairwise metabolic correlations, large overview network graphs easily get incomprehensible. Therefore, the subnetwork of polar metabolites was further investigated to emphasize the effects on primary metabolism and to exemplify the idea of using connectivity topology ranking (Fig. 4). Strong correlations between trehalose and sugar alcohols were observed in leaves but not in tuber networks. Besides its proposed role in sugar sensing in plants (39), trehalose also affects sucrose synthase and invertase activities (40) and is thus directly related to basic sucrose metabolism. Generally, the partial network graph in Fig. 4 points to a more general way to assess genotypic differences: intriguingly, trehalose and sugar alcohols had far fewer numbers of network connections in SS2 than in WT leaves. We propose to use such differences of "metabolite connectivities" as a general tool to study perturbation in metabolic networks. First, the connectivity count of each metabolite has to be normalized to the total number of correlations in the metabolic network. At any given relevance level of r_{xy}, the average number of correlations per metabolite was found to be higher in SS2 tubers compared with WT tubers, and lower in SS2 leaves than in WT leaves (Fig. 11, which is published as supporting information on the PNAS web site). This result indicates robustness of network calculations to the choice of the actual

Table 3. Differences in ratios of polar metabolite pairs comparing SS2 and WT tubers and leaves

Metabolite ratio	*P* value	x-fold WT/SS2
Tuber		
Arg/homoserine	0.000	1.8
Arg/lys	0.000	1.6
Asn/gln	0.000	1.5
Asn/homoserine	0.000	1.8
Gly/homoserine	0.000	1.9
Met/lys	0.000	1.4
Ser/homoserine	0.000	1.6
Shikimate/homoserine	0.000	1.7
Thr/homoserine	0.000	1.9
Met/GABA	0.001	1.5
Met/homoserine	0.001	1.7
Succinate/phe	0.001	1.5
Thr/lys	0.001	1.7
Arg/citrulline	0.002	1.3
Thr/GABA	0.002	1.6
Val/lys	0.003	1.3
Gly/GABA	0.004	1.5
Val/homoserine	0.004	1.6
Leu/lys	0.005	1.3
Ser/GABA	0.005	1.6
Shikimate/met	0.005	1.3
Ile/lys	0.007	1.2
Thr/gln	0.008	1.5
Malate/fumarate	0.010	0.8
Inositol/fructose-6p	0.013	0.6
Ser/citrulline	0.017	1.2
Met/citrulline	0.018	1.2
Ser/val	0.018	1.2
Glycolate/homoserine	0.024	1.7
Arg/val	0.026	1.2
Benzoate-4OH/caffeate	0.027	1.2
Gly/gln	0.028	1.4
Met/val	0.034	1.1
Ser/phe	0.042	1.4
Putrescine/phe	0.044	1.4
Thr/val	0.044	1.3
Gln/homoserine	0.046	1.4
Met/phe	0.047	1.3
Leaf		
Trehalose/arabinitol	0.000	2.8
Mannitol/arabinitol	0.000	2.4
Galactitol/arabinitol	0.000	1.3
Glucose-6deoxy/erythritol	0.000	0.7
Glucose-6deoxy/mannitol	0.000	0.5
Erythritol/arabinitol	0.001	1.6
Erythritol/trehalose	0.001	0.6
Tetronate/mannitol	0.001	0.6
Glycerate/malate	0.005	1.4
Erythritol/mannitol	0.009	0.7
Trehalose/mannitol	0.016	1.1
Galactitol/mannitol	0.017	0.6
Sorbitol/galactitol	0.021	0.9
Beta-alanine/asp	0.023	4.8
Tagatose/arabinose	0.023	1.4
Glycerate/tetronate	0.025	1.2
Glucose-6deoxy/trehalose	0.025	0.4
Beta-alanine/arabinitol	0.032	1.9
Galactose/arabinitol	0.032	0.6
Quinate/glucose-6deoxy	0.034	1.6
Tetronate/malate	0.037	1.2
Glu/asp	0.037	0.8

Only polar metabolites with known chemical structure are shown. Proteinogenic amino acids are abbreviated by three-letter code. Tuber: 1,181 pairs tested; 157 pairs with $P < 0.05$; 62 pairs with $P < 0.01$. Leaf: 1,733 pairs tested; 317 pairs with $P < 0.05$; 149 pairs with $P < 0.01$.

