Research Article

Proteomic Analysis of Human Esophageal Cancer Using Tandem Mass Tag Quantifications

Suofeng Sun,1 Huijuan Zhang,2 Yu Wang,1,3 Jing Gao,1 Shen Zhou,1 Yuan Li,4 Shuangyin Han,1 Xiuling Li,1 and Jian Li1

1Department of Gastroenterology, Henan Provincial People’s Hospital and the People’s Hospital of Zhengzhou University, Zhengzhou, China 450000
2Medical Care Department, First Affiliated Hospital of Hainan Medical College, Haikou, Hainan, China 570102
3Xinxiang Medical University, Zhengzhou, China 450000
4Department of Traditional Chinese Medicine, The Third Affiliated Hospital Affiliated of Henan University of Traditional Chinese Medicine, Zhengzhou, Henan, China 450000

Correspondence should be addressed to Xiuling Li; zlxiiuling@yeah.net and Jian Li; jiuiwei@163.com

Suofeng Sun, Huijuan Zhang, and Yu Wang contributed equally to this work.

Received 22 April 2020; Revised 11 June 2020; Accepted 1 July 2020; Published 8 August 2020

Guest Editor: Jialiang Yang

Copyright © 2020 Suofeng Sun et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Esophageal cancer (EC) is a type of extremely aggressive gastrointestinal cancer with high incidences in China and other Asian countries. EC does not have specific symptoms and is relatively easy to metastasize, which makes it difficult in early diagnosis. Thus, novel noninvasive diagnostic method is urgently needed in clinical practice. In this study, mass spectrometry with tandem mass tags and differential protein analysis were applied for identifying esophageal cancer-related proteins. The identified proteins were annotated based on their enrichment in Gene Ontology (GO) terms and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways. In addition, hierarchical clustering was applied based on differentially expressed proteins. As a result, a total of 5131 quantifiable proteins were identified from our liquid chromatography-tandem mass spectrometry with tandem mass tags (LC-MS/MS-TMT) method with 63 upregulated and 97 downregulated differential proteins between esophageal cancer and controlled normal samples. The differentially expressed proteins were highly enriched in GO terms associated with mitochondrial dissemble and apoptosis, and blood vessel regulation, and the upregulated differentially expressed proteins in EC samples were significantly enriched in major histocompatibility complex MHC-class I/II pathway of immune system. The functional clustering analysis revealed potential protein-protein interactions among tetraspanin, myosin, and S-100. In summary, our study provided a practical technological procedure of proteomic analysis for discovering novel biomarkers of a specific cancer type.

1. Introduction

Esophageal cancer (EC) processes a significant health risk due to increasing incidence and poor prognosis [1]. As an extremely aggressive neoplasm, approximately 45,000 people are diagnosed with EC each year, while the overall 5-year survival rate is less than 10%. Although chemotherapy and adjuvant chemotherapy are widely used in the treatment of esophageal cancer, the prognosis remains poor due to high possibility of tumor relapse or distant metastasis. Beginning in the mucosa of the esophagus and spreading through a deeper tissue layer, such as the submucosa, muscular layer, and serosa, esophageal cancer cells are able to metastatic progress through lymphatic and homogametic vessels. The most common pathologically histological types are squamous cell and adenocarcinoma, which usually occur at the lower (distal) part of the esophagus [2–4]. Endoscopy typically shows a tumor mass at the distal or gastroesophageal junction. One hypothesis of the esophageal cancer model is the squamous epithelium undergoing chronic inflammatory changes, leading the changes in the cell structure and shape and in situ malignant changes as well [5–7].
The risk factors of esophageal carcinoma include history of symptomatic gastroesophageal reflux disease, tobacco use, and alcohol intake. Barium esophagography is widely used as the initial assessment in patients with symptoms of esophageal carcinoma. However, the confirmed diagnosis based on biopsy tissue required an invasive procedure. Dysphagia with weight loss is the only symptom of esophageal cancer in early stage; esophageal cancers are commonly diagnosed at advanced stages, highlighting the need for improved detection and prediction methods with novel biomarkers. Furthermore, esophageal cancer in advanced stage was insensitive to chemoradiotherapy making therapeutic even challenging; diagnosis based on biomarkers in the early stage will benefit the prognosis and overall survival rate of EC patients. Previous studies have demonstrated that many proteins play an important role in tumor development and their abnormal expressions are associated with cancer cell proliferation and migration, such ATP-binding cassette protein E1 and loss of PAR4 gene expression.

