The Transcriptional Program of Sporulation in Budding Yeast

S. Chu, *et al.*

Science **282**, 699 (1998);
DOI: 10.1126/science.282.5389.699

The Transcriptional Program of Sporulation in Budding Yeast

S. Chu,* J. DeRisi,* M. Eisen, J. Mulholland, D. Botstein, P. O. Brown,† I. Herskowitz†

Diploid cells of budding yeast produce haploid cells through the developmental program of sporulation, which consists of meiosis and spore morphogenesis. DNA microarrays containing nearly every yeast gene were used to assay changes in gene expression during sporulation. At least seven distinct temporal patterns of induction were observed. The transcription factor Ndt80 appeared to be important for induction of a large group of genes at the end of meiotic prophase. Consensus sequences known or proposed to be responsible for temporal regulation could be identified solely from analysis of sequences of coordinately expressed genes. The temporal expression pattern provided clues to potential functions of hundreds of previously uncharacterized genes, some of which have vertebrate homologs that may function during gametogenesis.

All sexually reproducing organisms have a specialized developmental pathway for gametogenesis, in which diploid cells undergo meiosis to produce haploid germ cells. Gametogenesis in yeast (sporulation) involves two overlapping processes, meiosis and spore morphogenesis (Fig. 1), and results in four haploid spores. Each spore is capable of germinating and fusing with a cell of the opposite mating type, analogous to the fusion of egg and sperm.

Sporulation in yeast is characterized by sequential transcription of at least four sets of genes—early, middle, mid-late, and late (*1*). Most of the known early genes are involved in meiotic prophase (pairing of homologous chromosomes and recombination). The Ume6/Ime1 complex, which recognizes a conserved site (URS1) found in the upstream region of many of the known early genes, appears to be the major transcriptional regulator of this class (*2*, *3*). Products of the known middle genes are required for the concomitant events of meiotic nuclear division and spore formation (*4*–*6*). Ndt80, a meiosis-specific transcription factor, has been shown to be important in inducing transcription of middle genes at the end of meiotic prophase, binding to the middle gene sporulation element (MSE) motif found upstream of many of these genes (*4*, *7*–*9*). The mid-late class includes genes necessary for formation of the outer layer of the spore wall (*10*). The late genes are thought to have a role in spore maturation (*11*). The factors that regulate transcription of these latter two classes of genes remain to be discovered.

Numerous genes necessary for sporulation have been found in genetic screens by means of visual assays for spore formation (*7*, *10*, *12*) or assays for meiotic recombination or other sporulation-specific processes (*13*). A variety of approaches have been used to identify genes that are differentially expressed during sporulation (*2*, *14*). All told, these efforts have identified roughly 150 genes (*2*).

We used DNA microarrays containing 97% of the known or predicted genes of *Saccharomyces cerevisiae* (*15*) to explore the temporal program of gene expression during meiosis and spore formation. The entire data set, including raw images and quantitation is available at http://cmgm.stanford.edu/pbrown/sporulation. Electron and light microscopy were used to monitor the accompanying morphological changes (Fig. 2), allowing us to relate the gene expression program to the observable sequence of events in the sporulation process (*16*). Changes in the concentrations of the mRNA transcripts from each gene were measured at seven successive intervals after transfer of wild-type (strain SK1) diploid yeast cells to a nitrogen-deficient medium that induces sporulation (*17*, *18*). The samples were taken at times based on the expression pattern of known early, middle, mid-late, and late genes (*DMC1*, *SPS1*, *DIT1*, and *SPS100*, respectively) (*19*). The expression patterns of these four genes were independently monitored by Northern analysis (Fig. 3A). To display the quantitative expression patterns for hundreds of genes in a compact graphical format suitable for this report, measured changes in mRNA levels were shown in a tabular form (Fig. 3B), with rows corresponding to individual genes and columns corresponding to the successive intervals during the sporulation program at which mRNA levels were measured. The changes in expression of each gene are represented in the table not as numbers, but by mapping the numerical values onto a color scale. Increases in expression relative to vegetative cells are represented as graded shades of red, and decreases as graded shades of green. The pattern of expression as assayed by Northern analysis (Fig. 3A) was very similar to that determined by microarray analysis (Fig. 3B).

Sequential Induction of Genes During Sporulation

Of the about 6200 protein-encoding genes in the yeast genome, more than 1000 showed significant changes in mRNA levels during sporulation (*20*). About half of these genes were induced during sporulation, and half were repressed. To facilitate the visualization and interpretation of the gene expression program represented in this very large body of data, we have used the method of Eisen *et al.* (*21*, *22*) to order genes on the basis of similarities in their expression patterns and display the results in a compact graphical format (Fig. 4A).

The relatively small number of genes (about 50) whose transcription has been studied previously had defined four temporal classes of sporulation-specific genes (*1*). These classes were evident in this analysis but were not sufficient to represent the diversity of observed expression patterns. We found it useful to distinguish seven temporal patterns of induced transcription that reflect sequential progression through this program, even though well-defined boundaries between temporal classes could not be determined. (Increased synchrony and more frequent time points might sharpen these boundaries and reveal more classes.) For each of these seven temporal patterns, a small, representative set of genes was hand-picked and used to define a model expression profile (Fig. 4B). A variety of temporal expression patterns were also observed for the genes whose mRNA transcripts decreased during sporulation.

To display the results as shown in Fig. 5A, correlation coefficients were computed, relating the expression profiles of each induced gene to each of the seven model profiles in Fig. 4B. Genes were then grouped according to the model profile that gave the highest correlation coefficient. The seven groups were placed in a sequence that reflected the time of initial induction. Genes assigned to each group were then further ordered on the basis of the

S. Chu and I. Herskowitz, Department of Biochemistry and Biophysics, University of California, San Francisco, San Francisco, CA 94143–0448, USA. J. DeRisi, Department of Biochemistry, Stanford University School of Medicine, Stanford, CA 94305–5428, USA. M. Eisen, Departments of Biochemistry and Genetics, Stanford University School of Medicine, Stanford, CA 94305–5428, USA. J. Mulholland and D. Botstein, Department of Genetics, Stanford University School of Medicine, Stanford, CA 94305–5428, USA. P. O. Brown, Department of Biochemistry and Howard Hughes Medical Institute, Stanford University School of Medicine, Stanford, CA 94305–5428, USA.

*These authors contributed equally to this work
†To whom correspondence should be addressed. E-mail: pbrown@cmgm.stanford.edu and ira@cgl.ucsf.edu

similarity of their expression patterns to the model expression patterns that defined the two neighboring groups. On average, 70% of the genes within each class had a correlation coefficient of >0.9. However, genes were binned according to the best fit, and often the expression profile for a given gene was also highly correlated with other classes as well. The order of the genes in Fig. 5 reflects this fact—a gene that correlated well to a neighboring class was placed closer to the border between the classes. For additional information, see cmgm.stanford.edu/pbrown/sporulation.

