Aberystwyth University

Opposite variations in fumarate and malate dominate metabolic phenotypes of
Arabidopsis salicylate mutants with abnormal biomass under chilling

Scott, Ian Morris; Ward, Jane L.; Miller, Sonia J.; Beale, Michael H.

Opposite variations in fumarate and malate dominate metabolic phenotypes of *Arabidopsis* salicylate mutants with abnormal biomass under chilling

Ian M. Scott[a,*], Jane L. Ward[b], Sonia J. Miller[b] and Michael H. Beale[b]

[a]Institute of Biological, Environmental and Rural Sciences, Aberystwyth University, SY23 3DA, UK

[b]National Centre for Plant and Microbial Metabolomics, Rothamsted Research, Harpenden, AL5 2JQ, UK

Correspondence

*Corresponding author, e-mail: ias@aber.ac.uk

Abstract

In chilling conditions (5ºC), salicylic acid (SA)-deficient mutants (*sid2*, *eds5*, *NahG*) of *Arabidopsis thaliana* produced more biomass than wild-type (Col-0), while the SA-overproducer *cpr1* was extremely stunted. The hypothesis that these phenotypes were reflected in metabolism was explored using 600 MHz ^{1}H NMR analysis of unfractionated polar shoot extracts. Biomass-related metabolic phenotypes were identified as multivariate data models of these NMR 'fingerprints'. These included principal components that correlated with biomass. Also, partial least squares-regression models were found to predict the relative size of plants in previously unseen experiments in different light intensities, or relative size of one genotype from the others. The dominant signal in these models was fumarate, which was high in SA-deficient mutants, intermediate in Col-0, and low in *cpr1* at 5ºC. Among signals negatively correlated with biomass, malate was prominent. Abundance of transcripts of the FUM2 cytosolic fumarase (At5g50950) showed strong positive correlation with fumarate levels and with biomass, whereas no significant differences were found for the FUM1 mitochondrial fumarase (At2g47510). It was confirmed that the morphological effects of SA under chilling find expression in the metabolome, with a role of fumarate highlighted.

Abbreviations - CV, coefficient of variation; DW, dry weight; N_0, transcript quantity measure; NMR, nuclear magnetic resonance; PCA, principal component analysis; PLS, partial least squares; qRT-PCR, quantitative reverse transcription-polymerase chain reaction; RGR, relative growth rate; SA, salicylic acid; TSP, trimethylsilylpropionate; **X**, predictor variables matrix; **y**, response variables vector.

Introduction

Understanding temperature responses of biomass and metabolism is important for modifying seasonal and geographical crop limits to meet production challenges of demographic and climate change (Humphreys et al. 2006), and in predicting biosphere impacts of global change (Enquist et al. 2007). The research model *Arabidopsis thaliana* belongs to the Brassicaceae, many members of which are adapted to seasonal cold (Griffith et al. 2007). It is typically a winter-annual, germinating in the fall and overwintering as a rosette to reproduce in spring, though this varies among populations (Griffith et al. 2004).

Chilling (low but non-freezing) temperatures influence biomass allocation, though the mechanisms are poorly understood. A hormonal component in chilling-induced growth inhibition in *Arabidopsis* was found by Scott *et al.* (2004), who observed that salicylic acid (SA) and its glucoside slowly accumulated at 5°C. Transgenic and mutant plants unable to accumulate SA had biomass 2.7-fold that of wild-type after 2 months at 5°C. SA-deficient genotypes exhibiting such responses include: *NahG*, in which SA accumulation is prevented by a bacterial hydroxylase transgene; the *eds5* (or *sid1*) mutant, which is defective in a putative plastid membrane protein homologous to MATE transporters; and the *sid2* mutant, which is defective in the ICS1 isochorismate synthase that mediates SA biosynthesis (Kim et al. 2013, Scott et al. 2004). The *ICS1* gene is induced by chilling (Kim et al. 2013), and the enzyme has unusually high activity at chilling temperature (Strawn et al. 2007).

Conversely, SA-overaccumulation and stunted growth at 5°C were seen in mutants of the *cpr1* gene (Scott et al. 2004), which maps to a *Resistance* cluster and affects defense and growth via a temperature-responsive signaling pathway (Yang and Hua 2004). Kim *et al*. (2013) found the same high-SA, low-growth chilling syndrome in mutants of the CAMTA family of calmodulin-binding transcriptional activators, which may transduce the cold-induced spike in cytoplasmic calcium levels into *CBF* gene activation. CAMTA factors repressed *ICS1*, and hence SA biosynthesis at warm temperature. Microarray analysis suggested a major role for SA in configuring the low temperature transcriptome (Kim et al. 2013).

Metabolism, the theme of this paper, is a major determinant of biomass accumulation. Interpreting the complexity of metabolite composition (or, the metabolome), however, remains a challenge in the field of metabolomics (Steinfath et al. 2008). There is evidence that metabolomics has predictive power for growth rate. Correlations of biomass with metabolites (Meyer et al. 2007, Sulpice et al. 2013, Sulpice et al. 2009), or coincidence of quantitative trait loci for biomass and metabolites (Lisec et al. 2008), have been sought in *Arabidopsis*. A motivation of these studies has been improved crop yields for bioenergy and food (Meyer et al. 2007).

The hypothesis of the present study was that the remarkable biomass phenotypes of SA mutants under chilling could be mathematically modeled using their metabolomes. Irrespective of the nature and complexity of the underlying cause-and-effect mechanisms, demonstration that biomass could be

predicted from metabolite data confirms a relationship in the context of both low temperature and hormonal regulation. This has relevance for understanding stress-related geographical, seasonal or genetic limits to growth of crops or natural vegetation.

Materials and methods

Plant material

All *Arabidopsis thaliana* (L.) lines were in the Col-0 wild-type background (Scott et al. 2004). Growth stages were monitored as in Boyes *et al.* (2001). Seedlings germinated in compost (Scott et al. 2004) to growth stage 1.0 were placed in a Fisons 600G3/THTL chamber (Loughborough, UK) at 23ºC mean air temperature. At a mean growth stage of 1.04, plants were either kept in this chamber, or transferred to a similar one at 5ºC, and incubated up to a further 100 d. Within each chamber, genotypes were distributed in mixed assortments across six trays whose positions were exchanged twice weekly. No additional nutrients were provided. Daily light periods (16 h) were provided by Sylvania F36W/135 fluorescent tubes (Raunheim, Germany), monitored with a Skye Instruments SKP 215 PAR Quantum light sensor (Llandrindod Wells, UK). Approximate maximum irradiances (400-700 nm) were 100, 55 or 25 μmol m^{-2} s^{-1}, this range being low to minimize stress symptoms apparent in stronger light at 5ºC, especially for *cpr1* (Scott et al. 2004).

Growth conditions in the metabolomics experiments (coded A-G) are in Table 1. Shoots were harvested at stage 5.10 (first unopened flower buds), except for experiment G, in which flowers were allowed to open (growth stage 6.00-6.10). These growth stages were reached at the same time by all plants apart from the *cpr1* mutants at 5ºC, which were infertile, forming minute, highly abnormal inflorescence-like structures before the other genotypes.

Shoot vegetative tissues were sampled over 90 min in mid-light-period. Roots and inflorescence stems were quickly detached with a blade, and the rosette immediately plunged into liquid N_2. Samples were freeze-dried for 48 h, then weighed to obtain 'biomass' as DW. Shoot biomass varied 86-fold, so differing numbers were batched for NMR analysis: 46% of samples were single shoots, 35% 2-4 shoots, and 19% 12 or more. Every sample had at least four biological replicates, 91% five or more, and 60% ten or more. Samples were stored with desiccant at -80ºC until NMR analysis, which entailed shipment and laboratory processing over several days at ambient temperature.

Plants for the quantitative real-time PCR analyses were similarly grown (62 d at 5ºC in 51 μmol m^{-2} s^{-1}), harvested (freeze-dried at stage 5.10), and stored (desiccated at -80ºC).

Nuclear magnetic resonance (NMR) fingerprinting

Samples (15 ± 0.03 mg) were extracted (50 ºC, 10 min) in 1 ml 80:20 D_2O-CD_3OD with 0.05% d_4-TSP (trimethylsilylpropionate, sodium salt). After cooling and centrifugation, supernatants (850 μl) were re-heated (90ºC, 2 min), refrigerated 45 min and re-centrifuged. ^{1}H NMR spectra of supernatants (750 μl) were acquired at 300 ºK on an Avance Spectrometer (Bruker Biospin, Coventry, UK) at 600.05 MHz, with a 5 mm selective inverse probe. A water suppression pulse sequence with 5 s relaxation delay was used. Spectra, acquired using 128 scans of 64 000 data points with 7 309.99 Hz

spectral width, were Fourier transformed using an exponential window with 0.5 Hz line broadening. Phasing and baseline correction used instrument software. ^{1}H chemical shifts were referenced to d_4-TSP at $\delta 0.0$. Spectra were reduced in Amix software (Bruker Biospin) to ASCII files containing integrated ('binned') regions of equal width (0.01 ppm), and intensities scaled to d_4-TSP ($\delta 0.05$ to -0.05). Residual water ($\delta 4.865$-4.775), d_4-MeOH ($\delta 3.335$-3.285) and d_4-TSP signals were removed. Where sample quantities permitted (95% of cases), experimental variation was reduced by averaging spectra of 2-3 replicate extractions. The resultant spectra are hereafter called 'fingerprints'. Metabolite assignments of spectral peaks were made using standard compounds. Individual metabolites were quantified relative to the d_4-TSP internal standard using the following spectral ranges: fumarate ($\delta 6.535$-6.505); sucrose ($\delta 5.435$-5.405); malate ($\delta 4.355$-4.295); glutamine ($\delta 2.465$-2.425); citrate plus malate ($\delta 2.605$-2.475).

