Identification of key genes and pathways associated with esophageal squamous cell carcinoma development based on weighted gene correlation network analysis

Mingrui Shao, Wenya Li*, Shiyang Wang, Zhenghua Liu

1. Department of Thoracic Surgery, The First Affiliated Hospital of China Medical University, Shenyang, Liaoning Province, China.
2. Department of Geriatric Surgery, The First Affiliated Hospital of China Medical University, Shenyang, Liaoning Province, China.

*Corresponding author: Prof Wen-ya Li, E-mail: saint5288@hotmail.com, Phone: 86-13514288257, Zip Code: 110001. Department of Thoracic Surgery, The First Affiliated Hospital of China Medical University, NO.155 North Nanjing Street, Heping District, Shenyang, Liaoning Province, China.

© The author(s). This is an open access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/).

See http://ivyspring.com/terms for full terms and conditions.

Received: 2018.10.15; Accepted: 2019.12.01; Published: 2020.01.13

Abstract

**Background:** As one of the most aggressive malignancies, esophageal squamous cell carcinoma (ESCC) remains one of the leading causes of cancer related death worldwide. The majority of ESCCs are diagnosed at advanced stages with poor five-year survival rate, making it urgent to identify specific genes and pathways associated with its initiation and prognosis.

**Materials and Methods:** The differentially expressed genes in TCGA were analysed to construct a co-expression network by WGCNA. Gene ontology (GO) terms and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways analysis were performed for the selected genes. Module–clinical trait relationships were analyzed to explore the genes and pathways that associated with clinicopathological parameters of ESCC. Log-rank tests and COX regression were used to identify the prognosis-related genes.

**Results:** The brown module containing 716 genes which most significantly contributed to ESCC. GO analysis suggested enrichment of adaptive immune response, cyclin–dependent protein serine, regeneration and mRNA metabolic process. KEGG analysis indicated pathways including Cellular senescence, Ribosome biogenesis, Proteasome, Base excision repair and p53 signaling pathway. Clinical stage was associated with cyan module; clinical M was associated with grey60 module; clinical T was associated with darkturquoise module; while clinical N, histological type and cancer location were associated with turquoise module. Key genes of TCP1, COQ3, PTMA and MAPRE1 might be potential prognostic markers for ESCC.

**Discussion:** Differentially expressed genes and key modules contributing to initiation and progression in ESCC were identified by WGCNA. These findings provide novel insights into the mechanisms underlying the initiation, prognosis and treatment of ESCC.

Key words: esophageal squamous cell carcinoma, weighted gene correlation network analysis, risk, prognosis.

Introduction

As one of the most aggressive malignancies, esophageal cancer remains one of the leading causes of cancer related death worldwide [1, 2]. Among the two main histological types of esophageal cancer including esophageal squamous cell carcinoma (ESCC) and esophageal adenocarcinoma (EAC), ESCC represents the dominant subtype in Asians and EAC incidence is higher in Western countries [3]. These two subtypes have different pathogenesis, biology features and clinical outcomes [4]. At present, the majority of ESCCs are diagnosed at advanced stages, the five-year survival rate of which still remains unfavourable with frequently metastasis and recurrence [5, 6]. Therefore, comprehensive researches
are urgently needed to fully clarify critical molecular mechanisms related to the initiation, progression and prognosis of ESCC.

It has been widely accepted that alcohol drinking and tobacco consumption are certain risk factors for ESCC with synergistic effects [7, 8]. There is also evidence suggesting that possible association of human papillomavirus (HPV) with ESCC risk according to meta-analysis [9]. Studies have found that polymorphisms in aldehyde dehydrogenase 2 (ALDH2) gene also lead to the development of ESCC especially in Asian populations [10-12]. Multiple studies including high-throughput analysis revealed that genes and pathways involved in ESCC were related with cell cycle, differentiation, and Epidermal Growth Factor Receptor signalling [13]. In addition, epigenetic alterations including DNA methylation of such as APC, RB1 and CDKN2A, histone modification, and loss of genome imprinting also contribute to ESCC [14-16].

Although a number of genes and mechanisms have been proved to be closely implicated in the development of ESCC, the comprehensive picture of the whole genes and regulations of ESCC is still unclear. In recent years, bioinformatic methods become increasingly effective in exploration and analysis of multiple genes or proteins of complicated diseases. Weighted gene co-expression network analysis (WGCNA), a new gene co-expression network-based method, has been successfully used to screen biomarkers and pathways that could be applied in susceptibility genes, diagnose and treatment of cancer [17]. In this study, WGCNA was conducted to analyse data of TCGA data repository of ESCC to identify gene modules and biomarkers (hub genes) implicated in the pathogenesis, progression and prognosis of ESCC.

