RESEARCH ARTICLE

Protein-Protein Interaction Network Analysis for a Biomarker Panel Related to Human Esophageal Adenocarcinoma

Majid Rezaei-Tavirani[1], Sina Rezaei-Tavirani[2], Vahid Mansouri[2], Mohammad Rostami-Nejad[3], Mostafa Rezaei-Tavirani[2]*

Abstract

Background: Esophageal adenocarcinoma (EAC) is one of the mostlethal cancers in the world with a very poor prognosis. Identification of molecular diagnostic methods is an important goal. Since protein-protein interaction (PPI) network analysis is a suitable method for molecular assessment, in the present research a PPI network related to EAC was targeted. **Material and Method:** Cytoscape software and its applications including STRING DB, Cluster ONE and ClueGO were applied to analyze the PPI network. **Result:** Among 182 EAC-related proteins which were identified, 129 were included in a main connected component. Proteins based on centrality analysis of characteristics such as degree, betweenness, closeness and stress were screened and key nodes were introduced. Two clusters were determined of which only one was significant statistically. Gene ontology revealed 50 terms in three groups associated with EAC. **Conclusion:** The findings indicate nine crucial proteins could form a candidate biomarker panel for EAC. Furthermore, an important cluster with 27 proteins related to the disease was identified. Gene ontology analysis of this cluster showed main related terms to closely correspond with those for colorectal cancer.

Keywords: Esophageal adenocarcinoma (EAC)- protein-protein interaction (PPI) network analysis- cluster analysis

Introduction

Esophageal adenocarcinoma is one of the most common malignancies in the world with a very poor prognosis. EAC ranked as the sixth leading cause of cancer death in the world (Parkin et al., 2001). In 2012, it is estimated that 455,800 esophageal cancer cases and 400,200 deaths are occurred in the word. Based on selected registries, the highest esophageal cancer incidence rates are in Malawi, South Africa, and Iran (Torre et al., 2016). The incidence of esophageal adenocarcinoma has been increased dramatically over the past 25 years (Brown and Devesa, 2002) and it perhaps represents a true rise in disease burden (Pohl and Welch, 2005). It has been reported the rapid increament in the incidence of EAC among white men (Brown et al., 2008). EAC is the sixth most common cancer among males and ninth most common cancer among females globally (Guo and Jiang, 2009). Henan province of northern china has the highest incidence and mortality of EAC in the world (Ke, 2002). The majority of EAC patients present with advanced metastatic disease upon diagnosis (Rubenstein and Shaheen, 2015). It is important to understand the molecular mechanism in the tumor invasion process to assist finding new biomarkers for early diagnosis and

prognostic evaluation (Fu et al., 2007). There has been a lack of systematic analysis of the molecular markers that characterize protein of EAC. Attempts to clarify the mechanism underlying esophageal cancer by searching for novel proteins and proteins interactions are attractive recently. Undrestanding a network containing proteins, and generalized an existing computational methods were used to identify related disease proteins simultaneously (Gao et al., 2015). PPI network analysis as an attracting method is used for molecular assesment of many diseases (Shannon et al., 2003; Massoud and Gambhir, 2007). In this method, many proteins implicated to a certain disease reterived and atribute in an interactive structure (PPI network) (Liu et al., 2014; Safari-Alighiarloo et al., 2014; Abbaszadeh et al., 2017). Topological analysis of network provide useful information about roles of the initial proteins (nodes) in the nework (Safari-Alighiarloo et al., 2014; Moreau and Tranchevent, 2012; Safari-Alighiarloo et al., 2016). Crucial nodes of network are the important proteins related to the studied disease (Safari-Alighiarloo et al., 2016; Zamanian-Azodi et al., 2015). These key proteins interact with the large numbers of the nodes of network and considerable types of their connections are as shortest lengh paths so they are called hub-bottleneck proteins (Goh et al., 2007; Safaei et al.,

[1]*Faculty of Medicine, Iran University of Medical Sciences,* [2]*Proteomics Research Center,* [3] *Gastroenterology and Liver Diseases Research Center, Research Institute for Gastroenterology and Liver Diseases, Shahid Beheshti University of Medical Sciences, Tehran, Iran. *For Correspondence: Tavirany@yahoo.com*

2016). Gene ontology analysis can be used to determine underlying biological processes, molecular function, cellular component and biochemical pathways relative to the proteins. These findings can be used in interpretation of played roles of the proteins in onset and development of any diseases (Goh et al., 2007; Barabási et al., 2011; Özgür et al, 2008; Safaei et al., 2016). In this study, we aimed to clarify protein-protein interactions of EAC by network analysis. It can lead to purpose a panel of informative biomarkers specific to esophageal adenocarcinoma.

