Int J Clin Exp Med 2018;11(3):2032-2040
www.ijcem.com /ISSN:1940-5901/IJCEM0060517

Original Article

Gene expression analysis predicts the pathological response of esophageal squamous cell carcinomas to neoadjuvant chemoradiotherapy

Guoqing Yang[1], Qingfang Wang[1], Dongsheng Hu[1], Qingfang Wang[2], Xuejun Yu[1]

Departments of [1]Thoracic, [2]Operating Room, Shouguang People's Hospital, Weifang 262700, China

Received June 29, 2017; Accepted October 10, 2017; Epub March 15, 2018; Published March 30, 2018

Abstract: Purpose: The aim of this study was to research biomarkers for ESCC diagnosis and chemosensitivity related differently expressed genes (DEGs). Methods: Gene microarray data of GSE45670 was obtained and DEGs were screened by Significant analysis of microarray method. series test of cluster (STC) was processed to research the expression trend features among three groups. Based on GO and KEGG database, go functions and KEGG pathways were enriched, gene signal network and gene co-expression network were constructed. Results: A total of 9061 DEGs were screened and mainly grouped in 16 profiles. Among of these profiles, there were 6 significant profiles. These screened DEGs were enriched in various functions, including mitotic cell cycle and transcription. Besides, these genes were also participate in different pathways, such as metabolic pathway and pathways in cancer. Pathway relationship network with 96 nodes and 336 edges were constructed. Thereinto, MAPK signaling pathway, apoptosis, pathways in cancer and cell cycle were hub nodes with degree of 43, 31, 28 and 23, respectively. Gene signal network was constructed with 992 nodes and 3015 edges. The hub nodes included PRKACB, PLCB4, MAPK8 and PIK3R1. PRKACB and MAPK8 were with indirect activation relationships, while PIK3R1 and PLCB4 were with metabolic relationships. Gene co-expression network was also constructed with 456 nodes and 1713 edges. Top 5 nodes were NEGR1, MITF, TNFSF12 and DAAM2. Conclusions: Screened genes including NEGR1, MITF, PRKACB and PLCB4 might be potential key genes for ESCC diagnosis. Importantly, PRKACB and PLCB4 were closely related with chemosensitivity.

Keywords: Gene expression analysis, pathological response, esophageal squamous cell carcinomas, neoadjuvant chemoradiotherapy

Introduction

Esophageal carcinomas is always arising from esophagus with symptoms including difficulty in swallowing, weight loss, enlarged lymph nodes around the collarbone and even vomiting blood [1]. Esophageal squamous cell carcinoma (ESCC) was a main sub-type of esophageal carcinomas, which was common in the developing world [2]. Chemotherapy was an effective treatment for ESCC in a tumor type-dependent manner. Moreover, neoadjuvant chemoradiotherapy was confirmed to improve survival and recommended in the guidelines for ESCC management [3].

The effects of neoadjuvant chemoradiotherapy for ESCC treatment was a controversial issue. Liu et al. showed that clinical complete response, pathologic complete response and neoadjuvant chemoradiotherapy were with significant correlations [4]. Some studies have confirmed that neoadjuvant chemoradiotherapy might provide more survival benefit to ESCC patients than patients with adenocarcinoma [5]. Besides, Huang et al. found that age, smoking history and tumor length were key predictors for pathologic complete response in ECSS patients [6]. Interestingly, pathological complete response to neoadjuvant chemoradiotherapy played important role in postoperative survival rate of ESCC patients, especially in patients of stage of II-III [7]. In addition, the expression level of Rad51 was confirmed to be an effective predictive factor in this process. Besides, Okumura et al. [8] summarized many literatures and selected numbers of biomarkers for response to neoadjuvant chemoradiation therapy, such as cell

division cycle 25B (CDC25B), ribonucleotide reductase regulatory TP53 inducible subunit M2B (p53R2), ERCC excision repair 1, endonuclease non-catalytic subunit (ERCC1) and vascular endothelial growth factor (VEGF).

However, not all ESCC patients showed pathological complete response to neoadjuvant chemoradiotherapy. Understanding the sensitivity of tumor to chemotherapy might contribute to select postoperative chemotherapy scheme. Thereby, differently expressed genes (DEGs) among normal, pathological complete response and not pathological complete response groups were screened in this study. Furthermore, DEGs related functions and pathways were also researched. The research of this study might lay a foundation for ESCC disease prediction and chemosensitivity prediction.

Methods

Accession numbers

Gene microarray data of GSE45670 was deposited in gene expression omnibus (GEO) database by Wen *et al*. [9] on Aug 1, 2014 (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE45670). The platform of this microarray was A-AFFY-44-Affymetrix GeneChip Human Genome U133 Plus 2.0 [HG-U133_Plus_2]. The raw data included 10 normal esophageal epithelium samples, 17 esophageal squamous cell carcinoma (ESCC) patients not pathological complete response to preoperative chemoradiotherapy (CRT) and 11 ESCC patients pathological complete response to preoperative CRT.

