# 10

# *Saccharomyces pombe* and *Saccharomyces cerevisiae* Gene Regulatory Network Inference Using the Fuzzy Logic Network

Yingjun Cao[1], Paul P. Wang[2], and Alade Tokuta[1]

[1] Department of Mathematics and Computer Science, North Carolina Central University,
1801 Fayetteville Street, Durham, NC 27707.
{ycao,atokuta}@nccu.edu

[2] Department of Electrical and Computer Engineering, P.O. Box 90291, Duke University,
Durham, NC 27708.
ppw@ee.duke.edu

**Summary.** In this chapter, a novel gene regulatory network inference algorithm based on the fuzzy logic network theory is proposed and tested. The key motivation for this algorithm is that genes with regulatory relationships may be modeled via fuzzy logic, and the strength of regulations may be represented as the length of accumulated distance during a period of time intervals. One unique feature of this algorithm is that it makes very limited a priori assumptions concerning the modeling. Hence the algorithm is categorized as a data-driven algorithm. With the theoretical guidelines to quantify the upper limits of parameters, the algorithm is implemented to infer gene regulatory networks for *Saccharomyces cerevisiae* and *Saccharomyces pombe*. The computation results not only prove the validity of the data-driven algorithm, but also offer a possible explanation concerning the difference of network stabilities between the budding yeast and the fission yeast.

## 10.1 Introduction

One of the most challenging problems in bioinformatics is to determine how genes inter-regulate in a systematic manner which results in various translated protein products and phenotypes. To find the causal pathways that control the complex biological functions, previous work have modeled gene regulatory mechanisms as a network topologically [1]. The importance of networking models is that normal regulatory pathways are composed of regulations resulting from many genes, RNAs, and transcription factors (TFs). The complicated inter-connections among these controlling chemical complexes are the driving forces in maintaining normal organism functions. A precise structural presentation of components should illustrate the key properties of the system.

Based on the network representation of gene regulations, a number of inference models have been proposed. They include Bayesian networks [2], hybrid Petri

Y. Cao et al.: *Saccharomyces pombe and Saccharomyces cerevisiae Gene Regulatory Network Inference Using the Fuzzy Logic Network*, Studies in Computational Intelligence (SCI) **94**, 237–256 (2008)
www.springerlink.com

net [3], growth network [4], genetic circuits [5], inductive logic programming [6], and hybrid networks [7]. These models have focused on different aspects of gene regulatory behaviors, and each model has contributed good inference results in certain aspects. The ongoing research on these models is focused on the challenges of data integration, non-linear data processing, noise tolerance, synchronization, and model over fitting [8].

Research on genome-wide gene regulations has used dynamic microarray data which quantify the genomic expression levels at each sample time. Given a series of microarray data, researchers have attempted to find the spatial and temporal modes of regulations regarding different conditions or different stages of cell cycles on different species [8]. But because of the hybridization process and the synchronization issues of time-series microarray, the data, very often contain missing, noisy, or unsynchronized data subsets. Thus data normalization and pre-processing techniques have become necessary to reduce the noise. Other techniques like SAGE [9], TFs mappings [10], and antibiotic arrays [11] have been designed to discover the regulatory mechanisms.

In this chapter, a novel network model, the fuzzy logic network (FLN), is proposed and thoroughly examined. The feasibility of applying this model to inferring gene regulatory networks is investigated. The FLN is a generalization of the Boolean network, but it is capable of overcoming the unrealistic constraints of Boolean values (ON/OFF symbolically). With distinctive properties in processing real life incomplete data and uncertainties, researchers have applied fuzzy logic to gene expression analysis, and by the use of specific scoring matric, *Saccharomyces cerevisiae* gene regulatory networks with biological verifications were inferred [12–14]. This chapter expands these previous research work.

The rest of the chapter is organized as follows: In Section 10.2, the definition of the FLN is introduced, and the critical connectivity of the FLN is deduced using the anneal approximation. Then, the structure of the inference algorithm is discussed in Section 10.3. Finally, in Section 10.4, the algorithm is used to infer the gene regulatory networks for *Saccharomyces cerevisiae* and *Saccharomyces pombe*. The inference results are compared and analyzed in this section. The chapter concludes in Section 10.5.

## 10.2 Fuzzy Logic Network Theory

The proposed FLN theory is based on theoretical deductions at the second level of fuzzy uncertainty. This means that the variables have been normalized into $[0,1]$ interval, and statistical methods are built on the basis of fuzzy variables as well as their relationships.

The FLN is defined as follows:

Given a set of $N$ variables (genes), $\Sigma(t) = (\sigma_1(t), \sigma_2(t), \ldots, \sigma_N(t))$, $(\sigma_i(t) \in [0,1]$, $(i = 1,2,\ldots,N)$, index $t$ represents time), the variables are to be updated by means of the dynamic equations:

$$\sigma_i(t+1) = \Lambda_i(\sigma_{i_1}(t), \sigma_{i_2}(t), \ldots, \sigma_{i_K}(t)),\ (1 \le i \le N) \tag{10.1}$$

**Table 10.1.** Four commonly used fuzzy logical functions including their AND ($\wedge$), OR ($\vee$), and NOT (-).

| Fuzzy Logical Functions | $a \wedge b$ | $a \vee b$ | $\overline{a}$ |
|---|---|---|---|
| Max-Min | min($a,b$) | max($a,b$) | 1-$a$ |
| GC | $a \times b$ | min(1, $a+b$) | 1-$a$ |
| MV | max(0, $a+b-$ 1) | min(1, $a+b$) | 1-$a$ |
| Probabilistic | $a \times b$ | $a+b-a \times b$ | 1-$a$ |

where $\Lambda_i$ is a fuzzy logical function, and $K$ represents the number of regulators for $\sigma_i$.

For an FLN, the logical functions may be constructed using the combinations of AND ($\wedge$), OR ($\vee$), and NOT (-). The total number of choices for fuzzy logical functions is determined only by the number of inputs. If a node has $K$ $(1 \leq K \leq N)$ inputs, then there are $2^K$ different logical functions. In the definition of the FLN, each node, $F_i(t)$, has $K$ inputs. But this fixed connectivity will be relaxed later.

To apply the FLN to modeling gene regulatory networks, each fuzzy variable will represent a gene, and genetic regulatory relationships will be modeled as fuzzy logical functions. A fuzzy logical function is defined as a function $\Lambda : U \rightarrow [0,1]$ where $\Lambda(u)$ is the degree of the membership. Usually, it has to satisfy the requirement of the t-norm/t-co-norm, which is a binary operation that satisfies the identity, commutative, associative, and increasing properties [15]. Table 10.1 shows the commonly used fuzzy logical functions with distinctive dynamics [16].

Although the logical functions are expressed via simple algebraic expressions, they have their own distinctive properties. The Max-Min logical function which is closely related to Boolean logic, is one of the classical fuzzy logical functions. This logical function uses the maximum of two values to replace the Boolean OR, whereas the minimum replaces the Boolean AND. GC logical function is a combination of MV and Probabilistic logical functions. The MV logical function follows the trivalent logic whereas Probabilistic does not. In this chapter, all four fuzzy logical functions are tested on the *S. cerevisiae* dataset.

