Pacific Symposium on Biocomputing 6:446-458 (2001)

DEVELOPMENT OF A SYSTEM FOR THE INFERENCE OF LARGE SCALE GENETIC NETWORKS

Y. Maki, D. Tominaga
Dept. of Biochemical Engineering & Science, Kyushu Institute of Technology, 680-4, Kawazu, Iizuka-city, Fukuoka 820-8502, Japan

M. Okamoto*
Laboratory for Control Design of Biological Functions, Graduate School of Bioresource and Bioenvironmental Sciences, Kyushu University, 6-10-1, Hakozaki, Higashi-ku, Fukuoka 812-8581, Japan

S. Watanabe, Y. Eguchi
Research Institute, Mitsui Knowledge Industry Co., Ltd, 2-7-14, Higashinakano, Nakano-ku, Tokyo 164-8555, Japan

We propose a system named **AIGNET** (Algorithms for Inference of Genetic Networks), and introduce two top-down approaches for the inference of interrelated mechanism among genes in genetic network that is based on the steady state and temporal analyses of gene expression patterns against some kinds of gene perturbations such as disruption or overexpression. The former analysis is performed by a static Boolean network model based on multi-level digraph, and the latter one is by S-system model. By integrating these two analyses, we show our strategy is flexible and rich in structure to treat gene expression patterns; we applied our strategy to the inference of a genetic network that is composed of 30 genes as a case study. Given the gene expression time-course data set under the conditions of wild-type and the deletion of one gene, our system enabled us to reconstruct the same network architecture as original one.

1. Introduction

Powerful new technologies, such as DNA microarrays, provide simple and economical ways to explore gene expression patterns on a genomic scale[1,2]. Using observed gene expression data, recent advances of technology in bioinformatics have made gene expression comprehensive and several approaches have been proposed to infer the genetic networks[3-7].

We previously introduced two top-down approaches for the inference of interrelated mechanism among genes in genetic network that is based on the steady state and temporal analyses of gene expression patterns against some kinds of gene perturbations such as disruption or over expression. The former analysis is performed by a static Boolean network model based on a multi-level digraph approach[8] that can treat a large number of expression data. The latter one is by a dynamic network model such as S-system[9] that can infer the interrelated mechanism in genetic network including even loop structure (interdependent structure) among genes. We show our strategy is flexible and

* To whom all correspondence should be addressed

rich in structure to treat gene expression patterns[8]. We have already demonstrated that these models can infer a simple but large scale genetic network architecture[8,10], however, the reliability of these models obviously depends on the structure of the data given to the system. Estimation of the interrelated mechanism among genes by using experimentally observed expression data is generally referred to as "inverse problem" and the expression data correspond to the restricted conditions for solving inverse problem[10,11]. Since observed gene expression patterns at steady-state (or stationary-state) are given as restricted conditions to a static Boolean network model based on a multi-level digraph approach and observed time-course data of gene expression are to a dynamic network model such as an S-system, it is highly expected that these two models should work in a supplementary manner to cover the disadvantage and the limitation of individual models. Thus in this study, in order to improve the reliability and efficiency of the inference of genetic networks, we introduce the revised strategy which is integrated these two network models. We shall demonstrate that this strategy is useful and powerful to infer a large scale genetic network.

2. Methods

Here, we briefly describe two network models. One is a static Boolean network model based on a multi-level digraph approach[8], which mainly relies on the analysis of state changes of gene expression patterns at steady-state (or stationary state) resulting from deletion or forcible expression of one gene. The other one is a dynamic network model such as an S-system model[9], which relies on the analysis of temporal responses (time-courses) of gene expression patterns against perturbations (e.g., heat shock, hormone stimulus) or internal changes (e.g., development).

2.1 Static Boolean network model based on the multi-level digraph

A static Boolean network model based on the multi-level digraph approach treats the data representing binary relations of gene expressions. These relations describe the effects of one gene on the expression of the other genes and are mainly provided by the changes of the state of gene expression patterns. Systematical analysis of the binary relations between pairs of genes enables us to reconstruct a possible minimum architecture of the genetic network that is consistent for all of the data.

