

Integrated analysis of DNA methylation and microRNA regulation of the lung adenocarcinoma transcriptome

JIANG DU and LIN ZHANG

Department of Thoracic Surgery, Chinese Medical University Affiliated No. 1 Hospital,
Shenyang, Liaoning 110001, P.R. China

Received January 27, 2015; Accepted April 20, 2015

DOI: 10.3892/or.2015.4023

Abstract. Lung adenocarcinoma, as a common type of non-small cell lung cancer (40%), poses a significant threat to public health worldwide. The present study aimed to determine the transcriptional regulatory mechanisms in lung adenocarcinoma. Illumina sequence data GSE 37764 including expression profiling, methylation profiling and non-coding RNA profiling of 6 never-smoker Korean female patients with non-small cell lung adenocarcinoma were obtained from the Gene Expression Omnibus (GEO) database. Differentially methylated genes, differentially expressed genes (DEGs) and differentially expressed microRNAs (miRNAs) between normal and tumor tissues of the same patients were screened with tools in R. Functional enrichment analysis of a variety of differential genes was performed. DEG-specific methylation and transcription factors (TFs) were analyzed with ENCODE ChIP-seq. The integrated regulatory network of DEGs, TFs and miRNAs was constructed. Several overlapping DEGs, such as v-ets avian erythroblastosis virus E26 oncogene homolog (*ERG*) were screened. DEGs were centrally modified by histones of tri-methylation of lysine 27 on histone H3 (H3K27me3) and di-acetylation of lysine 12 or 20 on histone H2 (H2BK12/20AC). Upstream TFs of DEGs were enriched in different ChIP-seq clusters, such as glucocorticoid receptors (GRs). Two miRNAs (miR-126-3p and miR-30c-2-3p) and three TFs including homeobox A5 (*HOXA5*), Meis homeobox 1 (*MEIS1*) and T-box 5 (*TBX5*), played important roles in the integrated regulatory network conjointly. These DEGs, and DEG-related histone modifications, TFs and miRNAs may be important in the pathogenesis of lung adenocarcinoma. The present results may indicate directions for the next step in the study of the further elucidation and targeted prevention of lung adenocarcinoma.

Introduction

Lung cancer is presently the leading cause of global cancer-related death, with an increasing prevalence and mortality. Smoking is the predominant risk factor for lung cancer. While, in East Asia, ~30% of patients suffering from lung cancer were never smokers (1,2), and non-smoking-related lung cancer can also occur in current and former smokers (3). Unfortunately, lung cancer has not been solved in regards to prevention or diagnosis or treatment.

In addition to small cell lung cancer (SCLC) accounting for 10-15% of lung cancer cases, non-small cell lung cancer (NSCLC) represents ~85-90% of overall lung cancer cases (4). NSCLC is also subdivided into three histologic types, including adenocarcinoma, squamous cell carcinoma and large cell undifferentiated carcinoma. Lung adenocarcinoma accounts for almost 40% of NSCLC. Due to the relatively high incidence of lung adenocarcinoma, much research has been conducted to elucidate its nature and mechanisms.

Previous studies have found various genes related to lung adenocarcinoma. Su *et al* (5) found that a higher level of cyclooxygenase-2 decreased the survival rate of patients through many mechanisms, such as a corresponding higher level of vascular endothelial growth factor that stimulated the growth and migration of cancer cells (6), a higher lymphatic vessel density that reduced the restriction of cancer cell invasion (7), and enhanced lymph node metastasis that accelerated the metastasis of cancer cells (8). Mutations of oncogene *K-ras* and tumor-suppressor gene *TP53* have a strong link with lung adenocarcinoma (9). Other fusion genes have been further studied concerning the correlation with lung adenocarcinoma. Fusion of the kinesin family member 5B and *RET* proto-oncogene was found to occur in a subset of NSCLC (10). Fusion genes of echinoderm microtubule associated protein like 4 - anaplastic lymphoma receptor tyrosine kinase and kinesin light chain 1 - anaplastic lymphoma receptor tyrosine kinase were also found in lung adenocarcinoma (11).

To date, the pathogenesis of NSCLC and lung adenocarcinoma is difficult to determine. To reduce the enormous morbidity and mortality of lung adenocarcinoma, it is critical to identify lung adenocarcinoma-associated genes and mechanisms. Integrated analysis of full DEGs and the expression of regulatory factors such as methylation, mRNA splicing, transcription factors (TFs) and microRNAs (miRNAs) is an

Correspondence to: Dr Lin Zhang, Department of Thoracic Surgery, Chinese Medical University Affiliated No. 1 Hospital, Shenyang, Liaoning 110001, P.R. China
E-mail: linzhangzhl@163.com

Key words: lung adenocarcinoma, differentially expressed genes, microRNA, transcription factors, methylation

effective method for disease pathogenesis study. In the present study, DEGs, exons and isoforms, as well as DEG-related methylation, TFs and miRNAs were integrated and analyzed.

Materials and methods

Datasets. The raw experimental data under accession no. GSE37764 (12) used in the present study are publically available in the Gene Expression Omnibus (GEO) database (<http://www.ncbi.nlm.nih.gov/geo>). These data, which include expression profiling, methylation profiling and non-coding RNA profiling of 6 never-smoker Korean female patients, were produced by high throughput sequencing. The histologic origins were cancer tissues and adjacent normal tissues of non-small cell lung adenocarcinoma. In the present study, using normal tissues as control, the molecular variations in tumor tissues were identified. The platform of this data is GPL10999 (Illumina Genome Analyzer IIX, *Homo sapiens*).

Methylation profiling and differentially methylated gene screening. Trimmomatic (13) software package, a flexible, pair-aware and efficient preprocessing tool for Illumina sequence data, is often used to remove low quality reads and trim adaptor sequences. In the present study, the methylated DNA immunoprecipitation-sequencing (MeDIP-seq) was preprocessed with the Trimmomatic (13). During the preprocessing of the Illumina reads, a minimum quality cutoff on the first and last bases using LEADING: 3 (trim the leading nucleotides until quality >3) and TRAILING: 3 (trim the trailing nucleotides until quality >3) was imposed, and a minimum sliding window quality using SLIDINGWINDOW: 4:15 (trim the window of size four for reads with local quality below a score of 15) was subjected. In addition, the resulting reads shorter than 25 bases were discarded. Then the Bowtie (14) alignment algorithm (with default parameters) was used to align the Illumina reads to the human reference genome (hg19), and SAM tools (15) was applied to remove PCR duplicates. The differentially methylated regions (DMRs) were identified by MEDIPS (16) in R. with false discovery rate (FDR) <0.1. Each DMR contains multiple methylated loci, and the determination of overlaps between methylation loci and the adjacent genes were computed using the BED Tools (17) software. Briefly, differentially methylated loci between -2,000 and +1,000 bp around transcription start site (TSS) were selected, and the adjacent genes were defined as differentially methylated genes (DMGs).

Gene expression profile analysis. RNA-seq reads were cleaned to remove low quality regions and sequencing adaptors utilizing Trimmomatic (13) software package (LEADING: 3, TRAILING: 3, SLIDINGWINDOW: 4:15, MINLEN: 36). These massively parallel short reads were subsequently mapped to a reference genome with TopHat (18) (no >5 bases mismatch). Since multi-exon genes can encode different transcripts and multiple transcript variants encode different isoforms, differentially expressed exons were analyzed by DEXSeq (19) in R. and differential expression analysis of genes and transcript isoforms were performed with Cufflinks (20) algorithm. The parameters of DEXSeq and Cufflinks were default values. The thresholds were q-value <0.05 and fold-change (FC) >2. Comparing the cancer tissues and control, genes and exons with average

expression levels >10 FPKM (fragments/kilobase of transcript/million mapped reads) were defined as differentially expressed.

