Exploring the Metabolic and Genetic Control of Gene Expression on a Genomic Scale

Joseph L. DeRisi, Vishwanath R. Iyer, Patrick O. Brown*  

DNA microarrays containing virtually every gene of *Saccharomyces cerevisiae* were used to carry out a comprehensive investigation of the temporal program of gene expression accompanying the metabolic shift from fermentation to respiration. The expression profiles observed for genes with known metabolic functions pointed to features of the metabolic reprogramming that occur during the diauxic shift, and the expression patterns of many previously uncharacterized genes provided clues to their possible functions. The same DNA microarrays were also used to identify genes whose expression was affected by deletion of the transcriptional co-repressor *TUP1* or overexpression of the transcriptional activator *YAP1*. These results demonstrate the feasibility and utility of this approach to genomewide exploration of gene expression patterns.

The complete sequences of nearly a dozen microbial genomes are known, and in the next several years we expect to know the complete genome sequences of several metazoa, including the human genome. Defining the role of each gene in these genomes will be a formidable task, and understanding how the genome functions as a whole in the complex natural history of a living organism presents an even greater challenge.

Knowing when and where a gene is expressed often provides a strong clue as to its biological role. Conversely, the pattern of genes expressed in a cell can provide detailed information about its state. Although regulation of protein abundance in a cell is by no means accomplished solely by regulation of mRNA, virtually all differences in cell type or state are correlated with changes in the mRNA levels of many genes. This is fortuitous because the only specific reagent required to measure the abundance of the mRNA for a specific gene is a cDNA sequence. DNA microarrays, consisting of thousands of individual gene sequences printed in a high-density array on a glass microscope slide (1, 2), provide a practical and economical tool for studying gene expression on a very large scale (3–6).

*Saccharomyces cerevisiae* is an especially favorable organism in which to conduct a systematic investigation of gene expression. The genes are easy to recognize in the genome sequence, cis regulatory elements are generally compact and close to the transcription units, much is already known about its genetic regulatory mechanisms, and a powerful set of tools is available for its analysis.

A recurring cycle in the natural history of yeast involves a shift from anaerobic (fermentation) to aerobic (respiration) metabolism. Inoculation of yeast into a medium rich in sugar is followed by rapid growth fueled by fermentation, with the production of ethanol. When the fermentable sugar is exhausted, the yeast cells turn to ethanol as a carbon source for aerobic growth. This switch from anaerobic growth to aerobic respiration upon depletion of glucose, referred to as the diauxic shift, is correlated with widespread changes in the expression of genes involved in fundamental cellular processes such as carbon metabolism, protein synthesis, and carbohydrate storage (7). We used DNA microarrays to characterize the changes in gene expression that take place during this process for nearly the entire genome, and to investigate the genetic circuitry that regulates and executes this program.

Yeast open reading frames (ORFs) were amplified by the polymerase chain reaction (PCR), with a commercially available set of primer pairs (8). DNA microarrays, containing approximately 6400 distinct DNA sequences, were printed onto glass slides by using a simple robotic printing device (9). Cells from an exponentially growing culture of yeast were inoculated into fresh medium and grown at 30°C for 21 hours. After an initial 9 hours of growth, samples were harvested at seven successive 2-hour intervals, and mRNA was isolated (10). Fluorescently labeled cDNA was prepared by reverse transcription in the presence of Cy3(green)- or Cy5(red)-labeled deoxyuridine triphosphate (dUTP) (11) and then hybridized to the microarrays (12). To maximize the reliability with which changes in expression levels could be discerned, we labeled cDNA prepared from cells at each successive time point with Cy5, then mixed it with a Cy3-labeled “reference” cDNA sample prepared from cells harvested at the first interval after inoculation. In this experimental design, the relative fluorescence intensity measured for the Cy3 and Cy5 fluoros at each array element provides a reliable measure of the relative abundance of the corresponding mRNA in the two cell populations (Fig. 1). Data from the series of seven samples (Fig. 2), consisting of more than 43,000 expression-ratio measurements, were organized into a database to facilitate efficient exploration and analysis of the results. This database is publicly available on the Internet (13).

