FUNCTIONAL ANALYSIS OF THE GENES IN THE GLUCOSE METABOLIC NETWORK BY A SYSTEM BASED MODULAR APPROACH

by

Okan Yüzüak

B.S., Chemical Engineering, Yıldız Technical University, 2008

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
Boğaziçi University
2012

to my family

ACKNOWLEDGEMENTS

Firstly, I would like to express my sincere thanks to my thesis supervisor Prof. Betül Kırdar and coadvisor Assoc. Prof. Elif Özkirimli Ölmez. Their advices helped me to improve my work and their guidance was invaluable throughout my thesis. It was a great honor for me to work with them during my graduate thesis.

I am very grateful to my thesis committee members; Prof. Zeynep İlsen Önsan, Assoc. Prof. Kazım Yalçın Arga and Assoc. Prof. Ebru Toksoy Öner.

I wish to express my gratitude to Duygu Dikicioğlu and Erkan Karabekmez, who deserve special thanks for their guidance and help throughout my work.

Very special thanks goes to my friends Melek Selcen Başar, Aybüke Leba and Caner Ülgüel for their friendship and endless support throughout my master education without any hesitation.

I also would like to express my gratitude to other members of Biosystems Group; Ayça Cankorur Çetinkaya, Elif Dereli, Serpil Eraslan, Esra Börklü Yücel, Ceyda Kasavi, Şebnem Öç and Saliha Durmuş Tekir for their valuable advice, kind support and help. I would also like to thank all my friends in Chemical Engineering Department at Boğaziçi University.

Cordial thanks are for Bilgi Dedeoğlu, Nurettin Bektaş and Yakup Bal for their technical assistance and help during my thesis.

I would like to express my very great appreciation to my family. They always believed in me, even when I didn’t. This thesis would not have been possible without their everlasting support.

Finally, I cannot find the right words for the one person who I like to thank the most, my girlfriend Burcu Yoğurtçu for her endless patience, support and encouragement during my thesis study.

ABSTRACT

FUNCTIONAL ANALYSIS OF THE GENES IN THE GLUCOSE METABOLIC NETWORK BY A SYSTEM BASED MODULAR APPROACH

In this study, all proteins associated with the Process Gene Ontology (GO) term “Glucose Metabolic Process” were collected and a network of all reported interactions among these entities was constructed with a stringent confidence threshold. Two networks were constructed at the beginning of the study; one general network consisting of a global interaction network of yeast obtained from the STRING database (Global Network, GN) and another network that was constructed by including only the interactions for the nodes of the Glucose Metabolic Process Gene Ontology Term (Glucose Metabolic Network, GMN). These two networks were then compared and the GMN was determined to be a scale free biological network. The parameters used in module identification were optimized in order to increase module specificity. Modular structures of GN and GMN were investigated by using Molecular Complex Detection (MCODE) plug-in of Cytoscape. The modules of GN had more members and these modules were associated with very general GO terms but the modules that were identified in GMN were more compact and functionally more distinctive with significant gene ontological associations. 65 modules were identified in GMN in the present study and eight five membered modules were selected and analyzed. These modules were associated with a specific function in the cell, either as a part of a complex or as a part of a signaling cascade. The extension of the modules provided additional nodes that were associated with the same function or other functions. All associations were verified from the available literature on the investigated pathways and complexes. The functional robustness of the method in identification of module members associated with the same cellular process suggests that this modular approach can be used to assign probable functions to the unannotated or uncharacterized loner genes present, or to assign novel additional functions to the proteins with already known functional associations. As a result of this investigation, YGR067c was potentially related to “Ubiquitin Dependent Protein Catabolic Process”, Roy1p was potentially related to “Interphase” and Nnk1p was potentially related to “Proteolysis” GO Process Terms.

ÖZET

GLİKOZ METABOLİK AĞINDAKİ GENLERİN SİSTEM BAZLI MODÜLER YAKLAŞIM İLE iŞLEVSEL ANALİZİ

Bu çalışmada, Glikoz Metabolik Proses ile ilintili bütün proteinler ve bunlar arasındaki en yüksek değerlikli etkileşimler belirlenmiştir. Çalışmanın başlangıcında Glikoz Metabolik Ağ ve Genel Ağ adı altında iki adet protein-protein etkileşim ağı oluşturulmuştur. Genel Ağ (GN) mayada STRING veri tabanındaki etkileşimler temel alınarak oluşturulmuş, Glikoz Metabolik Ağı (GMN) ise sadece Glikoz Metabolik Proses gen ontoloji terimine ait olan proteinler arasındaki etkileşimler temel alınarak oluşturulmuştur. Bu iki ağın topolojik nitelikleri belirlenmiş ve karşılaştırılmıştır. Glikoz metabolik ağının, diğer biyolojik ağlarda olduğu gibi, ölçeksiz bir ağ olduğu saptanmıştır. Modül tanımlamada kullanılan ve modüllerin özgünlüğünü belirlemeye yardımcı olan parametreler optimize edilmiştir. Genel Ağ modül üyelerinin sayısının daha fazla ve bu modüllerin daha genel GO süreç terimleri ile ilintili olduğu belirlenmiştir. Glikoz Metabolik Ağ'ında toplam 65 adet modül bulunmuştur ve bu ağ içerisinde saptanan modül üyelerinin daha kompakt ve daha özgün işlevlerle ilgili GO süreç terimleri içermesi sebebiyle Glikoz Metabolik Ağı'nda bulunan beş elemanlı modüller incelenmiştir. Bu modüllerin belirli görevlere sahip olduğu ve modül elemanlarının ya kompleksin bir parçası ya da bir sinyalizasyon yolunun elemanı olduğu belirlenmiştir. Modüllerin genişletilmesi ayrıca aynı işleve ya da farklı fonksiyonel ilişkiye sahip olan proteinlerin görülmesini sağlamıştır. Bütün bu ilişkiler mevcut literatürdeki araştırılmış sinyal yolları ve kompleksler aracılığı ile doğrulanmıştır. Modül incelemede kullanılan metodun güvenilirliği, aynı modül içerisinde bulunan proteinler hakkında olası fonksiyonlar atamamızı sağlamıştır. Bu şekilde tanımlanmamış ya da karakterize edilmemiş genler hakkında tanımlama yapılabilir ya da daha önceden fonksiyonu tanımlanmış proteinlere ek fonksiyonlar atanabilir. YGR067c, Roy1p, Nnk1p proteinlerinin sırasıyla "Ubiquitin'e Bağlı Protein Katabolik Süreç", "İnterfaz" ve "Proteoliz" GO süreç terimleri ile ilintili oldukları bulunmuştur.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF SYMBOLS

b_m	Betweenness
C_n	Clustering coefficient
e_n	Number of connected pairs between all neighbors of *n*
K	Degree
k_n	Number of neighbors of *n*
M	Maximum number of edges
N	Number of edges between the neighbors of *n*
N	Node
P(k)	Degree distribution
γ	Degree exponent
$\Gamma(i,m,j)$	Number of shortest paths between i^{th} and j^{th} nodes

LIST OF ACRONYMS/ABBREVIATIONS

Act	Actin
ATP	Adenosine Triphosphate
CS	Confidence Score
CSRE	Carbon Source Responsive Element
DNA	Deoxyribonucleic Acid
FS	Functional Subset
GMN	Glucose Metabolic Network
GN	General Network
GO	Gene Ontology
GTP	Guanosine Triphosphate
MAP	Mitogen-Activated Protein
MCODE	Molecular Complex Detection
mRNA	Messenger RNA
PKA	AMP-Dependent Protein Kinase
PPI	Protein-Protein Interaction
rRNA	Ribosomal RNA
S. cerevisiae	Saccharomyces Cerevisiae
SGD	Saccharomyces Genome Database
snRNA	Small Nuclear RNA
STRING	Search Tool for the Retrieval of Interacting Genes/Proteins
SUMO	Small Ubiquitin-Like Modifier
TF	Transcription Factor

TOR	Target of Rapamycin
TORC1	TOR Complex 1
TREX	Transcription Export
tRNA	Transfer RNA
UDP	Uridine Diphosphate
V-ATPase	Vacuolar-type H(+)-ATPase

1. INTRODUCTION

Saccharomycess cerevisiae, a species of budding yeast, is one of the most studied eukaryotic model organisms in molecular and cellular biology. "*Saccharomyces*" is the name of the "sugar mold" or the "sugar fungus" and *cerevisiae* means "beer" stemming from Latinized Greek.

The genome of the yeast *Saccharomyces cerevisiae* was completely sequenced through an international effort involving nearly 600 scientists from Europe, North America, and Japan. It is a large collaborative effort to determine the first complete genome sequence of a eukaryote. The genome sequence was released in the public domain on April 24th, 1996 and updates have been performed on the Saccharomyces Genome Database (SGD) regularly (Goffeau *et al*., 1996).

Saccharomyces cerevisiae, unlike many more complex eukaryotes, is a unicellular organism that may be grown on defined medium allowing the researchers to have control over its chemical and physical environment (Goffeau *et al*., 1996). This species is a well-studied model organism for which a large amount of interaction data is available (Uetz *et al*., 2000).

Post-genomic biology has focused on systems biology for the investigation of large-scale organization of the cell due to the shortcomings of classical molecular biology, which could not handle such sophisticated systems.

The recent increase in the number of sequenced genomes and the amount of genome-scale experimental data lead motivated the researchers to use computational techniques for investigations by the way many algorithms, programs and databases constructed analyzing the available data.

Computational studies of biological networks are milestones of systems biology. A variety of modeling environments have been developed to simulate biochemical reactions and gene transcription kinetics, cellular physiology, and metabolic control. Such models promise to transform biological research by providing a framework to systematically examine and experimentally verify knowledge of a pathway; manage the complexity of hundreds or potentially thousands of cellular components and interactions; and show properties and sudden consequences of different pathway configurations (Shannon *et al.,* 2003).

1.1. Glucose Metabolism in *Saccharomyces cerevisiae*

Nutrient availability is the major factor that controls growth and development in living organisms. For yeasts, as many other organisms, glucose is the preferred carbon and energy source (Rolland *et al.,* 2002). Availability of glucose or other easily fermentable sugars are important for the rearrangement of the metabolism to utilize its resources for aerobic respiration, anaerobic respiration, or fermentation (Livas *et al*., 2011).

Glucose causes various changes in yeast due to the efficient use of available carbon sources. These changes include regulation of gene expression at the transcriptional, posttranscriptional, translational, and posttranslational levels. For these adaptations to occur properly; three glucose sensing systems are well documented for yeast. Each system detects and transmits the glucose signal differently. One mechanism operates through the Snf1p kinase to cause a repression of gene expression when glucose levels are high. Another mechanism works through the Snf3p and Rgt2p glucose sensors to induce expression of genes encoding glucose transporters. A third glucose sensing mechanisms employs the G-protein-coupled receptor; Gpr1p and cyclic AMP as a second messenger (Kaniak *et al.,* 2004). The cell must sense glucose and transmit a signal to the appropriate targets at first, for these adaptations to take place.

It was previously reported that the yeast Saccharomyces cerevisiae shows a great variety of cellular responses to glucose via several glucose-sensing and signaling pathways (Gelade *et al.,* 2003). The Glucose Signal Transduction Pathway was explained and its connection with the highly interconnected regulatory network of glucose sensing pathways was described previously by Johnston and Kim (Johnston and Kim, 2005). The effect of glucose depletion on protein synthesis was investigated previously and this investigation underlined the close connection between the nutrient status of the cell and its translational capacity (Ashe *et al*., 1999). The effect of glucose on the regulation of carbon metabolism, advances in the understanding of glucose sensing and signaling in yeast for the elucidation of nutrient-sensing mechanisms in other eukaryotic organisms, and the identification of several new components of glucose-induction pathway was previously discussed by Rolland *et al*., 2002. A study related to the early elements in glucose signaling and the similarities and differences related to the response of yeasts and mammalian cells to glucose was performed previously. It was shown that *Saccharomyces cerevisiae,* sensing systems for other nutrients share some of the characteristics of the glucose-sensing pathways (Gancedo, 2008).

The requirement of both extracellular and intracellular glucose signaling for properly setting the cell size in glucose media was mentioned previously (Vanoni *et al*., 2005). In the study of Zaman *et al.* it was shown that; yeast cells set their transcriptional growth pattern and their growth rate only on their sensing of the nutritional sources, rather than on the products and activity relation on metabolism of those nutrients (Zaman *et al*., 2009). The difference in the regulation of Nrg1p and Nrg2p in response to carbon source at both the RNA and protein levels was shown previously and the differences in expression and function of these two repressors contributed to the complex regulation of the large set of glucose-repressed genes (Berkey *et al*., 2004).

The effects of glucose were previously discussed on the yeast physiology. The effects of high (>1%) and low (0.01%) glucose were studied and these levels (very low (0.01%), low (0.1%) and high glucose signals (1.0%)) were characterized by transcript profiling. It was shown that yeast is more sensitive to very low glucose signals than was previously thought, and that yeast displays different responses to these different glucose signals (Yin *et al*., 2003).

A study on the two high and low affinity glucose sensing and signaling genes; Rgt2p and Snf3p; which were the sensors of extracellular glucose and involved in generation of an intracellular glucose signal that triggers the induction of HXT gene expression, was performed previously (Ozcan *et al*., 1998). The ability of Snf3p to transduce a signal in the complete absence of extracellular glucose was reported previously (Dlugai *et al*., 2001). A study related to interaction of Mth1p with glucose sensors Snf3p and Rgt2p was performed previously. The effectors to this interaction such as mutations in Mth1p and glucose concentration in the medium were observed. A model independent of Snf3p and Rgt2p related to glucose signaling was presented (Lafuente *et al*., 2000). Another study proposes that glucose acts via Grr1p to promote the degradation of Mth1p, which leads to phosphorylation and dissociation of Rgt1p from HXT promoters thus activating HXT gene expression (Flick *et al*., 2003).

A global analysis of gene expression to assess the roles of Ras2p and Gpa2p was resulted that the activation of these two genes was important and their mediation were controlled by protein kinase A (PKA), which also has a role in the glucose signaling (Wang *et al*., 2004). A study regarding PKA suggested different classes of genes that can be induced or repressed by glucose in the absence of PKA. Additionally, involvement of PKA and variety of pathways alone or in combination in the triggering of the transcriptional responses to glucose was told (Livas *et al*., 2011). The regulation of Snf1 Kinase was proposed to be achieved by the beta subunits and a novel signaling pathway was suggested due to this regulation in response to glucose phosphorylation (Vincent *et al*., 2001). The regulatory role of glucose related to interaction between Snf1p and Snf4p was reported previously. This regulation was related to glucose level; as Snf4p binds to the Snf1 regulatory domain in low glucose, whereas in high glucose the regulatory domain binds to the kinase domain of Snf1p itself (Jiang and Carlson, 1996). A previous study showed that the Snf1p kinase controls glucose repression in yeast by modulating interactions between the Mig1p repressor and the Cyc8p-Tup1p co-repressor. Regulation of this interaction was proposed as the molecular switch that controls transcriptional repression and de-repression (Chronakis *et al*., 2002).

The requirement of the nuclear exportin Msn5 for nuclear export of the Mig1 glucose repressor of *Saccharomyces cerevisiae* was studied previously and it was reported that Mig1p contains a nuclear export signal, which was phosphorylated by Snf1p, and this signal was recognized by Msn5p (Devit and Johnston, 1999). The main role of Hxk2p in the glucose signaling pathway was proposed to be the interaction with Mig1p to form a complex located in the *S. cerevisiae* nucleus (Moreno *et al*., 2005).

1.1.1. Glucose Metabolic Process

The Glucose Metabolic Process consists of chemical reactions and pathways involving glucose. The child terms of the parent Glucose Metabolic Process term which are used in the present study are gluconeogenesis, glycolysis, glycogen biosynthetic process and glycogen catabolic process, ethanol biosynthetic process, glucose-1-phosphate metabolic process, glucose-6-phosphate metabolic process, UDP-glucose metabolic process, glucose catabolic process to ethanol and pentose-phosphate shunt.

1.2. Properties of Biological Networks

With increasing use of computational techniques in systems biology, the investigation of biological molecules and their behaviors became easier and faster in comparison to the experimental studies of biological systems.

A dense network of molecular interactions is formed inside a cell by proteins, nucleic acids and small molecules. The architecture of molecular networks can reveal important principles of cellular organization and function similar to the way that protein structure tells us about the function and organization of a protein. Maps of several biological networks such as metabolic networks, protein-protein and protein-DNA interactions are generated by large-scale experiments and integration of published data (Spirin *et al*., 2003).

For better understanding of biological networks, the graph theoretical analyzes are being performed. The proteins in the graphs are called the nodes and the interactions between them are called the edges (Spirin *et al*., 2003). The degree specifies the number of edges that a node has to connect to other nodes. It is symbolized by the letter; k. In directed networks, in which edges have selected directions, nodes have k_{in} degrees to specify the edges pointing them and k_{out} degrees to specify the edges outgoing from the nodes.

If there is a high diversity of node degrees and no typical node in the network, which would be used to characterize the rest of the nodes; these networks are described as ‘scale-free’ due to the absence of a typical degree (Albert, 2005). A few highly connected nodes exist in scale-free networks and that is the distinguishing feature of scale-free networks, which participate in a very large number of metabolic reactions (Barabasi *et al.*, 2002). Biological networks have a scale-free network structure and the scale free networks have some nodes, which have high connectivity with other nodes (Figure 1.1).

Scale-free networks follow a power law distribution where

$P(k) \sim k-\gamma$, γ is degree exponent

and they are extremely heterogeneous, where a few highly connected nodes (hubs) are dominant and rest of the less connected nodes are linked to network through these hubs (Barabási and Oltvai, 2004) (Figure 1.2).

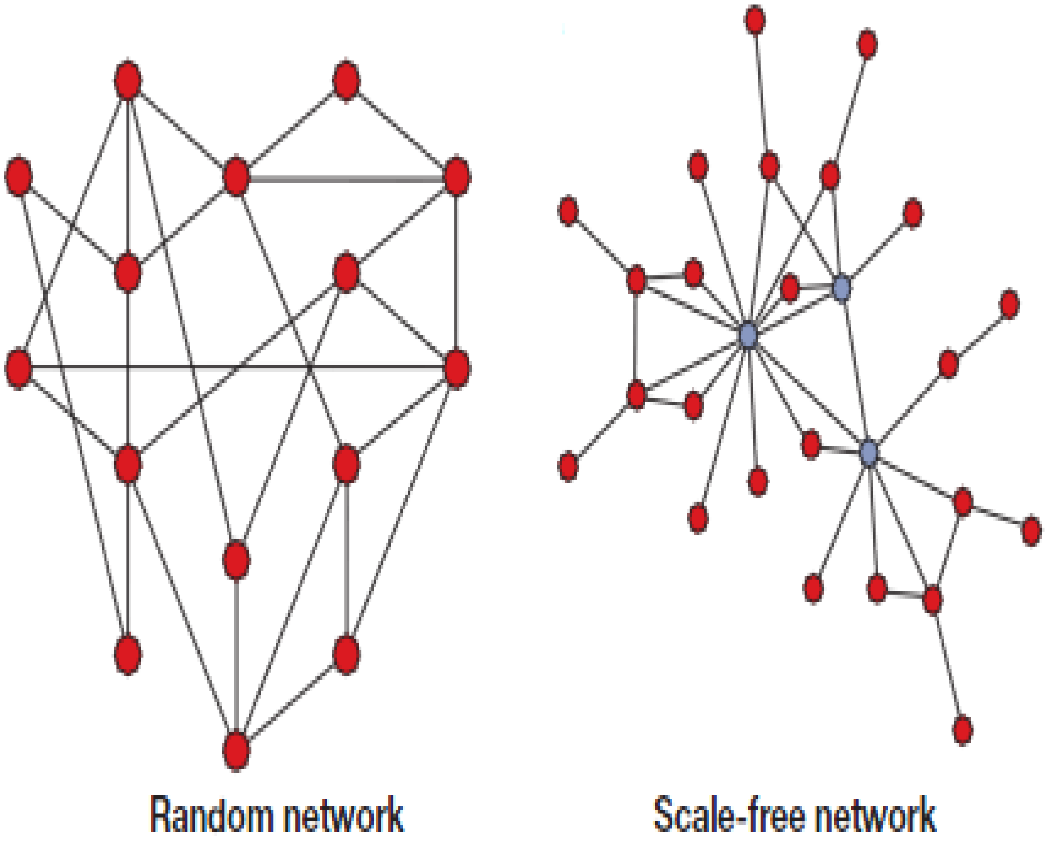


Figure 1.1. Representation of random and scale free networks (Barabási & Oltvai, 2004).

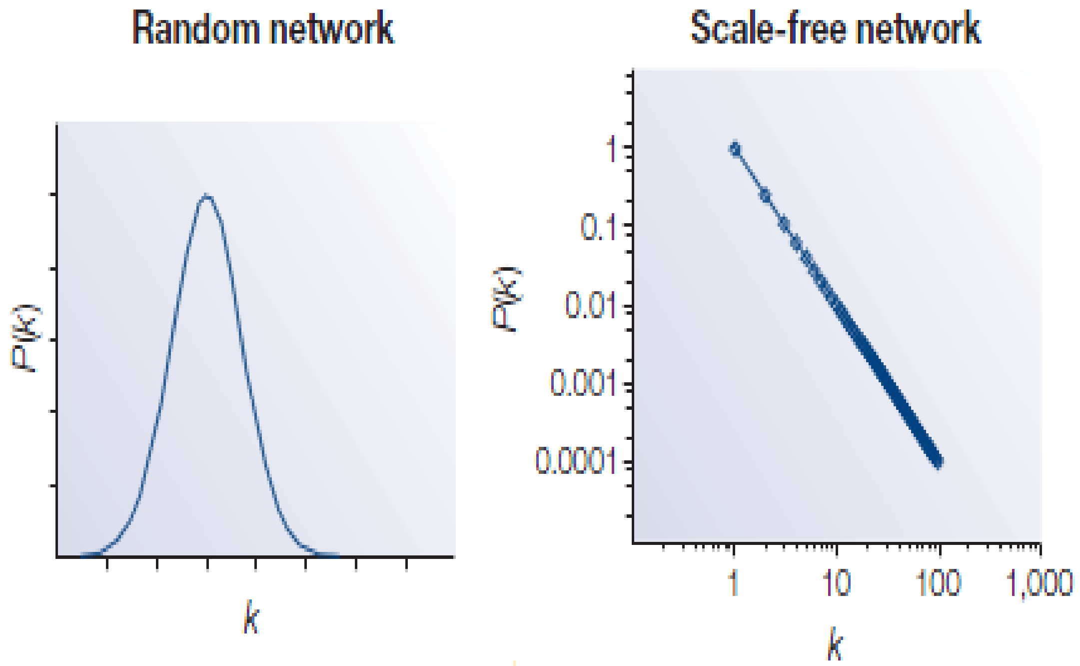


Figure 1.2. Distribution of random and scale free networks (Barabási & Oltvai, 2004).

1.2.1. Topological Analysis of Networks

Analyzing the connections between biological interpretations and topological properties is one of the major aims of systems biology. The distribution of the components and their relations with each other can be revealed by topological analysis of complex biological networks. Barabasi and Oltvai showed that changes in clustering coefficient and connectivity are measures of biological activity in a PPI (Barabási & Oltvai, 2004). Several properties of networks would help revealing novel biological information from the constructed biological networks.

1.2.1.1. Clustering Coefficient. In undirected networks, the clustering coefficient C_n of a node n is defined as;

$$C_n = 2e_n/(k_n(k_n-1)) \tag{1.1}$$

where k_n is the number of neighbors of n and e_n is the number of connected pairs between all neighbors of n.

In directed networks the relation becomes;

$$C_n = e_n/(k_n(k_n\text{-}1)) \tag{1.2}$$

In both cases, the clustering coefficient is a ratio:

$$N / M \tag{1.3}$$

where N is the number of edges between the neighbors of n and M is the maximum number of edges that could possibly exist between the neighbors of n.

The network clustering coefficient is the average of the clustering coefficients for all nodes in the network and the clustering coefficient of a node is always a number between 0 and 1. Nodes with less than two neighbors are assumed to have a clustering coefficient of 0.

1.2.1.2. Betweenness. The number of shortest paths between two nodes which passes through a particular node is defined as betweenness (b_i). It is the measure of centrality within a network. "(i,m,j)" in the Equation 1.4 is the number of shortest paths between i^{th} and j^{th} nodes that passes through m^{th} node.

$$b_m = \sum_{i \neq j} \Gamma(i,m,j) \tag{1.4}$$

1.2.1.3. Network Diameter. Network diameter is the largest distance between two nodes. If a network is disconnected, its diameter is the maximum of all diameters of its connected components.

1.2.1.4. Network Radius. The Network Radius is the minimum among the non-zero eccentricities (Eccentricity: The maximum non-infinite length of a shortest path between n and another node in the network. If n is an isolated node, the value of this attribute is zero) of the nodes in the network.

1.2.1.5. Network Density. The value of Network Density changes between 0 and 1 and it shows how densely the network is populated with edges (self-loops and duplicated edges are ignored). If the networks do not contain edges and have only isolated nodes, its density is 0. On the other hand; the density of a clique is 1.

1.2.1.6. Characteristic Path Length. It is also known as the average shortest path length, which gives the expected distance between two connected nodes.

1.2.1.7. Shortest Paths. It is the distance between two nodes n and m which is denoted by $L(n,m)$.

1.2.1.8. Average Number of Neighbors. It indicates the average connectivity of a node in the network. A normalized version of average number of neighbors is the network density.

1.2.1.9. Isolated Nodes. The number of isolated nodes gives information about how the network density is distributed. For example; isolated nodes has a density of 0.

1.2.1.10. Network Heterogeneity. It reflects the tendency of a network to contain hub nodes.

1.2.1.11. Multi-Edge Node Pairs. Number of multi-edge node pairs indicates how often neighboring nodes are linked by more than one edge.

1.2.1.12. Network Centralization. It provides information about density of network. The networks of centralization value close to 1 resemble a star; whereas decentralized networks are characterized by having a centralization value close to 0.

1.2.2. Network Reconstruction

Interactions among biological entities would be obtained from publicly available databases such as STRING (Jensen *et al*., 2008) and the Biological General Repository for Interaction Datasets; BioGRID (Stark *et al*., 2011). The STRING database is a web resource dedicated to protein–protein interactions, including both physical and functional interactions (Jensen *et al*., 2008). The interactions include direct (physical) and indirect (functional) associations, which are derived from four sources (Genomic Context, High-throughput Experiments, Coexpression, Previous Knowledge). STRING gets the interaction data from these sources for a large number of organisms, and transfers information between these organisms where applicable. The database currently covers 2,590,259 proteins from 630 organisms. Additionally, the neighbors of a protein can be manually found by entering its name and selecting at which confidence score the interaction of that protein is preferred to be visualized.

Several data visualization tools exist in order to visualize and analyze the properties of a network. Osprey, which was developed by the Sun Micro Systems Java Standard Development Kit and used for the visualization and manipulation of complex interaction networks (Breitkreutz *et al*., 2003), GeneMANIA, which is a web-interface used for the predictions about gene functions, analyzing gene lists and prioritizing genes for functional assays (Morris *et al*., 2010), ProHits, which is developed for the MS-based interaction proteomics and which may convert the MS data files to protein-protein interaction data sets (Breitkreutz *et al*., 2010) and Cytoscape, which is an open source software to analyze the biomolecular interaction networks with high-throughput expression data (Shannon *et al*., 2003) could be mentioned among these frequently utilized tools.

Cytoscape was developed as a collaborative project between the Institute for Systems Biology, University of California San Diego, Memorial Sloan-Ketterin Cancer Center, Institut Pasterabasiur, Agilent Technologies, University of California San Francisco, Unilever, University of Toronto and National Center for Integrative Biomedical Informatics. Cytoscape is the most powerful source when used in conjunction with large databases of protein-protein, protein-DNA, and genetic interactions. Cytoscape's software cores provides basic functionality to layout and investigate the network; to visually integrate the network with expression profiles, phenotypes, and other molecular states; and to link the network to databases of functional annotations. The modification properties of Cytoscape and the plug-ins make it a powerful tool for the network visualization. There are many lay-out options developed in the Cytoscape in order to better comprehend the studied networks. Plug-ins may be downloaded and updated for Cytoscape in order to perform network analysis, to find modules and to attribute biological descriptions to a group of biological entity.

1.3. Functional Modularity in Biological Networks

Modularity is the cellular functionality that can be used to uniquely partition a network into a collection of smaller entities. Each module achieves an identifiable task, separable from the functions of other modules (Barabasi *et al.*, 2002). The molecular modules in a biological are usually densely connected within them but sparsely connected with the rest of the network. The extraction of functional modules from biological interaction networks would be achieved by cluster analysis. Clustering analysis may be broadly defined as the grouping of objects based on their sharing of diverse, measurable properties (Pereira-Leal *et al.*, 2004; Spirin and Mirny, 2003).

Several studies were conducted in order to develop methods for the identification of functional modules. Zhang et al. introduced a novel algorithm to identify overlapping communities in complex networks, which was associated with an approximation mapping of network nodes into Euclidean space and fuzzy c-means clustering. This algorithm was also verified with experimental studies (Zhang et al., 2006). A study related to identifying modules was used to predict functions of uncharactarized proteins. This study utilized the method of Statistical-Algorithmic Method for Bicluster Analysis (SAMBA), which enables the processing of current and future sources of biological information and it could be used in conjunction with experimental techniques and higher organisms (Tanay et al., 2003).

Functional modularity of model organisms including bacteria and earthworm as well as the baker’s yeast was analyzed in several studies. In a study, module collections of seven bacterial organisms were analyzed and it was observed that modules related to membrane transport and cell motility were conserved among multiple organisms. The study was based on the spectral clustering of the protein-protein functional association networks and it was proposed that the parameter settings would provide correct solutions (Wu et al., 2011). A study on the Caenorhabditis elegans core neuronal protein–DNA interaction network indicated that this network was organized into two TF modules and that they were related to different functional aspects of the complete network. It was also indicated that neuronal function of paired homeodomains was evolutionarily conserved (Vermeirssen et al., 2006). A comprehensive study was performed on the condition-dependent expression of genes coding for the components of hand curated multi-protein complexes of Saccharomyces cerevisiae and as a result it was suggested that transcriptional modules would represent basic functional and evolutionary building blocks of protein complexes (Simonis et al., 2006). In a study, the yeast PPI network was analyzed including the protein complexes (PIC network) and excluding the complexes (PEC network). It was seen that PIC and PEC networks displayed greater structural modularity that random rewired networks. The results of the study proposed that the structural modules in the PPI network originated as an evolutionary byproduct without biological significance (Wang and Zhang, 2007). In another previous study the yeast protein-protein interaction network was represented as a weighted graph. The betweenness-based partition algorithm was used to identify 266 functional modules in the yeast proteome network. It was shown that the genes in the same functional module conferred a similar phenotype and they could be useful for gene annotation (Chen and Yuan, 2006).

Functional modularity was also used to shed light onto the mechanisms of disease related associations in higher organisms. A previous study indicated that due to the network representations of phenotypes, it was observed that genetic disease phenotypes appeared in similar modular styles. The correlations between the modularity and functional genomics as well as its connection to drug-target associations were also reported (Jiang et al., 2008). The modular approach related to cardiovascular disease, which is a complex disorder, was proposed to provide support for melanoma and cardiovascular disease association (Rende *et al*., 2011).

Several clustering algorithms have been developed and are successfully being used in diverse fields such as mathematics, physics, math-programming, statistics, computer sciences, artificial intelligence and classic disciplines as psychology and business including pattern recognition, compression and classification (Fung G., 2001; Pereira-Leal *et al*., 2004). FAG-EC is a fast agglomerate algorithm that identifies modules in the basis of edge clustering coefficients. The most important property of FAG-EC is to be extremely fast and thus be useful in the analysis of large protein interaction networks (Li *et al*., 2008). LeMoNe is a software package for identifying the modules inside networks and the software is implemented in Java. It utilizes the bottom-up Bayesian Hierarchical Clustering for the construction of regulatory programs and conditional entropy measure to assign regulators to the regulation program nodes. The learning process is faster for larger data sets due to the assignment of the regulators. While the partitioning of genes inside a module is not dependent on the expression profiles of the potential regulators, the regulatory tree structures learned by LeMoNe are dependent on the expression data for the genes in a module (Michoel *et al.*, 2007). ENIGMA is a software tool, which uses combinatorial statistics and graph-based clustering to extract modules from perturbational microarray data. The modules are formed by simultaneous use of GO annotations, protein interactions and transcription factor binding information and the software also suggests regulators, which may affect the expression of some of the genes in the module (Maere *et al.*, 2008). MCODE, another software for the identification of modules in a network, is also available as a plug-in for Cytoscape. The abbreviated name MCODE stands for Molecular Complex Detection and it detects densely connected regions in a large interaction network, which represent molecular complexes. The algorithm of MCODE operates in three stages based on vertex weightings (Bader and Houge, 2003).

In order to identify functional relationships among the module members and between modules through the investigation of functional cross talk, gene ontology assignments are frequently utilized. Ontologies are used when the relationships are described in well-defined terms. The Gene Ontology (GO) project provides dynamic and controlled sets of descriptions for biological entities that can be applied to all of the eukaryotes. There are 3 types of descriptions that were developed for the use of ontologies, which would be associated with biological processes, cellular components and molecular functions. Biological process gene ontology describes the biological objective with the contribution of gene or gene product. Chemical or physical transformation is usually called a process, which means that something goes into a process and something different comes out of it. Molecular function is defined as the biochemical activity of a gene product, which also includes specific binding to ligands or structures. It describes only what is done without specifying where or when the event actually occurs. The active place of a gene product in the cell refers to its cellular component. These terms, when used alone or in conjunction, reflect our understanding of eukaryotic cell structure (Ashburner *et al.*, 2000). Several tools are available for the identification of gene ontology enrichment of a particular group of genes or of modules that are suitable for use in the Cytoscape software environment. PiNGO, which is implemented in Java (Maere *et al.*, 2008), and its sister BiNGO, which is again a Java-based plug-in of Cytoscape are among them. BiNGO helps to detect the set of genes that are statistically overrepresented in a group of genes. It is an open source of bioinformatics software platform for visualizing and integrating molecular interaction networks. BiNGO maps the predominant functional themes of a given gene set on the GO hierarchy and outputs this mapping as a Cytoscape graph. Getting the full advantage of Cytoscape visualization environment is the main advantage of this tool over other GO tools and BiNGO can be used directly or indirectly on molecular interaction graphs (Maere *et al.*, 2005).

1.4. Assigning Functions to Biological Entities

Several computational methodologies were used in recent studies in order to be able to assign probable functions to biological entities such as genes, or modules or a group of genes. The global optimization method (GOM) was used to analyze the yeast Saccharomyces cerevisiae protein-protein interaction network on the basis of their network of physical interactions and assigns functions to unclassified proteins based on their positions in a physical interactions network. The numbers of protein interactions were minimized among different functional categories. The methodology resulted in equivalent multiple functional assignments (Vazquez *et al*., 2003). The modified and faster global optimization as a modified form of the Global Optimization Method was used in the study of Sun *et al.* (2006). This method employs local optimal repetition method to reduce calculation time and also takes into consideration the topological structure information to achieve a more accurate prediction. This method was also confirmed by the experimental studies (Sun *et al*., 2006).

In a very recent study, the protein function prediction was carried out for Mycobacterium tuberculosis (MTB) proteins using the STRING database and Gene Ontology Terms. This study yielded a new functionally characterized MTB strain CDC1551 proteome and was reported to contribute to research into the development of effective anti-tubercular drugs (Mazandu and Mudler, 2011).

A classification technique known as support vector machines was used for protein function prediction previously. A support vector classifier was constructed for in method that obtained values of specificity and sensitivity higher than 90% in prediction of PPIs and also provided a confidence score in interaction prediction of each pair of proteins (Urquiza *et al*., 2011). A three stage graphical method was used previously for protein function prediction. The proposed method was applied to two separate problems, which were the prediction of enzymes and the prediction of DNA-binding proteins (Alvarez and Yan, 2012). Another prediction approach; GEST, which is based on gene expression similarity and taxonomy similarity of the functional classes, was used previously and mentioned as a powerful method for predicting protein function to very specific terms (Zhu *et al*., 2007).

A method for integrating multiple graphs within a framework of semi-supervised learning was proposed previously. The method alternates between minimizing the objective function with respect to network output and with respect to combining weights. We apply the method to the task of protein functional class prediction in yeast. This method was reported to be much useful than a single graph (Shin *et al*., 2009).

