Systematic Identification of TOR Downstream Effectors Using Random-Walks on the Yeast Interactome

Shahin Mohammadi[1]*, Shankar Subramaniam[2,3], and Ananth Grama[1]

[1]Department of Computer Sciences, Purdue University, West Lafayette, IN 47906.

[2]Department of Chemistry and Biochemistry, University of California at San Diego, La Jolla, CA 92093.

[3]Department of Bioengineering, University of California at San Diego, La Jolla, CA 92093.

E-mail: mohammadi@purdue.edu, shankar@sdsc.edu, ayg@cs.purdue.edu

*Corresponding author

Abstract

Motivation: Calorie restriction (CR), without malnutrition, is one of the most conserved non-genetic interventions that extends both the mean, and the maximum life-span in evolutionarily distant species, ranging from yeast to mammals. The target of rapamycin (TOR) has been shown to play a key role in mediating life-span extension in response to the CR, by modulating cellular response to nutrient-availability and orchestrating various components of cellular machinery, including cell growth, translation initiation, ribosome biogenesis, and autophagy. Furthermore, both genetic and pharmacological interventions inhibiting the TOR pathway exhibit a similar phenotype, which can not be further extended by CR. These observations have motivated experimental investigations of downstream effectors of TOR, which are responsible for mediating life-span extension.

Results: In this paper, we derive the first comprehensive computational map of downstream effectors of TOR. We adopt a systematic approach, based on the known random-walk method for tracing information flow in the yeast interactome. Using a rigorous statistical framework, we identify targets carrying significant amounts of TOR signaling. Our approach, unlike experimental methods, is not limited to specific aspects of cellular response. Rather, it is shown to predict transcriptional changes, as well as post-translational modifications in response to TOR signaling. GO enrichment analysis of the identified effectors not only sheds light on the functional mechanisms downstream of TOR, but also provides mechanistic understanding of the crosstalk among them. In addition to identifying several known effectors, our method also identifies a number of other targets, whose roles have been hypothesized in literature but not confirmed; as well as potential new targets.

Availability: All datasets and supplementary materials are available for download from http://compbio.soihub.org/projects/torc1.

1 Introduction

Cellular aging is a multi-factorial complex phenotype, characterized by the accumulation of damaged cellular components over the organism's life-span (Fontana et al., 2010). The progression of aging depends on both the increasing rate of damage; to DNA, RNA, proteins, and cellular organelles; as well as the gradual decline of the cellular defense mechanisms against stress, which can ultimately lead to a dysfunctional cell with a higher risk factor for disease.

Limiting caloric intake without causing malnutrition, also known as *calorie restriction (CR)*, is one of the most conserved non-genetic interventions, which extends life-span in evolutionarily distant species, ranging from yeast to mammals(Bishop and Guarente, 2007; Fontana et al., 2010; Kaeberlein, 2010). Inhibition of the nutrient-sensing pathways using either genetic or pharmacological intervention also results in a similar phenotype(Bishop and Guarente, 2007; Fontana et al., 2010). More importantly, CR not only extends life-span, but also delays both the progression and the increasing risk-factor of a wide range of age-related diseases, including cancers(Colman et al., 2009), cardiovascular disease(Mattson and Wan, 2005; Cruzen and Colman, 2009), and multiple neurodegenerative disorders(Maswood et al., 2004; Wu et al., 2008). Motivated by these observations, considerable effort has been invested in understanding the downstream effectors of the nutrient-sensing pathways that orchestrate DR-mediated life-span extension.

Ethical concerns, combined with a longer typical life-span make it difficult to study aging in humans. Budding yeast, *Saccharomyces cerevisiae*, has been used extensively as a model organism in aging research, due to its rapid growth and ease of manipulation(Kaeberlein et al., 2007; Kaeberlein, 2010). Having two different aging paradigms – replicative life-span (RLS), defined as "the number of buds a mother cell can produce before senescence occurs," and chronological life-span (CLS), defined as "duration of viability after entering the stationary-phase," yeast provides a unique opportunity for modeling

both proliferating and post-mitotic cells. These can ultimately be used to shed light on the progression of cancers and neurodegenerative diseases.

Yeast cells are typically cultured in growth media containing 2% glucose. Reducing glucose concentration to 0.5% or less has been shown to increase both CLS and RLS(Lin et al., 2000; Kaeberlein et al., 2004; Smith et al., 2007). Both Sir2, the leading member of the sirtuins family, and the target of rapamycin (TOR) play key roles in CR-mediated life-span extension in yeast. Regulation of Sir2 in response to CR has been attributed to the metabolic shift towards respiration. This elevates the cellular NAD^+/NADH ratio and triggers the Sir2 protein, which is a NAD^+-dependent histone deacetylase(Lin et al., 2002). The most commonly investigated molecular mechanism by which Sir2 mediates life-span is suppressing homologous recombination around rDNA loci. This prevents the accumulation of extrachromosomal rDNA circles (ERC). The precise role of Sir2 in CR has been a matter of debate in recent years. Kaeberlein et al. (2004) posit that Sir2 acts in parallel with CR to promote longevity. In a subsequent study(Kaeberlein et al., 2005), they use a pool of 564 single-gene mutants and identify TOR pathway as a key element in life-span regulation by CR. TOR acts as a hub, integrating various nutrient and stress-related signals, and regulates both spatial and temporal aspects of cell growth(De Virgilio and Loewith, 2006b; Zaman et al., 2008; Loewith and Hall, 2011; Wei and Zheng, 2011). Inhibiting TOR pathway using rapamycin helps in identifying downstream effectors of TOR. However, these targets may be regulated in different levels, including, but not limited to transcription regulation and post-translational modifications. Capturing various changes that happen during rapamycin treatment to create a comprehensive systems view of the cell is a complex and onerous task.

In this paper, we propose a computational approach to the identification of downstream effectors of TOR. Our method is based on the known random walk technique for identifying key players in mediation of lifespan. We develop a rigorous statistical methodology for reverse engineering downstream effectors

from their random walk distances. Using this framework, we build a comprehensive interaction map of the TOR targets. We validate this map using enrichment analysis, as well as through experimental data, and identify a number of new targets for future investigations.

2 Target of rapamycin (TOR): A central arbitrator of nutrient and stress related signals

The target of rapamycin (TOR) is a serine/threonine protein kinase, which belongs to the conserved family of PI3K-related kinases (PIKKs). It was first identified using genetic studies in yeast while searching for mutants that confer rapamycin-resistance(Heitman et al., 1991). Rapamycin, a lipophilic macrolide originally purified as an antifungal agent and then re-discovered as an immunosuppressive drug, forms a toxic complex with its intracellular receptor FKBP12, encoded by Fpr1 gene in yeast, and directly binds to TOR in order to perform its inhibitory action and mediate cell growth(Loewith and Hall, 2011). It was later discovered that TOR protein kinases, encoded by TOR1 and TOR2 genes in yeast, form two structurally and functionally distinct multiprotein complexes(Loewith et al., 2002; Wedaman et al., 2003; Reinke et al., 2004; Wullschleger et al., 2006). TOR Complex 1 (TORC1) is rapamycin-sensitive and consists of either TOR1 or TOR2, together with KOG1, LST8, and TCO89. On the other hand, TOR Complex 2 (TORC2) does not contain TOR1, is not inhibited by rapamycin, and contains AVO1, AVO2, AVO3, LST8, BIT2, and BIT61. These two complexes correspond to two separate branches of the TOR signaling network, controlling the temporal and the spatial aspects of cell growth(Loewith and Hall, 2011), which are conserved from yeast to humans(Wullschleger et al., 2006). TORC1 has a critical role in aging and age-related pathologies(Kapahi et al., 2010; McCormick et al., 2011). Additionally, many of the known oncoproteins act as upstream activators of TORC1, while several tumor suppressor

proteins inhibit its activity(De Virgilio and Loewith, 2006a). These observations, together with more recent studies, have motivated many researchers to identify both the upstream regulators and downstream effectors of TORC1.