Proteomics has been identified as a powerful approach for novel disease biomarkers discovery, and mass spectrometry with tandem mass tags (TMT) for proteomics profiling has been reported in various cancer types. Zhang et al. reported plasma exosomes from an ovarian cancer patient from detection by LC-MS/MS with TMT containing tumor-specific proteins relevant to tumorigenesis and metastasis, while Hou et al. developed a drug anti-HCC efficacy method for hepatocellular carcinoma.

In this study, we applied the tandem mass tag- (TMT-) based quantitative proteomic technique for esophageal cancer proteomics profiling annotation biological meaning and comparing protein-protein interaction. Our study identified potential biomarker for early diagnosis to discover relative effective therapy in the future clinical practices.

2. Materials and Methods

2.1. Protein Extraction and TMT/iTRAQ Labeling. Tissue samples of three esophageal cancer patients were collected and sonicated for 5 min (30 s on and 30 s off) by a high-intensity ultrasonic processor. The protein solution was centrifuged for 15 min and reduced at 56°C with dithiothreitol, diluted with 100 mM TEAB, and digested with trypsin for 5 hours. The proteomic workflow included desalting and vacuum centrifugation. The protein solution was then analyzed by tandem mass spectrometry in Q Exactive Plus (Thermo Scientific, MA) with full-scan setting of m/z from 300 to 1800 and resolution of 70000 for Orbitrap. NCE setting of 28 was selected for intact fraction detection and selected for further 20 MS scans with 10 s dynamic exclusion.

2.3. Proteomics Analysis and Annotations. The MaxQuant search engine was used for LC-MS/MS-TMT result analysis, and tandem mass spectra were searched against human UniProt database concatenated with reverse decoy database. The mass tolerance was set as follows: 0.02 Da for fragment ions and 5 ppm of main search for precursor ions. Fixed modifications and variable modifications were specified with carboxymethyl on Cys and oxidation on Met. GO annotation was applied for differentially expressed proteins with biological process, cellular compartment, and molecular function, and enrichment was tested with two-tailed Fisher's exact. Each category protein was also searched in InterPro database for biological function analysis. Kyoto Encyclopedia of Genes and Genomes (KEGG) database was applied for enriched pathway identification. All annotation analyses with a corrected p value < 0.05 were considered significant.

2.4. Enrichment-Based Clustering and Protein-Protein Interaction Network. We used WoLF PSORT to predict subcellular localization of proteins. WoLF PSORT is an updated version of PSORT/PSORT II for the prediction of eukaryotic sequences. Special for proteokaryon species, subcellular localization prediction soft CELLO was used. We applied hierarchical clustering based on differential expressed protein functional classification and collated all the categories obtained after enrichment along with their -log10 (p value) and then normalized with z-score for hierarchical clustering in Genesis. Cluster membership was visualized by “heatmap.2” function from the “gplots” R-package. STRING database (version 10.1) was used for the prediction of protein-protein interactions with differentially expressed proteins, and external candidates were excluded. The confidence score was calculated for the connection evaluation of protein-protein interaction, and 0.7 was set as the cut-off for PP interaction.