Materials and Methods

Acquiring and preparing genetic and clinical data

The RNA sequencing and clinical data of esophageal squamous cell carcinoma patients were downloaded from TCGA data repository (https://cancergenome.nih.gov/). The level of gene expression was measured as fragments per kilobase of transcript per million mapped reads (FPKM). Clinical data included the sample type, clinical TNM stage, histologic grade, cancer position and survival information. Samples with incomplete pathologic stage or histologic grade information were not included. As genes with little variation in expression usually represent noise, the most variant genes were filtered for network construction. Gene variabilities were measured by median absolute deviation (MAD).

Constructing gene co-expression network

Gene co-expression network was constructed by the WGCNA package in R and was visualized by Cytoscape software [18]. Power values were screened out by WGCNA algorithm in the construction of co-expression modules. Scale independence and average connectivity analysis of modules with different power value were performed by gradient test (power value ranging from 1 to 20). Appropriate power value was determined when the scale independence value was equal to 0.9. WGCNA algorithm was then used to construct the co-expression network and extract the gene information in the most relevant module. The criterion of co-expression weight> 2 was used to select the candidate network. Heatmap tool package in R language was adopted to describe the strength of the relationship among modules (strong or weak degree).

Relating modules to cancer risk and identifying the prognosis related genes

One of the advantages of WGCNA is that the correlation between modules and clinical parameters can be analyzed. The module eigengene (ME), which can be regarded as a representative of the gene expression profiles from a module, is defined as the first principal component of a given module. Given that the ME can summarize the gene expression profiles, we calculated the correlation between MEs and external sample type data. The gene in the most relevant module was chosen as the risk-related gene. Log-rank tests and COX regression were used to select the prognosis-related genes in the risk-related genes. R package survival was used to carry out log-rank tests and survminer package was used to plot Kaplan–Meier survival curves. In order to validate our findings, microarray dataset of GSE53625 from Gene Expression Omnibus (GEO, https://www.ncbi.nlm.nih.gov/geo/) in NCBI (The National Center for Biotechnology Information) containing 179 esophageal squamous cell carcinoma patients with prognosis information were used to confirm our results.

Gene ontology and pathway Enrichment analysis

To explore the potential biological themes and pathways of genes in risk-related module, the clusterprofiler package in R was used to annotate and visualize gene ontology (GO) terms [19] and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways [20].

http://www.jcancer.org
Screening for candidate module with clinical significance and KEGG analysis

We calculated the correlation between MEs and external clinical data. $p < 0.05$ suggested significant correlation. We selected the most relevant module in each clinical data for analysis. KEGG pathway analysis was used to enrich the related genes.

Results

Gene co-expression network of ESCC

Clinical and level-3 RNA sequencing data for 162 esophageal squamous cell carcinoma samples and 11 normal samples were obtained from TCGA database. For module detection, the 15121 most variant genes were selected according to MAD value for further analysis. When the value of soft thresholding power $\beta$ was 4, the connectivity between genes met a scale-free network distribution (Fig. S1). Altogether 21 modules were identified by hierarchical clustering and the Dynamic branch Cutting. Each module was assigned a unique color as an identifier. Interaction relationship analysis of co-expression genes was shown in Figure 1. The number of genes in modules ranged from 34 to 2353. The grey module represented a gene set that was not assigned to any of the modules.

The risk related module in the WGCNA algorithm

We then use the cancer type as the Clinical phenotype to select the most relevant module with the occurrence of ESCC. Finally, the brown module was selected (Fig. S2). The brown module contained 716 genes. Then we constructed the co-expression network with the criterion of co-expression weight $\geq 2$. As was shown in Figure 2, a total of 161 genes were selected and their multiple interactions were visualized. The size represents the connection numbers and the color represents the fold change compared by the cancer and normal samples.

