COMPARATIVE ANALYSIS OF TEAFS AND NP ANALYSIS TO INTEGRATE INTERACTOME AND TRANSCRIPTOME DATA TO REVEAL RESPONSE TO C-PULSE IN *SACCHAROMYCES CEREVISIAE*

by

Muhammed Erkan Karabekmez

B.S. in Chemical Engineering, Boğaziçi University, 2004

Submitted to the Institute for Graduate Studies in
Science and Engineering in partial fulfillment of
the requirements for the degree of
Master of Science

Graduate Program in Chemical Engineering

Bogaziçi University

2010

ACKNOWLEDGEMENT

Before all, I have to thank my supervisor Prof. Betül Kırdar for her encouragements during all my works. It was a privilege to work with her during my graduate study as she made possible this thesis to be created with her support, trust and patience.

I would like to sincerely thank to Assist. Prof. Kazım Yalçın Arğa who introduced me the computational systems biology at the beginning. I would also like to thank to Duygu, Dicle, Ayça and all other members of Biosystems Engineering Research Group for their contributions to my thesis and for their companionship.

I have a deep gratitude to İsa for his friendship during all this process and even before; for more than one and an half decades. I am also grateful to Bayram Ali for his intimate friendliness and scientific zeal. I would also like to thank my housemate İsmail and all my other friends from boundinos, spistforeignaffairs and bura for their sincere comradeship.

Finally, I have to express that I am forever in debt to my parents for providing me with affection, peace, support and trust. I would like to deeply thank to my sister, Nurgül, my brothers, Furkan and Hüseyin and my cute nieces Fatma Sueda and Esma Reyyan for their heartfelt support to make life more bearable. I am also sincerely grateful to my love, my fiancée, Meryem who rose into my life and turned it into a miracle. I am giving praise to God for my family, without them definitely I could not achieve anything in my life.

ABSTRACT

COMPARATIVE ANALYSIS OF TEAFS AND NP ANALYSIS TO INTEGRATE INTERACTOME AND TRANSCRIPTOME DATA TO REVEAL RESPONSE TO C-PULSE IN *SACCHAROMYCES CEREVISIAE*

Sudden glucose introduction to *S. cerevisiae* cells grown in carbon limited medium triggers a complex response mechanism. Time series expression data can be used to identify differentially active genes after the C-pulse. However, in order to reveal dynamics of the response and formulate more meaningful hypotheses, integration of interactome data and transcriptome data can be useful. In this study, two previously proposed integrative methods, Topological Enrichment Analysis of Functional Subnetworks (TEAFS) and NP analyses were applied using the time series expression data. In TEAFS, condition specific networks corresponding to each time point were constructed based on presence/absence criteria. In order to analyze local dynamic topological changes in the interaction network biological process modules as group of connected proteins with the same GO biological process term were identified. Then standard deviations of topological parameters of the modules were scored to identify significantly active biological processes after the C-pulse. 216 distinct modules were identified as active. The most active modules during overall time span were found to be related to ribosome biogenesis, transcription, rRNA processing, regulation of transcription and translation. Module related to regulation of transcription from RNA polymerase II promoter was identified in all distinct time spans. Nine, six and 96 active modules were identified as seconds, minutes and hours specific modules respectively. In NP analysis, protein-protein interaction network were reduced to a global active subnetwork (NP network) by classifying interactions as correlated, anti-correlated or not correlated based on expression profiles. Four modules of similar expression profiles were determined by hierarchical clustering of expression profiles. Gene Ontology annotations of these modules revealed down-regulated or up-regulated biological processes

and their possible interactions after a glucose pulse. The expression level of module M decreased and that of module P increased significantly with time until the 8th minute. The expression levels of these two modules changed in opposite directions after 8th minute. The expression level of D module displayed similar trends with M module but in a more smooth way. The expression level of module S was almost constant over entire time. The genes of P, M, D and S modules were found to be enriched in 147, 97, 169 and 137 distinct GO biological process terms, respectively. The module P was enriched in GO annotations related to ribosome biogenesis, transcription, and rRNA processing. Metabolic process, oxidation reduction, cellular response to heat and transport were found to be enriched terms in module M. The module S was enriched in translation, DNA repair, transport, cell cycle, chromatin modification, and glycolysis. The module D was enriched in the terms transport, protein amino acid phosphorylation, and ubiquitin-dependent protein catabolic process. Enrichment in transcription in both S and D modules were significant. PSA1 which is related to cell wall biosynthesis was determined to be contributing to the communication between P and D modules and HHF1 which is related to chromatin assembly was identified to be contributing to the communication between P and M modules. The active biological processes which were identified by two distinct approaches using the same transcriptome data set were compared in order to interpret the advantages /disadvantages of these two approaches.

ÖZET

ETKİLEŞİM VE ANLATIM VERİSİNİN BÜTÜNLEŞTİRİLMESİ YÖNTEMLERİ OLAN TEAFS VE NP ANALİZİ İLE *SACCHAROMYCES CEREVISIAE*'NİN GLİKOZ VURUMUNA TEPKİSİNİN KARŞILAŞTIRMALI ANALİZİ

Kısıtlı glikoz ortamında yetiştirilen *S. cerevisiae* hücrelerinde ani glikoz girdisi karmaşık bir tepki mekanizmasını tetikler. Glikoz girdisinden sonra zamansal seri anlatım verisi kullanılarak değişime en anlamlı yanıt veren genler belirlenebilir. Fakat seri anlatım verisi ile etkileşim verisi bütünleştirilerek tepkinin dinamiği açığa çıkarılabilir ve aynı zamanda daha anlamlı önermeler düzenlenebilir. Bu çalışmada daha önce önerilen TEAFS ve NP analizi isimli etkileşim ve anlatım verilerini bütünleştirmeye yarayan iki farklı yöntem kullanılarak glikoz girdisinden sonra gözlenen zamansal seri anlatım verisi analiz edilmiştir. TEAFS yönteminde, var/yok ölçütüne göre her zaman noktası için koşula bağlı ağlar oluşturuldu. Etkileşim ağının bölgesel dinamik topolojisini analiz edebilmek için aynı GO biyolojik süreç terimleri ile ilintili ve etkileşim içerisinde olan protein gurupları biyolojik süreç birimleri olarak tanımlandı. Daha sonra bu birimlerin topolojik özelliklerinin standart sapmaları puanlanarak, glikoz girdisinden sonra etkin olan biyolojik süreçler belirlendi. Zamansal seri anlatım verisinin tamamı kullanılarak bu şekilde tespit edilen birimler içinde en etkin olanlar ribozom oluşumu, transkripsiyon, rRNA üretimi, transkripsiyon düzenlenmesi ve translasyon düzenlenmesi birimleri olarak tespit edilmiştir. RNA polimeraz II ilerleticisi ile transkripsiyon düzenlenmesi ile ilintili birim zamansal seri anlatım verisinin parçalı olarak analizlerinin hepsi ile tespit edilmiştir. Bu parçalı analizlerde sırasıyla dokuz, altı ve 96 adet etkin birim saniyelere, dakikalara ve saatlere özgü birimler olarak tespit edilmişlerdir. NP analizinde ise etkileşimleri ilişkili, ters ilişkili veya ilişkisiz olarak sınıflandırarak protein-protein etkileşim ağı genel bir etkin ağa indirgendi. Anlatım görüntülerinin hiyerarşik olarak sınıflandırılması ile benzer anlatım görüntülerine sahip dört birim belirlendi. Bu birimlerin gen varlıkbilimsel açıklamaları,

glikoz girdisinden sonra azaltılmış ve artırılmış biyolojik süreçleri ve bunların olası etkileşimlerini açığa çıkardı. Sekizinci dakikaya kadar M biriminin anlatım seviyesi anlamlı bir şekilde azalmış, P biriminin ki ise artmıştır. Sekizinci dakikadan sonra ise bu iki birimin anlatım seviyeleri ters yönde değişmeye başlamıştır. D biriminin anlatım seviyesi M biriminin ki ile benzer bir görüntüyü daha düz bir şekilde göstermiştir. S biriminin anlatım seviyesi ise zaman içinde neredeyse hiç değişmemiştir. P, M, D ve S birimlerindeki genlerin sırası ile 147, 97, 169 ve 137 farklı GO biyolojik süreçle ilintili oldukları bulundu. P biriminin ilintili olduğu GO terimleri ribozom oluşumu, transkripsiyon ve rRNA üretimidir. Metabolik üretim, yükseltgenme indirgenme, ısıya hücresel tepki ve taşınım terimleri M birimi ile ilintili olarak bulundu. S biriminin translasyon, DNA tamiri, taşınım, hücre döngüsü, kromatin değişikliği ve glikoliz terimleri ile ilintili olduğu tespit edildi. D biriminin ise taşınım, protein aminoasit fosforlaşması ve ubikutine bağlı protein parçalanması süreci terimleri ile ilintili olduğu tespit edildi. S ve D birimlerinin ikisinin de transkripsiyonla önemli ölçüde ilintili olduğu bulundu. Hücre duvarı oluşumu ile ilgili bir gen olan PSA1'nın P ve D birimleri arasındaki iletişimi sağladığı ve kromatin toparlanması ile ilgili bir gen olan HHF1'in ise P ve M birimleri arasındaki iletişimi sağladığı tespit edildi. Aynı anlatım verisinden iki farklı yöntemle tespit edilen etkin biyolojik süreçler karşılaştırılarak iki yöntemin avantaj ve dezavantajları yorumlandı.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

TEAFS	Topological enrichment analysis of functional subnetworks
NP	Negative/Positive
GO	Gene ontology
TF	Transcriptional factor
dp	Detection parameter
EDA-LCC	Extend of differential activity of local clustering coefficient
EDA-LCON	Extend of differential activity of local connectivity
PPI	Protein-protein interaction
PCC	Pearson correlation coefficient
HCE	Hierarchical clustering explorer
AvgPCC	Average Pearson correlation coefficient
FDR	False discovery rate

1. INTRODUCTION

Cellular processes operate in a complex mechanism of actions of molecular components and their interactions. Classical molecular biology approaches are not capable of handling such sophisticated systems. In order to investigate cell's large-scale organization and reveal transcriptional regulatory mechanism, post-genomic biology has focused on systems biology. Fortunately, with rapid developments in hardware technologies, recently, high throughput methodologies have been used to enlighten biological systems. Then, instead of reductionist approach of studying particular components of cell, more global and integrative approaches of studying all genes or proteins of cell at once were started to get more interest. It was stated that a more complete understanding of cell will require not only the description of individual proteins or genes but also the description of all components involved in a biological process. This conceptual shift in the strategy of post-genomic studies has triggered the development of mathematical models at the molecular level to predict responses of these systems in to various treatments (Vidal M., 2001). The mathematical models mentioned can rely solely on interactome, transcriptome or other –omes and also integration of these –omes as well.

Biological roles of topological properties of protein-protein networks have been investigated in many studies (Barabasi and Oltvai 2004). Dynamically organized modularity (Han et al., 2004), (Bertin et al. 2007), scale-free structure with embedded hierarchical modularity (Yook et al., 2004), non-hub-centric hierarchical organization (Valente and Cusick 2006) in the protein-protein interaction networks have been proposed. Visualization of interactome networks has been studied (Lu *et al.*, 2004).

In response to stimuli, the expressions of a large fraction of the genes change either by increasing or decreasing its levels (Gasch et al, 2000). Therefore, co-expressed modules over time or under different environmental conditions or gene deletion mutants have been tried to be extracted by using large-scale gene expression data. In addition to various clustering approaches which have been proposed to uncover transcriptional regulatory models (Ihmels *et al.*, 2004), (Segal *et al.*, 2003), (Herrgard *et al.*, 2003), (Tavazoie *et al.*,

1999), it has been shown that hierarchical clustering of time course expression profiles could reveal groups of genes of similar functions (Eisen *et al.,* 1998), (Wen *et al.,* 1998).

Interactome data is crucial for cellular communication and it is necessary to be taken into consideration, as transcriptome data alone cannot provide a complete understanding of a biological process. For instance proteins with many links may not have a significant differential expression level but still, they are more likely to play important roles in the cellular organization. However, topological properties of a static interactome data alone cannot sufficiently explain biological meaning of a network neither (Guet et al. 2002). Then, Integrative modeling approaches will be inevitable to use the information in predictive models. It is to say that, integrating the static interactome together with the expression profiles can help to understand the dynamics of a cellular network. Even visualization of integrated networks alone can be a useful tool to interpret biological relations (Gopalacharyulu *et al.,* 2008).

In order to accomplish the task of data integration many methodologies have been suggested in diverse studies. Integrating interactome data with transcriptome data of perturbated states have been studied as genes with similar expression profiles across various conditions are more likely to encode linked proteins (Ideker *et al.,* 2001). Transcriptome data of different states were also integrated with protein-protein interaction networks to elucidate the biological processes associated with disorders (Liu et al., 2007). By integrating transcriptome and interactome data in different ways various biological information has been revealed, such as potential protein-protein interactions or regulatory models (Yeang *et al.,* 2004), (Ge *et al.,* 2001), (Jansen *et al.,* 2002), (Xia *et al,* 2006), (Gopalacharyulu *et al,* 2009). Luscombe *et al.* has showed that topological properties of active subnetworks of the protein-protein network, which created by using transcriptome data, differ under external stimuli from internal stimuli (Luscombe *et al.,* 2004). Lee *et al.,* had used transcriptome and interactome data with a Bayesian statistics approach to establish a functional gene network (Lee *et al.,* 2004). Time series transcriptome data have been used to improve interactome data (Bader et al., 2004), (Jansen *et al.,* 2003). It has been proposed to integrate also phenotype profiles with interactome and transcriptome data as well (Gunsalus *et al.,* 2005), (Walhout *et al.,* 2002).

Two distinct approaches for this integration, called as ‘Topological Enrichment Analysis of Functional Subnetworks’ (TEAFS) which was recently proposed by Gopalacharyulu *et al,* in 2009 (Gopalacharyulu *et al,* 2009) and ‘NP analysis’ which was proposed by Xia *et al,* in 2006 (Xia *et al,* 2006), were used to investigate dynamic topology changes in functional modules and time dependent behavior of regulatory networks.

TEAFS based on the assumption that environmental challenges are tightly dynamically regulated, with the aim to maintain the functionality of the system. Therefore, they described functional modules by using the Gene Ontology and studied dynamic topological changes of these modules by integrating time course genome-wide gene expression data, with integrated networks (i.e., protein–protein interaction, metabolic, and regulatory networks). It has been shown that this approach is effective in finding the activity of mechanisms which classical methods could not identify (Gopalacharyulu *et al,* 2009).

Protein–protein interaction (PPI) networks have been shown to have dynamic modular structures (Han *et al,* 2004). But Xia *et al,* first, used time course expression data to find out a global active subnetwork (NP network) and then tried to split the subnetwork into artificial modules, to be able to identify proliferation/differentiation switch in the cellular network of multicellular organisms (Xia *et al*, 2006). Moreover, Xue *et al,* in 2007, used the same methodology to analyze aging phenomena for human and fruit fly (Xue *et al*, 2007).

1.1. TEAFS

‘Topological Enrichment Analysis of Functional Subnetworks’ (TEAFS) proposed by Gopalacharyulu *et al,* in 2009 (Gopalacharyulu *et al.,* 2009) based on differential activity of local topological parameters, namely connectivity and clustering coefficient. After identifying condition specific networks, scoring the topological parameters of modules, defined according to Gene Ontology (GO) biological process terms, is followed by a permutation step.

1.1.1. Presence/absence criteria

Luscombe *et al,* in 2004, proposed a methodology to eliminate nodes in a discrete on/off manner (Luscombe *et al,* 2004). Based on this approach, Gopalacharyulu *et al.,* detailed presence/absence criteria in order to construct condition specific networks (Table 1.1.). To do so each node was classified as present or absent based on their corresponding gene's reference expression level and whether it was up, constant or down-regulated after a C-pulse (Gopalacharyulu *et al,* 2009). The reference expression levels were clustered into three groups, low, medium, and high, by k-mean clustering. Nodes which are encoded by genes, whose expression level was clustered as high at the reference condition, were assumed to be present in each condition specific network. Whereas, proteins, whose reference expression levels of encoding genes were clustered as medium, were classified as present if the expression is up-regulated or remained constant. On the other hand, when an initial expression level was clustered as low, only if its change with respect to the initial condition is up, its corresponding node was said to be present (Table 1.1.).

Table 1.1. Presence/Absence Criteria

Reference expression level	Change in expression level	Conclusion on Presence or Absence
High	Up	Present
	Constant	Present
	Down	Present
Medium	Up	Present
	Constant	Present
	Down	Absent
Low	Up	Present
	Constant	Absent
	Down	Absent

According to this methodology, all genes in each condition were classified as present or absent. Then, links of absent nodes were also removed from the networks to get condition specific networks.

1.1.2. Gene Ontology (GO)

To be able to analyze dynamic behaviors of local topological parameters it is needed to define functional subnetworks or modules. At this point Gopalacharyulu *et al,* decided to define modules to be consisted of nodes whose corresponding genes has the same GO biological process term as GO concept based on the annotations of homologous gene and protein sequences in multiple eukaryotic organisms (Ashburner *et al,* 2000).

GO biological process terms form a network tree which has 'biological process' at the top. Each GO term can have more than one parent and also more than one children. Genes have direct associations to GO terms and indirect associations to all parental GO terms of that GO term. One gene can have multiple GO term annotations, thus the modules can overlap to each other (Ashburner *et al,* 2000).

1.1.3. Topological parameters of a network

One of long term goals of systems biology is to match biological interpretations to topological properties. It has been shown that changes in clustering coefficient and connectivity are measures of biological activity in a PPI (Barabasi and Oltvai, 2004). Then, in TEAFS, standard deviations of local total clustering coefficients and degrees over time are described as extend of differential activity (EDA) scores of the modules. Therefore, for each module two kinds of EDA scores can be calculated; EDA-LCON and EDA-LCC based on local connectivities and local clustering coefficients respectively. To check statistical significance of calculated EDA scores a permutation step was developed as well (Gopalacharyulu *et al*, 2009).

1.2. NP analysis

Genes which response to an environmental perturbation in the same way or, also, in an opposite manner were shown to be more likely exist in the same mechanisms (Dhillon *et al*, 2003). Therefore, in this case, NP analyses, defined by Xia *et al*, based on correlation and anti-correlations of expression profiles of genes. They eliminated not-correlated links from a PPI to reduce global network to a NP network where N stands for negative, anti-

correlation, and P stands for positive, correlation. Then the largest modules with opposite expression profiles were obtained by manual dissection. GO biological process terms were used to give biological meanings to these large modules and global subnetwork NP which is active after an environmental perturbation (Xia *et al*, 2006). The procedure was visualized (Figure 1.1.).

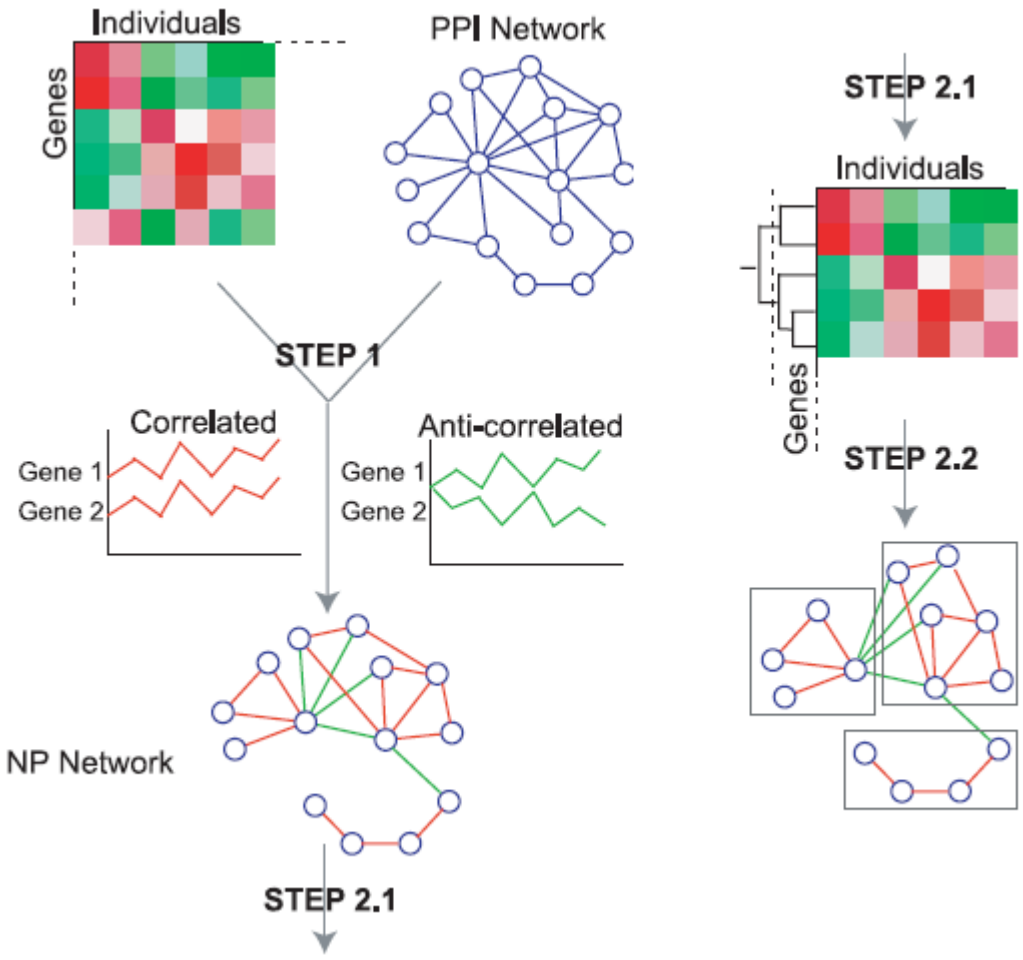

Figure 1.1. NP analysis pipeline used to reveal the anti-correlated modules (Xia *et. al.*, 2006)

1.2.1. Pearson Correlation Coefficient (PCC)

For each gene pair in 'step 1' of Fig. 1.1., identification of being correlated or anti-correlated is based on PCC values. PCC captures linear relationship between any pair of observations. It focuses on change trends rather than values of expression levels. It is chosen not to skip genes with little expression levels. With PCC based clustering it has

shown that genes with opposite expression profiles are often functionally related (Dhillon *et al*, 2003).

PCC value of 1 corresponds to the perfect correlation whereas that of -1 stands for the perfect anti-correlation. Whereas, getting perfect correlations or anti-correlations for experimental measurements are impossible. Therefore it was needed to set cut-off values to identify correlations and anti-correlations (Xia *et al*, 2006).

Any edge of PPI network, that links proteins whose encoding genes' absolute PCC value is less than selected cut-off, was removed from the reference network (Figure 1.1.). After edge removal, nodes remained alone were removed as well to construct global active subnetwork of NP network.

1.2.2. Hierarchical Clustering

It has been mentioned that PPI networks have modular structures (Han *et al*, 2004). In NP analysis, Xia *et al*, proposed to split NP network by hierarchical clustering largest possible modules to have less than 1% of anti-correlated links within the modules (Figure 1.1).

Hierarchical clustering explorer (HCE) (http://www.cs.umd.edu/hcil/hce/) applies without predetermined number of cluster which is suitable for manual dissecting. Dendogram view obtained by using HCE 3.0 can be used to split into smaller modules step by step until reaching largest possible modules with less than 1% of anti-correlated links. Then, modules which enclose correlated edges and interact with each other mostly with anti-correlated edges would be refined.

1.2.3. Betweenness

It has shown that there is a strong relation between nodes' betweenness and their essentiality for the interaction network (Yu *et al*, 2007). As a topological parameter, a node's betweenness is a measure of number of shortest paths that passes through that node. Betweenness is also highly related to degree as can be seen in the illustration of Figure 1.2.

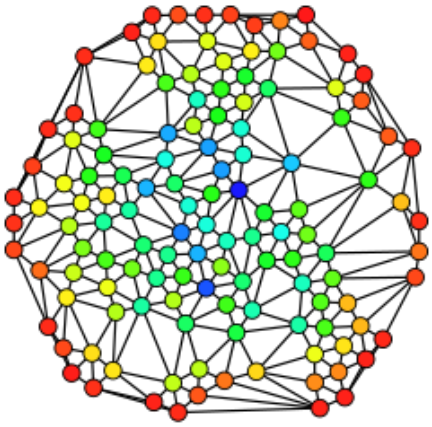

Figure 1.2. Simplified visualization of betweenness (Node colors with ascending order of betweenness are red, yellow, green, and blue) (http://www.answers.com/)

While studying aging, Xue *et al*, had showed that the increase of probability of being a regulatory gene with increasing betweenness in NP network is sharper than that of reference network. Then, to interpret the NP network, betweenness distribution of transcriptional factors can be a useful parameter.

Analyzing betweenness can help to capture possible hidden intrinsic hierarchy in networks. Degree analyses alone cannot take into account essential non-hub nodes with high betweenness (Figure 1.3.).

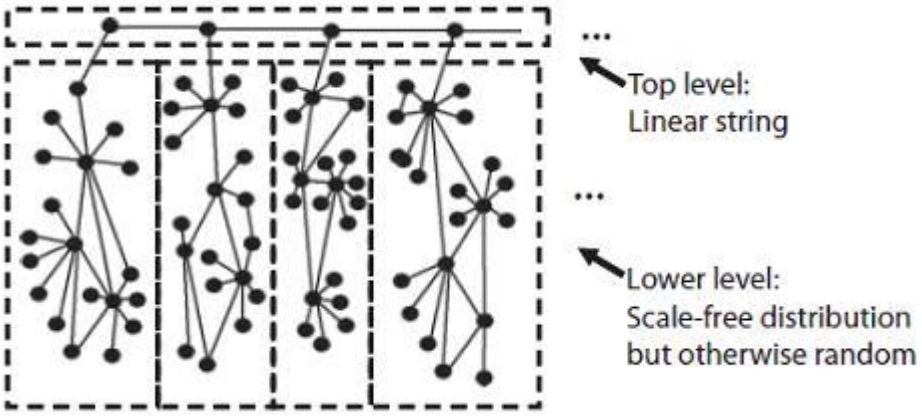

Figure 1.3. Non-hub-centric hierarchical organization (Valente and Cusick 2006)

1.2.4. Date and party hubs

Date and party hubs described by Han *et al,* as hubs whose neighbors' average PCC values are low and high were called as date hubs and party hubs respectively (Figure 1.4.) (Han *et al.,* 2004). Consistency of this classification across different datasets was also verified (Bertin *et al.,* 2007). This distinction suggests a model of organized modularity with modules connected through regulators, mediators or adaptors, the date hubs. Party hubs represent integral elements within distinct modules (Han *et al.,* 2004).

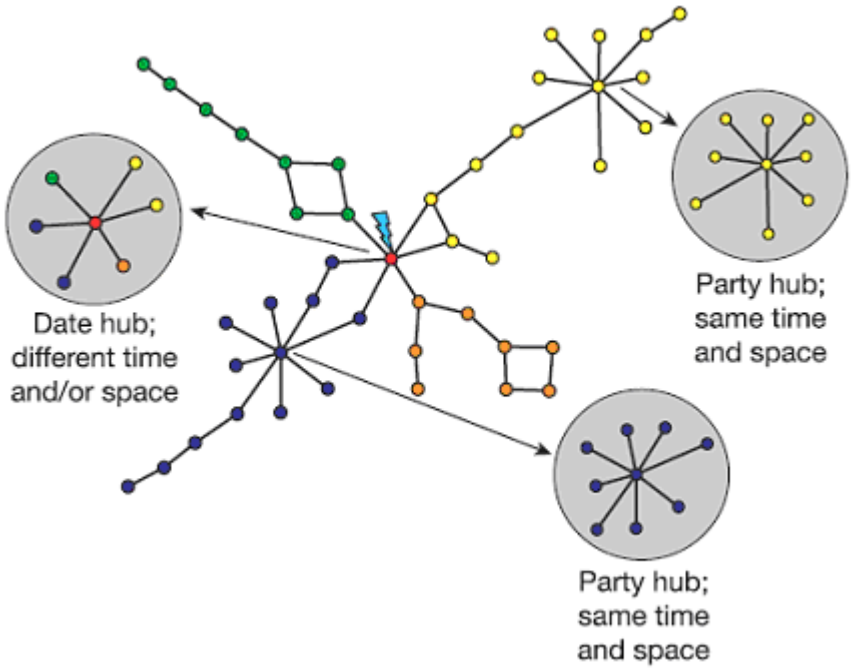

Figure 1.4. Date and party hubs (Han *et al.,* 2004)

1.3. The aim of the Thesis

In this study two distinct approaches of integrating transcriptome and interactome data were analyzed in a comparative analysis. Time series transcriptome data, which was collected after a sudden C-pulse for yeast which has grown in a carbon limited media, was integrated with protein-protein interaction data by using the previously proposed methods, namely TEAFS and NP analysis.

