

# Inference of Gene Regulatory Networks Using Time-Series Data: A Survey

Chao Sima\*, Jianping Hua and Sungwon Jung

Computational Biology Division, Translational Genomics Research Institute, Phoenix, AZ 85004, USA

**Abstract:** The advent of high-throughput technology like microarrays has provided the platform for studying how different cellular components work together, thus created an enormous interest in mathematically modeling biological network, particularly gene regulatory network (GRN). Of particular interest is the modeling and inference on time-series data, which capture a more thorough picture of the system than non-temporal data do. We have given an extensive review of methodologies that have been used on time-series data. In realizing that validation is an impartible part of the inference paradigm, we have also presented a discussion on the principles and challenges in performance evaluation of different methods. This survey gives a panoramic view on these topics, with anticipation that the readers will be inspired to improve and/or expand GRN inference and validation tool repository.

Received on: December 04, 2008 - Revised on: February 28, 2009 - Accepted on: March 02, 2009

## 1. INTRODUCTION

Biological system has been traditionally studied *via* reductionism approach, that is, explaining cell behaviors by studying functions of individual cellular components. Though the knowledge being insightful, it has been increasingly apparent that the understanding of the complex cellular system requires understanding of how different components work together. The advent of high-throughput technology like microarrays, where cellular activities can be measured at genome-wide scale, has provided just this platform and thus created an enormous interest in mathematically modeling biological network, particularly gene regulatory network (GRN). The goal is to mimic the biological network in some abstract level, and a better understanding of the underlying biological system could be achieved through the analysis on the resulted mathematical model. To do so, it is critical to have a reliable modeling and inference procedure.

Amid the explosion of efforts in inferring biological network models that has been witnessed by the last decade, there have been accordingly a number of reviews on these different methodologies. Some of these reviews are focused on one specific type of modeling, for example, on Bayesian Network and Dynamic Bayesian Network (graphical model) [1, 2], or topologies (random *v.s.* scale-free *v.s.* hierarchical) [3]. Others describe models in different categories: Cho *et al.* reviewed methods that incorporate prior knowledge as well as machine learning approaches (e.g., Genetic Algorithm or Neural Net) [4], in addition to the general methods we reviewed in this paper; Schlitt *et al.* ordered the modeling methods according to their complexity [5], from simple ones that are just an aggregation of different functional components, to those that are trying to model the dynamics and have a myriad of parameters to tune. van Someren *et al.*

took a unique approach, listing the models in a chronological order along with pointing out the key differences between them [6]. Lastly there is a good review by de Jong [7] which provided a concise mathematical background and formulation for many widely used models.

A lot of those early efforts have been focused on non-temporal data, largely due to the paucity of time-series data. As the biological system is inherently complex and GRN is essentially an ensemble of genes which evolve over time, time-series data will clearly capture a more thorough picture of the system than non-temporal data do, which only take *snapshots* of the system at one given time point. Table 1 lists some of the time-series data sets that have been used. (Bar-Joseph has a general review of gene expression time-series data [26]). It is therefore our interest to present a survey of study on models and inference methodologies for time-series data, its promises and unique challenges. In addition to our focus of time-series data inference, there are two other distinctive aspects of our review work that are different from others: (1) we will take a practitioner's view, reviewing different implementations within one modeling framework as well as different modeling frameworks; (2) we have a special interest in performance evaluation: the validation of the methodologies and comparison studies where it is possible. It is our hope that at the end of paper, readers will have a rather complete knowledge of the methods in the literature, and be inspired to improve and/or expand this tool repository.

This paper is organized as follows: modeling and inference methods focusing on inferring the structure of the network will be reviewed in Section 2, which include the *Relevance Network*, *Bayesian Network* and *Dynamic Bayesian Network* (DBN). This is followed by Section 3 in which methods inferring both structure and dynamics will be reviewed. Among them, *Boolean Network* and *Probabilistic Boolean Network* (PBN) have their states in discrete space and *Markov Model* (MM), *State Space Model* (SSM), and *Ordinary Differential Equation* (ODE) typically have their states in continuous space. The prominent characteristics of

\*Address correspondence to this author at the Computational Biology Division, Translational Genomics Research Institute, Phoenix, AZ 85004, USA; Tel: 1(602)343-8485; Fax: 1(602)343-8740; E-mail: csima@tgen.org

**Table 1. Time Series Data Sets Used to Infer GRNs**

Organism	Brief Description	Ref.	Time Points	Year of Publication
B. subtilis	Expression in MMGE environment	[8]	8	1995
Yeast	Diauxie shift; response to medium growth conditions	[9]	7	1997
Yeast	Cell cycle synchronization	[10]	18,24,14*	1998
Yeast	Cell cycle synchronization	[11]	17	1998
Yeast	Yeast Sporulation	[12]	7	1998
House Mouse	Development of the central nervous system of rats	[13]	8	1998
Human	Human Fibroblasts	[14]	13	1999
Yeast	Yeast meiotic expression	[15]	9	2000
E. coli	Tryptophan rich, starving	[16]	8	2000
Drosophila	Life cycle	[17]	74	2002
Human	T-cell activation	[18]	10	2004
Synechocystis	Light intensity experiment	[19]	47,27*	2004
Rat	Circadian rhythmicity of gene expression	[20]	12	2005
B. subtilis	During feed-batch protease production process	[21]	20	2005
E. coli	Perturbation of the SOS system	[22]	6	2006
Human	Endothelial cell apoptosis in blood vessel	[23]	7	2007
Mouse	IL-2-stimulated immune response	[24]	12	2007
Human	Migration of skin keratinocyte	[25]	7	2008

\*multiple data sets.

these models are shown in Fig. (1). After a discussion on performance evaluation of the models in Section 4, we conclude the paper in Section 5.

We need to point out that the placement of DBN in Section 2 is due to its close kinship with *Bayesian Network* and we feel it is more natural to introduce them together. It would have well fallen into Section 3 otherwise. In fact DBN was shown to be related to a lot of models in Section 3: Lähdesmäki *et al.* showed that variables in both discrete-valued DBN and PBN can have the same joint probability distribution [27]; DBN can also be considered as a generalization of Boolean Network [28], or SSM [29].

## 2. INFERRING STRUCTURE OF THE NETWORK

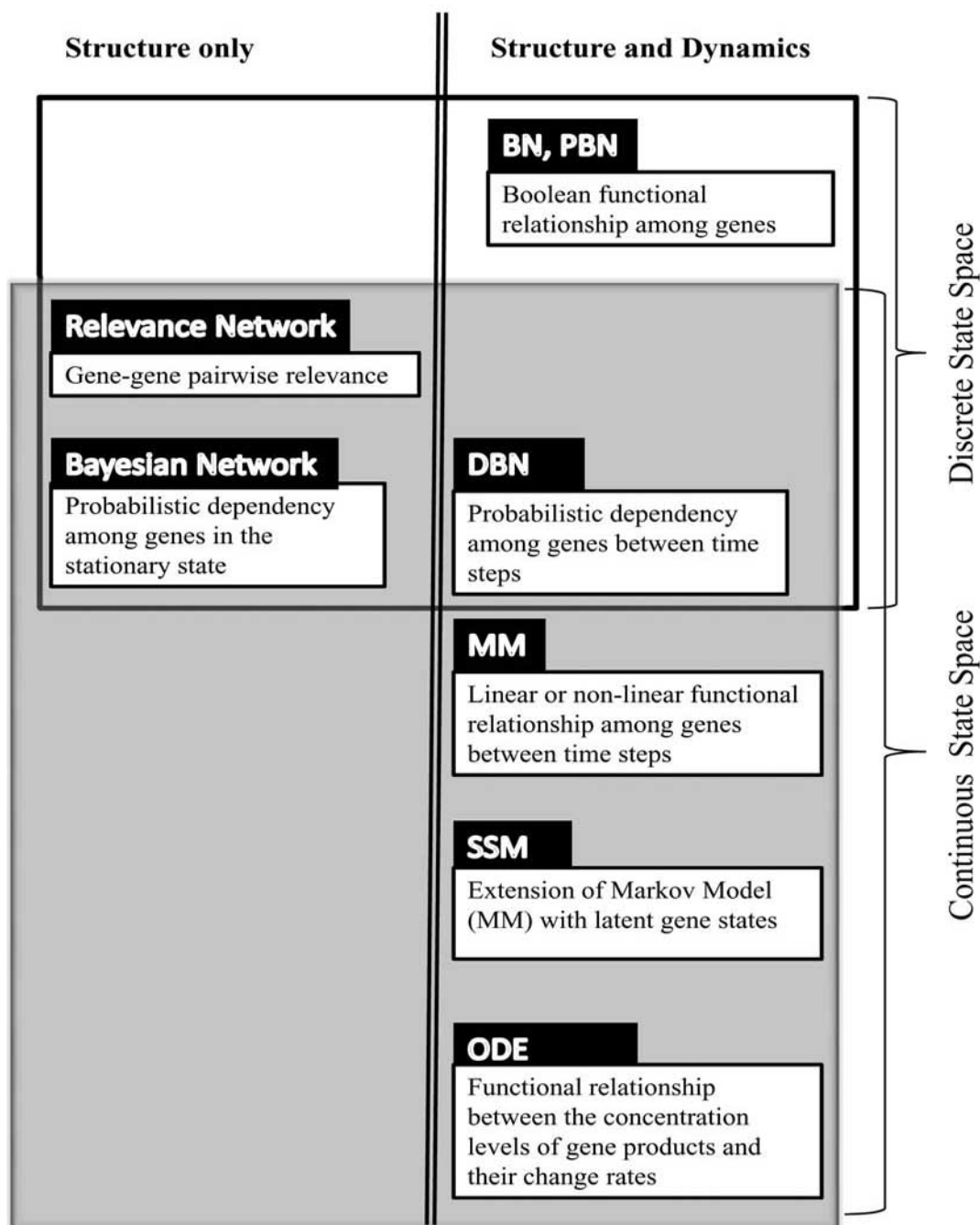
Given a set of genes as nodes for a network, the structure is basically the assemble of all the interconnections among the nodes. Depending on the models, these connections usually take on different meanings, but generally specify the relationships between one gene with another gene, or another set of genes. *Relevance Network*, argued by some to be *ad hoc* network, is the simplest model in this category.