Figure 1. The connectivity analysis of critical genes in different module
Enrichment analysis of the brown module

GO and KEGG enrichment analysis was performed on the genes in the constructed cancer-risk-related module. Altogether 984 terms showed difference in GO enrichment (Table 1). As was illustrated in Figure 3, this module in GO analysis was related with biological process of regulation of adaptive immune response, regulation of cyclin-dependent protein serine, regeneration and positive regulation of mRNA metabolic process. According to the KEGG analysis, 19 pathways were associated with brown module including Cellular senescence, Ribosome biogenesis in eukaryotes, Proteasome, Base excision repair and p53 signaling pathway.
Prognosis analysis of the brown module

Since the brown module were selected from the Gene Ontology analysis and KEGG pathway enrichment analysis for genes in the risk-related brown module of ESCC. The brown module included the following genes:

- **TCPI**: t-complex 1
- **MTHFD2**: methylenetetrahydrofolate dehydrogenase (NADP+ dependent) 2
- **COQ3**: coenzyme Q3, methylenetetrahydrofolate reductase
- **AC016205.1**: Chromosome 7 open reading frame 17
- **MRPL18**: mitochondrial ribosomal protein L18
- **LSM5**: LSM5 homolog, U6 small nuclear RNA and mRNA degradation associated
- **TLT**: Tubulin Tyrosine Ligase
- **PTGES3**: prostaglandin E synthase 3
- **SRRF2**: serine and arginine rich splicing factor 2
- **HSPD1**: heat shock protein family D (Hsp60) member 1
- **SFA3**: splicing factor 3a subunit 3
- **PTMA**: prothymosin, alpha
- **CTPS1**: CTP synthase 1
- **MAGOH**: mago homolog B, exon junction complex core component
- **MAPRE1**: microtubule associated protein RP/EB family member 1
- **MCMB**: minichromosome maintenance 8 homologous recombination repair factor
- **TRIM37**: tripartite motif containing 37
- **HNRNPA/B**: heterogeneous nuclear ribonucleoprotein A/B
- **STRAP**: serine/threonine kinase receptor associated protein
- **BOLA3**: Bola Family Member 3
- **LSM6**: LSM6 homolog, U6 small nuclear RNA and mRNA degradation associated

The genes in the brown module were found to be significantly associated with various functions and pathways, including:

- Positive regulation of adaptive immune response
- Positive regulation of mRNA metabolic process
- Negative regulation of DNA replication

The significance of these genes in the brown module was evaluated using a variety of statistical methods, including logistic regression analysis and Kaplan-Meier survival analysis. The results suggested that these genes are potential prognostic biomarkers for patients with ESCC.

