Supplementary Figures

Figure S1: Response of selected genes to osmotic shock. X-axis: time (min) after exposure to 0.5M KCL. Y-axis: Transcription level (expression ratios in $\log_2$ scale from O'Rourke and Herskowitz 2004 normalized to standard Gaussians). **(A)** Hog1 downregulates transcription of known Ste12 target genes (FUS3, FUS1, KSS1 and TEC1) in WT. The same downregulation is observed in *sho1* strains, indicating Hog1-dependent downregulation which is not dependent on Sho1. **(B)** Hog1 upregulates transcription of its known downstream targets (STL1, HOR2 and GRE2) in WT. There is no upregulation in *hog1* and *pbs2* mutants, but high levels are observed in the double mutants *ssk1ste11* and *ssk1sho1*, although the two inputs from the HOG pathway upstream branches are blocked. Hence, our method predicts that either Pbs2 or Ssk2/22 are activated by an alternative third input.

Figure S2: The regulatory units used for network expansion. In principle, we wish to consider every possible regulatory unit of one or two regulators out of 29 variables (435 regulatory units in total). However, given our specific dataset of 106 conditions, many of these regulatory units are equivalent: The specific genetic perturbations performed in the network (brown x-s) cause some downstream variables to have similar predicted activities. For example, no distinction between the effect of Ssk1 and Ssk2/22 is possible given our data. Hence, under our dataset, the model is partitioned into 12 sets of equivalent regulators (circumscribed by dashed green lines). The putative regulatory units can be determined up to these sets only, and thus we used a single representative regulator for each of the sets (marked in green or yellow oval), and got only 78 (one or two out of 12) different regulatory units. Among them, six regulatory units have a single known logic (black diamonds). The other 72 putative units can have any possible logic (3^9 logics in case of two regulators with 3 states). Units containing only environmental stimuli (yellow ovals) are model-independent, and all other units are model-dependent. In order to test the hypothesis that Hog1 has two different activity modes, two alternative Hog1/Pbs2 variables were included in possible regulatory units during the expansion analysis: *Hog1* is activated (through Pbs2) by the two upstream branches of the HOG pathway Sho1-Ste11 and Sln1-Ssk1 and a third uncharacterized input (this is the logic suggested by the refinement procedure), and $Hog1^{(2)}$ is activated (through $Pbs2^{(2)}$) only by these two upstream branches.

Figure S3: Expression levels of CTT1 and HSP12 as a function of their regulators Msn2/4.
A: Predicted activity levels of Msn2/4 (upper row) and observed expression levels of their targets CTT1 and HSP12 (lower matrix) in each of the 106 conditions (columns). Expression levels are shown in $\log_2$ scale. **B,C,D**: Similar plots with observed expression levels of MSN2 and MSN4. In plots A,B,C, columns are sorted by increasing levels of the upper row. In plot D, columns are sorted by increasing level of MSN4 for three MSN2 expression level ranges: low (<-0.2, left), medium (>-0.2 and <0.2, middle) and high (> 0.2, right). In plot A, where predicted activities are computed based only on the network model and ignoring Msn2/4 data, the expression levels of CTT1 and HSP12 are predicted well by the model. However, expression levels of CTT1/HSP12 cannot be predicted well by either MSN2 expression level (B), MSN4 expression level (C), or a combinatorial logic of both expression levels (D). Hence, it is hard to decipher the Msn2/4 module based only on the dataset used in this work. Instead, our approach integrates the prior knowledge into a model that accounts for post-transcriptional effects (i.e. phosphorylation by PKA and interaction with Hog1 in the nucleus) in order to predict the activity of Msn2/4. We assign genes to the Msn2/4 module based on the fit between the target's measured expression levels and Msn2/4 predicted activities.

Figure S4: **The expansion improvement score.** Distributions of the model-dependent normalized Bayesian score (left), model-independent normalized Bayesian score (middle) and the improvement score (right). The distributions present the scores of all 5700 genes analyzed. Each plot gives the score distribution on true (green) and shuffled data (blue), generated by randomly permuting the experimental procedure labels while keeping the measurements intact. The random data are used to estimate the significance of the normalized Bayesian scores. To avoid a bias in the original Bayesian score (i.e., the maximal likelihood score of a regulatory unit), we standardize scores by subtracting from this original score the highest score obtained for a single Gaussian model (i.e., a model that best fits the distribution of the candidate gene to a single Gaussian, without any regulators). The *normalized Bayesian score* of a candidate gene assignment is the highest standardized score obtained for this gene with any regulatory unit. Both true model-dependent and model-independent normalized Bayesian score distributions are enriched with highly scored genes. Hence, in order to discriminate between model-dependent and model-independent genes, we used an improvement score, which is the difference between model-dependent and model-independent Bayesian score. Indeed, significant improvements are observed in the true data only.

