High participation ratio genes in the interaction network structure

Nastaran Allahyari¹, Nima Abedpour², Ali Hosseiny¹, G. R. Jafari¹,³*

¹ Department of Physics, Shahid Beheshti University, Evin, Tehran, Iran, ² Department of Translational Genomics, University of Cologne, Weyertal, Cologne, Germany, ³ Institute of Information Technology and Data Science, Irkutsk National Research Technical University, Lermontova, Irkutsk, Russia

* g.jafari@sbu.ac.ir (G. R. J)

Abstract

Genes have specific functional roles, however, since they are dependent on each other, they can play a structural role within a network structure of their interactions. In this study, we analyze the structure of the gene interaction network and detect the most contributing genes through the random matrix theory. Specifically, we compare the interaction network of essential and nonessential genes of the yeast Saccharomyces cerevisiae. Most remarkably, this well-established combined framework by measuring the node participation ratio (NPR) index helps detect important genes, which control the insightful structural patterns in the underlying networks. Results indicate that the essential genes have higher values of NPR rather than the nonessential ones which means that they have the most contribution to the network structure. It is worth mentioning that among all essential genes, the NPR value of 5 significant ones is considerably higher than the other essential genes, and also the same is for 15 significant nonessential genes compared to the others. Thus, the significant essential genes strongly manage their network structure, while the significant nonessential genes, besides their global contributions, have weak effects on their network structure. Most strikingly, these genes existing in a limited number of structural patterns are responsible for the specific bioprocesses which are the signature of their networks.

Introduction

The abundance of available data at the genomic level and the considerable progress in high-tech methods in data collection, instead of gene-based biology, has resulted in network biology [1,2]. When a gene interaction network is analyzed, the interaction between two genes is interestingly affected by the other interconnections they participate within [3,6]. Indeed, the dependency of interactions through a network can create a structure where the emergence of collective behavior may occur. Thus, uncovering the mechanism leading to collective behavior in networks is considered a significant subject in biology. To this aim, we explore the weighted, signed, and undirected networks of genetic interaction profile similarities of Saccharomyces cerevisiae. In this paper, to obtain a deeper understanding of the role and importance of interaction patterns in our networks, a time-efficient and cost-effective mathematical tool, random matrix theory (RMT) is applied. Through this study, the crucial questions are: Do the genetic interaction networks display specific evidence to deviate from randomness and clarify heterogeneous structures [7]? How to detect genes that
play a leading role in forming heterogeneities in the network? This work, by RMT techniques, not only deals with the importance of structural patterns in the networks but also helps in detecting genes, which are significant for the networks' identity.

There is a long history in which the non-randomness of gene interaction networks has been studied by considering motifs as the simplest units of network structure. Here, we explore the genetic interaction networks compared to the shuffled ones in the context of RMT. This method was initially introduced by Eugene Wigner to study the spectrum of the nuclei of heavy atoms where a large number of nucleons are interacting. It has recently demonstrated its outstanding success in analyzing complex systems in various fields such as quantum chaos, mobility, ecological, and financial networks. Moreover, it has provided a valuable framework to study complex biological networks. One application of RMT in biological networks is to study the spectral fluctuation of interaction matrices indicating the Gaussian orthogonal ensemble (GOE) of RMT. Also, it helps the overall functioning of altered complex molecules and predicting the connections among the molecules as one of the main goals in modern biology. Moreover, through RMT, the cellular roles of some unknown genes are predicted in terms of performance. Besides, due to inadequacy in the theoretical study to analyze gene interaction networks in cancer cells, RMT plays a crucial role to predict the behavior of a large complex interaction network from the behavior of a finite size network. Furthermore, as another application, information about the structure and dynamics of a network, through the localization properties of the gene network is achieved via the eigenvalues of its adjacency matrix. Additionally, RMT has helped solve the challenge of revealing all of the gene-gene interactions that are large in number. As a recent application, we can refer to where eigenvalues and eigenvectors denoise single-cell data, which provides the opportunity to identify new cellular states. Margaliot et al. have developed a new theoretical framework using tools from RMT to analyze the steady-state production rate affected by many stochastic “local” factors, as well.

Accordingly, we are interested in determining the heterogeneities of our gene interaction networks and identifying the genes that play a significant role in this structure. In this work, we explore the genome of Saccharomyces cerevisiae, as a whole system to understand the involved complexity. Here, we focus on the gene interaction similarity networks of about 1000 essential genes and 4500 nonessential genes. Costanzo and his colleagues have produced the data. Thus, we lead our analysis in two groups of genes, namely, essential and nonessential. By considering the threshold taken in their study to plot the networks, the essential gene similarity network is more densely connected compared to the corresponding nonessential network. Thus, it can be deduced that essential genes play the role of hubs in the global interaction network. Moreover, functional relationships in the gene similarity network for essential genes are stronger and their power of function prediction provided higher accuracy.

The rest of the paper is organized as follows. First, we demonstrate through the structural parameters namely degree, clustering coefficient, and assortativity coefficient that the structure of both essential and nonessential gene networks is far from randomization. Second, we explore the genetic interaction networks via eigenvalue distributions in comparison with their corresponding shuffled versions to analyze the existence of heterogeneous structures. Since in random networks, the spectrum of eigenvalues fits a semi-circle, we discuss that the deviation from such behavior is a sign of structure in the pairwise interactions. Next, we study the nearest neighbor spacing distribution for both networks to understand the universality and to check if we are allowed to apply RMT in case they follow the GOE of RMT. Then, we investigate the localization properties based on the participation of all genes in each eigenvector via calculating the inverse participation ratio in the gene networks. We discuss that the
analysis of the nodes with the top contributions in the localized eigenvectors demonstrates their importance in the structural patterns of the networks. Finally, as a novel index in RMT, we measure node participation ratio (NPR) of the important genes in all eigenvectors to separate significant genes which have higher NPR than the boundary of the shuffled networks. Interestingly, the selected significant genes leading to the structural patterns of each network have the most responsibility for the specific bioprocesses which are merely appointed to the same network.

