Protein Signaling Pathway: Network Centrality Analysis

Inva Koçiaj[a], Eliana Ibrahimi[b], and Dode Prenga[a]

[a] *University of Tirana, Faculty of Natural Sciences, Department of Physics, Tirana, 1001, Albania*
[b] *University of Tirana, Faculty of Natural Sciences, Department of Biology, Tirana, 1001, Albania*

Abstract

The topology of a biological network may differ, depending on the type of elements and the interactions between them. A protein signaling pathway considered as a directed network can be studied through the centralities analysis. The knowledge about the centralities is crucial to determine the most important network nodes, which can somehow define the further analysis needed for that network. Here we make a structural analysis of two networks, the AMPK-signaling pathway, and the mTOR-signaling pathway. The commonalities between these two networks are visible, and calculation of the centralities for all the nodes of the two networks show that they both have the same most important nodes, meaning that the signal inside the networks passes mostly through the same most important proteins such as AMPK, mTOR, RHEB, Akt, ELK1.2, and TSC1/2. However, we aim to go beyond this process. Our purpose is to make a dynamic study of a new bigger network composed of the most important nodes of these two elementary nodes. In this way, we can better predict all these vital proteins' effects on other elements, inside or outside the network.

Keywords: Network, AMPK-mTOR Pathways, Topology, Analysis, Centrality

1. Introduction

Knowledge about the topology of any network is very important not only for understanding how its elements are arranged but also to know the relations between them. As there are different types of networks, there exist different types of network topologies, and to determine the right one, network analysis is required. This analysis corresponds to the graph theory application, according to which the network components are modeled as nodes and the connections between them are modeled as edges that show the type of relationships that exist between these nodes [1]. Thus, the graph corresponding to a specific system (network) is a set of vertices (nodes) and a set of edges (links) that connect a pair of nodes.

Different networks such as computer networks, social networks, and disease networks have been studied for years, while biological systems have only been studied for a few years from now. Theoretical methods used for this purpose have continuously increased the interest among scientists to investigate more about the topology of a biological network and not only.

Biological systems can be introduced by undirected graphs, directed graphs, or weighted graphs. In each case, nodes represent genes, proteins, enzymes, or other metabolic and transcription elements, whereas links represent several interactions such as physical interactions, signaling pathways, co-expression, activation, and inhibition. Here, we present a study of the structure analysis of two signaling pathways; AMP-activated protein kinase (AMPK) signaling pathway, and target of mammalian rapamycin (mTOR) signaling pathway, both created and presented by SIGNOR [2]. We make the structural analysis of these two different protein signaling pathways by using Cytoscape [3]. We analyze both networks' centralities, compare them to each other, and go further with the analysis. For all elements, in both

Proccedings of RTA-CSIT 2021, May 2021, Tirana, Albania
EMAIL: inva.kociaj@gmail.com (I. Kociaj);
eliana.ibrahimi@fshn.edu.al (E. Ibrahimi);
dode.prenga@fshn.edu.al (D. Prenga)

networks, we observe all centralities with the purpose to find out which are the most important elements. Two protein kinases, such as AMPK and mTOR, are on the focus of both signaling pathway networks, and interestingly it is noticed that in both of them, the biological signal is mostly transmitted through the same elements (proteins). In these circumstances, our purpose is not only to show the network analysis but what is more important is that we aim to prove that the role and the importance of these common proteins, found in these two networks, are the same in both of them. We believe that this outcome enables us to build a new bigger network, composed of all these examined nodes, and that can help us look for further answers. Continuing further, it is important to emphasize that despite the differences that result in the outputs of these two signaling pathways, there are also some bond connections between them because of the biochemical processes that they represent. For this reason, we believe that all of these outputs, presented also as phenotypes of the biological systems, can be considered for the new aimed network, that we want to proceed with, in the upcoming research work.

