Metabolites **2012**, *2*, 553-566; doi:10.3390/metabo2030553

OPEN ACCESS

metabolites

ISSN 2218-1989

www.mdpi.com/journal/metabolites/

Review

Mathematical Modeling of Plant Metabolism—From Reconstruction to Prediction

Thomas Nägele * and Wolfram Weckwerth

Department of Molecular Systems Biology, University of Vienna, Althanstraße 14, 1090 Vienna, Austria; E-Mail: Wolfram.Weckwerth@univie.ac.at (W.W.)

* Author to whom correspondence should be addressed; E-Mail: Thomas.Naegele@univie.ac.at; Tel.: +43-1-4277-57702.

Received: 7 July 2012; in revised form: 22 August 2012 / Accepted: 28 August 2012 / Published: 6 September 2012

Abstract: Due to their sessile lifestyle, plants are exposed to a large set of environmental cues. In order to cope with changes in environmental conditions a multitude of complex strategies to regulate metabolism has evolved. The complexity is mainly attributed to interlaced regulatory circuits between genes, proteins and metabolites and a high degree of cellular compartmentalization. The genetic model plant *Arabidopsis thaliana* was intensely studied to characterize adaptive traits to a changing environment. The availability of genetically distinct natural populations has made it an attractive system to study plant-environment interactions. The impact on metabolism caused by changing environmental conditions can be estimated by mathematical approaches and deepens the understanding of complex biological systems. In combination with experimental high-throughput technologies this provides a promising platform to develop *in silico* models which are not only able to reproduce but also to predict metabolic phenotypes and to allow for the interpretation of plant physiological mechanisms leading to successful adaptation to a changing environment. Here, we provide an overview of mathematical approaches to analyze plant metabolism, with experimental procedures being used to validate their output, and we discuss them in the context of establishing a comprehensive understanding of plant-environment interactions.

Keywords: plant systems biology; mathematical modeling; *Arabidopsis thaliana*; natural variation; plant-environment interaction; stoichiometric modeling

---

1. Introduction

The performance and distribution of plants is significantly affected by several environmental factors, like for example temperature, drought and soil salinity. Due to their sessile lifestyle, plants have to cope with spontaneous or seasonal changes in these environmental factors in order to survive. To a certain degree, numerous plant species are able to respond to such changes by re-adjustment of their metabolism—a process that is in general termed as acclimation. It comprises a multitude of biochemical and physiological changes, ultimately leading to an increase in the capacity of the plants to cope with environmental stress. One prominent example, which was amongst others intensely studied in the model plant *Arabidopsis thaliana*, is the acclimation to low temperature leading to an improved cold tolerance. For many herbaceous plant species it was shown that they can grow at low temperature and even survive freezing [1–3]. It was demonstrated that freezing tolerance is a multigenic trait influenced by multiple factors comprising changes in gene expression, protein abundance, enzyme activity, metabolite concentrations and membrane structure [4–9]. Particularly, reprogramming of primary metabolism affects photosynthetic activity, accumulation of soluble sugars, certain amino acids and polyamines, indicating a complex relationship between metabolic consequences of low temperature. Therefore, it is not surprising that although the presence of certain sugars like e.g. sucrose or raffinose is well known to correlate with winter hardiness in many plant species [9,10], this accumulation alone is insufficient to explain the development of freezing tolerance [11]. Additionally, it is not clear whether sugars accumulate as cryoprotective substances or whether they are substrates for the synthesis of cryoprotectants, or even just as a consequence of growth retardation, which is stronger than reduction of photosynthetic activity at low temperature [12,13].

The analysis of complex metabolic processes involved in acclimation of plant metabolism to environmental stress significantly benefits from the availability of genetically distinct natural populations of the model plant *Arabidopsis thaliana*. This species is natural to Europe and central Asia, and the climate on a global scale was shown to be sufficient for shaping its range boundaries [14]. Arabidopsis has a comparatively large climatic amplitude and is spread over a latitudinal range from 68°N to 0°N, which makes it suitable for the analysis of variation in adaptive traits [15,16]. Due to being a predominantly selfing species, most individual Arabidopsis plants collected in nature are homozygous inbred lines and are commonly referred to as accessions which are specialized to particular sets of environmental conditions. The usefulness of accessions in exploring plant cold acclimation mechanisms was exemplified by Hannah and co-workers who demonstrated that the freezing tolerance of nine natural accessions, originating from Scandinavia to the Cape Verde Islands, correlates with habitat winter temperatures [6]. Additionally, they showed that low temperature represents an important selective pressure for Arabidopsis.

Besides such complex plant-environment interactions, latest developments in bioanalytical research comprising shotgun and next-generation genome sequencing as well as molecular analysis using OMICS technologies have driven the need for computer-assisted analysis and modeling of biological data. Systems biology has evolved in a research field focusing on the system wide understanding of biological networks, like for example the cellular metabolism in a photosynthetically active plant cell. In a systems biology approach, network elements, such as genes, proteins or metabolites, are considered as interacting components rather than isolated entities in order to deepen the comprehensive

understanding of the organization of a complex biological system. A promising way to analyze such complex biological and biochemical networks is formal representation by mathematical models enabling their computer based handling and making biological data accessible to theoretical methods originating from applied mathematics and systems theory. Numerous mathematical approaches to model plant metabolic networks have been suggested and discussed, both relying on and emphasizing the importance of the iterative processes of model development, simulation and validation by experimental data [17–22]. An overview of several computational approaches to study metabolic networks with respect to features like topology, stability and dynamical behavior was provided previously [23], classifying the approaches by their mathematical structure. Thus, qualitative models for a static network description containing no kinetic parameters (network/stoichiometric analysis) were distinguished from quantitative approaches applying a dynamic system description using kinetic parameters (kinetic models) [23].

