IPSJ Transactions on Bioinformatics Vol. 3 24–39 (Mar. 2010)

Original Paper

Monte Carlo-based Mouse Nuclear Receptor Superfamily Gene Regulatory Network Prediction: Stochastic Dynamical System on Graph with Zipf Prior

Yusuke Kitamura,[†1] Tomomi Kimiwada,[†2]
Jun Maruyama,[†1] Takashi Kaburagi,[†1]
Takashi Matsumoto[†1] and Keiji Wada[†2]

A Monte Carlo based algorithm is proposed to predict gene regulatory network structure of mouse nuclear receptor superfamily, about which little is known although those genes are believed to be related with several difficult diseases. The gene expression data is regarded as sample vector trajectories from a stochastic dynamical system on a graph. The problem is formulated within a Bayesian framework where the graph prior distribution is assumed to follow a Zipf distribution. Appropriateness of a graph is evaluated by the graph posterior mean. The algorithm is implemented with the Exchange Monte Carlo method. After validation against synthesized data, an attempt is made to use the algorithm for predicting network structure of the target, the mouse nuclear receptor superfamily. Several remarks are made on the feasibility of the predicted network from a biological viewpoint.

1. Introduction

1.1 Introduction

Genes in cells code for one or more proteins, many of which, in turn, regulate expression of genes through regulatory pathways. Deciphering such regulatory networks from experimental data is extremely important for understanding biological processes. Since the behaviors of genes are interrelated in complex ways rather than acting in isolation, the deciphering method needs to consider experimental data as a whole, which is far from trivial. This is currently one of the exciting challenges for machine learning. As such, there is much literature on gene regulatory network prediction [1)–18)].

This study regards gene expression time-series data as trajectories of stochastic dynamical systems on a graph G defined by genes. A gene is represented by a *node* in the graph. An arc between nodes is present if a gene influences another gene, and an arc is not present otherwise.

Let $x(t) := (x_1(t), \cdots, x_N(t))$ be the expression vector at time t, where N denotes the number of nodes (genes), and let $x := (x(0), \cdots, x(T))$. The Bayes rule gives

$$P(G|x) = \frac{P(x|G)P(G)}{\sum_{G' \in \mathcal{G}} P(x|G')P(G')} \tag{1}$$

where $\mathcal{G}$ stands for the set of all possible graph configurations with a given number of nodes, and $P(G)$ is the prior distribution for the graph. The first factor of the numerator $P(x|G)$, the marginal likelihood for G, typically comes from

$$P(x|G) = \int P(x|\theta, G)P(\theta)d\theta, \tag{2}$$

where θ is a parameter vector describing the likelihood function for the data, and $P(\theta)$ is the prior distribution for θ.

Three novel aspects of this study should be noted:

(i) Recent studies on topological structures of a variety of networks, including gene regulatory networks, revealed the presence of the Zipf law, that is, a power law [16),17)]. This can be thought of as a particular type of sparseness of the network topology. This study naturally incorporates such findings as prior information within a Bayesian framework and uses Eq. (1) for making predictions.

(ii) An important issue to be addressed in general prediction problems within a graphical setting, particularly gene regulatory network prediction, is the computational complexity of evaluating the performance criteria [1)–15)]. Most of the proposed algorithms, if not all, evaluate performance criteria to select optimum or good models for making predictions. Because of the nature of the problem, the performance criterion for gene regulatory network prediction is naturally a function of the graphical structure of the underlying model.

To be more specific, first note that both Eqs. (1) and (2) are exact. Evaluation of Eq. (1) is performed in two steps. The first step is the marginalization (2).

†1 Waseda University
†2 National Center of Neurology and Psychiatry

Generally, this is non-trivial; however, if one assumes a conjugate prior $P(\theta)$ for θ, then analytical marginalization in closed form is possible. Evaluation of Eq. (1) is harder than it looks. The computational cost of evaluating the denominator of Eq. (1) is often enormous. For instance, it is known [20] that if the number of nodes of a graph is 40, the number of all possible directed non-cyclic graphs is in the order of 10^{276}. Our formulation presented in Section 2 does not exclude cyclic graphs. Therefore, exhaustive evaluation will be even more difficult. This study attempts to make predictions by computing the posterior mean instead of a single graph via a Monte Carlo method which avoids exhaustive evaluation while automatically searching for regions where probabilities are high.

More precisely, the proposed algorithm draws posterior samples of arcs from Eq. (1) via the Exchange Monte Carlo method without knowing the denominator and computes the Monte Carlo posterior mean for the graph structure prediction:

$$\sum_{G \in \mathcal{G}} GP(G|x) \approx \frac{1}{S} \sum_{k=1}^{S} G^{(k)}. \tag{3}$$

where $G^{(k)}$ stands for the k-th sample from the graph posterior (1).

After testing the prediction capabilities of the proposed algorithm, predictions are made for the target data.

A third novel aspect of this study is the target material:

(iii) Target experimental data used in this study is gene expression time-series data of nuclear receptors in proliferating neural progenitor cells (NPCs) derived from adult mouse brain, where little is known about the regulatory network structure. Those nuclear receptors are understood to be involved in several cancers, diabetes mellitus, hyperlipidemia, atherosclerosis, and immune system disorders, among others.

The organization of this paper is as follows. After describing related work in the next subsection, Section 2 describes the proposed prediction algorithm and its implementations. Section 3 examines the prediction capabilities of the algorithm against synthesized data. Section 4 describes attempted network structure predictions on experimental data obtained from the mouse nuclear receptor superfamily, followed by a discussion from a biological perspective. Section 5 concludes the paper.

1.2 Related Work

Friedman and Goldszmidt [1] observed that gene regulatory network predictions can be viewed as a problem of inferring probabilistic interdependencies between random variables, so that this particular class of prediction problems can be formulated as a Bayesian network. Since the model is static, in this formulation, it is crucial that there be conditional independence between various quantities so that the joint distribution of all the random variables involved is defined in a consistent manner, including the fact that the model graphs should be Directed Acyclic Graphs (DAGs).

Friedman, et al. [2] proposed a dynamic version of their Bayesian network model and examined the prediction capabilities. Since the model is dynamic, topological constraints are much less stringent than the static version.

Murphy and Mian [3] proposed dynamic Bayesian networks with and without hidden states. They also considered continuous state models with Gaussian transition densities and derived posterior distributions of the parameters with the so-called "weight decay" prior on the parameters.

Friedman, et al. [4] applied their Bayesian network model to the data of Spellman, et al. [5], which contains 76 gene expression measurements of the mRNA levels of 6,177 *Saccharomyces cerevisiae* open reading frames (ORFs).

Kim, et al. [11] proposed a continuous state dynamic Bayesian network with B-splines for data fitting. They proposed the criterion function BNRC-dynamic, derived from the graph posterior, where parameters are integrated out via Laplace approximation, with the graph prior being the reciprocal of the logarithm of the number of parent genes.

Beal, et al. [15] formulated the problem as a continuous-state linear dynamical system with observation noise and evaluated the marginal likelihood of a model via a variational Bayes scheme.

In parallel with the developments described above, there has been recent interest in the structure of various networks in general, including social networks, computer networks, epidemiology networks, and biological networks. There has been particular interest in their *scale free* properties, meaning that the number of elements with rank order k with respect to some ordering is proportional to $(k)^{-\gamma}$ for some real number $\gamma > 0$. Jeong, et al. [16] reports that a metabolic network

follows the power law with $\gamma = 2.2$, whereas Yook, et al. [21] shows that protein-protein network of Saccharomyces cerevisiae exhibits $\gamma = 2.1$–2.5. Basso, et al. [17] proposed an algorithm to reconstruct a gene regulatory network of human B cells taking into account this scale-free property.

2. Algorithm

2.1 Formulation

This study regards gene expression time-series data as trajectories of stochastic dynamical systems on a graph G defined by genes. A gene is represented by a *node* in the graph. An arc with an arrow directed from node j to node i represents the fact that gene j influences gene i. No arc means no influence.