The methodology was explained in the Methods section. All results were presented in the Results section while the comparison of the results of distinct approaches was interpreted in Discussion section. Further work was also proposed to improve the study.

2. METHODS

2.1. Raw data

Interactome data downloaded (http://acgt.cs.tau.ac.il/matisse/) rely on following interaction identification methods; Affinity Capture-MS, Affinity Capture-RNA, Affinity Capture-Western, Affinity Chromatography, Affinity Precipitation, ChIP, Chip on-chip, Co-crystal Structure, Co-purification, Phosphorylation Array, Purified Complex, Two hybrid and literature as well. These interaction data had collected from diverse sources (Ulitsky and Shamir, 2007). The raw protein-protein network consists of 6220 nodes and 63990 edges.

Time series transcriptome data of yeast after a C-pulse which was measured for 5668 genes at 13 different time points between initial and final steady states were kindly provided by Dikicioglu. Wild type strain BY4743 (MATa/MATα his3Δ/his3Δ leu2Δ/leu2Δ LYS2/lys2Δ MET15/met15Δ ura3Δ/ura3Δ), which is derived from S288C, was cultivated under aerobic conditions in glucose limited media (2% w/v) before pulse injection. After 150 hours of fermentation 50 ml of 40% (w/v) glucose injected to 1 L of the media. After injection samples were collected after 20, 40, and 60 seconds, then after 8, 16, 24, and 32 minutes, and finally after 1, 2, 3, 4, 5, and 7 hours (Dikicioglu *et al.*, unpublished data).

2.2. Topological Enrichment Analysis of Functional Subnetworks (TEAFS)

Dynamic topological changes of the functional modules were analyzed by a method called "TEAFS" described by Gopalacharyulu *et al.,* in 2009 (Gopalacharyulu *et al.,* 2009).

2.2.1. Construction of condition specific networks

To identify condition specific networks for each time point, presence/absence criteria based on studies of Luscombe *et al.,* in 2004 (Luscombe *et al.,* 2004), were used. These criteria based on initial expression level of a gene and change in its expression level at a

condition specific point. At first, two reference points' expression level data clustered into three groups (high, medium and, low) by k-means clustering by using Matlab. And then, ratio of expression level at condition specific points to that of initial point is classified as up, constant and, down as well (Gopalacharyulu *et al,* 2009). Initial expression level was taken as averages of two reference data points and they were rounded to one digit for constancy.

According to the initial expression level and the change in expression level, presence/absence criteria were applied as detailed in Table 1.1. and condition specific networks for 13 time points were determined.

2.2.2. Identification of Functional Modules

Gene Ontology (GO) biological process terms were used for the identification of the functional modules. From *Saccharomyces cerevisiae* Genome Database (http://www.yeastgenome.org/), GO biological process term annotations for all genes were downloaded. Groups of genes with the same GO term annotation with an average clustering coefficient which is significantly greater than zero were defined as functional modules. For this significance, student t-test was applied to the average clustering coefficients of each module (calculated in Matlab, the code was provided in Appendix B1) over all time points and the ones with higher *p-value* than 0.05 were identified as functional modules.

2.2.3. Scoring the Functional Modules

At first, for each module (*m*), at each time point (*t*), total local connectivity; *LCON(m,t)* and total local clustering coefficient; *LCC(m,t)* parameters were calculated. By using this calculations extent of differential activity (*EDA*) scores, which is defined as the standard deviation of the local topological parameters over time, were calculated as in Eqn. 2.1. and Eqn. 2.2. Two distinct Matlab codes were written to carry out these calculations (provided in Appendix B2 and B3).

$$EDA-LCON(m)=\sigma_t(LCON(m,t)) \qquad (2.1.)$$

$$EDA-LCC(m)=\sigma_t(LCC(m,t)) \tag{2.2.}$$

To check statistical significance of calculated EDA scores a permutation step was used (Gopalacharyulu *et al*, 2009). Basic assumption of this permutation test is that total topological properties, across all time point networks, decreases uniformly with a deactivation ratio of *x(t)* which is defined as ratio of sum of topological properties of nodes turning off from time point *t* to *t+1* with sum of topological properties of all nodes present at time point *t*.

Within 1000 accordingly generated sets of networks, *p-value* of a module was defined as ratio of EDA scores found to be greater or equal to the original EDA scores. These calculations were executed by using Matlab (Appendix B4 and B5). Then to be able to control falsely rejected cases *q-values* of FDR were calculated as in Eqn. 2.3.

$$q(i)=\frac{m\times p(i)}{i} \tag{2.2.}$$

p-value of module *i* is *p(i)* as *p-values* of *m* modules in ascending order and *q(i)* is the FDR *q-value* of module *i* (Benjamini and Hochberg, 1995). Gopalacharyulu *et al*, 2009 called a module active if its EDA scores are high and its *q*-value is less than chosen cut-off value of *0.05*.

2.2.4. TEAFS for distinct time spans

By hierarchical clustering explorer we clustered time points to analyze distinct time spans of the original data. Original time course expression data is consist of 3 time points in second base, 4 time points in minute base and, 6 time points in hour base. Hierarchical clustering led the same groups. Hierarchical Clustering Explorer (HCE) software was used for this analysis. By analyzing these time points as three different time spans, according to TEAFS defined in Subsection 2.5.1, three sets of results were obtained as active modules within seconds, minutes, and hours respectively.

2.3. NP analysis

To reduce the initial global interaction network to a global active subnetwork NP analysis defined by Xia *et al,* in 2006 (Xia *et al,* 2006) was used, where NP stands for negative and positive correlations. NP analysis relies on expression correlations and anti-correlations between pairs of genes which are measured by the Pearson correlation coefficient (PCC). PCC was chosen to focus on trend of expression changes rather than the expression values. To calculate PCC matrices a Matlab code was used (Appendix B6).

2.3.1. Identification of an active subnetwork of the global network

PCC value of *1* means perfect correlation and *-1* means perfect anti-correlation between any couple of nodes. Here, $PCC>0.7$ was accepted for correlation and $PCC<-0.7$ for anti-correlation.

From the reference PPI network, interactions between nodes that are not correlated or anti-correlated were eliminated. After this elimination the nodes with no interaction were eliminated as well. The remaining network is called as "NP network" and assumed to be a global active subnetwork of the reference global PPI network.

2.3.2. Expression levels of modules and Average PCC values

Arithmetic average of expression levels of genes in a module at a time point was defined as expression level of a module at this time point. Average PCC values between any pair of modules were calculated by using expression levels of the modules.

To investigate trends of expression levels of modules slopes and *R-square* were calculated by Microsoft Excel.

2.3.3. Identification of modules

From the NP network, two largest possible modules with an opposite expression profiles were extracted by hierarchical clustering methods. By using Hierarchical

Clustering Explorer (HCE) software, nodes of the NP network clustered. Hierarchical centered correlation and average linkage were used after log transforms, median center and, randomization steps, for clustering. Then, the largest anti-correlated clusters, so that the clusters will have less than *1%* intra-cluster anti-correlated interactions, were manually dissected.

2.3.4. Biological meanings of the modules

From Saccharomyces cerevisiae Genome Database, GO biological process term annotations for the genes of each module, identified in Subsection 2.2.2, were downloaded. To eliminate insignificant GO terms a detection parameter (*dp*) which is defined as in Eqn. 2.3., were calculated.

$$dp = \left(\frac{n}{t}\right)^2 \tag{2.3.}$$

Where n is the number of genes associated with the GO term and t is the total number of genes in the species. A GO term was used only if *dp* was greater than a significance threshold (*0.05/g*), where g is the total number of GO terms associated with all the genes in a module. Bonferroni correction for multiple hypotheses testing on the total number of GO terms was applied to detection parameters. Calculations were carried out by using Microsoft Excel.

To be able to consider indirect associations as well, GO slim term cluster frequencies obtained from http://www.yeastgenome.org/.

2.3.5. Interface and Core

Interface was defined as described by Xue *et al.*, 2007 (Xue *et al.*, 2007). For each pair of modules, interface and core genes were identified separately. Any node of module *X* with at least one edge with the module *Y* were defined as in the interface of module *X* with module Y and interface genes of *X* and interface genes of *Y* were defined to compose interface of module pair *X*-*Y*. The remaining genes of these two modules were defined as

the core of the *X*-*Y* couple. Interface genes for each couple of modules were extracted by Matlab (Appendix B7). The remaining genes of the cores were extracted in Microsoft Excel by using result of the Matlab code.

The TFs which were used to analyze regulatory roles of interfaces and cores were downloaded from online YEASTRACT database (http://www.yeastract.com/) (27.04.2009). The number of known TFs for *Saccharomyces cerevisiae* which was downloaded was 201. Among them, 192 TFs were found to be present in the reference network and were used for further investigations.

2.3.6. Betweenness

Betweenness is a measure of centrality within a network. Betweenness (b_i) for a node is defined as the number of shortest paths between two nodes which passes through a particular node. Specifically, betweenness is defined for the m^{th} node in the network as in Eqn. 2.4. where $\Gamma(i,m,j)$ is the number of shortest paths between i^{th} and j^{th} nodes that passes through m^{th} node. Python software was used for betweenness calculations.

$$b_m = \sum_{i \neq j} \Gamma(i, m, j) \tag{2.4.}$$

2.3.7. AvgPCC

AvgPCC value of a node was defined as the arithmetic average of all PCC values with its neighbors. Specifically, AvgPCC is defined for the m^{th} node in the network as in Eqn. 2.5. where *PCC(m,i)* is the PCC value between expression profiles of m^{th} and i^{th} nodes and *n* is the number of neighbors of m^{th} node.

$$AvgPCC_m = \frac{\sum_{i}^{n} PCC(m, i)}{n} \tag{2.5.}$$

For a hub, low absolute AvgPCC values indicate that it is a “date hub” and high absolute AvgPCC values indicate that it is a “party hub” (Han *et al.*, 2004). The cut off value for this distinction was devised as 0.1 as shown in Eqn. 2.6.

$$\begin{aligned} |AvgPCC_m| > 0.1 &\Rightarrow \text{"}party\text{"} \\ |AvgPCC_m| < 0.1 &\Rightarrow \text{"}date\text{"} \end{aligned} \qquad (2.6.)$$

3. RESULTS

3.1. Topological Enrichment Analysis of Functional Subnetworks (TEAFS)

Topological Enrichment Analysis of Functional Subnetworks (TEAFS) was used to identify dynamic response of *S. cerevisiae* to the relaxation of carbon limiting conditions by injection of a carbon pulse. TEAFS was developed by Gopalacharyulu et al (2009) to analyze integrated interactome data and time series expression data.

3.1.1. Networks

In response to C-pulse, expression data available at 13 time points and a protein-protein interaction network, consisting of 5539 nodes (whose both expression and interaction data were available) and 59902 bidirectional edges, were integrated to construct reference and condition specific networks corresponding to each the time point. Condition specific networks were determined using presence/absence criteria obtained from the gene expression data (Gopalacharyulu *et al.*, 2009). The numbers of nodes which were considered to be present or absent in a condition specific network were provided in Table 3.1. The numbers of the genes which were not included into the condition specific networks at each time point was presented in shaded boxes (Table 3.1.).

Table 3.1. Number of present/absent genes in condition specific networks

Reference expression levels	Number of nodes	Expression change	20 sec	40 sec	60 sec	8 min	16 min	24 min	32 min	1 hr	2 hr	3 hr	4 hr	5 hr	7 hr
High	1415	Up	639	607	693	686	700	693	746	625	662	656	567	563	543
		Constant	100	77	76	56	47	57	67	59	83	85	83	97	85
		Down	676	731	646	673	668	665	602	731	670	674	765	755	787
Medium	3292	Up	1442	1400	1625	1547	1545	1522	1680	1388	1509	1496	1391	1325	1295
		Constant	238	167	152	122	129	141	167	169	194	183	182	191	195
		Down	1612	1725	1515	1623	1618	1629	1445	1735	1589	1613	1719	1776	1802
Low	961	Up	427	408	466	466	465	475	507	416	468	450	421	412	385
		Constant	73	69	61	40	42	30	45	64	63	79	62	66	51
		Down	461	484	434	455	454	456	409	481	430	432	478	483	525
Total	5539	Present	3522	3390	3658	3550	3554	3553	3769	3388	3586	3544	3409	3343	3290
		Absent	2146	2278	2010	2118	2114	2115	1899	2280	2082	2124	2259	2325	2378

Sizes, average connectivities and clustering coefficients of condition specific global networks at each time point are provided in Table 3.2. We found that the reference network has higher numbers of nodes and edges, and higher values of average clustering coefficient and connectivity than all time specific networks.

Among the time specific networks the one at 32 minute has the highest number of nodes and edges (3595 and 30857 respectively). The highest average clustering coefficient was present at 16 minute which is 0.3492 and the highest average connectivity was found to be 8.6675 at 8 minute. On the other hand, the network at 7 hour has the lowest number of nodes and average clustering coefficient; 3106 and 0.2981 respectively. Similarly the network at 4 hours has that of edges and connectivity; 21467 and 6.7126 (Table 3.2.). These observations suggest that between 8 and 32 minutes, condition specific networks were larger and more connected and they become smaller and less connected after 4th hour as the system reaches again a steady state.

Table 3.2. Number of nodes and edges, average connectivities and clustering coefficients of the reference and condition specific networks

Time	Nodes	Edges	Average Clustering Coefficient	Average Connectivity
20 sec.	3336	27282	0.3285	8.1781
40 sec.	3226	25531	0.3250	7.9141
60 sec.	3478	28958	0.3307	8.3260
8 min.	3398	29452	0.3435	8.6675
16 min.	3397	28854	0.3492	8.4940
24 min.	3399	29126	0.3385	8.5690
32 min.	3595	30857	0.3480	8.5833
1 hr.	3206	25515	0.3212	7.9585
2 hr.	3423	27954	0.3368	8.1665
3 hr.	3400	28863	0.3329	8.4891
4 hr.	3198	21467	0.3152	6.7126
5 hr.	3161	23203	0.3057	7.3404
7 hr.	3106	22468	0.2981	7.2337
Reference	5539	59902	0.3584	10.8146

3.1.2. Hubs

Hubs in a network are the nodes with larger degrees compared to all other nodes and therefore, they are more likely to play an important role within an interaction network. Top 20 highest degree nodes in the reference network were selected as hubs in the present study. The time points where these hubs become absent (removed from the network) were provided with the hubs (Table 3.3.). Descriptions of the hubs were obtained from http://www.ncbi.nlm.nih.gov/.

RNA polymerase I enhancer binding protein; REB1p, transcriptional factor that mediates drug resistance and salt tolerance; CIN5p, regulatory protein that regulates sterol biosynthesis; ECM22p and transcriptional factors that regulate transcriptions at G2/M and M/G1 transitions, FKH2p and SWI5p respectively, were identified as hubs in all conditional networks. On the other hand, DNA binding protein SWI4p and its transcriptional cofactor SWI6p that regulate transcription at G1/S transition, were identified as absent in all networks except in condition specific network at 7^{th} hours. In this 7^{th} hour network, SWI4p was present while SWI6p was still absent.

A DNA binding transcriptional activator CHA4p that mediates serine/threonine activation of the catabolic L-serine (L-threonine) deaminase (CHA1) was absent in all conditional networks. SSA1p which is encoding an ATPase involved in protein folding and nuclear localization signal was only present in the last network (Table 3.3.).

The half of the reference network hub proteins were absent in 40 seconds network (Table 3.3.). This result is in good agreement with the observed sharp decrease in average connectivity values at 40^{th} seconds (Table 3.2.). DNA replication regulator SMT3p and proteins which are regulated by nitrogen presence, like GLN3p, TPK1p, and DAL81p were identified as hub proteins in 40sec network (Table A1).

Table 3.3. Hubs in the reference network.

Gene Name	Degree	Description	Absent at time points
MCM1	500	Transcription factor involved in cell-type-specific transcription and pheromone response; plays a central role in the formation of both repressor and activator complexes	4hr, 7hr
SWI4	479	DNA binding component of the SBF complex (Swi4p-Swi6p), a transcriptional activator that in concert with MBF (Mbp1-Swi6p) regulates late G1-specific transcription of targets including cyclins and genes required for DNA synthesis and repair	20sec-5hr
GAT1	466	Transcriptional activator of genes involved in nitrogen catabolite repression; contains a GATA-1-type zinc finger DNA-binding motif; activity and localization regulated by nitrogen limitation and Ure2p	20-60sec
PSA1	440	GDP-mannose pyrophosphorylase (mannose-1-phosphate guanyltransferase), synthesizes GDP-mannose from GTP and mannose-1-phosphate in cell wall biosynthesis; required for normal cell wall structure	40sec, 8-32min
REB1	421	RNA polymerase I enhancer binding protein; DNA binding protein which binds to genes transcribed by both RNA polymerase I and RNA polymerase II; required for termination of RNA polymerase I transcription	-
TEF2	420	Translational elongation factor EF-1 alpha; also encoded by TEF1; functions in the binding reaction of aminoacyl-tRNA (AA-tRNA) to ribosomes	1hr
URA2	417	Bifunctional carbamoylphosphate synthetase (CPSase)-aspartate transcarbamylase (ATCase), catalyzes the first two enzymatic steps in the de novo biosynthesis of pyrimidines; both activities are subject to feedback inhibition by UTP	20-40sec, 1hr-7hr
FKH2	414	Forkhead family transcription factor with a major role in the expression of G2/M phase genes; positively regulates transcriptional elongation; negative role in chromatin silencing at HML and HMR; substrate of the Cdc28p/Clb5p kinase	-
PHO85	399	Cyclin-dependent kinase, with ten cyclin partners; involved in regulating the cellular response to nutrient levels and environmental conditions and progression through the cell cycle	20sec
FKH1	359	Forkhead family transcription factor with a minor role in the expression of G2/M phase genes; negatively regulates transcriptional elongation; positive role in chromatin silencing at HML and HMR; regulates donor preference during switching	40sec, 1hr
SKO1	354	Basic leucine zipper (bZIP) transcription factor of the ATF/CREB family, forms a complex with Tup1p and Ssn6p to both activate and repress transcription; cytosolic and nuclear protein involved in osmotic and oxidative stress responses	40-60sec, 16min
CIN5	346	Basic leucine zipper transcriptional factor of the yAP-1 family that mediates pleiotropic drug resistance and salt tolerance; localizes constitutively to the nucleus	-
RAP1	346	DNA-binding protein involved in either activation or repression of transcription, depending on binding site context; also binds telomere sequences and plays a role in telomeric position effect (silencing) and telomere structure	2hr
ECM22	326	Sterol regulatory element binding protein, regulates transcription of the sterol biosynthetic genes ERG2 and ERG3; member of the fungus-specific Zn[2]-Cys[6] binuclear cluster family of transcription factors; homologous to Upc2p	-
SWI5	323	Transcription factor that activates transcription of genes expressed at the M/G1 phase boundary and in G1 phase; localization to the nucleus occurs during G1 and appears to be regulated by phosphorylation by Cdc28p kinase	-

SWI6	322	Transcription cofactor, forms complexes with DNA-binding proteins Swi4p and Mbp1p to regulate transcription at the G1/S transition; involved in meiotic gene expression; localization regulated by phosphorylation; potential Cdc28p substrate	20sec-7hr
CBF1	314	Helix-loop-helix protein that binds the motif CACRTG, which is present at several sites including MET gene promoters and centromere DNA element I (CDEI); required for nucleosome positioning at this motif; targets Isw1p to DNA	4hr
HSF1	311	Trimeric heat shock transcription factor, activates multiple genes in response to stresses that include hyperthermia; recognizes variable heat shock elements (HSEs) consisting of inverted NGAAN repeats; posttranslationally regulated	20sec-32min
CHA4	302	DNA binding transcriptional activator, mediates serine/threonine activation of the catabolic L-serine (L-threonine) deaminase (CHA1); Zinc-finger protein with Zn[2]-Cys[6] fungal-type binuclear cluster domain	20sec-7hr
SSA1	300	ATPase involved in protein folding and nuclear localization signal (NLS)-directed nuclear transport; member of heat shock protein 70 (HSP70) family; forms a chaperone complex with Ydj1p; localized to the nucleus, cytoplasm, and cell wall	20sec-5hr

3.1.3. Modules

In order to analyze local dynamic topological changes in the interaction network biological process modules as group of connected proteins with the same GO biological process term were identified using AmiGO (http://amigo.geneontology.org/). A total of 1797 distinct biological process terms were captured for all 5539 genes. To filter out the small and less connected modules, clustering coefficients were calculated for these 1797 modules at each time point and a t-test was performed over time to select modules with average clustering coefficients significantly greater than zero (*p-value* < 0.005). A total of 294 modules whose average clustering coefficient over time was significantly greater than zero were selected for further topological analysis.

In order to identify differentially active modules extend of differential activity scores *EDA-LCON* and *EDA-LCC*, based on two separate topological parameters clustering coefficient and connectivity respectively, were calculated as described in the previous section. Criterion for differential activity of a module in TEAFS method is based on false discovery rate (FDR) *q-value*. A cut-off of *0.05* for the *q-value* was selected to identify active modules. It should be noted that high *EDA* scores and low FDR *q-values* indicate modules' differential activity or significant fluctuations of their topological properties. A

highly active module identified in the present study indicates the most changing module in time as a response to the relaxation from C-limiting conditions.

Out of 294 modules 200 modules were found to be active using *EDA-LCON* score based on connectivity fluctuations. The use of *EDA-LCC* score based on clustering coefficient values revealed 167 differentially active modules. These two separate results, based on two separate topological parameters, overlaps for 151 modules. Cumulatively, 216 distinct modules were identified as active among the 294 modules. All modules which were found to be active and their *EDA* scores, *p-values, q-values* and average clustering coefficient were tabulated (Table A2). Modules with the top 20 *EDA-LCON* scores were selected for further investigation (Table 3.4.). The most active module was found to be related to ribosome biogenesis. Transcription, rRNA processing, regulation of transcription and translation were also identified as active modules. These modules were also identified equally active on basis of *EDA-LCC* scores.

Table 3.4. Modules with the highest EDA-LCON scores.

Module names	Number of Proteins in the Module	Average Clustering Coefficient in the Reference Network	EDA - LCON	EDA – LCC	p-value Connectivity	p-value Clustering Coefficient	FDR q-value Connectivity	FDR q-value Clustering Coefficient
ribosome biogenesis	697	0.25	1440.1	29.49	0	0	0	0
Transcription	598	0.29	701.4	15.68	0	0	0	0
rRNA processing	267	0.52	639.4	20.73	0	0	0	0
regulation of transcription, DNA-dependent	629	0.27	481.1	9.825	0	0	0	0
Translation	633	0.23	360.0	6.907	0	0	0	0
mRNA processing	175	0.61	291.3	8.979	0	0	0	0
RNA splicing	114	0.62	236.8	6.58	0	0	0	0
nuclear mRNA splicing, via spliceosome	119	0.67	205.7	6.736	0	0	0	0
regulation of translation	404	0.15	200.4	7.529	0	0	0	0
ubiquitin-dependent protein catabolic process	266	0.52	187.4	7.699	0	0	0	0
transcription from RNA polymerase II promoter	66	0.63	122.7	3.907	0	0	0	0
endonucleolytic cleavage in ITS1 to separate SSU-rRNA from 5.8S rRNA and LSU-rRNA from tricistronic rRNA transcript (SSU-rRNA, 5.8S rRNA, LSU-rRNA)	42	0.63	112.0	5.92	0	0	0	0

maturation of SSU-rRNA from tricistronic rRNA transcript (SSU-rRNA, 5.8S rRNA, LSU-rRNA)	66	0.41	102.5	4.187	0	0	0	0
rRNA metabolic process	304	0.15	101.1	6.104	0	0	0	0
protein transport	411	0.41	91.4	6.921	0.003	0.003	0.005	0.006
cell cycle	323	0.32	89.5	8.367	0.017	0	0.026	0
endonucleolytic cleavage to generate mature 5'-end of SSU-rRNA from (SSU-rRNA, 5.8S rRNA, LSU-rRNA)	32	0.62	77.2	3.945	0	0	0	0
ribosomal large subunit biogenesis	50	0.65	75.7	5.106	0	0	0	0
ribosomal large subunit assembly	50	0.47	74.7	7.262	0	0	0	0
chromatin modification	128	0.66	72.5	3.49	0	-	0	-

In order to visualize the fluctuations in the topological parameters within the top 20 modules which were identified as the modules with the highest EDA-LCON scores, average clustering coefficients at each sampling point were plotted (Figure 3.1A., Figure 3.1B. and Figure 3.1C.). Average clustering coefficient of 13 modules, out of top 20 (Table 3.4.), have shown a sudden drop immediately after the relaxation from carbon limited condition upon pulse injection (Fig. 3.1A. and Fig. 3.1B.). The remaining modules which were related to ribosome biogenesis term were found to display more fluctuations after 1st hour (sample point 8) and a peak at 3rd hour (sample point 10) (Figure 3.1C.).

