

The core regulation module of stress-responsive regulatory networks in yeast

Dongsan Kim, Man-Sun Kim and Kwang-Hyun Cho*

Department of Bio and Brain Engineering, Korea Advanced Institute of Science and Technology (KAIST), 291 Daehak-ro, Yuseong-gu, Daejeon 305-701, Republic of Korea

Received February 19, 2012; Revised May 22, 2012; Accepted June 11, 2012

ABSTRACT

How does a cell respond to numerous external stresses with a limited number of internal molecular components? It has been observed that there are some common responses of yeast to various stresses, but most observations were based on gene-expression profiles and only some part of the common responses were intensively investigated. So far there has been no system-level analysis to identify commonly responsive or regulated genes against various stresses. In this study, we identified a core regulation module (CRM), a commonly involved regulation structure in the regulatory networks of yeast, which cells reuse in response to an array of environmental stresses. We found that regulators in the CRM constitute a hierarchical backbone of the yeast regulatory network and that the CRM is evolutionarily well conserved, stable against genetic variations and crucial for cell growth. All these findings were consistently held up to considerable noise levels that we introduced to address experimental noise and the resulting false positives of regulatory interactions. We conclude that the CRM of yeast might be an evolutionarily conserved information processing unit that endows a cell with enhanced robustness and efficiency in dealing with numerous environmental stresses with a limited number of internal elements.

INTRODUCTION

Cells have developed stress-responsive strategies which are dynamically implemented through molecular regulatory networks in order to survive the various stresses they may experience (1–3). A fundamental question arises as to how cells direct specific strategies against many possible stresses given a limited number of molecular components.

Cells may realize stress-specific responses by combinatorial usage of regulatory molecules and their interactions, which suggests that common regulatory molecules might be involved in various specific responses.

Microarray experiments in yeast (4–5) have revealed that hundreds of genes, referred to as environmental stress response (ESR) or common environmental response (CER) genes, are commonly induced or repressed in response to a variety of stresses. These observations indicate that yeast has a common protective mechanism against various types of stress. However, very few regulatory molecules have been identified among ESR or CER genes, perhaps due to their tendency of inducing stable gene expressions. We should also note that many regulatory molecules play their roles in post-translational modifications. Thus, cellular responses may not be fully captured by gene-expression profiles (6), which suggests a fundamental limitation in identifying common regulators based on microarray experiments alone. On the other hand, other relevant studies revealed that signal transduction pathways such as the mitogen-activated protein kinase (MAPK) and cAMP/PKA pathways respond to a variety of stresses (7–10). However, these findings alone are also insufficient for the identification of common regulators because they lack genome-wide perspective.

To overcome the aforementioned limitations, we have employed in the present study the integrated information about genome, signaling networks and transcriptional regulatory networks of yeast and taken a network-based approach to identify common regulators. It is likely that common regulators would not work alone, but act in harmony with their interacting molecules. So, the common regulators would not be equally distributed in the whole yeast regulatory network but would probably exist as a sub-network by forming a core regulation module (CRM) (Figure 1). With this motivation and background, we have investigated such a CRM and identified its presence by solving a maximum-weight connected subgraph (MWCS) problem. From a network perspective, we have further investigated its topological

*To whom correspondence should be addressed. Tel: +82 42 350 4325; Fax: +82 42 350 4310; Email: ckh@kaist.ac.kr

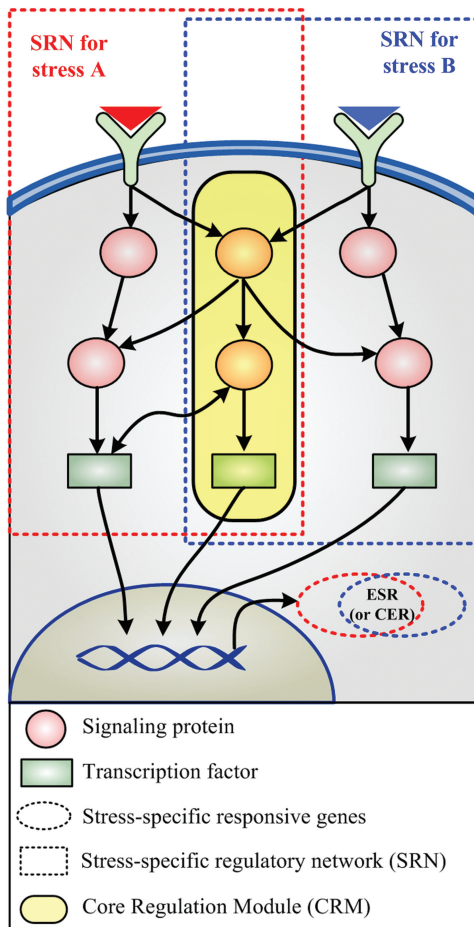


Figure 1. Schematic diagram of a CRM. Dotted red and blue squares indicate the SRN for stresses (A) and (B), respectively. Dotted red and blue ellipses denote the set of stress-responsive genes against stresses (A) and (B), respectively. ESR (or CER) genes are denoted at the intersection between dotted red and blue ellipses.

structure and genetic properties. As a result, we found that the regulators in the CRM form a hierarchical backbone of the yeast-signaling regulatory network and that the CRM is evolutionarily well-conserved, robust against genetic variations and crucial for cell growth. From these, we infer that cells might have developed the CRM as a core information processing unit to respond to numerous stresses with a limited number of internal molecular components, which is an important organizing principle achieved evolutionarily.

MATERIALS AND METHODS

Construction of the global regulatory network of yeast

We collected data describing a broad range of regulatory (signaling and transcriptional) networks of yeast (11–17). All the regulations were collected with stringent cutoff ($P = 0.001$) (17) or validated through the information from the literature (11–16). Using the regulatory information, we constructed a global regulatory network containing 523 nodes and 2093 links where all nodes are regulatory molecules and links are directed regulatory

interactions. In the global regulatory network, regulatory molecules are defined as genes which have at least one out-degree. All the regulations are protein modifications except transcriptional regulations and therefore the regulatory molecules include transcription factors, kinases, phosphatases, guanine nucleotide exchange factor, GTPases, histone modifying enzymes, chromatin modifiers and ubiquitin modifiers.

Gold standards

We identified several stress-regulated pathways or stress-related functions for each of the seven stresses from literature (Supplementary Table S1). Based on this information, we further identified stress-specific gene ontology (GO) terms for each stress and then defined the gold standards of positives from these stress-specific GO terms (Supplementary Table S1). The rest of genes are considered as gold standards of negatives that are not responding to any specific stress (see Supplementary Material S1 for details).

