

RESEARCH

Open Access

# A U-system approach for predicting metabolic behaviors and responses based on an alleged metabolic reaction network

Kansuporn Sriyudthsak<sup>1,2</sup>, Yuji Sawada<sup>1</sup>, Yukako Chiba<sup>3,4</sup>, Yui Yamashita<sup>3</sup>, Shigehiko Kanaya<sup>5</sup>, Hitoshi Onouchi<sup>2,6</sup>, Toru Fujiwara<sup>7</sup>, Satoshi Naito<sup>3,6</sup>, Ebernard O Voit<sup>8</sup>, Fumihide Shiraishi<sup>9\*</sup>, Masami Yokota Hirai<sup>1,2\*</sup>

From The 25th International Conference on Genome Informatics (GIW/ISCB-Asia) Tokyo, Japan. 15-17 December 2014

## Abstract

**Background:** Progress in systems biology offers sophisticated approaches toward a comprehensive understanding of biological systems. Yet, computational analyses are held back due to difficulties in determining suitable model parameter values from experimental data which naturally are subject to biological fluctuations. The data may also be corrupted by experimental uncertainties and sometimes do not contain all information regarding variables that cannot be measured for technical reasons.

**Results:** We show here a streamlined approach for the construction of a coarse model that allows us to set up dynamic models with minimal input information. The approach uses a hybrid between a pure mass action system and a generalized mass action (GMA) system in the framework of biochemical systems theory (BST) with rate constants of 1, normal kinetic orders of 1, and -0.5 and 0.5 for inhibitory and activating effects, named Unity (U)-system. The U-system model does not necessarily fit all data well but is often sufficient for predicting metabolic behavior of metabolites which cannot be simultaneously measured, identifying inconsistencies between experimental data and the assumed underlying pathway structure, as well as predicting system responses to a modification of gene or enzyme. The U-system approach was validated with small, generic systems and implemented to model a large-scale metabolic reaction network of a higher plant, *Arabidopsis*. The dynamic behaviors obtained by predictive simulations agreed with actually available metabolomic time-series data, identified probable errors in the experimental datasets, and estimated probable behavior of unmeasurable metabolites in a qualitative manner. The model could also predict metabolic responses of *Arabidopsis* with altered network structures due to genetic modification.

**Conclusions:** The U-system approach can effectively predict metabolic behaviors and responses based on structures of an alleged metabolic reaction network. Thus, it can be a useful first-line tool of data analysis, model diagnostics and aid the design of next-step experiments.

## Background

Systems biology have advanced far enough that it is becoming possible to gain a comprehensive understanding of a metabolic system, and to use this knowledge for

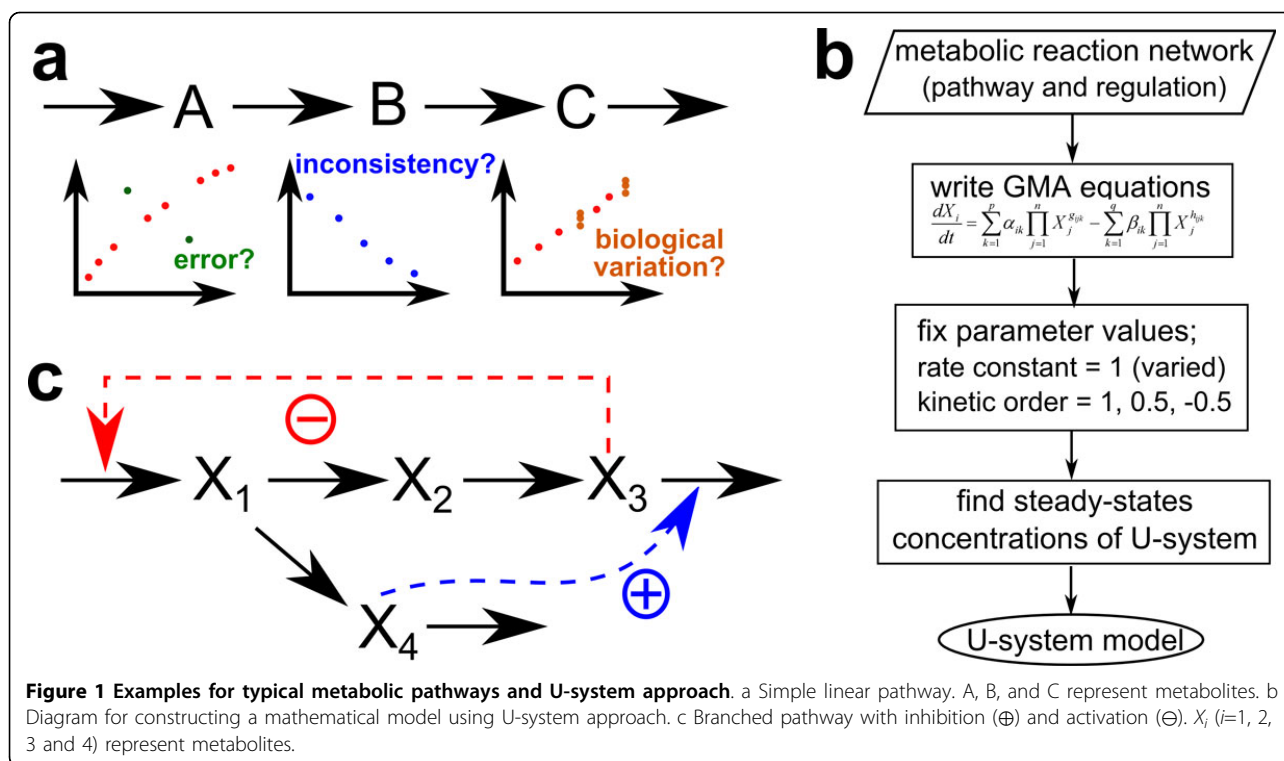
developing rational treatment options for diseases or devising strategies for increasing the productivity of foods and chemicals. While the main hallmark of systems biology is often portrayed as a means for predicting systems responses to external or internal stimuli, a similarly important aspect is its abilities to envisage all relevant information and integrate experimental data into logical and computable structures. As an example, suppose that three metabolites, A, B, and C, form a linear pathway (Figure 1a).

\* Correspondence: fumishira@brs.kyushu-u.ac.jp; masami.hirai@riken.jp

<sup>1</sup>RIKEN Center for Sustainable Resource Science, Yokohama 230-0045, Japan

<sup>9</sup>Graduate School of Bioresource and Bioenvironmental Sciences, Kyushu University, Fukuoka 812-8581, Japan

Full list of author information is available at the end of the article



Supposing further that only experimental data for A and C are available and the data show A and C increasing after stimulating the reaction producing A. On the basis of the network structure together with the available data, even though B may not be ably measured, it is logically predictable that B is probably increased as well. On the other hand, supposing that the data for all three metabolites, A, B, and C, are available and the data show that A and C increasing whereas B decreasing. Without any information on the structure of pathway, one can only take the data of B at face value and deem them correct. However, if there is an indication showing that A, B, and C, form a linear pathway like the previous case, it becomes immediately clear that there is some inconsistencies. According to the logic of a conceptual or mathematical model of the pathway, either some branches or regulatory signals are missing, or the data are unreliable. To be more precise, if the data behave in a reasonable trend with some biological fluctuations like those in orange dots (Figure 1a), one may assume that these data just contain intrinsic biological variations and some missing pathways or regulations make the data inconsistent. In contrast, if some data split out of the trend in an unusual way like those in green dots (Figure 1a), one may judge that these data are prone to be errors and probably unreliable. This implies that the logical concepts allow us to assess associated data and possibly identify inconsistencies.

The exploitation of logical concepts for the prediction of probable behaviors of unmeasurable data and identification

of uncertainties in datasets becomes much more complicated and important if the pathway systems are large. Large-scale datasets, and in particular time series data characterizing metabolic systems, are also becoming more and more commonplace to assist the investigation. To acquire such datasets with reasonable efforts, the high-throughput analytical instruments have been developed in recent times [1-3]. Given the nature of experiments, however, it is not even always certain that all information for interested variables can be gained from experiments and not all metabolites in the large-scale pathway can be measured. It also cannot be guarantee that the replicated samples in terms of biological, technical and analytical aspects reproduce the same results, even in a qualitative sense. Two major questions thus arise. The first question is whether computational methods of systems biology which are cheap and straightforward may be able to predict or analyze such unmeasurable data with respect to system identification based on our understanding of the system. The second question is how reliable or accurate biological data are, especially if they were obtained from different cellular compartments or from different developmental stages in plants and animals.

It is therefore useful to explore to what degree it might be possible to employ the logic of the pathway to address these issues. Such an analysis seems to be feasible in principle, because the biologist executing the experiments usually has a relatively well-supported concept of the topology of the pathway system. The challenge is that

even small pathways quickly become too complicated for intuitive assessments, and larger and more complicated pathway systems simply prevent us from testing assumptions or hypotheses without computational support. Such computational support of course requires mathematical models, which immediately leads to the challenge of setting up detailed, quantitatively appropriate models with a minimum of input requirements. The goal of such models is not necessarily to provide a quantitatively perfect picture of a pathway or to fit all data with great accuracy. Instead, these models should be easy to construct, robust, and merely able to identify how data within a large dataset could probably be.

Power-law models within Biochemical Systems Theory (BST; [4]) possess two great advantages that can be leveraged toward predicting the unmeasurable data as well as testing the consistency between the measured data and an alleged pathway structure and its regulation. First, symbolic BST models can be designed immediately for a pathway system of any complexity. These symbolic models do not include parameter values, but they do define a model's potential repertoire of responses, rather than addressing specific data fits. Second, the estimation of optimal parameter values for large systems is generally fraught with technical difficulties [5]. This issue is greatly ameliorated for BST models, because even relatively coarse numerical settings of their parameters are often sufficient to capture the behavior of a metabolic pathway system in a semi-quantitative fashion.

Here we capitalize on these two features of BST and propose a coarse test for the prospect of metabolic time-series data and consistency between the data and an alleged pathway diagram. The purpose is not to establish optimally fitting models but to identify how the unmeasurable data of metabolites located in a focused pathway probably behave, which time-series data within the dataset may be inconsistent with the understanding of pathway, as well as which data within a large dataset are probably reliable and which, if any, are most likely imprecise. We use a hybrid between a pure mass action system and a generalized mass action (GMA) system in the framework of BST with kinetic orders of 1, which allows for inhibitory and activating effects that are modeled with kinetic orders of -0.5 and 0.5. The rate constants are at first set to 1, but may be later subjected to order-of-magnitude adjustments, which obviously improve the model representation of the data. Once the parameter values are fixed, their effects are to some degree compensated by adjustments in metabolite concentrations according to the network structure so that we named this simplified method unity (U)-system. The diagram for U-system approach is also exhibited in Figure 1b. We begin the demonstration of

the U-system approach with artificial "data" from fully known, representative models and show that coarse values for kinetic orders and rate constants retain much of the qualitative behavior of true model responses. These U-system solutions are rather coarse, but can be improved with order-of-magnitude adjustments of flux split ratios at branched points. We then demonstrate the feasibility of the method with actual data describing amino acid synthesis in *Arabidopsis*. This pathway system contains 351 metabolites and 441 fluxes, which are subject to various regulatory mechanisms. Apart from these, we have measurements from two different analytical methods, one comprising 268 metabolites, including 21 amino acids, and the other one accounting for 16 amino acids. As expected, although a large number of data can be measured, it is clear that not all data can be obtained. Also, none of the obtained data contain exactly the same values even in the replicated samples, mainly due to high biological variation among samples. In most cases, the measured time series are qualitatively the same, but in some cases they differ. We show that the U-system analysis estimates the probable behaviors of unmeasurable data and identifies at least some of the unreliable data. Besides, the U-system model is also applicable to predict metabolic responses when a metabolic reaction is altered.

