universität wien

DISSERTATION / DOCTORAL THESIS

Titel der Dissertation /Title of the Doctoral Thesis

„Reprogramming of plant metabolism in a changing environment: mathematical analysis and experimental quantification“

verfasst von / submitted by

Mag. rer. nat. Lisa Fürtauer

angestrebter akademischer Grad / in partial fulfilment of the requirements for the degree of

Doctor of Philosophy (PhD)

Wien, 2018 / Vienna 2018

Studienkennzahl lt. Studienblatt:	A 794 685 437
Dissertationsgebiet lt. Studienblatt:	Biologie
Betreut von / Supervisor:	Univ.-Prof. Dr. Wolfram Weckwerth
	Univ.-Prof. Dr. Thomas Nägele

Acknowledgements

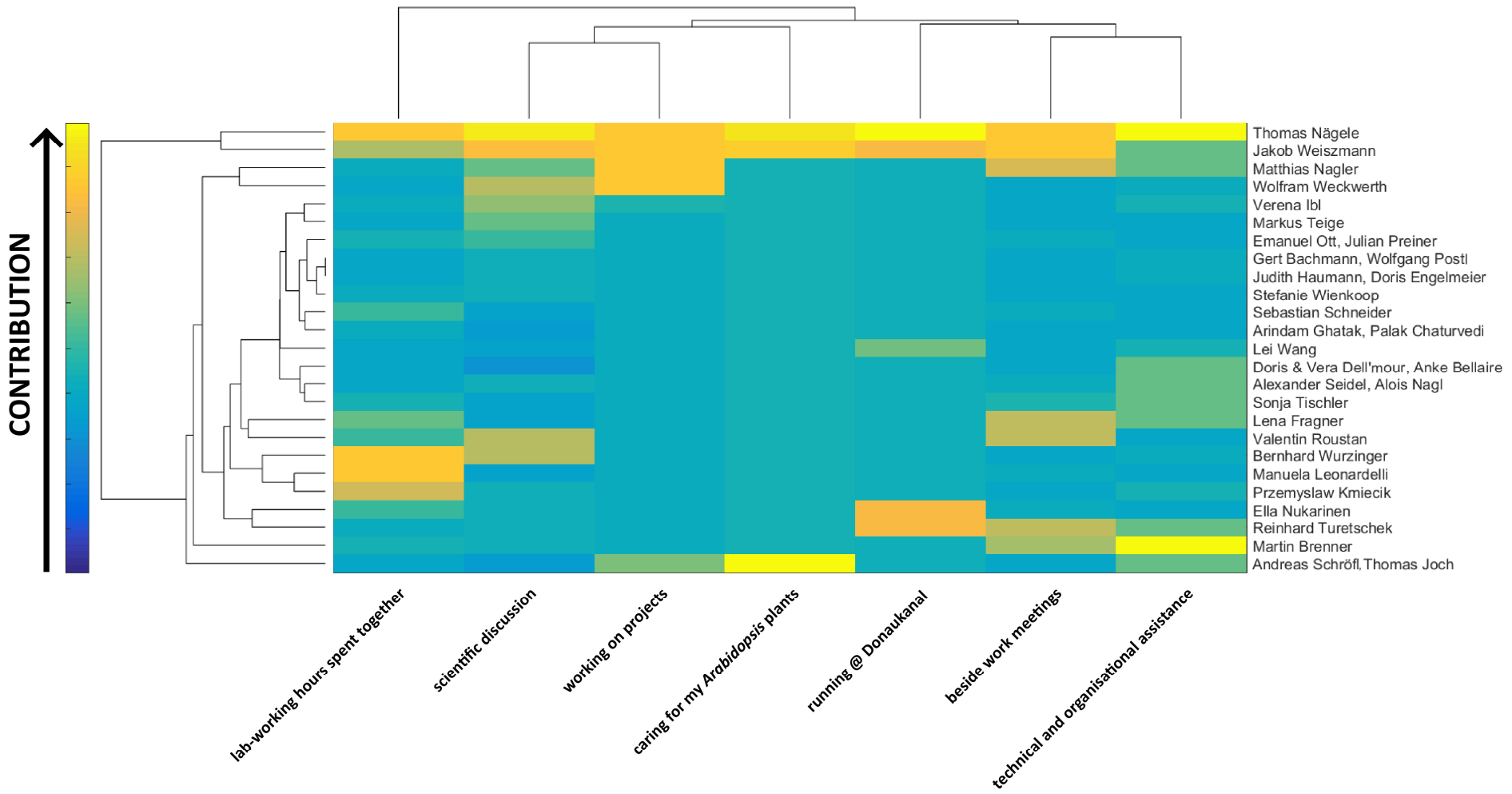
I would like to thank everyone who supported me throughout the whole thesis. First of all I want to express my gratitude to my both supervisors Univ.-Prof. Dr. Wolfram Weckwerth and Univ.-Prof. Dr. Thomas Nägele. I thank especially Wolfram for his support, for providing me the opportunity to work in his laboratory and to be a member of his department. I thank Thomas for his support, motivation, scientific viewpoints, (hard) biological-mathematical discussions, for trusting in me and my working style, and overall for our friendship.

I want to thank Univ.-Prof. Dr. Thomas Sauter and Univ.-Prof. Dr. Raimund Tenhaken for their offer to read and evaluate this thesis.

I thank Jakob Weiszmann for being the best PhD-buddy one could wish, for many hours of discussions about ideas, presentations, data, manuscripts (or mansucripts?) and for climate chamber visits. I will miss to share an office, a bench and a Jakob-mum-cake coffee break with you. I thank Martin Brenner who took so much care about my proteomics samples in the night as well as on the weekends. I thank all MoSys members who accompanied me during my thesis – in detail expressed in the MoSys-clustergram on the next page.

Thanks to our collaboration partners from Stuttgart, Univ.-Prof. Dr. Arnd G. Heyer, Dr. Simon Stutz and especially Lisa Wolf for the great subcellular project. Consequently, I want to thank the FWF (Fonds wissenschaftlicher Förderung) and the DFG (Deutsche Forschungsgemeinschaft) who provided funds for this work by the project I-2071.

Finally, I would like to thank Siegi for his overall support, his appreciation for my work and for the willingness to discuss and share chemistry and scientific related topics with me. Additionally I thank our family and friends for being there.



Thank-you-Figure The MoSys clustergram for this thesis.

Table of content

I. Abstract 1

II. Zusammenfassung 3

III. Introduction 5

IV. Results and Discussion 14

V. Results in form of publications and manuscripts 26

1. A Benchtop Fractionation Procedure for Subcellular Analysis of the Plant Metabolome 27
2. Mathematical modelling approaches in plant metabolomics 42
3. Approximating the stabilization of cellular metabolism by compartmentalization 69
4. A Strategy for Functional Interpretation of Metabolomic Time Series Data in Context of Metabolic Network Information 85
5. Combined multivariate analysis and machine learning reveals a predictive module of metabolic stress response in *Arabidopsis thaliana* 94

VI. References 127

I. Abstract

Various environmental factors affect plant distribution, growth and yield. Hence, a central aim of biological research is to quantify plant-environment interactions. In higher plants, abiotic factors like temperature and light intensity are well described to induce a reprogramming of metabolism. Further, plants are able to efficiently adapt to a changing environment, comprising a reprogramming of the transcriptome, proteome and metabolome as well as communication and signalling between subcellular organelles. Since higher plants possess one of the most compartmentalized cells across all kingdoms of life, it is particularly challenging to elucidate regulatory strategies. Mathematical models, i.e. abstract representation of plant metabolism, have been shown to be suitable to overcome this limitation and to facilitate quantitative analysis of plant-environment interactions.

The present work comprises different attempts to unravel reprogramming of metabolism in *Arabidopsis thaliana* upon abiotic stress factors. First, an experimental method for resolving and assigning metabolites to their subcellular compartment is described. Separation of cellular fractions via density gradients combined with marker enzyme assays and applied mathematical correlation strategies revealed metabolite distributions across compartments. The method is applicable to elucidate metabolome dynamics in a fast and statistically robust manner. Applied to a cold acclimation experiment different strategies of metabolic reprogramming in a cold sensitive (Cvi) and cold tolerant (Rschew) accession were observed. While the Rschew accession was characterized by a stable subcellular metabolic constitution resulting in an accumulation of primary metabolites, especially amino acid metabolism was strongly deregulated in the Cvi accession. To quantify the contribution of subcellular compartmentation to stabilization of a metabolic homeostasis, stability characteristics during environmental fluctuations were simulated by a mathematical model. Simulation of several millions of possible enzyme kinetic parameter constellations revealed diverse stabilizing contribution of different subcellular compartments. In summary, cytosolic and plastidial control of sucrose metabolism was found to stabilize metabolism more efficiently than under vacuolar control.

To make mathematical analysis applicable to nonlinear time series experiments, a strategy for the connection of dynamic metabolic functions with biochemical network structure was developed and applied to a set of experimental time course data. Mathematical analysis of diurnal sucrose dynamics and stress-induced flavonoid biosynthesis revealed time points of metabolic regulation. Additionally, a combined cold and high light experiment of mutants being perturbed in the central

carbohydrate metabolism of sucrose and starch was performed. Stress induced dynamics of primary metabolites and proteins were recorded which were applied to generate a statistical model for pattern recognition in plant stress response. This approach revealed a molecular network with a highly significant stress reaction across all analyzed genotypes. The identified network comprised 23 proteins with diverse molecular functions connecting transcriptional regulation with primary and secondary metabolism.

In conclusion, interconnected reprogramming of plant metabolism during abiotic stress affects diverse molecular levels. The combination of several experimental, methodological and mathematical strategies presented in this work provide new insights into complex plant-environment interactions.

II. Zusammenfassung

Zahlreiche Umweltbedingungen beeinflussen die Verbreitung, das Wachstum und den Ertrag von Pflanzen. Daher ist die Untersuchung von Pflanze-Umweltinteraktionen von zentraler Bedeutung für die Grundlagenforschung als auch Biotechnologie. In höheren Pflanzen führen sowohl spontane als auch langfristige Änderungen abiotischer Faktoren, wie zum Beispiel Temperatur und Lichtintensität, zu einer Anpassung des Stoffwechsels. Diese Anpassung betrifft oftmals gleichzeitig die molekularen Ebenen des Transkriptoms, Proteoms und Metaboloms sowie der Kommunikation und Signalgebung zwischen Zellorganellen. Pflanzliche Zellen besitzen einen stark kompartimentierten Stoffwechsel, weshalb regulatorische Prinzipien oftmals nur mit sehr hohem Zeit- und Arbeitsaufwand erfassbar sind. Mathematische Modelle des pflanzlichen Stoffwechsels haben sich bei der Erfassung sowie der quantitativen Analyse komplexer Fragen zu Pflanze-Umwelt Interaktionen als sehr hilfreich erwiesen.

Die vorliegende Arbeit befasst sich mit der stressinduzierten Stoffwechselanpassung in *Arabidopsis thaliana*. Zunächst wird eine experimentelle Methode zur Auflösung subzellulären Metabolitkonzentrationen vorgestellt. Die Trennung zellulärer Fraktionen aufgrund von Dichtegradienten in Kombination mit Aktivitätsbestimmung spezifischer Markerenzyme und mathematischen Korrelationsstrategien ermöglichte eine Aussage über Metabolitverteilungen in den unterschiedlichen Kompartimenten. Diese Methode ermöglichte zudem die Auflösung von Stoffwechseldynamiken in einer reproduzierbaren und statistisch robusten Art und Weise. In einer Kälteakklimatisierungsstudie konnten verschiedene regulatorische Prinzipien der Stoffwechselanpassung zwischen einer kältesensitiven (Cvi) und kältetoleranten (Rschew) natürlichen Akzession von *Arabidopsis thaliana* festgestellt werden. Während der subzelluläre Stoffwechsel der toleranten Akzession durch kälteinduzierte Akkumulation von Zuckern, organischen Säuren und Aminosäuren geprägt war, war vor allem der Aminosäurestoffwechsel der sensitiven Akzession durch eine signifikante subzelluläre Verschiebung gekennzeichnet. Für die Quantifizierung der Beiträge einzelner subzellulärer Kompartimente zur Stabilisierung einer metabolischen Homöostase während umweltbedingter Fluktuationen wurden verschiedene mathematische Modelle erstellt und simuliert. Zudem wurden Charakteristika bezüglich des Stabilitätsverhaltens nach Auslenkung untersucht. Modellsimulationen deuteten darauf hin, dass der Saccharosestoffwechsel im Blatt durch plastidäre als auch zytosolische Regulation effizienter stabilisiert wird als durch vakuoläre Regulation.

Um die Anwendbarkeit von mathematischen Analysen mit nicht-linearen Zeitserien-Experimenten zu ermöglichen, wurde eine Strategie zur Verbindung von dynamischen metabolischen Funktionen mit biochemischen Netzwerkstrukturen entwickelt. Diese Methode wurde zur Analyse experimenteller Datensätze verwendet, wobei regulatorisch bedeutende Zeitpunkte in der täglichen Saccharose-Dynamik sowie der stressinduzierten Flavonoid-Biosynthese identifiziert werden konnten. Zusätzlich wurden die Auswirkungen eines kombinierten Kälte- und Hochlichtstresses auf den Stoffwechsel des *Arabidopsis*-Wildtyps Columbia-0 analysiert und mit Stoffwechselmutanten des zentralen Kohlenhydratmetabolismus verglichen. Hierzu wurden Verfahren der statistischen Mustererkennung verwendet, welche eine Klassifizierung von Metabolit-, Protein- und physiologischen Chlorophyllfluoreszenzdaten ermöglichte. In Kombination mit multivariater Datenanalyse konnten Komponenten eines zentralen molekularen Netzwerks identifiziert werden, welches Teil der Stressreaktion aller analysierten Genotypen war. Dieses Netzwerk umfasste 23 Proteine und verknüpfte transkriptionelle Regulation mit Stoffwechselwegen des Primär- und Sekundärmetabolismus. Diese Befunde belegen die Komplexität pflanzlicher Anpassung an sich ändernde Umweltbedingungen. Gleichzeitig beschreibt diese Arbeit experimentelle und theoretische Ansätze, welche zu neuen Erkenntnissen im Gebiet pflanzlicher Stressreaktionen führen.

III. Introduction

Plant-environment interactions

As sessile organisms, plants must cope with various environmental conditions and have to adapt fast for surviving unfavourable conditions. Abiotic stress is defined as environmental conditions affecting growth and yield. The response of plants to abiotic stress is highly dynamic and complex (Cramer *et al.*, 2011). The ability to adapt to environmental conditions limits the dispersion of plants in different geographical regions, and low temperature is a major factor to account for significant losses in agricultural plant productivity. Naturally occurring inbred lines of the model plant *Arabidopsis thaliana* constitute locally adapted populations (Provart *et al.*, 2016), called accessions. Therefore, *Arabidopsis* represents a suitable system to analyse adaptation to changing environmental conditions. Freezing tolerance in *Arabidopsis* is important as some of the rosettes overwinter before flowering (Nordborg & Bergelson, 1999) and freezing temperatures can occur at the beginning of their vegetative growth phase. Supplementary, the analysis of over 70 naturally occurring *Arabidopsis* accessions has shown a positive correlation between freezing tolerance and latitude of origin (Zhen & Ungerer, 2008), and additionally, a correlation between the latitude with the longitude and freezing tolerance was observed (Zuther *et al.*, 2012). Environmental factors seem to differ across the species range, but besides temperature especially the combination with precipitation might play a crucial role in freezing tolerance (Horton *et al.*, 2016). Cold slows down various processes like enzyme activities, metabolite mobility, energy metabolism and membrane fluidity. Additionally, the structure of biomolecules, membrane channel conductivity and plant anatomy can be affected (Thomashow, 1999; van Buer *et al.*, 2016).

Physically, in most plant tissues, freezing injury occurs upon ice formation which leads to cellular dehydration affecting cellular membrane systems (Thomashow, 2001). Cold acclimation is defined as the ability to enhance freezing tolerance upon exposure to chilling/non-freezing temperature (Thomashow, 1999). Acclimation is induced by environmental changes, is initiated with a stress response and can be characterized by transient, physiological, biochemical and molecular perturbations. Finally, a stable long-term adjustment of metabolism is gained (Huner *et al.*, 1998). The process of cold acclimation requires light (Wanner & Junttila, 1999) and is initiated already after a relatively short-time period of exposure to low but non-freezing temperature (Ristic & Ashworth, 1993) in the presence of (required) light. The high mortality at -8°C of non-acclimated plants compared to acclimated plants was already described for over 70 *Arabidopsis* accessions grown in growth chambers (Zhen & Ungerer, 2008). In cold stress field studies it was shown that

accessions with lowest cold acclimation potential benefited most from a cold pre-treatment regarding seed yield (Cvetkovic *et al.*, 2017). After cold acclimation, especially younger leaves were observed to develop freezing tolerance more rapidly than mature leaves. Interestingly, as fast as freezing tolerance increases, it can be lost again within 1 or 2 days (Wanner & Junttila, 1999) though other studies showed that the timeframe of deacclimation might differ between natural *Arabidopsis* accessions (Zuther *et al.*, 2015). Moreover, this suggests that *Arabidopsis* can memorize an earlier (priming) cold stress for several days over a stress-free period. It was suggested, that especially the plastid antioxidant system transmits information on a previous cold stress over time (van Buer *et al.*, 2016).

Regulation of metabolism during exposure to cold

Primary effects of cold sensing are changes in calcium fluxes, mobilization of transcription factors and photosynthetic signals. Those processes activate signal transduction processes, modify gene expression and might improve plant performance under low temperature ((van Buer *et al.*, 2016) and references therein). Plasma membrane fluidity is handled as one primary signal for cold perception in plants (Örvar *et al.*, 2000) as well as plasma membrane localized kinases which transduce the signal from the plasma membrane into the nucleus (Liu *et al.*, 2017). On the level of transcription, one of the best characterized pathways being involved in cold acclimation and freezing tolerance is the C-repeat binding factor (CBF) pathway of *Arabidopsis* (Thomashow, 1999; Knight & Knight, 2012). The CBF locus (also known as DREB; drought-responsive element binding protein) includes three genes CBF1, CBF2 and CBF3, transcription factors which are induced within minutes of exposure to low non-freezing temperatures (Gilmour *et al.*, 1998). After induction, CBF-targeted cold-regulated genes (COR genes) are altered in their expression. Those more than hundred COR genes are also known as the CBF regulon (Thomashow, 2010). Overexpression of the CBF loci results in increased freezing tolerance without cold acclimation (Jaglo-Ottosen *et al.*, 1998; Gilmour *et al.*, 2004). There is additional evidence for a particular role of CBF2 which might contribute to local adaptation in natural populations of *Arabidopsis* and, therefore, the CBF pathway might be an evolutionary factor of those populations (Gehan *et al.*, 2015). However, other studies do find just weak or no evidence for a correlation of natural variation of freezing tolerance and expression of CBF genes (McKhann *et al.*, 2008; Zuther *et al.*, 2012). The CBF pathway itself is regulated by phytohormones, other transcription factors, Ca^{2+} binding receptor kinases and key components of the circadian clock and light as well as light quality via the phytochrome system (Kurepin *et al.*, 2013; Shi *et al.*, 2015). Nevertheless, cold

acclimation seems to modulate gene expression up to one third of the total *Arabidopsis* genome (Hannah *et al.*, 2005) and transcriptome analysis has shown that only ~12% of the cold-responsive genes are controlled by CBFs indicating an interaction of diverse low-temperature regulons in cold signalling (Fowler & Thomashow, 2002). Subsequently, the contribution of genes to freezing tolerance variation within *Arabidopsis* remains largely unknown (Horton *et al.*, 2016).

The adaptation of plants to a fluctuating environment is not only visible in different expression pattern of genes, but also in their protein products. As proteins are one of the major players in most cellular events, they are directly involved in plant low temperature response. Proteome analysis is recommended as an appropriate strategy for complementing transcriptome level changes (Janmohammadi *et al.*, 2015). Additionally, post-transcriptional mechanisms (e.g. alternative splicing, RNA silencing, mRNA export) and post-translational modifications (e.g. phosphorylation, ubiquitination, sumoylation) and protein isoforms do play a crucial role during cold response in plants (Kosová *et al.*, 2011; Miura & Furumoto, 2013; Kosová *et al.*, 2018). Subsequently, it has been proven that modification of gene expression at the transcript level frequently does not correlate with the protein level (Maier *et al.*, 2009). Low temperature stress affects proteins involved in carbohydrate metabolism, photosynthesis, polyamine synthesis, ROS scavenging, protein folding, stabilizing cell structure and cell membrane integrity (Janmohammadi *et al.*, 2015) and those proteins are often coded by COR genes (Hannah *et al.*, 2005). Well characterized low temperature responsive proteins are late embryogenesis abundant (LEA) proteins, heat shock proteins (HSP), pathogen related (PR) proteins (Janská *et al.*, 2010; Kosová *et al.*, 2011; Miura & Furumoto, 2013; Janmohammadi *et al.*, 2015; Kosová *et al.*, 2018). Shortly, LEA proteins (e.g. COR15a, KIN1) might participate in the stabilization of membranes against freeze induced injury and protect proteins from denaturation when cytoplasm becomes dehydrated. One group of LEA proteins are dehydrins which possess a high number of charged amino acids, are heat-stable and might be crucial for plants freezing tolerance (Thomashow, 1999; Puhakainen *et al.*, 2004; Hundertmark & Hincha, 2008; Janská *et al.*, 2010). Low temperature increases the potential risk of protein misfolding, resulting in non-functional proteins. Therefore, it is not surprising that low temperature enhances the accumulation of proteins with chaperone functions (Kosová *et al.*, 2011). Several heat shock proteins take part in the refolding of denatured proteins and in prevention of their aggregation. Additionally, they have cryoprotective effects, and participate in membrane protection, and for example HSP90 regulates CBF genes and members of the circadian clock (Thomashow, 1999; Yan *et al.*, 2006; Timperio *et al.*, 2008; Noren *et al.*, 2016). Some PR proteins are known to have antifreeze activities, they inhibit the recrystallization of

intercellular ice in the apoplastic space and even prevent intracellular ice formation (Janská *et al.*, 2010).

Low temperature and/or high light leads to a potential energy imbalance between photochemistry, electron transport and metabolism. This leads to increased PSII excitation pressure (Huner *et al.*, 1998). Plants transferred to low temperature showed an immediate loss of the effective photosynthetic capacity i.e. the maximum rate of photosynthesis is decreased (Strand *et al.*, 1999). Correlated with this, a transcript suppression of genes coding for photosynthetic proteins was described (Strand *et al.*, 1997). Hence, during cold acclimation photosynthetic capacity recovers and might reach similar levels again like under ambient temperature. Several Calvin cycle enzyme activities increase during acclimation to cold, with the most pronounced increase in ribulose-1,5-bisphosphate-carboxylase/oxygenase (Rubisco) activity (Strand *et al.*, 1999). Carbohydrates are primary photosynthetic products playing a crucial role in energy metabolism, developmental processes, stress signalling and many other processes in the subcellular and whole plant level (Graf *et al.*, 2010; Pommerrenig *et al.*, 2018). Confirmed by many studies, the central carbohydrate metabolism is a particularly prominent component of the reprogramming of the metabolome at low temperatures (summary in (Guy *et al.*, 2008)). The response and acclimation to cold stress in various *Arabidopsis* accessions and primary metabolism has been focused by many studies (Kaplan *et al.*, 2007; Guy *et al.*, 2008; Nägele *et al.*, 2011; Nägele *et al.*, 2012). Sugars and starch play a crucial role in entrainment of the circadian clock (Haydon *et al.*, 2013) and clock components are significantly affected by abiotic stress conditions (Miura & Furumoto, 2013). Thus, with respect to freezing tolerance, especially glucose, fructose and sucrose belong to the most strongly positive correlated metabolites from the central carbohydrate metabolism (Hannah *et al.*, 2006) as well as raffinose and various amino acids, organic acids and sugar alcohols (Kaplan *et al.*, 2004). Diverse saccharides are capable to directly stabilize biological membranes under stress conditions. Cell membranes are protected by, e.g. sucrose and raffinose family oligosaccharides (RFOs) members (Tarkowski & Van den Ende, 2015). The amount of transcripts and activity of enzymes from the cytosolic pathway of sugar biosynthesis show an upregulation at low temperatures. Especially the two enzymes involved in sucrose synthesis, sucrose phosphate synthase (SPS) and cytosolic fructose-1,6-bisphosphatase (cFBPase) activities increase during cold exposure (Strand *et al.*, 1997; Strand *et al.*, 1999) and SPS is known to be rate limiting (Stitt *et al.*, 1988). Furthermore, SPS overexpressing plants were observed to have an improved photosynthetic performance and increased freezing tolerance after cold acclimation (Strand *et al.*, 2003). Sugars are well known to play pivotal roles as signalling cascades and a prominent and

conserved involved enzyme is Hexokinase1 (Moore *et al.*, 2003; Rolland *et al.*, 2006). In addition to accumulation of soluble carbohydrates, also the content of starch is known to significantly increase during cold acclimation (Guy *et al.*, 2008; Nagler *et al.*, 2015; Thalmann & Santelia, 2017). Previously, starch degradation during initial cold response was suggested to augment the accumulation of carbohydrates (Sicher, 2011). Starch is a major storage compound in plants, and *Arabidopsis* partitions approximately 30-50% of its photoassimilates into transitory starch. During the night, starch is degraded to release maltose and glucose. A plastidial enzyme involved in the formation of starch in chloroplasts is phosphoglucomutase (PGM). The Calvin-Benson cycle is linked to starch biosynthesis by the enzymes phosphoglucoisomerase (PGI), PGM and ADP-glucose pyrophosphorylase (AGPase) resulting in the formation of ADP-glucose. ADP-glucose is the substrate for biosynthesis of starch (Stitt & Zeeman, 2012). PGM catalyses the reversible interconversion of glucose-6-phosphate (G6P) to glucose-1-phosphate (G1P) and is present in the cytosol as well as in the plastid (Malinova *et al.*, 2014).

Besides carbohydrates, also other compounds of primary metabolism, e.g. amino acids, are a central part of plant stress response. Proline is well known to accumulate during various biotic and abiotic stress conditions. Protective functions of proline during stress exposure are diverse, and it has been suggested that proline protects membranes, enzymes, polyribosomes, serves as osmoticum, is a substrate for the TCA cycle upon relief from stress situations, regulates cellular pH and affects cellular redox potential (Hare & Cress, 1997; Szabados & Savoure, 2010). However, a cold acclimation study of over 50 natural accession showed that there are some exceptions regarding the content of proline and freezing tolerance (Zuther *et al.*, 2012). These authors speculated that a functional connection of high sugar content rather than a high proline content is a prerequisite for an enhanced freezing tolerance in *Arabidopsis*.

Gene expression related to secondary metabolism is well correlated with freezing tolerance (Hannah *et al.*, 2006) and in *Arabidopsis* cold stress induces the biosynthesis of flavonoids and anthocyanins (Kaplan *et al.*, 2007). Hence, several flavonoid biosynthesis mutants with reduced flavonoid content showed impaired freezing tolerance, though the contribution of flavonoids to freezing tolerance was shown to be genotype dependent (Schulz *et al.*, 2016).

Abiotic stress factors are often occurring at the same time, like drought and heat or cold and high-light. Chilling temperatures combined with high light conditions occur for example in the morning hours in spring. Combined effects result in various molecular changes, and frequently photosystem II is strongly affected (Van Hasselt, 1990). In many cases, stress responses are antagonistic, and the predictability of a combined stress output from experiments with single stress application is

questionable (Mittler, 2006; Prasch & Sonnewald, 2015). Beyond, numerous other molecular compounds contribute to stress acclimation and increasing evidence is provided for a crucial role of subcellular metabolite concentrations in shaping acclimation output (Hurry, 2017; Pommerrenig *et al.*, 2018).

Subcellular metabolite dynamics

All eukaryotic cells are characterized by compartmentalization. Compartmentalization within cells creates different biochemical reaction conditions. Particularly plant cells show a highly compartmentalized cell structure, and the analysis of metabolic pathways is challenging due to a high diversity of subcellular biochemical reactions (Lunn, 2007). Stress-induced changes of metabolite concentrations determined from crude extracts provide only limited information about the roles of specific metabolites and proteins (Tiessen & Padilla-Chacon, 2012; Hurry, 2017). In contrast, information about changes of subcellular metabolite concentrations may provide pivotal information about metabolic stress response (Lunn, 2007). This knowledge can contribute to development of crop plants which are tolerant and/or even are able to maintain productivity under chronic stress (Hurry, 2017). The interconnection by various transport systems of the subcellular compartments chloroplast, cytosol and vacuole enables a regulated exchange of metabolites across biological membrane systems (Linka & Weber, 2010). The activity and/or expression and regulation of soluble sugar transporters like plastidial triose-phosphate translocator (TPT) (Flügge, 1999; Lundmark *et al.*, 2006), tonoplast monosaccharide transporters (TMTs) (Wormit *et al.*, 2006) and “Sugars Will Eventually be Exported Transporters” (SWEETs) (Klemens *et al.*, 2013) play a crucial role in metabolic stress response (Tarkowski & Van den Ende, 2015). Several previous studies have focused on the subcellular analysis of metabolite dynamics and underlying fluxes (see e.g. (Masakapalli *et al.*, 2010; Klie *et al.*, 2011; Krueger *et al.*, 2011; Nägele & Heyer, 2013; Szecowka *et al.*, 2013; Arrivault *et al.*, 2014; Hoermiller *et al.*, 2017; Hossain *et al.*, 2017)). A recent study has analyzed subcellular dynamics of primary metabolism during cold acclimation in two mutants impaired in plastidial phosphoglucomutase (*pgm1*) and sucrose phosphate synthase (*spsa1*) (Hoermiller *et al.*, 2017). The authors highlighted the central role of plastid metabolism and the preparation for the continuation of growth under the new temperature regime (Hoermiller *et al.*, 2017). Additionally there is evidence, that especially invertases, which control the compartment specific sucrose/hexoses ratios, are key controllers in response to abiotic and biotic stresses (Tarkowski & Van den Ende, 2015), besides they are located in several compartments e.g.

cytosol, vacuole, cell wall, mitochondria (Sturm, 1999; Xiang *et al.*, 2011; Tiessen & Padilla-Chacon, 2012). Subcellular analysis combined with kinetic modelling showed that there is a different contribution between vacuolar and cytosolic invertase in a sensitive and tolerant natural accession upon cold stress. For the *inv4* mutants (impaired vacuolar invertase activity) it was shown that photosynthesis was lowered and whole (subcellular) energy metabolism was perturbed in stress conditions (Weiszmann *et al.*, 2017). In addition to these examples, it is demanded that metabolome analysis has to consider subcellular localization, including modelling and simulation, for a full understanding of metabolism (Dietz, 2017).

Mathematical modelling of plant metabolism

Typically, experimental analysis of plant metabolism before, during and after stress exposure results in a multidimensional data set. Hence, dimensionality reduction and efficient variable extraction is pivotal for suitable hypothesis generation and testing. In addition to multivariate statistics, strategies of mathematical modelling have been shown to be essential for gaining new insights into complex metabolic networks (Rohwer, 2012). Representation and computational simulation of plant metabolic models potentially reveals detailed mechanistic insights (Heinig *et al.*, 2013). Mathematical modelling is also frequently applied in biotechnological applications, e.g. for metabolic engineering (Mintz-Oron *et al.*, 2012; Heinig *et al.*, 2013). Frequently, modelling results reveal new hypotheses and/or provide hints which biochemical reaction or regulatory interaction might be a suitable candidate for further detailed experimental analysis.

Dynamics of metabolism can be described mathematically using ordinary differential equations, ODEs. Metabolic reactions are based on various non-linear elements, e.g. substrate saturation of Michaelis-Menten kinetics or feedback-inhibition terms. Particularly with regard to experimental validation, solving and testing such nonlinear systems is challenging. Instead, linearization at a metabolic (quasi) steady state, at which (almost) no significant changes of metabolite concentrations can be observed, frequently helps to unravel complex regulatory principles which hardly accessible by intuition (Nägele, 2014). The steady state assumption is defined by no change in metabolite concentrations [**M**] over a certain time period (Eq. 1):

$$\frac{d[\boldsymbol{M}]}{dt} = \mathbf{0} \qquad \textbf{Eq. 1}$$

Steady state conditions imply that reaction rates of the system are constant, and consuming and synthetizing reactions are in equilibrium. Experimental analysis of metabolome dynamics, e.g. during a day/night cycle, hardly agrees with the assumption of a steady state. Yet, instead of a

steady state, a *quasi* steady state can frequently be assumed and **no change** might be replaced by **no significant change** of metabolite concentration, $\frac{d[\boldsymbol{M}]}{dt} \approx \mathbf{0}$**.** In general, a system of ODEs can be described by the product of the stoichiometric matrix $\boldsymbol{S}$ (entries of $\boldsymbol{S}$ represent the participation of a metabolite in a reaction) and the vector $\boldsymbol{v}$ (contains reaction rates) (Eq. 2):

$$\frac{d[\boldsymbol{M}]}{dt} = \boldsymbol{S} * \boldsymbol{v} = \boldsymbol{f}(\boldsymbol{M}, \boldsymbol{p}, t) = f_{\boldsymbol{M}}(t) = \mathbf{0} \qquad \textbf{Eq. 2}$$

Metabolic functions constitute the right side of the equation, $\boldsymbol{f}(\boldsymbol{M}, \boldsymbol{p}, t)$, and depend on metabolite concentrations $\boldsymbol{M}$, parameters $\boldsymbol{p}$ (e.g. kinetic parameters) and time t. Linearization of the metabolic system at time point t_0 of the quasi steady state and building partial derivatives of metabolic functions with respect to metabolite concentrations potentially characterize regulatory interactions, i.e. trajectories, within a metabolic network. Representation of all partial derivatives with respect to metabolite concentrations is summarized by the Jacobian matrix $\boldsymbol{J}$ (Eq. 3)

$$\boldsymbol{J} = \begin{pmatrix} \frac{\delta f_1}{\delta M_1} & \frac{\delta f_1}{\delta M_2} & \cdots & \frac{\delta f_1}{\delta M_m} \\ \frac{\delta f_2}{\delta M_1} & \frac{\delta f_2}{\delta M_2} & \cdots & \frac{\delta f_2}{\delta M_m} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\delta f_m}{\delta M_1} & \frac{\delta f_m}{\delta M_2} & \cdots & \frac{\delta f_m}{\delta M_m} \end{pmatrix} \qquad \textbf{Eq. 3}$$

Jacobian entries report on the response of metabolic functions to (slight) perturbations of the metabolic (steady) state, and high absolute values in Jacobian entries indicate a strong effect of concentration changes on metabolic functions. However, even within strongly simplified biochemical reaction networks, many kinetic parameters and enzymatic parameters have to be determined for numerical approximation of $\boldsymbol{J}$, which limits the application of mathematical modelling to metabolomics data (Schaber *et al.*, 2009). To circumvent this limitation, approaches like structural kinetic modelling approach (Steuer *et al.*, 2006) or the inverse approximation of $\boldsymbol{J}$ from experimental metabolomics data (Nägele *et al.*, 2014) have been developed. In addition, metabolic time series analysis combined with modelling strategies potentially supports the description of nonlinear system behavior and might predict concentration dynamics over large time intervals. For example, Dutta and coworkers developed an algorithm for identification of differentially expressed genes in time series experiments (Dutta *et al.*, 2007). Dutta an co-workers have applied their algorithm to transcriptome and metabolome data which were determined within a time-series experiment growing liquid cultures of *Arabidopsis* under elevated CO_2 (Dutta *et al.*, 2009). In addition to time-series analysis, approaches of pattern recognition and machine learning are increasingly applied in (plant) biology (Ma *et al.*, 2014b). Machine learning (ML) techniques

extract characteristic pattern information from examples (training dataset) to enable a prediction or search for associations and/or patterns in, hitherto, uncharacterized data (Mjolsness & DeCoste, 2001; Ma *et al.*, 2014b). ML-driven analysis of *Arabidopsis* salt stress transcriptome data pointed towards two previously unreported genes, where mutants showed indeed as predicted salt-sensitive phenotypes (Ma *et al.*, 2014a). Recently, Sperschneider and colleagues used ML techniques to predict apoplastic localization, whereas this might help to facilitate functional studies if plant pathogens localize to apoplast or (already) entered plant cells (Sperschneider *et al.*, 2018). Thus, in summary, a combination of mathematical modelling, time series analysis and machine learning techniques promises to support hypothesis generation and experimental validation within complex metabolic systems.

Objective of the present work

This thesis aimed to quantitatively analyse metabolic reprogramming during abiotic stress exposure in the model organism *Arabidopsis thaliana*. To elucidate regulatory instances of plant stress response and acclimation, subcellular metabolite analysis, protein analysis and chlorophyll fluorescence measurements were combined with mathematical modelling strategies, regression analysis and pattern recognition. Finally, all applied methods and techniques could elucidate a central molecular network of (sub)cellular abiotic stress response in plants.

IV. Results and Discussion

Publication

A Benchtop Fractionation Procedure for Subcellular Analysis of the Plant Metabolome

Lisa Fürtauer, *Wolfram Weckwerth and Thomas Nägele*

Published in Frontiers in Plant Science (2016) DOI: 10.3389/fpls.2016.01912

While crude whole cell extracts of metabolites and proteins are suitable to record the overall stress response of metabolism, its information content about organelle-specific subcellular processes, e.g. biosynthetic pathways, is strongly limited. Further, the role of specific metabolites and proteins might be overseen or, even worse, might be misleading (Hurry, 2017). To overcome this limitation, the method of non-aqueous fractionation (NAF) was developed to separate vacuolar, plastidial and cytosolic marker enzymes and metabolites in spinach leaves (Gerhardt & Heldt, 1984). This method was successfully applied for decades (e.g. (Gerhardt & Heldt, 1984; Stitt *et al.*, 1989; Farré *et al.*, 2001; Klie *et al.*, 2011; Nägele & Heyer, 2013; Arrivault *et al.*, 2014; Hoermiller *et al.*, 2017; Hossain *et al.*, 2017; Medeiros *et al.*, 2017)). Combining NAF with experimental high-throughput analyses methods, e.g. mass spectrometry coupled to chromatography, is deemed to be important for biological research (Geigenberger *et al.*, 2011; Kueger *et al.*, 2012; Tiessen & Padilla-Chacon, 2012) and metabolic engineering (Heinig *et al.*, 2013). Hence, technical challenges have frequently limited data reproducibility and sample throughput by significantly affecting statistical robustness of subcellular data. In the presented study, a new fractionation method was developed being adapted to a benchtop standard equipment enhancing the applicability of the NAF technique. This method needs only a fraction of plant material compared to the original methods (~10 % from (Gerhardt & Heldt, 1984)). This makes it suitable for studies which are strongly limited by the amount of material, e.g. seedling or field experiments. Additionally, the developed method replaces gradient pumps by pipetting of solvents with a pre-defined density facilitating the reliable generation of linear, and also non-linear, density gradients. Ultracentrifugation steps of the original protocol were replaced by benchtop centrifugation due to reduction of the liquid column height which every particle has to pass through. The protocol developments enabled the simultaneous handling of more than 20 samples at once. Furthermore, sonication steps of each single density fraction helped to raise the efficiency of compartment separation. By downscaling of compartment-specific marker enzyme assays to microtiter plate volume, sample throughput was additionally increased and made it applicable for automatized high throughput pipetting platforms.

Based on marker enzyme fractionation across the density gradient, subcellular metabolite distribution can be calculated applying correlation strategies. This step is crucial to reveal the *in vivo* situation in subcellular compartments, yet finding criteria for selection of suitable correlation strategies is critical (Dietz, 2017). Frequently, minimization of summed squared errors between a theoretical (optimum) and experimentally observed marker enzyme distribution is applied to resolve subcellular metabolite distribution between two or three compartments ((Stitt *et al.*, 1989; Riens *et al.*, 1991; Nägele & Heyer, 2013; Hoermiller *et al.*, 2017). However, it was also shown that repetition of marker enzyme measurements might lead to distinct results due to technical errors (Geigenberger *et al.*, 2011). To account for such technical standard deviation, an algorithm was developed comparing dynamics of metabolites and marker enzyme activities across the density gradient within predefined technical errors, e.g. within a 5% or 10% margin. While this might lead to higher standard deviations of calculated mean metabolite distributions, it reduces the probability of over-interpreting technical artefacts.

The method was developed, tested and validated in a cold acclimation experiment comprising the *Arabidopsis* accessions Col-0 (Columbia), Cvi (Cap Verde Island) and Rsch (Rschew) with differential cold acclimation capacity (Cvi<Col<Rsch, see e.g. (Hannah *et al.*, 2006)). At bolting stage, non-cold acclimated plants were harvested at ambient temperature (22°C) at midday while cold acclimated plants were harvested after 7 days at 5°C. In non-cold acclimated plants, all analyzed accessions showed similar relative distributions of sugars and TCA cycle intermediates across plastids, cytosol and the vacuole. In contrast, cold exposure induced a decrease of the relative proportion of amino acids in the plastids of Cvi and an increase in the vacuole. In Col-0, a decrease of the relative proportion of amino acids was observed in the cytosol accompanied by an increase in the vacuole. The most freezing tolerant accession Rsch showed nearly the same relative distribution of primary metabolites before and after cold acclimation. These data suggested the vacuole to serve as a buffer compartment for subcellular reprogramming of primary metabolism during environmental fluctuations which might contribute to the stabilization of plastidial amino acid metabolism. Furthermore, observations indicated a negative correlation between subcellular re-arrangement of primary metabolites and cold acclimation capacity. Absolute metabolite concentrations significantly increased due to cold exposure irrespective of the genotype. Compared to previous studies, similar subcellular patterns of, potentially cold protective, substances were observed, e.g. a significant increase of plastidial raffinose concentration during cold acclimation (Schneider & Keller, 2009; Knaupp *et al.*, 2011). Raffinose has previously been reported to be involved in photosystem II protection against damage during

freeze-thaw cycles (Knaupp *et al.*, 2011). As expected, a more pronounced increase of plastidial raffinose was observed in cold tolerant Rsch (~88-fold) than in in Cvi (~1.5-fold). In the cold, proline, which is prominently involved in plant cold response, accumulated in different compartments to a different extent like presented earlier (Hoermiller *et al.*, 2017) providing evidence for the applicability of the whole fractionation procedure. In detail, after cold acclimation a stronger accumulation of proline was observed in plastids and cytosol of Rsch than in Cvi. Thus, due to a potential role of proline in protecting membrane systems against freezing damage, it is hypothesised that membranes of subcellular compartments in Rsch are better protected against membrane fusion than in Cvi.

Irrespective of the genotype, the percentage of significantly changing metabolite concentrations due to cold exposure was ~70 % per compartment and this indicates strong environmentally induced reprogramming of subcellular pathways. This metabolic adaptation results from a multigenic and highly coordinated reprogramming comprising changes in gene expression, translational processes and enzymatic reactions. Conclusively, the presented findings highlight the importance of subcellular analysis of metabolic reprogramming to reveal acclimation strategies of plants to a changing environment.

Publications

Approximating the stabilization of cellular metabolism by compartmentalization

Lisa Fürtauer *and Thomas Nägele*

Published in Theory in Biosciences (2016), Vol. 135, Issue 1-2, pp 73-87. DOI: 10.1007/s12064-016-0225-y

and

Mathematical modelling approaches in plant metabolomics.

Lisa Fürtauer*, Jakob Weiszmann*, Wolfram Weckwerth and Thomas Nägele.* In: Plant Metabolomics.

In press (2018)*: Plant Metabolomics: Methods and Protocols, Methods in Molecular Biology, vol. 1778, © Springer Science+Business Media, DOI: https://doi.org/10.1007/978-1-4939-7819-9_24*

Fluctuation of environmental factors, e.g. light intensity or temperature, induce reprogramming of plant metabolism to establish a new and adjusted metabolic homeostasis. This homeostasis is the results of an interplay of numerous enzymatic reactions and transport processes on a subcellular level. Due to the diversity of involved regulatory processes, the quantitative characterization of environmentally induced metabolic reprogramming relies on theoretical, i.e. mathematical,

approaches which enable the abstract representation of biochemical reaction networks. Such abstract representations enable the computationally assisted analysis of complex networks and, thus, promise to unravel strategies of metabolic reprogramming which can hardly be traced back by intuition. The two presented publications of this paragraph summarize the procedure of model generation, mathematical description by ordinary differential equations (ODEs) and the computational analysis stability characteristics after simulated environmental perturbation. Due to its central role in whole plant biochemistry, central leaf carbohydrate metabolism (Fig. 1) was translated into a mathematical model. The model comprised three intracellular compartments plastid, cytosol, vacuole and an extracellular environment.

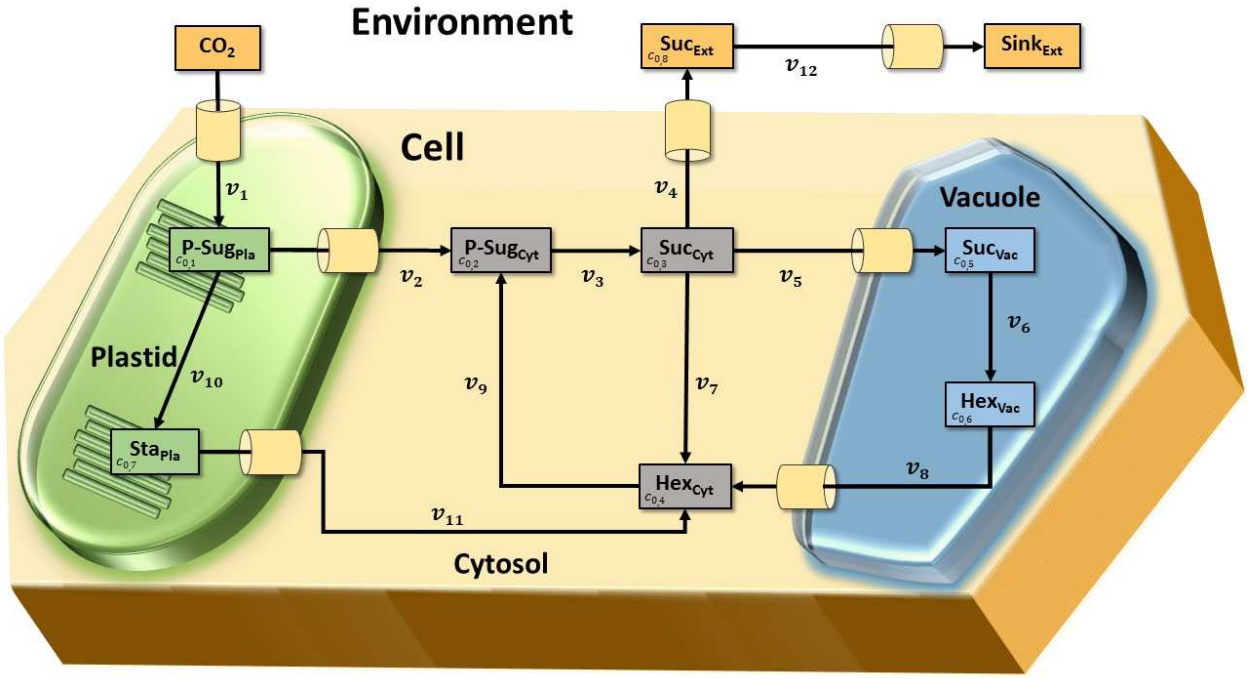


Figure 1 *Subcellular biological model including reactions (v) of enzymatic interconversion and transport reactions for a leaf mesophyll cell.* $c_{0,i}$ *– steady state concentrations with respective number* i *of metabolite, Pla – Plastid, Cyt – Cytosol, Vac – Vacuole, Ext – Extracellular, P-Sug – Phosphorylated Sugars, Sta – Starch, Suc – Sucrose, Hex – Hexoses. Cylinder represent transport processes.*

Photosynthetic carbon fixation results in plastidial sugar phosphates ($P\text{-}Sug_{Pla}$, reaction v_1). Environmental fluctuations were simulated to directly affect the flux into sugar phosphates. Plastidial sugar phosphates were substrate for two reactions, the export into the cytosol ($P\text{-}Sug_{Cyt}$, v_2) and intraplastidial conversion to starch (Sta_{Pla}, v_{10}). Cytosolic sucrose (Suc_{Cyt}) was either exported to extracellular space (v_4, Suc_{Ext} leading via v_{12} to $Sink_{Ext}$), imported into the vacuole (v_5) or remained in the cytosol. Sucrose cleavage was catalysed by invertase enzymes either in the cytosol or the vacuole (v_6, v_7, Hex_{Vac}, Hex_{Cyt}). The cytosolic hexose pool was additionally supplemented by the export of vacuolar hexoses (v_8) and the degradation of starch

(v_{11}). Phosphorylation (v_9), catalysed by hexokinase, interconverted Hex_{Cyt} into $P\text{-}Sug_{Cyt}$ and finalized the metabolic cycle of sucrose metabolism. This biological model was translated into a system of ODEs, where within each reaction (v) several enzymatic steps and/or membrane transports were included. Experimentally determined metabolite concentrations were derived from previous non-aqueous fractionation studies (Nägele *et al.*, 2012; Nägele & Heyer, 2013). To evaluate the impact of extracellular sucrose (Suc_{Ext}) on intracellular sugar metabolism, the effect of different extracellular sucrose concentrations (low, medium and high compared to cytosolic sucrose concentrations) was computationally evaluated by model simulations.

The approach of structural kinetic modelling (SKM) (Steuer *et al.*, 2006) was applied to analyze the re-stabilization of the metabolic system after a simulated perturbation. In the SKM procedure, kinetic parameters were normalized to a steady-state condition enabling a fast simulation of millions of normalized enzyme kinetic parameter constellations. This method has already been applied before in several studies (e.g. (Steuer *et al.*, 2006; Reznik & Segre, 2010; Henkel *et al.*, 2011; Carbonaro & Thorpe, 2017)). The assumed steady state is defined as no change in metabolite concentrations [$\boldsymbol{M}$] over a change of time thus consuming and synthesizing reactions of each metabolite are in equilibrium (compare Eq. 1). In the model 12 reactions ($v_j;\ j = 1, \dots, r\ here\ r = 12$), 8 steady state metabolite concentrations ($c_{0,i}\ i = 1, \dots, m\ here\ m = 8$) and their fluxes $v_j(c_{0,i})$ were defined. To simulate a steady state equilibrium, the input flux (v_1) and output flux (v_4) were set to an equal value. The flux F ($F = 1$) into the system was perturbed by multiplication with a term α, which simulated environmental fluctuations. Two proportion characters were used to quantify the relative proportion of carbon flux from chloroplast to cytosol (β, proportionated reaction v_2 and v_{10}) as well as the carbon flux from cytosol to vacuole (γ, proportionated reaction v_5 and v_7). Based on the SKM approach, the Jacobian matrix was defined as the product of matrices $\boldsymbol{\Lambda}$ and $\boldsymbol{\theta}$ (Eq. 4)

$$\boldsymbol{J} \coloneqq \boldsymbol{\Lambda\theta} \qquad \textbf{Eq. 4}$$

Here, $\boldsymbol{\Lambda}$ is the stoichiometric matrix ($\boldsymbol{N}$) normalized to steady-state fluxes $v(c_0)$ and steady state metabolite concentrations (c_0) (Eq. 5)

$$\boldsymbol{\Lambda}_{i,j} \coloneqq \boldsymbol{N}_{i,j}\frac{v_j(c_{0,i})}{c_{0,i}} \qquad \textbf{Eq. 5}$$

Thus, in a metabolic context, rows of $\boldsymbol{\Lambda}$ indicated the participation of metabolites $c_{0,i}$ in reactions v_j which are described column-wise:

$$\Lambda_{i,j} = \begin{pmatrix} \frac{\alpha F}{[PSuc_{Pla}]_0} & -\frac{\beta\alpha F}{[PSuc_{Pla}]_0} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -\frac{(1-\beta)\alpha F}{[PSuc_{Pla}]_0} & 0 & 0 \\ 0 & \frac{\beta\alpha F}{[PSuc_{Cyt}]_0} & -\frac{F}{[PSuc_{Cyt}]_0} & 0 & 0 & 0 & 0 & 0 & \frac{F-\beta\alpha F}{[PSuc_{Cyt}]_0} & 0 & 0 & 0 \\ 0 & 0 & \frac{F}{[Suc_{Cyt}]_0} & -\frac{\alpha F}{[Suc_{Cyt}]_0} & -\frac{\gamma(1-\alpha)F}{[Suc_{Cyt}]_0} & 0 & -\frac{(1-\gamma)(1-\alpha)F}{[Suc_{Cyt}]_0} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \frac{(1-\gamma)(1-\alpha)F}{[Hex_{Cyt}]_0} & \frac{\gamma(1-\alpha)F}{[Hex_{Cyt}]_0} & -\frac{F-\beta\alpha F}{[Hex_{Cyt}]_0} & 0 & \frac{(1-\beta)\alpha F}{[Hex_{Cyt}]_0} & 0 \\ 0 & 0 & 0 & 0 & \frac{\gamma(1-\alpha)F}{[Suc_{Vac}]_0} & -\frac{\gamma(1-\alpha)F}{[Suc_{Vac}]_0} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \frac{\gamma(1-\alpha)F}{[Hex_{Vac}]_0} & 0 & -\frac{\gamma(1-\alpha)F}{[Hex_{Vac}]_0} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{(1-\beta)\alpha F}{[Sta_{Pla}]_0} & -\frac{(1-\beta)\alpha F}{[Sta_{Pla}]_0} & 0 \\ 0 & 0 & 0 & \frac{\alpha F}{[Suc_{Ext}]_0} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -\frac{\alpha F}{[Suc_{Ext}]_0} \end{pmatrix}$$

Further, normalized elasticities were defined as the entries of matrix $\boldsymbol{\theta}$, where every $\theta_{j,i}$ consisted of μ (normalized reaction rates) normalized to x (normalized metabolite concentrations) (Eq. 6)

$$\boldsymbol{\theta} := \frac{d\mu}{dx} = \frac{d\frac{v(c)}{v(c_0)}}{d\frac{c(t)}{c_0}} \qquad \textbf{Eq. 6}$$

Finally, Eq. 6 represents the degree of saturation of the normalized flux μ with regard to the normalized substrate concentration x. Non-zero elements $\theta_{j,i}$ in matrix $\boldsymbol{\theta}$ indicate the participation of a metabolite in a reaction, e.g. as a substrate, as an activating or inhibiting compound. Conclusively, variation of entries in the matrix $\boldsymbol{\theta}$ resulted in a simulated deflection of the Jacobian matrix which was characterized by its eigenvalues.