Materials and Methods

Since the significant proteins related to a certain disease are selected and collected in the string data base and this data bank is up dated periodically, a number of 200 possible proteins correlated to esophageal adenocarcinoma were retrived from disease query of STRING DB V 10.5 (http://string-db.org/) (Coates, 2015). This database is one of the Cytoscape software 3.4.0 applications (Tian et al., 2017) that provides interaction information from three different panels including disease query, protein query, and PubMed query. The strenght of protein interactions can be fitted for the network construction (Szklarczyk et al., 2016). Here, it is set as the 0.4, the default option. The proteins were analyzed via the undirected edges methods by Cytoscape software. The main connected component of PPI network was layout by degree values. Power law is a line that correlates the number of the nodes as a function of degree values. The suitable fitted power law is corresponded to scale free network (Barabási and Bonabeau, 2003). The constructed network was scale free network. The top 10% of the nodes based on degree value were selected as hub nodes. Degree, betweenness, closeness, stress distributions were provided for more screening of the introduced hub nodes due to identify potent hub genes. However grouping of the hub nodes is leaded to detemine super hub nodes and the other potent

hub genes (the top potent hub nodes are introduced as super hub genes). The main connected component (the network) was analyzed by Cluster ONE V.1.0. (clustering with overlapping neighborhood expansion), the other application of Cytoscape software (Saito et al., 2012) that found one significant cluster. In fact, this plug-in derives densely interconnected reigons in the PPI network. Presence of at least 10 proteins in the cluster and P-value less than 0.05 were used to identify significant cluster.

Furthermore, the obtained cluster as a sub-network was analyzed and the contributing nodes were classified regarding degree values. The top hub proteins of cluster and the main connected component were compared and the major candidated biomarker panel was introduced. The connections of the proteins of panel were illustrited. The elements of the cluster were assested by ClueGO (Bindea et al., 2009) for enrichment analysis. The statistical criteria for this examination (pathway analysis) are as follow: Kappa score (term/pathway connectivity) = 0.2. Gene per term was set to 5 and the percentage of atributing proteins was set to 3. P-Value correction method was Bonferroni step down (Bindea et al., 2009; Bindea et al., 2013). P-value less than 0.05 was considered as significant parameter in all steps of analysis.

Results

A number of 200 possible related proteins for esophageal adenocarcinoma were searched via STRING database but only 182 proteins were found in database and included in network study. The protein-protein interaction network of disease was constructed with 182 proteins. The number of 50 nodes were isolated and one component of three connected nodes were formed. The main connected component including 129 nodes and 562 edges was constructed (see Figure 1). As it shown in Figure 1 each node is distinguishable from the other nodes based on degree value. The network was analyzed and the nodes

Table 1. The Top 10% of the Nodes (Including 13 Proteins) Based on Degree Value were Selected as Hub Nodes. The top eight genes are introduced as related biomarkers to EAC. D, BC, CC, S, DS and ASPL correspond to degree, betweenness centrality, closeness centrality, stress, disease score and average shortest paths length respectively.

| R	name	Description	D	BC	CC	S	DS	ASPL
1	TP53	tumor protein p53	58	0.26	0.6	15192	1.83	1.67
2	EGFR	epidermal growth factor receptor	45	0.12	0.55	9250	1.21	1.83
3	AKT1	v-akt murine thymoma viral oncogene homolog 1	43	0.1	0.53	7606	0.84	1.88
4	ERBB2	v-erb-b2 erythroblastic leukemia viral oncogene homolog 2, neuro/glioblastoma derived oncogene homolog (avian)	39	0.06	0.52	5894	1.55	1.91
5	MYC	v-myc myelocytomatosis viral oncogene homolog (avian)	35	0.08	0.51	5902	1.27	1.95
6	CCND1	cyclin D1	35	0.06	0.52	5630	1.14	1.92
7	CTNNB1	catenin (cadherin-associated protein), beta 1, 88kDa	34	0.07	0.53	5986	1.04	1.89
8	CDH1	cadherin 1, type 1, E-cadherin (epithelial)	29	0.06	0.51	4596	1.04	1.96
9	BCL2L1	BCL2-like 1	23	0.03	0.47	2464	0.69	2.11
10	WWOX	WW domain containing oxidoreductase	22	0.06	0.46	4504	1.32	2.16
11	FOXM1	forkhead box M1	21	0.05	0.46	4318	0.78	2.16
12	CDKN2A	cyclin-dependent kinase inhibitor 2A	21	0.01	0.47	1082	1.37	2.12
13	RAC1	ras-related C3 botulinum toxin substrate 1 (rho family, small GTP binding protein Rac1)	18	0.04	0.46	2826	0.71	2.17
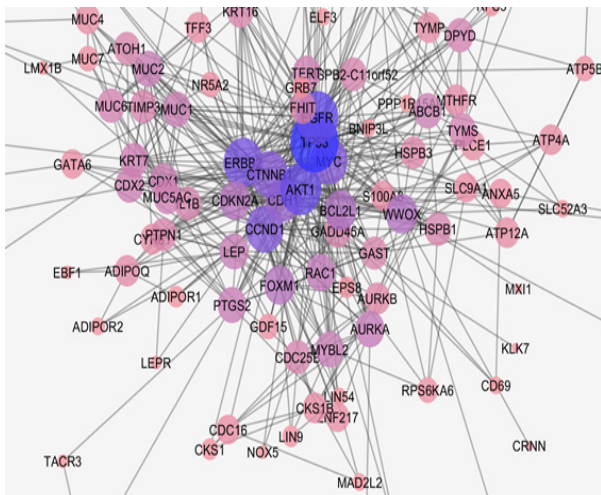
Figure 1. Main Connected Component of protein-protein interaction network of esophageal adenocarcinoma. The component includes 129 nodes and 562 edges. The network is layout by degree value of the nodes. The bigger size of node corresponds to higher value of degree. The color change from orange to blue indicates increment of degree value.
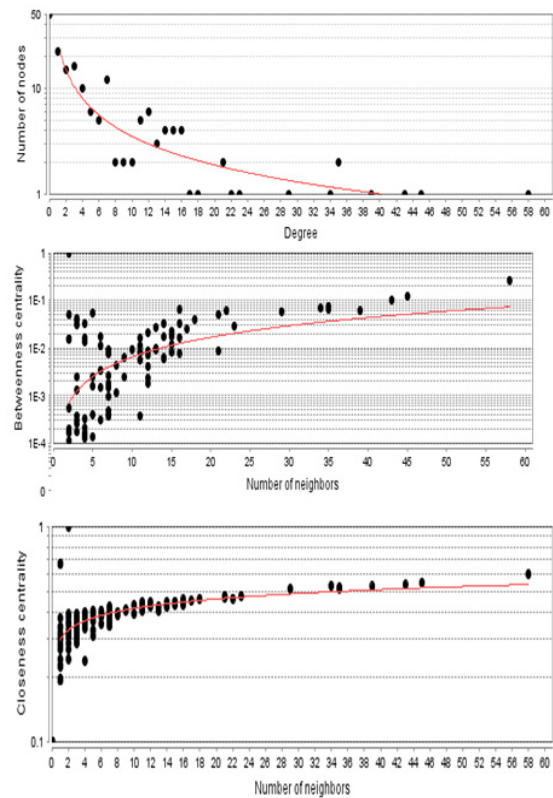