## Results and discussion

### Testing of feasibility of assumptions in U-system

Obviously, a system with arbitrary parameter values of 1 is quite different from a system with diverse parameter values. The performances and characteristics of U-system approach were simply approved using a simple model of a linear pathway by comparing with the typical model constructed in Michaelis-Menten format where the enzymatic reaction velocity is usually determined by an *in vitro* experiment (Supplementary Information - additional file 1). The result showed that U-system model can produce qualitatively similar results to the original. Thus, a simple generic, yet representative model was used to test to what degree the U-system model can produce qualitatively comparable results.

Many studies in recent years have used a simple, representative model of a branched pathway with inhibition and activation (Figure 1c), for instance, to test new parameter estimation methods [6-8]. We used this same system but slightly modified it to allow a wider range of responses. Equations (1)-(4) show the differential equations together with actual parameters for the rate constants and kinetic orders of this system in GMA-system format.

$$\frac{dX_1}{dt} = \alpha_1 X_3^{g13} - \beta_{11} X_1^{h111} - \beta_{12} X_1^{h112} = 12X_3^{-0.8} - 8X_1^{0.5} - 2X_1^{0.5} \quad (1)$$

$$\frac{dX_2}{dt} = \alpha_2 X_1^{g_{21}} - \beta_2 X_2^{h_{22}} = 8X_1^{0.5} - 3X_2^{0.75} \quad (2)$$

$$\frac{dX_3}{dt} = \alpha_3 X_2^{g_{32}} - \beta_3 X_3^{h_{33}} X_4^{h_{34}} = 3X_2^{0.75} - 5X_3^{0.5} X_4^{0.2} \quad (3)$$

$$\frac{dX_4}{dt} = \alpha_4 X_1^{g_{41}} - \beta_4 X_4^{h_{44}} = 2X_1^{0.5} - 6X_4^{0.8} \quad (4)$$

#### Effect of rate constants on metabolic behaviors

Figure 2a and 2b show patterns of normalized concentrations for each metabolite over time resulting from changes in rate constants  $\alpha_1$  and  $\beta_3$ , respectively. It is clear that when  $\alpha_1$  is increased, the time-course curves of metabolite concentrations tend to stretch along the x-axis, whereas the extension along the y-axis is only slight. Although the results for  $\beta_3$  are clearly different from those for  $\alpha_1$ , the dynamic responses for each  $X_i$  are qualitatively similar. Overall, changes in the rate constants tend to affect the dynamics in the direction of the x-axis. The individual degree of elongation depends on the specifics of the system. This result is not surprising, because multiplying the entire system with a positive constant corresponds to changing the unit of time and thereby the scale of the x-axis (see, e.g., [9]). Additional results are presented in Supplementary Figures S3 and S4 (additional file 1).

#### Effect of kinetic orders on metabolic behaviors

Figure 2c and 2d show patterns of normalized concentrations of each metabolite over time resulting from changes in kinetic orders  $g_{13}$  and  $h_{34}$ , respectively. When  $g_{13}$  is changed, only the magnitudes of normalized metabolite concentrations are changed but almost no difference is observed in temporal changes in time-course patterns of metabolite concentrations. Similar results are also observed for  $g_{21}$ ,  $g_{32}$ ,  $g_{41}$ ,  $h_{111}$ ,  $h_{112}$ ,  $h_{22}$ ,  $h_{33}$  and  $h_{44}$  (Supplementary Figures S5 and S6 - additional file 1). This consistency may be due to the fact that the ranges of kinetic orders are typically between -1 and 1 [9], so that kinetic orders do not play an overly strong role in the variation of metabolite concentrations, and the corresponding fluxes do not change much. The results for  $h_{34}$  (Figure 2d) and  $h_{44}$  (Supplementary Figure S6 - additional file 1) are different, as the trajectories noticeably elongate in the direction of the x-axis with changes in kinetic orders. According to metabolic reaction network diagram in Figure 1c, one sees that these kinetic orders are involved in the degradation of  $X_3$ , so that an increase in kinetic order strongly increases the degradation flux of  $X_3$ . However, this situation is somewhat unusual because of the high activation of the efflux. In this case,  $X_4$  is produced from  $X_1$  which is inhibited by  $X_3$  so that a strong activation of the efflux of  $X_3$  by  $X_4$  quite effectively slows down changes in  $X_3$ . Also,  $X_3$  has a low initial value and

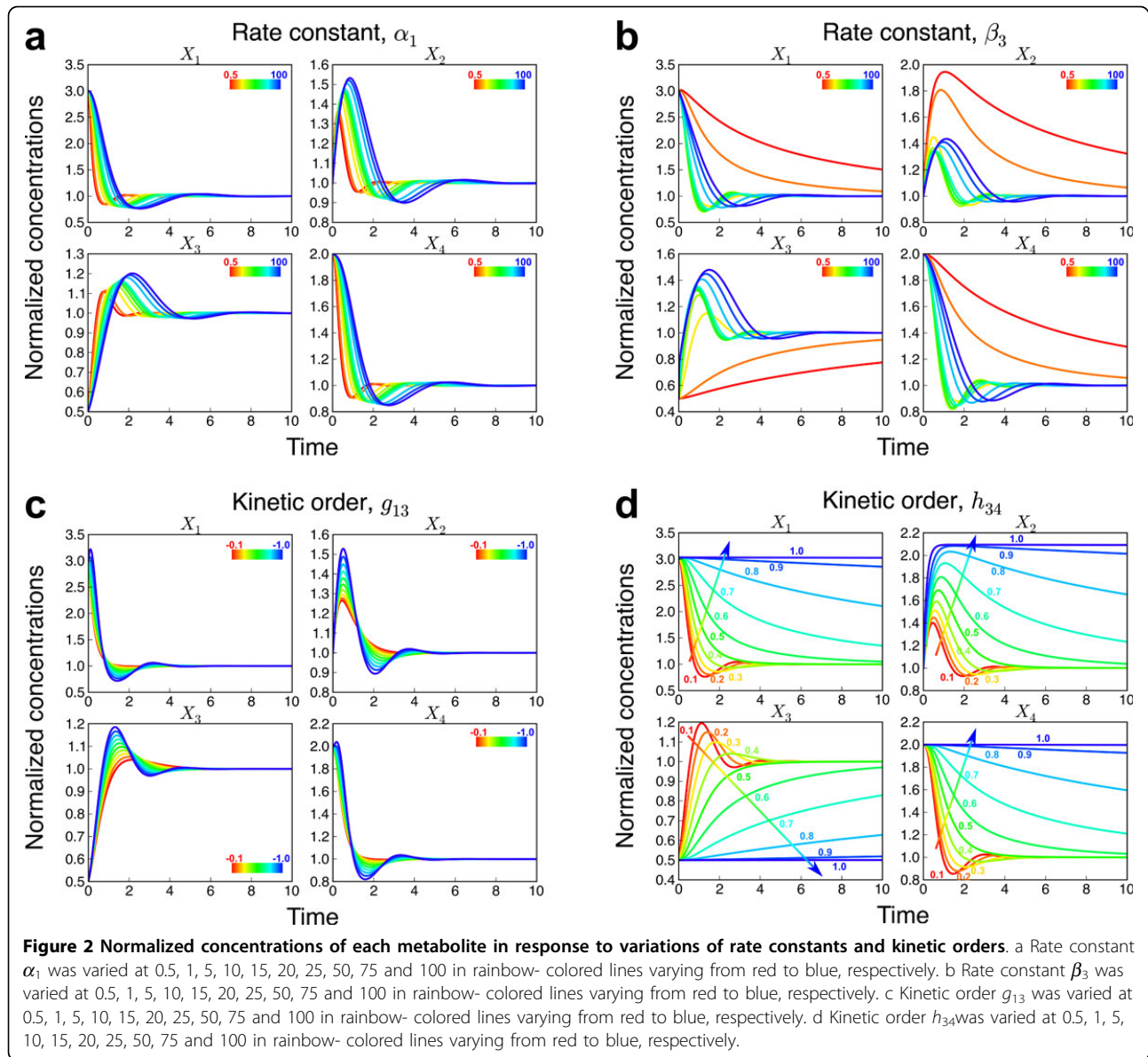
kinetic order compared to other metabolites. Thus, the dynamics of  $X_3$  is driven by other metabolites and  $X_3$  concentration does not change much unless the other parameters vary. Furthermore, the input  $X_1$  is strongly affected by  $X_3$  which does not change much. Accordingly, other metabolites hardly change throughout the period of observed time. In general, the kinetic parameters for activation are seldom higher than kinetic parameters for the metabolite itself so that this particular scenario is rare. As a consequence, it is still able to deduce that changes of kinetic orders tend to affect the trajectories in the direction of the y-axis rather than the x-axis.

#### Effect of random parameters on metabolic behaviors using Monte-Carlo simulation

Figure 3 shows a selection of Monte-Carlo simulation results associated with changes in the rate constants between 0.2 and 20, and kinetic orders between 0.2 (-0.2) and 0.8 (-0.8) for the branched pathway model with inhibition and activation as well as the simulations for U-system model (blue lines) and original GMA-system model (green lines). Parameter combinations not leading to a stable system were removed. For the U-system model, all rate constants and kinetic orders of substrates and enzymes were set to 1, while the kinetic orders of parameters for inhibitions and activations were set to be -0.5 and 0.5, respectively. As a result, the observed metabolite concentrations are no longer the real concentrations, but simply indications of the shapes of their trajectories. The present study calls these concentrations as U-system concentrations (Figure 1b).

The initial values of metabolite concentrations of  $X_1$ ,  $X_2$ ,  $X_3$ , and  $X_4$  were set to 3-, 1-, 0.5- and 2-fold their steady-state values. The simulated metabolite concentrations for all cases were normalized by their steady-state values for the observation and comparison. To facilitate inspection of the calculated behaviors of the metabolite concentrations, the scales of both axes were adjusted by the maximum and minimum values of each simulation. The result shows that most Monte-Carlo simulations exhibit qualitatively comparable dynamic responses regardless of the parameter values chosen. Each metabolite concentration changes throughout the period of time and returns to its steady-state value in a similar manner. This observation implies that both models provide qualitatively comparable dynamic patterns of changes in metabolite concentrations. Furthermore, the GMA and U-system responses are surrounded by the Monte-Carlo simulation results for 1000 parameter sets.

