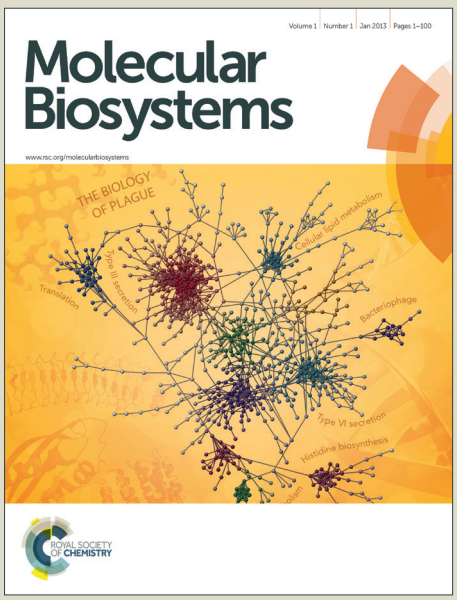
Molecular BioSystems

Accepted Manuscript



ROYAL SOCIETY OF CHEMISTRY

www.rsc.org/molecularbiosystems

Bacterial Metabolism in Immediate Response to Nutritional Perturbation with Temporal and Network View of Metabolites

Daichi Yukihira[1], Yoshinori Fujimura[2], Hiroyuki Wariishi[2, 3, 4], and Daisuke Miura[2 *]

[1]Graduate School of Bioresource and Bioenvironmental Sciences, Kyushu University, 6-10-1 Hakozaki, Higashi-ku, Fukuoka 812-8581, Japan

[2]Innovation Center for Medical Redox Navigation, Kyushu University, 3-1-1 Maidashi, Higashi-ku, Fukuoka 812-8582, Japan

[3]Bio-architecture Center, Kyushu University, 6-10-1 Hakozaki, Higashi-ku, Fukuoka 812-8581, Japan

[4]Faculty of Arts and Science, Kyushu University, 6-10-1 Hakozaki, Higashi-ku, Fukuoka 812-8581, Japan.

[*]Corresponding authors

Email addresses:

DY: yukihira@agr.kyushu-u.ac.jp

YF: fujimu@redoxnavi.med.kyushu-u.ac.jp

HW: hirowari@agr.kyushu-u.ac.jp

DM: daipon@agr.kyushu-u.ac.jp

Key words

metabolome, correlation network, temporal correlation, nutritional fluctuations

Abbreviations

9-AA 9-Aminoacridine

CRA Centering resonance analysis

GGM Gaussian graphical model

m/z Mass to charge

MALDI-MS Matrix-assisted laser/desorption ionization mass spectrometry

SPL Shortest path length

Abstract

In this study, the initial propagation of metabolic perturbation in *Escherichia coli* was visualized to understand the dynamic characteristics of the metabolic pathways without the associations of transcription alterations. *E. coli* cells were exposed to the sudden relief of glucose starvation, and time-dependent variances in metabolite balances were traced in the second scale. The acquired time-course data were represented by structural variations of the metabolite-metabolite correlation network. The initial correlation structure was altered immediately by the glucose pulse, followed by further structural variations within a few minutes. It was demonstrated that one metabolite temporally correlated with distinct metabolites with different timings, and such a behavior could imply a regulatory role for the metabolite in the metabolic network. Centrality analysis of the networks and partial correlation analysis indicated that preparation for growth and oxidative stress could be coupled as a structural property of the metabolic pathways.

1. Introduction

Adapting to fluctuations in carbon source availability is crucial for microorganisms to survive. Adaptation systems comprise molecular regulatory networks, and are invoked by their ability to recognize fluctuations. Previously, we demonstrated that variations in the intracellular metabolite levels in *Escherichia coli* progressed on a sub-minute scale following a glucose pulse.[1] This preliminary analysis indicated clearly that the variations in the intracellular metabolite levels progressed on a second scale. This time scale is relatively short compared with the time scale of alterations in gene expression, where Dikicioglu *et al.* detected a strong response in mRNA levels in *E. coli* a few minutes after changes in nutrient availability.[2] A longer time scale was also reported for variations in protein levels, where a temporal surge of transcription immediately promoted an increase in protein levels in yeast, resulting in minute- to hour-scale variations.[3] It has been suggested that the sensing system is associated not only with the abundance of carbon sources themselves, but also with the consequent perturbations in the intracellular metabolism.[4, 5] These studies indicated that that, although such immediate metabolic fluctuations progressed in a passive manner, the metabolic pathways must be organized to buffer and recognize metabolic fluctuations with certain specificity to the source of perturbation so as to realize the appropriate cellular responses. However, the temporal and structural characteristics of fluctuation propagation in an intrinsic time-scale are still unclear because studies of biological responses that focused on the time scales of biochemical processes have been emerged only recently.[6-9] Dynamic behaviors of molecular networks are of particular importance because cellular systems cannot be understood through static network structures alone.[10] Given these circumstances, we were encouraged to further investigate the time-dependent structural variation of a metabolic network in an early phase of metabolic fluctuations.

In addition to mRNAs and proteins, the structural properties of metabolite networks have been investigated.[11] Metabolomic network analysis has recently received much attention because the distribution of metabolites through a metabolic pathway (or flux) was not accessible by analyses at the mRNA or protein levels. Correlation analysis of metabolite levels is one approach to exploiting metabolome data efficiently, and this approach has been used to discover novel pathways[12] and to infer unknown metabolic networks.[13] Graph-based clustering techniques on metabolite networks helped to gather metabolic pathways that changed in response to different regulations and that were dependent on the tissue and/or genotype.[14] Szymanski *et al.* conducted metabolite pair-wise correlation analysis under various stress conditions to allow advanced observation beyond the change in metabolite concentration.[15] They found that a stable network, a commonly observed network under various stresses, had some components enriched for functionally-related biochemical pathways. On the other hand, Müller-Linow *et al.* reported that closeness in metabolomic correlation was not always an indicator of closeness in biochemical networks.[16] These findings imply that it is difficult to understand a correlation profile solely based on known metabolic pathways, and inversely that metabolite correlation provides important information to estimate previously unidentified functionality in a metabolic system.

In the present study, we integrated dynamic metabolome data and metabolite correlation analysis to understand the characteristics of the metabolic system of *E. coli* in response to environmental fluctuations. Starving *E. coli* cells were exposed to an instant glucose relief to induce a strong metabolic fluctuation, and time-course metabolome data were obtained in a time scale of seconds. It should be noted that a correlation itself does not provide temporal or contextual information about observed relationships, because such a correlation only statically represents the interdependency of two variables under a given condition. To address this issue, we formulated

the temporal behavior of metabolic fluctuation as an evolving network of metabolite correlation, where a sliding window was used to evaluate the transient correlation structure of metabolites in response to the fluctuation. The sequential propagation of the metabolic fluctuation was then investigated based on the time-dependent structural alteration of the metabolite correlation network. We have discussed possible ways to interpret the observed metabolite-metabolite correlations for a better understanding of the dynamics in cellular regulatory systems.

