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

Dongsan Kim, Man-Sun Kim, 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

Supplementary Material 1

*Corresponding author, E-mail: ckh@kaist.ac.kr, Phone: +82-42-350-4325, Fax: +82-42-350-4310, Web address: http://sbie.kaist.ac.kr

1. Data acquisition

Experimental data

We downloaded seven different mRNA expression profiles from Gene Expression Omnibus (GEO) database (1) and Stanford Microarray Database (SMD) (2). The seven different stresses are Adenine dropout (3), DNA damage (gamma radiation) (4), Glycerol (5), H_2O_2 (6), Heat shock (3), NaCl (7) and Sorbitol treatments (3). We neglected all the ORFs that have at least one missing value in their time-series data. Then, we have converted the time-series data into one dimensional data using principle component analysis (PCA) (8). We also downloaded growth fitness defect score data from (9). The growth fitness score indicates the ratio of the mean control intensity to the chemical (stress) treatment intensity of yeast homozygous deletion (9). Sensitivity and specificity of mRNA and growth fitness defect score data were compared to those of our score (see Computation section and Fig. S1).

Network data

We have constructed the yeast regulatory network from BioGRID (10) (http://www.thebiogrid.org/downloads.php). We downloaded BioGRID Version 3.0.64 and collected the information on molecular interactions based on biochemical activity. Then, we have incorporated the phosphorylation/dephosphorylation network acquired from the literature (11). In the next, we have integrated the transcriptional networks acquired from high-throughput ChIP-chip experiments (12) (http://fraenkel.mit.edu/improved_map/latest_maps.html, orfs_by_factor_p0.001_cons0.txt). Then, we have also integrated the network from Science Signaling database (13-15) (http://stke.sciencemag.org/cgi/collection/specific_pathways), and other molecular regulation based on manual curation (16). Finally, we have constructed the whole regulatory network composed of 9,438 links and 3,170 nodes.

Gold standards

We downloaded Gene Ontology terms from Saccharomyces Genome Database (SGD) (17). We assumed that those ORFs containing only GO: 0003674 (molecular function) or GO: 0008150 (biological process) (or containing no GO term) are functionally ‘unknown’. 5,412 ORFs in total were found to be related to specific biological processes. We also found several stress-regulated pathways or stress-related functions from literature (Table S1). From these, we have identified the gold standards of positives that have stress-specific GO terms (Table S1). The rest of genes are considered as gold standards of negatives that are not responding to any specific stress. The list of gold standards of positives is in Supplementary file 2.

YeastNet score

We can have better inference results about hidden functional characteristics by using multiple genomic data sources than single data (18-19). Because most of the genomic data are noisy and, moreover, each genomic data have its own biological features, Jansen *et al*. showed that protein interactions can be more accurately predicted when mixed multiple genomic data are used than only single experimental measurement is used (18). There are many genome-wide experimental data of yeast, but most experiments are not performed in a specific stress condition. Thus, those genome-wide experimental data cannot be directly used to extract stress-specific information. To resolve this problem, we have employed functional linkage data integrated with multiple genomic data (20). The functional linkage data 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. To calculate the 'activation level' of an ORF in a specific stress condition, 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.

In this study, our goal is to understand how a cell processes information through complex molecular interactions against various stresses. One might imagine that something similar to the CRM can be obtained using the YeastNet, but that is very different from our CRM and the goal of our study cannot be achieved with this since the YeastNet is a functional network and therefore does not include the information on molecular regulatory interactions. This is the reason why we integrated several types of molecular regulatory networks and reconstructed a global regulatory network.

2. Computation

Computing the log likelihood ratio

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. P(f|positive) is the probability with which an ORF included in the gold standards of positives (positive) has a certain feature (f). In practice, we can discretize the continuous feature space (e.g., log-ratio of mRNA expression levels) for this

purpose. In this paper, we divided the log-ratio of mRNA expression levels into five bins. For instance, if there are 100 gold standards of positives (positive) and 10 of them are included in a certain bin, P(f|positive) of the ORFs in this bin will be 0.1. P(f|negative) can be computed similarly. In our study, we have randomly divided the feature space into five bins and computed the probabilities 100 times iteratively. Then, we assigned the average P(f|positive) (or P(f|negative)) value to each ORF. Finally, the log likelihood ratio is calculated using these P(f|positive) and P(f|negative).