2.1.1 Identification problem

We assume that genetic network is expressed by a directed graph and that each symbol of gene and a relation between two-paired genes represent a 'node' and an 'arc' in a directed graph, respectively. A set of genes is defined as $S=\{a,b,c,...\}$. We assume here the experiments of deletion or forcible expression of one gene and that the measurements of intensities for many genes are

performed simultaneously. Examine the intensity of each resulting from the deletion or forcible expression of one target gene and check that it gets higher, lower or stays than its intensity at normal condition (wild-type of genes). A gene expression matrix E is created from a set of gene disruption experiments, in which each matrix element represents the real-valued intensity of gene expression. For instance, the value of matrix element $E(a,b)$ indicates the relative change in intensity of gene 'b' to the normal condition, which is caused by the deletion of gene 'a'. Thus the E is defined as $E=\{(a,b),...\}$. The inference procedures of this network model are as follows:

(0)Obtain the gene expression matrix E using several sets of the gene expression patterns resulting from disruption or forcible expression of one gene.

(1)Using the gene expression matrix E, for instance, if the intensity of gene 'b' is changed higher than a given threshold value θ (more than θ-times higher than the normal), or is changed lower than a given threshold $1/\theta$ (less than $1/\theta$-times lower than the normal) resulting from the disruption of gene 'a', it is defined that gene 'a' affects gene 'b' in directly or indirectly (see Fig.1 (A)) and the value of element (a,b) in the binary matrix R is set to 1; $R(a,b)=1$. Thus the binary matrix R is created by cutting the value of each element in the gene expression matrix E at the threshold (θ or $1/\theta$).

(2)In the binary matrix R, if there is the relation that gene 'a' and 'b' affect each other, that is $R(a,b)=R(b,a)=1$, we cannot decide which gene is located at the upper stream. This is the limitation or disadvantage of this method, however, we introduce an *equivalence set*, which makes a set of group consisting of genes affecting each other and the group is assumed to be one gene. The procedure for finding equivalence sets in the binary matrix R is as follows:

(2)-1 To make partition genes into equivalence sets, we use accessibility matrix R^* (see Fig. 1(B)). This matrix R^* is a reflective transitive closure of binary relation matrix.

$$R^* = \bigcup_{n=0}^{\infty} R^n \quad \begin{cases} R^0(i,j) = \begin{cases} 1: & i = j \\ 0: & i \neq j \end{cases} \\ R^{n+1} = R^n \circ R \end{cases} \tag{1}$$

$$(G \circ F)(i,j) = \min\left(1, \sum_k G(i,k) \cdot F(k,j)\right)$$

where, matrix $R^*(a,b)$ means that gene 'a' finally affects gene 'b' or not. If $R^*(a,b)=1$ and $R^*(b,a)=1$, genes 'a' and 'b' consist a closed-loop.

(2)-2 Definition of equivalence relation ER and equivalence set $[a]_{ER^*}$;

$$ER(a,b) = \begin{cases} 1: & R^*(a,b) = 1 \wedge R^*(b,a) = 1 \\ 0: & R^*(a,b) \neq 1 \vee R^*(b,a) \neq 1 \end{cases} \tag{2}$$

$$[a]_{ER} = \left\{ b \mid ER(a,b) = 1 \right\}$$

Make partition genes into equivalence set. All genes are included in only one group.

$$\begin{aligned} [a]_{ER} &= [b]_{ER} & b &\in [a]_{ER} \\ [a]_{ER} \cap [b]_{ER} &= \phi & b &\notin [a]_{ER} \end{aligned} \tag{3}$$

(2)-3 Definition of relationship between equivalence sets CR^*;
First, select one target gene from each equivalence set. Second, pick up elements of an accessibility matrix R^* associated with that target gene.

(3)Ordering genes (topological sort); The relation between equivalence sets can be described as follows; “Set A affects set B”, “Set B affects set A”, ”Set A and B are independent”. Therefore, equivalence sets have the semi-order relation and we can be drawn up equivalence set in semi-order (topological sort) to infer the network (see Fig. 1(C)).