The critical connectivity of the FLN is crucial in the data-driven algorithm's application to gene regulatory network inference. It quantifies the algorithm's search strategy, and the computational complexity of the algorithm is determined by it. To study the detailed dynamics and the connectivity of the FLN, the annealed approximation [17, 18] has been used. Consider the following two FLN configurations at time $t$: $\Sigma(t)$ and $\widetilde{\Sigma}(t)$, where

$$\begin{aligned} \Sigma(t) &= \{\sigma_1(t), \sigma_2(t), \ldots, \sigma_N(t)\} \\ \widetilde{\Sigma(t)} &= \{\widetilde{\sigma_1}(t), \widetilde{\sigma_2}(t), \ldots, \widetilde{\sigma_N}(t)\} \end{aligned} \tag{10.2}$$

Assume that logical functions selected by the two configurations are not time variant throughout the dynamic process. Then the distance between the two configurations may be computed as the accumulated Hamming distance (AHD):

$$AHD(t) = \sum_{i=1}^{N} Hamming(\sigma_i(t), \widetilde{\sigma}_i(t)) \tag{10.3}$$

and

$$Hamming(\sigma_i(t), \widetilde{\sigma}_i(t)) = \begin{cases} 1 & \text{if } |\sigma_i(t) - \widetilde{\sigma}_i(t)| > \delta \\ 0 & \text{if } |\sigma_i(t) - \widetilde{\sigma}_i(t)| \leq \delta \end{cases} \tag{10.4}$$

The Hamming distance uses $\delta \in [0,1]$ (Hamming threshold) as a parameter to differentiate the closeness of two values. The distance between two Boolean values may also be computed using (10.4) with $\delta \equiv 0$. Thus, the AHD of the FLN is the extension of the Boolean distance. One may easily see that the maximum distance between $\Sigma(t)$ and $\widetilde{\Sigma}(t)$ is $N$, while the minimum distance is 0. In comparison with the distance, another quantity, $a_t \in [0,1]$, may be defined as the similarity of the two networks, i.e.,

$$a_t = 1 - \frac{AHD(t)}{N} \tag{10.5}$$

Suppose at time $t$, $\Sigma(t)$ and $\widetilde{\Sigma}(t)$ are at distance $l_t$. Then the probability of the two configurations having a distance $l_{t+1}$ at time $t+1$ may be found. This change in distances represents the dynamic paths of the two configurations. Denote this probability as $P_t(l_{t+1}, l_t)$. Suppose $\Sigma(t)$ and $\widetilde{\Sigma}(t)$ have the same logical function selections for their corresponding variables but different initial values for each variable, and the variables in the two systems can select one out of $S$ values ($S$ is finite, $S \geq 2N$). The requirement of $S \geq 2N$ is to guarantee that different fuzzy logical functions may be used by the FLN [19]. The probability of selecting each of the $S$ values is assumed to be the same, i.e. $\frac{1}{S}$.

Suppose $A$ is the set of variables which are identical in $\Sigma(t)$ and $\widetilde{\Sigma}(t)$ at time $t$. Obviously, set $A$ has $N - l_t$ variables. Define $Q(N_0)$ as the probability that $N_0$ variables have all their $K$ parents from set $A$. Then, $Q(N_0)$ is a discrete random variable following the binomial distribution with parameter $\left(\frac{N-l_t}{N}\right)^K$. By definition,

$$\frac{N - l_t}{N} = 1 - \frac{l_t}{N} = a_t \tag{10.6}$$

so,

$$\begin{aligned} Q(N_0) &= \binom{N}{N_0} \left[\left(\frac{N-l_t}{N}\right)^K\right]^{N_0} \left[1 - \left(\frac{N-l_t}{N}\right)^K\right]^{N-N_0} \\ &= \binom{N}{N_0} \left[{a_t}^K\right]^{N_0} \left[1 - {a_t}^K\right]^{N-N_0} \end{aligned} \tag{10.7}$$

It is obvious that these $N_0$ variables will be the same at time $t+1$ in both $\Sigma(t+1)$ and $\widetilde{\Sigma}(t+1)$. For the remaining $N - N_0$ variables, since at least one of their parents will be different, there is a probability of $p = \frac{S(S-1)}{S^2}$ that a variable will be different in two networks at the next step, while $1-p$ is the probability that it will be the same.

More generally, let $P$ be the probability that a function produces different values from different inputs. If a variable can take $S$ values, and the probability of selecting one of these values is the same, then $P$ may be expressed using the Bayesian rule.

$$P = p(S-1)\frac{1-p}{S-1} + (S-1)\frac{1-p}{S-1}(1 - \frac{1-p}{S-1}) \tag{10.8}$$

Thus, through deductions using the annealed approximation, the following equation may be found:

$$P(l_{t+1}, l_t) = \frac{N!}{l_{t+1}!(N-l_{t+1})!}\left(P(1-a_t^K)\right)^{l_{t+1}} \quad \left[1-P(1-a_t^K)\right]^{N-l_{t+1}} \tag{10.9}$$

As can be seen, (10.9) follows binomial distribution. Thus the possibility of the coverage at the next step will peak at the current mean. The dynamic recursive equation, then, may be expressed as

$$a_{t+1} = 1 - P(1 - {a_t}^K) \tag{10.10}$$

A general situation is considered in which $P$ is uniformly distributed. Then $P$ can be computed as,

$$P = P(|\sigma_i(t) - \sigma_j(t)| \geq \delta) = (1-\delta)^2 \tag{10.11}$$

If the two networks converge, then the following marginal stability should be imposed:

$$\begin{array}{ll} \frac{\partial a_{t+1}}{\partial a_t} < 1 & \text{(implies that the coverage does not decrease with time)} \\ \lim_{t\to\infty} a_t = 1 & \text{(the condition for a full coverage in the steady state)} \end{array} \tag{10.12}$$

If a network does not have uniform connectivity for all nodes, we may assume that the nodes may have different number of parents with a discrete distribution $\rho_k$, where

$$\rho_k = Prob(\text{a node has } k \text{ parents}) \text{ and } \sum_{k=1}^{N} \rho_k = 1 \tag{10.13}$$

By applying (10.11), (10.12), and (10.13) to (10.10), the following relationship may be found.

$$\overline{K} = \frac{1}{(1-\delta)^2} \tag{10.14}$$

It has been found that, in yeast protein-protein networks, as well as in the Internet and social networks, the distribution of connectivity follows the Zipf's law [20], i.e.,

$$P((\text{number of inputs}) = K) \propto \frac{1}{K^\gamma}, 1 \leq K \leq N \tag{10.15}$$

where $\gamma$ is a real number, usually between 2 and 3.

Hence, according to (10.15), the mean connectivity may be computed as

$$\overline{K} = \varepsilon \sum_{K=1}^{N} K \frac{1}{K^{\gamma}} = \varepsilon \sum_{K=1}^{N} \frac{1}{K^{\gamma-1}} \tag{10.16}$$

where $\varepsilon$ is a constant to guarantee that the sum of distribution equals 1.