Figure S5: Modules size distribution. The number of large modules obtained on true data (red) is significantly higher than the number of modules obtained using shuffled data (blue). Based on the true data, 71 modules contain more than three genes. In contrast, no module of size > 3 was detected using the shuffled data (for readability, the bottom left corner of the distribution is enlarged). We focus our analysis on five known modules with more than 10 genes (solid red arrows) and five novel modules with more than 20 genes (dashed red arrows)**.**

Figure S6: Functional coherence and separation of modules. (A,B) Enrichment of the genes in each module (rows) in various experimental conditions (columns). Enrichment is represented by a distinct behavior of the genes of the module compared to the rest of the genome (hyper-geometric p-value) in ChIP profiles (A, Harbison et al. 2004) and gene expression profiles (B, Gasch et al. 2000). The profiles used for the enrichment tests were not part of our original dataset. Grey: Bonferroni-corrected p-value $\leq 10^{-5}$ (A) and $\leq 10^{-3}$ (B). **(C)** Separation between expression profiles of different modules. The same stress conditions from Gasch et al. (2000) as in B were used. Modules separation is computed by KS-test with Bonferroni-correction. **(D)** The distribution of distance between genes of the same module (red) and of different modules (green), for modules Ssk2/22 and Hog1B. Distance between genes is the Euclidian distance between their normalized expression vectors across the 106 conditions. Since the inter-module and intra-module distance distributions are very similar, it is very hard to identify the modules using the data alone. In contrast, our expansion procedure distinguishes the modules by utilizing both data and model. **(E)** The predicted separation is corroborated in a high temperature stress experiment that was not used by the expansion procedure (KS-test p-value$<10^{-3}$).

Figure S7: Hog1 plays a role in inhibition of expression that is not related to the mating/pseudohyphae pathways. Expression of three Hog1-dependent repressed modules (the known Ste12 module, and the novel modules Hog1B and Hog1/Ca) in *ste7* mutant defective in mating/pseudohyphae response (**A**, Madhani HD et al. 1999) and exposure to pheromone that induces the mating response (**B**, O'rourke and Herskowitz 2004, these profiles were not part of our original dataset). The expression of mating/pseudohyphae-dependent genes is expected to decrease in A and increase in B, compared to the expression distribution of all 5700 genes included in our analysis (red

curve). It can be clearly seen that only the known Ste12 module responds to the mating pathway, but the novel modules do not show response. Hence, although the Hog1-dependent repressed response is commonly attributed to the inhibition of the mating pathway, the novel repressed modules Hog1B and Hog1/Ca are probably regulated via another mechanism.

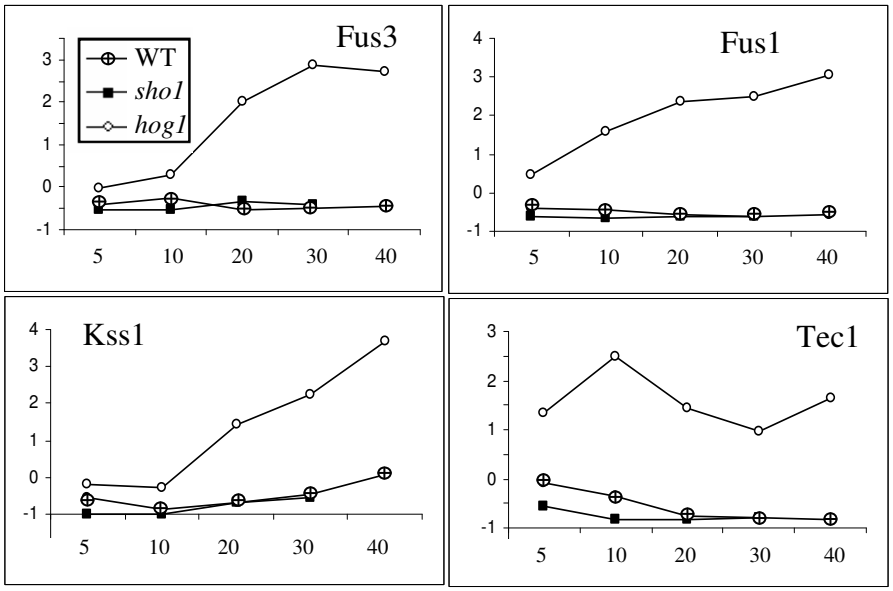
Figure S8: Response of predicted Hog1-dependent genes (A) and predicted Ste12-dependent genes (B) to pheromone. Shown are average and standard deviation of four expression profiles (in $\log_2$ scale) taken10-40 minutes after exposure to pheromone (O'Rourke and Herskowitz 2004). Although these experiments were not used in the expansion, they support our predictions: only the predicted Ste12-dependent genes are induced by pheromone that specifically activates the mating pathway.