2. Materials and Methods

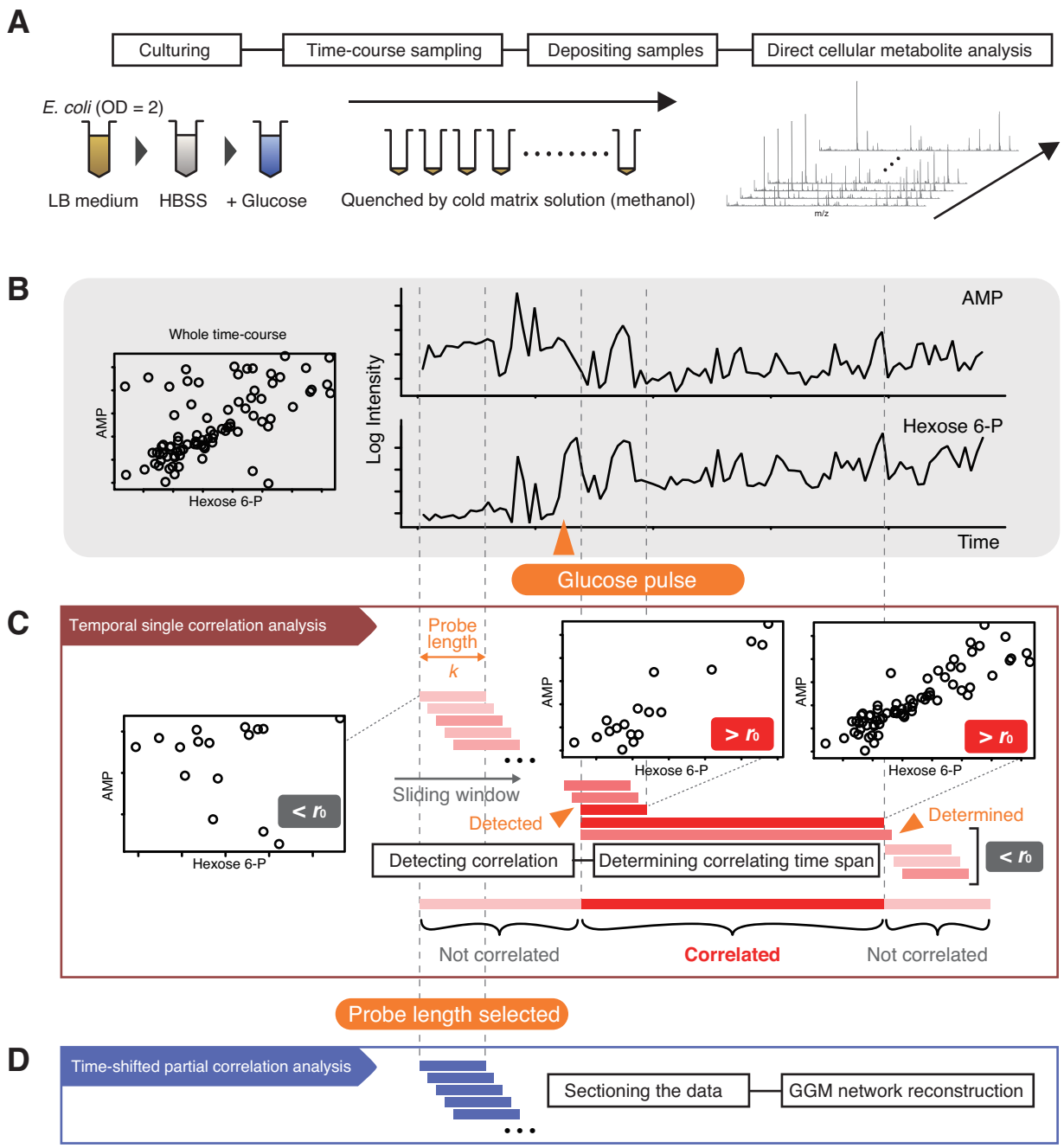
The workflow of the experiments and data analysis is summarized in Figure 1.

2.1. Acquisition of time-course metabolite data

Time-course metabolite data was acquired by a MALDI-MS-based high-throughput method introduced in the previous study.[1] *E. coli* cells were exposed to a sudden relief from glucose depletion by a pulse of glucose addition (a final concentration of 5% (w/v)), and cell samples were harvested from the suspension both before and after glucose addition. The sampling interval was fixed at 10 s. For each time-course sample acquisition, 24 samples were taken prior to the nutritional perturbation induction and 72 post-induction, resulting in a sample set of 96 time points over 16 min. Sample collection was repeated three times, resulting in three measured value per one time point. Step-by-step experimental methods can be found in Supplementary Information.

2.2. Genome-scale model of *E. coli* metabolic network

A previously reported genome-scale network of *E. coli, i*JO1366 was used as the reference network of metabolism.[17] This network model comprised 1136 compounds and 2551 reactions, reconstructed based on 1366 genes. The systems biology markup language (SBML)-type genome-scale network data was converted into a stoichiometric matrix using COBRA Toolbox[18] on MATLAB version R2012b. We excluded ubiquitous compounds that incorporate various biochemical reactions, which serve as a hub and reduce the path lengths among metabolites, based on their degree of connectivity in the genome-scale metabolic network. The matrix was imported to R environment and treated as a graph. Shortest path length (SPL) was calculated using the igraph package.[19]



2.3. Construction of temporal correlation profile matrix

A matrix $\mathbf{X}$ is the time course data of M metabolites with T discrete time points for observation. Consider the time-course data of two metabolite levels denoted as $\boldsymbol{x} = (\boldsymbol{x}_1, \boldsymbol{x}_2)$, where $\boldsymbol{x}_n = (x_1, x_2, \ldots, x_T)^T$. We evaluated the maximum time span in which a significant correlation was observed (Figure 1B and 1C). The subset of $\boldsymbol{x}$ with successive t time points starting from t_0-th time point

Figure 1. Workflows of experiments and data processing.

A. *E. coli* cells were exposed to a sudden relief from starvation. The cell suspension was continuously harvested and mixed with cold matrix solution (9-AA/methanol). Intracellular metabolites were detected by analyzing the cells directly by MALDI-MS. LB medium, Luria-Bertani medium, HBSS, Hank's balanced salt solution. **B.** Representative time-course data of fructose 6-phosphate and AMP are presented. Hexose 6-P, fructose 6-phosphate. The scatter plot (left) was constructed using the overall time point of the data set. **C.** Temporal single correlation analysis. The aim of analysis was to find the widest time range in which the levels of the metabolites showed as significant a correlation as possible. A short-span correlation was detected using a sliding window with a width of *k* time points. The width *k* and correlation threshold r_0 was tuned using a graph-theory evaluation (Scheme S1). **D.** Time-shifted partial correlation analysis. Temporal correlation network analysis based on graphical Gaussian modeling (GGM) was performed to illustrate the shift in the correlation structure of metabolite levels. Because GGM uses a partial correlation, the direct correlation network is provided.

