Transcriptional regulatory networks in human lung adenocarcinoma

XIANGRUI MENG1*, PENG LU2*, HUA BAI3, PENG XIAO3 and QINGXIA FAN1

1Department of Oncology, The First Affiliated Hospital of Zhengzhou University; Departments of 2Oncological Surgery and 3Medical Oncology, People’s Hospital of Zhengzhou, Zhengzhou, P.R. China

Received May 5, 2012; Accepted July 27, 2012

DOI: 10.3892/mmr.2012.1034

Abstract. Lung adenocarcinoma (AC) is the most common histological subtype of lung cancer worldwide and its absolute incidence is increasing markedly. Transcriptional regulation is one of the most fundamental processes in lung AC development. However, high-throughput functional analyses of multiple transcription factors and their target genes in lung AC are rare. Thus, the objective of our study was to interpret the mechanisms of human AC through the regulatory network using the GSE2514 microarray data. Our results identified the genes peroxisome proliferator activated receptor-γ (PPARG), CCAAT/enhancer binding protein β (CEBPB), ets variant 4 (ETV4), Friend leukemia virus integration 1 (FLI1), T-cell acute lymphocytic leukemia 1 (TAL1) and nuclear factor of kappa light polypeptide gene enhancer in B-cells 1 (NFκB1) as hub nodes in the transcriptome network. Among these genes, it appears that: PPARG promotes the PPAR signaling pathway via the upregulation of lipoprotein lipase (LPL) expression, but suppresses the cell cycle pathway via downregulation of growth arrest and DNA-damage-inducible, γ (GADD45G) expression; ETV4 stimulates matrix metallopeptidase 9 (MMP9) expression to induce the bladder cancer pathway; FLI upregulates transforming growth factor, β receptor II (TGFBR2) expression to activate TGF-β signaling and upregulates cyclin D3 (CCND3) expression to promote the cell cycle pathway; NFκB1 upregulates interleukin 1, β (IL-1B) expression and initiates the prostate cancer pathway; CEBPβ upregulates IL-6 expression and promotes pathways in cancer; and TAL1 promotes kinase insert domain receptor (KDR) expression to promote the TGF-β signaling pathway. This transcriptional regulation analysis may provide an improved understanding of the molecular mechanisms and potential therapeutic targets in the treatment of lung AC.

Introduction

Lung carcinoma is the most common cancer in the world and the leading cause of cancer-related mortality, with over one million cases annually (1). Lung carcinomas are usually classified as small-cell lung carcinomas (SCLCs) or non-small-cell lung carcinomas (NSCLCs). NSCLCs are histopathologically and clinically distinct from SCLCs and are further subcategorized as adenocarcinomas (ACs), squamous cell carcinomas or large-cell carcinomas, of which ACs are the predominant form (2).

Several molecular changes are characteristic of lung AC. These include mutations in the tyrosine kinase domain of the epidermal growth factor receptor (EGFR), v-Ki-ras Kirsten rat sarcoma viral oncogene homolog (KRAS) (3,4), tumor protein 53 (TP53) (5), serine/threonine kinase 11 (STK11) (6) and cyclin-dependent kinase inhibitor 2A (CDKN2A) (7) genes. Among them, somatic-activating EGFR mutations, particularly deletions in exon 19 and L858R point mutations in exon 21, may activate the gp130/JAK/STAT3 pathway by means of interleukin 6 (IL-6) upregulation in primary human lung AC, thereby promoting cell cycle progression, cell growth and tumorigenesis (8). Further studies have revealed that thyroid transcription factor-1 (TTF-1) expression is positively associated with EGFR mutations in lung AC (9). These results indicate that transcriptional regulation is a fundamental process in lung AC development.

However, high-throughput functional analyses of multiple transcription factors (TFs) and their target genes in lung AC are rare. Therefore, the objective of this study was to identify potential transcription regulation correlations between TFs and differentially expressed genes (DEGs) in lung AC using microarray data and transcriptional network analysis. In addition, the underlying molecular mechanisms were explored by KEGG pathway enrichment.

Materials and methods

Affymetrix microarray data analysis. The transcription profiles of human AC GSE2514 (10) were obtained from a public functional genomics data repository GEO.
The 10 most enriched TFs, target genes and their underlying networks of molecular interactions include FLI1 (http://www.ncbi.nlm.nih.gov/geo/) and are based on the Affymetrix GPL8300 platform data (Affymetrix Human Genome U95 Version 2 Array) (11). Only 20 AC chips and 19 control chips were useable. Each pair of samples (one derived from cancer cells and the other from normal cells) represented a single patient with AC.