2. Background: AMPK and mTOR signaling Pathways

Cell growth is regulated by maintaining the balance between the positive regulation of anabolic pathways and the negative regulation of catabolic pathways. The mTOR and AMPK signaling pathways regulate growth and metabolism, with mTOR activating the anabolic processes and AMPK inducing a catabolic response when cells have low energy levels [4]. In this section, we give brief information on these two kinases and their role in breast cancer and the regulation of cell growth.

AMP-activated protein kinase regulates cell energy homeostasis. AMPK is activated when there is a fall in ATP level, resulting in the activation of catabolic processes and the inhibition of anabolic processes [3, 4]. This crucial metabolic sensor regulates protein and lipid metabolism based on the alterations of energy levels.

In the past decade, more studies have focused on AMPK since it appeared to be a targeting molecule for cancer therapy. Many research studies have focused on understanding the role of AMPK signaling pathways in the regulation of growth and the development of drug resistance in triple-negative breast cancer [5].

Expression of AMPK is correlated with breast cancer stage and distant metastasis, patients with positive expression of AMPK exhibit shorter overall survival and disease-free survival [6]. These results suggest AMPK as a possible prognostic biomarker for triple-negative breast cancer. Recent research has reported that AMPK is reduced by 90% in cancer tissues of primary breast cancer patients than normal breast epithelial cells [7] Decreased AMPK signaling and the negative correlation with cancer grade/metastasis shows that AMPK reactivation can prevent breast cancer [5].

The mTOR complex is part of the PI3K-related protein kinase family, and it is located on chromosome 1p36.2 [8, 9]. Several studies have reported that rapamycin is involved in antitumor activities and can inhibit cell division [10]. It has two protein complexes, mTORC1 and mTORC2, that have differences in elements and functions [11, 12, 13].

mTOR has an important role in gene transcription, protein translation, ribosome synthesis, and a fundamental regulatory role in cell growth, cell division, differentiation, and apoptosis [14]. It is also reported that mTOR has an important role in tumor growth and metabolism, and plays a crucial role in breast cancer. The protein components of the mTOR are encoded by oncogenes or tumor suppressor genes. The mTOR pathway depends on the activation or the inhibition of the pathway signaling. In breast cancer, activation of the mTOR pathway is evaluated to be as common as 70% of breast cancer overall [15].

3. Experimental Setup

Here, we work on two signaling pathway networks, firstly generated and presented from SIGNOR [2], and then imported and represented for further analysis in Cytoscape (Figure 1) [16]. Cytoscape is an open-source platform, for general-purpose modeling, that is used to visualize and analyze molecular interaction networks and signaling pathways of large-scale complex networks (especially biological systems).

Cytoscape's Core is Java-based that provides basic functionalities for integrating arbitrary data of different formats, imported from several sources. As previously mentioned, the two networks on focus here, are directly imported into Cytoscape, whose data were downloaded in "*sbml*" format from a public database such as SIGNOR.

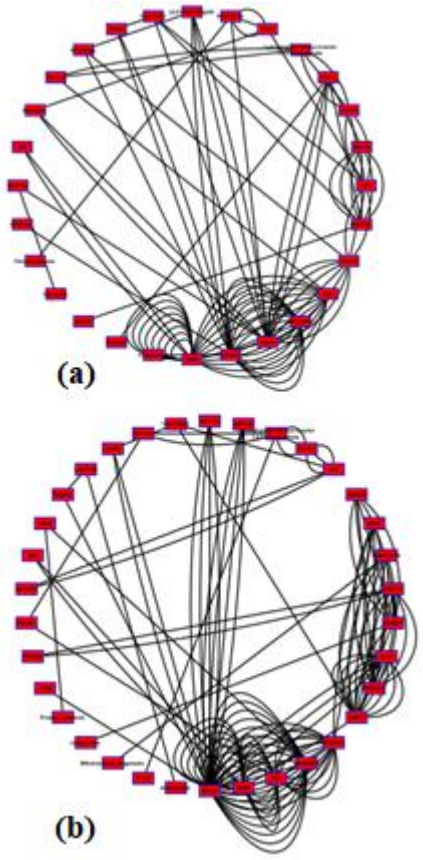


Figure 1: Two protein kinases signaling pathway networks. (a) AMPK – Signaling Pathway Network, composed of 28 nodes and 94 edges; (b) mTOR – Signaling Pathway Network, composed of 34 nodes and 131 edges. Both networks are imported from SIGNOR via NDE-x and reconstructed by Cytoscape for further analysis.

