European Review for Medical and Pharmacological Sciences 2014; 18: 217-228

Gene expression analysis of lung cancer

W.-Y. SHI, K.-D. LIU, S.-G. XU, J.-T. ZHANG, L.-L. YU, K.-Q. XU, T.-F. ZHANG

Department of Respiratory Disease, Dachang Hospital, Baoshan District, Shanghai, China
Weiyi Shi and Kedi Liu should be regarded as co-first authors

Abstract. – OBJECTIVE: We aim to explore the expression difference between lung cancer cells and normal lung cells, and to investigate the mechanism of lung cancer development. Besides, we predicted the potential target site of transcriptional factors and microRNAs for differentially expressed genes (DEGs), which may help to regulate expression of DEGs. Small molecules were also identified to cure lung cancer.

MATERIALS AND METHODS: Gene expression profiles we used were downloaded from Gene Expression Omnibus (GEO) using accession number of GSE2378. Firstly, we identified differential genes between lung cancer cells and normal lung cells by using R package limma. Then, we detected the processes and pathways that changed in lung cancer cells by Gene Ontology (GO) and KEGG pathway enrichment analysis. Potential target sites of transcriptional factors and microRNAs were also detected based on gene annotation data in MSigDB. Finally, small molecule drugs were screened via querying Connectivity Map database.

RESULTS: We obtained 2961 differentially expressed genes between lung cancer cells and normal lung cells. Besides changes in cell cycle, metabolic processes and proteasome were also dramatically disordered. Some DEGs shared target sites of the transcription factor such as E2F, ETS and CEBPB. Target sites of hsa-miR-196a and hsa-miR-200c were also significantly enriched by DEGs. Iloprost simulated the state of normal cells, while MS-275 might be potential pathogenic substances.

CONCLUSIONS: We investigate the lung cancer from Gene Ontology, pathway, transcription factors and microRNAs based on gene expression profiles. All these results may facilitate lung cancer treatment with a new breakthrough.

Key Words:
Lung cancer, Transcription factor, microRNA, Pathway, Gene ontology, Small molecules.

Introduction

Lung cancer is the uncontrolled growth of abnormal cells in one or both lungs[1]. These abnormal cells do not carry out the functions of normal cells and do not develop into healthy lung tissue. As they grow, the abnormal cells can form tumors[2]. 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[3]. 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[4]. Therefore, the research and treatment of lung cancer is of great significance to human health.

Genome-wide transcriptional expression analysis is a powerful strategy for characterizing the biological activity of disease phenotypes. Gene ontology (GO) is utilized to characterize the function categories affected by lung cancer. It determines the functional categories of lung cancer related genes. Genes that differentially expressed between normal lung tissue and cancer show enrichment in gene ontology terms associated with mitosis and proliferation[5]. 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"[6]. Methylated genes are significantly enriched as transcription factors (TFs) and in processes of neuronal differentiation[7]. Gene ontology characterize that epigenetic genes shows are almost exclusively involved in morphogenetic differentiation processes in lung cancer[8].

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 significantly related with lung cancer susceptibility[9]. DNA copy number aberrations in small-cell lung cancer are significantly activated by the focal adhesion pathway[10]. MicroRNA signatures in tumor tissue are related to pathway "angiogenesis" in non-small cell lung cancer evaluated by Gene Set Enrichment Analysis (GSEA)[11]. In the final integrative analysis of lung cancer related miR-21-targets analysis, 24 hub genes are identified by overlap

Corresponding Author: Tiefeng Zhang, MD; e-mail: zhangtief@hotmail.com

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[12].

Based on the genes related with disease phenotype, researchers screen drugs to repress these genes. The Connectivity Map is a collection of genome-wide transcriptional expression data from cultured human cells treated with bioactive small molecules and simple pattern-matching algorithms that together enable the discovery of functional connections between drugs, genes and diseases through the transitory feature of common gene-expression changes[13]. Drug versus Disease (DvD) provides a pipeline for the comparison of drug and disease gene expression profiles from public microarray repositories[14]. Yeh et al[15] screen drugs that target cancer stem cells (CSCs) to improve the current treatment and overcome drug resistance. They use gene signatures between embryonic stem cell and CSC to identify potential drugs that can reverse the gene expression profile of CSCs based on Connectivity Map. They repurpose trifluoperazine as a potential anti-cancer stem cell agent that could overcome EGFR-tyrosine kinase inhibitor and chemotherapy resistance. There is no standard therapy for cholangiocarcinoma (CCA). Chen et al[16] hypothesize that, if a drug could reverse the gene expression signature of CCA, then it may inhibit the carcinogenesis of CCA. After comparing the expression pattern of CCA-related genes and genes perturbed by small molecules in connectivity map, NVP-AUY922 is regarded as an effective treatment option for patients with CCA. Combining the query signature with correlating profiles allows for the definition of a tight neurodegeneration signature that successfully highlights many neuroprotective drugs in the Broad connectivity map[17]. Using microarray technology, Claerhout et al[18] generate a gene expression profile of human gastric cancer-specific genes from human gastric cancer tissue samples. Then they use profiles in the Broad Institute's Connectivity Map analysis to identify vorinostat as a candidate therapy for gastric cancer.