3. Results

3.1. Identification of Differentially Expressed Proteins by LC-MS/MS-TMT. Extracted proteins from esophageal cancer samples were quantified by LC/LC-MS with TMT as shown in Figure 1(a). A total of 367174 secondary spectrograms obtained were analyzed by mass spectrometry, and 98555 (26.8%) effective spectrograms obtained were filtered by MaxQuant. 47185 peptides were identified by spectrogram analysis with 96.9% unique peptides. Overall, 5823 proteins were identified with database and 5131 (88.12%) proteins could be quantifiable with comparable information in the control group (Figure 1(b) and Supplementary Table 1). Principal component analysis (PCA) was performed to validate the quantitative results between the biological triplicate samples, and the EC109 group shows better repeatability (Figure 1(c)). Relative standard deviation (RSD) distribution shows that protein quantification evaluation was comparable between replicates and no significant difference.
in overall protein quantification between EC109 and EC109_T group (Figure 1(d)). The proteins of all samples are quantitatively detected, and differentially expressed protein was defined with $p$ value < 0.05 and log2 fold change > 1.2 based on mean expression. We detected 63 upregulated proteins, 97 downregulated proteins between two groups with 57/62/43 upregulated and 60/112/69 downregulated proteins separately (Figure 1(e)).

3.2. Functional Analysis of Identified Differentially Expressed Proteins. To further analyze differentially expressed proteins, systematic bioinformatics analysis was applied including functional classification, function enrichment, and enrichment-based clustering analysis. The top three significant biological processes of the differentially expressed protein in this project were (1) cellular process, (2) single-organism process, and (3) biological regulation. For molecular function, the results indicated that many proteins are involved in binding and catalytic activities, meaning those proteins participated in cell-to-cell signaling transduction and facilitated signal pathway regulation. Gene ontology terms of quantified proteins also revealed a response to stimulus and immune system processes are enriched in the biological process, while catalytic activity of differentially expressed protein may be involved in development (Figure 2(a)). Subcellular localization analysis illustrated that cytoplasm accounted for 24.38% in the overall cellular structure, while 13.75% proteins from the mitochondria suggested potential mitochondria-mediated oxidative stress regulation between EC109 and EC109_T groups (Figure 2(b)). We further checked the number of proteins in COG/KOG categories (Clusters of Orthologous Groups of proteins) in all the comparable groups; 28 proteins are associated with signal transduction mechanism and post-translational modification, revealing that signaling between cells is important in esophageal cancer. Furthermore, 16 other proteins are categorized in lipid transport and metabolism and the protein metabolic pathway, while the metabolic product from lipid metabolism served as long-term

Figure 1: Overall design and protein identification. (a) Overall study design and analysis process. (b) Protein identification from 367174 secondary spectrograms. (c) PCA analysis between EC109 and EC109_T. (d) Boxplot of RSD distribution for protein quantification evaluation between replicates. (e) Differentially expressed protein analysis between EC tumor and control samples.
energy for cell development and maintaining accurate functions. The energy from lipid metabolism is usually stored in the specific cell and plays a critical role in stress response and survival (Figure 2(c)).

3.3. GO Analysis for Esophageal Cancer-Related Proteins. Furthermore, we applied enrichment analysis on three GO domains, including biological process cellular component and molecular function of differentially expressed proteins. The biological analysis reveals that most differentially expressed proteins participated in cell regulation and metabolism, such as mitochondrial dissemble and apoptosis, blood vessel regulation as well as downregulation of cell migration and motility. Our results indicate that in an EC patient, the cellular immune response was decreased, cell migration and motility were dysregulated, and the cell circle checkpoint was dysfunctional, which may lead to abnormal cell biological function (Figure 3(a)).

Cellular component analysis shows that MHC protein complex was upregulated, and the anchored component of
the membrane, actin cytoskeleton, and intrinsic component of plasma membrane were all downregulated, suggesting that extracellular interaction has been increased and the cytoskeleton remains in unstable status with the risk of collapsing (Figure 3(b)). Maintenance of membrane integrity was found in the molecular function analysis, which is the most important to facilitate extracellular activity such as binding to a specific protein or enzyme functioning (Figure 3(c)).