During experimental growth in glucose-rich medium, the global pattern of gene expression was remarkably stable. Indeed, when gene expression patterns between the first two cell samples (harvested at a 2-hour interval) were compared, mRNA levels differed by a factor of 2 or more for only 19 genes (0.3%), and the largest of these differences was only 2.7-fold (14). However, as glucose was progressively depleted from the growth media during the course of the experiment, a marked change was seen in the global pattern of gene expression. mRNA levels for approximately 710 genes were induced by a factor of at least 2, and the mRNA levels for approximately 1030 genes declined by a factor of at least 2. Messenger RNA levels for 183 genes increased by a factor of at least 4, and mRNA levels for 203 genes diminished by a factor of at least 4. About half of these differentially expressed genes have no currently recognized function and are not yet named. Indeed, more than 400 of the differentially expressed genes have no apparent homology.  

---

*To whom correspondence should be addressed. E-mail: pibrown@cmgm.stanford.edu*  


---

37. M. Ho et al., Cell 77, 869 (1994).  
39. We thank H. Skalitsky and F. Lewitt for help with sequence analysis; Lawrence Livermore National Laboratory for the flow-sorted Y cosmid library; and P. Bain, A. Bortvin, A. de la Chapelle, G. Fink, K. Jegalian, T. Kawaguchi, E. Lander, H. Lodish, P. Matsudaira, D. Menke, U. Rajbhandary, R. Rejo, S. Rozen, A. Schwartz, C. Sun, and C. Tifford for comments on the manuscript. Supported by NIH.  
28 April 1997; accepted 9 September 1997.
The global view of changes in expression of genes with known functions provides a vivid picture of the way in which the cell adapts to a changing environment. Figure 3 shows a portion of the yeast metabolic pathways involved in carbon and energy metabolism. Mapping the changes we observed in the mRNAs encoding each enzyme onto this framework allowed us to infer the redirection in the flow of metabolites through this system. We observed large inductions of the genes coding for the enzymes aldehyde dehydrogenase (ALD2) and acetyl–coenzyme A(CoA) synthase (ACS1), which function together to convert the products of alcohol dehydrogenase into acetyl-CoA, which in turn is used to fuel the tricarboxylic acid (TCA) cycle and the glyoxylate cycle. The concomitant shutdown of transcription of the genes encoding pyruvate decarboxylase and induction of pyruvate carboxylase rechannels pyruvate away from acetaldehyde, and instead to oxalacetate, where it can serve to supply the TCA cycle and gluconeogenesis. Induction of the pivotal genes PCK1, encoding phosphoenolpyruvate carboxykinase, and FBP1, encoding fructose 1,6-bisphatase, switches the directions of two key irreversible steps in glycolysis, reversing the flow of metabolites along the reversible steps of the glycolytic pathway toward the essential biosynthetic precursor, glucose-6-phosphate. Induction of the genes coding for the trehalose synthase and glycogen synthase complexes promotes channeling of glucose-6-phosphate into these carbohydrate storage pathways.

Just as the changes in expression of genes encoding pivotal enzymes can provide insight into metabolic reprogramming, the behavior of large groups of functionally related genes can provide a broad view of the systematic way in which the yeast cell adapts to a changing environment (Fig. 4). Several classes of genes, such as cytochrome c-related genes and those involved in the TCA/glyoxylate cycle and carbohydrate storage, were coordinately induced by glucose exhaustion. In contrast, genes devoted to protein synthesis, including ribosomal proteins, tRNA synthetases, and translation, elongation, and initiation factors, exhibited a coordinated decrease in expression. More than 95% of ribosomal genes showed at least twofold decreases in expression during the diauxic shift (Fig. 4) (13). A noteworthy and illuminating exception was that the genes encoding mitochondrial ribosomal genes were generally induced rather than repressed after glucose limitation, highlighting the requirement for mitochondrial biogenesis (13). As more is learned about the functions of every gene in the yeast genome, the ability to gain insight into a cell’s response to a changing environment through its global gene expression patterns will become increasingly powerful.