**Materials and methods**

**Data description**

**Essential and nonessential genes.** Saccharomyces cerevisiae is known as a helpful yeast since nearly all bioprocesses in eukaryotes can exist in it [45]. This paper, to achieve a deeper understanding of the complexity of its genome, analyzes its interaction similarity networks. Costanzo and his colleagues have provided and published the data [35] as three gene interaction similarity matrices, for essential genes, nonessential genes, and the combination of them as the global form [41,43]. What follows is the specific characteristics that based on them, the genes are categorized into two groups of essential and nonessential ones. The type of mutation generating these mutants for essential and nonessential genes are temperature-sensitive and deletion mutations, respectively. Besides, essential genes reveal a stronger functional connection and higher-accuracy gene function predictions for biological processes. Moreover, the biological processes specifically annotated by the essential genes are “cell polarity and morphogenesis”, “protein degradation”, and “ribosomal RNA (rRNA) and non-coding RNA (ncRNA) processing”. While, the bioprocesses such as “nuclear-cytoplasmic transport”, “ribosome biogenesis”, “peroxisome”, “metabolism and fatty acid biosynthesis”, “respiration, oxidative phosphorylation, mitochondrial targeting”, “MVB sorting and pH−dependent signaling”, and “tRNA wobble modification” are identified by nonessential genes. In addition, there are six common bioprocesses between both essential and nonessential gene interaction networks that are presented in Table 1 beside specific bioprocesses in each network [35].

**Table 1. Classification of bioprocesses in essential and nonessential gene networks.**

| GO biological process in: | Essential network | Nonessential network | Both networks |
|---------------------------|------------------|----------------------|--------------|
| cell polarity and morphogenesis | respiration, oxidative phosphorylation, mitochondrial targeting | mitosis and chromosome segregation |
| protein degradation / turnover | MVB sorting and pH−dependent signaling | vesicle traffic |
| rRNA and ncRNA processing | tRNA wobble modification | transcription and chromatin organization |
| – | ribosome biogenesis | glycosylation, protein folding/targeting, cell wall biosynthesis |
| – | peroxisome | DNA replication and repair |
| – | nuclear−cytoplasmic transport | mRNA processing |
| – | metabolism and fatty acid biosynthesis | – |

**Gene interaction matrices.** The data analyzed during the current study comprehensively covers ∼ 90% of all yeast genes and is publicly available at [http://boonelab.ccbr.utoronto.ca/supplement/costanzo2016/](http://boonelab.ccbr.utoronto.ca/supplement/costanzo2016/). We have worked
with data file S3 titled “Genetic interaction profile similarity matrices”. The steps taken to create this data are as follows:

1. Based on the growth rate of a colony size including two specific mutated genes, the genetic interaction score (epsilon) of them has been obtained.
2. Each mutated gene is crossed to an ordered array of other mutated genes. Then, a genetic interaction profile for that gene is constructed.
3. By calculating the Pearson correlation coefficient, the similarity between each two genetic interaction profiles has been produced.

Furthermore, the preprocessing procedure on data to reach the gene interaction networks is algorithmically plotted in Fig 1. To be more specific, in the genetic interaction matrix, the positive value for every two genes means how much their profiles are functionally similar to each other, and vice versa. Moreover, the larger the positive (negative) interaction value of the two genes, the more similarity (dissimilarity) between them. Also, elements with zero value indicate that there is no relation between those two genes functionally. Besides, the size of essential, nonessential, and global matrices are ~1000, 4500, and 5500, respectively. Thus, all our matrices here are undirected, signed, and weighted adjacency matrices. To have a vision of the construction of the networks, in Fig 1 they are graphed by using the spring-embedded layout algorithm in the Cytoscape software [46]. Only here, the same threshold taken previously by Costanzo et al. [35] has been used, i.e., if the interaction between two genes is more than 0.2, they are connected, otherwise, there is no connection between them. We then investigate the structural properties of these networks.

### Structural measures

**Degree** $k$. Several statistical indicators are proposed to comprehend specific features of the network [47]. The most primary structural measure of a network is the degree of a node $k_i$, which is defined in a weighted and signed network as the sum over the absolute of links’ weights of a node ($k_i = \sum_j |w_{ij}|$). Besides, mean degree $\langle k \rangle$ shows the degree of nodes in a network on average.

**Assortativity Coefficient.** To understand that high-degree components tend to be connected with similar (different) counterparts, the assortativity (disassortativity) coefficient is calculated. If similar (different) counterparts are connected, then the sign of this coefficient is positive (negative).

**Clustering Coefficient** $CC$. Another significant parameter is the clustering coefficient $CC$ of the network. The $CC$ of a network characterizes the overall tendency of nodes to form clusters or groups. It is defined as the ratio of the number of closed triplets it has by the number of open triplets it has. These closed triplets are complete subgraphs in the network, which are also known as cliques-3. As proposed in [48], in a signed, undirected, and weighted network, the clustering coefficient of a node $CC_i$ and of the network $CC$ can be calculated, respectively as

$$ CC_i = \frac{\sum_j \sum_k w_{ij} w_{jk} w_{ki}}{\sum_j \sum_k |w_{ij} w_{ik}|}, $$

$$ CC = \frac{\sum_i \sum_j \sum_k w_{ij} w_{jk} w_{ki}}{\sum_i \sum_j \sum_k |w_{ij} w_{ik}| + \sum_i \sum_j w_{ij}^2}, $$

where $w$ stands for the weight of the link between every two genes in the network.

In the following, to obtain a deeper understanding of the complexity and importance of structural patterns in our networks, the use of random matrix theory ($RMT$) as a
time-efficient and cost-effective mathematical tool is explained. Moreover, RMT as a powerful and frequently used method can provide an approach to identify significant nodes implicated in each network structure. Therefore, the global properties such as spectral properties, local properties such as eigenvalue fluctuations, and detection of significant genes appearing in the localized eigenvectors are studied.

Spectral techniques

Spectral distribution of eigenvalues $P(\lambda)$. Our approach to studying the genetic interaction similarity networks is RMT that provides a framework to go beyond the assumption that pair interactions are independent of each other in networks. When a large number of elements interact, and a random relation mainly drives the interactions, and the distribution of the components of the interaction matrix is Gaussian, then the $P(\lambda)$ has a semi-circle shape. Specifically, in a random network with zero mean value of its eigenvalues, all eigenvalues are collected around zero. But in a random network with a nonzero mean value of its eigenvalues, there is an eigenvalue just out of the bulk. Similarly, in the shuffled network, the distribution of eigenvalues has the same form as the bulk \[25\]. In the shuffling process, two elements of the matrix are chosen randomly
and are exchanged. As a result, neither the value of links nor their distribution is modified, but the correlation between links dissolves. Therefore, when the process is complete, the shuffled network has random connections with no specific structure [49]. Despite the shuffled networks, the original networks have a large class of eigenvalues out of the bulk. To define the boundary of the bulk, we shuffle the matrix, then put the largest eigenvalue aside. This experiment is repeated 100 times. Then, the mean value of the borders and the standard deviations of them are worked out. Accordingly, all eigenvalues at least three standard deviations away from the border are considered as the eigenvalues outside of the bulk.