Supplementary References

1. 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**, D760-765.
2. 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**, D766-770.
3. 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.
4. 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.
5. Tirosh, I., Weinberger, A., Carmi, M. and Barkai, N. (2006) A genetic signature of interspecies variations in gene expression. *Nat Genet*, **38**, 830-834.
6. 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.
7. 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.
8. 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.
9. 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.

10. 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**, D637-640.
11. 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.
12. 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.
13. Henrik Dohlman, J.E.S. (2009) Pheromone Signaling Pathways in Yeast. *Sci. Signal.*
14. Jeremy Thorner, D.M.T., Lindsay S. Garrenton. (2009) Filamentous Growth Pathway in Yeast. *Sci. Signal.*
15. Jeremy Thorner, P.J.W., Daniel R. Ballon. (2009) High Osmolarity Glycerol (HOG) Pathway in Yeast. *Sci. Signal.*
16. 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.
17. 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**, D577-581.
18. 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.
19. Lee, I., Date, S.V., Adai, A.T. and Marcotte, E.M. (2004) A probabilistic functional network of yeast genes. *Science*, **306**, 1555-1558.
20. 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.
21. Kvitek, 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.
22. 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.
23. 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.

24. Servant, G., Pennetier, C. and Lesage, P. (2008) Remodeling yeast gene transcription by activating the Ty1 long terminal repeat retrotransposon under severe adenine deficiency. *Mol Cell Biol*, **28**, 5543-5554.
25. Erlich, R.L., Fry, R.C., Begley, T.J., Daee, 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**, e3717.
26. 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.
27. 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.
28. Toivari, M.H., Salusjarvi, L., Ruohonen, L. and Penttila, M. (2004) Endogenous xylose pathway in Saccharomyces cerevisiae. *Appl Environ Microbiol*, **70**, 3681-3686.
29. Berdichevsky, A. and Guarente, L. (2006) A stress response pathway involving sirtuins, forkheads and 14-3-3 proteins. *Cell Cycle*, **5**, 2588-2591.
30. 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.
31. 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.
32. 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.
33. 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.
34. 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.
35. 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.
36. 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.
37. 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.

38. Levin, D.E. (2005) Cell wall integrity signaling in Saccharomyces cerevisiae. *Microbiol Mol Biol Rev*, **69**, 262-291.

39. Silva, R.D., Sotoca, R., Johansson, B., Ludovico, P., Sansonetty, 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.

Supplementary Figures

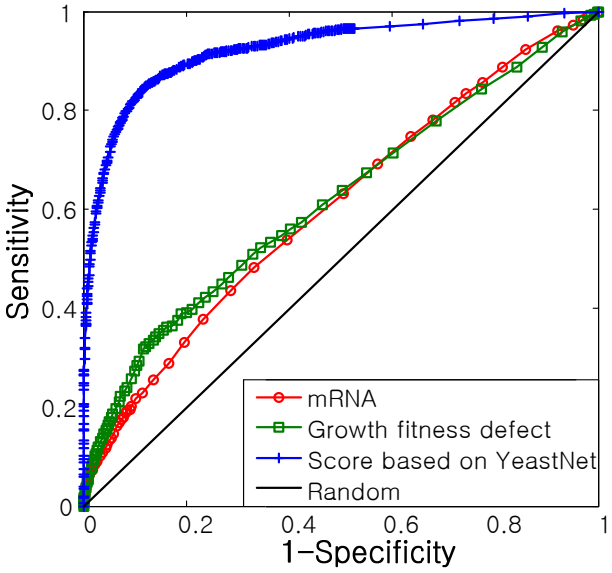


Figure S1. The sensitivity and specificity of three different scores of log likelihood ratio. We plotted ROC curves using the scores of the log likelihood ratio based on each dataset for DNA damage stress. For all the other stress conditions, those based on YeastNet showed always the best performance (data not shown).