As a central hub that integrates various nutrient and stress related signals, and regulates different aspects of cell growth in response, TORC1 is not only regulated by the quality and the quantity of both carbon and nitrogen sources(Zaman et al., 2008; Binda et al., 2009; Neklesa and Davis, 2009; Smets et al., 2010), but also in response to noxious stressors, such as caffeine(Kuranda et al., 2006; Wanke et al., 2008). In response, TORC1 coordinately orchestrates various aspects of cellular machinery to mediate cell growth, including protein synthesis (by regulating ribosome biogenesis(Jorgensen et al., 2004), translation initiation(Barbet et al., 1996), and nutrient uptake(Schmidt et al., 1998; Beck and Hall, 1999a)), autophagy(Chang et al., 2009), and stress response(Gasch and Werner-Washburne, 2002; Wanke et al., 2008) (please see Smets et al. (2010) and Loewith and Hall (2011) for comprehensive reviews).

3 Methods

3.1 Tracing information flow in the interactome

Living cells are complex systems comprising of numerous pathways, that orchestrate various aspects of the cellular machinery. Understanding their structure, interactions, and crosstalk is essential for uncovering the cellular organization. We propose a computational approach, using a discrete-time random walk process, to trace information flow in the interactome. Similar formulations have been previously used to prioritize candidate disease genes(Köhler et al., 2008; Navlakha and Kingsford, 2010), discover network bio-markers for cancer(Nibbe et al., 2010), and identify protein complexes (Macropol et al., 2009;

Maruyama and Chihara, 2011). Additionally, there is a well-known correspondence between random-walk methods and formulations based on circuit network models(Doyle and Snell, 1984). To the best of our knowledge, our study represents the first computational effort aimed at comprehensively identifying downstream effectors of TOR using random walks. Our methods are supported by rigorous statistical models for ranking various effectors, resulting in a complete interaction map of TOR effectors.

Let $G = (V, E)$ be a (directed) graph with n vertices and m edges, where $(u, v) \in E$ iff vertex u is connected to vertex v in graph G. A *random walk* on G, initiated from vertex v, is defined as a sequence of transitions among vertices, starting from v. At each step, the random walker randomly chooses the next vertex from among the neighbors of the current node. The sequence of visited vertices generated by this random process is a Markov chain, since the choice of next vertex depends only on the current node. We can represent the transition matrix of this Markov process as a column-stochastic matrix, P, where $p_{ij} = Prob(S_{n+1} = v_i | S_n = v_j)$ and random variable S_n to represent the state of the random walk at the time step n.

Random walk with restart (RWR) is a modified Markov chain in which, at each step, a random walker has the choice of either continuing along its path, with probability α, or jump (teleport) back to the initial vertex, with probability $1 - \alpha$. Given the transition matrix of the original random walk process, P, the transition matrix of the modified chain, M, can be computed as $M = \alpha P + (1 - \alpha)\boldsymbol{e}_v\mathbf{1}$, where $\boldsymbol{e}_v$ is a stochastic vector of size n having zeros everywhere, except at index v. The stationary distribution of the modified chain, $\boldsymbol{\pi}_v(\alpha)$, defines the portion of time spent on each node in an infinite random walk with restart initiated at node v, with parameter α. This stationary distribution can be computed as follows:

$$\begin{aligned} \boldsymbol{\pi}_v(\alpha) &= M\boldsymbol{\pi}_v(\alpha) \\ &= (\alpha P + (1 - \alpha)\boldsymbol{e}_v)\boldsymbol{\pi}_v(\alpha) \end{aligned} \tag{1}$$

Enforcing a unit norm on the dominant eigenvector to ensure its stochastic property, $\| \boldsymbol{\pi}_v(\alpha) \|_1 = 1$, we have the following iterative form:

$$\boldsymbol{\pi}_v(\alpha) = \alpha P \boldsymbol{\pi}_v(\alpha) + (1-\alpha)\boldsymbol{e}_v, \tag{2}$$

which is a special case of the personalized Pagerank(Brin and Page, 1998; Page et al., 1999; Jeh and Widom, 2002; Haveliwala, 2002), with preference vector $\boldsymbol{e}_v$. Alternatively, we can compute $\boldsymbol{\pi}_v$ directly by solving the following linear system:

$$\boldsymbol{\pi}_v(\alpha) = \underbrace{(1-\alpha)(I-\alpha P)^{-1}}_{Q}\boldsymbol{e}_v, \tag{3}$$

where the right-multiplication with $\boldsymbol{e}_v$ simply selects column v of the matrix Q. A major benefit of computing Q directly is that it can be reused to compute the stationary distribution of random walks starting from different vertices, or to compute random samples of initial vertices with known distribution. For example, if we have a set of k vertices, S, that have equal chance of selection as the initial vertex (uniform distribution over set S), then we can linearly integrate the corresponding columns of Q, each with weight $\frac{1}{k}$, which is the average of the corresponding columns. To interpret $\boldsymbol{\pi}_v(\alpha)$, we expand the right-hand side of Equation 3 using the Neumann series:

$$\boldsymbol{\pi}_v(\alpha) = (1-\alpha)\sum_{i=0}^{\infty}(\alpha P)^i \boldsymbol{e}_v \tag{4}$$

We construct a model in which node v propagates a unit signal, and at every step, each node equally distributes its accumulated received signal among its out-going neighbors. Using this model, $P^i\boldsymbol{e}_v$ measures the strength of the received signal by each node after i steps of transmissions. This assumes that

transmissions are perfect, i.e., the signal does not attenuate on transmission. However, if we assume that only $\alpha\%$ of the signal is propagated during each transmission, then we can compute the accumulated received signal by each node after exactly i transmissions as $(\alpha P)^i \boldsymbol{e}_v$. We can view $\boldsymbol{\pi}_v(\alpha)$ as the strength of the signal received by each node, initially originated from v, after zero or more transmission steps. Using this method, we note that the closer a node is to the source node (v), the stronger the signal it receives from node v. Additionally, we can observe that since we integrate signals over all possible paths, unlike shortest-path based methods that only take a single path into account, we prioritize pairs of nodes that have multiple parallel pathways between them.

As noted, the stationary distribution of the random walk, $\boldsymbol{\pi}_v$, depends on the parameter α. This parameter can be viewed as the *decay factor* of the signal; the higher the parameter α, the further the signal can propagate. Let us denote by X the number of hops taken by random walker before it jumps back to source node v. Then, X follows a geometric distribution with probability of success $(1-\alpha)$ and the expected (average) length of paths taken by random walker $E(X) = \frac{\alpha}{1-\alpha}$. In other words, if we let $\alpha = \frac{l}{1+l}$, for a given parameter l, we expect the average length of paths taken by such random walk to be equivalent to l, thus we call l the depth of the random walk.

We conclude this discussion by describing the process of computation of transition matrix P for a given directed graph G with adjacency matrix A, where $A_{ij} = 1$, if node i has a directed edge to node j, and is 0 otherwise. We first create a column sub-stochastic matrix, $\tilde{P}$, in which the sum of each column is either zero, if the corresponding node has no out-going edges, or one, otherwise. Formally:

$$\tilde{P}_{ji} = \begin{cases} 0 & \text{if } A_{ik} = 0; \forall 1 \leq k \leq n, \\ \frac{A_{ij}}{\sum_{k=1}^{n} A_{ik}} & \text{O.W.} \end{cases} \tag{5}$$

We then construct the column stochastic matrix P, from $\tilde{P}$, by adding self-loops to the *dangling*

nodes, whose out-degree is zero. Mathematically, this can be formulated as $P = \tilde{P} + diag(\mathbf{1}^T - \mathbf{1}^T\tilde{P})$, where $diag(.)$ operator takes a vector and constructs a diagonal matrix from it.

A similar approach, based on the random-walk method, is the emitting model of ITMProbe (Stojmirović and Yu, 2007, 2012). The main theoretical difference is that in the emitting diffusion model, unlike our formulation, information in the form of a random walker cannot return to the respective sources. However, since the set of sources is relatively small in our study (5 nodes), the computed results using these two seemingly different methods are correlated (please see Section 3.1 for details).