Figure 2. up) Distribution of the Nodes versus Degree, middle) Betweenness Centrality versus Number of Neighbors for the Nodes, and down) Closeness Centrality versus Number of Neighbors for the Nodes are Illustrated. The power law is drowning in red color.

were sorted by degree value. The top 10% of the nodes (including 13 proteins) were selected as hub nodes. The characteristic properties of the hub nodes based on analysis of the network are tabulated in Table 1. These properties are including degree, betweenness, closeness, stress, disease score and average shortest paths length (ASPL). Distribution of the nodes centralities is a useful tool for determination of the potent key nodes and also property of the network (Flórez et al., 2010). This distriution for degree, betweenness and closeness is shown in the Figure 2. These informative graphs are used to determine suitable cut off to identify the crucial nodes (Yu et al., 2007; Aittokallio and Schwikowski, 2006). Grouping of found data is useful tool to screen and classify them. The classified 13 hub nodes based on degree value in four groups is shown in Figure 3. Since clustering provides important information relative to the some crucial parts of a network (Berkhin, 2006), two clusters restricted by presence at least 10 nodes were identified, however only one of them is valided by statistical test. The sub-network related to this cluster is

shown in Figure 4. Distribution of the nodes of cluster-1 sub-network versus the ranges of degree value is shown in Figuer 5. Based on this distribution, the top 10 nodes of this sub-network were selected as the hub nodes of the analyzed cluster (see Tabel 2). As it is depicted in the Table 3, there are nine common genes between Tables 1 and 2. The closed relationship between the nine key genes is presented in the Figure 6. Gene ontology is a powerful method for assessment of biological pathways, molecular function, biological processes and cellular component related to the set of proteins (Thomas et al., 2003; Vastrik et al., 2007). Since cluster-1 includes all key proteins of the main network, gene ontology analysis of 27 proteins

Table 2. Top 10 hub nodes of the Cluster-1. The Properties of the Nodes are Determined

R	Name	description	D	BC	CC	S	DS
1	TP53	tumor protein p53	24	0.08	0.93	260	1.83
2	EGFR	epidermal growth factor receptor	24	0.1	0.93	276	1.21
3	MYC	v-myc myelocytomatosis viral oncogene homolog (avian)	23	0.07	0.9	230	1.27
4	AKT1	v-akt murine thymoma viral oncogene homolog 1	23	0.05	0.9	218	0.84
5	ERBB2	v-erb-b2 erythroblastic leukemia viral oncogene homolog 2, neuro/ glioblastoma derived oncogene homolog (avian)	23	0.1	0.9	252	1.55
6	CCND1	cyclin D1	21	0.04	0.84	166	1.14
7	CTNNB1	catenin (cadherin-associated protein), beta 1, 88kDa	19	0.02	0.79	114	1.04
8	BCL2L1	BCL2-like 1	18	0.03	0.76	100	0.69
9	CDKN2A	cyclin-dependent kinase inhibitor 2A	17	0.01	0.74	72	1.37
10	CDH1	cadherin 1, type 1, E-cadherin (epithelial)	16	0.01	0.72	52	1.04

