# Module of Cellular Networks in Saccharomyces cerevisiae

Yueying Yang
Northeast Forestry University Postdoctoral Programme
Heilongjiang Acamedy of Agricultural Sciences
Postdoctoral Programme
Department of Information and Computation Science
Northeast Agricultural University
Harbin, China
yangyueying2008@yahoo.com.cn

Di Liu
Heilongjiang Acamedy of Agricultural Sciences
Harbin, China

Jun Meng
Department of Information and Computation Science
Northeast Agricultural University
Harbin, China

***Abstract*—The focus of the network of research is to determine their community or module, it helps the functional organization and evolution of the network. Modular can be seen as a function of a dynamic cell system executing complex functions in a living cell. How to identify the precious knowledge resources to build a more reliable module is still one of the most important and difficult problems in bioinformatics. We put forward a state space model combining the topological method to describe the time and space module in the cell cycle of the process. Not only our module function sets of genes related to identify a condition to activate or suppress in the cell cycle process in S.cerevisiae, but also have many different solutions, which have evolved into different molecular components will be the assembly at the right time in the cell cycle. The resulting module mapping analysis showed several assumptions connection biological process to a particular cell cycle conditions.**

***Keywords- cell cycle; dynamics; modules; state-space model***

## I. INTRODUCTION

System application to be automatic high-throughput molecular biology technology provides an unprecedented opportunity to understand countless functional genomics[1-3]. With more and more high throughput molecular biology experimental data of gene expression and regulation[4,5], there are more and more need to identify the data related to biology.

Biological systems can be described as dynamic and distributed programs[6] of a large number of different material under the unified arrangement of the interaction and adjust the cell behavior and function. Molecular machinery of cell cycle control has been described as a whole dynamic system, including signaling pathways[7] and interconnection of transcription programs[8]. In the classic case of the S.cerevisiae, several beautiful measurements have been taken[9,10].

Systematic models[11-13] that link the behaviour of a system to the interactions between its components will play an increasingly important role in post-genomic biology. Since the 1990s, modeling has become a new kind of tools to deal with the fast growth the potential genetic information architecture and interactive complex circuit of the overwhelming signal networks[14-20]. The quantitative data demands quantitative models on the molecular level, from regulation of gene expression and gene function to cellular mechanisms.

The different areas of the genome data to continue to grow and mature, this is an opportunity, the development of bioinformatics, its purpose is to rebuilding the biological system through the modeling and components. In short, the field of biological systems is still not fully explored, and especially the area of multi-time delay biological systems needs directed research efforts. The state-space model[21-24] can be used as a framework for calculating the biological system. In this paper, we put forward a kind of combination state space model of topological method in the process of cell cycle, in which the periodically expressed genes are viewed as the observation variables and the periodically expressed genes' expression dynamics in cell cycle are governed by a group of the internal variables. We focus on trying to explore the cell cycle related biological characteristics, it provides a description of the network as a continuous time dynamic system, it can be used to infer that the main mechanisms of the cell cycle network.

## II. METHODS

Now most of the cell cycle is a sign in a specific time, provide only a snapshot of the gene expression, it is difficult to understand and the causes and effects of the differentiation from gene signature. Time series or condition-specific data need to further understand cell kinetics. Comprehensive time and space-dependent method should be developed to deduce causality and define the directional pathway activation in the cell cycle's signature.

The combination of the state space model of topological method we adopted the integration of the different characteristics of space and time of cell cycle process and clarify the dependence of hidden variables. We can replace the hard task of constructing gene regulation network based on time as building blocks of the gene expression regularly topology. More accurately, this kind of method is based on the idea, the instantaneous character gene interest can be from module, established including moleculars endowed with their relevant functions.


978-1-4673-4398-5/12/$31.00 ©2012 IEEE

In this paper, we put the periodic expressed genes as the expression of the observation variable depending on whether the current value of the internal state variables and any external investment. In fact, for a lot of species of cell cycle system, the change in internal state totally depends on the current state of the existence of the internal and external input. Although each internal variable has a different combination of the observed variables, and share the same internal characteristics can be grouped together.

