

In search of functional association from time-series microarray data based on the change trend and level of gene expression

Feng He and An-Ping Zeng*

Address: Research Group Systems Biology, GBF – German Research Center for Biotechnology, Mascheroder Weg 1, 38124 Braunschweig, Germany

Email: Feng He - fhe@gbf.de; An-Ping Zeng* - aze@gbf.de

* Corresponding author

Published: 15 February 2006

Received: 24 June 2005

BMC Bioinformatics 2006, 7:69 doi:10.1186/1471-2105-7-69

Accepted: 15 February 2006

This article is available from: <http://www.biomedcentral.com/1471-2105/7/69>

© 2006 He and Zeng; licensee BioMed Central Ltd.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Abstract

Background: The increasing availability of time-series expression data opens up new possibilities to study functional linkages of genes. Present methods used to infer functional linkages between genes from expression data are mainly based on a point-to-point comparison. Change trends between consecutive time points in time-series data have been so far not well explored.

Results: In this work we present a new method based on extracting main features of the change trend and level of gene expression between consecutive time points. The method, termed as trend correlation (TC), includes two major steps: 1, calculating a maximal local alignment of change trend score by dynamic programming and a change trend correlation coefficient between the maximal matched change levels of each gene pair; 2, inferring relationships of gene pairs based on two statistical extraction procedures. The new method considers time shifts and inverted relationships in a similar way as the local clustering (LC) method but the latter is merely based on a point-to-point comparison. The TC method is demonstrated with data from yeast cell cycle and compared with the LC method and the widely used Pearson correlation coefficient (PCC) based clustering method. The biological significance of the gene pairs is examined with several large-scale yeast databases. Although the TC method predicts an overall lower number of gene pairs than the other two methods at a same p-value threshold, the additional number of gene pairs inferred by the TC method is considerable: e.g. 20.5% compared with the LC method and 49.6% with the PCC method for a p-value threshold of $2.7E-3$. Moreover, the percentage of the inferred gene pairs consistent with databases by our method is generally higher than the LC method and similar to the PCC method. A significant number of the gene pairs only inferred by the TC method are process-identity or function-similarity pairs or have well-documented biological interactions, including 443 known protein interactions and some known cell cycle related regulatory interactions. It should be emphasized that the overlapping of gene pairs detected by the three methods is normally not very high, indicating a necessity of combining the different methods in search of functional association of genes from time-series data. For a p-value threshold of $1E-5$ the percentage of process-identity and function-similarity gene pairs among the shared part of the three methods reaches 60.2% and 55.6% respectively, building a good basis for further experimental and functional study. Furthermore, the combined use of methods is important to infer more complete regulatory circuits and network as exemplified in this study.

Conclusion: The TC method can significantly augment the current major methods to infer functional linkages and biological network and is well suitable for exploring temporal relationships of gene expression in time-series data.

Background

Gene expression profiling has gained a tremendous importance in functional genomic research. The presently often used approach is to compare gene expression in discrete time points resulting for example of different genotypes or cell lines, morbid and healthy (control) objects or under different physiological conditions. This type of static gene expression profiling can already give useful information on the patterns of significantly differentiated expression of genes. However, in order to achieve a more complete picture of significantly differentiated gene expression, especially in order to capture and understand the dynamics of the altered gene expression, it is desirable to measure time-series gene expression [1].

Large efforts have been made in recent years to develop bioinformatics methods to study gene expression patterns [2-5] and/or to infer functional linkages and even regulatory network from microarray data [6-11]. For inference of functional associations among genes two major classes of methods are presently in use. One class of the methods is based on graphical modeling, which include Bayesian network [7] and Gaussian graphical model [12]. Recently, dynamic Bayesian networks have been proposed to model temporal gene expression and represent a promising direction. However, most of the current work in this area is limited to the analysis of a relatively small set of genes due to computational complexity [9-11]. Another class of methods infers functional association from large-scale gene expression data by defining a statistic threshold for the association. The main measure used for defining association is the Pearson correlation coefficient (PCC) [6,13,14]. PCC is widely used for detecting co-expressed genes from both static and time-series expression data. However, several important issues are not specifically addressed in PCC based methods when applied to time-series expression data. A major issue is that the PCC clustering method treats its input as a vector of independent samples and hence doesn't take into account the temporal relationship between consecutive time points. In addition, time-shifted and/or inverted expression of certain gene pairs are not considered. Time-shifted and inverted relationships are important features of gene expression regulation [8]. For example, a gene may activate or inhibit another gene or even several related genes downstream in a regulatory pathway, resulting in time-delayed positive or negative response in the transcription of the downstream gene(s). To consider these phenomena, Qian et al. [8] proposed a local clustering (LC) method. As demonstrated with the expression profiling data of yeast cell cycle this method can identify new, biologically relevant interactions that could not be found by the conventional PCC clustering method. However, the method of Qian et al. [8] is principally still based on a point-to-point comparison or local clustering of expression levels of genes although it

explicitly considers time-shifted and inverted gene expression profile. Kwon et al. [15] proposed an 'event-based' edge detection method to consider the change trend of gene expression between consecutive time points. By simplifying a profile of time series into a sequence of decrease or increase events this method is more robust to noises. However, it does not fully make use of the information contained in the gene expression levels in the original data. Filkov et al. [16] proposed a similar method called 'edge detection'. Recently Balasubramaniyan et al. [17] proposed a method to use Spearman rank correlation based on the rank of expressional values. The rank of expressional values is more insensitive to noises or outliers but this method is still based on the point-to-point comparison *per se*.

To more comprehensively consider the temporal relationships of gene expression in time-series microarray data we propose here a new method that is based on extracting the main features of the change trend and the change level of gene expression between consecutive time points. We not only consider the qualitative information (i.e. the change trend) but also the quantitative information (the change level) in the original data. We seek to make the methods more noise-tolerant and at the same time to keep more useful information in the expression values. This new method, termed here as trend correlation (TC), is demonstrated with the microarray data from cell cycle of yeast [18]. The biological significance of functional associations of inferred gene pairs is examined with several large-scale yeast databases. We also extensively compare our method with the LC method and the PCC based clustering method of Eisen et al. [6]. It is shown that a significant number of functionally associated gene pairs, which have well-documented biological interactions and relationships but cannot be significantly detected by the LC and PCC methods, can be inferred by the new method with high statistic significance. The biological significance of the functional association pairs inferred by our method is generally higher than that of the LC method and similar to that of the PCC clustering method which detects however only simultaneous co-expression gene pairs. Furthermore, it is shown that the overlapping of gene pairs detected by the three methods is normally not very high, indicating a necessity of combining the different methods in search of functional association of genes from time-series microarray data.

Principle and scheme of the proposed method

The principle of our method is to use information in the change trend and the change level of gene expression between consecutive time points for the inference of functional linkages among genes. To consider not only the positive correlation but also time-shifted and/or inverted expression of certain gene pairs we employ a similar algo-

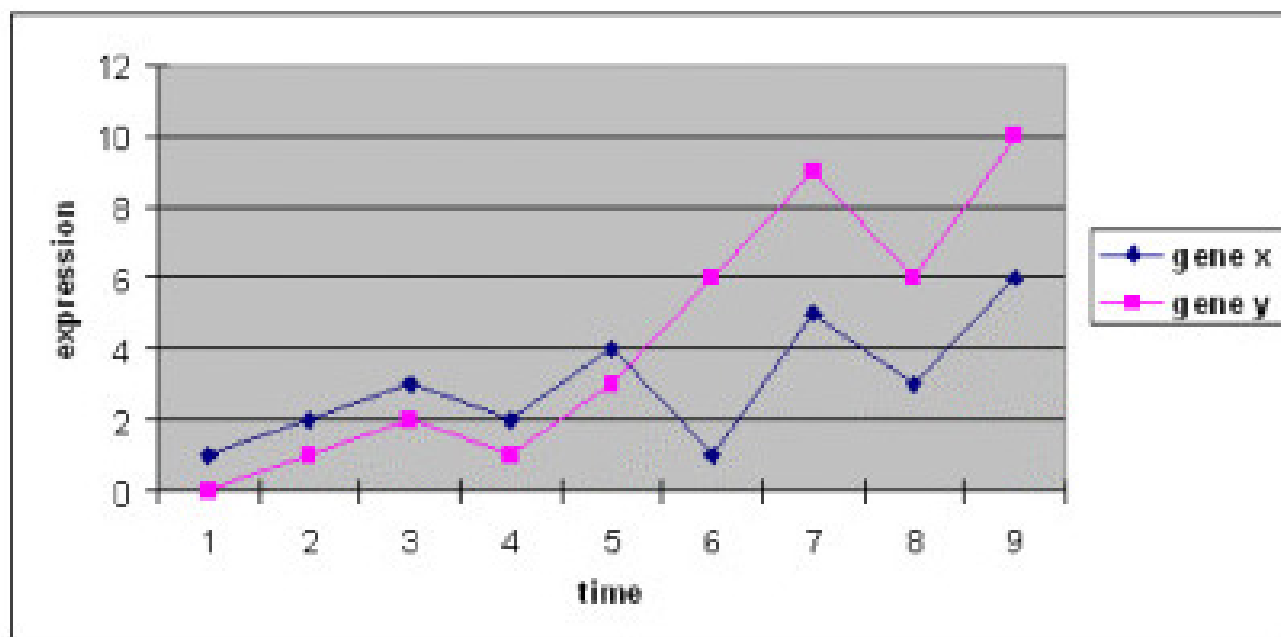


Figure 1

Illustration of calculating the change trend score sc and the correlation coefficient cc of the match change trend. The two genes X and Y in this example have a positive relationship (here expression levels are not normalized). Only the change trend between time points 5 and 6 is different. The number of matched change after time shifts (by dynamic programming) is less than 7. Thus, the maximal match change trend score sc is 7 in this example with 9 time points. To calculate the correlation coefficient, we first extract the change levels of the matched change trends:

$Xmc_1^2 = 1, Xmc_2^3 = 1, Xmc_3^4 = -1, Xmc_4^5 = 2, Xmc_5^6 = 4, Xmc_6^7 = -2, Xmc_7^8 = 3$ The change level between time points 5 and 6 is excluded
 $Ymc_1^2 = 1, Ymc_2^3 = 1, Ymc_3^4 = -1, Ymc_4^5 = 2, Ymc_5^6 = 3, Ymc_6^7 = -3, Ymc_7^8 = 4$

because the change (inverted) trend is different from the main (positive) change trend. Using the above matched change levels we can obtain a correlation coefficient $cc = 0.9597$ according to Eq. 1.

algorithm as the local sequence alignment [19] to calculate a maximal local alignment of change trend (sc) between each gene pair (Fig. 1). We noticed that a significant number of gene pairs that have a same sc value may in fact have a large difference in the change levels of their expression. The distinction of their change levels may lead to the different degrees of similarity of functional association. To solve this problem we make use of the quantitative information in the change levels between consecutive time points and calculate a correlation coefficient (cc) between the maximal alignment.

The method consists of following major steps:

1. Generating a random dataset by shuffling the normalized expression levels at different time points among each gene expression profile in the original dataset;

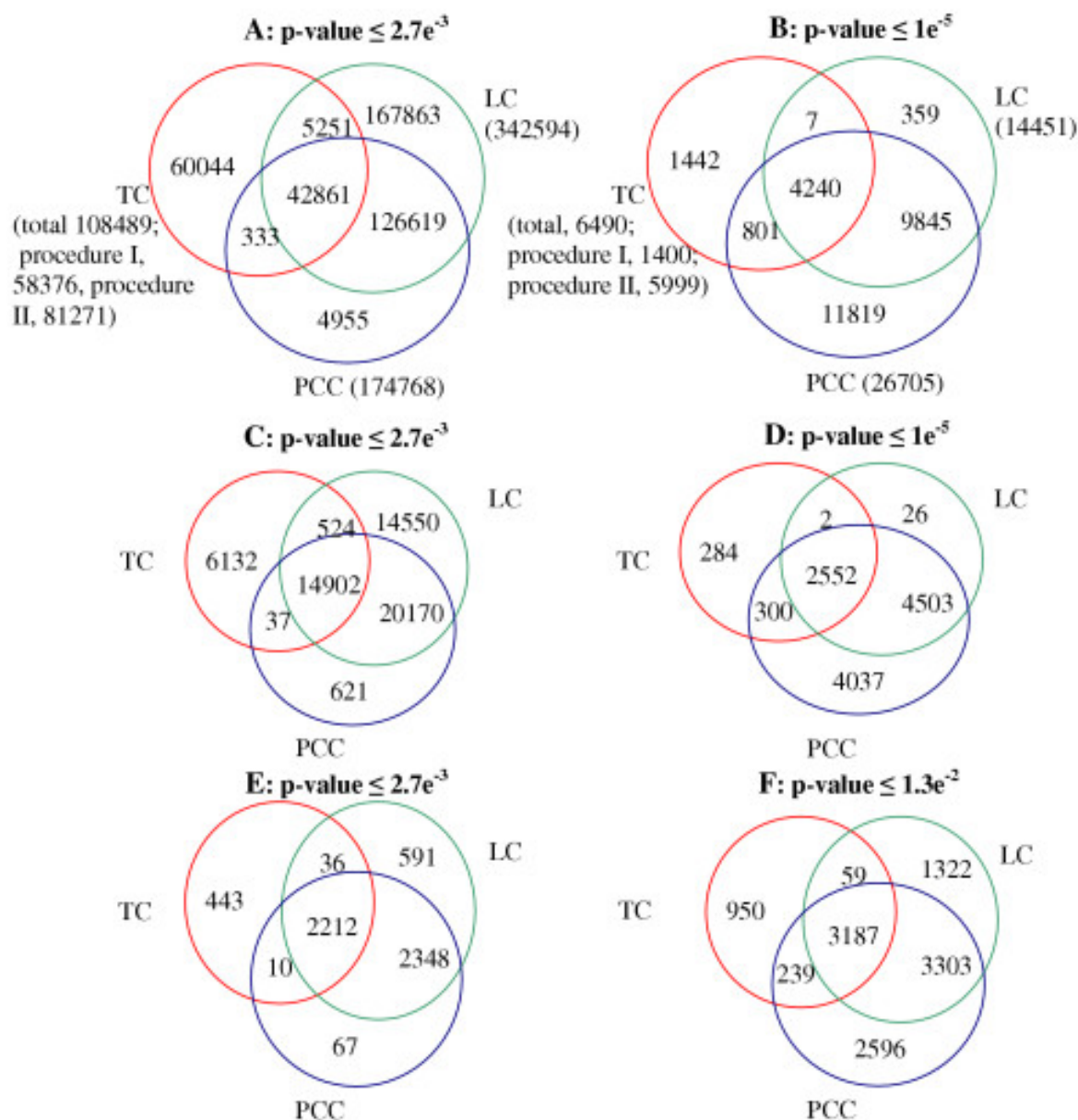
2. Calculating a maximal local alignment of change trend (sc) between each gene pair in the random dataset as illustrated in Fig. 1 for a simple case;

3. Calculating a correlation coefficient (cc) between the maximal alignment for each gene pair in the random dataset (Fig. 1);

4. Tabulating the frequency of sc (i.e. $f(sc)$) as function of sc as shown in Fig. S1A [see Additional file 1]; followed by tabulating the distribution of cc for gene pairs which have the same sc ;

5. Calculating the conventional p-values for the two scores sc and cc ($P_{sc}(s \geq sc), P_{cc}(c \geq cc)$) through integration of the frequency distributions (Fig. S1B);

6. Calculating sc and cc between each gene pair in the original dataset;

**Figure 2**

A and B, number of gene pairs detected by the TC method in comparison to those of the LC method and the PCC method inferred from the yeast cell cycle dataset of Cho et al. [18] at two different p-value thresholds. The numbers in parentheses are the whole number of gene pairs detected by the corresponding method. The percentage of the additional number of gene pairs inferred only by the TC method in the text is calculated in a way exemplified in the following by a comparison with those resulted merely from the PCC method at the p-value threshold of $2.7e^{-3}$: $\frac{60044 + 5251}{4955 + 126619} = 49.6\%$. C and D, number of

process-identity pairs among the detected pairs from the three methods; E and F, number of known protein interactions (including protein complexes) in collection dataset of protein interactions [21] among the inferred gene pairs from the three methods.

7. Extraction of functional linkages using Procedure I proposed: extract gene pairs with significantly high *sc* values with a certain preset *p*-value. The correlation coefficient *cc* is regarded as a second index when the gene pairs have the same score *sc*;

8. Extraction of functional linkages using Procedure II: extract gene pairs with statistically significantly high value of combined scores of *sc* and *cc*.

The key steps mentioned above are described in the section of Methods in more detail. The main programs TC_linkage_infer created in this work can be freely downloaded from our website <http://www.gbf.de/SystemsBiology>.

Results

Dataset of yeast cell-cycle

We tested our algorithm with the time series microarray data (17 time points) generated by Cho et al. [18] for yeast cell cycle with a whole genome yeast oligonucleotide chip which included over 6000 ORFs. After removing all the negative expression levels in the scaled measurements and all the dubious and genes now deleted in the SGD database [20], 5680 genes were included in our calculation. We examined all the possible pairs among them. The values of the two scores *sc* and *cc* and the type of possible relationship (simultaneous, time-shift or inverted) were calculated and assigned for each gene pair.

Functional associations inferred by different methods

Figs. 2A and 2B show the number of gene pairs with possible functional association inferred from the yeast time-series data by the TC method (including separate results of the two extraction procedures), the PCC clustering method and the LC method at two different cut-off *p*-values. The number of gene pairs extracted by Procedure II of our method is generally higher than that from Procedure I, especially at the lower *p*-value. Note that there is some overlapping between these two extraction procedures so that the total number of the gene pairs inferred is lower than the sum of the two procedures. For comparing the TC method with the other two methods the combined results of the two extraction procedures are considered in the following if not otherwise mentioned. *P*-values of $2.7E-3$ and $1E-5$ in the TC method are equivalent to Pearson correlated coefficients of 0.76 and 0.89 in the PCC clustering method and scores of 13 [8] and 15.6 in the local clustering method respectively.