**Table 2. Prognosis analysis of the brown module**

<table>
<thead>
<tr>
<th>Gene</th>
<th>Gene full name</th>
<th>Location</th>
<th>HR</th>
<th>Lower limit</th>
<th>Upper limit</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>TCPI</td>
<td>t-complex 1</td>
<td>6q25.3-q26</td>
<td>2.427</td>
<td>1.439</td>
<td>4.098</td>
<td>0.001</td>
</tr>
<tr>
<td>MTHFD2</td>
<td>methylenetetrahydrofolate dehydrogenase (NADP+ dependent) 2</td>
<td>2p13.1</td>
<td>2.179</td>
<td>1.302</td>
<td>3.650</td>
<td>0.003</td>
</tr>
<tr>
<td>COQ3</td>
<td>coenzyme Q3, methylenetetrahydrofolate reductase</td>
<td>6q16.2</td>
<td>2.247</td>
<td>1.318</td>
<td>3.831</td>
<td>0.003</td>
</tr>
<tr>
<td>AC016205.1</td>
<td>Chromosome 7 open reading frame 17</td>
<td>18q12.2</td>
<td>2.227</td>
<td>1.284</td>
<td>3.861</td>
<td>0.004</td>
</tr>
<tr>
<td>MRPL18</td>
<td>mitochondrial ribosomal protein L18</td>
<td>Xq28</td>
<td>2.088</td>
<td>1.230</td>
<td>3.546</td>
<td>0.006</td>
</tr>
<tr>
<td>LSM5</td>
<td>LSM5 homolog, U6 small nuclear RNA and mRNA degradation associated</td>
<td>8q24.11</td>
<td>2.110</td>
<td>1.208</td>
<td>3.690</td>
<td>0.009</td>
</tr>
<tr>
<td>TLT</td>
<td>Tubulin Tyrosine Ligase</td>
<td>6p21.31</td>
<td>1.976</td>
<td>1.178</td>
<td>3.311</td>
<td>0.010</td>
</tr>
<tr>
<td>PTGES3</td>
<td>prostaglandin E synthase 3</td>
<td>1q32</td>
<td>1.805</td>
<td>1.074</td>
<td>3.040</td>
<td>0.026</td>
</tr>
<tr>
<td>SRRF2</td>
<td>serine and arginine rich splicing factor 2</td>
<td>1q24.34</td>
<td>1.832</td>
<td>1.005</td>
<td>3.098</td>
<td>0.002</td>
</tr>
<tr>
<td>HSPD1</td>
<td>heat shock protein family D (Hsp60) member 1</td>
<td>18q12.21</td>
<td>1.795</td>
<td>1.063</td>
<td>3.030</td>
<td>0.029</td>
</tr>
<tr>
<td>SFA3</td>
<td>splicing factor 3a subunit 3</td>
<td>12q21.2</td>
<td>1.828</td>
<td>1.055</td>
<td>3.165</td>
<td>0.031</td>
</tr>
<tr>
<td>PTMA</td>
<td>prothymosin, alpha</td>
<td>17q25.3</td>
<td>1.887</td>
<td>1.138</td>
<td>3.125</td>
<td>0.014</td>
</tr>
<tr>
<td>CTPS1</td>
<td>CTP synthase 1</td>
<td>17q25.3</td>
<td>1.887</td>
<td>1.138</td>
<td>3.125</td>
<td>0.014</td>
</tr>
<tr>
<td>MAGOH</td>
<td>mago homolog B, exon junction complex core component</td>
<td>11p15.1</td>
<td>1.783</td>
<td>1.073</td>
<td>2.959</td>
<td>0.026</td>
</tr>
<tr>
<td>MAPRE1</td>
<td>microtubule associated protein RP/EB family member 1</td>
<td>1q32</td>
<td>1.805</td>
<td>1.074</td>
<td>3.040</td>
<td>0.026</td>
</tr>
<tr>
<td>MCM8</td>
<td>mini-chromosome maintenance 8 homologous recombination repair factor</td>
<td>12q13.3</td>
<td>1.733</td>
<td>1.040</td>
<td>2.890</td>
<td>0.035</td>
</tr>
<tr>
<td>TRIM37</td>
<td>tripartite motif containing 37</td>
<td>17q25.1</td>
<td>1.718</td>
<td>1.040</td>
<td>2.841</td>
<td>0.035</td>
</tr>
<tr>
<td>HNRNPA/B</td>
<td>heterogeneous nuclear ribonucleoprotein A/B</td>
<td>12q21.2</td>
<td>1.795</td>
<td>1.063</td>
<td>3.030</td>
<td>0.029</td>
</tr>
<tr>
<td>STRAP</td>
<td>serine/threonine kinase receptor associated protein</td>
<td>17q21.2</td>
<td>1.718</td>
<td>1.031</td>
<td>2.865</td>
<td>0.038</td>
</tr>
<tr>
<td>BOLA3</td>
<td>Bola Family Member 3</td>
<td>12p13.1</td>
<td>1.706</td>
<td>1.026</td>
<td>2.833</td>
<td>0.039</td>
</tr>
<tr>
<td>LSM6</td>
<td>LSM6 homolog, U6 small nuclear RNA and mRNA degradation associated</td>
<td>2p13.1</td>
<td>1.684</td>
<td>1.022</td>
<td>2.770</td>
<td>0.041</td>
</tr>
</tbody>
</table>

The HR values show the relative risk of cancer progression or survival, with a higher HR indicating a greater risk. The P values indicate the statistical significance of the association between each gene and cancer outcome. Values less than 0.05 are considered statistically significant.
genes, 39 genes were associated with the prognosis of esophageal squamous cell carcinoma (Table 2). As was summarized in Figure 4, the top four up-regulated genes which indicated a poor prognosis included TCP1 (HR=2.427, \( P=0.001 \)), COQ3 (HR=2.247, \( P=0.003 \)), AC016205.1 (HR=2.227, \( P=0.004 \)) and MTHFD2 (HR=2.179, \( P=0.003 \)). Of the top four down-regulated genes, PTMA (HR=1.706, \( P=0.039 \)), MAPRE1 (HR=1.698, \( P=0.043 \)) and BOLA3 (HR=1.667, \( P=0.048 \)) significantly associated with poor prognosis while TTL predicted better survival (HR=0.572, \( P=0.035 \)).

Altogether 179 esophageal squamous cell carcinoma patients with prognosis information of microarray GSE53625 were used to validate our results. As was shown in Table 4, AC016205.1 (\( P=0.013 \)), PTMA (\( P=0.023 \)), TTL (\( P=0.013 \)) still showed significant relation with prognosis of ESCC patients, while MTHFD2 (\( P=0.082 \)) indicated borderline significant association (Table 4).

**Module–clinical trait relationships**

Identifying genes associated with a certain clinical trait is of great value to explore the molecular mechanisms behind the trait. In the present study, the clinical parameters of ESCC patients, including clinical T, clinical N, clinical M, clinical stage, recrudescence stage, histological type and cancer location were involved in the module–trait relationship (MTR) analysis. As was suggested in Figure 5, clinical stage was associated with cyan module; clinical M was associated with grey60 module; clinical T was associated with darkturquoise module; while clinical N, histological type and cancer location were associated with turquoise module. No module correlated with recrudescence.

**KEGG analysis of clinical related module**

We then performed KEGG pathway analysis of the candidate module. The results indicated that genes of cyan module were enriched in Cell cycle; grey60 module was associated with Ribosome; darkturquoise module was related with RNA transport; turquoise module was associated with Fat digestion and absorption, Tight junction, Maturity onset diabetes of the young, Protein digestion and absorption and Fructose and mannose metabolism (Table 3).

**Figure 4.** The correlation between the expression levels of key genes of risk-associated brown module and the survival of ESCC patients.
Table 3. KEGG analysis of clinical related module

<table>
<thead>
<tr>
<th>Phenotype</th>
<th>Module description</th>
<th>ID</th>
<th>Description</th>
<th>GeneRatio</th>
<th>Adjusted P</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clinical stage cyan</td>
<td>Cell cycle</td>
<td>hsa04110</td>
<td>hsa04110</td>
<td>4/29</td>
<td>0.0012644</td>
<td>4</td>
</tr>
<tr>
<td>Clinical M grey60</td>
<td>Ribosome</td>
<td>hsa00100</td>
<td>hsa00100</td>
<td>4/24</td>
<td>0.0013257</td>
<td>4</td>
</tr>
<tr>
<td>Clinical T darkturquoise</td>
<td>RNA transport</td>
<td>hsa00133</td>
<td>hsa00133</td>
<td>6/45</td>
<td>0.0005229</td>
<td>6</td>
</tr>
<tr>
<td>Clinical N turquoise</td>
<td>Fat digestion and absorption</td>
<td>hsa04975</td>
<td>hsa04975</td>
<td>14/607</td>
<td>2.21E-06</td>
<td>14</td>
</tr>
<tr>
<td>Histological type</td>
<td>Tight junction</td>
<td>hsa04530</td>
<td>hsa04530</td>
<td>31/607</td>
<td>1.499E-05</td>
<td>31</td>
</tr>
<tr>
<td>Cancer location</td>
<td>Maturity onset diabetes of the young</td>
<td>hsa04950</td>
<td>hsa04950</td>
<td>10/607</td>
<td>1.936E-05</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td>Protein digestion and absorption</td>
<td>hsa04974</td>
<td>hsa04974</td>
<td>20/607</td>
<td>2.754E-05</td>
<td>20</td>
</tr>
<tr>
<td></td>
<td>Fructose and mannose metabolism</td>
<td>hsa0051</td>
<td>hsa0051</td>
<td>11/607</td>
<td>3.566E-05</td>
<td>11</td>
</tr>
</tbody>
</table>

Figure 5. The module–clinical trait relationships of genes involved in ESCC.
Discussion

Although significant progresses have been made in studies concerning the risk and development of ESCC, our understanding of the complex mechanisms of ESCC is still limited. In this study, we conducted a differential expression analysis followed by WGCNA to detect genes and pathways associated with the occurrence, clinical parameters and prognosis of ESCC. Finally, key modules were identified which contribute to the risk (brown module) and clinicopathological parameters (cyan module, grey60 module, darkturquoise module and turquoise module) of ESCC. In addition, enrichment analysis of the genes in core modules suggested significant involvement of pathways such as Cellular senescence, Ribosome biogenesis in eukaryotes, Proteasome, Base excision repair and p53 signalling pathway.