Univariate statistics

For full spectral datasets, one-way ANOVA (without assumption of equal group variances) was done in the R package FIEmspro (Enot et al. 2008), and correlation by the *corrcoef* function in MATLAB v.6.5 (The MathWorks, Natick, MA). Standard Bonferroni assessments of *P* values were made (Broadhurst and Kell 2006). Smaller datasets were analyzed using PAST v.1.91 (Hammer et al. 2001) for one-way ANOVA with Tukey *post-hoc* tests (where data passed Shapiro-Wilk, Jarque-Bera and Levene tests for normality and homogeneity of variance), Pearson and rank correlations, and non-parametric Kruskal-Wallis and Mann-Whitney tests.

Multivariate data modeling

Principal component analysis (PCA) and partial least squares (PLS)-regression (Steinfath et al. 2008) were performed in SIMCA-P v.11.0 (Umetrics, Umeå, Sweden) on unscaled, mean-centered data matrices, whose columns were NMR fingerprints of 901 binned signals (of which 762 were non-zero), and whose rows were plant samples (numbering up to 183 in the full dataset). All reported PCs were 'significant' in the standard SIMCA-P cross-validation. Cross-validation can also be used in SIMCA-P to optimize the number of components in PLS-regression models (Wold et al. 2001). Optimization avoids over-fitting variation spuriously correlated with **y**, the vector of 'response' variables (here, shoot biomass). In SIMCA-P, the data are divided into seven subsets, and a series of models developed with each subset missing is used to predict **y** for the held-out subset. Optimal models require a series of components that successively reduce the overall prediction error. We found one to three PLS components were optimal in exploratory analyses of individual growth experiments. In predictive modeling, however, optimization is problematical unless the test data (whose **y** variables are to be predicted) are excluded from this step of model development (Broadhurst and Kell 2006). To avoid this issue, all reported PLS analyses were conducted using a fixed number of two components.

PCR primer selection and design

Four genes potentially suitable as internal references were identified by stable expression in the following cold- or salicylate-related experiments: EMBL-EBI ArrayExpress (http://www.ebi.ac.uk/arrayexpress/) E-MEXP-1345, E-TABM-51, E-TABM-52; NCBI Gene Expression Omnibus (http://www.ncbi.nlm.nih.gov/geo/) GSE Series 5534, 5620-21, 5745, 6177, 9955, 10522, 10646, 19255; NASCArrays (http://affy.arabidopsis.info/) 404. These yielded 83 Affymetrix ATH1 Genome Array microarray plate records, which were normalized by robust multi-array averaging in Robin (Lohse et al. 2010). ATH1 probe sets were ranked by coefficient of variation (CV) of normalized expression values across all plates, then compared with Czechowski *et al*. (2005). Stable expression (99th percentile of CV values) was confirmed for At1g13320, At4g32470 and At5g08290 (Czechowski et al. 2005). Also chosen was the traditional reference gene *EF-1α* (ATH1 probe set 247644_s_at, 96th CV percentile), for which At1g07940 was used for primer design.

Sequences for the reference genes, and the following genes of interest, were obtained from NCBI (http://www.ncbi.nlm.nih.gov/): At1g53310, At2g22500, At2g47510, At4g00570, At5g01340, At5g47560, At5g50950, At5g66760. Choices of genes over similar functional candidates involved preliminary assessment of their levels and inducibility in meta-analyses of the microarray experiments listed above, in addition to literature evidence.

Parameters of primers, designed using Primer-BLAST (Ye et al. 2012), PrimerExpress 2.0 (Applied Biosystems, Foster City, CA, USA) and NetPrimer (Premier Biosoft, Palo Alto, CA, USA), were: T_m 59 ± 1ºC, amplicons 69-83 base pairs, primers 18-25 bases, GC contents 40-61%. Each primer pair was tested in PCR with and without cDNA template, and over a concentration range of 50 to 900 nM. Primer pair 5'-3' sequences (forward and reverse) for each transcript were:
At1g07940 (ACCGGAGCCAAGGTTACCA and ACCTGCGGCAGATAGAGTTTTG),
At1g13320 (GATTCTTCGTGCAGTATCGCTTCT and TTACCGCAGGTAAGAGTTTGGAA),
At1g53310 (GTTCTGGACACGTTTCATGTCATAG and
GAAGGTGCAGTTGCCATAGAGATAA), At2g22500 (TGCGAGCGTTGCGAGTAAT and
ACTCCAGCCACCACCTTCATATT), At2g47510 (CTCCGGCGGAACAACTGT and
TCCCTAAACGAGGTCGAATAAGATC), At4g00570 (CTTCTCTCTGGTGCACGTATCG and
GTCATGTAGGACGCAAGGCATT), At4g32470 (ATGGACATCAAAGAGGCTATGAACA and
CACGCTTGAGACGCTGGTTA),
At5g01340 (GAGACCGGAAAGGTCAGCAA and TGGCGAGAGCTTCAAGAACAC), At5g08290
(GCGTCTCGTCGTCATTCGTT and CGCAAGCACCTCATCCATCT), At5g47560
(GACAAACGGTGGGACAACAAC and GAGGACCAAGAGCGATGTAACAG),
At5g50950 (AAGATGTGGTCTTGGTGAACTTTCT and TGTAGGATTTACCTTTCCAGGCATA),
At5g66760 (CTCTTTGATGGCGTCTCTGGAT and GAGGATCTCGTATCGGTAGAACCA).

Quantitative reverse transcription (qRT)-PCR procedure

Freeze-dried shoots were minced with a blade, and random portions (10 mg) sampled for total RNA extraction with the RNeasy Plant Mini Kit (Qiagen, Valencia, CA, USA). RNA quantity and quality were assessed by UV spectra and agarose gels. The Qiagen QuantiTect Reverse Transcription Kit was used for genomic DNA elimination and cDNA synthesis from 9 µg RNA samples; controls without reverse transcriptase were included to confirm absence of genomic DNA.

PCR reactions were performed in optical 96-well plates on an Applied Biosystems 7500 Real-Time PCR System. Reactions contained 2 µL cDNA, 7.9 µl master mix of SYBR Green PCR Core Reagents (Applied Biosystems), and 400 nM each of forward and reverse primers in a final volume of 25 µl. The thermal program was 40 cycles of 95ºC (15 s) and 60ºC (1 min), after initial steps of 50ºC (2 min) and 95ºC (10 min). Dissociation curves (60-95ºC) were recorded after each 40-cycle program.

Amplication data were exported from the 7500 System SDS v.1.2.3 software as fluorescence ratios (of the SYBR Green reporter and ROX passive reference dyes) for each cycle. Transcript quantities (N_0) per sample, in arbitrary fluorescence units, were obtained in LinRegPCR v.2012.2, using mean PCR efficiency per amplicon (Ruijter et al. 2009).

Normalization of qRT-PCR data

Potential reference genes At1g07940 (*EF-1α*), At1g13320 (*PP2A3*), At4g32470 (*QCR7-1*) and At5g08290 (*YLS8*) were evaluated, by average pairwise variations (*M*) in the cDNA samples between one gene and the others, calculated in geNorm v.3.5 (Vandesompele et al. 2002). Successive elimination of the less stable genes in geNorm indicated that the two references *EF-1α* and *PP2A3* were optimal ($M = 0.20$, $r = 0.97$, $n = 16$). All N_0 values were therefore normalized using geNorm-calculated geometric means of *EF-1α* and *PP2A3* measurements in the same qRT-PCR plates and samples. Although *YLS8* correlated with both these genes (r 0.91, 0.93), its inclusion caused insufficient variation of geNorm normalization factors (0.097) to justify an additional reference (Vandesompele et al. 2002). Quantifications of each transcript of interest were made on four biological replicates (different shoots), each measured in two qRT-PCR plates.