As can be clearly seen in Figs. 2A and 2B, the number of inferred gene pairs depends much on the *p*-value in all the three methods. In general, the TC method infers a significantly lower number of gene pairs compared to the other two methods. 39.8% and 44.3% of the pairs inferred by

the TC method are also found by the PCC and LC methods respectively. The common part of the three methods is somewhat lower and accounts only 12.5% for the LC method, 32.5% for the PCC method and 39.5% for the TC method. The number of gene pairs merely inferred by the TC method is considerable compared with those merely from the LC method (20.5%) and the PCC clustering method (49.6%). By decreasing the *p*-value threshold to $1E-5$, the number of inferred functional pairs decreases remarkably, especially for the TC and LC methods (Fig. 2B). However, the shared part of gene pairs predicted by the TC method increases up to 65.4% and 77.7% compared to the LC and PCC methods respectively, indicating an increased reliability of the prediction (see below for biological significance). The number of additional gene pairs merely inferred by the TC method amounts 1442 and is still significant (Fig. 2B).

Biological significance of the gene pairs inferred

To assess the biological significance of the inferred functional associations, the gene pairs are compared to known biological processes and protein functions, known protein interactions and regulatory interactions in yeast respectively. The results from the three methods are also compared to each other. The results show that the TC method can significantly enhance the LC and PCC methods to infer functional linkages and biological network, and is well suited to explore temporal relationships of gene expression in time-series data as detailed below.

Biological process and protein function-similarity gene pairs

In order to generally assess the biological significance of the gene pairs inferred, we first use two databases of biological processes and protein functions classification (supplementary Table S1) [see Additional file 1]. The *S. cerevisiae* Genome Database (SGD) mainly utilizes the Gene Ontology (GO) annotations [20]. We use 32 main biological processes (i.e. conjugation). In this work, if two genes in the pair inferred are involved in the same biological process, we consider the gene pair as a process-identity one. 19.9% of the 108489 gene pairs inferred with the TC method at a *p*-value threshold of $2.7E-3$ (Fig. 2A) are found to be process-identity pairs (Fig. 2C). The detailed distribution of the process-identity pairs in each biological process is listed in Table S3. A similar ratio (20.4%) of process homology gene pairs is found for the 174768 gene pairs inferred by the PCC method. Only 14.6% of the 342594 gene pairs inferred by the LC method with the same *p*-value cutoff are process-identity pairs. If the results inferred by procedure I and procedure II of the TC method are separately considered, 22.9% of the 58376 gene pairs detected by procedure I are process-identity pairs. The percentage (17.8%) of process-identity pairs among the 81271 gene pairs detected by procedure II is slightly lower

than that by procedure I but still somewhat higher than that (14.6%) of the LC method. The additional number (Fig. 2C) of process-identity pairs inferred by the TC method is considerable compared to those resulted only from the LC method (17.8%) and the PCC method (32%) respectively.

Among the 6490 gene pairs (Fig. 2B) predicted by the TC method with a p-value threshold of $1E-5$ there are 3138 (48.3%) process-identity pairs (Fig. 2D). Separately, only 32.6% of the gene pairs (1400) detected by procedure I are process-identity pairs, in contrast to as high as 50.8% for the gene pairs by procedure II. Considering the results from $p\text{-value} \leq 2.7E-3$ mentioned above no general conclusion can be drawn with regard to the question which extraction procedure is more relevant. The percentage of the process-identity pairs among the genes pairs resulted from the LC method and the PCC method is 49% and 42% respectively. Thus, the lower p-value threshold can significantly increase the portion of gene pairs involved in the same biological processes in all the three methods. At this low p-value the additional number (584, Fig. 2D) of process-identity pairs merely inferred by the TC method amounts to 12.9% of those only resulted from the LC method. Compared with those resulted by the PCC method this number declines to 286 (3.3%), suggesting that the gene pairs with a higher ranked functional association inferred by the TC method is more similar to those resulted by the PCC method.

With the percentage of process-identity pairs in the range of 14.6–20.4 at $p\text{-value} \leq 2.7E-3$ the gene pairs inferred by the three methods seem to have a relatively low biological significance. If the common part of the gene pairs inferred by all the three methods is considered (Fig. 2A), the percentage of process-identity pairs (Fig. 2C) increases to 34.8%, resulting in fairly good biological significance for an *in silico* method of biological function inference. The biological significance can be significantly increased by lowering the p-value. At $p\text{-value} \leq 1E-5$ the percentage of process-identity pairs ranges from 42 to 49% for the three methods. If the common part of the gene pairs inferred by the three methods at this p-value is considered (Fig. 2B), the percentage of process-identity pairs (Fig. 2D) increases to 60.2%, resulting in a satisfactorily high biological significance.

The second source used in this work for assessing biological relevance of the gene pairs is the Munich Information Center for Protein Sequences (MIPS, [22]) functional catalogue database. For protein functional classification, the MIPS database contains up to 6 different levels within the hierarchy (i.e. metabolism in the first level). We use here the second level of MIPS (i.e. respiration) as Qian et al. [8] did. Altogether 158 function classes are used. It should be

mentioned that some functional classes merely belong to protein cellular functions of plants and animals. Here if two genes in the pair inferred have the same protein cellular function, we term the pair as a function-similarity pair. The results of comparison among the three methods (Fig. S2) are similar to those of process-identity pairs, and detailed distribution of the function-similarity pairs in each protein cellular function by the trend correlation method is provided in Table S4 [see Additional file 1]. As in the case of function-similarity pairs, if the common part of the inferred gene pairs (Figs. S2A and S2B) from the three methods is considered the percentage of function-similarity pairs (Figs. S2A and S2B) increases to 31.7% and 55.6% at the two p-values respectively, resulting in a good basis for further experimental and functional studies of the gene pairs inferred.

Comparison of inferred gene pairs with known protein interactions

To further assess the biological significance of the gene pairs inferred and especially for comparing the three methods we examine here the known protein-protein interactions (including protein complexes) in current databases of yeast among the gene pairs inferred from the yeast cell cycle data by the different methods. Four databases and published high quality datasets (Table S2) are chosen for this purpose.

The protein-protein interactions collection of Yu et al [21] integrates datasets from the databases of MIPS [22], the Database of Interacting Proteins (DIP, [23]), the Biomolecular Interaction Network Database (BIND, [24]) and the experimental datasets of yeast two-hybrid [25,26] and high-throughput mass spectrometry measurements [27,28]. Many of the interactions are manually curated beyond the experimentally derived protein-protein interactions in the three databases mentioned above [21].

With a p-value threshold of $2.7E-3$ the TC method detects 2701 gene pairs among the 65160 known protein interactions. Separately, extraction procedure I detects 1152 protein interaction pairs, compared to 2542 pairs by extraction procedure II. With the same p-value threshold 5187 and 4637 such gene pairs can be detected by the LC method and the PCC method respectively. The relatively low number of protein interactions among the gene pairs inferred by the TC method is a consequence of the overall lower number of gene pairs inferred by the TC method (Fig. 2A). In fact, the percentage of protein interaction pairs (Fig. 2E) in the overall gene pairs (Fig. 2A) inferred by the TC method ($2701/108489 = 2.49\%$) is comparable to that of the PCC method (2.65%) but higher than that of LC method (1.51%). Especially the extraction procedure II of the TC method achieves a relatively higher percentage ($2542/81271 = 3.13\%$). It should be mentioned

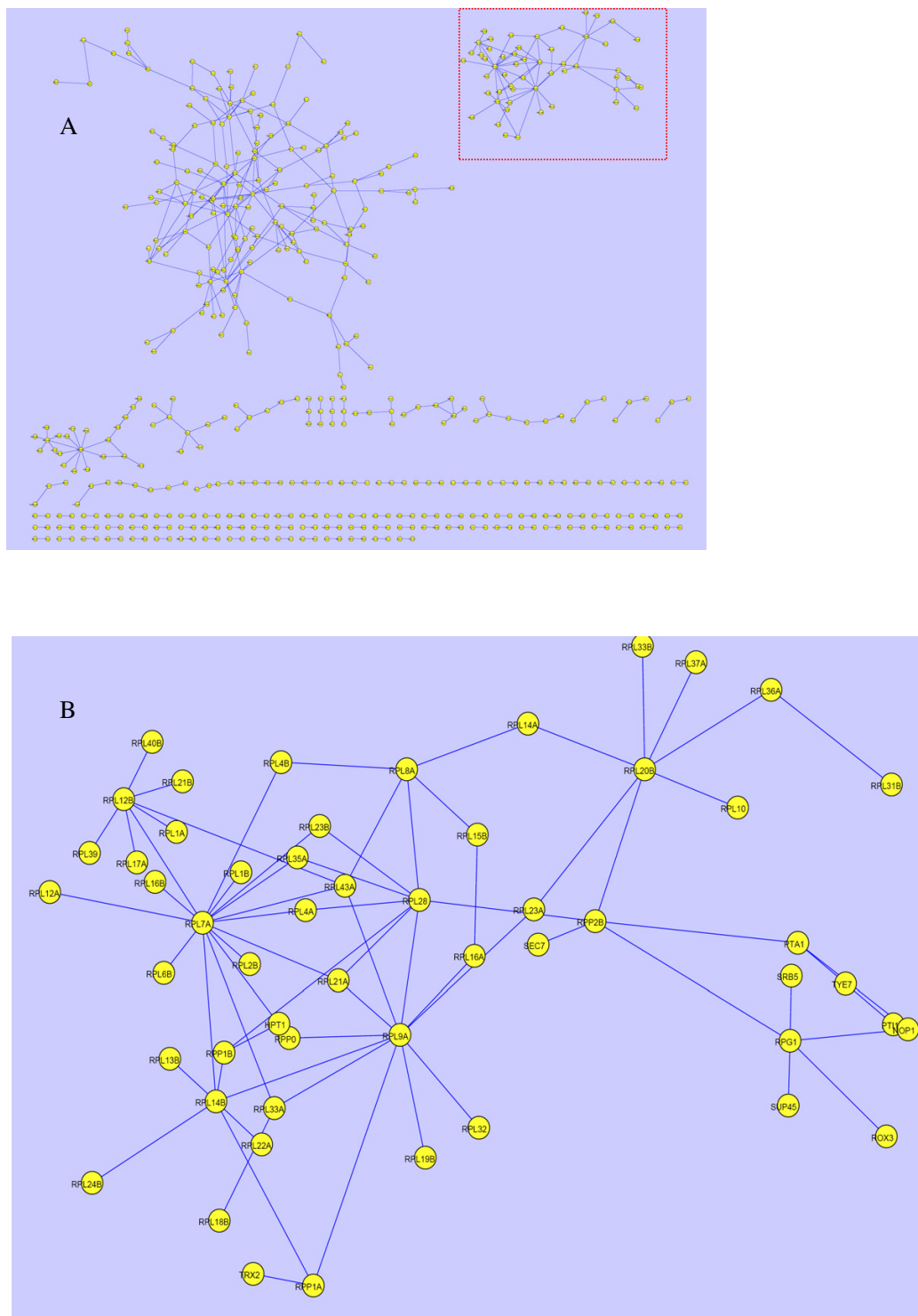


Figure 3

A, Protein interactions (including protein complexes) only detected by the trend correlation method (443 interactions) [see Additional file 2] with a p-value threshold of $2.7E-3$ which are known in the protein interactions collection dataset [21]. The network was visualized by Cytoscape <http://www.cytoscape.org/>. B, the zooming-in part (red dashed rectangle in the Fig. A) of protein components of the large ribosomal subunit.

Table 1: Results by the trend correlation (TC) method compared to those resulted from the local clustering method (LC) and PCC based clustering method in four protein interactions datasets (p-value threshold 2.7E-3 if not otherwise mentioned).

		Collection dataset	Collection Dataset with p-value $\leq 1.3E-2$	MIPS	DIP	BIND
Compared to the LC method	Only TC	453	1189	62	58	111
	Both	2248	3246	84	64	91
	Only LC	2939	4625	393	362	555
	Additional by TC	15.4%	25.7%	15.8%	16.0%	20.0%
Compared to the PCC method	Only TC	479	1009	66	60	112
	Both	2222	3426	80	62	90
	Only PCC	2415	5899	228	243	373
	Additional by TC	19.8%	17.1%	28.9%	24.7%	30.0%

that the low percentage in all these cases is not surprising since only data from cell cycle is used in the inference and the cell cycle represents only a part of cellular functions involving protein-protein interactions in yeast. The results presented in this section are thus more suitable for comparing the different methods rather than for quantitatively assessing the biological significance of the gene pairs inferred. Fig. 2E presents a comparison of the numbers of unique and common pairs of protein-protein interactions inferred by the three methods. The additional pairs (453–479) merely (Fig. 3A) detected by the TC method are significant compared with those resulted merely from the LC method (15.4%) and the PCC method (19.8%). Remarkably, 58 protein-protein interactions among 42 (Fig. 3B) protein components of the large ribosomal subunit of yeast are only significantly detected by the TC method. The fact that the 42 proteins connect with each other by the detected interactions is consistent with the assembling phenomena of components of the large ribosomal subunit. Similarly, 22 interactions among 22 protein components of the small ribosomal subunit of yeast are also only significantly detected by the TC method. The 22 proteins also connect with each other by the 22 interactions (Fig. 3A). Therefore, to have a more complete coverage of the protein interactions it is obvious that the TC method should be combined with the LC and PCC methods.

By loosening the p-value threshold from 2.7E-3 to 1.3E-2 the TC method can detect significantly more pairs (4435) of known protein-protein interactions (Fig. 2F). A p-value of 1.3E-2 in the TC method is equivalent to a Pearson correlated coefficient of 0.6 in the PCC method and a score of 12 in the LC method [29]. With this p-value threshold, 7871 and 9325 pairs can be detected by the LC method and the PCC respectively. The number of protein-protein interactions additionally inferred by the TC method is also considerable (Table 1). When the MIPS, DIP and BIND updated databases are used individually for comparison similar results are obtained and summarized in

Table 1. Though the majority of the three databases is consistent some differences exist among them. We chose them to cover more known protein interactions and could therefore obtain more reliable statistics results.

Comparison with known regulatory interactions

To examine the biological significance of functional associations inferred by the different methods it is also interesting to know if the gene pairs inferred cover some of the known regulatory interactions, especially those involved in the regulation of cell cycle. For this purpose, we use two regulatory datasets (Table S2) as a comparison basis. We first use the dataset including regulatory interactions which are confirmed with a p-value threshold of $1e^{-3}$ by genome wide location analysis (GWLA) [30]. With a p-value threshold of 2.7E-3, only a relative low number of the regulatory interactions is detected by the LC method (127), the PCC method (47) and the TC (24) (Fig. S3A). Regulatory interactions detected by the TC method are significantly less than those by the other two methods. Nevertheless, the additional number of regulatory interactions detected by the TC method is considerable compared with those only resulted from the LC method (about 12%) and the PCC method (about 37.8%).

Similar results are obtained when the regulatory interaction collection dataset of Luscombe et al. [31] is used (Fig. S3B). The additional number of interactions predicted by the TC method is 21 (11%) and 22 (about 32%) compared with those resulted only from the LC method and the PCC with a p-value threshold of 2.7E-3. With a p-value threshold of 1.3E-2, the results of comparison are similar to those at a p-value cutoff of 2.7E-3. Some of the typical interactions between transcriptional regulators and target genes which are detected by the TC method but cannot be significantly detected by the LC and/or the PCC method are summarized in Table S5.

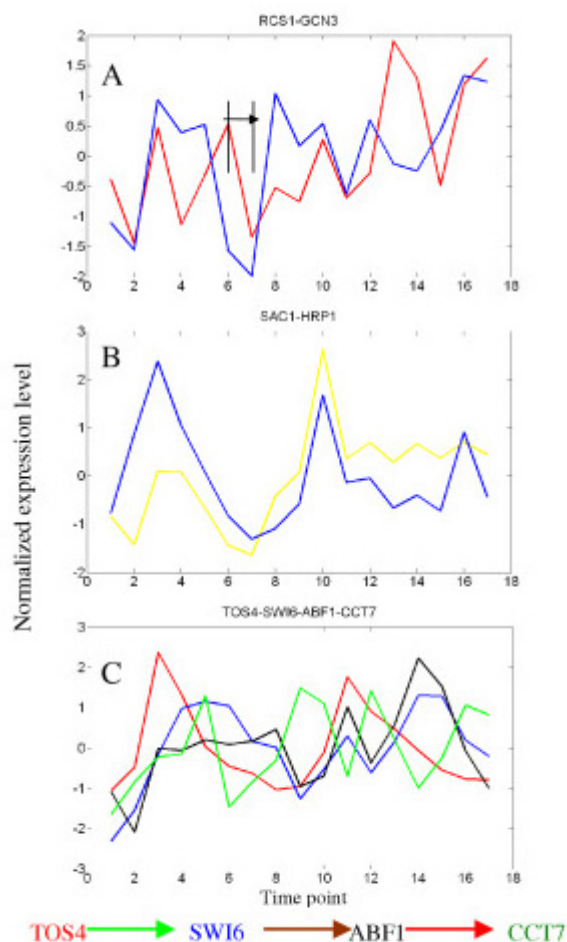


Figure 4

A and B, examples of functional linkages detected only by the new method with a p-value cutoff of $2.7E-3$. C, an example of more complete regulatory motifs detected by combining the three methods. Red, regulator; yellow, SAC1 in Fig.B; the legend of linkages in Fig.C is same to that of Fig.5.

