

# Target hub proteins serve as master regulators of development in yeast

Anthony R. Borneman,<sup>1</sup> Justine A. Leigh-Bell,<sup>1</sup> Haiyuan Yu,<sup>2</sup> Paul Bertone,<sup>1,2</sup> Mark Gerstein,<sup>2</sup> and Michael Snyder<sup>1,2,3</sup>

<sup>1</sup>Department of Molecular, Cellular, and Developmental Biology, <sup>2</sup>Department of Molecular Biophysics and Biochemistry, Yale University, New Haven, Connecticut 06511, USA

To understand the organization of the transcriptional networks that govern cell differentiation, we have investigated the transcriptional circuitry controlling pseudohyphal development in *Saccharomyces cerevisiae*. The binding targets of Ste12, Tec1, Sok2, Phd1, Mga1, and Flo8 were globally mapped across the yeast genome. The factors and their targets form a complex binding network, containing patterns characteristic of autoregulation, feedback and feed-forward loops, and cross-talk. Combinatorial binding to intergenic regions was commonly observed, which allowed for the identification of a novel binding association between Mga1 and Flo8, in which Mga1 requires Flo8 for binding to promoter regions. Further analysis of the network showed that the promoters of *MGA1* and *PHD1* were bound by all of the factors used in this study, identifying them as key target hubs. Overexpression of either of these two proteins specifically induced pseudohyphal growth under noninducing conditions, highlighting them as master regulators of the system. Our results indicate that target hubs can serve as master regulators whose activity is sufficient for the induction of complex developmental responses and therefore represent important regulatory nodes in biological networks.

[*Keywords:* Binding network; ChIP chip; genomics; master regulator; yeast]

Supplemental material is available at <http://www.genesdev.org>.

Received October 31, 2005; revised version accepted December 20, 2005.

Investigating the mechanisms of cellular differentiation is fundamental to the understanding of eukaryotic development. Transcriptional regulatory networks play a key role in mediating developmental pathways through the coordinated activation of specific downstream genes (Levine and Davidson 2005). Thus, there is considerable interest in elucidating the hierarchy of networks formed by transcription factors, the combinations in which these factors associate, and the key regulatory elements within the network. Determination of the components that function together, particularly those that are likely to be weakly associated as binary interactions is particularly difficult, and new techniques are required to identify and characterize these physiologically important associations.

In some systems, key “master regulators” have been identified whose action is sufficient to induce an entire complex developmental pathway. For example, in *Drosophila melanogaster* the expression of the transcription factor *eyeless* induces ectopic eye formation, while in vertebrates, expression of the proteins MyoD, Myf5, and NeuroD can induce muscle and neural development, re-

spectively (Braun et al. 1989; Weintraub et al. 1989; Halder et al. 1995; Lee et al. 1995). Identification of these key regulatory components has provided useful insight into these developmental processes and has enormous potential for controlling them. Thus, the development of approaches to identify master regulators and understand where they reside in regulatory hierarchies is expected to be extremely valuable.

We have studied the transcriptional control of development using *Saccharomyces cerevisiae* as a model system by investigating the networks formed by several transcription factors involved in the pseudohyphal developmental pathway. When exposed to an environment with limited available nitrogen, diploid strains of *S. cerevisiae* switch from their normal, vegetative growth pattern and begin growing as pseudohyphae, in which cells form branched chains of connected cells that are capable of invading their growth substrate (Gimeno et al. 1992). Previous work identified many genes that affect pseudohyphal growth, and several of these encode transcription factors (for reviews, see Pan et al. 2000; Gancedo 2001). The genes *STE12* and *TEC1* are predicted to encode homeodomain and TEA/ATTS transcription factors, respectively, and both are required for the formation of pseudohyphae (Liu et al. 1993; Gavrias et al. 1996). These two proteins bind cooperatively to activate the expression of downstream targets under pseudohyphal growth

<sup>3</sup>Corresponding author.

E-MAIL [michael.snyder@yale.edu](mailto:michael.snyder@yale.edu); FAX (203) 432-6161.

Article published online ahead of print. Article and publication date are at <http://www.genesdev.org/cgi/doi/10.1101/gad.1389306>.

conditions (Madhani and Fink 1997). The activity of Ste12 is dependent on a MAPK cascade, which also affects the formation of pseudohyphae (Liu et al. 1993; Cook et al. 1997).

In addition to the MAPK signaling pathway, a cAMP-responsive cascade is also required for pseudohyphal growth (Kubler et al. 1997; Lorenz and Heitman 1997; Robertson and Fink 1998; Pan and Heitman 1999). The protein products of the *SOK2*, *PHD1*, *FLO8*, and *MGA1* genes are all predicted to be transcription factors controlled by the cAMP pathway during pseudohyphal growth (Ward et al. 1995; Robertson and Fink 1998; Pan and Heitman 1999; Rupp et al. 1999). Sok2 and Phd1 both contain basic helix-loop-helix DNA-binding motifs; Mga1 has a heat-shock factor (HSF) DNA-binding domain, whereas the binding domain of Flo8 is yet to be characterized (Gimeno and Fink 1994; Ward et al. 1995; Kobayashi et al. 1996; Feroli et al. 1997). Phd1 and Sok2 modulate the pseudohyphal response either positively (Phd1) or negatively (Sok2), while Flo8 and Mga1, like Ste12 and Tec1, are both required for pseudohyphal formation (Ward et al. 1995; Liu et al. 1996; Lorenz and Heitman 1998b). The targets bound by each of these factors during pseudohyphal growth have not been characterized previously, and the role of Mga1 in particular, is poorly understood. Of high interest is understanding the hierarchy of the regulatory network, the relationships of its components, and whether key elements exist in the network that can serve as master regulators.

In this study we have sought to address these issues through the investigation of the networks formed by six key regulators of pseudohyphal growth—Tec1, Ste12, Sok2, Phd1, Flo8, and Mga1—and their DNA-binding targets using chromatin immunoprecipitation and DNA microarrays (ChIP chip). ChIP chip allows for the identification of the complete genomic complement of in vivo binding sites for a DNA-binding protein, allowing for entire downstream networks to be determined without a priori knowledge of the binding specificity of that factor (Ren et al. 2000; Iyer et al. 2001). Our results demonstrate that pseudohyphal differentiation is mediated by a complex pathway involving autoregulation, cross-factor control, and feedback and feed-forward loops. Moreover, we find that Mga1 and Phd1 are key target hubs for the network and overexpression of either of these factors resulted in pseudohyphal growth under noninducing conditions. We therefore demonstrate that target hubs can serve as master regulators, with their expression being sufficient to ectopically drive complex development, making them attractive as possible control points in biological pathways.

## Results

*The transcription factors Tec1, Ste12, Sok2, Phd1, Flo8, and Mga1 bind to a large number of downstream genes*

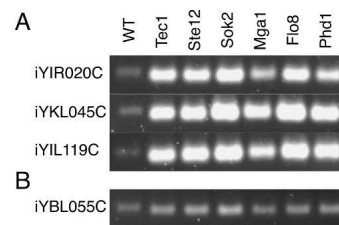
To identify the targets of Tec1, Ste12, Sok2, Phd1, Flo8, and Mga1, we tagged each protein by creating C-terminal

fusions to the *c-myc* epitope by a PCR-based approach (Longtine et al. 1998). These fusions were made independently in both  $\alpha$  and  $\alpha$  haploid strains in the pseudohyphal-competent  $\Sigma$ 1278b background, and homozygous diploids were formed by mating. The tagged strains were shown to be phenotypically indistinguishable from wild type with respect to pseudohyphal growth, and immunoblot analysis showed the presence of the tagged protein in each strain (data not shown).

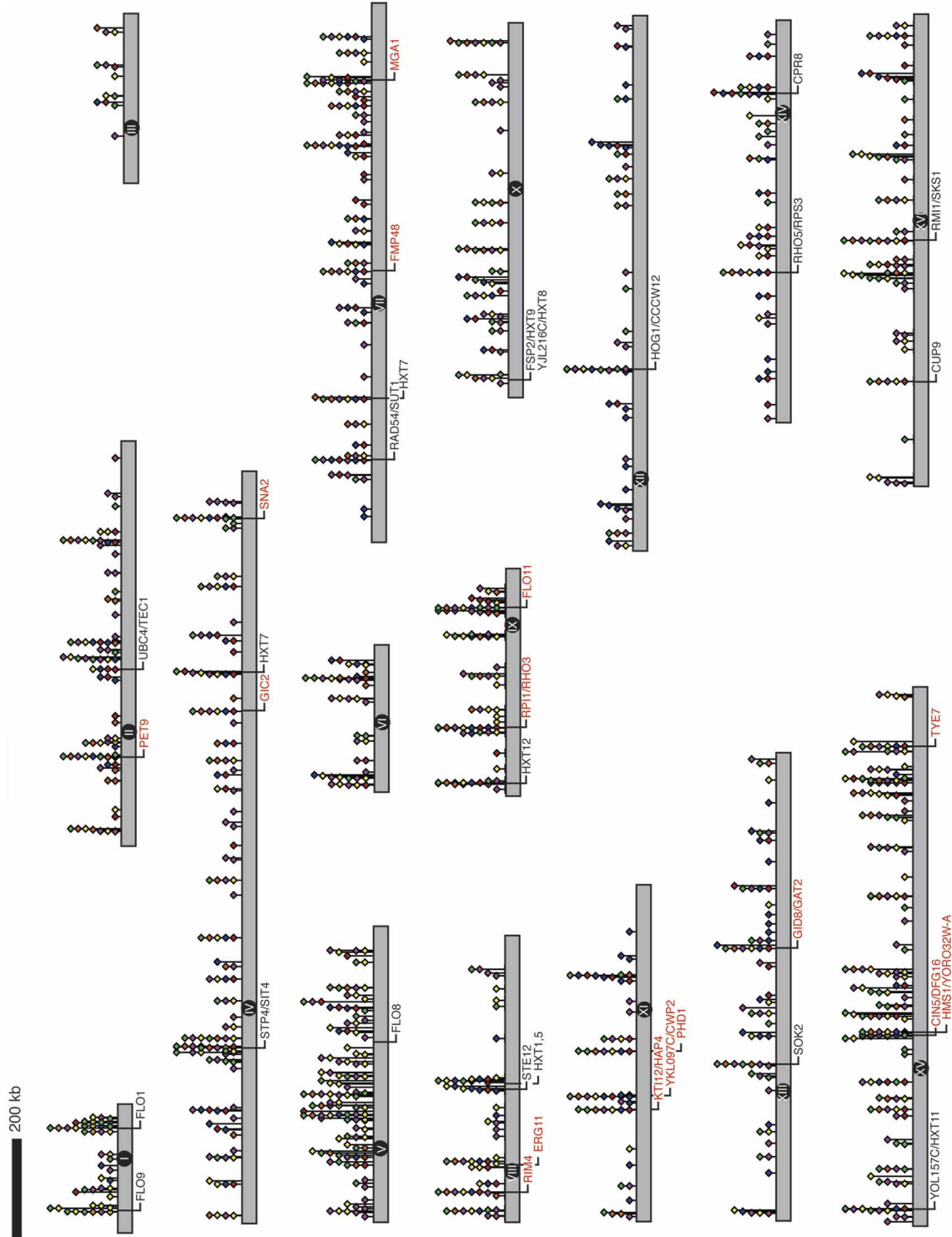
The binding targets for each of the transcription factors were determined in cells grown in liquid low nitrogen (SLAD) medium by ChIP, followed by DNA isolation and the probing of microarrays (see Materials and Methods). At least five independent biological replicates were performed for each factor, and following normalization, a stringent cut-off was used to identify intergenic regions whose tagged versus untagged ratios were  $>4$  standard deviations (SD) from the mean (Horak et al. 2002). At least 40 binding targets for each transcription factor from across the range of standard deviations were also shown to be enriched in the immunoprecipitations by PCR with an  $\sim 90\%$  success rate, with PCR results for three intergenic regions bound by all six factors shown in Figure 1.