Recall that a graph G consists of a set of nodes $V := \{i\}_{i=1}^{N}$, and a set of arcs $B := \{b_{ij}\}_{i,j=1}^{N}$. We regard B, and hence G, as a set of random variables. An arc b_{ij} is regarded as a random variable with three discrete values: $b_{ij} \in \{-1, 0, +1\}$, where "0" means no influence, "+" means that gene j influences gene i, and "−" means that gene i influences gene j.

The expression value at time t of gene i is represented as the state variable $x_i(t)$. A gene i may influence other genes by influencing their expression via proteins, in addition to influencing gene i itself. Such an influencing mechanism is generally dynamic in the sense that the past expression values may influence the current expression values. Thus, $x_i(t)$ may depend on $x_j(t-\tau)$, for some j not equal to i, where $\tau > 0$, and also on the value $x_i(t-\tau)$, where $\tau > 0$, the past value of the expression of gene i itself. An *arc* between nodes is present if such an influencing mechanism exists, and an arc is not present otherwise. One of the important distinctions between the dynamic model assumed here and static models is that the underlying graph structure of the former can contain *loops*, whereas the latter does not allow loops. In particular, the dynamic model allows *self loops*. A dynamical system without self loops is a very restricted class of dynamics. Note that the exclusion of loops from static models is necessary for data consistency for statistical inference.

There could be at least two uncertainties associated with gene regulatory network prediction problems. One is the uncertainty incurred from measurement uncertainty in biological experiments. Another is the possible stochastic nature of the gene expression itself [19]. In order to take into account these uncertainties, transition of the states is assumed to be stochastic. This study assumes a first-order Markov process where the current expression value $x_i(t)$ depends on the values $x_j(t-1)$, as well as on $x_i(t-1)$, in a probabilistic manner, where time is discretized with an appropriate unit length, e.g., four hours. Generalizations are possible to multiple time delay cases.

The expression values, at least in this study, are discretized into finite discrete values so that nonlinearity is captured by state transition probabilities associated with stochastic dynamics. The prediction is formulated as a graph structure prediction problem within a Bayesian framework, where a score is computed using a graph posterior distribution.

Experimental data in this study consist of gene expression data $\{x_i(t)\}$ which extend over genes, $i = 1, \cdots, N$, as well as over time, $t = 0, \cdots, T$. Consider a graph G containing N nodes where node i represents gene i. Associated with node i is its state variable $x_i(t)$, which represents the expression value of gene i at time t. A gene i may influence other genes by influencing their expression via proteins, in addition to influencing gene i itself. We assign an arc between nodes i and j if such an influencing mechanism exists, and an arc is not present otherwise. Each arc carries an arrow. If the arrow is directed from node j to node i, it represents the fact that gene j influences gene i.

Under such a graph structure, this study attempts to capture two other possible structures behind gene expression time-series data. First is the *dynamics*. We assume that $x_i(t)$ comes from a dynamical system, so that it could be influenced by the past expression values of other genes as well as its own. One of the important distinctions of this formulation from a static formulation, is that this formulation allows loops in the underlying graph, whereas a static formulation excludes loops. In particular, the current formulation considers self loops.

Second is *uncertainty* associated with the data. One of the uncertainties is measurement uncertainty of expression values, and another is the stochastic nature of the expression process itself [19].

In this study, we treat the target problem within a Bayesian framework, for which we need to define a likelihood function and prior distributions.

2.2 Likelihood: Stochastic Dynamical System on a Graph

Recall that we regard each arc b_{ij} of a graph G as a ternary random variable $b_{ij} \in \{-1, 0, +1\}$, where "+" means that gene j influences gene i, "−" means that gene i influences gene j, and "0" means no influence. This study assumes the likelihood function of the form

$$P(x(0), x(1), \ldots, x(T)|G) := \prod_{t=1}^{T} P(x(t)|x(t-1), G) P(x(0)|G) \tag{4}$$

where

$$P(x(t)|x(t-1), G) \tag{5}$$

is the conditional probability of $x(t)$, the set of expression values at time t, given $x(t-1)$, those at the previous time $t-1$. Equation (4) amounts to considering the target gene expression time-series data as a realization of a first-order Markov process with transition probability (5) and initial state probability $P(x(0)|G)$. To be more precise, we need to consider the representation of $x_i(t)$. Although the gene expression values obtained from experiments are real valued, we will discretize them, at least in this paper, into K discrete values:

$x_i(t) \in \{1, 2, \ldots, K\}$.

If $x_i(t)$ is influenced by genes $x_j(t-1)$ for j belonging to some index set, the state variables associated with such indexes may be called the *parent* states of x_i. Define

$pa_i := \{\text{parent states of } x_i\}$,

let pa_i^j be the j-th configuration of pa_i, $j = 1, \cdots, q_i$ with q_i being the *number* of configurations [34] and set

$$P(x_i(t) = k|pa_i^j, \theta_i; G) := \theta_{i,j,k} \tag{6}$$

where

$$\theta_i := \{(\theta_{i,j,k})_{k=1}^{K}\}_{j=1}^{q_i}, \qquad \sum_{k=1}^{K} \theta_{i,j,k} = 1 \tag{7}$$

where G stands for the underlying graph structure.

In order to explain Eq. (6), consider a simple example illustrated in **Fig. 1** where there are only three nodes. At node 1, the only arrow which comes in is its self loop so that

$P(x_1(t) = k|x_1(t-1) = l; G) := \theta_{1,j,k}$,

where subscript 1 stands for node 1, j stands for the configurations of parent state $x_1(t-1) = l$, and k indicates that the value that $x_1(t)$ takes is k. At node 2, one arrow comes from node 1, in addition to the self loop, so that

$P(x_2(t) = k|x_2(t-1) = l, x_1(t-1) = m; G) := \theta_{2,j,k}$

where j represents the configurations of the parent states. Finally at node 3, two arrows come from the other nodes in addition to its self loop; thus

$P(x_3(t) = k|x_3(t-1) = l, x_2(t-1) = m, x_1(t-1) = n; G) := \theta_{3,j,k}$

where j represents the configurations of the parent states.

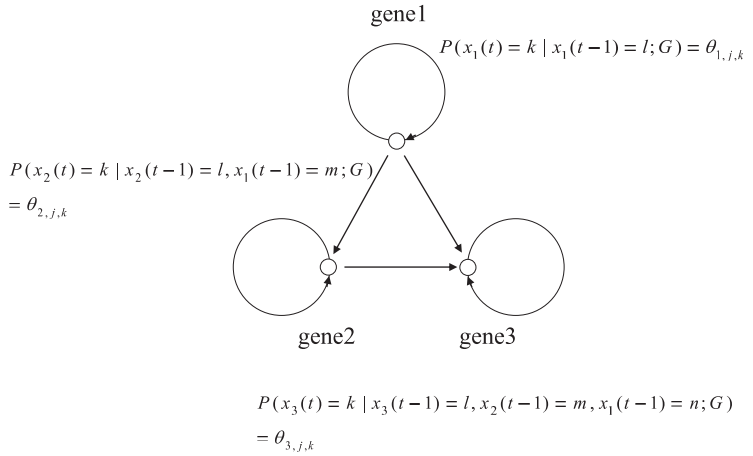

Fig. 1 A simple example with three nodes.

Remarks

1. Note that this model is not a Hidden Markov Model because state is directly observable. Note also that, in contrast with static cases, cyclic graphs and self loops $\{b_{ii}\}_{i=1}^{N}$ are allowed. In fact, a dynamical system without self loops is highly restricted. Even without topological constraints, this class of dynamic models is sometimes called dynamic Bayesian networks [2),3),11)].