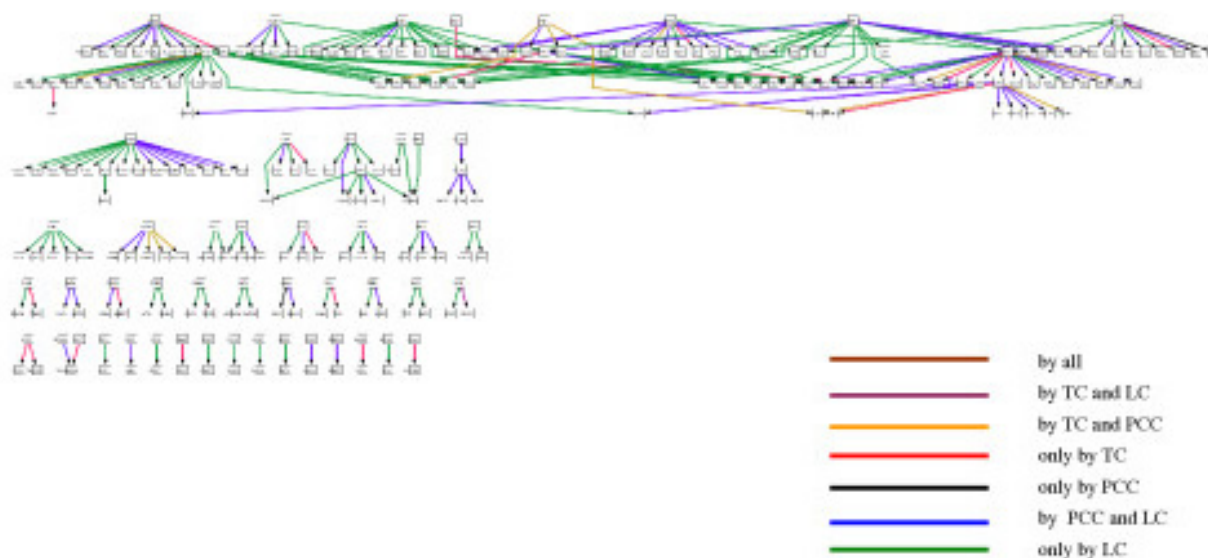
A detailed investigation of the results from a p-value threshold of $2.7E-3$ shows that the transcriptional factors and/or the targeted genes of 29 [see Additional file 3] among 34 known regulatory interactions inferred by the TC method are involved in cell cycle regulation according to Luscombe et al. [31]. Among the 29 interactions only six interactions have both the activated regulator and the differentially expressed targeted gene according to the gene set of cell cycle regulation given in Luscombe et al. [31]. The fact that 23 of the known interactions have either an inactivated regulator or a non-differentially expressed targeted gene in the cell cycle condition as detected by the TC method with a high p-value cutoff

seems to indicate controversies in the designation of activated regulators and differentially expressed genes in the cell cycle condition [31]. For example, the gene FKH2 is annotated in Hollenhorst et al. [32] as a transcriptional factor of the forkhead family that regulates the cell cycle according to SGD. Furthermore, many interactions known to involve FKH2 are detected by at least one of the three methods. However, FKH2 is not regarded as a regulator in the cell cycle condition by Luscombe et al. [31].

We further present two examples of predicted functional linkages which have well-documented biological relationships but cannot be significantly detected by the PCC method and the LC method. The regulatory relationship between RCS1 and GCN3 (Fig. 4A) was confirmed by GWLA with a p-value of $9.2E-6$. The Pearson correlation coefficient for this gene pair is as low as 0.40 and the score of the LC is only 6.88 (an example of detailed calculation is provided in Table S6). We detect this gene pair with a maximal matched inverted change trend sc of 15 after RCS1 is shifted forward by one time point and obtain a cc value of 0.70 between the matched change levels by the TC method. Another example (Fig. 4B) is the proteins of the genes SAC1 and HRP1 that form a protein complex [21]. Though the score of the LC method is only 10.35 and the Pearson correlation coefficient is 0.44, we could detect this gene pair with a maximal matched similar change trend sc of 15 and obtained a cc value of 0.85 by the TC method.

The inference of regulatory circuit and network

As shown above, functional linkages inferred by the individual method do not give a complete picture. It is conceivable that this would be especially the case for inferring regulatory circuits and further for reconstructing regulatory network. Fig. 5 shows comparisons of inferred functional relationships by the individual methods and their combination which are mapped into regulatory circuits (motifs) and networks based on the datasets [30,31]. In consistence with the results of detected regulatory interactions (Fig. S3) the LC method is superior to the other two methods but still gives a very much incomplete picture. The TC and PCC methods can augment the LC method significantly. To infer more complete regulatory circuits and network the three methods should be therefore combined. This is exemplified with two relatively simple regulatory circuits. For the regulatory chain (Fig. 4C) from the regulator TOS4 to the target gene CCT7 the negative relationship between TOS4 and SWI6 can only be detected by the LC method with a shift of two-time-points. The interaction between SWI6 and ABF1 can be detected by all the three methods. However, the interaction between ABF1 and CCT7 can only be detected by the TC method. The three interactions can form one regulatory chain among the complex combination of several motifs (Fig. 5).

**Figure 5**

Regulatory network of yeast resulted by combining the three methods with a p-value threshold of $2.7E-3$ in the dataset of cell cycle. The network was layout by Cytoscape <http://www.cytoscape.org/>. For the interactions in this network [see Additional file 3].

Another example about a single input motif is explained in Fig. S4 [see Additional file 1].

Discussion

From a practical point of view there are two key general issues in the analysis of gene expression data. First, we ought to infer functional relationships with high statistic significance and as completely as possible. Second, the functional relationships inferred should have a high biological significance. We proposed a new method in this work and evaluated it mainly regarding these two aspects with the microarray data of yeast cell cycle [18]. The new method is also compared with the local clustering method and the Pearson correlation coefficient based clustering method. The number of functional gene pairs inferred depends very much on the p-value threshold in all the three methods (Fig. 2). With the two p-value cutoffs ($2.7E-3$ and $1E-5$) applied the TC method detects general significantly lower number of gene pairs. Nevertheless, a considerable number of gene pairs is only detected by the TC method, ranging from 20.5% of the number merely detected by the LC method to as high as 49.6% of that merely detected by the PCC method at p-value $\leq 2.7E-3$. The ratio of additional gene pairs inferred by our method remains similar (22%) at p-value $\leq 1E-5$ compared with the LC method but decreases to a relatively lower value

(6.7%) compared with the PCC method. This can be explained by the fact that the statistically higher ranked correlations have mostly simultaneous relationships that can be well detected by all the three methods. Since p-value $\leq 2.7E-3$ represents a relatively high statistic significance it can be concluded that the new method detects a significantly high portion of additional functional relationships compared to the other two methods. Since the shared part of gene pairs of the three methods at p-value $\leq 2.7E-3$ is less than 50% (12.5–39.5% for the individual method pair) it is also obvious that these methods should be combined to have a more complete exploitation of functional associations of genes buried in time-series expression data.

Concerning the biological significance 19.9% and 48.3% of the gene pairs (Figs. 2A and 2B) inferred by our method (at p-value $\leq 2.7E-3$ and $1E-5$ respectively) are process-identity ones (Figs. 2C and 2D) by comparing with the known biological processes in the *S. cerevisiae* Genome Database (SGD). This is compared with a process-identity ratio in the range of 14.6–49% for the LC method and 20.4–42 % for the PCC method. Similar results are obtained when the MIPS protein functional catalogue database is used to assess the biological significance (Fig. S2). These results suggest that the gene pairs detected by

the three methods achieve fairly good and comparable biological significance. As in the case of the total gene pairs the additional gene pairs with biological significance inferred merely by the trend correlation method is considerable compared to the other two methods (Figs. 2 and S2).

We also examined the gene pairs inferred by the three methods with regard to known protein interactions (Figs. 2E and 2F and Table 1) and known regulatory interactions (Fig. S3) in yeast. In general, the percentage of protein interaction pairs in the overall gene pairs is low and comparable for all the methods (2.5% for our method (3.1% for procedure II), 1.5% for the LC method and 2.7% for the PCC method). This can be understood in view of the fact that only cell cycle data under very specific conditions are used here and the cell cycle represents only a small portion of the cellular activities. Furthermore, only protein interaction pairs with significantly changed expression levels of all the involved partner proteins can be theoretically detected. This applies also to the known regulatory interactions. It is shown that most of the detected regulatory interactions are indeed involved in the cell cycle regulation. Nevertheless, the number of protein and regulatory interactions additionally inferred merely by the trend correlation method is significant in all the cases. It is obvious that more molecular interactions can be obtained if more microarray datasets under different conditions are considered [14].

Given the large number of functionally associated gene pairs inferred by the different methods an important question arises as to how we can find gene pairs which have really a high biological significance and thus would be best candidates for further experimental and functional studies. To this end, the shared part of gene pairs inferred by all the three methods is of particular interest, especially at low p-values. It is found for example that the percentage of process and function-similarity gene pairs of the shared pairs at p-value $\leq 1E-5$ can be as high as 60.2% and 55.6%, building a very good basis for experimental study. The common part of the TC and LC methods would be also of particular interest for finding time-delayed and/or inverted functional relationships which have received so far less attention. It should be mentioned that all the functional associated pairs predicted have a high probability to be true and can thus serve as hypotheses for further study in view of the high statistic significance criteria applied. We would also like to emphasize that the combined use of the different methods is not only useful for finding more and highly possible potential candidates of functional association but also very important to infer more complete regulatory circuits and network as exemplified in this study (Figs. 4, 5 and 4).

Conclusion

The major difference between our method and the other current methods is that the trend correlation method is based on the main change trend and comprehensively considers correlation coefficient between the main change trend of two genes, whereas the other methods are mainly based on the correlation of point-to-point expression levels of two genes. Hence the trend correlation method can reveal additional gene pairs with same function or in the same biological process but yet not significantly co-expressed. It therefore also can infer additional protein-protein and regulatory network as demonstrated. As mentioned above, a combined use of different methods is presently necessary for the analysis of time-series microarray data. As clearly demonstrated in this work the proposed new method can significantly augment the currently major methods and is well suitable for exploring temporal relationships of gene expression in time-series data.

Methods

Maximal local alignment of expression change trend

The calculation of the maximal local alignment of expression change trend is similar to the algorithm of local sequence alignment [19] and the local clustering method of Qian et al. [8] for gene expression level. Different from the local clustering method that merely compares the expression levels at each time point, the change trends between time points are used in our method. Considering an expression profiling dataset of n time points the expression ratios at the n time points are first normalized in the "z-score" fashion, resulting in an average expression ratio of zero and a standard deviation of 1 for each gene. The normalized expression level at time point i for gene X is denoted as X_i . The change level between time points i and $i+1$ for gene X is denoted as Xc_i^{i+1} . The change trend for Xc_i^{i+1} is denoted as Xct_i^{i+1} . Xct_i^{i+1} can have only one of the three possible values 1, -1 and 0, corresponding to increase, decrease and no change of gene expression between time points i and $i+1$ respectively. A matrix of possible alignment between the change trends for gene X and gene Y can then be formed. In our algorithm two matrices P and N are calculated in the following way:

if $Xc_i^{i+1} = Yct_j^{j+1}$, $P_{i,j} = P_{i-1,j-1} + 1$; else $P_{i,j} = P_{i-1,j-1}$

and

if $Xc_i^{i+1} * Yct_j^{j+1} < 0$, $N_{i,j} = N_{i-1,j-1} + 1$; else $N_{i,j} = N_{i-1,j-1}$

The initial conditions are $P_{0,j} = 0$ and $P_{i,0} = 0$, and the same initial conditions are also applied to the matrix of N. The purpose of calculating P and N is to find a local segment that has a maximal aggregated score, namely a maximal match of change trend between two expression patterns. This can be accomplished by using standard dynamic programming [19] as in the local clustering method and results in an alignment of lc matched change trend, where $lc \leq n-1$ (altogether $n-1$ changes among n time points).

Finally, an overall maximal value sc is found by comparing the maximums for matrices P and N. The maximal value is the matched change trend score sc for the two expression patterns. A maximal value from the matrix P means a positively correlated expression pattern of the two genes, whereas a maximal value from the matrix N indicates that these two patterns have an inverted relationship. If the maximum is off-diagonal in its corresponding matrix, then the two expression patterns have a time-shifted relationship. When the number of matched change trends between two genes is relatively small, it is possible that several repeated maximal values exist with different number of shifted-time-points. In this case, we choose the maximal matched change trends with the shortest time shift in the algorithm. Normally these gene pairs are not regarded as inferred linkages because of the small matched change trends based on the p-value threshold.

Correlation coefficient between the maximal matched change trend of two genes

After obtaining the maximal matched change trend, we need to calculate the correlation coefficient of the maximal matched change trend of the gene pair (Fig. 1). It is calculated according to the following algorithm. Assuming that the final matched change trend for gene X and gene Y are Xct_{xf-1}^{xf} and Yct_{yf-1}^{yf} respectively in the maximal alignment, where xf and yf refer to the last time points of genes X and Y in the maximal alignment and we assign $xf \leq yf$. We then give a new index for the time points of match trends, where the matched change levels for genes X and Y are now denoted as Xmc and Ymc in the following algorithm:

Do while $2 \leq i \leq xf$ and $yf - xf + 2 \leq j \leq yf$

If sc from the matrix P and $Xct_{i-1}^i = Yct_{j-1}^j$ then

$$Xmc_{p-1}^p = Xc_{i-1}^i$$

$$Ymc_{p-1}^p = Yc_{j-1}^j$$

Endif

If sc from the matrix N and $Xct_{i-1}^i * Yct_{j-1}^j < 0$ then

$$Xmc_{n-1}^n = Xc_{i-1}^i$$

$$Ymc_{n-1}^n = Yc_{j-1}^j$$

Endif

Loop

In the following equation k corresponds to p or n in the above loop.

$$cc_{(x,y)} = \left| \frac{1}{sc} * \sum_{k=2,sc+1} \left(\frac{Xmc_{k-1}^k - \overline{Xmc}}{\sigma_{Xmc}} \right) \left(\frac{Ymc_{k-1}^k - \overline{Ymc}}{\sigma_{Ymc}} \right) \right| \quad \text{Eq.1}$$

Where

$cc_{(x,y)}$ is the correlation coefficient between the maximal matched change trends of genes X and Y. $\overline{Xmc}$ and $\overline{Ymc}$ are the mean of the maximal matched change levels of genes X and Y respectively σ_{Xmc} and σ_{Ymc} are the standard deviation of the maximal matched change levels.

Significance statistics

In order to estimate the p-value for a given score a dataset of random expression patterns was generated by shuffling the normalized expression levels of the original data at different time points. To assure the reliability of p-value we generated the random expression patterns three times, resulting in calculations for about $4.9e7$ gene pairs. Using the algorithm described above we calculated the maximal matched change trend scores sc and the correlation coefficient cc between the maximal matched change trend for each gene pair in the random expression dataset. We first tabulated the frequency ($f(sc)$) of sc and then the distribution of cc for pairs which have the same sc value. Through integration we could calculate the conventional p-value for the two scores sc (Fig. S1B) and cc (Fig. S1C) ($P_{sc}(s \geq sc)$, $P_{cc}(c \geq cc)$). They are defined as the probability of obtaining a score (s or c) larger than or equivalent to sc or cc from the random patterns. The higher the matched change trend score is, the more likely the gene pair is correlated (positively or negatively). The higher the correlated coefficient, the more likely the gene pairs are correlated for a given value of sc .

Extraction procedure I

As shown in Fig. S1B the probability to achieve a high sc score in the random dataset is very low. For example, the p-value for gene pairs with $sc \geq 15$ is as low as $1.7e^{-3}$ and with $sc \geq 16$ as low as $1.1e^{-4}$. Hence, if two genes in the original expression data have a sc score ≥ 15 it is very likely that these two genes have a functional association that would be expected with a low probability by chance. For a given significant p-value threshold such as $2.7E^{-3}$, because the p-value of gene pairs with $sc \geq 15$ is only $1.7e^{-3}$ that is lower than the threshold, all the gene pairs with $sc \geq 15$ can be regarded as having functional relationships. But P_{sc} for gene pairs with $sc \geq 14$ is about $1.27e^{-2}$ that is higher than the threshold $2.7E^{-3}$, only parts of the gene pairs with $sc = 14$ should be extracted. Among the gene pairs with $sc = 14$ the pairs with a higher cc value (Fig. S1C) should have a higher possibility to have a functional association. Considering these facts and the distribution frequency of sc we propose to define the following overall p-value to better reflect the significance statistics for inferring functional associations among genes:

$$P = P_{sc}(sc + 1) + (P_{sc}(sc) - P_{sc}(sc + 1)) * P_{cc} \text{ or } P = P_{sc}(sc + 1) + f(sc) * P_{cc}$$

where if sc is the highest possible score (here 16 in the yeast cell cycle dataset with 17 time points), then $P_{sc}(sc + 1) = 0$. This overall p-value combines the p-values for both sc and cc and the distribution frequency of sc. If the sc and cc values of a gene pair in original expression dataset result in an overall p-value less than a certain threshold, this gene pair is considered as functionally associated. This extraction procedure is called here Procedure I.

Extraction procedure II

In applying extraction Procedure I we noticed that there are many gene pairs the change trend score (sc) of which is not very high, but the correlation coefficient between the maximal change levels of which is significantly high. These gene pairs will not be extracted according to Procedure I because of the relatively high P_{sc} values at low sc values (Fig. S1B). However, they may in fact also have a high possibility of functional association since the main change levels are correlated well. The difference of some of the change trends from the main trend between the two genes may be caused by the involvement of multi-regulators in some time regions or by measurement errors and expression noises. For these reasons, we propose a second procedure (Procedure II) to extract gene pairs with possible functional relationships. In Procedure II, if the correlation coefficient cc is not considered, the likelihood of functional association of all the pairs increases with increasing sc if sc is higher than or equal to a characteristic score sc_m . This characteristic score sc_m is defined as the sc that has the highest frequency in the random dataset (Fig.

S1A, here $sc_m = 10$). If sc is smaller than sc_m , the frequency of pairs with sc reduces along with the decrease of sc, as does also the possibility for a functional association of the gene pair. Because the possibility for functional association should have a reverse relationship with the frequency of the corresponding sc, one should not consider gene pairs with sc smaller than sc_m . We therefore propose to extract gene pairs through the formulas:

P-value of cc cutoff for each sc:

$$P_{cc}(i) = \frac{f}{f(sc_i)} (i = sc_m \text{ to } sc_{max})$$

Where $f = \frac{p}{m}$

sc_m , sc with the highest frequency.

sc_{max} , the maximal possible sc score in the time series dataset

m, the number between sc_m and sc_{max}

p, the given significant p-value threshold

f, the p-value threshold for each sc (between sc_m and sc_{max})

$f(sc_i)$, the frequency of gene pairs with sc_i

If the p-value of cc from a gene pair with sc_i in the original expression dataset as determined in Fig. S1C is less than the corresponding $P_{cc}(i)$, this gene pair is considered as functionally associated in the extraction procedure II.

Authors' contributions

FH initially proposed the method, wrote the programs and drafted the manuscript. APZ made suggestions to improve the method, supervised the study and revised the manuscript.

Additional material

Additional File 1

Supplementary material four supplementary figures and six supplementary tables.

Click here for file

[<http://www.biomedcentral.com/content/supplementary/1471-2105-7-69-S1.doc>]

Additional File 2

Table of protein-protein interactions in Fig. 3 all the protein-protein interactions and comparison among three methods in the figure 3.

Click here for file

[<http://www.biomedcentral.com/content/supplementary/1471-2105-7-69-S2.xls>]

Additional File 3

Table of regulatory interactions in Fig. 5 all the regulatory interactions in the figure 5 and the corresponding scores by the three methods.

