

Transcriptional response of steady-state yeast cultures to transient perturbations in carbon source

Michal Ronen* and David Botstein†‡

*Department of Genetics, Stanford University School of Medicine, Stanford, CA 94305; and †Lewis–Sigler Institute for Integrative Genomics, Carl Icahn Laboratory, Princeton University, Princeton, NJ 08544

Contributed by David Botstein, November 17, 2005

To understand the dynamics of transcriptional response to changing environments, well defined, easily controlled, and short-term perturbation experiments were undertaken. We subjected steady-state cultures of *Saccharomyces cerevisiae* in chemostats growing on limiting galactose to two different size pulses of glucose, well known to be a preferred carbon source. Although these pulses were not large enough to change growth rates or cell size, $\approx 25\%$ of the genes changed their expression at least 2-fold. Using DNA microarrays to estimate mRNA abundance, we found a number of distinguishable patterns of transcriptional response among the many genes whose expression changed. Many of these genes were already known to be regulated by particular transcription factors; we estimated five potentially relevant transcription factor activities from the observed changes in gene expression (i.e., Mig1p, Gal4p, Cat8p, Rgt1p, Adr1p, and Rcs1p). With these estimates, for two regulatory circuits involving interaction among multiple regulators we could generate dynamical models that quantitatively account for the observed transcriptional responses to the transient perturbations.

regulatory dynamics | steady-state growth | transcription factor interactions

Yeast has evolved to respond rapidly and effectively to fluctuations in temperature, nutrients, and other environmental changes. When environmental conditions change abruptly, cells must rapidly adjust their genomic expression program to adapt to the new conditions. Most environmental perturbations studied to date are complex and appear to provoke multiple regulatory systems simultaneously; Gasch *et al.* (1) found thousands of genes affected, directly or indirectly, by a great variety of stresses, including such things as starvation, oxidative stress, and heat shock. Furthermore, long-term perturbations, such as introduction of mutations or shifts of medium (2, 3), are inevitably complicated by long-term adaptations as well as short-term responses.

We developed an experimental design intended to focus on the immediate response to environmental change, by making very small and specific perturbations to an otherwise rigorously controlled environment. Cultures growing at steady state in galactose-limited chemostats were subjected to pulses of added glucose, well known to be a preferred source of carbon and energy (4, 5). A single bolus of glucose was added to the growth vessel, and the concentration was allowed to fall naturally by the combination of the constant dilution characteristic of the chemostat and metabolism by the cells. Two sizes of pulse (0.2 or 2.0 g/liter) were administered, so that we could observe a kind of “dose–response” to the same perturbation. The perturbations were small enough to prevent noticeable growth rate alteration and other major physiological changes; however, they proved sufficient to provoke substantial changes in transcript levels. By pulsing with glucose, a metabolite expected to be preferentially consumed by the cells, we expected to capture, over time, two opposite events: response to the addition of the preferred metabolite and then to its disappearance over time.

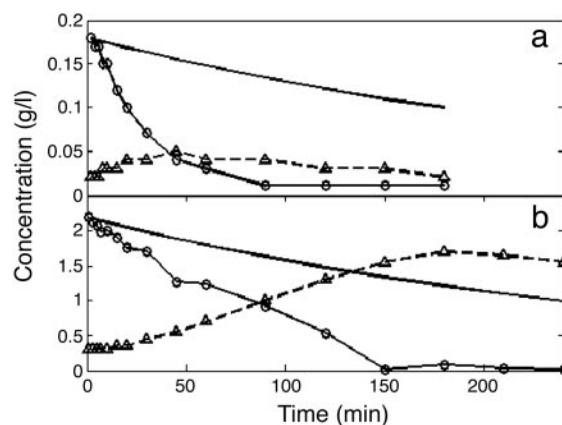


Fig. 1. Glucose (○) and ethanol (△) concentrations observed after the 0.2 g/liter (a) and 2.0 g/liter (b) pulses of glucose. The glucose concentrations predicted solely from dilution in the chemostat are shown by the solid line.