Node scoring based on YeastNet

In order to identify stress-specific regulators, we first calculated the stress-specific node scores of all the yeast open reading frames (ORFs). To do this, we transformed functional linkage scores computed by integrating heterogeneous data (YeastNet) (18) into node scores based on *a priori* knowledge (Supplementary Table S1). The YeastNet score is more useful to infer a new biological insight than a single dataset due to its integrated nature. The functional linkage data in YeastNet provides us with the information about a network composed of genes (nodes) and the functional similarity (link) of each pair of genes. The link strength indicates the degree of functional similarity between genes. We have assigned the sum of link strengths to the ORF where only those links having gold standards of positives with the ORF are considered. Thus, an ORF having more strong functional linkages with gold standards of positives will have a larger stress-specific score.

The YeastNet is a functional similarity network which is an undirected graph. So, it does not include the information on molecular regulatory interactions. As our major goal is to understand how a cell processes information through complex molecular interactions, we have constructed a directed global regulatory network of yeast separately as described in the previous subsection. Note that we did not use YeastNet to construct our regulatory network of yeast, but only utilized it to calculate the node scores for each stress.

Log likelihood ratio

We have employed a log likelihood ratio to normalize the various stress-specific scores [mRNA expression data, growth fitness defect score data and YeastNet (18) based score data] and to evaluate the sensitivity and specificity. Log likelihood ratio, L is defined as follows:

$$L = \ln \left(\frac{P(f|positive)}{P(f|negative)} \right)$$

where f , *positive* and *negative* indicate the target dataset, gold standards of positives and gold standards of negatives, respectively. If the log likelihood ratio of a gene is larger (smaller) than zero, the gene can be considered active (inactive, respectively) given the specific stress condition. L is computed from contingency tables by binning the dataset values into N intervals (19). So, scores of the likelihood ratio can strongly depend on how the score interval is divided. To avoid this problem, we calculated the average likelihood ratio over 100 times of randomly divided score intervals (see Supplementary Material S1 for details). Here, we set N to 5 as it showed a smooth likelihood ratio distribution while maintaining the data specificity (data not shown).

For overall assessment of the log likelihood ratio, we have divided the gold standards into a training set and a test set. We conducted the 7-fold cross-validation (19). Then, we evaluated the sensitivity and specificity of the log likelihood ratio. We have also computed the sensitivity and specificity of mRNA expression data and growth fitness defect score data (Supplementary Figure S1). We found that the log likelihood ratio based on YeastNet is much more reliable than other log likelihood ratio based on mRNA or growth fitness defect score data (Supplementary Figure S1). This indicates that our stress-specific score faithfully reflects the stress-specific responses. In total, we obtained 6243 ORF specific log likelihood ratios as node scores for each stress condition (adenine dropout, DNA damage, glycerol, H₂O₂, heat shock, NaCl and sorbitol treatment).

MWCS problem formulation and optimization

Given a connected and undirected node-weighted graph $G = (V, E, w)$ with nodes V , edges E and weights w , the MWCS problem is defined as finding a connected subgraph $T = (V_T, E_T)$ of G , $V_T \in V$, $E_T \in E$ that maximizes the score $w(T) = \sum_{v \in V_T} w(v)$ (20). In case both positive and negative node weights are present, solving the MWCS problem is nontrivial. It is known as an NP-complete problem, but a mathematical programming-based algorithm to search for an optimal solution has been proposed (20,21). Therefore, we can obtain an optimal sub-network by solving the MWCS problem using this algorithm. We obtained seven stress-specific regulatory networks (SRNs) based on seven stress-specific regulator scores and a global regulatory network ($N = 523$, $E = 2093$) where all nodes are regulatory molecules and links are directed regulatory interactions. We employed the algorithm provided by (21), and IBM ILOG CPLEX callable library version 12.1 was used for programming.

This approach provides us with a global optimal solution and therefore it is advantageous compared to other heuristic approaches that can only identify non-unique local optimal solutions (22,23). For instance, if we employ heuristic approaches, we cannot assess whether the obtained solutions are globally optimal. Moreover, if a poor solution is obtained, we cannot determine whether it is caused by optimality gap, parameter setting, or score function (20). On the other hand, by

solving the MWCS problem, we can easily evaluate whether the scoring function is adequate for obtaining biologically relevant solution.

Identification of a CRM against various noise levels

It is hard to determine the most biologically relevant SRN since every experimental measurement contains a certain level of inherent noise. To identify SRNs that are irrelevant of different noise levels, we added a random noise following the normal distribution $\sim N(0, \sigma^2)$ to the score of each node where σ denotes a noise level ($\sigma = 0.30$ in Figure 2). Next, we computed SRNs by formulating the MWCS problem. Finally, we found out the CRM from the identified seven SRNs (Figure 2). The CRM is the intersection of both nodes and edges of seven SRNs. The whole procedure is described in Supplementary Figure S2 in detail. We examined whether the gold standards of positives for each stress are overlapped and found that none of them overlaps with each other. This means that the genes in the CRM we defined are not simply originated from the overlapping of gold standards of positives.

We defined the regulators in the CRM as CRs (core regulators) and the regulators not in the CRM as NCRs (non-core regulators). This classification has been done by conservative selection of only consistent results over 100 times repeated simulations. Then we further examined whether all the topological and genetic properties of the CRM are still consistent against different CRs/NCRs classifications. For this purpose, we carried out the same statistical tests for different CRs/NCRs classifications under both node and link perturbations. It turns out that most of the properties are consistently significant (Supplementary Tables S2–S7). Thus, we confirm that the properties of the CRM do not depend on particular CRs/NCRs classifications.

Computation of the hierarchical order and hierarchy destruction score

Hierarchical order of each node is computed based on the global regulatory network using the vertex sort algorithm (24). If a node does not have a specific hierarchical order, we assigned a median hierarchical order to it. Hierarchy destruction score of node X (D_X) is defined as follows:

$$D_X = \sum_{i=1}^N |H_{O,i} - H_{X,i}|$$

where $H_{O,i}$, $H_{X,i}$ and N denote the hierarchical order of node i in the original network, that in the network after removing node X , and the number of nodes in the network, respectively. Removed and isolated nodes are not considered for computation of D_X (see Supplementary Figure S3 for details). Nodes which are included in the hierarchical backbone are expected to have larger D_X .