2. The initial state distribution $P(x_i(0)|G)$ in this paper will be uniform over $\{1, 2, \ldots, K\}$, where

$$P(x(0)|G) := \prod_{i=1}^{N} P(x_i(0)|G).$$

2.3 Prior Distributions

There are two unknown quantities of interest: $\theta := \{\theta_i\}_{i=1}^N$ and G. Due to the formulation of the problem, the prior distribution should be of the form

$$P(\theta, G) = P(\theta|G)P(G). \tag{8}$$

(A) Prior for θ

Observe that Eq. (6) gives rise to a multinomial distribution with parameter $\theta_{i,j} := \{\theta_{i,j,k}\}_{k=l}^K$, so that a Dirichlet distribution

$$Dir(\theta_{i,j}; \alpha_{i,j}), \tag{9}$$

with hyperparameter $\alpha_{i,j} := \{\alpha_{i,j,k}\}_{k=l}^K$, is a conjugate prior distribution. With this, one can analytically marginalize our likelihood (5). To this end, let

$\theta := (\theta_1, \cdots, \theta_N)$

$\theta_{i,j} := (\theta_{i,j,1}, \cdots, \theta_{i,j,K})$

and assume independence among $\theta_{i,j}$, as well as independence among θ_i. Let s_{ijk} stand for the number of cases in which

$$x_i(t) = k \text{ with } pa(t-1)_i^j.$$

It follows from Eqs. (4), (6), and (9) that the marginal likelihood (2) is given by

$$\prod_{i=1}^{N} \int \prod_{j=1}^{q_i} \prod_{k=1}^{K} (\theta_{i,j,k})^{s_{ijk}} P(\theta_{i,j}|G) d\theta_{ij}.$$

Because of the conjugacy, one sees that the analytical marginalization is possible [34]:

$$P(x(t), \cdots, x(1)|G) = \prod_{i=1}^{N} \prod_{j=1}^{q_i} \frac{\Gamma(M_{ij})}{\Gamma(M_{ij} + s_{ij})} \prod_{k=1}^{K} \frac{\Gamma(\alpha_{ijk} + s_{ijk})}{\Gamma(\alpha_{ijk})} \tag{10}$$

where α_{ijk} is the hyperparameter associated with the Dirichlet distribution,

$$s_{ij} := \sum_{k=1}^{K} s_{ijk}, \qquad M_{ij} := \sum_{k=1}^{K} \alpha_{ijk},$$

and $\Gamma(\cdot)$ denotes a gamma function.

Note that one does not infer θ since it is integrated out in Eq. (10). Note also that Eq. (10) is computable via counting the number of cases that the state variables take.

(B) Prior for G

This study assumes the Zipf prior distribution for G with hyperparameter γ:

$$P(G) = P(G|\gamma) = \prod_{i=1}^{N} P_{Zipf}(k_i^{(G)}; \gamma) = \prod_{i=1}^{N} \frac{(k_i^{(G)})^{-\gamma}}{\sum_{z=1}^{k_{\max}^{(G)}} z^{-\gamma}}, \tag{11}$$

where $k_i^{(G)}$ denotes the degree, or the number of arcs with node i, $i = 1, \cdots, N$, and $k_{\max}^{(G)}$ is the maximum number of arcs of a single node.

Generally, the Zipf law states that the number of elements with rank order k with respect to some ordering is proportional to $k^{-\gamma}$. This law has been reported in a variety of disciplines, including biological networks. See, for example, Ref. 21). Observe that the Zipf law is a form of sparseness of arcs in a graph. While many of the nodes have small degrees, there are nodes that have large degrees, although they are few. **Figure 2** is a simple illustration of the Zipf law, and **Fig. 3** is a schematic diagram showing a "random" graph, where the degree distribution is centered around a particular value.

2.4 Posterior Distributions

Given a time-series data set $(x(0), x(1), \ldots, x(T)), x(t) := (x_1(t), x_2(t), \ldots, x_N(t))$, the posterior distribution of G is given by

$$P(G|(x(0), x(1), \ldots, x(T)), \gamma) = \frac{P((x(0), x(1), \ldots, x(T))|G)P(G|\gamma)}{\sum_{G' \in \mathcal{G}} P((x(0), x(1), \ldots, x(T))|G')P(G'|\gamma)} \tag{12}$$

where the first factor in the numerator is available from Eq. (10), and the second factor in the numerator is the Zipf prior alluded earlier.

2.5 Implementation

Our goal in this paper is to evaluate Eq. (12) for predicting plausible G. In order to achieve this goal, generally speaking, we need three different quantities: (a) the marginal likelihood for G, which is the denominator of Eq. (12); (b) the hyperparameter γ for the Zipf prior (11); and (c) the hyperparameters $\alpha_{i,j}$ for

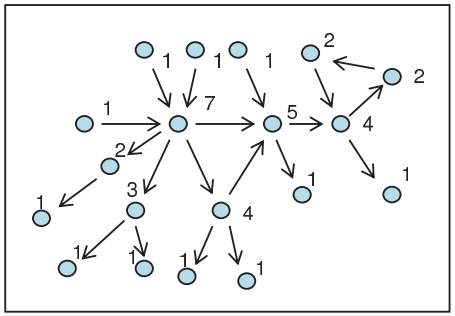

Fig. 2 A schematic diagram demonstrating the Zipf law. While many of the nodes have small degrees, there are nodes that have large degrees, although they are few. The numerals show degrees at nodes. Self loops are omitted for simplicity.

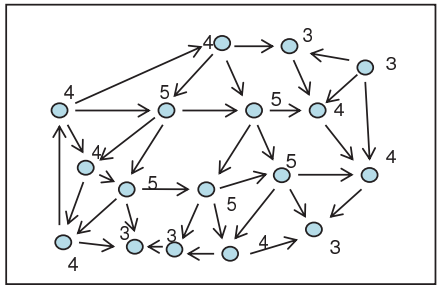

Fig. 3 A schematic for a random graph. The degree distribution is centered around a particular value. The numerals show degrees at nodes. Self loops are omitted for simplicity.

the Dirichlet prior (9). Each of them is generally nontrivial to set. Note that θ has been integrated out in Eq. (10). In this paper, we adopted a Markov Chain Monte Carlo (MCMC) method to evaluate Eq. (12).

The first and probably the most difficult quantity: the marginal likelihood for G. To obtain the marginal likelihood, we need to evaluate the sum,

$$\sum_{G' \in \mathcal{G}} P((x(0), x(1), \ldots, x(T))|G')P(G'|\gamma),$$

where $\mathcal{G}$ stands for the set of all graph configurations with the number of nodes fixed. Note that the sum must run through all possible G's, which is difficult due to the magnitude of cardinality of $\mathcal{G}$. One of the greatest advantages of the class of MCMC methods which we use in this paper is the fact that the normalization constant vanishes within the procedure. Exchange Monte Carlo designs a family of target distributions instead of a single target distribution. While running different Markov chains, it exchanges the states of different Markov chains with some probability so that the chains will mix together and eventually give rise to reasonable samples. Since MCMC is a widely used method, our explanation here is brief and heuristic.

Hyperparameter γ is another quantity which is difficult to set. While there may be methods of learning its value from the data, this paper uses an informative prior, i.e., the value available from previous studies. In most of the literature on the Zipf structure of biological networks reported so far, γ appears to be between 2 and 3. This paper uses $\gamma = 2.2$ based on the result of Ref. 21), where the authors studied protein-protein interaction networks and reported that the γ values lie between 2.1 and 2.5.

Although the hyperparameters for the Dirichlet prior $\alpha_{i,j}$ may be learned from the data, we will set, at least in this work, all of them to unity. Our future research will include learning these hyperparameters. This is equivalent to saying that there is no prior information on the relative frequencies of $x_i(t) = k$.