In summary of this analysis, the U-system approach indicates that metabolic behaviors mainly depend on the network structure of a metabolic system. Unlike most of modeling methods, it does not need the process of parameter estimation in model construction, although it cannot offer actual quantitative concentrations of

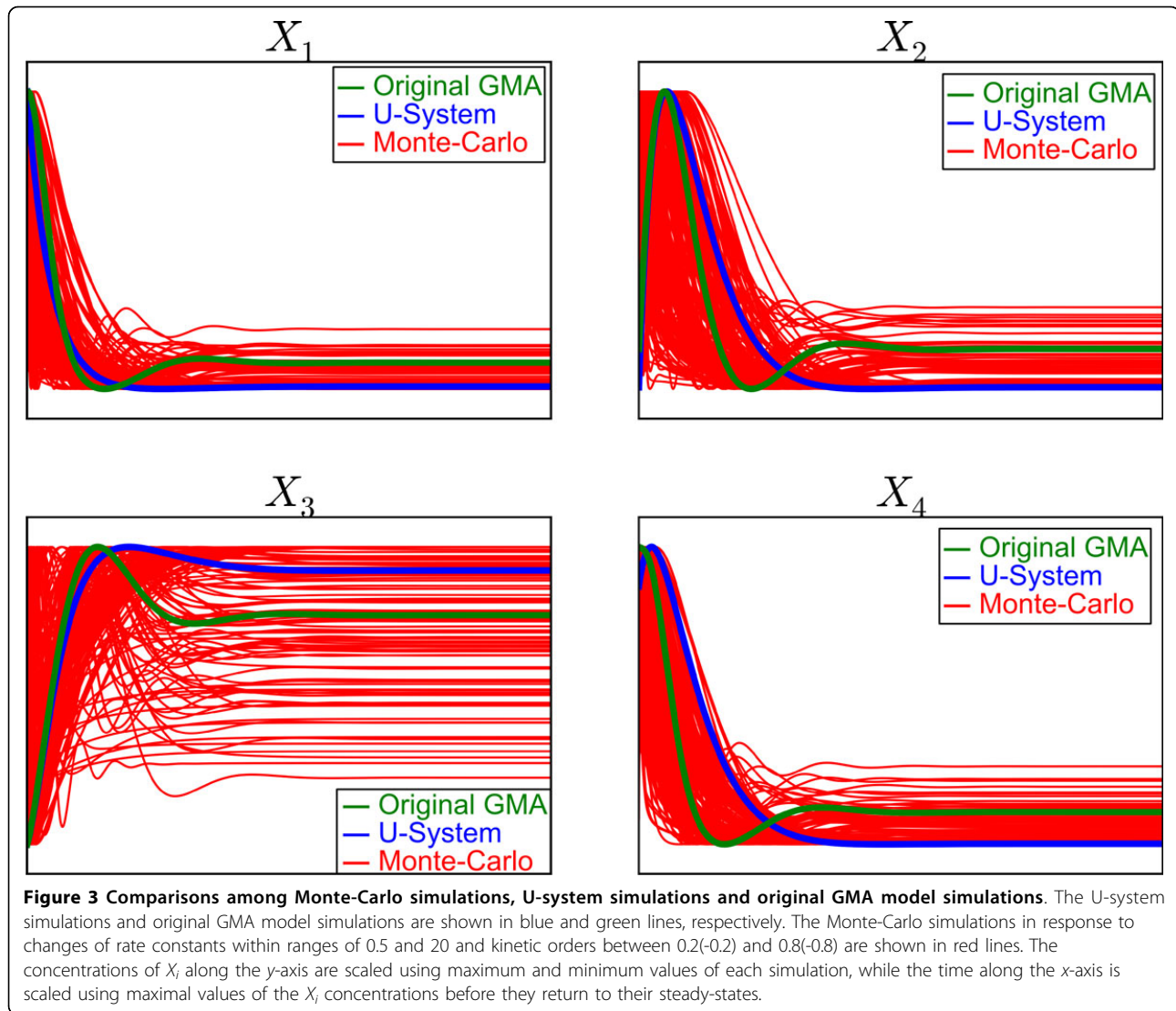


metabolic behaviors. This insight leads to several consequences. First, the U-system approach provides a coarse time-transient behavior of the system regardless of the requirements or dependencies of experimental data. Accordingly, it will not confront difficulties in parameter fitting, especially in a large-scale system. Second, the U-system approach is simple. Plausible parameters for inhibition and activation can be easily set into the model to grasp their possible effects, which makes it possible to design additional experiments. Third, the U-system approach is suitable for practical applications because of its simplicity and flexibility. One can simply construct a mathematical model based on only experimental facts regarding the network structure and regulation. If the U-system simulations are inconsistent with

experimental data, other possible systems may be simply constructed to find out unknown candidates. As a result, an appropriate experiment may be designed to validate or refute the predictions from the simulation and to characterize biological details. Lastly, the U-system approach allows us to acquire tendencies of metabolic behaviors to assess experimental data within the context of their actual network structure.

**Application to metabolic reaction network of *Arabidopsis***  
 Methionine, lysine and threonine are essential amino acids for non-ruminant animals including humans. The methionine derivative, *S*-adenosyl methionine (AdoMet), is a major primary methyl-group donor and also a precursor for a plant hormone ethylene and polyamines. To increase

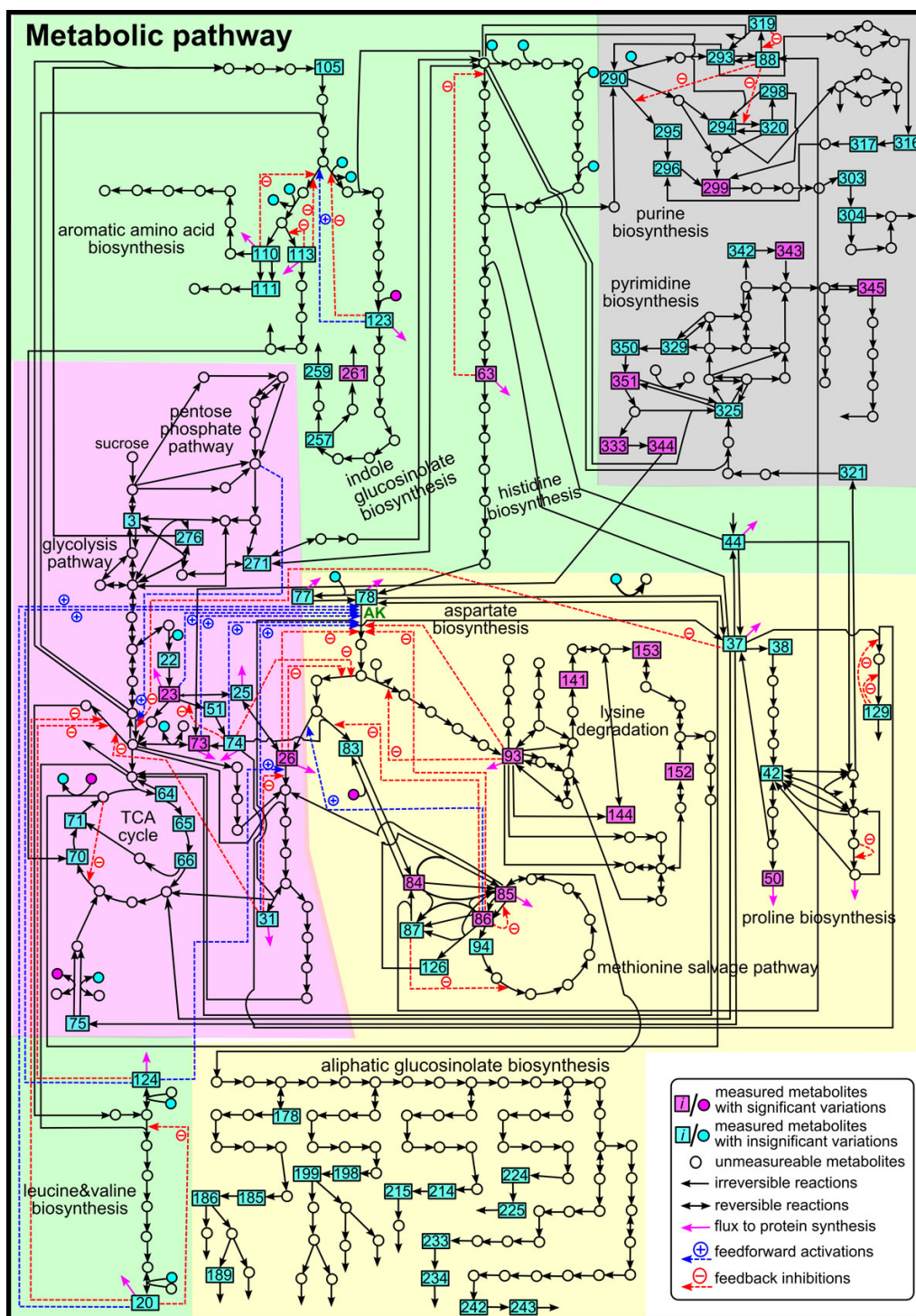




**Figure 3 Comparisons among Monte-Carlo simulations, U-system simulations and original GMA model simulations.** The U-system simulations and original GMA model simulations are shown in blue and green lines, respectively. The Monte-Carlo simulations in response to changes of rate constants within ranges of 0.5 and 20 and kinetic orders between 0.2(-0.2) and 0.8(-0.8) are shown in red lines. The concentrations of  $X_i$  along the y-axis are scaled using maximum and minimum values of each simulation, while the time along the x-axis is scaled using maximal values of the  $X_i$  concentrations before they return to their steady-states.

the sulfur-containing amino acids methionine and cysteine in crops, numerous researches in the field of genetics and metabolic engineering have been carried out for more than two decades, and the objectives have been partially realized. However, the results are not quite satisfactory yet, because the amino acids and related metabolites in the obtained transgenic crops are not well balanced. One of the major causes for not achieving a good balance is an incomplete understanding of the whole regulatory mechanisms of amino acid accumulation, as well as the regulatory feedback structure of the system, although it is well known that the accumulation of methionine, a member of the aspartate-family amino acids together with lysine and threonine, is tightly controlled by negative feedback by lysine and threonine [10]. Another cause is an ambiguity in the characteristics of the metabolic flow

pattern, which affects productivity and metabolic balance [11]. To elucidate their correlations among metabolites, there are various researches focusing on the integrations of *in vitro* data into kinetic models for specific pathways [12,13] and the reconstructions of metabolic fluxes at the genome scale [14,15]. However, no large-scale kinetic models including regulatory mechanisms are known for the *Arabidopsis* metabolic system. In this study, therefore, we exploited the U-system approach to construct a large-scale mathematical model based on available metabolic reaction networks (Figure 4) for describing important characteristics of its metabolic system and predicting metabolic behaviors in whole. A perturbation experiment was also conducted with *Arabidopsis* callus by exogenous application of lysine and threonine to obtain metabolic time-series data for validating the U-system model.