Many diseases are caused by hereditary mutations. An increasing number of the identified disease-related mutations occur in gene regulatory sequences. Laurila and Lähdesmäki[19] 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[20]. Mutation of HBD also affect its binding with transcriptional factor GATA1, which finally leads to δ-thalassemia[21]. Respiratory and related diseases are affected by modified transcription regulation programs. Many TFs are significantly enriched in the target disease groups[22]. MicroRNA-183 regulates Ezrin expression in lung cancer cells[23]. miR-93, miR-98, and miR-197 regulate expression of tumor suppressor gene FUS1[24]. MiR-21 is an EGFR-regulated anti-apoptotic factor in lung cancer in never-smokers[25].

In this study, we aim to integrate information of functional annotation, transcription factors and microRNAs to elaborate the mechanism of lung cancer development. Moreover, we screened small molecules for lung cancer treatment based on the differential genes. The potential target sites of microRNAs and transcriptional factors were also detected, helping to regulate expression of differential genes.

Materials and Methods

Data Source

To gain insight into the changes in lung cancer cells, we intended to explore disease mechanism from the gene level by identifying differentially expressed genes (DEGs). We collected cells from lung cancer samples and normal lung samples. The cells are then used for microarray samples to get gene expression profiles. The microarray data were downloaded from Gene Expression Omnibus (GEO, http://www.ncbi.nlm.nih.gov/geo/, accession number GSE2378[26]). In this dataset, it includes five normal samples and five lung cancer samples. Microarray platform is GPL96 [HG_U133A].

Screen of Differential Genes

We used R package (Geoquery[27] and limma[28]) to analyze chip data (version 2.13.0). Geoquery can quickly query microarray data from GEO (Gene Expression Omnibus) database. Limma is the most popular method to identify differentially expressed genes. First, the preprocessed microarray data were obtained by Geoquery. Then log2 transformation was performed on these data. Up- and down-regulated genes were identified between lung cancer cells and normal lung cells using R package limma[28].

Detecting Significant Gene Ontology (GO) Terms of DEGs

To explore changes of DEGs occurring at the cellular level and detect the function of DEGs, we elaborated gene function and cellular location based on Gene Otology (GO)[29]. GO enrichment analysis was performed using GOEAST[30]. If DEGs contain valid IDs of *k* genes from a microarray with a total of *t* genes, for a given GO term, there are *q* genes within *k* and *m* genes within *t*; then, the possibility that whether genes associated with this GO term is enriched by hypergeometric test.

$$p\,(X = x > q) = \sum_{x=q}^{m} \frac{\binom{m}{x}\binom{t-m}{k-x}}{\binom{t}{k}}$$

We analyzed how DEGs to impact cells from three ontologies of GO, that are biological process, molecular function and cellular component. In this way, we investigated how the DEGs affected the lung cancer cells.

Detecting Significant Pathways

After elaborating gene function and cellular location based on Gene Ontology, we further explored how these DEGs interact with other genes in a pathway. We obtained all the metabolic and non-metabolic pathways from kegg pathway database. Then, we identified the enriched pathways of DEGs using Gene Set Analysis Toolkit V2[31,32]. Pathway enrichment analysis was performed using hypergeometric test. If DEGs contain valid IDs of *k* genes from a microarray with a total of *t* genes, for a given pathway, there are *q* genes within *k* and *m* genes within *t*, the possibility is calculated as following,

$$p\,(X = x > q) = \sum_{x=q}^{m} \frac{\binom{m}{x}\binom{t-m}{k-x}}{\binom{t}{k}}$$

In this way, we identified pathways that were disturbed in lung cancer cells.

Potential Transcription Factor Target Sites

We performed gene enrichment analysis based on gene annotation data in MSigDB (http://www.broadinstitute.org/gsea/msigdb/index.jsp). The statistical method used was hypergeometric algorithm, and BH (Benjamini-Hochberg) adjustment was used for multiple test correction[33]. Finally, we obtained the potential target sites that regulate transcription factors.

Potential microRNA Target Sites

We performed gene enrichment analysis based on gene annotation data in MSigDB (http://www.broadinstitute.org/gsea/msigdb/index.jsp). The statistical method used was hypergeometric algorithm, and BH adjustment was used for multiple test correction. Finally, we obtained the potential target sites of microRNAs that regulate expression of DEGs.

Expression Profiles of Cell Lines Perturbed by Small Molecules

Connectivity Map[13] (http://www.broad.mit.edu/Connectivity Map/) consists of more than 7,000 gene expression profiles treated with 1,309 small molecules. These expression profiles represent about 6,000 instances, each of which comprises a treatment and vehicle pair. By comparing the expression pattern similarity of DEGs and the genes perturbed in Connectivity Map instances, a list of molecules related to the input genes will be identified.

Differentially expressed genes were partitioned into up- or down-regulated groups. KS scores for both up () and down () were calculated as follows:

$$a = \max_{j=1}^{t}\left[\frac{j}{t} - \frac{V(j)}{N}\right], \quad b = \max_{j=1}^{t}\left[\frac{V(j)}{N} - \frac{j-1}{t}\right]$$

$$KS_{up/down} = \begin{cases} a, (a > b) \\ -b, (b > a) \end{cases}$$

Where, *t* is the number of genes in either the up- or down-regulated gene group, *N* is the total number genes in array, *j* denotes the jth gene in the rank ordered up- or down-regulated groups according to the extent of differential expression, *V(j)* denotes the position of the jth gene in the rank ordered whole gene list (also ranked according to the extent of differential expression). The connectivity *S* score is set to zero where KS_{up} and KS_{down} have the same sign. Otherwise, set *S* to be KS_{up}-KS_{down}.