Several distinct temporal patterns of expression could be recognized, and sets of genes could be grouped on the basis of the similarities in their expression patterns. The characterized members of each of these groups also showed important similarities in their function. Moreover, in most cases, common regulatory mechanisms could be inferred for sets of genes with similar expression profiles. For example, seven genes showed a late induction profile, with mRNA levels increasing by more than ninefold at the last timepoint but less than threefold at the preceding timepoint (Fig. 5B). All of these genes were known to be glucose-repressed, and five of the seven were previously noted to share a common upstream activating sequence (UAS), the carbon source response element (CSRE) (16–20). A search in the promoter regions of the remaining two genes, IDP2 and ACR1, revealed that ACRI, a gene essential for ACS1 activity, also possessed a consensus CSRE motif, but interestingly, IDP2 did not. A search of the entire yeast genome sequence for the consensus CSRE motif revealed only four additional candidate genes, none of which showed a similar induction.

Examples from additional groups of genes that shared expression profiles are illustrated in Fig. 5, C through F. The sequences upstream of the named genes in Fig. 5C all contain stress response elements (STRE), and with the exception

---

**Fig. 1.** Yeast genome microarray. The actual size of the microarray is 18 mm by 18 mm. The microarray was printed as described (9). This image was obtained with the same fluorescent scanning confocal microscope used to collect all the data we report (49). A fluorescein-labeled cDNA probe was prepared from mRNA isolated from cells harvested shortly after inoculation (culture density of <5 × 10^6 cells/ml and media glucose level of 19 g/liter) by reverse transcription in the presence of Cy3-dUTP (Fig. 5B). Similarly, a second probe was prepared from mRNA isolated from cells taken from the same culture 9.5 hours later (culture density of ~2 × 10^6 cells/ml, with a glucose level of <0.2 g/liter) by reverse transcription in the presence of Cy5-dUTP. In this image, hybridization of the Cy3-dUTP-labeled cDNA (that is, mRNA expression at the initial timepoint) is represented as a green signal, and hybridization of Cy5-dUTP-labeled cDNA (that is, mRNA expression at 9.5 hours) is represented as a red signal. Thus, genes induced or repressed after the diauxic shift appear in this image as red and green spots, respectively. Genes expressed at roughly equal levels before and after the diauxic shift appear in this image as yellow spots.
of HSP42, have previously been shown to be controlled at least in part by these elements (21–24). Inspection of the sequences upstream of HSP42 and the two uncharacterized genes shown in Fig. 5C, YKL026c, a hypothetical protein with similarity to glutathione peroxidase, and YGR043c, a putative transaldolase, revealed that each of these genes also possesses repeated upstream copies of the stress-responsive CCCCCT motif. Of the 13 additional genes in the yeast genome that shared this expression profile [including HSP30, ALD2, OM45, and 10 uncharacterized ORFs (25)], nine contained one or more recognizable STRE sites in their upstream regions.

The heterotrimeric transcriptional activator complex HAP2,3,4 has been shown to be responsible for induction of several genes important for respiration (26–28). This complex binds a degenerate consensus sequence known as the CCAAT box (26). Computer analysis, using the consensus sequence TNRYTGGB (29), has suggested that a large number of genes involved in respiration may be specific targets of HAP2,3,4 (30). Indeed, a putative HAP2,3,4 binding site could be found in the sequences upstream of each of the seven cytochrome c-related genes that showed the greatest magnitude of induction (Fig. 5D). Of 12 additional cytochrome c-related genes that were induced, HAP2,3,4 binding sites were present in all but one. Significantly, we found that transcription of HAP4 itself was induced nearly ninefold coincident with the diauxic shift.

Control of ribosomal protein biogenesis is mainly exerted at the transcriptional level, through the presence of a common upstream-activating element (UAS$_{rs}$) that is recognized by the Rap1 DNA-binding protein (31, 32). The expression profiles of seven ribosomal proteins are shown in Fig. 5F. A search of the sequences upstream of all seven genes revealed consensus Rap1-binding motifs (33). It has been suggested that declining Rap1 levels in the cell during starvation may be responsible for the decline in ribosomal protein gene expression (34). Indeed, we observed that the abundance of RAP1 mRNA diminished by 4.4-fold, at about the time of glucose exhaustion.

Of the 149 genes that encode known or putative transcription factors, only two, HAP4 and SIP4, were induced by a factor of more than threefold at the diauxic shift. SIP4 encodes a DNA-binding transcriptional activator that has been shown to interact with Sml1, the “master regulator” of glucose repression (35). The eightfold induction of SIP4 upon depletion of glucose strongly suggests a role in the induction of downstream genes at the diauxic shift.