Results

NMR fingerprints of SA mutants

Over several weeks at 5ºC, relative shoot biomass of *Arabidopsis* genotypes correlated inversely with ability to accumulate SA (Fig. 1a). Thus, *cpr1* mutants, which over-accumulate SA (Scott et al. 2004), were very stunted (Fig. 1a). Abnormally large shoots, on the other hand, were formed by the *sid2*, *eds5* and *NahG* genotypes, which fail to accumulate SA at 5°C (Kim et al. 2013, Scott et al. 2004).

Chilling treatments in *Arabidopsis* are generally conducted in low light (Gilmour et al. 2000, Klotke et al. 2004). The maximum intensity that could be used without causing visible stress in every genotype at 5ºC was 100 µmol m^{-2} s^{-1} (Scott et al. 2004). Light levels were experimentally varied down from this limit, and in fact the harvest growth stage was reached sooner at 55 µmol m^{-2} s^{-1} (Table 1). When levels were reduced to 25 µmol m^{-2} s^{-1}, the rate of development was sub-optimal, but long-term growth still occurred (Table 1). To evaluate the consistency of metabolic phenotypes, plants subjected to chilling at each light level were analyzed. Mutant biomass at 5ºC was significantly different to Col-0 in all light levels tested, but differentials were greater in the higher light range (Fig. 1b). At 23ºC, *NahG* shoots were slightly smaller than Col-0 (Fig. 1b), as seen previously (Scott et al. 2004).

Shoot metabolites were analyzed by 600 MHz ^{1}H NMR of unfractionated polar extracts. Spectra were binned to 0.01 ppm, which reduces resolution but resolves minor sample-to-sample peak shifts due to small pH changes and thus aids data mining. The resultant fingerprints were comparable to other NMR studies on *Arabidopsis* (Ward et al. 2003), with strong signals for organic acids, amino acids, and a complex sugar-dominated region (Fig. 2a). A sharp singlet around δ 6.5, from resonance of olefinic protons of fumarate (Ward et al. 2003), was prominent in an otherwise sparse spectral region (Fig. 2a), and showed striking variation among the 5ºC fingerprints. Fumarate was maximal in SA-deficient mutants (Fig. 2b), intermediate in Col-0 (Fig. 2c), and low in *cpr1* (Fig. 2d).

Phenotypic variation of each NMR signal (i.e., non-zero intensities of a 0.01 ppm bin) was tested by its *F* value in ANOVA of SA-deficient, wild-type, and *cpr1* classes (Fig. 2). In ANOVA of all 5ºC fingerprints, many signals differed in SA-deficient/wild-type and wild-type/SA-overproducer comparisons; of these, 21% were stronger in the smaller phenotypes, and only 3% the converse. Among NMR signals showing correlation (Bonferroni-corrected significance, $P < 1.5 \times 10^{-5}$) with shoot biomass in each of the 5ºC experiments A-D, 31% were negatively correlated and only 2% positively.

NMR signals were not numerically equivalent to metabolites, which mostly give multiple signals that can overlap in complex regions describing several metabolites. Prominent signals showing negative associations with biomass were, however, attributable to malate and glutamine (Fig. 2). Signals positively associated with biomass included putative sucrose peaks, though their statistical significance did not rival fumarate (Fig. 2).

Metabolomic comparisons between chilling or optimal temperatures were limited in this study. Growth at 5 and 23ºC differed so greatly that plants at similar growth stages were very different in age (Table 1). Low temperature affected many metabolites: 62% of NMR signals for Col-0 at 5ºC were stronger than at 23ºC, while the converse applied to only 2.0% (growth stage 5.10, 55 µmol m^{-2} s^{-1} light, ANOVA, $n \geq 9$ plants, $P < 10^{-4}$). Proline peaks, for example, were eight-fold larger at 5ºC in Col-0 (Fig. 2a,c).

Quantifications of major metabolites, using spectral ranges of their NMR peaks (Table 2), confirmed that fumarate, sucrose, malate and glutamine showed significant (positive or negative) correlations with biomass in each of the 5ºC experiments. Citrate (not shown) was a major metabolite that did not correlate with biomass (mean values: 1.2-1.4 mg g^{-1} DW at 5ºC). At 23ºC, none of these metabolites differed significantly between Col-0 and *NahG*. Comparing Col-0 shoots at 5ºC and 23ºC, sucrose, malate and glutamine were more abundant at low temperature. Unlike studies with shorter cold-incubations (Cook et al. 2004, Kaplan et al. 2004), however, this was not the case for fumarate (Table 2). Fumarate was noteworthy for opposite extremes of variance according to phenotype. For the 23ºC Col-0 and *NahG* shoots, fumarate measurements had proportionately greater SDs than sucrose, malate or glutamine, and showed even more variability for Col-0 shoots at 5ºC. In contrast, fumarate measurements in the SA-deficient mutants at 5ºC were remarkably uniform, across three genotypes in four experiments (Table 2), considering the varied growth conditions (Table 1).

Biomass-oriented metabolic models of SA mutants

Global comparison of metabolic phenotypes was made by PCA, which produces a series of linear combinations of correlated original variables (the PCs), ranked by the data variance each encapsulates. If variance in a PC reflects differences between sample classes, these may separate by their coordinates in PC space or 'scores' (Steinfath et al. 2008). Scores on the first PC separated 23ºC and 5ºC plants, indicating temperature effects accounted for the greatest single source of variance (49%) in the NMR fingerprints (Fig. 3a). The second, PC[2], represented 30% of variance and, most relevant for the hypothesis, separated SA mutants at 5ºC (Fig. 3a).

Data variables (here, NMR chemical shifts) have weight coefficients or 'loadings' on each PC. High positive or negative loadings mean strong contributions to positive or negative scores. Most PC[1] loadings were negative (Fig. 3b) and so contributed to the negative or low positive scores of 5ºC fingerprints (Fig. 3a). The positive PC[1] loadings of fumarate reflected the lack of a cold-induced increase in Col-0 (Table 2), in contrast to other prominent peaks including malate, sucrose, and several amino acids (Fig. 3b). PC[2] loadings were dominated by the positive contribution of fumarate (Fig. 3c), reflecting its dramatic variation (Table 2) between the phenotypes that separated on this PC.

The separation of phenotypes on PC[2] (Fig. 3) appeared to follow a biomass trend. Indeed, when PC[2] scores were plotted against log$_{10}$-biomass of all experimental plants, the correlation was highly significant (Fig. 4a). PC[2] scores were thereby composite biomarkers for biomass at the harvested

developmental stage, despite the wide differences in growth periods and conditions in which plants of different experiments reached this stage (Table 1). The only outliers were the 5°C Col-0 shoots grown in 100 µmol m^{-2} s^{-1} light (Fig. 4a), which reflected the variability of fumarate contents noted in the 5°C experiments (Table 2).

To test more rigorously the hypothesis of a biomass-related component of metabolic phenotype, predictive data modeling was applied. NMR fingerprints were used to model biomass for subsets of the plant samples, e.g. for particular experiments or genotypes. Accurate prediction of biomass for experiments or genotypes *not* involved in model construction would then require meaningful consistency in the hypothetical phenotype.

Properties of chemical samples can be related to their composition by regression, where a vector, **y**, of 'response' variables (here, biomass) is modeled from a matrix, **X**, of 'predictor' variables (here, NMR fingerprints). Where **X** comprises many intercorrelated variables, as here, traditional multiple linear regression is unsuitable. An alternative is PLS-regression, which assumes **X** and **y** are joint realizations of relatively few 'latent' factors underlying the investigated system (Wold et al. 2001). It seeks multivariate components (hypothetically reflecting latent factors) via coordinated algebraic decompositions of **X** and **y** into scores and weights matrices, which are constrained to maximize covariance of **X** and **y**. PLS-regression, like PCA, is suited to the systems biology philosophy (Janes and Yaffe 2006).

Alternative tests were performed to validate PLS modeling of biomass data (expressed as rankings) from NMR fingerprints:

(1) **Cross-experiment predictions**. Models were built on one 5°C experiment, and used to predict the biomass data of the other three. Figure 4b shows the use of experiment D (in the lowest light) to predict the biomass data of experiments A, B and C (in the three light levels). Actual and predicted biomass ranks correlated well using PLS models of any experiment ($r = 0.82$-0.89; $P < 10^{-28}$). This demonstrated consistency in the biomass-related metabolic 'signature' from one experiment to another, despite different growth conditions.

(2) **Cross-genotype predictions.** For each 5°C experiment, one genotype was held out in turn, and its biomass data predicted by models built on the other four genotypes (Fig. 4c). Prediction of *cpr1* was demanding, as its phenotype was far more extreme than the others. Models consequently tended to overestimate *cpr1* biomass, but predictions nonetheless correlated significantly with actual values (Fig. 4c). This supported a relationship between metabolome and biomass in different genotypes.

Consistency was also seen in the NMR signals that were most important in both the PCA and PLS models of different 5°C experiments. Mean Pearson correlation coefficients ($r \pm$ SD) for pairwise between-experiment correlations of relevant PC loadings were 0.80 ± 0.15, and those for PLS regression coefficients were 0.82 ± 0.10. Moreover, the PC loadings correlated with the PLS coefficients (mean rank correlation = 0.72 ± 0.16). Thus, the multivariate analyses indicated a biomass-correlated, NMR-detectable metabolic signature.