Click here for file

[<http://www.biomedcentral.com/content/supplementary/1471-2105-7-69-S3.xls>]

Acknowledgements

We are grateful to Dr. Hong-Wu Ma for useful discussions and to Marcio Rosa da Silva for help with network visualization. We acknowledge the support by the Deutsche Forschungsgemeinschaft through SFB 578 and by the Federal Ministry for Education and Research (BMBF) through the bioinformatics project "Intergeomics".

References

- Bar-Joseph Z: **Analyzing time series gene expression data.** *Bioinformatics* 2004, **20**:2493-2503.
- D'haeseleer P, Wen X, Fuhrman S, Somogyi R: **Linear modeling of mRNA expression levels during CNS development and injury.** *Pac Symp Biocomput* 1999:41-52.
- Bar-Joseph Z, Gerber G, Jaakkola T, Gifford D, Simon I: **Comparing the continuous representation of time series expression profiles to identify differentially expressed genes.** *Proc Natl Acad Sci USA* 2003, **100**:10146-10151.
- Schliep A, Schonhuth A, Steinhoff C: **Using hidden Markov models to analyze gene expression time course data.** *Bioinformatics* 2003, **19**:1264-1272.
- Guthke R, Möller U, Hoffmann M, Thies F, Töpfer S: **Dynamic network reconstruction from gene expression data applied to immune response during bacterial infection.** *Bioinformatics* 2005, **21**:1626-1634.
- Eisen MB, Spellman PT, Brown PO, Botstein D: **Cluster analysis and display of genome-wide expression patterns.** *Proc Natl Acad Sci USA* 1998, **95**:14863-14868.
- Friedman N, Linial M, Nachman I, Pe'er D: **Using Bayesian network to analyze expression data.** *J Comput Biol* 2000, **7**:601-620.
- Qian J, Filhart MD, Lin J, Yu HY, Gerstein M: **Beyond synexpression relationships: local clustering of time-shifted and inverted gene expression profiles identifies new, biologically relevant interactions.** *J Mol Biol* 2001, **314**:1053-1066.
- Ong I, Glasner J, Page D: **Modelling regulatory pathways in E.Coli from time series expression profiles.** *Bioinformatics* 2002, **18**:S241-248.
- Perrin BE, Ralavivola L, Mazurie A, Bottani S, Mallet J, D'Alche-Buc F: **Gene network inference using dynamic bayesian networks.** *Bioinformatics* 2003, **19**:1138-1148.
- Zou M, Conzen SD: **A new dynamic Bayesian network(DBN) approach for identifying gene regulatory networks from time course microarray data.** *Bioinformatics* 2005, **21**:71-79.
- Magwene PM, Kim J: **Estimating genomic coexpression networks using first-order conditional independence.** *Genome Biol* 2004, **5**:R100.
- Stuart JM, Segal E, Koller D, Kim SK: **A gene-coexpression network for global discovery of conserved genetic modules.** *Science* 2003, **302**:249-255.
- Lee I, Data VS, Adai AT, Marcotte EM: **A probabilistic functional network of yeast genes.** *Science* 2004, **306**:1555-1558.
- Kwon AT, Hoos HH, Ng R: **Inference of transcriptional regulation relationships from gene expression data.** *Bioinformatics* 2003, **19**:905-912.
- Filkov V, Skiena S, Zhi JZ: **Analysis techniques for microarray time-series data.** *J Comput Biol* 2002, **9**:317-330.
- Balasubramanian R, Hüllermeier E, Weskamp N, Kämper J: **Clustering of gene expression data using a local shape-based similarity measure.** *Bioinformatics* 2005, **21**:1069-1077.
- Cho RJ, Campbell MJ, Winzler EA, Steinmetz L, Conway A, Wodicka L, Wolfsberg TG, Gabriellian AE, Landsman D, Lockhart DJ, Davis RW: **A genome-wide transcriptional analysis of the mitotic cell cycle.** *Mol Cell* 1998, **2**:65-73.
- Smith TF, Waterman MS: **Identification of common molecular subsequences.** *J Mol Biol* 1981, **147**:195-197.
- Dwight SS, Harris MA, Dolinski K, Ball CA, Binkley G, Christie KR, Fisk DG, Issel-Tarver L, Schroeder M, Sherlock G, Sethuraman A, Weng S, Botstein D, Cherry JM: **Saccharomyces Genome Database (SGD) provides secondary gene annotation using the Gene Ontology (GO).** *Nucleic Acids Res* 2002, **30**:69-72.
- Yu HY, Zhu XW, Greenbaum D, Karro J, Gerstein M: **TopNet: a tool for comparing biological sub-networks, correlating protein properties with topological statistics.** *Nucleic Acids Res* 2004, **32**:328-337.
- Mewes HW, Frishman D, Guldener U, Mannhaupt G, Mayer K, Mokrejs M, Morgenstern B, Munsterkötter M, Rudd S, Weil B: **MIPS: a database for genomes and protein sequences.** *Nucleic Acids Res* 2002, **30**:31-34.
- Xenarios I, Salwinski L, Duan XJ, Higney P, Kim SM, Eisenberg D: **DIP, the Database of Interacting Proteins: a research tool for studying cellular networks of protein interactions.** *Nucleic Acids Res* 2002, **30**:303-305.
- Bader GD, Betel D, Hogue CW: **BIND: the Biomolecular Interaction Network Database.** *Nucleic Acids Res* 2003, **31**:248-250.
- Ito T, Tashiro K, Muta S, Ozawa R, Chiba T, Nishizawa M, Yamamoto K, Kuhara S, Sakaki Y: **Toward a protein-protein interaction map of the budding yeast: a comprehensive system to examine two-hybrid interactions in all possible combinations between the yeast proteins.** *Proc Natl Acad Sci USA* 2000, **97**:1143-1147.
- Uetz P, Giot L, Cagney G, Mansfield TA, Judson RS, Knight JR, Lockshon D, Narayan V, Srinivasan M, Pochart P, Qureshi-emili A, Li Y, Godwin B, Conover D, Kalbfleisch T, Vijayadamar G, Yang M, Johnston M, Fields S, Rothberg JM: **A comprehensive analysis of protein-protein interactions in Saccharomyces cerevisiae.** *Nature* 2000, **403**:623-627.
- Ho Y, Gruhler A, Heilbut A, Bader GD, Moore L, Adams SL, Millar A, Taylor P, Bennett K, Boutilier K, Yang LY, Wolting C, Donaldson L, Schandorff S, Shewnarane J, Vo M, Taggart J, Goudreau M, Muskat B, Alfaro C, Dewar D, Lin Z, Michalickova K, Willems AR, Sassi H, Nielsen PA, Rasmussen KJ, Andersen JR, Johansen LE, Hansen LH, Jespersen H, Podtelejnikov A, Nielsen E, Crawford J, Poulsen V, Sørensen BD, Matthies J, Hendrickson RC, Gleeson F, Pawson T, Moran MF, Durocher D, Mann M, Hogue CW, Figeys D, Tyers M: **Systematic identification of protein complexes in Saccharomyces cerevisiae by mass spectrometry.** *Nature* 2002, **415**:180-183.
- Gavin AC, Bosche M, Krause R, Grandi P, Marzioch M, Bauer A, Schultz J, Rick JM, Michon AM, Cruciat CM, Remor M, Höfert C, Schelder M, Brajenovic M, Ruffner H, Merino A, Klein K, Hudak M, Dickson D, Rudi T, Gnau V, Bauch A, Bastuck S, Huhse B, Leutwein C, Heurtier M, Copley RR, Edelmann A, Querfurth E, Rybin V, Drewes G, Raida M, Bouwmeester T, Bork P, Seraphin B, Kuster B, Neubauer G, Superti-Furga G: **Functional organization of the yeast proteome by systematic analysis of protein complexes.** *Nature* 2002, **415**:141-147.
- Yu HY, Luscombe NM, Qian J, Gerstein M: **Genomic analysis of gene expression relationship relationships in transcriptional regulatory networks.** *Trends in Genetics* 2003, **19**:422-427.
- Lee TI, Rinaldi NJ, Robert F, Odom DT, Bar JZ, Gerber GK, Hannett NM, Harbison CT, Thompson CM, Simon I, Zeitlinger J, Jennings EG, Murray HL, Gordon DB, Ren B, Wyrick JJ, Tagne JB, Volkert TL, Fraenkel E, Gifford DK, Young RA: **Transcriptional Regulatory Networks in Saccharomyces cerevisiae.** *Science* 2002, **298**:799-804.
- Luscombe NM, Babu MM, Yu HY, Snyder M, Teichmann SA, Gerstein M: **Genomic analysis of regulatory network dynamics reveals large topological changes.** *Nature* 2004, **431**:308-312.
- Hollenhorst PC, Pietz G, Fox CA: **Mechanisms controlling differential promoter-occupancy by the yeast forkhead proteins**

Fkh1p and Fkh2p: implications for regulating the cell cycle and differentiation. *Genes Dev* 2001, **15**:2445-2456.

Publish with **BioMed Central** and every scientist can read your work free of charge

"BioMed Central will be the most significant development for disseminating the results of biomedical research in our lifetime."

Sir Paul Nurse, Cancer Research UK

Your research papers will be:

- available free of charge to the entire biomedical community
- peer reviewed and published immediately upon acceptance
- cited in PubMed and archived on PubMed Central
- yours — you keep the copyright

Submit your manuscript here:
http://www.biomedcentral.com/info/publishing_adv.asp



Supplementary materials to the manuscript

**In search of functional association from time-series microarray data based on the
change trend and level of gene expression**

By Feng He and An-Ping Zeng

2005-06

List

Four supplementary figures

Six supplementary tables

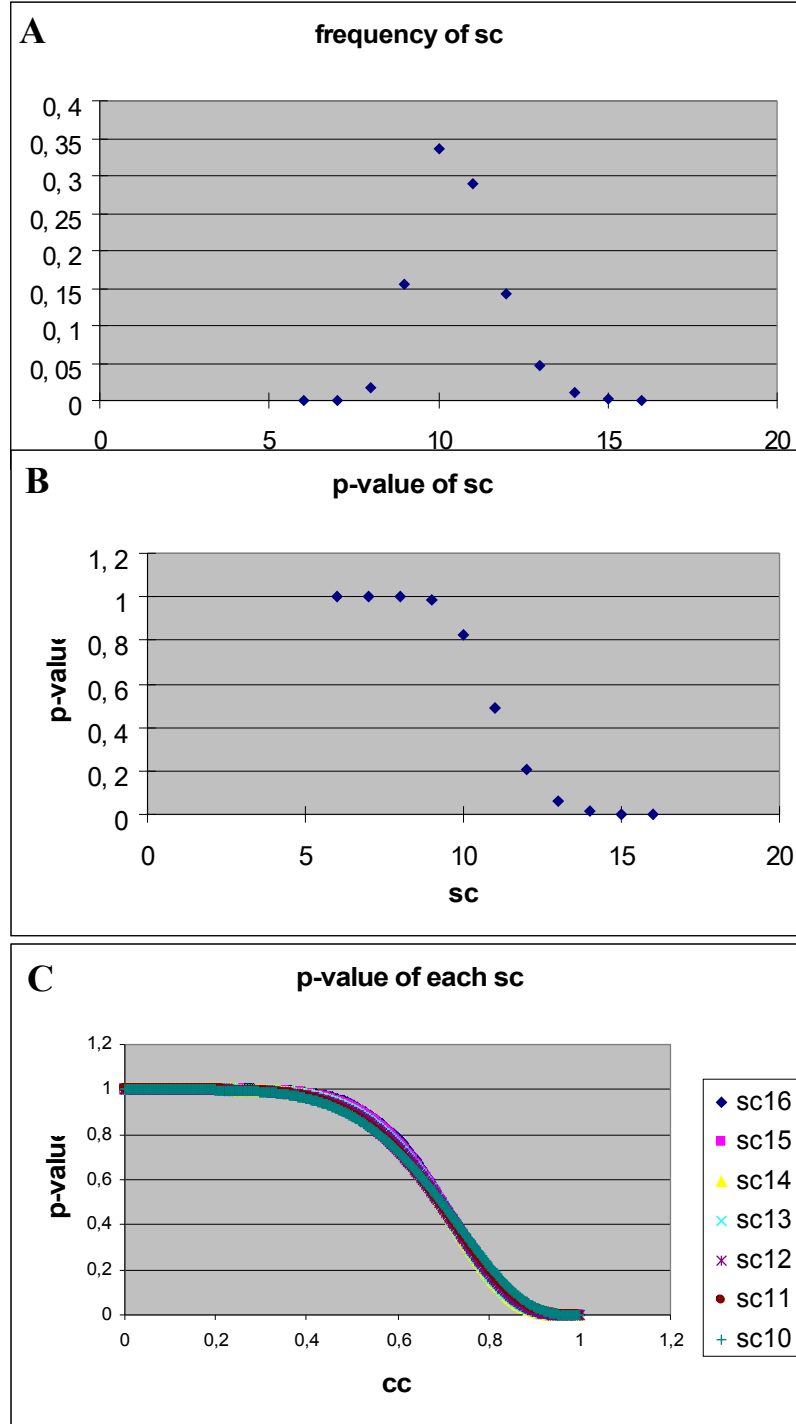


Fig.S1. A. Frequency of sc; B: p-value for sc; C: p-value for cc at each sc in the randomly shuffled expression data of yeast cell cycle (Cho et al.[18]. If a gene pair has an sc value of 14 and a cc value of 0.86, an overall p-value is calculated as $2.3e^{-3}$ (with $P_{sc}(15) = 0.0017$ and $P_{sc}(14) = 0.0127$ and $P_{cc} = 0.0573$). With a threshold p-value of $2.7e^{-3}$, this gene pair is considered to be functionally associated with a statistically high probability in the extraction procedure I proposed.

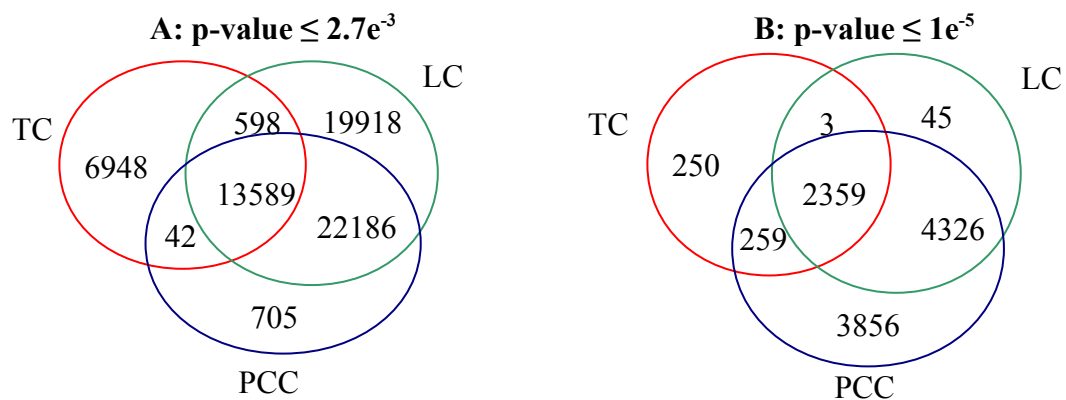
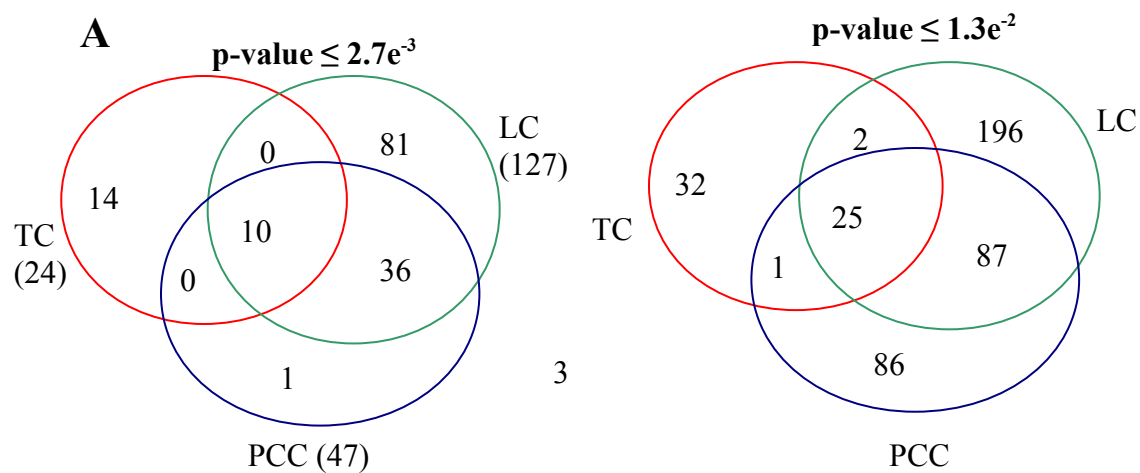


Fig.S2. Function-similarity pairs (based on MIPS database) inferred by the TC method versus those resulted from the LC method and the conventional PCC clustering method.



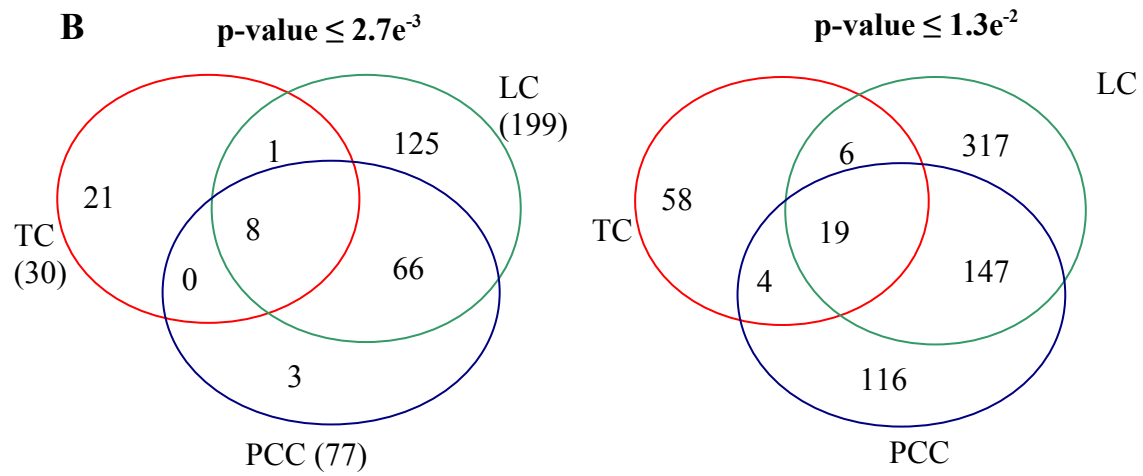


Fig.S3. A, results by the TC method vs. the LC method and the PCC clustering method respectively according to the dataset of genome wide location analysis (Lee et al. [14]); B, results by the TC method vs. the LC method and the PCC clustering method respectively according to the regulatory interactions collection dataset (Luscombe et al. [31]). The number in parenthesis is the whole number of regulatory interactions detected by the corresponding method with a p-value threshold of $2.7e^{-3}$.

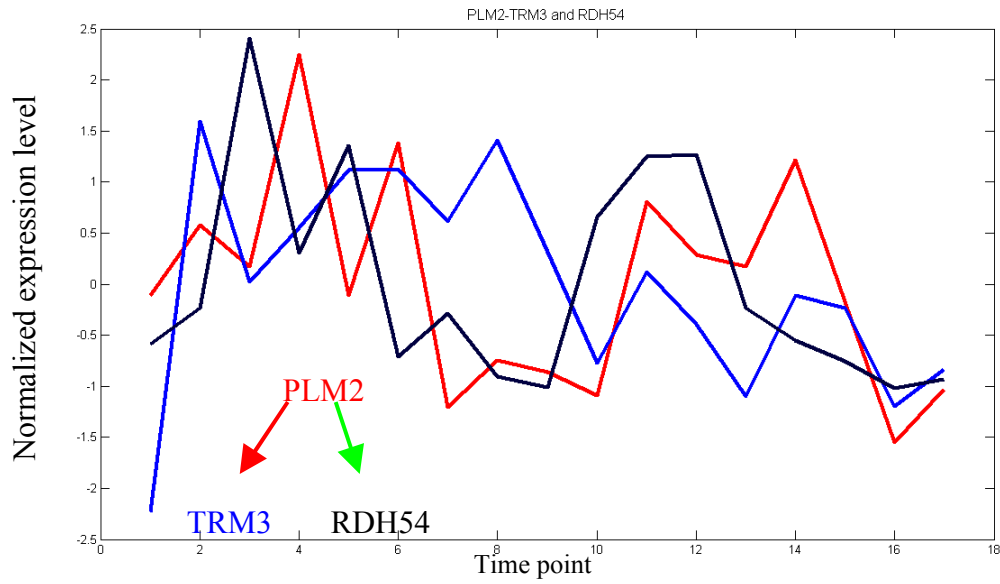


Fig.S4. Another example of more complete regulatory motifs detected by combining the three methods. The legend of linkages is same to that of Fig.5 (For details see text). The transcriptional regulator PLM2 is known to regulate TRM3 and RDH54 forming one part of a single input motif (Lee et al., [14]; Luscombe et al., [31]). But the two interactions between the regulator PLM2 and target genes TRM3 and RDH54 can only be significantly detected by TC method and LC method, respectively.

Table S1. Databases of biological processes and protein cellular function classification.

Database	Number of terms	Downloaded date
SGD	32	01-20-2005
MIPS	158	02-15-2005

Note: In this work, we have chosen all the biological processes in the list of advanced search in SGD except for the class biological process unknown.

Table S2. Databases or datasets of protein-protein interactions and regulatory interactions.

Type	Datasets	Number of gene pairs	Downloaded date
Protein interactions	Collection dataset (Yu et al., [21])	65160	Published 01-14-2004
	MIPS	13895	01-18-2005
	DIP	14187	02-06-2005
	BIND	27480	02-02-2005
Regulatory interactions	Genome wide location analysis (Lee et al., [14])	3760	Published 10-25-2002
	Collection dataset (Luscombe et al., [31])	6105	Published 09-16-2004

Note: in the six datasets, we exclude the pairs with two same genes and the pairs with genes which don't exist in the used Cho cell cycle dataset.

Table S3. Distribution of process-identity pairs inferred by the proposed method (with a p-value threshold of $2.7e^{-3}$) in each biological process class of the database SGD.

Biological process	Number of genes	Number of pairs
DNA metabolism	430	711
RNA metabolism	426	514
amino acid and derivative metabolism	188	153
carbon metabolism	190	124
cell budding	77	44
cell cycle	508	855
cell homeostasis	106	57
cell wall organization and biogenesis	138	90
cellular respiration	86	19
conjugation	100	24
cytokinesis	96	44
cytoskeleton organization and biogenesis	290	242
electron transport	21	4
generation of precursor and energy	222	245
lipid metabolism	201	73
meiosis	127	29
membrane organization and biogenesis	29	6
morphogenesis	140	82
nuclear organization and biogenesis	60	2
organelle organization and biogenesis	944	3964
protein biosynthesis	461	13117
protein catabolism	156	68
protein modification	390	326
pseudohyphal growth	48	9
response of stress	347	501
ribosome biogenesis and assembling	226	828
signal transduction	155	45
sporulation	94	36
transcription	465	535
transport	851	1907
vesicle mediated transport	256	223
vitamin metabolism	72	14

Table S4. Distribution of function-similarity pairs inferred by the proposed method (with a p-value threshold of $2.7e^{-3}$) in each protein cellular function class of the database MIPS (only genes existing in the chosen Cho dataset are included in the table).

Protein function	Number of genes	Number of pairs
amino acid metabolism	242	165
nitrogen and sulfur metabolism	91	26
nucleotide metabolism	225	197
phosphate metabolism	416	408
C-compound and carbohydrate metabolism	498	694
lipid, fatty acid and isoprenoid metabolism	265	119
metabolism of vitamins, cofactors, and prosthetic groups	161	80
secondary metabolism	70	8
extracellular metabolism	1	0
glycolysis and gluconeogenesis	54	41
glyoxylate cycle	9	0
Entner-Doudoroff pathway	0	0
pentose-phosphate pathway	23	1
pyruvate dehydrogenase complex	4	0
anaplerotic reactions	0	0
tricarboxylic-acid pathway (citrate cycle, Krebs cycle, TCA cycle)	31	3
electron transport and membrane-associated energy conservation	48	20
respiration	122	48
fermentation	48	15
chemolithotrophie (e.g. sulfide, nitrogenous compounds)	0	0
metabolism of energy reserves (e.g. glycogen, trehalose)	56	5
oxidation of fatty acids	6	0
photosynthesis	0	0
energy conversion and regeneration	35	3
storage facilitating proteins	0	0
stored proteins	0	0
DNA processing	500	867
cell cycle	638	1023
RNA synthesis	621	854
RNA processing	378	419
RNA modification	59	10
ribosome biogenesis	364	11314
translation	88	188
translational control	63	80
aminoacyl-tRNA-synthetases	37	10
protein folding and stabilization	90	51
protein targeting, sorting and translocation	280	245
protein modification	612	771
assembly of protein complexes	196	127
protein degradation	251	161
protein binding	372	322
peptide binding	3	0
nucleic acid binding	345	618
polysaccharide binding	0	0

motor protein	5	0
structural protein	52	10
lipid binding	17	0
amino acid binding	3	0
sulfate binding	0	0
C-compound binding	9	0
metal binding	28	4
nucleotide binding	221	107
complex cofactor/cosubstrate binding	55	9
mechanism of regulation	30	0
target of regulation	223	92
transported compounds (substrates)	567	1140
transport facilitation	184	113
transport routes	691	1316
intracellular signalling	193	86
transmembrane signal transduction	43	2
stress response	456	897
disease, virulence and defense	33	1
detoxification	111	50
degradation of foreign (exogenous) compounds	1	0
ionic homeostasis	171	108
membrane excitability	0	0
cell motility	0	0
cell adhesion	13	1
cellular sensing and response	284	157
nutrients uptake and absorption (e.g. digestion)	0	0
osmoregulation and excretion	2	0
gas and metabolite distribution	0	0
systemic temperature regulation	0	0
systemic rhythm control	0	0
plant / fungal specific systemic sensing and response	3	0
animal specific systemic sensing and response	3	0
LTR retroelements (retroviral)	0	0
non-LTR retroelements	0	0
transposons	0	0
viral proteins	0	0
phage proteins	0	0
proteins necessary for the integration or inhibition of transposon movement	8	0
cell growth / morphogenesis	245	192
cell differentiation	1	0
dedifferentiation	0	0
cell death	18	2
cell aging	28	1
fungal/microorganismic development	67	11
plant development	0	0
animal development	0	0
cell wall	218	140
eukaryotic plasma membrane	7	1
cytoplasm	1	0
cytoskeleton	261	184

centrosome	2	0
cell junction	0	0
endoplasmic reticulum	10	0
Golgi	7	0
intracellular transport vesicles	8	0
nucleus	158	94
mitochondrion	157	100
peroxisome	34	2
endosome	1	0
vacuole or lysosome	43	8
plastid	0	0
extracellular / secretion proteins	1	0
periplasmatic space	0	0
bud / growth tip	43	22
prokaryotic cytoplasmic membrane	5	0
flagellum	0	0
pilus/fimbria	0	0
prokaryotic cell envelope structures	0	0
prokaryotic intracytoplasmic membrane	0	0
prokaryotic cell inclusions	0	0
prokaryotic nucleoid	0	0
fungus/microorganismic cell type differentiation	453	591
plant cell type differentiation	0	0
animal cell type differentiation	0	0
fungus/microorganismic tissue	0	0
plant tissue	0	0
animal tissue	0	0
fungus organ	0	0
plant organ	0	0
animal organ	0	0
cell wall	0	0
eukaryotic plasma membrane / membrane attached	0	0
cytoplasm	0	0
cytoskeleton	0	0
centrosome	0	0
cell junction	0	0
endoplasmic reticulum	0	0
Golgi	0	0
intracellular transport vesicles	0	0
nucleus	0	0
mitochondrion	0	0
peroxisome	0	0
endosome	0	0
vacuole or lysosome	0	0
plastid	0	0
extracellular / secretion proteins	0	0
periplasmatic space	0	0
bud / growth tip	0	0
prokaryotic cytoplasmic membrane	0	0
flagellum	0	0

pilus/fimbria	0	0
prokaryotic cell envelope component	0	0
prokaryotic intracytoplasmic membrane	0	0
prokaryotic cell inclusions	0	0
prokaryotic nucleoid	0	0
fungal / microorganismic cell type	0	0
plant cell type	0	0
animal cell type	0	0
fungal/microorganismic tissue	0	0
plant tissue	0	0
animal tissue	0	0
fungal organ	0	0
plant organ	0	0
animal organ	0	0

Table S5. List of part of the regulatory interactions that cannot be significantly detected by the LC method and/or the PCC method but are detected by TC method with a p-value threshold of $1.3e^{-2}$.

Regulator	Target gene	sc	cc	Relationship	Local clustering score	Pearson correlated coefficient	P-value of GWLA
YOL089 C	YOL121 C	16	0.83112 2	Negative	9.940745 8	-0.55043	4.60E- 04
YER111 C	YJL196 C	15	0.80223 2	Positive	11.53894 8	0.64534	1.10E- 04
YGL071 W	YHL013 C	15	0.70216 4	Negative	8.941140 1	-0.30789	6.50E- 04
YGL071 W	YKR026 C	15	0.69694 2	Negative shift 1	6.879076 6	0.40465	9.20E- 06
YER111 C	YPL024 W	15	0.62385 2	Positive	7.532260 9	0.41849	7.90E- 06
YKL043 W	YCR09 8C	15	0.55109 3	Negative	9.028982 5	-0.44967	2.30E- 06
YJL056 C	YCR03 9C	14	0.85819 4	Negative	11.26858 5	-0.66286	7.80E- 04
YPL089 C	YGR21 8W	14	0.83304 4	Negative shift 2	10.46728 1	0.12039	5.80E- 04
YDR423 C	YLL060 C	14	0.81614 7	Negative	9.596128 8	-0.39143	6.10E- 09
YMR164 C	YPR082 C	14	0.79078 6	Negative shift 1	9.148612	-	7.40E- 04
YMR164 C	YKL008 C	14	0.78633 9	Negative	9.572666 6	-0.48135	4.80E- 04
YBR049 C	YGL089 C	14	0.78623 1	Negative	10.32126 7	-0.24474	7.40E- 04
YBL021 C	YLL027 W	14	0.72055 7	Positive	9.248479	0.36266	2.70E- 11

YKL043 W	YKL063 C	14	0.72003 3	Positive	9.714021 6	0.55578	1.70E- 04
YMR043 W	YKL058 W	14	0.70879 8	Positive	8.660197 2	-0.13215	7.40E- 04
YMR043 W	YLR189 C	14	0.70342 9	Negative shift 1	8.366505 5	-0.27692	1.90E- 06
YBL021 C	YNL009 W	14	0.70217 1	Negative	10.23064 5	-0.59078	1.70E- 04
YBR182 C	YBL037 W	14	0.60968 3	Positive	9.193838 6	0.45403	1.40E- 04
YER111 C	YOR31 5W	14	0.48810 2	Negative shift 1	11.87096 4	-0.55708	1.20E- 05
YKL112 W	YDL012 C	14	0.47158 5	Negative shift 1	10.06697 5	-0.21901	3.20E- 04
YKL043 W	YMR02 0W	14	0.43634 4	Negative shift 1	11.36833 2	-0.48792	3.90E- 07
YER111 C	YOL019 W	13	0.93845 8	Positive	10.94639	0.54274	8.60E- 04
YNL027 W	YGL038 C	12	0.97318 6	Negative	10.51321 3	-0.5013	8.60E- 04
YPR065 W	YML056 C	12	0.94203 9	Positive shift 2	8.324737 5	-0.39056	8.50E- 04
YDL170 W	YEL023 C	12	0.93304 7	Positive	9.912442 4	0.50061	5.80E- 04
YMR043 W	YML053 C	12	0.92586 3	Positive	10.04718 3	0.58961	1.20E- 06
YPL049 C	YGR01 4W	12	0.92270 6	Negative shift 1	8.310545 2	0.26673	3.60E- 04
YLR131 C	YML007 W	12	0.91955 5	Positive shift 1	10.98320 8	0.23229	5.30E- 04

YLR131 C	YJR147 W	11	0.95057 9	Negative shift 2	7.648163 7	0.26497	2.00E- 05
YPR104 C	YBR126 C	10	0.98278	Negative shift 1	10.32143 9	-0.60714	4.20E- 04
YMR042 W	YDR43 4W	10	0.97825 6	Negative	8.915609 6	-0.52445	9.20E- 04
YKL112 W	YJL111 W	10	0.96758 6	Negative	7.201133 2	-0.20003	2.30E- 04
YOR028 C	YOL116 W	10	0.95591 4	Negative shift 1	8.170286 4	-0.32806	9.60E- 04
YDL056 W	YER111 C	14	0.74252 1	Negative	13.86825 2	-0.78333	3.60E- 05
YNL068 C	YPL117 C	13	0.92780 5	Negative	12.28981 3	-0.72293	2.30E- 05
YDR207 C	YDR28 5W	15	0.78350 1	Negative shift 1	8.769669	-0.15794	
YML027 W	YKL091 C	15	0.71825 3	Negative	10.00014 5	-0.58824	
YGL035 C	YBR050 C	15	0.62789 8	Positive	8.348476 8	0.33049	
YDR501 W	YDL112 W	15	0.47143 5	Positive	8.285615 4	0.29687	
YDR207 C	YGR15 7W	14	0.78277 6	Negative	10.26711 1	-0.58183	
YGL096 W	YAL058 W	14	0.78003	Positive	10.54976 1	0.61108	
YGL013 C	YDR40 6W	14	0.77768 7	Positive shift 1	8.929366 3	0.26907	
YBR049 C	YGL026 C	14	0.75888 8	Negative	11.51836 1	-0.13874	

YML027 W	YFR011 C	14	0.72331 2	Negative	9.568071 9	-0.45911	
YER040 W	YJR152 W	14	0.65830 5	Positive shift 1	8.225701 6	- 0.062523	
YIL122 W	YBR161 W	14	0.60161 7	Negative	9.016046 6	-0.41813	
YML027 W	YML006 C	14	0.59789 4	Positive	7.848425 6	0.22884	
YOR344 C	YBR195 C	14	0.58462 4	Negative shift 1	8.403437 7	0.12185	
YIL122 W	YGR01 5C	14	0.57232 3	Negative	10.6481	-0.59923	
YOR372 C	YGL226 W	14	0.57127 7	Positive shift 1	9.990383	0.29061	
YER040 W	YFL021 W	14	0.55537	Positive	10.96989 8	0.64529	
YOR344 C	YAR073 W	14	0.54508 1	Negative shift 1	5.513845 9	0.22707	
YBL021 C	YOR37 5C	14	0.48016 3	Negative shift 1	9.176307 5	0.13332	
YOR372 C	YCL028 W	14	0.48002 3	Negative	10.32776 5	-0.55964	
YOR372 C	YDR47 1W	14	0.40490 9	Negative	7.180547 3	-0.28519	
YGL096 W	YDL057 W	13	0.91877 8	Positive shift 1	10.75116 3	0.61501	
YLR183 C	YNR03 9C	12	0.94848 3	Negative	10.72115 8	-0.59372	
YDL106 C	YGL234 W	11	0.96057 7	Negative shift 1	11.49753 3	-0.3548	

YDR451 C	YJL115 W	11	0.95627 6	Positive shift 2	11.79978 6	-0.11328	
YDR501 W	YDR32 8C	11	0.95196 2	Positive shift 1	8.609263 3	-0.11041	
YDR123 C	YNR01 6C	11	0.94582 4	Negative	8.611166 9	0.012985	
YGL096 W	YOL077 C	11	0.94557 8	Negative	8.226414 2	-0.41821	
YKL038 W	YGL062 W	10	0.95499 5	Positive shift 1	10.79755 7	-0.25253	
YLR183 C	YCL027 W	10	0.95234 4	Negative shift 1	9.049978 8	0.19728	
YML007 W	YHR00 8C	10	0.95026 5	Positive shift 1	9.143307 9	0.28109	
YDR451 C	YGL114 W	15	0.59454 9	Negative shift 1	12.30028 6	-0.29749	
YDR207 C	YAL054 C	14	0.66963 1	Negative	12.17364 2	-0.70959	
YDR123 C	YBR093 C	11	0.95808 6	Negative	13.90686 1	-0.81805	
YLR183 C	YPR139 C	11	0.93834 2	Negative shift 2	12.14910 1	0.34538	

Note: There are some blank in the column p-value of GWLA because the corresponding interactions are found by the other methods and from regulatory interaction collection dataset.