Each transcription factor bound to a large number of intergenic regions, ranging from 116 for Mga1 to 306 for Phd1. Collectively, 625 unique intergenic regions were bound by at least one of the six factors (see Supplementary Table 1); the targets were distributed across all 16 *S. cerevisiae* chromosomes (Fig. 2). The enriched intergenic regions were subsequently mapped to the promoter regions of 699 *S. cerevisiae* open reading frames (ORFs). Only ORFs whose 5' ends were downstream of, and directly flanking, the intergenic target region were scored as gene targets for that factor (Supplementary Table 2).

Once the lists of target genes were established, "GO term finder" was used to search for specific Gene Ontology (GO) terms that were enriched in the gene targets of these factors relative to the entire yeast genome (Table 1; Boyle et al. 2004). Several classes of genes were enriched by the ChIP chip procedure, including the "filamentous growth" ( $p = 2.8 \times 10^{-3}$ ) and "growth" ( $p = 7.5 \times 10^{-3}$ ) categories, which contain large numbers of genes known to be involved in pseudohyphal growth. Interestingly,



**Figure 1.** PCR confirmation of factor binding. Three intergenic regions that were shown to be bound by all six factors (iYIR020C, iYKL045C and iYIL119C), in addition to a region shown to be unbound (iYBL055C), were amplified from ChIP samples from either an untagged control strain (WT) or strains containing *c-myc*-tagged Tec1, Ste12, Sok2, Mga1, Flo8, or Phd1.



**Figure 2.** Tec1, Ste12, Sok2, Phd1, Mga1, and Flo8 bind to discrete sites across all 16 chromosomes of *S. cerevisiae*. The binding sites for Tec1 (red), Ste12 (blue), Sok2 (yellow), Phd1 (pink), Mga1 (orange), and Flo8 (green) were mapped across the intergenic regions of *S. cerevisiae*. When binding was observed by multiple factors to the same intergenic region or to closely adjacent regions, diamonds were stacked above each other to ensure visibility. The position of key genes (black text), in addition to those genes whose 5' regions are bound by all six factors (red text), are indicated below each chromosome.

**Table 1.** GO term enrichment by ChIP chip

GO term	ChIP	Gen	P value	Genes
Hexose transport	2.1%	0.3%	$1.6 \times 10^{-7}$	<i>HXT9, HXT1, HXT5, HXT2, SKS1, HXT6, HXT4, HXT3, HXT16, HXT11, HXT8, HXT7, HXT13</i>
Amine transport	2.8%	0.7%	$1.5 \times 10^{-4}$	<i>CAN1, AGP1, ORT1, GAP1, YMR088C, GNP1, DUR3, PTK2, TPO4, HIP1, TPO1, PUT4, UGA4, PTK1, BAP2, NPR2, SAM3</i>
Alcohol metabolism	5.3%	2.2%	$6.2 \times 10^{-4}$	<i>ERG2, GID8, ERG5, ERG11, SOR2, DAK2, RHR2, PSA1, MUQ1, ERG28, TDH2, MDH2, ERG6, NDE1, HOR2, ENO2, STD1, TDH3, HMG1, RKI1, TOS3, ERG25, GAL2, PDR16, PGI1, MIG2, ARE2, ATG26, HXK1, ZWF1, SOR1, NCP1, TYE7</i>
Filamentous growth	3.6%	1.2%	$2.8 \times 10^{-3}$	<i>SHO1, HMS1, FLO11, BMH1, SOK2, PEA2, MSS11, DIA3, MGA1, BUD8, PHD1, GPA2, MEP2, DFG16, MSB2, TEC1, NRG2, STE12, DOT6, DIA2, RGA2, DIG1</i>
Growth	4.1%	1.6%	$7.5 \times 10^{-3}$	<i>SHO1, HMS1, FLO11, BMH1, SOK2, KCC4, PEA2, ACT1, MSS11, DIA3, MGA1, PHD1, BUD8, GPA2, MEP2, DFG16, MSB2, TEC1, NRG2, STE12, DOT6, DIA2, TOS2, RGA2, DIG1</i>
Ion transport	3.9%	1.5%	$8.3 \times 10^{-5}$	<i>FRE4, ATO3, PMP1, SIT1, GIT1, ARN1, FET4, ATP1, HIP1, ENA1, PMP2, INH1, PIC2, MEP1, FET3, VCX1, FTR1, NHX1, MEP2, PHO89, PHO91, RCS1, COT1, FRE1</i>

other categories such as “hexose transport” ( $p = 1.6 \times 10^{-7}$ ) and “alcohol metabolism” ( $p = 6.2 \times 10^{-4}$ ) were also highly enriched in the target gene list. This indicates that there are very strong links between pseudohyphal growth and carbon metabolism, in agreement with previous studies (Lambrechts et al. 1996).

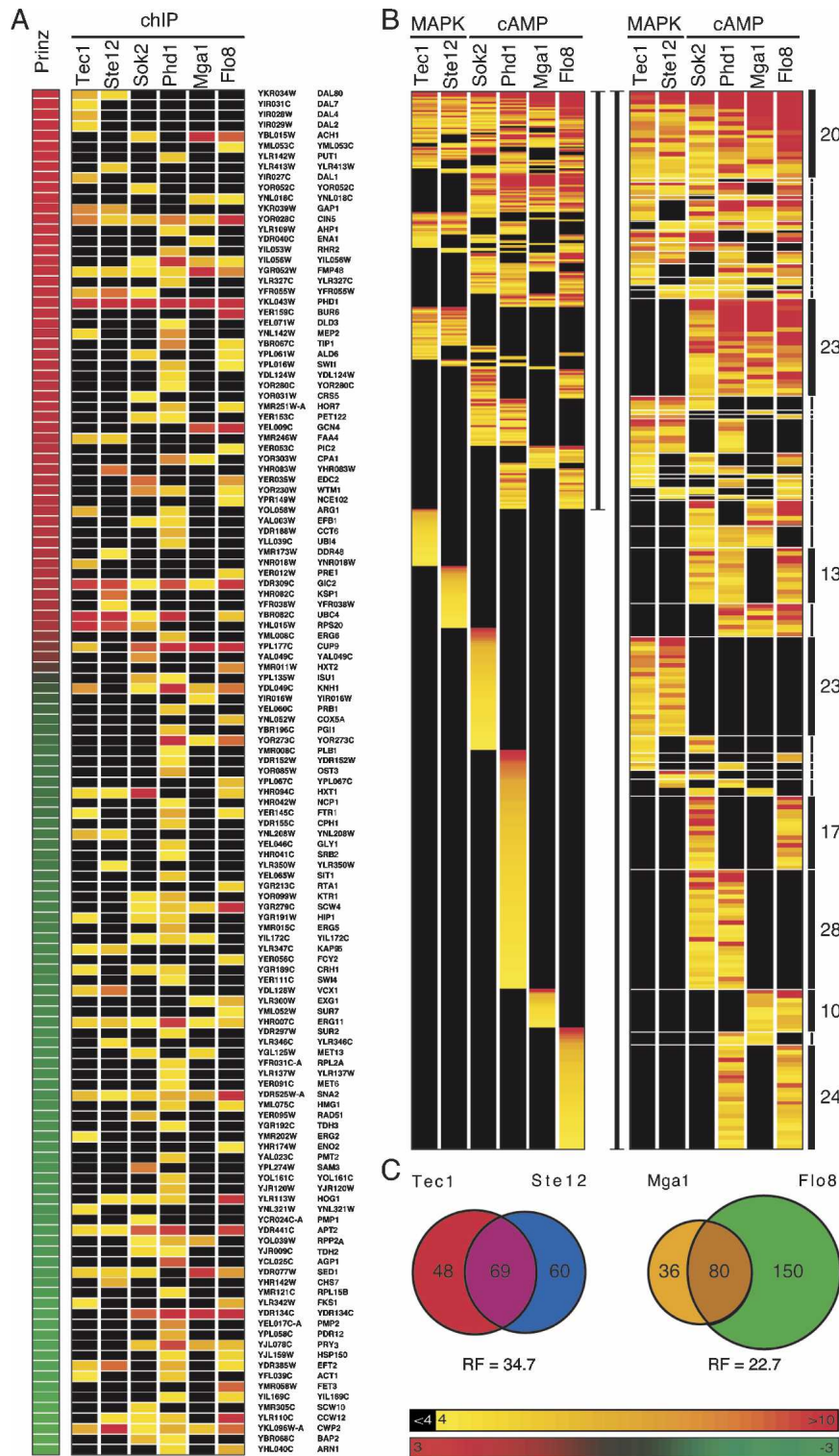
To further investigate the validity of the targets of the transcription factors identified by ChIP chip, the list of target genes was compared with expression microarray experiments performed with pseudohyphal cells (Prinz et al. 2004). Of the 699 genes identified by the ChIP chip, direct partners could be found for 638 in the expression data set (Ty elements were excluded as these were scored differently by the two studies); 131 (20.5%) of these genes were shown to also be significantly regulated either positively or negatively (from 873 regulated genes) (Prinz et al. 2004) during pseudohyphal growth (Fig. 3A). GO category enrichment showed that many genes involved in nitrogen compound metabolism were bound and up-regulated, especially the *DAL* genes that are involved in allantoin metabolism, while those involved in cell wall organization and sterol metabolism were bound and down-regulated.