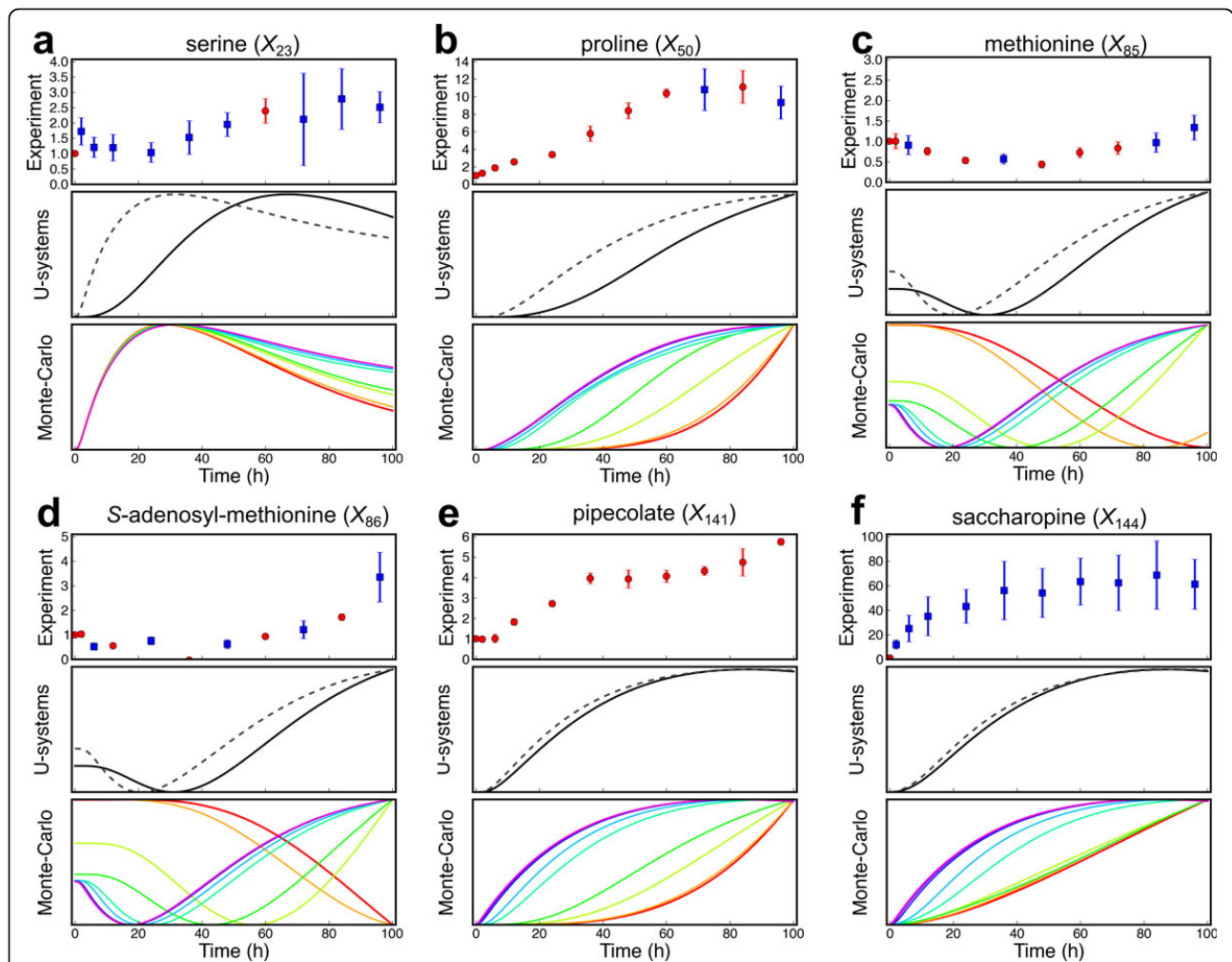
**Figure 4 Metabolic reaction network of *Arabidopsis* focused in this study.** Green zone, aromatic amino acid biosynthesis, indole GSL biosynthesis, histidine biosynthesis and leucine and valine biosynthesis; gray zone, purine and pyrimidine biosynthesis; pink zone, glycolysis pathway, TCA cycle and pentose phosphate pathway; yellow zone, aspartate-family amino acid biosynthesis, proline biosynthesis, lysine degradation, methionine salvage pathway, aliphatic GSL biosynthesis. The number *i* in square boxes indicate the metabolite numbers arbitrarily given for simulations (Supplementary Model - additional file 2).

**Modified U-system approach with rough adjustments of rate constants**

The mathematical model including nonlinear ordinary differential equations (ODEs) for 351 metabolites was constructed by U-system approach on the basis of only structures of metabolic reaction networks available in databases (Supplementary Model - additional file 2). Fick's laws of diffusion [16] was used for representing the uptake rate of lysine plus threonine from the callus culture medium to the inside of the cells from  $t=0$ , and the parameters for uptake rates were estimated using experimental data. For simulations, all ODEs representing changes of metabolite concentrations inside the cells were calculated

together with equations for the supplementation throughout a period of 100 h.

Figure 5 shows the time courses of metabolite concentrations from the U-system approach (gray dotted lines) compared with those from metabolome analysis covering 268 metabolites measured by liquid chromatography-mass spectrometry (LC-MS) (red dots) and those with a coefficient of variations more than 20% (blue dots). It is clear that the actual data contained some biological fluctuations and/or analytical errors. The results indicated that simulated metabolic behaviors changed in similar ways to actual metabolic behaviors but in different time-scales. This is because all rate constants were fixed to 1 so that some rate



**Figure 5 Comparisons of metabolome data with U-system and Monte-Carlo simulations.** Experimental data from metabolome analysis with coefficient of variations less than and more than 20% are represented in red and blue dots, respectively (upper boxes). The U-system and modified U-system simulations for metabolic system of *Arabidopsis* are represented in gray dotted and black lines, respectively (middle boxes), whereas Monte-Carlo simulations with changes of rate constants between branch points around proline biosynthesis, lysine degradation, methionine salvage pathway, aliphatic GSL biosynthesis (yellow zone in Figure 5), in which parameters are 0.005, 0.01, 0.05, 0.1, 0.5, 1, 5, 10, and 50, are represented in rainbow-colored lines varying from red (0.005) to blue (50), respectively (lower boxes). a, serine ( $X_{23}$ ); b, proline ( $X_{50}$ ); c, methionine ( $X_{85}$ ); d, S-adenosyl-methionine ( $X_{86}$ ); e, pipecolate, ( $X_{141}$ ); f, saccharopine ( $X_{144}$ ). The results for other metabolites are shown in Supplementary Figures (additional file 3).



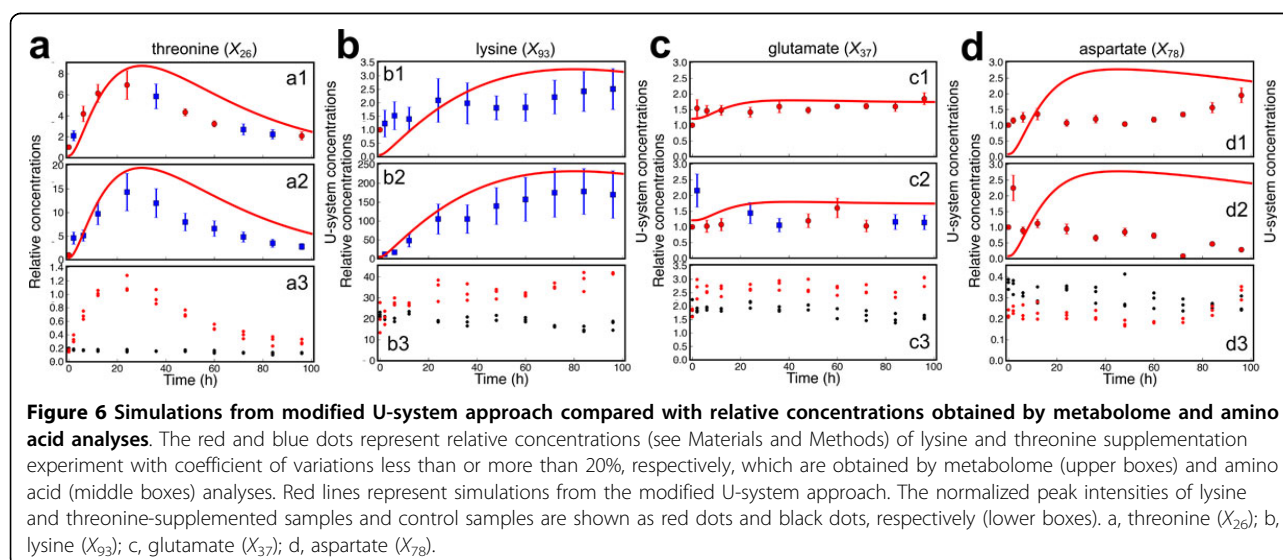
constants may be too large whereas others may be too small compared with the real ones. To prove the hypothesis that it is possible to roughly adjust rate constants, the flux-split ratios at the branched point associated with methionine salvage pathway, aliphatic glucosinolate (GSL) biosynthesis, lysine degradation and proline biosynthesis (yellow zone in Figure 4) were varied. Specifically, the rate constants from these branched points to other branched points were varied at 0.005, 0.01, 0.05, 0.1, 0.5, 1, 5, 10, and 50, respectively, on the basis of Monte-Carlo simulations. Some calculation results satisfactorily agreed with the measured data although there were some differences. The calculation results indicated that simulated concentrations of serine ( $X_{23}$ ) (Figure 5a) does not vary whereas those of proline ( $X_{50}$ ) (Figure 5b) clearly alter. This is because the rate constants of fluxes nearby serine ( $X_{23}$ ) (pink zone in Figure 4) were kept constant whereas those nearby proline ( $X_{50}$ ) (yellow zone in Figure 4) were varied. Similarly, with high rate constants (between 0.5 and 50), the simulated concentrations of methionine ( $X_{85}$ ) (Figure 5c) and Ado-Met ( $X_{86}$ ) (Figure 5d) temporarily decreased and increased back, whereas those of pipercolate ( $X_{141}$ ) (Figure 5e) and saccharopine ( $X_{144}$ ) (Figure 5f) increased and remained constant during the observation period. At the same time, their metabolic responses seemed to slow down with lower values of rate constants, although they would behave in the same manner if longer period of time were observed. These findings implied that it is possible to adjust the time scale of calculated metabolite concentrations in the U-system model by adjusting the rate constants of fluxes at branched points. Thus, rate constants between branched points were roughly adjusted using some of the time course of relative metabolite concentrations for the coarse prediction of the metabolites concentration and this system was named modified U-system. The

results for modified U-system (black lines in Figure 5) showed reasonable metabolic behaviors useful for further analysis.