The cellular response, in terms of transcriptional control, can be represented by the transcription factor activity (TFA) (6, 7). Transcription factors (TFs) become active through series of signal transduction events triggered by extracellular or intercellular metabolites. Activation by translocation, phosphorylation, ligand binding, or other biochemical modifications renders the TF capable of DNA binding and results in gene expression. Direct measurement of TFA profiles is not usually possible; therefore, we treated them as intrinsic (hidden) state variables. Global methods for TFA estimation (8–10) require assumptions of steady-state mRNA production and constant mRNA degradation rate, which do not hold in our case. Therefore, we estimated each TFA locally using mRNA levels of sets of downstream genes thought to be directly regulated by the factor(s). In our analysis we also were able to combine TFA profiles to capture combinatorial regulation of the transcripts.

Results

Response to Glucose Pulses. When glucose was added to the cultures growing at steady state in limiting galactose, its concentration immediately began to fall at a rate substantially higher than the rate of dilution (Fig. 1), as expected if the cells immediately began to metabolize it. The observed $t_{1/2}$ was three times shorter than that predicted for dilution alone for the high-glucose pulse and approximately nine times shorter for the low-glucose pulse. During each pulse, ethanol concentration increased until the glucose supply was exhausted, ≈ 45 min for the low pulse and 150 min for the high pulse. No significant change was found in cell number and size ($<10\%$), confirming

Conflict of interest statement: No conflicts declared.

Abbreviations: TF, transcription factor; TFA, TF activity.

†To whom correspondence should be addressed. E-mail: botstein@princeton.edu.

© 2005 by The National Academy of Sciences of the USA

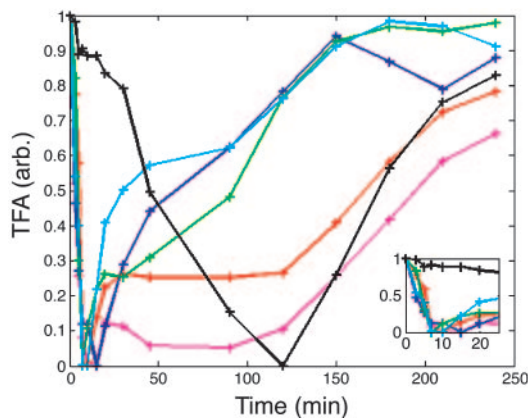


Fig. 4. The estimated TFAs with the 2 g/liter glucose pulse, 1-Mig1 (blue), Gal4 (green), Cat8 (red), Rgt1 (pink), Adr1 (cyan), and Rcs1 (black). (Inset) A zoom-in to the initial time points.

10 min, followed by a gradual increase to higher than initial levels, reaching their peak when glucose was depleted and then dropping down. For the low pulse the same pattern appeared, with a compressed time scale. Two repressors, Mig1p and Rgt1p, which are oppositely responsive to glucose (Fig. 5a), have been reported to regulate both genes (15, 18). When glucose levels increase, Rgt1p repression is relieved and repression by Mig1p is enhanced.

The TFA plot (Fig. 5b) reveals the different thresholds of the repressors, with Rgt1p at ≈ 0.03 g/liter and Mig1p at ≈ 0.7 g/liter, so that, when the glucose pulse begins, there is Mig1p-dependent repression, which is relieved when glucose concentration drops below 0.7; then, when it drops below 0.03 g/liter, Rgt1p is activated to reestablish repression. As described in *Materials and Methods*, a quantitative model of regulation of the mRNA level for the *HXX2* and *MTH1* genes by combined TFs was generated. As shown in Fig. 5c, the model succeeded in producing the observed bidirectional mRNA profile, including amplitude and duration, for both the low-glucose and high-glucose pulses.

Lasting Response Model for Genes Regulated by Cat8p, Mig1p, and the Snf1p Kinase. The lasting response genes that are regulated under the transcriptional activator Cat8p were used to calculate its active profile; the *CAT8* gene itself is under Mig1p regulation (19) and was used to calculate active Mig1p's profile. Both Mig1p and Cat8p receive signal from glucose