Screening of differentially expressed miRNAs. High-throughput sequencing data for miRNA expression were cleaned to remove low quality regions and sequencing adaptors utilizing Trimmomatic [2] software package (LEADING: 3, TRAILING: 3, SLIDINGWINDOW: 4:15, MINLEN: 25). Reads of millions of miRNA sequences were aligned to the genome and the expression value of each miRNA was measured using miRExpress (21). Subsequently, the screening of differentially expressed miRNAs was performed with DESeq2 (22) in R/Bioconductor FDR <0.01; FC >2; base mean >10). Conserved miRNA targets were retrieved from public websites (TargetScan, <http://targetscan.org>) (23).

Functional analysis of a variety of differential genes. Gene Ontology (GO) functional enrichment and annotation of differential genes, including differentially methylated genes and DEGs, were computed using the database for annotation, visualization and integration discovery (DAVID) (24). The annotation of miRNA-target DEGs was performed with TarBase 6.0 database (capturing the exponential growth of miRNA targets with experimental support) (25). Then the DIANA miRPath v.2.0 (26) was used to determine molecular pathways potentially altered by differentially expressed miRNAs based on the Kyoto Encyclopedia of Genes and Genomes (KEGG) database. The ChIP-X Enrichment Analysis (ChEA) and the ENCODE ChIP-seq (28) were utilized to search for enriched TFs located upstream of the DEGs. TF-target genes were predicted and combined with differentially expressed miRNAs and DEGs using mirConnX (29) with Pearson's correlation coefficient >0.96 and then, the integrated network of TFs, miRNAs and TF-target DEGs were constructed and analyzed.

Results

Differentially methylated regions and genes. After comparison of the MeDIP-Seq data between cancer tissues and paracarcinoma tissues of 6 non-small cell lung carcinoma patients, DMRs and DMGs were obtained. Most of the DMGs (>90%) were detected in one patient only (Fig. 1A). Only 82 genes were found in 2 or more patients and ~1/3 of the DMGs (34) were located in the mitochondrial genome. The numbers of DMGs in 5 patients were similar except in 1 patient (P3) (Fig. 1B). In patient P3, the hypermethylated genes were significantly more than in the others.

Functional GO analysis showed that the DMGs were mostly associated with metabolic pathways (Fig. 2A). The most commonly enriched GO terms were cell morphogenesis, mitochondrial ATP synthesis coupled electron transport and ATP synthesis coupled electron transport.

KEGG pathway analysis identified two enriched pathways, respectively: Parkinson's disease (hsa: 05012) and oxidative phosphorylation (hsa: 00190) (Fig. 2A).

Gene expression profile analysis. DEG screening analysis found that a total of 1,498 genes were differentially expressed between the cancer tissues and adjacent normal tissues, and 1207 isoforms of 1,103 genes were differentially expressed.

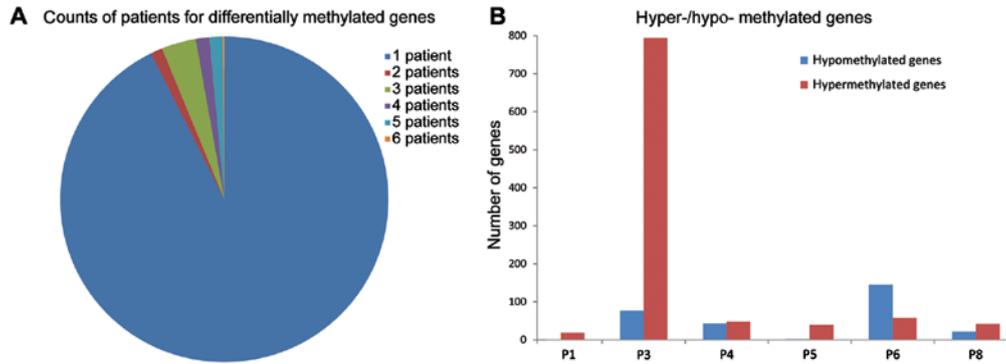


Figure 1. Differentially methylated genes. (A) Gene frequency of methylation in 6 patients. (B) Number of methylated genes in 6 patients. Red column indicates hypermethylated genes and blue column indicates hypomethylated genes. The x-coordinate represents 6 patients and the y-coordinate represents the gene number.