Although most of the transcriptional responses that we observed were not previously known, the responses of many genes during the diauxic shift have been described. Comparison of the results we obtained by DNA microarray hybridization with previously reported results therefore provided a strong test of the sensitivity and accuracy of this approach. The expression patterns we observed for previously characterized genes showed almost perfect concordance with previously published results (36). Moreover, the differential expression measurements obtained by DNA microarray hybridization were reproducible in duplicate experiments. For example, the remarkable changes in gene expression between cells harvested immediately after inoculation and immediately after the diauxic shift (the first and sixth intervals in this time series) were measured in duplicate, independent DNA microarray hybridizations. The correlation coefficient for two complete sets of expression ratio measurements was 0.87, and for more than 95% of the genes, the expression ratios measured in these duplicate experiments differed by less than a factor of 2. However, in a few cases, there were discrepancies between our results and previous results, pointing to technical limitations that will need to be addressed as DNA microarray technology advances (37, 38). Despite the noted exceptions, the high concordance between the results we obtained in these experiments and those of previous studies provides confidence in the reliability and thoroughness of the survey.

The changes in gene expression during this diauxic shift are complex and involve integration of many kinds of information about the nutritional and metabolic state of the cell. The large number of genes whose expression is altered and the diversity of temporal expression profiles observed in this experiment highlight the challenge of understanding the underlying regulatory mechanisms. One approach to defining the contributions of individual regulatory genes to a complex program of this kind is to use DNA microarrays to identify genes whose expression is affected

---

*Fig. 2.* The section of the array indicated by the gray box in Fig. 1 is shown for each of the experiments described here. Representative genes are labeled. In each of the arrays used to analyze gene expression during the diauxic shift, red spots represent genes that were induced relative to the initial timepoint, and green spots represent genes that were repressed relative to the initial timepoint. In the arrays used to analyze the effects of the tup1Δ mutation and YAP1 overexpression, red spots represent genes whose expression was increased, and green spots represent genes whose expression was decreased by the genetic modification. Note that distinct sets of genes are induced and repressed in the different experiments. The complete images of each of these arrays can be viewed on the Internet (13). Cell density as measured by optical density (OD) at 600 nm was used to measure the growth of the culture.
by mutations in each putative regulatory gene. As a test of this strategy, we analyzed the genomewide changes in gene expression that result from deletion of the TUP1 gene. Transcriptional repression of many genes by glucose requires the DNA-binding repressor Mig1 and is mediated by recruiting the transcriptional co-repressors Tup1 and Cyc8/Ssn6 (39). Tup1 has also been implicated in repression of oxygen-regulated, mating-type–specific, and DNA-damage-inducible genes (40).

Wild-type yeast cells and cells bearing a deletion of the TUP1 gene (tup1Δ) were grown in parallel cultures in rich medium containing glucose as the carbon source. Messenger RNA was isolated from exponentially growing cells from the two populations and used to prepare cDNA labeled with Cy3 (green) and Cy5 (red), respectively (11). The labeled probes were mixed and simultaneously hybridized to the microarray. Red spots on the microarray therefore represented genes whose transcription was induced in the tup1Δ strain, and thus presumably repressed by Tup1 (41). A representative section of the microarray (Fig. 2, bottom middle panel) illustrates that the genes whose expression was affected by the tup1Δ mutation were, in general, distinct from those induced upon glucose exhaustion [complete images of all the arrays shown in Fig. 2 are available on the Internet (13)]. Nevertheless, 34 (10%) of the genes that were induced by a factor of at least 2 after the diauxic shift were similarly induced by deletion of TUP1, suggesting that these genes may be subject to TUP1-mediated repression by glucose. For example, SUC2, the gene encoding invertase, and all five hexose transporter genes that were induced during the course of the diauxic shift were similarly induced, in duplicate experiments, by the deletion of TUP1.

The set of genes affected by Tup1 in this experiment also included α-glucosidases, the mating-type–specific genes MFA1 and MFA2, and the DNA damage–inducible RNR2 and RNR4, as well as genes involved in flocculation and many genes of unknown function. The hybridization signal corresponding to expression of Tup1 itself was also severely reduced because of the (incomplete) deletion of the transcription unit in the tup1Δ strain, providing a positive control in the experiment (42).