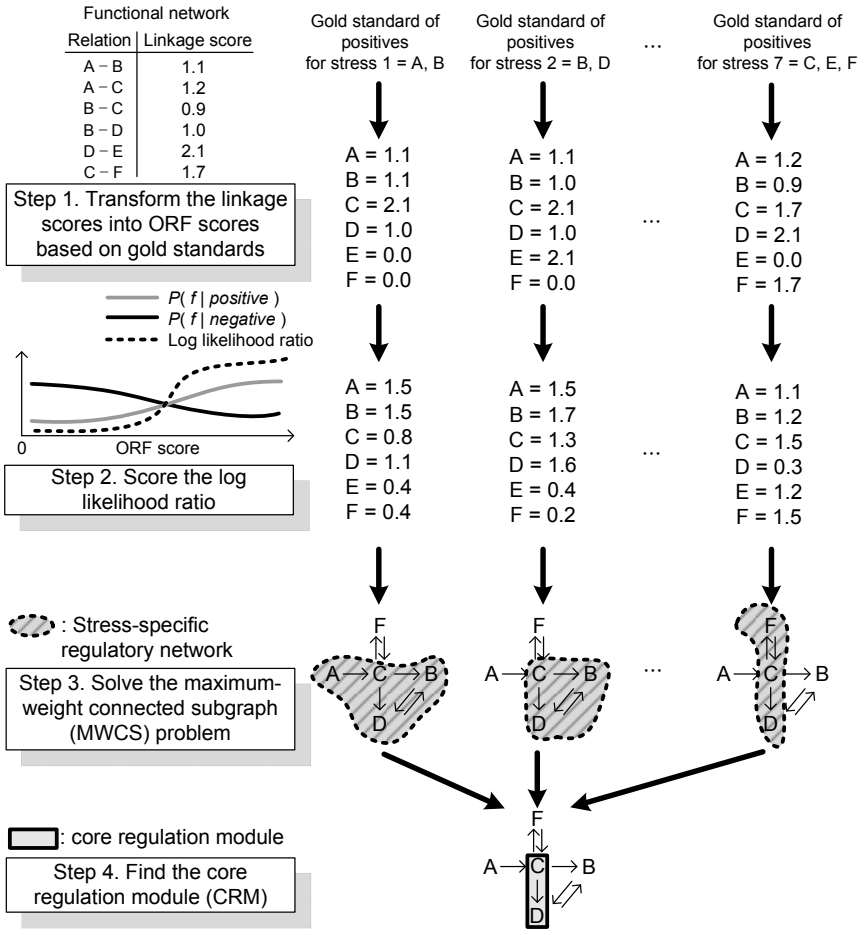
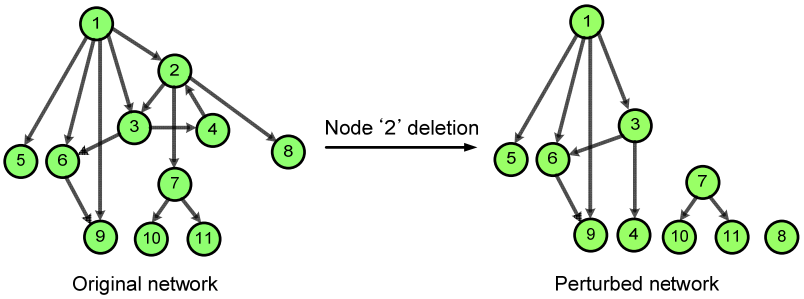


Figure S2. A schematic diagram illustrating the algorithm for identification of a CRM. (1) As a first step, we computed the stress-specific ORF scores by transforming the functional linkage score between ORFs of the "YeastNet". In this example, the node score of C for stress 1 is 2.1 since C has functional linkage with the gold standard of positives A and B, for stress 1, and the linkage scores of C with A and B are 1.2 and 0.9, respectively. (2) In the second step, we computed the stress-specific ORF scores based on log likelihood ratios. (3) Then, we found SRN for each stress by solving the MWCS problem based on the global regulatory network (523 nodes and 2,093 links where all nodes are regulatory molecules and links are directed molecular regulatory interactions) and the stress-specific ORF scores obtained by utilizing the "YeastNet". (4) Finally, the CRM is obtained by investigating the common subset of seven SRNs. The key concept of this procedure is to find biologically active regulatory molecules from various stresses based on a priori knowledge from a network perspective.