3.4. Protein Structure Domain and KEGG Analysis for Esophageal Cancer-Related Proteins. Proteins were mostly aggregated in the mucin (specific secretion protein in esophagus) biosynthesis pathway and downregulation of steroid biosynthesis. Furthermore, a recent study reported that mucin synthesis was strongly associated with esophageal cancer progress. Meanwhile, complement and coagulation cascades have been upregulated, meaning angiogenesis has been promoted and the blood circulation has been relatively blocked around the tumor mass. In addition, there were small amounts of proteins involved in autoimmune defense or immune function which may have the ability to increase the risk of self-immunological defense. Moreover, some amino acid metabolisms were downregulated, leading the corresponding protein dysfunction (Figure 4(a)).

Enrichment of protein structural domain analysis was applied in Figure 5(a), showing downregulation of protein belonging to the cellular matrix domain, such as myosin and myosin-like IQ motif; therefore, the cells will be more likely to collapse. Figure 5(b) shows that protein domains were more related to cellular recognition and facilitate immune function. MHC-class I/II was an important domain which can recognize antigen-antibody before the immunological reaction happens. Moreover, proteins facilitate cellular communication and the immune function domain was upregulated, indicating inflammation would more likely happen, as the theory of carcinogenesis could be recognized as chronic inflammation which happened in cell-induced subcellular changes and may have the risk of cancer occurrence (Figure 4(b)).

3.5. The Functional Clustering for Association between Specific Protein and Esophageal Cancer. In addition, we conducted a clustering analysis to find the correlation between the functions of differentially expressed proteins among the
comparison groups by the hierarchical clustering method. Humoral immune response (leukocyte, lymphocyte, and adaptive immune response) and regulation of receptor binding have been slightly upregulated in the biological process (p < 0.05). For cellular components, the MHC protein complex/MHC-1 protein complex has been upregulated in the control group meaning that T-cell immunity has been upregulated. The muscle-related cell components (myosin and myosin filament) have been significantly downregulated (p < 0.05), which would increase the risk of cellular scaffold collapsing. The specific peptide or structure protein binding function was significantly upregulated (including mRNA-3-UTR, peptide antigen, beta-2-microglobulin, and heparin). Moreover, the muscle-related binding activities such as structural constituent and actin filament were significantly downregulated with p < 0.01 (Figure 5(a)).

The FoxO pathway has been abnormally regulated which played a role in cell apoptosis, stress response, glucose metabolism, and posttranslational methylation. It may involve the carcinogenic process since the glucose is the primary energy for both normal cells and tumor cells and change the signaling pathways as well. Moreover, the autoimmune pathway which enhances cell immune function and the PI3K/Akt signaling pathway were both upregulated significantly serving as the key regulator in cellular survival and energy metabolism. The cholesterol metabolism pathway has been significantly downregulated, meaning the metabolic pathway for energy production shifted from lipid to glucose (p < 0.05), Protein-protein interaction network.

All differentially expressed protein database accession or sequence were compared against the STRING for protein-protein interaction. As shown in Figure 5(b), tetraspanin, myosin, and S-100 domains are in the top protein structure of differentially expressed protein, suggesting the association between specific protein and esophageal cancer.

4. Discussion

Esophageal carcinoma (EC) affects more than 450,000 worldwide in 2013, which is extremely aggressive and has a poor survival rate. The overall 5-year survival is lower than 10% worldwide. People with early diagnosis and locally advanced disease have better disease prognosis. The top five risk factors of EC are smoking, alcohol consumption, injury to the esophagus, chronic esophageal irritation, and obesity [21]. EC ranks fourth in China and occurs at a rate of 20-30 times than in the United States. The incidence in males is higher than females, aged over 40 years old. It is the second cancer from gastrointestinal cancers, only less than gastric cancer. Although endoscopic biopsy is a gold standard for diagnosis, however, it does not benefit patients in the early stage. Since early diagnosis is associated with disease outcomes, it is critical to find novel biomarkers with high sensitivity and specificity for detecting EC in the early stage. With advanced technology in biochemical, genetic, and imaging, the early screening and diagnostic method are critical. In this project, we performed a quantitative proteome analysis on four comparable groups of EC samples using TMT labeling and LC-
MS/MS to quantify the changes in the whole proteome of EC patients, followed by comprehensive bioinformatics analysis. The bioinformatics analysis showed that most of the differentially expressed proteins were involved in maintaining cell structure cytoskeletal integrity, intracellular trafficking, secretion, vesicular transport and signal transduction, metabolic process, stress-response, and single organism processes such as renal system and urogenital system development.