Eigenvalues of Jacobian matrices characterize the stability of a metabolic steady state (Steuer *et al.*, 2006; Reznik & Segre, 2010). The time dependent concentration of a metabolite ($c_i(t)$) can be described as the sum of steady state concentration and a fluctuation term τ (Eq. 7)

$$c_i(t) = c_{0,i} + \tau_i(t) \qquad \textbf{Eq. 7}$$

After linearization (Taylor expansion) and integration at the considered steady state the general solution of $\tau_i(t)$ is given by (Eq. 8):

$$\tau_i(t) = \sum_{p=1}^{m} C_{i,p} e^{\lambda_p t} \qquad \textbf{Eq. 8}$$

Here, $\boldsymbol{C}$ represents constants which depend on the initial condition and λ_i represents the eigenvalues of the Jacobian matrix $\boldsymbol{J}$. Thus, the solution of the fluctuation term (Eq. 8) is characterized by the eigenvalue λ_i which generally represents a complex number consisting of a real and an imaginary part. If the real part is negative ($re(\lambda_i) < 0$) the fluctuation term decays

exponentially and leads to a stable solution, while for $re(\lambda_i) > 0$ the fluctuation term increases exponentially resulting in instability. Therefore, analysing eigenvalue real part distributions for simulated Jacobian matrices is sufficient to analyse whether a metabolic steady state remains stable or not (Steuer, 2007). To evaluate the contribution of feedback-inhibition or activation steps to stability characteristics of subcellular carbohydrate metabolism, eigenvalue distributions of Jacobian matrices were calculated for a diverse set of model variants with different combinations of enzymatic activation and inhibition. For each model, 10^6 parameter sets, i.e. variations of $\boldsymbol{\theta}$, were calculated to reduce the probability of mathematical coincidences. Based on the resulting eigenvalue real part distribution, models were classified as stable or unstable. In every iteration step, the environmental fluctuation term (α), the proportion characters (β, γ) and normalized elasticities $\theta_{j,i}$ were randomly chosen within a pre-defined interval while activation and inhibition terms remained constant.

In a first scenario, terms of metabolic activation and feedback-inhibition of cytosolic carbohydrate metabolism were defined as described earlier (Henkel *et al.*, 2011). In detail, hexoses feedforward-activated the biosynthesis of sucrose and feedback-inhibited the cleavage of sucrose. In addition, hexose phosphorylation was feedback-inhibited by reaction products, i.e. phosphorylated sugars. In comparison to the small metabolic network analyzed by Henkel and colleagues which did not account for subcellular compartmentation (Henkel *et al.*, 2011), the subcellular model confirmed a strong stabilizing effect of feedback inhibition of hexose phosphorylation. Variation of subcellular localization of metabolic feedback inhibition finally indicated most efficient stabilization of carbohydrate metabolism to arise from plastidial and cytosolic parts. Vacuolar feedback inhibition was observed to less efficiently stabilize the metabolic steady state after perturbation of the photosynthetic input function. Based on these observations it was hypothesized that metabolic regulation in the vacuolar compartment induces oscillation in metabolic trajectories, i.e. entries of Jacobian matrices, which might connect metabolic information in the vacuole with cytosol and plastids. Supporting this hypothesis, oscillations have earlier been described to encode and transfer information in time and space (Cheong & Levchenko, 2010).

Although normalization of concentrations and/or parameters to steady state conditions supports the analysis of complex metabolic networks, extrapolation and interpretation of derived solutions is strongly limited by the steady state assumption which rarely corresponds to observed system dynamics, e.g. due to developmental or diurnal metabolic changes. Therefore, the development and application of dynamic, time continuous and non-linear mathematical strategies is essential to support the analysis of metabolic regulation in plants.

Publication

A Strategy for Functional Interpretation of Metabolomic Time Series Data in Context of Metabolic Network Information

Thomas Nägele, **Lisa Fürtauer***, Matthias Nagler*, Jakob Weiszmann* and Wolfram Weckwerth*

*(*these authors contributed equally)*

Published in Frontiers in Molecular Biosciences (2016) DOI: 10.3389/fmolb.2016.00006

Experimental analysis of diurnal or developmental dynamics of metabolism results in a comprehensive and multidimensional data matrix including metabolite concentrations, enzyme activities and/or protein levels. Regression and correlation analysis is widely used to statistically characterize complex system dynamics. However, the interpretation of regression and/or correlation output in context of metabolic networks is frequently limited by a missing link to enzymatic reactions. In this study, a method was developed to link regression analysis of experimental data to metabolic network information. Method development was based on a system of ODEs describing time-dependent dynamics of metabolite concentrations in a biochemical network (Eq. 9):

$$\frac{d}{dt}\boldsymbol{M}(t) = \boldsymbol{N}\boldsymbol{v}(\boldsymbol{M}, \boldsymbol{p}, t) = f(\boldsymbol{M}, \boldsymbol{p}, t) \qquad \textbf{Eq. 9}$$

$\boldsymbol{M}$ represents an n-dimensional vector of metabolite concentrations, $\boldsymbol{N}$ $(n \times k)$ the stoichiometric matrix of the metabolic network, $\boldsymbol{v}$ is a k-dimensional vector of reaction rates which depend on metabolite concentrations $\boldsymbol{M}$, $\boldsymbol{p}$ comprises (enzyme) parameters, and t represents the time. Thus, metabolic functions (f_i) are first-order derivatives of metabolite concentrations with respect to time. To (implicitly) derive metabolic functions from dynamics of metabolite concentrations, experimental data were interpolated using time continuous spline interpolations with a weight function based on the inverse variance information. Thus, time points with a low variance of a variable (here: metabolite concentrations) had a strong impact on regression analysis while high variance decreased the impact. Finally, the first derivative of the spline interpolation of metabolite concentrations represented a metabolic function, i.e. the right side of Eq. 9. Based on the stoichiometric matrix of a reaction network, an interaction matrix $\boldsymbol{Y}$ was derived which indicated if a metabolic function depended on a metabolite concentration. For example, substrate molecule *a* affects its own metabolic function and is consumed by the reaction. In this case the interaction matrix indicates that *a* affects its metabolic function, f_a. After variance weighed regression analysis and differentiation the ω equation was derived, in which reaction substrates (here: *a*) and products

(here: b) were combined. Hence, metabolic functions f_a and the time-dependent derivative of f_b were related to each other (Eq. 10):

$$\omega(a \rightarrow b, t) = \frac{\frac{d}{dt} f_b(\boldsymbol{M}, \boldsymbol{p}, t)}{f_a(\boldsymbol{M}, \boldsymbol{p}, t)}, D = \{\mathbb{R} \setminus f_a(\boldsymbol{M}, \boldsymbol{p}, t) = 0\} \qquad \textbf{Eq. 10}$$

Eq. 10 implies that changes in metabolic functions of a reaction product ($\frac{d}{dt} f_b$) are related to dynamics in substrate concentration, f_a. Based on the metabolic network information, ω ratios were only determined if a corresponding entry was present in the interaction matrix $\boldsymbol{Y}$. The numerator showed high absolute entries as soon as point of inflection occurred in the corresponding metabolic function. In a biochemical context this was interpreted as a change in flux and/or enzymatic activity. Hence, ω functions potentially reveal time points of changing reaction dynamics, e.g. due to enzymatic reprogramming. Subsequently, validation of predicted biochemical changes e.g. by enzyme activity assays, transcriptional approaches, proteomics or phosphorylation analyses are needed at the considered time points. Additionally, there might be a connection of $\omega(t)$ with entries of the Jacobian matrix $\boldsymbol{J}$ for infinitesimal time points. The entries of $\boldsymbol{J}$ characterize dynamic properties at a steady state while $\omega(t)$ refers to non-infinitesimal time frames. To ease steps of variance weighed regression analysis, differentiation and division, a MATLAB®-based graphical user interface was programmed (FEMTO, ***F**unctional **E**valuation of **M**etabolic **T**ime-series **O**bservations;* http://mosys.univie.ac.at/resources/software/).

For validation of the mathematical assumptions, a previously published diurnal time-course of carbohydrate concentrations and enzyme activities in *Arabidopsis thaliana* was analyzed before and after cold acclimation (Nägele *et al.*, 2012). Regression analysis and building of $\omega(t)$ functions indicated a strong cold-induced deregulation of SPS-driven sucrose biosynthesis during the initial day phase (Sucrose Phosphate Synthase, $\omega(SP \rightarrow Suc, t)$). This observation coincided with kinetic simulations which had previously shown that during the initial day phase, i.e. 3-4 hours in the light, rates of sucrose biosynthesis differed most significantly between control and cold acclimated plants (Nägele *et al.*, 2012). Conclusively, the approach of building $\omega(t)$ functions to reveal time points of enzymatic regulation in a metabolic network was supported by enzyme kinetic measurements. Additionally, a second published data set on stress-induced dynamics of primary and secondary metabolism was analyzed by building $\omega(t)$ functions (Doerfler *et al.*, 2013). The analysis revealed a fast stress-induced response of sucrose metabolism to combined cold/high light stress ($\omega(Suc \rightarrow Suc, t)$). In contrast, at the interface to secondary metabolism, a late response was detected between phenylalanine (substrate) and flavonoid (product) dynamics

($\omega(Phe \rightarrow Flav, t)$). While, in this study, no enzyme kinetic data were available to validate the prediction of $\omega(t)$ functions, findings still agreed with the output of a previous modelling approach (see (Doerfler *et al.*, 2013)) indicating the suitability of the developed method to analyse reprogramming of metabolic networks.

Publication (under review)

Combined multivariate analysis and machine learning reveals a predictive module of metabolic stress response in Arabidopsis thaliana

Lisa Fürtauer*, *Alice Pschenitschnigg*, Helene Scharkosi1*, Wolfram Weckwerth and Thomas Nägele*

*(*these authors contributed equally)*

Under Review (20.04.2018): Molecular Omics

Stress-induced metabolic reprogramming is necessary to maintain growth and developmental processes of plants. Carbohydrates, as primary products of photosynthesis, play a crucial role in energy metabolism, developmental processes, signalling and many other processes (Pommerrenig *et al.*, 2018). Particularly, regulation of starch and sucrose metabolism affects carbon and energy metabolism on a whole plant level, and mutations in both biosynthetic pathways are known to affect photosynthesis, other metabolite concentrations and stress tolerance (Sicher, 2011). Here, *Arabidopsis* Col-0 wildtype plants and mutant lines with enzymatic deficiency either in the sucrose (*spsa1*) or starch (*pgm1*) biosynthetic pathway were exposed to a combined low temperature/high light stress in order to induce a broad and multi-stress responsive metabolic reprogramming. To reveal a central molecular network of combined abiotic stress response across different metabolic constitutions, i.e. metabolic mutants, primary metabolites and proteins were quantified before and after three days of combined cold (5 °C) and high light (6-fold increase, 300 µmol m^{-2} s^{-1}) application. Control and stress plants were harvested in the early day phase (3h in light) in order to quantify the effect before a daily metabolic equilibrium was reached. Subsequently, stress induced dynamics in photosynthetic performance, primary metabolites, SPS activity, starch levels and proteins were determined. To evaluate whether quantified metabolic patterns were predictable across all genotypes, wildtype data was used to train a machine learning algorithm which was then applied to data of mutant lines in order to differentiate control and stress samples.

All genotypes showed similar significantly changed chlorophyll fluorescence parameters upon stress exposure. In detail, no significant difference between quantum yield of photosystem II,

photochemical quenching and non-photochemical quenching was detected between Col-0, *pgm1* and *spsa1*. Starchless *pgm1* plants showed an increased v_{max} of SPS already under control conditions, and a significantly positive correlation with sucrose concentrations was revealed across all genotypes. Under combined stress, SPS activity decreased in *pgm1* while SPS protein levels increased. The reason for this discrepancy remains elusive, yet it is hypothesized that SPS activity was reduced by phosphorylation (Huber & Huber, 1992; Kaiser & Huber, 2001).

A goal of pattern recognition in experimental plant data sets is to classify and predict developmental stages, stress tolerance or growth conditions based on molecular markers. Such predictive patterns potentially comprise conserved molecular mechanisms in molecular interaction networks. Here, experimental metabolomics and proteomics data sets were classified by support vector machines with linear, quadratic and cubic kernel functions. Molecular wildtype data was used as a training data set and applied to classify molecular data of *pgm1* and *spsa1*. Interestingly, predictability of metabolite data of *pgm1* was weak (~67% accuracy), whereas the metabolome of *spsa1* was accurately classified into control and stress samples (100% accuracy). While also the full experimentally quantified protein data set, comprising 1644 proteins, was not sufficient to accurately classify *pgm1* samples across all SVM kernel functions, a stress responsive core set of 23 proteins, which was identified by an ANOVA, enabled the high accuracy classification of all samples across all genotypes. Analysis of protein-protein interactions, using the String database (https://string-db.org, (Szklarczyk *et al.*, 2014)), revealed a potential interaction network comprising 18 out 23 proteins (Fig. 2). Within this core set, protein functions ranged from stress responsive proteins, molecular chaperones, transcription factors, proteins involved in biosynthesis of proline, primary/secondary and cell wall metabolism. Proteins with most interactions were COR78 (cold regulated 78), P5CS2 (delta 1-pyrroline5-carboxylate synthase 2), F3H (flavanone 3-hydroxylase), TT4 (chalcone and stilbene synthase family protein), ADH1 (alcohol dehydrogenase 1) and PAL1 (phenylalanine ammonia-lyase 1). COR78 is a well described stress responsive protein (Horvath *et al.*, 1993) and a member of the CBF regulon (Thomashow, 1999).

P5CS2 is involved in biosynthesis of proline, which is known to accumulate during stress response in many plants for redox balance, cryoprotection, signalling and many more (Szabados & Savoure, 2010).

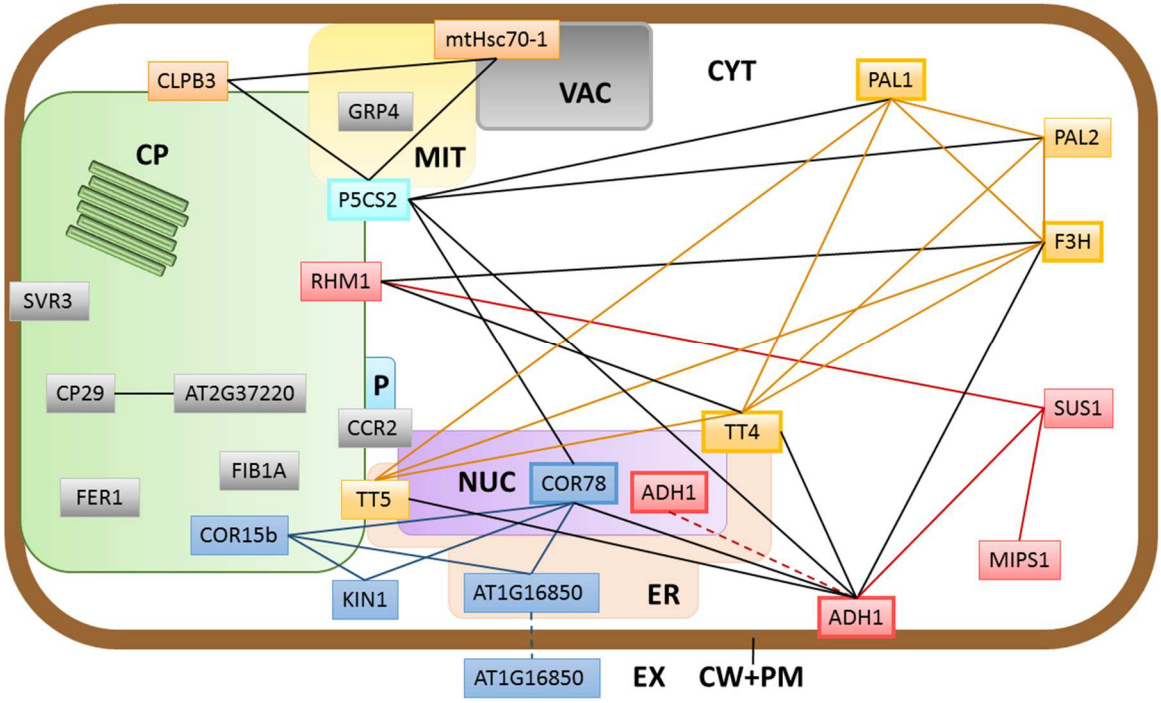


Figure 2 *A stress-induced molecular network with potential protein-protein interaction. CP – Chloroplast, CW+PM – Cell Wall + Plasma Membrane, MIT – Mitochondrion, VAC – Vacuole, NUC – Nucleus, ER – Endoplasmic Reticulum, CYT – Cytosol, P – Peroxisome, EX – Extracellular Region.* ***COR78*** *(cold regulated 78), KIN1 (cold inducible protein kin1), COR15b (Cold regulated 15b),* ***P5CS2*** *(delta 1- pyrroline-5-carboxylate synthase 2), mtHsc70-1 (mitochondrial heat shock protein 70-1), CLPB3 (casein lytic proteinase B3),* ***PAL1****/2 (phenylalanine ammonia lyase 1/2),* ***F3H*** *(flavanone 3-hydroxylase),* ***TT4*** *(chalcone and stilbene synthase family protein), TT5 (chalcone-flavanone isomerase family protein),* ***ADH1*** *(alcohol dehydrogenase 1), SUS1 (sucrose synthase 1), MIPS1 (myo-inositol-1/3-phosphate synthase 1), RHM1 (rhamnose biosynthesis 1), CP29 (chloroplast RNA binding protein 29), GRP4 (glycine-rich RNA binding protein 4), FIB1A (fibrillin precursor protein), FER1 (ferritin 1), SVR3 (elongation factor family protein), CCR2 (cold, circadian rhythm RNA binding 2).*

In the interaction network, P5CS2 connects heat-shock proteins with amino acid metabolism and is linked to the interface between primary and secondary metabolism via PAL1/PAL2 (phenylalanine ammonia-lyase 1/2). Regarding secondary metabolism, TT4 is a key enzyme involved in the biosynthesis of flavonoids, and F3H is deemed to regulate flavonoid biosynthesis (Winkel-Shirley, 2001; Jiang *et al.*, 2015). RHM1 (rhamnose biosynthesis 1) is known to play a major role in supplying UDP-rhamnose for flavonol modifications (Yonekura-Sakakibara *et al.*, 2008), and cell wall synthesis. Additionally, SUS1 (sucrose synthase 1) and MIPS1 (myo-Inositol-1-phosphate synthase), belonged to the identified stress-responsive network. The specific role of SUS in context of cellulose and starch biosynthesis is still under discussion (Baroja-Fernández *et al.*, 2012; Sweetlove & Fernie, 2013) and MIPS1, which plays a central role in inositol biosynthesis, has been found to be directly activated by far red light signalling proteins due to oxidative stress (Ma *et al.*, 2016).

V. Results in form of publications and manuscripts

1. ***A Benchtop Fractionation Procedure for Subcellular Analysis of the Plant Metabolome***

   Lisa Fürtauer*, Wolfram Weckwerth and Thomas Nägele*

   Published in Frontiers in Plant Science (2016)

   DOI: 10.3389/fpls.2016.01912

2. ***Mathematical modelling approaches in plant metabolomics***

   Lisa Fürtauer*, Jakob Weiszmann*, Wolfram Weckwerth and Thomas Nägele*

   *(*these authors contributed equally)*

   In press (04/2018): Plant Metabolomics: Methods and Protocols, Methods in Molecular Biology, vol. 1778, © Springer Science+Business Media

3. ***Approximating the stabilization of cellular metabolism by compartmentalization***

   Lisa Fürtauer *and Thomas Nägele*

   Published in Theory in Biosciences (2016), Vol. 135, Issue 1-2, pp 73-87

   DOI: 10.1007/s12064-016-0225-y

4. ***A Strategy for Functional Interpretation of Metabolomic Time Series Data in Context of Metabolic Network Information***

   Thomas Nägele,* ***Lisa Fürtauer****, Matthias Nagler*, Jakob Weiszmann* and Wolfram Weckwerth*

   *(*these authors contributed equally)*

   Published in Frontiers in Molecular Biosciences (2016)

   DOI: 10.3389/fmolb.2016.00006

5. ***Combined multivariate analysis and machine learning reveals a predictive module of metabolic stress response in Arabidopsis thaliana***

   Lisa Fürtauer*, Alice Pschenitschnigg*, Helene Scharkosi1*, Wolfram Weckwerth and Thomas Nägele*

   *(*these authors contributed equally)*

   Under Review (20.04.2018): Molecular Omics

1. A Benchtop Fractionation Procedure for Subcellular Analysis of the Plant Metabolome

Lisa Fürtauer, Wolfram Weckwerth and Thomas Nägele

Published in Frontiers in Plant Science (2016)

DOI: 10.3389/fpls.2016.01912

Declaration of Authorship:

LF: performed experiments, data evaluation and wrote the paper

WW: performed data evaluation, wrote the paper

TN: conceived the study, performed data evaluation, wrote the paper

frontiers in Plant Science

METHODS
published: 22 December 2016
doi: 10.3389/fpls.2016.01912

A Benchtop Fractionation Procedure for Subcellular Analysis of the Plant Metabolome

Lisa Fürtauer[1], *Wolfram Weckwerth*[1,2] *and Thomas Nägele*[1,2]*

[1] Department of Ecogenomics and Systems Biology, University of Vienna, Vienna, Austria, [2] Vienna Metabolomics Center, University of Vienna, Vienna, Austria

Although compartmentation is a key feature of eukaryotic cells, biological research is frequently limited by methods allowing for the comprehensive subcellular resolution of the metabolome. It has been widely accepted that such a resolution would be necessary in order to approximate cellular biochemistry and metabolic regulation, yet technical challenges still limit both the reproducible subcellular fractionation and the sample throughput being necessary for a statistically robust analysis. Here, we present a method and a detailed protocol which is based on the non-aqueous fractionation technique enabling the assignment of metabolites to their subcellular localization. The presented benchtop method aims at unraveling subcellular metabolome dynamics in a precise and statistically robust manner using a relatively small amount of tissue material. The method is based on the separation of cellular fractions via density gradients consisting of organic, non-aqueous solvents. By determining the relative distribution of compartment-specific marker enzymes together with metabolite profiles over the density gradient it is possible to estimate compartment-specific metabolite concentrations by correlation. To support this correlation analysis, a spreadsheet is provided executing a calculation algorithm to determine the distribution of metabolites over subcellular compartments. The calculation algorithm performs correlation of marker enzyme activity and metabolite abundance accounting for technical errors, reproducibility and the resulting error propagation. The method was developed, tested and validated in three natural accessions of *Arabidopsis thaliana* showing different ability to acclimate to low temperature. Particularly, amino acids were strongly shuffled between subcellular compartments in a cold-sensitive accession while a cold-tolerant accession was characterized by a stable subcellular metabolic homeostasis. Finally, we conclude that subcellular metabolome analysis is essential to unambiguously unravel regulatory strategies being involved in plant-environment interactions.

Keywords: subcellular analysis, compartmentalization, non-aqueous fractionation, metabolome, cold acclimation, *Arabidopsis*

OPEN ACCESS

Edited by:
Basil J. Nikolau,
Iowa State University, USA

Reviewed by:
Damien L. Callahan,
Deakin University, Australia
Ruth Welti,
Kansas State University, USA

***Correspondence:**
Thomas Nägele
thomas.naegele@univie.ac.at

Specialty section:
This article was submitted to Technical Advances in Plant Science, a section of the journal Frontiers in Plant Science

Received: *12 September 2016*
Accepted: *02 December 2016*
Published: *22 December 2016*

Citation:
Fürtauer L, Weckwerth W and Nägele T (2016) A Benchtop Fractionation Procedure for Subcellular Analysis of the Plant Metabolome. Front. Plant Sci. 7:1912. doi: 10.3389/fpls.2016.01912

INTRODUCTION

Eukaryotic cells are characterized by a high degree of compartmentalization establishing a variety of biochemical reaction conditions. In particular, plant cells possess one of the most compartmentalized cell structures across all kingdoms of life. Thus, although true for all eukaryotes, particularly metabolism of plant cells is challenging to analyse due to the high diversity of metabolic pathways (Lunn, 2007). The interconnection of subcellular compartments by various transport and shuttle systems enables a regulated exchange of metabolites across biological membrane systems (Linka and Weber, 2010). A limiting step to unravel such a non-intuitive metabolic system is the lack of knowledge about distribution and dynamics of metabolite concentrations and enzyme activities (Masakapalli et al., 2010; Nägele, 2014). These data are essential for modeling approaches (Nägele, 2014) but also for biotechnological applications, e.g., metabolic engineering (Mintz-Oron et al., 2012). To gain knowledge about compartment specific metabolite levels, Gerhardt and Heldt developed the so-called non-aqueous fractionation (NAF) technique and applied it, e.g., in spinach leaves (Gerhardt and Heldt, 1984) where a density gradient allows the separation of plastids, cytosol and vacuole. Although this technique is time consuming and technically challenging (Stitt et al., 1989; Geigenberger et al., 2011) it is still the method of choice for the assignment of subcellular location of metabolites (Arrivault et al., 2014). The combination of subcellular compartmentation with experimental high-throughput analyses, e.g., mass spectrometry-coupled chromatography techniques, is deemed to be an important part of current and future biological research (Kueger et al., 2012; Tiessen and Padilla-Chacon, 2012). The basic principle of the NAF technique is the subdivision of a cell in small fractions comprising metabolites, lipids, proteins, enzymes, and all other cellular components (**Figure 1A**). These fractions are derived from a snap-freeze procedure of the biological material, typically using liquid nitrogen, followed by the lyophilization of a finely ground powder of the material. With this, all metabolic reactions are immediately stopped and the composition of the fractions remains unaltered due to the loss of reaction solvent. Consequently, quenched marker enzymes cluster together with metabolites, lipids and proteins finally resulting in a compartment-specific fraction density (**Figure 1B**). Typically, vacuolar fractions of plant leaf material possess the highest density followed by cytosolic fragments and plastidial fractions comprising lipid-rich thylakoids. Finally, these fractions are separated in a non-aqueous density gradient via ultracentrifugation. The gradient with the different fractions is then evaluated measuring marker enzyme activities and correlating the relative compound abundances with the activity distribution (Gerhardt and Heldt, 1984; Geigenberger et al., 2011; Nägele and Heyer, 2013; Arrivault et al., 2014).

Although the NAF procedure represents a comprehensive and powerful technique, its application is technically challenging frequently affecting the reproducibility and statistical robustness of the output. Additionally, ultracentrifugation steps over hours are necessary which dramatically limit the throughput capacity of the method.

To overcome these limitations and to ease the accessibility of subcellular metabolite distributions, we re-designed the experimental workflow finally leading to a protocol which enables the fractionation procedure on-the-bench. Additionally, we developed an algorithm to correlate marker enzyme activities and other compounds of interest accounting for technical error propagation. We confirmed our method in the genetic model plant *Arabidopsis thaliana* and provide new insights into the stress-induced subcellular re-organization of the primary metabolome.

A BENCHTOP FRACTIONATION PROCEDURE

The following section comprises a detailed description of the benchtop fractionation protocol. To support the application, the fractionation procedure is illustrated in **Figure 2**. In a first step,

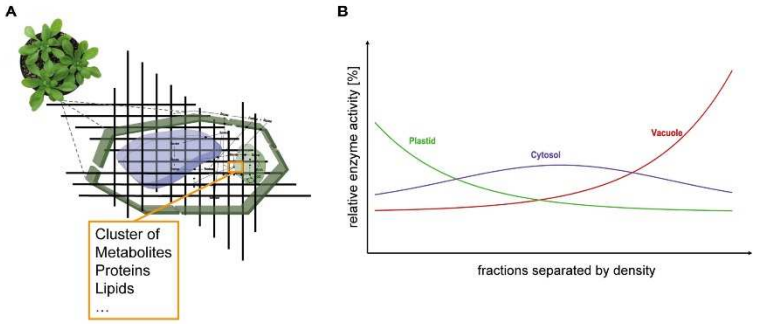


FIGURE 1 | Subcellular analysis of plant leaf tissue. (A) Schematic representation of the dissection of leaf cells. Fractions have specific densities according to their composition of various components (e.g., proteins, metabolites, lipids etc.) and can be separated within a density gradient. **(B)** Ideal distribution of the compartments plastid (green), cytosol (purple), and vacuole (red) separated according to their densities.

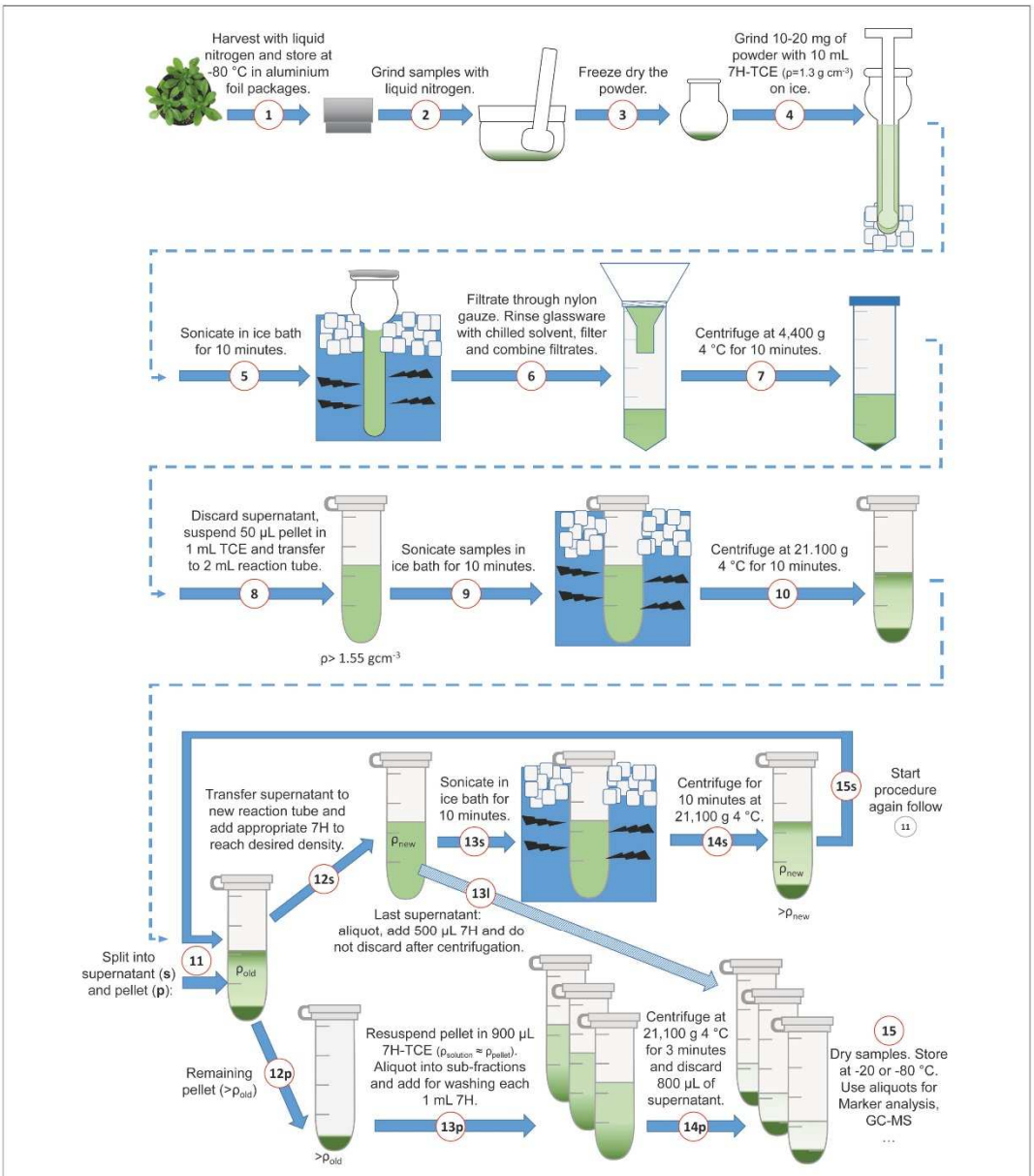


FIGURE 2 | Schematic overview of the benchtop fractionation procedure. Harvesting, grinding of leaf material and sonication (Steps **1–5**); Filtration, collection of material and separation of fractions (Steps **6–9**); Centrifugation (Step **10**); Iterative supernatant ("s") workflow with new density, sonication and centrifugation (Steps **12–15s**); Break-off from the supernatant loop (Step **13l**); Pellet ("p") resuspension, generation of aliquots and washing (Steps **12–24p**); Sample drying (Step **15**). **7H**, n-Heptane; TCE, Tetrachlorethylene.

approximately 10 to 20 mg of finely ground and lyophilized leaf material was suspended in 10 mL of a pre-chilled (4°C) heptane (C_7H_{16}; "7H")-tetrachlorethylene (C_2Cl_4; "TCE")-mixture with a density of $\rho = 1.3\,g\,cm^{-3}$(ratio 7H:TCE = 0.484:1). In a second step, the suspended plant material was ground on ice in a 15 mL glass tissue grinder with a tight mortar (Dounce Tissue Grinder, Wheaton USA) in order to further reduce the average particle size of the plant material (**Figure 2**; Steps **1–4**). After homogenization, the glass vials were placed in an ice water filled ultrasonic bath (Bandelin Sonorex, Typ RK 100, 50/60 Hz), and sonicated for 10 min. Every 2 min, samples were checked to assure that they remained ice-cold. After sonication, the average particle size was in the range of nanometers to low micrometers (**Supplementary Image 1**). The sonicated suspension was filtered through 22–25 μm pore nylon gauze (Miracloth, Calbiochem) into a 50 mL reaction tube. The glass grinder was rinsed with 10 mL of prechilled heptane which was then filtered through the gauze and pooled with the filtered sample. The pooled filtered suspension was centrifuged with 4400 g at 4°C for 10 min (**Figure 2**; Steps **5–7**). The supernatant was discarded as efficient as possible and the remaining pellet, still containing approximately 50 μL of the supernatant, was re-suspended in 1 mL of tetrachlorethylene and transferred to a 2 mL reaction tube (**Figure 2**; Step **8**). The estimation of the supernatant remaining in the pellet was performed using a pipette and is necessary to reliably approximate the first gradient density (fraction "f," indicated as $\rho > 1.55$). The samples were sonicated for 10 min in an ice water ultrasonic bath, and afterwards centrifuged in a benchtop microcentrifuge with 21,100 g at 4°C for 10 min (**Figure 2**; Steps **9–11**; pictures in **Supplementary Images 1**, **2**). After centrifugation, there was a split into supernatant ("s") and pellet ("p"). In step **12s** (**Figure 2**), the supernatant was transferred to a new 2 mL reaction tube and diluted with heptane according to the next gradient step. The amount of the added heptane depended on the next desired density and was calculated as described in Equation (1). The formula contains the sum of mass concentrations of *i* components of a solution divided by the total volume:

$$\rho_{solution} - \frac{1}{V_{total}} \sum_i m_i - \frac{\sum_i \rho_i V_i}{V_{total}}, \qquad (1)$$

Here, m_i represents the mass, V_i the volume and V_{total} the total volume of solution.

For example, given a supernatant volume of 1.050 mL with a density of $1.55\,g\,cm^{-3}$ and a desired density of $1.45\,g\,cm^{-3}$ in the next centrifugation step, inserting all volumes and densities in Equation (1) reveals the volume x of heptane ($\rho = 0.68\,g\,cm^{-3}$) which has to be added:

$$1.45\,g\,cm^{-3} = \frac{1.55\,g\,cm^{-3} \cdot 1.05\,mL + 0.68\,g\,cm^{-3} \cdot x\,mL}{1.05\,mL + x\,mL}$$

Solving for x reveals $x = 0.137$ mL, i.e., 0.137 mL of heptane has to be added to the supernatant.

The reaction tube with the new density was again sonicated in an ice bath for 10 min and afterwards centrifuged with 21,100 g at 4°C for 10 min (**Figure 2**, Steps **13–14s**; pictures in **Supplementary Image 2**). Then, the iterative procedure started again for the second fraction and continued until the final density was reached (**Figure 2**, Step **15s**). Examples for different density gradients are provided in the supplements (**Supplementary Data 1**). The remaining pellets after Step 12s were dissolved in 900 μL heptane-tetrachlorethylene according to the density of the corresponding fraction and were aliquoted into three sub-fractions. Those three sub-fractions were washed with 1 mL heptane and centrifuged for 3 min with 21,100 g at 4°C. Afterwards, 800 μL of the supernatant were discarded (**Figure 2**, Steps **12–14p**). The supernatant of the last fraction "l" (fraction number 1) was aliquoted into sub-fractions, e.g., 3, and 500 μL heptane were added, centrifuged at 21,100 g 4°C for 3 min and the supernatant was kept (**Figure 2**, Step **13l**). Samples were dried in a vacuum concentrator (LaboGene™, Denmark) and stored at −20°C until use for enzyme activity measurement and metabolome analysis (**Figure 2**, Step **15**).

Following the fractionation procedure, marker enzyme activities were determined photometrically, i.e., activity of plastidial pyrophosphatase activity, cytosolic uridine 5′-diphosphoglucose pyrophosphorylase and vacuolar acidic phosphatase. Finally, the relative distribution of marker enzyme activities for plastid, cytosol, and vacuole was determined which allowed for the correlation with metabolite levels.

At this point we would like to make some general technical remarks which we think are of central importance for a successful workflow: all used solvents have been pre-chilled, the work was always performed on ice and under the fume hood and special attention has to be given that samples are not contaminated with water. All open glass grinders were covered with aluminum foil when they were not used. The design and application of density gradients might differ significantly between species, organs, tissues and cell types. While it is hardly predictable which gradient has to be applied to an uncharacterized tissue or cell type, we recommend to begin with a linear gradient covering a large density range, e.g., a gradient ranging from 1.2 to $1.55\,g\,cm^{-3}$ (for detailed example gradients, please refer to **Supplementary Data 1**). Results of the *first* enzyme marker measurements will then indicate whether the chosen linear gradient is suitable for a clear separation or whether more density steps are necessary to resolve a certain density range.

CALCULATION ALGORITHM FOR THE DETERMINATION OF SUBCELLULAR METABOLITE DISTRIBUTIONS

The relative distribution of marker enzyme activities (EA) was assumed to represent the relative distribution of corresponding subcellular compartments across density gradients. To correlate the relative distribution of marker enzyme activities with relative metabolite abundance, changes of both enzyme activities and metabolite abundance were compared pairwise between all measured fractions of one sample. The algorithm is described in detailed steps in the supplements (**Supplementary Data 2**) and is exemplarily shown in **Figure 3**. In brief, extraction and measurements resulted in m Metabolites ($x = \{1, \ldots, m\}$) and f fractions ($i, j = \{1, \ldots, f\}$ *and* $i < j$) for each sample. Slopes

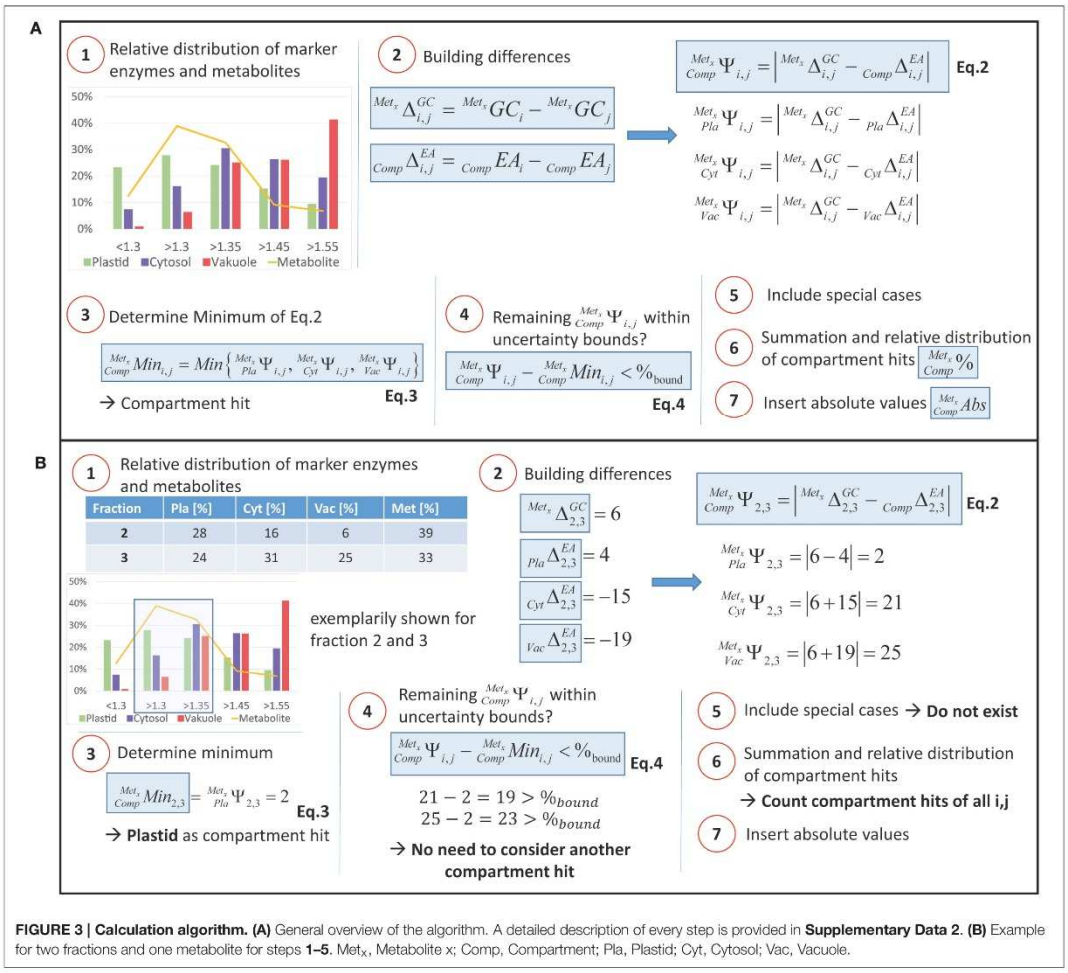


FIGURE 3 | Calculation algorithm. (A) General overview of the algorithm. A detailed description of every step is provided in **Supplementary Data 2**. **(B)** Example for two fractions and one metabolite for steps **1–5**. Met_x, Metabolite x; Comp, Compartment; Pla, Plastid; Cyt, Cytosol; Vac, Vacuole.

between two fractions were built for all marker enzymes and metabolites, resulting in $\binom{f}{2}$ possibilities for every compartment ($Comp^{\Delta^{EA}_{i,j}}$, **Figure 3**, Step **2**) and every metabolite ($^{Met_x}\Delta^{GC}_{i,j}$, **Figure 3**, Step **2**). With this, the absolute ***difference***, i.e., distance, of every metabolite to the compartment was determined, $^{Met_x}_{Comp}\Psi_{i,j}$ (Equation 2):

$$\begin{aligned} {}^{Met_x}_{Pla}\Psi_{i,j} &= \left|{}^{Met_x}\Delta^{GC}_{i,j} - {}_{Pla}\Delta^{EA}_{i,j}\right| \\ {}^{Met_x}_{Cyt}\Psi_{i,j} &= \left|{}^{Met_x}\Delta^{GC}_{i,j} - {}_{Cyt}\Delta^{EA}_{i,j}\right| \\ {}^{Met_x}_{Vac}\Psi_{i,j} &= \left|{}^{Met_x}\Delta^{GC}_{i,j} - {}_{Vac}\Delta^{EA}_{i,j}\right| \end{aligned} \quad (2)$$

Then, the ***minimum*** of Equation (2) was built, $^{Met_x}_{Comp}\mathrm{Min}_{i,j}$ (Equation 3):

$$^{Met_x}_{Comp}\mathrm{Min}_{i,j} = Min\left\{{}^{Met_x}_{Pla}\Psi_{i,j},\ {}^{Met_x}_{Cyt}\Psi_{i,j},\ {}^{Met_x}_{Vac}\Psi_{i,j}\right\}, \quad (3)$$

The minimum was counted as a ***hit*** for the metabolite within the respective compartment (see **Figure 3A**, Step **3**).

Afterwards, all other differences between slopes of metabolites and compartments, i.e., $^{Met_x}_{Comp}\Psi_{i,j}$, were subtracted from the solution of Equation (3) ($^{Met_x}_{Comp}\mathrm{Min}_{i,j}$) (see **Figure 3**, **Step4**). Finally, these differences were compared to pre-defined (uncertainty) ***bounds*** of 5, 7.5, and 10% introducing a methodological standard error in the slope interpretation (Equation 4).

$$ {}^{Met_x}_{Comp}\Psi_{i,j} - {}^{Met_x}_{Comp}\mathrm{Min}_{i,j} < \%_{bound} \quad (4) $$

These bounds were chosen according to the (average) technical errors of photometric marker enzyme activity measurements and metabolite quantification using gas chromatography coupled to mass spectrometry (data not shown).

If inequation (Equation 4) was true, the corresponding compartment was additionally counted as a compartment hit. By including three of these error bounds (5, 7.5, and 10%) instead of only one error bound (e.g., 10%) slight differences between slopes could be differentiated from strong differences because, depending on the extend of the error bound, a different number of hits was assigned to metabolites.

A ***special case*** was assumed when a metabolite was only detected in a single fraction (**Figure 3**, Step **5**). Then, only the compartment with the highest marker enzyme activity in this fraction was taken into account.

Finally, after all these steps, all hits from the compartments were summed up for each metabolite and a ***relative distribution*** was built. The arithmetical mean and standard deviation for every metabolite and compartment ${}^{Met_x}_{Comp}\%$ (**Figure 3**, Step **6**) was determined. To estimate ***absolute subcellular metabolite levels***, values for relative metabolite distribution were multiplied with absolute metabolite levels derived from non-fractionated samples (${}^{Met_x}_{Comp}Abs$; **Figure 3**, Step **7**). A fill-in form for the calculation algorithm is provided in the supplements (**Supplementary Data 3**).

RESULTS

Resolving the Relative Subcellular Distribution of the Primary Metabolome

To demonstrate the output of the fractionation procedure, leaf material of the *Arabidopsis* accession Col-0 was analyzed combining the NAF method with GC-MS analysis of the primary metabolome. The subcellular distribution of the primary leaf metabolome in Col has been analyzed in previous studies (see e.g., Szecowka et al., 2013; Arrivault et al., 2014), and, hence, represents a suitable model tissue for new methodological development. In addition to plants grown under ambient conditions (22°C), we also analyzed leaf material of cold acclimated plants (7 days at 5°C) in order to confirm the capability of the method to resolve stress-induced metabolic reprogramming (Hoermiller et al., 2016). Activities of marker enzymes for the three compartments chloroplast, cytosol, and vacuole showed a characteristically different trend across all fractions being a prerequisite for a reliable correlation with associated metabolite levels (**Figure 4**). Three replicates were found to be sufficient to yield a reproducible and representative mean value for each marker enzyme activity (**Figure 4**). In summary, this provided evidence for a reproducible procedure for plant leaf tissue under control and stress conditions.

To reveal the subcellular metabolome distribution, metabolite levels were correlated with the marker enzyme distribution applying the calculation algorithm described in the previous section. This resulted in the relative distribution of metabolites per compartment, ${}^{Met_x}_{Comp}\%$ (**Table 1**), comprising sugars, sugar alcohols, polyamines, organic acids and amino acids. Comparing means of glucose and fructose the distribution indicated a similar trend among compartments under both conditions. Comparing the relative distribution under both conditions revealed a shift of glucose, fructose, melibiose, and threitol from the vacuolar part into the chloroplast. Additionally, melibiose showed a relative increase in the cytosolic fraction. Significant shifts were observed for the relative distribution of sucrose, raffinose, and galactinol, predominantly from the cytosol into the vacuole. For maltose and myo-inositol we observed a cold-induced shift from the chloroplast into the vacuole while the cytosolic fraction remained similar between the two conditions. The polyamine spermidine was significantly shifted from the plastidial compartment into the vacuole, while for putrescine the relative fraction in chloroplasts and cytosol decreased and the vacuolar fraction significantly increased. Under both conditions, most of the relative amount of organic acids was found to be located in the vacuole, while oxaloacetate and 2-oxoglutarate were rather located in the cytosol. Particularly, the TCA cycle intermediates citric acid, fumarate, malate, and succinate were found to be distributed similarly. Pyruvate was shifted from the vacuole into the plastid during cold

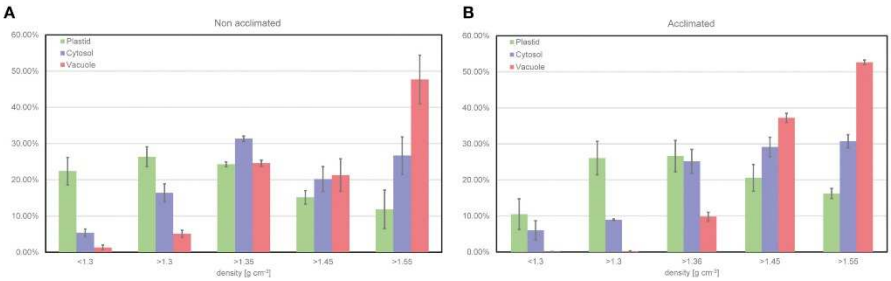


FIGURE 4 | Relative distribution of compartment-specific marker enzyme activities. (A) Distribution resolved for non-acclimated plant leaf material of Col-0 ($n = 3$, MW ± SE), **(B)** Distribution resolved for acclimated plant leaf material of Col-0 ($n = 3$, MW ± SE). Plastid (green), Cytosol (purple) and Vacuole (red).

TABLE 1 | Relative distribution of metabolites in Col-0 before and after cold acclimation.

