INTEGRATION OF INTERACTOME AND TRANSCRIPTOME DATA TO REVEAL GLUCOSE REPRESSION PATHWAY IN *SACCHAROMYCES CEREVISIAE*

by

Dicle Hasdemir

B.S. in Chemical Engineering, Boğaziçi University, 2006

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

2008

ACKNOWLEDGEMENT

I want to thank my supervisor Prof. Betül Kırdar for guiding me all through this research and encouraging me to extract biologically meaningful results out of my numbers and graphs. Without her, I would definitely be lost and hopelessly trying to plan what to do next.

I also would like to thank my co-supervisor Prof. Z. İlsen Önsan for her kind support during my graduate study and Prof. Jens Nielsen who initiated this project and kindly provided the data.

I feel deeply grateful to Ayça, Esra, Güray, Duygu, Mustafa and Seda for spending hours together with me not only for academic reasons but also for friendship. I am also grateful to Yalçın for his kind contribution to this thesis no matter what the problem is and to all the biosystems laboratory residents with whom I worked together anywhere, any time.

I believe I could not write this thesis without the priceless support of my love Yunus, my precious friend Ceylan and my brother Barış who were always with me.

Lastly, I want to express my deepest gratitude to my mother Pürnaz and my father Yılmaz who were always my dearest friends regardless of all the kilometers that has been separating us for years.

ABSTRACT

INTEGRATION OF INTERACTOME AND TRANSCRIPTOME DATA TO REVEAL GLUCOSE REPRESSION PATHWAY IN *SACCHAROMYCES CEREVISIAE*

Due to glucose's role as a primary signaling molecule, a change in glucose level immediately stimulates certain elements of the glucose sensing and signaling pathway. This stimulation of the pathway activates the associated regulatory network. In yeast *S. cerevisiae*, some of the proteins involved in the glucose repression pathway are clearly identified. A key element of the glucose repression pathway; Snf1 kinase is composed of a catalytic subunit Snf1p, a regulatory subunit Snf4p and a scaffolding β-subunit (Sip1p, Sip2p or Gal83p). Due to its key role, it is important to predict how the signaling mechanism is affected in the absence of this complex. In this study the genome wide expression data from SNF1Δ, SNF4Δ and SNF1ΔSNF4Δ mutants were mapped onto the pre-processed protein-protein interaction network of yeast. All the linear paths starting from the glucose importer protein Hxk2p and ending at the transcriptional repressor of gluconeogenic genes Mig1p were identified via NetSearch algorithm. The linear paths were then scored based on the expression levels of the genes encoding the proteins involved. Thus, the key signaling elements of the significantly responsive linear paths were determined. The identified proteins were predicted to be the most likely elements to play a role in the glucose repression mechanism of yeast in the absence of the components of the Snf1 complex. In addition to that, the transcriptome data from these mutants were integrated with the regulatory network of yeast. As the final objective, a complete view of the glucose repression mechanism in the absence of the two constituent proteins of the Snf1 complex was predicted and putative elements with similar roles to this important complex were proposed.

ÖZET

SACCHAROMYCES CEREVISIAE 'DE GLİKOZ BASKILANMA YOLAĞINI AÇIĞA ÇIKARMAK İÇİN ETKİLEŞİM VE ANLATIM VERİSİNİN BÜTÜNLEŞTİRİLMESİ

Glikozun birincil bir sinyal molekülü olarak rolü sebebiyle, glikoz düzeyinde bir değişiklik hemen, glikoz algılama ve sinyal ileti yolağının belirli öğelerini uyarır. Yolağının bu uyarımı, ilgili düzenleyici ağı etkin hale getirir. *S. cerevisiae*'de, glikozun baskılanması sinyal ileti yolağında görev alan proteinlerin bir kısmı saptanmıştır. Glikozun baskılanması sinyal ileti yolağının anahtar bir öğesi olan Snf1 kinazı, katalitik bir alt birim olan Snf1 proteini, bir düzenleyici alt birim olan Snf4 proteini ve bir iskele malzemesi olan β alt biriminden (Sip1 proteini, Sip2 proteini veya Gal83 proteini) oluşmaktadır. Kompleksin anahtar rolü sebebiyle, bu kompleksin yokluğunda sinyal ileti yolağının nasıl etkilendiğini öngörmek önemlidir. Bu çalışmada, Snf1, Snf4 ve Snf1Snf4 delesyon mutantlarından alınan genom bazında anlatım verisi, mayanın önceden işleme tabi tutulan protein-protein etkileşim ağı üzerine işlendi. Hxk2 glikoz almaç işlevi gören proteininden başlayan ve glukoneojenik genlerin yazılımsal baskılayıcısı olan Mig1 proteinde biten bütün doğrusal yollar NetSearch algoritması yardımıyla belirlendi. Saptanan bu doğrusal yollar, içerdikleri proteinleri sentezleyen genlerin anlatım düzeyleri temel alınarak skorlandı. Böylece, genlerin delesyonuna en belirgin biçimde yanıt veren doğrusal yolların anahtar konumundaki öğeleri saptandı. Bu proteinler, Snf1 kompleksinin bileşenlerinin yokluğunda mayanın glikoz bastırılma mekanizmasında rol oynaması olasılığı yüksek öğeler olarak öngörüldü. Diğer taraftan, bu mutantlardan alınan anlatım verisi, mayanın yazılımsal kontrol ağı ile bütünleştirildi. Son hedef olarak, Snf1 protein kompleksinin iki öğesi olan proteinlerin birisi ya da her ikisinin de yokluğunda glikoz bastırılma mekanizması modellendi ve bu önemli kompleksle benzer rol taşıması olası görülen proteinler aday olarak saptandı.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF SYMBOLS/ABBREVIATIONS

BioGrid	The general repository for interaction datasets
GO	Gene ontology
SGD	Saccharomyces genome database
TF	Transcription factor
TRANSFAC	The gene regulation databases
YEASTRACT	Yeast search for transcriptional regulators and consensus tracking
Z_n	Z score of a gene
Z_{path}	Average path score of a path
Δ	Deletion
μ_κ	Mean of the z scores of randomly chosen groups of k genes
σ_k	Standard deviation of the z scores of randomly chosen groups of k genes

1. INTRODUCTION

1.1. Signaling Pathways in *S. cerevisiae*

1.1.1. Signal Transduction

Signal transduction is the main process which initiates a cell's metabolic, morphologic, genetic responses to environmental factors such as growth factors, hormones, nutrients, osmolarity and other chemical and tactile stimuli (Steffen *et al*., 2002). A signal initiated by one of these stimulators is transduced typically by the elements of the associated signaling pathway. The main elements of a typical signaling pathway are as follows:

- Cell surface receptor molecules which are activated by external signals
- Intracellular receptors which capture internal signals
- Signaling pathway components (kinases, phosphatases)
- The transcription factor which determines the cell's response

The steps of the signal flow consist of:

- The binding of a ligand
- The subsequent phosphorylation of an intracellular enzyme
- Amplification and passage of the signal
- The resultant change in cellular function

The final aim of modeling the signal transduction pathways of different organisms is the quantitative prediction of the dynamics of the signal transduction mechanism. But this is partly achieved only for causal and stoichiometric reconstructions of smaller cellular signaling network reconstructions, due to difficulties in obtaining kinetic data (Bailey, 2001) and the unclear structure of the pathways.

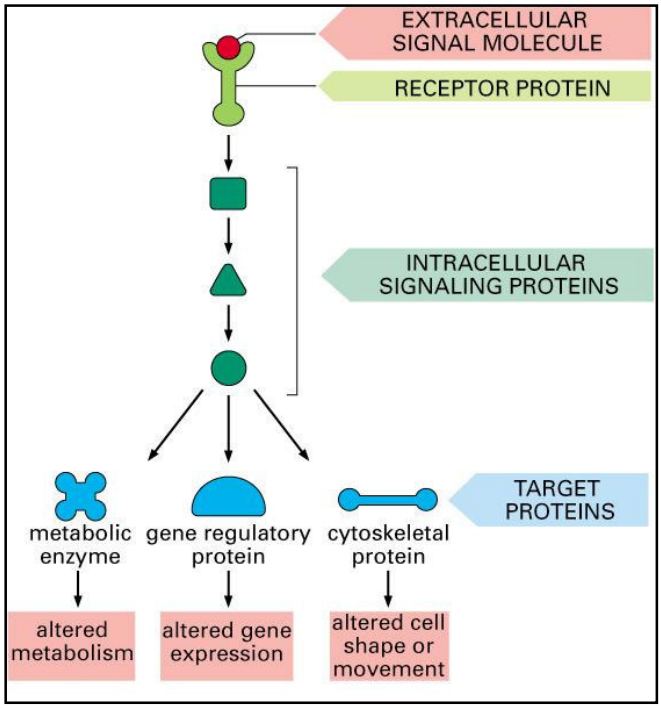

Figure 1.1. Simplified view of an intracellular signaling pathway (Alberts *et al.*, 2002)

Unlike the metabolic and the transcriptional regulatory networks, the notions of motifs and modules are still developing for signaling networks. This relatively slow progress in revealing the signaling pathways are partly due to the extent of crosstalk between different signaling pathways. Via a reductionistic approach, these pathways are formulated in a linear fashion whereas the reality is much more complex. So the first step is to reconstruct to the network structure of these pathways; the question is which proteins interact with each other to transduce the signal to the final element (Arga *et al.*, 2007).

Traditionally, the discovery of the components of signaling pathways is carried out by the use of gene knockouts and epistasis analysis. These methods worked well for describing the specific linear pathways in detail but the knowledge of complex signaling networks and their interactions are still incomplete. Nowadays, the availability of high throughput genomic data makes it easier to see the overall behavior of the signaling elements in a cell and thus to capture a more comprehensive picture of the interactions between these elements (Steffen *et al.*, 2002).

1.1.2. Reconstructing Signaling Networks

Signaling networks can be reconstructed in three ways as demonstrated in Figure 1.2. The first way is to reconstruct the highly connected nodes in the network. This type of reconstruction covers the comprehensive list of compounds and reactions which are associated with a given protein or signaling molecule. The second approach involves the formation of linear paths starting from the receptor protein and ending at the signaling output. For example, the signaling output might be the activation or deactivation of a transcription factor that induces the expression of target genes. The third approach identifies the signaling modules which are the sets of proteins working together under certain conditions (Papin *et al*., 2005).

All types of reconstruction methods require detailed collection of known protein-protein interactions from various sources. Experimental techniques are still being developed for the identification of the interactions between proteins and thus the components of the cellular signaling networks. Most recently, high throughput genome sequencing provided information for the prediction of putative elements of the signaling mechanism. Recent progress in ChIP-chip assays allows extensive investigation on the control of expression of these elements. The most widely used technique for discovering protein-protein interactions involves yeast two hybrid assays. As the imaging technologies are being more advance, FRET (Fluorescence Resonance Energy Transfer) is also being commonly used for deciphering signaling mechanisms as well as the GFP (Green Fluorescent Protein) enabled localization of proteins. In Figure 1.3, a scheme for the process of data integration and reconstruction of a signaling network is shown.

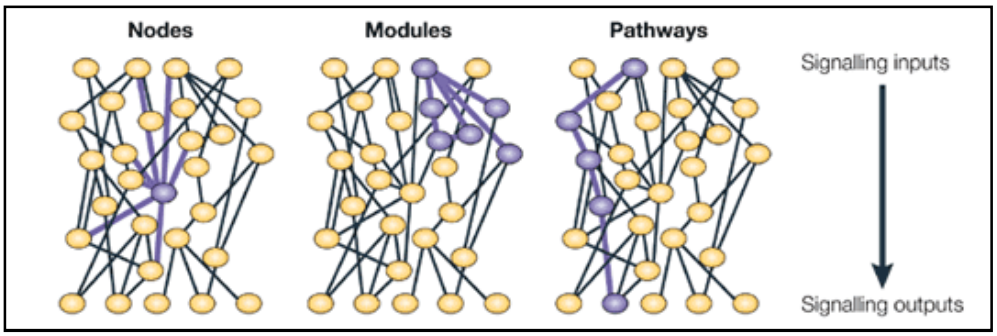

Figure 1.2. Three approaches of signaling network reconstruction (Papin *et al*., 2005)

1.1.3. Selective Permissibility Algorithm (SPA)

Recently, extensive knowledge on protein-protein, protein-DNA interactions, gene expression profiles and protein characterization is gained via experimental studies reporting high-throughput functional genomics data. Extensive use of computers and the internet make it easy to collect these data in databases. But a key challenge is to integrate these data in a proper way and predict the biological models behind the scene. Due to high noise in these data, it is a hard challenge to filter it. For a proper filtering of data, one must be able to filter out false positives and false negatives as much as possible to extract network structures from the genome-wide data.

The previous signaling pathway reconstruction schemes (Begley *et al*., 2002; Haugen *et al*., 2004; Ideker *et al*., 2001; Liu and Zhao, 2004; Steffen *et al*., 2002*)* are mostly focused on integration of protein–protein interaction data with microarray gene expression profiles and do not take into account the GO annotations. The Selective Permissibility Algorithm (SPA) is a convenient tool for the preliminary network reconstruction from candidate proteins for signal transduction mechanisms and it makes use of the GO annotations of the proteins which makes the method different from previously developed methods. (Arga *et al*., 2007)

The SPA algorithm has two distinct elements; the input proteins and the selection criterion. In the first step, all interactions of the input proteins are extracted from interactome databases. In the next step, the relevance of these interacting proteins is tested through a selection criterion which employs Gene Ontology (GO) Annotations. At this step, the interacting proteins are either accepted or rejected. The cycle continues with the accepted proteins as input proteins until no new interacting partners were identified.

The GO annotations which were used as the selection criterion are the annotations used to describe the cellular communication and signal transduction mechanism category in the MIPS Comprehensive Yeast Genome Database (Guldener *et al*., 2005) in terms of component, process and function. This collection covers 376 annotations. If all three GO Annotations (component, process and function) are present in this list of 376 annotations, the candidate protein is included in the preliminary network.

1.2. Glucose Metabolism in *Saccharomyces cerevisiae*

Glucose is the basic carbon source for most living organisms and therefore these organisms evolved sophisticated mechanisms for sensing it and responding to it. This is especially obvious in the yeast *S. cerevisiae*, where these regulatory mechanisms determine the distinctive fermentative metabolism of yeast, a lifestyle it shares with many kinds of tumor cells (Johnston and Kim, 2005).

S. cerevisiae has an unusual lifestyle; it prefers to ferment rather than oxidize glucose, even when oxygen is abundant. Glucose is metabolized through glycolysis to pyruvate, via two distinct chains of metabolic reactions initiated by two distinct ways of signal transduction. These two distinct pathways are named as glucose repression and glucose induction pathways (See Figure 1.4).These paths are linked to each other via numerous elements.

In the presence of oxygen, most organisms convert pyruvate to carbon dioxide and water (via the tricarboxylic acid cycle), generating many ATPs via oxidative phosphorylation. Only when oxygen becomes limiting do most cells resort to fermentation, because it yields only 2 ATPs per molecule of glucose. The propensity of *S.cerevisiae* to carry out aerobic fermentation is called the 'Crabtree effect' (Johnston and Kim, 2005).

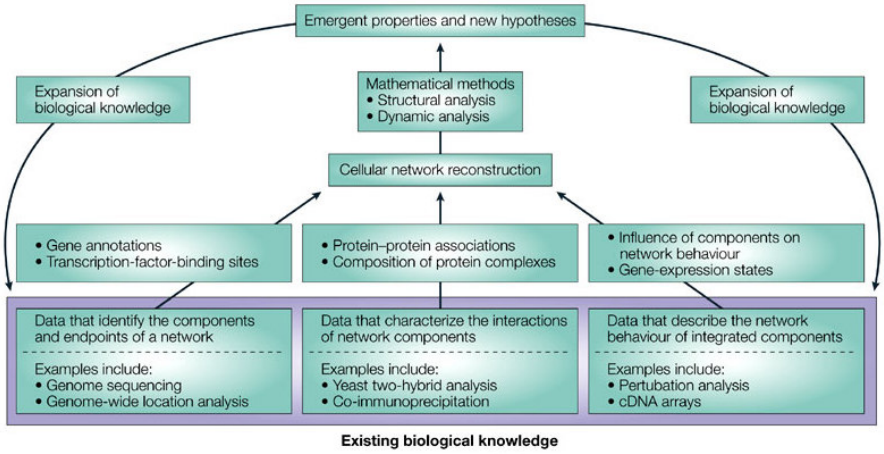

Figure 1.3. Steps of signaling network reconstruction process (Papin *et al*., 2005)

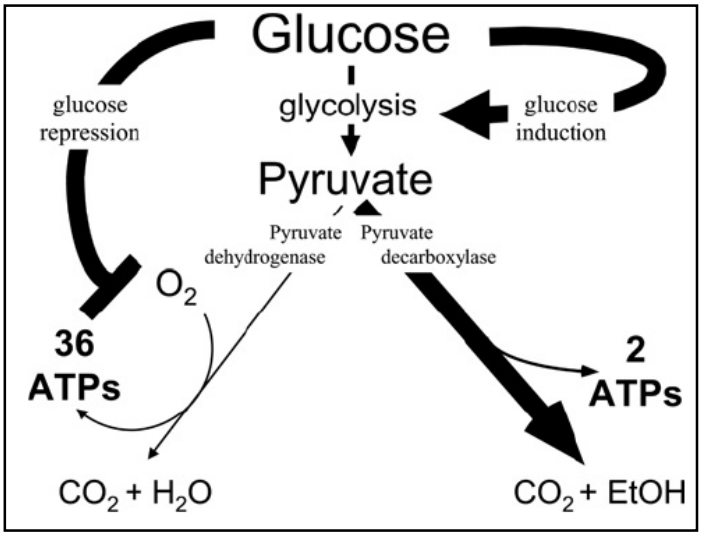

Figure 1.4. Glucose metabolism in yeast (Johnston and Kim, 2005)

1.2.1. Glucose Induction Pathway in *Saccharomyces cerevisiae*

S. cerevisiae's sophisticated glucose-sensing and signaling mechanisms enable it to sense a wide range of glucose concentrations and utilize glucose efficiently (Kim *et al*., 2006). *S. cerevisiae* is able to detect glucose levels of the extracellular medium and to generate intracellular signals that result in an adequate cellular response to variations in the glucose medium composition. Many of these responses involve changes in gene expression at the transcriptional, post-transcriptional, translational and post-translational levels (Arga *et al*., 2007). The majority of these changes occur at the level of mRNA transcription by the repression or activation of the transcription of many genes that encode enzymes implicated in carbon metabolism (Moreno *et al*., 2005).

The starting elements of the glucose induction pathway are two yeast membrane glucose sensors Snf3p and Rgt2p with very long cytoplasmic tails that sense low and high concentrations of glucose, respectively. After binding of glucose to these sensors, a conformational change occurs and thus the membrane bound casein kinase, Yck1p becomes activated. This step is followed by the phosphorylation of Mth1p and Std1p by Yck1p. These two proteins interact with the cytoplasmic tails in glucose grown cells. Later, phosphorylated Mth1p and Std1p are ubiquitinated by the Grr1p. These ubiquitinated proteins are recognized by the proteasomal machinery and degraded. The above mentioned protein-protein interactions occur in the cytoplasm of the cell, and resultins in a signal

being sent to the nucleus, where the transcription factor Rgt1p becomes hyperphosphorylated. When Rgt1p becomes hyperphosphorylated, it is believed that an unknown protein activates the transcription of the HXT genes. On the other hand, when glucose is not available in the medium, the Mth1p and Std1p move to the nucleus, where they interact with the active Rgt1 protein and repress the transcription of HXTs (Raghevendran *et al*., 2005).

The importance of glucose to yeast is underscored by the large number of hexose transporters it possesses. The presence of 18 such proteins has been reported to date. At least six of these (Hxt1, 2, 3, 4, 6 and 7) are known to be glucose transporters. Each of the glucose transporters has a different affinity and capacity for glucose. Hxt1 is a low-affinity, high-capacity transporter that is more useful when glucose is abundant (1 per cent); Hxt2 is a high-affinity, low-capacity glucose transporter that is necessary when glucose is scarce (0.2 per cent). The other hexose transporters have evolved for dealing with different concentrations of glucose and under different conditions. Thus, yeast cells have many different glucose transporters to deal with many different conditions.

1.2.2. Glucose Repression Pathway in *Saccharomyces cerevisiae*

Saccharomyces cerevisiae and many other types of yeast may thrive on a variety of carbon sources, but glucose and fructose are the preferred ones. When one of these sugars is present, the enzymes required for the utilization of alternative carbon sources are synthesized at low rates or not at all (Gancedo, 1998).

The role of the Mig1 complex in the glucose repression pathway is to repress the gluconeogenic genes when glucose is abundant in the media. When glucose is readily available in the medium; Mig1p (a DNA-binding protein with a zinc finger domain) is dephoshorylated by the Glc7 phosphatase via its regulatory subunit Reg1 and is inactive. There is also evidence showing the contribution of the metabolic enzyme; Hxk2p's to positive regulation of this step. It is however not clear if this contribution occurs in a direct or indirect manner. Active, unphosphorylated Mig1p is located in the nucleus where it interacts with the co-repressors Ssn6 and Tup1 and binds to the promoters of various genes and represses their expression. These genes involve mostly genes which encode enzymes

of the tricarboxylic acid (TCA) cycle, electron transport chain, alternative carbon sources consumption and gluconeogenesis.

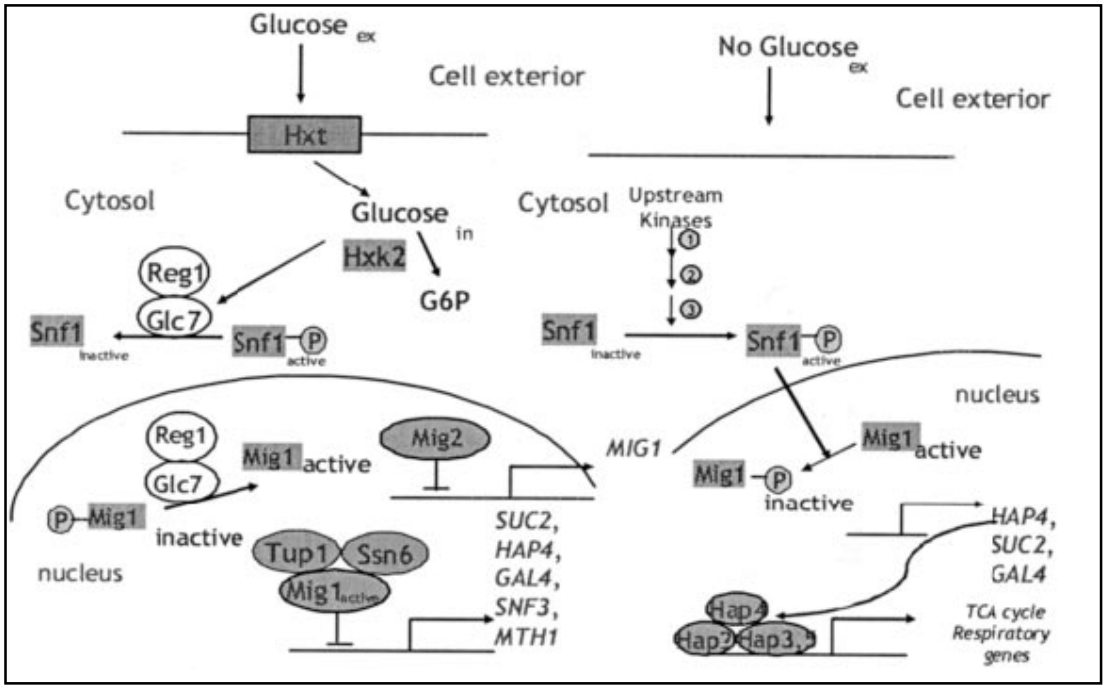

Figure 1.5. Glucose Repression Pathway in *S. cerevisiae* (Raghevendran *et al*., 2005)

When glucose becomes depleted, Mig1p must be repressed for the derepression of its target genes which will encode the enzymes needed for oxidative growth, like in glucose-limited chemostat. In order to achieve this, Snf1p is phosphorylated and thus activated by one of the three upstream kinases. Active Snf1 phosphorylates the transcriptional repressor Mig1p which moves from the nucleus to the cytoplasm and its interaction with the co-repressors Ssn6p and Tup1p are inhibited. The action of the protein kinase Snf1 is regulated via Reg1p by Glc7p, which dephosphorylates Snf1p. The dephosphorylation of Snf1p deactivates it under glucose-excess conditions.

The glucose induction and repression pathways in yeast have to function in close association with each other for efficient utilization of both glucose and other carbon sources. There are some proteins which refer to a high degree of cross-talk between two pathways. Some of the genes in induction pathway, such as SNF3 and MTH1, which directly regulate the expression of hexose transporters, are regulated by the repression

pathway. On the other hand, there is also evidence indicating that genes whose protein products function in the repression pathway like MIG2, MIG3 and HXK2, are regulated by the induction pathway (Raghevendran *et al*., 2005).