The cellular ability to prevent oxidative damage was decreased; superoxidized substances accumulate inside the cell, leading to cell apoptosis or death. Furthermore, we also found that a large amount of proteins participated in ion and inorganic hemostasis and immune regulation (adaptive immune response and lymphocyte-mediated immunity). Some cellular membrane proteins were identified; the ECM-receptor interaction pathway can influence tissue/organ morphogenesis [22]. ECM contains a large number of molecules, such as collagen, elastin, microfilaria proteins, and proteoglycans which can be involved in multiple pathways. ECM molecules also play a critical role in regulating cell adhesion, migration, differentiation, proliferation, and apoptosis [23]. The angiogenesis process is an important pathway in oncogenesis. In this study, we identified proteins regulating cytokines in the angiogenesis process, response to promoting and initiating the inflammatory process. This study results should be validated by an immunohistochemistry exam, such as ELISA to confirm if the protein is present in the EC cells.

5. Conclusion

This study applied TMT labeling proteomics followed by mass spectrometric analysis to study the EC cellular proteomic profile and identified proteins involving in oncogenesis-related pathways. TMT labeling proteomics is a new approach to engage cancer diagnosis and therapy, such as the identified protein could be used in biomarker designing, as well as the target drug therapy. Although the results of this study were promising, further study regarding sensitivity and specificity should be validated as well. Large cohort studies are also needed to evaluate the clinical significance.

Data Availability

The data of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declared that they have no conflicts of interest to this work.

Authors’ Contributions

Suofeng Sun, Huijuan Zhang, and Yu Wang contributed equally to this work.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (No. 81772670).

Supplementary Materials

Supplementary Table 1: the identified differentially expressed proteins and the annotation information for each 299 related protein and genes. (Supplementary Materials)

References

[1] A. Pennathur, M. K. Gibson, B. A. Jobe, and J. D. Luketich, “Oesophageal carcinoma,” Lancet, vol. 381, no. 9864, pp. 400–412, 2013.
[2] J. Zhang, D. Jiang, X. Li et al., “Establishment and characterization of esophageal squamous cell carcinoma patient-derived xenograft mouse models for preclinical drug discovery,” Laboratory Investigation, vol. 94, no. 8, pp. 917–926, 2014.
[3] D.-C. Lin, J.-J. Hsu, Y. Nagata et al., “Genomic and molecular characterization of esophageal squamous cell carcinoma,” Nature Genetics, vol. 46, no. 5, pp. 467–473, 2014.
[4] Y. Song, L. Li, Y. Ou et al., “Identification of genomic alterations in esophageal squamous cell cancer,” Nature, vol. 509, no. 7498, pp. 91–95, 2014.
[5] Y.-B. Gao, Z.-L. Chen, J.-G. Li et al., “Genetic landscape of esophageal squamous cell carcinoma,” Nature Genetics, vol. 46, no. 10, pp. 1097–1102, 2014.
[6] L. Zhang, Y. Zhou, C. Cheng et al., “Genomic analyses reveal mutational signatures and frequently altered genes in esophageal squamous cell carcinoma,” American Journal of Human Genetics, vol. 96, no. 4, pp. 597–611, 2015.
[7] J. Sun, X. Cai, M. M. H. Yung et al., “miR-137 mediates the functional link between c-Myc and EZH2 that regulates cisplatin resistance in ovarian cancer,” Oncogene, vol. 38, no. 4, pp. 564–580, 2019.
[8] The Cancer Genome Atlas Research Network, “Integrated genomic characterization of oesophageal carcinoma,” Nature, vol. 541, no. 7636, pp. 169–175, 2017.
[9] X. Zhan, X. Wang, and D. M. Desiderio, “Mass spectrometry analysis of nitrotyrosine-containing proteins,” Mass Spectrometry Reviews, vol. 34, no. 4, pp. 423–448, 2015.
[10] D. Cruz-Topete, B. Christensen, L. Sackmann-Sala, S. Okada, J. O. L. Jorgensen, and J. J. Kopchick, “Serum proteome changes in acromegalic patients following transsphenoidal surgery: novel biomarkers of disease activity,” European Journal of Endocrinology, vol. 164, no. 2, pp. 157–167, 2011.
[11] Z. Li, W. Zhou, Y. Zhang et al., “ERK regulates HIF1α-mediated platinum resistance by directly targeting PHD2 in ovarian cancer,” Clinical Cancer Research, vol. 25, no. 19, pp. 5947–5960, 2019.
[12] S. Qian, X. Zhan, M. Lu et al., “Quantitative analysis of ubiquitinated proteins in human pituitary and pituitary adenoma tissues,” Frontiers in Endocrinology, vol. 10, p. 328, 2019.
[13] Y. Meng, C. W. Chen, M. M. H. Yung et al., “DUOX1-medi- ated ROS production promotes cisplatin resistance by activating ATR-Chk1 pathway in ovarian cancer,” Cancer Letters, vol. 428, pp. 104–116, 2018.
[14] C. Cheng, Y. Zhou, H. Li et al., “Whole-genome sequencing reveals diverse models of structural variations in esophageal squamous cell carcinoma,” American Journal of Human Genetics, vol. 98, no. 2, pp. 256–274, 2016.
[15] H.-D. Qin, X.-Y. Liao, Y.-B. Chen et al., “Genomic characterization of esophageal squamous cell carcinoma reveals critical genes underlying tumorigenesis and poor prognosis,”
[16] J. Sun, X. Fu, Y. Wang et al., "Erianin inhibits the proliferation of T47D cells by inhibiting cell cycles, inducing apoptosis and suppressing migration," *American Journal of Translational Research*, vol. 8, no. 7, pp. 3077–3086, 2016.