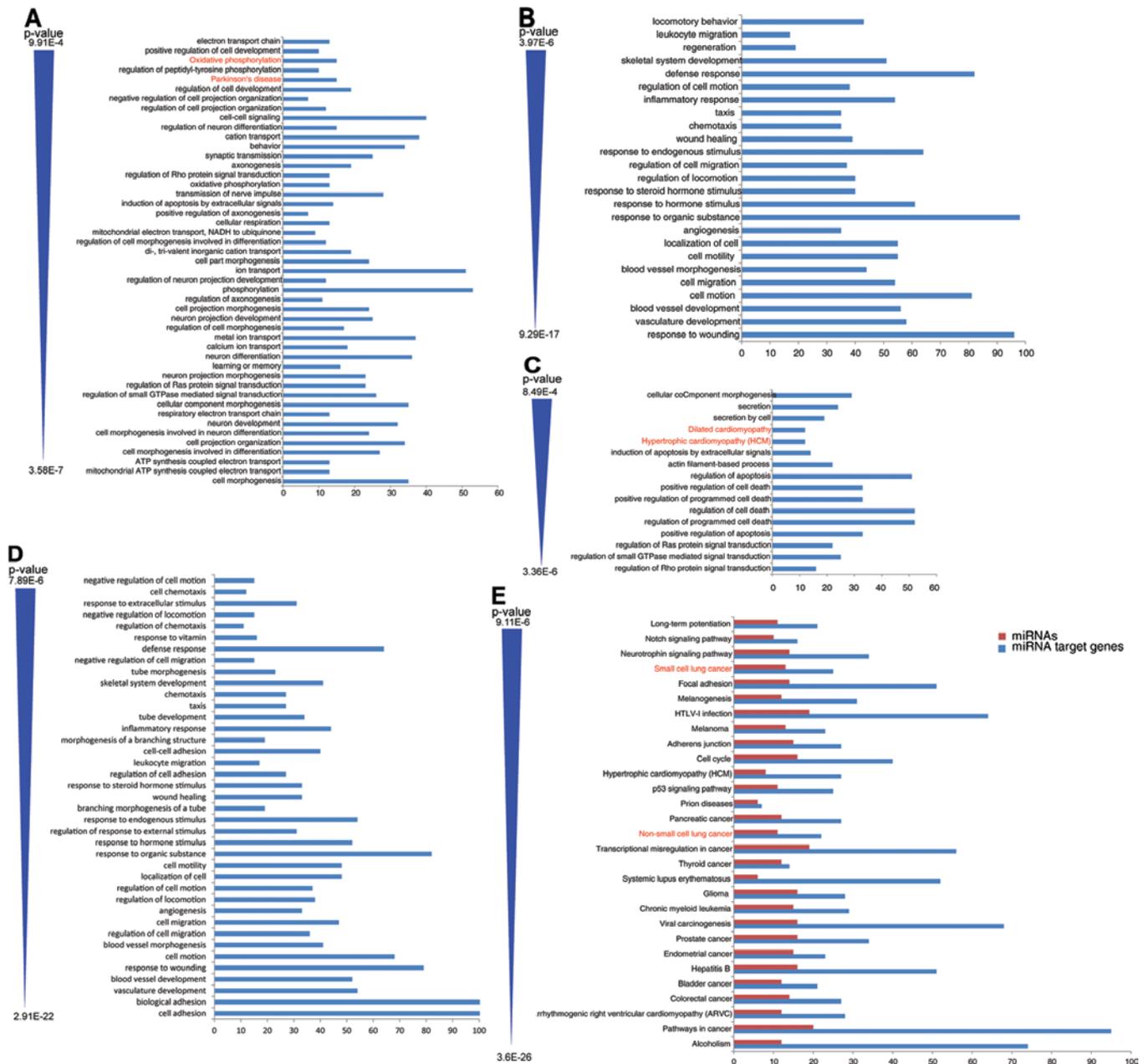


Figure 2. Functional annotations of the differential genes. (A) GO and KEGG analysis of the DMGs. Red indicates the KEGG pathways. (B) GO analysis of the DEGs. Red indicates KEGG pathways. (C) GO and KEGG analysis of differentially expressed exon-related genes. (D) GO analysis of differentially expressed isoform-related genes. (E) GO and KEGG analysis of differentially expressed microRNAs. The x-coordinate represents p-value and the y-coordinate represents gene number.

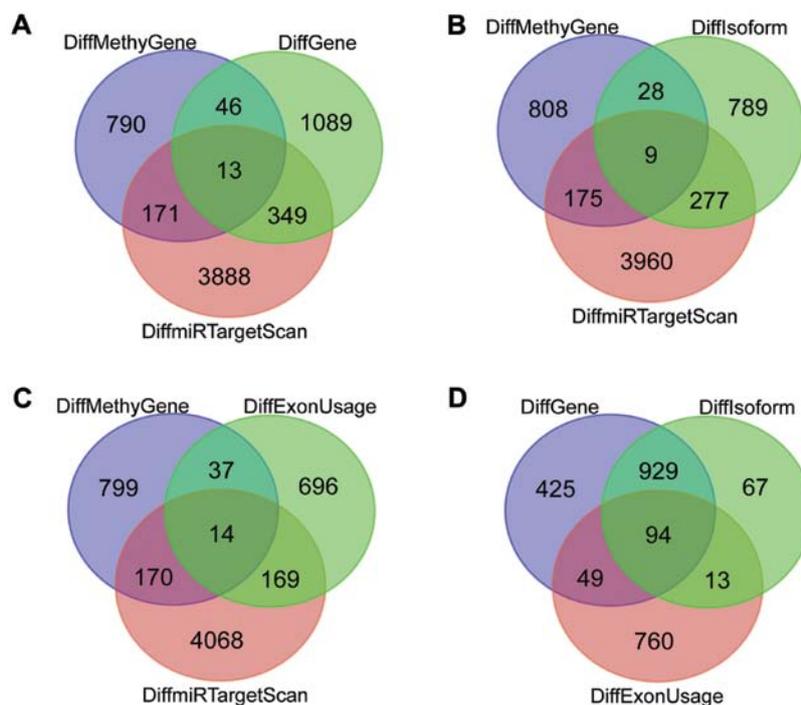


Figure 3. Comparison of DMGs, differentially expressed transcripts and differentially expressed microRNA-target genes. (A) Comparison of DMGs, DEGs and differentially expressed miRNA-target genes. (B) Comparison of DMGs, differentially expressed isoform-related genes and differentially expressed miRNA-target genes. (C) Comparison of DMGs, differentially expressed exon-related genes and differentially expressed miRNA-target genes. (D) DEGs, differentially expressed isoform-related genes and differentially expressed exon-related genes.

Table I. Thirteen differentially expressed genes under the regulation of differentially expressed miRNAs and methylation.

Name	miRNA	Expression	N	T	log2FC (T/N)	Hyper-methylated sample
ERG	hsa-miR-9	1	30.2622	7.6856	-1.97729	1T
FLI1	hsa-miR-200b	1	35.8052	11.0298	-1.69876	1T
RPS6KA2	hsa-miR-200b	1	39.4554	19.498	-1.01689	1N1T
STARD8	hsa-miR-144	-1	15.475	4.3143	-1.84274	1T
PDZD2	hsa-miR-182	1	17.9635	5.28059	-1.76629	1T
GPR56	hsa-miR-139-5p	-1	32.3344	65.4073	1.01638	1T
SDC1	hsa-miR-182	1	67.1969	206.511	1.61975	1T
SDC1	hsa-miR-708	1	67.1969	206.511	1.61975	1T
SDC1	hsa-miR-9	1	67.1969	206.511	1.61975	1T
EGLN3	hsa-miR-9	1	5.38757	30.8381	2.51701	1N
COL5A1	hsa-miR-182	1	25.2564	57.9578	1.19835	1T
COL5A1	hsa-miR-9	1	25.2564	57.9578	1.19835	1T
AATK	hsa-miR-182	1	15.6319	2.48414	-2.65367	1T
AATK	hsa-miR-503	1	15.6319	2.48414	-2.65367	1T
AATK	hsa-miR-9	1	15.6319	2.48414	-2.65367	1T
GAB2	hsa-miR-486-5p	-1	17.3317	8.056	-1.10527	1N
GAB2	hsa-miR-9	1	17.3317	8.056	-1.10527	1N
DOT1L	hsa-let-7a	-1	14.0112	6.96755	-1.00786	1N1T
DOT1L	hsa-miR-200b	1	14.0112	6.96755	-1.00786	1N1T
DOT1L	hsa-miR-224	1	14.0112	6.96755	-1.00786	1N1T
DOT1L	hsa-miR-9	1	14.0112	6.96755	-1.00786	1N1T
THBS2	hsa-miR-182	1	10.5524	64.5186	2.61214	1N1T
THBS2	hsa-miR-30a	-1	10.5524	64.5186	2.61214	1N1T
THBS2	hsa-miR-9	1	10.5524	64.5186	2.61214	1N1T

N, normal; T, tumor; -1, downregulated in tumor; 1, upregulated in tumor; FC, fold change.

