

Genetic Network Reconstruction in Tryptophan Metabolism from Gene Expression Profiles

Akira Imaizumi^{1,2}
akira_imaizumi@ajinomoto.com

Hisashi Yasueda¹
hisashi_yasueda@ajinomoto.com

Sachiyo Aburatani¹
sachiyo@ims.u-tokyo.ac.jp

Katsuhisa Horimoto²
khorimot@ims.u-tokyo.ac.jp

¹ Laboratories of Fermentation and biotechnology, AJINOMOTO.CO.,Inc., 1-1 Suzukicho, Kawasaki-ku, Kawasaki-city, Kanagawa 210-8681, Japan

² Laboratory of Biostatistics, Institute of Medical Science, University of Tokyo, 4-6-1 Shirokanedai, Minato-ku, Tokyo 108-8639, Japan

Keywords: Tryptophan operons, genetic network, hierarchical clustering, graphical Gaussian modeling, binding site

1 Introduction

Tryptophan metabolism is one of the amino acid metabolisms that have been extensively studied. The extensive information on the genes, operons, and proteins for tryptophan biosynthesis and degradation and their regulatory mechanisms are known in diverse microorganisms, especially *Escherichia coli* [4, 5].

Here, the expression profiles of genes that were related with tryptophan metabolism in a previous study were subjected to the network reconstruction.

2 Method and Results

2.1 Gene Expression Data

The expression profile data analyzed in present study are cited from [4]. According to the previous study, we analyzed the 169 gene profiles to infer a network between them.

2.2 Clustering and Network Inference

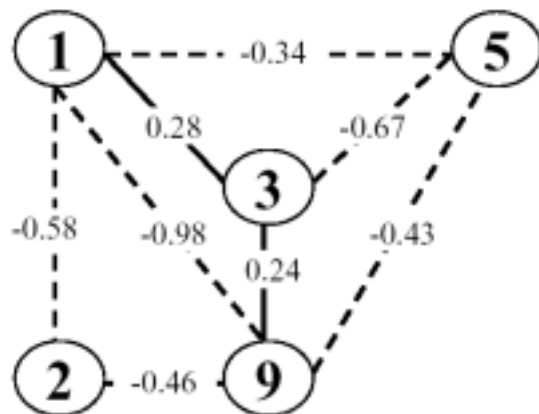
The expression profiles were analyzed by the tools on the recently constructed ASIAN web site (<http://eureka.ims.u-tokyo.ac.jp/asian>) [1, 2, 3]. The tool in the web site performs simultaneously the different types of hierarchical clusterings, the estimation of cluster number, and the network inference. To progress the profile analyses beyond the clustering, we inferred a network between clusters by the graphical Gaussian modeling (GGM) [6].

2.3 Reconstruction of Genetic Network

The network revealed more concrete and quantitative relationship between *trp* related operons and the influencing genes on them, in comparison with the previous analyses by the statistical classification of up- and down regulated genes and the clustering for gathering profiles with similar patterns (Figure 1).

Furthermore, the inferred network suggests some new insights into the genetic relationship in tryptophan metabolism, especially between tryptophan and aromatic amino acids metabolisms, leading to propose the candidates of binding sites for transcription factors by the sequence analyses of upstream regions.

A



B

Cluster No.	TrpR regulon	TyrR regulon	ArgH regulon	Others
1		<i>aroL</i> (M) <i>aroP</i> <i>aroF-tyrA</i> <i>tyrB</i>		
2				<i>roaC</i> , <i>AN</i>
3	<i>trpL</i> , <i>JEDCBA</i> <i>aroH</i> <i>trpH</i>	<i>avr</i> <i>aroG</i>		
5			<i>argCB(DG)</i> <i>argD</i> <i>argE</i> <i>argF</i> <i>argI</i>	
9		<i>tyrR</i>		

Figure.1 Inferred network between clusters (A) and correspondence between clusters and principal operons and regulating factors (B). A: Solid and dotted lines indicate positive and negative correlations between two clusters. The partial correlation coefficients between two clusters are shown on the lines. B: Parentheses show the genes which are not detected as the “significantly changed gene” in [4].

3 Discussion

The inferred network between clusters that were automatically estimated by the hierarchical clustering was well consistent with the previous study in that the gene were grouped in terms of up- and down-regulated genes in the measured conditions. In particular, the relationship between the main operons for tryptophan metabolism was clearly described by the network. Further relationship was also inferred between the operons for the tryptophan metabolism operons and the genes for the aromatic amino acid metabolism, beyond the estimation of the gene grouping. Thus, the network representation between clusters is useful to elucidate the ample and complex information including the expression profiles measured under well-controlled conditions.

References

- [1] Aburatani, S., Kuhara, S., Toh, H., and Horimoto, K., Deduction of a gene regulatory relationship framework from gene expression data by the application of graphical Gaussian modeling, *Signal Processing*, 83: 777-788, 2003.
- [2] Aburatani, S., Goto, K., Saito, S., Fumoto, M., Imaizumi, A., Sugaya, N., Murakami, H., Sato, M., Toh, H. and Horimoto, K., ASIAN: A web site for network inference, *Bioinformatics*, (in press).
- [3] Horimoto, K., Aburatani, S., Kuhara, S. and Toh, H., ASIAN – automatic system for inferring a network from gene expression profiles, *Res. Commun. Biochem. Cell Mol. Biol.*, 5: 192-207, 2001.
- [4] Khodursky A. B., Peter, B. J., Cozzarelli, N. R., Botstein, D., Brown, P. O., and C. Yanofsky., DNA microarray analysis of gene expression in response to physiological and genetic changes that affect tryptophan metabolism in *Escherichia coli*. *Proc. Natl. Acad. Sci.*, 97:12170-12175, 2000.
- [5] Pittard, A. J., Biosynthesis of aromatic amino acids, p. 458-484. In Neidhardt, F.C. et al. (ed). *Escherichia coli and Salmonella; cellar and molecular biology*, 2nd ed. ASM press, Washington D.C., 1996.
- [6] Whittaker, J. *Graphical Models in Applied Multivariate Statistics*. Wiley, Chichester, 1990.