		Col-0, Relative Distribution: *Mean ± SD* [%]								
		CHLOROPLAST			CYTOSOL			VACUOLE		
METABOLITES		Non Acc	Acc	*p*	Non Acc	Acc	*p*	Non Acc	Acc	*p*
SUGARS/SUGAR ALCOHOLS	Fructose	17.2 ± 5.7	22.6 ± 4.5		29.6 ± 10.4	28.5 ± 3.7		53.2 ± 14.5	48.9 ± 5.3	
	Galactinol	27.7 ± 8	22.1 ± 6.7		45.3 ± 8.4	35.7 ± 5.5	↓*	27 ± 11.9	42.2 ± 6.3	↑**
	Glucose	18.1 ± 4.1	23.7 ± 4.7	↑*	28.9 ± 8.3	28.1 ± 3.7		53 ± 10.4	48.3 ± 6.2	
	Melibiose	23.5 ± 15.4	38.2 ± 5	↑*	21.8 ± 9.4	35.5 ± 2.8	↑**	47.2 ± 32	26.3 ± 7	
	myo-Inositol	40.7 ± 4.8	32.5 ± 6	↓*	38.1 ± 2.6	39.1 ± 5		17 ± 2.4	28.5 ± 2.8	↑***
	Raffinose	25.4 ± 6.5	19.7 ± 8.2		43.9 ± 9.3	32.7 ± 4.4	↓*	26.8 ± 5	43.6 ± 6.5	↑**
	Sucrose	29.7 ± 5.2	27.6 ± 5.2		47.2 ± 2.1	39.4 ± 3.2	↓***	24.6 ± 6.1	33 ± 3.3	↑**
	Threitol	19.9 ± 14.5	23 ± 14.4		27.5 ± 7.7	29.4 ± 15.4		53.8 ± 20.8	48.5 ± 26	
ORGANIC ACIDS	2-Oxoglutarate	32.3 ± 1.8	26.4 ± 14.6		47.9 ± 9.1	39.3 ± 7.4		25.5 ± 2.8	25.1 ± 12.7	
	Citrate	21.2 ± 10.3	20.3 ± 7.8		32.3 ± 10.1	31.9 ± 5.1		46.5 ± 12.3	47.8 ± 9.8	
	Fumarate	19.1 ± 13.7	19.3 ± 6		28.7 ± 11.3	28.4 ± 5.1		52.3 ± 18.8	52.3 ± 5.8	
	Gluconate	33.3 ± 23.8	34.1 ± 1.8		31.8 ± 2.1	34.1 ± 2.5		38.9 ± 30.8	30.7 ± 4.1	
	Malate	23.7 ± 12	21.1 ± 8.9		34.3 ± 11	34.3 ± 5.4		42.1 ± 10.6	44.6 ± 10.8	
	Oxaloacetate	38.3 ± 18.1	34.9 ± 5.1		41.9 ± 14.1	34.6 ± 3.5		19.8 ± 9.2	29.8 ± 6.1	↑*
	Pyruvate	29.1 ± 5.2	33.3 ± 0	↑*	30 ± 4.5	33.3 ± 0		41 ± 8.9	33.3 ± 0	↓*
	Succinate	22.6 ± 12.3	21.2 ± 7.6		31.1 ± 2.8	28.3 ± 4.9		51 ± 8.6	53.7 ± 5.2	
	Threonate	38.3 ± 4.2	35.9 ± 6.4		42.6 ± 5.6	37.6 ± 4.5		18.3 ± 1.8	26.5 ± 5.7	↑*
AMINO ACIDS/POLYAMINES	Alanine	26.9 ± 12	33.2 ± 6.6		51.2 ± 6.2	38 ± 6.4	↓**	27.2 ± 2.8	28.8 ± 5.2	
	Asparagine	46.7 ± 1	31.3 ± 15.4		37.5 ± 5.7	26.4 ± 13.3		14.6 ± 1.2	22.8 ± 16.4	
	Aspartate	45.1 ± 6.3	47.7 ± 4		36.2 ± 3.1	31.1 ± 4.1	↓*	16.8 ± 10.1	22.9 ± 3.8	
	Glutamate	44.4 ± 8.8	47.3 ± 1.1		38.8 ± 5.8	30.9 ± 4.4	↓*	16.7 ± 8.9	22.1 ± 4	
	Glutamine	39.8 ± 11.6	21.8 ± 8.4	↓**	37.8 ± 4.8	33.7 ± 8.5		19.8 ± 9.9	44.5 ± 14	↑**
	Glycine	22.6 ± 11.2	19.2 ± 4		37.1 ± 11.7	32.6 ± 7.7		40.3 ± 10.4	45.6 ± 10.1	
	Isoleucine	26.1 ± 13.2	19.4 ± 4.5		47.7 ± 4.3	31.9 ± 7.7	↓**	34.6 ± 6.4	46.6 ± 10.4	↑*
	Leucine	27 ± 11.8	20.8 ± 3.8		43.4 ± 13	34 ± 5.6		29.6 ± 10	43.5 ± 8.1	↑*
	Lysine	29.3 ± 6.3	18.6 ± 7.6	↓*	41.3 ± 12.9	23.2 ± 10.2	↓*	29.4 ± 11.4	58.2 ± 16.8	↑**
	Methionine	38.4 ± 1.3	33.7 ± 6.1		43.2 ± 4.9	38 ± 4.1	↓*	19.5 ± 7.2	28.9 ± 5.7	↑*
	Ornithine	35.5 ± 3.9	40.2 ± 25.2		45.9 ± 6.9	29.9 ± 15	↓*	18.5 ± 7.5	22 ± 12.9	
	Phenylalanine	25.6 ± 10.1	18.5 ± 7.8		47.4 ± 8.4	26.7 ± 4.1	↓***	32.6 ± 10.1	53.2 ± 9.2	↑**
	Proline	27.6 ± 8.5	29 ± 2.6		47.1 ± 10.2	37.5 ± 4.9	↓*	25.3 ± 9.3	31.6 ± 5.1	
	Serine	28.9 ± 7.7	24.1 ± 2.8		47.7 ± 10.3	37.8 ± 5	↓*	23.4 ± 8.7	35.7 ± 6.1	↑**
	Threonine	30.6 ± 13	25.9 ± 3.9		45.5 ± 15.9	37.9 ± 5.8		20 ± 4.6	34.1 ± 9.4	↑*
	Tryptophan	20.8 ± 8.5	30.9 ± 3.4	↑**	30.1 ± 16	33.2 ± 0.4		49.1 ± 23.4	34.6 ± 2.5	
	Tyrosine	25.4 ± 5.7	16.7 ± 10.5		38.9 ± 14.8	19.9 ± 10.9	↓*	35.7 ± 14.4	64.2 ± 20.6	↑*
	Valine	25.8 ± 12	19.6 ± 3.3		48.3 ± 4.2	32.5 ± 6.3	↓***	30.4 ± 11.2	45.8 ± 9.4	↑*
	Putrescine	38.2 ± 4	37.8 ± 7.7		40.5 ± 3.5	38 ± 5.3		19.1 ± 5	24.2 ± 3.5	↑*
	Spermidine	46.7 ± 4.2	33.4 ± 0.1	↓***	33.6 ± 5.7	33.5 ± 0.3		20.7 ± 1.3	32.6 ± 3.4	↑***

*Mean values of non-acclimated plants (Non Acc) and acclimated plants (Acc). (n ≥ 6); Asterisks indicate significant changes during acclimation per compartment (ANOVA and Tukey, *p < 0.05; **p < 0.01; ***p < 0.001) and arrows (↑; ↓) indicate an increase or decrease.*

acclimation, and threonate showed a significant shift into the vacuole.

Cold-induced differences in the relative distribution were also detected for various amino acids (**Table 1**). Comparing mean values of relative distribution under control condition, two thirds, i.e., 12 out of 18, of the amino acids were found to have their maximum portion in the cytosolic compartment and only 2 out of 18 had their maximum portion in the vacuolar part of the cell. During cold acclimation, the maximum relative portion of most amino acids was found to be vacuolar. In detail, only tryptophan reached a significantly higher percentage in the chloroplast. A slight increase of the plastidial portion was observed for alanine, aspartate, glutamic acid, ornithine and proline, while all others decreased. In the cytosol, only the relative amount of tryptophan increased while the relative amount of all other amino acids decreased during cold acclimation.

Cold-Induced Subcellular Reprogramming in Context of Stress-Tolerance

The subcellular fractionation technique was applied to reveal significant differences between the cold-induced metabolic reprogramming in a cold sensitive (Cvi, origin: Cape Verde Islands) and tolerant (Rsch, origin: Russia) accession of *A. thaliana*. In general, and except for melibiose levels in Cvi, sugars were found to accumulate significantly during the acclimation process. Those differences between absolute levels in Cvi and Rsch have been described before for the same experimental setup (Nagler et al., 2015). Here, we now provide the relative distribution of metabolites among cellular compartments (for details see **Supplementary Data 4**) allowing for the calculation of absolute levels within these compartments (**Table 2**; **Supplementary Data 5**). During cold acclimation, soluble carbohydrates as well as sugar alcohols increased significantly in both accessions. Particularly, sucrose, raffinose, glucose and fructose levels increased during the cold exposure and this increase was more pronounced across all compartments in the tolerant accession Rsch. In contrast, levels of melibiose, which is a degradation product of raffinose, significantly decreased only in the cytosol and vacuole of the sensitive accession Cvi. In addition, the observed fold-change of myo-inositol was higher in Cvi than in Rsch for all compartments (**Table 2**).

For organic acids the picture of cold-induced changes was more diverse. In Cvi, pyruvate levels decreased significantly in all compartments while they increased in plastids and cytosol of Rsch. Citrate significantly increased in chloroplasts of both accessions, yet this increase was more pronounced in Rsch. Additionally, citrate levels increased significantly in the cytosol and vacuole of Rsch which was not observed for Cvi. 2-oxoglutarate was found to increase significantly only in the cytosol of Rsch. Succinate significantly decreased in Cvi and increased in Rsch across all compartments. Fumarate and malate significantly increased in both accessions and again this increase was found to be more pronounced in Rsch (**Table 2**).

Accession-specific dynamics of subcellular absolute levels were also detected for several amino acids. The strongest cold-induced alteration was detected for proline levels in Rsch which were increased 60-fold in the cytosol of acclimated plants (**Table 2**). In Cvi, the strongest proline accumulation was found to occur in the vacuole with a 14-fold increase during cold acclimation. A significantly different picture for both accessions was observed for glutamate which significantly decreased in all compartments of Cvi while it significantly increased in Rsch, pointing to a central role in metabolic reprogramming in the cold tolerant accession.

Hierarchical cluster analysis, which was based on Euclidean distance information of relative subcellular metabolite distribution (**Figure 5A**), revealed a separation of the vacuolar from the plastidial and cytosolic compartment in both acclimation states. In both genotypes, this separation was due to a high vacuolar percentage of metabolites like threitol, glucose, fructose, fumaric acid, citric acid, and malic acid. *Vice versa*, most amino acids and polyamines showed a higher percental distribution in cytosol and chloroplasts across both analyzed genotypes and conditions (**Figure 5A**).

Clustering of absolute subcellular metabolite levels revealed a different pattern where acclimated samples of the cold tolerant accession Rsch showed a much stronger increase of metabolite levels across all compartments than Cvi (**Figure 5B**). Additionally, in chloroplasts and cytosol of Rsch, amino acids, and polyamines showed a much more pronounced accumulation than in Cvi. Also cryoprotective substances like sucrose, raffinose and proline accumulated to higher levels in the cytosolic and plastidial compartments of the tolerant accession Rsch compared to Cvi (**Figure 5B**; **Table 2**; **Supplementary Data 5 Tables SIII, SIV**). Yet, again, and similar to the relative distribution of metabolites, the metabolites glucose, fructose, citric acid, malic acid and fumaric acid increased much stronger in the vacuoles than in other compartments of both genotypes.

DISCUSSION

Due to the high degree of subcellular compartmentation, the interpretation of metabolome data derived from eukaryotic whole-cell extracts is limited. While protein data can often be interpreted in context of subcellular compartments due to characteristic amino acid sequences and enrichment techniques (Millar and Taylor, 2014), accompanying dynamics of the metabolome cannot be resolved equivalently. Here, we present a method which is capable of resolving the subcellular metabolome in a high-throughput benchtop manner by correlating metabolite abundances with compartment-specific marker enzyme activities. The original methodology of non-aqueous fractionation, which our methodology is based on, has been developed and applied successfully for decades (see e.g., Gerhardt and Heldt, 1984; Stitt et al., 1989; Farre et al., 2008; Klie et al., 2011; Krueger et al., 2011; Nägele and Heyer, 2013; Szecowka et al., 2013; Arrivault et al., 2014), yet several technical difficulties still limit its high-throughput application. In this context, our presented technique allows a benchtop fractionation which replaces technically difficult and time-consuming steps, e.g., ultracentrifugation and density gradient mixing of organic solvents. This is particularly enabled by changing the height of the organic liquid column which the sample has to pass through in each centrifugation step. Classically, to separate the cellular fractions by their specific density, ultracentrifugation is applied using tubes with volumes >10 mL (Stitt et al., 1989). In the presented method, this procedure is replaced by step-wise resuspension of tissue material and centrifugation through a small volume of organic solvent (1 mL) which enables the application of benchtop centrifuges and reduces centrifugation periods significantly. Further, in a typical benchtop centrifuge, more than 20 samples can be processed simultaneously while this is not possible in an ultracentrifuge. Hence, this increases the potential sample throughput considerably and, particularly in combination with microplate readers, pipetting robot platforms and hyphenated chromatography-mass spectrometry methods a

TABLE 2 | Comparison of subcellular metabolite levels before and after cold acclimation.

		Ratio of Subcellular Absolute Levels: acc vs. non-acc					
		CHLOROPLAST		CYTOSOL		VACUOLE	
METABOLITES		Cvi	Rsch	Cvi	Rsch	Cvi	Rsch
SUGARS/SUGAR ALCOHOLS	Fructose	26.59***	81.08***	27.75***	89.57***	35.35***	107.36***
	Galactinol	1.08	2.21*	1.51*	3.08***	1.08	6.12*
	Glucose	17.18***	20.68***	20.02***	21.02***	20.58***	22.81***
	Melibiose	0.96	1.25	0.92***	1.05	0.89**	1.06
	myo-Inositol	6.40***	4.83***	9.70***	8.49***	12.51***	9.09**
	Raffinose	1.32	87.81***	1.22	109.70***	2.88*	199.74*
	Sucrose	3.63***	8.77***	3.47***	13.08***	3.54***	18.81**
	Threitol	0.28**	1.28	0.28**	1.09	1.53*	0.81
ORGANIC ACIDS	2-Oxoglutarate	0.96	0.83	0.87	1.34*	0.79	0.81
	Citrate	1.43*	6.09***	1.13	7.78***	0.99	5.88***
	Fumarate	6.10***	8.13***	6.23***	5.91***	8.86***	4.46***
	Gluconate	0.94***	1.47	0.94***	1.42*	0.94***	0.73
	Malate	3.18***	8.54***	2.48***	9.49***	2.63***	7.00***
	Oxaloacetate	1.47	1.13	1.47	1.04	0.51*	1.05
	Pyruvate	0.70***	1.31*	0.64***	1.59**	0.64***	1.02
	Succinate	0.24***	1.74*	0.24***	2.02**	0.37***	1.97***
	Threonate	0.56***	2.09***	1.05	2.78***	1.28	4.15**
AMINO ACIDS/POLYAMINES	Alanine	0.67**	1.47*	0.86	1.70**	1.10	1.43
	Asparagine	0.70	0.80	1.48	1.22	1.84	0.92
	Aspartate	1.64***	3.25***	1.76***	3.79***	2.07***	2.89*
	Glutamate	0.27***	1.65**	0.32***	2.22***	0.36***	2.72**
	Glutamine	0.52***	1.74*	1.32*	2.90***	3.44***	5.57**
	Glycine	1.35	1.48**	1.69***	2.19***	3.66***	3.23***
	Isoleucine	0.73*	1.38	1.29**	1.71**	1.47**	1.87
	Leucine	0.72*	1.27	1.16	1.37*	1.28	1.31
	Lysine	0.25**	0.43***	0.47**	0.52**	2.83*	2.54**
	Methionine	0.54*	0.84	1.24	1.27	1.46	1.00
	Ornithine	0.52	0.81	1.77	1.21	1.84**	1.69
	Phenylalanine	0.32***	0.85	0.65**	1.02	1.48**	1.08
	Proline	7.65***	41.55***	14.86***	60.44***	13.97***	56.76**
	Serine	0.56***	0.99	1.14	1.48*	1.22*	1.70
	Threonine	1.54*	1.72*	1.46***	2.40***	1.49**	2.10*
	Tryptophan	0.68	0.91	0.59*	0.94	1.41	1.31
	Tyrosine	0.46*	0.80	0.51	1.18	2.99*	2.53***
	Valine	0.74*	1.58**	1.18	1.84***	1.20	1.64*
	Putrescine	0.88	1.38**	1.88***	1.76***	3.54***	2.22***
	Spermidine	0.93***	0.82	0.93***	0.77*	0.93***	2.49**

Ratios were built from mean values of cold acclimated and non-cold acclimated samples (acc/non-acc). A detailed summary of all values including information about standard deviation is provided in ***Supplementary Data 4, 5****. Asterisks indicate significant ratios (ANOVA and Tukey, *$p < 0.05$; **$p < 0.01$; ***$p < 0.001$).*

comprehensive and statistically robust insight into subcellular metabolome dynamics becomes possible.

A second major advancement is the possibility to adjust gradients of organic solvent in a non-linear manner which is hardly possible by (most) gradient mixings systems. Additionally, testing the quality of mixed colorless organic solvent gradients in an ultracentrifugation tube is not possible which dramatically impacts the quality control of the fractionation method before the time consuming ultracentrifugation step. In our presented procedure such a control of correct solvent density is automatically given as the solution is prepared in the desired density directly before the sample is re-suspended in it.

Third, compared to the original method, which suggests the application of 200–300 mg of dry leaf powder (Gerhardt and Heldt, 1984; Stitt et al., 1989), only a fraction of this is needed in the presented method (e.g., 10–20 mg of dried leaf material).

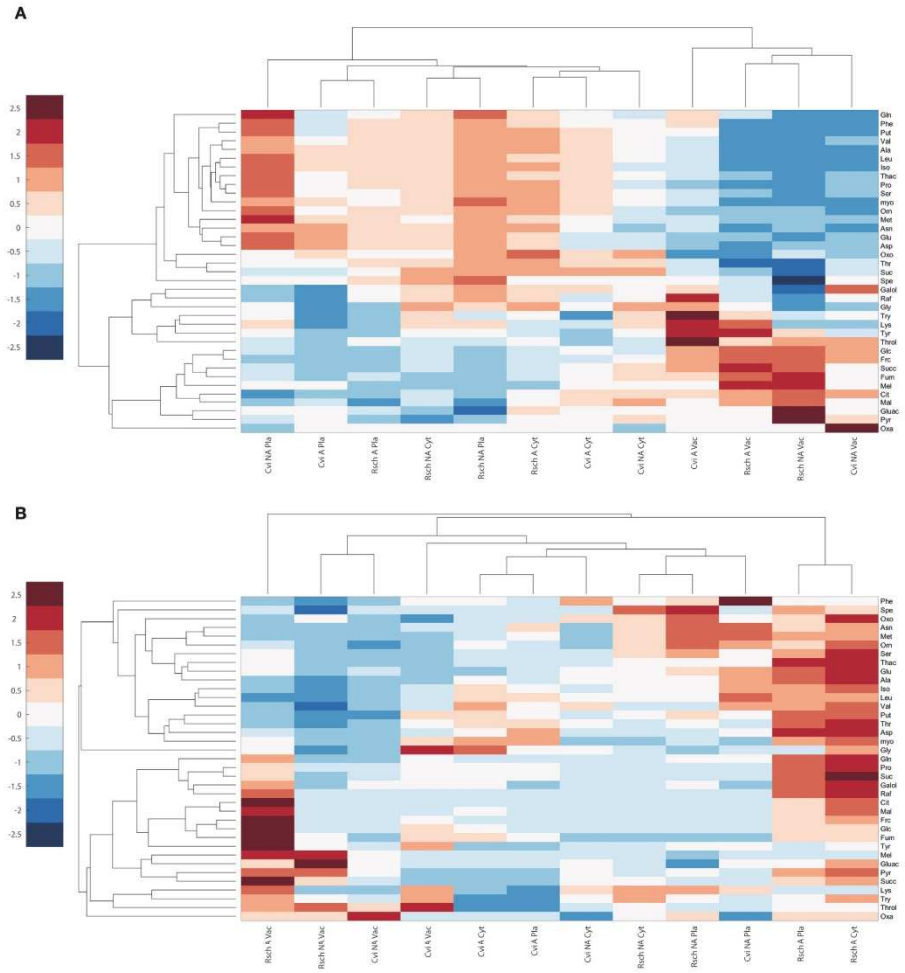


FIGURE 5 | Cluster analysis of relative and absolute subcellular metabolite distributions in Cvi and Rsch. (A) Relative distribution of metabolites, **(B)** Absolute levels per compartment. Mean values of relative and absolute levels of non-acclimated (NA) and acclimated (A) samples ($n \geq 6$) were scaled by zero mean-unit variance (z-scores). Frc, Fructose; Galol, Galactinol; Glc, Glucose; Mel, Melibiose; myo, myo-Inositol; Raf, Raffinose; Suc, Sucrose; Throl, Threitol; Cit, Citrate; Fum, Fumarate; Gluac, Gluconate; Mal, Malate; Oxa, Oxaloacetate; Oxo, 2-Oxoglutarate; Pyr, Pyruvate; Succ, Succinate; Thac, Threonate; Ala, Alanine; Asn, Asparagine; Asp, Aspartate; Glu, Glutamate; Gln, Glutamine; Iso, Isoleucine; Leu, Leucine; Lys, Lysine; Met, Methionine; Orn, Ornithine; Phe, Phenylalanine; Pro, Proline; Ser, Serine; Thr, Threonine; Try, Tryptophan; Tyr, Tyrosine; Val, Valine; Spe, Spermidine; Put, Putrescine; Pla, Plastid; Cyt, Cytosol; Vac, Vacuole.

This makes the whole procedure applicable to studies which are limited by material, e.g., *in situ* or "in-field" samples, rare organ or tissue material.

Finally, the step-wise centrifugation of material suspensions allows for additional supporting ultrasonication steps which contribute to a more efficient lysis of cell component fractions being a prerequisite for a successful separation. The efficiency of sonication was already shown for tobacco cells, where sonication cycles of 6 min lead to more than 99% of disrupted cells (Hu and Brown, 1994). This fine-tunes the separation efficiency of cellular compartments and contributes to a higher resolution capability of the method. Particularly with regard to the separation of small organelles, e.g., peroxisomes or mitochondria, which are frequently associated with other subcellular compartments and, hence, difficult to resolve (Arrivault et al., 2014), this might play a crucial role.

When applying the fractionation technique to non-acclimated and acclimated *Arabidopsis* plants of the accession Col-0, marker enzyme activities indicated a reproducible separation of compartments irrespective of the acclimation state (**Figure 4**). Hence, although leaves might undergo changes in their physical characteristics during cold exposure (Strand et al., 1999), these changes do not seem to affect the quality of separation. However, while the presented method accounts for the subcellular fractionation of homogenized leaf material, it does not resolve the heterogeneity of leaf tissue. Hence, if the portion of epidermal, palisade or mesophyll cells undergoes developmentally or environmentally-induced dynamics, the direct comparison of resolved subcellular metabolite levels will become difficult or, in the worst case, even misleading. A strategy to reduce the resulting ambiguousness of the experimental data could be the application of the presented method to isolated protoplasts or a combination of the presented method with single-cell methods, e.g., imaging methods (Wuyts et al., 2010) or laser ablation electrospray ionization (LAESI) mass spectrometry (MS) (see e.g., Li et al., 2015).

Applying the suggested calculation algorithm as described (see **Figure 3**), several significant cold-induced shifts of metabolite distribution were detected. A dominating effect was the shift of relative amounts of sugars, sugar alcohols and amino acids from the cytosol into the vacuole (see **Table 1**). Hence, these findings provide evidence for a central role of the cytosol-vacuole interaction during cold-acclimation in Col-0. The importance of the vacuolar compartment in cold-induced metabolic reprogramming has also been indicated in a previous tonoplast proteome study which could identify several membrane proteins which were altered in their abundance during cold acclimation (Schulze et al., 2012). Based on their observations, Schulze and colleagues concluded that cold-induced vacuolar solute accumulation occurs due to increased acidification, and transport activity across the tonoplast was suggested to be modulated by protein amounts as well as phosphorylation states. Another study provided evidence for a central role of proton-coupled vacuolar glucose transport in the development of freezing tolerance (Klemens et al., 2014). While this substantiates our indications for a significant role of vacuolar sugar metabolism during cold acclimation, the role of vacuolar reprogramming of amino acid metabolism seems less clear. The relative distribution of amino acids in Col-0 was shifted toward the vacuole during cold acclimation whereas the plastid-cytosol interaction was less affected (see **Table 1**). Due to the fact that the chloroplast plays a central role in the regulation of various amino acid-related pathways, e.g., the biosynthetic pathways of glutamate, glutamine, branched chain and aromatic amino acids (Kleffmann et al., 2004; Maeda and Dudareva, 2012), a possible explanation for our observation would be that the vacuole buffers subcellular re-arrangements against environmental fluctuations. This could contribute to the stabilization of the plastidial amino acid metabolism playing a crucial role in the whole plant C/N homeostasis.

To reveal whether this hypothesis of vacuolar buffering of amino acid metabolism might also be evident for other natural accessions of *Arabidopsis*, we further analyzed the cold-induced subcellular metabolome dynamics in the cold-sensitive and tolerant accessions Cvi and Rsch, respectively. Indeed, we observed a more significant shift of amino acids between plastid, cytosol and vacuole in the sensitive Cvi accession, while in Rsch only a few metabolites were affected in their relative distribution during cold acclimation. Hence, to summarize our findings about relative shifts of metabolite levels during cold acclimation with respect to the freezing tolerance of the considered accessions (Cvi<Col<Rsch, see e.g., Hannah et al., 2006) we suggest that the intensity of significant subcellular re-arrangement of primary metabolism is negatively correlated with freezing tolerance. In contrast, the absolute amount of substance of many sugars, amino acids and organic acids has been shown to positively correlate with freezing tolerance, not only on a cellular (Klotke et al., 2004; Nagler et al., 2015) but also on a subcellular level (Nägele and Heyer, 2013). These significant differences between absolute subcellular metabolite levels became also clear in the present study (**Figure 5B**; **Table 2**; **Supplementary Data 5 Tables SIII, SIV**). While the molecular reasons for these differences remain elusive it is tempting to speculate that the regulation of the energy homeostasis plays a crucial role and might differ between Cvi and Rsch. While the cold sensitive accession Cvi seems to be characterized by a strong cold-induced shuffling of metabolites between subcellular compartments, the cold- and freezing-tolerant accession Rsch might coordinate its metabolism rather in direction of biosynthesis than intracellular transport of cryoprotective substances. Of course, it is not surprising that energy balance plays a central role in cold acclimation and stress tolerance as it has already been nicely summarized almost two decades ago (Huner et al., 1998). However, recent advances in understanding a plant's energy homeostasis and its reprogramming due to environmental fluctuation and stress conditions have revealed a highly complex regulatory and biochemical network (for an overview see e.g., Tome et al., 2014; Provart et al., 2016). These metabolic and regulatory networks comprise various cellular organelles and their interaction via signaling cascades and shuttles. Hence, it can be expected that a comprehensive, reliable and realistic picture of eukaryotic metabolism and its control can only arise if the subcellular compartmentation of metabolism is considered in experimental studies. Our presented methodology for cellular fractionation aims at supporting these biochemical and physiological studies. Conclusively, the output gained from subcellular studies combining relative with absolute metabolite levels can profoundly enhance the output of metabolomics studies and helps to unravel cellular players and regulatory strategies which cannot be inferred from whole-cell approaches.

EXPERIMENTAL PROCEDURES

Plant Cultivation and Sampling

Plants of the *A. thaliana* accessions Col-0, Cvi and Rsch were cultivated in a growth chamber under controlled conditions. The substrate was composed of Einheitserde ED63 and perlite. Plants were watered daily and fertilized once with NPK fertilization

solution (WUX-AL Super; MANNA°-Dünger, Ammerbuch). Light intensity was set to 75 μmol m^{-2} s^{-1} in an 8/16 h day/night cycle and relative humidity to 70 % with a temperature of 22°/16°C. Following 28 days after of sowing light intensity was increased to 125 μmol m^{-2} s^{-1} in a 16/8 h day/night cycle. Bolting stage was reached 43 days after sowing and samples of non-acclimated (non-acc, NA) plants were collected from the three accessions at midday, i.e., after 8 h in the light. One biological sample consisted of 3 leaf rosettes which were immediately quenched in liquid nitrogen and stored until further use at −80°C. Non-sampled plants were transferred to 5°C for acclimation (acc, A) with conditions of light intensity, cycle and humidity as before. After 7 days at 5°C, leaf rosettes were sampled as described for non-acclimated plants.

Measurements of Marker Enzyme Activities

One fraction aliquot was used for photometric (Multiscan Spectrum, Thermo Scientific) determination of three marker enzymes. Samples were extracted in 1 mL extraction buffer consisting of 50 mM Tris-HCl pH 7.3, 5 mM $MgCl_2$, and 1 mM DTT. The marker for the plastidial compartment was alkaline pyrophosphatase as described previously (Jelitto et al., 1992) with modifications of amounts to down-scale the assay to a total volume of 200 μL to make it suitable for photometric measurement in a 96 well-plate (Crystal Clear, Greiner Bio One). The cytosolic marker was uridinediphosphate glucose pyrophosphorylase (UGPase) as described in Zrenner et al. (1993) with modifications regarding recalculations of amounts in a 96 well-plate (μClear, Greiner Bio One). Acid phosphatase was used as a marker for the vacuolar compartment as described in Boller and Kende (1979), again adapted for a 96 well plate (Crystal Clear, Greiner Bio One).

GC-MS Analysis of Subcellular Metabolite Levels

Metabolites were extracted from subfractions according to Weckwerth et al. (2004) with slight modifications. 1 mL of −20°C cold methanol/chloroform/H_2O (2.5/1/0.5 v/v/v) mixture was added to each sample fraction. Samples were vortexed, incubated on ice for 10 min and centrifuged for 4 min at 4°C and 20.000 g. 300 μL H_2O were added to the supernatant, followed by brief vortexing and 2 min of centrifugation. Following the separation of the polar methanol/water phase from the unpolar chloroform phase, samples were dried in a vacuum concentrator (LaboGene™, Denmark) for derivatization. The dried pellets were resolved at 30°C for 90 min in 20 μL of a 40 mg mL^{-1} methoxyamine hydrochloride in pyridine solution. Afterwards, 80 μL of N-methyl-N-trimethylsilyltrifluoroacetamid (MSTFA), spiked with 30 μL mL^{-1} of a mix of even-numbered alkanes (C_{10}-C_{40}), were added, and the samples were incubated for 30 min at 37°C under constant shaking, followed by 2 min of centrifugation. The supernatant was transferred into a glass vial for measurement. GC-MS measurements were performed on an Agilent 6890 gas chromatograph coupled to a LECO Pegasus® 4D GCxGC-TOF mass spectrometer (LECO® Corporation, Michigan, USA). For GC analysis, the initial oven temperature was set to 70°C for 1 min, followed by a 9°C min^{-1} ramp with 330°C end temperature which was set constant for 8 min. In the MS method, the data acquisition rate was set to 20 spectra sec^{-1} at a detector voltage of 1550 V. The acquisition delay was set to 5.5 min and the detected mass range was set from 40 to 600 m/z. Raw data were processed with the LECO Chroma-TOF® software (LECO® Corporation, Michigan, USA). Measurements of non-acclimated samples were performed in a splitless measurement mode (1:1) while cold acclimated samples were measured in split mode (1:2). Relative distribution of metabolites was calculated based on the quantified area in the chromatograms.

Data Analysis and Statistics

Data evaluation was performed in Microsoft Excel (http://www.microsoft.com). Outlier identification and hierarchical cluster analysis was performed within the numerical software Matlab (www.mathworks.com). Analysis of Variance (ANOVA) and Tukey *post-hoc* test were done in the R software environment (The R Project for Statistical Computing; http://www.r-project.org/).

AUTHOR CONTRIBUTIONS

LF performed experiments, data evaluation and wrote the paper. WW performed data evaluation and wrote the paper. TN conceived the study, performed data evaluation and wrote the paper.

FUNDING

This work was supported by the Austrian Science Fund (FWF Project number I2071).

ACKNOWLEDGMENTS

We thank the members of the Department of Ecogenomics and Systems Biology at the University of Vienna, particularly Matthias Nagler and Ella Nukarinen, Jakob Weiszmann, Reinhard Turetschek, and Lena Fragner for fruitful discussion and constructive advice. We thank our collaboration partners from the Department of Plant Biotechnology, University of Stuttgart, headed by Prof. Arnd G. Heyer, for constructive discussions and advice. Finally, we would like to thank the gardeners Andreas Schröfl and Thomas Joch for excellent plant cultivation.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: http://journal.frontiersin.org/article/10.3389/fpls.2016.01912/full#supplementary-material

Supplementary Image 1 | Suspension of lyophilized leaf material. (A,B) Sample before and **(C,D)** after sonication. **(A,C)** Leaf material suspension; **(B,D)** light microscopy pictures reveal the effect of sonication on leaf cells.

Supplementary Image 2 | Results of subcellular fractionation steps. (A–C) Fractions of different density after step 11 (see **Figure 2**). **(D–F)** Fractions after step 14 before drying (see **Figure 2**).

Supplementary Data 1 | Examples for density gradient modifications.

Supplementary Data 2 | Description of calculation algorithm.

Supplementary Data 3 | Fill-in form for measurements.

Supplementary Data 4 | Relative compartmental distribution of metabolites in Cvi and Rsch.

Supplementary Data 5 | Absolute compartmental distribution of metabolites in Cvi and Rsch.

REFERENCES

Arrivault, S., Guenther, M., Florian, A., Encke, B., Feil, R., Vosloh, D., et al. (2014). Dissecting the subcellular compartmentation of proteins and metabolites in *Arabidopsis* leaves using non-aqueous fractionation. *Mol. Cell Proteomics* 13, 2246–2259. doi: 10.1074/mcp.M114.038190

Boller, T., and Kende, H. (1979). Hydrolytic enzymes in the central vacuole of plant cells. *Plant Physiol.* 63, 1123–1132. doi: 10.1104/pp.63.6.1123

Farre, E. M., Fernie, A. R., and Willmitzer, L. (2008). Analysis of subcellular metabolite levels of potato tubers (*Solanum tuberosum*) displaying alterations in cellular or extracellular sucrose metabolism. *Metabolomics* 4, 161–170. doi: 10.1007/s11306-008-0107-5

Geigenberger, P., Tiessen, A., and Meurer, J. (2011). Use of non-aqueous fractionation and metabolomics to study chloroplast function in *Arabidopsis*. *Methods Mol. Biol.* 775, 135–160. doi: 10.1007/978-1-61779-237-3_8

Gerhardt, R., and Heldt, H. W. (1984). Measurement of subcellular metabolite levels in leaves by fractionation of freeze-stopped material in nonaqueous media. *Plant Physiol.* 75, 542–547. doi: 10.1104/pp.75.3.542

Hannah, M. A., Wiese, D., Freund, S., Fiehn, O., Heyer, A. G., and Hincha, D. K. (2006). Natural genetic variation of freezing tolerance in *Arabidopsis*. *Plant Physiol.* 142, 98–112. doi: 10.1104/pp.106.081141

Hoermiller, I. I., Naegele, T., Augustin, H., Stutz, S., Weckwerth, W., and Heyer, A. G. (2016). Subcellular reprogramming of metabolism during cold acclimation in *Arabidopsis thaliana*. *Plant Cell Environ.* doi: 10.1111/pce.12836. [Epub ahead of print].

Hu, H., and Brown, P. H. (1994). Localization of boron in cell walls of squash and tobacco and its association with pectin (evidence for a structural role of boron in the cell wall). *Plant Physiol.* 105, 681–689. doi: 10.1104/pp.105.2.681

Huner, N. P. A., Öquist, G., and Sarhan, F. (1998). Energy balance and acclimation to light and cold. *Trends Plant Sci.* 3, 224–230. doi: 10.1016/S1360-1385(98)01248-5

Jelitto, T., Sonnewald, U., Willmitzer, L., Hajirezeai, M., and Stitt, M. (1992). Inorganic pyrophosphate content and metabolites in potato and tobacco plants expressing *E. coli* pyrophosphatase in their cytosol. *Planta* 188, 238–244. doi: 10.1007/BF00216819

Kleffmann, T., Russenberger, D., von Zychlinski, A., Christopher, W., Sjölander, K., Gruissem, W., et al. (2004). The *Arabidopsis thaliana* chloroplast proteome reveals pathway abundance and novel protein functions. *Curr. Biol.* 14, 354–362. doi: 10.1016/j.cub.2004.02.039

Klemens, P. A., Patzke, K., Trentmann, O., Poschet, G., Büttner, M., Schulz, A., et al. (2014). Overexpression of a proton-coupled vacuolar glucose exporter impairs freezing tolerance and seed germination. *New Phytol.* 202, 188–197. doi: 10.1111/nph.12642

Klie, S., Krueger, S., Krall, L., Giavalisco, P., Flügge, U. I., Willmitzer, L., et al. (2011). Analysis of the compartmentalized metabolome - a validation of the non-aqueous fractionation technique. *Front. Plant Sci.* 2:55. doi: 10.3389/fpls.2011.00055

Klotke, J., Kopka, J., Gatzke, N., and Heyer, A. G. (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. doi: 10.1111/j.1365-3040.2004.01242.x

Krueger, S., Giavalisco, P., Krall, L., Steinhauser, M.-C., Büssis, D., Usadel, B., et al. (2011). A topological map of the compartmentalized *Arabidopsis thaliana* leaf metabolome. *PLoS ONE* 6:e17806. doi: 10.1371/journal.pone.0017806

Kueger, S., Steinhauser, D., Willmitzer, L., and Giavalisco, P. (2012). High-resolution plant metabolomics: from mass spectral features to metabolites and from whole-cell analysis to subcellular metabolite distributions. *Plant J.* 70, 39–50. doi: 10.1111/j.1365-313X.2012.04902.x

Li, H., Smith, B. K., Shrestha, B., Márk, L., and Vertes, A. (2015). "Automated cell-by-cell tissue imaging and single-cell analysis for targeted morphologies by laser ablation electrospray ionization mass spectrometry," in *Mass Spectrometry Imaging of Small Molecules*, ed L. He (New York, NY: Springer), 117–127.

Linka, N., and Weber, A. P. (2010). Intracellular metabolite transporters in plants. *Mol. Plant* 3, 21–53. doi: 10.1093/mp/ssp108

Lunn, J. E. (2007). Compartmentation in plant metabolism. *J. Exp. Bot.* 58, 35–47. doi: 10.1093/jxb/erl134

Maeda, H., and Dudareva, N. (2012). The shikimate pathway and aromatic amino acid biosynthesis in plants. *Annu. Rev. Plant Biol.* 63, 73–105. doi: 10.1146/annurev-arplant-042811-105439

Masakapalli, S. K., Le Lay, P., Huddleston, J. E., Pollock, N. L., Kruger, N. J., and Ratcliffe, R. G. (2010). Subcellular flux analysis of central metabolism in a heterotrophic *Arabidopsis* cell suspension using steady-state stable isotope labeling. *Plant Physiol.* 152, 602–619. doi: 10.1104/pp.109.151316

Millar, A. H., and Taylor, N. L. (2014). Subcellular proteomics-where cell biology meets protein chemistry. *Front. Plant Sci.* 5:55. doi: 10.3389/fpls.2014.00055

Mintz-Oron, S., Meir, S., Malitsky, S., Ruppin, E., Aharoni, A., and Shlomi, T. (2012). Reconstruction of *Arabidopsis* metabolic network models accounting for subcellular compartmentalization and tissue-specificity. *Proc. Natl. Acad. Sci. U.S.A.* 109, 339–344. doi: 10.1073/pnas.1100358109

Nägele, T. (2014). Linking metabolomics data to underlying metabolic regulation. *Front. Mol. Biosci.* 1:22. doi: 10.3389/fmolb.2014.00022

Nägele, T., and Heyer, A. G. (2013). Approximating subcellular organisation of carbohydrate metabolism during cold acclimation in different natural accessions of *Arabidopsis thaliana*. *New Phytol.* 198, 777–787. doi: 10.1111/nph.12201

Nagler, M., Nukarinen, E., Weckwerth, W., and Nägele, T. (2015). Integrative molecular profiling indicates a central role of transitory starch breakdown in establishing a stable C/N homeostasis during cold acclimation in two natural accessions of *Arabidopsis thaliana*. *BMC Plant Biol.* 15:284. doi: 10.1186/s12870-015-0668-1

Provart, N. J., Alonso, J., Assmann, S. M., Bergmann, D., Brady, S. M., Brkljacic, J., et al. (2016). 50 years of *Arabidopsis* research: highlights and future directions. *New Phytol.* 209, 921–944. doi: 10.1111/nph.13687

Schulze, W. X., 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. doi: 10.1111/j.1365-313X.2011.04812.x

Stitt, M., Lilley, R. M., Gerhardt, R., and Heldt, H. W. (1989). Metabolite levels in specific cells and subcellular compartments of plant leaves. *Meth. Enzym.* 174, 518–582. doi: 10.1016/0076-6879(89)74035-0

Strand, A., Hurry, V., Henkes, S., Huner, N., Gustafsson, P., Gardeström, P., et al. (1999). Acclimation of *Arabidopsis* leaves developing at low temperatures. Increasing cytoplasmic volume accompanies increased activities of enzymes in the Calvin cycle and in the sucrose-biosynthesis pathway. *Plant Physiol.* 119, 1387–1397. doi: 10.1104/pp.119.4.1387

Szecowka, M., Heise, R., Tohge, T., Nunes-Nesi, A., Vosloh, D., Huege, J., et al. (2013). Metabolic fluxes in an illuminated *Arabidopsis rosette*. *Plant Cell* 25, 694–714. doi: 10.1105/tpc.112.106989

Tiessen, A., and Padilla-Chacon, D. (2012). Subcellular compartmentation of sugar signaling: links among carbon cellular status, route of sucrolysis,

sink-source allocation, and metabolic partitioning. *Front. Plant Sci.* 3:306. doi: 10.3389/fpls.2012.00306

Tome, F., Nägele, T., Adamo, M., Garg, A., Marco-Llorca, C., Nukarinen, E., et al. (2014). The low energy signaling network. *Front. Plant Sci.* 5:353. doi: 10.3389/fpls.2014.00353

Weckwerth, W., Wenzel, K., and Fiehn, O. (2004). Process for the integrated extraction, identification and quantification of metabolites, proteins and RNA to reveal their co-regulation in biochemical networks. *Proteomics* 4, 78–83. doi: 10.1002/pmic.200200500

Wuyts, N., Palauqui, J. C., Conejero, G., Verdeil, J. L., Granier, C., and Massonnet, C. (2010). High-contrast three-dimensional imaging of the Arabidopsis leaf enables the analysis of cell dimensions in the epidermis and mesophyll. *Plant Methods* 6:17. doi: 10.1186/1746-4811-6-17

Zrenner, R., Willmitzer, L., and Sonnewald, U. (1993). Analysis of the expression of potato uridinediphosphate-glucose pyrophosphorylase and its inhibition by antisense RNA. *Planta* 190, 247–252. doi: 10.1007/BF00196618

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

2. Mathematical modelling approaches in plant metabolomics

Lisa Fürtauer*, Jakob Weiszmann*, Wolfram Weckwerth and Thomas Nägele

(*these authors contributed equally)

In press (04/2018): Plant Metabolomics: Methods and Protocols, Methods in Molecular Biology, vol. 1778, © Springer Science+Business Media

Declaration of Authorship:

LF & JW: wrote the paper

WW: proofread the paper

TN: conceived the study, wrote the paper

Mathematical modelling approaches in plant metabolomics

Lisa Fürtauer[1,*], Jakob Weiszmann[1,2,*], Wolfram Weckwerth[1,2] and Thomas Nägele[1,2,3]

[1] Department of Ecogenomics and Systems Biology, Faculty of Life Sciences, University of Vienna, Althanstr. 14, 1090 Vienna, Austria

[2] Vienna Metabolomics Center, University of Vienna, Althanstr. 14, 1090 Vienna, Austria

[3] Department Biology I, Ludwig-Maximilians-Universität München, Planegg-Martinsried, Germany

* Authors contributed equally

Abstract

The experimental analysis of a plant metabolome typically results in a comprehensive and multidimensional data set. To interpret metabolomics data in context of biochemical regulation and environmental fluctuation, various approaches of mathematical modelling have been developed and proven useful. In this chapter, a general introduction to mathematical modelling is presented and discussed in context of plant metabolism. A particular focus is laid on the suitability of mathematical approaches to functionally integrate plant metabolomics data in a metabolic network and combine it with other biochemical or physiological parameters.

Key Words

Metabolomics; Plant biochemistry; Mathematical modelling; ODE; Jacobian matrix; Multivariate statistics; Time series analysis; Granger causality

Introduction

Due to the need for theoretical methods to combine the comprehensive output of plant systems biology studies, approaches of multivariate statistics, pattern recognition and mathematical modelling have become more and more attractive to plant biologists during the last decades. Ultimately, many modelling approaches in plant biology aim at linking genotype-specific information to an observed phenotype, e.g. by modelling plant metabolic pathways [1], growth and development [2] or by functionally integrating multilevel high-throughput results [3-5]. However, deriving predictive mathematical models of plant metabolism is challenging due to several reasons. First, plants are sessile organisms having to cope with a strongly fluctuating environment. As a consequence, plant metabolism has a complex, interlaced and nested structure with numerous redundant pathways being responsive to specific stress conditions. Hence, to unambiguously dissect the contribution of a metabolic pathway to plant stress response numerous experimental setups and mutants are necessary. Second, plant cells are highly compartmentalized and it has become evident that subcellular resolution of plant metabolism is necessary to fully understand stress-induced metabolic reprogramming [6-8]. Finally, it is the pure amount and complexity of multidimensional high-throughput data, such as plant metabolomic, proteomic or transcriptomic studies, preventing an intuitive recognition of regulatory principles.

Conclusively, mathematical methods are preliminary for a functional and causal interpretation of experimental high-throughput studies, e.g. plant metabolomics studies, by enabling the efficient analysis of multidimensional data. Hence, the following paragraphs intend to provide the reader with basic principles of selected mathematical procedures for the analysis of plant metabolomics data.

Representation of Plant Metabolism by Differential Equations

Differential equations are commonly applied in natural sciences for the quantitative description of dynamic systems being also utilized for the description of natural laws, e.g. Newton laws. A differential equation describes the relationship between functions and related derivative functions, where a derivative function describes a rate of change. There are two main types of differential equations, (i) ordinary differential equations (ODEs), and (ii) partial differential equations (PDEs). Ordinary differential equations are such equations where the searched function depends only on one variable and just *ordinary* derivatives appear. On the contrary, partial differential equations comprise *partial* derivatives, and the searched function depends on several variables. PDEs and related solution strategies are nowadays still in the focus of mathematical research and will not be discussed in this chapter.

In a biological context, a dynamic system can be defined by a set of ordinary differential equations (Eq. 1):

$$\frac{dx_i(t)}{dt} = influx - efflux = synthesis - degradation \pm transport \pm$$
$$chemical\ interconversion \pm shuttling \pm chemical\ modification\ \pm \cdots \textbf{(Eq. 1)}$$

The state x_i of the system, which might represent a metabolite concentration in the plant metabolome, depends on other state variables, i.e. can be re-written by $x_i(t) = \{x_1(t), x_2(t), x_3(t), \dots\}$, finally resulting in a *coupled* system of ODEs. To solve the system, the entire set of differential equations has to be integrated simultaneously. In context of plant metabolism, this integration step represents the summation of all metabolic fluxes and reactions affecting a metabolite concentration x_i over a time period, predominantly being dependent on enzymatic reactions [9]. The mathematical integration of a metabolite's ODE finally results in the actual metabolite concentration at time point t_0, $x_i(t_0)$.

Mathematically, ODEs are defined as follows (Eq. 2):

$$\begin{aligned} \dot{x}(t) &= f\big(t, x(t)\big) \qquad for\ all\ t \in \mathbb{R} \\ \dot{x}(t) &= f(t, x) \qquad for\ all\ t \in \mathbb{R} \end{aligned} \textbf{(Eq. 2)}$$

Here, $x(t)$ represents the searched variable within function f, and $\dot{x}(t)$ is the derivative function of $x(t)$.

In a system of (ordinary) differential equations, the state of the system, i.e. the value of $x(t)$, is set in context of its time-dependent derivative function $\dot{x}(t)$. The power of ODEs in mathematical modelling stems from the sufficiency to describe the relationship of a system's state and its derivative within a given time interval, and, based on this temporally local definition, to be able to predict the behavior of the system beyond this time interval. Solving ODEs means to find the searched function $x(t)\ t \in \mathbb{R}$ (i.e. the solution) for which the differential equation is fulfilled at any time point. In context of computational biology, analytical solutions of ODEs are hardly definable and, hence, most studies rely on numerical approximation performed by optimization algorithms. While, therefore, optimization theory and its application plays a crucial role in computational biology, the present chapter does not focus on this comprehensive topic and we recommend the interested reader to refer to more specialized literature (e.g. [10-12]).

Example: Solving a single linear ordinary differential equation

To consider general solution strategies for ODEs, we start with a simple linear example:

$$\dot{x}(t) = \frac{dx(t)}{dt} = k\ x(t) \textbf{(Eq. 3)}$$

Separation of variables, here $x(t)$ and t, yields (Eq. 4):

$$\frac{dx(t)}{x(t)} = k\ dt \textbf{ (Eq. 4)}$$

To solve this ODE, both sides of the equation are integrated (Eq. 5):

$$\int \frac{1}{x(t)}\, dx(t) = \int k\, dt$$ **(Eq. 5)**

This finally yields (Eq. 6):

$$ln|x(t)| = kt + c_1$$ **(Eq. 6)**

Applying the rule that, for a > 0, $e^{\ln a} = a$ (where e is Euler's number, e=2.718281…), we can rewrite (Eq. 6) by (Eq. 7):

$$|x(t)| = e^{kt+c_1}$$ **(Eq. 7)**

Rewriting the absolute value of $x(t)$ applying the substitution $c = e^{c_1}$, two cases can be distinguished:

$$|x(t)| = ce^{kt} \quad \left\{ \begin{matrix} I. for\ x(t) > 0 \rightarrow x(t) = ce^{kt} \\ II. for\ x(t) < 0 \rightarrow x(t) = -ce^{kt} \end{matrix} \right\}$$ **(Eq. 8)**

If $x(t)$, in context of plant metabolism, represents a (time-dependent) metabolite concentration, the second case defining a negative value for $x(t)$ is (biologically) irrelevant. In contrast, the first case implies that, if c and k are known, $x(t)$ can be calculated at any time point.

Additionally, we can solve this equation for the integration constant c by inserting the initial (known) condition at time point t_0 (Eq. 9):

$$c = x(t_0)e^{-kt_0}$$ **(Eq. 9)**

Plugging (Eq. 9) in (Eq.8, I.) yields (Eq. 10):

$$x(t) = x(t_0)e^{k(t-t_0)}$$ **(Eq. 10)**

Hence, based on a given concentration $x(t_0)$ at a time point t_0 and the parameter k, metabolite concentrations at any time point t_1, t_2, t_3, etc. can be calculated.

5

ODEs in Mathematical Modelling of Plant Metabolism

Mathematical models of plant metabolism have been developed and reviewed for decades (for some selected examples see [13-19]), and numerous research questions have benefitted from the application of ODEs to quantitatively describe, analyse and predict complex and non-intuitive system dynamics in plant biochemistry and physiology.

Pettersson and Ryde-Pettersson developed a kinetic model of the Calvin cycle [20]. This model takes into consideration that Calvin cycle activity is affected by metabolite concentrations outside the chloroplast, e.g. the concentration of external orthophosphate which was experimentally confirmed with isolated chloroplasts. Additionally, the dependency of photosynthetic rate on CO_2 and O_2 concentrations and the rate of starch production was predicted and experimentally confirmed. Besides, this model was used and extended to study properties like feedback regulation or Rubisco control in the Calvin cycle [21].

Another example for the application of ODEs to study plant metabolic pathways is a mathematical model of the plastidial 2-C-methyl-D-erythritol 4-phosphate (MEP) pathway. The model was developed to simulate and analyze diurnal pathway regulation in *Arabidopsis thaliana* [22]. Comprising several mutant lines, the authors validated their model assumptions according to substrate availability, circadian regulation and feedback regulation. Finally, the authors unraveled a feedback regulation mechanism stabilizing MEP pathway fluxes against fluctuations in substrate concentration.

While there are numerous further examples for a successful application of ODE-based mathematical models, unravelling non-intuitive and interlaced regulatory networks, the two selected studies already provide evidence for the suitability of ODE models to analyze pathway regulation and enzyme activity. Conclusively, mathematical modelling of plant metabolism potentially provides insights into complex network regulation which might be overlooked without such approaches.

Notes on an Example: How to develop a model of carbohydrate metabolism

The development of a biochemical ODE model typically starts with a graphical representation of metabolites and (enzymatic) reactions which will be analysed. To provide an example, the central carbohydrate metabolism of leaf mesophyll cells is graphically summarized in Figure1.

Frequently, model construction comprises several critical steps of structural and/or kinetic simplification and, hence, has to be considered very carefully. In the provided example, cellular carbohydrate metabolism has been simplified by ignoring aspects like subcellular compartmentation, which, under *in vivo* conditions, would split sucrose cleavage, catalysed by different invertase isoforms, into cytosolic and vacuolar reactions. Consequently, model simulations will not reveal information about subcellular partitioning of biochemistry which has to be considered when interpreting model output. A graphical model can be developed by hand with pen and paper, in a graphical program or in cellular simulator software, e.g. CellDesigner™ [23], which allows a direct combination of graphical representation, ODEs and simulations.

