Dna methylation in lung cancer biology essay

Science, Biology

Abstract

Objective We aim 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. Methods DNA methylation and gene expression profiles we used were from lung cancer samples and non-caner lung samples. Firstly, we identified differentially methylated genes (DMGs). Then we detected the biological processes and pathways that 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. Conclusion DNA methylation participates in carcinogenesis from 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. Keyword 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 . These aberrant cells do not carry out the functions of normal cells and do not develop into healthy lung tissue. As they grow, the aberrant cells can form tumors . 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, as of 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 cancers. Aberrant DNA methylation of OLIG1 is regarded as a novel prognostic factor in non-small cell lung cancer . Park et al. investigate the significance of DNA methylation in SLC5A8 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 . Expression of many tumor-specifically methylated genes is regulated by methylation in NSCLC patients. Moreover, HOXA2 and HOXA10 methylation may serve as prognostic parameters in squamous cell carcinoma (SCC) patients . Many diseases are caused by hereditary mutations. An increasing number of the identified disease-related mutations occur in gene regulatory sequences. Laurila and Lähdesmäki investigate the effect of mutations on transcription factor binding affinity

computationally . For example, the mutation in ALOX changes its binding status with transcriptional factor SPI1, which results in inflammatory effects . Mutation of HBD also affect its binding with transcriptional factor GATA1, which finally leads to δ-thalassemia . Respiratory and related diseases are affected by modified transcription regulation programs. Many TFs are significantly enriched in the target disease groups . Several histone lysine methyltransferases (HKMTs) have been identified and histone lysine methylation is now considered to be a critical regulator of transcription . Hypomethylation of intragenic LINE-1 represses transcription in cancer cells through AGO2 . Ldhc gene expression in cancer cells is regulated by transcription factor Sp1, CREB, and CpG island methylation . MicroRNAs (miRNAs) are short, non-coding RNAs (~22 nt) that play important roles in the pathogenesis of human diseases by negatively regulating gene expression. MicroRNA-183 regulates Ezrin expression in lung cancer cells . miR-93, miR-98, and miR-197 regulate expression of tumor suppressor gene FUS1 . MiR-21 is an EGFR-regulated anti-apoptotic factor in lung cancer in never-smokers . Brueckner et al. identify let-7a-3 as an epigenetically regulated miRNA gene with oncogenic function and suggest that aberrant miRNA gene methylation might contribute to the human cancer epigenome . Gene Ontology is utilized to characterize the function categories affected by lung cancer. It determines the functional categories of lung cancer related genes. Genes that differently expressed between normal lung tissue and cancer show enrichment in gene ontology terms associated with mitosis and proliferation . RBM5/H37 tumor suppressor, located at the lung cancer hot spot 3p21. 3, alters expression of genes involved in metastasis. Overall gene

set of the gene ontology group is " proteinaceous extracellular matrix" . Methylated genes are significantly enriched as transcription factors and in processes of neuronal differentiation . Gene ontology characterize that epigenetic genes shows are almost exclusively involved in morphogenetic differentiation processes in lung cancer . Many of lung cancer related genes are enriched in biologic pathways. Using an improved gene-set-enrichment analysis approach, the Fas signaling pathway and the antigen processing and presentation pathway are most significant related with lung cancer susceptibility . DNA copy number aberrations in small-cell lung cancer are significantly activated by the focal adhesion pathway . MicroRNA signatures in tumor tissue are related to pathway " angiogenesis" in non-small cell lung cancer evaluated by Gene Set Enrichment Analysis (GSEA) . In the final integrative analysis of lung cancer related miR-21-targets analysis, 24 hub genes were identified by overlap calculation, suggesting that miR-21 may play an important role in the development and progression of lung cancer through JAK/STAT signal pathway, MAPK signaling pathway, Wnt signaling pathway, cell cycle, PPAR signaling pathway, apoptosis pathway and other pathways . 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 is downloaded from NCBI Gene Expression Omnibus (GEO) (http://www. ncbi. nlm. nih. gov/geo/) under accession number GSE32867. All genome-scale data is generated by

Selamat et al. . They detected the DNA methylation status of 27, 578 CpG dinucleotides CpGs sites on 28 lung cancer samples and 27 adjacent non-lung cancer tissue samples. The DNA methylation profiles use Illumina Infinium HumanMethylation27 platform on 28 lung adenocarcinoma and 27 adjacent non-tumor lung samples.

Gene expression data

The gene expression data we used is downloaded from NCBI Gene Expression Omnibus (GEO) (http://www. ncbi. nlm. nih. gov/geo/) under accession number GSE32863 . The gene expression profiles are detected on HumanWG6 v3. 0 microarray, the platform is Illumina HumanWG-6 v3. 0 expression beadchip. The 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, 27, 578 probes map to14, 495 protein-coding genes. When multiple probes are mapped to the same gene, the level of DNA methylation βfor this gene is represented by the mean value of all the probes. For gene expression, when one probe is mapped to multiple genes, these probes are removed. For multiple probes that are mapped to the same gene, expression level of this gene is 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 identify differentially methylated genes (DMGs).