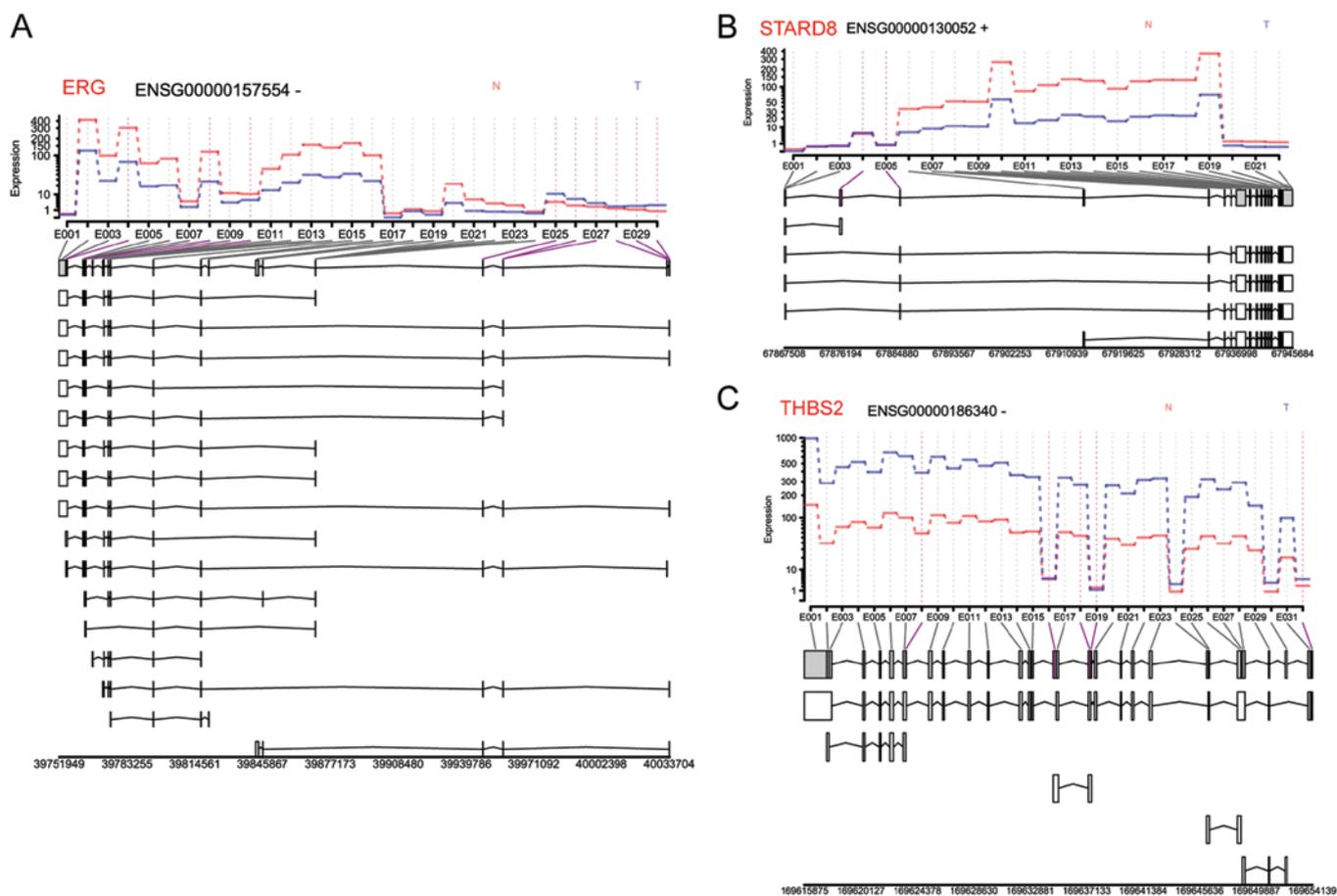


Figure 4. (A-C) Alternative splicing of three overlapping genes. The x-coordinate represents the location on the chromosome and the y-coordinate represents the gene expression level.

Additionally, 1,286 exons of 916 genes were also differentially expressed.

Functional GO analysis showed that most of these differential genes were related to cell migration and apoptosis (Fig. 2B-D). The most commonly enriched terms of genes were response to wounding, vasculature development and blood vessel development. The most commonly enriched terms of the exons were regulation of Rho protein signal transduction, regulation of small GTPase mediated signal transduction and regulation of Ras protein signal transduction. The most commonly enriched terms of transcripts were cell adhesion, biological adhesion and vasculature development.

KEGG pathway analysis of the differentially expressed exons found two enriched pathways named hypertrophic cardiomyopathy (HCM) (hsa: 05410) and dilated cardiomyopathy (hsa: 05414) (Fig. 2C).

A total of 541 genes among these differential genes possessed only 1 differentially expressed exon. As shown in Figs. 3 and 4, the expression levels of different exons corresponding to the same one gene had dozens of time variations. Several exons with a lower expression level appeared almost unanimously between the cases and control, while others displayed a significant difference between these two groups.

There were 94 common genes differentially expressed at the levels of genes, isoforms and exons (Fig. 3D).

Differential expression of miRNAs. Within 50 precursors, 44 miRNAs were detected as differentially expressed miRNAs. The most overexpressed miRNAs were miR-6510, miR-301b and miR-183. miR-144, miR-486 and miR-451a were the most commonly downregulated miRNAs.

GO and pathway analysis shown that the differentially expressed miRNAs were correlated with many disease-related pathways (Fig. 2E). The top three enriched pathways, respectively, were alcoholism (hsa: 05034), pathways in cancer (hsa: 05200) and arrhythmogenic right ventricular cardiomyopathy (ARVC) (hsa: 05412). NSCLC (hsa: 05223) and small cell lung cancer (hsa: 05222) were also enriched pathways.

Nine differentially expressed miRNA-target genes with methylation were differentially expressed at levels of isoforms (Fig. 3B) and 14 at the level of exons (Fig. 3C).

Integrated analysis of the differential genes. Thirteen differentially expressed miRNA-target genes included differentially methylated genes, and also differentially expressed exon-related genes. Furthermore, they were regulated by differentially expressed miRNAs (Table I, Fig. 3A). Yet, differential methylation of these 13 genes was observed only in one patient. In addition, ribosomal protein S kinase, 90 kDa, polypeptide 2 (*RPS6KA2*), DOT1-like histone H3K79 methyltransferase (*DOT1L*) and thrombospondin 2 (*THBS2*),

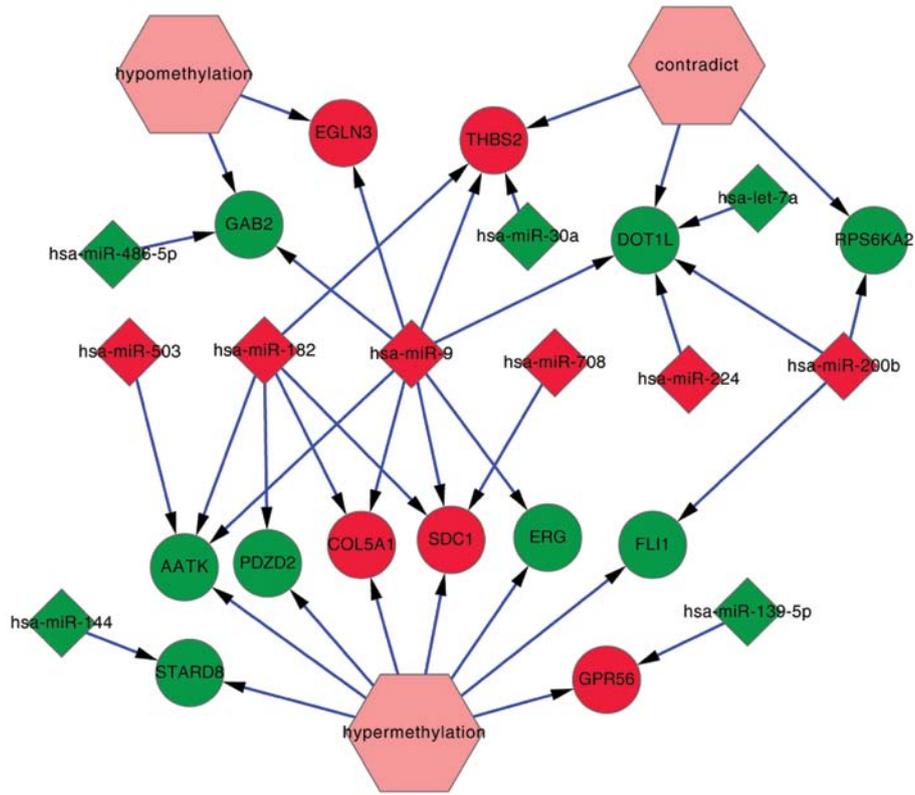


Figure 5. Methylation and microRNA regulatory network of 13 DEGs. Downregulated and upregulated genes are shown in green and red circles, respectively; overexpressed and downregulated miRNAs are displayed in red and green rhombuses, respectively. DEGs, differentially expressed genes.

