Identification of Temporal Association Rules from Time-series Microarray Data Set

Hojung Nam  
Department of Bio and Brain Engineering, KAIST, 373–1 Guseong-dong, Yuseong-gu, Daejeon, Korea.  
+82-42-869-5356  
hjnam@kaist.ac.kr

KiYoung Lee  
Department of Bioengineering  
University of California at San Diego, La Jolla, California 92093, USA  
+1-858-822-4665  
kylee@bioeng.ucsd.edu

Doheon Lee  
Department of Bio and Brain Engineering, KAIST, 373–1 Guseong-dong, Yuseong-gu, Daejeon, Korea.  
+82-42-869-4316  
dhlee@kaist.ac.kr

ABSTRACT
One of the most challenging problems in mining gene expression data is to identify how the expression of any particular gene affects the expression of other genes. To elucidate the relationships between genes, an association rule mining (ARM) method has been applied to microarray gene expression data. A conventional ARM method, however, has a limit on extracting temporal dependencies between genes, though the temporal information is indispensable to discover underlying regulation mechanisms in biological pathways. In this paper, therefore, we propose a novel method, referred to as temporal association rule mining (TARM), which can extract temporal dependencies among related genes. A temporal association rule has the form \([\text{gene A} \downarrow, \text{gene B} \uparrow] \rightarrow (7 \text{ min})[\text{gene C}],\) which represents that high expression level of \(\text{gene A}\) and significant repression of \(\text{gene B}\) followed by significant expression of \(\text{gene C}\) after 7 minutes. The proposed TARM method is tested with \(\text{Saccharomyces cerevisiae}\) cell cycle time-series microarray gene expression data set. In the parameter fitting phase of TARM, the best parameter set \([\text{threshold} = \pm 0.8, \text{support cutoff} = 3 \text{ transactions, confidence cutoff} = 90\%],\) which extracted the most number of correct associations in KEGG cell cycle pathway, has been chosen for rule mining phase. Furthermore, comparing the precision scores of TARM (0.38) and Bayesian network (0.16), TARM method showed better accuracy. With the best parameter set, numbers of temporal association rules with five transcriptional time delays \((0, 7, 14, 21, 28 \text{ minutes})\) are extracted from gene expression data of 799 genes which are pre-identified cell cycle relevant genes, while comparably small number of rules are extracted from random shuffled gene expression data of 799 genes. From the extracted temporal association rules, associated genes which play same role of biological processes within short transcriptional time delay and some temporal dependencies between genes with specific biological processes are identified.

Categories and Subject Descriptors  
J.3 [Life and medical sciences]: Biology and genetics

General Terms  
Algorithms.

Keywords  
Data mining, association rule mining, gene expression.

1. INTRODUCTION
The genome of an organism plays a central role in the control of cellular processes such as genetic regulation, metabolic pathway, and signal transduction. Because these processes are very complex and comprised of many genetic interacting elements, it is hard to discover those interacting elements in the complex biological regulations. Since microarray technique allows researchers to simultaneously observe the expression levels of thousands of genes in a single experiment, there have been many studies to discover global genetic regulation from microarray gene expression data by using various computational methods to uncover the hidden roles of genetic elements, such as clustering techniques to identify clusters of co-expressed genes [1-3], network inference techniques to construct the genome-wide regulatory network models [4-9].