Then, define

$$H_N^{(\gamma)} = \sum_{i=1}^{N} \frac{1}{K^{\gamma}} \tag{10.17}$$

as the partial sum of the generalized harmonic series. It may be proved that

$$\varepsilon = \frac{1}{H_N^{\gamma}} \quad \text{and} \quad \overline{K} = \frac{H_N^{\gamma-1}}{H_N^{\gamma}} \tag{10.18}$$

Since there is no general formula for (10.17), approximations for the sum may be used if $N$ is large enough, which is true for the application to gene regulatory network inference. The approximation of $H_N^{(\gamma)}$ is

$$H_N^{(\gamma)} \approx \begin{cases} \infty & \text{if } \gamma = 1 \\ \frac{\pi^2}{6} & \text{if } \gamma = 2 \\ 1.202 & \text{if } \gamma = 3 \\ \frac{\pi^4}{90} & \text{if } \gamma = 4 \\ 1.036 & \text{if } \gamma = 5 \\ \frac{\pi^6}{945} & \text{if } \gamma = 6 \end{cases} \tag{10.19}$$

By substituting (10.19) into (10.18), the mean connectivity of the network may be found as

$$\overline{K} \approx \begin{cases} \infty & \text{if } \gamma = 2 \\ 1.3685 & \text{if } \gamma = 3 \\ 1.1106 & \text{if } \gamma = 4 \\ 1.0447 & \text{if } \gamma = 5 \\ 1.0183 & \text{if } \gamma = 6 \end{cases} \tag{10.20}$$

By applying (10.8), (10.12), and (10.13) to (10.10), the relationship between $S$ and $\overline{K}$ is

$$S \geq \frac{\overline{K}}{\overline{K} - 1} \tag{10.21}$$

Therefore, by substituting $\overline{K}$ in (10.20) to (10.21), the value of $S$ may be found to have a lower bound, i.e.

$$S > \begin{cases} 1 & \text{if } \gamma = 2 \\ 3.7137 & \text{if } \gamma = 3 \\ 10.0416 & \text{if } \gamma = 4 \\ 23.3714 & \text{if } \gamma = 5 \\ 55.6448 & \text{if } \gamma = 6 \end{cases} \tag{10.22}$$

In general, the connectivity of a real network should be greater than the critical connectivity because real networks are usually much more complicated. In other words, the critical connectivity only serves as a lower bound. In addition, more relaxed criteria may find more possible regulations, and provide a much smaller search space for further investigations. Thus an $S$ with $\gamma \geq 3$ should be chosen in initial searches, which means $S$ should be more than 2. This triggers the question of whether a crisp Boolean network is powerful enough to infer the genetic network structure.

If (10.20) is substituted into (10.14), the requirement that $\delta$ must satisfy may be found as

$$\delta > \begin{cases} 0.1452 & \text{if } \gamma = 3 \\ 0.0511 & \text{if } \gamma = 4 \\ 0.0216 & \text{if } \gamma = 5 \\ 0.0090 & \text{if } \gamma = 6 \end{cases} \tag{10.23}$$

When $\gamma$ increases, the network has to adjust itself by adopting stricter criteria. The result also agrees with the relationship between $S$ and $\gamma$ from (10.22). However, when the FLN is used on a real dataset, $\delta$ must also be increased to account for the noise inside.

## 10.3 Algorithm

Let $\boldsymbol{G} \in R^{n\times m}$ be the time-series microarray data where $n$ is the number of genes in the data and $m$ is the number of time slots in the microarray set. The algorithm will first randomly select $G_r = (G_{r_1}, G_{r_2}, ..., G_{r_i}, ...G_{r_K})$, a group of regulators that regulates $G_t$ ($G_{r_i}, G_t \in R^{1\times m}, t \neq r_1, r_2, ...r_K$). Then the algorithm will filter the regulators through a fuzzy logic mask, *FLogic*[1], to generate a pseudo-gene-time-series, $G_s \in R^{1\times m}$ where

$$G_s^j = Flogic(G_{r_1}^j, G_{r_2}^j, ..., G_{r_i}^j, ...G_{r_K}^j), j \in \underline{m} \tag{10.24}$$

The distance between $G_r$ and $G_t$ is then computed as:

$$Distance(G_r, G_t) = AHD(G_s, G_t) = \sum_{j=1}^{m-1} Hamming(G_t^{j+1}, G_s^j) \tag{10.25}$$

where the Hamming distance is computed according to (10.4). The value of $\delta$ should base on (10.23), on the noise level, and on the data completeness. As shown in (10.25), the AHD between $G_t$ and $G_s$ is computed with a time shift throughout the time series, which is a reasonable assumption that regulations happen with one time delay.

For each group of possible regulators and the regulated gene, the algorithm determines its AHD, and records it. In the end, the algorithm will infer regulatory groups

[1] *Flogic* is one of the possible fuzzy logical functions that is applied on $K$ variables

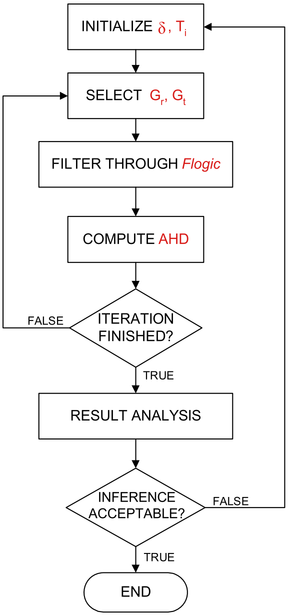

**Fig. 10.1.** Algorithm flowchart

with $\frac{AHD}{m}$ less than another threshold, $T_i, (T_i \in [0,1])$, which is the inference threshold. This threshold is defined as the percentage of time periods that a regulation persists with respect to the total time slots in the microarray. The flow chart of the algorithm is depicted in Fig.10.1

The complexity of the algorithm is $O(n^{\overline{K}})$ where $n$ is the number of genes, and $\overline{K}$ is the critical connectivity of the FLN. As shown in (10.20), the mean connectivity should be less than 1.3685. Thus the complexity of the algorithm is almost linear with the number of genes. Based on the result shown in (10.20) and [21], we assume the maximum number of regulators for each inference group is 2. In this scenario, there may be more than 2 regulators for a regulated gene but we assume that regulators are correlated in a pair-wise manner. This assumption, which assumes that two regulators are more efficient to deliver regulatory mechanisms, ensures the critical connectivity. One should note that the total number of regulators is not limited by this assumption (any gene in the network may have $N$ regulators maximally).

## 10.4 Applications to Yeast Datasets

To test the algorithm, the $\alpha$-factor arrest subset of *S. cerevisiae* cell cycle dataset [22] and the *S. pombe* cell cycle dataset [23] are selected. Although the budding yeast, *S. cerevisiae*, and the fission yeast, *S. pombe*, are generally named as yeast, they diverged about 1.1 billion years ago [24]. It has been shown that these two fungi have different chromosome structures, cell cycle synergy, and different gene functions. In addition, it has been reported that the S and M phases of *S. cerevisiae* are very hard to disrupt while the disrupt of *S. pombe* cell cycles can be relatively easier [25]. The inference results shown later in this chapter confirm the differences of gene regulation strategies in the two yeast, and offer a possible explanation on why the *S. cerevisiae* cell cycle is more stable.