We intend to identify small molecules with similar or opposite effect with disease DEGs. *S*

ranges from -1 to 1. If the *S* is close to 1, it indicates that this small molecule can simulate the state of normal lung cells. In contrary, if the *S* is close to -1, it indicates that this small molecule can simulate the disease state of lung cancer cells.

Results

Differentially Expressed Genes Between Lung Cancer Cells and Normal Lung Cells

We identified differential genes between lung cancer cells and normal lung cells using linear models and empirical Bayes methods to assess differentially expressed genes[28]. Benjamini-Hochberg false discovery rate (BH-FDR)[33] was performed for multiple testing correction. All genes were ranked by *p*-value. Genes with $p < 0.05$ were regarded as differential genes. Finally, 3730 probes exhibited differential expression, which involved 2961 genes.

Enriched GO Terms of DEGs

Differentially expressed genes were enriched into GO terms using GOEAST. DEGs enriched cellular components are shown in Figure 1. DEGs enriched molecular function is shown in Figure 2. DEGs enriched biological process is shown in Figure 3. The GO enrichment results preliminarily interpret the impact of the differentially expressed gene on biological function. From the perspective of molecule function, the DEGs influence protein-protein binding and the activity of the respiratory chain, including the transmission of oxygen as well as of NAD + metabolism. These indicate that lung cancer cells differ from norm lung cells in the energy metabolism.

From the perspective of the cellular component, changes occur in microtubule, centromere and chromosome relevant parts. These parts are closely related to cell proliferation, which is also consistent with the fact that cancer cell proliferation is abnormal and uncontrollable. Besides, we also identified the mitochondrial changes, which is consistent with the result of the molecular function.

From the perspective of biological process, the changes of lung cancer cells involves multiple aspects, including cell division, cell cycle, organ development and regeneration, lymphocyte differentiation and maturation, cell signal transduction and metabolism of sugar, protein and fat. These suggest that differences do exist in lung cancer cells and normal cells.

Pathways Disturbed in Lung Cancer Cells

In order to explore the local changes, we further investigated the changes in pathways. We utilized hypergeometric algorithm for KEGG pathway enrichment and BH method was used for *p* value correction. All the pathways with corrected $p < 0.05$ and at least two DEGs were regarded as significant signaling pathways. These pathways are listed in Table I.

KEGG pathway enrichment further confirmed the speculation of GO enrichment. DNA replication and cell cycle confirmed the strong proliferation of lung cancer cells. Changes in mismatch repair indicated that autonomous DNA damage repair ability was changed in lung cancer cells.

In addition, changes in KEGG pathway, such as proteasome, metabolic pathways and glutathione metabolism, further confirmed the inference of GO enrichment that cell metabolism changed in lung cancer cells. Wherein it is worthy noted that changes in pathway proteasome may lead to degradation of intracellular tumor suppressor proteins (e.g., p53), or the degradation of the cell cycle inhibitory factor. All these implications provide some clues for lung cancer development. Changes in pathway glutathione metabolism indicated alteration occurred in glutathione or its metabolites, which makes it possible to early diagnose lung cancer.

Potential Transcription Factor Target Sites

Studies on molecular differences between lung cancer cells and normal lung cells offer the possibility to understand lung cancer. The analysis of the lung cancer DEGs also facilitates to elucidate the mechanism of lung cancer. Exploring targets of transcription factor may fundamentally describe the occurrence of lung cancer, because transcription factors are important regulatory elements for downstream gene. Here we analyzed upstream sequences of DEGs and explored the potential target sites of the transcription factors. The results are shown in Table II. For example, the target site of transcription factor E2F, ETS and CEBPB were significantly enriched by the DEGs.

Potential microRNAs that Target DEGs

Cells generally regulate gene translation in two ways. One is via regulatory elements such as transcription factors regulating DNA transcription into RNA. The other is via elements such as microRNAs regulating RNA stability. Therefore,

Figure 1. Enriched cellular components of DEGs. Significantly enriched terms are colored yellow (FDR < 0.05). The more yellow the color is, the more significant the term is.