One of the most challenging problems in analyzing gene expression data is to determine how the expression of any particular gene might affect the expression of other genes. To find the relationships among different genes, an association rule mining (ARM) method has been applied to gene expression data set because the method can identify associations among genes even when the genes are not co-expressed [10-14]. An association rule has the form \(\text{LHS} \left( \text{Left Hand Side} \right) \rightarrow \text{RHS} \left( \text{Right Hand Side} \right),\) where \(\text{LHS}\) and \(\text{RHS}\) are sets of items, and it represents that the \(\text{RHS}\) set being likely to occur whenever the \(\text{LHS}\) set occurs. In case of analyzing gene expression data, the items in an association rules are represented as genes which are highly expressed or highly repressed. An example of an association rule from gene expression data might be \([\text{gene A}], \text{gene B}] \rightarrow [\text{gene C}],\) which represents that when \(\text{gene A}\) is measured as highly expressed and \(\text{gene B}\) is highly repressed then it is also likely to observe and \(\text{gene C}\) is highly expressed. From the result of the ARM method, it is possible to discover interactions between correlated expressions of genes in microarray experiments. Despite of the usefulness of ARM [12], the time dependency between associated genes cannot be extracted by using the conventional ARM method even though the temporal information is indispensable to discover regulation mechanisms.
Previous studies which identify time-dependent regulatory relations among genes can be grouped into two general categories. The first approach constructs cellular dynamic models to observe the response of cells by using dynamic Bayesian network (DBN) [15-18] and ordinary differential equation (ODE). However, these approaches have fundamental problems: They need a huge amount of computational time to infer the temporal dependency among genes and show relatively low accuracies analyzing in microarray gene expression data [16, 18]. These drawbacks are mainly caused by the fact that the currently available time-series microarray data is not suited for such complex models of genetic regulation. Most of microarray gene expression data sets have relatively small number of experiments compared to the number of genes and they have relatively large regular time intervals between experiment time points. The second approach identifies pair-wise temporal dependency between genes by clustering with local patterns of gene expression [19], by measuring the Pearson correlation coefficient of two genes, by detecting the major changes in expression level [20], by scoring the expression patterns with several defined events [21], and by matching the expression patterns with shifted patterns [2, 3]. Although such methods can identify pair-wise temporal relations, it cannot identify combinatorial temporal relations which are regarded an important characteristic of regulation [22, 23]. For example, the meaning of \([\text{gene } A, \text{gene } B] \rightarrow (7 \text{ min}) [\text{gene } C]\), and \([\text{gene } A] \rightarrow (7 \text{ min}) [\text{gene } C]\) AND \([\text{gene } B] \rightarrow (7 \text{ min}) [\text{gene } C]\) is completely different: In the case of \([\text{gene } A, \text{gene } B] \rightarrow (7 \text{ min}) [\text{gene } C]\), gene A and gene B play a role as combinatorial regulators in a single regulation. On the other hand, \([\text{gene } A] \rightarrow (7 \text{ min}) [\text{gene } C]\) AND \([\text{gene } B] \rightarrow (7 \text{ min}) [\text{gene } C]\), gene A and gene B are independent regulators.

Even though there are some previous studies related to extraction of association rules from time series data in other application domains [24, 25], they do not provide temporal dependent information among items within different time (e.g. time shifted, time delayed). To address the problem, we propose a new mining method for gene expression data sets which can extract temporal dependency among genes by applying temporal association rule mining (TARM) method. The temporal association rules represent various transcriptional time delays between associated genes. An example of a temporal association rule is \([\text{gene } A↑, \text{gene } B↓] \rightarrow (7 \text{ min}) [\text{gene } C↑]\), which represents that high expression level of gene A and significant repression of gene B followed by significant expression of gene C after 7 minutes. Hence, the temporal association rule can tell us the size of transcriptional time delay (7 minutes) between associated genes (gene A, gene B and gene C), activation and inhibition relationship (\(\text{gene } A↑ \rightarrow \text{gene } C↑\)), and sets of co-regulators (\(\text{gene } A↑ \text{gene } B↓\)).

The overall process of the proposed method is depicted in Fig. 1. The proposed method consists of two main phases. First, temporal association rule mining phase. With an obtained fitted parameter set, the steps of temporal association mining method is applied to time-series gene expression data: (i) converting gene expression values into discrete values, (ii) generating temporal transaction sets with various sizes of transcriptional time delay \(\Delta\), (iii) generating temporal frequent item sets, (iv) and finally, extracting temporal association rules. The proposed method is tested with public microarray experiments of Saccharomyces cerevisiae cell cycle alpha factor arrest synchronization data set.

Second, parameters fitting phase. In this phase, external known regulation information (KEGG cell cycle regulation information) is used to choose the best parameter set from all possible combinations of parameter sets. Three parameters are selected for the proposed temporal association rule mining (TARM) method. Among every possible combination of three parameter values, the best parameter set that has the highest overlap degree with previously known biological regulation relationships is selected as the fitted parameter set.