### 2.1. Relevance Network

Relevance networks can be categorized as networks in which relationship among genes can be defined using a pairwise measure of relevance. Given a set of  $n$  genes

$\mathbf{G} = \{g_1, g_2, \dots, g_n\}$  and a set of observations  $\mathbf{D}$  on genomic profiling (e.g., gene expression) for  $m$  time points, a relevance between  $g_i$  and  $g_j$  can be evaluated using their time-series profiles,  $[g_{i,1} \ g_{i,2} \ \dots \ g_{i,m}]$  and  $[g_{j,1} \ g_{j,2} \ \dots \ g_{j,m}]$ . Various relevance measures have been used to infer relationships between two genes, from simple correlation measures to biologically motivated relevance measures.

One example of using a simple Pearson correlation measure is the work of Remondini *et al.*, where they used two sets of gene expression data from rat fibroblast cell lines to construct correlation-based networks [30, 31]. Even though such a correlation measure can be useful in many cases, it cannot provide causal information. To overcome this limitation, Gupta *et al.* used a slope metric (SR) to elucidate not only the presence of a relationship between two genes but also its directionality [32], which is used to represent causality. In their study, the structure was determined using a correlation measure with a preferred threshold. The directionality of relevance was determined using the following SR, based on the assumption that a gene  $g_i$  is linearly dependent on another gene  $g_j$ , i.e.,  $g_i = a_{ij} + b_{ij}g_j$ .



**Fig. (1).** A summary of the models in the paper. Some abbreviations: BN - Boolean Network; PBN - Probabilistic Boolean Network; DBN - Dynamic Bayesian Network; MM - Markov Model; SSM - State Space Model and ODE - Ordinary Differential Equation.

$$SR = \frac{\min(|b_{ij}|, |b_{ji}|)}{\max(|b_{ij}|, |b_{ji}|)} \quad (1)$$

If  $SR = |b_{ij}| / |b_{ji}|$ , then the directionality is determined as  $g_i \rightarrow g_j$  and if  $SR = |b_{ji}| / |b_{ij}|$ , then it becomes  $g_j \rightarrow g_i$ . This interpretation of the directed edge between two genes is

based on the biological assumption that a small change in the source gene is associated with a large change in the target gene. This method was applied to a time-series microarray data of *B. subtilis* [21] and showed some correspondence with already known biological information.

Besides the approaches based on relevance measures for the same time point, there are approaches to incorporate

time-delay in measuring the relevance between genes. Schmitt *et al.* used a time-lagged correlation to infer gene regulatory networks [19]. For a transcription profile represented by a series of  $m$  measurements taken at equally spaced time points, the correlation between genes  $g_i$  and  $g_j$  with a time lag,  $\tau$ , is  $\mathbf{R}(\tau) = r_{ij}(\tau)$ , defined by

$$S_{ij} = \sum_{t=\tau+1}^m (g_{i,t} - \bar{g}_i)(g_{j,t-\tau} - \bar{g}_j) \quad (2)$$

$$r_{ij}(\tau) = \frac{S_{ij}(\tau)}{\sqrt{S_{ii}(\tau) S_{jj}(\tau)}} \quad (3)$$

where  $g_{i,t}$  denotes the expression of a gene  $g_i$  at time  $t$ ,  $\bar{g}_i$  is the averaged expression value of a gene  $g_i$  across all time points, and  $S_{ij}$  is essentially the inner product between the time-shifted profiles. This measure was used to analyze whole genome DNA microarrays of *Synechocystis* under various light intensity conditions [19]. Seed genes were selected first based on the time-lagged correlation between the expression profile and the light intensity profile across the time points. These seed genes were expanded using the time-lagged correlation between the average expression profile of seeds and the expression profile of each gene.

Ma *et al.* and Barker *et al.* used a biologically motivated approach to measure the dependency between two genes using temporal information in time-series profiles [33-35]. After ternarizing the expression data into highly/averagely/lowly expressed states (denoted by H/A/L respectively), Ma *et al.* looked for dependencies between  $g_i$  being 'H' at one time and  $g_j$  also being 'H' at next time point [33, 34]. Specifically, they computed a statistics on the difference between the number of occurrences of  $g_i$  being 'H' at one time AND  $g_j$  also being 'H' at next time point, and the expected number of these occurrences. The latter was computed by averaging the product of the number of  $g_i$  being 'H' at one time (regardless of  $g_j$ 's state at next time point) and number of  $g_j$  being 'H' at next time point (regardless of  $g_i$ 's state at previous time point). A 95 percent confidence level was then used to determine whether the dependency is deemed significant. Using this method, the authors constructed a gene interaction diagram from a yeast data set [10]. Similarly, Barker *et al.* proposed a method to determine whether  $g_i$  is regulated by  $g_j$ , based on three ratios of samples, the ratio of  $g_j$  activating  $g_i$ , the ratio of  $g_j$  repressing  $g_i$  and the ratio of  $g_j$  doing nothing on  $g_i$  [35]. With threshold values to those three ratios, the final structure of a GRN is constructed. These methods investigate the relationship between two time-series profiles with single epoch-delay, thus assuming Markov condition regarding the time-point, where an observation at some time point is dependent on that of the previous time-point only.

Kwon *et al.* proposed a method based on string alignment to infer transcriptional regulation relationships from time-series gene expression data, in which relevance with arbitrary time-delay can be considered [36]. They converted the time-course of each gene into a string composed of rising(R), constant(C) or falling(F). The similarity between two event strings is evaluated by string alignment algorithm using scoring matrix that describes similarity score between each pair of event characters. This method was applied to a yeast cell-cycle data [10] and showed it could find some already known transcriptional regulation relationships.

There were also efforts to enhance the quality of inferred relevance network through post processing using additional criteria. Bickel used decisive false discovery rate (dFDR) to estimate the probability of spurious connections between genes in GRNs [37]. After building a network using a time-lagged correlation measure, three kinds of probabilities were evaluated for each edge, i) the probability of being a false positive connection, ii) the probability of being a connection with wrong time order and iii) the probability of being a connection with a time-delay while there is actually no time-delay. This method was applied to yeast cell-cycle data [10] and showed it could successfully find already known genetic relationships.

## 2.2. Bayesian Network

A Bayesian network model is a graphical representation of a joint probability distribution of random variables. A Bayesian network  $B$  is defined as a tuple of  $(G, \Phi)$ , where  $G$  is a graph structure that represents conditional dependency relationships between random variables and  $\Phi$  is a set of parameters describing conditional probability distribution. In modeling GRNs,  $G$  corresponds to the topology of a GRN, where each node represents a gene as a random variable and each edge represents dependency between genes. With Markov assumption on dependency relationships, the joint probability distribution of genes  $\mathbf{G} = \{g_1, g_2, \dots, g_n\}$  is described as follows:

$$P(g_1, g_2, \dots, g_n) = \prod_{i=1}^n P(g_i | \mathbf{Pa}(g_i), \theta) \quad (4)$$

where  $\mathbf{Pa}(g_i)$  is a set of parents of  $g_i$  in  $G$  and  $\theta$  is a statistics from  $\mathbf{D}$ . In this framework, a Bayesian network represents a static joint probability distribution of a set of random variables. For this reason, most of applications of Bayesian networks do not incorporate the temporal information in time-series data. Usual assumption in using Bayesian networks for time-series data is that the time-series is from the stationary state of the target biological system. One may use the temporal information to determine the direction of edges in Bayesian networks. However, it is important to notice that the direction of an edge in Bayesian networks does not necessarily represent causality between random variables.

Learning a Bayesian network  $B = (G, \Phi)$  from observed data  $\mathbf{D}$  implies learning the dependency structure  $G$  and learning the set of probabilistic parameters  $\Phi$ . Learning  $\Phi$  is a relatively easy problem once  $G$  and  $\mathbf{D}$  are given, and

learning  $G$  can be done by finding  $G^*$  with maximum  $P(G|\mathbf{D})$ . However, learning  $G$  given  $\mathbf{D}$  is a hard problem because the number of possible graph structures increase exponentially as the size of a network increases. For this reason, most of approaches using Bayesian networks focus on small problems or take heuristics to handle large problems. One popular heuristic approach is restricting the  $G$  to a certain category.

One of the first applications of Bayesian networks for genetic networks is the work by Friedman *et al.* with the strategy of restricting  $G$  [38]. In this study, learning Bayesian networks was applied to microarray gene expression data for cell-cycle of *S. cerevisiae* [10]. By using Bayesian network learning, the authors analyzed gene expression data for Markov relation between genes and the coverage of influence for each gene, where the coverage of influence was measured by counting the number of descendants in the graph structure for that gene. They used sparse candidate algorithm, which restricts the candidate parents of each node in  $G$  during the search. In their study, 800 genes related to cell-cycle were considered as random variables. Their result was evaluated through literature mining and statistical significance test. Even though they used time-series expression data, temporal information was not used in their study due to the previously mentioned reason.

Considering the problem of learning Bayesian networks as an optimization problem for the objective function  $P(G|\mathbf{D})$ , search algorithms for large solution spaces can be also used. One of such methods is an estimation of distribution algorithm (EDA) and Dai *et al.* used the EDA for learning genetic networks with a Bayesian network model [39]. With EDA, they evolved a population of  $G$ s that have high Bayesian information criterion (BIC) scores. After some iterated evolution process,  $k$  graph structures with highest scores were chosen to build final aggregated genetic networks. Their method was applied to two sets of time-series data [10, 11], and was evaluated by Gene Ontology (GO) search and literature mining.

### 2.3. Dynamic Bayesian Network

A Dynamic Bayesian Network (DBN) is an extension of a Bayesian network model to incorporate temporal concept. Compared to conventional Bayesian networks, DBNs include random variables  $\{g_{1,t-1}, g_{2,t-1}, \dots, g_{n,t-1}\}$  of time step  $t-1$  in addition to  $\{g_{1,t}, g_{2,t}, \dots, g_{n,t}\}$  of time step  $t$ . A transition network is composed of those  $2n$  random variables with no edges from time step  $t$  to  $t-1$ . It is assumed that the transition probability  $P(\vec{g}_t | \vec{g}_{t-1})$ , where  $\vec{g}_t$  represents the values of  $n$  genes at time  $t$ , is homogeneous across entire observation. Learning DBNs can be done using the same idea of learning Bayesian networks. The only difference is that we need to consider additional random variables of time  $t-1$ . From this perspective, Friedman *et al.* extended some scoring rules for learning structures from Bayesian networks to the case of DBNs [40].

