Analysis of Transcriptional Factors and Regulation Networks in Laryngeal Squamous Cell Carcinoma Patients with Lymph Node Metastasis

Geng-Ming Cai,†‡ Dong-Hai Huang,† Yao-Zhang Dai,† Yong Liu,† Lei-Ming Pi,† Hao-Lei Tan,† Li-Jun Liu,† Yun-Xia Lv,† Chang-Yun Yu,† Ping-Qing Tan,† Yong-Quan Tian,† Georgia Z. Chen,§ and Xin Zhang*,†

†Department of Otolaryngology-Head Neck Surgery, Xiangya Hospital, Central South University, 87 Xiangya Road, Changsha, Hunan, P.R. China 410008
‡Department of Otolaryngology-Head Neck Surgery, First Affiliated Hospital of Quanzhou, Fujian Medical University, 248 Dong Street, Quanzhou, Fujian, P.R. China 362000
§Winship Cancer Institute, Emory University School of Medicine, Atlanta, Georgia 30322, United States

ABSTRACT: The present study was to identify and quantitate differentially expressed proteins in laryngeal squamous cell carcinoma (LSCC) tissues with or without lymph node metastasis and to explore transcriptional factors and regulation networks associated with the process. Tissue specimens were taken from 20 patients with LSCC, including 10 cases of LSCC without metastasis LSCC (N0) and 10 cases of LSCC with metastasis LSCC (Nx). Among the 643 unique proteins identified by using iTRAQ labeling and quantitative proteomic technology, 389 proteins showed an abundance change in LSCC (Nx) as compared to LSCC (N0). Cytoskeleton remodeling, cell adhesion, and immune response activation were found to be the main processes in LSCC metastasis. The construction of transcription regulation networks identified key transcription regulators for lymph node metastasis of LSCC, including Sp1, c-myc, and p53, which may affect LSCC metastasis through the epithelial—mesenchymal transition. Furthermore, our results suggest that ubiquitination may be a critical factor in the networks. The present study provides insights into transcriptional factors and regulation networks involved in LSCC metastasis, which may lead to new strategies for treatment of LSCC metastasis.

KEYWORDS: Laryngeal squamous cell carcinoma, metastasis, proteome regulation network, transcriptional factor, ubiquitination

INTRODUCTION

Squamous-cell carcinoma of the head and neck (HNSCC) is the sixth most frequently occurring malignancy worldwide, which represents a major international health problem. Annually, there are more than 500,000 new cases and 200,000 deaths around the world.† Despite improved detection and aggressive treatment approaches, including surgical resection and chemoradiation protocols, SCCHN is still a great threat to human life, and limited improvement in 5-year survival has been achieved during the last few decades. The poor outcome has mainly been attributed to local and distant lymph node metastasis.‡ Laryngeal squamous cell carcinoma (LSCC) is a very common malignant tumor in head and neck cancers.§ In China, the incidence of laryngeal carcinoma has been rising gradually, especially in the Northeast.§ Lymph node metastasis and exterior expansion of LSCC are the major reasons for incomplete incision of the tumor and postoperative failure. Gene expression profiling of LSCC has shown that genes with aberrant expressions possibly contributed to the pathogenesis of LSCC.¶ However, as proteins are the ultimate effector of phenotype and have been shown to not always correlate very well with mRNA levels,‖ it is necessary to study the changes of proteins related with the metastasis of LSCC. Using tissue samples from patients may be the most direct and persuasive way to find biomarkers and therapeutic targets for cancers by a proteomic approach.¶ The functional molecules in cells are proteins, proteome analysis is believed to have an advantage over cDNA microarray for clinical use.‖ In the present study, we employed isobaric tags for relative and absolute quantitation (iTRAQ) proteomic technology, off-line two-dimensional (2D) LC-MS/MS, and MetaCore bioinformatic analyses to identify and quantitate differentially expressed proteins in LSCC tissues with or without lymph node metastasis, and we explored transcriptional factors and regulation networks associated with the process.