The most prominent features among the model coefficients were a positive contribution by fumarate, and negative ones by malate and glutamine (Figs. 3c & 5). Other positive signals (albeit minor relative to fumarate) were sucrose, glucose and other peaks in the spectral region attributable to sugars, and sinapoyl malate.

The 23ºC experiments, in different lighting for different periods (Table 1), yielded a four-fold range in maximal biomass, due to treatment rather than genotype. Shoots reached growth stage 5.10 earlier at 23ºC than at 5ºC (Table 1), and their harvest biomass was less (Fig. 4a). In one experiment (G), plants at 23ºC were left until flowers opened, so they were within the age range of the 5ºC plants (Table 1). All these 23ºC fingerprints, somewhat unexpectedly, fitted the PC[2] biomass-correlation (Fig. 4a). This was primarily due to fumarate, whose correlation to $\log_{10}$(biomass) at 23ºC was significant ($r = 0.60$, $P < 0.001$), albeit weaker than at 5ºC. Neither malate nor glutamine, on the other hand, exhibited comparable trends in relation to biomass at the two temperatures. Similarities between the biomass-correlations of 5ºC and 23ºC fingerprints were therefore incomplete.

Expression of fumarate-related transcripts in SA mutants under chilling

In addition to the formation of fumarate via mitochondrial succinate dehydrogenase (SDH), its diurnal accumulation in *Arabidopsis* is catalyzed by a recently characterized cytosolic fumarase, FUM2 (Pracharoenwattana et al. 2010). qRT-PCR analyses of relevant transcripts were consistent with the latter being responsible for the genotype-dependent fumarate levels at 5°C. Expression of *FUM2* (At5g50950) was strongly elevated in the high-fumarate *NahG* and *sid2* phenotypes, and strongly reduced in the low-fumarate *cpr1* phenotype, relative to the wild-type (Fig. 6). Consequently, *FUM2* transcript levels correlated strongly with biomass of the shoots in which they were quantified. This pattern was not typical of the genes examined (Fig. 6). Negative correlation with biomass, for example, was observed for the *AtPPC1* (At1g53310) transcripts of phosphoenolpyruvate carboxylase (Gregory et al. 2009), which represents a cytosolic route to malate production.

Substantial proportions of cellular fumarate and malate occur in *Arabidopsis* vacuoles, and the tonoplast dicarboxylate transporter *At*TDT (At5g47560) is one of the transport mechanisms (Hurth et al. 2005, Schulze et al. 2012). However, no significant differences or trends in *AtTDT* expression were found (Fig. 6). Regulation of this transporter might, of course, be post-translational.

Transcript analyses did not implicate mitochondrial metabolism in the variations in fumarate. Expression of *SDH1-1* (At5g66760), the major transcript for the succinate-binding flavoprotein subunit of SDH (León et al. 2007), was greater in the low-fumarate *cpr1* mutants, and negatively correlated with biomass (Fig. 6). The *FUM1* mitochondrial fumarase gene, At2g47510, showed no significant differences in expression (Fig. 6). The mitochondrial *SFC1* succinate/fumarate (At5g01340) and *DIC1* dicarboxylate (At2g22500) carrier transcripts (Palmieri et al. 2011) showed negative correlation with biomass (Fig. 6). The mitochondrial NAD-malic enzyme was studied, as effects on fumarate have been found in *Arabidopsis* plants lacking this function (Tronconi et al. 2008),

or over-expressing a maize malic enzyme (Fahnenstich et al. 2007). Significant variation was not seen for *AtNAD-ME2* (At4g00570), which encodes a subunit of this dimeric enzyme (Fig. 6).

Discussion

NMR fingerprinting revealed distinct metabolic phenotypes associated with the divergent morphologies of SA mutants under chilling. Fumarate and malate, prominent as NMR-detectable metabolites in *Arabidopsis*, differed in opposite directions in high- or low-SA shoots. SA-deficient phenotypes, with greater biomass at 5ºC (Scott et al. 2004), had abundant fumarate and low malate, relative to the Col-0 wild-type. The opposite held for the high-SA *cpr1* phenotype.

Of technical interest for breeding crops with increased biomass (Meyer et al. 2007, Steinfath et al. 2010, Sulpice et al. 2013) was our use of NMR fingerprints to model shoot biomass. This supports the credibility of spectral fingerprinting (Scott et al. 2013, Scott et al. 2010), as a alternative to more time-consuming chromatography (Lisec et al. 2008, Sulpice et al. 2013).

Our metabolomic models were biomass-correlated PCs, and PLS models of relative biomass. Models should not be uncritically seen as biologically meaningful (Broadhurst and Kell 2006). Validity of our models was supported by consistent relations to biomass in different experiments, and by predictive testing. PLS models of NMR fingerprints from one experiment predicted the relative biomass of shoots in other experiments. Relative biomass of shoots of one genotype could be predicted from other genotypes.

There is insufficient current evidence to generalize about particular metabolites in relation to different growth conditions (Sulpice et al. 2013). Other studies modeling *Arabidopsis* biomass from metabolites have focused on natural genetic variation or heterosis (Meyer et al. 2012, Sulpice et al. 2013), instead of dramatic hormonal mutants. This may be why we found relatively strong biomass-correlations of individual metabolites. One commonality with other studies was the preponderance of negative biomass-correlations among metabolites (Sulpice et al. 2013, Sulpice et al. 2009). Prominent metabolites negatively associated with biomass in this study and others were malate (Meyer et al. 2007, Sulpice et al. 2013) and glutamine (Meyer et al. 2007). Most compounds with negative biomass-correlation identified by Meyer *et al.* (2007) were intermediates of central metabolism, which they inferred was depleted by faster growth, rather than being driven by increased substrates. Accordingly, Cross *et al.* (2006) found accessions with larger rosettes had higher central metabolism enzyme activities, perhaps contributing to faster carbon utilization and growth. SA-deficient mutants at 5ºC had higher net assimilation rates than wild-type (Scott et al. 2004), so their generally weaker metabolite signals may have reflected faster conversion to structural biomass.

Evidence has emerged for negative impacts of apoplastic malate (and fumarate) on growth, via reduction of stomatal apertures (Araújo et al. 2011). Malate contents were about 45% higher in *cpr1* than the wild-type at 5°C, while combined levels of malate and fumarate were about 19% higher. Given the more potent stomatal effect of malate (Araújo et al. 2011), and depending on how these whole-shoot levels manifested in the relevant apoplastic compartment, it is conceivable that stomatal effects of metabolites might contribute to the biomass phenotypes.

There is also evidence that carboxylates, particularly citrate, can influence the *Arabidopsis* leaf transcriptome via unknown signaling pathways (Finkemeier et al. 2013). The lack of correlation with biomass suggested citrate was not a determinant of the SA chilling phenotypes. Potentially more relevant to the present study was the finding of Finkemeier *et al.* (2013) that malate treatment had effects on the transcriptome.

Whereas previous studies reported a negative biomass-correlation for sucrose, we found the opposite. This difference may have been related to the chilling conditions in our study, given the roles of sucrose in low temperature physiology (Rekarte-Cowie et al. 2008).

Our most distinctive finding, however, was the strong positive biomass-correlation of fumarate. Other studies found negative or non-significant biomass-correlation for fumarate (Cross et al. 2006, Meyer et al. 2007, Sulpice et al. 2013, Sulpice et al. 2009). Given the precursor/product relations of malate and fumarate, their pronounced opposite variation between the SA chilling phenotypes suggested a single causal enzyme. Several potentially relevant transcripts were therefore examined.

The positive biomass-correlation of 0.84 for transcripts of the cytosolic fumarase FUM2 matched that of fumarate (0.90), its proposed product, and was the reverse of malate (-0.83), its proposed substrate. There were evident parallels in the characterization of FUM2 by Pracharoenwattana *et al.* (2010): in *fum2* mutants, they found ten-fold less fumarate, but twice as much malate compared to the wild-type. Further evidence for the importance of *FUM2* is that At5g50950 mapped within the sole quantitative trait locus for fumarate content identified in *Arabidopsis* by Lisec *et al.* (2008).

While transcript abundance does not necessarily reflect enzyme activity, the *FUM2* pattern we observed was distinct from other transcripts quantified. We found no comparable evidence, for causal roles in the variations in malate and fumarate, either for the FUM1 mitochondrial fumarase, or for SDH, the mitochondrial fumarate source. Transcriptome evidence to compare the two fumarases of *Arabidopsis* was not available, as *FUM2* and *FUM1* are detected by the same microarray probe (e.g., '248461_s_at' on the Affymetrix ATH1 GeneChip).