3.2 Validation and statistical analysis of ranked-list of genes

Given a ranked-list of genes (in our case, genes ranked in terms of their random-walk proximity to TOR), we are interested in assessing the enrichment of different functional annotations among top-ranked genes. The classical approach to this problem is to select a cutoff, l, which separates the top-ranked genes (target set) from the rest (background set), and to compute the enrichment p-value using a hypergeometric distribution. Let us denote the number of genes by N and the total number of annotations (positives) by A. Using notation similar to Eden et al. (2007), we encode functional annotations for genes in the ranked-list using a binary vector, $\lambda = \lambda_1, \lambda_2, ...\lambda_N \in \{0,1\}^N$, having A ones and $N - A$ zeros. Let the random variable T denote the number of positive genes in the target set, if we distribute genes randomly. In this formulation, the hypergeometric p-value is:

$$\begin{aligned} P - value(T = b_l(\lambda)) &= Prob(b_l(\lambda) \leq T) = HGT(b_l(\lambda)|N, A, l) \\ &= \sum_{t=b_l(\lambda)}^{min(A,l)} \frac{C(A,t)C(N-A, l-t)}{C(N,l)}, \end{aligned} \tag{6}$$

where HGT is the tail of hypergeometric distribution, and $b_l(\lambda) = \sum_{i=1}^{l} \lambda_i$, is the number of ob-

served positives in the target set. The drawback of this approach is that we need a predefined cutoff value, l. To remedy this, Eden et al. (2007) propose a two-step method for computing the exact enrichment p-value, called *mHG p-value*, without a predefined l. In the first step of this process, we identify an optimal cut, over all possible cuts, which minimizes the hypergeometric p-value. The value computed in this manner is called the *minimum hypergeometric (mHG) score*, and is defined as:

$$mHG(\lambda) \quad = \quad min_{1 \leq l \leq N} HGT(b_l(\lambda)|N, A, l) \tag{7}$$

Next, a dynamic programming method is employed to compute the exact p-value of the observed mHG score, in the state space of all possible λ vectors (please refer to Eden et al. (2007) for algorithmic details, and Eden (2007) for an efficient implementation). We use this strategy to compute the mHG p-value of different GO terms, as well as for cross-validating our results with experimental datasets on rapamycin treatment. For the latter case, we also report $-log_{10}(HGT(b_l(\lambda)|N, A, l))$ as a function of top-ranked score percentage, which provides us additional insights into the distribution of positives among the ranked list of genes. It is worth noting that the only theoretical difference in GO enrichment analysis versus cross-validating with experimental rapamycin treatment observations is in the construction of λ vector.

3.3 Dissecting transcriptional regulatory network (TRN) using random-walk scores

As described in Section 3.2, we compute the mHG p-value for the rapamycin treatment dataset, which partitions the random-walk scores into l top-ranked genes and $N - l$ bottom-ranked genes. In this section, we establish the necessary background for analyzing the regulatory elements responsible for the

transcriptional changes observed during rapamycin treatment. Given a transcription factor (TF), with k targets, we denote the number of its positive and top-ranked targets by k_P and k_T, respectively. The first question that we are interested in answering is: how well do computationally predicted positive targets predict experimentally observed targets of a given TF? In other words, given that a target gene is ranked high in terms of its random-walk score, how confident are we that it is a true positive target? To address this question, we define two random variables. Let X be the number of top-ranked targets, if we were uniformly distributing k targets of the given TF among all genes. Similarly, let Y be the number of positive targets of the TF. Then, we can compute the following p-values:

$$\begin{aligned} p-value(X=k_T) &= Prob(k_T \leq X) \\ &= HGT(k_T|N,l,k) \\ &= \sum_{x=k_T}^{min(l,k)} \frac{C(l,x)C(N-l,k-x)}{C(N,k)} \end{aligned} \tag{8}$$

$$\begin{aligned} p-value(Y=k_P) &= Prob(k_P \leq Y) \\ &= HGT(k_P|N,A,k) \\ &= \sum_{y=k_P}^{min(A,k)} \frac{C(A,y)C(N-A,k-y)}{C(N,k)} \end{aligned} \tag{9}$$

We define *sensitivity* and *specificity* for the entire set of transcription factors as:

$$\begin{aligned} sensitivity &= Prob(pval(X=k_T) : significan|pval(Y=k_P) : significant) \\ specificity &= Prob(pval(X=k_T) :!significant|pval(Y=k_P) :!significant) \end{aligned} \tag{10}$$

Next, we are interested in identifying the relevant subset of transcription factors that are responsible

for significant transcriptional changes among top-ranked genes. Let us denote the number of top-ranked positive targets of a given TF by k_{TP}. If we compute the probability of observing k_{TP} or more positive targets among top-ranked genes, purely by chance, we can identify the associated subset of transcription factors. Let the random variable Z denote the number of top-ranked positive targets, if we were randomly distributing all targets for a given TF. We can compute the p-value of Z by conditioning it on the number of top-ranked targets as follows:

$$\begin{aligned} p-value(Z=k_{TP}) &= Prob(k_{TP} \leq Z) \\ &= \sum_{x=k_{TP}}^{min(l,k)} Prob(k_{TP} \leq Z|X=x)Prob(X=x) \\ &= \sum_{x=k_{TP}}^{min(l,k)} \sum_{z=k_{TP}}^{min(x,b_l(\lambda))} Prob(Z=z|X=x)Prob(X=x) \\ &= \sum_{x=k_{TP}}^{min(l,k)} HG(x|N,l,k) \sum_{z=k_{TP}}^{min(x,b_l(\lambda))} HT(z|l,b_l(\lambda),x) \end{aligned} \tag{11}$$

We define $-log_{10}(p-value(Z=k_{TP}))$ as the *confidence score* of a TF being involved in TORC1-related gene regulation. This allows us to identify the subset of TF(s) responsible for the transcriptional changes during rapamycin treatment.

4 Results

4.1 Constructing yeast interactome

We obtained the yeast protein-protein interactions (PPI) from the Biogrid(Stark et al., 2006) database, update 2011(Stark et al., 2011), version 3.1.83, by extracting all physical interactions, except for protein-RNA interactions, and excluding interspecies and self interactions. This dataset consists of 103,619

(63,395 non-redundant) physical interactions among 5,691 proteins. We then identified the subset of post-translational modifications (PTM), associated with "biochemical activity" evidence code in Biogrid, and used them to assign direction to the feasible subset of physical interactions, resulting in 5,791 (5,443 non-redundant) biochemical activities among proteins in yeast. After unifying different modifications among similar pairs of proteins, we obtained 5,421 directional edges among 2,002 proteins in the yeast proteome. The bulk of these interactions (over 4,000) are the phosphorylation events identified by (Ptacek et al., 2005) using proteome chip technology. Finally, we constructed the integrated network of yeast, the yeast *interactome*, by superimposing protein-protein interactions and post-translational modifications over the set of known yeast genes, and removing isolated vertices. The final constructed network, available in the Supplementary Network 1, is shown in Figure 1. Nodes in the network are color-coded using the family of the corresponding protein. The list of transcription factors was retrieved from YEASTRACT(Abdulrehman et al., 2011), while the list of all kinases and phosphatases was downloaded from YeastKinome(Breitkreutz et al., 2010). Figure 2 illustrates an example of the integration process around *Sch9* protein, which is a well-documented substrate of TORC1. The set of interactions around *Sch9* has been extracted from the protein-protein interaction and the post-translational modification networks, respectively.

4.2 Information flow analysis from TORC1 in the yeast interactome

Using the yeast interactome, represented as its adjacency matrix A, we compute information flow scores from TORC1 using the approach described in Section 3.1. In order to give all nodes in the interactome a chance of being visited, we choose parameter α, controlling the depth of the random walk process, to be equal to $\frac{d}{1+d}$, where d is the diameter of the interactome. For the yeast interactome, we determined the diameter to be equal to 6 and we set the $\alpha = \frac{6}{7} \sim 0.85$, correspondingly. We used TORC1

members, namely *Tor1, Tor2, Kog1, Lst8,* and *Tco89*, as sources of information flow, and averaged the corresponding columns of Q by letting $\boldsymbol{e}_v(i) = \frac{1}{5}$, if $v_i \in$ TORC1, and zero otherwise.

Figure 3 illustrates the distribution of computed random-walk distances, starting from TORC1, as a function of node distance from TORC1. As mentioned in Section 3.1, random-walk scores are functions of both distance from source nodes, as well as multiplicity of paths between source and sink nodes. This can be verified from the figure by noting the overlapping tails of distributions for nodes at different distances, as well as varied distribution of random-walk scores among nodes with same distance from TORC1.