The general form of the state-space model can be represented by the general form:

$$\begin{cases} z(t+\Delta t) = A \cdot z(t) + n_1(t) \\ x(t) = C \cdot z(t) + n_2(t), \end{cases} \quad (1)$$

We constructed an observation variables matrix $x(t)$ whose rows are genes and columns are time points, in which the it-th element is the expression level of gene i at time t. The matrix $z(t)$ consists of the internal state variables of the system, and the jt-th element is used to mimic the expression value of internal element j at time t. Once the matrix $z(t)$ characterizing the internal state variables is determined, it is a simple matter to deduce a cluster of observation variables in every internal state variable.

A certain cluster of observation variables in each internal state variable $z_i(t)$, i=1,…,p (p is the number of the internal state variables) is represent by $O_j^i(t) = [o_1^i(t), \cdots, o_k^i(t)]^T$, $o_j^i(t) \in \{x_1^i(t) \cdots x_n^i(t)\}$, $j = 1, \cdots, k$ and $k < n$. When $j \neq 1$ then $o_j^i(t) \neq o_l^i(t)$, k is the number of observation variables in a certain internal state variable.

We proposed the following linear temporal topology equations to model the regulation of a certain cluster of observation variables:

$$X(t+1) = \sum_{\tau=0}^{\tau_{\max}} W_\tau \cdot O_j^i(t-\tau) \quad (2)$$

where the vector $X(t+1)=[x_1(t+1), \cdots, x_n(t+1)]^T$ contains the expression levels of overall periodically expressed gene at time point t+1. The regulatory relationships and degrees among genes are thus captured by the $W_\tau = ([\omega_{uv}]_{n \times k})_\tau, (\tau = 0, \cdots, \tau_{\max})$ in (2) where the matrix element $\omega_{uv}$ indicates the power of gene v at time $t-\tau$ influences the expression level of gene u at time $t+1$. A positive matrix element leads to the u-th gene being positively reinforced by the v-th gene expression level at a pretty time, vice versa.

It is well known that there should be a maximum time delay interval in a cellular network since the time of a cell cycle is limited. In this limit, an activator/inhibitor gene can regulate another gene either instantly, or in the next time slice or up to τmax time slices (The parameter τmax indicates the maximum delay within which a gene can regulate another gene). So the gene expression signal of each gene is correlated with other genes expression signals with a maximum lag of τmax time slices. Then we used the least squares method combined with the control of time-delay to resolve the problem of parameter estimation, i.e. the activation and repression are characterized by these parameters $W_\tau$.

The analysis results show that, through the quantitative measurement observed variables and the overall regular delay expressed genes, we can realize the two important goal: 1) that the state space model, will get a quantitative representation of the internal state variables and every internal state variables corresponding observation variable, this is our analysis is an important information. As discussed before, quantitative study gene module is very meaningful. 2) through the topological method, we can generate mechanism of temporal topology module. So, the topological module form can be efficient discusses the cell cycle system in regulating gene expression mechanism.

We think that this model has a guiding significance not only, and can forecast and provide a chance to learn the cell cycle. In short, the state space model combining topological method (SSTT) will be integrating different space time dynamic signal from the different components provide new perspective of the cell cycle system. SSTT model will provide the opportunity to have cell network establishment mechanism. Cell network establishment mechanism once breakthrough, in the network of key control factors will identify possible.

## III. RESULTS

The composition of the cell cycle network features the function of topological model, these functions related components usually form module. It for precise understanding of the dynamic characteristics of the cell activity module has played a significant role. In our work, we used the S. cerevisiae dataset of timecourses microarray. We identified these modules with the SSTT approach. We have determined the parameters of an approach (as described in Methods) for the S. cerevisiae dataset of periodically expressed genes.