The difficulty in learning DBNs is its heavy computational complexity. Because additional  $n$  random variables are considered compared to conventional Bayesian networks, learning algorithms should consider much more candidate graph structures and probabilistic parameters. For this reason, most of applications of DBN usually target smaller systems compared to the study of Bayesian networks. Further, heuristics to restrict candidate graph structures are widely used in the applications of DBNs.

Ong *et al.* used DBNs to infer regulatory network for tryptophan metabolism in *E. coli* [41]. By using prior knowledge (operon map), they restricted possible network structures into predetermined category. Then a DBN was learned using Expectation-Maximization (EM) method with gene expression data of 8 time points [16]. Missal *et al.* used mutual information between two gene expression profiles and applied  $\chi^2$ -test for the significance of the mutual information to determine the structure of a DBN [42]. Zhao *et al.* took a similar approach of using mutual information [43], but they used minimum description length (MDL) principle [44] to determine the threshold of the significance, and to remove indirect or false links in a post-pruning process. However, by using different encoding schemes, this method can generate non-unique results that need *ad hoc* adjustment. Dougherty *et al.* overcome this drawback by measuring the description length based on a universal model: normalized maximum likelihood model [45]. Zou *et al.* used DBNs with various time-delay, by shifting time-series profiles with properly predicted amount of time steps [46], and applied their method to the yeast cell-cycle data of Chou *et al.* [11].

Variables in DBNs can take continuous values as well as discrete values. When random variables are continuous in a DBN, conditional probability tables, which are used in the discrete case, cannot be used. To model  $P(g_{i,t} | \mathbf{Pa}(g_{i,t}))$  in a continuous domain, Kim *et al.* assumed a nonparametric additive regression model with Gaussian noise [47-49],

$$g_{i,t} = m_{i1}p_{1,t-1} + \dots + m_{iq}p_{iq,t-1} + \varepsilon_i(t) \quad (5)$$

where  $q$  is the number of parents of  $g_{i,t}$ ,  $p_{ij}$  is the  $j$ th parent of  $g_{i,t}$  and  $\varepsilon_i(t)$  is a Gaussian noise of  $g_i$  at time  $t$ . A scoring measure for DBN structures was proposed based on the regression model. This method was applied to yeast cell-cycle data of Spellman *et al.* [10]. A nonlinear regression extension was proposed by Ferrazzi *et al.* in which  $p_{ij,t-1}$  in Eqn. (5) is replaced by  $\tanh(p_{ij,t-1})$ , where  $\tanh(\cdot)$  is the hyperbolic tangent function [50].

Several recent studies focused on using different types of gene expression data. Dojer *et al.* proposed a method to handle perturbed gene expression data in using DBNs [51]. In their approach, candidate regulators for each gene were inferred from only a subset of entire data, where the target gene was not perturbed. The motivation of this approach is based on the assumption that data with a specific gene perturbed may not be used in the process of inferring

regulators of that gene, because knocking out a gene can disable regulations *toward* that gene. But since knocking out a gene may represent under-expression of that gene, the targets of a perturbed gene may still have regulation effect *from* the perturbed gene. The effectiveness of this method was shown using simulated data from an ordinary differential equation model. Lähdesmäki *et al.* proposed a method for learning DBNs from mixture of steady state data and time-series data [52]. If a steady state data is  $\mathbf{D}_A$  and a time-series data is  $\mathbf{D}_\pi$ , learning a DBN structure  $G$  requires the evaluation of the marginal likelihood  $P(\mathbf{D}_A, \mathbf{D}_\pi | G)$ . By assuming  $\mathbf{D}_A$  and  $\mathbf{D}_\pi$  are independent given  $(G, \Phi)$ , where  $\Phi$  is a set of probabilistic parameters, the marginal likelihood can be evaluated as follows:

$$P(\mathbf{D}_A, \mathbf{D}_\pi | G) = \int_{\Phi} P(\mathbf{D}_A, \mathbf{D}_\pi | G, \Phi) P(\Phi | G) d\Phi \quad (6)$$

$$= \int_{\Phi} P(\mathbf{D}_A | G, \Phi) P(\mathbf{D}_\pi | G, \Phi) P(\Phi | G) d\Phi \quad (7)$$

Evaluation of  $P(\mathbf{D}_A | G, \Phi)$  can be done in the same way of static Bayesian networks and the evaluation of  $P(\mathbf{D}_\pi | G, \Phi)$  is done using the steady-state distribution of the DBN.

### 3. INFERRING STRUCTURE AND DYNAMICS

Structure alone does not completely describe the network. Often we are interested in the evolution of the system from a given condition, or the response to a particular perturbation, which require a network model that is able to characterize the dynamics and describe the system transitions into future time. The states, defined as the values for the vector of genes, can be either in continuous domain or constrained to be in discrete space.

#### 3.1. Discrete State Space

A discrete state space model characterizes a system using quantized data. The most popular approaches are Boolean network and probabilistic Boolean network (PBN).

Boolean network assumes that the gene expression takes just two levels: ON/1 and OFF/0, and the functional relationship between the genes is determined by logical rules. A Boolean network consists of  $n$  nodes  $\mathbf{G} = \{g_1, g_2, \dots, g_n\}$  and a list of Boolean functions  $\mathbf{F} = \{f_1, f_2, \dots, f_n\}$ . Each node  $g_i \in \{0, 1\}$  is a binary variable that represents the state (expression) of gene  $i$ . The Boolean function  $g_{i,t+1} = f_i(\mathbf{Pa}(g_{i,t}))$  specifies how the value of node  $g_i$  at next time point  $t+1$  is determined by the values of its input nodes  $\mathbf{Pa}(g_i)$  at current time point  $t$ .

REVEAL [53] proposed by Liang *et al.* is one of the first Boolean-network-based inference scheme. Based on information theory, if the mutual information of the input and output is equal to the entropy of the output, the input fully determines the output. Hence for each node, REVEAL searches for the minimal input node set that can fully

determine the output. Rather than using mutual information, Akutsu *et al.* chose to check consistency for all logic rules [54], and later extended the inference to noisy data [55]. Furthermore, Akutsu *et al.* proved that, if the maximum number of input nodes is bounded by a constant and transition pairs are randomly picked, the sample size needed to reconstruct the original Boolean network is in  $O(\log n)$  [54]. However, for limited sample size, there exist considerable number of valid networks that are consistent with the given data. Rather than pick one or a few networks, Martin *et al.* enumerate all the consistent networks and aggregate them by network's attractors [24]. Their results show that most networks share only a few common attractors, which indicate similar network dynamics.

Some view the binary-state system like Boolean network as oversimplification that has significant information loss. Laubenbacher and Stigler introduced a multi-state system where the gene-gene relationship is determined by the computational algebra of the finite field [56]. Normally the solution is not unique, and the inference scheme will pick the minimal network functions by removing all redundant terms. Like REVEAL, this approach assumes noiseless data and the performance suffers when noise presents.

The Boolean network and the multi-state extension are based on the assumption of deterministic gene-gene relationship. Probabilistic Boolean network adds a sense of randomness into the Boolean network by allowing the nodes to have more than one associated Boolean functions. So for a PBN, the nodes  $\mathbf{G}$ 's associated functions are now denoted as  $\mathbf{F} = \{\mathbf{F}_1, \mathbf{F}_2, \dots, \mathbf{F}_n\}$ , where  $\mathbf{F}_i = \{f_1^{(i)}, f_2^{(i)}, \dots, f_{l_i}^{(i)}\}$ , i.e., each node's output is now associated with  $l_i$  possible Boolean functions. At any time point, PBN allows each node  $i$  to take only one Boolean function from  $\mathbf{F}_i$ . Hence there are altogether  $\prod_{i=1}^n l_i$  realizations of a PBN. Perturbation probability and selection probability were later introduced to allow the network be perturbed in current realization or switched between realizations, respectively.

It was suggested in Shmulevich *et al.* that *Coefficient of Determination* can be used to infer the network [57]. Later Marshall *et al.* implemented an inference procedure for PBN that successfully infers all the constituent Boolean networks, as well as all the perturbation and selection probabilities associated with them [58]. Unfortunately, as they pointed out in the paper, the amount of temporal data needed for inference is huge. A more practical way to infer PBN is to approximate the network by multivariate Markov model, as shown by Ching *et al.* [59]. In this model, the state of gene  $i$  at time point  $t$  takes a binary probability distribution denoted by vector  $\vec{p}_{i,t} = [P(g_{i,t} = 0), P(g_{i,t} = 1)]^T$ . The model assumes

$$\vec{p}_{i,t+1} = \sum_{j=1}^n \lambda_{ij} T_{ij} \vec{p}_{j,t}, \quad (8)$$

where  $T_{ij}$  is the probability transition matrix from gene  $j$  to gene  $i$ , and  $\lambda_{ij}$  the non-negative weight factor that has  $\sum_{j=1}^n \lambda_{ij} = 1$ .

### 3.2. Continuous State Space

A continuous state space model can characterize a system without discretizing the data, a step argued to result loss of information. One of the straightforward modeling strategies here is to describe the system evolution as a *Markov Process*, and this gives the *Markov Model* networks.

#### 3.2.1. Markov Model

Dewey *et al.* studied a simple linear Markov model in the form of

$$\vec{g}_{t+1} = T\vec{g}_t \quad (9)$$

in which  $\vec{g}_t$  denotes the gene expression levels at time  $t$ , and the matrix  $T$  is the transition for genes between two time points [60, 61]. Assuming a total of  $m$  time points, this model can be written in another form

$$G_{+1} = TG \quad (10)$$

where  $G_{+1} = [\vec{g}_2 \ \vec{g}_3 \ \dots \ \vec{g}_m]$  and  $G = [\vec{g}_1 \ \vec{g}_2 \ \dots \ \vec{g}_{m-1}]$ , and the transition matrix  $T$  can be calculated as  $G_{+1}G^+$ ,  $G^+$  being the pseudo-inverse of  $G$  obtained by using *singular value decomposition* (SVD). They further extended it to include non-linear terms that capture both between-time (contained in  $GG^T$ ) and between-gene (contained in  $G^TG$ ) correlations by considering the form  $G_{+1} = T_1G + T_2G^TG + GG^TT_3$ , as well as higher order Markov terms [60].