![Figure 1. Method overview. (a) The overall phase of proposed method. (b) Parameter fitting phase.](image)

### 2. METHODS

#### 2.1 Conventional association rule mining (Apriori Algorithm)

| No. | Item purchased          |
|-----|-------------------------|
| 1   | Bread, Butter, Cereal, Juice, Milk |
| 2   | Cereal, Juice, Milk     |
| 3   | Bagels, Butter, Cereal, Juice, Milk |
| 4   | Bread, Cereal, Jelly, Juice, Milk |
| 5   | Bagels, Jelly, Juice, Milk |
| 6   | Jelly, Juice, Milk     |

To explain the basic concepts of association rule mining, we use the definitions and the examples of supermarket data shown in [26]. Consider a small store that sells the following set of items: [Bagels, Bread, Butter, Cereal, Juice, Milk]. List of items bought by six hypothetical customers are shown in Table 1. This table will be used to illustrate the concepts presented in this section.
Definition 1:
(1) An association rule is a pair of disjoint item sets. If \( \text{LHS} \) (Left Hand Side) and \( \text{RHS} \) (Right Hand Side) denote the two disjoint item sets, the association rule is written as \( \text{LHS} \rightarrow \text{RHS} \).
(2) The support of the association rule \( \text{LHS} \rightarrow \text{RHS} \) with respect to a transaction set \( T \) is the support of the item set \( \text{LHS} \cup \text{RHS} \) with respect to \( T \).
(3) The confidence of the rule \( \text{LHS} \rightarrow \text{RHS} \) with respect to a transaction set \( T \) is the ratio support \( (\text{LHS} \cup \text{RHS}) / \text{support(\text{LHS})} \).

Example: Consider the item sets \( A_1 = \{ \text{Juice, Milk} \} \) and \( A_2 = \{ \text{Cereal} \} \). Since \( A_1 \) and \( A_2 \) are disjoint, \( A_1 \rightarrow A_2 \) (or equivalently, \( \{ \text{Juice, Milk} \} \rightarrow \{ \text{Cereal} \} \)) is an association rule. Let \( R_1 \) denote this association rule. The support of \( R_1 \) is the support of the item set \( \{ \text{Juice, Milk, Cereal} \} \). From Table 1, it can be seen that this support value is 4. Also from Table 1, the support of the item set \( \{ \text{Juice, Milk} \} \) is 6. Therefore, the confidence of Rule \( R_1 \) is 4/6 or 66.67%.

2.2 Temporal association rule mining (TARM)

In this work, we propose a temporal association rule mining (TARM) method which is based on Apriori algorithm. Following two sub-sections will explain the detailed methodology of temporal association rule mining phase (Fig 1(b)), and parameter fitting phase (Fig. 1(a)).

To explain the concept of the proposed TARM method, we first define new terminologies.

Definition 2:
(1) A temporal item is an item which has a time stamp.
(2) A temporal item set \( I \) is a non-empty set of temporal items.
(3) Given a temporal item set \( I \), a set \( T \) of transactions on \( I \), and a positive integer \( a \), \( I \) is a temporal frequent item set with respect to \( T \) and \( a \) if support \( T(I) >= a \). (\( a \) is the support threshold.)
(4) A temporal association rule is a pair of disjoint temporal item sets. If \( \text{LHS} \) and \( \text{RHS} \) denote the left and right temporal item sets respectively, then the time stamp of each temporal item in \( \text{LHS} \) is ahead of those of all temporal items in \( \text{RHS} \). A temporal association rule is written as \( \text{LHS} \rightarrow (\Delta) \text{RHS} \), where \( \Delta \) is the interval of different two time stamps.

Fig. 2 shows an illustration of temporal association rule mining process. First, continuous gene expression values are converted into discrete values (up, down, and none) (Fig. 2(a)). Second, to find temporally associated genes, we first assume that all related genes may have various sizes of transcriptional time delay. Therefore, our method searches associated genes in all possible sets of different time point experiments where the time interval is from 0 to \( n \) (Fig. 2(b)). In this illustration, \( \Delta = 2 \). For example, Temporal transaction set \( t_0 + t_2 = \{ g_{1L}\uparrow, g_{2L}\downarrow, g_{1R}\downarrow, g_{2R}\uparrow, g_{3R}\downarrow \} \) consists of up or down regulated genes at time stamps \( t_0 \) and \( t_2 \) with the size of transcriptional time delay \( \Delta = 2 \). Note that, for \( g_1 \), it is up regulated in both cases of \( t_0 \) and \( t_2 \), but we marked them as two different genes like \( g_{1L} \) (in Left hand side) and \( g_{1R} \) (in Right hand side). Third, Fig. 2(c) indicates the extracted temporal frequent item sets with support threshold 50%. And finally, two temporal association rules are discovered with confidence threshold 50% as shown in Fig. 2(d).

Figure 2. An illustration of temporal association rule mining process. An illustration of temporal association rule mining process with transcriptional time delay \( \Delta = 2 \), support cutoff = 50%, confidence cutoff = 50%.

2.3 Parameter extraction

This section shows the phase for obtaining three different parameters which are necessary when mining temporal association rules: (1) a cutoff value for binning transcriptional expression values, (2) a support value for mining temporal frequent item sets, and (3) a confidence value for extracting temporal association rules. Since the performance of the proposed method is dependent on the parameter set, the parameter set should be chosen very carefully. If the ground truths of cell cycle regulation are known, the regulation information can be used to fit the parameters. However, absence of such kinds of information, alternative information source is used. In this study, we utilize KEGG cell cycle regulation path as known information set to find the best parameter set which can extract the most number of accurate temporal association rules. The KEGG cell cycle regulation path is a collection of manually drawn pathway maps representing the regulation knowledge on the molecular interaction, and the pathway contains interaction information which are relevant to cell cycle of yeast [27, 28]. (The URL of yeast cell cycle pathway map: http://www.genome.jp/kegg/pathway/sce/sec04111.html)

The KEGG regulation information is used for a measure of correctness of the extracted candidate rules with various combinations of parameters. If an extracted temporal association
rule is matched with KEGG regulation information, then we regard
the rule as a correctly extracted rule. Namely, the validation score is
calculated by the following equation:

\[
\text{precision} = \frac{\text{(# of matched rules)}}{\text{(# of extracted rules)}}
\] (1)

To select a fitted parameter set among the various combinations,
we select a parameter set which shows the highest validation
score.

3. Results and Discussion

3.1 Data sets

To check the performance of the proposed method, we used
*S. cerevisiae* cell cycle alpha factor arrest synchronization
microarray data set [29]. This time-series microarray data set has
18 time points with relatively small regular time intervals (7
minutes) between every sampling time point.

3.2 Results

In the parameter fitting phase, combination sets of parameters are
generated within binning cutoff values from 0.2 to 1.4, support
cutoff values from 2 to 6 transaction, and confidence cutoff values
from 80 to 100%. With these parameter sets, TARM method is
applied on cell cycle expression data of 57 genes which are nodes
of KEGG yeast cell cycle regulation pathway. Extracted temporal
association rules with every parameter set are validated with
KEGG cell cycle regulation information. The precision scores of
parameter sets are summarized in Table 2. To determine the best
parameter set, extracted rules with several sets of parameters
which show relatively high precision scores are examined
(precision scores with 0.25, 0.28, and 0.38). The temporal
association rules extracted with three selected parameter sets are
listed in Fig. 3. Finally, [threshold = ±0.8, support cutoff =3,
confidence cutoff = 90%] set is selected as the fitted parameter set
which shows the highest precision score (0.38). Although the
precision score of the fitted parameter set seems not significant,
the score is satisfactory in the case of microarray analysis.
Because it is reported that when inferring linkages of regulatory
proteins in KEGG pathway only from microarray gene expression
data set, the accuracy of inferred results were not high owing to
the property of microarray itself [30]. Furthermore, we also
applied Bayesian network inference (using deal package
implemented in R) on cell cycle expression data of 57 genes. The
precision score of Bayesian network inference is 0.16.
Comparing the precision scores of TARM (0.38) and Bayesian
network (0.16), TARM method identified more number of
accurate cell cycle regulation relationships among genes.

Using the selected parameter set, we applied TARM method to 799
genes which are pre-identified as cell cycle relevant genes in [29]
and extracted numbers of temporal association rules with various
sizes of transcriptional time delay. To test the significance of the
temporal association rules, TARM is also applied to random
shuffled cell cycle expression data of 799 genes. Fig. 5 is the
comparison result of both the real cell cycle data set and the shuffled
cell cycle data set. As the Fig. shows, the extracted numbers of rules
from real cell cycle data set and random data set are comparably
different. That means TARM method is able to catch significant
temporal regulation relationships among genes.

### Table 2: A summary of precision scores of 70 different parameter sets.

| Confidence | 90%, 100% | 80% |
|------------|-----------|-----|
| Support    | 2 3 4 5 6 | 2 3 4 5 6 |
| ± 0.2      | 0.05 0.05 0.06 0.08 0.10 | 0.05 0.06 0.06 0.07 0.08 |
| ± 0.4      | 0.04 0.06 0.08 0.10 0.15 | 0.05 0.07 0.08 0.08 0.11 |
| ± 0.6      | 0.05 0.14 0.16 0.17 0.0 | 0.06 0.14 0.15 0.15 0.13 |
| ± 0.8      | 0.17 0.38 0.25 - - | 0.15 0.25 0.23 0.0 - |
| ± 1.0      | 0.28 0.0 - - - | 0.17 0.28 0.0 - - |
| ± 1.2      | 0.18 0.0 - - - | 0.18 0.0 - - - |
| ± 1.4      | 0.0 - - - - | 0.0 - - - - |

24
Figure 3. Extracted temporal association rules with the selected three parameter sets. Best three parameter sets are selected to compare results of extracted rules on cell cycle expression data of 57 genes with association delay 0 ~ 28 minutes. Set A = [threshold = ±0.8, support cutoff = 3 transactions, confidence cutoff = 80%], set B = [threshold = ±0.8, support cutoff = 3 transactions, confidence cutoff = 90%], set C = [threshold = ±1.0, support cutoff = 3 transactions, confidence cutoff = 80%]. The intersection area of a Venn diagram stands for the commonly extracted rules with different parameter sets. Rules written in Italic font denote known regulation relations in KEGG Cell cycle pathway data.

From the extracted temporal association rules, rules with significant support (S ≥ 5) are chosen for further Gene Ontology (GO) term [31] analysis and represented in a directed graph structure (Fig. 6). By this analysis, interesting features are found. First, associated genes which play same role of biological phase with relatively short transcriptional time delay are identified. For example, HTB2, HTA2, HHT1, HHT1, HTB1, HHF2, and HHT2 those who share same annotation term (Organelle organization and biogenesis, DNA metabolic process) are complexly associated with one another within 0 ~ 7 minutes and these associated genes are known as having protein interactions with each other. HTA1 interacts with HTA2 [32], HTB1 [33], HTB2 [34, 35], HHF1[33], HHT1 [34-36]. HTA2 interacts with HHF1 [37], HHT1[32], HHT2 [32], HTA1 [32], HHF2 [32]. Second, some temporal dependencies between genes with specific biological processes are detected. Like POL30, YLR183C (RNA metabolic process, Transcription, Cell cycle) and HTA1, HTA2, HTB1, HHF2 (Organelle organization and biogenesis, DNA metabolic process) have temporal association with Δ = 14 minutes. PIR1, PIR3 (Cell wall organization and biogenesis) and HTB2 (Organelle organization and biogenesis, DNA metabolic process) are temporally associated with Δ = 21 minutes.

| A | B | C |
|---|---|---|
| Precision |
| (% of known relations) | 7/27 | 7/18 | 2/7 |
| (% of total rules) | 0.25 | 0.38 | 0.28 |

Figure 4. The result of Bayesian network inference. from cell cycle expression data of 57 genes. A red arrow represents known regulation relations in KEGG Cell cycle pathway data. The precision score of Bayesian network is 0.16.

4. Conclusions
We developed the TARM method that can extract temporal association rules in time-series gene expression data, and validated the proposed method with yeast cell cycle gene expression data set. A temporal association rule can describe how the expression of one gene might be associated with the expression of other genes with the related temporal dependency.