Many of the transcriptional targets of Tup1 fell into sets of genes with related biochemical functions. For instance, although only about 3% of all yeast genes appeared to be TUP1-repressed by a factor of more than 2 in duplicate experiments under these conditions, 6 of the 13 genes that have been implicated in flocculation (15) showed a reproducible increase in expression of at least twofold when TUP1 was deleted. Another group of related genes that appeared to be subject to TUP1 repression encodes the serine-rich cell wall mannoproteins, such as Tip1 and Tir1/Srp1 which are induced by cold shock and other stresses (43), and similar, serine-poor proteins, the seripauperins (44). Messenger RNA levels for 23 of the 26 genes in this group were reproducibly elevated by at least 2.5-fold in the tup1Δ

---

**Fig. 3.** Metabolic reprogramming inferred from global analysis of changes in gene expression. Only key metabolic intermediates are identified. The yeast genes encoding the enzymes that catalyze each step in this metabolic circuit are identified by name in the boxes. The genes encoding succinyl-CoA synthase and glycogen-debranching enzyme have not been explicitly identified, but the ORFs YGR244 and YPR184 show significant homology to known succinyl-CoA synthase and glycogen-debranching enzymes, respectively, and are therefore included in the corresponding steps in this figure. Red boxes with white lettering identify genes whose expression increases in the diauxic shift. Green boxes with dark green lettering identify genes whose expression diminishes in the diauxic shift. The magnitude of induction or repression is indicated for these genes. For multimeric enzyme complexes, such as succinate dehydrogenase, the indicated fold-induction represents an unweighted average of all the genes listed in the box. Black and white boxes indicate no significant differential expression (less than twofold). The direction of the arrows connecting reversible enzymatic steps indicate the direction of the flow of metabolic intermediates, inferred from the gene expression pattern, after the diauxic shift. Arrows representing steps catalyzed by genes whose expression was strongly induced are highlighted in red. A broad gray arrow between the flow of metabolites after the diauxic shift, inferred from the indicated changes in gene expression.
strain, and 18 of these genes were induced by more than sevenfold when TUP1 was deleted. In contrast, none of 83 genes that could be classified as putative regulators of the cell division cycle were induced more than twofold by deletion of TUP1. Thus, despite the diversity of the regulatory systems that employ Tup1, most of the genes that it regulates under these conditions fall into a limited number of distinct functional classes.

Because the microarray allows us to monitor expression of nearly every gene in yeast, we can, in principle, use this approach to identify all the transcriptional targets of a regulatory protein like Tup1. It is important to note, however, that in any single experiment of this kind we can only recognize those target genes that are normally repressed (or induced) under the conditions of the experiment. For instance, the experiment described here analyzed a MATα strain in which MFA1 and MFA2, the genes encoding the α-factor mating pheromone precursor, are normally repressed. In the isogenic tup1Δ strain, these genes were inappropriately expressed, reflecting the role that Tup1 plays in their repression. Had we instead carried out this experiment with a MATα strain (in which expression of MFA1 and MFA2 is not repressed), it would not have been possible to conclude anything regarding the role of Tup1 in the repression of these genes. Conversely, we cannot distinguish indirect effects of the chronic absence of Tup1 in the mutant strain from effects directly attributable to its participation in repressing the transcription of a gene.

Another simple route to modulating the activity of a regulatory factor is to overexpress the gene that encodes it. YAP1 encodes a DNA-binding transcription factor belonging to the bZIP class of DNA-binding proteins. Overexpression of YAP1 in yeast confers increased resistance to hydrogen peroxide, o-phenanthroline, heavy metals, and osmotic stress (45). We analyzed differential gene expression between a wild-type strain bearing a control plasmid and a strain with a plasmid expressing YAP1 under the control of the strong GAL1-10 promoter, both grown in galactose (that is, a condition that induces YAP1 overexpression). Complementary DNA from the control and YAP1 overexpressing strains, labeled with Cy3 and Cy5, respectively, was prepared from mRNA isolated from the two strains and hybridized to the microarray. Thus, red spots on the array represent genes that were induced in the strain overexpressing YAP1.

Of the 17 genes whose mRNA levels increased by more than threefold when YAP1 was overexpressed in this way, five bear homology to aryl-alcohol oxidoreductases (Fig. 2 and Table 1). An additional four of the genes in this set also belong to the general class of dehydrogenases/oxidoreductases. Very little is known about the role of aryl-alcohol oxidoreductases in S. cerevisiae, but these enzymes have been isolated from liginolytic fungi, in which they participate in coupled redox reactions, oxidizing aromatic, and alphabetic unsaturated alcohols to aldehydes with the production of hydrogen peroxide (46, 47). The fact that a remarkable fraction of the targets identified in this experiment belong to the same small, functional group of oxidoreductases suggests that these genes might play an important protective role during oxidative stress. Transcription of a small number of genes was reduced in the strain overexpressing Yap1. Interestingly, many of these genes encode sugar permeases or enzymes involved in inositol metabolism.

We searched for Yap1-binding sites (TTACTAA or TGACTAA) in the sequences upstream of the target genes we identified (48). About two-thirds of the genes that were induced by more than threefold upon Yap1 overexpression had one or more binding sites within 600 bases upstream of the start codon (Table 1), suggesting that they are directly regulated by Yap1. The absence of canonical Yap1-bind-

<table>
<thead>
<tr>
<th>ORF</th>
<th>Distance of Yap1 site from ATG</th>
<th>Gene Description</th>
<th>Fold-increase</th>
</tr>
</thead>
<tbody>
<tr>
<td>YNL331C</td>
<td>162–222 (5 sites)</td>
<td>Putative aryl-alcohol reductase</td>
<td>12.9</td>
</tr>
<tr>
<td>YKL071W</td>
<td>98</td>
<td>Similarity to bacterial csqA protein</td>
<td>10.4</td>
</tr>
<tr>
<td>YML007W</td>
<td>223, 242</td>
<td>Transcriptional activator involved in oxidative stress response</td>
<td>9.8</td>
</tr>
<tr>
<td>YFL056C</td>
<td>98</td>
<td>Homology to aryl-alcohol dehydrogenases</td>
<td>9.0</td>
</tr>
<tr>
<td>YLL060C</td>
<td>266</td>
<td>Putative glutathione transferase</td>
<td>7.4</td>
</tr>
<tr>
<td>YOL165C</td>
<td>409</td>
<td>Putative aryl-alcohol dehydrogenase (NADP+−)</td>
<td>7.0</td>
</tr>
<tr>
<td>YCR107W</td>
<td>142, 167, 364</td>
<td>Putative aryl-alcohol reductase</td>
<td>6.5</td>
</tr>
<tr>
<td>YML116W</td>
<td>142, 167, 364</td>
<td>Aminotiazole and 4-nitroquinoline resistance protein</td>
<td>6.5</td>
</tr>
<tr>
<td>YBR008C</td>
<td>98</td>
<td>Homology to benomyl/methotrexate resistance protein</td>
<td>6.1</td>
</tr>
<tr>
<td>YCLX08C</td>
<td>98</td>
<td>Hypothetical protein</td>
<td>6.1</td>
</tr>
<tr>
<td>YJR155W</td>
<td>148, 212</td>
<td>Putative aryl-alcohol dehydrogenase</td>
<td>6.0</td>
</tr>
<tr>
<td>YPL171C</td>
<td>167, 317</td>
<td>NAPDH dehydrogenase (old yellow enzyme), isoform 3</td>
<td>5.8</td>
</tr>
<tr>
<td>YLR460C</td>
<td>167, 317</td>
<td>Homology to hypothetical proteins</td>
<td>4.7</td>
</tr>
<tr>
<td>YKRO76W</td>
<td>178</td>
<td>YOR102c and YNL134c</td>
<td>4.5</td>
</tr>
<tr>
<td>YHR179W</td>
<td>327</td>
<td>Homology to hypothetical protein</td>
<td>4.1</td>
</tr>
<tr>
<td>YML131W</td>
<td>507</td>
<td>NAD(P)H oxidoreductase (old yellow enzyme), isoform 1</td>
<td>3.7</td>
</tr>
<tr>
<td>YOL126C</td>
<td>507</td>
<td>Malate dehydrogenase</td>
<td>3.3</td>
</tr>
</tbody>
</table>
ing sites upstream of the others may reflect an ability of Yap1 to bind sites that differ from the canonical binding sites, perhaps in cooperation with other factors, or less likely, may represent an indirect effect of Yap1 overexpression, mediated by one or more intermediary factors. Yap1 sites were found only four times in the corresponding region of an arbitrary set of 30 genes that were not differentially regulated by Yap1.

Use of a DNA microarray to characterize the transcriptional consequences of mutations affecting the activity of regulatory molecules provides a simple and powerful approach to dissection and characterization of regulatory pathways and networks. This strategy also has an important practical application in drug screening. Mutations in specific genes encoding candidate drug targets can serve as surrogates for the ideal chemical inhibitor or modulator of their activity. DNA microarrays can be used to define the resulting signature pattern of alterations in gene expression, and then subsequently used in an assay to screen for compounds that reproduce the desired signature pattern.

DNA microarrays provide a simple and economical way to explore gene expression patterns on a genomic scale. The hurdles to extending this approach to any other organism are minor. The equipment required for fabricating and using DNA microarrays (9) consists of components that were chosen for their modest cost and simplicity. It was feasible for a small group to accomplish the amplification of more than 6000 genes in about 4 months and, once the amplified gene sequences were in hand, only 2 days were required to print a set of 110 microarrays of 6400 elements each. Probe preparation, hybridization, and fluorescent imaging are also simple procedures. Even conceptually simple experiments, as we described here, can yield vast amounts of information. The value of the information from each experiment of this kind will progressively increase as more is learned about the functions of each gene and as additional experiments define the global changes in gene expression in diverse other natural processes and genetic perturbations. Perhaps the greatest challenge now is to develop efficient methods for organizing, distributing, interpreting, and extracting insights from the large volumes of data these experiments will provide.

Fig. 5. Distinct temporal patterns of induction or repression help to group genes that share regulatory properties. (A) Temporal profile of the cell density, as measured by OD at 600 nm and glucose concentration in the media. (B) Seven genes exhibited a strong induction (greater than ninefold) only at the last timepoint (20.5 hours). With the exception of IDP2, each of these genes has a CSRE UAS. There were no additional genes observed to match this profile. (C) Seven members of a class of genes marked by early induction with a peak in mRNA levels at 18.5 hours. Each of these genes contain STRE motif repeats in their upstream promoter regions. (D) Cytochrome c oxidase and ubiquinol cytochrome c reductase genes. Marked by an induction coincident with the diauxic shift, each of these genes contains a consensus binding motif for the HAP2,3,4 protein complex. At least 17 genes shared a similar expression profile. (E) SAM1, GPP1, and several genes of unknown function are repressed before the diauxic shift, and continue to be repressed upon entry into stationary phase. (F) Ribosomal protein genes comprise a large class of genes that are repressed upon depletion of glucose. Each of the genes profiled here contains one or more RAP1-binding motifs upstream of its promoter. RAP1 is a transcriptional regulator of most ribosomal proteins.

REFERENCES AND NOTES

2. D. Shalon, S. J. Smith, P. O. Brown, Genome Res. 6, 639 (1996).
8. Primers for each known or predicted protein coding sequence were supplied by Research Genetics. PCR was performed with the protocol supplied by Research Genetics, using genomic DNA from yeast strain S288C as a template. Each PCR product was verified by agarose gel electrophoresis and was deemed correct if the lane contained a single band of appropriate mobility. Failures were marked as such in the database. The overall success rate for a single-pass amplification of 6116 ORFs was ~94.5%.
9. Glass slides (Gold Seal) were cleaned for 2 hours in a solution of 2 N NaOH and 70% ethanol. After rinsing in distilled water, the slides were then treated with a 1:5 dilution of poly-L-lysine adhesive solution (Sigma) for 1 hour, and then dried for 5 min at 40°C in a vacuum oven. DNA samples from 100-μl PCR reactions were purified by ethanol purification in 96-well microtiter plates. The resulting precipitates were resuspended in 3× standard saline citrate (SSC) and transferred to new plates for arraying. A custom-built arraying robot was used to print on a batch of 110 slides. Details of the design of the microarrayer are available at cgrng.stanford.edu/pbrown. After printing, the microarrays were rehydrated for 30 s in a hummid chamber and then snap-dried for 2 s on a hot plate (100°C). The DNA was then ultraviolet (UV)-crosslinked to the surface by subjecting the slides to 162 J/cm² of energy. STRATEGIES: The rest of the poly-L-lysine surface was blocked by a 15-min incubation in a solution of 70 mM sucinic anhydride dissolved in a solution consisting of 0.1 M 1-methylimidazole (Aldrich) and 35 ml of 1 M boric acid (pH 8.0). Directly after the blocking reac-
tion, the bound DNA was denatured by a 2-min incubation in distilled water at –95°C. The slides were then transferred into a bath of 100% ethanol at room temperature, rinsed, and then spun dry in a clinical centrifuge. Slides were stored in a closed box at room temperature until used.

10.YPD medium (8 liters), in a 10-liter fermentation vessel, was inoculated with 2 ml of a fresh overnight culture of yeast strain DBY7286 (MATa, ura3, gal2). The fermentor was maintained at 30°C with constant agitation and aeration. The glucose content of the media was measured with a UV test kit (Boehringer Mannheim, catalog number 716251). Cell density was measured by OD at 600-nm wavelength. Aliquots of culture were rapidly withdrawn from the fermentation vessel by peristatic pump, spun down at room temperature, and then flash frozen with liquid nitrogen. Frozen cells were stored at –80°C.

11. Cycl-dUTP or Cy5-dUTP (Amersham) was incorporated during reverse transcription of 1.25 μg of polyadenylated [poly(A)] RNA, primed by a dT(16) oligomer. This mixture was heated to 70°C for 10 min, and then transferred to ice. A premixed solution, consisting of 200 U Superscript II ( Gibco), buffer, dexteroxyribonucleoside triphosphates, and fluorescein isothiocyanate (FITC) added to the RNA, was used to determine the completeness of the reverse transcription. The reaction was then incubated at 42°C for 2 hours. UV-irradiated fluorescent nucleotides were removed by first diluting the reaction mixture with 470 μl of 10 mM MOPS (pH 7.0)/1 mM EDTA and then incubating at 42°C for 2 hours. Unincorporated fluorescent nucleotides were removed by first diluting the reaction mixture with 470 μl of 10 mM MOPS (pH 7.0)/1 mM EDTA and then subsequently concentrating the mix to ~5 μl, using Centricon-30 microconcentrators (Amicon).

12. Purified, labeled cDNA was resuspended in 1 ml of 3.5× SSC containing 10 μg poly(dA) and 0.3 μl of 10% SDS. Before hybridization, the solution was boiled for 2 min and then allowed to cool to room temperature. The solution was applied to the microarray under a cover slip, and the slide was placed in a custom hybridization chamber which was subsequently incubated for ~8 to 12 hours in a water bath at 62°C. Before scanning, slides were washed in 2× SSC, 0.2% SDS for 5 min, and then 0.05× SSC for 1 min. After washing, slides were dried before scanning by centrifugation at 500 rpm in a Beckman CS-6R centrifuge.

13. The complete data set is available on the Internet at the following public databases: Saccharomyces Genome Database (genome-www.stanford.edu), Munich Information Centre for Protein Sequences (speedy.mips.biochem.mpg.de), and Munich Information Centre for Protein Sequences, should relieve this problem in many cases.

14. For 95% of all the genes analyzed, the mRNA levels measured in cells harvested at the first and second interval after inoculation differed by a factor of less than 2. The correlation coefficient for the comparison between mRNA levels measured for each gene in these two different mRNA samples was 0.98. When duplicate mRNA preparations from the same cell sample were compared in the same way, the correlation coefficient between the expression levels measured for the two samples by comparative hybridization was 0.91.

15. The numbers and identities of known and putative genes, and their homologies to other genes, were gathered from the following public databases: Saccharomyces Genome Database (genome-www.stanford.edu), Munich Information Centre for Protein Sequences (speedy.mips.biochem.mpg.de), and Munich Information Centre for Protein Sequences (speedy.mips.biochem.mpg.de) (https://index/index.html).