2.6 MCMC Procedure for Graph Posterior Samples

Let

$$P(G|x, \gamma) \propto P(x|G)P(G|\gamma) \tag{13}$$

be the target distribution with $x := (x(0), \ldots, x(T))$, and consider the parameterized family of Q different distributions

$$\begin{aligned} &P_q(G|x, \gamma) \propto (P(x|G)P(G|\gamma))^{1/T_q}, \\ &\quad q = 1, \ldots, Q, \\ &\quad 0 < T_1, < T_2, \ldots < T_Q = 1. \end{aligned} \tag{14}$$

And set

$$\begin{aligned} &P_q(x|G, \gamma) := (P(x|G, \gamma))^{1/T_q}, \\ &\quad q = 1, \ldots, Q. \end{aligned} \tag{15}$$

Then, implementation is given as shown in **Fig. 4**.

(1) for $q = 1,\ldots,Q$

A. Initializing Procedure

(2) for $k = 1,\cdots,S$

(i) for $q = 1,\cdots,Q$

B. Graph Transition Procedure: Metropolis Sampler

C. Exchange Procedure

·set $k := k+1$ and go to (2)

Fig. 4 The overview of exchange Monte Carlo procedures for graph posterior samples. Precise steps for Procedures A. through C. are presented below.

A. Initializing Procedure

consider Q Markov Chains with target $\propto P_q\left(x|G\right)P_q\left(G|\gamma\right)$
at time $k = 0$
·set $b_{ij}^{(0)} = 0$, i, j $= 1,\ldots,N$, i.e., all nodes are isolated:
·call this graph $G_q^{(0)}$

Fig. 4 A. The details of initializing procedure.

Remarks

(1) Our prediction in this paper is performed using the Monte Carlo posterior mean of the graph:

$$\frac{1}{S}\sum_{n=\tau+1}^{\tau+S} G_n, \tag{16}$$

where G_n is the posterior sample obtained by the Monte Carlo method described above with a burn-in period τ.

B. Graph Transition Procedure: Metropolis Sampler

•*Graph Transition: Metropolis*
given $G_q^{(k-1)}$,
(a)
·choose two nodes i and j randomly
if $b_{ij}^{(k-1)} \neq 0$, then either set $b_{ij}^* := -b_{ij}$, or $b_{ij}^* := 0$ with equal probability
else set $b_{ij}^* := 1$, or $b_{ij}^* := -1$ with equal probability
endif
(b) set the resulting graph G_q^*
•*Sampler*
(c) compute $\beta_q := \min\left(1, \frac{P_q\left(x|G_q^*\right)P_q\left(G_q^*|\gamma\right)}{P_q\left(x|G_q^{(k)}\right)P_q\left(G_q^{(k)}|\gamma\right)}\right)$
(d) generate $u_1 \sim\ U(0,1)$, the uniform distribution over (0, 1)
if $u_1 \leq \beta_q$, accept G_q^*, i.e., set $G_q^{(k)} := G_q^*$
else reject G_q^* , i.e., set $G_q^{(k)} := G_q^{(k-1)}$
endif

Fig. 4 B. The details of graph transition procedure: Metropolis Sampler.

C. Exchange Procedure

(ii) compute

$$w_q := \min\left(1, \frac{P_q(G_q^{(k)}|x)P_{q'}(G_{q'}^{(k-1)}|x)}{P_q(G_{q'}^{(k-1)}|x)P_{q'}(G_q^{(k)}|x)}\right)$$
$$= \min\left(1, \frac{P_q(x|G_q^{(k)})P_q(G_q^{(k)}|\gamma)P_{q'}(x|G_{q'}^{(k-1)})P_{q'}(G_{q'}^{(k-1)}|\gamma)}{P_q(x|G_{q'}^{(k-1)})P_q(G_{q'}^{(k-1)}|\gamma)P_{q'}(x|G_q^{(k)})P_{q'}(G_q^{(k)}|\gamma)}\right)$$

(iii) generate $u_2 \sim\ U(0,1)$
if $u_2 \leq w_q$, exchange q with q'
else no exchange

Fig. 4 C. The details of exchange procedures.

(2) It should be noted that the graph posterior consists of

$$P(\{b_{ij}\}_{i,j=1}^{N}|(x(0),x(1),\ldots,x(T)),\gamma), \tag{17}$$

where $b_{ij} \in \{-1,0,+1\}$ together with the set of nodes V.

(3) We do not need to know the normalization constants for $P(G|x,\gamma)$, $P_q(G|x,\gamma)$, $q=1,\ldots,Q$, because β_q and w_q are the ratios.

(4) Hyperparameter for the Dirichlet distribution in principle, may be learned from the data. Because of the scarcity of the data for learning, we select not to perform learning, at least in the present paper.

3. Experiment 1: Synthesized Data

This section examines the prediction performance of the proposed algorithm against synthesized data. **Figure 5** shows the target graph with the number of nodes $N = 20$; many of the nodes have small degrees, and some have larger degrees. Time-series data is generated according to Eqs. (4)–(6) and the number of discrete values of the state $x_i(t)$ is $K = 3$. Since the purpose of the experiment with simulated data is to validate the performance of our proposed algorithm, the number of data points is set to emulate the real data experiment described in Section 4. Generally, the more the number of nodes the network has, the more difficult it is to estimate its topology. To have equal footing among the experiments, the ratio of the number of nodes N to the number of time points T is fixed, i.e., $\frac{T}{N} = const.$ The number of nodes N is restricted by the selection of genes one is interested in. There is more flexibility in the selection of time points T, as well as the number of experiments. In the biological experiment, $N = 35$, and *three sets* of $T = 19$ data points are obtained. Therefore, for our experiment in the synthesized data, given the number of nodes $N = 20$, we set *three sets* of $T = \frac{19}{35} \times 20 \fallingdotseq 11$ as the number of data points. To explain the implementation of our prediction algorithm, first let $b_{ij}^{(k)}$ be an arc of a posterior sample graph defined by Eq. (17) and note that it has one of three possible values: 0, +1, and −1. We predict b_{ij} according to

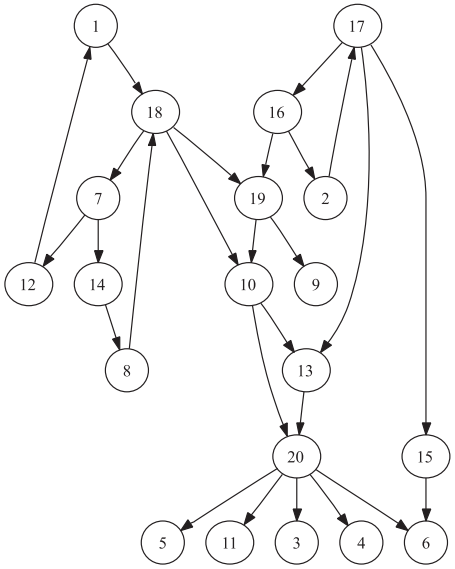

Fig. 5 Target graph for prediction experiment with synthesized data. Self loops are omitted for simplicity.

$$b_{ij} = \begin{cases} 0, & \text{if } \frac{1}{S}\{\#b_{ij}^{(k)}|b_{ij}^{(k)} \neq 0, k=1,\ldots,S\} < \xi \\ +1, & \text{if } \frac{1}{S}\{\#b_{ij}^{(k)}|b_{ij}^{(k)} \neq 0, k=1,\ldots,S\} \geq \xi \\ & \text{and } \frac{1}{S}\{\#b_{ij}^{(k)}|b_{ij}^{(k)} = +1, k=1,\ldots,S\} \\ & \qquad < \frac{1}{S}\{\#b_{ij}^{(k)}|b_{ij}^{(k)} = -1, k=1,\ldots,S\} \\ -1, & \text{else.} \end{cases} \tag{18}$$

where S is the number of samples and ξ is a certain threshold.

Figure 6 shows our prediction with

- γ (the Zipf prior hyperparameter) = 2.2
- number of samples: 500,000 (burn in 1,000,000)
- number of replicas: 5 with temperature 1/1.0, 1/0.9, 1/0.8, 1/0.7, 1/0.6

The predicted result has only one error: a false negative indicated by the dashed arrow. The remaining solid arrows indicate true positives.