A main focus of mathematical modeling in biochemistry and plant science has been on the construction of kinetic models where metabolic states are simulated based on the knowledge about network topology, stoichiometry, rate equations and kinetic parameters. Typically, a system of ordinary differential equations (ODEs) is used to describe the time-dependent changes in state variables, *i.e.,* metabolite concentration, protein abundance or the amount of transcripts. In the context of metabolic systems, the sum of synthesizing and degrading rates of enzyme reactions defines the time-dependent change in metabolite concentrations. The representation of biological systems by sets of ODEs has successfully been applied to various processes in plant biology and also in the comprehensive analysis of plant-environment interactions, as was already outlined in [24]. While kinetic modeling represents an attractive method to study and beyond that, to potentially predict the behavior of complex metabolic networks, plenty of information about the network topology and the kinetics of metabolite interconverting steps is required for model development and its experimental validation [25]. For example, although rates of enzymatic steps can in general be reliably approximated by Michaelis-Menten kinetics, in many cases a probable influence of allosteric effectors is not confirmed but may have a dramatic effect on systems dynamics, and negligence will have a significant impact on the model output, *i.e.*, the result of model simulation [26]. Additionally, in many cases mechanisms of allosteric regulation are known but quantitative experiments on parameters like substrate affinity (K_M) or inhibitory constants (K_i) are lacking. This enforces the application of parameter estimation to calculate parameter values which are either completely unknown or can be estimated within numerical bounds based on published data on a different condition or organism. Indeed, such assumptions cause uncertainties, which have to be discussed carefully when interpreting the model output. However, although there might be several uncertainties with respect to regulatory instances involved in every single reaction of metabolism, numerous studies have proven kinetic modeling to be a promising approach to comprehensively analyze complex processes in plant biology. An overview of applications is given by Schallau and Junker [27] exemplarily comprising the process of photosynthesis [28], leaf carbon metabolism [29], sucrose metabolism in sugar cane (*Saccharum officinarum*) [30] or the aspartic acid-derived amino acid pathway in *Arabidopsis thaliana* [31].

In contrast to kinetic modeling, the approach of structural modeling is based on the idea of constructing models without kinetic information. This modeling approach refers only to the stoichiometry of the reactions within the system which is summarized in the stoichiometric matrix **N**.

Considering a metabolic reaction network, each column of **N** represents a reaction while rows represent metabolites. Hence, the elements of **N** describe the stoichiometric coefficients of metabolites in the reactions. Positive entries indicate that the metabolite is produced by the reaction, while negative values indicate consumption. Entries of zero indicate that the metabolite is not involved in this reaction. The definition of a vector *v* containing the rates of metabolite interconversions allows for the description of the steady state of the metabolic reaction network by a set of differential equations:

$$\frac{dM}{dt} = Nv = 0 \tag{1}$$

where **M** represents a matrix containing metabolite concentrations and **t** is time. Solutions of this equation can be calculated applying linear algebraic rules. The advantage of this approach becomes obvious when considering the large number of reactions in a metabolic system, which can be predicted from an annotated genome sequence. The workflow from metabolic reconstruction to modeling of a metabolic network based on an annotated genome sequence was previously described in detail [32]. Sequence data of a genome are derived from techniques like next-generation sequencing, which are capable of sequencing a whole prokaryotic or eukaryotic genome within days. The assembled complete genome sequence is analyzed for intron/exon structure, start and stop codons as well as homology with known sequences. A genome-scale functional annotation based on the homology with functionally characterized genes from other organisms results in a gene list, which is used to predict enzymatic reactions. Based on the substrates and products of the enzymatic reactions, pathways are structured and a stoichiometric matrix can be derived. Applying this approach, it becomes possible to analyze a large set of metabolic interactions simultaneously, and results of experimental high-throughput studies on the metabolome and proteome are a promising way to validate the model output in metaproteogenomic studies as demonstrated for *Chlamydomonas reinhardtii* [33]. Yet, in context of metabolism of higher plants the main challenges of this modeling approach result from the genome content whose function remains undiscovered, and also from characteristics like subcellular compartmentation and tissue differentiation making the analysis of higher plants much more complex than in prokaryotic organisms [34].

Focusing now again on the complexity of plant-environment interactions and the variability of stress responses in natural accessions of *Arabidopsis thaliana*, both kinetic and stoichiometric modeling represent promising approaches to comprehensively study regulatory instances in plant metabolism, its re-adjustment after environmental perturbations or even the impact of changes in transcriptional control on the metabolome. On the other hand, both applications of mathematical modeling are significantly limited in their ability to reconstruct and predict the behavior of plant metabolism *in vivo*. Kinetic modeling typically focuses on a relatively small part of metabolism and aims at simplification to constrain the complexity and amount of kinetic information, which is needed to simulate network dynamics. In contrast, approaches of stoichiometric modeling focus on the comprehensive compilation of network interaction, ultimately aiming at the complete representation of metabolism, yet neglecting kinetic information and the estimation of non-linear network dynamics. Motivated by the limitations of each of these methods, the following sections are intended to summarize current progress in mathematical modeling of plant metabolism and to figure out its potential to analyze and predict complex plant-environment interactions.