was denoted as $\boldsymbol{x}'^{t}_{t_0} = (\boldsymbol{x}'_1, \boldsymbol{x}'_2)$, where $\boldsymbol{x}'_{\mathrm{n}} = \left(x_{t_0}, x_{t_0+1}, ..., x_{t_0+t-1}\right)^{\mathrm{T}}$. Function $\mathrm{cor}\left(\boldsymbol{x}\right)$ returns Spearman's rank correlation coefficient (r) of $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$. The statistical significance of the correlation coefficient was tested using a null hypothesis that $|r| = r_0$ and an alternative hypothesis that $|r| > r_0$. A one-way *t*-test was performed on the *t*-statistic $Z = \dfrac{\left|\xi_r - \xi_p\right|}{1/\sqrt{t-3}}$ with the alpha level at 0.05 ($P\left(Z\right) \leq 0.05$). When the correlation coefficient r was calculated by $\mathrm{cor}\left(\boldsymbol{x}^{t}_{t_0}\right)$, the resulting statistic Z was denoted as $Z\left(\boldsymbol{x}^{t}_{t_0}\right)$. ξ_r and ξ_p denote a Z-transformed score of the

sample correlation coefficient r and the population correlation coefficient r_0, calculated as $\zeta = \frac{1}{2}\ln\frac{1+|r|}{1-|r|}$. Here, the starting time point t_1 was determined with given r_0 and k as follows:

$$t_1 = \min\left\{t_a \middle| k \le t_a \le T-k+1,\ P\left(Z\left(x_{t_a}^{k}\right)\right) \le 0.05\right\}$$

Next, the maximum time span k' was determined as follows:

$$k' = \max\left\{k_a \middle| k_a \ge k,\ t_1 + k_a - 1 \le T,\ P\left(Z\left(x_{t_1}^{k_a}\right)\right) \le 0.05,\ P\left(Z\left(x_{t_1}^{k_a+1}\right)\right) > 0.05\right\}$$

Then t_1 and k' were represented as a binary vector with length T.

$$\left(\boldsymbol{b}\right)_t = \begin{cases} 1 & t_1 \le t \le t_1 + k' - 1 \\ 0 & \text{otherwise} \end{cases}$$

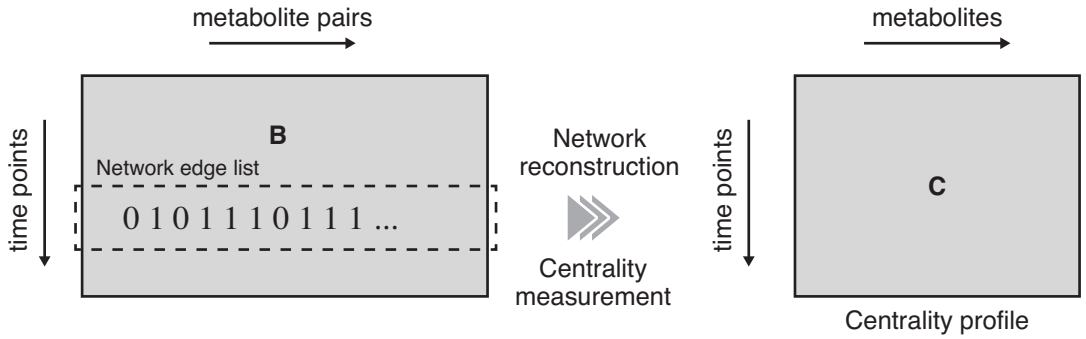
These evaluations were performed for each possible p metabolite pair, and the resulting series of $\boldsymbol{b}$ was combined to form matrix **B** with T rows and p columns, referred to as the correlation indicator matrix.

2.4. Temporal network analysis on single correlation profile

After parameter optimization, each row in **B** was reconstructed as a graph and its property was evaluated (Scheme 1). These graphs were considered as correlation networks at a given time point. Centrality was calculated for each of the time points, producing a centrality profile with length M. The resulting centrality profile matrix **C** thus has T rows and M columns. **C** was applied to correspondence analysis (CA) to visualize which compound has greater centrality at which time point. CA is one of dimension reduction methods, which is basically similar to principal component analysis, but scales data to treat rows and columns equivalently.[20] This approach is called centering resonance analysis (CRA), a method of network analysis that was originally designed to study complex discourse systems derived from a wide range of

sociological or psychological phenomena.[21] CRA has also been applied to feature extraction of evolving networks.[22] In the present study, CA was conducted using the MASS package for R .[20]

2.5. Transient GGM network analysis



Scheme 1. Centrality analysis of the single correlation profile matrix.

The graphical Gaussian modeling (GGM) framework[23] was employed to eliminate indirect interrelations. GGMs, also known as covariance selection models, are undirected graphical models where each relationship is conditioned on all remaining metabolites simultaneously. GGM modeling is based on partial Pearson correlation scores, calculated simply by inversion and normalization of the Pearson correlation matrix.[24] Partial correlation analysis was conducted using a sliding window with a length of k time points to illustrate a GGM network in a transient context (Figure 1D). A temporal subset of $\mathbf{X}$ starting from time point t with length k was denoted as $\mathbf{X}_t^k = (\boldsymbol{x}_t, \boldsymbol{x}_{t+1}, ..., \boldsymbol{x}_{t+k-1})$, where $0 \le t \le T - k + 1$. Vector $\boldsymbol{x}_t$ represents the peak intensities of M metabolites in a mass spectrum observed at time point t, $\boldsymbol{x} = (x_1, x_2, ..., x_M)^T$. The parameter k was set in accordance with the single correlation analysis. We estimated the partial correlations using GeneNet,[25] an R package that employs a shrinkage approach, which is suitable for data with a small sample size and a large number of variables. The calculation result for each possible metabolite pairs is a local false discovery rate (FDR) vector $\boldsymbol{q}_t$, whose elements are FDRs of the

corresponding partial correlations evaluated at t-th time window. $\boldsymbol{q}_t$ was then converted into a vector $\boldsymbol{b}'_t$ using an FDR cutoff threshold F_t.

$$\left(\boldsymbol{b}'\right)_t = \begin{cases} 1 & \left(\boldsymbol{q}\right)_t \le F_t \\ 0 & \text{otherwise} \end{cases}$$

F_t was adjusted for each t to achieve a consistent visualization of the correlation profile (see Supplementary Information). A series of $\boldsymbol{b}'$ was reconstructed into a correlation indicator matrix $\mathbf{B}'$.

3. Results and Discussion

3.1. Experimental design for high-throughput metabolite analysis

Previously, we developed a high-throughput analytical method that employed MALDI-MS using its semi-quantitative performance as well as its high-throughput characteristic to acquire detailed time course data on intracellular metabolites. With minimal experimental work, this method was able to trace the levels of phosphorylated metabolites such as sugar phosphates, nucleotides, nucleotide sugars, and cofactors, which play important roles in cellular metabolism.[26] In the present study, *E. coli* was exposed to a nutritional perturbation and its response was characterized by the temporal variations in metabolite levels. In the direct cell analysis using MALDI-MS, about 100 mass peaks were frequently detected. Of these, we identified 28 metabolites that were detected reproducibly throughout the time course (Table 1).

Table 1. Identified or estimated metabolites detected in the direct cell analysis using MALDI-MS.

Observed *m/z*	Metabolite	Abbreviation
259.0	Fructose 6-phosphate (hexose phosphate)	F6P
275.0	6-Phosphogluconate	6PG
306.1	Glutathione (reduced form)	GSH
321.0	Thymidine monophosphate	dTMP
322.1	Cytidine monophosphate	CMP
323.1	Uridine monophosphate	UMP
339.0	Fructose 1,6-bisphosphate	F16P
346.1	Adenosine monophosphate	AMP
362.1	Guanosine monophosphate	GMP
401.1	Thymidine diphosphate	dTDP
402.1	Cytidine diphosphate	CDP
403.0	Uridine diphosphate	UDP
426.1	Adenosine diphosphate	ADP

442.0	Guanosine diphosphate	GDP
481.0	Thymidine triphosphate	dTTP
482.0	Cytidine triphosphate	CTP
483.0	Uridine triphosphate	UTP
506.0	Adenosine triphosphate	ATP
522.0	Guanosine triphosphate	GTP
540.1	Nicotinamide adenine dinucleotide	NADH
545.1	Thymidine diphosphate 4-oxo-6-deoxy-glucose	dTDPg
565.1	Uridine diphosphate glucose	UDPG
588.1	Adenosine diphosphate glucose	ADPG
606.1	Uridine diphosphate *N*-acetylglucosamine	UDPGN
611.1	Glutathione (oxidized form)	GSSG
620.1	Nicotinamide adenine dinucleotide phosphate	NADPH
766.1	Coenzyme A	CoA
808.2	Acetyl-coenzyme A	AcCoA

3.2. Cellular energy state

We checked the time-dependent alteration of the ATP-ADP ratio in response to a glucose pulse. The ATP-ADP ratio represents the adenylate energy charge that should indicate the actual free energy of ATP hydrolysis that is available for cellular reactions.[27] The concept of the ratio is based on the assumption that, although the absolute individual amounts of ATP, ADP, and AMP might vary widely, the ratios of ATP and ADP, or ATP and AMP are more reliable indicators of metabolism. The time course of the ATP-ADP ratio indicated that the metabolic state of the cells changed within 30–40 sec in response to a nutritional fluctuation (Figure 1), which confirmed that the induced metabolic fluctuation was propagated in such a time scale as measured by the energetic state in the cells. It was of interest to estimate the relative variations in the levels of other metabolites. Therefore, we further investigated the variations in the metabolomic profile using graph theory, or a network view.

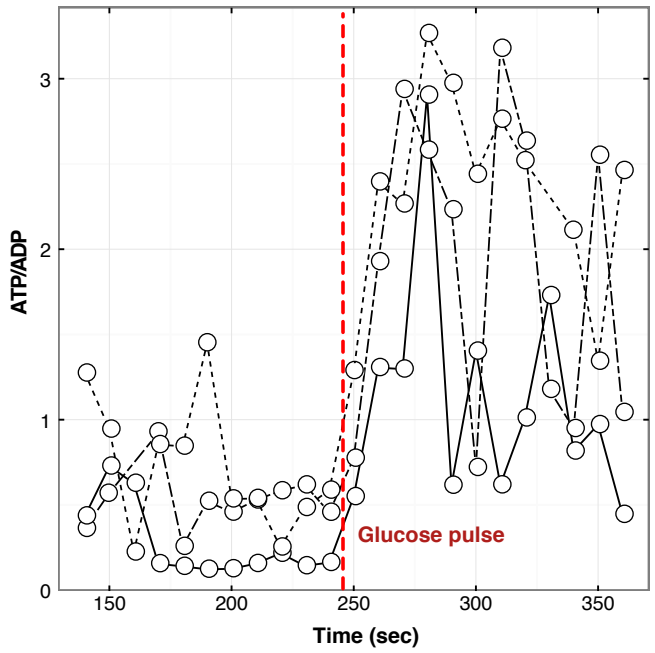


Figure 2. Time course of the ATP-ADP ratio before and after the glucose pulse. The peak intensity of ATP was divided by that of ADP at each time point. Triplicate data are shown. Not available points are omitted. Following the glucose pulse, the ATP-ADP ratio reached a maximum within three to four time points (corresponding to 30–40 s). The time points are in seconds.

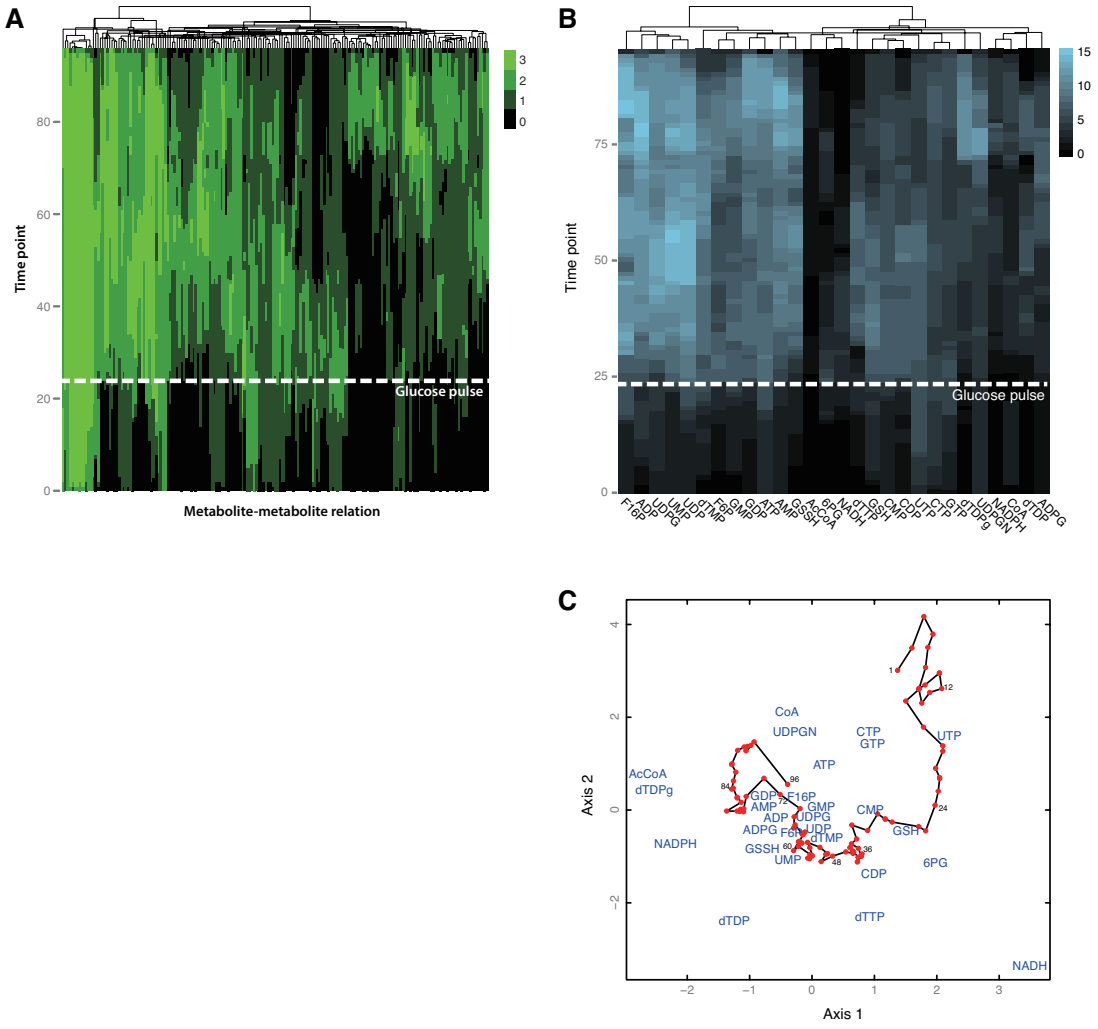
3.3. Formulation of transient metabolite correlation networks