### 10.4.1 Data Pre-processing

The *S. cerevisiae* dataset includes 6178 commonly accepted open reading frames, and for the $\alpha$-arrest subset there are 18 time slots. The *S. pombe* dataset includes 99.5 percent of the total open reading frames, but we choose to use the 407 genes that have been inferred to be cell-cycle regulated [23]. For the *S. pombe* dataset, three *elutriation* and two *cdc*25 block-release subsets are combined to produce a series of 115 time slots for each gene. To reduce errors introduced by noise, and to reduce computational complexity, both datasets are pre-processed with 3 criteria. Genes that do not satisfy all three criteria are deleted. The criteria are stated as follows:

- Only genes having more than two thirds valid time slots, with respect to the entire time span in the microarray, are considered above the noise level. The reason for this requirement is that some genes do not have expression strong enough to counter the background noise at some time slots. Totally 115 genes do not fit this criterion in the *S. pombe* dataset. The number of genes deleted in the *S. cerevisiae* dataset is 125.
- For *S. cerevisiae*, the maximum value of each gene's expression must be at least three times greater than its minimum value in the time series. For *S. pombe*, the ratio is limited to 1. If not, the gene is excluded from the dataset. This requirement guarantees that genes running inside the algorithm have a dynamic range of expression. Thus it reduces the computational time by limiting the search space. *S. cerevisiae* dataset has 5366 genes deleted, but no gene is deleted from *S. pombe* dataset due to this criterion.
- Genes with spikes in the time series are not included. The signal-to-noise ratio of the spike is defined as five. 4 genes in the *S. pombe* dataset and 8 genes in the *S. cerevisiae* dataset have spikes, and are deleted.

After the pre-processing and filtering, 680 genes in the *S. cerevisiae* subset are found to satisfy all three criteria, and, as far as the *S. pombe* dataset is concerned, 286 genes have survived the cut. The values of gene expression are then normalized into $[0,1]$ interval throughout the time series. For *S. cerevisiae*, the values of the dataset are changed from log-odds into true values. Then, the maximum value of each gene series is found, and used to divide the expression of that gene in the series. For *S. pombe*, every gene series in each subset is normalized to have zero median. In the next step, the maximum value of a gene's expression inside each subset is found, and used to divide the values of that gene's expression in the same subset (the five subsets are normalized separately). After these steps, the values of each gene in the dataset have been normalized into $[0,1]$ interval.

### 10.4.2 Inference Results of the *S. cerevisiae* Dataset

We investigate the inference results when regulators apply controls to the regulated gene in pairs. All four logical functions in Table 10.1 are tested with same parameter settings ($\delta = 0.01, T_i = 21\%$), and they have inferred four different regulatory networks. The MV logical function do not introduce as many false positives as that from using other commonly used fuzzy logical functions. Furthermore, MV logical function causes the algorithm to be less sensitive to small variations of parameters i.e., $\delta$ from (10.4) and the inference threshold of the algorithm, $T_i$. The inferred network based on the MV logical function is shown in Fig. 10.2 and Table 10.2.

Out of $4.3 \times 10^9$ possible groups, the algorithm locates 51 regulatory groups (32 regulatory pairs) involving 21 genes with average connectivity of 1.5238. There are 17 verified regulations, 5 unknown regulations, and 10 dubious regulations. Table 10.2 shows the verified regulations with functions of the regulator and regulated gene. One interesting finding is that 15 out of the 21 genes in the network have been proved to be involved in yeast mating or the cell cycle, and most of them are downstream mating regulatory genes. In addition, the backbone of the network (nodes with high connectivities) is made up of 9 out of these 15 genes. The clustering attribute in the result, although unexpected, may explain why 14 out of the 17 verified regulations are based on close relationships.

The inferred network also shows network motifs. The network includes seven feed-forward loops, three single-input modules, as well as the dense overlapping and bi-fan modules [33, 34]. Through comparative studies on complex biological, technological, and sociological networks, it has been shown that these modules share different evolutionary properties and advantages [35, 36]. The feed-forward loop is believed to play a functional role in information processing. This motif may serve as a circuit-like function to activate output only if the input signal is persistent, and allows a rapid deactivation if the input signal is off. Further, the bi-fan structure of (*PRM1*, *FUS2*) and (*FIG1*, *ASG7*) are coupled with a number of feed-forward motifs. The inferred network also includes two internal cycles (*FIG1* $\longleftrightarrow$ *PRM1*, *FIG1* $\longleftrightarrow$ *FIG2*) and one feedback loop among *FIG1*, *FIG2* and *ASG7*. All the genes in the cycles or feedback loop are involved in the signaling for yeast mating, and the close regulations among them are integral to yeast mating. Although network motif studies on *E. Coli* have not found cyclic structures [34, 37], the feedback loop is believed to be the most important regulatory mechanism for cells to adapt to new environments. The inferred network shows that while preserving specific regulatory strategies, different species share a striking similarity of regulation mechanisms.

### 10.4.3 Inference Results of the *S. pombe* Dataset

The algorithmic parameters for the *S. pombe* dataset are set as $\delta = 0.018$ because the network is selected to be a more general network according to (10.23), and $T_i = 71\%$ because the combination of five different subsets lowers the percentage of time that a regulation may persist. The quantifications of the two thresholds are also based on previous investigations on the algorithm's behaviors [12]. The algorithm uses the

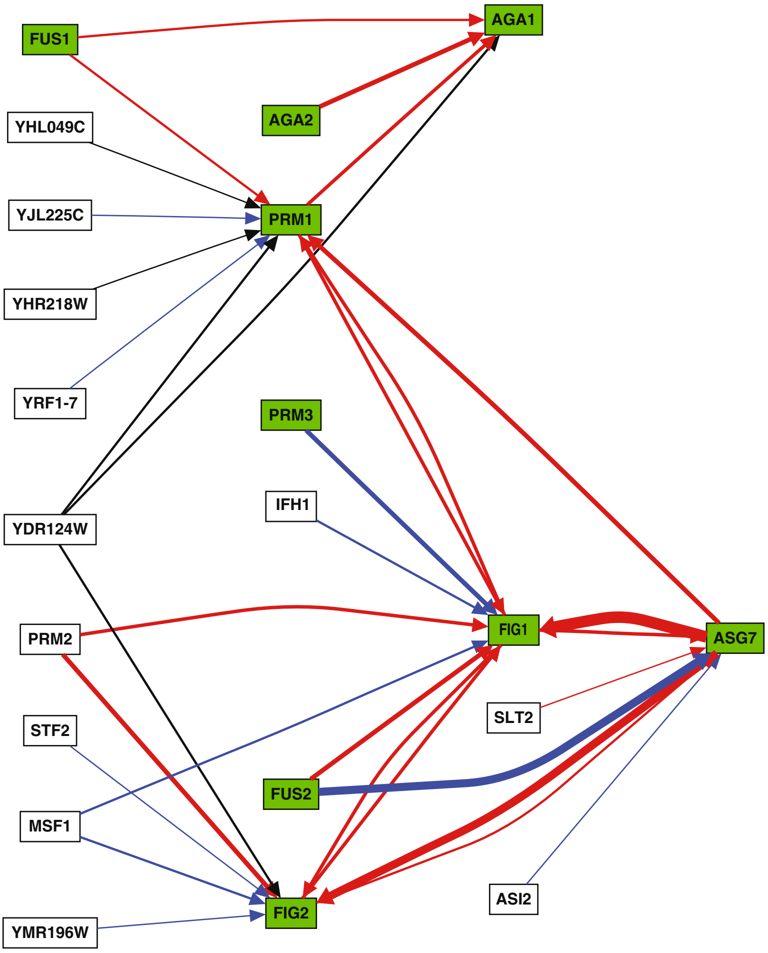

**Fig. 10.2.** The $\alpha$-factor gene regulatory network using the MV fuzzy logical function. In the network, there are 21 genes and 32 regulatory arcs. Nodes with high connectivities are green. The colors of arcs are coded as: functionally verified regulations (red), regulations involving genes with unknown functions (black), and dubious regulations (blue). The width of the arcs represents the regulatory strength between the regulator and regulated gene, and it is calculated as $(m - AHD)$

MV logical function to find 105 regulations (125 regulatory pairs) among 108 genes. The regulatory network is shown in Fig. 10.3.

