

# Using Bayesian Networks to Analyze Expression Data

NIR FRIEDMAN,<sup>1</sup> MICHAL LINIAL,<sup>2</sup> IFTACH NACHMAN,<sup>3</sup> and DANA PE'ER<sup>1</sup>

## ABSTRACT

DNA hybridization arrays simultaneously measure the expression level for thousands of genes. These measurements provide a “snapshot” of transcription levels within the cell. A major challenge in computational biology is to uncover, from such measurements, gene/protein interactions and key biological features of cellular systems. In this paper, we propose a new framework for discovering interactions between genes based on multiple expression measurements. This framework builds on the use of *Bayesian networks* for representing statistical dependencies. A Bayesian network is a graph-based model of joint multivariate probability distributions that captures properties of conditional independence between variables. Such models are attractive for their ability to describe complex stochastic processes and because they provide a clear methodology for learning from (noisy) observations. We start by showing how Bayesian networks can describe interactions between genes. We then describe a method for recovering gene interactions from microarray data using tools for learning Bayesian networks. Finally, we demonstrate this method on the *S. cerevisiae* cell-cycle measurements of Spellman *et al.* (1998).

**Key words:** gene expression, microarrays, Bayesian methods.

## 1. INTRODUCTION

**A** CENTRAL GOAL OF MOLECULAR BIOLOGY is to understand the regulation of protein synthesis and its reactions to external and internal signals. All the cells in an organism carry the same genomic data, yet their protein makeup can be drastically different both temporally and spatially, due to regulation. Protein synthesis is regulated by many mechanisms at its different stages. These include mechanisms for controlling transcription initiation, RNA splicing, mRNA transport, translation initiation, post-translational modifications, and degradation of mRNA/protein. One of the main junctions at which regulation occurs is mRNA transcription. A major role in this machinery is played by proteins themselves that bind to regulatory regions along the DNA, greatly affecting the transcription of the genes they regulate.

In recent years, technical breakthroughs in spotting hybridization probes and advances in genome sequencing efforts lead to development of *DNA microarrays*, which consist of many species of probes, either oligonucleotides or cDNA, that are immobilized in a predefined organization to a solid phase. By using

---

<sup>1</sup>School of Computer Science and Engineering, Hebrew University, Jerusalem, 91904, Israel.

<sup>2</sup>Institute of Life Sciences, Hebrew University, Jerusalem, 91904, Israel.

<sup>3</sup>Center for Neural Computation and School of Computer Science and Engineering, Hebrew University, Jerusalem, 91904, Israel.

DNA microarrays, researchers are now able to measure the abundance of thousands of mRNA targets simultaneously (DeRisi *et al.*, 1997; Lockhart *et al.*, 1996; Wen *et al.*, 1998). Unlike classical experiments, where the expression levels of only a few genes were reported, DNA microarray experiments can measure *all* the genes of an organism, providing a “genomic” viewpoint on gene expression. As a consequence, this technology facilitates new experimental approaches for understanding gene expression and regulation (Iyer *et al.*, 1999; Spellman *et al.*, 1998).

Early microarray experiments examined few samples and mainly focused on differential display across tissues or conditions of interest. The design of recent experiments focuses on performing a larger number of microarray assays ranging in size from a dozen to a few hundreds of samples. In the near future, data sets containing thousands of samples will become available. Such experiments collect enormous amounts of data, which clearly reflect many aspects of the underlying biological processes. An important challenge is to develop methodologies that are both statistically sound and computationally tractable for analyzing such data sets and inferring biological interactions from them.

Most of the analysis tools currently used are based on *clustering* algorithms. These algorithms attempt to locate groups of genes that have similar expression patterns over a set of experiments (Alon *et al.*, 1999; Ben-Dor *et al.*, 1999; Eisen *et al.*, 1999; Michaels *et al.*, 1998; Spellman *et al.*, 1998). Such analysis has proven to be useful in discovering genes that are co-regulated and/or have similar function. A more ambitious goal for analysis is to reveal the structure of the transcriptional regulation process (Akutsu, 1998; Chen *et al.*, 1999; Somogyi *et al.*, 1996; Weaver *et al.*, 1999). This is clearly a hard problem. The current data is extremely noisy. Moreover, mRNA expression data alone only gives a partial picture that does not reflect key events, such as translation and protein (in)activation. Finally, the amount of samples, even in the largest experiments in the foreseeable future, does not provide enough information to construct a fully detailed model with high statistical significance.

In this paper, we introduce a new approach for analyzing gene expression patterns that uncovers properties of the transcriptional program by examining statistical properties of *dependence* and *conditional independence* in the data. We base our approach on the well-studied statistical tool of *Bayesian networks* (Pearl, 1988). These networks represent the dependence structure between multiple interacting quantities (e.g., expression levels of different genes). Our approach, probabilistic in nature, is capable of handling noise and estimating the confidence in the different features of the network. We are, therefore, able to focus on interactions whose signal in the data is strong.

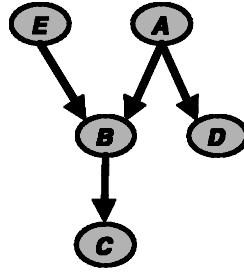
Bayesian networks are a promising tool for analyzing gene expression patterns. First, they are particularly useful for describing processes composed of *locally* interacting components; that is, the value of each component *directly* depends on the values of a relatively small number of components. Second, statistical foundations for learning Bayesian networks from observations, and computational algorithms to do so, are well understood and have been used successfully in many applications. Finally, Bayesian networks provide models of causal influence: Although Bayesian networks are mathematically defined strictly in terms of probabilities and conditional independence statements, a connection can be made between this characterization and the notion of *direct causal influence*. (Heckerman *et al.*, 1999; Pearl and Verma, 1991; Spirtes *et al.*, 1993). Although this connection depends on several assumptions that do not necessarily hold in gene expression data, the conclusions of Bayesian network analysis might be indicative of some causal connections in the data.

The remainder of this paper is organized as follows. In Section 2, we review key concepts of Bayesian networks, learning them from observations and using them to infer causality. In Section 3, we describe how Bayesian networks can be applied to model interactions among genes and discuss the technical issues that are posed by this type of data. In Section 4, we apply our approach to the gene-expression data of Spellman *et al.* (1998), analyzing the statistical significance of the results and their biological plausibility. Finally, in Section 5, we conclude with a discussion of related approaches and future work.

## 2. BAYESIAN NETWORKS

### 2.1. Representing distributions with Bayesian networks

Consider a finite set  $\mathcal{X} = \{X_1, \dots, X_n\}$  of random variables where each variable  $X_i$  may take on a value  $x_i$  from the domain  $\text{Val}(X_i)$ . In this paper, we use capital letters, such as  $X, Y, Z$ , for variable names and lowercase letters, such as  $x, y, z$ , to denote specific values taken by those variables. Sets of variables are denoted by boldface capital letters  $\mathbf{X}, \mathbf{Y}, \mathbf{Z}$ , and assignments of values to the variables in these sets



**FIG. 1.** An example of a simple Bayesian network structure. This network structure implies several conditional independence statements:  $I(A; E)$ ,  $I(B; D \mid A, E)$ ,  $I(C; A, D, E \mid B)$ ,  $I(D; B, C, E \mid A)$ , and  $I(E; A, D)$ . The network structure also implies that the joint distribution has the product form:  $P(A, B, C, D, E) = P(A)P(B|A, E)P(C|B)P(D|A)P(E)$ .

are denoted by boldface lowercase letters  $\mathbf{x}, \mathbf{y}, \mathbf{z}$ . We denote  $I(\mathbf{X}; \mathbf{Y} \mid \mathbf{Z})$  to mean  $\mathbf{X}$  is independent of  $\mathbf{Y}$  conditioned on  $\mathbf{Z}$ :  $P(\mathbf{X}|\mathbf{Y}, \mathbf{Z}) = P(\mathbf{X}|\mathbf{Z})$ .

A *Bayesian network* is a representation of a joint probability distribution. This representation consists of two components. The first component,  $G$ , is a *directed acyclic graph* (DAG) whose vertices correspond to the random variables  $X_1, \dots, X_n$ . The second component,  $\theta$  describes a conditional distribution for each variable, given its parents in  $G$ . Together, these two components specify a unique distribution on  $X_1, \dots, X_n$ .

The graph  $G$  represents conditional independence assumptions that allow the joint distribution to be decomposed, economizing on the number of parameters. The graph  $G$  encodes the *Markov Assumption*:

(\*) Each variable  $X_i$  is independent of its nondescendants, given its parents in  $G$ .

By applying the chain rule of probabilities and properties of conditional independencies, any joint distribution that satisfies (\*) can be decomposed into the *product form*

$$P(X_1, \dots, X_n) = \prod_{i=1}^n P(X_i | \mathbf{Pa}^G(X_i)), \quad (1)$$

where  $\mathbf{Pa}^G(X_i)$  is the set of parents of  $X_i$  in  $G$ . Figure 1 shows an example of a graph  $G$  and lists the Markov independencies it encodes and the product form they imply.

A graph  $G$  specifies a product form as in (1). To fully specify a joint distribution, we also need to specify each of the conditional probabilities in the product form. The second part of the Bayesian network describes these conditional distributions,  $P(X_i | \mathbf{Pa}^G(X_i))$  for each variable  $X_i$ . We denote the parameters that specify these distributions by  $\theta$ .

In specifying these conditional distributions, we can choose from several representations. In this paper, we focus on two of the most commonly used representations. For the following discussion, suppose that the parents of a variable  $X$  are  $\{U_1, \dots, U_k\}$ . The choice of representation depends on the type of variables we are dealing with:

- **Discrete variables.** If each of  $X$  and  $U_1, \dots, U_k$  takes discrete values from a finite set, then we can represent  $P(X \mid U_1, \dots, U_k)$  as a table that specifies the probability of values for  $X$  for each joint assignment to  $U_1, \dots, U_k$ . Thus, for example, if all the variables are binary valued, the table will specify  $2^k$  distributions.

This is a general representation which can describe any discrete conditional distribution. Thus, we do not lose expressiveness by using this representation. This flexibility comes at a price: The number of free parameters is exponential in the number of parents.

- **Continuous variables.** Unlike the case of discrete variables, when the variable  $X$  and its parents  $U_1, \dots, U_k$  are real valued, there is no representation that can represent all possible densities. A natural choice for multivariate continuous distributions is the use of Gaussian distributions. These can be represented in a Bayesian network by using *linear Gaussian* conditional densities. In this representation, the conditional density of  $X$  given its parents is given by:

$$P(X \mid u_1, \dots, u_k) \sim N(a_0 + \sum_i a_i \cdot u_i, \sigma^2).$$

That is,  $X$  is normally distributed around a mean that depends *linearly* on the values of its parents. The variance of this normal distribution is independent of the parents' values. If all the variables in a network have linear Gaussian conditional distributions, then the joint distribution is a multivariate Gaussian (Lauritzen and Wermuth, 1989).

- **Hybrid networks.** When the network contains a mixture of discrete and continuous variables, we need to consider how to represent a conditional distribution for a continuous variable with discrete parents and for a discrete variable with continuous parents. In this paper, we disallow the latter case. When a continuous variable  $X$  has discrete parents, we use *conditional Gaussian* distributions (Lauritzen and Wermuth, 1989) in which, for each joint assignment to the discrete parents of  $X$ , we represent a linear Gaussian distribution of  $X$  given its continuous parents.

Given a Bayesian network, we might want to answer many types of questions that involve the joint probability (e.g., what is the probability of  $X = x$  given observation of some of the other variables?) or independencies in the domain (e.g., are  $X$  and  $Y$  independent once we observe  $Z$ ?). The literature contains a suite of algorithms that can answer such queries efficiently by exploiting the explicit representation of structure (Jensen, 1996; Pearl, 1988)).

## 2.2. Equivalence classes of Bayesian networks

A Bayesian network structure  $G$  implies a set of independence assumptions in addition to (\*). Let  $\text{Ind}(G)$  be the set of independence statements (of the form  $X$  is independent of  $Y$  given  $Z$ ) that hold in all distributions satisfying these Markov assumptions.

More than one graph can imply exactly the same set of independencies. For example, consider graphs over two variables  $X$  and  $Y$ . The graphs  $X \rightarrow Y$  and  $X \leftarrow Y$  both imply the same set of independencies (i.e.,  $\text{Ind}(G) = \emptyset$ ). Two graphs  $G$  and  $G'$  are *equivalent* if  $\text{Ind}(G) = \text{Ind}(G')$ . That is, both graphs are alternative ways of describing the same set of independencies.

This notion of equivalence is crucial, since when we examine observations from a distribution, we cannot distinguish between equivalent graphs.<sup>1</sup> Pearl and Verma (1991) show that we can characterize *equivalence classes* of graphs using a simple representation. In particular, these results establish that equivalent graphs have the same underlying undirected graph but might disagree on the direction of some of the arcs.

**Theorem 2.1** (Pearl and Verma, 1991). *Two DAGs are equivalent if and only if they have the same underlying undirected graph and the same v-structures (i.e., converging directed edges into the same node, such that  $a \rightarrow b \leftarrow c$ , and there is no edge between  $a$  and  $c$ ).*

Moreover, an equivalence class of network structures can be uniquely represented by a *partially directed graph* (PDAG), where a directed edge  $X \rightarrow Y$  denotes that all members of the equivalence class contain the arc  $X \rightarrow Y$ ; an undirected edge  $X-Y$  denotes that some members of the class contain the arc  $X \rightarrow Y$ , while others contain the arc  $Y \rightarrow X$ . Given a DAG  $G$ , the PDAG representation of its equivalence class can be constructed efficiently (Chickering, 1995).

## 2.3. Learning Bayesian networks

The problem of learning a Bayesian network can be stated as follows. Given a *training set*  $D = \{\mathbf{x}^1, \dots, \mathbf{x}^N\}$  of independent instances of  $\mathcal{X}$ , find a network  $B = \langle G, \Theta \rangle$  that *best matches*  $D$ . More precisely, we search for an equivalence class of networks that best matches  $D$ .

The theory of learning networks from data has been examined extensively over the last decade. We only briefly describe the high-level details here. We refer the interested reader to Heckerman (1998) for a recent tutorial on the subject.

The common approach to this problem is to introduce a statistically motivated scoring function that evaluates each network with respect to the training data and to search for the optimal network according

<sup>1</sup>To be more precise, under the common assumptions in learning networks, which we also make in this paper, one cannot distinguish between equivalent graphs. If we make stronger assumptions, for example, by restricting the form of the conditional probability distributions we can learn, we might have a preference for one equivalent network over another.

to this score. One method for deriving a score is based on Bayesian considerations (see Cooper and Herskovits (1992) and Heckerman *et al.* (1995) for complete description). In this score, we evaluate the posterior probability of a graph given the data:

$$\begin{aligned} S(G : D) &= \log P(G \mid D) \\ &= \log P(D \mid G) + \log P(G) + C \end{aligned}$$

where  $C$  is a constant independent of  $G$  and

$$P(D \mid G) = \int P(D \mid G, \Theta) P(\Theta \mid G) d\Theta$$

is the *marginal likelihood* which averages the probability of the data over all possible parameter assignments to  $G$ . The particular choice of priors  $P(G)$  and  $P(\Theta \mid G)$  for each  $G$  determines the exact Bayesian score. Under mild assumptions on the prior probabilities, this scoring rule is asymptotically consistent: Given a sufficiently large number of samples, graph structures that exactly capture all dependencies in the distribution, will receive, with high probability, a higher score than all other graphs (see, for example, Friedman and Yakhini (1996)). This means, that given a sufficiently large number of instances, learning procedures can pinpoint the exact network structure up to the correct equivalence class.

In this work, we use a prior described by Heckerman and Geiger (1995) for hybrid networks of multinomial distributions and conditional Gaussian distributions. (This prior combines earlier works on priors for multinomial networks (Buntine, 1991; Cooper and Herskovits, 1992; Heckerman *et al.*, 1995) and for Gaussian networks (Geiger and Heckerman, 1994).) We refer the reader to Heckerman and Geiger (1995) and Heckerman (1998) for details on these priors.

In the analysis of gene expression data, we use a small number of samples. Therefore, care should be taken in choosing the prior. Without going into details, each prior from the family of priors described by Heckerman and Geiger can be specified by two parameters. The first is a *prior network*, which reflects our prior belief on the joint distribution of the variables in the domain, and the second is an *effective sample size* parameter, which reflects how strong is our belief in the prior network. Intuitively, setting the effective sample size to  $K$  is equivalent to having seen  $K$  samples from the distribution defined by the prior network. In our experiments, we choose the prior network to be one where all the random variables are independent of each other. In the prior network, discrete random variables are uniformly distributed, while continuous random variables have an a priori normal distribution. We set the equivalent sample size 5 in a rather arbitrary manner. This choice of the prior network is to ensure that we are not (explicitly) biasing the learning procedure to any particular edge. In addition, as we show below, the results are reasonably insensitive to this exact magnitude of the equivalent sample size.

In this work we assume *complete data*, that is, a data set in which each instance contains the values of all the variables in the network. When the data is complete and the prior satisfies the conditions specified by Heckerman and Geiger, then the posterior score satisfies several properties. First, the score is *structure equivalent*, i.e., if  $G$  and  $G'$  are equivalent graphs they are guaranteed to have the same posterior score. Second, the score is *decomposable*. That is, the score can be rewritten as the sum

$$S(G : D) = \sum_i \text{ScoreContribution}(X_i, \mathbf{Pa}^G(X_i) : D),$$

where the contribution of every variable  $X_i$  to the total network score depends only on the values of  $X_i$  and  $\mathbf{Pa}^G(X_i)$  in the training instances. Finally, these local contributions for each variable can be computed using a closed form equation (again, see Heckerman and Geiger (1995) for details).

Once the prior is specified and the data is given, learning amounts to finding the structure  $G$  that maximizes the score. This problem is known to be NP-hard (Chickering, 1996). Thus we resort to heuristic search. The decomposition of the score is crucial for this optimization problem. A *local* search procedure that changes one arc at each move can efficiently evaluate the gains made by adding, removing, or reversing a single arc. An example of such a procedure is a greedy hill-climbing algorithm that at each step performs the local change that results in the maximal gain, until it reaches a local maximum. Although this procedure does not necessarily find a global maximum, it does perform well in practice. Examples of other search methods that advance using one-arc changes include beam-search, stochastic hill-climbing, and simulated annealing.

## 2.4. Learning causal patterns

Recall that a Bayesian network is a model of dependencies between multiple measurements. However, we are also interested in modeling the mechanism that generated these dependencies. Thus, we want to model the flow of causality in the system of interest (e.g., gene transcription in the gene expression domain). A *causal network* is a model of such causal processes. Having a causal interpretation facilitates predicting the effect of an *intervention* in the domain: setting the value of a variable in such a way that the manipulation itself does not affect the other variables.

While at first glance there seems to be no direct connection between probability distributions and causality, causal interpretations for Bayesian networks have been proposed (Pearl and Verma, 1991; Pearl, 2000). A causal network is mathematically represented similarly to a Bayesian network, a DAG where each node represents a random variable along with a local probability model for each node. However, causal networks have a stricter interpretation of the meaning of edges: the parents of a variable are its *immediate causes*.

A causal network models not only the distribution of the observations, but also the effects of *interventions*. If  $X$  causes  $Y$ , then manipulating the value of  $X$  affects the value of  $Y$ . On the other hand, if  $Y$  is a cause of  $X$ , then manipulating  $X$  will not affect  $Y$ . Thus, although  $X \rightarrow Y$  and  $X \leftarrow Y$  are equivalent Bayesian networks, they are not equivalent causal networks.

A causal network can be interpreted as a Bayesian network when we are willing to make the *Causal Markov Assumption*: given the values of a variable's immediate causes, it is independent of its earlier causes. When the causal Markov assumption holds, the causal network satisfies the Markov independencies of the corresponding Bayesian network. For example, this assumption is a natural one in models of genetic pedigrees: once we know the genetic makeup of the individual's parents, the genetic makeup of her ancestors is not informative about her own genetic makeup.

The central issue is: When can we learn a causal network from observations? This issue received a thorough treatment in the literature (Heckerman *et al.*, 1999; Pearl and Verma, 1991; Spirtes *et al.*, 1993; Spirtes *et al.*, 1999). We briefly review the relevant results for our needs here. For a more detailed treatment of the topic, we refer the reader to Pearl (2000) and to Cooper and Glymour (1999).

First it is important to distinguish between an *observation* (a passive measurement of our domain, i.e., a sample from  $\mathcal{X}$ ) and an *intervention* (setting the values of some variables using forces outside the causal model, e.g., gene knockout or over-expression). It is well known that interventions are an important tool for inferring causality. What is surprising is that some causal relations can be inferred from observations alone.

To learn about causality, we need to make several assumptions. The first assumption is a modeling assumption: we assume that the (unknown) causal structure of the domain satisfies the Causal Markov Assumption. Thus, we assume that causal networks can provide a reasonable model of the domain. Some of the results in the literature require a stronger version of this assumption, namely, that causal networks can provide a perfect description of the domain (that is, an independence property holds in the domain if and only if it is implied by the model). The second assumption is that there are no *latent* or hidden variables that affect several of the observable variables. We discuss relaxations of this assumption below.

If we make these two assumptions, then we essentially assume that one of the possible DAGs over the domain variables is the "true" causal network. However, as discussed above, from observations alone, we cannot distinguish between causal networks that specify the same independence properties, i.e., belong to the same equivalence class (see Section 2.2). Thus, at best we can hope to learn a description of the equivalence class that contains the true model. In other words, we will learn a PDAG description of this equivalence class.

Once we identify such a PDAG, we are still uncertain about the true causal structure in the domain. However, we can draw some causal conclusions. For example, if there is a directed path from  $X$  to  $Y$  in the PDAG, then  $X$  is a causal ancestor of  $Y$  in *all* the networks that could have generated this PDAG including the "true" causal model. Moreover, by using Theorem 2.1, we can predict what aspects of a proposed model would be detectable based on observations alone.

When data is sparse, we cannot identify a unique PDAG as a model of the data. In such a situation, we can use the posterior over PDAGs to represent *posterior* probabilities over causal statements. In a sense the posterior probability of " $X$  causes  $Y$ " is the sum of the posterior probability of all PDAGs in which this statement holds. (See Heckerman *et al.* (1999) for more details on this Bayesian approach.) The situation is somewhat more complex when we have a combination of observations and results of different

interventions. From such data we might be able to distinguish between equivalent structures. Cooper and Yoo (1999) show how to extend the Bayesian approach of Heckerman *et al.* (1999) for learning from such mixed data.

A possible pitfall in learning causal structure is the presence of latent variables. In such a situation the observations that  $X$  and  $Y$  depend on each other probabilistically might be explained by the existence of an unobserved common cause. When we consider only two variables we cannot distinguish this hypothesis from the hypotheses “ $X$  causes  $Y$ ” or “ $Y$  causes  $X$ ”. However, a more careful analysis shows that one can characterize all networks with latent variables that can result in the same set of independencies over the observed variables. Such equivalence classes of networks can be represented by a structure called a *partial ancestral graphs* (PAGs) (Spirtes *et al.*, 1999). As can be expected, the set of causal conclusions we can make when we allow latent variables is smaller than the set of causal conclusions when we do not allow them. Nonetheless, causal relations often can be recovered even in this case.

The situation is more complicated when we do not have enough data to identify a single PAGs. As in the case of PDAGs, we might want to compute posterior scores for PAGs. However, unlike for PDAGs the question of scoring a PAG (which consists of many models with different number of latent variables) remains an open question.

### 3. ANALYZING EXPRESSION DATA

In this section we describe our approach to analyzing gene expression data using Bayesian network learning techniques.

#### 3.1. Outline of the proposed approach

We start with our modeling assumptions and the type of conclusions we expect to find. Our aim is to understand a particular *system* (a cell or an organism and its environment). At each point in time, the system is in some *state*. For example, the state of a cell can be defined in terms of the concentration of proteins and metabolites in the various compartments, the amount of external ligands that bind to receptors on the cell's membrane, the concentration of different mRNA molecules in the cytoplasm, etc.

The cell (or other biological systems) consists of many interacting components that affect each other in some consistent fashion. Thus, if we consider random sampling of the system, some states are more probable. The likelihood of a state can be specified by the joint probability distribution on each of the cells components.

Our aim is to estimate such a probability distribution and understand its structural features. Of course, the state of a system can be infinitely complex. Thus, we resort to a partial view and focus on some of the components. Measurements of attributes from these components are random variables that represent some aspect of the system's state. In this paper, we are dealing mainly with random variables that denote the mRNA expression level of specific genes. However, we can also consider other random variables that denote other aspects of the system state, such as the phase of the system in the cell-cycle process. Other examples include measurements of experimental conditions, temporal indicators (i.e., the time/stage that the sample was taken from), background variables (e.g., which clinical procedure was used to get a biopsy sample), and exogenous cellular conditions.

Our aim is to model the system as a joint distribution over a collection of random variables that describe system states. If we had such a model, we could answer a wide range of queries about the system. For example, does the expression level of a particular gene depend on the experimental condition? Is this dependence direct or indirect? If it is indirect, which genes mediate the dependency? Not having a model at hand, we want to learn one from the available data and use it to answer questions about the system.

In order to learn such a model from expression data, we need to deal with several important issues that arise when learning in the gene expression domain. These involve statistical aspects of interpreting the results, algorithmic complexity issues in learning from the data, and the choice of local probability models.

Most of the difficulties in learning from expression data revolve around the following central point: Contrary to most situations where one attempts to learn models (and in particular Bayesian networks), expression data involves transcript levels of thousands of genes while current data sets contain at most a few dozen samples. This raises problems in both computational complexity and the statistical significance of the resulting networks. On the positive side, genetic regulation networks are believed to be sparse,

i.e., given a gene, it is assumed that no more than a few dozen genes directly affect its transcription. Bayesian networks are especially suited for learning in such sparse domains.

### 3.2. Representing partial models

When learning models with many variables, small data sets are not sufficiently informative to significantly determine that a single model is the “right” one. Instead, many different networks should be considered as reasonable explanations of the given data. From a Bayesian perspective, we say that the posterior probability over models is not dominated by a single model (or an equivalence class of models).<sup>2</sup>

One potential approach to deal with this problem is to find all the networks that receive high posterior score. Such an approach is outlined by Madigan and Raftery (1994). Unfortunately, due to the combinatoric aspect of networks the set of “high posterior” networks can be huge (i.e., exponential in the number of variables). Thus, in a domain such as gene expression with many variables and diffused posterior, we cannot hope to explicitly list all the networks that are plausible given the data.

Our solution is as follows. We attempt to identify properties of the network that might be of interest. For example, are  $X$  and  $Y$  “close” neighbors in the network. We call such properties *features*. We then try to estimate the posterior probability of features given the data. More precisely, a feature  $f$  is an indicator function that receives a network structure  $G$  as an argument and returns 1 if the structure (or the associated PDAG) satisfies the feature and 0 otherwise. The posterior probability of a feature is

$$P(f(G) \mid D) = \sum_G f(G) P(G \mid D). \quad (2)$$

Of course, exact computation of such a posterior probability is as hard as processing all networks with high posteriors. However, as we shall see below, we can estimate these posteriors by finding representative networks. Since each feature is a binary attribute, this estimation is fairly robust even from a small set of networks (assuming that they are an unbiased sample from the posterior).

Before we examine the issue of estimating the posterior of such features, we briefly discuss two classes of features involving pairs of variables. While at this point we handle only pairwise features, it is clear that this type of analysis is not restricted to them, and in the future we are planning to examine more complex features.

The first type of feature is *Markov relations*: Is  $Y$  in the *Markov blanket* of  $X$ ? The Markov blanket of  $X$  is the minimal set of variables that *shield*  $X$  from the rest of the variables in the model. More precisely,  $X$  given its Markov blanket is independent from the remaining variables in the network. It is easy to check that this relation is symmetric:  $Y$  is in  $X$ ’s Markov blanket if and only if there is either an edge between them or both are parents of another variable (Pearl, 1988). In the context of gene expression analysis, a Markov relation indicates that the two genes are related in some joint biological interaction or process. Note that two variables in a Markov relation are directly linked in the sense that no variable *in the model* mediates the dependence between them. It remains possible that an unobserved variable (e.g., protein activation) is an intermediate in their interaction.

The second type of feature is *order relations*: Is  $X$  an ancestor of  $Y$  in all the networks of a given equivalence class? That is, does the given PDAG contain a path from  $X$  to  $Y$  in which all the edges are directed? This type of feature does not involve only a close neighborhood, but rather captures a global property. Recall that under the assumptions discussed in Section 2.4, learning that  $X$  is an ancestor of  $Y$  would imply that  $X$  is a cause of  $Y$ . However, these assumptions are quite strong (in particular the assumption of no latent common causes) and thus do not necessarily hold in the context of expression data. Thus, we view such a relation as an indication, rather than evidence, that  $X$  might be a causal ancestor of  $Y$ .