In addition to our random-walk method, we also computed the ITMprobe scores, using the emitting model, with the termination probability set to 0.85. The computed ITMprobe scores are highly correlated with our scores in this experiment ($\rho = 0.9993$), and thus, we only provide further analysis for our method. The final computed random-walk scores, both for our method and the ITMProbe, are available for download as Supplementary Table 1.

4.3 Information flow-based prediction of transcriptional changes in response to rapamycin treatment

In order to further analyze the random-walk neighborhood of TORC1, we ranked all nodes in the yeast interactome based on their computed information flow scores. We hypothesized that if the information flow-based method faithfully reconstructs the TORC1 signaling network, it should be able to predict transcriptional changes due to rapamycin treatment, which inhibits TORC1 *in vivo*. To validate this hypothesis, we used a recent mRNA expression profile of yeast in response to rapamycin treatment(Fournier et al., 2010). We extracted the set of genes that have significant expression change at a minimum threshold of 2-fold change, resulting in a total of 366 repressed and 291 induced genes. We constructed the

set of true positives from the rapamycin-induced genes by filtering out genes that do not have a corresponding vertex in the yeast interactome. Using this set of true-positives, we computed the mHG score to assess the frequency of rapamycin-induced genes among top-ranked genes identified by our information flow method. Figure 4 shows the tail of the hypergeometric distribution as a function of cut-size from the top of random-walk scores. The peak of the plot, corresponding to the minimum hypergeometric (mHG) score, occurs at the index $l = 906$ from the top of random walk scores, which covers approximately the top 15% of scores. There are 181 positive genes in this partition, from a total of 579, yielding a mHG score of 1.1078e-22. We computed the exact p-value corresponding to this mHG score, using the dynamic-programming method of Eden et al. (2007), resulting in the significant enrichment p-value of 3.2526e-19. This in turn suggests that the random-walk neighborhood of TORC1 is highly correlated with rapamycin-induced gene expression changes and can be used as a predictor of genes that are regulated by TORC1.

4.4 Information flow scores and transcriptional regulatory networks

As shown in the previous section, top-ranked random-walk scores are highly enriched with rapamycin-impacted genes. However, TORC1 does not directly regulate transcription of these genes. This observation led us to ask: which transcription factors are responsible and which intermediary elements are involved in these regulations?

To address this problem, we constructed the yeast transcriptional regulatory network (TRN) from the documented regulations in YEASTRACT(Abdulrehman et al., 2011), consisting of 48082 interactions between 183 transcription factors (TF) and 6403 target genes (TG). Among these 183 TFs, 179 of them have a corresponding node in the yeast interactome. For each of these transcription factors, we computed all the statistics described in Section 3.3. These are summarized in Supplementary Table 2. We used the

threshold of 0.01, after correcting for multiple hypothesis testing using Bonferroni method, in order to find significant p-values. At this threshold level, sensitivity and specificity are 0.2619 and 0.9854, respectively. The observed specificity value is high, meaning that if a TF is not experimentally validated (the number of positive targets are not significant), with high probability our computational model also reports it as negative (it will not have significant number of top-ranked targets). On the other hand, the relatively low sensitivity value means that even if a TF has many positive targets, our computational method may miss it. In conclusion, transcription factors that have significant numbers of top-ranked targets are high-confidence, but not comprehensive, candidate(s) for downstream effectors of TORC1.

In order to identify most relevant subset of transcription factors in a systematic manner, we introduce the *confidence score* for each TF, defined as $-log_{10}(p-value(Z = k_{TP}))$, which assesses both positivity and high rank of its targets (please see Section 3.3 for details). Using this approach, we identified 16 TFs with very high confidence scores, which are postulated to be responsible for transcriptional changes in a TORC1-dependent manner. From a total of 16 high-confidence TFs, five are also highly ranked themselves (Table 1).

Among these top-ranked, high confidence, transcription factors, Sfp1, Gln3, and Gcn4 are well-documented downstream effectors of TORC1(Lempiäinen et al., 2009; Beck and Hall, 1999b; Bertram et al., 2000; Steffen et al., 2008)(please see Zaman et al. (2008), Smets et al. (2010), and Loewith and Hall (2011) for a more comprehensive review). *Sfp1* is a stress- and nutrient-sensitive regulator of cell growth, responsible for mediating the expression of genes involved in ribosome biogenesis, such as RP genes and RiBi factors(Jorgensen et al., 2002; Marion et al., 2004). TORC1 mediates Sfp1-related genes by phosphorylating Sfp1 and regulating its nuclear localization(Lempiäinen et al., 2009). *Gln3* is a GATA-family transcription factor, which positively regulates the expression of nitrogen catabolite repression (NCR)-sensitive genes(Courchesne and Magasanik, 1988; Bertram et al., 2000). TORC1-dependent

regulation of Gln3 is mediated by promoting its association with its cytoplasmic anchor protein Ure2, which is dependent on Tap42-PPase branch(Loewith and Hall, 2011), but not the Sch9 branch(Urban et al., 2007). *Gcn4* is a nutrient-responsive transcription factor, activated upon amino acid starvation, which is noted for its role in regulating amino acid biosynthesis(Hinnebusch, 2005). Additionally, Steffen et al. (2008) proposed a critical role for Gcn4 in mediating life-span in yeast. TORC1 regulates Gcn4 by mediating its translation level in a eIF2α-dependent manner(Loewith and Hall, 2011).

In contrast, Ste12 and Yap1 have not been previously positioned downstream of TORC1, to the best of our knowledge. Ste12 is best known as a downstream target of mitogen-activated protein kinase (MAPK) signaling cascade and is responsible for regulating genes involved in mating or pseudohyphal/invasive growth(Madhani and Fink, 1997). Rutherford et al. (2008) showed that over-expression of the ammonium permease Mep2 induces the transcription of known targets of Ste12. A more recent study by Santos et al. (2012) additionally positions TORC1 downstream of Mep2, which, taken together with the link between Mep2-Ste12, suggests Ste12 as a potential downstream effector of TORC1. *Yap1* is an AP-1 family transcription factor required for inducing oxidative(Stephen et al., 1995; Temple et al., 2005) and carbon(Wiatrowski and Carlson, 2003) stress responses, the latter of which is proposed to be independent of TORC1. Additionally, Yap1 expression has been shown to increase significantly during replicative aging(Yiu et al., 2008). Finally, it has been suggested recently that spermidine, a conserved longevity factor (Eisenberg et al., 2009), mediates macroautophagy in a Yap1 and Gcn4 dependent manner(Teixeira et al., 2011). Finally, there is a diverse set of age-related functions associated with Yap1, many of which are also attributed to TORC1. These observations suggest Yap1 as a potential candidate downstream effector of TORC1.

An interesting observation, from Figure 4, is that the highest-ranked genes (approximately the top 150 genes) are not enriched in terms of rapamycin-induced genes. This can be explained by noting that

the regulatory elements, including TFs, do not typically change their expression level in response to TOR signaling. Instead, they are targeted for post-translational modifications (typically, phosphorylation). We consequently hypothesize that the top genes should be enriched in terms of phosphorylation events. To verify this hypothesis, we extracted the signaling and regulatory elements responsible for *Gap1*, a general amino acid permease regulated by NCR. More specifically, Gap1 is positively regulated via Gln3 and Gat1 and is repressed by Gzf3 and Dal80(Zaman et al., 2008; Smets et al., 2010). Surprisingly, all four of these regulators are among top-ranked transcription factors, yet none of these has significant expression change in response to rapamycin treatment. Using a recent phosphoproteome of yeast in response to rapamycin treatment(Huber et al., 2009), we identified that both of the transcriptional activators of Gap1, *Gln3* and *Gal80*, are highly phosphorylated in response to rapamycin treatment. Finally, Tap42-Sit4 is the upstream regulator of Gcn4 and a direct substrate of TORC1. Figure 5 illustrates this signaling pathway with each element annotated using its random-walk rank. All signaling elements upstream Gap1 are ranked among top-ranked scores, yet none of these change their expression levels in response to rapamycin treatment. This supports the notion that the top-ranked genes in the random-walk are primarily involved in TOR signaling.

Finally, we constructed the *TORC-specific TRN*, a subgraph of the yeast TRN in which we only included top-ranked transcription factors and their top-ranked targets. This network consists of 1960 interactions, between 25 transcription factors and 778 target genes, and is available as Supplementary Network 2.