Table S6. Normalized expression value at each time point in the gene RCS1 and GCN3 in Fig.4A (in the text).

Time point	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
------------	---	---	---	---	---	---	---	---	---	----	----	----	----	----	----	----	----

GCN3	RCS1
-1.101	-0.37976
-1.5485	-1.4387
0.93343	0.4742
0.39092	-1.1312
0.52655	-0.31144
-1.5621	0.54252
-1.9825	-1.3362
1.0419	-0.5164
0.17392	-0.7555
0.54011	0.26925
-0.62628	-0.68719
0.59436	-0.27729
-0.12446	1.9089
-0.24652	1.294
0.41805	-0.48224
1.3403	1.1915
1.2318	1.6356

Pearson correlation coefficient (PCC) is 0.40465 according to $\frac{1}{17} \sum_{i=1}^{17} X_i Y_i$ because here the expression levels are normalized to X_i, Y_i in the “z score” fashion (X represents RCS1).

The results of local clustering (LC) according to the algorithm (Qian, et al., [8]) are

	max_score	startx	starty	len	relationship
	6.8791	17	17	17	1

The max_score is the final score of LC; 1 means positive relationship (details in Qian, et al., [8]).

The maximal local alignment of expression change trend between the two genes is 15 and cc is 0.70 according to the algorithm of TC in the supplementary materials and text.

systematic	systematic	gene1	gene2	sc	cc	score of locpcc		
YBL093C	YBR079C	ROX3	RPG1		16	0,86	9,76	0,49
YOR341W	YOR224C	RPA190	RPB8		16	0,85	12,55	0,74
YLR175W	YOR361C	CBF5	PRT1		16	0,85	11,66	0,68
YPL211W	YPL093W	NIP7	NOG1		16	0,82	12,36	0,73
YDR036C	YDL051W	EHD3	LHP1		16	0,82	12,99	-0,76
YOR224C	YNL113W	RPB8	RPC19		16	0,8	11,96	0,7
YMR284W	YFL022C	YKU70	FRS2		16	0,77	10,13	-0,56
YNL021W	YPR010C	HDA1	RPA135		16	0,73	10,66	-0,62
YNL110C	YKL009W	NOP15	MRT4		16	0,69	10,59	0,62
YDR201W	YKL179C	SPC19	COY1		16	0,67	12,54	0,74
YGL019W	YIL094C	CKB1	LYS12		16	0,66	10,41	-0,28
YDL014W	YBR079C	NOP1	RPG1		15	0,92	12,81	0,74
YGL245W	YBL045C	GUS1	COR1		15	0,92	12,32	0,71
YMR146C	YDR429C	TIF34	TIF35		15	0,91	11,95	0,7
YPR088C	YPL243W	SRP54	SRP68		15	0,89	11,98	0,7
YNR016C	YLR180W	ACC1	SAM1		15	0,89	10,38	0,58
YBR142W	YFR001W	MAK5	LOC1		15	0,89	12,31	0,67
YBR025C	YGL105W	YBR025C	ARC1		15	0,88	10,54	0,62
YPR187W	YNR016C	RPO26	ACC1		15	0,88	7,23	0,35
YER155C	YER082C	BEM2	UTP7		15	0,87	12,7	0,75
YMR229C	YNL175C	RRP5	NOP13		15	0,86	10,75	0,63
YDR023W	YDR350C	SES1	TCM10		15	0,86	11,4	-0,36
YNR016C	YJL138C	ACC1	TIF2		15	0,86	9,86	0,35
YBR142W	YLR106C	MAK5	MDN1		15	0,86	11,18	0,64
YDL213C	YGR081C	NOP6	SLX9		15	0,85	8,24	0,45
YDR347W	YDR041W	MRP1	RSM10		15	0,85	12,01	0,65
YGL008C	YBR011C	PMA1	IPP1		15	0,84	12,4	0,71
YPL043W	YOL041C	NOP4	NOP12		15	0,83	7,52	0,38
YOR361C	YMR308C	PRT1	PSE1		15	0,82	10,29	0,6
YNL147W	YDR240C	LSM7	SNU56		15	0,82	12,12	-0,71
YHR085W	YKL014C	IPI1	URB1		15	0,82	11,43	0,67
YFL039C	YLR449W	YFL039C	FPR4		15	0,81	12,74	0,75
YLL036C	YGL049C	PRP19	TIF4632		15	0,8	8,25	0,19
YAL043C	YDR382W	PTA1	RPP2B		15	0,8	10,9	0,6
YBR009C	YLR106C	HHF1	MDN1		15	0,8	9,13	0,34
YML065W	YLL004W	ORC1	ORC3		15	0,79	10,76	0,63
YLR449W	YLR221C	FPR4	RSA3		15	0,79	10,69	0,62
YLR175W	YGR081C	CBF5	SLX9		15	0,79	9,85	0,27
YNL030W	YLR150W	HHF2	STM1		15	0,79	12,46	0,73
YBR267W	YKL014C	REI1	URB1		15	0,79	12,54	0,72
YMR308C	YGR081C	PSE1	SLX9		15	0,78	9,95	0,59
YKL009W	YOR361C	MRT4	PRT1		15	0,78	11,11	0,62
YFR010W	YOL038W	UBP6	PRE6		15	0,78	12,26	0,72
YML092C	YPL020C	PRE8	ULP1		15	0,77	11,11	0,65
YDR299W	YBR247C	BFR2	ENP1		15	0,77	10,02	0,04
YNL030W	YGR081C	HHF2	SLX9		15	0,77	12,41	0,05
YML069W	YDR224C	POB3	HTB1		15	0,76	8,92	0,52
YNL030W	YLR221C	HHF2	RSA3		15	0,76	8,25	0,48
YCR046C	YLR439W	IMG1	MRPL4		15	0,76	12,23	0,67
YMR049C	YDL098C	ERB1	SNU23		15	0,76	10,34	-0,27
YNL110C	YPL012W	NOP15	RRP12		15	0,76	11,79	0,66
YLL034C	YLL008W	RIX7	DRS1		15	0,76	12,24	0,67
YKR059W	YGR081C	TIF1	SLX9		15	0,75	10,81	0,12

YLR276C	YFL002C	DBP9	SPB4	15	0,75	11,27	0,66
YDR155C	YJR090C	CPR1	GRR1	15	0,74	9,13	0,46
YKL172W	YKL009W	EBP2	MRT4	15	0,74	9,85	0,58
YNL186W	YPL013C	UBP10	MRPS16	15	0,73	10,7	0,61
YLR418C	YML010W	CDC73	SPT5	15	0,73	11,99	0,71
YDR405W	YNL236W	MRP20	SIN4	15	0,73	8,22	0,38
YMR049C	YML124C	ERB1	TUB3	15	0,72	9,28	0,54
YLR175W	YNL317W	CBF5	PFS2	15	0,72	11,53	-0,6
YOR212W	YOR281C	STE4	PLP2	15	0,72	8,11	0,37
YDL208W	YLR106C	NHP2	MDN1	15	0,72	8,18	0,35
YMR049C	YLR249W	ERB1	YEF3	15	0,71	9,3	0,54
YLR439W	YDR237W	MRPL4	MRPL7	15	0,71	7,18	0,29
YDR405W	YLR439W	MRP20	MRPL4	15	0,71	8,33	0,31
YDL060W	YDR071C	TSR1	PAA1	15	0,71	8,93	0,04
YPL082C	YIL126W	MOT1	STH1	15	0,71	9,7	0,57
YNL317W	YKR059W	PFS2	TIF1	15	0,71	11,18	-0,59
YJL138C	YEL026W	TIF2	SNU13	15	0,71	11,04	0,65
YER139C	YOR206W	YER139C	NOC2	15	0,7	10,37	0,17
YKL009W	YPL012W	MRT4	RRP12	15	0,7	7,94	0,34
YLR186W	YGL171W	EMG1	ROK1	15	0,7	12,38	0,73
YDL213C	YLR003C	NOP6	YLR003C	15	0,7	6,61	0,36
YNL025C	YER022W	SSN8	YER022W	15	0,7	11,54	0,53
YER082C	YKL143W	UTP7	LTV1	15	0,69	8,77	0,52
YHR027C	YBR039W	RPN1	ATP3	15	0,69	10,12	0,6
YDR448W	YIL077C	ADA2	YIL077C	15	0,69	7,15	0,32
YLR044C	YML124C	PDC1	TUB3	15	0,69	9,97	0,4
YNL110C	YLR449W	NOP15	FPR4	15	0,69	9,42	0,55
YNL308C	YLL050C	KRI1	COF1	15	0,69	8,52	0,18
YDL166C	YOR206W	FAP7	NOC2	15	0,69	8,29	0
YDR405W	YFR051C	MRP20	RET2	15	0,69	9,53	-0,53
YNL244C	YDR427W	SUI1	RPN9	15	0,69	8,43	0,12
YDR023W	YDL192W	SES1	ARF1	15	0,68	11,4	0,67
YMR116C	YMR049C	ASC1	ERB1	15	0,68	9,06	0,51
YIL061C	YER172C	SNP1	BRR2	15	0,67	8,33	-0,4
YGR091W	YDR416W	PRP31	SYF1	15	0,67	11,09	0,64
YNL113W	YDL192W	RPC19	ARF1	15	0,66	8,34	0,47
YMR116C	YMR146C	ASC1	TIF34	15	0,65	8,93	0,52
YBR274W	YDR217C	CHK1	RAD9	15	0,64	10,08	0,26
YER029C	YCL037C	SMB1	SRO9	15	0,63	10,41	0,61
YKL056C	YDR224C	YKL056C	HTB1	15	0,63	6,95	0,38
YER012W	YEL030W	PRE1	ECM10	15	0,63	8,43	-0,21
YPL218W	YNL113W	SAR1	RPC19	15	0,63	8,24	0,48
YLR268W	YLR026C	SEC22	SED5	15	0,63	7,68	-0,42
YOL086C	YKR008W	ADH1	RSC4	15	0,62	11,41	-0,6
YOR341W	YDL192W	RPA190	ARF1	15	0,62	8,68	0,5
YMR309C	YDL007W	NIP1	RPT2	15	0,6	11,21	-0,41
YMR012W	YOR086C	CLU1	TCB1	15	0,6	9,8	0,53
YJL138C	YLR423C	TIF2	ATG17	15	0,6	10,29	-0,6
YOL010W	YLR003C	RCL1	YLR003C	15	0,57	8,84	-0,1
YAL043C	YOR344C	PTA1	TYE7	15	0,56	9,38	0,35
YJR050W	YKL173W	ISY1	SNU114	15	0,56	9,93	-0,15
YDL160C	YER029C	DHH1	SMB1	15	0,55	7,42	0,44
YOR159C	YOR319W	SME1	HSH49	15	0,55	9,85	0,58
YHR076W	YLR304C	PTC7	ACO1	15	0,55	8,81	0,31

YDR382W	YDR170C	RPP2B	SEC7	15	0,55	8,04	-0,34
YNL030W	YLR003C	HHF2	YLR003C	15	0,55	9,61	-0,12
YLR180W	YLR003C	SAM1	YLR003C	15	0,53	11,86	0,37
YKL009W	YOR341W	MRT4	RPA190	15	0,52	8,57	0,49
YPL247C	YLR109W	YPL247C	AHP1	15	0,52	12,04	0,7
YLR241W	YLR109W	YLR241W	AHP1	15	0,51	11,19	0,66
YPR179C	YOR323C	HDA3	PRO2	15	0,5	6,9	0,23
YOR344C	YGR156W	TYE7	PTI1	15	0,49	7,62	-0,34
YGL171W	YGL252C	ROK1	RTG2	15	0,48	8,88	-0,44
YGL068W	YDR101C	MNP1	ARX1	15	0,43	8,07	-0,44
YGL011C	YGR090W	SCL1	UTP22	15	0,41	8,67	-0,47
YKR081C	YLR233C	RPF2	EST1	15	0,35	8,95	0,06
YOR159C	YMR268C	SME1	PRP24	14	0,96	9,74	0,17
YDL185W	YDR226W	TFP1	ADK1	14	0,94	12,5	0,62
YOR361C	YEL026W	PRT1	SNU13	14	0,94	10,17	0,59
YNL308C	YNL175C	KRI1	NOP13	14	0,93	12,29	0,72
YAL033W	YBL018C	POP5	POP8	14	0,93	12,05	0,65
YPL154C	YOR332W	PEP4	VMA4	14	0,92	12,79	0,68
YMR128W	YDR324C	ECM16	UTP4	14	0,91	8,15	0,33
YNL002C	YMR310C	RLP7	YMR310C	14	0,91	8,51	0,28
YFL039C	YNL079C	YFL039C	TPM1	14	0,91	12,73	0,75
YGR282C	YGR192C	BGL2	TDH3	14	0,9	12,98	0,76
YPR016C	YKL172W	TIF6	EBP2	14	0,9	10,9	0,64
YOR341W	YNL308C	RPA190	KRI1	14	0,9	12,58	0,74
YNL030W	YOR361C	HHF2	PRT1	14	0,9	11,66	0,68
YGR266W	YEL051W	YGR266W	VMA8	14	0,9	8,02	-0,18
YPR016C	YLR449W	TIF6	FPR4	14	0,89	12,61	0,71
YNL308C	YKL172W	KRI1	EBP2	14	0,89	9,17	0,54
YOR061W	YOL021C	CKA2	DIS3	14	0,89	10,63	0,63
YDL031W	YGR103W	DBP10	NOP7	14	0,89	12,45	0,73
YOR224C	YPR187W	RPB8	RPO26	14	0,89	10,98	0,6
YBR025C	YDR023W	YBR025C	SES1	14	0,88	11	0,6
YBR025C	YBR039W	YBR025C	ATP3	14	0,88	9,83	0,55
YLR074C	YLR221C	BUD20	RSA3	14	0,88	9,03	0,53
YPR016C	YHR089C	TIF6	GAR1	14	0,88	12,13	0,69
YDL031W	YLR106C	DBP10	MDN1	14	0,88	12,51	0,74
YNL040W	YPL043W	YNL040W	NOP4	14	0,87	7,52	-0,37
YNL175C	YBR009C	NOP13	HHF1	14	0,87	12,42	0,71
YHR058C	YBR193C	MED6	MED8	14	0,87	11,23	0,65
YKL172W	YNL175C	EBP2	NOP13	14	0,87	12,62	0,74
YHR216W	YIL035C	IMD2	CKA1	14	0,87	8,93	0,45
YMR049C	YMR310C	ERB1	YMR310C	14	0,87	12,43	0,73
YBR011C	YEL034W	IPP1	HYP2	14	0,87	12,69	0,74
YPR086W	YPL093W	SUA7	NOG1	14	0,86	8,21	0,35
YBR079C	YDR382W	RPG1	RPP2B	14	0,86	11,49	0,67
YEL051W	YFL045C	VMA8	SEC53	14	0,86	10,88	0,64
YGR094W	YBR011C	VAS1	IPP1	14	0,86	12,6	0,71
YER029C	YER006W	SMB1	NUG1	14	0,86	11,72	0,69
YOR078W	YGR081C	BUD21	SLX9	14	0,86	7,38	0,4
YEL026W	YLR027C	SNU13	AAT2	14	0,86	11,53	0,62
YPL012W	YNL308C	RRP12	KRI1	14	0,86	12,67	0,75
YPR016C	YBR142W	TIF6	MAK5	14	0,86	12,22	0,71
YNL175C	YDR224C	NOP13	HTB1	14	0,86	11,89	0,69
YOR039W	YBL003C	CKB2	HTA2	14	0,86	7,71	0,43

YHR060W	YBR260C	VMA22	RGD1	14	0,86	7,99	-0,19
YNR016C	YJR064W	ACC1	CCT5	14	0,86	10,05	0,3
YNL308C	YOL139C	KRI1	CDC33	14	0,85	9,11	0,51
YPL228W	YNL030W	CET1	HHF2	14	0,85	8,98	-0,42
YPR016C	YDL213C	TIF6	NOP6	14	0,85	12,64	0,71
YNL023C	YBR081C	FAP1	SPT7	14	0,85	9,54	0,24
YGL111W	YOR272W	NSA1	YTM1	14	0,85	12,21	0,72
YNL110C	YML124C	NOP15	TUB3	14	0,85	11,14	0,65
YDL014W	YAL043C	NOP1	PTA1	14	0,85	9,57	0,5
YBR065C	YCR046C	ECM2	IMG1	14	0,85	10,68	-0,63
YJL138C	YDR224C	TIF2	HTB1	14	0,85	9,28	0,55
YGL049C	YMR310C	TIF4632	YMR310C	14	0,85	12,62	0,74
YDL185W	YBR039W	TFP1	ATP3	14	0,84	10,64	0,5
YMR049C	YDL213C	ERB1	NOP6	14	0,84	11,41	0,65
YNL132W	YHR148W	KRE33	IMP3	14	0,84	11,42	0,67
YLR196W	YLR397C	PWP1	AFG2	14	0,84	11,85	0,7
YBR079C	YGR104C	RPG1	SRB5	14	0,84	12,84	-0,76
YDR405W	YDR237W	MRP20	MRPL7	14	0,84	8,5	0,31
YDR324C	YDR365C	UTP4	ESF1	14	0,84	10,96	0,64
YML124C	YLR249W	TUB3	YEF3	14	0,84	9,99	0,59
YOL041C	YPR137W	NOP12	RRP9	13	0,96	12,98	0,75
YDL084W	YCL011C	SUB2	GBP2	13	0,96	12,39	0,73
YDR322W	YMR225C	MRPL35	MRPL44	13	0,96	8,89	0,49
YBL047C	YDL220C	EDE1	CDC13	13	0,95	9,28	-0,55
YFR028C	YDL160C	CDC14	DHH1	13	0,95	12,27	0,72
YLR175W	YER029C	CBF5	SMB1	13	0,95	11,42	0,67
YDL213C	YMR310C	NOP6	YMR310C	13	0,94	8,13	0,38
YNL055C	YLL050C	POR1	COF1	13	0,94	12,19	0,68
YBL004W	YLR222C	UTP20	UTP13	13	0,94	12,8	0,74
YMR226C	YDL185W	YMR226C	TFP1	13	0,94	11,71	0,61
YKL172W	YGR103W	EBP2	NOP7	13	0,94	12,87	0,76
YBR079C	YBR143C	RPG1	SUP45	13	0,94	12,88	0,76
YMR226C	YBR009C	YMR226C	HHF1	13	0,94	12,64	0,7
YGL195W	YGR234W	GCN1	YHB1	13	0,94	9,16	0,54
YNL005C	YGR013W	MRP7	SNU71	13	0,93	7,29	0,05
YLR197W	YML124C	SIK1	TUB3	13	0,93	11,16	0,66
YBR025C	YCL035C	YBR025C	GRX1	13	0,93	11,41	0,67
YNL132W	YOR341W	KRE33	RPA190	13	0,93	12,75	0,75
YDR188W	YDR381W	CCT6	YRA1	13	0,93	10	0,59
YPR010C	YGR103W	RPA135	NOP7	13	0,93	12,45	0,71
YMR214W	YDR238C	SCJ1	SEC26	13	0,93	7,89	-0,27
YJR072C	YGR040W	NPA3	KSS1	13	0,93	11,43	-0,67
YBR211C	YJR135C	AME1	MCM22	13	0,93	9,13	0,4
YNL079C	YDL072C	TPM1	YET3	13	0,93	11,28	0,66
YGR282C	YLR249W	BGL2	YEF3	13	0,92	11,83	0,69
YCL037C	YER006W	SRO9	NUG1	13	0,92	11,21	0,52
YJL122W	YDR224C	YJL122W	HTB1	13	0,92	10,37	0,29
YBR011C	YGL105W	IPP1	ARC1	13	0,92	8,79	0,5
YGR103W	YPL012W	NOP7	RRP12	13	0,92	11,97	0,62
YJR043C	YNL102W	POL32	POL1	12	0,97	12,06	0,71
YDL185W	YPR187W	TFP1	RPO26	12	0,97	11,87	0,61
YGL245W	YLL050C	GUS1	COF1	12	0,97	12,76	0,72
YNR035C	YOR361C	ARC35	PRT1	12	0,97	9	0,45
YKL172W	YER002W	EBP2	NOP16	12	0,96	10,9	0,64

YKR048C	YLR002C	NAP1	NOC3	12	0,96	8,95	-0,08
YDL185W	YOL139C	TFP1	CDC33	12	0,96	10,71	0,47
YNL250W	YDR032C	RAD50	PST2	12	0,96	11,15	-0,62
YGR103W	YLR106C	NOP7	MDN1	12	0,96	12,46	0,68
YNL002C	YER002W	RLP7	NOP16	12	0,96	11,15	0,63
YGR128C	YLR222C	UTP8	UTP13	12	0,96	12,98	0,74
YNL094W	YDR155C	APP1	CPR1	12	0,95	12,22	-0,71
YER043C	YER091C	SAH1	MET6	12	0,95	12,15	0,71
YPR103W	YGL245W	PRE2	GUS1	12	0,95	11,73	0,55
YDR443C	YNL189W	SSN2	SRP1	12	0,95	8,05	-0,38
YNL207W	YMR128W	RIO2	ECM16	12	0,95	8,87	0,5
YGL245W	YNL135C	GUS1	FPR1	12	0,95	12,43	0,71
YPL037C	YLR027C	EGD1	AAT2	12	0,95	12,86	0,76
YER146W	YIL048W	LSM5	NEO1	12	0,95	9,02	0,17
YNL021W	YOR361C	HDA1	PRT1	12	0,94	8,45	-0,38
YGR090W	YLR216C	UTP22	CPR6	11	0,98	9,41	0,12
YNL110C	YDL051W	NOP15	LHP1	11	0,98	10,07	0,47
YKL144C	YKL085W	RPC25	MDH1	11	0,98	12,19	-0,05
YLR430W	YGR040W	SEN1	KSS1	11	0,98	11,73	-0,54
YGL245W	YML123C	GUS1	PHO84	11	0,98	12,36	0,7
YGL245W	YPR187W	GUS1	RPO26	11	0,98	12,82	0,72
YMR027W	YGR013W	YMR027W	SNU71	11	0,97	7,23	0,37
YKL210W	YPL037C	UBA1	EGD1	11	0,97	9,4	0,34
YCR057C	YLR180W	PWP2	SAM1	11	0,97	7,32	-0,11
YMR290C	YPL093W	HAS1	NOG1	11	0,97	10,68	0,63
YPL082C	YGL241W	MOT1	KAP114	11	0,97	8,77	0,5
YBR162C	YMR214W	TOS1	SCJ1	11	0,97	9,08	-0,38
YOR261C	YML124C	RPN8	TUB3	11	0,97	8,04	0,13
YML123C	YGR040W	PHO84	KSS1	11	0,97	11,6	-0,67
YLR234W	YDL225W	TOP3	SHS1	11	0,97	11,68	0,26
YMR309C	YOR272W	NIP1	YTM1	11	0,97	10,05	0,58
YPL011C	YGR094W	TAF3	VAS1	11	0,96	11,66	0,69
YBL074C	YOL139C	AAR2	CDC33	11	0,96	12,12	-0,71
YPL001W	YER021W	HAT1	RPN3	11	0,96	12,15	-0,71
YEL060C	YEL030W	PRB1	ECM10	11	0,96	6,67	0,39
YML028W	YGL245W	TSA1	GUS1	11	0,96	12,93	0,71
YGL245W	YBR009C	GUS1	HHF1	11	0,96	11,57	0,64
YMR116C	YGR254W	ASC1	ENO1	11	0,96	12,58	0,74
YKR048C	YJR064W	NAP1	CCT5	11	0,96	10,08	0,52
YML069W	YBR009C	POB3	HHF1	11	0,96	12,58	0,73
YDL166C	YHR196W	FAP7	UTP9	11	0,96	10,59	0,59
YOL004W	YMR236W	SIN3	TAF9	11	0,96	11,18	-0,66
YMR246W	YHL030W	FAA4	ECM29	11	0,96	11,11	-0,38
YGR193C	YPL140C	PDX1	MKK2	11	0,96	8,11	-0,15
YMR276W	YIL075C	DSK2	RPN2	11	0,96	10,21	0,56
YDL097C	YJR064W	RPN6	CCT5	11	0,96	7,06	0,19
YBL003C	YNL075W	HTA2	IMP4	11	0,96	8,47	-0,01
YJR138W	YBL023C	IML1	MCM2	10	0,99	10,85	-0,01
YER091C	YCL035C	MET6	GRX1	10	0,99	7,3	0,07
YJR065C	YOR361C	ARP3	PRT1	10	0,98	7,22	-0,02
YMR049C	YAL035W	ERB1	FUN12	10	0,98	12,64	0,7
YPR187W	YCL028W	RPO26	RNQ1	10	0,98	12,86	0,76
YOR332W	YLR109W	VMA4	AHP1	10	0,98	11,67	0,53
YER077C	YKL172W	YER077C	EBP2	10	0,98	10,91	0,2

YIR009W	YOR017W	MSL1	PET127	10	0,98	7,79	0,46
YLL040C	YMR105C	VPS13	PGM2	10	0,98	11,63	0,42
YMR226C	YML022W	YMR226C	APT1	10	0,97	11,97	0,64
YNL175C	YLR397C	NOP13	AFG2	10	0,97	7,72	0,21
YCL059C	YBR084W	KRR1	MIS1	10	0,97	12,58	0,62
YJR138W	YBL088C	IML1	TEL1	10	0,97	7,68	-0,2
YLR109W	YLR304C	AHP1	ACO1	10	0,97	11,06	0,53
YER091C	YNL135C	MET6	FPR1	10	0,97	11,66	0,69
YER091C	YDR155C	MET6	CPR1	10	0,97	9	0,53
YHR196W	YLR222C	UTP9	UTP13	10	0,97	11	0,6
YDR462W	YML049C	MRPL28	RSE1	10	0,97	8,6	-0,17
YBR154C	YGR063C	RPB5	SPT4	10	0,97	11,44	0,67
YER174C	YBR085W	GRX4	AAC3	10	0,97	11,19	-0,12
YML092C	YGL245W	PRE8	GUS1	10	0,97	10,74	0,55
YFR019W	YMR112C	FAB1	MED11	10	0,97	11,88	-0,29
YPR180W	YDL115C	AOS1	IWR1	10	0,97	9,02	0,2
YNL132W	YDR168W	KRE33	CDC37	10	0,97	7,5	-0,27
YGR048W	YOR116C	UFD1	RPO31	10	0,97	10,7	-0,63
YML016C	YER054C	PPZ1	GIP2	10	0,97	8,94	-0,31
YBL072C	YIL069C	RPS8A	RPS24B	16	0,81	12,22	0,71
YHR052W	YER043C	CIC1	SAH1	16	0,8	10,85	0,64
YBR189W	YIL069C	RPS9B	RPS24B	16	0,79	11,81	0,69
YDL083C	YIL069C	RPS16B	RPS24B	16	0,75	11,45	0,66
YDR450W	YIL069C	RPS18A	RPS24B	16	0,75	10,94	0,64
YGL200C	YIL109C	EMP24	SEC24	16	0,71	11,43	-0,59
YCR031C	YIL069C	RPS14A	RPS24B	15	0,89	11,37	0,6
YKR081C	YGL049C	RPF2	TIF4632	15	0,85	12,43	0,73
YKL212W	YOL123W	SAC1	HRP1	15	0,85	10,35	0,44
YBR181C	YIL069C	RPS6B	RPS24B	15	0,84	11,57	0,68
YOR234C	YOR312C	RPL33B	RPL20B	15	0,83	12,74	0,75
YLR185W	YOR312C	RPL37A	RPL20B	15	0,82	12,69	0,75
YLR406C	YMR194W	RPL31B	RPL36A	15	0,81	12,84	0,72
YLR075W	YOR312C	RPL10	RPL20B	15	0,8	12,49	0,73
YDR064W	YIL069C	RPS13	RPS24B	15	0,79	10,42	0,61
YMR308C	YEL055C	PSE1	POL5	15	0,79	8,9	0,51
YJL014W	YJL111W	CCT3	CCT7	15	0,79	10,84	0,03
YJR141W	YLR295C	YJR141W	ATP14	15	0,79	9,8	-0,51
YMR229C	YDL213C	RRP5	NOP6	15	0,77	8,95	0,45
YKL116C	YLR389C	PRR1	STE23	15	0,77	9,18	0,42
YKL006W	YOR312C	RPL14A	RPL20B	15	0,77	11,69	0,68
YMR194W	YOR312C	RPL36A	RPL20B	15	0,76	11,13	0,65
YMR236W	YNL308C	TAF9	KRI1	15	0,75	10,09	0,23
YLR427W	YML056C	MAG2	IMD4	15	0,75	7,9	0,32
YMR205C	YDR158W	PFK2	HOM2	15	0,74	11,61	0,68
YCL043C	YBR010W	PDI1	HHT1	15	0,74	10,71	0,62
YCR002C	YGR061C	CDC10	ADE6	15	0,73	8,63	0,45
YNL308C	YMR116C	KRI1	ASC1	15	0,73	8,87	0,42
YER006W	YER120W	NUG1	SCS2	15	0,72	10,89	0,64
YJL109C	YKL012W	UTP10	PRP40	15	0,7	6,34	-0,31
YFR002W	YGL172W	NIC96	NUP49	15	0,69	8,45	0,22
YMR123W	YPL169C	PKR1	MEX67	15	0,69	7,43	0,23
YBL076C	YBR249C	ILS1	ARO4	15	0,68	11,44	0,67
YMR229C	YOL139C	RRP5	CDC33	15	0,68	6,19	0,24
YGR005C	YKL009W	TFG2	MRT4	15	0,65	6,29	0,22

YDR429C	YJL026W	TIF35	RNR2	15	0,64	9,52	0,02
YJR070C	YOR362C	LIA1	PRE10	15	0,64	7,28	0,28
YBL076C	YGR061C	ILS1	ADE6	15	0,64	9,84	0,36
YMR286W	YNL185C	MRPL33	MRPL19	15	0,63	11,27	0,66
YGR214W	YJL136C	RPS0A	RPS21B	15	0,56	12,08	0,66
YPL235W	YER179W	RVB2	DMC1	15	0,55	10,05	0,3
YNL308C	YDR174W	KRI1	HMO1	15	0,53	7,79	-0,23
YKL056C	YPL204W	YKL056C	HRR25	15	0,38	11,54	0,14
YBL076C	YAL012W	ILS1	CYS3	14	0,95	12,9	0,76
YGR148C	YHL001W	RPL24B	RPL14B	14	0,93	12,86	0,76
YGL076C	YGL135W	RPL7A	RPL1B	14	0,92	11,33	0,66
YDL081C	YGR209C	RPP1A	TRX2	14	0,92	12,29	0,64
YDR418W	YPR043W	RPL12B	RPL43A	14	0,91	12,72	0,75
YDR382W	YGL103W	RPP2B	RPL28	14	0,91	12,43	0,64
YDL061C	YIL069C	RPS29B	RPS24B	14	0,9	12,09	0,71
YDR037W	YEL034W	KRS1	HYP2	14	0,9	12,58	0,73
YBR010W	YNL030W	HHT1	HHF2	14	0,89	11,85	0,69
YJR091C	YKL002W	JSN1	DID4	14	0,88	7,16	-0,22
YKR096W	YKR060W	YKR096W	UTP30	14	0,88	9,13	-0,48
YBR010W	YNL031C	HHT1	HHT2	14	0,88	11,79	0,69
YCR031C	YGR118W	RPS14A	RPS23A	14	0,88	12,68	0,75
YBR048W	YIL069C	RPS11B	RPS24B	14	0,88	10,75	0,63
YDR418W	YPL220W	RPL12B	RPL1A	14	0,88	12,67	0,74
YOR017W	YMR290C	PET127	HAS1	14	0,88	9,63	-0,48
YPR016C	YOL139C	TIF6	CDC33	14	0,87	12,24	0,71
YGR027C	YGR214W	RPS25A	RPS0A	14	0,87	12,3	0,72
YPR016C	YBR009C	TIF6	HHF1	14	0,87	10,77	0,58
YMR229C	YEL055C	RRP5	POL5	14	0,86	12,85	0,76
YLR447C	YNL175C	VMA6	NOP13	14	0,86	10,12	0,56
YHR052W	YER165W	CIC1	PAB1	14	0,86	12,61	0,74
YNL301C	YPL143W	RPL18B	RPL33A	14	0,86	12,9	0,76
YDR224C	YNL031C	HTB1	HHT2	14	0,86	12,67	0,74
YPL016W	YPL195W	SWI1	APL5	14	0,86	11,12	-0,64
YHR203C	YIL069C	RPS4B	RPS24B	14	0,86	12,7	0,74
YNL169C	YGR282C	PSD1	BGL2	14	0,86	8,85	-0,22
YDL060W	YAR073W	TSR1	IMD1	14	0,86	9,65	0,57
YBR009C	YNL031C	HHF1	HHT2	14	0,85	12,53	0,73
YPL012W	YDR091C	RRP12	RLI1	14	0,85	6,78	0,14
YPR016C	YEL034W	TIF6	HYP2	14	0,85	11,74	0,69
YGR234W	YFL045C	YHB1	SEC53	14	0,85	11,33	0,62
YDL130W	YDR399W	RPP1B	HPT1	14	0,85	11,7	0,69
YFL002C	YFR028C	SPB4	CDC14	14	0,84	9,67	0,48
YJR091C	YMR049C	JSN1	ERB1	14	0,84	8,77	-0,46
YMR241W	YOL123W	YHM2	HRP1	14	0,84	11,26	0,39
YDL191W	YGL076C	RPL35A	RPL7A	13	0,96	12,01	0,67
YDR418W	YKL180W	RPL12B	RPL17A	13	0,96	12,58	0,72
YDL130W	YGL103W	RPP1B	RPL28	13	0,95	12,74	0,68
YER117W	YGL076C	RPL23B	RPL7A	13	0,95	12,01	0,69
YKR028W	YKR048C	SAP190	NAP1	13	0,95	12,03	0,71
YGL147C	YPL143W	RPL9A	RPL33A	13	0,95	12,37	0,73
YDR012W	YGL076C	RPL4B	RPL7A	13	0,95	11,64	0,63
YIL125W	YIL034C	KGD1	CAP2	13	0,94	8,38	0,49
YJR136C	YMR047C	YJR136C	NUP116	13	0,94	10,43	-0,52
YBL030C	YGR282C	PET9	BGL2	13	0,94	12,76	0,74