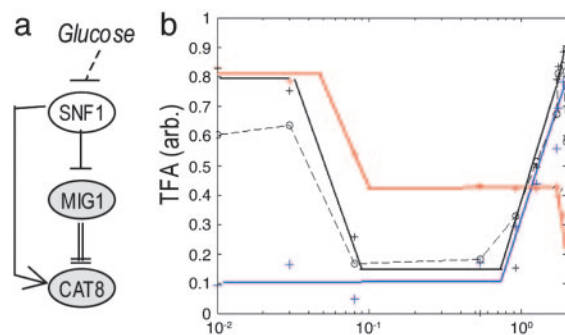


Fig. 6. Lasting response. (a) Mig1-Cat8 circuit. (b) The estimated active TFAs during derepression after a high-glucose pulse. Blue, Mig1p; red, Cat8p; black solid line, Rcs1p; black dashed line, predicted Snf1p activity.

throughout Snf1p (19) in a feed-forward manner. High glucose initiates the repression of *CAT8* by Mig1p (decreasing the number of activator molecules) and also directly decreases the activity of Cat8p itself (Fig. 6a).

The TFA curves (Fig. 6b) show that Mig1p and Cat8p have different glucose (or Snf1p) thresholds; Cat8p's final threshold is at an ≈ 10 -fold lower concentration than that of Mig1p (from the high pulse data), indicating the short-term Mig1p activity. Therefore, the lasting response of Cat8p might be attributed to Snf1p and not to Mig1p. Although Mig1p's activity is also regulated by Snf1p (20, 21), it has different kinetics. This finding could be explained either by different thresholds or by different Snf1p complexes involved in each interaction. The pattern of Snf1p signaling activity is estimated by substituting Mig1p activity from that of Cat8p; it correlates well with the pattern of Rcs1p transcriptional activity, downstream of Snf1p (22).

Comparing the TFAs for the Two Pulses. Rgt1p's activities have similar dependency on the concentration of glucose for the two pulses (Fig. 5b); Cat8p displays similar behavior (data not shown). However, Mig1p's activity does not depend linearly on the glucose level; its thresholds are different for the different pulses (≈ 0.8 g/liter for the high pulse and ≈ 0.08 g/liter for the low pulse) (Fig. 5b). In the glucose ranges that we tested (both commonly regarded as "low-glucose") it seems to act as a digital pulse, with the same amplitude and width, regardless of glucose level. Direct measurements of Mig1p localization at different glycolytic rates (23) support this interpretation.

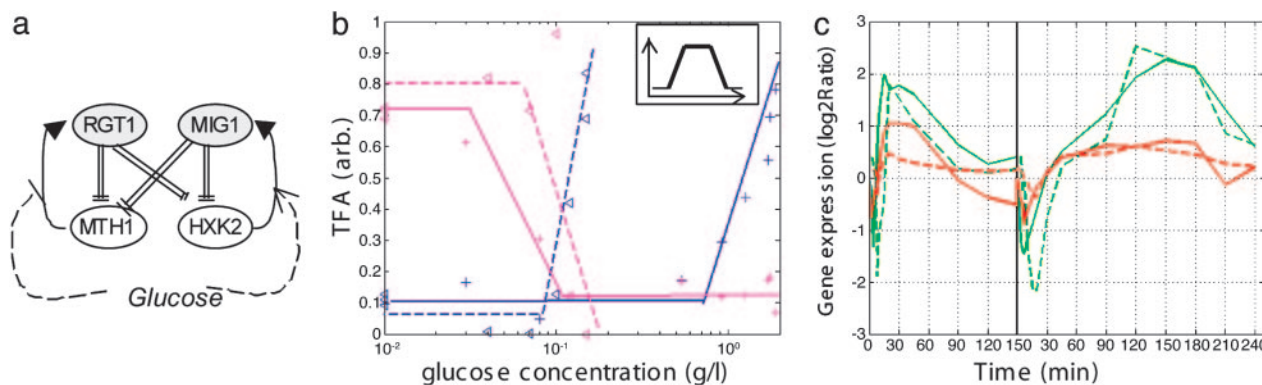


Fig. 5. Bidirectional response model. (a) Mig1-Rgt1 circuits. (b) The estimated TFA during derepression. Blue, Mig1p; pink, Rgt1p; solid line, high pulse; dashed line, low pulse. (Inset) The combined factors pattern of gene activation. (c) The expression of *MTH1* (green) and *HXX2* (red) genes, measured (solid line) and model-predicted (dashed line)

Discussion

We designed our experiment in the hope that the stimulus and response could be observed cleanly in a procedure that minimizes the introduction of confounding physiological changes. To this end, we arranged for the cells to be in a steady state, so that gene expression pattern and metabolic fluxes could be safely assumed to be constant and any changes could be attributable to the perturbation. The addition of a small amount of glucose, although a major factor for the cell's metabolism, as indicated by the immediate responses of so many genes, nevertheless did not alter the cell growth rate and size.

Thus, we are reasonably confident that all responses that we observe are, directly or indirectly, consequences of just the brief glucose pulse and not any other component in the system. Moreover, the small stimulation allowed us to capture interesting patterns of regulation that could be masked with high/saturated doses, and, by following the kinetics, new correlations in gene expression could be revealed. Cells grown in chemostat are "poor but not starving" (24, 25), and, indeed, no explicitly starvation-related and stress-response genes were activated (except for "stress" genes that were previously known to be induced by glucose-related pathways).

Most responding genes have a burst-like expression pattern, relatively uncorrelated with the measured glucose and ethanol concentrations and consumption rates. Even a low added-glucose level triggered the subsequent transcriptional responses, although the duration sometimes was truncated relative to a longer pulse. One of the TFs involved, Mig1p, also reflects this behavior, because it seems to act as a digital pulse. The initial response to very low levels of glucose added to the culture imply the high importance of fast adaptation to it. We suggest that there is more than one discrete wave of transcriptional response to glucose appearance, where each wave "reevaluates" the level of glucose, and maybe also its derivatives, and responds appropriately. Similar discrete regulation was found with responses to stress with the SOS DNA repair system in *Escherichia coli* (26) and the human p53-mdm2 (27). To test this suggestion, analysis of single-cell response with high temporal resolution is required.

Bidirectional response is generated by repression of two TFs with different activity thresholds (Mig1p and Rgt1p). Their downstream genes, *HXK2* and *MTH1* (28–30), participate in the regulation of two of the glucose pathways (*HXK2* may also have a role in the cAMP-related pathway). They are regulated in opposing directions by glucose, with low expression at high and at very low glucose levels and high expression at an intermediate level, which implies adaptivity. For Mth1p, which is a sensor for glucose, protein production is the mean to achieve adaptivity, because the protein is degraded in the presence of glucose. The high gene expression at intermediate glucose levels increases its sensitivity to glucose at that range.

The expression pattern of Mth1 and Hxk2 was captured by using a simple model. We suggest that contradicting signals also regulate *SUC2*, *HXK1*, and the additional genes with similar expression patterns.

Iron ion homeostasis genes respond to glucose signals through Snf1p regardless of iron deprivation signals. We suggest that this signaling pathway is an evolutionary adaptation to successive cellular events; i.e., if a consequence of glucose addition is the impending depletion of Fe ions, a cell that would start expressing the iron homeostasis genes once glucose is added will have an advantage over a cell that will do that only when iron is depleted. This kind of anticipatory coordinated expression could explain the large number of mRNAs changing in response to glucose throughout glucose signaling pathways.

The sharp drop in mRNA levels of the repressed genes is due to a high degradation rate. The mRNA's half-lives represent-

ing degradation under glucose exponential growth (31) are ≈ 3 -fold higher than those detected in our experiments. It was recently reported that feeding glucose to cells growing on other carbon sources increases the mRNA degradation rate for a few genes (*JEN1* and *FBP1*) (32). It is possible that the same phenomenon occurs with the larger set of genes we see in our experiments.

To conclude, we believe that experiments designed around short perturbations of cells growing in steady state may be of general use in understanding metabolic regulatory networks. Estimated active TF response curves generated as described above provide a complex picture of the cells' dynamical regulation, but one that can readily be used to produce models that can be compared with experimental observations. In turn, successes in such modeling can provide experimental support for proposed regulatory modules. Extending the scope of the analysis by additional estimated TFAs would provide a broader view of the cellular response and enable additional modules to be supported by experimental evidence.

Materials and Methods

Strain and Growth. DBY10085 is a haploid prototrophic derivative of CEN.PK122 (*Mata*; *URA*; *LEU*; *HIS*; *TRP*; *MAL2-8^c*; *SUC2*). Minimal defined medium contains, in addition to carbon source, CaCl₂·H₂O (0.1 g/liter), NaCl (0.1 g/liter), Mg₂(SO₄)₂·7H₂O (0.5 g/liter), K₂HPO₄ (1 g/liter), (NH₄)₂SO₄ (5 g/liter), boric acid (0.5 mg/liter), CuSO₄·5H₂O (0.04 mg/liter), KI (0.1 mg/liter), FeCl₃·5H₂O (0.2 mg/liter), MnSO₄·H₂O (0.4 mg/liter), Na₂MoO₄·2H₂O (0.2 mg/liter), ZnSO₄·5H₂O (0.4 mg/liter), biotin (2 μ g/liter), calcium pantothenate (400 μ g/liter), folic acid (2 μ g/liter), inositol (2 mg/liter), niacin (400 μ g/liter), *p*-aminobenzoic acid (200 μ g/liter), pyridoxine HCl (400 μ g/liter), riboflavin (200 μ g/liter), and thiamine HCl (400 μ g/liter).

Chemostat. A 300-ml fermenter vessel was adapted as a chemostat by using minimal defined medium with 2 g/liter galactose, 5 standard liters/min air flow, and stirring at 400 rpm at 30°C. The dilution rate was ≈ 0.2 vol/h. Once the chemostat culture achieved a constant cell number for at least 3 vol changes, a glucose pulse (enough to bring the transient concentration in the vessel to 0.2 or 2.0 g/liter) was injected into the vessel. In experiment set A, samples were collected at $t = 0, 5, 10, 20, 30, 45, 60,$ and 90 min. In experiment set B, samples were collected at $t = 0, 1, 3, 5, 7, 10, 15, 20, 30, 45, 60, 90, 120, 150, 180, 210,$ and 240 min for the high-level pulse, and at $t = 0, 2, 4, 6, 8, 10, 15, 20, 30, 45, 60, 90, 120, 150,$ and 180 min for the low-level pulse. The same samples were used to measure glucose and ethanol concentrations, cell number and size, and mRNA levels. The results presented here are from set B; the reference for the DNA microarrays was the $t = 0$ sample. All of the features we describe for set B appeared also in set A.

Physiological Measurements. At each time point a sample was withdrawn and divided for the following analyses.

Cell density, number, and size. A 1-ml sample was sonicated (a 10-s 50% duty cycle). The density was measured by light scattering at 600 nm, and cell number and size were measured in a Coulter counter (model Z2, Beckman Coulter).

Glucose and ethanol. A 1-ml sample was centrifuged at high speed to remove cells; the supernatant was stored at -20°C . The residual glucose and ethanol concentration in the growth medium were assayed by enzyme-coupled NADH oxidation reactions (assay kits from R-Biopharm, Darmstadt, Germany).

RNA Isolation, Labeling, Microarray Hybridization, and Analysis. Five-milliliter samples were vacuum-filtered onto a nylon filter, which was quickly placed into a tube and stored in liquid

nitrogen at -80°C . RNA was extracted from culture filtrate by the acid-phenol method followed by ethanol precipitation (33), amplified (MessageAmp aRNA Kit, Ambion), and labeled with Cy3 (time points) or Cy5 (reference). Cy5- and Cy3-labeled probes were hybridized together to microarrays printed with PCR-amplified fragments containing all known and predicted *Saccharomyces cerevisiae* ORFs (12). Microarrays were washed and then scanned with a GenePix 4000B laser scanner (Axon Instruments, Foster City, CA), and the images were analyzed by using GENEPIX PRO software. Resulting microarray data were submitted to the Stanford MicroArray Database (34) (<http://genome-www5.stanford.edu>).

Results for each gene and time point are expressed as \log_2 of the ratio of sample signal divided by the reference signal. Spots were filtered by correlation (0.6), level over background (at least 80% of pixels $>$ background + 1 SD), and by manual gridding of the image.

Models. The mRNA signal is modeled as

$$\frac{dM}{dt} = \frac{\beta}{1 + \left(\frac{R}{k}\right)^h} - k_{\text{deg}} \times M, \quad [1]$$

for a TF that functions as a repressor, where M is the mRNA level, R is the effective repressor concentration, β is the maximal production rate, k is the affinity of the TF (concentration at half maximal repression), h is Hill coefficient (indicating the level of cooperativity), and k_{deg} is a rate constant for mRNA degradation. The model assumes that cell growth rate remains constant.

Rewriting Eq. 1, we get the mRNA production term X :

$$X = \frac{dM}{dt} + k_{\text{deg}} \times M = \frac{\beta}{1 + \left(\frac{R}{k}\right)^h}. \quad [2]$$

For TF that functions as an activator we get

$$X = \frac{\beta \times \left(\frac{A}{k}\right)^h}{1 + \left(\frac{A}{k}\right)^h}, \quad [3]$$

where A is the activating TF concentration.

For a combined activator–repressor model we get

$$X = \frac{\beta \times \left(\frac{A}{k}\right)^h}{1 + \left(\frac{A}{k}\right)^h + \left(\frac{R}{k}\right)^h}. \quad [4]$$

More complex combined models were tested, but the resulting predictions were not significantly altered.

TFA Estimation. The relative transcription activity of Mig1, Cat8-Sip4, Rgt1, Gal4, Adr1, and Rcs1 were calculated. mRNA \log_2 (ratio) values of the genes were used with missing values estimated by KNNIMPUTE (35). The left side of Eq. 2 was calculated as $k_{\text{deg}} = \log_2/\text{minimal } t_{1/2}$; it is assumed that k_{deg} values are constant over the experiment's time, except at $t = 0$, where values from ref. 31 were used. We transformed the ratio model to a bilinear form and determined the parameter R (or A) using singular-value decomposition (for details see ref. 7). Their levels were set in the range [0, 1]. Nonlinear least squares was used to calculate the parameters k and h (LSQNONLIN MATLAB 7, Mathworks, Natick, MA). Cat8p activity is combined with that of Sip4, because they are inseparable, and, similarly, Mig2p is combined with Mig1p. Rcs1p activity was calculated for the high pulse only, because the data for the low pulse was too noisy. Estimation of other TFs (Sfp1/Fhl1/Ihf, ribosomal biogenesis TFs) was not attempted here.

We thank G. Sherlock and P. O. Brown for their generous help and M. Shapira for valuable discussions. This work was supported by National Institutes of Health Grant GM46406.

- Gasch, A. P., Spellman, P. T., Kao, C. M., Carmel-Harel, O., Eisen, M. B., Storz, G., Botstein, D. & Brown, P. O. (2000) *Mol. Biol. Cell* **11**, 4241–4257.
- Hughes, T. R., Marton, M. J., Jones, A. R., Roberts, C. J., Stoughton, R., Armour, C. D., Bennett, H. A., Coffey, E., Dai, H., He, Y. D., et al. (2000) *Cell* **102**, 109–126.
- Ideker, T., Thorsson, V., Ranish, J. A., Christmas, R., Buhler, J., Eng, J. K., Bumgarner, R., Goodlett, D. R., Aebersold, R. & Hood, L. (2001) *Science* **292**, 929–934.
- Gelade, R., Van de Velde, S., Van Dijck, P. & Thevelein, J. M. (2003) *Genome Biol.* **4**, 233.
- Johnston, M. (1999) *Trends Genet.* **15**, 29–33.
- Kao, K. C., Yang, Y. L., Boscolo, R., Sabatti, C., Roychowdhury, V. & Liao, J. C. (2004) *Proc. Natl. Acad. Sci. USA* **101**, 641–646.
- Ronen, M., Rosenberg, R., Shraiman, B. I. & Alon, U. (2002) *Proc. Natl. Acad. Sci. USA* **99**, 10555–10560.
- Bussemaker, H. J., Li, H. & Siggia, E. D. (2000) *Proc. Natl. Acad. Sci. USA* **97**, 10096–10100.
- Liao, J. C., Boscolo, R., Yang, Y. L., Tran, L. M., Sabatti, C. & Roychowdhury, V. P. (2003) *Proc. Natl. Acad. Sci. USA* **100**, 15522–15527.
- Tran, L. M., Brynildsen, M. P., Kao, K. C., Suen, J. K. & Liao, J. C. (2005) *Metab. Eng.* **7**, 128–141.
- Boyle, E. I., Weng, S., Gollub, J., Jin, H., Botstein, D., Cherry, J. M. & Sherlock, G. (2004) *Bioinformatics* **20**, 3710–3715.
- DeRisi, J. L., Iyer, V. R. & Brown, P. O. (1997) *Science* **278**, 680–686.
- Reifenberger, E., Boles, E. & Ciriacy, M. (1997) *Eur. J. Biochem.* **245**, 324–333.
- Diderich, J. A., Schepper, M., van Hoek, P., Luttik, M. A., van Dijken, J. P., Pronk, J. T., Klaassen, P., Boelens, H. F., de Mattos, M. J., van Dam, K. & Kruckeberg, A. L. (1999) *J. Biol. Chem.* **274**, 15350–15359.
- Kaniak, A., Xue, Z., Macool, D., Kim, J. H. & Johnston, M. (2004) *Eukaryot. Cell* **3**, 221–231.
- Polish, J. A., Kim, J. H. & Johnston, M. (2005) *Genetics* **169**, 583–594.
- Ozcan, S. (2002) *J. Biol. Chem.* **277**, 46993–46997.
- Moreno, F., Ahuatz, D., Riera, A., Palomino, C. A. & Herrero, P. (2005) *Biochem. Soc. Trans.* **33**, 265–268.
- Randez-Gil, F., Bojunga, N., Proft, M. & Entian, K. D. (1997) *Mol. Cell. Biol.* **17**, 2502–2510.
- Treitel, M. A. & Carlson, M. (1995) *Proc. Natl. Acad. Sci. USA* **92**, 3132–3136.
- De Vit, M. J., Waddle, J. A. & Johnston, M. (1997) *Mol. Biol. Cell* **8**, 1603–1618.
- Haurie, V., Boucherie, H. & Saggiocco, F. (2003) *J. Biol. Chem.* **278**, 45391–45396.
- Elbing, K., Stahlberg, A., Hohmann, S. & Gustafsson, L. (2004) *Eur. J. Biochem.* **271**, 4855–4864.
- Brauer, M. J., Saldanha, A. J., Dolinski, K. & Botstein, D. (2005) *Mol. Biol. Cell* **16**, 2503–2517.
- Saldanha, A. J., Brauer, M. J. & Botstein, D. (2004) *Mol. Biol. Cell* **15**, 4089–4104.
- Friedman, N., Vardi, S., Ronen, M., Alon, U. & Stavans, J. (2005) *PLoS Biol.* **3**, e238.
- Lahav, G., Rosenfeld, N., Sigal, A., Geva-Zatorsky, N., Levine, A. J., Elowitz, M. B. & Alon, U. (2004) *Nat. Genet.* **36**, 147–150.
- Schmidt, M. C., McCartney, R. R., Zhang, X., Tillman, T. S., Solimeo, H., Wolf, S., Almonte, C. & Watkins, S. C. (1999) *Mol. Cell. Biol.* **19**, 4561–4571.
- Lakshmanan, J., Mosley, A. L. & Ozcan, S. (2003) *Curr. Genet.* **44**, 19–25.
- Palomino, A., Herrero, P. & Moreno, F. (2005) *Biochem J.* **388**, 697–703.
- Wang, Y., Liu, C. L., Storey, J. D., Tibshirani, R. J., Herschlag, D. & Brown, P. O. (2002) *Proc. Natl. Acad. Sci. USA* **99**, 5860–5865.
- Andrade, R. P., Kotter, P., Entian, K. D. & Casal, M. (2005) *Biochem. Biophys. Res. Commun.* **332**, 254–262.
- Spellman, P. T., Sherlock, G., Zhang, M. Q., Iyer, V. R., Anders, K., Eisen, M. B., Brown, P. O., Botstein, D. & Futcher, B. (1998) *Mol. Biol. Cell* **9**, 3273–3297.
- Sherlock, G., Hernandez-Boussard, T., Kasarskis, A., Binkley, G., Matese, J. C., Dwight, S. S., Kaloper, M., Weng, S., Jin, H., Ball, C. A., et al. (2001) *Nucleic Acids Res.* **29**, 152–155.
- Troyanskaya, O., Cantor, M., Sherlock, G., Brown, P., Hastie, T., Tibshirani, R., Botstein, D. & Altman, R. B. (2001) *Bioinformatics* **17**, 520–525.