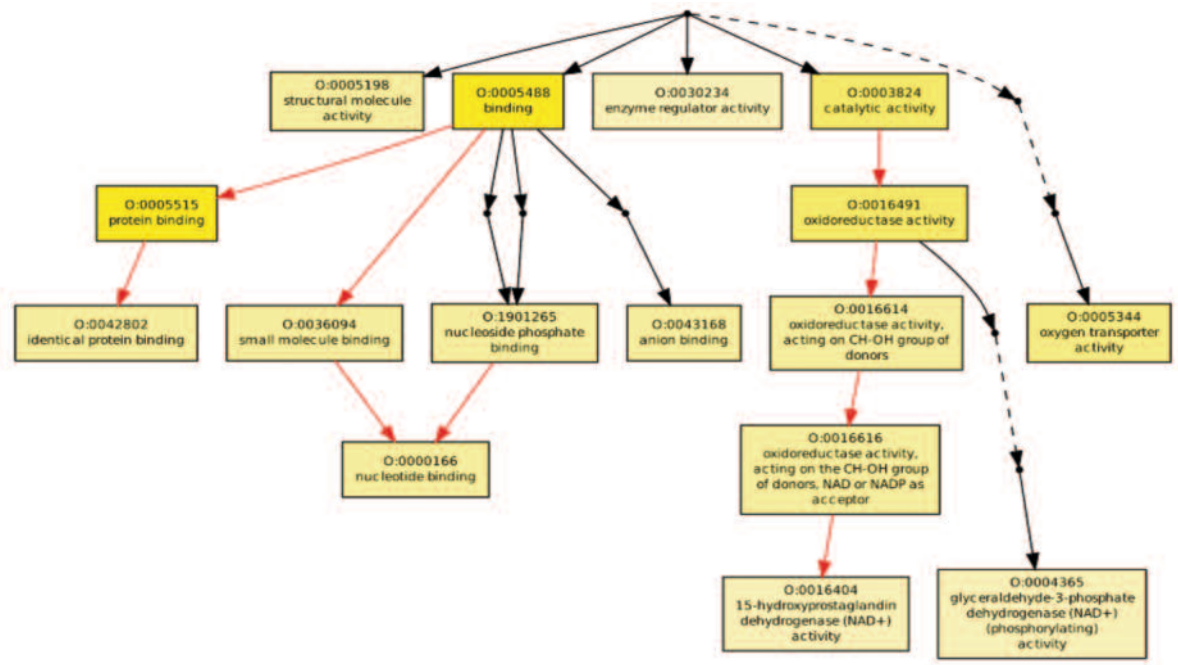

Figure 2. Enriched molecular function of DEGs. Significantly enriched terms are colored yellow (FDR < 0.05). The more yellow the color is, the more significant the term is.

it is of particular significance to identify the microRNAs that regulate DEGs. Here we analyze potential microRNAs based on the gene sequence of DEGs. The results are shown in Table III. For example, the target site of hsa-miR-196a and hsa-miR-200 c were significantly enriched by the DEGs.

Lung Cancer Related Small Molecules

We hope our studies on lung cancer help the treatment of lung cancer, which include exploring possible small molecule drugs for lung cancer treatment. DEGs were divided into two types, up-regulated and down-regulated. By comparing the expression pattern of DEGs and that of genes perturbed by small molecules, we obtained lung cancer related small molecules. The most correlated 20 small molecules are listed in Table IV.

From Table IV, small molecules iloprost (enrichment = 0.947) could simulate the state of normal cells, that is to say, it may be potentially used for lung cancer treatment. At the same time, small molecule MS-275 (enrichment = –0.998) could simulate the cell state of lung cancer cells, that is to say MS-275 may be potential pathogenic substances. They may provide good model to study the mechanism of lung cancer.

Discussion

Lung cancer is the most common primary pulmonary malignant tumors, generally originate in the bronchial epithelium[34]. During the past 50 years, all over the world particularly in industrial countries, the incidence and mortality of lung cancer has rapidly rise[35]. Therefore, it is of great significance to research on treatment of lung cancer.

In this study, we identified 2961 differentially expressed genes (DEGs) based on gene expression profile of lung cancer cells and normal lung cells. It indicates that the DEGs imply the occur-

Table I. Significant pathways that change in lung cancer cells.

KEGG pathway	*p* value
DNA replication	5.17E-06
Cell cycle	3.61E-05
Mismatch repair	0.0056
Proteasome	0.0091
Systemic lupus erythematosus	0.0109
Metabolic pathways	0.0243
Glutathione metabolism	0.026
Alzheimer's disease	0.0683
Type I diabetes mellitus	0.0829
Glycolysis/gluconeogenesis	0.0893

Figure 3. Enriched biological processes of DEGs. Significantly enriched terms are colored yellow (FDR < 0.05). The more yellow the color is, the more significant the term is.

Table II. Potential transcription factor target sites.

Target	*p*-value	Target	*p*-value
a_V$E2F_Q6_01	0.0012	hsa_V$CP2_01	0.0123
hsa_V$ZF5_01	0.0012	hsa_V$E2F1DP2_01	0.0123
hsa_TTGTTT_V$FOXO4_01	0.0012	hsa_V$HNF3ALPHA_Q6	0.0136
hsa_V$TCF11MAFG_01	0.0026	hsa_RYTTCCTG_V$ETS2_B	0.0136
hsa_V$E2F_Q4_01	0.0026	hsa_GGGYGTGNY_UNKNOWN	0.0175
hsa_V$E2F_03	0.0039	hsa_V$ETS_Q4	0.0184
hsa_TGANTCA_V$AP1_C	0.0046	hsa_V$E2F_Q6	0.0184
hsa_V$E2F1DP1RB_01	0.0068	hsa_V$CEBPB_01	0.0198
hsa_V$E2F1_Q6	0.0068	hsa_V$CEBP_Q2	0.0241
hsa_GGGCGGR_V$SP1_Q6	0.0102	hsa_V$TFIIA_Q6	0.0241
hsa_V$HFH4_01	0.0102	hsa_TGANNYRGCA_V$TCF11MAFG_01	0.0288
hsa_CTTTGT_V$LEF1_Q2	0.0102	hsa_V$FOXD3_01	0.0315
hsa_V$NFE2_01	0.0123	hsa_GGGAGGRR_V$MAZ_Q6	0.0324
hsa_V$E2F_Q4	0.0123	hsa_V$AP1_Q2_01	0.0331
hsa_TGGAAA_V$NFAT_Q4_01	0.0123	hsa_V$FOXO4_01	0.0337
hsa_V$E2F4DP2_01	0.0123	hsa_MGGAAGTG_V$GABP_B	0.0337
hsa_V$E2F_Q3_01	0.0123	hsa_V$ELF1_Q6	0.0337
hsa_V$E2F4DP1_01	0.0123	hsa_V$AHR_01	0.0389
hsa_V$E2F1_Q4_01	0.0123	hsa_V$ELK1_01	0.0409
hsa_V$E2F1DP1_01	0.0123	hsa_.V$CP2_02	0.0428
hsa_WGGAATGY_V$TEF1_Q6	0.0123	hsa_V$SRF_C	0.0432
hsa_SGCGSSAAA_V$E2F1DP2_01	0.0123	hsa_GGGTGGRR_V$PAX4_03	0.045
hsa_V$E2F_02	0.0123		