(4)Skeleton matrix; Semi-ordered accessibility matrix between equivalence sets includes indirect affections. In order to remove them and to make skeleton matrix, we set the rank to each equivalence set defined as follows: Equivalence set belonging to rank 1 gives no *indirect* affection to another equivalence sets. Equivalence set belonging to rank 3 gives *direct* affection to the sets with rank 2 and does *indirect* affection to ones with rank 1. After setting the rank to each equivalence set, remove all indirect affection from semi-ordered accessibility matrix (See Fig. 1(D)).

(5)Draw multi-level digraph; Draw the lines between nodes based on the value of each element in the skeleton matrix. As shown in Fig. 1(E), the genes with parentheses indicate an equivalence set of genes.

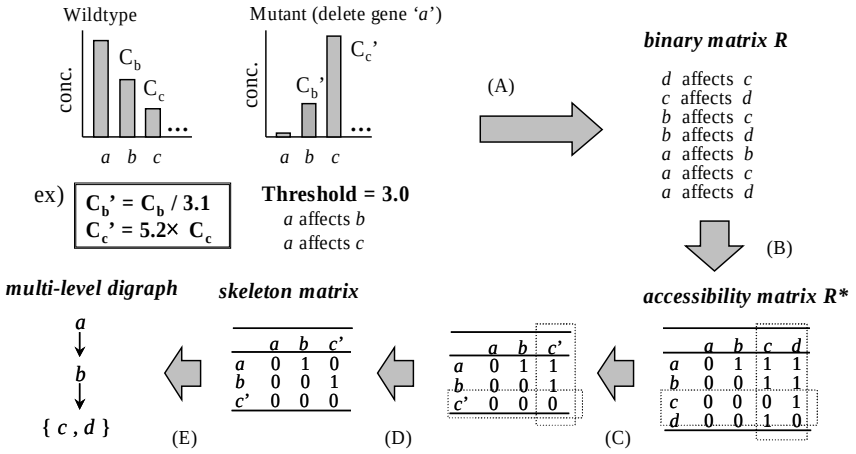

Figure 1: Inference process of Boolean network model based on multi-level digraph approach.

2.2 Dynamic network model based on S-system

Genetic networks are complex nonlinear system and the details of the interrelated mechanism at molecular level that govern interactions among system components are generally well not known. The S-system[9] is one of the best formalisms to estimate such interaction mechanisms among system components, and enables us to reconstruct genetic network architectures with the experimentally observed time-courses of the patterns of gene expression[10,11].

The S-system belongs to the type of power-law formalism because it is based on a particular type of ordinary differential equation in which the component processes are characterized by power-law functions;

$$\frac{d}{dt}X_i = \alpha_i \prod_{j=1}^{n} X_j^{g_{ij}} - \beta_i \prod_{j=1}^{n} X_j^{h_{ij}} \qquad (4)$$

where n is the total number of state variables or reactants (X_i), i, j ($1 \le i,\ j \le n$) are suffixes of state variables. The terms g_{ij} and h_{ij} are interactive effectivity of X_j to X_i. The first term represents all influences that increase X_i, whereas the second term represents all influences that decrease X_i. In a genetic network context, the non-negative parameters α_i and β_i are called relative inflow and outflow of gene X_i, and real-valued exponents g_{ij} and h_{ij} are referred to as the interrelated coefficients between genes X_j and X_i. The S-system formalism has a major disadvantage in that this formalism includes a large number of parameters that must be estimated (α_i, β_i, g_{ij} and h_{ij}) [10]; the number of estimated parameters in S-system formalism is $2n(n+1)$, where n is the number of state variables (X_i). We describe here an algorithm and procedures for the estimation (optimization) of large numbers of parameters[10,11]. The basic idea is as follows: the Genetic Algorithm (GA)[12] is introduced as a nonlinear numerical optimization method which is much less likely to be stranded in local minima. Furthermore, in order to find the skeletal structure (small-size system) of S-system formalism that matches the experimentally observed responses, some of the parameters (g_{ij} and h_{ij}), absolute values of which are less than a given threshold value, are to be removed (reset to 0) during optimization procedures. By introducing this algorithm referred to as structure skeletalizing[10,11], that optimized essential S-system model that matches the experimentally observed responses should be possible.