In the given example, the translation of the graphical into an ODE model results in the following equation system (Table).

Table I. Summary of ODEs to describe the graphical model shown in Figure 1. Reactions of SPS and SPP have been summarized by the SPS reaction and S6P is not explicit part of the ODE structure

$\frac{d\,[F6P]}{dt} = Import + FrcK - PGI - SPS$	$\frac{d\,[Suc]}{dt} = SPS - I\boldsymbol{nv} - Export$
$\frac{d\,[G6P]}{dt} = PGI + GlcK - PGM$	$\frac{d\,[Glc]}{dt} = Inv - GlcK$
$\frac{d\,[G1P]}{dt} = PGM - UGPase$	$\frac{d\,[Frc]}{dt} = Inv - FrcK$
$\frac{\mathrm{d}\,[\mathrm{UDPG}]}{\mathrm{dt}} = \mathrm{UGPase} - \mathrm{SPS}$	$\frac{\mathrm{d}\,[\mathrm{Sink}]}{\mathrm{dt}} = Export$

In the next step, all reaction rates have to be defined by kinetic equations. For example, the ODE of G6P depends on the reactions *PGI* and *GlcK* (influx) and *PGM* (efflux). To quantify the in- and efflux,

kinetic laws are defined to resemble *in vivo* kinetics as closely as possible with the available enzyme kinetic parameters, e.g. assuming mass action kinetics (Eq. 11,12) or Michaelis-Menten kinetics which might also comprise inhibition terms (Eq. 13, *here:* competitive inhibition of GlcK by G6P).

$$PGI = [F6P] * k_{PGI} \textbf{ (Eq. 11)}$$

$$PGM = [G6P] * k_{PGM} \textbf{ (Eq. 12)}$$

$$\mathrm{GlcK} = \frac{v_{max} * [Glc]}{[Glc] + K_M * (1 + \frac{[G6P]}{K_{i,G6P}})} \textbf{ (Eq. 13)}$$

Finally, the system of ODEs has to be solved, i.e. integrated, revealing the time course of the metabolite concentrations. As explained before, analytical solutions of such ODE systems fulfilling all experimental preliminaries, i.e. kinetic parameters and metabolite concentrations, can hardly be assigned. Instead, numerical solutions are determined applying local and/or global optimization algorithms, searching for locally or globally optimal solutions. Optimal solutions are searched within a pre-defined range of kinetic parameters being derived from experiments and/or database entries. A common optimality criterion is the so-called cost function which, for example, quantifies the squared error between simulated and experimentally determined metabolite concentrations. Optimization algorithms and solvers are implemented in modelling software like CellDesigner™ [23] or COPASI [24], or are available for other numerical software environments [25].

A limiting step of kinetic (ODE) modelling is the availability of sufficient kinetic information being necessary to develop and solve ODEs within biochemically and physiologically meaningful boundaries. Although enzyme databases, e.g. BRENDA (http://www.brenda-enzymes.org/) [26], contribute comprehensive kinetic information across numerous organisms, pathways and enzyme isoforms, in many studies additional experimental assays are necessary to reveal diurnal or stress-induced dynamics in enzyme activities (v_{max}), protein abundance, substrate affinity (K_M) and inhibition/activation of enzymes by metabolic compounds (K_i). Particularly, large-scale modelling approaches comprising hundreds, or even thousands, of metabolites and proteins are inherently

8

limited by experimental data. To (partly) overcome these limitations, alternative approaches have been developed, frequently comprising steps of system linearization, or parameter normalization [27-30]. A mathematical term playing a central role in these approaches is the so-called Jacobian matrix being introduced in the following chapter.

Steady states and the biochemical Jacobian matrix

Due to the rapid development of genome sequencing technologies, numerous whole genome sequences have become available for a diverse set of organisms, comprising more than 180 plant reference sequences [31]. Further, protocols have been developed using genome sequence information to deduct metabolic networks attempting to resemble *in vivo* reaction networks [32,33]. To combine these genome-scale networks with results from transcriptomics, proteomics and metabolomics studies, mathematical methods are needed being able to functionally, i.e. causally, connect metabolite concentrations and/or enzyme activities with biochemical network information. ODE-based kinetic modelling of such large-scale networks is currently precluded by the lack of kinetic information being necessary to numerically integrate ODEs within a physiologically feasible range. Further, finding a global optimum for kinetic parameters and metabolite concentrations of large metabolic systems is challenging due to the nonlinear dynamics being particularly characteristic of plant-environment interactions. Hence, instead of solving the complete nonlinear system, deriving solutions at a steady state of the system is a way to deal with these limitations.

The steady state assumption is expressed by the term $\frac{d[\boldsymbol{M}]}{dt} = \mathbf{0}$ stating that no change in metabolite concentrations $[\mathbf{M}]$ can be observed over a change of time. Geometrically spoken, $\frac{d[\boldsymbol{M}]}{dt}$ represent the slopes of tangents of all functions $\boldsymbol{M}(t)$, i.e. of the time course of metabolite concentrations. At steady state conditions, slopes of these tangents are zero. Hence, steady state conditions imply that reaction rates of the system are constant, and consuming and synthesizing reactions of each metabolite are in equilibrium. While these steady state conditions are hardly met by any plant biological study focusing

9

on metabolic reprogramming, e.g. due to stress response or diurnal rhythms, two additional requirements can be defined which still justify the biochemical interpretation:

(i) Instead of a strict steady state, a *quasi* steady state might be considered by changing the preliminaries from ($\frac{d[\boldsymbol{M}]}{dt} = \mathbf{0}$; **no** change) to ($\frac{d[\boldsymbol{M}]}{dt} \approx \mathbf{0}$; **no significant** change).

(ii) The interpretation of results is only valid within a very small (or infinitesimal small) time period.

Mathematically, $\frac{d[\boldsymbol{M}]}{dt}$ can be described via the stoichiometric matrix $\boldsymbol{S}$ (mxn; m number of metabolites, n number of reactions), where the entries represent the participation of a metabolite in a reaction, multiplied with the vector $\boldsymbol{v}$ containing the reaction rates. Resulting from this, a system of linear algebraic equations can be defined (Eq. 14).

$$\frac{d\boldsymbol{M}}{dt} = \boldsymbol{S} * \boldsymbol{v} = \mathbf{f}(\mathbf{M}, \mathbf{p}, \mathrm{t}) = 0 \quad \textbf{(Eq. 14)}$$

Here, metabolic functions, $\boldsymbol{f}(\boldsymbol{M}, \boldsymbol{p}, t)$, summarize the right side of the ODEs depending on metabolite concentrations $\mathbf{M}$, parameters $\boldsymbol{p}$ (e.g. kinetic parameters) and time t.

Explicitly applying this equation to the previously defined model (see Figure 1), the reaction network can be written as a steady state equation matrix system (Eq. 15).

$$\frac{d}{dt}\begin{bmatrix} F6P \\ G6P \\ G1P \\ UDPG \\ Suc \\ Glc \\ Frc \\ Sink \end{bmatrix} = \begin{bmatrix} 1 & 1 & -1 & -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & -1 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & -1 & -1 \\ 0 & 0 & 0 & 0 & -1 & 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} * \begin{bmatrix} Import \\ FrcK \\ PGI \\ SPS \\ GlcK \\ PGM \\ UGPase \\ INV \\ Export \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \quad \textbf{(Eq. 15)}$$

To mathematically analyse system behaviour at a *quasi* steady state condition, metabolic functions can be linearized at time point t_0 of the steady state. In context of metabolic networks, this linearization leads to tangents of metabolic functions at $(t_0, \boldsymbol{f}(\boldsymbol{M}, \boldsymbol{p}, t_0))$.

To characterize (regulatory) effects of metabolites on metabolic functions, partial derivatives are built quantifying changes in metabolic functions with respect to changes in metabolite concentrations. Summarizing all partial derivatives in a matrix results in the so-called *Jacobian matrix* ***J*** (Eq. 16).

$$J = \begin{pmatrix} \frac{\delta f_1}{\delta M_1} & \frac{\delta f_1}{\delta M_2} & \cdots & \frac{\delta f_1}{\delta M_m} \\ \frac{\delta f_2}{\delta M_1} & \frac{\delta f_2}{\delta M_2} & \cdots & \frac{\delta f_2}{\delta M_m} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\delta f_m}{\delta M_1} & \frac{\delta f_m}{\delta M_2} & \cdots & \frac{\delta f_m}{\delta M_m} \end{pmatrix} \textbf{(Eq. 16)}$$

As described before, metabolic functions reflect all internal and external factors affecting the dynamics of a metabolite. Therefore, the Jacobian entries report on the response of metabolic functions to (slight) perturbations of the metabolic (steady) state. Hence, high absolute values of Jacobian entries indicate a strong effect of concentration changes on metabolic functions.

In the following paragraph, examples are depicted to visualize the effect of different metabolic constitutions on Jacobian entries. The previously introduced network of central carbohydrate metabolism will serve as an example, in which all kinetic parameters and metabolite concentrations are arbitrarily chosen.

The first scenario describes a metabolic state of the system in which every reaction is unidirectional and no activating or inhibitory influences are included in the network (Figure 2). The corresponding Jacobian matrix is nonzero in all entries where a metabolite represents either a reaction substrate or a reaction product.

For example, the metabolic function of F6P (f_{F6P}; derivatives δf_{F6P} first row in table of Figure 2) depends on changes in the concentration of Fru being a substrate for the formation of F6P by the activity of

FrcK. Further, f_{F6P} depends on changes in the concentration of UDPG because F6P and UDPG are substrates for the SPS reaction.

In the given example, highest absolute values of Jacobian entries describe the dependency of f_{F6P} and f_{Fru} on fructose dynamics. This indicates a strong effect of (slight) dynamics in fructose concentration on the rate of fructose phosphorylation, affecting the metabolic function f_{F6P} to increase (F6P: reaction product) and affecting f_{Fru} to decrease (Fru: reaction substrate). Another example of such a strong product-substrate dependency can be observed for the reaction catalysed by UGPase where [G1P] dynamics decrease f_{G1P} and increase f_{UDPG}. Hence, both reactions have in common that reaction substrates are affecting reaction rates stronger than observed for other reactions, for example compared to the effect of Glc on GlcK reaction or G6P on the PGM reaction. This is due to different ratios of substrate concentrations to kinetic parameters and resulting fluxes: while the ratio of the FrcK flux to Fru concentration is ~300, the ratio of UGPase flux to G1P concentration is ~37 and PGM flux related to G6P concentration results in a ratio of (only) ~3. Further, these three reactions differ in the kinetic law which was assumed for flux simulation. While FrcK was assumed to follow Michaelis-Menten kinetic, UGPase and PGM driven reactions were described by mass-action kinetics. While Michaelis-Menten kinetics are characterized by a saturation level above which higher substrate concentrations do not affect the reaction rate anymore, mass-action law (for mono-substrate reactions) states a linear dependence of reaction rate on substrate concentration. The consequence for dynamics of Jacobian entries is exemplarily shown for both reaction kinetics (Figure 3). While the mass action-related Jacobian entry $\delta(f_{UDPG})/\delta[G1P]$ is constant over G1P concentrations (Fig. 3 A, B), the Michaelis-Menten related entry $\delta(f_{F6P})/\delta[Fru]$ is asymptotically decreasing over increasing Fru concentrations (Fig. 3 C, D).

While the above discussed metabolic network (see Figure 2) does not contain any regulatory effectors, a modified network structure contains feedback inhibition by reaction products of invertase, fructo- and glucokinase reactions (Figure 4).

Inhibitions are directly reflected in the Jacobian entries. Invertase inhibition by Fru and Glc results in positive entries of $\delta(f_{Suc})/\delta[Fru]$ and $\delta(f_{Suc})/\delta[Glc]$ due to their potential reduction of sucrose consumption. Analogously, inhibition of FrcK by F6P and GlcK by G6P are reflected by positive entries $\delta(f_{Fru})/\delta[F6P]$ and $\delta(f_{Glc})/\delta[G6P]$. In contrast, product inhibition of invertase by Fru and Glc results in negative entries for $\delta(f_{Fru})/\delta[Glc]$ and $\delta(f_{Glc})/\delta[Fru]$ because both inhibitors (negatively) affect each other's production by sucrose cleavage.

In summary, a high value of Jacobian entries of a metabolic system does not necessarily reflect a high flux through an enzymatic reaction, yet rather delivers information about the influence of changes in a metabolite concentration on a metabolic function. Considering the relation between substrates and products, this influence is especially large if the enzymatic reaction is far away from saturation, so a small change in substrate concentration will have a large effect on the product's metabolic function. This is of particular importance for biochemical systems, which usually do not follow linear kinetics, and, hence, a biochemical Jacobian matrix allows insight in the system's capabilities to cope with intracellular or environmental fluctuations.

The analysis of metabolic networks at (*quasi*) steady state conditions significantly simplifies mathematical modelling and optimization procedures. Further, only a fraction of experimental data is needed when compared to time dynamic and non-linear systems. However, even under these simplified conditions many kinetic parameters have to be determined, and particularly for large-scale networks this still limits the application of mathematical modelling to metabolomics data. To overcome such limitations, approaches have been developed reducing the amount of kinetic information by normalisation procedures (see e.g. [34]) or by inference of biochemical Jacobian matrices from experimental covariance data (see e.g. [35-37,9]). Such approaches may then be used for the comparison of two observed steady states under different conditions to get insight in the regulatory state of a biochemical system. Finally, it still has to be investigated how far *quasi* steady state-derived information about Jacobian matrices might be extrapolated to gain (still) valid information about

dynamic properties of the metabolic system. These properties might be experimentally analysed within metabolomic time course experiments being dealt with in the next section.

Time series analysis of multidimensional metabolomics data

The previous paragraphs of this chapter introduced strategies to deal with the challenge of functionally connecting experimental plant metabolomics data with biochemical network information. While solving ODE systems using enzyme kinetic parameters and absolute metabolite concentrations is desirable as it yields a causal model to simulate and predict reprogramming of plant metabolism due to environmental cues, particularly for large-scale models it is highly challenging, if possible at all, to derive a biochemically feasible solution due to the lack of knowledge about kinetic parameters. On the contrary, approximation of large-scale systems by linearization methods, e.g. resulting in the Jacobian matrix ***J***, is limited by its steady state assumption hampering the extrapolation and interpretation of solutions for ***J*** in a dynamic and nonlinear context. Hence, to complement the presented toolbox of mathematical procedures, methods are needed coping with the identification and quantification of dynamic, and often nonlinear, signatures in metabolomics data sets. Methods of mathematical time series analysis are able to decipher such signatures in reiterated measurements over a time period which might comprise diurnal rhythms, stress-induced dynamics or developmental reprogramming of metabolism. Mathematical procedures of time series analysis are diverse and numerous textbooks are available providing detailed information about algorithms and mathematical theory (see e.g. [38] or [39]). In addition, statistical software tools are providing a comprehensive compilation of ready-to-use algorithms with a detailed documentation. For example, the freely available language and environment for statistical computing and graphics, R (https://www.r-project.org/), provides a collection of algorithms, procedures and documentation summarized by the so-called *CRAN Task View: Time Series Analysis* (https://CRAN.R-project.org/view=TimeSeries; Date: 06/2017). While many of these algorithms might have been developed in context other than metabolomics, e.g. finance and econometrics, they prove to be very useful for the research field of metabolomics (see the example given below, *Granger causality*). Moreover, tutorial reviews and web-based tools are available

particularly focusing on the analysis of dynamic metabolomics data (see e.g. [40-42]). Finally, commercial software, being frequently used for mathematical modelling, provides a comprehensive library of methodologies in the fields of statistical time series analysis and/or pattern recognition, e.g. the numerical programming environment MATLAB® (https://www.mathworks.com). In addition to commercially available toolboxes, there are numerous freely available graphical user interfaces, based on the MATLAB® software, supporting the application of multivariate statistics, mathematical modelling and time series analysis (see e.g. [43-47]). Previously, we have applied the MATLAB®-based multivariate statistics toolbox COVAIN [43] to analyse stress-induced dynamics between primary and secondary metabolism in *Arabidopsis thaliana* [37]. In this study, Granger causations were calculated describing a time series correlation between experimental data collected on a GC-MS and LC-MS platform. Selecting the most significant correlations between both data sets revealed an abiotic stress-induced initiation of cyanidine glucoside biosynthesis being accompanied by a deregulation of the shikimic acid pathway.

Granger causality analysis of metabolomics data: Procedure and Limitations

Previously, we have applied the MATLAB -based multivariate statistics toolbox COVAIN [43] to analyze stress-induced dynamics between primary and secondary metabolism in Arabidopsis thaliana [37]. In this study, Granger causations were calculated describing a time series correlation between experimental data collected on a GC-MS and LC-MS platform. Selecting the most significant correlations between both data sets revealed an abiotic stress-induced initiation of cyanidine glucoside biosynthesis being accompanied by a deregulation of the shikimic acid pathway.

Granger causality analysis was developed several decades ago to investigate causal relations by econometric models [48]. It considers the time series of two variables, here X and Y, which, in a plant metabolomics context, might represent concentrations of two different metabolites (Eqs. 17 and 18)

$$X(t) = \sum_{i=1}^{n} C_{X,i} X(t-i) + \sum_{i=1}^{n} C_{XY,i} Y(t-i) + R_X(t)$$

(Eq. 17)

$$Y(t) = \sum_{i=1}^{n} C_{YX,i} X(t-i) + \sum_{i=1}^{n} C_{Y,i} Y(t-i) + R_Y(t)$$

(Eq. 18)

$X(t)$ and $Y(t)$ are metabolite concentrations at time point t. In Eq. 17, $C_{X,i}$ and $C_{XY,i}$ represent the regression coefficients between $X(t)$ and $X(t-i)$ and between $X(t)$ and $Y(t-i)$, respectively. In Eq. 18, $C_{YX,i}$ is the regression coefficient between $Y(t)$ and $X(t-i)$ while $C_{Y,i}$represents the regression coefficient between $Y(t)$ and $Y(t-i)$. $R_X(t)$and $R_Y(t)$are residual errors of the regression. The variable n describes the maximum time lag between the variables X and Y, i.e. metabolite concentrations. Hence, significant cross-coefficients $C_{XY,i}$and $C_{YX,i}$indicate a putatively causal relationship between the two variables X and Y within the recorded time series and across a time lag n. In general, the regression coefficients might be positive or negative indicating how X influences Y (and/or *vice versa*). In context of metabolite levels derived from a metabolomics experiment, positive cross-coefficients indicate a parallel trend of the metabolites X and Y, shifted by the lime lag n. This might either be represented by a decrease or an increase of both metabolite concentrations. If the cross-coefficients are negative, this indicates an opposite trend of X and Y, i.e. a decrease of the concentration of X and an increase of the concentration of Y, or *vice versa*. The maximum time lag n corresponds to the maximum time period which is considered for building the regression coefficients. If adequately chosen, n resembles the minimal time frame which it takes to yield an observable output of the potentially causal interaction of X and Y. In a biochemical context this time lag might be related to signalling processes or enzymatic reactions rates interconverting X into Y. Although only discussed for two variables, this procedure is also applicable to multiple variables (e.g. m variables) using a model of m^{th}order [48].

Limitations of applying Granger causality, and also other strategies of time series analysis, to metabolomics time course data are mainly attributable to

(i) comparability of experimental data from different platforms,

(ii) interpretability of putatively causal relationships in a biochemical context.

To be able to compare data from different platforms, e.g. relative metabolite levels derived from GC-MS and LC-MS measurements, it is crucial to normalise and scale the data before applying the time series analysis. This is, of course, not only necessary for Granger causality analysis but for various statistics and modelling approaches. Commonly, chromatographic peak areas are normalised to internal or external standards, sample protein amount, sample fresh weight or sample dry weight. If calibration solutions of substances are available for data normalisation, this enables the absolute quantification of metabolite concentrations which allows a direct comparison of metabolomics data among different analytical platforms. If calibration solutions are not available, for example in (untargeted) metabolite profiling approaches, it is crucial to scale normalised data before performing statistical analysis. A frequently used scaling method is *autoscaling* (also known as *zero mean-unit variance* or *z-scores*) which subtracts the mean of each variable, i.e. metabolite, from each single value and divides by its standard deviation. Finally, metabolites have the same standard deviation of one making them equally important. While this scaling method is advantageous for making data comparable between different experimental platforms, measurement errors are strongly weighted, and, hence, might enhance data scattering due to technical limitations. Alternatively, experimental results might be normalised, centered, scaled or transformed in various ways for meaningful and convincing data analysis. However, a careful choice of these mathematical operations is crucial to reduce the chance of misinterpretation. For this, excellent reviews and methodological summaries are available supporting the experimentalist to choose an adequate method (see e.g. [49]).

Conclusions

While many statistical approaches yield information on correlations between metabolites, this does not necessarily resemble realistic, i.e. biochemically observable, interactions. Hence, it is crucial to combine the statistical output with biochemical information in order to enhance the reliability and predictive power of the output. In context of metabolomic studies, this biochemical information can typically be derived from a genome-scale metabolic network which connects cellular metabolites by biochemical reactions which are predicted by the genome sequence of an organism [9,33]. Numerous genome-scale metabolic network models are available, in particular for plant model organisms, e.g. Arabidopsis [50], maize [51] and rice [52]. With the list of enzymatically catalysed reactions in each of those models it is possible to select the biochemically interpretable correlations between two metabolites derived from statistical time series analysis. In a previously published approach we automatized this connection of biochemical network information and time-continuous metabolomics data analysis within a MATLAB®-based graphical user interface, FEMTO (Functional Evaluation of Metabolic Time series Observations) [53]. FEMTO unravels time points of interests, where perturbed dependencies in metabolic functions indicate a de-regulated biochemical interaction. A regularly updated version of FEMTO, including a user manual and a test data set, is freely available at http://mosys.univie.ac.at/resources/software/.

References

1. Morgan JA, Rhodes D (2002) Mathematical modeling of plant metabolic pathways. Metab Eng 4 (1):80-89. doi:10.1006/mben.2001.0211

2. Prusinkiewicz P (2004) Modeling plant growth and development. Current Opinion in Plant Biology 7 (1):79-83. doi:http://dx.doi.org/10.1016/j.pbi.2003.11.007

3. Fiehn O (2002) Metabolomics--the link between genotypes and phenotypes. Plant Mol Biol 48 (1-2):155-171

4. Weckwerth W (2011) Unpredictability of metabolism-the key role of metabolomics science in combination with next-generation genome sequencing. Anal Bioanal Chem 400 (7):1967-1978. doi:10.1007/s00216-011-4948-9

5. Liberman LM, Sozzani R, Benfey PN (2012) Integrative systems biology: an attempt to describe a simple weed. Curr Opin Plant Biol 15 (2):162-167. doi:10.1016/j.pbi.2012.01.004

6. Hoermiller, II, Naegele T, Augustin H, Stutz S, Weckwerth W, Heyer AG (2017) Subcellular reprogramming of metabolism during cold acclimation in Arabidopsis thaliana. Plant Cell Environ 40 (5):602-610. doi:10.1111/pce.12836

7. Hurry V (2017) Metabolic reprogramming in response to cold stress is like real estate, it's all about location. Plant, Cell & Environment 40 (5):599-601. doi:10.1111/pce.12923

8. Fürtauer L, Weckwerth W, Nägele T (2016) A Benchtop Fractionation Procedure for Subcellular Analysis of the Plant Metabolome. Front Plant Sci 7 (1912):1912. doi:10.3389/fpls.2016.01912

9. Nägele T (2014) Linking metabolomics data to underlying metabolic regulation. Front Mol Biosci 1:22. doi:10.3389/fmolb.2014.00022

10. Wang Y, Zhang X-S, Chen L (2010) Optimization meets systems biology. BMC Systems Biology 4 (2):S1. doi:10.1186/1752-0509-4-s2-s1

11. Banga JR (2008) Optimization in computational systems biology. BMC Systems Biology 2:47. doi:10.1186/1752-0509-2-47

12. Reali F, Priami C, Marchetti L (2017) Optimization Algorithms for Computational Systems Biology. Frontiers in Applied Mathematics and Statistics 3 (6). doi:10.3389/fams.2017.00006

13. Loomis RS, Rabbinge R, Ng E (1979) Explanatory models in crop physiology. Annual Review of Plant Physiology 30 (1):339-367

14. Rios-Estepa R, Lange BM (2007) Experimental and mathematical approaches to modeling plant metabolic networks. Phytochemistry 68 (16-18):2351-2374. doi:10.1016/j.phytochem.2007.04.021

15. Klipp E, Liebermeister W (2006) Mathematical modeling of intracellular signaling pathways. BMC Neurosci 7 Suppl 1:S10. doi:10.1186/1471-2202-7-S1-S10

16. Chew YH, Smith RW, Jones HJ, Seaton DD, Grima R, Halliday KJ (2014) Mathematical models light up plant signaling. Plant Cell 26 (1):5-20. doi:10.1105/tpc.113.120006

17. Rohwer JM (2012) Kinetic modelling of plant metabolic pathways. Journal of experimental botany 63 (6):2275-2292. doi:10.1093/jxb/ers080

18. Gombert AK, Nielsen J (2000) Mathematical modelling of metabolism. Current Opinion in Biotechnology 11 (2):180-186. doi:http://dx.doi.org/10.1016/S0958-1669(00)00079-3

19

19. Giersch C (2000) Mathematical modelling of metabolism. Curr Opin Plant Biol 3 (3):249-253

20. Pettersson G, Ryde-Pettersson U (1988) A mathematical model of the Calvin photosynthesis cycle. Eur J Biochem 175 (3):661-672

21. Pettersson G (1997) Control properties of the Calvin photosynthesis cycle at physiological carbon dioxide concentrations. Biochimica et Biophysica Acta (BBA) - Bioenergetics 1322 (2):173-182. doi:http://dx.doi.org/10.1016/S0005-2728(97)00080-7

22. Pokhilko A, Bou-Torrent J, Pulido P, Rodríguez-Concepción M, Ebenhöh O (2015) Mathematical modelling of the diurnal regulation of the MEP pathway in Arabidopsis. New Phytologist 206 (3):1075-1085. doi:10.1111/nph.13258

23. Funahashi A, Morohashi M, Kitano H, Tanimura N (2003) CellDesigner: a process diagram editor for gene-regulatory and biochemical networks. Biosilico 1 (5):159-162. doi:10.1016/s1478-5382(03)02370-9

24. Hoops S, Sahle S, Gauges R, Lee C, Pahle J, Simus N, Singhal M, Xu L, Mendes P, Kummer U (2006) COPASI - a COmplex PAthway SImulator. Bioinformatics 22 (24):3067-3074. doi:10.1093/bioinformatics/btl485

25. Schmidt H (2007) SBaddon: high performance simulation for the Systems Biology Toolbox for MATLAB. Bioinformatics 23 (5):646-647. doi:10.1093/bioinformatics/btl668

26. Schomburg I, Hofmann O, Baensch C, Chang A, Schomburg D (2000) Enzyme data and metabolic information: BRENDA, a resource for research in biology, biochemistry, and medicine. Gene Function & Disease 1 (3-4):109-118. doi:10.1002/1438-826X(200010)1:3/4<109::AID-GNFD109>3.0.CO;2-O

27. Steuer R., Gross T., Selbig J., Blasius B. (2006) Structural kinetic modeling of metabolic networks. Proceedings of the National Academy of Sciences of the United States of America 103 (32):11868-11873. doi:10.1073/pnas.0600013103

28. Reznik E., Segre D. (2010) On the stability of metabolic cycles. Journal of theoretical biology 266 (4):536-549. doi:10.1016/j.jtbi.2010.07.023

29. Henkel S., Nägele T., Hörmiller I., Sauter T., Sawodny O., Ederer M., and Heyer A.G. (2011) A systems biology approach to analyse leaf carbohydrate metabolism in Arabidopsis thaliana.

30. Fürtauer L, Nägele T (2016) Approximating the stabilization of cellular metabolism by compartmentalization. Theory in Biosciences. doi:10.1007/s12064-016-0225-y

31. Jiao W-B, Schneeberger K (2017) The impact of third generation genomic technologies on plant genome assembly. Current Opinion in Plant Biology 36:64-70. doi:https://doi.org/10.1016/j.pbi.2017.02.002

32. Vivek-Ananth RP, Samal A (2016) Advances in the integration of transcriptional regulatory information into genome-scale metabolic models. Biosystems 147:1-10. doi:http://dx.doi.org/10.1016/j.biosystems.2016.06.001

33. Thiele I, Palsson BO (2010) A protocol for generating a high-quality genome-scale metabolic reconstruction. Nat Protoc 5 (1):93-121. doi:10.1038/nprot.2009.203

34. Steuer R, Gross T, Selbig J, Blasius B (2006) Structural kinetic modeling of metabolic networks. Proc Natl Acad Sci U S A 103 (32):11868-11873. doi:10.1073/pnas.0600013103

35. Nägele T, Mair A, Sun X, Fragner L, Teige M, Weckwerth W (2014) Solving the differential biochemical Jacobian from metabolomics covariance data. PLoS One 9 (4):e92299. doi:10.1371/journal.pone.0092299

36. Kügler P, Yang W (2014) Identification of alterations in the Jacobian of biochemical reaction networks from steady state covariance data at two conditions. J Math Biol 68 (7):1757-1783. doi:10.1007/s00285-013-0685-3

37. Doerfler H, Lyon D, Nägele T, Sun X, Fragner L, Hadacek F, Egelhofer V, Weckwerth W (2013) Granger causality in integrated GC-MS and LC-MS metabolomics data reveals the interface of primary and secondary metabolism. Metabolomics 9 (3):564-574. doi:10.1007/s11306-012-0470-0

38. Schelter B (2006) Handbook of time series analysis recent theoretical developments and applications. WILEY-VCH, Weinheim

39. Derryberry DR (2014) Basic data analysis for time series with R. Wiley, Hoboken, N.J.

40. Smilde AK, Westerhuis JA, Hoefsloot HCJ, Bijlsma S, Rubingh CM, Vis DJ, Jellema RH, Pijl H, Roelfsema F, van der Greef J (2010) Dynamic metabolomic data analysis: a tutorial review. Metabolomics 6 (1):3-17. doi:10.1007/s11306-009-0191-1

41. Xia J, Sinelnikov IV, Wishart DS (2011) MetATT: a web-based metabolomics tool for analyzing time-series and two-factor datasets. Bioinformatics 27 (17):2455-2456. doi:10.1093/bioinformatics/btr392

42. Xia J, Wishart DS (2002) Using MetaboAnalyst 3.0 for Comprehensive Metabolomics Data Analysis. In: Current Protocols in Bioinformatics. John Wiley & Sons, Inc. doi:10.1002/cpbi.11

43. Sun XL, Weckwerth W (2012) COVAIN: a toolbox for uni- and multivariate statistics, time-series and correlation network analysis and inverse estimation of the differential Jacobian from metabolomics covariance data. Metabolomics 8 (1):S81-S93. doi:Doi 10.1007/S11306-012-0399-3

44. Girbig D, Selbig J, Grimbs S (2012) A MATLAB toolbox for structural kinetic modeling. Bioinformatics 28 (19):2546-2547. doi:10.1093/bioinformatics/bts473

45. Schmidt H, Jirstrand M (2006) Systems Biology Toolbox for MATLAB: a computational platform for research in systems biology. Bioinformatics 22 (4):514–515. doi:10.1093/bioinformatics/bti799

46. Aurich MK, Fleming RMT, Thiele I (2016) MetaboTools: A Comprehensive Toolbox for Analysis of Genome-Scale Metabolic Models. Frontiers in physiology 7 (327). doi:10.3389/fphys.2016.00327

47. Fitzpatrick MA, McGrath CM, Young SP (2014) Pathomx: an interactive workflow-based tool for the analysis of metabolomic data. BMC Bioinformatics 15 (1):396. doi:10.1186/s12859-014-0396-9

48. Granger CWJ (1969) Investigating causal relations by econometric models and cross-spectral methods. Econometrica 37:414–426

49. van den Berg RA, Hoefsloot HC, Westerhuis JA, Smilde AK, van der Werf MJ (2006) Centering, scaling, and transformations: improving the biological information content of metabolomics data. BMC Genomics 7 (1):142. doi:10.1186/1471-2164-7-142

50. Mintz-Oron S, Meir S, Malitsky S, Ruppin E, Aharoni A, Shlomi T (2012) Reconstruction of Arabidopsis metabolic network models accounting for subcellular compartmentalization and tissue-specificity. Proc Natl Acad Sci U S A 109 (1):339-344. doi:10.1073/pnas.1100358109

51. Bogart E, Myers CR (2016) Multiscale Metabolic Modeling of C4 Plants: Connecting Nonlinear Genome-Scale Models to Leaf-Scale Metabolism in Developing Maize Leaves. PLOS ONE 11 (3):e0151722. doi:10.1371/journal.pone.0151722

52. Shaw R, Kundu S (2015) Flux balance analysis of genome-scale metabolic model of rice (Oryza sativa): Aiming to increase biomass. Journal of biosciences 40 (4):819-828. doi:10.1007/s12038-015-9563-z

53. Nägele T, Fürtauer L, Nagler M, Weiszmann J, Weckwerth W (2016) A Strategy for Functional Interpretation of Metabolomic Time Series Data in Context of Metabolic Network Information. Front Mol Biosci 3:6. doi:10.3389/fmolb.2016.00006

Figures

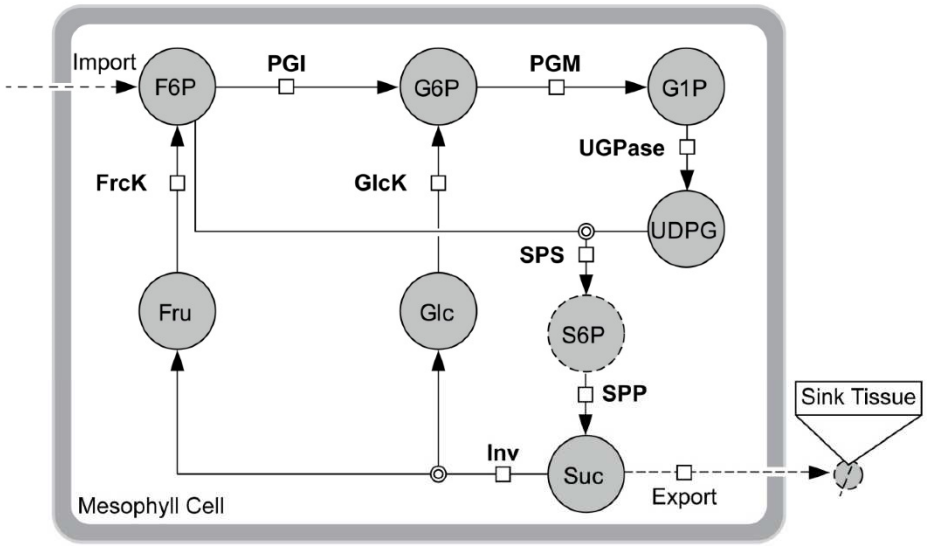
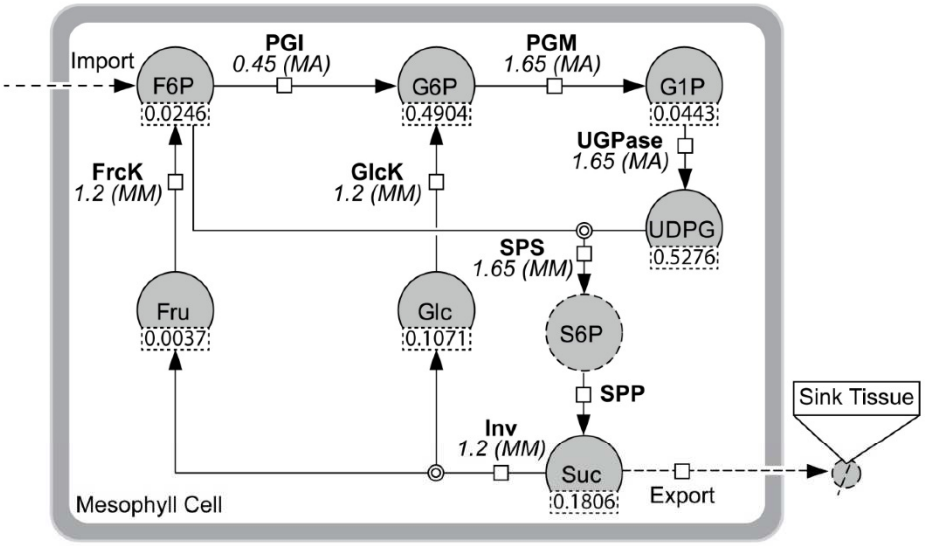


Figure 1. A schematic overview of the central carbohydrate metabolism of plant leaf cells. Metabolites are described in circles, enzymatic reactions are indicated by arrows. F6P: fructose-6-phosphate; G6P: glucose-6-phosphate; G1P: glucose-1-phosphate; UDPG: uridinediphosphate glucose; S6P: sucrose-6-phosphate; Suc; sucrose; Glc: glucose; Fru: fructose; PGI: phosphoglucoisomerase; PGM: phosphoglucomutase; SPS: sucrose phosphate synthase; SPP: sucrose phosphate phosphatase; Inv: Invertase; FrcK: fructokinase; GlcK: glucokinase.



		δ[M]						
		F6P	G6P	G1P	UDPG	Suc	Fru	Glc
δ(*f*)	F6P	-44.87	0	0	-1.59	0	277.33	0
	G6P	18.46	-4.04	0	0	0	0	9.05
	G1P	0	4.04	-37.33	0	0	0	0
	UDPG	-26.41	0	37.33	-1.59	0	0	0
	Suc	26.41	0	0	1.59	-8.08	0	0
	Fru	0	0	0	0	5.62	-277.33	0
	Glc	0	0	0	0	5.62	0	-9.05

Figure 2. A steady state model of the central carbohydrate metabolism. The graphical model on the left side contains information about an arbitrarily chosen steady state comprising metabolite concentrations (below circles, in mM) and fluxes (below reaction names, in mM h-1). Reactions follow either mass-action (MA) or Michaelis-Menten (MM) kinetics. The table summarizes entries of the Jacobian matrix of this steady state. Entries correspond to the definition in Eq. 16. For example, the entry in first column and first row represents the value of (δfF6P/ δ[F6P]). Abbreviations refer to Figure 1.

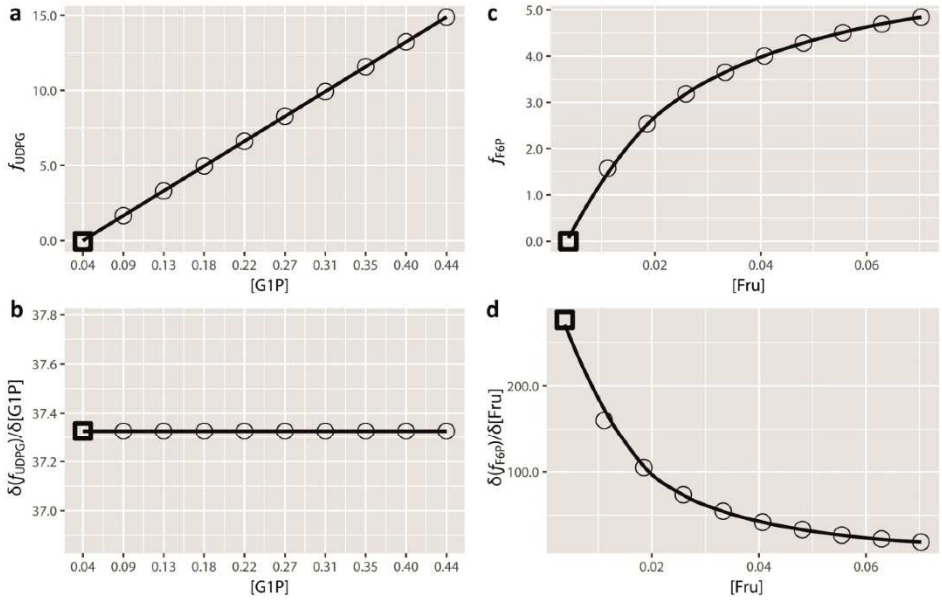
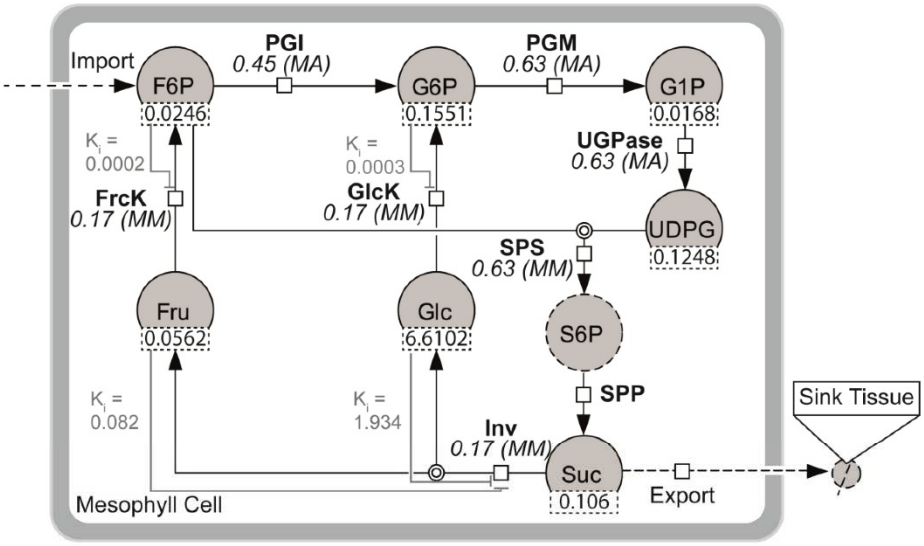


Figure 3. Metabolic functions in relation to Jacobian entries for mass-action kinetics (A,B) and Michaelis-Menten kinetics (C,D). The metabolic function of UDPG (A) and its Jacobian entry (B) are related to the substrate concentration for its biosynthesis, [G1P]. The metabolic function of F6P (C) and its Jacobian entry (D) are related to the substrate concentration for its biosynthesis, [Fru]. Squares indicate the concentration of G1P and Fru at the simulated steady state (see Figure 2).

25



		δ[M]						
		F6P	G6P	G1P	UDPG	Suc	Fru	Glc
δ(*f*)	F6P	-35.22	0	0	-4.09	0	2.99	0
	G6P	18.46	-5.11	0	0	0	0	0.03
	G1P	0	4.04	-37.33	0	0	0	0
	UDPG	-10.01	0	37.33	-4.09	0	0	0
	Suc	10.01	0	0	4.09	-3.31	1.12	0.02
	Fru	6.75	0	0	0	0.86	-4.11	-0.02
	Glc	0	1.07	0	0	0.86	-1.12	-0.05

Figure 4. A steady state model of the central carbohydrate metabolism comprising feedback inhibition of reactions Inv, FrcK and Glck by their products. The graphical model on the left side contains information about an arbitrarily chosen steady state comprising metabolite concentrations (below circles, in mM), fluxes (below reaction names, in mM h^{-1}) and inhibitory constants, K_i (in mM). Reactions follow either mass-action (MA) or Michaelis-Menten (MM) kinetics. The table summarizes entries of the Jacobian matrix of this steady state. Entries correspond to the definition in Eq. 16. For example, the entry in first column and first row represents the value of (δf_{F6P}/ δ[F6P]). Grey coloured entries are related to feedback inhibition. Abbreviations refer to Figure 1.

3. Approximating the stabilization of cellular metabolism by compartmentalization

Lisa Fürtauer and Thomas Nägele

Published in Theory in Biosciences (2016), Vol. 135, Issue 1-2, pp 73-87

DOI: 10.1007/s12064-016-0225-y

Declaration of Authorship:

LF: performed modelling, data evaluation, wrote the paper

TN: conceived the study, performed modelling, data evaluation, wrote the paper

Theory Biosci.
DOI 10.1007/s12064-016-0225-y

ORIGINAL PAPER

Approximating the stabilization of cellular metabolism by compartmentalization

Lisa Fürtauer[1] · **Thomas Nägele**[1,2]

Received: 22 December 2015 / Accepted: 21 March 2016

Abstract Biochemical regulation in compartmentalized metabolic networks is highly complex and non-intuitive. This is particularly true for cells of higher plants showing one of the most compartmentalized cellular structures across all kingdoms of life. The interpretation and testable hypothesis generation from experimental data on such complex systems is a challenging step in biological research and biotechnological applications. While it is known that subcellular compartments provide defined reaction spaces within a cell allowing for the tight coordination of complex biochemical reaction sequences, its role in the coordination of metabolic signals during metabolic reprogramming due to environmental fluctuations is less clear. In the present study, we numerically analysed the effects of environmental fluctuations in a subcellular metabolic network with regard to the stability of an experimentally observed steady state in the genetic model plant *Arabidopsis thaliana*. Applying a method for kinetic parameter normalization, several millions of probable enzyme kinetic parameter constellations were simulated and evaluated with regard to the stability information of the metabolic homeostasis. Information about the stability of the metabolic steady state was derived from real parts of eigenvalues of Jacobian matrices. Our results provide evidence for a differential stabilizing contribution of different subcellular compartments. We could identify stabilizing and destabilizing network components which we could classify according to their subcellular localization. The findings prove that a highly dynamic interplay between intracellular compartments is preliminary for an efficient stabilization of a metabolic homeostasis after environmental perturbation. Further, our results provide evidence that feedback-inhibition originating from the cytosol and plastid seem to stabilize the sucrose homeostasis more efficiently than vacuolar control. In summary, our results indicate stabilizing and destabilizing network components in context of their subcellular organization.

Keywords Stability · Subcellular network · Eigenvalue · Jacobian matrix · Compartmentalization · Systems biology

Introduction

A characteristic feature of eukaryotic cells is a compartmentalized subcellular structure resulting in numerous cellular reaction environments. Due to this compartmentalization, various biochemical reaction conditions are established in a cell enabling highly coordinated and tightly regulated reactions and reaction sequences. Particularly, such highly compartmentalized subcellular structures are a characteristic feature of plant cells (Lunn 2007). The first steps of converting light to chemically usable energy, i.e. the reaction sequence of photosynthetic carbon fixation and the central carbohydrate metabolism, are tightly linked and regulated which allows plants to efficiently respond and acclimate to environmental fluctuations (Yamori et al. 2014). The subcellular compartments chloroplast, cytosol and vacuole are connected by various transporters and

Electronic supplementary material The online version of this article (doi:10.1007/s12064-016-0225-y) contains supplementary material, which is available to authorized users.

✉ Thomas Nägele
thomas.naegele@univie.ac.at

[1] Department of Ecogenomics and Systems Biology, University of Vienna, Althanstr. 14, 1090 Vienna, Austria

[2] Vienna Metabolomics Center (VIME), University of Vienna, Althanstr. 14, 1090 Vienna, Austria

Published online: 05 April 2016

channels which enable the regulated exchange of metabolic compounds across biological membrane systems (Linka and Weber 2010). Finally, this results in a highly complex metabolic network which has been analysed by various approaches of mathematical modelling (see e.g. Grafahrend-Belau et al. 2013; Morgan and Rhodes 2002; Nägele and Weckwerth 2013; Pokhilko et al. 2014).

As sessile organisms, plants have to cope with fast and fluctuating changes in environmental conditions, e.g. a sudden increase/decrease in light intensity or temperature. To prevent a disbalance of primary and secondary photosynthetic reactions, which could easily cause the generation of reactive oxygen species (ROS) and the irreversible damage of membrane systems and cells, a new and stable metabolic homeostasis has to be established within a relatively short time period. In biochemical networks, the metabolic homeostasis results from a concerted orchestration of enzymatic activities which determine metabolic fluxes, i.e. rates of metabolic interconversion and membrane transport. The coordinate reprogramming of a metabolic homeostasis, which finally results in a new acclimated stable homeostasis, is a complex process comprising numerous enzymatic reactions and regulatory molecular circuits. Decades ago it was outlined that by linear approximation of enzymatic reaction chains, i.e. the assumption that enzyme velocities linearly depend on substrate concentrations, analytical solutions can be obtained for metabolite concentrations and metabolic fluxes at steady state conditions (Heinrich and Rapoport 1974). Defining the cardinal terms control strength, control matrix and effector strength, Heinrich and Rapoport defined the dependence of metabolic flux and metabolite concentrations on enzyme kinetic properties as well as on effector concentrations. Together with the independently developed work of Kacser and Burns (Kacser and Burns 1973) this provided the basic concept for the research field of Metabolic Control Analysis as it was summarized more than two decades later (Kacser et al. 1995).