1.2.3. Snf1 Complex

The SNF1 gene (also called CAT1 or CCR1) is required for the derepression of genes repressed by glucose (Carlson *et al*.,1981; Ciriacy, 1975; Zimmermann *et al.*, 1971). SNF1 gene encodes the first member of a family of protein kinases to be identified across several organisms. Its mammalian homolog is the catalytic subunit of AMPK (AMP-activated protein kinase)(Carling *et al.*, 1994; Mitchelhill *et al.*, 1994; Recht *et al.*, 1996).

In yeast cells, the Snf1 protein is found to be associated with other proteins: Snf4p, Sip1p, Sip2p, and Gal83p (Celenza *et al.,* 1989; Yang *et al.,* 1992; Yang *et al.,* 1994). The earliest experiments with Snf4 mutants showed that these deletion mutants were unable to derepress the genes controlled by catabolite repression (Entian *et al.,* 1982; Neigeborn *et al*., 1984), while Sip1, Sip2, and Gal83 mutants and even the Sip1Δ Sip2 Δ Gal83 Δ triple mutant have no defect in the expression of a GAL gene (Erickson *et al.,* 1993) or the SUC2 gene (Yang *et al.,* 1994). However all the mutants including the Snf4 mutant are viable. So there is a possibility that Snf1p is still active in some manner without the Snf4 protein which is defined as the activating sub-unit of the complex. Under these conditions, it is important to predict the proteins involved in the alternative mechanism when one or more constituents of the complex are absent.

Snf1p participates, together with Snf4p, in a family of complexes containing either Sip1p, Sip2p, or Gal83p as mentioned above. These complexes may be, at least to some extent, functionally equivalent, but the observation that a Sip1ΔSip2ΔGal83Δ mutant is able to maintain normal regulation of a set of glucose-repressed genes, would suggest the existence of still unidentified additional regulatory proteins.

The subunits of the complex are homologs of the α–β-γ–subunits of the AMPK in mammalian cells. SNF4 gene encodes the γ–subunit of the Snf1 complex whereas the β-subunits may be the Sip1p, Sip2p and Gal83p proteins interchangeably. Although the

redundancy of the β-subunit is not understood well, it was reported that the β-subunit can mediate subcellular localization of the Snf1 protein kinase under derepressing conditions, for example, low glucose (Vincent *et al.*, 2001) and therefore may indicate distinct functions of specific β-subunits in signaling transduction and regulation.

It is shown that Snf1p and Snf4p do not interact in two-hybrid experiments in glucose grown yeast cells (Jiang *et al.,* 1996) whereas Sip1p and Sip2p interact with Snf1p (Yang *et al.,* 1994). Even in the absence of Snf4p, coimmunoprecipitation of Sip1p and Sip2p with Snf1p is observed. When the glucose concentration is low in the medium, meaning that the cells are grown in glucose limited medium; a direct interaction between Snf1p and Snf4p is observed. (Jiang *et al.*, 1996). This observation helps to recognize the Snf4 protein as the activating subunit while Sip1, Sip2, or Gal83 proteins are predicted to act as bridging proteins between Snf1p and Snf4p.

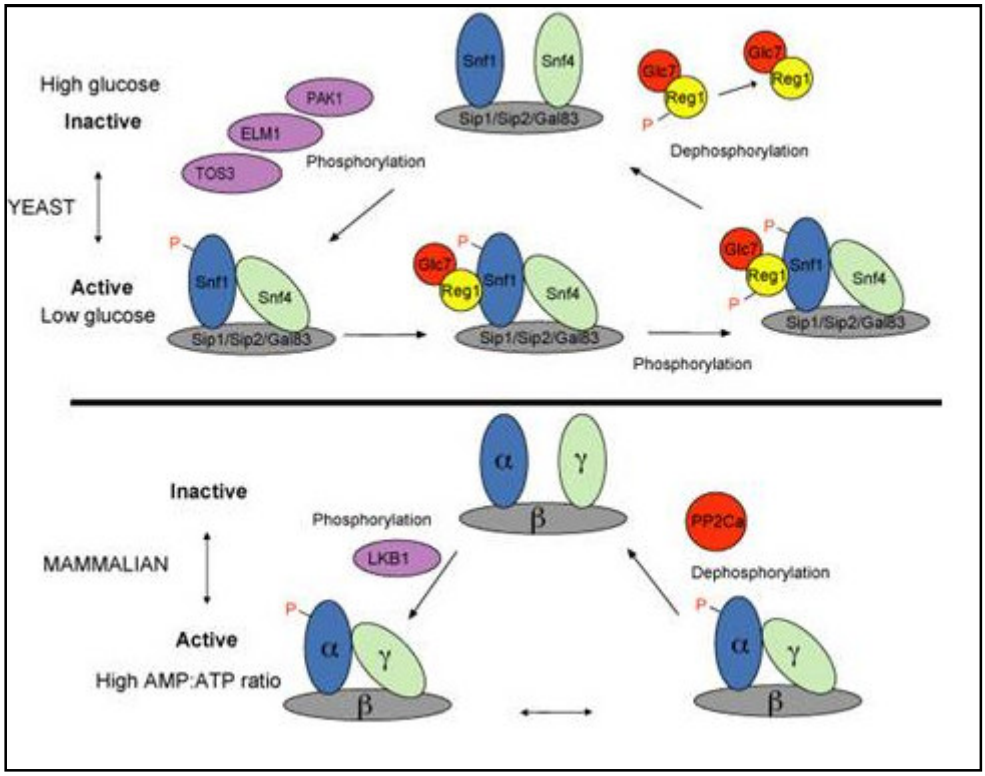

Figure 1.6. Active and Inactive Snf1 and AMPK complexes. (Hohmann, 2006)

Elm1p, Pak1p and Tos3p are a subfamily of yeast kinases that can phosphorylate the Snf1p. At least one of these is required for the activation of the Snf1 complex in the

glucose limited condition (Sutherland *et al.*; 2003). There is also evidence that show interaction between Snf1 complex and Cka1 protein indicating a potential kinase of the complex (Nath *et al.,* 2003). On the other hand, Glc7 and Reg1 proteins form a complex which dephosphorylates the Snf1 complex and thus deactivates it.

Snf1 protein kinase also plays an important role in several other processes such as stress resistance, invasive growth and ageing (Hong *et al.,* 2007; Cullen *et al.,* 2000; Ashrafi *et al.,* 2000). In addition to these, there is also evidence showing that Snf1p is directly involved in nitrogen signaling. Nitrogen limitation improves the phosphorylation of Snf1p and activates the complex (Orlova *et al.,* 2006).

1.2.4. Trancriptional Repressor of Gluconeogenic Genes: Mig1p

It is known that there are several activating and repressing transcription factors which play a role in the adjustment of mRNA expression levels of the related genes as a response to the glucose level change in the medium. Their control mechanism is via the activity of the Snf1 kinase complex. Mig1p is the first discovered transcriptional repressor acting downstream of the Snf1 complex. It was first identified as a multicopy inhibitor of expression of genes related with the galactose utilization pathway, namely the GAL genes (Nehlin and Ronne, 1990).

A major function of Snf1 kinase is to inhibit repression by Mig1 when glucose is limiting as mentioned in Section 1.2.3. Snf1p appears to regulate the localization of Mig1p. At low levels of glucose, Snf1 complex is activated through phosphorylation by upstream kinases, and the active Snf1p then catalyzes the phosphorylation of Mig1, causing Mig1p to translocate from the nucleus to the cytosol. As a result, Mig1p-associated repression of several different genes is released. On the other hand, at high levels of glucose, when Snf1 complex is not phosphorylated, Mig1p remains in the nucleus and here it represses the transcription of genes involved in the metabolism of carbon sources other than glucose and fructose (Carlson, 1999; Westergaard et al., 2007). (See Figure 1.7)

Major role of Mig1p in the glucose repression pathway of yeast is to repress transcription of the genes encoding the enzymes required for the utilization of carbon

sources other than glucose. When glucose is abundant, it is deactivated so the utilization of glucose as the primary carbon source is facilitated. The target genes of Mig1p involve those encoding enzymes for utilization of the sugars maltose (MAL), sucrose (SUC), or galactose (GAL) (Schüller, 2003; Carlson, 1999). There is also evidence showing the transcriptional activator activity of Mig1p in some cases (Schuller, 2003; Santangelo, 2006). As a result of these facts, deletion of the MIG1 gene relieves glucose repression of numerous target genes (Flick and Johnston, 1992; Schuller, 2003) and it is thought that Mig1p interacts with the intergenic regions of at least 90 genes in vivo (Lutfiyya *et al.*, 1998; Mukherjee, 2004).

Mig1p is a Zinc-finger protein, of the C_2H_2 type, that binds specifically to DNA molecules with a GC-rich consensus sequence and a flanking AT sequence (Lundin *et al.*, 1994). The fingers of Mig1p are very similar to those present in the mammalian Egr finger proteins, which are induced during early growth response, and also to the finger protein encoded by a human gene that is deleted in Wilms' tumor cells (Nehlin and Ronne, 1990).

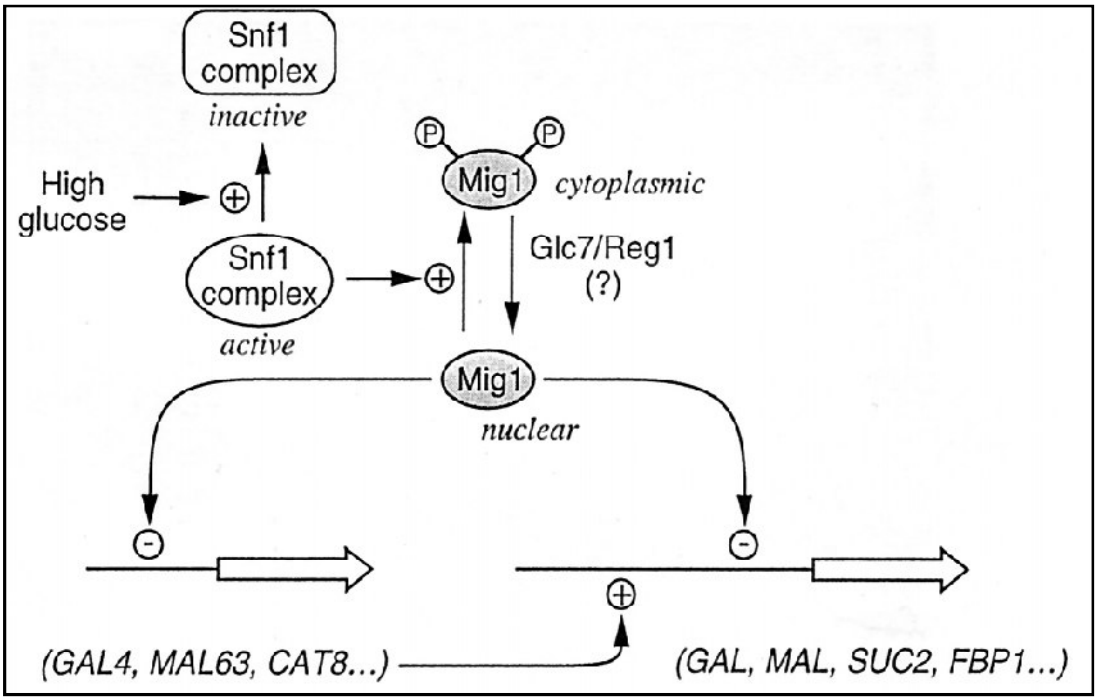

Figure 1.7. Schematic view of the mode of action of Mig1p and its regulation (Gancedo, 1998)

1.3. The Aim of this Thesis

In this study, the active signaling pathways in ΔSnf1, ΔSnf4 and ΔSnf1 ΔSnf4 deletion mutants of yeast were identified through integration of transcriptome data with the reported protein-protein interactions and protein- DNA interactions collected from literature. The collection of the protein-protein interactions were initially preprocessed via SPA algorithm. The "Average Path Score" and "Reporter Features Algorithm" were used to determine the most active linear signaling paths starting from Hxk2p, ending at the transcriptional repressor Mig1p and to identify the most responsive key elements of these paths in the deletion mutants of various subunits of the Snf1 complex.

The reporter features algorithm was also used to detect the most responsive key transcription factors to the genetic disturbances made. Finally, a complete view of the glucose repression pathway in yeast was predicted in the absence of the constituent proteins of the Snf1 kinase complex. The method is explained in detail in the Methods section while the comparisons of the results with each other and with the wild type are made in the Discussion section. The study is summarized with important key points in the Conclusions and Recommendations section. Some new techniques and additional work to improve the study are also suggested.

2. METHODS

2.1. Reconstruction of the Global Signaling Network of Yeast

The candidate proteins were selected by the Selective Pemissibility Algorithm (Arga *et al.*, 2007) which meant that these proteins were predicted to be associated with the ‘‘cellular communication and signal transduction mechanism’’ category in the MIPS Comprehensive Yeast Genome Database (Guldener *et al.*, 2005). For a subset of these candidate proteins, the transcriptome data was available for the subject deletion mutants, so this subset of proteins were included in the network.

The interactions between these proteins were downloaded from the BioGrid database and filtered according to the criterion of reliability of the experiment type by which the interaction is detected.

The experiment types which are accepted as realiable are listed below:

1. Affinity capture-MS
2. Affinity capture-RNA
3. Affinity capture-Western
4. Biochemical Activity
5. Co-crystal Structure
6. Far Western
7. FRET
8. Protein-peptide
9. Two Hybrid

The topological properties such as the diameter of the network and the connectivity of the proteins in the network were determined using the Bioinformatics package of Matlab.

2.2. Reconstruction of the Transcriptional Regulatory Network of Yeast

In this study, the genome-scale transcriptional regulatory network of S. *cerevisiae* was reconstructed by assembling known regulatory interactions between known transcription factors and their target genes. The transcription factors list consists of 142 transcription factors which are common in both TRANSFAC and YEASTRACT databases by July 2007. 138 of these proteins have the transcriptional regulation activity in also SGD database by April 2008. The regulatory interactions of these transcription factors with their target genes are collected from YEASTRACT database. The regulatory network of yeast was represented as a bipartite undirected graph. In this graph both TFs and the target genes were represented as nodes and the genetic interactions between them were represented as edges. The regulatory network contained nodes and edges.

2.3. Analysis of Transcriptome Data

Microarray data from the deletion mutants were normalized using the *Affy* package of Bioconductor which is a tool for complete microarray analysis written in the statistical computing language R and the p values for the whole genome (the probability that the significant change in the expression levels of the genes with respect to wild type is due to chance only) were calculated using the *Limma* Package of Bioconductor. Then they were corrected for multiple testing by the Benjamin- Hochberg correction method. The (1-p) values for the candidate proteins which denoted the probability significant change in expression levels were converted to Z scores by using *qnorm* function of R which is the inverse of the cumulative standard normal probability distribution function.

$$Z_n = \theta^{-1}(1-p) \tag{2.1}$$

Where θ denotes the cumulative standard normal probability distribution function and Z_n denotes the Z score of one gene.

2.4. Identification of the Linear Paths in Glucose Repression Pathway

The linear paths with a maximum path length of nine and 10 were determined by using the NetSearch algorithm. The paths started with the hexose transporter protein Hxk2 and ended in Mig1, the transcriptional repressor of the gluconeogenic genes.

2.5. Scoring of the Linear Paths

The score of each path identified were formulated as follows as in Eqn 2.2.

$$Z_{path} = \sum_{1}^{k} Z / n \tag{2.2}$$

Where n is the number of proteins in the path and Z is the Z score corresponding to each protein. Then the top scoring paths above certain thresholds for each case were determined by using histograms and the proteins involved were further investigated.

2.6. Identification of the Key Transcription Factors

The approach used to identify key transcription factors is the modified version of the approach previously developed for identification of reporter metabolites in yeast. The algorithm developed for the identification of the reporter proteins is basically a simplification of the algorithm presented by Ideker et al., in 2001. (Ideker *et al.*, 2002) By this algorithm, the most responsive transcription factors to the perturbation given to the organism are identified. They are named as the reporter transcription factors around which the most significant changes occur due to the perturbation given. The reporter features algorithm was first developed based on metabolites (Patil and Nielsen, 2005) and then extended to cover reporter proteins, protein complexes, GO terms and transcription factors (Oliveira *et al.*, 2008).

In this method each transcription factor (TF) is scored based on its neighbors' Z scores calculated according to Eqn. 2.1. The input for the algorithm is the transcriptional regulatory network described in Chapter 2.2. The algorithm consists of the following steps.

- Each TF node was scored based on the normalized transcriptional response of its neighboring genes. For that purpose, Z scores of the genes connected to the TF were summed and this total score was multiplied by the inverse square root of the number of neighbors that each TF has.

$$Z_{TF} = 1/k \sum_{i=1}^{n} Z_i \tag{2.3}$$

- Z_{TF} scores were then corrected for the background distribution. 1000 sets of k genes were randomly selected from the graph and new random Z_{TF} scores were calculated for each TF. Then, the mean of the resultant Z_{TF} scores were subtracted from the previously calculated score and divided by their standard deviation.

$$Z_{correctedTF} = \frac{Z_{TF} - \mu_k}{\sigma_k} \tag{2.4}$$

- These corrected Z scores were ranked and the high scoring TFs above a threshold were accepted as the reporter (key) TFs. The threshold determining process is discussed in detail in the discussion section.

2.7. Identification of the Reporter Proteins

The interface implemented in C++ was used for calculations. (Patil and Nielsen, 2008). The inputs for this algorithm consisted of the SPA-preprocessed protein-protein interaction network of yeast and the p-values which show the significance of the biologically meaningful differential expression of the genes. The algorithm is basically the same as explained in the previous section.

3. RESULTS AND DISCUSSION

3.1. Properties of the Reconstructed Global Signaling Network

The global signaling network of yeast which was preprocessed based on the SPA algorithm was reconstructed as a unipartite graph following the method described in Section 2.1. A unipartite graph means a graph in which all nodes might be connected with each other. In this unipartite graph, the nodes denote the proteins whereas the edges are the interactions between each other. The graph consisted of 11242 directed edges and 1521 nodes. Diameter of the network was determined as 10 which is the measure of the length of the maximum of the shortest paths between each node. An earlier study on SPA based reconstruction of the signaling network of yeast had resulted in a network of 1388 nodes and 4640 edges. (Arga *et al.*, 2007) The increase in the number of proteins and the interactions clearly indicates the updated information on signaling networks.

3.2. Identification of the Significant Linear Paths with a Maximum Path Length of Nine

All possible linear paths starting from the membrane-bound receptor and ending at the transcription factor were identified through pathway analysis using the NetSearch Algorithm (Steffen *et al.*, 2002). The starting membrane-bound receptor is chosen as HXK2 (Hexokinase Isoenzyme 2; systematic name: YGL253W). HXK2 is involved in the glucose uptake process. Its product catalyzes the phosphorylation of glucose at C6. Hxk2p is phosphorylated *in vivo* and can be autophosphorylated *in vitro*. It was also shown that it is more highly phosphorylated when the glucose concentration in the medium is low. (Herrero *et al.*, 1989; Vojtek *et al.*, 1990). Therefore it has been proposed that Hxk2p plays a role as a receptor in the glucose repression pathway of yeast. Due to these evidence indicating the relation of HXK2 with the carbon catabolite repression pathway, the linear paths starting from this receptor protein are considered to be important paths for the transduction of signal in glucose repression mechanism in both wild type cells and in deletion mutants of the Snf1 complex.

The final protein was also assumed to be Mig1p. Mig1p is the repressing transcription factor involved in glucose repression. The operation of Mig1p appears to be controlled by the protein kinase Snf1. (Östling *et al.*, 1996, Treitel *et al.*, 1995) There is no clear evidence indicating an alternative control mechanism for Mig1p.

The SPA based pre-processed network, the starting protein, the final protein, minimum and maximum path length values were given as input and the NetSearch algorithm gave the linear paths between the specified proteins in specified path length values as the output. Since the aim in using the algorithm is to find the linear signal transduction paths in three different deletion mutants, the global network lacked the protein encoded by the deleted gene in each case. The number of linear paths with a maximum path length of nine is shown in Table 3.1.

Table 3.1. The number of identified linear paths for maximum path length of nine

Strain	Number of paths
snf1Δ	1,051,148
snf4Δ	4,050,224
snf1Δ snf4Δ	1,038,734

As seen in Table 3.1, the number of identified paths was approximately four-fold higher in Snf4 deletion mutant with respect to the other mutants. The reason of this finding is that most of the interactions of the Snf1 complex are documented as interactions with Snf1p in the databases. When the Snf1p is removed from the global signaling network of yeast, most of the interactions of the Snf1 complex are also removed. In other words, in the Snf1 and Snf1Snf4 deletion mutants, the identified linear paths did not contain Snf1p and thus the Snf1 complex. This resulted in a clear decrease in the number of identified linear paths in the Snf1 and Snf1Snf4 deletion mutants.

3.3. Average Path Score Algorithm

The linear paths consisted of minimum two (HXK2 and MIG1 only) and maximum nine proteins. The total scores of the paths were calculated as the averages of the Z scores of the genes encoding the proteins involved in the paths. (See Equations 2.1 and 2.2 and Sections 2.3 and 2.5 for details of the algorithm) Due to the immense amount of interactions between proteins, namely the high connectivity of the network, the number of the potential linear paths in each case was also very high. Under these conditions, it was a hard task to set a threshold for the paths which would be accepted as significant. Because setting a very high threshold value, would cause loss of information coming from the big amount of paths below the threshold whereas setting a low threshold would make the biological explanation of the resulting sub-networks harder. The maximum and minimum average path scores of the paths are shown in Table 3.2.

Table 3.2. Path scores for maximum path length of nine

Strain	Max. Score	Min. Score
snf1Δ	2.085	-0.557
snf4Δ	1.875	-0.783
snf1Δ snf4Δ	2.757	-0.299

The process of setting thresholds was accomplished by using histograms. An example for the histograms showing the distribution of the scores of the paths can be seen in Figure 3.1 for the Snf1 deletion mutant. The paths scoring above the threshold value which is marked by the arrow, are accepted as significant. In Figure 3.2 and Figure 3.3, the histograms showing the distribution of the average path scores are plotted for Snf1Snf4 and Snf4 deletion mutants respectively.

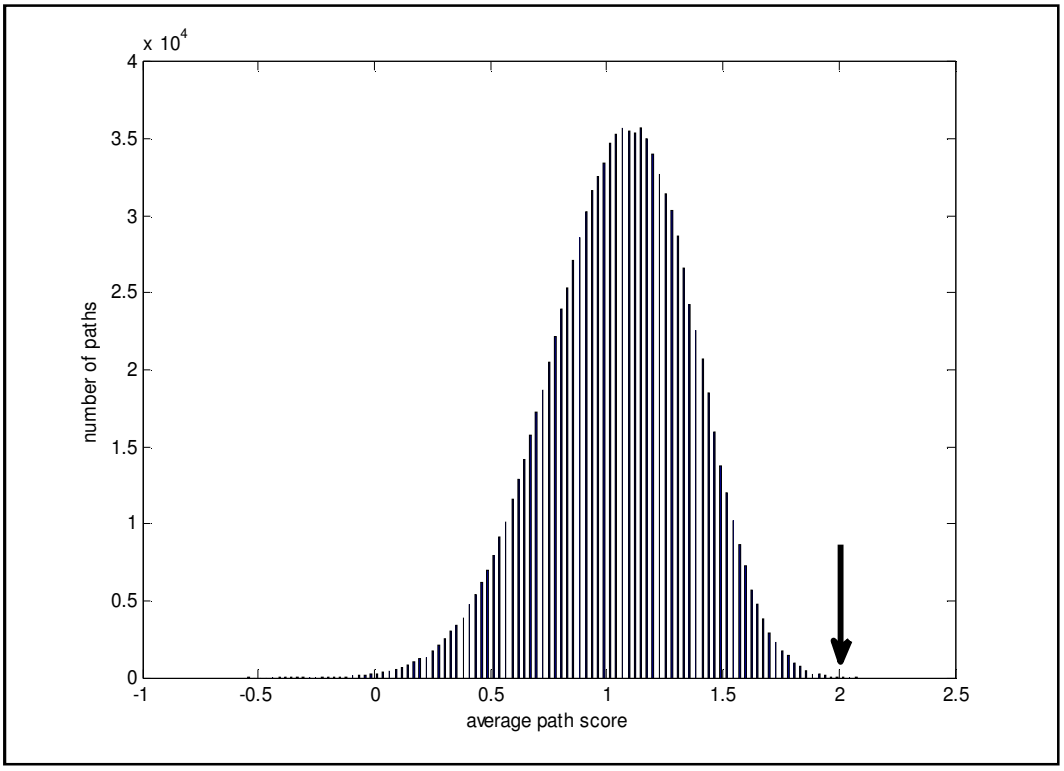

Figure 3.1. Distribution of the path scores in *snf1Δ* strain for maximum path length of nine

Finally, the sub-networks were reconstructed through a collection of the significant linear paths in each case. The number of these significant linear paths and the total number of the proteins involved in each sub-network are tabulated. (See Table 3.3) The graphs in Figure 3.4, Figure 3.5 and Figure 3.6 show the resulting sub-networks in Snf1, Snf1Snf4 and Snf4 deletion mutants respectively. The legend for the color coding is explained in Table 3.4. Graphviz software was used for the visualization of the resulting sub-networks.