2.2.1 Optimization procedure

Since the S-system is a formalism of ordinary nonlinear differential equation, the system can easily be solved numerically by using a numerical calculation program to be customized specifically for this structures[13].

However, when an adequate time-course of relevant state variable is given, a set of parameter values α_i, β_i, g_{ij} and h_{ij}, in many cases, will not be uniquely determined, because it is highly possible that the other sets of parameter values will also show a similar time-course. Therefore, even if one set of parameter values that matches the observed time-courses is obtained, this set is still one of the best candidates that explain the observed time-courses. Our strategy is to explore and exploit these candidates within the immense huge searching space of parameter values.

In this optimization problem, each set of parameter values to be estimated is evaluated using the following procedure: Suppose that $X_{i,cal,t}$ is numerically calculated time-course at time t of state variable X_i and $X_{i,exp,t}$ represents the experimentally observed time-course at time t of X_i. Sum the relative error between $X_{i,cal,t}$ and $X_{i,exp,t}$ to get the total error f

$$f = \sum_{i=1}^{N} \sum_{t=1}^{T} \left\{ \left(\frac{X_{i,cal,t} - X_{i,exp,t}}{X_{i,exp,t}} \right)^2 \right\} \quad (5)$$

where N is the number of experimentally observable state variables, T is the number of sampling points of the experimental data. The problem is to find a set of parameters that minimizes f.

The proposed method is based on simple GA, and the structure of the genome (design code) of each individual (each set of parameter values) is shown in Fig.2. A genome (corresponds to one individual) contains a set of S-system parameters (n α_is and β_is, and $n \times n$ g_{ij}s and h_{ij}s) which forms an $n \times (2n+2)$ matrix. An individual represents one S-system model. Each small square in Fig.2 corresponds to each parameter that has a real value. We introduced a 32 bit unsigned integer format within a given searching region for representing real numbers; each dimensional region to be searched is divided into 2^{32} discrete points and is numbered using a unsigned integer[10,11]. A real value within a searching region is represented by scaling a unsigned integer with offset. The optimization procedure in GA has been described elsewhere[10,11].

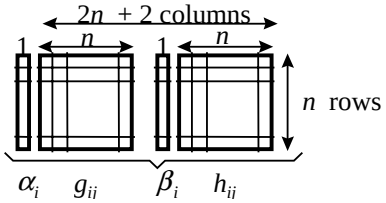

Figure 2: Design code of an individual; two n vectors of α_i and β_i, and two $n \times n$ matrices of $n \times n$ g_{ij} and h_{ij} form $n \times (2n+2)$ matrix. This matrix represents one S-system model.

2.3 Strategy for the inference of large scale genetic networks

Our strategy for the inference of large scale genetic networks is as follows: (1) Set a gene expression matrix E, which was described in 2.1.1, by using gene expression time-course data sets. (2) By using the data of the gene expression matrix, the binary matrix R is created, which is followed by the drawing of multi-level digraph described in 2.1.1. The static Boolean network model based on multi-level digraph reconstructs the genetic network classified into the independent genes and some equivalence sets. This model cannot infer the interactions of the genes involved in the equivalence set. (3) Focused on the transient time-courses of the genes belonging to each equivalence set and of the gene effecting to its set, the S-system model is applied in order to infer the network architecture of the equivalence set. Our strategy can be summarized in Fig. 3.