The stability of an observed metabolic steady-state can be numerically characterized by the eigenvalues of the related Jacobian matrix (Reznik and Segrè 2010). One major obstacle in determination of the Jacobian is the lack of experimental data on enzyme kinetic parameters, e.g. the maximum velocity, v_{max}, or the substrate affinity, K_M. To overcome this limitation, Steuer and co-workers have developed the approach of Structural Kinetic Modelling,

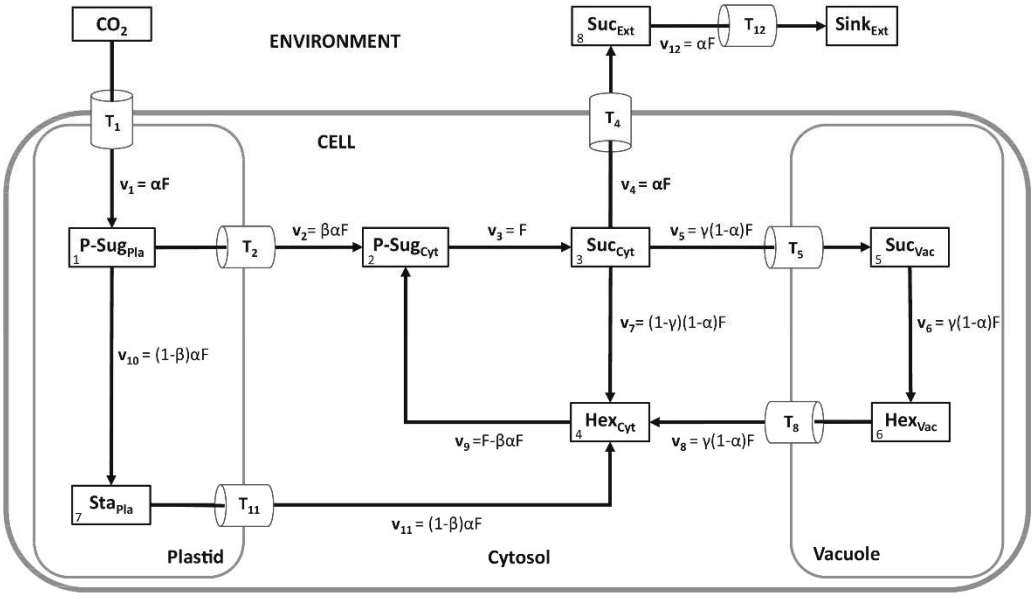


Fig. 1 Schematic model of the central subcellular carbohydrate metabolism in plant leaf mesophyll cells. The assimilation of CO_2 supplies the pool of plastidial sugar phosphates ($P\text{-}Sug_{Pla}$) which is the first metabolic substrate for all other reactions. Metabolites are written in *boxes*, transport systems are indicated by T_j and enzyme-driven interconverting reactions by v_j. Steady state fluxes for all reactions are additionally shown (e.g. αF). *Pla* plastidial, *Cyt* cytosolic, *Vac* vacuolar, *Ext* extracellular, *P-Sug* sugar-phosphates, *Sta* starch, *Suc* sucrose, *Hex* hexoses

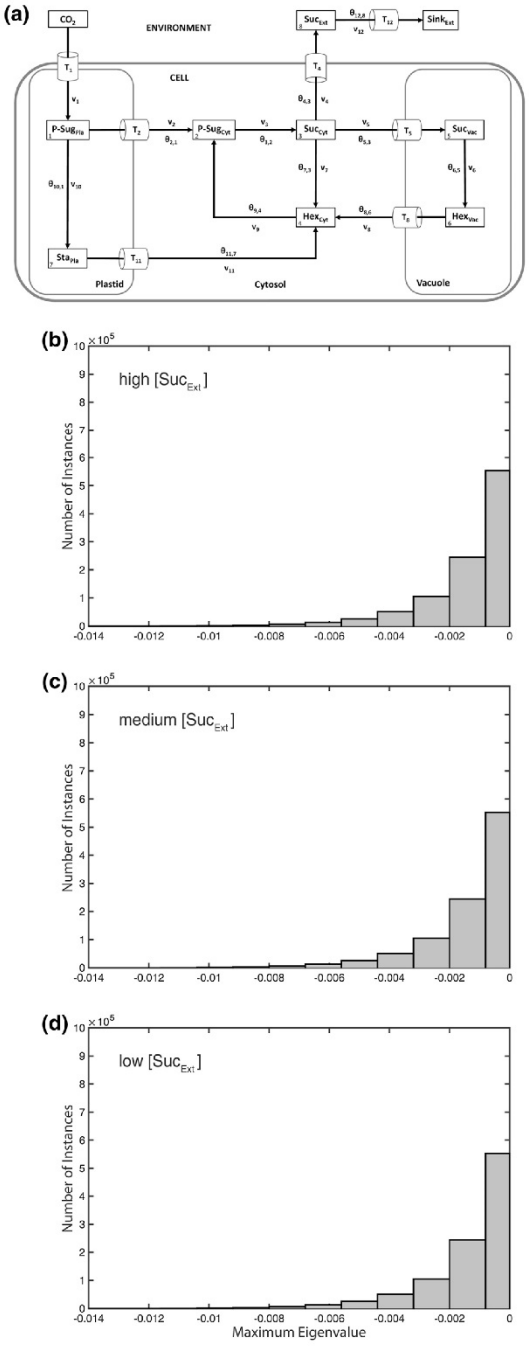
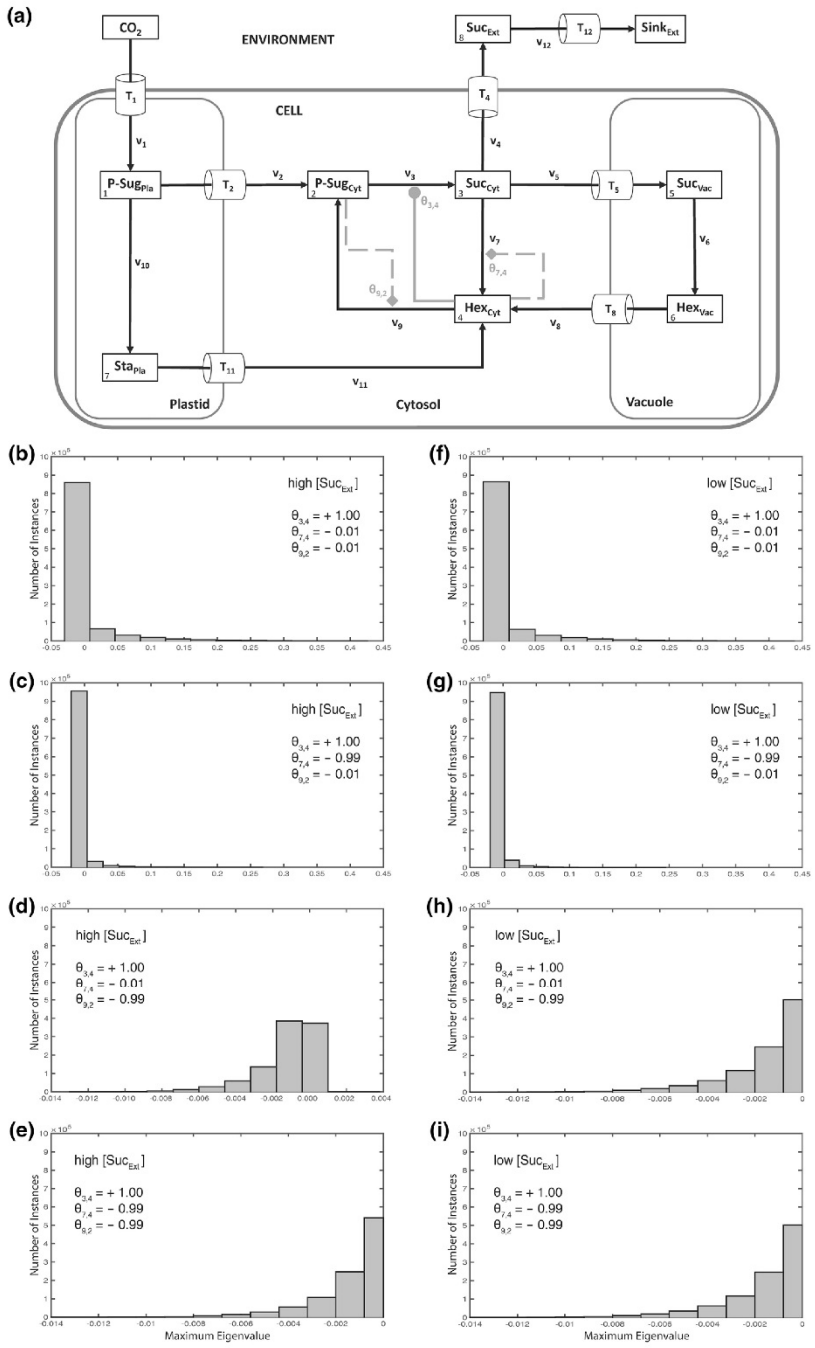


Fig. 2 Model structure and histograms of maximum eigenvalue real parts without activation and inhibition. Calculations for the shown model configuration (**a**) were performed 10^6 times for high (**b**), medium (**c**) and low (**d**) concentration of extracellular/apoplastic sucrose concentrations. Further detailed information, e.g. about maximum values of the eigenvalue real part distribution, is provided in Supplementary Table S4


(a)
ENVIRONMENT
CELL
Plastid
Cytosol
Vacuole
(b)
high [Suc_Ext]
θ3,4 = + 1.00
θ7,4 = − 0.01
θ9,2 = − 0.01
(f)
low [Suc_Ext]
θ3,4 = + 1.00
θ7,4 = − 0.01
θ9,2 = − 0.01
(c)
high [Suc_Ext]
θ3,4 = + 1.00
θ7,4 = − 0.99
θ9,2 = − 0.01
(g)
low [Suc_Ext]
θ3,4 = + 1.00
θ7,4 = − 0.99
θ9,2 = − 0.01
(d)
high [Suc_Ext]
θ3,4 = + 1.00
θ7,4 = − 0.01
θ9,2 = − 0.99
(h)
low [Suc_Ext]
θ3,4 = + 1.00
θ7,4 = − 0.01
θ9,2 = − 0.99
(e)
high [Suc_Ext]
θ3,4 = + 1.00
θ7,4 = − 0.99
θ9,2 = − 0.99
(i)
low [Suc_Ext]
θ3,4 = + 1.00
θ7,4 = − 0.99
θ9,2 = − 0.99
Number of Instances
Maximum Eigenvalue

◀**Fig. 3** Model structure and histograms of maximum eigenvalue real parts with cytosolic activation and inhibition. Calculations for the shown model configuration (**a**) were performed 10^6 times for high (**b–e**), and low (**f–i**) concentration of extracellular/apoplastic sucrose concentrations. Steps of metabolic activation are indicated by *grey filled circles*, steps of inhibition are indicated by *grey filled diamonds* and *dashed lines*. Particular settings in the $\boldsymbol{\theta}$ matrix are indicated within the *single histograms*. Further detailed information, e.g. about maximum values of the eigenvalue real part distribution, is provided in Supplementary Table S4

SKM, in which normalized enzyme kinetic parameters replace conventional parameters (Steuer et al. 2006). This approach has been proven to be suitable for identification of stability and robustness of metabolic states (Grimbs et al. 2007), but also for specific network structures like metabolic cycles (Reznik and Segrè 2010). In a previous approach we have applied the SKM approach to determine stability properties of a highly simplified reaction core of the central carbohydrate metabolism in plant leaves (Henkel et al. 2011). While we could already provide evidence for a differential stabilization effect of feedback inhibition originating from different metabolic sites in a small metabolic cycle belonging to the central carbohydrate metabolism (Henkel et al. 2011), we did not consider a possible effect of subcellular compartmentalization. Here, we enlarge our previous approach aiming at a comprehensive understanding of how a subcellular allocation of central metabolic compounds affects the stability properties of a metabolic homeostasis. Finally, our approach intends to shed light on one of the most characteristic features of eukaryotic cells in context of a successful acclimation strategy to a fluctuating environment.

Biological model preliminaries

Based on approaches of mathematical modelling in previous publications (Nägele et al. 2010; Nägele and Heyer 2013; Nägele et al. 2012), we derived a simplified, but still representative, metabolic model of the central carbohydrate metabolism in leaf mesophyll cells (Fig. 1). The model comprised the three intracellular compartments plastid, cytosol and vacuole, as well as an extracellular environment. Transporters (T_j) were interpreted as a pool of all involved membrane transporters, even though some of these transporters might still have to be characterized (Lunn 2007; Linka and Weber 2010; Bräutigam and Weber 2011). The influx of the system, T_1 and v_1, represented the complete process of CO_2 transport and fixation resulting in the pool of plastidial sugar phosphates ($P\text{-}Sug_{Pla}$). Environmental fluctuations α, e.g. occurring due to sudden changes in light intensity or temperature, directly affected the flux F of v_1 as a constant scalar ($v_1 = \alpha F$). The pool of plastidial sugar phosphates was substrate for two reactions: v_2, the export to the cytosol ($P\text{-}Sug_{Cyt}$), catalysed by membrane transporters, and v_{10}, representing the intraplastidial conversion to starch (Sta_{Pla}). Cytosolic sucrose, Suc_{Cyt}, was the substrate for the export reaction via T_4 supplying the pool of extracellular sucrose, Suc_{Ext}. Reaction v_5 indicated the transport of sucrose from the cytosol (Suc_{Cyt}) to the vacuole (Suc_{Vac}). In both compartments, hydrolytic cleavage of sucrose via invertase yielded free hexoses [v_6 (Hex_{Vac}) and v_7 (Hex_{Cyt})]. Vacuolar hexose transport to the cytosol was described via transport process T_8 resulting in a circular structure. Degradation of plastidial starch (Sta_{Pla}) to hexose equivalents and their transport to the cytosol was described by transport T_{11} and reaction v_{11}. Reaction v_9 described the phosphorylation of cytosolic hexoses via hexokinase activity (Claeyssen and Rivoal 2007), supplying the pool of sugar phosphates ($P\text{-}Sug_{Cyt}$). Finally, reaction v_{12} which was coupled to transport T_{12}, described the transport of apoplastic sucrose, Suc_{Ext}, to sink organs ($Sink_{Ext}$).

Mathematical model preliminaries

The model comprised 12 reactions (v_j; $j = 1,\ldots,r$ here: $r = 12$), steady state concentrations of different metabolites ($c_{0,i}$ $i = 1,\ldots,m$ here: $m = 8$) and their fluxes $v_j(c_{0,i})$. Experimentally determined steady state concentrations of metabolites, which are provided in the supplements (Supplementary Table S1), were derived from previous studies (Nägele and Heyer 2013; Nägele et al. 2012). Steady state concentrations of extracellular sucrose (Suc_{Ext}), which were not experimentally determined in these studies, were set to a low (0.01 mM), medium (15 mM) and a high (45 mM) concentration when compared to the cytosolic sucrose concentration (15 mM). With these settings, three possible scenarios were simulated: (1) a sucrose concentration gradient from the intracellular to the extracellular space (Suc_{Ext} low), (2) no sucrose gradient (Suc_{Ext} medium), and (3) a sucrose concentration gradient from the extracellular to the intracellular space (Suc_{Ext} high). To define a steady state equilibrium, the input flux v_1 and output flux v_4 were set to an equal value (αF). The flux F was set to be constant ($F = 1$). For simulation of random environmental fluctuations and the response of the system, alpha (α) and the proportion characters beta (β) and gamma (γ), were randomly varied in the interval (0; 1). While β quantified the relative proportion of carbon flux from chloroplast to cytosol, γ quantified the proportion of carbon flux from cytosol to vacuole. Based on the SKM approach (Steuer et al. 2006), the Jacobian matrix $\mathbf{J}$ was defined as the product of matrices $\boldsymbol{\Lambda}$ and $\boldsymbol{\theta}$ (Eq. 1):

$$J := \Lambda\theta \tag{1}$$

Here, $\mathbf{\Lambda}$ is the stoichiometric matrix (**N**) of the considered metabolic system, normalized to steady-state fluxes $v(c_0)$ and metabolite concentrations $c_{0,i}$ (Eq. 2):

$$\Lambda_{i,j} := N_{i,j}\frac{v_j(c_{0,i})}{c_{0,i}} \tag{2}$$

Entries of the matrix $\boldsymbol{\theta}$ represent normalized elasticities, i.e. the degree of saturation of a normalized flux μ with regard to the normalized substrate concentration x (Eq. 3):

$$\theta := \frac{\mathrm{d}\mu}{\mathrm{d}x} = \frac{\mathrm{d}\frac{v(c)}{v(c_0)}}{\mathrm{d}\frac{c(t)}{c_0}} \tag{3}$$

A non-zero element $\theta_{j,i}$ in matrix $\boldsymbol{\theta}$ indicated the involvement of a metabolite as a substrate, as an activating or inhibiting compound. In case of activating compounds, elasticities in $\boldsymbol{\theta}$ were defined within the interval (0; 1], while for inhibitory compounds the interval was (−1; 0). A detailed explanation of these entries in context of enzyme kinetics, such as Michaelis–Menten kinetics, was provided previously (Reznik and Segrè 2010).

A detailed and explicit formulation of the $\mathbf{\Lambda}$ and $\boldsymbol{\theta}$ matrices derived in the present study is provided in the Supplements (Supplementary files S2 and S3) while a short and very fundamental explanation of eigenvalues in context of a differential equation is provided in the following paragraph.

Eigenvalues of the Jacobian **J** characterize a metabolic steady state with regard to its stability. The steady state concentration $c_{0,\mathrm{i}}$ of a metabolite $\mathrm{c_i}$, is related to the time-dependent concentration $c_i(t)$ by a fluctuation term $\tau_i(t)$ (Eq. 4):

$$c_i(t) = c_{0,i} + \tau_i(t) \tag{4}$$

The first element of a Taylor expansion reveals the linearization around the considered steady state (Eq. 5):

$$\frac{\mathrm{d}\tau_i(t)}{\mathrm{d}t} = \sum_{p=1}^{m} J_{i,p}\tau_p(t) \tag{5}$$

Integration yields the general solution (Eq. 6):

$$\tau_i(t) = \sum_{p=1}^{m} C_{i,p}e^{\lambda_p t} \tag{6}$$

here, **C** represents constants which depend on the initial fluctuation conditions. The complex number λ_i represents the eigenvalues of the Jacobian matrix **J** (Eq. 7):

$$det(\boldsymbol{J} - \lambda\boldsymbol{I_m}) = 0 \tag{7}$$

$\mathbf{I_m}$ represents the unit matrix, and Eq. 7 is known as the characteristic equation.

Hence, the eigenvalue λ_i characterizes the solution of Eq. 6: for $\mathrm{re}(\lambda_i) < 0$, the fluctuation term $\tau_i(t)$ decays exponentially (stability), while $\mathrm{re}(\lambda_i) > 0$ results in an exponential increase (instability). While the imaginary part of λ_i, $\mathrm{im}(\lambda_i)$, is necessary to exactly classify the local dynamics and to differentiate a node from a focus or a saddle (Steuer 2007), the present study only accounted for the real parts of λ_i which is sufficient to differentiate stable from instable solutions.

Stability characteristics of steady states were evaluated by the maximum real part of the eigenvalue ($\lambda_{\max}$). To simulate different activation and inhibition scenarios, the $\boldsymbol{\theta}$ matrix was adapted accordingly. In general, for an activation, the corresponding $\theta_{j,i}$ value was set to 1 to obtain a maximal perturbation. Exceptions from this are directly indicated in the text or in the figures. Inhibition scenarios were simulated with a weak ($\theta_{j,i} = -0.01$) and strong ($\theta_{j,i} = -0.99$) effect. For all models, 10^6 iterations were calculated to reduce the probability of numerical coincidences. For reactions v_2 to v_{12}, the enzyme-substrate interaction was described by entries in the elasticity matrix $\boldsymbol{\theta}$ indicated by the subscript index (first index number, j: reaction; second index number, i: metabolite pool). Elasticities were chosen randomly in the interval of (0; 1) for every iteration process to reveal the system behaviour over all 10^6 iterations. The source code which was applied for calculations is provided in the supplements (Supplementary file S5).

Subcellular stability analysis in context of environmental perturbation

Simulation of 10^6 variations, i.e. implicit normalized parameter sets, in the $\boldsymbol{\theta}$ matrix without any metabolic activation or inhibition yielded only stable systems, irrespective of a variation in external sucrose concentration (Fig. 2a–d). Additionally, in the three $\mathrm{Suc_{Ext}}$ concentration scenarios, no differences between resulting $\lambda_{\max}$ values were observed (Fig. 2b–d). Conclusively, when steps of metabolic activation and inhibition were absent, a variation of external sucrose levels did not affect the stability properties of the considered intracellular metabolic system.

In the next step, instances of metabolic activation and feedback-inhibition were incorporated into the model, belonging to the central cytosolic carbohydrate metabolism as described previously in a simplified model structure (Henkel et al. 2011). In this model structure, cytosolic hexoses feedforward-activated the biosynthesis of sucrose and feedback-inhibited the cytosolic cleavage of sucrose (Fig. 3a). Additionally, cytosolic sugar phosphates feedback-inhibited the phosphorylation of free cytosolic

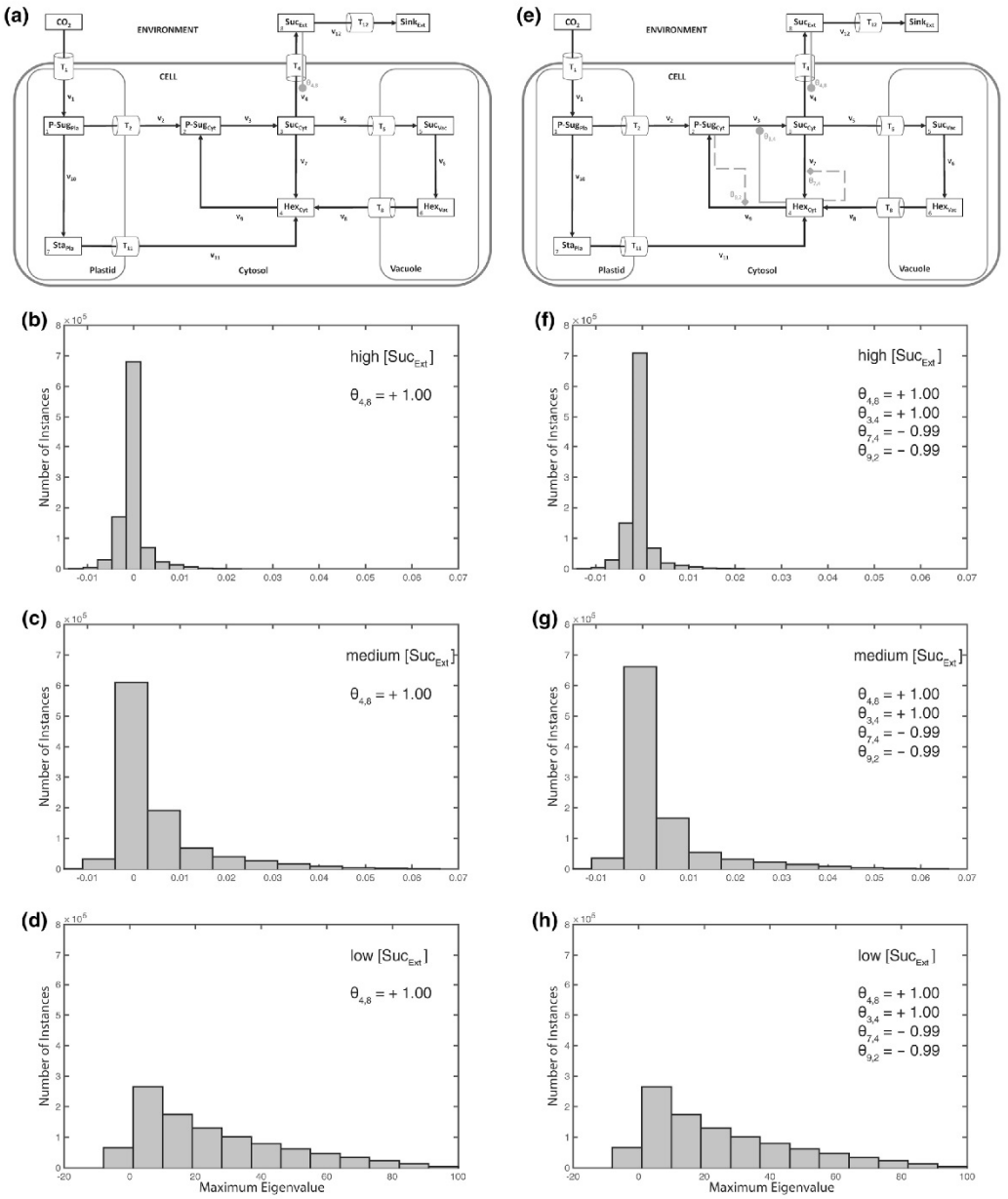


Fig. 4 Activation of sucrose export by extracellular sucrose. Calculations for the shown model configuration (**a**) were performed 10^6 times for high (**b**, **f**), medium (**c**, **g**), and low (**d**, **h**) concentration of extracellular/apoplastic sucrose concentrations. Steps of metabolic activation are indicated by *grey filled circles*, steps of inhibition are indicated by *grey filled diamonds* and *dashed lines*. Particular settings in the $\boldsymbol{\theta}$ matrix are indicated within the *single histograms*. Further detailed information, e.g. about maximum values of the eigenvalue real part distribution, is provided in Supplementary Table S4

hexoses. In contrast to the previous observation without any regulatory instance, stability properties were now observed to depend on external sucrose concentrations (Fig. 3b–i). While no difference in system stability was observed as long as feedback-inhibition of cytosolic sugar phosphates was set to be weak ($\theta_{9,2} = -0.01$; Fig. 3b, c, f, g), this pattern changed with a strong feedback of cytosolic sugar phosphates and a weak feedback-inhibition of cytosolic hexoses ($\theta_{7,4} = -0.01$; Fig. 3d, h): in this constitution, high external sucrose concentrations resulted in instabilities (Fig. 3d). Instabilities, i.e. positive eigenvalue real parts, disappeared again when both feedback-inhibitions, i.e. $\theta_{9,2}$ and $\theta_{7,4}$, were defined to be strong (Fig. 3e, i). These stability characteristics were also observed when the steady state flux was divided in half ($F = 0.5$) or doubled ($F = 2$; see Supplementary Figure S6).

To further analyse the possible role and effect of an extracellular pool on the intracellular system stability, we combined the previously discussed metabolic scenarios (Figs. 2, 3) with an activation of the export reaction v_4 ($Suc_{Cyt} \rightarrow Suc_{Ext}$) by extracellular sucrose ($\theta_{4,8}$; Fig. 4). In a biological context, such a hypothetical scenario might be interpreted as the result of a molecular signalling process which connects sink organs, e.g. root, with the source organs, e.g. leaves. The experimental and theoretical analysis of such sink-source interactions represents a challenging and difficult to assess research question in current plant biology which has been focused by numerous studies during the last decades (e.g. see Bancal and Soltani 2002; Brauner et al. 2014; Milne et al. 2013; Paul and Foyer 2001; Sonnewald and Willmitzer 1992; Tiessen and Padilla-Chacon 2012; Turgeon and Wolf 2009). In our stability analysis, the activation of sucrose export from the intracellular to the extracellular environment resulted in a destabilization of the intracellular metabolic homeostasis (Fig. 4a–d) which could not be stabilized by strong feedback-inhibitions of the cytosolic sugar metabolism (Fig. 4e–h). Yet, although instabilities occurred under all considered external sucrose concentrations, we detected an increase of maximum positive real parts of eigenvalues with lower external sucrose concentrations.

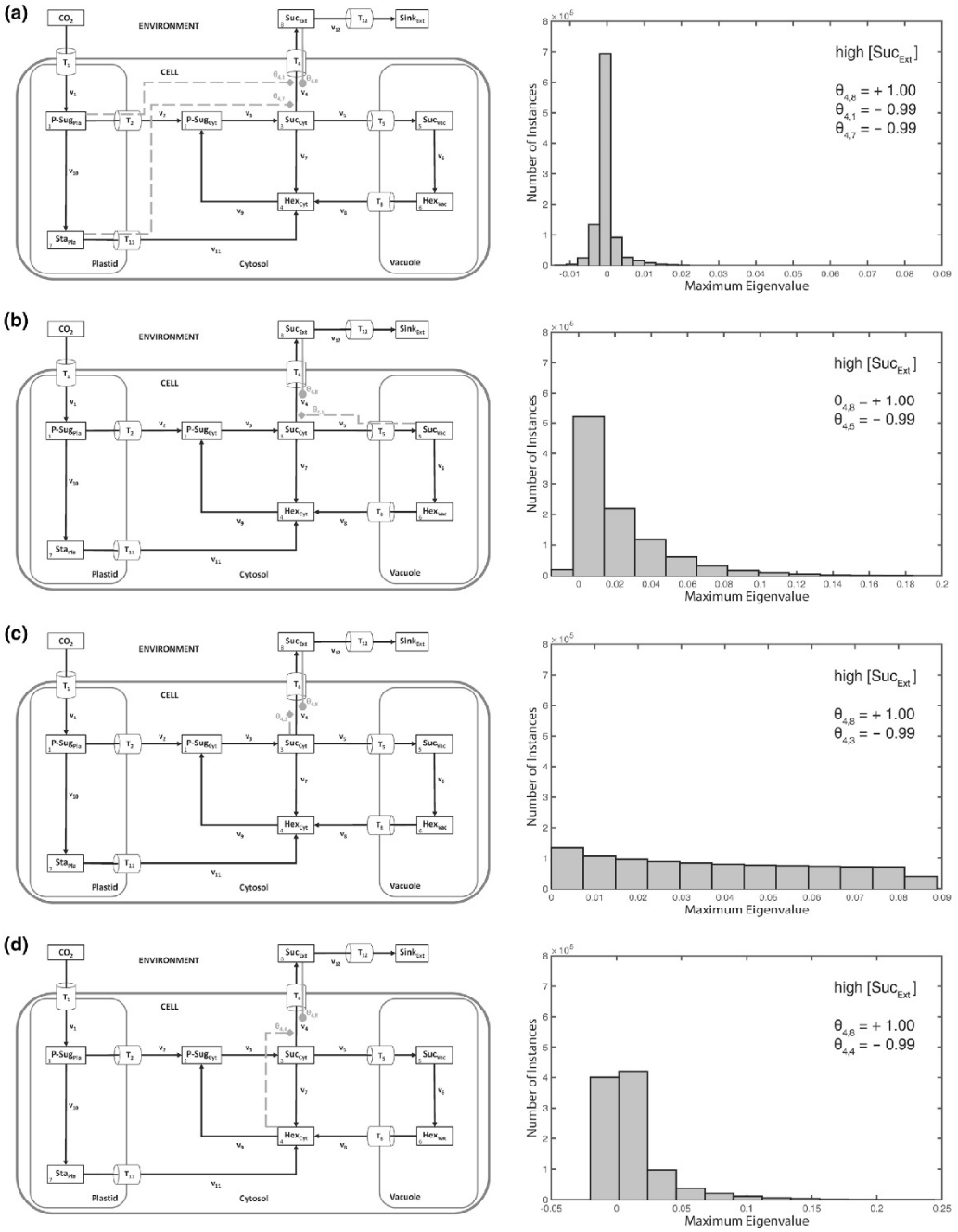
Based on these observations, we explored various combinations of subcellular feedback-inhibition with regard to their capacity to re-stabilize the metabolic homeostasis after perturbation by a high external sucrose concentration (Fig. 5). In summary, neither a sole feedback-inhibition from plastidial (Fig. 5a) nor from vacuolar (Fig. 5b) nor from cytosolic (Fig. 5c, d) metabolic compounds could fully stabilize the system under all explored parameter constellations, i.e. elasticity variations. Further, also for the combination of plastidial and cytosolic regulatory instances, instable solutions were derived (Fig. 6a, b). Yet, in case of only a weak activation of sucrose export by extracellular sucrose ($\theta_{4,8} = 0.001$), the combination of cytosolic and plastidial feedback-inhibition was found to yield a stable system (Fig. 6c). In contrast, a combination of plastidial with vacuolar (instead of cytosolic) regulatory instances was neither found to stabilize the system with strong nor weak activation of sucrose export by extracellular sucrose (Figs. 6d–f, 7a–c). In this context, we hypothesised that, following an extracellular perturbation, the vacuolar reactions of the model were negligible or even destabilizing the intracellular system. We proved this hypothesis by ignoring the vacuolar reaction cycle (Fig. 7d) which finally resulted in a stabilized system, even after strong activation by extracellular sucrose (Fig. 7e). This system stability was not persistent when vacuolar reactions were included (see Fig. 6a, b).

These observed outputs on steady state stability were related to the model assumption of an activation of sucrose export by extracellular sucrose, which, in a biological context, might be interpreted as a regulation of source metabolism by sink demand. To test a contrasting regulatory pattern, we replaced the activation term for extracellular sucrose on sucrose export by an inhibition term ($\theta_{4,8} \in (-1; 0)$; Fig. 8a). The combination of plastidial and cytosolic feedback-inhibitions resulted in a system stabilization irrespective of extracellular sucrose concentration (Fig. 8b, c). Even enhancing the previous observations, the replacement of cytosolic by vacuolar regulation yielded (purely) unstable solutions (Fig. 9) for both high and low extracellular sucrose concentrations.

Fig. 5 Differential regulatory stabilization of externally induced perturbation. Calculations for the shown model configurations (**a**–**d**) were performed 10^6 times for high extracellular/apoplastic sucrose concentration. Distributions of resulting maximum eigenvalue real parts are shown in histograms. Steps of metabolic activation are indicated by *grey filled circles*, steps of inhibition are indicated by *grey filled diamonds* and *dashed lines*. Particular settings in the $\boldsymbol{\theta}$ matrix are indicated within the *single histograms*. Further detailed information, e.g. about maximum values of the eigenvalue real part distribution, is provided in Supplementary Table S4

Discussion and conclusion

Metabolic reprogramming of biochemical networks due to environmental perturbations is a complex and multifaceted process. We have analysed a highly simplified, but still representative, biochemical network of the central carbohydrate metabolism in plant leaf cells with regard to its capability to respond and stabilize after a perturbation. Our results indicate that different subcellular organelles potentially contribute differentially to the reprogramming process, yielding a stable metabolic homeostasis only under certain preconditions. While it might not be


(a)
ENVIRONMENT
CELL
Plastid
Cytosol
Vacuole
high [Suc$_{Ext}$]
$\theta_{4,8}$ = + 1.00
$\theta_{4,1}$ = − 0.99
$\theta_{4,7}$ = − 0.99
Number of Instances
Maximum Eigenvalue
(b)
ENVIRONMENT
CELL
Plastid
Cytosol
Vacuole
high [Suc$_{Ext}$]
$\theta_{4,8}$ = + 1.00
$\theta_{4,5}$ = − 0.99
Number of Instances
Maximum Eigenvalue
(c)
ENVIRONMENT
CELL
Plastid
Cytosol
Vacuole
high [Suc$_{Ext}$]
$\theta_{4,8}$ = + 1.00
$\theta_{4,3}$ = − 0.99
Number of Instances
Maximum Eigenvalue
(d)
ENVIRONMENT
CELL
Plastid
Cytosol
Vacuole
high [Suc$_{Ext}$]
$\theta_{4,8}$ = + 1.00
$\theta_{4,4}$ = − 0.99
Number of Instances
Maximum Eigenvalue

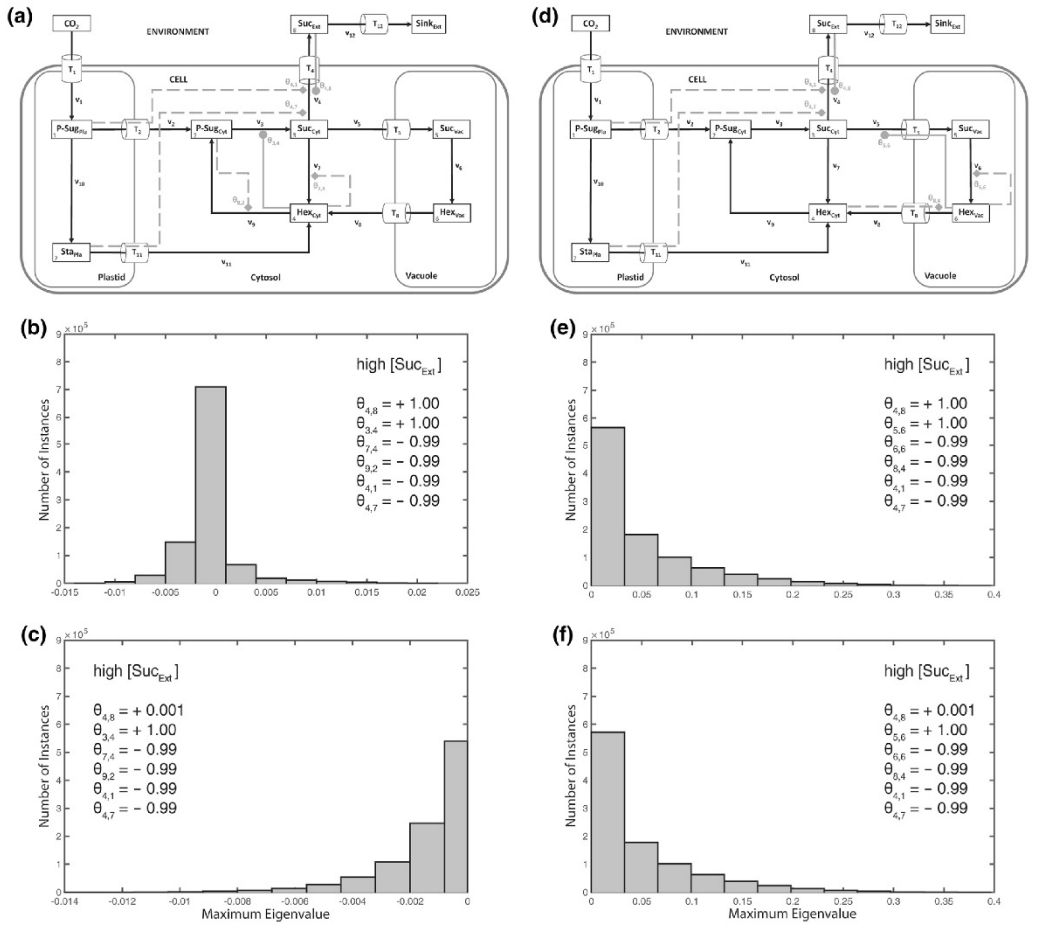


Fig. 6 Differential stabilization output of cytosolic and vacuolar regulation. Calculations for the shown model configurations (**a**, **d**) were performed 10^6 times for high extracellular/apoplastic sucrose concentration. Distributions of resulting maximum eigenvalue real parts are shown in *histograms* (**b**, **c**, **e**, **f**). Steps of metabolic activation are indicated by *grey filled circles*, steps of inhibition are indicated by *grey filled diamonds* and *dashed lines*. Particular settings in the $\boldsymbol{\theta}$ matrix are indicated within the *single histograms*. Further detailed information, e.g. about maximum values of the eigenvalue real part distribution, is provided in Supplementary Table S4

surprising that different cellular organelles contribute differentially to metabolic regulation due to their different pH, volumes or proteomes (Ito et al. 2014; Joyard et al. 2010; Lunn 2007; Millar and Taylor 2014), the quantification of these contributions remains challenging. One example indicating the regulatory complexity on a subcellular and molecular level is the finding that vacuolar sugar carriers contribute differentially to metabolic reprogramming in different environmental conditions (Klemens et al. 2013). Klemens and co-workers showed that the metabolic consequences in the central carbohydrate metabolism, which were due to an increased activity of the vacuole-located carrier SWEET16, depended on the type of environmental stress. Based on their experimental observations the authors finally concluded that SWEET16 activity has to be tightly linked to developmental regulation in Arabidopsis under stress conditions (Klemens et al. 2013). Together with other classes of sugar transporters, SWEETs have also been suggested to be localized in the plasma membrane of phloem parenchyma cells and to play a central role in phloem loading (Chen 2014). While our proposed and analysed model of central carbohydrate metabolism does

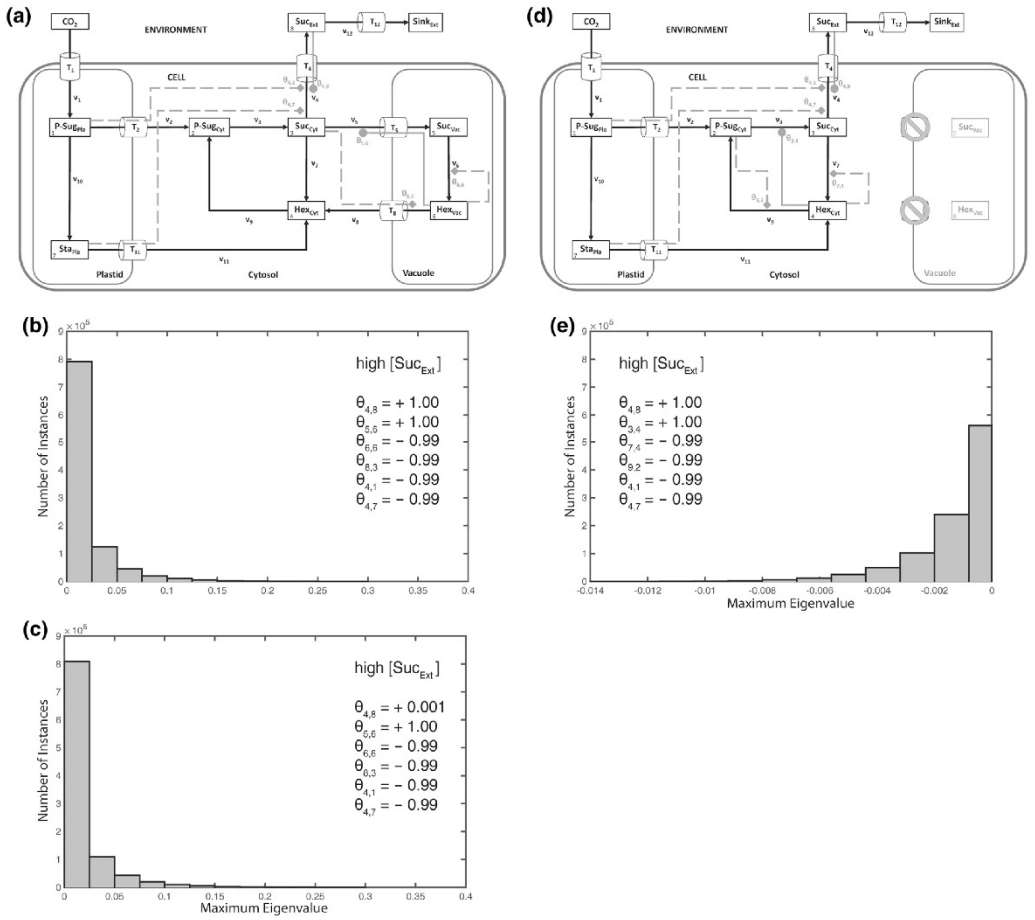


Fig. 7 Contribution of vacuolar regulation to stabilization of a perturbed metabolic homeostasis. All calculations for the shown model configurations (**a**, **d**) were performed 10^6 times for high extracellular/apoplastic sucrose concentration. Distributions of resulting maximum eigenvalue real parts are shown in histograms (**b**, **c**, **e**). Steps of metabolic activation are indicated by *grey filled circles*, steps of inhibition are indicated by *grey filled diamonds* and *dashed lines*. Particular settings in the **θ** matrix are indicated within the *single histograms*. Further detailed information, e.g. about maximum values of the eigenvalue real part distribution, is provided in Supplementary Table S4

not account for different cell types, such as mesophyll and phloem parenchyma cells, but only differentiates between cellular metabolism and environment in general, our results might still be interpretable in this molecular context. One evident output of our stability analysis was the finding that feedback-inhibition originating from or directed to vacuolar carbohydrate metabolism had a less stabilizing effect on plastidial or cytosolic metabolism than we found it to occur vice versa. Although our approach is based on an abstract and highly simplified representation of these organelles, derived stability properties might still be representative for larger and more detailed networks as they arise from such small network modules which can be expected to occur frequently in large-scale networks. Hypothesising about a possible functional context of the observed destabilisation by the vacuolar part of the reaction network, we speculate that oscillatory processes, which might arise from instabilities, are possibly part of signal transduction mechanisms interconnecting carbohydrate metabolism of the vacuole with the cytosol. Previously, it has been discussed

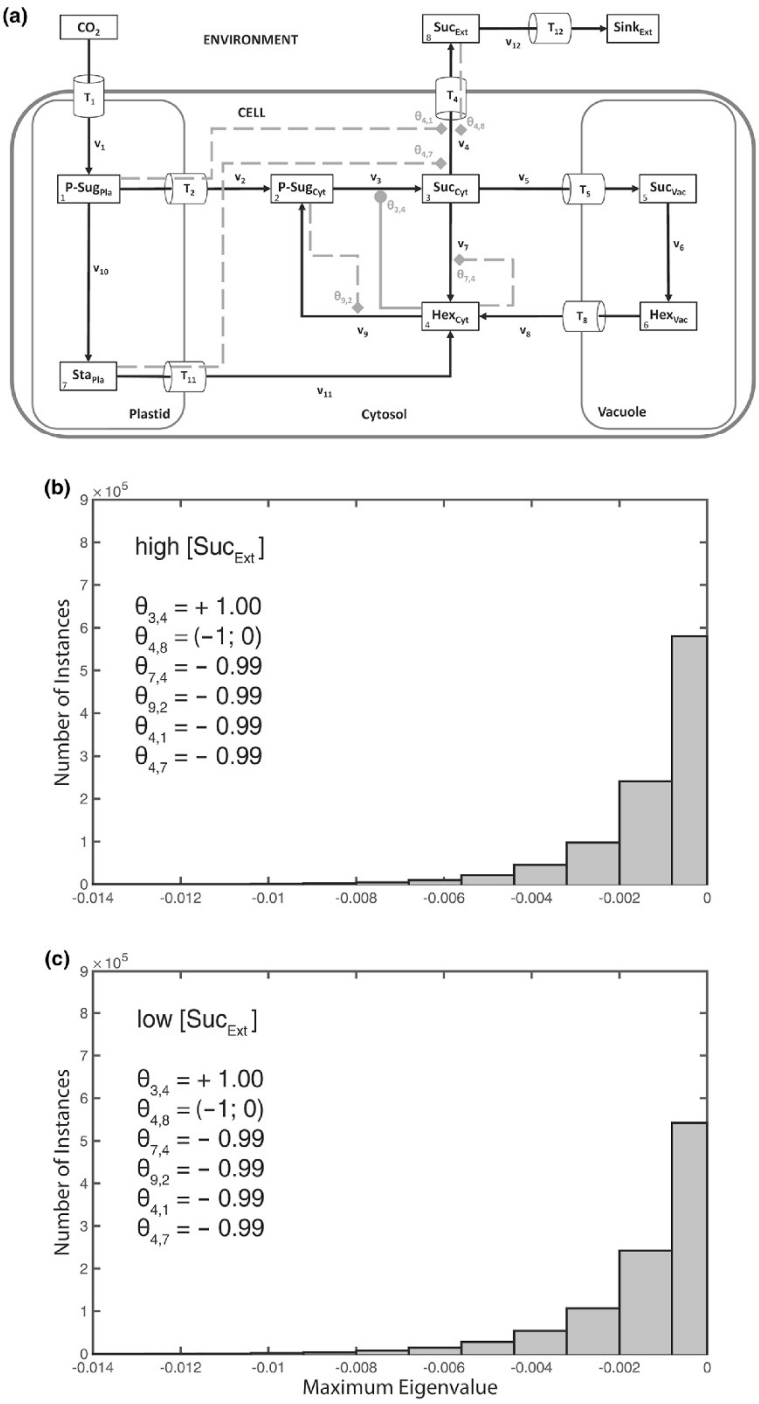


Fig. 8 Stabilization output of external feedback inhibition. Calculations for the shown model configuration (**a**) were performed 10^6 times for high (**b**) and low (**c**) extracellular/apoplastic sucrose concentration. Distributions of resulting maximum eigenvalue real parts are shown in both histograms. Steps of metabolic activation are indicated by *grey filled circles*, steps of inhibition are indicated by *grey filled diamonds* and *dashed lines*. Particular settings in the $\boldsymbol{\theta}$ matrix are indicated within the *single histograms*. Further detailed information, e.g. about maximum values of the eigenvalue real part distribution, is provided in Supplementary Table S4

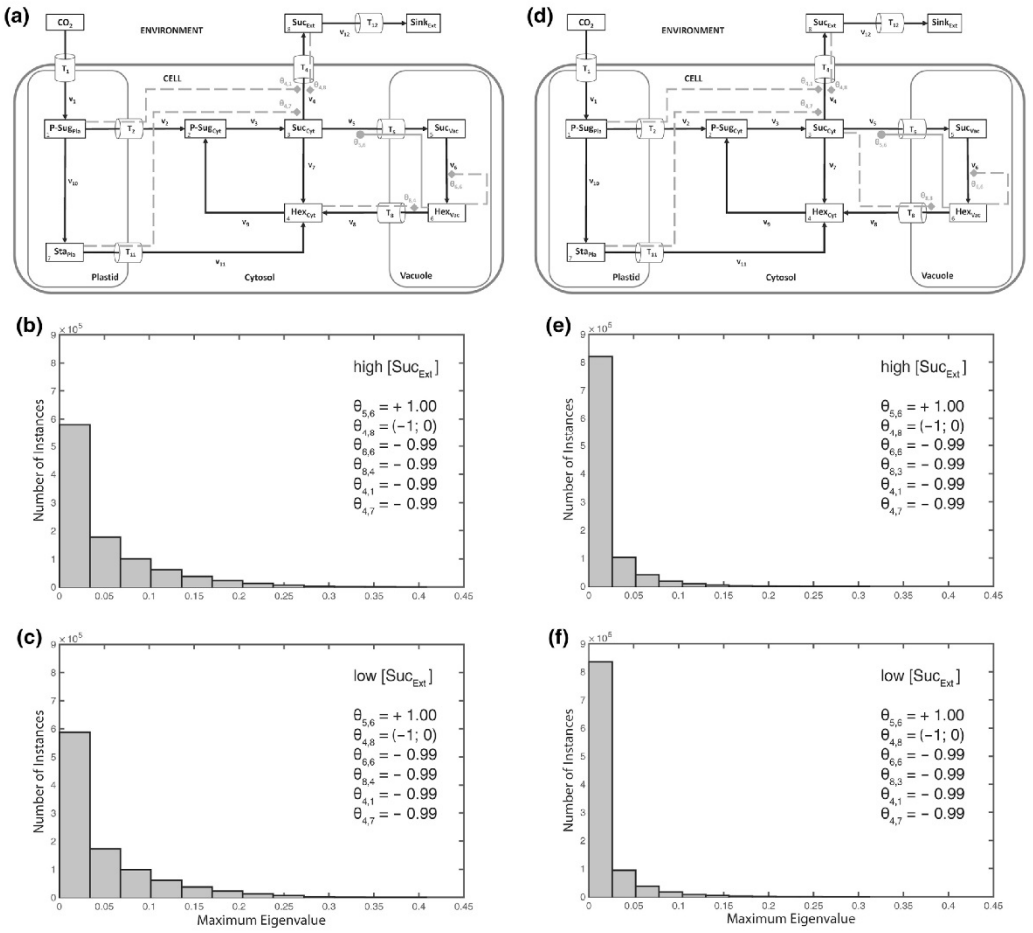


Fig. 9 Destabilization by vacuolar regulation. Calculations for the shown model configurations (**a**, **d**) were performed 10^6 times for high (**b**, **e**) and low (**c**, **f**) extracellular/apoplastic sucrose concentration. Distributions of resulting maximum eigenvalue real parts are shown in *histograms*. Steps of metabolic activation are indicated by *grey filled circles*, steps of inhibition are indicated by *grey filled diamonds* and *dashed lines*. Particular settings in the **θ** matrix are indicated within the *single histograms*. Further detailed information, e.g. about maximum values of the eigenvalue real part distribution, is provided in Supplementary Table S4

that oscillations in genetic or biochemical networks might represent a powerful mechanism to encode and transfer information both in time and space (Cheong and Levchenko 2010). This oscillatory information transfer might particularly play a role in vacuolar metabolism due to the diverse roles which it plays in plant leaf metabolism. First, it represents a cellular storage compartment for primary metabolites, but also for proteins or pigment molecules. Second, tonoplast monosaccharide transporters (TMTs) have been shown to be regulated by a mitogen-activated triple kinase (MAPKKK) which directly connects the vacuolar sugar metabolism to a central whole-cell signalling network affecting plant development, yield and stress tolerance (Neuhaus and Trentmann 2014; Wingenter et al. 2010). Third, in a previous study we have shown that vacuolar invertase activity, catalysing hydrolytic cleavage of vacuolar sucrose, affects whole plant carbon metabolism in *Arabidopsis thaliana* (Nägele et al. 2010). Although these examples comprise only a fraction of the in vivo metabolic and regulatory processes of vacuolar carbohydrate metabolism, it sums up to a complex regulatory picture which has to be tightly coupled to other cellular

processes occurring in different compartments. Further, as sucrose represents the main transport sugar in the phloem (Lemoine et al. 2013), and, hence, links carbohydrate metabolism of source and sink tissue, the homeostasis of its intracellular levels and the concentration gradient across the plasma membrane are of central importance for carbon and energy metabolism of the whole plant. In contrast to carbohydrate metabolism of chloroplasts, which directly results in the pool of sugar phosphates being substrate for the whole-plant carbon homeostasis, vacuolar carbohydrate metabolism is attached to cytosolic metabolism which rather indirectly connects vacuolar metabolism to reactions of sugar biosynthesis and the transport of carbohydrates to sink organs. This might necessitate a differential regulatory strategy than in other compartments which could result in the observed characteristic stabilization pattern.

In conclusion, the presented findings indicate that, depending on extracellular fluctuations of metabolite concentrations, differential regulatory consequences might be necessary to efficiently stabilize intracellular networks. Particularly, in context of the very general question how sources interact with sinks, this indicates a highly dynamic interplay and allows for the speculation of the existence of multiple regulatory strategies which enable plants, and organisms in general, to respond to environmental fluctuations. Finally, it can be concluded that the diversity of molecular interactions between structural and regulatory elements of subcellular compartments contributes significantly to the stabilization of a cellular metabolic homeostasis.

Acknowledgments We especially would like to thank Wolfram Weckwerth, Jakob Weiszmann, Matthias Nagler, Reinhard Turetschek and Lena Fragner for critical discussions and computational assistance. Further, we thank our collaboration partners at the Department of Plant Biotechnology, University of Stuttgart, headed by Prof. Arnd G. Heyer, for general support and advice in context of subcellular compartmentalization. This work was supported by the Austrian Science Fund (FWF), Project I 2071 and P 26342.