Table 3.3. Properties of the Significant Paths for maximum path length of nine)

Strain	Number of significant paths	Significance threshold for average path score	Number of proteins
snf1Δ	500	1.894	98
snf4Δ	2000	1.829	131
snf1Δ snf4Δ	500	2.594	90

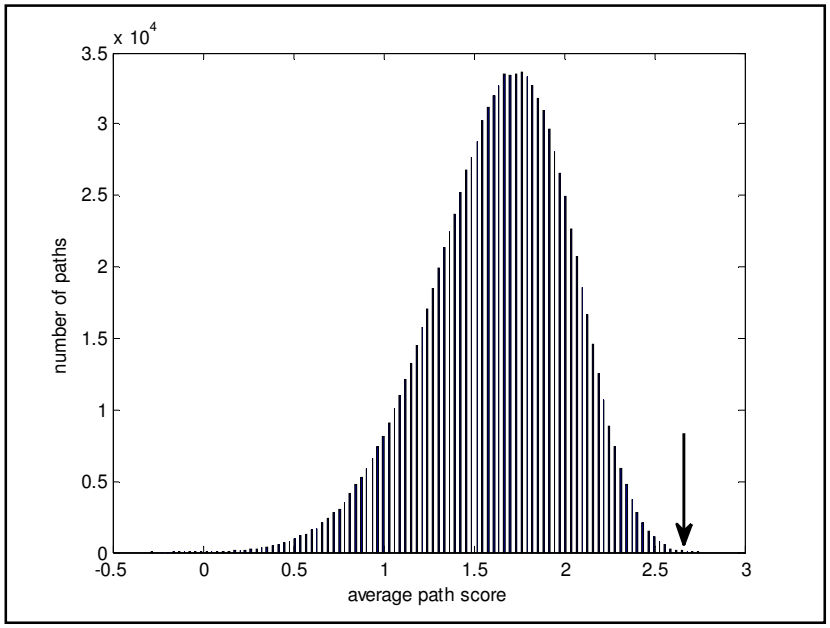

Figure 3.2. Distribution of the path scores in *snf1Δsnf4Δ* strain for maximum path length of nine

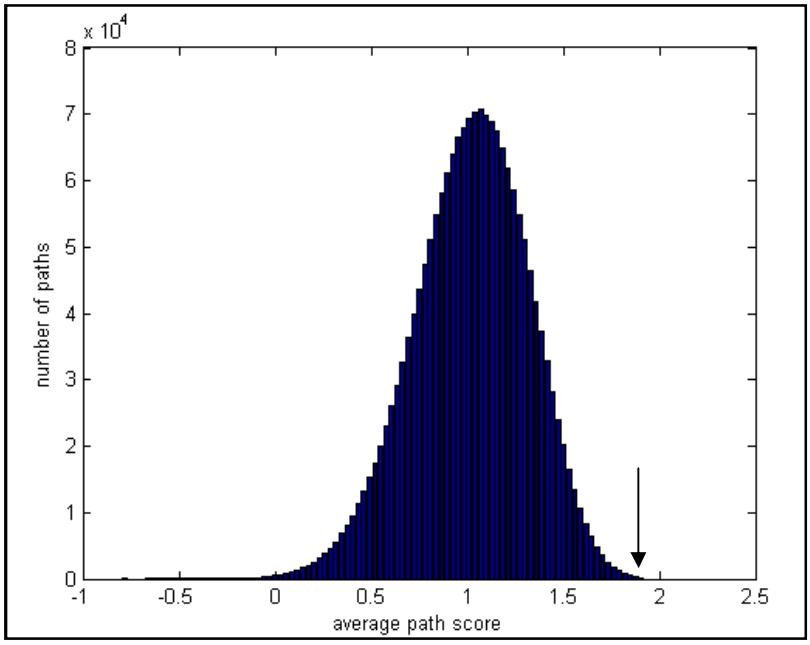

Figure 3.3. Distribution of the path scores in *snf4Δ* strain for maximum path length of nine

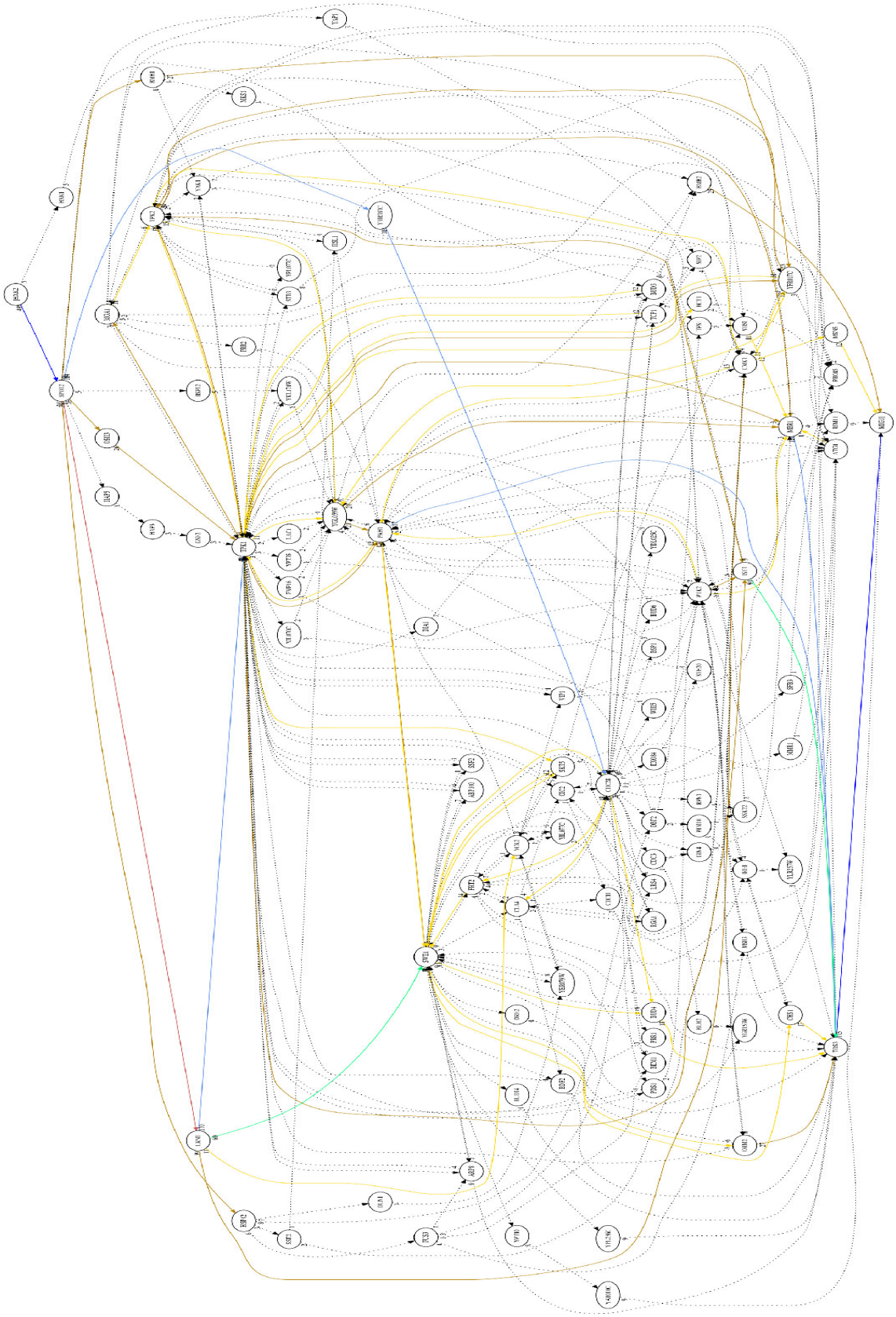

Figure 3.4. Full dimension sub-network in Snf1 deletion mutant

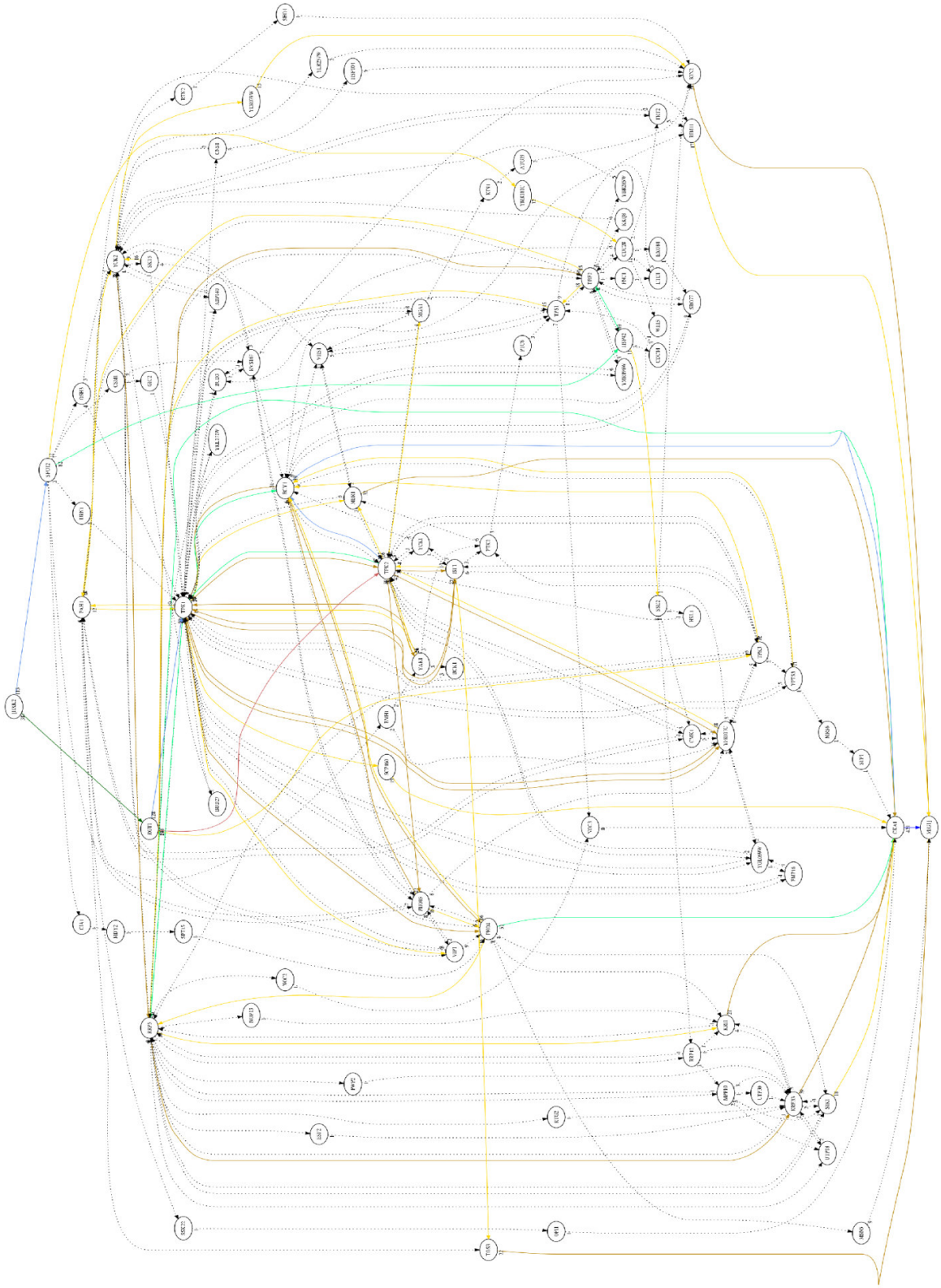

Figure 3.5. Full dimension sub-network in Snf1Snf4 deletion mutant

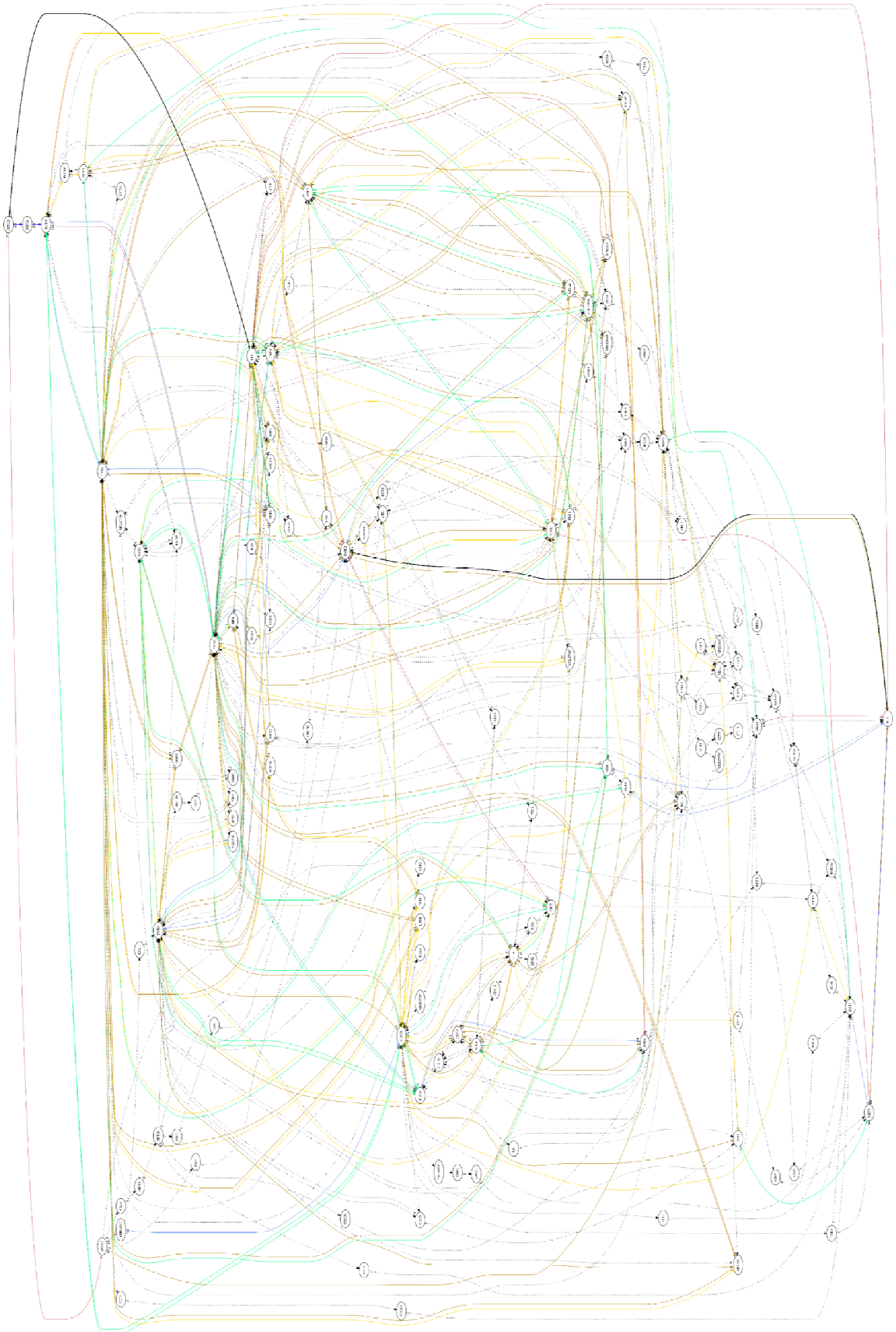

Figure 3.6. Full dimension sub-network in Snf4 deletion mutant

Table 3.4. Color coding for the edges in the graphs

The number of observations of the binary signal transducing elements in the top scoring paths	The properties of the edge in the graph
0-10	Dotted, black
10-20	Gold
20-50	Dark gold
50-100	Light green
100-200	Light blue
200-300	Red
300-400	Dark green
400-600	Dark Blue
600-800	Black
>800	Bold, black

When the linear paths which were accepted as significant were investigated in more detail; it was seen that the number of times that a certain protein transmitted signal to another certain protein changed in a range between 1 and over 70 per cent of the total number of paths. In addition to that, some proteins were observed in approximately 80 per cent of all the significant paths while others were seen only once. This finding resulted in another question about how to attribute importance to a protein predicted to have a role in the glucose repression signaling mechanism in the three deletion mutants. Two approaches were developed to answer these questions and to reduce the dimension of the resulting sub-networks by incorporating only the proteins which are more likely to contribute to the signaling process with respect to the others.

3.4. Reduction of the dimension of the resulting sub-networks

3.4.1. Approach 1: Significant Binary Signal Transducing Elements

In this approach, the number of the observations of a single signal transduction step between two certain proteins was taken into account. In other words, if a certain protein couple was detected significantly in a higher frequency than other couples in the significant linear paths, it was assumed that these proteins are more likely to participate in the signaling process in the absence of the Snf1 complex components. The threshold value

for the number of the detection of the protein couples' in the significant paths was determined by using the histograms. The histograms in Figure 3.7, Figure 3.8 and Figure 3.9 show the distribution of the number of detection of the protein couples. These protein couples are named as “Binary Signal Transducing Elements”.

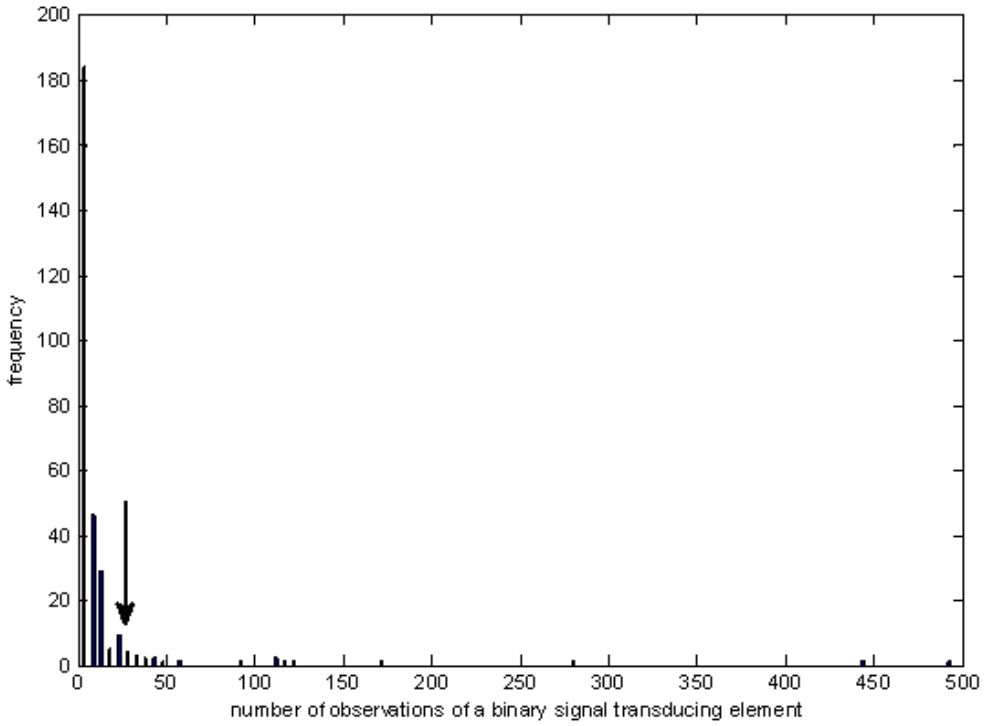

Figure 3.7. Distribution of the number of observations of the binary signal transducing elements in *snf1Δ*

The threshold values for the number of appearances of the binary signal transducing elements are tabulated in Table 3.5. The constituents of binary signal transducing elements which are above this threshold value are accepted as significantly more likely to have a role in the glucose repression mechanism of Snf1 kinase complex mutants. The threshold values are also shown by arrows on Figure 3.7, Figure 3.8 and Figure 3.9.

This approach led to eliminate the enormous number of edges which are rarely observed with respect to other edges. The selected binary transducing elements were considered significantly important in transmitting the signal and to have a possible role in the signal transduction between HXK2 and MIG1.

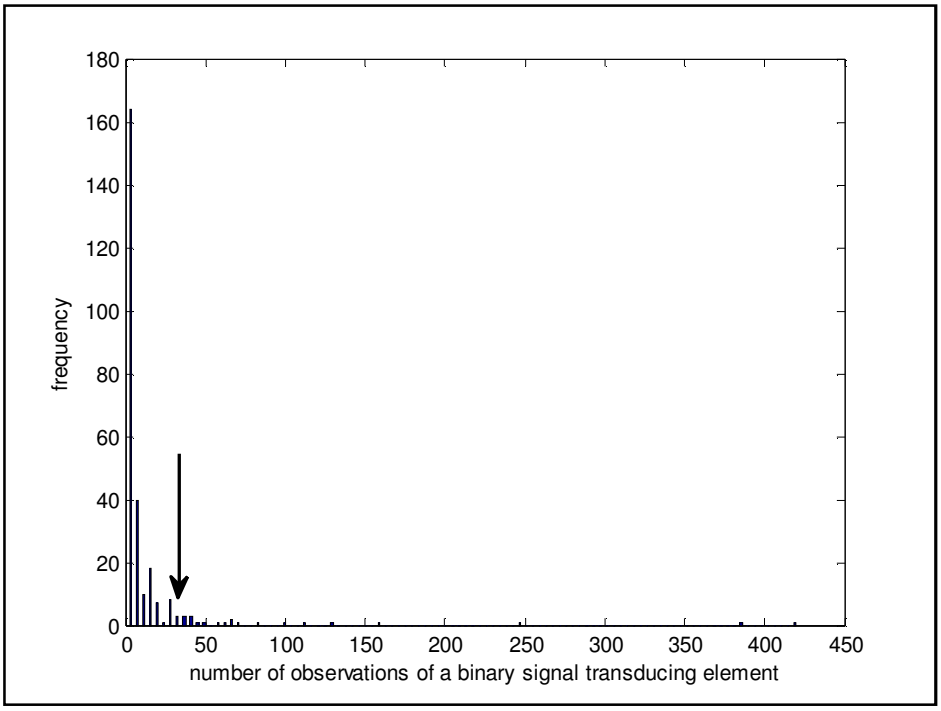

Figure 3.8. Distribution of the number of observations of the binary signal transducing elements in *snf1Δ snf4Δ*

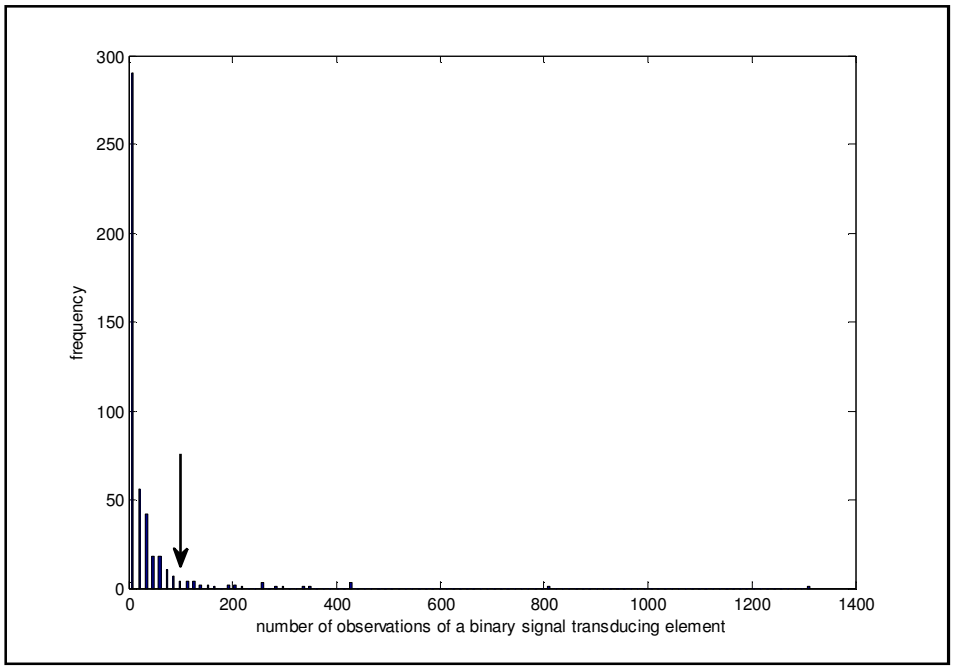

Figure 3.9. Distribution of the number of observations of the binary signal transducing elements in *snf4Δ*

Table 3.5. Threshold values for the detection of significant binary signal transducing elements

Strain	Threshold value
snf1Δ	25
snf4Δ	70
snf1Δ snf4Δ	40

The resulting sub-networks are presented in Figure 3.10, Figure 3.11 and Figure 3.12. The numbers on the edges show the number of observations of each edge. The color coding legend given in Table 3.4 is valid also for Figure 3.10, Figure 3.11 and Figure 3.12.