Figure S9: A Model-independent module. The predicted regulators of the target genes in this module are Turgor pressure and Calcium stress, and thus the analysis suggests that this response is independent of the model. Indeed, one can clearly see that the target genes are affected by changes in osmotic concentration, but the genetic perturbations have only minor effect. This module is enriched with ribosomal proteins (hyper geometric p-value $\leq 10^{-10}$), and enriched in ChIP profiles of the TFs that are known to regulate ribosome biogenesis genes: RAP1 ($p \leq 10^{-17}$), FHL1($p \leq 10^{-26}$), SFP1($p \leq 10^{-13}$) from Harbison et al. (2004), and FKH1 ($p \leq 10^{-15}$) from Zhu et al (2000).

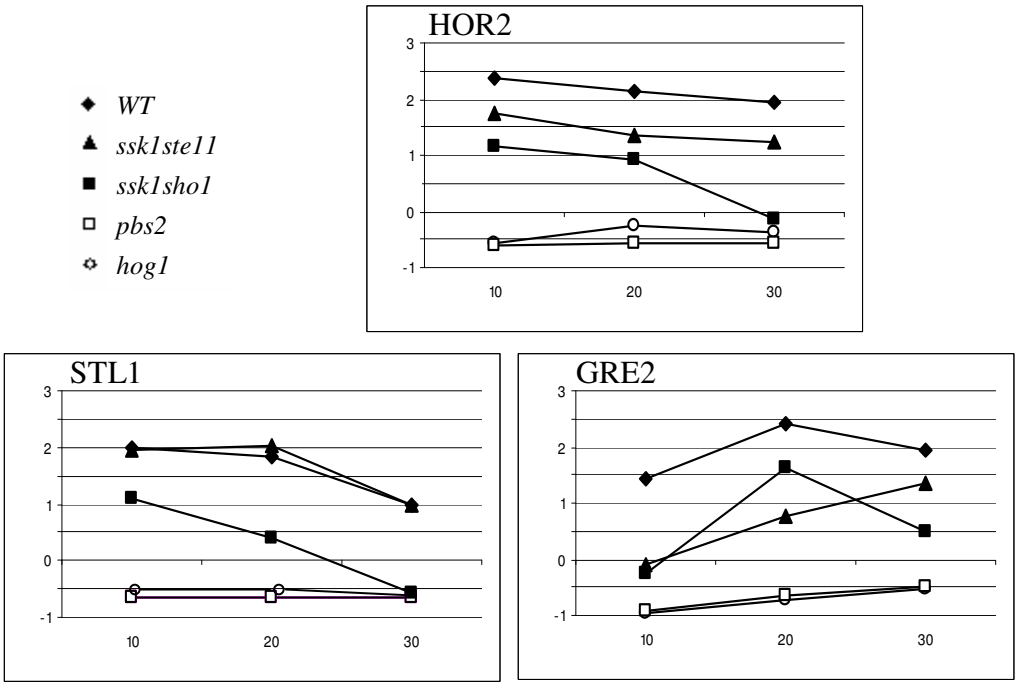
Table S1: Interactions predicted by the refinement procedure. The right hand column lists the experimental support in the literature for the predicted interaction.

Table S2: The quality of known targets assignment. For each transcription factor (column 1), we collected from the literature (column 2) a set of known targets (column 3). Columns 4-5: The known targets that were assigned to the model. These include genes that were correctly (column 4) or incorrectly (column 5) assigned to the model as detailed in Supplement C. Assignments to modules (either known or novel) are in bold. Assignments to very small gene sets that were filtered from our analysis are in parentheses. Only one gene was incorrectly assigned to a known module: HOR2 is in the Sko1 module instead of its correct Hot1/Msn1 module (Hohmann 2002). (While we cannot explain biologically this incorrect assignment, it is a direct consequence from our dataset: the expression profile of HOR2 shows high similarity to the predicted expression of the Sko1 module and even to the expression profile of GRE2, a known target of Sko1, and shares less similarity with the predicted expression of Hot1/Msn1 module and with STL1, a known target of Hot1/Msn1). All the other 12 incorrect genes were filtered from the analysis since they were assigned to very small gene sets. [a]Eleven Msn2/4 known targets were assigned to the Msn2/4 known module. Six additional genes were assigned to the Hog1A novel module, which is also hypothesized to be regulated by Msn2/4. [b]Sixteen genes were assigned to the Ste12 module. Five additional genes were assigned to various small (filtered) modules that are also predicted as Ste12 targets.

Sup Figure 1

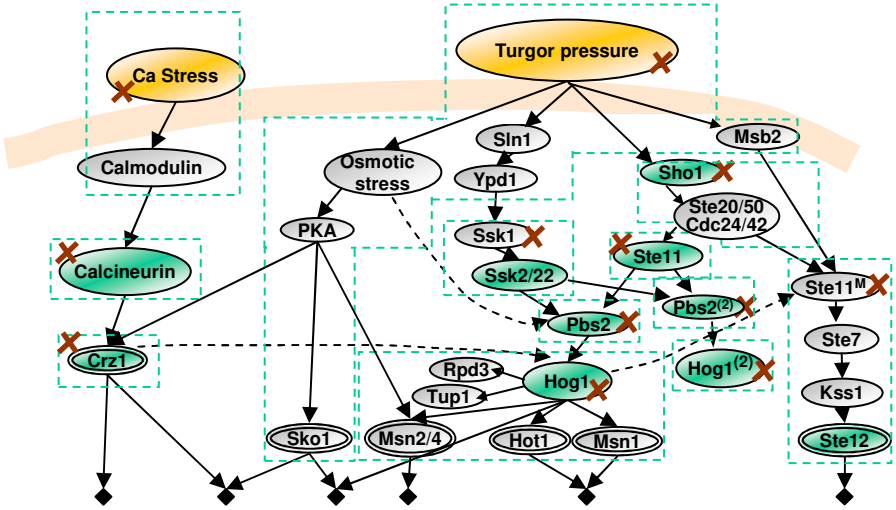


A

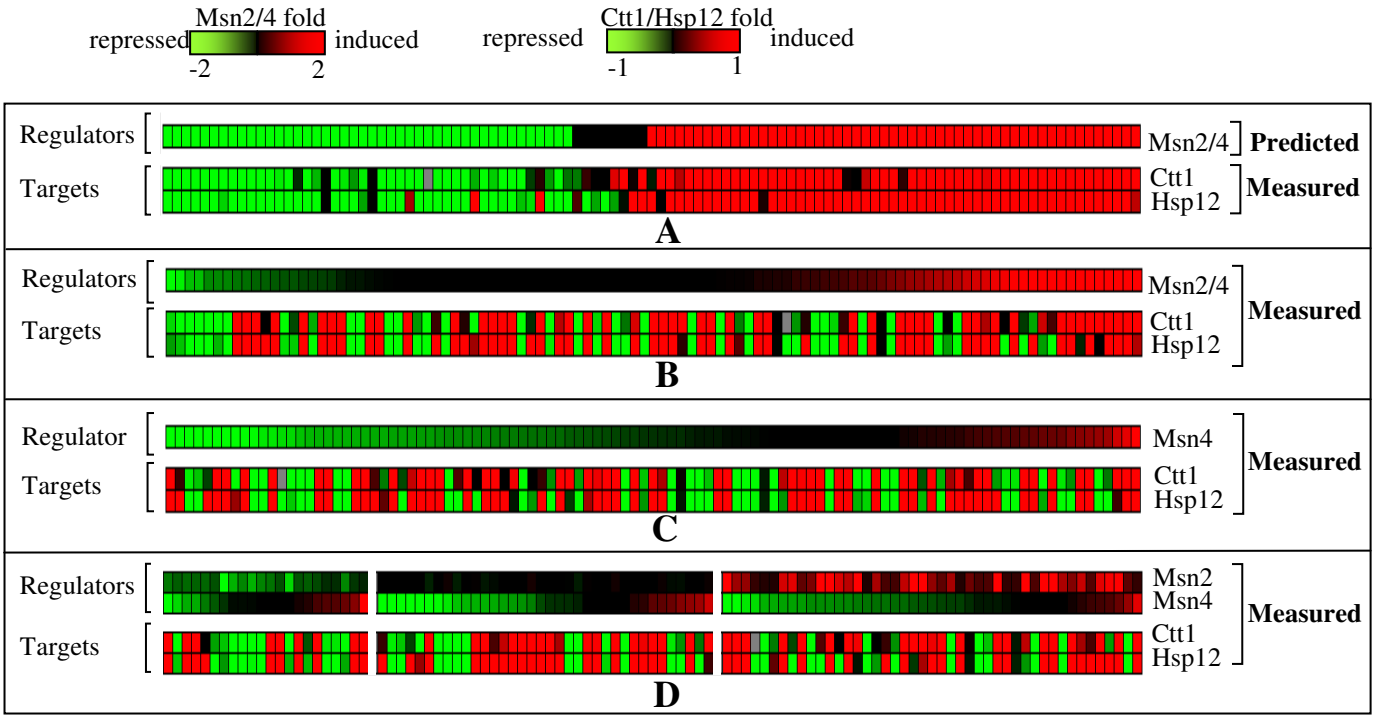


B

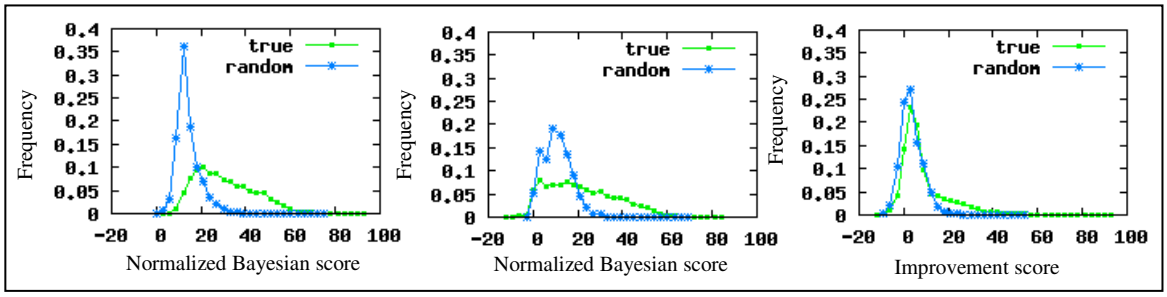
Sup Figure 2


Turgor pressure
Ca Stress
Sln1
Msb2
Calmodulin
Osmotic stress
Ypd1
Sho1
Ste20/50 Cdc24/42
PKA
Ssk1
Ste11
Calcineurin
Ssk2/22
Ste11M
Pbs2(2)
Pbs2
Crz1
Ste7
Rpd3
Hog1(2)
Hog1
Tup1
Kss1
Sko1
Msn2/4
Hot1
Msn1
Ste12

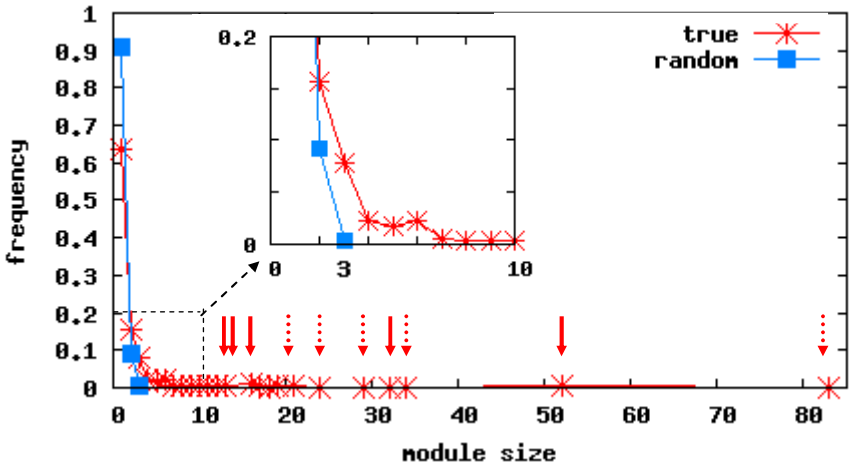
Sup Figure 3


Msn2/4 fold
repressed
induced
-2
2
Ctt1/Hsp12 fold
repressed
induced
-1
1
Regulators
Targets
Msn2/4
Predicted
Ctt1
Hsp12
Measured
A
Regulators
Targets
Msn2/4
Ctt1
Hsp12
Measured
B
Regulator
Targets
Msn4
Ctt1
Hsp12
Measured
C
Regulators
Targets
Msn2
Msn4
Ctt1
Hsp12
Measured
D

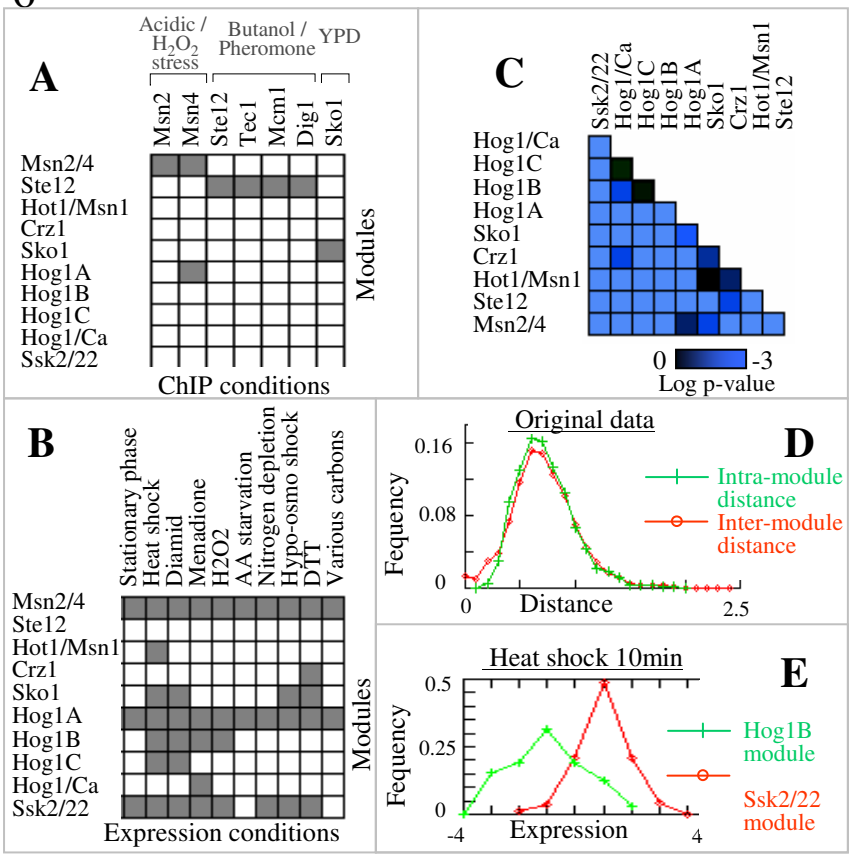
Sup Figure 4


Frequency
true
random
Normalized Bayesian score
Frequency
true
random
Normalized Bayesian score
Frequency
true
random
Improvement score

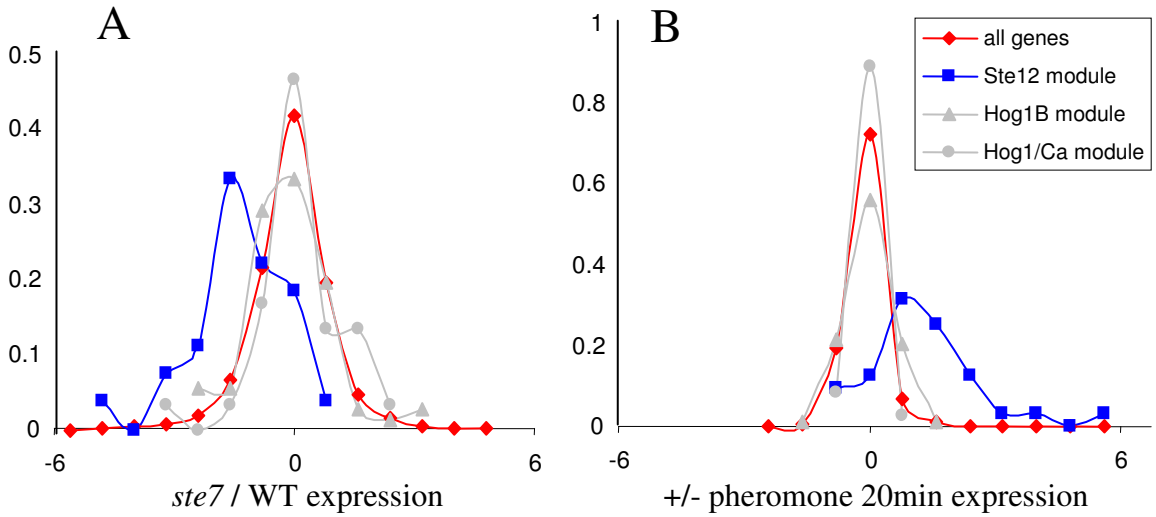
Sup Figure 5


frequency
module size
true
random
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
0
10
20
30
40
50
60
70
80
0.2
0
0
3
10

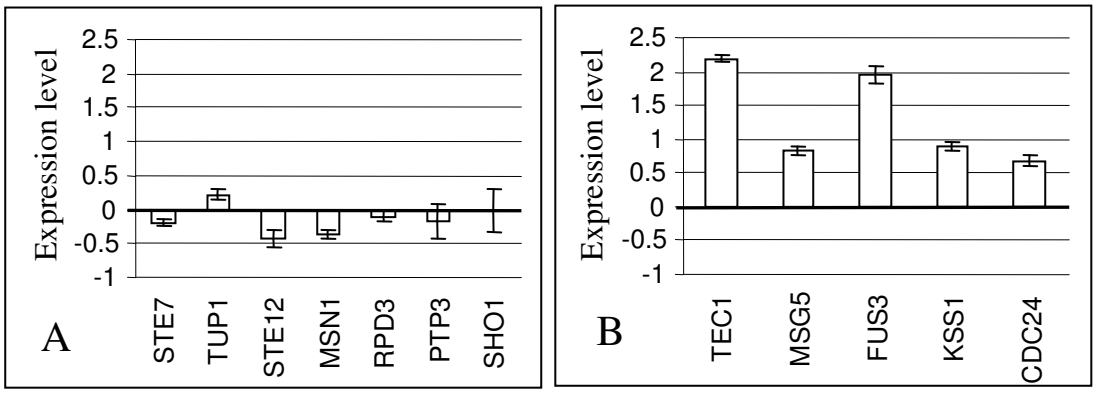
Sup Figure 6


A
Acidic / H_2O_2 stress
Butanol / Pheromone
YPD
Msn2
Msn4
Ste12
Tec1
Mcm1
Dig1
Sko1
Msn2/4
Ste12
Hot1/Msn1
Crz1
Sko1
Hog1A
Hog1B
Hog1C
Hog1/Ca
Ssk2/22
Modules
ChIP conditions
C
Ssk2/22
Hog1/Ca
Hog1C
Hog1B
Hog1A
Sko1
Crz1
Hot1/Msn1
Ste12
Msn2/4
0
-3
Log p-value
B
Stationary phase
Heat shock
Diamid
Menadione
H2O2
AA starvation
Nitrogen depletion
Hypo-osmo shock
DTT
Various carbons
Msn2/4
Ste12
Hot1/Msn1
Crz1
Sko1
Hog1A
Hog1B
Hog1C
Hog1/Ca
Ssk2/22
Modules
Expression conditions
Original data
D
0.16
0.08
0
Fequency
0
Distance
2.5
Intra-module distance
Inter-module distance
Heat shock 10min
E
0.5
0.25
0
Fequency
-4
Expression
4
Hog1B module
Ssk2/22 module

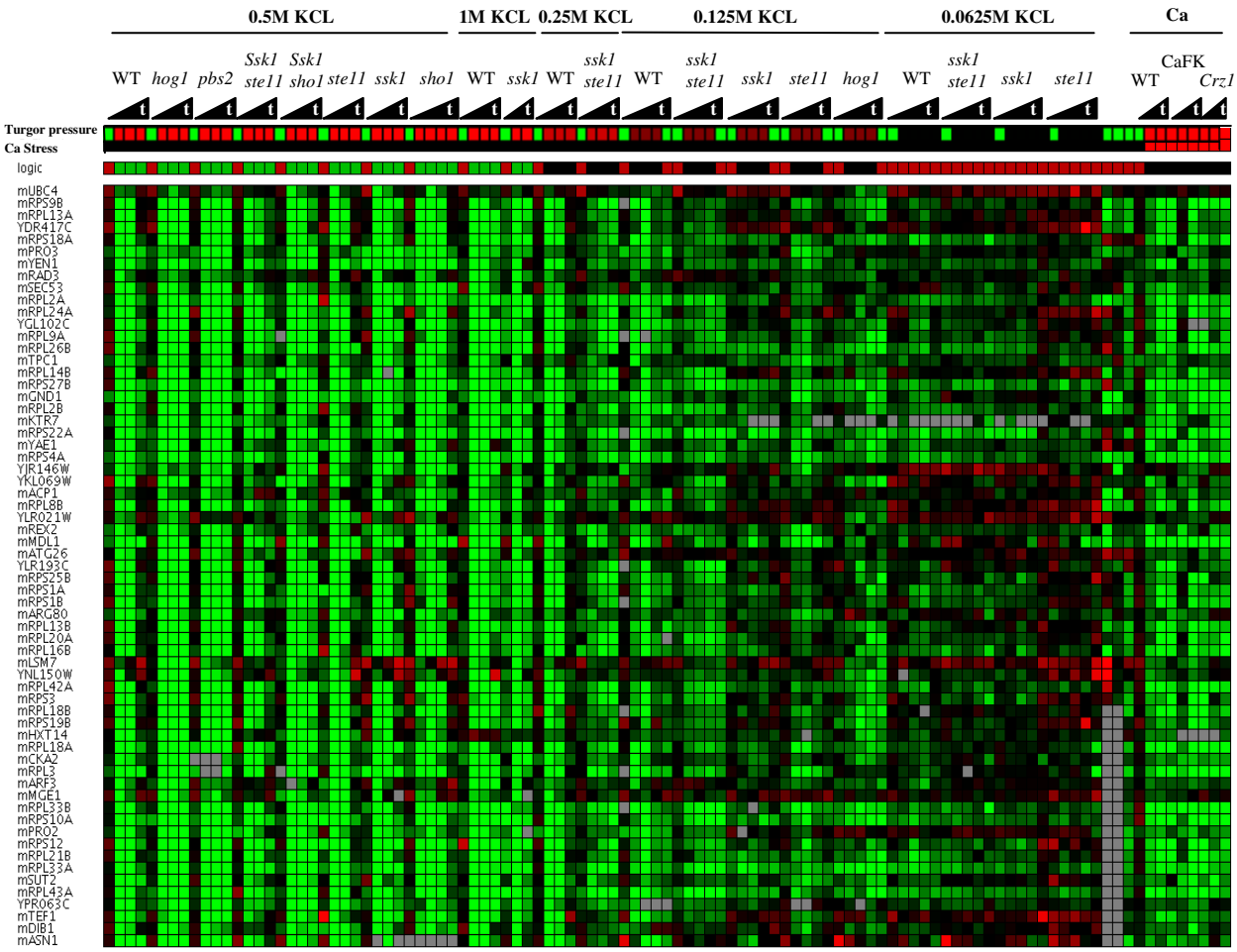
Sup Figure 7


A
0.5
0.4
0.3
0.2
0.1
0
-6
0
6
ste7 / WT expression
B
1
0.8
0.6
0.4
0.2
0
-6
0
6
+/- pheromone 20min expression
all genes
Ste12 module
Hog1B module
Hog1/Ca module

Sup Figure 8


Expression level
2.5
2
1.5
1
0.5
0
-0.5
-1
A
STE7
TUP1
STE12
MSN1
RPD3
PTP3
SHO1
Expression level
2.5
2
1.5
1
0.5
0
-0.5
-1
B
TEC1
MSG5
FUS3
KSS1
CDC24

Sup Figure 9


0.5M KCL
1M KCL
0.25M KCL
0.125M KCL
0.0625M KCL
Ca
WT
hog1
pbs2
Ssk1 ste11
Ssk1 sho1
ste11
ssk1
sho1
WT
ssk1
WT
ssk1 ste11
WT
ssk1 ste11
ssk1
ste11
hog1
WT
ssk1 ste11
ssk1
ste11
WT
CaFK
Crz1
t
Turgor pressure
Ca Stress
logic
mUBC4
mRPS9B
mRPL13A
YDR417C
mRPS18A
mPRO3
mYEN1
mRAD3
mSEC53
mRPL2A
mRPL24A
YGL102C
mRPL9A
mRPL26B
mTPC1
mRPL14B
mRPS27B
mGND1
mRPL2B
mKTR7
mRPS22A
mYAE1
mRPS4A
YJR146W
YKL069W
mACP1
mRPL8B
YLR021W
mREX2
mMDL1
mATG26
YLR193C
mRPS25B
mRPS1A
mRPS1B
mARG80
mRPL13B
mRPL20A
mRPL16B
mLSM7
YNL150W
mRPL42A
mRPS3
mRPL18B
mRPS19B
mHXT14
mRPL18A
mCKA2
mRPL2
mARF3
mMGE1
mRPL33B
mRPS10A
mPRO2
mRPS12
mRPL21B
mRPL33A
mSUT2
mRPL43A
YPR063C
mTEF1
mDIB1
mASN1

Sup Table 1

Predicted interaction	Improvement P-value	Support in the literature
Crz1 inhibits Hog1	10^{-3}	Calcium ions induce Hog1 hyperphosphorilation in *crz1* mutant (Shitamukai et al. 2004).
Third input of the HOG pathway	10^{-3}	Phosphorylation of Hog1 and Pbs2 under different osmotic stress conditions and mutants such as *sho1ssk2ssk22* (Van Wuytswinkel et al. 2000).
Hog1 inhibits the mating pathway	10^{-2}	Mutation in Hog1 and Pbs2 induces activation of pheromone response phenotype (morphological changes, mating, and induction of FUS1 reporter). This requires Sho1, Ste11, Ste7, Fus3/Kss1, and Ste12 (O'Rourke and Herskowitz 1998).

Sup Table 2

TF/Module	literature source	Known targets	Assigned to the model		**specificity** correctly assigned / total assigned	**sensitivity** correctly assigned / known
			Correctly assigned	assigned to small modules or incorrectly assigned		
Msn2/4	Rep et al. 2000	48	**11**+ **6**[a]	[6]	1	0.35
Ste12	Zeitlinger et al. 2003	67	**16**+[5][b]	[4]	1	0.24
Sko1	Rep et al. 2001	5	**2**	[2]	1	0.4
Crz1	Stathopoulos and Cyert 1997	3	0	0	-	0
Hot1+Msn1	Rep et al. 1999	3	**1**	**1**	0.5	0.33
		126	**36** [5]	**1** [12]	0.97	0.29