Table 3. The Top Nodes in Table 1 that are Presented in the Table 2 are shown

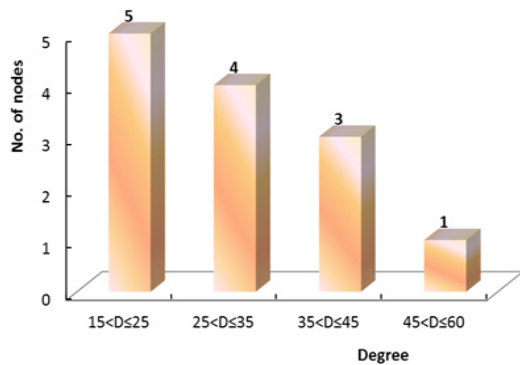| Gene	TP53	EGFR	AKT1	ERBB2	MYC	CCND1	CTNNB1	CDH1	BCL2L1
Row in Table 1	1	2	3	4	5	6	7	8	9
Row in Table 2	1	2	4	5	3	6	7	10	8

Figure 3. Classification of 13 Hub Nodes Based on Degree Value
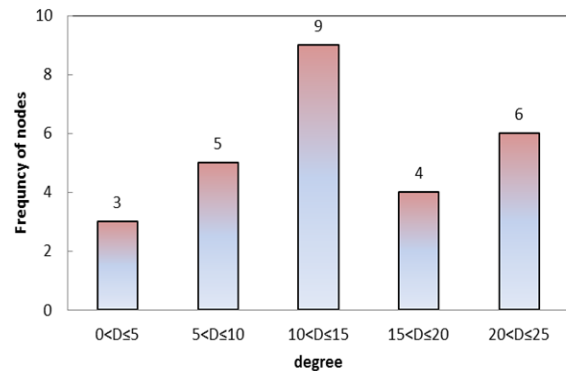
Figure 5. Distribution of the Nodes of Cluster-1 Sub-Network versus the Ranges of Degree Value is Presented
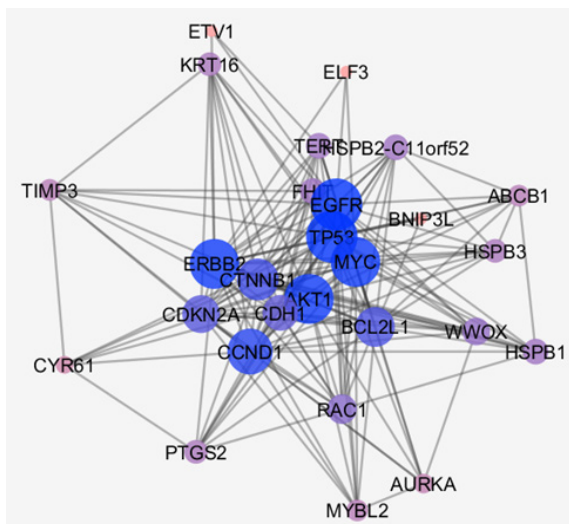

Figure 4. The Sub- Network Corresponds to the Significant Cluster. The sub-network includes 27 nodes and 179 edges. This cluster was determined by cluster ONE application of Cystoscape software. The cluster identification was restricted by presence of at least 10 nodes in cluster (p<0.001).
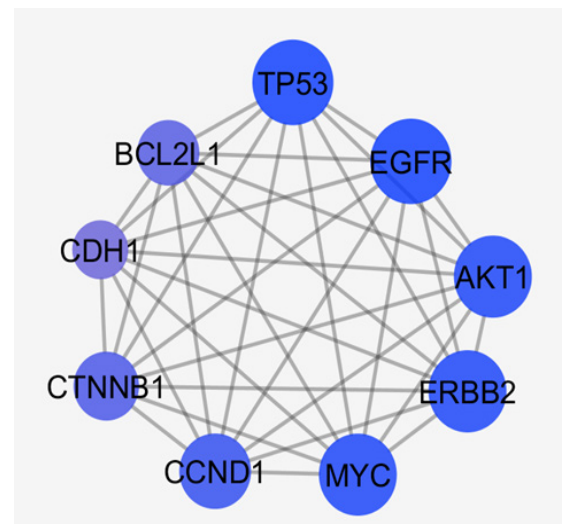

Figure 6. Final Candidate Biomarker Panel Related to Esophageal Adenocarcinoma Including Proteins and Their Interactions. The size and color of the nodes are layout by degree value. The bigger size and highlighted color is corresponded to higher value of degree. All of the nodes are connected to each other.

of this cluster is performed by ClueGO application of Cytoscape. The results are shown in Figures 7-8.

Discussion

Protein molecules play a major role in different diseases and stress condition that their expression contributes to the onset and advances of diseases (Lamb et al., 2006; Zamanian–Azodi et al., 2016). These related proteins to a certain type of disease or condition are present in a whole interacting system (PPI), that worth further analysis for interpreting centrality roles (Fraser et al., 2002; Khayyer et al., 2017). In a way that, the associated proteins are screened and ranked in this manner and finally among many initial proteins a few proteins will be presented as highlighted proteins (Rual et al.,

2005). These key proteins can be potentially differential biomarkers. In this investigation, 200 proteins related to esophagus adenocarcinoma from STRING DB were retrieved; however, only 182 proteins were extracted. At last, the constructed PPI network of EAC was by 182 proteins. The network contains 50 isolated proteins, a triple connected nodes and a main connected component. The main component contains 129 nodes and 562 edges (see Figure 1). In this component (network), the nodes are distincted based on the role of the nodes in the network. This role is depended to the central properties of the nodes in the network. Degree, betweenness, closeness and stress are four important central parameters in the network analysis.