#### *A high degree of combinatorial binding exists for downstream targets genes*

The intergenic targets of the six factors were examined to determine the proportion of genes bound by multiple factors. A large number of promoter regions ( $n = 246$ , 40%) were bound by at least two factors, and 20 intergenic regions, many of which are upstream of genes associated with pseudohyphal growth, including *FLO11*, *PHD1*, *MGA1*, and *HMS1*, were bound by all six (Figs. 2,

3B; Gimeno and Fink 1994; Lo and Dranginis 1998; Lorenz and Heitman 1998b). These data therefore support previous work which showed that *Tec1*, *Ste12*, *Sok2*, and *Flo8* each control the expression of *Flo11* (Rupp et al. 1999; van Dyk et al. 2005). Our work demonstrates that *Phd1* and *Mga1* also bind *Flo11* and, in addition, that large numbers of transcription factors often bind concurrently to the promoters of pseudohyphal genes.

Following this initial analysis, the factors were clustered by the similarity of their target lists, resolving the proteins into two main groups. Group 1 contained *Ste12* and *Tec1*, while the second group contained *Sok2*, *Phd1*, *Flo8*, and *Mga1*. This distribution of factors is particularly interesting as it follows the particular signaling pathway that lies upstream of each protein, with *Tec1* and *Ste12* lying downstream of the MAPK cascade and *Sok2*, *Phd1*, and *Flo8* all associated with the cAMP pathway (Liu et al. 1993; Madhani and Fink 1997; Robertson and Fink 1998; Pan and Heitman 1999, 2000; Rupp et al. 1999). To examine the differences between these two pathways, the downstream targets unique to either group were also examined for GO category enrichment. The major category enriched in the *Ste12* and *Tec1* group was “pseudohyphal growth” ( $p = 1.4 \times 10^{-4}$ ), followed distantly by “allantoin metabolism” ( $p = 0.02$ ). In contrast, the major category enriched in the cAMP-regulated group was “hexose transport” ( $p = 7.9 \times 10^{-4}$ ), followed by “amine transport” ( $p = 6 \times 10^{-3}$ ). Interestingly, the cAMP group also contained many members of several other categories related to the metabolism and catabolism of various compounds, including alcohol, hexose sugars, steroid compounds, and organic acids. Thus, the MAPK and cAMP pathways transcriptionally activate distinct functional groups of proteins.



**Figure 3.** Binding by Tec1, Ste12, Sok2, Phd1, Mga1, and Flo8 correlate with pseudo-hyphal-expressed genes and also reveal pathway-specific signatures and associated binding pairs. (A) The 131 genes shown to be both regulated during the induction of pseudo-hyphal growth (Prinz et al. 2004) and bound by Tec1, Ste12, Sok2, Phd1, Mga1, or Flo8. Expression changes are indicated on a red-green log<sub>2</sub> color scale (up-regulated in pseudo-hyphal cells, red; down-regulated, green). ChIP binding significance is indicated by the red-yellow color scale, (black, binding less than SD = 4 statistical cut-off). (B) The intergenic regions bound by the MAPK factors Tec1 and Ste12 and the cAMP factors Sok2, Phd1, Mga1, and Flo8 were clustered according to the combination of factors that bound to each specific region. Horizontal lines represent intergenic regions, with the probability of binding indicated by the yellow-red color scale. Those regions that were bound by at least one factor are shown on the left, while only those bound by two or more factors were enlarged and are shown to the right. Those combinatorial classes that contain  $\geq 10$  intergenic regions are highlighted by the thick black bars. (C) Venn diagrams of the overlap in the intergenic regions bound by Ste12 and Tec1 (left) and Flo8 and Mga1 (right). The RF, which calculates the amount of overlap above that due to random sampling, and the probability for this enrichment are indicated below each Venn diagram.

Further examination of the target clustering revealed that certain factors showed a specific preference for binding to the same intergenic regions. Pair-wise combinatorial binding in the data set was investigated to determine if significant overlap existed for any pairs of factors. Ste12 and Tec1, the two components of the MAPK

group, had previously been shown to bind cooperatively to promoter regions under pseudo-hyphal conditions (Madhani and Fink 1997). We found that these two factors shared 69 common promoter target regions. A representation factor of 34.7 (RF; a measure of the observed number of overlapping genes compared with the ex-

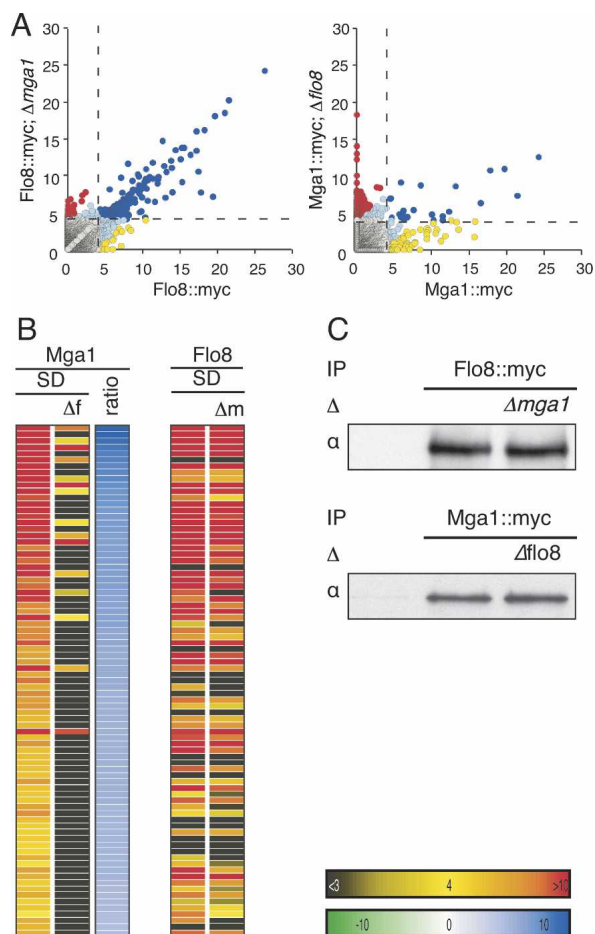
pected number) was calculated for this overlap ( $p < 1.98 \times 10^{-99}$ ), supporting their established cooperative interaction (Fig. 3C). However, there were also numerous intergenic regions that uniquely bound either Tec1 ( $n = 60$ ) or Ste12 ( $n = 48$ ), suggesting that this codependence may not be absolute.

In addition to the Ste12–Tec1 combination, there was also significant enrichment among all of the transcription factors pairs, reflecting their common function in pseudohyphal growth (lowest RF = 9.5,  $p < 2.9 \times 10^{-33}$ ). However, the only factors whose pair-wise enrichment approached that of Ste12 and Tec1 were Flo8 and Mga1 (RF = 22.7,  $p < 1.5 \times 10^{-96}$ ), two factors that had not previously been shown to associate. Flo8 and Mga1 bound to 80 common promoter regions, representing 69% of the total binding sites for Mga1 (Fig. 3C). These two factors therefore represented an attractive candidate for forming a novel binding pair and were selected for further study.

#### *Mga1 and Flo8 display a novel binding association*

To investigate the relationship between Mga1 and Flo8, reciprocal deletions were made in the diploid *MGA1::MYC* and *FLO8::MYC* backgrounds. These new strains (*MGA1::MYC/MGA1::MYC flo8Δ/flo8Δ*, and *FLO8::MYC/FLO8::MYC mga1Δ/mga1Δ*) were then used for ChIP chip. Comparison of the results obtained with the *FLO8::MYC/FLO8::MYC mga1Δ/mga1Δ* strain with those of the wild type revealed no significant difference in the binding of Flo8 to target promoters ( $R^2 = 0.6$ ) (Fig. 4A; Supplementary Table 3). In contrast, the binding of Mga1 to intergenic regions was severely reduced in the *flo8Δ/flo8Δ* strain relative to wild type ( $R^2 = 0.06$ ), with the majority of the wild-type Mga1 targets either falling below the  $SD \geq 4$  cut-off or showing a greatly reduced binding strength (average reduction of 4.8 SD units) (Fig. 4A; see Supplementary Table 4). In addition to the intergenic regions that showed a reduction in binding by Mga1 in the *flo8Δ/flo8Δ* strain, there are a small number of regions with faint array signals that showed modest increases in binding strength. Further testing of these regions by PCR indicated that they were not bound by Mga1 and thus represent noise in the array comparisons. This is in direct opposition to the regions that showed a reduction in binding in the *flo8Δ* strain, which were reproducibly confirmed by PCR as yet (data not shown).