References

Bancal P, Soltani F (2002) Source-sink partitioning. Do we need Münch? J Exp Bot 53:1919–1928

Brauner K, Hörmiller II, Nägele T, Heyer AG (2014) Exaggerated root respiration accounts for growth retardation in a starchless mutant of Arabidopsis thaliana. Plant J Cell Mol Biol 79:82–91. doi:10.1111/tpj.12555

Bräutigam A, Weber APM (2011) Do metabolite transport processes limit photosynthesis? Plant Physiol 155:43–48. doi:10.1104/pp.110.164970

Chen LQ (2014) SWEET sugar transporters for phloem transport and pathogen nutrition New Phytol 201:1150–1155

Cheong R, Levchenko A (2010) Oscillatory signaling processes: the how, the why and the where. Curr Opin Genet Dev 20:665–669. doi:10.1016/j.gde.2010.08.007

Claeyssen E, Rivoal J (2007) Isozymes of plant hexokinase: occurrence, properties and functions. Phytochemistry 68:709–731. doi:10.1016/j.phytochem.2006.12.001

Grafahrend-Belau E, Junker A, Eschenroder A, Muller J, Schreiber F, Junker BH (2013) Multiscale metabolic modeling: dynamic flux balance analysis on a whole-plant scale. Plant Physiol 163:637–647. doi:10.1104/pp.113.224006

Grimbs S, Selbig J, Bulik S, Holzhutter HG, Steuer R (2007) The stability and robustness of metabolic states: identifying stabilizing sites in metabolic networks. Mol Syst Biol 3:146. doi:10.1038/msb4100186

Heinrich R, Rapoport TA (1974) A linear steady-state treatment of enzymatic chains. General properties, control and effector strength. Eur J Biochem 42:89–95

Henkel S, Nägele T, Hörmiller I, Sauter T, Sawodny O, Ederer M, Heyer AG (2011) A systems biology approach to analyse leaf carbohydrate metabolism in Arabidopsis thaliana. EURASIP J Bioinform Syst Biol 2011:2. doi:10.1186/1687-4153-2011-2

Ito J, Parsons HT, Heazlewood JL (2014) The Arabidopsis cytosolic proteome: the metabolic heart of the cell Front. Plant Sci 5:21. doi:10.3389/fpls.2014.00021

Joyard J, Ferro M, Masselon C, Seigneurin-Berny D, Salvi D, Garin J, Rolland N (2010) Chloroplast proteomics highlights the subcellular compartmentation of lipid metabolism. Prog Lipid Res 49:128–158. doi:10.1016/j.plipres.2009.10.003

Kacser H, Burns JA (1973) The control of flux. Symp Soc Exp Biol 27:65–104

Kacser H, Burns JA, Kacser H, Fell DA (1995) The control of flux. Biochem Soc Trans 23:341–366. doi:10.1042/bst0230341

Klemens PA et al (2013) Overexpression of the vacuolar sugar carrier AtSWEET16 modifies germination, growth, and stress tolerance in Arabidopsis. Plant Physiol 163:1338–1352. doi:10.1104/pp.113.224972

Lemoine R et al (2013) Source-to-sink transport of sugar and regulation by environmental factors. Front Plant Sci 4:272. doi:10.3389/fpls.2013.00272

Linka N, Weber AP (2010) Intracellular metabolite transporters in plants Mol Plant 3:21–53. doi:10.1093/mp/ssp108

Lunn JE (2007) Compartmentation in plant metabolism. J Exp Bot 58:35–47. doi:10.1093/jxb/erl134

Millar AH, Taylor NL (2014) Subcellular proteomics-where cell biology meets protein chemistry. Front Plant Sci 5:55. doi:10.3389/fpls.2014.00055

Milne RJ, Byrt CS, Patrick JW, Grof CPL (2013) Are sucrose transporter expression profiles linked with patterns of biomass partitioning in Sorghum phenotypes? Front Plant Sci 4:223. doi:10.3389/fpls.2013.00223

Morgan JA, Rhodes D (2002) Mathematical modeling of plant metabolic pathways. Metab Eng 4:80–89. doi:10.1006/mben.2001.0211

Nägele T, Heyer AG (2013) Approximating subcellular organisation of carbohydrate metabolism during cold acclimation in different natural accessions of *Arabidopsis thaliana*. New Phytol 198:777–787. doi:10.1111/nph.12201

Nägele T, Weckwerth W (2013) A workflow for mathematical modeling of subcellular metabolic pathways in leaf metabolism of *Arabidopsis thaliana*. Front Plant Sci 4:541. doi:10.3389/fpls.2013.00541

Nägele T, Henkel S, Hörmiller I, Sauter T, Sawodny O, Ederer M, Heyer AG (2010) Mathematical modeling of the central carbohydrate metabolism in Arabidopsis reveals a substantial regulatory influence of vacuolar invertase on whole plant carbon metabolism. Plant Physiol 153:260–272. doi:10.1104/pp.110.154443

Nägele T, Stutz S, Hörmiller I, Heyer AG (2012) Identification of a metabolic bottleneck for cold acclimation in *Arabidopsis thaliana*. Plant J 72:102–114. doi:10.1111/j.1365-313X.2012.05064.x

Neuhaus HE, Trentmann O (2014) Regulation of transport processes across the tonoplast. Front Plant Sci 5:460. doi:10.3389/fpls.2014.00460

Paul MJ, Foyer CH (2001) Sink regulation of photosynthesis. J Exp Bot 52:1383–1400

Pokhilko A, Flis A, Sulpice R, Stitt M, Ebenhoh O (2014) Adjustment of carbon fluxes to light conditions regulates the daily turnover of starch in plants: a computational model. Mol BioSyst. doi:10.1039/c3mb70459a

Reznik E, Segrè D (2010) On the stability of metabolic cycles. J Theor Biol 266:536–549. doi:10.1016/j.jtbi.2010.07.023

Sonnewald U, Willmitzer L (1992) Molecular approaches to sink-source interactions. Plant Physiol 99:1267–1270

Steuer R (2007) Computational approaches to the topology, stability and dynamics of metabolic networks. Phytochemistry 68:2139–2151. doi:10.1016/j.phytochem.2007.04.041

Steuer R, Gross T, Selbig J, Blasius B (2006) Structural kinetic modeling of metabolic networks. Proc Natl Acad Sci USA 103:11868–11873. doi:10.1073/pnas.0600013103

Tiessen A, Padilla-Chacon D (2012) Subcellular compartmentation of sugar signaling: links among carbon cellular status, route of sucrolysis, sink-source allocation, and metabolic partitioning. Front Plant Sci 3:306. doi:10.3389/fpls.2012.00306

Turgeon R, Wolf S (2009) Phloem transport: cellular pathways and molecular trafficking. Annu Rev Plant Biol 60:207–221. doi:10.1146/annurev.arplant.043008.092045

Wingenter K et al (2010) Increased activity of the vacuolar monosaccharide transporter TMT1 alters cellular sugar partitioning, sugar signaling, and seed yield in Arabidopsis. Plant Physiol 154:665–677. doi:10.1104/pp.110.162040

Yamori W, Hikosaka K, Way DA (2014) Temperature response of photosynthesis in C3, C4, and CAM plants: temperature acclimation and temperature adaptation. Photosynth Res 119:101–117. doi:10.1007/s11120-013-9874-6

4. A Strategy for Functional Interpretation of Metabolomic Time Series Data in Context of Metabolic Network Information

Thomas Nägele*, **Lisa Fürtauer***, Matthias Nagler*, Jakob Weiszmann*
and Wolfram Weckwerth
(*these authors contributed equally)

Published in Frontiers in Molecular Biosciences (2016)
DOI: 10.3389/fmolb.2016.00006

Declaration of Authorship:

TN, **LF**, MN, JW: performed data analysis, wrote the source code, wrote the paper
TN & WW: conceived the study and wrote the paper

frontiers in Molecular Biosciences

ORIGINAL RESEARCH
published: 07 March 2016
doi: 10.3389/fmolb.2016.00006

A Strategy for Functional Interpretation of Metabolomic Time Series Data in Context of Metabolic Network Information

Thomas Nägele[1,2*†], *Lisa Fürtauer*[1†], *Matthias Nagler*[1†], *Jakob Weiszmann*[1†] and *Wolfram Weckwerth*[1,2]

[1] Department of Ecogenomics and Systems Biology, University of Vienna, Vienna, Austria, [2] Vienna Metabolomics Center, University of Vienna, Vienna, Austria

The functional connection of experimental metabolic time series data with biochemical network information is an important, yet complex, issue in systems biology. Frequently, experimental analysis of diurnal, circadian, or developmental dynamics of metabolism results in a comprehensive and multidimensional data matrix comprising information about metabolite concentrations, protein levels, and/or enzyme activities. While, irrespective of the type of organism, the experimental high-throughput analysis of the transcriptome, proteome, and metabolome has become a common part of many systems biological studies, functional data integration in a biochemical and physiological context is still challenging. Here, an approach is presented which addresses the functional connection of experimental time series data with biochemical network information which can be inferred, for example, from a metabolic network reconstruction. Based on a time-continuous and variance-weighted regression analysis of experimental data, metabolic functions, i.e., first-order derivatives of metabolite concentrations, were related to time-dependent changes in other biochemically relevant metabolic functions, i.e., second-order derivatives of metabolite concentrations. This finally revealed time points of perturbed dependencies in metabolic functions indicating a modified biochemical interaction. The approach was validated using previously published experimental data on a diurnal time course of metabolite levels, enzyme activities, and metabolic flux simulations. To support and ease the presented approach of functional time series analysis, a graphical user interface including a test data set and a manual is provided which can be run within the numerical software environment Matlab®.

Keywords: metabolic network, data integration, metabolomics, time series analysis, systems biology, network dynamics

OPEN ACCESS

Edited by:
Guowang Xu,
Chinese Academy of Sciences, China

Reviewed by:
Michal Jan Markuszewski,
Medical University of Gdansk, Poland
Huub C. J. Hoefsloot,
University of Amsterdam, Netherlands

***Correspondence:**
Thomas Nägele
thomas.naegele@univie.ac.at

† These authors have contributed equally to this work.

Specialty section:
This article was submitted to Metabolomics, a section of the journal Frontiers in Molecular Biosciences

Received: *11 December 2015*
Accepted: *19 February 2016*
Published: *07 March 2016*

Citation:
Nägele T, Fürtauer L, Nagler M, Weiszmann J and Weckwerth W (2016) A Strategy for Functional Interpretation of Metabolomic Time Series Data in Context of Metabolic Network Information. Front. Mol. Biosci. 3:6. doi: 10.3389/fmolb.2016.00006

INTRODUCTION

The functional interpretation of experimental data in context of biochemical network information represents one of the central challenges in current biological research. While genome sequencing projects have enabled the reconstruction of genome-scale metabolic networks, their high dimensionality precludes a direct and intuitive application to interpret experimental data. Hence, although genome sequence information and metabolic networks have become available for

numerous organisms, tissues, or cell types (Herrgard et al., 2008; Chang et al., 2011; De Oliveira Dal'Molin and Nielsen, 2013; Thiele et al., 2013), functional metabolic data interpretation still represents a major obstacle in systems biology. Various mathematical and computational strategies from the fields of multivariate statistics, ordinary, and partial differential equations (ODEs/PDEs), optimization or statistical time series analysis have been developed and applied to reveal a biologically meaningful interpretation of comprehensive and multidimensional experimental data sets. For example, a computational model of starch metabolism in plants enabled the analysis of starch kinetics through diurnal metabolic and circadian sensors (Pokhilko et al., 2014). The authors developed a model of 28 ODEs which were numerically simulated in order to analyze diurnal kinetics of carbon metabolism *in silico*. In another study, the response of *Escherichia coli* to varying oxygen concentrations was analyzed applying a mathematical model of the central metabolism (Ederer et al., 2014). Here, the authors derived a prediction about the impact of product formation on biomass concentration using steady state simulations at varying environmental conditions.

Both examples for mathematical modeling differ in organism and application. Besides, the dynamic approach can be distinguished from the steady state approach. However, in both approaches, dynamics of metabolic systems can be described by sets of ODEs. If sufficient kinetic information is available, such ODEs can be numerically integrated revealing simulated metabolic concentrations depending on time, enzyme parameters, thermodynamic constraints, etc. Yet, statistically robust experimental enzyme kinetic information often limits the applicability of such modeling approaches. Particularly, the resolution of enzyme activities, substrate affinities, or inhibitory constants is very laborious and only possible if well-established experimental assays and sufficient biochemical knowledge are available. Additionally, uncertainties about model structures and reaction kinetics complicate the interpretation of a numerically simulated output (Schaber et al., 2009). Such limitations have been addressed by different theoretical approaches, for example by structural kinetic modeling, SKM (Steuer et al., 2006). In the SKM approach, local linear models are applied to explore and statistically analyze a given parameter space without the need for explicit information about functional forms of enzyme kinetics and rate equations. Finally, a Jacobian matrix is derived which characterizes the dynamic capabilities of a metabolic system at a certain steady state. In previous publications, we have developed a procedure to determine Jacobian matrices directly from experimental metabolomics data (Nägele, 2014; Nägele et al., 2014). Based on experimental metabolic (co)variance information a metabolic regulator was identified indicating a strategy how plant metabolism is reprogrammed during exposure to energy limiting conditions. In a different context, other studies have also shown that it is possible to infer regulatory information about metabolic steady states from experimental data with such approaches (see e.g., Steuer et al., 2003; Sun and Weckwerth, 2012; Kügler and Yang, 2014).

Beyond these approaches of dynamic and steady state modeling, time series analysis and related regression models offer another mathematical strategy to reveal information about molecular system dynamics (Schelter, 2006). For example, Dutta and co-workers developed an algorithm for identification of differentially expressed genes in a time series experiment (Dutta et al., 2007), which they also applied to integrate transcriptome and metabolome data (Dutta et al., 2009). In another study, statistical modeling and regression analysis revealed a nitrogen-dependent modulation of root system architecture in the genetic model plant *Arabidopsis thaliana* (Araya et al., 2015). While these exemplarily mentioned studies present only a very small fraction of possible statistical applications, it already becomes evident that these are promising and necessary mathematical approaches to reveal biologically meaningful information from comprehensive experimental data sets being preliminary for hypothesis generation and experimental validation. However, a common problem of regression and correlation approaches in a biochemical context is a missing functional linkage of the results to causal biochemical interrelations, i.e., enzymatically driven reactions. To overcome this limitation and to facilitate the biochemical interpretation of the statistical results, the present study derives a theoretical connection between mathematical approaches of ODE-based dynamic modeling and statistical time series analysis. Based on the stoichiometric matrix information of a metabolic network, ratios of time-dependent derivative functions were built providing an estimate for the strength and probability of a metabolic interaction during the time course. The suggested strategy was tested using previously published experimental data sets on diurnal and stress-induced dynamics of metabolite concentrations and related enzyme kinetic information. Finally, a graphical user interface for Matlab is provided which intends to facilitate the application of the presented strategy.

RESULTS

Deriving Metabolic Functions by Inverse Variance-Weighted Regression Analysis

Time-dependent dynamics of metabolite concentrations in a biochemical network can be described by a set of ODEs:

$$\frac{d}{dt}\boldsymbol{M}(t) = \boldsymbol{N}\boldsymbol{v}\left(\boldsymbol{M},\boldsymbol{p},t\right) = f(\boldsymbol{M},\boldsymbol{p},t) \tag{1}$$

Here, $\boldsymbol{M}$ represents an n-dimensional vector of mean metabolite concentrations (c_n), $\boldsymbol{N}$ is the $n \times k$ stoichiometric matrix and $\boldsymbol{v}$ describes the k-dimensional vector of reaction rates which depend on metabolite concentrations $\boldsymbol{M}$, enzyme kinetic parameters $\boldsymbol{p}$ and time t. The right side of the ODE system can be summarized by metabolic functions, $f(\boldsymbol{M},\boldsymbol{p},t)$. Hence, these metabolic functions define the time-dependent changes in metabolite concentration as a sum of all biochemical reactions either consuming or producing a metabolite. A metabolic steady state is described by ODEs which equal zero:

$$\frac{d}{dt}\boldsymbol{M}(t) = 0 \tag{2}$$

Linearization enables the investigation of the system behavior close to a steady state. The linearization process results in the so-called Jacobian matrix $\boldsymbol{J}$ which characterizes the dynamic properties of the system at a steady state:

$$J = \begin{pmatrix} \frac{\partial f_1}{\partial c_1} & \cdots & \frac{\partial f_1}{\partial c_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_n}{\partial c_1} & \cdots & \frac{\partial f_n}{\partial c_n} \end{pmatrix} \quad (3)$$

Hence, in a biochemical context, the Jacobian matrix $\boldsymbol{J}$ describes the behavior of metabolic functions f_i (for $i = 1,...,n$) of a metabolic network with regard to small changes of variables c_i (for $i = 1,...,n$), i.e., metabolite concentrations at the considered steady state. The information if a metabolic function f_i biochemically depends on the concentration of a metabolite c_i is provided by the stoichiometric matrix $\boldsymbol{N}$ of a metabolic network (see Equation 1).

To derive, i.e., predict, time continuous information from time discrete experimental observations, interpolation methods can be applied. To prevent unrealistic oscillations of high-degree polynomial interpolation, intervals of approximation can be partitioned in subintervals which can be approximated, for example, by cubic polynomials which form a cubic spline $S_{c_i}(t)$ (see e.g., Bronstein et al., 2008):

$$S_{c_i}(t) = \alpha_{ij} + \beta_{ij}\left(t - t_j\right) + \gamma_{ij}\left(t - t_j\right)^2 + \delta_{ij}(t - t_j)^3 \quad (4)$$

Here, it is $t \in [t_j, t_{j+1}]$ with $(j = 1, 2,..., z-1)$, where z represents the number of interpolation nodes (t_j, g_{ij}), and it is $S_{c_i}\left(t_j\right) = g_{ij}$. Interpolation coefficients are represented by α, β, γ, and δ. Due to the occurrence of experimental errors, the requirement of $S_{c_i}\left(t_j\right) = g_{ij}$ is not fulfilled which prevents the suitability of such a type of interpolation. Instead, a smoothing element can be introduced accounting for those experimental errors:

$$min\left(\sum_{j=1}^{z} w_{ij}[g_{ij} - S_{c_i}(t_j)]^2 + \lambda \int_{t=t_1}^{t_z} \left[S_{c_i}''(t)\right]^2 dt\right) \quad (5)$$

Here, w_{ij} represents a weighting factor, $S_{c_i}''(t)$ is the second derivative of $S_{c_i}(t)$, and λ (with $\lambda \geq 0$) represents a smoothing factor. For $\lambda = 0$, one obtains the cubic spline interpolation, while the degree of smoothing increases with the value of λ. To connect the smoothing spline generation to experimentally observed errors we defined the weighting factor w_{ij} to equal the inverse variance information, i.e., the inverse squared standard deviation:

$$w_{ij} = {\sigma_{ij}}^{-2} = \left(\frac{1}{r-1}\sum_{k=1}^{r}(c_{ij,k} - \bar{c}_{ij})^2\right)^{-1} \quad (6)$$

Here, r represents the number of experimental and independent replicates.

Merging Equations (1), (5) and (6) and replacing g_{ij} by the mean concentration of metabolite i at time point j, $\bar{c}_{ij}$, reveals a description of metabolic functions by inverse variance-weighted regression analysis:

$$\begin{aligned} \frac{d}{dt}M_i(t) &= f_i\left(\boldsymbol{M},\boldsymbol{p},t\right) \\ &= \frac{d}{dt}\left[min\left(\sum_{j=1}^{z}\left[\left(\frac{1}{r-1}\sum_{k=1}^{r}(c_{ij,k} - \bar{c}_{ij})^2\right)^{-1}\right.\right.\right. \\ &\quad \left.\left.\left.[\bar{c}_{ij} - S_{c_i}(t_j)]^2\right] + \lambda \int_{t=t_1}^{t_z}\left[S_{c_i}''(t)\right]^2 dt\right)\right] \quad (7) \end{aligned}$$

Hence, building the first derivative of the smoothed interpolation of experimental time-course data reveals information about the connected metabolic function. In the present study, this approach was applied to evaluate a diurnal time course of previously published metabolite concentrations (Nägele et al., 2012) belonging to the central carbohydrate metabolism in leaves of the genetic model plant *A. thaliana*. Diurnal dynamics of metabolic functions are shown exemplarily (**Figure 1**) for the metabolite pools of sucrose (Suc) and sugar phosphates (SP) in a control experiment (non-cold acclimated, na) and after exposure to low temperature (acc).

To characterize time-dependent changes in metabolic functions, the second time-dependent derivative was built from the approximated diurnal time course of metabolite concentrations:

$$\begin{aligned} \frac{d^2}{dt^2}M_i(t) &= \frac{d}{dt}f_i\left(\boldsymbol{M},\boldsymbol{p},t\right) \\ &= \frac{d^2}{dt^2}\left[min\left(\sum_{j=1}^{z}\left[\left(\frac{1}{r-1}\sum_{k=1}^{r}(c_{ij,k} - \bar{c}_{ij})^2\right)^{-1}\right.\right.\right. \\ &\quad \left.\left.\left.[\bar{c}_{ij} - S_{c_i}(t_j)]^2\right] + \lambda \int_{t=t_1}^{t_z}\left[S_{c_i}''(t)\right]^2 dt\right)\right] \quad (8) \end{aligned}$$

As described for **Figure 1**, diurnal dynamics of those time-dependent changes of metabolic functions are also shown exemplarily for Suc and SP (**Figure 2**).

Connecting Metabolic Functions Based on Biochemical Network Information

While the metabolic time-course information derived before characterizes time-dependent rates of changes in each considered metabolite concentration separately (see Equations 7 and 8), information of biochemical interdependencies, i.e., the information about a substrate - product relationship between two or more metabolites, is only contained implicitly. To explicitly analyze and visualize these biochemical interdependencies with regard to the time-dependent rates of concentrations changes, a metabolic $n \times n$ interaction matrix, $\boldsymbol{Y}$, was derived where n represents the number of metabolites comprised by the model. In $\boldsymbol{Y}$, each entry indicates whether two metabolites are biochemically connected (entry: 1) or not (entry: 0). The interaction is characterized analogous to entries of the Jacobian

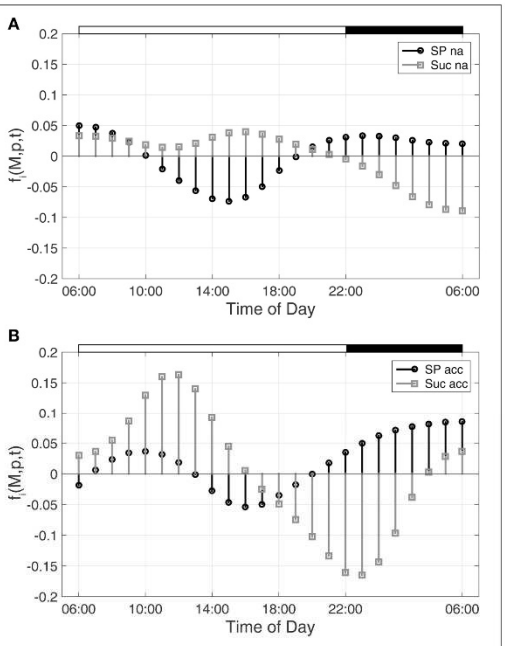


FIGURE 1 | Metabolic functions derived from inverse variance-weighted regression analysis for (A) non-cold acclimated and (B) cold acclimated plants. Metabolic functions $f_i(\boldsymbol{M},\boldsymbol{p},t)$ were derived for sugar phosphates (SP) and sucrose (Suc) as described in the text (see Equation 7). Experimental data were derived from a previous study comprising metabolite levels of non-cold acclimated (na) and cold acclimated (acc) leaf material of *Arabidopsis thaliana*, accession Col-0 (Nägele et al., 2012). White and black bars on the top indicate light and dark phase of a diurnal cycle.

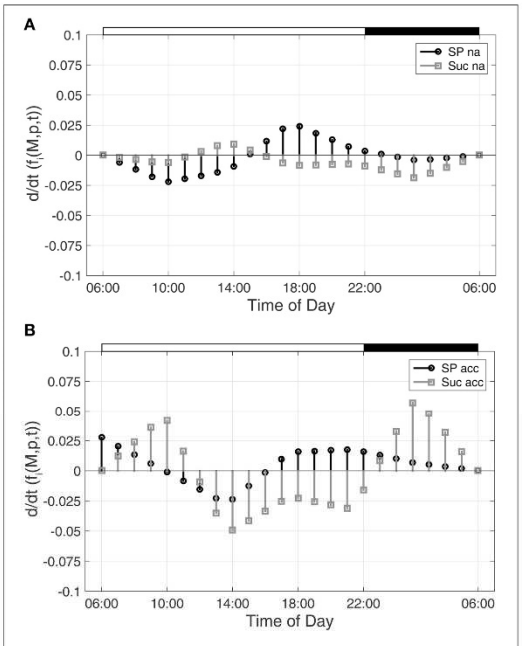


FIGURE 2 | Time-dependent dynamics of metabolic functions derived from inverse variance-weighted regression analysis for (A) non-cold acclimated and (B) cold acclimated plants. Function dynamics were derived for sugar phosphates (SP) and sucrose (Suc) as described in the text (see Equation 8). Experimental data were derived from a previous study comprising metabolite levels of non-cold acclimated (na) and cold acclimated (acc) leaf material of *A. thaliana*, accession Col-0 (Nägele et al., 2012). White and black bars on the top indicate light and dark phase of a diurnal cycle.

matrix (Equation 3): if the metabolic function of metabolite A is biochemically connected to changes in concentrations of metabolite B, the corresponding entry in $\boldsymbol{Y}$ is 1. Information about metabolic functions is given row-wise, while biochemically connected metabolites are indicated column-wise for each function. In a simple example, containing three reactions (r_1–r_3) and four metabolites (A–D), the construction and content of $\boldsymbol{Y}$ is exemplified (**Figure 3**). The diagonal entries indicate the biochemical dependencies of metabolic functions on substrate concentrations. For example, $\boldsymbol{Y}_{11} = 1$ indicates that metabolic function f(A,t) depends on the concentration of A(t). The non-diagonal entries describe interdependencies between different metabolite pools. For example, $\boldsymbol{Y}_{21} = 1$ indicates that metabolic function f(B,t) depends on the concentration of A(t).

Based on a previously published metabolic model (Nägele et al., 2012), an interaction matrix $\boldsymbol{Y}$ was derived for the central carbohydrate metabolism in leaves of *A. thaliana*. The metabolic functions (Equation 7) and their time-dependent derivatives (Equation 8) were related to each other according to the entries of $\boldsymbol{Y}$. This finally resulted in functions $\omega(a \rightarrow b, t)$ indicating changes in metabolic functions of b in context of concentration changes of a which might represent substrates, inhibitors or activators:

$$\omega\left(a \rightarrow b, t\right) = \frac{\frac{d}{dt} f_b(\boldsymbol{M},\boldsymbol{p}, t)}{f_a(\boldsymbol{M},\boldsymbol{p}, t)},\ D = \left\{\mathbb{R} \setminus f_a\left(\boldsymbol{M},\boldsymbol{p}, t\right) = 0\right\} \quad (9)$$

With regard to the analyzed time-course of sugar phosphate (SP) and sucrose (Suc) concentrations (see **Figures 1**, **2**), ω(SP→Suc, t) revealed information about the reaction of sucrose biosynthesis, catalyzed by the enzyme sucrose phosphate synthase (SPS):

$$SugarPhosphates\,(SP) \overset{SPS}{\rightarrow} Sucrose \quad (10)$$

In detail, ω(SP→Suc, t) described changes in the metabolic function of sucrose in context of concentration changes of its biochemical substrate sugar phosphates:

$$\omega\,(SP \rightarrow Suc, t) = \frac{\frac{d}{dt} f_{Suc}(\boldsymbol{M},\boldsymbol{p}, t)}{f_{SP}(\boldsymbol{M},\boldsymbol{p}, t)} \quad (11)$$

Comparing ω(SP→Suc, t) for na and acc plants revealed a noticeable difference between both conditions within the first

4 h of the day (**Figure 4**). Interestingly, in the same time period, simulations of sucrose biosynthesis, based on a system of ordinary differential equations (ODEs), revealed a similar picture in which rates of sucrose biosynthesis were decreased only in acc plants (Nägele et al., 2012).

Characterization of ω(t)

$\omega(t)$ represents a ratio of metabolic functions and related derivatives. Hence, the unit of ω(t) is derived from the flux unit of a metabolic function, [mM s^{-1}]:

$$\left[\omega\left(a \rightarrow b, t\right)\right] = \frac{\left[\frac{d}{dt} f_b(\boldsymbol{M},\boldsymbol{p}, t)\right]}{\left[f_a(\boldsymbol{M},\boldsymbol{p}, t)\right]} = \frac{\left[mM\ s^{-2}\right]}{\left[mM\ s^{-1}\right]} = \left[\frac{1}{s}\right] \quad (12)$$

Here, concentrations are given in mM (mmol l^{-1}) and the time unit is seconds (s).

This results in the unit of a rate or frequency. Hence, $|\omega(a \rightarrow b,t)|$ was interpreted as oscillations of a metabolic function per time-period with reference to a biochemical effector.

In the case of $|\omega(a \rightarrow b,\ t)| \rightarrow \infty$ for $t \rightarrow \tau$, the influence of the biochemical effector on a metabolic function was defined to be strong, while $|\omega(a \rightarrow b,\ t)| \rightarrow 0$ for $t \rightarrow \tau$ indicated a weak effect. In detail, $|\omega(a \rightarrow b,\ t)| \rightarrow \infty$ for $t \rightarrow \tau$ indicates that it is $|d/dt\ (f_b(\boldsymbol{M},\boldsymbol{p},t))| >> |f_a(\boldsymbol{M},\boldsymbol{p},t)|$. *Vice versa*, $|\omega(a \rightarrow b,\ t)| \rightarrow 0$ for $t \rightarrow \tau$ indicates that $|d/dt\ (f_b(\boldsymbol{M},\boldsymbol{p},t))| << |f_a(\boldsymbol{M},\boldsymbol{p},t)|$.

Application Example: Stress-Induced Metabolic Reprogramming in *Arabidopsis thaliana*

While in the above mentioned example, the calculation and interpretation of ω(t) was demonstrated in context of a previously published kinetic ODE model, another published data set was analyzed by this strategy comprising metabolite levels of the primary and secondary metabolism in *A. thaliana* (Doerfler et al., 2013). In the experiment performed by Doerfler and co-workers, a combined strategy of gas chromatography and liquid chromatography coupled to mass spectrometry was applied in order to reveal a comprehensive picture of metabolic reprogramming during exposure to low temperature and high light intensity. The time period of stress exposure comprised more than 2 weeks which allowed for the analysis of a short- and long-term acclimation response in the metabolome of *A. thaliana,* accession Col-0. A central output of the study was the characterization of metabolomic and regulatory dynamics at the interface of primary and secondary metabolism. The authors observed a fast increase of stress-responsive compounds, e.g., sucrose, which became significant already after 2 days of stress exposure, while the interaction with the secondary metabolism, resulting in biosynthesis of flavonoids, became most significant after 8 days of stress exposure.

To prove the suitability of deriving the absolute value function $|\omega(a \rightarrow b,\ t)|$ in order to reveal steps of metabolic regulation within a considered time interval, regression analysis and metabolic functions were calculated for the dataset of Doerfler et al. (2013) and compared to their observations. The metabolic interaction matrix $\boldsymbol{Y}$ was derived from the metabolic network model which was previously suggested and applied for inverse approximations of the Jacobian matrix (Doerfler et al., 2013). For regression analysis and for integration of metabolic network information we developed and applied a graphical user interface (*FEMTO*, ***F***unctional ***E***valuation of ***M***etabolic ***T***ime series ***O***bservations) which is based on the numerical software environment Matlab® (http://www.mathworks.com), and which is provided in the supplements together with a user manual (**Supplementary Files S1**, **S2**).

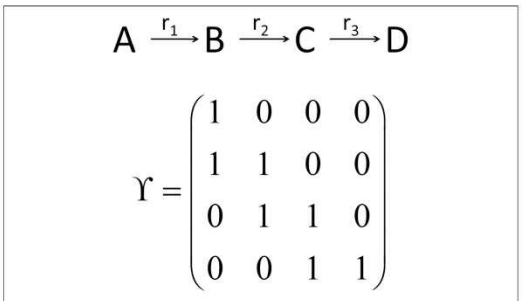


FIGURE 3 | Schematic reaction chain and the derived interaction matrix Y. Rows (metabolic functions) and columns (metabolites) of **Y** describe biochemical interactions of metabolites A (first row/column), B (second row/column), C (third row/column), and D (fourth row/column). Entries, i.e., 0 and 1, indicate if two metabolites interact (entry 1) or not (entry 0).

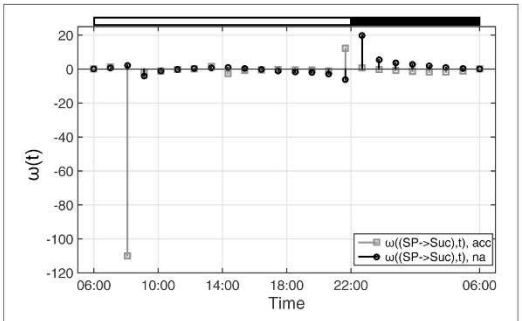


FIGURE 4 | Functions ω(t) indicating changes in metabolic functions in context of concentration changes of biochemical interaction partners. ω(t) was calculated as described in the main text (see Equation 9). Results of ω(t) are shown for the reaction of sucrose biosynthesis for non-cold acclimated (na) and cold acclimated (acc) leaf material of *A. thaliana,* accession Col-0 (Nägele et al., 2012). White and black bars on the top indicate light and dark phase of a diurnal cycle.

To characterize sucrose metabolism, changes of the metabolic function of sucrose were related to changes in sucrose concentrations:

$$|\omega\ (Suc \rightarrow Suc, t)| = \left|\frac{\frac{d}{dt} f_{Suc}(\boldsymbol{M},\boldsymbol{p}, t)}{f_{Suc}(\boldsymbol{M},\boldsymbol{p},\ t)}\right| \quad (13)$$

For time-dependent characterization of flavonoid dynamics, changes in the flavonoid (Flav) metabolic function were related to substrate concentration changes, i.e., phenylalanine (Phe) dynamics:

$$\left|\omega\left(Phe \rightarrow Flav, t\right)\right| = \left|\frac{\frac{d}{dt}f_{Flav}(\boldsymbol{M},\boldsymbol{p}, t)}{f_{Phe}(\boldsymbol{M},\boldsymbol{p}, t)}\right| \quad (14)$$

Results of metabolic function analysis and the resulting time course of $|\omega(t)|$ revealed an early de-regulation of sucrose metabolism during the first 2 days of stress exposure (**Figure 5A**), while the peak value of $|\omega(t)|$ for flavonoid biosynthesis occurred delayed after 8 days (**Figure 5B**).

These findings coincide with the previous findings described by Doerfler and colleagues who applied the method of Granger causality time-series correlation and a covariance-based inverse approximation of Jacobian matrices to reveal strategies of metabolic regulation (Doerfler et al., 2013). Conclusions which have been drawn from the $|\omega(t)|$ calculation were found to be highly similar to the output of other statistical methods, finally substantiating the validity of the suggested workflow and the derived method to unravel time points of regulatory perturbation in a biochemical system.

DISCUSSION

Mathematical analysis of biochemical system dynamics represents a central focus of current biomathematical, biochemical and biotechnological research due to the need for methods and algorithms enabling a functional interpretation of experimental data in context of a biochemical network. Particularly, system dynamics which arise due to circadian regulation (Harmer, 2009; Kumar Jha et al., 2015), diurnal metabolic adjustment (Geiger and Servaites, 1994; Pokhilko et al., 2014) or stress-induced metabolic reprogramming (Jozefczuk et al., 2010; Kanshin et al., 2015) are hardly traceable by intuition. Hence, this indicates a strong need for suitable theoretical approaches being capable of resolving and functionally connecting molecular moieties with underlying biochemical regulation.

Various theoretical strategies have addressed this complex issue, providing a comprehensive methodological platform for time-series analysis, dynamic flux balance analysis, kinetic and Boolean modeling (see e.g., Mahadevan et al., 2002; Schelter, 2006; Rohwer, 2012; Steinway et al., 2015). In a recent approach, Willemsen and colleagues have modified the approach of dynamic flux balance analysis by incorporating time-resolved metabolomics measurements (Willemsen et al., 2015). With their extended method, the authors derived an estimate of dynamic flux profiles which allowed them to generate and test hypotheses related to environmentally induced molecular dynamics. In another recent study, a computational approach was suggested to translate metabolomics data into flux information (Cortassa et al., 2015). One main methodological difference between the studies of Willemsen et al. and Cortassa et al. was the extent of kinetic information which was needed to estimate cellular behavior and metabolic fluxes. While Willemsen et al. focused on minimalistic kinetic information, the study of Cortassa and co-workers used a detailed kinetic model of glucose catabolic pathways to derive flux information.

In our presented approach, flux information, which was implicitly derived from spline interpolation, was interpreted only indirectly by comparing time-dependent changes in metabolic functions to concentration changes of biochemical reaction partners. This procedure revealed information about a rate which was interpreted in terms of metabolic functions related to concentration changes in a substrate or co-substrate. Comparing derived results to other methods, it was shown that changes in ratios of second- to first-order derivatives between functionally connected variables potentially reveal time points of regulatory perturbation within a biochemical interaction. Hence, these observed perturbations might indicate a change in enzymatic activity, protein abundance, or allosteric regulation ultimately leading to a change in the metabolic functions.

The information content of the introduced time-dependent functions $\omega(t)$ is related to entries of the Jacobian matrix $\boldsymbol{J}$ (see Equation 3) indicating the dynamics of metabolic functions with respect to (small) concentration changes at a certain steady state. This theoretical connection of $\boldsymbol{J}$ and $\omega(t)$ at a considered time point t_0 might be illustrated in a simple first-order reaction scheme.

$$A \xrightarrow{k} B \quad (15)$$

Here, substance A is interconverted into substance B, and the reaction velocity is characterized by the rate constant k. The time-dependent change in concentration of A equals $dA/dt = -k{\cdot}A$. Hence, a general solution of this ODE is given by $A(t) = A_0 e^{-\mathrm{kt}}$ which finally yields $J_{11}(t_0) = \omega(A \rightarrow A, t_0) = -\mathrm{k}$.

With this, the information of $\omega(t)$ becomes comparable to entries of the Jacobian matrix $\boldsymbol{J}$. Yet, in contrast to entries of $\boldsymbol{J}$, characterizing dynamic properties of a metabolic steady state ($d/dt\ \boldsymbol{M}(t) = 0$), functions $\omega(t)$ were derived from a time series of experimental data and might rather be valid for a non-infinitesimal than for an infinitesimal time frame. While for $\lim_{t \rightarrow t_0} \left|\omega(t)\right|$, $|\omega(t)|$ might be assumed to approach entries of $\boldsymbol{J}$, this was not tested in the present study and would need experimental validation. In addition, while a connection, and probable correlation, to other molecular levels, such as the proteome or transcriptome, was not experimentally analyzed, this might be a promising target for analysis in future studies. However, the incorporation of an interaction matrix, which, in the present study, was derived from a previously published reaction network, and which might be derived from genome-scale metabolic reconstruction works in future studies (Weckwerth, 2011; King et al., 2015), provides direct evidence for the biochemical and physiological relevance of the performed theoretical analysis.

While our results indicate a realistic and biochemically interpretable output of the presented method, limitations of application might occur due to several reasons. First, the presented method significantly depends on the knowledge about the biochemical network structure and involved regulatory interactions, e.g., feedback inhibition or feedforward activation. Although regression analysis of time series data might be

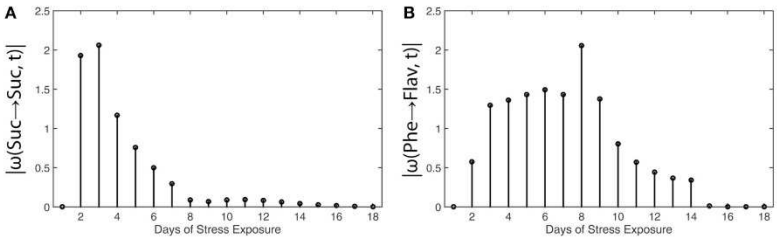


FIGURE 5 | Absolute value functions of ω(t) for (A) sucrose metabolism and (B) flavonoid biosynthesis in leaves of *Arabidopsis thaliana*. Abscissae indicate days of exposure to low temperature and high-light stress conditions. Detailed information about the calculation is provided in the main text (see Equations 13 and 14). Experimental data were derived from a previous study (Doerfler et al., 2013).

performed for all network components independently, deriving a reliable biochemical interaction matrix $\boldsymbol{Y}$ is essential to reveal realistic information about time-dependent changes in metabolic interactions. A second central prerequisite for a meaningful regression analysis is the design of an adequate experimental setup. This comprises the number of biological (independent) replicates as well as the number and interval of sampling points. It is hardly possible to generalize a number of replicates or sampling points due to heterogeneous technical or environmental fluctuations which are introduced by different analytical techniques, growth conditions or sample types. Yet, spanning various experimental scenarios, it might be generalized that the interval of sampling points is crucial to be able to discriminate between metabolic fast or short-term responses and slow or long-term responses. Particularly to resolve fast metabolic regulation, a narrow sampling interval is needed in order to prevent any over-interpretation of regression analysis and related derivatives. Comparing the presented approach to methods of metabolic modeling, a third major limitation is the missing predictive output by model simulations. For example, enzyme kinetic models of metabolism aim at going beyond the time interval of measured rate constants or metabolite concentrations to predict changes in system dynamics under changing environmental conditions or due to a mutated gene. However, although our presented method cannot afford this simulation output, time-dependent changes within the considered time interval might indicate regulatory bottlenecks and kinetic information supporting the numerical solution and simulation of metabolic ODE models.

In summary, the suggested approach intends to promote the functional interpretability of metabolic time series data in context of metabolic network information. Particularly with regard to multidimensional metabolomics data sets, this might unravel strategies of complex biochemical regulation and might overcome some limitations in the generation of testable hypotheses as we have discussed previously (Nägele and Weckwerth, 2012). Finally, the direct integration of biochemical network information with experimental data promises to enable the functional interpretation and the causal connection of various levels of molecular organization.

MATERIALS AND METHODS

The described procedure of data analysis, spline interpolation and graphical representation was performed within the numerical software environment Matlab®. A Matlab-based graphical user interface (FEMTO, *F*unctional *E*valuation of *M*etabolic *T*ime series *O*bservations) was developed and is provided, together with a user manual, in the supplements (**Supplementary Files S1**, **S2**).

AUTHOR CONTRIBUTIONS

TN, LF, MN, and JW performed data analysis, statistics, wrote the source code of the graphical user interface and wrote the paper. TN and WW conceived the study and wrote the paper. All authors read and approved the final version of the manuscript.

FUNDING

This work was supported by the Austrian Science Fund, FWF, Project P 26342 and Project I 2071.

ACKNOWLEDGMENTS

We would like to thank the members of the MoSys Department for constructive advice and fruitful discussions. Further, we thank Arnd G. Heyer from the Department of Plant Biotechnology at the University of Stuttgart, Germany, for providing experimental data.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: http://journal.frontiersin.org/article/10.3389/fmolb.2016.00006

Supplementary File S1 | Matlab code-file and test data file.

Supplementary File S2 | Manual for the graphical user interface FEMTO 1.0.

REFERENCES

Araya, T., Kubo, T., von Wirén, N., and Takahashi, H. (2015). Statistical modeling of nitrogen-dependent modulation of root system architecture in *Arabidopsis thaliana*. *J. Integr. Plant. Biol.* doi: 10.1111/jipb.12433. [Epub ahead of print].

Bronstein, I. J. N., Semendjajew, K. A., Musiol, G., and Mühlig, H. (2008). *Taschenbuch der Mathematik*. Frankfurt am Main: Wissenschaftlicher Verlag Harri Deutsch GmbH.

Chang, R. L., Ghamsari, L., Manichaikul, A., Hom, E. F., Balaji, S., Fu, W., et al. (2011). Metabolic network reconstruction of Chlamydomonas offers insight into light-driven algal metabolism. *Mol. Syst. Biol.* 7, 518. doi: 10.1038/msb.2011.52

Cortassa, S., Caceres, V., Bell, L. N., O'Rourke, B., Paolocci, N., and Aon, M. A. (2015). From metabolomics to fluxomics: a computational procedure to translate metabolite profiles into metabolic fluxes. *Biophys. J.* 108, 163–172. doi: 10.1016/j.bpj.2014.11.1857

De Oliveira Dal'Molin, C. G., and Nielsen, L. K. (2013). Plant genome-scale metabolic reconstruction and modelling. *Curr. Opin. Biotechnol.* 24, 271–277. doi: 10.1016/j.copbio.2012.08.007

Doerfler, H., Lyon, D., Nägele, T., Sun, X., Fragner, L., Hadacek, F., et al. (2013). Granger causality in integrated GC-MS and LC-MS metabolomics data reveals the interface of primary and secondary metabolism. *Metabolomics* 9, 564–574. doi: 10.1007/s11306-012-0470-0

Dutta, B., Kanani, H., Quackenbush, J., and Klapa, M. I. (2009). Time-series integrated "omic" analyses to elucidate short-term stress-induced responses in plant liquid cultures. *Biotechnol. Bioeng.* 102, 264–279. doi: 10.1002/bit.22036

Dutta, B., Snyder, R., and Klapa, M. I. (2007). Significance analysis of time-series transcriptomic data: a methodology that enables the identification and further exploration of the differentially expressed genes at each time-point. *Biotechnol. Bioeng.* 98, 668–678. doi: 10.1002/bit.21432

Ederer, M., Steinsiek, S., Stagge, S., Rolfe, M. D., Ter Beek, A., Knies, D., et al. (2014). A mathematical model of metabolism and regulation provides a systems-level view of how *Escherichia coli* responds to oxygen. *Front. Microbiol.* 5:124. doi: 10.3389/fmicb.2014.00124

Geiger, D. R., and Servaites, J. C. (1994). Diurnal regulation of photosynthetic carbon metabolism in C3 plants. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* 45, 235–256. doi: 10.1146/annurev.pp.45.060194.001315

Harmer, S. L. (2009). The circadian system in higher plants. *Annu. Rev. Plant Biol.* 60, 357–377. doi: 10.1146/annurev.arplant.043008.092054

Herrgard, M. J., Swainston, N., Dobson, P., Dunn, W. B., Arga, K. Y., Arvas, M., et al. (2008). A consensus yeast metabolic network reconstruction obtained from a community approach to systems biology. *Nat. Biotechnol.* 26, 1155–1160. doi: 10.1038/nbt1492

Jozefczuk, S., Klie, S., Catchpole, G., Szymanski, J., Cuadros-Inostroza, A., Steinhauser, D., et al. (2010). Metabolomic and transcriptomic stress response of *Escherichia coli*. *Mol. Syst. Biol.* 6, 364. doi: 10.1038/msb.2010.18

Kanshin, E., Kubiniok, P., Thattikota, Y., D'Amours, D., and Thibault, P. (2015). Phosphoproteome dynamics of *Saccharomyces cerevisiae* under heat shock and cold stress. *Mol. Syst. Biol.* 11, 813. doi: 10.15252/msb.20156170

King, Z. A., Lloyd, C. J., Feist, A. M., and Palsson, B. O. (2015). Next-generation genome-scale models for metabolic engineering. *Curr. Opin. Biotechnol.* 35, 23–29. doi: 10.1016/j.copbio.2014.12.016

Kügler, P., and Yang, W. (2014). Identification of alterations in the Jacobian of biochemical reaction networks from steady state covariance data at two conditions. *J. Math. Biol.* 68, 1757–1783. doi: 10.1007/s00285-013-0685-3

Kumar Jha, P., Challet, E., and Kalsbeek, A. (2015). Circadian rhythms in glucose and lipid metabolism in nocturnal and diurnal mammals. *Mol. Cell Endocrinol.* *418*(Pt 1), 74–88. doi: 10.1016/j.mce.2015.01.024

Mahadevan, R., Edwards, J. S., and Doyle, F. J. III (2002). Dynamic flux balance analysis of diauxic growth in *Escherichia coli*. *Biophys. J.* 83, 1331–1340. doi: 10.1016/S0006-3495(02)73903-9

Nägele, T. (2014). Linking metabolomics data to underlying metabolic regulation. *Front. Mol. Biosci.* 1:22. doi: 10.3389/fmolb.2014.00022

Nägele, T., Mair, A., Sun, X., Fragner, L., Teige, M., and Weckwerth, W. (2014). Solving the differential biochemical Jacobian from metabolomics covariance data. *PLoS ONE* 9:e92299. doi: 10.1371/journal.pone.0092299

Nägele, T., Stutz, S., Hörmiller, I., and Heyer, A. G. (2012). Identification of a metabolic bottleneck for cold acclimation in *Arabidopsis thaliana*. *Plant J.* 72, 102–114. doi: 10.1111/j.1365-313X.2012.05064.x

Nägele, T., and Weckwerth, W. (2012). Mathematical modeling of plant metabolism—from reconstruction to prediction. *Metabolites* 2, 553–566. doi: 10.3390/metabo2030553

Pokhilko, A., Flis, A., Sulpice, R., Stitt, M., and Ebenhöh, O. (2014). Adjustment of carbon fluxes to light conditions regulates the daily turnover of starch in plants: a computational model. *Mol. Biosyst.* 10, 613–627. doi: 10.1039/c3mb70459a

Rohwer, J. M. (2012). Kinetic modelling of plant metabolic pathways. *J. Exp. Bot.* 63, 2275–2292. doi: 10.1093/jxb/ers080

Schaber, J., Liebermeister, W., and Klipp, E. (2009). Nested uncertainties in biochemical models. *IET Syst. Biol.* 3, 1–9. doi: 10.1049/iet-syb:20070042

Schelter, B. (2006). *Handbook of Time Series Analysis: Recent Theoretical Developments and Applications*. Weinheim: WILEY-VCH.

Steinway, S. N., Biggs, M. B., Loughran, T. P. Jr., Papin, J. A., and Albert, R. (2015). Inference of network dynamics and metabolic interactions in the gut microbiome. *PLoS Comput. Biol.* 11:e1004338. doi: 10.1371/journal.pcbi.1004338

Steuer, R., Gross, T., Selbig, J., and Blasius, B. (2006). Structural kinetic modeling of metabolic networks. *Proc. Natl. Acad. Sci. U.S.A.* 103, 11868–11873. doi: 10.1073/pnas.0600013103

Steuer, R., Kurths, J., Fiehn, O., and Weckwerth, W. (2003). Observing and interpreting correlations in metabolomic networks. *Bioinformatics* 19, 1019–1026. doi: 10.1093/bioinformatics/btg120

Sun, X. L., and Weckwerth, W. (2012). COVAIN: a toolbox for uni- and multivariate statistics, time-series and correlation network analysis and inverse estimation of the differential Jacobian from metabolomics covariance data. *Metabolomics* 8, S81–S93. doi: 10.1007/s11306-012-0399-3

Thiele, I., Swainston, N., Fleming, R. M., Hoppe, A., Sahoo, S., Aurich, M. K., et al. (2013). A community-driven global reconstruction of human metabolism. *Nat. Biotechnol.* 31, 419–425. doi: 10.1038/nbt.2488

Weckwerth, W. (2011). Unpredictability of metabolism—the key role of metabolomics science in combination with next-generation genome sequencing. *Anal. Bioanal. Chem.* 400, 1967–1978. doi: 10.1007/s00216-011-4948-9

Willemsen, A. M., Hendrickx, D. M., Hoefsloot, H. C., Hendriks, M. M., Wahl, S. A., Teusink, B., et al. (2015). MetDFBA: incorporating time-resolved metabolomics measurements into dynamic flux balance analysis. *Mol. Biosyst.* 11, 137–145. doi: 10.1039/C4MB00510D

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

5. Combined multivariate analysis and machine learning reveals a predictive module of metabolic stress response in Arabidopsis thaliana

Lisa Fürtauer*, Alice Pschenitschnigg*, Helene Scharkosi1*, Wolfram Weckwerth and Thomas Nägele

(*these authors contributed equally)

Under Review (20.04.2018): Molecular Omics

Declaration of Authorship:

LF, AP, HS: performed experiments, data evaluation, wrote the paper

WW: revised the paper

TN: designed the study, performed experiments, data evaluation, wrote the paper

Combined multivariate analysis and machine learning reveals a predictive module of metabolic stress response in *Arabidopsis thaliana*

Authors and Affiliations:

Lisa Fürtauer[1,*], Alice Pschenitschnigg[1,*], Helene Scharkosi[1,*], Wolfram Weckwerth[1,2] and Thomas Nägele[1,2,3,§]

[1] Department of Ecogenomics and Systems Biology, University of Vienna, Vienna, Austria

[2] Vienna Metabolomics Center, University of Vienna, Vienna, Austria

[3] Ludwig-Maximilians-Universität München, Department Biology I, Planegg-Martinsried, Germany

[*] These authors contributed equally

[§] Corresponding Author

Thomas Nägele
Biozentrum der LMU München
Department Biology I
Großhadernerstr. 2-4
D-82152 Planegg-Martinsried
Germany
T.: +49-89-2180-74660
F.: +49-89-2180-74661

Email: Thomas.Naegele@lmu.de

Key words: abiotic stress; *Arabidopsis thaliana*; metabolic reprogramming; mass spectrometry; machine learning

Abstract

Abiotic stress exposure of plants induces metabolic reprogramming which is tightly regulated by signalling cascades connecting transcriptional with translational and metabolic regulation. Complexity of such interconnected metabolic networks impedes the functional understanding of molecular plant stress response compromising the design of breeding strategies and biotechnological processes. Thus, defining a molecular network to enable the prediction of a plant's stress mode promises to promote the understanding of stress responsive biochemical regulation and its technological application. *Arabidopsis* wild type plants and two mutant lines with deficiency in sucrose or starch metabolism were grown under ambient and cold/high light stress conditions. Stress-induced dynamics of the primary metabolome and the proteome were quantified in a mass spectrometry-based high-throughput experiment. Wild type data were used to train a machine learning algorithm to classify mutant lines under control and stress conditions. Multivariate analysis and classification identified a module consisting of 23 proteins enabling the reliable prediction of coupled temperature and light stress conditions. 18 of these 23 proteins displayed putative protein-protein interactions connecting transcriptional regulation with regulation of primary and secondary metabolism under stress. The identified stress-responsive core module provides evidence for predictability of complex biochemical regulation during environmental fluctuation.