Holter *et al.* performed SVD [62-64] on the data  $D = [\vec{g}_1 \ \vec{g}_2 \ \dots \ \vec{g}_m] = U\Sigma V^T$  and took the first  $r$  rows in  $\Sigma V^T$ ,  $r$  being the number of non-zero eigenvalues of  $DD^T$ , as the *dominant patterns*, or *modes* [65]. The temporal expression for each gene is therefore a linear combination of these *modes*  $\vec{X}_i = [x_{i,1} \ x_{i,2} \ \dots \ x_{i,m}]$ ,  $i = 1, \dots, r$ . The linear Markov model they considered is in the form of

$$\begin{pmatrix} x_{1,t+1} \\ x_{2,t+1} \\ \vdots \\ x_{r,t+1} \end{pmatrix} = T \begin{pmatrix} x_{1,t} \\ x_{2,t} \\ \vdots \\ x_{r,t} \end{pmatrix} \quad (11)$$

The *modes* transition matrix  $T$  is then estimated by minimizing the divergence of the trajectory from the observed values.

Wiggins *et al.* took a Bayesian approach and rewrote the transition dynamics with added term of Gaussian noise  $\vec{\epsilon}_t$  as [66]:

$$\vec{g}_{t+1} - \vec{g}_t = (T - I)\vec{g}_t + \vec{\epsilon}_t \quad (12)$$

where  $I$  is the identity matrix. Given the biological prior which was modeled in  $T$ , they were able to derive the posterior probability for  $T$  and used the expectation of  $T$  to represent the transition dynamics. They also considered an augmented model which treated latent variables as additional "hidden degrees of freedom" and derived the expected  $T$  after integrating out these latent variable.

Another variation of Markov model was studied by Li *et al.* in which the transition of gene expression  $g_{i,t}$  is modeled as power functions:

$$g_{i,t+1} = \alpha_i \prod_j g_{j,t}^{w_{ij}} + (1 - \beta_i)g_{i,t} + \epsilon_{i,t} \quad (13)$$

where  $\alpha_i$  is the rate of transcription and  $\beta_i$  is the rate of degradation [67, 68]. To find the most-likely structure of the network, the authors used either guided simulated annealing method [67, 69] or Genetic Algorithm [68] to optimize a score function which is essentially the likelihood function with a penalty term that penalizes complex models (complex in the sense that number of parameters being too large). Used on a set of yeast *Saccharomyces Cerevisiae* microarray data [10], this procedure recovered 31 out of 42 and 17 out of 22 regulation relationships which are consistent with those found experimentally [67, 68].

#### 3.2.2. State Space Models

State Space Model (SSM) can be viewed as an extension to Markov models, based on the assumption that gene expression levels are hidden states and cannot be directly observed, and are related with the observed values by some transformation. A general linear *SSM with input* takes the form as:

$$\vec{g}_{t+1} = T\vec{g}_t + A\vec{u}_t + \vec{\epsilon}_{g,t} \quad (14)$$

$$\vec{y}_t = C\vec{g}_t + B\vec{v}_t + \vec{\epsilon}_{y,t} \quad (15)$$

where  $\vec{g}$  denotes the state of the genes for the system,  $\vec{y}$  the observed data for  $\vec{g}$ ,  $T$  the state transition matrix,  $C$  the state to observation matrix, and  $A$  and  $B$  are the inputs influence matrices for inputs  $\vec{u}_t$  and  $\vec{v}_t$ , respectively. Here  $\vec{\epsilon}_{g,t}$  and  $\vec{\epsilon}_{y,t}$  are white noise terms. If  $A = 0$  and  $B = 0$ , this is reduced to the *basic linear SSM*, or *standard SSM*.

Rangel *et al.* and Beal *et al.* used *SSM with input* and set the inputs as the observations from previous time point [18, 70]. Specifically, their system is described as:

$$\vec{g}_{t+1} = T\vec{g}_t + A\vec{y}_t + \vec{\epsilon}_{g,t} \quad (16)$$

$$\vec{y}_t = C\vec{g}_t + B\vec{y}_{t-1} + \vec{\epsilon}_{y,t} \quad (17)$$

Notice the above equation can be rewritten as

$$\vec{y}_t = C(T\vec{g}_{t-1} + A\vec{y}_{t-1} + \vec{\epsilon}_{g,t-1}) + B\vec{y}_{t-1} + \vec{\epsilon}_{y,t} \quad (18)$$

$$= (CA+B)\vec{y}_{t-1} + C(T\vec{g}_{t-1} + C\vec{\epsilon}_{g,t-1}) + \vec{\epsilon}_{y,t} \quad (19)$$

Therefore the matrix  $T' = CA + B$  captures the transition in the *observation* domain over time, through the hidden states

$\vec{g}_t$ , and the authors focused their interests in  $T'$ . Rangel *et al.* estimated the model parameters using EM algorithm and constructed confidence interval on  $T'$  by using bootstrap, while Beal *et al.* used what they called *Variational Bayesian EM Algorithm*, which can be considered as a Bayesian extension of the standard EM algorithm, to derive a posterior estimation on  $T'$ .

Both Hirose *et al.* and Yoshida *et al.* argued that the dimension  $k$  of the hidden states  $\vec{g}_t$ , which they called “modules”, is less than the number of dimension for observations  $\vec{y}_t$  [71, 72]. Using *standard SSM* and assuming the noise term  $\vec{\epsilon}_{y,t}$  has a diagonal covariance matrix  $R$  and some other constraints, the authors carefully designed a projection transformation matrix  $H$  such that the denoised observation vector  $\vec{y}_t^\diamond = R^{-1/2}(\vec{y}_t - \vec{\epsilon}_{y,t})$  can be projected to lower dimension  $k$  ( $\vec{g}_t = H\vec{y}_t^\diamond$  to be exact) and the system follows a dynamic as described in:

$$\vec{y}_t^\diamond = T_M \vec{y}_{t-1}^\diamond + \vec{\epsilon}_{g,t}^\diamond \quad (20)$$

where  $T_M$  and  $\vec{\epsilon}_{g,t}^\diamond$  are appropriately transformed from  $T$  and  $\vec{\epsilon}_{g,t}$ , respectively. All module-module interactions are presented in the transition matrix  $T_M$ . Yoshida *et al.* further argued that the network structure may not be the same over time, and they proposed what they called “Markov switching” [71]. In essence, this is an inhomogeneous SSM, where at each time point the system is allowed to change its structure. They put the inference in a Bayesian framework, and introduced an additional vector of hidden variables which served as indicators of whether the system is switching at each time points. Posterior distribution of model parameters were estimated by using Gibbs sampling. Hirose *et al.* on the other hand, did a comparison study with the models in the work by Rangel *et al.* and Beal *et al.* [18, 70] and investigated the benefits of using multiple replicates of time course data.

Kasabov *et al.* also used the *standard SSM* and proceeded to estimate  $T$  using Kalman filter, and constructed the network from two sets of human leukemic cell line data, which have 32 pre-selected genes and 4 time points [73]. They also showed the potential application in larger data sets with more genes by implementing Genetic Algorithm (fitness function evaluated by using Kalman Filter estimated likelihood) for gene selection. The validation of the networks is rather weak, however, as they merely showed that the observed data fall on the estimated trajectories for four of the genes they selected for the network.

### 3.2.3. Ordinary Differential Equations

Ordinary differential equations can be used to describe the gene products, e.g., mRNA and protein, and their interactions. A general form of an  $n$ -node ODE system can be written as:

$$\frac{d\vec{g}(t)}{dt} = f(\vec{g}(t), t) + S(t), \quad (21)$$

where  $\vec{g}(t)$  is the gene product concentrations of  $n$  nodes,  $f(\vec{g}(t), t)$  the regulation function, and  $S(t)$  the external stimulus. For genetic regulation network, the most popular ODE model is the linear time-invariant model:

$$\frac{d\vec{g}(t)}{dt} = R\vec{g}(t) + W\vec{s}(t), \quad (22)$$

where  $R$  is an  $n \times n$  matrix denoting the direct regulation among  $n$  nodes, and  $\vec{s}$  a  $k \times 1$  constant vector whose effects on  $n$  nodes are intermediated by the  $n \times k$  constant matrix  $W$ . In practice the available data are sampled at limited time points, so most methods actually solve *Difference Equations* like

$$\vec{g}_{t+1} - \vec{g}_t = R\vec{g}_t + W\vec{s}_t. \quad (23)$$

In reality, noise, whether significant or slight, is always present in the observed raw data. Some approaches explicitly incorporate the noise into the system as error terms  $\vec{\epsilon}_t$ :

$$\vec{g}_{t+1} - \vec{g}_t = R\vec{g}_t + W\vec{s}_t + \vec{\epsilon}_t, \quad (24)$$

Note that this form closely resembles Eqn. (12) in Markov model.

For any node  $i$ , its regulation is defined by the  $i$ -th row of  $R$ , i.e.,  $R_i = [r_{i1}, \dots, r_{im}]$ . The signs of  $[r_{i1}, \dots, r_{im}]$  determine the network structure: for node  $j$ ,  $r_{ij} \neq 0$  means it is an input of node  $i$ , and depending on whether  $r_{ij}$  is positive or negative, node  $j$  activates or inhibits node  $i$ , respectively. If  $r_{ii} \neq 0$ , then the node  $i$  is self-regulated. The actual values of  $[r_{i1}, \dots, r_{im}]$  determine the production/degradation rates of node  $i$ . For most algorithms, the objective is to estimate  $R$ , especially the network structure. Furthermore, they normally decouple the problem and work on the regulation of one gene at a time based on  $dg_i(t)/dt = R_i\vec{g}(t) + W_i\vec{s}(t)$ .

Taking  $\vec{s} \equiv 0$ , Chen *et al.* provided two estimation solutions of  $R$  [74]. One solution used Fourier transform. It assumed that several cell cycles are observed, and the system is stable enough to be approximated by  $\vec{g}(t) = Qe^{\lambda t}$ , where  $Q$  is constant matrix and  $\lambda$  the eigenvalues of  $R$ . Assuming that each node's input set size is small and fixed, the other solution used Minimal Weight Solutions to Linear Equations [75] to solve the difference equations.

Instead of fixing input set size, de Hoon *et al.* allowed the network to have input sets of different sizes [76]. They assumed that the error terms are normally distributed so log-likelihood of proposed network can be calculated. The search for the network of maximal log-likelihood was regulated by Akaike's Information Criterion (AIC):  $AIC = -2[\log$



likelihood] + 2[number of parameters]. The approach limited the size of nodes it can handle to be smaller than the number of time points. Bourque and Sandkoff preferred a forward search approach [77]. For gene  $i$ , the fitness of its input set is measured by the sum of squared errors (SSE). To add a new input to the existing input set, the decrease of SSE must pass F-test at given significance level. The authors also extended the inference to multiple related networks, where the fitness includes not only SSE, but also evolution cost, which is the sum of pairwise symmetric differences of all networks. Chan *et al.* put connection constraints on linear model through sparse Bayesian learning [78]. In this approach, SSE is regulated by a parameter magnitude function, in which the magnitude of each connection  $r_{ij}$  is weighted by a hyperparameter. The optimization is conducted in a recursive way where the regulation matrix  $R$  and hyperparameters are estimated alternately.