The network was analyzed and central parameters of the nodes were determined. These parameters were
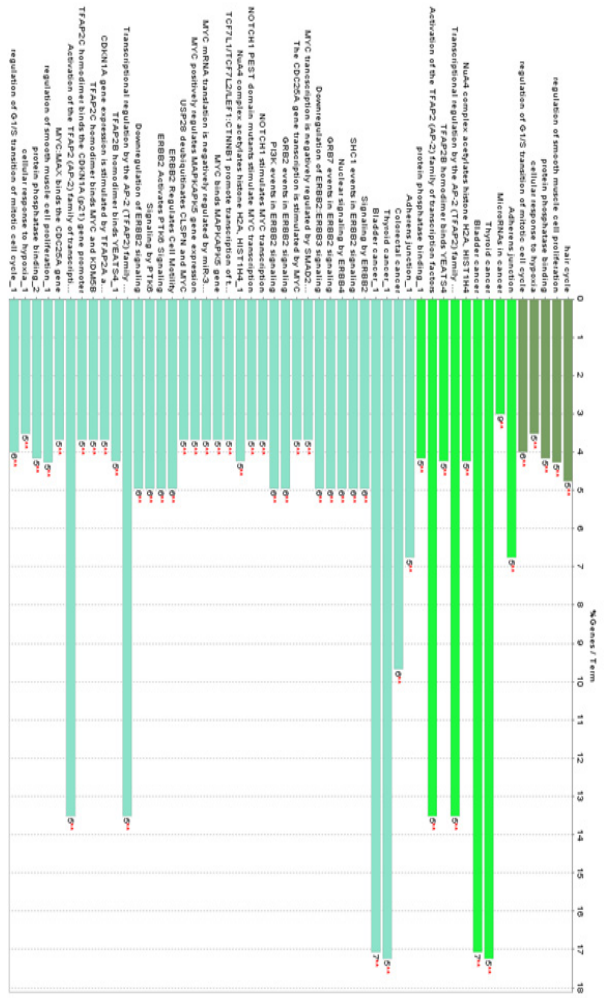
Figure 7. Gene Ontology by Using ClueGO for 27 Proteins of Cluster-1 is Presented. The selected terms were limited to the terms that at least were included 5 initial proteins and 3% Proteins/Term. .Kappa score (Term/Pathway connectivity) was 0.2. The bars that are labeled by numbers corresponded to the attributed proteins in terms. The terms are categorized in three groups (the groups are shown in the three colors).

sorted based on degree value and the top10% of the nodes (including 13 nodes) were selected as hub nodes and tabulated in Table 1. Typically, limited number of proteins can be considered as a suggestive panel for a specific disease. Therefore, it is tried to restrict the number of the hub nodes. As it is shown in the Figure 2, there are several methods to assign a cutoff for identification of the important nodes. In some cases mean +2 standard deviation or top 5% of the nodes are selected as important nodes and also the top nodes that deviated from power law are introduced as the key nodes (Safari-Alighiarloo et al., 2016; Zamanian-Azodi et al., 2015; Safari-Alighiarloo et al., 2017). Based on distribution of degree value the vital hub nodes can be specified. The power law of this distribution implies 40 is a suitable cutoff (see the Figure 2). Similarly, betweenness distribution can be used to define bottleneck nodes . The betweenness value (0.02) is a critical value in the Figure 2. Closeness distribution indicates that 0.5 is the starting point of deviation from power law (see Figure 2). As it was described the three
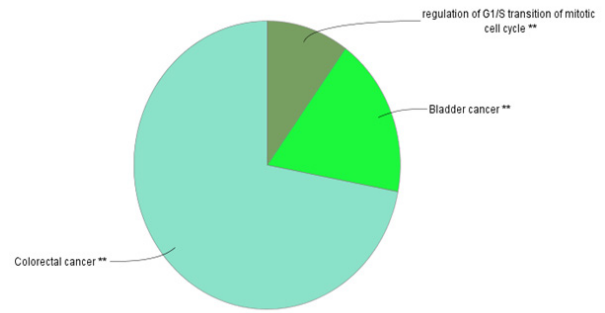

Figure 8. The Identified Terms in Figure 7, are Classified in Three Groups. Frequencies of the member of colorectal cancer, bladder cancer and regulation of G1/S transition of mitotic cell cycle groups are 72%, 18.5% and 9.5% respectively.

proteins including TP53, EGFR and AKT1 are the super-hub proteins. The all hub nodes except CDKN2A are bottleneck nodes. Closeness centrality analysis showed that the top eight nodes are essential nodes. Since degree distribution (Figure 2) provided the limited information about the hub nodes, the degree values were classified in four groups (see Figure 3). The top three groups including eight nodes correspond to closeness distribution indicate that the top eight nodes in the Table 1 are the potent hub proteins. As it is shown in the Table 1, the first eight nodes are characterized by the higher betweenness values in compare to the other nodes. It can be concluded that, the first eight nodes in Table 1 including TP53, EGFR, AKT1, ERBB2, MYC, CCND1, CTNNB1 and BCL2L1, are a suitable candidate biomarker panel related to esophageal adenocarcinoma disease; yet, more analysis is needed to support this claim.