In order to reveal the possible role of the Zipf prior, we drew a False Positive Rate (FPR)-Sensitivity Receiver Operating Characteristic (ROC) curve. More specifically, we define

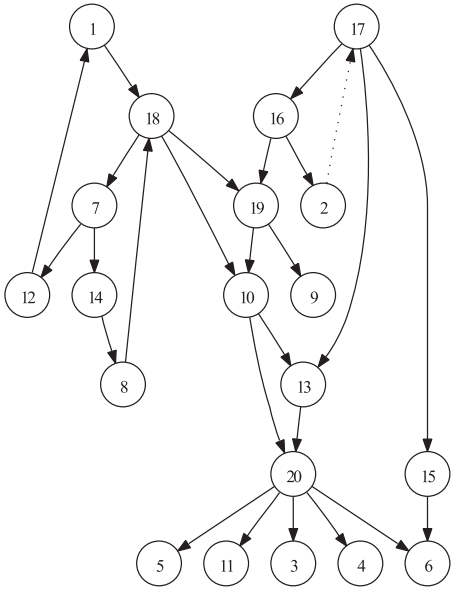

Fig. 6 Predicted network with the proposed algorithm. There is one false negative indicated by a dotted arrow. The solid arrows indicate that the predictions are correct. Self loops are omitted for simplicity.

TP (True Positive): number of correctly predicted nonzero arcs

FP (False Positive): number of predicted nonzero arcs where true graph does not have nonzero arcs

TN (True Negative): number of correctly predicted zero arcs

FN (False Negative): number of predicted zero arcs where true graph has nonzero arcs

Then, we have

$$\begin{aligned} &\text{Sensitivity} := \tfrac{TP}{TP+FN}, \qquad &&\text{Specificity} := \tfrac{TN}{TN+FP} \\ &\text{FPR} := 1 - \text{Specificity}, \qquad &&\text{PPV} := \tfrac{FN}{TN+FP}. \end{aligned} \tag{19}$$

Let ξ be the threshold value below which the algorithm predicts that a target arc assumes a zero value. Varying ξ from 0 to 1 produces a curve in the FPR-Sensitivity plane, which is the ROC curve. A natural performance criterion will be the area under this curve.

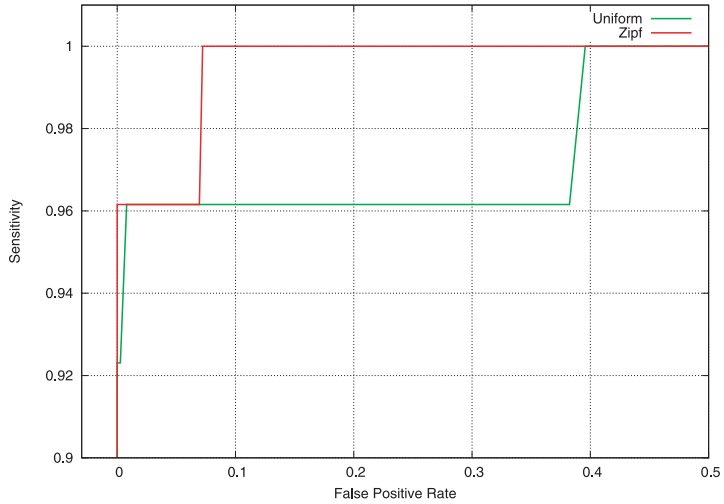

Fig. 7 ROC curves for performance comparison between uniform prior and Zipf prior for the graph. The larger the area under a curve, the better the prediction performance. Prediction with the Zipf prior appears functional.

Figure 7 shows the ROC curves for the target graph shown in Fig. 5, as before, with the proposed Zipf prior with $\gamma = 2.2$ and the uniform prior shown on the graph. The proposed Zipf prior appears to be better than the uniform prior.

4. Experiment 2: Real Data

Nuclear receptors represent a superfamily of ligand-dependent transcription factors that regulate essential biological processes, including development, reproduction, and metabolism. The classical endocrine receptors that mediate the actions, such as steroid hormones, thyroid hormones, vitamins A and D, and orphan receptors whose endogeneous ligands are unknown, are included in this superfamily.

In neural progenitor cells, which have unique capabilities of self-renewal and pluripotency to differentiate into neurons, astrocytes and oligodendrocytes, it has been shown that several nuclear receptors are critically required for their differentiation or proliferation. For example, retinoids, the active metabolites

of vitamin A, regulate complex gene networks through heterodimers between the retinoic acid receptor (RAR) and retinoid X receptor (RXR). Despite the importance of these diverse aspects, the signal transduction and transcriptional regulatory pathways of nuclear receptors are still not completely understood.

In this study, nesural progenitor cells derived from adult mouse brains were applied as an experimental sample of particular interest and we aimed to investigate the possible network of nuclear receptors by using a bioinformatics approach. The number of members in the nuclear receptor superfamily has been shown to be 48 in the human genome database [35]. We selected the 48 genes reported in the human genome database out of 49 genes in the mouse genome database [36].

Progenitor cells are important target for the challenge because of their multipotency to produce various types of cells. Among them, neural progenitor cells in the adult brain are of particular interest, since it had been believed for long time that neurogenesis never happens in adult.

Shown in **Fig. 8** is a coronal section of adult mouse. Neural progenitor cells locates in the region surrounding anterior lateral ventricle (aLV). Closed area drawn by a pen shows a representative region including aLV.

Neural progenitor cells are prepared from the region surrounding aLV according to the procedure in the Methods. Cultured neural progenitor cell forms neurospheres as shown in **Fig. 9**.

In general, progenitor cells are designated as the cells with the ability for self renewal and multipotency. From neural progenitor cells, neurons and glial cells are generated (see **Fig. 10**). To achieve the proliferation and the differentiation, neural progenitor cells have the strict gene expression system, though its detailed mechanism is unknown.

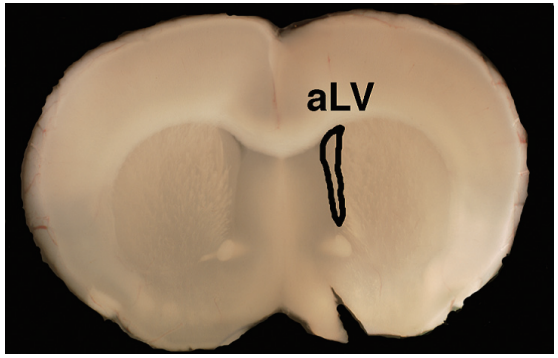

Fig. 8 Coronal section of an adult mouse.

4.1 Animals and Cell Culture

Adult male C57BL/6J mice, 6 weeks old, were purchased from CLEA Japan

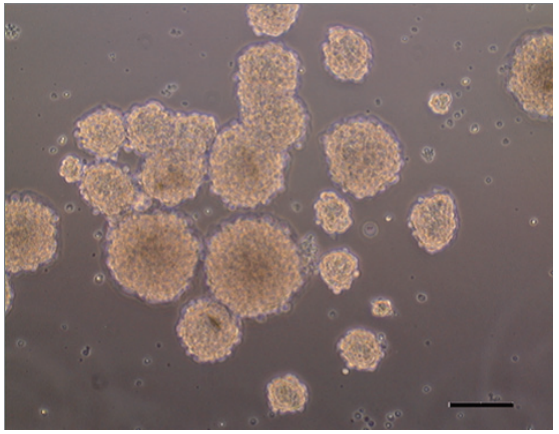

Fig. 9 Cultured neural progenitor cells form neurospheres.

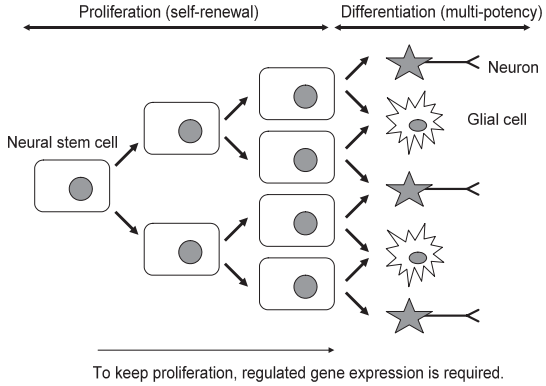

Fig. 10 Proliferation and differentiation.

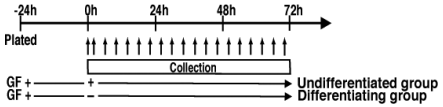

Fig. 11 Proliferation experiments.

(Tokyo, Japan). All experiments were performed in a laboratory designated for animal experiments according to the *NIH Standards for the Treatment of Laboratory Animals*.

Adult mouse NPCs were prepared as previously reported [23),24)], with some modifications. Briefly, adult C57BL/6J mice were sacrificed by cervical dislocation. Their brains were removed and covered with a 1.5% solution of low-melting-temperature agarose (Cambrex, Rockland, ME) at 37°C. Each brain was sliced on a vibratome into viable 500 μm coronal sections in cold oxygenated phosphate buffered saline (PBS). Rostral sections containing the anteriorlateral sub-ventricular zone (SVZ) were dissected under a microscope (SZ61, Olympus, Tokyo, Japan) and enzymatically dissociated with 0.25% trypsin-EDTA (Invitrogen, Carlsbad, CA) for 15 minutes at 37°C. The tissue was then mechanically triturated into a single cell suspension and serum-free DMEM/F12 (Invitrogen) containing 0.5 mg/ml trypsin inhibitor (Invitrogen), and DNase I (Sigma) was added. The solution was left for 15 minutes at 37°C. The cell suspension was then centrifuged at 1,500 rpm for 5 minutes and the supernatant was removed. The isolated primary cells were plated at a density of 20 viable cells/μl in HydroCell low cell-binding culture plates (CellSeed, Tokyo, Japan). The cells were expanded for 7 days in serum-free DMEM/F12 containing B-27 supplement (Invitrogen) diluted 50-fold, 20 ng/ml of EGF (PeproTech EC, London, UK), 20 ng/ml of bFGF (PeproTech EC), 100 U/ml of penicillin, and 100 μg/ml of streptomycin (P/S; Invitrogen).

For the secondary culture, cells were dissociated with 0.25% trypsin-EDTA and mechanically triturated into a single-cell suspension. The cells were reseeded and cultured for a further 7 days under the same conditions as the primary culture. We used cells derived from the secondary culture for experiments.

4.2 Adult NSPC Studies and Sampling

For proliferation experiments, the cells were replated on poly-L-ornithine and fibronectin-coated dishes or plates at a density of 300,000 cells/ml in serum-free DMEM/F12 supplemented with B27, EGF, bFGF, and P/S in a monolayer culture. At this stage, $> 90\%$ of the cells were Nestin-positive, a neural progenitor marker. On the following day, the medium was replaced with fresh medium and the cells were separated into two groups: undifferentiated NPCs with growth factor, and differentiating NPCs without growth factor. It is generally accepted that NPCs initiate differentiation upon removing the growth factor. Indeed, neuronal and glial cells were produced in our study (data not shown). In the present study, cells from the undifferentiated proliferating group were collected every 4 hours over a 70-hour period and the expression levels of 17 clock and clock-related genes were examined by quantitative RTPCR (**Fig. 11**).

4.3 Real-time Quantitative RT-PCR

Real-time quantitative RT-PCR was carried out using the SYBR Green-based method (ABI PRISM 7700 Sequence Detection System; Applied Biosystems, Foster City, CA), as previously described in Ref. 25). Briefly, total RNAs were isolated from cultured cells using a RNeasy Mini Kit (Qiagen, Hilden, Germany), and subjected to first-strand cDNA synthesis^{TM} (Invitrogen) according to the manufacturer's instructions. PCR amplifications were performed using the following protocol: 2 minutes at 50°C, 10 minutes at 95°C, and 40 cycles of 15 seconds at 95°C and 1 minute at 60°C. Dissociation reaction plots were produced to confirm the specificity of the PCR. The products were analyzed using the sequence detection system software developed by Applied Biosystems (version 1.7). The SYBR Green signal for the glyceraldehyde-3-phosphate dehydrogenase (Gapdh) gene amplicon was used as a reference. Samples derived from three independent experiments were used for analysis of the relative gene expression data using the comparative Ct method (User Bulletin #2; Applied Biosystems).

Information regarding the primer sequences for each member of the nuclear receptor superfamily are available upon request.

4.4 Prediction

4.4.1 Preprocessing

There were 48 genes in our target superfamily. Three time-series data sets were measured, each covering a 72-hour period with a sampling frequency of 4 hours, so that each data set consisted of 19 points. Of the 48 genes, 13 consistently

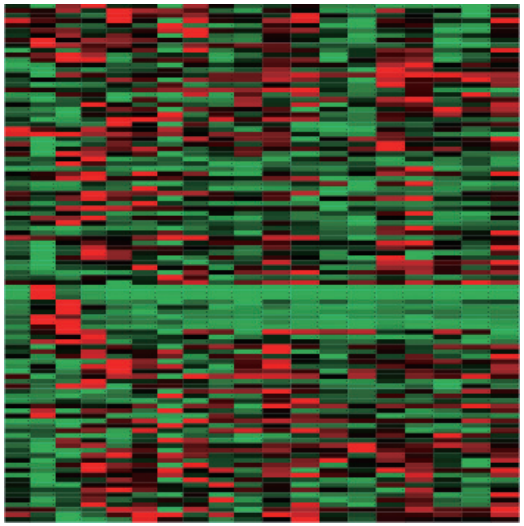

Fig. 12 Target time series data for 72 hours with sampling period 4 hours consisting of 19 points. Of the 48 genes, 13 consistently exhibited very small amplitudes in their expression values so that they were discarded. The expression values are normalized between 0 and 1 at least in this study.

exhibited very small amplitudes in their expression values. Those 13 genes were discarded. Each expression time series was normalized between 0 and 1 and was then discretized into three equally spaced discrete values, at least in this study. The resulting time-series data is shown in **Fig. 12**. We will perform prediction experiments with other types of preprocessing in our future work.

4.4.2 Parameter Setting

The parameters were set as in our prediction described in Section 3.

4.4.3 Prediction Implementation

In order to cope with the sparseness of the data, we drew posterior samples from 100 independent MCMC simulations, in which parameters were set as described in Section 3. We ranked ordered arcs b_{ij} by

$$\frac{1}{S^*}\{\#b_{ij}^{(k)}|b_{ij}^{(k)} \neq 0, k = 1, \ldots, S^*\}$$

where S^* denotes the total number of posterior samples from the 100 MCMC simulations, namely, $100S$. Those arcs ranked less than R were discarded and the remaining arcs were predicted in a manner similar to that defined in Eq. (18). **Figure 13** is our prediction with $R = 60$.

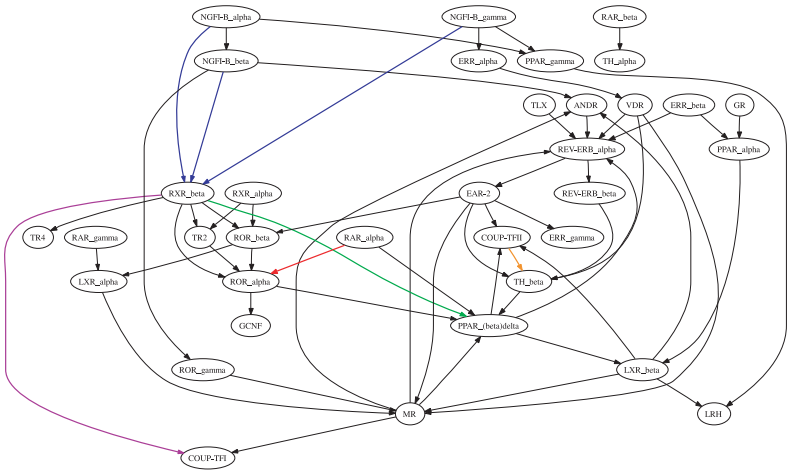

Fig. 13 Prediction of the target network with the proposed algorithm.

4.5 Discussion

There is no ground truth for this network. There are, however, several well known biologically plausible scenarios in the target genes of this study. For instance, a previous biological study reported synergistic regulation of a cerebellum-specific gene by ROR_alpha and RAR [26]. The predicted network (Fig. 13) appears to indicate the functional relationship between ROR_alpha and RAR, as indicated by the red arrow. ROR_alpha has been shown to be a key regulator for the cerebellum development [27]. Another scenario is the RXR heterodimer arc between PPAR and LXR, as revealed in Ref. 28). The green arrow in Fig. 13 appears to indicate this arc. It has also been verified that RXR with COUP-TF interactions modulates rethinoic acid signaling [29]. This relationship appears to be indicated by the purple arrow in Fig. 13. It is also known [30] that the orphan nuclear receptor NGFI-B (also called Nur77) can heterodimerize with RXR. This appears to be indicated by the blue arrow. Finally, COUP-TFII with TH beta heterodimer is also reported in Ref. 31). The orange arrow of Fig. 13 appears to

indicate this.

It seems reasonable to conclude that NGFI-B genes are located upstream of the gene regulation we predicted, since these genes are known to be immediate early genes. Besides NGFI-B genes, our prediction suggests that GR, TLX, and ERR_alpha genes are located upstream of some nuclear receptor genes, including PPAR_delta and LXR_beta. Both GR and TLX are suggested to play some role in neurogenesis [32),33)]. Our prediction may provide further clues to identify molecular cascades regulating the fate and function of neural progenitor cells.

For comparison, we tested the same data set using a greedy search method. The greedy method could not find any of the relationship biologically verified in the literature.

5. Conclusion

A Monte Carlo based algorithm was proposed for predicting a gene regulatory network from time-series expression data by formulating the problem as a graph prediction problem on which a stochastic dynamical system is defined. The prediction was performed by evaluating the graph posterior mean within a Bayesian framework, where the graph prior distribution is assumed to follow the Zipf prior, which is incorporated by taking into account several recent biological studies. The algorithm was first tested against synthesized data for which the ground truth is available, and the prediction was shown to be reasonable. Based on this, an attempt was made to predict the gene regulatory network structure of the mouse nuclear receptor superfamily. The prediction was discussed from a biological perspective.

The graphical representation of gene regulatory networks proposed in this study is from a bioinformatics view point. An arrow from gene j to gene i represents the fact that gene j influences the expression values of gene i via proteins. Even though further concreteness is not currently possible, the reported predictions can serve as guidelines for biological experiment design to verify the predicted interactions, which may lead to new biological findings.

It is a challenging problem to infer γ, the hyperparameter for the Zipf prior, from the available data. It will be interesting to incorporate as a part of the graph prior, if any, partial prior information about how genes are related with each other from biologically verified data, besides those arcs discussed in Section 4.5. A possible means of evaluating a graph structure other than our posterior mean is the graph posterior mode, which could be our next topic of investigation. Another project which is already underway is to make predictions with the continuous data where the gene expression values are not discretized. An advantage of this approach is that there is no information loss associated with discretization, while a disadvantage is the difficulty of considering the tractable and yet reasonably flexible transition probability density of stochastic dynamics.

The prediction results reported in this paper was on the expression time series data in the proliferation mode. One of our future research topics will be predictions of the gene regulatory network in the differentiation mode and will be reported elsewhere.

Acknowledgments We would like to thank the reviewers for the comments. A part of this research is supported by Grant-in-Aid for Scientific Research (20810034).

References

1) Friedman, N. and Goldszmidt, M.: Learning Bayesian networks with local structure, *Proc. 12th Conference on Uncertainty in Artificial Intelligence*, pp.252–262 (1996).
2) Friedman, N., Murphy, K. and Russell, S.: Learning the Structure of Dynamic Probabilistic Networks, *Proc. 14th Conference on Uncertainty in Artificial Intelligence* (*UAI 98*), pp.139–147 (1998).
3) Murphy, K. and Mian, S.: Modelling Gene Expression Data using Dynamic Bayesian Networks, Technical Report, Computer Science Division, University of California, Berkeley (1999).
4) Friedman, N., Linial, M., Nachman, I., and Pe'er, D.: Using Bayesian Networks to Analyze Expression Data, *J. Comput. Biol.*, Vol.7, pp.601–620 (2000).
5) Spellman, P., Sherlock, G., Zhang, M.Q., Iyer, V.R., Anders, K., Eisen, M.B., Brown, P.O., Botstein, D. and Futcher, B.: Comprehensive Identification of Cell Cycle-regulated Genes of the Yeast *Saccharomyces cerevisiae* by Microarray Hybridization, *Molecular Biology of the Cell*, Vol.9, pp.3273–3297 (1998).
6) Moler, E.J., Radisky, D.C. and Mian, I.S.: Integrating naive Bayes models and external knowledge to examine copper and iron homeostasis in *S. cerevisiae*, *Physiol. Genomic*, Vol.4, pp.127–135 (2000).
7) van Someren, E.P., Wesselsa, L.F.A. and Reinders, M.J.T.: Linear Modeling of Genetic Networks from Experimental Data, *Proc. 8th International Conference on*

Intelligent Systems for Molecular Biology, pp.355–366 (2000).
8) Eisen, M., Spellman, P., Brown, P. and Botstein, D.: Cluster Analysis and Display of Genome-wide Expression Patterns, *Proc. Nat. Acad. Sci. USA*, Vol.95, No.25, pp.14863–14868 (1998).
9) Zhou, X., Wang, X., Pal, R., Ivanov, I., Bittner, M. and Dougherty, E.: A Bayesian Connectivity-based Approach to Constructing Probabilistic Gene Regulatory Networks, *Bioinformatics*, Vol.20, No.17, pp.2918–2927 (2004).
10) Bittner, M., Meltzer, P., Chen, Y., Jiang, Y., Seftor, E., Hendrix, M., Radmacher, M., Simon, R., Yakhini, Z., Ben-Dor, A., Sampas, N., Dougherty, E., Wang, E., Marincola, F., Gooden, C., Lueders, J., Glatfelter, A., Pollock, P., Carpten, J., Gillanders, E., Leja, D., Dietrich, K., Beaudry, C., Berens, M., Alberts, D., Sondak, V., Hayward, N. and Trent, J.: Molecular Classification of Cutaneous Malignant Melanoma by Gene Expression Profiling, *Nature*, Vol.406, pp.536–540 (2000).
11) Kim, S.Y., Imoto, S. and Miyano, S.: Dynamic Bayesian Network and Nonparametric Regression for Nonlinear Modeling Gene Networks from Time Series Gene Expression Data, *Proc. 1st Computational Methods in Systems Biology*, pp.104–113 (2003).
12) Mao, L. and Resat, H.: Probabilistic Representation of Gene Regulatory Networks, *Bioinformatics*, Vol.20, No.14, pp.2258–2269 (2004).
13) Guet, C.C., Elowitz, M.B., Hsing, W. and Leibler, S.: Combinatorial Synthesis of Genetic Networks, *Science*, Vol.296, pp.1466–1470 (2002).
14) Middendorf, M., Kundaje, A., Wiggins, C., Freund, Y. and Leslie, C.: Predicting Genetic Regulatory Response using Classification, *Bioinformatics*, Vol.20, No.1, pp.232–240 (2004).
15) Beal, M.J., Falciani, F., Ghahramani, Z., Rangel, C. and Wild, D.L.: A Bayesian Approach to Reconstructing Genetic Regulatory Networks with Hidden Factors, *Bioinformatics*, Vol.21, No.3, pp.349–356 (2005).
16) Jeong, H., Tombor, B., Albert, R., Oltvai, Z.N. and Barabasi, A.L.: The large-scale organization of metabolic networks. *Nature*, Vol.407, pp.651–654 (2000).
17) Basso, K., Margolin, A., Stolovizky, G. and Klein, U.: Reverse Engineering of Regulatory Networks in Human B cells, *Nature Genetics*, Vol.37, pp.382–390 (2005).
18) Wildenhain, J. and Crampin, E.J.: Reconstructing Gene Regulatory Networks: From Random to Scale-free Connectivity, *IEE Proc. Systems Biology*, Vol.153, pp.247–256 (2006).
19) Chen, K., Wang, T., Tseng, H., Huang, C. and Kao, C.: A stochastic differential equation model for quantifying transcriptional regulatory network in *Saccharomyces cerevisiae*, *Bioinformatics*, Vol.21, No.12, pp.2883–2890 (2005).
20) http://www.research.att.com/˜njas/sequences/table?a=3024&fmt=5
21) Yook, S.H., Oltvai, Z.N. and Barabasi, A.L.: Functional and Topological Characterization of Protein Interaction Networks, *Proteomics*, Vol.4, pp.928–942 (2004).
22) Gelman, A., Carlin, J.B., Stern, H.S. and Rubin, D.: *Bayesian Data Analysis*, 2nd edition, Chapman and Hall (2004).
23) Reynolds, B.A. and Weiss, S.: Clonal and Population Analyses Demonstrate that an EGF-responsive Mammalian Embryonic CNS Precursor is a Stem Cell, *Dev. Biol.*, Vol.175, pp.1–13 (1996).
24) Seaberg, R.M. and van der Kooy, D.: Stem and Progenitor Cells: The Premature Desertion of Rigorous Definitions, *Trends Neurosci.*, Vol.26, pp.125–31. Review (2003).
25) Aoki, K., Sun, Y.J., Aoki, S., Wada, K. and Wada, E.: Cloning, Expression, and Mapping of a Gene that is Upregulated in Adipose Tissue of Mice Deficient in Bombesin Receptor Subtype-3, *Biochem. Biophys. Res. Commun.*, Vol.290, pp.1282–1288. (2002).
26) Matsui, T.: Transcriptional Regulation of a Purkinje Cell Specific Gene through a Functional Interaction between ROR Alpha and RAR, *Genes Cells*, Vol.2, pp.263–72 (1997).
27) Jetten, A.M., Kurebayashi, S. and Ueda, E.: The ROR nuclear orphan receptor subfamily: Critical regulators of multiple biological processes, *Prog. Nucleic Acid. Res. Mol. Biol.*, Vol.69, pp.205–47 (2001).
28) Mangelsdorf, D.J. and Evans, R.M.: The RXR Heterodimers and Orphan Receptors, *Cell*, Vol.83, pp.841–50 (1995).
29) Kliewer, S.A., Umesono, K., Heyman, R.A., Mangelsdorf, D.J., Dyck, J.A. and Evans, R.M.: Retinoid X receptor-COUP-TF Interactions Modulate Retinoic Acid Signaling, *Proc. Nat. Acad. Sci. USA*, Vol.89, pp.1448–1452 (1992).
30) Katagiri, Y., Takeda, K., Yu, Z.X., Ferrans, V.J., Ozato, K. and Guroff, G.: Modelling of Retinoid Signalling Through NGF-induced Nuclear Receptor Export of NGFI-B, *Nature Cell Biology*, Vol.2, pp.435–440 (July 2000).
31) Berrodin, T.J., Marks, M.S., Ozato, K., Linney, E., Lazar, M.A.: Heterodimerization among Thyroid Hormone Receptor, Retinoic Acid Receptor, Retinoid X Receptor, Chicken Ovalbumin Upstream Promoter Transcription Factor, and an Endogenous Liver Protein, *Mol. Endocrinol.*, Vol.6, pp.1468–1478 (1992).
32) Shi, Y., Sun, G., Zhao, C. and Stewart, R.: Neural Stem Cell Self-renewal, *Crit. Rev. Oncol. Hematol.*, 2007 Jul 20, Epub ahead of print (2007).
33) Kim, J.B., Ju, J.Y., Kim, J.H., Kim, T.Y., Yang, B.H., Lee, Y.S. and Son, H.: Dexamethasone Inhibits Proliferation of Adult Hippocampal Neurogenesis in vivo and in vitro, *Brain Res.*, Vol.1027, pp.1–10 (2004).
34) Heckerman, D.: Bayesian Networks for Data Mining, *Data Mining and Knowledge Discovery*, Vol.1, pp.79–119 (1997).
35) Nuclear Receptors Nomenclature Committee: A unified nomenclature system for the nuclear receptor superfamily, *Cell*, Vol.97, No.2, pp.161–163 (1999).
36) Zhang, Z., Burch, P.E., Cooney, A.J., Lanz, R.B., Pereira, F.A., Wu, J., Gibbs, R.A., Weinstock, G. and Wheeler, D.A.: Genomic analysis of the nuclear receptor family: New insights into structure, regulation, and evolution from the rat genome,

Genome Res., Vo.14, No.4, pp.580–590 (2004).

(Received October 22, 2009)
(Accepted December 28, 2009)
(Released March 15, 2010)

(Communicated by *Jun Sese*)

Yusuke Kitamura
Biography and photo were not available at the time of the publication.

Tomomi Kimiwada was born in 1975. She received her M.D. degree in 2001, and Ph.D. degree in 2007 from Tohoku University, Japan, respectively. She has been working in Tohoku University Hospital since 2001 and now is working in Children's Hospital of Miyagi as a neurosurgeon. Since 2005 until 2007 she had been a visiting researcher of NCNP, Japan. She has been engaging in the research areas of neuroscience (neurogenesis, nuclear receptors, clock genes and epilepsy). She is a member of JNS, JES, ESSJ and JSS.

Jun Maruyama
Biography and photo were not available at the time of the publication.

Takashi Kaburagi was born in 1980. He received his BSEE, ME and Ph.D. degrees from Waseda University, Japan in 2003, 2005, and 2009 respectively. He is a Research Associate at the Department of Electrical Engineering and Bioscience, Waseda University. His current research interests are transmembrane protein structure predictions, gene regulatory netwrok structure predictions, and information efficient text input systems. He is a member of the IPSJ, JSBi, and ISCB.

Takashi Matsumoto received his BSEE from Waseda University, Tokyo, Japan, MS in Applied Mathematics from Harvard University, Cambridge, Massachusetts, and Ph.D. from Waseda University. Currently he is a professor at the Faculty of Advanced Science and Engineering, Waseda University. He has had visiting positions at U.C. Berkeley (1977–1979) and Cambridge University, U.K. (2003–2004). His research interests are in Bayesian machine learning with Monte Carlo implementations. Target data includes biological data, image data, time series data among others. He also has interests in non-parameteric priors for Bayesian leaning. He is a member of IPSJ, IEEE (fellow), among others.

Keiji Wada was born in 1955. He received his M.D. from Osaka University Medical School, Japan in 1979 and Ph.D. from Graduate School of Medicine, Osaka University in 1984, respectively. He has been working in National Institute of Neuroscience, National Center of Neurology and Psychiatry. Since 1985 until 1988 he had been a Postdoctoral Fellow of Molecular Neurobiology Laboratory, The Salk Institute in San Diego, California, USA. Since 1988 until 1992 he had been a Visiting Scientist in the Section on Neurotransmitter Receptor Biology, Laboratory of Neurochemistry, National Institute on Deafness and Other Communication Disorders, NIH, USA. Since 1992 until present, he has been a Director of Department of Degenerative Neurological Diseases, National Institute of Neuroscience, National Center of Neurology and Psychiatry. He has been engaging in the research area of molecular neuroscience. He is a member of Society for Neuroscience, Japan Neuroscience Society, The Japanese Society for Neurochemistry and Societas Neurologica Japonica. He is an editorial board member of Neuroscience Research, Neurochemistry International and Glia.