Introduction

The exposure of plants to abiotic stress conditions induces metabolic reprogramming to stabilize growth and developmental processes and to prevent chronic damage of tissue and organs. Typically, plant stress tolerance and adaptation mechanisms represent multigenic and complex physiological traits [1].While intensity and extent of metabolic reprogramming depends on stress type, intensity and duration, plant stress response can generally be subdivided into an early (alarm) phase, an acclimation phase and the maintenance phase of the newly acclimated metabolic homeostasis [2]. Additionally, metabolic reprogramming affects various cellular processes on multiple molecular levels comprising regulation of transcriptional, translational and enzymatic processes [3]. For example, the C-repeat binding factor (CBF) pathway of *Arabidopsis* plays a crucial role in freezing tolerance [4]. Only minutes after transferring *Arabidopsis* plants from warm to cold temperature induces the CBF1, CBF2 and CBF3 genes [5] to alter the expression of about hundred cold-regulated (COR) genes, also known as the CBF regulon [6]. This results in an increase of freezing tolerance [7] and recent work indicated extensive co-regulation of the CBF regulon by other transcription factors providing evidence for a complex and highly interconnected low-temperature regulatory network [8]. Further, comparison of CBF regulon gene induction in natural *Arabidopsis* accessions with a differing freezing tolerance revealed a potential contribution of the CBF pathway to the adaptive evolution of those populations [9].

Transcriptional and translational reprogramming significantly affects protein levels under stress conditions [10], and, *vice versa*, changes in protein concentrations, e.g. transcription factors, affect transcriptional and translational processes. Previous work has revealed stress responsive proteins, e.g. dehydrins [11, 12] and RNA-binding protein CP29 and GRP7 [13-15], to strongly accumulate during exposure to low temperature. Stress-responsive proteins are distributed across various cell compartments involved in numerous signalling cascades, metabolic pathways of primary and secondary metabolism, protein folding, membrane stabilization, energy and redox regulation [16]. Central carbohydrate metabolism is a prominent target of stress-induced metabolic reprogramming. Carbohydrates are the primary photosynthetic products playing crucial roles in energy metabolism, developmental processes, stress signalling and many other processes ranging from the cellular to the whole plant level. Starch and soluble sugars have been found to be tightly linked by the circadian clock ensuring a continuous carbohydrate availability [17, 18]. Further, sugars play a crucial role in entrainment of the circadian clock [19], and clock components have been found to be significantly influenced by abiotic stress conditions, e.g. low temperature [20]. Previous work indicated an important role of starch

degradation during initial response to cold stress augmenting hexose and raffinose accumulation [21]. Yet, also in cold acclimated plants reprogramming of the starch degradation machinery was found to be a characteristic part of naturally occurring cold and freezing tolerance of *Arabidopsis thaliana* [22]. Regulation of starch and sucrose metabolism affects carbon and energy metabolism on a whole plant level and, hence, its reprogramming is central to plant abiotic stress response. Starch biosynthesis is directly linked to the Calvin-Benson cycle by a sequential action of phosphoglucoisomerase (PGI), phosphoglucomutase (PGM1) and ADP-glucose pyrophosphorylase (AGPase) yielding ADP-glucose, the direct substrate for biosynthesis of the starch granule [23]. Sucrose synthesis is regulated by sucrose-phosphate synthase (SPS) catalysing the cytosolic synthesis of sucrose-6-phosphate (S6P) from UDP-glucose and fructose-6-phosphate [24]. Sucrose phosphate phosphatase (SPP) releases inorganic phosphate from S6P, yielding sucrose. Mutant lines, deficient in PGM1 (*pgm1*) or SPS (*spsa1*) activity, have previously been reported to be significantly affected not only in the central carbohydrate metabolism but also in metabolism of organic and amino acids [25]. In starchless *pgm1* plants, increased sugar concentrations were observed in root and shoot tissue and were accompanied by significantly affected growth [26].

Due to the pivotal role of sugars in plant stress response, analysing mutants with a deficiency in the central carbohydrate metabolism promises to unravel metabolic network components with a crucial role in stress tolerance. However, deriving patterns of stress-induced metabolic regulation from experimental data sets is challenging due to the vast amount of involved molecular and regulatory processes, affecting diverse transcripts, proteins and metabolites simultaneously. Particularly, due to the multidimensional output of high-throughput experiments theoretical approaches are needed to support the detection of characteristic metabolic patterns and the generation of predictive models. Machine learning techniques are suitable to classify data sets based on patterns recognized in multivariate large-scale data sets [27]. Typically, training data sets comprising predictor and response variables are used to generate a trained model which can be applied to predict structures and classes in a new and, hitherto, unknown data set. Recently, a machine learning method was developed to predict whether proteins localize to the apoplast independent of the presence of a signal peptide [28]. Potentially, this method improves the accuracy of predicting whether plant pathogen-derived effectors localize still to the apoplast or entered already the plant cell. Other studies have applied machine learning techniques for stress phenotyping in plants from high-resolution images or the classification of proteomics data of field-grown potato cultivars [29, 30]. Conclusively, developing and applying machine learning techniques for classification, regression or clustering of

biological systems represents a promising approach to deriving conclusive metabolic patterns crucial to generate predictive models.

The present study aimed to identify central molecular components involved in abiotic stress-induced reprogramming of plant metabolism. Frequently, cold and high light occur simultaneously under natural conditions and extrapolation from each stress response to a combined one is not possible [31, 32]. Therefore, a merged abiotic stress treatment was chosen to induce stress response which potentially reflects a broad scenario of plant growth conditions. Stress-induced dynamics in protein and metabolite levels were recorded in wild type and mutant plants being affected in starch or sucrose biosynthesis. Multivariate statistics and machine learning techniques were combined to identify a molecular network enabling the prediction of metabolic stress response across all genotypes.

Materials and Methods

Plant material

Plants of *Arabidopsis thaliana,* accession Columbia (Col-0), and mutant lines *pgm1* (plastidial PGM, At5g51820, TAIRstock CS3092), and *spsa1* (At5g20280, SALK line 148643C) were grown within a growth cabinet under controlled conditions (Conviron® Adaptis). Light intensity under control condition was 50 µmol m^{-2} s^{-1} in a 12/12h day/night cycle. Relative air humidity was 60% and temperature was 22°C during the day and 18°C during the night. All plants were grown on soil which was composed of Einheitserde® ED63 and perlite. Plants were watered daily and fertilized once with NPK fertilization solution (WUXAL ®Super; MANNA°-Dünger, Ammerbuch). After 42 days, leaf rosettes were sampled after 3 hours in the light (= control samples). All other plants were transferred to 5°C and 300 µmol m^{-2} s^{-1} for 72 hours. Like control samples, stress samples were collected after 3 hours in the light and immediately quenched in liquid nitrogen. Each sample (control and stress) consisted of a whole leaf rosette which was ground to a fine powder and lyophilized.

Chlorophyll fluorescence

To quantify the impact of stress exposure on the photosynthetic apparatus, parameters of chlorophyll fluorescence were recorded using a WALZ MINI-PAM II/B (Heinz Walz GmbH, Germany). Rapid light curves were recorded after 5 minutes of dark adaptation. Intensities of

151 actinic illumination were increased as follows: 0, 24, 44, 64, 89, 124, 189, 284, 419, 630, 822,
152 1149, and 1498 µmol m^{-2} s^{-1}.

153

154 **Starch quantification**

155 Leaf starch content was quantified photometrically as described earlier [33]. In brief, soluble
156 sugars were extracted from ground plant material by adding 80% ethanol and heating to 80°C
157 for 30 minutes. The supernatant was discarded and the pellet was used for starch quantification.
158 Pellets were suspended in 0.5M NaOH at 95°C for 45 min. Following acidification with 1M
159 CH_3COOH, the suspension was digested with amyloglucosidase for 2h at 55°C. Glucose
160 content of the supernatant was quantified photometrically in a coupled glucose
161 oxidase/peroxidase/dianisidine assay.

162

163 **SPS activity measurement**

164 SPS activity was determined as described previously with slight modification [34]. Leaf tissue
165 was suspended in 50mM Hepes /KOH (pH 7.5), 10mM $MgCl_2$, 1mM EDTA, 2.5mM
166 dithiothreitol, 10% glycerol and 0.1% Triton X-100. Suspensions were centrifuged at 4°C and
167 SPS activity was assayed in the supernatant. The reaction buffer consisted of 50mM Hepes
168 /KOH (pH 7.5), 15mM $MgCl_2$, 2.5mM dithiothreitol, 35mM UDP-glucose, 35mM fructose 6-
169 phosphate and 140mM glucose 6-phosphate; 30% KOH was added to the control of each assay.
170 Reactions were incubated for 30 min at 25°C, and then for 5 min at 95°C. Anthrone 0.14% in
171 H_2SO_4 was added, and the samples were incubated for 30 min at 40°C. Extinction was
172 determined photometrically at 620 nm.

173

174 **Metabolite analysis**

175 Metabolite concentrations were absolutely quantified via gas chromatography coupled to time-
176 of-flight mass spectrometry applying a previously published protocol with slight modifications
177 [22, 35]. Primary metabolites were extracted twice with a methanol-chloroform-water mixture
178 (MCW, 5/2/1, v/v/v) followed by an extraction step with 80% ethanol in which the samples
179 were heated to 80°C for 30 minutes. Addition of water to the MCW supernatant induced
180 separation of polar and apolar phases. The polar phase was mixed with the ethanol extract and
181 dried in a vacuum concentrator (ScanVac, LaboGene). The dried extracts were derivatized
182 applying methoximation (Methoxyamine hydrochloride in pyridine) and silylation (N-Methyl-

183 N-(trimethylsilyl) trifluoroacetamide). For methoximation, samples were incubated for 90
184 minutes at 30°C. For silylation, samples were incubated for 30 minutes at 37°C. Derivatized
185 samples were transferred into glass vials and sealed with a crimp cap. GC-ToF-MS analysis
186 was performed on an Agilent 6890 gas chromatograph (Agilent Technologies®, Santa Clara,
187 USA) coupled to a LECO Pegasus® GCxGC-TOF mass spectrometer (LECO Corporation, St.
188 Joseph, USA). Compounds were separated on an Agilent column HP5MS (length: 30 m,
189 diameter: 0.25 mm, film: 0.25 µm). Deconvolution of the total ion chromatogram and peak
190 integration was performed using the software LECO Chromatof®.

191

192 **Proteome analysis**

193 Plant material was solubilized in 8M urea, 50mM Hepes KOH (pH=7.8) on ice. Samples were
194 precipitated in acetone with 0.5% beta-mercaptoethanol. Pellets were washed 2 times with
195 methanol and acetone, and afterwards again solubilized in 8M urea, 50mM Hepes KOH (pH
196 7.8). Protein concentration was determined via Bio-Rad Bradford-assay using BSA as standard
197 (Bio-Rad, USA). Equal amounts of protein (15µg) were reduced with dithiothreitol (DTT) at a
198 concentration of 5mM for 45 minutes at 37°C, and alkylated at a concentration of 10mM with
199 iodacetamid (IAA) and dark incubated for 60 minutes at 23°C. Alkylation was stopped by
200 increasing DTT concentration to 10mM and dark incubated for 15 minutes. Samples were
201 diluted 2-fold by 20% acetonitrile (ACN) and 100mM ammonium bicarbonate (AmBic),
202 proteins were predigested with Lys-C (1:1000 w:w; Sigma-Aldrich, USA) at 30 °C for 2.5 hours
203 dark. Samples were diluted 2-fold to 2M urea by 10% ACN, 25mM AmBic, 10mM $CaCl_2$ and
204 digested with sequencing grade modified trypsine (Poroszyme, immobilized trypsin; 1:100 v:w)
205 for 12 hours. Digested proteins were acidified with formic acid (pH ~ 3.0) and desalted with
206 C18 materials (Agilent, Bond Elut SPEC) and dried in a vacuum concentrator (ScanVac,
207 LaboGene). Peptides were solved in 2% ACN, 0.1% formic acid and the same amount of total
208 protein was loaded and separated on a PepMap RSLC 75µm, 50cm column (Thermo Fisher
209 Scientific Inc., Waltham, USA). The flowrate was set to 300nl/min with 2% to 40% in 120
210 minutes of mobile phase B (mobile phase A: 0.1% formic acid (FA) in water [v/v]; mobile
211 phase B: 0.1% FA in 90% ACN [v/v]). The run ended with 60 minutes of equilibrium.
212 Subsequently, mobile phase B was increased from 40% to 90% within 1 minute and held stable
213 for 10 minutes, followed by decreased mobile phase B from 90% to 2%. MS analysis was
214 performed with an Orbitrap Elite instrument (Thermo Fisher Scientific Inc., Waltham, USA) in
215 positive mode and full scan in FT with a resolution of 60.000 in profile mode. Precursor masses

ranged between 360-1800 m/z. MS/MS analysis was done in the linear ion trap with CID fragmentation and rapid scan mode for the 20 most intense ions with a minimal signal threshold of 500 counts. Prediction of ion injection time was enabled ($5*10^2$ ions for up to 10 ms). Dynamic exclusion was enabled with repeat count 1 and repeat duration of 30 seconds. Exclusion list size was set to 500 and exclusion duration to 30 seconds. Excluded mass was set to +/- 10ppm relative to reference mass, early expiration enabled with 1 count and the s/n threshold was 2.0.

Peptide identification and protein quantification was performed with MaxQuant (http://www.maxquant.org) and implemented algorithms of version 1.5.5.1 [36] against the TAIR10 (www.arabidopsis.org) protein database [37]. A maximum of 2 missed cleavages was applied. Maximally 5 variable modifications per peptide were allowed for N-terminal acetylation and methionine oxidation. Carbamidomethylation was set as a fixed modification (due to previous methylation). For identification a minimum of 2 peptides and 2 minimum razor + unique peptides were requested.

Data analysis

Statistical analysis (Pearson correlation, 2-way ANOVA, PCA) was performed in R (The R Project for Statistical Computing; http://www.r-project.org/) [38]. Venn diagrams were created using the online tool *Venny 2.1* [39]. Data classification by supervised machine learning was performed in MATLAB® (R2017a, 9.2.0.538062) (www.mathworks.com) applying the Classification Learner app. Classification was performed with linear, quadratic and cubic support vector machines (SVM). SVMs classify data by finding the best hyperplane for separating data points of one class (*here*: control) from the other (*here*: stress) [40]. Classification training was performed using the wild type (Col-0) data set. The trained wild type model was exported into the MATLAB® workspace and applied to predict the classification of control and stress samples of *pgm1* and *spsa1* mutants.

Results

Stress-induced effects on chlorophyll fluorescence parameters

The exposure of wild type and mutant plants to combined cold and high light stress resulted in a decreased effective photochemical quantum yield of PSII, ϕ_{PSII} (Fig. 1). This effect

248 significantly differed between control and stress condition (ANOVA, $p<0.001$) and was similar
249 in all genotypes. ϕ_{PSII} measures the proportion of light absorbed by chlorophyll associated with
250 PSII being used in photochemistry, thus being an indicator of overall photosynthesis [41].
251 Together with significant condition-dependent effects on coefficients of photochemical (qP)
252 and non-photochemical quenching (qN; Supplemental Figure S1) these observations indicated
253 a similar stress-induced effect on photobiochemistry in all genotypes.
254
255 **Deficiency of phosphoglucomutase affects the capacity of sucrose biosynthesis**
256 Under ambient conditions, maximum sucrose phosphate synthase activity was significantly
257 elevated in the plastidial *pgm1* mutant line compared to the wild type (Fig. 2a). Under stress,
258 SPS activity of the wild type slightly increased while it decreased significantly in *pgm1*
259 (ANOVA, $p<0.01$). Compared to Col-0, stress exposure resulted in lower SPS activity in *pgm1*
260 while SPS activity in *spsa1* was 10-20% of the wild type under both conditions (Fig. 2a; for
261 comparison of SPSA1 protein levels see Supplemental Figure S2).
262 *pgm1* was starch deficient under both analysed conditions (Fig. 2b) while all other genotypes
263 displayed a significant stress-induced increase of starch levels (ANOVA, $p<0.001$). Under both
264 ambient and stress conditions starch levels were highest in *spsa1*. Under stress, starch level of
265 *spsa1* was almost twice as high as in Col-0 ($p<0.001$).
266 Pearson correlation of SPS activity, starch levels and concentrations of primary metabolites
267 across all genotypes and conditions revealed a significantly positive correlation of SPS activity
268 with sucrose concentration (Fig. 3; $p<0.05$). Additionally, SPS activity was negatively
269 correlated to levels of starch, serine, malate and proline (Fig. 3). Further, starch was negatively
270 correlated with sucrose concentration while it showed positive correlation with central
271 compounds at the interface of carbon/nitrogen metabolism, e.g. glutamate, glutamine, glycine
272 and serine. Strongest correlation of TCA cycle intermediates was observed between citrate,
273 succinate and malate while fumarate was only weakly correlated with succinate (Fig. 3).
274 Interestingly, raffinose was found to positively correlate with its metabolic precursor galactinol,
275 yet not sucrose, pointing to a substrate limitation of raffinose biosynthesis by galactinol
276 biosynthesis [42]. Further, raffinose concentration was negatively correlated ($p<0.05$) with
277 pyruvate concentration which might indicate trade-off between synthesis of stress protective
278 substances and energy metabolism.
279
280
281

PCA reveals a weak stress-induced effect on the proteome of *pgm1*

To compare stress-induced reprogramming of the metabolome and proteome across all conditions and genotypes, principal components were calculated distinctly separating stress samples from control samples in all genotypes but the starch deficient *pgm1* (Fig. 4). While PCA of the primary metabolome, including starch, clearly separated stressed from control samples of *pgm1* (Fig. 4a), stress-induced reprogramming of the proteome of *pgm1* was less distinct than in all other genotypes (Fig. 4b). In the PCA of the primary metabolome, largest proportion of variance was explained by stress-induced reprogramming of Col-0 and *spsa1* (PC1; 35%), while PC2 explained the separation of *pgm1* from the other genotypes as well as the stress reaction of *pgm1* (Fig. 4a). Associated loadings, i.e. metabolites, showing a significant increase during stress in all genotypes but staying constant in *pgm1* were citrate, succinate, threonine, glutamate and starch (Fig. 2b and Fig 5 a-d). Remarkably, pyruvate was the only metabolite which decreased in its concentration during stress in Col-0 ($p=0.07$) but stayed constant in both mutant lines (Fig. 5 e).

On PC2, describing the *pgm1* stress reaction, starch and sucrose were the loadings with most influence on the observed variance (Supplemental Table I). Starch was deficient in *pgm1* under both conditions while it significantly increased under stress in all other genotypes ($p<0.001$; Fig. 2 b). While sucrose decreased in *pgm1,* it slightly increased in Col-0 and *spsa1* (Fig. 5f). Further, glucose and fructose concentrations increased significantly in all genotypes except for *pgm1* where it decreased (Fig. 5 g, h). Similar dynamics were observed for maltose and asparagine (Fig. 5 k, m). Under stress, myo-inositol significantly decreased in *pgm1* mutants while it increased in *spsa1* and did not change in Col-0 (Fig. 5 n).

In contrast to the primary metabolome, stress-induced reprogramming of the *pgm1* proteome was less pronounced (Fig. 4 b). Genotype-wise comparison of stress-induced significant changes in the proteome ($p<0.05$) revealed a much higher number of changing proteins in Col-0 (199) and *spsa1* (120) than in *pgm1* (46) (Fig. 6). Also, the number of genotype-specific and uniquely changing proteins was lowest in *pgm1* (15) followed by *spsa1* (58), Col-0 (141). These observations indicated a strong interference of the central carbohydrate metabolism with reprogramming of the proteome during combined cold/high light stress. Except for that, 23 proteins (out of 1644 proteins) were observed to be significantly reprogrammed in all genotypes (see Fig. 6) indicating a central stress-responsive set of proteins which seemed to respond independently of the cellular carbohydrate status. Remarkably, levels of all 23 proteins were significantly increased under stress condition in all genotypes while none of them decreased ($p<0.001$; see Supplemental Figure 3).

316 To analyse whether this set of proteins potentially constituted a stress-induced interaction
317 and/or signalling network, information about potential protein-protein interactions were derived
318 from the STRING database (https://string-db.org/; Minimum interaction score: 0.4 (medium
319 confidence); [43]; (Table I)). Interaction analysis revealed a protein-protein interaction network
320 comprising 18 out of the 23 selected proteins. The interaction network comprised well-known
321 cold-induced proteins like COR15b and COR78 being part of the CBF regulon. Interestingly,
322 both COR proteins showed potential interaction with an unknown transmembrane protein
323 predicted to be located in the ER or in extracellular region (AT1G16850). Further, both COR
324 proteins showed interaction with the ABA and cold-inducible protein KIN1, a potential anti-
325 freeze protein. Beyond molecular chaperones in mitochondria (mtHsc70-1; AT4G37910) and
326 plastids (CLPB3; AT5G15450), also delta 1-pyrroline-5-carboxylate synthase 2 (P5CS2),
327 involved in proline biosynthesis, was part of the identified stress-responsive network with
328 potential interaction with COR78. Further proteins of the network were a central part of
329 carbohydrate metabolism (RHM1; SUS1 and MIPS1) or belonged to the interface of primary
330 and secondary metabolism, e.g. phenylalanine ammonia-lyase 1 and 2, flavanone 3-
331 hydroxylase and chalcone synthase/isomerase family proteins (PAL1 and PAL2, F3H, TT4,
332 TT5; see Table II). Finally, two plastidial RNA-binding proteins, CP29 and AT2G37220, were
333 identified as potential protein-protein interaction partners. Proteins which were part of the
334 genotype-independent stress-responsive proteome but did not show any potential interaction
335 within this set were CCR2 (AT2G21660), fibrillin precursor protein (FIB1A, AT4G04020),
336 ferritin 1 (FER1, AT5G01600), glycine-rich RNA-binding protein 4 (GRP4, AT3G23830) and
337 SVR3 (AT5G13650) which is involved in the elongation process during protein biosynthesis.
338
339 **Predicting metabolic stress response using support vector machines**
340 The predictability of metabolic stress response was analysed using support vector machines
341 (SVMs) for pattern recognition. In general, SVMs classify data by optimizing a hyperplane
342 separating data points of two classes. Here, the two classes were defined by the conditions,
343 control and stress. The best hyperplane SVM was defined by the largest margin between those
344 two classes. Three different kernel functions were applied for data classification, a linear,
345 quadratic and cubic function. Models were trained using experimental metabolomics and
346 proteomics data of Col-0. Subsequently, trained models, which showed 100% accuracy for Col-
347 0, were applied to experimental data of *pgm1* and *spsa1* to predict whether sample data belong
348 to the class "control" or "stress". Rate of success of SVM classification was calculated as the
349 percentage of accurate predictions (Table II).

In general, using experimental training data of Col-0 allowed for the accurate prediction of control and stress samples in *spsa1* mutants. In contrast, *pgm1* stress state was not accurately predictable by the Col-0 model if metabolite concentrations were applied (Table II, A-F). Prediction of condition in *pgm1* based on the primary metabolome including starch resulted in maximally 87% accuracy. To test the effect of starch dynamics on prediction accuracy in *pgm1*, training and prediction was performed excluding starch levels (Table II, B). However, prediction accuracy of *pgm1* samples could not be rescued to 100% as it was observed for *spsa1*. Next, sucrose, hexoses and maltose were excluded separately (Table II, C-E) and in combination with each other from training and prediction due to their significantly inverse stress dynamics in *pgm1* compared to all other genotypes (see Fig. 5). Excluding both sucrose and hexoses from the data sets most efficiently raised the prediction accuracy from 67% to 87% across all applied SVM kernel functions (Table II, F).

In contrast to the full metabolome information, the full proteome information showed a similar prediction accuracy in all genotypes (Table II, G-I). Except for the linear SVM kernel function prediction, where prediction accuracy of *pgm1* was still at 67% (Table II, G lSVM), the accuracy of condition was successful to 100% in the quadratic and cubic SVM kernel function, even for *pgm1*. Further, using only the set of 23 genotype-independent stress responsive proteins a predictive SVM model with 100% accuracy was curated across all kernel functions and genotypes (Table II, H). Finally, using this set of proteins together with the full metabolome information increased the predictability of *pgm1* samples from 67% (Table II, full metabolome, lSVM) to 100%.

Discussion

Sucrose biosynthesis capacity is affected by plastidial phosphoglucomutase during the initial light phase

Starch and sucrose represent two major products of photosynthetic carbon assimilation. Regulation of their metabolism crucially affects carbon partitioning and carbon balance on a whole-plant level [44]. Thus, perturbing biosynthetic pathways of both carbohydrates, i.e. by T-DNA insertion, results in dramatic alteration of photosynthesis, metabolite concentrations and stress tolerance [21, 45]. In the present study, metabolomic and proteomic consequences of altered sucrose and starch metabolism were analysed during the early day phase, i.e. after 3 hours in the light. Previous studies have shown strong dynamics of carbohydrate metabolism during this day period (see e.g. [26, 46]). Hence, analysis at such an early point of time in the day promises to

384 reveal the role of sucrose and starch metabolism in establishing a daily equilibrium. While the
385 applied stress combination of increased light intensity and decreased temperature cannot
386 resolve the individual stress response, it provides insight into the synergistic stress response
387 which might not be predictable by single stress application [47]. Chlorophyll fluorescence
388 parameters revealed a significant stress effect (ANOVA, $p<0.001$) on the effective
389 photochemical quantum yield as well as on photochemical and non-photochemical quenching,
390 yet no significant genotype-effect was detected. This implies a similar constitution of
391 photosystems in all genotypes and allows for an interpretation of observed differences in
392 metabolism independent of (potential) photosystem damage.

393 Deficiency of plastidial phosphoglucomutase, resulting in, more or less, starchless plants, has
394 been shown in many previous studies to enhance diurnal and stress-induced dynamics of soluble
395 carbohydrates due to a loss of buffering capacity for carbon acquisition rates affecting both
396 source and sink tissue [26]. However, maximum capacity of sucrose biosynthesis ($v_{max, SPS}$) in leaf
397 tissue of Arabidopsis was not found to correlate with increased sucrose concentrations, neither
398 before nor after cold acclimation [25]. In contrast, SPS activity correlated weakly, yet
399 significantly, positive with sucrose concentrations in the present study. Already under control
400 conditions a significantly increased v_{max} of SPS was observed in *pgm1* plants. Similarly,
401 previous analyses indicated slightly increased SPS activity in *pgm1* compared to Col-0 under
402 12/12 day/night growth conditions [46]. However, in the present study the difference to Col-0 was
403 more pronounced which might be due to a different light regime. Here, plants were grown at
404 PAR 50 µmol $m^{-2}s^{-1}$ while Gibon and co-workers applied a light regime which was almost
405 threefold higher (140 µmol $m^{-2}s^{-1}$). Also, on the protein level, a significant increase of SPSA1
406 protein was observed in *pgm1* mutants under control conditions (Supplemental Figure S2),
407 indicating that activity reflected the protein level of the SPS enzyme. Interestingly, activity
408 under combined stress decreased in the *pgm1* mutant while protein levels increased, pointing to
409 an inactivation by phosphorylation [48-50]. While the exact mechanism of SPS inactivation
410 remains speculative in the present study, it seems probable that the interplay of cytosolic protein
411 kinase/phosphatase activity and SPS is differentially regulated in *pgm1* compared to Col-0.

412

413 **A central core proteome is sufficient for accurate predictions of plant growth conditions**
414 **using machine learning techniques**

415 Identification of patterns in experimental data sets suitable for prediction of growth conditions,
416 developmental stages or stress tolerance is a central aim of quantitative (plant) biology. Such

patterns potentially reveal conserved molecular mechanisms and molecular interaction networks fundamental for complex cellular organisation. Promisingly, machine learning techniques enable such pattern recognition in large, heterogeneous and multidimensional data sets, hence being suitable for biological data integration [51]. Using support vector machines with linear, quadratic and cubic kernel functions to classify experimental proteomics and metabolomics data using the Col-0 data set as a training data set, *spsa1* growth conditions were predictable with high accuracy. For *pgm1*, predictability of metabolome data was weak (~67% accuracy) due to the strong deviation from patterns observed in Col-0. Yet, when combining the metabolome with a statistically selected subset of the proteome which was observed to be stress-induced independent of the genotype, all applied machine learning techniques predicted the growth condition with high accuracy (see Table I). Remarkably, predictions based solely on the proteome resulted in much higher accuracy of predictions than predictions based on metabolomics data. Together with multivariate analysis of metabolomics and proteomics data this indicates that stress-induced reprogramming of the *pgm1* proteome occurs in higher similarity to Col-0 and *spsa1* than reprogramming of the metabolome. Previous findings on cold response of *pgm1* revealed a similar multivariate pattern of integrated metabolite/protein data after 3 days at 4°C [52]. In this study, the authors detected metabolic processes which distinguish genotype- and temperature-specific effects.

Conclusively, these findings suggest proteomics data to enable high-accuracy predictions of growth conditions with a varying light intensity and temperature regime. In particular, the identified core proteome, resulting from the overlap of all significantly changed proteins across all genotypes (see Fig. 6), strongly contributed to these predictions and was sufficient to increase prediction accuracy of the metabolome from 67% to 100%. The resulting potential protein-protein interaction network comprised several well-described stress responsive proteins, e.g. COR15b, COR78 or molecular chaperones like CLPB3 [53-55]. Proteins with most interactions were COR78 (n=5), P5CS2 (delta 1-pyrroline-5-carboxylate synthase 2, n=6), F3H (n=6) and TT4 (n=6), PAL1 (n=5), ADH1 (n=6). P5CS2 is involved in biosynthesis of proline which is well-known to accumulate during stress response in many plant species playing diverse roles in signalling, cryoprotection and redox balance [56]. F3H, a flavanone 3-hydroxylase, hydroxylates naringenin to form dihydrokaempferol which can further be hydroxylated to form dihydroquercetin, the substrates for flavonol and anthocyanin biosynthesis [57]. TT4, a chalcone and stilbene synthase family protein, and TT5, a chalcone-flavanone isomerase family protein, catalyse the biosynthesis of naringenin and, hence, provide the substrate for F3H enzymes [57]. Previously, mathematical modelling of metabolomics data identified biochemical reactions

being strongly involved in metabolic reprogramming during simultaneous application of cold and light stress [58]. One of the most significantly reprogrammed reactions in this study was the entry point of flavonoid and anthocyanin biosynthesis being directly related to PAL1/PAL2 abundance and activity [58, 59]. In the present study, several additional proteins which are involved in this pathway were identified indicating the central role of its regulation during abiotic stress response. Furthermore, a predicted potential interaction between PAL1, PAL2 and P5CS2 establishes an interface between primary and secondary metabolism and indicates how redox balance might affect stress-induced branches of secondary metabolism.

An additional branch of the genotype-independent stress-responsive protein network comprised steps of central carbohydrate metabolism (MIPS1, SUS1) and cell wall synthesis (RHM1). MIPS1, a myo-Inositol-1-phosphate synthase, catalyses the limiting step of inositol biosynthesis and in response to stress, the transcription of MIPS1 is induced promoting the biosynthesis of inositol and derivatives. Previously, the light signalling protein FAR-RED ELONGATED HYPOCOTYL3 (FHY3) and its homolog FAR-RED IMPAIRED RESPONSE1 (FAR1) were shown to regulate light-induced inositol biosynthesis and oxidative stress responses by directly binding to the promoter of MIPS1 and activating its expression [60]. Hence, findings of the present study suggest a central role of MIPS1 under combined cold/high light stress conditions where oxidative stress response might be even more pronounced than under single stress conditions. Together with previous reports about the potential involvement of sucrose synthase (SUS1) in starch and cellulose biosynthesis [61], the finding that RHM1 protein levels are significantly increased suggests reprogramming of carbon allocation towards cell wall biosynthesis. Another central core protein was ADH1, which reduces acetaldehyde to ethanol to regenerate NAD^+ to maintain energy-generating glycolysis. ADH1 is known to accumulate during various stresses like hypoxia, salt, dehydration and cold [62-64]. The interaction of ADH1 with the cold response genes (COR78), amino acid metabolism (P5CS2), secondary metabolites (F3H, TT4, TT5) and sucrose synthase (SUS1) integrates all main targets of metabolism found in the core set.

Six further proteins involved in transcriptional and circadian regulation, cell organisation and protein biosynthesis were part of the identified stress responsive core proteome. Remarkably, all of them shared the chloroplast as a subcellular localization. FIB1A, a fibrillin precursor protein, has earlier been shown to play a role in abscisic acid (ABA)-mediated photoprotection [65]. These authors reported enhanced tolerance of PSII towards light-stress induced photoinhibition due to ABA treatment and fibrillin accumulation. In the present study, no significant genotype effect was detected for chlorophyll fluorescence parameters. Although the

485 contribution of FIB1A to photosystem stabilization remains to be elucidated in the background
486 of *pgm1* and *spsa1*, the significant increase of FIB1A is likely to contribute to the observed
487 similar photosystem constitution in all genotypes.
488 In summary, the stress-responsive core proteome identified in this study interconnects cellular
489 processes across various subcellular compartments and biological functions. Due to their
490 consistent stress-response across a large range of metabolic constitutions, it is likely that the
491 identified 23 protein candidates are dominantly involved in light and temperature stress
492 response across a wide genetic background. Finally, resolving the evolutionary conservation
493 and the ecological role of this core proteome will potentially provide novel insights into
494 complex stress tolerance mechanisms in plants.

Acknowledgements

We thank the whole MoSys Team at the University of Vienna for support and fruitful discussion. Particularly, we would like to thank Jakob Weiszmann and Martin Brenner for critical discussion and technical assistance. Finally, we thank the gardeners, Thomas Joch and Andreas Schröfl, for excellent plant cultivation. This work was supported by the Austrian Science Fund, FWF (Project Number I2071).

Author Contribution

LF, AP and HS performed experiments, data evaluation and wrote the manuscript. WW revised the manuscript. TN designed the study, performed experiments, data evaluation and wrote the manuscript.

507 **Tables**
508

509 ***Table I.*** **Molecular network components of genotype-independent stress response.**
510 Synonyms, function, subcellular localization and potential interaction was derived from the
511 databases TAIR (https://www.arabidopsis.org/) [37], MapMan (https://mapman.gabipd.org/) [66],
512 SUBA4 (http://suba.live/) [67], and String (https://string-db.org/) [68]. Further detailed
513 information is provided in the supplements (Supplemental Table II).
514

AGI	Synonym	Function	Subcellular localization	(Potential) Interaction with	Processes involved in
AT5G52310	**COR78** LTI78 LTI140 RD29A	Low-temperature-induced 78	Nucleus	KIN1 COR15B At1g16850 P5CS2 ADH1	Cold stress response/ has cis-acting regulatory elements that can impart cold-regulated gene expression
AT2G42530	**COR15B**	cold regulated 15b	Chloroplast	COR78 KIN1 AT1g16850	Cold stress response/protects chloroplast membranes during freezing
AT5G15960	**KIN1**	stress-induced protein (KIN1)	Cytosol	COR78 COR15B	Hormone metabolism/ABA cold and ABA inducible protein kin1/ possibly functions as an anti-freeze protein. Transcript level of this gene is induced by cold, ABA, dehydration and osmoticum (mannitol).
AT1G16850	**x**	Unknown protein transmembrane protein	Endoplasmic reticulum/ Extracellular region	COR15B COR78	not assigned/response to salt stress
AT3G55610	**P5CS2**	delta 1-pyrroline-5-carboxylate synthase 2	Cytosol/ Chloroplast/ Mitochondria/ Plasmodesma	COR78 mtHsc70-1 CLPB3 PAL2 ADH1 PAL1	Amino acid metabolism/gene expression is induced by dehydration, high salt and ABA
AT4G37910	**mtHsc70-1**	mitochondrial heat shock protein 70-1	Mitochondrion/ Cell wall/ Vacuolar membrane	CLPB3 P5CS2	Heat stress response/mitochondrial heat shock protein
AT5G15450	**CLPB3** CLPB-P APG6 ATCLPB3	casein lytic proteinase B3	Chloroplast (Stroma)/ Cytoplasm	mtHsc70-1 P5CS2	Heat stress response/functions as a molecular chaperone/involved in plastid differentiation mediating internal thylakoid membrane formation/conferring thermotolerance to chloroplasts during heat stress
AT1G77120	**ADH1** ATADH	alcohol dehydrogenase 1	Cytosol/Nucleus /Plasma membrane	F3H COR78 P5CS2 SUS1 TT4 TT5	Cellular respiration/oxidation-reduction process/positive regulation of cellular response to hypoxia/response to abscisic acid
AT2G37040	**PAL1** ATPAL1 PHE Ammonia Lyase 1	phenylalanine ammonia-lyase 1	Cytosol	F3H P5CS2 PAL2 TT4 TT5	L-phenylalanine catabolic process/cinnamic acid biosynthetic process/defense response, drought recovery/lignin catabolic process/
AT3G53260	**PAL2** ATPAL2	phenylalanine ammonia-lyase 2	Cytosol	P5CS2 F3H TT4 PAL1	Secondary metabolism/Phenylpropanoids/Lignin synthesis
AT3G51240	**F3H** TT6	flavanone 3-hydroxylase	Cytosol	TT4 TT5 PAL2 RHM1 ADH1	Secondary Metabolism/Flavonoids/ Dihydroflavonols/regulates flavonoid biosynthesis

				PAL1	
AT5G13930	**TT4** CHS ATCHS	Chalcone and stilbene synthase family protein	Endoplasmic reticulum/ Nucleus/ Cytoplasm	F3H PAL2 TT5 RHM1 ADH1 PAL1	Secondary metabolism/Flavonoids/Chalcones, Encodes chalcone synthase (CHS), a key enzyme involved in the biosynthesis of flavonoids/ Required for the accumulation of purple anthocyanins in leaves and stems.
AT3G55120	**TT5** A11 CFI CHI ATCHI	Chalcone-flavanone isomerase family protein	Endoplasmic reticulum/ Nucleus/ Chloroplast	TT4 F3H ADH1 PAL1	Secondary metabolism/Flavonoids/Chalcones, Catalyzes the conversion of chalcones into flavanones/Required for the accumulation of purple anthocyanins in leaves and stems. Co-expressed with CHS.
AT1G78570	**RHM1** ROL1 ATRHM1	rhamnose biosynthesis 1	Cytosol/ Chloroplast/ Plasmodesma	F3H TT4 SUS1	Cell wall/precursor synthesis/UDP-glucose 4,6-dehydratase/Encodes a UDP-L-Rhamnose synthase involved in the biosynthesis of rhamnose, a major monosaccharide component of pectin.
AT5G20830	**SUS1** ASUS1 atsus1	sucrose synthase 1	Cytosol	RHM1 MIPS1 ADH1	Major CHO metabolism/UDP-glycosyltransferase activity/sucrose synthase activity
AT4G39800	**MIPS1** ATMIPS1 ATIPS1 MI-1-P SYNTHASE	myo-inositol-1-phosphate synthase 1 myo-inositol-**3**-phosphate synthase 1	Cytosol	SUS1	Minor CHO metabolism/myo-inositol/InsP Synthases
AT3G53460	**CP29**	chloroplast RNA-binding protein 29	Chloroplast	AT2G37220	Regulation of transcription/RNA binding/ Encodes a nuclear gene with a consensus RNA-binding domain that is localized to the chloroplast
AT2G37220	x	RNA-binding (RRM/RBD/RNP motifs) family protein	Chloroplast	CP29	RNA binding/Encodes a chloroplast RNA binding protein.
AT3G23830	**GRP4** RBGA4	glycine-rich RNA-binding protein 4	Mitochondria	-	Response to cold/response to osmotic stress/ response to salt stress/response to water deprivation
AT2G21660	**CCR2** ATGRP7 GR-RBP7 GRP7 RBGA3	cold, circadian rhythm, and RNA binding 2	Nucleus/ Cytosol/ Peroxisome/ Chloroplast	-	RNA binding/Encodes a small glycine-rich RNA binding protein that is part of a negative-feedback loop through which AtGRP7 regulates the circadian oscillations of its own transcript/Gene expression is induced by cold.
AT4G04020	**FIB1A** FIB PGL35	Fibrillin precursor protein	Chloroplast/ Stroma	-	Cell organisation/Fibrillin precursor protein. The fibrillin preprotein, but not the mature protein interacts with ABI2/Regulated by ABA response regulators/involved in ABA-mediated photoprotection.
AT5G01600	**FER1** ATFER1	Ferritin 1	Chloroplast	-	Metal handling, binding/chelation/storage ferric iron/iron binding Encodes a ferretin protein that is targeted to the chloroplast.
AT5G13650	**SVR3**	elongation factor family protein	Chloroplast/ Plasma Membrane	-	Protein synthesis/Elongation/Encodes SVR3, a putative chloroplast TypA translation elongation GTPase.

515

516

19

517 Table II. **Support vector machine classification of control and stress samples.** Percentage
518 indicates the accuracy of predictions for genotypes based on a model trained with Col-0 data.
519 Kernel functions: lSVM – linear SVM; qSVM: quadratic SVM; cSVM: cubic SVM.
520 Abbreviations: Suc: Sucrose; Hex: Hexoses, Glucose and Fructose; Fum: Fumarate.
521

	Prediction accuracy (%)					
Experimental Data used for Training	***pgm1***			***spsa1***		
	lSVM	**qSVM**	**cSVM**	**lSVM**	**qSVM**	**cSVM**
A) Col-0, full Metabolome	67	87	73	100	100	100
B) Col-0, Metabolome, Starch excluded	73	67	67	100	100	100
C) Col-0, Metabolome, Suc excluded	67	80	73	100	100	100
D) Col-0, Metabolome, Hex excluded	80	87	87	100	100	100
E) Col-0, Metabolome, Maltose excluded	87	87	67	100	100	100
F) Col-0, Metabolome, Suc/Hex excluded	87	87	87	100	100	100
G) Col-0, full Proteome (1644 proteins)	67	100	100	100	100	100
H) Col-0, Stress-responsive Core Proteome (23 proteins)	100	100	100	100	100	100
I) Col-0, Stress-responsive Core Proteome (23 proteins) + full Metabolome	100	100	100	100	100	100

522

References

1. B. Vinocur and A. Altman, *Curr Opin Biotechnol*, 2005, **16**, 123-132.
2. K. Kosova, P. Vitamvas, I. T. Prasil and J. Renaut, *Journal of Proteomics*, 2011, **74**, 1301–1322.
3. D. C. Haak, T. Fukao, R. Grene, Z. Hua, R. Ivanov, G. Perrella and S. Li, *Frontiers in Plant Science*, 2017, **8**.
4. M. F. Thomashow, *Plant Physiol*, 2010, **154**, 571-577.
5. S. J. Gilmour, D. G. Zarka, E. J. Stockinger, M. P. Salazar, J. M. Houghton and M. F. Thomashow, *Plant J*, 1998, **16**, 433-442.
6. S. Fowler and M. F. Thomashow, *Plant Cell*, 2002, **14**, 1675–1690.
7. K. R. Jaglo-Ottosen, S. J. Gilmour, D. K. Zarka, O. Schabenberger and M. F. Thomashow, *Science*, 1998, **280**, 104–106.
8. S. Park, C. M. Lee, C. J. Doherty, S. J. Gilmour, Y. Kim and M. F. Thomashow, *Plant J*, 2015, **82**, 193-207.
9. M. A. Gehan, S. Park, S. J. Gilmour, C. An, C. M. Lee and M. F. Thomashow, *Plant J*, 2015, **84**, 682-693.
10. K. Nakaminami, A. Matsui, H. Nakagami, A. Minami, Y. Nomura, M. Tanaka, T. Morosawa, J. Ishida, S. Takahashi, M. Uemura, K. Shirasu and M. Seki, *Molecular & Cellular Proteomics*, 2014, **13**, 3602-3611.
11. S. A. Campbell and T. J. Close, *New Phytologist*, 1997, **137**, 61-74.
12. S. Amme, A. Matros, B. Schlesier and H.-P. Mock, *Journal of experimental botany*, 2006, **57**, 1537-1546.
13. K. Kosová, P. Vítámvás and I. T. Prášil, *Frontiers in plant science*, 2014, **5**, 711.
14. J. S. Kim, H. J. Jung, H. J. Lee, K. Kim, C. H. Goh, Y. Woo, S. H. Oh, Y. S. Han and H. Kang, *The Plant Journal*, 2008, **55**, 455-466.
15. C. Kupsch, H. Ruwe, S. Gusewski, M. Tillich, I. Small and C. Schmitz-Linneweber, *The Plant cell*, 2012, **24**, 4266-4280.
16. K. Kosová, P. Vítámvás, M. O. Urban, I. T. Prášil and J. Renaut, *Frontiers in Plant Science*, 2018, **9**.
17. A. Graf, A. Schlereth, M. Stitt and A. M. Smith, *Proceedings of the National Academy of Sciences*, 2010, **107**, 9458-9463.
18. B. Pommerrenig, F. Ludewig, J. Cvetkovic, O. Trentmann, P. A. W. Klemens and H. E. Neuhaus, *Plant and Cell Physiology*, 2018, DOI: 10.1093/pcp/pcy037, pcy037-pcy037.
19. M. J. Haydon, O. Mielczarek, F. C. Robertson, K. E. Hubbard and A. A. Webb, *Nature*, 2013, **502**, 689.
20. K. Miura and T. Furumoto, *International journal of molecular sciences*, 2013, **14**, 5312-5337.
21. R. Sicher, *Plant Sci*, 2011, **181**, 167-176.
22. M. Nagler, E. Nukarinen, W. Weckwerth and T. Nägele, *BMC Plant Biol*, 2015, **15**, 284.
23. M. Stitt and S. C. Zeeman, *Current Opinion in Plant Biology*, 2012, **15**, 282-292.
24. Y. L. Ruan, *Annu Rev Plant Biol*, 2014, **65**, 33-67.

25. Hoermiller, II, T. Naegele, H. Augustin, S. Stutz, W. Weckwerth and A. G. Heyer, *Plant Cell Environ*, 2017, **40**, 602-610.

26. K. Brauner, I. Hörmiller, T. Nägele and A. G. Heyer, *Plant J*, 2014, **79**, 82-91.

27. C. Ma, H. H. Zhang and X. Wang, *Trends in Plant Science*, 2014, **19**, 798-808.

28. J. Sperschneider, P. N. Dodds, K. B. Singh and J. M. Taylor, *New Phytologist*, 2018, **217**, 1764-1778.

29. A. Singh, B. Ganapathysubramanian, A. K. Singh and S. Sarkar, *Trends in Plant Science*, 2016, **21**, 110-124.

30. W. Hoehenwarter, A. Larhlimi, J. Hummel, V. Egelhofer, J. Selbig, J. T. van Dongen, S. Wienkoop and W. Weckwerth, *Journal of Proteome Research*, 2011, **10**, 2979-2991.

31. R. Mittler, *Trends in Plant Science*, 2006, **11**, 15-19.

32. C. M. Prasch and U. Sonnewald, *Environmental and Experimental Botany*, 2015, **114**, 4-14.

33. T. Nägele, S. Stutz, I. Hörmiller and A. G. Heyer, *Plant J*, 2012, **72**, 102-114.

34. T. Nägele, B. A. Kandel, S. Frana, M. Meissner and A. G. Heyer, *FEBS J*, 2011, **278**, 506-518.

35. W. Weckwerth, K. Wenzel and O. Fiehn, *Proteomics*, 2004, **4**, 78-83.

36. J. Cox and M. Mann, *Nature Biotechnology*, 2008, **26**, 1367.

37. P. Lamesch, T. Z. Berardini, D. Li, D. Swarbreck, C. Wilks, R. Sasidharan, R. Muller, K. Dreher, D. L. Alexander, M. Garcia-Hernandez, A. S. Karthikeyan, C. H. Lee, W. D. Nelson, L. Ploetz, S. Singh, A. Wensel and E. Huala, *Nucleic acids research*, 2012, **40**, D1202-D1210.

38. R Core Team, *R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0*, 2008.

39. J. C. Oliveros, 2007-2015.

40. G. James, D. Witten, T. Hastie and R. Tibshirani, *An introduction to statistical learning*, Springer, 2013.

41. K. Maxwell and G. N. Johnson, *Journal of Experimental Botany*, 2000, **51**, 659–668.

42. F. Keller and D. M. Pharr, *Photoassimilate distribution in plants and crops: source-sink relationships. Marcel Dekker, New York*, 1996, 157-183.

43. C. von Mering, L. J. Jensen, B. Snel, S. D. Hooper, M. Krupp, M. Foglierini, N. Jouffre, M. A. Huynen and P. Bork, *Nucleic acids research*, 2005, **33**, D433-437.

44. A. M. Smith and M. Stitt, *Plant, Cell and Environment*, 2007, **30**, 1126–1149.

45. J. Sun, J. Zhang, C. T. Larue and S. C. Huber, *Plant Cell Environ*, 2011, **34**, 592-604.

46. Y. Gibon, O. E. Bläsing, N. Palacios-Rojas, D. Pankovic, Hendriks, J. Fisahn, M. Höhne, M. Günther and M. Stitt, *Plant Journal*, 2004, **39**, 847–862.

47. C. M. Prasch and U. Sonnewald, *Environ Exp Bot*, 2015, **114**, 4-14.

48. R. W. McMichael, R. R. Klein, M. E. Salvucci and S. C. Huber, *Archives of Biochemistry and Biophysics*, 1993, **307**, 248-252.

49. S. C. Huber and J. L. Huber, *Plant Physiology*, 1992, **99**, 1275–1278.

50. W. M. Kaiser and S. C. Huber, *Journal of Experimental Botany*, 2001, **52**, 1981–1989.

602 51. Y. Li, F.-X. Wu and A. Ngom, *Briefings in bioinformatics*, 2016, DOI: 10.1093/bib/bbw113,
603 bbw113-bbw113.

604 52. S. Wienkoop, K. Morgenthal, F. Wolschin, M. Scholz, J. Selbig and W. Weckwerth,
605 *Molecular & cellular proteomics : MCP*, 2008, **7**, 1725–1736.

606 53. D. P. Horvath, B. K. McLarney and M. F. Thomashow, *Plant Physiology*, 1993, **103**, 1047-
607 1053.

608 54. A. Thalhammer, M. Hundertmark, A. V. Popova, R. Seckler and D. K. Hincha, *Biochimica et*
609 *Biophysica Acta (BBA) - Biomembranes*, 2010, **1798**, 1812-1820.

610 55. J. Larkindale and E. Vierling, *Plant Physiology*, 2008, **146**, 748-761.

611 56. L. Szabados and A. Savouré, *Trends in Plant Science*, 2010, **15**, 89-97.

612 57. W. Jiang, Q. Yin, R. Wu, G. Zheng, J. Liu, R. A. Dixon and Y. Pang, *Journal of Experimental*
613 *Botany*, 2015, **66**, 7165-7179.

614 58. H. Doerfler, D. Lyon, T. Nägele, X. Sun, L. Fragner, F. Hadacek, V. Egelhofer and W.
615 Weckwerth, *Metabolomics*, 2013, **9**, 564-574.

616 59. A. Rohde, K. Morreel, J. Ralph, G. Goeminne, V. Hostyn, R. De Rycke, S. Kushnir, J. Van
617 Doorsselaere, J.-P. Joseleau and M. Vuylsteke, *The Plant cell*, 2004, **16**, 2749-2771.

618 60. L. Ma, T. Tian, R. Lin, X. W. Deng, H. Wang and G. Li, *Mol Plant*, 2016, **9**, 541-557.

619 61. E. Baroja-Fernandez, F. J. Munoz, J. Li, A. Bahaji, G. Almagro, M. Montero, E. Etxeberria,
620 M. Hidalgo, M. T. Sesma and J. Pozueta-Romero, *Proceedings of the National Academy of*
621 *Sciences of the United States of America*, 2012, **109**, 321–326.

622 62. R. Dolferus, G. De Bruxelles, E. S. Dennis and W. J. Peacock, *Annals of Botany*, 1994, **74**,
623 301-308.

624 63. J. A. Jarillo, A. Leyva, J. Salinas and J. M. Martinez-Zapater, *Plant physiology*, 1993, **101**,
625 833-837.

626 64. H. Shi, W. Liu, Y. Yao, Y. Wei and Z. Chan, *Plant Science*, 2017, **262**, 24-31.

627 65. Y. Yang, R. Sulpice, A. Himmelbach, M. Meinhard, A. Christmann and E. Grill, *Proceedings*
628 *of the National Academy of Sciences*, 2006, **103**, 6061-6066.

629 66. O. Thimm, O. Bläsing, Y. Gibon, A. Nagel, S. Meyer, P. Krüger, J. Selbig, L. A. Müller, S. Y.
630 Rhee and M. Stitt, *The Plant Journal*, 2004, **37**, 914-939.

631 67. C. M. Hooper, S. K. Tanz, I. R. Castleden, M. A. Vacher, I. D. Small and A. H. Millar,
632 *Bioinformatics*, 2014, **30**, 3356-3364.

