AMPARs in Bergmann glia are necessary for preventing surplus CFs from producing synapses onto single Purkinje cells. Conversion of Ca^2+–permeable AMPARs into Ca^2+–impermeable receptors in Bergmann glia elicited morphological changes in fine glial processes wrapping Purkinje cell synapses, prolonged the kinetics of glutamatergic synaptic transmission, and caused multiple innervation of Purkinje cells by CFs. Thus, the morphology of glial processes and the synaptic activities would be interdependent. The Ca^2+–permeable AMPARs in glial cells probably play key roles in such interactions between glia and glutamatergic synapses.

We demonstrate an integrated approach to build, test, and refine a model of a cellular pathway, in which perturbations to critical pathway components are analyzed using DNA microarrays, quantitative proteomics, and databases of known physical interactions. Using this approach, we identify 997 messenger RNAs responding to 20 systematic perturbations of the yeast galactose-utilization pathway, provide evidence that approximately 15 of 289 detected proteins are regulated posttranscriptionally, and identify explicit physical interactions governing the cellular response to each perturbation. We refine the model through further iterations of perturbation and global measurements, suggesting hypotheses about the regulation of galactose utilization and physical interactions between this and a variety of other metabolic pathways.

For organisms with fully sequenced genomes, DNA microarrays are an extremely powerful technology for measuring the mRNA expression responses of practically every gene (1). Technologies for globally and quantitatively measuring protein expression are also becoming feasible (2), and developments such as the two-hybrid system are enabling construction of a map of interactions among proteins (3). Although such large-scale data have proven invaluable for distinguishing cell types and biological states, new approaches are needed which, by integrating these diverse data types and assimilating them into biological models, can predict cellular behaviors that can be tested experimentally. We propose and apply one such strategy here, consisting of four distinct steps:

(i) Define all of the genes in the genome and the subset of genes, proteins, and other small molecules constituting the pathway of interest. If possible, define an initial model of the molecular interactions governing pathway function, drawn from previous genetic and biochemical research.

(ii) Perturb each pathway component through a series of genetic (e.g., gene deletions or overexpressions) or environmental (e.g., changes in growth conditions or temperature) manipulations. Detect and quantify the corresponding global cellular response to each perturbation with technologies for large-scale mRNA- and protein-expression measurement.

(iii) Integrate the observed mRNA and protein responses with the current, pathway-specific model and with the global network of protein-protein, protein-DNA, and other known physical interactions.

(iv) Formulate new hypotheses to explain
Observations not predicted by the model. Design additional perturbation experiments to test these, and iteratively repeat steps (ii), (iii), and (iv).

As proof-of-principle, we now implement this integrated approach to explore the process of galactose utilization (GAL) in the yeast *Saccharomyces cerevisiae*. The GAL pathway is a classic example of a genetic regulatory switch, in which enzymes required specifically for transport and catabolism of galactose are expressed only when galactose is present and repressing sugars such as glucose are absent. Extensive biochemical studies (4) and saturating mutant screens (5) have defined the genes, gene products, and metabolic substrates required for function of this process and have elucidated the key molecular interactions that lead to pathway activation or inhibition. Thus, by combining this prior work with the complete sequence of the yeast genome, step (i) above has in large part already been accomplished. In steps (ii) through (iv) that follow, we report the first large-scale comparison of mRNA and protein responses, describe an ongoing attempt to systematically explain these responses using existing databases of regulatory and other physical interactions, and explore a number of refinements to the GAL model suggested by these integrative studies.

Step (i): As shown in Fig. 1, galactose utilization consists of a biochemical pathway that converts galactose into glucose-6-phosphate and a regulatory mechanism that controls whether the pathway is on or off. This process has been reviewed extensively (4, 6) and involves at least three types of proteins. A transporter gene (GAL2) encodes a permease that transports galactose into the cell; several other hexose transporters (HXTs) may also have this ability (7). A group of enzymatic genes encodes the proteins required for conversion of intracellular galactose, including galactokinase (GAL1), uridylyltransferase (GAL7), epimerase (GAL10), and phosphoglucomutase (GAL5/PGM2). The regulatory genes GAL3, GAL4, and GAL80 exert tight transcriptional control over the transporter, the enzymes, and to a certain extent, each other. GAL4p is a DNA-binding factor that can strongly activate transcription, but in the absence of galactose, GAL80p binds GAL4p and inhibits its activity. When galactose is present in the cell, it causes GAL3p to associate with GAL80p. This association causes GAL80p to release its repression of GAL4p, so that the transporter and enzymes are expressed at a high level.

Although these genes and interactions form the core of the GAL pathway, the complete regulatory mechanism is more complex (8–11) and involves genes whose roles in galactose utilization are not entirely clear (12, 13). For instance, the gene GAL6 (LAP3) functions predominantly in a drug-resistance pathway, but can suppress transcription of the GAL transporter and enzymes under certain conditions and may itself be transcriptionally controlled by GAL4 (14).

Step (ii): Guided by the current model, we applied 20 initial perturbations to the GAL pathway. Wild-type (wt) and nine genetically altered yeast strains were examined (15), each with a complete deletion of one of the nine GAL genes: transport (gal2Δ), enzymatic (gal1Δ, gal5Δ, gal7Δ, or gal10Δ), or regulatory (gal3Δ, gal4Δ, gal6Δ, or gal80Δ). These strains were perturbed environmentally by growth in the presence (+gal) or absence (–gal) of 2% galactose, with 2% raffinose provided in both media (16).

We examined global changes in mRNA
expression resulting from each perturbation, with DNA microarrays of approximately 6200 nuclear yeast genes as described (17). In each experiment, fluorescently labeled cDNA from a perturbed strain was hybridized against labeled cDNA from a reference strain (wt, grown in + gal media). To obtain robust estimates of fluorescent intensity, four replicate hybridizations were performed for each perturbation. Using a statistical method based on maximum-likelihood estimation (18), we identified 997 genes whose mRNA levels differed significantly from reference under one or more perturbations. This set was then divided into 16 clusters using self-organizing maps (19), where each cluster contained genes with similar expression responses over all perturbations. Figure 2 displays a matrix summarizing the effects of perturbation on mRNA expression of the GAL genes and gene clusters [complete data provided in Web table 1 (20)].

Are the observed changes in mRNA expression also reflected at the level of protein abundance? To address this question, we examined differences in protein abundance between wt+gal and wt–gal conditions using isotope-coded affinity tag (ICAT) reagents and tandem mass spectrometry (MS/MS) (21). Equal amounts of protein extracts from wt+gal and wt–gal cultures were labeled with isotopically heavy and normal ICAT reagents, respectively, then combined and digested with trypsin. The resulting peptide mixture was fractionated by multidimensional chromatography and analyzed by MS/MS. Computational analysis of the tandem mass spectra was used to identify the proteins from which specific peptides originated and to indicate relative abundances of the heavy and normal ICAT isoforms of each of these peptides.

We obtained protein-abundance ratios for a total of 289 proteins [Web table 1 (20)], including all of the GAL enzymes and the transporter. Figure 3 shows protein-abundance ratios versus the corresponding mRNA-expression ratios obtained with DNA microarrays: as a whole, protein-abundance ratios were moderately correlated with their mRNA counterparts (r = 0.61, P < 1.3 × 10 ^{-30}). Although approximately 30 proteins displayed clear changes in abundance between the wt+gal and wt–gal conditions (log_{10} ratio > 0.25), mRNA levels for 15 of these did not change significantly in response to any perturbation, suggesting that these proteins may be regulated posttranscriptionally. In addition, many ribosomal-protein genes increased three- to fivefold in mRNA but not in protein abundance in response to galactose addition. These results underscore the importance of integrated mRNA- and protein-expression measurements for understanding biological systems.

Step (iii): Can we attribute the observed mRNA and protein changes to underlying regulatory interactions in the cell? Although we already have a model of interactions among the GAL genes, it does not address changes in expression observed for the hundreds of other genes appearing in Figs. 2 and 3. To supplement this model, we assembled a catalog of previously observed physical interactions in yeast by combining a published list of 2709 protein-protein interactions (3) with 317 protein–DNA interactions recorded in the transcription-factor databases (22). A total of 348 genes associated with interactions in this catalog were affected in mRNA or protein expression by at least one perturbation or involved in two or more interactions with affected genes. Figure 4A displays these genes graphically, along with their 362 associated interactions, as a physical-interaction network.

Genes linked by physical interactions in the network tend to have more strongly correlated expression profiles than genes chosen at random (P < 0.001). We believe these correlations identify network interactions that are likely to have transmitted a change in expression from one gene (or protein) to another over our 20 perturbations. Most straightforwardly, a protein–DNA interaction may be responsible for directly transmitting an expression change from a transcription factor to a highly correlated target gene (e.g., Mcm1 → Far1 and Mig1 → Fbp1; mRNA expression profile correlations are r_{Mcm1,Far1} = 0.82 and r_{Mig1,Fbp1} = 0.63). Alternatively, genes A and B may be under control of a common transcription factor C―(A,B); coexpression of A and B provides evidence that C transmits these changes, regardless of whether C itself changes detectably in expression. This is the case for the GAL enzymes regulated by Gal4 (Fig. 4B), amino acid synthesis genes regulated by Gcn4 (Fig. 4C), and a class of gluconeogenic genes controlled by Sip4 (Sip4 → Fbp1, Pck1, Iel1). Finally, we may scan the network for indirect effects, such as a change in A transmitted to B through a protein-protein interaction with a signaling protein (e.g., Ger2–Gcr1 → Tpi1; r_{Ger2,Tpi1} = -0.86). Many other physically interacting, strongly correlated genes are listed in Web table 2 (20); each of these associates an observed change in gene expression with the regulatory interaction(s) likely to have caused it.

Ultimately, we wish to determine paths through the network connecting perturbed GAL genes to every other affected gene. This is not always possible, because many of the required interactions linking galactose utilization to other metabolic processes are still unknown. However, analysis of our expression data suggests that Gal4p directly regulates genes in several of these processes through novel protein–DNA interactions. To identify putative interactions, we looked for the well-characterized Gal4p-binding site (23) upstream of genes in expression clusters 1, 2, and 3, which together contained all seven genes with established Gal4p-binding sites. Of the 87 remaining genes in these three clusters, nine had Gal4p-binding sites not previously identified [Web table 3 (20)], a significantly greater proportion than were
found in clusters 4 through 16 (10.3% versus 2.8%; \( P < 0.002 \)). This set of nine contained genes involved in glycogen accumulation and protein metabolism as well as several genes of unknown function (e.g., YMR318C, a gene shown in Fig. 3 to have strong mRNA and protein responses to galactose induction) (24). As shown in Fig. 1, we suggest that Gal4p may regulate these genes by direct binding.

Fig. 4. Integrated physical-interaction network. Nodes represent genes, a yellow arrow directed from one node to another signifies that the protein encoded by the first gene can influence the transcription of the second by DNA binding (protein\( \rightarrow \)DNA), and a blue line between two nodes signifies that the corresponding proteins can physically interact (protein-protein). Highly interconnected groups of genes tend to have common biological function and are labeled accordingly. (A) Effects of the gal4\( \Delta \)gal perturbation are superimposed on the network, with GAL4 colored red and the gray scale intensity of other nodes representing changes in mRNA as in Fig. 2 (node diameter also scales with the magnitude of change). Regions corresponding to (B) galactose utilization and (C) amino acid synthesis are detailed at right. Graphical layout and network display were performed automatically using software based on the LEDA toolbox (37). An enlarged version of (A) is provided in (20).

Fig. 5. Tree comparing gene-expression changes resulting from different perturbations to the GAL pathway. We used the Neighbor and Drawtree programs (38) to construct a hierarchical-clustering tree (39) based on Euclidean distance between perturbation profiles, where each profile consists of log\(_{10}\) mRNA expression ratios over the set of 997 significantly affected genes. The closer two perturbations are to each other through the branches of the tree, the more similar their observed changes in gene expression. Leaves of the tree are labeled with the relevant genetic perturbation (wild-type or gene deletion) followed by the environmental perturbation (+/− gal). Twenty initial perturbations (solid branches) and three follow-up perturbations are shown (dotted branches). As in Fig. 2, profiles for all genetic perturbations are relative to that of the wild type, with both strains grown in identical media (+gal or −gal).
Step (iv): Lastly, how do the observed responses of GAL genes compare to their predicted behavior? Figure 2B shows the qualitative changes (+ and −) in mRNA expression that we predicted based on the model shown in Fig. 1 and from current knowledge of galactose utilization as summarized in Step (ii). In general, these were in good agreement with the observed changes. For example, growth of wild-type cells in +gal versus −gal media significantly induced GAL1, GAL2, GAL7, GAL10, and GAL80 as expected, while deleting the positive regulators GAL3 and GAL4 led to a significant expression decrease in many of these genes. In −gal media, deletion of the repressor GAL80 caused a dramatic increase in GAL-enzyme expression; in +gal, this deletion had little or no effect on these genes, presumably because they were already highly expressed.

A number of observations were not predicted by the model and are listed in Web Table 4 (20); in many cases, these suggest new regulatory phenomena that may be tested by hypothesis-driven approaches. For example, in the presence of galactose, gal7 and gal10 deletions unexpectedly reduced the expression levels of other GAL enzymes. Because the metabolite Gal-1-P is known to accumulate in cells lacking other GAL enzymes. Because the metabolite Gal-1-P is known to accumulate in cells lacking other GAL enzymes, this deletion had little or no effect on these genes, presumably because they were already highly expressed.

To test the hypothesis that the effects of gal7Δ and gal10Δ are dependent on increased levels of Gal-1-P or a derivative molecule, we examined the expression profile of a gal1Δgal10Δ double deletion growing in +gal conditions (relative to the wt+gal reference). We predicted that in this strain, the absence of GAL1 activity would prevent build-up of Gal-1-P and the changes in GAL gene expression would not occur. Conversely, if the expression changes did not depend on Gal-1-P (e.g., are caused by chromosomal interactions at the GAL1-10-7 locus), they would also be likely to occur in the gal1Δgal10Δ strain. Consistent with our initial hypothesis, GAL-enzyme expression was not significantly affected by this perturbation, and as shown in Fig. 5, the expression profile of gal1Δgal10Δ over all affected genes was more similar to the profile of gal1Δ+gal than to that of gal10Δ+gal or any other perturbation.

Another unanticipated observation was the slow growth of the gal80Δ mutant in −gal conditions (Fig. 2D), the large number of gene clusters affected by this perturbation (Fig. 2C, compare rightmost column to the other eight columns in the −gal set), and the corresponding large decrease in the gal80Δ/gal1 expression profile and every other profile in Fig. 5. Since this perturbation leads to constitutive expression of the GAL enzymes and transporter, we wished to determine whether the widespread expression changes were dependent on these genes. Accordingly, we measured the expression profile of a gal1Δgal80Δ gal double deletion, in which the GAL enzymes and transporter are not expressed. Both the doubling time (144 min) and overall expression profile of this strain (Fig. 5) were more similar to those of gal1Δ (129 min) than gal80Δ (205 min), suggesting that the effects of the gal80Δ perturbation are indeed mediated by other GAL genes. To further determine which GAL genes were important for this effect, we measured the expression profile of a gal1Δgal80ΔΔ gal double deletion, in which the GAL transporter was absent. This profile was more similar to that of gal1Δ than gal80ΔΔ, providing evidence that the transporter is necessary to produce the slow growth and expression changes seen for the gal80ΔΔ perturbation.

We expect that more directed experimental approaches (i.e., biochemistry, genetics, cell biology) will be required to test these ideas and further deepen our understanding of galactose utilization and its interacting networks. Even so, global and integrated analyses are extremely powerful for suggesting new hypotheses, especially with regard to the regulation of a pathway and its interconnections with other pathways. As technologies for cellular perturbation and global measurement mature, these approaches will soon become feasible in higher eukaryotes.

References and Notes
15. Strains were derived from the wild-type haploid MATa strain BY4741 (ATCC #201388, MATa his3Δ1 leu2Δ0 met15Δ0 ura3Δ0). The mutants gal1A and gal10ΔA were constructed by complete replacement of the corresponding genes with kanR using the loxP-kanR- loxP cassette (26), while gal2ΔA, gal4ΔA, gal6ΔA, and gal80Δ were obtained from the Saccharomyces Genome Deletion Project [27] and were constructed analogously. Strains gal2Δgal80Δ and gal4Δgal80Δ were obtained by mating and sporulation of the single-deletion mutants; gal1Δgal10Δ (#RR146) was a generous gift from C. Roberts of Rosen lab.
16. Yeast were inoculated in 1 ml of either of GAL-inducing “−gal” media (1% yeast extract, 2% peptone, 2% raffinose, 2% galactose) or noninducing “−gal” media (1% yeast extract, 2% peptone, 2% raffinose). Cultures were grown at 30°C to a density of 1 to 2 OD600 washed in 5 ml H2O, and snap-frozen.
18. A likelihood statistic λ was computed for each gene to determine whether its mRNA expression level differed between the two cell populations compared by a microarray (17); genes having λ < 45 were selected as differentially expressed. This value was approximately the maximum obtained in control experiments in which the two mRNA populations were derived from identical strains and growth conditions (wt vs +gal).
19. The 997 affected genes were clustered based on Euclidean distance between their log10 expression ratios over all perturbation conditions, using a 4 by 4 column self-organizing map (SOM) implement- ed by the GeneCluster application (28), with 100 epochs.
20. Supplementary material is available at www.science mag.org/cgi/content/full/292/5518/929/DC1
21. Cells were grown and harvested for mRNA measure- ment, with proteins extracted according to Futcher (29). Extracts were desalted (Biorad 1000 columns) in 50 mM tris 8.3, 1 mM EDTA, and 0.5% SDS. The ICAT method (2) was applied to 30 μg of protein from each extract, with the following modifications. After trypsin digestion, peptides were fractionated on a 2.1 mm by 200 mm polySULFOETHYL A Column (PolyLC) by running a salt gradient from 0 to 25% B (5 mM KH2PO4, pH 3.0, 350 mM KCl, 25% CH3CN) over 30 min, followed by 25% to 100% Buffer B over 20 min at 0.2 ml/min. Labeled peptides were isolated from each cation-exchange-chromatography fraction by monomeric avidin (Pierce) affinity chromatography. Then, 10 to 80% of the peptide mixture was analyzed by capillary-liquid-chromatogra- phy MS/MS.
22. Protein-protein interactions, as reported in (3), were derived predominantly through two-hybrid assays and interactions culled from the literature. Protein-DNA interactions were obtained with permission from TRANSEQ (CFBDB) and represent all interactions in these databases as of July 2000. A more comprehensive analysis of the physical-interaction network will be provided in a future publication.
23. A nucleotide weight matrix model of the binding site (TRANSEQ site MM00049) was used to identify potential binding sites in the promoter regions of genes in the 997-gene set. Nucleotide sequences of up to 800 bp upstream of translation start sites, terminating at the nearest upstream ORFs, were scored against the weight matrix using MatInspector (32) with core similarity 0.7 and matrix similarity 0.8.
24. While this work was in review, one of these genes (PCL10) was identified as a target of Gal4p by direct DNA-binding assay (33). Also, YHR318C was impli- cated in a previous binding-site prediction study (34).
Vital Involvement of a Natural Killer Cell Activation Receptor in Resistance to Viral Infection

Michael G. Brown, 1* Ayotunde O. Dokun, 1,2* Jonathan W. Heusel, 1 Hamish R. C. Smith, 1 Diana L. Beckman, 1 Erika A. Blattenberger, 1 Chad E. Dubbelde, 1 Laurie R. Stone, 1 Anthony A. Scalzo, 2 Wayne M. Yokoyama 1,3

Natural killer (NK) cells are lymphocytes that can be distinguished from T and B cells through their involvement in innate immunity and their lack of rearranged antigen receptors. Although NK cells and their receptors were initially characterized in terms of tumor killing in vitro, we have determined that the NK cell activation receptor, Ly-49H, is critically involved in resistance to murine cytomegalovirus (MCMV) in vivo. Ly-49H requires an immunoreceptor tyrosine-based activation motif (ITAM)–containing transmembrane molecule for expression and signal transduction. Thus, NK cells use receptors functionally resembling ITAM-coupled T and B cell antigen receptors to provide vital innate host defense.

Natural killer (NK) cells were first identified because of their “natural” ability to kill tumors in vitro, an ability that is now known to occur through activation receptors that trigger the release of perforin-containing cytolytic granules [reviewed in (1)]. These lymphocytes can be distinguished from T and B cells because they do not express rearranged antigen receptors and are not directly involved in acquired immunity. However, NK cells participate in early innate host defense against pathogens and are generally thought to counter infections through a nonspecific response to inflammatory cytokines that induce their production of interferon-γ (2). Yet, NK cells appear to respond specifically against certain pathogens. For example, in humans, selective NK cell deficiency is associated with recurrent systemic infections, especially with herpesviruses such as cytomegalovirus (3). This is closely paralleled by the susceptibility of NK cell–depleted mice to murine cytomegalovirus (MCMV) but not to lymphocytic choriomeningitis virus (4). Although the mechanisms underlying this susceptibility are incompletely understood, NK cell receptors that activate tumor cytotoxicity may play important roles in innate defense against specific infections (5).

The critical involvement of NK cell activation receptors in defense against pathogens is highlighted by the expression of virus-encoded proteins that interfere with natural killing (6). In many cases, these proteins enhance the function of inhibitory major histocompatibility complex (MHC) class I–specific NK cell receptors that potentially interfere with signals from activation receptors, such as Ly-49D and Ly-49H, that are coupled to immunoreceptor tyrosine-based activation motif (ITAM)–containing transmembrane proteins.

A physical map of the Ly-49 gene cluster is expanded at right (36 tested loci) and between D6Mit19 and D6Mit370 (36 tested loci) and between Iva2 and Xmmv54 (10 loci) contain DBA/2 alleles. C57BL/6 alleles account for all BXD-8 loci reported to reside between the centromere and telomere (10 loci). Chromosomal regions derived from C57BL/6, BXD-8, except for the Xmmv54 locus (9, 25). A physical linkage map of the NKC is depicted in the center, with selected loci that have been useful to distinguish alleles (19). A physical map of the Ly-49 gene cluster is expanded at right (26, 27). BXD-8 and C57BL/6 NKC alleles are identical sized for all NKC loci shown at center and for D6Mit18, D6Mit198, 6 Mit16, and D6Mit198 (22). Surrounding Ly-49h are the Ly-49k and Ly-49n pseudogenes (28). MCMV replication in F1 hybrid offspring from DBA/2 and BXD-8. Three days after infection (with MCMV Smith strain, 2 × 10^4 plaque-forming units [PFUs]), organ viral titers were assessed in tissue homogenates collected from C57BL/6, DBA/2, BXD-8, and (DBA/2 × BXD-8) F1 hybrid mice (five mice per group), as indicated. Spleen titers are shown here; liver titers are available online (8). Each point represents the average titer determined for an individual mouse. In the spleens of two C57BL/6-infected mice, viral replication was below the level of detection by this assay and is indicated with asterisks. Mean viral titers for each group are depicted as horizontal bars. For mice with titers below the level of detection of the assay, the minimum number of detectable PFUs (10^2) was assumed to determine the mean. This assumption overestimates the mean for the group having titers below detectable levels.