In the network, there are 108 genes and 125 regulatory arcs. The colors of arcs are coded as: functionally verified regulations (red), regulations involving genes with unknown functions (black), and dubious regulations (blue). The width of the arcs represents the regulatory strength between the regulator and regulated gene, and it is calculated as $(m - AHD)$. The network includes 59 functionally verified regulations,

**Table 10.2.** Functionally verified regulations in the inferred *S. cerevisiae* gene regulatory network. The regulations are grouped by regulators and the criteria of verifications are categorized as functional (verified by gene functions from Saccharomyces Genome Database, 4 regulations), close relationship (regulators and regulated genes are usually co-expressed, co-regulated, 14 regulations). Some of the verifications are also based on the included references

| Regulator | Regulated gene | Verification |
|---|---|---|
| **AGA2**: adhesion subunit of a-agglutinin of a-cells | **AGA1**: anchorage subunit of a-agglutinin of a-cells | close relationship [26] |
| **ASG7**: regulates signaling from Ste4p | **FIG1**: integral membrane protein for efficient mating | close relationship [27] |
| | **FIG2**: cell wall adhesin specifically for mating | functional |
| | **PRM1**: SUN family gene involved in cell separation | close relationship [28] |
| **FIG1**: integral membrane protein for efficient mating | **ASG7**: regulates signaling from Ste4p | close relationship [27] |
| | **FIG2**: cell wall adhesin specifically for mating | close relationship [29] |
| | **PRM1**: pheromone-regulated protein for membrane fusion during mating | close relationship [27] |
| **FIG2**: cell wall adhesin specifically for mating | **FIG1**: integral membrane protein for efficient mating | close relationship [29] |
| | **ASG7**: regulates signaling from Ste4p | functional |
| **FUS1**: membrane protein required for cell fusion | **AGA1**: anchorage subunit of a-agglutinin of a-cells | close relationship [30] |
| | **PRM1**: pheromone-regulated protein for membrane fusion during mating | close relationship [31] |
| **PRM1**: pheromone-regulated protein for membrane fusion during mating | **AGA1**: anchorage subunit of a-agglutinin of a-cells | close relationship [30] |
| | **FIG1**: integral membrane protein for efficient mating | close relationship [27] |
| **PRM2**: pheromone-regulated protein regulated by Ste12p | **FIG1**: integral membrane protein for efficient mating | close relationship [27] |
| | **FIG2**: cell wall adhesin specifically for mating | close relationship [27] |
| **SLT2**: suppressor of lyt2 | **ASG7**: regulates signaling from Ste4p | close relationship [32] |
| **FUS2**: cytoplasmic protein for the alignment of parental nuclei before nuclear fusion | **FIG1**: integral membrane protein for efficient mating | close relationship |

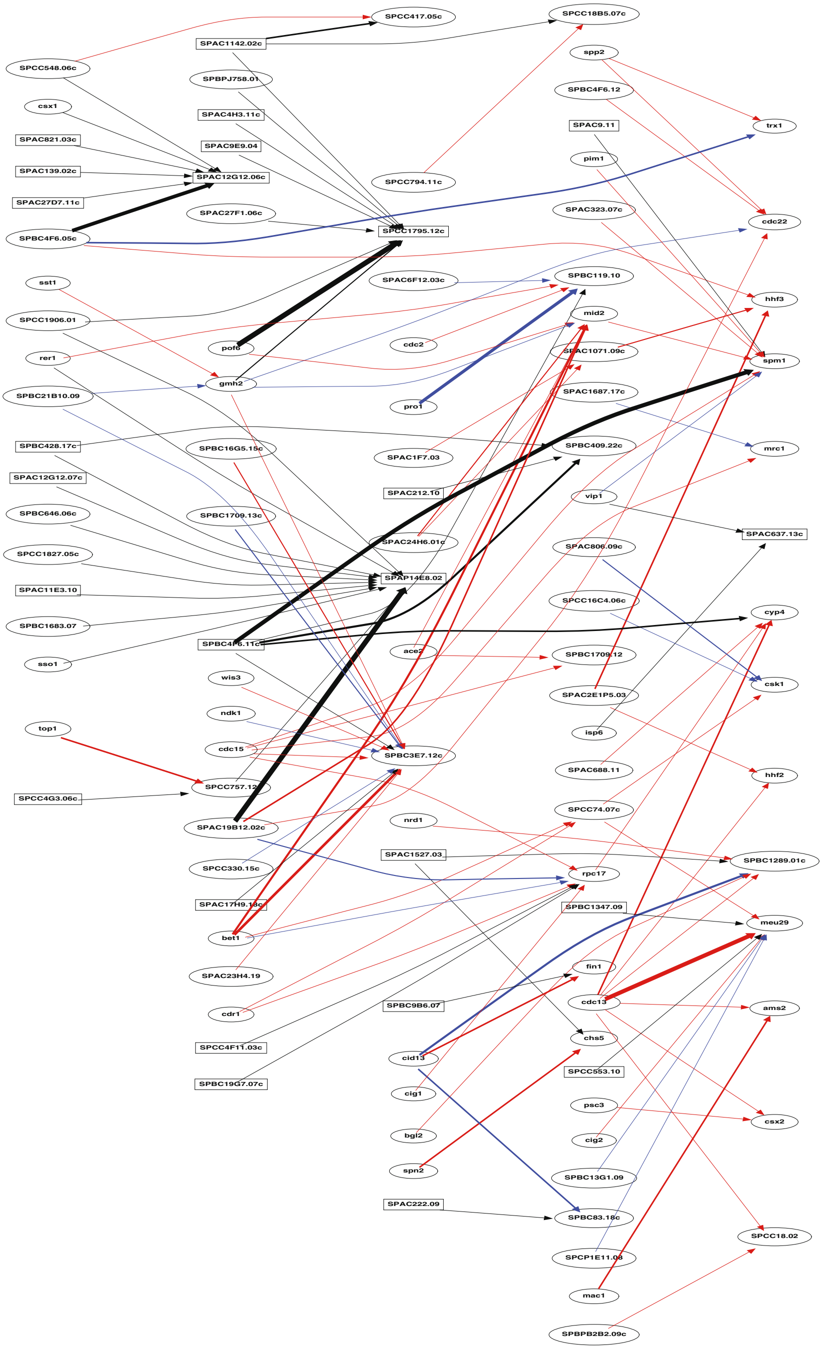

**Fig. 10.3.** The *S. pombe* gene regulatory network using the MV fuzzy logical function

47 regulations involving genes with unknown functions, and 19 dubious regulations. The 59 functionally verified regulations are listed in Table 10.3 with the functions of regulators and regulated genes.