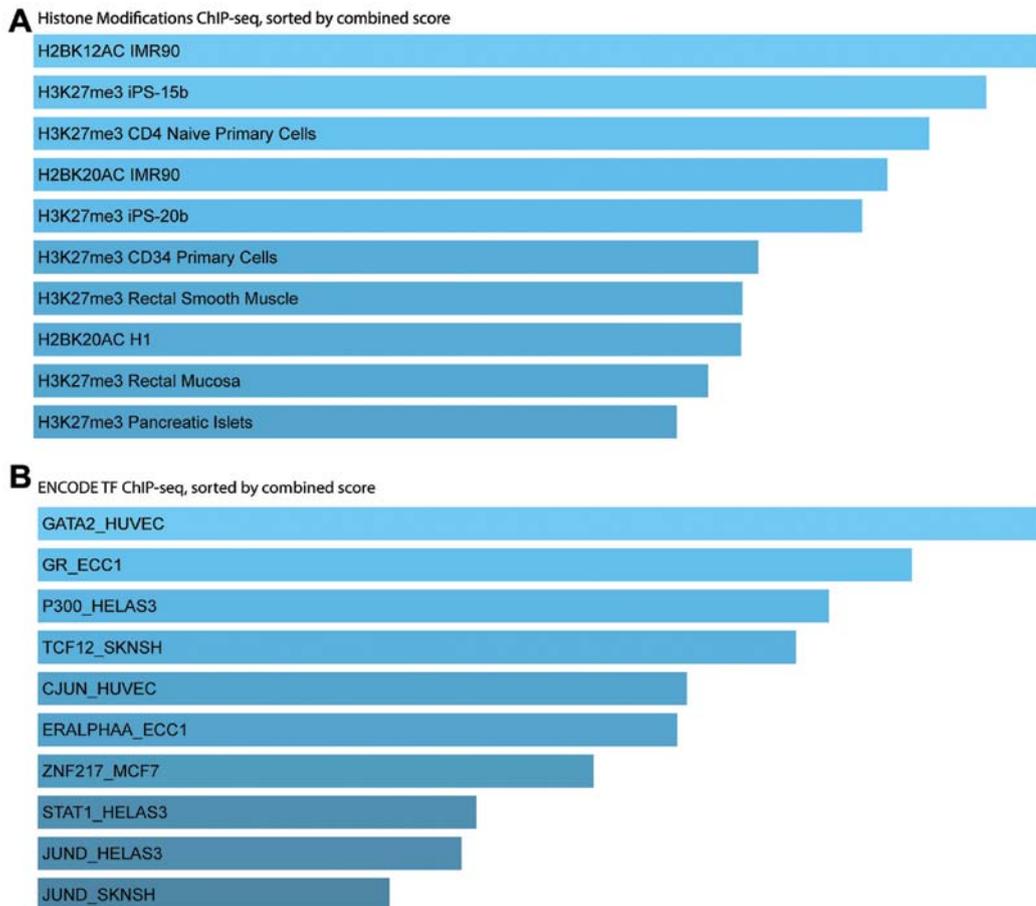


Figure 6. Enrichment analyses of (A) histone modifications and (B) upstream transcription factors of the differentially expressed genes.

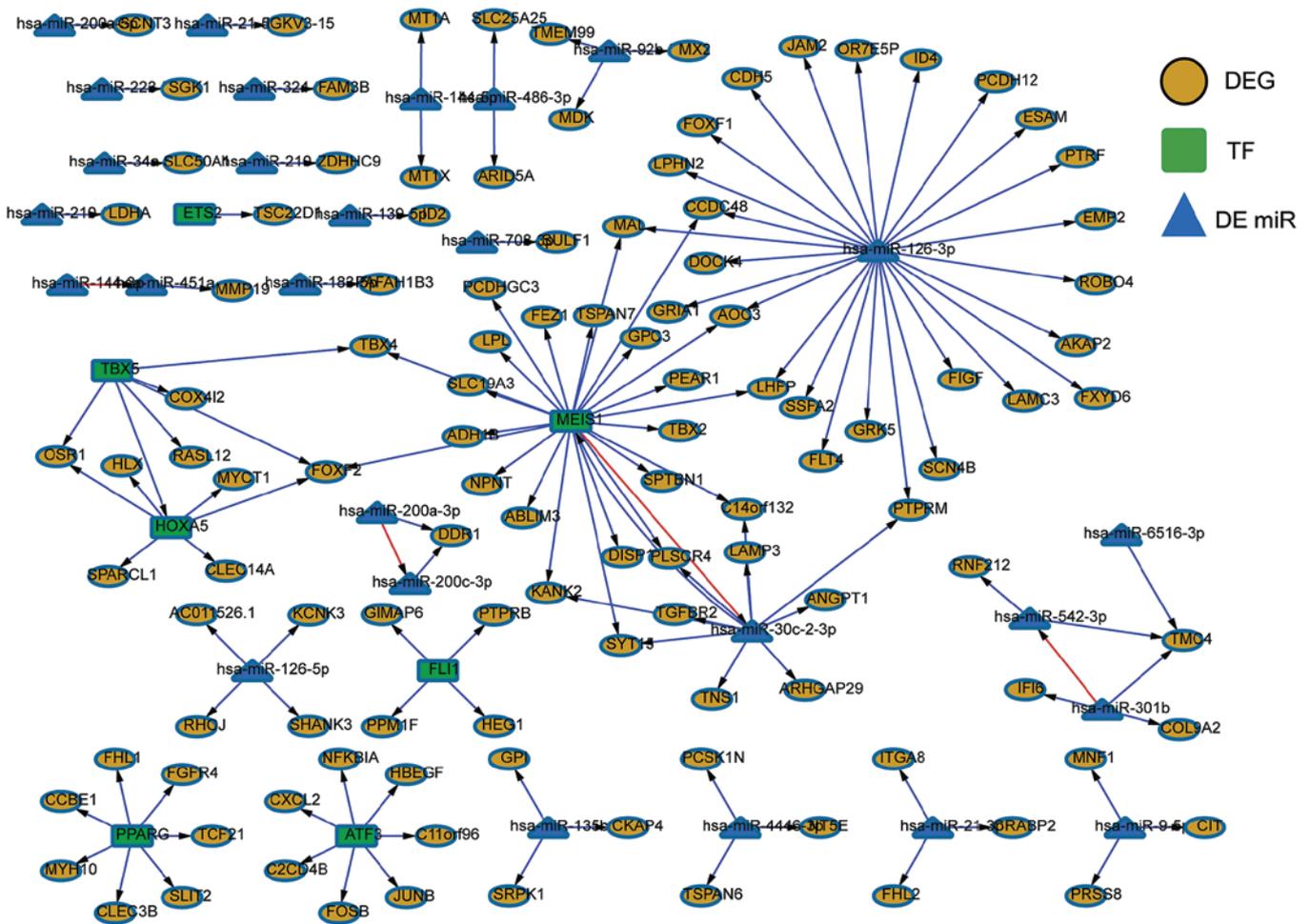


Figure 7. Transcription factor-microRNA regulatory network of the differentially expressed genes. Differentially expressed genes, transcript factors and differentially expressed miRNAs are shown in yellow, green and blue, respectively.

included in these 13 genes, were detected with both hypomethylation and hypermethylation.