The linear assumption limits the range of networks the model can emulate. To add nonlinearity to the model, Perkins *et al.* introduced a hybrid model by coupling the concentration value of each node with its logical state [79]. The approach normalizes the concentration to a range of [0, 1], discretizes the production rate to either on or off, and fixes the degradation rate. By this means, the regulation function  $f_i$  of gene  $i$  is reduced to a logical function plus the degradation term.

Another popular non-linear differential equation model is the S-system form, where the production/degradation rate is a product of power-law functions:

$$dg_i(t)/dt = \alpha_i \prod_{j=1}^n X_j^{u_{ij}} - \beta_i \prod_{j=1}^n X_j^{v_{ij}}, \quad (25)$$

where  $\alpha_i$  and  $\beta_i$  are rate constants,  $u_{ij}$  and  $v_{ij}$  are kinetic orders. The earlier work of Akutsu *et al.* was based on qualitative modeling and used linear programming which can only determine  $\log \alpha_i$ ,  $-\log \beta_i$  and  $u_{ij} - v_{ij}$  [55]. Later work like Marino and Voit used Levenberg-Marquardt method to find the parameters that minimize SSE [80]. Daisuke and Horton used distributed genetic algorithm to overcome the local minima [81]. The algorithm is seeded with networks that follow the scale-free property and results in multiple candidates that are later aggregated to determine the network structure. Novikov and Barillot converted differential equations into integral equations :

$$g_i(t) = \sum_{x=1}^n \int_{t_0}^t r_{ij}(t, x) g_j(x) dx + s_i(t, t_0), \quad (26)$$

where the  $r_{ij}$  and  $s_i$  are modeled by nonlinear time-variant kernel functions: polynomial, exponential or delta-function models [82]. Similar to Bourque and Sandkoff [77], the network is inferred by using forward search, but with  $\chi^2$  criterion value replacing F-test's p-value.

Recurrent Neural Network (RNN) is another popular model that can capture the nonlinear dynamics of various systems. Xu *et al.* used the model of the following form:

$$dg_i(t)/dt = (\sigma(R_i \vec{g}(t) + W_i \vec{s} + b_i) - \lambda_i g_i(t)) / \tau_i, \quad (27)$$

where  $\sigma(\cdot)$  is the sigmoid function,  $b_i$  the bias term,  $\lambda_i$  the self-degradation rate, and  $\tau_i$  the time constant [83]. The network is learned with particle swarm optimization [84]. Busch *et al.* studied the gene regulation network of the migration of human skin keratinocytes with RNN [25]. The model used in their work follows the form

$$dg_i(t)/dt = \left( \sum_{j=1}^n r_{ij} \sigma(g_j(t - \Delta\tau_j) + b_i) - g_i(t) + s_i(t) \right) / \tau_i, \quad (28)$$

where  $\Delta\tau_j$  is the time delay associated with gene  $j$ . The network is learned with Genetic Algorithm.

Cavelier and Anastassiou pursued a hybrid of linear and nonlinear functions, where linear functions are used for translation and nonlinear functions for transcription [85]. The authors considered three nonlinear functions of increasing complexity: sigmoid-type function, Hill function and thermodynamically derived function. The authors further assumed that the network structure is available through prior knowledge in the literature, so the algorithm focuses on the estimation of the production/degradation parameters through the so-called evolution strategies. Another way to add nonlinearity is through time-delay, as described by Kim *et al.* [86]:

$$dg_i(t)/dt = \sum_{j=1}^n r_{ij} (g_j(t - \tau_j) + \varepsilon_{ij}), \quad (29)$$

where  $\varepsilon_{ij}$  is a noise term.

Rather than inferring the exact chemical kinetic equations regulating every node, Sontag *et al.* estimated the influence of node  $i$  on the node  $j$  by measuring the change of node  $i$  production rate relative to the change of node  $j$  concentration at any give network status [87]. To do so, they proposed an experiment protocol in which a series of perturbations is applied and the unperturbed and perturbed time-series data are compared and evaluated: for an  $n$ -gene network, the protocol needs about  $n^2$  perturbations with perfect measurement to fully infer all the influences. Bansal *et al.* also proposed an inference scheme based on perturbation [22]. The difference equation is now written as

$$G_{+1} = [R + I, W] \begin{bmatrix} G \\ S \end{bmatrix}, \quad (30)$$

where  $I$  is the identity matrix,  $G = [\vec{g}_1, \vec{g}_2, \dots]$ ,  $G_{+1} = [\vec{g}_2, \vec{g}_3, \dots]$ , and  $S = [\vec{s}_1, \vec{s}_2, \dots]$ . The network structure is then estimated through the dominated singular values of the PCA decomposition of  $\begin{bmatrix} G \\ S \end{bmatrix}$ . With PCA decomposition,

this approach can handle a lot more genes with limited observations.

Arguing that the accuracy of single best regulation parameter set is prone to error due to limited data, Nam *et al.* aggregated the most likely regulator sets through voting [88]. Similarly, Kim *et al.* used noise injection to improve inference robustness [89]. In their model,  $r_{ij}$  is a time-varying function  $r_{ij}(t) = \alpha_{ij} \sin(\omega t + \phi_{ij}) + \beta_{ij}$ , where  $\alpha_{ij}$ ,  $\omega$ ,  $\phi_{ij}$  and  $\beta_{ij}$  are unknown parameters. The effects of gene  $j$  on gene  $i$  are assessed on the 2D trajectory map of  $\{g_i(t), r_{ij}(t)g_j(t)\}$ . By injecting random noise to the original data to generate several slightly different data sets, only the connections that shows certain degree of stability across all data sets are picked.

#### 4. PERFORMANCE EVALUATION

With the abundance of proposed models and inference algorithms, it is essential to have a validation protocol so that the merit of each proposed method can be assessed, and guideline can be established in aiding practitioners to choose the right modeling and inference procedure. Validation therefore should be considered as an impartible part of the complete inference scheme.

##### 4.1. Distance Measure

Ideally one would compare the inferred network  $N_I$ , with ‘ground truth’ network  $N_G$ , the one from which data used for inference are derived. The validation is given quantitatively in the form of the distance measure  $D(N_I, N_G)$ , as shown in Fig. (2a). Two critical issues arise from here, as discussed below.

##### 4.1.1. Lack of ‘Ground Truth’

Although  $N_G$  as the underlying biological network is the ultimate golden standard that the inferred network should be evaluated against, unfortunately it is almost guaranteed that we will not be blessed with this knowledge for any biological system of non-trivial size (as a matter of fact, we lack the knowledge for even trivial-sized system, for example, the three gene oscillating network of *E. Coli* [90]).

To circumvent this problem, two common approaches have been used.

The first approach is to use ‘partial ground truth’, which is most often in the form of regulation relationships gleaned from literature. This is demonstrated in Fig. (2b), in which  $D(N_I, N_G^p)$  can be thought as an approximation of  $D(N_I, N_G)$ . This approach has being widely utilized. For example: Hirose *et al.* found genes in each “module” are related to the same molecular function according to Gene Ontology (GO) [72]; in addition to GO, Dai *et al.* also searched *Saccharomyces Genome Database* to find relationships which are consistent to their findings; both Kim *et al.* and Novikov *et al.* compared their yeast cell cycle pathway networks with those selected from KEGG (Kyoto Encyclopedia of Genes and Genomes) [47, 82], and Kim *et al.* further compared with the metabolic pathway reported by DeRisi *et al.* [9]. Although versatile and bearing immediate biological interpretation, this approach is limited by the thoroughness and accuracy of reports from literature, and subject to the bias in mining the literature.

The second approach is the use of synthetic network generator, so we know every aspect of the underlying network  $N_G^s$ . Illustrated in Fig. (2c), the synthetic network  $N_G^s$  serves as a surrogate of the original  $N_G$ , and ideally, should mimic the real biological system. In reality,  $N_G^s$  can only model a subset of the properties in  $N_G$  and this leads to a inherent bias where certain class or classes of inference methods could be favored unintentionally. Setting up the synthetic network that is free of this bias is not trivial, but sometimes this obstacle can be overcome by carefully aligning synthetic network properties with those in the study objectives.

One immediate advantage of using  $N_G^s$  is that we can easily study the properties of the proposed inference algorithms. As an example, robustness property can be analyzed by perturbing and adding noise to the network. These learned properties often lead to improvement on the inference. For example, Bansal *et al.* randomly generated 100 networks for each of the two sizes, one with 10 gene and 5 time points, and the other with 1000 gene and 10 time

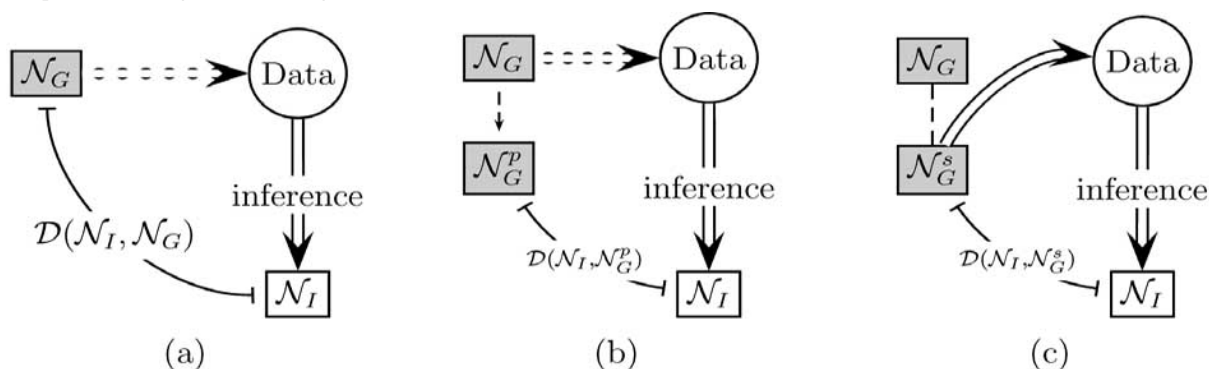


Fig. (2). Validation paradigms.  $N_G^p$  and  $N_G^s$  represent the ‘partial ground truth’ and synthetic network, respectively.

points, and added white Gaussian noise of different standard deviations to the generated time series data [22]. The tuning of parameters in inferring network from real data was facilitated by examining the performances on these synthetic network data. Perkins *et al.* fixed input set size for each node for their ODE model, and randomly selected the elements in input node sets and the associated Boolean functions [79]. They were able to derive the number of data points needed to fully infer the network structure and regulation functions in different scenarios. An intensive simulation study on DBN by Yu *et al.* evaluated various combinations of Bayesian scoring metrics, search heuristics, discretization levels, sampling interval, quantity of data and data interpolation using simulated data from a stochastic linear Markov model [91]. As a result, they published a series of guidelines for various factors in DBN learning. The benefits associated with using synthetic network in validating inference algorithms make it a practical and fruitful choice as surrogate of  $N_G$ .