This is realized via NDE-x project (The Network Data Exchange) which provides to the researchers an open-source framework to store, modify, share, etc their networks. Because of its features, the networks were directly imported to Cytoscape without having the necessity of processing their data before. Furthermore, as we focus on centralities, the analysis of these networks was performed by running CentiScaPe2.2 (Figure 2 and Appendix), which is one of the Apps, incorporated into Cytoscape. As we firstly make a simple network analysis, we can see the networks' general characteristics such as the number of nodes, the number of edges, diameter, network radius, etc., that allows us to make a quick comparison between them (See Appendix).

Continuing further, a deeply structural analysis is made (for both networks), and the path followed toward the understanding of the most important nodes of the networks is through the analysis of the centralities for each element, such as diameter, average path, in-degree, and out-degree, eccentricity, radiality, closeness, stress, betweenness, centroid, eigenvector and bridging parameter [3, 17].

4. Results and Discussion

The evaluation of the centralities for each node of the network gives us a better understanding of its functionalities. Structural analysis through Cytoscape is based on several algorithms related to those hidden data that a simple general view of the network cannot find out [17, 19]. As running the CentiScaPe2.2 package, we reach the results about the centralities we are interested in and find out that there are some differences between the two networks, and there are also many other commonalities between them.

While network diameter and average path are two parameters that are more reliable for big networks, other centrality parameters can give important and reliable information applied even for small networks as our networks are. Thus, we analyze the results performed to see if the same elements have the same importance or not. Since both signaling pathways are directed networks, we can easily notice the interaction between nodes. According to the function that each node has (activation or inhibition) and the number of interactions it has with other nodes, we can pre-predict the most important nodes of the networks.

Generally speaking, these centralities are used to determine not only the importance of the network but also the importance of each node found in the network. However, it is recommended that to define the importance of

each node it is a necessity to compare several centralities simultaneously. In this way, we get more accurate information, and consequently, we can remove the less important nodes from the network without losing any important information [17].
Figure 2 shows the values of those general centrality parameters that are related to the networks, not a specific node, but on the other hand, obviously are different. Despite these differences, our analysis includes all general centralities, and the information we achieve from this is as follows:

The degree is a parameter determined as in-degree and out-degree and shows the number of the directed links connecting two nodes. While in-degree shows the links entering a node, out-degree shows the number of links going out from the target node. The nodes characterized by a high value of degree are considered central nodes as they play a key role in transmitting the signal in the network [18]. The central nodes seem to be the same elements in both networks, such as AMPK, mTOR, EIF4EBP1, INSR, ERK1/2.

Stress is another very important parameter that shows the importance of a node compared to another one. A stressed node is considered the one which is mostly reached from other nodes following the shortest paths. A stressed node can be an important node in a signaling pathway network that connects all other regulatory nodes, but it can be even a very involved node in cells' processes [17, 19]. So, from our observations, we realize that the most stressed common nodes are: IRS1, mTOR, RHEB, TSC1/2, PDPK1, Akt, PIK3CA, and phosphate group.

Betweenness and Centroid are two centralities that are strongly related to stress centrality [20]. They are complementary to each other that together give a better understanding of the required information. Nevertheless, a high value of the betweenness parameter is mostly related to a protein that keeps together the other communicated proteins of the network. In contrast, a node with a high value of centroid (higher than the average value of the network) is considered very important. It represents a protein involved in the coordination of other proteins' activities, leading to the participation in a cell regulatory activity. Proteins highlighted here for both networks are AMPK, mTOR, RHEB, TSC1/2, Akt, ULK1, RPS6KB1.

Eigenvector Centrality is a parameter that shows the importance of a node based on the assumption that high-scored nodes perform better than low-scored ones. According to this parameter, the importance of a specific node is determined not only by the number of its first neighbors connected to it but also from the value of the Eigenvector of each of these first neighbors [21]. In the biological meaning, this refers to a protein that interacts with several regulatory proteins simultaneously. Such a protein is considered a central node of the network, and the bigger the Eigenvector value is, the more this protein can generate other biological effects in the network. According to this parameter, the nodes in both networks that seem to be more important are ERK1/2, ULK1, AMPK, mTOR.

Bridging Centrality is another parameter that is different in both signaling pathway networks. This parameter indicates the ability of a node to position itself as a connecting bridge between two other nodes [18]. The bridging centrality parameter is bigger if a high degree parameter characterizes the nodes connected to the bridge node, and the proteins highlighted here are RHEB, PIK3CA, TSC1/2, ULK1.

It is very imortant to emphasize that all average values of all centralities are considered meaningful only when they are compared to the minimum and the maximum values that correspond to the nodes characterized by these values [20].

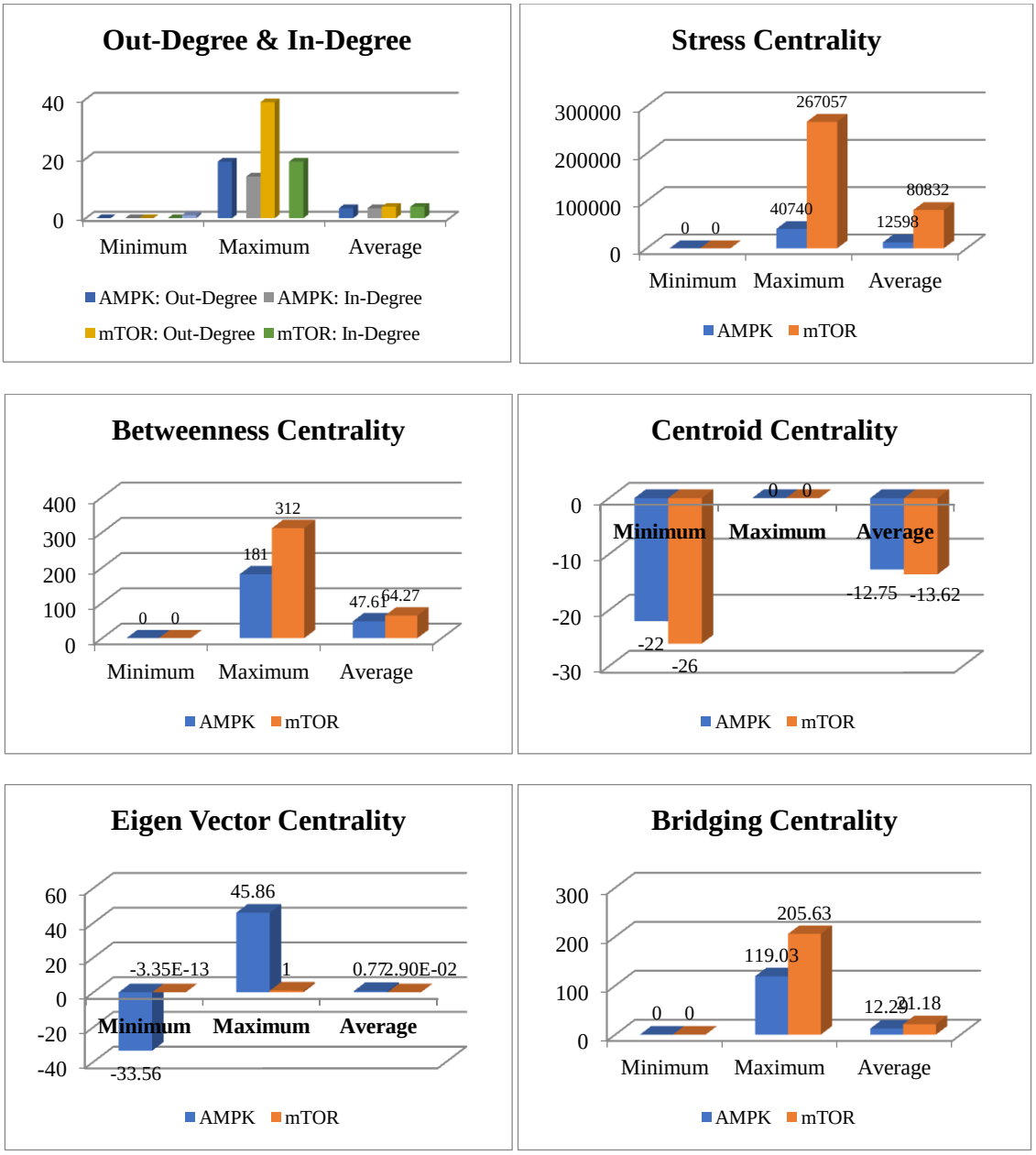


Figure 2: Main centralities analyzed by Cytoscape for both networks; AMPK-signaling pathway and mTOR-signaling pathway. Here are presented only those parameters that result to be different in both cases. Each graph gives the minimum, maximum, and average values of these centralities. The average value for each parameter is the one considered for the whole network whereas the extremities (minimum and maximum) are values corresponding to those nodes of the network that correspond to the minimum and the maximum values, respectively.

5. Conclusions and Future Work

The structure analysis realized through Cytoscape was made to define the most important nodes of the network, and this procedure is made for two different signaling pathway networks, such as the AMPK-signaling pathway and mTOR-signaling pathway. Both networks are characterized by direct connections between proteins whose signaling pathways are two protein kinases, AMPK, and mTOR. The relation between these two proteins is very strong, and usually, they are found together in most of the proteins' networks. Even here, we see many elements found in both networks and not only that, but we see that their function toward other

proteins is the same. To show the commonalities between networks, we analyze all the centralities via structural analysis. It is found that some of the centralities such as eccentricity, radiality, closeness give almost the same values for both networks, whereas some other centralities such as in/out-degree, stress, betweenness, centroid, Eigenvector, and bridging parameter show different values in both networks. Nevertheless, the fact that these networks are almost composed of the same elements, the differences between them are not so big. Moreover, since that mTOR-signaling pathway network has more elements inside it, it justifies these differences.

The network is better described if all these centralities are analyzed for each node of the network separately. To determine which the most important nodes are, we should pay the same attention to all these centralities at the same time. From this analysis, we define several nodes that are considered the most important ones, and interestingly some nodes are highlighted in almost all these centralities. Thus, we suggest that all those proteins highlighted from the analysis of the centralities are the most important nodes of the network. Even more interesting is the fact that we find the same most important proteins in both networks, giving us the permission to marginalize both networks and build a new one composed of the most important nodes/proteins so that the new network can be powerful in transmitting the signaling.

Our ongoing future work is the study of the biological network's dynamical evolution. The dynamical analysis, this time will be based on Boolean modeling, whose logical functions will be written only for the most important nodes of the network determined by this structure analysis.

6. References

[1] T. J. Grant, R. H. P. Janssen, and H. Monsuur, Network Topology in Command and Control: Organization, Operation and Evolution. 1st edition, IGI Global, US, (2014).

[2] L. Licata, P. L. Surdo, M. Iannuccelli, A. Palma, E. Micarelli, L. Perfetto, D. Peluso, A. Calderone, L. Castagnoli, G. Cesareni. SIGNOR 2.0, the Signaling Network Open Resource 2.0: 2019 update, J. Nucleic Acids Research, 48, (2020), 504–510. https://doi.org/10.1093/nar/gkz949

[3] M. Cline, M. Smoot, E. Cerami. Integration of biological networks and gene expression data using Cytoscape. Nat Protoc 2 (2007) 2366–2382 https://doi.org/10.1038/nprot.2007.324

[4] SK. Hindupur, A. González, MN. Hall. The opposing actions of target of rapamycin and AMP-activated protein kinase in cell growth control. Cold Spring Harb Perspect Biol 3. (2015) 7(8). doi: 10.1101/cshperspect.a019141.

[5] W. Cao, J. Li, Q. Hao, J V. Vadgama Y. Wu. Okay AMP-activated protein kinase: a potential therapeutic target for triple-negative breast cancer. Breast Cancer Res 21, (2019) 29. https://doi.org/10.1186/s13058-019-1107-2

[6] X. Huang, X. Li, X. Xie, F. Ye, B. Chen, C. Song, H. Tang. High expressions of LDHA and AMPK as prognostic biomarkers for breast cancer. Breast.,30 (2016) 39–46.

[7] SM. Hadad, L. Baker, PR. Quinlan, KE. Robertson, SE. Bray, G. Thomson, D. Kellock, LB. Jordan, CA. Purdie, DG. Hardie. Histological evaluation of AMPK signaling in primary breast cancer. BMC Cancer. (2009) 9:307.

[8] N.D. Golberg, A.M. Druzhevskaya, V.A. Rogozkin, I.I. Ahmetov Role of mTOR in the regulation of skeletal muscle metabolism. Hum. Physiol., 40 (2014). 580-588

[9] G.J. Wiederrecht, C.J. Sabers, G.J. Brunn, M.M. Martin, F.J. Dumont, R.T. Abraham. Mechanism of action of rapamycin: new insights into the regulation of G1-phase progression in eukaryotic cells Prog. Cell Cycle Res., 1 (1995), 53-71.

[10] W.M. Flanagan, G.R. Crabtree Rapamycin inhibits p34cdc2 expression and arrests T lymphocyte proliferation at the G1/S transition Ann. Ny. Acad. Sci., 696 (2006), 31-36.

[11] D. Sarbassov, S.M. Ali, D.H. Kim, D.A. Guertin, R.R. Latek, H. Bromage, D.M. Sa batini. A novel binding partner of mTOR, defines a rapamycin insensitive and raptor-independent pathway that regulates the

cytoskeleton Curr. Biol., 14 (2004), 1296-1302.

[12] R. Peterson, M. Laplante, C.C. Thoreen, Y. Sancak, Seong A. Kang, W.M. Kuehl, N.S. Gray, D.M. Sabatini. DEPTOR is an mTOR inhibitor frequently overexpressed in multiple myeloma cells and required for their survival Cell, 137 (2009), 873-886.

[13] D. Gao, L. Wan, H. Inuzuka, A.H. Berg, A. Tseng, B. Zhai, S. Shaik, E. Bennett, A. E. Tron, J.A. Gasser, A. Lau, S.P. Gygi, J. W. Harper, J.A. DeCaprio, A. Toker, W. Wei Rictor forms a complex with cullin-1 to promote SGK1 ubiquitination and destruction Mol. Cell., 39 (2010), 797-808.

[14] T. Xu, D. Sun, Y. Chen, L. Ouyang, Targeting mTOR for fighting diseases: A revisited review of mTOR inhibitors, European Journal of Medicinal Chemistry, 199, (2020), 112391, https://doi.org/10.1016/j.ejmech.2020.112391.

[15] QB. She, SK. Gruvberger-Saal, M. Maurer. Integrated molecular pathway analysis informs a synergistic combination therapy targeting PTEN/PI3K and EGFR pathways for basal-like breast cancer. BMC Cancer 16(1), (2016), 587.

[16] P. Shannon, A. Markiel, O. Ozier, NS. Baliga, JT. Wang, D. Ramage, N. Amin, B. Schwikowski, T. Ideker. Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res, 13:11 (2498-504). 2003 Nov. PubMed ID: 14597658.

[17] SP. Borgatti. "Centrality and Network Flow". *Social Networks*. (2005), **27**: 55–71. doi:10.1016/j.socnet.2004.11.008

[18] M.E.J. Newman. Networks: An Introduction. (2010) Oxford, UK: Oxford University Press.

[19] A. R. Mashaghi, A. Ramezanpour, and V. Karimipour. Investigation of a protein complex network. The European Physical Journal B-Condensed Matter and Complex Systems 41.1 (2004): 113-121.

[20] G. Scardoni, and C. Laudanna. Centralities based analysis of complex networks. New Frontiers in Graph Theory (2012): 323-348.

[21] N.F.A. Christian, N.M. Uriel, P.H. Heidi, P. Rhitankar, P. George, J. Lisi, L. Patrick, I. Rivalta, H. Junming, S.B Victor. Eigenvector centrality for characterization of protein allosteric pathways. Proceedings of the National Academy of Sciences. **115** (52) (2018), E12201– E12208. doi:10.1073/pnas.1810452115.

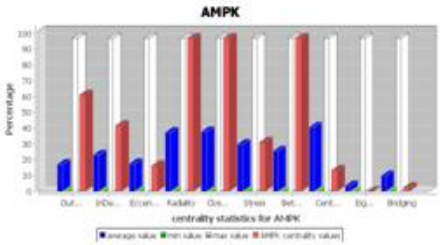
Appendix

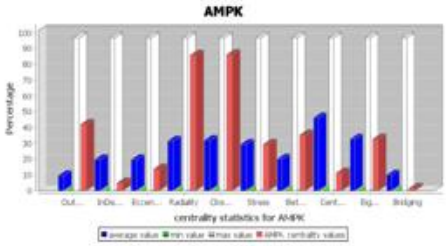
General characteristics of both networks generated by Cytoscape

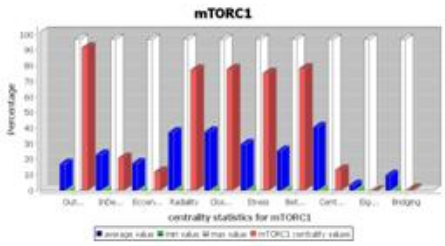
AMPK Signaling (directed)	
Summary Statistics	
Number of nodes	28
Number of edges	94
Avg. number of neighbors	2.571
Network diameter	13
Network radius	1
Characteristic path length	4.187
Clustering coefficient	0.043
Network density	0.050
Connected components	1
Multi-edge node pairs	15
Number of self-loops	8
Analysis time (sec)	0.156

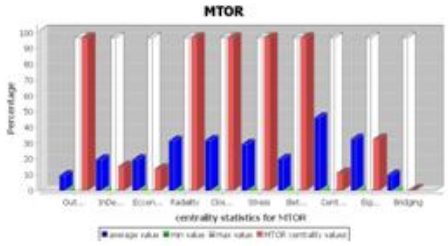
MTOR Signaling (directed)	
Summary Statistics	
Number of nodes	34
Number of edges	131
Avg. number of neighbors	2.235
Network diameter	14
Network radius	1
Characteristic path length	5.356
Clustering coefficient	0.000
Network density	0.035
Connected components	1
Multi-edge node pairs	21
Number of self-loops	8
Analysis time (sec)	0.083

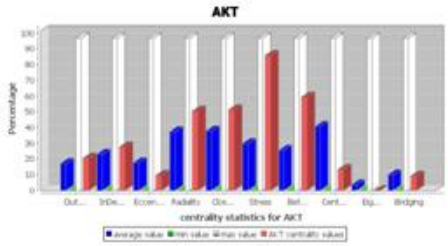
The centralities for the most important nodes, in both networks, are given below. The left side corresponds to the nodes' analysis AMPK-signaling pathway network whereas the right side corresponds to the nodes' analysis in the mTOR-signaling pathway.

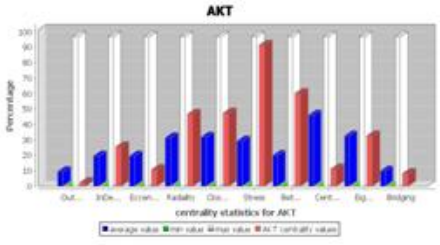


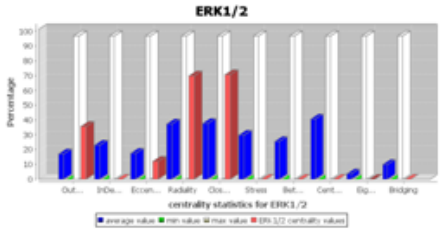


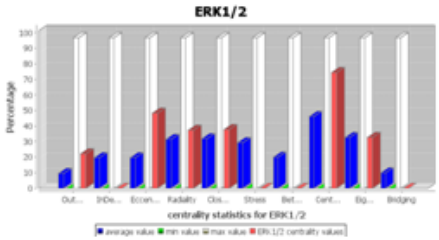


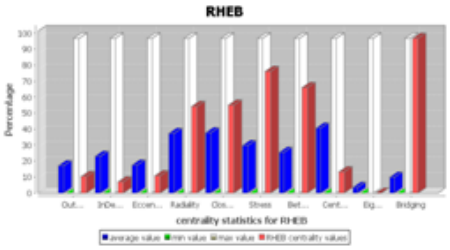


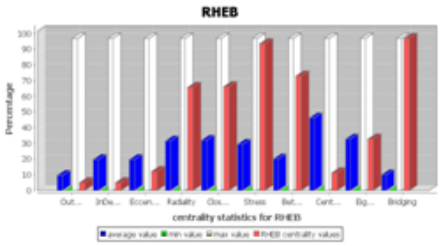


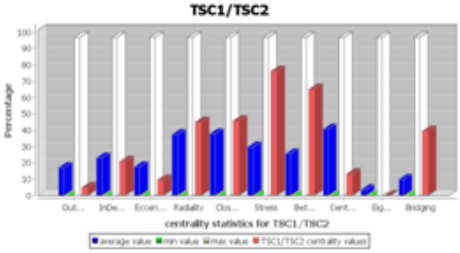


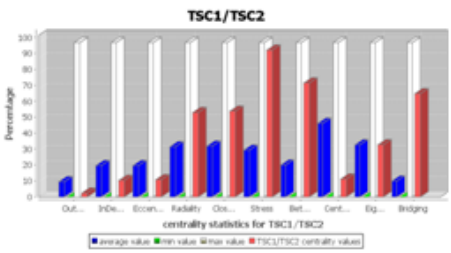

ERK1/2
Percentage
centrality statistics for ERK1/2

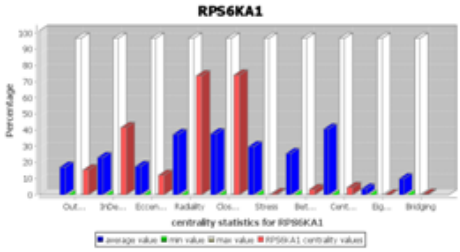

ERK1/2
Percentage
centrality statistics for ERK1/2

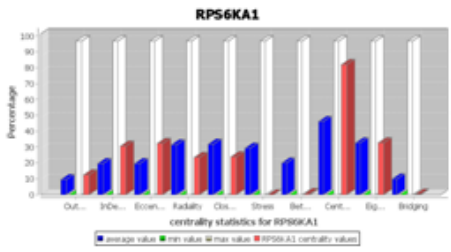

RHEB
Percentage
centrality statistics for RHEB


RHEB
Percentage
centrality statistics for RHEB


TSC1/TSC2
Percentage
centrality statistics for TSC1/TSC2


TSC1/TSC2
Percentage
centrality statistics for TSC1/TSC2


RPS6KA1
Percentage
centrality statistics for RPS6KA1


RPS6KA1
Percentage
centrality statistics for RPS6KA1