SUMMARY OF Ph.D. DISSERTATION

School	Student Identification Number	SURNAME, First name
Fundamental Science and Technology		KYODA, Koji

Title

Algorithms for Deducing Gene Networks Modeled as Signed Directed Graphs from Gene Expression Profiles of Gene Deletion Mutants.

Abstract

Identification of gene regulatory networks is essential for understanding cellular functions. Large-scale gene expression profiles measured in gene deletion mutants are invaluable sources for identifying gene networks. Signed directed graph is the most common representation of gene networks in genetics and cell biology. However, no practical procedure that deduces signed directed graphs consistent with such profiles has been developed. In this study, I developed new methods for deducing gene networks represented as signed directed graphs consistent with given expression profiles by an assumption that is commonly used in genetics and cell biology.

First, I developed the DBRF (difference-based regulation finding) method in which an algorithm deduces a non-redundant signed directed graph consistent with expression profiles of gene deletion mutants. Positive (or negative) directed edges representing positive (or negative) gene regulations are deduced by comparing the gene expression level between the wild type and mutant. A non-redundant signed directed graph is deduced by removing redundant edges from signed directed edges deduced above. A main advantage of the method is its applicability to continuous-value gene expression profiles without any translation. The performance of the method was evaluated by using continuous-value and binary artificial gene expression profiles obtained from artificial gene networks by varying the network size and indegree of each gene. The results indicate that the method is superior to the other methods that require binary translation of continuous-value gene expression profiles.

Next, I developed the DBRF-MEGN (minimum equivalent gene network) method in which an algorithm deduces the most parsimonious signed directed graphs consistent with expression profiles of gene deletion mutants. Although the DBRF method successfully deduces a non-redundant signed directed graph consistent with given expression profiles in polynomial time, the deduced signed directed graph may not be the most parsimonious signed directed graph consistent with those profiles. Because combinatorial search algorithm for deducing the most parsimonious signed directed graph requires exponential time, it may not be applicable to large-scale gene expression profiles. Use of reduction and division of the search space greatly reduces the computational cost, thus making the method applicable to large-scale expression profiles. I confirmed the applicability of the DBRF-MEGN method by applying it to the gene expression profiles of 265 *Saccharomyces cerevisiae* deletion mutants, and I confirmed the method's validity by comparing the pheromone response pathway, general amino acid control system, copper and iron homeostasis system deduced by the method with those reported in the literature. Interpretation of the gene network deduced from the *S. cerevisiae* expression profiles by using the method led to prediction of 132 transcriptional targets and modulators of transcriptional activity of 18 transcriptional regulators.

In this thesis, I present the two methods, the DBRF method and the DBRF-MEGN method. Although the DBRF-MEGN method is desirable to be applied to the expression profiles obtained from real organism, its applicability depends on the scale and complexity of given gene expression profiles. By adapting either of those two methods to suit given gene expression profiles, fruitful information for understanding cellular functions can be obtained.