### 3.3. Estimating statistical confidence in features

We now face the following problem: To what extent does the data support a given feature? More precisely, we want to estimate the posterior of features as defined in (2). Ideally, we would like to sample networks from the posterior and use the sampled networks to estimate this quantity. Unfortunately, sampling from the posterior is a hard problem. The general approach to this problem is to build a *Markov Chain Monte*

<sup>2</sup>This observation is not unique to Bayesian network models. It equally well applies to other models that are learned from gene expression data, such as clustering models.



*Carlo* (MCMC) sampling procedure (see Madigan and York, 1995, and Gilks *et al.*, 1996, for a general introduction to MCMC sampling). However, it is not clear how these methods scale up for large domains.

Although recent developments in MCMC methods for Bayesian networks, such as Friedman and Koller (2000), show promise for scaling up, we choose here to use an alternative method as a “poor man’s” version of Bayesian analysis. An effective, and relatively simple, approach for estimating confidence is the *bootstrap* method (Efron and Tibshirani, 1993). The main idea behind the bootstrap is simple. We generate “perturbed” versions of our original data set and learn from them. In this way we collect many networks, all of which are fairly reasonable models of the data. These networks reflect the effect of small perturbations to the data on the learning process.

In our context, we use the bootstrap as follows:

- For  $i = 1 \dots m$ .  
Construct a data set  $D_i$  by sampling, with replacement,  $N$  instances from  $D$ .  
Apply the learning procedure on  $D_i$  to induce a network structure  $G_i$ .
- For each feature  $f$  of interest calculate

$$\text{conf}(f) = \frac{1}{m} \sum_{i=1}^m f(G_i)$$

where  $f(G)$  is 1 if  $f$  is a feature in  $G$ , and 0 otherwise.

See Friedman, Goldszmidt, and Wyner (1999) for more details, as well as for large-scale simulation experiments with this method. These simulation experiments show that features induced with high confidence are rarely false positives, even in cases where the data sets are small compared to the system being learned. This bootstrap procedure appears especially robust for the Markov and order features described in Section 3.2. In addition, simulation studies by Friedman and Koller (2000) show that the confidence values computed by the bootstrap, although not equal to the Bayesian posterior, correlate well with estimates of the Bayesian posterior for features.

### 3.4. Efficient learning algorithms

In Section 2.3, we formulated learning Bayesian network structure as an optimization problem in the space of directed acyclic graphs. The number of such graphs is superexponential in the number of variables. As we consider hundreds of variables, we must deal with an extremely large search space. Therefore, we need to use (and develop) efficient search algorithms.

To facilitate efficient learning, we need to be able to focus the attention of the search procedure on relevant regions of the search space, giving rise to the *Sparse Candidate* algorithm (Friedman, Nachman, and Pe’er, 1999). The main idea of this technique is that we can identify a relatively small number of *candidate* parents for each gene based on simple local statistics (such as correlation). We then restrict our search to networks in which only the candidate parents of a variable can be its parents, resulting in a much smaller search space in which we can hope to find a good structure quickly.

A possible pitfall of this approach is that early choices can result in an overly restricted search space. To avoid this problem, we devised an iterative algorithm that adapts the candidate sets during search. At each iteration  $n$ , for each variable  $X_i$ , the algorithm chooses the set  $C_i^n = \{Y_1, \dots, Y_k\}$  of variables which are the most promising *candidate parents* for  $X_i$ . We then search for  $G_n$ , a high scoring network in which  $\mathbf{Pa}^{G_n}(X_i) \subseteq C_i^n$ . (Ideally, we would like to find the highest-scoring network given the constraints, but since we are using a heuristic search, we do not have such a guarantee.) The network found is then used to guide the selection of better candidate sets for the next iteration. We ensure that the score of  $G_n$  monotonically improves in each iteration by requiring  $\mathbf{Pa}^{G_{n-1}}(X_i) \subseteq C_i^n$ . The algorithm continues until there is no change in the candidate sets.

We briefly outline our method for choosing  $C_i^n$ . We assign to each variable  $X_j$  some score of relevance to  $X_i$ , choosing variables with the highest score. The question is then how to measure the relevance of potential parent  $X_j$  to  $X_i$ . Friedman *et al.* (1999) examine several measures of relevance. Based on their experiments, one of the most successful measures is simply the improvement in the score of  $X_i$  if we add  $X_j$  as an additional parent. More precisely, we calculate

$$\text{ScoreContribution}(X_i, \mathbf{Pa}^{G_{n-1}}(X_i) \cup \{X_j\} : D) - \text{ScoreContribution}(X_i, \mathbf{Pa}^{G_{n-1}}(X_i) : D).$$

This quantity measures how much the inclusion of an edge from  $X_j$  to  $X_i$  can improve the score associated with  $X_i$ . We then choose the new candidate set to contain the previous parent set  $\mathbf{Pa}^{G_{n-1}}(X_i)$  and the variables that are most informative given this set of parents.

We refer the reader to Friedman, Nachman, and Pe'er (1999) for more details on the algorithm and its complexity, as well as empirical results comparing its performance to traditional search techniques.

### 3.5. Local probability models

In order to specify a Bayesian network model, we still need to choose the type of the local probability models we learn. In the current work, we consider two approaches:

- **Multinomial model.** In this model we treat each variable as discrete and learn a multinomial distribution that describes the probability of each possible state of the child variable given the state of its parents.
- **Linear Gaussian model.** In this model we learn a linear regression model for the child variable given its parents.

These models were chosen since their posterior can be efficiently calculated in closed form.

To apply the multinomial model we need to discretize the gene expression values. We choose to discretize these values into three categories: *underexpressed* ( $-1$ ), *normal* ( $0$ ), and *overexpressed*  $1$ , depending on whether the expression rate is significantly lower than, similar to, or greater than control, respectively. The control expression level of a gene either can be determined experimentally (as in the methods of DeRisi *et al.* (1997)) or it can be set as the average expression level of the gene across experiments. We discretize by setting a threshold to the ratio between measured expression and control. In our experiments, we choose a threshold value of  $0.5$  in logarithmic (base 2) scale. Thus, values with ratio to control lower than  $2^{-0.5}$  are considered underexpressed, and values higher than  $2^{0.5}$  are considered overexpressed.

Each of these two models has benefits and drawbacks. On one hand, it is clear that by discretizing the measured expression levels we are losing information. The linear-Gaussian model does not suffer from the information loss caused by discretization. On the other hand, the linear-Gaussian model can only detect dependencies that are close to linear. In particular, it is not likely to discover combinatorial effects (e.g., a gene is over-expressed only if several genes are jointly over-expressed, but not if at least one of them is not under-expressed). The multinomial model is more flexible and can capture such dependencies.

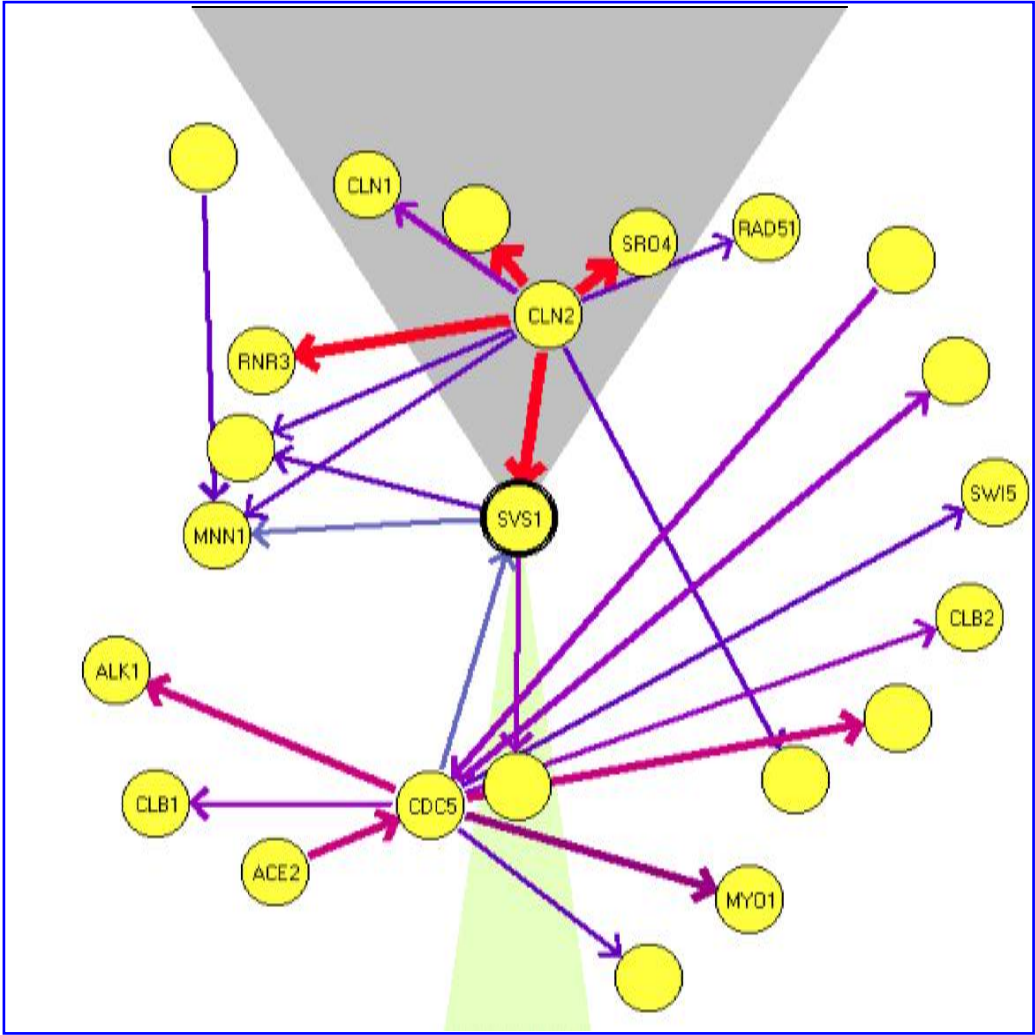
## 4. APPLICATION TO CELL CYCLE EXPRESSION PATTERNS

We applied our approach to the data of Spellman *et al.* (1998). This data set contains 76 gene expression measurements of the mRNA levels of 6177 *S. cerevisiae* ORFs. These experiments measure six time series under different cell cycle synchronization methods. Spellman *et al.* (1998) identified 800 genes whose expression varied over the different cell-cycle stages.

In learning from this data, we treat each measurement as an independent sample from a distribution and do not take into account the temporal aspect of the measurement. Since it is clear that the cell cycle process is of a temporal nature, we compensate by introducing an additional variable denoting the cell cycle phase. This variable is forced to be a root in all the networks learned. Its presence allows one to model dependency of expression levels on the current cell cycle phase.<sup>3</sup>

We used the Sparse Candidate algorithm with a 200-fold bootstrap in the learning process. We performed two experiments, one with the discrete multinomial distribution, the other with the linear Gaussian distribution. The learned features show that we can recover intricate structure even from such small data sets. It is important to note that our learning algorithm uses *no prior biological knowledge nor constraints*. All learned networks and relations are based solely on the information conveyed in the measurements themselves. These results are available at our WWW site: <http://www.cs.huji.ac.il/labs/compbio/expression>. Figure 2 illustrates the graphical display of some results from this analysis.

<sup>3</sup>We note that we can learn temporal models using a Bayesian network that includes gene expression values in two (or more) consecutive time points (Friedman *et al.*, 1998). This raises the number of variables in the model. We are currently perusing this issue.

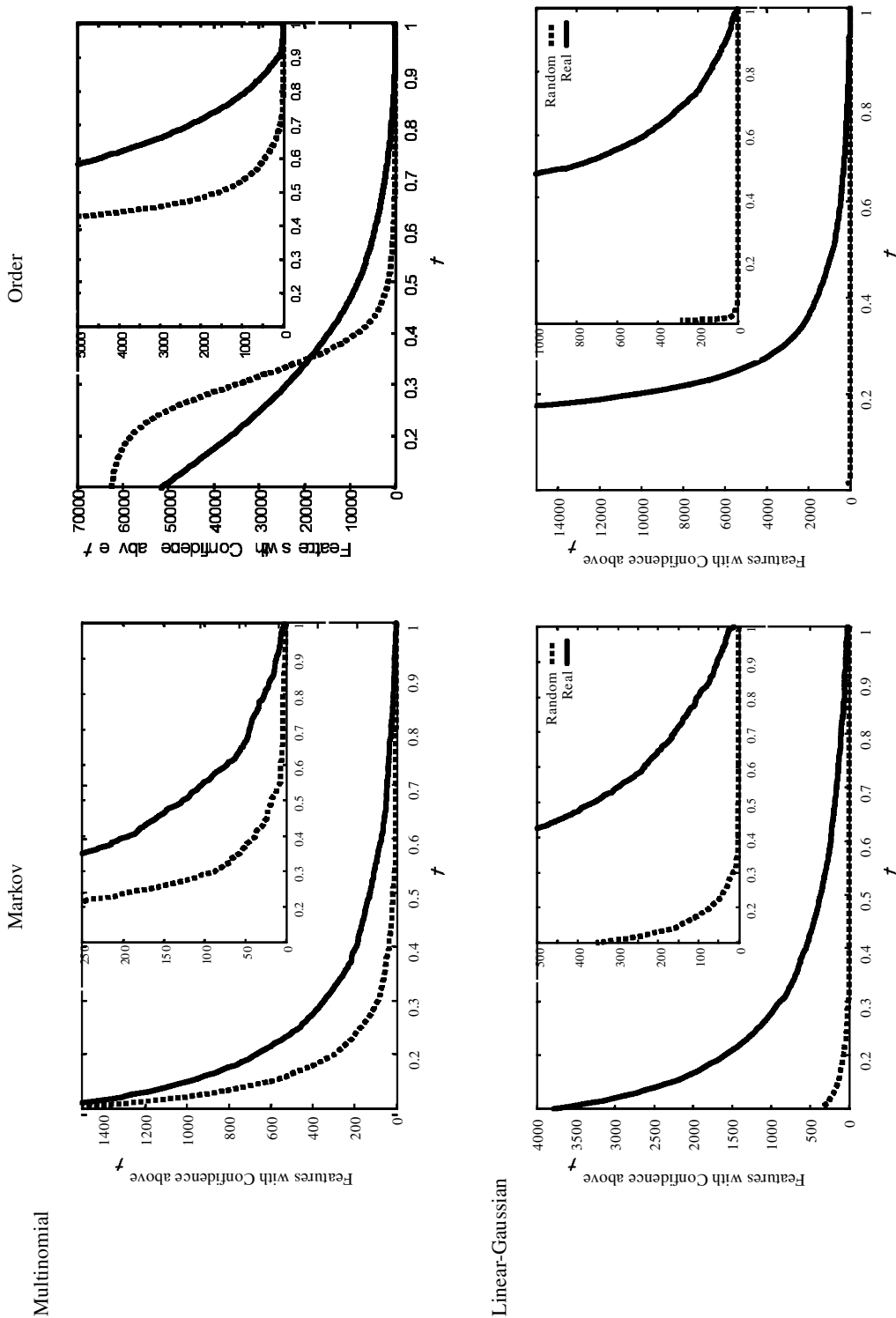


**FIG. 2.** An example of the graphical display of Markov features. This graph shows a “local map” for the gene SVS1. The width (and color) of edges corresponds to the computed confidence level. An edge is directed if there is a sufficiently high confidence in the order between the genes connected by the edge. This local map shows that CLN2 separates SVS1 from several other genes. Although there is a strong connection between CLN2 to all these genes, there are no other edges connecting them. This indicates that, with high confidence, these genes are conditionally independent given the expression level of CLN2.

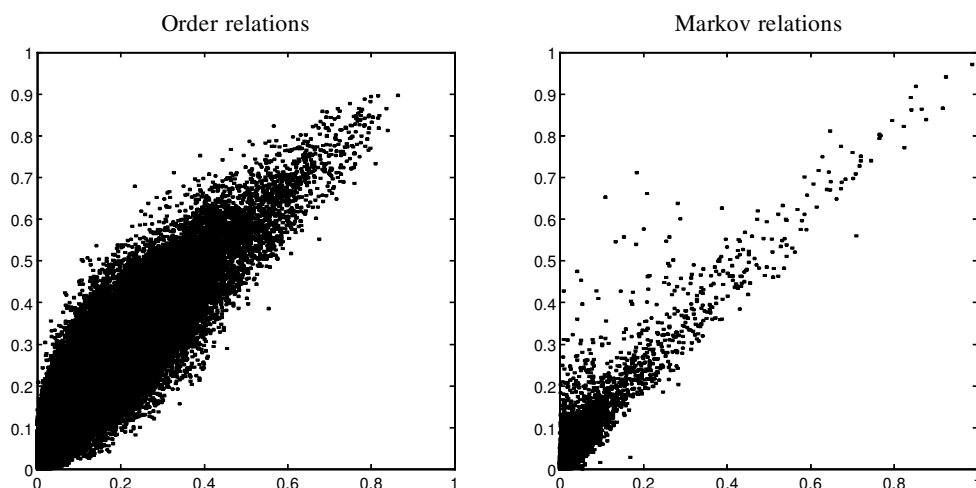
4.1. Robustness analysis

We performed a number of tests to analyze the statistical significance and robustness of our procedure. Some of these tests were carried out on a smaller data set with 250 genes for computational reasons.

To test the credibility of our confidence assessment, we created a random data set by randomly permuting the order of the experiments independently for each gene. Thus for each gene the order was random, but the composition of the series remained unchanged. In such a data set, genes are independent of each other, and thus we do not expect to find “real” features. As expected, both order and Markov relations in the random data set have significantly lower confidence. We compare the distribution of confidence estimates between the original data set and the randomized set in Figure 3. Clearly, the distribution of confidence estimates in the original data set have a longer and heavier tail in the high confidence region. In the linear-Gaussian model we see that random data does not generate any feature with confidence above 0.3. The multinomial model is more expressive and thus susceptible to over-fitting. For this model, we see a smaller gap between the two distributions. Nonetheless, randomized data does not generate any feature with confidence above 0.8, which leads us to believe that most features that are learned in the original data set with such confidence are not artifacts of the bootstrap estimation.



**FIG. 3.** Plots of abundance of features with different confidence levels for the cell cycle data set (solid line), and the randomized data set (dotted line). The x-axis denotes the confidence threshold, and the y-axis denotes the number of features with confidence equal or higher than the corresponding x-value. The graphs on the left column show Markov features, and the ones on the right column show Order features. The top row describes features found using the multinomial model, and the bottom row describes features found by the linear-Gaussian model. Inset in each graph is a plot of the tail of the distribution.



**FIG. 4.** Comparison of confidence levels obtained in two data sets differing in the number of genes, on the multinomial experiment. Each relation is shown as a point, with the  $x$ -coordinate being its confidence in the 250 genes data set and the  $y$ -coordinate the confidence in the 800 genes data set. The left figure shows order relation features, and the right figure shows Markov relation features.

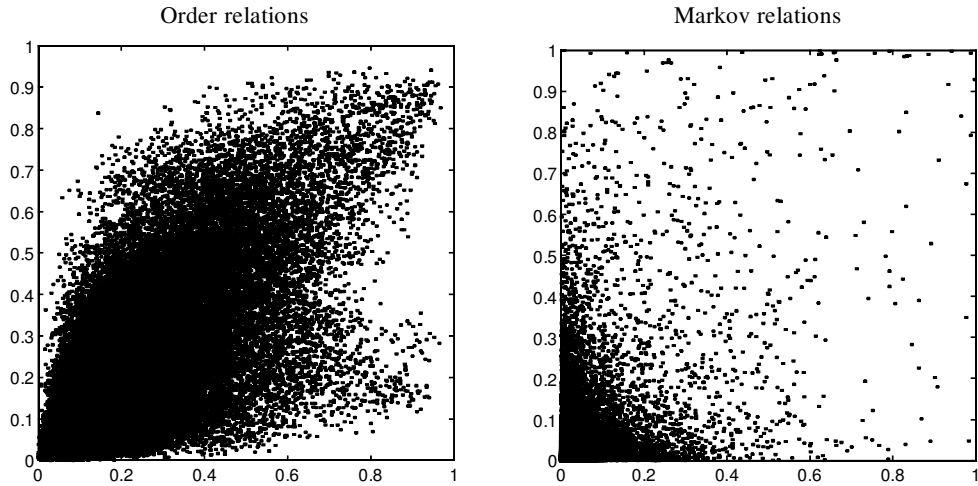
To test the robustness of our procedure for extracting dominant genes, we performed a simulation study. We created a data set by sampling 80 samples from one of the networks learned from the original data. We then applied the bootstrap procedure on this data set. We counted the number of descendents of each node in the synthetic network and ordered the variables according to this count. Of the 10 top dominant genes learned from the bootstrap experiment, 9 were among the top 30 in the real ordering.

Since the analysis was not performed on the whole *S. cerevisiae* genome, we also tested the robustness of our analysis to the addition of more genes, comparing the confidence of the learned features between the 800-gene data set and a smaller 250-gene data set that contains genes appearing in eight major clusters described by Spellman *et al.* (1998). Figure 4 compares feature confidence in the analysis of the two data sets for the multinomial model. As we can see, there is a strong correlation between confidence levels of the features between the two data sets. The comparison for the linear-Gaussian model gives similar results.

A crucial choice for the multinomial experiment is the threshold level used for discretization of the expression levels. It is clear that by setting a different threshold we would get different discrete expression patterns. Thus, it is important to test the robustness and sensitivity of the high-confidence features to the choice of this threshold. This was tested by repeating the experiments using different thresholds. The comparison of how the change of threshold affects the confidence of features show a definite linear tendency in the confidence estimates of features between the different discretization thresholds (graphs not shown). Obviously, this linear correlation gets weaker for larger threshold differences. We also note that order relations are much more robust to changes in the threshold than are Markov relations.

A valid criticism of our discretization method is that it penalizes genes whose natural range of variation is small: since we use a fixed threshold, we would not detect changes in such genes. A possible way to avoid this problem is to *normalize* the expression of genes in the data. That is, we rescale the expression level of each gene so that the relative expression level has the same mean and variance for all genes. We note that analysis methods that use *Pearson correlation* to compare genes, such as those of Ben-Dor *et al.* (1999) and Eisen *et al.* (1999), implicitly perform such a normalization.<sup>4</sup> When we discretize a normalized data set, we are essentially rescaling the discretization factor differently for each gene, depending on its variance in the data. We tried this approach with several discretization levels, and got results comparable

<sup>4</sup>An undesired effect of such a normalization is the amplification of measurement noise. If a gene has fixed expression levels across samples, we expect the variance in measured expression levels to be noise either in the experimental conditions or the measurements. When we normalize the expression levels of genes, we lose the distinction between such noise and true (i.e., significant) changes in expression levels. In the Spellman *et al.* data set we can safely assume this effect will not be too grave, since we only focus on genes that display significant changes across experiments.



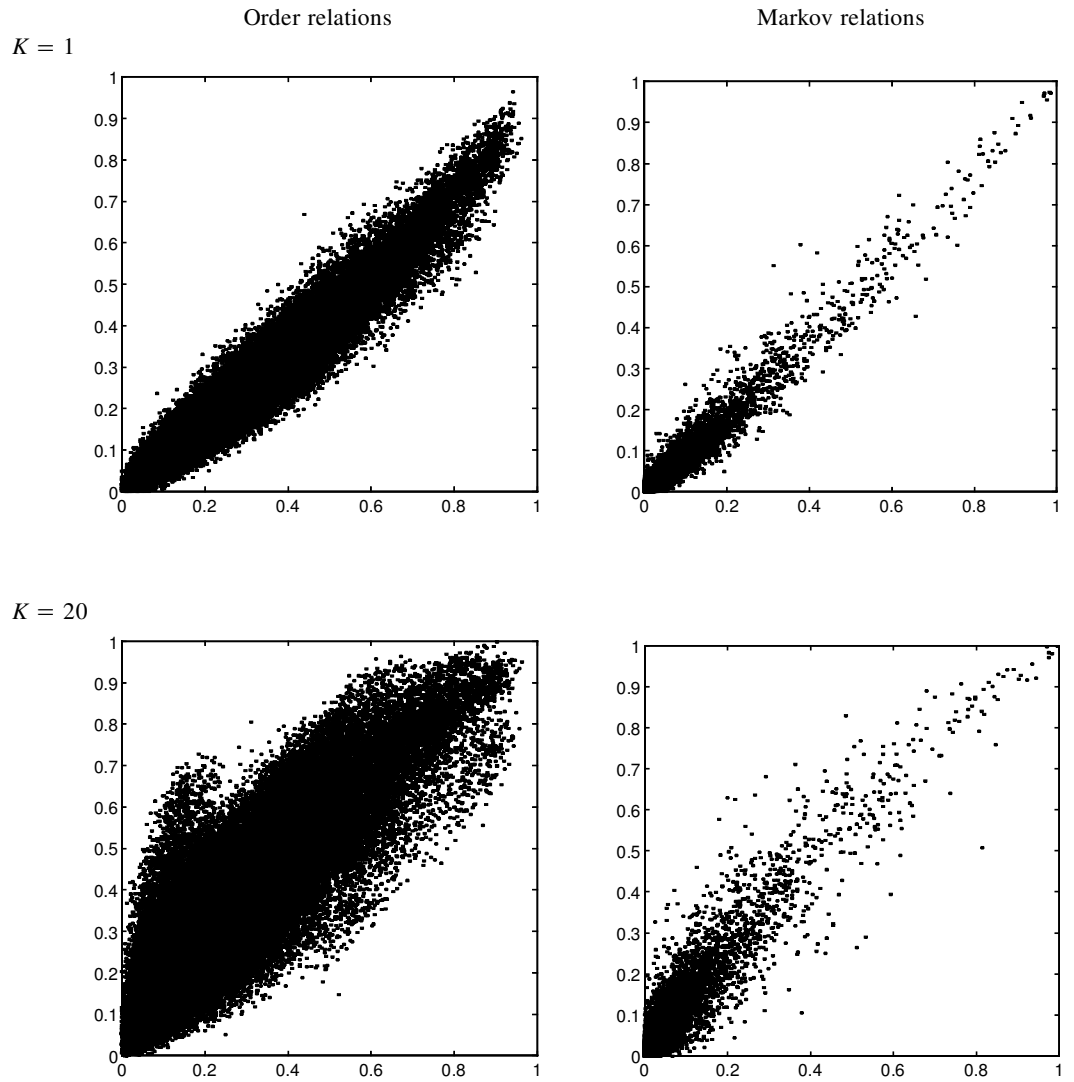
**FIG. 5.** Comparison of confidence levels between the multinomial experiment and the linear-Gaussian experiment. Each relation is shown as a point, with the  $x$ -coordinate being its confidence in the multinomial experiment, and the  $y$ -coordinate its confidence in the linear-Gaussian experiment. The left figure shows order relation features, and the right figure shows Markov relation features.

to our original discretization method. The 20 top Markov relations highlighted by this method were a bit different, but interesting and biologically sensible in their own right. The order relations were again more robust to the change of methods and discretization thresholds. A possible reason is that order relations depend on the network structure in a global manner and thus can remain intact even after many local changes to the structure. The Markov relation, being a local one, is more easily disrupted. Since the graphs learned are extremely sparse, each discretization method “highlights” different signals in the data, which are reflected in the Markov relations learned.

A similar picture arises when we compare the results of the multinomial experiment to those of the linear-Gaussian experiment (Figure 5). In this case, there is virtually no correlation between the Markov relations found by the two methods, while the order relations show some correlation. This supports our assumption that the two methods highlight different types of connections between genes.

Finally, we consider the effect of the choice of prior on the learned features. It is important to ensure that the learned features are not simply artifacts of the chosen prior. To test this, we repeated the multinomial experiment with different values of  $K$ , the effective sample size, and compared the learned confidence levels to those learned with the default value used for  $K$ , which was 5. This was done using the 250-gene data set and discretization level of 0.5. The results of these comparisons are shown in Figure 6. As can be seen, the confidence levels obtained with a  $K$  value of 1 correlate very well with those obtained with the default  $K$ , while when setting  $K$  to 20 the correlation is weaker. This suggests that both 1 and 5 are low enough values compared to the data set size of 76, making the prior’s affect on the results weak. An effective sample size of 20 is high enough to make the prior’s effect noticeable. Another aspect of the prior is the prior network used. In all the experiments reported here, we used the empty network with uniform distribution parameters as the prior network. As our prior is noninformative, keeping down its effect is desired. It is expected that once we use more informative priors (by incorporating biological knowledge, for example) and stronger effective sample sizes, the obtained results will be biased more toward our prior beliefs.

In summary, although many of the results we report below (especially order relations) are stable across the different experiments discussed in the previous paragraph, it is clear that our analysis is sensitive to the choice of local model and, in the case of the multinomial model, to the discretization method. It is probably less sensitive to the choice of prior, as long as the effective sample size is small compared to the data set size. In all the methods we tried, our analysis found interesting relationships in the data. Thus, one challenge is to find alternative methods that can recover all these relationships in one analysis. We are currently working on learning networks with semi-parametric density models (Friedman and Nachman, 2000; Hoffman and Tresp, 1996) that would circumvent the need for discretization on one hand and allow nonlinear dependency relations on the other.