#### 4.1.2. Choice of Criterion

When comparing two networks  $N_I$  and  $N_G$ ,  $D(N_I, N_G)$  can take one of many possible forms. For example, the most commonly used measures  $D(N_I, N_G^p)$  for structures are *true positive* (TP), *true negative* (TN), *false positive* (FP), *false negative* (FN), or some derivation from them (e.g., *Receiver Operating Characteristic*, or *ROC* curves). Dougherty has proposed a list of “semi-metrics” as validation measures for goodness of the inference algorithms [92]. The list can be grown to accommodate different objectives of studies. It is important to notice that there doesn't exist a one-size-fit-all criterion that works as a universal validation measure due to two prominent reasons: (1) the goal and focus vary widely from study to study, and the criterion has to be chosen to be consistent with the objective of the study. For example, if the purpose of the experiment is to discover pairwise gene-gene interactions, then a proper measure could be to compare the difference in connections (either directed or undirected) in  $N_I$  and  $N_G$ . Or if the interest is in how system evolves, a trajectory-based measure may be more appropriate [92]. (2) Inference algorithms are typically designed for a certain subtype of models, which in turn are proposed for some specific aims of the study. Using a validation measure which is more in line with the same goal will inevitably bias favorably towards these methods. This makes comparing across models particularly difficult, and is part of the reason that we see very few comparative studies on inference methodology. Even when such studies are carried out, they are practically limited to one type of models [93].

#### 4.2. Comparative Studies

Some efforts have been made to address the above issues. Brun *et al.* proposed a steady-state trajectory based metric between networks that is independent of nature of networks [94], hence has immediate application in comparative studies. Trajectory of each gene of the network is decomposed into a transient part and a steady-state part, the latter of which could be either constant or periodic, and

assessed with an *amplitude cumulative distribution* [95].  $D(N_I, N_G)$  is therefore the average (across all genes) of distances, which are computed as some norm between *amplitude cumulative distributions*. This metric could be useful if it is the steady-state behavior of the network that is of interest.

For comparison studies, Hartemink suggested that DBN seems to work better than Bayesian Network [2]. Note the author conceded that this is hardly a conclusion due to the “different properties” of the data used for inference. A more detailed study by Werhli *et al.* [93] compared the Relevance Network and Bayesian Network models, as well as the Graphical Gaussian Model. The last model is based on the assumption that genes are multivariate Gaussian distributed and the *partial correlation*, calculated on the estimated covariance matrix (through some stabilizing techniques), describes the correlation between two genes. In addition to a *Raf* signaling pathway protein expression data, they run the study using two synthetic data generators (one linear and the other nonlinear) so  $N_G^s$  is known. The evaluation was carried out using 2 criteria: ROC, and comparison of TP given fixed FP value, both of which are based on directed or undirected edges, or connections, in the networks  $N_I$  and  $N_G^s$ . Though the edges carry different meanings in these three different network models, the validation is appropriate in a broader ‘regulation relation’ sense.

Comparing networks of different natures has been attempted by Bansal *et al.* [96]. Three inference procedures were chosen: BANJO (Bayesian network) [97], ARACHE (relevance network) [98], and NIR [99] and MNI [100] (ODE), along with hierarchical clustering and *random* inference which served as references. The choice of these methods was due partially to the availability of their software code. Data were generated from a linear ODE model, and merits of inference were evaluated on the networks using *positive predictive value* (PPV), a ratio of TP and TP+FP, and *sensitivity*, a ratio of TP and TP+FN. Note PPV and *sensitivity* are also referred to as *precision* and *recall* and used by other researchers [101, 102]. It is very interesting to notice the lackluster performances, as demonstrated by being not far from the random method, for all three inference algorithms, particularly on time-series data, using the authors' model setups.

An applaudable endeavor by the DREAM initiative team during the past couple of years allows researchers to validate their networks  $N_I$  inferred from the data, which are generated from the  $N_G$  the team provides, in a competitive setting [102].  $N_I$  is in the form of pairwise regulation relationships, each tagged with a confidence probability. Much like the work of Werhli *et al.* [93], each participant's inferred network is evaluated against  $N_G$  using the area-under-ROC-curve, but with an additional area-under-prediction-versus-recall criterion. It is worth noting that both of these scores are still structure based, and it is expected that the team will have measures that target on the dynamics of inferred networks in the future.

### 4.3. Validation by New Experiments

As a completely different paradigm, the validation can be done experimentally, and the protocol usually runs like this:

1. Inference of network from experiment data;
2. Prediction of certain response using the inferred model;
3. Verification using new experiment.

If the confirmation from new experiment comes back as negative, the model is expected to *learn* from the error and revised. The feedback helps improving the inference procedure, but from a pure validation point of view, this protocol is particularly useful if we are interested in biological network intervention, where the exact correctness of the inferred network structure is of less importance; the ultimate criterion there is whether it can successfully predict the response given a perturbation. In practice the experimental data typically are the responses to selective or systematic perturbation (e.g., stimuli like starvation or drugs), or system behaviors after gene knockout/knock down, and are well suited for network intervention studies.

Such approach has already been taken by researchers. In the study of keratinocyte migration [25], the authors used gene *ptgs2* as migration indicator and built a *recurrent neural network* (RNN) model with nine genes to predict the migration behavior upon hepatocyte growth factor stimulus. They were able to follow it by *in vitro* experiment and the discrepancy with *in silico* predictions helped them build a second RNN model that is more consistent with experimental findings.

### 5. CONCLUSION

We have reviewed the modeling and inference of Gene Regulatory Network (GRN) from time-series data, categorized into those focused on structure, or those on both structure and dynamics, the latter of which is further bifurcated into discrete or continuous space models. The richness of proposed methods calls for comparative performance studies that can be used to establish merit of each inference procedure and appropriateness for a given application. These studies as we see are still lacking despite some modest efforts, due to the fact that ‘ground truth’ network required by various validation schemes is not readily available, and it is not always easy to find a criterion that can effectively evaluate a motley of model types which all have different design goals. On the other hand, the validation can be carried out experimentally where the requirement for ‘ground truth’ network is relaxed. This is especially useful in perturbation-response experiments, and the feedback from new experiments will aid the inferences as well. This survey gives a panoramic view on these topics, with anticipation that the readers will be inspired to improve and/or expand GRN inference and validation tool repository.

### 6. ACKNOWLEDGEMENTS

This study was funded by Translational Genomics Research Institute (TGen) institutional support (CS and JH) and partially by NIH LM009706-01 (SJ).

### REFERENCES

- [1] Friedman, N. Inferring Cellular Networks Using Probabilistic Graphical Models. *Science*, **2004**, 303(5659), 799-805.
- [2] Hartemink, A.J. Reverse engineering gene regulatory networks. *Nat. Biotech.*, **2005**, 23(5), 554-555.
- [3] Barabasi, A.-L.; Oltvai, Z.N. Network biology: understanding the cell's functional organization. *Nat. Rev. Genet.*, **2004**, 5(2), 101-113.
- [4] Cho, K. H.; Choo, S. M.; Jung, S.H.; Kim, J.R.; Choi, H.S.; Kim, J. Reverse engineering of gene regulatory networks. *IET Syst. Biol.*, **2007**, 1(3), 149-163.
- [5] Schlitt, T.; Brazma, A. Current approaches to gene regulatory network modelling. *BMC Bioinform.*, **2007**, 8(Suppl 6), S9.
- [6] van Someren, E.P.; Wessels, L.F.A.; Backer, E.; Reinders, M.J.T. Genetic network modeling. *Pharmacogenomics*, **2002**, 3(4), 507-525.
- [7] de Jong, H. Modeling and Simulation of Genetic Regulatory Systems: A Literature Review. *J. Comput. Biol.*, **2002**, 9, 67-103.
- [8] Biauudet, V.; Samson, F.; Bessieres, P. In Second Meeting on the Interconnection of Molecular Biology Databases; Cambridge, UK **1995**.
- [9] DeRisi, J.L.; Iyer, V.R.; Brown, P.O. Exploring the Metabolic and Genetic Control of Gene Expression on a Genomic Scale. *Science*, **1997**, 278(5338), 680-686.
- [10] Spellman, P.T.; Sherlock, G.; Zhang, M.Q.; Iyer, V.R.; Anders, K.; Eisen, M.B.; Brown, P.O.; Botstein, D.; Futcher, B. Comprehensive Identification of Cell Cycle-regulated Genes of the Yeast *Saccharomyces cerevisiae* by Microarray Hybridization. *Mol. Biol. Cell*, **1998**, 9(12), 3273-3297.
- [11] Cho, R.J.; Campbell, M.J.; Winzler, E.A.; Steinmetz, L.; Conway, A.; Wodicka, L.; Wolfsberg, T.G.; Gabrielian, A.E.; Landsman, D.; Lockhart, D.J.; Davis, R.W. A genome-wide transcriptional analysis of the mitotic cell cycle. *Mol. Cell*, **1998**, 2, 65-73.
- [12] Chu, S.; DeRisi, J.; Eisen, M.; Mulholland, J.; Botstein, D.; Brown, P.O.; Herskowitz, I. The Transcriptional Program of Sporulation in Budding Yeast. *Science*, **1998**, 282(5389), 699-705.
- [13] Wen, X.; Fuhrman, S.; Michaels, G.S.; Carr, D.B.; Smith, S.; Barker, J.L.; Somogyi, R. Large-scale temporal gene expression mapping of central nervous system and development. *Proc. Natl. Acad. Sci. USA*, **1998**, 95, 334-339.
- [14] Iyer, V.R.; Eisen, M.B.; Ross, D.T.; Schuler, G.; Moore, T.; Lee, J.C.; Trent, J.M.; Staudt, L. M.; Hudson, J.; Jr, Boguski, M. S.; Lashkari, D.; Shalon, D.; Botstein, D.; Brown, P.O. The Transcriptional Program in the Response of Human Fibroblasts to Serum. *Science*, **1999**, 283(5398), 83-87.
- [15] Primig, M.; Williams, R.M.; Winzler, E.A.; Tevzadze, G.G.; Conway, A.R.; Hwang, S.Y.; Davis, R.W.; Esposito, R.E. The core meiotic transcriptome in budding yeasts. *Nat. Genet.*, **2000**, 26(4), 415-423.
- [16] Khodursky, A.B.; Peter, B.J.; Cozzarelli, N.R.; Botstein, D.; Brown, P.O.; Yanofsky, C. DNA microarray analysis of gene expression in response to physiological and genetic changes that affect tryptophan metabolism in *Escherichia coli*. *PNAS*, **2000**, 97(22), 12170-12175.
- [17] Arbeitman, M.N.; Furlong, E.E.M.; Imam, F.; Johnson, E.; Null, B.H.; Baker, B.S.; Krasnow, M.A.; Scott, M.P.; Davis, R.W.; White, K.P. Gene expression during the life cycle of *Drosophila melanogaster*. *Science*, **2002**, 297(5590), 2270-2275.
- [18] Rangel, C.; Angus, J.; Ghahramani, Z.; Lioumi, M.; Sotharan, E.; Gaiba, A.; Wild, D.L.; Falciani, F. Modeling T-cell activation using gene expression profiling and state-space models. *Bioinformatics*, **2004**, 20(9), 1361-1372.
- [19] Schmitt, W.A.; Raab, R.M.; Stephanopoulos, G. Elucidation of gene interaction networks through time-lagged correlation analysis of transcriptional data. *Genome Res.*, **2004**, 14(8), 1654-1663.
- [20] Rudic, R.D.; McNamara, P.; Reilly, D.; Grosser, T.; Curtis, A.-M.; Price, T.S.; Panda, S.; Ho-genesch, J.B.; FitzGerald, G.A. Bioinformatic Analysis of Circadian Gene Oscillation in Mouse Aorta. *Circulation*, **2005**, 112(17), 2716-2724.
- [21] Gupta, A.; Varner, J.D.; Maranas, C.D. Large-scale inference of the transcriptional regulation of *Bacillus subtilis*. *Comput. Chem. Eng.*, **2005**, 29(3), 565-576.
- [22] Bansal, M.; Della Gatta, G.; di Bernardo, D. Inference of gene regulatory networks and compound mode of action from time