#### Evaluation of the modified U-system approach for metabolic responses in Arabidopsis

The modified U-system includes various approximations so that it is necessary to validate whether it is practicability and conformable to experimental data. Figure 6 (upper boxes) shows the time courses after lysine and threonine application of the relative concentrations of four amino acids measured in metabolome analysis. The threonine concentration ( $X_{26}$ ) exhibited small biological fluctuations among sample replicates, increasing from its steady-state level to a maximum at around 24 h, from where it decreased back to its steady-state level. The lysine concentrations ( $X_{93}$ ) contained more biological fluctuations among the sample replications, but it is still possible to detect a clear pattern over time. The glutamate concentration ( $X_{37}$ ) increased within 2 h and then seemed to be constant. The aspartate concentration ( $X_{78}$ ) increased slightly, then decreased, and finally tended to increase considerably.

The red lines in Figure 6 (a1 and b1) indicate that the simulated concentrations of threonine ( $X_{26}$ ) and lysine ( $X_{93}$ ) agreed with the metabolic behaviors of experimental data obtained by metabolome analysis with significant correlations ( $p$ -values < 0.005; detailed in Supplementary Model - additional file 2). However, the simulated concentration of aspartate ( $X_{78}$ ) was not consistent with metabolome data (Figure 6d1). It could be that, although the metabolic reaction network used for model construction is correct, the U-system approach might fail to predict the appropriate metabolic behavior. To assess this possibility, a specific measurement only for amino acids using another analytical instrument gas chromatography-mass



spectrometry (GC-MS) was performed using the identical samples (Figure 6, middle boxes). The results from amino acid analysis seemed to provide better metabolite trajectories with lower fluctuations. With respect to the concentrations of threonine ( $X_{26}$ ), lysine ( $X_{93}$ ) and glutamate ( $X_{37}$ ), both methods yielded similar results in terms of the metabolic trajectories, although the simulations were in better agreement with the time-courses of metabolite concentrations measured by amino acid analysis. For example, lysine concentration ( $X_{93}$ ) measured by metabolome analysis (Figure 6b1) increased and maintained its level from  $t=24$  to  $t=60$  h before increasing again. The U-system simulation pinpointed that four data at  $t=24, 36, 48$  and  $60$  h obtained by metabolome analysis may contain errors. Actually, the data at  $t=24$  and  $36$  h (shown as blue dots) exhibited high standard deviations. This assumption may be indirectly supported by the results of correlation analysis. For example, the correlation between the U-system simulation and the time-series data obtained by the amino acid analysis (Figure 6b2) is 0.991, which is higher than the correlation of the value, 0.881, between the U-system simulation and the data obtained by metabolome analysis. The result indicates that if a relatively reliable metabolic reaction network diagram is available, the prediction of time-dependent changes in metabolite concentrations using the U-system is beneficial for judging the likely correctness of data measured by metabolome analysis (see, e.g., green dots in Figure 1a), which is designed to provide massive datasets of metabolite concentrations sometimes at the expense of absolute concentration accuracies for all metabolites. By contrast, if the predicted behaviors are significantly different in some of the metabolites (see, e.g., blue dots in Figure 1a), there may be an unknown metabolic pathway associated with these metabolites or an unidentified regulatory signal. Thus, one may attempt to test a possible system by modifying a pathway or regulation, which can provide a consistent result to predict an unknown candidate before performing a next experiment.

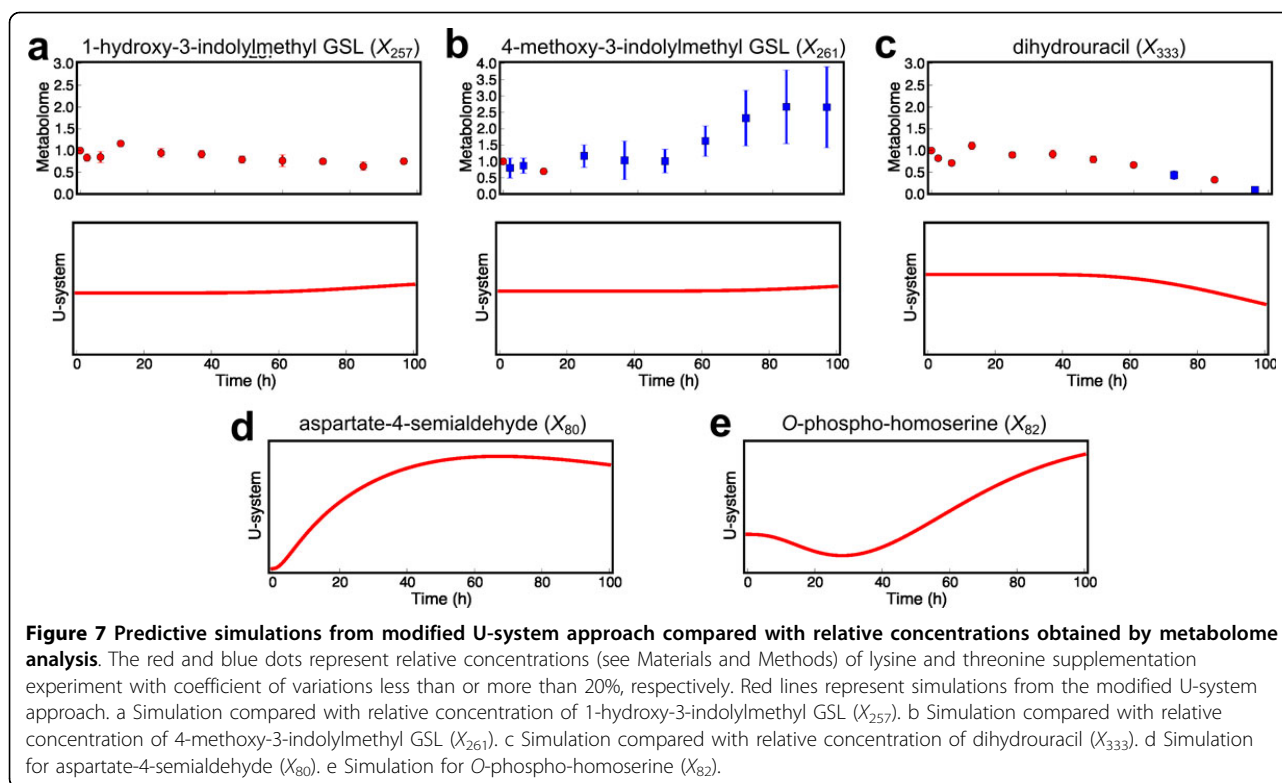
As mentioned above, the relative concentrations observed for several amino acids such as aspartate and glutamate were scattered mainly because of biological variability and possibly due to analytical limitations, and it was difficult to precisely capture the time courses of their concentrations (Figure 6c and 6d). In fact, significant fluctuations in actual metabolite concentrations were observed even in the control samples without lysine and threonine supplementation (Figure 6c3 and 6d3, black dots). However, the computational results suggested that when lysine and threonine were applied as supplements, the aspartate concentration ( $X_{78}$ ) slightly increased and then decreased to its initial concentration level or steady-state (Figure 6d1 and 2). This estimation can be supported by the previous experiment related to aspartate kinase (AK) [11]. AK is the first committed

enzyme of aspartate-family amino acid biosynthesis and regulated by the feedback inhibition by lysine and threonine (Figure 4). The flux analysis in *Lemna* revealed that the deregulation of AK, i.e., the increase of AK activity, caused accumulation of lysine, threonine, methionine, and AdoMet and decrease in the aspartate concentration [11]. Thus, in our study, a reduction in the AK activity through the feedback inhibition by supplementation of lysine and threonine might lead to an initial increase in the aspartate concentration as predicted in modified U-system. The transient increases in glutamate and aspartate concentrations ( $X_{37}$  and  $X_{78}$ , respectively) after lysine and threonine supplementation were also indirectly supported by the increase of proline ( $X_{50}$ ) observed both in simulation and metabolome analysis (Figure 5b). One could assume that if the simulated behaviors of dead-end metabolite concentrations (proline ( $X_{50}$ ) in this case) agree well with the experimental data, those of intermediate metabolites (glutamate ( $X_{37}$ ) and aspartate ( $X_{78}$ ) in this case) are likely to be correct. This consistency implies that the prediction of metabolic behavior by our approach is reliable even if the quality of the experimental data for some metabolites is not very high.

#### **Prediction of dynamic behaviors of unmeasurable metabolite concentrations**

The 86 out of 268 metabolite concentrations measured by metabolome analysis (purple and blue boxes in Figure 4) were included in the U-system model. The 66 out of 86 metabolite concentrations (blue boxes in Figure 4) showed unclear dynamic patterns changing throughout the period of 96 h after the supplementation of lysine and threonine (Supplementary Figures - additional file 3). One may note that these metabolite concentrations showing no significant variations do not change or change only slightly to the extent that cannot be experimentally observed due to biological fluctuations and analytical constraints. Thus, although a large number of metabolite concentrations in metabolic pathways can be simultaneously measured, the metabolic behaviors assumed from measured metabolite concentrations may sometimes not be entirely reliable. Also, the 265 out of 351 metabolites incorporated in the U-system model could not be detected by this metabolome analysis. Given a computational model, we may be able to predict the dynamic responses of metabolites that cannot be measured reliably nor detected at all. The computational results also have a potential to provide some types of theoretical validation on the accuracy of the experimental data.