Genetic measurements

Evolutionary rate indicates the rank of the protein evolutionary rates (25) where a low (high) rank means a low

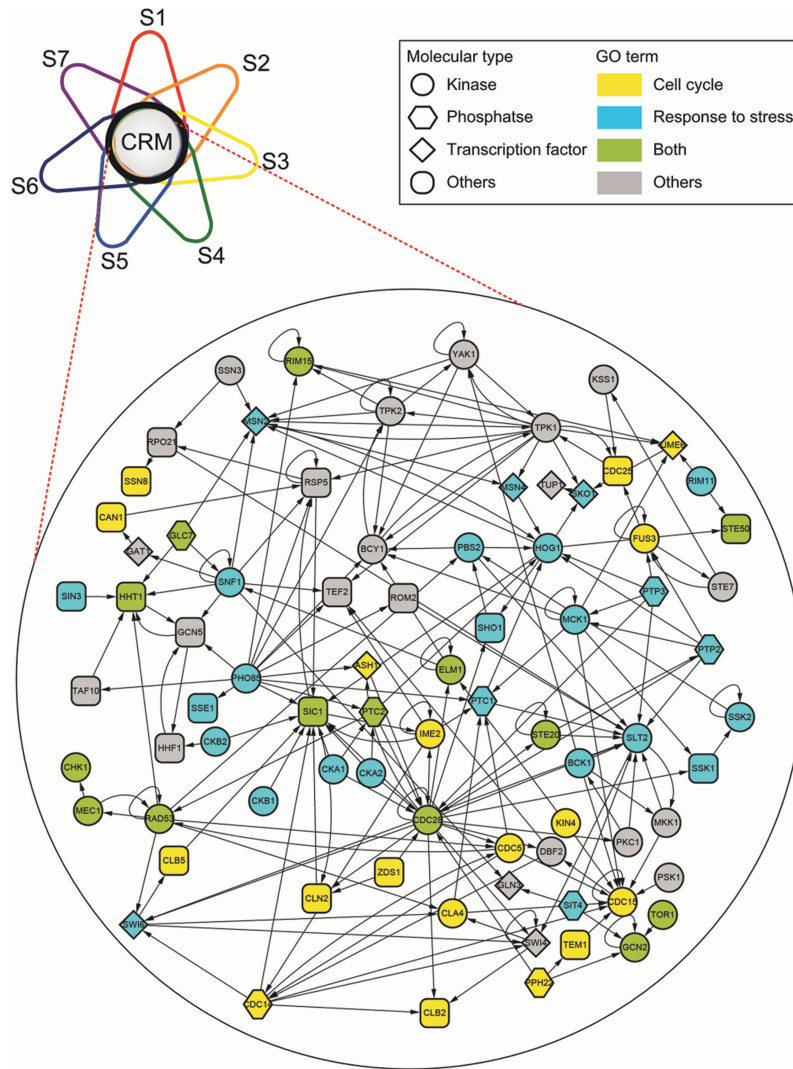


Figure 2. The CRM identified from the global regulatory network of yeast. S1–7 indicate the seven SRNs. The CRM is determined by the intersection of seven SRNs. Nodes are named according to *Saccharomyces Genome Database* nomenclature (73). The CRM includes 31 ‘Cell cycle (GO:0007049)’ genes and 39 ‘Response to stress (GO:0033554)’ genes. ‘Both’ indicates those genes having both GO terms. The network is created by Cytoscape (74).

(high, respectively) evolutionary rate. Interstrain variance was calculated as the variance of gene expressions across different yeast *Saccharomyces cerevisiae* strains with and without stress (26–28). From these data, we excluded clinical and laboratory strains because they have very different phenotypic outcomes among yeast strains and thus can be considered as phenotypic extremes (28,29). From this regard, we used the gene-expression differences with respect to YPS163 (neither clinical nor laboratory strain) in (27). Synthetic lethal pairs were obtained from (30) and only the dataset with stringent cutoff was considered. We computed the growth fitness defect as the average ratio of the mean control intensity to the chemical treatment intensity of yeast strains with homozygous gene deletions (31). Each gene is assigned with a growth fitness defect score which indicates how slow the cell growth will be if the gene is deleted. If the score of a gene is large, then the

gene is essential against various stresses. In the computation, we excluded essential genes since their growth rates cannot be measured.

RESULTS

Identifying stress-specific regulatory networks in yeast

Since not all regulators respond to each stress, sub-network identification methods (20,22–23) were used to identify connected sets of stress-specific regulatory molecules. We identified seven SRNs by formulating an MWCS problem based on the global regulatory network and stress-specific regulator scores (see ‘Materials and Methods’ section for details). The SRNs we identified are mathematically optimal, but they might not be biologically optimal since all experimental measurements

employed here to determine the stress-specific regulator scores could contain inherent noise and this might have induced a certain level of bias. To exclude such a bias, we extracted only those SRNs that are not affected by different noise levels (see ‘Materials and Methods’ section for details). The seven stresses considered in this study are adenine dropout, DNA damage, glycerol, H₂O₂, heat shock, NaCl and sorbitol treatment.

Validation of the CRM based on biological evidences

From the intersection of the SRNs, we identified a CRM (Figure 2) composed of 79 regulators and 210 regulatory interactions (isolated regulators are excluded). In order to validate our finding, we have investigated the elements of the CRM and examined whether they coincide with the established knowledge. We found that the CRM includes the most well-known general stress responsive regulators, MSN2 and MSN4. We have identified a set of genes related to the MAPK-signaling pathway (e.g. FUS3, HOG1 and SLT2) which is known to be central in response to oxidative, osmotic and chemical stresses and in diverse functions such as osmoadaptation, filamentous growth and cell-cycle control (32–36). We have also identified those genes related to the cAMP/PKA pathways (e.g. BCY1, TPK1 and TPK2) that are primarily associated with the response to environmental nutrient or growth factor signaling and with the regulation of cell-cycle, growth and stress responses (10,37,38). We have further identified those genes related to the AMPK-signaling pathway (e.g. SNF1 and ELM1) which controls energy homeostasis and various ESRs (39). Other identifications include 31 ‘regulation of cell cycle (GO:0051726)’ genes ($P = 0.001$) and 39 ‘response to stress (GO:0033554)’ genes ($P = 0.001$) which were significantly enriched in the CRM (Figure 2 and Supplementary Table S2). In particular, the cell-cycle machinery was found closely linked to the aforementioned general stress-regulating pathways. We note that yeast is known to control the cell cycle for energy balance when it encounters environmental stresses (40).

If a cell encounters an environmental stress, it could gain tolerance against subsequent stresses of not only the same type but also different ones (41). This phenomenon called ‘cross-protection’ leads us to infer that a cell might have a common strategy against various types of environmental stresses. To examine this, we analyzed whether the target genes of the CRM have specific GO terms. We found that most of the target genes have ‘cell cycle (GO:0007049)’ ($P = 6E-11$) and ‘negative regulation of biological process (GO:0048519)’ ($P = 5E-10$) (Supplementary Table S3). This implicates that, when a cell perceives a stress signal, it would first shut down all the processes including cell-cycle progression which is the most crucial event of a living organism. This result has in common with the fact that one of the main responses against various stresses is the cell-cycle arrest (41). So, we confirm that our findings are well supported from previous evidenced knowledge.