A dense part of a network including some nodes that are closely connected together is called cluster (Milenkovic and Przulj, 2008). The elements of a cluster participate in a closed biological function and processes (Ashburner et al., 2000). The main cluster of the network (Figure 4) provided novel information. Analysis of this sub-network leads to introduce some important proteins. Based on degree value of the 27 nodes of cluster network, it is possible that the proteins to be categorized in five groups (see Figure 5). As it is shown in the Figure 5, the top two groups include 10 hub nodes of the cluster. The central parameters of these 10 hub nodes of cluster network are tabulated in Table 2. Comparison of Tables 1 and 2 is a useful tool for better understanding of the role of the agents of Table 1. In the first glance, it is depicted that the eight key proteins in Table 1, are presented in Table 2 and the more information is shown in Table 3. The second point is presence of the ninth node of Table 1 in the eighth row of Table 2. It seems that considering the role of BCL2L1 in cluster-1, implies that it can be added to the elements of the introduced biomarker panel. In a transcriptome study, two genes including SPARC and SPP1 were introduced as prognostic biomarkers for EAC (Kim et al., 2010), but these proteins are not included in our panel. However there are documents that indicate to the regulatory role of SPARC on TP53 inactivation via AKT mediated process (Fenouille et al.,

2011). In the other report regulation of SPP1-PI3K-AKT signaling pathway in cancer is emphasized (Yang et al., 2012). The assessed patients are mainly (above 50%) men and are in stages II and III. These conditions may affect the expression of the studied proteins. Expression change of EGFR and ERBB2 in EAC patients is reported by (Al-Kasspooles et al., 1993). The prominent role of TP53 in cancer and EAC is discussed in several studies (Wang et al., 1993; Verma et al., 2016). Relationship between MYC, AKT1, CTNNB1, CDH, BCL2L1, and CCND1 proteins is investigated and emphasized (Carneiro et al., 2008; Essack, 2009). The main role of TP53 in early event of esophagus cancer is confirmed by (Gao et al., 1994). Amplification of c-myc, EGFR and the 20q12 loci related to early event of esophagus cancer that are detectable by using FISH on brush cytology specimens is reported by previously (Lu et al., 1988). The role of AKT signaling pathway in chemoresistance induction in esophageal cancers is reported and discussed in details (Hamano et al., 2011; Hildebrandt et al., (2009). It seems that a biomarker panel including TP53, EGFR, and AKT1 can be used as diagnostic agent in early stage of esophagus cancer. It is reasonable idea that the amount of expression change of these genes in correlation with development of disease be increased. Genomic study of oesophageal carcinoma showed expression change of CCND1 in squamous cell carcinomas, whereas ERBB2 amplified in adenocarcinomas (Network, (2017). Comparative analysis of the Tables 1 and 2, expresses that the identified cluster is the core part of the network of esophageal adenocarcinoma. Closed connections between the nine key proteins are shown in Figure 6. Each protein is connected to the other eight proteins. Since the all-key nodes are common between the main network and cluster-1, gene ontology of the elements of cluste-1 can provide useful information about related biological processes to EAC. Gene ontology finding indicated that, the nodes of cluster-1 play fundamental roles in esophageal adenocarcinoma malignancy. The identified terms (see Figures 7-8) are related closely to the cancer especially colorectal cancers. It is a potent confirmation for the findings of the analyzed network. Therefore, it is a novel combination and ranked gene set that is introduced as biomarker panel of EAC.

In conclusion, it can be concluded that, there is a cluster including 27 proteins related to esophageal adenocarcinoma disease. By consider the elements of cluster-1 and the key proteins in the PPI network of this disease, a panel including TP53, EGFR, AKT1, ERBB2, MYC, CCND1, CTNNB1, CDH1, AND BCL2L1 was introduced as a candidate biomarker panel for esophageal adenocarcinoma. Still, more investigation in the field is required for validation of these findings.

Acknowledgements

Conflict of interest
There is no conflict of interest.

References

Abbaszadeh H-A, Peyvandi AA, Sadeghi Y, et al (2017). Er: YAG laser and cyclosporin a effect on cell cycle regulation of human gingival fibroblast cells. *J Laser Med Sci*, **8**, 143-9.

Aittokallio T, Schwikowski B (2006). Graph-based methods for analysing networks in cell biology. *Brief Bioinformatics*, **7**, 243-55.

Al-Kasspooles M, Moore JH, Orringer MB, et al (1993). Amplification and over-expression of the EGFR and erbB-2 genes in human esophageal adenocarcinomas. *Int J Cancer*, **54**, 213-9.

Ashburner M, Ball CA, Blake JA, et al (2000). Gene ontology: tool for the unification of biology. *Nature Genet*, **25**, 25-9.

Barabási A-L, Gulbahce N, Loscalzo J(2011). Network medicine: a network-based approach to human disease. *Nature Rev Genet*, **12**, 56-68.

Barabási BA-L, Bonabeau E (2003). Scale-free networks. *Sci Am*, **288**, 50-9.

Berkhin PA (2006). Survey of clustering data mining techniques. Grouping multidimensional data: Springer, pp 25-71.