rence mechanism of lung cancer or even the key points of the lung cancer treatment. After conducting functional analysis from the perspective of Gene Ontology (GO) and KEGG pathway, a number of changes occur in cell cycle and metabolic processes. Some DEGs shared target sites of the transcription factor and microRNAs. Further, small molecule drugs, which repair the disordered metabolic pathways, were identified for lung cancer treatment.

From GO enrichment analysis, the DEGs influence cell cycle and metabolic process[36,37], which is reported to be related to cancer progression. The Gene Ontology term "protein binding" becomes the predominant reaction with phenethyl isothiocyanate (PEITC) that accounts for 87% of the total uptake but only 12% for sulforaphane (SFN)[38]. These indicate that lung cancer cells differ from norm lung cells in the energy metabolism. From the aspect of the cellular component, changes occur in microtubule. Taxol, a chemotherapeutic drug for cancer treatment, binds to microtubules and abolish their dynamic behavior[39]. Centromere protein or protein of centromere complexes are also related to carcinogenesis and progression of lung cancer[40,41]. Changes also occur in the connecting portion between cells, which may explain morphology changes of cancer cells.

KEGG pathway analysis confirms that DNA replication and cell cycle are correlated with lung cancer. Changes in mismatch repair (MMR) indicate that autonomous DNA damage repair ability changes in lung cancer cells. Hansen et al[42] investigate the role of MMR in small-cell lung cancer cells. They propose two possibilities to explain the controversial findings in different samples. It is worthy noted that changes occur in pathway Proteasome. The proteasome constitutes the main protein waste disposal and recycling system lung cells. In lung cancer cells, the proteasome is a promising therapeutic target[43,44]. Moreover, the proteasome is involved in lung pathogenesis[45]. Common genetic variations in

Table III. Potential microRNAs that target differentially expressed genes.

Target sequence	Potential microRNA	*p*-value
hsa_ACTACCT	MIR-196A, MIR-196B	0.021
hsa_CAGTATT	MIR-200B, MIR-200C, MIR-429	0.021
hsa_TGCCTTA	MIR-124A	0.021
hsa_AATGTGA	MIR-23A,MIR-23B	0.021
hsa_GTGCCAA	MIR-96	0.021
hsa_TTGGGAG	MIR-150	0.024
hsa_CATTTCA	MIR-203	0.024

Table IV. List of lung cancer related small molecules.

cmap name	Enrichment	p
Vorinostat	-0.73	0
Trichostatin A	-0.708	0
Thioridazine	-0.57	0
LY-294002	-0.446	0
Sirolimus	-0.358	0
MS-275	-0.998	0.00004
Prestwick-675	0.918	0.00004
Prestwick-857	0.902	0.0001
Prochlorperazine	-0.532	0.0001
Trifluoperazine	-0.525	0.00012
Tanespimycin	-0.276	0.00012
Iloprost	0.947	0.00018
Mebendazole	-0.82	0.00046
Atractyloside	0.813	0.00058
PHA-00745360	0.65	0.00084
Puromycin	-0.851	0.00092
Scriptaid	-0.91	0.00128
Phenoxybenzamine	-0.83	0.00151
Niclosamide	-0.757	0.00154
Cefamandole	0.818	0.00193

genes of EGFR and glutathione pathways may be associated with overall survival among patients with advanced stage NSCLC treated with platinum, taxane, and/or gemcitabine combinations[46]. Changes in pathway glutathione metabolism indicate alteration occurs in glutathione or its metabolites, which makes it possible to early diagnose of lung cancer.

Moreover, the DEGs significantly share target sites of transcription factors and microRNA. These sites play an important role for the regulation of gene expression. The putative control of Gpr19 expression by E2F transcription factors is verified by chromatin immunoprecipitation in small cell lung cancer[47]. Transcription factor CEBPB and ets variant 4 (ETV4) are the hub nodes in the transcriptome network of human lung cancer[48]. Rs11614913 is a polymorphism in hsa-miR-196a2 reported to alter mature microRNA expression and function. It is reported to be associated with severe toxicity in lung cancer patients[49]. The expression level of miR-200c in non-small-cell lung cancer A549 cells is low in contrast to normal human bronchial epithelial cells[50]. The transcription factor and microRNA analysis may provide an improved understanding of the molecular mechanisms and potential therapeutic targets in the treatment of lung cancer.

