

Gene Expression Analysis of Orthologs in *Synechocystis* sp. with Operons in *Escherichia coli*

Yoshinori K. Okuji¹

okuji@kuicr.kyoto-u.ac.jp

Yukako Hihara²

hihara@molbiol.saitama-u.ac.jp

Ayako Kamei³

44kame@bio.c.u-tokyo.ac.jp

Masahiko Ikeuchi³

mikeuchi@ims.u-tokyo.ac.jp

Iwane Suzuki⁴

iwane@nibb.ac.jp

Norio Murata⁴

murata@nibb.ac.jp

Minoru Kanehisa¹

kanehisa@kuicr.kyoto-u.ac.jp

¹ Institute for Chemical Research, Kyoto University, Uji, Kyoto 611-0011, Japan

² Department of Molecular Biology, Faculty of Science, Saitama University, 255 Shimo-ookubo, Urawa-shi, Saitama 338-8570, Japan

³ College of Arts and Sciences, University of Tokyo, 3-8-1 Komaba, Meguro-ku, Tokyo 153-8902, Japan

⁴ National Institute of Basic Biology, Okazaki, Aichi, Japan

Keywords: ortholog, microarray, expression profile, operon

1 Introduction

The number of bacterial genomes sequenced has grown quickly, and it turned out that the structures of these genomes are not always conserved. In particular, between *Synechocystis* sp. and *Escherichia coli*, the gene clusters are rarely conserved [1]. But does this mean that transcriptional regulations are not conserved, either?

In this study, we examined how conserved the regulatory controls of orthologs are, using expression profiles in *Synechocystis* sp. and operons in *E. coli*.

2 Method and Results

To determine potential orthologs, we made use of ssearch. The similarity searches were performed using all annotated ORFs in KEGG database [2], with both *Synechocystis* sp. against *E. coli* and *E. coli* against *Synechocystis* sp. The results were classified into three groups by the number of genes similar to each gene (E value ≤ 0.05). The statistics is shown in Table 1.

Table 1: The statistics of the number of genes by the number of similar genes. *E. coli* operons are only experimentally known ones.

Organism	No. similar genes		
	0	1	> 1
<i>E. coli</i>	1917	980	1392
<i>Synechocystis</i> sp.	1317	671	1178
<i>E. coli</i> (operons)	311	285	416

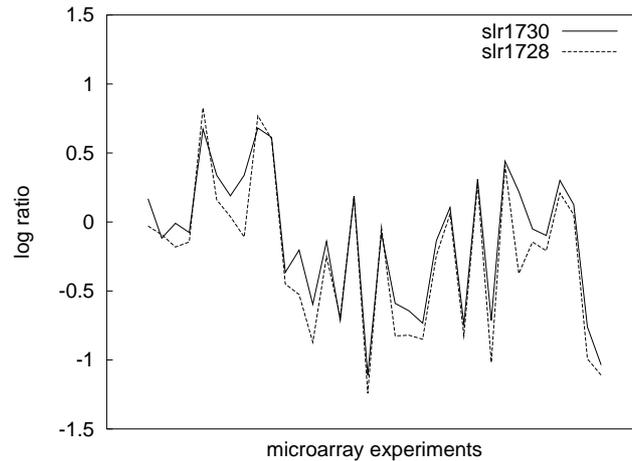


Figure 1: An example of conserved expression profiles.

Operon data in *E. coli* were gathered from RegulonDB [3]. The number of experimentally known operons was 457 and the number of the ORFs contained in them was 1012, and the number of potential orthologs is also shown in Table 1.

If a gene has a single homolog, it is trivial to identify the ortholog, so we examined the distribution of the 254 *Synechocystis* sp. genes uniquely mapped by 285 *E. coli* genes on the genomes. The result was that these genes were included in 129 operons in *E. coli*, and those genes were included in 170 groups in *Synechocystis* sp. (Genes were clustered into a group when the positions on the genome were close (≤ 3).)

We used 48 microarray experiments in *Synechocystis* sp. and computed expression similarities between each pair of ORFs, using the Euclid distance between log ratios. Then, potential orthologs in *Synechocystis* sp. were clustered using UPGMA, and the clusters were mapped to operon data in *E. coli*.

Figure 1 shows a pair of orthologs whose expression profiles are very conserved. The ORFs slr1730 and slr1728 in *Synechocystis* sp. are orthologous with *kdpC* and *kdpA* in *E. coli*, respectively.

3 Acknowledgement

We acknowledge the Synechocystis DNA chip Consortium for productions of *Synechocystis* sp. expression data, T. Kamiya for the registration of the data, and M. Kato, M. S. Nakao, K. Nakai, and S. Goto, for helpful discussions.

References

- [1] Fujibuchi, W., Ogata, H., Matsuda, H., and Kanehisa, M., Automatic detection of conserved gene clusters in multiple genomes by graph comparison and P-quasi grouping, *Nucleic Acids Research*, 28(20):4029–4036, 2000.
- [2] Kanehisa M. and Goto S., KEGG: Kyoto Encyclopedia of Genes and Genomes, *Nucleic Acids Research*, 28(1):27–30, 2000.
- [3] Salgado, H., Santos-Zavaleta, A., Gama-Castro, S., Millan-Zarate, D., Blattner, F.R., and Collado-Vides, J., RegulonDB (version 3.0): transcriptional regulation and operon organization in *Escherichia coli* K-12, *Nucleic Acids Research*, 28(1):65–67, 2000.