Available online at www.sciencedirect.com

ScienceDirect

Current Opinion in Biotechnology

Using metabolome data for mathematical modeling of plant metabolic systems

Masami Yokota Hirai[1,2] and Fumihide Shiraishi[3]

Plant metabolism is characterized by a wide diversity of metabolites, with systems far more complicated than those of microorganisms. Mathematical modeling is useful for understanding dynamic behaviors of plant metabolic systems for metabolic engineering. Time-series metabolome data has great potential for estimating kinetic model parameters to construct a genome-wide metabolic network model. However, data obtained by current metabolomics techniques does not meet the requirement for constructing accurate models. In this article, we highlight novel strategies and algorithms to handle the underlying difficulties and construct dynamic *in vivo* models for large-scale plant metabolic systems. The coarse but efficient modeling enables the prediction of unknown mechanisms regulating plant metabolism.

Addresses
[1] RIKEN Center for Sustainable Resource Science, 1-7-22, Suehiro-cho, Tsurumi-ku, Yokohama 230-0045, Japan
[2] Graduate School of Bioagricultural Sciences, Nagoya University, Furo-cho, Chikusa-ku, Nagoya 464-8601, Japan
[3] Section of Bio-Process Design, Department of Bioscience and Biotechnology, Graduate School of Bioresource and Bioenvironmental Sciences, Kyushu University, West #5 Bldg., Moto-oka 744, Nishi-ku, Fukuoka 819-0395, Japan

Corresponding author: Hirai, Masami Yokota (masami.hirai@riken.jp)

Current Opinion in Biotechnology 2019, **54**:138–144

This review comes from a themed issue on **Analytical biotechnology**

Edited by **Hiroshi Shimizu** and **Fumio Matsuda**

https://doi.org/10.1016/j.copbio.2018.08.005

Introduction — metabolomics has accelerated plant functional genomics

In the past two decades, metabolomics, in combination with transcriptomics, has been efficiently utilized for plant functional genomics. An integrated metabolomics strategy has enabled discovery and characterization of new genes involved in the biosynthesis of plant natural products or specialized metabolites (previously called secondary metabolites) [1]. This strategy is based on omics-based 'hypothesis generation,' in other words, 'prediction of gene function.' As large-scale transcriptome and metabolome datasets have revealed that many genes that encode enzymes involved in a pathway are coexpressed with each other and with accumulation of pathway metabolites [2–4], one can predict gene function based only on sequence similarity and coexpression relationships among genes [5]. Metabolic profile of a knockout mutant of a gene is a direct *in vivo* evidence of its predicted function; therefore, metabolomics provides a tool for confirmation of predicted gene function.

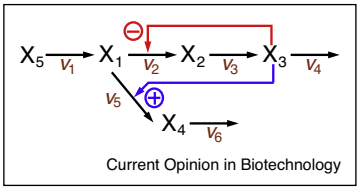
Because metabolomes directly represent *in vivo* metabolic status, metabolomics also helps us reveal relationships among metabolites and metabolic fluxes in a metabolic reaction network or metabolic system (Figure 1). In the following sections, we introduce a novel strategy for metabolomics-based mathematical modeling and discuss its technical difficulty and the solution, aiming at an understanding of complicated plant metabolic reaction networks.

Predictive metabolic engineering is a challenge in plant science

For sustainable production of food, chemical compounds, and biofuels, metabolic engineering of plants is becoming increasingly important. Plant metabolic engineering has two purposes. One is to utilize plant cells as green factories for the production of chemicals or plant natural products, because they have been widely used as medicines, pigments, flavors, etc. since ancient times and often accumulate in very small quantities in plant tissue [6–14,15••]. Because plants are also used as food, feed, and fuels, another purpose of metabolic engineering is to develop a plant with customized characteristics, such as staple crops with improved nutritional value and energy crops with enhanced biomass or oil content [14,16–19]. A number of successful examples can be found in microbial metabolic engineering, which is based on an extensive systems-level understanding of metabolic and regulatory networks [20••]. Many tools and strategies enable microbial metabolic engineering in more sophisticated and diverse ways [21•]. On the other hand, plant metabolic engineering is still challenging; it often yields results different from what was expected, presumably because plant metabolic networks are spread across multiple organelles and are extremely complex as compared to microorganisms [15••,20••].