(A)



(B)

Node ID	Original network	Perturbed network
1	4	4
2	3	Deleted
3	3	3
4	3	1.5
5	2	2
6	2	2
7	2	2
8	1.5	Isolated
9	1	1
10	1	1
11	1	1

Figure S3. An illustrative example of computing the hierarchy destruction score. (A) Diagrams showing an original network and a perturbed network where node '2' is removed. (B) A list of hierarchical orders of each node in the original network and the perturbed network. In this example, the hierarchy destruction score for node 2 is 1.5.

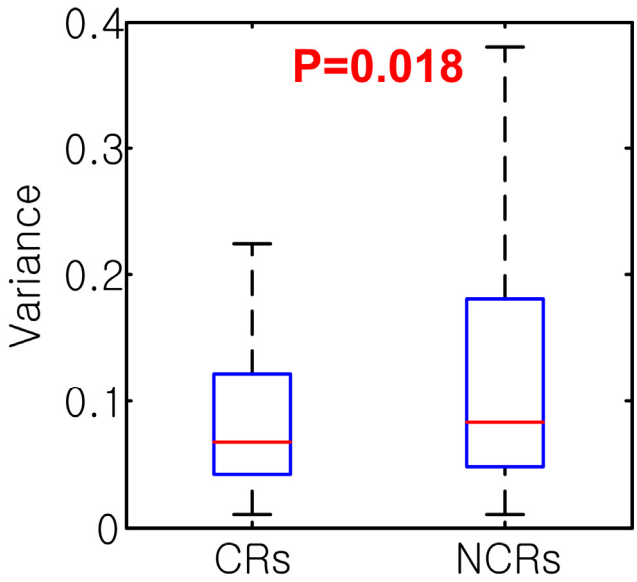


Figure S4. Variance of gene expression profiles between CRs and NCRs. Gene expression profiles were employed from (21). A box plot show a five-number summary of data: the 25th, 50th, 75th percentiles of the samples, and the two most extreme values within 1.5 times of the interquartile range (distance between 25th and 75th percentiles). P-values were computed using Wilcoxon's rank-sum test.