A series of small molecules were identified based on expression pattern of the DEGs and genes perturbed by small molecules in Connectivity Map. These small molecules can simulate the states of normal lung cells compared to lung cancer cells, which means that these small molecules might repair the disordered pathways. Small molecules iloprost (enrichment = 0.947) could simulate the state of normal cells. Iloprost is regarded as a promising agents for lung cancer[51]. The trial of the prostacyclin analog iloprost recently reported 20%-30% lung cancer mortality reductions[52]. Keith et al[54] carry out a multicenter double-blind, randomized, phase II placebo-controlled trial of oral iloprost in current or former smokers with sputum cytologic atypia or endobronchial dysplasia. When compared with placebo, former smokers receiving oral iloprost exhibited a significantly greater improvement[53]. Iloprost also can improve bronchial histology. iloprost is able to inhibit transformed growth of human non-small cell lung cancer (NSCLC)[55]. At the same time, small molecule MS-275 (enrichment = –0.998) could simulate the cell state of lung cancer cells. The histone deacetylase (HDAC) inhibitor MS-275 reverses CSC-induced epithelial-to-mesenchymal transition (EMT), migration, and invasion through the restoration of E-cadherin expression[56]. MS-275 restored E-cadherin expression and moderate sensitivity to erlotinib in HCC4006ER cells[57]. Treatment with MS-275 was most effective in E-cadherin up-regulation and persistence in non-small cell lung cancers. As with other tumor types and HDAC inhibitors, MS-275 inhibited growth and induced apoptosis. Importantly, blocking E-cadherin induction by short hairpin RNA resulted in less inhibition by MS-275, implicating the epithelial to mesenchymal phenotype process as a contributing factor[58]. MS-275 effect is related to an HDAC tethered to the HLA-DRA promoter by the transcription factor YY1[59]. The combination of 5AZA-CdR and MS-275 has been reported in patients with advanced Ewing's sarcoma[60]. Such small molecules provide new possibility for lung cancer treatment. It is also found many DEGs have the same target sites of the transcription factor. These sites play a vital role in the regulation of gene expression.

Conclusions

Overall, through the functional analysis of DEGs, we identified cell cycle, metabolic pathways and proteasome as the significant terms for lung cancer. The target sites of transcriptional

factors and microRNAs play vital roles in the regulation of gene expression. Although it may be premature to suggest that these screened small molecules might be ready for lung cancer clinical trials, it is clearly a direction that warrants additional consideration. All these results may facilitate lung cancer treatment with a new breakthrough.

Conflict of Interest

The Authors declare that there are no conflicts of interest.

Reference

1) Sterner-Kock A, Thorey IS, Koli K, Wempe F, Otte J, Bangsow T, Kuhlmeier K, Kirchner T, Jin S, Keski-Oja J, von Melchner H. Disruption of the gene encoding the latent transforming growth factor-beta binding protein 4 (LTBP-4) causes abnormal lung development, cardiomyopathy, and colorectal cancer. Genes Dev 2002; 16: 2264-2273.
2) Dang SP, Liberman BA, Shepherd FA, Messner H, Tweeddale M, Gardner HA, Colgan T, Rose TH, Evans WK. Therapy-related leukemia and myelodysplasia in small-cell lung cancer. Report of a case and results of morphologic, cytogenetic, and bone marrow culture studies in long-term survivors. Arch Intern Med 1986; 146: 1689-1694.
3) Collins LG, Haines C, Perkel R, Enck RE. Lung cancer: diagnosis and management. Am Fam Phys 2007; 75: 56-63.
4) Ferlay J, Shin HR, Bray F, Forman D, Mathers C, Parkin DM. Estimates of worldwide burden of cancer in 2008: GLOBOCAN 2008. Int J Cancer 2010; 127: 2893-2917.
5) Urgard E, Vooder T, Võsa U, Välk K, Liu M, Luo C, Hoti F, Roosipuu R, Annilo T, Laine J, Frenz CM, Zhang L, Metspalu A. Metagenes associated with survival in non-small cell lung cancer. Cancer Inform 2011; 10: 175-183.
6) Oh JJ, Taschereau EO, Koegel AK, Ginther CL, Rotow JK, Isfahani KZ, Slamon DJ. RBM5/H37 tumor suppressor, located at the lung cancer hot spot 3p21.3, alters expression of genes involved in metastasis. Lung Cancer 2010; 70: 253-262.
7) Kalari S, Jung M, Kernstine KH, Takahashi T, Pfeifer GP. The DNA methylation landscape of small cell lung cancer suggests a differentiation defect of neuroendocrine cells. Oncogene 2013; 32: 3559-3568.
8) Helman E, Naxerova K, Kohane IS. DNA hypermethylation in lung cancer is targeted at differentiation-associated genes. Oncogene 2012; 31: 1181-1188.
9) Wei S, Wang LE, McHugh MK, Han Y, Xiong M, Amos CI, Spitz MR, Wei QW. Genome-wide gene-environment interaction analysis for asbestos exposure in lung cancer susceptibility. Carcinogenesis 2012; 33: 1531-1537.
10) Ocak S, Yamashita H, Udyavar AR, Miller AN, Gonzalez AL, Zou Y, Jiang A, Yi Y, Shyr Y, Estrada L, Quaranta V, Massion PP. DNA copy number aberrations in small-cell lung cancer reveal activation of the focal adhesion pathway. Oncogene 2010; 29: 6331-6342.
11) Donnem T, Fenton CG, Lonvik K, Berg T, Eklo K, Andersen S, Stenvold H, Al-Shibli K, Al-Saad S, Bremnes RM, Busund LT. MicroRNA signatures in tumor tissue related to angiogenesis in non-small cell lung cancer. PLoS One 2012; 7: e29671.
12) Gao W, Xu J, Liu L, Shen H, Zeng H, Shu Y. A systematic-analysis of predicted miR-21 targets identifies a signature for lung cancer. Biomed Pharmacother 2012; 66: 21-28.
13) Lamb J, Crawford ED, Peck D, Modell JW, Blat IC, Wrobel MJ, Lerner J, Brunet JP, Subramanian A, Ross KN, Reich M, Hieronymus H, Wei G, Armstrong SA, Haggarty SJ, Clemons PA, Wei R, Carr SA, Lander ES, Golub TR. The Connectivity Map: using gene-expression signatures to connect small molecules, genes, and disease. Science 2006; 313: 1929-1935.
14) Pacini C, Iorio F, Gonçalves E, Iskar M, Klabunde T, Bork P, Saez-Rodriguez J. DvD: An R/Cytoscape pipeline for drug repurposing using public repositories of gene expression data. Bioinformatics 2013; 29: 132-134.
15) Yeh CT, Wu AT, Chang PM, Chen KY, Yang CN, Yang SC, Ho CC, Chen CC, Kuo YL, Lee PY, Liu YW, Yen CC, Hsiao M, Lu PJ, Lai JM, Wang LS, Wu CH, Chiou JF, Yang PC, Huang CY. Trifluoperazine, an antipsychotic agent, inhibits cancer stem cell growth and overcomes drug resistance of lung cancer. Am J Respir Crit Care Med 2012; 186: 1180-1188.
16) Chen MH, Lin KJ, Yang WL, Kao YW, Chen TW, Chao SC, Chang PM, Liu CY, Tzeng CH, Chao Y, Chen MH, Yeh CN, Huang CY. Gene expression-based chemical genomics identifies heat-shock protein 90 inhibitors as potential therapeutic drugs in cholangiocarcinoma. Cancer 2013; 119: 293-303.
17) Williams G. A searchable cross-platform gene expression database reveals connections between drug treatments and disease. BMC Genomics 2012; 13; 12 doi: 10.1186/1471-2167-13-12.
18) Claerhout S, Lim JY, Choi W, Park YY, Kim K, Kim SB, Lee JS, Mills GB, Cho JY. Gene expression signature analysis identifies vorinostat as a candidate therapy for gastric cancer. PLoS One 2011; 6(9): e24662.
19) Laurila K, Lahdesmaki H. Systematic analysis of disease-related regulatory mutation classes reveals distinct effects on transcription factor binding. In Silico Biol 2009; 9: 209-224.
20) Wittwer J, Marti-Jaun J, Hersberger M. Functional polymorphism in ALOX15 results in increased allele-specific transcription in macrophages through binding of the transcription factor SPI1. Hum Mutat 2006; 27: 78-87.