- course gene expression profiles. *Bioinformatics*, **2006**, *22*(7), 815-822.
- [23] Affara, M.; Dunmore, B.; Savoie, C.; Imoto, S.; Tamada, Y.; Araki, H.; Charnock-Jones, D. S.; Miyano, S.; Print, C. Understanding endothelial cell apoptosis: what can the transcriptome, glycome and proteome reveal? *Phil. Trans. R. Soc. B*, **2007**, *362*(1484), 1469-1487.
- [24] Martin, S.; Zhang, Z.; Martino, A.; Faulon, J.-L. Boolean dynamics of genetic regulatory networks inferred from microarray time series data. *Bioinformatics*, **2007**, *23*(7), 866-874.
- [25] Busch, H.; Camacho-Trullio, D.; Rogon, Z.; Breuhahn, K.; Angel, P.; Eils, R.; Szabowski, A. Gene network dynamics controlling keratinocyte migration. *Mol. Syst. Biol.*, **2008**, *4*, 199.
- [26] Bar-Joseph, Z. Analyzing time series gene expression data. *Bioinformatics*, **2004**, *20*(16), 2493-2503.
- [27] L'ahdesm'aki, H.; Hautaniemi, S.; Shmulevich, I.; Yli-Harja, O. Relationships between probabilistic Boolean networks and dynamic Bayesian networks as models of gene regulatory networks. *Signal Processing*, **2006**, *86*(4), 814-834.
- [28] Murphy, K.; Mian, S. Modelling gene expression data using dynamic Bayesian networks. Tech. rep., Computer Science Division, University of California, Berkeley, CA **1999**.
- [29] Murphy, K.P. Dynamic Bayesian Networks: Representation, Inference and Learning. PhD thesis, University of California, Berkeley **2002**.
- [30] Remondini, D.; O'Connell, B.; Intrator, N.; Sedivy, J.M.; Neretti, N.; Castellani, G.C.; Cooper, L.N. Targeting c-Myc-activated genes with a correlation method: Detection of global changes in large gene expression network dynamics. *PNAS*, **2005**, *102*(19), 6902-6906.
- [31] Remondini, D.; Neretti, N.; Franceschi, C.; Tieri, P.; Sedivy, J.M.; Milanesi, L.; Castellani, G.C. Networks from Gene Expression Time Series: Characterization of Correlation Patterns. *Int. J. Bifurcat. Chaos Appl. Sci. Eng.*, **2007**, *17*(7), 2477-2483.
- [32] Gupta, A.; Maranas, C.D.; Albert, R. Elucidation of directionality for co-expressed genes: predicting intra-operon termination sites. *Bioinformatics*, **2006**, *22*(2), 209-214.
- [33] Ma, P.C.H.; Chan, K.C.C. Inference of Gene Regulatory Networks from Time Series Expression Data: A Data Mining Approach. In Sixth IEEE International Conference on Data Mining - Workshops (ICDMW'06), Volume 0, Los Alamitos, CA, USA **2006**, 109-113.
- [34] Ma, P.C.H.; Chan, K.C.C. An Effective Data Mining Technique for Reconstructing Gene Regulatory Networks from Time Series Expression Data. *J. Bioinform. Comput. Biol.*, **2007**, *5*(3), 651-668.
- [35] Barker, N.; Myers, C.; Kuwahara, H.: Learning Genetic Regulatory Network Connectivity from Time Series Data. In The 9<sup>th</sup> International Conference on Industrial Engineering and other Application of Applied Intelligent Systems, **2006**.
- [36] Kwon, A.T.; Hoos, H.H.; Ng, R. Inference of transcriptional regulation relationships from gene expression data. *Bioinformatics*, **2003**, *19*(8), 905-912.
- [37] Bickel, D. R. Probabilities of spurious connections in gene networks: application to expression time series. *Bioinformatics*, **2005**, *21*(7), 1121-1128.
- [38] Friedman, N.; Lital, M.; Nachman, I.; Pe'er, D. Using Bayesian networks to analyze expression data. *J. Comput. Biol.*, **2000**, *7*(3-4), 601-620.
- [39] Dai, C.; Liu, J. Inducing Pairwise Gene Interactions from Time-Series Data by EDA Based Bayesian Network. In IEEE Engineering in Medicine and Biology Society, Shanghai, China **2005**, 7746-7749.
- [40] Friedman, N.; Murphy, K.; Russell, S. Learning the Structure of Dynamic Probabilistic Networks. In 14th Conference on Uncertainty in Artificial Intelligence. Edited by Cooper, G.; Moral, S., Madison, WI: Morgan Kaufmann **1998**, 139-147.
- [41] Ong, I.M.; Glasner, J.D.; Page, D. Modelling regulatory pathways in *E. coli* from time series expression profiles. *Bioinformatics*, **2002**, *18*(Suppl 1), S241-248.
- [42] Missal, K.; Cross, M.A.; Drasdo, D. Gene network inference from incomplete expression data: transcriptional control of hematopoietic commitment. *Bioinformatics*, **2006**, *22*(6), 731-738.
- [43] Zhao, W.; Serpedin, E.; Dougherty, E.R. Inferring gene regulatory networks from time series data using the minimum description length principle. *Bioinformatics*, **2006**, *22*(17), 2129-2135.
- [44] Rissanen, J. Modeling by shortest data description. *Automatica*, **1978**, *14*, 465-471.
- [45] Dougherty, J.; Tabus, I.; Astola, J. Inference of Gene Regulatory Networks Based on a Universal Minimum Description Length. *EURASIP J. Bioinform. Syst. Biol.*, **2008**, Article ID 482090, 11 pages.
- [46] Zou, M.; Conzen, S.D. A new dynamic Bayesian network (DBN) approach for identifying gene regulatory networks from time course microarray data. *Bioinformatics*, **2005**, *21*, 71-79.
- [47] Kim, S.; Imoto, S.; Miyano, S. Dynamic Bayesian network and nonparametric regression for nonlinear modeling of gene networks from time series gene expression data. *Biosystems*, **2004**, *75*(1-3), 57-65.
- [48] Kim, S.Y.; Imoto, S.; Miyano, S. In Computational Methods in Systems Biology, First International Workshop; Priami, C., Eds.; Springer: Roverto, Italy, **2003**, Vol. 2602, pp. 104-113.
- [49] Kim, S.Y.; Imoto, S.; Miyano, S. Inferring gene networks from time series microarray data using dynamic Bayesian networks. *Brief. Bioinformatics*, **2003**, *4*(3), 228-235.
- [50] Ferrazzi, F.; Sebastiani, P.; Ramoni, M.; Bellazzi, R. Bayesian approaches to reverse engineer cellular systems: a simulation study on nonlinear Gaussian networks. *BMC Bioinform.*, **2007**, *8*(Suppl 5), S2.
- [51] Dojer, N.; Gambin, A.; Mizera, A.; Wilczynski, B.; Tiurnyn, J. Applying dynamic Bayesian networks to perturbed gene expression data. *BMC Bioinform.*, **2006**, *7*, 249.
- [52] L'ahdesm'aki, H.; Shmulevich, I. Learning the structure of dynamic Bayesian networks from time series and steady state measurements. *Mach. Learn.*, **2008**, *71*(2), 185-217.
- [53] Liang, S.; Fuhrman, S.; Somogyi, R. In Pacific Symposium Biocomputing; **1998**, Vol. 3, pp. 18-29.
- [54] Akutsu, T.; Miyano, S.; Kuhara, S. In Proc. Pacific Symposium on Biocomputing: World Scientific: Singapore, **1999**, Vol. 4, pp. 17-28.
- [55] Akutsu, T.; Miyano, S.; Kuhara, S. Inferring qualitative relations in genetic networks and metabolic pathways. *Bioinformatics*, **2000**, *16*(8), 727-734.
- [56] Laubenbacher, R.; Stigler, B. A computational algebra approach to the reverse engineering of gene regulatory networks. *J. Theor. Biol.*, **2004**, *229*(4), 523-537.
- [57] Shmulevich, I.; Dougherty, E.R.; Kim, S.; Zhang, W. Probabilistic Boolean networks: a rule-based uncertainty model for gene regulatory networks. *Bioinformatics*, **2002**, *18*(2), 261-274.
- [58] Marshall, S.; Yu, L.; Xiao, Y.; Dougherty, E.R. Inference of a Probabilistic Boolean Network from a Single Observed Temporal Sequence. *EURASIP J. Bioinform. Syst. Biol.*, **2007**, Article ID 32454, 15 pages.
- [59] Ching, W.; Ng, M.M.; Fung, E.S.; Akutsu, T. On construction of stochastic genetic networks based on gene expression sequences. *Int. J. Neural. Syst.*, **2005**, *15*(4), 297-310.
- [60] Dewey, T.G.; Galas, D.J. Dynamic models of gene expression and classification. *Funct. Integr. Genomics*, **2001**, *1*(4), 269-278.
- [61] Dewey, T.G. From microarrays to networks: mining expression time series. *Drug Discov. Today*, **2002**, *7*(20), s170-s175.
- [62] Raychaudhuri, S.; Stuart, J.; Altman, R. In Pacific Symposium on Biocomputing; Oahu, HI, **2000**, Vol. 5, pp. 452-463.
- [63] Alter, O.; Brown, P.O.; Botstein, D. Singular value decomposition for genome-wide expression data processing and modeling. *PNAS*, **2000**, *97*(18), 10101-10106.
- [64] Holter, N.S.; Mitra, M.; Maritan, A.; Cieplak, M.; Banavar, J. R.; Fedoroff, N.V. Fundamental patterns underlying gene expression profiles: Simplicity from complexity. *Proc. Natl. Acad. Sci. USA*, **2000**, *97*(15), 8409-8414.
- [65] Holter, N.S.; Maritan, A.; Cieplak, M.; Fedoroff, N.V.; Banavar, J.R. Dynamic modeling of gene expression data. *Proc. Natl. Acad. Sci. USA*, **2001**, *98*(4), 1693-1698.
- [66] Wiggins, C.H.; Nemenman, I. Process pathway inference via time series analysis. *Exp. Mech.*, **2003**, *43*(3), 361-370.
- [67] Li, S.P.; Tseng, J.J.; Wang, S.C. Reconstructing gene regulatory networks from time-series microarray data. *Phys. A*, **2005**, *350*, 63-69.
- [68] Wang, S.C.; Li, S.P. In Biologically Inspired Approaches to Advanced Information Technology; Springer-Verlag, **2004**, Vol. 3141, pp. 49-62.
- [69] Chou, C.I.; Han, R.S.; Li, S.P.; Lee, T.-K. Guided simulated annealing method for optimization problems. *Phys. Rev. E*, **2003**, *67*(6), 066704.

