A bstract. – OBJECTIVES: We aimed to explore the DNA methylation difference between lung cancer samples and non-cancer lung samples, and to investigate the role of DNA methylation in the mechanism of lung cancer development. Besides, we analyzed the transcriptional regulation network of DNA methylation and the miRNAs regulated by DNA methylation. This study provides a framework for DNA methylation in other tumors or diseases.

MATERIALS AND METHODS: DNA methylation and gene expression profiles were obtained from Gene Expression Omnibus. Firstly, we identified differentially methylated genes (DMGs) by Student’s t-test. Then we detected the biological processes and pathways changed in lung cancer by Gene Ontology (GO) and KEGG pathway enrichment analysis. The transcriptional factors in differential genes were identified and the microRNAs regulated by them were also obtained in TransmiR.

RESULTS: We obtained 108 DMGs between lung cancer samples and non-cancer samples. Besides development related biological processes and pathways were dramatically disordered. For the DMGs, we identified 11 transcriptional factors regulating them. Moreover, we screened out 21 relationships between DMGs and their transcriptional targets. Five microRNAs are reported to be regulated by DNA methylation genes. Finally a regulation network of DNA methylation was constructed.

CONCLUSIONS: DNA methylation participates in carcinogenesis at the transcriptional and post-transcriptional level. Aberrant DNA methylation will prevent its binding with the upstream regulatory proteins, inhibit the function of downstream target genes and regulate the expression of downstream miRNA, and consequently affect cell development, immuneresponse and apoptosis.

Key Words: Lung cancer, Transcription factor, MicroRNA, Pathway, Gene Ontology, DNA methylation.

Introduction

Lung cancer is the uncontrolled growth of aberrant cells in one or both lungs. Survival of lung cancer depends on stage, overall health, and other factors. Overall, 15% of people in the United States diagnosed with lung cancer survive five years after the diagnosis. Worldwide, lung cancer is the most common cause of cancer-related death in men and women, and is responsible for 1.38 million deaths annually in 2008. Therefore, the research and treatment of lung cancer is of great significance to human health.

Aberrant DNA hypermethylation has been implicated as a component of an epigenetic mechanism that silences genes in lung cancer. Park et al. investigate the significance of DNA methylation in SLC5A8 (salute carrier family 5 iodide transporter member 8) expression in lung cancer cell lines and tissues and conclude that DNA methylation in the SLC5A8 promoter region may suppress the expression of SLC5A8 in lung tumor. Tekpli et al. show DNA methylation at promoter regions of interleukin (IL) 1B, IL-6, and IL-8 which are involved in the inflammatory response during lung cancer development. Carvalho et al. further report that more than 75% of the hypermethylated genes belong to the class of transcriptional regulators. Aberrant DNA methylation of OLIG1 (oligodendrocyte transcription factor 1) is regarded as a novel prognostic factor in NSCLC (non-small cell lung cancer). Hypomethylation of intragenic LINE-1 (long interspersed nuclear element-1) represses transcription in cancer cells through AGO2 (argonaute 2). LDHC (lactate dehydrogenase C) gene expression in cancer cells is regulated by Sp1, CREB, and CpG island methylation. In addition, it is reported that miRNA expression also can be regulated by DNA methylation. Let-7a-3 belongs to the archetypal let-7 miRNA gene family and is found to be methylated by the DNA methyltransferases DNMT1 and DNMT3B, resulting enhanced lung tumor phenotypes.

Corresponding Author: Hezhong Chen, MD; e-mail: chenhezhongchh@hotmail.com; Zhiyun Xu, MD; e-mail: xuzhiyunchh@hotmail.com
Gene Ontology (GO) and gene-set-enrichment analysis approach (GSEA) are utilized to characterize the underlying function of identified genes in lung cancer. Genes that differentially expressed between normal lung tissue and cancer show enrichment in gene ontology (GO) terms associated with mitosis and proliferation. The Fas signaling pathway and the antigen processing and presentation pathway are most significantly related with lung cancer susceptibility by GSEA. DNA copy number aberrations in SCLC are significantly activated by the focal adhesion pathway. MicroRNA signatures in tumor tissue are related to pathway “angiogenesis” in NSCLC evaluated by GSEA.