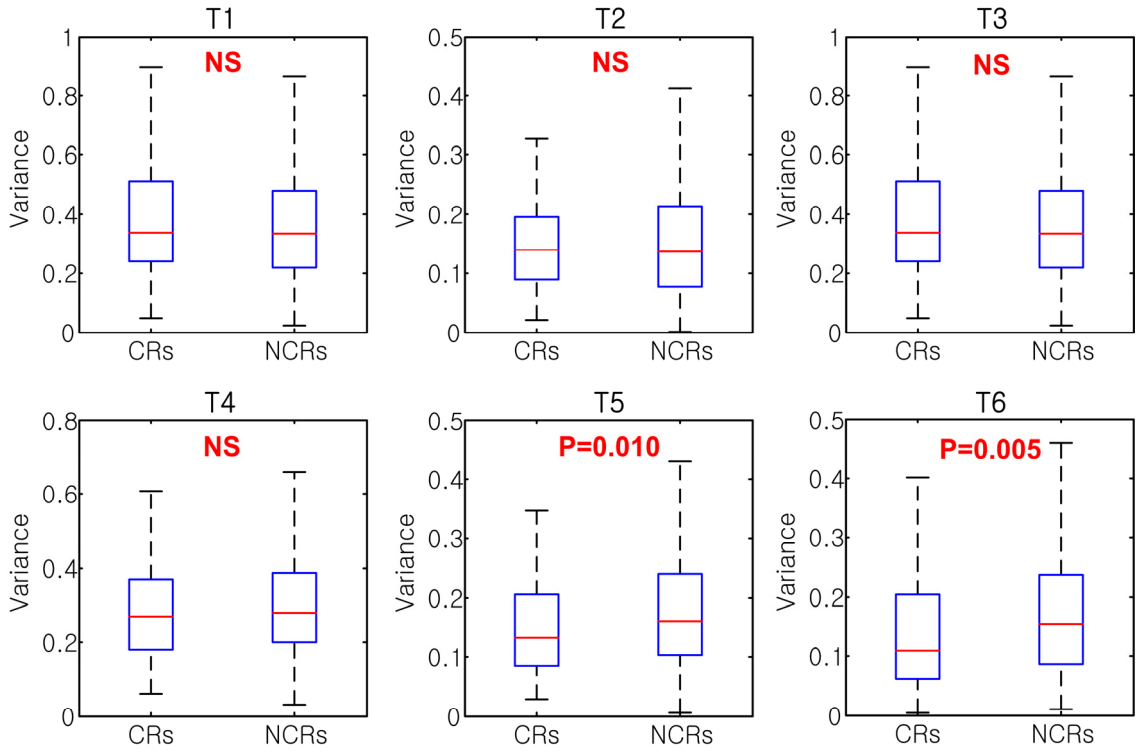


Figure S5. Variance of gene expression profiles between CRs and NCRs. Gene expression profiles were employed from (22) and T1-6 indicate each time point described therein. Each box plot shows a five-number summary of data: the 25th, 50th, 75th percentiles of the samples, and the two most extreme values within 1.5 times of the interquartile range (distance between 25th and 75th percentiles). P-values were computed using Wilcoxon's rank-sum test. NS indicates 'not significant at $P<0.05$.'

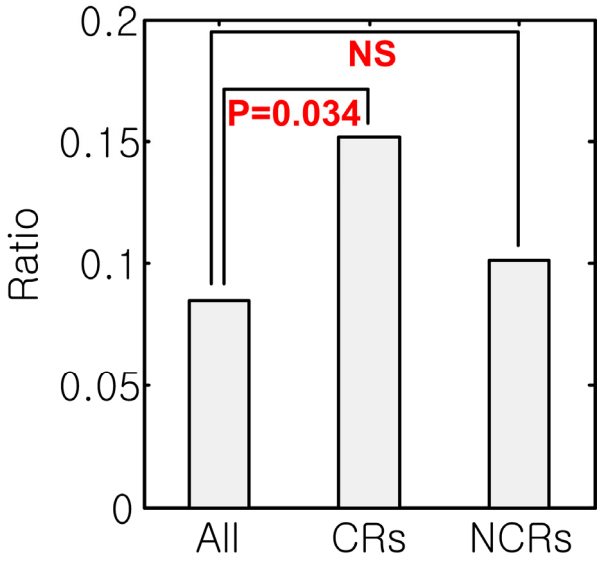


Figure S6. The ratio of MDR genes from all genes, CRs, and NCRs. The MDR gene lists were employed from (9). We computed the 'All' fixing 6,000 genes in total. *P*-values were computed using Hypergeometric test. NS indicates 'not significant at P<0.05.'

Table S1. Gold standards of positives and negatives for a particular stress condition. Abbreviations: Adenosine monophosphate; AMP, Ribonucleotide-diphosphate reductase; RNR, High osmolarity glycerol; HOG, Glutathione; GSH, Target of rapamycin; TOR, Protein kinase C; PKC, Protein kinase A; PKA

Stress type	# of gold standards of positives	# of gold standards of negatives	Regulated pathways or related functions	References
Adenine dropout	205	5,207	AMP biosynthesis, adenine nucleotide synthesis, Ty1 transposition, Proton transport, Pyridine nucleotide metabolic process, glycine, adenine, response to starvation	(23-24)
DNA damage	303	5,109	DNA repair pathway, RNR pathway, DNA polymerase activity, response to radiation, response to DNA damage	(25-26)
Glycerol	469	4,943	HOG pathway, MAPK pathway, Glycerol pathway, Glycolytic pathway, Energy production pathway, cellular carbohydrate catabolic process, glycerol	(27-28)
H_2O_2	226	5,186	Thiolredoxin pathway, GSH pathway, Yap1 pathway, Hydroperoxide pathway, TOR pathway, Pentose phosphate pathway, 14-3-3 pathway, Sir 2 pathway, Oxidative stress	(29-31)
Heat shock	798	4,614	Cell integrity pathway, PKC pathway, MAPK pathway, ubiquitin-proteasome pathway, Heat shock pathway, HOG pathway, response to heat	(32-34)
NaCl	230	5,182	HOG pathway, Calcineurin pathway, TOR pathway, PKA pathway, snf1p pathway, Rim101p pathway, Sodium ion transport	(35)
Sorbitol	953	4,459	HOG pathway, polyol pathway, Cell integrity pathway, PKC pathway, Xylose pathway, Sorbitol pathway	(28,36-39)

Table S2. GO term enrichment analysis of the CRM. CRM is computed according to various CRs/NCRs classifications based on different noise levels (σ). N=±5 (10)% indicates the networks obtained by randomly adding (or removing) 5 (10) % of links. Sample frequency is defined as the genes in the CRM and background frequency is denoted with respect to the 523 regulators. *P*-values were computed using Hypergeometric test and all the *P*-values were Bonferroni corrected. *P*-values smaller than 0.05 are shown in boldface.