YGL147C	YPR043W	RPL9A	RPL43A	13	0,94	12,38	0,73
YEL054C	YGL076C	RPL12A	RPL7A	13	0,93	12,08	0,71
YCR031C	YNL096C	RPS14A	RPS7B	13	0,93	12,82	0,75
YGL147C	YLR340W	RPL9A	RPP0	13	0,93	12,21	0,72
YDR418W	YPL079W	RPL12B	RPL21B	13	0,93	12,41	0,73
YGL076C	YPR043W	RPL7A	RPL43A	13	0,93	12,74	0,69
YLR293C	YGR282C	GSP1	BGL2	13	0,93	12,19	0,72
YBR173C	YMR314W	UMP1	PRE5	13	0,92	11,84	0,67
YGR214W	YOL040C	RPS0A	RPS15	13	0,92	12,81	0,75
YGL076C	YHL001W	RPL7A	RPL14B	12	0,97	12,34	0,71
YHR089C	YGR282C	GAR1	BGL2	12	0,97	12,84	0,75
YDL191W	YGL103W	RPL35A	RPL28	12	0,97	12,86	0,7
YBR031W	YGL076C	RPL4A	RPL7A	12	0,96	12,08	0,68
YHL001W	YMR142C	RPL14B	RPL13B	12	0,96	12,68	0,74
YBL092W	YGL147C	RPL32	RPL9A	12	0,96	12,75	0,75
YDR418W	YJL189W	RPL12B	RPL39	12	0,96	12,67	0,73
YCR031C	YLR264W	RPS14A	RPS28B	12	0,96	12,84	0,75
YDL130W	YHL001W	RPP1B	RPL14B	12	0,96	12,73	0,73
YER117W	YGL103W	RPL23B	RPL28	12	0,96	12,72	0,72
YBR191W	YGL076C	RPL21A	RPL7A	12	0,96	12,97	0,76
YBL087C	YOR312C	RPL23A	RPL20B	12	0,95	10,91	0,6
YHL033C	YPR043W	RPL8A	RPL43A	12	0,95	12,95	0,76
YDR012W	YHL033C	RPL4B	RPL8A	12	0,95	12,48	0,7
YDR418W	YKR094C	RPL12B	RPL40B	12	0,95	12,86	0,76
YOR049C	YOR302W	RSB1	YOR302W	12	0,95	9,3	-0,4
YNR054C	YOL139C	ESF2	CDC33	12	0,95	12,11	0,7
YBL087C	YGL147C	RPL23A	RPL9A	12	0,95	11,2	0,66
YDL081C	YHL001W	RPP1A	RPL14B	12	0,95	12,93	0,73
YBR191W	YGL147C	RPL21A	RPL9A	12	0,95	12,28	0,68
YGL076C	YIL018W	RPL7A	RPL2B	12	0,95	11,76	0,68
YIL008W	YLR288C	URM1	MEC3	12	0,95	10,78	-0,47
YHL001W	YLR061W	RPL14B	RPL22A	12	0,95	12,91	0,75
YBL072C	YGR214W	RPS8A	RPS0A	12	0,95	12,82	0,75
YDR447C	YER102W	RPS17B	RPS8B	12	0,95	11,6	0,62
YDL061C	YER102W	RPS29B	RPS8B	12	0,95	11,85	0,63
YDL081C	YGL147C	RPP1A	RPL9A	12	0,94	12,33	0,73
YCR031C	YHL015W	RPS14A	RPS20	12	0,94	12,23	0,69
YGL147C	YHL001W	RPL9A	RPL14B	12	0,94	12,72	0,75
YGL076C	YLR340W	RPL7A	RPP0	11	0,99	12,54	0,72
YBR191W	YGL103W	RPL21A	RPL28	11	0,98	12,84	0,69
YBR031W	YGL103W	RPL4A	RPL28	11	0,98	11,97	0,63
YBR025C	YNR016C	YBR025C	ACC1	11	0,97	9,86	0,43
YDR418W	YGL076C	RPL12B	RPL7A	11	0,97	10,37	0,52
YFL016C	YOL139C	MDJ1	CDC33	11	0,97	10,06	0,42
YIL133C	YMR121C	RPL16A	RPL15B	11	0,97	12,93	0,71
YGL103W	YGL147C	RPL28	RPL9A	11	0,97	10,59	0,6
YLR328W	YPR016C	NMA1	TIF6	11	0,97	7,58	-0,36
YGL147C	YIL133C	RPL9A	RPL16A	11	0,97	12,46	0,73
YGR267C	YGR155W	FOL2	CYS4	11	0,96	11,25	0,66
YGL076C	YNL069C	RPL7A	RPL16B	11	0,96	11,08	0,62
YGL076C	YPL143W	RPL7A	RPL33A	11	0,96	12,57	0,73
YHL033C	YMR121C	RPL8A	RPL15B	11	0,96	12,76	0,68
YJR006W	YLR103C	HYS2	CDC45	11	0,96	9,07	0,52
YBL027W	YGL147C	RPL19B	RPL9A	11	0,96	10,68	0,63

YCR031C	YER131W	RPS14A	RPS26B	11	0,96	11,92	0,63
YDR382W	YOR312C	RPP2B	RPL20B	11	0,96	11,16	0,66
YGL076C	YLR448W	RPL7A	RPL6B	11	0,96	12,57	0,72
YHR193C	YLR304C	EGD2	ACO1	11	0,96	11,71	0,69
YGL103W	YHL033C	RPL28	RPL8A	10	0,98	11,88	0,67
YOL078W	YDR342C	AVO1	HXT7	10	0,98	10,29	-0,39
YOR232W	YBR009C	MGE1	HHF1	10	0,97	11,35	0,63
YLR125W	YLR288C	YLR125W	MEC3	10	0,97	9,33	0,1
YBL072C	YER102W	RPS8A	RPS8B	10	0,97	12,38	0,68
YAL028W	YDL239C	FRT2	ADY3	10	0,97	11,65	0,6
YHL033C	YKL006W	RPL8A	RPL14A	10	0,97	12,3	0,67
YDL047W	YLR304C	SIT4	ACO1	10	0,97	11,89	0,63

detected known regulatory interactions with a p-value cutoff of 2.7e-3 by the three methods.

"1" means the corresponding interaction can be detected by the corresponding method

"p" means the corresponding gene exists in the activated transcriptional factors or the differentially expressed genes

systemic nar	systemic nar	inferred by T	inferred by P	inferred by LC	regulator	regulated gene
YOL089C	YOL121C	1			HAL9	RPS19A
YER111C	YJL196C	1			SWI4	ELO1
YGL071W	YHL013C	1			RCS1	YHL013C
YGL071W	YKR026C	1			RCS1	GCN3
YER111C	YPL024W	1			SWI4	NCE4
YKL043W	YCR098C	1			PHD1	GIT1
YER111C	YML102C-	1	1	1	SWI4	YML102C-
YER111C	YBR071W	1	1	1	SWI4	YBR071W
YNL309W	YOL007C	1	1	1	STB1	YOL007C
YNL309W	YNL289W	1	1	1	STB1	PCL1
YIL131C	YDR179C	1			FKH1	CSN9
YNL309W	YPL256C	1	1	1	STB1	CLN2
YJL056C	YCR039C	1			ZAP1	MATALPH
YNL068C	YMR198W	1	1	1	FKH2	CIK1
YER111C	YOL019W	1			SWI4	YOL019W
YEL009C	YHR162W	1	1	1	GCN4	YHR162W
YNL068C	YPL117C	1			FKH2	IDI1
YNL027W	YGL038C	1			CRZ1	OCH1
YEL009C	YER055C	1	1	1	GCN4	HIS1
YLR182W	YKL112W	1	1	1	SWI6	ABF1
YPR104C	YBR126C	1			FHL1	TPS1
YMR042W	YDR434W	1			ARG80	GPI17
YEL009C	YBR249C	1	1	1	GCN4	ARO4
YKL112W	YJL111W	1			ABF1	CCT7
YDR207C	YDR285W	1			UME6	ZIP1
YML027W	YKL091C	1			YOX1	YKL091C
YGL035C	YBR050C	1			MIG1	REG2
YDR451C	YGL114W	1			YHP1	YGL114W
YDR501W	YDL112W	1			PLM2	TRM3
YML027W	YKL042W	1	1	1	YOX1	SPC42
YLR183C	YNR039C	1			TOS4	ZRG17
YML027W	YPL256C	1	1	1	YOX1	CLN2
YDL106C	YGL234W	1			PHO2	ADE5,7
YDR123C	YBR093C	1		1	INO2	PHO5
YER111C	YGL038C		1	1	SWI4	OCH1
YEL009C	YLR355C		1	1	GCN4	ILV5
YNL068C	YDR150W		1	1	FKH2	NUM1
YNL068C	YPL155C		1	1	FKH2	KIP2
YER111C	YDL127W		1	1	SWI4	PCL2
YEL009C	YDL182W		1	1	GCN4	LYS20
YER111C	YMR179W		1	1	SWI4	SPT21

YER111C YGR109C	1	1	SWI4	CLB6
YER111C YOL011W	1	1	SWI4	PLB3
YDL056W YMR215W	1	1	MBP1	GAS3
YNL309W YGR221C	1	1	STB1	TOS2
YNL068C YPR119W	1	1	FKH2	CLB2
YDR146C YGL108C	1	1	SWI5	YGL108C
YDL056W YOR246C	1	1	MBP1	YOR246C
YIL131C YPL227C	1	1	FKH1	ALG5
YNL068C YNL172W	1	1	FKH2	APC1
YER111C YPL267W	1	1	SWI4	YPL267W
YPL248C YBR019C	1	1	GAL4	GAL10
YIL131C YMR215W	1	1	FKH1	GAS3
YNL216W YPL016W	1	1	RAP1	SWI1
YER111C YML027W	1	1	SWI4	YOX1
YFL044C YCR021C	1	1	YOD1	HSP30
YDR423C YJL100W	1	1	CAD1	LSB6
YKL043W YMR015C	1	1	PHD1	ERG5
YNL068C YJR092W	1	1	FKH2	BUD4
YIL131C YPL139C	1	1	FKH1	UME1
YNL068C YBR138C	1	1	FKH2	YBR138C
YNL068C YJL051W	1	1	FKH2	YJL051W
YDL106C YDL058W	1	1	PHO2	USO1
YER111C YGR189C	1	1	SWI4	CRH1
YLR182W YJL092W	1	1	SWI6	HPR5
YDL056W YNR009W	1	1	MBP1	YNR009W
YKL109W YPR020W	1	1	HAP4	ATP20
YNL068C YBR038W	1	1	FKH2	CHS2
YDL056W YDR451C	1	1	MBP1	YHP1
YBR049C YPR164W	1	1	REB1	MMS1
YIL131C YLR210W	1		FKH1	CLB4
YGL096W YFR012W	1	1	TOS8	YFR012W
YML027W YGR109C	1	1	YOX1	CLB6
YIR017C YFR030W	1	1	MET28	MET10
YJR094C YMR133W	1	1	IME1	REC114
YIR017C YKL001C	1	1	MET28	MET14
YIR017C YJR010W	1	1	MET28	MET3
YML027W YFR027W	1	1	YOX1	ECO1
YCR096C YCR039C	1	1	HMRA2	MATALPH
YLR256W YDR044W	1	1	HAP1	HEM13
YML027W YER095W	1	1	YOX1	RAD51
YLR183C YGR015C	1	1	TOS4	YGR015C
YGL096W YDL223C	1	1	TOS8	HBT1
YGL096W YGL158W	1	1	TOS8	RCK1
YGL096W YDR058C	1	1	TOS8	TGL2
YLR183C YBR042C	1	1	TOS4	YBR042C

YDR253C YIR017C	1	1	MET32	MET28
YCR065WYOR144C	1	1	HCM1	ELG1
YLR183C YKR090W	1	1	TOS4	PXL1
YGL096WYEL020C	1	1	TOS8	YEL020C
YGL035C YMR011W	1	1	MIG1	HXT2
YOR372C YPR034W	1	1	NDD1	ARP7
YLR183C YPR174C	1	1	TOS4	YPR174C
YML027WYLR381W	1	1	YOX1	CTF3
YML027WYNL072W	1	1	YOX1	RNH201
YDR451C YNR028W	1	1	YHP1	CPR8
YPR065WYML075C	1	1	ROX1	HMG1
YCR065WYHR149C	1	1	HCM1	SKG6
YJR094C YHL022C	1	1	IME1	SPO11
YKL109W YJR048W	1	1	HAP4	CYC1
YBR297WYBR298C	1	1	MAL33	MAL31
YML027WYGL207W	1	1	YOX1	SPT16
YGL096WYHR182W	1	1	TOS8	YHR182W
YPL248C YOR120W	1		GAL4	GCY1
YLR183C YDR480W	1		TOS4	DIG2
YLR182W YKR013W		1	SWI6	PRY2
YLR182W YNL312W		1	SWI6	RFA2
YFR034C YPL133C		1	PHO4	RDS2
YLR182W YKL113C		1	SWI6	RAD27
YLR182W YNL289W		1	SWI6	PCL1
YLR182W YLR103C		1	SWI6	CDC45
YPR065WYMR017W		1	ROX1	SPO20
YLR182W YDL018C		1	SWI6	ERP3
YLR182W YIL026C		1	SWI6	IRR1
YLR182W YJL187C		1	SWI6	SWE1
YKR099WYFL060C		1	BAS1	SNO3
YIL131C YPL141C		1	FKH1	YPL141C
YOR372C YJL051W		1	NDD1	YJL051W
YML007WYJL098W		1	YAP1	SAP185
YNL068C YDL018C		1	FKH2	ERP3
YLR182W YMR076C		1	SWI6	PDS5
YIL131C YHR143W		1	FKH1	DSE2
YOR372C YGR230W		1	NDD1	BNS1
YER111C YER001W		1	SWI4	MNN1
YDL056W YJL074C		1	MBP1	SMC3
YLR182W YPL267W		1	SWI6	YPL267W
YPL049C YMR052W		1	DIG1	FAR3
YDL056W YER111C		1	MBP1	SWI4
YDL056W YOR074C		1	MBP1	CDC21
YKL043W YMR195W		1	PHD1	ICY1
YDL056W YNL312W		1	MBP1	RFA2

YOR028C YLL061W	1	CIN5	MMP1
YLR131C YOR138C	1	ACE2	RUP1
YOR372C YJL079C	1	NDD1	PRY1
YLR182W YNL300W	1	SWI6	YNL300W
YIR023W YGR120C	1	DAL81	COG2
YBR049C YMR039C	1	REB1	SUB1
YOR372C YGL114W	1	NDD1	YGL114W
YKL043W YMR017W	1	PHD1	SPO20
YHR206WYOL154W	1	SKN7	ZPS1
YDL056W YDR097C	1	MBP1	MSH6
YDR146C YOR138C	1	SWI5	RUP1
YCR018C YDR042C	1	SRD1	YDR042C
YLR182W YPL163C	1	SWI6	SVS1
YLR182W YER070W	1	SWI6	RNR1
YLR182W YHR154W	1	SWI6	RTT107
YER088C YBR156C	1	DOT6	SLI15
YDL056W YDL018C	1	MBP1	ERP3
YLR131C YER152C	1	ACE2	YER152C
YDL056W YKL113C	1	MBP1	RAD27
YIL131C YBR138C	1	FKH1	YBR138C
YNL068C YER070W	1	FKH2	RNR1
YER111C YGR014W	1	SWI4	MSB2
YOR372C YER078C	1	NDD1	YER078C
YML007WYHR048W	1	YAP1	YHR048W
YNL216W YIL177C	1	RAP1	YIL177C
YCR018C YMR257C	1	SRD1	PET111
YIL131C YGR092W	1	FKH1	DBF2
YDL056W YDL003W	1	MBP1	MCD1
YDL056W YDR113C	1	MBP1	PDS1
YPR104C YHR204W	1	FHL1	MNL1
YLR182W YER095W	1	SWI6	RAD51
YOR032C YDR042C	1	HMS1	YDR042C
YGR044C YLR439W	1	RME1	MRPL4
YOR372C YGL116W	1	NDD1	CDC20
YLR182W YOR074C	1	SWI6	CDC21
YER111C YER078C	1	SWI4	YER078C
YDL056W YBR071W	1	MBP1	YBR071W
YPR065WYDR042C	1	ROX1	YDR042C
YDL056W YNL273W	1	MBP1	TOF1
YKR099WYKR010C	1	BAS1	TOF2
YDR421WYDR523C	1	ARO80	SPS1
YIR023W YCL014W	1	DAL81	BUD3
YDR146C YDL173W	1	SWI5	YDL173W
YDL056W YHR153C	1	MBP1	SPO16
YOR372C YPR119W	1	NDD1	CLB2

YIL131C YPL032C	1	FKH1	SVL3
YNL068C YGR092W	1	FKH2	DBF2
YLR182W YJL074C	1	SWI6	SMC3
YER111C YMR305C	1	SWI4	SCW10
YMR016C YBL029W	1	SOK2	YBL029W
YGL035C YPR058W	1	MIG1	YMC1
YOR372C YPL242C	1	NDD1	IQG1
YER111C YJL187C	1	SWI4	SWE1
YNL068C YGL021W	1	FKH2	ALK1
YOR372C YDR451C	1	NDD1	YHP1
YLR182W YNL102W	1	SWI6	POL1
YGL096W YDR186C	1	TOS8	YDR186C
YLR256W YOR065W	1	HAP1	CYT1
YLR183C YLR182W	1	TOS4	SWI6
YOR372C YOR126C	1	NDD1	IAH1
YCR065W YOL014W	1	HCM1	YOL014W
YJR094C YNL210W	1	IME1	MER1
YGL096W YBR129C	1	TOS8	OPY1
YMR016C YOR187W	1	SOK2	TUF1
YDL056W YNL102W	1	MBP1	POL1
YDR123C YLR153C	1	INO2	ACS2
YPR065W YIL111W	1	ROX1	COX5B
YNL216W YGR254W	1	RAP1	ENO1
YLR182W YDL164C	1	SWI6	CDC9
YGL096W YPL049C	1	TOS8	DIG1
YOR372C YFR042W	1	NDD1	YFR042W
YGL096W YER026C	1	TOS8	CHO1
YDR451C YPL267W	1	YHP1	YPL267W
YLR182W YGR109C	1	SWI6	CLB6
YDR451C YER140W	1	YHP1	YER140W
YIL122W YGR166W	1	POG1	KRE11
YMR016C YOR070C	1	SOK2	GYP1
YBL005W YDR011W	1	PDR3	SNQ2
YGL096W YER014W	1	TOS8	HEM14
YBL005W YOL156W	1	PDR3	HXT11
YLR182W YPR120C	1	SWI6	CLB5
YLR182W YKL045W	1	SWI6	PRI2
YDR451C YDL156W	1	YHP1	YDL156W
YDR501W YBR073W	1	PLM2	RDH54
YCR065W YDR113C	1	HCM1	PDS1
YMR016C YJL011C	1	SOK2	RPC17
YMR016C YIL173W	1	SOK2	VTH1
YGL096W YDL174C	1	TOS8	DLD1
YCR065W YDR356W	1	HCM1	SPC110
YLR183C YPR175W	1	TOS4	DPB2

YGL096WYER161C	1	TOS8	SPT2
YLR183C YIL023C	1	TOS4	YIL023C
YDR451C YKL165C	1	YHP1	MCD4
YGL096WYFL009W	1	TOS8	CDC4
YLR256WYEL039C	1	HAP1	CYC7
YOL004WYAL005C	1	SIN3	SSA1
YOR372C YKL048C	1	NDD1	ELM1
YLR256WYPR065W	1	HAP1	ROX1
YOL004WYHR014W	1	SIN3	SPO13
YPR065WYDR044W	1	ROX1	HEM13

pressed genes sets of the cell cycle condition according to Luscombe et al.(2004,Nature,431,308).
 ne present regupresent regulated gene in cell cycle condition

	p	p
	p	
	p	
	p	
	p	
-A	p	
,	p	p
		p
	p	p
		p
IA2	p	
		p
,	p	p
	p	
	p	
	p	
	p	
	p	
	p	
	p	
	p	p
	p	
	p	p
	p	
	p	
		p
	p	p
	p	
	p	p

p
p
p

p

p

p
p
p

p

p

p
p
p

p
p
p

p

p

p

p
p

p
p
p
p

p
p
p

p
p
p

p
p

p
p
p
p
p
p
p
p
p
p

p

p

p

p
p

p

p
p
p
p

p
p
p
p
p

p
p
p
p
p
p
p

p
p
p
p
p

p

p
p

p
p
p

p
p
p
p

p
p

p
p
p

p
p
p
p
p
p

p
p
p

p

p
p

p
p
p
p

p
p

p

p	
p	p
p	
p	
	p
p	
p	