Regulators in the CRM form a hierarchical backbone of the global regulatory network

What is the characteristic feature of the structural organization of CRM in the global regulatory network of yeast? We investigated the topological basis of CRM as it is closely related to transmitting and processing of information through multiple molecular interactions. Hierarchy is one of the crucial topological characteristics of directed networks since the hierarchical order is known to be strongly related to many biological properties (24,42). We have constructed the hierarchical structure of the global regulatory network by using the ‘vertex sort algorithm’ (24). We identified first the strongly connected components (SCC) defined as a set of nodes connected through feedback regulation. Next, we collapsed each SCC into a single node to transform the complex global regulatory network into a directed acyclic graph (Figure 3A). Then we assigned the hierarchical order to each node of the network. We classified the nodes into five groups: Middle, Top–middle, Top, Bottom–middle and Bottom defined, respectively, as a giant SCC, nodes that only regulate the SCC, the remaining nodes which control other nodes, nodes which are only regulated by the SCC, and the rest of the nodes which are regulated by other nodes (Figure 3A). From this transformed network, we identified that the global regulatory network has a bow-tie structure with the Middle group as a knot (43). In this structure, the knot can integrate various inputs and regulate multiple outputs, so the Middle group is a central regulator of yeast. Intriguingly, we found that most of genes in the CRM are located in the Middle (Figure 3B) and this is very significant compared to random samples (Figure 3C). Therefore, we conclude that the CRM locates in the center of the global regulatory network, which leads us to infer that the CRM might tune the relationship between input (=environmental stress) and output (=cellular response) as a unified information processing unit of the yeast regulatory system.

To unravel the topological importance of each node, we computed the hierarchy destruction score defined as how much each node perturbs the original hierarchical order when it is removed (see ‘Materials and Methods’ section). We found that nodes in the CRM (CRs) are more structurally crucial than nodes outside the CRM (NCRs). This suggests that the CRM constitutes a hierarchical backbone of the yeast regulatory system (Figure 3D). In addition, we investigated the enrichment of feedback and feedforward network motifs in the CRM as they are important basic building blocks in constructing SCC and multi-layered hierarchical order, respectively, and therefore are crucial in determining the hierarchy of a network. We found that the CRM contains a significant number of these network motifs compared to random samples (Figure 3E, F, Supplementary Tables S4 and S5), which also supports that the CRM is a structural core. Taken together, we infer that the regulatory system of yeast might have been designed to let all the environmental signals flow into the CRM, a core decision processor, such that common stress responses can be induced and utilized.

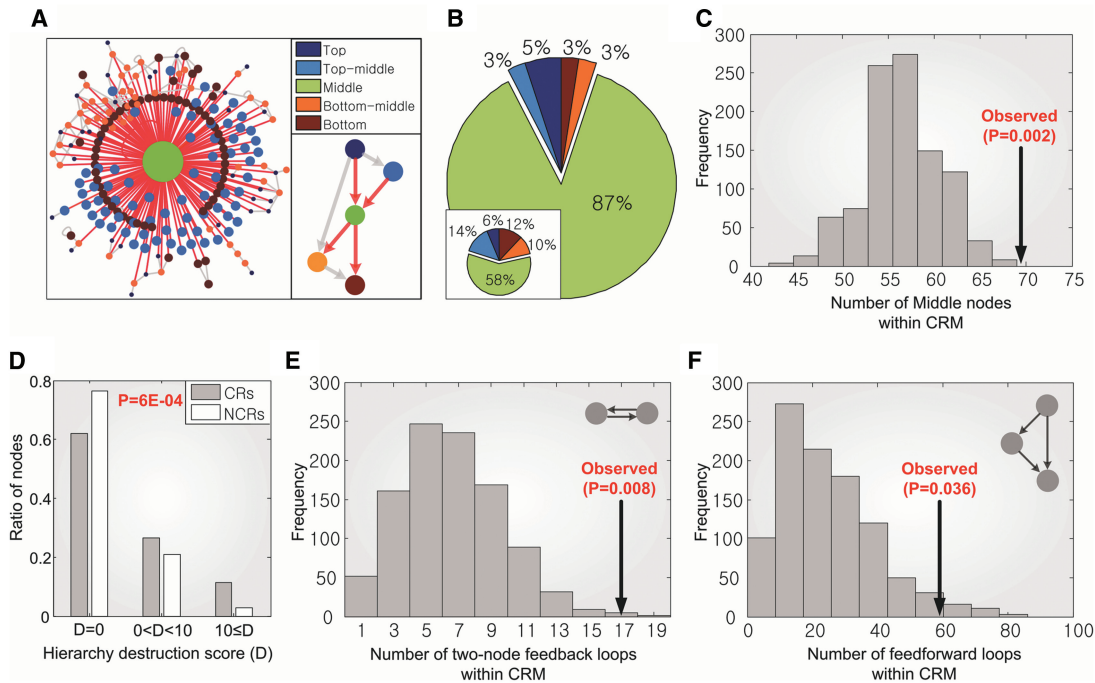


Figure 3. Topological properties of the CRM. (A) A schematic view of the global regulatory network of yeast. All the SCC are reduced to single nodes. Node size is proportional to the number of genes included in each group. Red and grey links denote the regulation of the genes in the Middle group and those in the others, respectively. (B) The ratio of nodes in the CRM for each group (the inlet shows those of the global regulatory network of yeast). (C) The number of Middle nodes in the CRM (Observed) and randomly selected sub-networks (histogram). (D) Hierarchy destruction score of CRs and NCRs. (E), (F) the numbers of two-node feedback loops and feedforward loops within the CRM and those in randomly selected sub-networks, respectively. In (C), (E) and (F), empirical P -values were computed using 1000 random samples. In (D), P -value was computed using Wilcoxon's rank-sum test.

CRM is well conserved, robust against genetic variations and crucial for cell growth

We compared various genetic properties (see 'Materials and Methods' section) between CRs and NCRs to find out functional characteristics of the CRM. We assessed whether CRs or NCRs are evolutionarily more conserved by comparing the evolutionary rate of each group, and found that CRs tend to be more conserved than NCRs (Figure 4A). In addition, we found that CRs have less variation in their gene expressions than NCRs across different wine yeast strains (Figure 4B). We further compared the gene-expression variance of different yeast strains and found the same result (Supplementary Figure S4) (27). However, we noted that all these expression profiles were measured without stress. So, we have further analyzed the gene-expression profiles of various yeast strains under stress (28). We found again that the expressions of CRs are significantly less variable than those of NCRs after adaptation to new environments (Supplementary Figure S5). Therefore, we conclude that a cell can generate robust steady-state responses through CRs under stress. These results indicate that CRs might have undergone stronger selective pressure to maintain their gene-expression levels during evolution. Overall, we conclude that CRs are evolutionarily well conserved in terms of molecular function.

