

### ***Microarray data***

As microarray benchmark sets, the Spellman [1] and Gasch [1] datasets were used. The Spellman dataset [1] contains 77 experiments describing the dynamic changes of yeast genes during the cell cycle. The Gasch dataset [2] consists of 177 experiments, examining gene expression behavior during various stress conditions. Expression profiles were normalized (subtracting the mean of each profile and dividing by the standard deviation across the time points) and stored in a gene expression data matrix, denoted by  $A$ , with a row for each gene expression profile and a column for each condition.

### ***Location data***

Genome-wide location data performed by Harbison et al. [3] were downloaded from [http://web.wi.mit.edu/young/regulatory\\_code](http://web.wi.mit.edu/young/regulatory_code). These contain information regarding the binding of 204 regulators (although Harbison et al. [3] only describe 203 regulators) to their respective target genes in rich medium (the 106 regulators initially profiled by Lee et al. [4] and 98 new regulators). Besides rich medium, 84 regulators were profiled in at least one environmental condition other than rich medium. The ChIP-chip data matrix (denoted by  $R$ ) used in our study consists of one minus the *ChIP-chip p-values* for each gene, obtained from the combined ratio's of immunoprecipitated and control DNA using an error model (see Harbison et al. [3])

### ***Motif data***

The motif data used in this study were obtained from a comparative genome analysis between distinct yeast species (phylogenetic shadowing) performed by Kellis et al. [5]. These motifs, available online as regular expressions, were transformed into their corresponding weight matrices by selecting for each motif, the 20 *Saccharomyces cerevisiae* genes in which the motif was most reliably detected according to the scoring scheme of Kellis et al. [5]. Subsequently, the intergenic sequences of these genes were subjected to motif detection based on Gibbs sampling. If the statistically overrepresented motif in this set of putatively co-expressed genes corresponded to the motif that was detected by the comparative motif search of Kellis et al. [5] the motif model was retained. As such, out of the 71 putative motifs described by Kellis et al. [5] 53 were retained. The weight matrices corresponding to these motifs was subsequently used to screen all intergenic sequences of yeast using MotifLocator. The higher the score of a motif hit in a gene, the more likely it will be a true instance. The score distribution of the motif hits depends on the motif length and the degree of conservation of the motif i.e. in general motif hits for longer motifs with higher degree of conservation will score better than motif hits for shorter motifs with smaller degree of conservation. Because this distribution of scores

differs between motifs, scores were converted into percentile values. As such, all motif hits scores are on the same scale. The matrix containing these percentile values is the motif data matrix M that will be used in this work.

#### Reference List

1. Spellman PT, Sherlock G, Zhang MQ, Iyer VR, Anders K, Eisen MB, Brown PO, Botstein D, Futcher B: **Comprehensive identification of cell cycle-regulated genes of the yeast *Saccharomyces cerevisiae* by microarray hybridization.** *Mol Biol Cell* 1998, **9**:3273-3297.
2. Gasch AP, Spellman PT, Kao CM, Carmel-Harel O, Eisen MB, Storz G, Botstein D, Brown PO: **Genomic expression programs in the response of yeast cells to environmental changes.** *Mol Biol Cell* 2000, **11**:4241-4257.
3. Harbison CT, Gordon DB, Lee TI, Rinaldi NJ, Macisaac KD, Danford TW, Hannett NM, Tagne JB, Reynolds DB, Yoo J et al.: **Transcriptional regulatory code of a eukaryotic genome.** *Nature* 2004, **431**:99-104.
4. Lee TI, Rinaldi NJ, Robert F, Odom DT, Bar-Joseph Z, Gerber GK, Hannett NM, Harbison CT, Thompson CM, Simon I et al.: **Transcriptional regulatory networks in *Saccharomyces cerevisiae*.** *Science* 2002, **298**:799-804.
5. Kellis M, Patterson N, Endrizzi M, Birren B, Lander ES: **Sequencing and comparison of yeast species to identify genes and regulatory elements.** *Nature* 2003 May 15 ;423 (6937 ):241 -54 2003, **423**:241-254.