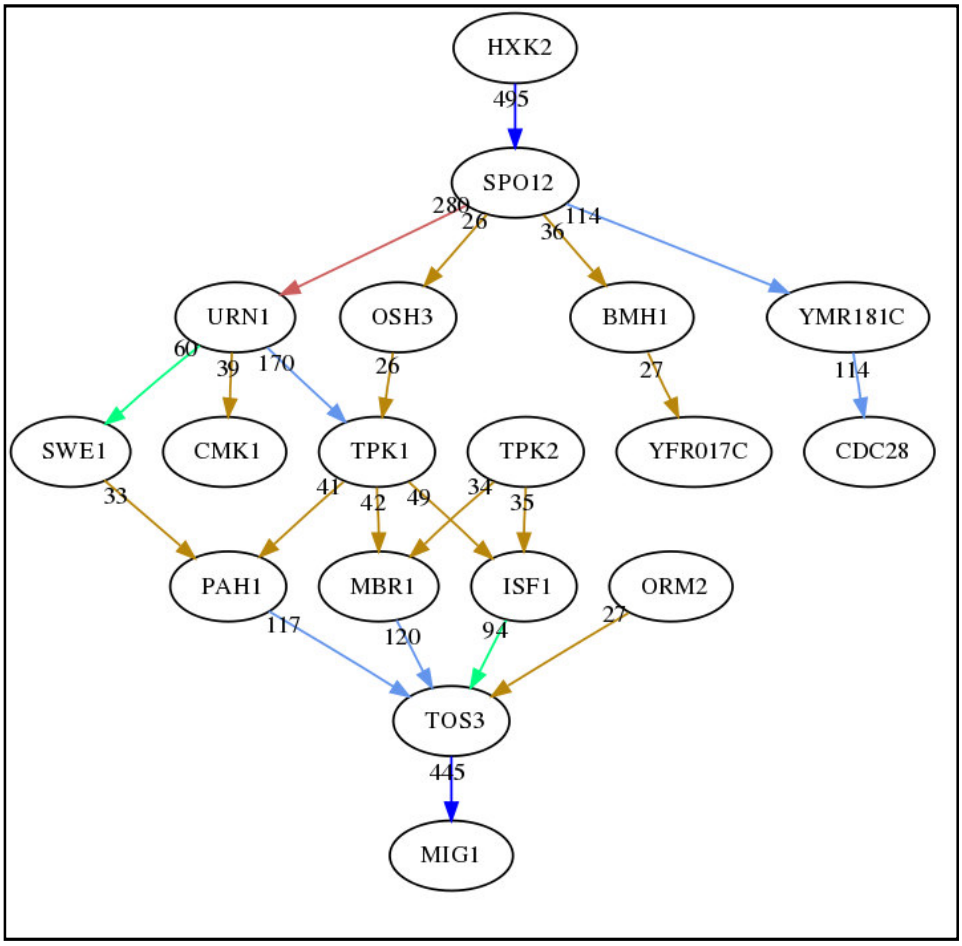

Figure 3.10. Sub-network in *snf1Δ* strain predicted by the first approach

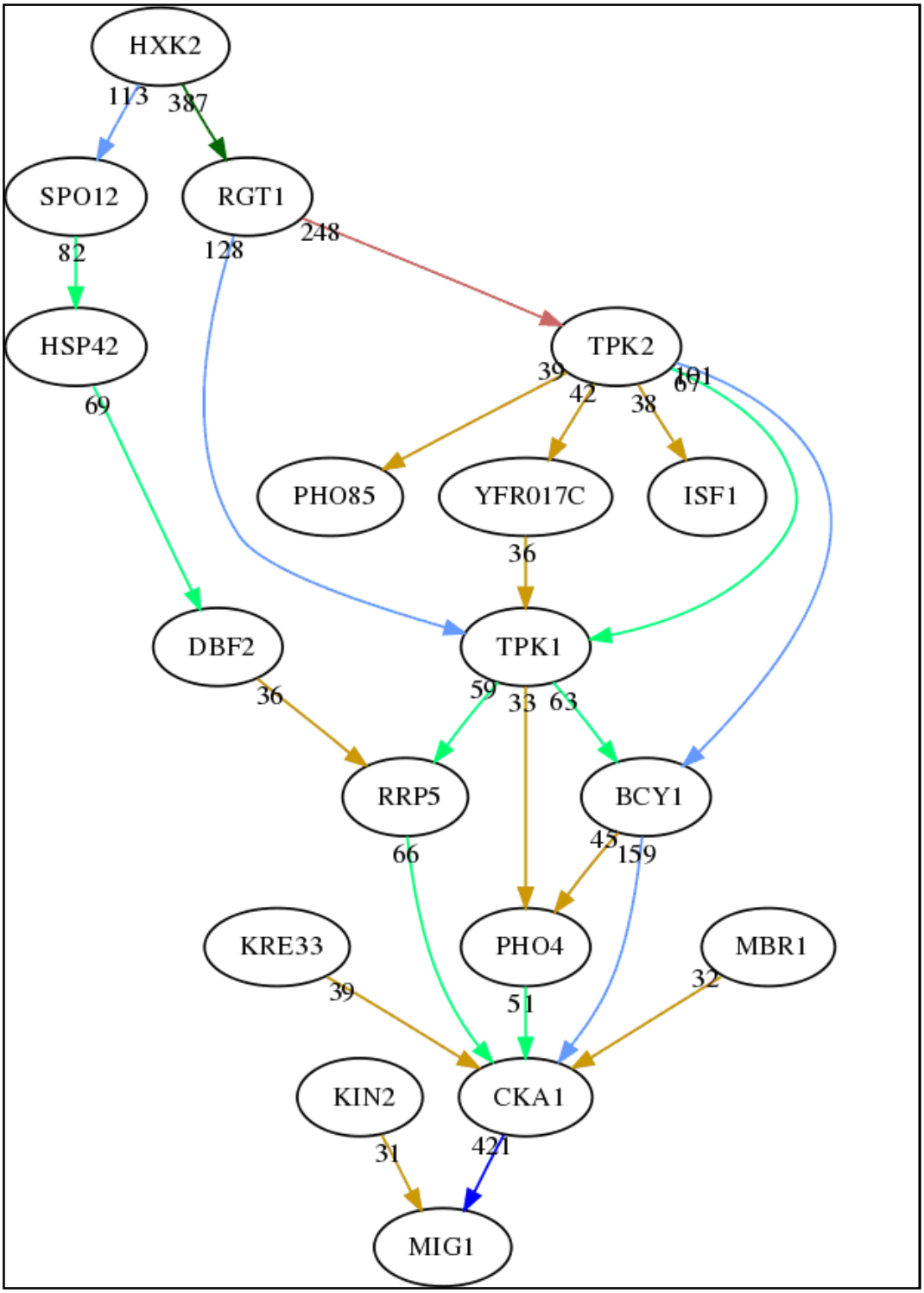

Figure 3.11. Sub-network in *snf1Δ snf4Δ* strain predicted by the first approach

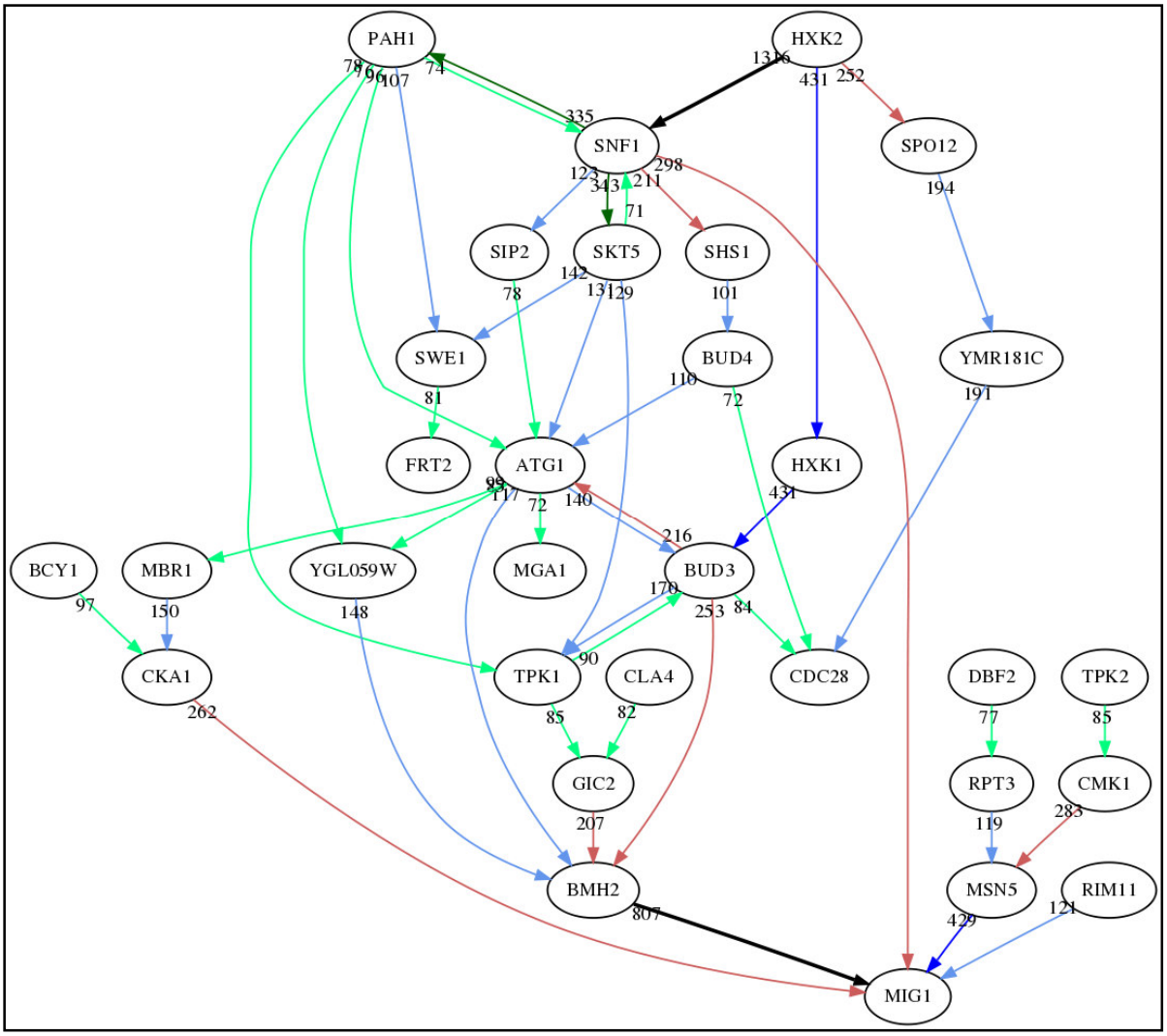

Figure 3.12. Sub-network in *snf4Δ* strain predicted by the first approach

3.4.2. Approach 2: Significant Individual Signal Transducing Elements

In this approach, the proteins which are more likely to play a role in the glucose repression mechanism of the three deletion mutants were determined directly based on their number of appearances in the significant linear paths identified in the deletion mutants. The proteins which were detected significantly in higher frequency were accepted as individual signal transducing elements involved in the pathway in the absence of one or more components of the Snf1 complex. The total number of signal coming in to and out from each protein in the sub-networks presented in Figure 3.4, Figure 3.5 and Figure 3.6

are calculated and their distribution is presented in histograms for setting the proper threshold. (Figure 3.13, Figure 3.14 and Figure 3.15)

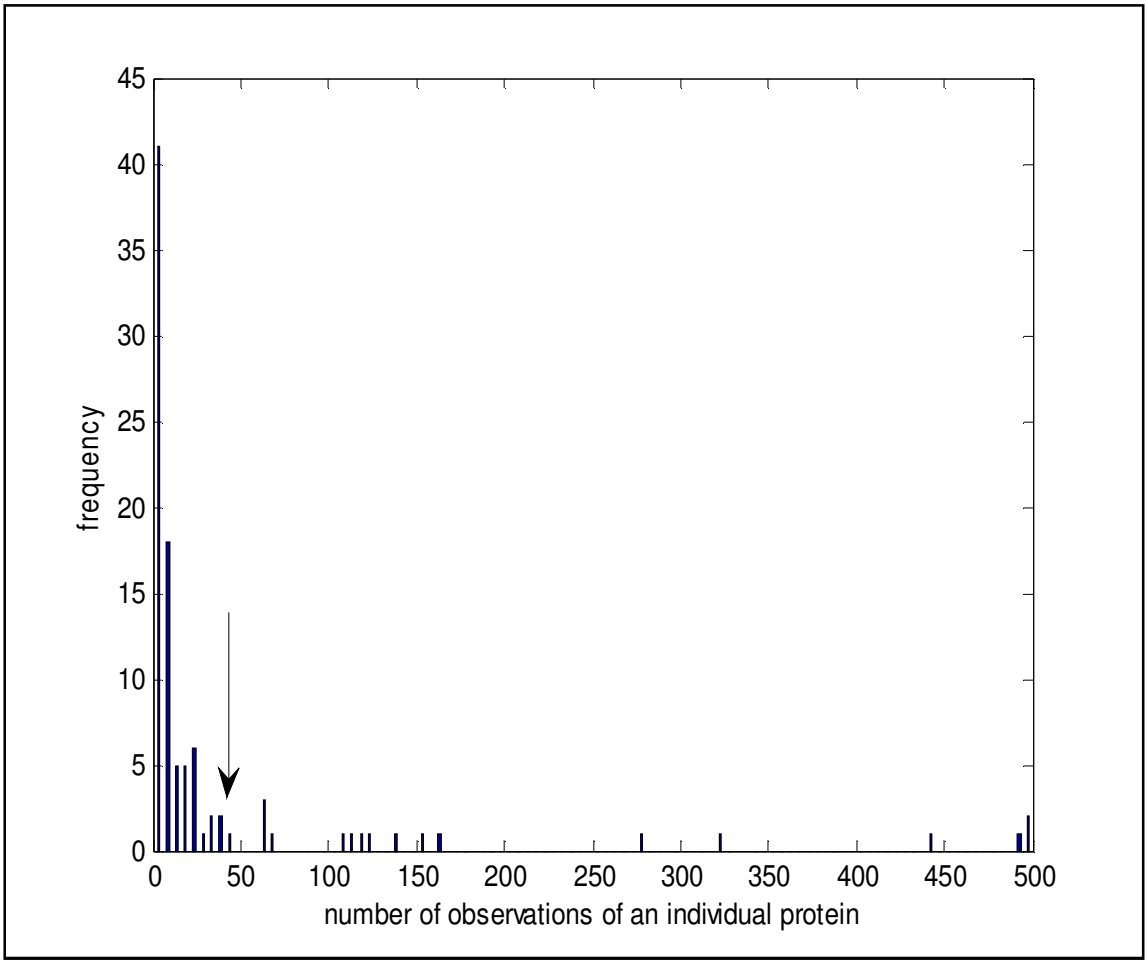

Figure 3.13. Distribution of the number of observations of individual signal transducing elements in *snf1Δ* strain for maximum path length of nine

These sub-networks contained only the proteins which are more likely to play a role in the signaling mechanism, as the nodes and the signaling incidences between these proteins as edges. The resulting sub-networks are visualized in Figure 3.16, Figure 3.17 and Figure 3.18. The color coding legend given in Table 3.4 is valid also for Figure 3.16, Figure 3.17 and Figure 3.18.

The thresholds for significant number for signals coming in and out to a protein can be seen in Table 3.6.

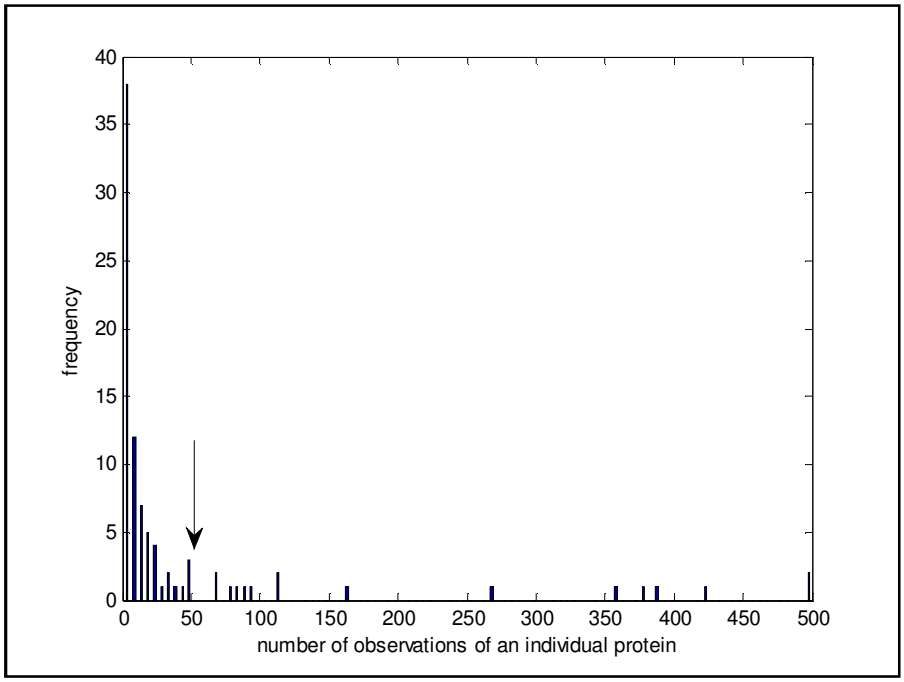

Figure 3.14. Distribution of the number of observations of individual signal transducing elements in *snf1Δ snf4Δ* strain for maximum path length of nine

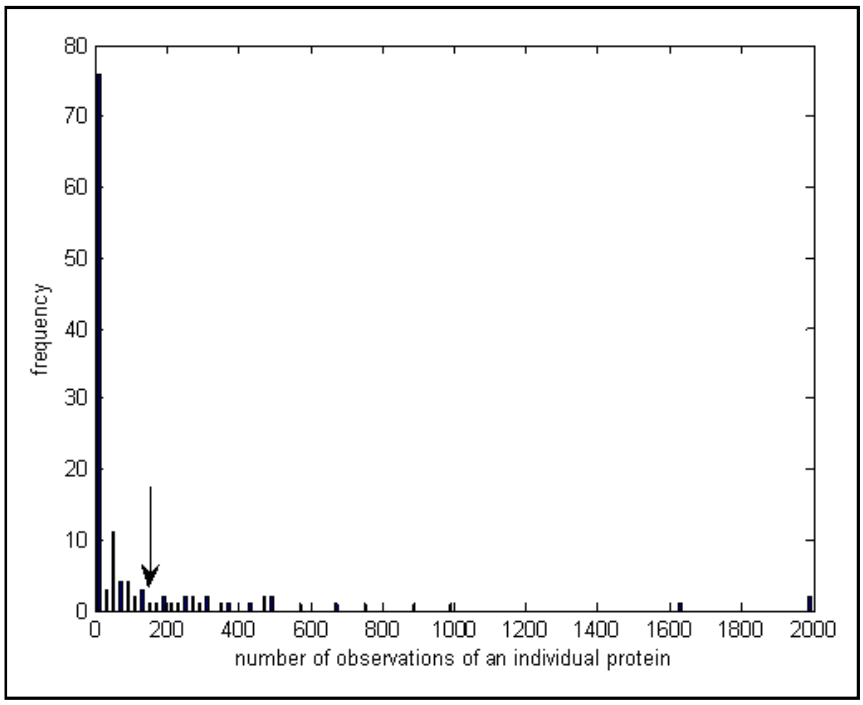

Figure 3.15. Distribution of the number of observations of individual signal transducing elements in *snf4Δ* strain for maximum path length of nine

Table 3.6. Threshold values for the detection of significant individual signal transducing elements (Maximum path length: nine)

Strain	Threshold value
Snf1 Δ	40
Snf1 Δ	140
Snf1 Δ	50

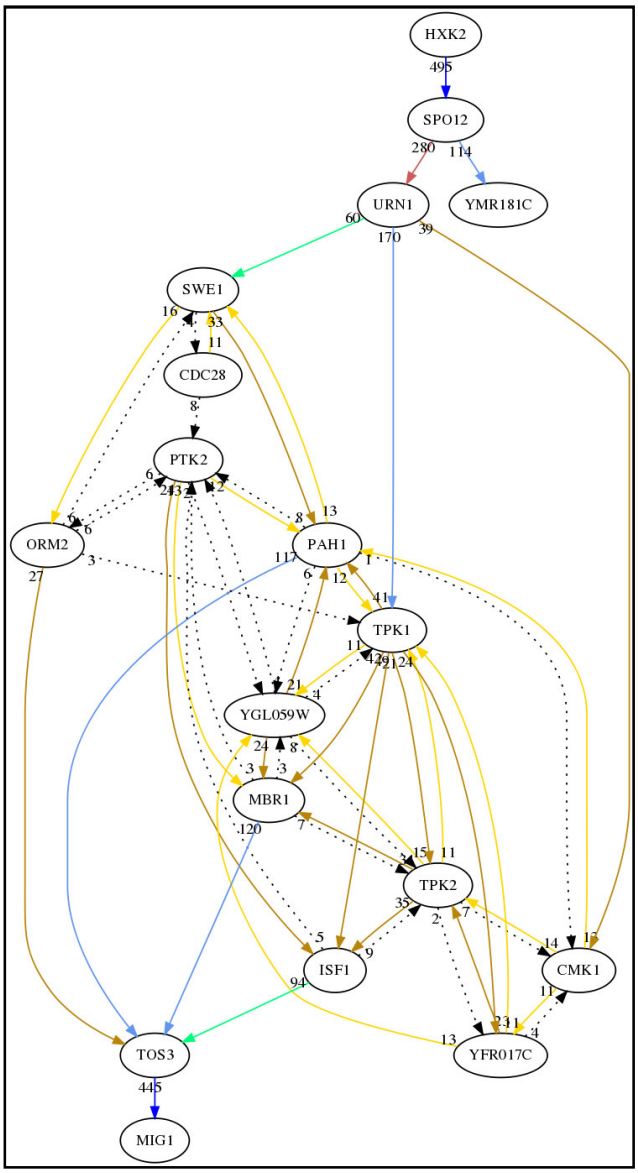

Figure 3.16. Sub-network in *snf1Δ* strain predicted by the second approach for maximum path length of nine

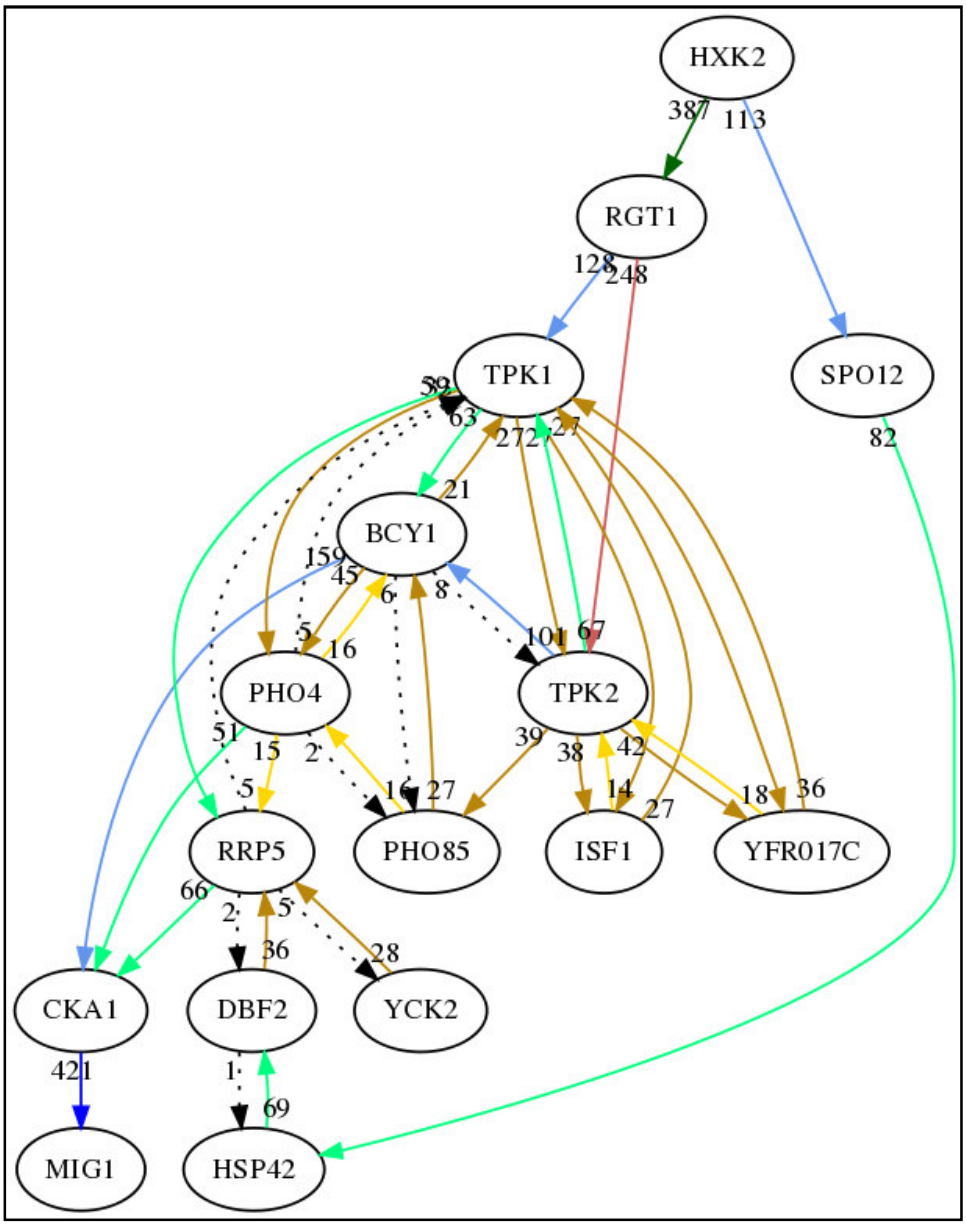

Figure 3.17. Sub-network in *snf1Δ snf4Δ* strain predicted by the second approach for maximum path length of nine