We further investigated whether the synthetic lethal pairs are enriched in the CRM and found that the ratio of synthetic lethal pairs within the CRM is >2 -fold of that

within the non-CRM (Figure 4C). Finally, we found that deletion of CRs leads to slower cell growth than the deletion of NCRs upon exposure to various stresses (Figure 4D). In order to further examine whether CRs are crucial against multiple stress, we have investigated whether MDR (multiple drug resistance) genes are enriched in CRs or NCRs. MDR genes are defined as essential genes to grow in the presence of multiple drugs (31). We found that MDR genes are significantly enriched in CRs ($P = 0.034$, Hypergeometric test), but are not overlapped with NCRs (Supplementary Figure S6). All the above results are consistent with various CRs/NCRs classifications (Supplementary Tables S6 and S7). Therefore, we can conclude that deletion of CRs leads to slower cell growth than the deletion of NCRs upon exposure to various stresses. These are the direct evidences showing that knockout mutants of CRs result in greatly increased sensitivity to a wide range of stress conditions. It also leads us to infer that CRs are more critical to cell growth than NCRs. Taken together, the genes in the CRM are not only evolutionarily conserved, but also they are central regulators of cellular function.

DISCUSSION

In this article, we have identified the CRM of *S. cerevisiae* and uncovered its topological and genetic properties. Biological data used for determination of gene function and construction of a molecular interaction network

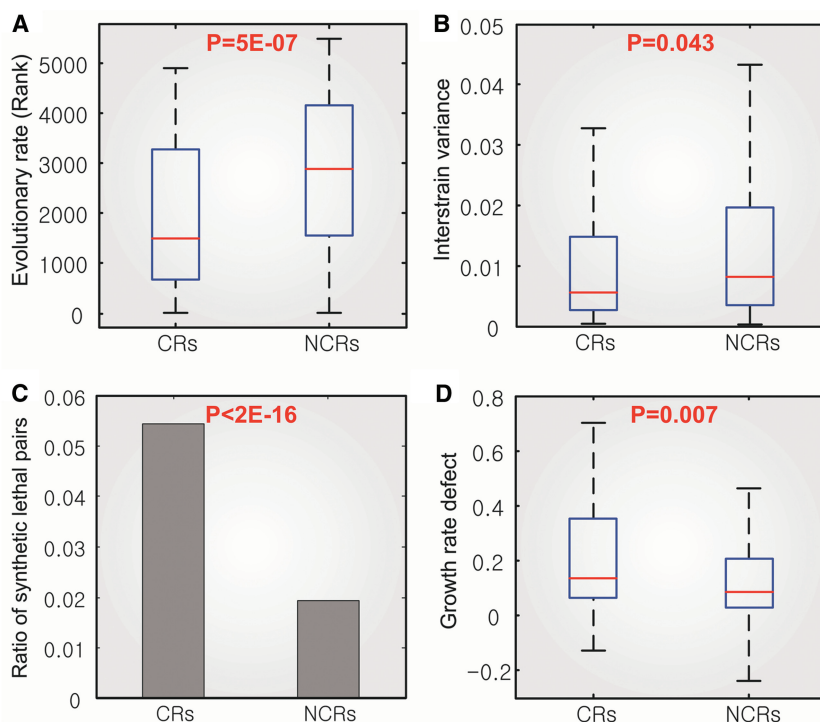


Figure 4. Genetic properties of the CRM. In (A), (B) and (D), box plots show a five-number summary of data: the 25th, 50th, 75th percentiles of the samples, and the two most extreme values within 1.5 times of the interquartile range (distance between 25th and 75th percentiles). *P*-values were computed using Wilcoxon's rank-sum test. In (C), *P*-value was computed using Fisher's exact test.

usually contain various inherent noises. So, in this study, we intentionally added random perturbations to the node scoring as well as to the construction of a network and identified only those properties consistently held irrespective of different noise levels (see Supplementary Tables S2–S7 for details). Therefore, we confirm that our results do not depend on such random biases that might be introduced during the integration of uncertain biological data.

We wondered whether the genes in the CRM might be overlapped with the ESR genes, but such an overlapping was very little. This is because most ESR genes are not regulatory molecules. In addition, we noted that the functional role of ESR genes and that of the genes in the CRM are considerably different. In contrast with the genes in the CRM, cell-cycle genes were not identified in the ESR genes which are mostly involved with the transcriptional (e.g. RNA processing), translational (e.g. ribosome biogenesis), or other cellular processes. ESR genes perform cellular protection by preserving energy, balancing internal osmolarity and stabilizing biomolecular structure (2). However, these functions were not represented in the CRM. How can we interpret such differential functions? For instance, we can think that cell-cycle control is the most crucial event regulated by the CRM and the ESR gene expressions are the resulting responses. In other words, the genes in the CRM are involved in the stress-induced cell-cycle arrest and the ESR gene expressions are the subsequent processes followed by the growth arrest. Brauer *et al.* also suggested that many ESR genes

may not be reacting to stress but responding to a growth reduction (44). In summary, our network-based approach provides a complementary insight into the common stress responses of yeast compared to previous studies on genome-wide gene expressions.

What might be the reason why yeast cells have developed the CRM? There is no doubt that the most prominent activity of yeast as a unicellular organism is to determine whether it should run a cell-cycle or not in consideration of all environmental stresses. For a wild yeast, to survive under full of environmental stresses, there might be a necessity to develop some central information processing unit like the CRM that is to be frequently and robustly utilized. We note that the expression variation of those genes in the CRs is lower than that in the NCRs. From this, we speculate that yeast cells can have constant transmission of various kinds of environmental information as the molecular level of CRs is kept invariant across various yeast strains. This also suggests that the CRM might be designed to maintain the robustness of cellular information processing. Furthermore, we infer that the CRM is the most important information processing unit of a cell achieved evolutionarily since the CRM is both topologically and genetically essential parts of the yeast regulatory network. Taken together, we summarize that a cell can efficiently interpret numerous external inputs using the CRM, a unified information processor, and thereby the cell can choose an appropriate decision from the most important and problematic choices of yeast: to run a cell cycle, or not.

In order to identify the CRM for a specific species, we need three kinds of information: a genome-wide functional association network, gene lists that are known to respond to specific stresses, and a molecular interaction network. Since the first (e.g. SPRING (45)) and third (e.g. Bio-GRID (46)) are well-established and widely available in general, our method can be applied to many other species with different physiological conditions.

A similar concept of the CRM has also been proposed in developmental regulatory networks (47), in which the developmental role and regulatory interactions of a CRM have been argued to be most impervious to change during evolutionary processes and suggested to affect the planning of major body morphology. Why does a cell have such a regulatory system? Reuse of the CRM in response to various stimuli allows cells to minimize resource utilization while the robustness of the CRM maximizes the efficiency and accuracy of information processing. Thus, the CRM endows a cell with an enhanced efficiency in dealing with numerous external stimuli with a limited number of internal molecular components. We speculate that the CRM might be a ubiquitous organizing principle of various regulatory systems by shaping central responses in an efficient and accurate manner.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online: Supplementary Tables 1–7, Supplementary Figures 1–6, Supplementary Data and Supplementary References [48–72].