GO terms	Noise level	Sample frequency	Background frequency	*P*-value
Regulation of cell cycle (GO:0051726)	σ=0.00	42/130 (32.3%)	94/523 (18.0%)	**0.004**
	σ=0.05	42/127 (33.1%)	94/523 (18.0%)	**0.002**
	σ=0.10	40/119 (33.6%)	94/523 (18.0%)	**0.002**
	σ=0.15	38/109 (34.9%)	94/523 (18.0%)	**0.002**
	σ=0.20	38/102 (37.3%)	94/523 (18.0%)	**2E-04**
	σ=0.25	33/92 (35.9%)	94/523 (18.0%)	**0.005**
	σ=0.30	31/79 (39.2%)	94/523 (18.0%)	**0.001**
	σ=0.35	26/64 (40.6%)	94/523 (18.0%)	**0.005**
	σ=0.30, N=+5%	26/72 (36.1%)	94/523 (18.0%)	0.067
	σ=0.30, N=-5%	26/69 (37.7%)	94/523 (18.0%)	**0.027**
	σ=0.30, N=+10%	28/77 (36.4%)	94/523 (18.0%)	**0.027**
	σ=0.30, N=-10%	25/69 (36.2%)	94/523 (18.0%)	0.090
Response to stress (GO:0006950)	σ=0.00	52/130 (40.0%)	136/523 (26.0%)	0.055
	σ=0.05	50/127 (39.4%)	136/523 (26.0%)	0.146
	σ=0.10	49/119 (41.2%)	136/523 (26.0%)	**0.038**
	σ=0.15	47/109 (43.1%)	136/523 (26.0%)	**0.012**
	σ=0.20	45/102 (44.1%)	136/523 (26.0%)	**0.010**
	σ=0.25	40/92 (43.5%)	136/523 (26.0%)	0.058
	σ=0.30	39/79 (49.4%)	136/523 (26.0%)	**0.001**
	σ=0.35	30/64 (46.9%)	136/523 (26.0%)	0.111
	σ=0.30, N=+5%	34/72 (47.2%)	136/523 (26.0%)	**0.028**
	σ=0.30, N=-5%	34/69 (49.3%)	136/523 (26.0%)	**0.008**
	σ=0.30, N=+10%	36/77 (46.8%)	136/523 (26.0%)	**0.020**
	σ=0.30, N=-10%	34/69 (49.3%)	136/523 (26.0%)	**0.008**

Table S3. GO term enrichment analysis of the regulated genes by the CRM. CRM is computed according to various CRs/NCRs classifications based on different noise levels (σ). N=±5 (10)% indicates the networks obtained by randomly adding (or removing) 5 (10)% of links. Sample frequency is defined as the ratio of genes regulated by the CRM and background frequency is denoted as the ratio of genes regulated by all 523 regulators. *P*-values were computed using Hypergeometric test and all the *P*-values were Bonferroni corrected. *P*-values smaller than 0.05 are shown in boldface.