In this study, we aim to integrate information of functional annotation, transcription factors and microRNAs to elaborate the role of DNA methylation in the mechanism of lung cancer development.

Materials and Methods

DNA Methylation Data
The DNA Methylation data we used was downloaded from National Center for Biotechnology Gene Expression Omnibus (NCBI GEO, http://www.ncbi.nlm.nih.gov/geo/) under accession number GSE32867, which detects the DNA methylation status of 27,578 CpG (cytosine-phosphate-guanine) dinucleotides CpG sites on 28 lung cancer samples and 27 adjacent non-lung cancer tissue samples. The number of probes is 27,578, corresponding to 14,495 protein-coding genes.

Gene Expression Data
The gene expression data we used was downloaded from NCBI GEO (http://www.ncbi.nlm.nih.gov/geo/) under accession number GSE32863, which is based on the platform of Illumina HumanWG-6 v3.0 expression beadchip. This profile is compromised by 58 lung cancer samples and 58 adjacent non-cancer samples. The number of probes is 48,803, corresponding to 25,188 protein-coding genes.

Date Preprocessing
For DNA methylation data, when multiple probes were mapped to the same gene, the level of DNA methylation β for this gene was represented by the mean value of all the probes. For gene expression, when one probe was mapped to multiple genes, these probes were removed. For multiple probes that were mapped to the same gene, expression level of this gene was represented by the median expression value of all the probes.

Differential Analysis
By comparing the methylation pattern of lung cancer samples and non-cancer lung samples, we identified differentially methylated genes (DMGs). The statistical method used was Student’s t-test (FDR: False Discovery Rate = 0.01), with the threshold of differential methylation level at 10%. In this manner, we identified genes and methylation loci with significant DNA methylation under disease status.

Gene Ontology and Pathway Analysis
In order to analyze the biological function that is regulated by methylation, we performed Gene Ontology (biological processes, BP) analysis for DMGs. These genes were also analyzed from the perspective of KEGG (Kyoto Encyclopedia of Genes and Genomes) pathway. Both processes utilize DAVID (Database for Annotation, Visualization and Integration Discover), the parameter “count” is set to 2 and corrected p-value is set to 0.01 for multiple testing.

Results

DNA Methylation Analysis
After mapping the DNA methylation probes to the genome, 72.3% of the methylated loci were less than 500 bp (base pair) from TSS (Transcription Start Site) of genes and 92% of the methylated loci were less than 1,000 bp from TSS of genes (Figure 1). Significant methylation difference was identified in 108 genes between lung cancer and non-cancer samples with multiple testing corrected FDR < 0.01. As shown in Figure 2, 84 genes had high methylation level near their TSS. The DNA methylation level was low for 14 genes in tumor samples.

Functional Analysis of Differentially Methylated Genes
The differentially methylated genes (DMGs) were annotated to biological processes and pathways. As shown in Table I, 108 DMGs were significantly enriched in 19 biological processes (including apoptosis process, the cell death process, ion transfer and the immune response regulation process) or four pathways (Calcium
signaling pathway, Cardiac muscle contraction, Alzheimer’s disease, and Huntington’s disease).

**Transcriptional Regulation of Methylation**

In mammals, DNA methylation is widespread in the genome epigenetic modification. A methyl (-CH3), which can block the binding of the protein, is added to cytosine of CG diad on DNA by DNA methylation transferase enzyme. It is reported that DNA methylation near TSS can regulate gene expression. Thus, aberrant DNA methylation has been found in embryonic development and genomic imprinting19,20. In this study, we obtained gene expression of 47 DMGs between lung samples and adjacent non-cancer lung samples. Among these 47 genes, we further identified 18 genes that differentially expressed in lung cancer based on KS (Kolmogorov-Smirnov)-test method (p < 0.05).

**Regulation Network of DNA Methylation**

The TRANSFAC (Transcription Factor database) provides abundant data on eukaryotic transcription factors, their experimentally-proven binding sites, consensus binding sequences and regulated genes21. For the DMGs, we picked out the transcriptional factors and extracted their target genes based on TRANSFAC version 11.4. Finally we obtained 21 relationships between DMGs and their corresponding transcriptional targets, involving five DMGs. Moreover, we identified the transcriptional factors that target DMGs. As result, 13 relationships between upstream transcriptional factors and DMGs were found to regulate aberrant DNA methylation, involving 11 transcriptional factors. Since the initial transcript of miRNA is unknown, it is not possible to directly determine its upstream proteins that regulate the transcription process. TransmiR (transcription factor-microRNA regulation) database stores the relationship between miRNA and transcriptional factors by literature mining22. From TransmiR database we identified the transcriptional relationship between two DNA hypermethylated genes and six miRNAs (pre-mir-302a/b/c/d, pre-mir-367, and let-7a). We constructed a lung cancer-related methylation regulatory network by integrating three types of DNA methylation regulation. As shown in Figure 3, as epigenetic modification markers on the genome, DNA methylation participates in carcinogenesis at the transcriptional and post-transcriptional level. In the process of carcinogenesis aberrant DNA methylation will prevent its binding with the upstream regulatory proteins and in-
Table I. GO and KEGG annotation of differentially methylated genes (DMGs).

<table>
<thead>
<tr>
<th>Category</th>
<th>GO ID</th>
<th>p value</th>
<th>Term</th>
</tr>
</thead>
<tbody>
<tr>
<td>GO TERM_BP</td>
<td>GO:0065008</td>
<td>3.18E-04</td>
<td>Regulation of biological quality</td>
</tr>
<tr>
<td>GO TERM_BP</td>
<td>GO:006816</td>
<td>7.82E-04</td>
<td>Calcium ion transport</td>
</tr>
<tr>
<td>GO TERM_BP</td>
<td>GO:0015674</td>
<td>0.002029</td>
<td>di-, tri-valent inorganic cation transport</td>
</tr>
<tr>
<td>GO TERM_BP</td>
<td>GO:0003013</td>
<td>0.00258</td>
<td>Circulatory system process</td>
</tr>
<tr>
<td>GO TERM_BP</td>
<td>GO:0008015</td>
<td>0.00258</td>
<td>Blood circulation</td>
</tr>
<tr>
<td>GO TERM_BP</td>
<td>GO:0002026</td>
<td>0.00294</td>
<td>Regulation of the force of heart contractio</td>
</tr>
<tr>
<td>GO TERM_BP</td>
<td>GO:0006915</td>
<td>0.003402</td>
<td>Apoptosis</td>
</tr>
<tr>
<td>GO TERM_BP</td>
<td>GO:0012501</td>
<td>0.003752</td>
<td>Programmed cell death</td>
</tr>
<tr>
<td>GO TERM_BP</td>
<td>GO:0006811</td>
<td>0.005588</td>
<td>Ion transport</td>
</tr>
<tr>
<td>GO TERM_BP</td>
<td>GO:0048518</td>
<td>0.00542</td>
<td>Positive regulation of biological process</td>
</tr>
<tr>
<td>GO TERM_BP</td>
<td>GO:0050776</td>
<td>0.005901</td>
<td>Regulation of immune response</td>
</tr>
<tr>
<td>GO TERM_BP</td>
<td>GO:0048584</td>
<td>0.00706</td>
<td>Positive regulation of response to stimulus</td>
</tr>
<tr>
<td>GO TERM_BP</td>
<td>GO:0002684</td>
<td>0.007309</td>
<td>Positive regulation of immune system process</td>
</tr>
<tr>
<td>GO TERM_BP</td>
<td>GO:0006959</td>
<td>0.007627</td>
<td>Humoral immune response</td>
</tr>
<tr>
<td>GO TERM_BP</td>
<td>GO:0032501</td>
<td>0.007766</td>
<td>Multicellular organisal process</td>
</tr>
<tr>
<td>GO TERM_BP</td>
<td>GO:0051240</td>
<td>0.00895</td>
<td>Positive regulation of multicellular organi</td>
</tr>
<tr>
<td>GO TERM_BP</td>
<td>GO:0007188</td>
<td>0.009322</td>
<td>G-protein signaling, coupled to cAMP nucleo</td>
</tr>
<tr>
<td>GO TERM_BP</td>
<td>GO:0006958</td>
<td>0.00953</td>
<td>Complement activation, classical pathway</td>
</tr>
<tr>
<td>KEGG_PATHWAY</td>
<td>hsa04020</td>
<td>1.45E-05</td>
<td>Calcium signaling pathway</td>
</tr>
<tr>
<td>KEGG_PATHWAY</td>
<td>hsa04260</td>
<td>4.31E-04</td>
<td>Cardiac muscle contraction</td>
</tr>
<tr>
<td>KEGG_PATHWAY</td>
<td>hsa05010</td>
<td>5.08E-04</td>
<td>Alzheimer’s disease</td>
</tr>
<tr>
<td>KEGG_PATHWAY</td>
<td>hsa05016</td>
<td>7.83E-04</td>
<td>Huntington’s disease</td>
</tr>
</tbody>
</table>