4.5 GO Enrichment analysis: constructing a functional map of the information flow scores

We used GOrilla(Eden et al., 2009) to compute the mHG p-value of each gene ontology (GO)(Ashburner et al., 2000) term, under biological process, molecular function, and cellular component branches. We identified the set of significant terms using the p-value threshold of $p < 10^{-3}$. The comprehensive list of enriched terms for each branch of the GO is available for download as Supplementary Tables 3-5.

In order to visualize the high-level organization of enriched terms, we extracted the subset of yeast GO terms marked by the Saccharomyces Genome Database (SGD)(Cherry et al., 2012) as GO slim. We then used EnrichmentMap(Merico et al., 2010), a recent Cytoscape(Smoot et al., 2011) plugin, to construct the network (map) of enriched GO slim terms, shown in Figure 6. Here, nodes represent GO terms and edges represent the extent of overlap between the geneset of terms. GO terms under BP, MF, and CC branches of GO are color-coded using red, green, and blue colors, respectively. The p-value of each term determines the opacity of both the node and its label; the bolder a term looks, the more significant its enrichment score is. Finally, the total number of enriched genes for each GO term is shown using the size the corresponding node. The final enrichment map is available for download as Supplementary Network 3.

We first note that *all* previously known targets of TORC1 are also identified using our random-walk method. For example, many of the terms in the enrichment map, such as transcription from RNA polymerase I and III promoter(s), rRNA processing, ribosomal subunit export from nucleus, and ribosomal large subunit biogenesis, describe different aspects of the ribosome biogenesis. Additionally, in order to gain a mechanistic understanding of each GO term, we project them back to the original network and construct the induced subgraph of the yeast interactome. As a case study, we extracted the set of enriched

genes represented by the *transcription initiation* GO term and constructed its corresponding induced subgraph, which is shown in Figure 7. Here, nodes are grouped and/or annotated based on their functional role in forming the transcription pre-initiation complex (PIC), as well as the RNA polymerase (RNAP). The basal level of transcription in Eukaryotic cells by RNAP needs a family of general transcription factors (GTF), prior to the formation of PIC. The TATA-binding protein (TBP), encoded by Spt15 gene in yeast, is a universal GTF that is involved in transcription by all three types of nuclear RNAP. As a component of TFIIIB complex, it forms the PIC complex and recruits RNAPIII to the transcriptional start site(TSS) of tRNAs, 5S rRNA, and most snRNAs. As a part of TFIID, it forms a complex together with TBP-associated factors (TAF), and binds to the core promoter region of the protein-coding genes, as well as some snRNAs. The correct assembly of PIC, required for directing RNAPII to the TSS, needs additional GTFs, namely TFIIA, -B, -D, -E, -F, and TFIIH, as well as the Mediator (MED) complex. These components are assembled in an orderly fashion to form the PIC and mediate the transcription initiation by RNAPII (please see Hampsey (1998) and Maston et al. (2006) for a comprehensive review). These complex interactions are shown in Figure 7, which provides a mechanistic understanding of transcription initiation in the yeast cells.

5 Conclusion

We have proposed a random-walk based method for tracing information flow in the yeast interactome and applied it to identify downstream effectors of TORC1. We showed that our information flow scores faithfully predict transcriptional changes in response to rapamycin-treatment, which supports accuracy of predicted effectors. Additionally, we proposed a rigorous statistical framework to identify the transcription factors that are responsible for the bulk of these transcriptional changes. Our framework identifies

previously known targets of the TOR signaling pathway, as well as providing new targets for future research.

6 Acknowledgement

This work supported by the Center for Science of Information (CSoI), an NSF Science and Technology Center, under grant agreement CCF-0939370, and by NSF grants DBI 0835677 and 0641037.

7 Disclosure statement

No competing financial interests exist.

References

Abdulrehman, D., Monteiro, P.T., Teixeira, M.C., et al. 2011. YEASTRACT: providing a programmatic access to curated transcriptional regulatory associations in Saccharomyces cerevisiae through a web services interface. *Nucleic acids research* 39(Database issue), D136–40.

Ashburner, M., Ball, C.A., Blake, J.A., et al. 2000. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nature genetics* 25(1), 25–9.

Barbet, N.C., Schneider, U., Helliwell, S.B., et al. 1996. TOR controls translation initiation and early G1 progression in yeast. *Molecular biology of the cell* 7(1), 25–42.

Beck, T. and Hall, M.N. 1999a. The TOR signalling pathway controls nuclear localization of nutrient-regulated transcription factors. *Nature* 402(6762), 689–92.

Beck, T. and Hall, M.N. 1999b. The TOR signalling pathway controls nuclear localization of nutrient-regulated transcription factors. *Nature* 402(6762), 689–92.

Bertram, P.G., Choi, J.H., Carvalho, J., et al. 2000. Tripartite regulation of Gln3p by TOR, Ure2p, and phosphatases. *The Journal of biological chemistry* 275(46), 35727–33.

Binda, M., Péli-Gulli, M.P., Bonfils, G., et al. 2009. The Vam6 GEF controls TORC1 by activating the EGO complex. *Molecular cell* 35(5), 563–73.

Bishop, N.A. and Guarente, L. 2007. Genetic links between diet and lifespan: shared mechanisms from yeast to humans. *Nature reviews. Genetics* 8(11), 835–44.

Breitkreutz, A., Choi, H., Sharom, J.R., et al. 2010. A global protein kinase and phosphatase interaction network in yeast. *Science (New York, N.Y.)* 328(5981), 1043–6.

Brin, S. and Page, L. 1998. The Anatomy of a Large-Scale Hypertextual Web Search Engine. In *Seventh International World-Wide Web Conference (WWW 1998).*

Chang, Y.Y., Juhász, G., Goraksha-Hicks, P., et al. 2009. Nutrient-dependent regulation of autophagy through the target of rapamycin pathway. *Biochemical Society transactions* 37(Pt 1), 232–6.

Cherry, J.M., Hong, E.L., Amundsen, C., et al. 2012. Saccharomyces Genome Database: the genomics resource of budding yeast. *Nucleic acids research* 40(Database issue), D700–5.

Colman, R.J., Anderson, R.M., Johnson, S.C., et al. 2009. Caloric restriction delays disease onset and mortality in rhesus monkeys. *Science (New York, N.Y.)* 325(5937), 201–4.

Courchesne, W.E. and Magasanik, B. 1988. Regulation of nitrogen assimilation in Saccharomyces cerevisiae: roles of the URE2 and GLN3 genes. *Journal of bacteriology* 170(2), 708–13.

Cruzen, C. and Colman, R.J. 2009. Effects of caloric restriction on cardiovascular aging in non-human primates and humans. *Clinics in geriatric medicine* 25(4), 733–43, ix–x.

De Virgilio, C. and Loewith, R. 2006a. Cell growth control: little eukaryotes make big contributions. *Oncogene* 25(48), 6392–415.

De Virgilio, C. and Loewith, R. 2006b. The TOR signalling network from yeast to man. *The international journal of biochemistry & cell biology* 38(9), 1476–81.

Doyle, P.G. and Snell, J.L. 1984. *Random walks and electric networks*. Carus mathematical monographs. Mathematical Association of America.

Eden, E. 2007. *Discovering Motifs in Ranked Lists of DNA Sequences*. Ph.D. thesis Technion - Israel Institute of Technology.

Eden, E., Lipson, D., Yogev, S., et al. 2007. Discovering motifs in ranked lists of DNA sequences. *PLoS computational biology* 3(3), e39.

Eden, E., Navon, R., Steinfeld, I., et al. 2009. GOrilla: a tool for discovery and visualization of enriched GO terms in ranked gene lists. *BMC bioinformatics* 10, 48.

Eisenberg, T., Knauer, H., Schauer, A., et al. 2009. Induction of autophagy by spermidine promotes longevity. *Nature cell biology* 11(11), 1305–14.

Fontana, L., Partridge, L., and Longo, V.D. 2010. Extending healthy life span–from yeast to humans. *Science (New York, N.Y.)* 328(5976), 321–6.