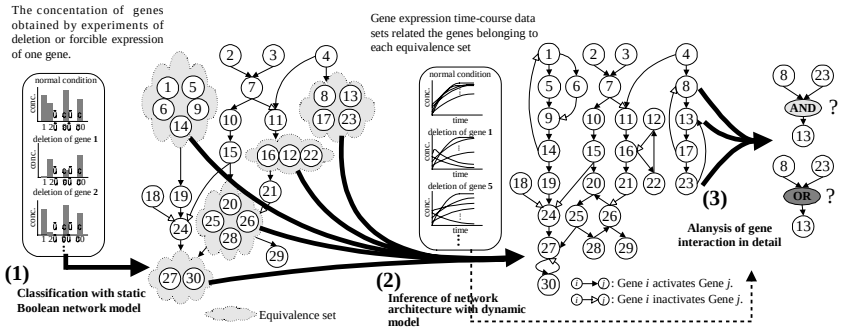

Figure 3: Strategy for the inference of interrelated mechanism of a large scale genetic network. Given the gene expression time-course data sets corresponding to the deletion or forcible expression of one gene, we obtain a gene expression matrix which is provided by the change from the state of normal condition (wild type) at an arbitrary time. (1)Given this matrix, the static Boolean network model reconstructs the network architecture classified into the genes and some equivalence sets. (2)Using the time-course of the genes related to each equivalence set, the dynamic model (S-system) enables us to infer the genetic network architecture. (3)Still more, our system can analyze the gene interaction in detail with various network models.

3. Applications

3.1 Genetic network model

In order to examine the effectiveness of our strategy, as a case study, we supposed a genetic network composed of 30 genes as shown in Fig. 4 and formulated this network in S-system formalism. Under the condition of Table 1, we prepared 31 sets of time-course data of 30 genes as experimentally observed data, that is, 1 set of time-course data in wild-type (normal condition) and 30 sets of time-course data resulting from disruption of one gene. For instance, the time-course data under the condition of the disruption of gene '1' is numerically created by setting α_1 to zero in the S-system.

3.2 Analysis with multi-level digraph approach

According to the procedure in the section 2.1, Boolean network model based on multi-level digraph was applied to the inference of network structure shown in Fig. 3 only based on prepared 31 sets of gene expression patterns at the steady state (stationary state). The threshold value θ to create the binary matrix R was set to 1.5. The result is summarized in Fig. 5. As shown in Fig. 6, which is equivalent to the inferred network structure based on the result in Fig. 5, five

equivalence sets (abbreviated A – E) were obtained. The Boolean network model based on multi-level digraph could not infer the interrelated mechanism among genes in each equivalence set. According to the step (2) in Fig. 3, the S-system model is applied to each equivalence set in order to infer the interrelated mechanism among genes belonging to the equivalence set.

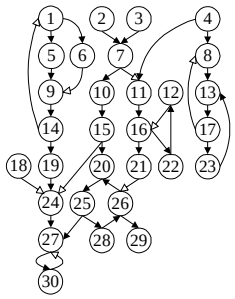

Figure 4: Genetic network model composed of 30 genes.

Table 1: Given S-system parameters in Fig 4.

α_i	1.0
β_i	1.0
g_{ij}	$g_{5.1}$=1.0, $g_{6.1}$=1.0, $g_{7.2}$=0.5, $g_{7.3}$=0.4, $g_{8.4}$=0.2, $g_{11.4}$=0.4, $g_{9.5}$=1.0, $g_{9.6}$=-0.1, $g_{10.7}$=0.3, $g_{11.7}$=-0.2, $g_{13.8}$=0.6, $g_{14.9}$=1.0, $g_{15.10}$=0.2, $g_{16.11}$=0.5, $g_{16.12}$=-0.2, $g_{17.13}$=0.5, $g_{1.14}$=-0.1, $g_{19.14}$=0.1, $g_{20.15}$=0.7, $g_{24.15}$=-0.2, $g_{21.16}$=0.6, $g_{22.16}$=0.5, $g_{8.17}$=-0.2, $g_{23.17}$=0.2, $g_{24.18}$=-0.1, $g_{24.19}$=0.3, $g_{25.20}$=0.4, $g_{26.21}$=-0.2, $g_{11.22}$=0.4, $g_{12.23}$=0.1, $g_{27.24}$=0.6, $g_{27,25}$=0.3, $g_{28,25}$=0.5, $g_{20,26}$=0.3, $g_{29,26}$=0.4, $g_{30,27}$=0.6, $g_{26,28}$=0.1, $g_{27,30}$=-0.2, other g_{ij}=0.0
h_{ij}	1.0: ($i = j$), 0.0: ($i \neq j$)

3.3 Analysis with S-system approach

By using the time-course data of the genes belonging to each equivalence set and the gene effecting to its set, the S-system approach is applied in order to infer the network architecture of the set; for example, for the case of subnet D in Fig. 6, we prepared 4 kinds of time-course data of the genes 15, 20, 21, 25, 26, 28 to infer the interrelationship among genes; one is given for the case of wild-type (normal condition), the other 3 kinds of time-course data are given for the case of the deletion of gene 15, gene 20 and gene 21, respectively. These time-course data are shown in Fig. 7; the time-course data in Fig. 7(A) is obtained by setting α_1 to the value in Table 1, 7(B), (C) and (D) are obtained by setting α_{15}=0, α_{20}=0 and α_{21}=0, respectively.