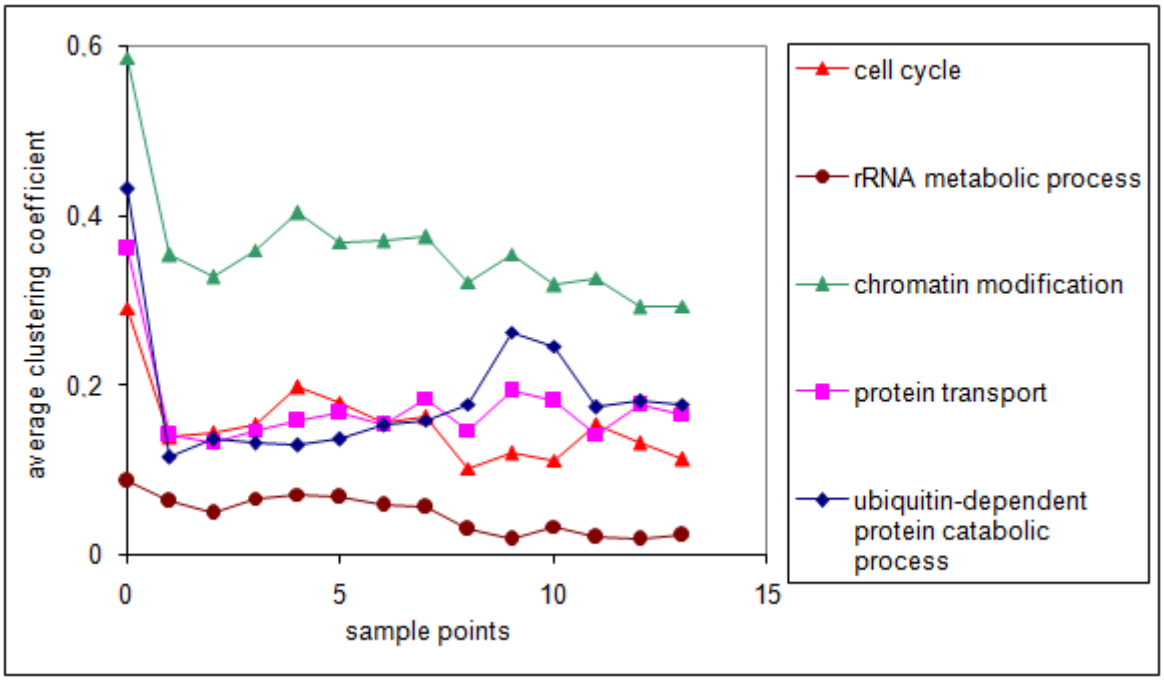

Figure 3.1A. Clustering coefficients of modules as a function of time after pulse injection

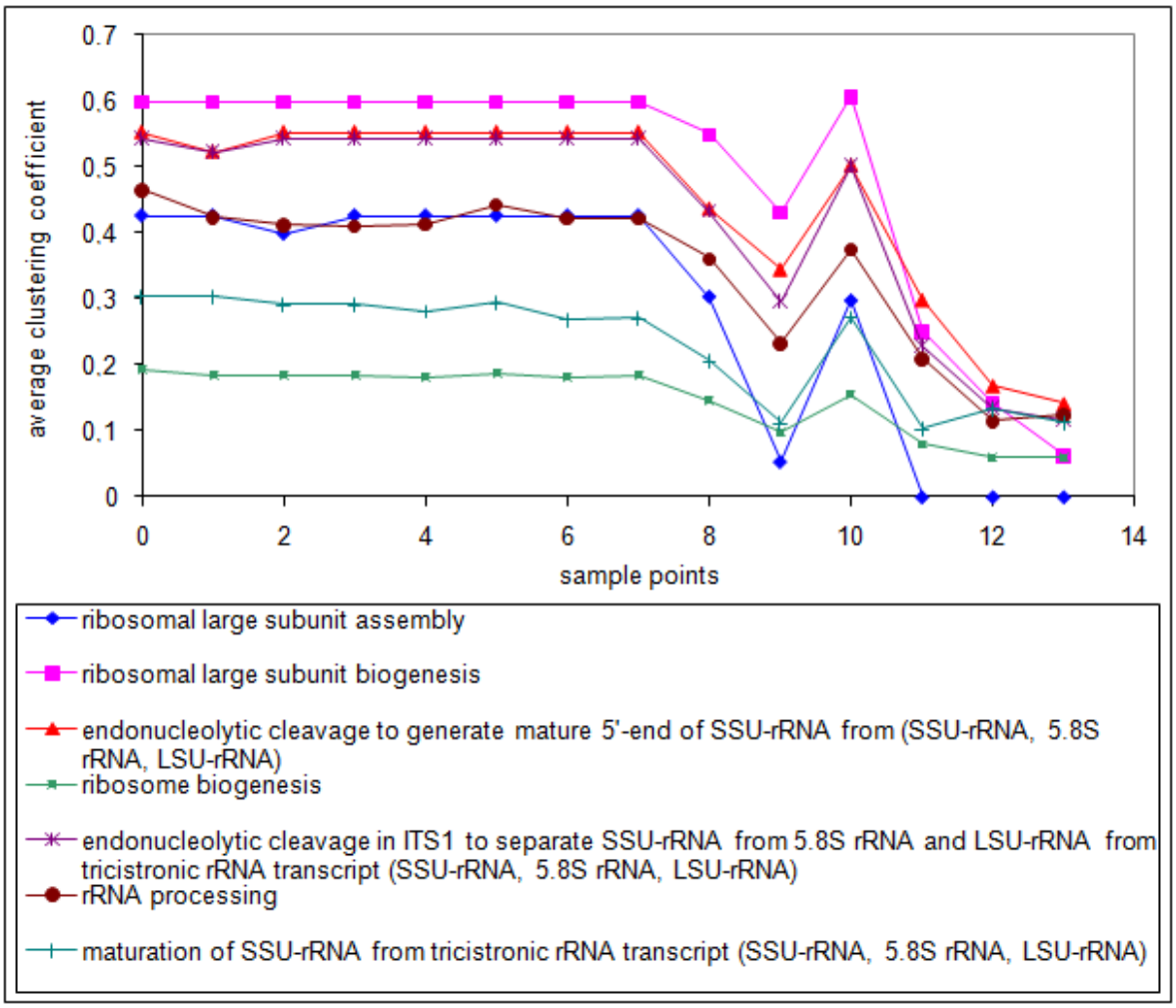

Figure 3.1B. Clustering coefficients of modules as a function of time. after pulse injection

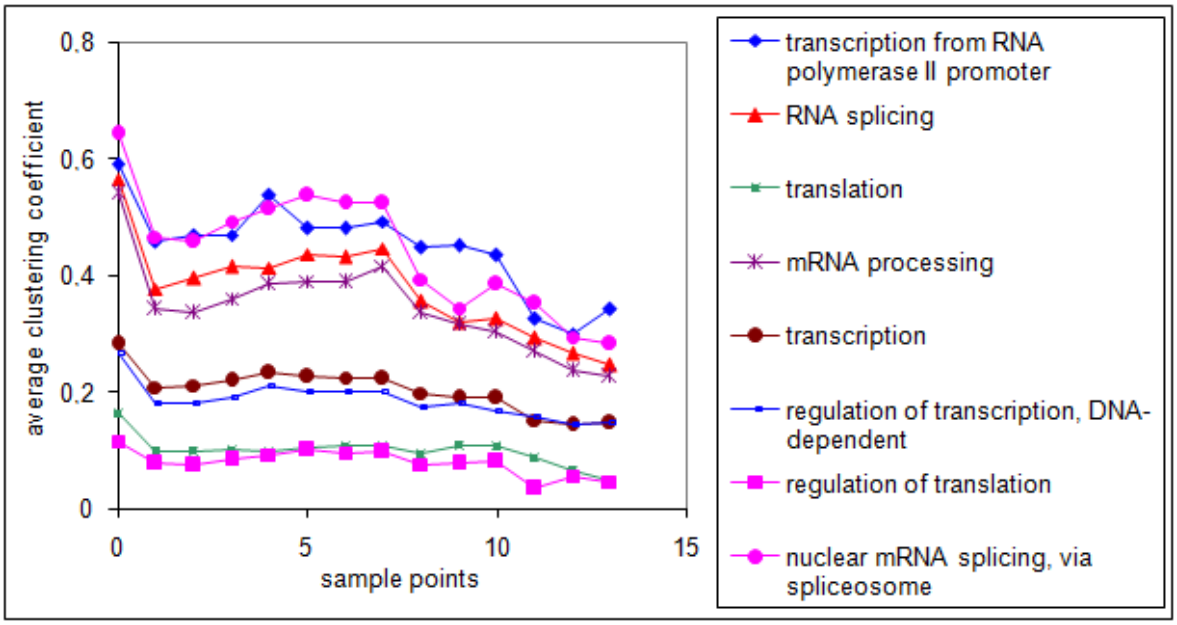

Figure 3.1C. Clustering coefficients of modules as a function of time after pulse injection

3.1.4. Distinct Time Spans

The time series expression data have already been analyzed by Dikicioglu *et al.,* and shown that expression levels at closer time points were similar by using hierarchical clustering analysis (Figure 3.2.) (Dikicioglu *et al.,* unpublished data). Samples collected within the first minute after carbon pulse, have similar expression levels. Similarly samples of the first hour composed another group.

In order to analyze the dynamic topological changes within seconds, minutes and hours, TEAFS analyses were applied as described in Section 2.2.3. over three distinct time spans. Time series expression data counted for the first of these three time spans were composed only of data points within the first minute. The second time span consists of 8, 16, 24, and 32 minutes. The last time span was formed by the remaining last 6 data points at which the samples were collected after the first hour. Hereafter these time spans will be called as seconds, minutes and hours.

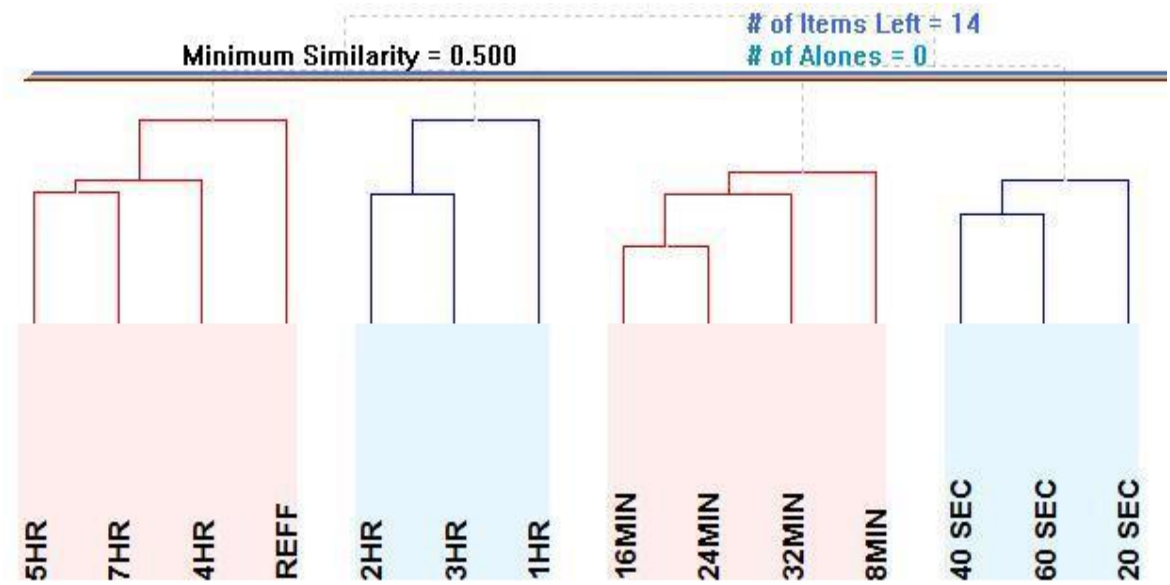

Figure 3.2. Pearson correlation based hierarchical clustering results of time points. 'REFF' is for reference time point (Dikicioglu *et al.,* unpublished data)

85, 107, and 210 active modules were identified by using EDA-LCON and 49, 77, and 170 active modules by using EDA-LCC scores, were identified as condition specific active modules at seconds, minutes, and hours respectively. A total of 94, 122, and 229 distinct active modules were identified as condition specific active modules in seconds,

minutes and hours respectively (Table 3.5.). In Table 3.5. the number of active modules identified by connectivity and clustering coefficient based EDA scores are under the columns ‘EDA-LCON’ and ‘EDA-LCC’, respectively. ‘Overlap’ is the percentage of the modules in the third column that exist also in the second column. ‘Total’ is the number of distinct modules in the cumulative list of the second and the third columns.

Table 3.5. Numbers of modules identified using the data at entire data points and over distinct time spans (rows).

	EDA-LCON	EDA-LCC	Overlap	Total
Entire data points	200	167	90.40%	216
Seconds	85	49	81.60%	94
Minutes	107	77	80.50%	122
Hours	210	170	88.80%	229

216 distinct modules were identified as active using the data at entire time points as described in the Section 3.1.3. These 216 modules were also tabulated in Table A2 together with their identification status over distinct time spans.

A comparative analysis indicated that out of 94, 122, and 229 active modules identified as seconds, minutes and hour specific 88, 112, and 199 modules were also present in 216 modules identified using all data points.

Six active modules related to biological process terms namely flocculation via cell wall protein-carbohydrate interaction, regulation of transcription from RNA polymerase III promoter, biosynthetic process, histone acetylation, protein targeting to ER, and transcriptional preinitiation complex assembly were only identified as second specific active modules. And we can easily presume that these modules react in short term as their topological fluctuations loss their significance after 60 seconds (Figure 3.3.).

Tree active modules which were identified as condition specific modules in minutes, namely, transcription initiation from RNA polymerase II promoter, histone methylation, and RNA 3'-end processing could not be identified among the second and hour specific

modules. It can be concluded that topological properties of these modules fluctuates significantly (FDR *q-value<0.05*) between 8 and 32 minutes as a response to relaxation from carbon limitation after the injection of C-pulse. Four modules, related to G1/S transition of mitotic cell cycle, DNA packaging, chromatin silencing at silent mating-type cassette and posttranslational protein targeting to membrane, translocation display significant (FDR *q-value<0.05*) fluctuations between 20 seconds and 32 minutes (Figure 3.3.)

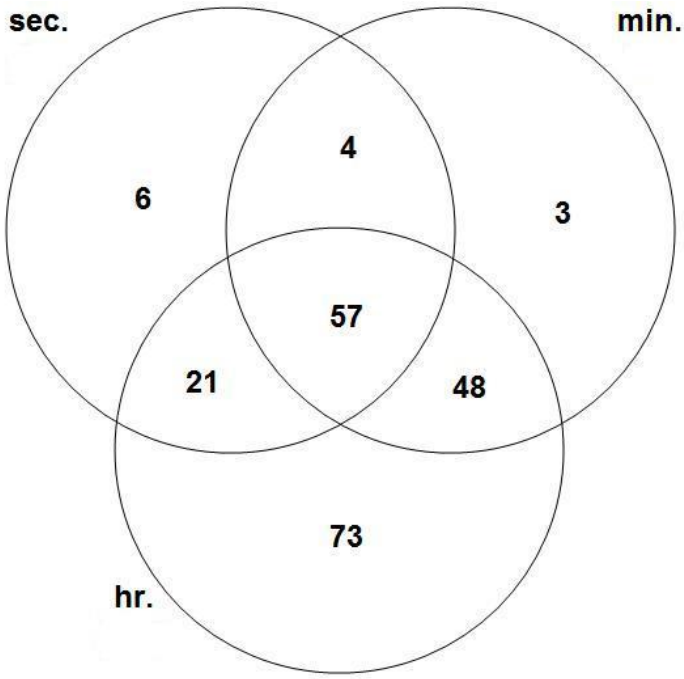

Figure 3.3. Overlapping active modules identified at both distinct time spans and overall time series analysis.

4 modules out of 216 active modules identified using the entire time points could not be detected over any of these time spans. These are methionine biosynthetic process, protein amino acid N-linked glycosylation via asparagine, interphase, and nuclear-transcribed mRNA catabolic process, 3'-5' exonucleolytic nonsense-mediated decay. This might be due to that their insignificant (FDR *q-value>0.05*) fluctuations of topological parameters over three distinct time spans add up to a relatively significant (FDR *q-value<0.05*) fluctuation in the overall analysis and thus they were detected.

A total of 37 modules, which were not identified by the overall time series analysis, were identified by TEAFS as distinct time span specific modules (Table A3 and Figure 3.4.). Module related to regulation of transcription from RNA polymerase II promoter was identified in all distinct time spans. Biological process term protein ubiquitination during ubiquitin-dependent protein catabolic process related module however, was identified as seconds and minutes specific and could not be identified in hours. Biopolymer biosynthetic process related module was identified as active in seconds and hours but not active in minutes. Three modules related to ergosterol biosynthetic process, steroid biosynthetic process, and Golgi to endosome transport were identified as seconds specific active modules. Another three modules related to axial cellular bud site selection, lipid biosynthetic process and protein insertion into ER membrane were detected as minutes specific active modules. Five modules related to establishment of cell polarity, protein folding, filamentous growth, sulfur amino acid metabolic process and rRNA pseudouridine synthesis were detected as minute and hour specific active modules that could not be identified in seconds and by overall time series analysis (Figure 3.4.)

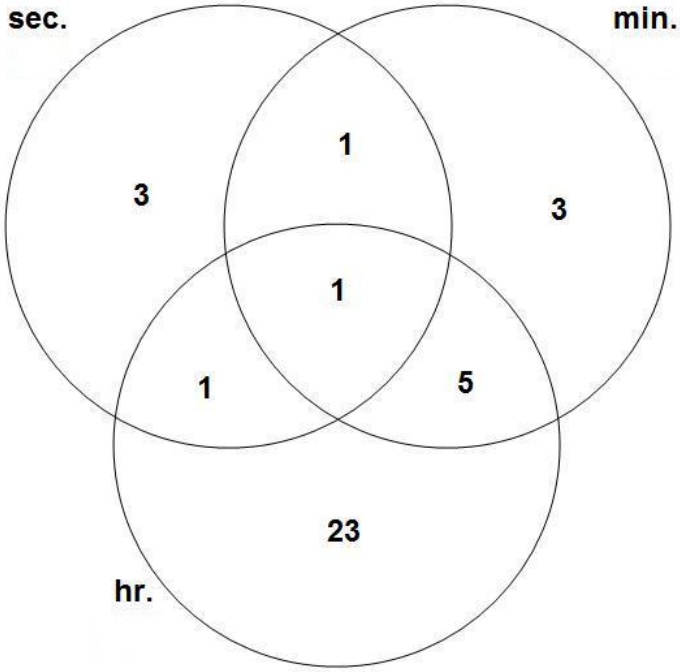

Figure 3.4. The overlapping second- minute- and hour- specific active modules which could not be detected by overall time series analysis.

3.2. NP analysis

Dynamic modular structure of the protein-protein interaction networks during the relaxation from C-limitation as a response to C-pulse were also analyzed using NP analysis developed by Xia *et al.* 2006. This Analysis permits also integration of both transcriptome and interactome information and serves to identify sub-network that is active during a specific process. NP analysis uses correlation or anti-correlation concepts to classify expression profiles of genes. To investigate these concepts Pearson Correlation Coefficient (PCC) was used. PCC was chosen to focus on expression profiles rather than expression levels. Any pair of expression profiles with similar trends but at completely different levels could be detected as correlated by using PCC.

3.2.1. Construction of NP Network

For 5668 genes, for which time course expression data available, correlations and anti-correlations in the expression levels of any pair of genes were measured by the Pearson correlation coefficient (PCC) as described in Section 2.3. PCC value which is equal to 1 indicates a perfect correlation and -1 for perfect anti-correlation in the expression levels of two genes. PCC values of 0.3 and -0.3 were selected as cutoffs for correlations and anti-correlations, respectively. In other words any gene pair with a PCC value which was less than -0.7 were defined as anti-correlated and a PCC value which was greater than 0.7 was classified as correlated. The interactions in PPI network if they link two nodes whose corresponding genes were not correlated or anti-correlated ($-0.7 < PCC < 0.7$) were eliminated. Nodes that were remained edgeless after this process were also eliminated. Remaining network was called NP network (N:negative, P:positive) with two type of interactions: Positive interactions which connect correlated nodes and negative interactions which connect anti-correlated nodes.

Reference PPI network consisted of 5539 nodes and 59902 bidirectional edges. Constructed NP network contained 3419 nodes and 12286 edges. 8762 of these edges were correlated edges and the remaining 3524 edges were anti-correlated edges (Table 3.6.).

Table 3.6. Properties of reference network and NP network.

	Reference network	NP network
Nodes	5539	3419
Edges	59902	12286
Co-regulated edges	-	8762
Anti-regulated edges	-	3524

3.2.2. Extraction of Modules

NP analyses are based on the partition of NP network into the largest possible co-regulated modules. Hierarchical Clustering Explorer was used to cluster genes according to their expression pattern (Figure 3.5.). Two main clusters identified by this program (shown in blue and red in Figure 3.5.) manually narrowed until each module were found to have less than *1%* intra cluster anti-correlated interactions and then these two modules were named as 'P' and 'M'. Numbers of nodes, percent of anti-correlated edges between the two modules as well as percent of anti-correlated edges within the modules were presented in Table 3.7. Among red colored genes of Figure 3.5. excluded genes after elucidating 'P' module, were named as module 'S'. Similarly, from blue colored genes of Figure 3.5. excluded part was named as module 'D' (Figure 3.5.).

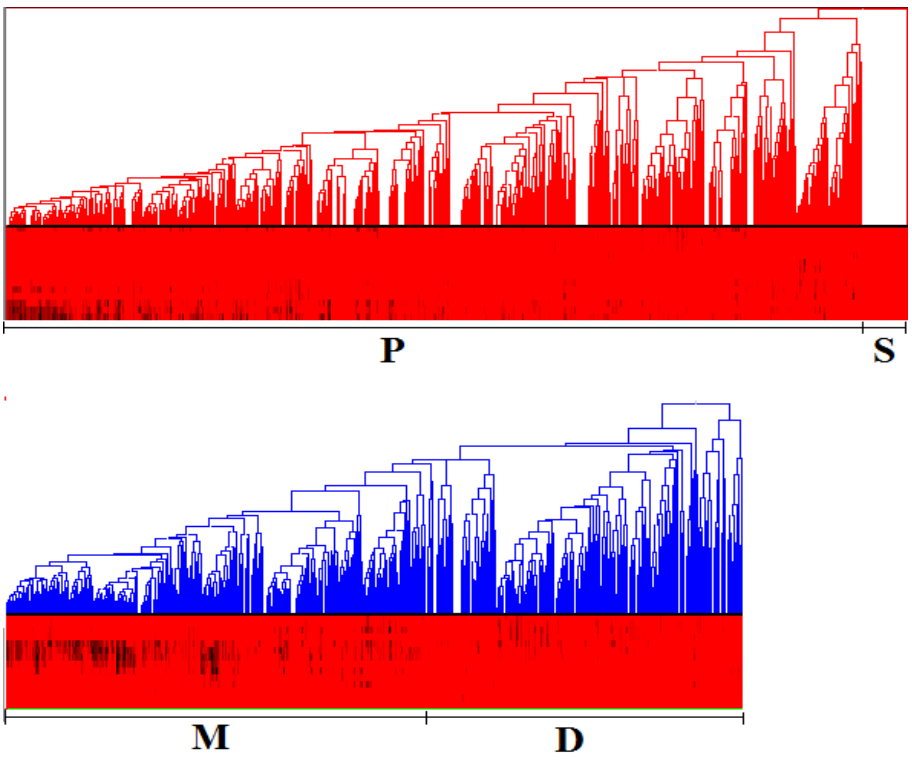

Figure 3.5. Clustering of NP network nodes

Table 3.7. Distribution of Nodes and Anti-correlated edges

	P module	M module
Nodes	1782	1021
Percent of Anti-correlated interactions within clusters	0.98%	0.17%
Percent of Anti-correlated interactions between clusters	97.67%	

3.2.3. Average PCC Values

In order to understand transcriptional relationships between the modules, average PCC values between the modules were calculated as described in Section 2.3.2. An average PCC value *-0.89* between P and M modules indicates an almost perfect anti-correlation. Module P has also a significant (*<-0.7*) anti-correlation with module D as the average PCC between them is *-0.81*. The average PCC value between modules P and S consisting of red colored genes, from the same side of initial hierarchical clustering results were calculated as *+0.45*. Similarly, average PCC between M and D was *+0.58* as expected. Whereas the average PCC between S and D was not significantly minus (*>-0.7*), it is even slightly greater than zero (*+0.11*) (Figure 3.6.). That is why excluding the genes of S and D helped us to get more uniform correlated modules of P and M. In Figure 3.6. Numbers under module names are the number of nodes in the modules. Average PCC values between modules are written on the corresponding lines, within parenthesis. Red lines for negative average PCC values whereas green ones are for positive average PCC values. Dashed lines indicate insignificant correlation or anti-correlation (*-0.7 < PCC < 0.7*)

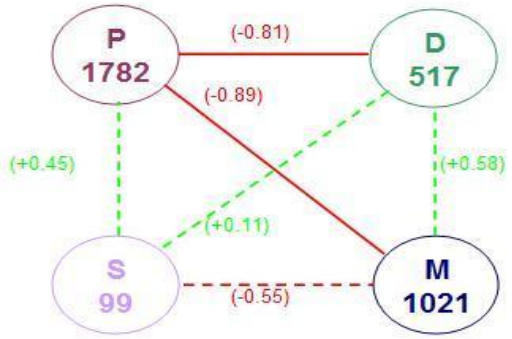

Figure 3.6. Transcriptional relationships among the modules of the NP network.

3.2.4. Average Expression Levels of the Modules

Genes of a module were expected to have similar expression profiles by definition. Therefore expression level of a module was defined in Section 2.3.2. as the arithmetic average of expression levels of genes in the module at a particular sampling point. The expression levels of the modules were used in Section 3.2.3. to calculate average PCC values. It could be useful to visualize expression profiles of the modules to understand correlations and anti-correlations among them (Figure 3.7.).

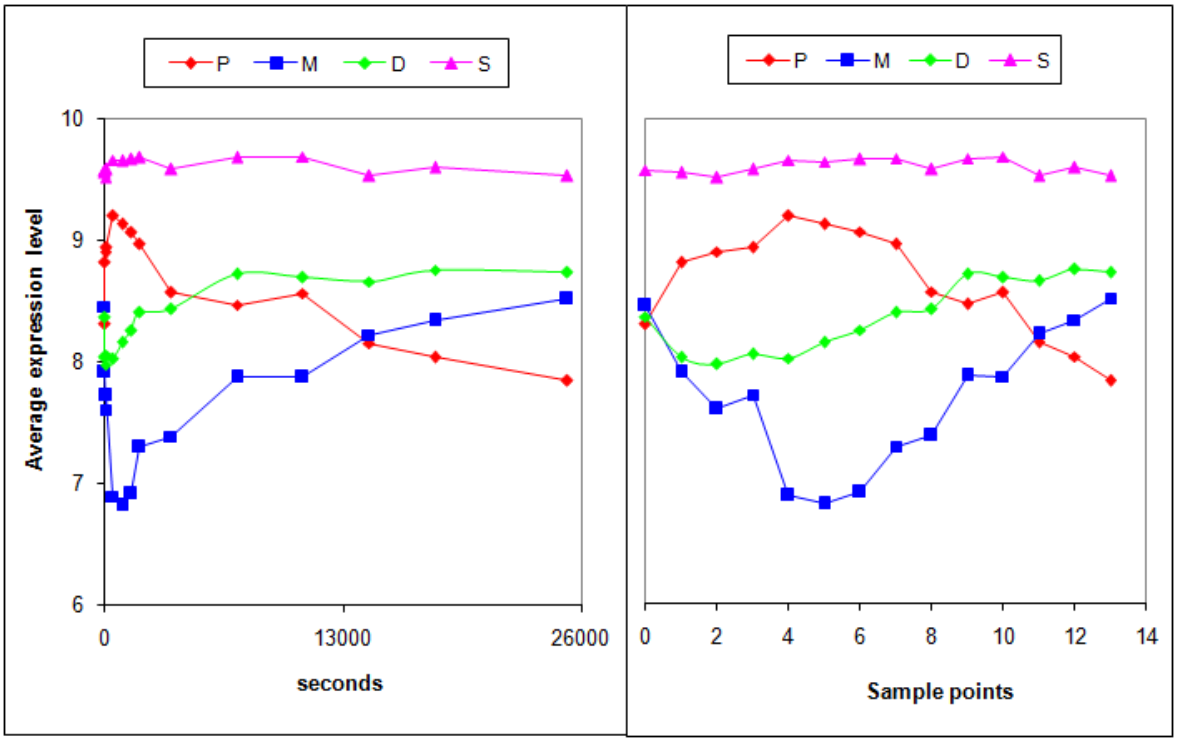

Figure 3.7. Variations in average expression levels of modules with time

To investigate trends of expression levels for modules, slopes and R-square values were calculated. Slopes of fitted trendlines to the expression profiles of modules were used to represent increase or decrease in the expression levels. Slopes were calculated for sample points collected within the first minute, after the first minute and entire sample points to observe changes in the trends. R-square which is a statistical measure of goodness of fit, had closer values to 1 when trendline fits better to data and thus when slopes were significant and it lose sensitivity at low slopes. R^2 values, larger than 0.7, indicate

significance of the slopes and R^2 values less than 0.7 but larger than 0.5 indicate marginally significant slopes ($0.5 < R^2 < 0.7$) (Table 3.8.).

The expression level of M decreased and that of P increased significantly with time until the 8^{th} minute (Figure 3.6. and Table 3.8.). The expression levels of these two modules changed in opposite directions after 8 minutes (sample point 5 in Figure 3.7.). The expression level of D module displayed similar trends with M module but in a more smooth way (Figure 3.7.). Its increase after 8 minutes was statistically significant whereas its decline at the beginning was marginally significant (Table 3.8.). The expression level of module S was almost constant over entire time. This result may possible indicate equal distribution of correlated and anti-correlated genes with P and M modules within this module. Indeed the average PCC values between S module and other modules were closer to zero compared to other average PCC values (Figure 3.6.), i.e. correlation or anti-correlation were not obvious.

Table 3.8. Statistical significance of the changes in the gene expression levels of the yeast modules.

module	slope (10^{-3})	R-square	condition
P	-0.0453	0.7145	entire period
M	0.0452	0.3986	
S	-0.0018	0.0589	
D	0.03	0.6683	
P	-0.0461	0.6884	minutes and hours
M	0.0522	0.5059	
S	-0.0039	0.3094	
D	0.0244	0.6425	
P	9.87	0.755	seconds
M	-12.557	0.7419	
S	-0.1076	0.009	
D	-4.9802	0.534	

3.2.5. GO Terms and Biological Meanings of the Modules

In order to investigate biological meanings of the NP network and the modules GO biological process terms associated to their genes were determined. Assuming that NP network genes are playing a key role in response to a C-pulse, all GO Biological Process Terms which were associated to any of 3419 genes of the NP network were identified.

Since the identification of significantly enriched GO terms using the tools such as Gene-Mapper was difficult due to the huge sizes of the NP network and its modules. A detection parameter (dp) which was defined by Xia *et al.*, 2006, was used to assign a GO biological process terms to these sub-networks.

A detection parameter larger than the significance threshold which was identified separately for each module was used to define the GO terms by which the module was enriched. Calculation of the detection parameters and significance thresholds were described in Section 2.3.4.

The genes of P, M, D and S modules were found to be enriched in 147, 97, 169 and 137 distinct GO biological process terms, respectively. Then NP network genes were found to be enriched in a total of 299 distinct GO biological process terms (Table A4).