Figure 7 illustrates that the relative concentration of 1-hydroxy-3-indolylmethyl GSL ( $X_{257}$ ) measured by metabolome analysis (Figure 7a) did not change throughout the period of 96 h whereas that of 4-methoxy-3-indolylmethyl GSL ( $X_{261}$ ) (Figure 7b) started increasing at  $t=60$  h. Interestingly, in the U-system

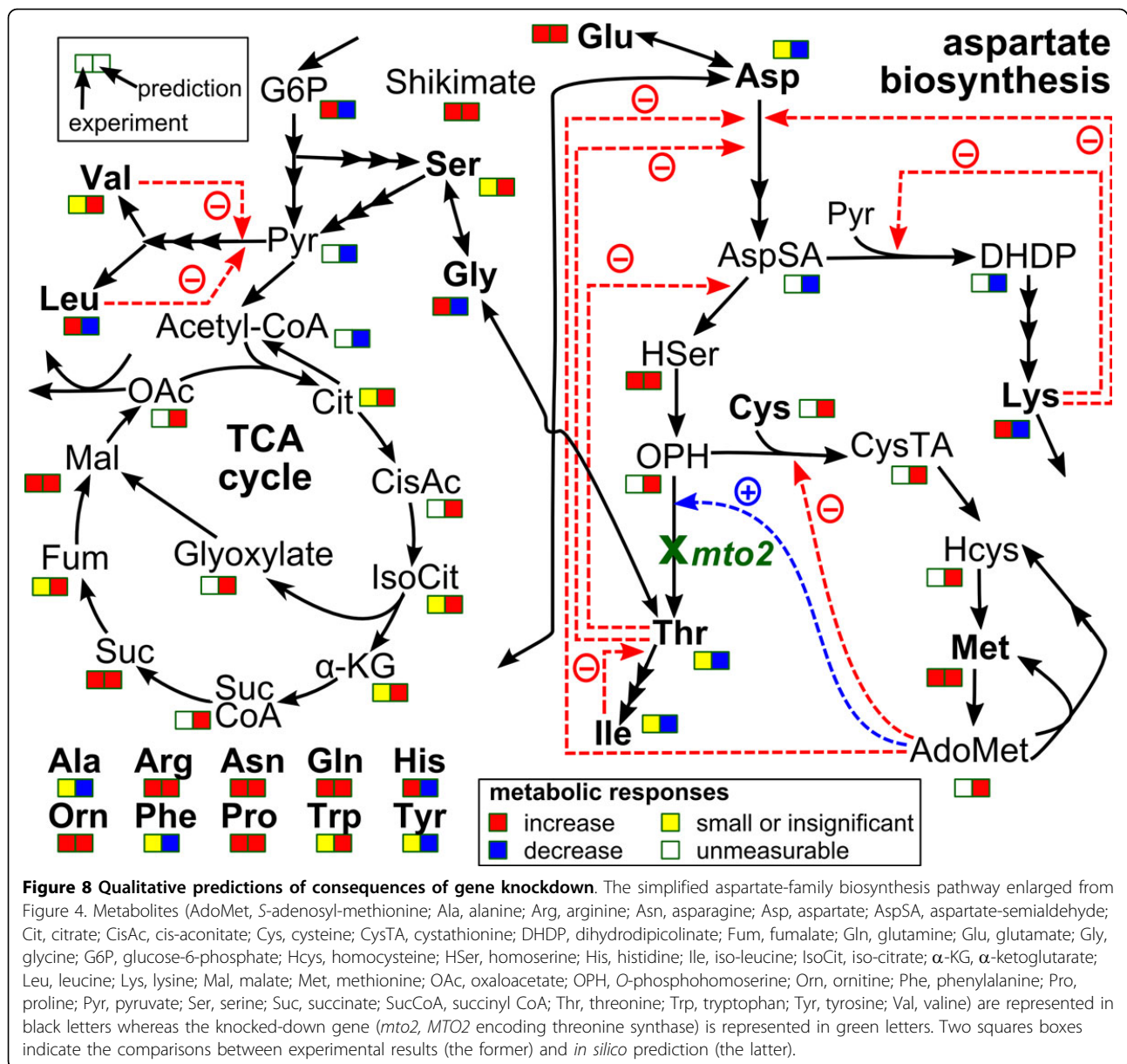


simulations 1-hydroxy-3-indolylmethyl GSL concentration ( $X_{257}$ ) started increasing at around  $t=55$  h and then 4-methoxy-3-indolylmethyl GSL concentration ( $X_{261}$ ) increased around  $t=60$  h. In this case, the dynamic behaviors of 1-hydroxy-3-indolylmethyl GSL ( $X_{257}$ ) from model prediction were inconsistent with the relative concentrations obtained by metabolome analysis. In the metabolic map, the reactions between  $X_{257}$  and  $X_{261}$  are linearly connected without any regulations, *i.e.*, inhibition and activation (green zone in Figure 4) so that  $X_{257}$  and  $X_{261}$  are expected to behave in the similar manner as observed in the predictive simulations. This inconsistency between actual data and computational prediction suggests two possibilities; that is, an unknown regulatory mechanism may exist or 1-hydroxy-3-indolylmethyl GSL concentration ( $X_{257}$ ) might have not be reliably analyzed.

Moreover, we were able to predict the behaviors of concentrations of aspartate-4-semialdehyde ( $X_{80}$ ) (Figure 7d) and O-phospho-homoserine ( $X_{82}$ ) (Figure 7e), which can hardly be detected due to technical limitations. They are not only located at important branched points in the metabolic map but also related to many regulations including both inhibitions and activations (Figure 8). Thus, the information of these metabolites could allow us to comprehend metabolic system. Again, the aspartate-family amino acid biosynthesis includes various inhibitions and activations, so that these dynamic behaviors will not be reasonably observed by normal mass action equations.

#### Prediction of “unpredictable” metabolic behaviors

The modified U-system enables us to mathematically model a relatively large-scale metabolic system, and hence simulations based on the modified U-system provide predictive dynamic behaviors of metabolite concentrations which are hardly provided by existing methods. The modified U-system model for the *Arabidopsis* metabolic system predicted that the supplementation of lysine and threonine did not affect 4-methoxy-3-indolylmethyl GSL concentration ( $X_{261}$ ) until  $t=60$  h (Figure 7b), nor dihydrouracil concentration ( $X_{333}$ ) in pyrimidine biosynthesis until  $t=40$  h (Figure 7c). These results suggested that the supplementation of lysine and threonine influenced  $X_{261}$  and  $X_{333}$  in a delayed fashion due to metabolic flow passing through various reactions in long and complicated pathways (see Figure 4). The U-system predictions of the time-transient behaviors of metabolite concentrations were reasonable and this kind of metabolic relationship between distant metabolites on metabolic map cannot be examined by either statistical methods or typical mathematical modeling without delay functions. Furthermore, similar predictive simulations of several metabolite located in long and complicated pathways such as xanthine ( $X_{299}$ ) and ureidopropionate ( $X_{344}$ ) in purine and pyrimidine biosynthesis (gray zone in Figure 4), and amino adipate ( $X_{153}$ ), glutarate ( $X_{152}$ ), and pipecolate ( $X_{141}$ ) in lysine degradation (yellow zone in Figure 4) reasonably agreed with the experimental results (Supplementary Figures - additional file 3). The



results could be an example for a large-scale system to indirectly validate the U-system approach. They also imply that predicted behaviors of unmeasurable metabolites located between detectable metabolites are probably correct, at least, qualitatively.

#### Qualitative predictions of consequences of gene modifications

A challenge in systems biology, in addition to the prediction of metabolic behaviors, is the prediction of metabolic responses of an altered network structure, in which, for instance, gene expression or enzyme activities are modified. To assess whether the modified U-system model is applicable for this kind of prediction, an *in silico* knockdown for

*MTO2* encoding threonine synthase gene of *Arabidopsis* in the biosynthetic pathway of aspartate-family amino acids (Figure 8 enlarged from Figure 4). When the flux controlled by *MTO2* was 90% reduced *in silico*, the threonine concentration was predicted to decrease whereas the methionine concentration was predicted to increase. In overall, the 6 out of 7 amino acids or 85% prediction agreed with the original result from biological experiment using an *Arabidopsis mto2* mutant [17]. The prediction was further compared to the result from metabolome analysis [18]. Figure 8 showed that 73% correctness from 11 metabolites showing the significant changes comparing with the wild-type whereas the metabolome and amino acid data showed



only 50% agreement among 6 amino acids. The incorrect prediction might be because of the insufficient information for model input which could be considered to be an unidentified pathway. Nevertheless, changes of all unmeasurable metabolites can also be qualitatively predicted using our model. The result suggests that the U-system approach is capable of predicting metabolic levels when the system is modified, even if specific information on complicated regulatory mechanisms is not available. This information can aid the design of next-step experiments.

## Conclusions

The U-system approach is simple, flexible and can be practically exploited to predict the coarse time-transient behavior of metabolites in a large-scale metabolic system. It is also capable of predicting metabolic behaviors when the system is modified. The U-system approach has two major characteristics: 1) the U-system model is presented in a normalized form, and 2) the U-system model constructed only using partial, currently-available metabolic pathways. Recent advances in metabolomics technology have enabled us to obtain large-scale semi-quantitative datasets for hundreds of metabolite concentrations. The first characteristic is essential to establish mathematical models for actual metabolic systems consisting of hundreds to thousands of metabolites, whose concentrations cannot be measured practically by existing methods of measurement other than metabolomics. The second characteristic is desirable for the coarse characterization of complicated systems, because metabolic responses mainly depend on the network structure, rather than parameter values, as we have shown in this study. Thus, the U-system approach is not only informatively useful for data analysis and model diagnostics but it also can aid the design of next-step experiments.

## Methods

### Background on Biochemical Systems Theory (BST)

#### 1. Structure of equations

BST uses ordinary differential equation systems of a specific format: Each process is represented with a product of power-law functions. This format is the result of Taylor approximation in a logarithmic space [4,9,19,20]. The change in each variable of a system is thus represented as a so-called Generalized Mass Action (GMA)-system as follows:

$$\frac{dX_i}{dt} - \sum_{k=1}^p \alpha_{ik} \prod_{j=1}^n X_j^{g_{ijk}} \prod_{j=1}^m Y_j^{G_{ijk}} - \sum_{k=1}^q \beta_{ik} \prod_{j=1}^n X_j^{h_{ijk}} \prod_{j=1}^m Y_j^{H_{ijk}} \quad i = 1, 2, \dots, n \quad (5)$$

Here,  $X_i$  ( $i=1, \dots, n$ ) are dependent variables (typically metabolite concentrations),  $Y_j$  ( $j=1, \dots, m$ ) are independent variables (typically external substrates or enzyme activities),  $\alpha_{ik}$  ( $i=1, \dots, n; k=1, \dots, p$ ) and  $\beta_{ik}$  ( $i=1, \dots, n; k=1, \dots, q$ )

are the non-negative rate constants for influxes and effluxes, respectively,  $g_{ijk}$  ( $i=1, \dots, n; j=1, \dots, m; k=1, \dots, p$ ) and  $h_{ijk}$  ( $i=1, \dots, n; j=1, \dots, m; k=1, \dots, q$ ) are real-valued kinetic orders associated with dependent variables in influxes and effluxes, respectively, while  $G_{ijk}$  ( $i=1, \dots, n; j=1, \dots, m; k=1, \dots, p$ ) and  $H_{ijk}$  ( $i=1, \dots, n; j=1, \dots, m; k=1, \dots, q$ ) are those associated with independent variables.  $n$  and  $m$  are the numbers of dependent and independent variables, respectively,  $p$  and  $q$  are the maximum numbers of influxes and effluxes, respectively, and  $t$  is time. Details have been documented in the literature many times, e.g. in [6,7,9,21].

#### 2. Validity and accuracy of BST for modeling biological systems

Since the power-law format in BST appears to be quite restrictive, the accuracy of different model variants within BST has been assessed many times. As Taylor approximations, the models are mathematically guaranteed to be correct at an operating point of choice, but their ranges of validity were also tested against results from more traditional formulations like the Michaelis-Menten formalism. These comparisons demonstrated sufficient accuracy of BST models over quite large ranges in variation of the involved variables [22-25]. Applications of BST models were also demonstrated with various large-scale systems, including the TCA cycle [26], ethanol fermentation [27], purine metabolism [28], and sphingolipid metabolism [29].

#### Meaning of parameters and U-system approach

Model parameters of pathway systems are usually estimated from kinetic information or from metabolic time series data. In Eq. (5), the familiar kinetic parameters involved in Michaelis-Menten kinetics (maximum reaction rate  $V_{max}$ , Michaelis constant  $K_M$ , etc.) are not explicitly visible, but they are instead distributed to the rate constants and kinetic orders. The BST parameters have their own meaning either as turn-over rates or as the specific effect that a variable has on the flux term in which it is involved. Thus, if a kinetic order is negative, the variable with which it is associated has a negative (inhibitory) effect on the flux term. This effect is given exclusively by the one appropriate kinetic order, thereby reducing the number of kinetic parameters in the system, which can be much higher in traditional rate functions. Still, the simplified BST parameters are able to capture time-transient behaviors of metabolite concentrations and other characteristics of metabolic reaction systems very well.

A consequence of the direct meaning of kinetic orders in BST is that it is possible to set up differential equations in symbolic notation, solely based on the structure of a metabolic pathway and its regulatory structure, whereas deep knowledge of their regulatory mechanisms is not needed. Furthermore, the kinetic parameters in Eq. (5) can be

estimated directly from time-series data of metabolite concentrations, and subsequently convey important information about the system characteristics. In some cases, traditional kinetic parameters from the literature, such as  $V_{max}$ , can be converted into BST parameters, but this is seldom possible on the basis of metabolome data, which are often presented as relative rather than absolute concentrations.

It was recently shown [30] that a coarse grid of kinetic orders is sufficient to fit time developments in pathway systems with surprisingly high accuracy. Expanding on this idea, we here fix all kinetic orders to 1, 0.5 or -0.5 to represent enzyme kinetics, positive or negative effects of metabolites, respectively. In other words, we use a hybrid between a pure mass action system and a GMA system, which allows for inhibitory and activating effects that are modeled with kinetic orders of -0.5 and 0.5. The rate constants are at first also set to 1, but may be later subjected to order-of-magnitude adjustments, which obviously improve the model representation of the data. Thus, at branched points, the equations are adjusted based on stoichiometry. For example, in a branched reaction of the type shown in Figure 1c the  $X_1$  is converted into  $X_2$  and  $X_4$ , so that the equation is  $dX_1/dt = V_{inX1} - V_{outX1-1} - V_{outX1-2}$ . Since most of the parameters are set at unity in the proposed approach, we name this system as Unity-system or U-system.

Fixing the parameters values is to some degree compensated by adjustments in metabolite concentrations. In other words, if the kinetic orders are forced to be 1, +0.5, or -0.5, a data fit will lead to adjusted metabolite concentrations.

#### Monte-Carlo simulations

Monte-Carlo simulations were employed to test in an unbiased manner whether the qualitative behavior of a model is consistent with the U-system representation. Most Monte-Carlo simulations were calculated for 1,000 loops. The parameter values were generated randomly from uniform distributions. The ranges for rate constants and kinetic orders were between 0.5 and 20, and  $\pm 0.2$  and  $\pm 0.8$ , respectively.

#### Calculation method

Differential equations were solved using automatic switching for stiff and non-stiff problems (LSODA) [31] using the ODE integration package from SciPy (Scientific Tools for Python) [32]. All visualizations were created using Matplotlib [33].

#### Details of the *Arabidopsis* system

##### *Arabidopsis callus culture*

*Arabidopsis thaliana* liquid callus culture derived from accession Col-0 was prepared as described in Murota et al.

[34] with slight modifications. For callus induction, minced seedlings were incubated in RM28 medium under constant light. The medium was changed every 6 days. For a metabolic perturbation experiment, RM28 medium supplemented with 10 mM L-lysine and 1 mM L-threonine [35] was used at the third medium change. For a control experiment, RM28 without supplementation was used. Sucrose in RM28 medium was a sole carbon source for callus culture. The experiments were carried out in triplicate.

For both metabolome and amino acid analyses, calli were collected prior to lysine and threonine treatment (0 h), and 2, 6, 12, 24, 36, 48, 60, 72, 84 and 96 h after the treatment. The calli were immediately frozen in liquid nitrogen and stored at  $-80^\circ\text{C}$ . Prior to analyses, the frozen samples were lyophilized using a freeze dryer (FDU-2100, EYELA) in a vacuum.

##### *Metabolome analysis*

Metabolites were extracted by homogenizing lyophilized callus in 500  $\mu\text{L}$  80% methanol in 0.1% formic acid per 2 mg dry weight callus with 5 mm zirconia beads (no.5-4060-13, AS ONE Co. Ltd.) in 2.0 mL sampling tubes (no.132-620C, WATSON Co., LTD) for 5 min using shake master NEO (Bio Medical Science, Tokyo, Japan). After centrifugation using a high speed refrigerated micro centrifuge (TOMY MX-300) at 14,000 r.p.m. at  $4^\circ\text{C}$ , 250  $\mu\text{L}$  supernatant was dried up in 96 well plate and the residue was dissolved in 120  $\mu\text{L}$  ultrapure water (no. 210-01303, Wako Pure Chemical Industries, Ltd.). One  $\mu\text{L}$  of the solution was subjected to widely targeted metabolome analysis [36] by LC-MS using UPLC-TQD system (Waters, Milford, MA, USA). The peak intensities were normalized by those of internal standards (10-campher sulfonate and lidocaine). For mathematical modeling, the normalized peak intensities of treated samples were subtracted from those of control samples to eliminate effects of cell growth and obtain relative metabolite concentrations.

##### *Amino acid analysis*

Metabolites were extracted by homogenizing lyophilized callus in 500  $\mu\text{L}$  of the extraction solution (methanol: milliQ water = 4:1) per 2 mg dry weight in 2.0 mL sampling tubes using Mixer Mill MM300. After centrifugation at 15,000 r.p.m. at  $25^\circ\text{C}$  for 10 min, the 400  $\mu\text{L}$  supernatant was used for the analysis performed according to the protocol based on the EZ:faast amino acid derivatization technique for GC-MS (Phenomenex, Torrance, CA) with slight modification. The reagent 1 (internal standard solution) was diluted to 10-fold while the reagent 6 (re-dissolution solvent) was added only 50  $\mu\text{L}$  to concentrate amino acid concentrations in samples. Then, 1  $\mu\text{L}$  of solution was subjected to amino acid analysis by GC-MS using GCMS-QP2010 Plus (Shimadzu, Kyoto, Japan).

### Statistical evaluation

The correlation coefficients, testing for the significance of the correlation coefficients, and student's *t* continuous random variable based on a survival function were calculated for comparing model simulations with experimental data.

### *Arabidopsis callus model*

Information for metabolic pathways including their inhibition and activation was assembled from KEGG (<http://www.genome.jp/kegg/>) [37,38] and AraCyc (<http://arabidopsis.org/tools/aracyc/>) [39] databases. A mathematical model was constructed based on the metabolic reaction network including 351 metabolites and 441 fluxes, which cover a wide variety of metabolic pathways related to amino acid biosynthesis, glycolysis pathway, TCA cycle, pentose phosphate pathway, GSL biosynthesis, and purine and pyrimidine biosynthesis (Figure 4; Supplementary Model - additional file 2). The relative concentrations of 86 metabolites among 268 analyzed in metabolome analysis were included in the model. Twenty (square purple boxes in Figure 4) out of the 86 metabolite concentrations included in the model (23.2%) showed considerable changes.

The present study employs a GMA-system model to which the U-system approach was applied for characterizing metabolic behaviors in a large-scale metabolic reaction network related to aspartate-derived amino acids, namely, aspartate, lysine, threonine, and methionine. The effects of feedback inhibition and activation were also included based on information available from the literature and pertinent databases [12,40]. Kinetic orders of metabolites were set to 1, while kinetic orders for inhibition and activation signals were set to -0.5 and 0.5, respectively. In the experiments to obtain relative metabolite concentrations, lysine and threonine was supplemented to the culture medium and transported into the callus cells. Then Fick's law of diffusion [16] was included in the equations for the supplemented metabolites lysine and threonine. Also, their parameters were appropriately fitted with available data of Lys and Thr obtained by metabolome analysis. Metabolite concentrations at the steady state were calculated by simultaneously solving the ordinary differential equations for all 351 metabolites (Figure 4). Since these concentrations were optimized according to the mass balance of the system, the magnitude of the concentration values become smaller as the location of metabolites on the metabolic map were farther away from glucose, which was the main initial metabolite in the present study. It is noted that time-series metabolome data from the experiment were given in relative rather than absolute concentrations. Thus, it is reasonable to take into account the magnitudes of metabolite concentrations while considering predicted behaviors of the metabolite concentrations.

### Additional material

**Additional file 1:** Ussystem\_SupplementaryInformation.pdf. Additional figures for the branched pathway model. Figure S1. Linear pathway with inhibition ( $\ominus$ ).  $X_1$  and  $X_2$  represent metabolites whereas  $Y_1$  and  $Y_2$  represent enzyme activities for the  $X_1$  influx and  $X_2$  efflux, respectively. Figure S2. Comparisons of time courses of  $X_1$  and  $X_2$  obtained by the Michaelis-Menten model and those by the simplified U-system approach of the corresponding GMA model. The x-axis represents time, while the y-axis represents the concentrations of  $X_i$ . The red lines pertain to the actual system in left and bottom axes whereas the blue dotted lines are for U-system in right and top axes. a The  $X_1$  U-system concentration compared with real concentration when the values of  $X_1$  was increased two-fold at  $t=0$ . b The  $X_2$  U-system concentration compared with real concentration when the values of  $X_1$  was increased two-fold at  $t=0$ . c The  $X_1$  U-system concentration compared with real concentration when the values of  $X_2$  was increased two-fold at  $t=0$ . d The  $X_2$  U-system concentration compared with real concentration when the values of  $X_2$  was increased two-fold at  $t=0$ . Figure S3. Normalized concentrations of  $X_i$  with variations of parameters  $\alpha_i$ , which value from 0.5 (red), 1, 5, 10, 15, 20, 25, 50, 75 to 100 (blue). Figure S4. Normalized concentrations of  $X_i$  with variations of parameters  $\beta_i$ , which value from 0.5 (red), 1, 5, 10, 15, 20, 25, 50, 75 to 100 (blue). Figure S5. Normalized concentrations of  $X_i$  with variations of parameters  $g_{ij}$  which range from 0.1 (or -0.1; red) to 1.0 (or -1.0; blue). Figure S6. Normalized concentrations of  $X_i$  with variations of parameters  $h_{ij}$  which range from 0.1 (red) to 1.0 (blue).

**Additional file 2:** Ussystem\_SupplementaryModel.pdf. The details for *Arabidopsis* model.

**Additional file 3:** Ussystem\_SupplementaryFigures.pdf. The U-system simulations comparing with experimental data from metabolome and amino acids analysis.

### Competing interests

The authors declare that they have no competing interests.

### Authors' contributions

SN and MYH conceived the experimental concepts and designed callus culture experiment. KS, EOV, FS and MYH conceived the theoretical concepts and designed modeling approach. YC, YY and HO planned and performed callus culture experiments. YS conducted metabolome analysis by LC-MS. KS conducted amino acid analysis by GC-MS, and performed modeling and simulations. KS, TF, SN, EOV, FS and MYH performed data analyses. KS, SK, HO, TF, SN, EOV, FS and MYH performed data interpretation. KS, YS, SN, EOV, FS, MYH wrote and edited the manuscript. All authors read and approved the final manuscript.

### Acknowledgements

We thank Mss. Saeko Yasokawa, Eriko Tanaka and Hitomi Sekihara for technical assistance in callus cultures, Ms. Akane Sakata and Mrs. Muneo Sato and Yutaka Yamada for technical assistance in metabolome analysis, and Mr. Hiroshi Kiyota for technical assistance in amino acid analysis. This work was supported in part by the Japan Science and Technology Agency, CREST to MYH, MEXT KAKENHI grant numbers 25119719 to FS and 22119006 to SN, JSPS KAKENHI grant number 25660084 to MYH, Grant-in-Aid from the NC-CARP project of MEXT to MYH, and a grant from the U.S. National Science Foundation (Project MCB-0946595) to EOV.

### Declarations

Publication charges for this article were funded by KAKENHI grant number 25660084.

This article has been published as part of *BMC systems Biology* Volume 8 Supplement 5, 2014: Proceedings of the 25th International Conference on Genome Informatics (GIW/ISCB-Asia): Systems Biology. The full contents of the supplement are available online at <http://www.biomedcentral.com/bmcsystbiol/supplements/8/S5>.

#### Authors' details

<sup>1</sup>RIKEN Center for Sustainable Resource Science, Yokohama 230-0045, Japan. <sup>2</sup>JST, CREST, Kawaguchi 332-0012, Japan. <sup>3</sup>Graduate School of Life Science, Hokkaido University, Sapporo 060-0810, Japan. <sup>4</sup>Faculty of Science, Hokkaido University, Sapporo 060-0810, Japan. <sup>5</sup>Graduate School of Information Science, NARA Institute of Science and Technology, Nara 630-0192, Japan. <sup>6</sup>Graduate School of Agriculture, Hokkaido University, Sapporo 060-8586, Japan. <sup>7</sup>Graduate School of Agricultural and Life Sciences, University of Tokyo, Tokyo 113-8657, Japan. <sup>8</sup>The Wallace H. Coulter Department of Biomedical Engineering, Georgia Institute of Technology and Emory University, Atlanta, GA 30332-0535, USA. <sup>9</sup>Graduate School of Bioresource and Bioenvironmental Sciences, Kyushu University, Fukuoka 812-8581, Japan.

Published: 12 December 2014

#### References

1. Fiehn O: **Metabolomics - the link between genotypes and phenotypes.** *Plant Mol Biol* 2002, **48**:155-171.
2. Schauer N, Fernie AR: **Plant metabolomics: towards biological function and mechanism.** *Trends Plant Sci* 2006, **11**:508-516.
3. Weckwerth W: **Metabolomics in systems biology.** *Annu Rev Plant Biol* 2003, **54**:669-689.
4. Savageau MA: **Biochemical systems analysis I: Some mathematical properties of the rate law for the component enzymatic reactions.** *J Theor Biol* 1969, **25**:365-369.
5. Voit EO: **What if the fit is unfit? Criteria for biological systems estimation beyond residual errors.** In *Applied Statistics for Biological Networks*. New York: J. Wiley and Sons; M. Dehmer, F. Emmert-Streib and A. Salvador 2011:183-200.
6. Jia G, Stephanopoulos GN, Gunawan R: **Parameter estimation of kinetic models from metabolic profiles: Two-phase dynamic decoupling method.** *Bioinformatics* 2011, **27**:1964-1970.
7. Kutalik Z, Tucker W, Moulton V: **S-system parameter estimation for noisy metabolic profiles using Newton-flow analysis.** *IET Syst Biol* 2007, **1**:174-180.
8. Voit EO, Almeida J: **Decoupling dynamical systems for pathway identification from metabolic profiles.** *Bioinformatics* 2004, **20**:1670-1681.
9. Voit EO: **Computational analysis of biochemical systems: A practical guide for biochemists and molecular biologists.** United Kingdom: Cambridge University Press; 2000.
10. Buchanan BB, Gruissem W, Jones RL: **Biochemistry and molecular biology of plants.** Courier Companies, Inc; 2000.
11. Giovanelli J, Mudd SH, Datko AH: **Regulatory structure of the biosynthetic pathway for the aspartate family of amino acids in *Lemna paucicostata* Hegelm. 6746, with special reference to the role of aspartokinase.** *Plant Physiol* 1989, **90**:1584-1599.
12. Curien G, Bastien O, Robert-Genthon M, Cornish-Bowden A, Cardenas ML, Dumas R: **Understanding the regulation of aspartate metabolism using a model based on measured kinetic parameters.** *Mol Syst Biol* 2009, **5**.
13. Nagele T, Henkel S, Hormiller I, Sauter T, Sawodny O, Ederer M, Heyer AG: **Mathematical modeling of the central carbohydrate metabolism in *Arabidopsis* reveals a substantial regulatory influence of vacuolar invertase on whole plant carbon metabolism.** *Plant Physiol* 2010, **153**:260-272.
14. 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 USA* 2012, **109**:339-344.
15. Poolman MG, Miguet L, Sweetlove LJ, Fell DA: **A genome-scale metabolic model of *Arabidopsis* and some of its properties.** *Plant Physiol* 2009, **151**:1570-1581.
16. Faghri A, Zhang Y, Howell J: **Advanced Heat and Mass Transfer.** Global Digital Press; 2010.
17. Bartlem D, Lambein I, Okamoto T, Itaya A, Uda Y, Kijima F, Tamaki Y, Nambara E, Naito S: **Mutation in the threonine synthase gene results in an over-accumulation of soluble methionine in *Arabidopsis*.** *Plant Physiol* 2000, **123**:101-110.
18. Kusano M, Fukushima A, Redestig H, Kobayashi M, Otsuki H, Onouchi H, Naito S, Hirai MY, Saito K: **Comparative metabolomics charts the impact of genotype-dependent methionine accumulation in *Arabidopsis thaliana*.** *Amino Acids* 2010, **39**:1013-1021.
19. Savageau MA: **Biochemical systems analysis II: The steady-state solutions for an n-pool systems using a power-law approximation.** *J Theor Biol* 1969, **25**:370-379.
20. Savageau MA: **Biochemical systems analysis III: Dynamic solutions using a power-law approximation.** *J Theor Biol* 1970, **26**:215-226.
21. Voit EO: **Biochemical systems theory: A review.** *ISRN Biomath* 2013, **2013**:1-53.
22. Voit EO, Savageau MA: **Accuracy of alternative representations for integrated biochemical systems.** *Biochem* 1987, **26**:6869-6880.
23. Sorribas A, Savageau MA: **Strategies for representing metabolic pathways within biochemical systems theory: Reversible pathways.** *Math Biosci* 1989, **94**:239-269.
24. Sorribas A, Savageau MA: **A comparison of variant theories of intact biochemical systems. I. Enzyme-enzyme interactions and biochemical systems theory.** *Math Biosci* 1989, **94**:161-193.
25. Sorribas A, Savageau MA: **A comparison of variant theories of intact biochemical systems. II. Flux-oriented and metabolic control theories.** *Math Biosci* 1989, **94**:195-238.
26. Shiraishi F, Savageau M: **The tricarboxylic acid cycle in *Dictyostelium discoideum*. I. Formulation of alternative kinetic representations.** *J Biol Chem* 1992, **267**:22912-22918.
27. Cascante M, Curto R, Sorribas A: **Comparative characterization of the fermentation pathway of *Saccharomyces cerevisiae* using biochemical systems theory and metabolic control analysis: Steady-state analysis.** *Math Biosci* 1995, **130**:51-69.
28. Curto R, Voit EO, Sorribas A, Cascante M: **Mathematical models of purine metabolisms in man.** *Math Biosci* 1998, **151**:1-49.
29. Alvarez-Vasquez F, Sims KJ, Cowart LA, Okamoto Y, Voit EO, Hannun YA: **Simulation and validation of modelled sphingolipid metabolism in *Saccharomyces cerevisiae*.** *Nature* 2005, **433**:425-430.
30. Iwata M, Shiraishi F, Voit EO: **Coarse but efficient identification of metabolic pathway system.** *Int J of Syst Biol* 2013, **4**:57-72.
31. Hindmarsh AC: **ODEPACK, A systematized collection of ODE solvers.** In *IMACS Transactions on Scientific Computation*. North-Holland, Amsterdam: R. S. S. E. 1983:55-64.
32. SciPy: **Open Source Scientific Tools for Python**: [http://www.scipy.org/].
33. Hunter JD: **Matplotlib: a 2D graphics environment.** *Comput Sci Eng* 2007, **9**:90-95.
34. Murota K, Hagiwara-Komoda Y, Komoda K, Onouchi H, Ishikawa M, Naito S: ***Arabidopsis* cell-free extract, ACE, a new in vitro translation system derived from *Arabidopsis* callus cultures.** *Plant Cell Physiol* 2011, **52**:1443-1453.
35. Datko AH, Mudd SH: **Methionine biosynthesis in *Lemna*: inhibitor studies.** *Plant Physiol* 1982, **69**:1070-1076.
36. Sawada Y, Akiyama K, Sakata A, Kuwahara A, Otsuki H, Sakurai T, Saito K, Hirai MY: **Widely targeted metabolomics based on large-scale MS/MS data for elucidating metabolite accumulation patterns in plants.** *Plant Cell Physiol* 2009, **50**:37-47.
37. Kanehisa M, Goto S, Sato Y, Furumichi M, Tanabe M: **KEGG for integration and interpretation of large-scale molecular datasets.** *Nucleic Acids Res* 2012, **40**:109-114.
38. Kanehisa M, Goto S: **KEGG: Kyoto Encyclopedia of Genes and Genomes.** *Nucleic Acids Res* 2000, **28**:27-30.
39. Mueller LA, Zhang P, Rhee SY: **AraCyc: A Biochemical Pathway Database for *Arabidopsis*.** *Plant Physiol* 2003, **132**:453-460.
40. Goto DB, Onouchi H, Naito S: **Dynamics of methionine biosynthesis.** *Plant Biotechnol* 2005, **22**:379-388.

doi:10.1186/1752-0509-8-S5-S4

**Cite this article as:** Sriyudthsak et al.: A U-system approach for predicting metabolic behaviors and responses based on an alleged metabolic reaction network. *BMC Systems Biology* 2014 **8**(Suppl 5):S4.