The aim of this present thesis is to construct a sub-network of the yeast global interaction network by obtaining a specialized Glucose Metabolic Process network through the application of GO Term Elimination criteria and to investigate this specialized network from a modular point of view in order to identify functional entities that are present within the highly specialized modules and to assign probable functions to the unannotated or uncharacterized genes, or to assign novel additional functions to the proteins with already known functional associations. The methods utilized during this process are presented in the next chapter. The comparative analysis of the newly constructed specialized network and the global interaction network in yeast, the identification and analysis of the modules that are present in the specialized network and their functional analyses are presented in the Results and Discussion chapter. Several suggestions are made regarding the function of various yet uncharacterized or verified yeast proteins in the same chapter. The conclusions, which could be drawn from these analysis as well some recommendations regarding future studies on this topic are presented in Conclusion and Recommendations chapter.

2. METHODS

2.1. Construction of Networks

STRING Database was used for the construction of networks. The interaction data was downloaded from the official web-site of STRING with the version of STRING 8.3. The interaction data with the confidence score 999 was used due to the higher probability of the occurrence for the listed interactions.

Two networks were constructed, one was the Global Network (GN) and the second was the Glucose Metabolic Network (GMN), which was obtained by the application of network reduction to the data set of "Confidence Score 999" by using Amigo Gene Ontology (GO) annotation terms associated with Glucose Metabolic Process in yeast.

The pool of GO terms constituted the major criterion for scaling down the GN and the Selective Permissibility Algorithm (SPA) was inspired for this process (Arga *et al.*, 2006). A scoring criterion was established for the reduction of the network. Each of the process, function and component terms was scored 1 in the functional subset (FS) such that for each gene, the total score would be 3. Every protein in CS999 was also scored.

All proteins possessing the same process, function and component GO annotations as those in FS would get a score of 3 and would thus be spared and the remaining proteins with scores 2 or less were eliminated from the GN (Figure 2.1).

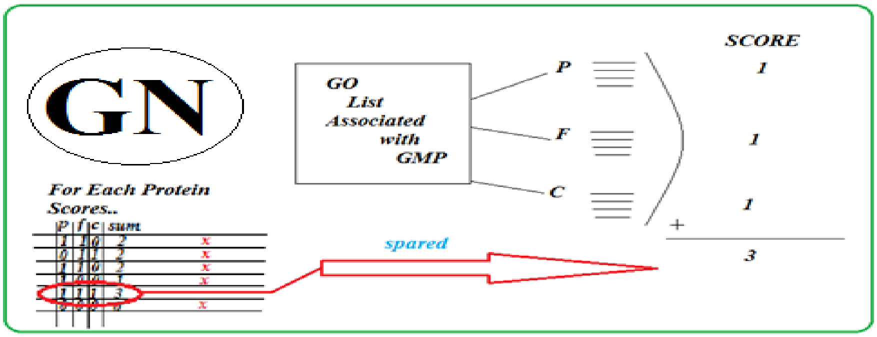


Figure 2.1. Representation of GO filtering.

2.2. Identification of Hubs

Nodes, which are present both in the top 10 percent of the highest degree nodes and the top 10 percent of the highest betweenness centrality nodes, were identified and the intersection of these two sets was determined as hub protein set for each of the networks (Figure 2.2).

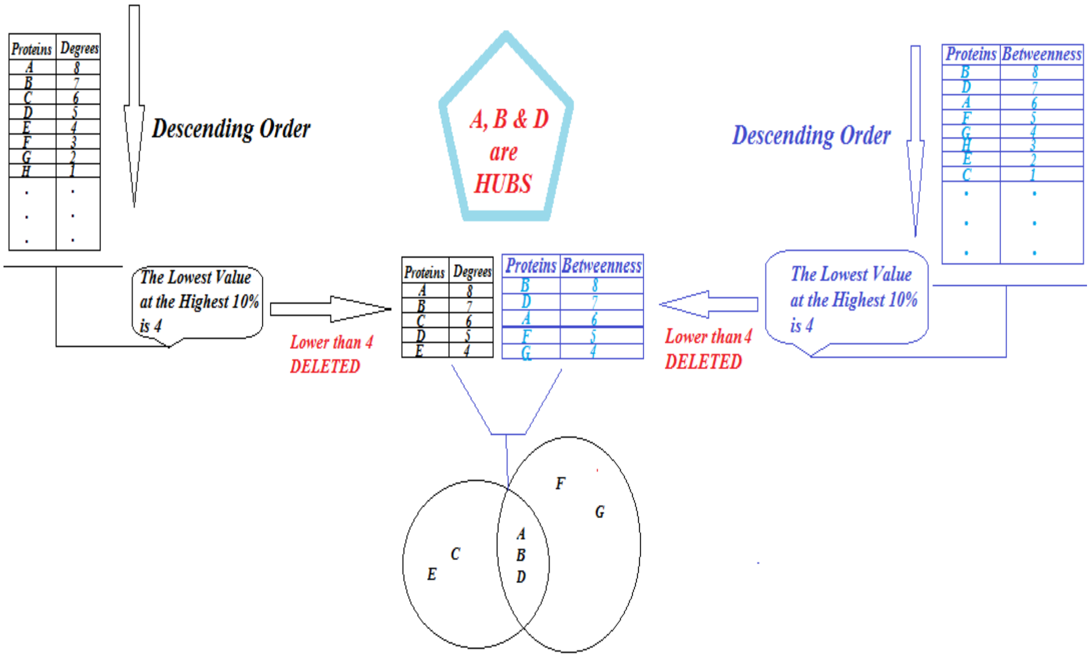


Figure 2.2. Representation of hub identification.

2.3. Analysis of the Networks

In order to analyze and visualize the networks constructed with “Confidence Score 999” (CS999), Cytoscape (Cytoscape, verison 2.7.0), which is a java based software to analyze the biomolecular interaction networks, was used (Shannon *et. al.*, 2003).

Plug-ins were downloaded for Cytoscape for performing the network analyses, finding the modules and providing the functional descriptions of these modules. In order to analyze the network; the network analyzer plug-in was used and in order to determine the modules the MCODE plug-in was utilized and the BINGO plug-in was used for determining the functional description of the identified modules.

2.3.1. The Topological Analyses of the Networks

The topological properties of GN and GMN were analyzed using the Network Analyzer plug-in (version 2.7) of Cytoscape and the number of nodes and edges, the network diameter, network radius, clustering coefficient, characteristic path length, betweenness and closeness, as well as the distributions of the degrees, neighborhood connectivities, average clustering coefficients and the shortest path lengths were determined using this plug-in.

2.3.2. Module Detection

Modules within the Global and the Glucose Metabolic Networks were detected by MCODE, which is another application that is implemented in Cytoscape, using either the fluff on or the fluff off modes. The plug-in named ClusterViz was downloaded from the official web-site of Cytoscape for this purpose.

The “Fluff” option, which expands the core cluster by one neighbor shell, and the “Node Score Threshold”, which sets the acceptable score deviance from the seed cluster node’s score for expanding a cluster, were set in order to detect the modules. The analyses were conducted turning the fluff option on and off in order to identify the core and the extended modules. Several thresholds were used as the node score threshold and it was set to 0.1 in the study.

2.3.3. Determination of the GO Annotations for the Modules

GO annotations were obtained for the modules using the BINGO 2.3 plug-in, which is a Java-based tool to determine which Gene Ontology categories are statistically overrepresented in a set of genes or a sub graph of a biological network. The plug-in was downloaded from the official web-site of Cytoscape for this purpose.

Two significance tests that are used by BINGO for determining the significance of the gene ontology representation are the hypergeometric and the binomial tests. The significance of the gene ontology assignments of the modules were provided using the hypergeometric test and it was previously reported that the hypergeometric test provides results faster than when the binominal test is used for determining the p-values (Maere *et al.*, 2005).

3. RESULTS AND DISCUSSION

3.1. Construction of Glucose Metabolic Network

The PPI data were downloaded from the official web-site of STRING database with the version of STRING 8.3. The interactions having the highest confidence score of 999 were taken into account for the present study.

In order to analyze and visualize the PPI network constructed with "Confidence Score 999" (CS999), Java based software Cytoscape was used. The circular layout option in Cytoscape was enabled in order to visualize and remove the disconnected proteins (Figure 3.1). Then, the largest connected graph of CS999 was obtained to be used for further investigations. This network which is constituted of 1792 nodes and 6919 edges was named as Global Network (GN).

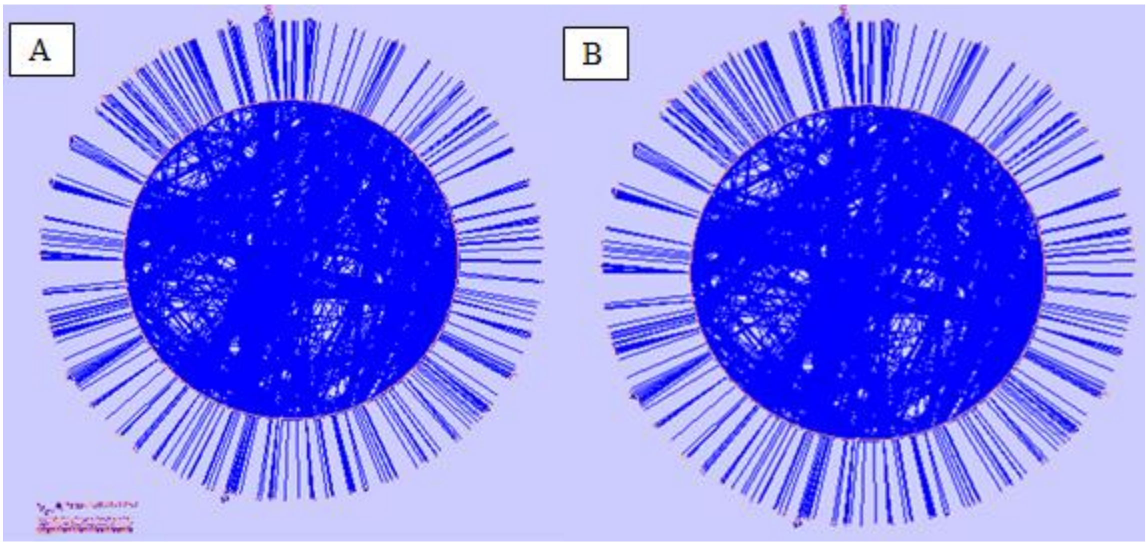


Figure 3.1. A. CS999 network and B. GN network.

The next step of the study involved a scaling down of GN through the use of Gene Ontology terms.

First, 108 genes which have annotations to “Glucose Metabolic Process” GO biological process term were identified. Their annotations to child terms of “Glucose Metabolic Process” including “Negative and Positive Regulation of Gluconeogenesis”, “Glucose 1-phosphate Metabolic Process”, “Glucose 6-phosphate Metabolic Process”, “Udp-glucose Metabolic Process”, “Ethanol Biosynthetic Process During Fermentation”, “Glucose Catabolic Process To Ethanol”, “Pentose-phosphate Shunt”, “Regulation of Glycolysis” and “Glycogen Metabolic-Biosynthetic-Catabolic Processes” was tabulated in Table 3.1. The descriptions of 108 proteins and the GO-IDs are provided in APPENDIX-A. This set of genes was named as functional subset of interest (FS).

Table 3.1. GO terms related to glucose metabolic process and number of genes associated to these terms.

Major GO	Sub-GO	Gene Number
Glucose Metabolic Process	Glycolysis	31
	Gluconeogenesis	19
	Glycogen Metabolic Process	18
	Glycogen Biosynthetic Process	12
	Negative Regulation of Gluconeogenesis	9
	Pentose Phosphate Shunt	7
	Glucose Catabolic Process to Ethanol	5
	Pentose Phosphate Shunt Oxidative Branch	5
	Positive Regulation of Gluconeogenesis	5
	UDP-Glucose Metabolic Process	3
	Glycogen Catabolic Process	3
	Regulation of Glycogen Biosynthetic Process	3
	Negative Regulation of Glycogen Biosynthetic Process	3
	Regulation of Glycolysis	3
	Positive Regulation of Glycolysis	3
	Ethanol Biosynthetic Process During Fermentation	2
	Glucose 1-phosphate Metabolic Process	2
	Glucose 6-phosphate Metabolic Process	2
	Regulation of Glycogen Catabolic Process	2
	Pentose Phosphate Shunt Non-oxidative Branch	1

Second, process, function and component Gene Ontology terms of the genes of FS were identified by using AmiGO to form three pools of reference GO terms (APPENDIX-A). Then a scoring criterion was established for the reduction of GN in order to ensure that each node of the reduced subnetwork have annotations to at least one of the terms of each pool.

The interactions among the spared proteins were visualized again by Cytoscape. The disconnected nodes were removed and the largest connected network constituted of 864 nodes and 2495 edges was named as Glucose Metabolic Network (GMN) (Figure 3.2).

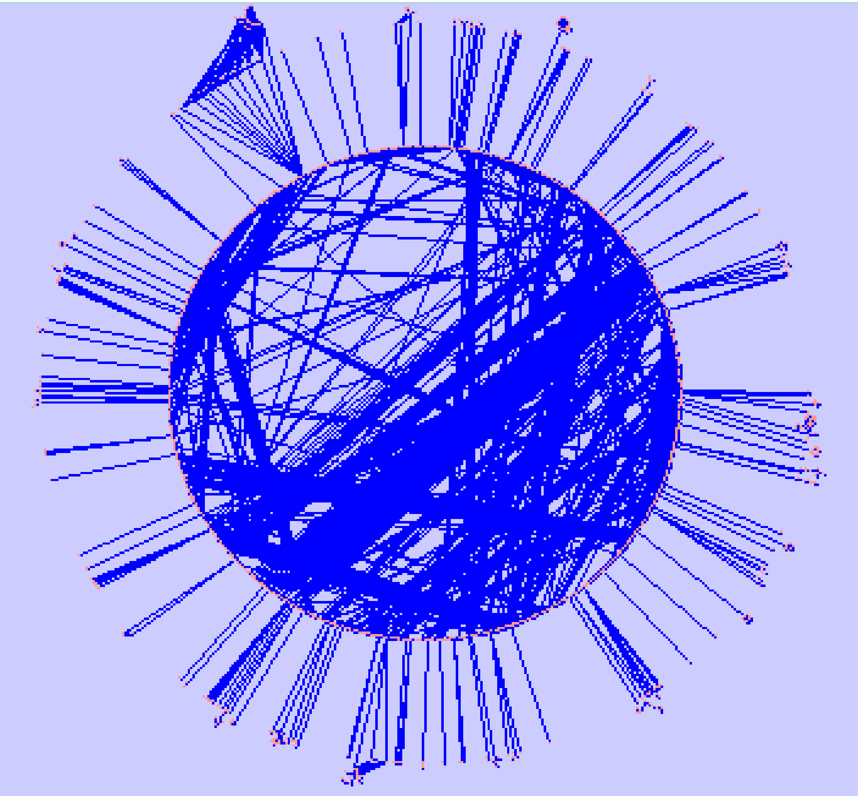


Figure 3.2. Reduced network (GMN) of GN by GO filtering.

The topological properties of GN and GMN were analyzed using Network Analyzer plug-in (version 2.7) of Cytoscape and tabulated in Table 3.2. While the number of nodes was reduced from 1792 to 864, that of edges was reduced down to 2495 from 6919. Network diameters of two networks were nearly the same indicating that the GO term filtering sparsified the network in a random manner, not causing a drastic change in the diameter of the network. The network radius was 11 for both networks also due to the same reason. The network density increased from 0.004 to 0.007 while the network heterogeneity was reduced from 1.074 to 0.9 which indicated that one or more hub proteins were removed as a result of GO filtering. These results indicated that the network became denser and more homogenous while maintaining the diameter.

Table 3.2. Comparison of the network properties of GN and GMN.

	GN	**GMN**
Number of Nodes	1792	864
Number of Edges	6919	2495
Clustering Coefficient	0.494	0.482
Network Diameter	21	20
Network Radius	11	11
Network Density	0.004	0.007
Characteristic Path Length	7.388	8.089
Shortest Paths	3209472 (100%)	745632 (100%)
Average Number of Neighbors	7.722	5.775
Isolated Nodes	0	0
Network Heterogeneity	1.074	0.900
Number of Self Loops	0	0
Multi-edge Node Pairs	0	0
Network Centralization	0.029	0.032

The GN and GMN networks were identified as scale-free which nearly follow a power law distribution and were heterogeneous with a few highly connected nodes (hubs) that are dominant and the rest of the less connected nodes were linked to the network through these hubs.

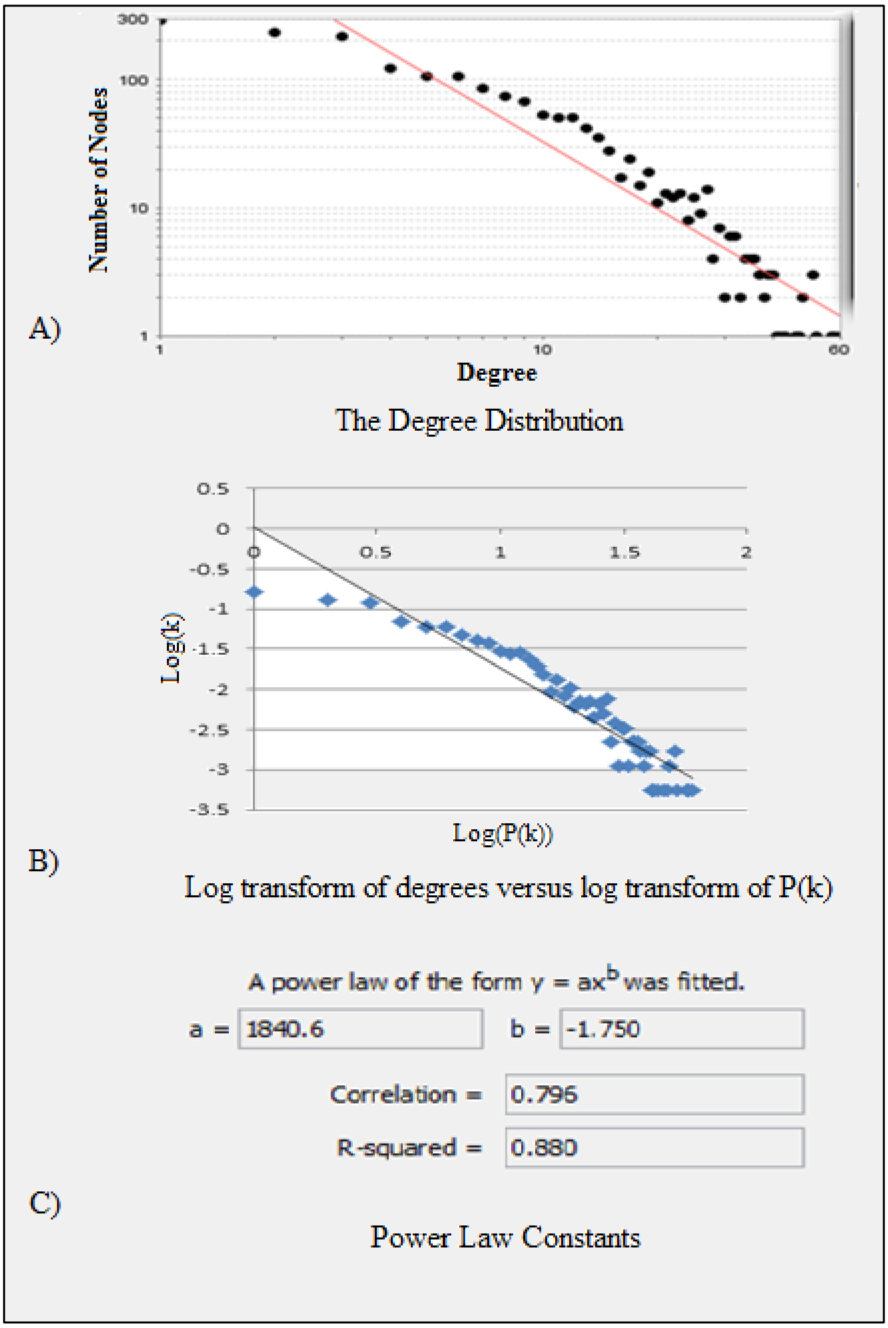
The degree distribution of GN indicated that the constructed network follows nearly a power law distribution with γ=1.75 and R^2=0.91values (Figure 3.3).

The degree distribution of GMN also follows nearly a power law distribution with γ=1.65 and R^2=0.82. The node-degree distributions indicate that only few proteins have high degrees while the rest do not (Figure 3.4)

3.2. Identification of Hubs

Hubs were identified in order to find the proteins which have important roles in the network and have high connectivity compared to the other nodes. It has been shown that hubs play biologically important roles in cells (Vallabhajosyula *et al.*, 2009).

The degree specifies the number of connections that a node has and betweenness centrality reflects the amount of control that a node exerts over the interactions of other nodes in the network (Yoon *et al*., 2006). Thus, nodes which are present both in the top 10 percent of the highest degree nodes and the top 10 percent of the highest betweenness centrality nodes were identified as hubs. 42 and 26 hubs were identified for GN and GMN, respectively. 18 out of these hubs were common for the both of the networks (Table 3.3).



The Degree Distribution

Log transform of degrees versus log transform of P(k)

A power law of the form $y = ax^b$ was fitted.

a = 1840.6

b = -1.750

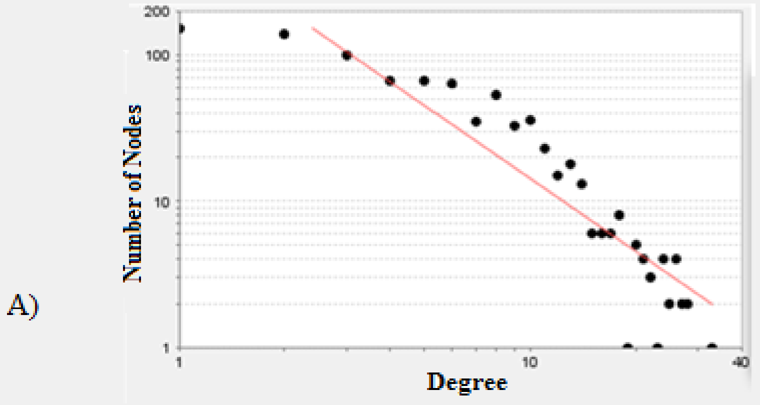
Correlation = 0.796

R-squared = 0.880

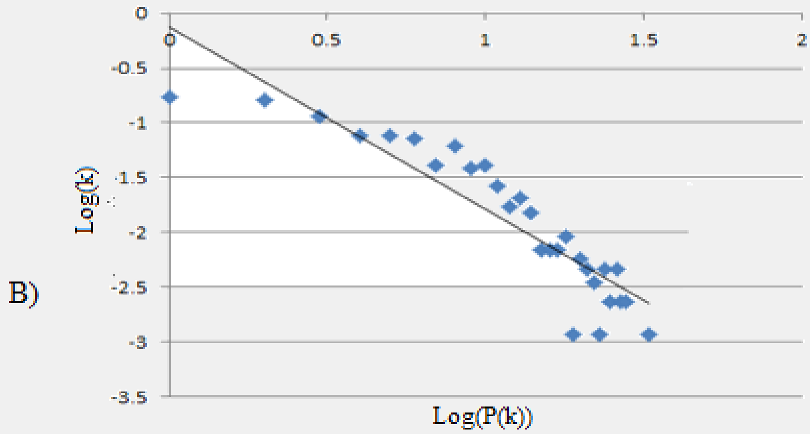
Power Law Constants

Figure 3.3. Topological analysis of GN.

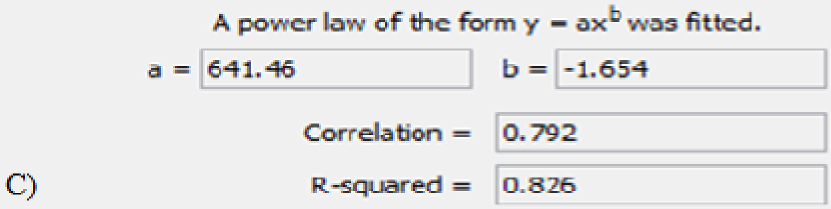
The descriptions of all hubs are obtained from the official web-site of SGD and shown in APPENDIX-A for both networks.



The Degree Distribution



Log transform of degrees versus log transform of P(k)



A power law of the form $y = ax^b$ was fitted.

a = 641.46 b = -1.654

Correlation = 0.792

R-squared = 0.826

Power Law Constants

Figure 3.4. Topological analysis of GMN.

The degrees of the hubs were found to be decreased at GMN due to the network reduction. On the other hand an increase in the betweennes centrality of hubs other than Act1p was observed in GMN, indicating that the number of shortest paths among hub proteins increased and their control over the network was enhanced by the reduction of GN to GMN (Yoon *et al*., 2006). Act1p is a structural protein involved in cell polarization, endocytosis and other cytoskeletal functions. That might be the reason why it lost its centrality whereas all other common hubs improved their betweenness centrality (Table 3.3). Its first neighbors were mostly eliminated.

Table 3.3. Betweenness centrality and degree values of common hubs.

Protein	**Degree at GMN**	**Degree at GN**	**Betweenness at GMN**	**Betweenness at GN**
Brr2p	33	51	0.0330672	0.02441875
Rpt6p	28	39	0.06340406	0.02953092
Rpt3p	28	38	0.04425468	0.01349897
Rpo21p	27	40	0.16047889	0.10374282
Pre2p	25	35	0.06220004	0.03481828
Cdc28p	24	28	0.09953085	0.03811149
Act1p	22	42	0.05689529	0.14394221
Spt15p	20	36	0.11897851	0.04979029
Pre8p	20	27	0.05096359	0.02550309
Gcn5p	18	26	0.13966844	0.03262445
Orc2p	18	19	0.1344025	0.02057212
Kap95p	18	19	0.09750838	0.05789308
Glc7p	18	27	0.0954692	0.02549189
Pol30p	18	23	0.03586795	0.02815044
Sgs1p	17	25	0.08944007	0.03163846
Tra1p	16	23	0.06466947	0.01286875
Isw1p	14	19	0.03318369	0.01607012
Skp1p	13	19	0.06213871	0.05060455

Top five hubs of the GN were identified as, Rps3p; a component of the small (40S) ribosomal subunit; Brr2; an RNA-dependent ATPase RNA helicase (DEIH box); Rpl3p; involved in the replication and maintenance of killer double stranded RNA virus; Rpl10p; regulates translation initiation and Act1p; a structural protein involved in cell polarization, endocytosis, and other cytoskeletal functions. 42 hub proteins of GN were found to be significantly associated with 5 GO process terms (Table 3.4). The most significant enrichment was to the term Macromolecule Metabolic Process term.

Table 3.4. GO term annotations of the group of hub proteins of GN.

Gene Ontology Term	**Cluster Frequency**	**P-value**
Macromolecule Metabolic Process	41 out of 42 genes, 97.6%	1.20E-13
Cellular Macromolecule Metabolic Process	40 out of 42 genes, 95.2%	2.02E-12
Cellular Component Organization or Biogenesis at Cellular Level	35 out of 42 genes, 83.3%	9.23E-12
Cellular Component Biogenesis	26 out of 42 genes, 61.9%	4.48E-11
Cellular Component Organization or Biogenesis	35 out of 42 genes, 83.3%	6.24E-11

The top 5 hubs of GMN were Brr2p (RNA-dependent ATPase), Rpt6p (involved in the degradation of ubiquitinated substrates), Rpt3p (involved in the degradation of ubiquitinated substrates), Rpo21p (RNA polymerase II largest subunit B220) and Pre2p (responsible for the chymotryptic activity of the proteasome).

26 hub proteins of GMN were also analyzed for their GO enrichment (Table 3.5). The most significantly associated GO term (p-value 1.38E-09) was the Cellular Component Organization at cellular level and 22 of 26 hub proteins were annotated to this term.

Table 3.5. GO term annotations of the group of hub proteins of GMN.

Gene Ontology Term	**Cluster Frequency**	**P-value**
Cellular Component Organization at Cellular Level	22 out of 26 genes, 84.6%	1.38E-09
Chromosome Organization	14 out of 26 genes, 53.8%	3.53E-09
Cellular Macromolecule Metabolic Process	26 out of 26 genes, 100.0%	4.07E-09
Macromolecule Metabolic Process	26 out of 26 genes, 100.0%	6.40E-09
Cellular Component Organization or Biogenesis at Cellular Level	23 out of 26 genes, 88.5%	2.17E-08

The hubs present in the GN but not in the GMN due to GO filtering were; Taf14p; involved in RNA polymerase II transcription initiation and in chromatin modification, Sic1p; controls G1/S phase transition; Hhf1p; a core histone protein required for chromatin assembly and chromosome function, Rpn10p; a Non-ATPase base subunit of the 19S regulatory particle (RP) of the 26S proteasome, Rpl3p; involved in the replication and maintenance of killer double stranded RNA virus; Rad23p; a Protein with ubiquitin-like N terminus and subunit of Nuclear Excision Repair Factor 2 (Nef2p) with Rad4p, Cdc31p; a Calcium-binding component of the spindle pole body (SPB) half-bridge, Rpl10p; responsible for joining the 40S and 60S subunits and regulates translation initiation, Rps3p; a component of the small (40S) ribosomal subunit and has apurinic/apyrimidinic (AP) endonuclease activity also essential for viability, Enp1p; required for pre-rRNA processing and 40S ribosomal subunit synthesis; Sec22p; cycles between the ER and Golgi complex and Rad52p; involved in the repair of double-strand breaks in DNA during vegetative growth and meiosis (Saccharomyces Genome Database).

Hubs of GN which are not common with GMN but survived from GO filtering were tabulated in Table 3.6. The drastic fall in degree or betweenness centrality is the reason for not to be identified as hub in GMN. The degree or betweenness values fall to the hubs of GMN with the lowest degree and betweenness centrality were marked with bold font (Table 3.6).

Table 3.6. Hubs eliminated due to drastic fall in degree or betweenness centrality. Bold values are below the lowest degree or the lowest betweenness centrality values of GMN hubs.

Proteins	Degree	Betweenness Centrality
Sto1p	**10**	0.09818303
Utp22p	**12**	0.09660349
Ada2p	19	**0.02120251**
Nop7p	13	**0.00401984**
Egd1p	**2**	0.00923773
Pre4p	24	**0.01468526**
Nop1p	**12**	0.03710381
Smd2p	**11**	0.05828899
Nas6p	14	**0.00000342**
Rvb2p	14	**0.00740842**

Most of the hubs specific to GMN were found to be a complex or a part of a complex related to ATPases which are energy suppliers or regulators of cells (APPENDIX-A) in accordance with the assumption that GMN was supposed to be a reduced sub-network of GN related to Glucose Metabolic Process.

3.3. Identification of Modules

Modular structures of GN and GMN were investigated by using Molecular Complex Detection (MCODE) (through ClusterViz v.1.2) plug-in of Cytoscape.

The modules were identified by using different settings of MCODE (Table 3.7). While the other parameters were kept constant, node score thresholds (0.2 or 0.1) and fluff mode (with fluff or without fluff) were manipulated. The number of modules in GN and GMN identified by using Mode 1 and 2 with fluff off and on options were presented in Table 3.8. The number of the modules of GN was found to be 117 and 119 by Mode1 and Mode2, respectively. Manipulating fluff settings does not change identified number of modules. The number of modules identified for GMN was found to be 65 for all settings (Table 3.8).

Table 3.7. Module detection settings.

Option	**Mode1**	**Mode2**
Include Loop	False	False
Degree Threshold	2	2
Haircut	True	True
Fluff	On/Off	On/Off
Node Score Threshold	0.2	0.1
K-Core Threshold	2	2
Max Depth	100	100

Table 3.8. Number of modules with changing options.

Network	**Option**	**Fluff**	**Number of Modules**
GN	Mode 1	Off	117
GN	Mode 1	On	117
GN	Mode 2	Off	119
GN	Mode 2	On	119
GMN	Mode 1	Off	65
GMN	Mode 1	On	65
GMN	Mode 2	Off	65
GMN	Mode 2	On	65

Numbers of the nodes identified by the application of different modes/settings in each module were compared to determine the effect of these settings on the module identification. The largest module was found to be consisted of 55 nodes in GN with mode1/fluff off. However the size of this module increased to 118 nodes when the fluff option applied. If the option changed to Mode2, the numbers of nodes in the largest module were reduced to 26 from 55 without fluff. When the fluff option applied to Mode2, the number of nodes in the largest module increased from 26 to 68 (Table 3.9). The sizes of the modules in Mode2 are smaller than that of the modules in Mode1. Moreover the largest module identified in GN by Mode1 was found to be split into two distinct modules by Mode2 (Table 3.9).

Table 3.9. Modules identified in GN. The modules in bold represent the same module.

Nodes within Modules	**Option**	**Fluff**	**Number of Module**	**Nodes within Modules**	**Option**	**Fluff**	**Number of Module**
55	**Mode 1**	**Off**	**1**	**26**	**Mode 2**	**Off**	**1**
38	Mode 1	Off	1	20	Mode 2	Off	2
33	Mode 1	Off	1	19	Mode 2	Off	1
32	Mode 1	Off	1	18	Mode 2	Off	3

Table 3.9. Modules identified in GN. The modules in bold represent the same module (cont.).

29	Mode 1	Off	2	17	Mode 2	Off	1
17	Mode 1	Off	1	16	Mode 2	Off	1
16	Mode 1	Off	1	14	Mode 2	Off	2
14	Mode 1	Off	4	13	Mode 2	Off	3
13	Mode 1	Off	4	12	Mode 2	Off	2
12	Mode 1	Off	3	11	Mode 2	Off	3
11	Mode 1	Off	2	10	Mode 2	Off	2
10	Mode 1	Off	5	9	Mode 2	Off	1
9	Mode 1	Off	4	8	Mode 2	Off	9
8	Mode 1	Off	7	7	Mode 2	Off	7
7	Mode 1	Off	4	6	Mode 2	Off	9
6	Mode 1	Off	10	5	Mode 2	Off	10
5	Mode 1	Off	8	4	Mode 2	Off	24
4	Mode 1	Off	19	3	Mode 2	Off	38
3	Mode 1	Off	39				
118	**Mode 1**	**On**	**1**	82	Mode 2	On	1
97	Mode 1	On	1	**68**	**Mode 2**	**On**	**2**
84	Mode 1	On	1	58	Mode 2	On	1
70	Mode 1	On	2	54	Mode 2	On	1
59	Mode 1	On	1	48	Mode 2	On	1
55	Mode 1	On	1	44	Mode 2	On	1
52	Mode 1	On	1	42	Mode 2	On	1
44	Mode 1	On	2	41	Mode 2	On	2
43	Mode 1	On	1	39	Mode 2	On	2
40	Mode 1	On	1	38	Mode 2	On	1
38	Mode 1	On	1	37	Mode 2	On	1
36	Mode 1	On	1	35	Mode 2	On	4
35	Mode 1	On	3	33	Mode 2	On	2
34	Mode 1	On	1	31	Mode 2	On	2
31	Mode 1	On	1	27	Mode 2	On	3
30	Mode 1	On	1	26	Mode 2	On	1
29	Mode 1	On	2	23	Mode 2	On	1

Table 3.9. Modules identified in GN. The modules in bold represent the same module (cont.).

27	Mode 1	On	2	22	Mode 2	On	3
26	Mode 1	On	2	20	Mode 2	On	3
24	Mode 1	On	1	18	Mode 2	On	2
23	Mode 1	On	2	17	Mode 2	On	2
22	Mode 1	On	3	16	Mode 2	On	4
20	Mode 1	On	1	15	Mode 2	On	5
19	Mode 1	On	1	14	Mode 2	On	2
18	Mode 1	On	1	13	Mode 2	On	3
17	Mode 1	On	1	12	Mode 2	On	8
16	Mode 1	On	2	11	Mode 2	On	5
15	Mode 1	On	4	10	Mode 2	On	3
14	Mode 1	On	2	9	Mode 2	On	3
13	Mode 1	On	4	8	Mode 2	On	10
12	Mode 1	On	6	7	Mode 2	On	4
11	Mode 1	On	3	6	Mode 2	On	4
10	Mode 1	On	6	5	Mode 2	On	10
9	Mode 1	On	4	4	Mode 2	On	13
8	Mode 1	On	8	3	Mode 2	On	8
7	Mode 1	On	4				
6	Mode 1	On	4				
5	Mode 1	On	9				
4	Mode 1	On	16				
3	Mode 1	On	9				

The largest modules were identified to be related with Transcription and Ubiquitin-dependent Protein Catabolic Process with the GO-IDs of 6350 and 6511 respectively (Table 3.10).

Table 3.10. The largest module identified in GN.