Take the *ace2* - *mid2* and *cdc15* - *spm1* regulatory pairs as examples. Gene *ace2* is a transcription factor that regulates the transcription of genes required for cell separation; *mid2* is essential for the central positioning of the division septum before the cell divides and in organizing the septin ring during late mitosis. Gene *cdc15* mediates cytoskeletal rearrangements required for cytokinesis on the onset of mitosis, and *spm1* involves in the regulation of cell wall structure. The average connectivity of the inferred network is 1.157 which fits the requirement of the Zipf's law as shown in (10.20) with $\gamma$ between 2 and 3.

The regulations among genes in *S. pombe* are different from *S. cerevisiae* largely because the regulatory network of *S. pombe* does not include any feed-forward structure or cycle. As shown before, the feed-forward loop is a stable motif for the network, and this might offer an explanation why *S. pombe* cell cycles are less stable. We also found that the regulatory network of *S. pombe* is sparser than that of the *S. cerevisiae*. As for the regulatory logic in the two networks, we found striking similarities between the two yeast. For *S. pombe*, 75 percent of regulations are OR logic while the majority of the remaining 25 percent are single-regulator situations. For *S. cerevisiae*, 63 percent are OR logic and 17 percent are single-regulator scenarios.

## 10.5 Conclusions and Discussion

The biologically plausible results from the applications of our FLN algorithm to the *S. cerevisiae* and *S. pombe* datasets suggest that the data-driven algorithm is potentially a powerful tool to decipher the causal pathways involved in complex biological systems. In this work, the focus has been on the theoretical deduction of the FLN's dynamic behavior and on the computational aspects of the inference algorithm. The theory of the FLN not only provides a sound theoretical guarantee for algorithmic parameter adjustments, but also is also a novel proposal for a new network model with potentially broad applications in modeling complex networks. From the computation results, the algorithm has provided detailed and insightful causal relationships among various genes. Thus, we believe that, given less noisy data, the FLN algorithm may be applied to a large range of biological systems having different spatial or temporal scales.

Regarding future research on the theoretical aspects of the FLN, we think that the dynamics and the steady-state properties of the FLN are important. Further effort should also focus on the effect of other distance metrics as means of comparing the performance of the modeling. It is also our belief that there is still room for improvement in computational complexity via heuristic search. Although the theory of the FLN is still in its infancy, particularly with respect to the details of network evolution, we think that the FLN, in the future, can model the real world uncertainty and mimic the behaviors of complex systems.

**Table 10.3.** Functionally verified regulations in the inferred *S. pombe* gene regulatory network. The regulations are grouped by regulators and the criteria of verifications are categorized as functional (verified by gene functions from GeneDB database, 49 regulations), close relationship (regulators and regulated genes are usually co-expressed or co-regulated, 9 regulations), and homolog (1 homolog). Some of the verifications are also based on the included references

| Regulator | Regulated gene | Verification |
|---|---|---|
| **ace2**: regulator of cell separation genes | **mid2**: positioning of the division septum before the cell divides<br>**SPBC1709.12**: Rho GTPase binding signaled by cell cycle | functional<br>functional |
| **bet1**: controls intracellular protein transport and cell wall formation | **mid2**: positioning of the division septum before the cell divides<br>**SPBC3E7.12c**: chitin biosynthesis<br>**SPCC74.07c**: SUN family gene involved in cell separation | close relationship<br>functional<br>functional |
| **bgl2**: regulates cell expansion during growth and cell-cell fusion during mating | **SPBC1289.01c**: involves in septum formation | functional [38] |
| **cdc13**: controls the cell cycle at the G2/M (mitosis) transition | **ams2**: required for proper chromosome segregation<br>**csx2**: involves in cell-cycle regulated ADP-ribosylation<br>**cyp4**: peptidyl-prolyl cistrans isomerase involved in mitosis<br>**hhf2**: histone H4<br>**meu29**: up-regulated in meiotic<br>**SPBC1289.01c**: cell wall chitin protein<br>**SPCC18.02**: involves hydrogen anti-porter activity in cell cycle | functional<br>functional [39]<br>functional [40]<br>functional<br>functional<br>functional<br>functional |
| **cdc15**: mediates cytoskeletal rearrangements required for cytokinesis | **mrc1**: mediator of replication checkpoint 1<br>**rpc17**: RNA polymerase<br>**SPBC1709.12**: Rho GTPase binding signaled by cell cycle<br>**SPBC3E7.12c**: chitin biosynthesis<br>**spm1**: involves in cell separation | functional<br>functional<br>functional [41]<br>functional<br>functional |
| **cdc2**: controls the eukaryotic cell cycle | **SPBC119.10**: asparagine synthase involved in glutamine-hydrolyzing | functional |
| **cdr1**: mitotic inducer | **rpc17**: RNA polymerase<br>**SPCC74.07c**: involve in beta-glucosidase activity at cell separation | functional<br>functional |
| **cid13**: creates the 3’ poly(A) tail of suc22 mRNA affecting DNA replication | **fin1**: Promotes chromosome condensation and nuclear envelope dynamics during mitosis | functional [42] |

(*continued*)

**Table 10.3** (Continued)

| Regulator | Regulated gene | Verification |
|---|---|---|
| **cig1**: cyclin regulating G1/S transition | **rpc17**: RNA polymerase | functional |
| **cig2**: cyclin regulating G2/M transition | **meu29**: up-regulated in meiotic | functional |
| **gmh2**: affects Golgi membrane and chitin synthase | **SPBC3E7.12c**: chitin biosynthesis | functional [43] |
| **mac1**: required for cell separation | **ams2**: required for proper chromosome segregation | functional |
| **mid2**: positioning of the division septum before the cell divides | **spm1**: involves in cell separation | functional |
| **nrd1**: negative regulator of sexual differentiation affecting chitin biosynthesis | **SPBC1289.01c**: cell wall chitin protein | functional [44] |
| **pim1**: involves in the control of mitosis | **spm1**: involves in cell separation | functional |
| **pof6**: involves in cell division | **mid2**: positioning of the division septum before the cell divides | functional |
| **psc3**: required for normal mitosis | **csx2**: involves in cell-cycle regulated ADP-ribosylation | functional |
| **rer1**: COPI-coated vesicle | **SPBC119.10**: asparagine synthase | close relationship [45] |
| **rpc17**: RNA polymerase | **cyp4**: PPIases to accelerate the folding of proteins | close relationship [46] |
| **SPAC1071.09c**: contains a DnaJ domain which mediates interactions with histone-modifying heat shock proteins | **hhf3**: histone 4 | functional |
| **SPAC19B12.02c**: high similarity to 1,3-beta-glucanosyl transferase | **cdc22**: provides the precursors necessary for DNA synthesis<br>**mid2**: positioning of the division septum before the cell divides | close relationship [47]<br>close relationship |
| **SPAC1F7.03**: involves in calcium transport that affects heat shock genes | **SPAC1071.09c**: interacts with heat shock proteins | functional [48] |
| **SPAC23H4.19**: putative cell wall biogenesis protein | **SPBC3E7.12c**: chitin biosynthesis | functional |
| **SPAC24H6.01c**: involved in phospholipid biosynthesis affecting cell division | **mid2**: positioning of the division septum before the cell divides<br>**SPAC1071.09c**: interacts with heat shock proteins | functional [49]<br>functional [50] |