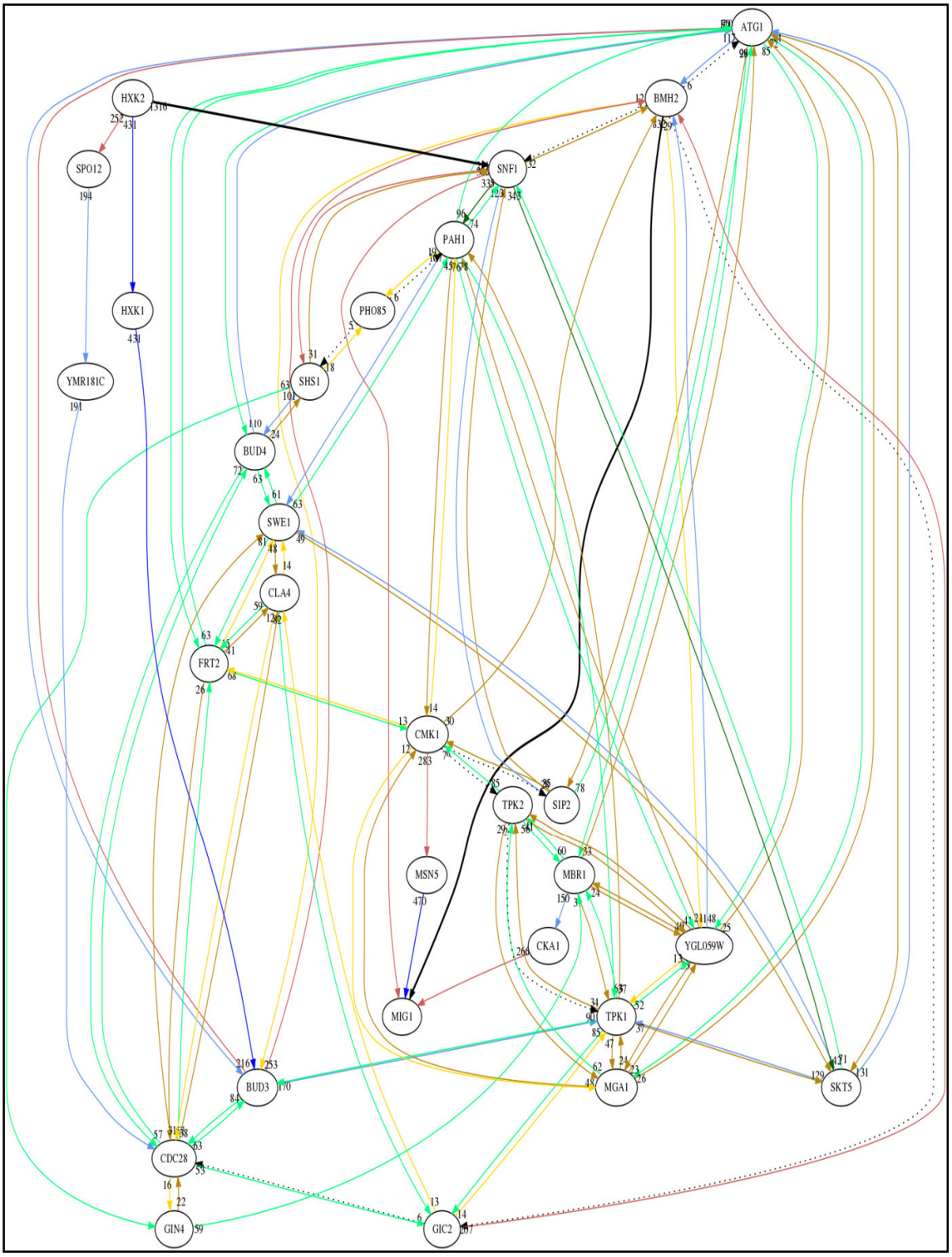

Figure 3.18. Sub-network in *snf4Δ* strain predicted by the second approach for maximum path length of nine

Table 3.7. Predicted common proteins by the two approaches

Snf1Δ strain	*Snf4Δ* strain	*Snf1Δ Snf4Δ* strain
SPO12	SPO12	SPO12
URN1	PAH1	RGT1
YMR181C	SNF1	TPK1
SWE1	SIP2	BCY1
CMK1	SKT5	PHO4
TPK1	SHS1	TPK2
TPK2	SWE1	RRP5
YFR017C	BUD4	PHO85
CDC28	FRT2	ISF1
PAH1	ATG1	YFR017C
MBR1	HXK1	CKA1
ISF1	YMR181C	DBF2
ORM2	BCY1	HSP42
TOS3	MBR1	
	YGL059W	
	MGA1	
	BUD3	
	CKA1	
	TPK1	
	CLA4	
	CDC28	
	TPK2	
	GIC2	
	CMK1	
	BMH2	
	MSN5	

The common and different proteins predicted by the two approaches are tabulated in Table 3.7, Table 3.8, Table 3.9 and Table 3.10.

Table 3.8. Differences in the predicted proteins by the two approaches and their function and process GO terms in *Snf1Δ*

Snf1Δ strain					
1st approach			2nd approach		
Gene Name	Process GO Term	Function GO Term	Gene Name	Process GO Term	Function GO Term
OSH3	pseudohyphal growth, organelle organization and biogenesis, conjugation, cell wall organization and biogenesis, membrane organization and biogenesis, transport, vesicle-mediated transport, nuclear organization and biogenesis, anatomical structure morphogenesis	Lipid Binding	PTK2	Cell cycle, transport, cellular homeostasis	Protein kinase activity, transferase activity
BMH1	pseudohyphal growth, signal transduction, cell cycle, response to stress, carbohydrate metabolic process, generation of precursor metabolites and energy, sporulation, cell wall organization and biogenesis	Protein binding, DNA binding	YGL059W	Protein modification process	Protein kinase activity, transferase activity

When the resulting active sub-networks in the Snf1 deletion mutant predicted by the two approaches are considered together, it is easily seen that the sub-networks predicted by this second approach were only slightly different from those predicted by the previous approach. When two approaches are compared, most of the resulting proteins constituting the active sub-networks are the same differing only four genes for each deletion mutant. (See Table 3.8, Table 3.9 and Table 3.10)

Table 3.9. Differences in the predicted proteins by the two approaches and their function and process GO terms in *Snf4Δ*

Snf4Δ strain					
1st approach			2nd approach		
Gene Name	Process GO Term	Function GO Term	Gene Name	Process GO Term	Function GO Term
DBF2	Protein modification process	Protein kinase activity, transferase activity	GIN4	protein modification process, cytoskeleton organization and biogenesis, cytokinesis, organelle organization and biogenesis, cell budding, anatomical structure morphogenesis	Protein kinase activity, transferase activity
RPT3	Protein catabolic process	Peptidase activity, hydrolase activity			
RIM11	Protein modification process, response to stress, sporulation	Protein kinase activity, transferase activity			

However the advantage of the second approach over the first approach can be summarized as follows: when the binary signal transducing elements are considered as unit elements, the resulting sub-network contains couple nodes which are connected between each other but not connected to the remaining elements of the sub-network. This type of a view of the sub-network is not capable of giving an integral picture of the sub-network. The disjointed elements seem to be biologically meaningless. (See Figure 3.16, Figure 3.17 and Figure 3.18)

Table 3.10. Differences in the predicted proteins by the two approaches and their function and process GO terms in *Snf1Δ Snf4Δ*

Snf1ΔSnf4Δ strain					
1st approach			2nd approach		
Gene Name	Process GO Term	Function GO Term	Gene Name	Process GO Term	Function GO Term
KRE33	ribosome biogenesis and assembly, organelle organization and biogenesis	unknown	YCK2	protein modification process, transport, vesicle-mediated transport, cytokinesis, membrane organization and biogenesis, response to chemical stimulus, anatomical structure morphogenesis	Protein kinase activity, transferase activity
MBR1	cellular respiration, generation of precursor metabolites and energy	unknown			
KIN2	protein modification process, transport, vesicle-mediated transport	Protein kinase activity, transferase activity			

Table 3.11. Function GO terms of the 14 commonly predicted proteins in *Snf1Δ* strain

GO term	Frequency	Gene(s)
molecular function unknown	7 out of 14 genes, 50 per cent	YFR017C,SPO12,MBR1,ORM2,ISF1, YMR181C, URN1
protein kinase activity	6 out of 14 genes, 42.9 per cent	CDC28,CMK1,TOS3,TPK1,SWE1,TPK2
transferase activity	6 out of 14 genes, 42.9 per cent	CDC28,CMK1,TOS3,TPK1,SWE1,TPK2
hydrolase activity	1 out of 14 genes, 7.1 per cent	PAH1

Table 3.12. Process GO terms of the 14 commonly predicted proteins in *Snf1Δ* strain

GO term	Frequency	Gene(s)
protein modification process	5 out of 14 genes, 35.7 per cent	CDC28,CMK1,TOS3,TPK1 TPK2
signal transduction	3 out of 14 genes, 21.4 per cent	CMK1,TPK1,TPK2
cellular respiration	3 out of 14 genes, 21.4 per cent	MBR1,ISF1,PAH1
generation of precursor metabolites and energy	3 out of 14 genes, 21.4 per cent	MBR1,ISF1,PAH1
biological process unknown	3 out of 14 genes, 21.4 per cent	YFR017C,YMR181C, URN1
cell cycle	2 out of 14 genes, 14.3 per cent	SPO12,SWE1
response to stress	1 out of 14 genes, 7.1 per cent	ORM2
meiosis	1 out of 14 genes, 7.1 per cent	SPO12
carbohydrate metabolic process	1 out of 14 genes, 7.1 per cent	TOS3
response to chemical stimulus	1 out of 14 genes, 7.1 per cent	ORM2
anatomical structure morphogenesis	1 out of 14 genes, 7.1 per cent	SWE1

Table 3.13. Function GO terms of the 26 commonly predicted proteins in *Snf4Δ* strain

GO term	Frequency	Gene(s)
transferase activity	12 out of 26 genes, 46.2 per cent	CDC28,SNF1,CMK1,HXK1,PKP2,ATG1, SIP2,CKA1,TPK1,SWE1,CLA4,TPK2
protein kinase activity	11 out of 26 genes, 42.3 per cent	CDC28,SNF1,CMK1,PKP2,ATG1,SIP2, CKA1,TPK1,SWE1,CLA4,TPK2
molecular function unknown	5 out of 26 genes, 19.2 per cent	FRT2,BUD3,SPO12,MBR1,YMR181C
enzyme regulator activity	3 out of 26 genes, 11.5 per cent	SKT5,GIC2,BCY1
protein binding	2 out of 26 genes, 7.7 per cent	BMH2,MSN5
DNA binding	2 out of 26 genes, 7.7 per cent	BMH2,MGA1
transcription regulator activity	1 out of 26 genes, 3.8 per cent	MGA1
structural molecule activity	1 out of 26 genes, 3.8 per cent	SHS1
lipid binding	1 out of 26 genes, 3.8 per cent	GIC2
transporter activity	1 out of 26 genes, 3.8 per cent	MSN5
hydrolase activity	1 out of 26 genes, 3.8 per cent	PAH1

Instead, when individual elements are considered as unit elements and all the instances of signal transduction occurring between them are taken into account, this approach gives the whole topology of the resulting sub-network. So, in this thesis second approach is used for further studies with changing maximum protein number. (See Section 3.7)

When the process GO terms of the commonly predicted proteins by the two approaches are investigated, it can be easily seen in Table 3.12, Table 3.14 and Table 3.16, the most abundant process term is the “protein modification process” term. This finding is biologically meaningful because the protein modification process covers the post-translational modifications which mostly depend on protein-protein interactions. In the whole yeast genome approximately 10 per cent are involved in the protein modification process whereas at least 30 per cent of the commonly predicted genes has this process GO term; namely the CMK1, TOS3, CDC28, SNF1, PKP2 (YGL059W), ATG1, SIP2, CKA1,

CLA4, TPK2, DBF2, TPK1, PHO85 genes encode the proteins involved in the protein modification process.

Table 3.14. Process GO terms of the 26 commonly predicted proteins in *Snf4Δ* strain

GO term	Frequency	Gene(s)
protein modification process	10 out of 26 genes, 38.5 per cent	CDC28,SNF1,CMK1,PKP2,ATG1,SIP2,CKA1 TPK1,CLA4,TPK2
signal transduction	8 out of 26 genes, 30.8 per cent	BMH2,GIC2,SNF1,CMK1,SIP2,TPK1,CLA4, TPK2
anatomical structure morphogenesis	7 out of 26 genes, 26.9 per cent	BUD3,SHS1,GIC2,CKA1,SWE1,BUD4,CLA4
cytokinesis	6 out of 26 genes, 23.1 per cent	SKT5,BUD3,SHS1,GIC2,BUD4,CLA4
response to stress	6 out of 26 genes, 23.1 per cent	FRT2,SKT5,BMH2,SNF1,SIP2,CKA1
cell budding	4 out of 26 genes, 15.4 per cent	BUD3,GIC2,BUD4,CLA4
cell cycle	4 out of 26 genes, 15.4 per cent	BMH2,SPO12,CKA1,SWE1
carbohydrate metabolic process	3 out of 26 genes, 11.5 per cent	SKT5,BMH2,HXK1
generation of precursor metabolites and energy	3 out of 26 genes, 11.5 per cent	BMH2,MBR1,PAH1
pseudohyphal growth	2 out of 26 genes, 7.7 per cent	BMH2,SNF1
cellular respiration	2 out of 26 genes, 7.7 per cent	MBR1,PAH1
cell wall organization and biogenesis	2 out of 26 genes, 7.7 per cent	SKT5,BMH2
transport	2 out of 26 genes, 7.7 per cent	MSN5,HXK1
response to chemical stimulus	1 out of 26 genes, 3.8 per cent	CLA4
organelle organization and biogenesis	1 out of 26 genes, 3.8 per cent	ATG1
meiosis	1 out of 26 genes, 3.8 per cent	SPO12

Table 3.15. Function GO terms of the 13 commonly predicted proteins in *Snf1Δ Snf4Δ* strain

GO term	Frequency	Gene(s)
protein kinase activity	5 out of 13 genes, 38.5 per cent	DBF2,CKA1,TPK1,PHO85, TPK2
transferase activity	5 out of 13 genes, 38.5 per cent	DBF2,CKA1,TPK1,PHO85, TPK2
molecular function unknown	3 out of 13 genes, 23.1 per cent	YFR017C,SPO12,ISF1
DNA binding	2 out of 13 genes, 15.4 per cent	PHO4,RGT1
protein binding	2 out of 13 genes, 15.4 per cent	HSP42,RGT1
transcription regulator activity	2 out of 13 genes, 15.4 per cent	PHO4,RGT1

Table 3.16. Process GO terms of the 13 commonly predicted proteins in *Snf1Δ Snf4Δ* strain

GO term	Frequency	Gene(s)
protein modification process	5 out of 13 genes, 38.5 per cent	DBF2,CKA1,TPK1, PHO85, TPK2
response to stress	4 out of 13 genes, 30.8 per cent	HSP42,PHO4,CKA1 PHO85
organelle organization and biogenesis	3 out of 13 genes, 23.1 per cent	HSP42,RRP5,PHO85
cell cycle	3 out of 13 genes, 23.1 per cent	SPO12,CKA1,PHO85
signal transduction	2 out of 13 genes, 15.4 per cent	TPK1,TPK2
ribosome biogenesis and assembly	1 out of 13 genes, 7.7 per cent	RRP5
cytoskeleton organization and biogenesis	1 out of 13 genes, 7.7 per cent	HSP42
cellular respiration	1 out of 13 genes, 7.7 per cent	ISF1
carbohydrate metabolic process	1 out of 13 genes, 7.7 per cent	RGT1
RNA metabolic process	1 out of 13 genes, 7.7 per cent	RRP5
generation of precursor metabolites and energy	1 out of 13 genes, 7.7 per cent	ISF1
cell wall organization and biogenesis	1 out of 13 genes, 7.7 per cent	PHO85
response to chemical stimulus	1 out of 13 genes, 7.7 per cent	PHO85
meiosis	1 out of 13 genes, 7.7 per cent	SPO12
cellular homeostasis	1 out of 13 genes, 7.7 per cent	CKA1

The “signal transduction” process GO category is also very significant for Snf1 and Snf4 deletion strains. The Cmk1p, Tpk1p, Tpk2p, Bmh2p, Gic2p, Snf1p, Cmk1p, Sip2p, Tpk1p, Cla4p, Tpk2p are proved to be involved in signal transduction mechanism. The appearances of these proteins may indicate cross talk between signal transduction pathways since most of them are known to play a role in response to various types of disturbances and in various processes. This finding is also consistent with another significant process GO term. The “response to stress” GO term is enriched in all deletion mutants with a minimum frequency of 21.4 per cent.

In the Snf1 deletion mutant, the “cellular respiration” GO term is also enriched with a frequency of 21.4 per cent. The MBR1, ISF1 and PAH1 genes are involved in respiration and they are significantly down regulated in the Snf1 deletion case. This observation indicates a deficiency in respiration in the Snf1 deletion. In the other deletion mutants, this category is not enriched as much as the Snf1 deletion case.

Another observation is that products of genes involved in sporulation play role in the predicted glucose repression sub-networks in each deletion mutant. The Spo12p is directly involved in meosis and it is the second protein to which the signal is transduced from Hxk2p in all deletion mutants. When the cell cannot respond to stress efficiently, the cell holds for a while until the stress conditions alter and doesn’t reproduce in the meantime. That’s why it is meaningful to observe cross talk between the sporulation, cell cycle processes and the glucose repression pathway in a degree when the components of the genes encoding the Snf1 kinase complex are deleted. The Spo12p also has the “cell cycle” process GO term together with the Swe1p, Bmh2p, Cka1p and Pho85p. These cell cycle proteins are quite enriched with a frequency of 23.1 per cent in the Snf1Snf4 deletion mutant where the glucose repression pathway is expected to be damaged the most in all three deletion mutants.

While Snf1p still plays a role in the Snf4 deletion mutant, the Tos3p also appears in the predicted sub-network in the Snf1 deletion mutant indicating that some of the pathway elements are the same in the deletion mutants of the constituents of the Snf1 complex. Tos3p has been identified as the upstream activator of Snf1p. In the absence of Snf1p, the signal is transduced from Tos3p to the final element Mig1p whereas this role belongs to

Snf1p itself in wild type yeast cells. This result is obtained by both approaches. On the other hand, the Cka1p, Bmh2p and Msn5p also seem to play the role of the Snf1p in the other two deletion mutants.

Referring to Table 3.11, Table 3.13 and Table 3.15, the "protein kinase activity" functional GO category is enriched in all three deletion mutants with a minimum frequency of 38.5 per cent. This finding indicates the potential kinase activity of the most of the proteins predicted in the sub-network which give makes their appearance in the signaling mechanism biologically meaningful.

The enriched category "transferase activity" (with a frequency of 42.9 per cent in Snf1 deletion mutant), and other significant categories such as enzyme regulator activity (with a frequency of 11.5 per cent in Snf4 deletion mutant) and protein binding activity (with a frequency of 7.7 per cent in Snf4 deletion mutant) also indicates that the results are biologically meaningful. Because, these categories are directly relevant to signaling mechanism.

3.5. Identification of the Key Proteins by the Reporter Features Approach

The reporter features algorithm was also used for the detection of key proteins. The key proteins, namely the hot spot proteins are defined as the proteins around which the most significant changes occur when shifted from one condition to another.

The physical interaction network which is used as the input for this algorithm is not exactly the same with the network used for the path scoring algorithm. In this network the protein encoded by the deleted gene is not removed since the algorithm investigates the proteins around which the significant activation occur not the proteins which are directly involved in transducing the signal.

The "Reporter Features Algorithm" implemented in C++ was used for the analysis. (Oliveira *et al.*, 2008) The cut off values for the Z scores of the proteins are determined by the help of the histograms showing the distribution of the Z scores against the proteins in

the network. The arrows in the figures shows the cut off values above which the proteins are accepted as significantly key proteins. (See Figure 3.19, Figure 3.20 and Figure 3.21)

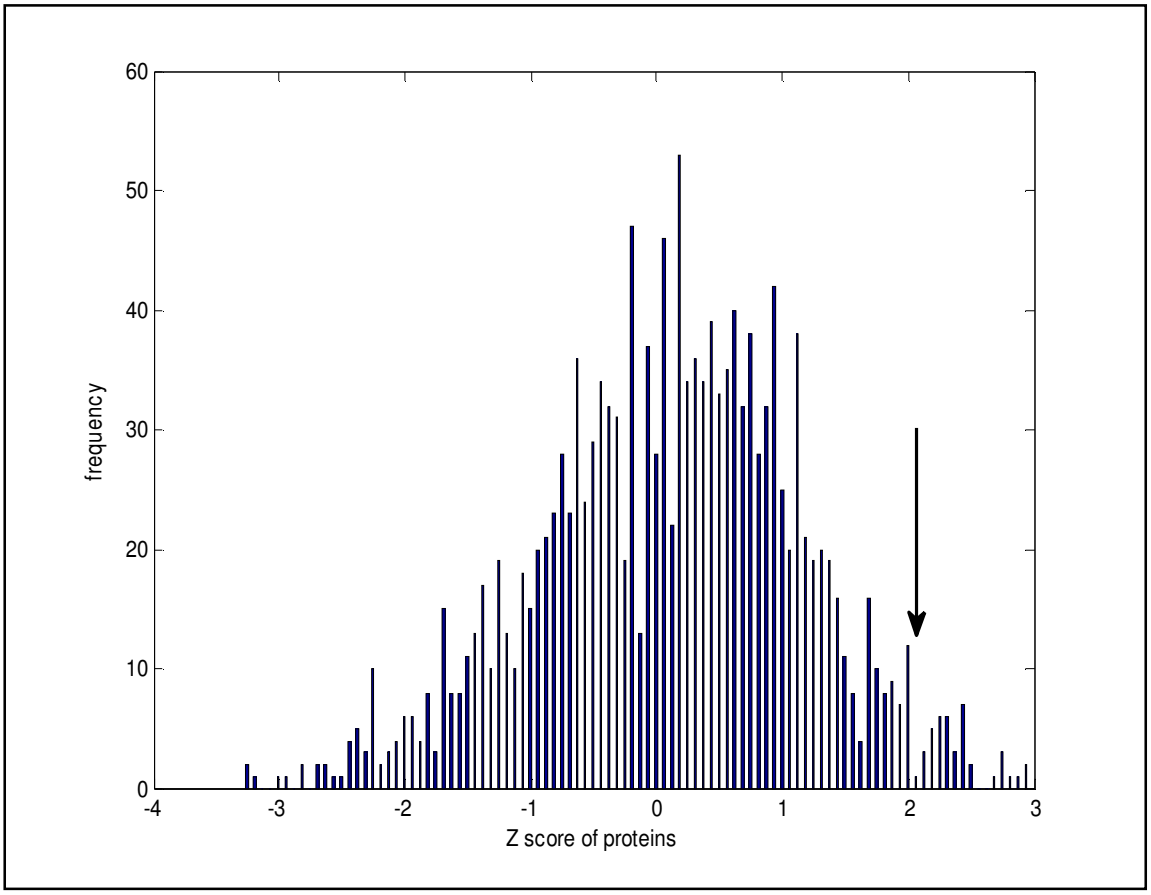

Figure 3.19. Distribution of the Z scores of the proteins in *snf1Δ* strain.

The list of the reporter proteins in each case are given in Table 3.17.

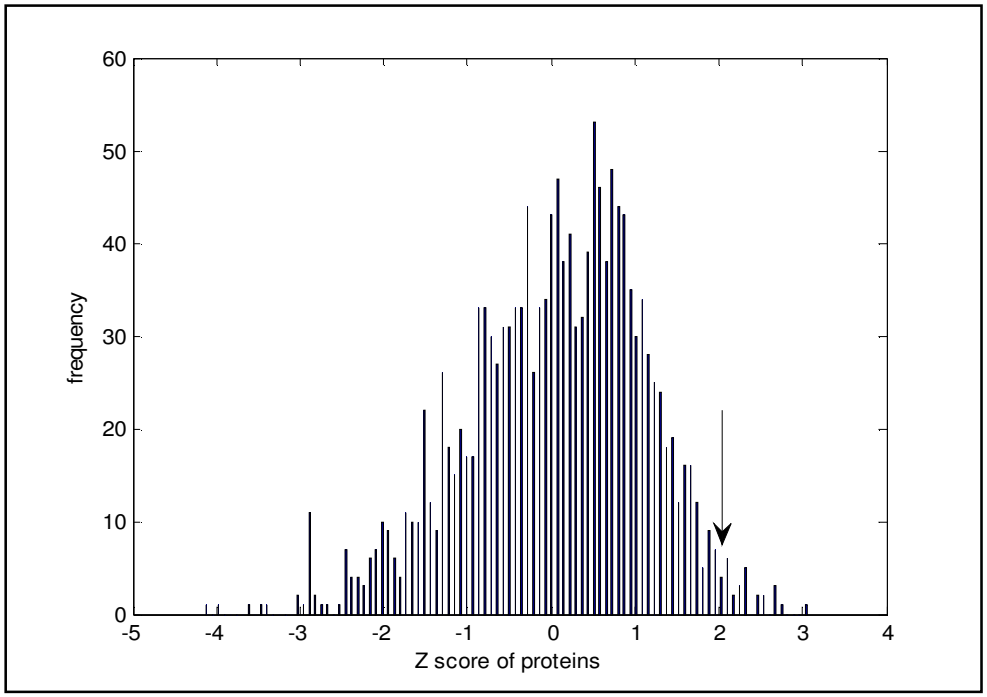

Figure 3.20. Distribution of the Z scores of the proteins in *snf1Δ snf4Δ* strain.