MATERIALS AND METHODS

Tissues

From June 2009 to June 2010, tumor tissues were collected from 20 LSCC patients treated at Xiangya Hospital, Central South University. All diagnoses was based on biopsy before surgery, including 10 cases without metastasis (N0) and 10 cases with metastasis (N1).
Table 1. Baseline Characteristics of Patients

<table>
<thead>
<tr>
<th>subtype</th>
<th>grade</th>
<th>N</th>
<th>age range (years)</th>
<th>male N (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>LSCC (N0) supraglottis</td>
<td>total</td>
<td>3</td>
<td>56–62</td>
<td>3 (100)</td>
</tr>
<tr>
<td></td>
<td>I</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>II</td>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>LSCC (N0) glottis</td>
<td>total</td>
<td>7</td>
<td>55–70</td>
<td>7 (100)</td>
</tr>
<tr>
<td></td>
<td>I</td>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>II</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>III</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>LSCC (Nx) supraglottis</td>
<td>total</td>
<td>4</td>
<td>43–61</td>
<td>4 (100)</td>
</tr>
<tr>
<td></td>
<td>III</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>IV</td>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>LSCC (Nx) glottis</td>
<td>total</td>
<td>6</td>
<td>55–81</td>
<td>6 (100)</td>
</tr>
<tr>
<td></td>
<td>III</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>IV</td>
<td>4</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*LSCC (N0), laryngeal squamous cell carcinoma without metastasis; LSCC (Nx), laryngeal squamous cell carcinoma with metastasis.

with metastasis (Nx). Patient baseline characteristics are shown in Table 1. The study was approved by the Ethical Committee of Xiangya Hospital, Central South University. Informed consent was obtained from all patients.

Reagents

Trypsin was purchased from Sigma (St. Louis, MO, USA). The iTRAQ Reagents Kit was purchased from Applied Biosystems (Foster City, CA, USA). Formic acid was purchased from TEDIA (Fairfield, OH, USA). Acetone and trichloroacetic acid (TCA) were purchased from Sinopharm Chemical Reagent Co., Ltd. (Shanghai, China). All buffers were prepared with Milli-Q water (Millipore; Billerica, MA, USA).

Protein Extraction

Protein was extracted as previously described. Briefly, LSCC tissues were frozen in liquid nitrogen immediately after surgical removal and stored at −80 °C in the tissue bank of MMCI until sample preparation. Normal adjacent tissue (NAT) 5 mm from the tumor was collected in each patient. The tissue samples were categorized in four groups: NAT, LSCC (N0), LSCC (Nx), and the metastatic lymph node of LSCC (Nx) (LN). The tissue samples were cut to pieces (8 mm³), rinsed with PBS, and briefly disintegrated with a mortar and pestle in liquid nitrogen. A volume of 250 μL of lysis solution containing 0.5 M triethylammonium bicarbonate and 0.1% (w/v) SDS was added immediately after disintegration. The sample was homogenized with needle sonication (30 × 0.1 s, power set to 150 W) and further solubilized for 1 h at room temperature. The solubilized samples were centrifuged at 12,000 rpm for 15 min at 4 °C, and the total protein in the supernatant material was quantified using Bradford assay.

Protein Digestion and Labeling with iTRAQ Reagents

TCA—acetone precipitation was performed to purify the protein extracts as previously described. Briefly, a total of 100 μg of each sample was added with four times the volume of cold (−20 °C) acetone containing 12% (w/v) TCA, and the mixture was incubated overnight at −20 °C. Then the mixture was centrifuged at 12,000g at 4 °C for 30 min, the supernatant was removed without disturbing the visible pellet, and four times the sample volume of cold (−20 °C) acetone was added to the pellet. After incubation for 2 h at −20 °C, with intermittent vibration and centrifugation as mentioned above, the supernatant was removed and the pellet was air-dried. The protein pellets were then dissolved in the solution buffer, reduced, and blocked with cysteins, as described in the iTRAQ protocol. Each sample was digested with sequencing-grade modified trypsin with a protein–to-enzyme ratio of 20:1 at 37 °C overnight and then labeled with the iTRAQ tags as follows: NAT, iTRAQ 116 (IT116); LSCC (N0), iTRAQ 117 (IT117); LSCC (Nx), iTRAQ 113 (IT113); LN, iTRAQ 114 (IT114). The labeled digests were then mixed and dried using a rotary vacuum concentrator (Christ RVC 2-25; Osterode am Harz, Germany).

Peptide Fractionation and Off-line 2D LC-MS/MS Analysis

The combined peptide mixture was fractionated by strong cation exchange chromatography on a 20AD high-performance liquid chromatography (HPLC) system (Shimadzu, Kyoto, Japan). The sample was added to loading buffer (10 mM KH₂PO₄ in 25% acetonitrile, pH 2.6) and loaded onto the column. Separation was performed at a flow rate of 200 μL/min for 60 min. The absorbance at 214 and 280 nm was monitored, and SCX fractions were collected, dried down by the rotary vacuum concentrator, dissolved, and analyzed on a QSTAR XL system (Applied Biosystems, Carlsbad, CA, USA) interfaced with a 20AD HPLC system (Shimadzu). Peptides were separated on a Zorbax 300SB-C18 column (Agilent Technologies, Santa Clara, CA, USA). Survey scans were acquired from m/z 400–1800 with up to four precursors selected for MS/ MS from m/z 100–2000. Each SCX fraction was analyzed in duplicate.

Data Analysis

The MS/MS spectra were extracted and searched against the International Protein Index (IPI) database (version 3.45, HUMAN) using ProteinPilot software (version 3.0, revision 114732, Applied Biosystems). The reports included two types of scores for each protein: unused ProtScore and total ProtScore. The criteria were set as the following: unused ProtScore > 1.2 with at least one peptide with 95% confidence per repetition. The candidate proteins were examined in the Protein ID of the Protein Pilot software. Protein expression ratios were computed on the basis of the peak area ratios of the peptides accounting for the same protein. The bias correction algorithm was applied to correct for unequal mixing during the combination of the different labeled samples, based on the assumption that most proteins do not change in expression. All quant ratios (both the average ratio for proteins and the individual peptide ratios) were corrected for the bias.

Analysis of Regulation Networks

MetaCore (GeneGo; St. Joseph, MI, USA) is an integrated software suite for functional analysis of experimental data. To functionally annotate the differentially expressed proteins identified in this study, the proteins (≥2.0-fold change) were entered into GeneGo’s MetaCore for analysis. The biological process enrichment was analyzed according to Gene Ontology processes. The genes encoding expressed overabundant protein were used as the input list for generation of regulation networks using Transcription Regulation algorithm, which generated subnetworks centered on transcription factors. Subnetworks are ranked by a p-value and interpreted in terms of Gene Ontology.
RESULTS

As shown in Table 1, the clinical stages of LSCC (N0) patients were mostly stage I/II, compared to those of LSCC (Nx) patients. Data combination of the two MS repetitions resulted in identification of 19,882 distinct peptides corresponding to 853 unique proteins, of which 432 unique proteins were identified by at least two peptides, and 643 unique proteins were quantified by the iTRAQ labels (Table 2; Table S1 of the Supporting Information). The two MS repetitions shared an overlap of 74% of the total unique proteins of the combined data, and 643 proteins were identified with a global false discovery rate from f values of 1%. A total of 389 proteins showed an abundance change in LSCC (Nx) compared to LSCC (N0), with 244 proteins overexpressed (more than 1.2-fold) and 145 proteins underexpressed (less than 0.8-fold) (Table S2 of the Supporting Information). A total of 459 proteins showed an abundance change in LN compared to LSCC (N0), with 241 proteins overexpressed and 218 proteins underexpressed (Table S3 of the Supporting Information). To further interpret the differentially expressed proteins in metastasis, we used the MetaCore mapping tool to analyze and build the regulation networks related with the differentially expressed proteins. Multiple sets of proteins were mapped into different process pathways, including cytoskeleton remodeling, immune response, cell adhesion, and blood coagulation. As shown in Figure 1, the top score biological processes were associated with cytoskeleton remodeling, cell adhesion, and immune response.

To further discover unknown pathways and associations involved in metastasis, clusters of up- and down-regulated proteins were submitted to MetaCore and subjected to the Analyze Networks (AN) of Transcription Regulation and Transcriptional Factor algorithm (see Figure 2). The algorithms generated a set of regulation networks from the input list. The proteins were also submitted to the AN algorithm with default settings for

---

Table 2. Summary Statistics of Proteomic Results of International Protein Index Data

<table>
<thead>
<tr>
<th>unused (con) cutoff</th>
<th>proteins detected</th>
<th>proteins before grouping</th>
<th>distinct peptides</th>
<th>spectra identified</th>
<th>% total spectra</th>
</tr>
</thead>
<tbody>
<tr>
<td>&gt;2.0 (99)</td>
<td>640</td>
<td>2555</td>
<td>18074</td>
<td>40925</td>
<td>31.0</td>
</tr>
<tr>
<td>&gt;1.2 (95)</td>
<td>853</td>
<td>3263</td>
<td>19882</td>
<td>43673</td>
<td>33.1</td>
</tr>
<tr>
<td>&gt;0.47 (66)</td>
<td>1099</td>
<td>4530</td>
<td>23046</td>
<td>48178</td>
<td>36.5</td>
</tr>
</tbody>
</table>

This row shows the unused Protscore threshold applied in this paper and the corresponding statistical result.

---

Figure 1. Top biological processes of expressed proteins in laryngeal squamous cell carcinoma (LSCC) metastasis. (A) The gene content is aligned between all uploaded experiments. The intersection set of the experiments is defined as “common” and marked as a blue/white striped bar. The unique genes for the experiments are marked as colored bars. The genes from the “similar” set are present in all but one (any) file: LSCC (N0) (113:117) are marked as orange bars; metastatic lymph node of LSCC (Nx) (LN) compared with LSCC (N0) (114:117) are marked as blue bars. (B) GeneGo Pathway Maps. Sorting is done for the “common” set. Canonical pathway maps represent a set of about 650 signaling and metabolic maps covering human biology (signaling and metabolism) in a comprehensive way. All maps are drawn with GeneGo annotators and manually curated and edited.
generation of biological networks. It generated a set of networks from the input list. The top five score regulation networks were composed of 48—135 proteins altered in expression with \( p \)-value ranging from \( 8.07 \times 10^{-293} \) to \( 2.73 \times 10^{-101} \). These top score networks included the regulation initiated through activation of Sp1 (Figure 3), c-myc (Figure 4), p53 (Figure 5), androgen receptor, and ESR1 (nuclear). The proteins were connected through association with other objects, including transcription factors (e.g., YB-1, STAT1), binding proteins (e.g., ubiquitin), receptors (e.g., alpha6/beta4 integrin), receptor ligands (e.g., fibronectin), enzymes (e.g., SOD1), and GPCR (e.g., CXCR4), etc. All the interactions were based on MetaCore’s curated database. Therefore, any objects that contain the most connections to the root objects may represent key regulators in the clustered proteins. As shown in Figure 3, Sp1 plays a key role in orchestrating observed changes of protein abundance, by up- or down-regulating fibronectin, STAT1, and many other proteins associated with cytoskeleton remodeling and cell adhesion. In addition, c-myc (Figure 4) and p53 (Figure 5) were also noted as key network objects. The differentially expressed proteins belong to a diverse set of pathways and processes, which may help to illustrate the pathogenesis and the metastasis mechanisms of LSCC. The GO processes analysis indicated that most transcription factors were involved in cytoskeleton remodeling and cell adhesion. A set of proteins related with the cytoskeleton remodeling and cell adhesion process, such as keratin, collagen I, fibronectin, CD44, and alpha6/beta4 integrin, were altered. As shown in Figures 3—5, the ubiquitin system plays an important role in cytoskeleton remodeling and cell adhesion.

**DISCUSSION**

In the present study, we used iTraq-2DLC-MS/MS and MetaCore bioinformatic analyses to identify and quantitate differentially expressed proteins in LSCC tissues with or without
Figure 3. Sp1 (*Homo sapiens*), the first lowest p-value (by the number of pathways) transcription regulation network from active experiments. Thick cyan lines indicate the fragments of canonical pathways. Up-regulated genes are marked with red circles, and down-regulated ones with blue circles. The “checkerboard” color indicates mixed expression for the gene between files or between multiple tags for the same gene.

lymph node metastasis, and we explored transcriptional factors and regulation networks associated with the process. This is the first study to have investigated transcriptional regulation networks related to metastasis of LSCC at protein levels. The iTRAQ-2DLC-MS/MS and MetaCore bioinformatic analysis allowed the construction of complex networks that provide a new framework for understanding the molecular basis of LSCC metastasis. Using the Panther Database and Analyze Networks (AN), we found that the top score biological processes were associated with cytoskeleton remodeling, cell adhesion, and immune response.

In our previous studies, epithelial—mesenchymal transition (EMT) and anoikis-resistance were identified as two major features of metastatic LSCC cells. In the present study, transcriptional factors Sp1, c-myc, and p53 were found to play significant roles in regulating cytoskeleton remodeling and cell adhesion, which were critical for epithelial—mesenchymal transition (EMT). It is the first time that Sp1 has been found as a critical transcriptional factor in proteomics studies on head and neck tumor metastasis. Sp1 reportedly plays an important role in tumor growth and metastasis by regulating the activation of many genes involved in tumor growth, apoptosis, and angiogenesis. It has potential therapeutic use to prevent tumor dissemination.

Our findings in the present study demonstrate that Sp1 can directly or indirectly regulate the extracellular matrix (ECM) components, including keratins, collagens, fibronectin, and vitronectin, which were related with cytoskeleton remodeling and cell adhesion. Thus, Sp1 may be a key regulator of the pathologic process of EMT.

Our results show that c-myc and p53 also play important roles in lymph node metastasis of LSCC. C-myc was deregulated in most tumor types and stages, and it has been implicated in most aspects of tumor biology, including cellular proliferation, apoptosis, metastases, and angiogenesis. P53 was considered as a tumor inhibitor, but mutant p53 may promote tumor progression and induce metastasis. Our results suggest that c-myc and p53 protein may also affect the metastasis of LSCC by influencing proteins related with EMT.

Ubiquitination is reportedly related with EMT. In the present study, we found that up-regulation of ubiquitin by down-regulating USP5 could regulate proteins (such as STAT1, Sp1, c-myc, p53, NF-κB, PCNA, and CXCR4) related with cytoskeleton remodeling and cell adhesion in metastatic LSCC, which indicates that ubiquination may be critical for EMT and
metastasis of LSCC. Ubiquitination is a type of post-translational modification (PTM) that alters the properties of a protein by the addition of a modifying chemical group or another protein to one or more of its amino acid residues. Proteins are modified not only by the addition of small chemical moieties but also by covalent attachment of short modifier proteins such as ubiquitin, small ubiquitin-like modifier (SUMO), or other ubiquitin-like proteins (UBLs). Ubiquitination has different functions: receptor endocytosis, DNA-damage repair, ERAD (endoplasmic reticulum-associated degradation), regulation in cell-cycle, lysosomal degradation, kinase modification, protein degradation, and endocytosis. It also plays a role in activation of protein kinases and in signaling.26,27 The degradation accommodation of some transcription factors (such as p53, NF-κB, c-fos, c-jun, c-mos, and c-myc) was regulated by the ubiquitin-proteasome pathway (UPP), which is reportedly closely associated with the occurrence and development of malignant tumors.28 We propose that ubiquitination of transcriptional factors and many other proteins coordinates the EMT of LSCC, which ultimately leads to tumor metastasis.

Most of the results in the present study are in agreement with our previous studies using cDNA microarrays. However, many proteins identified in our previous studies6–9 (such as MMPs and E-cadherin) were not identified in the present study. This may be due to the following reasons: (1) Proteins are not always correlated very well with mRNA levels.10 (2) Some proteins were inactivated or degraded with time. (3) Ubiquitination might alter the properties of some proteins and, thus, lead to different results between the proteomics analysis and the DNA microarray analysis. (4) While we used cell lines for cDNA microarrays in previous studies, we used tissue samples in the present study, which may lead to more realistic results.

In conclusion, we identified and quantitated differentially expressed proteins in LSCC tissues with or without lymph node metastasis, and we explored transcriptional factors and regulation networks associated with the process, using iTRAQ-2DLC-MS/MS and MetaCore bioinformatic analyses. Cytoskeleton remodeling, cell adhesion and immune response activation were found to be the main processes in LSCC metastasis. The construction of transcription regulation networks identified key transcription regulators for lymph node metastasis of LSCC, including Sp1, c-myc, and p53, which may affect LSCC metastasis through EMT. In addition, our results suggest that ubiquitination may be a critical factor in the networks. The present study provides insights into transcriptional factors and regulation networks involved in LSCC metastasis, which may lead to new strategies for treatment of LSCC metastasis.

Figure 4. C-myc (Homo sapiens), the second lowest p-value (by the number of pathways) transcription regulation network from active experiments. Thick cyan lines indicate the fragments of canonical pathways. Up-regulated genes are marked with red circles, and down-regulated ones with blue circles. The "checkerboard" color indicates mixed expression for the gene between files or between multiple tags for the same gene.
ASSOCIATED CONTENT

Supporting Information
Table of proteins by category. This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author
*Telephone: 86-731-84327623. E-mail: xinzh99@yahoo.com.

ACKNOWLEDGMENT

This work was supported by the National Natural Science Foundation of China (Grants #81071757 and #30901664), Key Project of Science and Technology Office of Hunan Province (Grants #2010TP4012-1), the Research Fund for the Doctoral Program of Higher Education of China (201001621), and the Graduate Degree Thesis Innovation Foundation of Hunan province (CX2011B059). We thank the Department of Chemistry and the Institute of Biomedical Science of Fudan University for technical support.

REFERENCES


