A System for Visualizing Gene Expressions Using Metabolic Networks

Ryoko Tamaru r-tamaru@is.aist-nara.ac.jp
Toshiyuki Amagasa amagasa@is.aist-nara.ac.jp
Shigehiko Kanaya skanaya@is.aist-nara.ac.jp
Shunsuke Uemura uemura@is.aist-nara.ac.jp
Graduate School of Information Science, Nara Institute of Science and Technology

Keywords: gene expression, metabolic network, XML, SVG

1 Introduction

It has become possible to monitor gene expressions comprehensively with the rapid progress of the DNA microarray technique. Microarray data is useful for estimating the functions of a gene because a set of genes is considered to have similar functions if the genes have similar gene expression profiles. However, it is difficult to extract biologically meaningful data from such data. One promising approach for browsing gene expression data is to map them onto genetic networks. It is a well-known fact that genes interact with each other and form complex networks. Genetic networks are modeled as graphs where a node represents a gene and an edge represents an interaction of the genes at the ends. Estimating genetic networks using gene expression data is one of the most important tasks in the field of bioinformatics.

There have been several related studies to visualize gene expression data onto metabolic networks, such as KEGG (Kyoto Encyclopedia of Genes and Genomes) [2], ExPASy [1], and EcoCyc[3]. However, most studies only provide static images. Obviously, the cost of maintaining massive images of metabolic networks is expensive, and it is also difficult to update the images based on the latest information. Although some studies support automatic graph drawing, their functionalities are limited; that is, they only support one particular species or they cannot handle data based on multiple pathways. To cope with this problem, we have implemented a system that supports automatic drawing of metabolic pathways. It allows us to visualize microarray data by overlaying gene expression statuses without complex operations.

In this paper, we propose a system for visualizing DNA microarray data by mapping them onto metabolic pathways. In our implementation, we employ SVG for drawing graphs to support dynamic redrawing of metabolic networks. As a result, when we have a set of gene expression data, we can quickly spot the distinctions among those data by browsing them contrastively.

2 Our Proposed System

We exploit the KEGG pathway data of *Bacillus subtilis* written in XML format. Specifically, KEGG's metabolic pathway data comprises a number of subnetworks, and to obtain a complete network as a whole, we concatenate those subnetworks. We use SVG (Scalable Vector Graphics) for drawing the complete metabolic pathways, which enable us to draw dynamic vector graphics in XML.

Next, we explain how to draw metabolic pathways: 1) We obtain the original pathway data from the KEGG database; 2) we concatenate the pathway data, and convert it to the GraphViz format, which is an open-source software for graph drawing; and 3) we draw a complete metabolic pathway by GraphViz. Actually, we need to adjust some of GraphViz's parameters to make its outputs more readable.

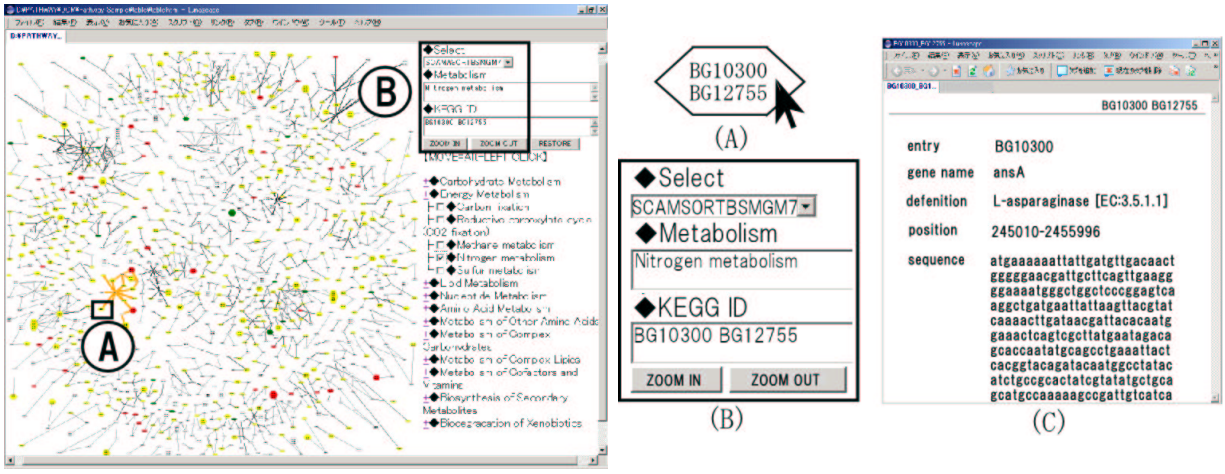

Figure 1: Operations of our system.

Next, we map gene expression data onto the metabolic pathways. Our system takes gene expression data represented in XML as its input, and makes some modifications to the pathways by changing the colors of the nodes that have positive (or negative) feedbacks. Such dynamic aspects of the graphs are implemented using Javascript.

3 Implementation

Figure 1 shows an example of our proposed system's user interface. A pathway map is displayed on the left, and the console is on the right of the figure. A user can select the experiment that he/she would like to browse by the list on the top of the console. The system than analyzes the experimental data and displays the result. In addition, when placing the mouse cursor on a node, detailed information, such as KEGG ID and metabolic name, are displayed in the textbox. When the user clicks the node, a new window is created. If a user wants to monitor gene expressions on particular pathways, he/she can select the pathways to be highlighted by the checkbox below.

4 Conclusions

We have proposed a system for visualizing DNA microarray data by mapping them onto metabolic pathways. In the future, we plan to improve the placement of the nodes in generated metabolic pathways. We also adapt to make our system to species other than *Bacillus subtilis*.

Acknowledgments

This work was supported in part by a Grant-in-Aid for the 21st COE Research Program from the Ministry of Education, Culture, Sports, Science and Technology of Japan.

References

[1] Gasteiger, E., Gattiker, A., Hoogland, C., Ivanyi, I., Appel, R.D., and Bairoch, A., ExPASy: the proteomics server for in-depth protein knowledge and analysis, *Nucleic Acids Res.*, 31:3784–3788, 2003.

[2] Kanehisa, M. and Goto, S., KEGG: kyoto encyclopedia of genes and genomes, *Nucleic Acids Res.*, 28:27–30, 2000.

[3] Karp, P.D., Riley, M., Paley, S.M., Pellegrini-Toole, A., and Krummenacker, M., EcoCyc: encyclopedia of *Escherichia coli* genes and metabolism. *Nucleic Acids Res.*, 25:43–51, 1997.