Number of Nodes	**Settings**	**GO Term**	**p-Value**
55	Mode1- Fluff off	Transcription	2.3638E-53
118	Mode1- Fluff on	Transcription	2.4765E-95
26	Mode2 - Fluff off	Ubiquitin-dependent Protein Catabolic Process	7.6476E-41
68	Mode2 - Fluff on	Ubiquitin-dependent Protein Catabolic Process	9.3908E-51

The modular structure of the GMN was also investigated by using similar settings (Table 3.11). The largest module was identified to be consisted of 23 nodes with mode1/fluff off. When the fluff option applied, the size increased to 50 nodes. If the option changed to Mode2, the numbers of nodes in the largest module were reduced to 19 from 23 without fluff. When the fluff option applied to Mode2, the number of nodes within the largest module increased from 23 to 43.

Table 3.11. Modules identified in GMN using different modes.

Nodes within Modules	**Option**	**Fluff**	**Number of Module**	**Nodes within Modules**	**Option**	**Fluff**	**Number of Module**
23	**Mode 1**	**Off**	**1**	**19**	**Mode 2**	**Off**	**1**
20	Mode 1	Off	1	16	Mode 2	Off	1
17	Mode 1	Off	1	13	Mode 2	Off	1
14	Mode 1	Off	1	11	Mode 2	Off	2
13	Mode 1	Off	1	10	Mode 2	Off	1
11	Mode 1	Off	4	9	Mode 2	Off	1
10	Mode 1	Off	1	8	Mode 2	Off	2
9	Mode 1	Off	3	7	Mode 2	Off	5
8	Mode 1	Off	5	6	Mode 2	Off	11
7	Mode 1	Off	5	5	Mode 2	Off	8

Table 3.11. Modules identified in GMN using different modes (cont.).

5	Mode 1	Off	14	4	Mode 2	Off	16
4	Mode 1	Off	14	3	Mode 2	Off	16
3	Mode 1	Off	14				
50	**Mode 1**	**On**	**1**	**43**	**Mode 2**	**On**	**1**
43	Mode 1	On	1	42	Mode 2	On	1
39	Mode 1	On	2	39	Mode 2	On	1
33	Mode 1	On	1	33	Mode 2	On	1
29	Mode 1	On	2	29	Mode 2	On	2
28	Mode 1	On	2	26	Mode 2	On	1
27	Mode 1	On	1	22	Mode 2	On	1
22	Mode 1	On	3	20	Mode 2	On	3
21	Mode 1	On	1	18	Mode 2	On	2
20	Mode 1	On	2	17	Mode 2	On	1
18	Mode 1	On	2	15	Mode 2	On	2
17	Mode 1	On	2	14	Mode 2	On	4
15	Mode 1	On	1	13	Mode 2	On	2
14	Mode 1	On	2	12	Mode 2	On	2
13	Mode 1	On	2	11	Mode 2	On	1
12	Mode 1	On	3	10	Mode 2	On	3
11	Mode 1	On	1	9	Mode 2	On	6
10	Mode 1	On	2	8	Mode 2	On	5
9	Mode 1	On	4	7	Mode 2	On	1
8	Mode 1	On	4	6	Mode 2	On	5
7	Mode 1	On	2	5	Mode 2	On	10
6	Mode 1	On	4	4	Mode 2	On	6
5	Mode 1	On	10	3	Mode 2	On	4
4	Mode 1	On	6				
3	Mode 1	On	4				

The largest modules were identified to be related with “Transcription, DNA dependent” and “Ubiquitin-dependent Protein Catabolic Process” with the GO-IDs of 6350 and 6511 respectively; these modules are present in the Table 3.12 with their p-Values.

Table 3.12. The largest modules identified in GMN by mode1 and mode2.

Number of Nodes	**Settings**	**Description**	**p-Value**
23	Mode1 – Fluff off	Transcription, DNA Dependent	3.67E-31
50	Mode1 – Fluff on	Transcription, DNA Dependent	2.42E-48
19	Mode2 – Fluff off	Ubiquitin-dependent Protein Catabolic Process	7.15E-30
43	Mode2 – Fluff on	Ubiquitin-dependent Protein Catabolic Process	2.99E-39

Since the application of Mode 1 resulted in larger modules which are associated with parental GO terms, further studies on GMN were carried on the modules identified by Mode 2 with or without fluff (APPENDIX-B).

3.4. Analysis of the Selected Modules

In order to determine the functional linkage between the members of the modules identified within the framework of this thesis, eight modules with five nodes and 10 edges detected by Mode2 (fluff off) within GMN (M5.1-M5.8) were selected and further analyzed.

The significantly (p-values<0.01) associated GO-Process terms with each of these modules were first determined to assign a biological process for each module. “Transcription from RNA Polymerase I Promoter”, “pH Reduction”, “Protein Amino Acid Phosphorylation”, “Ras Protein Signal Transduction”, “Chromatin Remodeling”, “Protein Amino Acid Acetylation”, “Nuclear mRNA Splicing via Splicesome” and “Nuclear mRNA Splicing via Splicesome” terms were identified for the Modules M5.1-M5.8 respectively (Table 3.13).

The biological functions of each member within a module were then further analyzed and the supplementary information provided by the enlargement of the module through application of fluff were also assessed and compared with the information obtained without fluff.

Table 3.13. GO-Process terms significantly associated with modules.

Module Name	GO-Process Term	p-Value
M5.1	Transcription from RNA I Polymerase I Promoter	2.70E-10
M5.2	pH Reduction	7.65E-13
M5.3	Protein Amino Acid Phosphorylation	6.24E-9
M5.4	Ras Protein Signal Transduction	4.27E-9
M5.5	Chromatin Remodeling	2.08E-6
M5.6	Protein Amino Acid Acetylation	3.54E-8
M5.7	Nuclear mRNA Splicing via Splicesome	9.88E-10
M5.8	Histone Deacetylation	7.65E-13

3.4.1. Module M5.1

Module M5.1 consisted of 5 proteins: Rpa135p; RNA polymerase I largest subunit A135, Rpa34p; RNA polymerase I subunit A34.5 (Gadal *et al*., 1997), Rpa12p; RNA polymerase I subunit A12.2 (Nogi *et al*., 1993), Rpa49p; RNA polymerase I subunit A49 (Liljelund *et al*., 1992) and Rpa190p; RNA polymerase I largest subunit A190 (Memet *et al*., 1988) (Figure 3.5) and identified to be significantly associated with “Transcription from RNA Polymerase I Promoter” GO Process Term (p-Value: 2.7062E-10).

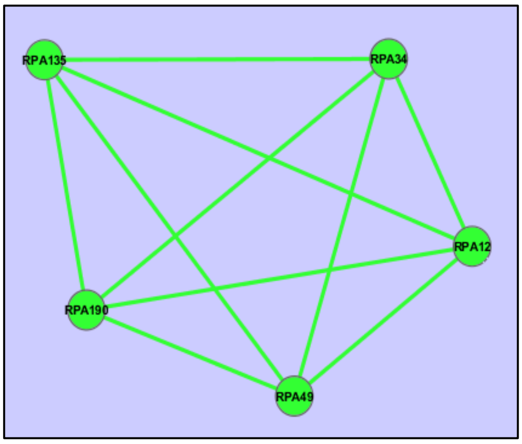


Figure 3.5. Core module M5.1.

Proteins in the core module were reported to be members of the RNA Polymerase I Complex. RNA Polymerase I was identified to have a molecular weight of 590 kDa and consisted of 14 subunits (Table 3.14). RNA Polymerase I specific proteins were determined as Rpa190p (A190), Rpa135p (A135), Rpa12p (A12.2), Rpa14p (A14), Rpa43p (A43), Rpa49p (A49) and Rpa34p (A34.5) in yeast. Rpc40p, Rpc19p were both identical in RNA polymerase I and III; Rpb5p, Rpb6p, Rpb8p, Rpb10p and Rpb12p were identical in all three polymerases (Russell and Zomerdijk, 2006, Kuhn *et al.*, 2007).

Table 3.14. RNA polymerase I subunits (Russell and Zomerdijk, 2006).

Polymerase Part	**Pol I Subunit**	**MW (kDa)**	**Corresponding Pol II Subunit**	**Subunit Type**	**Sequence Identitiy (%)**	**Conserved Pol II Fold (%)**
Core	A190	186.4	Rpb1	homolog	22.3	47.8
	A135	135.7	Rpb2	homolog	26.0	62.1
	AC40	37.7	Rpb3	homolog	21.2	53.5
	AC19	16.2	Rpb11	homolog	17.6	77.5
	A12.2	13.7	Rpb9	homolog	19.2	35.2
	Rpb5 (ABC27)	25.1	Rpb6	common	100	100
	Rpb6 (ABC23)	17.9	Rpb6	common	100	100
	Rpb8 (ABC14.5)	16.5	Rpb8	common	100	100
	Rpb10 (ABC10β)	8.3	Rpb10	common	100	100
	Rpb12 (ABC10α)	7.7	Rpb12	common	100	100
Subcomplex A14/43	A14	14.6	Rpb4	counterpart	4.5	25
	A43	36.2	Rpb7	counterpart	8.0	78.4
Subcomplex A49/34.5	A49	46.7	RAP74	specific	7.6	57.2
	A34.5	26.9	RAP30	specific	8.3	80.5
Total	-	589.6	--	-	29.5	60.8

Two (Rpa43p (A43), Rpa14p (A14)) of seven RNA Polymerase I specific proteins were not part of this module. Rpa43p was eliminated based on the GO Term Elimination Criteria during the construction of GMN and Rpa14p was not included in the module M5.1 but captured with the extension of M5.1 to M5.1E (by fluff option). Rpa14p and Rpa43p form the heterodimer (blue: Rpa43p, red: Rpa14p) (Figure 3.6) and distantly related to Rpb4p, Rpb7p in Polymerase II and Rpc17p, Rpc25p in Polymerase III.

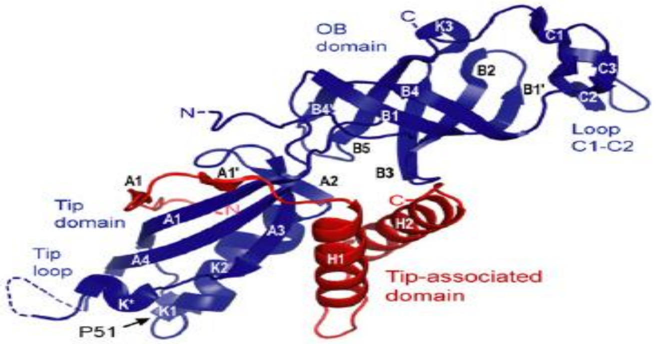


Figure 3.6. RPA14 and RPA43 subunits (Kuhn *et al*., 2007).

The subunits; Rpa14p and Rpa43p are connected to the Polymerase I core as seen in the Figure 3.7 with a stalk-like structure. There is accumulating evidence that Rpa43p forms an essential bridge for the conserved initiation factor Rrn3p and Rpa14p forms two helices that packed on the Rpa43p domain (Kuhn *et al*., 2007). These two proteins are responsible for the initiation procedure and may be considered slightly independent from the RNA Polymerase I Complex. Therefore this special localization of the heterodimer may be responsible for the exclusion of these proteins from M5.1.

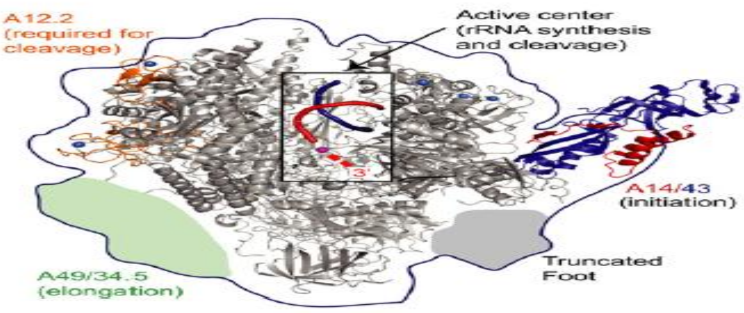


Figure 3.7. RNA polymerase I complex (Kuhn *et al*., 2007).

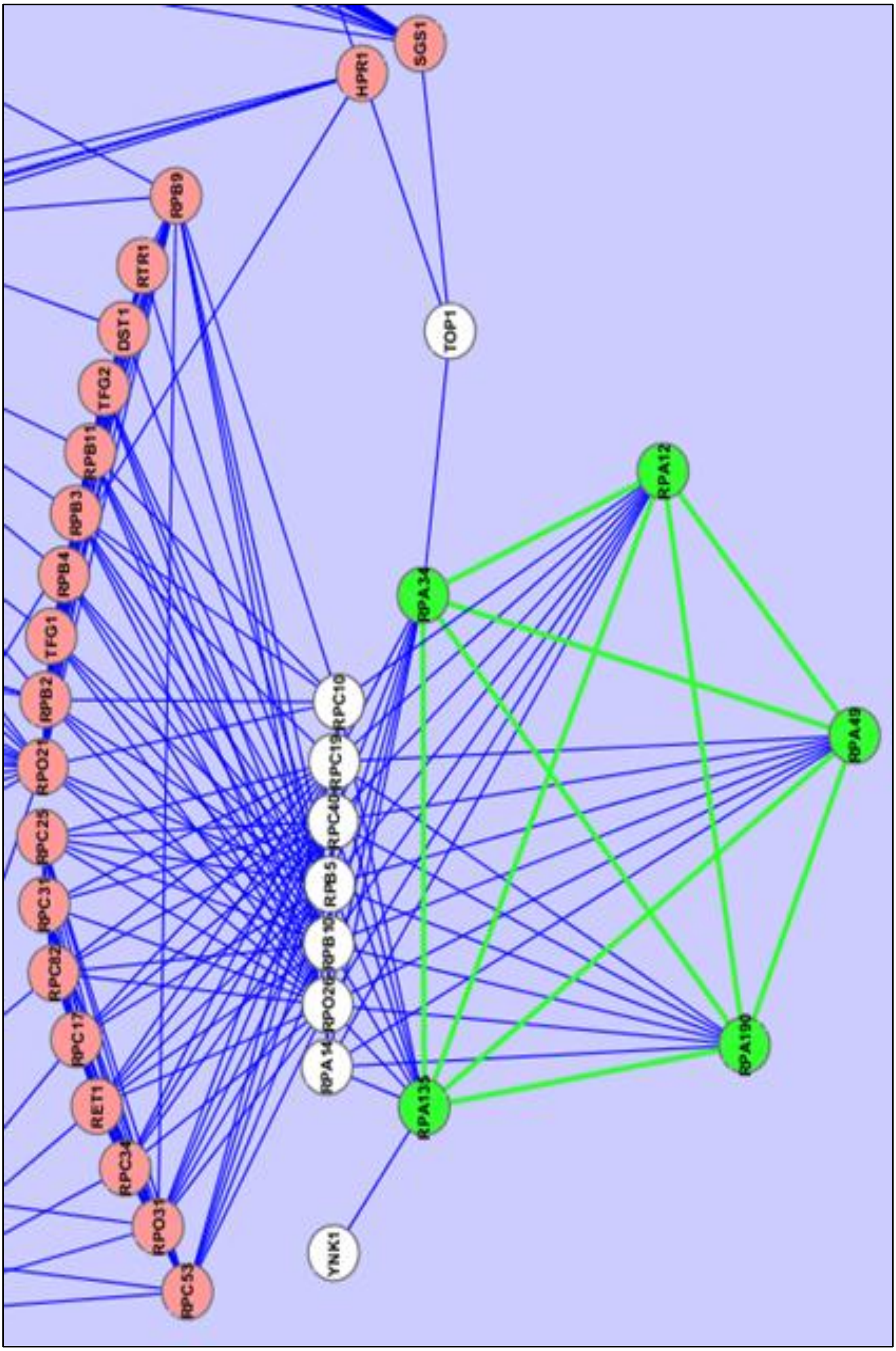


Figure 3.8. Extended module M5.1E and neighbors.

Figure 3.8 represents the extension of M5.1 with the first neighbors (white nodes). This module was extended by the application of fluff and called M5.1E. M5.1E consisted of 14 nodes, 60 edges and identified to be significantly associated with “Transcription from RNA Polymerase I Promoter” GO-Process Term as it was the case for the original module (p-Value: 5.2740 E-22).

All the RNA Polymerase I members except Rpa43p and Rpb8p were captured in M5.1E. Rpa43p and Rpb8p were eliminated due to GO Term Elimination Criteria. Ynk1p and Top1p were also present in M5.1E, which were described in SGD as nucleoside diphosphate kinase (catalyzes the transfer of gamma phosphates from nucleoside triphosphates, usually ATP, to nucleoside diphosphates by a mechanism that involves formation of an autophosphorylated enzyme intermediate) and Topoisomerase I (nuclear enzyme that relieves torsional strain in DNA by cleaving and re-sealing the phosphodiester backbone; relaxes both positively and negatively supercoiled DNA; functions in replication, transcription, and recombination) respectively. A further investigation is required to understand the presence of these proteins in M5.1E.

Rpc53p, Rpo31, Rpc34p, Ret1p, Rpc17p, Rpc82p, Rpc31p, Rpc25p, Rpo21p, Rpb2p, Tfg1p, Rpb4p, Rpb3p, Rpb11p, Tfg2p, Dst1p, Rtr1p, Rpb9p, Hrr1p and Sgs1p (red nodes in Figure 3.8) were captured as the first neighbors of the extended module. All the proteins except Sgs1p; nucleolar DNA helicase of the RecQ family which have interaction with Top1p and potential role as repressor of a subset of rapamycin responsive genes (Kusano *et al*., 1999, Marakhci *et al*., 2011), were subunits of RNA Polymerase II or RNA Polymerase III Complex.

3.4.2. Module M5.2

M5.2 consisted of 5 proteins (Figure 3.9): Vma13p; subunit H of the eight-subunit V_1 peripheral membrane domain of the vacuolar H^+-ATPase (Ho *et al*., 1993), Vma1p; subunit A of the eight-subunit V_1 peripheral membrane domain of the vacuolar H^+-ATPase (Hirata *et al*., 1990), Vma8p; subunit D of the eight-subunit V_1 peripheral membrane domain of the vacuolar H^+-ATPase (Graham *et al*., 1994), Vma7p; subunit F of the eight-subunit V_1 peripheral membrane domain of vacuolar H^+-ATPase (Graham *et al*., 1994) and

Vma10p; subunit G of the eight-subunit V_1 peripheral membrane domain of the vacuolar H^+-ATPase and identified to be significantly associated with “pH Reduction” GO Process Term (p-Value: 7.6580E-13). These proteins were reported to be members of V-ATPase Complex and involved in obtaining the pH balance through organelles (Table 3.11).

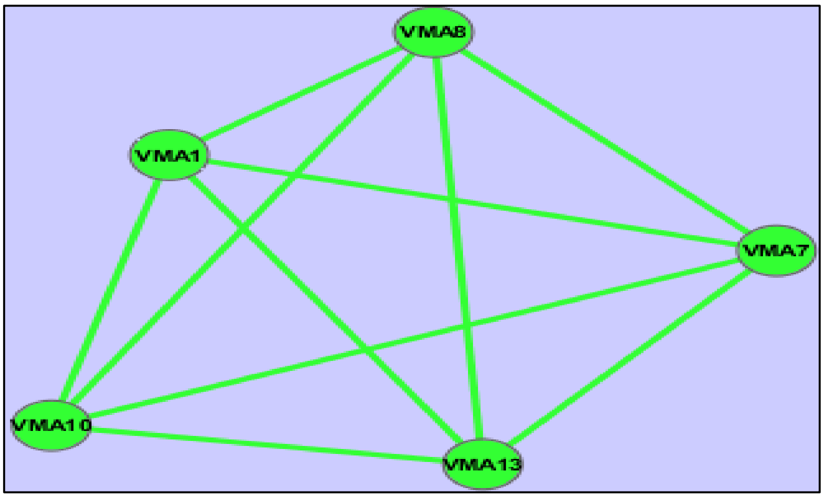


Figure 3.9. Core module M5.2.

V-ATPase Complex is composed of V_1 and V_0 domains which work together as a rotary machine and consisted of 8 and 6 different subunits respectively (Figure 3.10, Table 3.15). These subunits are organized into an ATP-hydrolytic domain (V_1) and a proton-translocation domain (V_0) (Forgac, 2007).

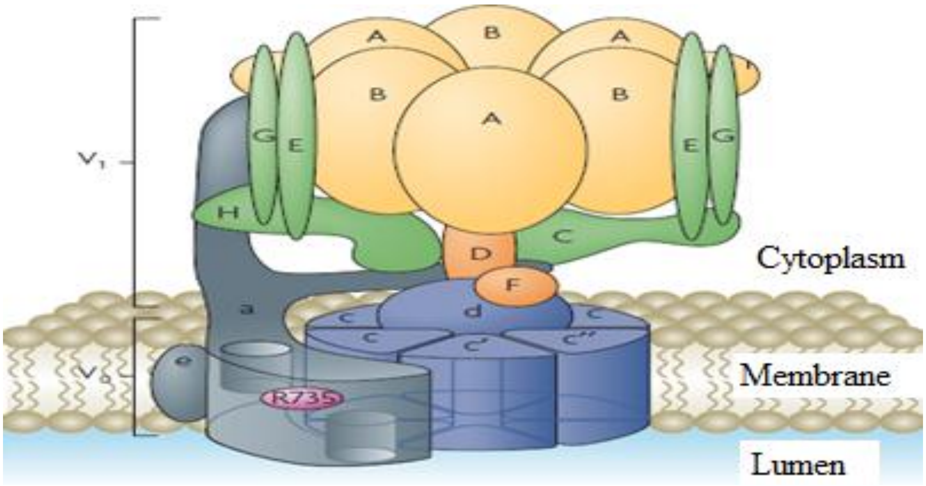


Figure 3.10. Structure of V-ATPase complex (Forgac, 2007).

Table 3.15. Characteristics of V-ATPase subunits (Forgac, 2007).

Subunit	**Molecular Mass (kDa)**	**Yeast Gene**	**Subunit Function**
V_1 Domain			
A	70	Vma1p	ATP hyrolitic site, regulation via non-homologous domain, stator subunit
B	60	Vma2p	Non-catalytic ATP site, binds actin and aldolase, stator subunit
C	40	Vma5p	Regulatory, stator subunits, binds actin
D	34	Vma8p	Rotary Subunit
E	33	Vma4p	Stator Subunit, binds RAVE and aldolase
F	14	Vma7p	Rotary Subunit
G	13	Vma10p	Stator Subunits, binds NEF
H	50	Vma13p	Regulatory, stator subunit, binds NEF
V_0 Domain			
a	100	Vph1p Stv1p	H^+ transport, targeting, binds aldolase, stator subunit
d	38	Vma6p	Coupling, rotary subunit
e	9	Vma9p	unknown
c	17	Vma3p	H^+ transport, rotary subunit
c’	17	Vma11p	H^+ transport, binds Vma21p assembly factor, rotary subunit
c’’	21	Vma16p	H^+ transport, rotary subunit

The proteins captured in M5.2 were members of V_1 domain of V-ATPase, which was responsible for ATP hydrolysis (Forgac, 2007). Three of the V_1 domain proteins (Vma2p, Vma4p and Vma5p) were not included in M5.2.

When fluff mode was applied; the module M5.2 extended to the module M5.2E which consisted of 11 nodes and 43 edges. M5.2E was identified to be significantly associated with "pH Reduction" GO Process Term (p-Value: 3.8833E-28) as it was the case for the original module. Figure 3.11 represents the module M5.2E; the green nodes form the module M5.2, the white nodes in the figure are the first neighbors of this module captured by the application of fluff option and the red nodes are the first neighbors of the extended module.

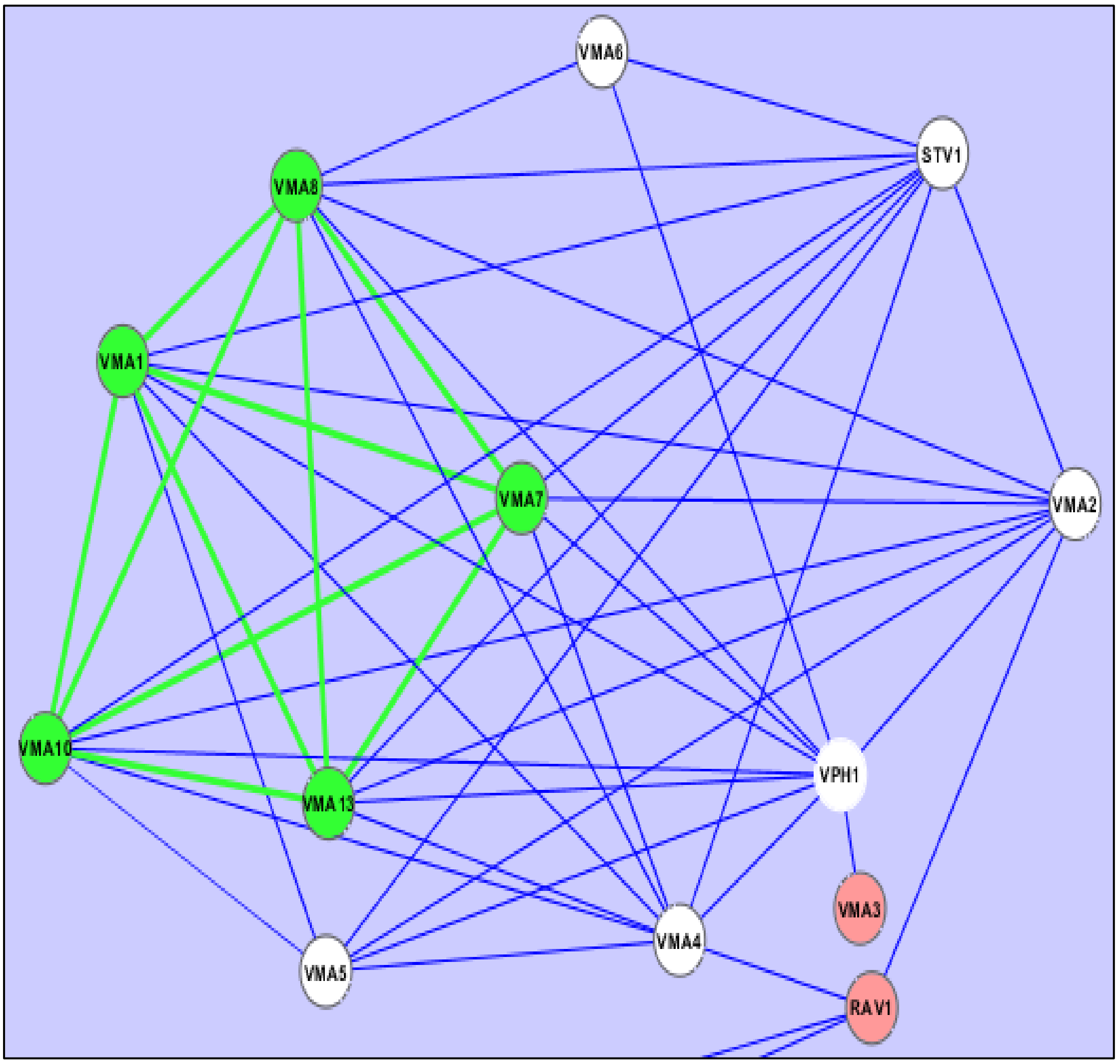


Figure 3.11. Extended module M5.2E and neighbors.

When V_1 part of Figure 3.10 is taken into account, the proteins Vma5p, Vma6p, Vma2p, Stv1p, Vph1p, Vma4p, Vma13p, Vma1p, Vma8p, Vma7p and Vma10p were all obtained by the application of fluff.

Indeed “a” and “d” in the Figure 3.12 were members of V_0 domain but they have direct interaction with some of V_1 domain subunits so it is not surprising to get them with fluff option on. Additionally; there is not much information present for the protein interactions within the membrane till now. Vma9p, Vma11p and Vma19p which are buried inside the membrane could not be captured by this approach (possible due to the lack of information about these membrane imbedded proteins). Only Vma3p, which was captured as a first neighbor of the extended module, was captured among the membrane embedded proteins. Rav1p; a subunit of the RAVE Complex, which promotes the assembly of the V-ATPase holoenzyme was also captured as a neighbor of extended module.

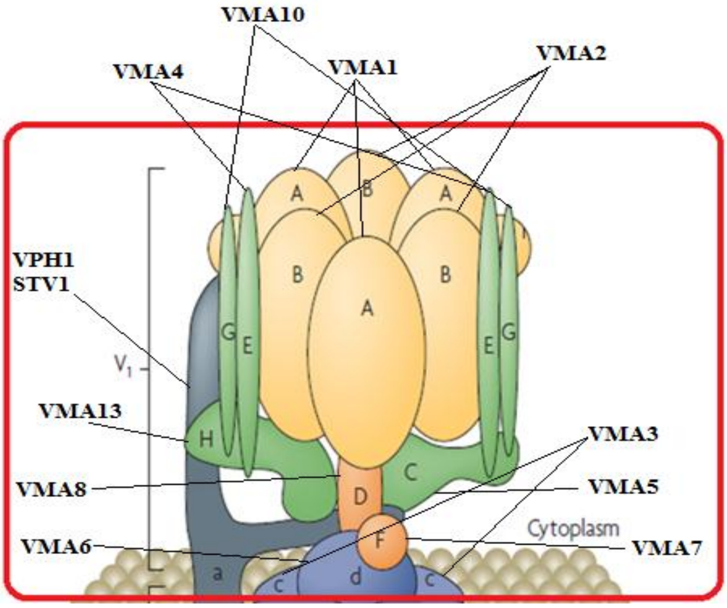


Figure 3.12. V_1 domain of V-ATPase (Forgac, 2007).

3.4.3. Module M5.3

Module M5.3 consisted of 5 proteins (Figure 3.13): Snf1p; AMP-activated serine/threonine protein kinase, Snf4p; activating gamma subunit of the AMP-activated SNF1 Kinase Complex, Gal83p; one of three possible β subunits of the SNF1 Kinase

Complex, Sip1p; another possible β subunits of the SNF1 Kinase Complex and Sak1p; upstream serine/threonine kinase for the SNF1 Complex (Hardie *et al.*, 1998, Wiatrowski *et al.*, 2003, Vincent *et al.*, 2001, Nath *et al.*, 2003). The proteins identified in this module were significantly associated with "Protein Amino Acid Phosphorylation" GO Process Term (pValue: 6.24E-9).

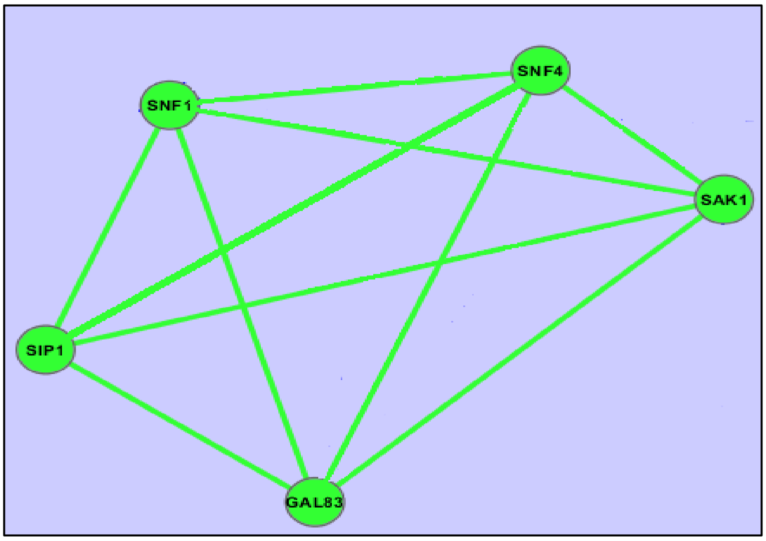


Figure 3.13. Core module M5.3.

Snf1p, Snf4p, Gal83p and Sip1p are components of the SNF1 Complex, which is composed of a catalytic α subunit (Snf1p) and two regulatory subunits γ (Snf4p) and β (either one of Gal83p, Sip1p and Sip2p). This complex plays an important role in response to nutritional and environmental stresses. The SNF1 Complex is inactive when the glucose level is high. However Snf4p binds to the regulatory domain of Snf1p and activates its catalytic function as the glucose is depleted. β subunits of the SNF1 Kinase Complex was reported to be composed of either one of the scaffolding proteins Gal83p, Sip1p or Sip2p and is required for the kinase function and substrate definition (Elbing *et al.*, 2006; Lovas *et al.*, 2003; Carlson *et al.*, 2010).

Two of the three interchangeable members of these subunits; Sip1p and Gal83p were identified in the present module M5.3. It has been shown that the Sip1p β subunit relocalizes from the cytosol to the vacuolar membrane in response to various types of carbon stress (Hedbacker *et al.*, 2004). Sip2p was reported to be functionally associated with the role of Snf1p and Snf4p in the modification of chromatin structure (Ashrafi *et al.*, 2000, Lin *et al.*, 2003).

Sak1p was reported to be the major kinase responsible for the phosphorylation and the activation of Snf1p. Furthermore Sak1p was identified as the sole kinase controlling the nuclear enrichment of Snf1p in which Gal83p constituted β subunit of the SNF1 Complex in response to carbon stress (Liu *et al.*, 2011). The identification of Gal83p, Snf1p and Sak1p simultaneously in the same module supports this finding.

When the module was extended to M5.3E, the newly formed module consisted of 12 proteins with 22 interactions identified among them. The genes encoding these proteins in the extended module were determined to be significantly associated with the "Protein Amino Acid Phosphorylation" GO Process Term (p Value: 2.23E-9) as it was the case for the original module. The figure below graphically represents the extended module M5.3E with green nodes indicating members of the core module, the white nodes indicating the new members of the extended module, the red nodes indicating first neighbor of the extended module that are associated with SNF1 Complex and the cyan node indicating the protein Tom71p with its special localization (Figure 3.14).

Mig1p; the transcription factor involved in glucose repression which is regulated by the Snf1p kinase was included in the extended module (Schüller, 2003). Mig2p; a very close homolog of Mig1p was captured as a first neighbor of extended network rather than directly being associated with the core module. Adr1p; a protein kinase positively regulated by Snf1p through binding in the absence of glucose (Young *et al.*, 2002). Sip4p; a zinc cluster transcriptional activator that binds to the carbon source-responsive element (CSRE) of gluconeogenic genes; involved in the positive regulation of gluconeogenesis is also regulated by Snf1p protein kinase (Lesage et al., 1996). These three proteins identified in the extended network appear in the lower portion of the glucose signaling pathway passing through SNF1 Complex.

Nrg1p; a transcriptional repressor that recruits the Cyc8p-Tup1p Complex to promoters; mediating glucose repression was also reported to interact with the catalytic domain of SNF1 (Park et al., 1999) and this protein was also identified in the extended

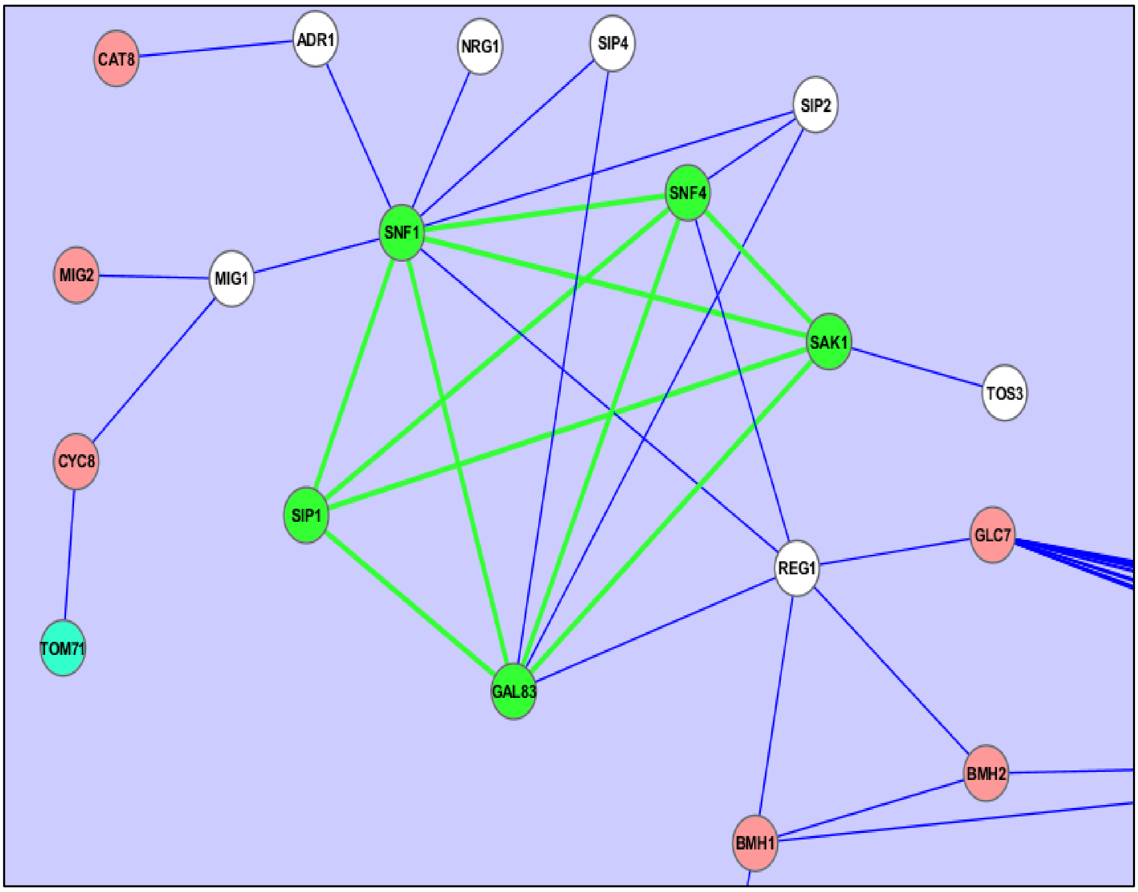


Figure 3.14. Extended module M5.3E and neighbors.

network. Although Cyc8p was also captured as first neighbor of the extended module, Tup1p could not be detected since it was initially removed from the core proteins according the GO elimination criteria.