ACKNOWLEDGEMENTS

The authors thank Jeong-Rae Kim for his valuable comments on this article. The authors also thank Ivana Ljubić for the provision of software and Namsu An for his helpful technical support.

FUNDING

Funding for open access charge: National Research Foundation of Korea (NRF) grants funded by the Korea Government, the Ministry of Education, Science and Technology (MEST) [2009-0086964 and 2010-0017662]; NRF grants funded by the MEST through the WCU (World Class University) program [R32-2008-000-10218-0].

Conflict of interest statement. None declared.

REFERENCES

- Estruch, F. (2000) Stress-controlled transcription factors, stress-induced genes and stress tolerance in budding yeast. *FEMS Microbiol. Rev.*, **24**, 469–486.
- Gasch, A.P. (2003) The environmental stress response: a common yeast response to diverse environmental stresses. *Top. Curr. Genet.*, **1**, 11–70.
- Gasch, A.P. (2007) Comparative genomics of the environmental stress response in ascomycete fungi. *Yeast*, **24**, 961–976.
- Gasch, A.P., Spellman, P.T., Kao, C.M., Carmel-Harel, O., Eisen, M.B., Storz, G., Botstein, D. and Brown, P.O. (2000) Genomic expression programs in the response of yeast cells to environmental changes. *Mol. Biol. Cell*, **11**, 4241–4257.
- Causton, H.C., Ren, B., Koh, S.S., Harbison, C.T., Kanin, E., Jennings, E.G., Lee, T.I., True, H.L., Lander, E.S. and Young, R.A. (2001) Remodeling of yeast genome expression in response to environmental changes. *Mol. Biol. Cell*, **12**, 323–337.
- Birrell, G.W., Brown, J.A., Wu, H.I., Giaever, G., Chu, A.M., Davis, R.W. and Brown, J.M. (2002) Transcriptional response of *Saccharomyces cerevisiae* to DNA-damaging agents does not identify the genes that protect against these agents. *Proc. Natl. Acad. Sci. USA*, **99**, 8778–8783.
- Schaeffer, H.J. and Weber, M.J. (1999) Mitogen-activated protein kinases: specific messages from ubiquitous messengers. *Mol. Cell. Biol.*, **19**, 2435–2444.
- Thevelein, J.M. and de Winde, J.H. (1999) Novel sensing mechanisms and targets for the cAMP-protein kinase A pathway in the yeast *Saccharomyces cerevisiae*. *Mol. Microbiol.*, **33**, 904–918.
- Ikner, A. and Shiozaki, K. (2005) Yeast signaling pathways in the oxidative stress response. *Mutat. Res.*, **569**, 13–27.
- Zurita-Martinez, S.A. and Cardenas, M.E. (2005) Tor and cyclic AMP-protein kinase A: two parallel pathways regulating expression of genes required for cell growth. *Eukaryot. Cell*, **4**, 63–71.
- Breitkreutz, B.J., Stark, C., Reguly, T., Boucher, L., Breitkreutz, A., Livstone, M., Oughtred, R., Lackner, D.H., Bahler, J., Wood, V. *et al.* (2008) The BioGRID Interaction Database: 2008 update. *Nucleic Acids Res.*, **36**, D637–D640.
- Fiedler, D., Braberg, H., Mehta, M., Chechik, G., Cagney, G., Mukherjee, P., Silva, A.C., Shales, M., Collins, S.R., van Wageningen, S. *et al.* (2009) Functional organization of the *S. cerevisiae* phosphorylation network. *Cell*, **136**, 952–963.
- Henrik Dohlman, J.E.S. (2009) Pheromone signaling pathways in yeast. *Sci. Signal* (Connections Map in the Database of Cell Signaling), http://stke.sciencemag.org/cgi/cm/stkecm;CMP_13999 (21 June 2012, date last accessed).
- Jeremy Thorner, D.M.T. and Garrenton, L.S. (2009) Filamentous growth pathway in yeast. *Sci. Signal* (Connections Map in the Database of Cell Signaling), http://stke.sciencemag.org/cgi/cm/stkecm;CMP_14554 (21 June 2012, date last accessed).
- Jeremy Thorner, P.J.W. and Ballon, D.R. (2009) High osmolarity glycerol (HOG) pathway in yeast. *Sci. Signal* (Connections Map in the Database of Cell Signaling), http://stke.sciencemag.org/cgi/cm/stkecm;CMP_14620 (21 June 2012, date last accessed).
- Kaizu, K., Ghosh, S., Matsuoka, Y., Moriya, H., Shimizu-Yoshida, Y. and Kitano, H. (2010) A comprehensive molecular interaction map of the budding yeast cell cycle. *Mol. Syst. Biol.*, **6**, 415.
- MacIsaac, K.D., Wang, T., Gordon, D.B., Gifford, D.K., Stormo, G.D. and Fraenkel, E. (2006) An improved map of conserved regulatory sites for *Saccharomyces cerevisiae*. *BMC Bioinformatics*, **7**, 113.
- Lee, I., Li, Z. and Marcotte, E.M. (2007) An improved, bias-reduced probabilistic functional gene network of baker's yeast, *Saccharomyces cerevisiae*. *PLoS One*, **2**, e988.
- Jansen, R., Yu, H., Greenbaum, D., Kluger, Y., Krogan, N.J., Chung, S., Emili, A., Snyder, M., Greenblatt, J.F. and Gerstein, M. (2003) A Bayesian networks approach for predicting protein-protein interactions from genomic data. *Science*, **302**, 449–453.
- Dittrich, M.T., Klau, G.W., Rosenwald, A., Dandekar, T. and Muller, T. (2008) Identifying functional modules in protein-protein interaction networks: an integrated exact approach. *Bioinformatics*, **24**, i223–i231.
- Ljubić, I., Weiskircher, R., Pferschy, U., Klau, G.W., Mutzel, P. and Fischetti, M. (2006) An algorithmic framework for the exact solution of the prize-collecting Steiner tree problem. *Math. Program.*, **105**, 427–449.
- Ideker, T., Ozier, O., Schwikowski, B. and Siegel, A.F. (2002) Discovering regulatory and signalling circuits in molecular interaction networks. *Bioinformatics*, **18**(Suppl. 1), S233–S240.
- Guo, Z., Wang, L., Li, Y., Gong, X., Yao, C., Ma, W., Wang, D., Zhu, J., Zhang, M., Yang, D. *et al.* (2007) Edge-based scoring and