Given these 4 kinds of time-course data, we examined whether the S-system approach can infer the network architecture of equivalence set D (subnet D) or not. We attempted to estimate a part of the system parameters shown in Table 2. The targets of the optimization are the following 20 parameters

(underlined ones in Table 2 are the target); $g_{20,15}$, $g_{20,21}$, $g_{20,25}$, $g_{20,26}$, $g_{20,28}$, $g_{25,15}$, $g_{25,20}$, $g_{25,21}$, $g_{25,26}$, $g_{25,28}$, $g_{26,15}$, $g_{26,20}$, $g_{26,21}$, $g_{26,25}$, $g_{26,28}$, $g_{28,15}$, $g_{28,20}$, $g_{28,21}$, $g_{28,25}$, $g_{28,26}$. The searching range for these g_{ij} are [-3.0, 3.0] and the structure skeletalizing is performed at every generation and threshold value of which is 0.05. The values of other kinetic parameters are the same as in Table 1. The optimization procedure were described in the section 2.2

At the average 146th generation (standard deviation; 69.5) (average CPU-time is 0.94 hours (standard deviation; 0.41))(processor: Alpha 21164A, 600MHz, SPECfp95: 21.3, SPECint95: 18.6)), we found the parameter set, which is quite identical to that shown in Table 1. The average relative error

```
Set:[ G1 G2 G3 G4 G5 G6 G7 G8 G9 G10 G11 G12 G13 G14 G15 G16 G17 G18
G19 G20 G21 G22 G23 G24 G25 G26 G27 G28 G29 G30 ]
* [ G1 G5 G6 G9 G14 ]
  ->[ G19 ]
* [ G2 ]
  ->[ G7 ]
* [ G3 ]
  ->[ G7 ]
* [ G4 ]
  ->[ G8 G13 G17 G23 ]
  ->[ G11 ]
* [ G7 ]
  ->[ G10 ]
  ->[ G11 ]
* [ G8 G13 G17 G23 ]
* [ G10 ]
  ->[ G15 ]
* [ G11 ]
  ->[ G12 G16 G22 ]
* [ G12 G16 G22 ]
  ->[ G21 ]
* [ G15 ]
  ->[ G20 G25 G26 G28 ]
  ->[ G24 ]
* [ G18 ]
  ->[ G24 ]
* [ G19 ]
  ->[ G24 ]
* [ G20 G25 G26 G28 ]
  ->[ G27 G30 ]
  ->[ G29 ]
* [ G21 ]
  ->[ G20 G25 G26 G28 ]
* [ G24 ]
  ->[ G27 G30 ]
* [ G27 G30 ]
* [ G29 ]
```

Figure 5: Inferred interrelation of genes obtained by Boolean network based on multi-level digraph. The numeral followed by 'G' corresponds to the number of genes in Fig. 4, respectively

between calculated value and experimental value per sampling point is 0.86%. For another equivalence sets A, B, C and E, the S-system also enabled us to reconstruct the network completely (not shown here).

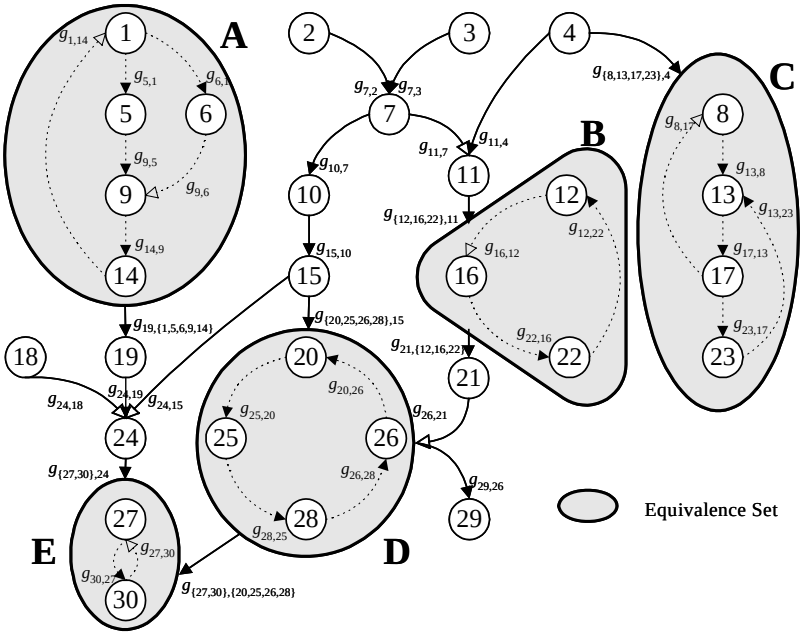

Figure 6: Obtained network architecture with multi-level digraph approach.

4. Discussions

As described already, Boolean network model based on multi-level digraph can analyze gene interactions of large scale genetic network with high speed; As for the prototype genetic network which is composed of 10,000 genes, the time for analyzing is less than a second (processor: Pentium II 300MHz), while this model cannot infer the interaction of genes belonging to an equivalence set. On the other hand, dynamic model based on the S-system can infer the network architecture even if it has closed-loop structure (corresponds to equivalence set), while the number of parameters which must be estimated increases with the order of n^2 ($O(n^2)$), where n indicates the number of system components in the network. This means that the S-system practically cannot analyze a large scale network. Thus neither model can independently analyze a large scale genetic network having many closed-loop structures. Our strategy to infer the

interactions among genes in a large scale network is the integration of both models. In this study, we could verify the effectiveness of this integration by analyzing artificial genetic network shown in Fig. 4. For practical use, the experimentally observed data generally include at least more than 10% measurement error and the number of time-course data with changing experimental conditions (deletion or forcible expression of gene) will not be enough and limited. Under these situations we have to revise our system to be interactive with users, in which the system can show several better (not best one) candidates for network architecture and propose further experimental conditions for the discrimination of these candidates; for instance, which gene or genes must be deleted in order to pick up the best candidate and so on.

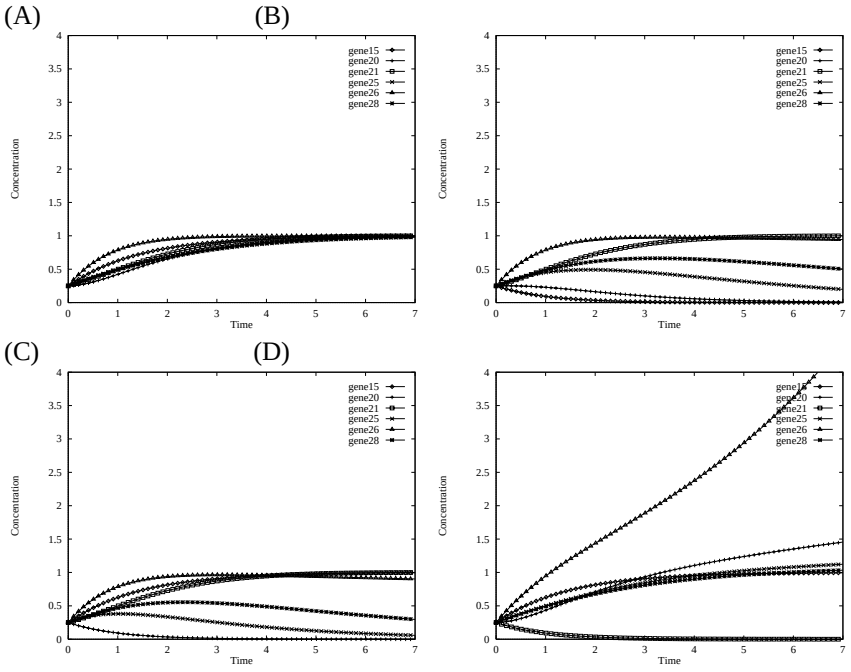

Figure 7: Time-course data related to the subnet D in Fig. 6, which are given to the S-system approach. (A)wild-type, (B)deletion of gene 15, (C)deletion of gene 20, (D)deletion of gene 21.