Pracharoenwattana *et al.* (2010) found a fumarate-related effect on biomass in high-nitrogen conditions, where *fum2* mutants grew abnormally slowly. High-nitrogen treatment increased fumarate several-fold in wild-type but not *fum2* leaves, while malate increased to a greater extent in *fum2* mutants. Moreover, nitrogen limitation has been found to reduce the fumarate:malate ratio in wild-type *Arabidopsis* (Kant et al. 2008, Tschoep et al. 2009). Pracharoenwattana *et al.* (2010) suggested fumarate's greater acidity, and involvement in fewer metabolic processes, than malate might be advantages for a function in balancing proton consumption during nitrate reduction. Fumarate accumulation via FUM2 activity was thereby proposed to promote growth by facilitating nitrogen assimilation. If nitrogen assimilation were a factor in the differential growth of SA mutants at 5°C, the higher glutamine levels in the low-biomass *cpr1* mutants would not be inconsistent with observations on *Arabidopsis* in sustained nitrogen-limitation (Tschoep et al. 2009).

Evidence supports a role for fumarate as an alternative transient carbon store to starch. Fumarate-deficient *fum2* leaves accumulated twice as much starch as wild-type (Pracharoenwattana et al. 2010), while starchless *pgm* mutants accumulated excess fumarate (Chia et al. 2000). In wild-type *Arabidopsis*, respiratory quotients indicated a shift to organic acid substrates by night end, which was seen in short days from the low-biomass, carbon-starvation phenotype of plants unable to accumulate fumarate or malate due to a maize NADP-malic enzyme transgene (Zell et al. 2010).

Schulze *et al*. (2012) found cold acclimation in *Arabidopsis* led to higher vacuolar concentrations of fumarate and malate, with increased abundance of the tonoplast dicarboxylate transporter *At*TDT. We did not, however, find relevant variation in *AtTDT* transcripts between SA mutants at 5°C.

Another transcript that did not vary significantly between genotypes was *AtNAD-ME2*. *Arabidopsis* mutants lacking mitochondrial NAD-malic enzyme activity have reduced fumarate levels, though on the other hand no effects on growth were observed (Tronconi et al. 2008).

Several transcripts showed negative biomass-correlation, due primarily to higher expression in the stunted, high-malate *cpr1* phenotype. One was a phosphoenolpyruvate carboxylase (*AtPPC1*), which might have contributed to cytosolic malate production, subject to the multiple metabolic controls of this enzyme (Gregory et al. 2009).

The negative biomass-correlation of *SDH1-1* transcripts was not consistent with SDH being directly responsible for the opposite trend in its fumarate product. It has been observed in *Arabidopsis* (Fuentes et al. 2011) and tomato (Araújo et al. 2011) that genetic impairment of SDH resulted in increased biomass. Evidence in tomato suggested this was mediated by the effects of apoplastic malate and fumarate on stomatal aperture (Araújo et al. 2011).

Given the strong profile of fumarate in our metabolic phenotypes, we investigated the mitochondrial carrier SFC1, which exchanges fumarate for succinate (Catoni et al. 2003). Developmental expression of *SFC1* has suggested a role in import of succinate produced outside the mitochondrion (Catoni et al. 2003). Generally, *SFC1* was less highly expressed than the other transcripts analyzed, but a negative biomass-correlation, with higher expression in the stunted *cpr1* phenotype, was again observed.

Negative biomass-correlation was also seen for transcripts of *DIC1*, a mitochondrial dicarboxylate carrier (Palmieri et al. 2008). DIC1 substrates include malate and oxaloacetate, making it a mitochondrion/cytosol 'redox shuttle' candidate. It may also be involved in mitochondrial import of dicarboxylic acids as respiratory substrates, though not fumarate (Palmieri et al. 2008).

DIC1 expression was high and variable, in accordance with its reported strong regulation in mitochondrial stress (Palmieri et al. 2011, Van Aken et al. 2009). The broad stress responses of this and other mitochondrial carriers suggest increased exchange of substrates under stress (Van Aken et al. 2009). The extremely high salicylate levels in *cpr1* mutants at 5°C (Scott et al. 2004) might provoke a mitochondrial stress response. The roles of SA as a stress signal (Rivas-San Vicente and Plasencia 2011) include induction of the mitochondrial alternative oxidase pathway (Norman et al.

2004). Moreover, at sufficient concentrations, SA might act as a mitochondrial stress factor in its own right, via uncoupling of respiration and inhibition of electron transport from SDH to the ubiquinone pool (Norman et al. 2004). Direct evidence of stress, as oxidative damage, has been described in the *cpr1* chilling phenotype (Scott et al. 2004).

In conclusion, identifiable metabolic phenotypes were found to characterize the abnormal morphologies of SA mutants under chilling. This was established using metabolomic models that could predict shoot biomass. Such models are, of course, correlative and do not identify underlying mechanisms. Different metabolomic techniques (Scott et al. 2010) would, moreover, undoubtedly characterize further metabolites showing biomass-related variation in this system. Nonetheless, the power of NMR fingerprint-modeling to guide future research was evident in its vivid demonstration of the importance of fumarate in *Arabidopsis* physiology.

References

Araújo WL, Nunes-Nesi A, Osorio S, Usadel B, Fuentes D, Nagy R, Balbo I, Lehmann M, Studart-Witkowski C, Tohge T, Martinoia E, Jordana X, DaMatta FM and Fernie AR (2011) Antisense inhibition of the iron-sulphur subunit of succinate dehydrogenase enhances photosynthesis and growth in tomato via an organic acid-mediated effect on stomatal aperture. *Plant Cell* 23: 600-627

Boyes DC, Zayed AM, Ascenzi R, McCaskill AJ, Hoffman NE, Davis KR 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

Broadhurst DI and Kell DB (2006) Statistical strategies for avoiding false discoveries in metabolomics and related experiments. *Metabolomics* 2: 171-196

Catoni E, Schwab R, Hilpert M, Desimone M, Schwacke R, Flügge U-I, Schumacher K and Frommer WB (2003) Identification of an *Arabidopsis* mitochondrial succinate-fumarate translocator. *FEBS Lett* 534: 87-92

Chia DW, Yoder TJ, Reiter W-D and Gibson SI (2000) Fumaric acid: an overlooked form of fixed carbon in *Arabidopsis* and other plant species. *Planta* 211: 743-751

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

Cross JM, von Korff M, Altmann T, Bartzetko L, Sulpice R, Gibon Y, Palacios N and Stitt M (2006) Variation of enzyme activities and metabolite levels in 24 Arabidopsis accessions growing in carbon-limited conditions. *Plant Physiol* 142: 1574-1588

Czechowski T, Stitt M, Altmann T, Udvardi MK and Scheible WR (2005) Genome-wide identification and testing of superior reference genes for transcript normalization in Arabidopsis. *Plant Physiol* 139: 5-17

Enot DP, Lin W, Beckmann M, Parker D, Overy DP and Draper J (2008) Preprocessing, classification modeling and feature selection using flow injection electrospray mass spectrometry metabolite fingerprint data. *Nature Protoc* 3: 446-470

Enquist BJ, Kerkhoff AJ, Huxman TE and Economo EP (2007) Adaptive differences in plant physiology and ecosystem paradoxes: insights from metabolic scaling theory. *Global Change Biol* 13: 591-609

Fahnenstich H, Saigo M, Niessen M, Zanor MI, Andreo CS, Fernie AR, Drincovich MF, Flügge U-I and Maurino VG (2007) Alteration of organic acid metabolism in Arabidopsis overexpressing the maize C_4 NADP-malic enzyme causes accelerated senescence during extended darkness. *Plant Physiol* 145: 640-652

Finkemeier I, König A-C, Heard W, Nunes-Nesi A, Pham PA, Leister D, Fernie AR and Sweetlove LJ

(2013) Transcriptomic analysis of the role of carboxylic acids in metabolite signaling in Arabidopsis leaves. *Plant Physiol* 162: 239-253

Fuentes D, Meneses M, Nunes-Nesi A, Araújo WL, Tapia R, Gómez I, Holuigue L, Gutiérrez RA, Fernie AR and Jordana X (2011) A deficiency in the flavoprotein of Arabidopsis mitochondrial complex II results in elevated photosynthesis and better growth in nitrogen-limiting conditions. *Plant Physiol* 157: 1114-1127

Gilmour SJ, Sebolt AM, Salazar MP, Everard JD and Thomashow MF (2000) Overexpression of the Arabidopsis *CBF3* transcriptional activator mimics multiple biochemical changes associated with cold acclimation. *Plant Physiol* 124: 1854-1865

Gregory AL, Hurley BA, Tran HT, Valentine AJ, She YM, Knowles VL and Plaxton WC (2009) *In vivo* regulatory phosphorylation of the phosphoenolpyruvate carboxylase AtPPC1 in phosphate-starved *Arabidopsis thaliana*. *Biochem J* 420: 57-65

Griffith C, Kim E and Donohue K (2004) Life-history variation and adaptation in the historically mobile plant *Arabidopsis thaliana* (Brassicaceae) in North America. *Am J Bot* 91: 837-849