In this present study, RNA sequencing data for 162 esophageal squamous cell carcinoma samples and 11 normal samples from TCGA were systematically analyzed. According to the WGCNA analysis of most variant genes, altogether 21 modules were identified and each module was assigned a unique color as an identifier. The number of genes in modules ranged from 34 to 2353. The brown module containing 716 genes was selected when we used the cancer type as the Clinical phenotype, indicating its close involvement in the development of ESCC. GO enrichment analysis were then performed on the genes in the constructed cancer-risk-related module, suggesting significance of biological process including adaptive immune, cyclin-dependent protein serine, regeneration and mRNA metabolic process. It has been reported that immune response and immune escape might play a critical role in ESCC progression and therapy [21, 22]. In addition, key mRNA metabolic procedure such as alternative splicing has also been detected in ESCC. Alternative splicing isoforms of LOXL2, VIL2, OSMRβ and MUC1 have been reported to contribute to ESCC development and progression [23-26]. From this point of view, the specific role of immune response and altered mRNA metabolic process in ESCC require future investigations to elucidate. As for key pathways of KEGG enrichment, aberrant base excision repair capacity and altered p53 signaling pathways have been found to be associated with ESCC development by a number of studies [27, 28]. Pathways of cellular senescence, ribosome biogenesis in eukaryotes and proteasome might be novel research directions for physicians and surgeons for ESCC.

The five-year survival rate of ESCC currently remains unsatisfactory despite recent improvements in treatments of surgical resection and adjuvant chemotherapy. It was therefore of great significance to evaluate the potentiality of key factors to serve as prognostic biomarkers. We eventually screened 39 prognosis-related genes in ESCC after analyzing the risk-associated brown module. Aberrant TCP1 gene status has been detected in hepatocellular carcinoma, breast cancer and colorectal cancer, but no study has been performed for its effect in ESCC [29-31]. MTHFD2 RNA and protein are remarkably increased in many types of cancers and correlated with worse survival in breast cancer according to a comprehensive analysis of RNA profiles of 1,454 metabolic enzymes across 1,981 tumors spanning 19 cancer types [32]. Although some of the prognosis-related genes we identified have been investigated in various types of cancers, limited information was known about their roles in ESCC until now. One limitation of this study is that most results were generated by analysis of publicly available data, further studies based on larger populations and molecular mechanism researches are therefore needed to confirm the results of our study.

In order to explore the genes and modules that associated with clinicopathological parameters of ESCC, we also performed module-clinical trait relationships analysis. We finally detected specific modules which determine the clinical stage, TNM classification, histological type and cancer location. Pathway analysis of the selected modules indicated that these genes enriched in Cell cycle, Ribosome, RNA transport, Fat digestion and absorption, Tight junction, Maturity onset diabetes of the young, Protein digestion and absorption and Fructose and mannose metabolism. Remodeling of energetic metabolism is a hallmark of malignant tumor, by
which tumor cells change themselves and the environment into more suitable conditions for growth and invasion [33]. According to our findings, aberrant metabolism of carbohydrate, fat, protein and might be closely related to the progression of ESCC. In addition, alternations in functions of ribosome, RNA transport might also be promising research directions for ESCC.

In summary, differentially expressed genes and key modules contributing to initiation and progression in ESCC were identified by means of WGCNA. Pathways including cellular senescence, ribosome biogenesis in eukaryotes, Proteasome, base excision repair, fat digestion and absorption and p53 signalling might be closely related with ESCC development. Key genes of TCP1, COQ3, PTMA and MAPRE1 might be potential prognostic markers for ESCC. These findings provide novel insights into the mechanisms underlying the etiology, prognosis and treatment of ESCC.

Supplementary Material
Supplementary figures and tables.
http://www.jcancer.org/v11p1393s1.pdf

Acknowledgements
This study was supported by Natural Science Foundation of Liaoning Province (grant no. 2015020561), Wu JiePing Medical Found (grant no. 320.6750.18293) and the Fund for Scientific Research of The First Hospital of China Medical University (grant no. fsfh1514).

Competing Interests
The authors have declared that no competing interest exists.

References
3. Siewert JR, Ott K. Are squamous and adenocarcinomas of the esophagus closely related to the progression of ESCC. In addition, alternations in functions of ribosome, RNA transport might also be promising research directions for ESCC. Metabolism of carbohydrate, fat and protein might be closely related with ESCC development. Key genes of TCP1, COQ3, PTMA and MAPRE1 might be potential prognostic markers for ESCC. These findings provide novel insights into the mechanisms underlying the etiology, prognosis and treatment of ESCC.

31. Guest ST, Kretche ZR, Bollig-Fischer A, Haddad R, Ethier SP: Two members of the TRiC chaperonin complex, CCT2 and TCP1 are essential for survival of breast cancer cells and are linked to driving oncogenes. Experimental cell research 2015, 332(2):223-235.