All patients participating in this study were enrolled in a local Colorado Multiple Institutional Review Board (COMIRB)-approved protocol for use of remnant tissue with anonymization and analysis of specimens and clinical data. Informed consent was obtained from all the patients. All but one of the patients had a history of smoking. Patients ranged in age from 45 to 73 years of age. Tumors from five males and five females were used in the study. All specimens for microarray analysis were obtained at surgery with nine patients undergoing lobectomy and one wedge resection. Specimens were examined immediately following removal from the patient and grossly visible solid tumor tissue was snap-frozen for RNA extraction. The tumors were all invasive ACs, but five specimens exhibited evidence of bronchoalveolar differentiation at the edge of tumor nests. Most tumors were low to intermediate grade and low stage, although two stage III tumors were included in the analysis.

The limma method (12) was used to identify DEGs. The original expression datasets from all conditions were extracted into expression estimates using the robust multiarray average (RMA) method (13) with the default settings implemented in Bioconductor and the linear model was constructed. Only the DEGs with a fold-change >1.5 and p-value <0.05 were selected. Regulatory relationships having an absolute PCC >0.7 were considered to be significant.

Using the limma package, a total of 915 DEGs with p<0.05 and fold-change >1.5 were selected. Regulatory relationships with a PCC >0.7 were considered to be significant. A regulatory network for human AC comprising TFs and their target genes was constructed (Fig. 1). In this network, p53 downregulates transforming growth factor, beta (TGFBR2) and cyclin D3 (CCND3) expression; FLI1 can upregulate transforming growth factor, beta receptor II (TGFBR2) and cyclin D3 (CCND3) expression; ETV4 can downregulate TGFBR2 expression, but stimulate matrix metalloproteinase 9 (MMP9) expression; NFkB1 can upregulate IL6 expression; CEBPB can upregulate IL6 expression; TAL1 could promote kinase insert domain receptor (KDR) expression.

To further investigate the regulatory relationships between TFs and significant pathways, we mapped target genes in the network to pathways and created a regulatory network comprising TFs and pathways.

**Results**

**Microarray data and regulatory network analysis.** Using the limma package, a total of 915 DEGs with p<0.05 and fold-change >1.5 were selected. Regulatory relationships with a PCC >0.7 were considered to be significant. A regulatory network for human AC comprising TFs and their target genes was constructed (Fig. 1). In this network, p53 downregulates transforming growth factor, beta (TGFBR2) and cyclin D3 (CCND3) expression; FLI1 can upregulate transforming growth factor, beta receptor II (TGFBR2) and cyclin D3 (CCND3) expression; ETV4 can downregulate TGFBR2 expression, but stimulate matrix metalloproteinase 9 (MMP9) expression; NFkB1 can upregulate IL6 expression; CEBPB can upregulate IL6 expression; TAL1 could promote kinase insert domain receptor (KDR) expression.

**Significant pathways and TF pathway regulatory network analysis.** Using the KEGG pathways to describe the function of the regulatory network, several KEGG pathways among the pathways in the regulatory network were revealed to be enriched, including pathways in cancer (hsa05200), the PPAR signaling pathway (hsa03320), cell cycle (hsa04110) and the TGF-β signaling pathway (hsa04350). The 10 most enriched KEGG pathways are listed in Table I.

To further investigate the regulatory relationships between TFs and pathways, we mapped DEGs to significant pathways and obtained a regulatory network comprising TFs and pathways (Fig. 2). In the network, 9 TFs regulated 6 pathways. The network indicates that the PPAR pathway is upregulated by PPAR and CEBPB; the cell cycle is upregulated by FLI and STAT5B, but downregulated by PPAR; the TGF-β signaling pathway is upregulated by FLI and TAL1, but downregulated by ETV4; ETV4 promotes the bladder cancer pathway; NFkB1 promotes the prostate cancer pathway. Pathways in cancer may be upregulated by CEBPD and CEBPB.

**Discussion**

We investigated the comprehensive regulatory network of lung AC comprising TFs, target genes and their underlying
molecular pathways. In our transcriptosome network, the genes PPARG, CEBPB, ETV4, FLI1, TAL1, and NFκB1 are hub nodes. PPARG may promote the PPAR signaling pathway via upregulation of LPL expression, but suppress the cell cycle pathway via downregulation of GADD45G expression; ETV4 can stimulate MMP9 expression to induce the bladder cancer pathway; FLI can upregulate TGFβ2 expression to activate TGF-β signaling, and upregulate CCND3 expression to promote the cell cycle pathway; NFκB1 can upregulate ILIB expression and initiate the prostate cancer pathway; CEBPB
Table I. Pathway significance analysis.