Sip2p; the third β subunit of the SNF1 Complex was included in the extended module. Thus the extended module could capture all of the physical subunits of the SNF1 Complex. The extended module could identify Tos3p another one of the three Snf1p-activating kinases Sak1p, Tos3p, Elm1p (Liu *et al.*, 2011) The only member, that was not a part of the core nor the extended module Elm1p which was initially eliminated from the glucose network core proteins based on the GO elimination criteria.

Std1p; a protein involved in the control of glucose-regulated gene expression which interacts with kinase Snf1p (Schmidt *et al.*, 1999) was excluded from the network core proteins based on the GO elimination criteria.

Reg1p; regulatory subunit of type 1 protein phosphatase Glc7p, dephosphorylates Snf1p in the presence of glucose (Tu and Carlson, 1995, Dombek *et al*., 2004) and this protein was included in the extended module whereas Glc7p; the other subunit of the complex was included as a first neighbor to the extended module.

Cat8p; a zinc cluster transcriptional activator necessary for derepression of a variety of genes under non-fermentative growth conditions which acts together with Adr1p (Haurie *et al*., 2001, Tachibana *et al*., 2005). Cat8p is a member of the extended network in performing glucose related functions and this protein was identified as a first neighbor to the extended network although Cat8p also regulates the transcription of other genes related to encoding the transcriptional factor Sip4p these two proteins were not shown to be associated in the present network.

Bmh1p and Bmhp2p; 14-3-3 proteins that function as adaptors in signal transduction pathways by binding to phosphorylated proteins to activate, inactivate, or sequester their substrates and have an important role in glucose repression by binding to Reg1 (Dombek *et al*., 2004, Arms *et al*., 2010). Bmh1p and Bmh2p were captured as neighbors of the extended module.

It is interesting to note that Tom71p (cyan in Figure 3.14); a mitochondrial outer membrane protein (Schlossmann *et al*., 1996) with an unknown biological process ontology, was identified as a lone second neighbor to extended network. Although the function of this protein is indirectly related to the SNF1 Complex, its presence as a neighbor may indicate yet undiscovered functional connections to these proteins

3.4.4. Module M5.4

Module M5.4 consisted of 5 proteins (Figure 3.15): Tpk1p, Tpk2p, Tpk3p; cAMP-dependent protein kinase catalytic subunits that promotes vegetative growth in response to nutrients via the Ras-cAMP signaling pathway, Bcy1p; regulatory subunit of the cyclic AMP-dependent protein kinase (PKA) (Santangelo, 2006) and Ras2p; GTP-binding protein that regulates the nitrogen starvation response, sporulation, and filamentous growth

(Bhattacharya *et al*., 1995) and identified to be significantly associated with “Ras Protein Signal Transduction” GO-Process Term (p-Value: 4.2701E-9).

Tpk1p, Tpk2p, Tpk3p, Bcy1p and Ras2p were reported to be in relation with Ras-cAMP (PKA) Signaling that coordinates cell growth and proliferation with nutritional sensing (Xiaojia and Jian, 2010).

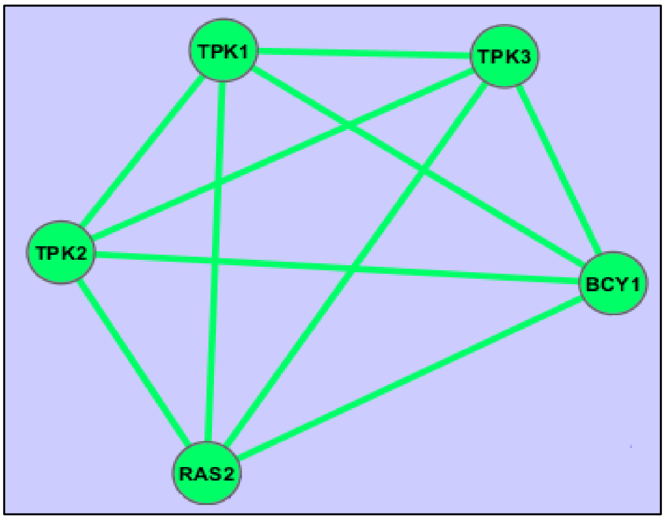


Figure 3.15. Core module M5.4.

PKA which has several functions in the cell, including regulation of glycogen, sugar, and lipid metabolism, is a heterotetramer consisting of catalytic subunits and regulatory subunits (Figure 3.16). The activity of PKA is dependent on cellular levels of cyclic AMP (cAMP) and known as cAMP-dependent Protein Kinase (Santangelo, 2006).

Tpk1p, Tpk2p, and Tpk3p encode isoforms of the catalytic subunit of cAMP-dependent protein kinase (PKA), which is the effector kinase of the Ras-cAMP signaling pathway. Bcy1p encodes the regulatory subunit of PKA and it is also the effector kinase of the Ras-cAMP signaling pathway (Broach, 2002, Santangelo, 2006).

Ras proteins (G proteins) are monomeric GTPases that function as switches. Their activation is depending on GDP and GTP. They are inactive when GDP is bound and active when GTP is bound (Figure 3.16) (Santangelo, 2006). Phosphorylation of Ras2p reported to be dynamic (occurred on roughly 60% of the Ras proteins) and was potentially serving a regulatory role (Whistler and Rine, 1997) thus capturing Ras2p in M5.4 supports this finding.

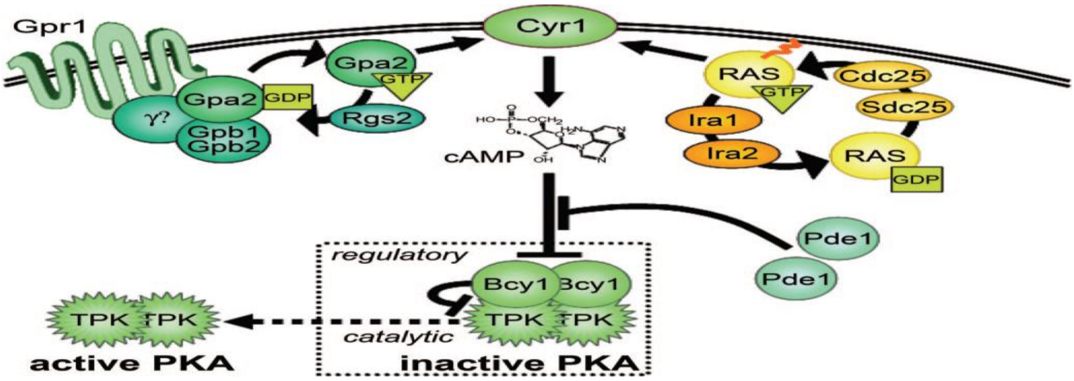


Figure 3.16. Cytoplasmic events in PKA signaling (Santangelo, 2006).

When the module was extended by fluff mode on (Figure 3.17), the newly formed module consisted of 10 proteins with 16 interactions identified among them. Genes encoding these proteins in the extended module were determined to be significantly associated with the "Ras Protein Signal Transduction" GO Process Term (p-Value: 1.91E-17) as it was the case for the original module.

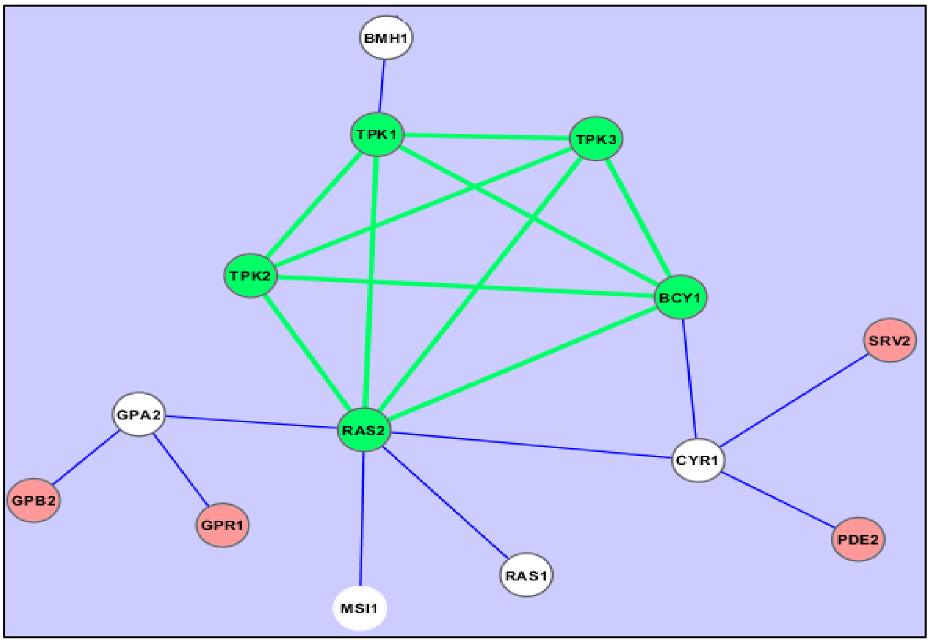


Figure 3.17. Extended module M5.4E and neighbors.

In the Figure 3.17, the green nodes forms the core module, the white nodes are the neighbors of extended module which were captured by fluff mode on and the red nodes are the first neighbors of the extended module.

Gpa2p; nucleotide binding alpha subunit of the heterotrimeric G protein that interacts with the receptor Gpr1p which was the first neighbor of the extended module and proposed to activate adenylate cyclase through Gpa2p (Colombo *et al*., 1998, Peeters *et al*., 2006), Gpb2p; one negative regulator of the Gpa2p signaling pathway that directly interact with the G protein alpha subunit Gpa2p (Niranjan *et al*., 2007) and captured as the neighbor of the extended module (Gpb1p was also stated as the other negative regulator of the Gpa2p signaling pathway but initially removed according to the GO elimination criteria additionally Gpb1p has the interaction data starting from the CF996 score), Msi1p; subunit of chromatin assembly factor I which has role in the suppression of Protein Kinase A Pathway (Pratt *et al*., 2007); Ras1p; GTPase involved in G-protein signaling in the adenylate cyclase activating pathway which plays a role in cell growth (Whistler and Rine, 1997), Cyr1p; adenylate cyclase which is required for cAMP production and cAMP-dependent protein kinase signaling (Kataoka, *et al*., 1985), Pde2p; high-affinity cyclic AMP phosphodiesterase whose product is a key regulator of the cAMP control cascade (Namy *et al*., 2002) and captured as the neighbor of the extended module, Srv2p; CAP (cyclase-associated protein) subunit of Adenylyl Cyclase Complex which is required for RAS-activated adenylate cyclase activity (captured as the first neighbor of extended module) and Bmh1p; 14-3-3 protein which interact with Tpk1p (Chaiken *et al*., 1995) were captured in the extended module.

Cdc25p, Ira1p, Ira2p, Pde1p, Cdc25p and Sdc25p are the other proteins involved in PKA signaling (Figure 3.16). Ira1p, Ira2p, Cdc25, Pde1p were eliminated due to the GO Elimination Criteria and also there was no available data for Sdc25p in STRING database.

3.4.5. Module M5.5

Module M5.5 consisted of 5 proteins (Figure 3.18): Ies1p; subunit of the INO80 Chromatin Remodeling Complex, Ies3p; subunit of the INO80 Chromatin Remodeling Complex, Ies5p; protein that associates with the INO80 Chromatin Remodeling Complex under low-salt conditions (Shen *et al*., 2003), Nhp10p; component of the INO80 Chromatin Remodeling Complex (Ray and Grove, 2009), Ino80p; ATPase and nucleosome spacing factor (Udugama *et al*., 2010) and identified to be significantly associated with "Chromatin Remodeling" GO Process Term (p-Value: 2.08E-6).

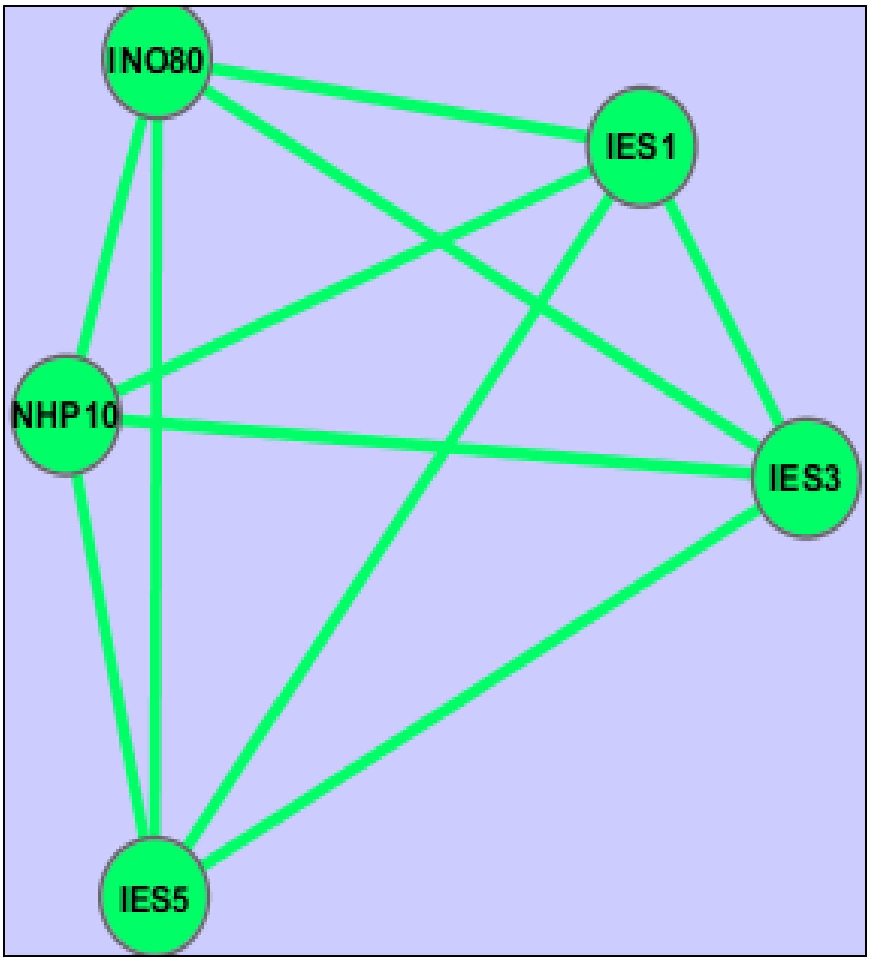


Figure 3.18. Core module M5.5.

Ies1p, Ies3p, Ies5p, Nhp10p and Ino80 were related to INO80 Complex (Table 3.16) which functions as an integral component of a multisubunit ATP-dependent Chromatin Remodeling Complex. DNA microarray studies showed that INO80 regulates about 20% of *Saccharomyces cerevisiae* genes both positively and negatively. Recent studies also revealed that INO80 Complex have roles in transcription, DNA Repair and DNA Replication (Conaway and Conaway, 2008, Udugama *et al*., 2010).

Table 3.16. Components of INO80 complex (Conaway and Weliky, 2008).

Yeast	**Human**	**D.melanogaster**	**Domain Structure**
Ino80p	INO80	dIno80	Snf2-like ATPase
Actin	B-actin (ACTB)	Actin 5C	
Arp4p	Baf53a (Arp4, ACTL6A)	ND	Actin-related proetin
Arp5p	Arp5 (ACTR5)	Arp5	Actin-related proetin
Arp8p	Arp8(ACTR8)	Arp8	Actin-related proetin
Rvb1p	RuvB-like 1 (Tip49, RUVBL1)	Pontin	AAA+ ATPase
Rvb2p	RuvB-like2 (Tip49, RUVBL2b)	Reptin	AAA+ ATPase
Ies2p	Ies2 (INO80B, PAPA-1)	ND	Zinc finger-HIT domain
Ies6p	Ies6 (INO80C, c18orf37)	ND	
Nhp10p	-	-	HMG type-II domain
Taf14p	-	-	
Ies1p	-	-	
Ies3p	-	-	
Ies4p	-	-	
Ies5p	-	-	
-	YY1	Pho	Gli-Kruppel zinc finger transcription deubiquitylating enzyme
-	Uch37 (UCHL5)	Uch37 (Uch-L3)	UCH family deubiqutylating enzyme
-	NFRKB (INO80G)	Nfrkb (CG11970)	
-	MCRS1 (MCRS2, MSP58, INO80Q)	ND	FHA domain
-	TFPT	ND	
-	INO80D (FLJ20309)	ND	
-	INO80E (CCDC95, FLJ90652)	ND	Coiled-coil domain

When the module was extended to M5.5E, the newly formed module consisted of 8 proteins with 14 interactions identified among them. The genes encoding these proteins in the extended module were determined to be significantly associated with the “Chromatin Remodeling” GO Process Term (p Value: 2.74E-5) as it was the case for the original module.

The figure below graphically represents the extended module M5.5E, the green nodes indicating members of the core module, the white nodes indicating the new members of the extended module and the red nodes indicating first neighbors of the extended module that are associated with INO80 Complex.

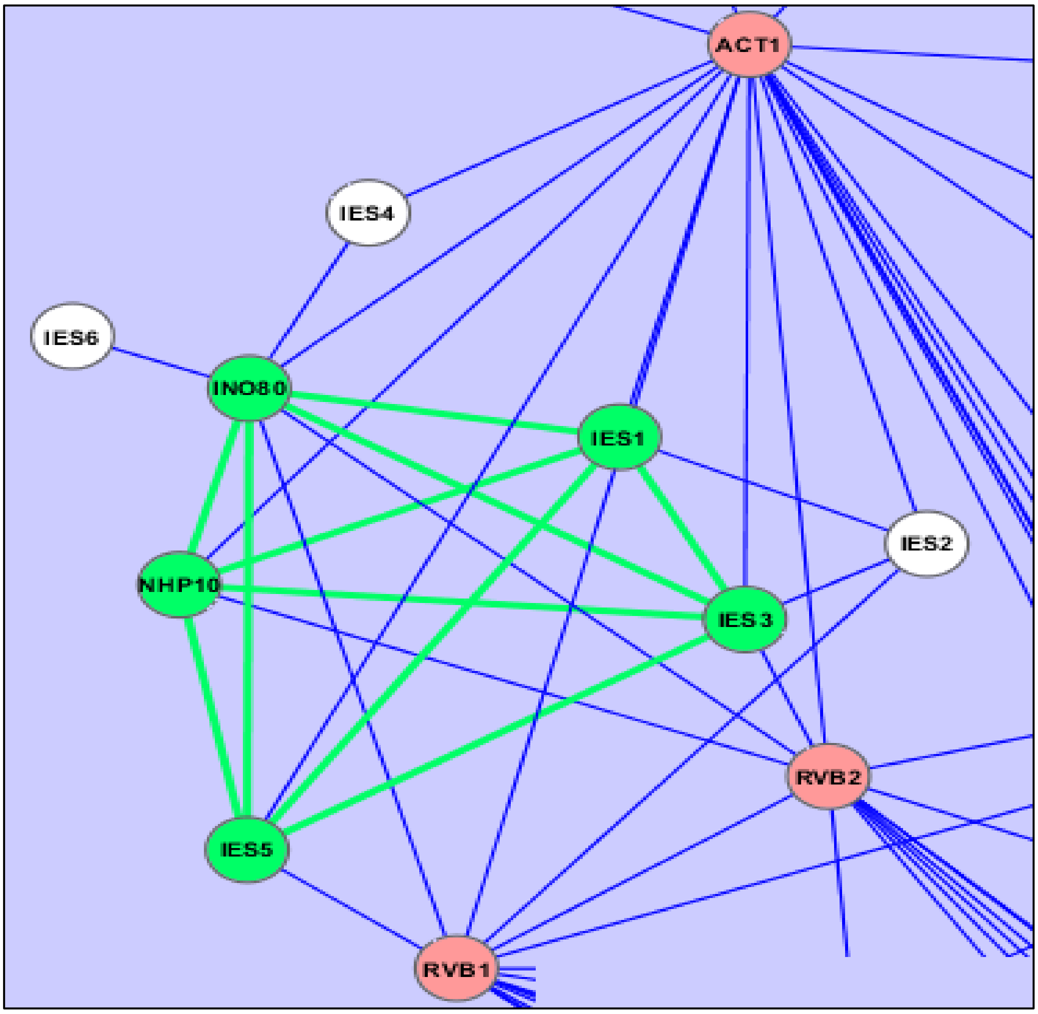


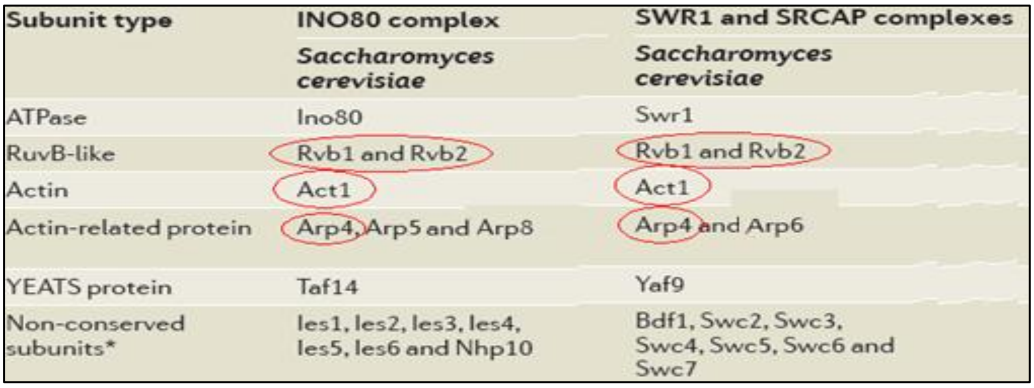
Figure 3.19. Extended module M5.5E and neighbors.

Ies1p, Ies2p, Ies3p, Ies4p, Ies5p, Ies6p and Nhp10p were stated to be identical subunits of yeast INO80 Complex (Shen and Morrisson, 2009). When the module was extended to M5.5E; Ies2p; protein that associates with the INO80 Chromatin Remodeling Complex under low-salt conditions, Ies4p; component of the INO80 Chromatin Remodeling Complex and Ies6p; protein that associates with the INO80 Chromatin Remodeling Complex under low-salt conditions (Shen *et al*., 2003) were captured thus all the identical subunits of INO80 Complex were captured.

Act1p (actin) is a structural protein involved in cell polarization, endocytosis, and other cytoskeletal functions, Rvb1p and Rvb2p were essential proteins involved in transcription regulation that are required for assembly and function of the INO80 Complex (Jonsson *et al*., 2001). These three proteins were captured as neighbors of the extended module which were also shown as components of INO80 Complex in Table 3.16. Arp4p, Arp5p, Arp8p (nuclear actin-related proteins involved in chromatin remodeling) (Shen *et al*., 2003) and Taf14p; subunit of TFIID, TFIIF, INO80, SWI/SNF, and NuA3 Complexes were eliminated due to the GO Term Elimination Criteria.

Further investigation related to module M5.5E leads to recognize SWR1 Complex which includes some of the INO80 Complex subunits. Rvb1p, Rvb2p and Act1p presented in Table 3.16 as INO80 subunits were captured in the module related to SWR1 Complex (Table 3.17 “proteins in the red circles are common”). This module consisted of 7 nodes and 20 edges with the GO-Process Term of “Histone Exchange” (p-Value: 4.76E-19).

Table 3.17. INO80, SWR1 and SRCAP chromatin remodeling complexes (Shen and Morisson, 2009).



Subunit type	**INO80 complex**	**SWR1 and SRCAP complexes**
	Saccharomyces cerevisiae	***Saccharomyces cerevisiae***
ATPase	Ino80	Swr1
RuvB-like	Rvb1 and Rvb2	Rvb1 and Rvb2
Actin	Act1	Act1
Actin-related protein	Arp4, Arp5 and Arp8	Arp4 and Arp6
YEATS protein	Taf14	Yaf9
Non-conserved subunits*	Ies1, Ies2, Ies3, Ies4, Ies5, Ies6 and Nhp10	Bdf1, Swc2, Swc3, Swc4, Swc5, Swc6 and Swc7

INO80 and SWR1 Complexes were investigated to function cooperatively with their histone substrates. Recent studies revealed that the INO80 and SWR1 Complexes have crucial functions in many essential processes, containing DNA repair, checkpoint regulation, DNA Replication, Telomere Maintenance and Chromosome Segregation (Shen and Morisson, 2009).

As a consequence; all the identical subunits (Ies1p, Ies2p, Ies3p, Ies4p, Ies5p, Ies6p and Nhp10p) and the main subunit (Ino80p) of yeast INO80 Complex were captured via our approach. Besides, Chromatin Remodeling Mechanisms of the INO80 and SWR1 Complexes in diverse nuclear processes were reported to be largely unknown (Shen and Morisson, 2009).

The absence of Act1p, Rvb1p and Rvb2p in M5.5E but presence in another module requires further investigation. To decipher the mechanisms of INO80 and SWR1 Complexes can be better understood in a deeper manner.

3.4.6. Module M5.6

Module M5.6 consisted of 5 proteins (Figure 3.20): Esa1p; catalytic subunit of the Histone Acetyltransferase Complex (NuA4) (Clarke *et al.*, 1999), Eaf1p; component of the NuA4 Histone Acetyltransferase Complex which acts as a platform for assembly of NuA4 subunits into the native complex, Eaf5p; Esa1p-associated factor which is non-essential subunit of the NuA4 Acetyltransferase Complex, Eaf7p; subunit of the NuA4 Histone Acetyltransferase Complex which acetylates the N-terminal tails of histones H4 and H2A (Krogan *et al.*, 2004), Yng2p; subunit of NuA4 and identified to be significantly associated with “Protein Amino Acid Acetylation” GO-Process Term (p-Value: 3.54E-6).

Esa1p, Eaf1p, Eaf5p, Eaf7p and Yng2p were reported to be members of NuA4 Histone Acetyltransferase Multisubunit Complex which have roles in the regulation of transcription (especially in DNA repair) (Doyon *et al.*, 2003). Additionally; Nua4 was stated to be the only essential Histone Acetyltransferase Complex in *Saccharomyces cerevisiae* which acetylates the N-terminal tails of histones H4 and H2A (Krogan *et al.*, 2004).

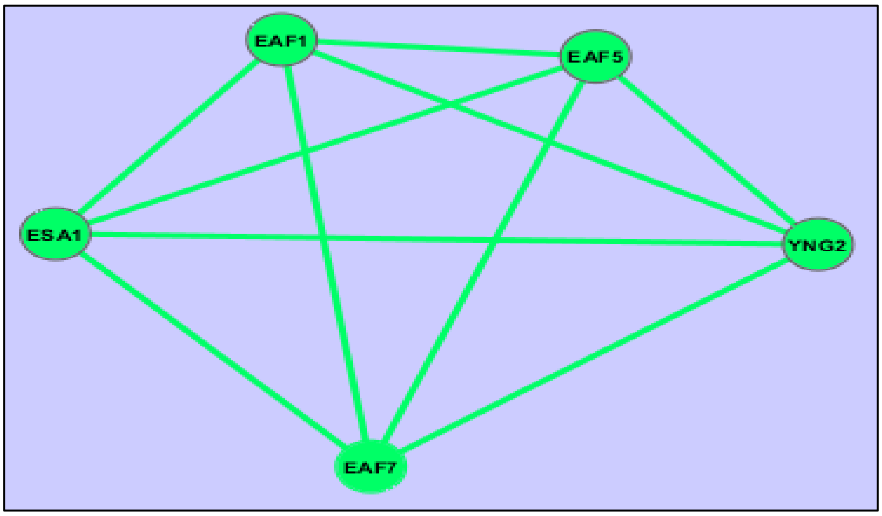


Figure 3.20. Core module M5.6.

The accumulating evidence suggest that more than 20 histone acetyltransferase and histone deacetylases present in *S. cerevisiae*, but only Esa1p (the catalytic enzyme of NuA4) is essential for cell viability among these subunits. It was also mentioned that "piccolo NuA4" core which acetylates the N-terminal tails of histones H2A, H4 and Htz1p in vivo is formed by the collaboration between Esa1p, Yng2p and Epl1p (eliminated due to GO Term Elimination Criteria) (Lin *et al*., 2009).

NuA4 Complex was contains 13 subunits (Eaf7p, Eaf3p, Eaf5p, Eaf6p, Eaf1p, Eaf2p, Tra1p, Eaf2p, Arp4p, Yaf9p, Act1p, Epl1p, Yng2p, Esa1). These subunits and their organization are shown in Figure 3.21 (Doyon and Cote, 2004).

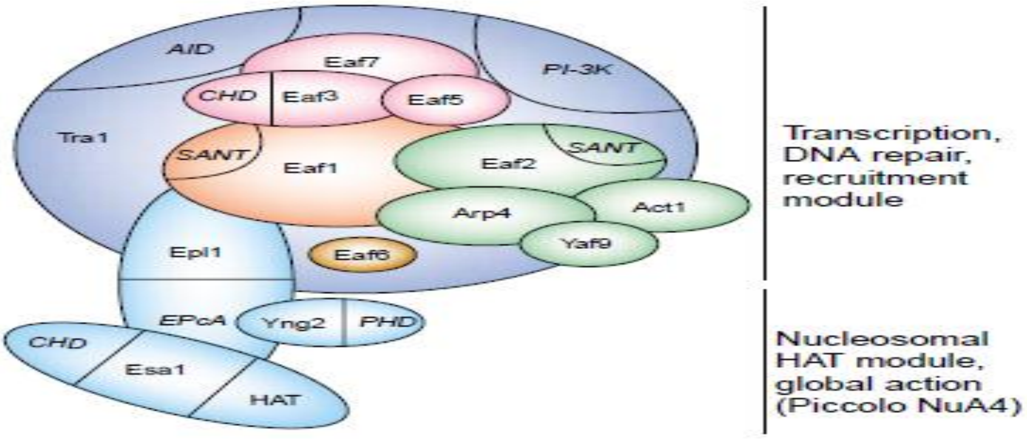


Figure 3.21. Yeast NuA4 HAT complex (Doyon and Cote, 2004).

Eaf6p, Eaf3p, Arp4p and Epl1p in Figure 3.21 were eliminated during GO Term Elimination Process so their absences in the interaction data lead results in their absence from M5.6.

When fluff mode applied to extend the module M5.6, only Swc4p (also known asEaf2p and a component of the NuA4 Histone Acetyltransferase Complex) obtained in the module. Thus M5.6E consisted of 6 nodes, 15 edges and associated with the “Protein Amino Acid Acetylation” GO Process Term (pValue: 3.7337E-10) as the case for the original module.

Figure 3.22 graphically represents the extended module M5.6E, the green nodes indicating members of the core module, the white node (Swc4p) indicating the new member of the extended module and the red nodes (Act1p, Yaf9p and Tra1p) indicating first neighbors of the extended module that are associated with Nua4 Complex.

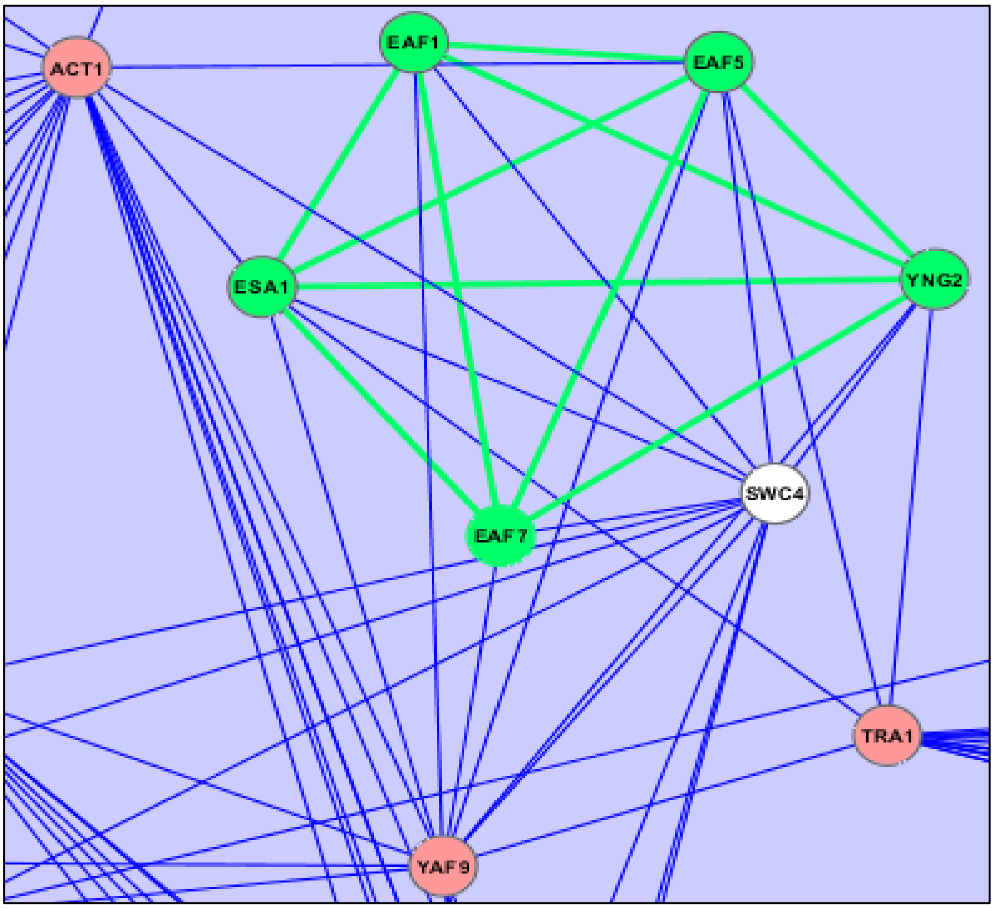


Figure 3.22. Extended module M5.6E and neighbors.

It is interesting to note that Tra1p and Yaf9p are also members of SAGA and SWR1 Complexes respectively. Therefore Tra1p, Act1p and Yaf9p could also be captured in different modules. Additionally; Act1p is a ubiquitous, conserved cytoskeletol element in many cellular processes.

3.4.7. Module M5.7

Module M5.7 consisted of 5 proteins (Figure 3.23): Cbc2p; small subunit of the Heterodimeric Cap Binding Complex (Shen *et al*., 2000), Mud1p; U1 snRNP protein which is involved in nuclear mRNA splicing (Neubauer *et al*., 1997), Nam8p; RNA binding protein which is a component of the U1 snRNP protein (Qiu *et al*., 2011), Snp1p; component of U1 snRNP which is required for mRNA splicing via spliceosome (Kao and Siliciano, 1992), Sto1p; large subunit of the nuclear mRNA Cap Binding Complex (Das *et al*., 2000) and identified to be significantly associated with “Nuclear mRNA Splicing via Splicesome” GO-Process Term (p-Value: 3.88E-10).

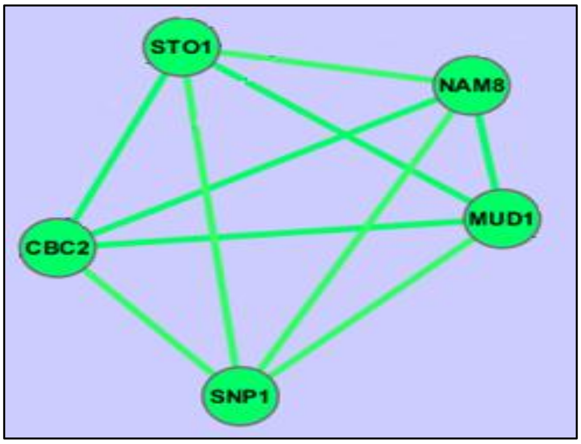


Figure 3.23. Core Module M5.7.

Spliceosome is composed of U1, U2, U3, U4, U5 and U6 snRNA. Splicing occurs simply as in the order seen in Figure 3.24 (x: branch point recognition sequence). U1 binds to 5’ splice site then U2 binds to branch point recognition sequence, U4, U5, U6 binds also to 5’ splice site, U1 is displaced, U6 binds to U2, finally U5 binds to exon sequence then the intron is removed (Beggs, 2001). The yeast Cap Binding Complex was also stated to play an important role in mRNA splicing (Lewis *et al*., 1996).