**Nearest neighbor spacing distribution** $P(s)$. There are three categories of RMT depending on their symmetries, the Gaussian orthogonal ensembles (GOE), the Gaussian unitary ensembles (GUE), and the Gaussian symplectic ensembles (GSE). Here, as we study the case that the gene interaction matrices are real symmetric, thus, GOE of RMT is analyzed. In RMT, besides the $P(\lambda)$ as a global property, eigenvalue fluctuation which is related to the symmetry of the system as local properties is considered. To this aim, one of the most powerful techniques in RMT, i.e., the nearest neighbor spacing distribution $P(s)$ is studied. In the large matrix size, the $P(s)$ becomes a universal function. Here, “universal” denotes that the distribution is independent of any detail of the systems and is only concerned with its symmetry. In the GOE case,

$$P(s) = \frac{\pi s}{2} \exp\left(\frac{-\pi s^2}{4}\right),$$

where $s$ is the spacing between two adjacent eigenvalues normalized by the mean eigenvalue spacing. It is called the Wigner distribution. The Wigner distribution implies that the eigenvalues have mutual correlations and repel each other. Moreover, where the eigenvalues have no correlation and are randomly distributed, $P(s)$ becomes

$$P(s) = \exp(-s),$$

this is known as the Poisson distribution in RMT. Since in real systems the matrix size is finite, the method of rescaling of the eigenvalues called unfolding has to be applied. For details about the unfolding method, see [27]. Additionally, we tested the $P(s)$ obtained by the one-sample Kolmogorov-Smirnov test. Also, the significance level $\alpha$ is set at 0.05. When the p-value is larger than the significance level $\alpha$, the null hypothesis is not rejected.

**Inverse participation ratio (IPR)**. We used the inverse participation ratio (IPR) to analyze the localization properties of the eigenvectors, which has been introduced by Bell and Dean, and initially was applied in the context of atomic physics [50]. Specifically, IPR is one of the most frequently used techniques in RMT to identify nodes appearing in the localized eigenvectors. Throughout this study, we have calculated the eigenvalues and eigenvectors to determine the IPR which tries to provide a measure for participation rates of all nodes in each eigenvalue. IPR of each eigenvalue $\lambda_k$ is given by

$$IPR_k = \sum_{n=1}^{N} (u_n^k)^4,$$

where the $u_n^k$ stands for components of the related eigenvector, and $N$ is the number of nodes. The summation is over all elements in each eigenvector. The $IPR_k$ shows two extreme cases: (1) a vector with equal elements ($u_n^k = \frac{1}{\sqrt{N}}$) has $IPR_k = \frac{1}{N}$, (2) a vector with zero elements except one element ($u_n^k = 1$) has $IPR_k = 1$. Thus, the IPR expresses the reciprocal of the number of eigenvector elements that contribute remarkably [28]. Note that IPR defined as above separates the top contributing genes by keeping the threshold. The threshold is taken via the following steps. First, we shuffle the network for 100 iterations. Then, the eigenvalues that are out of the bulk of
the shuffled networks are selected. Finally, the mean value of IPR of shuffled versions plus three standard deviations is taken as the boundary. We additionally calculate the average IPR to measure an overall localization of the network as

$$\langle IPR \rangle = \Sigma_{k=1}^{N} IPR_k. \quad (6)$$

**Node participation ratio (NPR).** Despite IPR, which deals with each eigenvector, the node participation ratio (NPR) is a characteristic of each node. To calculate the NPR of each gene, we calculate the sum of its contributions in all eigenvectors as

$$NPR_n = \Sigma_{k=1}^{N} (u^k_n)^4. \quad (7)$$

The summation is over the participation of a node in all eigenvalues. If a node has a random performance in all eigenvalues, then its NPR is small, that is in extreme cases equal to $1/N$. If, however, a node has unique roles in a few numbers of eigenvectors, then the value of its NPR is high, i.e., at extreme it equals to one. Thus, the great value of NPR means that a node plays a critical role in a specific biological process. To identify genes with high NPR among those with top contributions in the localized eigenvectors, we performed the same procedure of taking the threshold for the IPR as mentioned above.

**Results**

**Structural properties of the networks.** The structural parameters of the networks, which are in line with our main research question, are presented in Table 2a. Besides, the clustering coefficient of each node $CC_i$, which is calculated through Eq (1), the clustering coefficient of the network $CC$ is achieved through Eq (2). The point is that the higher value of $CC$, the more presence of functional modules in the network [51]. In other words, the higher value of $CC$ means the presence of a high number of clique structures in a network [52], which as building blocks of a network make it more robust [53] and stable [54]. What follows is that the nonessential network has fewer cliques of order three compared to the essential one. This indicates that there is a deficiency of building blocks in the nonessential network, which may be leading to a more unstable system. It should be remarked that besides the aforementioned reason about the significance of $CC$, it is a deterministic indicator in defining the network’s topology. Therefore, to obtain a deeper insight into our networks’ topology, which unifies the mentioned results above, we examine our networks’ structure within the analytical expression for $CC$ of theoretical networks. The values of analytical expressions for $CC$ of theoretical networks, and that of our real data are presented in Table 2b. The result indicates that the value of $CC$ in our networks, as in the small-world networks, is between the values of that in the random and lattice versions. Additionally, the essential and nonessential gene networks exhibit approximately similar statistics for the assortativity coefficient. That is, as exhibited by most of the biological systems investigated under the network theory framework, our gene networks are disassortative [55]. Specifically, components with high degrees tend to be connected with different ones. Moreover, the density of networks by summing over absolute magnitude values of links divided by all possible links shows high sparsity in our networks. Besides the analyzed structural features, the vital differences between them, are revealed through spectral analysis. In the following, the global spectral properties, eigenvalue fluctuations, and properties of genes appearing in the localized eigenvectors are analyzed. These are known as the most frequently used techniques in RMT to understand the underlying complex system comprehensively.
Table 2. Network analysis.