The module P was enriched in GO annotations related to ribosome biogenesis, transcription, and rRNA processing. Metabolic process, oxidation reduction, cellular response to heat and transport were found to be enriched terms in module M.

The module S was enriched in translation, DNA repair, transport, cell cycle, chromatin modification, and glycolysis. The module D was enriched in the terms transport, protein amino acid phosphorylation, and ubiquitin-dependent protein catabolic process. Enrichment in transcription in both S and D modules were significant (Table 3.9.).

Table 3.9. Top 10 detection parameters of GO biological process terms in each module.

Modules / Cut-offs	GO terms	*dp*
Module P dp > 0.000044	ribosome biogenesis	0.0650
	Transcription	0.0231
	regulation of transcription, DNA-dependent	0.0172
	biological_process	0.0137
	Transport	0.0135
	rRNA metabolic process	0.0135
	regulation of translation	0.0125
	Translation	0.0110
	rRNA processing	0.0077
	Cell cycle	0.0038
Module M dp > 0.000057	biological_process	0.0361
	Transport	0.0292
	protein transport	0.0086
	cellular response to heat	0.0079
	metabolic process	0.0062
	oxidation reduction	0.0048
	regulation of transcription, DNA-dependent	0.0032
	ubiquitin-dependent protein catabolic process	0.0031
	Transcription	0.0029
	Translation	0.0026
Module S dp > 0.000183	Translation	0.1253
	Transport	0.0745
	regulation of translation	0.0369
	regulation of transcription, DNA-dependent	0.0331
	Transcription	0.0262
	protein transport	0.0199
	Cell cycle	0.0172
	DNA repair	0.0146
	Glycolysis	0.0123
	chromatin modification	0.0102
Module D dp > 0.000076	Transport	0.0548
	regulation of transcription, DNA-dependent	0.0259
	protein transport	0.0182
	biological_process	0.0174
	Transcription	0.0169
	ubiquitin-dependent protein catabolic process	0.0102
	protein amino acid phosphorylation	0.0098
	regulation of translation	0.0094
	vesicle-mediated transport	0.0083
	Translation	0.0055

P module was found to be enriched specifically in 37 GO terms (Figure 3.8. and Table A4) and RNA modification and maturation of SSU-rRNA from tricistronic rRNA transcript (SSU-rRNA, 5.8S rRNA, LSU-rRNA) terms were found to have the highest detection parameters (Table A4). The four GO terms in which all modules except P module enriched in were glycolysis, replicative cell aging, mitochondrial translation and vacuolar protein catabolic process. There are only two GO terms, lipid metabolic process and cellular respiration, were found to be explicitly associated to modules P and M.

20 GO terms (Figure 3.8.) were identified as specific to module M and include mitochondrion inheritance, tricarboxylic acid cycle, ATP biosysnthetic process, and carbohydrate transport. Another 20 GO terms which were associated with all but except module M include transcription from RNA polymerase II promoter and ribosome biogenesis. There were five M and S module specific GO terms which were shown in red dashed circle in Figure 3.8., namely proton transport, ATP synthesis coupled proton transport, glucose metabolic process, protein targeting to vacuole and mitochondrial genome maintenance (Table A4). Among 14 GO terms, associated with only M and D modules, cell redox hemostasis and energy reserve metabolic process terms were identified.

44 GO terms were found to be specific to S module and at the top of these terms with the highest detection parameter was S phase of mitotic cell cycle (Figure 3.8.). Response to stress, meiosis, autophagy and positive regulation of transcription from RNA polymerase II promoter were found to be among the 15 GO terms in which S module genes were not enriched while the remaining modules were enriched. tRNA export from nucleus and aminoacid biosynthetic process terms were found to be specific to P and S modules within 19 overlapping GO terms.

D module was specifically associated with 57 GO terms including iron ion transport as the highest scoring term. Three GO terms namely protein complex assembly, chromosome segregation and aerobic respiration were associated with all except module D. 17 GO terms including negative regulation of transcription from RNA polymerase II promoter and tRNA metabolic process were identified to be both P and D specific GO terms (Figure 3.8.). The genes of modules S and D were found to be commonly enriched in

eight GO terms which include positive regulation of spindle pole body separation and mitotic sister chromatid segregation terms.

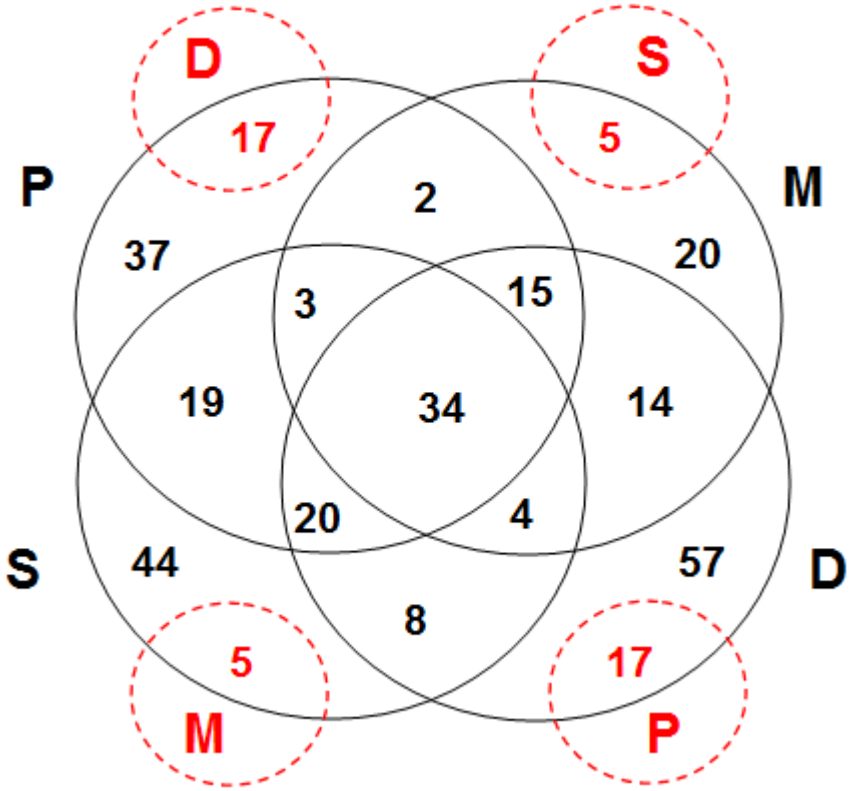

Figure 3.8. Distribution of the 299 identified GO terms among the modules. (Red dashed circles are the intersection of modules standing in opposite sides of the Venn diagram)

3.2.6. GO Slim Process Terms

The devised detection parameter definition used in the present study (Section 2.3.4.) could not consider indirect GO term associations to parental GO terms. Therefore, in order to account indirect associations between terms, GO slim terms were also identified in each module.

Databases do not provide p-values for slim terms by definition; instead they provide cluster frequency as a measure of significance. Cluster frequency is simply the percentage of genes in a module which have directly or indirectly associated to GO slim terms.

GO slim terms which were identified to be with more than 10% cluster frequencies in each module were presented in Table 3.10. Top scoring GO slim process terms for each module were RNA metabolic process in modules P and S, transport in modules M and D.

Table 3.10. GO-Slim terms in each module for cluster frequencies higher than 10%.

Modules	GO-Slim term	Cluster frequency
Module P	RNA metabolic process	622 out of 1782 genes, 34.9%
	transcription	300 out of 1782 genes, 16.8%
	transport	284 out of 1782 genes, 15.9%
	ribosome biogenesis	272 out of 1782 genes, 15.3%
	biological process unknown	206 out of 1782 genes, 11.6%
	translation	187 out of 1782 genes, 10.5%
	protein modification process	184 out of 1782 genes, 10.3%
Module M	transport	218 out of 1021 genes, 21.4%
	biological process unknown	203 out of 1021 genes, 19.9%
	response to stress	126 out of 1021 genes, 12.3%
Module S	RNA metabolic process	28 out of 99 genes, 28.3%
	transport	25 out of 99 genes, 25.3%
	translation	19 out of 99 genes, 19.2%
	transcription	18 out of 99 genes, 18.2%
	cell cycle	16 out of 99 genes, 16.2%
	response to stress	14 out of 99 genes, 14.1%
	DNA metabolic process	14 out of 99 genes, 14.1%
	vesicle-mediated transport	13 out of 99 genes, 13.1%
	protein modification process	12 out of 99 genes, 12.1%
Module D	transport	140 out of 517 genes, 27.1%
	RNA metabolic process	96 out of 517 genes, 18.6%
	transcription	77 out of 517 genes, 14.9%
	cell cycle	70 out of 517 genes, 13.5%
	protein modification process	70 out of 517 genes, 13.5%
	vesicle-mediated transport	65 out of 517 genes, 12.6%
	response to stress	61 out of 517 genes, 11.8%
	biological process unknown	53 out of 517 genes, 10.3%

3.2.7. Interface and Core

In order to investigate regulatory role of inter-modular connections, the term ‘interface’ (Xue *et. al.*, 2007) as described in Section 2.3.5. was used. For each pair of the modules, all nodes were classified into two groups. If a node has any edge which links to the other module of the pair, it was classified as interface node and if not, it was classified

as core gene. Then, distributions of known TFs between interfaces and cores were analyzed to interpret the possible contribution of cores and interfaces to the dynamic regulation of regulatory network. In other words, for each pair of modules, percentages of nodes which were TF were calculated for each interface and core (Figure 3.9.).

All pair-wise module interfaces were found to have higher proportions of TFs compared to the corresponding cores except for M-S and S-D module pairs. The gaps between percentages of TF nodes in interfaces and cores were more obvious for P-D and M-D couples. The only module pair which was found to have higher TF percentage in the core compared to its interface was S-D. TFs were found to be evenly distributed within only M-S modules (Figure 3.9.).

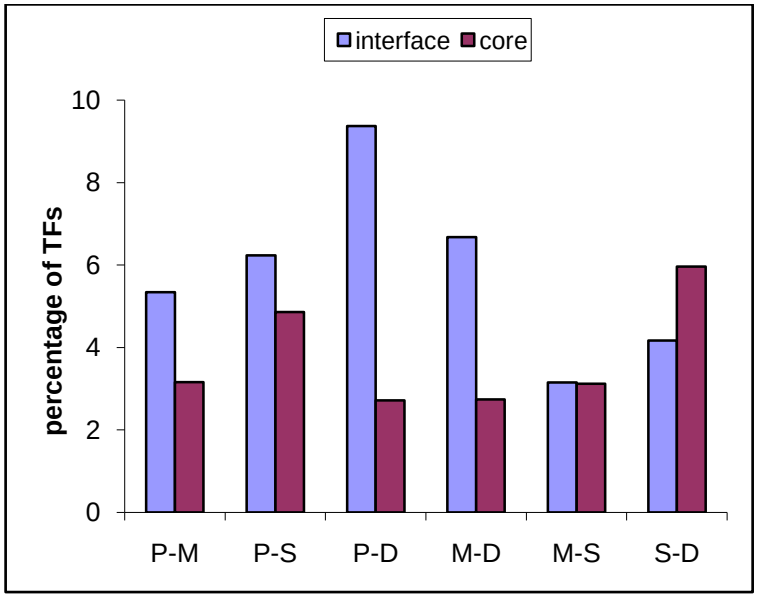

Figure 3.9. TF proportions of the interfaces and cores of module pairs.

The percentages of TFs within the modules were calculated to analyze the modular distribution of these regulatory proteins (Figure 3.10.). The highest TF percentage was found for the module D. The lowest percentages for TFs were found for the modules M and S. This is consistent with the previous results for module M that the detection parameters of transcription and regulation of transcription terms were found to be smaller than that of all other modules indicating also a poor enrichment of the module M for TFs (Table 3.9. and Table 3.10.).

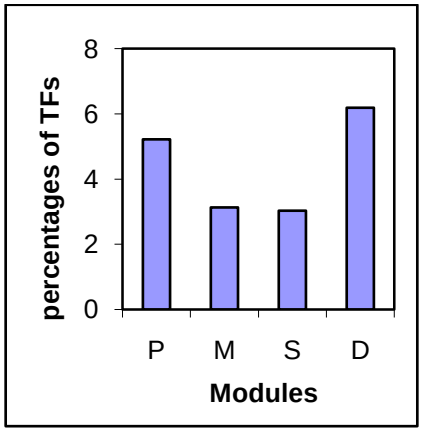

Figure 3.10. Modular distribution of TFs.

GO term analysis described in Section 3.2.5. was applied for interface and core genes to associate them with biological function. However, this analysis provided little information about distinctions between the interfaces and the cores. Similar GO terms for interfaces and cores were identified and the detection parameters for these GO terms were quite similar.

Table 3.11. The GO terms with detection parameters higher than 0.02 in P-M interface.

GO biological process term	*dp* in P-M interface	*dp* in P-M core
ribosome biogenesis	0.0395	0.0291
Translation	0.0255	0.0049
biological_process	0.0233	0.0242
Transport	0.0231	0.0261
Transcription	0.0211	0.0137
regulation of transcription, DNA-dependent	0.0206	0.0106

3.2.8. Degree and Betweenness

In order to analyze the distribution of TFs with higher degrees in both reference and NP networks the percentages of TFs were calculated for each degree and then TF percentages were plotted against degrees in reference and NP networks (Figure 3.11.). In both reference network and the active subnetwork in response to C-pulse (NP network), the percentage of known regulatory nodes – TFs increases with the increasing degree. The percentages of TFs with higher degrees in NP network were found to be larger than that of reference network (Figure 3.11.).

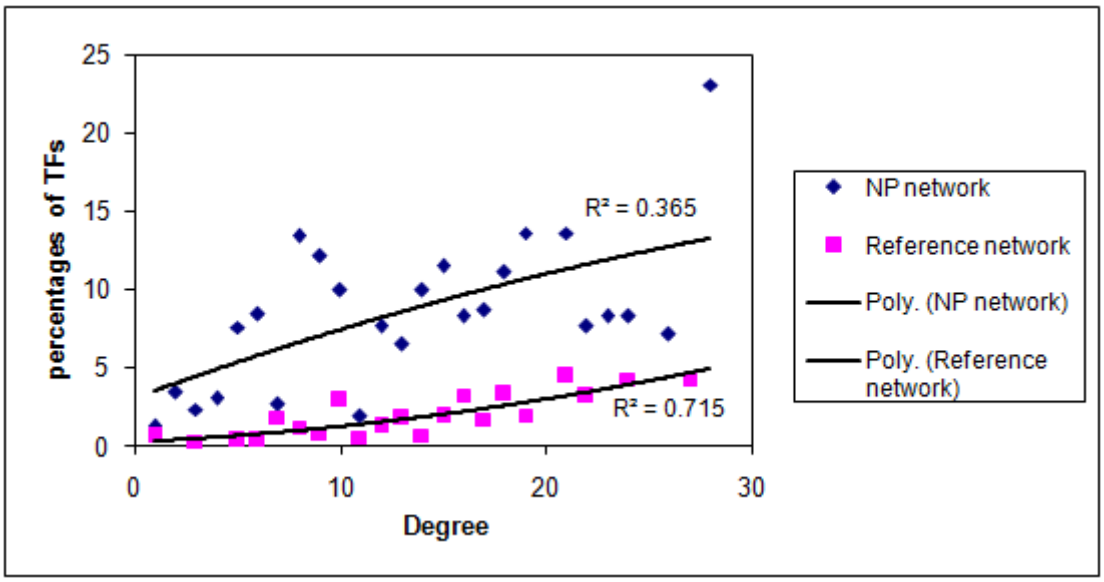

Figure 3.11. Correlation of the percentage of TFs to PPI degree.

Betweenness and degree are highly related with the essentiality of genes in a network. Betweenness for a node is the total number of shortest paths (between any two other nodes) in the network that pass through that node (Valente and Cusick 2006). Thus, percentages of TFs in each betweenness value interval of 2000 were calculated and then plotted against betweenness (Figure 3.12.). The percentage of transcription regulators increases with increasing node betweenness. However, the increase of percentage of TFs with betweenness in NP network is indistinguishable compared to that in reference network.

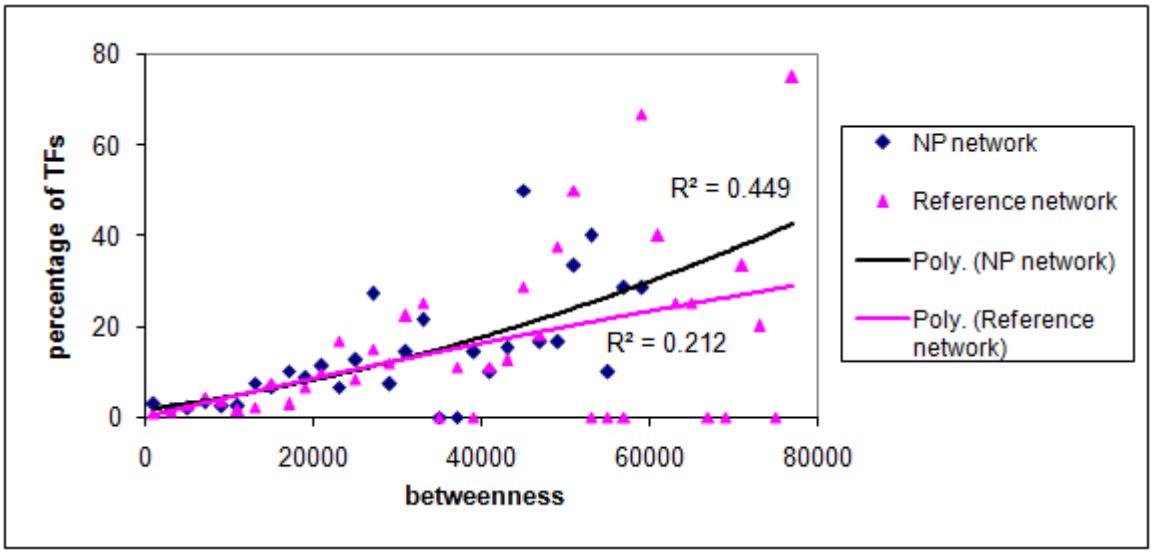

Figure 3.12. Correlation of the percentage of TFs to betweenness

3.2.9. Hubs

The distribution of TF-hubs and non-TF-hubs in reference and NP networks was also investigated to analyze the dynamic regulatory changes upon relaxation from C-limited conditions. Hubs were defined as the nodes with the top 20 highest degrees in the present study. However, here to analyze a wider range, hubs were defined as the top 10% of the highest degree nodes. 345 nodes were identified as hubs for NP network and 553 nodes were identified as hubs for the reference network. 19.7% and 27.7% of hubs were TFs in NP and reference networks respectively. Total degree of TF-hubs was found to be higher in reference as expected, that of NP network decreased by 76%. For the total degree of non-TF-hubs that drop is only 47% for NP network. These results may indicate that TF-hubs of the reference network have no links to neither correlated or anti-correlated nodes.

Hubs can also be categorized as ‘date’ and ‘party’ hubs. Party hubs interact with their links in short term or directly, whereas date hubs interact with their links in long term or indirectly. The types of the hubs can be estimated by their average PCC (AvgPCC) values which were defined in Section 2.3.7. (In Section 3.2.3. the term average PCC were used to represent relation between two modules whereas hereafter it will be noted as AvgPCC to prevent confusion). It has been shown that a high absolute AvgPCC value was sign of party hubs, and a low absolute AvgPCC value indicates date hubs (Han *et al*, 2004; Bertin *et al*, 2007).

In order to investigate the presence of date and party hubs in NP and reference networks, AvgPCC values for each hub was calculated and plotted against their degrees (Figure 3.13.).

AvgPCC values are spread over a wider range for NP network hubs. Average of AvgPCC values of hubs in NP network is (0.155) higher than that of reference network (0.039). This is also consistent with the fact that correlated edges are the majority of NP network edges (Table 3.6.).

Although the TF-hubs both in reference and NP networks display relatively lower absolute AvgPCC values with respect to non-TF-hubs, the distinction became more

obvious in NP network (Figure 3.13.). Therefore transcription regulators present in NP network which are preferentially associated with lower absolute AvgPCC can be considered as ‘date hubs’ (Figure 3.13.).

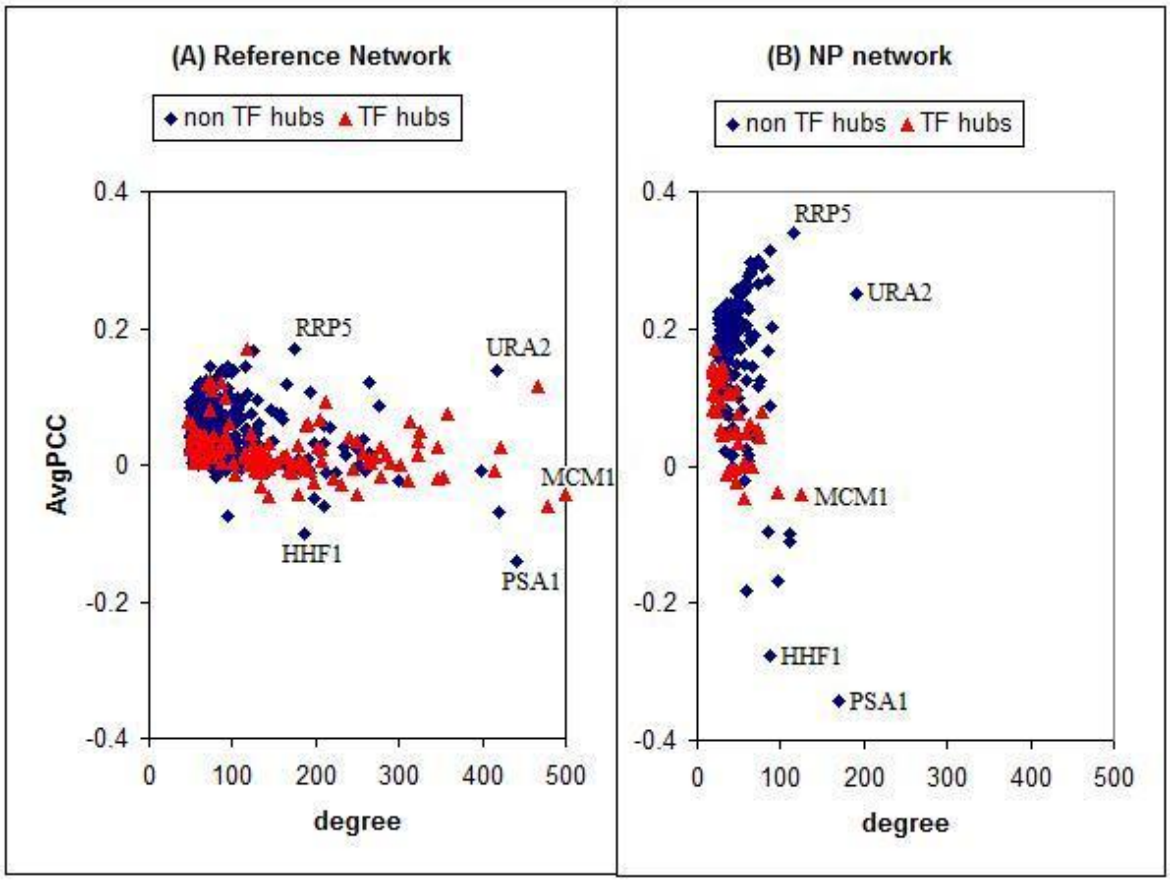

Figure 3.13. Average PCC distribution of TF hubs and non TF hubs with respect to degrees: (A) for Reference network, (B) for NP network.

3.2.10. Distribution of Hubs in NP Network

Distribution of hubs in the NP network and distribution of correlated and anti-correlated edges of the hubs were also analyzed in order to investigate the inter- or intra-modular function of hubs as a response to C-pulse. The top 26 NP network hubs with the highest total degrees were identified (Table 3.12.). Because of the large size of the module P, not surprisingly top 11 hubs with the highest correlated degrees were located in the module P. Whereas, M and D modules were enriched in hubs with higher anti-correlated degrees.

In module P, 12 party and 6 date hubs were identified (Table 3.12.). Three of date hubs and one of the party hubs were identified as TF. In module M, three date and three party hubs were determined. The only TF-hub located in module M was a date hub. Two party hubs were also determined to be located in modules M and S. Therefore, out of 5 TF-hubs which were identified among the top 26 hubs of the NP network, 4 TF-hubs were determined to be date hub.

Hubs with the top two highest anti-correlated degrees and the top two highest correlated degrees were found to be PSA1, HHF1 and URA2, RRP5 respectively (Table 3.12.).

PSA1, which was related to cell wall biosynthesis and has the highest anti-correlated degree (Table 3.12.), placed at the bottom of AvgPCC distribution of NP hubs with a value of *-0.34* (Figure 3.13.). This non-TF-hub was considered to be a party hub and was identified in the module D. These results may indicate that PSA1 is contributing to the communication between P and D modules.

HHF1, related to chromatin assembly, was another party hub with mostly anti-correlated edges – low AvgPCC (Figure 3.13.), located in the module M. These results may indicate that HHF1 is contributing to the communication between P and M modules.

URA2 which was found to have the highest correlated degree, also has relatively high AvgPCC value therefore considered as also a party hub (Figure 3.13.), was related to pyrimidine biosynthetic processes.

RRP5 has a high AvgPCC value (Figure 3.10.), related to rRNA transcription. URA2 and RRP5 were identified to be located in module P (Table 3.12.). These results may suggest that URA2 and RRP5 are contributing to the internal communication of module P.

These details suggest that the hubs of the modules P and S has less anti-correlated PPIs and higher AvgPCC values, whereas that of M and S has more anti-correlated PPIs and lower AvgPCC values.

Table 3.12. Definitions, degrees and module distribution of the top 26 hubs in NP network

Hubs	Total Degree	Correlated Degree	Anti-Correlated Degree	date/party	TF	modules				Definitions
						P	M	S	D	
URA2	190	125	65	party		+				Bifunctional carbamoylphosphate synthetase (CPSase)-aspartate transcarbamylase (ATCase), catalyzes the first two enzymatic steps in the de novo biosynthesis of pyrimidines; both activities are subject to feedback inhibition by UTP
PSA1	170	44	126	party					+	GDP-mannose pyrophosphorylase (mannose-1-phosphate guanyltransferase), synthesizes GDP-mannose from GTP and mannose-1-phosphate in cell wall biosynthesis; required for normal cell wall structure
MCM1	125	67	58	date	+	+				Transcription factor involved in cell-type-specific transcription and pheromone response; plays a central role in the formation of both repressor and activator complexes
RRP5	116	104	12	party		+				RNA binding protein with preference for single stranded tracts of U's involved in synthesis of both 18S and 5.8S rRNAs; component of both the ribosomal small subunit (SSU) processosome and the 90S preribosome
TPK1	112	52	60	party			+			cAMP-dependent protein kinase catalytic subunit; promotes vegetative growth in response to nutrients via the Ras-cAMP signaling pathway; inhibited by regulatory subunit Bcy1p in the absence of cAMP; partially redundant with Tpk2p and Tpk3p
SSA1	111	51	60	date			+			ATPase involved in protein folding and nuclear localization signal (NLS)-directed nuclear transport; member of heat shock protein 70 (HSP70) family; forms a chaperone complex with Ydj1p; localized to the nucleus, cytoplasm, and cell wall
ATG1	96	39	57	party			+			Protein ser/thr kinase required for vesicle formation in autophagy and the cytoplasm-to-vacuole targeting (Cvt) pathway; structurally required for phagophore assembly site formation; during autophagy forms a complex with Atg13p and Atg17p
FKH2	96	52	44	date	+	+				Forkhead family transcription factor with a major role in the expression of G2/M phase genes; positively regulates transcriptional elongation; negative role in chromatin silencing at HML and HMR; substrate of the Cdc28p/Clb5p kinase
PAB1	90	74	16	party		+				Poly(A) binding protein, part of the 3'-end RNA-processing complex, mediates interactions between the 5' cap structure and the 3' mRNA poly(A) tail, involved in control of poly(A) tail length, interacts with translation factor eIF-4G

Hubs	Total Degree	Correlated Degree	Anti-Correlated Degree	date/party	TF	modules				Definitions
						P	M	S	D	
HHF1	88	20	68	party			+			Histone H4, core histone protein required for chromatin assembly and chromosome function; one of two identical histone proteins (see also HHF2); contributes to telomeric silencing; N-terminal domain involved in maintaining genomic integrity
PWP2	87	84	3	party		+				Conserved 90S pre-ribosomal component essential for proper endonucleolytic cleavage of the 35 S rRNA precursor at A0, A1, and A2 sites; contains eight WD-repeats; PWP2 deletion leads to defects in cell cycle and bud morphogenesis
KAP123	87	60	27	date		+				Karyopherin beta, mediates nuclear import of ribosomal proteins prior to assembly into ribosomes and import of histones H3 and H4; localizes to the nuclear pore, nucleus, and cytoplasm; exhibits genetic interactions with RAI1
HSP82	86	39	47	date			+			Hsp90 chaperone required for pheromone signaling and negative regulation of Hsf1p; docks with Tom70p for mitochondrial preprotein delivery; promotes telomerase DNA binding and nucleotide addition; interacts with Cns1p, Cpr6p, Cpr7p, Sti1p
NOP1	84	79	5	party		+				Nucleolar protein, component of the small subunit processome complex, which is required for processing of pre-18S rRNA; has similarity to mammalian fibrillarin
RPL5	84	70	14	party		+				Protein component of the large (60S) ribosomal subunit with similarity to E. coli L18 and rat L5 ribosomal proteins; binds 5S rRNA and is required for 60S subunit assembly
NOP7	78	74	4	party		+				Component of several different pre-ribosomal particles; forms a complex with Ytm1p and Erb1p that is required for maturation of the large ribosomal subunit; required for exit from G0 and the initiation of cell proliferation
GCR2	78	53	25	date	+	+				Transcriptional activator of genes involved in glycolysis; interacts and functions with the DNA-binding protein Gcr1p
RIM101	76	48	28	date		+				Transcriptional repressor involved in response to pH and in cell wall construction; required for alkaline pH-stimulated haploid invasive growth and sporulation; activated by proteolytic processing; similar to A. nidulans PacC
SUA7	76	57	19	party	+	+				Transcription factor TFIIB, a general transcription factor required for transcription initiation and start site selection by RNA polymerase II
CIC1	74	73	1	party		+				Essential protein that interacts with proteasome components and has a potential role in proteasome substrate specificity; also copurifies with 66S pre-ribosomal particles

Hubs	Total Degree	Correlated Degree	Anti-Correlated Degree	date/party	TF	modules				Definitions
						P	M	S	D	
UTP22	73	70	3	party		+				Possible U3 snoRNP protein involved in maturation of pre-18S rRNA, based on computational analysis of large-scale protein-protein interaction data
MSN4	73	46	27	date	+		+			Transcriptional activator related to Msn2p; activated in stress conditions, which results in translocation from the cytoplasm to the nucleus; binds DNA at stress response elements of responsive genes, inducing gene expression
RPL4A	72	54	18	party		+				N-terminally acetylated protein component of the large (60S) ribosomal subunit, nearly identical to Rpl4Bp and has similarity to E. coli L4 and rat L4 ribosomal proteins
RPL7A	69	59	10	party				+		Protein component of the large (60S) ribosomal subunit, nearly identical to Rpl7Bp and has similarity to E. coli L30 and rat L7 ribosomal proteins; contains a conserved C-terminal Nucleic acid Binding Domain (NDB2)
KRR1	67	67	0	party		+				Essential nucleolar protein required for the synthesis of 18S rRNA and for the assembly of 40S ribosomal subunit
ADH1	67	41	26	date		+				Alcohol dehydrogenase, fermentative isozyme active as homo- or heterotetramers; required for the reduction of acetaldehyde to ethanol, the last step in the glycolytic pathway

4. DISCUSSION

In TEAFS 216 distinct modules were identified as active. The most active modules during overall time span were found to be related to ribosome biogenesis, transcription, rRNA processing, regulation of transcription and translation.

Module related to regulation of transcription from RNA polymerase II promoter was identified in all distinct time spans. Biological process term protein ubiquitination during ubiquitin-dependent protein catabolic process related module however, was identified as seconds and minutes specific. Three modules related to ergosterol biosynthetic process, steroid biosynthetic process, and Golgi to endosome transport were identified as seconds specific active modules. Another three modules related to axial cellular bud site selection, lipid biosynthetic process and protein insertion into ER membrane were detected as minutes specific active modules.

In NP analysis four modules of similar expression profiles were determined by hierarchical clustering of expression profiles. The expression level of module M decreased and that of module P found to be increased significantly with time until the 8^{th} minute. The expression levels of these two modules changed in opposite directions after 8^{th} minute. The expression level of D module displayed similar trends with M module but in a more smooth way. The expression level of module S was almost constant over entire time.

The genes of P, M, D and S modules were found to be enriched in 147, 97, 169 and 137 distinct GO biological process terms, respectively. Then NP network genes were found to be enriched in a total of 299 distinct GO biological process terms.

The module P was enriched in GO annotations related to ribosome biogenesis, transcription, and rRNA processing. Metabolic process, oxidation reduction, cellular response to heat and transport were found to be enriched terms in module M. The module S was enriched in translation, DNA repair, transport, cell cycle, chromatin modification, and glycolysis. The module D was enriched in the terms transport, protein amino acid phosphorylation, and ubiquitin-dependent protein catabolic process. Enrichment in

transcription in both S and D modules were significant. PSA1 which is related to cell wall biosynthesis was determined to be contributing to the communication between P and D modules and HHF1 which is related to chromatin assembly was identified to be contributing to the communication between P and M modules.

The total numbers of nodes in time dependent condition specific networks identified by TEAFS varied between 3106 and 3595. In NP analysis, number of nodes was determined as 3419 in the NP network. The numbers of edges were identified to be between 21467 and 30857 in the condition specific networks by TEAFS and 12286 edges were determined in the NP network.

216 and 299 distinct biological processes were identified as active after a carbon pulse by TEAFS and NP analysis respectively. 152 GO biological process terms was identified as active by using both approaches. Therefore, 70% of GO terms identified by TEAFS were also identified by NP analysis.

The most active modules during overall time span were found to be related to ribosome biogenesis, transcription, rRNA processing, regulation of transcription and translation by TEAFS. Ribosome biogenesis and rRNA processing were identified as up-regulated module P specific GO terms by NP analysis. Upregulation of ribosome biogenesis and rRNA processing have verified by various studies (Brink *et. al.,* 2008), (Fazio *et. al.,* 2008), (Zaman *et. al.,* 2009), (Wang *et. al.,* 2004), (Kresnowati *et. al.,* 2006).

All modules were determined to be enriched in transcription by NP analysis. However, enrichment of transcription in modules P, S and D were found to have higher detection parameters than that of module M. Most of the genes enriched in transcription were identified as up-regulated while some other down-regulated as reported in a previous study (Kresnowati *et. al.,* 2006).In another study it was reported that most of the genes enriched in transcription were identified as unchanged and remaining genes upregulated (Wang *et. al.,* 2004).Therefore it can be concluded that our results are in good agreement with both studies.

All modules were also determined to be enriched in translation by NP analysis and enrichment in translation was found to be the most significant in module S. Genes enriched in translation were identified as up-regulated in several studies (Fazio *et. al.*, 2008), (Wang *et. al.*, 2004), (Kresnowati *et. al.*, 2006). Wang *et. al.*, identified genes which are related to translation as up-regulated within only the first 20 minutes in response to glucose (Wang *et. al.*, 2004).

19.9% biological process unknown slim term enrichment of the down-regulated module M was identified by NP analysis. Fazio *et. al.*, reported that 22.7% of down-regulated genes identified with increasing growth rate which were also found to be related to unknown biological process (Fazio *et. al.*, 2008). These results are also in good agreement assuming that glucose introduction into a carbon limited mediun provokes growth of the yeast cells

Among GO terms identified by TEAFS, the top 55 GO terms with the highest EDA-LCON scores were also identified by NP analysis. Nucleosome disassembly, protein amino acid acetylation and positive regulation of transcription from RNA polymerase I promoter terms were determined only by TEAFS with the highest EDA-LCON scores as a response to carbon pulse.

The entire top ten GO terms which were identified by NP analysis for each module, except cellular response to heat, oxidation reduction and glycolysis, were also determined by TEAFS. Especially down-regulated modules M and D were found to be enriched in cellular response to heat and oxidation reduction and module S whose expression level was identified as almost unchanged over entire time span was determined to be significantly enriched in glycolysis. Cellular metabolic compound salvage, peptidyl-amino acid modification, biopolymer biosynthetic process, tRNA metabolic process, vacuolar protein catabolic process, energy reserve metabolic process, aerobic respiration, positive regulation of spindle pole body separation and ion transport were also determined by NP analysis with the highest detection parameters while also could not be identified by TEAFS.

Down-regulation of response to stress, vacuolar protein catabolic process, energy reserve metabolic process, aerobic respiration, oxidation reduction and glycolysis were

reported by various studies (Brink *et. al.,* 2008), (Fazio *et. al.,* 2008), (Zaman *et. al.,* 2009), (Wang *et. al.,* 2004), (Kresnowati *et. al.,* 2006).

17 GO terms out of 37 GO terms, which were identified by distinct time spans TEAFS but not by overall time span analysis, were also determined by NP analysis.

In this work, TCA cycle term was identified as module M specific GO term only by NP analysis. In TEAFS, eliminating modules may lead not to identify these terms. For instance, there are eight genes directly associated with carboxylic metabolic process term. The methodology does not focus on differential expression; instead, it requires eliminating less connected modules. In our PPI network there are only two interactions between proteins which correspond to these eight genes. As a result, carboxylic metabolic process eliminated as a module from further work, regardless of differentiality of expressions of related genes.

Module related to regulation of transcription from RNA polymerase II promoter was identified in all distinct time spans. Biological process term protein ubiquitination during ubiquitin-dependent protein catabolic process related module however, was identified as seconds and minutes specific. Module D was determined to be enriched in this term with the highest detection parameter. The study of Fazio *et. al.,* reported also down-regulation of protein ubiquitination during ubiquitin-dependent protein catabolic process (Fazio *et. al.,* 2008).

Yeast cells grown under carbon limited conditions were investigated after a C-pulse by two distinct integrative approaches namely TEAFS and NP analysis. TEAFS was concluded to be weak to identify down-regulated biological processes. TEAFS failed to identify biopolymer biosynthetic process, vacuolar protein catabolic process, energy reserve metabolic process, aerobic respiration and oxidation reduction.

5. RECOMMENDATIONS

Average number of nodes in the top 20 modules was 268.7, whereas for all 216 modules the average was 78.03. Obviously, as the size of a module increases, its EDA score increases as well. Therefore, differential activity is not the only factor which increases EDA score. Therefore in further studies TEAFS can be revised by a normalization procedure for EDA scores.

As discussed in the previous section some important modules were eliminated at the beginning of TEAFS because of the insufficient interaction data. Therefore, integrating metabolic interactions to PPI may prevent to miss this kind of modules.

The percentage of modules identified by EDA-LCC scores, which identified also by EDA-LCON scores was 90.4%. Execution time needed for EDA-LCC score calculations was much more than the other. Therefore, concentrating only to connectivity fluctuations in further studies may help to save time without losing information.

TEAFS results over distinct time spans suggest that only over hours we can get 92 % of the results of the overall data. This might be a result of larger number of sampling points in hours. Results of the NP analysis and previous works in the literature suggest that response to glucose relief happens within the first 30 minutes and then recovery to steady state takes place. Therefore, by using a shorter time series data (maximum 3 hours) noise in the results can be avoided.

Interpreting biological meanings of large groups of genes present in modules was problematic. Devised detection parameter was not considering indirect associations and it was not a probabilistic approach. GO slim term finder which uses a probabilistic approach was useful in these regards but GO slim terms were weak to give details. Therefore, detection parameter can be rearranged to have a probabilistic sense. Total number of known genes associated to GO terms can be used to get p-values instead of detection parameters.

Table A1. Hubs of condition specific networks (“-“ indicates presence but not being a hub for a node).

Node	reference network	20 sec	40 sec	60 sec	8min	16min	24min	32min	1hr	2hr	3hr	4hr	5hr	7hr
MCM1	MCM1	MCM1	MCM1	MCM1	MCM1	MCM1	MCM1	MCM1	MCM1	MCM1	MCM1	off	MCM1	Off
SWI4	SWI4	off	off	off	off	off	off	off	off	Off	off	off	Off	SWI4
GAT1	GAT1	off	off	off	GAT1	GAT1	GAT1	GAT1	GAT1	GAT1	GAT1	GAT1	GAT1	GAT1
PSA1	PSA1	PSA1	off	PSA1	off	off	off	off	PSA1	PSA1	PSA1	PSA1	PSA1	PSA1
REB1	REB1	REB1	REB1	REB1	REB1	REB1	REB1	REB1	REB1	REB1	REB1	REB1	REB1	REB1
TEF2	TEF2	TEF2	TEF2	TEF2	TEF2	TEF2	TEF2	TEF2	off	TEF2	TEF2	TEF2	TEF2	TEF2
URA2	URA2	off	off	URA2	URA2	URA2	URA2	URA2	off	Off	off	off	Off	Off
FKH2	FKH2	FKH2	FKH2	FKH2	FKH2	FKH2	FKH2	FKH2	FKH2	FKH2	FKH2	FKH2	FKH2	FKH2
PHO85	PHO85	off	PHO85	PHO85	PHO85	PHO85	PHO85	PHO85	PHO85	PHO85	PHO85	PHO85	PHO85	PHO85
FKH1	FKH1	FKH1	off	FKH1	FKH1	FKH1	FKH1	FKH1	Off	FKH1	FKH1	FKH1	FKH1	FKH1
SKO1	SKO1	SKO1	off	off	SKO1	off	SKO1	SKO1	SKO1	SKO1	SKO1	SKO1	SKO1	SKO1
RAP1	RAP1	RAP1	RAP1	RAP1	RAP1	RAP1	RAP1	RAP1	RAP1	Off	RAP1	RAP1	RAP1	RAP1
CIN5	CIN5	CIN5	CIN5	CIN5	CIN5	CIN5	CIN5	CIN5	CIN5	CIN5	CIN5	CIN5	CIN5	CIN5
ECM22	ECM22	ECM22	ECM22	ECM22	ECM22	ECM22	ECM22	ECM22	ECM22	ECM22	ECM22	ECM22	ECM22	ECM22
SWI5	SWI5	SWI5	SWI5	SWI5	SWI5	SWI5	SWI5	SWI5	SWI5	SWI5	SWI5	SWI5	SWI5	SWI5
SWI6	SWI6	off	off	off	off	off	off	off	off	Off	off	off	Off	Off
CBF1	CBF1	CBF1	CBF1	CBF1	CBF1	CBF1	CBF1	CBF1	CBF1	CBF1	CBF1	off	CBF1	CBF1
HSF1	HSF1	off	off	off	off	off	off	off	HSF1	HSF1	HSF1	HSF1	HSF1	HSF1

Table A1. Hubs of condition specific networks ("-" indicates presence but not being a hub for a node).

Node	reference network	20 sec	40 sec	60 sec	8min	16min	24min	32min	1hr	2hr	3hr	4hr	5hr	7hr
CHA4	CHA4	off	off	off	off	off	off	off	off	off	off	off	off	Off
SSA1	SSA1	off	off	off	off	off	off	off	off	off	off	off	off	SSA1
MIG1	-	-	-	-	-	off	off	off	off	-	-	MIG1	-	MIG1
UGA3	-	UGA3	UGA3	UGA3	UGA3	UGA3	UGA3	UGA3	UGA3	UGA3	-	off	off	UGA3
DAL81	-	-	DAL81	DAL81	-	-	-	DAL81	DAL81	-	-	DAL81	DAL81	-
DAL82	-	DAL82	DAL82	DAL82	DAL82	DAL82	DAL82	DAL82	DAL82	DAL82	DAL82	DAL82	DAL82	DAL82
SMT3	-	off	SMT3	SMT3	SMT3	SMT3	SMT3	SMT3	SMT3	SMT3	SMT3	SMT3	SMT3	off
YAP7	-	YAP7	YAP7	-	-	YAP7	-	-	-	YAP7	-	YAP7	off	-
STE12	-	STE12	STE12	STE12	STE12	STE12	STE12	STE12	STE12	-	-	STE12	STE12	STE12
TUB2	-	TUB2	TUB2	TUB2	TUB2	TUB2	TUB2	-	TUB2	TUB2	TUB2	-	-	-
MAL33	-	-	-	-	-	-	-	-	-	-	-	-	-	-
RPP0	-	RPP0	RPP0	RPP0	RPP0	RPP0	RPP0	RPP0	RPP0	RPP0	RPP0	RPP0	-	-
TPK1	-	TPK1	TPK1	-	-	-	-	-	-	-	-	-	-	-
TUB1	-	-	-	-	-	-	-	-	off	-	TUB1	-	TUB1	-
YAP1	-	off	off	off	off	off	off	off	off	off	off	off	off	off
JSN1	-	-	-	-	off	off	off	off	-	-	-	-	-	-
GLN3	-	GLN3	GLN3	-	-	-	-	-	-	-	-	-	-	-

Table A2. All modules identified based on TEAFS method with the order of descending EDA-LCON scores. Identification statuses over distinct time spans are also provided.

EDA – LCON	EDA – LCC	p-value Connectivity	p-value Clustering Coefficient	FDR q-value Connectivity	FDR q-value Clustering Coefficient	Number of Proteins in the Module	Average Clustering Coefficient in the Reference Network	Sec	min	hr	Biological Process Module Definition
1440.1	29.49	0	0	0	0	697	0.25	-	-	OK	ribosome biogenesis
701.4	15.68	0	0	0	0	598	0.29	-	-	OK	Transcription
639.4	20.73	0	0	0	0	267	0.52	-	-	OK	rRNA processing
481.1	9.82	0	0	0	0	629	0.27	OK	-	OK	regulation of transcription, DNA-dependent
360.0	6.91	0	0	0	0	633	0.23	-	-	OK	Translation
291.3	8.98	0	0	0	0	175	0.61	-	-	OK	mRNA processing
236.8	6.58	0	0	0	0	114	0.62	-	-	OK	RNA splicing
205.7	6.74	0	0	0	0	119	0.67	-	-	OK	nuclear mRNA splicing, via spliceosome
200.4	7.53	0	0	0	0	404	0.15	OK	-	OK	regulation of translation
187.4	7.70	0	0	0	0	266	0.52	OK	OK	OK	ubiquitin-dependent protein catabolic process
122.7	3.91	0	0	0	0	66	0.63	-	-	OK	transcription from RNA polymerase II promoter
112.0	5.92	0	0	0	0	42	0.63	-	-	OK	endonucleolytic cleavage in ITS1 to separate SSU-rRNA from 5.8S rRNA and LSU-rRNA from tricistronic rRNA transcript (SSU-rRNA, 5.8S rRNA, LSU-rRNA)
102.5	4.19	0	0	0	0	66	0.41	-	-	OK	maturation of SSU-rRNA from tricistronic rRNA transcript (SSU-rRNA, 5.8S rRNA, LSU-rRNA)
101.1	6.10	0	0	0	0	304	0.15	OK	OK	OK	rRNA metabolic process
91.4	6.92	0.003	0.003	0.005	0.006	411	0.41	-	OK	OK	protein transport
89.5	8.37	0.017	0	0.026	0	323	0.32	-	OK	OK	cell cycle
77.2	3.94	0	0	0	0	32	0.62	-	-	OK	endonucleolytic cleavage to generate mature 5'-end of SSU-rRNA from (SSU-rRNA, 5.8S rRNA, LSU-rRNA)

Table A2. All modules identified based on TEAFS method with the order of descending EDA-LCON scores. Identification statuses over distinct time spans are also provided.

EDA – LCON	EDA - LCC	p-value Connectivity	p-value Clustering Coefficient	FDR q-value Connectivity	FDR q-value Clustering Coefficient	Number of Proteins in the Module	Average Clustering Coefficient in the Reference Network	sec	min	hr	Biological Process Module Definition
75.7	5.11	0	0	0	0	50	0.65	-	-	OK	ribosomal large subunit biogenesis
74.7	7.26	0	0	0	0	50	0.47	-	-	OK	ribosomal large subunit assembly
72.5	-	0	-	0	-	128	0.66	-	OK	OK	chromatin modification
67.4	4.97	0	0	0	0	183	0.27	-	-	OK	response to stress
66.9	3.26	0	0	0	0	30	0.60	-	-	OK	endonucleolytic cleavage in 5'-ETS of tricistronic rRNA transcript (SSU-rRNA, 5.8S rRNA, LSU-rRNA)
65.9	4.74	0	0	0	0	57	0.57	OK	OK	OK	mRNA transport
60.3	6.25	0	0	0	0	68	0.51	-	OK	OK	intracellular protein transmembrane transport
58.8	3.60	0	0	0	0	51	0.50	-	-	OK	transcription from RNA polymerase III promoter
57.3	5.91	0	0	0	0	193	0.50	OK	OK	OK	vesicle-mediated transport
56.2	3.45	0	0	0	0	74	0.43	-	-	OK	transcription from RNA polymerase I promoter
55.1	4.99	0.015	0	0.024	0	217	0.37	-	OK	OK	cell division
54.7	3.81	0	0	0	0	58	0.55	-	-	OK	translational initiation
53.8	3.25	0	0	0	0	32	0.78	-	-	OK	exonucleolytic trimming to generate mature 3'-end of 5.8S rRNA from tricistronic rRNA transcript (SSU-rRNA, 5.8S rRNA, LSU-rRNA)
53.2	4.46	0	0	0	0	83	0.47	OK	OK	OK	mRNA export from nucleus
52.5	5.57	0	0	0	0	51	0.49	OK	OK	OK	establishment of protein localization
50.5	3.26	0	0	0	0	231	0.28	-	OK	OK	response to DNA damage stimulus
49.0	3.20	0	0.003	0	0.006	288	0.29	OK	OK	OK	DNA repair
48.2	4.19	0	0	0	0	61	0.70	OK	-	OK	RNA elongation from RNA polymerase II promoter
46.1	2.74	0	0	0	0	25	0.88	-	-	OK	RNA fragment catabolic process
45.2	3.63	0	0	0	0	128	0.47	OK	OK	OK	ER to Golgi vesicle-mediated transport
44.7	5.12	0	0	0	0	138	0.50	-	OK	OK	Mitosis

Table A2. All modules identified based on TEAFS method with the order of descending EDA-LCON scores. Identification statuses over distinct time spans are also provided.

EDA - LCON	EDA - LCC	p-value Connectivity	p-value Clustering Coefficient	FDR q-value Connectivity	FDR q-value Clustering Coefficient	Number of Proteins in the Module	Average Clustering Coefficient in the Reference Network	sec	min	hr	Biological Process Module Definition
43.0	3.47	0	0	0	0	36	0.56	OK	OK	OK	protein export from nucleus
42.8	4.03	0.001	0.001	0.002	0.002	74	0.72	-	-	OK	chromatin remodeling
41.3	3.65	0	0	0	0	34	0.60	OK	OK	OK	mRNA-binding (hnRNP) protein import into nucleus
39.8	1.90	0	0.01	0	0.019	104	0.28	-	OK	OK	Endocytosis
39.7	2.74	0	0.003	0	0.006	106	0.13	OK	-	OK	amino acid biosynthetic process
38.7	2.24	0.02	0.028	0.03	0.046	103	0.57	-	OK	-	transcription initiation from RNA polymerase II promoter
38.4	4.47	0	0	0	0	144	0.36	OK	OK	OK	DNA replication
38.3	3.82	0	0	0	0	31	0.67	OK	OK	OK	nuclear pore organization
35.7	3.37	0	0	0	0	47	0.54	OK	OK	OK	rRNA export from nucleus
35.0	2.29	0	0	0	0	48	0.43	-	OK	OK	protein import into nucleus
34.9	-	0	-	0	-	367	0.06	OK	-	OK	metabolic process
34.7	3.43	0	0	0	0	24	0.66	OK	OK	OK	snRNA export from nucleus
34.7	3.43	0	0	0	0	25	0.66	OK	OK	OK	NLS-bearing substrate import into nucleus
34.7	3.43	0	0	0	0	24	0.66	OK	OK	OK	snRNP protein import into nucleus
34.7	3.43	0	0	0	0	24	0.66	OK	OK	OK	ribosomal protein import into nucleus
34.4	3.00	0	0	0	0	37	0.49	OK	OK	OK	tRNA export from nucleus
34.2	2.08	0	0	0	0	18	1.00	-	-	OK	nucleosome disassembly
33.4	2.19	0	0	0	0	32	0.82	-	OK	OK	ATP-dependent chromatin remodeling
29.8	2.66	0	0	0	0	21	0.87	-	-	OK	nuclear-transcribed mRNA catabolic process, exonucleolytic, 3'-5'
28.9	2.73	0	0.003	0	0.006	163	0.22	OK	OK	OK	Meiosis
28.6	3.02	0	0.001	0	0.002	37	0.86	OK	OK	OK	mRNA cleavage
28.0	-	0.016	-	0.025	-	65	0.79	OK	-	-	histone acetylation

Table A2. All modules identified based on TEAFS method with the order of descending EDA-LCON scores. Identification statuses over distinct time spans are also provided.

EDA – LCON	EDA - LCC	p-value Connectivity	p-value Clustering Coefficient	FDR q-value Connectivity	FDR q-value Clustering Coefficient	Number of Proteins in the Module	Average Clustering Coefficient in the Reference Network	sec	min	hr	Biological Process Module Definition
26.8	3.83	0	0	0	0	48	0.59	OK	OK	OK	DNA replication initiation
24.9	-	0	-	0	-	94	0.55	-	OK	OK	mitochondrial translation
24.9	2.06	0	0.01	0	0.018	36	0.70	-	OK	OK	mRNA polyadenylation
21.6	2.26	0	0	0	0	129	0.16	OK	OK	OK	carbohydrate metabolic process
21.3	2.54	0	0	0	0	33	0.45	-	-	OK	RNA processing
20.8	2.95	0	0	0	0	56	0.34	OK	OK	OK	actin filament organization
19.3	4.04	0	0	0	0	96	0.49	-	OK	OK	intracellular protein transport
18.4	-	0.014	-	0.022	-	17	0.91	-	-	OK	protein amino acid acetylation
17.7	3.03	0	0	0	0	10	1.00	-	OK	OK	positive regulation of transcription from RNA polymerase I promoter
17.4	2.68	0	0	0	0	29	0.58	-	-	OK	protein catabolic process
17.1	-	0.03	-	0.044	-	40	0.47	-	-	-	methionine biosynthetic process
16.7	-	0	-	0	-	84	0.51	OK	OK	OK	chromatin silencing at telomere
16.7	2.90	0	0	0	0	83	0.47	-	OK	OK	chromosome segregation
16.5	1.21	0.004	0.012	0.007	0.022	10	1.00	OK	-	-	transcriptional preinitiation complex assembly
16.4	1.05	0.033	0.007	0.047	0.013	63	0.18	OK	OK	OK	negative regulation of transcription from RNA polymerase II promoter
16.3	-	0	-	0	-	148	0.01	-	-	OK	RNA modification
16.3	2.11	0	0.003	0	0.006	104	0.42	OK	OK	OK	tRNA processing
16.0	2.12	0	0	0	0	99	0.29	OK	OK	OK	positive regulation of transcription from RNA polymerase II promoter
15.9	3.40	0	0	0	0	25	0.72	OK	OK	OK	pre-replicative complex assembly
15.8	2.47	0	0.005	0	0.01	47	0.69	-	OK	OK	mitotic sister chromatid segregation
15.8	2.30	0	0	0	0	25	0.68	OK	OK	OK	mRNA catabolic process
15.7	-	0	-	0	-	43	0.74	-	OK	OK	retrograde vesicle-mediated transport, Golgi to ER

Table A2. All modules identified based on TEAFS method with the order of descending EDA-LCON scores. Identification statuses over distinct time spans are also provided.

EDA – LCON	EDA - LCC	p-value Connectivity	p-value Clustering Coefficient	FDR q-value Connectivity	FDR q-value Clustering Coefficient	Number of Proteins in the Module	Average Clustering Coefficient in the Reference Network	sec	min	hr	Biological Process Module Definition
15.1	2.61	0	0	0	0	113	0.22	OK	OK	OK	response to drug
15.0	-	0	-	0	-	39	0.48	OK	OK	OK	protein amino acid N-linked glycosylation
14.7	1.89	0	0	0	0	56	0.18	OK	OK	OK	response to osmotic stress
14.6	1.90	0	0.017	0	0.029	47	0.37	OK	OK	-	chromatin silencing at silent mating-type cassette
14.4	0.78	0.009	0.021	0.015	0.035	71	0.07	-	OK	OK	regulation of transcription
14.3	-	0.018	-	0.028	-	59	0.58	OK	-	OK	protein ubiquitination
13.7	2.43	0	0	0	0	30	0.67	OK	-	OK	nucleosome assembly
13.6	2.18	0	0	0	0	38	0.52	-	OK	OK	bipolar cellular bud site selection
13.5	-	0.02	-	0.03	-	32	0.51	-	-	OK	tRNA modification
13.3	2.81	0	0	0	0	12	0.91	-	OK	OK	snoRNA 3'-end processing
13.3	-	0	-	0	-	222	0.15	-	-	OK	cell wall organization
13.1	-	0.001	-	0.002	-	21	1.00	-	-	OK	anaphase-promoting complex-dependent proteasomal ubiquitin-dependent protein catabolic process
12.5	2.55	0.003	0.004	0.005	0.008	17	0.88	-	-	OK	U4 snRNA 3'-end processing
12.5	2.18	0.006	0.013	0.01	0.023	22	0.55	-	-	OK	ribosomal small subunit biogenesis
12.2	1.86	0	0	0	0	36	0.33	OK	OK	OK	telomere maintenance
12.0	-	0	-	0	-	17	0.76	-	-	OK	transcription initiation
11.1	1.64	0	0	0	0	57	0.34	OK	OK	OK	invasive growth in response to glucose limitation
11.0	2.03	0	0	0	0	63	0.30	OK	OK	OK	protein complex assembly
11.0	1.17	0.004	0.016	0.007	0.028	13	0.96	-	OK	OK	termination of RNA polymerase II transcription, poly(A)-coupled
10.7	2.58	0	0	0	0	44	0.59	-	OK	OK	nucleotide-excision repair
10.4	3.29	0	0	0	0	49	0.52	OK	OK	OK	response to pheromone

Table A2. All modules identified based on TEAFS method with the order of descending EDA-LCON scores. Identification statuses over distinct time spans are also provided.

EDA – LCON	EDA - LCC	p-value Connectivity	p-value Clustering Coefficient	FDR q-value Connectivity	FDR q-value Clustering Coefficient	Number of Proteins in the Module	Average Clustering Coefficient in the Reference Network	sec	min	hr	Biological Process Module Definition
10.3	-	0.001	-	0.002	-	81	0.35	OK	OK	OK	pseudohyphal growth
10.2	1.17	0.004	0.016	0.007	0.028	16	0.77	-	OK	OK	termination of RNA polymerase II transcription, poly(A)-independent
9.8	-	0.02	-	0.031	-	184	0.11	-	OK	OK	Proteolysis
9.7	-	0	-	0	-	18	0.50	OK	-	-	protein targeting to ER
9.2	1.46	0	0	0	0	48	0.33	OK	-	OK	pheromone-dependent signal transduction involved in conjugation with cellular fusion
9.2	-	0	-	0	-	38	0.48	-	-	OK	snoRNA metabolic process
9.1	1.59	0	0.001	0	0.002	7	0.93	-	-	OK	ncRNA 3'-end processing
8.9	-	0	-	0	-	31	0.26	-	OK	OK	nucleus organization
8.8	-	0	-	0	-	7	1.00	OK	-	OK	SRP-dependent cotranslational protein targeting to membrane, signal sequence recognition
8.7	1.54	0	0.015	0	0.027	35	1.00	OK	OK	OK	histone H3-K4 methylation
8.7	1.81	0	0	0	0	42	0.55	-	OK	OK	mismatch repair
8.5	1.84	0	0	0	0	21	0.73	-	OK	OK	S phase of mitotic cell cycle
8.5	-	0.003	-	0.005	-	16	0.31	-	-	OK	maturation of LSU-rRNA from tricistronic rRNA transcript (SSU-rRNA, 5.8S rRNA, LSU-rRNA)
8.3	1.26	0	0.01	0	0.019	42	0.19	OK	-	OK	actin cytoskeleton organization
8.3	-	0	-	0	-	22	0.67	OK	OK	OK	mitotic cell cycle
8.3	1.61	0	0	0	0	39	0.58	OK	-	OK	protein import into mitochondrial matrix
8.2	2.05	0	0	0	0	34	0.71	-	-	OK	regulation of translational initiation
8.1	2.57	0	0	0	0	73	0.22	-	OK	OK	ascospore formation
8.1	2.20	0	0	0	0	5	1.00	-	-	OK	regulation of actin filament polymerization

Table A2. All modules identified based on TEAFS method with the order of descending EDA-LCON scores. Identification statuses over distinct time spans are also provided.

EDA – LCON	EDA - LCC	p-value Connectivity	p-value Clustering Coefficient	FDR q-value Connectivity	FDR q-value Clustering Coefficient	Number of Proteins in the Module	Average Clustering Coefficient in the Reference Network	sec	min	hr	Biological Process Module Definition
8.0	0.69	0	0	0	0	56	0.19	OK	OK	OK	reciprocal meiotic recombination
7.9	1.59	0	0	0	0	12	0.69	OK	OK	-	posttranslational protein targeting to membrane, translocation
7.9	0.51	0	0	0	0	30	0.22	-	-	OK	RNA elongation
7.7	-	0.001	-	0.002	-	38	0.15	OK	OK	OK	nucleocytoplasmic transport
7.6	1.59	0	0	0	0	14	0.30	-	-	OK	actin cortical patch assembly
7.3	-	0	-	0	-	13	0.58	-	OK	-	RNA 3'-end processing
7.1	-	0	-	0	-	62	0.45	OK	-	OK	DNA recombination
7.1	1.07	0	0	0	0	9	0.85	-	OK	OK	intronic box C/D snoRNA processing
7.0	1.36	0	0	0	0	16	0.69	-	-	OK	leading strand elongation
7.0	-	0	-	0	-	23	0.29	OK	OK	OK	nuclear transport
6.9	2.11	0	0	0	0	23	0.91	-	-	OK	retrograde transport, endosome to Golgi
6.7	1.71	0	0	0	0	20	0.53	OK	OK	OK	protein sumoylation
6.5	-	0	-	0	-	40	0.38	OK	OK	OK	vacuole fusion, non-autophagic
6.5	-	0	-	0	-	13	0.19	-	-	OK	maturation of 5.8S rRNA from tricistronic rRNA transcript (SSU-rRNA, 5.8S rRNA, LSU-rRNA)
6.4	-	0.007	-	0.011	-	14	0.66	-	-	OK	mRNA metabolic process
6.3	-	0	-	0	-	22	0.77	-	OK	OK	microtubule nucleation
6.1	-	0.026	-	0.038	-	55	0.31	-	-	OK	protein amino acid dephosphorylation
5.9	-	0	-	0	-	14	0.57	-	-	OK	ribosome export from nucleus
5.9	-	0.014	-	0.022	-	12	0.80	-	-	OK	U5 snRNA 3'-end processing
5.7	-	0.033	-	0.047	-	42	0.47	-	-	OK	cytoskeleton organization
5.7	-	0	-	0	-	57	0.46	-	OK	OK	cellular respiration

Table A2. All modules identified based on TEAFS method with the order of descending EDA-LCON scores. Identification statuses over distinct time spans are also provided.

EDA – LCON	EDA - LCC	p-value Connectivity	p-value Clustering Coefficient	FDR q-value Connectivity	FDR q-value Clustering Coefficient	Number of Proteins in the Module	Average Clustering Coefficient in the Reference Network	sec	min	hr	Biological Process Module Definition
5.6	1.02	0	0	0	0	28	0.26	OK	-	OK	amino acid metabolic process
5.6	1.49	0.003	0	0.005	0	23	0.56	-	OK	OK	lagging strand elongation
5.5	1.55	0	0	0	0	15	1.00	OK	OK	OK	positive regulation of transcription, DNA-dependent
5.5	1.77	0	0.007	0	0.013	15	0.84	-	OK	OK	DNA strand elongation during DNA replication
5.3	-	0	-	0	-	9	0.71	-	OK	-	histone methylation
5.1	2.68	0.007	0	0.012	0	17	0.65	OK	OK	OK	cellular ion homeostasis
5.1	1.74	0	0	0	0	19	0.62	-	OK	OK	deadenylation-dependent decapping of nuclear-transcribed mRNA
5.0	-	0.003	-	0.005	-	39	0.18	OK	-	-	biosynthetic process
5.0	1.29	0.025	0	0.037	0	34	0.41	OK	OK	OK	double-strand break repair via nonhomologous end joining
4.9	-	0	-	0	-	9	0.60	-	-	OK	arginine metabolic process
4.6	-	0.013	-	0.021	-	17	0.50	-	-	OK	ubiquitin-dependent protein catabolic process via the multivesicular body sorting pathway
4.6	-	0	-	0	-	28	0.14	-	-	OK	branched chain family amino acid biosynthetic process
4.6	1.75	0	0	0	0	31	0.50	-	OK	OK	DNA-dependent DNA replication
4.5	1.43	0	0	0	0	30	0.74	OK	OK	OK	mitotic spindle organization in nucleus
4.5	1.52	0.003	0	0.005	0	52	0.55	-	-	OK	protein targeting to vacuole
4.4	1.61	0	0	0	0	22	0.55	OK	OK	-	DNA packaging
4.3	-	0	-	0	-	73	0.27	OK	OK	OK	GPI anchor biosynthetic process
4.3	1.62	0	0	0	0	14	0.78	-	-	OK	protein import into nucleus, docking
4.3	-	0.029	-	0.042	-	26	0.33	-	-	OK	budding cell bud growth
4.2	1.14	0	0	0	0	18	0.29	-	-	OK	ribosomal large subunit export from nucleus
4.0	1.92	0	0	0	0	21	0.78	-	OK	OK	vesicle docking during exocytosis

Table A2. All modules identified based on TEAFS method with the order of descending EDA-LCON scores. Identification statuses over distinct time spans are also provided.

EDA – LCON	EDA - LCC	p-value Connectivity	p-value Clustering Coefficient	FDR q-value Connectivity	FDR q-value Clustering Coefficient	Number of Proteins in the Module	Average Clustering Coefficient in the Reference Network	sec	min	hr	Biological Process Module Definition
3.9	1.43	0	0	0	0	20	0.93	OK	OK	OK	regulation of microtubule polymerization or depolymerization
3.9	-	0	-	0	-	28	0.57	-	-	OK	Golgi to vacuole transport
3.8	1.37	0	0.003	0	0.006	17	0.83	-	-	OK	regulation of translational fidelity
3.8	0.94	0	0	0	0	39	0.03	OK	-	OK	regulation of cell cycle
3.8	-	0	-	0	-	11	0.70	-	-	OK	SRP-dependent cotranslational protein targeting to membrane
3.8	1.37	0.003	0.01	0.005	0.018	7	0.67	OK	OK	OK	double-strand break repair via single-strand annealing
3.7	1.69	0.035	0.007	0.05	0.013	9	1.00	OK	-	-	flocculation via cell wall protein-carbohydrate interaction
3.7	1.69	0.011	0	0.018	0	27	0.56	-	-	OK	mitotic cell cycle spindle assembly checkpoint
3.6	-	0.01	-	0.016	-	28	0.53	OK	-	OK	proteasome assembly
3.6	1.13	0.03	0.005	0.043	0.01	14	0.50	-	OK	OK	telomere maintenance via recombination
3.4	1.49	0	0	0	0	6	0.90	-	OK	OK	negative regulation of transcription from RNA polymerase II promoter, mitotic
3.3	1.22	0.023	0.012	0.035	0.022	36	0.13	-	-	OK	protein modification process
3.0	1.14	0.024	0	0.036	0	28	0.23	-	OK	OK	regulation of cyclin-dependent protein kinase activity
2.9	1.58	0	0	0	0	5	1.00	OK	OK	OK	regulation of mitotic metaphase/anaphase transition
2.9	0.49	0	0	0	0	8	1.00	-	OK	OK	sister chromatid cohesion
2.8	-	0	-	0	-	14	0.59	-	-	OK	budding cell apical bud growth
2.7	1.45	0	0	0	0	12	0.68	-	OK	OK	transcription-coupled nucleotide-excision repair
2.6	-	0.02	-	0.03	-	7	1.00	-	-	OK	U1 snRNA 3'-end processing
2.6	1.46	0	0	0	0	11	1.00	OK	-	OK	spliceosome assembly
2.6	1.46	0	0	0	0	6	1.00	-	OK	OK	negative regulation of meiosis
2.5	1.26	0	0	0	0	15	1.00	-	-	OK	cellular metabolic process
2.4	-	0.025	-	0.037	-	12	0.39	-	-	-	interphase

Table A2. All modules identified based on TEAFS method with the order of descending EDA-LCON scores. Identification statuses over distinct time spans are also provided.

EDA – LCON	EDA - LCC	p-value Connectivity	p-value Clustering Coefficient	FDR q-value Connectivity	FDR q-value Clustering Coefficient	Number of Proteins in the Module	Average Clustering Coefficient in the Reference Network	sec	min	hr	Biological Process Module Definition
2.4	1.46	0	0	0	0	25	0.53	OK	-	OK	protein import into mitochondrial inner membrane
2.1	1.46	0	0	0	0	5	1.00	-	-	OK	receptor-mediated endocytosis
2.0	1.50	0	0	0	0	6	1.00	OK	OK	OK	meiotic recombination checkpoint
1.9	1.46	0.002	0	0.004	0	5	0.60	OK	-	OK	stress granule assembly
1.9	1.50	0.007	0	0.011	0	17	0.72	-	-	OK	adaptation to pheromone during conjugation with cellular fusion
1.8	1.38	0	0	0	0	7	1.00	-	OK	OK	ceramide biosynthetic process
1.7	1.08	0	0	0	0	73	1.00	-	OK	OK	lipid metabolic process
1.7	1.26	0	0	0	0	4	0.21	-	-	-	protein amino acid N-linked glycosylation via asparagine
1.5	1.26	0	0	0	0	27	0.48	OK	-	OK	telomere maintenance via telomerase
0.9	0.08	0	0	0	0	24	0.31	-	-	OK	phosphate transport
-	7.53	-	0.024	-	0.04	967	0.27	OK	OK	OK	Transport
-	2.47	-	0.006	-	0.012	52	0.62	OK	OK	OK	G2/M transition of mitotic cell cycle
-	2.30	-	0.005	-	0.01	40	0.43	-	-	OK	negative regulation of transposition, RNA-mediated
-	2.22	-	0.016	-	0.028	40	0.58	-	-	OK	proton transport
-	2.16	-	0.017	-	0.029	61	0.56	-	-	OK	ATP synthesis coupled proton transport
-	2.14	-	0	-	0	38	0.64	-	-	OK	intra-Golgi vesicle-mediated transport
-	2.03	-	0	-	0	241	0.21	-	-	OK	autophagy
-	1.88	-	0.003	-	0.006	33	0.72	-	OK	OK	Golgi to plasma membrane transport
-	1.87	-	0	-	0	1108	0.02	OK	OK	OK	biological_process
-	1.84	-	0	-	0	48	0.40	-	OK	OK	cytokinesis
-	1.73	-	0	-	0	18	0.70	-	-	-	nuclear-transcribed mRNA catabolic process, 3'-5' exonucleolytic nonsense-mediated decay

Table A2. All modules identified based on TEAFS method with the order of descending EDA-LCON scores. Identification statuses over distinct time spans are also provided.

EDA – LCON	EDA - LCC	p-value Connectivity	p-value Clustering Coefficient	FDR q-value Connectivity	FDR q-value Clustering Coefficient	Number of Proteins in the Module	Average Clustering Coefficient in the Reference Network	sec	min	hr	Biological Process Module Definition
-	1.69	-	0.007	-	0.013	10	0.67	OK	-	-	regulation of transcription from RNA polymerase III promoter
-	1.63	-	0.02	-	0.034	6	0.87	OK	-	OK	regulation of transcription from RNA polymerase I promoter
-	1.56	-	0.004	-	0.008	31	0.60	-	-	OK	vesicle fusion
-	1.47	-	0.003	-	0.006	280	0.13	OK	OK	OK	protein amino acid phosphorylation
-	1.44	-	0.02	-	0.034	42	0.84	-	OK	OK	histone deacetylation
-	1.29	-	0.023	-	0.038	11	1.00	-	-	OK	tubulin complex assembly
-	1.23	-	0.016	-	0.028	54	0.37	OK	OK	-	G1/S transition of mitotic cell cycle
-	0.80	-	0	-	0	38	0.33	OK	-	OK	protein amino acid glycosylation
-	0.64	-	0.003	-	0.006	16	0.84	-	OK	OK	SCF-dependent proteasomal ubiquitin-dependent protein catabolic process

Table A3. The modules identified by distinct time spans but not by overall time span. Listed in descending order of their EDA-LCON values.

Biological Process Module Definition	hr	min	sec	number of proteins in the module	Identification status by NP analysis
regulation of transcription from RNA polymerase II promoter	OK	OK	OK	91	OK
ion transport	OK	-	-	111	OK
vacuolar acidification	OK	-	-	47	OK
biopolymer biosynthetic process	OK	-	OK	136	OK
regulation of transcription during G1 phase of mitotic cell cycle	OK	-	-	28	OK
establishment of cell polarity	OK	OK	-	49	OK
establishment or maintenance of chromatin architecture	OK	-	-	38	-
protein folding	OK	OK	-	139	OK
regulation of exit from mitosis	OK	-	-	28	OK
mitotic spindle elongation	OK	-	-	24	-
histone exchange	OK	-	-	10	-
mitotic metaphase/anaphase transition	OK	-	-	13	-
nuclear-transcribed mRNA poly(A) tail shortening	OK	-	-	13	-
cyclin catabolic process	OK	-	-	12	-
one-carbon compound metabolic process	OK	-	-	44	-
nuclear migration along microtubule	OK	-	-	19	OK
Rho protein signal transduction	OK	-	-	22	-
signal transduction	OK	-	-	79	OK
RNA catabolic process	OK	-	-	22	-
filamentous growth	OK	OK	-	31	-
translational elongation	OK	-	-	28	OK
snRNA pseudouridine synthesis	OK	-	-	10	-
ergosterol biosynthetic process	-	-	OK	28	OK
steroid biosynthetic process	-	-	OK	26	OK
protein ubiquitination during ubiquitin-dependent protein catabolic process	-	OK	OK	31	-
axial cellular bud site selection	-	OK	-	25	OK
Golgi to endosome transport	-	-	OK	14	-
Cellular protein metabolic process	OK	-	-	11	-
chromatin silencing at rDNA	OK	-	-	21	-
Nuclear-transcribed mRNA catabolic process, non-stop decay	OK	-	-	18	-
conjugation with cellular fusion	OK	-	-	31	OK
late endosome to vacuole transport	OK	-	-	32	OK
sphingolipid biosynthetic process	OK	-	-	20	-
lipid biosynthetic process	-	OK	-	54	OK
sulfur amino acid metabolic process	OK	OK	-	5	-
rRNA pseudouridine synthesis	OK	OK	-	3	-
protein insertion into ER membrane	-	OK	-	3	-

Table A4. Biological process terms and their distribution over the modules.

Biological Process Module Definition	Detection parameter for module P	Detection parameter for module M	Detection parameter for module S	Detection parameter for module D	Identification status by TEAFS
translation	0.011025	0.002601	0.125316	0.005476	OK
transport	0.013456	0.029241	0.074529	0.054756	OK
regulation of translation	0.012544	0.000361	0.036864	0.009409	OK
regulation of transcription, DNA-dependent	0.017161	0.003249	0.033124	0.025921	OK
transcription	0.023104	0.002916	0.026244	0.016900	OK
protein transport	0.002601	0.008649	0.019881	0.018225	OK
cell cycle	0.003844	0.001600	0.017161	0.003844	OK
DNA repair	0.001024	0.000841	0.014641	0.001521	OK
chromatin modification	0.000841	0.000121	0.010201	0.000729	OK
vesicle-mediated transport	0.000324	0.001156	0.008281	0.008281	OK
biopolymer biosynthetic process	0.000729	0.000196	0.006561	0.004624	-
proteolysis	0.000400	0.001089	0.006561	0.002704	OK
response to DNA damage stimulus	0.001681	0.000784	0.006561	0.001936	OK
intracellular protein transport	0.000144	0.000400	0.005041	0.000729	OK
protein amino acid phosphorylation	0.000625	0.001225	0.003721	0.009801	OK
cell division	0.001521	0.000841	0.003721	0.002304	OK
metabolic process	0.003600	0.006241	0.003721	0.001936	OK
response to drug	0.000676	0.000196	0.003721	0.000729	OK
chromatin silencing at telomere	0.000256	0.000064	0.003721	0.000289	OK
ubiquitin-dependent protein catabolic process	0.000529	0.003136	0.002601	0.010201	OK
ER to Golgi vesicle-mediated transport	0.000144	0.000196	0.002601	0.003600	OK
ion transport	0.000100	0.000441	0.002601	0.003136	-
DNA replication	0.000400	0.000121	0.002601	0.000961	OK
mitosis	0.000400	0.000576	0.002601	0.000729	OK
biological_process	0.013689	0.036100	0.001600	0.017424	OK
cell wall organization	0.000256	0.000576	0.001600	0.003136	OK
mRNA processing	0.002209	0.000064	0.001600	0.002116	OK
regulation of transcription from RNA polymerase II promoter	0.000625	0.000064	0.001600	0.000529	-
endocytosis	0.000064	0.000729	0.000900	0.001369	OK
protein ubiquitination	0.000100	0.000100	0.000900	0.000100	OK
oxidation reduction	0.001089	0.004761	0.000400	0.005476	-
pseudohyphal growth	0.000169	0.000100	0.000400	0.000841	OK
signal transduction	0.000225	0.000289	0.000400	0.000361	-
small GTPase mediated signal transduction	0.000144	0.000064	0.000400	0.000225	-
protein complex assembly	0.000100	0.000196	0.000900	<0.000076	OK
chromosome segregation	0.000121	0.000100	0.000900	<0.000076	OK
aerobic respiration	0.000049	0.000841	0.000400	<0.000076	-
response to stress	0.000256	0.002025	<0.000183	0.001369	OK
positive regulation of transcription from RNA polymerase II promoter	0.000225	0.000144	<0.000183	0.001369	OK
Meiosis	0.000225	0.000676	<0.000183	0.000961	OK
carbohydrate metabolic process	0.000100	0.001369	<0.000183	0.000841	OK
cellular response to heat	0.000121	0.007921	<0.000183	0.000625	-

Table A4. Biological process terms and their distribution over the modules.

Biological Process Module Definition	Detection parameter for module P	Detection parameter for module M	Detection parameter for module S	Detection parameter for module D	Identification status by TEAFS
autophagy	0.000625	0.002601	<0.000183	0.000625	OK
protein folding	0.000361	0.000529	<0.000183	0.000625	-
response to oxidative stress	0.000100	0.000361	<0.000183	0.000529	-
response to osmotic stress	0.000064	0.000081	<0.000183	0.000361	OK
ascospore formation	0.000064	0.000324	<0.000183	0.000289	OK
transmembrane transport	0.000100	0.000361	<0.000183	0.000225	-
actin filament organization	0.000049	0.000324	<0.000183	0.000225	OK
protein amino acid dephosphorylation	0.000049	0.000081	<0.000183	0.000144	OK
sporulation resulting in formation of a cellular spore	0.000100	0.000625	<0.000183	0.000100	-
amino acid transport	0.000064	0.000081	<0.000183	0.000100	-
lipid metabolic process	0.000064	0.000169	<0.000183	<0.000076	OK
cellular respiration	0.000049	0.000100	<0.000183	<0.000076	OK
transcription from RNA polymerase II promoter	0.000361	<0.000057	0.006561	0.000144	OK
ribosome biogenesis	0.065025	<0.000057	0.005041	0.000841	OK
mRNA export from nucleus	0.000400	<0.000057	0.005041	0.000196	OK
cellular metabolic compound salvage	0.001849	<0.000057	0.002601	0.000961	-
rRNA processing	0.007744	<0.000057	0.002601	0.000196	OK
G2/M transition of mitotic cell cycle	0.000064	<0.000057	0.002601	0.000196	OK
nuclear mRNA splicing, via spliceosome	0.000729	<0.000057	0.001600	0.000625	OK
transcription initiation from RNA polymerase II promoter	0.000225	<0.000057	0.001600	0.000625	OK
histone acetylation	0.000100	<0.000057	0.001600	0.000441	OK
RNA splicing	0.001225	<0.000057	0.001600	0.000361	OK
RNA elongation from RNA polymerase II promoter	0.000361	<0.000057	0.001600	0.000289	OK
intracellular protein transmembrane transport	0.000400	<0.000057	0.000900	0.000289	OK
G1/S transition of mitotic cell cycle	0.000121	<0.000057	0.000900	0.000196	OK
chromatin remodeling	0.000121	<0.000057	0.000900	0.000100	OK
chromatin organization	0.000100	<0.000057	0.000900	0.000100	-
mRNA catabolic process	0.000049	<0.000057	0.000900	0.000100	OK
peptidyl-amino acid modification	0.001024	<0.000057	0.000400	0.001089	-
cell morphogenesis	0.000064	<0.000057	0.000400	0.000529	-
rRNA metabolic process	0.013456	<0.000057	0.000400	0.000361	OK
protein amino acid glycosylation	0.000049	<0.000057	0.000400	0.000225	OK
tRNA export from nucleus	0.000100	<0.000057	0.005041	<0.000076	OK
amino acid biosynthetic process	0.001024	<0.000057	0.001600	<0.000076	OK
mRNA transport	0.000400	<0.000057	0.001600	<0.000076	OK
rRNA export from nucleus	0.000361	<0.000057	0.001600	<0.000076	OK
translational initiation	0.000225	<0.000057	0.000900	<0.000076	OK
nuclear pore organization	0.000121	<0.000057	0.000900	<0.000076	OK
NLS-bearing substrate import into nucleus	0.000100	<0.000057	0.000900	<0.000076	OK
DNA replication initiation	0.000081	<0.000057	0.000900	<0.000076	OK
tRNA processing	0.000625	<0.000057	0.000400	<0.000076	OK

Table A4. Biological process terms and their distribution over the modules.

Biological Process Module Definition	Detection parameter for module P	Detection parameter for module M	Detection parameter for module S	Detection parameter for module D	Identification status by TEAFS
transcription from RNA polymerase I promoter	0.000576	<0.000057	0.000400	<0.000076	OK
protein export from nucleus	0.000144	<0.000057	0.000400	<0.000076	OK
mRNA-binding (hnRNP) protein import into nucleus	0.000121	<0.000057	0.000400	<0.000076	OK
tRNA aminoacylation for protein translation	0.000100	<0.000057	0.000400	<0.000076	-
methionine biosynthetic process	0.000100	<0.000057	0.000400	<0.000076	OK
snRNA export from nucleus	0.000100	<0.000057	0.000400	<0.000076	OK
snRNP protein import into nucleus	0.000100	<0.000057	0.000400	<0.000076	OK
ribosomal protein import into nucleus	0.000100	<0.000057	0.000400	<0.000076	OK
nucleotide-excision repair	0.000064	<0.000057	0.000400	<0.000076	OK
ATP-dependent chromatin remodeling	0.000049	<0.000057	0.000400	<0.000076	OK
negative regulation of transcription from RNA polymerase II promoter	0.000121	<0.000057	<0.000183	0.001369	OK
cytokinesis	0.000049	<0.000057	<0.000183	0.000529	OK
invasive growth in response to glucose limitation	0.000144	<0.000057	<0.000183	0.000289	OK
establishment of cell polarity	0.000121	<0.000057	<0.000183	0.000289	-
chromatin silencing at silent mating-type cassette	0.000049	<0.000057	<0.000183	0.000289	OK
protein amino acid N-linked glycosylation	0.000049	<0.000057	<0.000183	0.000289	OK
negative regulation of transposition, RNA-mediated	0.000049	<0.000057	<0.000183	0.000225	OK
reproduction	0.000064	<0.000057	<0.000183	0.000196	-
DNA recombination	0.000064	<0.000057	<0.000183	0.000196	OK
mRNA cleavage	0.000064	<0.000057	<0.000183	0.000196	OK
response to pheromone	0.000049	<0.000057	<0.000183	0.000196	OK
regulation of transcription	0.000441	<0.000057	<0.000183	0.000144	OK
response to toxin	0.000081	<0.000057	<0.000183	0.000144	-
tRNA metabolic process	0.000729	<0.000057	<0.000183	0.000100	-
establishment of protein localization	0.000289	<0.000057	<0.000183	0.000100	OK
purine nucleotide biosynthetic process	0.000064	<0.000057	<0.000183	0.000100	-
cytoskeleton organization	0.000049	<0.000057	<0.000183	0.000100	OK
RNA modification	0.003025	<0.000057	<0.000183	<0.000076	OK
maturation of SSU-rRNA from tricistronic rRNA transcript (SSU-rRNA, 5.8S rRNA, LSU-rRNA)	0.000676	<0.000057	<0.000183	<0.000076	OK
endonucleolytic cleavage in ITS1 to separate SSU-rRNA from 5.8S rRNA and LSU-rRNA from tricistronic rRNA transcript (SSU-rRNA, 5.8S rRNA, LSU-rRNA)	0.000400	<0.000057	<0.000183	<0.000076	OK
ribosomal large subunit assembly	0.000400	<0.000057	<0.000183	<0.000076	OK
transcription from RNA polymerase III promoter	0.000256	<0.000057	<0.000183	<0.000076	OK

Table A4. Biological process terms and their distribution over the modules.