21) Matsuda M, Sakamoto N, Fukumaki Y. Delta-thalassemia caused by disruption of the site for an erythroid-specific transcription factor, GATA-1, in the delta-globin gene promoter. Blood 1992; 80: 1347-1351.

22) Chowdhary R, Tan SL, Pavesi G, Jin J, Dong D, Mathur SK, Burkart A, Narang V, Glurich I, Raby BA, Weiss ST, Wong L, Liu JS, Bajic VB. A database of annotated promoters of genes associated with common respiratory and related diseases. Am J Respir Cell Mol Biol; 47: 112-119.

23) Wang G, Mao W, Zheng S. MicroRNA-183 regulates Ezrin expression in lung cancer cells. FEBS Lett 2008; 582: 3663-3668.

24) Du L, Schageman JJ, Subauste MC, Saber B, Hammond SM, Prudkin L, Wistuba II, Ji L, Roth JA, Minna JD, Pertsemlidis A. miR-93, miR-98, and miR-197 regulate expression of tumor suppressor gene FUS1. Mol Cancer Res 2009; 7: 1234-1243.

25) Seike M, Goto A, Okano T, Bowman ED, Schetter AJ, Horikawa I, Mathe EA, Jen J, Yang P, Sugimura H, Gemma A, Kudoh S, Croce CM, Harris CC. MiR-21 is an EGFR-regulated anti-apoptotic factor in lung cancer in never-smokers. Proc Natl Acad Sci U S A 2009; 106: 12085-12090.

26) Hernandez MR, Agapova OA, Yang P, Salvador-Silva M, Ricard CS, Aoi S. Differential gene expression in astrocytes from human normal and glaucomatous optic nerve head analyzed by cDNA microarray. Glia 2002; 38: 45-64.

27) Davis S, Meltzer PS. GEOquery: a bridge between the Gene Expression Omnibus (GEO) and BioConductor. Bioinformatics 2007; 23: 1846-1847.

28) Smyth GK. Linear models and empirical bayes methods for assessing differential expression in microarray experiments. Stat Appl Genet Mol Biol 2004; 3: N. 1, Article 3.

29) Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT, Harris MA, Hill DP, Issel-Tarver L, Kasarskis A, Lewis S, Matese JC, Richardson JE, Ringwald M, Rubin GM, Sherlock G. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. Nat Genet 2000; 25: 25-29.

30) Zheng Q, Wang XJ. GOEAST: a web-based software toolkit for Gene Ontology enrichment analysis. Nucleic Acids Res 2008; 36(Web Server issue): W358-363.