(a) Network parameters.

|                  | Essential | Nonessential |
|------------------|-----------|--------------|
| Nodes            | 1,040     | 4,430        |
| Density of links | 0.006     | 0.005        |
| $\langle k \rangle$ | 6.366     | 24.207       |
| $k_{\text{max}}$ | 28.120    | 78.333       |
| Assortativity    | $-0.160$  | $-0.139$     |
| Clustering Coefficient | 0.025     | 0.007        |
| (IPR)            | 0.0031    | 0.0007       |
| (NPR)_{imp genes} | 0.0040    | 0.0008       |
| $\lambda_{\text{max}}$ | 37.25     | 39.29        |

(b) Clustering Coefficient (CC) for canonical topologies.

| Network type | Expression | Essential | Nonessential |
|--------------|------------|-----------|--------------|
| Lattice      | $\frac{3(k-2)}{4(k-1)}$ | $CC_{\text{ess}} < 0.610$ | $CC_{\text{non}} < 0.718$ |
| Random       | $\frac{k}{N}$ | $0.006 < CC_{\text{ess}}$ | $0.005 < CC_{\text{non}}$ |
| Small – world | $\frac{3k-2}{4k-1}(1-p)^3$ | $CC_R < 0.445 < CC_L$ | $CC_R < 0.523 < CC_L$ |

(a) Number of nodes, mean degree, maximum degree, density of links, assortativity, clustering coefficient, mean IPR, mean NPR of important genes, and maximum eigenvalue in both essential and nonessential gene networks are presented.

(b) $CC_{\text{ess}} = 0.025$, $CC_{\text{non}} = 0.007$, $CC_L$, and $CC_R$ indicate the clustering coefficient of essential, nonessential, lattice, and random networks, respectively. The value of $CC_{\text{ess}}$ and $CC_{\text{non}}$ are between the values of that in random and lattice versions as small-world networks.

Eigenvalue distribution $P(\lambda)$ and nearest neighbour spacing distribution $P(s)$. Now, a couple of questions arise: How do we evaluate the heterogeneities of the structures [57–59]? What happens if we analyze the nonessential gene interaction network? Are we left with a homogenous network with a random structure, then? To answer these questions, we can use spectral methods to look over the spectrum of eigenvalues [60,61]. Indeed, if the pairwise interaction of genes does not have a higher level of structure, then its matrix should have properties of random matrices. From RMT, we know that for a system that consists of a large number of elements, if the structure is limited to the pairwise level without a higher scale structure and if the probability density function of the pairwise interactions comes from a Gaussian distribution, then the $P(\lambda)$ has a semi-circle shape that centers around zero. As the result, we end up with a distinct $P(\lambda)$ of the genetic interaction networks (in green bars) and their corresponding shuffled versions (in a nice purple semi-circle) in Fig2A and Fig2B. Despite the shuffled networks, the genetic interaction networks possess several eigenvalues that have large values which their related eigenvectors can carry meaningful information concerning the system [62]. The serious difference between the spectrum of the genetic interaction networks and their shuffled networks leads us to a clear conclusion that the genetic interaction networks have a structure with their communities and are far from randomness. Thus, significant information should be carried by the structure of the genetic interaction networks beyond the local pairwise
interactions. About the second question, we know that the role of essential genes is so important in the network that their sole annihilation leaves a phenotype footprint in the reproduction process. However, we are surprised that nonessential genes whose sole role is not so important compared to essential ones, which have a vital phenotype footprint, play a role in shaping the heterogeneities in their network from the RMT perspective.

Further, as depicted in Fig2C and Fig2D, both the essential and nonessential gene networks follow Gaussian orthogonal ensemble (GOE) statistics of RMT. The distribution of spacing of unfolded consecutive eigenvalues follows the Wigner distribution (Eq (3)), not the Poisson distribution (Eq (4)). This fitness reflects that there is a correlation between eigenvalues. Thus, they repel each other; otherwise, there was no correlation between them. Therefore, it can be noted that both networks have a minimal amount of randomness which is referred to as random connections between nodes [63]. The universal GOE statistics displayed by both networks lead to two main results. First, both networks, even the nonessential network have robustness. Second, it establishes that these networks can be modeled using the GOE of RMT. Therefore, we can apply all the techniques developed under the well-established framework of RMT. In the following, by analyzing the localization properties of eigenvectors, the set of genes that have the most contribution to the structure in each network is achieved.

**Important genes through eigenvector localization.** Based on the inverse participation ratio (IPR) values calculated using Eq (5), the eigenvectors can be categorized into two groups. Some of them follow RMT predictions of Porter-Thomas distribution which corresponds to random interactions [64]. The other ones deviate from this universality and show localization which gives the system-dependent information [65]. The deviating part reveals important genes as explained below. The essential (nonessential) gene network yields 26 (52) top contributing genes corresponding to the top 7 (11) most localized eigenvectors. The IPR of the eigenvalues of the networks compared to their corresponding shuffled versions have been depicted in Fig3A and Fig3B. Moreover, the average IPR calculated by using Eq (6) is higher for the essential network than for the nonessential one. It demonstrates that the nonessential network is more random than the essential network. In the following section, the functions of these selected genes are discussed briefly.

**Functional properties of top contributing genes in the most localized eigenvectors.** The most important outcome of the functional analysis of top contributing essential genes in the most localized eigenvectors is that most of them are involved in the specific bioprocesses which particularly exist in the essential gene network (first column in Table 1). The first localized eigenvector has five contributing genes, of which YIL118W and YLR166C are responsible for “cell polarity and morphogenesis” which occurs only in the essential gene network, YGL116W in the cell controls the “mitosis and chromosome segregation”, and YLR129W, YER127W help in “tRNA and ncRNA processing” in the essential network. Besides, the second localized eigenvector consists of four top contributing genes, among which YFL005W and YGL233W and YLR166C as above are responsible for “cell polarity and morphogenesis”, and YER136W is responsible for “glycosylation, protein folding/targeting, cell wall biosynthesis”. Also, among the top contributing genes in the third most localized eigenvector, YKR086W is generally involved in “mRNA and tRNA processing”, YNL163C is implicated in “ribosome biogenesis”. Similarly, YPR103W controls “protein degradation/turnover”, and YMR005W, is found to handle “transcription and chromatin organization”. Moreover, the fourth localized eigenvector has two top contributing genes of which YBR198C and YDR145W are responsible for “transcription and chromatin organization”. What follows is that all the top contributing genes, in the fifth most localized eigenvectors, i.e., YIR010W and YMR117C and YGL093W have a major contribution to “mitosis and chromosome
Figure 2. Eigenvalue distribution and nearest neighbour spacing distribution for both essential and nonessential gene networks. (A) $P(\lambda)$ for the essential gene network, (B) $P(\lambda)$ for the nonessential gene network. (The plots depict a clear difference between distribution for both networks (in green) and their corresponding shuffled versions (in purple).) (C) $P(s)$ for the essential gene network, (D) $P(s)$ for the nonessential gene network. (The spacing of unfolded eigenvalues follows GOE statistics of the Wigner distribution. The green bars represent data points, and the solid purple (black) line represents the Wigner (Poisson) distribution. Besides, the one-sample Kolmogorov-Smirnov test in each of the segments is applied which shows the p-values larger than the significance level $\alpha = 0.05$.)

Among the top contributing genes of the sixth localized eigenvector, YDR498C and YLR078C have been found in “vesicle traffic”, YHR036W accounts for “nuclear-cytoplasmic transport”, besides “mitosis and chromosome segregation”, and YBL097W appear in “mitosis and chromosome segregation”. At last, all top contributing genes in the seventh localized eigenvalue contribute in “protein degradation/turnover”. Regarding the functional analysis of the top contributing nonessential genes in the most localized eigenvectors, all of them interestingly are responsible for “respiration, oxidative phosphorylation, mitochondrial targeting”, bioprocess, which in particular exist in the nonessential gene network. Besides, most of the nonessential genes in the localized eigenvectors (33 of 52) take participate in one of the other specific bioprocesses in the nonessential gene network, which is “ribosome biogenesis”. At the third stage, 28 genes are responsible for “transcription and chromatin organization”, and “mitosis and chromosome segregation” which are common processes between essential and nonessential gene networks. What follows is the consecutive processes that have less number of responsible important genes, such as “vesicle traffic” (24 genes),
“glycosylation, protein folding/targeting, cell wall biosynthesis” (19 genes), “DNA replication and repair” (17 genes). Also, “peroxisome” as a specific process in the nonessential gene network is devoted to 16 genes. Moreover, “metabolism and fatty acid biosynthesis”, as another specific process is handled through 8 important genes in the nonessential gene network. The three remained bioprocesses which are “tRNA wobble modification” and “MVB sorting and pH-dependent signaling”, and “mRNA processing” are driven with 6, 5, and 2 important genes in the nonessential gene network, respectively. Also, on average, most of the genes in each eigenvector are helping to four processes, that half of them are specific for the nonessential gene network, and the other half of them are common between essential and nonessential gene networks.

**Node participation ratio (NPR).** As the last step, to indicate genes that have played a significant role in a small portion of eigenvectors among those who live in the structure, we calculate the node participation ratio (NPR) of each important gene through Eq (7). Indeed, NPR designates to identify if a gene has randomly taken part in a large number of eigenvectors or it has selected relatively a small portion of eigenvectors to take part. We calculate the NPR of important genes achieved from IPR analysis and depict the spectrums in Fig3C and Fig3D. As can be seen, the values of NPR of essential genes are higher than those of nonessential genes. This means that essential genes compared to nonessential genes are strongly managing their network’s structure. Moreover, this implies that besides a global contribution among those significant genes in nonessential genes, they weakly affect their structure. Particularly, 5 (15) specific essential (nonessential) genes indicate higher values of NPR than the boundary of the shuffled version, which is achieved via the mean value of NPR in the shuffled version after 100 iterations then added to three standard deviations.

**Structural properties of localized genes with high value of NPR.** The top contributing genes in the most localized eigenvectors with high values of NPR reveal interesting structural properties. Specifically, most of them in the essential gene network lie in the low degree regime compared to the mean degree, remarkably. But the importance of these genes based on the analysis of their interactions via CC strikingly exhibits the existence of robust local structural patterns against external perturbations. Unlike the significant essential genes, the significant nonessential genes have higher degrees than their mean degree on average. Though, they still indicate a robust structure within the clustering coefficient higher than the mean degree, remarkably. The importance of these genes in their essential gene network is based on the analysis of their interactions via CC strikingly exhibits the existence of robust local structural patterns against external perturbations. Interestingly these important genes imply information on handling more purely specific bioprocesses which are merely appointed to the same network (first and second columns in Table 1). All these significant genes, the specific bioprocesses handled by them, the values of their degrees, and the values of their clustering coefficients in both essential and nonessential networks are presented in Table 3a and Table 3b, respectively.

**DISCUSSION**

This study has analyzed the undirected, weighted, and signed networks of essential and nonessential gene interaction networks of Saccharomyces cerevisiae. It is worth mentioning that to consider the whole story of a gene interaction system, both the functional relation and the containing structural patterns include beneficiary information. Besides the importance of this structure, detecting the significant genes with top contributions in the most localized eigenvectors is the second step to be taken. Therefore, our study has shown the following results:

There is a rich literature in the network theory in order to make us well acquainted with the structure and to distinguish non-random patterns through measuring many
Figure 3. Inverse participation ratio (IPR) of eigenvalues and node participation ratio (NPR) of genes for both essential and nonessential gene networks. (A) IPR for the essential gene network, (B) IPR for the nonessential gene network. (C) NPR for the essential gene network, (D) NPR for the nonessential gene network. (The green dots represent IPR(NPR) of the real data, and the purple dots indicate that of shuffled versions.)

structural parameters such as clustering coefficients, mean degree, assortativity, modularity, etc. While the essential gene network has less number of nodes with less mean degree, it has a more robust and stable structure due to the higher clustering coefficient. But it doesn’t mean that the nonessential gene network doesn’t benefit from structure since the clustering coefficient of nonessential gene network follows the same behavior of small-world structure as being between random and lattice measures.

This paper, to achieve a deeper understanding of the global structural patterns in our gene interaction networks, under the RMT framework, focuses on the spectrum of eigenvalues. Specifically, the first major question in this paper is about the signs of structures, especially in the nonessential gene network. Our results interestingly suggest a clear difference between the spectrum of the eigenvalues of the interaction matrices for both essential and nonessential gene networks with their corresponding shuffled counterparts, which follow the semi-circle distribution of the random network. Moreover, the greatest eigenvalue in a larger (expectedly highly structural) network is bigger than that of a smaller one. While the greatest eigenvalue in the nonessential gene network has nearly the same value in the essential gene network. Thus, there is a weekly structure in the nonessential gene network, which takes it away from randomness. In addition, through the spectrum of eigenvalues, the fluctuation of them as a strong analysis in RMT is analyzed. The results indicate that both essential and nonessential gene networks follow GOE statistics of RMT. Thus, due to the universality, we are
allowed to apply the \textit{RMT} method. On the other hand, this fitness with Wigner distribution reflects that both the networks have a minimal amount of randomness between their genes, which means there is a correlation between their eigenvalues so that they repel each other.

Another main goal is to identify genes that play a significant role in the structure of the networks. The genes that contribute to the heterogeneous structure are meaningful to us, and the value of their contribution is proportional to their eigenvector component. Inverse participation ratio (\textit{IPR}) is a measurement for the participation rates of all nodes in each eigenvalue. Regarding the \textit{IPR} of original networks, we have observed the following result: The average of \textit{IPR} in both networks is higher than that of shuffled ones. This result supports the fact that the networks have a substructure. Additionally, the higher value of \textit{IPR} in the essential gene network than the nonessential one indicates the existence of a stronger structure and more heterogeneities in this network. Furthermore, in Costanzo’s study, besides all specific features of essential (nonessential) genes, they are responsible for three (seven) bioprocesses that are unique for themselves. In the analysis of the \textit{IPR}, we observed that most of these top contributing genes in the most localized eigenvectors are responsible for those specific bioprocesses. Last but not least, this is the first time that node participation ratio (\textit{NPR}) is utilized to calculate the participation of each gene in all eigenvectors. A high value of \textit{NPR} identifies nodes that play a significant role in a small portion of the eigenvectors. Our result indicates that in both networks, the average value of \textit{NPR} of important genes is larger than that of corresponding shuffled versions. Moreover, essential genes have a larger value of \textit{NPR} than nonessential genes, which means that essential genes manage the structure strongly, while nonessential genes have a global role in their structure and weakly control it. Finally, in the essential (nonessential) gene network, we observed 5 (15) top contributing genes in the most localized eigenvectors with the highest value of \textit{NPR} that despite lying at a low (high) degree regime, exhibit structural significance with high clustering coefficients. This interesting result turns out to be more fascinating in the light of the functional analysis of these important genes, which mostly are responsible for the specific bioprocesses appointed to their network.

**Author contributions**

**Conceptualization:** Nastaran Allahyari, Gholam Reza Jafari, Ali Hosseyni, Nima Abedpour.

**Formal analysis:** Nastaran Allahyari, Gholam Reza Jafari, Ali Hosseyni, Nima Abedpour.

**Methodology:** Nastaran Allahyari, Gholam Reza Jafari, Ali Hosseyni, Nima Abedpour.

**Resources:** Nastaran Allahyari, Nima Abedpour, Gholam Reza Jafari.

**Software:** Nastaran Allahyari.

**Supervision:** Gholam Reza Jafari, Ali Hosseyni.

**Visualization:** Nastaran Allahyari.

**Writing:** Nastaran Allahyari.

**Acknowledgments**

N.A. would like to express her appreciation to M. Ghanbarzadeh, A. Kargaran, and Z. Moradimanesh for constructive comments that improved the manuscript.
References

1. Battiston F., et al. Networks beyond pairwise interactions: Structure and dynamics. Physics Reports. 2021; 874:1-92. https://doi.org/10.1016/j.physrep.2020.08.004

2. Barabasi AL, Gulbahce N, Loscalzo J. Network medicine: a network-based approach to human disease. Nat Rev Gen. 2011; 12:56–68. https://doi.org/10.1038/nrg2918 PMID: PMC3140052

3. Szappanos B., et al. An integrated approach to characterize genetic interaction networks in yeast metabolism. Nat Gen. 2011; 43:656-662 https://doi.org/10.1038/ng.846

4. Cordell HJ. Detecting gene–gene interactions that underlie human diseases. Nat Rev Gen. 2009; 10:392-404. https://doi.org/10.1038/nrg2579

5. Tong AHY., et al. Global mapping of the yeast genetic interaction network. Sci. 2004; 303(5659):808-813. https://doi.org/10.1126/science.1091317

6. Marbach D., et al. Wisdom of crowds for robust gene network inference. Nat Meth. 2012; 9:796-804. https://doi.org/10.1038/nmeth.2016

7. Orsini C., et al. Quantifying randomness in real networks. Nat Com. 2015; 6(8627). https://doi.org/10.1038/ncomms9627

8. Shen-Orr SS, Milo R, Mangan S, Alon U. Network motifs in the transcriptional regulation network of Escherichia coli. Nat Gen. 2002; 31:64–68. https://doi.org/10.1038/ng881

9. Sanz J, Navarro J, Arbusa A, Martin C, Marijuan PC, Moreno Y. The Transcriptional Regulatory Network of Mycobacterium tuberculosis. 2011; PLOS ONE. 6(7):e22178. https://doi.org/10.1371/journal.pone.0022178

10. Sanz J, Cozzo E, Borge-Holthoefer J, Moreno Y. Topological effects of data incompleteness of gene regulatory networks. BMC Systems Biology. 2012; 6(110). https://doi.org/10.1186/1752-0509-6-110

11. Lee TI., et al. Transcriptional Regulatory Networks in Saccharomyces cerevisiae. Sci. 2002; 298(5594):799-804. https://doi.org/10.1126/science.1075090

12. Allahyari N, Kargaran A, Hosseyni A, Jafari GR. The structure of gene-gene networks beyond pairwise interactions. arXiv:2111.08436. [Preprint]. 2021 [cited 2021 Oct 14]. Available from: https://arxiv.org/pdf/2111.08436.pdf

13. Kargaran A, Jafari GR. Heider and coevolutionary balance: From discrete to continuous phase transition. Phys Rev E. 2020; 103(5):052302. https://doi.org/10.1103/PhysRevE.103.052302

14. Wigner EP. On a class of analytic functions from the quantum theory of collisions. Annals of mathematics. 1951; 53(1):36-67. https://doi.org/10.1037/1969342

15. Papenbrock T, Weidenmüller HA. Random matrices and chaos in nuclear spectra. Rev Mod Phys. 2007; 79(3):997. https://doi.org/10.1103/RevModPhys.79.997

16. Louail T., et al. Uncovering the spatial structure of mobility networks. Nat Com. 2015; 6(6007). https://doi.org/10.1038/ncomms7007
17. Stone L. The stability of mutualism. Nat Com. 2020; 11(2648).
   https://doi.org/10.1038/s41467-020-16474-4

18. Grilli J, Rogers T, Allesina S. Modularity and stability in ecological communities. Nat Com. 2016; 7(12031).
   https://doi.org/10.1038/ncomms12031

19. Allesina S, Grilli J, Barabas G, Tang S, Aljadeff J, Maritan A. Predicting the stability of large structured food webs. Nat Com. 2015; 6(7842).
   https://doi.org/10.1038/ncomms8842

20. Urama TC, Ezepele PO, Nnanwa CP. Analysis of cross-correlations in emerging markets using random matrix theory. Mathematical Finance. 2017; 7(2):291-307.
   https://doi.org/10.4236/jmf.2017.72015

21. Saiedian M, Jamal T, Kamali MZ, Bayani H, Yasseri T, Jafari GR. Emergence of world-stock-market network. Phys A. 2019; 526(120792).
   https://doi.org/10.1016/j.physa.2019.04.028

22. Jamali T, Jafari GR. Spectra of empirical autocorrelation matrices: A random-matrix-theory–inspired perspective. Epl. 2015; 111(1):10001
   https://doi.org/10.1209/0295-5075/111/10001

23. Luo F, Srimani PK, Zhou J. Application of Random Matrix Theory to Analyze Biological Data. 2011; (Springer, New York, NY)

24. Luo F, Zhong J, Yang Y, Zhou J. Application of random matrix theory to microarray data for discovering functional gene modules. Phys Rev E. 2006; 73(3):031924.
   https://doi.org/10.1103/PhysRevE.73.031924

25. Rai A, Pawar AK, Jalan S. Network spectra for drug-target identification in complex diseases: new guns against old foes. Applied Network Science. 2018; 3(51).
   https://doi.org/10.1007/s41109-018-0107-y

26. Gibson SM, Ficklin SP, Isaacsom S, Luo F, Feltus FA, Smith MC. Massive-Scale gene co-expression network construction and robustness testing using random matrix theory. PLOS ONE. 2013; 8(2):e55871.
   https://doi.org/10.1371/journal.pone.0055871 PMID: PMC3567026

27. Kikkawa A. Random Matrix Analysis for Gene Interaction Networks in Cancer Cells. Sci Rep. 2018; 8(10607).
   https://doi.org/10.1038/s41598-018-28954-1

28. Jalan S, Solymosi N, Vattay G, Li B. Random matrix analysis of localization properties of gene coexpression network. Phys Rev E. 2010; 81(4):046118.
   https://doi.org/10.1103/PhysRevE.81.046118

29. Rai A, Menon V, Jalan S. Randomness and preserved patterns in cancer network. Sci Rep. 2014; 4(6368).
   https://doi.org/10.1038/srep06368

30. Rai A, Shinde P, Jalan S. Prognostic interaction patterns in diabetes mellitus II: A random-matrix-theory relation. Phys Rev E. 2015; 92(2):022806.
   https://doi.org/10.1103/PhysRevE.92.022806

31. Kikkawa A. Spectral analysis for gene communities in cancer cells. Complex Networks. 2020; 8(1):cnaa005.
   https://doi.org/10.1093/comnet/cnaa005

32. Frost HR, Amos CI, Moore JH. A global test for gene-gene interactions based on random matrix theory. Genetic epidemiology. 2016; 40(8):689-701.
   https://doi.org/10.1002/gepi.21990
33. Aparicio L, Bordyuh M, Blumberg AJ, Rabdan R. A Random Matrix Theory Approach to Denoise Single-Cell Data. Patterns. 2020; 1(3):100035. 
https://doi.org/10.1016/j.patter.2020.100035

34. Margaliot M, Huleihel W, Tuller T. Variability in mRNA translation: a random matrix theory approach. Sci Rep. 2021; 11(5300). 
https://doi.org/10.1038/s41598-021-84738-0

35. Costanzo M., et al. A global genetic interaction network maps a wiring diagram of cellular function. Sci. 2016; 353(6306). 
https://doi.org/10.1126/science.aaf1420

36. Blomenand VA., et al. Gene essentiality and synthetic lethality in haploid human cells. Sci. 2015; 350(6264):1092-1096. 
https://doi.org/10.1126/science.aac7557

37. Winzeler EA., et al. Functional characterization of the S. cerevisiae genome by gene deletion and parallel analysis. Sci. 1999; 285(5429):901-906. 
https://doi.org/10.1126/science.285.5429.901

38. Giaever G., et al. Functional profiling of the Saccharomyces cerevisiae genome. Nat. 2002; 418:387–391. https://doi.org/10.1038/nature00935

39. Wang T., et al. Identification and characterization of essential genes in the human genome. Sci. 2015; 350(6264):1096-1101. 
https://doi.org/10.1126/science.aac7041

40. Costanzo M., et al. The Genetic Landscape of a Cell. Sci. 2010; 327(5964):425-431. 
https://doi.org/10.1126/science.1180823

41. Baryshnikova A., et al. Quantitative analysis of fitness and genetic interactions in yeast on a genome scale. Nat Met. 2010; 7:1017–1024. 
https://doi.org/10.1038/nmeth.1534

42. Deshpande R., et al. A comparative genomic approach for identifying synthetic lethal interactions in human cancer. AACR. 2013; 73(20). 
https://doi.org/10.1158/0008-5472.CAN-12-3956

43. Li Z., et al. Systematic exploration of essential yeast gene function with temperature-sensitive mutants. Nat Bio. 2011; 29:361-367. 
https://doi.org/10.1038/nbt.1832

44. Levine JM, Bascompte J, Adler PB, Allesina S. Beyond pairwise mechanisms of species coexistence in complex communities. Nat. 2017; 546:56–64. 
https://doi.org/10.1038/nature22898

45. Parapouli M, Vasileiadis A, Afendra AS, Hatziloukas E. Saccharomyces cerevisiae and its industrial applications. AIMS Microbiol. 2020; 6(1):1-31. 
https://doi.org/10.3934/microbiol.2020001

46. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T. Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res. 2003; 13:2498–2504. 
https://doi.org/10.1101/gr.1239303

47. Boccaletti S, Latora V, Moreno Y, Chavez M, Hwang DU. Complex networks: Structure and dynamics. Physics Reports. 2006; 424(4-5):175-308. 
https://doi.org/10.1016/j.physrep.2005.10.009
48. Costantini G, Perugini M. Generalization of Clustering Coefficients to Signed Correlation Networks. PLOS ONE. 2014; 9(2):e88669. https://doi.org/10.1371/journal.pone.0088669

49. Leskovec J, Huttenlocher D, Kleinberg J. Signed networks in social media. Proc. SIGCHI Conf. Hum. Factors Comput. Syst. 2010; 22:1361-1370. https://doi.org/10.1145/1753326.1753532

50. Bell R, Dean P. Atomic vibrations in vitreous silica. Discussions of the Faraday society. 1970; 50:55–61. https://doi.org/10.1039/DF9705000055

51. Barabasi AL, Oltavi ZN. Network biology: understanding the cell’s functional organization. Nat Rev Gen. 2004; 5(2):101-113. https://doi.org/10.1038/nrg1272 PMID: 14735121

52. Watts DJ, Strogatz SH. Collective dynamics of ‘small-world’ networks. Nat. 1998; 393(6684):440–442 https://doi.org/10.1038/30918 PMID: 9623998

53. Alon U. An introduction to Systems Biology: Design principles of Biological Circuits. Chapman and Hall/CRC, 2006; London.

54. Dwivedi SK, Jalan S. Emergence of clustering: Role of inhibition. Phys Rev E. 2014; 90(3):032803. https://doi.org/10.1103/PhysRevE.90.032803

55. van Wijk BCM, Stam CJ, Daffertshofer A. Comparing brain networks of different size and connectivity density using graph theory. PLOS ONE. 2010; 5(10): e13701. https://doi.org/10.1371/journal.pone.0013701 PMID: PMC2965659

56. Shinde P, Yadav A, Rai A., et al. Dissortativity and duplications in Oral cancer. Eur. Phys J. B. 2015; 88(197). https://doi.org/10.1140/epjb/e2015-60426-5

57. Cimini G, Squartini T, Saracco F, Garlaschelli D, Gabrielli A, Caldarelli G. The statistical physics of real-world networks. Nat Rev Phys. 2019; 1:58–71. https://doi.org/10.1038/s42254-018-0002-6

58. Moradimanesh Z, Khosrowabadi R, Eshaghi Gordji M, Jafari GR. Altered structural balance of resting-state networks in autism. Sci Rep. 2021; 11(1):1966. https://doi.org/10.1038/s41598-020-80330-0

59. Kargaran A, Ebrahimi M, Riazi M, Hosseiny A, Jafari GR. Quartic balance theory: Global minimum with imbalanced triangles. Phys Rev E. 2020; 102(1):012310. https://doi.org/10.1103/PhysRevE.102.012310

60. Rai A, Pradhan P, Nagaraj J, Lohitesh K, Chowdhury R, Jalan S. Understanding cancer complexome using networks, spectral graph theory and multilayer framework. Sci Rep. 2017; 7(41676). https://doi.org/10.1038/srep41676

61. Chan THH, Louis A, Tang ZG, Zhang C. Spectral Properties of Hypergraph Laplacian and Approximation Algorithms. ACM. 2018; 65(3):1–48. https://doi.org/10.1145/3178123

62. Safavi Mobarhana NS, Saeedi A, Rahnamay Roodposhti F, Jafari GR. Network trending: leadership, followership and neutrality among companies: A random matrix approach. Phys A. 2016; 462:858-863. https://doi.org/10.1016/j.physa.2016.06.067
63. Jalan S, Bandyopadhyay JN. Randomness of random networks: A random matrix analysis. EPL. 2009; 87(4):48010. [https://doi.org/10.1209/0295-5075/87/48010]

64. Zyczkowski K. [Quantum mechanical eigenvector statistics of chaotic systems]. [153–168] Quantum Chaos [Cerdeira, H. A., Ramaswami, R., Gutzwiller, M. C. and Casati, G. (ed.)] (World Scientific, Singapore, 1991).

65. Plerou V, Gopikrishnan P, Rosenow B, Nunes Amara, LA, Guhr T, Stanley HE. Random matrix approach to cross correlations in financial data. Phys. Rev. E. 2002; 65(6):066126. [https://doi.org/10.1103/PhysRevE.65.066126]
Table 3. Classification of important genes with high NPR values.

(a) Essential genes.

| Gene   | Biological process                                                                 | Degree | Clustering Coefficient |
|--------|------------------------------------------------------------------------------------|--------|------------------------|
| YGL116W | protein degradation, mitosis and chromosome segregation                            | 4.5    | 0.073                  |
| YLR166C | cell polarity and morphogenesis, protein degradation, vesicle traffic               | 5.7    | 0.040                  |
| YGL233W | cell polarity and morphogenesis, vesicle traffic                                    | 7.0    | 0.042                  |
| YDR145W | cell polarity and morphogenesis, protein degradation                               | 3.4    | 0.023                  |
| YLR094W | protein degradation/tunover, cell polarity and morphogenesis                        | 1.5    | 0.051                  |

(b) Nonsential genes.

| Gene   | Biological process                                                                 | Degree | Clustering Coefficient |
|--------|------------------------------------------------------------------------------------|--------|------------------------|
| YEL016W | Nuclear – cytoplasmic Transport, Mitosis and Chromosome Segregation, Ribosome Biogenesis, Respiration, DNA Replication and Repair | 32.9   | 0.009                  |
| YGL134W | Peroxisome, Glycolysis, Transcription and Chromatin Organization, Respiration, tRNA Wobble Modification, Nuclear – cytoplasmic Transport, Vesicle Traffic, Ribosome Biogenesis MVB Sorting and pH – dependent Signaling | 45.1   | 0.009                  |
| YCL016C | Respiration, DNA Replication and Repair, Mitosis and Chromosome Segregation         | 35.8   | 0.008                  |
| YJR030W | Mitosis and Chromosome Segregation, Respiration, Ribosome Biogenesis, Transcription and Chromatin Organization, DNA Replication and Repair | 40.3   | 0.010                  |
| YER016W | Respiration, Vesicle Traffic, Mitosis and Chromosome Segregation, Transcription and Chromatin Organization | 36.5   | 0.013                  |
| YDL062W | Respiration, Vesicle Traffic, Mitosis and Chromosome Segregation, Transcription and Chromatin Organization, Ribosome Biogenesis, DNA Replication and Repair | 22.7   | 0.013                  |
| YDR095C | Respiration, Transcription and Chromatin Organization                               | 3.9    | 0.008                  |
| YDR164C | Respiration, Vesicle Traffic, Mitosis and Chromosome Segregation, Ribosome Biogenesis, DNA Replication and Repair | 23.4   | 0.007                  |
| YMR071C | Ribosome Biogenesis, Mitosis and Chromosome Segregation, Respiration, Vesicle Traffic | 33.6   | 0.014                  |
| YHR012W | Respiration, Vesicle Traffic, Peroxisome, Ribosome Biogenesis, Glycolysis, Mitosis and Chromosome Segregation, MVB Sorting and pH – dependent Signaling, Nuclear – cytoplasmic Transport, tRNA Wobble Modification | 24.1   | 0.011                  |
| YGL136C | Ribosome Biogenesis, Respiration, Glycolysis, Transcription and Chromatin Organization, Ribosome Biogenesis, Vesicle Traffic, DNA Replication and Repair | 13.4   | 0.010                  |
| YHR141C | Ribosome Biogenesis, Respiration, Glycolysis, Transcription and Chromatin Organization, Metabolism and Fatty Acid Biosynthesis, DNA Replication and Repair | 22.0   | 0.011                  |
| YLR027W | Respiration, Glycolysis, Transcription and Chromatin Organization, Metabolism and Fatty Acid Biosynthesis, Ribosome Biogenesis, DNA Replication and Repair | 20.6   | 0.013                  |
| YPL275W | Respiration, Ribosome Biogenesis MVB Sorting and pH – dependent Signaling, DNA Replication and Repair, Metabolism and Fatty Acid Biosynthesis, tRNA Wobble Modification | 30.5   | 0.014                  |
| YJR159W | Respiration, Glycolysis, tRNA Wobble Modification, Ribosome Biogenesis              | 28.9   | 0.015                  |