SUPPLEMENTARY METHODS

Yeast Microarray Normalization and Quantification

Normalization of the raw spot intensities was performed using Linear Models for Microarray Data, LIMMA (Smyth & Speed, 2003). Three different array platforms were used in this study. Samples were imported in three batches according to their array platform. For each batch, background correction was performed using the LIMMA function *backgroundCorrect* with the method type *normexp* and an offset of 50. Print-tip loess normalization was performed using the LIMMA function *normalizeWithinArrays*. After print-tip correction, all batches were merged into a single *MAList* object, and treated in a platform independent manner. For each array, M and A values for duplicate probes were averaged. Individual mutants were hybridized against one of three wild type RNA samples (2 from BY4741 and 1 from S288C). Each mutant had a minimum of two biological replicates, although the replicates were often hybridized against different wild type samples. We handle this differently than (Reimand, Vaquerizas, Todd, Vilo, & Luscombe, 2010) by constructing a linear model which relates all mutants to one of the BY4741 samples using the LIMMA functions *modelMatrix* and *lmFit*. The coefficients returned by *lmFit*, which correspond to the log-ratio of each gene's expression in each mutant relative to the BY4741 wild type, were saved as the table of gene expression used by all subsequent network inference analyses. Differential expression of each deletion strain was assessed relative to the wild type strain (BY4741) using an empirical Bayesian moderated t-test (Smyth, 2004) implemented with the LIMMA function *eBayes*,

which yielded p-values (corrected for multiple testing) and log-odds statistics for differential expression.

Characterizing Direct and Functional Regulation in *S. cerevisiae*

To identify regulated genes in *S. cerevisiae*, microarray data from studies in which transcription factors were either disrupted (Hu, Killion, & Iyer, 2007) or overexpressed (Chua et al., 2006) were normalized as described above. Differential expression was assessed for each TF mutant relative to wild type. A total of 6,258 genes (97.5% of all yeast genes) were significantly differentially expressed (corrected p-val < 0.05) in one or more TF mutants. Genes that were both differentially expressed in response to a TF perturbation and bound by the TF in curated ChIP studies were considered direct and functional targets of the TF. 6,078 direct and functional interactions were identified between 152 TFs and 2,850 genes (45% of all yeast genes).

Yeast Network Inference

TF network inference was performed on the yeast microarray data set described above using three algorithms: Inferelator, GENIE3, and NetProphet. We used the same Inferelator pipeline (CLR + Inferelator +MCZ) that was applied to the DREAM4 data to infer the yeast transcriptional network. This pipeline was modified to restrict the set of allowed regulators to be transcription factors. Because the data do not include time courses, regular CLR was used instead of mixed CLR. For GENIE3 and Inferelator, the table of expression data for all

mutants was moved out of $\log_2$-space by exponentiation. Each gene's expression was then normalized (divided by its maximum over all measurements). This treatment conforms to the DREAM4 data standards on which Inferelator and GENIE3 were originally tested. For NetProphet, the values were left as log ratios, which were scaled in such a way that genes with low variance would be given less priority but genes with exceptionally high variance would not dominate the LASSO solution. The maximum scaling factor λ was computed from the distribution of gene-wise standard deviations in the unnormalized data:

$$\lambda = \sqrt{N}\left(\sigma(\{\sigma(Y_j) : j \in [1,N]\}) + \mu(\{\sigma(Y_j) : j \in [1..N]\})\right)^{-1}$$

where Y_j is the set of $\log_2$ fold changes for gene j over all measurements, N is the number of measurements; σ(X) and μ(X) are functions that map a set of real values to the population estimated standard deviation and mean respectively. Each gene j's log ratios were scaled (multiplied) by either λ or $\sqrt{N}\,\sigma(Y_j)^{-1}$, whichever was smaller.

To learn evidence weights for NetProphet, we estimated optimal region weights using five rounds of two-fold cross validation, partitioning the training data (interactions labeled as true or false by ChIP evidence) by regulator, and an additional five rounds of two-fold cross validation partitioning the training data by target. The region weights **ω** were allowed to take on the following values: [1e-3, 1e-2, 1e-1, 1, 2, 3]. The offset coefficients c_b and c_p were allowed to take on the

values: [1e-2, 1e-1, 1]. Weights and offsets were selected so as to maximize the average AU-PRC (area under the precision recall curve) over all rounds of cross validation. The weights selected by cross-validation were: $\boldsymbol{\omega}_{I} = 3$, $\boldsymbol{\omega}_{II} = 1$, $\boldsymbol{\omega}_{III} = 1$, $\boldsymbol{\omega}_{IV} = 1$, $\boldsymbol{\omega}_{B} = 2$, $\boldsymbol{\omega}_{D} = 2$; and the offset coefficients were $c_b = 0.1$ and $c_d = 0.01$.

Assessing gene expression with Quantigene

After cell lysis, gene expression was assessed using the QuantiGene 2.0 Plex Assay kit. Samples were prepared according to kit manufacturer's instructions. In brief, using an Affymetrix hybridization plate, 60µl of working bead mix and 40µl of lysed sample were added. The plate was incubated in a VorTemp 56 shaking incubator (Labnet International) at 54°C at 600rpm overnight. Afterwards, samples were transferred to a Magnetic Separation Plate (Affymetrix). The plate was placed on a Hand-Held Magnetic Plate Washer (Affymetrix), the beads were allowed to settle, and the working bead mix was removed. The plate was washed with Wash Buffer (Affymetrix) three times. Pre-Amplifier (Affymetrix) was added and allowed to bind to the sample. The plate was incubated at 50°C for 1hr and washed as before. The Label Probe (Affymetrix) was added to the beads, incubated, and washed as before. Then SAPE Working Reagent (Affymetrix) was added and samples were incubated at room temperature for 30 minutes. The plate was washed with SAPE Wash Buffer three times and then the beads were re-suspended in 130µl SAPE Wash Buffer. Finally, the plate was read on a Luminex instrument with the following specifications: Sample size 100µl, DD Gates 5000-25000, Bead Event/Bead

Region 100.

Analysis of DREAM4 expression data

Inference of the DREAM4 networks was performed using our method, NetProphet, and publicly available versions of Inferelator (http://err.bio.nyu.edu/inferelator/) and GENIE3 (http://www.montefiore.ulg.ac.be/~huynh-thu/software.html). Datasets for the DREAM4 in-silico 100 networks were obtained from http://wiki.c2b2.columbia.edu/dream/index.php/D4c2. These datasets covered 5 networks each containing 100 genes. The individual datasets for each network contained single measurements for wild type, all single knockouts, all single knockdowns (50% expression) and 10 time courses (each with 21 time points), for a total of 411 measurements. All expression data is provided in a normalized format such that each gene's expression lies on the interval [0,1]. We applied Inferelator to this dataset as described, using the CLR + Inferelator + MCZ pipeline . Similarly for GENIE3 we inferred the network structure using the GENIE3 functions: *read.expr.matrix* and *get.weight.matrix* and their default parameterizations.

To properly handle the time course data, a spline was fit to the expression values for each gene. For each time point the derivative of the spline was used to estimate the rate of change for each gene, and the transcription rate of each gene was estimated by adding the gene's concentration to the rate of change at

each time point (assuming a unit degradation rate constant). These estimated transcription rates were used instead of the expression measurements as the response matrix, *Y*, for LASSO regression. Note that a mixture of steady state and time course measurements in the response matrix is compatible under this formulation, because steady state concentrations are equal to transcription rates assuming a unit degradation rate constant. In addition to modifying the response variable to allow for a mixing of steady state and time course data, the covariate matrix, *X*, is also modified. Gene expression measurements for time course measurements in the covariate matrix are replaced with protein concentration estimates for each time point (which are effectively lagged expression measurements). We estimated a gene's protein concentrations using the spline fit to mRNA measurements, and integrated an ODE which defines a protein's rate of change as a function of the mRNA concentration, minus the protein concentration times a degradation rate constant. The degradation rate constant which we set to 0.01 for all genes defines the lag between the mRNA and protein species to be roughly one time point. Finally, a $\log_2$ transformation was performed on both the response matrix *Y* and covariate matrix *X* before applying LASSO regression.

Differential expression analysis of the knockout measurements was used to compute the DE rank scores D_{ij}. LIMMA was used to compute these scores by comparing each knockout (which consisted of a single measurement) to 11 measurements of wild type (one for each time 0 point of the 10 time courses,

which was a steady state wild type measurement, and one which was provided separate of the time courses).

Investigation of the effects of having smaller datasets

To investigate the effects of data set size, we carried out a number of experiments in which we ran NetProphet 4 times, each using a different set of TF deletion profiles (Hu, et al., 2007), and extracted the NetProphet scores of 25% of the TFs from each run. These scores were then combined, predictions were ranked according to score, and evaluation was carried out as before using the curated PWMs from the UNIPROBE database (Gordan et al., 2011; Robasky & Bulyk, 2011). To investigate the effects of data set size on targets of TFs for which no deletion profile was available, we ran NetProphet on subsamples consisting of 25% (66-67), 50% (135), or 75% (202-203) of the TF deletion profiles. For each data set size, NetProphet was run 4 times and scores for 25% of TFs *whose deletion profiles were not included in the run* were extracted, combined, and evaluated. To investigate the effects of data set size on targets of TFs for which a deletion profile *was* available, we carried out the same procedure but extracted the scores of 25% of TFs *whose deletion profiles were included in the run*.

REFERENCES

Chua, G, Morris, QD, Sopko, R, Robinson, MD, Ryan, O, Chan, ET, Frey, BJ, Andrews, BJ, Boone, C, & Hughes, TR. (2006). Identifying transcription factor functions and targets by phenotypic activation. *Proc Natl Acad Sci U S A,* ***103***(32), 12045-12050. doi: 10.1073/pnas.0605140103

Gordan, R, Murphy, K, McCord, RP, Zhu, C, Vedenko, A, & Bulyk, ML. (2011). Curated collection of yeast transcription factor DNA binding specificity data reveals novel structural and gene regulatory insights. *Genome Biology,* ***12***(12), R125. doi: 10.1186/gb-2011-12-12-r125

Hu, Z, Killion, PJ, & Iyer, VR. (2007). Genetic reconstruction of a functional transcriptional regulatory network. *Nat Genet,* ***39***(5), 683-687.

Reimand, J, Vaquerizas, JM, Todd, AE, Vilo, J, & Luscombe, NM. (2010). Comprehensive reanalysis of transcription factor knockout expression data in Saccharomyces cerevisiae reveals many new targets. *Nucleic Acids Res,* ***38***(14), 4768-4777.

Robasky, K, & Bulyk, ML. (2011). UniPROBE, update 2011: expanded content and search tools in the online database of protein-binding microarray data on protein-DNA interactions. *Nucleic Acids Research,* ***39***, D124-D128. doi: Doi 10.1093/Nar/Gkq992

Smyth, GK. (2004). Linear models and empirical bayes methods for assessing differential expression in microarray experiments. *Statistical applications in genetics and molecular biology,* ***3***, Article3. doi: 10.2202/1544-6115.1027

Smyth, GK, & Speed, T. (2003). Normalization of cDNA microarray data. *Methods,* ***31***(4), 265-273.