633 68. L. J. Jensen, M. Kuhn, M. Stark, S. Chaffron, C. Creevey, J. Muller, T. Doerks, P. Julien, A.
634 Roth and M. Simonovic, *Nucleic acids research*, 2008, **37**, D412-D416.

635

636

Figures

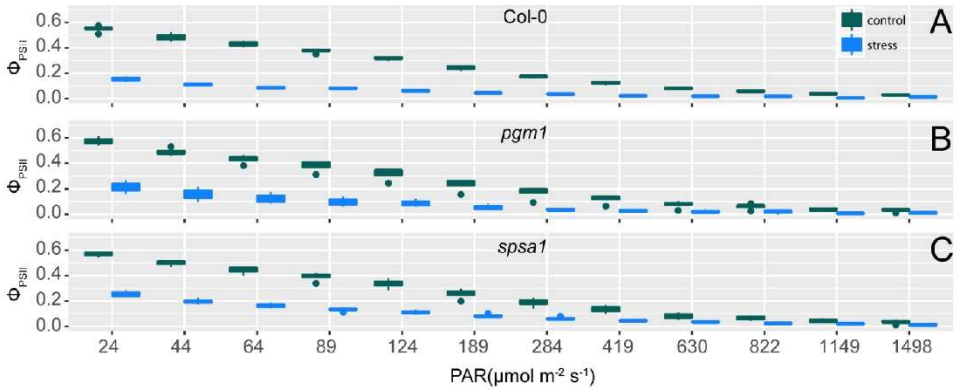


Figure 1. Effective photochemical quantum yield of PSII, ϕ_{PSII}, as a function of actinic illumination intensity. ϕPSII was measured in (**A**) Col-0, (**B**) *pgm1*, and (**C**) *spsa1* under control (green) and cold/high light stress (blue) conditions with an increasing actinic light intensity (PAR; n = 4). Conditions significantly differed within all four genotypes (ANOVA, p<0.001).

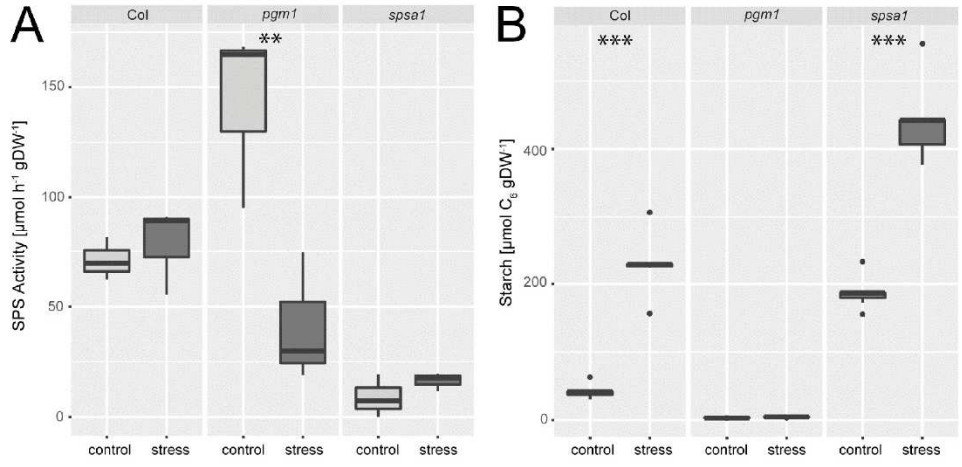
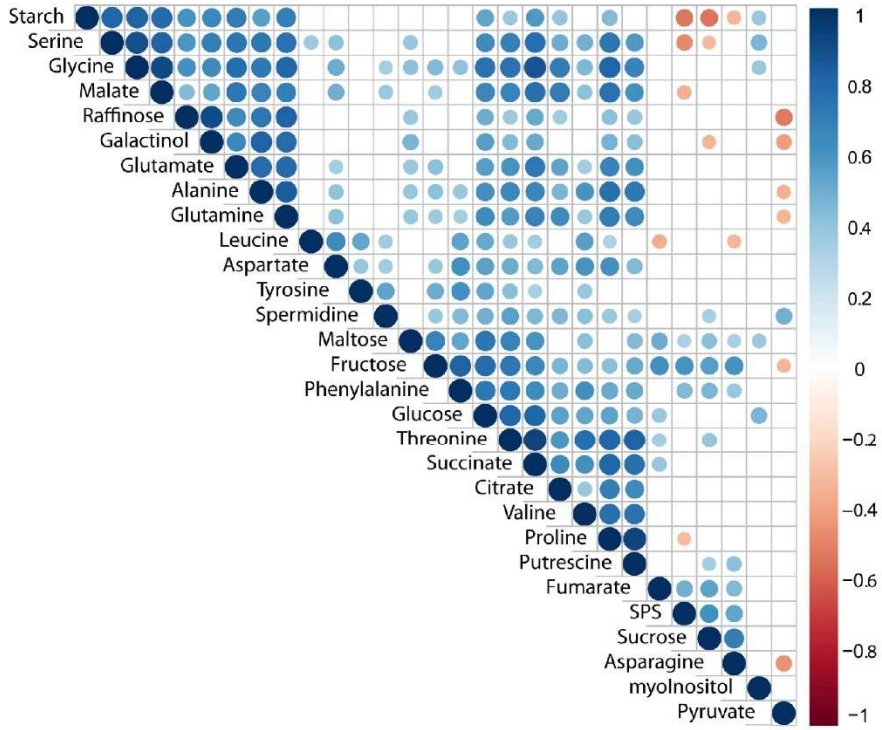


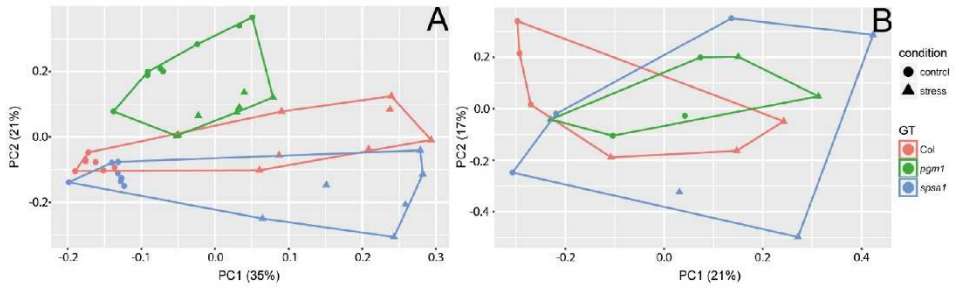
Figure 2. (A) SPS activity and (B) starch levels under control and combined cold/high light stress. Asterisks indicate significant differences within the genotypes revealed by an ANOVA. ** p<0.01; *** p<0.001 (SPS activity: n=3; Starch: n=5).



650

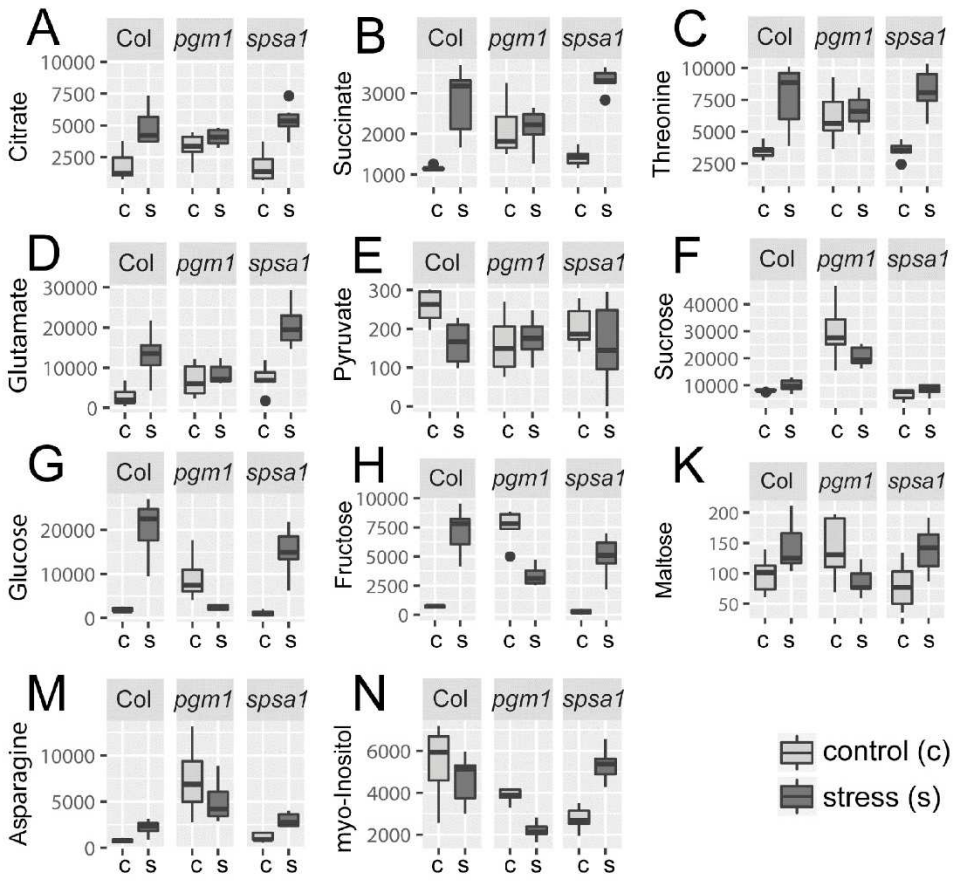
651 **Figure 3. Pearson correlation of primary metabolites, starch and SPS activity across all**
652 **genotypes and conditions**. Correlation coefficients are indicated by the colour (blue: positive
653 correlation; red: negative correlation). Only significant correlations ($p<0.05$) were colour-
654 coded, blank fields indicate non-significant correlations.

655



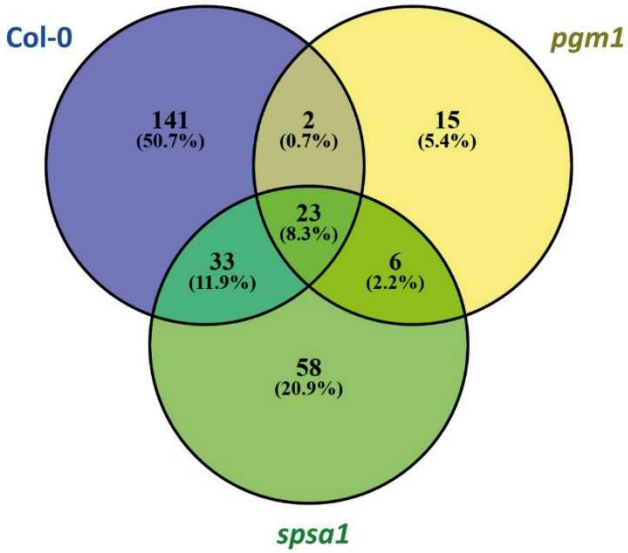
656

657 **Figure 4. Principal component analysis of the (A) primary metabolome (including starch)**
658 **and the (B) proteome**. Symbol colours indicate the genotype (red: Col-0; green: *pgm1*; blue:
659 *spsa1*). Symbol shape indicates conditions (control: filled circles; stress: filled triangles).



660

661 **Figure 5. Stress-induced reprogramming of the primary metabolome**. Concentration of
662 metabolites under control (c) and stress (s) conditions are summarized genotype-wise as
663 indicated on the top of each sub-figure. All levels are given in [a.u. gDW^{-1}]. Control samples
664 are shown in light grey, stress samples in dark grey (n=7). Only a selection of metabolites is
665 shown which increase significantly in all genotypes but stay constant in *pgm1* (A-D), decrease
666 only in Col-0 (E), increase in all genotypes but decrease in *pgm1* (F-N).



667

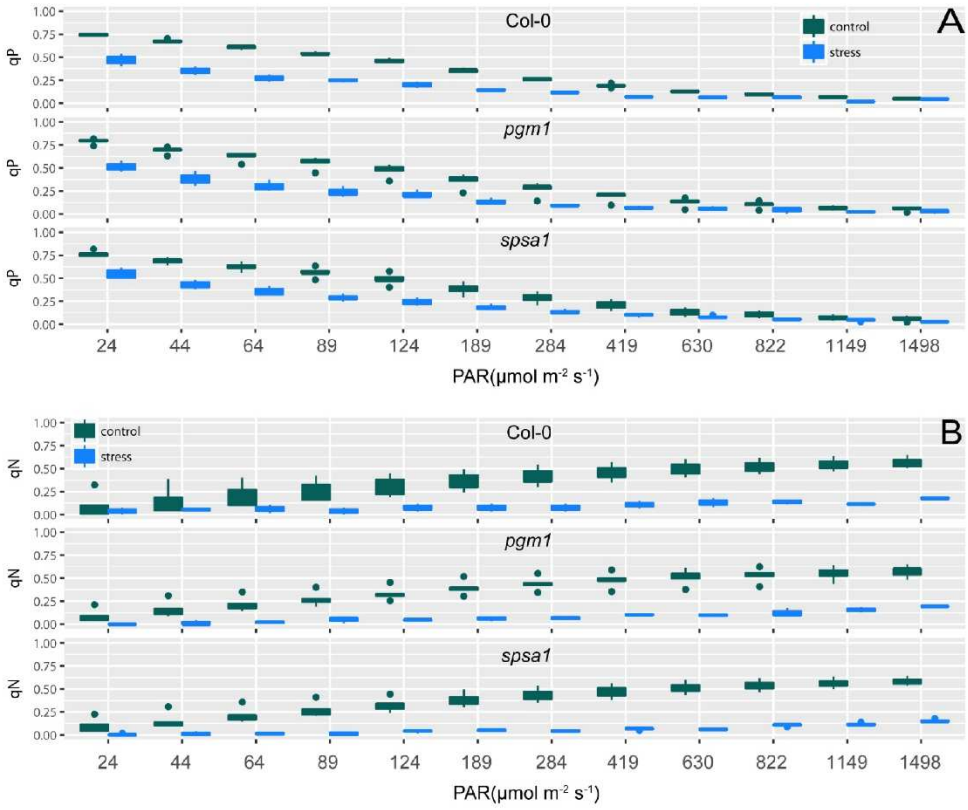
668 **Figure 6. Stress-induced reprogramming of the proteome.** The Venn diagram comprises
669 only significantly changed proteins due to stress treatment. Col-0: blue; *pgm1*: yellow; *spsa1*:
670 green. (The complete set of quantified proteins is provided in the Supplement, Table SIII).

671

27

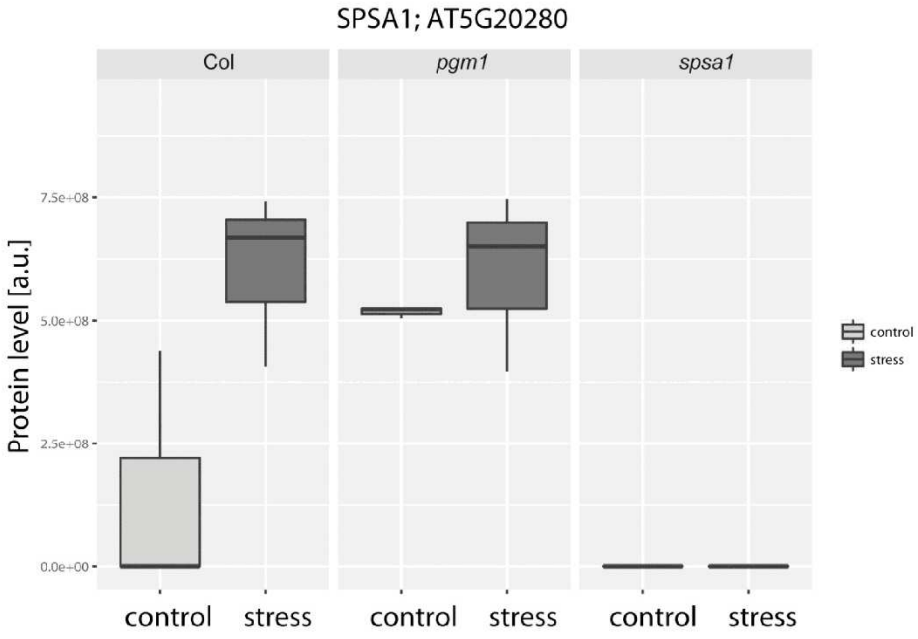
672 **Supporting Information**

673



674

675 **Supplementary Figure S1.** Coefficients of photochemical quenching (qP) and non-
676 photochemical quenching (qN) as a function of actinic illumination intensity. (A) qP and (B)
677 qN was measured in genotypes Col-0, *pgm1, spsa1* under control (green) and cold/high light
678 stress (blue) condition with an increasing actinic light intensity (PAR, n=4). Conditions
679 significantly differed within all four genotypes (ANOVA, $p<0.001$).

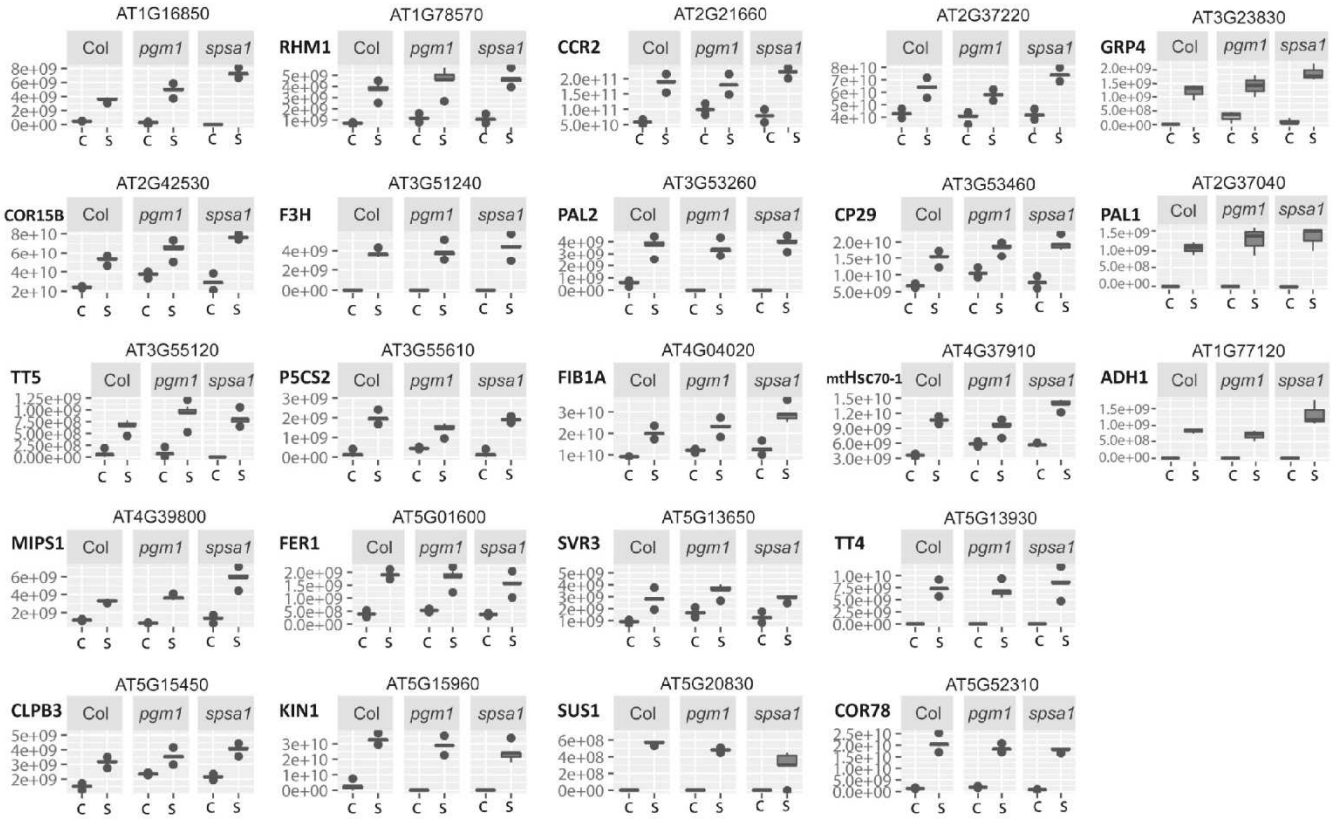


680

681 **Supplementary Figure S2.** Protein levels of SPSA1 measured by LC-MS shotgun proteomics
682 for all genotypes and conditions. Control samples are shown in light grey, stress samples in
683 dark grey (n=3). For further details on data analysis, please refer to the *Materials and Method*
684 section.

685

29



686

687 **Supplementary Figure S3.** Identified core proteins for all genotypes and conditions. c: control condition; s: stress condition (n=3). Abundance is given

688 in arbitrary units (MaxQuant LFQ values * protein amount [µg] * 1/gFW).

689 **Supplementary Table I.** PCA loadings of metabolome data.

690 Gol = galactinol, Raf = Raffinose, Gln = Glutamine, Glu = Glutamate, Ala = Alanine, myoIn =
691 myo-Inositol, Asp = Asparagine, Leu=Leucine, Glc = Glucose, Pyr = Pyruvate, Put =
692 Putrescine, Tyr = Tyrosine, Val = Valine, Phe = Phenylalanine, Spdine=Spermidine,
693 Frc=Fructose, Asn = Asparagine, Suc = Sucrose

	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9	PC10
Pyr	-0.0550	0.0402	0.4106	-0.2513	0.2318	0.3030	-0.1186	0.0752	0.0441	-0.4066
Suc	0.0159	0.4285	-0.1652	0.0009	0.1130	0.1466	-0.1088	-0.1890	0.1972	0.0504
Asn	0.0331	0.3054	-0.3544	0.1305	0.0587	0.0263	0.1334	0.3162	-0.3296	-0.0007
Fumarate	0.0660	0.2058	-0.2768	-0.4197	0.1055	-0.1413	0.0663	0.0376	0.1933	-0.1090
myoIn	0.0839	-0.0788	0.0717	-0.5069	-0.1119	0.1638	-0.0867	-0.0932	-0.6426	0.2399
Spdine	0.1089	0.2285	0.3331	-0.1800	0.2133	-0.1182	-0.2458	-0.1538	0.0185	0.2441
Tyr	0.1100	0.2019	0.2254	-0.0610	-0.2701	-0.2394	-0.4870	0.5050	0.0083	-0.1541
Leu	0.1135	0.0206	0.3877	0.1161	-0.3724	0.0255	0.2268	0.1097	0.2098	0.2110
Asp	0.1694	0.0353	0.2536	0.2355	-0.1432	-0.1976	0.2754	-0.3122	-0.3424	-0.3701
Maltose	0.1707	0.1606	-0.1322	-0.2149	-0.2188	0.3404	0.0665	-0.1716	0.1433	-0.4479
Raf	0.1776	-0.2196	-0.2426	0.0000	-0.2085	-0.0084	-0.3040	-0.3042	0.1459	0.0697
Starch	0.1796	-0.2936	0.0003	-0.1876	0.0414	-0.1260	0.1962	0.2376	0.2697	0.0254
Phe	0.1835	0.2753	0.0500	0.0972	-0.2558	-0.1563	0.1049	-0.0725	0.0471	0.1399
Frc	0.1860	0.2881	-0.1487	-0.0936	-0.1657	-0.1477	0.1342	0.0519	-0.1685	0.0346
Val	0.1884	0.1511	0.1725	0.2723	0.0171	0.3556	0.0089	0.0500	0.0461	-0.0313
Gol	0.1932	-0.2384	-0.1591	0.0438	-0.2655	0.1723	-0.1237	-0.1180	0.0398	-0.1333
Citrate	0.2014	0.0418	0.0227	0.0488	0.2553	-0.4600	-0.1301	-0.2690	0.0145	-0.2442
Glu	0.2248	-0.1353	-0.1414	0.0048	0.0768	-0.0335	0.1233	0.3882	-0.0638	-0.2395
Put	0.2257	0.0960	-0.0783	0.2382	0.2640	0.1682	-0.1434	-0.0321	-0.0853	0.0485
Malate	0.2332	-0.1291	0.0676	-0.0826	0.2620	-0.1280	0.2043	0.0426	-0.0165	-0.1282
Serine	0.2337	-0.2007	0.0873	-0.0822	0.1037	0.2023	0.0928	0.0232	-0.0362	0.1085
Ala	0.2347	-0.1242	-0.1083	0.1624	-0.0463	0.1541	-0.1884	0.1419	-0.1128	0.0053
Gln	0.2367	-0.1676	-0.0766	0.0481	-0.0413	-0.1142	-0.3880	-0.0073	-0.0112	-0.0204
Proline	0.2440	0.0029	0.0630	0.2339	0.2386	0.0489	-0.0712	0.0325	-0.0947	0.0831
Glc	0.2463	0.0826	0.0510	-0.1901	-0.2329	-0.0556	0.1049	-0.0785	0.0447	0.0079
Threonine	0.2559	0.1598	0.0319	-0.0358	0.0811	0.2007	0.0341	0.0012	0.0595	0.0965
Glycine	0.2599	-0.1183	0.0130	-0.0840	0.1479	-0.1025	0.1390	-0.0439	0.0024	0.1408
Succinate	0.2634	0.0595	-0.0143	-0.0739	0.0920	0.0511	0.1189	0.0417	0.2218	0.2612

694

695

696

697

698

699 **(Here) not included:**

700 **Supplementary Table II.** Molecular network components of stress response. Information
701 about synonyms, function, subcellular localization and potential interaction was derived from
702 the databases TAIR (https://www.arabidopsis.org/), MapMan (https://mapman.gabipd.org/),
703 SUBA4 (http://suba.live/), and String (https://string-db.org/).

704 **Supplementary Table III.** Identified and quantified protein levels and functions of all
705 genotypes and conditions (n=3). c: control condition; s: stress condition. Abundance is given in
706 arbitrary units (MaxQuant LFQ values * protein amount [µg] * 1/gFW).

VI. References

Arrivault S, Guenther M, Florian A, Encke B, Feil R, Vosloh D, Lunn JE, Sulpice R, Fernie AR, Stitt M, et al. 2014. Dissecting the subcellular compartmentation of proteins and metabolites in arabidopsis leaves using non-aqueous fractionation. *Mol Cell Proteomics* **13**(9): 2246-2259.

Baroja-Fernández E, Muñoz FJ, Li J, Bahaji A, Almagro G, Montero M, Etxeberria E, Hidalgo M, Sesma MT, Pozueta-Romero J. 2012. Sucrose synthase activity in the sus1/sus2/sus3/sus4 Arabidopsis mutant is sufficient to support normal cellulose and starch production. *Proceedings of the National Academy of Sciences* **109**(1): 321-326.

Carbonaro NJ, Thorpe IF. 2017. Using Structural Kinetic Modeling To Identify Key Determinants of Stability in Reaction Networks. *The Journal of Physical Chemistry A* **121**(26): 4982-4992.

Cheong R, Levchenko A. 2010. Oscillatory signaling processes: the how, the why and the where. *Curr Opin Genet Dev* **20**(6): 665-669.

Cramer GR, Urano K, Delrot S, Pezzotti M, Shinozaki K. 2011. Effects of abiotic stress on plants: a systems biology perspective. *BMC Plant Biol* **11**(1): 163.

Cvetkovic J, Müller K, Baier M. 2017. The effect of cold priming on the fitness of Arabidopsis thaliana accessions under natural and controlled conditions. *Sci Rep* **7**: 44055.

Dietz K-J. 2017. Subcellular metabolomics: the choice of method depends on the aim of the study. *J Exp Bot*: erx406-erx406.

Doerfler H, Lyon D, Nagele T, Sun X, Fragner L, Hadacek F, Egelhofer V, Weckwerth W. 2013. Granger causality in integrated GC-MS and LC-MS metabolomics data reveals the interface of primary and secondary metabolism. *Metabolomics* **9**(3): 564-574.

Dutta B, Kanani H, Quackenbush J, Klapa MI. 2009. Time-series integrated “omic” analyses to elucidate short-term stress-induced responses in plant liquid cultures. *Biotechnology and Bioengineering* **102**(1): 264-279.

Dutta B, Snyder R, Klapa MI. 2007. Significance analysis of time-series transcriptomic data: A methodology that enables the identification and further exploration of the differentially expressed genes at each time-point. *Biotechnology and Bioengineering* **98**(3): 668-678.

Farré EM, Tiessen A, Roessner U, Geigenberger P, Trethewey RN, Willmitzer L. 2001. Analysis of the Compartmentation of Glycolytic Intermediates, Nucleotides, Sugars, Organic Acids, Amino Acids, and Sugar Alcohols in Potato Tubers Using a Nonaqueous Fractionation Method. *Plant Physiol* **127**.

Flügge U-I. 1999. Phosphate translocators in plastids. *Annual Review of Plant Physiology and Plant Molecular Biology* **50**: 27–45.

Fowler S, Thomashow MF. 2002. Arabidopsis transcriptome profiling indicates that multiple regulatory pathways are activated during cold acclimation in addition to the CBF cold response pathway. *Plant Cell* **14**(8): 1675-1690.

Gehan MA, Park S, Gilmour SJ, An C, Lee CM, Thomashow MF. 2015. Natural variation in the C-repeat binding factor cold response pathway correlates with local adaptation of Arabidopsis ecotypes. *Plant J*.

Geigenberger P, Tiessen A, Meurer J. 2011. Use of non-aqueous fractionation and metabolomics to study chloroplast function in Arabidopsis. *Methods Mol Biol* **775**: 135-160.

Gerhardt R, Heldt HW. 1984. Measurement of Subcellular Metabolite Levels in Leaves by Fractionation of Freeze-Stopped Material in Nonaqueous Media. *Plant Physiol* **75**(3): 542-547.

Gilmour SJ, Fowler SG, Thomashow MF. 2004. Arabidopsis Transcriptional Activators CBF1, CBF2, and CBF3 have Matching Functional Activities. *Plant Molecular Biology* **54**(5): 767-781.

Gilmour SJ, Zarka DG, Stockinger EJ, Salazar MP, Houghton JM, Thomashow MF. 1998. Low temperature regulation of the Arabidopsis CBF family of AP2 transcriptional activators as an early step in cold-induced COR gene expression. *The Plant Journal* **16**(4): 433-442.

Graf A, Schlereth A, Stitt M, Smith AM. 2010. Circadian control of carbohydrate availability for growth in Arabidopsis plants at night. *Proceedings of the National Academy of Sciences* **107**(20): 9458-9463.

Guy C, Kaplan F, Kopka J, Selbig J, Hincha DK. 2008. Metabolomics of temperature stress. *Physiologia Plantarum* **132**(2): 220-235.

Hannah MA, Heyer AG, Hincha DK. 2005. A Global Survey of Gene Regulation during Cold Acclimation in Arabidopsis thaliana. *PLOS Genetics* **1**(2): e26.

Hannah MA, Wiese D, Freund S, Fiehn O, Heyer AG, Hincha DK. 2006. Natural genetic variation of freezing tolerance in Arabidopsis. *Plant Physiol* **142**(1): 98-112.

Hare PD, Cress WA. 1997. Metabolic implications of stress-induced proline accumulation in plants. *Plant Growth Regulation* **21**(2): 79-102.

Haydon MJ, Mielczarek O, Robertson FC, Hubbard KE, Webb AAR. 2013. Photosynthetic entrainment of the Arabidopsis thaliana circadian clock. *Nature* **502**: 689.

Heinig U, Gutensohn M, Dudareva N, Aharoni A. 2013. The challenges of cellular compartmentalization in plant metabolic engineering. *Current Opinion in Biotechnology* **24**(2): 239-246.

Henkel S, Nagele T, Hormiller I, Sauter T, Sawodny O, Ederer M, Heyer AG. 2011. A systems biology approach to analyse leaf carbohydrate metabolism in Arabidopsis thaliana. *EURASIP J Bioinform Syst Biol* **2011**(1): 2.

Hoermiller, II, Naegele T, Augustin H, Stutz S, Weckwerth W, Heyer AG. 2017. Subcellular reprogramming of metabolism during cold acclimation in Arabidopsis thaliana. *Plant Cell Environ* **40**(5): 602-610.

Horton MW, Willems G, Sasaki E, Koornneef M, Nordborg M. 2016. The genetic architecture of freezing tolerance varies across the range of Arabidopsis thaliana. *Plant Cell Environ* **39**(11): 2570-2579.

Horvath DP, McLarney BK, Thomashow MF. 1993. Regulation of Arabidopsis thaliana L.(Heyn) cor78 in response to low temperature. *Plant Physiol* **103**(4): 1047-1053.

Hossain MS, Persicke M, ElSayed AI, Kalinowski J, Dietz K-J. 2017. Metabolite profiling at the cellular and subcellular level reveals metabolites associated with salinity tolerance in sugar beet. *J Exp Bot* **68**(21-22): 5961-5976.

Huber SC, Huber JL. 1992. Role of sucrose-phosphate synthase in sucrose metabolism in leaves. *Plant Physiol* **99**(4): 1275-1278.

Hundertmark M, Hincha DK. 2008. LEA (Late Embryogenesis Abundant) proteins and their encoding genes in Arabidopsis thaliana. *BMC Genomics* **9**(1): 118.

Huner NPA, Öquist G, Sarhan F. 1998. Energy balance and acclimation to light and cold. *Trends in Plant Science* **3**(6): 224–230.

Hurry V. 2017. Metabolic reprogramming in response to cold stress is like real estate, it's all about location. *Plant Cell Environ* **40**(5): 599-601.

Jaglo-Ottosen KR, Gilmour SJ, Zarka DG, Schabenberger O, Thomashow MF. 1998. *Arabidopsis* CBF1 Overexpression Induces *COR* Genes and Enhances Freezing Tolerance. *Science* **280**(5360): 104-106.

Janmohammadi M, Zolla L, Rinalducci S. 2015. Low temperature tolerance in plants: Changes at the protein level. *Phytochemistry* **117**: 76-89.

Janská A, Maršík P, Zelenková S, Ovesná J. 2010. Cold stress and acclimation - what is important for metabolic adjustment? *Plant Biology* **12**(3): 395–405.

Jiang W, Yin Q, Wu R, Zheng G, Liu J, Dixon RA, Pang Y. 2015. Role of a chalcone isomerase-like protein in flavonoid biosynthesis in Arabidopsis thaliana. *J Exp Bot* **66**(22): 7165-7179.

Kaiser WM, Huber SC. 2001. Post-translational regulation of nitrate reductase: mechanism, physiological relevance and environmental triggers. *J Exp Bot* **52**(363): 1981-1989.

Kaplan F, Kopka J, Haskell DW, Zhao W, Schiller KC, Gatzke N, Sung DY, Guy CL. 2004. Exploring the Temperature-Stress Metabolome of Arabidopsis. *Plant Physiol* **136**(4): 4159-4168.

Kaplan F, Kopka J, Sung DY, Zhao W, Popp M, Porat R, Guy CL. 2007. Transcript and metabolite profiling during cold acclimation of Arabidopsis reveals an intricate relationship of cold-regulated gene expression with modifications in metabolite content. *Plant J* **50**(6): 967-981.

Klemens PA, Patzke K, Deitmer J, Spinner L, Le Hir R, Bellini C, Bedu M, Chardon F, Krapp A, Neuhaus HE. 2013. Overexpression of the vacuolar sugar carrier AtSWEET16

modifies germination, growth, and stress tolerance in Arabidopsis. *Plant Physiol* **163**(3): 1338-1352.

Klie S, Krueger S, Krall L, Giavalisco P, Flugge UI, Willmitzer L, Steinhauser D. 2011. Analysis of the compartmentalized metabolome - a validation of the non-aqueous fractionation technique. *Front Plant Sci* **2**: 55.

Knaupp M, Mishra KB, Nedbal L, Heyer AG. 2011. Evidence for a role of raffinose in stabilizing photosystem II during freeze-thaw cycles. *Planta* **234**(3): 477-486.

Knight MR, Knight H. 2012. Low-temperature perception leading to gene expression and cold tolerance in higher plants. *New Phytologist* **195**(4): 737-751.

Kosová K, Vitamvas P, Prasil IT, Renaut J. 2011. Plant proteome changes under abiotic stress--contribution of proteomics studies to understanding plant stress response. *J Proteomics* **74**(8): 1301-1322.

Kosová K, Vítámvás P, Urban MO, Prášil IT, Renaut J. 2018. Plant Abiotic Stress Proteomics: The Major Factors Determining Alterations in Cellular Proteome. *Front Plant Sci* **9**(122).

Krueger S, Giavalisco P, Krall L, Steinhauser MC, Bussis D, Usadel B, Flugge UI, Fernie AR, Willmitzer L, Steinhauser D. 2011. A topological map of the compartmentalized Arabidopsis thaliana leaf metabolome. *PLoS One* **6**(3): e17806.

Kueger S, Steinhauser D, Willmitzer L, Giavalisco P. 2012. High-resolution plant metabolomics: from mass spectral features to metabolites and from whole-cell analysis to subcellular metabolite distributions. *Plant J* **70**(1): 39-50.

Kurepin LV, Dahal KP, Savitch LV, Singh J, Bode R, Ivanov AG, Hurry V, Huner NP. 2013. Role of CBFs as integrators of chloroplast redox, phytochrome and plant hormone signaling during cold acclimation. *Int J Mol Sci* **14**(6): 12729-12763.

Linka N, Weber A. 2010. Intracellular metabolite transporters in plants. *Mol Plant* **3**(1): 21-53.

Liu Z, Jia Y, Ding Y, Shi Y, Li Z, Guo Y, Gong Z, Yang S. 2017. Plasma Membrane CRPK1-Mediated Phosphorylation of 14-3-3 Proteins Induces Their Nuclear Import to Fine-Tune CBF Signaling during Cold Response. *Molecular Cell* **66**(1): 117-128.e115.

Lundmark M, Cavaco AM, Trevanion S, Hurry V. 2006. Carbon partitioning and export in transgenic Arabidopsis thaliana with altered capacity for sucrose synthesis grown at low temperature: a role for metabolite transporters. *Plant Cell Environ* **29**(9): 1703-1714.

Lunn J. 2007. Compartmentation in plant metabolism. *J Exp Bot* **58**(1): 35-47.

Ma C, Xin M, Feldmann KA, Wang X. 2014a. Machine Learning–Based Differential Network Analysis: A Study of Stress-Responsive Transcriptomes in Arabidopsis. *Plant Cell* **26**(2): 520-537.

Ma C, Zhang HH, Wang X. 2014b. Machine learning for Big Data analytics in plants. *Trends in Plant Science* **19**(12): 798-808.

Ma L, Tian T, Lin R, Deng X-W, Wang H, Li G. 2016. Arabidopsis FHY3 and FAR1 regulate light-induced myo-inositol biosynthesis and oxidative stress responses by transcriptional activation of MIPS1. *Mol Plant* **9**(4): 541-557.

Maier T, Güell M, Serrano L. 2009. Correlation of mRNA and protein in complex biological samples. *FEBS Lett* **583**(24): 3966-3973.

Malinova I, Kunz HH, Alseekh S, Herbst K, Fernie AR, Gierth M, Fettke J. 2014. Reduction of the cytosolic phosphoglucomutase in Arabidopsis reveals impact on plant growth, seed and root development, and carbohydrate partitioning. *PLoS One* **9**(11): e112468.

Masakapalli SK, Le Lay P, Huddleston JE, Pollock NL, Kruger NJ, Ratcliffe RG. 2010. Subcellular flux analysis of central metabolism in a heterotrophic Arabidopsis cell suspension using steady-state stable isotope labeling. *Plant Physiol* **152**(2): 602-619.

McKhann HI, Gery C, Bérard A, Lévêque S, Zuther E, Hincha DK, De Mita S, Brunel D, Téoulé E. 2008. Natural variation in CBF gene sequence, gene expression and freezing tolerance in the Versailles core collection of Arabidopsis thaliana. *BMC Plant Biol* **8**(1): 105.

Medeiros DB, Barros K, Barros JA, Omena-Garcia RP, Arrivault S, Vincis Pereira Sanglard L, Detmann KC, Silva WB, Daloso DM, DaMatta F, et al. 2017. Impaired malate and fumarate accumulation due the mutation of tonoplast dicarboxylate transporter. *Plant Physiol* **175**(3): 1068-1081.

Mintz-Oron S, Meir S, Malitsky S, Ruppin E, Aharoni A, Shlomi T. 2012. Reconstruction of Arabidopsis metabolic network models accounting for subcellular compartmentalization and tissue-specificity. *Proc Natl Acad Sci U S A* **109**(1): 339-344.

Mittler R. 2006. Abiotic stress, the field environment and stress combination. *Trends in Plant Science* **11**(1): 15-19.

Miura K, Furumoto T. 2013. Cold Signaling and Cold Response in Plants. *International Journal of Molecular Sciences* **14**(3): 5312-5337.

Mjolsness E, DeCoste D. 2001. Machine Learning for Science: State of the Art and Future Prospects. *Science* **293**(5537): 2051-2055.

Moore B, Zhou L, Rolland F, Hall Q, Cheng W-H, Hwang I, Jones T, Sheen J. 2003. Role of the Arabidopsis Glucose Sensor HXK1 in Nutrient, Light, and Hormonal Signaling. *Science* **300**(5617): 332-336.

Nägele T. 2014. Linking metabolomics data to underlying metabolic regulation. *Front Mol Biosci* **1**: 22.

Nägele T, Heyer AG. 2013. Approximating subcellular organisation of carbohydrate metabolism during cold acclimation in different natural accessions of Arabidopsis thaliana. *New Phytol* **198**(3): 777-787.

Nägele T, Kandel BA, Frana S, Meissner M, Heyer AG. 2011. A systems biology approach for the analysis of carbohydrate dynamics during acclimation to low temperature in Arabidopsis thaliana. *FEBS J* **278**(3): 506-518.

Nägele T, Mair A, Sun X, Fragner L, Teige M, Weckwerth W. 2014. Solving the differential biochemical Jacobian from metabolomics covariance data. *PLoS One* **9**(4): e92299.

Nägele T, Stutz S, Hörmiller II, Heyer AG. 2012. Identification of a metabolic bottleneck for cold acclimation in Arabidopsis thaliana. *Plant J* **72**(1): 102-114.

Nagler M, Nukarinen E, Weckwerth W, Nägele T. 2015. Integrative molecular profiling indicates a central role of transitory starch breakdown in establishing a stable C/N homeostasis during cold acclimation in two natural accessions of Arabidopsis thaliana. *BMC Plant Biol* **15**: 284.

Nordborg M, Bergelson J. 1999. The effect of seed and rosette cold treatment on germination and flowering time in some Arabidopsis thaliana (Brassicaceae) ecotypes. *American Journal of Botany* **86**: 470–475.

Noren L, Kindgren P, Stachula P, Ruhl M, Eriksson ME, Hurry V, Strand A. 2016. Circadian and Plastid Signaling Pathways Are Integrated to Ensure Correct Expression of the CBF and COR Genes during Photoperiodic Growth. *Plant Physiol* **171**(2): 1392-1406.

Örvar BL, Sangwan V, Omann F, Dhindsa RS. 2000. Early steps in cold sensing by plant cells: the role of actin cytoskeleton and membrane fluidity. *The Plant Journal* **23**(6): 785-794.

Pommerrenig B, Ludewig F, Cvetkovic J, Trentmann O, Klemens PAW, Neuhaus HE. 2018. In concert: orchestrated changes in carbohydrate homeostasis are critical for plant abiotic stress tolerance. *Plant and Cell Physiology*: pcy037-pcy037.

Prasch CM, Sonnewald U. 2015. Signaling events in plants: stress factors in combination change the picture. *Environmental and Experimental Botany* **114**: 4-14.

Provart NJ, Alonso J, Assmann SM, Bergmann D, Brady SM, Brkljacic J, Browse J, Chapple C, Colot V, Cutler S, et al. 2016. 50 years of Arabidopsis research: highlights and future directions. *New Phytol* **209**(3): 921-944.

Puhakainen T, Hess MW, Makela P, Svensson J, Heino P, Palva ET. 2004. Overexpression of multiple dehydrin genes enhances tolerance to freezing stress in Arabidopsis. *Plant Mol Biol* **54**(5): 743-753.

Reznik E, Segre D. 2010. On the stability of metabolic cycles. *J Theor Biol* **266**(4): 536-549.

Riens B, Lohaus G, Heineke D, Heldt HW. 1991. Amino Acid and sucrose content determined in the cytosolic, chloroplastic, and vacuolar compartments and in the Phloem sap of spinach leaves. *Plant Physiol* **97**(1): 227-233.

Ristic Z, Ashworth EN. 1993. Changes in leaf ultrastructure and carbohydrates in Arabidopsis thaliana L. (Heyn) cv. Columbia during rapid cold acclimation. *Protoplasma* **172**: 111–123.

Rohwer JM. 2012. Kinetic modelling of plant metabolic pathways. *J Exp Bot* **63**(6): 2275-2292.

Rolland F, Baena-Gonzalez E, Sheen J. 2006. Sugar sensing and signaling in plants: conserved and novel mechanisms. *Annu Rev Plant Biol* **57**: 675-709.

Schaber J, Liebermeister W, Klipp E. 2009. Nested uncertainties in biochemical models. *IET systems biology* **3**(1): 1-9.

Schneider T, Keller F. 2009. Raffinose in Chloroplasts is Synthesized in the Cytosol and Transported across the Chloroplast Envelope. *Plant and Cell Physiology* **50**(12): 2174-2182.

Schulz E, Tohge T, Zuther E, Fernie AR, Hincha DK. 2016. Flavonoids are determinants of freezing tolerance and cold acclimation in Arabidopsis thaliana. *Sci Rep* **6**: 34027.

Shi Y, Ding Y, Yang S. 2015. Cold signal transduction and its interplay with phytohormones during cold acclimation. *Plant Cell Physiol* **56**(1): 7-15.

Sicher R. 2011. Carbon partitioning and the impact of starch deficiency on the initial response of Arabidopsis to chilling temperatures. *Plant Science* **181**(2): 167-176.

Sperschneider J, Dodds PN, Singh KB, Taylor JM. 2018. ApoplastP: prediction of effectors and plant proteins in the apoplast using machine learning. *New Phytologist* **217**(4): 1764-1778.

Steuer R. 2007. Computational approaches to the topology, stability and dynamics of metabolic networks. *Phytochemistry* **68**(16-18): 2139-2151.

Steuer R, Gross T, Selbig J, Blasius B. 2006. Structural kinetic modeling of metabolic networks. *Proc Natl Acad Sci U S A* **103**(32): 11868-11873.

Stitt M, Lilley RM, Gerhardt R, Heldt HW. 1989. Metabolite levels in specific cells and subcellular compartments of plant leaves. *Methods in Enzymology* **174**: 518–550.

Stitt M, Wilke I, Feil R, Heldt HW. 1988. Coarse control of sucrose-phosphate synthase in leaves: Alterations of the kinetic properties in response to the rate of photosynthesis and the accumulation of sucrose. *Planta* **174**(2): 217–230.

Stitt M, Zeeman SC. 2012. Starch turnover: pathways, regulation and role in growth. *Curr Opin Plant Biol* **15**(3): 282-292.

Strand A, Foyer CH, Gustafsson P, Gardeström P, Hurry V. 2003. Altering flux through the sucrose biosynthesis pathway in transgenic Arabidopsis thaliana modifies photosynthetic acclimation at low temperatures and the development of freezing tolerance. *Plant, Cell and Environment* **26**: 523–535.

Strand A, Hurry V, Gustafsson P, Gardeström P. 1997. Development of Arabidopsis thaliana leaves at low temperatures releases the suppression of photosynthesis and photosynthetic gene expression despite the accumulation of soluble carbohydrates. *Plant Journal* **12**(3): 605–614.

Strand A, Hurry V, Henkes S, Huner N, Gustafsson P, Gardeström P, Stitt M. 1999. Acclimation of Arabidopsis leaves developing at low temperatures. Increasing cytoplasmic volume accompanies increased activities of enzymes in the Calvin cycle and in the sucrose-biosynthesis pathway. *Plant Physiol* **119**: 1387–1397.

Sturm A. 1999. Invertases. Primary structures, functions, and roles in plant development and sucrose partitioning. *Plant Physiol* **121**(1): 1-8.

Sweetlove LJ, Fernie AR. 2013. The spatial organization of metabolism within the plant cell. *Annual Review of Plant Biology* **64**: 723-746.

Szabados L, Savoure A. 2010. Proline: a multifunctional amino acid. *Trends in Plant Science* **15**(2): 89-97.

Szecowka M, Heise R, Tohge T, Nunes-Nesi A, Vosloh D, Huege J, Feil R, Lunn J, Nikoloski Z, Stitt M, et al. 2013. Metabolic fluxes in an illuminated Arabidopsis rosette. *Plant Cell* **25**(2): 694-714.

Szklarczyk D, Franceschini A, Wyder S, Forslund K, Heller D, Huerta-Cepas J, Simonovic M, Roth A, Santos A, Tsafou KP. 2014. STRING v10: protein–protein interaction networks, integrated over the tree of life. *Nucleic Acids Research* **43**(D1): D447-D452.

Tarkowski LP, Van den Ende W. 2015. Cold tolerance triggered by soluble sugars: a multifaceted countermeasure. *Front Plant Sci* **6**: 203.

Thalmann M, Santelia D. 2017. Starch as a determinant of plant fitness under abiotic stress. *New Phytologist* **214**(3): 943-951.

Thomashow MF. 1999. PLANT COLD ACCLIMATION: Freezing Tolerance Genes and Regulatory Mechanisms. *Annual Review of Plant Physiology and Plant Molecular Biology* **50**: 571–599.

Thomashow MF. 2001. So What's New in the Field of Plant Cold Acclimation? Lots! *Plant Physiol* **125**(1): 89-93.

Thomashow MF. 2010. Molecular basis of plant cold acclimation: insights gained from studying the CBF cold response pathway. *Plant Physiol* **154**(2): 571-577.

Tiessen A, Padilla-Chacon D. 2012. Subcellular compartmentation of sugar signaling: links among carbon cellular status, route of sucrolysis, sink-source allocation, and metabolic partitioning. *Front Plant Sci* **3**: 306.

Timperio AM, Egidi MG, Zolla L. 2008. Proteomics applied on plant abiotic stresses: Role of heat shock proteins (HSP). *Journal of Proteomics* **71**(4): 391-411.

van Buer J, Cvetkovic J, Baier M. 2016. Cold regulation of plastid ascorbate peroxidases serves as a priming hub controlling ROS signaling in Arabidopsis thaliana. *BMC Plant Biol* **16**(1): 163.

Van Hasselt P. 1990. Light-induced damage during chilling. *Chilling injury of horticultural crops*: 113-128.

Wanner LA, Junttila O. 1999. Cold-induced freezing tolerance in Arabidopsis. *Plant Physiol* **120**(2): 391–400.

Weiszmann J, Fürtauer L, Weckwerth W, Nägele T. 2017. Vacuolar invertase activity shapes photosynthetic stress response of Arabidopsis thaliana and stabilizes central energy supply. *bioRxiv*.

Winkel-Shirley B. 2001. Flavonoid biosynthesis. A colorful model for genetics, biochemistry, cell biology, and biotechnology. *Plant Physiol* **126**(2): 485-493.

Wormit A, Trentmann O, Feifer I, Lohr C, Tjaden J, Meyer S, Schmidt U, Martinoia E, Neuhaus HE. 2006. Molecular Identification and Physiological Characterization of a Novel Monosaccharide Transporter from *Arabidopsis* Involved in Vacuolar Sugar Transport. *Plant Cell* **18**(12): 3476-3490.

Xiang L, Le Roy K, Bolouri-Moghaddam MR, Vanhaecke M, Lammens W, Rolland F, Van den Ende W. 2011. Exploring the neutral invertase-oxidative stress defence connection in Arabidopsis thaliana. *J Exp Bot* **62**(11): 3849-3862.

Yan S-P, Zhang Q-Y, Tang Z-C, Su W-A, Sun W-N. 2006. Comparative Proteomic Analysis Provides New Insights into Chilling Stress Responses in Rice. *Molecular & Cellular Proteomics* **5**(3): 484-496.

Yonekura-Sakakibara K, Tohge T, Matsuda F, Nakabayashi R, Takayama H, Niida R, Watanabe-Takahashi A, Inoue E, Saito K. 2008. Comprehensive Flavonol Profiling and Transcriptome Coexpression Analysis Leading to Decoding Gene–Metabolite Correlations in Arabidopsis. *Plant Cell* **20**(8): 2160-2176.

Zhen Y, Ungerer MC. 2008. Clinal variation in freezing tolerance among natural accessions of Arabidopsis thaliana. *New Phytologist* **177**(2): 419-427.

Zuther E, Juszczak I, Ping Lee Y, Baier M, Hincha DK. 2015. Time-dependent deacclimation after cold acclimation in Arabidopsis thaliana accessions. *Sci Rep* **5**: 12199.

Zuther E, Schulz E, Childs LH, Hincha DK. 2012. Clinal variation in the non-acclimated and cold-acclimated freezing tolerance of Arabidopsis thaliana accessions. *Plant Cell Environ* **35**(10): 1860-1878.