**FIG. 6.** Comparison of confidence levels between runs using different parameter priors. The difference in the priors is in the effective sample size,  $K$ . Each relation is shown as a point, with the  $x$ -coordinate being its confidence in a run with  $K = 5$ , and the  $y$ -coordinate its confidence in a run with  $K = 1$  (top row) and  $K = 20$  (bottom row). The left figure shows order relation features, and the right figure shows Markov relation features. All runs are on the 250 gene set, using discretization with threshold level 0.5.

4.2. Biological analysis

We believe that the results of this analysis can be indicative of biological phenomena in the data. This is confirmed by our ability to predict sensible relations between genes of known function. We now examine several consequences that we have learned from the data. We consider, in turn, the order relations and Markov relations found by our analysis.

**4.2.1. Order relations.** The most striking feature of the high-confidence order relations, is the existence of *dominant genes*. Out of all 800 genes, only a few seem to dominate the order (i.e., appear before many genes). The intuition is that these genes are indicative of potential causal sources of the cell-cycle process. Let  $C_o(X, Y)$  denote the confidence in  $X$  being an ancestor of  $Y$ . We define the *dominance score* of  $X$  as  $\sum_{Y, C_o(X, Y) > t} C_o(X, Y)^k$ , using the constant  $k$  for rewarding high confidence features and the threshold  $t$  to discard low confidence ones. These dominant genes are extremely robust to parameter selection for both  $t$  and  $k$ , the discretization cutoff of Section 3.5, and the local probability model used. A list of the highest-scoring dominating genes for both experiments appears in Table 1.

TABLE 1. LIST OF DOMINANT GENES IN THE ORDERING RELATIONS<sup>1</sup>

Gene/ORF	Score in experiment		Notes
	Multinomial	Gaussian	
MCD1	550	525	Mitotic Chromosome Determinant, null mutant is inviable
MSH6	292	508	Required for mismatch repair in mitosis and meiosis
CSI2	444	497	Cell wall maintenance, chitin synthesis
CLN2	497	454	Role in cell cycle START, null mutant exhibits G1 arrest
YLR183C	551	448	Contains forkheaded associated domain, thus possibly nuclear
RFA2	456	423	Involved in nucleotide excision repair, null mutant is inviable
RSR1	352	395	GTP-binding protein of the RAS family involved in bud site selection
CDC45	—	394	Required for initiation of chromosomal replication, null mutant lethal
RAD53	60	383	Cell cycle control, checkpoint function, null mutant lethal
CDC5	209	353	Cell cycle control, required for exit from mitosis, null mutant lethal
POL30	376	321	Required for DNA replication and repair, null mutant is inviable
YOX1	400	291	Homeodomain protein
SRO4	463	239	Involved in cellular polarization during budding
CLN1	324	—	Role in cell cycle START, null mutant exhibits G1 arrest
YBR089W	298	—	

<sup>1</sup>Included are the top 10 dominant genes for each experiment.

Inspection of the list of dominant genes reveals quite a few interesting features. Among them are genes directly involved in initiation of the cell cycle and its control. For example, CLN1, CLN2, CDC5, and RAD53 whose functional relation has been established (Cvrckova and Nasmyth, 1993; Drebot *et al.*, 1993). The genes MCD1, RFA2, CDC45, RAD53, CDC5, and POL30 were found to be essential (Guacci *et al.*, 1997). These are clearly key genes in essential cell functions. Some of them are components of prereplication complexes(CDC45,POL30). Others (like RFA2,POL30, and MSH6) are involved in DNA repair. It is known that DNA repair is associated with transcription initiation, and DNA areas which are more active in transcription are also repaired more frequently (McGregor, 1999; Tornaletti and Hanawalk, 1999). Furthermore, a cell cycle control mechanism causes an abort when the DNA has been improperly replicated (Eisen and Lucchesi, 1998).

Most of the dominant genes encode nuclear proteins, and some of the unknown genes are also potentially nuclear (e.g., YLR183C contains a forkhead-associated domain which is found almost entirely among nuclear proteins). A few nonnuclear dominant genes are localized in the cytoplasm membrane (SRO4 and RSR1). These are involved in the budding and sporulation processes which have an important role in the cell cycle. RSR1 belongs to the RAS family of proteins, which are known as initiators of signal transduction cascades in the cell.

**4.2.2. Markov relations.** We begin with an analysis of the Markov relations in the multinomial experiment. Inspection of the top Markov relations reveals that most are functionally related. A list of the top scoring relations can be found in Table 2. Among these, all those involving two known genes make sense biologically. When one of the ORFs is unknown, careful searches using Psi-Blast (Altschul *et al.*, 1997), Pfam (Sonnhammer *et al.*, 1998), and Protomap (Yona *et al.*, 1998) can reveal firm homologies to proteins functionally related to the other gene in the pair. For example YHR143W, which is paired to the endochitinase CTS1, is related to EGT2—a cell-wall-maintenance protein. Several of the unknown pairs are physically adjacent on the chromosome and thus presumably regulated by the same mechanism (see Blumenthal (1998)), although special care should be taken for pairs whose chromosomal location overlap on complementary strands, since in these cases we might see an artifact resulting from cross-hybridization. Such an analysis raises the number of biologically sensible pairs to nineteen out of the twenty top relations.

Some interesting Markov relations were found that are beyond the limitations of clustering techniques. Among the high-confidence Markov relations, one can find examples of conditional independence, i.e., a group of highly correlated genes whose correlation can be explained within our network structure. One such example involves the genes CLN2,RNR3,SVS1,SRO4, and RAD51. Their expression is correlated, and in Spellman *et al.* (1998) they all appear in the same cluster. In our network CLN2 is with high confidence a parent of each of the other 4 genes, while no links are found between them (see Figure 2). This agrees with biological knowledge: CLN2 is a central and early cell cycle control, while there is



TABLE 2. LIST OF TOP MARKOV RELATIONS, MULTINOMIAL EXPERIMENT

Confidence	Gene 1	Gene 2	Notes
1.0	YKL163W-PIR3	YKL164C-PIR1	Close locality on chromosome
0.985	PRY2	YKR012C	Close locality on chromosome
0.985	MCD1	MSH6	Both bind to DNA during mitosis
0.98	PHO11	PHO12	Both nearly identical acid phosphatases
0.975	HHT1	HTB1	Both are histones
0.97	HTB2	HTA1	Both are histones
0.94	YNL057W	YNL058C	Close locality on chromosome
0.94	YHR143W	CTS1	Homolog to EGT2 cell wall control, both involved in cytokinesis
0.92	YOR263C	YOR264W	Close locality on chromosome
0.91	YGR086	SIC1	Homolog to mammalian nuclear ran protein, both involved in nuclear function
0.9	FAR1	ASH1	Both part of a mating type switch, <b>expression uncorrelated</b>
0.89	CLN2	SVS1	Function of SVS1 unknown
0.88	YDR033W	NCE2	Homolog to transmembrane proteins suggest both involved in protein secretion
0.86	STE2	MFA2	A mating factor and receptor
0.85	HHF1	HHF2	Both are histones
0.85	MET10	ECM17	Both are sulfite reductases
0.85	CDC9	RAD27	Both participate in Okazaki fragment processing

no clear biological relationship between the others. Some of the other Markov relations are intercluster pairing genes with low correlation in their expression. One such regulatory link is FAR1–ASH1: both proteins are known to participate in a mating-type switch. The correlation of their expression patterns is low and Spellman *et al.* (1998) cluster them into different clusters. When looking further down the list for pairs whose Markov relation confidence is high relative to their correlation, interesting pairs surface, for example, SAG1 and MF-ALPHA-1, a match between the factor that induces the mating process and an essential protein that participates in the mating process. Another match is LAC1 and YNL300W. LAC1 is a GPI transport protein and YNL300W is most likely modified by GPI (based on sequence homology).

The Markov relations from the Gaussian experiment are summarized in Table 3. Since the Gaussian model focuses on highly correlated genes, most of the high-scoring genes are tightly correlated. When we checked the DNA sequence of pairs of physically adjacent genes at the top of Table 3, we found that there is significant overlap. This suggests that these correlations are spurious and due to *cross hybridization*. Thus, we ignore the relations with the highest score. However, in spite of this technical problem, few of the pairs with a confidence > 0.8 can be discarded as biologically false.

Some of the relations are robust and also appear in the multinomial experiment (e.g., STE2–MFA2, CST1–YHR143W). Most interesting are the genes linked through regulation. These include: SHM2 which converts glycine that regulates GCV2 and DIP5 which transports glutamate which regulates ARO9. Some pairs participate in the same metabolic process, such as: CTS1–YHR143 and SRO4–YOL007C, all which participate in cell wall regulation. Other interesting high-confidence (> 0.9) examples are: OLE1–FAA4 linked through fatty acid metabolism, STE2–AGA2 linked through the mating process, and KIP3–MSB1, both playing a role in polarity establishment.

5. DISCUSSION AND FUTURE WORK

In this paper we presented a new approach for analyzing gene expression data that builds on the theory and algorithms for learning Bayesian networks. We described how to apply these techniques to gene expression data. The approach builds on two techniques that were motivated by the challenges posed by this domain: a novel search algorithm (Friedman, Nachman, and Pe’er, 1999) and an approach for estimating statistical confidence (Friedman, Goldszmidt, and Wyner, 1999). We applied our methods to the real expression data of Spellman *et al.* (1998). Although, we did not use any prior knowledge, we managed to extract many biologically plausible conclusions from this analysis.

TABLE 3. LIST OF TOP MARKOV RELATIONS, GAUSSIAN EXPERIMENT<sup>1</sup>

Confidence	Gene 1	Gene 2	Notes
1.0	YOR263C	YOR264W	Close locality on chromosome
1.0	CDC46	YOR066W	YOR066W is totally unknown
1.0	CDC45	SPH1	No suggestion for immediate link
1.0	SHM2	GCV2	SHM2 interconverts glycine, GCV2 is regulated by glycine
1.0	MET3	ECM17	MET3 required to convert sulfate to sulfide, ECM17 sulfite reductase
1.0	YJL194W-CDC6	YJL195C	Close locality on chromosome
1.0	YGR151C	YGR152C	Close locality on chromosome
1.0	YGR151C	YGR152C-RSR1	Close locality on chromosome
1.0	STE2	MFA2	A mating factor and receptor
1.0	YDL037C	YDL039C	Both homologs to mucin proteins
1.0	YCL040W-GLK1	WCL042C	Close locality on chromosome
1.0	HTA1	HTA2	Two physically linked histones
...			
0.99	HHF2	HHT2	both histones
0.99	YHR143W	CTS1	Homolog to EGT2 cell wall control, both involved in cytokinesis
0.99	ARO9	DIP5	DIP5 transports glutamate which regulates ARO9
0.975	SRO4	YOL007C	Both proteins are involved in cell wall regulation at the plasma membrane

<sup>1</sup>The table skips over 5 additional pairs with which close locality.

Our approach is quite different from the clustering approach used by Alon *et al.* (1999), Ben-Dor *et al.* (1999), Eisen *et al.* (1999), Michaels *et al.* (1998), and Spellman *et al.* (1998) in that it attempts to learn a much richer structure from the data. Our methods are capable of discovering causal relationships, interactions between genes other than positive correlation, and finer intracluster structure. We are currently developing hybrid approaches that combine our methods with clustering algorithms to learn models over “clustered” genes.

The biological motivation of our approach is similar to the work on inducing *genetic networks* from data (Akutsu *et al.*, 1998; Chen *et al.*, 1999; Somogyi *et al.*, 1996; Weaver *et al.*, 1999). There are two key differences: First, the models we learn have probabilistic semantics. This better fits the stochastic nature of both the biological processes and noisy experiments. Second, our focus is on extracting features that are pronounced in the data, in contrast to current genetic network approaches that attempt to find a single model that explains the data.

We emphasize that the work described here represents a preliminary step in a longer-term project. As we have seen above, there are several points that require accurate statistical tools and more efficient algorithms. Moreover, the exact biological conclusions one can draw from this analysis are still not well understood. Nonetheless, we view the results described in Section 4 as definitely encouraging.

We are currently working on improving methods for expression analysis by expanding the framework described in this work. Promising directions for such extensions are: (a) developing the theory for learning local probability models that are suitable for the type of interactions that appear in expression data; (b) improving the theory and algorithms for estimating confidence levels; (c) incorporating biological knowledge (such as possible regulatory regions) as prior knowledge to the analysis; (d) improving our search heuristics; (e) learning temporal models, such as *Dynamic Bayesian Networks* (Friedman *et al.*, 1998), from temporal expression data, and (f) developing methods that discover *hidden variables* (e.g., protein activation).

Finally, one of the most exciting longer-term prospects of this line of research is discovering causal patterns from gene expression data. We plan to build on and extend the theory for learning causal relations from data and apply it to gene expression. The theory of causal networks allows learning from both observational and *interventional* data, where the experiment intervenes with some causal mechanisms of the observed system. In the gene expression context, we can model knockout/over-expressed mutants as such interventions. Thus, we can design methods that deal with mixed forms of data in a principled manner (see Cooper and Yoo (1999) for recent work in this direction). In addition, this theory can provide

tools for *experimental design*, that is, understanding which interventions are deemed most informative to determining the causal structure in the underlying system.

## ACKNOWLEDGMENTS

The authors are grateful to Gill Bejerano, Hadar Benyaminy, David Engelberg, Moises Goldszmidt, Daphne Koller, Matan Ninio, Itzik Pe'er, Gavin Sherlock, and the anonymous reviewer for comments on drafts of this paper and useful discussions relating to this work. We also thank Matan Ninio for help in running and analyzing the robustness experiments. This work was supported in part through the generosity of the Michael Sacher Trust and by Israeli Science Foundation grant 224/99-2. Nir Friedman was also supported by an Alon Fellowship.

## REFERENCES

- Akutsu, S., Kuhara, T., Maruyama, O., and Minyano, S. 1998. Identification of gene regulatory networks by strategic gene disruptions and gene over-expressions. *Proc. Ninth Annual ACM-SIAM Symposium on Discrete Algorithms*, ACM-SIAM.
- Alon, U., Barkai, N., Notterman, D., Gish, K., Ybarra, S., Mack, D., and Levine, A.J. 1999. Broad patterns of gene expression revealed by clustering analysis of tumor and normal colon tissues probed by oligonucleotide arrays. *Proc. Nat. Acad. Sci. USA* 96, 6745–6750.
- Altschul, S., Thomas, L., Schaffer, A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D. 1997. Gapped blast and psi-blast: A new generation of protein database search programs. *Nucl. Acids Res.* 25.
- Ben-Dor, A., Shamir, R., and Yakhini, Z. 1999. Clustering gene expression patterns. *J. Comp. Biol.* 6, 281–297.
- Blumenthal, T. 1998. Gene clusters and polycistronic transcription in eukaryotes. *Bioessays* 20, 480–487.
- Buntine, W. 1991. Theory refinement on Bayesian networks. *Proc. of the Seventh Annual Conference on Uncertainty in AI (UAI)*, 52–60.
- Chen, T., Filkov, V., and Skiena, S. 1999. Identifying gene regulatory networks from experimental data. *Proc. Third Annual International Conference on Computational Molecular Biology (RECOMB)*, 94–103.
- Chickering, D.M. 1995. A transformational characterization of equivalent Bayesian network structures. *Proc. Eleventh Conference on Uncertainty in Artificial Intelligence (UAI '95)*, 87–98.
- Chickering, D.M. 1996. Learning Bayesian networks is NP-complete, in D. Fisher and H.-J. Lenz, eds. *Learning from Data: Artificial Intelligence and Statistics V*, Springer-Verlag, New York.
- Cooper, G.F., and Herskovits, E. 1992. A Bayesian method for the induction of probabilistic networks from data. *Machine Learning* 9, 309–347.
- Cooper, G., and Glymour, C., eds. 1999. *Computation, Causation, and Discovery*, MIT Press, Cambridge, MA.
- Cooper, G., and Yoo, C. 1999. Causal discovery from a mixture of experimental and observational data. *Proc. Fifteenth Conference on Uncertainty in Artificial Intelligence (UAI '99)*, 116–125.
- Cvrckova, F., and Nasmyth, K. 1993. Yeast G1 cyclins CLN1 and CLN2 and a GAP-like protein have a role in bud formation. *EMBO J.* 12, 5277–5286.
- DeRisi, J., Iyer, V., and Brown, P. 1997. Exploring the metabolic and genetic control of gene expression on a genomic scale. *Science* 282, 699–705.
- Drebot, M.A., Johnston, G.C., Friesen, J.D., and Singer, R.A. 1993. An impaired RNA polymerase II activity in *Saccharomyces cerevisiae* causes cell-cycle inhibition at START. *Mol. Gen. Genet.* 241, 327–334.
- Efron, B., and Tibshirani, R.J. 1993. *An Introduction to the Bootstrap*, Chapman and Hall, London.
- Eisen, A., and Luchesi, J. 1998. Unraveling the role of helicases in transcription. *Bioessays* 20, 634–641.
- Eisen, M., Spellman, P., Brown, P., and Botstein, D. 1998. Cluster analysis and display of genome-wide expression patterns. *Proc. Nat. Acad. Sci. USA* 95, 14863–14868.
- Friedman, N., Goldszmidt, M., and Wyner, A. 1999. Data analysis with Bayesian networks: A bootstrap approach. *Proc. Fifteenth Conference on Uncertainty in Artificial Intelligence (UAI '99)*, 206–215.
- Friedman, N., and Koller, D. 2000. Being Bayesian about network structure. *Proc. Sixteenth Conference on Uncertainty in Artificial Intelligence (UAI '00)*, 201–210.
- Friedman, N., Murphy, K., and Russell, S. 1998. Learning the structure of dynamic probabilistic networks. *Proc. Fourteenth Conference on Uncertainty in Artificial Intelligence (UAI '98)*, 139–147.
- Friedman, N., and Nachman, D. 2000. Gaussian process networks. *Proc. Sixteenth Conference on Uncertainty in Artificial Intelligence (UAI '00)*, 211–219.
- Friedman, N., Nachman, I., and Pe'er, D. 1999. Learning Bayesian network structure from massive datasets: The “sparse candidate” algorithm. *Proc. Fifteenth Conference on Uncertainty in Artificial Intelligence (UAI '99)*, 196–205.

- Friedman, N., and Yakhini, Z. 1996. On the sample complexity of learning Bayesian networks. *Proc. Twelfth Conference on Uncertainty in Artificial Intelligence (UAI '96)*, 274–282.
- Geiger, D., and Heckerman, D. 1994. Learning Gaussian networks. *Proc. Tenth Conference on Uncertainty in Artificial Intelligence (UAI '94)*, 235–243.
- Gilks, W., Richardson, S., and Spiegelhalter, D. 1996. *Markov Chain Monte Carlo Methods in Practice*, CRC Press.
- Guacci, V., Koshland, D., and Strunnikov, A. 1997. A direct link between sister chromatid cohesion and chromosome condensation revealed through the analysis of MCD1 in *S. cerevisiae*. *Cell* 91(1), 47–57.
- Heckerman, D. 1998. A tutorial on learning with Bayesian networks, in M.I. Jordan, ed., *Learning in Graphical Models*, Kluwer, Dordrecht, Netherlands.
- Heckerman, D., and Geiger, D. 1995. Learning Bayesian networks: a unification for discrete and Gaussian domains. *Proc. Eleventh Conference on Uncertainty in Artificial Intelligence (UAI '95)*, 274–284.
- Heckerman, D., Geiger, D., and Chickering, D.M. 1995. Learning Bayesian networks: The combination of knowledge and statistical data. *Machine Learning* 20, 197–243.
- Heckerman, D., Meek, C., and Cooper, G. 1999. A Bayesian approach to causal discovery, in *Cooper and Glymour*, 141–166.
- Hoffman, R., and Tresp, V. 1996. Discovering structure in continuous variables using Bayesian networks, in *Advances in Neural Information Processing Systems 8 (NIPS '96)*, MIT Press, Cambridge, MA.
- Iyer, V., Eisen, M., Ross, D., Schuler, G., Moore, T., Lee, J., Trent, J., Staudt, L., Hudson, J., Boguski, M., Lashkari, D., Shalon, D., Botstein, D., and Brown, P. 1999. The transcriptional program in the response of human fibroblasts to serum. *Science* 283, 83–87.
- Jensen, F.V. 1996. *An introduction to Bayesian Networks*, University College London Press, London.
- Lauritzen, S.L., and Wermuth, N. 1989. Graphical models for associations between variables, some of which are qualitative and some quantitative. *Annals of Statistics* 17, 31–57.
- Lockhart, D.J., Dong, H., Byrne, M.C., Follettie, M.T., Gallo, M.V., Chee, M.S., Mittmann, M., Want, C., Kobayashi, M., Horton, H., and Brown, E.L. 1996. DNA expression monitoring by hybridization of high density oligonucleotide arrays. *Nature Biotechnology* 14, 1675–1680.
- Madigan, D., and Raftery, E. 1994. Model selection and accounting for model uncertainty in graphical models using Occam's window. *J. Am. Stat. Assoc.* 89, 1535–1546.
- Madigan, D., and York, J. 1995. Bayesian graphical models for discrete data. *Inter. Stat. Rev.* 63, 215–232.
- McGregor, W.G. 1999. DNA repair, DNA replication, and UV mutagenesis. *J. Investig. Dermatol. Symp. Proc.* 4, 1–5.
- Michaels, G., Carr, D., Askenazi, M., Fuhrman, S., Wen, X., and Somogyi, R. 1998. Cluster analysis and data visualization for large scale gene expression data. *Pac. Symp. Biocomputing*, 42–53.
- Pearl, J. 1988. *Probabilistic Reasoning in Intelligent Systems*, Morgan Kaufmann, San Francisco.
- Pearl, J. 2000. *Causality: Models, Reasoning, and Inference*, Cambridge University Press, London.
- Pearl, J., and Verma, T.S. 1991. A theory of inferred causation, in *Principles of Knowledge Representation and Reasoning: Proc. Second International Conference (KR '91)*, 441–452.
- Somogyi, R., Fuhrman, S., Askenazi, M., and Wuensche, A. 1996. The gene expression matrix: Towards the extraction of genetic network architectures, in *The Second World Congress of Nonlinear Analysts (WCNA)*.
- Sonnhammer, E.L., Eddy, S., Birney, E., Bateman, A., and Durbin, R. 1998. Pfam: multiple sequence alignments and hmm-profiles of protein domains. *Nucl. Acids Res.* 26, 320–322.
- Spellman, P., Sherlock, G., Zhang, M., Iyer, V., Anders, K., Eisen, M., Brown, P., Botstein, D., and Futcher, B. 1998. Comprehensive identification of cell cycle-regulated genes of the yeast *Saccharomyces cerevisiae* by microarray hybridization. *Mol. Biol. Cell* 9, 3273–3297.
- Spirtes, P., Glymour, C., and Scheines, R. 1993. *Causation, Prediction, and Search*, Springer-Verlag, New York.
- Spirtes, P., Meek, C., and Richardson, T. 1999. An algorithm for causal inference in the presence of latent variables and selection bias, in *Cooper and Glymour*, 211–252.
- Tornaletti, S., and Hanawalt, P.C. 1999. Effect of DNA lesions on transcription elongation. *Biochimie* 81, 139–146.
- Weaver, D., Workman, C., and Stormo, G. 1999. Modeling regulatory networks with weight matrices, in *Pac. Symp. Biocomputing*, 112–123.
- Wen, X., Fuhrmann, S., Micheals, G.S., Carr, D.B., Smith, S., Barker, J.L., and Somogyi, R. 1998. Large-scale temporal gene expression mapping of central nervous system development. *Proc. Nat. Acad. Sci. USA* 95, 334–339.
- Yona, G., Linial, N., and Linial, M. 1998. Protomap-automated classification of all protein sequences: A hierarchy of protein families, and local maps of the protein space. *Proteins: Structure, Function, and Genetics* 37, 360–378.

Address correspondence to:

Nir Friedman

School of Computer Science and Engineering

Hebrew University

Jerusalem, 91904, Israel

E-mail: nir@cs.huji.ac.il

## This article has been cited by:

1. Junichi Mori, Vladimir Mahalec, Jie Yu. 2014. Identification of probabilistic graphical network model for root-cause diagnosis in industrial processes. *Computers & Chemical Engineering* **71**, 171-209. [[CrossRef](#)]
2. Yazhuo Liu, José L. Zayas-Castro, Peter Fabri, Shuai Huang. 2014. Learning high-dimensional networks with nonlinear interactions by a novel tree-embedded graphical model. *Pattern Recognition Letters* **49**, 207-213. [[CrossRef](#)]
3. Y. Kim, J.-H. Jang, S. Choi, D. Hwang. 2014. TEMPI: probabilistic modeling time-evolving differential PPI networks with multiPle information. *Bioinformatics* **30**, i453-i460. [[CrossRef](#)]
4. N. Misra, E. Szczurek, M. Vingron. 2014. Inferring the paths of somatic evolution in cancer. *Bioinformatics* **30**, 2456-2463. [[CrossRef](#)]
5. A. Fukushima, M. Kusano. 2014. A network perspective on nitrogen metabolism from model to crop plants using integrated 'omics' approaches. *Journal of Experimental Botany* . [[CrossRef](#)]
6. Alireza Tamaddon-Nezhad, Dianhuan Lin, Hiroaki Watanabe, Jianzhong Chen, Stephen Muggleton Machine Learning of Biological Networks Using Abductive ILP 363-401. [[CrossRef](#)]
7. Nataly Maimari, Krysia Broda, Antonis Kakas, Rob Krams, Alessandra Russo Symbolic Representation and Inference of Regulatory Network Structures 1-48. [[CrossRef](#)]
8. Jianjun Yang, Yunhai Tong, Zitian Wang, Shaohua Tan. 2014. Efficient and effective Bayesian network local structure learning. *Frontiers of Computer Science* **8**, 527-536. [[CrossRef](#)]
9. Chengyi Tu. 2014. A Dynamical method to estimate gene regulatory networks using time-series data. *Complexity* n/a-n/a. [[CrossRef](#)]
10. Emmanuel S. Adabor, George K. Acquaaah-Mensah, Francis T. Oduro. 2014. SAGA: A hybrid search algorithm for Bayesian Network structure learning of transcriptional regulatory networks. *Journal of Biomedical Informatics* . [[CrossRef](#)]
11. Yiming Zuo, Guoqiang Yu, Mahlet G. Tadesse, Habtom W. Ressom. 2014. Biological network inference using low order partial correlation. *Methods* . [[CrossRef](#)]
12. Jinliang Liu, Engang Tian, Zhou Gu, Yuanyuan Zhang. 2014. State estimation for Markovian jumping genetic regulatory networks with random delays. *Communications in Nonlinear Science and Numerical Simulation* **19**, 2479-2492. [[CrossRef](#)]
13. Réka Albert, Juilee Thakar. 2014. Boolean modeling: a logic-based dynamic approach for understanding signaling and regulatory networks and for making useful predictions. *Wiley Interdisciplinary Reviews: Systems Biology and Medicine* n/a-n/a. [[CrossRef](#)]
14. Fabrício M. Lopes, David C. Martins, Junior Barrera, Roberto M. Cesar. 2014. A feature selection technique for inference of graphs from their known topological properties: Revealing scale-free gene regulatory networks. *Information Sciences* **272**, 1-15. [[CrossRef](#)]
15. Yongqun He. 2014. Ontology-supported research on vaccine efficacy, safety and integrative biological networks. *Expert Review of Vaccines* **13**, 825-841. [[CrossRef](#)]
16. Yunjun Zhang, Qi Ouyang, Zhi Geng. 2014. Refining network reconstruction based on functional reliability. *Journal of Theoretical Biology* **353**, 170-178. [[CrossRef](#)]
17. Tao Lu. 2014. Investigate Data Dependency for Dynamic Gene Regulatory Network Identification through High Dimensional Differential Equation Approach. *Communications in Statistics - Simulation and Computation* 00-00. [[CrossRef](#)]
18. Riten Mitra, Peter Müller, Yuan Ji, Yitan Zhu, Gordon Mills, Yiling Lu. 2014. A Bayesian hierarchical model for inference across related reverse phase protein arrays experiments. *Journal of Applied Statistics* 1-10. [[CrossRef](#)]
19. Zhi Li, Ke Chen. 2014. Exponential Stability of Stochastic Genetic Regulatory Networks with Interval Uncertainties and Multiple Delays. *Arabian Journal for Science and Engineering* . [[CrossRef](#)]