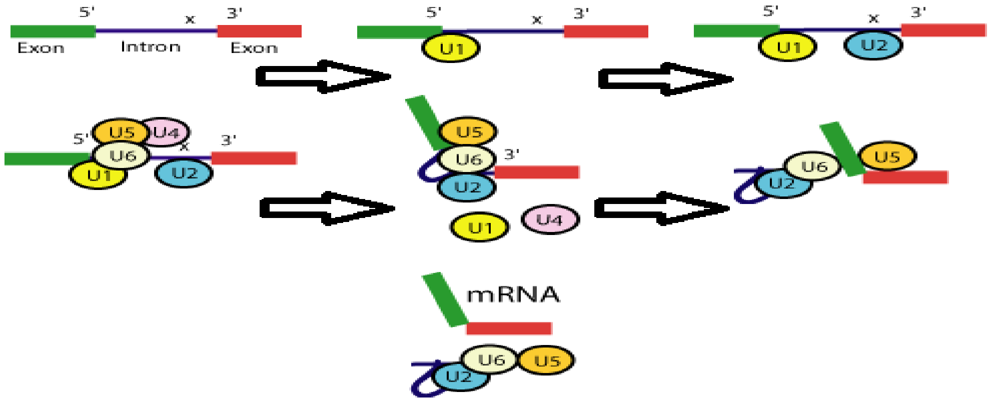


Figure 3.24. Splicing (Beggs, 2001).

When the module was extended to M5.7E, the newly formed module consisted of 10 proteins with 18 interactions identified among them. The genes encoding these proteins in the extended module were determined to be significantly associated with the "Nuclear mRNA Splicing via Splicesome" GO-Process Term (p-Value: 3.01E-9) as it was the case for the original module. Figure 3.25 graphically represents the extended module M5.7E (green nodes: members of the core module, white nodes: new members of the core module (green and white nodes forms the extended module), red nodes: first neighbors of the extended module).

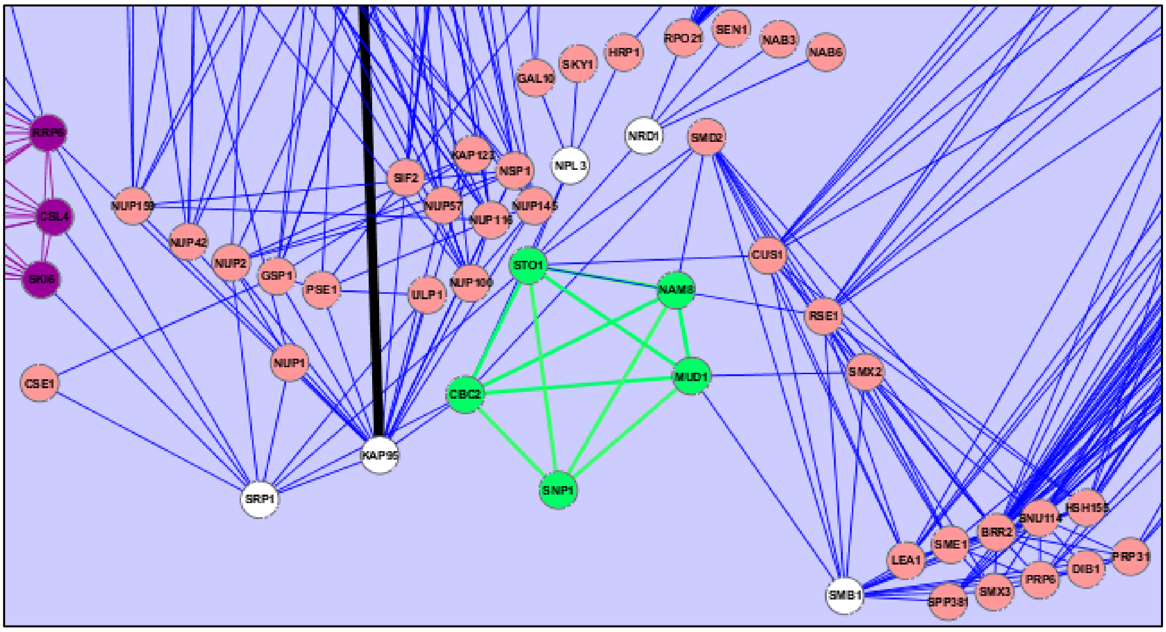


Figure 3.25. Extended module M5.7E and neighbors.

Srp1p; karyopherin alpha homolog which forms a dimer with Kap95p to mediate import of nuclear proteins (Tabb *et al*., 2000), Kap95p; karyopherin β which forms a complex with Srp1p/Kap60p (Enenkel *et al*., 1995, Mackinnon *et al*., 2009) and essential for nuclear accumulation of the U6 snRNP (Spiller *et al*., 2007), Smb1p; part of Heteroheptameric Complex (with Smd1p, Smd2p, Smd3p, Sme1p, Smx3p, and Smx2p) which is a part of the spliceosomal U1, U2, U4, and U5 snRNPs (Salgado-Garrido *et al*., 1999), Npl3p; RNA-binding protein, which is required for pre-mRNA splicing (Kress *et al*., 2008), Nrd1p; RNA-binding protein which interacts with the C-terminal domain of the RNA polymerase II large subunit (Conrad *et al*., 2000), were the members of the extended module and captured by turning the fluff option on. All the proteins that are members of M5.7E have a role in the splicing mechanism.

Nup1p, Nup2p, Nup42p, Nup57p, Nup100p, Nup116p, Nup145p, Nup159p, Nsp1p and Pse1p are components of the Nuclear Pore Complex (NPC). These proteins were bound to M5.7E via the karyopherin; Kap95p that was captured in M5.7E. Karyopherins were previously reported to bind to specific import or export signals and mediate substrate docking to the NPC (Rout *et al*., 2000). Gsp1p, which is regulated by Kap95p (Floer and Blobel, 1996), was also captured as a neighbor of the extended module.

Several other first neighbors of the extended module; Smd2p, Cus1p, Rse1p, Smx2p, Lea1p, Sme1p, Brr2p, Snu114p, Hsh155p, Spp381p, Smx3p, Prp6p, Dip1p and Prp31p, were reported as components of U1, U2, U3, U4, U5 or U6 snRNPs (Saccharomyces Genome Database, www.yeastgenome.org).

Cse1p; a nuclear envelope protein that mediates the nuclear export of importin alpha (Srp1p) was captured as a first neighbor of M5.7E (Hood and Silver, 1998). Sen1p and Nab3p also form a complex with the M5.7E member Nrd1p constituting the NRD1 Complex (Vasiljeva *et al*., 2008).

Additionally Ulp1p, Sif2p, Gal10p, Sky1p, Hrp1p, Rpo2p and Nab6p were also identified as first neighbors of the extended module and these proteins were related to histone deacetylase, galactose or RNA metabolisms. Functions associated with the spliceosome were not previously reported to be linked to the functions of histone deacetylase, galactose or RNA metabolisms. Therefore this functional association may require further investigation.

3.4.8. Module M5.8

Module M5.8 consisted of 5 proteins (Figure 3.26): Rxt3p; the subunit of the RPD3L Complex, which is involved in histone deacetylation (Carrozza *et al*., 2005), Pho23p; a probable component of the RPD3 Histone Deacetylase Complex (Loewith *et al*., 2001), Rpd3p; a histone deacetylase, which regulates transcription, silencing, and other processes by influencing chromatin remodeling (Rundlett *et al*., 1996), Cti6p; a PHD domain protein, which relieves transcriptional repression by binding to the Cyc8p-Tup1p corepressor (Papamichos *et al*., 2002), Dep1p; a transcriptional modulator involved in the regulation of structural phospholipid biosynthesis genes and metabolically unrelated genes (Askree *et al*., 2004) and the module members were identified to be significantly associated with “Histone Deacetylation” GO-Process Term (p-Value: 7.65E-13).

Cti6p, Dep1p, Rxt3p, Pho23p, Rpd3p were stated to be the subunits of RPD3 Histone Deacetylase Complex and two distinct RPD3 Complexes (RPD3L and RPD3S) were characterized previously. Cti6p, Dep1p, Rxt3p and Pho23p were reported to be unique subunits of RPD3L whereas Rpd3p was a subunit of both RPD3L and RPD3S (Carrozza *et al*., 2005).

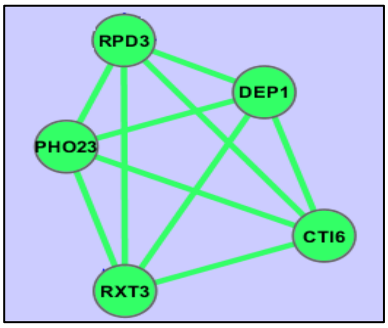


Figure 3.26. Core module M5.8.

When the module was extended (M5.8E), the newly formed module consisted of 9 proteins with 15 interactions identified among them. The genes encoding these proteins in the extended module were determined to be significantly associated with the "Protein Amino Acid Deacetylation" GO Process Term (p Value: 1.18E-18). This was the only case among the five-membered modules, in which the inclusion of additional proteins to the core module in order to obtain the extended module resulted in a change in the most significant GO Process association of the module members. The members of the extended module were less significantly associated with the child term "Histone Deacetylation" (p-Value: 2.77E-16) indicating a loss of specificity in the function of the module.

The figure below (Figure 3.27) graphically represents the extended module M5.8E with green nodes indicating the members of the core module, the white nodes indicating the new members of the extended module, the red nodes indicating first neighbor of the extended module.

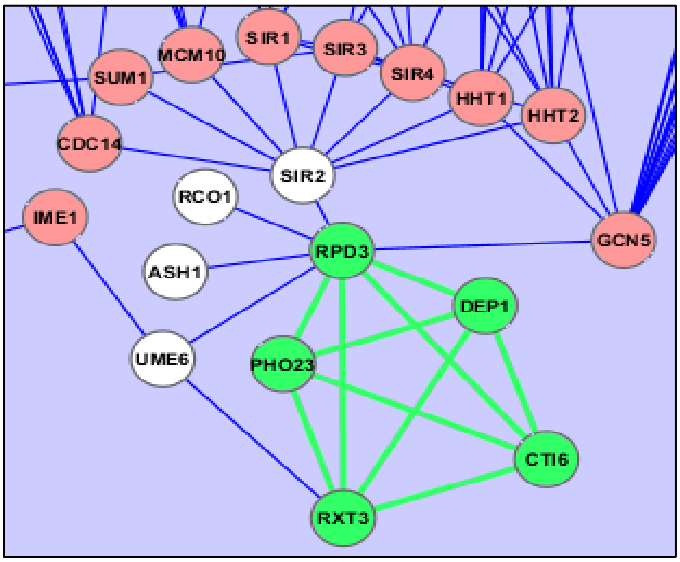


Figure 3.27. Extended module M5.8E and neighbors.

Ume6p; a key transcriptional regulator of early meiotic genes (Strich *et al*., 1994) and a unique subunit of the Rpd3L Complex (Carrozza *et al*., 2005), Ash1p; a unique subunit of the Rpd3L Complex (Carrozza *et al*., 2005), Rco1p; an essential subunit of the Histone Deacetylase Rpd3S Complex, Sir2p; a conserved NAD+ dependent histone deacetylase of the sirtuin family involved in regulation of lifespan (Blander and Guarente, 2004), were the new members of the extended module and they were captured by turning the fluff option on.

RPD3 Complex consisted of Rpd3p, Pho23p, Cti6p, Rxt3p, Dep1p, Dep1p, Ume6p, Ash1p, Rco1p, Sin3p, Ume1p, Sap30p, Sds3p, Rxt2p and Eaf1p (Carrozza *et al*., 2005). Sin3p, Ume1p, Sap30p, Sds3p, Rxt2p and Eaf1p were initially eliminated due to the GO Term Elimination Criteria and the rest of the RPD3 subunits were captured inM5.8E. A vague association between Sir2p and RPD3 Complex were previously reported (Zhou *et al*., 2009) and this was also captured in M5.8E.

Sir1p, Sir2p, Sir3p, Sir4p are named as the "Silent Information Regulator” genes in yeast and they form the SIR Complex (Sharp *et al*., 2003). Sir2p was captured in M5.8E and the other SIR proteins were first neighbors of the extended module.

Ime1p was reported to have a role in the transcription of early meiotic genes through an interaction with Ume6p; a member of the extended module (Kassir *et al*., 2003). The other neighbors of the extended module were either associated with mitotic processes as in the case of Cdc14p and Sum1p, or they were associated with the operation of a functional chromatin as in the case of Mcm10p, Hht1p, Hht2p and Gcn5p. It is known that DNA is packed into chromatins, which consists of DNA, structural histone proteins and non-histone proteins. Histones are responsible for the maintenance of the shape and structure of chromatins. Mitotic recombination was also reported to be affected by changes in chromatin architecture, which is mediated by RPD3 Complex (Dora *et al*., 1999). Thus the presence of the mitotic and chromatin related proteins as first neighbors to the Histone Deacetylation Complex related module M5.8E could be explained.

In this study, all proteins that were associated with the Process Gene Ontology (GO) term “Glucose Metabolic Process” were collected in order to construct a network of all reported interactions among these entities with a stringent confidence threshold. The network was then investigated from a modular point of view in order to identify small groups with distinct functional properties that would also be associated with the glucose metabolism in yeast. The identified core and extended modules as well as their first neighbors were investigated for their functional associations. As it may be observed from the analyses above, all 5 membered core modules that would be identified could be significantly associated with a specific function in the cell, either as a part of a complex or as a part of a signaling cascade. The extension of the modules provided additional nodes that would be associated with the same function. The first neighbors of the extended modules were either associated with the same function or would provide information on a related cellular process. These associations were all verified from the available literature on the investigated pathways and complexes. The functional robustness of the module members suggests that this modular approach can be used to assign probable functions to unannotated or uncharacterized loner genes that would be present in the identified functional modules.

3.5. Assigning a Probable Function to an Uncharacterized Member; YGR067c, of a Module involved in Ubiquitin Dependent Protein Catabolic Process

The core module M19.1 consisted of 19 proteins (Figure 3.28): Pre1p, Pre2p, Pre3p, Pre4p, Pre7p and Pup1p; the beta subunits of the 20S proteasome, Pre5p, Pre6p, Pre8p, Pre9p, Pre10p, Scl1p and Pup2p; the alpha subunits of the 20S proteasome, Rpt1p, Rpt2p, Rpt3p and Rpt6p; four of the six ATPases of the 19S proteasome, and Rpn11; the metalloprotease subunit of the 19S regulatory particle (Fischer *et al*., 1994, Groll et al., 1999, Kikuchi *et al*., 2010), Rpn7p; the non-ATPase regulatory subunit of the 26S proteasome (Glickman *et al*., 1998), with 170 interactions among them. This module was identified to be significantly associated with the "Ubiquitin Dependent Protein Catabolic Process" GO-Process Term (p-Value: 7.15E-30).

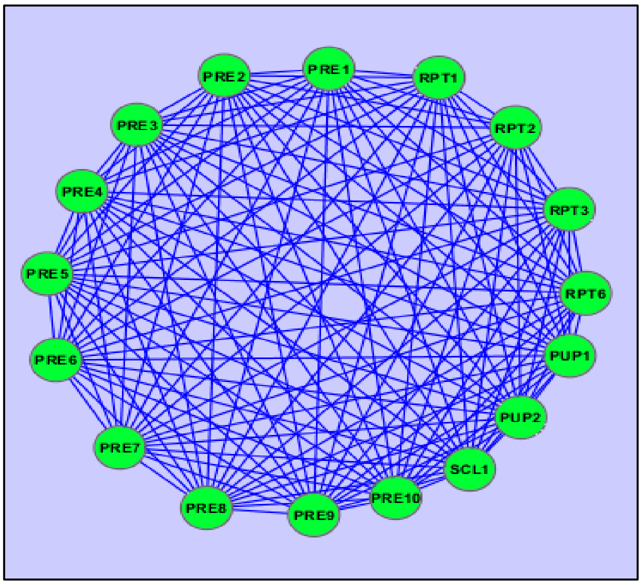


Figure 3.28. Core module M19.1.

The 26S proteasome consisted of two parts (19S regulatory particle and a 20S catalytic core particle) and is responsible for the ubiquitination of a protein (Hochstrasser, 1996). The 20S catalytic core also comprised of 7 alpha-type and 7 beta-type subunits (Hochstrasser, 1996, Heinemeyer *et al.*, 1994). 13 of the 14 subunits of 20S Catalytic Core Particle were captured except for the beta 3 subunit of the 20S catalytic core particle in the core module.

Additionally, the 19S regulatory particle is comprised of the lid and the base subcomplexes. The lid complex is comprised of Rpn3p, Rpn5p, Rpn6p, Rpn7p, Rpn8p, Rpn9p, Rpn11p, Rpn12p and Sem1p and the base complex is comprised of Rpt1p, Rpt2p, Rpt3p, Rpt4p, Rpt5p and Rpt6p (Lander *et al.*, 2012). Four members of the base complex; Rpt1p, Rpt2p, Rpt3p and Rpt6p, as well as two members of the lid complex; Rpn11p and Rpn7p were captured in the core module.

The extended module of this core complex comprised of 40 nodes with 300 interactions identified among them and the extended module was also identified to be significantly associated with “Ubiquitin Dependent Protein Catabolic Process” GO Process Term (p-Value: 2.99E-39). The figure below (Figure 3.29) graphically represents the extended module with green nodes indicating the members of the core module and the white nodes indicating the new members of the extended module.

Pup3p; beta 3 subunit of the 20S proteasome, which is involved in ubiquitin dependent catabolism (Groll *et al.*, 1999) , Rpn1p; a non-ATPase base subunit of the 19S regulatory particle of the 26S proteasome (Elsasser *et al.*, 2002), Rpn4p; a transcription factor that stimulates expression of proteasome genes also Rpn4p levels are in turn regulated by the 26S proteasome (Xie and Varshavsky, 2001), Rpn2p; subunit of the 26S proteasome (De Marini *et al.*, 1995), Rpn14p; a proteasome-interacting protein involved in the assembly of the base subcomplex of the 19S proteasome regulatory particle (Saeki *et al.*, 2009), Rpt4p and Rpt5p; two of six ATPases of the 19S proteasome (Glickman *et al.*, 1998), Nas2p; a proteasome-interacting protein involved in the assembly of the base subcomplex of the 19S proteasomal regulatory particle (Saeki *et al.*, 2009), Doa4; a ubiquitin isopeptidase which required for recycling ubiquitin from proteasome-bound ubiquitinated intermediates (Papa *et al.*, 1999), Ubp5p; a putative ubiquitin-specific

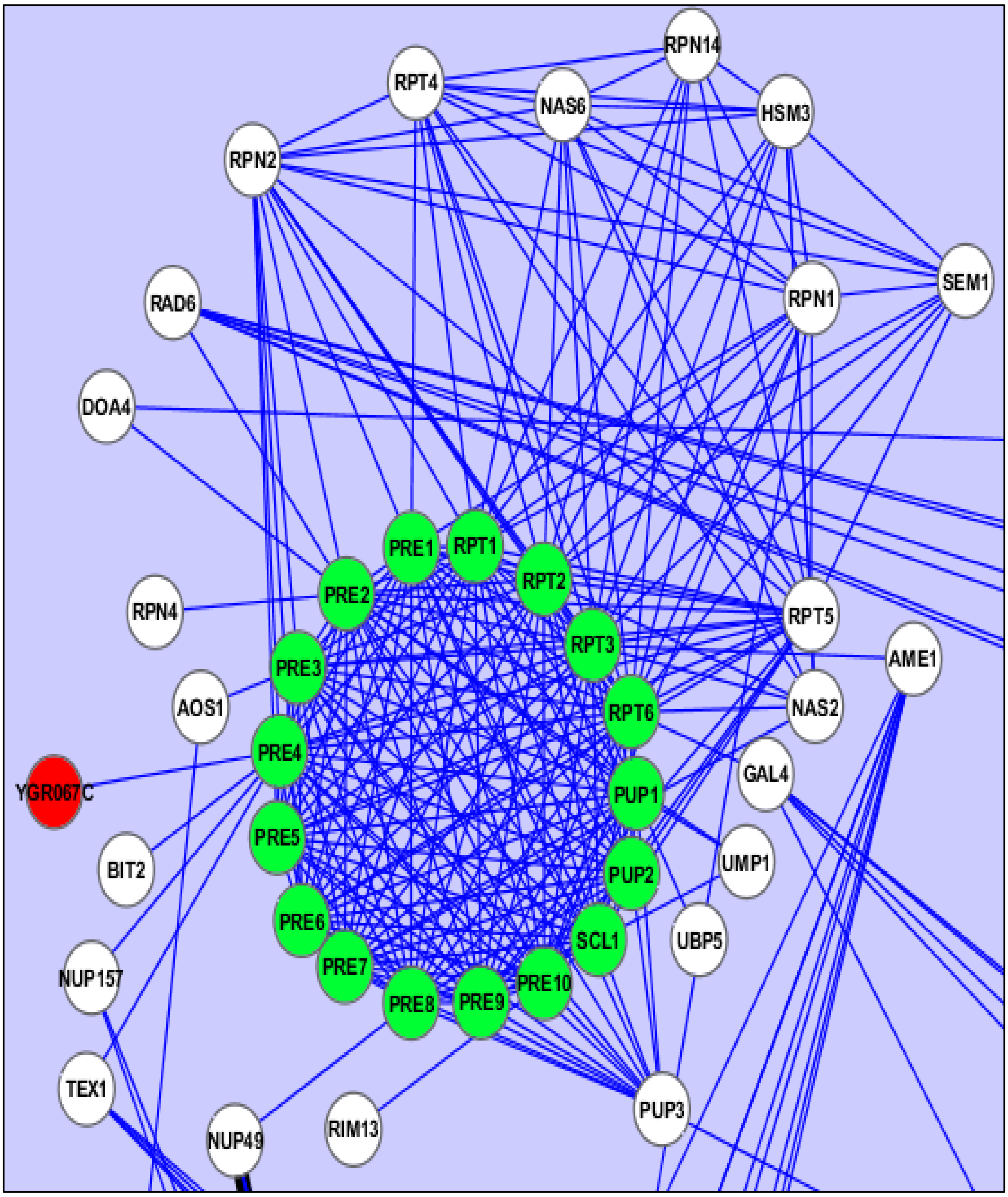


Figure 3.29. Extended module M19.1E.

protease, Ump1p; a short-lived chaperone required for correct maturation of the 20S proteasome (Ramos *et al*., 1998), Hsm3p; a proteasome-interacting protein involved in the assembly of the base subcomplex of the 19S proteasomal regulatory particle (Le Tallec *et al*., 2009), Sem1p; a component of the lid subcomplex of the regulatory subunit of the 26S proteasome (Funakoshi *et al*., 2004) and Nas6p; a proteasome interacting protein involved in the assembly of the base subcomplex of the 19S proteasomal regulatory particle (Saeki *et al*., 2009) were captured as proteins related to 26s proteasome in the extended module.

The two base components of the 19S regulatory particle, Rpt6p and Rpt4p, were reported to interact specifically with the Gal4-activation domain (Chang *et al*., 2001). Rad6p; a ubiquitin-conjugating enzyme was also reported to have a role in the ubiquitin dependent protein catabolic process via the N-end rule pathway and it can ubiquitinate histones *in vitro* (Robzyk *et al*., 2000). Aos1p; a subunit of a heterodimeric nuclear small ubiquitin-like modifier (SUMO) activating enzyme, whose role in cell viability is to conjugate Smt3p; a ubiquitin like protein of the SUMO family, was captured as a member of the extended module. SUMO proteins are similar to ubiquitin proteins and a subset of SUMO-2-conjugated proteins were reported to be subsequently ubiquitinated and degraded by the proteasome (Schimmel *et al*., 2008).

Nup157p, Nup49p were members of the nuclear pore complex. Bit2p; a subunit of TORC2, Tex1p; a component of the transcription export (TREX) complex, which is involved in mRNA export, Rim13p; a calpain like cysteine protease involved in proteolytic activation and Ame1p; an essential kinetochore protein, were also captured in the M19.1E. These proteins probably have functional roles that support the function of the module significantly associated with the “Ubiquitin Dependent Protein Catabolic Process” GO Process Term and they need further investigations.

The proteins captured in the M19.1 were only the subunits of the 26s proteasome. When the module extended to M19.1E, the functional related proteins were also included in the module. The beta 3 subunit of the 20S catalytic core particle was captured in the M19.1E completing the subunit composition of the 20S catalytic core particle. Rpt4p and Rpt5p were also captured in the extended module, completing the 6 ATPases of 19S regulatory particle. Rpt4 and Rpt5 were reported to form coiled coil which probably forms the flexibly attached ubiquitin interacting motif Rpn10p, which was eliminated during the application of the GO Term Elimination Criteria. Two of the four non-ATPase subunits were also captured in the extended module. The non-ATPase subunits were reported as

Rpn1p, Rpn2p, Rpn10p and Rpn13p. Rpn10p and Rpn13p were eliminated due to the GO Term Elimination Criteria. Rpn7p and Rpn11p that were captured in the core module and Sem1p that was captured in the extended module were the subunits of the lid complex of 19S regulatory particle. The other subunits of lid complex (Rpn3p, Rpn5p, Rpn6p, Rpn8p and Rpn12p) were eliminated as a result of the application of the GO Term Elimination Criteria.

The extension of the module provided additional nodes that would be associated with the same function. The uncharacterized member "YGR067c" was captured in the extended module M19.1E with one interaction to Pre4p (the beta 7 subunit of the 20S proteasome). The module was significantly associated with the "Ubiquitin Dependent Protein Catabolic Process" GO Process Term therefore; YGR067c would be potentially related to this function and would have a role in association of the proteasome complex.

3.6. Assigning Novel Functions to Characterized Proteins

3.6.1. Bridging Protein between Glucose and Nitrogen Metabolisms

One of the 6 membered modules consisted of Rpt4p; one of the six ATPases of the 19S regulatory particle of the 26S proteasome, which is involved in degradation of ubiquitinated substrates (Rubin *et al*., 1998), Sem1p; a component of the lid subcomplex of the regulatory subunit of the 26S proteasome (Sone *et al*., 2004), Hsm3p; a proteasome-interacting protein involved in the assembly of the base subcomplex of the 19S proteasomal regulatory particle (Le Tallec *et al*., 2009), Rpn14p; a proteasome interacting protein involved in the assembly of the base subcomplex of the 19S proteasome regulatory particle (Saeki *et al*., 2009), Nas6p; a proteasome interacting protein involved in the assembly of the base subcomplex of the 19S proteasomal regulatory particle (Saeki *et al*., 2009) and Rpn1p; a Non-ATPase base subunit of the 19S regulatory particle of the 26S proteasome, which may participate in the recognition of several ligands of the proteasome (Elsasser *et al*., 2002), were captured in the core module. The core module (Figure 3.30) consisted of 6 nodes with 14 interactions identified among them and the module was significantly associated with the "Proteolysis" GO Process Term (p-Value: 4.75E-6).

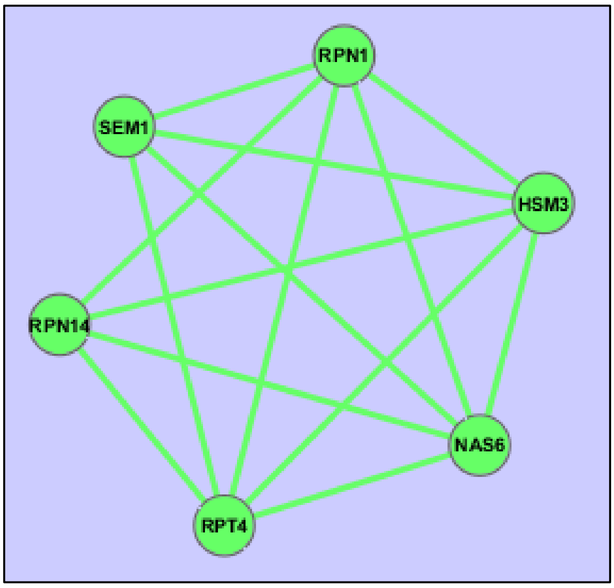


Figure 3.30. Core module M6.1.

Extending the module yielded a newly formed module consisting of 9 proteins with 23 interactions identified among them (Figure 3.31). The genes encoding these proteins in the extended module were determined to be significantly associated with the “Proteolysis” GO Process Term (p-Value: 1.55E-9) as it was the case for the original module.

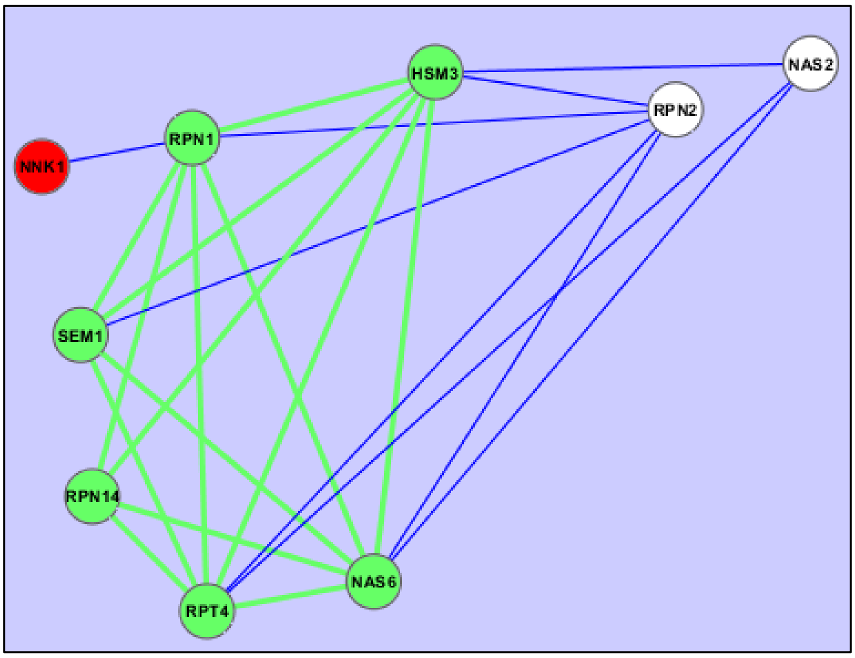


Figure 3.31. Extended module M6.1E.

Rpn2p; a subunit of the 26S proteasome (De Marini *et al*., 1995) and Nas2p; a proteasome-interacting protein involved in the assembly of the base subcomplex of the 19S proteasomal regulatory particle (Saeki *et al*., 2009) were captured in M6.1E.

Nnk1p was recently reported to be Nitrogen Network Kinase implicated in proteasome function (Breitkreutz *et al*., 2010). Although the core proteins of this network were associated with the “Glucose Metabolic Process” Gene Ontology Term, interestingly, Nnk1p, a Nitrogen Network Protein Kinase was captured in the extended module owing to its association with proteolysis. Nnk1p has recently been characterized and verified (2010). The identification of this protein in the extended module is interesting in the sense that this protein might present itself as a protein bridging glucose and nitrogen metabolisms through proteolysis and this would require further investigation.

All members of this extended module that were involved in Proteolysis were also captured in the M19.1E which was significantly associated with the “Ubiquitin Dependent Protein Catabolic Process” GO-Process Term. Ubiquitin is attached to proteins and label them for their degradation. The proteins are directed to proteasomes after the labeling.

3.6.2. Role of a GTPase Inhibitor in Cell Cycle

The 11 membered core module consisted of 11 proteins (Figure 3.32): Skp1p; an evolutionarily conserved kinetochore protein which have roles in the G1/S transition of mitotic cell cycle (Bai *et al*., 1996), Grr1p; F-box protein component of the SCF ubiquitin-ligase complex which have role in the G1 cyclin-controlled cell division (Barral *et al*., 1995), Cln3p; G1 cyclin involved in cell cycle progression, Cdc28p; the catalytic subunit of the main cell cycle cyclin-dependent kinase which alternately associates with G1 cyclins (Dirick *et al*., 1995), Swi4p and Swi6p comprising the SBF complex (Swi4p-Swi6p) and a transcriptional activator that in concert with MBF (Mbp1-Swi6p) regulating the late G1-specific transcription of targets including cyclins and genes required for DNA synthesis and repair (Nasmyth and Dirick, 1991), Cdc53p; a structural protein of the SCF complexes which promotes the G1-S transition by targeting G1 cyclins (Mathias *et al*., 1996), Hrt1p; a RING finger containing subunit of Skp1-Cullin-F-box ubiquitin protein ligases which reported to have function in G1/S transition of mitotic cell cycle (Ohta *et al*., 1999), Cln2p;

G1 cyclin involved in regulation of the cell cycle, Cdc34p; a Ubiquitin-conjugating enzyme (E2) and catalytic subunit of SCF ubiquitin-protein ligase complex, Cdc4p; F-box protein required for G1/S and G2/M transition (Hoyt, 1997, Goh and Surana, 1999), with 27 interactions present among them and the module was identified to be significantly associated with "G1/S transition of mitotic cell cycle" (p-Value: 1.37E-17).

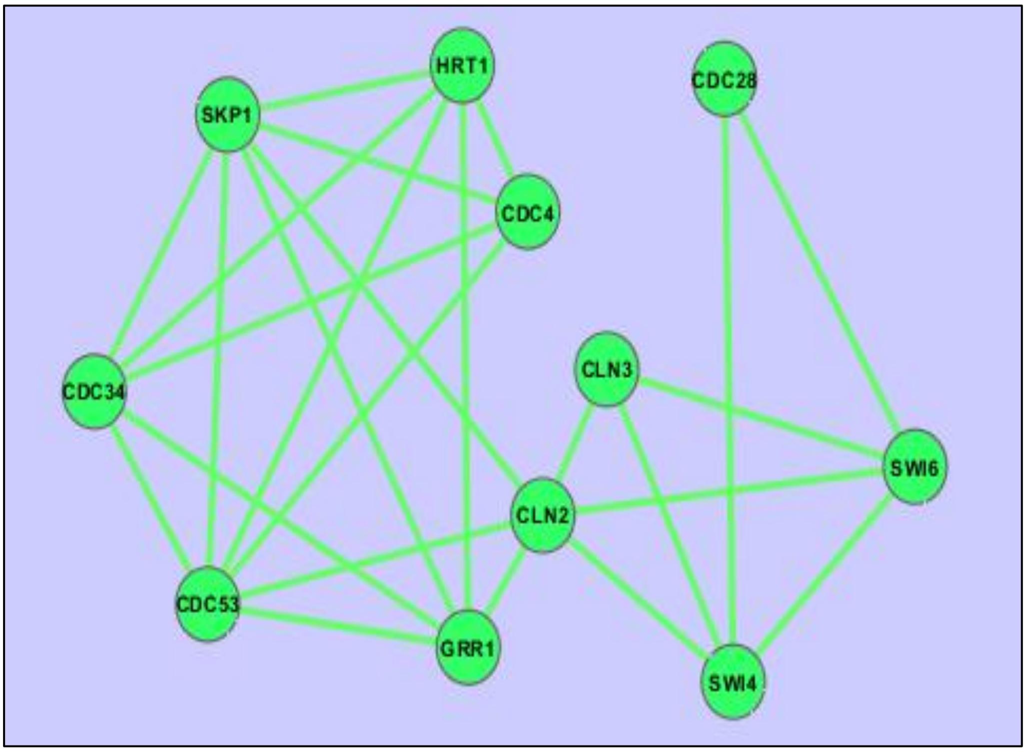


Figure 3.32. Core module M11.1.

When the module was extended (Figure 3.33), the newly formed module M11.1E consisted of 39 proteins with 83 interactions identified among them. The genes encoding these proteins in the extended module were determined to be significantly associated with the "Interphase" GO Process Term (p-Value: 4.34E-26). Inclusion of additional proteins to the core module in order to obtain the extended module resulted in a change in the most significant GO Process association of the module members. The members of the extended module were less significantly associated with the child term "G1/S transition of mitotic cell cycle" (p-Value: 1.37E-17) indicating a loss of specificity in the function of the module.