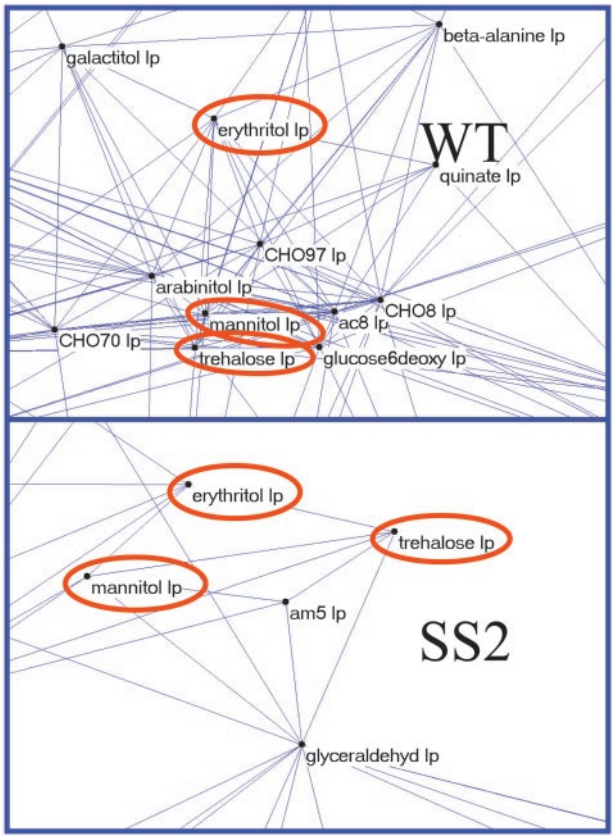

Fig. 4. Enlarged view of polar metabolite correlations in WT and SS2 leaves at $r_{xy} > 0.80$. am, amine; CHO, carbohydrate; ac, acid; extension lp, leaf polar fraction.

threshold taken for r_{xy}. Second, the connectivity distribution $P(k)$ of polar metabolites was shown to follow the typical power law for scale-free networks (41) (Fig. 5). In scale-free networks, most nodes are only sparsely connected with others whereas some few nodes are heavily associated with many others. This property is uniformly and

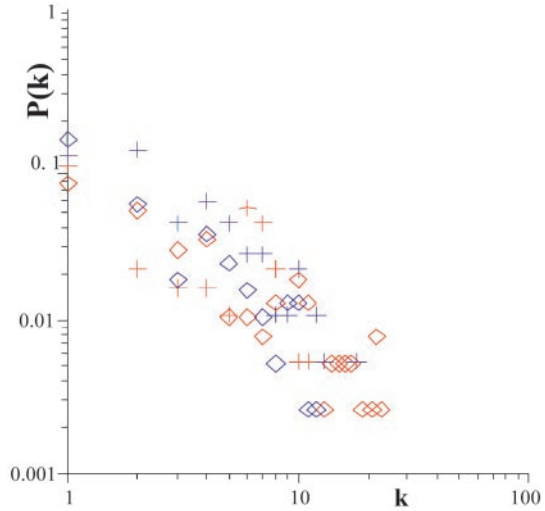

Fig. 5. Distribution function $P(k)$ of polar potato plant metabolite correlations at $r_{xy} > 0.7$, giving the probability that a randomly selected metabolite has exactly k correlations to other metabolites. Blue, SS2; red, WT plants; crosses, tubers; diamonds, leaves.

Table 4. Difference in relative connectivity in metabolic networks between SS2 and WT plants

Leaf		Tuber	
Metabolite	Rel. connectivity SS2 − WT, $r > 0.70$	Metabolite	Rel. connectivity SS2 − WT, $r > 0.70$
Arabinitol	−4.467	Phenylalanine	−8.08
Galactitol	−2.641	Glucose	−7.84
Glucose-6-deoxy	−4.435	Arginine	−6.28
Glyceraldehyde	−4.274	β-Alanine	−5.05
Mannose	−3.263	Galactose	−4.89
Normethyladrenaline	−3.344	Alanine	−4.88
Quinic acid	−3.246	Fructose	−4.61
Trehalose	−2.460	GABA	−4.37
Glutamate	2.218	Raffinose	4.23
Malate	2.327	Citrate	4.50
Raffinose	2.454	Sucrose	4.63
Fructose	2.661	Homoserine	6.68
Psicose	2.725	Aspartate	6.90
Fumaric acid	2.788	Glycerate	7.70
Isotetronate	3.154	Phosphate	8.47
Oxoproline	3.704	Nicotinate	11.08
Phenylalanine	4.231		
Tetronate	4.871		
Aspartate	5.576		
Tagatose	6.481		