The overlapping genes among the differentially methylated genes, DEGs and isoforms, differentially expressed exon-related genes, differentially expressed miRNA-targeted genes were v-ets avian erythroblastosis virus E26 Oncogene Homolog (*ERG*), StAR-related lipid transfer domain containing 8 (*STARD8*) and *THBS2*. Isoforms and expression levels of these 3 genes are shown in Fig. 4

The possible regulatory networks of methylation, miRNA expression and gene expression are shown in Fig. 5. There were 6 overexpressed and 5 downregulated miRNAs, 5 upregulated and 8 downregulated genes, 2 genes with hypomethylation, 8 genes with hypermethylation and 3 genes with contradiction. In the 11 miRNAs, miR-9 (degree= 8) and miR-182 (degree=5) had more degrees than the others. Genes including *DOT1L*, apoptosis-associated tyrosine kinase (*AATK*), syndecan 1 (*SDC1*) and *THBS2* had more degrees.

Transcription analysis of DEGs. ChEA2 analysis results indicated that the screened DEGs were modified and regulated by multi-cancer cell line histones including tri-methylation of lysine 27 on histone H3 (H3K27me3) and di-acetylation of lysine 12 or 20 on histone H2 (H2BK12/20AC), which

pertained to the ENCODE database (Fig. 6A). The upstream TF binding patterns were not as clustered as that of the histone modification; they were enriched in different ChIP-seq clusters of TFs in different cell lines (Fig. 6B), such as GATA2 and CJUN in human umbilical vein endothelial cells (HUVECs), glucocorticoid receptors (GRs) and estrogen receptor (*ER*) α in endometrial cells (ECC1), while, P300, signal transducer and activation of transcription (STAT1) and JUND in HeLaS3 cells.

DEGs were regulated by TFs and miRNAs, and the regulatory network was constructed as shown in Fig. 7. There were 116 DEGs, 72 TFs and 7 differentially expressed miRNAs. miR-126-3p served as a ‘hub’ in the gene regulatory network which regulated 26 DEGs. The TF MEIS1 was another ‘hub’, which also regulated 22 DEGs and miR-30c-2-3p. Several sub-networks with homeobox A5 (*HOXA5*), Meis homeobox 1 (*MEIS1*), T-box 5 (*TBX5*), miR-126-3p and miR-30c-2-3p as centers shared several nodes and then formed another greater regulatory network. The remaining sub-networks were detached from each other.

Discussion

NSCLC accounts for ~85% of all lung cancer cases (30) and remains the leading cause of cancer-related death worldwide

(31). Lung adenocarcinoma, the major subtype of NSCLC responsible for more than 500,000 mortalities/year worldwide (32), is associated with a poor prognosis. In the present study, differentially methylated regions, differentially expressed miRNAs and transcriptomics between different tissues from 6 non-small cell lung adenocarcinoma patients were analyzed. Several DEGs, miRNAs and TFs were screened, which were expected to be associated with metabolism, cell apoptosis or various diseases; thus, they may be important in the progression of lung adenocarcinoma. Kim *et al* (12) identified various novel genetic aberrations, gene network modules and miRNA-target interactions within the same dataset, yet, it is distinct from ours. In addition, the pathogenesis of lung adenocarcinoma is far from clear. With the different bioinformatics tools, the results of the same analysis were slightly different from those of Kim *et al* (12). The new information obtained from the present study may help to illuminate the molecular mechanisms of this disease.

The methylation analysis results of the 6 patients were relatively diverse. Since a relatively large number of methylation sites are concentrated in the mitochondrial genome, it is difficult to research the connections between methylation and gene expression. Additionally, *RPS6KA2*, *DOT1L* and *THBS2* were detected with both hypomethylation and hypermethylation. This may be relatively related to the great individual differences of the methylated regions in the patients. However, previous research has confirmed that DNA methylation is critical in lung cancer (33). Combined with the subsequent analysis of differential expression of genes, isoforms and exons, screening of miRNA-related genes, three overlapping genes were obtained. *ERG*, a member of the *ETS* oncogene family (34), is intimately involved in the development of multiple cancers including prostate cancer (35). The *TMPRSS2-ERG* gene fusion is now a specific biomarker of prostate cancer (36). There is little research on *ERG* and lung adenocarcinoma. As cancer develops from glands which are the same as prostate cancer, *ERG* may also be related with lung adenocarcinoma. *STARD8* was found downregulated and highly methylated in the present study. Durkin *et al* (37) suggested that *STARD8* is a tumor-suppressor gene encoding DLC-3 to suppress tumor cell growth. It has a higher level of methylation in colorectal cancer than other types of cancer (38). *THBS2*, an upregulated gene, encodes a protein belonging to the thrombospondin family. This protein has been shown to function as a potent inhibitor of tumor growth and angiogenesis, and it may be involved in cell adhesion and migration (39). *THBS2* also has CpG island methylation in malignant ovarian tumors (40). Therefore, *STARD8* and *THBS2* may also be involved in lung adenocarcinoma.

Gene expression is under the elaborate control of interrelated factors including TFs and histone modification. In the present study, comparative analysis of histone modifications in tumor and normal tissues was conducted. This revealed that DEGs were centrally regulated by H3K27me3 and H2BK12/20AC in several cancer cell lines. H3K27me3 is regarded as related to gene silencing (41). The H3K27me3 marker is associated with promoters of all hypermethylated genes associated with tumor suppressors in cancer cells (42). With the prevalent regulation executed by H3K27me3 on DEGs screened in the present study, abnormal modification of H3K27me3 may play an important

role in lung adenocarcinoma. Studies have shown that high expression of histone H3K27me3 is related with a good prognosis of patients with NSCLC; namely, the higher expression of histone H3K27me3, the better the prognosis of patients (43). Functional analysis of DEGs and isoforms revealed that they were enriched in the process of hormonal responses. Enrichment analysis of ChIP-seq presented that DEGs were enriched in different ChIP-seq clusters including GR and ER α of TFs. This indicated that regulation of ER and GR may be associated with lung adenocarcinoma. Studies have shown that ER α and ER β , especially ER β , are expressed in NSCLC to induce tumor cell proliferation (44). It was found that midkine plays a pivotal role in epithelial-mesenchymal transition in lung adenocarcinoma (45). Enhanced ER β -mediated estradiol dysregulates midkine expression (46). In previous research, GR, a member of the nuclear hormone receptor family, mediated cancer cell apoptosis and thereby slowed tumor growth (47). GR is down-regulated by increased promoter methylation, which is similar to mechanisms associated with common tumor-suppressor genes (48).

The constructed TF and miRNA regulatory networks showed that hub nodes including miR-126-3p, miR-30c-2-3p, HOXA5, MEIS1 and TBX5 were markedly different in two separate TF and miRNA enrichment analyses, and their levels were significantly decreased in cancer tissues. As crucial upstream genes, their significant change in the expression level may affect a plurality of downstream target genes. Expression of most of the HOX family member are altered in NSCLC cells significantly (49). Contradictory results were found by Abe *et al* (49) who detected the downregulation of HOXA5 in NSCLC. Whether HOXA5 regulates various lung cancer-related genes or what changes it undergoes in lung adenocarcinoma, remains to be elucidated. The relationships of two other TFs including MEIS1 and TBX5 with lung cancer are unclear. It is known that MEIS1 is one of the co-factors of the HOX family (especially for HOXA7 and HOXA9) (50). Together, they are involved in human leukemogenesis (51). TBX5 regulates cell proliferation during cardiogenesis (52) and it is related to cell migration as well as cell proliferation in cancers (53). miR-126 has a clearer relationship with NSCLC and could inhibit the proliferation and invasion of NSCLC cells (54). Downregulation of miR-30c promotes cell migration and invasion of NSCLC cells (55). Despite the fact that altered expression of these two miRNAs has long been known, the exact regulatory mechanisms remain to be studied. The present integrated analysis found that possible target expression levels of these two miRNAs may also undergo significant changes. This indicates one direction for further study.