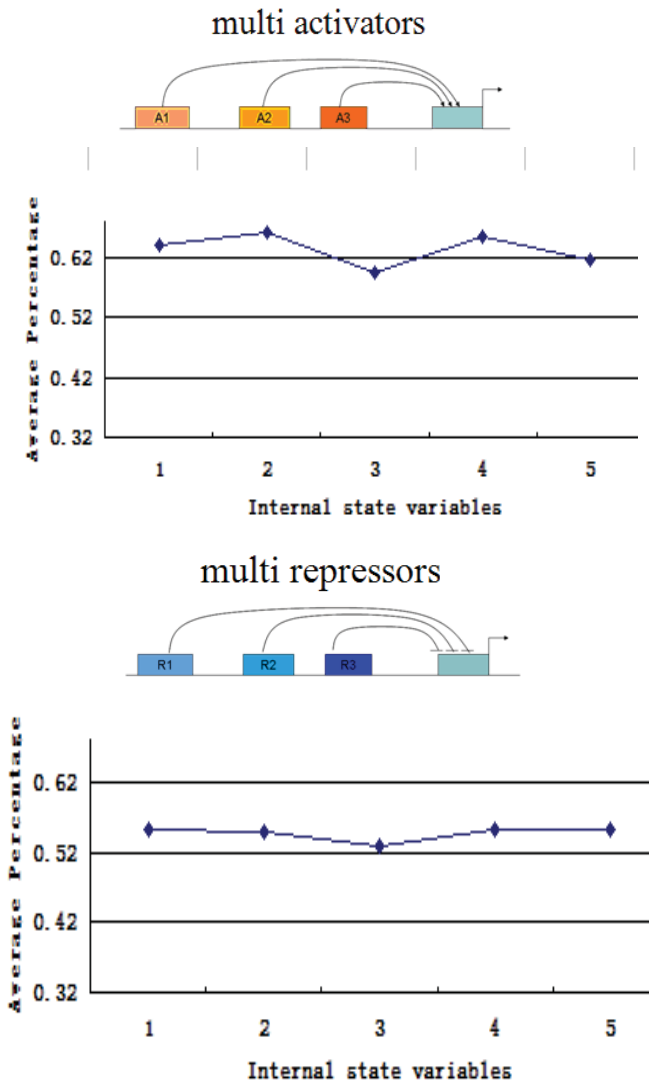
### *A. Temporal Module*

In the study of dynamic gene expression in describing the cell cycle, we describe the simple cell cycle regulation module formed when regulators and target genes in different internal state variables of the system. Find the basic cell cycle temporal modules, make all kinds of regulation relationship discussion in a uniform way, simple unit, which contains several components listed. These temporal modules on display as shown in figure 1.

The figure depicted several different dynamic temporal regulatory modules. As shown in figure 1, the inverted ‘T’ symbol indicates inhibition, arrows represented activation. The first listed in figure 1 includes several time control module. Other columns depict the average percentage of the activation and suppress the adjustment function target genes in the corresponding time management module of different internal state variables. These modules can maintain certain types of topological properties, such as the environment of the disturbance of robustness and evolution of the conservative, highlight the different dynamic performance requirements of every condition, and show some function of self control.


978-1-4673-4398-5/12/$31.00 ©2012 IEEE

These results seemed all activators should show strong activation rather than the average effect. We conclude that genes including in this module have the bistability and switch-like behavior to keep balance of the system. The activation effect should be dominant in the "simple repressor" module, but the results are on the contrary. The first and the fifth internal state variables of the system, the repression effect is almost stronger than activators compared with the common cases. This is a habit of cell cycle to perform adaptation, desensitization and preservation of homeostasis. It is the same with the "dual repressors" module. This is the reflection of organism self-monitoring control. Specifically, the "multi activators" and "multi repressors" modules are suitable for keeping long-lasting signals to drive multi-staged. The "simple repressor" and "dual repressors" modules are suitable for initiating a quick and coordinated response to external stimuli. It is a biologically pleasing result in the precise spatial and temporal control of gene expression. Temporal regulatory modules reveal the dynamic properties that contribute to cellular functions.



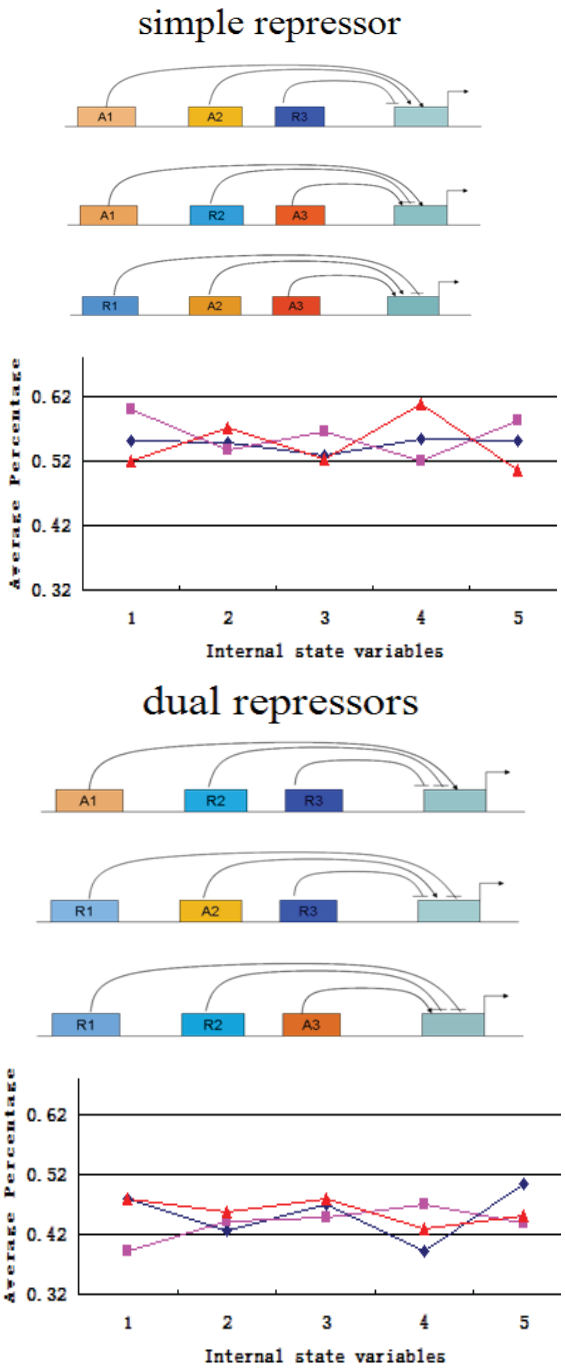




Figure 1. Several different temporal regulatory modules

## B. *Evaluation of Temporal Modules*

Using the approach as described in Methods section, we evaluate the performance of the proposed temporal regulatory modules. Each of them bears a distinct regulatory function in cellular networks. In our work, we showed time plays an important role in module regulatory behavior. Thus, analyzing how time affect on modules may shed some insights into understanding temporal regulatory modules' principles in signaling networks. We identified eight types of temporal regulatory modules (figure 1). We mapped target genes onto eight temporal regulatory modules, and discussed each type of


978-1-4673-4398-5/12/$31.00 ©2012 IEEE

module on different time points based on internal state variables above mentioned. For example, the network modules may have three activations to target genes at three time points, or may have two to target genes, or just one or none. For each module, we further calculated the ratio of positive actions to the negative actions (+/-) in each module for five internal state variables at three time point and compared them with the average (+/-) in all the modules, which is shown as a horizontal line in Figure 2.

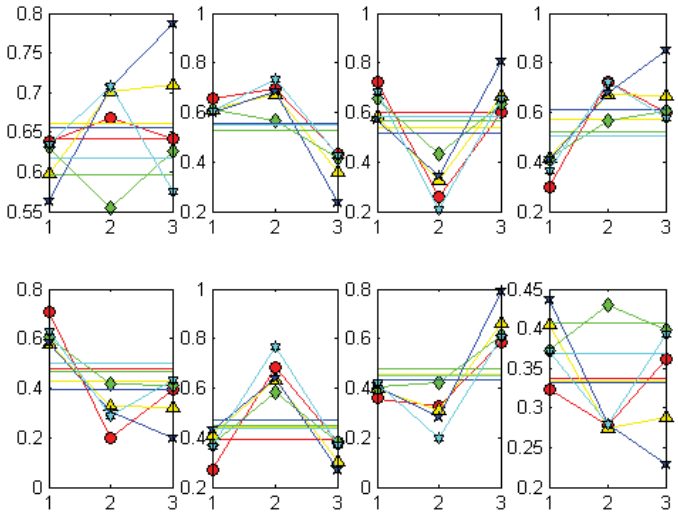


Figure 2. The figure shows the ratio of the activation to repression regulative power of target genes which belong to eight temporal regulatory modules under five internal state variables in three time points. (a) 1th internal state variable (red), (b) 2th internal state variable (yellow), (c) 3th internal state variable (green), (d) 4th internal state variable (blue), (e) 5th internal state variable (indigo).

The figure shows the ratio of the activation to repression regulative effect of target genes in eight temporal regulatory modules under five internal state variables in three time points. For most modules except the first and the eighth, the performance at the three time point is basically consistent. Remarkably, we found that the third internal state variable of the system (which were marked green diamond) in the first and eighth modules is significantly different with the rest internal state variables in the figure. In the first and eighth modules, the third internal state variable show clearly lower at the second time point than at the other time points. Genes including in this module seems like to have the bistability and switch-like behavior to keep balance of the system. Cln3 was identified as an important gene of the third internal state variable, consequently it belongs to the first module. CLN3 is expressed at the late M phase, Cln3 was identified as a positive regulator of target genes SWI4 and CLN2. Similar to the immense amount of evidence showing that Cln3 is the most prominent activator of SWI4, our result displayed that Cln3 significantly regulated the expression of several G1 phase genes.

It should be noted that, in cell cycle module identification, the influences in the module indicate different characteristics to five internal variables of system (see Materials and Methods). Along with the distribution of genes which were assigned positive action, we also plotted 1000 distributions in a certain internal variables. In this randomization we preserved the number of genes which were assigned positive action per element in a certain internal variables, yet assigned actions at random to each gene. The distributions obtained after the randomization differed markedly from the original distribution, both in terms of width and shape.

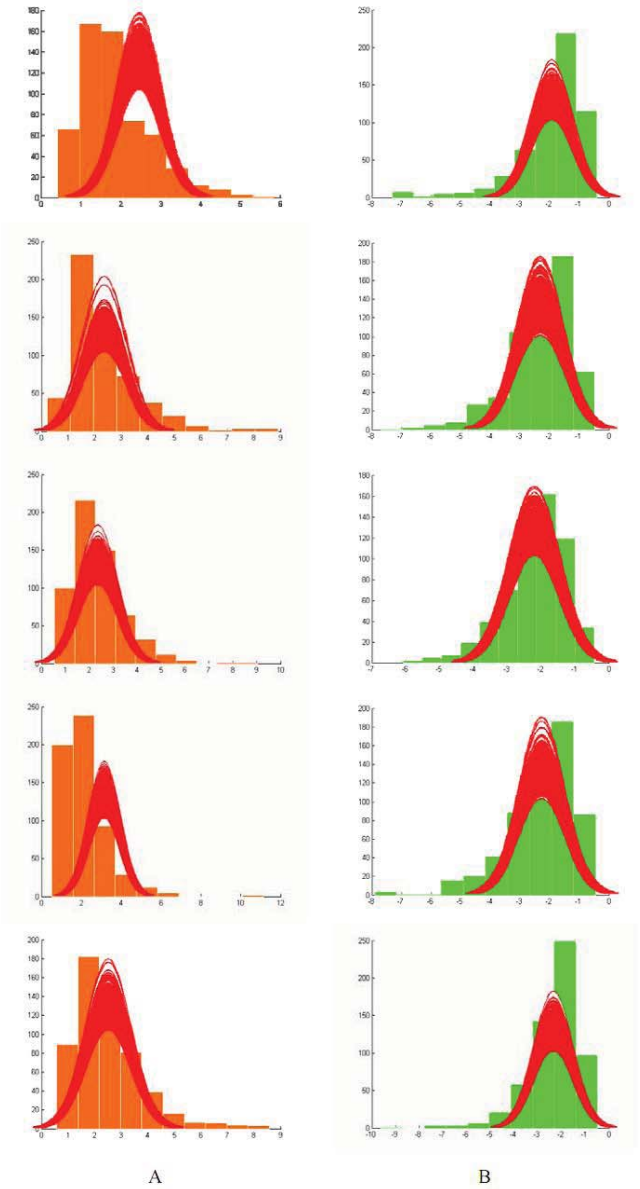



Figure 3. (A) Distribution of the number of genes which were assigned positive action. (B) Distribution of the number of genes which were assigned negative action. The red line represents the distribution obtained after randomizing each of the columns in a certain internal variables. In this randomization we preserved the number of genes which were assigned positive action per element in a certain internal variables, yet assigned actions at random to each gene.

The distribution of genes which were affected by different internal variables of system in the original distribution (Figure 3A) has a long tail in contrast to the distributions in the randomized matrices that looked Gaussian. Along with it is shown a set of distributions obtained by random, namely by


978-1-4673-4398-5/12/$31.00 ©2012 IEEE

randomly assigning effects to each gene, preserving the real number of effects to each target gene. Here, too, the original data significantly higher(or lower) than the number that would be obtained by merely preserving the statistics of number of effects to each target gene.

At last, it should be noted that, in cell cycle module identification, the influences in the module indicate physical attachment but not necessarily activation or repression. The fluctuation of mRNA expression levels may not effectively reflect regulator activities. In addition, post-transcriptional regulation and post-translational modification control protein activities as well. These effects, once they can be quantitatively determined, should be incorporated into the approach. However, actual realization of such modules remains to be explored.

## IV. DISCUSSION

A major challenge of the post-genomic research is to understand how cellular phenomena arise from the connectivity of genes and proteins. In this paper, we described a SSTT model that provided a compact description useful for cellular networks studies without the need to invoke the biochemical details of every component. We demonstrate our algorithm on the S.cerevisiae cell cycle datasets, showing that our framework does learn to predict topology-based time delay modules. The module suggested a general mechanism where only some genes are regulated during cell cycle process and the module consist of these genes controls the timing of complex assembly.

Specially, it can be used to systematically generate hypotheses on the likely mechanisms over all phases of the cell cycle. The old models are much simpler than ours, and fails to capture important aspects such as taking consider of time delay combinatorial effects on the expression data. Our algorithm could also find highly significant modules in different organisms that it asserts are conservational, providing a strong biological basis for this claim.

## ACKNOWLEDGMENT

This work was supported by the Heilongjiang Academy of Agricultural Sciences Project (No. LBH-Z10039), the Doctoral Research Foundation of Northeast Agricultural University.

## REFERENCES

[1] Daigo, Y. and Nakamura, Y. From cancer genomics to thoracic oncology: discovery of new biomarkers and therapeutic targets for lung and esophageal carcinoma, Gen Thorac Cardiovasc Surg, 56, pp.43-53, 2008.

[2] Ragoussis, J. and Elvidge, G. Affymetrix GeneChip system: moving from research to the clinic, Expert Rev Mol Diagn, 6, pp.145-152, 2006.

[3] Ryu, D.D. and Nam, D.H. Recent progress in biomolecular engineering, Biotechnol Prog, 16, pp. 2-16, 2000.

[4] Bailey,W.J.and Ulrich,R. Molecular profiling approaches for identifying novel biomarkers, Expert Opin Drug Saf, 3, pp.137-151, 2004.

[5] Fischer, H.P. Towards quantitative biology: integration of biological information to elucidate disease pathways and to guide drug discovery, Biotechnol Annu Rev, 11, pp.1-68, 2005.

[6] Xiong, M., Arnett, F.C., Guo, X., Xiong, H. and Zhou, X. Differential dynamic properties of scleroderma fibroblasts in response to perturbation of environmental stimuli, PLoS ONE, 3, pp.e1693, 2008.

[7] Li, X., Wang, H., Touma, E., Rousseau, E., Quigg, R.J. and Ryaby, J.T. Genetic network and pathway analysis of differentially expressed proteins during critical cellular events in fracture repair, J Cell Biochem, 100, pp.527-543, 2007.

[8] Luscombe, N.M., Babu, M.M., Yu, H., Snyder, M., Teichmann, S.A. and Gerstein, M. Genomic analysis of regulatory network dynamics reveals large topological changes, Nature, 431, pp. 308-312, 2004.

[9] Rubinstein, A., Gurevich, V., Kasulin-Boneh, Z., Pnueli, L., Kassir, Y. and Pinter, R.Y. Faithful modeling of transient expression and its application to elucidating negative feedback regulation, Proc Natl Acad Sci U S A, 104, pp.6241-6246, 2007.

[10] Vrabioiu, A.M. and Mitchison, T.J. Structural insights into yeast septin organization from polarized fluorescence microscopy, Nature, 443, pp. 466-469, 2006.

[11] Ishii, N., Robert, M., Nakayama, Y., Kanai, A. and Tomita, M. Toward large-scale modeling of the microbial cell for computer simulation, J Biotechnol, 113, pp.281-294, 2004.

[12] Li, H. and Zhan, M. Systematic intervention of transcription for identifying network response to disease and cellular phenotypes, Bioinformatics, 22, pp.96-102, 2006.

[13] Saucerman, J.J. and McCulloch, A.D. Mechanistic systems models of cell signaling networks: a case study of myocyte adrenergic regulation, Prog Biophys Mol Biol, 85, pp.261-278, 2004.

[14] Bhalla, U.S. and Iyengar, R. Functional modules in biological signalling networks, Novartis Found Symp, 239, pp.45-51, 2001.

[15] Bleakley, K., Biau, G. and Vert, J.P. Supervised reconstruction of biological networks with local models, Bioinformatics, 23, pp.57-65, 2007.

[16] Faeder, J.R., Blinov, M.L., Goldstein, B. and Hlavacek, W.S. Combinatorial complexity and dynamical restriction of network flows in signal transduction, Syst Biol (Stevenage), 2, pp.5-15, 2005.

[17] Galperin, M.Y. Bacterial signal transduction network in a genomic perspective, Environ Microbiol, 6, pp.552-567, 2004.

[18] James, A., Swann, K. and Recce, M. A dynamical model of the distributed interaction of intracellular signals, Int J Neural Syst, 7, pp.333-341, 1996.

[19] Kollmann, M., Lovdok, L., Bartholome, K., Timmer, J. and Sourjik, V. Design principles of a bacterial signalling network, Nature, 438, pp.504-507, 2005.

[20] Schamel, W.W. and Dick, T.P. Signal transduction: specificity of growth factors explained by parallel distributed processing, Med Hypotheses, 47, pp.249-255, 1996.

[21] Haavisto, O., Hyotyniemi, H. and Roos, C. State space modeling of yeast gene expression dynamics, J Bioinform Comput Biol, 5, pp.31-46, 2007.

[22] Wolkenhauer, O. Mathematical modelling in the post-genome era: understanding genome expression and regulation--a system theoretic approach, Biosystems, 65, pp.1-18, 2002.

[23] Wu, F.X. Gene Regulatory Network modelling: a state-space approach, Int J Data Min Bioinform, 2, pp.1-14, 2008.

[24] Wu, F.X., Zhang, W.J. and Kusalik, A.J. Modeling gene expression from microarray expression data with state-space equations, Pac Symp Biocomput, pp.581-592, 2004.


978-1-4673-4398-5/12/$31.00 ©2012 IEEE