20. S.- J. Dunn, G. Martello, B. Yordanov, S. Emmott, A. G. Smith. 2014. Defining an essential transcription factor program for naive pluripotency. *Science* **344**, 1156-1160. [[CrossRef](#)]
21. Tanja Zerenner, Petra Friederichs, Klaus Lehnertz, Andreas Hense. 2014. A Gaussian graphical model approach to climate networks. *Chaos: An Interdisciplinary Journal of Nonlinear Science* **24**, 023103. [[CrossRef](#)]
22. Rohit Kannan, Arun K. Tangirala. 2014. Correntropy-based partial directed coherence for testing multivariate Granger causality in nonlinear processes. *Physical Review E* **89**. . [[CrossRef](#)]
23. Judea Pearl Graphical Models for Probabilistic and Causal Reasoning 1-24. [[CrossRef](#)]
24. Lian En Chai, Swee Kuan Loh, Swee Thing Low, Mohd Saberi Mohamad, Safaai Deris, Zalmiyah Zakaria. 2014. A review on the computational approaches for gene regulatory network construction. *Computers in Biology and Medicine* **48**, 55-65. [[CrossRef](#)]
25. Gianfranco Politano, Alessandro Savino, Alfredo Benso, Stefano Di Carlo, Hafeez Ur Rehman, Alessandro Vasciaveo. 2014. Using Boolean networks to model post-transcriptional regulation in gene regulatory networks. *Journal of Computational Science* **5**, 332-344. [[CrossRef](#)]
26. Xiaoqing Yu, Hongyun Gao, Xiaoqi Zheng, Chun Li, Jun Wang. 2014. A computational method of predicting regulatory interactions in Arabidopsis based on gene expression data and sequence information. *Computational Biology and Chemistry* . [[CrossRef](#)]
27. Federico Schlüter, Facundo Bromberg, Alejandro Edera. 2014. The IBCMAP approach for Markov network structure learning. *Annals of Mathematics and Artificial Intelligence* . [[CrossRef](#)]
28. D Jayachandran, U Ramkrishna, J Skiles, J Renbarger, D Ramkrishna. 2014. Revitalizing Personalized Medicine: Respecting Biomolecular Complexities Beyond Gene Expression. *CPT: Pharmacometrics & Systems Pharmacology* **3**, e110. [[CrossRef](#)]
29. Hulin Wu, Tao Lu, Hongqi Xue, Hua Liang. 2014. Sparse Additive Ordinary Differential Equations for Dynamic Gene Regulatory Network Modeling. *Journal of the American Statistical Association* **109**, 700-716. [[CrossRef](#)]
30. Y.X. Rachel Wang, Haiyan Huang. 2014. Review on statistical methods for gene network reconstruction using expression data. *Journal of Theoretical Biology* . [[CrossRef](#)]
31. Spencer Angus Thomas, Yaochu Jin. 2014. Reconstructing biological gene regulatory networks: where optimization meets big data. *Evolutionary Intelligence* **7**, 29-47. [[CrossRef](#)]
32. Rui Li, Meng Yang, Tianguang Chu. 2014. State feedback stabilization for probabilistic Boolean networks. *Automatica* **50**, 1272-1278. [[CrossRef](#)]
33. S. Isci, H. Dogan, C. Ozturk, H. H. Otu. 2014. Bayesian network prior: network analysis of biological data using external knowledge. *Bioinformatics* **30**, 860-867. [[CrossRef](#)]
34. S Aburatani, H Toh. 2014. Network inference of AP pattern formation system in D.melanogaster by structural equation modeling. *Journal of Physics: Conference Series* **490**, 012145. [[CrossRef](#)]
35. Vikas Reddy, Anna Charisse Farr, Paul Wu, Kerrie Mengersen, Prasad K D V Yarlagadda. 2014. An Intuitive Dashboard for Bayesian Network Inference. *Journal of Physics: Conference Series* **490**, 012023. [[CrossRef](#)]
36. Y. Kim, S. Han, S. Choi, D. Hwang. 2014. Inference of dynamic networks using time-course data. *Briefings in Bioinformatics* **15**, 212-228. [[CrossRef](#)]
37. Ali Masoudi-Nejad, Gholamreza Bidkhori, Saman Hosseini Ashtiani, Ali Najafi, Joseph H. Bozorgmehr, Edwin Wang. 2014. Cancer Systems Biology and Modeling: microscopic scale and multiscale approaches. *Seminars in Cancer Biology* . [[CrossRef](#)]
38. Kassandra Fronczyk, Michele Guindani, Marina Vannucci, Annalisa Palange, Paolo Decuzzi. 2014. A Bayesian hierarchical model for maximizing the vascular adhesion of nanoparticles. *Computational Mechanics* **53**, 539-547. [[CrossRef](#)]

39. S. Dumcke, J. Brauer, B. Anchang, R. Spang, N. Beerenwinkel, A. Tresch. 2014. Exact likelihood computation in Boolean networks with probabilistic time delays, and its application in signal network reconstruction. *Bioinformatics* **30**, 414-419. [[CrossRef](#)]
40. Y. Rahmatallah, F. Emmert-Streib, G. Glazko. 2014. Gene Sets Net Correlations Analysis (GSNCA): a multivariate differential coexpression test for gene sets. *Bioinformatics* **30**, 360-368. [[CrossRef](#)]
41. Aarash Bordbar, Jonathan M. Monk, Zachary A. King, Bernhard O. Palsson. 2014. Constraint-based models predict metabolic and associated cellular functions. *Nature Reviews Genetics* . [[CrossRef](#)]
42. Ya-Li Hsiao, Pei-Chun Chang, Hung-Jin Huang, Chia-Chen Kuo, Calvin Yu-Chian Chen. 2014. Treatment of Acute Lymphoblastic Leukemia from Traditional Chinese Medicine. *Evidence-Based Complementary and Alternative Medicine* **2014**, 1-21. [[CrossRef](#)]
43. Alfredo Benso, Stefano Carlo, Gianfranco Politano, Alessandro Savino, Alessandro Vasciaveo. 2014. An extended gene protein/products boolean network model including post-transcriptional regulation. *Theoretical Biology and Medical Modelling* **11**, S5. [[CrossRef](#)]
44. Zhenzhen Zheng, Scott Christley, William T Chiu, Ira L Blitz, Xiaohui Xie, Ken WY Cho, Qing Nie. 2014. Inference of the *Xenopus tropicalis* embryonic regulatory network and spatial gene expression patterns. *BMC Systems Biology* **8**, 3. [[CrossRef](#)]
45. Changwon Yoo, Luis Ramirez, Juan Liuzzi. 2014. Big Data Analysis Using Modern Statistical and Machine Learning Methods in Medicine. *International Neurourology Journal* **18**, 50. [[CrossRef](#)]
46. Swarup Roy, Dhruba K Bhattacharyya, Jugal K Kalita. 2014. Reconstruction of gene co-expression network from microarray data using local expression patterns. *BMC Bioinformatics* **15**, S10. [[CrossRef](#)]
47. Jehandad Khan, Nidhal Bouaynaya, Hassan M Fathallah-Shaykh. 2014. Tracking of time-varying genomic regulatory networks with a LASSO-Kalman smoother. *EURASIP Journal on Bioinformatics and Systems Biology* **2014**, 3. [[CrossRef](#)]
48. Paola Vera-Licona, Abdul Jarrah, Luis D. Garcia-Puente, John McGee, Reinhard Laubenbacher. 2014. An Algebra-Based Method for Inferring Gene Regulatory Networks. *BMC Systems Biology* **8**, 37. [[CrossRef](#)]
49. Jun Li, Manfeng Hu, Jinde Cao, Yongqing Yang, Yinghua Jin. 2014. Stability of Uncertain Impulsive Stochastic Genetic Regulatory Networks with Time-Varying Delay in the Leakage Term. *Abstract and Applied Analysis* **2014**, 1-15. [[CrossRef](#)]
50. Seiya Imoto, Hiroshi Matsuno, Satoru MiyanoGene Networks: Estimation, Modeling, and Simulation 89-112. [[CrossRef](#)]
51. Xun Huang, Zhike Zi. 2014. Inferring cellular regulatory networks with Bayesian model averaging for linear regression (BMALR). *Molecular BioSystems* **10**, 2023. [[CrossRef](#)]
52. Natsu Nakajima, Tatsuya Akutsu. 2014. Exact and Heuristic Methods for Network Completion for Time-Varying Genetic Networks. *BioMed Research International* **2014**, 1-13. [[CrossRef](#)]
53. Hsin-Chieh Tang, Calvin Yu-Chian Chen. 2014. Investigation of the Novel Lead of Melanocortin 1 Receptor for Pigmentary Disorders. *Evidence-Based Complementary and Alternative Medicine* **2014**, 1-13. [[CrossRef](#)]
54. Natsu Nakajima, Tatsuya Akutsu. 2014. Network Completion for Static Gene Expression Data. *Advances in Bioinformatics* **2014**, 1-9. [[CrossRef](#)]
55. Tingting Yu, Jing Wang, Xian Zhang. 2014. A Less Conservative Stability Criterion for Delayed Stochastic Genetic Regulatory Networks. *Mathematical Problems in Engineering* **2014**, 1-11. [[CrossRef](#)]
56. Boris Kovatchev, Stephen PatekStochastic Models of Physiology 231-256. [[CrossRef](#)]



57. Ramanuja Simha, Hagit Shatkay. 2014. Protein (multi-)location prediction: using location inter-dependencies in a probabilistic framework. *Algorithms for Molecular Biology* **9**, 8. [[CrossRef](#)]
58. Edwin Villanueva, Carlos Dias Maciel. 2014. Efficient methods for learning Bayesian network super-structures. *Neurocomputing* **123**, 3-12. [[CrossRef](#)]
59. Lisette JA Kogelman, Haja N Kadarmideen. 2014. Weighted Interaction SNP Hub (WISH) network method for building genetic networks for complex diseases and traits using whole genome genotype data. *BMC Systems Biology* **8**, S5. [[CrossRef](#)]
60. Ståle Nygård, Trond Reitan, Trevor Clancy, Vegard Nygaard, Johannes Bjørnstad, Biljana Skrbic, Theis Tønnessen, Geir Christensen, Eivind Hovig. 2014. Identifying pathogenic processes by integrating microarray data with prior knowledge. *BMC Bioinformatics* **15**, 115. [[CrossRef](#)]
61. Ngoc Le, Tu Ho, Bich Ho, Dang Tran. 2014. A nucleosomal approach to inferring causal relationships of histone modifications. *BMC Genomics* **15**, S7. [[CrossRef](#)]
62. Federico M. Stefanini. 2014. Chain Graph Models to Elicit the Structure of a Bayesian Network. *The Scientific World Journal* **2014**, 1-12. [[CrossRef](#)]
63. Anwoy Mohanty, Aniruddha Datta, Vijayanagaram Venkatraj. 2014. Using the message passing algorithm on discrete data to detect faults in boolean regulatory networks. *Algorithms for Molecular Biology* **9**, 20. [[CrossRef](#)]
64. Cristian Andrés Gallo, Jessica Andrea Carballido, Ignacio Ponzoni. Inference of Gene Regulatory Networks Based on Association Rules 803-838. [[CrossRef](#)]
65. A. F. Villaverde, J. R. Banga. 2013. Reverse engineering and identification in systems biology: strategies, perspectives and challenges. *Journal of The Royal Society Interface* **11**, 20130505-20130505. [[CrossRef](#)]
66. I-Chi Hung, Su-Sen Chang, Pei-Chun Chang, Cheng-Chun Lee, Calvin Yu-Chian Chen. 2013. Memory enhancement by traditional Chinese medicine?. *Journal of Biomolecular Structure and Dynamics* **31**, 1411-1439. [[CrossRef](#)]
67. Hong Ling, Sandhya Samarasinghe, Don Kulasiri. 2013. Novel recurrent neural network for modelling biological networks: Oscillatory p53 interaction dynamics. *Biosystems* **114**, 191-205. [[CrossRef](#)]
68. Lucia Sacchi, Allan Tucker, Steve Counsell, David Garway-Heath, Stephen Swift. 2013. Improving predictive models of glaucoma severity by incorporating quality indicators. *Artificial Intelligence in Medicine* . [[CrossRef](#)]
69. Jack Kuipers, Giusi Moffa. 2013. Uniform random generation of large acyclic digraphs. *Statistics and Computing* . [[CrossRef](#)]
70. Mohammad J. Sadeh, Giusi Moffa, Rainer Spang. 2013. Considering Unknown Unknowns: Reconstruction of Nonconfoundable Causal Relations in Biological Networks. *Journal of Computational Biology* **20**:11, 920-932. [[Abstract](#)] [[Full Text HTML](#)] [[Full Text PDF](#)] [[Full Text PDF with Links](#)]
71. M. Catherine Duryea, Andrew D. Kern, Robert M. Cox, Ryan Calsbeek. 2013. A novel application of Approximate Bayesian Computation for detecting male reproductive advantages due to mating order. *Behavioral Ecology and Sociobiology* **67**, 1867-1875. [[CrossRef](#)]
72. Ilga Porth, Jaroslav Klápště, Oleksandr Skyba, Michael C. Friedmann, Jan Hannemann, Juergen Ehlting, Yousry A. El-Kassaby, Shawn D. Mansfield, Carl J. Douglas. 2013. Network analysis reveals the relationship among wood properties, gene expression levels and genotypes of natural *Populus trichocarpa* accessions. *New Phytologist* **200**:10.1111/nph.2013.200.issue-3, 727-742. [[CrossRef](#)]
73. Hao Chen, Jitao Sun. 2013. Global stability and stabilization of switched Boolean network with impulsive effects. *Applied Mathematics and Computation* **224**, 625-634. [[CrossRef](#)]
74. Michael Harris, Krithika Bhuvaneshwar, Thanmozhi Natarajan, Laura Sheahan, Difei Wang, Mahlet G. Tadesse, Ira Shoulson, Ross Filice, Kenneth Steadman, Michael J. Pishvaian, Subha



- Madhavan, John Deeken. 2013. Pharmacogenomic characterization of gemcitabine response – a framework for data integration to enable personalized medicine. *Pharmacogenetics and Genomics* 1. [\[CrossRef\]](#)
75. Liv Gingnell, Ulrik Franke, Robert Lagerström, Evelina Ericsson, Joakim Lilliesköld. 2013. Quantifying success factors for IT projects –An expert-based Bayesian model. *Information Systems Management* 131028101955000. [\[CrossRef\]](#)
  76. Yinghua Zhang, Wensheng Zhang, Yuan Xie. 2013. Improved heuristic equivalent search algorithm based on Maximal Information Coefficient for Bayesian Network Structure Learning. *Neurocomputing* 117, 186–195. [\[CrossRef\]](#)
  77. L. Lopez-Kleine, L. Leal, C. Lopez. 2013. Biostatistical approaches for the reconstruction of gene co-expression networks based on transcriptomic data. *Briefings in Functional Genomics* 12, 457–467. [\[CrossRef\]](#)
  78. V. Vembarasan, G. Nagamani, P. Balasubramaniam, Ju H. Park. 2013. State estimation for delayed genetic regulatory networks based on passivity theory. *Mathematical Biosciences* 244, 165–175. [\[CrossRef\]](#)
  79. Min Hee Hahn, Kun Chang Lee, Nam Yong Jo. 2013. Scenario-based management of individual creativity. *Computers in Human Behavior* . [\[CrossRef\]](#)
  80. Jennifer Clarke, Clara Penas, Chiara Pastori, Ricardo J. Komotar, Amade Bregy, Ashish H. Shah, Claes Wahlestedt, Nagi G. Ayad. 2013. Epigenetic pathways and glioblastoma treatment. *Epigenetics* 8, 785–795. [\[CrossRef\]](#)
  81. Alejandro Burga, Ben Lehner. 2013. Predicting phenotypic variation from genotypes, phenotypes and a combination of the two. *Current Opinion in Biotechnology* 24, 803–809. [\[CrossRef\]](#)
  82. Soheil Feizi, Daniel Marbach, Muriel Médard, Manolis Kellis. 2013. Network deconvolution as a general method to distinguish direct dependencies in networks. *Nature Biotechnology* 31, 726–733. [\[CrossRef\]](#)
  83. Lourdes A. Medina, Marija Jankovic, Gül E. Okudan Kremer, Bernard Yannou. 2013. An investigation of critical factors in medical device development through Bayesian networks. *Expert Systems with Applications* . [\[CrossRef\]](#)
  84. Roy H. Perlis. 2013. A Clinical Risk Stratification Tool for Predicting Treatment Resistance in Major Depressive Disorder. *Biological Psychiatry* 74, 7–14. [\[CrossRef\]](#)
  85. Assieh Saadatpour, Réka Albert. 2013. Boolean modeling of biological regulatory networks: A methodology tutorial. *Methods* 62, 3–12. [\[CrossRef\]](#)
  86. I. S. Jang, A. Margolin, A. Califano. 2013. hARACNe: improving the accuracy of regulatory model reverse engineering via higher-order data processing inequality tests. *Interface Focus* 3, 20130011–20130011. [\[CrossRef\]](#)
  87. Shuzhao Li, Helder I. Nakaya, Dmitri A. Kazmin, Jason Z. Oh, Bali Pulendran. 2013. Systems biological approaches to measure and understand vaccine immunity in humans. *Seminars in Immunology* . [\[CrossRef\]](#)
  88. Alejandro Villaverde, John Ross, Julio Banga. 2013. Reverse Engineering Cellular Networks with Information Theoretic Methods. *Cells* 2, 306–329. [\[CrossRef\]](#)
  89. Emmanuel Prestat, Sérgio Rodrigues de Moraes, Julie A. Vendrell, Aurélie Thollet, Christian Gautier, Pascale A Cohen, Alex Aussem. 2013. Learning the local Bayesian network structure around the ZNF217 oncogene in breast tumours. *Computers in Biology and Medicine* 43, 334–341. [\[CrossRef\]](#)
  90. Guanming Wu, Lincoln SteinNetwork-Module-Based Approaches in Cancer Data Analysis 173–192. [\[CrossRef\]](#)
  91. Binhua Tang, Fei Gu, Victor X. JinInference of Gene Regulatory Networks in Breast and Ovarian Cancer by Integrating Different Genomic Data 153–171. [\[CrossRef\]](#)

92. Nidhal Bouaynaya, Roman Shterenberg, Dan Schonfeld, Hassan M. Fathallah-Shaykh Intervention and Control of Gene Regulatory Networks: Theoretical Framework and Application to Human Melanoma Gene Regulation 215-238. [[CrossRef](#)]
93. Stephan Stahlschmidt, Helmut Tausendteufel, Wolfgang K. Härdle. 2013. Bayesian networks for sex-related homicides: structure learning and prediction. *Journal of Applied Statistics* 1-17. [[CrossRef](#)]
94. Changwon Yoo. 2013. Statistical methods in experimentation recommendation models for discovering gene regulation pathways. *Wiley Interdisciplinary Reviews: Computational Statistics* 5:10.1002/wics.v5.2, 121-134. [[CrossRef](#)]
95. Pablo Hernandez-Leal, Alma Rios-Flores, Santiago Ávila-Rios, Gustavo Reyes-Terán, Jesus A. Gonzalez, Lindsey Fiedler-Cameras, Felipe Orihuela-Espina, Eduardo F. Morales, L. Enrique Sucar. 2013. Discovering human immunodeficiency virus mutational pathways using temporal Bayesian networks. *Artificial Intelligence in Medicine* 57, 185-195. [[CrossRef](#)]
96. David C. Martins, Evaldo A. de Oliveira, Ulisses M. Braga-Neto, Ronaldo F. Hashimoto, Roberto M. Cesar. 2013. Signal propagation in Bayesian networks and its relationship with intrinsically multivariate predictive variables. *Information Sciences* 225, 18-34. [[CrossRef](#)]
97. Roberto Cordone, Guglielmo Lulli. 2013. An integer optimization approach for reverse engineering of gene regulatory networks. *Discrete Applied Mathematics* 161, 580-592. [[CrossRef](#)]
98. YUCEF DERBAL. 2013. ON MODELING OF LIVING ORGANISMS USING HIERARCHICAL COARSE-GRAINING ABSTRACTIONS OF KNOWLEDGE. *Journal of Biological Systems* 21, 1350008. [[CrossRef](#)]
99. Luis A. Fuente, Michael A. Lones, Alexander P. Turner, Susan Stepney, Leo S. Caves, Andy M. Tyrrell. 2013. Computational models of signalling networks for non-linear control. *Biosystems* . [[CrossRef](#)]
100. J. P. Faria, R. Overbeek, F. Xia, M. Rocha, I. Rocha, C. S. Henry. 2013. Genome-scale bacterial transcriptional regulatory networks: reconstruction and integrated analysis with metabolic models. *Briefings in Bioinformatics* . [[CrossRef](#)]
101. Qingying Meng, Ville-Petteri Mäkinen, Helen Luk, Xia Yang. 2013. Systems Biology Approaches and Applications in Obesity, Diabetes, and Cardiovascular Diseases. *Current Cardiovascular Risk Reports* 7, 73-83. [[CrossRef](#)]
102. L. Wang, X. Wang, A. P. Arkin, M. S. Samoilov. 2013. Inference of gene regulatory networks from genome-wide knockout fitness data. *Bioinformatics* 29, 338-346. [[CrossRef](#)]
103. T. Judeh, C. Johnson, A. Kumar, D. Zhu. 2013. TEAK: Topology Enrichment Analysis framework for detecting activated biological subpathways. *Nucleic Acids Research* 41, 1425-1437. [[CrossRef](#)]
104. Joke G. van Bommel, Guillaume J. Filion, Arantxa Rosado, Wendy Talhout, Marcel de Haas, Tibor van Welsem, Fred van Leeuwen, Bas van Steensel. 2013. A Network Model of the Molecular Organization of Chromatin in Drosophila. *Molecular Cell* 49, 759-771. [[CrossRef](#)]
105. Fang Yu, Jane Meza Design and Statistical Analysis of Mass Spectrometry-Based Quantitative Proteomics Data 179-204. [[CrossRef](#)]
106. Jörg Bartel, Jan Krumsiek, Fabian J. Theis. 2013. STATISTICAL METHODS FOR THE ANALYSIS OF HIGH-THROUGHPUT METABOLOMICS DATA. *Computational and Structural Biotechnology Journal* 4, 1-9. [[CrossRef](#)]
107. ClarLynda R Williams-DeVane, David M Reif, Elaine Cohen Hubal, Pierre R Bushel, Edward E Hudgens, Jane E Gallagher, Stephen W Edwards. 2013. Decision tree-based method for integrating gene expression, demographic, and clinical data to determine disease endotypes. *BMC Systems Biology* 7, 119. [[CrossRef](#)]
108. David A. Bohan, Alan Raybould, Christian Mulder, Guy Woodward, Alireza Tamaddon-Nezhad, Nico Bluthgen, Michael J.O. Pocock, Stephen Muggleton, Darren M. Evans, Julia Astegiano,

François Massol, Nicolas Loeuille, Sandrine Petit, Sarina Macfadyen Networking Agroecology 1-67. [\[CrossRef\]](#)

109. David Toubiana, Alisdair R. Fernie, Zoran Nikoloski, Aaron Fait. 2013. Network analysis: tackling complex data to study plant metabolism. *Trends in Biotechnology* **31**, 29-36. [\[CrossRef\]](#)
110. David M Simcha, Laurent Younes, Martin J Aryee, Donald Geman. 2013. Identification of direction in gene networks from expression and methylation. *BMC Systems Biology* **7**, 118. [\[CrossRef\]](#)
111. Fabrizio F Borelli, Raphael Y de Camargo, David C Martins, Luiz CS Rozante. 2013. Gene regulatory networks inference using a multi-GPU exhaustive search algorithm. *BMC Bioinformatics* **14**, S5. [\[CrossRef\]](#)
112. D. Muraro, H.M. Byrne, J.R. King, M.J. Bennett, A. Chavarria-Krauser, L. Dupuy, M. Ptashnyk. 2013. Mathematical Modelling Plant Signalling Networks. *Mathematical Modelling of Natural Phenomena* **8**, 5-24. [\[CrossRef\]](#)
113. Guy Karlebach. 2013. Inferring Boolean network states from partial information. *EURASIP Journal on Bioinformatics and Systems Biology* **2013**, 11. [\[CrossRef\]](#)
114. L. Chen, J. Xuan, R. B. Riggins, Y. Wang, R. Clarke. 2013. Identifying protein interaction subnetworks by a bagging Markov random field-based method. *Nucleic Acids Research* **41**, e42-e42. [\[CrossRef\]](#)
115. Céline Brouard, Christel Vrain, Julie Dubois, David Castel, Marie-Anne Debily, Florence d'Alché-Buc. 2013. Learning a Markov Logic network for supervised gene regulatory network inference. *BMC Bioinformatics* **14**, 273. [\[CrossRef\]](#)
116. Sungwon Jung, Michael Verdicchio, Jeff Kiefer, Daniel V Hoff, Michael Berens, Michael Bittner, Seungchan Kim. 2013. Learning contextual gene set interaction networks of cancer with condition specificity. *BMC Genomics* **14**, 110. [\[CrossRef\]](#)
117. G. Ian Gallicano. 2013. Modeling to Optimize Terminal Stem Cell Differentiation. *Scientifica* **2013**, 1-20. [\[CrossRef\]](#)
118. Nikolay Balov. 2013. A categorical network approach for discovering differentially expressed regulations in cancer. *BMC Medical Genomics* **6**, S1. [\[CrossRef\]](#)
119. Tobias C. Fuchs, Germaine L. Truise, Philip G. Hewitt Toxicogenomics in Preclinical Development 827-854. [\[CrossRef\]](#)
120. Ratana Lim, Martha Lappas, Clyde Riley, Niels Borregaard, Holger J Moller, Nuzhat Ahmed, Gregory E Rice. 2013. Investigation of human cationic antimicrobial protein-18 (hCAP-18), lactoferrin and CD163 as potential biomarkers for ovarian cancer. *Journal of Ovarian Research* **6**, 5. [\[CrossRef\]](#)
121. Atul J. Butte, David Chen Translational Bioinformatics for Genomic Medicine 272-286. [\[CrossRef\]](#)
122. Chengwei Su, Angeline Andrew, Margaret R Karagas, Mark E Borsuk. 2013. Using Bayesian networks to discover relations between genes, environment, and disease. *BioData Mining* **6**, 6. [\[CrossRef\]](#)
123. Thuc D Le, Lin Liu, Bing Liu, Anna Tsykin, Gregory J Goodall, Kenji Satou, Jiuyong Li. 2013. Inferring microRNA and transcription factor regulatory networks in heterogeneous data. *BMC Bioinformatics* **14**, 92. [\[CrossRef\]](#)
124. Rafael Silva-Rocha, Javier Tamames, Victor de Lorenzo The Logic of Decision Making in Environmental Bacteria 279-302. [\[CrossRef\]](#)
125. Baikang Pei, Dong-Guk Shin. 2012. Reconstruction of Biological Networks by Incorporating Prior Knowledge into Bayesian Network Models. *Journal of Computational Biology* **19**:12, 1324-1334. [\[Abstract\]](#) [\[Full Text HTML\]](#) [\[Full Text PDF\]](#) [\[Full Text PDF with Links\]](#)
126. Reiko Nagano, Hiromi Akanuma, Xian-Yang Qin, Satoshi Imanishi, Hiroyoshi Toyoshiba, Jun Yoshinaga, Seiichiroh Ohsako, Hideko Sone. 2012. Multi-Parametric Profiling Network Based on

Gene Expression and Phenotype Data: A Novel Approach to Developmental Neurotoxicity Testing. *International Journal of Molecular Sciences* **13**, 187-207. [[CrossRef](#)]

127. V Vembarasan, P Balasubramaniam, Kurunathan Ratnavelu, N Kumaresan. 2012. Robust stability analysis of delayed Takagi–Sugeno fuzzy genetic regulatory networks. *Physica Scripta* **86**, 065003. [[CrossRef](#)]
128. Hao Jiang, Wai-Ki Ching, Kiyoko F. Aoki-Kinoshita, Dianjing Guo. 2012. Modeling genetic regulatory networks: a delay discrete dynamical model approach. *Journal of Systems Science and Complexity* **25**, 1052-1067. [[CrossRef](#)]
129. Classification Techniques in Bioinformatics 247-298. [[CrossRef](#)]
130. Shuai Huang, Jing Li, Kewei Chen, Teresa Wu, Jieping Ye, Xia Wu, Li Yao. 2012. A transfer learning approach for network modeling. *IIE Transactions* **44**, 915-931. [[CrossRef](#)]
131. Andrea Rau, Florence Jaffrézic, Jean-Louis Foulley, R. W. Doerge. 2012. Reverse engineering gene regulatory networks using approximate Bayesian computation. *Statistics and Computing* **22**, 1257-1271. [[CrossRef](#)]
132. S. M. Hill, Y. Lu, J. Molina, L. M. Heiser, P. T. Spellman, T. P. Speed, J. W. Gray, G. B. Mills, S. Mukherjee. 2012. Bayesian Inference of Signaling Network Topology in a Cancer Cell Line. *Bioinformatics* **28**, 2804-2810. [[CrossRef](#)]
133. Sameer S. Bajikar, Kevin A. Janes. 2012. Multiscale Models of Cell Signaling. *Annals of Biomedical Engineering* **40**, 2319-2327. [[CrossRef](#)]
134. P.S. Wasan, M. Uttamchandani, S. Moomchala, V.B. Yap, P.H. Yap. 2012. Application of statistics and machine learning for risk stratification of heritable cardiac arrhythmias. *Expert Systems with Applications* . [[CrossRef](#)]
135. Binhua Tang, Pei-Yin Hsu, Tim H.-M. Huang, Victor X. Jin. 2012. Cancer omics: From regulatory networks to clinical outcomes. *Cancer Letters* . [[CrossRef](#)]
136. M. Aluru, J. Zola, D. Nettleton, S. Aluru. 2012. Reverse engineering and analysis of large genome-scale gene networks. *Nucleic Acids Research* . [[CrossRef](#)]
137. George Michailidis. 2012. Statistical Challenges in Biological Networks. *Journal of Computational and Graphical Statistics* **21**, 840-855. [[CrossRef](#)]
138. Roberto Cordone, Guglielmo Lulli. 2012. A GRASP metaheuristic for microarray data analysis. *Computers & Operations Research* . [[CrossRef](#)]
139. S. Vineetha, C. Chandra Shekara Bhat, Sumam Mary Idicula. 2012. Gene regulatory network from microarray data of colon cancer patients using TSK-type recurrent neural fuzzy network. *Gene* **506**, 408-416. [[CrossRef](#)]
140. Andrej Aderhold, Dirk Husmeier, Jack J. Lennon, Colin M. Beale, V. Anne Smith. 2012. Hierarchical Bayesian models in ecology: Reconstructing species interaction networks from non-homogeneous species abundance data. *Ecological Informatics* **11**, 55-64. [[CrossRef](#)]
141. Xing Liu, Zhong-ru Wu, Yang Yang, Jiang Hu, Bo Xu. 2012. Information fusion diagnosis and early-warning method for monitoring the long-term service safety of high dams. *Journal of Zhejiang University SCIENCE A* **13**, 687-699. [[CrossRef](#)]
142. Imme Ebert-Uphoff, Yi Deng. 2012. Causal Discovery for Climate Research Using Graphical Models. *Journal of Climate* **25**, 5648-5665. [[CrossRef](#)]
143. Li-Zhi Liu, Fang-Xiang Wu, Wen-Jun Zhang. 2012. Reverse engineering of gene regulatory networks from biological data. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery* **2**, 365-385. [[CrossRef](#)]
144. G. Morota, B.D. Valente, G.J.M. Rosa, K.A. Weigel, D. Gianola. 2012. An assessment of linkage disequilibrium in Holstein cattle using a Bayesian network. *Journal of Animal Breeding and Genetics* n/a-n/a. [[CrossRef](#)]

145. Jessica Kasza, Gary Glonek, Patty Solomon. 2012. ESTIMATING BAYESIAN NETWORKS FOR HIGH-DIMENSIONAL DATA WITH COMPLEX MEAN STRUCTURE AND RANDOM EFFECTS. *Australian & New Zealand Journal of Statistics* no-no. [[CrossRef](#)]
146. J. Gu, J. Xuan, R. B. Riggins, L. Chen, Y. Wang, R. Clarke. 2012. Robust identification of transcriptional regulatory networks using a Gibbs sampler on outlier sum statistic. *Bioinformatics* 28, 1990-1997. [[CrossRef](#)]
147. I-Ming Wang, Bin Zhang, Xia Yang, Jun Zhu, Serguei Stepaniants, Chunsheng Zhang, Qingying Meng, Mette Peters, Yudong He, Chester Ni, Deborah Slipetz, Michael A Crackower, Hani Houshyar, Christopher M Tan, Ernest Asante-Appiah, Gary O'Neill, Mingjuan Jane Luo, Rolf Thieringer, Jeffrey Yuan, Chi-Sung Chiu, Pek Yee Lum, John Lamb, Yves Boie, Hilary A Wilkinson, Eric E Schadt, Hongyue Dai, Christopher Roberts. 2012. Systems analysis of eleven rodent disease models reveals an inflammatoric signature and key drivers. *Molecular Systems Biology* 8. . [[CrossRef](#)]
148. Daniel Marbach, James C Costello, Robert Küffner, Nicole M Vega, Robert J Prill, Diogo M Camacho, Kyle R Allison, Andrej Aderhold, Kyle R Allison, Richard Bonneau, Diogo M Camacho, Yukun Chen, James J Collins, Francesca Cordero, James C Costello, Martin Crane, Frank Dondelinger, Mathias Drton, Roberto Esposito, Rina Foygel, Alberto de la Fuente, Jan Gertheiss, Pierre Geurts, Alex Greenfield, Marco Grzegorzczak, Anne-Claire Haury, Benjamin Holmes, Torsten Hothorn, Dirk Husmeier, Vân Anh Huynh-Thu, Alexandre Irrthum, Manolis Kellis, Guy Karlebach, Robert Küffner, Sophie Lèbre, Vincenzo De Leo, Aviv Madar, Subramani Mani, Daniel Marbach, Fantine Mordélet, Harry Ostrer, Zhengyu Ouyang, Ravi Pandya, Tobias Petri, Andrea Pinna, Christopher S Poultney, Robert J Prill, Serena Rezny, Heather J Ruskin, Yvan Saey, Ron Shamir, Alina Sirbu, Mingzhou Song, Nicola Soranzo, Alexander Statnikov, Gustavo Stolovitzky, Nicci Vega, Paola Vera-Licona, Jean-Philippe Vert, Alessia Visconti, Haizhou Wang, Louis Wehenkel, Lukas Windhager, Yang Zhang, Ralf Zimmer, Manolis Kellis, James J Collins, Gustavo Stolovitzky. 2012. Wisdom of crowds for robust gene network inference. *Nature Methods* 9, 796-804. [[CrossRef](#)]
149. Guillaume Gravier, Claire-Hélène Demarty, Siwar Baghdadi, Patrick Gros. 2012. Classification-oriented structure learning in Bayesian networks for multimodal event detection in videos. *Multimedia Tools and Applications* . [[CrossRef](#)]
150. M. Kotera, Y. Yamanishi, Y. Moriya, M. Kanehisa, S. Goto. 2012. GENIES: gene network inference engine based on supervised analysis. *Nucleic Acids Research* 40, W162-W167. [[CrossRef](#)]
151. Xiaoming He, Satoshi Imanishi, Hideko Sone, Reiko Nagano, Xian-Yang Qin, Jun Yoshinaga, Hiromi Akanuma, Junko Yamane, Wataru Fujibuchi, Seiichiroh Ohsako. 2012. Effects of methylmercury exposure on neuronal differentiation of mouse and human embryonic stem cells. *Toxicology Letters* 212, 1-10. [[CrossRef](#)]
152. C. F. Quo, C. Kaddi, J. H. Phan, A. Zollanvari, M. Xu, M. D. Wang, G. Alterovitz. 2012. Reverse engineering biomolecular systems using -omic data: challenges, progress and opportunities. *Briefings in Bioinformatics* 13, 430-445. [[CrossRef](#)]
153. Lipi Acharya, Thair Judeh, Dongxiao Zhu A Survey of Computational Approaches to Reconstruct and Partition Biological Networks 1-43. [[CrossRef](#)]
154. Federico Schlüter. 2012. A survey on independence-based Markov networks learning. *Artificial Intelligence Review* . [[CrossRef](#)]
155. A. O. Frolova. 2012. Overview of methods of reverse engineering of gene regulatory networks: Boolean and Bayesian networks. *Biopolymers and Cell* 28:10.7124/bc.28.3, 163-170. [[CrossRef](#)]
156. Yiqian Zhou, Rehman Qureshi, Ahmet Sacan. 2012. Data simulation and regulatory network reconstruction from time-series microarray data using stepwise multiple linear regression. *Network Modeling Analysis in Health Informatics and Bioinformatics* . [[CrossRef](#)]



157. Nimrod Bar-Yaakov, Zehava Grossman, Nathan Intrator. 2012. Using Iterative Ridge Regression to Explore Associations Between Conditioned Variables. *Journal of Computational Biology* **19**:5, 504-518. [[Abstract](#)] [[Full Text HTML](#)] [[Full Text PDF](#)] [[Full Text PDF with Links](#)]
158. Jun He, Syuan-Ming Guo, Mark Bathe. 2012. Bayesian Approach to the Analysis of Fluorescence Correlation Spectroscopy Data I: Theory. *Analytical Chemistry* 120415120145006. [[CrossRef](#)]
159. N. Kravchenko-Balasha, A. Levitzki, A. Goldstein, V. Rotter, A. Gross, F. Remacle, R. D. Levine. 2012. On a fundamental structure of gene networks in living cells. *Proceedings of the National Academy of Sciences* **109**, 4702-4707. [[CrossRef](#)]
160. Song Wu, Wei ZhaoCandidate Screening through Bioinformatics Tools 235-260. [[CrossRef](#)]
161. K. Kim, K. Jiang, S.-I. Melinda Teng, L. J. Feldman, H. Huang. 2012. Using biologically interrelated experiments to identify pathway genes in Arabidopsis. *Bioinformatics* . [[CrossRef](#)]
162. A. Tucker, D. Duplisea. 2012. Bioinformatics tools in predictive ecology: applications to fisheries. *Philosophical Transactions of the Royal Society B: Biological Sciences* **367**, 279-290. [[CrossRef](#)]
163. F. Marta L. Lascio, Simone Giannerini. 2012. A Copula-Based Algorithm for Discovering Patterns of Dependent Observations. *Journal of Classification* . [[CrossRef](#)]
164. Celine Lefebvre, Gabrielle Rieckhof, Andrea Califano. 2012. Reverse-engineering human regulatory networks. *Wiley Interdisciplinary Reviews: Systems Biology and Medicine* n/a-n/a. [[CrossRef](#)]
165. Fang Wang, Xing-wei Tong, Jing Xu, Jian-ping Chen. 2012. Causal analysis of futures sugar prices in Zhengzhou. *Acta Mathematicae Applicatae Sinica, English Series* **28**, 39-48. [[CrossRef](#)]
166. Damian Roqueiro, Lei Huang, Yang Dai. 2012. Identifying transcription factors and microRNAs as key regulators of pathways using Bayesian inference on known pathway structures. *Proteome Science* **10**, S15. [[CrossRef](#)]
167. Natsu Nakajima, Takeyuki Tamura, Yoshihiro Yamanishi, Katsuhisa Horimoto, Tatsuya Akutsu. 2012. Network Completion Using Dynamic Programming and Least-Squares Fitting. *The Scientific World Journal* **2012**, 1-8. [[CrossRef](#)]
168. Roozbeh Manshaei, Pooya Sobhe Bidari, Mahdi Aliyari Shoorehdeli, Amir Feizi, Tahmineh Lohrasebi, Mohammad Ali Malboobi, Matthew Kyan, Javad Alirezaie. 2012. Hybrid-Controlled Neurofuzzy Networks Analysis Resulting in Genetic Regulatory Networks Reconstruction. *ISRN Bioinformatics* **2012**, 1-16. [[CrossRef](#)]
169. Jianlong Qi, Tom Michoel, Gregory Butler. 2012. An Integrative Approach to Infer Regulation Programs in a Transcription Regulatory Module Network. *Journal of Biomedicine and Biotechnology* **2012**, 1-8. [[CrossRef](#)]
170. Chia-Ling Huang, John Lamb, Leonid Chindelevitch, Jarek Kostrowicki, Justin Guinney, Charles DeLisi, Daniel Ziemek. 2012. Correlation set analysis: detecting active regulators in disease populations using prior causal knowledge. *BMC Bioinformatics* **13**, 46. [[CrossRef](#)]
171. Na You, Peng Mou, Ting Qiu, Qiang Kou, Huaijin Zhu, Yuexi Chen, Xueqin Wang. 2012. Gene Expression Network Reconstruction by LEP Method Using Microarray Data. *The Scientific World Journal* **2012**, 1-6. [[CrossRef](#)]
172. Benjamin A Logsdon, Gabriel E Hoffman, Jason G Mezey. 2012. Mouse obesity network reconstruction with a variational Bayes algorithm to employ aggressive false positive control. *BMC Bioinformatics* **13**, 53. [[CrossRef](#)]
173. Jorg Linde, Peter Hortschansky, Eugen Fazius, Axel A Brakhage, Reinhard Guthke, Hubertus Haas. 2012. Regulatory interactionis for iron homoeostatis in *Aspergillus fumigatus* inferred by a Systems Biology approach. *BMC Systems Biology* **6**, 6. [[CrossRef](#)]
174. Anil Aswani, Mark D. Biggin, Peter Bickel, Claire TomlinNonparametric Variable Selection and Modeling for Spatial and Temporal Regulatory Networks 243-261. [[CrossRef](#)]

175. Anne-Claire Haury, Fantine Mordelet, Paola Vera-Licona, Jean-Philippe Vert. 2012. TIGRESS: Trustful Inference of Gene REgulation using Stability Selection. *BMC Systems Biology* **6**, 145. [[CrossRef](#)]
176. Fabio Fioravanti, Manuela Helmer-Citterich, Enrico Nardelli. 2012. Modeling gene regulatory network motifs using statecharts. *BMC Bioinformatics* **13**, S20. [[CrossRef](#)]
177. Nguyen Xuan, Madhu Chetty, Ross Coppel, Pramod P Wangikar. 2012. Gene regulatory network modeling via global optimization of high-order dynamic Bayesian network. *BMC Bioinformatics* **13**, 131. [[CrossRef](#)]
178. Yukito Watanabe, Shigeto Seno, Yoichi Takenaka, Hideo Matsuda. 2012. An estimation method for inference of gene regulatory network using Bayesian network with uniting of partial problems. *BMC Genomics* **13**, S12. [[CrossRef](#)]
179. Andrei S. Rodin, Grigoriy Gogoshin, Anatoliy Litvinenko, Eric Boerwinkle Exploring Genetic Epidemiology Data with Bayesian Networks 479-510. [[CrossRef](#)]
180. Christopher S. Poultney, Alex Greenfield, Richard Bonneau Integrated Inference and Analysis of Regulatory Networks from Multi-Level Measurements 19-56. [[CrossRef](#)]
181. Neil R Clark, Ruth Dannenfelser, Christopher M Tan, Michael E Komosinski, Avi Ma'ayan. 2012. Sets2Networks: network inference from repeated observations of sets. *BMC Systems Biology* **6**, 89. [[CrossRef](#)]
182. Kenneth Lo, Adrian E Raftery, Kenneth M Dombek, Jun Zhu, Eric E Schadt, Roger E Bumgarner, Ka Yeung. 2012. Integrating external biological knowledge in the construction of regulatory networks from time-series expression data. *BMC Systems Biology* **6**, 101. [[CrossRef](#)]
183. Adriano V Werhli. 2012. Comparing the reconstruction of regulatory pathways with distinct Bayesian networks inference methods. *BMC Genomics* **13**, S2. [[CrossRef](#)]
184. Yifang Zhao, Ming-Hui Chen, Baikang Pei, David Rowe, Dong-Guk Shin, Wangang Xie, Fang Yu, Lynn Kuo. 2011. A Bayesian Approach to Pathway Analysis by Integrating Gene-Gene Functional Directions and Microarray Data. *Statistics in Biosciences* . [[CrossRef](#)]
185. L. Acharya, T. Judeh, G. Wang, D. Zhu. 2011. Optimal Structural Inference of Signaling Pathways from Unordered and Overlapping Gene Sets. *Bioinformatics* . [[CrossRef](#)]
186. C. Parker, J.-H. Zhang High-Throughput Screening for Small-Molecule Drug Discovery . [[CrossRef](#)]
187. F.K. Ahmad, S. Deris, N.H. Othman. 2011. The inference of breast cancer metastasis through gene regulatory networks. *Journal of Biomedical Informatics* . [[CrossRef](#)]
188. Andrew Kasarskis, Xia Yang, Eric Schadt. 2011. Integrative genomics strategies to elucidate the complexity of drug response. *Pharmacogenomics* **12**, 1695-1715. [[CrossRef](#)]
189. K. Y. Yeung, K. M. Dombek, K. Lo, J. E. Mittler, J. Zhu, E. E. Schadt, R. E. Bumgarner, A. E. Raftery. 2011. Construction of regulatory networks using expression time-series data of a genotyped population. *Proceedings of the National Academy of Sciences* **108**, 19436-19441. [[CrossRef](#)]
190. Kumar Selvarajoo, Masa Tsuchiya In a Quest to Uncover Governing Principles of Cellular Networks 261-280. [[CrossRef](#)]
191. Bart S. Hendriks Applications of Systems Biology Approaches to Target Identification and Validation in Drug Discovery 93-133. [[CrossRef](#)]
192. Chien-Pang Lee, Yung-Ho Leu, Wei-Ning Yang. 2011. Constructing gene regulatory networks from microarray data using GA/PSO with DTW. *Applied Soft Computing* . [[CrossRef](#)]
193. Frank Dondelinger, Dirk Husmeier, Sophie Lèbre. 2011. Dynamic Bayesian networks in molecular plant science: inferring gene regulatory networks from multiple gene expression time series. *Euphytica* . [[CrossRef](#)]
194. Christopher K. Tuggle, Fadi Towfic, Vasant G. Honavar Introduction to Systems Biology for Animal Scientists 1-30. [[CrossRef](#)]

195. Austin H. Chen, Chenyin Yang. 2011. The improvement of breast cancer prognosis accuracy from integrated gene expression and clinical data. *Expert Systems with Applications* . [[CrossRef](#)]
196. Wenbing Zhang, Jian-an Fang, Yang Tang. 2011. Robust stability for genetic regulatory networks with linear fractional uncertainties. *Communications in Nonlinear Science and Numerical Simulation* . [[CrossRef](#)]
197. E. R. Morrissey, M. A. Juarez, K. J. Denby, N. J. Burroughs. 2011. Inferring the time-invariant topology of a nonlinear sparse gene regulatory network using fully Bayesian spline autoregression. *Biostatistics* **12**, 682-694. [[CrossRef](#)]
198. Marco Scutari, Korbinian Strimmer Introduction to Graphical Modelling 235-254. [[CrossRef](#)]
199. Dirk Husmeier, Adriano V. Werhli, Marco Grzegorzczuk Advanced Applications of Bayesian Networks in Systems Biology 270-289. [[CrossRef](#)]
200. Gaëlle Lelandais, Sophie Lèbre Recovering Genetic Network from Continuous Data with Dynamic Bayesian Networks 255-269. [[CrossRef](#)]
201. Andrei S Rodin, Grigoriy Gogoshin, Eric Boerwinkle. 2011. Systems biology data analysis methodology in pharmacogenomics. *Pharmacogenomics* **12**, 1349-1360. [[CrossRef](#)]
202. Rafael Silva-Rocha, Javier Tamames, Vitor Martins dos Santos, Víctor de Lorenzo. 2011. The logicome of environmental bacteria: merging catabolic and regulatory events with Boolean formalisms. *Environmental Microbiology* **13**:10.1111/emi.2011.13.issue-9, 2389-2402. [[CrossRef](#)]
203. Qing Zhang, Yong Cao, Yong Li, Yanming Zhu, Samuel S.M. Sun, Dianjing Guo. 2011. A sub-space greedy search method for efficient Bayesian Network inference. *Computers in Biology and Medicine* **41**, 763-770. [[CrossRef](#)]
204. Tao Lu. 2011. High-Dimensional ODEs Coupled With Mixed-Effects Modeling Techniques for Dynamic Gene Regulatory Network Identification. *Journal of the American Statistical Association* 1-17. [[CrossRef](#)]
205. Tuomas Eerola, Lasse Lensu, Joni-Kristian Kamarainen, Tuomas Leisti, Risto Ritala, Göte Nyman, Heikki Kälviäinen. 2011. Bayesian network model of overall print quality: Construction and structural optimisation. *Pattern Recognition Letters* **32**, 1558-1566. [[CrossRef](#)]
206. Brian J. Ross. 2011. The evolution of higher-level biochemical reaction models. *Genetic Programming and Evolvable Machines* . [[CrossRef](#)]
207. S. Roy, M. Werner-Washburne, T. Lane. 2011. A multiple network learning approach to capture system-wide condition-specific responses. *Bioinformatics* **27**, 1832-1838. [[CrossRef](#)]
208. A. P. Parikh, W. Wu, R. E. Curtis, E. P. Xing. 2011. TREEGL: reverse engineering tree-evolving gene networks underlying developing biological lineages. *Bioinformatics* **27**, i196-i204. [[CrossRef](#)]
209. Lei Yang, Jay Lee. 2011. Bayesian Belief Network-based approach for diagnostics and prognostics of semiconductor manufacturing systems. *Robotics and Computer-Integrated Manufacturing* . [[CrossRef](#)]
210. S. Isci, C. Ozturk, J. Jones, H. H. Otu. 2011. Pathway analysis of high-throughput biological data within a Bayesian network framework. *Bioinformatics* **27**, 1667-1674. [[CrossRef](#)]
211. Riccardo Bellazzi, Silvio Biciato, Claudio Cobelli, Barbara Di Camillo, Fulvia Ferrazzi, Paolo Magni, Lucia Sacchi, Gianna Toffolo Microarray Data Analysis: Gene Regulatory Networks 473-488. [[CrossRef](#)]
212. Marco Grzegorzczuk, Dirk Husmeier. 2011. Non-homogeneous dynamic Bayesian networks for continuous data. *Machine Learning* **83**, 355-419. [[CrossRef](#)]
213. Michael M. Zavlanos, A. Agung Julius, Stephen P. Boyd, George J. Pappas. 2011. Inferring stable genetic networks from steady-state data. *Automatica* **47**, 1113-1122. [[CrossRef](#)]
214. MÅRTEN SIMONSSON, PONTUS JOHNSON, MATHIAS EKSTEDT, WALDO ROCHA FLORES. 2011. IT GOVERNANCE DECISION SUPPORT USING THE IT



ORGANIZATION MODELING AND ASSESMENT TOOL. *International Journal of Innovation and Technology Management* **08**, 167-189. [[CrossRef](#)]

215. Rónán Daly, Qiang Shen, Stuart Aitken. 2011. Learning Bayesian networks: approaches and issues. *The Knowledge Engineering Review* **26**, 99-157. [[CrossRef](#)]
216. Ulrik Franke, Pontus Johnson, Johan König, Liv Marcks von Würtemberg. 2011. Availability of enterprise IT systems: an expert-based Bayesian framework. *Software Quality Journal* . [[CrossRef](#)]
217. Haseong Kim, Erol Gelenbe. 2011. G-Networks Based Two Layer Stochastic Modeling of Gene Regulatory Networks with Post-Translational Processes. *Interdisciplinary Bio Central* **3**, 1-6. [[CrossRef](#)]
218. U. Mansmann, V. Jurinovic. 2011. Biological feature validation of estimated gene interaction networks from microarray data: a case study on MYC in lymphomas. *Briefings in Bioinformatics* **12**, 230-244. [[CrossRef](#)]
219. Sergey M. Plis, Michael P. Weisend, Eswar Damaraju, Tom Eichele, Andy Mayer, Vincent P. Clark, Terran Lane, Vince D. Calhoun. 2011. Effective connectivity analysis of fMRI and MEG data collected under identical paradigms. *Computers in Biology and Medicine* . [[CrossRef](#)]
220. Rebecca Jörnsten, Tobias Abenius, Teresia Kling, Linnéa Schmidt, Erik Johansson, Torbjörn E M Nordling, Bodil Nordlander, Chris Sander, Peter Gennemark, Keiko Funa, Björn Nilsson, Linda Lindahl, Sven Nelander. 2011. Network modeling of the transcriptional effects of copy number aberrations in glioblastoma. *Molecular Systems Biology* **7** . [[CrossRef](#)]
221. Daniel Remondini, Gastone Castellani Multiscale Network Reconstruction from Gene Expression Measurements: Correlations, Perturbations, and “A Priori Biological Knowledge” 105-131. [[CrossRef](#)]
222. Matthew T. Weirauch Gene Coexpression Networks for the Analysis of DNA Microarray Data 215-250. [[CrossRef](#)]
223. Shigeru Saito, Katsuhisa Horimoto Network Screening: A New Method to Identify Active Networks from an Ensemble of Known Networks 281-298. [[CrossRef](#)]
224. Sandra Rodriguez-Zas, Younhee Ko Elucidation of General and Condition-Dependent Gene Pathways Using Mixture Models and Bayesian Networks 91-103. [[CrossRef](#)]
225. Ahmet Ay, David N. Arnosti. 2011. Mathematical modeling of gene expression: a guide for the perplexed biologist. *Critical Reviews in Biochemistry and Molecular Biology* **46**, 137-151. [[CrossRef](#)]
226. Austin H. Chen, Ching-Heng Lin. 2011. A novel support vector sampling technique to improve classification accuracy and to identify key genes of leukaemia and prostate cancers. *Expert Systems with Applications* **38**, 3209-3219. [[CrossRef](#)]
227. Henning Redestig, Jędrzej Szymanski, Masami Y. Hirai, Joachim Selbig, Lothar Willmitzer, Zoran Nikoloski, Kazuki Saito Data Integration, Metabolic Networks and Systems Biology 261-316. [[CrossRef](#)]
228. Harya Widiputra, Russel Pears, Nikola Kasabov. 2011. DYNAMIC INTERACTION NETWORKS VERSUS LOCAL TREND MODELS FOR MULTIPLE TIME-SERIES PREDICTION. *Cybernetics and Systems* **42**, 100-123. [[CrossRef](#)]
229. Y. Prat, M. Fromer, N. Linial, M. Linial. 2011. Recovering key biological constituents through sparse representation of gene expression. *Bioinformatics* **27**, 655-661. [[CrossRef](#)]
230. Guixia Liu, Lei Liu, Chunyu Liu, Ming Zheng, Lanying Su, Chunguang Zhou. 2011. Combination of Neuro-Fuzzy Network Models with Biological Knowledge for Reconstructing Gene Regulatory Networks. *Journal of Bionic Engineering* **8**, 98-106. [[CrossRef](#)]
231. P. Larrañaga, S. Moral. 2011. Probabilistic graphical models in artificial intelligence. *Applied Soft Computing* **11**, 1511-1528. [[CrossRef](#)]
232. Dana Pe'er, Nir Hacohen. 2011. Principles and Strategies for Developing Network Models in Cancer. *Cell* **144**, 864-873. [[CrossRef](#)]

233. Zaher Dawy, Elias Yaacoub, Marcel Nassar, Rami Abdallah, Hady Ali Zeineddine. 2011. A multiorganism based method for Bayesian gene network estimation. *Biosystems* **103**, 425-434. [[CrossRef](#)]
234. XinGuang Zhu, GuiLian Zhang, Danny Tholen, Yu Wang, ChangPeng Xin, QingFeng Song. 2011. The next generation models for crops and agro-ecosystems. *Science China Information Sciences* **54**, 589-597. [[CrossRef](#)]
235. Alberto Roverato, F. Marta L. Di Lascio. 2011. Wilks'  $\Lambda$  Dissimilarity Measures for Gene Clustering: An Approach Based on the Identification of Transcription Modules. *Biometrics* no-no. [[CrossRef](#)]
236. MARTIN GEBSER, TORSTEN SCHAUB, SVEN THIELE, PHILIPPE VEBER. 2011. Detecting inconsistencies in large biological networks with answer set programming. *Theory and Practice of Logic Programming* **11**, 323-360. [[CrossRef](#)]
237. S.I. Ao, V. Palade. 2011. Ensemble of Elman neural networks and support vector machines for reverse engineering of gene regulatory networks. *Applied Soft Computing* **11**, 1718-1726. [[CrossRef](#)]
238. Joshua W. K. Ho, Michael A. Charleston. 2011. Network modelling of gene regulation. *Biophysical Reviews* **3**, 1-13. [[CrossRef](#)]
239. Jagath C Rajapakse, Piyushkumar A Mundra. 2011. Stability of building gene regulatory networks with sparse autoregressive models. *BMC Bioinformatics* **12**, S17. [[CrossRef](#)]
240. Cristian A Gallo, Jessica A Carballido, Ignacio Ponzoni. 2011. Discovering time-lagged rules from microarray data using gene profile classifiers. *BMC Bioinformatics* **12**, 123. [[CrossRef](#)]
241. Mathukumalli Vidyasagar. 2011. Probabilistic Methods in Cancer Biology. *European Journal of Control* **17**, 483-511. [[CrossRef](#)]
242. Wei Feng, Simon Yang, Haixia Wu. 2011. On delayed uncertain genetic regulatory networks: robust stability analysis. *International Journal of Computer Mathematics* 1-16. [[CrossRef](#)]
243. Angel Alvarez, Peter J Woolf. 2011. RegNetB: Predicting Relevant Regulator-Gene Relationships in Localized Prostate Tumor Samples. *BMC Bioinformatics* **12**, 243. [[CrossRef](#)]
244. Shouguo Gao, Xujing Wang. 2011. Quantitative utilization of prior biological knowledge in the Bayesian network modeling of gene expression data. *BMC Bioinformatics* **12**, 359. [[CrossRef](#)]
245. Ritwik K. Layek, Aniruddha Datta, Edward R. Dougherty. 2011. From biological pathways to regulatory networks. *Molecular BioSystems* **7**, 843. [[CrossRef](#)]
246. Donny Soh, Difeng Dong, Yike Guo, Limsoon Wong. 2011. Finding consistent disease subnetworks across microarray datasets. *BMC Bioinformatics* **12**, S15. [[CrossRef](#)]
247. José A. Gámez, Juan L. Mateo, José M. Puerta. 2011. Learning Bayesian networks by hill climbing: efficient methods based on progressive restriction of the neighborhood. *Data Mining and Knowledge Discovery* **22**, 106-148. [[CrossRef](#)]
248. Brandon NS Ooi, Toan Phan. 2011. Insights gained from the reverse engineering of gene networks in keloid fibroblasts. *Theoretical Biology and Medical Modelling* **8**, 13. [[CrossRef](#)]
249. Alfredo A Kalaitzis, Neil D Lawrence. 2011. A Simple Approach to Ranking Differentially Expressed Gene Expression Time Courses through Gaussian Process Regression. *BMC Bioinformatics* **12**, 180. [[CrossRef](#)]
250. Carlos HA Higa, Vitor HP Louzada, Tales P Andrade, Ronaldo F Hashimoto. 2011. Constraint-based analysis of gene interactions using restricted boolean networks and time-series data. *BMC Proceedings* **5**, S5. [[CrossRef](#)]
251. Leonardo Vanneschi, Antonella Farinaccio, Giancarlo Mauri, Mauro Antoniotti, Paolo Provero, Mario Giacobini. 2011. A comparison of machine learning techniques for survival prediction in breast cancer. *BioData Mining* **4**, 12. [[CrossRef](#)]

252. Jia Kang, Wei Zheng, Lun Li, Joon Lee, Xiting Yan, Hongyu Zhao. 2011. Use of Bayesian networks to dissect the complexity of genetic disease: application to the Genetic Analysis Workshop 17 simulated data. *BMC Proceedings* 5, S37. [[CrossRef](#)]
253. Konstantina Dimitrakopoulou, Charalampos Tsimpouris, George Papadopoulos, Claudia Pommerenke, Esther Wilk, Kyriakos N Sgarbas, Klaus Schughart, Anastasios Bezerianos. 2011. Dynamic gene network reconstruction from gene expression data in mice after influenza A (H1N1) infection. *Journal of Clinical Bioinformatics* 1, 27. [[CrossRef](#)]
254. Haixiang Shi, Bertil Schmidt, Weiguo Liu, Wolfgang Mueller-Wittig. 2011. Parallel Mutual Information Estimation for Inferring Gene Regulatory Networks on GPUs. *BMC Research Notes* 4, 189. [[CrossRef](#)]
255. Ilana Lichtenstein, Albert Zomaya, Jennifer Gamble, Mathew Vadas Approaches to Construction and Analysis of MicroRNA-Mediated Networks 979-1006. [[CrossRef](#)]
256. Bhaskar Dasgupta, Paola Vera-Ilicona, Eduardo Sontag Reverse Engineering of Molecular Networks from a Common Combinatorial Approach 941-953. [[CrossRef](#)]
257. Mohamed Elati, Céline Rouveirol Unsupervised Learning for Gene Regulation Network Inference from Expression Data: A Review 955-978. [[CrossRef](#)]
258. Yonghui Sun, Gang Feng, Jinde Cao. 2010. A New Approach to Dynamic Fuzzy Modeling of Genetic Regulatory Networks. *IEEE Transactions on NanoBioscience* 9, 263-272. [[CrossRef](#)]
259. Mozaherul Hoque Abul Hasanat, Dhanesh Ramachandram, Rajeswari Mandava. 2010. Bayesian belief network learning algorithms for modeling contextual relationships in natural imagery: a comparative study. *Artificial Intelligence Review* 34, 291-308. [[CrossRef](#)]
260. Franziska Hinkelmann, Reinhard Laubenbacher. 2010. Boolean models of bistable biological systems. *Discrete and Continuous Dynamical Systems - Series S* 4, 1443-1456. [[CrossRef](#)]
261. Hubert Hackl, Gernot Stocker, Pornpimol Charoentong, Bernhard Mlecnik, Gabriela Bindea, Jerome Galon, Zlatko Trajanoski. 2010. Information technology solutions for integration of biomolecular and clinical data in the identification of new cancer biomarkers and targets for therapy. *Pharmacology & Therapeutics* 128, 488-498. [[CrossRef](#)]
262. Joel T. Dudley, Eric Schadt, Marina Sirota, Atul J. Butte, Euan Ashley. 2010. Drug Discovery in a Multidimensional World: Systems, Patterns, and Networks. *Journal of Cardiovascular Translational Research* 3, 438-447. [[CrossRef](#)]
263. Li-Hua JIANG, Yi-Xue LI, Qi LIU. 2010. Reconstruction of Gene Regulatory Networks by Integrating ChIP-chip, Knock out and Expression Data\*. *PROGRESS IN BIOCHEMISTRY AND BIOPHYSICS* 37, 996-1005. [[CrossRef](#)]
264. C. Bender, F. Henjes, H. Frohlich, S. Wiemann, U. Korf, T. Beissbarth. 2010. Dynamic deterministic effects propagation networks: learning signalling pathways from longitudinal protein array data. *Bioinformatics* 26, i596-i602. [[CrossRef](#)]
265. E. R. Morrissey, M. A. Juarez, K. J. Denby, N. J. Burroughs. 2010. On reverse engineering of gene interaction networks using time course data with repeated measurements. *Bioinformatics* 26, 2305-2312. [[CrossRef](#)]
266. Qi Luo, Rubei Zhang, Xiaoxin Liao. 2010. Unconditional global exponential stability in Lagrange sense of genetic regulatory networks with SUM regulatory logic. *Cognitive Neurodynamics* 4, 251-261. [[CrossRef](#)]
267. Lorenzo Farina, Alfredo Germani, Gabriella Mavelli, Pasquale Palumbo. 2010. Identification of Regulatory Network Motifs from Gene Expression Data. *Journal of Mathematical Modelling and Algorithms* 9, 233-245. [[CrossRef](#)]
268. James W. Baurley, David V. Conti, W. James Gauderman, Duncan C. Thomas. 2010. Discovery of complex pathways from observational data. *Statistics in Medicine* 29:10.1002/sim.v29:19, 1998-2011. [[CrossRef](#)]

269. SHUHEI KIMURA, YUICHI SHIRAISHI, MARIKO OKADA. 2010. INFERENCE OF GENETIC NETWORKS USING LPMs: ASSESSMENT OF CONFIDENCE VALUES OF REGULATIONS. *Journal of Bioinformatics and Computational Biology* **08**, 661-677. [[CrossRef](#)]
270. Christoph Echtermeyer, Tom V. Smulders, V. Anne Smith. 2010. Causal pattern recovery from neural spike train data using the Snap Shot Score. *Journal of Computational Neuroscience* **29**, 231-252. [[CrossRef](#)]
271. Xin-Guang Zhu, Lanlan Shan, Yu Wang, William Paul Quick. 2010. C4 Rice - an Ideal Arena for Systems Biology Research. *Journal of Integrative Plant Biology* **52**:10.1111/jipb.2010.52.issue-8, 762-770. [[CrossRef](#)]
272. A. Youn, D. J. Reiss, W. Stuetzle. 2010. Learning transcriptional networks from the integration of ChIP-chip and expression data in a non-parametric model. *Bioinformatics* **26**, 1879-1886. [[CrossRef](#)]
273. Sach Mukherjee, Terence P. Speed, Steven M. Hill Model Averaging for Biological Networks with Prior Information 347-372. [[CrossRef](#)]
274. André Fujita, João Ricardo Sato, Marcos Angelo Almeida Demasi, Satoru Miyano, Mari Cleide Sogayar, Carlos Eduardo Ferreira An Introduction to Time-Varying Connectivity Estimation for Gene Regulatory Networks 205-230. [[CrossRef](#)]
275. Bai Zhang, Huai Li, Robert Clarke, Leena Hilakivi-Clarke, Yue Wang Differential Dependency Network Analysis to Identify Topological Changes in Biological Networks 185-203. [[CrossRef](#)]
276. Franziska Matthäus, V. Anne Smith, Peter J. Gebicke-Haerter Some Useful Mathematical Tools to Transform Microarray Data into Interactive Molecular Networks 277-300. [[CrossRef](#)]
277. References 327-352. [[CrossRef](#)]
278. D. B. van Schalkwijk, K. van Bochove, B. van Ommen, A. P. Freidig, E. P. van Someren, J. van der Greef, A. A. de Graaf. 2010. Developing computational model-based diagnostics to analyse clinical chemistry data. *Briefings in Bioinformatics* **11**, 403-416. [[CrossRef](#)]
279. Youngchul Kim, Jae K. Lee, Haseong Kim, Annamalai Muthiah, Ginger Davis Statistical Network Analysis for Biological Systems and Pathways 249-282. [[CrossRef](#)]
280. Elena S. Dimitrova, M. Paola Vera Licona, John McGee, Reinhard Laubenbacher. 2010. Discretization of Time Series Data. *Journal of Computational Biology* **17**:6, 853-868. [[Abstract](#)] [[Full Text PDF](#)] [[Full Text PDF with Links](#)]
281. Zhiqiang (Eric) Zheng, Paul A. Pavlou. 2010. Research Note —Toward a Causal Interpretation from Observational Data: A New Bayesian Networks Method for Structural Models with Latent Variables. *Information Systems Research* **21**, 365-391. [[CrossRef](#)]
282. Mehmet Koyutürk. 2010. Algorithmic and analytical methods in network biology. *Wiley Interdisciplinary Reviews: Systems Biology and Medicine* **2**, 277-292. [[CrossRef](#)]
283. Tyler J. VanderWeele, Sonia Hernández-Díaz, Miguel A. Hernán. 2010. Case-only gene-environment interaction studies: when does association imply mechanistic interaction?. *Genetic Epidemiology* **34**:10.1002/gepi.v34:4, 327-334. [[CrossRef](#)]
284. Qi-Jun LIU, Zheng-Hua WANG, Wan-Lin LIU, Dong LI, Fu-Chu HE, Yun-Ping ZHU. 2010. A Novel Method to Identify The Condition-specific Regulatory Sub-network That Controls The Yeast Cell Cycle Based on Gene Expression Model\*. *PROGRESS IN BIOCHEMISTRY AND BIOPHYSICS* **37**, 402-415. [[CrossRef](#)]
285. Yandi Dharmadi, Ramon Gonzalez. 2010. Elementary network reconstruction: A framework for the analysis of regulatory networks in biological systems. *Journal of Theoretical Biology* **263**, 499-509. [[CrossRef](#)]
286. Ling-Yun Wu, Xiaobo Zhou, Stephen T. C. Wong Computational Imaging and Modeling for Systems Biology 381-401. [[CrossRef](#)]

287. Jean-Philippe Vert Reconstruction of Biological Networks by Supervised Machine Learning Approaches 163-188. [[CrossRef](#)]
288. A. M. Gholami, K. Fellenberg. 2010. Cross-species common regulatory network inference without requirement for prior gene affiliation. *Bioinformatics* **26**, 1082-1090. [[CrossRef](#)]
289. A. Reverter, N. J. Hudson, S. H. Nagaraj, M. Perez-Enciso, B. P. Dalrymple. 2010. Regulatory impact factors: unraveling the transcriptional regulation of complex traits from expression data. *Bioinformatics* **26**, 896-904. [[CrossRef](#)]
290. Luwen Zhang, Mei Xiao, Yong Wang, Wu Zhang. 2010. Reverse engineering large-scale genetic networks: synthetic versus real data. *Journal of Genetics* **89**, 73-80. [[CrossRef](#)]
291. Jianye Sun. 2010. Neural belief network. *Neural Networks* **23**, 334-340. [[CrossRef](#)]
292. D. Sahoo, J. Seita, D. Bhattacharya, M. A. Inlay, I. L. Weissman, S. K. Plevritis, D. L. Dill. 2010. MiDReG: A method of mining developmentally regulated genes using Boolean implications. *Proceedings of the National Academy of Sciences* **107**, 5732-5737. [[CrossRef](#)]
293. Aaron L. Vollrath, Christopher A. Bradfield An Introduction to Toxicogenomics . [[CrossRef](#)]
294. Eun Yong Kang, Chun Ye, Ilya Shpitser, Eleazar Eskin. 2010. Detecting the Presence and Absence of Causal Relationships between Expression of Yeast Genes with Very Few Samples. *Journal of Computational Biology* **17**:3, 533-546. [[Abstract](#)] [[Full Text PDF](#)] [[Full Text PDF with Links](#)]
295. Salih Tuna, Mahesan Niranjan. 2010. Inference from Low Precision Transcriptome Data Representation. *Journal of Signal Processing Systems* **58**, 267-279. [[CrossRef](#)]
296. Ondřej Čepěk, David Kronus, Petr Kučera. 2010. Analysing DNA microarray data using Boolean techniques. *Annals of Operations Research* . [[CrossRef](#)]
297. Liping Jing, M.K. Ng, Ying Liu. 2010. Construction of Gene Networks With Hybrid Approach From Expression Profile and Gene Ontology. *IEEE Transactions on Information Technology in Biomedicine* **14**, 107-118. [[CrossRef](#)]
298. Xiaohu Shen, Haris Vikalo. 2010. Inferring Parameters of Gene Regulatory Networks via Particle Filtering. *EURASIP Journal on Advances in Signal Processing* **2010**, 1-9. [[CrossRef](#)]
299. Mingyi Wang, Vagner Augusto Benedito, Patrick Xuechun Zhao, Michael Udvardi. 2010. Inferring large-scale gene regulatory networks using a low-order constraint-based algorithm. *Molecular BioSystems* **6**, 988. [[CrossRef](#)]
300. Anthony Almudevar. 2010. A Hypothesis Test for Equality of Bayesian Network Models. *EURASIP Journal on Bioinformatics and Systems Biology* **2010**, 1-10. [[CrossRef](#)]
301. Yusuke Kitamura, Tomomi Kimiwa, Jun Maruyama, Takashi Kaburagi, Takashi Matsumoto, Keiji Wada. 2010. Monte Carlo-based Mouse Nuclear Receptor Superfamily Gene Regulatory Network Prediction: Stochastic Dynamical System on Graph with Zipf Prior. *IPSJ Transactions on Bioinformatics* **3**, 24-39. [[CrossRef](#)]
302. Dimitrios H Roukos. 2010. Novel clinico-genome network modeling for revolutionizing genotype-phenotype-based personalized cancer care. *Expert Review of Molecular Diagnostics* **10**, 33-48. [[CrossRef](#)]
303. Uri David Akavia, Oren Litvin, Jessica Kim, Felix Sanchez-Garcia, Dylan Kotliar, Helen C. Causton, Panisa Pochanard, Eyal Mozes, Levi A. Garraway, Dana Pe'er. 2010. An Integrated Approach to Uncover Drivers of Cancer. *Cell* **143**:6, 1005. [[CrossRef](#)]
304. Diogo F. T. Veiga, Bhaskar Dutta, Gábor Balázs. 2010. Network inference and network response identification: moving genome-scale data to the next level of biological discovery. *Molecular BioSystems* **6**, 469. [[CrossRef](#)]
305. Yongqun He, Rino Rappuoli, Anne S. De Groot, Robert T. Chen. 2010. Emerging Vaccine Informatics. *Journal of Biomedicine and Biotechnology* **2010**, 1-26. [[CrossRef](#)]
306. M. F. Ochs. 2010. Knowledge-based data analysis comes of age. *Briefings in Bioinformatics* **11**, 30-39. [[CrossRef](#)]



307. Atul J. Butte Bioinformatic and Computational Analysis for Genomic Medicine 111-130. [[CrossRef](#)]
308. Nahoko Shikata, Yukihiro Maki, Masahiko Nakatsui, Masato Mori, Yasushi Noguchi, Shintaro Yoshida, Michio Takahashi, Nobuo Kondo, Masahiro Okamoto. 2010. Determining important regulatory relations of amino acids from dynamic network analysis of plasma amino acids. *Amino Acids* **38**, 179-187. [[CrossRef](#)]
309. T. M. Przytycka, M. Singh, D. K. Slonim. 2010. Toward the dynamic interactome: it's about time. *Briefings in Bioinformatics* **11**, 15-29. [[CrossRef](#)]
310. Rui Yamaguchi, Seiya Imoto, Satoru Miyano. 2010. Network-Based Predictions and Simulations by Biological State Space Models: Search for Drug Mode of Action. *Journal of Computer Science and Technology* **25**, 131-153. [[CrossRef](#)]
311. Anand K. Gavai, Yury Tikunov, Remco Ursem, Arnaud Bovy, Fred Eeuwijk, Harm Nijveen, Peter J. F. Lucas, Jack A. M. Leunissen. 2009. Constraint-based probabilistic learning of metabolic pathways from tomato volatiles. *Metabolomics* **5**, 419-428. [[CrossRef](#)]
312. GRAHAM J. HICKMAN, T. CHARLIE HODGMAN. 2009. INFERENCE OF GENE REGULATORY NETWORKS USING BOOLEAN-NETWORK INFERENCE METHODS. *Journal of Bioinformatics and Computational Biology* **07**, 1013-1029. [[CrossRef](#)]
313. S. Anjum, A. Doucet, C. C. Holmes. 2009. A boosting approach to structure learning of graphs with and without prior knowledge. *Bioinformatics* **25**, 2929-2936. [[CrossRef](#)]
314. Feng He, Rudi Balling, An-Ping Zeng. 2009. Reverse engineering and verification of gene networks: Principles, assumptions, and limitations of present methods and future perspectives. *Journal of Biotechnology* **144**, 190-203. [[CrossRef](#)]
315. A. Andrei, C. Kendzierski. 2009. An efficient method for identifying statistical interactors in gene association networks. *Biostatistics* **10**, 706-718. [[CrossRef](#)]
316. Y. Wang, X.-S. Zhang, Y. Xia. 2009. Predicting eukaryotic transcriptional cooperativity by Bayesian network integration of genome-wide data. *Nucleic Acids Research* **37**, 5943-5958. [[CrossRef](#)]
317. Xiao-bo Li. 2009. Mathematical modeling of carcinogenesis based on chromosome aberration data. *Chinese Journal of Cancer Research* **21**, 240-246. [[CrossRef](#)]
318. Emma J. Cooke, Richard S. Savage, David L. Wild. 2009. Computational approaches to the integration of gene expression, ChIP-chip and sequence data in the inference of gene regulatory networks. *Seminars in Cell & Developmental Biology* **20**, 863-868. [[CrossRef](#)]
319. Zhi-hua DU, Yi-wei WANG. 2009. Novel structure learning method for constructing gene regulatory network. *Journal of Computer Applications* **29**, 1539-1543. [[CrossRef](#)]
320. HASEONG KIM, JAE K. LEE, TAESUNG PARK. 2009. INFERENCE OF LARGE-SCALE GENE REGULATORY NETWORKS USING REGRESSION-BASED NETWORK APPROACH. *Journal of Bioinformatics and Computational Biology* **07**, 717-735. [[CrossRef](#)]
321. Hong-Qiang Wang, Hau-San Wong, Hailong Zhu, Timothy T.C. Yip. 2009. A neural network-based biomarker association information extraction approach for cancer classification. *Journal of Biomedical Informatics* **42**, 654-666. [[CrossRef](#)]
322. E. Steele, A. Tucker, P.A.C. t Hoen, M.J. Schuemie. 2009. Literature-based priors for gene regulatory networks. *Bioinformatics* **25**, 1768-1774. [[CrossRef](#)]
323. References 295-327. [[CrossRef](#)]
324. K Basso. 2009. Toward a systems biology approach to investigate cellular networks in normal and malignant B cells. *Leukemia* **23**, 1219-1225. [[CrossRef](#)]
325. Guy N. Brock, William D. Beavis, Laura Salter Kubatko. 2009. Fuzzy logic and related methods as a screening tool for detecting gene regulatory networks. *Information Fusion* **10**, 250-259. [[CrossRef](#)]
326. L. Song, M. Kolar, E. P. Xing. 2009. KELLER: estimating time-varying interactions between genes. *Bioinformatics* **25**, i128-i136. [[CrossRef](#)]

327. Ellen S Koster, Andrei S Rodin, Jan AM Raaijmakers, Anke-Hilse Maitland-vander Zee. 2009. Systems biology in pharmacogenomic research: the way to personalized prescribing?. *Pharmacogenomics* **10**, 971-981. [[CrossRef](#)]
328. G. De Micheli. 2009. An Outlook on Design Technologies for Future Integrated Systems. *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems* **28**, 777-790. [[CrossRef](#)]
329. Ahmet Yardimci. 2009. Soft computing in medicine. *Applied Soft Computing* **9**, 1029-1043. [[CrossRef](#)]
330. FERHAT AY, TAMER KAHVECI, VALÉRIE DE CRÉCY-LAGARD. 2009. A FAST AND ACCURATE ALGORITHM FOR COMPARATIVE ANALYSIS OF METABOLIC PATHWAYS. *Journal of Bioinformatics and Computational Biology* **07**, 389-428. [[CrossRef](#)]
331. Zuyi Huang, Juergen Hahn. 2009. Fuzzy modeling of signal transduction networks. *Chemical Engineering Science* **64**, 2044-2056. [[CrossRef](#)]
332. R. Armananzas, B. Calvo, I. Inza, M. Lopez-Hoyos, V. Martinez-Taboada, E. Ucar, I. Bernales, A. Fullaondo, P. Larranaga, A.M. Zubiaga. 2009. Microarray Analysis of Autoimmune Diseases by Machine Learning Procedures. *IEEE Transactions on Information Technology in Biomedicine* **13**, 341-350. [[CrossRef](#)]
333. B. Anchang, M. J. Sadeh, J. Jacob, A. Tresch, M. O. Vlad, P. J. Oefner, R. Spang. 2009. Complex Systems: From Chemistry to Systems Biology Special Feature: Modeling the temporal interplay of molecular signaling and gene expression by using dynamic nested effects models. *Proceedings of the National Academy of Sciences* **106**, 6447-6452. [[CrossRef](#)]
334. O. Litvin, H. C. Causton, B.-J. Chen, D. Pe'er. 2009. Modularity and interactions in the genetics of gene expression. *Proceedings of the National Academy of Sciences* **106**, 6441-6446. [[CrossRef](#)]
335. Yongjin Park, S. Shackney, R. Schwartz. 2009. Network-Based Inference of Cancer Progression from Microarray Data. *IEEE/ACM Transactions on Computational Biology and Bioinformatics* **6**, 200-212. [[CrossRef](#)]
336. Sangeeta B. English, Shou-Ching Shih, Marco F. Ramoni, Lois E. Smith, Atul J. Butte. 2009. Use of Bayesian networks to probabilistically model and improve the likelihood of validation of microarray findings by RT-PCR. *Journal of Biomedical Informatics* **42**, 287-295. [[CrossRef](#)]
337. B GENG, X ZHOU, Y HUNG. 2009. Growing enzyme gene networks by integration of gene expression, motif sequence, and metabolic information. *Pattern Recognition* **42**, 557-561. [[CrossRef](#)]
338. D. R. Bickel, Z. Montazeri, P.-C. Hsieh, M. Beatty, S. J. Lawit, N. J. Bate. 2009. Gene network reconstruction from transcriptional dynamics under kinetic model uncertainty: a case for the second derivative. *Bioinformatics* **25**, 772-779. [[CrossRef](#)]
339. J. Wang, C. Li, E. Wang, X. Wang. 2009. Uncovering the rules for protein-protein interactions from yeast genomic data. *Proceedings of the National Academy of Sciences* **106**, 3752-3757. [[CrossRef](#)]
340. Alexandre Haye, Yves Dehouck, Jean Marc Kwasigroch, Philippe Bogaerts, Marianne Rooman. 2009. Modeling the temporal evolution of the Drosophila gene expression from DNA microarray time series. *Physical Biology* **6**, 016004. [[CrossRef](#)]
341. Barbara Di Camillo, Gianna Toffolo, Claudio Cobelli. 2009. A Gene Network Simulator to Assess Reverse Engineering Algorithms. *Annals of the New York Academy of Sciences* **1158**, 125-142. [[CrossRef](#)]
342. D LI, Y LIN, Y HUANG. 2009. Constructing marketing decision support systems using data diffusion technology: A case study of gas station diversification. *Expert Systems with Applications* **36**, 2525-2533. [[CrossRef](#)]
343. Fabio Parisi, Heinz Koeppl, Felix Naef. 2009. Network Inference by Combining Biologically Motivated Regulatory Constraints with Penalized Regression. *Annals of the New York Academy of Sciences* **1158**, 114-124. [[CrossRef](#)]

344. Robert Castelo, Alberto Roverato. 2009. Reverse Engineering Molecular Regulatory Networks from Microarray Data with qp-Graphs. *Journal of Computational Biology* **16**:2, 213-227. [[Abstract](#)] [[Full Text PDF](#)] [[Full Text PDF with Links](#)]
345. KUMAR SELVARAJOO, MASARU TOMITA, MASA TSUCHIYA. 2009. CAN COMPLEX CELLULAR PROCESSES BE GOVERNED BY SIMPLE LINEAR RULES?. *Journal of Bioinformatics and Computational Biology* **07**, 243-268. [[CrossRef](#)]
346. AO YUAN, GUANJIE CHEN, CHARLES ROTIMI. 2009. GENETIC NETWORK ANALYSIS BY QUASI-BAYESIAN METHOD. *Journal of Bioinformatics and Computational Biology* **07**, 175-192. [[CrossRef](#)]
347. Sunil K. Mathur. 2009. A run-based procedure to identify time-lagged gene clusters in microarray experiments. *Statistics in Medicine* **28**:10.1002/sim.v28:2, 326-337. [[CrossRef](#)]
348. B. Zhang, H. Li, R. B. Riggins, M. Zhan, J. Xuan, Z. Zhang, E. P. Hoffman, R. Clarke, Y. Wang. 2009. Differential dependency network analysis to identify condition-specific topological changes in biological networks. *Bioinformatics* **25**, 526-532. [[CrossRef](#)]
349. Reinhard Laubenbacher, Abdul Salam Jarrah Algebraic Models of Biochemical Networks 163-196. [[CrossRef](#)]
350. Duncan C Thomas, David V Conti, James Baurley, Frederik Nijhout, Michael Reed, Cornelia M Ulrich. 2009. Use of pathway information in molecular epidemiology. *Human Genomics* **4**, 21. [[CrossRef](#)]
351. Leonardo Vanneschi, Francesco Archetti, Mauro Castelli, Ilaria Giordani. 2009. Classification of Oncologic Data with Genetic Programming. *Journal of Artificial Evolution and Applications* **2009**, 1-13. [[CrossRef](#)]
352. Francisco Ruiz Juan, María Elena García Montes, Manuel Gómez López. 2009. The Bayesians nets in the study of free time: determination of social-demographic profile of the Cuban population. *Journal of Human Sport and Exercise* **4**:10.4100/jhse.2009.42, 173-188. [[CrossRef](#)]
353. Hendrik Hache, Hans Lehrach, Ralf Herwig. 2009. Reverse Engineering of Gene Regulatory Networks: A Comparative Study. *EURASIP Journal on Bioinformatics and Systems Biology* **2009**, 1-12. [[CrossRef](#)]
354. Bing Han, Xue-Wen Chen, Xinkun Wang, Elias K. Michaelis. 2009. Integrating Multiple Microarray Data for Cancer Pathway Analysis Using Bootstrapping K-S Test. *Journal of Biomedicine and Biotechnology* **2009**, 1-11. [[CrossRef](#)]
355. Tae-Su Chung, Ji-Hun Kim, Kee-Won Kim, Ju-Han Kim. 2008. Biological Pathway Extension Using Microarray Gene Expression Data. *Genomics & Informatics* **6**, 202-209. [[CrossRef](#)]
356. Matthew V. Rockman. 2008. Reverse engineering the genotype-phenotype map with natural genetic variation. *Nature* **456**, 738-744. [[CrossRef](#)]
357. Eleni I. Vlahogianni, Nikolas Geroliminis, Alexander Skabardonis. 2008. Empirical and Analytical Investigation of Traffic Flow Regimes and Transitions in Signalized Arterials. *Journal of Transportation Engineering* **134**, 512-522. [[CrossRef](#)]
358. Dong Hyun Jeong, Alireza Darvish, Kayvan Najarian, Jing Yang, William Ribarsky. 2008. Interactive visual analysis of time-series microarray data. *The Visual Computer* **24**, 1053-1066. [[CrossRef](#)]
359. E STEELE, A TUCKER. 2008. Consensus and Meta-analysis regulatory networks for combining multiple microarray gene expression datasets. *Journal of Biomedical Informatics* **41**, 914-926. [[CrossRef](#)]
360. A Braunstein, A Pagnani, M Weigt, R Zecchina. 2008. Inference algorithms for gene networks: a statistical mechanics analysis. *Journal of Statistical Mechanics: Theory and Experiment* **2008**, P12001. [[CrossRef](#)]



361. D NAPOLETANI, T SAUER, D STRUPPA, E PETRICOIN, L LIOTTA. 2008. Augmented sparse reconstruction of protein signaling networks. *Journal of Theoretical Biology* **255**, 40-52. [[CrossRef](#)]
362. M. Andrecut, S. Huang, S.A. Kauffman. 2008. Heuristic Approach to Sparse Approximation of Gene Regulatory Networks. *Journal of Computational Biology* **15**:9, 1173-1186. [[Abstract](#)] [[Full Text PDF](#)] [[Full Text PDF with Links](#)]
363. Richard Bonneau. 2008. Learning biological networks: from modules to dynamics. *Nature Chemical Biology* **4**, 658-664. [[CrossRef](#)]
364. Xiaoning Qian, E.R. Dougherty. 2008. Effect of Function Perturbation on the Steady-State Distribution of Genetic Regulatory Networks: Optimal Structural Intervention. *IEEE Transactions on Signal Processing* **56**, 4966-4976. [[CrossRef](#)]
365. Guy Karlebach, Ron Shamir. 2008. Modelling and analysis of gene regulatory networks. *Nature Reviews Molecular Cell Biology* **9**, 770-780. [[CrossRef](#)]
366. J. W. Shin, R. Huggenberger, M. Detmar. 2008. Transcriptional profiling of VEGF-A and VEGF-C target genes in lymphatic endothelium reveals endothelial-specific molecule-1 as a novel mediator of lymphangiogenesis. *Blood* **112**, 2318-2326. [[CrossRef](#)]
367. M. Grzegorzcyk, D. Husmeier, K. D. Edwards, P. Ghazal, A. J. Millar. 2008. Modelling non-stationary gene regulatory processes with a non-homogeneous Bayesian network and the allocation sampler. *Bioinformatics* **24**, 2071-2078. [[CrossRef](#)]
368. Jing Li, Kuei-Ying Huang, Jionghua Jin, Jianjun Shi. 2008. A survey on statistical methods for health care fraud detection. *Health Care Management Science* **11**, 275-287. [[CrossRef](#)]
369. Zhaoxia Yu, Liewei Wang, Michelle A. T. Hildebrandt, Daniel J. Schaid. 2008. Testing whether genetic variation explains correlation of quantitative measures of gene expression, and application to genetic network analysis. *Statistics in Medicine* **27**:10.1002/sim.v27:19, 3847-3867. [[CrossRef](#)]
370. F. Mordellet, J.-P. Vert. 2008. SIRENE: supervised inference of regulatory networks. *Bioinformatics* **24**, i76-i82. [[CrossRef](#)]
371. Xiaogang Ruan, Jinlian Wang, Hui Li, Rhoda E. Perozzi, Edmund F. Perozzi. 2008. The use of logic relationships to model colon cancer gene expression networks with mRNA microarray data#. *Journal of Biomedical Informatics* **41**, 530-543. [[CrossRef](#)]
372. Ruiqi Wang, Chunguang Li, Luonan Chen, K. Aihara. 2008. Modeling and Analyzing Biological Oscillations in Molecular Networks. *Proceedings of the IEEE* **96**, 1361-1385. [[CrossRef](#)]
373. Marco Grzegorzcyk. 2008. Comparison of Two Different Stochastic Models for Extracting Protein Regulatory Pathways Using Bayesian Networks. *Journal of Toxicology and Environmental Health, Part A* **71**, 827-834. [[CrossRef](#)]
374. ADRIANO V. WERHLI, DIRK HUSMEIER. 2008. GENE REGULATORY NETWORK RECONSTRUCTION BY BAYESIAN INTEGRATION OF PRIOR KNOWLEDGE AND/OR DIFFERENT EXPERIMENTAL CONDITIONS. *Journal of Bioinformatics and Computational Biology* **06**, 543-572. [[CrossRef](#)]
375. Z CHAN, I HAVUKKALA, V JAIN, Y HU, N KASABOV. 2008. Soft computing methods to predict gene regulatory networks: An integrative approach on time-series gene expression data. *Applied Soft Computing* **8**, 1189-1199. [[CrossRef](#)]
376. Marco Grzegorzcyk, Dirk Husmeier. 2008. Improving the structure MCMC sampler for Bayesian networks by introducing a new edge reversal move. *Machine Learning* **71**, 265-305. [[CrossRef](#)]
377. Xue-Wen Chen, G. Anantha, Xiaotong Lin. 2008. Improving Bayesian Network Structure Learning with Mutual Information-Based Node Ordering in the K2 Algorithm. *IEEE Transactions on Knowledge and Data Engineering* **20**, 628-640. [[CrossRef](#)]
378. D. Marinazzo, M. Pellicoro, S. Stramaglia. 2008. Kernel-Granger causality and the analysis of dynamical networks. *Physical Review E* **77**. . [[CrossRef](#)]

379. Wentao Zhao, E. Serpedin, E.R. Dougherty. 2008. Inferring Connectivity of Genetic Regulatory Networks Using Information-Theoretic Criteria. *IEEE/ACM Transactions on Computational Biology and Bioinformatics* **5**, 262-274. [[CrossRef](#)]
380. Jing Li, Jianjun Shi, Devin Satz. 2008. Modeling and analysis of disease and risk factors through learning Bayesian networks from observational data. *Quality and Reliability Engineering International* **24**:10.1002/qre.v24:3, 291-302. [[CrossRef](#)]
381. Marc Zapatka, Yvonne Koch, Benedikt Brors. 2008. Ontological Analysis and Pathway Modelling in Drug Discovery. *Pharmaceutical Medicine* **22**, 99-105. [[CrossRef](#)]
382. Jae K. Lee, Paul D. Williams, Sooyoung Cheon. 2008. Data Mining in Genomics. *Clinics in Laboratory Medicine* **28**, 145-166. [[CrossRef](#)]
383. H. Xiong, Y. Choe. 2008. Structural systems identification of genetic regulatory networks. *Bioinformatics* **24**, 553-560. [[CrossRef](#)]
384. S L Rodriguez-Zas, Y Ko, H A Adams, B R Southey. 2008. Advancing the understanding of the embryo transcriptome co-regulation using meta-, functional, and gene network analysis tools. *Reproduction* **135**, 213-224. [[CrossRef](#)]
385. G TSAFNAT. 2008. The Field Representation Language. *Journal of Biomedical Informatics* **41**, 46-57. [[CrossRef](#)]
386. M. ANDRECUT, S. A. KAUFFMAN, A. M. MADNI. 2008. EVIDENCE OF SCALE-FREE TOPOLOGY IN GENE REGULATORY NETWORK OF HUMAN TISSUES. *International Journal of Modern Physics C* **19**, 283-290. [[CrossRef](#)]
387. Ludwig Feinendegen, Philip Hahnfeldt, Eric E. Schadt, Michael Stumpf, Eberhard O. Voit. 2008. Systems biology and its potential role in radiobiology. *Radiation and Environmental Biophysics* **47**, 5-23. [[CrossRef](#)]
388. Isabel M. Tienda-Luna, Yufang Yin, Yufei Huang, Diego P. Ruiz Padillo, Maria C. Carrion Perez, Yufeng Wang. 2008. Constructing Gene Networks Using Variational Bayesian Variable Selection. *Artificial Life* **14**, 65-79. [[CrossRef](#)]
389. Tzu-Tsung Wong, Ching-Han Hsu. 2008. Two-stage classification methods for microarray data. *Expert Systems with Applications* **34**, 375-383. [[CrossRef](#)]
390. Kuo-Ching Liang, Xiaodong Wang. 2008. Gene Regulatory Network Reconstruction Using Conditional Mutual Information. *EURASIP Journal on Bioinformatics and Systems Biology* **2008**, 1-14. [[CrossRef](#)]
391. David A. Winkler Network models in drug discovery and regenerative medicine 143-170. [[CrossRef](#)]
392. A Braunstein, A Pagnani, M Weigt, R Zecchina. 2008. Gene-network inference by message passing. *Journal of Physics: Conference Series* **95**, 012016. [[CrossRef](#)]
393. John Dougherty, Ioan Tabus, Jaakko Astola. 2008. Inference of Gene Regulatory Networks Based on a Universal Minimum Description Length. *EURASIP Journal on Bioinformatics and Systems Biology* **2008**, 1-11. [[CrossRef](#)]
394. Tatsuya Akutsu Bioinformatics . [[CrossRef](#)]
395. Lucia Sacchi, Cristiana Larizza, Paolo Magni, Riccardo Bellazzi. 2007. Precedence Temporal Networks to represent temporal relationships in gene expression data. *Journal of Biomedical Informatics* **40**, 761-774. [[CrossRef](#)]
396. F FERRAZZI, P MAGNI, L SACCHI, A NUZZO, U PETROVIC, R BELLAZZI. 2007. Inferring gene regulatory networks by integrating static and dynamic data. *International Journal of Medical Informatics* **76**, S462-S475. [[CrossRef](#)]
397. Samantha Kleinberg, Kevin Casey, Bud Mishra. 2007. Systems biology via redescription and ontologies (I): finding phase changes with applications to malaria temporal data. *Systems and Synthetic Biology* **1**, 197-205. [[CrossRef](#)]

398. Isabel M. Tienda-Luna, Yufang Yin, Maria C. Carrion, Yufei Huang, Hong Cai, Maribel Sanchez, Yufeng Wang. 2007. Inferring the skeleton cell cycle regulatory network of malaria parasite using comparative genomic and variational Bayesian approaches. *Genetica* **132**, 131-142. [[CrossRef](#)]
399. X. Xiao, Z. Wang, M. Jang, C. B. Burge. 2007. Coevolutionary networks of splicing cis-regulatory elements. *Proceedings of the National Academy of Sciences* **104**, 18583-18588. [[CrossRef](#)]
400. Jasmin Fisher, Thomas A Henzinger. 2007. Executable cell biology. *Nature Biotechnology* **25**, 1239-1249. [[CrossRef](#)]
401. R XU, G VENAYAGAMOORTHY, D WUNSCHII. 2007. Modeling of gene regulatory networks with hybrid differential evolution and particle swarm optimization. *Neural Networks* **20**, 917-927. [[CrossRef](#)]
402. Rui Xu, Donald Wunsch II, Ronald Frank. 2007. Inference of Genetic Regulatory Networks with Recurrent Neural Network Models Using Particle Swarm Optimization. *IEEE/ACM Transactions on Computational Biology and Bioinformatics* **4**, 681-692. [[CrossRef](#)]
403. Shinn-Ying Ho, Chih-Hung Hsieh, Fu-Chieh Yu, Hui-Ling Huang. 2007. An Intelligent Two-Stage Evolutionary Algorithm for Dynamic Pathway Identification From Gene Expression Profiles. *IEEE/ACM Transactions on Computational Biology and Bioinformatics* **4**, 648-704. [[CrossRef](#)]
404. I. Ponzoni, F.J. Azuaje, J.C. Augusto, D.H. Glass. 2007. Inferring Adaptive Regulation Thresholds and Association Rules from Gene Expression Data through Combinatorial Optimization Learning. *IEEE/ACM Transactions on Computational Biology and Bioinformatics* **4**, 624-634. [[CrossRef](#)]
405. J GEBERT, N RADDE, G WEBER. 2007. Modeling gene regulatory networks with piecewise linear differential equations. *European Journal of Operational Research* **181**, 1148-1165. [[CrossRef](#)]
406. Marco Grzegorzczuk. 2007. Extracting Protein Regulatory Networks with Graphical Models. *PROTEOMICS* **7**:10.1002/pmic.v7:18+, 51-59. [[CrossRef](#)]
407. M. Affara, B. Dunmore, C. Savoie, S. Imoto, Y. Tamada, H. Araki, D. S. Charnock-Jones, S. Miyano, C. Print. 2007. Understanding endothelial cell apoptosis: what can the transcriptome, glycome and proteome reveal?. *Philosophical Transactions of the Royal Society B: Biological Sciences* **362**, 1469-1487. [[CrossRef](#)]
408. Solveig K. Sieberts, Eric E. Schadt. 2007. Moving toward a system genetics view of disease. *Mammalian Genome* **18**, 389-401. [[CrossRef](#)]
409. Z HUANG, J LI, H SU, G WATTS, H CHEN. 2007. Large-scale regulatory network analysis from microarray data: modified Bayesian network learning and association rule mining. *Decision Support Systems* **43**, 1207-1225. [[CrossRef](#)]
410. TAESUNG PARK, KIWOONG KIM, SUNG-GON YI, JIN HYUK KIM, YONG-SUNG LEE, SEUNGYEOUN LEE. 2007. SPOT INTENSITY RATIO STATISTICS IN TWO-CHANNEL MICROARRAY EXPERIMENTS. *Journal of Bioinformatics and Computational Biology* **05**, 865-873. [[CrossRef](#)]
411. M. P. Brynildsen, T.-Y. Wu, S.-S. Jang, J. C. Liao. 2007. Biological network mapping and source signal deduction. *Bioinformatics* **23**, 1783-1791. [[CrossRef](#)]
412. L. Y. T. Inoue, M. Neira, C. Nelson, M. Gleave, R. Etzioni. 2007. Cluster-based network model for time-course gene expression data. *Biostatistics* **8**, 507-525. [[CrossRef](#)]
413. N. Soranzo, G. Bianconi, C. Altafini. 2007. Comparing association network algorithms for reverse engineering of large-scale gene regulatory networks: synthetic versus real data. *Bioinformatics* **23**, 1640-1647. [[CrossRef](#)]
414. Y. Pan, T. Durfee, J. Bockhorst, M. Craven. 2007. Connecting quantitative regulatory-network models to the genome. *Bioinformatics* **23**, i367-i376. [[CrossRef](#)]
415. A. Fujita, J.R. Sato, H.M. Garay-Malpartida, P.A. Morettin, M.C. Sogayar, C.E. Ferreira. 2007. Time-varying modeling of gene expression regulatory networks using the wavelet dynamic vector autoregressive method. *Bioinformatics* **23**, 1623-1630. [[CrossRef](#)]

416. J. Pandey, M. Koyuturk, Y. Kim, W. Szpankowski, S. Subramaniam, A. Grama. 2007. Functional annotation of regulatory pathways. *Bioinformatics* **23**, i377-i386. [[CrossRef](#)]
417. X. Sun, P. Hong. 2007. Computational modeling of *Caenorhabditis elegans* vulval induction. *Bioinformatics* **23**, i499-i507. [[CrossRef](#)]
418. K. Bleakley, G. Biau, J.-P. Vert. 2007. Supervised reconstruction of biological networks with local models. *Bioinformatics* **23**, i57-i65. [[CrossRef](#)]
419. J. Bongard, H. Lipson. 2007. Automated reverse engineering of nonlinear dynamical systems. *Proceedings of the National Academy of Sciences* **104**, 9943-9948. [[CrossRef](#)]
420. Jinki Kim, Doheon Lee, Gwan-Su Yi. 2007. Inferring genetic regulatory networks of the inflammatory bowel disease in human peripheral blood mononuclear cells. *Interdisciplinary Bio Central* **1**, 71-74. [[CrossRef](#)]
421. Claire J. Tomlin, Jeffrey D. Axelrod. 2007. Biology by numbers: mathematical modelling in developmental biology. *Nature Reviews Genetics* **8**, 331-340. [[CrossRef](#)]
422. S. Martin, Z. Zhang, A. Martino, J.-L. Faulon. 2007. Boolean dynamics of genetic regulatory networks inferred from microarray time series data. *Bioinformatics* **23**, 866-874. [[CrossRef](#)]
423. Jing Li, Jianjun Shi. 2007. Knowledge discovery from observational data for process control using causal Bayesian networks. *IEEE Transactions* **39**, 681-690. [[CrossRef](#)]
424. Amar V. Singh, Eric C. Rouchka, Greg A. Rempala, Caleb D. Bastian, Thomas B. Knudsen. 2007. Integrative database management for mouse development: Systems and concepts. *Birth Defects Research Part C: Embryo Today: Reviews* **81**:10.1002/bdrc.v81:1, 1-19. [[CrossRef](#)]
425. Heon-Gyu Lee, Keun-Ho Ryu, Doo-Young Joung. 2007. Constructing Gene Regulatory Networks using Frequent Gene Expression Pattern and Chain Rules. *The KIPS Transactions:PartD* **14D**, 9-20. [[CrossRef](#)]
426. Juhyun Jeong, Doheon Lee. 2007. Inferring candidate regulatory networks in human breast cancer cells. *Interdisciplinary Bio Central* **1**, 26-29. [[CrossRef](#)]
427. Mukesh Bansal, Vincenzo Belcastro, Alberto Ambesi-Impiombato, Diego di Bernardo. 2007. How to infer gene networks from expression profiles. *Molecular Systems Biology* **3**. . [[CrossRef](#)]
428. M. Thai, Z. Cai, D. Du. 2007. Genetic networks: processing data, regulatory network modelling and their analysis. *Optimization Methods and Software* **22**, 169-185. [[CrossRef](#)]
429. OLLI HAAVISTO, HEIKKI HYÖTYNIEMI, CHRISTOPHE ROOS. 2007. STATE SPACE MODELING OF YEAST GENE EXPRESSION DYNAMICS. *Journal of Bioinformatics and Computational Biology* **05**, 31-46. [[CrossRef](#)]
430. Florian Markowetz, Olga G. Troyanskaya. 2007. Computational identification of cellular networks and pathways. *Molecular BioSystems* **3**, 478. [[CrossRef](#)]
431. J. Z. Song, K. M. Duan, T. Ware, M. Surette. 2007. The Wavelet-Based Cluster Analysis for Temporal Gene Expression Data. *EURASIP Journal on Bioinformatics and Systems Biology* **2007**, 1-7. [[CrossRef](#)]
432. Nestor Walter Trepode, Hugo Aguirre Armelin, Michael Bittner, Junior Barrera, Marco Dimas Gubitoso, Ronaldo Fumio Hashimoto. 2007. A Robust Structural PGN Model for Control of Cell-Cycle Progression Stabilized by Negative Feedbacks. *EURASIP Journal on Bioinformatics and Systems Biology* **2007**, 1-11. [[CrossRef](#)]
433. Catherine T. Falk, Stephen J. Finch, Wonkuk Kim, Nitai D. Mukhopadhyay. 2007. Data mining of RNA expression and DNA genotype data: Presentation Group 5 contributions to Genetic Analysis Workshop 15. *Genetic Epidemiology* **31**:10.1002/gepi.v31:1+, S43-S50. [[CrossRef](#)]
434. Isabel Tienda Luna, Yufei Huang, Yufang Yin, Diego P. Ruiz Padillo, M. Carmen Carrion Perez. 2007. Uncovering Gene Regulatory Networks from Time-Series Microarray Data with Variational Bayesian Structural Expectation Maximization. *EURASIP Journal on Bioinformatics and Systems Biology* **2007**, 1-14. [[CrossRef](#)]

435. T. R. Lezon, J. R. Banavar, M. Cieplak, A. Maritan, N. V. Fedoroff. 2006. From the Cover: Using the principle of entropy maximization to infer genetic interaction networks from gene expression patterns. *Proceedings of the National Academy of Sciences* **103**, 19033-19038. [[CrossRef](#)]
436. Gérard Biau, Kevin Bleakley. 2006. Statistical inference on graphs. *Statistics & Decisions* **24**, 209-232. [[CrossRef](#)]
437. X. Chen, M. Chen, K. Ning. 2006. BNArray: an R package for constructing gene regulatory networks from microarray data by using Bayesian network. *Bioinformatics* **22**, 2952-2954. [[CrossRef](#)]
438. M Andrecut, S A Kauffman. 2006. A simple method for reverse engineering causal networks. *Journal of Physics A: Mathematical and General* **39**, L647-L655. [[CrossRef](#)]
439. Minnie M. Sarwal, Li Li. 2006. Designer Genes: Filling the Gap in Transplantation. *Transplantation* **82**, 1261-1272. [[CrossRef](#)]
440. Yonghong Peng. 2006. Integration of gene functional diversity for effective cancer detection. *International Journal of Systems Science* **37**, 931-938. [[CrossRef](#)]
441. A. V. Werhli, M. Grzegorzczak, D. Husmeier. 2006. Comparative evaluation of reverse engineering gene regulatory networks with relevance networks, graphical gaussian models and bayesian networks. *Bioinformatics* **22**, 2523-2531. [[CrossRef](#)]
442. Dougu Nam, Seunghyun Seo, Sangsoo Kim. 2006. An efficient top-down search algorithm for learning Boolean networks of gene expression. *Machine Learning* **65**, 229-245. [[CrossRef](#)]
443. Ioannis Tsamardinos, Laura E. Brown, Constantin F. Aliferis. 2006. The max-min hill-climbing Bayesian network structure learning algorithm. *Machine Learning* **65**, 31-78. [[CrossRef](#)]
444. Yujing Zeng, Javier Garcia-Frias, Adam G. MarshGene Distribution in the Human Genome . [[CrossRef](#)]
445. Hiroyoshi Toyoshiba, Hideko Sone, Takeharu Yamanaka, Frederick M. Parham, Richard D. Irwin, Gary A. Boorman, Christopher J. Portier. 2006. Gene interaction network analysis suggests differences between high and low doses of acetaminophen. *Toxicology and Applied Pharmacology* **215**, 306-316. [[CrossRef](#)]
446. James R. Bradford, Chris J. Needham, Andrew J. Bulpitt, David R. Westhead. 2006. Insights into Protein-Protein Interfaces using a Bayesian Network Prediction Method. *Journal of Molecular Biology* **362**, 365-386. [[CrossRef](#)]
447. Alireza Tamaddon-Nezhad, Raphael Chaleil, Antonis Kakas, Stephen Muggleton. 2006. Application of abductive ILP to learning metabolic network inhibition from temporal data. *Machine Learning* **64**, 209-230. [[CrossRef](#)]
448. W. Zhao, E. Serpedin, E. R. Dougherty. 2006. Inferring gene regulatory networks from time series data using the minimum description length principle. *Bioinformatics* **22**, 2129-2135. [[CrossRef](#)]
449. J. Li, X. Li, H. Su, H. Chen, D. W. Galbraith. 2006. A framework of integrating gene relations from heterogeneous data sources: an experiment on Arabidopsis thaliana. *Bioinformatics* **22**, 2037-2043. [[CrossRef](#)]
450. Dominique Haughton, Arnold Kamis, Patrick Scholten. 2006. A Review of Three Directed Acyclic Graphs Software Packages. *The American Statistician* **60**, 272-286. [[CrossRef](#)]
451. Avi Ma'ayan, Katherine Gardiner, Ravi Iyengar. 2006. The cognitive phenotype of Down syndrome: Insights from intracellular network analysis. *Neurotherapeutics* **3**, 396-406. [[CrossRef](#)]
452. Anthony Almudevar, Lev B. Klebanov, Xing Qiu, Peter Salzman, Andrei Y. Yakovlev. 2006. Utility of correlation measures in analysis of gene expression. *Neurotherapeutics* **3**, 384-395. [[CrossRef](#)]
453. Kwang-Hyun Cho, Jeong-Rae Kim, Songjoon Baek, Hyung-Seok Choi, Sang-Mok Choo. 2006. Inferring biomolecular regulatory networks from phase portraits of time-series expression profiles. *FEBS Letters* **580**, 3511-3518. [[CrossRef](#)]



454. Haja N. Kadarmideen, Peter Rohr, Luc L.G. Janss. 2006. From genetical genomics to systems genetics: potential applications in quantitative genomics and animal breeding. *Mammalian Genome* **17**, 548-564. [[CrossRef](#)]
455. IVAN IVANOV, EDWARD R. DOUGHERTY. 2006. MODELING GENETIC REGULATORY NETWORKS: CONTINUOUS OR DISCRETE?. *Journal of Biological Systems* **14**, 219-229. [[CrossRef](#)]
456. Siliang Zhang, Bang-Ce Ye, Ju Chu, Yingping Zhuang, Meijin Guo. 2006. From multi-scale methodology to systems biology: to integrate strain improvement and fermentation optimization. *Journal of Chemical Technology & Biotechnology* **81**:10.1002/jctb.v81:5, 734-745. [[CrossRef](#)]
457. Siobhan M. Brady, Philip N. Benfey. 2006. A systems approach to understanding root development. *Canadian Journal of Botany* **84**, 695-701. [[CrossRef](#)]
458. K OZBAY, N NOYAN. 2006. Estimation of incident clearance times using Bayesian Networks approach. *Accident Analysis & Prevention* **38**, 542-555. [[CrossRef](#)]
459. Taizo Hanai, Hiroyuki Hamada, Masahiro Okamoto. 2006. Application of bioinformatics for DNA microarray data to bioscience, bioengineering and medical fields. *Journal of Bioscience and Bioengineering* **101**, 377-384. [[CrossRef](#)]
460. Ali Abbas, Lei LiuBioinformatics . [[CrossRef](#)]
461. Xiaobo Zhou, Xiaodong Wang, Edward R. DoughertyGenomic Networks: Statistical Inference from Microarray Data . [[CrossRef](#)]
462. Lei Liu, Ali AbbasBioinformatics . [[CrossRef](#)]
463. J.H.G.M. van Beek. 2006. Channeling the Data Flood: Handling Large-Scale Biomolecular Measurements In Silico. *Proceedings of the IEEE* **94**, 692-709. [[CrossRef](#)]
464. STERGIOS PAPADIMITRIOU, SEFERINA MAVROUDI, SPIRIDON D. LIKOTHANASSIS. 2006. MUTUAL INFORMATION CLUSTERING FOR EFFICIENT MINING OF FUZZY ASSOCIATION RULES WITH APPLICATION TO GENE EXPRESSION DATA ANALYSIS. *International Journal on Artificial Intelligence Tools* **15**, 227-250. [[CrossRef](#)]
465. C YOO, G COOPER, M SCHMIDT. 2006. A control study to evaluate a computer-based microarray experiment design recommendation system for gene-regulation pathways discovery. *Journal of Biomedical Informatics* **39**, 126-146. [[CrossRef](#)]
466. XUTAO DENG, HUIMIN GENG, HESHAM H. ALI. 2006. JOINT LEARNING OF GENE FUNCTIONS — A BAYESIAN NETWORK MODEL APPROACH. *Journal of Bioinformatics and Computational Biology* **04**, 217-239. [[CrossRef](#)]
467. Z. Li, S. M. Shaw, M. J. Yedwabnick, C. Chan. 2006. Using a state-space model with hidden variables to infer transcription factor activities. *Bioinformatics* **22**, 747-754. [[CrossRef](#)]
468. Irit Gat-Viks, Amos Tanay, Daniela Raijman, Ron Shamir. 2006. A Probabilistic Methodology for Integrating Knowledge and Experiments on Biological Networks. *Journal of Computational Biology* **13**:2, 165-181. [[Abstract](#)] [[Full Text PDF](#)] [[Full Text PDF with Links](#)]
469. Chen-Hsiang Yeang, Tommi Jaakkola. 2006. Modeling the Combinatorial Functions of Multiple Transcription Factors. *Journal of Computational Biology* **13**:2, 463-480. [[Abstract](#)] [[Full Text PDF](#)] [[Full Text PDF with Links](#)]
470. MINGYI WANG, HUIJUAN LU, ZUOZHOU CHEN, PING WU. 2006. MINING GENE EXPRESSION DATABASES FOR LOCAL CAUSAL RELATIONSHIPS USING A SIMPLE CONSTRAINT-BASED ALGORITHM. *International Journal of Pattern Recognition and Artificial Intelligence* **20**, 311-327. [[CrossRef](#)]
471. EDWARD R. DOUGHERTY, ULISSES BRAGA-NETO. 2006. EPISTEMOLOGY OF COMPUTATIONAL BIOLOGY: MATHEMATICAL MODELS AND EXPERIMENTAL

PREDICTION AS THE BASIS OF THEIR VALIDITY. *Journal of Biological Systems* **14**, 65-90. [[CrossRef](#)]