[17] S. Sun, W. Wang, X. Luo et al., "Circular RNA circ-ADD3 inhibits hepatocellular carcinoma metastasis through facilitating EZH2 degradation via CDK1-mediated ubiquitination," *American Journal of Cancer Research*, vol. 9, no. 8, pp. 1695–1707, 2019.

[18] X. Zhan and D. M. Desiderio, "Nitroproteins from a human pituitary adenoma tissue discovered with a nitrotyrosine affinity column and tandem mass spectrometry," *Analytical Biochemistry*, vol. 354, no. 2, pp. 279–289, 2006.

[19] M. A. Collins, J. An, B. L. Hood, T. P. Conrads, and R. P. Bowser, "Label-free LC-MS/MS proteomic analysis of cerebrospinal fluid identifies protein/pathway alterations and candidate biomarkers for amyotrophic lateral sclerosis," *Journal of Proteome Research*, vol. 14, no. 11, pp. 4486–4501, 2015.

[20] Q. Chu, A. Rathore, J. K. Diedrich, C. J. Donaldson, J. R. Yates III, and A. Saghatelian, "Identification of microprotein-protein interactions via APEX tagging," *Biochemistry*, vol. 56, no. 26, pp. 3299–3306, 2017.

[21] J. Cox and M. Mann, "MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification," *Nature Biotechnology*, vol. 26, no. 12, pp. 1367–1372, 2008.

[22] N. Gershoni-Emek, A. Mazza, M. Chein et al., "Proteomic analysis of dynein-interacting proteins in amyotrophic lateral sclerosis synaptosomes reveals alterations in the RNA-binding protein staufen1," *Molecular & Cellular Proteomics*, vol. 15, no. 2, pp. 506–522, 2016.

[23] P. Barschke, P. Oeckl, P. Steinacker, A. Ludolph, and M. Otto, "Proteomic studies in the discovery of cerebrospinal fluid biomarkers for amyotrophic lateral sclerosis," *Expert Review of Proteomics*, vol. 14, no. 9, pp. 769–777, 2017.