The biosynthetic pathways of plant natural products seem comparatively simple; the main route of biosynthesis chiefly comprises linear pathways, and the compounds of interest are largely end products of the pathways. The

Figure 1



An example of a metabolic reaction network. X_i ($n = 1, \ldots, 4$) denotes metabolites. Black arrows represent metabolic fluxes denoted as v_i ($n = 1, \ldots, 6$). In this case, X_3 inhibits and activates the enzymes governing v_2 and v_5, respectively.

pathway enzymes are regulated mainly at the transcriptional level via a relatively small number of transcription factors, resulting in a simple relationship between flux size and metabolite pool (metabolite accumulation). In contrast, primary metabolism involving carbon fixation, central carbon metabolism, and energy metabolism is composed of complicated pathways including many branches and loops, and regulated at multiple omics-layer levels, including regulation of enzymatic activities by metabolites. This factor makes it difficult to anticipate the influence of changes in flux size on metabolite accumulation and *vice versa*. As a result, compared to metabolic engineering for chemical production, modification of primary metabolism seems far more difficult.

An example of plant metabolic engineering aiming at value-added staple crops targets the biosynthesis of aspartate (Asp)-family amino acids, because this pathway produces four essential amino acids (threonine, lysine, methionine, and isoleucine) for humans and non-ruminant livestock. Plants tightly regulate this pathway at the enzymatic level, where key enzymes are inhibited or activated by the pathway metabolites [17]. When a crop is used as a sole protein source for humans and livestock, the amounts of these amino acids are imbalanced. Thus, many transgenic studies have been conducted to realize desirable amino acid compositions in crops [16,17]. Mathematical modeling has also been exploited to understand this metabolism as a system and clarify relationships among metabolite contents and metabolic fluxes [22,23].

Construction of mathematical models of metabolic reaction networks

In general, the types of mathematical models of metabolic reaction networks vary from large-scale qualitative models to small-scale quantitative models. The former includes topological and stoichiometric models, while the latter includes kinetic models [24•]. Flux balance analysis (FBA) is one of the stoichiometric and static modeling methods [25]. Since FBA assumes a steady state and thus does not require kinetic information, it is widely used to construct genome-scale models of various plant species [26–32]. On the other hand, kinetic models are constructed to understand dynamic characteristics of a smaller metabolic system, such as Asp-family amino acid biosynthesis [22,23].

To construct kinetic models, flux is often expressed as a function of metabolite concentrations and kinetic properties of enzymes, using kinetic rate laws such as Michaelis–Menten kinetics. In this case, model parameters are Michaelis constants and maximum velocities of enzymes, whose numerical values are determined *in vitro*. As these values may not represent actual *in vivo* performance of enzymes, considerable experimental effort is put to make a constructed model biologically reasonable. For example, a model including 11 metabolites and 18 fluxes was constructed to understand interactions among allosteric regulations in Asp-family amino acid biosynthesis in *Arabidopsis thaliana*, while the kinetic parameters of involving enzymes were measured *in vitro* with taking utmost care to maintain biological relevance [22,23]. To capture flux redistributions caused by genetic manipulations, Colón *et al.* (2009) constructed a benzenoid biosynthesis model in petunia flower, which includes 17 metabolites and 35 fluxes, with the kinetic parameters determined from *in vivo* metabolite pool sizes and labeling patterns obtained via feeding experiments [33].

Alternatively, flux is expressed in power-law expressions in the framework of biochemical systems theory (BST) [34–37] to construct kinetic models. In BST, metabolic systems are mainly represented in the following two types of power-law formalisms: the saturable and synergistic (S)-system and generalized mass action (GMA)-system.

S-system:

$$\frac{dX_i}{dt} = \alpha_i \sum_{j=1}^{n+m} X_j^{g_{ij}} - \beta_i \prod_{j=1}^{n+m} X_j^{h_{ij}} = V_i - V_{-i} \qquad (1)$$

where X_i is the metabolite concentration, α_i and β_i are the rate constants for net influx V_i and efflux V_{-i}, respectively, g_{ij} and h_{ij} are the kinetic orders, n and m are the numbers of dependent and independent variables, respectively, and t is the time. It should be noted that influxes and effluxes in Equation 1 are individually grouped into one power-law form.

GMA-system:

$$\frac{dX_i}{dt} = \sum_{k=1}^{p} A_{ik} \prod_{j=1}^{n+m} X_j^{G_{ijk}} - \sum_{k=1}^{q} B_{ik} \prod_{j=1}^{n+m} X_j^{H_{ijk}} = \sum_{k=1}^{p} v_{ik} - \sum_{k=1}^{q} v_{-ik} \qquad (2)$$