In the parameter fitting phase, the best parameter set (threshold = ±0.8, support cutoff = 3 transactions, confidence cutoff = 90%, 100%), which extracted the most number of correct associations in KEGG cell cycle pathway among 70 combinations of parameters, has been chosen for rule mining. Furthermore, comparing the precision scores of TARM (0.38) and Bayesian...
network (0.16), TARM method showed better accuracy. With the best parameter set, numbers of temporal association rules are extracted among pre-identified 799 cell cycle relevant genes. From the extracted temporal association rules, temporally associated genes which play same role of biological processes (Organelle organization and biogenesis, DNA metabolic process) with short transcriptional time delay, and some temporal dependencies between genes with specific biological processes are detected. The strong points of our method are the detection abilities of (1) various sizes of transcriptional time delay between associated genes, (2) activation and inhibition relationship, (3) sets of co-regulators for the target genes.

Figure. 5 The number of extracted temporal association rules from cell cycle data set and random data set. The graph shows the number of extracted temporal association rules in five transcriptional time delays (0, 7, 14, 21, 28 minutes) from time-series gene expression of 799 cell cycle relevant genes and random shuffled cell cycle data set [threshold = ±0.8, support cutoff = 3 transactions, confidence cutoff = 90, 100%]. Black bar indicates the number of extracted rules in real data set and gray bar stands for the average number of extracted rules of 100 times of random tests.

5. Acknowledgments
This work was supported by the Korean Systems Biology Program from the Ministry of Education, Science and Technology through the Korea Science and Engineering Foundation (No. M10309020000-03B5002-00000) and this work was also supported by the second stage of the Brain Korea 21 Project in 2008. We would like to thank CHUNG Moon Soul Center for BioInformation and BioElectronics for providing research facilities.

Figure. 6 Validation of the extracted temporal association rules. Extracted temporal association rules with high support (support cutoff = 5) are represented in network structure (upper). A solid pointed arrow edge indicates ‘up → up’ relation; a solid blunt arrow indicates ‘down → up’; a dashed pointed arrow indicates ‘down → down’; a dashed blunt arrow indicates ‘up → down’ relation. Grey colored nodes denote genes whose biological function is known. White colored nodes stand for genes whose biological function is not discovered yet. The numeric values on each edge stands for transcriptional time delay (Δ) between genes. Biological process annotation terms of genes represented in network are summarized in Table.

6. References
[1] Kim, D. W., et al. 2005. Detecting clusters of different geometrical shapes in microarray gene expression data. Bioinformatics. 21 (May 1) 1927-34
[2] Ji, L. and Tan, K. L. 2004. Mining gene expression data for positive and negative co-regulated gene clusters. Bioinformatics. 20 (Nov 1) 2711-8
[3] Ji, L. and Tan, K. L. 2005. Identifying time-lagged gene clusters using gene expression data. Bioinformatics. 21 (Feb 15) 509-16
[4] Tamada, Y., et al. 2003. Estimating gene networks from gene expression data by combining Bayesian network model with promoter element detection. Bioinformatics. 19 Suppl 2 (Oct) ii227-36
Liang, S., et al. 1998. Reveal, a general reverse engineering algorithm for inference of genetic network architectures. Pac Symp Biocomput. 18-29

Akutsu, T., et al. 2000. Algorithms for inferring qualitative models of biological networks. Pac Symp Biocomput. 293-304

Friedman, N. 1999. Learning bayesian network structure from massive datasets: the 'sparse candidate' algorithm. Proc. of Fifteenth Conference on Uncertainty in Artificial Intelligence.

Segal, E., et al. 2003. Module networks: identifying regulatory modules and their condition-specific regulators from gene expression data. Nat Genet. 34 (Jun) 166-76

Lee, P. H. and Lee, D. 2005. Modularized learning of genetic interaction networks from biological annotations and mRNA expression data. Bioinformatics. 21 (Jun 1) 2739-47

Creighton, C. and Hanash, S. 2003. Mining gene expression databases for association rules. Bioinformatics. 19 (Jan) 79-86

Georgii, E., et al. 2005. Analyzing microarray data using quantitative association rules. Bioinformatics. 21 Suppl 2 (Sep 1) i123-9

Carmona-Saez, P., et al. 2006. Integrated analysis of gene expression by Association Rules Discovery. BMC Bioinformatics. 7 54

Becquet, C., et al. 2002. Strong-association-rule mining for large-scale gene-expression data analysis: a case study on human SAGE data. Genome Biol. 3 RESEARCH0067

Morgan, X. C., et al. 2007. Predicting combinatorial binding of transcription factors to regulatory elements in the human genome by association rule mining. BMC Bioinformatics. 8 (Nov 15) 445

Perrin, B. E., et al. 2003. Gene networks inference using dynamic Bayesian networks. Bioinformatics. 19 Suppl 2 (Oct) ii138-48

Zou, M. and Conzen, S. D. 2005. A new dynamic Bayesian network (DBN) approach for identifying gene regulatory networks from time course microarray data. Bioinformatics. 21 (Jan 1) 71-9

Kikuchi, S., et al. 2003. Dynamic modeling of genetic networks using genetic algorithm and S-system. Bioinformatics. 19 (Mar 22) 643-50

Husmeier, D. 2003. Sensitivity and specificity of inferring genetic regulatory networks from microarray experiments with dynamic Bayesian networks. Bioinformatics. 19 (Nov 22) 2271-82

Qian, J., et al. 2001. Beyond synexpression relationships: local clustering of time-shifted and inverted gene expression profiles identifies new, biologically relevant interactions. J Mol Biol. 314 (Dec 14) 1053-66

Filkov, V., et al. 2002. Analysis techniques for microarray time-series data. J Comput Biol. 9 317-30

Kwon, A. T., et al. 2003. Inference of transcriptional regulation relationships from gene expression data. Bioinformatics. 19 (May 22) 905-12

Kato, M., et al. 2004. Identifying combinatorial regulation of transcription factors and binding motifs. Genome Biol. 5 R56

Wang, W., et al. 2005. Inference of combinatorial regulation in yeast transcriptional networks: a case study of sporulation. Proc Natl Acad Sci U S A. 102 (Feb 8) 1998-2003

Ning, H., et al. 2006. Temporal Association Rules in Mining Method.

Li, Y., et al. 2001. Discovering Calendar-based Temporal Association Rules.

Doddi, S., et al. 2001. Discovery of association rules in medical data. Med Inform Internet Med. 26 (Jan-Mar) 25-33

Kanehisa, M., et al. 2006. From genomics to chemical genomics: new developments in KEGG. Nucleic Acids Res. 34 (Jan 1) D354-7

Kanehisa, M., et al. 2002. The KEGG databases at GenomeNet. Nucleic Acids Res. 30 (Jan 1) 42-6

Spellman, P. T., et al. 1998. Comprehensive identification of cell cycle-regulated genes of the yeast Saccharomyces cerevisiae by microarray hybridization. Mol Biol Cell. 9 (Dec) 3273-97

Lee, I., et al. 2004. A probabilistic functional network of yeast genes. Science. 306 (Nov 26) 1555-8

Ashburner, M., et al. 2000. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. Nat Genet. 25 (May) 25-9

Krogan, N. J., et al. 2006. Global landscape of protein complexes in the yeast Saccharomyces cerevisiae. Nature. 440 (Mar 30) 637-43

Ho, Y., et al. 2002. Systematic identification of protein complexes in Saccharomyces cerevisiae by mass spectrometry. Nature. 415 (Jan 10) 180-3

Fukuma, M., et al. 1994. Purification of yeast histones competent for nucleosome assembly in vitro. Yeast. 10 (Mar) 319-31

Grant, P. A., et al. 1999. Expanded lysine acetylation specificity of Gcn5 in native complexes. J Biol Chem. 274 (Feb 26) 5895-900

Gelbart, M. E., et al. 2001. Interactions of Isw2 chromatin remodeling complex with nucleosomal arrays: analyses using recombinant yeast histones and immobilized templates. Mol Cell Biol. 21 (Mar) 2098-106

Gavin, A. C., et al. 2006. Proteome survey reveals modularity of the yeast cell machinery. Nature. 440 (Mar 30) 631-6