Fournier, M.L., Paulson, A., Pavelka, N., et al. 2010. Delayed correlation of mRNA and protein expression in rapamycin-treated cells and a role for Ggc1 in cellular sensitivity to rapamycin. *Molecular & cellular proteomics : MCP* 9(2), 271–84.

Gasch, A.P. and Werner-Washburne, M. 2002. The genomics of yeast responses to environmental stress and starvation. *Functional & integrative genomics* 2(4-5), 181–92.

Hampsey, M. 1998. Molecular genetics of the RNA polymerase II general transcriptional machinery. *Microbiology and molecular biology reviews : MMBR* 62(2), 465–503.

Haveliwala, T.H. 2002. Topic-Sensitive PageRank. In *To appear in WWW-2002*.

Heitman, J., Movva, N.R., and Hall, M.N. 1991. Targets for cell cycle arrest by the immunosuppressant rapamycin in yeast. *Science (New York, N.Y.)* 253(5022), 905–9.

Hinnebusch, A.G. 2005. Translational regulation of GCN4 and the general amino acid control of yeast. *Annual review of microbiology* 59, 407–50.

Huber, A., Bodenmiller, B., Uotila, A., et al. 2009. Characterization of the rapamycin-sensitive phosphoproteome reveals that Sch9 is a central coordinator of protein synthesis. *Genes & development* 23(16), 1929–43.

Jeh, G. and Widom, J. 2002. Scaling Personalized Web Search. Technical Report 2002-12 Stanford InfoLab.

Jorgensen, P., Nishikawa, J.L., Breitkreutz, B.J., et al. 2002. Systematic identification of pathways that couple cell growth and division in yeast. *Science (New York, N.Y.)* 297(5580), 395–400.

Jorgensen, P., Rupes, I., Sharom, J.R., et al. 2004. A dynamic transcriptional network communicates growth potential to ribosome synthesis and critical cell size. *Genes & development* 18(20), 2491–505.

Kaeberlein, M. 2010. Lessons on longevity from budding yeast. *Nature* 464(7288), 513–9.

Kaeberlein, M., Burtner, C.R., and Kennedy, B.K. 2007. Recent developments in yeast aging. *PLoS genetics* 3(5), e84.

Kaeberlein, M., Kirkland, K.T., Fields, S., et al. 2004. Sir2-independent life span extension by calorie restriction in yeast. *PLoS biology* 2(9), E296.

Kaeberlein, M., Powers, R.W., Steffen, K.K., et al. 2005. Regulation of yeast replicative life span by TOR and Sch9 in response to nutrients. *Science (New York, N.Y.)* 310(5751), 1193–6.

Kapahi, P., Chen, D., Rogers, A.N., et al. 2010. With TOR, less is more: a key role for the conserved nutrient-sensing TOR pathway in aging. *Cell metabolism* 11(6), 453–65.

Köhler, S., Bauer, S., Horn, D., et al. 2008. Walking the interactome for prioritization of candidate disease genes. *American journal of human genetics* 82(4), 949–58.

Kuranda, K., Leberre, V., Sokol, S., et al. 2006. Investigating the caffeine effects in the yeast Saccharomyces cerevisiae brings new insights into the connection between TOR, PKC and Ras/cAMP signalling pathways. *Molecular microbiology* 61(5), 1147–66.

Lempiäinen, H., Uotila, A., Urban, J., et al. 2009. Sfp1 interaction with TORC1 and Mrs6 reveals feedback regulation on TOR signaling. *Molecular cell* 33(6), 704–16.

Lin, S.J., Defossez, P.A., and Guarente, L. 2000. Requirement of NAD and SIR2 for life-span extension by calorie restriction in Saccharomyces cerevisiae. *Science (New York, N.Y.)* 289(5487), 2126–8.

Lin, S.J., Kaeberlein, M., Andalis, A.A., et al. 2002. Calorie restriction extends Saccharomyces cerevisiae lifespan by increasing respiration. *Nature* 418(6895), 344–8.

Loewith, R. and Hall, M.N. 2011. Target of rapamycin (TOR) in nutrient signaling and growth control. *Genetics* 189(4), 1177–201.

Loewith, R., Jacinto, E., Wullschleger, S., et al. 2002. Two TOR complexes, only one of which is rapamycin sensitive, have distinct roles in cell growth control. *Molecular cell* 10(3), 457–68.

Macropol, K., Can, T., and Singh, A.K. 2009. RRW: repeated random walks on genome-scale protein networks for local cluster discovery. *BMC bioinformatics* 10, 283.

Madhani, H.D. and Fink, G.R. 1997. Combinatorial control required for the specificity of yeast MAPK signaling. *Science (New York, N.Y.)* 275(5304), 1314–7.

Marion, R.M., Regev, A., Segal, E., et al. 2004. Sfp1 is a stress- and nutrient-sensitive regulator of ribosomal protein gene expression. *Proceedings of the National Academy of Sciences of the United States of America* 101(40), 14315–22.

Maruyama, O. and Chihara, A. 2011. NWE: Node-weighted expansion for protein complex prediction using random walk distances. *Proteome science* 9 Suppl 1, S14.

Maston, G.A., Evans, S.K., and Green, M.R. 2006. Transcriptional regulatory elements in the human genome. *Annual review of genomics and human genetics* 7, 29–59.

Maswood, N., Young, J., Tilmont, E., et al. 2004. Caloric restriction increases neurotrophic factor levels and attenuates neurochemical and behavioral deficits in a primate model of Parkinson's disease. *Proceedings of the National Academy of Sciences of the United States of America* 101(52), 18171–6.

Mattson, M.P. and Wan, R. 2005. Beneficial effects of intermittent fasting and caloric restriction on the cardiovascular and cerebrovascular systems. *The Journal of nutritional biochemistry* 16(3), 129–37.

McCormick, M.A., Tsai, S.Y., and Kennedy, B.K. 2011. TOR and ageing: a complex pathway for a complex process. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences* 366(1561), 17–27.

Merico, D., Isserlin, R., Stueker, O., et al. 2010. Enrichment map: a network-based method for gene-set enrichment visualization and interpretation. *PloS one* 5(11), e13984.

Navlakha, S. and Kingsford, C. 2010. The power of protein interaction networks for associating genes with diseases. *Bioinformatics (Oxford, England)* 26(8), 1057–63.

Neklesa, T.K. and Davis, R.W. 2009. A genome-wide screen for regulators of TORC1 in response to amino acid starvation reveals a conserved Npr2/3 complex. *PLoS genetics* 5(6), e1000515.

Nibbe, R.K., Koyutürk, M., and Chance, M.R. 2010. An integrative -omics approach to identify functional sub-networks in human colorectal cancer. *PLoS computational biology* 6(1), e1000639.

Page, L., Brin, S., Motwani, R., et al. 1999. The PageRank Citation Ranking: Bringing Order to the Web. Technical Report 1999-66 Stanford InfoLab.

Ptacek, J., Devgan, G., Michaud, G., et al. 2005. Global analysis of protein phosphorylation in yeast. *Nature* 438(7068), 679–84.

Reinke, A., Anderson, S., McCaffery, J.M., et al. 2004. TOR complex 1 includes a novel component, Tco89p (YPL180w), and cooperates with Ssd1p to maintain cellular integrity in Saccharomyces cerevisiae. *The Journal of biological chemistry* 279(15), 14752–62.

Rutherford, J.C., Chua, G., Hughes, T., et al. 2008. A Mep2-dependent transcriptional profile links permease function to gene expression during pseudohyphal growth in Saccharomyces cerevisiae. *Molecular biology of the cell* 19(7), 3028–39.

Santos, J., Sousa, M.J.a., and Leão, C. 2012. Ammonium is toxic for aging yeast cells, inducing death and shortening of the chronological lifespan. *PloS one* 7(5), e37090.

Schmidt, A., Beck, T., Koller, A., et al. 1998. The TOR nutrient signalling pathway phosphorylates NPR1 and inhibits turnover of the tryptophan permease. *The EMBO journal* 17(23), 6924–31.

Smets, B., Ghillebert, R., De Snijder, P., et al. 2010. Life in the midst of scarcity: adaptations to nutrient availability in Saccharomyces cerevisiae. *Current genetics* 56(1), 1–32.

Smith, D.L., McClure, J.M., Matecic, M., et al. 2007. Calorie restriction extends the chronological lifespan of Saccharomyces cerevisiae independently of the Sirtuins. *Aging cell* 6(5), 649–62.