(*continued*)

**Table 10.3** (Continued)

| Regulator | Regulated gene | Verification |
|---|---|---|
| **SPAC2E1P5.03**: mediates interaction with heat shock proteins that interacts with histones | **hhf2**: histone 4<br>**hhf3**: histone 4 | functional [51]<br>functional [51] |
| **SPAC323.07c**: member of the MatE family of integral membrane proteins | **spm1**: involves in cell separation | close relationship |
| **SPAC688.11**: required for hyphal growth | **cyp4**: peptidyl-prolyl cis-trans isomerase involved in mitosis | close relationship [52] |
| **SPBC16G5.15c**: required for the correct timing, positioning and contraction of the division septum | **SPBC3E7.12c**: involves chitin biosynthesis | functional |
| **SPBC4F6.05c**: involves in sugar biding that affects histones | **hhf3**: histone 4 | functional [53] |
| **SPBC4F6.12**: regulates integrin or growth factor-mediated responses | **cdc22**: provides the precursors necessary for DNA synthesis | functional |
| **SPBPB2B2.09c**: involves in thiamine biosynthesis | **SPCC18.02**: involves hydrogen anti-porter activity | close relationship [54] |
| **SPCC548.06c**: involves in glucose transport that affects chitin synthesis | **SPCC417.05c**: stimulates chitin synthase III activity | functional [55] |
| **SPCC74.07c**: involves in beta-glucosidase activity at cell separation | **csk1**: cell cycle kinase<br>**meu29**: up-regulated in meiotic | functional<br>functional |
| **SPCC794.11c**: involves in formation of clathrin coats at the Golgi and endosomes | **SPCC18B5.07c**: nuclear pore protein (nucleoporin) | homolog [56] |
| **SPCP1E11.08**: nuclear protein involved in ribosome biogenesis | **meu29**: up-regulated in meiotic | close relationship |
| **spn2**: septin involved in cell separation | **chs5**: involves in chitin synthesis and also required for mating | functional |
| **spp2**: DNA primase, large (non-catalytic) subunit | **cdc22**: provides the precursors necessary for DNA synthesis<br>**trx1**: putative thioredoxin that affects DNA primase | functional<br>functional |
| **sst1**: member of sodium or calcium exchanger protein family of membrane transporters | **gmh2**: affects Golgi membrane and chitin synthase | functional [57] |
| **top1**: DNA topoisomerase I, involved in chromatin organization | **SPCC757.12**: alpha-amylase with special chromatin structure | functional [58] |
| **wis3**: regulates cell cycle progression | **SPBC3E7.12c**: involves in chitin biosynthesis | functional |

## 10.6 Acknowledgment

The authors want to thank the Pittsburgh Supercomputing Center for providing high-speed computer clusters for this work.

## References

1. Strogatz S (2001) Exploring complex networks. Nature 410:268–276
2. Nachman I, Regev A, Friedman N (2004) Inferring quantitative models of regulatory networks from expression data. Bioinformatics 20:i248–i256
3. Matsuno H, Doi A, Nagasaki M, Miyano S (2000) Hybrid Petri net representation of gene regulatory network. Pacific Symposium on Biocomputing 5:341–352
4. TeichMann S, Babu M (2004) Gene regulatory network growth by duplication. Nature Genetics 36:492–496
5. Sprinzak D, Elowitz M (2005) Reconstruction of genetic circuits. Nature 438:443–448
6. Ong I, Page D, and Costa V (2006) Inferring Regulatory Networks from Time Series Expression Data and Relational Data via Inductive Logic Programming. In: Proc. 16th International Conference on Inductive Logic Programming, Spain
7. Kwon A, Hoos H, Ng R (2003) Inference of transcriptional regulation relationships from gene expression data. Bioinformatics 19:905–912
8. Joseph Z (2004) Analyzing time series gene expression data. Bioinformatics 20:2493–2503
9. Georgantas R, Tanadve V, Malehorn M, Heimfeld S, Chen C, Carr L, Martinez-Murillo F, Riggins G, Kowalski J, Civin C (2004) Microarray and serial analysis of gene expression analyses indentify known and novel transcripts overexpressed in hematopoietic stem cells. Cancer Research 64:4434–4441
10. Lee T, Rinaldi N, Robert F, Odom D, Joseph Z, Gerber G, Hannettt N, Harbinson CT, Thompson C.M, Simon I, Zeitlinger J, Jennings E, Murray H, Gordon DB, Ren B, Wyrick J, Tagne J, Volkert T, Fraenkel E, Gifford D, Young R (2002) Transcriptional regulatory networks in *Saccharomyces cerevisiae*. Science 298:799–804
11. Fung E, Thulasiraman V, Weinberger SR, Dalmasso E (2001) Protein biochips for differential profiling. Current Opinion in Biotechnology 12:65–69
12. Cao Y, Wang P, Tokuta A (2005) 3. In: Gene regulating network discovery. Volume 5 of Studies in Computational Intelligence. Springer-Verlag GmbH 49–78
13. Woolf P, Wang Y (2000) A fuzzy logic approach to analyzing gene expression data. Physiological Genomics 3:9–15
14. Cao Y, Wang P, Tokuta A (2006) *S. pombe* regulatory network construction using the fuzzy logic network. Poster, LSS Computational Systems Bioinformatics Conference, Stanford University
15. Hajek P (1998) Metamathematics of fuzzy logic. Volume 4 of Trends in Logic. Kluwer academic publishers, Boston
16. Reiter C (2002) Fuzzy automata and life. Complexity 3:19–29
17. Derrida B, Pomeau Y (1986) Random networks of automata: A simple annealed approximation. Europhysics letters 1:45–49
18. Sole R, Luque B, Kauffman S (2000) Phase transition in random network with multiple states. Technical report, Santa Fe Institute
19. Resconi G, Cao Y, Wang P (2006) Fuzzy biology. In: Proc. 5th International Symposium on Intelligent Manufacturing Systems, Sajarya, Turkey 29–31

