# Phylogenetic Analysis of Essential Genes by Using ORI-GENE

**Hirotoshi Kawano**[1]

`f673026h@bio.kyutech.ac.jp`

**Hideaki Mizuno**[2]

`mizunohda@chugai-pharm.co.jp`

**Akinori Sarai**[1]

`sarai@bio.kyutech.ac.jp`

[1] Department of Bioscience and Bioinformatics, Kyushu Institute of Technology, 680-4 Kawazu, Iizuka, 820-8502, Japan

[2] Pharmaceutical technology Dept., Research Division, Chugai pharmaceutical Co., Ltd.

**Keywords**: phylogenetic tree, evolution phylogenetic system, essential gene

## 1    Introduction

Complete genome sequences of many organisms have been determined, and the functional analysis of genes is the subject of intensive research. Essential genes, which are necessary for the survival of organisms, have been studies for several organisms such as bacteria and yeast. Since these genes are absolutely necessary to sustain cellular life, their function must be distinct from others. It is interesting to examine the relationship among functional, evolutionary and other properties of these genes, and to clarify the common features among them. ORI-GENE is a computer tool we developed before [1] to classify genes and infer their function by combining the result of homology search (e.g. BLAST) and the phylogenetic analysis. In this study, we carried out the phylogenetic analysis of essential genes by using ORI-GENE and attempt to clarify the phylogenetic relation and the origin of these genes.

## 2    Method and Results

We used ORI-GENE for the phylogenetic analysis of essential genes. ORI-GENE identifies the species of genes from a result of homology search (e.g., BLAST), and it displays the distribution patterns of the genes on the phylogenetic tree (Fig. 1a). In addition, ORI-GENE predicts the origin of the genes from the distribution pattern, and classifies them based on their origin (Fig. 1b). Thus, this tool enables us to extract more biological information than homology search alone.



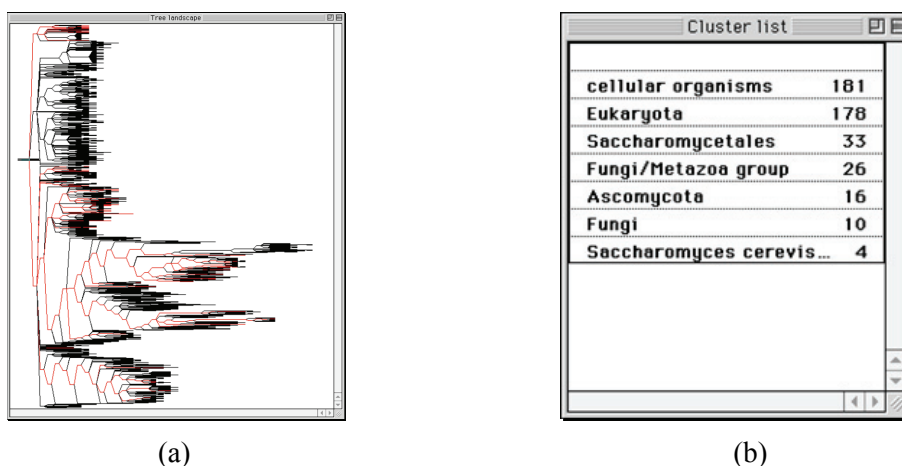(a)                                                                (b)

Figure 1: Gene distribution on the phylogenetic tree (a) and gene cluster based on the origin of genes (b).

In this study, we analyzed the phylogenetic profile of essential genes of *Saccharomyces cerevisiae*. The dataset of essential genes was obtained from DEG: Database of Essential Gene (http://tubic.tju.edu.cn/deg/) [2]. DEG is a database which contains information on essential genes of several organisms currently

available. For the reference purpose, we also analyzed non-essential genes. For the dataset of non-essential genes, we used those genes annotated as non-essential by SGD: Saccharomyces Genome Database [3]. We performed homology search using BLAST for essential genes and non-essential genes of *Saccharomyces cerevisiae*, which were obtained from these databases. Then, feeding the result of BLAST searches into ORI-GENE, we obtained the phylogenetic profiles and compared the results of essential genes and non-essential genes. So far, we have obtained a significant difference between the profiles of essential genes and non-essential genes. As expected, the origins of essential genes are closer to the root of the tree than those of non-essential genes. We are currently analyzing the relation of the phylogenetic profiles of essential genes with their functional properties for individual genes in detail.

## 3   Discussion

We have carried out phylogenetic analysis of essential genes by using ORI-GENE. This tool enables us to view the distributions of homologous genes on the phylogenetic tree consisting of an enormous number of species currently available, and identify their origin automatically. Essential genes are necessary for the survival of organisms, and thus their distributions on the phylogenetic tree are expected to spread over many organisms. This was indeed observed in our results. ORI-GENE can display detailed profile on the tree and automatically assign the origin of the genes. In order to analyze the relationship between the phylogenetic profile and functions of genes, we need to consider the information of orthologous and paralogous genes more carefully. It would be interesting to characterize essential genes from evolutionary aspect and clarify the relation with function and other properties for each gene. We will carry out more detailed analyses based on the present results.

## References

[1] Mizuno, H., Tanaka, Y., Nakai, K., and Sarai, A., ORI-GENE: Gene classification based on the evolutionary tree, *Bioinformatics*, 17:167-173, 2001.
URL: **http://dna01.bse.kyutech.ac.jp/jouhou/ORI-GENE3/**

[2] Zhang, R., Ou, H.Y., and Zhang, C.T., DEG: A database of essential gene, *Nucleic Acids Res.*, 32:271-272, 2004.
URL: **http://tubic.tju.edu.cn/deg/**

[3] SGD: Saccharomyces Genome Database: **http://www.yeastgenome.org/**