Smoot, M.E., Ono, K., Ruscheinski, J., et al. 2011. Cytoscape 2.8: new features for data integration and network visualization. *Bioinformatics (Oxford, England)* 27(3), 431–2.

Stark, C., Breitkreutz, B.J., Chatr-Aryamontri, A., et al. 2011. The BioGRID Interaction Database: 2011 update. *Nucleic acids research* 39(Database issue), D698–704.

Stark, C., Breitkreutz, B.J., Reguly, T., et al. 2006. BioGRID: a general repository for interaction datasets. *Nucleic acids research* 34(Database issue), D535–9.

Steffen, K.K., MacKay, V.L., Kerr, E.O., et al. 2008. Yeast life span extension by depletion of 60s ribosomal subunits is mediated by Gcn4. *Cell* 133(2), 292–302.

Stephen, D.W., Rivers, S.L., and Jamieson, D.J. 1995. The role of the YAP1 and YAP2 genes in the regulation of the adaptive oxidative stress responses of Saccharomyces cerevisiae. *Molecular microbiology* 16(3), 415–23.

Stojmirović, A. and Yu, Y.K. 2007. Information flow in interaction networks. *Journal of computational biology : a journal of computational molecular cell biology* 14(8), 1115–43.

Stojmirović, A. and Yu, Y.K. 2012. Information flow in interaction networks II: channels, path lengths, and potentials. *Journal of computational biology : a journal of computational molecular cell biology* 19(4), 379–403.

Teixeira, M.C., Cabrito, T.R., Hanif, Z.M., et al. 2011. Yeast response and tolerance to polyamine toxicity involving the drug : H+ antiporter Qdr3 and the transcription factors Yap1 and Gcn4. *Microbiology (Reading, England)* 157(Pt 4), 945–56.

Temple, M.D., Perrone, G.G., and Dawes, I.W. 2005. Complex cellular responses to reactive oxygen species. *Trends in cell biology* 15(6), 319–26.

Urban, J., Soulard, A., Huber, A., et al. 2007. Sch9 is a major target of TORC1 in Saccharomyces cerevisiae. *Molecular cell* 26(5), 663–74.

Wanke, V., Cameroni, E., Uotila, A., et al. 2008. Caffeine extends yeast lifespan by targeting TORC1. *Molecular microbiology* 69(1), 277–85.

Wedaman, K.P., Reinke, A., Anderson, S., et al. 2003. Tor kinases are in distinct membrane-associated protein complexes in Saccharomyces cerevisiae. *Molecular biology of the cell* 14(3), 1204–20.

Wei, Y. and Zheng, X.F.S. 2011. Nutritional control of cell growth via TOR signaling in budding yeast. *Methods in molecular biology (Clifton, N.J.)* 759, 307–19.

Wiatrowski, H.A. and Carlson, M. 2003. Yap1 accumulates in the nucleus in response to carbon stress in Saccharomyces cerevisiae. *Eukaryotic cell* 2(1), 19–26.

Wu, P., Shen, Q., Dong, S., et al. 2008. Calorie restriction ameliorates neurodegenerative phenotypes in forebrain-specific presenilin-1 and presenilin-2 double knockout mice. *Neurobiology of aging* 29(10), 1502–11.

Wullschleger, S., Loewith, R., and Hall, M.N. 2006. TOR signaling in growth and metabolism. *Cell* 124(3), 471–84.

Yiu, G., McCord, A., Wise, A., et al. 2008. Pathways change in expression during replicative aging in Saccharomyces cerevisiae. *The journals of gerontology. Series A, Biological sciences and medical sciences* 63(1), 21–34.

Zaman, S., Lippman, S.I., Zhao, X., et al. 2008. How Saccharomyces responds to nutrients. *Annual review of genetics* 42, 27–81.

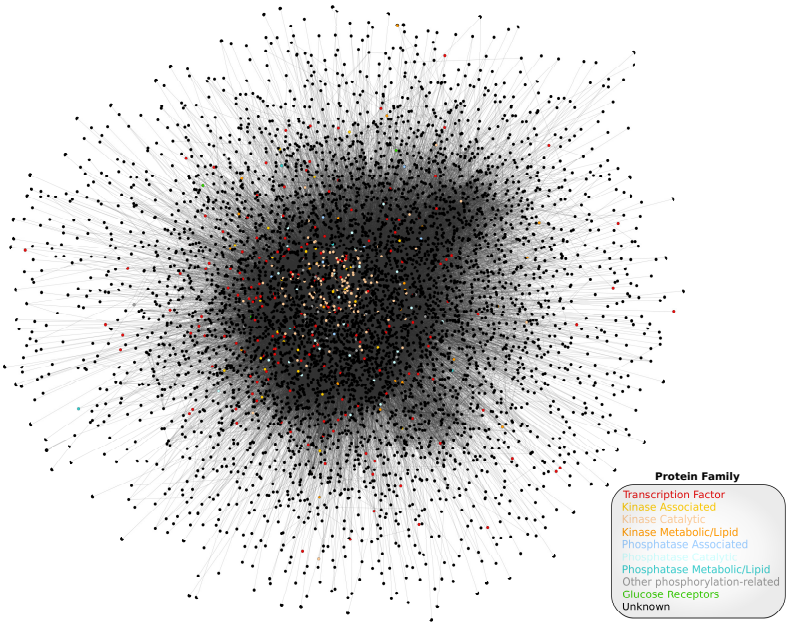

Figure 1: Yeast interactome, constructed using physical interactions extracted from the Biogrid dataset.

Nodes are color-coded using the family of their corresponding protein.

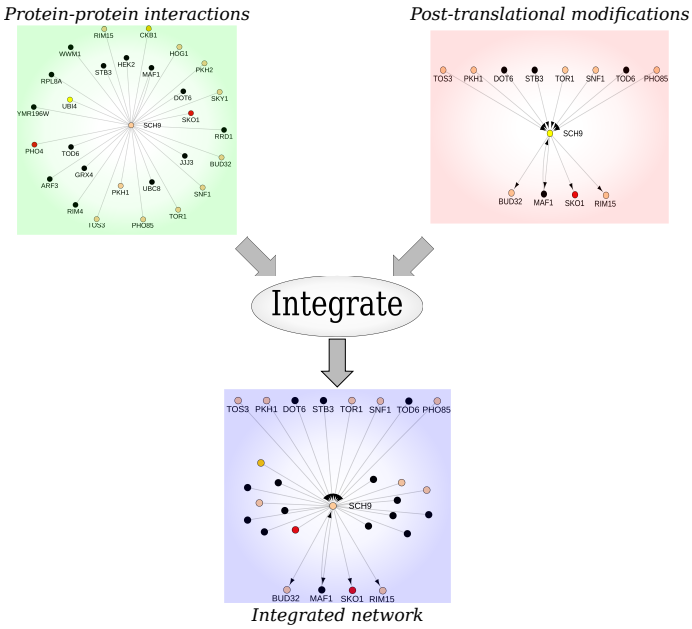

Figure 2: Example of network construction process around Sch9p. Protein-protein interactions (PPI) and post-translational modifications (PTM) were extracted from Biogrid dataset.

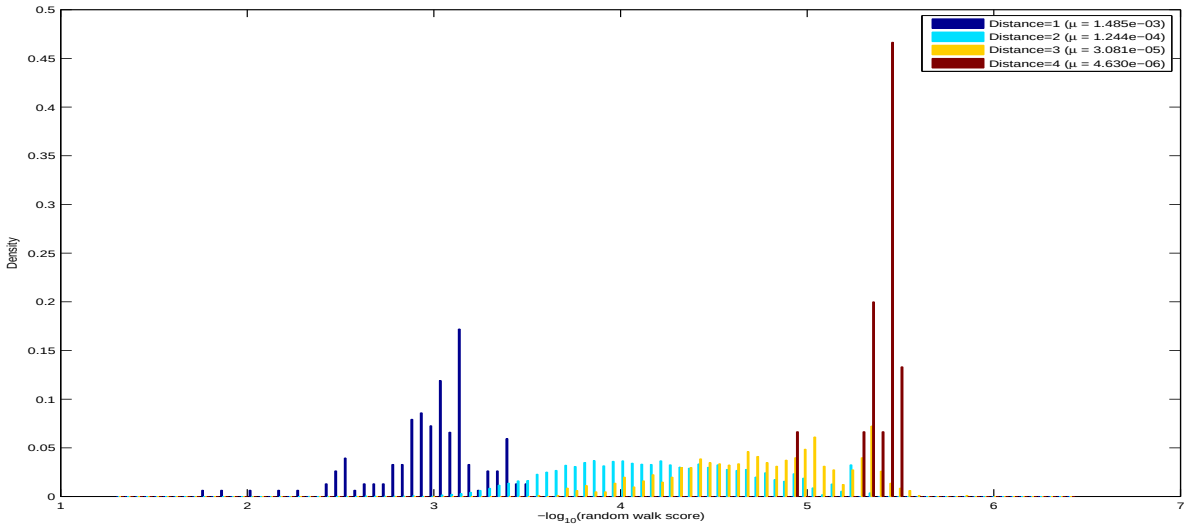

Figure 3: Distribution of random-walk scores, given the distance from TORC1, showing that random-walk scores are a function of distance, as well as multiplicity of paths.

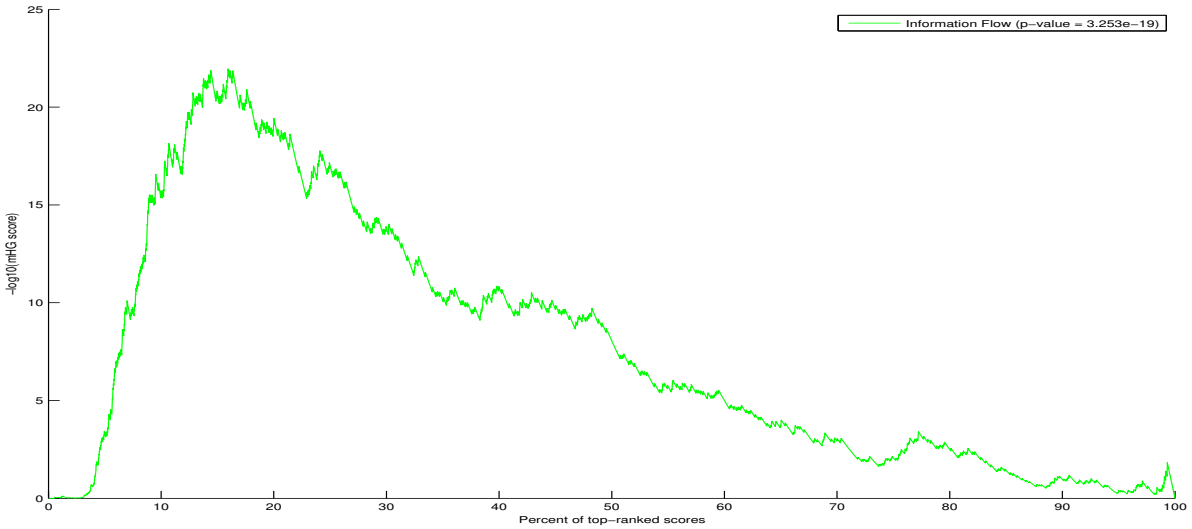

Figure 4: mHG score as a function of random-walk cut-off, computed using true-positives constructed from the set of genes that significantly change their expression level in response to rapamycin treatment. The peak of plot occurs at around top 10% of scores, with mHG p-value of 3.3e-19.

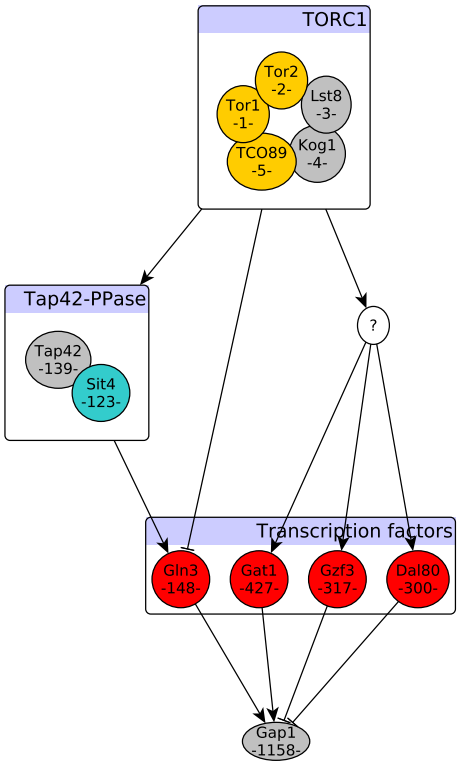

Figure 5: Signaling network from TORC1 to Gap1. Each node is annotated with its rank in random-walk scores from TORC1.

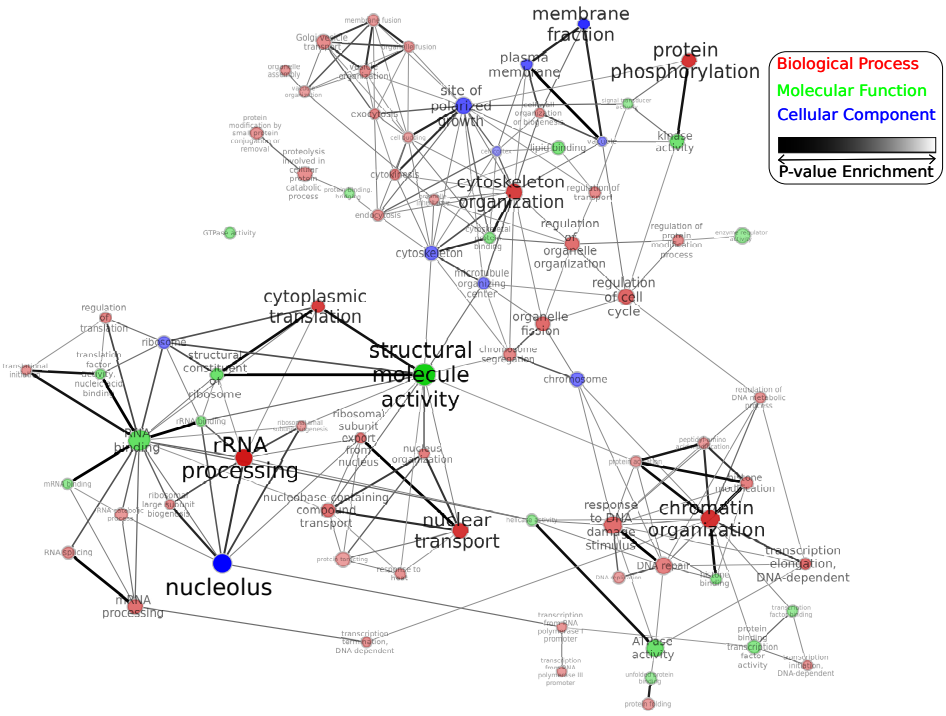

Figure 6: Enrichment map of yeast GOslim terms identified by mHG p-value, computed for ranked-list of genes from random-walk scores. Each node represents a GO term and edges represent the overlap between genesets of GO terms.

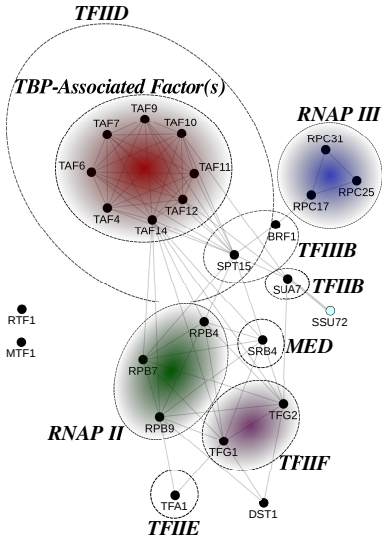

Figure 7: Induced subgraph in the yeast interactome by the enriched genes in the transcription initiation term.

Table 1: Top-ranked transcription factors with high confidence scores, postulated to be TORC1-dependent regulators.

TF ORF	TF name	TF rank	TF confidence
YLR403W	SFP1	22	43.5048
YER040W	GLN3	148	57.7734
YML007W	YAP1	618	24.3672
YEL009C	GCN4	638	4.822
YHR084W	STE12	825	2.9668