Metabolites with negative values have a higher connectivity in WT plants. Positive values indicate a higher connectivity in SS2 plants.

linearly distributed over the whole connectivity range, and it provides the possibility to compare SS2 and WT network connectivities by ranking the relative differences in correlations per metabolite. In Table 4, the extremes in connectivity differences are given for compounds with known chemical identity. Most interestingly, effects for primary carbohydrate metabolism in SS2 tubers could be observed more clearly than by the other statistical methods previously applied. Sucrose showed more correlations in SS2 tubers whereas fewer correlations were found for glucose, fructose, and galactose. This finding of opposite effects found for sucrose with respect to glucose and fructose is in accordance with the increased total activity of sucrose synthase in tubers. Additionally, large relative changes in aspartate connectivity in tuber metabolic networks were found. Together with the significant changes in metabolite/metabolite ratios for aspartate-derived amino acids such as homoserine and lysine and increased lysine average values, this result indicates a general shift in control of aspartate-derived biochemical pathways in SS2 tubers. This result is in good agreement with the recent finding that sucrose and amino acid metabolism are intimately connected in potato tubers (23).

Conclusions

The functional assignment of silent phenotypes is a necessity for today's genomic strategies, specifically when altering the expression of one of several enzyme isoforms. Here, it is shown how to distinguish the silent SS2 phenotype from its parental background line by comprehensive metabolome analysis combined with statistical tools. Particularly, topological differences in metabolic correlation networks proved useful to complement findings of subtle differences in variances and averages of metabolite levels. A mechanistic interpretation of such observations is still hampered by two important features: first, a high number of metabolic peaks in MS are as yet unknown, disabling a link to known biochemical networks. Second, regulations that can cause metabolic correlations can be based on all system levels, i.e., transcription, translation, and ultimately enzyme activities including protein interaction clusters. The utility of metabolic networks has to be complemented with classical biochemical studies, for instance by comprehensive profiling of metabolic fluxes (42) or enzyme activities as in this study. Nevertheless, analysis on the metabolome level is applicable to all kinds of perturbated biological systems at competitive costs, enabling the use of a wide variety of statistical tools for generating novel hypotheses.

We thank Gareth Catchpole for MATLAB calculations, Frank Kose for improving network algorithms, Alisdair Fernie for help with SS2 expression analysis, and Birgit Linkohr for carrying out DFA analyses. This project was funded through the Max Planck Society and a Fundaçao de Amparo à Pesquisa do Estado de Minas Gerais–Comissao de Aperfeiçoamento de Pessoal de Nival Superior–Conselho Nacional de Pesquisas–Deutscher Akademischer Austauschdienst fellowship (to M.E.L.).