Griffith M, Timonin M, Wong ACE, Gray GR, Akhter SR, Saldanha M, Rogers MA, Weretilnyk EA and Moffatt B (2007) *Thellungiella*: an *Arabidopsis*-related model plant adapted to cold temperatures. *Plant Cell Environ* 30: 529-538

Hammer Ø, Harper DAT and Ryan PD (2001) PAST: paleontological statistics software package for education and data analysis. *Palaeontol Electron* 4(1): art. 4

Humphreys MW, Yadav RS, Cairns AJ, Turner LB, Humphreys J and Skøt L (2006) A changing climate for grassland research. *New Phytol* 169: 9-26

Hurth MA, Suh SJ, Kretzschmar T, Geis T, Bregante M, Gambale F, Martinoia E and Neuhaus HE (2005) Impaired pH homeostasis in Arabidopsis lacking the vacuolar dicarboxylate transporter and analysis of carboxylic acid transport across the tonoplast. *Plant Physiol* 137: 901-910

Janes KA and Yaffe MB (2006) Data-driven modelling of signal-transduction networks. *Nat Rev Mol Cell Biol* 7: 820-828

Kant S, Bi Y-M, Weretilnyk E, Barak S and Rothstein SJ (2008) The Arabidopsis halophytic relative *Thellungiella halophila* tolerates nitrogen-limiting conditions by maintaining growth, nitrogen uptake, and assimilation. *Plant Physiol* 147: 1168-1180

Kaplan F, Kopka J, Haskell DW, Zhao W, Schiller KC, Gatzke N, Sung DY and Guy CL (2004) Exploring the temperature-stress metabolome of Arabidopsis. *Plant Physiol* 136: 4159-4168

Kim Y, Park S, Gilmour SJ and Thomashow MF (2013) Roles of CAMTA transcription factors and salicylic acid in configuring the low-temperature transcriptome and freezing tolerance of Arabidopsis. *Plant J* 75: 364-376

Klotke J, Kopka J, Gatzke N and Heyer AG (2004) Impact of soluble sugar concentrations on the acquisition of freezing tolerance in accessions of *Arabidopsis thaliana* with contrasting cold adaptation - evidence for a role of raffinose in cold acclimation. *Plant Cell Environ* 27: 1395-

1404

León G, Holuigue L and Jordana X (2007) Mitochondrial complex II Is essential for gametophyte development in Arabidopsis. *Plant Physiol* 143: 1534-1546

Lisec J, Meyer RC, Steinfath M, Redestig H, Becher M, Witucka-Wall H, Fiehn O, Törjék O, Selbig J, Altmann T and Willmitzer L (2008) Identification of metabolic and biomass QTL in *Arabidopsis thaliana* in a parallel analysis of RIL and IL populations. *Plant J* 53: 960-972

Lohse M, Nunes-Nesi A, Krüger P, Nagel A, Hannemann J, Giorgi FM, Childs L, Osorio S, Walther D, Selbig J, Sreenivasulu N, Stitt M, Fernie AR and Usadel B (2010) Robin: an intuitive wizard application for R-based expression microarray quality assessment and analysis. *Plant Physiol* 153: 642-651

Meyer RC, Steinfath M, Lisec J, Becher M, Witucka-Wall H, Torjek O, Fiehn O, Eckardt A, Willmitzer L, Selbig J and Altmann T (2007) The metabolic signature related to high plant growth rate in *Arabidopsis thaliana*. *Proc Natl Acad Sci USA* 104: 4759-4764

Meyer RC, Witucka-Wall H, Becher M, Blacha A, Boudichevskaia A, Dörmann P, Fiehn O, Friedel S, von Korff M, Lisec J, Melzer M, Repsilber D, Schmidt R, Scholz M, Selbig J, Willmitzer L and Altmann T (2012) Heterosis manifestation during early Arabidopsis seedling development is characterized by intermediate gene expression and enhanced metabolic activity in the hybrids. *Plant J* 71: 669-683

Norman C, Howell KA, Millar AH, Whelan JM and Day DA (2004) Salicylic acid is an uncoupler and inhibitor of mitochondrial electron transport. *Plant Physiol* 134: 492-501

Palmieri F, Pierri CL, De Grassi A, Nunes-Nesi A and Fernie AR (2011) Evolution, structure and function of mitochondrial carriers: a review with new insights. *Plant J* 66: 161-181

Palmieri L, Picault N, Arrigoni R, Besin E, Palmieri F and Hodges M (2008) Molecular identification of three *Arabidopsis thaliana* mitochondrial dicarboxylate carrier isoforms: organ distribution, bacterial expression, reconstitution into liposomes and functional characterization. *Biochem J* 410: 621-629

Pracharoenwattana I, Zhou W, Keech O, Francisco PB, Udomchalothorn T, Tschoep H, Stitt M, Gibon Y and Smith SM (2010) Arabidopsis has a cytosolic fumarase required for the massive allocation of photosynthate into fumaric acid and for rapid plant growth on high nitrogen. *Plant J* 62: 785-795

Rekarte-Cowie I, Ebshish OS, Mohamed KS and Pearce RS (2008) Sucrose helps regulate cold acclimation of *Arabidopsis thaliana*. *J Exp Bot* 59: 4205-4217

Rivas-San Vicente M and Plasencia J (2011) Salicylic acid beyond defence: its role in plant growth and development. *J Exp Bot* 62: 3321-3338

Ruijter JM, Ramakers C, Hoogaars WMH, Karlen Y, Bakker O, van den Hoff MJB and Moorman AFM (2009) Amplification efficiency: linking baseline and bias in the analysis of quantitative PCR data. *Nucleic Acids Res* 37: 12

Schulze WX, Schneider T, Starck S, Martinoia E and Trentmann O (2012) Cold acclimation induces changes in Arabidopsis tonoplast protein abundance and activity and alters phosphorylation of tonoplast monosaccharide transporters. *Plant J* 69: 529-541

Scott IM, Clarke SM, Wood JE and Mur LAJ (2004) Salicylate accumulation inhibits growth at chilling temperature in Arabidopsis. *Plant Physiol* 135: 1040-1049

Scott IM, Lin W, Liakata M, Wood JE, Vermeer CP, Allaway D, Ward JL, Draper J, Beale MH, Corol DI, Baker JM and King RD (2013) Merits of random forests emerge in evaluation of chemometric classifiers by external validation. *Anal Chim Acta* 801: 22-33

Scott IM, Vermeer CP, Liakata M, Corol DI, Ward JL, Lin W, Johnson HE, Whitehead L, Kular B, Baker JM, Walsh S, Larson TR, Graham IA, Wang TL, King RD, Draper J and Beale MH (2010) Enhancement of plant metabolite fingerprinting by machine learning. *Plant Physiol* 153: 1506-1520

Steinfath M, Gärtner T, Lisec J, Meyer R, Altmann T, Willmitzer L and Selbig J (2010) Prediction of hybrid biomass in *Arabidopsis thaliana* by selected parental SNP and metabolic markers. *Theor Appl Genet* 120: 239-247

Steinfath M, Groth D, Lisec J and Selbig J (2008) Metabolite profile analysis: from raw data to regression and classification. *Physiol Plant* 132: 150-161

Strawn MA, Marr SK, Inoue K, Inada N, Zubieta C and Wildermuth MC (2007) *Arabidopsis* isochorismate synthase functional in pathogen-induced salicylate biosynthesis exhibits properties consistent with a role in diverse stress responses. *J Biol Chem* 282: 5919-5933

Sulpice R, Nikoloski Z, Tschoep H, Antonio C, Kleessen S, Larhlimi A, Selbig J, Ishihara H, Gibon Y, Fernie AR and Stitt M (2013) Impact of the carbon and nitrogen supply on relationships and connectivity between metabolism and biomass in a broad panel of Arabidopsis accessions. *Plant Physiol* 162: 347-363

Sulpice R, Pyl E-T, Ishihara H, Trenkamp S, Steinfath M, Witucka-Wall H, Gibon Y, Usadel B, Poree F, Piques MC, Von Korff M, Steinhauser MC, Keurentjes JJB, Guenther M, Hoehne M, Selbig J, Fernie AR, Altmann T and Stitt M (2009) Starch as a major integrator in the regulation of plant growth. *Proc Natl Acad Sci USA* 106: 10348-10353

Tronconi MA, Fahnenstich H, Gerrard Wheeler MC, Andreo CS, Flügge U-I, Drincovich MF and Maurino VG (2008) Arabidopsis NAD-malic enzyme functions as a homodimer and heterodimer and has a major impact on nocturnal metabolism. *Plant Physiol* 146: 1540-1552

Tschoep H, Gibon Y, Carillo P, Armengaud P, Szecowka M, Nunes-Nesi A, Fernie AR, Koehl K and Stitt M (2009) Adjustment of growth and central metabolism to a mild but sustained nitrogen-limitation in *Arabidopsis*. *Plant Cell Environ* 32: 300-318