<table>
<thead>
<tr>
<th>Term</th>
<th>Description</th>
<th>Count</th>
<th>p-value</th>
<th>FDR</th>
</tr>
</thead>
<tbody>
<tr>
<td>hsa05200</td>
<td>Pathways in cancer</td>
<td>16</td>
<td>1.01e-9</td>
<td>1.05e-6</td>
</tr>
<tr>
<td>hsa05220</td>
<td>Chronic myeloid leukemia</td>
<td>7</td>
<td>9.72e-6</td>
<td>0.010121</td>
</tr>
<tr>
<td>hsa05216</td>
<td>Thyroid cancer</td>
<td>5</td>
<td>3.95e-4</td>
<td>0.041159</td>
</tr>
<tr>
<td>hsa05210</td>
<td>Colorectal cancer</td>
<td>6</td>
<td>2.40e-4</td>
<td>0.249483</td>
</tr>
<tr>
<td>hsa03320</td>
<td>PPAR signaling pathway</td>
<td>5</td>
<td>0.001184</td>
<td>1.225733</td>
</tr>
<tr>
<td>hsa04110</td>
<td>Cell cycle</td>
<td>6</td>
<td>0.001489</td>
<td>1.538652</td>
</tr>
<tr>
<td>hsa05222</td>
<td>Small cell lung cancer</td>
<td>5</td>
<td>0.002457</td>
<td>2.527709</td>
</tr>
<tr>
<td>hsa04350</td>
<td>TGF-β signaling pathway</td>
<td>5</td>
<td>0.002793</td>
<td>2.86916</td>
</tr>
<tr>
<td>hsa05219</td>
<td>Bladder cancer</td>
<td>4</td>
<td>0.002854</td>
<td>2.931146</td>
</tr>
<tr>
<td>hsa05215</td>
<td>Prostate cancer</td>
<td>5</td>
<td>0.003034</td>
<td>3.113205</td>
</tr>
</tbody>
</table>

Term represents the pathway ID. Description is the pathway symbol. Count is the number of enrichment pathways. The p-value is the probability of obtaining a test statistic; the smaller the p-value, the more enriched the pathway. False discovery rate (FDR) control is a statistical method used in multiple hypothesis testing to correct for multiple comparisons; the smaller the FDR, the higher the correctness.

can upregulate IL-6 expression and promote pathways in cancer; TAL1 could upregulate KDR expression to promote the TGF-β signaling pathway.

PPARG is a ligand-activated TF, whose activation has been implicated in the pathology of numerous diseases, including lung AC. High PPARG expression levels have been detected in lung AC patients (18) and PPARG-positive has been identified more frequently in well-differentiated AC cases than in moderately and poorly differentiated ones (19). The treatment of lung AC cells with PPARG ligands induces a dose-dependent inhibition of lung AC cell growth, that is, a cell cycle arrest at G0/G1 (18,20). GADD45 is a cell cycle-regulated nuclear protein that reaches maximal levels in the G1 phase of the cycle (21). Through its association with Cdc2, GADD45 disrupts the interactions of Cdc2 with cyclin B1 and, thus, may induce G2/M arrest (22). Previous studies indicate that the activation of PPARG may lead to apoptosis and growth arrest, at least in part, by inducing the Oct-1-mediated transcription of GADD45 (23). However, GADD45G methylation is significantly frequent in lung cancer patients and results in lung tumorigenesis (24). Similarly, we also found that GADD45G was downregulated by PPARG in lung AC. Moreover, PPARG is a transcriptional factor which mediates pleiotropic effects, including the regulation of genes involved in lipid metabolism, such as LPL, which is a component of the PPAR signaling pathway. PPARG and the 9-cis retinoic acid receptor (RXR) heterodimers bind to the promoter sequence (-169 TGCCCTTTCCCCC -157) of the LPL gene and thus promote the transcriptional activation of the LPL gene (25,26). Higher LPL levels accelerate the growth of cancer cells (27) and predict shorter NSCLC patient survival times (28).

ETV4 (also known as E1AF), which binds to the enhancer elements of the adenovirus type 5 E1A gene, is a TF of the ets oncogene family (29). ETV4 is expressed in NSCLC cells. Significantly, ETV4-transfected NSCLC cells show a 2-fold increase in cell motility and invasion compared with parental and vector-transfected control cells (30). ETV4 is able to upregulate multiple MMP genes that contribute to the malignant phenotype of cancer cells by inducing invasive and metastatic activities (31). A previous study has shown that the ERK-ETV4-MMP1 axis is upregulated in esophageal AC cells and is a potentially significant driver of the metastatic progression of esophageal ACs (32). In the current study, our results revealed that MMP9 expression was upregulated by ETV4, which may be involved in the development and invasion of lung AC.

FLI1 is a TF of the ETS family, defined by a highly conserved DNA-binding domain (33). Its clinical role is most evident in human Ewing's sarcoma in which it is fused with EWS. It has been shown that transduction of the gene EWSR1-FLI1 transforms NIH3T3 cells and that mutants containing a deletion in either the EWS domain or the DNA-binding domain in FLI1 lose this ability (34,35). Significantly, EWS-FLI1 binds to the second positive regulatory element of the TGF-b RI promoter, a putative tumor suppressor gene in the TGF-β signaling pathway, and suppresses transcription of the TGF-b RII gene at the mRNA and protein levels (36). Antisense to EWSR1-FLI1 in ES cell lines positive for this gene fusion restores TGF-β RII expression (37) which blocks tumorigenicity. However, the expression of FLI1 in lung cancer cells is not well characterized. FLI1 expression is usually detected at lower levels in certain non-hematopoietic tissues, including the lung. Weak nuclear immunoreactivity is observed in lung AC (38). Therefore, TGF-β RII expression may be upregulated to block tumorigenicity. In addition, levels of CCND3 and FLI1 are also positively correlated, since FLI1 maintains high levels of CCND3 in erythroblasts, thereby promoting proliferation over differentiation (39). Our results also suggest that FLI1 upregulates CCND3 expression. Therefore, we also suggest that CCND3 expression is associated with the proliferation of lung AC cells.

NFκB1 is a subunit of the TF NFκB which is derived by proteolytic cleavage from the N-terminal of a 105-kDa precursor protein (40). Activated NFκB stimulates the expression of genes involved in a wide variety of biological functions; for example, NFκB target genes, including chemokine (C-C motif) ligand 19 (CCL19), CCL21, chemokine (C-X-C motif) ligand 12 (CXCL12), CXCL13 and B-cell-activating factor of the tumour-necrosis-factor family (BAFF), were markedly
upregulated in pancreatic ductal AC cell lines (41). In addition, the -300 region of the IL-1B promoter contains a functional NFkB binding site composed of the decoamer sequence 5’-GGGAAAATCC-3’ (42). Therefore, activated Nfkb may also induce the expression of IL-1B, which is an important cytokine involved in inflammatory and immune diseases, including various cancers (43).

CEBPβ is a bZIP TF which may bind as a homodimer to certain DNA regulatory regions or form heterodimers with the related proteins CEBP-α, CEBP-β, and CEBP-γ. CEBPB protein is important in the regulation of genes involved in immune and inflammatory responses. It has been shown to activate the IL-6 promoter and induce elevated IL-6 expression levels (44), which are frequently observed in human lung ACs (45). A further study has revealed that the regulation of the IL-6 promoter by CEBPB is completely dependent upon co-operative functions with Nfkb in autocrine human prostate cancer cells (46), suggesting a model in which the bZIP protein primarily functions to augment the activity of NF-κB.

The TALI (or SCL) gene encodes a basic helix-loop-helix (bHLH) TF that has been demonstrated to be significant in hematopoiesis and vasculogenesis (47). Previous studies have revealed that the expression of the SCL interrupting locus gene is increased in lung ACs and promotes metastatic spread, which may result from the controlling effect on its downstream gene, SCL (48). KDR (VERFR2) is one of the two VEGF receptors which is critical for mediating angiogenic endothelial cell responses via the VEGF pathway (49). Higher levels of soluble VEGFR2 have been observed in NSCLC patients compared with healthy controls (50). E-box protein E2-2 blocks endothelial cell activation via perturbation of VEGFR2 promoter activity (51). However, TALI/SCL relieves the E2-2-mediated repression of VEGFR2 reporter activity in endothelial cells by interacting with certain DNA sequences of E2-2 (52).

A basic understanding of the mechanisms underlying AC is valuable. A deeper understanding of TFS and their regulated genes remains an area of intense study. Our present findings shed new light on the complex interacting mechanisms of TFS and their regulated genes in lung AC. These results may provide potential therapeutic targets for lung AC treatment.

References