GO terms	Noise level	Sample frequency	Background frequency	*P*-value
Cell cycle (GO:0007049)	σ=0.00	354/2033 (17.4%)	428/3006 (14.2%)	**1.50E-10**
	σ=0.05	354/2025 (17.5%)	428/3006 (14.2%)	**6.03E-11**
	σ=0.10	339/1957 (17.3%)	428/3006 (14.2%)	**2.30E-08**
	σ=0.15	331/1835 (18.0%)	428/3006 (14.2%)	**4.17E-11**
	σ=0.20	318/1718 (18.5%)	428/3006 (14.2%)	**6.80E-12**
	σ=0.25	307/1656 (18.5%)	428/3006 (14.2%)	**8.06E-11**
	σ=0.30	275/1420 (19.4%)	428/3006 (14.2%)	**6.32E-11**
	σ=0.35	268/1383 (19.4%)	428/3006 (14.2%)	**2.62E-10**
	σ=0.30, N=+5%	273/1405 (19.4%)	428/3006 (14.2%)	**8.96E-11**
	σ=0.30, N=-5%	276/1431 (19.3%)	428/3006 (14.2%)	**1.54E-10**
	σ=0.30, N=+10%	281/1435 (19.6%)	428/3006 (14.2%)	**3.84E-12**
	σ=0.30, N=-10%	275/1391 (19.8%)	428/3006 (14.2%)	**3.24E-12**
Negative regulation of biological process (GO:0048519)	σ=0.00	244/2033 (12.0%)	293/3006 (9.7%)	**7.35E-07**
	σ=0.05	244/2025 (12.0%)	293/3006 (9.7%)	**3.95E-07**
	σ=0.10	239/1957 (12.2%)	293/3006 (9.7%)	**2.30E-07**
	σ=0.15	234/1835 (12.8%)	293/3006 (9.7%)	**1.44E-09**
	σ=0.20	226/1718 (13.2%)	293/3006 (9.7%)	**1.90E-10**
	σ=0.25	219/1656 (13.2%)	293/3006 (9.7%)	**8.99E-10**
	σ=0.30	198/1420 (13.9%)	293/3006 (9.7%)	**4.65E-10**
	σ=0.35	192/1383 (13.9%)	293/3006 (9.7%)	**4.43E-09**
	σ=0.30, N=+5%	198/1405 (14.1%)	293/3006 (9.7%)	**1.64E-09**
	σ=0.30, N=-5%	202/1431 (14.1%)	293/3006 (9.7%)	**4.17E-10**
	σ=0.30, N=+10%	204/1435 (14.2%)	293/3006 (9.7%)	**8.98E-11**
	σ=0.30, N=-10%	198/1391 (14.2%)	293/3006 (9.7%)	**4.68E-10**

Table S4. Topological properties of the CRM compared to random CRs/NCRs classifications. *P*-values of the topological properties of the CRM according to various CRs/NCRs classifications based on different noise levels (σ). N=±5 (10)% indicates the networks obtained by randomly adding (or removing) 5 (10)% of links. *P*-values smaller than 0.05 are shown in boldface. For hierarchy destruction score, *P*-values were computed using Wilcoxon's Rank-sum test and, for the others, *P*-values were computed based on randomly selected sub-networks with the same number of nodes (1,000 times).

Noise level	Number of the Middle nodes in the CRM	Hierarchy destruction score	Number of two-node feedback loops in the CRM	Number of feed-forward loops in the CRM
σ=0.00	**0.008**	**2E-04**	**0.001**	**0.028**
σ=0.05	**0.021**	**3E-04**	**0.002**	**0.030**
σ=0.10	**0.013**	**2E-04**	**0.001**	0.054
σ=0.15	**0.016**	**0.002**	**0.005**	0.063
σ=0.20	**0.003**	**0.002**	**0.001**	0.050
σ=0.25	**0.006**	**2E-04**	**0.001**	0.064
σ=0.30	**0.002**	**6E-04**	**0.008**	**0.036**
σ=0.35	**0.001**	**2E-04**	**0.001**	**0.001**
σ=0.30, N=+5%	**0.001**	**3E-04**	**0.007**	**0.015**
σ=0.30, N=-5%	**0.021**	**3E-04**	**0.001**	**0.015**
σ=0.30, N=+10%	**0.001**	**1E-04**	**0.001**	**0.005**
σ=0.30, N=-10%	**0.001**	**1E-04**	**0.001**	**0.032**

Table S5. Number of feedback loops in the CRM. *P*-values of the topological properties of the CRM according to various CRs/NCRs classifications based on different noise levels (σ). N=±5 (10)% indicates the networks obtained by randomly adding (or removing) 5 (10)% of links. *P*-values smaller than 0.05 are shown in boldface. *P*-values were computed based on randomly selected sub-networks with the same number of nodes (1,000 times).

Noise level	Number of three-node feedback loops in the CRM	Number of four-node feedback loops in the CRM	Number of five-node feedback loops in the CRM
σ=0.00	**0.001**	**0.004**	**0.001**
σ=0.05	**0.004**	**0.011**	**0.014**
σ=0.10	**0.003**	**0.008**	**0.009**
σ=0.15	**0.009**	**0.013**	**0.013**
σ=0.20	**0.004**	**0.007**	**0.008**
σ=0.25	**0.032**	**0.113**	0.226
σ=0.30	**0.017**	0.060	0.141
σ=0.35	**0.002**	**0.004**	**0.004**
σ=0.30, N=+5%	**0.013**	**0.049**	0.097
σ=0.30, N=-5%	**0.001**	**0.006**	**0.004**
σ=0.30, N=+10%	**0.003**	**0.006**	**0.006**
σ=0.30, N=-10%	**0.005**	**0.028**	0.143

Table S6. Genetic properties of the CRM compared with random CRs/NCRs classifications. *P*-values of the genetic properties of the CRM according to various CRs/NCRs classifications based on different noise levels (σ). N=±5 (10)% indicates the networks obtained by randomly adding (or removing) 5 (10)% of links. *P*-values smaller than 0.05 are shown in boldface. For the ratio of synthetic lethal pairs, *P*-values were computed using Fisher's exact test and, for the others, *P*-values were computed based on Wilcoxon's Rank-sum test.

Noise level	Evolutionary rate (Rank)	Interstrain variance	Ratio of synthetic lethal pairs	Growth rate defect
σ=0.00	**6E-08**	**0.008**	**<2E-16**	**0.002**
σ=0.05	**4E-07**	**0.009**	**<2E-16**	**0.002**
σ=0.10	**1E-05**	**0.026**	**<2E-16**	**0.003**
σ=0.15	**1E-06**	**0.023**	**<2E-16**	**0.004**
σ=0.20	**9E-07**	**0.009**	**<2E-16**	**0.002**
σ=0.25	**7E-07**	**0.021**	**<2E-16**	**6E-04**
σ=0.30	**5E-07**	**0.043**	**<2E-16**	**0.007**
σ=0.35	**2E-06**	**0.034**	**<2E-16**	**0.008**
σ=0.30, N=+5%	**8E-09**	**0.020**	**<2E-16**	**0.009**
σ=0.30, N=-5%	**3E-07**	**0.005**	**<2E-16**	**0.003**
σ=0.30, N=+10%	**1E-07**	**0.014**	**<2E-16**	**0.019**
σ=0.30, N=-10%	**2E-07**	**0.021**	**<2E-16**	**0.004**

Table S7. Genetic properties of the CRM compared with random CRs/NCRs classifications. *P*-values of the genetic properties of the CRM according to various CRs/NCRs classifications based on different noise levels (σ). N=±5 (10)% indicates the networks obtained by randomly adding (or removing) 5 (10)% of links. *P*-values smaller than 0.05 are shown in boldface. For the ratio of MDR genes in CRs, *P*-values were computed using Hypergeometric test and, for the others, *P*-values were computed based on Wilcoxon's Rank-sum test.

Noise level	Ratio of MDR genes in CRs	Expression variance (T5) (22)	Expression variance (T6) (22)	Expression variance (21)
σ=0.00	**0.026**	0.269	**0.007**	**0.017**
σ=0.05	**0.021**	0.278	**0.012**	**0.021**
σ=0.10	**0.023**	0.253	**0.020**	**0.018**
σ=0.15	**0.021**	0.212	**0.026**	**0.015**
σ=0.20	**0.012**	0.099	**0.025**	**0.010**
σ=0.25	**0.010**	0.069	**0.012**	**0.005**
σ=0.30	**0.034**	**0.010**	**0.005**	**0.018**
σ=0.35	0.090	0.188	**0.003**	0.106
σ=0.30, N=+5%	**0.007**	0.071	**0.006**	**0.014**
σ=0.30, N=-5%	**0.030**	**0.035**	**6E-04**	**0.024**
σ=0.30, N=+10%	**0.028**	0.068	**0.002**	**0.042**
σ=0.30, N=-10%	**0.030**	0.072	**0.013**	0.103