Preprocessing of raw data

Firstly, Robust Multi-chip Average (RMA) algorithm was used to calculate expression level of probe sets with 3 steps: background correction, normalization and summarizing [10]. Then, all probes were annotated based on annotation files (*.transcript.csv) of affymetrix official website (http://www.affymetrix.com/support/technical/annotationfilesmain.affx). Afterwards, the value of normalized unscaled standard errors (NUSE) was regarded as the evaluation criteria for the feasibility and reliability of design scheme. The threshold was defined as (1-0.2) < median {NUSE} < (1+0.2). And on this basis, relative log expression (RLE) was applied to assess the change rule of each probe set. Samples met standards of (-0.25) < median {RLE} < (0.25) were included.

Screening of differently expressed genes

Significant analysis of microarray (SAM) method [11] was used to screen DEGs among three groups of samples. The threshold of DEGs was $P < 0.05$.

Series test of cluster for DEGs

In this study, series test of cluster (STC) was processed to research the expression trend features among three groups [12]. Based on grouping and amplitude of gene expression variation, trending models were defined. In addition, relationship between gene expression and trending models were established, and the significance of each model was examined. Multiple comparative tests were applied to control family-wise error rate.

Functional and pathway enrichment analysis

Gene Ontology (GO), which was built by Gene Ontology Consortium, was widely used for functional enrichment of genes and proteins. Pathway analysis showed much biochemical information, including metabolism, signal transmission and cell cycle. In this study, Fisher exact test was used for GO and pathway enrichment as following contingency table:

	Difference Gene	Non-difference Gene	Total
Genes in GO/pathway	n_f	$n - n_f$	n
Genes out of GO/pathway	$N_f - n_f$	$(N - N_f) - (n - n_f)$	$N - n$
Total	N_f	$N - N_f$	N

In addition, *p* value was calculated by formula as follows:

$$p_F(n_f, n, N_f, N) = 2 * p_h(X \leq n_f) = 2 * \sum_{x=1}^{n_f} \frac{\binom{n}{x}\binom{N-n}{Nf-x}}{\binom{N}{N_f}}.$$

Benjamini-Hochberg was used to correct *p* value to FDR, and enrichment score was calculated to evaluate the enrichment level of DEGs in each function or pathway. The computing method was shown as follows:

$$R_e = \frac{n_f/n}{N_f/N}.$$

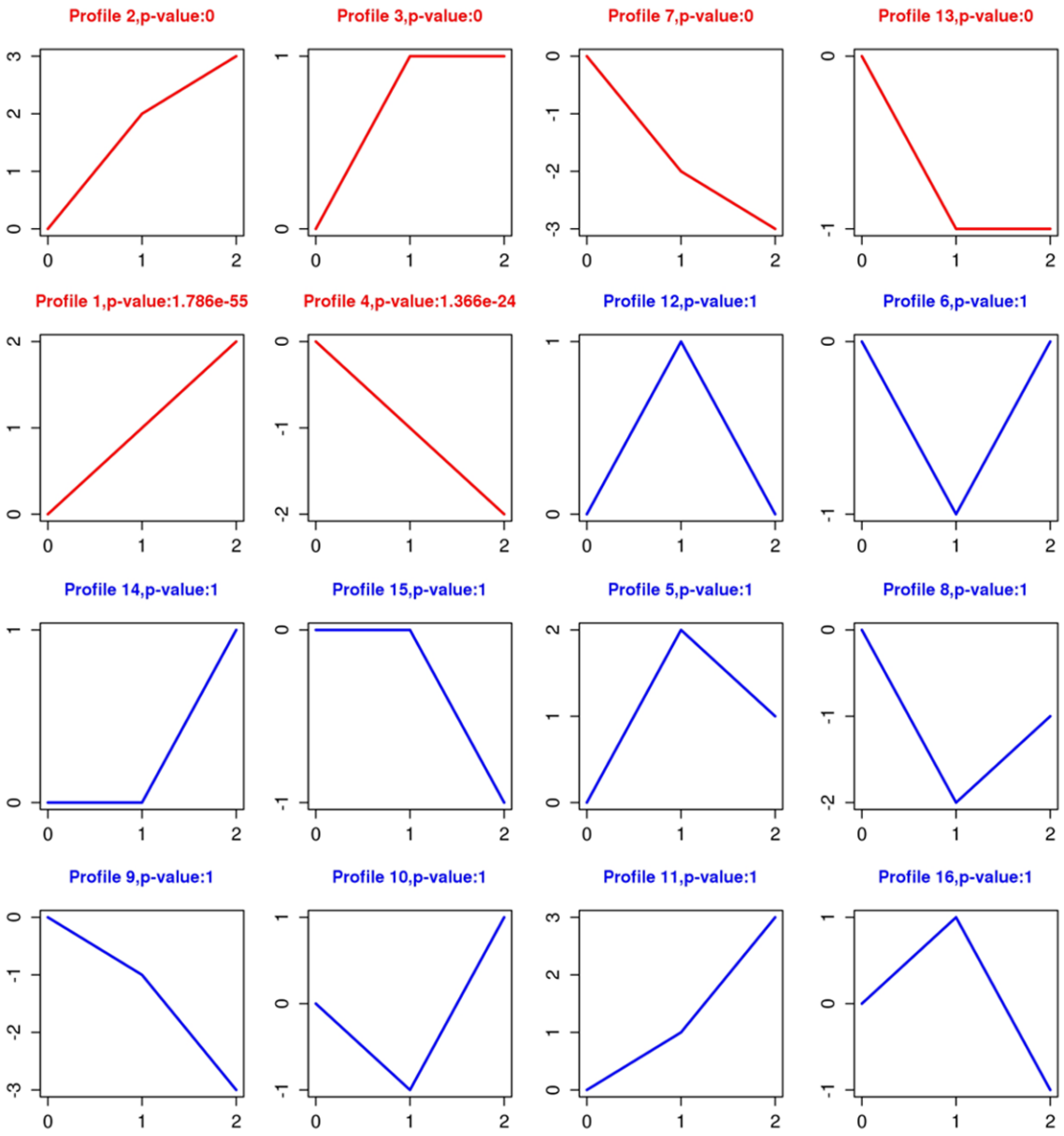

Figure 1. Top 16 series test of clusters: The red cluster was with significance. 0, 1 and 2 in a-axis represent normal esophageal epithelium samples, esophageal squamous cell carcinoma (ESCC) patients not pathological complete response to preoperative chemoradiotherapy (CRT) and ESCC patients pathological complete response to preoperative CRT, respectively.

Enriched functions and pathways with $P < 0.05$ were regarded to be with significance.

Construction of pathway relationship network

Pathway relationship network was constructed based on KEGG database. This network could intuitively reflect the signal conduction relationships among significant pathways. The degree of node i was calculated by formula of, while the betweeness was calculated by formula of. Thereinto, was the number of shortest path from node s to node t, while was the number of shortest path through node i.

Gene signal network and gene co-expression network construction

DEGs in GO terms and KEGG pathways were inserted and numbers of key DEGs were obtained. Then gene co-expression network and gene signal network were constructed based

Table 1. Top 5 GO terms of differently expressed genes

GO ID	GO Name	Diff Gene Counts in GO	Gene Amount in GO	Enrichment Score	*p*-value	FDR
GO:0000278	Mitotic cell cycle	192	363	4.501151895	1.20E-81	6.96E-78
GO:0006351	Transcription, DNA-dependent	477	1827	2.221820128	1.81E-66	5.22E-63
GO:0006915	Apoptotic process	237	654	3.083895568	4.18E-60	7.43E-57
GO:0051301	Cell division	149	295	4.298266287	5.14E-60	7.43E-57
GO:0010467	Gene expression	237	668	3.019263026	3.87E-58	4.48E-55

Table 2. Top 5 KEGG pathways of differently expressed genes

Pathway ID	Pathway Name	Diff Gene Counts in Pathway	Gene Amount in Pathway	Enrichment Score	*p*-value	FDR
1100	Metabolic pathways	304	1189	2.175809126	3.12E-40	8.19E-38
5200	Pathways in cancer	117	327	3.044858915	1.57E-29	2.06E-27
4141	Protein processing in endoplasmic reticulum	75	167	3.821851931	1.51E-26	1.33E-24
4110	Cell cycle	61	124	4.186366196	1.96E-24	1.29E-22
230	Purine metabolism	72	173	3.541730067	4.27E-23	2.24E-21

on these key DEGs. Gene co-expression network were constructed by learning from scale-free property of mass data, and scale-free relationship of network was fitted by correlation coefficient between key DEGs. Correlation coefficient of gene co-expression network should be approximated to 1. Gene signal network was constructed to screen up- and down-stream protein. In this study, up- and down-stream genes were filtrated in KEGG database. These genes were connected by various regulatory relationships, such as binding, phosphorylation, ubiquination and indirect relationships.

Results

DEGs screening

With the threshold of $P < 0.05$, a total of 9061 DEGs were screened in this study, such as angiopoietin-like 1 (ANGPTL1), chordin-like 1 (CHRDL1), chromosome 2 open reading frame 40 (C2orf40) and complement component 7 (C7).

Series test of cluster for DEGs

Screened DEGs were processed by series test of cluster and mainly grouped in 16 profiles. Among of these profiles, there were 6 significant profiles, including profile 2 (p = 0), profile 3 (p = 0), profile 7 (p = 0), profile 13 (p = 0), profile 1 (p = 1.786e-55) and profile 4 (p = 1.366e-24). Genes in profile 2 was increased in ESCC patients not pathological complete response to CRT and increased higher in ESCC patients pathological complete response to preoperative CRT. However, after preoperative CRT treatment, genes in profile 3 were showed increased trending, but with no different between pathological complete response and not pathological complete response groups (**Figure 1**).

Functional and pathway enrichment analysis

These screened DEGs were enriched in various functions, including mitotic cell cycle (FDR = 6.69E-78), transcription, DNA-dependent (FDR = 5.22E-63), apoptotic process (FDR = 7.43E-57) and cell division (FDR = 7.43E-57) (**Table 1**). Besides, these genes were also participate in different pathways, such as metabolic pathway (FDR = 8.19E-38), pathways in cancer (FDR = 2.06E-27), protein processing in endoplasmic reticulum (FDR = 1.33E-24) and cell cycle (FDR = 1.29E-22) (**Table 2**).