Figure 3. Regulatory network of DNA methylation. Red nodes represent transcription factors that are significantly hypermethylated in lung cancer. Blue nodes represent microRNA and target genes of transcriptional factors. Edges represent transcriptional regulation of DNA methylation.
hibit the function of gene regulated by them. At the same time, the aberrant DNA methylation will regulate the expression of downstream genes and miRNAs, and consequently affect cell development, differentiation and apoptosis.

**Discussion**

DNA methylation profiles of lung cancer have suggested that the methylation is an important marker in genome modification. In this study, we identified 108 DMGs based on DNA methylation profiles of lung cancer samples and non-cancer lung samples. These DMGs were predicted to be involved in lung cancer by cell development, immuneresponse and apoptosis pathways. Further we constructed the regulatory network of DNA methylation by integrating transcriptional and posttranscriptional information. Some DMGs regulated the downstream target genes and microRNAs and on the contrary they were regulated by upstream transcriptional factors. We will discuss them based on previous studies.

Paired box protein Pax-5 (PAX5) encodes a potent transcription factor that plays a key role in B-cell development and carcinogenic events. PAX5 expression is regulated by its upstream factor PRDM1 or methylation in NSCLC patients. The anti-tumorigenic function of PAX5 is mediated by upregulating downstream targets of tumor protein 53 (p53) and mesenchymal-epithelial transition factor (MET).

Adenosine deaminase (ADA) is an enzyme of purine catabolism which catalyzes adenosine into inosine and the deficit of ADA causesthe inhibition of DNA synthesis and repair. ADA is also demonstrated to be a transcriptional target of p73 and p63 that sustain cellular growth by transcriptional activation of cell cycle progression genes. ASD is inversely associated with the proliferation and differentiation of lymphocytes, especially T lymphocytes and thus is also important for immune response. Patients suffering from lung cancer have lower sputum ADA activity compared with pulmonary tuberculosis.

GATA-binding protein 2 (GATA-2) is shown to be directly recruited to the promoter region of the androgen receptor upon androgen stimulation of prostate cancer cells and induce androgen receptor and key androgen-regulated gene AZGP1 expression which is associated with the progression to metastatic disease. In this study, we also found that GATA2 was methylated in lung cancer and GATA2 might be involved in lung cancer progression by regulating gonadotropin-releasing hormone receptor (GNRHR) expression, which is in line with results described by Schang et al. In addition, recent study indicates that KRAS mutant lung cancer cells depend on the transcription factor GATA2 and GATA2 depletion inhibits suppresses KRAS mutant cell growth by suppressing proteasome activity and inactivating expression of the IL-1 signaling pathway or Rho signaling.

Friend leukemia virus integration 1 (FLI1), an Ets transcription factor family member, plays a central and rate-limiting role in the pathogenesis of Ewing’s sarcoma (EWS) by fusion with EWS gene. Recent studies also suggest that FLI1 is a candidate gene for squamous lung cancer by the transcriptome network analysis. Increasing evidence indicates that EWS-FLI-1 can regulate miRNA expression. For example, De Vito show that EWS-FLI-1 can directly suppress let-7a expression in EWS family tumors development and subsequently, let-7a expression regulates EWS growth by its target gene HMGA2, an oncogene in a variety of human tumors. In this study, we also found that FLI-1 might be involved in lung cancer by modulating let-7a expression.

However, there are some limits in this study. Since the current regulatory relationships between miRNAs and transcriptional factors are limited, therefore, additional data in existing databases is needed to construct the regulating network of DNA methylation. Moreover, apart from the markers of DNA methylation, histone modification is also an important dynamic marker of the genome. Subsequent analysis will integrate data from multiple aspects and, thus, interpret lung cancer at a system perspective.

**Conflict of Interest**

None declared.

**References**


