

On the Potential for Integrating Gene Expression and Metabolic Flux Data

M.A. Ovacik¹ and I.P. Androulakis^{*,1,2}

¹Chemical & Biochemical Engineering Department, Rutgers University, NJ 08854; ²Biomedical Engineering Department, Rutgers University, NJ 08854, USA

Abstract: Computational strategies, that integrate genomic-level information and metabolic flux data, have improved both prediction of metabolic fluxes and metabolic network identification. Due to the tight interplay between hierarchical (transcriptional) and metabolic control it is not clear how changes in gene expression drive changes in cellular phenotypes manifested through changes in metabolic fluxes. This raises the questions to what extent a change in a metabolic flux should be attributed to changes in gene expression and/or changes in metabolite concentrations and what kind of conclusions can be drawn by comparing gene expression profiles with the associated metabolic fluxes. This review addresses issues related to modeling approaches that attempt to integrate gene expression and metabolic flux data.

Keywords: Systems biology, gene expression, metabolic flux, modeling.

INTRODUCTION

The cellular phenotype is the decisive manifestation of the state of a living system. As such, it is defined as an emergent property where a number of external signals, internal mechanisms and control structures converge [1]. Thus the phenotype cannot be determined from information generated at a single scale. Rather, information at multiple levels, including gene expression, enzymatic activity, mRNA, protein, and metabolite concentrations, contribute to the metabolic state of the cell. Therefore, the ultimate goal in systems biology is to systematically integrate information from all these different levels and provide an integrated view of the structural and functional organization of living systems [2]. In this complex web of interacting elements the metabolic flux, defined as the rate of conversion of biochemical molecules in a metabolic network [3], has been identified as one of the most critical parameters and along with metabolite concentrations define a basic determinant of cellular physiology [4]. Metabolite changes, by and large, quantify systemic responses and are driving evolving phenotypes. Understanding the mechanisms underlying changes in metabolic fluxes, i.e. relative changes in the rates of biochemical reactions, is a critical enabler towards understanding systemic cellular responses and emerging phenotypes. As such, metabolic engineering has developed as a powerful approach that aims at the introduction of targeted manipulations and modifications of metabolic networks for the purpose of improving cellular functions [5]. Recent studies have indeed demonstrated how the modeling of metabolic pathways can have significant impact towards understanding cellular fate [6].

The central dogma of molecular biology [7] has largely influenced the view of how information is transferred from the genome to the phenotype. The brilliance in its statement is that within its simplicity hides the underlying complexity of cellular physiology. Beneath the apparent "linear" flow of

information from the DNA, to the mRNA, to protein, to enzymatic activity and eventually to the changes in biochemical rates, lies intricate regulation machinery that creates an amazing web of feedback interaction mechanisms, Fig. (1). Complexities in the regulation of cellular metabolism include DNA-protein interactions [8] and post-transcriptional regulation of mRNA [9, 10]. Despite such regulation complexities, combination of different types of information across various components of the overall machinery, such as between proteome and enzyme activity [11, 12], transcriptome and proteome [13-15], proteome and metabolic fluxes [16], have shown that a critical level of coordination is achieved and can be rationalized. In this complex web of interacting components, the metabolic flux analysis and the expression of genes can be thought of as the two end points of regulation of cellular metabolism. Yet, it has been shown that metabolic genes are differentially expressed under different environmental and intrinsic conditions such as nutrient availability or enzyme knock-outs [17, 18] and metabolite concentrations can have an effect on the activities of transcription factors, proteins that regulate gene expression [19]. These studies suggest that metabolic reactions and gene expression are, potentially, indirectly regulated by each other [20]. Of particular importance is the possibility of elucidating metabolic enzyme changes in the context of the corresponding metabolic networks [21-24].

In this review, we will attempt to explore whether and how information at these two important levels, gene expression of metabolic enzymes and metabolic fluxes, can be concurrently rationalized, and we will explore the potential for integrating these sources of information with the ultimate goal of understanding the relationship, if any, between changes in expression of metabolic genes and changes in metabolic fluxes. We introduce briefly how genomic and metabolic information are integrated to improve both metabolic flux prediction and metabolic network identification. It is important to realize that the identification of direct relations between the two is an elusive target and it is highly likely that a positive answer will never be obtained, simply because a straightforward correlation between the two does not exist [25]. Nevertheless, it is important to explore such a

*Address correspondence to this author at the Biomedical Engineering Department, Rutgers University, NJ 08854, USA;
E-mail: yannis@rci.rutgers.edu

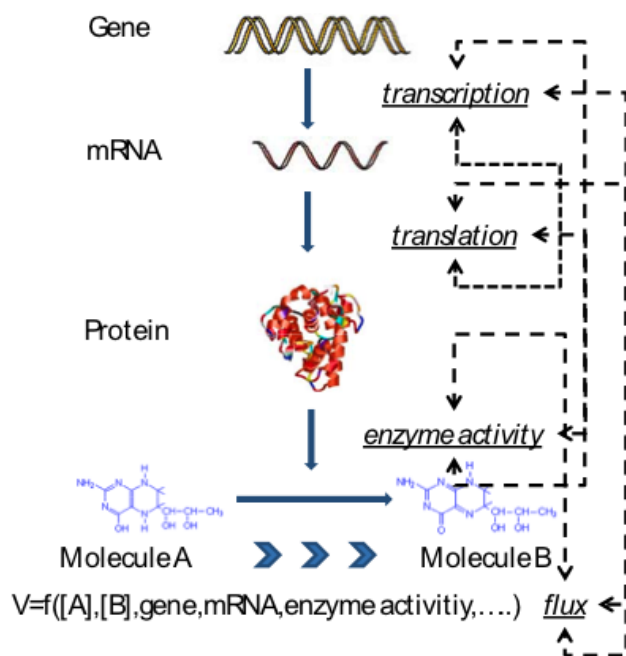


Fig. (1). The information of DNA is deciphered into mRNA levels, and subsequently transcribed to proteins. Protein-metabolite interactions regulate metabolic fluxes as well as the metabolite concentrations. The interplay between different stages of cellular metabolism is unknown.

possibility in order to eventually develop an integrative framework of the multiple sources of *-omics* information [26].

THE USE OF METABOLIC NETWORK MODELS TO INTEGRATE GENE EXPRESSION AND CELLULAR PHENOTYPE ANALYSIS

High-throughput genotyping technologies that generate genomic sequences and genome-wide transcription analyses raise expectations of high-throughput phenotyping assessments [27]. In principle, the network of metabolic reactions in an organism can be reconstructed from the annotated genome sequence, biochemical data and the experimentally determined characteristics of cell physiology [28]. After mapping the known metabolic genes on to the metabolic network, biochemical data is incorporated to assign relevant biochemical reactions to the gene products and to determine the existence of reactions or pathways not specified by genomic data. Knowledge of the physiology of a cell gives indirect verification of the existence or non-existence of certain metabolic reactions. The final reconstructed network models, i.e. for *Helicobacter pylori* 26695 and *Saccharomyces cerevisiae* [29, 30], can be analyzed to assess whether and how gene expression affects metabolic fluxes.

One of the methods that have been proposed to evaluate metabolic fluxes is to construct kinetic models, which are defined by metabolite and enzyme concentrations, and the system parameters such as rate constants. Downer *et al.* used data from mRNA expression levels of lactate dehydrogenase (LDH) isoforms, assuming that gene expression changes can predict changes in enzyme concentration, and the kinetic information to evaluate the fluxes of lactate production from

pyruvate in *erythroleukemia* cell lines [31]. The drawback of kinetic modeling approaches is that a detailed description of kinetic parameters and the complete set of biochemical reactions are required and are not always available. Therefore, Downer *et al.* performed robustness analysis to evaluate the effects of unknown parameters on the kinetic model [31]. Large scale phenotype prediction, on the other hand, establishes a combination of mechanistic and heuristic information, when the complete mechanistic detail is not available for kinetic modeling. The heuristic part of the model is derived based on the assumption that the regulatory program of an organism can be formulated as different rules, i.e., metabolic functions, that are developed from experimental observations [32]. These are then appropriately invoked in response to a stimulus [32, 33]. As such, Varner *et al.* were able to predict the shifts of metabolic fluxes based on the mRNA level changes in a given set of reactions in *E.coli* [34], yet only partially.

Due to large number of components and interactions involved in biological systems, it is unrealistic to expect to have a complete description of a cellular system. Therefore, modeling approaches are required to reduce the dimensionality to identify key components of cellular processes. *Constraint based models* [35] rely on certain fundamental physical and chemical constraints such as mass balance and thermodynamics. One can never of course expect to be able to predict exactly how the cells behave with a limited set of parameters; however the possibilities can be decreased so that we can analyze the capacity of the cell to perform a certain function under well defined conditions.

An interesting reconstruction problem arises due to the fact that, because of the limited amount of available information, there is potentially large number of valid flux distributions. Therefore, the problem is solved for a flux distribution that optimizes a particular cellular function, such as to maximize growth rate, which is usually expressed as a ratio of particular metabolites. This approach is known as flux balance analysis (FBA), which analyzes the performance of a framework to determine optimal cellular behaviors under diverse environmental and genetic conditions [27, 36]. Edwards *et al.* were predicted the growth rate of *E.coli* using acetate and succinate as single-carbon sources [37] via FBA. However, FBA lacks the regulatory information and consequently, deviations between models and experimental observations are inevitable when regulatory effects have major impact on the cell physiology [38]. Therefore, the use of gene expression measurements combined with constraint based modeling lessens the gap between gene regulation and metabolic processes. Akesson *et al.* [39] demonstrated how incorporating genome-wide measurements of transcription activity into FBA further improved metabolic behavior predictions in yeast.

Covert *et al.* [40] integrated gene expression data into metabolic network reconstructions by iteratively constructing the network model based on FBA and experimental results to identify new components and interactions in microbial biological networks. Yet, constraint based models can identify neither transcriptional regulation nor the key regulatory points of the systems but they can predict the effects of internal and external disturbances in the system.

INTEGRATION OF GENE EXPRESSION DATA WITH METABOLIC FLUX DATA

The phenotype of the cell is defined by the metabolic fluxes, whose complete characterization can be obtained via *metabolic flux analysis* (MFA) [41]. The complete characterization of the two end points, metabolic flux and gene expression, can be used to identify key regulatory points. However, integration of gene expression and metabolic flux data has been critically questioned because of the disparity between mRNA abundance and enzyme activity [42]. This discrepancy arises from a number of possibilities including posttranscriptional regulation of protein synthesis, protein regulation by posttranslational modification, and possible functional requirements for protein binding [42]. In addition to the discrepancy between mRNA abundance and enzyme activity, the timescales of various biological processes being unknown or very different complicates the studies that integrate gene expression data and metabolic flux data [43]. Despite concerns, a number of studies, which integrate gene expression and metabolic flux data, have indeed provided useful insights in terms of understanding relations between genotype, environment and phenotype. This section provides the synthesis of these studies. Specifically, how to analyze the results of the integration studies and their part in understanding the regulatory state of the cell are investigated.

Gene expression and their associated metabolic flux changes can be compared in qualitatively and quantitatively ways at the individual reaction level [44-48]. Interesting computational issues arise often times due to the required normalization of both gene expression and metabolic flux data. Among the simplest normalization methods is the evaluation of ratios of fluxes and transcript levels across different experiment conditions; [44, 45, 49] such that the changes can be compared based on ratios of fold change. However, metabolic fluxes can be shut down under some experimental conditions, and as a result some metabolic fluxes may be zero. In this case, the ratio of fluxes cannot be evaluated. Lequex *et al.* [47] and Wong *et al.* [25] calculated the mean of the fluxes at each experimental condition and this mean value was used as the reference state. Afterwards, each flux was divided by the reference state. Subsequently, the ratio of a given metabolic flux in a given experimental condition against another experimental condition was evaluated. The same normalization framework was applied for gene expression levels in order that a normalized quantity was calculated for both datasets. Evaluating ratios was used generally in studies where replicates are not available. Where replicates are available, Banta *et al.* [46] assigned certain values, such as -1, 0, and 1 that correspond significant decrease, no significant change and significant increase, respectively. Significance or non-significance were determined by statistical tests, such as t-test and ANOVA [25, 46, 47]. Consequently, one can compare the changes in gene expression and metabolic fluxes based on the assigned values.

The comparison between gene expression and metabolic flux data at the individual reaction level may be performed qualitatively by evaluating the trends in the gene expression changes and associated metabolic flux changes as increasing or decreasing [44, 48-50]. Oh *et al.* [44] and Yang *et al.* [45], on the other hand, matched the ratios of gene expression and metabolic fluxes across different experiment conditions.

However, qualitative analysis provides only general characterization; therefore quantitative analysis would be more useful to describe the precise level of correlation between gene expression and metabolic flux data. Wong *et al.* [25] and Lequex *et al.* [47] computed the Pearson correlation coefficient that provides a quantifiable metric of comparison between normalized gene expression and metabolic fluxes as positive correlation, negative correlation and no correlation. The reactions were then categorized based on the correlation between gene expression and metabolic flux [25, 47]. Moreover, Banta *et al.* [46] applied the Mann-Whitney test to compare the distributions of flux and gene expression changes across the network.

A critical issue associated with the analysis of gene expression and metabolic flux at the individual reaction level is the presence of isoenzymes that catalyze the same reaction [51]. Isoenzymes may catalyze reverse reactions, catalyze reactions mutually or be present at different cellular compartments. In other words, there is more than one gene products that catalyze the same reaction and consequently there are more than one gene expressions. Isoenzymes are critical for the function of the cell and have been associated with robustness through redundancy [52]. Cakir *et al.* [51] handled the issue via summing the transcript levels for all genes coding for the same reaction, however this is not an efficient way to consider the presence of isoenzymes, because, if the isoenzymes catalyze reverse reactions and gene expression contributes to the metabolic flux changes, most likely, these gene expressions should show opposite changes and summing the transcript level would eliminate the effect of individual isoenzymes. Yet, the multiple gene expressions are still compared to the same flux changes by the majority of the studies [44, 46, 47, 49]. In the presence of isoenzymes, it is not proficient to compare gene expression changes with metabolic flux changes unless functional information of the isoenzymes is integrated in the comparison analysis.

When the fluxes of individual reactions cannot be determined independently because of observability issues it is suggested to lump reaction series together [4, 53]. In addition, the metabolic networks can be constructed in a way that best reflects the major metabolic pathways and for simplicity reasons few reactions may be integrated into one reaction, for instance in [46] the reaction series from phosphoenolpyruvate to glyceraldehyde-3-P was characterized as one reaction, however the series includes at least four different reactions, that are defined in KEGG database [54]. Therefore, four different gene expression levels corresponded to only one metabolic flux. In order to handle this type of situations, summing the transcript levels for all genes coding for the same reaction [51] would give an insight about how the gene expressions of the reaction series would affect the corresponding flux of the reaction series unless there are any isoenzymes.

A simultaneous characterization of gene expression and metabolic flux data was performed by Kromer *et al.* [50] at consecutive time intervals; flux measurements were sampled following gene expression measurements. This work was testing the hypothesis that transcriptional events precede metabolic changes and as such better correlation could be obtained if the measurements were not in tandem. Even though sensitive time sampling may improve the correlation

results between gene expression data and metabolic flux data where the enzymatic activity is correlated with gene expression, in general the time scale between transcription and enzymatic activity is not known [43].

WHAT DOES the integration of gene expression and metabolic flux changes HAVE to offer?

The comparison of gene expression and flux changes, at the individual reaction level, may in general suffer from the lack of any type of correlation between measurements [39, 44-49]. Although there are some reactions that exhibit correlation between gene expression and metabolic flux changes, it is still questionable whether gene expression changes can predict the associated metabolic flux changes. Despite the fact that generalizable correlations between gene expression and metabolic fluxes cannot be easily identified, hypotheses on the regulatory mechanisms can be still formulated based on the comparison analyses, Fig. (2).

Significant transcriptional changes are not always reflected in the corresponding metabolic flux changes. Banta *et al.* [46] reported that changes in gene expression of metabolic enzymes in the urea cycle were not accompanied by related changes in urea synthesis following severe injury in rats. Similarly, Schilling *et al.* [48] demonstrated that in yeast the most significant changes in gene expressions (10-fold) were not accompanied by changes of the associated fluxes. The reason for this type of discrepancy is that mRNA abundance is not a good proxy of enzymatic activity, because it depends on post-transcriptional regulation [55], i.e. phosphorylation, and is affected by post transcriptional mechanisms. However, Oh *et al.* [44] suggested that changes in transcript level can potentially reveal the existence of significant regulation that can be used to develop strategies for genetic modifications.

Interesting issues arise when significant changes in metabolic fluxes are not accompanied by changes in the expression of metabolic enzymes. Evidence suggests that for highly expressed genes there is a relationship between mRNA and protein levels [56]. Therefore for metabolic genes, one could argue that we should not expect significant correlations between fluxes and gene expression, unless the gene expression measurements indicate significant up- or-down regulation [49]. However, Yang *et al.* [45] evaluated protein level changes along with the gene expression and metabolic flux changes in *Synechocystis* and showed that even if the transcript ratio and the protein ratio agreed, the flux ratio change may be significantly different from the other two. Tummala

et al. [49] stated that for the reactions exhibiting significant metabolic flux change associated with non-significant changes in gene expression, the relationship between enzyme activity and flux, such as allosteric changes or limitation of some metabolite concentrations, may be the determining factor on the significant change of the metabolic flux. Similarly, Banta *et al.* [46] related negative correlation between flux and gene expression to substrate availability, which would cause fluxes to go down, even when the genetic regulation mechanism was up regulated. Lequex *et al.* [47], on the other hand, associated the negative correlation between gene expression and metabolic flux as a compensation mechanism. In carbon limited culture, ptsG, which is the enzyme for glucose conversion to glucose 6-phosphate, was expressed more than in carbon abundant cultures in *E.coli*. This way the cell might aim to maximize the carbon uptake under carbon limited environment by expressing more ptsG.

Since the flux through a pathway equals the sum of the rate of the reaction series, it is common to define a combined flux of a pathway such as TCA cycle flux and glycolytic flux. These combined fluxes and gene expression can be further analyzed [25, 50, 57]. Kromer *et al.* investigated the different stages in *Corynebacterium glutamicum* cultures, such as growth and lysine production phases and concluded that for the most part gene expression changes in TCA cycle were correlated with TCA cycle flux, but most of the genes in the lysine biosynthetic pathway were unaffected in *E.coli* [50]. Wong *et al.* [25] stated that in mouse hepatoma cells, TCA and glycolytic flux changes due to glutamine depletion were not accompanied by genes related in these pathways.

It is probably fair to say that the relationship between gene expression and metabolic flux changes is an obscure feature, because any type of correlation between the two cannot be identified in general terms. However, gene expression changes in different functional groups can be compared to combined fluxes to understand the functional group contribution to the overall cell function. Gonzalez *et al.* and Tao *et al.* compared gene expression changes in the Pentose-Phosphate (PPP) and Embden-Meyerhof-Parnas (EMP) pathways to glycolytic flux [57, 58], defined by the conversion rate of glucose, and reported that changes in gene expression in EMP pathway were in agreement with the glycolytic flux whereas there were no significant changes in gene expression in PPP [57]. These results suggested that the control of the glycolytic flux was distributed within EMP whereas it was limited by PPP.

REGULATION ANALYSIS

It should be clear by now that due to tight interplay between transcriptional and metabolic control it is not obvious how the changes in gene expression influence changes in metabolic fluxes [59]. Ter Kuile *et al.* introduced *regulation analysis* as a means to quantitatively determine to what extent a change in a metabolic flux could be attributed to changes in the expression of metabolic genes and to what extent it should be attributed to changes in metabolite levels [60, 61]. *Regulation analysis* defines the rate of an enzymatic reaction as a function of substrates, products, modifiers and gene expression, which is the basic determinant of enzyme concentration. Consequently, hierarchical and meta-

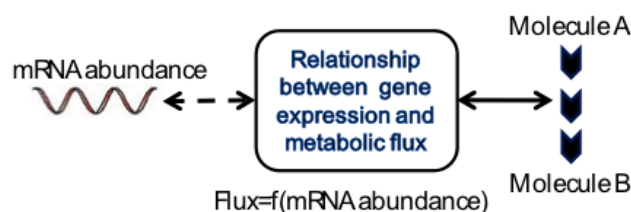


Fig. (2). The metabolic flux analysis and the expression of genes can be thought of as the two end points of regulation of cellular metabolism. The relationship between gene expression and metabolic flux cannot be generalized; however characterization of the relationship may help to generate new hypotheses.

bolic regulation coefficients of a metabolic flux change are evaluated as all changes caused by substrates, products, modifiers and changes caused by enzyme concentration alterations via translation or degradation, respectively. Cakir *et al.* [51] considered the reactions as metabolically regulated if there were coordinated metabolite level changes in the substrates and the products of a given reaction. Similarly, the reactions where there was significant gene expression changes were referred as hierarchically regulated. Subsequently, a categorization was performed of the reactions in terms hierarchically and metabolically regulation. As a result, in industrial yeast strain metabolic regulation was determined as predominant for secretion reactions and amino acid pathways, whereas in laboratory yeast strain the hierarchical regulation and metabolic regulation was equally dominated [51]. The less important degree of hierarchical regulation in the industrial strain may help the cells by decreasing the use of resources for the transcriptional regulatory mechanism.

BLACK BOX MODELING OF GENE EXPRESSION AND METABOLIC FLUX DATA

It is reasonable to anticipate that gene expression changes should affect metabolic flux changes, but the question is how to capture this relationship quantitatively. Since the general conclusion is that gene expression cannot predict changes in metabolic fluxes, the aforementioned analyses attempt to rationalize the regulatory mechanisms of an organism. Li *et al.* [62] hypothesize that gene expression changes can, in principle, be used to predict the cellular function. The framework includes an optimization algorithm (Genetic Algorithm, GA) coupled with partial least squares (PLS) analysis framework to select subset of genes. As a result, an optimally selected set of genes is identified that can be used to predict metabolic fluxes which are treated as the response variables of the system. Following the basic principles of model selection and building, appropriate training and test sets based on the data are constructed to test the accuracy of the model prediction. Li *et al.* [62] used rat hepatocytes cultivated in different insulin conditions and the GA/PLS model was applied to urea and triglyceride(TGL) synthesis. The framework showed that there was a subset of genes whose expression levels could predict the metabolic functions and subsequently pathways involved in urea and triglyceride synthesis were reconstructed with the selected subset of genes [62]. Interestingly, the framework was able identify transcription factors that were known to regulate TGL and urea synthesis. The framework did not incorporate translational and post-translational effects in the gene selection processes, rather the latent relationship between gene expression and metabolic fluxes was considered. Integration of such a disparate data, i.e. metabolite level and gene expression [63] or cytotoxicity and gene expression[64], revealed underlying pathways and mechanisms of cadmium toxicity in rat hepatocytes thus potentially leading to the formulation of control hypotheses.

CONCLUSION

Sauer *et al.* stated that the reductionist approach has successfully defined many concepts in biology yet fails to propose any generalizable understanding to how systemic properties emerge[65]. The variety of reasons and outcomes in

biological systems is better understood by monitoring multiple events simultaneously and by systematic data integration with theoretical models[65]. Therefore systems biology has evolved towards the direction of integration rather than reduction.

It is well established that the regulation of the cellular physiology is fairly complex and every level of the regulatory mechanism contributes to the interplay between the two endpoints of cellular metabolism, gene expression and metabolic phenotypes. However, it is unlikely to identify the contributions of each level generically and precisely. Rather, integration of the different levels of information is required before it is possible to develop any type of descriptive, let alone predictive, relationships between genotype, environment and phenotype.

As we understand more of the cellular interactions, we realize how complex the entire picture is. Despite such complexities, integration of gene expression and metabolic flux data can potentially offer the possibility of generating new hypotheses regarding how cells reach their final state or dynamically respond to internal and/or external signals [66].

ACKNOWLEDGEMENTS

The authors acknowledge financial support from the EPA under EPA GAD R 832721-010.

REFERENCES

- [1] Hellerstein MK. In vivo measurement of fluxes through metabolic pathways: the missing link in functional genomics and pharmaceutical research. *Annu Rev Nutr* **2003**; 23: 379-402.
- [2] Zhang S, Ye B, Chu J, Zhuang Y, Guo M. From multi-scale methodology to systems biology: to integrate strain improvement and fermentation optimization. *J Chem Technol Biotechnol* **2006**; 81: 734-745.
- [3] Stephanopoulos G, Aristidou AA, Nielsen JH, Metabolic engineering : principles and methodologies. Academic Press, San Diego, 1998
- [4] Stephanopoulos G. Metabolic fluxes and metabolic engineering. *Metab Eng* **1999**; 1: 1-11.
- [5] Yarmush ML, Berthiaume F. Metabolic engineering and human disease. *Nat Biotechnol* **1997**; 15: 525-8.
- [6] Yarmush ML, Banta S. Metabolic engineering: advances in modeling and intervention in health and disease. *Annu Rev Biomed Eng* **2003**; 5: 349-81.
- [7] Crick F. Central dogma of molecular biology. *Nature* **1970**; 227: 561-3.
- [8] Vemuri GN, Aristidou AA. Metabolic engineering in the -omics era: elucidating and modulating regulatory networks. *Microbiol Mol Biol Rev* **2005**; 69: 197-216.
- [9] Vanderpool CK, Gottesman S. Involvement of a novel transcriptional activator and small RNA in post-transcriptional regulation of the glucose phosphoenolpyruvate phosphotransferase system. *Mol Microbiol* **2004**; 54: 1076-89.
- [10] Bernstein JA, Lin PH, Cohen SN, Lin-Chao S. Global analysis of Escherichia coli RNA degradosome function using DNA microarrays. *Proc Natl Acad Sci U S A* **2004**; 101: 2758-63.
- [11] Peng L, Shimizu K. Global metabolic regulation analysis for Escherichia coli K12 based on protein expression by 2-dimensional electrophoresis and enzyme activity measurement. *Appl Microbiol Biotechnol* **2003**; 61: 163-78.
- [12] Wang W, Sun J, Hartlep M, Deckwer WD, Zeng AP. Combined use of proteomic analysis and enzyme activity assays for metabolic pathway analysis of glycerol fermentation by Klebsiella pneumoniae. *Biotechnol Bioeng* **2003**; 83: 525-36.
- [13] Yoon SH, Han MJ, Lee SY, Jeong KJ, Yoo JS. Combined transcriptome and proteome analysis of Escherichia coli during high cell density culture. *Biotechnol Bioeng* **2003**; 81: 753-67.

- [14] Lee JH, Lee DE, Lee BU, Kim HS. Global analyses of transcriptomes and proteomes of a parent strain and an L-threonine-overproducing mutant strain. *J Bacteriol* **2003**; 185: 5442-51.
- [15] Mader U, Homuth G, Scharf C, Buttner K, Bode R, Hecker M. Transcriptome and proteome analysis of *Bacillus subtilis* gene expression modulated by amino acid availability. *J Bacteriol* **2002**; 184: 4288-95.
- [16] Han MJ, Yoon SS, Lee SY. Proteome analysis of metabolically engineered *Escherichia coli* producing Poly(3-hydroxybutyrate). *J Bacteriol* **2001**; 183: 301-8.
- [17] Siddiquee KA, Arauzo-Bravo MJ, Shimizu K. Effect of a pyruvate kinase (pykF-gene) knockout mutation on the control of gene expression and metabolic fluxes in *Escherichia coli*. *FEMS Microbiol Lett* **2004**; 235: 25-33.
- [18] Oh MK, Rohlin L, Kao KC, Liao JC. Global expression profiling of acetate-grown *Escherichia coli*. *J Biol Chem* **2002**; 277: 13175-83.
- [19] Natarajan K, Meyer MR, Jackson BM, Slade D, Roberts C, Hinnebusch AG, Marton MJ. Transcriptional profiling shows that Gcn4p is a master regulator of gene expression during amino acid starvation in yeast. *Mol Cell Biol* **2001**; 21: 4347-68.
- [20] Yeang CH, Vingron M. A joint model of regulatory and metabolic networks. *BMC Bioinformatics* **2006**; 7: 332.
- [21] Shlomi T, Herrgard M, Portnoy V, Naim E, Palsson BO, Sharan R, Ruppin E. Systematic condition-dependent annotation of metabolic genes. *Genome Res* **2007**; 17: 1626-33.
- [22] Kharchenko P, Chen L, Freund Y, Vitkup D, Church GM. Identifying metabolic enzymes with multiple types of association evidence. *BMC Bioinformatics* **2006**; 7: 177.
- [23] Kharchenko P, Vitkup D, Church GM. Filling gaps in a metabolic network using expression information. *Bioinformatics* **2004**; 20 Suppl 1: i178-85.
- [24] Segre D, Zucker J, Katz J, Lin X, D'Haeseleer P, Rindone WP, Kharchenko P, Nguyen DH, Wright MA, Church GM. From annotated genomes to metabolic flux models and kinetic parameter fitting. *Omics* **2003**; 7: 301-16.
- [25] Wong MS, Raab RM, Rigoutsos I, Stephanopoulos GN, Kelleher JK. Metabolic and transcriptional patterns accompanying glutamine depletion and repletion in mouse hepatoma cells: a model for physiological regulatory networks. *Physiol Genomics* **2004**; 16: 247-55.
- [26] Morel NM, Holland JM, van der Greef J, Marple EW, Clish C, Loscalzo J, Naylor S. Primer on medical genomics. Part XIV: Introduction to systems biology--a new approach to understanding disease and treatment. *Mayo Clin Proc* **2004**; 79: 651-8.
- [27] Covert MW, Schilling CH, Famili I, Edwards JS, Goryanin II, Selkov E, Palsson BO. Metabolic modeling of microbial strains in silico. *Trends Biochem Sci* **2001**; 26: 179-86.
- [28] Reed JL, Palsson BO. Thirteen years of building constraint-based in silico models of *Escherichia coli*. *J Bacteriol* **2003**; 185: 2692-9.
- [29] Forster J, Famili I, Fu P, Palsson BO, Nielsen J. Genome-scale reconstruction of the *Saccharomyces cerevisiae* metabolic network. *Genome Res* **2003**; 13: 244-53.
- [30] Schilling CH, Covert MW, Famili I, Church GM, Edwards JS, Palsson BO. Genome-scale metabolic model of *Helicobacter pylori* 26695. *J Bacteriol* **2002**; 184: 4582-93.
- [31] Downer J, Sevinsky JR, Ahn NG, Resing KA, Betterton MD. Incorporating expression data in metabolic modeling: a case study of lactate dehydrogenase. *J Theor Biol* **2006**; 240: 464-74.
- [32] Varner J, Ramkrishna D. Metabolic engineering from a cybernetic perspective. 1. Theoretical preliminaries. *Biotechnol Prog* **1999**; 15: 407-25.
- [33] Varner J, Ramkrishna D. Metabolic engineering from a cybernetic perspective. 2. Qualitative investigation of nodal architectures and their response to genetic perturbation. *Biotechnol Prog* **1999**; 15: 426-38.
- [34] Varner JD. Large-scale prediction of phenotype: concept. *Biotechnol Bioeng*, **2000**; 69: 664-78.
- [35] Palsson B. The challenges of in silico biology. *Nat Biotechnol* **2000**; 18: 1147-50.
- [36] Papin AP, Price ND, Palsson B. In Silico Cells: Studying Genotype-Phenotype Relationships Using Constraint-based Models, in *Metabolic Engineering in Post Genomic Era*. Horizon Bioscience, Norfolk. 2004; p. 209-236.
- [37] Edwards JS, Ibarra RU, Palsson BO. In silico predictions of *Escherichia coli* metabolic capabilities are consistent with experimental data. *Nat Biotechnol* **2001**; 19: 125-30.
- [38] Covert MW, Schilling CH, Palsson B. Regulation of gene expression in flux balance models of metabolism. *J Theor Biol* **2001**; 213: 73-88.
- [39] Akesson M, Forster J, Nielsen J. Integration of gene expression data into genome-scale metabolic models. *Metab Eng*, **2004**; 6: 285-93.
- [40] Covert MW, Knight EM, Reed JL, Herrgard MJ, Palsson BO. Integrating high-throughput and computational data elucidates bacterial networks. *Nature* **2004**; 429: 92-6.
- [41] Stephanopoulos G, Aristidou AA, Nielsen J. *Metabolic engineering: principles and methodologies*. Academic Press, San Diego, 1998
- [42] Park SJ, Lee SY, Cho J, Kim TY, Lee JW, Park JH, Han MJ. Global physiological understanding and metabolic engineering of microorganisms based on omics studies. *Appl Microbiol Biotechnol* **2005**; 68: 567-79.
- [43] Nicholson JK, Holmes E, Lindon JC, Wilson ID. The challenges of modeling mammalian biocomplexity. *Nat Biotechnol* **2004**; 22: 1268-74.
- [44] Oh MK, Liao JC. Gene expression profiling by DNA microarrays and metabolic fluxes in *Escherichia coli*. *Biotechnol Prog* **2000**; 16: 278-86.
- [45] Yang C, Hua Q, Shimizu K. Integration of the information from gene expression and metabolic fluxes for the analysis of the regulatory mechanisms in *Synechocystis*. *Appl Microbiol Biotechnol* **2002**; 58: 813-22.
- [46] Banta S, Vemula M, Yokoyama T, Jayaraman A, Berthiaume F, Yarmush ML. Contribution of gene expression to metabolic fluxes in hypermetabolic livers induced through burn injury and cecal ligation and puncture in rats. *Biotechnol Bioeng* **2007**; 97: 118-137.
- [47] Lequeux G, Johansson L, Maertens J, Vanrolleghem PA, Liden G. MFA for overdetermined systems reviewed and compared with RNA expression data to elucidate the difference in shikimate yield between carbon- and phosphate-limited continuous cultures of *E. coli* W3110.shik1. *Biotechnol Prog* **2006**; 22: 1056-1070.
- [48] Schilling O, Frick O, Herzberg C, Ehrenreich A, Heinzle E, Wittmann C, Stulke J. Transcriptional and metabolic responses of *Bacillus subtilis* to the availability of organic acids: Transcription regulation is important but not sufficient to account for metabolic adaptation. *Appl Environ Microbiol* **2007**; 73: 499-507.
- [49] Tummala SB, Junne SG, Paredes CJ, Papoutsakis ET. Transcriptional analysis of product-concentration driven changes in cellular programs of recombinant *Clostridium acetobutylicum* strains. *Biotechnol Bioeng* **2003**; 84: 842-54.
- [50] Kromer JO, Sorgenfrei O, Klopprogge K, Heinzle E, Wittmann C. In-depth profiling of lysine-producing *Corynebacterium glutamicum* by combined analysis of the transcriptome, metabolome, and fluxome. *J Bacteriol* **2004**; 186: 1769-84.
- [51] Cakir T, Patil KR, Onsan Z, Ulgen KO, Kirdar B, Nielsen J. Integration of metabolome data with metabolic networks reveals reporter reactions. *Mol Syst Biol* **2006**; 2: 50.
- [52] Bailey JE. Lessons from metabolic engineering for functional genomics and drug discovery. *Nat Biotechnol* **1999**; 17: 616-8.
- [53] dos Santos MM, Akesson M, Nielsen J. Metabolic Flux Analysis in the Post Genomic Era, in *Metabolic Engineering in Post Genomic Era*. Horizon Bioscience, Norfolk. 2004; p. 90-105.
- [54] Kanehisa M, Goto S. KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res* **2000**; 28: 27-30.
- [55] Levine A, Vannier F, Absalon C, Kuhn L, Jackson P, Scrivener E, Labas V, Vinh J, Courtney P, Garin J, Seror SJ. Analysis of the dynamic *Bacillus subtilis* Ser/Thr/Tyr phosphoproteome implicated in a wide variety of cellular processes. *Proteomics* **2006**; 6: 2157-73.
- [56] Ideker T, Thorsson V, Ranish JA, Christmas R, Buhler J, Eng JK, Bumgarner R, Goodlett DR, Aebersold R, Hood L. Integrated genomic and proteomic analyses of a systematically perturbed metabolic network. *Science* **2001**; 292: 929-34.
- [57] Gonzalez R, Tao H, Shanmugam KT, York SW, Ingram LO. Global gene expression differences associated with changes in glycolytic flux and growth rate in *Escherichia coli* during the fermentation of glucose and xylose. *Biotechnol Prog* **2002**; 18: 6-20.
- [58] Tao H, Gonzalez R, Martinez A, Rodriguez M, Ingram LO, Preston JF, Shanmugam KT. Engineering a homo-ethanol pathway in *Escherichia coli*: increased glycolytic flux and levels of expression of

- glycolytic genes during xylose fermentation. *J Bacteriol* **2001**; 183: 2979-88.
- [59] Nielsen J. It is all about metabolic fluxes. *J Bacteriol* **2003**; 185: 7031-5.
- [60] ter Kuile BH, Westerhoff HV. Transcriptome meets metabolome: hierarchical and metabolic regulation of the glycolytic pathway. *FEBS Lett* **2001**; 500: 169-171.
- [61] Westerhoff HV, ter Kuile BH, Reijenga KA, Snoep JL, Kholodenko BN, A new millenium: functional genomics, cellular bioinformatics, regulation analysis, in *Animating the Cellular Map*. Stellenbosch University Press, South Africa. 2000; p. 1-7.
- [62] Li Z, Chan C. Integrating gene expression and metabolic profiles. *J Biol Chem* **2004**; 279: 27124-37.
- [63] Li Z, Srivastava S, Yang X, Mittal S, Norton P, Resau J, Haab B, Chan C. A hierarchical approach employing metabolic and gene expression profiles to identify the pathways that confer cytotoxicity in HepG2 cells. *BMC Syst Biol* **2007**; 1: 21.
- [64] Tan Y, Shi L, Hussain SM, Xu J, Tong W, Frazier JM, Wang C. Integrating time-course microarray gene expression profiles with cytotoxicity for identification of biomarkers in primary rat hepatocytes exposed to cadmium. *Bioinformatics* **2006**; 22: 77-87.
- [65] Sauer U, Heinemann M, Zamboni N. Genetics. Getting closer to the whole picture. *Science* **2007**; 316: 550-1.
- [66] Cornish-Bowden A, Cardenas ML. From genome to cellular phenotype--a role for metabolic flux analysis? *Nat Biotechnol* **2000**; 18: 267-8.

Received: July 05, 2007

Revised: July 05, 2008

Accepted: July 09, 2008