31) Zhang B, Kirov S, Snoddy J. WebGestalt: an integrated system for exploring gene sets in various biological contexts. Nucleic Acids Res 2005; 33(Web Server issue): W741-748.

32) Dexter Duncan NPaBZ. WebGestalt2: an updated and expanded version of the Web-based Gene Set Analysis Toolkit. BMC Bioinformatics 2010; Suppl 4: 10.

33) Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. Journal of the Royal Statistical Society Series B 1995; 57: 289-300.

34) Smythe WR, Williams JP, Wheelock MJ, Johnson KR, Kaiser LR, Albelda SM. Cadherin and catenin expression in normal human bronchial epithelium and non-small cell lung cancer. Lung Cancer 1999; 24: 157-168.

35) Kort EJ, Paneth N, Vande Woude GF. The decline in U.S. cancer mortality in people born since 1925. Cancer Res 2009; 69: 6500-6505.

36) Takahashi Y, Forrest AR, Maeno E, Hashimoto T, Daub CO, Yasuda J. MiR-107 and MiR-185 can induce cell cycle arrest in human non small cell lung cancer cell lines. PLoS One 2009; 4: e6677.

37) Zhang F, Chen JY. Discovery of pathway biomarkers from coupled proteomics and systems biology methods. BMC Genomics 2010; 11 Suppl 2: S12.

38) Leysen JE, W Gommeren, L Heylen, WH Luyten, I Van de Weyer, P Vanhoenacker, G Haegeman, A Schotte, P Van Gompel, R Wouters, AS Lesage. Alniditan, a new 5-hydroxytryptamine1D agonist and migraine-abortive agent: ligand-binding properties of human 5-hydroxytryptamine1D alpha, human 5-hydroxytryptamine1D beta, and calf 5-hydroxytryptamine1D receptors investigated with [3H]5-hydroxytryptamine and [3H]alniditan. Mol Pharmacol 1996; 50: 1567-1580.

39) Chao SK, J Lin, J Brouwer-Visser, AB Smith, 3rd, SB Horwitz, HM McDaid. Resistance to discodermolide, a microtubule-stabilizing agent and senescence inducer, is 4E-BP1-dependent. Proc Natl Acad Sci U S A 2011; 108: 391-396.

40) Liao WT, X Wang, LH Xu, QL Kong, CP Yu, MZ Li, L Shi, MS Zeng, LB Song. Centromere protein H is a novel prognostic marker for human nonsmall cell lung cancer progression and overall patient survival. Cancer 2009; 115: 1507-1517.

41) Hayama S, Y Daigo, T Kato, N Ishikawa, T Yamabuki, M Miyamoto, T Ito, E Tsuchiya, S Kondo, Y Nakamura. Activation of CDCA1-KNTC2, members of centromere protein complex, involved in pulmonary carcinogenesis. Cancer Res 2006; 66: 10339-10348.

42) Hansen LT, T Thykjaer, TF Orntoft, LJ Rasmussen, P Keller, M Spang-Thomsen, TB Edmonston, C Schmutte, R Fishel, LN Petersen. The role of mismatch repair in small-cell lung cancer cells. Eur J Cancer 2003; 39: 1456-1467.

43) Dogra N, T Mukhopadhyay. Impairment of the ubiquitin-proteasome pathway by methyl N-(6-phenylsulfanyl-1H-benzimidazol-2-yl)carbamate leads to a potent cytotoxic effect in tumor cells: a novel antiproliferative agent with a potential therapeutic implication. J Biol Chem 2012; 287: 30625-30640.

44) Xu L, L Su, X Liu. PKCdelta regulates death receptor 5 expression induced by PS-341 through ATF4-ATF3/CHOP axis in human lung cancer cells. Mol Cancer Ther 2012; 11: 2174-2182.

45) Meiners S, O Eickelberg. What shall we do with the damaged proteins in lung disease? Ask the proteasome! Eur Respir J 2012; 40: 1260-1268.

46) Li Y, Z Sun, JM Cunningham, MC Aubry, JA Wampfler, GA Croghan, C Johnson, D Wu, JA Aakre, J Molina, L Wang, VS Pankratz, P Yang. Genetic variations in multiple drug action pathways and survival in advanced stage non-small cell lung cancer treated with chemotherapy. Clin Cancer Res 2011; 17: 3830-3840.

47) Kastner S, T Voss, S Keuerleber, C Glockel, M Freissmuth, W Sommergruber. Expression of G protein-coupled receptor 19 in human lung cancer cells is triggered by entry into S-phase and supports G(2)-m cell-cycle progression. Mol Cancer Res 2012; 10: 1343-1358.

48) Meng X, P Lu, H Bai, P Xiao, Q Fan. Transcriptional regulatory networks in human lung adenocarcinoma. Mol Med Report 2012; 6: 961-966.

49) Zhan X, W Wu, B Han, G Gao, R Qiao, J Lv, S Zhang, W Zhang, W Fan, H Chen, T Zhang, A Gu, J Shen, Q Wu, D Lu. Hsa-miR-196a2 Functional SNP is associated with severe toxicity after platinum-based chemotherapy of advanced nonsmall cell lung cancer patients in a chinese population. J Clin Lab Anal 2012; 26: 441-446.

50) Hurteau GJ, JA Carlson, SD Spivack, GJ Brock. Overexpression of the microRNA hsa-miR-200c leads to reduced expression of transcription factor 8 and increased expression of E-cadherin. