Genome-Scale Gene Expression Profiles Mapped onto the Pathway and Genome Maps in KEGG

Mitsuteru Nakao [1] **Kazushige Sato** [2]
nakao@kuicr.kyoto-u.ac.jp kazs@scl.kyoto-u.ac.jp
Tomomi Kamiya [1] **Yoshiko Kimura** [1] **Minoru Kanehisa** [1]
tomomi@scl.kyoto-u.ac.jp yoshiko@scl.kyoto-u.ac.jp kanehisa@kuicr.kyoto-u.ac.jp

[1] Institute for Chemical Research, Kyoto University, Uji, Kyoto 611-0011, Japan
[2] Nihon SiliconGraphics K.K., Umeda, Kita-ku, Osaka 530-0001, Japan

1 Introduction

The emerging technology of DNA chips and microarrays makes it possible to simultaneously analyze the expression of many genes, such as the whole set of genes in the completely sequenced genome. We have been developing a system for network oriented analysis and visualization of such genome-scale gene expression data. It projects the time-series data of gene expression profiles on the functional relations map (KEGG pathway map) and the positional neighbors map (KEGG genome map), and uncovers underlying events in differential gene expression patterns between the reference and the perturbed state (e.g., the wild-type and a disruptant, or the control and an environmental shift)

2 Data and Methods

We have developed interfaces to the tools to adapt to any gene expression profiles data. In this report we use the numerical data for the microarray gene expression profiles of *Saccharomyces cerevisiae* [1] obtained by anonymous ftp [2]. The experiments involve two types of data, the environmental change of diauxic shift and the genetic change of deletion or overexpression of a transcription factor.

We use a part of KEGG (Kyoto Encyclopedia of Genes and Geneomes) [3] database; the positional neighbor data as represented by the GENOME graphical maps and the functional relation data as represented by the PATHWAY graphical maps. The basic colorization tool has been developed and it is made available in KEGG.

URL `http://www.genome.ad.jp/kegg-bin/mk_point_pathway_multi_html` for pathway maps

URL `http://www.genome.ad.jp/kegg-bin/mk_point_genome_multi_html` for genome maps

The coloring strategy is as follows. The qualitative information in the data is represented by the difference in color, and the quantitative information is represented by the color depth. The change of two gene expression profiles between the reference state and the perturbed state is represented by two alternative colors for increased and decreased expression, a third color for no change, and a fourth color for no expression in both states. The time-series data representation is investigated such as GIF animation.

3 Results and Discussions

The publicly released microarray data of *Saccharomyces cerevisiae* was investigated. It contained a metabolic shift with genetic reprogramming, referred to as the diauxic shift from fermentation to respiration. This genetic reprogramming involved in basic cellular processes such as carbohydrate metabolism, nucleic acid and protein synthesis and energy metabolism. Fig. 1 shows the diauxic

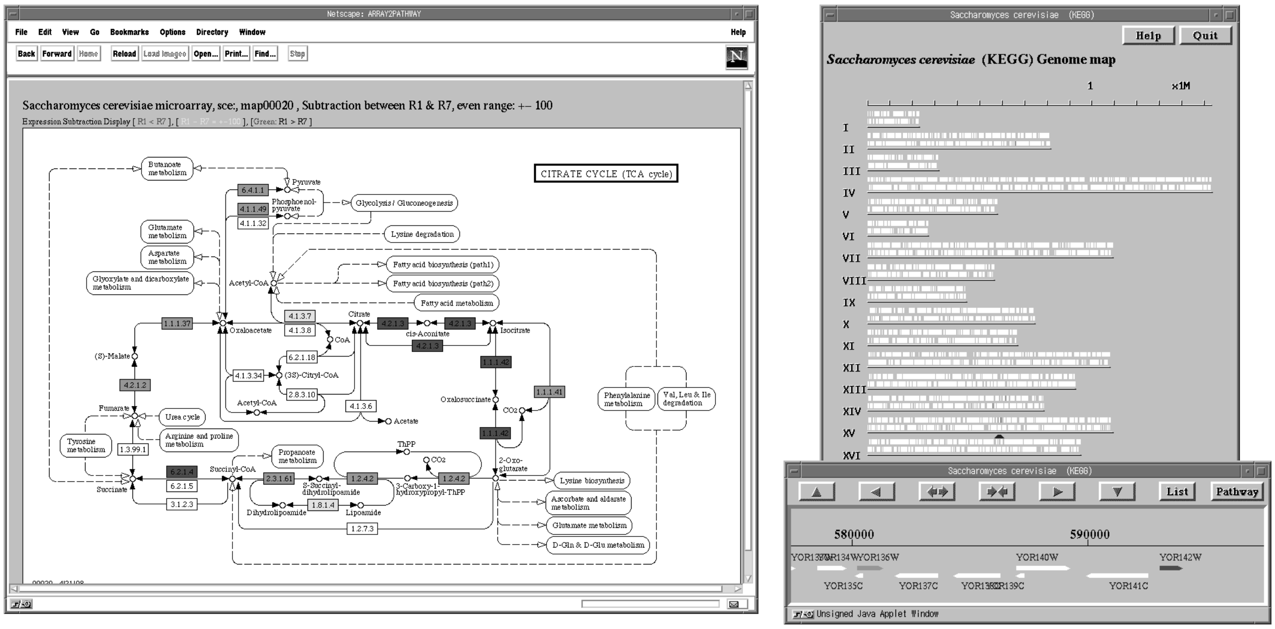

Figure 1: Gene expression profile mapped onto the PATHWAY (left) map and GENOME (right) map.

change that is automatically mapped on the pathway map of TCA cycle in KEGG. It is interesting to note that although there is no operon structure for the TCA cycle genes in *Saccharomyces cerevisiae* the bacterial genomes contain two separate operons, whenever they do have operons, that correspond to the upper and lower portions of the cycle as represented by different colors in Fig. 1. This suggests that there is a mechanism to cooperatively regulate the expression of genes that are far apart in the genome.

In general the correlation of gene expression patterns suggests functional relations and would assist to assign functions to hypothetical genes. By analyzing the positional correlation of synchronized expression of genes, common upstream activating sequences may also be found.

Acknowledgements

This work was supported in part by a Grant-in-Aid "Genome Science" for Scientific Research on Priority Areas from the Ministry of Education, Science, Sports and Culture in Japan. The computation time was provided by the Supercomputer Laboratory, Institute for Chemical Research, Kyoto University.

References

[1] DeRisi, J.L., Lyer, V.R., and Brown, P.O., Exploring the Metabolic and Genetic Control of Gene Expression on a Genomic Scale, *Science*, 278:680–686, 1997.

[2] URL http://cmgm.stanford.edu/pbrown/explore/index.html

[3] Kanehisa, M., KEGG - from genes to biochemical pathways, *Molecular Biology Databases* (Letovsky, S., ed.), in press, Kluwer Academic Press, 1998.