Differentially methylated regions, differentially expressed miRNAs and transcriptomics of normal and cancer tissues were analyzed. Three possible lung adenocarcinoma-related DEGs including *ERG*, *STARD8* and *THBS2* were identified. Moreover, DEG-related histone modifications and TFs were screened and underwent integrated analysis. Lung adenocarcinoma-related DEGs may be under comparable regulation of histones. Moreover, several TFs and miRNAs may play critical roles in the tumorigenesis of lung adenocarcinoma. These results provide the foundation for further lung adenocarcinoma research, and these results must be confirmed through additional experiments.

References

- Toh CK, Gao F, Lim WT, Leong SS, Fong KW, Yap SP, Hsu AA, Eng P, Koong HN, Thirugnanam A, *et al*: Never-smokers with lung cancer: Epidemiologic evidence of a distinct disease entity. *J Clin Oncol* 24: 2245-2251, 2006.
- Subramanian J and Govindan R: Lung cancer in never smokers: A review. *J Clin Oncol* 25: 561-570, 2007.
- Yano T, Haro A, Shikada Y, Maruyama R and Maehara Y: Non-small cell lung cancer in never smokers as a representative 'non-smoking-associated lung cancer': Epidemiology and clinical features. *Int J Clin Oncol* 16: 287-293, 2011.
- Molina JR, Yang P, Cassivi SD, Schild SE and Adjei AA: Non-small cell lung cancer: Epidemiology, risk factors, treatment, and survivorship. *Mayo Clin Proc* 83: 584-594, 2008.
- Su JL, Shih JY, Yen ML, Jeng YM, Chang CC, Hsieh CY, Wei LH, Yang PC and Kuo ML: Cyclooxygenase-2 induces EP1-and HER-2/Neu-dependent vascular endothelial growth factor-C up-regulation: A novel mechanism of lymphangiogenesis in lung adenocarcinoma. *Cancer Res* 64: 554-564, 2004.
- Skobe M, Hawighorst T, Jackson DG, Prevo R, Janes L, Velasco P, Riccardi L, Alitalo K, Claffey K and Detmar M: Induction of tumor lymphangiogenesis by VEGF-C promotes breast cancer metastasis. *Nat Med* 7: 192-198, 2001.
- Heaney AP, Singson R, McCabe CJ, Nelson V, Nakashima M and Melmed S: Expression of pituitary-tumour transforming gene in colorectal tumours. *Lancet* 355: 716-719, 2000.
- Bevilacqua G, Sobel ME, Liotta LA and Steeg PS: Association of low nm23 RNA levels in human primary infiltrating ductal breast carcinomas with lymph node involvement and other histopathological indicators of high metastatic potential. *Cancer Res* 49: 5185-5190, 1989.
- Billah S, Stewart J, Staerckel G, Chen S, Gong Y and Guo M: EGFR and KRAS mutations in lung carcinoma: Molecular testing by using cytology specimens. *Cancer Cytopathol* 119: 111-117, 2011.
- Ju YS, Lee WC, Shin JY, Lee S, Bleazard T, Won JK, Kim YT, Kim JI, Kang JH and Seo JS: A transforming KIF5B and RET gene fusion in lung adenocarcinoma revealed from whole-genome and transcriptome sequencing. *Genome Res* 22: 436-445, 2012.
- Togashi Y, Soda M, Sakata S, Sugawara E, Hatano S, Asaka R, Nakajima T, Mano H and Takeuchi K: KLC1-ALK: A novel fusion in lung cancer identified using a formalin-fixed paraffin-embedded tissue only. *PLoS One* 7: e31323, 2012.
- Kim SC, Jung Y, Park J, Cho S, Seo C, Kim J, Kim P, Park J, Seo J, Kim J, *et al*: A high-dimensional, deep-sequencing study of lung adenocarcinoma in female never-smokers. *PLoS One* 8: e55596, 2013.
- Bolger AM, Lohse M and Usadel B: Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics* 30: 2114-2120, 2014.
- Langmead B and Salzberg SL: Fast gapped-read alignment with Bowtie 2. *Nat Methods* 9: 357-359, 2012.
- Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G and Durbin R; 1000 Genome Project Data Processing Subgroup: The sequence alignment/map format and SAMtools. *Bioinformatics* 25: 2078-2079, 2009.
- Lienhard M, Grimm C, Morkel M, Herwig R and Chavez L: MEDIPS: Genome-wide differential coverage analysis of sequencing data derived from DNA enrichment experiments. *Bioinformatics* 30: 284-286, 2014.
- Quinlan AR and Hall IM: BEDTools: A flexible suite of utilities for comparing genomic features. *Bioinformatics* 26: 841-842, 2010.
- Trapnell C, Pachter L and Salzberg SL: TopHat: Discovering splice junctions with RNA-Seq. *Bioinformatics* 25: 1105-1111, 2009.
- Anders S, Reyes A and Huber W: Detecting differential usage of exons from RNA-seq data. *Genome Res* 22: 2008-2017, 2012.
- Trapnell C, Williams BA, Pertea G, Mortazavi A, Kwan G, van Baren MJ, Salzberg SL, Wold BJ and Pachter L: Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat Biotechnol* 28: 511-515, 2010.
- Wang WC, Lin FM, Chang WC, Lin KY, Huang HD and Lin NS: miRExpress: Analyzing high-throughput sequencing data for profiling microRNA expression. *BMC Bioinformatics* 10: 328, 2009. doi:10.1186/1471-2105-10-328.
- Anders S, McCarthy DJ, Chen Y, Okoniewski M, Smyth GK, Huber W and Robinson MD: Count-based differential expression analysis of RNA sequencing data using R and Bioconductor. *Nat Protoc* 8: 1765-1786, 2013.
- Lewis BP, Burge CB and Bartel DP: Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets. *Cell* 120: 15-20, 2005.
- Huang DW, Sherman BT, Tan Q, Kir J, Liu D, Bryant D, Guo Y, Stephens R, Baseler MW, Lane HC, *et al*: DAVID Bioinformatics Resources: Expanded annotation database and novel algorithms to better extract biology from large gene lists. *Nucleic Acids Res* 35 (Web Server issue): W169-W175, 2007.
- Vergoulis T, Vlachos IS, Alexiou P, Georgakilas G, Maragkakis M, Reczko M, Gerangelos S, Koziris N, Dalamagas T and Hatzigeorgiou AG: TarBase 6.0: Capturing the exponential growth of miRNA targets with experimental support. *Nucleic Acids Res* 40: D222-D229, 2012.
- Vlachos IS, Kostoulas N, Vergoulis T, Georgakilas G, Reczko M, Maragkakis M, Paraskevopoulou MD, Prionidis K, Dalamagas T and Hatzigeorgiou AG: DIANA miRPath v.2.0: Investigating the combinatorial effect of microRNAs in pathways. *Nucleic Acids Res* 40: W498-W504, 2012.
- Lachmann A, Xu H, Krishnan J, Berger SI, Mazloom AR and Ma'ayan A: ChEA: Transcription factor regulation inferred from integrating genome-wide ChIP-X experiments. *Bioinformatics* 26: 2438-2444, 2010.
- Consortium EP; ENCODE Project Consortium: The ENCODE (ENCyclopedia Of DNA Elements) Project. *Science* 306: 636-640, 2004.
- Huang GT, Athanassiou C and Benos PV: mirConnX: Condition-specific mRNA-microRNA network integrator. *Nucleic Acids Res* 39 (suppl): W416-W423, 2011.
- Zhang WC, Shyh-Chang N, Yang H, Rai A, Umashankar S, Ma S, Soh BS, Sun LL, Tai BC, Nga ME, *et al*: Glycine decarboxylase activity drives non-small cell lung cancer tumor-initiating cells and tumorigenesis. *Cell* 148: 259-272, 2012.
- Joshi M, Liu X and Belani CP: Taxanes, past, present, and future impact on non-small cell lung cancer. *Anticancer Drugs* 25: 571-583, 2014.
- Joseph P, Bowman B, Galban S and Rehemtulla A: A Kras driven murine model elucidates oncogenic role of FADD in lung cancer. Department of Radiation Oncology, University of Michigan, Ann Arbor, MI, 2013 (Honor's Bachelor's Thesis). <http://deepblue.lib.umich.edu/bitstream/handle/2027.42/98916/parijo.pdf?sequence=1>.
- Rauch TA, Wang Z, Wu X, Kernstine KH, Riggs AD and Pfeifer GP: DNA methylation biomarkers for lung cancer. *Tumour Biol* 33: 287-296, 2012.
- Reddy ES, Rao VN and Papas TS: The erg gene: A human gene related to the ets oncogene. *Proc Natl Acad Sci USA* 84: 6131-6135, 1987.
- Thoms JA, Birger Y, Foster S, Knezevic K, Kirschenbaum Y, Chandrakanthan V, Jonquieres G, Spensberger D, Wong JW, Oram SH, *et al*: ERG promotes T-acute lymphoblastic leukemia and is transcriptionally regulated in leukemic cells by a stem cell enhancer. *Blood* 117: 7079-7089, 2011.
- Leyten GH, Hessels D, Jannink SA, Smit FP, de Jong H, Cornel EB, de Reijke TM, Vergunst H, Kil P, Knipscheer BC, *et al*: Prospective multicentre evaluation of PCA3 and TMPRSS2-ERG gene fusions as diagnostic and prognostic urinary biomarkers for prostate cancer. *Eur Urol* 65: 534-542, 2012.
- Durkin ME, Ullmannova V, Guan M and Popescu NC: Deleted in liver cancer 3 (DLC-3), a novel Rho GTPase-activating protein, is downregulated in cancer and inhibits tumor cell growth. *Oncogene* 26: 4580-4589, 2007.
- Mokarram P, Kumar K, Brim H, Naghibalhossaini F, Saberi-firoozi M, Nouraei M, Green R, Lee E, Smoot DT and Ashktorab H: Distinct high-profile methylated genes in colorectal cancer. *PLoS One* 4: e7012, 2009.
- Bornstein P: Thrombospondins function as regulators of angiogenesis. *J Cell Commun Signal* 3: 189-200, 2009.
- Brezillon S, Zeltz C, Schneider L, Terryn C, Vuillermoz B, Ramont L, Perrau C, Pluot M, Diebold MD, Radwanska A, *et al*: Lumican inhibits B16F1 melanoma cell lung metastasis. *J Physiol Pharmacol* 60 (Suppl 4): 15-22, 2009.
- Barski A, Cuddapah S, Cui K, Roh TY, Schones DE, Wang Z, Wei G, Chepelev I and Zhao K: High-resolution profiling of histone methylations in the human genome. *Cell* 129: 823-837, 2007.