2. Estimating Dynamics of Plant Metabolism Due to Environmental Perturbation

The mathematical analysis of plant metabolism first of all relies on the representation by a model, which is constructed based on information on biochemical pathways and the interaction of pathway components gained from numerous previous experimental studies. Typically, the first step of model construction consists of a graphical representation of the pathway or network of interest. In case of a kinetic model of metabolism, this may result in a map, which connects nodes by lines. Every node now represents a metabolite and every line describes a metabolite interconversion, *i.e.,* enzymatic reaction. Based on this graphical representation, the mathematical model is then derived by translating nodes and lines in metabolite concentrations and enzymatic rate laws, for example the Michaelis-Menten equation. These rate equations are characterized by kinetic parameters like enzymatic substrate affinity, *i.e.,* the Michaelis-Menten constant K_M, and the maximum enzyme activity v_{max}. This process is crucial for the successful modeling approach as all further steps of mathematical analysis rely on these assumptions: if the interaction between two network components is described by equations or parameters which do not agree with confirmed experimental results, validation of simulation results by experimental data is not reliable anymore and the model becomes unfeasible. Although a vast number of metabolic interactions have intensively been characterized and many underlying laws of interaction are well known, like for example the Michaelis-Menten kinetics, deriving the most realistic model structure of a metabolic network becomes difficult when assumptions about simplification have to be made. This is frequently the case for kinetic models, based on systems of ODEs, which are intended to provide an insight into the dynamics of metabolism. These dynamics are predominantly nonlinear and model systems are often characterized by a high-dimensional parameter space. Kinetic parameters, characterizing substrate affinity (K_M) or inhibition (K_i), are often not directly accessible to experimental measurements. In addition to the everlasting question how results of *in vitro* measurements differ from *in vivo* data, experimental conditions, like the temperature or pH, significantly constrain their validity. Hence, besides the determination of a model structure, the process of mathematical identification of unknown kinetic parameters represents another crucial step in building a realistic ODE-based model to simulate dynamics of plant metabolism. To reduce the complexity and also the number of unknown kinetic parameters, individual enzymatic steps might be summarized in blocks of interconversions directly linking the metabolite concentrations that have been quantified. These blocks of interconversion are confined by the rate-limiting steps, *i.e.,* the enzymatic reaction representing a regulatory bottleneck for the synthesis/degradation of a metabolite. Measurement on the kinetic parameters of the corresponding enzymes then allows for the estimation of the kinetic characteristics of this metabolic pathway. This approach was recently applied to the analysis of diurnal dynamics of the central carbohydrate metabolism in leaves of *Arabidopsis thaliana* [35]. For modeling, the authors used a simplified scheme of central carbohydrate metabolism, focusing on the most abundant sugars and phosphorylated intermediates at which pathways branch. In this simplified model, subcellular compartmented pathways were lumped together and aerial organs of the vegetative plant were collectively interpreted as a source of carbon fixation that exports carbon to sink organs, for example roots [35]. The modeling approach was applied to wild type plants as well as mutant plants defective in the dominating vacuolar invertase AtβFruct4 (At1G12240) to analyze the physiological role of AtβFruct4 on whole plant carbon metabolism. Although this approach is based on assumptions

significantly simplifying the *in planta* metabolic network, the authors were successful in reproducing the experimentally observed data on metabolite concentrations. Furthermore, simulation results allowed for the evaluation of flux rates in the central carbohydrate metabolism revealing a significant impact of invertase activity on sink-source interaction and buffering metabolite concentrations against changes in environmental conditions [35]. The usefulness of this modeling approach to study dynamics of metabolism induced by environmental perturbations was underpinned by application to the analysis of the regulation of the central carbohydrate metabolism during cold acclimation [22]. In this study, the authors applied an extended version of the mathematical model presented in [35] to analyze diurnal dynamics in carbohydrate metabolism of natural accessions of *Arabidopsis thaliana*, now also comprising the central steps of raffinose interconversion. Based on results of model simulation, a critical temperature for sucrose synthesis in a cold sensitive accession could be predicted at which an imbalance in photosynthetic carbon fixation is caused, ultimately resulting in oxidative stress [22]. It was concluded that metabolic capacities contribute to the ability of accessions of *Arabidopsis thaliana* to cope with changes in environmental conditions at low temperature.

As exemplified by these approaches of mathematical modeling, realistic output of model simulations can be expected despite a significant simplification of the model structure. This was also proven by many other kinetic modeling approaches (for an overview of several approaches please refers to [20]). Such simplifications may be performed in order to reduce the number of unknown model parameters and to minimize ambiguity of the model output. This ambiguity occurs due to uncertainties concerning model parameters as well as experimental data, kinetics and model structure. These different types of uncertainty are interlaced because uncertain network structures contain uncertain reaction kinetics that are characterized by uncertain parameters [36]. Assuming that both the model structure and kinetic laws are known, then parameter values have to be estimated allowing for the simulation of experimental data, for example metabolite concentrations. While uncertainty of unknown or marginally characterized parameters will be huge, extensive experiments on parameters will significantly contribute to unambiguous simulation results. Additionally, experimental data on dynamics of the model components, like for example time courses of metabolites, will contribute to minimization of parameter uncertainty: only parameter sets allowing for the successful simulation of the time-course will be approved. The parameter estimation of nonlinear dynamic modeling approaches can be classified as a nonlinear programming problem being subject to nonlinear differential-algebraic constraints [37]. In general, this mathematical problem can be formulated as follows:

$$Z = \int_0^t \left(y_{exp}(t) - y_{pred}(p,t) \right)^T W(t) \left(y_{exp}(t) - y_{pred}(p,t) \right) dt \qquad (2)$$

where Z represents the cost function to be minimized, y_{exp} contains experimentally determined state variables (for example metabolite concentrations), $y_{pred}(p,t)$ is the model prediction of state variables depending on estimated parameters p and time t, and W(t) is the weighting matrix containing information about the level of importance of single state variables and determining their influence on the cost function. This optimization problem of minimization of Z is subject to the differential/algebraic equality constraints describing the systems dynamics and additional requirements

for system performance. Additionally, the estimation of model parameter p is subject to lower (p^{low}) and upper (p^{up}) bounds:

$$p^{low} \leq p \leq p^{up} \tag{3}$$

Due to nonlinearities in objective function and constraints, solving these optimization problems frequently means having to cope with multimodality, *i.e.,* the potential existence of multiple local solutions [37,38]. This implies the application of algorithms, which are able to overcome local minima to ultimately yield the best solution, *i.e.,* the global optimum. Gradient-based local optimization methods fail to reliably determine the global optimum in multimodal problems because of nonconvexity arising from the previously mentioned nonlinearities. A graphical representation of this problem is shown in [38]. In a simple example it was demonstrated that even with only two decision variables, e.g., unknown kinetic model parameters, multimodal surfaces may result from optimization problems, *i.e.,* surfaces of the cost function with multiple peaks and valleys, which do not allow for the determination of one unique optimal solution by local optimization methods. Solving such multimodal problems is the goal of global optimization [39], which was discussed and reviewed in the context of parameter estimation in biochemical pathways [37]. One example for a global optimization method is the particle swarm pattern search method for bound constrained global optimization [40]. This algorithm was shown to be highly competitive with other global optimization methods and is a demonstrative example of how possible nonconvexity of the objective function can be globally explored. The basic idea behind this approach is to construct a hybrid of a pattern search method and a particle swarm search [40]. A particle swarm algorithm attempts to simulate the social behavior of a population of particles to explore a given problem space [41]. For every iteration step of the optimization process, the particles are associated with a stochastic velocity vector indicating the particles' direction of movement. The velocity vector for each particle is a linear stochastic combination of the velocity at the previous time instant, of the direction to the particle's best position, and of the direction to the best swarm position. The new position of every particle is calculated by adding the current velocity vector to the old particle position. The stopping criterion for the algorithm may then be defined by a tolerance level of velocities, which has to be reached for all particles. While pattern search methods are designed to achieve convergence from arbitrary starting points to points satisfying necessary conditions for local optimality [42], the incorporation of a particle swarm search in the search step of a pattern search method enables the attraction of local optima and the identification of global optima to be overcome [40].

Due to its capability to develop methods for comprehensive analysis of complex data sets and provide strategies of how to solve nonlinear problems, optimization theory represents an essential component for mathematical modeling of plant metabolism and other biological systems. Beyond that, the prediction of metabolism from first principles only becomes possible by application of optimization approaches [43].

3. Modeling on a Large Scale—Reconstruction of Metabolic Networks and Validation of Predictions by Metabolomics Science

Reconstruction of metabolic networks is based on information about whole genome sequences finally resulting in the stoichiometric matrix **N** of the network, which provides the basis for all modeling approaches [32]. As described in the previous section, particularly in kinetic modeling approaches, this information is frequently reduced in order to minimize complexity and unambiguous model outputs. In contrast, stoichiometric modeling approaches aim at the compilation and integration of the entire stoichiometric information of the metabolic network. Numerous missing enzyme parameters prevent comprehensive analysis by kinetic modeling, yet determination of steady-state solutions for the metabolic network is possible by solving equation (1) numerically. Compared to the complex analysis of nonlinear dynamical systems, this system of linear equations can easily be solved. However, the complexity of such an approach is indicated by the comprehensive reconstruction process as well as the experimental validation, revealing the need for permanent improvement of published metabolic network reconstructions by biochemist experts' knowledge and proteogenomic methods [33,44,45]. In a detailed protocol, Thiele and Palsson described the complex reconstruction process within four major steps leading to a metabolic network model [46]. Based on an automated draft reconstruction, the network model is iteratively refined, converted into a computable format and evaluated by the comparison of model predictions with experimental results on physiology, biochemistry or genetics. Although the process of reconstruction is identical for prokaryotic and eukaryotic metabolic networks, the authors emphasize that in eukaryotic systems, e.g., metabolism of higher plants, it is more challenging due to the size of genomes and cellular compartmentation [46]. Additional complexity arises from network gaps and mass-balance errors resulting from incomplete genome annotation and reaction stoichiometry errors which severely affect the predictive power of network models [47]. Beyond that, model simulations provide only information about a steady state, *i.e.,* a snapshot, of the system, which is pre-defined by the experimental design. Recently, in several studies genome-scale metabolic modeling in *Arabidopsis thaliana* was applied to address questions like ATP demand for growth and maintenance [21], the metabolic activity of key enzymes responsible for the supply of redox equivalents in plastids during the photorespiratory cycle [48] or to predict the design of genetic manipulations that are expected to increase vitamin E content in metabolically engineered seed strains [49].

With respect to such comprehensive metabolic network simulations, quantitative measurement of metabolism is necessary to validate the output of such simulations, which can be accomplished applying bioanalytical methods in metabolomics science [50]. Mass spectrometry is one of the crucial technologies in this field, and an overview of different techniques in context with their characteristic features has recently been presented [32]. A recent development is the use of two-dimensional gas chromatography coupled with fast acquisition rate time-of-flight mass spectrometry (GC x GC-TOF-MS). The coupling of two gas chromatography columns with different characteristics, for example a hydrophobic and a polar column, increases the separation efficiency of a complex metabolomics sample. A complete strategy to perform a convenient data extraction and alignment using two-dimensional gas chromatography coupled with mass spectrometry (GC x GC-MS) technology is already available [51]. Another important extension of current metabolomics platforms for

metabolomics is the integration of gas chromatography coupled to mass spectrometry (GC-MS) with liquid chromatography coupled to mass spectrometry (LC-MS) [52]. This approach enables the analysis of components of the primary metabolism by GC-MS, for example carbohydrates and amino acids, and higher molecular masses by LC-MS, e.g., secondary metabolites [53,54]. Beyond the development of techniques and new platforms, the improvement of databases, experimental standards and data compatibility among different laboratories is crucial for efficient metabolomics science [55].

The analysis of metabolomics results on numerous biological replicates under different environmental conditions or with genotypic variation represents a multidimensional task and results in a complex data matrix. The covariance matrix C results from multivariate statistics representing a central result of the experiments [32,56–59]. The observed covariance matrix C of metabolite concentrations is linked to the underlying biochemical system and the corresponding genotype by a systematic approach, which is characterized by the following equation [60]:

$$CJ^T + JC = -2D \tag{4}$$

In this equation, J represents the Jacobian matrix and D is the fluctuation/diffusion matrix. The diagonal entries D_{ii} characterize the magnitude of fluctuations of each metabolite, whereas off-diagonal entries D_{ij} ($i \neq j$) represent the fluctuation of metabolites caused by the interaction between enzymes i and j. The interconnection between metabolic networks and the Jacobian Matrix as well as the fluctuation matrix is described in detail elsewhere [32,60,61]. In general, the Jacobian matrix characterizes the local dynamics at a steady state condition. In the context of metabolic networks, the entries of the Jacobian J represent the elasticities of reaction rates to any change of the metabolite concentrations being characterized by the following equation:

$$J = N\frac{\partial r}{\partial M} \tag{5}$$

Here, N is the stoichiometric matrix, r represents the rates for each reaction and M is the metabolite concentration. Based on equations (4) and (5), an approach of inverse calculation of a Jacobian from metabolomics covariance data was recently derived [59]. Additionally, the authors developed the differential Jacobian, dJ_{ij}, defining the relative change of two Jacobians J_a and J_b which are associated with different treatments, *i.e.,* environmental conditions:

$$dJ_{ij} = \log_2\left(abs\left(\frac{J_{a,ij}}{J_{b,ij}}\right)\right) \tag{6}$$

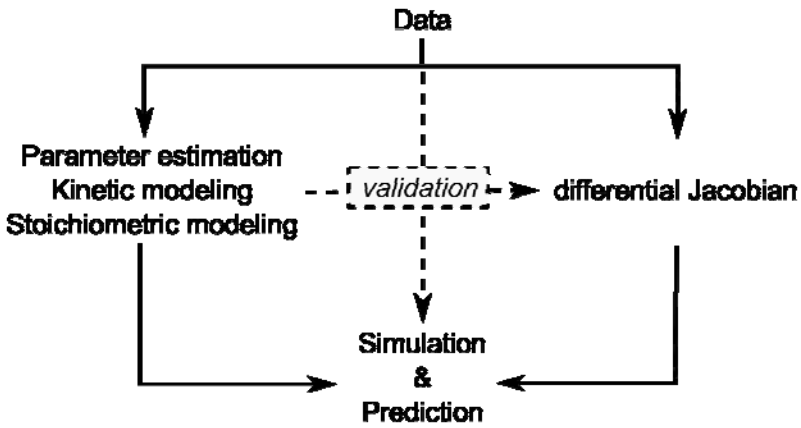
Calculation of the differential Jacobian reveals perturbation sites between two different metabolic states hinting at a significant regulatory event, e.g., the change of enzymatic reaction rates due to environmental perturbations. In principle, using this approach it is possible to conveniently connect a large metabolomics experiment with many samples and thousands of variables directly with the predicted genome-scale metabolic network to calculate biochemical regulation in the investigated biological system (for more detail see [32]). The approach relies on the assumption that regulation of metabolism becomes observable in the significant changes of the local dynamics around a steady state condition, e.g., rates of metabolite synthesis and degradation. Due to the redundancy of pathways and multiple isoforms of numerous enzymes, such calculations and predictions need to be confirmed and

validated by further biochemical experiments. Limitations to this approach are currently the low quality knowledge of N and the low number of detected metabolites in measurements compared to the number of predicted metabolites in a metabolome, necessitating the simplification of N in accordance with the data matrix [32,59].