Construction of pathway relationship network

Pathway relationship network with 96 nodes and 336 edges were constructed. Thereinto, MAPK signaling pathway, apoptosis, pathways in cancer and cell cycle were hub nodes with degree of 43, 31, 28 and 23, respectively. It was worth mentioning that pathways in cancer (outdegree = 28, indegree = 0) was an up-stream pathway, while cytokine-cytokine receptor interaction (outdegree = 0, indegree = 14) was a down-stream pathway (**Figure 2**).

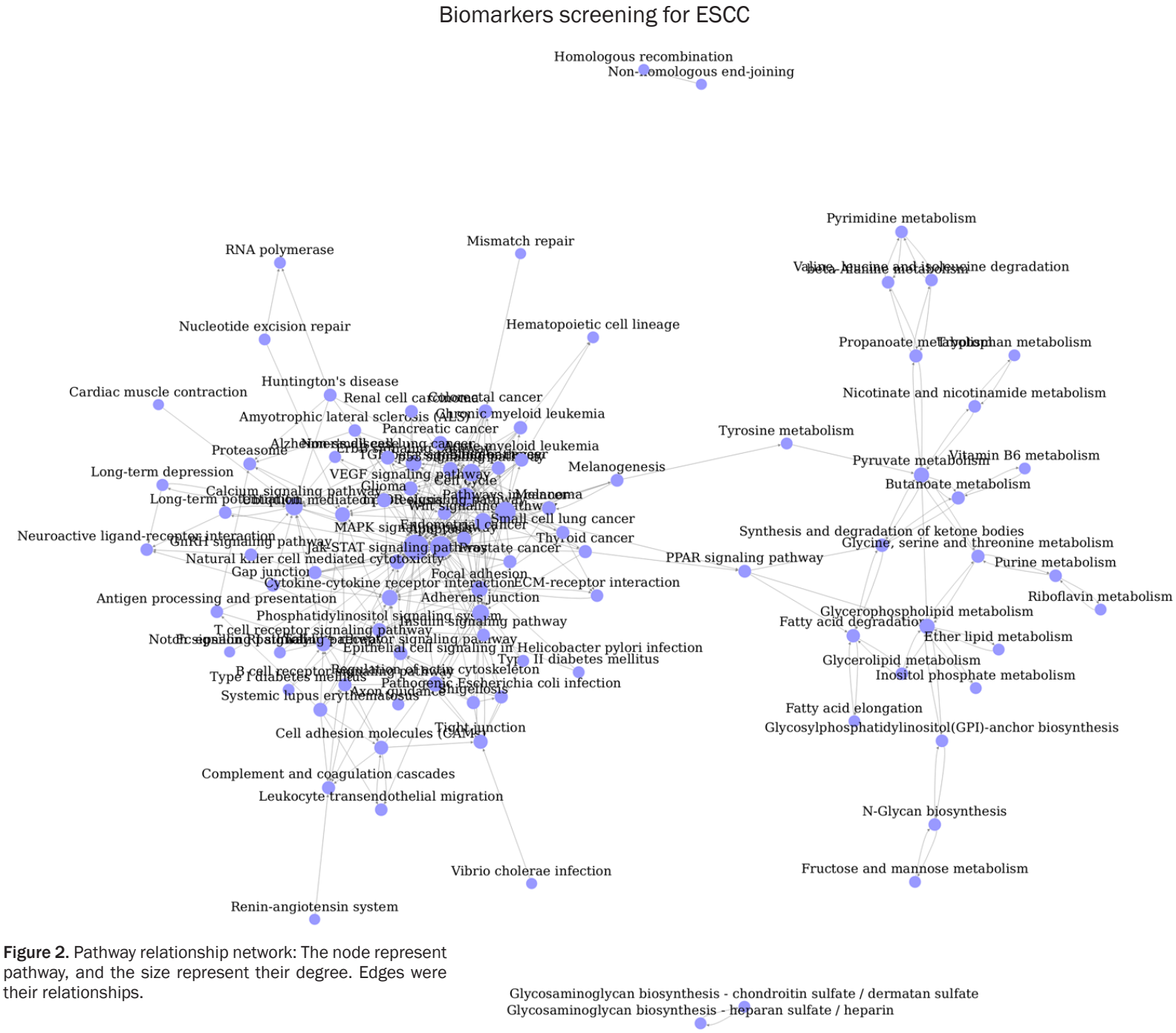

Figure 2. Pathway relationship network: The node represent pathway, and the size represent their degree. Edges were their relationships.

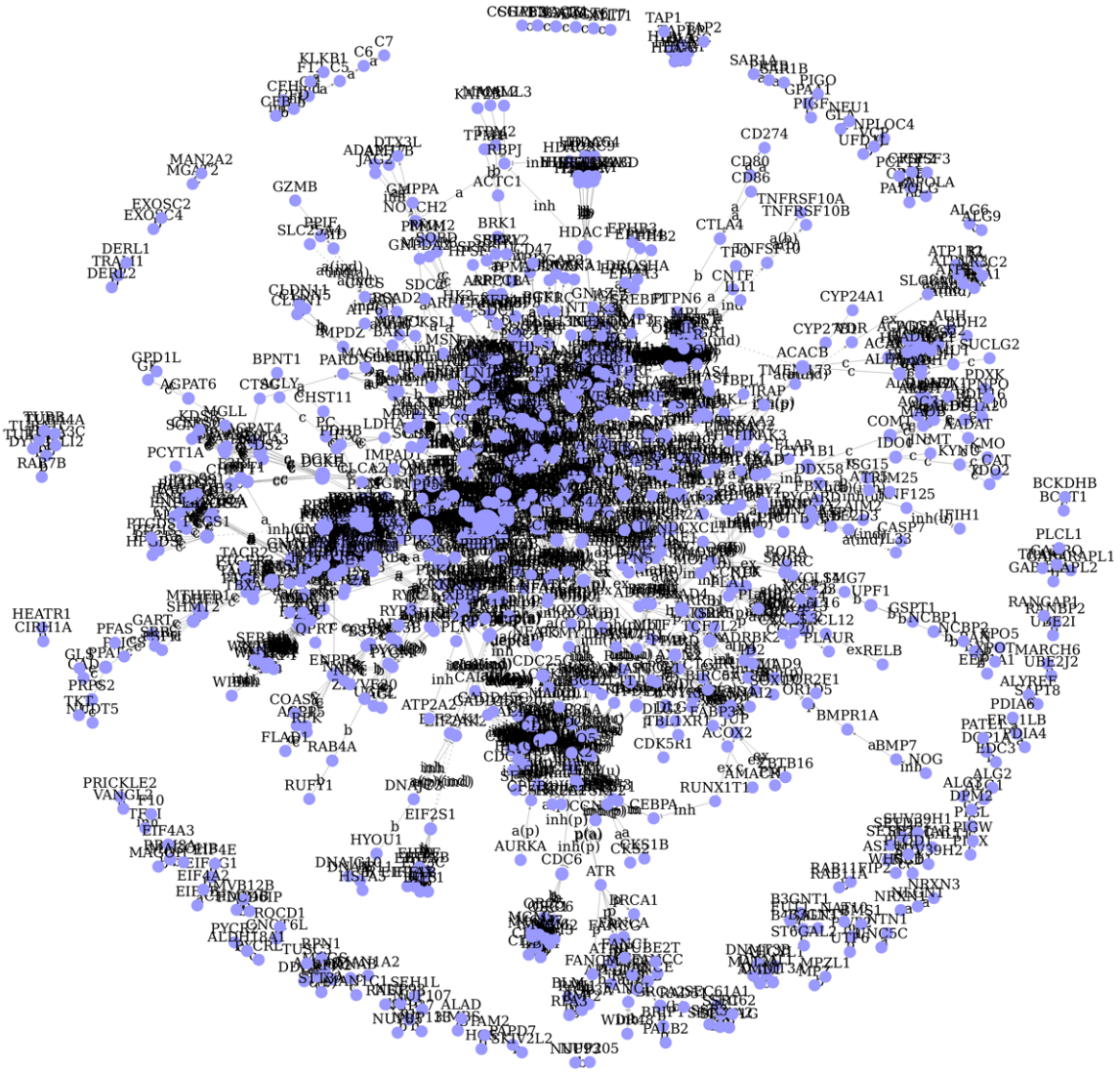

Figure 3. Gene signal expression network: The node represent gene, and the size represent betweenness centrality value. Edges and arrows represent regulatory relationship.

Gene signal network and gene co-expression network construction

Gene signal network was constructed with 992 nodes and 3015 edges. The hub nodes included PRKACB (Betweenness = 73087), PLCB4 (Betweenness = 48412), MAPK8 (Betweenness = 38854) and PIK3R1 (Betweenness = 28792). PRKACB and MAPK8 were with indirect activation relationships, while PIK3R1 and PLCB4 were with metabolic relationships (**Figure 3**).

Gene co-expression network was also constructed with 456 nodes and 1713 edges. As shown in **Figure 4**, top 5 nodes were NEGR1 (degree = 52), MITF (degree = 49), TNFSF12 (degree = 47) and DAAM2 (degree = 47). Interestingly, NEGR1 and TNFSF12 were with positive relationship. Similarly, NEGR1 and DAAM2 were also with positive relationship (**Figure 4**).

Discussion

In order to screen biomarkers for ESCC disease prediction and chemosensitivity prediction, several DEGs were screened including NEGR1, MITF, PRKACB and PLCB4. These genes might be potential important genes in pathogenesis and therapeutic mechanism of ESCC.

NEGR1 was an important gene which regulated dendritic arborization. Kim et al. [13] confirmed that NEGR1 participated in cell recognition and

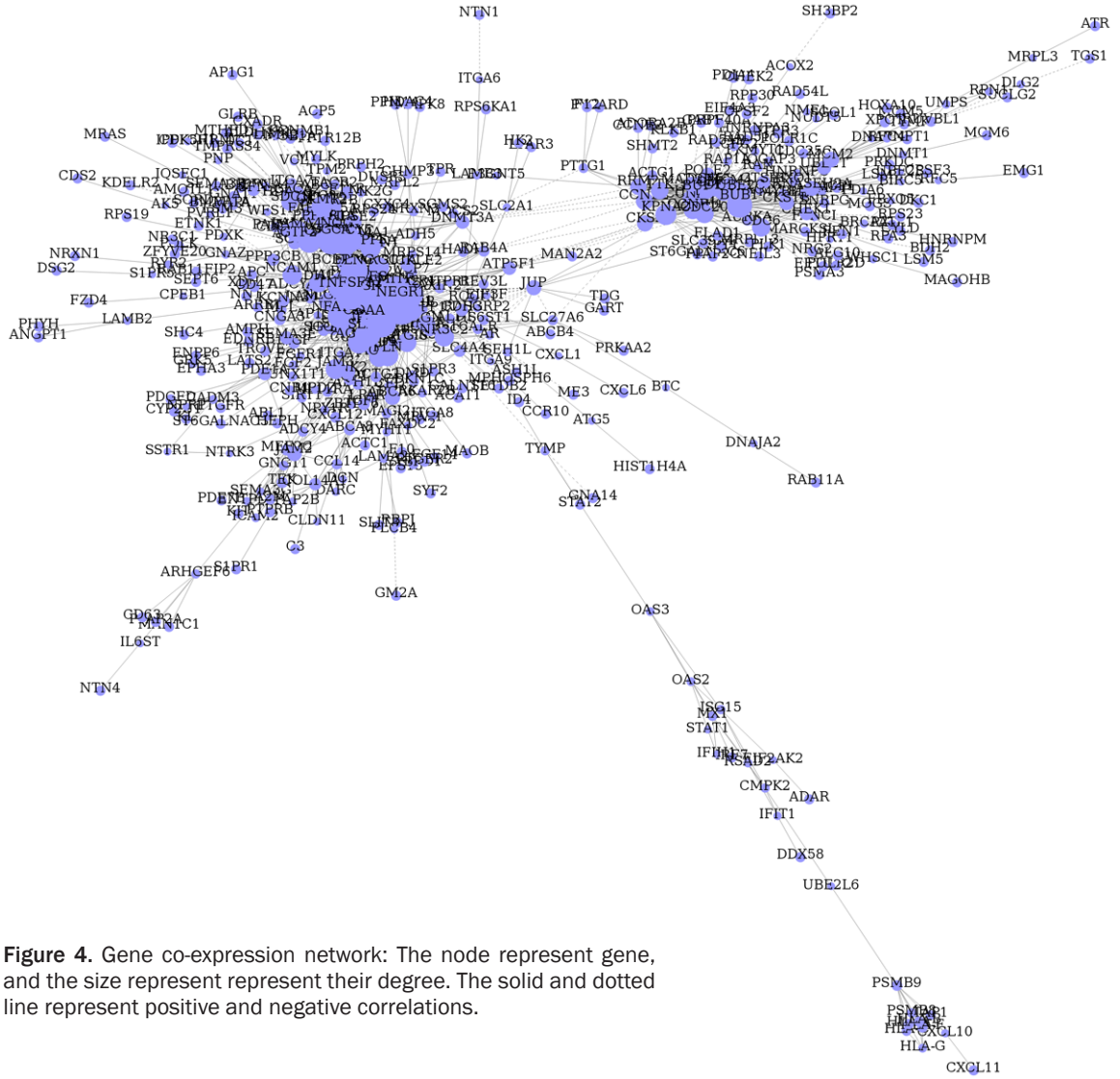

Figure 4. Gene co-expression network: The node represent gene, and the size represent represent their degree. The solid and dotted line represent positive and negative correlations.

interaction, and further contribute to tumor suppression. In the current study, NEGR1 was enriched in cell adhesion and cell adhesion molecules. In addition, NEGR1 also belonged in profile 13, which expressed lower in both ESCC treatment groups. However, the expression of this gene was with no difference between complete response and pathologic complete response groups. Similar results were obtained in study of Nishimori *et al*. [14], downregulated cell adhesion protein was detected in primary esophageal squamous cell carcinoma. Besides, several cell adhesion molecules such as CD44V6 were closely associated with pathogenesis of ESCC [15]. Thereby, we inferred that NEGR1 might be a biomarker for ESCC, but not related with sensitivity of chemotherapy.

MITF was screened with higher degree in gene co-expression network, and also belonged to profile 13 in this study. In addition, it enriched in positive regulation of transcription from RNA polymerase II promoter, negative regulation of apoptotic process and melanogenesis. As described in previous study, regulation of MITF occurred the downstream of various signaling, including FGF2, POMC and EDN1 [16]. The cell growth of ESCC was depended on different receptor tyrosine kinase signaling, and inhibition of these signals was closely related with the treatment of ESCC [17]. Besides, MITF was found to be with relationship with several genes in this study, such as CREB1, DCT and TYRP1. Wang et al. [18] found that CREB1 could participate in the process of DNA damage and apop-

tosis. Collectively, CREB1 might also be a biomarker for ESCC diagnosis.

PRKACB, a member of serine/threonine protein kinase family, encoded a protein which mediates signaling though cAMP [19]. It was well known that cAMP signaling was important to cell proliferation and differentiation [20]. In the present study, this gene was enriched in functions and pathways of blood coagulation, MAPK signaling pathway and small molecule metabolic process. In Kazakh patients, MAPK signaling pathway was found to be important in the early pathogenesis of ESCC [21]. By regulating MAPK/ERK signaling pathway, FAT1 was confirmed to promote the process of epithelial mesenchymal transition in ESCC [22]. It was worth noting that the changes of energy metabolism such as glucose metabolism play a critical role in malignant transformation of ESCC [23]. Importantly, PRKACB in this study was found to belong to profile 7, which expressed lower in complete response group than that in pathologic complete response group. PRKACB might be associated with sensitivity of chemotherapy. Thereby, this gene could be regarded as biomarkers for ESCC diagnosis and treatment.

PLCB4 encoded a protein which catalyzed the formation of inositol 1,4,5-trisphosphate and diacylglycerol from phosphatidylinositol 4,5-bisphosphate. Similarly with PRKACB, PLCB4 was also belonged to profile 7. Furthermore, this gene was also enriched in lipid metabolic process, chemokine signaling pathway and Wnt signaling pathway. In 2009, Diao et al. [24] found that the serum levels of lipid was associated with the different stages of ESCC. A new study from Germany found that ESCC cells could modulate hyaluronan synthesis and chemokine expression in fibroblasts [25]. Simultaneously, numbers of genes and proteins, such as RAP1B, NDRG1 and WNT5A have been confirmed to activate Wnt signaling pathway and further affect the development of ESCC [26-28]. Therefore, PLCB4 was also a potential gene for ESCC diagnosis and treatment.

In conclusions, screened genes including NEGR1, MITF, PRKACB and PLCB4 might be potential key genes for ESCC diagnosis. Importantly, PRKACB and PLCB4 were closely related with chemosensitivity, which might provide direction and help for postoperative chemotherapy.

Disclosure of conflict of interest

None.

Address correspondence to: Xuejun Yu, Department of Thoracic, Shouguang People's Hospital, No.47, Jiankang Street, Weifang 262700, China. Tel: +86-13869686111; Fax: +86-13869686111; E-mail: yuxuejun123abc@163.com

References

[1] Lin H, Li SD, Hu XG and Li ZS. Primary pancreatic lymphoma: report of six cases. World J Gastroenterol 2006; 12: 5064.

[2] Vaiphei K, Sinha SK and Kochhar R. Comparative analysis of Oct4 in different histological subtypes of esophageal squamous cell carcinomas in different clinical conditions. Asian Pac J Cancer Prev 2014; 15: 3519-24.

[3] Men Y, Hui ZG, Liang J, Feng QF, Chen DF, Zhang HX, Xiao ZF, Zhou ZM and Wang LH. [IMRT-based preoperative neoadjuvant chemoradiotherapy for thoracic esophageal squamous cell carcinoma (ESCC): an analysis of outcome and prognosis]. Zhonghua Zhong Liu Za Zhi 2016; 38: 682.

[4] Liu SL, Xi M, Yang H, Yang YD, Wu YJ, Zhao L, Zhang P, Luo LL and Liu MZ. Is there a correlation between clinical complete response and pathological complete response after neoadjuvant chemoradiotherapy for esophageal squamous cell cancer? Ann Surg Oncol 2016; 23: 273-81.

[5] Saeki H, Nakashima Y, Zaitsu Y, Tsuda Y, Kasagi Y, Ando K, Yu I, Ohgaki K, Ito S and Kimura Y. Current status of and perspectives regarding neoadjuvant chemoradiotherapy for locally advanced esophageal squamous cell carcinoma. Surgery Today 2015; 46: 1-7.

[6] Huang RW, Chao YK, Wen YW, Chang HK, Tseng CK, Chan SC and Liu YH. Predictors of pathological complete response to neoadjuvant chemoradiotherapy for esophageal squamous cell carcinoma. World J Surg Oncol 2014; 12: 170.

[7] Saeki H, Nakanoko T, Ohtsu H, Kawano H, Ando K, Oki E, Ohga T, Morita M, Kakeji Y and Toh Y. Neoadjuvant chemoradiotherapy for potentially resectable esophageal squamous cell carcinoma and the significance of Rad51 expression as a factor predictive of the treatment response. J Clin Oncol 2012.

[8] Okumura H, Uchikado Y, Setoyama T, Matsumoto M, Owaki T, Ishigami S and Natsugoe S. Biomarkers for predicting the response of esophageal squamous cell carcinoma to neoadjuvant chemoradiation therapy. Surgery Today 2014; 44: 421.

[9] Wen J, Yang H, Liu MZ, Luo KJ, Liu H, Hu Y, Zhang X, Lai RC, Lin T and Wang HY. Gene expression analysis of pretreatment biopsies predicts the pathological response of esophageal squamous cell carcinomas to neo-chemoradiotherapy. Ann Oncol 2014; 25: 1769.

[10] McGee M and Chen Z. Parameter estimation for the exponential-normal convolution model for background correction of affymetrix genechip data. Stat Appl Genet Mol Biol 2006; 1-27.

[11] Larsson O, Wahlestedt C and Timmons JA. Considerations when using the significance analysis of microarrays (SAM) algorithm. BMC Bioinformatics 2005; 6: 1-6.

[12] Hirano S and Tsumoto S. Cluster analysis of time-series laboratory test data based on the trajectory representation and multiscale comparison techniques. J Am Chem Soc 1950; 72: 2130-2132.

[13] Kim H, Hwang JS, Lee B, Hong J and Lee S. Newly identified cancer-associated role of human neuronal growth regulator 1 (NEGR1). J Cancer 2014; 5: 598.

[14] Nishimori T, Tomonaga TK, Oh IM, Kodera Y, Maeda T, Nomura F, Matsubara H, Shimada H and Ochiai T. Proteomic analysis of primary esophageal squamous cell carcinoma reveals downregulation of a cell adhesion protein, periplakin. Proteomics 2006; 6: 1011-1018.

[15] Zhang HZ, Hai gang LI and Huang HZ. Expression and clinical significance of cell adhesion molecules in esophageal squamous carcinoma. Chin J Cancer 1999.

[16] Haider AS, Duculan J, Whynot JA and Krueger JG. Increased JunB mRNA and protein expression in psoriasis vulgaris lesions. J Invest Dermatol 2006; 126: 912-4.

[17] Saito S, Morishima K, Ui T, Hoshino H, Matsubara D, Ishikawa S, Aburatani H, Fukayama M, Hosoya Y and Sata N. The role of HGF/MET and FGF/FGFR in fibroblast-derived growth stimulation and lapatinib-resistance of esophageal squamous cell carcinoma. BMC Cancer 2015; 15: 82.

[18] Wang Y, Hu Z, Liu Z, Chen R, Peng H, Guo J, Chen X and Zhang H. MTOR inhibition attenuates DNA damage and apoptosis through autophagy-mediated suppression of CREB1. Autophagy 2013; 9: 2069-86.

[19] Chen Y, Gao Y, Tian Y and Tian DL. PRKACB is downregulated in non-small cell lung cancer and exogenous PRKACB inhibits proliferation and invasion of LTEP-A2 cells. Oncol Lett 2013; 5: 1803-1808.

[20] Monje PV. To myelinate or not to myelinate: fine tuning cAMP signaling in Schwann cells to balance cell proliferation and differentiation. Neural Regen Res 2015; 10: 1936-1937.

[21] Zheng ST, Liu T, Middottuersun A, Huo Q, Liu Q, Huang CG, Feng JG, Lü GD, Wang X and Lin RY. [Expression of ERK1/2 MAPK signaling transduction pathway in esophageal cancers in Kazakh patients]. Zhonghua Zhong Liu Za Zhi 2011; 33: 421-425.

[22] Hu X, Zhai Y, Kong P, Cui H, Yan T, Yang J, Qian Y, Ma Y, Wang F and Li H. FAT1 promotes epithelial mesenchymal transition (EMT) via MAPK/ERK signaling pathway in esophageal squamous cell cancer. Cancer Lett 2017; 397: 83-93.

[23] de Andrade Barreto E, de Souza Santos PT, Bergmann A, de Oliveira IM, Wernersbach Pinto L, Blanco T, Rossini A, Pinto Kruel CD, Mattos Albano R, Ribeiro Pinto LF. Alterations in glucose metabolism proteins responsible for the Warburg effect in esophageal squamous cell carcinoma. Exp Mol Pathol 2016; 101: 66-73.

[24] Diao Y, Li H, Zhou Y, Ma Q, Wang Y and Li D. Association of serum levels of lipid and its novel constituents with the different stages of esophageal carcinoma. Lipids Health Dis 2009; 8: 48-48.

[25] Kretschmer I, Freudenberger T, Twarock S, Yamaguchi Y, Grandoch M and Fischer JW. Esophageal squamous cell carcinoma cells modulate chemokine expression and hyaluronan synthesis in fibroblasts. J Biol Chem 2016; 291: 4091.

[26] Li J, Ying J, Fan Y, Wu L, Ying Y, Chan AT, Srivastava G and Tao Q. WNT5A antagonizes WNT/β-catenin signaling and is frequently silenced by promoter CpG methylation in esophageal squamous cell carcinoma. Cancer Biol Ther 2010; 10: 617.

[27] Ai R, Sun Y, Guo Z, Wei W, Zhou L, Fang L, Hendricks DT, Yang X and Zhao X. NDRG1 overexpression promotes the progression of esophageal squamous cell carcinoma through modulating Wnt signaling pathway. Cancer Biol Ther 2016; 17: 943-954.

[28] Zhao J, Yang Y, Zhu D, Zhang C, Liu D, Wu K and Zhao S. RAP1B, a DVL2 binding protein, activates Wnt/beta-catenin signaling in esophageal squamous cell carcinoma. Gene 2017; 611: 15-20.