A metabolite correlation approach is an unsupervised method that has been used to characterize the effects of environmental or gene variations as a fingerprint.[28] However, a non-stationary metabolic system could lead to a transient correlation structure, where significant correlations observed at a given time might disappear with in a short period of time, and *vice versa*. In terms of transient state, microbial metabolic fluctuation has been focused in precedent studies. [8, 9, 29-31] Compared with these studies, where metabolic variation was investigated *e.g.* within a few seconds to sub-seconds, the time span we used is significantly longer. It should be noted that, however, metabolic alteration in different time span would focus on different biological event,

and a relatively longer metabolic shift observed in the present study was associated with transcription alteration. Furthermore, the major difference of the present study could be found in the presumption; the correlation analysis was not dependent on any known metabolic pathways. Although our subject model, which employed bacterial metabolic fluctuation by a glucose pulse, has no doubt been characterized rationally based on the context of metabolic pathways, the presented methodology is fundamentally applicable to totally unknown metabolic (sub-)systems. Furthermore, considering that even with well-known metabolic pathways unexpected interactions are often found as crucial parts for regulation of metabolic system, the method is expected to facilitate initial prototyping the system in some transient states of interest. However, metabolite correlation analysis has rarely been discussed with dynamic contexts; therefore, we investigated several possible approaches to formulate correlational characteristics of the metabolite time-course data. .

The result of temporal single correlation analysis was reconstructed as a matrix of temporal correlation profile, whose rows served as the edge list for temporal correlation networks (Figure 3A). This representation of networks is often termed an evolving network, and is a natural extension of network analysis onto a temporal context. In constructing a temporal correlation profile, the adjustment of parameters should be considered systematically. Although the width of sliding window should be as short as possible to evaluate a short-term correlation, the sample size itself influences the quality of the detected correlation. The threshold level of the correlation coefficients is also critical for the resulting correlation network. These two parameters were optimized to give an ideal balance of the graph theory properties (graph density and modularity) of a temporal similarity network, which was reconstructed from the temporal correlation profile based on the similarity of the columns (Figure S1). With the adjusted parameters (k = 16 and r_0 =

0.85, which means that correlation coefficients of significantly greater than 0.85 were selected), numerous correlations emerged immediately in response to a glucose pulse at various temporal durations, reflecting immediate shifts in the cellular metabolic balance. This approach for parameter setting was intended to extract a module structure of the interdependency concealed within a complex correlation network by comparing its temporal traits; namely, the simultaneity of relationships, which was one of the properties of the temporal profiles that could not be investigated by static methods. With the parameter $k = 16$, metabolite level correlation was evaluated for a time range of 160 sec. While the significant metabolite correlations observed at a given time range do not directly imply any relationship of the metabolites in a biological context, the temporally clustered alteration of the metabolic network was expected to be associated with a similar phase of the regulatory system. We further investigated the variations in the temporal correlation profile using graph theory, or a network view.



3.4. Dynamics of the metabolic rearrangement with a network view

The temporal correlation profile based on the single correlation was highly complicated and required further analysis from different perspectives for better interpretation. To characterize the temporal trend of the metabolic shift from the viewpoint of the metabolites themselves, we examined the time-dependent variation of graph centrality of the nodes (metabolites) in the temporal correlation networks, reconstructed from the temporal correlation profile. Centrality is one of the measures of the importance of given nodes in the network, and can be measured by the degree of each metabolite node (the number of significant correlations that the metabolite has with other metabolites), referred to as degree centrality. Time-dependent variances in centrality

Figure 3. Temporal profile of a single correlation structure represented by a time course of correlation indicator and centrality analysis.

A. Each slot indicates the maximum time span when a significant correlation was detected for a pair of metabolites. The white dashed line indicates the time point when glucose was added. Triplicate results are overlaid. The profiles were clustered by hierarchical clustering using a complete linkage method. **B.** Comprehensive view of the time course variation of centrality. Because centrality can be evaluated when a network is given, time-dependent correlation networks were constructed for each time point of the correlation indicator matrix. **C.** Centering resonance analysis plot of the degree centrality profile. The connected points represent the time points in sequence. Metabolite names are plotted to indicate relative contribution to specific time points.

were calculated for all nodes, and the results are shown in Figure 3B. Several metabolites exhibited significant variation of centrality in response to the glucose pulse across a variety of durations. Considered inversely, the variation of the correlation network could be compressed into the variation of node centrality. To obtain a better view of this profile, we employed CRA. Briefly, CRA can be conducted as a CA of network centrality on a set of nodes evaluated under different network structures. Although it can be difficult to determine which relationship appears at each time point because the distances in the correspondence analysis are not Euclidean,[32] CRA can illustrate evolving structural characteristics of the correlation network, rather than the component associations themselves.[33] The centrality matrix (Figure 3B) was alternatively represented as a CRA plot, where time points and metabolites were mapped on a 2D space to associate metabolic rearrangement with transient correlation network structure (Figure 3C). The centrality profile can be regarded alternatively as a metric of influence. The closeness between a compound symbol and a network symbol represents the degree of influence that the compound has in the network. Distance between given two network symbols represents the similarity of

influences that individual compounds have in these networks. It was understandable that the first 24 time points prior to the glucose pulse remained within a narrow region on Axis 1, because the cells should stay at a steady state prior to the glucose pulse. The plots of post-induction shifted away along the negative direction of Axis 1. Especially, the plots of 25-29th time points exhibited remarkable shifts toward that direction, which corresponded to the time scale of the responsive variance of the adenylate energy charge (Figure 1). It was thus indicated that significant alteration of the cellular energetic state was temporally related with the variances of the metabolite balances. While the metabolic shift was initiated in accordance with a surge in the degree centrality of the number of metabolites, shifts in the early phase were related to variations in the centrality of glutathione (reduced form) (GSH), cytidine monophosphate (CMP), and cytidine diphosphate (CDP). The widespread correlating behavior of GSH is reasonable in terms of the maintenance of redox balance during oxidative respiration, which coincided with the active utilization of the supplied nutrient. It is known that pyrimidine nucleotides are more responsive to the growth phase than purine nucleotides in *E. coli*.[34, 35] Indeed, UDP, UMP, and UDP-glucose showed higher centrality in the earlier phase while purine nucleotides, including ADP, AMP, GDP, and GMP, showed rather delayed variations in their centrality, in response to the glucose pulse. While nucleoside triphosphates themselves do not exhibit any correlative behavior in response to the fluctuation, the synthesis of pyrimidine nucleotides and their phosphorylation levels could be sensitively coupled with the growth conditions. Although alterations in transcription levels were likely to cooperate with metabolite levels, the variations in protein levels following transcription alterations should have little effect on the metabolite-level correlation in this time phase. Considering the time scale of biological events associated with gene and protein expression,[2, 3] the initial variations in the correlation network indicate a passive

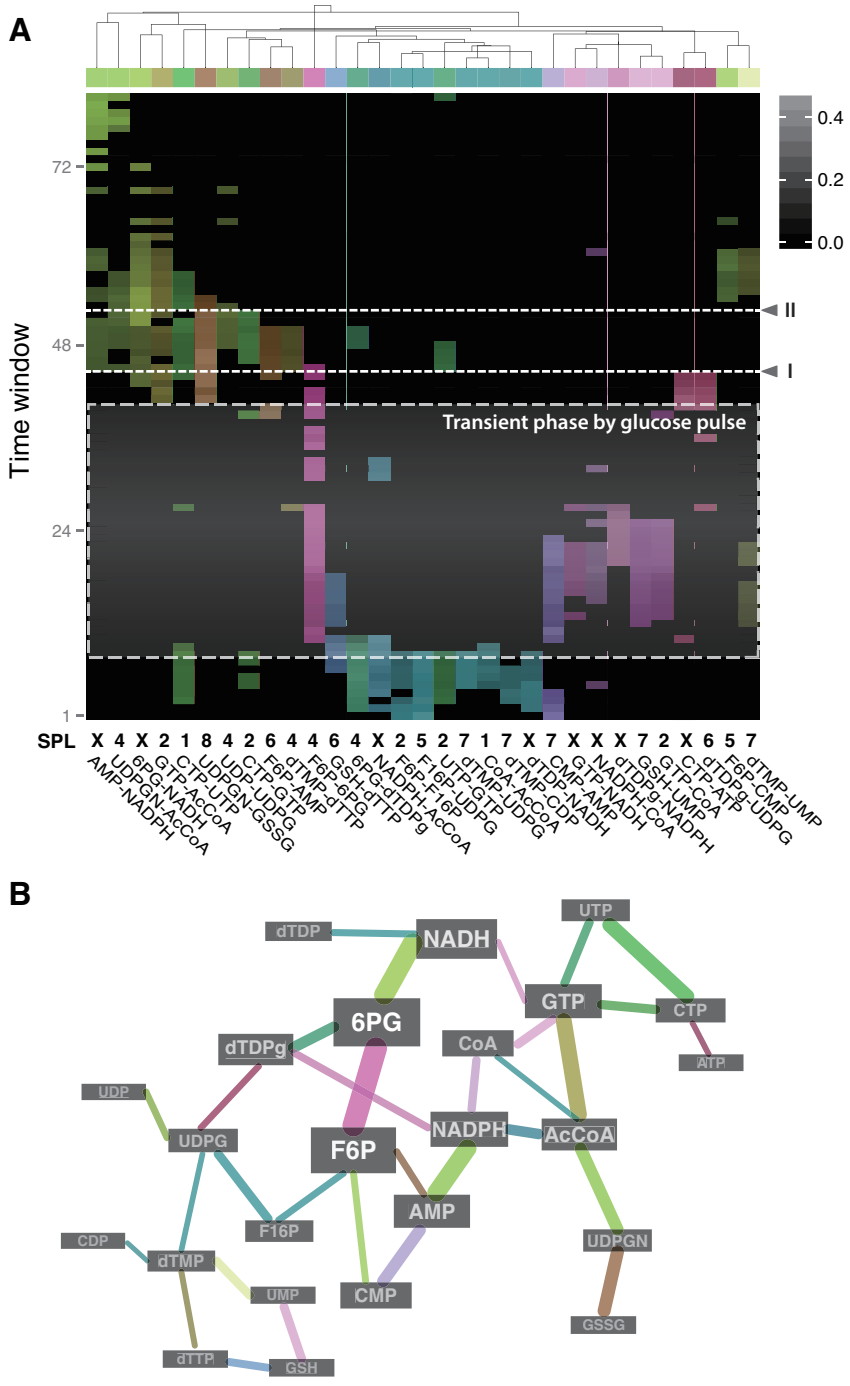
fluctuation in the metabolic system, which was associated dominantly with the metabolites. In this phase, a collapse of metabolic equilibrium could be buffered to prepare a new state of metabolic balance.

3.5. Structural similarity between metabolic pathways and metabolite correlation network

Partial correlation is one of several possibilities for estimating global regulatory interaction structures.[36] When partial correlations are measured, indirect correlations are explicitly excluded, which is preferable to reasonable interpretation of the correlation network. This approach is recommended to reveal the molecular interaction of cellular regulatory networks.[37] We thus constructed GGM networks using time-shifted sequential data, and reconstructed a temporal partial correlation profile. As a result, 31 out of 378 pairs of metabolites were significantly correlated with specific time-range windows (Figure 4A). Obviously, The correlation profile time-dependently altered in response to the glucose pulse, resulting in different profiles compared to those before the nutritional fluctuation. A significant alteration could be found between 44th and 45th time window (Figure 4A, line I), which corresponded to the timing when the adenylate energy charge reached a maximum. Similarly to the single correlation analysis, it was indicated that an alteration of the metabolite network structure could be associated with the temporal transience of the energy state. A moderately remarkable alteration in the temporal correlation profile was also observed at around 52th time window (Figure 4A, line II). Since this window was equivalent to 120-270 sec after the glucose pulse, the following transitions in the profile could be involved with transcriptional regulations. These findings indicated that, as we presumed, a significant metabolic rearrangement was induced in a time scale of seconds in

response to nutritional fluctuation preceding transcriptional alterations, and could be characterized through an evolving metabolite-metabolite network.

To compare the observed correlation network with the biologically reconstructed metabolic network, SPL in a genome-scale metabolic pathway was calculated for each metabolite pair that temporally correlated. If observed correlation is attributed to the 'closeness' in the metabolic pathways, it is expected that average SPL of metabolites in the correlation network is less than that in the genome-scale network. However, the observed SPLs, ranging from 1 to 8, indicated no significant difference from the SPLs between the detected metabolites ($t = 0.327$, $df = 23.228$, p-value = 0.746; Welch's t test), implying that the closeness of metabolites in the metabolic pathway was not necessarily a direct indicator for the correlative variances of the metabolite levels. Also, there were no obvious differences in the SPLs of the correlation profile with regard to the timings of the correlations. Nevertheless, many of the correlations observed before the glucose pulse involved fructose 6-phosphate (F6P), F16P, UDPG, CoA, and acetyl-CoA, and, therefore, could be characterized simply by the glycolytic pathways. An expressive correlation was observed between F6P and 6-phosphogluconate during the transition phase in response to the glucose pulse. Because these metabolites are intermediates of glycolysis and the pentose phosphate pathway respectively, their correlation could be interpreted as the coordinated supply of a carbon source to both pathways. This result agreed with a report that the back-flux from the pentose phosphate pathway scaled with the glycolytic flux.[38]



The whole temporal correlation profile was then reconstructed as a metabolite network to review the evolving structure of the temporal correlation network (Figure 4B). Generally, edges (shown in purple) were involved with immediate variance of the metabolic balance in response to the glucose pulse. Several metabolites had different colored edges, which indicated the temporal

Figure 4. Temporal profile of the partial correlation structure and network representation.

A. Partial correlation structure of pairs of metabolites in specific time-range windows. A partial correlation coefficient for each relationship was calculated using the sectioned time-shifted longitudinal data of metabolite levels. The profiles are visualized as a heat map and clustered by hierarchical clustering using a complete linkage method. A glucose pulse was applied just prior to the 25th time point, and time windows that included the time point (time windows 9–40) are highlighted (dashed line). The color of each profile was determined by mapping the three-dimensional coordination of each relationship in the similarity space of the temporal profile onto a red-green-blue color space. The brightness of the color corresponds the values of the partial correlation coefficients. The numbers along the bottom of the profile indicate the shortest path lengths (SPLs) of the corresponding metabolites in the genome-scale network of *E. coli*. The SPL of X indicates that no distance was evaluated because, in this case, representative cofactores (ATP, ADP, NADH and NADPH) were removed from the metabolic network to prevent the underestimation of the SPL. **B.** Metabolite correlation network based on the temporal correlation profile. The thickness of the edges indicates the time span of the correlation.

pattern of the correlation and confirmed that a single metabolite could participate in more than one correlation at distinct timings. Such correlations with different appearance times, which could be derived from different factors, at least in a temporal context, would have been overlooked in an ordinary correlation analysis.

F6P also showed correlations with various metabolites. Assuming that F6P was the initial indicator of glucose utilization, its correlation partners that were specific to the phase of metabolic shift might provide a simple indicator for the activation of related metabolic pathways. Similarly, GTP showed phase-dependent correlations with distinct metabolites. Because

guanidine phosphates are well-known alarmones,[39] such behaviors in the temporal correlation network might be linked with the stringent response system of *E. coli*. Although further investigations are required, such a structural pattern in temporal correlation network may indicate the relevance of metabolites that are sensitive to metabolic fluctuations, and possibly responsible for the regulatory system dominated by metabolites.

The metabolite pairs of GSH and nucleotides, oxidized glutathione (GSSG) and UDP-*N*-acetylglucosamine (UDPGN) showed significant temporal correlations, which agreed with the result of the single correlation analysis. These pairs showed relatively higher SPL values (≥ 7). It was indicated that the nucleotide metabolism could be coupled with the oxidative state or antioxidant capacity in the cells during the transition phase, despite the relatively distant metabolic pathways between them. RNA is vulnerable to oxidative damages,[40] which could cause disruption of transcriptional or translational fidelity. Observed relations thus might imply that accumulating nucleotides were prevented from oxidative insult induced by an immediate consumption of nutrients. On the other hand, since UDPGN serves as the source for peptidoglycan, the correlation between UDPGN and GSSG could be a counterpart of the GSH-correlated nucleotide metabolism. When the antioxidant capacity decreased, the distribution of carbon source could be directed to the synthesis of the cell wall, not to the nucleotide biosynthesis in excess. The correlations of GSH and nucleotides were observed at the transient phase induced by the glucose pulse, while GSSG and UDPGN correlated at the later phase. It was thus supposed that, in the transition phase, the metabolic system of *E. coli* showed an entirely passive behavior toward the nutritional fluctuations, maintaining the balance of the nucleotide pool and the oxidative capacity. Such a behavior might not be optimal, but just safer with regard to the potential risk of oxidation of newly synthesized nucleotides. Tracking

alterations in metabolite balances could serve as the information for the cellular decision-making toward optimal adaptation, which would be realized by a drastic metabolic rearrangement through transcriptional regulations.

4. Conclusion

This study illustrated a temporal and structural behavior of the metabolite correlation network in a dynamic context under a nutritional perturbation. The evolving network structure provided unique information for interpreting the metabolite correlation that would have been obscured had a static correlation structure alone been evaluated. From a broad perspective, the observed correlations between pyrimidine nucleotides and GSH (or GSSG) might imply a possible program for strategic nutrition utilization by the cells, whereby the biological resources are distributed to invest in further growth and to deal with the risks that usually accompany growth. From a more focused perspective, on the other hand, time-dependent multi-partner behavior in a temporal correlation network may be relevant to biological regulation systems. The temporal profiling of metabolite levels introduced in the present study should be considered as a methodological proposal that can complement metabolomic studies, which are aimed at developing a broader view of quantitative metabolite levels at the price of a limited sample size. Because cellular systems are dynamic, temporal information will also benefit advanced interpretations in other omics studies that deal with multivariate data individually influenced by a variety of factors. Temporal specific patterns of variation in biological networks remain to be investigated, but they should be a rich source of novel perspectives for understanding network dynamics.

5. Authors' contributions

DY and DM conceived the experiments. DY performed the experiments and analyzed the data. DY, DM, and YF wrote the manuscript. DM supervised the research. All authors read and approved the final manuscript.

6. Acknowledgements

This research was supported by the Science and Technology Incubation Program in Advanced Region from the funding program "Creation of Innovation Centers for Advanced Interdisciplinary research Areas" from the Japan Science and Technology Agency, commissioned by the Ministry of Education, Culture, Sports, Science and Technology, and by a Grant-in-Aid for JSPS Fellows to DY. This work was also supported in part by JSPS KAKENHI Grants Number 26713020 for DM.

7. References

1 D. Yukihira, D. Miura, K. Saito, K. Takahashi and H. Wariishi, *Anal. Chem.,* 2010, **82**, 4278-4282.

2 D. Dikicioglu, E. Karabekmez, B. Rash, P. Pir, B. Kirdar and S. G. Oliver, *BMC Syst. Biol.,* 2011, **5**, 148-148 (DOI:10.1186/1752-0509-5-148).

3 M. V. Lee, S. E. Topper, S. L. Hubler, J. Hose, C. D. Wenger, J. J. Coon and A. P. Gasch, *Mol. Syst. Biol.,* 2011, **7**, 514 (DOI:10.1038/msb.2011.48).

4 K. Kochanowski, B. Volkmer, L. Gerosa, B. R. Haverkorn van Rijsewijk, A. Schmidt and M. Heinemann, *Proc. Nat. Acad. Sci. U. S. A.,* 2012, (DOI:10.1073/pnas.1202582110).

5 O. Kotte, J. B. Zaugg and M. Heinemann, *Mol. Syst. Biol.,* 2010, **6**.

6 R. Steuer, *Brief. Bioinform.,* 2006, **7**, 151-158 (DOI:10.1093/bib/bbl009).

7 H. Link, K. Kochanowski and U. Sauer, *Nat. Biotech.,* 2013, **31**, 357-361 (DOI:10.1038/nbt.2489; http://www.nature.com/nbt/journal/v31/n4/abs/nbt.2489.html#supplementary-information).

8 H. Taymaz-Nikerel, M. De Mey, G. Baart, J. Maertens, J. J. Heijnen and W. van Gulik, *Metab. Eng.,* 2013, **16**, 115-129 (DOI:http://dx.doi.org/10.1016/j.ymben.2013.01.004).

9 H. Taymaz-Nikerel, W. M. van Gulik and J. J. Heijnen, *Metab. Eng.,* 2011, **1**3, 307-318 (DOI:http://dx.doi.org/10.1016/j.ymben.2011.03.003).

10 H. Kitano, *Nature,* 2002, **420**, 206-210 (DOI:10.1038/nature01254).

11 A. Wagner and D. A. Fell, *Proc. R. Soc. London, B,* 2001, **268**, 1803-1810 (DOI:10.1098/rspb.2001.1711).

12 W. Weckwerth and O. Fiehn, *Curr. Opin. Biotechnol.,* 2002, **13**, 156-160 (DOI:http://dx.doi.org/10.1016/S0958-1669(02)00299-9).

13 R. Steuer, J. Kurths, O. Fiehn and W. Weckwerth, *Bioinformatics,* 2003, **19**, 1019-1026 (DOI:10.1093/bioinformatics/btg120).

14 A. Fukushima, M. Kusano, H. Redestig, M. Arita and K. Saito, *BMC Syst. Biol.,* 2011, **5**, 1 (DOI:10.1186/1752-0509-5-1).

15 J. Szymanski, S. Jozefczuk, Z. Nikoloski, J. Selbig, V. Nikiforova, G. Catchpole and L. Willmitzer, *PloS one,* 2009, **4**, e7441 (DOI:10.1371/journal.pone.0007441).

16 M. Müller-Linow, W. Weckwerth and M. Hutt, *BMC Syst. Biol.,* 2007, 1, 44.

17 J. D. Orth, T. M. Conrad, J. Na, J. A. Lerman, H. Nam, A. M. Feist and B. O. Palsson, *Mol. Syst. Biol.,* 2011, **7** (DOI:http://www.nature.com/msb/journal/v7/n1/suppinfo/msb201165_S1.html).

18 J. Schellenberger, R. Que, R. M. T. Fleming, I. Thiele, J. D. Orth, A. M. Feist, D. C. Zielinski, A. Bordbar, N. E. Lewis, S. Rahmanian, J. Kang, D. R. Hyduke and B. O. Palsson, *Nat. Protocols,* 2011, **6**, 1290-1307 (DOI:http://www.nature.com/nprot/journal/v6/n9/abs/nprot.2011.308.html#supplementary-information).

19 G. Csárdi and T. Nepusz, *InterJ. Complex Syst.,* 2006, **1695**, 1-9.

20 W. N. Venables and B. D. Ripley, *Modern Applied Statistics with* S, Springer, New York, 2002.

21 S. R. Corman, T. Kuhn, R. D. Mcphee and K. J. Dooley, *Hum. Commun. Res.,* 2002, **28**, 157-206 (DOI:10.1111/j.1468-2958.2002.tb00802.x).

22 U. Brandes and S. R. Corman, *Info. Vis.,* 2002, **2**, 40-50.

23 D. Edwards, *Introduction to graphical modellin*g, Springer-Verlag, Berlin, 2000.

24 J. Schäfer and K. Strimmer, *Bioinformatics,* 2005, **21**, 754-64 (DOI:10.1093/bioinformatics/bti062).

25 J. Schäfer, R. Opgen-Rhein and K. Strimmer., GeneNet: Modeling and Inferring Gene Networks., R package version 1.2.8., 2013.

26 D. Miura, Y. Fujimura, H. Tachibana and H. Wariishi, *Anal. Chem.,* 2010, **82**, 498-504 (DOI:10.1021/ac901083a).

27 D. E. Atkinson, *Biochemistry,* 1968, **7**, 4030-4034.

28 R. Görke, A. Meyer-Bäse, D. Wagner, H. He, M. R. Emmett and C. a. Conrad, *BMC Syst. Biol.,* 2010, **4**, 126 (DOI:10.1186/1752-0509-4-126).

29 C. Chassagnole, N. Noisommit-Rizzi, J. W. Schmid, K. Mauch and M. Reuss, *Biotechnol. Bioeng.,* 2002, **79**, 53-73 (DOI:10.1002/bit.10288).

30 R. Steuer, T. Gross, J. Selbig and B. Blasius, *Proc. Nat. Acad. Sci. U. S. A.,* 2006, **103**, 11868-11873 (DOI:10.1073/pnas.0600013103).

31 L. Mišković and V. Hatzimanikatis, *Biotechnol. Bioeng.,* 2011, **108**, 413-423 (DOI:10.1002/bit.22932).

32 S. P. Borgatti and M. G. Everett, *Soc. Networks,* 1997, **19**, 243-269.

33 J. M. Roberts Jr., *Soc. Networks,* 2000, **22**, 65-72 (DOI:10.1016/S0378-8733(00)00017-4).

34 L. Huzyk and D. J. Clark, *J. Bacteriol.,* 1971, **108**, 74-81.

35 M. H. Buckstein, J. He and H. Rubin, *J. Bacteriol.,* 2008, **190**, 718-726 (DOI:10.1128/JB.01020-07).

36 S. Andorf, J. Selbig, T. Altmann, K. Poos, H. Witucka-Wall and D. Repsilber, *TAG: Theor. Appl. Genet.,* 2010, **120**, 249-259 (DOI:10.1007/s00122-009-1214-z).

37 A. V. Werhli, M. Grzegorczyk and D. Husmeier, *Bioinformatics,* 2006, **22**, 2523-2531 (DOI:10.1093/bioinformatics/btl391).

38 B. Haverkorn van Rijsewijk R.B., A. Nanchen, S. Nallet, R. J. Kleijn and U. Sauer, *Mol. Syst. Biol.,* 2011, **7**, 477.

39 K. Potrykus and M. Cashel, *Annu. Rev. Microbiol.,* 2008, **6**2, 35-51 (DOI:10.1146/annurev.micro.62.081307.162903).

40 J. Barciszewski, M. Z. Barciszewska, G. Siboska, S. I. S. Rattan and B. F. C. Clark, *Mol. Biol. Rep.,* 1999, **2**6, 231-238 (DOI:10.1023/A:1007058602594).