42. Hussain M, Rao M, Humphries AE, Hong JA, Liu F, Yang M, Caragacianu D and Schrupp DS: Tobacco smoke induces polycomb-mediated repression of Dickkopf-1 in lung cancer cells. *Cancer Res* 69: 3570-3578, 2009.
43. Chen X, Song N, Matsumoto K, Nanashima A, Nagayasu T, Hayashi T, Ying M, Endo D, Wu Z and Koji T: High expression of trimethylated histone H3 at lysine 27 predicts better prognosis in non-small cell lung cancer. *Int J Oncol* 43: 1467-1480, 2013.
44. Bogush TA, Dudko EA, Beme AA, Bogush EA, Kim AI, Polotsky BE, Tjuljandin SA and Davydov MI: Estrogen receptors, antiestrogens, and non-small cell lung cancer. *Biochemistry (Mosc)* 75: 1421-1427, 2010.
45. Mak P, Chang C, Pursell B and Mercurio AM: Estrogen receptor β sustains epithelial differentiation by regulating prolyl hydroxylase 2 transcription. *Proc Natl Acad Sci USA* 110: 4708-4713, 2013.
46. Zhao G, Nie Y, Lv M, He L, Wang T and Hou Y: ER β -mediated estradiol enhances epithelial mesenchymal transition of lung adenocarcinoma through increasing transcription of midkine. *Mol Endocrinol* 26: 1304-1315, 2012.
47. Schlossmacher G, Stevens A and White A: Glucocorticoid receptor-mediated apoptosis: Mechanisms of resistance in cancer cells. *J Endocrinol* 211: 17-25, 2011.
48. Koutsimpelas D, Pongsapich W, Heinrich U, Mann S, Mann WJ and Brieger J: Promoter methylation of MGMT, MLH1 and RASSF1A tumor suppressor genes in head and neck squamous cell carcinoma: Pharmacological genome demethylation reduces proliferation of head and neck squamous carcinoma cells. *Oncol Rep* 27: 1135-1141, 2012.
49. Abe M, Hamada J, Takahashi O, Takahashi Y, Tada M, Miyamoto M, Morikawa T, Kondo S and Moriuchi T: Disordered expression of HOX genes in human non-small cell lung cancer. *Oncol Rep* 15: 797-802, 2006.
50. Svingen T and Tonissen KF: Altered HOX gene expression in human skin and breast cancer cells. *Cancer Biol Ther* 2: 518-523, 2003.
51. Argiropoulos B, Yung E and Humphries RK: Unraveling the crucial roles of Meis1 in leukemogenesis and normal hematopoiesis. *Genes Dev* 21: 2845-2849, 2007.
52. Hatcher CJ, Kim M-S, Mah CS, Goldstein MM, Wong B, Mikawa T and Basson CT: TBX5 transcription factor regulates cell proliferation during cardiogenesis. *Dev Biol* 230: 177-188, 2001.
53. Aherne NJ, Rangaswamy G, and Thirion P: Prostate cancer in a male with Holt-Oram syndrome: First clinical association of the TBX5 mutation. *Case Rep Urol* 2013: 405343, 2013.
54. Miko E, Margitai Z, Czimmerer Z, Várkonyi I, Dezsó B, Lányi A, Bacsó Z and Scholtz B: miR-126 inhibits proliferation of small cell lung cancer cells by targeting SLC7A5. *FEBS Lett* 585: 1191-1196, 2011.
55. Zhou H, Xu X, Xun Q, Yu D, Ling J, Guo F, Yan Y, Shi J and Hu Y: microRNA-30c negatively regulates endometrial cancer cells by targeting metastasis-associated gene-1. *Oncol Rep* 27: 807-812, 2012.