Van Aken O, Zhang B, Carrie C, Uggalla V, Paynter E, Giraud E and Whelan J (2009) Defining the mitochondrial stress response in *Arabidopsis thaliana*. *Mol Plant* 2: 1310-1324

Vandesompele J, De Preter K, Pattyn F, Poppe B, Van Roy N, De Paepe A and Speleman F (2002)

Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biol* 3: 12

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

Wold S, Sjöström M and Eriksson L (2001) PLS-regression: a basic tool of chemometrics. *Chemom Intell Lab Syst* 58: 109-130

Yang S and Hua J (2004) A haplotype-specific *Resistance* gene regulated by *BONZAI1* mediates temperature-dependent growth control in Arabidopsis. *Plant Cell* 16: 1060-1071

Ye J, Coulouris G, Zaretskaya I, Cutcutache I, Rozen S and Madden T (2012) Primer-BLAST: A tool to design target-specific primers for polymerase chain reaction. *BMC Bioinformatics* 13: 134

Zell MB, Fahnenstich H, Maier A, Saigo M, Voznesenskaya EV, Edwards GE, Andreo C, Schleifenbaum F, Zell C, Drincovich MF and Maurino VG (2010) Analysis of Arabidopsis with highly reduced levels of malate and fumarate sheds light on the role of these organic acids as storage carbon molecules. *Plant Physiol* 152: 1251-1262

Supporting Information

Additional Supporting Information may be found in the online version of this article:

Appendix S1. NMR data for experiments A-G.

Table 1. Plant growth conditions for metabolomics experiments. Incubation period was from mean growth stage 1.04 (when seedlings were transferred to 5ºC, or kept at 23ºC), until harvest at mean growth stage 5.10 (coincident for all genotypes at 5°C except the infertile *cpr1* mutants). Experiment G was harvested at growth stage 6.00-6.10.

Experiment	Temperature	Light (μmol m^{-2} s^{-1})	Incubation (d)
A	5ºC	100	82
B		55	42
C		25	100
D		25	98
E	23ºC	55	22
F		25	31
G		25	50

Table 2. Metabolite quantifications in relation to biomass and genotype. Correlations of metabolite content and $\log_{10}$(biomass) shown as Pearson coefficients (*r*) for 5°C experiments (asterisks: minimum significance among experiments A-D). Metabolite contents: asterisks denote significant difference from Col-0 (in Kruskal-Wallis and Bonferroni-corrected Mann-Whitney tests on pooled data of all experiments at stated temperature). ***$P < 0.001$; **$P < 0.01$; *$P < 0.05$.

	Fumarate	Sucrose	Malate	Glutamine
	Mean *r* $\log_{10}$(biomass) at 5°C ± SD (*n* = 4, experiments A-D)			
	0.90 ± 0.05***	0.71 ± 0.19*	-0.83 ± 0.21***	-0.91 ± 0.11***
Genotype	Mean contents (mg g^{-1} DW) at 5°C ± SD (*n* = 4, experiments A-D)			
NahG	0.95 ± 0.09***	0.65 ± 0.12	1.18 ± 0.27***	0.31 ± 0.06*
eds5	0.97 ± 0.03***	0.74 ± 0.11**	1.36 ± 0.27***	0.33 ± 0.06*
sid2	0.95 ± 0.06***	0.70 ± 0.08**	1.40 ± 0.44*	0.33 ± 0.09
Col-0	0.45 ± 0.29	0.61 ± 0.10	1.77 ± 0.43	0.36 ± 0.09
cpr1	0.07 ± 0.02***	0.45 ± 0.14***	2.57 ± 0.25***	0.68 ± 0.15***
	Mean contents (mg g^{-1} DW) at 23°C ± SD (*n* = 3, experiments E-G)			
NahG	0.62 ± 0.22	0.16 ± 0.02	0.50 ± 0.06	0.17 ± 0.03
Col-0	0.49 ± 0.14	0.15 ± 0.01	0.50 ± 0.01	0.16 ± 0.03

Figure legends

Fig. 1. Morphological chilling-phenotypes. (a) SA over-accumulator (*cpr1*) and SA-deficient (*sid2*, *eds5*, *NahG*) mutants, with Col-0, after 60 d at 5ºC in 55 μmol m^{-2} s^{-1}. Bar: 1 cm. (b) Mean biomass of shoot vegetative tissues relative to Col-0 in light (μmol m^{-2} s^{-1}) and temperatures indicated. Symbols: 5°C, colored (100 μmol m^{-2} s^{-1}, blue; 55 μmol m^{-2} s^{-1}, green; 25 μmol m^{-2} s^{-1}, red); 23°C, black and white (55 μmol m^{-2} s^{-1}, large; 25 μmol m^{-2} s^{-1}, small); Col-0, filled circles; *NahG*, open circles; *eds5*, triangles; *sid2*, squares; *cpr1*, crosses. Bars: ± SE (*n* = 10). Data from experiments A, B, C, E, F (Table 1). Asterisks: significant difference from Col-0 (Kruskal-Wallis and Bonferroni-corrected Mann-Whitney tests): ***P < 0.001; *P < 0.05.

Fig. 2. Effects of temperature and genotype on metabolic phenotypes revealed by ^{1}H-NMR. Spectra are mean (*n* ≥ 9) fingerprints (binned to 0.01 ppm) of 23 or 5ºC shoots (experiments B and E). Spectral intensities scaled to d$_4$-TSP internal standard (δ0.0). Colored labels: significant associations with 5ºC phenotype in ANOVA (higher in greater-biomass phenotypes, blue; higher in low-biomass phenotypes, red), among 100 strongest signals. (a) Col-0 at 23ºC. (b) SA-deficient mutants (*sid2*, *eds5*, *NahG*) at 5ºC. *F* values compare fingerprints of these mutants with Col-0 in experiments A-D. (c) Col-0 at 5ºC. (d) *cpr1* at 5ºC. *F* values compare *cpr1* and Col-0 fingerprints in experiments A-D. Bonferroni-corrected significances: ***P < 10^{-5}; **P < 10^{-4}; *P < 5×10^{-4}. Ala, alanine; Glc, glucose; Gln, glutamine; Glu, glutamate; Ile, isoleucine; Leu, leucine; Pro, proline; Suc, sucrose; Thr, threonine; Val, valine. Experiment conditions are in Table 1.

Fig. 3. PCA of genotype and environment effects on metabolic phenotypes. (a) Scores of first two PCs for NMR fingerprints of shoots grown at 5ºC in 100 (blue), 55 (green) or 25 (red) μmol m^{-2} s^{-1}, or at 23ºC in 55 (black and white, large symbols) or 25 (black and white, small symbols) μmol m^{-2} s^{-1}. Variance accounted for by PCs is on axis labels. (b) PC[1] loadings (arrows: directions separating 23ºC and 5ºC plants). (c) PC[2] loadings (arrow: direction of higher biomass). Glc, glucose; Gln, glutamine; Glu, glutamate; Pro, proline; Suc, sucrose. Data from experiments A, B, C, E, and F. See Fig. 1b for relative biomass of shoots.

Fig. 4. Biomass-correlated models of metabolic phenotypes. (a) Correlation between biomass, at mean growth stage 5.10, and PC[2] scores (30% of variance) from PCA of fingerprints from experiments A-G. Abscissa shows shoot DW transformed to log$_{10}$, without any normalization between experiments. (b) Cross-experiment predictions of 5ºC biomass by PLS-regression. A PLS model (two components) was built on fingerprints from experiment D, and its predictions of biomass rankings in each of experiments A-C plotted against actual data. Ranks are descending and normalized by number of samples in experiment (largest shoots = 1.0). (c) Cross-genotype predictions of 5ºC biomass by PLS-regression. Each genotype in each of experiments A-D was held out in turn, and its biomass data

predicted by models of the other four genotypes. The plot of predicted and actual relative biomass (largest shoots = 1.0) for all genotypes and experiments derives from 20 models. Key: 100 (blue), 55 (green) or 25 (red) μmol m^{-2} s^{-1} at 5°C; or 55 (black and white, large symbols) or 25 (black and white, small symbols) μmol m^{-2} s^{-1} at 23°C. r, Pearson coefficients (***$P < 10^{-32}$); when PLS models were recalculated with randomized biomass data, r values were < 0.1 ($P > 0.3$).

Fig. 5. Coefficients of PLS-regression model of biomass built on NMR fingerprints from 5°C experiment D. Arrow: higher biomass. Glc, glucose; Gln, glutamine; Suc, sucrose. See Figure 4b for predictions by this model.

Fig. 6. Mean relative abundances of transcripts in Col-0, SA-deficient (*NahG*, *sid2*), and SA over-accumulator (*cpr1*) shoots, after 62 d at 5°C. Values (arbitrary units) were normalized using geometric means of two endogenous reference genes (*EF-1α*, *PP2A3*). Error bars show ± SE ($n = 4$ shoots), with asterisks for significant difference to Col-0 (one-way ANOVA with Tukey tests). Pearson coefficients (r): correlation of transcript abundance with $\log_{10}$(biomass) of shoots, with asterisks indicating significance. ***$P < 0.001$; **$P < 0.01$; *$P < 0.05$.
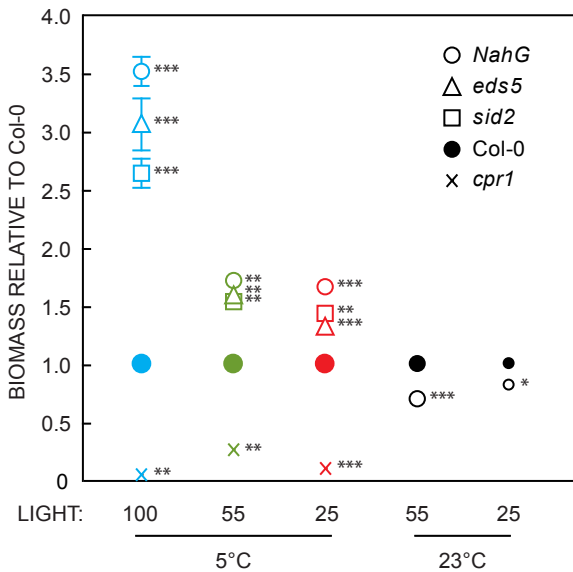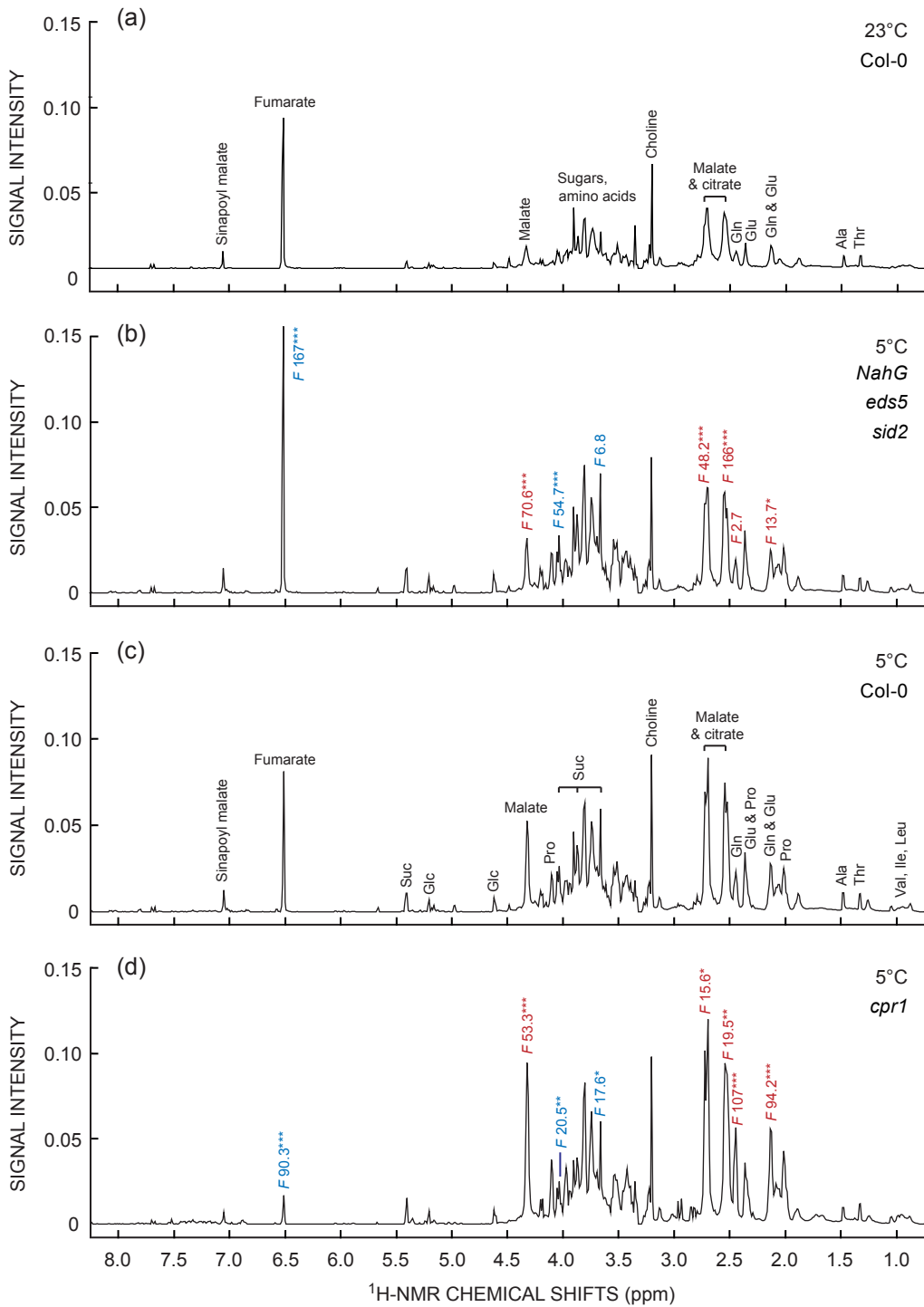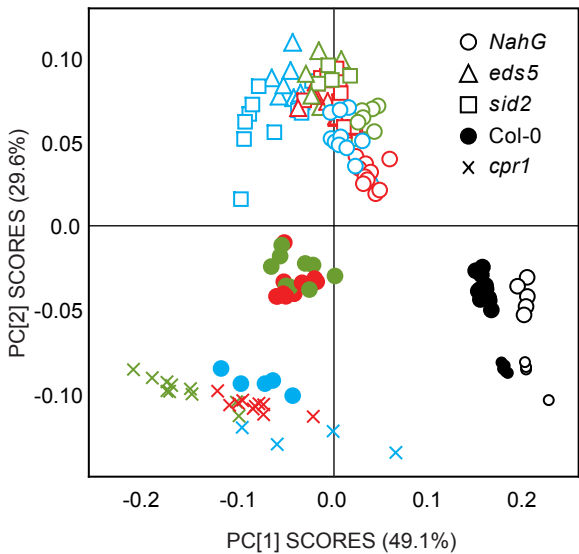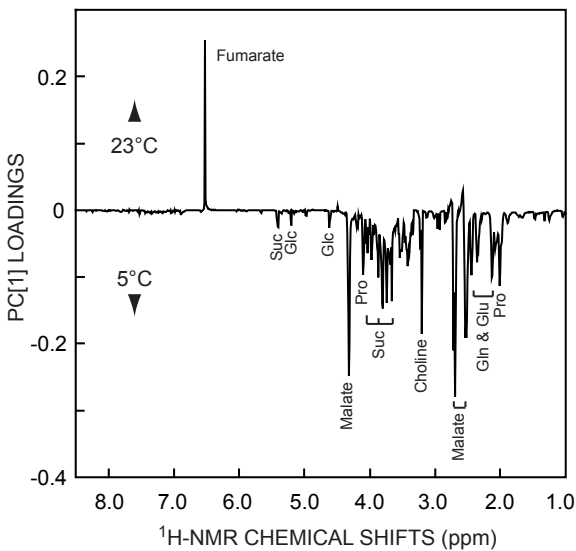
Figure 1

(a)

(b)
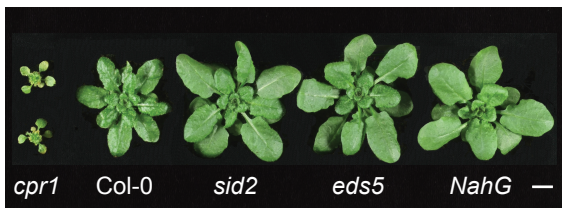
Figure 2

Figure 3
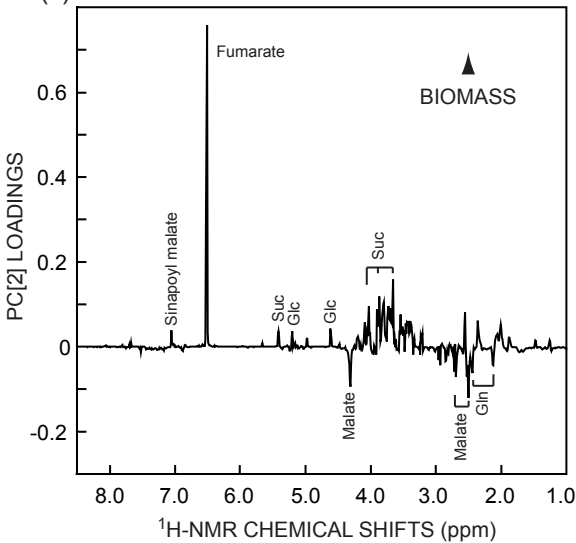
Figure 4

Figure 5

Figure 6