Biological Process Module Definition	Detection parameter for module P	Detection parameter for module M	Detection parameter for module S	Detection parameter for module D	Identification status by TEAFS
endonucleolytic cleavage to generate mature 5'-end of SSU-rRNA from (SSU-rRNA, 5.8S rRNA, LSU-rRNA)	0.000225	<0.000057	<0.000183	<0.000076	OK
ribosomal large subunit biogenesis	0.000225	<0.000057	<0.000183	<0.000076	OK
RNA processing	0.000225	<0.000057	<0.000183	<0.000076	OK
endonucleolytic cleavage in 5'-ETS of tricistronic rRNA transcript (SSU-rRNA, 5.8S rRNA, LSU-rRNA)	0.000196	<0.000057	<0.000183	<0.000076	OK
protein import into nucleus	0.000144	<0.000057	<0.000183	<0.000076	OK
cellular aromatic compound metabolic process	0.000144	<0.000057	<0.000183	<0.000076	-
nucleocytoplasmic transport	0.000144	<0.000057	<0.000183	<0.000076	OK
tRNA modification	0.000121	<0.000057	<0.000183	<0.000076	OK
one-carbon metabolic process	0.000121	<0.000057	<0.000183	<0.000076	-
biosynthetic process	0.000100	<0.000057	<0.000183	<0.000076	OK
snoRNA metabolic process	0.000100	<0.000057	<0.000183	<0.000076	OK
telomere maintenance	0.000100	<0.000057	<0.000183	<0.000076	OK
nucleus organization	0.000100	<0.000057	<0.000183	<0.000076	OK
regulation of cell size	0.000100	<0.000057	<0.000183	<0.000076	-
ribosomal small subunit biogenesis	0.000100	<0.000057	<0.000183	<0.000076	OK
exonucleolytic trimming to generate mature 3'-end of 5.8S rRNA from tricistronic rRNA transcript (SSU-rRNA, 5.8S rRNA, LSU-rRNA)	0.000081	<0.000057	<0.000183	<0.000076	OK
tRNA methylation	0.000081	<0.000057	<0.000183	<0.000076	-
RNA fragment catabolic process	0.000064	<0.000057	<0.000183	<0.000076	OK
maturation of LSU-rRNA from tricistronic rRNA transcript (SSU-rRNA, 5.8S rRNA, LSU-rRNA)	0.000064	<0.000057	<0.000183	<0.000076	OK
ribosomal large subunit export from nucleus	0.000064	<0.000057	<0.000183	<0.000076	OK
RNA elongation	0.000064	<0.000057	<0.000183	<0.000076	OK
nitrogen compound metabolic process	0.000064	<0.000057	<0.000183	<0.000076	-
nuclear transport	0.000064	<0.000057	<0.000183	<0.000076	OK
cellular bud site selection	0.000049	<0.000057	<0.000183	<0.000076	-
amino acid metabolic process	0.000049	<0.000057	<0.000183	<0.000076	OK
glutamine metabolic process	0.000049	<0.000057	<0.000183	<0.000076	-
nucleoside metabolic process	0.000049	<0.000057	<0.000183	<0.000076	-
iron-sulfur cluster assembly	0.000049	<0.000057	<0.000183	<0.000076	-
maturation of 5.8S rRNA from tricistronic rRNA transcript (SSU-rRNA, 5.8S rRNA, LSU-rRNA)	0.000049	<0.000057	<0.000183	<0.000076	OK
pre-replicative complex assembly	0.000049	<0.000057	<0.000183	<0.000076	OK
regulation of exit from mitosis	0.000049	<0.000057	<0.000183	<0.000076	-
nuclear-transcribed mRNA catabolic process, exonucleolytic, 3'-5'	0.000049	<0.000057	<0.000183	<0.000076	OK
glycolysis	<0.000044	0.000121	0.012321	0.001521	-

Table A4. Biological process terms and their distribution over the modules.

Biological Process Module Definition	Detection parameter for module P	Detection parameter for module M	Detection parameter for module S	Detection parameter for module D	Identification status by TEAFS
replicative cell aging	<0.000044	0.000144	0.000900	0.000196	-
mitochondrial translation	<0.000044	0.000529	0.000400	0.000961	OK
vacuolar protein catabolic process	<0.000044	0.001764	0.000400	0.000841	-
proton transport	<0.000044	0.000196	0.002601	<0.000076	OK
ATP synthesis coupled proton transport	<0.000044	0.000144	0.000900	<0.000076	OK
glucose metabolic process	<0.000044	0.000064	0.000900	<0.000076	-
protein targeting to vacuole	<0.000044	0.000169	0.000400	<0.000076	OK
mitochondrial genome maintenance	<0.000044	0.000081	0.000400	<0.000076	-
cell redox homeostasis	<0.000044	0.000081	<0.000183	0.001369	-
lipid biosynthetic process	<0.000044	0.000121	<0.000183	0.000729	-
ER-associated protein catabolic process	<0.000044	0.000121	<0.000183	0.000625	-
vacuolar transport	<0.000044	0.000064	<0.000183	0.000625	-
energy reserve metabolic process	<0.000044	0.001024	<0.000183	0.000441	-
peroxisome organization	<0.000044	0.000100	<0.000183	0.000289	-
exocytosis	<0.000044	0.000196	<0.000183	0.000225	-
bipolar cellular bud site selection	<0.000044	0.000121	<0.000183	0.000225	OK
electron transport chain	<0.000044	0.000121	<0.000183	0.000225	-
negative regulation of protein metabolic process	<0.000044	0.000064	<0.000183	0.000225	-
glycogen metabolic process	<0.000044	0.000064	<0.000183	0.000144	-
alcohol metabolic process	<0.000044	0.000064	<0.000183	0.000144	-
chromatin assembly or disassembly	<0.000044	0.000064	<0.000183	0.000100	-
actin cytoskeleton organization	<0.000044	0.000064	<0.000183	0.000100	OK
mitochondrion inheritance	<0.000044	0.000225	<0.000183	<0.000076	-
tricarboxylic acid cycle	<0.000044	0.000196	<0.000183	<0.000076	-
ATP biosynthetic process	<0.000044	0.000169	<0.000183	<0.000076	-
phosphorus metabolic process	<0.000044	0.000169	<0.000183	<0.000076	-
carbohydrate transport	<0.000044	0.000169	<0.000183	<0.000076	-
piecemeal microautophagy of nucleus	<0.000044	0.000169	<0.000183	<0.000076	-
glycogen biosynthetic process	<0.000044	0.000121	<0.000183	<0.000076	-
ascospore wall assembly	<0.000044	0.000121	<0.000183	<0.000076	-
CVT pathway	<0.000044	0.000100	<0.000183	<0.000076	-
vesicle fusion	<0.000044	0.000081	<0.000183	<0.000076	OK
reciprocal meiotic recombination	<0.000044	0.000081	<0.000183	<0.000076	OK
protein catabolic process	<0.000044	0.000081	<0.000183	<0.000076	OK
retrograde transport, endosome to Golgi	<0.000044	0.000081	<0.000183	<0.000076	OK
lipid transport	<0.000044	0.000064	<0.000183	<0.000076	-
mitochondrion organization	<0.000044	0.000064	<0.000183	<0.000076	-
Ras protein signal transduction	<0.000044	0.000064	<0.000183	<0.000076	-
cell death	<0.000044	0.000064	<0.000183	<0.000076	-
nucleotide metabolic process	<0.000044	0.000064	<0.000183	<0.000076	-
protein refolding	<0.000044	0.000064	<0.000183	<0.000076	-
proteasomal ubiquitin-dependent protein catabolic process	<0.000044	0.000064	<0.000183	<0.000076	-

Table A4. Biological process terms and their distribution over the modules.

Biological Process Module Definition	Detection parameter for module P	Detection parameter for module M	Detection parameter for module S	Detection parameter for module D	Identification status by TEAFS
positive regulation of spindle pole body separation	<0.000044	<0.000057	0.003721	0.000100	-
vacuolar acidification	<0.000044	<0.000057	0.001600	0.000529	-
mitotic sister chromatid segregation	<0.000044	<0.000057	0.000900	0.000625	OK
retrograde vesicle-mediated transport, Golgi to ER	<0.000044	<0.000057	0.000400	0.000441	OK
vacuole fusion, non-autophagic	<0.000044	<0.000057	0.000400	0.000289	OK
regulation of cyclin-dependent protein kinase activity	<0.000044	<0.000057	0.000400	0.000196	OK
histone deacetylation	<0.000044	<0.000057	0.000400	0.000196	OK
mismatch repair	<0.000044	<0.000057	0.000400	0.000100	OK
S phase of mitotic cell cycle	<0.000044	<0.000057	0.003721	<0.000076	OK
nuclear mRNA branch site recognition	<0.000044	<0.000057	0.001600	<0.000076	-
translational elongation	<0.000044	<0.000057	0.001600	<0.000076	-
calcium ion transport	<0.000044	<0.000057	0.001600	<0.000076	-
nuclear-transcribed mRNA catabolic process, nonsense-mediated decay	<0.000044	<0.000057	0.000900	<0.000076	-
telomere maintenance via recombination	<0.000044	<0.000057	0.000900	<0.000076	OK
glycyl-tRNA aminoacylation	<0.000044	<0.000057	0.000900	<0.000076	-
tRNA 3'-trailer cleavage, exonucleolytic	<0.000044	<0.000057	0.000900	<0.000076	-
regulation of gene expression, epigenetic	<0.000044	<0.000057	0.000900	<0.000076	-
double-strand break repair via homologous recombination	<0.000044	<0.000057	0.000400	<0.000076	-
recombinational repair	<0.000044	<0.000057	0.000400	<0.000076	-
karyogamy during conjugation with cellular fusion	<0.000044	<0.000057	0.000400	<0.000076	-
leading strand elongation	<0.000044	<0.000057	0.000400	<0.000076	OK
intron homing	<0.000044	<0.000057	0.000400	<0.000076	-
DNA packaging	<0.000044	<0.000057	0.000400	<0.000076	OK
transcription initiation	<0.000044	<0.000057	0.000400	<0.000076	OK
glutamyl-tRNA aminoacylation	<0.000044	<0.000057	0.000400	<0.000076	-
methionine metabolic process	<0.000044	<0.000057	0.000400	<0.000076	-
phosphatidylinositol biosynthetic process	<0.000044	<0.000057	0.000400	<0.000076	-
nicotinamide metabolic process	<0.000044	<0.000057	0.000400	<0.000076	-
intra-Golgi vesicle-mediated transport	<0.000044	<0.000057	0.000400	<0.000076	OK
receptor-mediated endocytosis	<0.000044	<0.000057	0.000400	<0.000076	OK
apoptosis	<0.000044	<0.000057	0.000400	<0.000076	-
vacuole organization	<0.000044	<0.000057	0.000400	<0.000076	-
sister chromatid cohesion	<0.000044	<0.000057	0.000400	<0.000076	OK
mitotic cell cycle spindle assembly checkpoint	<0.000044	<0.000057	0.000400	<0.000076	OK
budding cell bud growth	<0.000044	<0.000057	0.000400	<0.000076	OK
intracellular mRNA localization	<0.000044	<0.000057	0.000400	<0.000076	-
threonine biosynthetic process	<0.000044	<0.000057	0.000400	<0.000076	-
RNA metabolic process	<0.000044	<0.000057	0.000400	<0.000076	-
gene silencing	<0.000044	<0.000057	0.000400	<0.000076	-

Table A4. Biological process terms and their distribution over the modules.

Biological Process Module Definition	Detection parameter for module P	Detection parameter for module M	Detection parameter for module S	Detection parameter for module D	Identification status by TEAFS
protein deubiquitination	<0.000044	<0.000057	0.000400	<0.000076	-
negative regulation of translation	<0.000044	<0.000057	0.000400	<0.000076	-
cellular cation homeostasis	<0.000044	<0.000057	0.000400	<0.000076	-
protein import into mitochondrial matrix	<0.000044	<0.000057	0.000400	<0.000076	OK
protein splicing	<0.000044	<0.000057	0.000400	<0.000076	-
regulation of microtubule polymerization or depolymerization	<0.000044	<0.000057	0.000400	<0.000076	OK
intra-S DNA damage checkpoint	<0.000044	<0.000057	0.000400	<0.000076	-
positive regulation of protein amino acid autophosphorylation	<0.000044	<0.000057	0.000400	<0.000076	-
late endosome to vacuole transport via multivesicular body sorting pathway	<0.000044	<0.000057	0.000400	<0.000076	-
protein amino acid autophosphorylation	<0.000044	<0.000057	0.000400	<0.000076	-
negative regulation of nuclear mRNA splicing, via spliceosome	<0.000044	<0.000057	0.000400	<0.000076	-
'de novo' cotranslational protein folding	<0.000044	<0.000057	0.000400	<0.000076	-
chromosome organization	<0.000044	<0.000057	0.000400	<0.000076	-
iron ion transport	<0.000044	<0.000057	<0.000183	0.001225	-
cellular iron ion homeostasis	<0.000044	<0.000057	<0.000183	0.000841	-
signal peptide processing	<0.000044	<0.000057	<0.000183	0.000729	-
GPI anchor biosynthetic process	<0.000044	<0.000057	<0.000183	0.000625	OK
zinc ion transport	<0.000044	<0.000057	<0.000183	0.000441	-
ergosterol biosynthetic process	<0.000044	<0.000057	<0.000183	0.000361	-
mRNA polyadenylation	<0.000044	<0.000057	<0.000183	0.000289	OK
fatty acid biosynthetic process	<0.000044	<0.000057	<0.000183	0.000289	-
steroid biosynthetic process	<0.000044	<0.000057	<0.000183	0.000289	-
microautophagy	<0.000044	<0.000057	<0.000183	0.000289	-
late endosome to vacuole transport	<0.000044	<0.000057	<0.000183	0.000289	-
vacuole inheritance	<0.000044	<0.000057	<0.000183	0.000225	-
copper ion transport	<0.000044	<0.000057	<0.000183	0.000225	-
Golgi to vacuole transport	<0.000044	<0.000057	<0.000183	0.000225	OK
sterol biosynthetic process	<0.000044	<0.000057	<0.000183	0.000225	-
positive regulation of histone acetylation	<0.000044	<0.000057	<0.000183	0.000225	-
regulation of cell cycle	<0.000044	<0.000057	<0.000183	0.000225	OK
gluconeogenesis	<0.000044	<0.000057	<0.000183	0.000196	-
regulation of carbohydrate metabolic process	<0.000044	<0.000057	<0.000183	0.000196	-
double-strand break repair via nonhomologous end joining	<0.000044	<0.000057	<0.000183	0.000196	OK
protein amino acid O-linked glycosylation	<0.000044	<0.000057	<0.000183	0.000196	-
high-affinity iron ion transport	<0.000044	<0.000057	<0.000183	0.000196	-
response to unfolded protein	<0.000044	<0.000057	<0.000183	0.000196	-
axial cellular bud site selection	<0.000044	<0.000057	<0.000183	0.000196	-
response to heat	<0.000044	<0.000057	<0.000183	0.000196	-
nuclear migration along microtubule	<0.000044	<0.000057	<0.000183	0.000196	-

Table A4. Biological process terms and their distribution over the modules.

Biological Process Module Definition	Detection parameter for module P	Detection parameter for module M	Detection parameter for module S	Detection parameter for module D	Identification status by TEAFS
posttranslational protein targeting to membrane, translocation	<0.000044	<0.000057	<0.000183	0.000196	OK
nuclear migration during conjugation with cellular fusion	<0.000044	<0.000057	<0.000183	0.000144	-
pheromone-dependent signal transduction involved in conjugation with cellular fusion	<0.000044	<0.000057	<0.000183	0.000144	OK
nucleobase, nucleoside, nucleotide and nucleic acid metabolic process	<0.000044	<0.000057	<0.000183	0.000144	-
DNA-dependent DNA replication	<0.000044	<0.000057	<0.000183	0.000144	OK
nucleosome assembly	<0.000044	<0.000057	<0.000183	0.000144	OK
dolichol-linked oligosaccharide biosynthetic process	<0.000044	<0.000057	<0.000183	0.000144	-
SRP-dependent cotranslational protein targeting to membrane, translocation	<0.000044	<0.000057	<0.000183	0.000144	-
deoxyribonucleotide biosynthetic process	<0.000044	<0.000057	<0.000183	0.000144	-
protein import into peroxisome matrix	<0.000044	<0.000057	<0.000183	0.000144	-
homologous chromosome segregation	<0.000044	<0.000057	<0.000183	0.000144	-
positive regulation of transcription	<0.000044	<0.000057	<0.000183	0.000144	-
actin cortical patch localization	<0.000044	<0.000057	<0.000183	0.000144	-
regulation of transcription during G1 phase of mitotic cell cycle	<0.000044	<0.000057	<0.000183	0.000100	-
mitotic cell cycle	<0.000044	<0.000057	<0.000183	0.000100	OK
conjugation with cellular fusion	<0.000044	<0.000057	<0.000183	0.000100	-
mitotic recombination	<0.000044	<0.000057	<0.000183	0.000100	-
termination of RNA polymerase II transcription	<0.000044	<0.000057	<0.000183	0.000100	-
cation transport	<0.000044	<0.000057	<0.000183	0.000100	-
microtubule-based process	<0.000044	<0.000057	<0.000183	0.000100	-
mitotic chromosome condensation	<0.000044	<0.000057	<0.000183	0.000100	-
spindle pole body duplication in nuclear envelope	<0.000044	<0.000057	<0.000183	0.000100	-
donor selection	<0.000044	<0.000057	<0.000183	0.000100	-
isoprenoid biosynthetic process	<0.000044	<0.000057	<0.000183	0.000100	-
phospholipid biosynthetic process	<0.000044	<0.000057	<0.000183	0.000100	-
protein import into peroxisome matrix, docking	<0.000044	<0.000057	<0.000183	0.000100	-
response to starvation	<0.000044	<0.000057	<0.000183	0.000100	-
proteasome assembly	<0.000044	<0.000057	<0.000183	0.000100	OK
response to arsenic	<0.000044	<0.000057	<0.000183	0.000100	-
phosphoinositide phosphorylation	<0.000044	<0.000057	<0.000183	0.000100	-
organelle inheritance	<0.000044	<0.000057	<0.000183	0.000100	-

APPENDIX B: MATLAB CODES FOR CALCULATIONS

Table B1. Calculation of average clustering coefficients

```
% ============================================================
%            Average Clustering Coefficients for each Module (ACCM) calculations
%          by MUHAMMED ERKAN KARABEKMEZ, BOGAZICI UNIVERSITY
% ============================================================
% load protein-protein interaction matrices
load S8
S=S8;
% load GO term - protein interaction network
load G
%m stands for module order
for m=1:1797
   n=0;
   % n is the number of neighbors of protein x in module m
   for x=1:length(S)
      if G(m,x)>0
         for y=1:length(S)
            if G(m,y)>0
               n=n+1;
            end
         end
         N(x,m)=(n*(n-1))/2;
         % N(x,m) is the number of maximum possible connections
         % among the neighbors of protein x in module m
         n=0;
      else
         n=0;
      end
   end
end
for m=1:1797
   for x=1:length(S)
      i=1;
      K=0;
      if G(m,x)>0
         for y=1:length(S)
            if S(x,y)>0
               if G(m,y)>0
                  K(i)=y;
                  i=i+1;
               end
            end
         end
      end
      z=0;
      j=1;
```

```
        i=1;
        for i=1:length(K)
            for j=1:length(K)
                if K(i)>0
                    if K(j)>0
                        if K(i)-K(j)<1
                            z=z;
                        else
                            z=z+S(K(i),K(j));
                        end
                    end
                end
            end
        end
        C(x,m)=z;
    end
end
for m=1:1797
    for x=1:length(S)
        if N(x,m)<1
            CC(x,m)=0;
        else
            CC(x,m)=C(x,m)/N(x,m);
        end
    end
end
% CC(x,m) is the local clustering coefficient of protein x in module m
for m=1:1797
    TOT=0;
    z=0;
    for x=1:length(S)
        if G(m,x)>0
            TOT=TOT+CC(x,m);
            z=z+1;
        else
            TOT=TOT;
         end
    end
    if z>0
        ACC(m)=TOT/z;
        % ACC(m) is the overall average clustering coefficient in module m
    else
        ACC(m)=0;
    end
end
```

Table B2. Calculation of EDA-LCC scores

```
% ============================================================
%       Extend of Differential Activity Score calculations for Clustering Coefficients
%                                   (EDA-LCC)
%       by MUHAMMED ERKAN KARABEKMEZ, BOGAZICI UNIVERSITY
% ============================================================
load S1
load G1
t=1;
for m=1:294
   TOT(m,t)=0;
   for x=1:length(S1)
      i=1;
      K=0;
      n=0;
      if G1(m,x)<1, continue; end
      for y=1:length(S1)
         if S1(x,y)<1, continue; end
         if G1(m,y)<1, continue; end
         n=n+S1(x,y);
         K(i)=y;
         i=i+1;
      end
      N(x,m)=(n*(n-1))/2;
      z=0;
      for i=1:length(K)
         for j=1:length(K)
            if ((K(i)-K(j))^2)<1, continue; end
            z=z+S1(K(i),K(j));
         end
      end
      C(x,m)=z/2;
      if N(x,m)>0
         CC(x,m)=C(x,m)/N(x,m);
      else
         CC(x,m)=0;
      end
      TOT(m,t)=TOT(m,t)+CC(x,m);
   end
end
clear S1
load S2
t=2;
disp('1')
for m=1:294
   TOT(m,t)=0;
   for x=1:length(S2)
      i=1;
      K=0;
```

```
        n=0;
        if G1(m,x)<1, continue; end
        for y=1:length(S2)
            if S2(x,y)<1, continue; end
            if G1(m,y)<1, continue; end
            n=n+S2(x,y);
            K(i)=y;
            i=i+1;
        end
        N(x,m)=(n*(n-1))/2;
        z=0;
        for i=1:length(K)
            for j=1:length(K)
                if ((K(i)-K(j))^2)<1, continue; end
                z=z+S2(K(i),K(j));
            end
        end
        C(x,m)=z/2;
        if N(x,m)>0
            CC(x,m)=C(x,m)/N(x,m);
        else
            CC(x,m)=0;
        end
        TOT(m,t)=TOT(m,t)+CC(x,m);
    end
end
clear S2
load S3
t=3;
disp('2')
for m=1:294
    TOT(m,t)=0;
    for x=1:length(S3)
        i=1;
        K=0;
        n=0;
        if G1(m,x)<1, continue; end
        for y=1:length(S3)
            if S3(x,y)<1, continue; end
            if G1(m,y)<1, continue; end
            n=n+S3(x,y);
            K(i)=y;
            i=i+1;
        end
        N(x,m)=(n*(n-1))/2;
        z=0;
        for i=1:length(K)
            for j=1:length(K)
                if ((K(i)-K(j))^2)<1, continue; end
```

```
            z=z+S3(K(i),K(j));
          end
       end
       C(x,m)=z/2;
       if N(x,m)>0
          CC(x,m)=C(x,m)/N(x,m);
       else
          CC(x,m)=0;
       end
       TOT(m,t)=TOT(m,t)+CC(x,m);
    end
end
clear S3
load S4
t=4;
disp('3')
for m=1:294
    TOT(m,t)=0;
    for x=1:length(S4)
       i=1;
       K=0;
       n=0;
       if G1(m,x)<1, continue; end
       for y=1:length(S4)
          if S4(x,y)<1, continue; end
          if G1(m,y)<1, continue; end
          n=n+S4(x,y);
          K(i)=y;
          i=i+1;
       end
       N(x,m)=(n*(n-1))/2;
       z=0;
       for i=1:length(K)
          for j=1:length(K)
             if ((K(i)-K(j))^2)<1, continue; end
             z=z+S4(K(i),K(j));
          end
       end
       C(x,m)=z/2;
       if N(x,m)>0
          CC(x,m)=C(x,m)/N(x,m);
       else
          CC(x,m)=0;
       end
       TOT(m,t)=TOT(m,t)+CC(x,m);
    end
end
clear S4
load S5
```

```
t=5;
disp('4')
for m=1:294
   TOT(m,t)=0;
   for x=1:length(S5)
      i=1;
      K=0;
      n=0;
      if G1(m,x)<1, continue; end
      for y=1:length(S5)
         if S5(x,y)<1, continue; end
         if G1(m,y)<1, continue; end
         n=n+S5(x,y);
         K(i)=y;
         i=i+1;
      end
      N(x,m)=(n*(n-1))/2;
      z=0;
      for i=1:length(K)
         for j=1:length(K)
            if ((K(i)-K(j))^2)<1, continue; end
            z=z+S5(K(i),K(j));
         end
      end
      C(x,m)=z/2;
      if N(x,m)>0
         CC(x,m)=C(x,m)/N(x,m);
      else
         CC(x,m)=0;
      end
      TOT(m,t)=TOT(m,t)+CC(x,m);
   end
end
clear S5
load S6
t=6;
disp('5')
for m=1:294
   TOT(m,t)=0;
   for x=1:length(S6)
      i=1;
      K=0;
      n=0;
      if G1(m,x)<1, continue; end
      for y=1:length(S6)
         if S6(x,y)<1, continue; end
         if G1(m,y)<1, continue; end
         n=n+S6(x,y);
         K(i)=y;
```

```
        i=i+1;
      end
      N(x,m)=(n*(n-1))/2;
      z=0;
      for i=1:length(K)
        for j=1:length(K)
          if ((K(i)-K(j))^2)<1, continue; end
          z=z+S6(K(i),K(j));
        end
      end
      C(x,m)=z/2;
      if N(x,m)>0
        CC(x,m)=C(x,m)/N(x,m);
      else
        CC(x,m)=0;
      end
      TOT(m,t)=TOT(m,t)+CC(x,m);
    end
end
clear S6
load S7
t=7;
disp('6')
for m=1:294
    TOT(m,t)=0;
    for x=1:length(S7)
      i=1;
      K=0;
      n=0;
      if G1(m,x)<1, continue; end
      for y=1:length(S7)
        if S7(x,y)<1, continue; end
        if G1(m,y)<1, continue; end
        n=n+S7(x,y);
        K(i)=y;
        i=i+1;
      end
      N(x,m)=(n*(n-1))/2;
      z=0;
      for i=1:length(K)
        for j=1:length(K)
          if ((K(i)-K(j))^2)<1, continue; end
          z=z+S7(K(i),K(j));
        end
      end
      C(x,m)=z/2;
      if N(x,m)>0
        CC(x,m)=C(x,m)/N(x,m);
      else
```

```
            CC(x,m)=0;
        end
        TOT(m,t)=TOT(m,t)+CC(x,m);
    end
end
clear S7
load S8
t=8;
disp('7')
for m=1:294
    TOT(m,t)=0;
    for x=1:length(S8)
        i=1;
        K=0;
        n=0;
        if G1(m,x)<1, continue; end
        for y=1:length(S8)
            if S8(x,y)<1, continue; end
            if G1(m,y)<1, continue; end
            n=n+S8(x,y);
            K(i)=y;
            i=i+1;
        end
        N(x,m)=(n*(n-1))/2;
        z=0;
        for i=1:length(K)
            for j=1:length(K)
                if ((K(i)-K(j))^2)<1, continue; end
                z=z+S8(K(i),K(j));
            end
        end
        C(x,m)=z/2;
        if N(x,m)>0
            CC(x,m)=C(x,m)/N(x,m);
        else
            CC(x,m)=0;
        end
        TOT(m,t)=TOT(m,t)+CC(x,m);
    end
end
clear S8
load S9
t=9;
disp('8')
for m=1:294
    TOT(m,t)=0;
    for x=1:length(S9)
        i=1;
        K=0;
```

```
        n=0;
        if G1(m,x)<1, continue; end
        for y=1:length(S9)
            if S9(x,y)<1, continue; end
            if G1(m,y)<1, continue; end
            n=n+S9(x,y);
            K(i)=y;
            i=i+1;
        end
        N(x,m)=(n*(n-1))/2;
        z=0;
        for i=1:length(K)
            for j=1:length(K)
                if ((K(i)-K(j))^2)<1, continue; end
                z=z+S9(K(i),K(j));
            end
        end
        C(x,m)=z/2;
        if N(x,m)>0
            CC(x,m)=C(x,m)/N(x,m);
        else
            CC(x,m)=0;
        end
        TOT(m,t)=TOT(m,t)+CC(x,m);
    end
end
clear S9
load S10
t=10;
disp('9')
for m=1:294
    TOT(m,t)=0;
    for x=1:length(S10)
        i=1;
        K=0;
        n=0;
        if G1(m,x)<1, continue; end
        for y=1:length(S10)
            if S10(x,y)<1, continue; end
            if G1(m,y)<1, continue; end
            n=n+S10(x,y);
            K(i)=y;
            i=i+1;
        end
        N(x,m)=(n*(n-1))/2;
        z=0;
        for i=1:length(K)
            for j=1:length(K)
                if ((K(i)-K(j))^2)<1, continue; end
```

```
            z=z+S10(K(i),K(j));
         end
      end
      C(x,m)=z/2;
      if N(x,m)>0
         CC(x,m)=C(x,m)/N(x,m);
      else
         CC(x,m)=0;
      end
      TOT(m,t)=TOT(m,t)+CC(x,m);
   end
end
clear S10
load S11
t=11;
disp('10')
for m=1:294
   TOT(m,t)=0;
   for x=1:length(S11)
      i=1;
      K=0;
      n=0;
      if G1(m,x)<1, continue; end
      for y=1:length(S11)
         if S11(x,y)<1, continue; end
         if G1(m,y)<1, continue; end
         n=n+S11(x,y);
         K(i)=y;
         i=i+1;
      end
      N(x,m)=(n*(n-1))/2;
      z=0;
      for i=1:length(K)
         for j=1:length(K)
            if ((K(i)-K(j))^2)<1, continue; end
            z=z+S11(K(i),K(j));
         end
      end
      C(x,m)=z/2;
      if N(x,m)>0
         CC(x,m)=C(x,m)/N(x,m);
      else
         CC(x,m)=0;
      end
      TOT(m,t)=TOT(m,t)+CC(x,m);
   end
end
clear S11
load S12
```

```
t=12;
disp('11')
for m=1:294
   TOT(m,t)=0;
   for x=1:length(S12)
      i=1;
      K=0;
      n=0;
      if G1(m,x)<1, continue; end
      for y=1:length(S12)
         if S12(x,y)<1, continue; end
         if G1(m,y)<1, continue; end
         n=n+S12(x,y);
         K(i)=y;
         i=i+1;
      end
      N(x,m)=(n*(n-1))/2;
      z=0;
      for i=1:length(K)
         for j=1:length(K)
            if ((K(i)-K(j))^2)<1, continue; end
            z=z+S12(K(i),K(j));
         end
      end
      C(x,m)=z/2;
      if N(x,m)>0
         CC(x,m)=C(x,m)/N(x,m);
      else
         CC(x,m)=0;
      end
      TOT(m,t)=TOT(m,t)+CC(x,m);
   end
end
clear S12
load S13
t=13;
disp('12')
for m=1:294
   TOT(m,t)=0;
   for x=1:length(S13)
      i=1;
      K=0;
      n=0;
      if G1(m,x)<1, continue; end
      for y=1:length(S13)
         if S13(x,y)<1, continue; end
         if G1(m,y)<1, continue; end
         n=n+S13(x,y);
         K(i)=y;
```

```
        i=i+1;
      end
      N(x,m)=(n*(n-1))/2;
      z=0;
      for i=1:length(K)
        for j=1:length(K)
          if ((K(i)-K(j))^2)<1, continue; end
          z=z+S13(K(i),K(j));
        end
      end
      C(x,m)=z/2;
      if N(x,m)>0
        CC(x,m)=C(x,m)/N(x,m);
      else
        CC(x,m)=0;
      end
      TOT(m,t)=TOT(m,t)+CC(x,m);
    end
end
disp('13')
for m=1:294
    AVG(m)=0;
    for t=1:13
      AVG(m)=AVG(m)+TOT(m,t);
    end
    AVG(m)=AVG(m)/13;
    dev(m)=0;
    for t=1:13
      dev(m)=dev(m)+((AVG(m)-TOT(m,t))^2);
    end
    dev(m)=dev(m)/13;
    dev(m)=sqrt(dev(m));
end
dev=transpose(dev);
```

Table B3. Calculation of EDA-LCON scores

```
% ============================================================
%       Extend of Differential Activity Score calculations for Connectivity (EDA-LCON)
%           by MUHAMMED ERKAN KARABEKMEZ, BOGAZICI UNIVERSITY
% ============================================================
load S1
t=1;
load G1
for m=1:294
   for x=1:length(S1)
      Conn(m,x)=0;
      if G1(m,x)<1, continue; end
      for y=1:length(S1)
         if G1(m,y)<1, continue; end
         Conn(m,x)=Conn(m,x)+S1(x,y);
      end
   end
   ACO(m,t)=0;
   for x=1:length(S1)
      ACO(m,t)=ACO(m,t)+Conn(m,x);
   end
end
clear S1
load S2
disp('1')
t=2;
for m=1:294
   for x=1:length(S2)
      Conn(m,x)=0;
      if G1(m,x)<1, continue; end
      for y=1:length(S2)
         if G1(m,y)<1, continue; end
         Conn(m,x)=Conn(m,x)+S2(x,y);
      end
   end
   ACO(m,t)=0;
   for x=1:length(S2)
      ACO(m,t)=ACO(m,t)+Conn(m,x);
   end
end
clear S2
load S3
disp('2')
t=3;
for m=1:294
   for x=1:length(S3)
      Conn(m,x)=0;
      if G1(m,x)<1, continue; end
      for y=1:length(S3)
```

```
            if G1(m,y)<1, continue; end
            Conn(m,x)=Conn(m,x)+S3(x,y);
        end
    end
    ACO(m,t)=0;
    for x=1:length(S3)
        ACO(m,t)=ACO(m,t)+Conn(m,x);
    end
end
clear S3
load S4
disp('3')
t=4;
for m=1:294
    for x=1:length(S4)
        Conn(m,x)=0;
        if G1(m,x)<1, continue; end
        for y=1:length(S4)
            if G1(m,y)<1, continue; end
            Conn(m,x)=Conn(m,x)+S4(x,y);
        end
    end
    ACO(m,t)=0;
    for x=1:length(S4)
        ACO(m,t)=ACO(m,t)+Conn(m,x);
    end
end
clear S4
load S5
disp('4')
t=5;
for m=1:294
    for x=1:length(S5)
        Conn(m,x)=0;
        if G1(m,x)<1, continue; end
        for y=1:length(S5)
            if G1(m,y)<1, continue; end
            Conn(m,x)=Conn(m,x)+S5(x,y);
        end
    end
    ACO(m,t)=0;
    for x=1:length(S5)
        ACO(m,t)=ACO(m,t)+Conn(m,x);
    end
end
clear S5
load S6
disp('5')
t=6;
```

```
for m=1:294
   for x=1:length(S6)
      Conn(m,x)=0;
      if G1(m,x)<1, continue; end
      for y=1:length(S6)
         if G1(m,y)<1, continue; end
         Conn(m,x)=Conn(m,x)+S6(x,y);
      end
   end
   ACO(m,t)=0;
   for x=1:length(S6)
      ACO(m,t)=ACO(m,t)+Conn(m,x);
   end
end
clear S6
load S7
disp('6')
t=7;
for m=1:294
   for x=1:length(S7)
      Conn(m,x)=0;
      if G1(m,x)<1, continue; end
      for y=1:length(S7)
         if G1(m,y)<1, continue; end
         Conn(m,x)=Conn(m,x)+S7(x,y);
      end
   end
   ACO(m,t)=0;
   for x=1:length(S7)
      ACO(m,t)=ACO(m,t)+Conn(m,x);
   end
end
clear S7
load S8
disp('7')
t=8;
for m=1:294
   for x=1:length(S8)
      Conn(m,x)=0;
      if G1(m,x)<1, continue; end
      for y=1:length(S8)
         if G1(m,y)<1, continue; end
         Conn(m,x)=Conn(m,x)+S8(x,y);
      end
   end
   ACO(m,t)=0;
   for x=1:length(S8)
      ACO(m,t)=ACO(m,t)+Conn(m,x);
   end
```

```
end
clear S8
load S9
disp('8')
t=9;
for m=1:294
   for x=1:length(S9)
      Conn(m,x)=0;
      if G1(m,x)<1, continue; end
      for y=1:length(S9)
         if G1(m,y)<1, continue; end
         Conn(m,x)=Conn(m,x)+S9(x,y);
      end
   end
   ACO(m,t)=0;
   for x=1:length(S9)
      ACO(m,t)=ACO(m,t)+Conn(m,x);
   end
end
clear S9
load S10
disp('9')
t=10;
for m=1:294
   for x=1:length(S10)
      Conn(m,x)=0;
      if G1(m,x)<1, continue; end
      for y=1:length(S10)
         if G1(m,y)<1, continue; end
         Conn(m,x)=Conn(m,x)+S10(x,y);
      end
   end
   ACO(m,t)=0;
   for x=1:length(S10)
      ACO(m,t)=ACO(m,t)+Conn(m,x);
   end
end
clear S10
load S11
disp('10')
t=11;
for m=1:294
   for x=1:length(S11)
      Conn(m,x)=0;
      if G1(m,x)<1, continue; end
      for y=1:length(S11)
         if G1(m,y)<1, continue; end
         Conn(m,x)=Conn(m,x)+S11(x,y);
      end
```

```
    end
    ACO(m,t)=0;
    for x=1:length(S11)
        ACO(m,t)=ACO(m,t)+Conn(m,x);
    end
end
clear S11
load S12
disp('11')
t=12;
for m=1:294
    for x=1:length(S12)
        Conn(m,x)=0;
        if G1(m,x)<1, continue; end
        for y=1:length(S12)
            if G1(m,y)<1, continue; end
            Conn(m,x)=Conn(m,x)+S12(x,y);
        end
    end
    ACO(m,t)=0;
    for x=1:length(S12)
        ACO(m,t)=ACO(m,t)+Conn(m,x);
    end
end
clear S12
load S13
disp('12')
t=13;
for m=1:294
    for x=1:length(S13)
        Conn(m,x)=0;
        if G1(m,x)<1, continue; end
        for y=1:length(S13)
            if G1(m,y)<1, continue; end
            Conn(m,x)=Conn(m,x)+S13(x,y);
        end
    end
    ACO(m,t)=0;
    for x=1:length(S13)
        ACO(m,t)=ACO(m,t)+Conn(m,x);
    end
end
disp('13')
for m=1:294
    AVG(m)=0;
    for t=1:13
        AVG(m)=AVG(m)+ACO(m,t);
    end
    AVG(m)=AVG(m)/13;
```

```
    dev(m)=0;
    for t=1:13
       dev(m)=dev(m)+((AVG(m)-ACO(m,t))^2);
    end
    dev(m)=dev(m)/13;
    dev(m)=sqrt(dev(m));
end
dev=transpose(dev);
```

Table B4. Permutation test for EDA-LCC scores

```
% ============================================================
%                  EDA-LCC scores for 1000 randomly deactivating networks
%          by MUHAMMED ERKAN KARABEKMEZ, BOGAZICI UNIVERSITY
% ============================================================
disp(datestr(now))
w=0.019;
hata=0;
load G1
u=0;
for b=1:1000
   u=u+1;
   load S1
   SS=S1;
   TOTS1=0;
   for m=1:294
      TOT(m,1)=0;
      for x=1:length(S1)
         i=1;
         K=0;
         n=0;
         if G1(m,x)<1, continue; end
         for y=1:length(S1)
            if S1(x,y)<1, continue; end
            if G1(m,y)<1, continue; end
            n=n+S1(x,y);
            K(i)=y;
            i=i+1;
         end
         N(x,m)=(n*(n-1))/2;
         z=0;
         for i=1:length(K)
            for j=1:length(K)
               if ((K(i)-K(j))^2)<1, continue; end
               z=z+S1(K(i),K(j));
            end
         end
         C(x,m)=z/2;
         if N(x,m)>0
            CC(x,m)=C(x,m)/N(x,m);
         else
            CC(x,m)=0;
         end
         TOT(m,1)=TOT(m,1)+CC(x,m);
      end
      TOTS1=TOTS1+TOT(m,1);
   end
   for e=2:13
      l=0;
```

```
for x=1:length(S1)
   subtot=0;
   for y=1:length(S1)
      if x==y, continue; end
      subtot=subtot+S1(x,y);
   end
   if subtot<1, continue; end
   l=l+1;
end
t=0.019;
diff=0;
for r=1:100
   p=randint(1,round(l*t),[1,l]);
   for x=1:length(p)
      for y=1:length(SS)
         SS(p(x),y)=0;
         SS(y,p(x))=0;
      end
   end
   TOTO=0;
   for m=1:294
   TOT(m,e)=0;
      for x=1:length(SS)
         i=1;
         K=0;
         n=0;
         if G1(m,x)<1, continue; end
         for y=1:length(SS)
            if SS(x,y)<1, continue; end
            if G1(m,y)<1, continue; end
            n=n+SS(x,y);
            K(i)=y;
            i=i+1;
         end
         N(x,m)=(n*(n-1))/2;
         z=0;
         for i=1:length(K)
            for j=1:length(K)
               if ((K(i)-K(j))^2)<1, continue; end
               z=z+SS(K(i),K(j));
            end
         end
         C(x,m)=z/2;
         if N(x,m)>0
            CC(x,m)=C(x,m)/N(x,m);
         else
            CC(x,m)=0;
         end
         TOT(m,e)=TOT(m,e)+CC(x,m);
```

```
            end
            TOTO=TOTO+TOT(m,e);
        end
        diff=w-((TOTS1-TOTO)/TOTS1);
        if abs(diff)<0.001, break, end
        if diff<0
            t=t-.001;
        else
            t=t+.001;
        end
        SS=S1;
        clear p
    end
    if r==100
        hata=hata+1;
    else
        r=r;
    end
    S1=SS;
    TOTS1=TOTO;
end
for m=1:294
    AVG(m)=0;
    for e=1:13
        AVG(m)=AVG(m)+TOT(m,e);
    end
    AVG(m)=AVG(m)/13;
    dev(m,b)=0;
    for e=1:13
        dev(m,b)=dev(m,b)+((AVG(m)-TOT(m,e))^2);
    end
    dev(m,b)=dev(m,b)/13;
    dev(m,b)=sqrt(dev(m,b));
end
clear S1
save dev dev
if u==50
    disp(b)
    disp(datestr(now))
    u=u-50;
else
    continue
end
end
dev=transpose(dev);
save dev dev
disp(datestr(now))
```

Table B5. Permutation test for EDA-LCON scores

```
% ========================================================
%                    EDA-LCON scores for 1000 randomly deactivating networks
%          by MUHAMMED ERKAN KARABEKMEZ, BOGAZICI UNIVERSITY
% ========================================================
disp(datestr(now))
w=0.028;
hata=0;
load G1
u=0;
for b=1:1
   u=u+1;
   load S1
   SS=S1;
   TOTS1=0;
   for m=1:294
      for x=1:length(S1)
         Conn(m,x)=0;
         if G1(m,x)<1, continue; end
         for y=1:length(S1)
            if G1(m,y)<1, continue; end
            Conn(m,x)=Conn(m,x)+S1(x,y);
         end
      end
      ACO(m,1)=0;
      for x=1:length(S1)
         ACO(m,1)=ACO(m,1)+Conn(m,x);
      end
      TOTS1=TOTS1+ACO(m,1);
   end
   for e=2:13
      i=0;
      for x=1:length(S1)
         subtot=0;
         for y=1:length(S1)
            if x==y, continue; end
            subtot=subtot+S1(x,y);
         end
         if subtot<1, continue; end
         i=i+1;
      end
      t=0.028;
      diff=0;
      for r=1:100
         c=randint(1,round(i*t),[1,i]);
         for x=1:length(c)
            for y=1:length(SS)
               SS(c(x),y)=0;
               SS(y,c(x))=0;
```

```
            end
        end
        TOT=0;
        for m=1:294
            for x=1:length(SS)
                Conn(m,x)=0;
                if G1(m,x)<1, continue; end
                for y=1:length(SS)
                    if G1(m,y)<1, continue; end
                    Conn(m,x)=Conn(m,x)+SS(x,y);
                end
            end
            ACO(m,e)=0;
            for x=1:length(SS)
                ACO(m,e)=ACO(m,e)+Conn(m,x);
            end
            TOT=TOT+ACO(m,e);
        end
        diff=w-((TOTS1-TOT)/TOTS1);
        if abs(diff)<0.001, break, end
        if diff<0
            t=t-.001;
        else
            t=t+.001;
        end
        SS=S1;
        clear c
    end
    if r==100
        hata=hata+1;
    else
        r=r;
    end
    S1=SS;
    TOTS1=TOT;
end
for m=1:294
    AVG(m)=0;
    for e=1:13
        AVG(m)=AVG(m)+ACO(m,e);
    end
    AVG(m)=AVG(m)/13;
    dev(m,b)=0;
    for e=1:13
        dev(m,b)=dev(m,b)+((AVG(m)-ACO(m,e))^2);
    end
    dev(m,b)=dev(m,b)/13;
    dev(m,b)=sqrt(dev(m,b));
end
```

```
    clear S1
    save dev dev
    if u==100
       disp(b)
       disp(datestr(now))
       u=u-100;
    else
       continue
    end
end
dev=transpose(dev);
save dev dev
disp(datestr(now))
```

Table B6. Creation of PCC based distance matrices

```
% ============================================================
%                Distance matrices creation base on Pearson correlation coefficient (PCC)
%                by MUHAMMED ERKAN KARABEKMEZ, BOGAZICI UNIVERSITY
% ============================================================
load E
% expression matrices is loaded
for x=1:5668
   Eavg(x)=0;
   for t=1:14
      Eavg(x)=Eavg(x)+E(x,t);
   end
   Eavg(x)=Eavg(x)/14;
end
disp('1')
x=0;
for x=1:5668
   for y=1:5668
      totup=0;
      totdown1=0;
      totdown2=0;
      totdown=0;
      r=0;
      for t=1:14
         totup=totup+((E(x,t)-Eavg(x))*((E(y,t)-Eavg(y))));
         totdown1=totdown1+((E(x,t)-Eavg(x))*(E(x,t)-Eavg(x)));
         totdown2=totdown2+((E(y,t)-Eavg(y))*(E(y,t)-Eavg(y)));
      end
      totdown=totdown1*totdown2;
      totdown=sqrt(totdown);
      r=totup/totdown;
      g(x,y)=1-r;
   end
   disp(x)
end
save g g
```

Table B7. Determination of interface nodes

```
% ==========================================================
%                    Extracting interfaces
%        by MUHAMMED ERKAN KARABEKMEZ, BOGAZICI UNIVERSITY
% ==========================================================
load NP2
load k
% k is the array of orders of P module genes in NP network
load m
% m is the array of orders of M module genes in NP network
i=1;
disp('loaded')
for x=1:length(k)
   for y=1:length(m)
      if NP2(k(x),m(y))~=0
         faceP(i)=x;
         faceM(i)=y;
         i=i+1;
         disp(i)
      end
   end
end
faceP=transpose(faceP);
faceM=transpose(faceM);
% faceP and faceM are orders of interface genes in k and m arrays
```

REFERENCES

Ashburner, M., C. A. Ball, J. A. Blake, D. Botstein, H. Butler, J. M. Cherry, A. P. Davis, K. Dolinski, S. S. Dwight, J. T. Eppig, M. A. Harris, D. P. Hill, L. Issel-Tarver, A. Kasarskis, S. Lewis, J. C. Matese, J. E. Richardson, M. Ringwald, G. M. Rubin, G. Sherlock, 2000, "Gene ontology: tool for the unification of biology. The Gene Ontology Consortium", *Nat Genet* Vol. 25, pp. 25–29.

Bader, J. S., A. Chaudhuri, J. M. Rothberg & J. Chant, 2004, "Gaining confidence in high-throughput protein interaction networks", *Nature Biotechnol.* Vol. 22, pp. 78–85.

Barabasi, A-L., and Z.N. Oltvai, 2004, "Network biology: understanding the cell's functional organization", *Nat. Rev. Genet*. Vol. 5, pp.101–113.

Batada, N. N., T. Reguly, A. Breitkreutz, L. Boucher, B. J. Breitkreutz, D. H. Laurence, M. Tyers, 2006, "Stratus not altocumulus: A new view of the yeast protein interaction network", *PLoS Biol* Vol. 4(10), e317.

Benjamini, Y. and Y. Hochberg, 1995, "Controlling the false discovery rate—A practical and powerful approach to multiple testing", *J Royal Stat Soc Ser B Methodol.* Vol. 57, pp. 289–300.

Bertin, N., N. Simonis, D. Dupuy, M. E. Cusick, J. D. Han, H. B. Fraser, F. P. Roth, M. Vidal, 2007, "Confirmation of organized modularity in the yeast interactome", *PLoS Biol* Vol. 5, e153.

Brink, J. V. D., P. Daran-Lapujade, J. T. Pronk and J. H. de Winde, 2008, "New insight into the Saccharomyces cerevisiae fermentation switch: Dynamic transcriptional response to anaerobicity and glucose-excess", *BMC Genomics* Vol. 9, art. 100.

Dhillon, I. S., E. M. Marcotte, U. Roshan, 2003, "Diametrical clustering for identifying anti-correlated gene clusters", *Bioinformatics* Vol. 19, pp. 1612–1619.

Eisen, M. B., P. T. Spellman, P. O. Brown & D. Botstein, 1998, "Cluster analysis and display of genome-wide expression patterns", *Proc. Natl. Acad. Sci. USA* Vol. 95, pp. 14863-14868.

Fazio, A., M. C. Jewett, P. Daran-Lapujade, R. Mustacchi, R. Usaite, J. T. Pronk, C. T. Workman, J. Nielsen, 2008, "Transcription factor control of growth rate dependent genes in *Saccharomyces cerevisiae*: a three factor design", *BMC Genomics* Vol. 9, art.341.

Ge, H., Z. Liu, G. M. Church & M. Vidal, 2001, "Correlation between transcriptome and interactome mapping data from Saccharomyces cerevisiae", *Nature Genet.* Vol. 29, pp. 482–486.

Gopalacharyulu, P. V., V. R. Velagapudi, E. Lindfors, E. Halperin and M. Oresic, 2009, "Dynamic network topology changes in functional modules predict responses to oxidative stress in yeast", *Mol. BioSyst.*Vol. 5, pp. 276-287.

Guet, C. C., M. B. Elowitz, W. Hsing & S. Leibler, 2002, "Combinatorial synthesis of genetic networks", *Science* Vol. 296, pp. 1466-1470.

Gunsalus, K. C., H. Ge, A. J. Schetter, D. S. Goldberg, J. D. Han, T. Hao, G. F. Berriz, N. Bertin, J. Huang, L. S. Chuang, N. Li, R. Mani, A. A. Hyman, B. Sönnichsen, C. J. Echeverri, F. P. Roth, M. Vidal & F. Piano, 2005, "Predictive models of molecular machines involved in Caenorhabditis elegans early embryogenesis" *Nature* Vol. 436, pp. 861–865.

Han, J. D., N. Bertin, T. Hao, D. S. Goldberg, G. F. Berriz, L. V. Zhang, D. Dupuy, A. J. Walhout, M. E. Cusick, F. P. Roth, M. Vidal, 2004, "Evidence for dynamically organized modularity in the yeast protein–protein interaction network", *Nature* Vol. 430, pp. 88–93.

Ideker, T., V. Thorsson, J. Ranish, R. Chirstmas, J. Buhler, J. Eng, R. Bumgarner, D. Goodlett, R. Aebersold, L. Hood, 2001, "Integrated genomic and proteomic analysis of a systematically perturbed metabolic network", *Science* Vol. 292, pp. 929–934.

Jansen, R., D. Greenbaum & M. Gerstein, 2002, "Relating whole-genome expression data with protein-protein interactions", *Genome Res.* Vol. 12, pp. 37–46.

Jansen, R., H. Yu, D. Greenbaum, Y. Kluger, N. J. Krogan, S. Chung, A. Emili, M. Snyder, J. F. Greenblatt, M. Gerstein, 2003, "A Bayesian networks approach for predicting protein-protein interactions from genomic data", *Science* Vol. 302, pp. 449–453.

Kresnowati, M. T. A. P., W. A. van Winden, M. J. H. Almering, A. ten Pierick, C. Ras, T. A. Knijnenburg, P. Daran-Lapujade, J. T. Pronk, J. J. Heijnen and J. M. Daran, 2006, "When transcriptome meets metabolome: fast cellular responses of yeast to sudden relief of glucose limitation", *Molecular Systems Biology* art. 49.

Lee, I., S. V. Date, A. T. Adai, & E. M. Marcotte, 2004, "A probabilistic functional network of yeast genes", *Science* Vol. 306, pp. 1555–1558.

Lu, H., X. Zhu, H. Liu, G. Skogerbo, J. Zhang, Y. Zhang, L. Cai, Y. Zhao, S. Sun, J. Xu, D. Bu, R. Chen, 2004, "The interactome as a tree—an attempt to visualize the protein-protein interaction network in yeast", *Nucleic Acids Res* Vol. 32, pp. 4804–4811.

Luscombe, N. M., M. M. Babu, H. Yu, M. Snyder, S. A. Teichmann, M. Gerstein, 2004, "Genomic analysis of regulatory network dynamics reveals large topological changes" *Nature* Vol. 431, pp. 308–312.

Segal, E., M. Shapira, A. Regev, D. Pe'er, D. Botstein, D. Koller, and N. Friedman, 2003, "Module networks: Identifying regulatory Reconciling Expression and Regulatory Networks modules and their condition- specific regulators from gene expression data", *Nat. Genet.* Vol. 34, pp. 166–176.

Vidal, M., 2001, “A biological atlas of functional maps”, *Cell* Vol. 104, pp. 333–339.

Walhout, A. J., J. Reboul, O. Shtanko, N. Bertin, P. Vaglio, H. Ge, H. Lee, L. D. Stamm, K. C. Gunsalus, A. J. Schetter, D. G. Morton, K. J. Kempheus, V. Reinke, S. K. Kim, F. Piano, and M. Vidal, 2002, “Integrating interactome, phenome, and transcriptome mapping data for the C. elegans germline”, *Curr. Biol.* Vol. 12, pp. 1952–1958.

Wang, Y., M. Pierce, L. Schneper, C. G. Güldal, X. Zhang, S. Tavazoie, J. R. Broach, 2004, “Ras and Gpa2 mediate one branch of a redundant glucose signaling pathway in yeast”, *PLoS Biol* Vol. 2, pp. 610-622.

Wen, X., S. Fuhrman, G. S. Michaels, D. B. Carr, S. Smith, J. L. Barker, R. Somogyi, 1998, “Large-scale temporal gene expression mapping of central nervous system development”, *Proc. Natl Acad. Sci. USA* Vol. 95, pp. 334–339.

Xia, K., D. Dong, H. Xue, S. Zhu, J. Wang, Q. Zhang, L. Hou, H. Chen, R. Tao, Z. Huang, Z. Fu, Y. G. Chen, J. D. Han, 2006, “Identification of the proliferation/differentiation switch in the cellular network of multicellular organisms”, *PLoS Comput Biol* Vol. 2, e145.

Xue, H., B. Xian, D. Dong, K. Xia, S. Zhu, Z. Zhang, L. Hou, Q. Zhang, Y. Zhang and J. D. Han, 2007, “A modular network model of aging”, *Molecular Systems Biology* Vol. 3, art.147.

Yu, H., P. M. Kim, E. Sprecher, V. Trifonov, M. Gerstein, 2007, “The importance of bottlenecks in protein networks: correlation with gene essentiality and expression dynamics”, *PLoS Comput Biol* Vol. 3, e59.

Yeang, C. H., T. Ideker, T. Jaakkola, 2004, “Physical network models”, *J Comput Biol* Vol. 11**,** pp. 243-262.

Zaman, S., S. I. Lippman, L. Schneper, N. Slonim and J. R. Broach, 2009, “Glucose regulates transcription in yeast through a network of signaling pathways”, *Molecular Systems Biology* Vol. 5, art.245.