- [70] Beal, M.J.; Falciani, F.; Ghahramani, Z.; Rangel, C.; Wild, D.L. A Bayesian approach to reconstructing genetic regulatory networks with hidden factors. *Bioinformatics*, **2005**, *21*(3), 349-356.
- [71] Yoshida, R.; Yoshida, R.; Imoto, S.; Higuchi, T. In Computational Systems Bioinformatics Conference, 2005. *Proceedings. 2005 IEEE*; Imoto, S., Eds.; **2005**, pp. 289-298.
- [72] Hirose, O.; Yoshida, R.; Imoto, S.; Yamaguchi, R.; Higuchi, T.; Charnock-Jones, D.S.; Print, C.; Miyano, S. Statistical inference of transcriptional module-based gene networks from time course gene expression profiles by using state space models. *Bioinformatics*, **2008**, *24*(7), 932-942.
- [73] Kasabov, N.K.; Chan, Z.S.H.; Jain, V.; Sidorov, I.; Dimitrov, D.S. In Neural Information Processing; Springer Berlin/Heidelberg, **2004**, Vol. 3316, pp. 1344-1353.
- [74] Chen, T.; He, H.L.; Church, G.M. In Pacific Symposium Biocomputing; Big Island, HI, **1999**, Vol. 4, pp. 29-40.
- [75] Garey, M.R.; Johnson, D.S. Computers and intractability: a guide to the theory of NP-completeness. W.H. Freeman, **1979**.
- [76] de Hoon, M.; Imoto, S.; Kobayashi, K.; Ogasawara, N.; Miyano, S. In Pacific Symposium on Biocomputing; **2003**, Vol. 8, pp. 17-28.
- [77] Bourque, G.; Sankoff, D. Improving Gene Network Inference by Comparing Expression Time-Series across Species, Developmental Stages or Tissues. *J. Bioinform. Comput. Biol.*, **2004**, *2*(4), 765-83.
- [78] Chan, Z.S.H.; Collins, L.; Kasabov, N. Bayesian learning of sparse gene regulatory networks. *Biosystems*, **2007**, *87*(2-3), 299-306.
- [79] Perkins, T.J.; Hallett, M.; Glass, L. Inferring models of gene expression dynamics. *J. Theor. Biol.*, **2004**, *230*(3), 289-299.
- [80] Marino, S.; Voit, E.O. An Automated Procedure for the Extraction of Metabolic Network Information from Time Series Data. *J. Bioinform. Comput. Biol.*, **2006**, *4*(3), 665-691.
- [81] Daisuke, T.; Horton, P. Inference of Scale-Free Networks from Gene Expression Time Series. *J. Bioinform. Comput. Biol.*, **2006**, *4*(2), 503-514.
- [82] Novikov, E.; Barillot, E. Regulatory network reconstruction using an integral additive model with flexible kernel functions. *BMC Syst. Biol.*, **2008**, *2*, 8.
- [83] Xu, R.; Venayagamoorthy, G.; Wunsch, D. In Advances in Neural Networks - ISNN 2006; Springer Berlin / Heidelberg, **2006**, Vol. 3973, pp. 648-653.
- [84] Kennedy, J. F.; Eberhart, R.C.; Shi, Y. Swarm intelligence. Morgan Kaufmann Publishers: San Francisco, **2001**.
- [85] Cavelier, G.; Anastassiou, D. Data-based model and parameter evaluation in dynamic transcriptional regulatory networks. *Proteins-Structure Funct. Bioinform.*, **2004**, *55*(2), 339-350.
- [86] Kim, S.; Kim, J.; Cho, K.H. Inferring gene regulatory networks from temporal expression profiles under time-delay and noise. *Comput. Biol. Chem.*, **2007**, *31*(4), 239-245.
- [87] Sontag, E.; Kiyatkin, A.; Kholodenko, B.N. Inferring dynamic architecture of cellular networks using time series of gene expression, protein and metabolite data. *Bioinformatics*, **2004**, *20*(12), 1877-1886.
- [88] Nam, D.; Yoon, S.H.; Kim, J.F. Ensemble learning of genetic networks from time-series expression data. *Bioinformatics*, **2007**, *23*(23), 3225-3231.
- [89] Kim, J.; Bates, D.G.; Postlethwaite, I.; Heslop-Harrison, P.; Cho, K.-H. Linear time-varying models can reveal non-linear interactions of biomolecular regulatory networks using multiple time-series data. *Bioinformatics*, **2008**, *24*(10), 1286-1292.
- [90] Elowitz, M. B.; Leibler, S. A synthetic oscillatory network of transcriptional regulators. *Nature*, **2000**, *403*(6767), 335-338.
- [91] Yu, J.; Smith, V.A.; Wang, P.P.; Hartemink, A.J.; Jarvis, E.D. Advances to Bayesian network inference for generating causal networks from observational biological data. *Bioinformatics*, **2004**, *20*(18), 3594-3603.
- [92] Dougherty, E.R. Validation of Inference Procedures for Gene Regulatory Networks. *Curr. Genomics*, **2007**, *8*, 351-359.
- [93] Werhli, A.V.; Grzegorzczak, M.; Husmeier, D. Comparative evaluation of reverse engineering gene regulatory networks with relevance networks, graphical gaussian models and bayesian networks. *Bioinformatics*, **2006**, *22*(20), 2523-2531.
- [94] Brun, M.; Kim, S.; Choi, W.; Dougherty, E.R. Comparison of Gene Regulatory Networks via Steady-State Trajectories. *EURASIP J. Bioinform. Syst. Biol.*, **2007**, Article ID 82702, 11 pages.
- [95] Lynn, P. An Introduction to the Analysis and Processing of Signals. John Wiley&Sons, **1973**.
- [96] Bansal, M.; Belcastro, V.; Ambesi-Impiomato, A.; di Bernardo, D. How to infer gene networks from expression profiles. *Mol. Syst. Biol.*, **2007**, *3*, 78.
- [97] Hartemink, A.: BANJO: Bayesian Network Inference with Java Objects, URL [<http://www.cs.duke.edu/~amink/software/banjo/>].
- [98] Margolin, A.; Nemenman, I.; Basso, K.; Wiggins, C.; Stolovitzky, G.; Favera, R.; Califano, A. ARACNE: An Algorithm for the Reconstruction of Gene Regulatory Networks in a Mammalian Cellular Context. *BMC Bioinformatics*, **2006**, *7*(Suppl 1), S7.
- [99] Gardner, T.S.; di Bernardo, D.; Lorenz, D.; Collins, J.J. Inferring Genetic Networks and Identifying Compound Mode of Action via Expression Profiling. *Science*, **2003**, *301*(5629), 102-105.
- [100] di Bernardo, D.; Thompson, M.J.; Gardner, T.S.; Chobot, S.E.; Eastwood, E.L.; Wojtovich, A.P.; Elliott, S.J.; Schaus, S.E.; Collins, J.J. Chemogenomic profiling on a genome-wide scale using reverse-engineered gene networks. *Nat. Biotech.*, **2005**, *23*(3), 377-383.
- [101] Andreucot, M.; Kauffman, S. On the Sparse Reconstruction of Gene Networks. *J. Comput. Biol.*, **2008**, *15*, 21-30.
- [102] [http://wiki.c2b2.columbia.edu/dream/index.php/The\\_DREAM\\_Project](http://wiki.c2b2.columbia.edu/dream/index.php/The_DREAM_Project).