4. Conclusions

With regard to currently emerging discussions about world population feeding, global climate change and limited energy resources with fossil fuels, plant biology and biotechnology are central topics of life sciences in the coming decades [53]. Latest developments in bioanalytical research aim at the understanding of organisms at a systems level and within their ecosystemic context. Characterized by nonlinearities and multidimensionality, the comprehensive analysis of plant-environment interactions is non-intuitive. Thus, the application of methods, which are capable of coping with this complexity, is necessary. Mathematical modeling and computer-assisted data analysis are powerful and adequate approaches used to exploit entire data sets provided by experimental high-throughput technologies in order to derive a new hypothesis about regulation of biological systems. Although every single mathematical approach is limited by underlying assumptions, the combination of different modeling approaches may yield the ultimate amount of information available from experimental data sets (Figure 1).

Figure 1. Overview of modeling approaches and their interaction by validation. Data represent results of experiments on the metabolome, proteome, enzyme activities or transcriptome.



Kinetic modeling approaches are limited by lack of kinetic information and stoichiometric modeling approaches are limited by their reference to a steady state. Yet, if a stoichiometric modeling approach delivers information about potential perturbation sites in metabolism, this will enable systematic in-depth analysis, for example by kinetic modeling, promoting a comprehensive understanding of how plant metabolism is composed functionally.

Acknowledgments

We would like to thank the members of the Department Molecular Systems Biology for fruitful discussions. We would also like to thank the reviewers of this article for their constructive advice to improve its quality and coverage. TN is funded by a Marie Curie ITN project of the European Union, Grant Agreement number 264474.

Conflict of Interest

The authors declare no conflict of interest.

References

1. Strand, A.; Hurry, V.; Henkes, S.; Huner, N.; Gustafsson, P.; Gardeström, P.; Stitt, M. 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*. **1999**, *119*, 1387–1397.
2. Martindale, W.; Leegood, R.C. Acclimation of photosynthesis to low temperature in *Spinacia oleracea* L. Acclimation of photosynthesis to low temperature in *Spinacia oleracea* L. I. Effects of acclimation on CO_2 assimilation and carbon partitioning. *J. Exp. Bot.* **1997**, *48*, 1865–1872.
3. Zhen, Y.; Ungerer, M.C. Clinal variation in freezing tolerance among natural accessions of *Arabidopsis thaliana*. *New Phytol.* **2008**, *177*, 419–427.
4. Stitt, M.; Hurry, V. A plant for all seasons: alterations in photosynthetic carbon metabolism during cold acclimation in Arabidopsis. *Curr. Opin. Plant. Biol.* **2002**, *5*, 199–206.
5. Kaplan, F.; Kopka, J.; Haskell, D.W.; Zhao, W.; Schiller, K.C.; Gatzke, N.; Sung, D.Y.; Guy, C.L. Exploring the temperature-stress metabolome of Arabidopsis. *Plant Physiol.* **2004**, *136*, 4159–4168.
6. Hannah, M.A.; Wiese, D.; Freund, S.; Fiehn, O.; Heyer, A.G.; Hincha, D. Natural genetic variation of freezing tolerance in Arabidopsis. *Plant Physiol*. **2006**, *142*, 98–112.
7. Maruyama, K.; Takeda, M.; Kidokoro, S.; Yamada, K.; Sakuma, Y.; Urano, K.; Fujita, M.; Yoshiwara, K.; Matsukura, S.; Morishita, Y.*; et al.* Metabolic pathways involved in cold acclimation identified by integrated analysis of metabolites and transcripts regulated by DREB1A and DREB2A. *Plant Physiol.* **2009**, *150*, 1972–1980.
8. Davey, M.P.; Woodward, F.I.; Quick, W.P. Intraspecfic variation in cold-temperature metabolic phenotypes of *Arabidopsis lyrata* ssp. *petraea*. *Metabolomics* **2009**, *5*, 138–149.
9. Nägele, T.; Kandel, B.A.; Frana, S.; Meißner, M.; Heyer, A.G. A systems biology approach for the analysis of carbohydrate dynamics during acclimation to low temperature in *Arabidopsis thaliana*. *FEBS J.* **2011**, *278*, 506–518.
10. Scarth, G.W.; Levitt, J. The frost-hardening mechanism of plant cells. *Plant Physiol.* **1937**, *12*, 51–78.
11. Hincha, D.K.; Sonnewald, U.; Willmitzer, L.; Schmitt, J. The role of sugar accumulation in leaf frost hardiness: investigations with transgenic tobacco expressing a bacterial pyrophosphatase or a yeast invertase gene. *J. Plant Physiol*. **1996**, *147*, 604–610.
12. Huner, N.; Öquist, G.; Sarhan, F. Energy balance and acclimation to light and cold. *Trends Plant Sci.* **1998**, *3*, 224–230.

13. Klotke, J.; Kopka, J.; Gatzke, N.; Heyer, A.G. 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.* **2004**, *27*, 1395–1404.
14. Hoffmann, M.H. Biogeography of *Arabidopsis thaliana* (L.) Heynh. (Brassicaceae). *J. Biogeogr.* **2002**, *29*, 125–134.
15. Laibach, F. *Arabidopsis thaliana* (L.) Heynh. als Objekt für genetische und entwicklungsphysiologische Untersuchungen. *Bot. Archiv.* **1943**, *44*, 439–455.
16. Koornneef, M.; Alonso-Blanco, C.; Vreugdenhil, D. Naturally occurring genetic variation in *Arabidopsis thaliana*. *Annu. Rev. Plant Biol.* **2004**, *55*, 141–172.
17. Giersch, C. Mathematical modelling of metabolism. *Curr. Opin. Plant. Biol.* **2000**, *3*, 249–253.
18. Morgan, J.A.; Rhodes, D. Mathematical modeling of plant metabolic pathways. *Metab. Eng.* **2002**, *4*, 80–89.
19. Poolman, M.G.; Assmus, H.E.; Fell, D.A. Applications of metabolic modelling to plant metabolism. *J. Exp. Bot.* **2004**, *55*, 1177–1186.
20. Rios-Estepa, R.; Lange, B.M. Experimental and mathematical approaches to modeling plant metabolic networks. *Phytochemistry* **2007**, *68*, 2351–2374.
21. Poolman, M.G.; Miguet, L.; Sweetlove, L.J.; Fell, D.A. A genome-scale metabolic model of Arabidopsis and some of its properties. *Plant Physiol.* **2009**, *151*, 1570–1581.
22. Nägele, T.; Stutz, S.; Hörmiller, I.I.; Heyer, A.G. Identification of a metabolic bottleneck for cold acclimation in *Arabidopsis thaliana*. *Plant J.* **2012**, doi: 10.1111/j.1365-313X.2012.05064.x.
23. Steuer, R. Computational approaches to the topology, stability and dynamics of metabolic networks. *Phytochemistry* **2007**, *68*, 2139-2151.
24. Nägele, T.; Heyer, A.G. Acclimation to a changing environment in *Arabidopsis thaliana* - an overview of temperature responses in primary metabolism. In *Advances in Environmental Research*; Justin A. Daniels, Ed.; Nova Science Publishers: New York, USA, 2012; Volume 24.
25. Rohwer, J.M. Kinetic modelling of plant metabolic pathways. *J. Exp. Bot.* **2012**, *63*, 2275-2292.
26. Henkel, S.; Nägele, T.; Hörmiller, I.; Sauter, T.; Sawodny, O.; Ederer, M.; Heyer, A.G. 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.
27. Schallau, K.; Junker, B.H. Simulating plant metabolic pathways with enzyme-kinetic models. *Plant Physiol.* **2010**, *152*, 1763–1771.
28. Farquhar, G.; von Caemmerer, S.; Berry, J. A biochemical model of photosynthetic CO_2 assimilation in leaves of C_3 species. *Planta* **1980**, *149*, 78–90.
29. Zhu, X.G.; de Sturler, E.; Long, S.P. Optimizing the Distribution of Resources between Enzymes of Carbon Metabolism Can Dramatically Increase Photosynthetic Rate: A Numerical Simulation Using an Evolutionary Algorithm. *Plant Physiol.* **2007**, *145*, 513–526.
30. Rohwer, J.M.; Botha, F.C. Analysis of sucrose accumulation in the sugar cane culm on the basis of *in vitro* kinetic data. *Biochem J.* **2001**, *358*, 437–445.
31. Curien, G.; Bastien, O.; Robert-Genthon, M.; Cornish-Bowden, A.; Cárdenas, M.L.; Dumas, R. Understanding the regulation of aspartate metabolism using a model based on measured kinetic parameters. *Mol. Syst. Biol.* **2009**, *5*, 271.

32. Weckwerth, W. Unpredictability of metabolism—The key role of metabolomics science in combination with next-generation genome sequencing. *Anal. Bioanal. Chem.* **2011**, *400*, 1967–1978.
33. May, P.; Wienkoop, S.; Kempa, S.; Usadel, B.; Christian, N.; Rupprecht, J.; Weiss, J.; Recuenco-Munoz, L.; Ebenhöh, O.; Weckwerth, W.*; et al.* Metabolomics- and proteomics-assisted genome annotation and analysis of the draft metabolic network of *Chlamydomonas reinhardtii*. *Genetics* **2008**, *179*, 157–166.
34. Seaver, S.M.D.; Henry, C.S.; Hanson, A.D. Frontiers in metabolic reconstruction and modeling of plant genomes. *J. Exp. Bot.* **2012**, *63*, 2247–2258.
35. Nägele, T.; Henkel, S.; Hörmiller, I.; Sauter, T.; Sawodny, O.; Ederer, M.; Heyer, A.G. Mathematical modelling of the central carbohydrate metabolism in *Arabidopsis thaliana* reveals a substantial regulatory influence of vacuolar invertase on whole plant carbon metabolism. *Plant Physiol.* **2010**, *153*, 260–272.
36. Schaber, J.; Liebermeister, W.; Klipp, E. Nested uncertainties in biochemical models. *IET Syst. Biol.* **2009**, *3*, 1–9.
37. Moles, C.G.; Mendes, P.; Banga, J.R. Parameter estimation in biochemical pathways: a comparison of global optimization methods. *Genome Res.* **2003**, *13*, 2467–2474.
38. Banga, J.R. Optimization in computational systems biology. *BMC Syst. Biol.* **2008**, *2*, 47.
39. Esposito, W.R.; Floudas, C.A. Global optimization for the parameter estimation of differential-algebraic systems. *Ind. Eng. Chem. Res.* **2000**, *39*, 1291–1310.
40. Vaz, A.I.F.; Vicente, L.N. A particle swarm pattern search method for bound constrained global optimization. *J. Global Optim.* **2007**, *39*, 197–219.
41. Van den Bergh, F.; Engelbrecht, A. A study of particle swarm optimization particle trajectories. *Inf. Sci.* **2006**, *176*, 937–971.
42. Kolda, T.; Lewis, R.; Torczon, V. Optimization by direct search: new perspectives on some classical and modern methods. *SIAM Rev.* **2003**, *45*, 385–482.
43. Sutherland, W.J. The best solution. *Nature* **2005**, *435*, 569.
44. Castellana, N.E.; Payne, S.H.; Shen, Z.; Stanke, M.; Bafna, V.; Briggs, S.P. Discovery and revision of Arabidopsis genes by proteogenomics. *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 21034–21038.
45. Wienkoop, S.; Weiss, J.; May, P.; Kempa, S.; Irgang, S.; Recuenco-Munoz, L.; Pietzke, M.; Schwemmer, T.; Rupprecht, J.; Egelhofer, V.*; et al.* Targeted proteomics for *Chlamydomonas reinhardtii* combined with rapid subcellular protein fractionation, metabolomics and metabolic flux analyses. *Mol. Biosyst.* **2010**, *6*, 1018–1031.
46. Thiele, I.; Palsson, B.Ø. A protocol for generating a high-quality genome-scale metabolic reconstruction. *Nat. Protoc.* **2010**, *5*, 93–121.
47. Sweetlove, L.J.; Ratcliffe, R.G. Flux-balance modeling of plant metabolism. *Front Plant Sci.* **2011**, *2,* doi: 10.3389/fpls.2011.00038.
48. de Oliveira Dal'Molin, C.G.; Quek, L.E.; Palfreyman, R.W.; Brumbley, S.M.; Nielsen, L.K. AraGEM, a genome-scale reconstruction of the primary metabolic network in Arabidopsis. *Plant Physiol.* **2010**, *152*, 579–589.
49. Mintz-Oron, S.; Meir, S.; Malitsky, S.; Ruppin, E.; Aharoni, A.; Shlomi, T. Reconstruction of Arabidopsis metabolic network models accounting for subcellular compartmentalization and tissue-specificity. *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 339–344.

50. Goodacre, R.; Vaidyanathan, S.; Dunn, W.B.; Harrigan, G.G.; Kell, D.B. Metabolomics by numbers: acquiring and understanding global metabolite data. *Trends Biotechnol.* **2004**, *22*, 245–252.
51. Kempa, S.; Hummel, J.; Schwemmer, T.; Pietzke, M.; Strehmel, N.; Wienkoop, S.; Kopka, J.; Weckwerth, W. An automated GCxGC-TOF-MS protocol for batch-wise extraction and alignment of mass isotopomer matrixes from differential 13C-labelling experiments: a case study for photoautotrophic-mixotrophic grown *Chlamydomonas reinhardtii* cells. *J. Basic Microbiol.* **2009**, *49*, 82–91.
52. Weckwerth, W. Metabolomics: an integral technique in systems biology. *Bioanalysis* **2010**, *2*, 829–836.
53. Scherling, C.; Roscher, C.; Giavalisco, P.; Schulze, E.D.; Weckwerth, W. Metabolomics unravel contrasting effects of biodiversity on the performance of individual plant species. *PLoS One* **2010**, *5*, e12569.
54. Doerfler, H.; Lyon, D.; Nägele, T.; Sun, X.; Fragner, L.; Weckwerth, W. Granger causality in integrated GC/MS and LC/MS metabolomics data reveals the interface of primary and secondary metabolism. *Metabolomics,* **2012**, *submitted.*
55. Sansone, S.-A.; Fan, T.; Goodacre, R.; Griffin, J.L.; Hardy, N.W.; Kaddurah-Daouk, R.; Kristal, B.S.; Lindon, J.; Mendes, P.; Morrison, N.*; et al.* The metabolomics standards initiative. *Nat. Biotechnol.* **2007**, *25*, 846–848.
56. Weckwerth, W. Integration of metabolomics and proteomics in molecular plant physiology - coping with the complexity by data-dimensionality reduction. *Physiol. Plant* **2008**, *132*, 176–189.
57. Wienkoop, S.; Morgenthal, K.; Wolschin, F.; Scholz, M.; Selbig, J.; Weckwerth, W. Integration of metabolomic and proteomic phenotypes: analysis of data covariance dissects starch and RFO metabolism from low and high temperature compensation response in *Arabidopsis thaliana*. *Mol. Cell Proteomics* **2008**, *7*, 1725–1736.
58. Weckwerth, W. Green systems biology - From single genomes, proteomes and metabolomes to ecosystems research and biotechnology. *J. Proteomics* **2011**, *75*, 284–305.
59. Sun, X.; Weckwerth, W. 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* **2012**, *8*, 81–93.
60. Steuer, R.; Kurths, J.; Fiehn, O.; Weckwerth, W. Observing and interpreting correlations in metabolomic networks. *Bioinformatics* **2003**, *19*, 1019–1026.
61. Heinrich, R.; Schuster, S. *The Regulation of Cellular Systems*; Chapman & Hall: New York, USA 1996.

© 2012 by the authors; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution license (http://creativecommons.org/licenses/by/3.0/).