1. Thorneycroft, D., Sherson, S. M. & Smith, S. M. (2001) *J. Exp. Bot.* **52,** 1593–1601.
2. Bouche, N. & Bouchez, D. (2001) *Curr. Opin. Plant Biol.* **4,** 111–117.
3. Nelson, D. R. (1999) *Arch. Biochem. Biophys.* **369,** 1–10.
4. Lunn, J. E. & MacRae, E. (2003) *Curr. Opin. Plant Biol.* **6,** 208–214.
5. Doblin, M. S., Kurek, I., Jacob-Wilk, D. & Delmer, D. P. (2002) *Plant Cell Physiol.* **43,** 1407–1420.
6. Graham, I. A. & Eastmond, P. J. (2002) *Prog. Lipid. Res.* **41,** 156–181.
7. Fiehn, O., Kopka, J., Dormann, P., Altmann, T., Trethewey, R. N. & Willmitzer, L. (2000) *Nat. Biotechnol.* **18,** 1157–1161.
8. Weckwerth, W. (2003) *Annu. Rev. Plant Biol.* **54,** 669–689.
9. Steuer, R., Kurths, J., Fiehn, O. & Weckwerth, W. (2003) *Bioinformatics* **19,** 1019–1026.
10. Fell, D. (1997) Understanding the Control of Metabolism (Portland Press, London).
11. Giersch, C. (1995) *Eur. J. Biochem* **227,** 194–201.
12. Raamsdonk, L. M., Teusink, B., Broadhurst, D., Zhang, N. S., Hayes, A., Walsh, M. C., Berden, J. A., Brindle, K. M., Kell, D. B., Rowland, J. J., *et al.* (2001) *Nat. Biotechnol.* **19,** 45–50.
13. Allen, J., Davey, H. M., Broadhurst, D., Heald, J. K., Rowland, J. J., Oliver, S. G. & Kell, D. B. (2003) *Nat. Biotechnol.* **21,** 692–696.
14. Geigenberger, P. & Stitt, M. (1993) *Planta* **189,** 329–339.
15. Fu, H. Y. & Park, W. D. (1995) *Plant Cell* **7,** 1369–1385.
16. Barratt, D. H. P., Barber, L., Kruger, N. J., Smith, A. M., Wang, T. L. & Martin, C. (2001) *Plant Physiol.* **127,** 655–664.
17. Sturm, A. & Tang, G. Q. (1999) *Trends Plant Sci.* **4,** 401–407.
18. Sergeeva, L. I. & Vreugdenhil, D. (2002) *J. Exp. Bot.* **53,** 361–370.
19. Zrenner, R., Salanoubat, M., Willmitzer, L. & Sonnewald, U. (1995) *Plant J.* **7,** 97–107.
20. Chen, Y. C. & Chourey, P. S. (1989) *Theor. Appl. Genet.* **78,** 553–559.
21. Koch, K. E., Nolte, K. D., Duke, E. R., Mccarty, D. R. & Avigne, W. T. (1992) *Plant Cell* **4,** 59–69.
22. Guerin, J. & Carbonero, P. (1997) *Plant Physiol.* **114,** 55–62.
23. Roessner-Tunali, U., Urbanczyk-Wochniak, E., Czechowski, T., Kolbe, A., Willmitzer, L. & Fernie, A. R. (2003) *Plant Physiol.* **133,** 683–692.
24. Werner, R. A. & Schmidt, H. L. (2002) *Phytochemistry* **61,** 465–484.
25. Kofler, H., Hausler, R. E., Schulz, B., Groner, F., Flugge, U. I. & Weber, A. (2000) *Mol. Gen. Genet.* **263,** 978–986.
26. Chourey, P. S. & Taliercio, E. W. (1994) *Proc. Natl. Acad. Sci. USA* **91,** 7917–7921.
27. Winter, H. & Huber, S. C. (2000) *Crit. Rev. Biochem. Mol. Biol.* **35,** 253–289.
28. Dalluge, J., Vreuls, R. J. J., Beens, J. & Brinkman, U. A. T. (2002) *J. Separation Sci.* **25,** 201–214.
29. Hoglund, T. (1979) *Z. Wahrscheinl. Verw. Geb.* **49,** 105–117.
30. Fiehn, O. & Weckwerth, W. (2003) *Eur. J. Biochem.* **270,** 579–588.
31. Weckwerth, W., Wenzel, K. & Fiehn, O. (2004) *Proteomics* **4,** 78–83.
32. Ideker, T., Thorsson, V., Ranish, J. A., Christmas, R., Buhler, J., Eng, J. K., Bumgarner, R., Goodlett, D. R., Aebersold, R. & Hood, L. (2001) *Science* **292,** 929–934.
33. Schuster, S., Fell, D. A. & Dandekar, T. (2000) *Nat. Biotechnol.* **18,** 326–332.
34. Vance, W., Arkin, A. & Ross, J. (2002) *Proc. Natl. Acad. Sci. USA* **99,** 5816–5821.
35. Gavin, A. C., Bosche, M., Krause, R., Grandi, P., Marzioch, M., Bauer, A., Schultz, J., Rick, J. M., Michon, A. M., Cruciat, C. M., *et al.* (2002) *Nature* **415,** 141–147.
36. Lee, T. I., Rinaldi, N. J., Robert, F., Odom, D. T., Bar-Joseph, Z., Gerber, G. K., Hannett, N. M., Harbison, C. T., Thompson, C. M., Simon, I., *et al.* (2002) *Science* **298,** 799–804.
37. Ravasz, E., Somera, A. L., Mongru, D. A., Oltvai, Z. N. & Barbasi, A. L. (2002) *Science* **297,** 1551–1555.
38. Fell, D. A. & Wagner, A. (2000) *Nat. Biotechnol.* **18,** 1121–1122.
39. Rolland, F., Moore, B. & Sheen, J. (2002) *Plant Cell* **14,** S185–S205.
40. Muller, J., Boller, T. & Wiemken, A. (1998) *J. Plant Physiol.* **153,** 255–257.
41. Jeong, H., Mason, S. P., Barabasi, A. L. & Oltvai, Z. N. (2001) *Nature* **411,** 41–42.
42. Kiefer, P., Heinzle, E., Zelder, O. & Wittmann, C. (2004) *Appl. Environ. Microbiol.* **70,** 229–239.