20. Kauffman S, Peterson C, Samuelsson B, Troein C (2004) Genetic networks with canalyzing Boolean rules are always stable. PNAS 101:17102–17107
21. Cao Y (2006) Fuzzy Logic Network Theory with Applications to Gene Regulatory Networks. PhD thesis, Duke University
22. Spellman P, Sherlock G, Zhang M, Iyer V, Anders K, Eisen M, Brown P, Botstein D, Futcher B (1998) Comprehensive identification of cell cycle regulated genes of the yeast saccharomyces cerevisiae by microarray hybridization, molecular biology of the cell. Molecular Biology of the Cell 9:3273–3297
23. Rustici G, Mata J, Kivinen K, Lio P, Penkett C, Burns G, Hayles J, Brazma A, Nurse P, Bahler J (2004) Periodic gene expression program of the fission yeast cell cycle. Nature genetics 36:809–817
24. Heckman D, Geiser D, Eidell B, Stauffer R, Kardos N, Hedges S (2001) Molecular evidence for the early colonization of land by fungi and plants. Science 293:1129–1133
25. Forsburg S (1999) The best yeast. Trends in genetics 15:340–344
26. Cappellaro C, Baldermann C, Rachel R, Tanner W (1994) Mating type-specific cell-cell recognition of saccharomyces cerevisiae: cell wall attachment and active sites of a- and alpha-agglutinin. EMBO Journal 13:4737–4744
27. Lockhart S, Zhao R, Daniels K, Soll D (2003) Alpha-pheromone-induced "shmooing" and gene regulation require white-opaque switching during candida albicans mating. Eukaryotic cell 2:847–855
28. Stone E, Heun P, Laroche T, Pillus L, Gasser S (2000) Map kinase signaling induces nuclear reorganization in budding yeast. Current Biology 10:373–382
29. Erdman S, Lin L, Malczynski M, Snyder M (1998) Pheromone-regulated genes required for yeast mating differentiation. Journal of Cell Biology 140:461–483
30. Oehlen L, McKinney J, Cross F (1996) Ste12 and mcm1 regulate cell cycle-dependent transcription of far1. Molecular Cell Biology 16:2830–2837
31. Jin H, Carlile C, Nolan S, Grote E (2004) Prm1 prevents contact-dependent lysis of yeast mating pairs. Eukaryotic Cell 3:1664–1673
32. Wang K, Vavassori S, Schweizer L, Schweizer M (2004) Impaired prpp-synthesizing capacity compromises cell integrity signalling in saccharomyces cerevisiae. Microbiology 150:3327–3339
33. Milo R, Shen-Orr S, Itzkovitz S, Kashtan N, Chklovskii D, Alon U (2002) Network motifs: Simple building blocks of complex networks. Science 298:824–827
34. Shen-Orr S, Milo R, Mangan S, Alon U (2002) Network motifs in the transcriptional regulation network of Escherichia coli. Nature Genetics 31:64–68
35. Thieffry D, Romero D (1999) The modularity of biological regulatory networks. Biosystems 50:49–59
36. Milo R, Itzkovitz S, Kashtan N, Levitt R, Shen-Orr S, Ayzenshtat I, Sheffer M, Alon U (2004) Superfamilies of evolved and designed networks. Science 303:1538–1542
37. Thieffry D, Huerta A, Perez-Rueda E, Collado-Vides J (1998) From specific gene regulation to genomic networks: A global analysis of transcriptional regulation in Escherichia coli. BioEssays 20:433–440
38. Kapteyn J, Montijn R, Vink E, de la Cruz J, Llobell A, Douwes J, Shimoi H, Lipke P, Klis F (1996) Retention of *Saccharomyces cerevisiae* cell wall proteins through a phosphodiester-linked beta-1,3-/beta-1,6-glucan heteropolymer. Glycobiology 6:337–345
39. Jacobson E, Meadows R, Measel J (1985) Cell cycle perturbations following DNA damage in the presence of ADP-ribosylation inhibitors. Carcinogenesis 6:711–714

40. Morris D, Phatnani H, Greenleaf A (1999) Phospho-carboxyl-terminal domain binding and the role of a prolyl isomerase in pre-mRNA 3'-end formation. Journal of Biological Chemistry 274:31583–31587
41. Richnau N, Aspenstrom P (2001) RICH, a rho GTPase-activating protein domain-containing protein involved in signaling by cdc42 and rac1. Journal of Biological Chemistry 276:35060–35070
42. Read R, Martinho R, Wang S, Carr A, Norbury C (2002) Cytoplasmic poly(A) polymerases mediate cellular responses to S phase arrest. PNAS 99:12079–12084
43. Trautwein M, Schindler C, Gauss R, Dengjel J (2006) Arf1p, Chs5p and the CHAPs are required for export of specialized cargo from the Golgi. the EMBO Journal 25:943–954
44. Lee J, Choi J, Park B, Park Y, Lee M, Park H, Maeng P (2004) Differential expression of the chitin synthase genes of Aspergillus nidulans, chsA, chsB, and chsC, in response to developmental status and environmental factors. Fungal Genetics and Biology 41:635–646
45. Moussalli M, Pipe S, Hauri H, Nichols W, Ginsburg D, Kauffman R (1999) Mannose-dependent endoplasmic reticulum (ER)-golgi intermediate compartment-53-mediated ER to Golgi trafficking of coagulation factors V and VIII. Journal of Biological Chemistry 274:32539–32542
46. Chao S, Greenleaf A, Price D (2001) Juglone, an inhibitor of the peptidyl-prolyl isomerase pin1, also directly blocks transcription. Nucleic Acids Research 29:767–773
47. Mouyna I, Hartland R, Fontaine T, Diaquin M, Simenel C, Delepierre M, Henrissat B, Latge J (1998) A 1,3-beta-glucanosyltransferase isolated from the cell wall of Aspergillus fumigatus is a homologue of the yeast Bgl2p. Microbiology 144:3171–3180
48. Mosser D, Kotzbauer P, Sarge K, Morimoto R (1990) In vitro activation of heat shock transcription factor DNA-binding by calcium and biochemical conditions that affect protein conformation. PNAS 87:3748–3752
49. Pierucci O, Rickert M (1985) Duplication of Escherichia coli during inhibition of net phospholipid synthesis. Journal of Bacteriology 162:374–382
50. Glonek J, Lipinska B, Krzewski K, Zolese G, Bertoli E, Tanfani F (1997) HtrA heat shock protease interacts with phospholipid membranes and undergoes conformational changes. Journal of Biolgocial chemistry 272:8974–8982
51. Solomon M, Larsen P, Varshavsky A (1988) Mapping protein-DNA interactions in vivo with formaldehyde: evidence that histone H4 is retained on a highly transcribed gene. Cell 53:937–947
52. Derkx P, Madrid S (2001) The Aspergillus niger cypA gene encodes a cyclophilin that mediates sensitivity to the immunosuppressant cyclosporin A. Molecular Genetics and Genomics 266:527–536
53. Riou-Khamlichi C, Menges M, Healy J, Murray J (2000) Sugar control of the plant cell cycle: Differential regulation of Arabidopsis D-type cyclin gene expression. Molecular and Cellular Biology 20:4513–4521
54. Gastaldi G, Cova E, Verri A, Laforenza U, Faelli A, Rindi G (2000) Transport of thiamin in rat renal brush border membrane vesicles. Kidney International 57:2043–2054
55. Bulik D, Olczak M, Lucero H, Osmond B, Robins P, Specht C (2003) Chitin synthesis in *Saccharomyces cerevisiae* in response to supplementation of growth medium with glucosamine and cell wall stress. Eukaryotic Cell 2:886–900
56. Jekely G, Arendt D (2006) Evolution of intraflagellar transport from coated vesicles and autogenous origin of the eukaryotic cilium. BioEssays 28:191–198
57. Corven E, Os C, Mircheff A (1986) Subcellular distribution of ATP-dependent calcium transport in rat duodenal epithelium. Biochimica et biophysica acta 861:267–276
58. Levy-Wilson B (1983) Chromatin structure and expression of amylase genes in rat pituitary tumor cells. DNA 2:9–13