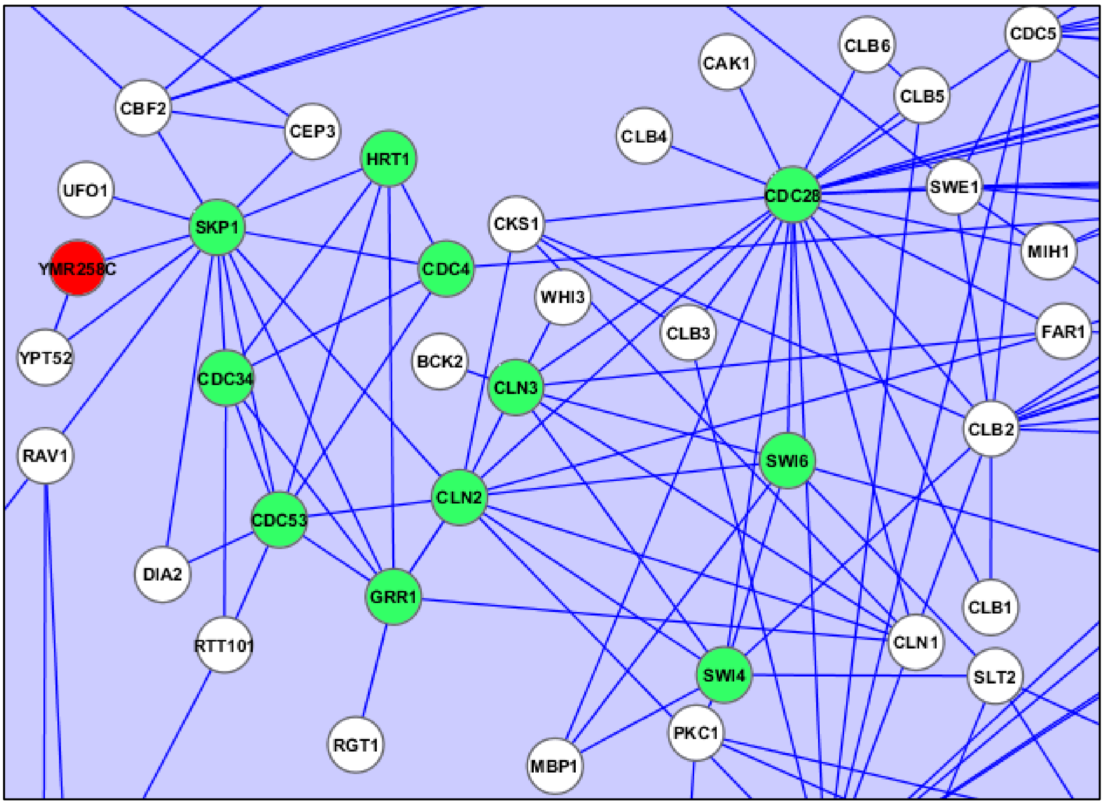


Figure 3.33. Extended module M11.1E.

Clb1p, Clb2p, Clb3p and Clb4p were reported as B-type cyclins involved in cell cycle progression, which activate Cdc28p to promote the G2/M transition. Clb5p (a B-type cyclin involved in DNA replication during S phase) and Clb6p (a B-type cyclin involved in DNA replication during S phase) were also reported to have roles in G1/S transition of mitotic cell cycle (Schwob and Nasmyth, 1993).

Cyclins are known to be encoded by Cln1p, Cln2p, and Cln3p (Sia *et al.*, 1997). Cln2p, Cln3p were members of the core module and Cln1p was captured in the extended module. Whi3p; an RNA binding protein that sequesters Cln3p in cytoplasmic foci (Gari *et al.*, 2001) was also captured in the extended module.

Dia2p was reported to be an origin binding F-box protein, which forms an SCF ubiquitin ligase complex with Skp1p and Cdc53p (Andress *et al.*, 2011). Skp1p and Cdc53p were captured as members of the core module previously and Dia2p was included in the extended module.

Cak1p; a cyclin dependent activating kinase, which activates Cdc28p (promotes the G2/M transition) (Enke *et al*., 2000), Cks1p; a cyclin dependent protein kinase regulatory subunit, which was required for G1 cyclin-cyclin-dependent kinase activity (Reynard *et al*., 2000), Cdc5p; a polo like kinase and a possible Cdc28p substrate, which has functions in mitosis and cytokinesis through substrate phosphorylation (Iacovella *et al*., 2010), Mih1p; a protein tyrosine phosphatase involved in cell cycle control and regulates the phosphorylation state of Cdc28p (Sia *et al*., 1996), Swe1p; a protein kinase that regulates the G2/M transition by inhibition of Cdc28p kinase activity (Booher *et al*., 1993), Mbp1p; a transcription factor involved in regulation of cell cycle progression from G1 to S phase (Koch *et al*., 1993), Far1p; a cyclin dependent kinase inhibitor, which is able to inactivate Cdc28p/Clnp complexes, and thus helps stop the cell cycle at start in G1 (Peter and Herskowitz, 1994, Mendenhall, 1998), Rav1p; a subunit of the RAVE complex (Rav1p, Rav2p, Skp1p all captured in the core module), which associates with the V_1 domain of the vacuolar membrane and which is a potential Cdc28p substrate (Seol *et al*., 2001), Dia2p; an origin-binding F-box protein which forms an SCF ubiquitin ligase complex (SCF promotes the G1-S transition by targeting G1 cyclins) with Skp1p and Cdc53p (Skp1p and Cdc53p were captured in the core module) (Andress *et al*., 2011), Ufo1p; an F-box receptor protein, which is a subunit of the Skp1-Cdc53-F-box receptor (SCF) E3 ubiquitin ligase complex (Kaplun *et al*., 2003) were captured in the extended module, which were all related to interphase process ontology.

Cdh1p; an activator of the anaphase promoting complex (Zachariae and Nasmyth, 1999) and Rtt101p; a cullin subunit of a Roc1p dependent E3 ubiquitin ligase complex with a role in anaphase progression (Michel *et al*., 2003) were also captured in the extended module. These proteins were related with the processes associated with anaphase during the cell cycle.

Cbf2p and Cep3p were essential kinetochore proteins, which are known to have roles during cell division pulling sister chromatids apart (Strunnikov *et al*., 1995, Espelin *et al*., 2003).

Rgt1p, which is the glucose responsive transcription factor that regulates expression of several glucose transporter (HXT) genes in response to glucose, was also present in the extended module. This protein is previously known to have an association with Grr1p, which was captured in the core module and the F box protein Grr1 antagonizes the activity of the transcriptional repressor Rgt1 (Flick *et al*., 2003).

Bck2p, Pkc1p and Slt2p were associated with cell integrity, cell wall remodeling and cell wall integrity, respectively, and Ypt52p was reported to have role in the endocytosis previously in SGD. These proteins probably have functional roles that support the function of the module significantly associated with the cell cycle and they need further investigations.

YMR258c (Roy1p) was an unknown protein at the beginning of the study and it was previously reported as an uncharacterized non–SCF-type F-box protein, which negatively modulates cell viability and intracellular transport by suppressing Ypt52p and Skp1p is necessary for the association of Roy1p with Ypt52p. However, the function of the non–SCF-type F-box proteins remains largely unknown. Roy1p was captured in a module significantly associated with “Interphase” GO Process Term thus this protein would have role in the mentioned GO Process Term.

4. CONCLUSIONS AND RECOMMENDATIONS

4.1. Conclusions

In this study, all proteins that were associated with the Process Gene Ontology (GO) term “Glucose Metabolic Process” were collected in order to construct a network of all reported interactions among these entities with a stringent confidence threshold. After establishing that the constructed network was indeed biological, the network was then investigated from a modular point of view in order to identify small groups with distinct functional properties that would also be associated with the glucose metabolism in yeast. For this purpose, modules with more than four members were investigated. As a major contribution, the highly specific functional nature of the identified modules allowed the assignment of a possible function to an uncharacterized open reading frame and possible novel functions to characterized and verified genes.

At the beginning of the study, two networks were constructed; one general network consisting of a global interaction network of yeast obtained from the STRING database (Global Network, GN) and another network that was constructed by including only the interactions for the nodes of the Glucose Metabolic Process Gene Ontology Term (Glucose Metabolic Network, GMN). These two networks were then compared in terms of their network properties such as number of nodes, edges, the network diameter, network radius, clustering coefficient, characteristic path length, betweenness, closeness, as well as the distributions of degrees, neighborhood connectivities, average clustering coefficients, shortest path lengths. As a result, the GMN was also determined to be a scale free biological network. Modularity analysis was conducted on both networks optimizing the parameters that were used in module identification in order to increase module specificity. The modules of GN had more members and these modules were associated with very general GO terms. On the other hand, the modules that were identified in GMN were more compact and functionally more distinctive with significant gene ontological associations ($p\text{-value} < 10^{-5}$).

In order to determine the functional linkage among the members of the modules that were identified within the framework of this thesis, eight modules with five nodes were selected and further analyzed. All 5-member core modules that were identified were significantly associated with a specific function in the cell, either as a part of a complex or as a part of a signaling cascade. The extension of the modules provided additional nodes that were associated with the same function or other functional associations that would be related to the functions of the module. The first neighbors of the extended modules were associated with the same function or provided information on a related cellular process. All associations were verified from the available literature on the investigated pathways and complexes. Some genes were not captured in the core module, although they were members of the same complex or they were in the same signaling cascade, because of the stringent threshold for the reliability of the interactions between the nodes. A node was included in the module only if that node was connected to other members of the module and the lower confidence score of some interactions resulted in the exclusion of some nodes from the core modules. However, these interactions and nodes were captured in the extended modules. The extended modules were enriched with either the same GO Process as that of the core module or its parent GO Process. The functional robustness of the method in identification of module members associated with the same cellular process suggests that this modular approach can be used to assign probable functions to the unannotated or uncharacterized lone genes present in the identified functional modules.

Three proteins with unknown function were identified as a result of this modularity analysis. In the extended module of a 19-member core module, an uncharacterized ORF, YGR067c was identified. Both the core module and the extended module were significantly associated with ubiquitin-dependent protein catabolic process. Therefore YGR067c may be potentially related to Ubiquitin Dependent Protein Catabolic Process and may have a role in association of the proteasome complex. Nnk1p, which was an uncharacterized protein at the beginning of this study, was identified in the extended module of a 6-member core module, which were both significantly enriched with Proteolysis Process Gene Ontology. Nnk1p has recently been reported to be a Nitrogen Network Kinase implicated in proteasome function. The identification of this protein is interesting in the sense that this protein might be bridging glucose and nitrogen metabolisms through proteolysis and this would warrant further investigation. Roy1p was also an unknown protein at the beginning of the study and it has been reported to be an uncharacterized non–SCF-type F-box protein. Roy1p was captured in the extended module of an 11 member core module, which was significantly associated with the G1/S transition of mitotic cell cycle GO Process Term. The extended module was significantly enriched with Interphase Process. The identification of this protein in the extended module might indicate that Roy1p might have an additional role in cell cycle, which has not been reported yet.

In this study, the Glucose Metabolic Network was investigated from a modular point of view in order to identify functional entities among the network, which would implicate related functionalities, and to assign probable functions to the unannotated or uncharacterized genes, or to assign novel additional functions to the proteins with already known functional associations.

4.2. Recommendations

The analysis of the 5-member modules allowed the identification of a protein, which did not have previously reported associations with the function that was assigned to a particular core module, either as a member of the extended module or as a first degree neighbor to the extended module. These novel functional linkages, which were suggested in the present study based on computational results, can be used as a starting point in the design of experiments for the verification of these suggestions. The 5-member modules were thoroughly analyzed in the present thesis. A similar analysis conducted on other modules would suggest more novel functional linkages that would be similar to those discussed above.

This study indicated that a central metabolic process such as the Glucose Metabolic Process allows the identification of functional modules directly associated with this process, allows assigning probable roles to uncharacterized proteins, which were identified in the modules, or assigning novel roles to characterized proteins, which would be identified in modules with different functional associations than those of that particular protein. Similar type of studies can be conducted using other central metabolic processes as the starting point. This would then enable a cross-comparison of the identified functional modules in different sub-networks.

This study was conducted using Saccharomyces cerevisiae as the model organism. The methodology can be applied to other model or higher organisms depending on the scope of the study as long as detailed information is available on the interaction network of the organism.

APPENDIX A: INFORMATION ABOUT PROTEINS

Table A.1. Core proteins.

Standard Name	**Systematic Name**	**GO Process Terms**
Fyv10p	YIL097W	negative regulation of gluconeogenesis; GO:0045721
Gid7p	YCL039W	negative regulation of gluconeogenesis; GO:0045721
Gid8p	YMR135C	negative regulation of gluconeogenesis; GO:0045721
Rmd5p	YDR255C	negative regulation of gluconeogenesis; GO:0045721
Ubc8p	YEL012W	negative regulation of gluconeogenesis; GO:0045721
Ubp14p	YBR058C	negative regulation of gluconeogenesis; GO:0045721
Vid24p	YBR105C	negative regulation of gluconeogenesis; GO:0045721
Vid28p	YIL017C	negative regulation of gluconeogenesis; GO:0045721
Vid30p	YGL227W	negative regulation of gluconeogenesis; GO:0045721
Cat8p	YMR280C	positive regulation of gluconeogenesis; GO:0045722
Rds2p	YPL133C	positive regulation of gluconeogenesis; GO:0045722
Sip4p	YJL089W	positive regulation of gluconeogenesis; GO:0045722
Snf1p	YDR477W	positive regulation of gluconeogenesis; GO:0045722
Snf4p	YGL115W	positive regulation of gluconeogenesis; GO:0045722
Acn9p	YDR511W	regulation of gluconeogenesis; GO:0006111,gluconeogenesis; GO:0006094
Cat5p	YOR125C	gluconeogenesis; GO:0006094
Eno1p	YGR254W	gluconeogenesis; GO:0006094, glycolysis; GO:0006096
Eno2p	YHR174W	gluconeogenesis; GO:0006094, glycolysis; GO:0006096
Fba1p	YKL060C	gluconeogenesis; GO:0006094, glycolysis; GO:0006096
Fbp1p	YLR377C	gluconeogenesis; GO:0006094
Gpm1p	YKL152C	gluconeogenesis; GO:0006094, glycolysis; GO:0006096
Mdh2p	YOL126C	gluconeogenesis; GO:0006094, glycolysis; GO:0006096
Pck1p	YKR097W	gluconeogenesis; GO:0006094
Pgi1p	YBR196C	gluconeogenesis; GO:0006094, glycolysis; GO:0006096; pentose-phosphate shunt; GO:0006098
Pgk1p	YCR012W	gluconeogenesis; GO:0006094, glycolysis; GO:0006096
Pyc1p	YGL062W	gluconeogenesis; GO:0006094
Pyc2p	YBR218C	gluconeogenesis; GO:0006094
Sdl1p	YIL167W	gluconeogenesis; GO:0006094
Tdh1p	YJL052W	gluconeogenesis; GO:0006094, glycolysis; GO:0006096
Tdh2p	YJR009C	gluconeogenesis; GO:0006094, glycolysis; GO:0006096

Table A.1. Core proteins (cont.).

Standard Name	Systematic Name	GO Process Terms
TDH3	YGR192C	gluconeogenesis; GO:0006094, glycolysis ; GO:0006096
TPI1	YDR050C	gluconeogenesis; GO:0006094, glycolysis; GO:0006096 pentose-phosphate shunt; GO:0006098
SDL1	YIL168W	gluconeogenesis; GO:0006094
PGM1	YKL127W	glucose 1-phosphate metabolic process, GO:0019255 glucose 6-phosphate metabolic process, GO:0051156 glycogen biosynthetic process; GO:0005978 UDP-glucose metabolic process; GO:0006011
PGM2	YMR105C	glycogen biosynthetic process; GO:0005978 glucose 1-phosphate metabolic process, GO:0019255 UDP-glucose metabolic process, GO:0006011 glucose 6-phosphate metabolic process, GO:0051156
ADH1	YOL086C	ethanol biosynthetic process during fermentation; GO:0043458
ADH5	YBR145W	ethanol biosynthetic process during fermentation; GO:0043458
NDE1	YMR145C	glucose catabolic process to ethanol; GO:0019655
NDE2	YDL085W	glucose catabolic process to ethanol; GO:0019655
PDC1	YLR044C	glucose catabolic process to ethanol; GO:0019655
PDC2	YDR081C	glucose catabolic process to ethanol; GO:0019655
PDC5	YLR134W	glucose catabolic process to ethanol; GO:0019655
CDC1	YAL038W	glycolysis; GO:0006096
EMI2	YDR516C	glycolysis; GO:0006096
ERR1	YOR393W	glycolysis; GO:0006096
ERR2	YPL281C	glycolysis; GO:0006096
ERR3	YMR323W	glycolysis; GO:0006096
GLK1	YCL040W	glycolysis; GO:0006096
GPM2	YDL021W	glycolysis; GO:0006096
GPM3	YOL056W	glycolysis; GO:0006096
HXK1	YFR053C	glycolysis; GO:0006096
HXK2	YGL253W	glycolysis; GO:0006096
KGD1	YIL125W	glycolysis; GO:0006096
MDH1	YKL085W	glycolysis; GO:0006096
MDH3	YDL078C	glycolysis; GO:0006096
PDA1	YER178W	glycolysis; GO:0006096
PDB1	YBR221C	glycolysis; GO:0006096

Table A.1. Core proteins (cont.).

Standard Name	**Systematic Name**	**GO Process Terms**
PFK1	YGR240C	glycolysis; GO:0006096
PFK2	YMR205C	glycolysis; GO:0006096
PYK2	YOR347C	glycolysis; GO:0006096
GND1	YHR183W	pentose-phosphate shunt; GO:0006098 pentose-phosphate shunt oxidative branch; GO:0009051
GND2	YGR256W	pentose-phosphate shunt; GO:0006098, pentose-phosphate shunt oxidative branch; GO:0009051
NQM1	YGR043C	pentose-phosphate shunt; GO:0006098
RKI1	YOR095C	pentose-phosphate shunt; GO:0006098 pentose-phosphate shunt non-oxidative branch ; GO:0009052
RPE1	YJL121C	pentose-phosphate shunt; GO:0006098
SOL1	YNR034W	pentose-phosphate shunt; GO:0006098
SOL2	YCR073W-A	pentose-phosphate shunt; GO:0006098
SOL3	YHR163W	pentose-phosphate shunt; GO:0006098 pentose-phosphate shunt oxidative branch; GO:0009051
SOL4	YGR248W	pentose-phosphate shunt; GO:0006098 pentose-phosphate shunt oxidative branch; GO:0009051
TAL1	YLR354C	pentose-phosphate shunt; GO:0006098
TKL1	YPR074C	pentose-phosphate shunt; GO:0006098
TKL2	YBR117C	pentose-phosphate shunt; GO:0006098
ZWF1	YNL241C	pentose-phosphate shunt oxidative branch; GO:0009051
GCR1	YPL075W	positive regulation of glycolysis; GO:0045821
GCR2	YNL199C	positive regulation of glycolysis; GO:0045821
TYE7	YOR344C	positive regulation of glycolysis; GO:0045821
PFK26	YIL107C	regulation of glycolysis; GO:0006110
PFK27	YOL136C	regulation of glycolysis; GO:0006110
	YLR345W	regulation of glycolysis; GO:0006110
AAP1	YHR047C	glycogen metabolic process; GO:0005977
BMH1	YER177W	glycogen metabolic process; GO:0005977
BMH2	YDR099W	glycogen metabolic process; GO:0005977
GAC1	YOR178C	glycogen metabolic process; GO:0005977 glycogen biosynthetic process; GO:0005978
GIP2	YER054C	glycogen metabolic process; GO:0005977
GLC7	YER133W	glycogen metabolic process; GO:0005977

Table A.1. Core proteins (cont.).

Standard Name	**Systematic Name**	**GO Process Terms**
GLC8	YMR311C	glycogen metabolic process; GO:0005977 glycogen biosynthetic process; GO:0005978
GPH1	YPR160W	glycogen metabolic process; GO:0005977 glycogen catabolic process; GO:0005980
PCL10	YGL134W	glycogen metabolic process; GO:0005977 negative regulation of glycogen biosynthetic process; GO:0045719
PCL6	YER059W	regulation of glycogen biosynthetic process; regulation of glycogen catabolic process; GO:0005979 glycogen metabolic process; GO:0005977
PCL7	YIL050W	regulation of glycogen catabolic process; regulation of glycogen biosynthetic process; GO:0005979 glycogen metabolic process; GO:0005977
PCL8	YPL219W	glycogen metabolic process; GO:0005977 negative regulation of glycogen biosynthetic process; GO:0045719
PPG1	YNR032W	glycogen metabolic process; GO:0005977
REG1	YDR028C	glycogen metabolic process; GO:0005977
SHP1	YBL058W	glycogen metabolic process; GO:0005977
VMA10	YHR039C-A	glycogen metabolic process; GO:0005977
VMA22	YHR060W	glycogen metabolic process; GO:0005977
YPI1	YFR003C	glycogen metabolic process; GO:0005977
GDB1	YPR184W	glycogen biosynthetic process; GO:0005978, glycogen catabolic process; GO:0005980
GLC3	YEL011W	glycogen biosynthetic process; GO:0005978
GLG1	YKR058W	glycogen biosynthetic process; GO:0005978
GLG2	YJL137C	glycogen biosynthetic process; GO:0005978
GSY1	YFR015C	glycogen biosynthetic process; GO:0005978
GSY2	YLR258W	glycogen biosynthetic process; GO:0005978
PIG1	YLR273C	glycogen biosynthetic process; GO:0005978 regulation of glycogen biosynthetic process; GO:0005979
UGP1	YKL035W	glycogen biosynthetic process; GO:0005978 UDP-glucose metabolic process; GO:0006011
SGA1	YIL099W	glycogen catabolic process; GO:0005980
PHO85	YPL031C	negative regulation of glycogen biosynthetic process; GO:0045719
PIG2	YIL045W	regulation of glycogen biosynthetic process; GO:0005979

Table A.2. Reference GO terms pool.

Component	Function	Proceess	Function	Proceess	Function	Proceess	Proceess
GO:0005829	GO:0004743	GO:0006096	GO:0019904	GO:0070623	GO:0003983	GO:0000903	GO:0006559
GO:0001950	GO:0003824	GO:0006090	GO:0050815	GO:0010552	GO:0016779	GO:0006364	GO:0006569
GO:0005634	GO:0000287	GO:0031134	GO:0003688	GO:0034221	GO:0004332	GO:0051726	GO:0000949
GO:0005737	GO:0005524	GO:0034727	GO:0004842	GO:0051436	GO:0004614	GO:0007114	GO:0009083
GO:0034657	GO:0000166	GO:0000045	GO:0004679	GO:0000077	GO:0004612	GO:0007059	GO:0030003
GO:0019898	GO:0016301	GO:0005977	GO:0004672	GO:0006270	GO:0004611	GO:0006873	GO:0006874
GO:0031410	GO:0016740	GO:0030437	GO:0004674	GO:0006267	GO:0016831	GO:0006397	GO:0007089
GO:0016023	GO:0030955	GO:0043161	GO:0003844	GO:0007124	GO:0017076	GO:0006000	GO:0043666
GO:0030659	GO:0046872	GO:0016579	GO:0004553	GO:0007265	GO:0004737	GO:0032445	GO:0009966
GO:0016020	GO:0008599	GO:0006511	GO:0016757	GO:0001402	GO:0004331	GO:0006357	GO:0045821
GO:0005739	GO:0004221	GO:0045721	GO:0043169	GO:0030435	GO:0042578	GO:0007031	GO:0045013
GO:0005967	GO:0008233	GO:0016192	GO:0019787	GO:0007126	GO:0042132	GO:0006350	GO:0042542
GO:0042645	GO:0008234	GO:0051469	GO:0019888	GO:0042787	GO:0008047	GO:0045719	GO:0006113
GO:0005759	GO:0016787	GO:0006810	GO:0016538	GO:0017148	GO:0004345	GO:0000079	GO:0016558
GO:0005777	GO:0008270	GO:0015031	GO:0004721	GO:0045722	GO:0019170	GO:0006808	GO:0008615
GO:0005782	GO:0004843	GO:0007039	GO:0004722	GO:0007165	GO:0004751	GO:0008361	GO:0009052
GO:0005758	GO:0003674	GO:0006623	GO:0016624	GO:0042710	GO:0046914	GO:0046015	GO:0006744
GO:0000164	GO:0004802	GO:0008152	GO:0004864	GO:0001403	GO:0004497	GO:0006800	GO:0019538
GO:0031588	GO:0005488	GO:0006098	GO:0004373	GO:0006468	GO:0005198	GO:0006915	GO:0016311
GO:0005641	GO:0004022	GO:0000947	GO:0043539	GO:0007155	GO:0031072	GO:0009051	GO:0000083
GO:0000324	GO:0016491	GO:0055114	GO:0004801	GO:0006995	GO:0016791	GO:0032889	GO:0016242
GO:0031965	GO:0004347	GO:0043458	GO:0051287	GO:0001302	GO:0030528	GO:0015992	GO:0016239
GO:0000307	GO:0016853	GO:0006116	GO:0008943	GO:0015976	GO:0004693	GO:0006811	GO:0032880
GO:0000778	GO:0009374	GO:0006094	GO:0004365	GO:0006111	GO:0004871	GO:0007035	GO:0045936
GO:0005935	GO:0004736	GO:0006740	GO:0003872	GO:0045944	GO:0008184	GO:0006508	GO:0050849
GO:0005730	GO:0016874	GO:0006086	GO:0016829	GO:0005978	GO:0004645	GO:0070072	GO:0043433
GO:0005816	GO:0004739	GO:0006013	GO:0004634	GO:0006513	GO:0030170	GO:0034599	GO:0006974
GO:0001400	GO:0004340	GO:0046323	GO:0004616	GO:0051246	GO:0004135	GO:0031647	GO:0032878
GO:0005847	GO:0016773	GO:0006006	GO:0050661	GO:0000209	GO:0004134	GO:0006916	GO:0031505
GO:0043231	GO:0004396	GO:0005975	GO:0050662	GO:0016567		GO:0005980	GO:0009410
GO:0005886	GO:0004618	GO:0006409	GO:0016820	GO:0043687		GO:0000272	GO:0007047
GO:0009277	GO:0017057	GO:0008150	GO:0046961	GO:0006470		GO:0005976	
GO:0005811	GO:0004619	GO:0044262	GO:0004177	GO:0005981		GO:0006003	
GO:0005945	GO:0016868	GO:0006108	GO:0008237	GO:0005979		GO:0006110	
GO:0000015	GO:0016616	GO:0006097	GO:0051082	GO:0034501		GO:0006103	
GO:0016471	GO:0030060	GO:0006635	GO:0004339	GO:0051301		GO:0009069	
GO:0000221	GO:0016615	GO:0006735	GO:0016798	GO:0007067		GO:0006355	
GO:0042406	GO:0003954	GO:0006099	GO:0003873	GO:0007049		GO:0006011	
GO:0009898	GO:0008137	GO:0019655	GO:0004591	GO:0031297		GO:0005992	
GO:0005625	GO:0050660	GO:0006109	GO:0030976	GO:0030846		GO:0006078	
GO:0009353	GO:0004807	GO:0042149	GO:0003941	GO:0030847		GO:0009060	
GO:0005575	GO:0003677	GO:0000122	GO:0043565	GO:0007094		GO:0001300	
GO:0005635	GO:0016563	GO:0008610	GO:0003700	GO:0009408		GO:0019255	
GO:0005743	GO:0003704	GO:0006633	GO:0004750	GO:0000076		GO:0019388	
GO:0000790	GO:0003676	GO:0045449	GO:0008466	GO:0016576		GO:0051156	

Table A.3. Hubs of GN.

Standard Name	Descriptions	Degree	Betweenness
Rps3p	Protein component of the small (40S) ribosomal subunit, has apurinic/apyrimidinic (AP) endonuclease activity; essential for viability; has similarity to E. coli S3 and rat S3 ribosomal proteins	58	0.0284195
Brr2p	RNA-dependent ATPase RNA helicase (DEIH box); required for disruption of U4/U6 base-pairing in native snRNPs to activate the spliceosome for catalysis; homologous to human U5-200kD	51	0.0244188
Rpl3p	Protein component of the large (60S) ribosomal subunit, has similarity to E. coli L3 and rat L3 ribosomal proteins; involved in the replication and maintenance of killer double stranded RNA virus	48	0.043392
Rpl10p	Protein component of the large (60S) ribosomal subunit, responsible for joining the 40S and 60S subunits; regulates translation initiation; has similarity to rat L10 ribosomal protein and to members of the QM gene family	43	0.0284289
Act1p	Actin, structural protein involved in cell polarization, endocytosis, and other cytoskeletal functions	42	0.1439422
Rpo21p	RNA polymerase II largest subunit B220, part of central core; phosphorylation of C-terminal hepta peptide repeat domain regulates association with transcription and splicing factors; similar to bacterial beta-prime	40	0.1037428
Rpn10p	Non-ATPase base subunit of the 19S regulatory particle (RP) of the 26S proteasome; N-terminus plays a role in maintaining the structural integrity of the RP; binds selectively to polyubiquitin chains; homolog of the mammalian S5a protein	40	0.0446939
Rpt6p	One of six ATPases of the 19S regulatory particle of the 26S proteasome involved in the degradation of ubiquitinated substrates; bound by ubiquitin-protein ligases Ubr1p and Ufd4p; localized mainly to the nucleus throughout the cell cycle	39	0.0295309
Rpt3p	One of six ATPases of the 19S regulatory particle of the 26S proteasome involved in the degradation of ubiquitinated substrates; substrate of N-acetyltransferase B	38	0.013499
Spt15p	TATA-binding protein, general transcription factor that interacts with other factors to form the preinitiation complex at promoters, essential for viability	36	0.0497903
Pre2p	Beta 5 subunit of the 20S proteasome, responsible for the chymotryptic activity of the proteasome	35	0.0348183
Cdc31p	Calcium-binding component of the spindle pole body (SPB) half-bridge, required for SPB duplication in mitosis and meiosis II; homolog of mammalian centrin; binds multiubiquitinated proteins and is involved in proteasomal protein degradation	34	0.0359473
Taf14p	Subunit of TFIID, TFIIF, INO80, SWI/SNF, and NuA3 complexes, involved in RNA polymerase II transcription initiation and in chromatin modification; contains a YEATS domain	32	0.0872927

Table A.3. Hubs of GN (cont.).

Standard Name	Descriptions	Degree	Betweenness
Pre4p	Beta 7 subunit of the 20S proteasome	31	0.0192391
Prp8p	Component of the U4/U6-U5 snRNP complex, involved in the second catalytic step of splicing; mutations of human Prp8 cause retinitis pigmentosa	31	0.0106603
Cdc28p	Catalytic subunit of the main cell cycle cyclin-dependent kinase (CDK); alternately associates with G1 cyclins (CLNs) and G2/M cyclins (CLBs) which direct the CDK to specific substrates	28	0.0381115
Pre8p	Alpha 2 subunit of the 20S proteasome	27	0.0255031
Glc7p	Type 1 serine/threonine protein phosphatase catalytic subunit, involved in many processes (eg: glycogen metabolism, sporulation, mitosis); accumulates at mating projections by interaction with Afr1p; interacts with many regulatory subunits	27	0.0254919
Rpl5p	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	27	0.0102252
Ada2p	Transcription coactivator, component of the ADA and SAGA transcriptional adaptor/HAT (histone acetyltransferase) complexes	26	0.0381025
Gcn5p	Acetyltransferase, modifies N-terminal lysines on histones H2B and H3; acetylates Rsc4p, a subunit of the RSC chromatin-remodeling complex, altering replication stress tolerance; catalytic subunit of the ADA and SAGA histone acetyltransferase complexes; founding member of the Gcn5p-related N-acetyltransferase superfamily	26	0.0326245
Utp22p	Possible U3 snoRNP protein involved in maturation of pre-18S rRNA, based on computational analysis of large-scale protein-protein interaction data	25	0.0416002
Nop7p	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 G_0 and the initiation of cell proliferation	25	0.0344314
Sgs1p	Nucleolar DNA helicase of the RecQ family; involved in genome integrity maintenance; regulates chromosome synapsis and meiotic joint molecule/crossover formation; potential role as repressor of a subset of rapamycin responsive genes; rapidly lost in response to rapamycin in Rrd1p-dependent manner; similar to human BLM and WRN proteins implicated in Bloom and Werner syndromes	25	0.0316385
Enp1p	Protein associated with U3 and U14 snoRNAs, required for pre-rRNA processing and 40S ribosomal subunit synthesis; localized in the nucleus and concentrated in the nucleolus	25	0.0276559
Rad52p	Protein that stimulates strand exchange by facilitating Rad51p binding to single-stranded DNA; anneals complementary single-stranded DNA; involved in the repair of double-strand breaks in DNA during vegetative growth and meiosis	25	0.0198986

Table A.3. Hubs of GN (cont.).

Standard Name	**Descriptions**	**Degree**	**Betweenness**
Nas6p	Proteasome-interacting protein involved in the assembly of the base subcomplex of the 19S proteasomal regulatory particle (RP); ortholog of human oncoprotein gankyrin, which interacts with the Rb tumor suppressor and CDK4/6	24	0.0166815
Hhf1p	Histone H4, core histone protein required for chromatin assembly and chromosome function; one of two identical histone proteins; contributes to telomeric silencing; N-terminal domain involved in maintaining genomic integrity	23	0.0597488
Pol30p	Proliferating cell nuclear antigen (PCNA), functions as the sliding clamp for DNA polymerase delta; may function as a docking site for other proteins required for mitotic and meiotic chromosomal DNA replication and for DNA repair	23	0.0281504
Smd2p	Core Sm protein Sm D2; part of heteroheptameric complex (with Smb1p, Smd1p, Smd3p, Sme1p, Smx3p, and Smx2p) that is part of the spliceosomal U1, U2, U4, and U5 snRNPs; homolog of human Sm D2	23	0.0168906
Tra1p	Subunit of SAGA and NuA4 histone acetyltransferase complexes; interacts with acidic activators (e.g., Gal4p) which leads to transcription activation; similar to human TRRAP, which is a cofactor for c-Myc mediated oncogenic transformation	23	0.0128688
Egd1p	Subunit beta1 of the nascent polypeptide-associated complex (NAC) involved in protein targeting, associated with cytoplasmic ribosomes; enhances DNA binding of the Gal4p activator; homolog of human BTF3b	21	0.0290012
Sto1p	Large subunit of the nuclear mRNA cap-binding protein complex, interacts with Npl3p to carry nuclear poly(A)+ mRNA to cytoplasm; also involved in nuclear mRNA degradation and telomere maintenance; orthologous to mammalian CBP80	20	0.0617224
Kap95p	Karyopherin beta, forms a complex with Srp1p/Kap60p; interacts with nucleoporins to mediate nuclear import of NLS-containing cargo proteins via the nuclear pore complex; regulates PC biosynthesis; GDP-to-GTP exchange factor for Gsp1p	19	0.0578931
Skp1p	Evolutionarily conserved kinetochore protein that is part of multiple protein complexes, including the SCF ubiquitin ligase complex, the CBF3 complex that binds centromeric DNA, and the RAVE complex that regulates assembly of the V-ATPase	19	0.0506046
Rad23p	Protein with ubiquitin-like N terminus, subunit of Nuclear Excision Repair Factor 2 (NEF2) with Rad4p that binds damaged DNA; enhances protein deglycosylation activity of Png1p; also involved, with Rad4p, in ubiquitylated protein turnover	19	0.0387346
Orc2p	Subunit of the origin recognition complex, which directs DNA replication by binding to replication origins and is also involved in transcriptional silencing; interacts with Spp1p and with trimethylated histone H3; phosphorylated by Cdc28p	19	0.0205721

Table A.3. Hubs of GN (cont.).

Standard Name	Descriptions	Degree	Betweenness
Isw1p	Member of the imitation-switch (ISWI) class of ATP-dependent chromatin remodeling complexes; ATPase that forms a complex with Ioc2p and Ioc4p to regulate transcription elongation, and a complex with Ioc3p to repress transcription initiation	19	0.0160701
Rvb2p	Essential protein involved in transcription regulation; component of chromatin remodeling complexes; required for assembly and function of the INO80 complex; also referred to as reptin; member of the RUVB-like protein family	19	0.0141269
Sic1p	Inhibitor of Cdc28-Clb kinase complexes that controls G1/S phase transition, preventing premature S phase and ensuring genomic integrity; phosphorylation targets Sic1p for SCF(CDC4)-dependent turnover; functional homolog of mammalian Kip1	18	0.0755423
Sec22p	R-SNARE protein; assembles into SNARE complex with Bet1p, Bos1p and Sed5p; cycles between the ER and Golgi complex; involved in anterograde and retrograde transport between the ER and Golgi; synaptobrevin homolog	18	0.0205017
Nop1p	Nucleolar protein, component of the small subunit processome complex, which is required for processing of pre-18S rRNA; has similarity to mammalian fibrillarin	18	0.0171684

Table A.4. Hubs of GMN.

Standard Name	Description	Degree	Betweenness
Brr2p	RNA-dependent ATPase RNA helicase (DEIH box); required for disruption of U4/U6 base-pairing in native snRNPs to activate the spliceosome for catalysis; homologous to human U5-200kD	33	0.0330672
Rpt6p	One of six ATPases of the 19S regulatory particle of the 26S proteasome involved in the degradation of ubiquitinated substrates; bound by ubiquitin-protein ligases Ubr1p and Ufd4p; localized mainly to the nucleus throughout the cell cycle	28	0.0634041
Rpt3p	One of six ATPases of the 19S regulatory particle of the 26S proteasome involved in the degradation of ubiquitinated substrates; substrate of N-acetyltransferase B	28	0.0442547
Rpo21p	RNA polymerase II largest subunit B220, part of central core; phosphorylation of C-terminal heptapeptide repeat domain regulates association with transcription and splicing factors; similar to bacterial beta-prime	27	0.1604789
Pre2p	Beta 5 subunit of the 20S proteasome, responsible for the chymotryptic activity of the proteasome	25	0.0622
Cdc28p	Catalytic subunit of the main cell cycle cyclin-dependent kinase (CDK); alternately associates with G1 cyclins (CLNs) and G2/M cyclins (CLBs) which direct the CDK to specific substrates	24	0.0995309

Table A.4. Hubs of GMN (cont.).

Standard Name	Description	Degree	Betweenness
Act1p	Actin, structural protein involved in cell polarization, endocytosis, and other cytoskeletal functions	22	0.0568953
SPt15p	TATA-binding protein, general transcription factor that interacts with other factors to form the preinitiation complex at promoters, essential for viability	20	0.1189785
Pre8p	Alpha 2 subunit of the 20S proteasome	20	0.0509636
Gcn5p	Acetyltransferase, modifies N-terminal lysines on histones H2B and H3; acetylates Rsc4p, a subunit of the RSC chromatin-remodeling complex, altering replication stress tolerance; catalytic subunit of the ADA and SAGA histone acetyltransferase complexes; founding member of the Gcn5p-related N-acetyltransferase superfamily	18	0.1396684
Pol30p	Proliferating cell nuclear antigen (PCNA), functions as the sliding clamp for DNA polymerase delta; may function as a docking site for other proteins required for mitotic and meiotic chromosomal DNA replication and for DNA repair	18	0.035868
Sgs1p	Nucleolar DNA helicase of the RecQ family; involved in genome integrity maintenance; regulates chromosome synapsis and meiotic joint molecule/crossover formation; potential role as repressor of a subset of rapamycin responsive genes; rapidly lost in response to rapamycin in Rrd1p-dependent manner; similar to human BLM and WRN proteins implicated in Bloom and Werner syndromes	17	0.0894401
Yaf9p	Subunit of both the NuA4 histone H4 acetyltransferase complex and the SWR1 complex, may function to antagonize silencing near telomeres; interacts directly with Swc4p, has homology to human leukemogenic protein AF9, contains a YEATS domain	17	0.0299198
Snu114p	GTPase component of U5 snRNP involved in mRNA splicing via spliceosome; binds directly to U5 snRNA; proposed to be involved in conformational changes of the spliceosome; similarity to ribosomal translocation factor EF-	17	0.0290083
Taf6p	Subunit (60 kDa) of TFIID and SAGA complexes, involved in transcription initiation of RNA polymerase II and in chromatin modification, similar to histone H4	17	0.0251001
Tra1p	Subunit of SAGA and NuA4 histone acetyltransferase complexes; interacts with acidic activators (e.g., Gal4p) which leads to transcription activation; similar to human TRRAP, which is a cofactor for c-Myc mediated oncogenic transformation	16	0.0646695
Cdc6p	Essential ATP-binding protein required for DNA replication, component of the pre-replicative complex (pre-RC) which requires ORC to associate with chromatin and is in turn required for Mcm2-7p DNA association; homologous to S. pombe Cdc18p	16	0.0525317
Clb2p	B-type cyclin involved in cell cycle progression; activates Cdc28p to promote the transition from G2 to M phase; accumulates during G2 and M, then targeted via a destruction box motif for ubiquitin-mediated degradation by the proteasome	14	0.0558748

Table A.4. Hubs of GMN (cont.)

Standard Name	**Description**	**Degree**	**Betweenness**
Isw1p	Member of the imitation-switch (ISWI) class of ATP-dependent chromatin remodeling complexes; ATPase that forms a complex with Ioc2p and Ioc4p to regulate transcription elongation, and a complex with Ioc3p to repress transcription initiation	14	0.0331837
Rrp6p	Nuclear exosome exonuclease component; has 3'-5' exonuclease activity; involved in RNA processing, maturation, surveillance, degradation, tethering, and export	14	0.02791
Skp1p	Evolutionarily conserved kinetochore protein that is part of multiple protein complexes, including the SCF ubiquitin ligase complex, the CBF3 complex that binds centromeric DNA, and the RAVE complex that regulates assembly of the V-ATPase	13	0.0621387
Has1p	ATP-dependent RNA helicase; localizes to both the nuclear periphery and nucleolus; highly enriched in nuclear pore complex fractions; constituent of 66S pre-ribosomal particles	13	0.0607447
Srs2p	DNA helicase and DNA-dependent ATPase involved in DNA repair, needed for proper timing of commitment to meiotic recombination and transition from Meiosis I to II; blocks trinucleotide repeat expansion; affects genome stability	13	0.036849

APPENDIX B: INFORMATION ABOUT MODULES

Table B.1. Modules of GMN.

Module Name	Member Number	Members	p-Value	GO Process Term
M19.1	19	Pre4p Pre7p Rpt2p Rpt3p Pre1p Rpn11p Scl1p Rpt6p Pre9p Pup2p Pre3p Rpt1p Pre8p Pre5p Pre6p Pup1p Pre10p Pre2p Rpn7p	7.15E-30	Ubiquitin-dependent protein catabolic process
M16.1	16	Rpb4p Rpo21p Rpc53p Rpc17p Rpc25p Rpc34p Ret1p Rpc31p Rpc82p Rpo31p Rpb9p Tfg2p Tfg1p Rpb3p Rpb11p Rpb2p	8.17E-22	Transcription, DNA-dependent

Table B.1. Modules of GMN (cont.).

Module Name	Member Number	Members	p-Value	GO Process Term
M13.1	13	Tra1p Gcn5p Spt7p Taf5p Sgf29p Taf12p Spt3p Ada2p Sgf73p Taf6p Ubp8p Taf9p Sgf11p	1.04E-27	Protein amino acid acetylation
M11.1	11	Grr1p Cln3p Cdc28p Swi4p Swi6p Cln2p Cdc53p Cdc34p Skp1p Cdc4p Hrt1p	1.37E-17	G1/S transition of mitotic cell cycle
M11.2	11	Mnd2p Cdc27p Apc4p Swm1p Cdc26p Cdc23p Cdc16p Apc9p Apc2p Apc1p Apc5p	2.12E-31	Anaphase-promoting complex-

Table B.1. Modules of GMN (cont.).

Module Name	**Member Number**	**Members**	**p-Value**	**GO Process Term**
M10.1	10	Rrp6p Rrp43p Rrp42p Rrp45p Rrp46p Mtr3p Ski6p Rrp4p Csl4p Dis3p	1.24E-20	mRNA catabolic process
M9.1	9	Hat2p Cdc6p Orc2p Hat1p Orc6p Orc3p Orc1p Orc5p Orc4p	5.17E-18	Pre-replicative complex assembly
M8.1	8	Yth1p Glc7p Ref2p Cft1p Pti1p Mpe1p Pap1p Ysh1p	9.53E-15	RNA 3’-end processing
M8.2	8	Not5p Ccr4p Cdc39p Cdc36p Mot2p Not3p Caf40p Pop2p	1.43E-21	Nuclear-transcribed mRNA poly(A) tail shortening

Table B.1. Modules of GMN (cont.).

Module Name	**Member Number**	**Members**	**p-Value**	**GO Process Term**
M7.1	7	Rsc30p Rsc3p Rsc8p Sth1p Rsc58p Rsc2p Rsc1p	3.45E-15	RNA elongation from RNA polymerase II promoter
M7.2	7	Bni5p Gin4p Cdc10p Shs1p Cdc12p Cdc11p Cdc3p	6.67E-13	Cytokinesis
M7.3	7	Vps71p Swc3p Swc5p Vps72p Act1p Yaf9p Rvb2p	4.76E-14	Histone exchange
M7.4	7	Nkp2p Ame1p Iml3p Chl4p Mcm21p Mcm16p Okp1p	4.91E-12	Chromosome segregation
M7.5	7	Pho80p Pho4p Glg2p Gsy1p Glg1p Gsy2p Pho85p	3.61E-10	Glycogen biosynthetic process

Table B.1. Modules of GMN (cont.).

Module Name	Member Number	Members	p-Value	GO Process Term
M6.1	6	Rpt4p Sem1p Hsm3p Rpn14p Nas6p Rpn1p	4.75E-6	Proteolysis
M6.2	6	Elp4p Elp2p Iki1p Iki3p Elp6p Elp3p	6.03E-13	tRNA modification
M6.3	6	Tfc8p Tfc3p Tfc1p Tfc6p Tfc4p Tfc7p	1.71E-17	Transcription initiation from RNA polymerase III promoter
M6.4	6	Prp4p Snu114p Prp6p Prp3p Prp31p Snu66p	1.51E-11	Nuclear mRNA splicing, via spliceosome
M6.5	6	Tex1p Sub2p Hpr1p Thp2p Mft1p Rlr1p	1.48E-10	mRNA export from nucleus
M6.6	6	Ste4p Fus3p Ste7p Ste12p Kss1p Ste5p	2.28E-13	Invasive growth in response to glucose limitation

Table B.1. Modules of GMN (cont.).

Module Name	**Member Number**	**Members**	**p-Value**	**GO Process Term**
M6.7	6	Nug1p Erb1p Nsa1p Nop15p Has1p Nop2p	6365	rRNA processing
M6.8	6	Nup116p Nup145p Seh1p Sec13p Nup120 Nup100p	8.83E-15	Nuclear pore organization and biogenesis
M6.9	6	Lea1p Sme1p Clf1p Hsh49p Ysf3p Hsh155p	1.52E-10	RNA splicing
M6.10	6	Mad3p Cdc20p Cin8p Slk19p Mad1p Bub1p	1.10E-7	Mitosis
M6.11	6	Psy2p Hta2p Kap114p Hta1p Psy4p Pph3p	1.67E-4	Response to drug
M5.1	5	Rpa135p Rpa34p Rpa12p Rpa49p Rpa190	2.70E-10	Transcription from RNA polymerase I promoter

Table B.1. Modules of GMN (cont.).

Module Name	**Member Number**	**Members**	**p-Value**	**GO Process Term**
M5.2	5	Vma13p Tfp1p Vma8p Vma7p Vma10p	7.65E-13	pH reduction
M5.3	5	Esa1p Eaf1p Eaf5p Yng2p Eaf7p	3.54E-8	Protein aminoacid acetylation
M5.4	5	Ies3p Nhp10p Ies5p Ies1p Ino80p	2.08E-6	Chromatin remodeling
M5.5	5	Tpk2p Ras2p Tpk1p Bcy1p Tpk3p	4.27E-9	Ras protein signal transduction
M5.6	5	Sak1p Snf1p Gal83p Snf4p Sip1p	6.24E-9	Protein amino acid phosphorylation
M5.7	5	Cbc2p Mud1p Nam8p Snp1p Sto1p	9.88E-10	Nuclear mRNA splicing, via splicesome
M5.8	5	Cti6p Dep1p Rxt3p Pho23p Rpd3p	7.65E-13	Histone deacetylation

Table B.1. Modules of GMN (cont.).

Module Name	**Member Number**	**Members**	**p-Value**	**GO Process Term**
M4.1	4	Utp21p Utp22p Utp10p Utp4p	1.52E-6	rRNA processing
M4.2	4	Rix1p Nsa2p Sda1p Nog1p	1.51E-8	Ribosomal large subunit biogenesis
M4.3	4	Rfc4p Rfc5p Rfc2p Rfc3p	2.09E-11	Leading strand elongation
M4.4	4	Bud7p Chs6p Chs5p Bch1p	3.54E-7	Nucleotide-excision repair, DNA incision, 5’-to lesion
M4.5	4	Swi1p Snf2p Swp82p Swi3p	3.59E-7	Chromosome segregation
M4.6	4	Rad14p Apn1p Rad10p Rad1p	2.84E-9	Exocytosis
M4.7	4	Dls1p Itc1p Isw2p Dpb4p	6.82E-9	Negative regulation of meiosis
M4.8	4	Tps3p Tps1p Tps2p Tsl1p	7.33E-13	Trehalose biosynthetic process
M4.9	4	Hos2p Sif2p Snt1p Cpr1p	6.82E-9	Negative regulation of meiosis

Table B.1. Modules of GMN (cont.).

Module Name	**Member Number**	**Members**	**p-Value**	**GO Process Term**
M4.10	4	Sec8p Exo84p Sec5p Sec10p	2.84E-9	Exocytosis
M4.11	4	Nsl1p Mtw1p Dsn1p Nnf1p	3.59E-7	Chromosome segregation
M4.12	4	Sir1p Sir4p Sir2p Rap1p	5.85E-8	Heterochromatin formation
M4.13	4	Fks3p Rho1p Gsc2p Fks1p	4.87E-10	1,3-beta-glucan biosynthetic process
M4.14	4	Ptp3p Ptc1p Hog1p Ptp2p	1.53E-10	Osmosensory signaling pathway
M4.15	4	Vma4p Stv1p Vma2p Vph1p	2.22E-10	pH reduction
M4.16	4	Cef1p Prp46p Prp19p Cwc2p	1.53E-6	rRNA processing

REFERENCES

Albert, R., 2005, “Scale-free Networks in Cell Biology”, *Journal of Cell Science*, Vol. 118, pp. 4947-4957.

Allard, S., J. Y. Masson, and J. Cote, 2004, “Chromatin Remodeling and the Maintenance of Genome Integrity”, *Biochimica et Biophysica Acta*, Vol. 1677, No. 1-3, pp. 158-164.

Andress, E. J., R. Holic, M. J. Edelmann, B. M. Kessler, and V. P. Yu, 2011, “Dia2 Controls Transcription by Mediating Assembly of the RSC Complex”, *PLoS One*, Vol. 6, No. 6, pp. 1-9.

Arga K. Y., Z. İ. Önsan, B. K. Kırdar, K. Ülgen, and J. Nielsen, 2006, “Understanding Signalling in Yeast: Insights From Network Analysis”, *Biotechnology and Bioengineering*, Vol. 97, No. 5, pp. 1246-1258.

Arms, P. M. Ryan, E. T. Young, P. K. Parua, S. Ratnakumar, K. A. Braun, and K. M. Dombek, 2010, “14-3-3 (Bmh) Proteins Inhibit Transcription Activation by Adr1 through Direct Binding to Its Regulatory Domain”, *Molecular and Cellular Biology*, Vol.30, pp. 5273-5283.

Arnold, K. M., S. Lee, and J. M. Denu, 2011, “Processing Mechanism and Substrate Selectivity of the Core Nua4 Histone Acetyltransferase Complex”, *Biochemistry*, Vol. 50, No. 5, pp. 727-737.

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, and G. Sherlock, 2000, “Gene Ontology: Tool for the Unification of Biology. The Gene Ontology Consortium”, *Nature Genetics*, Vol. 25, No. 1, pp. 25-29.

Ashrafi, K., S. S. Lin, J. K. Manchester, and J. I. Gordon, 2000, “Sip2p and Its Partner Snf1p Kinase Affect Aging in Saccharomyces Cerevisiae”, *Genes and Development*, Vol. 14, No. 15, pp. 1872-1885.

Askree, S. H., T. Yehuda, S. Smolikov, R. Gurevich, J. Hawk, C. Coker, A. Krauskopf, M. Kupiec, and M. J. McEachern, 2004, “A Genome-wide Screen for Saccharomyces Cerevisiae Deletion Mutants that Affect Telomere Length”, *Proceedings of the National Academy of Sciences of the Unites States of America*, Vol. 101, No. 23, pp. 8658-8663.

Assenov, Y., F. Ramirez, S. Schelhorn, T. Lengauer, and M. Albrecht, 2007, “Computing Topological Parameters of Biological Networks”, *Bioinformatics*, Vol. 24, No. 2, pp. 282-284.

Auger, A., L. Galarneau, M. Altaf, A. Nourani, Y. Doyon, R. T. Utley, D. Cronier, S. Allard, and J. Cote, 2008, “Eaf1 Is the Platform for NuA4 Molecular Assembly That Evolutionarily Links Chromatin Acetylation to ATP-Dependent Exchange of Histone H2A Variants”, *Molecular and Cellular Biology*, Vol. 28, No. 8, p. 2257-2270.

Bader, G. D. and C. W. V. Hogue, 2003, “An Automated Method for Finding Molecular Complexes in Large Protein Interaction Networks”, *BMC Bioinformatics*, Vol. 4., No. 2, pp. 1-27.

Bai, C., P. Sen, K. Hofmann, L.Ma, M. Goebl, J. W. Harper, and S.J. Elledge, 1996, “SKP1 Connects Cell Cycle Regulators to the Ubiquitin Proteolysis Machinery through a Novel Motif, the F-Box”, *Cell*, Vol. 86, pp. 263-274.

Bao, Y. and X. Shen, 2007, “INO80 Subfamily of Chromatin Remodeling Complexes”, *Mutation Research*, Vol. 618, No. 1-2, pp. 18-29.

Barabasi, A. L. and Z. N. Oltvai, 2004, “Network Biology: Understanding the Cells Functional Organization”, *Nature Reviews Genetics*, Vol. 5, pp. 101-113.

Barral, Y., S. Jentsch, and C. Mann, 1995, “G1 Cyclin Turnover and Nutrient Uptake Are Controlled by a Common Pathway in Yeast”, *Genes and Development*, Vol. 9, No. 4, pp. 399-409.

Beggs, J. D., 2001, *Pre-mRNA Splicing*, http://dx.doi.org/10.1006/rwgn.2001.1137, Academic Press, accessed at November 2011.

Bhattacharya, S., L. Chen, J. R. Broach, and S. Powers, 1995, “Ras Membrane Targeting Is Essential for Glucose Signaling But Not for Viability in Yeast”, *Proceedings of the National Academy of Sciences of the Unites States of America*, Vol. 92, No. 7, pp. 2984-2988.

Blander G. and L. Guarente, 2004, “The Sir2 Family of Protein Deacetylases”, *Annual Review of Biochemistry*, Vol. 73, pp. 417-435.

Booher, R. N., R. J. Deshaies, and M. W. Kirschner, 1993, “Properties of Saccharomyces cerevisiae wee1 and Its Differential Regulation of p34CDC28 in Response to G1 and G2 Cyclins”, *The EMBO Journal*, Vol. 12, No. 9, pp. 3417-3426.

Breitkreutz, A., H. Choi, J. R. Sharom, L. Boucher, V. Neduva, B. Larsen, Z. Y. Lin, B. J. Breitkreutz, C. Stark, G. Liu, J. Ahn, D. Dewar-Darch, T. Reguly, X. Tang, R. Almeida, Z. S. Qin, T. Pawson, A. C. Gingras, A. I. Nesvizhskii, and M. Tyers, 2010, “A Global Protein Kinase and Phosphatase Interaction Network in Yeast”, *Science*, Vol. 328, No. 5981, pp. 1043-1046.

Breitkreutz, B. J., C. Stark, and M. Tyers, 2003, “Osprey: A Network Visualization System”, *Genome Biology*, Vol. 4, No. 3, pp.101-113.

Broach, J. R., 1991, “RAS Genes in Saccharomyces cerevisiae: Signal Transduction in Search of a Pathway”, *Trends in Genetics*, Vol. 7, No. 1, pp. 28-33.

Carrozza M. J., L. Florens, S. K. Swanson, W. J. Shia, S. Anderson, J. Yates, M. P. Washburn, J. L. Workman, 2005, “Stable Incorporation of Sequence Specific

Repressors Ash1 and Ume6 into the Rpd3L Complex", *Biochimica et Biophysica Acta,* Vol. 1731, No. 2, pp. 77-87.

Chang, C., F. Gonzalez, B. Rotherme, L. Sun, S. A. Johnston, and T. Kodadek, 2001, "The Gal4 Activation Domain Binds Sug2 Protein, a Proteasome Component", in Vivo and in Vitro, *Journal of Biological Chemistry,* Vol. 276, No. 3, pp. 30956-30963.

Cho, J. H., Y. K. Lee, and C. B. Chae, 2001, "The Modulation of the Biological Activities of Mitochondrial Histone Abf2p by Yeast PKA and its Possible Role in the Regulation of Mitochondrial DNA Content during Glucose Repression", *Biochimica et Biophysica Acta,* Vol. 1522, No. 3, pp. 175-186.

Clarke, A. S., J. E. Lowell, S. J. Jacobson, and L. Pillus, 1999, "Esa1p Is an Essential Histone Acetyltransferase Required for Cell Cycle Progression", *Molecular and Cellular Biology*, Vol. 19, No. 4, pp. 2515-2526.

Colombo, S., P. Ma, L. Cauwenberg, J. Winderickx, M. Crauwels, A. Teunissen, D. Nauwelaers, J. H.de Winde, M.Gorwa, D. Colavizza and J. M.Thevelein, 1998, "Involvement of Distinct G-proteins, Gpa2 and Ras, in Glucose- and Intracellular Acidification-induced cAMP Signalling in the Yeast Saccharomyces Cerevisiae", *The EMBO Journal,* Vol. 17, No. 12 pp. 3326-3341.

Conrad N. K., S. M. Wilson, E. J. Steinmetz, M. Patturajan, D. A. Brow, M. S. Swanson, and J. L. Corden, 2000, "A Yeast Heterogeneous Nuclear Ribonucleoprotein Complex Associated with RNA Polymerase II", *Genetics*, Vol. 154, No. 2, pp. 557-571.

Conaway, R. C., and J. W. Conaway, 2008, "The INO80 Chromatin Remodeling Complex in Transcription, Replication and Repair", *Trends in Biochemical Sciences,* Vol. 34 No. 2, pp. 71-77.

Cytoscape Consortium, 2001, *Cytoscape: An Open Source Platform for Complex Network Analysis and Visualization,*http://www.cytoscape.org, accessed at December 2010.

Das B, Z. Guo, P. Russo, P. Chartrand, and F. Sherman, 2000, “The Role of Nuclear Cap Binding Protein Cbc1p of Yeast in mRNA Termination and Degradation”, *Molecular and Cellular Biology,* Vol. 20, No. 8, pp. 2827-2838.

De Las Rivas, J. and C. Fontanillo, 2010, “Protein-Protein Interactions Essentials: Key Concepts to Building and Analyzing Interactome Networks”, *PLoS Computational Biology*, Vol. 6, No. 6, pp. 1-8.

De Marini, D. J., F. R. Papa, S. Swaminathan, D. Ursic, T. P. Rasmussen, M. R. Culbertson, and M. Hochstrasser, 1995, “The Yeast SEN3 Gene Encodes a Regulatory Subunit of the 26S Proteasome Complex Required for Ubiquitin-dependent Protein Degradation in Vivo”, *Molecular and Cellular Biology*, Vol. 15, No. 11, pp. 6311-6321.

DeVit, M. J. and M. Johnston, 1999, “The Nuclear Exportin Msn5 Is Required for Nuclear Export of the Mig1 Glucose Repressor of Saccharomyces cerevisiae”, *Current Biology,* Vol. 9, No. 21, pp. 1231-1241.

Dirick, L., T. Böhm, and K. Nasmyth, 1995, “Roles and Regulation of Cln-Cdc28 Kinases at the Start of the Cell Cycle of Saccharomyces cerevisiae”, *The EMBO Journal*, Vol. 14, No. 19, pp. 4803-4813.

Dlugai, S., S. Hippler, R. Wieczorke, and E. Boles, 2001, “Glucose Dependent and Independent Signaling Functions of the Yeast Glucose Sensor Snf3”, *FEBS Letters*, Vol. 505, pp. 389-392.

Dombek, K. M., N. Kacherovsky, and E. T. Young, 2004, “The Reg1-interacting Proteins, Bmh1, Bmh2, Ssb1, and Ssb2, Have Roles in Maintaining Glucose Repression in Saccharomyces Cerevisiae”, *Journal of Biological Chemistry,* Vol. 279, No. 37, pp. 39165-39174.

Dora E. G., N. Rudin, J. R. Martell, M. S. Esposito, R. M. Ramirez, 1999, “RPD3 (REC3) mutations Affect Mitotic Recombination in Saccharomyces cerevisiae”, *Current Genetic*s, Vol. 35, No. 2, pp. 68-76.

Doyon, Y. and J. Cote, 2004, “The Highly Conserved and Multifunctional Nua4 HAT Complex”, *Current Opinion in Genetics and Development,* Vol. 14, No. 2, pp. 147-154.

Doyon, Y., W.Selleck, W. S. Lane, S. Tan, and J. Cote, 2004, “Structural and Functional Conservation of the NuA4 Histone Acetyltransferase Complex from Yeast to Humans”, *Molecular and Cellular Biology*, Vol. 24, No. 5, pp. 1884-1896.

Elbing, K., R. R. McCartney, and M. C. Schmidt, 2006, “Purification and Characterization of the Three Snf1-activating Kinases of Saccharomyces cerevisiae”, *The Biochemical Journal*, Vol. 393, pp. 797-805.

Elsasser, S., R. R. Gali, M. Schwickart, C. N. Larsen, D. S. Leggett, B. Müller, M. T. Feng, F. Tübing, G. A. Dittmar, and D. Finley, 2002, “Proteasome Subunit Rpn1 Binds Ubiquitin-like Protein Domains”, *Nature Cell Biology*, Vol. 4, No. 9, pp. 725-730.

Enenkel, C., G. Blobel, and M. Rexach, 1995, “Identification of a Yeast Karyopherin Heterodimer that Targets Import Substrate to Mammalian Nuclear Pore Complexes”, *Journal of Biological Chemistry,* Vol. 270, No. 28, 16499-16502.

Enke, D. A., P. Kaldis, and M. J. Solomon, 2000, “Kinetic Analysis of the Cyclin-dependent Kinase-activating Kinase (Cak1p) from Budding Yeast”, *Journal of Biological Chemistry,* Vol. 275, No. 43, pp. 33267-33271.

Espelin, C.W., K. T. Simons, S.C. Harrison, and P. K. Sorger, 2003, “Binding of the Essential Saccharomyces cerevisiae Kinetochore Protein Ndc10p to CDEII”, *Molecular Biology of the Cell,* Vol. 14, pp. 4557-4568.

Fischer, M., W. Hilt, B. Richter-Ruoff, H. Gonenb, A. Ciechanoverb, and D. H. Wolfap, 1994, “The 26s Proteasome of The Yeast Saccharomyces Cerevisia”, *FEBS Letters*, Vol. 355, pp.69-75.

Farrants, A. O., 2008, “Chromatin Remodeling and Actin Organization”, *FEBS Letters*, Vol. 582, pp. 2041-2050.

Fedor Chaiken, M., R. J. Deschenes, and J. R. Broach, 1990, “SRV2, a Gene Required for RAS Activation of Adenylate cyclase in Yeast”, *Cell*, Vol. 61, No. 2, pp. 329-340.

Flick, K. M., N. Spielewoy, T. I. Kalashnikova, M. Guaderrama, Q. Zhu, H.C. Chang, and C. Wittenberg, 2003, “Grr1-dependent Inactivation of Mth1 Mediates Glucose-Induced Dissociation of Rgt1 from HXT Gene Promoters”, *Molecular Biology of the Cell*, Vol. 14, No. 8, pp. 3230-3241.

Floer, M. and G. Blobel, 1996, “The Nuclear Transport Factor Karyopherin Beta Binds Stoichiometrically to Ran-GTP and Inhibits the Ran GTPase Activating Protein”, *The Journal of Biological Chemistry*, Vol. 271, No. 10, pp. 5313-5316.

Forgac, M., 2007, “Vacuolar ATPases: Rotary Proton Pumps in Physiology and Pathophysiology”, *Molecular Cell Biology*, Vol. 8, pp. 916-929.

Fung G., 2001, “A Comprehensive Overview of Basic Clustering Algorithms”, Technical Report, University of Wisconsin, Madison, pp. 1-37.

Funakoshi1, M., X. Li, I. Velichutina, M. Hochstrasser, and H. Kobayashi, 2004, “Sem1, The yeast Ortholog of a Human BRCA2-Binding Protein, is a Component of the Proteasome Regulatory Particle that Enhances Proteasome Stability”, *Journal of Cell Science*, Vol. 117, pp. 6447-6457.

Gadal, O., S. Mariotte Labarre, S. Chedin, E. Quemeneur, C. Carles, A. Sentenac, and P. Thuriaux, 1997, “A34.5, a Nonessential Component of Yeast RNA Polymerase I, Cooperates with Subunit A14 and DNA Topoisomerase I to Produce a Functional rRNA Synthesis Machine”, *Molecular and Cellular Biology,* Vol. 17, No. 4, pp. 1787-1795.

Gari, E., T. Volpe, H. Wang, C. Gallego, B. Futcher, and M. Aldea, 2001, “Whi3 Binds the mRNA of the G1 Cyclin CLN3 to Modulate Cell Fate in Budding Yeast”, *Genes and Development,* Vol. 15, No. 21, pp. 2803-2808.

Glickman, M. H., D. M. Rubin, V. A. Fried, and D. Finley, 1998, “The Regulatory Particle of the Saccharomyces cerevisiae Proteasome”, *Molecular and Cellular Biology,* Vol. 18, No. 6, pp. 3149-3162.

Goes Neto, A., M. V. C. Diniz, L. B. L. Santos, S. T. R. Pinho, J. G. V. Miranda, T. P. Lobao, E. P. Borges, C. N. El-Hani, and R. F. S. Andrade, 2010, “Comparative Protein Analysis of the Chitin Metabolic Pathway in Extant Organisms: A Complex Network Approach”, *Biosystems,* Vol. 101, No. 1, pp. 59-66.

Goffeau, A., B. G. Barrell, H. Bussey, R. W. Davis, B. Dujon, H. Feldmann, F. Galibert, J. D. Hoheisel, C. Jacq, M. Johnston, E. J. Louis, H. W. Mewes, Y. Murakami, P. Philippsen, H. Tettelin, and S. G. Oliver, 1996, “Life with 6000 Genes”, *Science,* Vol. 274, No. 5287, pp. 546-567.

Goh, P. Y. and U. Surana, 1999, “Cdc4, a Protein Required for the Onset of S Phase, Serves an Essential Function during G(2)/M Transition in Saccharomyces cerevisiae”, *Molecular and Cellular Biology,* Vol. 19, No. 8, pp. 5512-5522.

Graham, L. A., K. J. Hill, and T. H. Stevens, 1994, “VMA7 Encodes a Novel 14-kDa Subunit of the Saccharomyces Cerevisiae Vacuolar H(+)-Atpase Complex”, *Journal of Biological Chemistry,* Vol. 269, No. 42, pp. 25974-25977.

Graham, L. A., K. J. Hill, and T. H. Stevens, 1995, “VMA8 Encodes A 32-Kda V1 Subunit of the Saccharomyces Cerevisiae Vacuolar H(+)-Atpase Required for Function and Assembly of the Enzyme Complex”, *Journal of Biological Chemistry,* Vol. 270, No. 25, pp. 15037-15044.

Greig, D. and J. Y. Leu, 2009, “Natural History of Budding Yeast”, *Current Biology,* Vol. 19, No. 19, pp. 886-890.

Grigorov, M. G., 2005, “Global Properties of Biological Networks”, *Drug Discovery Today,* Vol. 10, No. 5, pp. 365-372.

Groll, M., W. Heinemeyer, S. Jäger, T. Ullrich, M. Bochtler, D. H. Wolf, and R. Huber, 1999, “The Catalytic Sites of 20S Proteasomes and Their Role in Subunit Maturation: A Mutational and Crystallographic Study”, *Proceedings of the National Academy of Sciences of the Unites States of America,* Vol. 96, No. 20, pp. 10976-10983.

Grummt, I., 2003, “Life on a Planet of Its Own: Regulation of RNA Polymerase I Transcription in the Nucleolus”, *Genes and Development,* Vol. 17, No. 14, pp.1691-1702.

Guterman, A. and M. H. Glickman, 2004, “Complementary Roles for Rpn11 and Ubp6 in Deubiquitination and Proteolysis by the Proteasome”, *Journal of Biological Chemistry,* Vol. 279, No. 3, pp. 1729-1738.

Hallows, K. R., R. Alzamora, H. Li, F. Gong, C. Smolak, D. Neumann, and N. M. Pastor-Soler, 2009, “AMP-activated Protein Kinase Inhibits Alkaline pH- and PKA-induced Apical Vacuolar H(+)-ATPase Accumulation in Epididymal Clear Cells”, *American Journal of Physiology. Cell Physiology,* Vol. 296, No. 4, pp. 672-681.

Hardie, D. G., D. Carling, and M. Carlson, 1998, “The AMP-activated/SNF1 Protein Kinase Subfamily: Metabolic Sensors of the Eukaryotic Cell?”, *Annual Review of Biochemistry,* Vol. 67, pp. 821-855.

Haurie, V., M. Perrot, T. Mini, P. Jeno, F. Sagliocco, and H. Boucherie, 2001, “The Transcriptional Activator Cat8p Provides a Major Contribution to the Reprogramming of Carbon Metabolism during the Diauxic Shift in Saccharomyces Cerevisiae”, *Journal of Biological Chemistry* , Vol. 276, No. 1, pp. 76-85.

Hedbacker, K., and R. Townley, and M. Carlson, 2004, “Cyclic AMP-dependent Protein Kinase Regulates the Subcellular Localization of Snf1-Sip1 Protein Kinase”, *Molecular and Cellular Biology,* Vol. 24, No. 5, pp. 1836-1843.

Heinemeyer, W., N.Tröndle, G.Albrecht, and D.H.Wolf, 1994, “PRE5 and PRE6, The Last Missing Genes Encoding 20S Proteasome Subunits from Yeast Indication for a Set of 14 Different Subunits in the Eukaryotic Proteasome Core”, *Biochemistry,* Vol. 33, pp. 12229-12237.

Hirata, R., Y. Ohsumk, A. Nakano, H. Kawasaki, K. Suzuki, and Y. Anraku, 1990, “Molecular Structure of a Gene, VMA1, Encoding the Catalytic Subunit of H(+)-Translocating Adenosine Triphosphatase from Vacuolar Membranes of Saccharomyces Cerevisiae”, *Journal of Biological Chemistry,* Vol. 265, No. 12, pp. 6726-6733.

Ho, M. N., R. Hirata, N. Umemoto, Y. Ohya, A. Takatsuki, T. H. Stevens, and Y. Anraku, 1993, “VMA13 Encodes a 54-kDa Vacuolar H(+)-ATPase Subunit Required for Activity But Not Assembly of the Enzyme Complex in Saccharomyces Cerevisiae”, *Journal of Biological Chemistry,* Vol. 268, No. 24, pp. 18286-18292.

Hochstrasser, M., 1996, “Ubiquitin-dependent Protein Degradation”, *Annual Review of Genetics,* Vol. 30, pp. 405-439.

Hood, J. K. and P. A. Silver, 1998, “Cse1p Is Required for Export of Srp1p/importin-alpha from the Nucleus in Saccharomyces cerevisiae”, *The Journal of Biological Chemistry,* Vol. 273, No. 52, pp. 35142-35146.

Hoyt, M. A., 1997, “Eliminating All Obstacles: Regulated Proteolysis in the Eukaryotic Cell Cycle”, *Cell,* Vol. 91, No. 2, pp. 149-151.

Iacovella, M. G., C. N. Daly, J. S. Kelly, A. J. Michielsen, and R. K. Clyne, 2010, “Analysis of Polo-like Kinase Cdc5 in the Meiosis Recombination Checkpoint”, *Cell Cycle,* Vol. 9, pp. 1182-1193.

Jensen, L. J., M. Kuhn, M. Stark, S. Chaffron, C. Creevey, J. Muller, T. Doerks, P. Julien, A. Roth, M. Simonovic, P. Bork, and C. von Mering, 2008, “STRING 8-A Global View on Proteins and their Functional Interactions in 630 Organisms”, *Nucleic Acids Research*, Vol. 37, pp. 412-416.

Jiang, X., B. Liu, J. Jiang, H. Zhao, M. Fan, J. Zhang, Z. Fan, and T. Jiang, 2008, Modularity in the Genetic Disease-phenotype Network”, *FEBS Lett*ers, Vol. 582, No. 17, pp. 2549-2554.

Jin, J., Y. Cai, T. Yao, A. J. Gottschalk, L. Florens, S. K. Swanson, J. L. Gutierrez, M. K. Coleman, J. L. Workman, A. Mushegian, M. P. Washburn, R. C. Conaway, and J. W. Conaway, 2005, “A Mammalian Chromatin Remodeling Complex with Similarities to the Yeast INO80 Complex”, *Journal of Biological Chemistry*, Vol. 280, No. 50, pp. 41207-41212.

Johnson, E. S., I. Schwienhorst, R. J. Dohmen, and G. Blobel, 1997, “The Ubiquitin-like Protein Smt3p Is Activated for Conjugation to Other Proteins by an Aos1p/Uba2p Heterodimer”, *The EMBO Journal*, Vol. 16, No. 18, pp. 5509-5519.

Johnston, M. and J. H. Kim, 2005, “Glucose as a Hormone: Receptor-mediated Glucose Sensing in the Yeast Saccharomyces cerevisiae”, *Biochemical Society Transactions*, Vol. 33, pp. 247-252.

Jonsson, Z. O., S. K. Dhar, G. J. Narlikar, R. Auty, N. Wagle, D. Pellman, R. E. Pratt, R. Kingston, and A. Dutta, 2001, “Rvb1p and Rvb2p Are Essential Components of a Chromatin Remodeling Complex That Regulates Transcription of over 5% of Yeast Genes”, *Journal of Biological Chemistry*, Vol. 276, No. 19, pp. 16279-16288.

Kao, H. Y. and P. G. Siliciano, 1992, “The Yeast Homolog of the U1 snRNP Protein 70K Is Encoded by the SNP1 Gene”, *Nucleic Acids Research*, Vol. 20, No. 15, pp. 4009-4013.

Kaplun, L., Y. Ivantsiv, A. Bakhrat, and D. Raveh, 2003, “DNA Damage Response-mediated Degradation of Ho Endonuclease via the Ubiquitin System Involves Its Nuclear Export”, *Journal of Biological Chemistry*, Vol. 278, No. 49, pp. 48727-48734.

Kassir, Y., N. Adir, E. Boger-Nadjar, N. G. Raviv, I. Rubin-Bejerano, S. Sagee, and G. Shenhar, 2003, “Transcriptional Regulation of Meiosis in Budding Yeast”, *International of Review of Cytology*, Vol. 224, pp. 111-171.

Kataoka, T., S. Powers, C. McGill, O. Fasano, J. Strathern, J. Broach, and M. Wigler, 1984, “Genetic Analysis of Yeast RASI and RAS2 Genes”, *Cell*, Vol. 37, pp. 437-445.

Kataoka, T., D. Broek, and M. Wigler, 1985, “DNA Sequence and Characterization of the S. Cerevisiae Gene Encoding Adenylate Cyclase”, *Cell*, Vol. 43, pp. 493-505.

Kikuchi, J., Y. Iwafune, T.Akiyama, A. Okayama, H. Nakamura, N. Arakawa, Y. Kimura and H. Hirano, 2010, “Co- and Post-translational Modifications of the 26S Proteasome in Yeast”, *Proteomics*, Vol. 10, pp. 2769-2779.

Koch, C., T. Moll, M. Neuberg, H. Ahorn H, and K. Nasmyth, 1993, “A Role for the Transcription Factors Mbp1 and Swi4 in Progression from G1 to S Phase”, *Science*, Vol. 261, No. 5128, pp. 1551-1557.

Kraakman, L., K., Lemaire, P., Ma, A. W., Teunissen, M. C., Donaton, P., Van Dijck, J., Winderickx, J. H. de Winde, and J. M., Thevelein, 1999, “A Saccharomyces cerevisiae G-protein Coupled Receptor, Gpr1, Is Specifically Required for Glucose Activation of the cAMP Pathway during the Transition to Growth on Glucose”, *Molecular Microbiology.*, Vol. 32, No. 5, pp.1002-1012.

Kress, T.L., N. J. Krogan, and C. Guthrie, 2008, “A Single SR-like Protein, Npl3, Promotes Pre-mRNA Splicing in Budding Yeast”, *Molecular Cell*, Vol. 32, No. 5, pp. 727-734.

Krogan, N. J., K. Baetz, M. Keogh, N. Datta, C.Sawa, T. C. Y. Kwok, N. J. Thompson, M. G. Davey, J. Pootoolal, T.R. Hughes, A. Emili, S. Buratowski, P. Hieter, and J.F. Greenblatt, 2004, “Regulation of Chromosome Stability by the Histone H_2A Variant Htz1, the Swr1 Chromatin Remodeling Complex, and the Histone Acetyltransferase NuA_4”, *Proceedings of the National Academy of Sciences of the Unites States of America*, Vol. 101, No. 37, pp. 13513-13518.

Kuchin, S., V. K. Vyas, and M. Carlson, 2002, “Snf1 Protein Kinase and the Repressors Nrg1 and Nrg2 Regulate FLO11, Haploid Invasive Growth, and Diploid Pseudohyphal Differentiation”, *Molecular and Cellular Biology*, Vol. 22, No. 12, pp. 3994-4000.

Kuhn, C. D., S. R. Geiger, S. Baumli, M. Gartmann, J. Gerber, S. Jennebach, T. Mielke, H. Tschochner, R. Beckmann, and P. Cramer, 2007, “Functional Architecture of RNA Polymerase I”, *Cell*, Vol. 131, No. 7, pp. 1260-1272.

Kusano, K., M. E. Berres, and W. R. Engels, 1999, “Evolution of the RECQ Family of Helicases: A drosophila Homolog, Dmblm, Is Similar to the Human Bloom Syndrome Gene”, *Genetics*, Vol. 151, No. 3, pp. 1027-1039.

Lander, G. C., E. Estrin, M. E. Matyskiela, C. Bashore, E. Nogales, and A. Martin, 2012, “Complete Subunit Architecture of the Proteasome Regulatory Particle”, *Nature*, Vol. 482, No. 7384, pp. 186-191.

Le Tallec, B., M. B. Barrault, R. Guerois, T. Carre, and A. Peyroche, 2009, “Hsm3/S5b Participates in the Assembly Pathway of the 19S Regulatory Particle of the Proteasome”, *Molecular Cell*, Vol. 33, No. 3, pp. 389-399.

Lee, P., S. Paik, C. Shin, W. Huh, and J. Hahn, 2011, “Regulation of Yeast Yak1 Kinase by PKA and Autophosphorylation-dependent 14-3-3 Binding”, *Molecular Microbiology*, Vol. 79, No. 3, pp. 633-646.

Lesage, P., X. Yang, and M. Carlson, 1996, “Yeast SNF1 Protein Kinase Interacts with SIP4, a C6 Zinc Cluster Transcriptional Activator: a New Role for SNF1 in the Glucose Response”, *Molecular and Cellular Biology*, Vol. 16, No. 5, pp. 1921-1928.

Lewis, J. D., D. Görlich, and W. Mattaj, 1996, “A Yeast Cap Binding Protein Complex (yCBC) Acts at an Early Step in Pre-mRNA Splicing”, *Nucleic Acids Research,* Vol. 24, No. 1, pp. 3332-3336.

Liljelund, P., S. Mariotte, J. M. Buhler, and A. Sentenac, 1992, “Characterization and Mutagenesis of the Gene Encoding the A49 Subunit of RNA Polymerase A in Saccharomyces Cerevisiae”, *Proceedings of the National Academy of Sciences of the Unites States of America*, Vol. 89, No. 19, pp. 9302-305.

Lin, S. S., J. K. Manchester, and J. I. Gordon, 2003, “Sip2, an N-myristoylated Beta Subunit of Snf1 Kinase, Regulates Aging in Saccharomyces cerevisiae by Affecting Cellular Histone Kinase Activity, Recombination at rDNA Loci, and Silencing”, *Journal of Biological Chemistry,* Vol. 278, No. 15, pp. 13390-13397.

Lin, Y. Y., J. Y. Lu, J. Zhang, W. Walter, W. Dang, J. Wan, S. C. Tao, J. Qian, Y. Zhao, J. D. Boeke, S. L. Berger, and H. Zhu, 2009, “Protein Acetylation Microarray Reveals that Nua4 Controls Key Metabolic Target Regulating Gluconeogenesis”, *Cell*, Vol. 136, No. 6, pp. 1073-1084.

Liou, G. G., J. C. Tanny, R. G. Kruger, T. Walz, and D. Moazed, 2005, “Assembly of the SIR Complex and Its Regulation by O-acetyl-ADP-ribose, a Product of NAD-dependent Histone Deacetylation”, *Cell,* Vol. 121, No. 4, pp. 515-527.

Liu, Y., K. Nakatsukasa, M. Kotera, A. Kanada, T. Nishimura, T. Kishi, S. Mimura, and T. Kamura, 2011, “Non-SCF-type F-box Protein Roy1/Ymr258c Interacts with a Rab5-like GTPase Ypt52 and Inhibits Ypt52 Function”, *Molecular Biology of the Cell*, Vol. 22, No. 9, pp. 1575-1584.

Liu, Y., X. Xu, and M. Carlson, 2011, “Interaction of SNF1 Protein Kinase with Its Activating Kinase Sak1”, *Eukaryotic Cell,* Vol. 10, No. 3, pp. 313-319.

Loewith, R., J. S. Smith, M. Meijer, T. J. Williams, N. Bachman, J. D. Boeke, and D. Young, 2001, “Pho23 Is Associated with the Rpd3 Histone Deacetylase and Is Required for Its Normal Function in Regulation of Gene Expression and Silencing in Saccharomyces Cerevisiae”, The Journal of Biological Chemistry, Vol. 276, No. 26, pp. 24068-24074.

Lovas, A., A. Sos-Hegedus, A. Bimbo, and Z. Banfalvi, 2003, “Functional Diversity of Potato SNF1-related Kinases Tested in Saccharomyces cerevisiae”, *Gene,* Vol. 321, pp. 123-129.

MacKinnon, M. A., A. J. Curwin, G. J. Gaspard, A. B. Suraci, J. P. Fernandez-Murray, and C. R. McMaster, 2009, “The Kap60-Kap95 Karyopherin Complex Directly Regulates Phosphatidylcholine Synthesis”, *Journal of Biological Chemistry,* Vol. 284, No. 11, pp. 7376-7384.

Maere, S., K. Heymans, and M. Kuiper, 2005, “*BINGO*: a Cytoscape Plugin to Assess Overrepresentation of Gene Ontology Categories in Biological Networks”, *Bioinformatics,* Vol. 21, No. 16, pp. 3448-3449.

Maere, S., P. Van Dijck, and M. Kuiper, 2008, “Extracting Expression Modules from Perturbational Gene Expression Compendia”, *BMC Systems Biology,* 2:33.

Marrakchi, R., C. Chouchani, M. Cherif, A. Boudabbous, and D. Ramotar, 2011, “The Isomerase Rrd1 Mediates Rapid Loss of the Sgs1 Helicase in Response to Rapamycin”, *Biochemistry and Cell Biology,* Vol. 89, No. 3, pp. 332-340.

Mathias, N., S. L. Johnson, M. Winey, A.E. M. Adams, L. Goetsch, J. R. Pringle, B. Byers, and M. G. Goebl, , 1996, “Cdc53p Acts in Concert with Cdc4p and Cdc34p To Control the G1-to-S-Phase Transition and Identifies a Conserved Family of Proteins”, *Molecular and Cellular Biology,* Vol. 16, No. 12, pp. 6634-6643.

Max Planck Institut Informatik, 2011, *Cytoscape Plug-in: Network Analyzer*, http://med.bioinf.mpi-inf.mpg.de/netanalyzer, accessed at December 2010.

Memet, S., M. Gouy, C. Marck, A. Sentenac, and J. M. Buhler, 1988, "RPA190, the Gene Coding for the Largest Subunit of Yeast RNA Polymerase A", *Journal of Biological Chemistry,* Vol. 263, No. 6, pp. 2830-2839.

Mendenhall, M. D., 1998, "Cyclin-dependent Kinase Inhibitors of Saccharomyces cerevisiae and Schizosaccharomyces Pombe", *Current Topics in Microbiology and Immunology,* Vol. 227, pp. 1-24.

Michel, J. J., J. F. McCarville, and Y. Xiong, 2003, "A Role for Saccharomyces cerevisiae Cul8 Ubiquitin Ligase in Proper Anaphase Progression", *The Journal of Biological Chemistry,* Vol. 278, No. 25, pp. 22828-22837.

Mitchell, A. P., 1994, "Control of Meiotic Gene Expression in Saccharomyces cerevisiae", *Microbiological Reviews,* Vol. 58, No. 1, pp. 56-70.

Moreno, F., D. Ahuatzi, A. Riera, C. A. Palomino, and P. Herrero, 2005, "Glucose Sensing through the Hxk2-dependent Signaling Pathway", *Biochemical Society Transactions,* Vol. 33, pp. 265-268.

Morrison, A. J. and X. Shen, 2009, "Chromatin Remodeling beyond Transcription: the INO80 and SWR1 Complexes", *Nature Reviews. Molecular Cell Biology,* Vol. 10, No. 6, pp. 373-384.

Namy, O, G. Duchateau Nguyen, and J. P. Rousset, 2002, "Translational Readthrough of the PDE2 Stop Codon Modulates cAMP Levels in Saccharomyces cerevisiae", *Molecular Microbiology,* Vol. 43, No. 3, pp. 641-652.

Nasmyth, K. and L. Dirick, 1991, "The Role of SWI4 and SWI6 in the Activity of G1 Cyclins in Yeast", *Cell,* Vol. 66, No. 5, pp. 995-1013.

Nath, N., R. R. McCartney, and M. C. Schmidt, 2003, “Yeast Pak1 Kinase Associates with and Activates Snf1”, *Molecular and Cellular Biology,* Vol. 23, No. 11, pp. 3909-3917.

Neubauer, G, A. Gottschalk, P. Fabrizio, B. Seraphin, R. Lührmann, and M. Mann, 1997, “Identification of the Proteins of the Yeast U1 Small Nuclear Ribonucleoprotein Complex by Mass Spectrometry”, *Proceedings of the National Academy of Sciences of the Unites States of America*, Vol. 21, No. 2, pp. 385-390.

Niranjan, T., X. Guo, J. Victor, A. Lu, and J. P. Hirsch, 2007, “Kelch Repeat Protein Interacts with the Yeast Gα Subunit Gpa2p at a Site That Couples Receptor Binding to Guanine Nucleotide Exchange”, *Journal of Biological Chemistry* , Vol. 282, No. 33, pp. 24231-24238.

Nogi, Y., R. Yano, J. Dodd, C. Carles, and M. Nomura, 1993, “Gene RRN4 in Saccharomyces Cerevisiae Encodes the A12.2 Subunit of RNA Polymerase I and Is Essential Only at High Temperatures”, *Molecular Biology of the Cell,* Vol. 13, No. 1, pp. 114-122.

Ohta, T., J. J. Michel, A. J. Schottelius, and Y. Xiong, 1999, “ROC1, a Homolog of APC11, Represents a Family of Cullin Partners with an Associated Ubiquitin Ligase Activity”, *Molecular Cell,* Vol. 3, No. 4, pp. 535-541.

Osprey Community, 2001, *Osprey: Network Visualization System,* http://biodata.mshri.on.ca/osprey, accessed at December 2010.

Papa, F. R., A. Y. Amerik, and M. Hochstrasser, 1999, “Interaction of the Doa4 Deubiquitinating Enzyme with the Yeast 26S Proteasome”, *Molecular Biology of the Cell,* Vol. 10, No. 3, pp. 741-756.

Papamichos-Chronakis, M., T. Petrakis, E. Ktistaki, I. Topalidou, and D. Tzamarias, 2002, “Cti6, a PHD Domain Protein, Bridges the Cyc8-Tup1 Corepressor and the SAGA

Coactivator to Overcome Repression at GAL1", *Molecular Cell*, Vol. 9, No. 6, pp. 1297-1305.

Park, S. H., S. S. Koh, J. H. Chun, H. J. Hwang, and H. S. Kang, 1999, "Nrg1 Is a Transcriptional Repressor for Glucose Repression of STA1 Gene Expression in Saccharomyces cerevisiae", *Molecular and Cellular Biology*, Vol. 19, No. 3, pp. 2044-2050.

Peeters, T., W. Louwet, R. Gelade, D. Nauwelaers, J. M. Thevelein, and M. Versele, 2006, "Kelch-repeat Proteins Interacting with the Gα Protein Gpa2 Bypass Adenylate Cyclase for Direct Regulation of Protein Kinase A in Yeast", *Proceedings of the National Academy of Sciences of the Unites States of America*, Vol. 103, No. 35, pp. 13034-13039.

Pereira-Leal, J. B., A. J. Enright, and C. A. Ouzounis, 2004, "Detection of Functional Modules from Protein Interaction Networks", *Proteins*, Vol. 54, No. 1, pp. 49-57.

Peter, M. and I. Herskowitz, 1994, "Direct Inhibition of the Yeast Cyclin-dependent Kinase Cdc28-Cln by Far1", *Science*, Vol. 265, No. 5176, pp. 1228-1231.

Pratt, Z. L., B. J. Drehman, M. E. Miller, and S. D. Johnston, 2007, "Mutual Interdepence of MSI1 (CAC3) and YAK1 in Saccharomyces cerevisiae", *Journal of Molecular Biology*, Vol. 368, No. 1, pp. 30-43.

Qiu, Z. R., B. Schwer, and S. Shuman, 2011, "Defining the Mer1 and Nam8 Meiotic Splicing Regulons by cDNA Rescue", *RNA*, Vol. 17, No. 9, pp. 1648-1654.

Ramos, P. C., J. Höckendorff, E. S. Johnson, A. Varshavsky, and R. J. Dohmen, 1998, "Ump1p Is Required for Proper Maturation of the 20S Proteasome and Becomes Its Substrate upon Completion of the Assembly", *Cell*, Vol. 92, No. 4, pp. 489-499.

Ray, S. and A. Grove, 2009, “The Yeast High Mobility Group Protein HMO2, a Subunit of the Chromatin-Remodeling Complex INO80, Binds DNA Ends”, *Nucleic Acids Research,* Vol. 37, No. 19, pp. 6389-6399.

Rende, D., N. Baysal, and B. Kırdar, 2011, “A Novel Integrative Network Approach to Understand the Interplay between Cardiovascular Disease and Other Complex Disorders”, *Molecular Biosystems,* Vol. 7, pp. 2219.

Reynard, G. J., W. Reynolds, R. Verma, and R. J. Deshaies, 2000, “Cks1 Is Required for G(1) Cyclin-cyclin-dependent Kinase Activity in Budding Yeast”, *Molecular and Cellular Biology,* Vol. 20, No. 16, pp. 5858-5864.

Robzyk, K., J. Recht, and M. A. Osley, 2000, “Rad6-dependent Ubiquitination of Histone H2B in Yeast”, *Science,* Vol. 287, No. 5452, pp. 501-504.

Rolland, F., J. Winderickx, and J. M. Thevelein, 2002, “Glucose Sensing and Signaling Mechanisms in Yeast”, *FEMS Yeast Research,* Vol. 2, No. 2, pp. 183-201.

Rout, M. P., J. D. Aitchison, A. Suprapto, K. Hjertaas, Y. Zhao, and B. T. Chait, 2000, “The Yeast Nuclear Pore Complex: Composition, Architecture, and Transport Mechanism”, *Journal of Cell Biology,* Vol. 148, No. 4, pp. 635-651.

Rubin, D. M., M. H. Glickman, C. N. Larsen, S. Dhruvakumar, and D. Finley, 1998, “Active Site Mutants in the Six Regulatory Particle Atpases Reveal Multiple Roles for ATP in the Proteasome”, *The EMBO Journal,* Vol. 17, No. 17, pp. 4909-4919.

Rundlett, S. E., A. A. Carmen, R. Kobayashi, S. Bavykin, B. M. Turner, and M. Grunstein, 1996, “HDA1 and RPD3 are Members of Distinct Yeast Histone Deacetylase Complexes that Regulate Silencing and Transcription”, *Proceedings of the National Academy of Sciences of the Unites States of America*, Vol. 93, No. 25, pp. 14503-14508.

Russell, J. and J. C. Zomerdik, 2006, “The RNA Polymerase I Transcription Machinery”, *Biochemical Society Symposia*, Vol. 73, pp. 203-216.

Saeki, Y, A. Toh-E, T. Kudo, H. Kawamura, and K. Tanaka, 2009, “Multiple Proteasome-interacting Proteins Assist the Assembly of the Yeast 19S Regulatory Particle”, *Cell*, Vol. 137, No. 5, pp. 900-913.

Salgado-Garrido, J., E. Bragado-Nilsson, S. Kandels-Lewis, and B. Seraphin, 1999, “Sm and Sm-like Proteins Assemble in two Related Complexes of Deep Evolutionary Origin”, *The EMBO Journal*, Vol. 18, No. 12, pp. 3451-3462.

Santangelo, G. M., 2006, “Glucose Signaling in Saccharomyces Cerevisiae”, *Molecular and Cellular Biology*, Vol. 70, No. 1, p. 253-282.

Schimmel, J., K. M. Larsen, I. Matic, M. Van Hagen, J. Cox, M. Mann, J. S. Andersen, and A. C. Vertegaal, 2008, “The Ubiquitin-Proteasome System Is a Key Component of the SUMO-2/3 Cycle”, *Molecular and Cellular Proteomics*, Vol. 7, No. 11, pp. 2107-2122.

Schlossmann, J., R. Lill, W. Neupert, and D. A. Court, 1996, “Tom71, a Novel Homologue of the Mitochondrial Preprotein Receptor Tom70”, *Journal of Biological Chemistry*, Vol. 271, No. 30, pp. 17890-17895.

Schmidt, M. C., R. R. McCartney, X. Zhang, T. S. Tillman, H. Solimeo, S. Wölfl, C. Almonte, and S. C. Watkins, 1999, “Std1 and Mth1 Proteins Interact with the Glucose Sensors to Control Glucose-Regulated Gene Expression in Saccharomyces Cerevisiae”, *Molecular and Cellular Biology*, Vol. 19, No. 7, pp. 4561-4571.

Schüller, H. J., 2003, “Transcriptional Control of Non-fermentative Metabolism in the Yeast Saccharomyces cerevisiae”, *Current Genetics*, Vol. 43, No. 3, pp. 139-160.

Schwob, E. and K. Nasmyth, 1993, “CLB5 and CLB6, a New Pair of B Cyclins Involved in DNA Replication in Saccharomyces cerevisiae”, *Genes and Development,* Vol. 7, pp. 1160-1175.

Seol, J. H., A. Shevchenko, and R. J. Deshaies, 2001, “Skp1 Forms Multiple Protein Complexes, Including RAVE, a Regulator of V-ATPase Assembly”, *Nature Cell Biology,* Vol. 3, No. 4, pp. 384-391.

Shannon, P., A. Markiel, O. Ozier, N. S. Baliga, J. T. Wang, D. Ramage, N. Amin, B. Schwikowski, and T. Ideker, 2003, “Cytoscape: A Software Environment for Integrated Models of Biomolecular Interaction Networks”, *Genome Research,* Vol. 13, No. 11, pp. 2498-2504.

Sharan, R., I. Ulitsky and R. Shamir, 2007, “Network-based Prediction of Protein Function”, *Molecular Systems Biology,* Vol. 3:88.

Sharp J. A., D. C. Krawitz, K. A. Gardner, C. A. Fox, and P. D. Kaufman, 2003, “The Budding Yeast Silencing Protein Sir1 is a Functional Component of Centromeric Chromatin”, *Genes and Development,* Vol. 17, No. 19, pp. 2356-2361.

Shen, E. C., T. Stage-Zimmermann, P. Chui, and P. A. Silver, 2000, “7The Yeast mRNA-binding Protein Npl3p Interacts with the Cap-binding Complex”, *Journal of Biological Chemistry* , Vol. 275, No. 31, pp. 23718-23724.

Shen, X., G. Mizuguchi, A. Hamiche, and C. Wu, 2000, “A Chromatin Remodeling Complex Involved in Transcription and DNA Processing”, *Nature,* Vol. 406, No. 6795, pp. 541-544.

Shen, X., R. Ranallo, E. Choi, and C. Wu, 2003, “Involvement of Actin-Related Proteins in ATP-Dependent Chromatin Remodeling”, *Molecular Cell,* Vol. 12, pp. 147-155.

Shin, H., K. Tsuda, and B. Schölkopf, 2009, “Protein Functional Class Prediction with a Combined Graph”, *Expert System with Applications,* Vol. 36, pp. 3284-3292.

Sia, R. A., H. A. Herald, and D. J. Lew, 1997, "Cdc28 Tyrosine Phosphorylation and the Morphogenesis Checkpoint in Budding Yeast", *Molecular Biology of the Cell,* Vol. 7, No. 11, pp. 1657-1666.

Simonis, N., D. Gonze, C. Orsi, J. Van Helden, and S. J. Wodak, 2006, "Modularity of the Transcriptional Response of Protein Complexes in Yeast", *Journal of Molecular Biology*, Vol. 363, pp. 589-610.

Smardon, A. M. and P. M. Kane, 2007, "RAVE Is Essential for the Efficient Assembly of the C Subunit with the Vacuolar H(+)-ATPase", *Journal of Biological Chemistry*, Vol. 282, No. 36, pp. 26185-26194.

Sone, T., Y. Saeki, A. Tohe, and H. Yokosawa, 2004, "Sem1p is a Novel Subunit of the 26 S Proteasome from Saccharomyces Cerevisiae", *Journal of Biological Chemistry*, Vol. 279, No. 27, pp. 28807-28816.

Spiller, M. P., K. L. Boon, M. A. Reijns, and J. D. Beggs, 2007, "The Lsm2-8 Complex Determines Nuclear Localization of the Spliceosomal U6 snRNA", *Nucleic Acids Research*, Vol. 35, No. 3, pp. 923-929.

Spirin, V. and L. A. Mirny, 2003, "Protein Complexes and Functional Modules in Molecular Networks", *Proceedings of the National Academy of Sciences of the Unites States of America*, Vol. 100, No. 21, pp. 12123-12128.

Stark, C., B. J. Breitkreutz, A. Chatr-Aryamontri, L. Boucher, R. Oughtred, M. S. Livstone, J. Nixon, K. Van Auken, X. Wang, X. Shi, T. Reguly, J. M. Rust, A. Winter, K. Dolinski, and M. Tyers, 2011, "The BioGRID Interaction Database: 2011 Update", *Nucleic Acid Research*, Vol. 39, pp. 698-704.

Strich, R., R. T. Surosky, C. Steber, E. Dubois, F. Messenguy, and R. E. Esposito, 1994, "UME6 is a Key Regulator of Nitrogen Repression and Meiotic Development", *Genes and Development*, Vol. 8, No. 7, pp. 796-810.

String Database, 2000, *STRING: Known and Predicted Protein Protein Interactions*, http://string-db.org, accessed at September 2010.

Sudol, M. and K. F. Harvey, 2010, "Modularity in the Hippo Signaling Pathway", *Trends in Biochemical Sciences,* Vol. 35, No. 11, pp. 627-633.

Sun, S., Y. Zhao, Y. Jiao, Y. Yin, L. Cai, Y. Zhang, H. Lu, R. Chen, and D. Bu, 2006, "Faster and More Accurate Global Protein Function Assignment from Protein Interaction Networks Using the MFGO Algorithm", *FEBS Letters*, Vol. 580, No. 7, pp. 1891-1896.

Strunnikov, A. V., J. Kingsbury, and D. Koshland, 1995, "CEP3 Encodes a Centromere Protein of Saccharomyces Cerevisiae", *The Journal of Cell Biology*, Vol. 128, pp. 749-760.

Tabb, M. M., P. Tongaonkar, L. Vu, and M. Nomura, 2000, "Evidence for Separable Functions of Srp1p, the Yeast Homolog of Importin Alpha (Karyopherin alpha): Role for Srp1p and Sts1p in Protein Degradation", *Molecular and Cellular Biology*, Vol. 20, No. 16, pp. 6062-6073.

Tachibana, C., J. Y. Yoo, J. Tagne, N. Kacherovsky, T. I. Lee, and E. T. Young, 2005, "CombinGlobal Localization Analysis and Transcriptome Data Identify Genes That Are Directly Coregulated by Adr1 and Cat8", *Molecular and Cellular Biology*, Vol. 25, p. 2138-2146.

The Gene Ontology Consortium, 1999, *The Gene Ontology Database*, http://www.geneontology.org, accessed at September 2010.

Thevelein, J. M. and J. H. de Winde, 1999, "Novel Sensing Mechanisms and Targets for the cAMP-protein Kinase: A Pathway in the Yeast Saccharomyces cerevisiae", *Molecular Microbiology*, Vol. 33, No. 5, pp. 904-918.

Toda, T., S. Cameron, P. Sass, M. Zoller and M. Wigler, 1987, “Three Different Genes in S. cerevisiae Encode the Catalytic Subunits of the cAMP-dependent Protein Kinase”, *Cell,* Vol. 50, No. 2, pp. 277-287.

Traven, A., B. Jelicic, and M. Sopta, 2006, “Yeast Gal4: A Transcriptional Paradigm Revisited”, *EMBO Reports,* Vol. 7, No. 5, pp. 496-499.

Tu, J. and M. Carlson, 1995, “REG1 Binds To Protein Phosphatase Type 1 and Regulates Glucose Repression in Saccharomyces Cerevisiae”, *The EMBO Journal,* Vol. 14, No. 23, pp. 5939-5946.

Udugama, M., A. Sabri, and B. Bartholomew, 2010, “The INO80 ATP-Dependent Chromatin Remodeling Complex Is a Nucleosome Spacing Factor”, *Molecular and Cellular Biology,* Vol. 31, No. 4, p. 662-673.

Uetz, P., L. Giot, G. Cagney, T. A. Mansfield, R. S. Judson, J. R. Knight, D. Lockshon, V. Narayan, M. Srinivasan, P. Pochart, A. Qureshi-Emili, Y. Li, B. Godwin, D. Conover, T. Kalbfleisch, G. Vijayadamodar, M. Yang, M. Johnston, S. Fields, and J. M. Rothberg, 2000, “A Comprehensive Analysis of Protein Protein Interactions in Saccharomyces Cerevisiae”, *Nature,* Vol. 403, pp. 623-627.

Urquiza, J. M., I. Rojas, H. Pomares, L. J. Herrera, J. Ortega and A. Prieto, 2011, “Method for Prediction of Protein-Protein Interactions in Yeast Using Genomics/Proteomics Information and Feature Selection”, *Neurocomputing,* Vol .74, pp. 2683-2690.

Vincent, O., R. Townley, and S. Kuchin, 2001, “Subcellular Localization of the Snf1 Kinase is Regulated by Specific b Subunits and a Novel Glucose Signaling Mechanism”, *Genes and Development,* Vol. 15, pp. 1104-1114.

Xiaojia, B. and D. Jian, 2010, “Serine214 of Ras2p Plays a role in the Feedback Regulation of the Ras-cAMP Pathway in the Yeast Saccharomyces cerevisiae”, *FEBS Letters,* Vol. 584, No. 11, pp. 2333-2338.

Xie Y. and A. Varshavsky, 2001, “RPN4 is a Ligand, Substrate, and Transcriptional Regulator of the 26S Proteasome: A Negative Feedback Circuit”, *Proceedings of the National Academy of Sciences of the Unites States of America*, Vol. 98, No. 6, pp. 3056-3061.

Vallabhajosyula, R. R., D. Chakravarti, S. Lutfeali, A. Ray, and A. Raval, 2009, “Identifying Hubs in Protein Interaction Networks”, *PLoS One,* Vol. 4, No. 4.

Vanoni, M., R. L. Rossi, L. Querin, V. Zinzalla, and L. Alberghina, 2005, “Glucose modulation of cell size in yeast”, *Biochemical Society Transactions,* Vol. 33, pp. 294-296.

Vasiljeva L., M. Kim, H. Mutschler, S. Buratowski, and A. Meinhart, 2008, ”The Nrd1-Nab3-Sen1 Termination Complex Interacts with the Ser5-phosphorylated RNA Polymerase II C-terminal Domain”, *Nature Structural and Molecular Biology,* Vol. 15, No. 8, pp. 795-804.

Vazquez, A., A. Flammini, A. Maritan, and A. Vespignani, 2003, “Global Protein Function Prediction from Protein-Protein Interaction Networks”, *Nature Biotechnology,* Vol. 21, No. 6, pp. 697-700.

Vermeirssen, V., M. I. Barrasa, C. A. Hidalgo, J.A. Babon, R. Sequerra, L. Doucette-Stamm, A. L. Barabasi, and A. J. Walhout, 2007, “Transcription Factor Modularity in a Gene-centered C. elegans Core Neuronal Protein DNA Interaction Network”, *Genome Research*, Vol. 17, No. 7, pp. 1061-1071.

Vincent, O. and M. Carlson, 1999, “Gal83 Mediates the Interaction of the Snf1 Kinase Complex with the Transcription Activator Sip4”, *The EMBO Journal,* Vol. 18, No. 23, pp. 6672-6681.

Vincent, O., R. Townley, S. Kuchin, and M. Carlson, 2001, “Subcellular Localization of the Snf1 Kinase Is Regulated by Specific Beta Subunits and a Novel Glucose Signaling Mechanism”, *Genes and Development*, Vol. 15, No. 9, pp. 1104-1114.

Vyas, V. K., S. Kuchin, and M. Carlson, 2001, “Interaction of the Repressors Nrg1 and Nrg2 with the Snf1 Protein Kinase in Saccharomyces cerevisiae”, *Genetics,* Vol. 158, No. 2, pp. 563-572.

Wang, C., C. Skinner, E. Easlon, and S. Lin, 2009, “Deleting the 14-3-3 Protein Bmh1 Extends Life Span in Saccharomyces Cerevisiae by Increasing Stress Response”, *Genetics*, Vol. 183, pp. 1373-1384.

Westergaard, S. L., A. P. Oliveira, C. Bro, L. Olsson, and J. Nielsen, 2007, “A Systems Biology Approach to Study Glucose Repression in the Yeast Saccharomyces cerevisiae”, *Biotechnology and Bioengineering*, Vol. 96, No. 1, pp. 134-145.

Whistler, J. L. and J. Rine, 1997, “Ras2 and Ras1 Protein Phosphorylation in Saccharomyces cerevisiae”, *Journal of Biological Chemistry*, Vol. 272, No. 30, pp. 18790-18800.

Wiatrowski, H. A., B. J. W. van Denderen, C. D. Berkey, B. E. Kemp, D. Stapleton, and M. Carlson, 2005, “Mutations in the Gal83 Glycogen-Binding Domain Activate the Snf1/Gal83 Kinase Pathway by a Glycogen-Independent Mechanism”, *Molecular and Cellular Biology*, Vol. 24, No. 1, p. 352-361.

Yano, R. and M. Nomura, 1991, “Suppressor Analysis of Temperature-sensitive Mutations of the Largest Subunit of RNA polymerase I in Saccharomyces Cerevisiae: A Suppressor Gene Encodes the Second Largest Subunit of RNA polymerase I”, *Molecular and Cellular Biology*, Vol. 11, No. 2, pp. 754-64.

Yoon, J., A. Blumer, and K. Lee, 2006, “An Algorithm for Modularity Analysis of Directed and Weighted Biological Networks based on Edge-betweenness Centrality”, *Bioinformatics*, Vol. 22, No. 24, pp. 3106-3108.

Young, E. T., N. Kacherovsky, and K. V. Riper, 2002, “Snf1 Protein Kinase Regulates Adr1 Binding to Chromatin but Not Transcription Activation”, *Journal of Biological Chemistry*, Vol. 277, No. 41, pp. 38095-38103.

Zachariae, W. and K. Nasmyth, 1999, “Whose End Is Destruction: Cell Division and the Anaphase-promoting Complex”, *Genes and Development*, Vol. 13, No. 16, pp. 2039-2058.

Zhang, J., L. Olsson, and J. Nielsen, 2010, “The b-subunits of the Snf1 Kinase in Saccharomyces Cerevisiae, Gal83 and Sip2, but not Sip1, Are Redundant in Glucose Depression and Regulation of Sterol Biosynthesis”, *Molecular Microbiology*, Vol.77, pp. 371-383.

Zhang, S., R. Wang, and X. Zhang, 2007, “Identification of Overlapping Community Structure in Complex Networks Using Fuzzy c-means Clustering”, *Physica A: Statistical Mechanics and Its Applications*, Vol. 374, No. 1, pp. 483-490.

Zhou, J., B. O. Zhou, B. A. Lenzmeier, and J. Q. Zhou, 2009, “Histone Deacetylase Rpd3 Antagonizes Sir2-dependent Silent Chromatin Propagation”, *Nucleic Acids Research*, Vol. 37, No. 11, pp. 3699-3713.

Zhou, H. and F. Winston, 2001, “NRG1 Is Required for Glucose Repression of the SUC2 and GAL Genes of Saccharomyces Cerevisiae”, *BMC Genetics*, Vol. 2:5.

Zhu, M., L. Gao, Z. Guo, Y. Li, D. Wang, J. Wang, and C. Wang, 2007, “Globally Predicting Protein Functions Based on Co-expressed Protein Protein Interaction Networks and Ontology Taxonomy Similarities”, *Gene*, Vol. 391, No. 1-2, pp. 113-119.