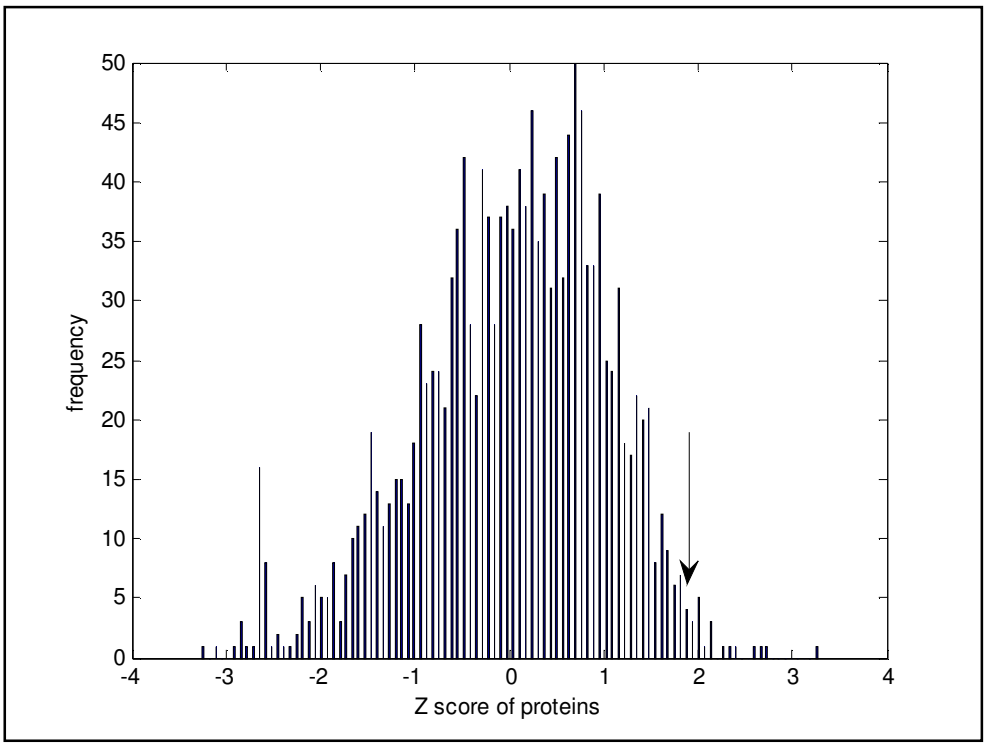

Figure 3.21. Distribution of the Z scores of the proteins in *snf4Δ* strain.

Table 3.17. Reporter proteins accepted as significant and their Z Scores

snf1Δ	Z Score	*snf1Δ (cont'd)*	Z Score	*snf4Δ*	Z Score	*snf1Δ snf4Δ*	Z Score
ARV1	2.96696	CTK1	2.17975	YCK2	3.28971	NOC4	3.07761
RUD3	2.96526	TPS1	2.16496	SWE1	2.71986	SAK1	2.77974
BUL2	2.87238	SKT5	2.16003	SIP2	2.63955	SMY1	2.70354
TPO3	2.8186	PFA4	2.1253	LEM3	2.61434	CCT4	2.67894
PTK2	2.77883	SHO1	2.12015	PHO4	2.38173	ARV1	2.67083
YMR086W	2.75328	MAK11	2.09321	REG1	2.31883	PHO4	2.55733
SMY1	2.74004	YFR017C	2.1987	BMH2	2.30099	ASC1	2.51971
FET3	2.68238			PDR5	2.17414	HSL1	2.47989
HSF1	2.52166			YMR196W	2.12018	BCY1	2.43325
GPI2	2.51566			FRE2	2.11662	SHS1	2.35448
PDR5	2.45812			BMH1	2.05021	SWE1	2.34214
FPS1	2.45474			SAK1	2.02785	BUL2	2.31131
PTP3	2.45242			GAL83	1.99982	RUD3	2.30621
SWE1	2.4315			FPS1	1.99465	PHO2	2.30405
JJJ1	2.42672			ACF4	1.98691	UTP18	2.27338
YKE4	2.41967			NOC4	1.97877	KRE33	2.24179
PHO4	2.40906			HSL1	1.96905	DIA2	2.2237
MGA1	2.37932			GTR2	1.93754	FRE2	2.18411
YPR115W	2.35564			SNT2	1.9122	VAB2	2.15045
SPO7	2.35282			TOS3	1.90668	EMG1	2.1351
DRE2	2.34109			PNC1	1.88348	VIP1	2.09432
YHR140W	2.33949			MEK1	1.87093	KRR1	2.08155
OSH3	2.33088			YFR017C	1.85096	ESF2	2.08105
RSP5	2.30957			VHS1	1.8409	MAK11	2.07002
SSF2	2.2994			TPK2	1.83646	CDC14	2.06821
CAN1	2.28447			KSP1	1.83592	PWP2	2.03804
YAK1	2.27523			CDC3	1.82628		
YGR153W	2.26963			CDC11	1.81227		
LEM3	2.25713			CMK2	1.81205		
CCT4	2.24589			PRR2	1.80874		
VHS1	2.2365			YMR086W	1.7748		
MBR1	2.22848			HSF1	1.771		
CDC3	2.20261			BCK1	1.75302		

3.6. Identification of the Key Transcription Factors by the Reporter Features Approach

The key transcription factors are also predicted by the reporter features algorithm discussed in the previous chapter. The key transcription factors and their overall Z scores are listed in Table 3.18. (See Section 2.6 for details of the algorithm). The lower threshold values are again determined by the use of histograms. The distribution of the overall Z scores of the transcription factors are shown in Figure 3.22, Figure 3.23 and Figure 3.24 for Snf1, Snf1Snf4 and Snf4 deletion mutants respectively.

The "Reporter Features Algorithm" was implemented in Matlab for the analysis. One disadvantage of the reporter features algorithm appeared to be that usually the TFs with high connectivity, namely the hubs of the network tend to appear as high scoring proteins. In other words, the aggregated, normalized Z score of a hub tends to be high due to the high number of its neighboring proteins. (See Section 2.6 for details) The hubs are usually TFs which play a role in response to stress conditions. Deletion of the genes encoding the components of one of the main elements of the glucose repression signaling mechanism is clearly a stress condition. Due to this reason, there is significant activation around these proteins. Thus, the high scoring hubs are trivial results. The main purpose is to find the TFs around which the significant activation is due to deletion of the subject genes and which have the possibility to change across a smaller set of conditions. These findings could be more easily related with the response of the pathway to the genetic perturbations made.

There are several treatments made to this method. Ideker *et al.* developed an algorithm for finding the most responsive sub-networks to genetic and environmental perturbations, by using simulated annealing. In their study, they also mentioned the disadvantage of the Z score assembling discussed above. They also proposed a treatment to the method by getting rid of the neighbors of the hubs with low z scores. (Ideker *et al.*, 2002) On the other hand, another treatment could be removing the hubs from the network prior to the calculations. But it is complicated how to define a hub, meaning this, how many neighbors does a TF need to be defined as a hub and what if the hub removed is indeed affected by the perturbation made and has an undiscovered relation with the glucose repression pathway.

In this study, the network is kept with the hubs but a treatment is made in the threshold determining step. The proteins are listed from the highest to the lowest score. When a TF with a lower number of neighbors than the TF below it appeared in the list for the first time, it was accepted as the first protein around which significant changes occurred. The lower threshold values are again determined by the use of histograms.

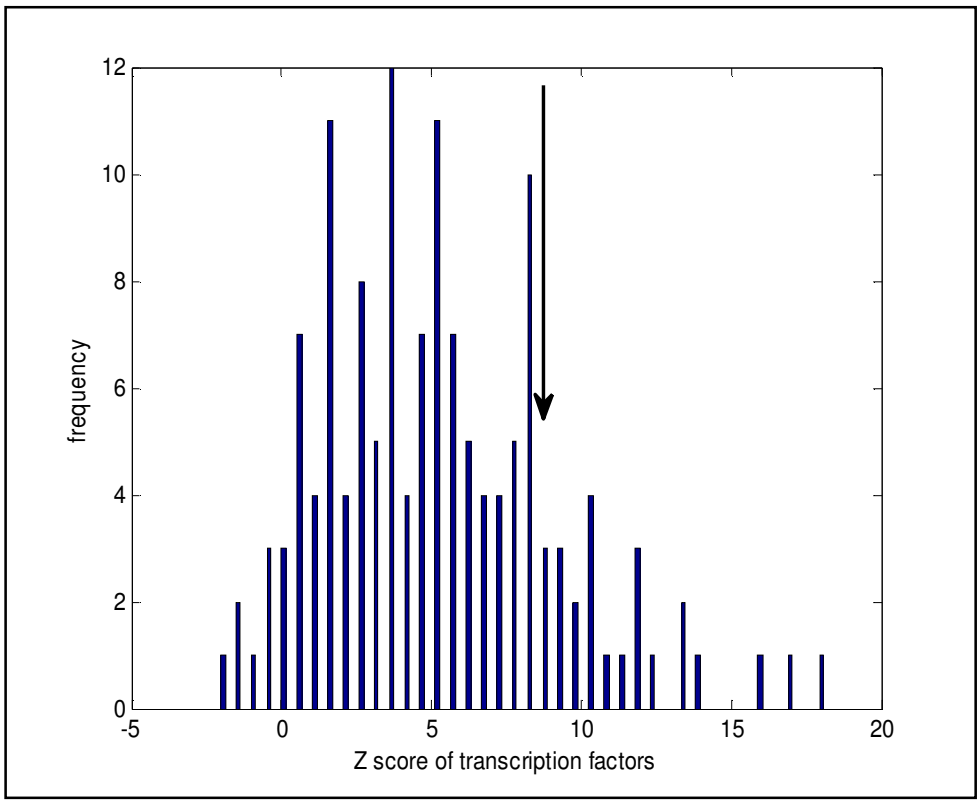

Figure 3.22. Distribution of the Z scores of the transcription factors in *snf1Δ*.

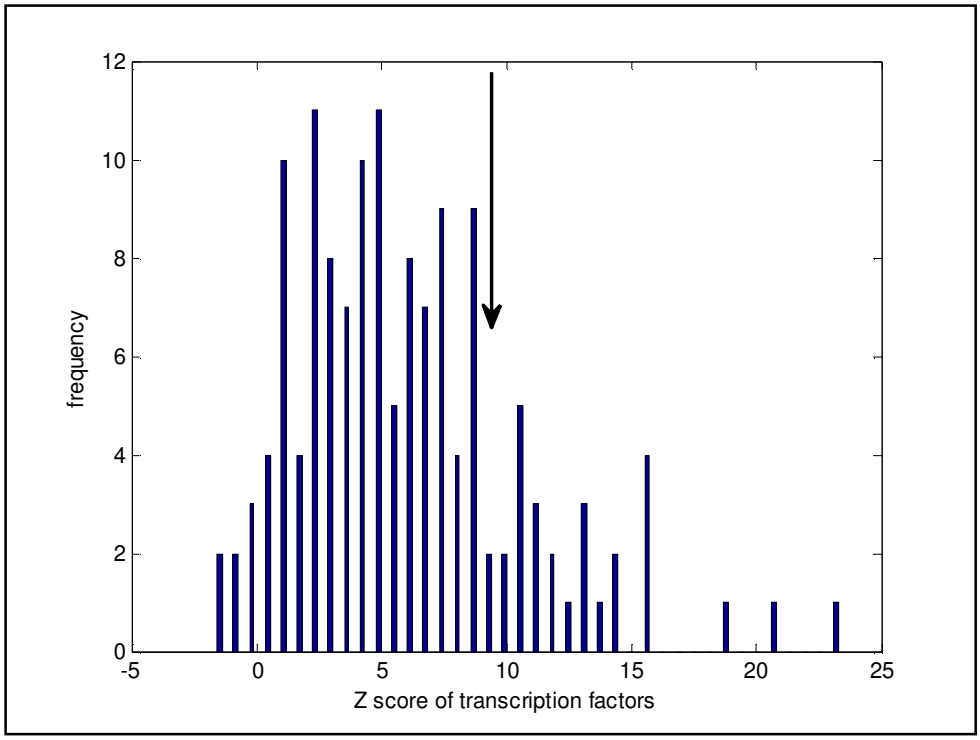

Figure 3.23. Distribution of the Z scores of the transcription factors in *snf1Δ snf4Δ*

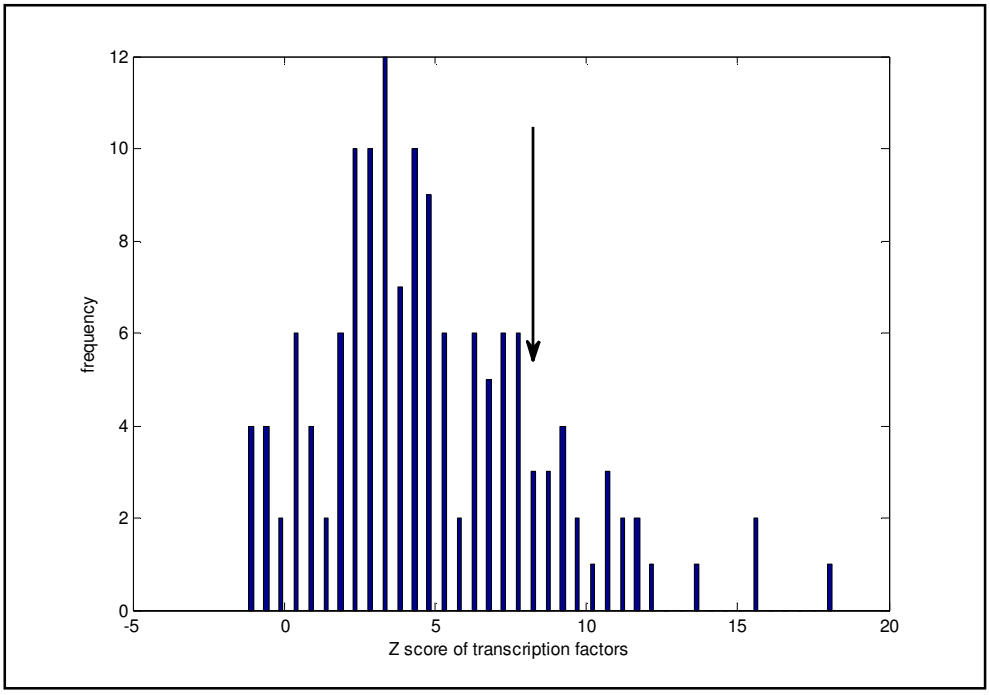

Figure 3.24. Distribution of the Z scores of the transcription factors in *snf4Δ*

Table 3.18. Reporter Transcription Factors accepted as significant and their Z Scores

snf1Δ	Z Score	*snf4Δ*	Z Score	*snf1Δ snf4Δ*	Z Score	*snf1Δ snf4Δ* (cont'd)	Z Score
HSF1	13.6622	HSF1	13.4408	MSN2	15.6758	XBP1	8.4405
STE12	13.4576	STE12	11.96	ARR1	15.5807	UPC2	8.4077
SWI4	13.2664	RPN4	11.7958	AFT1	15.5106	STP2	8.3795
MSN2	12.4793	INO4	11.5974	RPN4	15.5073	TOS8	8.3096
RPN4	12.0814	SWI4	11.1943	GCN4	14.4677	NRG1	7.9934
ARR1	11.8413	MSN2	11.1097	HSF1	14.1284	ECM22	7.8252
AFT1	11.6108	ARR1	10.8254	STE12	13.5779	FKH2	7.7234
MSN4	11.2965	AFT1	10.7049	RAP1	13.1388		
ADR1	10.7924	GCN4	10.6342	LEU3	12.9975		
TEC1	10.4652	ABF1	10.0748	MSN4	12.952		
PDR1	10.2624	ADR1	9.8268	INO4	12.5986		
INO4	10.1694	MSN4	9.5367	TEC1	12.0278		
RAP1	10.1415	RAP1	9.1932	PDR1	11.8048		
SKN7	9.7399	LEU3	9.1689	ABF1	11.4809		
GIS1	9.7362	MBP1	9.0516	SWI4	11.0731		
MBP1	9.5057	TEC1	8.9912	CAD1	10.9833		
GCN4	9.4599	YHP1	8.8177	ADR1	10.7442		
CAD1	9.1567	CRZ1	8.8067	MBP1	10.6799		
LEU3	8.9066	NRG1	8.612	ROX1	10.435		
ABF1	8.7732	CAD1	8.3021	SKN7	10.2574		
CST6	8.5993	GIS1	8.1968	CST6	10.2541		
PHD1	8.4247	MIG1	8.0937	STB5	10.1892		
		ROX1	7.9118	GIS1	9.7598		
		PHD1	7.8546	PHD1	9.4298		
		TOS8	7.716	YAP5	8.9905		
		MCM1	7.6982	CRZ1	8.8301		
		CBF1	7.6389	CIN5	8.7756		
		PDR1	7.5178	YHP1	8.7508		
				PDR3	8.6052		
				HAP4	8.5293		
				CBF1	8.4421		

3.7. Identification of the Linear Paths with a Maximum Path Length of 10

The calculations explained in Sections 3.2-3.4 are repeated by setting the maximum path length to 10 this time. This increased the number of linear paths identified and let a number of extra proteins to be involved in the sub-networks extracted. The number of the linear paths identified and the maximum and minimum average path scores are tabulated. (Table 3.19 and Table 3.20)

Table 3.19. The number of identified paths with maximum path length of 10

Strain	Number of paths
snf1Δ	17,942,301
snf4Δ	64,217,103
snf1Δ snf4Δ	18,181,783

Table 3.20. Path scores for maximum path length of 10

Strain	Max. Z score	Min. Z score
snf1Δ	2.095	-0.780
snf4Δ	2.040	-0.905
snf1Δ snf4Δ	2.813	-0.401

The histograms showing the distribution of the number of paths against the average path scores are plotted for each deletion mutant. One example for the cut-off detection process can be followed via Figure 3.25 and Figure 3.26 for the Snf1 deletion mutant. In Figure 3.25, the histogram covers all the 17,942,301 identified linear paths so it is not easy to see at which average paths score, the number of linear paths with that score or higher, significantly decreases. So the tail of the histogram is zoomed in and plotted in Figure 3.26. In Figure 3.26, the arrow shows the cut-off score 2.008 above which 222 linear paths are accepted as significant.

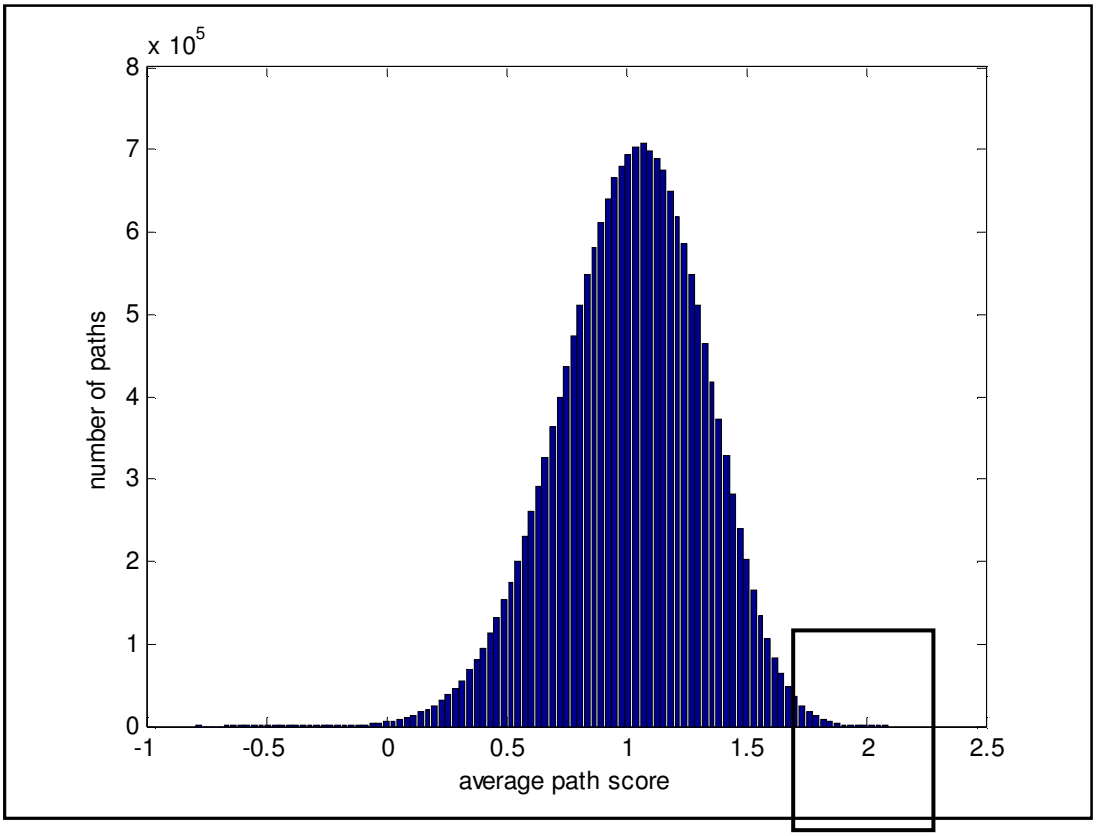

Figure 3.25. Distribution of the path scores in *snf1Δ* for a maximum path length of 10

The region shown inside the rectangular border is zoomed in and plotted in Figure 3.26. The zoomed in versions of the histograms plotted for *snf1Δ snf4Δ* and *snf4Δ* strains can also be seen in Figure 3.27 and Figure 3.28. The number of linear paths accepted as significant in each case are tabulated below. (See Table 3.21)

Table 3.21. The number of paths accepted as significant for maximum path length of 10

Strain	Number of significant paths
snf1Δ	222
snf4Δ	254
snf1Δ snf4Δ	237

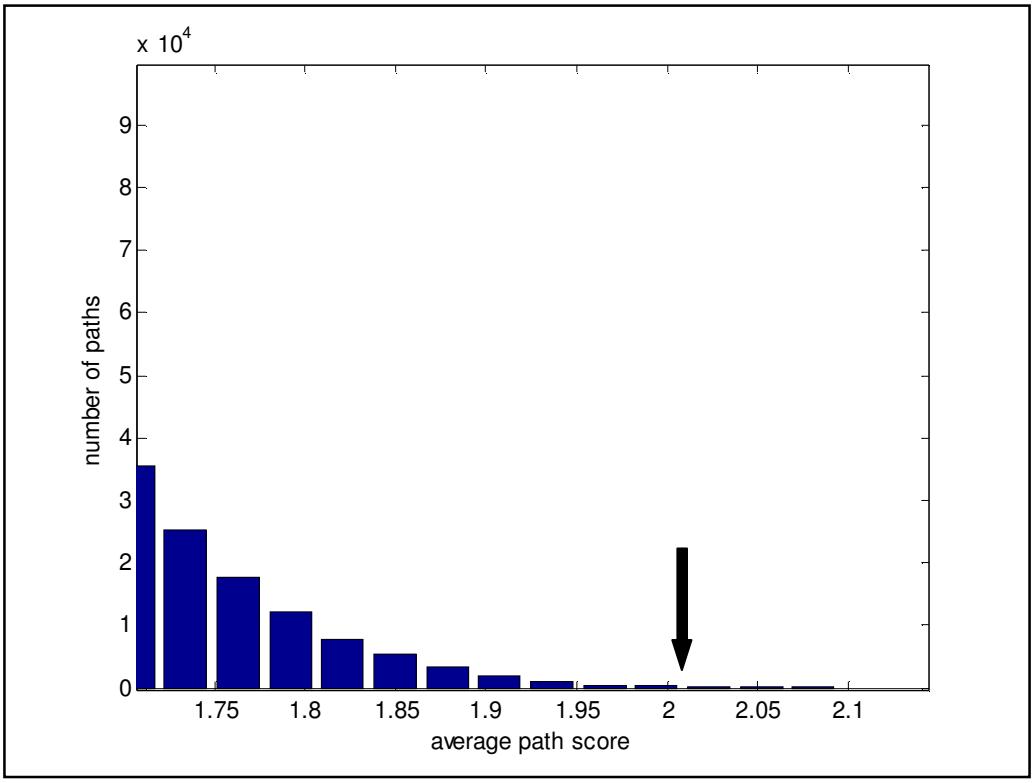

Figure 3.26. Distribution of the path scores in *snf1Δ* for a maximum path length of 10 (magnified version)

The resulting sub-networks needed to be narrowed for biologically meaningful further analysis. So, the individual signal transducing elements were determined as explained in Section 3.4. Basically, when a protein appeared in the significant paths significantly more than others, it is accepted as a key individual signal transducing element. The cut-off values for significant elements can be seen in Table 3.22 for each case.