Interestingly, when the Mga1 target regions were then examined for Flo8 binding, 73% of all of the fragments that showed a  $>3$  SD unit change in binding strength between the Mga1-tagged wild-type and Mga1-tagged *flo8Δ* strains were also shown to be binding targets of Flo8 (Fig. 4B). This difference was not due to alterations in the protein levels between the strains (Fig. 4C). Thus, the dependence of Mga1 on Flo8 for target binding is not simply due to an alteration in the levels of the Mga1 protein, but rather to a loss in binding by Mga1; this is likely due to interactions between the two factors as the effect is only observed to occur



**Figure 4.** Mga1 requires Flo8 for binding to DNA. (A) ChIP chip was performed on strains containing either Flo8::myc protein in both wild-type and *mga1Δ/mga1Δ* backgrounds or Mga1::myc protein in both wild-type and *flo8Δ/flo8Δ* backgrounds. The SD values for enrichment of each intergenic region from both of these strains were plotted against each other with the  $SD = 4$  cut-off (dotted lines) used to calculate enrichment in each experiment. Individual intergenic regions were classified as either being enriched in both strain backgrounds (dark blue), enriched only in the mutant strain with  $>3$  SD units difference between strains (red), enriched only in the wild-type strain with  $\geq 3$  SD units difference (yellow), or not enriched in either experiment or enriched only in one strain but with  $<3$  SD units difference between experiments (gray). (B, left) Intergenic regions were sorted by the difference in SD values (blue–green scale) obtained from ChIP chip experiments using *MGA1::MYC/MGA1::MYC* (Mga1) and *MGA1::MYC/MGA1::MYC flo8Δ/flo8Δ* (Mga1 Δf) strains. (Right) The SD values obtained for the same intergenic regions using both *FLO8::MYC/FLO8::MYC* (Flo8) and *FLO8::MYC/FLO8::MYC mga1Δ/mga1Δ* (Flo8 Δm) (red–yellow scale), in addition to their ratio (blue–green scale), were also aligned with those of Mga1. (C) Western blot analysis of Flo8::myc and Mga1::myc protein levels in wild-type and mutant strains as indicated. Two-hundred-fifty OD<sub>600</sub> units of cells grown under nitrogen-limiting conditions were lysed, immunoprecipitated, and subjected to Western blot analysis using anti-c-myc antibodies (mouse monoclonal 9E10).

at promoter regions where both proteins are normally present.

*A complex binding network is formed by Tec1, Ste12, Sok2, Phd1, Mga1, and Flo8*

To investigate the network hierarchies in which the six transcription factors operate, we determined how their binding relationships were organized among themselves and with other transcription factors. Analysis of the downstream targets of the six factors revealed that they often bound to the promoter regions of the other factors used in this study, such that the six factors form a highly interconnected network (Fig. 5A). In recent years there has been significant progress in characterizing biological networks mathematically to assist in the study of these complex systems (for review, see Barabasi and Oltvai 2004). Networks can therefore be represented by a series of nodes (in this case transcription factors and their targets) and edges (representing binding by a factor to its target). Nodes can then be classified by a number of properties, with the most basic being the number of outgoing edges ( $k_{out}$ ), representing the number of targets bound by each factor, and by the number of incoming edges ( $k_{in}$ ), representing the number of factors that bind to upstream of a particular gene.

When this analysis was applied to the network formed by the pseudohyphal transcription factors, we found that the  $k_{out}$  was relatively uniform (between three and four), whereas the  $k_{in}$  varied significantly between factors, ranging from zero for Flo8 to six for Mga1 and Phd1 (Fig. 5A). In addition, many of the downstream targets of the pseudohyphal factors are also transcription factors. For many of these proteins, binding data are available from other high-throughput studies, although not from diploid cells exposed to nitrogen starvation conditions (Harbison et al. 2004). Despite these differences in growth environment, the targets of Tec1, Ste12, Sok2, Phd1, Mga1, and Flo8 were overlaid with these other factors to obtain a broad approximation of the total network environment that could be encountered by each factor. This expanded list of factors was shown to form an even more connected binding network than that observed for the pseudohyphal factors alone (average  $k_{out} = 8.5$ ; average  $k_{in} = 4$ ) (Fig. 5B).

To facilitate the visualization of the interactions between the factors, this expanded network was broken down into several simple network patterns (autoregulation, cross-factor control, and feedback and feed-forward loops) (Fig. 5C). Four of the factors (eight in the extended network) displayed binding to their own promoter region (autoregulation), including Tec1, Ste12, Mga1, and Phd1. There were also nine cases of cross-factor control, where pairs of factors each bound to the promoter region of the reciprocal factor. Of the factors used in this study, Phd1 and Mga1 were the most likely to form this type of interaction. In addition to direct cross-factor control, there were 23 instances of feedback loops, where three proteins form a closed loop; although due to the reciprocal binding nature of many of the factors, 22 of these may

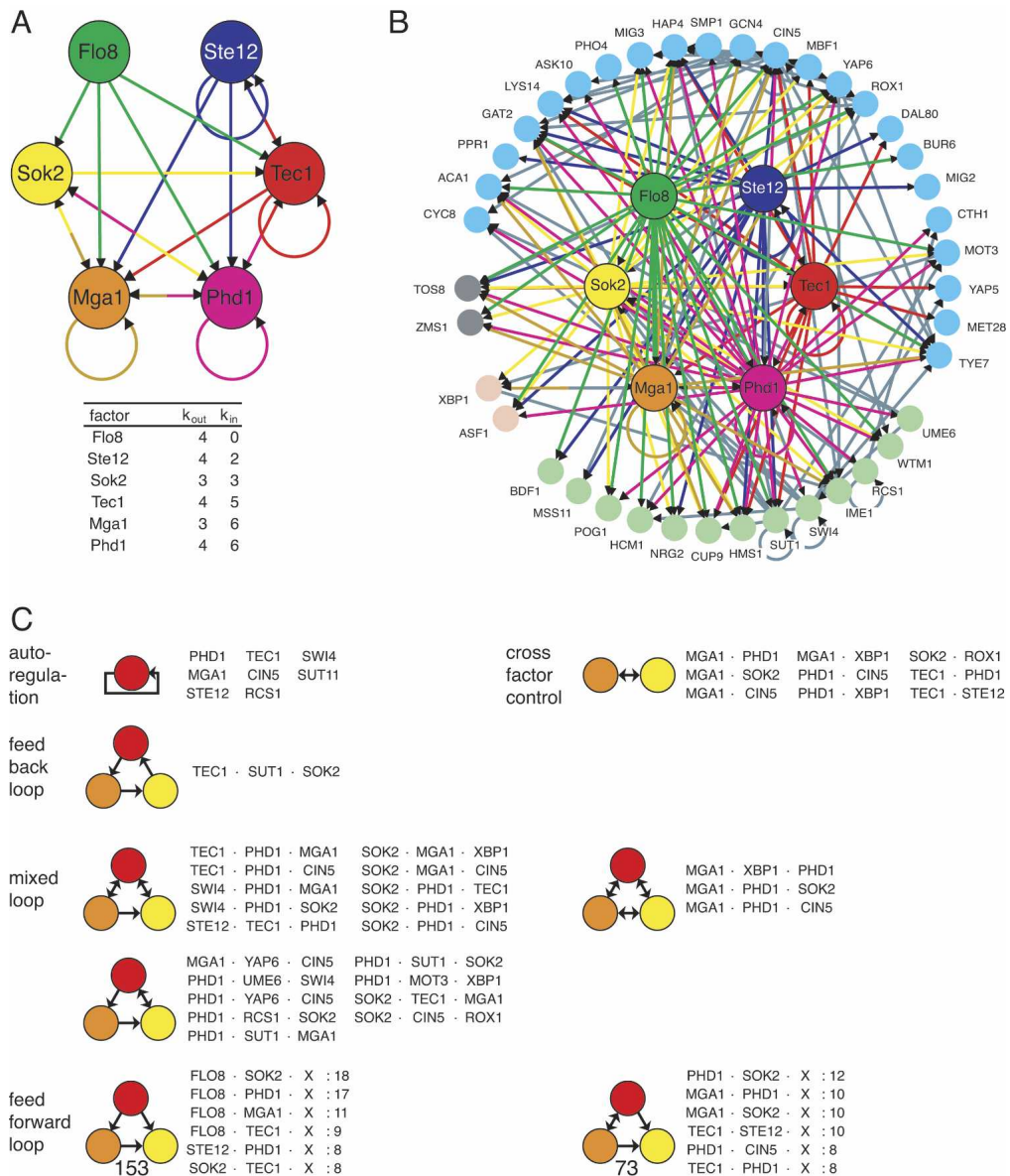
also function as feed-forward loops and were therefore classified as mixed loops. Again, Phd1 and Mga1 were the predominate factors participating in the feedback and mixed loop categories. Finally, true feed-forward loops were also observed in the network, with this pattern classified by one protein binding to the promoter of a second through a tertiary intermediate as well as binding it directly. This was the most common type of pattern, with 226 observed in the expanded network and with Flo8 representing the most prevalent member.

*Constitutive expression of the target hub proteins Mga1 and Phd1 induces pseudohyphal growth under noninducing conditions*

Network analysis revealed that Mga1 and Phd1 had the most incoming connections ( $k_{in} = 6$  for both proteins when using only the factors from this study,  $k_{in} = 9$  for Phd1,  $k_{in} = 10$  for Mga1 when other factors are included). These two factors therefore represent key "target hubs" in this network as their numerous interactions make them significant points for coordinating pseudohyphal growth signals. We hypothesized that Mga1, Phd1, and possibly Tec1 (which had five incoming connections) may be master regulators whose expression would be capable of ectopically activating the developmental pathway.

To test this possibility, the ORFs of Phd1, Mga1, and Tec1 along with Ste12, Sok2, and Flo8 were cloned into plasmid vectors that expressed the proteins from the constitutive *ADH3* promoter (Melcher 2000). Strains were grown under high nitrogen conditions (SC-URA), which normally suppress the pseudohyphal pathway. Ectopic expression of Ste12, Sok2, and Flo8 produced cells with morphologies indistinguishable from the empty vector control strain, although the expression of Flo8 made cells clump together significantly more than observed with the vector control (Fig. 6A,B). However, the expression of either Mga1, Phd1, or Tec1 caused cells to acquire pseudohyphal characteristics as evidenced by several criteria. First, overexpression of Mga1, Phd1, and Tec1 all stimulated agar invasion, although expression of Tec1 produced far less invasiveness than either Phd1 or Mga1 (Fig. 6A). Second, strains overexpressing Mga1, Phd1, and Tec1 were elongated and formed chains of pseudohyphal cells when grown in high nitrogen liquid media. The chains and clumps of Mga1 and Phd1 overexpressing strains were much larger than those observed with Tec1, correlating with their higher levels of invasiveness (Fig. 6B). Neither of these morphological features was observed for Sok2, Ste12, and Flo8 overexpressing cells.

Finally, to examine the ectopic induction of pseudohyphal growth at the molecular level, strains overexpressing Mga1 and Phd1 were subjected to expression microarray analysis. Overexpression of Mga1 and Phd1 were shown to induce 211 and 214 genes, respectively (twofold change cut-off). There was a high degree of correlation between the results obtained with the two factors, indicating that the overexpression of either factor



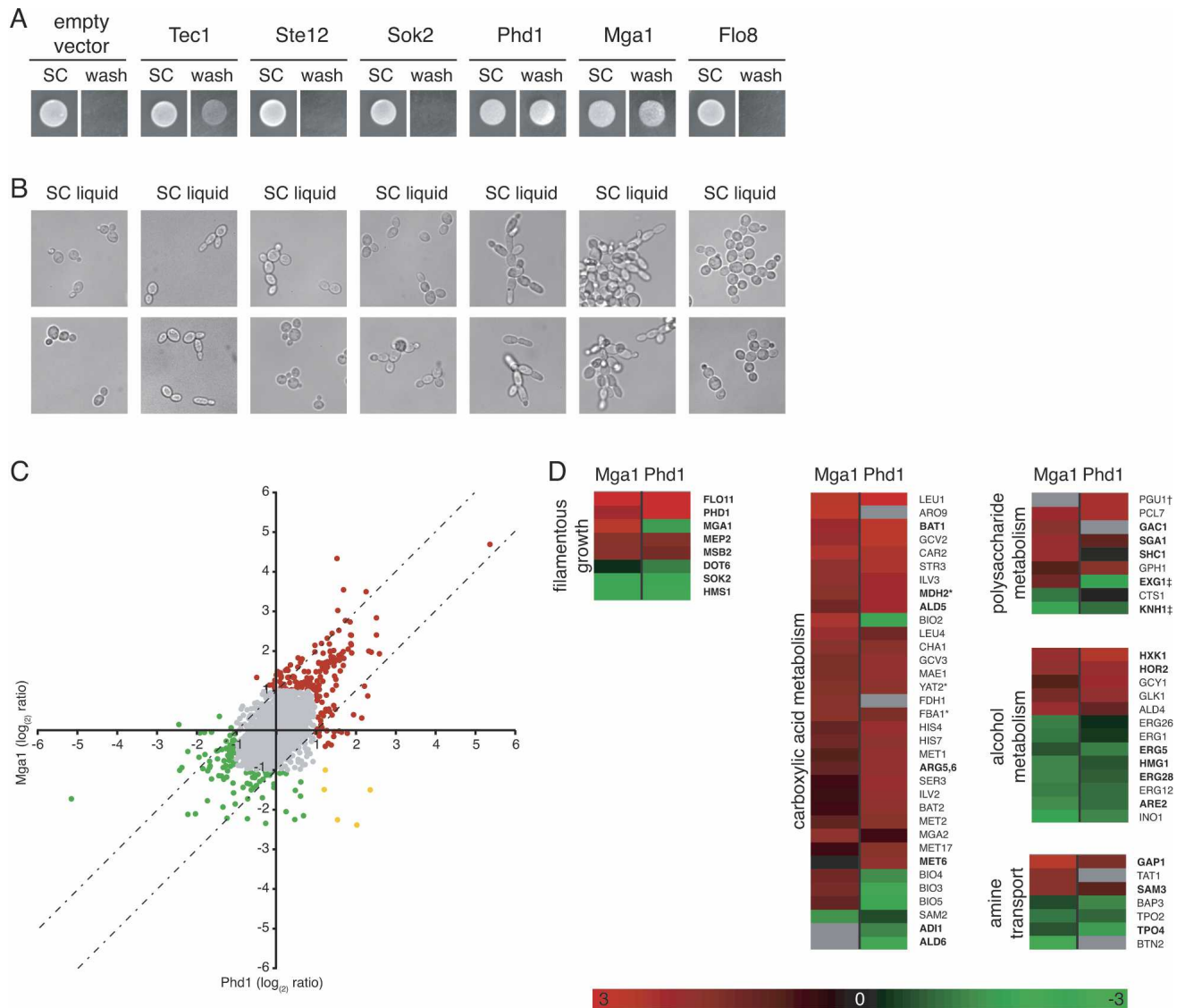
**Figure 5.** A complex network is formed by the pseudohyphal factors and their downstream targets. (A) The binding network was calculated for Tec1, Ste12, Sok2, Phd1, Mga1, and Flo8. Incoming edges (direction indicated by the arrowhead) represent binding to the promoter region of that factor, while the color of the edge indicates the factor that was shown to bind to this region. The number of outgoing ( $k_{out}$ ) and incoming edges ( $k_{in}$ ) for each factor are listed below the network schematic. (B) The binding network for the six factors was expanded by identifying all of the downstream targets that are also transcription factors and then integrating their binding target data into the network (Harbison et al. 2004). Downstream factors are shaded by their major GO term (light blue, transcription; light green, cell growth and/or maintenance; pink, stress; gray, unknown), with their outgoing edges shaded gray. (C) The numbers of five distinct types of simple regulatory patterns in the expanded network (auto-regulation, cross-factor control, feedback loops, feed-forward loops, and mixed loops). The members of each class were identified using Mfinder1.1 (Milo et al. 2002) and are listed to the right of each motif diagram. Due to the large number of feed-forward loops observed in the network, the total number of individual patterns are listed below each pattern diagram, with the top six pairs of factors participating in each pattern and their level of occurrence listed (X = any protein).

produces very similar phenotypes at both the phenotypic and mRNA expression levels (Fig. 6C).

The genes that were shown to be regulated in either experiment were analyzed for GO category enrichment to determine the correlation between those genes affected by the overexpression of Mga1 and Phd1 and those

bound by the pseudohyphal factors. The major GO categories that were shown to be enriched in the overexpression gene list were those involved in metabolism, with “carboxylic acid metabolism” being the most enriched member. This category also contains several genes that were bound as determined by the ChIP chip





**Figure 6.** Overexpression of Phd1 and Mga1 induce pseudohyphal growth. (A) Suspensions of diploid cells expressing Tec1, Ste12, Sok2, Phd1, Mga1, and Flo8 from the constitutive *ADH3* promoter were plated on SC-URA and incubated for 24 h at 30°C. Colonies were photographed (SC), and then the cells growing on the surface of the plate were removed by gentle agitation, after which the colonies were photographed again (wash). (B) Diploid cells expressing each of the factors used in A were grown for 16 h in liquid SC-URA and then examined microscopically using DIC optics. (C) The log<sub>2</sub> ratios of mRNA levels between either *ADH3::MGA1* (X-axis) or *ADH3::PHD1* (Y-axis) and the vector control were determined by expression microarray analysis. The ratios obtained for each ORF were plotted against each other to correlate gene expression during the overexpression of the two different factors. ORFs are shaded by their response to overexpression of the factors; twofold or more up-regulation under at least one condition (red), twofold or more down-regulation under at least one condition (green), twofold or more up-regulation under one condition, twofold or more down-regulation under the other (yellow), twofold or less up- or down-regulation (gray). (D) GO category enrichment of genes regulated twofold or more by the overexpression of either Mga1 or Phd1. Individual genes are listed next to their log<sub>2</sub> ratios, determined during overexpression of both Mga1 and Phd1 (red–green scale, bottom), with those also identified as a binding target of Tec1, Ste12, Sok2, Phd1, Mga1 or Flo8 highlighted in bold text. Several genes were found to be members of more than one GO category, and these are indicated as follows: (\*) also in “alcohol metabolism”, (†) also in “filamentous growth”.

experiments (Fig. 6D). In addition, the list was examined for those genes that were regulated upon the overexpression of either Mga1 and Phd1 and that were bound by at least one factor in the CHIP chip experiments. Of these 83 genes, the major GO category shown to be enriched was “filamentous growth” (Fig. 6D). Interestingly, all of

the genes shown to be involved in filamentous growth in the overexpression list except one (*PGU1*) were also shown to be bound in the CHIP. There were also a number of other GO categories that, while not being significantly enriched as determined by our very strict criteria, did contain a large number of genes that were regulated

upon expression of Mga1 or Phd1 and that were shown to be bound by the ChIP chip experiments. These also represented metabolic genes, strengthening the links between morphogenesis and metabolism in this pathway (Fig. 6D).

Comparison of the Mga1 and Phd1 overexpression targets with those found to be regulated during “normal” pseudohyphal growth (Prinz et al. 2004), revealed that many of the overexpression-regulated genes (20%–69% depending on the precise scoring method applied to the expression data) (see Supplementary Table 5) are also regulated during the pseudohyphal response. Thus, using several criteria, including cell morphology, invasiveness, and expression patterns, it was shown that the ectopic expression of Mga1 and Phd1 was driving pseudohyphal growth under noninducing conditions. This occurred through the regulation of pathways similar to those expressed under normal pseudohyphal growth conditions.

## Discussion

### *The targets of transcription factors involved in pseudohyphal growth reflect their upstream signaling pathways*

Although transcriptional regulation is critical for mediating many cellular and developmental processes, there is still a great deal to be learned concerning how transcriptional networks are organized to control gene expression. In particular, our understanding of the degree to which factors cooperate to regulate downstream genes and how the hierarchy of networks is structured is quite limited. Pseudohyphal growth is one example of this type of complex developmental process, requiring a large array of gene products, including several transcription factors (Gimeno and Fink 1994; Ward et al. 1995; Gavrias et al. 1996; Liu et al. 1996; Chandralapaty and Errede 1998; van Dyk et al. 2005). While a large number of genes have been identified as having important roles in signaling and transcriptional regulation during pseudohyphal growth, the downstream targets of these factors remained largely undiscovered. This work has sought to address this shortcoming by applying ChIP chip to globally identify the downstream transcriptional networks that exist for several factors that regulate pseudohyphal growth. The targets of Ste12, Tec1, Sok2, Flo8, and Mga1 were shown to be enriched for a variety of specific classes of genes, including many that had previously been shown to be involved in this developmental process.

It was possible to categorize transcription factors into two broad groups based on the clustering of their gene targets; Ste12 and Tec1 formed one group and Sok2, Phd1, Mga1, and Flo8 represented a second. This target-based classification mirrors the upstream signaling pathways that have been established for these factors, with Ste12 and Tec1 lying downstream of the MAPK pathway, whereas Sok2, Phd1, Flo8, and Mga1 all have links to the cAMP pathway (Liu et al. 1993; Madhani and Fink

1997; Robertson and Fink 1998; Pan and Heitman 1999, 2000, 2002; Rupp et al. 1999). We identified several groups of genes that were bound specifically in either the MAPK or cAMP groups, showing that each pathway has a unique function(s) required for different aspects of pseudohyphal growth. The targets of the MAPK group primarily focused on morphology, while the cAMP group contained many genes involved in catabolism and metabolism, suggesting that the cAMP pathway may help to coordinate developmental changes with the metabolic alterations that accompany the nutrient-limiting environment encountered by pseudohyphal cells. This division of function between the two pathways may therefore account for why two independent signaling pathways have evolved to induce pseudohyphal growth, with the MAPK pathway and Tec1/Ste12 regulating morphogenesis, whereas the cAMP pathway with Sok2/Phd1/Flo8/Mga1 regulates the many metabolic changes that accompany the pseudohyphal transition.

While the targets of three factors examined in this study, Tec1, Ste12, and Sok2, have been investigated by ChIP chip previously (Lee et al. 2002; Zeitlinger et al. 2003), those studies used haploid cells grown in rich media or treated with N-butanol, a condition with uncertain physiological significance with respect to pseudohyphal growth (Martinez-Anaya et al. 2003). Consequently these studies failed to identify many known targets of Ste12, Tec1, and Sok2 during pseudohyphal growth, including Flo11 for Tec1 and Sok2 (Rupp et al. 1999). Nonetheless, we did observe some overlap between those studies and ours (10%–15% overlap each for Ste12, Tec1, or Sok2). As our study analyzed both more factors and used conditions more physiologically associated with pseudohyphal growth, it is expected to provide a more comprehensive and accurate assessment of the gene targets involved in this process.

Interestingly, a comparison of the ChIP chip data with expression microarray data obtained using pseudohyphal cells produced only a small, although significant, overlap. While the discrepancies may be due in part to the difference between the solid and liquid medium, other studies have also noted that binding is not always reflective of alterations in gene expression as determined by microarrays (Martone et al. 2003; Euskirchen et al. 2004). Most likely, multiple factors help to regulate gene expression, and the specific combination of bound proteins ultimately controls the output (positive or negative) of gene expression.

### *A complex transcriptional network controls pseudohyphal development*

Ste12, Tec1, Sok2, Flo8, and Mga1 and their downstream targets form a highly connected network that contains binding patterns consistent with the presence of autoregulatory loops, cross-talk between factors, and feedback and feed-forward loops. While the complexity of the pathway may suggest robustness in the network to loss of individual factors, this appears to not be the case as strains lacking any of four of the factors examined in this

study (Tec1, Ste12, Flo8, and Mga1) fail to undergo the pseudohyphal response (Liu et al. 1993, 1996; Gavrias et al. 1996; Lorenz and Heitman 1998a). This may be due to the combinatorial binding displayed by many of the factors. Loss of one factor may actually disrupt the formation of tertiary complexes at promoter regions. However the connectedness of the network does allow for precise control and coordination of events required for mediating this complex process. This was highlighted by the expression assays that showed that overexpression of either Mga1 or Phd1 led to very similar expression profiles. The interconnectedness of the pathway therefore allows for similar overall results to be obtained from the activation of different key upstream factors.

#### *Binding by Mga1 to promoter regions requires the presence of Flo8*

Tec1 and Ste12 bind cooperatively to promoter regions to activate expression during the induction of pseudohyphal growth (Madhani and Fink 1997). In addition, Flo8 has been shown to bind cooperatively with the protein Mss11 to activate Flo11 expression (van Dyk et al. 2005). Cooperation between multiple factors therefore appears to be a recurring theme during pseudohyphal growth. By comparing the target lists of the different factors, we were able to confirm the association between Ste12 and Tec1 as well as uncover a novel association between Flo8 and Mga1. The Mga1–Flo8 association was subsequently confirmed by performing ChIP chip using *mga1Δ* and *flo8Δ* mutant strains. Interestingly, the effect observed between Flo8 and Mga1 is unlike the situations observed for either Tec1 and Ste12 or for Flo8 and Mss11 as it appears that the association between the factors is non-reciprocal, since Mga1 requires Flo8 for binding whereas Flo8 binds DNA independently of Mga1.

While Mga1 required Flo8 for DNA binding as assayed by ChIP, the formaldehyde cross-linking cannot discriminate between direct DNA binding and interactions facilitated through a secondary factor. Mga1 may therefore bind to DNA directly, with Flo8 stabilizing this binding, or alternatively, Mga1 may bind indirectly through Flo8. Yeast two-hybrid experiments and coimmunoprecipitation studies have so far failed to detect a stable interaction between Mga1 and Flo8, suggesting that the proteins likely require a tertiary intermediate, perhaps DNA, for their apparent association at promoter regions (A.R. Borneman and M. Snyder, unpubl.).

Regardless of whether Mga1 requires Flo8 to stabilize binding to DNA or whether Mga1 interacts with DNA indirectly through Flo8, this study has shown for the first time that it is possible to detect novel interactions in transcriptional networks through the global study of their specific target genes. This type of study could therefore be applied to large-scale studies of transcription factor networks, by identifying pairs of factors that share significant numbers of common gene targets and then examining their binding using ChIP chip in relevant mutant strains to highlight possible cooperativity. This technique also provides the major benefit of being

able to uncover protein associations, such as those discovered for Mga1 and Flo8, which would be difficult to define using classical techniques, given the failure of two-hybrid and coimmunoprecipitation to detect any biochemical interaction, despite a clear association between the two factors.

#### *Target hubs as key regulators of cellular and developmental processes*

Master regulators are defined as proteins whose ectopic expression is capable of inducing developmental processes. Interestingly, these proteins do not have to be essential for that process, as shown previously for MyoD and Phd1 (Rudnicki et al. 1993; Gimeno and Fink 1994). Likewise, other components of the pathway can be essential for a process but not serve as master regulators. Master regulators are therefore presumed to be key components whose ectopic activation affects a number of critical components that can then mediate the entire developmental process.

In this study we identified two highly connected hubs that are targets of many transcription factors and postulated that these may serve as master regulators of the pathway. Overexpression of *PHD1* has been shown previously to induce pseudohyphal growth (Gimeno and Fink 1994). We have confirmed this in our assays and further demonstrated that overexpression of Mga1, and to a lesser extent Tec1, can induce the pseudohyphal developmental pathway in nitrogen-rich medium, whereas the nontarget hubs proteins could not. Several other master regulators (MyoD, Myf5, and NeuroD) have also shown to be highly regulated in their respective developmental networks (Chae et al. 2004; Tapscott 2005). Thus, we suggest that in many cases, target hubs may be key regulators of developmental pathways and that our approach of extensively mapping transcription factor-binding sites may be a useful way to identify them. Global mapping of transcription factor-binding sites therefore represents a useful approach for identifying such hubs in a variety of other systems. The subsequent identification and activation of these factors could therefore be used to potentially control deficiencies in upstream components, and have enormous therapeutic potential for the treatment of a variety of medical disorders.

## Materials and methods

### *Yeast strains and growth conditions*

*S. cerevisiae* strains were all derived from Y825 (MATa, *ura3-52*,  $\Delta$ *leu2*) and Y826 (MAT $\alpha$ , *ura3-52*,  $\Delta$ *leu2*), which both have a  $\Sigma$ 1278b background. For the induction of pseudohyphal gene expression, *S. cerevisiae* strains were grown to OD<sub>600</sub> = 0.3 at 30°C in liquid minimal medium (SD; 2% glucose, 0.5% [38 mM] ammonium sulphate, 0.17% yeast nitrogen base, 5 mM uracil), harvested by vacuum filtration and transferred to liquid SLAD medium (2% glucose, 50  $\mu$ M ammonium sulphate, 0.17% yeast nitrogen base, 5 mM uracil) for an additional 4-h growth at 30°C (Madhani et al. 1999).

*Gene deletion, epitope tagging, and Western blotting*

*S. cerevisiae* genes were either deleted or epitope-tagged by the long-primer PCR approach using the plasmid pFA6a-KanMX (deletion) or pFA6a-13myc-KanMX (tagging)—which both contain the KanMX gene (G418<sup>R</sup>) as a selectable marker—as a template (Lorenz et al. 1995; Longtine et al. 1998). The sequence of the primers used for the deletions and tagging are listed in Supplementary Table 6 (tagging primers are indicated by “tag”, deletion primers by “Δ”). For the *Mga1::myc*; *Δflo8* and *Flo8::myc*; *Δmga1* strains, the KanMX cassette of the deletion allele was replaced by the NatMX (nourseothricin [NAT]<sup>R</sup>) cassette (Goldstein and McCusker 1999) by transformation with a PCR fragment of the NatMX cassette.

PCR products were transformed into the haploid strains Y825 and Y826 using a lithium acetate (LiAc) technique (Gietz and Woods 2002). Transformants were selected by growth on YPD plates containing 200 mg/mL of G418 or 150 mg/mL NAT. Strains were tested for correct integration by PCR.

Western blots were performed on protein lysates prepared from induced (see above) cells. Following SDS-PAGE and transfer to nitrocellulose, fusion proteins were detected using anti-myc antibody (monoclonal 9E10 1:2000 dilution) and visualized by chemiluminescence.

*ChIP*

ChIPs were performed by a method modified from Horak et al. (2002) and Horak and Snyder (2002). Briefly, 500 mL cultures of both an epitope-tagged and an untagged control were induced and then fixed by the addition of formaldehyde (1% final concentration). Cell pellets were resuspended in 2.5 mL lysis buffer (0.1% deoxycholic acid, 1 mM EDTA, 50 mM HEPES-KOH at pH 7.5, 140 mM NaCl, 1% Triton X-100 with the addition of protease inhibitors) per gram of cells. Three milliliters of this suspension was disrupted using Zirconia beads and then sonicated to shear chromatin (Branson sonifier, 3 × 30 sec, power setting 5, 50% duty cycle). Lysates were clarified, and anti-myc antibody was added (monoclonal antibody 9E10, 1:250 dilution). Samples were incubated overnight at 4°C with nutation. Protein A/G-coupled beads (Pierce) were then added, and samples were incubated for an additional 2 h at 4°C. Beads were sequentially washed for 5 min at 4°C in lysis buffer, high-salt lysis buffer (0.1% deoxycholic acid, 1 mM EDTA, 50 mM HEPES-KOH at pH 7.5, 500 mM NaCl, 1% Triton X-100), LiCl detergent solution (0.5% deoxycholic acid, 1 mM EDTA, 250 mM LiCl, 0.5% NP-40, 10 mM Tris-HCl at pH 8), and finally TBS (20 mM Tris-HCl at pH 7.5, 150 mM NaCl). Immunoprecipitated protein-DNA complexes were eluted from the beads by incubation in TES (10 mM Tris, 1 mM EDTA, 0.75% SDS) for 20 min at 65°C. Eluted samples were incubated overnight at 65°C to reverse the formaldehyde cross-linking, treated with proteinase K, and phenol-chloroform extracted to remove the proteins.

*DNA labeling and array hybridization*

DNA obtained from the ChIP was purified using the MinElute PCR purification kit (Qiagen). The ChIP chip samples were then labeled using Klenow polymerase and random primers (Bioprime kit, Invitrogen) in the presence of 240 μM amino-allyl dUTP (aa-dUTP). aa-dUTP-incorporated samples were then chemically coupled to either Cy5 (tagged) or Cy3 (untagged) dyes (amino-allyl cDNA labeling kit, Ambion). Following coupling, Cy5 and Cy3 sample pairs were mixed and applied to microarrays containing PCR products representing the intergenic regions of the *S. cerevisiae* genome (Iyer et al. 2001)

printed on UltraGAPS slides (Corning), and then processed according to Corning UltraGAPS hybridization and washing protocols.

*Data acquisition and analysis*

Slides were scanned using a Genepix 4000A scanner (Axon Instruments) and Genepix 3.0 (Axon Instruments). Following scanning, Genepix result (.gpr) files were uploaded to the ExpressYourself Web site (<http://array.mbb.yale.edu/analysis>; Luscombe et al. 2003) for normalization. Enriched spots were determined using the Cy3 noise method (Horak et al. 2002). A statistically stringent cut-off ( $SD \geq 4$ ) was used for this scoring (false-positive rate of ~0.001). Intensity shading and clustering were performed using the Genespring microarray analysis platform (Silicon Genetics).

GO enrichment of the ChIP chip target gene lists was calculated using “GO term finder” (Boyle et al. 2004), which uses a very stringent *p*-value test that corrects for multiple testings. Pair-wise binding enrichment values and probability calculations were performed using the “Statistical significance of the overlap between two groups of genes” resource ([http://elegans.uky.edu/MA/progs/overlap\\_stats.html](http://elegans.uky.edu/MA/progs/overlap_stats.html)), which calculates RFs and hypergeometric probabilities for any two groups of genes (for details of the calculations, see [http://elegans.uky.edu/MA/progs/representation\\_stats.html](http://elegans.uky.edu/MA/progs/representation_stats.html)).

*PCR verification*

PCR verification of the ChIP chip results was performed by designing oligonucleotides to 40–45 enriched and three nonenriched intergenic regions for each factor and then performing 30 cycles of PCR on 20 ng of DNA from both the tagged and untagged strains. Band intensities from each PCR reaction were compared by gel electrophoresis.

*Transcription factor overexpression*

For the overexpression of the transcription factors, each ORF was PCR amplified and cloned into the expression plasmid pVTU260 by gap repair (Melcher 2000). DNA for each ORF was amplified from the strains used for the ChIP chip assays such that the expressed protein would contain the same *c-myc* epitope tag at the C terminus. The primers used for this PCR are listed in Supplementary Table 5 (“oe” primers). PCR products were cotransformed into *S. cerevisiae* *NheI*, *NcoI* digested pVTU260. Plasmid DNA was rescued from *URA*<sup>+</sup> transformants for each factor into *Escherichia coli*, and the resulting plasmids were screened for correct integration by restriction digestion. Plasmids were then retransformed into a diploid strain created from the haploid strains Y825 and Y826. Strains were tested for protein production by Western blot using the *c-myc* epitope tag for detection, with two independent transformants that were shown to produce tagged protein used for further analysis.

*Expression microarrays*

Poly-A RNA was isolated from diploid cells containing either the pVTU::MGA1::MYC and pVTU::PHD1::MYC plasmids or the empty pVTU260 vector that had been grown for two generations in SC-URA (Ambion). Dye-labeled cDNA was prepared from 1 μg of poly-A RNA (amino-allyl cDNA labeling kit, Ambion) and hybridized to microarrays that were spotted with 70-mer oligonucleotide probes representing all of the ORFs of *S. cerevisiae* (Yeast AROS version 1.1, Qiagen).

## Acknowledgments

We thank Tom Royce for assistance with the ExpressYourself Web site; Ken Nelson for assistance in array printing; and Stephen Hartman, Daniel Gelperin, Lee Kung, and Michael Smith for critical reading of this manuscript. This work was supported by Burroughs Wellcome and NIH grants to M.S and M.G.

## References

- Barabasi, A.L. and Oltvai, Z.N. 2004. Network biology: Understanding the cell's functional organization. *Nat. Rev. Genet.* **5**: 101–113.
- Boyle, E.I., Weng, S., Gollub, J., Jin, H., Botstein, D., Cherry, J.M., and Sherlock, G. 2004. GO::TermFinder: Open source software for accessing Gene Ontology information and finding significantly enriched Gene Ontology terms associated with a list of genes. *Bioinformatics* **20**: 3710–3715.
- Braun, T., Bober, E., Buschhausen-Denker, G., Kohtz, S., Grzeschik, K.H., and Arnold, H.H. 1989. Differential expression of myogenic determination genes in muscle cells: Possible autoactivation by the Myf gene products. *EMBO J.* **8**: 3617–3625.
- Chae, J.H., Stein, G.H., and Lee, J.E. 2004. NeuroD: The predicted and the surprising. *Mol. Cell* **18**: 271–288.
- Chandarlapaty, S. and Errede, B. 1998. Ash1, a daughter cell-specific protein, is required for pseudohyphal growth of *Saccharomyces cerevisiae*. *Mol. Cell Biol.* **18**: 2884–2891.
- Cook, J.G., Bardwell, O., and Thorner, J. 1997. Inhibitory and activating functions for MAPK Kss1 in the *S. cerevisiae* filamentous-growth signalling pathway. *Nature* **390**: 85–88.
- Euskirchen, G., Royce, T.E., Bertone, P., Martone, R., Rinn, J.L., Nelson, F.K., Sayward, F., Luscombe, N.M., Miller, P., Gerstein, M., et al. 2004. CREB binds to multiple loci on human chromosome 22. *Mol. Cell Biol.* **24**: 3804–3814.
- Feroli, F., Carignani, G., Pavanello, A., Guerreiro, P., Azevedo, D., Rodrigues-Pousada, C., Melchiorretto, P., Panzeri, L., and Agostoni Carbone, M.L. 1997. Analysis of a 17.9 kb region from *Saccharomyces cerevisiae* chromosome VII reveals the presence of eight open reading frames, including BRF1 (TFIIB70) and GCN5 genes. *Yeast* **13**: 373–377.
- Gancedo, J.M. 2001. Control of pseudohyphae formation in *Saccharomyces cerevisiae*. *FEMS Microbiol. Rev.* **25**: 107–123.
- Gavrias, V., Andrianopoulos, A., Gimeno, C.J., and Timberlake, W.E. 1996. *Saccharomyces cerevisiae* *TEC1* is required for pseudohyphal growth. *Mol. Microbiol.* **19**: 1255–1263.
- Gietz, R.D. and Woods, R.A. 2002. Transformation of yeast by lithium acetate/single-stranded carrier DNA/polyethylene glycol method. *Methods Enzymol.* **350**: 87–96.
- Gimeno, C.J. and Fink, G.R. 1994. Induction of pseudohyphal growth by overexpression of *PHD1*, a *Saccharomyces cerevisiae* gene related to transcriptional regulators of fungal development. *Mol. Cell Biol.* **14**: 2100–2112.
- Gimeno, C.J., Ljungdahl, P.O., Styles, C.A., and Fink, G.R. 1992. Unipolar cell divisions in the yeast *S. cerevisiae* lead to filamentous growth: Regulation by starvation and RAS. *Cell* **68**: 1077–1090.
- Goldstein, A.L. and McCusker, J.H. 1999. Three new dominant drug resistance cassettes for gene disruption in *Saccharomyces cerevisiae*. *Yeast* **15**: 1541–1553.
- Halder, G., Callaerts, P., and Gehring, W.J. 1995. Induction of ectopic eyes by targeted expression of the *eyeless* gene in *Drosophila*. *Science* **267**: 1788–1792.
- Harbison, C.T., Gordon, D.B., Lee, T.I., Rinaldi, N.J., Macisaac, K.D., Danford, T.W., Hannett, N.M., Tagne, J.B., Reynolds, D.B., Yoo, J., et al. 2004. Transcriptional regulatory code of a eukaryotic genome. *Nature* **431**: 99–104.
- Horak, C.E. and Snyder, M. 2002. ChIP–chip: A genomic approach for identifying transcription factor binding sites. *Methods Enzymol.* **350**: 469–483.
- Horak, C.E., Luscombe, N.M., Qian, J., Bertone, P., Piccirillo, S., Gerstein, M., and Snyder, M. 2002. Complex transcriptional circuitry at the G1/S transition in *Saccharomyces cerevisiae*. *Genes & Dev.* **16**: 3017–3033.
- Iyer, V.R., Horak, C.E., Scafe, C.S., Botstein, D., Snyder, M., and Brown, P.O. 2001. Genomic binding sites of the yeast cell-cycle transcription factors SBF and MBF. *Nature* **409**: 533–538.
- Kobayashi, O., Suda, H., Ohtani, T., and Sone, H. 1996. Molecular cloning and analysis of the dominant flocculation gene *FLO8* from *Saccharomyces cerevisiae*. *Mol. Gen. Genet.* **251**: 707–715.
- Kubler, E., Mosch, H.U., Rupp, S., and Lisanti, M.P. 1997. Gpa2p, a G-protein  $\alpha$ -subunit, regulates growth and pseudohyphal development in *Saccharomyces cerevisiae* via a cAMP-dependent mechanism. *J. Biol. Chem.* **272**: 20321–20323.
- Lambrechts, M.G., Bauer, F.F., Marmur, J., and Pretorius, I.S. 1996. Muc1, a mucin-like protein that is regulated by Mss10, is critical for pseudohyphal differentiation in yeast. *Proc. Natl. Acad. Sci.* **93**: 8419–8424.
- Lee, J.E., Hollenberg, S.M., Snider, L., Turner, D.L., Lipnick, N., and Weintraub, H. 1995. Conversion of *Xenopus* ectoderm into neurons by NeuroD, a basic helix–loop–helix protein. *Science* **268**: 836–844.
- Lee, T.I., Rinaldi, N.J., Robert, F., Odom, D.T., Bar-Joseph, Z., Gerber, G.K., Hannett, N.M., Harbison, C.T., Thompson, C.M., Simon, I., et al. 2002. Transcriptional regulatory networks in *Saccharomyces cerevisiae*. *Science* **298**: 799–804.
- Levine, M. and Davidson, E.H. 2005. Gene regulatory networks for development. *Proc. Natl. Acad. Sci.* **102**: 4936–4942.
- Liu, H., Styles, C.A., and Fink, G.R. 1993. Elements of the yeast pheromone response pathway required for filamentous growth of diploids. *Science* **262**: 1741–1744.
- . 1996. *Saccharomyces cerevisiae* S288C has a mutation in *FLO8*, a gene required for filamentous growth. *Genetics* **144**: 967–978.
- Lo, W.S. and Dranginis, A.M. 1998. The cell surface flocculin Flo11 is required for pseudohyphae formation and invasion by *Saccharomyces cerevisiae*. *Mol. Biol. Cell.* **9**: 161–171.
- Longtine, M.S., McKenzie III, A., Demarini, D.J., Shah, N.G., Wach, A., Brachat, A., Philippsen, P., and Pringle, J.R. 1998. Additional modules for versatile and economical PCR-based gene deletion and modification in *Saccharomyces cerevisiae*. *Yeast* **14**: 953–961.
- Lorenz, M.C. and Heitman, J. 1997. Yeast pseudohyphal growth is regulated by *GPA2*, a G protein  $\alpha$  homolog. *EMBO J.* **16**: 7008–7018.
- . 1998a. The *MEP2* ammonium permease regulates pseudohyphal differentiation in *Saccharomyces cerevisiae*. *EMBO J.* **17**: 1236–1247.
- . 1998b. Regulators of pseudohyphal differentiation in *Saccharomyces cerevisiae* identified through multicopy suppressor analysis in ammonium permease mutant strains. *Genetics* **150**: 1443–1457.
- Lorenz, M.C., Muir, R.S., Lim, E., McElver, J., Weber, S.C., and Heitman, J. 1995. Gene disruption with PCR products in *Saccharomyces cerevisiae*. *Gene* **158**: 113–117.
- Luscombe, N.M., Royce, T.E., Bertone, P., Echols, N., Horak, C.E., Chang, J.T., Snyder, M., and Gerstein, M. 2003. ExpressYourself: A modular platform for processing and visualizing microarray data. *Nucleic Acids Res.* **31**: 3477–3482.

- Madhani, H.D. and Fink, G.R. 1997. Combinatorial control required for the specificity of yeast MAPK signaling. *Science* **275**: 1314–1317.
- Madhani, H.D., Galitski, T., Lander, E.S., and Fink, G.R. 1999. Effectors of a developmental mitogen-activated protein kinase cascade revealed by expression signatures of signaling mutants. *Proc. Natl. Acad. Sci.* **96**: 12530–12535.
- Martinez-Anaya, C., Dickinson, J.R., and Sudbery, P.E. 2003. In yeast, the pseudohyphal phenotype induced by isoamyl alcohol results from the operation of the morphogenesis checkpoint. *J. Cell. Sci.* **116**: 3423–3431.
- Martone, R., Euskirchen, G., Hartman, S., Royce, T.E., Luscombe, N.M., Rinn, J.L., Nelson, F.K., Miller, P., Gerstein, M., Weissman, S., et al. 2003. Distribution of NF- $\kappa$ B-binding sites across human chromosome 22. *Proc. Natl. Acad. Sci.* **100**: 12247–12252.
- Melcher, K. 2000. A modular set of prokaryotic and eukaryotic expression vectors. *Anal. Biochem.* **277**: 109–120.
- Milo, R., Shen-Orr, S., Itzkovitz, S., Kashtan, N., Chklovskii, D., and Alon, U. 2002. Network motifs: Simple building blocks of complex networks. *Science* **298**: 763–764.
- Pan, X. and Heitman, J. 1999. Cyclic AMP-dependent protein kinase regulates pseudohyphal differentiation in *Saccharomyces cerevisiae*. *Mol. Cell Biol.* **19**: 4874–4887.
- . 2000. Sok2 regulates yeast pseudohyphal differentiation via a transcription factor cascade that regulates cell–cell adhesion. *Mol. Cell Biol.* **20**: 8364–8372.
- . 2002. Protein kinase A operates a molecular switch that governs yeast pseudohyphal differentiation. *Mol. Cell Biol.* **22**: 3981–3993.
- Pan, X., Harashima, T., and Heitman, J. 2000. Signal transduction cascades regulating pseudohyphal differentiation of *Saccharomyces cerevisiae*. *Curr. Opin. Microbiol.* **3**: 567–572.
- Prinz, S., Avila-Campillo, I., Aldridge, C., Srinivasan, A., Dimitrov, K., Siegel, A.F., and Galitski, T. 2004. Control of yeast filamentous-form growth by modules in an integrated molecular network. *Genome Res.* **14**: 380–390.
- Ren, B., Robert, F., Wyrick, J.J., Aparicio, O., Jennings, E.G., Simon, I., Zeitlinger, J., Schreiber, J., Hannett, N., Kanin, E., et al. 2000. Genome-wide location and function of DNA binding proteins. *Science* **290**: 2306–2309.
- Robertson, L.S. and Fink, G.R. 1998. The three yeast A kinases have specific signaling functions in pseudohyphal growth. *Proc. Natl. Acad. Sci.* **95**: 13783–13787.
- Rudnicki, M.A., Schnegelsberg, P.N.J., Stead, R.H., Braun, T., Arnold, H.H., and Jaenisch, R. 1993. MyoD or Myf-5 is required for the formation of skeletal muscle. *Cell* **75**: 1351–1359.
- Rupp, S., Summers, E., Lo, H.J., Madhani, H., and Fink, G. 1999. MAP kinase and cAMP filamentation signaling pathways converge on the unusually large promoter of the yeast *FLO11* gene. *EMBO J.* **18**: 1257–1269.
- Tapscott, S.J. 2005. The circuitry of a master switch: MyoD and the regulation of skeletal muscle gene transcription. *Development* **132**: 2685–2695.
- van Dyk, D., Pretorius, I.S., and Bauer, F.F. 2005. Mss11p is a central element of the regulatory network that controls *FLO11* expression and invasive growth in *Saccharomyces cerevisiae*. *Genetics* **169**: 91–106.
- Ward, M.P., Gimeno, C.J., Fink, G.R., and Garrett, S. 1995. SOK2 may regulate cyclic AMP-dependent protein kinase-stimulated growth and pseudohyphal development by repressing transcription. *Mol. Cell Biol.* **15**: 6854–6863.
- Weintraub, H., Tapscott, S.J., Davis, R.L., Thayer, M.J., Adam, M.A., Lassar, A.B., and Miller, A.D. 1989. Activation of muscle-specific genes in pigment, nerve, fat, liver, and fibroblast cell lines by forced expression of MyoD. *Proc. Natl. Acad. Sci.* **86**: 5434–5438.
- Zeitlinger, J., Simon, I., Harbison, C.T., Hannett, N.M., Volkert, T.L., Fink, G.R., and Young, R.A. 2003. Program-specific distribution of a transcription factor dependent on partner transcription factor and MAPK signaling. *Cell* **113**: 395–404.