- searching method for identifying condition-responsive protein-protein interaction sub-network. *Bioinformatics*, **23**, 2121–2128.
24. Jothi,R., Balaji,S., Wuster,A., Grochow,J.A., Gspomer,J., Przytycka,T.M., Aravind,L. and Babu,M.M. (2009) Genomic analysis reveals a tight link between transcription factor dynamics and regulatory network architecture. *Mol. Syst. Biol.*, **5**, 294.
 25. Xia,Y., Franzosa,E.A. and Gerstein,M.B. (2009) Integrated assessment of genomic correlates of protein evolutionary rate. *PLoS Comput. Biol.*, **5**, e1000413.
 26. Townsend,J.P., Cavalieri,D. and Hartl,D.L. (2003) Population genetic variation in genome-wide gene expression. *Mol. Biol. Evol.*, **20**, 955–963.
 27. Kvittek,D.J., Will,J.L. and Gasch,A.P. (2008) Variations in stress sensitivity and genomic expression in diverse *S. cerevisiae* isolates. *PLoS Genet.*, **4**, e1000223.
 28. Carreto,L., Eiriz,M.F., Domingues,I., Schuller,D., Moura,G.R. and Santos,M.A. (2011) Expression variability of co-regulated genes differentiates *Saccharomyces cerevisiae* strains. *BMC Genomics*, **12**, 201.
 29. Warringer,J., Zorgo,E., Cubillos,F.A., Zia,A., Gjuvsland,A., Simpson,J.T., Forsmark,A., Durbin,R., Omholt,S.W., Louis,E.J. *et al.* (2011) Trait variation in yeast is defined by population history. *PLoS Genet.*, **7**, e1002111.
 30. Costanzo,M., Baryshnikova,A., Bellay,J., Kim,Y., Spear,E.D., Sevier,C.S., Ding,H., Koh,J.L., Toufighi,K., Mostafavi,S. *et al.* (2010) The genetic landscape of a cell. *Science*, **327**, 425–431.
 31. Hillenmeyer,M.E., Fung,E., Wildenhain,J., Pierce,S.E., Hoon,S., Lee,W., Proctor,M., St Onge,R.P., Tyers,M., Koller,D. *et al.* (2008) The chemical genomic portrait of yeast: uncovering a phenotype for all genes. *Science*, **320**, 362–365.
 32. Bermejo,C., Rodriguez,E., Garcia,R., Rodriguez-Pena,J.M., Rodriguez de la Concepcion,M.L., Rivas,C., Arias,P., Nombela,C., Posas,F. and Arroyo,J. (2008) The sequential activation of the yeast HOG and SLT2 pathways is required for cell survival to cell wall stress. *Mol. Biol. Cell*, **19**, 1113–1124.
 33. Escote,X., Zapater,M., Clotet,J. and Posas,F. (2004) Hog1 mediates cell-cycle arrest in G1 phase by the dual targeting of Sic1. *Nat. Cell Biol.*, **6**, 997–1002.
 34. Haghazari,E. and Heyer,W.D. (2004) The Hog1 MAP kinase pathway and the Mec1 DNA damage checkpoint pathway independently control the cellular responses to hydrogen peroxide. *DNA Repair*, **3**, 769–776.
 35. Proft,M., Mas,G., de Nadal,E., Vendrell,A., Noriega,N., Struhl,K. and Posas,F. (2006) The stress-activated Hog1 kinase is a selective transcriptional elongation factor for genes responding to osmotic stress. *Mol. Cell*, **23**, 241–250.
 36. Schwartz,M.A. and Madhani,H.D. (2004) Principles of MAP kinase signaling specificity in *Saccharomyces cerevisiae*. *Annu. Rev. Genet.*, **38**, 725–748.
 37. Gimeno,C.J., Ljungdahl,P.O., Styles,C.A. and Fink,G.R. (1992) Unipolar cell divisions in the yeast *S. cerevisiae* lead to filamentous growth: regulation by starvation and RAS. *Cell*, **68**, 1077–1090.
 38. Rohde,J.R., Bastidas,R., Puria,R. and Cardenas,M.E. (2008) Nutritional control via Tor signaling in *Saccharomyces cerevisiae*. *Curr. Opin. Microbiol.*, **11**, 153–160.
 39. Hedbacker,K. and Carlson,M. (2008) SNF1/AMPK pathways in yeast. *Front. Biosci.*, **13**, 2408–2420.
 40. Lai,L.C., Kissinger,M.T., Burke,P.V. and Kwast,K.E. (2008) Comparison of the transcriptomic “stress response” evoked by antimycin A and oxygen deprivation in *Saccharomyces cerevisiae*. *BMC Genomics*, **9**, 627.
 41. Lu,C., Brauer,M.J. and Botstein,D. (2009) Slow growth induces heat-shock resistance in normal and respiratory-deficient yeast. *Mol. Biol. Cell*, **20**, 891–903.
 42. Bhardwaj,N., Kim,P.M. and Gerstein,M.B. (2010) Rewiring of transcriptional regulatory networks: hierarchy, rather than connectivity, better reflects the importance of regulators. *Sci. Signal.*, **3**, ra79.
 43. Kitano,H. (2004) Biological robustness. *Nat. Rev. Genet.*, **5**, 826–837.
 44. Brauer,M.J., Huttenhower,C., Airoidi,E.M., Rosenstein,R., Matese,J.C., Gresham,D., Boer,V.M., Troyanskaya,O.G. and Botstein,D. (2008) Coordination of growth rate, cell cycle, stress response, and metabolic activity in yeast. *Mol. Biol. Cell*, **19**, 352–367.
 45. Szklarczyk,D., Franceschini,A., Kuhn,M., Simonovic,M., Roth,A., Minguéz,P., Doerks,T., Stark,M., Müller,J., Bork,P. *et al.* (2011) The STRING database in 2011: functional interaction networks of proteins, globally integrated and scored. *Nucleic Acids Res.*, **39**, D561–D568.
 46. Stark,C., Breitkreutz,B.J., Chatr-Aryamontri,A., Boucher,L., Oughtred,R., Livstone,M.S., Nixon,J., Van Auken,K., Wang,X., Shi,X. *et al.* (2011) The BioGRID Interaction Database: 2011 update. *Nucleic Acids Res.*, **39**, D698–D704.
 47. Davidson,E.H. and Erwin,D.H. (2006) Gene regulatory networks and the evolution of animal body plans. *Science*, **311**, 796–800.
 48. Barrett,T., Troup,D.B., Wilhite,S.E., Ledoux,P., Rudnev,D., Evangelista,C., Kim,I.F., Soboleva,A., Tomashevsky,M. and Edgar,R. (2007) NCBI GEO: mining tens of millions of expression profiles—database and tools update. *Nucleic Acids Res.*, **35**, D760–D765.
 49. Berdichevsky,A. and Guarente,L. (2006) A stress response pathway involving sirTuins, forkheads and 14-3-3 proteins. *Cell Cycle*, **5**, 2588–2591.
 50. Delaunay,A., Pflieger,D., Barrault,M.B., Vinh,J. and Toledano,M.B. (2002) A thiol peroxidase is an H₂O₂ receptor and redox-transducer in gene activation. *Cell*, **111**, 471–481.
 51. Demeter,J., Beauheim,C., Gollub,J., Hernandez-Boussard,T., Jin,H., Maier,D., Matese,J.C., Nitzberg,M., Wymore,F., Zachariah,Z.K. *et al.* (2007) The Stanford Microarray Database: implementation of new analysis tools and open source release of software. *Nucleic Acids Res.*, **35**, D766–D770.
 52. Erlich,R.L., Fry,R.C., Begley,T.J., Dae,D.L., Lahue,R.S. and Samson,L.D. (2008) Anc1, a protein associated with multiple transcription complexes, is involved in postreplication repair pathway in *S. cerevisiae*. *PLoS One*, **3**, e31717.
 53. Fry,R.C., DeMott,M.S., Cosgrove,J.P., Begley,T.J., Samson,L.D. and Dedon,P.C. (2006) The DNA-damage signature in *Saccharomyces cerevisiae* is associated with single-strand breaks in DNA. *BMC Genomics*, **7**, 313.
 54. Grauslund,M., Lopes,J.M. and Ronnow,B. (1999) Expression of GUT1, which encodes glycerol kinase in *Saccharomyces cerevisiae*, is controlled by the positive regulators Adr1p, Ino2p and Ino4p and the negative regulator Opi1p in a carbon source-dependent fashion. *Nucleic Acids Res.*, **27**, 4391–4398.
 55. Hao,N., Behar,M., Parnell,S.C., Torres,M.P., Borchers,C.H., Elston,T.C. and Dohlman,H.G. (2007) A systems-biology analysis of feedback inhibition in the Sho1 osmotic-stress-response pathway. *Curr. Biol.*, **17**, 659–667.
 56. Kuranda,K., Leberre,V., Sokol,S., Palamarczyk,G. and Francois,J. (2006) Investigating the caffeine effects in the yeast *Saccharomyces cerevisiae* brings new insights into the connection between TOR, PKC and Ras/cAMP signalling pathways. *Mol. Microbiol.*, **61**, 1147–1166.
 57. Lee,I., Date,S.V., Adai,A.T. and Marcotte,E.M. (2004) A probabilistic functional network of yeast genes. *Science*, **306**, 1555–1558.
 58. Levin,D.E. (2005) Cell wall integrity signaling in *Saccharomyces cerevisiae*. *Microbiol. Mol. Biol. Rev.*, **69**, 262–291.
 59. Lopez-Mirabal,H.R., Winther,J.R. and Kielland-Brandt,M.C. (2008) Oxidant resistance in a yeast mutant deficient in the Sit4 phosphatase. *Curr. Genet.*, **53**, 275–286.
 60. Medicherla,B. and Goldberg,A.L. (2008) Heat shock and oxygen radicals stimulate ubiquitin-dependent degradation mainly of newly synthesized proteins. *J. Cell. Biol.*, **182**, 663–673.
 61. Raychaudhuri,S., Stuart,J.M. and Altman,R.B. (2000) Principal components analysis to summarize microarray experiments: application to sporulation time series. *Pac. Symp. Biocomput.*, 455–466.
 62. Rebora,K., Desmoucelles,C., Borne,F., Pinson,B. and Daignan-Fornier,B. (2001) Yeast AMP pathway genes respond to adenine through regulated synthesis of a metabolic intermediate. *Mol. Cell. Biol.*, **21**, 7901–7912.
 63. Servant,G., Pennetier,C. and Lesage,P. (2008) Remodeling yeast gene transcription by activating the Tyl long terminal repeat

- retrotransposon under severe adenine deficiency. *Mol. Cell. Biol.*, **28**, 5543–5554.
64. Shalem, O., Dahan, O., Levo, M., Martinez, M.R., Furman, I., Segal, E. and Pilpel, Y. (2008) Transient transcriptional responses to stress are generated by opposing effects of mRNA production and degradation. *Mol. Syst. Biol.*, **4**, 223.
65. Shaner, L., Gibney, P.A. and Morano, K.A. (2008) The Hsp110 protein chaperone Sse1 is required for yeast cell wall integrity and morphogenesis. *Curr. Genet.*, **54**, 1–11.
66. Shrivastav, M., De Haro, L.P. and Nickoloff, J.A. (2008) Regulation of DNA double-strand break repair pathway choice. *Cell Res.*, **18**, 134–147.
67. Silva, R.D., Sotoca, R., Johansson, B., Ludovico, P., Sansonetti, F., Silva, M.T., Peinado, J.M. and Corte-Real, M. (2005) Hyperosmotic stress induces metacaspase- and mitochondria-dependent apoptosis in *Saccharomyces cerevisiae*. *Mol. Microbiol.*, **58**, 824–834.
68. Tirosh, I., Weinberger, A., Carmi, M. and Barkai, N. (2006) A genetic signature of interspecies variations in gene expression. *Nat. Genet.*, **38**, 830–834.
69. Toivari, M.H., Salusjarvi, L., Ruohonen, L. and Penttila, M. (2004) Endogenous xylose pathway in *Saccharomyces cerevisiae*. *Appl. Environ. Microbiol.*, **70**, 3681–3686.
70. Winkler, A., Arkind, C., Mattison, C.P., Burkholder, A., Knoche, K. and Ota, I. (2002) Heat stress activates the yeast high-osmolarity glycerol mitogen-activated protein kinase pathway, and protein tyrosine phosphatases are essential under heat stress. *Eukaryot. Cell*, **1**, 163–173.
71. Ye, T., Elbing, K. and Hohmann, S. (2008) The pathway by which the yeast protein kinase Snf1p controls acquisition of sodium tolerance is different from that mediating glucose regulation. *Microbiology*, **154**, 2814–2826.
72. Yoshimoto, H., Saltsman, K., Gasch, A.P., Li, H.X., Ogawa, N., Botstein, D., Brown, P.O. and Cyert, M.S. (2002) Genome-wide analysis of gene expression regulated by the calcineurin/Crz1p signaling pathway in *Saccharomyces cerevisiae*. *J. Biol. Chem.*, **277**, 31079–31088.
73. Hong, E.L., Balakrishnan, R., Dong, Q., Christie, K.R., Park, J., Binkley, G., Costanzo, M.C., Dwight, S.S., Engel, S.R., Fisk, D.G. et al. (2008) Gene Ontology annotations at SGD: new data sources and annotation methods. *Nucleic Acids Res.*, **36**, D577–D581.
74. Smoot, M.E., Ono, K., Ruscheinski, J., Wang, P.L. and Ideker, T. (2011) Cytoscape 2.8: new features for data integration and network visualization. *Bioinformatics*, **27**, 431–432.