Bindea G, Galon J, Mlecnik B (2013). CluePedia cytoscape plugin: pathway insights using integrated experimental and in silico data. *Bioinformatics*, **29**, 661-3.

Bindea G, Mlecnik B, Hackl H, et al (2009). ClueGO: a Cytoscape plug-in to decipher functionally grouped gene ontology and pathway annotation networks. *Bioinformatics*, **25**, 1091-3.

Brown LM, Devesa SS, Chow W-H (2008). Incidence of adenocarcinoma of the esophagus among white Americans by sex, stage, and age. *J Natl Cancer Inst*, **100**, 1184-7.

Brown LM, Devesa SS (2002). Epidemiologic trends in esophageal and gastric cancer in the United States. *Surg Oncol Clin N Am*, **11**, 235-56.

Carneiro A, Isinger A, Karlsson A, et al (2008). Prognostic impact of array-based genomic profiles in esophageal squamous cell cancer. *BMC Cancer*, **8**, 98.

Coates P (2015). p53–the gene that cracked the cancer code. *J Pathol*, **236**, 395

Essack M (2009). Transcription regulation and candidate diagnostic markers of esophageal cancer: University of the Western Cape, pp 9-16.

Fenouille N, Puissant A, Tichet M, et al (2011). SPARC functions as an anti-stress factor by inactivating p53 through Akt-mediated MDM2 phosphorylation to promote melanoma cell survival. *Oncogene*, **30**, 4887-900.

Flórez AF, Park D, Bhak J, et al (2010). Protein network prediction and topological analysis in Leishmania major as a tool for drug target selection. *BMC Bioinformatics*, **11**, 484.

Fraser HB, Hirsh AE, Steinmetz LM, et al (2002). Evolutionary rate in the protein interaction network. *Science*, **296**, 750-2.

Fu L, Qin YR, Xie D, et al (2007). Identification of alpha-actinin 4 and 67 kDa laminin receptor as stage-specific markers in esophageal cancer via proteomic approaches. *Cancer*, **110**, 2672-81.

Gao H, Wang L-D, Zhou Q, et al (1994). p53 tumor suppressor gene mutation in early esophageal precancerous lesions and carcinoma among high-risk populations in Henan, China. *Cancer Res*, **54**, 4342-6.

Gao Y-F, Yuan F, Liu J, et al (2015). Identification of new candidate genes and chemicals related to esophageal cancer using a hybrid interaction network of chemicals and proteins. *PLoS One*, **10**, e0129474.

Goh K-I, Cusick ME, Valle D, et al (2007).The human disease network. *Proc Natl Acad Sci U S A*, **104**, 8685-90.

Guo W, Jiang Y-G (2009). Current gene expression studies in esophageal carcinoma. *Curr Genomics*, **10**, 534-9.

Hamano R, Miyata H, Yamasaki M, et al (2011). Overexpression of miR-200c induces chemoresistance in esophageal cancers mediated through activation of the Akt signaling pathway. *Clin Cancer Res*, **17**, 3029-38.

Hildebrandt MA, Yang H, Hung M-C, et al (2009). Genetic variations in the PI3K/PTEN/AKT/mTOR pathway are associated with clinical outcomes in esophageal cancer patients treated with chemoradiotherapy. *J Clin Oncol*, **27**, 857-71.

Ke L (2002). Mortality and incidence trends from esophagus cancer in selected geographic areas of China circa 1970–90. *Int J Cancer*, **102**, 271-4.

Khayyer N, Azodi MZ, Mansouri V, et al (2017). Oral squamous cell Cancer protein-protein interaction network interpretation in comparison to esophagus adenocarcinoma. *Gastroenterol Hepatol Bed Bench*, **10**, 118-24.

Kim SM, Park Y-Y, Park ES, et al (2010). Prognostic biomarkers for esophageal adenocarcinoma identified by analysis of tumor transcriptome. *PLoS One*, **5**, e15074.

Lamb J, Crawford ED, Peck D, et al (2006). The connectivity map: using gene-expression signatures to connect small molecules, genes, and disease. *Science*, **313**, 1929-35.

Liu R, Wang X, Aihara K, et al (2014). Early diagnosis of complex diseases by molecular biomarkers, network biomarkers, and dynamical network biomarkers. *Med Res Rev*, **34**, 455-78.

Lu SH, Hsieh LL, Luo FC, et al (1988). Amplification of the EGF receptor and c-myc genes in human esophageal cancers. *Int J Cancer*, **42**, 502-5.

Massoud TF, Gambhir SS (2007). Integrating noninvasive molecular imaging into molecular medicine: an evolving paradigm. *Trends Mol Med*, **13**, 183-91.

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, e13984.

Milenkovic T, Przulj N (2008). Uncovering biological network function via graphlet degree signatures. *Cancer Inform*, **6**, 257-73

Moreau Y, Tranchevent L-C (2012). Computational tools for prioritizing candidate genes: boosting disease gene discovery. *Nature Rev Genet*, **13**, 523-36.

Nepusz T, Yu H, Paccanaro A (2012). Detecting overlapping protein complexes in protein-protein interaction networks. *Nat Methods*, **9**, 471-2.

Network CGAR (2017). Integrated genomic characterization of oesophageal carcinoma. *Nature*, **541**, 169-75.

Özgür A, Vu T, Erkan G, et al (2008). Identifying gene-disease associations using centrality on a literature mined gene-interaction network. *Bioinformatics*, **24**, 277-85.

Parkin DM, Bray F, Devesa S. (2001). Cancer burden in the year 2000. The global picture. *Eur J Cancer*, **37**, 4-66.

Pohl H, Welch HG (2005).The role of overdiagnosis and reclassification in the marked increase of esophageal adenocarcinoma incidence. *J Natl Cancer Inst*, **97**, 142-6.

Rual J-F, Venkatesan K, Hao T, et al (2005). Towards a proteome-scale map of the human protein–protein interaction network. *Nature*, **437**, 1173-8.

Rubenstein JH, Shaheen NJ (2015). Epidemiology, diagnosis, and management of esophageal adenocarcinoma. *Gastroenterol*, **149**, 302-17.

Safaei A, Tavirani MR, Oskouei AA, et al (2016). Protein-protein interaction network analysis of cirrhosis liver disease. *Gastroenterol Hepatol Bed Bench*, **9**, 114-23.

Safari-Alighiarloo N, Rezaei-Tavirani M, Taghizadeh M, et al (2016). Network-based analysis of differentially expressed genes in cerebrospinal fluid (CSF) and blood reveals new candidate genes for multiple sclerosis. *Peer J*, **4**, e2775.

Safari-Alighiarloo N, Taghizadeh M, Rezaei-Tavirani M, et al (2014). Protein-protein interaction networks (PPI) and complex diseases. *Gastroenterol Hepatol Bed Bench*, **7**, 17-31.

Safari-Alighiarloo N, Taghizadeh M, Tabatabaei SM, et al (2017). Identification of new key genes for type 1 diabetes through construction and analysis of protein–protein interaction networks based on blood and pancreatic islet transcriptomes. *J Diabetes*, **9**, 764-77

Saito R, Smoot ME, Ono K, et al (2012). A travel guide to Cytoscape plugins. *Nat Methods*, **9**, 1069-76.

Shannon P, Markiel A, Ozier O, et al (2003). Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res*, **13**, 2498-504.

Szklarczyk D, Morris JH, Cook H, et al (2017). The STRING database in 2017: quality-controlled protein–protein association networks, made broadly accessible. *Nucleic Acid Res*, **45**, 362-8.

Thomas PD, Campbell MJ, Kejariwal A, et al (2003). PANTHER: a library of protein families and subfamilies indexed by function. *Genome Res*, **13**, 2129-41.

Tian Y, Huang W, Yang J, et al (2017). Systematic identification of hepatitis E virus ORF2 interactome reveals that TMEM134 engages in ORF2-mediated NF-κB pathway. *Virus Res*, **228**, 102-8.

Torre LA, Siegel RL, Ward EM, et al (2016). Global cancer incidence and mortality rates and trends-an update. *Cancer Epidemiol Biomarkers Prev*, **25**, 16-27.

Vastrik I, D'Eustachio P, Schmidt E, et al (2007). Reactome: a knowledge base of biologic pathways and processes. *Genome Biol*, **8**, R39.

Verma A, Gunasekar S, Goel V, et al (2016). A molecular approach to Glioblastoma Multiforme. *Int J Mol Immuno Oncol*, **1**, 35-44.

Wang K, Li M, Hakonarson H (2010). Analysing biological pathways in genome-wide association studies. *Nature Rev Genet*, **11**,843-54.

Wang L-D, Hong J-Y, Qiu S-L, et al (1993). Accumulation of p53 protein in human esophageal precancerous lesions: a possible early biomarker for carcinogenesis. *Cancer Res*, **53**, 1783-7.

Yang Y-F, Tsia H-Y, Lee C-H, et al (2012). Overexpression of cholesterol metabolism regulatory gene farnesyl diphosphate farnesyltransferase is associated with poor prognosis of non-small cell lung cancer patient and promotes metastasis through feedback regulation of SPP1-PI3K-AKT signaling pathway. *Cancer Res*, **72**, 4612.

Yu H, Kim PM, Sprecher E, et al (2007). The importance of bottlenecks in protein networks: correlation with gene essentiality and expression dynamics. *PLoS Comput Biol*, **3**, e59.

Zali H, Rezaei-Tavirani M, Vafaee R, et al (2013). Gastric cardia adenocarcinoma pathway analysis. *Gastroenterol Hepatol Bed Bench*, **6**, 11-8.

Zamanian-Azodi M, Rezaei-Tavirani M, Rahmati-Rad S, et al (2015). Protein-protein interaction network could reveal the relationship between the breast and colon cancer. *Gastroenterol Hepatol Bed Bench*, **8**, 215-24.

Zamanian-Azodi M, Rezaei-Tavirani M, Rahmati-Rad S, et al (2016). Ethanol and cancer induce similar changes on protein expression pattern of human fibroblast cell. *Iran J Pharm Res*, **15**, 175-84.

Zhu X, Gerstein M, Snyder M (2007). Getting connected: analysis and principles of biological networks. *Genes Dev*, **21**, 1010-24.