Rapid, transient induction ("metabolic"). About 52 genes were induced rapidly and transiently after transfer to sporulation medium. The majority of these genes have known metabolic functions, many related to adaptation to nitrogen starvation. The promoters of 15 of these genes contain a URS1 motif (5′-DSGGCGGCND-3′) (*23*). Three of these genes are already known to be regulated by Ume6 (*24*, *25*). The difference in expression pattern between these 10 transiently induced genes and the more stably induced early (I) genes—both of which have URS1 sites—is striking and points to the importance of additional regulatory elements (*1*).

Early (I) induction. About 62 genes followed a pattern of expression characterized by early induction, detectable 0.5 hour after transfer to sporulation medium, and sustained expression through the rest of the time course (Fig. 5). About one-third of the genes that were expressed in this pattern have a URS1 consensus site within 600 base pairs upstream of their start codons and 43% have a core URS1 consensus site (5′-GGCGGC-3′). An unbiased search of the promoter regions for these genes with the pattern analysis tool MEME (*23*) revealed sequence motifs identical to the previously characterized URS1 consensus site. Many of the genes with known functions that were expressed in this pattern have roles in synapsis of homologous chromosomes (for example, *HOP1*) or recombination (for example, *DMC1*), suggesting that newly identified members of this class may have related roles.

Early (II) induction. About 47 genes followed this pattern of induction, which is distinguished from the early (I) profile by a slightly delayed increase in transcript levels (Fig. 5A). As was seen for early (I) genes, a large fraction of the characterized genes with an early (II) induction pattern have been implicated in meiotic chromosome pairing and recombination. Genes expressed in this pattern were less likely than the early (I) genes to have a putative URS1 site in their promoter regions.

Early-middle induction. About 95 genes showed an early-middle pattern of induction. As with early genes, these genes were initially induced in the first 2 hours after transfer to sporulation medium, but they were distinguished from the early genes by an additional increase in transcript levels measured at 5 or 7 hours after transfer, paralleling the induction of the middle and mid-late genes, respectively. Several of the genes with this expression pattern have roles in aspects of spindle and spindle pole body (SPB) dynamics (such as *SPC42*) or chromatid behavior (such as *PDS1*). How the expression of the early-middle genes is regulated is not known. Their early induction is unlikely to be directly mediated by Ime1/Ume6, as only five of the genes with this induction pattern contain a URS1 site in their promoters. A putative MSE (5′- DNCRCAAAW-3′) (*4*, *23*) site can be recognized in the promoter regions of about half of the early-mid genes, suggesting that the boost in their expression midway through the sporulation process may be mediated by Ndt80.

Middle induction. About 158 genes were strongly induced between 2 and 5 hours after transfer to sporulation medium, a time corresponding to the completion and exit from meiotic prophase (Fig. 2). Many genes of known function that showed this expression pattern are involved in the control or the mechanics of meiotic division (such as *CLB6*) or in spore morphogenesis (such as *DBI9*/*SPO20*). One or more putative MSE sites can be recognized in the sequences immediately upstream of 70% of the genes in this group. These sites are presumably recognized by Ndt80 to activate their transcription. Using the same approach (*23*) taken for the early genes, an unbiased search of the middle genes yielded sequence motifs perfectly matching the MSE, as previously defined (*4, 8, 9*). About 47 of the genes with this induction profile lack a perfect MSE. Their mid-sporulation induction may still be dependent on Ndt80, either directly or indirectly. For example, Ndt80 may recognize a noncanonical MSE, or it may activate synthesis of another transcription factor that stimulates transcription of these genes. Some middle genes, especially those without evident MSE sites, may have Ndt80-independent expression. Fifteen (31%) of these genes share a common sequence motif (5′-CWBYSCTTT-3′) (*23*) in their upstream regions that may be responsible for inducing their transcription at this time. The promoter regions of half of these genes contain more than one copy of this sequence.

Mid-late induction. About 61 genes were induced between 5 and 7 hours after transfer to sporulation medium. Among the genes in this class whose functions have been studied, most are involved in aspects of meiotic division (such as *CDC26*) or spore formation (such as *DIT1*). Thirty-six percent of the mid-late genes have one or more MSE sites in their promoter region. The delay in their induction relative to that of the middle genes suggests the possibility that a negative regulatory site may delay utilization of the MSE (*26*).

Late induction. Five genes included in this analysis were induced between 7 and 11.5 hours after transfer to sporulation medium. This set includes the prototypical late gene, *SPS100*, which is necessary for spore wall maturation (*11*). Consistent with a transcriptional cascade model, none of these genes were induced during sporulation in the absence of Ndt80 (*27*).

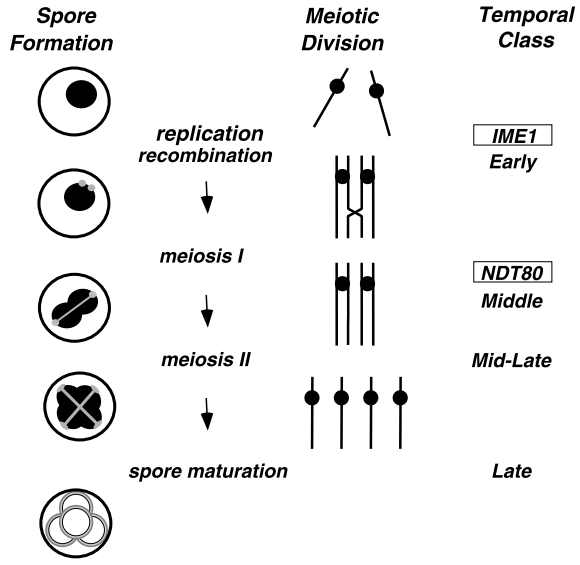

Fig. 1. Landmark events of sporulation. Sporulation involves chromosome distribution and spore morphogenesis. In meiosis, chromosomes first replicate, and then homologous chromosomes align and undergo recombination during prophase. Two consecutive nuclear divisions follow, in which, first, homologous chromosomes segregate apart (meiosis I), and then sister chromatids separate (meiosis II) (*29*). These divisions are preceded by duplication and separation of the spindle pole bodies (SPBs) and formation of the spindle. One of the first discernible steps of spore morphogenesis occurs after the SPBs have duplicated and formed the meiosis II spindles (*59*). At this point, each of their outer plaques becomes modified and nucleates the formation of a flattened membrane sac which grows to become a prospore membrane. Maturation of the spore walls then results in an ascus containing a tetrad of haploid spores. The gene expression program of sporulation has previously been characterized as a transcriptional cascade involving four temporal classes, indicated on the diagram. For review, see (*1, 2*). Ime1 is a transcriptional activator for some known early genes; Ndt80 is a transcriptional activator for some known middle genes.

Repression. Transcripts from more than 600 genes diminished in the course of sporulation. The temporal patterns of sporulation-specific repression varied, but at least three could be distinguished. Perhaps the most distinctive was the pattern shared by the genes encoding ribosomal proteins. Their transcripts diminished significantly within 0.5 hour of transfer to sporulation medium but by 11.5 hours had returned to their presporulation levels. The initial repression may reflect the cessation of growth upon nitrogen starvation. The renewed expression of these genes during spore maturation may serve to pack the maturing spores with ribosomes, allowing them to grow vigorously upon germination. Two other well-represented patterns were characterized by a sustained reduction in transcript levels beginning at 0.5 and 2 hours, respectively.

Regulation of Middle Genes by Ndt80

We found that more than 150 genes were induced between 2 and 5 hours after transfer to sporulation medium, in a pattern like that of the known targets of Ndt80. To further characterize the role of Ndt80 in their induction, we examined the consequences of expressing Ndt80 ectopically in vegetative cells (*28*) and of eliminating Ndt80 during sporulation (*17*). Previous work had shown that ectopic expression of Ndt80 in vegetative cells induces several known middle genes [*SPS1*, *SPC42*, *CLB3*, and *CLB6*, among others; (*4*)]. Microarray analysis revealed that more than 200 genes were induced at least threefold when Ndt80 was expressed ectopically. About 42% of these genes were expressed in the middle induction pattern during sporulation. In contrast, fewer than 20% of the genes induced by ectopic expression of Ndt80 exhibited the metabolic, early, early-mid, mid-late, or late induction patterns during sporulation. Two-thirds of the genes with a middle induction pattern were induced more than two-fold by ectopic expression of Ndt80 in vegetative cells. Despite this very strong association between a midsporulation pattern of induction and inducibility by Ndt80 in vegetative cells, there were some exceptions. For example, YDR065 and *BNR1* were induced during sporulation (>15-fold and >8-fold, respectively) but not in vegetative cells expressing Ndt80 (<2-fold). These genes contain presumptive Ndt80 binding sites in their promoters but may require an additional sporulation-specific transcription factor for induction or they may be subject to repression in vegetative cells.

About one-third of all the genes in the yeast genome and also one-third of all early genes have at least one sequence in their promoter regions that matches the consensus motif for an MSE site. In contrast, about 70% of all middle genes and 62% of all the genes that were induced at least threefold by ectopic Ndt80 expression contain at least one MSE site in their upstream regions. There is thus a strong, but by no means absolute, correlation between the presence of MSE sites in a gene's promoter, Ndt80-inducibility, and a mid-sporulation induction profile.

To identify genes for which Ndt80 was necessary for induction during sporulation, we analyzed gene expression in a strain lacking Ndt80 after 6 hours in sporulation medium. The magnitude of the induction of most middle genes was reduced in the Ndt80-deficient mutant to about one-third of the level in wild-type strains, indicating the importance of Ndt80 for their induction. There remained, however, significant Ndt80-independent induction of most of the middle genes. For several middle genes, such as *SPS2* and *SPR28*, the residual induction was comparable in magnitude and timing to that seen in wild-type cells. The promoters of both of these genes contain putative MSE sites, and they were strongly induced by ectopic Ndt80 expression in vegetative cells. Thus, it appears that both Ndt80 and another sporulation-specific factor participate in the mid-sporulation induction of these and perhaps most other middle genes.

Proposed Roles for Genes Induced During Sporulation

How can we begin to understand what contribution each of the induced genes makes to the efficiency of sporulation or fitness of the resulting spore? Because genes with related functions tend to be expressed in similar patterns, we can suggest possible roles for genes of unknown function based on their temporal association with genes of known function (a "guilt-by-association" argument). These hypotheses about gene function can then be tested by making mutations in the genes in question, and analyzing their effects.

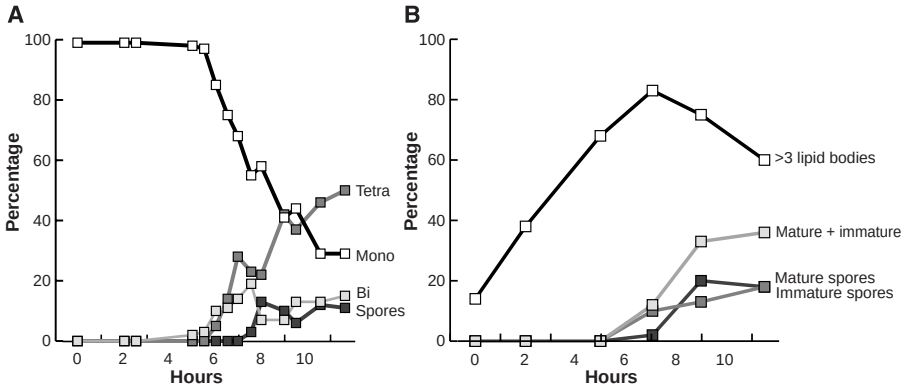

Fig. 2. Cytological assessment of the time course of sporulation. (**A**) Samples were taken at the indicated times and stained with DAPI (4′-6′diamidino-2-phenylindole) by standard methods (*61*). Percentages of cells containing mature spores, tetranucleate cells, binucleate cells, and mononucleate cells are plotted (n = 200). The final sporulation efficiency after more than 24 hours was 67%. (**B**) Samples were taken at the indicated times and subjected to sectioning, staining, and observation by electron microscopy as described (*60*). Because cell sections rather than entire cells were analyzed, percentages reflect relative rather than absolute frequencies of spores and lipid bodies. The curves represent the fraction of cell sections containing mature spores, the fraction containing immature spores (missing intact mature spore wall), the fraction of cell sections containing either of those two, and the fraction containing more than three "lipid bodies" (an ultrastructural term referring to an early progenitor of the spore wall).

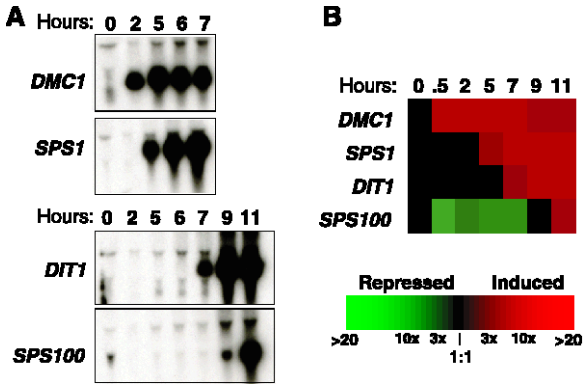

Fig. 3. Northern blot assay of gene expression and corresponding microarray data. (**A**) Samples from the indicated time points were assayed by Northern analysis (*19*). Genes were chosen to be representative of the four previously identified temporal classes. *DMC1*, *SPS1*, *DIT1*, and *SPS100* belong to the early, middle, mid-late, and late classes, respectively (*1*, *2*). (**B**) RNA samples from the same time course as in (A) were analyzed by microarray analysis. Data are graphically displayed with color to represent the quantitative changes. Increases in mRNA (relative to pre-sporulation levels) are shown as shades of red and decreases in mRNA levels are represented by shades of green.

Chromosome behavior in prophase. The meiosis-specific events of homologous chromosome pairing and recombination must be coordinated with DNA replication and sister chromatid cohesion. We found that most of the genes known to be involved in these processes showed early (I) or early (II) expression patterns. Conversely, many of the previously characterized genes that were expressed in either the early (I) or early (II) pattern are known to have roles in chromosome pairing and recombination. Examples include *ZIP1* and *RED1*, which encode coiled-coil proteins of the synaptonemal complex (*29*). It is therefore likely that some of the novel early (I) genes, for example, YOR177, which encode a putative coiled-coil protein, may also be involved in chromosome synapsis and meiotic recombination.

Many genes involved in DNA replication (*RNR1*, *CDC21*, *RFA2*, *RFA3*, *POL1*, *POL30*, *MCM3*, *CDC47*, and *DBF4*) (*30*) or sister chromatid cohesion (*SCC2*, *PDS1*, *PDS5*, *SMC1*, and *SMC3*) were also induced early during sporulation. Regulation of sister chromatid cohesion in meiosis is of special interest because sister chromatid cohesion must be maintained in anaphase I, in contrast to the situation in meiosis II or mitosis, where the sister chromatids must separate in anaphase (*31*). Although *SCC1* [which functions in mitotic sister chromatid cohesion (*32*)] is not induced during meiosis, it is possible that a novel sporulation-induced protein, for example, the YPR007 or YER106 gene products, performs this function during meiosis. The YER106 product is similar to the condensin, Smc3. YPR007, which showed an early (I) expression pattern, is predicted to encode a protein similar to Scc1 of *S. cerevisiae* (*32*) and Rec8 of *S. pombe* (*33*), both of which have roles in sister-chromatid cohesion. Deletion of YPR007 did not impair vegetative growth, but cells deleted for YPR007 did not undergo meiotic division and were unable to sporulate (*34*). We therefore designate this gene *SPO69*.

Exit from prophase. Meiotic prophase ends after the completion of recombination between synapsed homologous chromosomes, whereupon the duplicated SPBs undergo separation to establish the spindle for meiosis I. Mutants lacking Ndt80 arrest with fully synapsed homologs and duplicated but unseparated SPBs, suggesting that Ndt80 is important for triggering progression from prophase to the first meiotic division (*7*). Several putative targets of Ndt80 may be involved in desynapsis and later steps in anaphase. For example, *UBC9*, which is induced midway through sporulation, encodes a functional homolog of the mammalian ubiquitin-conjugating enzyme, Hsubc9. This protein has been implicated in disassembly of the synaptonemal complex during spermatogenesis (*35*). Another middle gene that may play a role in the disassembly of the synaptonemal complex is YOR339, which is similar to ubiquitin-conjugating enzymes and is strongly induced by ectopic Ndt80 expression. Therefore, new transcription at the end of prophase, perhaps mediated by Ndt80, may initiate desynapsis by stimulating degradation of components of the synaptonemal complex.

Meiotic nuclear division: machinery involved in chromosome distribution. Chromosome distribution during mitosis requires the functioning of cyclin-dependent kinases, motor proteins, and kinetochore proteins. Genes encoding proteins in each of these classes were induced during meiosis.

Of the B-type cyclins, five of the six *CLB* genes were induced just before cells initiated meiotic division (*4*). Genes for several known and presumptive kinesin-like (Cin8, Kip3, Kar3, and YGL075), dynein-like (Pac11), and myosin-like (Tid3 and YGR179) motor proteins were also induced at this time, suggesting that they may function in chromosome movement, cytoskeletal remodeling, or organelle movement.

Other genes required for mitotic chromosome segregation (for example, *CSE4* and *IPL1*) were also induced during sporulation. *CSE4* codes for a protein similar to a mammalian kinetochore protein, CENP-A, which is involved in centromere organization during spermatogenesis (*36*). Cse4 may play a similar role. *IPL1* encodes a serine/threonine kinase required for mitotic chromosome segregation (*37*). We have found that Ipl1 is required for normal spore viability (*38*). A mammalian homolog of *IPL1*, *AYK1*, is induced just prior to meiosis I (*39*), similar to the timing of *IPL1* expression during sporulation.

Anaphase of the meiotic divisions. Exit from mitosis is governed by proteolysis mediated by the anaphase promoting complex (APC) and its specificity factors, Cdc20 and Hct1 (*40*). We found that *CDC20*, *HCT1*, and many APC components (*APC4*, *APC5*, *APC9*, *APC11*, *CDC16*, *CDC23*, *CDC26*, and *CDC27*) were induced midway through sporulation, strongly suggesting that APC-mediated proteolysis also mediates exit from one or both of the meiotic divisions. YGR225, which was expressed in an early-mid pattern, is similar in sequence to the

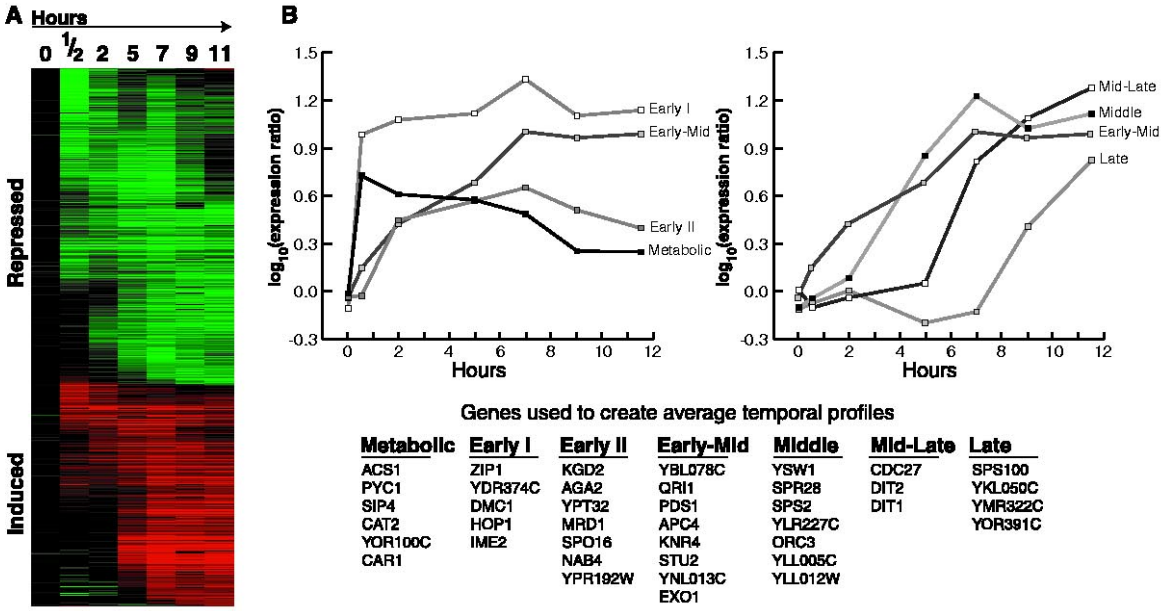

Fig. 4. The global pattern of gene expression during sporulation. (**A**) The genes that showed the most significant induction or repression during sporulation (*20*) (a total of 1116) were ordered so that genes with similar expression patterns were grouped together and so that the order roughly reflected the time of first induction or repression. Changes in mRNA concentration for each gene, relative to those in cells immediately prior to transfer to sporulation medium, are shown with a color to represent the quantitative changes as described in Fig. 3. Time after transfer to sporulation medium is shown on the horizontal axis. (**B**) A set of representative genes was chosen from each of these seven expression patterns, and average expression profiles for each set were calculated.

CDC20/FIZZY gene family (*40*) (Fig. 6). We therefore suggest that it might encode a meiosis-specific activator of the APC that is necessary for completion of one or both meiotic divisions. YGR225 was expressed in a pattern very similar to that of *APC4*, which encodes a known component of the proteolysis machinery (*41, 42*). We have inactivated YGR225 and found that it is not essential for vegetative growth but is required for sporulation (*34*) and therefore designate it *SPO70*. Regulation of Clb-associated kinase levels must be different from that occurring during the mitotic program to allow for the consecutive nuclear divisions that occur in meiosis. Perhaps a sporulation-specific protein, like Spo70, plays such a role.

Several other genes involved in exit from mitosis, including *CDC14* and *CDC5*, were induced during sporulation as previously reported (*43*). A mammalian homolog of *CDC5* is induced specifically during spermatogenesis (*44*). Two genes that are predicted to encode F-box proteins (*45*), YDR219 and YLR368, were induced midway through sporulation, raising the intriguing possibility that an SCF-like ubiquitin-conjugating complex may be necessary for progression through meiosis.

Spore morphogenesis. During formation of the prospore membrane, vesicles normally bound for the plasma membrane are thought to be redirected to the incipient prospore membrane (*46*). The genes that participate in the assembly of this specialized membrane-bound structure remain largely undefined. In parallel with the induction of *SPO20*, which codes for a sporulation-specific SNAP25 homolog necessary for prospore membrane growth (*46*), we observed increased expression by 7 hours of several other genes that may be necessary for vesicle fusion during this process: *SEC12*, *SED4*, *ROM2*, YPT1, YPT10, and YPT32. Reorganization of the cytoskeleton may also be required for redirection of vesicles. Several genes (such as *CDC42*, *BNR1*, *SPA2*, and *PEA2*) which affect the actin cytoskeleton were induced at the time of prospore membrane formation. We speculate that their proteins may localize to the leading edge of the prospore membrane, as has been observed for several of the septins (*46*), and may comprise the coated bulb-like lip structure at the leading edge (*47*).

The leading edge at the rim of each prospore membrane grows from its origin at the SPB around the nuclear membrane. It engulfs organelles and cytoplasm and eventually fuses to encapsulate each haploid nucleus. Genes involved in membrane fusion (such as *FUS2* and *KEL2*) were induced at the time (7 hours) when this process was occurring, suggesting that they may have roles in mediating the ultimate fusion of the rim of the prospore membrane to form a

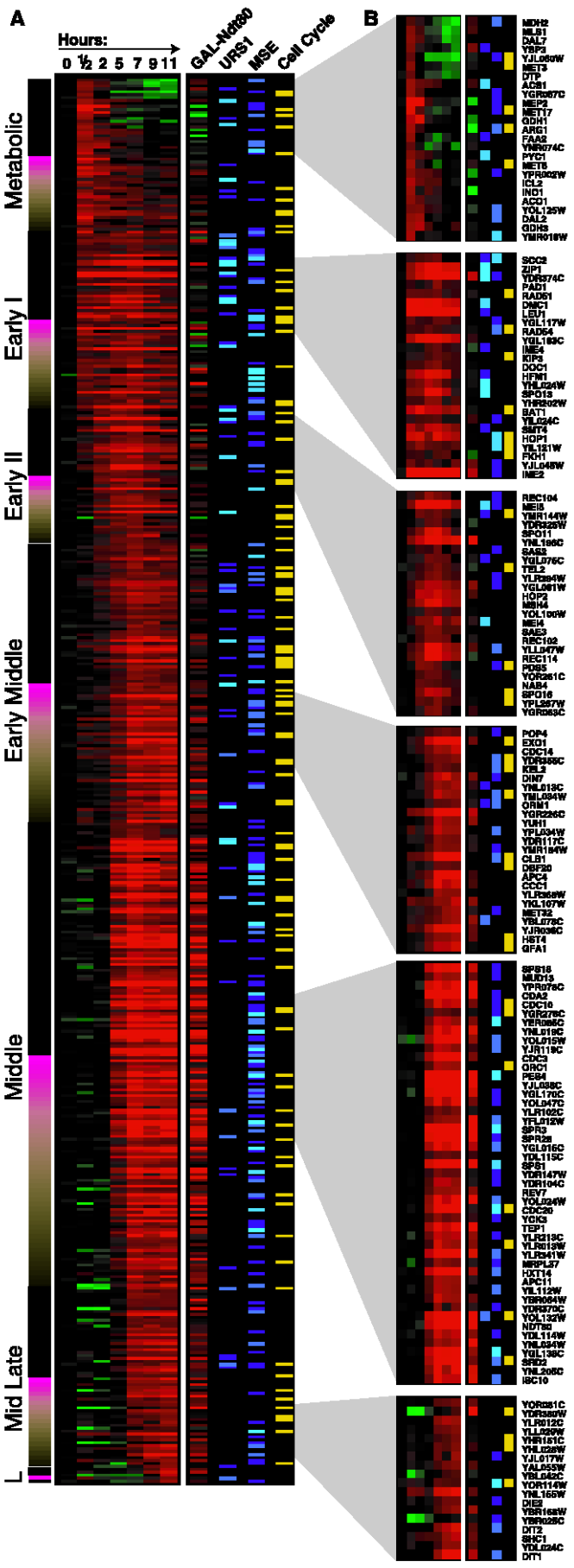

Fig. 5. Genes induced during sporulation. (**A**) The seven average expression profiles (Fig. 4B), arranged by time of first induction, were used to further classify the induced genes shown in Fig. 4A, by first assigning each gene to the expression class with which it had the highest correlation. Genes assigned to each class were then ordered with respect to correlation with adjacent temporal classes. In this graphical display, changes in mRNA level resulting from ectopic expression of Ndt80 in vegetative cells are shown in the eighth column (GAL-Ndt80). Two columns with blue bars identify genes whose promoters contain a putative URS1 or MSE regulatory sequence as labeled. The degree to which the sequence matches the consensus for each of these regulatory elements is indicated by the brightness of the bar—the best matches are represented by bright-blue bars, and less stringent matches by dark-blue bars. (*23*). Yellow bars in the rightmost column of this display identify genes that showed periodic expression during the mitotic cell cycle (*63*). (**B**) Representative sections from (A) are expanded to allow visualization and identification of individual genes.

closed structure.

We also observed induction of *STE5*, which encodes a scaffold protein for the MAP kinase signaling module used during mating (*48*), midway through sporulation. *STE5* may be redeployed during sporulation to hold together components of the proposed sporulation MAP kinase cascade (*5, 6*).

Several genes encoding putative GPI-linked proteins (*49*) were induced at the time of spore formation and thus represent candidate components of the prospore membrane or spore wall. These include the putative membrane proteins encoded by *SPS2*, YCL048, *GAS2*, and *GAS4*, and the putative cell wall proteins corresponding to *SPR2*, YPL130, *CWP1*, and *TIR2*. Other potential structural components of the spore wall include transmembrane proteins such as those predicted to be encoded by YDR104, YOL047, YAL018, and YCR061. YDR104 is a middle gene with a putative MSE in its promoter region that is predicted to have five transmembrane segments. Mutant cells deleted for the YDR104 gene had no vegetative phenotype. Upon sporulation, they underwent both meiotic divisions but failed to form a spore wall (*34*). We therefore designate YDR104 as *SPO71*. The expression kinetics of *SPO71* and the phenotype of the *SPO71*-deficient mutant were very similar to those of *SPS1*, another middle gene required for spore wall maturation (*5*), suggesting that *SPO71* may also act directly in this process.

Organelle morphogenesis. Changes in organelle structure have been noted during prospore membrane encapsulation (*50*). Several genes that might be involved in sporulation-associated alterations of mitochondria and the vacuole were identified by their time of expression. Some genes with known or suspected roles in mitochondrial function were induced by 7 hours of the onset of sporulation, including the yeast homolog (YBR179) of the *Drosophila FZO* gene, required for mitochondrial fusion (*51*). In addition, *VAM7* and *VAC8*, which are required in vegetative cells for vacuolar morphogenesis and inheritance, respectively (*52*), were induced in midsporulation. Their expression may therefore promote the synthesis of new vacuoles in a maturing spore (*53*).

Transcription factors. Our current understanding of the sporulation transcriptional cascade comes from knowledge of the two key transcriptional activators (Fig. 1), Ime1/Ume6, which activates early genes (*2*), and Ndt80, which activates middle genes (*4*). Several putative transcriptional regulators were induced during sporulation and might therefore function during this transcriptional cascade (*1*, *2*). For example, *FKH1* (which was induced early during sporulation) is homologous to Mei4, a protein of the forkhead family that is required for transcription during sporulation in fission yeast (*54*).

The regulation of mid-late and late gene transcription remains to be elucidated. Middle and mid-late genes that encode proteins with DNA-binding domains are natural candidates for regulating the later waves of transcription in sporulation. *SPS18*, whose sequence predicts a protein with a Zn-finger domain (*55*); YGL183, which is predicted to encode a putative HMG box protein; and YPR078, whose predicted product has homology to the C-terminal domain of the transcription factor, Mbp1, are among these candidate regulators.

Genome-Wide Studies of a Developmental Pathway

We have attempted to describe the complete transcriptional program of sporulation, independent of and unbiased by any preconceived ideas. As a result of these studies, the number of genes identified as being induced during sporulation has been increased from 50 to 500, distributed over the course of sporulation. The number of members of the middle gene class was increased more than 10-fold to at least 150 genes, which showed it to be an unexpectedly large and important class of induced genes. In addition, we identified a large class of underappreciated genes (more than 500) whose expression is repressed during sporulation.

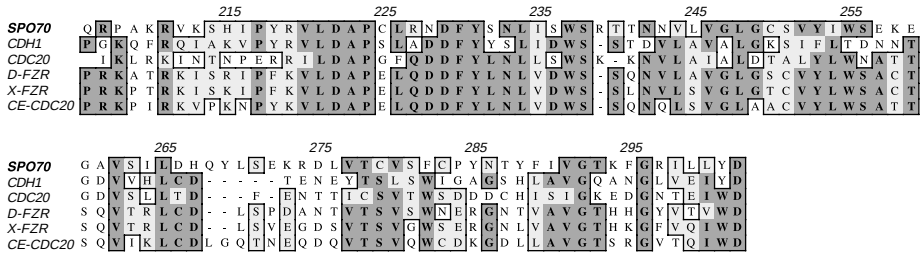

Fig. 6. Spo70 and Fzy/Fzr family. Spo70 (encoded by YGR225) is similar in the first two of seven distinctive WD repeats of the FZY/FZR family of ubiquitin-conjugating specificity factors (*40*). Comparison is shown for Spo70, Cdh1, and Cdc20 of *S. cerevisiae*, fizzy-related proteins of *Drosophila* and *Xenopus*, D-Fzr and X-Fzr, respectively, and Cdc20 of *Caenorhabditis elegans*.

This comprehensive characterization of the sporulation program makes a variety of insights and hypotheses possible. In particular, the genome-wide inventory of genes that are induced reinforces the idea that the multiple steps of sporulation can be viewed as comprising two major stages—first, meiotic prophase and second, meiotic division and gamete morphogenesis, with Ndt80 controlling entrance into the second phase, possibly in response to the recombination checkpoint (*56*). It is tempting to draw parallels with both spermatogenesis and human oogenesis, in which a natural arrest between these two stages can last for decades (*57*). We have seen several examples of genes induced during yeast meiosis that have homologs in vertebrates. Examples include yeast B-type cyclin (*CLB*) genes and cyclin genes of vertebrates, *UBC9* of yeast and its human homolog, *Hsubc9*, yeast *CSE4*, and mammalian *CENP-A*, yeast *IPL1* and mammalian *AYK1*, and yeast *CDC5* and its mammalian homolog. We expect that the yeast genes induced during sporulation may be a rich source of candidates for roles in gametogenesis in higher eukaryotes.

The approach we have used is readily applicable to genome-wide studies of other developmental processes, even in systems less experimentally tractable than yeast. Our studies indicate that for virtually any ordered process, potential regulatory sites such as the MSE can be identified exclusively on the basis of temporal expression data. It is particularly notable that we were able to identify sites such as the MSE and URS1 by searching the upstream regulatory regions of the middle and early gene set. (*24*). In principle, proteins such as Ndt80 that recognize specific sites can be identified by biochemical strategies (*62*) and then their biological relevance tested by ectopic expression and use of dominant negative derivatives.

Our studies serve as a reminder that even when simple models might appear to be sufficient to explain a biological process, multiple layers of complexity may be uncovered. In particular, we found strong evidence that additional proteins may modulate Ndt80 function (either to augment or inhibit its activity) or be partially redundant with Ndt80 in inducing middle gene transcription.

The method that we have used provides clues to potential functions of hundreds of previously uncharacterized genes. There is a strong correlation between the pattern in which a gene is expressed and the biological process in which its product participates. For example, genes involved in spore morphogenesis are induced late in sporulation, whereas genes involved in recombination are induced only in the first phase (*1*). Thus, the temporal expression pattern of a gene allows us to make inferences about its function. The

tools that we have used and developed for analyzing the expression data have facilitated this process. We know from prior studies [see (*14*)] that not all genes induced during sporulation will be essential for production of apparently normal spores under laboratory conditions. Nevertheless, we expect that for many of these genes, loss of function will result in a detectable defect in sporulation. We have provided three such examples here for genes with three different temporal expression profiles (*58*). The challenge remains to discover how each gene contributes to sporulation, and why the contribution has mattered in the evolution of *S. cerevisiae*.

References and Notes

1. A. P. Mitchell, *Microbiol. Rev.* **58**, 56 (1994).
2. M. Kupiec, B. Byers, R. E. Esposito, A. P. Mitchell, in *The Molecular and Cellular Biology of the Yeast* Saccharomyces: *Cell Cycle and Cell Biology*, J. R. Pringle, J. R. Broach, E. W. Jones, Eds. (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1997), pp. 889–1036.
3. I. Rubin-Bejerano, S. Mandel, K. Robzyk, Y. Kassir, *Mol. Cell. Biol.* **16**, 2518 (1996).
4. S. Chu and I. H. Herskowitz, *Mol. Cell* **1**, 685 (1998).
5. H. Friesen, R. Lunz, S. Doyle, J. Segall, *Genes Dev.* **8**, 2162 (1994).
6. L. Krisak *et al.*, *ibid.*, p. 2151.
7. L. Xu, M. Ajimura, R. Padmore, C. Klein, N. Kleckner, *Mol. Cell. Biol.* **15**, 6572 (1995).
8. S. R. Hepworth, L. K. Ebisuzaki, J. Segall, *ibid.*, p. 3934.
9. N. Ozsarac, M. J. Straffon, H. E. Dalton, I. W. Dawes, *ibid.* **17**, 1152 (1997).
10. P. Briza, M. Breitenbach, A. Ellinger, J. Segall, *Genes Dev.* **4**, 1775 (1990).
11. D. T. Law and J. Segall, *Mol. Cell. Biol.* **8**, 912 (1988).
12. R. E. Esposito, N. Frink, P. Bernstein, M. S. Esposito, *Mol. Gen. Genet.* **114**, 241 (1972).
13. R. E. Esposito and S. Klapholz, in *Molecular Biology of the Yeast* Saccharomyces: *Life Cycle and Inheritance*, J. N. Strathern, E. W. Jones, J. R. Broach, Eds. (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1981).
14. N. Burns *et al.*, *Genes Dev.* **8**, 1087 (1994), pp. 211–287.
15. J. L. DeRisi, V. R. Iyer, P. O. Brown, *Science* **278**, 680 (1997).
16. These microscopic observations indicate that meiosis I is initiated by about 5 hours and also provide a measure of the synchrony of the process in this cell population. Two-thirds of the cells in the population went through meiosis I and sporulated (Fig. 2A). The progression of the population through the morphologically defined stages of sporulation was only moderately synchronous, as indicated by the disappearance of mononucleate cells (Fig. 2A), and the synchrony diminished over time. Thus, the apparently sustained changes in expression that we observed may not reflect sustained expression in individual cells but rather the superposition of several stages in the process.
17. Cells from the wild-type SK1 background, YSC328 (a/α *ura3 leu2 lys2 ho::LSY2*), or the Ndt80-deficient SK1 background, YSC330 (a/α*ura3 leu2 lys2 ho:::LYS2 ndt80::LEU2*) were transferred to sporulation medium (SPM) at $t = 0$ with the protocol described in (*4*) to maximize the synchrony of sporulation. For YSC328, RNA was harvested at time $t = 0$, 0.5, 2, 5, 6, 7, 9, and 11.5 hours after transfer to SPM. (RNA from $t = 0.5$ hour was harvested from a separate time course with the same strain background and identical experimental conditions.) For YSC330, RNA was harvested at 0, 2, and 6 hours after transfer to SPM. Polyadenylated [poly(A)$^+$] RNA was prepared by purification with an oligo(dT) cellulose column as described (*15*).
18. Microarray analysis of poly(A)$^+$ RNA was carried out as described (*15*).
19. Portions from each time point were assayed by Northern (RNA) analysis following the protocol described in (*4*). Briefly, each lane was loaded with 30 μg of total RNA. Probe was made by random prime labeling of DNA template generated by polymerase chain reaction (PCR). Primers from Research Genetics were used to make probes corresponding to the open reading frames of *DMC1*, *SPS4*, *DIT1*, and *SPC100*.
20. Genes were included in this set if the root mean square of $\log_2 R$ at a given time point was greater than 1.13, where R is the measured ratio of each gene's mRNA levels to its mRNA level in vegetative cells just before transfer to sporulation medium. In practice, this criterion is essentially equivalent to a three-fold change for a single time point or an average 2.2-fold change across the entire time course.
21. M. B. Eisen, P. T. Spellman, P. O. Brown, D. Botstein, in preparation.
22. Briefly, a procedure based on a neural net algorithm was used to find an approximation to the optimal ordering of a chosen set of genes [see (*21*)] to minimize a parameter representing differences in the temporal pattern of expression between consecutive genes, summed over the entire ordered set [T. Kohonen, *Self-Organizing Maps* (Springer, Berlin, New York, 1997)].
23. The computer program Multiple EM for Motif Elicitation (MEME) was used to elicit sequence motifs, with 600 base pairs upstream of each gene as the input. MEME is available on the World Wide Web at www.sdsc.edu/MEME. The resulting degenerate motifs were used for Fig. 5. The least stringent sequence used for the MSE was 5′-DNCRCAAAWD and is shown as the darkest blue, followed by 5′-VKNCRCAAAWD and 5′-HDVKNCRCAAAWD. The most stringent consensus sequence was 5′-HDVGNCACAAAAD (brightest blue). Likewise, for the URS1 sequence, the least stringent sequence used to search promoters for Fig. 5 was 5′-GGCGGC (darkest blue), followed by 5′-DSGGCGGCND and TSGGCGGCTD. The most stringent was 5′-TCGGCGGCTDW (brightest blue). The International Union of Pure and Applied Chemistry nucleotide codes are as follows: D (A,G, or T); R (A or G); H (A, C, or T); V (A, C, or G); W (A or T); K (G or T); and S (C or G).
24. J. C. Jackson and J. M. Lopes, *Nucleic Acids Res.* **24**, 1322 (1996).
25. S. Kratzer and H. J. Schuller, *Mol. Microbiol.* **26**, 631 (1997); E. Dubois and F. Messenguy, *Mol. Gen. Genet.* **253**, 568 (1997).
26. H. Friesen, S. R. Hepworth, J. Segall, *Mol. Cell. Biol.* **17**, 123 (1997).
27. S. Chu *et al.*, unpublished data.
28. Cells from the wild-type W303 background (**a** *ura3 leu2 his3 trp1*) containing either the *GAL1-10* promoter integrated at *URA3* (YSC552) or the *GAL1-10* promoter fused to hemagglutinin-tagged *NDT80* (YSC553) were induced with 1.7% galactose for 3 hours as described in (*4*) before RNA was harvested. Poly(A)$^+$ RNA was prepared following the protocol of Rapid RNA Isolation from Invitrogen.
29. N. Kleckner, *Proc. Natl. Acad. Sci. U.S.A.* **93**, 8167 (1996); G. S. Roeder, *Genes Dev.* **11**, 2600 (1997).
30. D. Schild and B. Byers, *Chromosoma* **70**, 109 (1978).
31. W. Y. Miyazaki and T. L. Orr-Weaver, *Annu. Rev. Genet.* **28**, 167 (1994).
32. C. Michaelis, R. Ciosk, K. Nasmyth, *Cell* **91**, 35 (1997).
33. M. Molnar, J. Bahler, M. Sipiczki, J. Kohli, *Genetics* **141**, 61 (1995).
34. S. Chu, unpublished data.
35. O. V. Kovalenko *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **93**, 2958 (1996).
36. D. K. Palmer, K. O'Day, R. L. Margolis, *Chromosoma* **100**, 32 (1990).
37. C. S. Chan and D. Botstein, *Genetics* **135**, 677 (1993).
38. S. Biggins and S. Chu, unpublished data.
39. A. Yanai, E. Arama, G. Kilfin, B. Motro, *Oncogene* **14**, 2943 (1997).
40. F. M. Townsley and J. V. Ruderman, *Trends Cell Biol.* **8**, 238 (1998).
41. M. Glotzer, *Curr. Biol.* **5**, 970 (1995).
42. A precise deletion of only the YGR225 open reading frame prevented sporulation. Because the adjacent novel open reading frame, YGR226, is only 27 base pairs away and was also expressed as an early-mid gene, regulatory elements for YGR226 may also be affected in the YGR225 deletion mutant.
43. G. Simchen, *Genetics* **76**, 745 (1974).
44. N. Matsubara, M. Yanagisawa, Y. Nishimune, M. Obinata, Y. Matsui, *Mol. Reprod. Dev.* **41**, 407 (1995).
45. E. E. Patton, A. R. Willems, M. Tyers, *Trends Genet.* **14**, 236 (1998).
46. A. M. Neiman, *J. Cell Biol.* **140**, 29 (1998).
47. J. Mulholland *et al.*, unpublished data.
48. J. A. Printen and G. F. Sprague Jr., *Genetics* **138**, 609 (1994); K. Y. Choi, B. Satterberg, D. M. Lyons, E. A. Elion, *Cell* **78**, 499 (1994); S. Marcus, A. Polverino, M. Barr, M. Wigler, *Proc. Natl. Acad. Sci. U.S.A.* **91**, 7762 (1994).
49. L. H. Caro *et al.*, *Yeast* **13**, 1477 (1997).
50. I. Miyakawa, H. Aoi, N. Sando, T. Kuroiwa, *J. Cell Sci.* **66**, 21 (1984).
51. K. G. Hales and M. T. Fuller, *Cell* **90**, 121 (1997).
52. Y. Wada, Y. Ohsumi, Y. Anraku, *J. Biol. Chem.* **267**, 18665 (1992); Y. X. Wang, N. L. Catlett, L. S. Weisman, *J. Cell Biol.* **140**, 1063 (1998).
53. A. D. Roeder and J. M. Shaw, *Genetics* **144**, 445 (1996).
54. S. Horie *et al.*, *Mol. Cell. Biol.* **18**, 2118 (1998).
55. L. S. Ireland *et al.*, *EMBO J.* **13**, 3812 (1994).
56. D. Lydall, Y. Nikolsky, D. K. Bishop, T. Weinert, *Nature* **383**, 840 (1996).
57. A. W. Page and T. L. Orr-Weaver, *Curr. Opin. Genet. Dev.* **7**, 23 (1997); P. Sassone-Corsi, *Cell* **88**, 163 (1997).
58. Genes corresponding to YDR104, YPR007, and YGR225 were deleted by precise disruptions of each open reading frame with the *hisG-URA3* cassette. Methods were from M. Rose, F. Winston, and P. Hieter [*Methods in Yeast Genetics* (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1990)]; R. Rothstein [*Methods Enzymol.* **194**, 281 (1991)]; J. Sambrook, E. Frisch, and T. Maniatis [*Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, ed. 2, 1989)]; and F. Sherman, G. Fink, and J. Hicks [*Methods in Yeast Genetics: A Laboratory Manual* (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1986)].
59. B. Byers, in (*13*), pp. 59–96.
60. L. E. Buckingham *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **87**, 9406 (1990).
61. During the sporulation time course of YSC328, samples of cells were prepared every 30 min for microscopy. Cells were examined by Nomarski and fluorescent microscopy as described in (*3*). For electron microscopy, cells were immediately fixed in 4% glutaraldehyde [in phosphate-buffered saline at pH 6.7] for 1 to 2 hours. Fixed cells were then rapidly frozen in 20% glycerol in liquid nitrogen and held at −70°C for 48 hours. At time of use, cells were thawed on ice and processed for electron microscopy as described by B. Byers and L. Goetsch [*J. Bacteriol.* **124**, 511 (1975)].
62. J. T. Kadonaga and R. Tjian, *Proc. Natl. Acad. Sci. U.S.A.* **83**, 5889 (1986).
63. P. Spellman *et al.*, personal communication.

64. We are indebted to V. Iyer for help with genome array production. We also thank P. Spellman and S. Biggins for their help; F. Banuett, N. Kleckner, and D. Meling for useful discussions; and B. Byers, S. Klapholz, I. H. Herskowitz, Y. Kassir, and A. Mitchell for comments on the manuscript. Strains of the SK1 background and the *hisG-URA3* disruption cassette were kind gifts from N. Kleckner. S.C. and I.H. are supported by National Institutes of Health (NIH) grant AI18738. J.D. and P.O.B. are supported by the Howard Hughes Medical Institute (HHMI) and NIH grant HG00450. M.E. is supported by a postdoctoral fellowship from the Alfred E. Sloan Foundation. P.O.B. is an Associate Investigator of the HHMI. J.M. and D.B. are supported by NIH grant GM46406. This work was supported by grants from NIH and by HHMI.

13 August 1998; accepted 29 September 1998