The statistic method we use is Student's t-test (FDR = 0. 01), the threshold of differential methylation level is 10%. In this manner, we identify 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 are also analyzed from the perspective of KEGG pathway. Both processes utilize DAVID , 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 are less than 500bp from TSS of genes and 92% of the methylated loci are less than 1, 000bp 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(Supplementary Data). As shown in Figure 2, 94 genes have high methylation level near their TSS. The DNA methylation level is 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 1, 108 DMGs were significantly enriched in 19 biological processes and four pathways. They include multiple cancer-related processes, such as apoptosis process, the cell death process,

ion transfer and the immune response regulation process. The researchers believe that in the process of cancer progression, the normal homeostasis is damaged by endogenous factors, leading to the function decline of immune system. However, it has also been reported that the incidence of cancer drives immune system, because the rapid proliferation of cancer cells is resulted from that the immune barrier is broken down. The regulation of the immune system and the cancer occurrence is closely related to each other. The driving and driven relationship may co-exist in the process of cancer occurrence and development, but the dominant relationship varies at different stages. Therefore, in order to more accurately elaborate the mechanism of cancer, we urgently need additional experiments and data analysis systems.

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 imprinting . 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-test method ($p < 0.05$).

Regulation network of DNA methylation

The TRANSFAC provides abundant data on eukaryotic transcription factors, their experimentally-proven binding sites, consensus binding sequences and regulated genes . 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 regulating the transcription process. TransmiR database stores the relationship between miRNA and transcriptional factors by literature mining . From TransmiR database, we identified the transcriptional relationship between two DNA hypermethylated genes and five miRNAs (pre-mir-302a/b/c/d, and pre-mir-367). It is consistent with previous functional analysis. 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 participate in carcinogenesis from the transcriptional and post-transcriptional level. In the process of carcinogenesis, aberrant DNA methylation will prevent its binding with the upstream regulatory proteins and inhibit the function of gene regulated by them. At the same time, the aberrant DNA methylation will regulate the

expression of downstream genes and miRNA, and consequently affect cell development, differentiation and apoptosis.

Discussion

Lung cancer is the most common primary pulmonary malignant tumors, generally originate in the bronchial epithelium . During the past 50 years, all over the world particularly in industrial countries, the incidence and mortality of lung cancer has rapidly rise . Therefore, it is of great significance to research on treatment of lung cancer. DNA methylation profiles of lung cancer suggest that the methylation is an important marker in genome modification . In this study, we identified 108 differentially methylated genes (DMGs) based on DNA methylation profiles of lung cancer samples and non-caner lung samples. It indicates that the DMGs imply the occurrence mechanism of lung cancer or even the key points of the lung cancer treatment. After conducting functional analysis from biological processes of Gene Ontology (GO) and KEGG pathway, a number of changes occur in cell development, immuneresponse and apoptosis. Further we constructed the regulatory network of DNA methylation by integrating transcriptional and posttranscriptional information. Some DMGs regulate the downstream target genes and microRNAs and on the contrary they are regulated by upstream transcriptional factors. Paired box protein Pax-5 (PAX5) is differentially methylated in lung cancer. This is consistent with previous study that reported NSCLC patients exhibited methylation of multiple genes such as PAX5 . Moreover, the expression of PAX5 is significantly different between lung cancer samples and non-lung cancer samples. From GO enrichment analysis, the DMGs influence multiple cancer-related processes. Defects in

apoptosis process are reported to be implicated in tumorigenesis of lung cancer . Cancer could be explained, at least in part, as an abnormal immune system tolerance to uncontrolled cells. These important biological processes have already been reported to be involved in the occurrence of lung cancer . The malignancy degree of non-small cell lung cancer is vicious high, part of its high mortality rate originates from the brain metastase . We suppose that during the process of lung cancer development, lung cancer cells migrate into brain through spine with the pulmonary vascular and thus cause brain cancer. The lung cancer cells have good affinity with the central nervous cells. In this study, the DMGs are significantly enriched in the process of blood circulation and signaling pathways of neural diseases . KEGG pathway analysis presents that DMGs are closely enriched into pathway " Alzheimer’s Disease". Recent study demonstrates that a decreased incidence of overall cancers in patients with Alzheimer’s disease, and moreover patients with Alzheimer’s Disease have a significantly decreased risk of lung cancer . The activity of metabotropic glutamate receptors (mGluRs) is known to be altered in neurodegenerative diseases such as Alzheimer's and Huntington's disease. Recent reports indicate several somatic mutations in mGluR1a in lung cancer . Moreover, the DMGs influence multiple transcription factors, these DNA methylation loci play an important role for the regulation of gene expression. Zhu et al. demonstrate that methylation at CG sites outside of the consensus Sp1-binding site may directly reduce the ability of Sp1/Sp3 to bind its DNA recognition element in human lung cancer cells . Hypermethylation of CpG on the promoter is able to prevent the entry of the DNA binding protein, which is closely related with the gene silencing. In

addition, DNA methylation presents a dynamic landscape in the genome. Pre-mir-302a/b/c/d and pre-mir-367play key roles in development . The methylation status can be reversed by intervention. Therefore methylation pattern can be utilized to identify molecular markers. The methylation loci that potentially drive cancer could be considered as drug targets, which will effectively facilitate the prevention and treatment of cancer. DNA methylation is a critical epigenetic modification that silences gene transcription. Survivin overexpression is closely linked to lung oncogenesis. Silencing of survivin gene by molecular antagonists has shown promise as novel anticancer strategies. SurKex1 exerts its down-regulatory effects on survivin expression through the activation of DNMT1 . SLC5A8 was reactivated by treatment with DNA methyltransferase inhibitor, 5-Aza and/or HDAC inhibitor, trichostatin A (TSA) in lung cancer cell lines, which did not express SLC5A8 . The limitation of this study is the follow-up analysis of regulation network. Because 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.