Table 3.22. The cut-off values for significant individual elements

Strain	Threshold value
Snf1 Δ	46
Snf4 Δ	27
Snf1 Δ Snf4 Δ	38

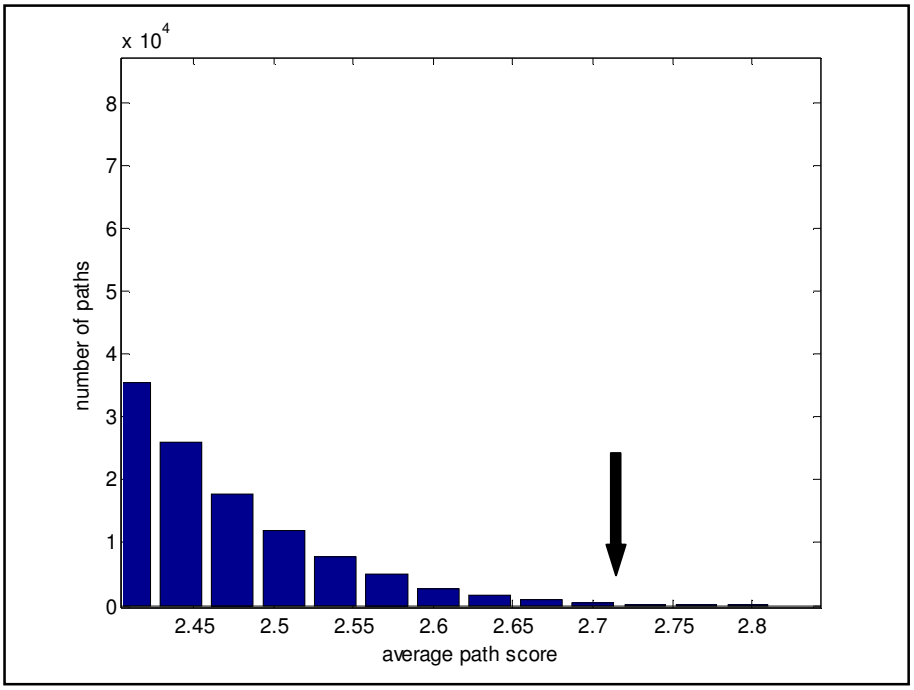

Figure 3.27. Distribution of the path scores in *snf1Δsnf4Δ* for a maximum path length of 10 (magnified version)

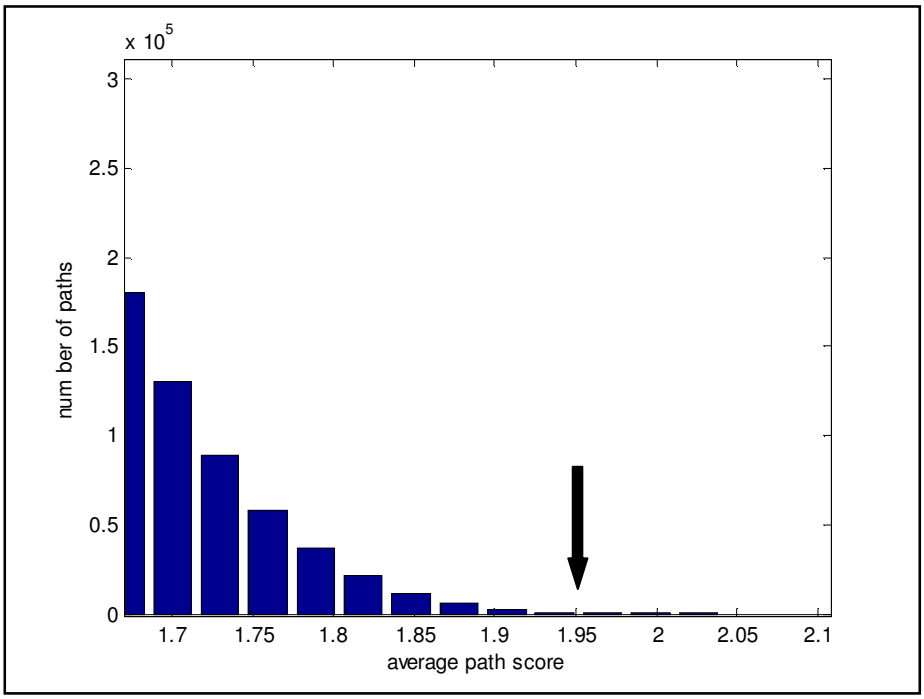

Figure 3.28. Distribution of the path scores in *snf4Δ* for a maximum path length of 10 (magnified version)

When constructing the resulting sub-networks shown in Figure 3.29, Figure 3.30 and Figure 3.31, the individual signal transducing elements are plotted as the nodes and the edges are their binary appearances in the significant linear signaling paths.

When the sub-networks extracted are investigated together, it is obvious that a number of proteins play role in the glucose repression signaling pathway common in all three deletion mutants when different path lengths are considered.

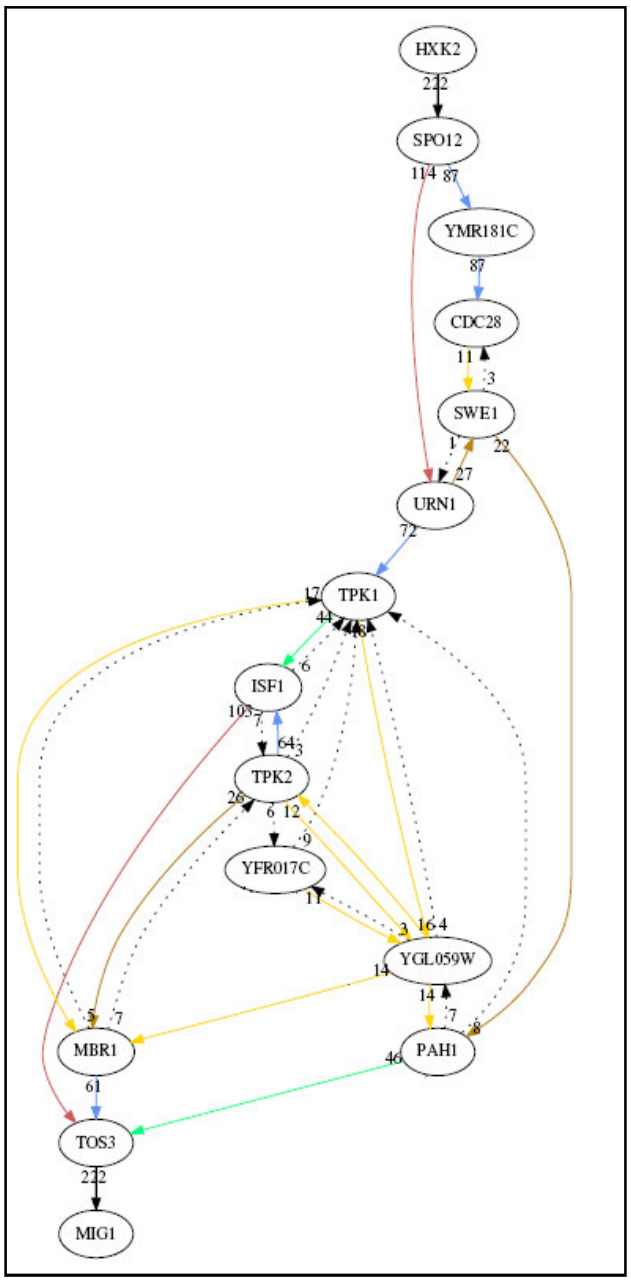

Figure 3.29. Glucose repression sub-network in *snf1Δ*

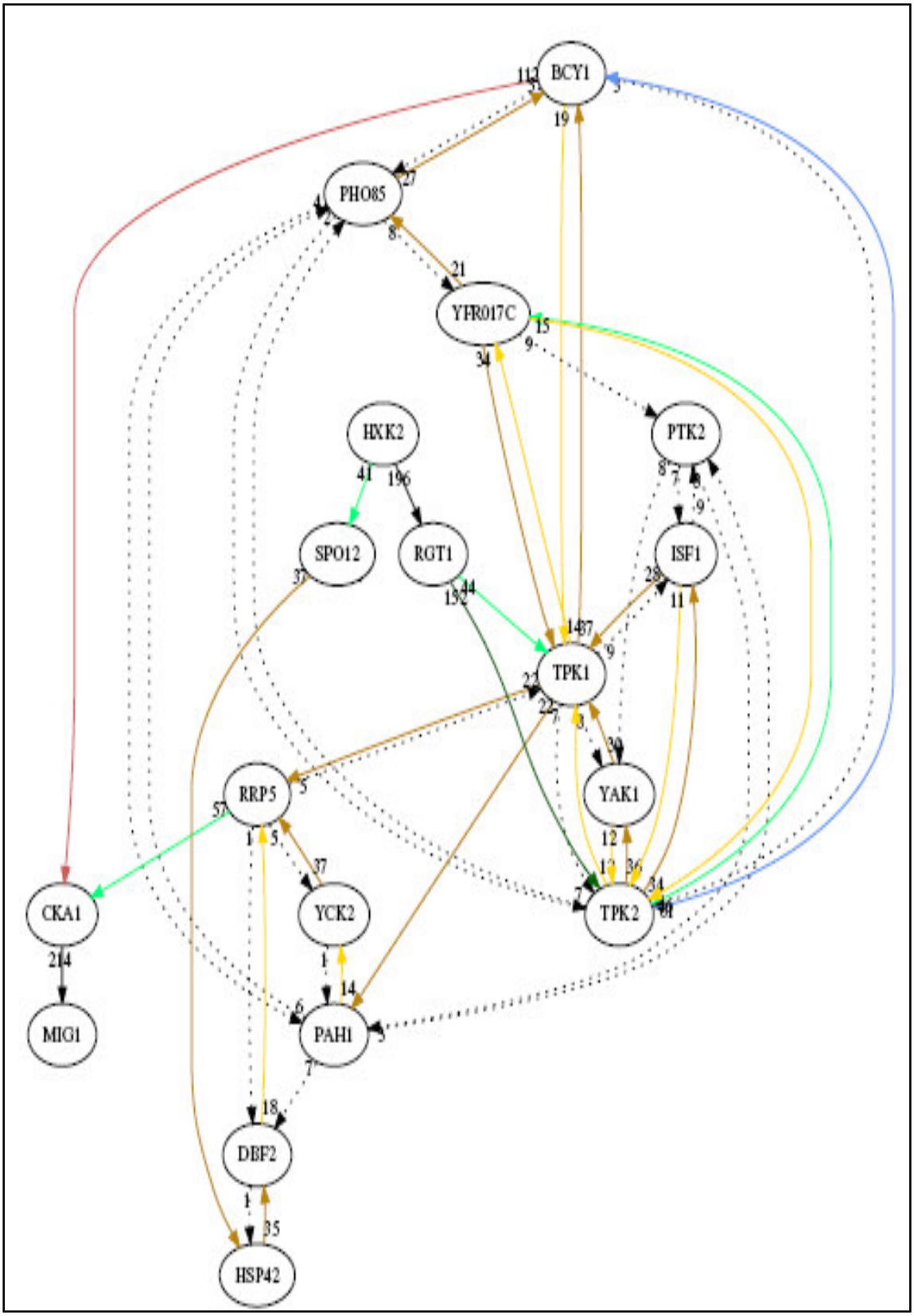

Figure 3.30. Glucose repression sub-network in *snf1Δ snf4Δ*

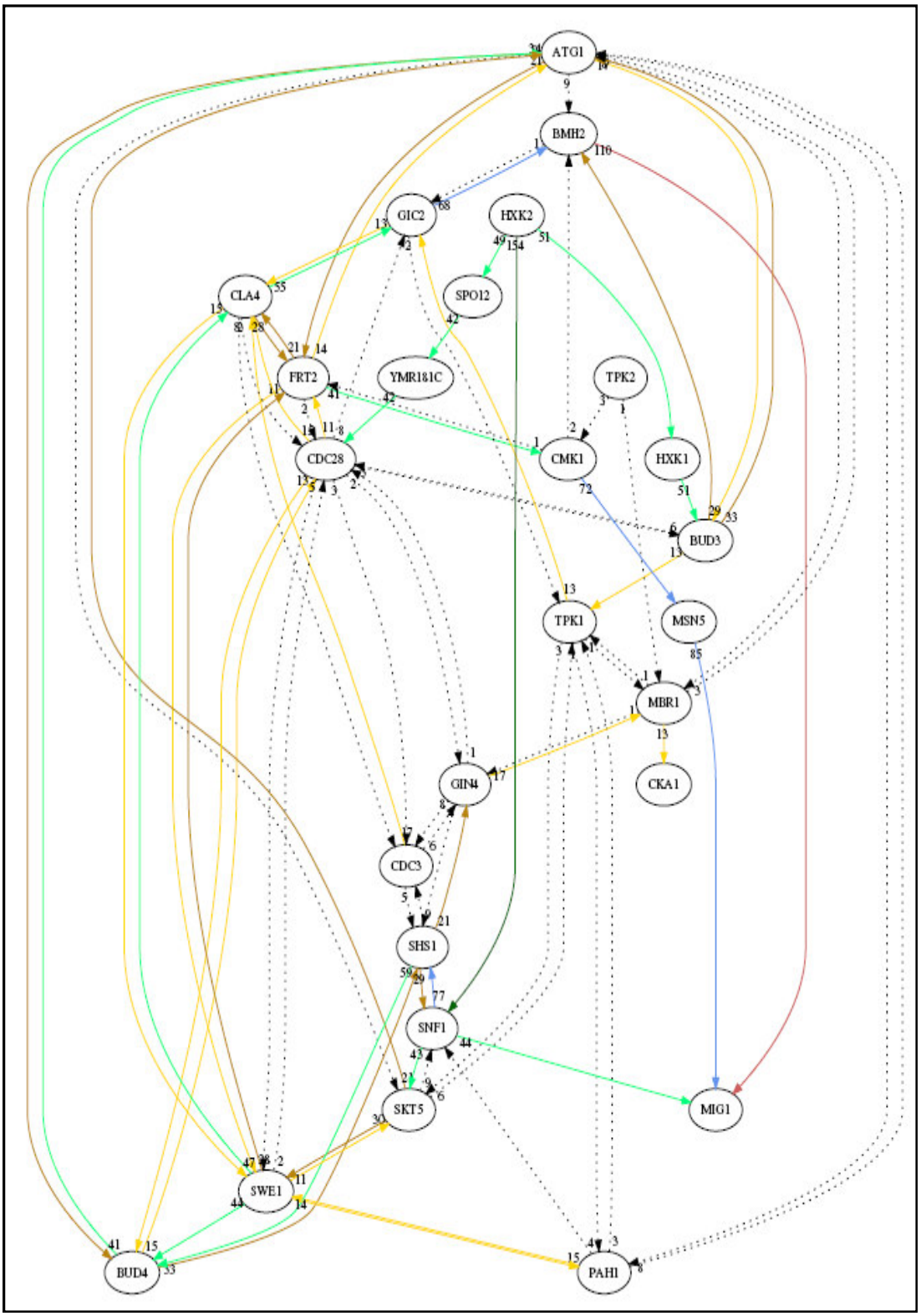

Figure 3.31. Glucose repression sub-network in *snf4Δ*

3.8. Comparison of the Sub-networks Predicted by Varying the Maximum Path Length Parameter

Table 3.23. Predicted common proteins by the individual signal transducing elements approach with varying maximum path length values

Snf1Δ strain	*Snf4Δ* strain	*Snf1Δ Snf4Δ* strain
SPO12	SPO12	SPO12
URN1	PAH1	RGT1
YMR181C	SNF1	TPK1
SWE1	SKT5	BCY1
TPK1	SHS1	TPK2
TPK2	SWE1	RRP5
YFR017C	BUD4	PHO85
CDC28	FRT2	ISF1
PAH1	ATG1	YFR017C
MBR1	HXK1	CKA1
ISF1	YMR181C	DBF2
TOS3	MBR1	HSP42
YGL059W	BUD3	
	CKA1	
	TPK1	
	CLA4	
	CDC28	
	TPK2	
	GIC2	
	CMK1	
	BMH2	
	MSN5	

Table 3.24. Differences in the predicted proteins for varying maximum path length values; their function and process GO terms in *Snf1Δ*

Max. Path Length: 9		
Gene Name	Process GO Term	Function GO Term
CMK1	Signal transduction, protein modification process	Protein kinase activity, transferase activity
ORM2	response to stress, response to chemical stimuli	molecular function unknown
PTK2	Cell cycle, transport, cellular homeostasis	Protein kinase activity, transferase activity

When the resulting active sub-networks in the Snf1 deletion mutant predicted by the second approach with a maximum path length of either 9 or 10, are considered together; it is easily seen that some proteins play a key role collecting and delivering the signal from various proteins to others. TPK1 and TPK2 are two proteins which show this type of behavior. TPK family genes, namely the TPK1, TPK2 and TPK3 genes encode isoforms of the catalytic subunit of cAMP-dependent protein kinase (PKA), the effector kinase of the Ras-cAMP signaling pathway.

Through phosphorylation of various targets, PKA activity regulates processes such as cell growth and response to nutrients and stress. (Santangelo, 2006; Estruch, 2000) Carbohydrate utilization process is one of these processes related to response to nutrients whereas the glucose depleted media in which the cells are cultured clearly indicates a stress condition. So it can be concluded that TPK1 and TPK2 proteins play key roles in glucose repression mechanism in the Snf1 deletion mutant. This is also the case for the other deletion mutants.

The transcription factors regulated by these proteins, in other words the substrates of PKA include Rap1p, Hsf1p, Adr1p, Msn2p, Msn4p and Ssn2p. (Klein and Struhl, 1994; Ferguson et al, 2005; Cherry et al., 1989; Gorner et al., 1998; Gorner et al., 2002, Chang et al., 2004)

Table 3.25. Differences in the predicted proteins for varying maximum path length values; their function and process GO terms in *Snf4Δ* strain

Snf4Δ strain					
Max. Path Length:9			Max. Path Length:10		
Gene Name	Process GO Term	Function GO Term	Gene Name	Process GO Term	Function GO Term
SIP2	Signal transduction, protein modification process, response to stress	Catalytic activity	MGA1	*Cannot be mapped to a term*	DNA binding, transcriptional regulator activity
BCY1	*Cannot be mapped to a term*	*Cannot be mapped to a term*	CDC3	cytokinesis , cell budding, anatomical structure morphogenesis, cell wall organization and biogenesis	Lipid binding, structural molecule activity
YGL059W	protein modification process	Catalytic activity	GIN4	cytokinesis , cell budding, anatomical structure morphogenesis, cytoskeleton organization and biogenesis, organelle organization and biogenesis, protein modification process	Protein kinase activity, transferase activity

When the key TFs predicted to response to the genetic perturbation made are also taken into account; as in Table 3.18, five of these TFs; Rap1p, Hsf1p, Adr1p, Msn2p and Msn4p are determined as key TFs which regulate the most differentially expressed genes in the Snf1 deletion mutant. Hsf1p is a trimeric heat shock TF. It is posttranslationally regulated by the PKA. It regulates hundreds of genes among which there are genes playing roles in energy generation and carbohydrate metabolism. (Hahn et al., 2004; Eastmond and Nelson, 2006) On the other hand, Msn2p and Msn4p show similar characteristics. They are TFs which are activated under stress conditions. (Schmitt and McEntee, 1996) DNA binding by Msn2p is thought to be regulated by stress and by the cAMP-dependent protein kinase. Under non-stress conditions, the Msn2p is located in the cytoplasm. (Gorner et al.,

1998) This localization is regulated by Bmh2p (Beck and Hall, 1999). It is also known that Msn2p is negatively regulated by also Snf1 phosphorylation. (Kaida et al., 2002)

Table 3.26. Differences in the predicted proteins for varying maximum path length values; their function and process GO terms in *Snf1ΔSnf4Δ* strain

Snf1ΔSnf4Δ strain					
Max. Path Length:9			Max. Path Length:10		
Gene Name	Process GO Term	Function GO Term	Gene Name	Process GO Term	Function GO Term
PHO4	Response to stress	DNA binding, transcriptional regulator activity	PTK2	Transport, cell cycle,cellular homeostasis	Protein kinase activity, transferase activity
			YAK1	protein modification process	Protein kinase activity, transferase activity
			YCK2	protein modification process, transport, vesicle-mediated transport, cytokinesis, membrane organization and biogenesis, response to chemical stimulus, anatomical structure morphogenesis	Protein kinase activity, transferase activity
			PAH1	Cellular respiration, generation of precursor metabolites and energy	Hydrolase activity

Rap1p is an activator/repressor protein targeting over 185 genes. Some of these genes encode glycolytic enzymes. (Lieb et al., 2001) Another substrate of PKA; Adr1p is required for transcription of the glucose repressed gene ADH2 and genes required for ethanol, glycerol and fatty acid utilization. (Verdone et al., 2002) It is known that Snf1p positively regulates Adr1p binding in the absence of glucose while the upstream phosphotase complex of Snf1p; the Glc7p/Reg1p complex inhibits the ability of Adr1p.

(Young et al., 2002) Direct phosphorylation by protein kinases, such as cAMP-dependent protein kinase, influences Adr1p expression and activity. (Cherry et al., 1989)Due to these evidences stated above, it is not surprising to observe TPK1 and TPK2 in most of the reconstructed sub-networks showing similar effects with Snf1p and thus being a potential alternative for it in deletion mutants of the Snf1 kinase complex.

An interesting binary signal transducing element in the Snf1 deletion mutant includes the Swe1p and Cdc28p. Cdc28p is the main cell cycle cyclin-dependent kinase (CDK). Its posttranslational modification is done by Cak1p and Swe1p on threonine and tyrosine sites respectively. Phosphorylation by Swe1p kinase inhibits mitotic CDK activity and hence entry into mitosis (Ross et al., 2000; Booher et al., 1993).

Swe1p itself also undergoes sequential phosphorylations by a variety of kinases including the Cla4p. (Harvey et al., 2005) This couple seems not to play a role in only the Snf1Snf4 deletion mutant where there is no Snf1 kinase complex activity. In this case the Cdc28p is not inhibited by the phosphorylation of Swe1p and thus the cell enters into mitotic cycle. The reason for delay during entry into mitosis is generally that the cell experiences an environmental disturbance and that disrupts the bud formation needed for mitosis. When both SNF1 and SNF4 genes are deleted, it is possible that there is a significant breakdown in the signaling mechanism which transduces the environmental disturbance (in this case; glucose starvation) so the cell cannot response to the disturbance.

A similar situation is observed in Snf4 mutant. The Cla4p kinase is known for down regulation of Swe1p and thus letting the cell into mitosis. In the Snf4 deletion mutant, the signal is transduced between Swe1p, Cla4p and Cdc28p which again means both activation and deactivation of Cdc28p. This phenomenon points a disruption in the signaling mechanism.

Other important proteins in the Snf1 deletion mutant are MBR1 and ISF1. They are not directly linked; instead they seem like substantial proteins playing roles in the same step of the signal transduction. In the Snf1Snf4 deletion mutant, Isf1p takes role whereas in the Snf4 deletion mutant Mbr1p plays a role in the repression mechanism. Mbr1p involves in mitochondrial functions and stress response. Over expression of this protein suppresses

the growth defects of Hap2, Hap3 and Hap4 mutants. (Daignan-Fornier et al., 1994; Reisdorf et al., 1997) Isf1p is not completely identified yet but it is similar to the Mbr1p playing a role in suppressing the growth defects in Hap complex mutants. Its expression is under glucose control. (Daignan-Fornier et al., 1994)

Actually the point on which the most of the curiosity is focused, was the proteins which may play a role in phosphorylation of Mig1p. This potential kinase was predicted to be different in the three cases. In a study where the phosphorylation ability of Pak1p, Tos3p and Cka1p of the Snf1 complex are determined, it was shown that Tos3p and Cka1p isolated from yeast were able to phosphorylate the Mig1p in vitro. (Nath et al., 2003) Up to date, Tos3p is known to be functionally redundant with Elm1p and Pak1p for the phosphorylation of Snf1p. These evidences also support the idea that Tos3p can be a potential alternative for phosphorylation of Mig1p in the absence of the Snf1p.

In the Snf4 deletion mutant, Snf1p also takes a role in the phosphorylation of Mig1p indicating that Snf1 still plays a role in phosphorylating the final protein, namely the Mig1p. In this mutant, Msn5p and Bmh2p also play a role in transducing the signal to the final element Mig1p. Msn5 is an importin β homolog implicated in glucose repression. Msn5p was initially identified as a high-copy suppressor of a Snf1 mutation (Estruch and Carlsson, 1990; Estruch and Carlsson, 1993). In earlier studies, a Snf1 deletion mutant was observed as unable to grow on sucrose and this observation was explained based on the deficiency of the cell in the phosphorylation of the Mig1p due to the deletion of the SNF1 gene. This caused Mig1p always to be in the nucleus, repressing gluconeogenic genes' expression (DeVit *et al*., 1997). High levels of the Mig1p export receptor would be expected to restore growth by exporting enough Mig1 to relieve repression. In another study, it was reported that Msn5p indeed mediated the export of Mig1p in response to glucose removal. In the same study, it was also reported that the recognition sites of the Snf1 kinase and the export receptor Msn5 were the same in the Mig1p. (DeVit and Johnston, 1999)

One interesting couple in the Snf1Snf4 deletion mutant is the Pho4p-Pho85p binary signal transducing element. Pho4p is a transcription factor which plays a role in response to phosphate starvation. It is activated by the phosphorylation by the Pho80p-Pho85p

complex and when phosphate is abundant it is exported from the nucleus by the export receptor protein Msn5p like the Mig1p. (Kaffman *et al.*, 1998)

In the Snf1Snf4 deletion mutant DBF2 protein takes place in both of the predicted sub-networks. It is known that DBF2 is a part of the CCR4 transcription factor complex which is regulated by glucose (Liu et al., 1997) The CCR4 protein from *Saccharoymces cerevisiae* affects the expression of a number of genes and processes. The target proteins of the cell cycle-regulated protein kinase DBF2 are not clearly identified yet. On the other hand; it has been reported that CCR4 is required for full derepression of *ADH2* and other non- fermentative genes under glucose-derepressed conditions (Denis, 1984; Denis and Malvar, 1990).

In the Snf4 deletion mutant, the highest number of appearance of signal transduction to Mig1p is observed from the Bmh2p. The Bmh2p is involved in the control of proteome at post-transcriptional level. Some of the process gene ontology terms related to Bmh2p involves glycogen metabolic process, Ras protein signal transduction, signal transduction during filamentous growth (van Hemert *et al.*, 2001)

4. CONCLUSIONS AND RECOMMENDATIONS

4.1. Conclusions

In this study, the signaling network of yeast was reconstructed based on SPA method. The SPA algorithm helped to remove the protein-protein interactions which are predicted to be irrelevant to the signaling mechanism in yeast. The criteria for a protein to be included in the network were based on their GO categories. The SPA pre-processed signaling network of yeast contained was represented as a directed unipartite graph in which the proteins were the nodes and the interaction between them were the edges. The graph contained 1521 nodes and 11242 edges. The maximum of the shortest path lengths between proteins was determined as 10, which corresponded to the diameter of the reconstructed network. All the linear paths starting from the Hxk2p and ending at the transcriptional repressor Mig1p were identified by NetSearch algorithm in the Snf1, Snf4 and Snf1Snf4 deletion mutants with the parameters minimum path length being 2 and the maximum path length being 9 and 10 respectively. The proteins which were encoded by the deleted gene were eliminated from the network in each case. The average path score algorithm was used for scoring the identified linear paths which was based on the Z scores of the genes encoding the proteins involved in the linear paths.

When the maximum path length parameter was given as 9, the highest scoring linear path was identified in the Snf1Snf4 deletion mutant. In the meantime, the average path score of the minimum scoring path in this mutant was higher than the other deletion mutants showing the degree of differential expression with respect to the wild type cell when two constituents of the Snf1 kinase complex were deleted together. The cut off values above which the linear paths would be accepted as significant were determined by the use of histograms in each case. The linear paths identified as significant in each case contained 98, 131 and 90 different proteins for the Snf1, Snf4 and Snf1Snf4 deletion mutants respectively. The dimensions of the predicted sub-networks were further reduced by two different approaches. In the first approach, binary signal transducing elements which transduced the signal from one to another were taken as unit elements of the sub-network. The binary elements that appeared in the significant linear paths significantly

more times than the other proteins couples were accepted as more likely to play role in the glucose repression pathway of the mutants. In the second approach, the unit elements were taken as the individual proteins. The sub-networks predicted by these two different approaches resulted in minor differences in terms of the proteins involved in the sub-networks. Spo12p, Tpk1p, Tpk2p were commonly predicted by two approaches in all the deletion mutants. Most of the predicted proteins were involved in protein modification process and had the kinase, hyrolase or transferase activity which indicated their potential signaling role. The second approach was used for further analysis with varying parameters as discussed in Section 3.8.

When the maximum path length was set to 10, the number of linear paths accepted as significant was 222, 254 and 237 in Snf1, Snf4 and Snf1Snf4 deletion mutants respectively. The cut off values for the individual signal transducing elements to be accepted as significant was determined to be 46, 27 and 38 for Snf1, Snf4 and Snf1Snf4 deletion mutants respectively. The proteins to which the signal is transduced from Hxk2p and the proteins from which the signal is tranduced to the Mig1p did not differ in any of the deletion mutants compared with the case in which the maximum path length was set as 9. Tpk1p, Tpk2p, Swe1p, Pah1p together with the protein encoded by the YGL059W (Pkp2p) are predicted to have important roles in the pathway in the Snf1 deletion mutant. In the Snf1Snf4 deletion mutant, Tpk1p, Tpk2p, Yak1p, Yck2p, Rrp5p, Pah1p, Isf1 and Bcy1p are the proteins which are involved in transducing the signal from various proteins to various targets whereas in the Snf4 deletion mutant, it is seen that Snf1p still exists in the predicted sub-network together with a number of different proteins more than the other cases.

The key transcription factors were also determined by the Reporter Features Algorithm around which the biggest amount of transcriptional response due to the deletion of the genes encoding the constituents of one of the key components of the glucose repression pathway in yeast has occurred. The key transcription factors involved Hsf1p, Ste12p, Swi4p, Msn2p, Adr1p, Rap1p, Cad1p and Msn4p among others. Some of these TFs' interactions with the predicted pathway elements were identified and documented in Section 3.8.

The Reporter Features algorithm was also used for the detection of the reporter proteins around which significant changes occurred due to the deletion of the subject genes. The key proteins predicted with this approach involved Ptk2p, Swe1p, Mbr1p, Yck2p, Bmh2p, Bcy1p, Shs1p, Fre2p, Tos3p, Tpk2p and Kre33p as the overlapping proteins with the ones predicted by the average path scoring algorithm.

4.2. Recommendations

Although Mig1p is known to be the repressor protein associated with the glucose repression signaling pathway of yeast, it is possible that other transcriptional repressor or activator proteins are also involved in the pathway. So the linear pathway analysis can be carried out by ending at other proteins than Mig1p. The potential final proteins can be predicted by transcriptome and binding site analysis. The transcription factors that are known to regulate the glucose repressible genes and that are also activated due to the deletion of the genes encoding the Snf1 kinase components can be predicted as potential final proteins.

The study can also be further developed by integration with proteome analysis. The phosphorylation levels and kinase activities of the predicted proteins can be monitored and the results of the current study can be validated.

REFERENCES

Alberts, B. A., A. Johnson, J. Lewis, M. Raff, K. Roberts and P. Walter, 2002, *Molecular Biology of the Cell*, 4th edition, Ch. 15, Garland Science, New York.

Arga, K. Y., Z. İ. Önsan, B. Kırdar, K. Ö. Ülgen, J. Nielsen., 2007, "Understanding Signaling in Yeast: Insights From Network Analysis", *Biotechnology and Bioengineering*, Vol. 97, No. 5, pp. 1246-1258.

Ashrafi, K., S. S. Lin, J. K. Manchester, J. I. Gordon, 2000, "Sip2p and Its Partner Snf1p Kinase Affect Aging in *S. cerevisiae*", *Genes Dev*, Vol. 14, pp. 1872-1885.

Bailey, J. E., 2001, "Complex Biology with no parameters", 2001, *Nature Biotechnology,* Vol. 19, pp. 503 - 504

Beck, T. and M. N. Hall, 1999, "The TOR signalling pathway controls nuclear localization of nutrient-regulated transcription factors", *Nature*, Vol. 402(6762), pp. 689-692

Begley, T. J., A. S. Rosenbach, T. Ideker , L. D. Samson, 2002, "Damage Recovery Pathways in Saccharomyces cerevisiae Revealed by Genomic Phenotyping and Interactome Mapping" *Mol Cancer Res.,* Vol. 1, pp. 103–112.

Booher R. N. et al, 1993, "Properties of Saccharomyces cerevisiae Wee1 and its differential regulation of p34CDC28 in response to G1 and G2 cyclins", *EMBO J.,* Vol. 12(9), pp. 3417-3426

Carling, D., K. Aguan, A. Woods, A. J. M. Verhoeven, R. K. Beri, C. H. Brennan, C. Sidebottom, M. D. Davison, J. Scott, 1994, "Mammalian Amp-Activated Protein Kinase is Homologous to Yeast and Plant Protein Kinases Involved in the Regulation Of Carbon Metabolism", *J. Biol. Chem.,* Vol. 289, pp. 11442–11448.

Carlson, M., 1999, "Glucose Repression in Yeast", *Current Opinion in Microbiology*, Vol.2, pp. 202-207.

Carlson, M., B. C. Osmond, L. Neigeborn, D. Botstein, 1981, "Mutants of Yeast Defective in Sucrose Utilization", *Genetics*, Vol. 98, pp. 25–40.

Celenza, J. L., F. J. Eng, M. Carlson, 1989, "Molecular Analysis of the SNF4 Gene of Saccharomyces cerevisiae: Evidence for Physical Association of the SNF4 Protein with the SNF1 Protein Kinase", *Mol. Cell. Biol.*, Vol. 9, pp. 5045– 5054.

Celenza, J. L., M. Carlson, 1989, "Mutational Analysis of the Saccharomyces cerevisiae SNF1 Protein Kinase and Evidence for Functional Interaction with the SNF4 Protein", *Mol. Cell. Biol.*, Vol. 9, pp. 5034–5044.

Chang YW, S. C. Howard and P. K. Herman, 2004, "The Ras/PKA signaling pathway directly targets the Srb9 protein, a component of the general RNA polymerase II transcription apparatus", *Mol. Cell*, Vol. 15(1), pp. 107-116

Cherry, J. R., T. R. Johnson, C. Dollard, J. R. Shuster and C. L. Denis, 1989, "Cyclic AMP-dependent protein kinase phosphorylates and inactivates the yeast transcriptional activator ADR1", *Cell*, Vol. 56(3), pp. 409-419

Ciriacy, M., 1977, "Isolation and Characterization of Yeast Mutants Defective in Intermediary Carbon Metabolism and in Carbon Catabolite Derepression", *Mol. Gen. Genet.*, Vol. 154, pp. 213–220.

Cullen, P. J., G. F. Sprague, Jr., 2000, "Glucose Depletion Causes Haploid Invasive Growth in Yeast", *Proc. Natl. Acad. Sci.*, Vol. 97, pp. 13619-13624.

Daignan-Fornier, B., C. C. Nguyen, P. Reisdorf, B. Lemeignan and M. Bolotin-Fukuhara,1994, "MBR1 and MBR3, two related yeast genes that can suppress the growth defect of hap2, hap3 and hap4 mutants", *Mol. Gen. Genet.*, Vol. 243(5), pp. 575-583

Denis, C. L. and T. Malvar, 1990, "The CCR4 gene from Saccharomyces cerevisiae is required for both nonfermentative and spt-mediated gene expression", *Genetics*, Vol. 124, pp. 283–291.

Denis, C. L., 1984, "Identification of new genes involved in the regulation of yeast alcohol dehydrogenase II", *Genetics*, Vol. 108, pp. 833–834.

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, pp. 1231-1241

DeVit, M. J., J. A. Waddle and M. Johnston, 1997, "Regulated nuclear translocation of the mig1 glucose repressor", *Mol Biol Cell*, Vol. 8, pp. 1603–1618.

Eastmond, D. L. and H. C. Nelson, 2006, "Genome-wide analysis reveals new roles for the activation domains of the Saccharomyces cerevisiae heat shock transcription factor (Hsf1) during the transient heat shock response", *J. Biol. Chem.*, Vol. 281(43), pp. 32909-32921

Entian, K.-D., F. K. Zimmermann, 1982, "New Genes Involved in Catabolite Repression and Derepression in the Yeast Saccharomyces cerevisiae", *J. Bacteriol.*, Vol. 151, pp. 1123–1128.

Erickson, J. R., M. Johnston, 1993, "Genetic and Molecular Characterization of GAL83: Its Interaction and Similarities with Other Genes Involved in Glucose Repression in Saccharomyces cerevisiae", *Genetics*, Vol. 135, pp. 655–664.

Estruch, F. and M. Carlson, 1990, "Increased dosage of the MSN1 gene restores invertase expression in yeast mutants defective in the SNF1 protein kinase", *Nucleic Acids Res*, Vol. 18, pp. 6959–6964.

Estruch, F. and M. Carlson, 1993, "Two homologous zinc finger genes identified by multicopy suppression in a SNF1 protein kinase mutant of Saccharomyces cerevisiae", *Mol Cell Biol*, Vol. 13, pp. 3872–3881.

Estruch, F., 2000, "Stress-controlled transcription factors, stress-induced genes and stress tolerance in budding yeast", *FEMS Microbiol Rev*., Vol. 24, pp. 469-486.

Ferguson, S. B., E. S. Anderson, R. B. Harshaw, T. Thate, N. L. Craig and H. C. M. Nelson, 2005, "Protein kinase A regulates constitutive expression of small heat-shock genes in an Msn2/4p-independent and Hsf1p-dependent manner in Saccharomyces cerevisiae", *Genetics*, Vol. 169(3), pp. 1203-1214

Flick, J. S. and M. Johnston, 1992, "Analysis of Ursg-Mediated Glucose Repression of the GAL1 Promoter of Saccharomyces cerevisiae", *Genetics*, Vol. 130, pp. 295–304.

Gancedo, J.M., 1998, "Yeast Carbon Catabolite Repression", *Microbiology and Molecular Biology Reviews*, Vol. 62, No. 2, pp. 334-361.

Gorner, W., E. Durchschlag, M. T. Martinez-Pastor, F. Estruch, G. Ammerer, B. Hamilton, H. Ruis and C. Schüller, 1998, "Nuclear localization of the C2H2 zinc finger protein Msn2p is regulated by stress and protein kinase A activity", *Genes Dev.*, Vol. 12(4), pp. 586-597

Gorner, W., E. Durchschlag, J. Wolf, E. L. Brown, G. Ammerer, H. Ruis and C. Schüller, 2002, "Acute glucose starvation activates the nuclear localization signal of a stress-specific yeast transcription factor", *EMBO J.*, Vol. 21(1-2), pp. 135-144

Guldener, U., M. Munsterkotter, G. Kastenmuller, N. Strack, J. van Helden, C. Lemer, J. Rischelles, S. J. Wodak, J. Garcia-Martinez, J. E. Perez-Ortin, H. Michael, A. Kaps, E. Talla, B. Dujon, B. Andre, J. L. Souciet, J. De Montigny, E. Bon, C. Gaillardin, H. W. Mewes, 2005, "CYGD: The Comprehensive Yeast Genome Database", *Nucleic Acids Res*, Vol. 33, pp. D364–D368.

Hahn, J. S., Z. Hu, D. J. Thiele, V. R. Iyer, 2004, "Genome-wide analysis of the biology of stress responses through heat shock transcription factor", *Mol. Cell Biol.*, Vol. 24(12), pp. 5249-5256

Hardie, D. G., C. M. Sutherland, S. A. Hawley, R. R. McCartney, A. Leech, M. J. Stark and M. C. Schmidt, 2003, "Elm1p is One of Three Upstream Kinases For the Saccharomyces Cerevisiae SNF1 Complex", *Current Biology*, Vol. 13, pp. 1299-1305.

Harvey, S. L., A. Charlet , W. Haas , S. Gygi and D. Kellogg, 2005, "Cdk1-dependent regulation of the mitotic inhibitor Wee1", *Cell*, Vol. 122(3), pp. 407-420

Haugen, A. C., R. Kelley, J. B. Collins, C. J. Tucker, C. Deng, C. A. Afshari, J. M. Brown, T. Ideker, B. Van Houten, 2004, "Integrating Phenotypic And Expression Profiles To Map Arsenic-Response Networks", *Genome Biol.*, Vol. 5:R95.

Hohmann, S., 2006, AMPKIN: *Systems biology of the AMP-activated protein kinase*, http://www.gmm.gu.se/AMPKIN

Hong, S. P., M. Carlson, 2007, "Regulation of Snf1 Protein Kinase in Response to Environmental Stress", *J. Biol. Chem.*, Vol. 282(23), pp. 16838-16845

Ideker, T., O. Ozier, B. Schiwikowski and A. F. Siegel, 2002, "Discovering Regulatory and Signalling Circuits in Molecular Interaction Networks", *Bioinformatics*, Vol.18, pp. S233-S240.

Ideker, T., T. Galitski, L. Hood, 2001, "A New Approach to Decoding Life: Systems Biology" *Annu Rev Genomics Hum Genet*, Vol.2, pp. 343–372.

Jiang, R., and M. Carlson, 1996, "Glucose Regulates Protein Interactions within the Yeast SNF1 Protein Kinase Complex", *Genes Dev.*, Vol. 10, pp. 3105–3115.

Johnston, M., 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.

Kaffman, A., N. M. Rank, E. M. O'Neill, L. S. Huang, E. K. O'Shea, 1998, "The receptor Msn5 exports the phosphorylated transcription factor Pho4 out of the nucleus", *Nature*, Vol. 396(6710), pp. 482-486

Kaida., D., H. Yashidora, A. Toh-e, Y. Kikuchi, 2002, "Yeast Whi2 and Psr1-phosphatase form a complex and regulate STRE-mediated gene expression", *Genes Cells*, Vol. 7(6), pp. 543-552

Kim, J-H., V. Brachet, H. Moriya and M. Johnston, 2006, "Integration of Transcriptional and Posttranslational Regulation in a Glucose Signal Transduction Pathway in Saccharomyces cerevisiae", *Eukaryotic Cell*, Vol. 5, No. 1, pp. 167-173.

Klein C. and K. Struhl, 1994, "Protein kinase A mediates growth-regulated expression of yeast ribosomal protein genes by modulating RAP1 transcriptional activity", *Mol Cell Biol.*, Vol. 14(3), pp. 1920-1928

Lieb J. D. et al, 2001, "Promoter-specific binding of Rap1 revealed by genome-wide maps of protein-DNA association", *Nat. Genet.*, Vol. 28(4), pp. 327-334

Liu, H., J. H. Toyn, Y. Chiang, M. P. Draper, L. H. Johnston and C. L. Denis, 1997, "DBF2, a cell cycle-regulated protein kinase, is physically and functionally associated with the CCR4 transcriptional regulatory complex", *EMBO J.*, Vol.16(17), pp. 5289–5298

Liu, Y., H. Zhao, 2004, "A Computational Approach for Ordering Signal Transduction Pathway Components from Genomics and Proteomics Data", *BMC Bioinformatics*, Vol. 5, pp. 158–163.

Lundin, M., J. O. Nehlin and H. Ronne, 1994, "Importance of a Flanking AT-Rich Region in Target Site Recognition by the GC Box-Binding Zinc Finger Protein Mig1", *Molecular and Cellular Biology*, Vol. 14, No. 3, pp. 1979-1985.

Lutfiyya, L. L., V. R. Iyer, J. DeRisi, M. J. DeVit, P. O. Brown and M. Johnston, 1998, "Characterization of Three Related Glucose Repressors and Genes They Regulate in Saccharomyces cerevisiae", *Genetics*, Vol. 150, pp. 1377–1391.

Mitchelhill, K. I., D. Stapleton, G. Gao, C. House, B. Michell, F. Katsis, L. A. Witters, B. E. Kemp, 1994, "Mammalian Amp-Activated Protein Kinase Shares Structural and Functional Homology with the Catalytic Domain of Yeast Snf1 Protein Kinase", *J. Biol. Chem*. Vol. 269, pp. 2361–2364.

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.

Mukherjee, S., M. F. Berger, G. Jona, X. S. Wang, D. Muzzey, M. Snyder, R. A. Young and M. L. Bulyk, 2004, "Rapid Analysis of the DNA-binding Specificities of Transcription Factors with DNA microarray", *Nat. Genet.*, Vol. 36, pp. 1331–1339.

Nath, N., R. R. Mccartney, and M. C. Schmidt, 2003, 'Yeast Pak1 Kinase Associates With and Activates Snf1', *Molecular and Cellular Biology*, Vol. 23, pp. 3909-3917.

Nehlin, J. O. and H. Ronne, 1990, "Yeast MIG1 Repressor is Related to The Mammalian Early Growth Response and Wilms' Tumour Finger Proteins", *The EMBO Journal*, Vol. 9, pp. 2891-2898.

Neigeborn, L., M. Carlson, 1984, "Genes Affecting the Regulation of SUC2 Gene Expression by Glucose Repression in Saccharomyces cerevisiae", *Genetics*, Vol. 108, pp. 845–858.

Oliveira, A. P., K. R. Patil and J. Nielsen, 2008, "Architecture of Transcriptional Regulatory Circuits is Knitted Over the Topology of Bio-Molecular Interaction Networks", *BMC Systems Biology*, Vol. 2:17.

Orlova, M., E. Kanter, D. Krakovich and S. Kuchin, 2006, "Nitrogen Availability and TOR Regulates the Snf1 Protein Kinase in Saccharomyces Cerevisia"', *Eukaryotic Cell*, Vol. 5, pp. 1831-1837.

Papin, J. A., T. Hunter, B. O. Palsson, S. Subramaniam, 2005, "Reconstruction of Cellular Signalling Networks and Analysis of Their Properties", *Nature Reviews Molecular Cell Biology*, Vol. 6, pp. 99-111.

Patil, K. R., J. Nielsen, 2005, "Uncovering Transcriptional Regulation of Metabolism by Using Metabolic Network Topology ", *PNAS*, Vol. 108, pp. 2685-2689.

Raghevendran, V., J. Nielsen and L. Olsson, 2005, "Teaching Microbial Physiology Using Glucose Repression in Baker's Yeast as an Example", *Biochemistry and Molecular Biology Education*, Vol. 33, No. 6, pp. 404-410.

Reisdorf, P., E. Boy-Marcotte, M. Bolotin-Fukuhara,1997, "The MBR1 gene from Saccharomyces cerevisiae is activated by and required for growth under sub-optimal conditions", *Mol. Gen. Genet.*, Vol. 255(4), pp. 400-409

Ross, K. E., P. Kaldis, M. J. Solomon, 2000, "Activating phosphorylation of the Saccharomyces cerevisiae cyclin-dependent kinase, cdc28p, precedes cyclin binding", *Mol. Biol. Cell*, Vol. 11(5), pp. 1597-1609

Santangelo, G. M., 2006, "Glucose Signaling in Saccharomyces Cerevisiae", *Microbial and Molecular Biology Reviews*, Vol. 70, No. 1, pp. 253-282.

Schmitt A. P. and K. McEntee, 1996, "Msn2p, a zinc finger DNA-binding protein, is the transcriptional activator of the multistress response in Saccharomyces cerevisiae", *Proc. Natl. Acad. Sci. USA*, Vol. 93(12), pp. 5777-5782.

Schuller, H. J., 2003, "Transcriptional Control of Nonfermentative Metabolism in the Yeast Saccharomyces Cerevisiae", *Current Genetics*, Vol. 43, No. 3, pp. 139-160.

Steffen, M., A. Petti, J. Aach, P. D'haeseleer and G. Church, 2002, "Automated Modeling of Signal Transduction Networks", *BMC Bioinformatics*, Vol. 3**:**34.

Sutherland, C. M., S. A. Hawley, R. R. McCartney, A. Leech, M. J.R. Stark, M. C. , D. G. van Hemert, M. J., et al, 2001, "Yeast 14-3-3 proteins", *Yeast*, Vol. 18(10), pp. 889-895

Verdone, L., J. Wu, K. van Riper, N. Kacherovsky, M. Vogelauer, E. T. Young, M. Grunstein, E. Di Mauro and M. Caserta, 2002, "Hyperacetylation of chromatin at the ADH2 promoter allows Adr1 to bind in repressed conditions", *EMBO J.*, Vol. 21(5), pp. 1101-1111

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

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.

Woods, A., M. R. Munday, J. Scott, X. Yang, M. Carlson, D. Carling, 1994, "Yeast Snf1 is Functionally Related to Mammalian Amp-Activated Protein Kinase and Regulates Acetyl-CoA Carboxylase in vivo", *J. Biol. Chem.*, Vol. 269, pp. 19509–19515.

Yang, X., J. A. Hubbard, M. Carlson, 1992, "A Protein Kinase Substrate Identified by the Two-Hybrid System", *Science*, Vol. 257, pp. 680–682.

Yang, X., R. Jiang, M. Carlson, 1994, "A Family of Proteins Containing a Conserved Domain That Mediates Interaction with the Yeast SNF1 Protein Kinase Complex", *EMBO J.*, Vol. 13, pp. 5878–5886.

Young E. T. et al., 2002, “Snf1 protein kinase regulates Adr1 binding to chromatin but not transcription activation”, *J. Biol. Chem.*, Vol. 277(41), pp. 38095-38103

Zimmermann, F. K., I. Kaufmann, H. Rasenberger, P. Haussmann, 1977, “Genetics of Carbon Catabolite Repression in Saccharomyces cerevisiae: Genes Involved in The Derepression Process”, *Mol. Gen. Genet.*, Vol. 151, pp. 95–103.