472. A.P. Quayle, S. Bullock. 2006. Modelling the evolution of genetic regulatory networks. *Journal of Theoretical Biology* **238**, 737-753. [[CrossRef](#)]
473. Alicja Gruźdź, Aleksandra Ihnatowicz, Dominik Ślęzak. 2006. Interactive Gene Clustering—A Case Study of Breast Cancer Microarray Data. *Information Systems Frontiers* **8**, 21-27. [[CrossRef](#)]
474. B. A. Novak, A. N. Jain. 2006. Pathway recognition and augmentation by computational analysis of microarray expression data. *Bioinformatics* **22**, 233-241. [[CrossRef](#)]
475. Sui Huang Multistability and Multicellularity 293-326. [[CrossRef](#)]
476. Irit Gat-Viks, Richard M. Karp, Ron Shamir, Roded Sharan. 2006. Reconstructing Chain Functions in Genetic Networks. *SIAM Journal on Discrete Mathematics* **20**, 727-740. [[CrossRef](#)]
477. Veronica Vinciotti, Allan Tucker, Paul Kellam, Xiaohui Liu. 2006. Robust Selection of Predictive Genes via a Simple Classifier. *Applied Bioinformatics* **5**, 1-11. [[CrossRef](#)]
478. H. Li, M. Zhan. 2006. Systematic intervention of transcription for identifying network response to disease and cellular phenotypes. *Bioinformatics* **22**, 96-102. [[CrossRef](#)]
479. Seiya Imoto, Hiroshi Matsuno, Satoru Miyano Gene Networks 205-228. [[CrossRef](#)]
480. M. J. BERRYMAN, A. ALLISON, C. R. WILKINSON, D. ABBOTT. 2005. REVIEW OF SIGNAL PROCESSING IN GENETICS. *Fluctuation and Noise Letters* **05**, R13-R35. [[CrossRef](#)]
481. Duncan C. Thomas. 2005. Discussion on "Statistical Issues Arising in the Women's Health Initiative". *Biometrics* **61**:10.1111/biom.2005.61.issue-4, 930-933. [[CrossRef](#)]
482. S WELCH, J ROE, S DAS, Z DONG, R HE, M KIRKHAM. 2005. Merging genomic control networks and soil-plant-atmosphere-continuum models. *Agricultural Systems* **86**, 243-274. [[CrossRef](#)]
483. YOSHINORI TAMADA, HIDEO BANNAI, SEIYA IMOTO, TOSHIAKI KATAYAMA, MINORU KANEHISA, SATORU MIYANO. 2005. UTILIZING EVOLUTIONARY INFORMATION AND GENE EXPRESSION DATA FOR ESTIMATING GENE NETWORKS WITH BAYESIAN NETWORK MODELS. *Journal of Bioinformatics and Computational Biology* **03**, 1295-1313. [[CrossRef](#)]
484. Junbai Wang, Leo Wang-Kit Cheung, Jan Delabie. 2005. New probabilistic graphical models for genetic regulatory networks studies. *Journal of Biomedical Informatics* **38**, 443-455. [[CrossRef](#)]
485. Atul J. Butte Relevance networks . [[CrossRef](#)]
486. Vassily Hatzimanikatis, Eleftherios T. Papoutsakis Extracting networks from expression data . [[CrossRef](#)]
487. F. Markowetz, J. Bloch, R. Spang. 2005. Non-transcriptional pathway features reconstructed from secondary effects of RNA interference. *Bioinformatics* **21**, 4026-4032. [[CrossRef](#)]
488. B. Xing, M. J. van der Laan. 2005. A causal inference approach for constructing transcriptional regulatory networks. *Bioinformatics* **21**, 4007-4013. [[CrossRef](#)]
489. Doheon Lee. 2005. Component-based software architecture for biosystem reverse engineering. *Biotechnology and Bioprocess Engineering* **10**, 400-407. [[CrossRef](#)]
490. Peter M. Bowers, Brian D. O'Connor, Shawn J. Cokus, Einat Sprinzak, Todd O. Yeates, David Eisenberg. 2005. Utilizing logical relationships in genomic data to decipher cellular processes. *FEBS Journal* **272**:10.1111/ejb.2005.272.issue-20, 5110-5118. [[CrossRef](#)]
491. CHEN-HSIANG YEANG, TOMMI JAAKKOLA. 2005. TIME SERIES ANALYSIS OF GENE EXPRESSION AND LOCATION DATA. *International Journal on Artificial Intelligence Tools* **14**, 755-769. [[CrossRef](#)]

492. A. S. Rodin, E. Boerwinkle. 2005. Mining genetic epidemiology data with Bayesian networks I: Bayesian networks and example application (plasma apoE levels). *Bioinformatics* **21**, 3273-3278. [[CrossRef](#)]
493. WAI-KI CHING, MICHAEL M. NG, ERIC S. FUNG, TATSUYA AKUTSU. 2005. ON CONSTRUCTION OF STOCHASTIC GENETIC NETWORKS BASED ON GENE EXPRESSION SEQUENCES. *International Journal of Neural Systems* **15**, 297-310. [[CrossRef](#)]
494. Kwang-Hyun Cho, Sung-Young Shin, Sang-Mok Choo. 2005. Unravelling the functional interaction structure of a cellular network from temporal slope information of experimental data. *FEBS Journal* **272**, 3950-3959. [[CrossRef](#)]
495. S. Rogers, M. Girolami. 2005. A Bayesian regression approach to the inference of regulatory networks from gene expression data. *Bioinformatics* **21**, 3131-3137. [[CrossRef](#)]
496. Isaac S. Kohane, Atul Butte *Bioinformatics in Functional Genomics* . [[CrossRef](#)]
497. S. Kaski, J. Nikkila, J. Sinkkonen, L. Lahti, J.E.A. Knuuttila, C. Roos. 2005. Associative Clustering for Exploring Dependencies between Functional Genomics Data Sets. *IEEE/ACM Transactions on Computational Biology and Bioinformatics* **2**, 203-216. [[CrossRef](#)]
498. Helge Langseth, Thomas D. Nielsen. 2005. Latent Classification Models. *Machine Learning* . [[CrossRef](#)]
499. Eran Segal, Nir Friedman, Naftali Kaminski, Aviv Regev, Daphne Koller. 2005. From signatures to models: understanding cancer using microarrays. *Nature Genetics* **37**, S38-S45. [[CrossRef](#)]
500. S. Bulashevskaya, R. Eils. 2005. Inferring genetic regulatory logic from expression data. *Bioinformatics* **21**, 2706-2713. [[CrossRef](#)]
501. I. Ben-Gal, A. Shani, A. Gohr, J. Grau, S. Arviv, A. Shmilovici, S. Posch, I. Grosse. 2005. Identification of transcription factor binding sites with variable-order Bayesian networks. *Bioinformatics* **21**, 2657-2666. [[CrossRef](#)]
502. P. H. Lee, D. Lee. 2005. Modularized learning of genetic interaction networks from biological annotations and mRNA expression data. *Bioinformatics* **21**, 2739-2747. [[CrossRef](#)]
503. Dabao Zhang, Martin T. Wells, Christine D. Smart, William E. Fry. 2005. Bayesian Normalization and Identification for Differential Gene Expression Data. *Journal of Computational Biology* **12**:4, 391-406. [[Abstract](#)] [[Full Text PDF](#)] [[Full Text PDF with Links](#)]
504. R. D. King, S. M. Garrett, G. M. Coghill. 2005. On the use of qualitative reasoning to simulate and identify metabolic pathways. *Bioinformatics* **21**, 2017-2026. [[CrossRef](#)]
505. D. Pe'er. 2005. Bayesian Network Analysis of Signaling Networks: A Primer. *Science's STKE* **2005**, pl4-pl4. [[CrossRef](#)]
506. J. Jeremy Rice, Aaron Kershenbaum, Gustavo Stolovitzky *Analyzing and reconstructing gene regulatory networks* . [[CrossRef](#)]
507. Michael J. Thompson, Michael E. Driscoll, Timothy S. Gardner, James J. Collins *Reverse engineering gene regulatory networks* . [[CrossRef](#)]
508. Kara Dolinski, Olga Troyanskaya *Functional genomics in Saccharomyces cerevisiae* . [[CrossRef](#)]
509. Katia Basso, Adam A Margolin, Gustavo Stolovitzky, Ulf Klein, Riccardo Dalla-Favera, Andrea Califano. 2005. Reverse engineering of regulatory networks in human B cells. *Nature Genetics* **37**, 382-390. [[CrossRef](#)]
510. Kathleen Marchal, Kristof Engelen, Frank De Smet, Bart De Moor *Computational Biology and Toxicogenomics* 37-92. [[CrossRef](#)]
511. J. Schafer, K. Strimmer. 2005. An empirical Bayes approach to inferring large-scale gene association networks. *Bioinformatics* **21**, 754-764. [[CrossRef](#)]
512. J. J. Rice, Y. Tu, G. Stolovitzky. 2005. Reconstructing biological networks using conditional correlation analysis. *Bioinformatics* **21**, 765-773. [[CrossRef](#)]



513. Lei M. Li, Henry Horng-Shing Lu. 2005. Explore Biological Pathways from Noisy Array Data by Directed Acyclic Boolean Networks. *Journal of Computational Biology* **12**:2, 170-185. [[Abstract](#)] [[Full Text PDF](#)] [[Full Text PDF with Links](#)]
514. Biao Xing, Mark J. van der Laan. 2005. A Statistical Method for Constructing Transcriptional Regulatory Networks Using Gene Expression and Sequence Data. *Journal of Computational Biology* **12**:2, 229-246. [[Abstract](#)] [[Full Text PDF](#)] [[Full Text PDF with Links](#)]
515. Sergej V. Aksenov, Bruce Church, Anjali Dhiman, Anna Georgieva, Ramesh Sarangapani, Gabriel Helmlinger, Iya G. Khalil. 2005. An integrated approach for inference and mechanistic modeling for advancing drug development. *FEBS Letters* **579**, 1878-1883. [[CrossRef](#)]
516. T. S. Mikkelsen, J. E. Galagan, J. P. Mesirov. 2005. Improving genome annotations using phylogenetic profile anomaly detection. *Bioinformatics* **21**, 464-470. [[CrossRef](#)]
517. J. Kasturi, R. Acharya. 2005. Clustering of diverse genomic data using information fusion. *Bioinformatics* **21**, 423-429. [[CrossRef](#)]
518. M STYCZYNSKI, G STEPHANOPOULOS. 2005. Overview of computational methods for the inference of gene regulatory networks. *Computers & Chemical Engineering* **29**, 519-534. [[CrossRef](#)]
519. A GUPTA, J VARNER, C MARANAS. 2005. Large-scale inference of the transcriptional regulation of. *Computers & Chemical Engineering* **29**, 565-576. [[CrossRef](#)]
520. Andrei Rodin, Thomas H. Mosley , Jr., Andrew G. Clark, Charles F. Sing, Eric Boerwinkle. 2005. Mining Genetic Epidemiology Data with Bayesian Networks Application to APOE Gene Variation and Plasma Lipid Levels. *Journal of Computational Biology* **12**:1, 1-11. [[Abstract](#)] [[Full Text PDF](#)] [[Full Text PDF with Links](#)]
521. JEONG-HO CHANG, KYU-BAEK HWANG, S. JUNE OH, BYOUNG-TAK ZHANG. 2005. BAYESIAN NETWORK LEARNING WITH FEATURE ABSTRACTION FOR GENE-DRUG DEPENDENCY ANALYSIS. *Journal of Bioinformatics and Computational Biology* **03**, 61-77. [[CrossRef](#)]
522. S. Ott, A. Hansen, S.-Y. Kim, S. Miyano. 2005. Superiority of network motifs over optimal networks and an application to the revelation of gene network evolution. *Bioinformatics* **21**, 227-238. [[CrossRef](#)]
523. M DASIKA, A GUPTA, C MARANAS. 2005. DEMSIM: a discrete event based mechanistic simulation platform for gene expression and regulation dynamics. *Journal of Theoretical Biology* **232**, 55-69. [[CrossRef](#)]
524. David B. Dunson Bayesian Biostatistics 743-761. [[CrossRef](#)]
525. Stephen M. Welch, Zhanshan Dong, Judith L. Roe, Sanjoy Das. 2005. Flowering time control: gene network modelling and the link to quantitative genetics. *Australian Journal of Agricultural Research* **56**, 919. [[CrossRef](#)]
526. B. de Bivort, S. Huang, Y. Bar-Yam. 2004. Dynamics of cellular level function and regulation derived from murine expression array data. *Proceedings of the National Academy of Sciences* **101**, 17687-17692. [[CrossRef](#)]
527. Mathäus Dejori, Martin Stetter. 2004. Identifying Interventional and Pathogenic Mechanisms by Generative Inverse Modeling of Gene Expression Profiles. *Journal of Computational Biology* **11**:6, 1135-1148. [[Abstract](#)] [[Full Text PDF](#)] [[Full Text PDF with Links](#)]
528. Irit Gat-Viks, Amos Tanay, Ron Shamir. 2004. Modeling and Analysis of Heterogeneous Regulation in Biological Networks. *Journal of Computational Biology* **11**:6, 1034-1049. [[Abstract](#)] [[Full Text PDF](#)] [[Full Text PDF with Links](#)]
529. Xiaojiang Xu, Lianshui Wang, Dafu Ding. 2004. Learning module networks from genome-wide location and expression data. *FEBS Letters* **578**, 297-304. [[CrossRef](#)]

530. 2004. Construction of Gene Interaction Networks from Gene Expression Data Based on Evolutionary Computation. *Journal of Control, Automation and Systems Engineering* **10**, 1189-1195. [[CrossRef](#)]
531. GUILLAUME BOURQUE, DAVID SANKOFF. 2004. IMPROVING GENE NETWORK INFERENCE BY COMPARING EXPRESSION TIME-SERIES ACROSS SPECIES, DEVELOPMENTAL STAGES OR TISSUES. *Journal of Bioinformatics and Computational Biology* **02**, 765-783. [[CrossRef](#)]
532. John M Hancock Bayesian Network (Belief Network, Probabilistic Network, Causal Network, Knowledge Map) . [[CrossRef](#)]
533. Radhakrishnan Nagarajan, Jane E. Aubin, Charlotte A. Peterson. 2004. Modeling genetic networks from clonal analysis. *Journal of Theoretical Biology* **230**, 359-373. [[CrossRef](#)]
534. S.D. Bay, L. Chrisman, A. Pohorille, J. Shrager. 2004. Temporal Aggregation Bias and Inference of Causal Regulatory Networks. *Journal of Computational Biology* **11**:5, 971-985. [[Abstract](#)] [[Full Text PDF](#)] [[Full Text PDF with Links](#)]
535. Sarika Mehra, Wei-Shou Hu, George Karypis. 2004. A Boolean algorithm for reconstructing the structure of regulatory networks. *Metabolic Engineering* **6**, 326-339. [[CrossRef](#)]
536. YUKIHIRO MAKI, YORIKO TAKAHASHI, YUJI ARIKAWA, SHOJI WATANABE, KEN AOSHIMA, YUKIHIRO EGUCHI, TAKANORI UEDA, SACHIYO ABURATANI, SATORU KUHARA, MASAHIRO OKAMOTO. 2004. AN INTEGRATED COMPREHENSIVE WORKBENCH FOR INFERRING GENETIC NETWORKS: VOYAGENE. *Journal of Bioinformatics and Computational Biology* **02**, 533-550. [[CrossRef](#)]
537. SUN-CHONG WANG. 2004. RECONSTRUCTING GENETIC NETWORKS FROM TIME ORDERED GENE EXPRESSION DATA USING BAYESIAN METHOD WITH GLOBAL SEARCH ALGORITHM. *Journal of Bioinformatics and Computational Biology* **02**, 441-458. [[CrossRef](#)]
538. M. Dejori, B. Schuermann, M. Stetter. 2004. Hunting Drug Targets by Systems-Level Modeling of Gene Expression Profiles. *IEEE Transactions on Nanobioscience* **3**, 180-191. [[CrossRef](#)]
539. Paul Helman, Robert Veroff, Susan R. Atlas, Cheryl Willman. 2004. A Bayesian Network Classification Methodology for Gene Expression Data. *Journal of Computational Biology* **11**:4, 581-615. [[Abstract](#)] [[Full Text PDF](#)] [[Full Text PDF with Links](#)]
540. R LAUBENBACHER. 2004. A computational algebra approach to the reverse engineering of gene regulatory networks. *Journal of Theoretical Biology* **229**, 523-537. [[CrossRef](#)]
541. Takeharu Yamanaka, Hiroyoshi Toyoshiba, Hideko Sone, Frederick M. Parham, Christopher J. Portier. 2004. The TAO-Gen Algorithm for Identifying Gene Interaction Networks with Application to SOS Repair in E. coli. *Environmental Health Perspectives* **112**, 1614-1621. [[CrossRef](#)]
542. Mathäus Dejori, Anton Schwaighofer, Volker Tresp, Martin Stetter. 2004. Mining Functional Modules in Genetic Networks with Decomposable Graphical Models. *OMICS: A Journal of Integrative Biology* **8**:2, 176-188. [[Abstract](#)] [[Full Text PDF](#)] [[Full Text PDF with Links](#)]
543. Hiroyoshi Toyoshiba, Takeharu Yamanaka, Hideko Sone, Frederick M. Parham, Nigel J. Walker, Jeanelle Martinez, Christopher J. Portier. 2004. Gene Interaction Network Suggests Dioxin Induces a Significant Linkage between Aryl Hydrocarbon Receptor and Retinoic Acid Receptor Beta. *Environmental Health Perspectives* **112**, 1217-1224. [[CrossRef](#)]
544. Yu Xia, Haiyuan Yu, Ronald Jansen, Michael Seringhaus, Sarah Baxter, Dov Greenbaum, Hongyu Zhao, Mark Gerstein. 2004. ANALYZING CELLULAR BIOCHEMISTRY IN TERMS OF MOLECULAR NETWORKS. *Annual Review of Biochemistry* **73**, 1051-1087. [[CrossRef](#)]
545. Ali E. Abbas, Susan P. Holmes. 2004. Bioinformatics and Management Science: Some Common Tools and Techniques. *Operations Research* **52**, 165-190. [[CrossRef](#)]

546. Michael A. Beer, Saeed Tavazoie. 2004. Predicting Gene Expression from Sequence. *Cell* **117**, 185-198. [[CrossRef](#)]
547. Amos Tanay, Ron Shamir. 2004. Multilevel Modeling and Inference of Transcription Regulation. *Journal of Computational Biology* **11**:2-3, 357-375. [[Abstract](#)] [[Full Text PDF](#)] [[Full Text PDF with Links](#)]
548. J.Jeremy Rice, Gustavo Stolovitzky. 2004. Making the most of it: pathway reconstruction and integrative simulation using the data at hand. *Drug Discovery Today: BIOSILICO* **2**, 70-77. [[CrossRef](#)]
549. SEIYA IMOTO, TOMOYUKI HIGUCHI, TAKAO GOTO, KOUSUKE TASHIRO, SATORU KUHARA, SATORU MIYANO. 2004. COMBINING MICROARRAYS AND BIOLOGICAL KNOWLEDGE FOR ESTIMATING GENE NETWORKS VIA BAYESIAN NETWORKS. *Journal of Bioinformatics and Computational Biology* **02**, 77-98. [[CrossRef](#)]
550. Ning Sun, Hongyu Zhao. 2004. Genomic approaches in dissecting complex biological pathways. *Pharmacogenomics* **5**:10.1517/pgs.2004.5.issue-2, 163-179. [[CrossRef](#)]
551. Xiaobo Zhou, Xiaodong Wang, Edward R. Dougherty, Daniel Russ, Edward Suh. 2004. Gene Clustering Based on Clusterwise Mutual Information. *Journal of Computational Biology* **11**:1, 147-161. [[Abstract](#)] [[Full Text PDF](#)] [[Full Text PDF with Links](#)]
552. STERGIOS PAPADIMITRIOU, SPIRIDON D. LIKOTHANASSIS. 2004. KERNEL-BASED SELF-ORGANIZED MAPS TRAINED WITH SUPERVISED BIAS FOR GENE EXPRESSION DATA ANALYSIS. *Journal of Bioinformatics and Computational Biology* **01**, 647-680. [[CrossRef](#)]
553. Elissa J. Chesler, Robert W. WilliamsBrain Gene Expression: Genomics and Genetics 59-95. [[CrossRef](#)]
554. J. M. PEÑA, J. A. LOZANO, P. LARRAÑAGA. 2004. UNSUPERVISED LEARNING OF BAYESIAN NETWORKS VIA ESTIMATION OF DISTRIBUTION ALGORITHMS: AN APPLICATION TO GENE EXPRESSION DATA CLUSTERING. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems* **12**, 63-82. [[CrossRef](#)]
555. Kazumi Hakamada, Taizo Hanai, Hiroyuki Honda, Takeshi Kobayashi. 2004. A preprocessing method for inferring genetic interaction from gene expression data using Boolean algorithm. *Journal of Bioscience and Bioengineering* **98**, 457-463. [[CrossRef](#)]
556. Andrei Dragomir, Seferina Mavroudi, Anastasios Bezerianos. 2004. Som-Based Class Discovery Exploring the ICA-Reduced Features of Microarray Expression Profiles. *Comparative and Functional Genomics* **5**, 596-616. [[CrossRef](#)]
557. Dongyan Yang, Stanislav O Zakharkin, Grier P Page, Jacob P L Brand, Jode W Edwards, Alfred A Bartolucci, David B Allison. 2004. Applications of Bayesian Statistical Methods in Microarray Data Analysis. *American Journal of Pharmacogenomics* **4**, 53-62. [[CrossRef](#)]
558. J Huang. 2003. Clustering gene expression pattern and extracting relationship in gene network based on artificial neural networks. *Journal of Bioscience and Bioengineering* **96**, 421-428. [[CrossRef](#)]
559. SUNG-BAE CHO, HONG-HEE WON. 2003. DATA MINING FOR GENE EXPRESSION PROFILES FROM DNA MICROARRAY. *International Journal of Software Engineering and Knowledge Engineering* **13**, 593-608. [[CrossRef](#)]
560. Yuk Fai Leung, Duccio Cavalieri. 2003. Fundamentals of cDNA microarray data analysis. *Trends in Genetics* **19**, 649-659. [[CrossRef](#)]
561. Seiya Imoto, Christopher J. Savoie, Sachiyo Aburatani, Sunyong Kim, Kousuke Tashiro, Satoru Kuhara, Satoru Miyano. 2003. Use of Gene Networks for Identifying and Validating Drug Targets. *Journal of Bioinformatics and Computational Biology* **01**, 459-474. [[CrossRef](#)]

562. Tero Aittokallio, Markus Kurki, Olli Nevalainen, Tuomas Nikula, Anne West, Riitta Lahesmaa. 2003. Computational Strategies for Analyzing Data in Gene Expression Microarray Experiments. *Journal of Bioinformatics and Computational Biology* **01**, 541-586. [[CrossRef](#)]
563. Jun Xie, Peter Bentler. 2003. Covariance Structure Models for Gene Expression Microarray Data. *Structural Equation Modeling: A Multidisciplinary Journal* **10**, 566-582. [[CrossRef](#)]
564. O. G. Troyanskaya, K. Dolinski, A. B. Owen, R. B. Altman, D. Botstein. 2003. A Bayesian framework for combining heterogeneous data sources for gene function prediction (in *Saccharomyces cerevisiae*). *Proceedings of the National Academy of Sciences* **100**, 8348-8353. [[CrossRef](#)]
565. Seiya Imoto, Sunyong Kim, Takao Goto, Sachiyo Aburatani, Kousuke Tashiro, Satoru Kuhara, Satoru Miyano. 2003. Bayesian Network and Nonparametric Heteroscedastic Regression for Nonlinear Modeling of Genetic Network. *Journal of Bioinformatics and Computational Biology* **01**, 231-252. [[CrossRef](#)]
566. Trey Ideker, Douglas Lauffenburger. 2003. Building with a scaffold: emerging strategies for high-to low-level cellular modeling. *Trends in Biotechnology* **21**, 255-262. [[CrossRef](#)]
567. J. Tegner, M. K. S. Yeung, J. Hasty, J. J. Collins. 2003. Reverse engineering gene networks: Integrating genetic perturbations with dynamical modeling. *Proceedings of the National Academy of Sciences* **100**, 5944-5949. [[CrossRef](#)]
568. X Ji. 2003. Mining gene expression data using a novel approach based on hidden Markov models. *FEBS Letters* **542**, 125-131. [[CrossRef](#)]
569. Denise M Wolf, Adam P Arkin. 2003. Motifs, modules and games in bacteria. *Current Opinion in Microbiology* **6**, 125-134. [[CrossRef](#)]
570. Javier Herrero, Ramón Díaz-Uriarte, Joaquín Dopazo. 2003. An Approach to Inferring Transcriptional Regulation Among Genes From Large-Scale Expression Data. *Comparative and Functional Genomics* **4**, 148-154. [[CrossRef](#)]
571. FARAMARZ VALAFAR. 2002. Pattern Recognition Techniques in Microarray Data Analysis. *Annals of the New York Academy of Sciences* **980**:10.1111/nyas.2002.980.issue-1, 41-64. [[CrossRef](#)]
572. Atul Butte. 2002. The use and analysis of microarray data. *Nature Reviews Drug Discovery* **1**, 951-960. [[CrossRef](#)]
573. Paul Brazhnik, Alberto de la Fuente, Pedro Mendes. 2002. Gene networks: how to put the function in genomics. *Trends in Biotechnology* **20**, 467-472. [[CrossRef](#)]
574. Grant S. Heffelfinger, Anthony Martino, Andrey Gorin, Ying Xu, Mark D. Rintoul III, Al Geist, Hashim M. Al-Hashimi, George S. Davidson, Jean Loup Faulon, Laurie J. Frink, David M. Haaland, William E. Hart, Erik Jakobsson, Todd Lane, Ming Li, Phil Locascio, Frank Olken, Victor Olman, Brian Palenik, Steven J. Plimpton, Diana C. Roe, Nagiza F. Samatova, Manesh Shah, Arie Shoshoni, Charlie E. M. Strauss, Edward V. Thomas, Jerilyn A. Timlin, Dong Xu. 2002. Carbon Sequestration in *Synechococcus* Sp.: From Molecular Machines to Hierarchical Modeling. *OMICS: A Journal of Integrative Biology* **6**:4, 305-330. [[Abstract](#)] [[Full Text PDF](#)] [[Full Text PDF with Links](#)]
575. 2002. Bayesian Network-Based Analysis on Clinical Data of Infertility Patients. *The KIPS Transactions:PartB* **9B**, 625-634. [[CrossRef](#)]
576. S Bay. 2002. Revising regulatory networks: from expression data to linear causal models. *Journal of Biomedical Informatics* **35**, 289-297. [[CrossRef](#)]
577. 2002. Discussion on the meeting on 'Statistical modelling and analysis of genetic data'. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* **64**:10.1111/rssb.2002.64.issue-4, 737-775. [[CrossRef](#)]
578. K. Sachs, D. Gifford, T. Jaakkola, P. Sorger, D. A. Lauffenburger. 2002. Bayesian Network Approach to Cell Signaling Pathway Modeling. *Science Signaling* **2002**, pe38-pe38. [[CrossRef](#)]

579. Alex Gilman, Adam P. Arkin. 2002. GENETIC "CODE": Representations and Dynamical Models of Genetic Components and Networks. *Annual Review of Genomics and Human Genetics* 3, 341-369. [[CrossRef](#)]
580. EP van Someren, LFA Wessels, E Backer, MJT Reinders. 2002. Genetic network modeling. *Pharmacogenomics* 3, 507-525. [[CrossRef](#)]
581. Nila Banerjee, Michael Q Zhang. 2002. Functional genomics as applied to mapping transcription regulatory networks. *Current Opinion in Microbiology* 5, 313-317. [[CrossRef](#)]
582. Yonatan Bilu, Michal Linial. 2002. The Advantage of Functional Prediction Based on Clustering of Yeast Genes and Its Correlation with Non-Sequence Based Classifications. *Journal of Computational Biology* 9:2, 193-210. [[Abstract](#)] [[Full Text PDF](#)] [[Full Text PDF with Links](#)]
583. John J Wyrick, Richard A Young. 2002. Deciphering gene expression regulatory networks. *Current Opinion in Genetics & Development* 12, 130-136. [[CrossRef](#)]
584. Hidde de Jong. 2002. Modeling and Simulation of Genetic Regulatory Systems: A Literature Review. *Journal of Computational Biology* 9:1, 67-103. [[Abstract](#)] [[Full Text PDF](#)] [[Full Text PDF with Links](#)]
585. Adam P. Arkin. 2001. Synthetic cell biology. *Current Opinion in Biotechnology* 12, 638-644. [[CrossRef](#)]
586. R.B. Altman. 2001. Challenges for intelligent systems in biology. *IEEE Intelligent Systems* 16, 14-18. [[CrossRef](#)]
587. Minoru Kanehisa. 2001. Prediction of higher order functional networks from genomic data. *Pharmacogenomics* 2, 373-385. [[CrossRef](#)]
588. Trey Ideker, Timothy Galitski, Leroy Hood. 2001. A NEW APPROACH TO DECODING LIFE : Systems Biology. *Annual Review of Genomics and Human Genetics* 2, 343-372. [[CrossRef](#)]
589. Sui Huang. 2001. Genomics, complexity and drug discovery: insights from Boolean network models of cellular regulation. *Pharmacogenomics* 2, 203-222. [[CrossRef](#)]
590. Marc Ladanyi, Wing C. Chan, Timothy J. Triche, William L. Gerald. 2001. Expression Profiling of Human Tumors: The End of Surgical Pathology?. *The Journal of Molecular Diagnostics* 3, 92-97. [[CrossRef](#)]
591. P Meltzer. 2001. Spotting the target: microarrays for disease gene discovery. *Current Opinion in Genetics & Development* 11, 258-263. [[CrossRef](#)]
592. R Altman. 2001. Whole-genome expression analysis: challenges beyond clustering. *Current Opinion in Structural Biology* 11, 340-347. [[CrossRef](#)]
593. Sarah Pollock, Hershel M. SaferChapter 20. Bioinformatics in the drug discovery process 201-210. [[CrossRef](#)]
594. 2001. Current Awareness on Comparative and Functional Genomics. *Comparative and Functional Genomics* 2, 187-194. [[CrossRef](#)]
595. Paola Lecca, Alida PalmisanoThe Present and the Future Perspectives of Biological Network Inference 118-140. [[CrossRef](#)]
596. Juan I. Alonso-Barba, Jens D. Nielsen, Luis de la Ossa, Jose M. PuertaLearning Probabilistic Graphical Models 223-236. [[CrossRef](#)]
597. Andrew KuznetsovSynthetic Biology as a Proof of Systems Biology 1981-1998. [[CrossRef](#)]
598. Sachiyo Aburatani, Hiroyuki TohStructural Equation Modeling for Systems Biology 458-467. [[CrossRef](#)]
599. Ramesh Ram, Madhu ChettyModelling Gene Regulatory Networks Using Computational Intelligence Techniques 333-352. [[CrossRef](#)]
600. Shruti Mishra, Debahuti MishraMethodologies for Modeling Gene Regulatory Networks 426-436. [[CrossRef](#)]