Table 2: Kinetic parameters (α_i, β_i, g_{ij}, h_{ij}) related to the subnet D in Fig. 6. The underlined parameters are the target of optimization.

	α_I	$g_{i,j}$						β_i	h_{ij}					
		15	20	21	25	26	28		15	20	21	25	26	28
15	1.0	0.0	0.0	0.0	0.0	0.0	0.0	1.0	1.0	0.0	0.0	0.0	0.0	0.0
20	1.0	**0.0**	0.0	**0.0**	**0.0**	**0.0**	**0.0**	1.0	0.0	1.0	0.0	0.0	0.0	0.0
21	1.0	0.0	0.0	0.0	0.0	0.0	0.0	1.0	0.0	0.0	1.0	0.0	0.0	0.0
25	1.0	**0.0**	**0.0**	**0.0**	0.0	**0.0**	**0.0**	1.0	0.0	0.0	0.0	1.0	0.0	0.0
26	1.0	**0.0**	**0.0**	**0.0**	**0.0**	0.0	**0.0**	1.0	0.0	0.0	0.0	0.0	1.0	0.0
28	1.0	**0.0**	**0.0**	**0.0**	**0.0**	**0.0**	0.0	1.0	0.0	0.0	0.0	0.0	0.0	1.0

Acknowledgements

This study was supported by Grant-in-Aid for Scientific Research on Priority Area (A) (No. 12024219) and (C) (No. 12208008) from the Ministry of Education, Science, Sports and Culture, Japan (MO) and by Grant-in-Aid for the research and development project of Industrial Science and Technology Program from NEDO (New Energy and Industrial Technology Development Organization), Japan (YE).

References

1. J. L, DeRisi et al., "Exploring the metabolic and genetic control of gene expression on a genomic scale", Science, 278:680-686 (1997).
2. X, Wen et al., "Large-scale temporal gene expression mapping of CNS development", Proc. Natl Acad Sci USA, 95:334-339 (1998).
3. T. Akutsu et al., "Identification of genetic networks from a small number of gene expression patterns under the Boolean network model", Proc. Pacific Symposium on Biocomputing, 17-28 (1999).
4. R, Somogyi et al., "Modeling the complexity of genetic networks: Understanding multistage and pleiotropic regulation", Complexity, 1:45-63 (1996).
5. M. Wahde, et al., "Coarse-grained reverse engineering of genetic regulatory networks", Information Processing in Cells and Tissues (IPCAT-99) (1999).
6. S, Liang, et al., "Reveal, a general reverse engineering algorithm for inference of genetic network architecture", Proc. Pacific Symposium on Biocomputing, 18-29 (1998).
7. M. A, Savageau, "Rules for the evaluation of gene circuitry", Proc. Pacific Symposium on Biocomputing, 54-65 (1998).
8. S. Watanabe et al., "Algorithms for Inference of Genetic Networks (AIGNET)", Intl. Workshop on Genome Informatics, (Tokyo, 1998).
9. M. A, Savageau, Biochemical Systems analysis: a study of function and design in molecular biology (Addison-Wesley, Reading), (1976).
10. D. Tominaga et al., "Nonlinear Numerical optimization technique based on a genetic algorithm for inverse Problems: Towards the inference of

genetic networks", Computer Science and Biology (Proc of the German Conf on Bioinformatics (GCB'99)) 127-140 (1999).
11. D. Tominaga et al., "Efficient Numerical Optimization Algorithm Based on Genetic Algorithm for Inverse Problem", Proc. of the Genetic and Evolutionary Computation Conf. (GECCO2000) 251-258 (2000).
12. L. Davis, Handbook of genetic algorithms, Van Nostrand Reinhold, NY (1991).
13. D.H. Irvine, M.A. Savageau, "Efficient solution of nonlinear ordinary differential equations expressed in S-system canonical form. SIAM J Numerical Analysis 27, 704-735 (1990).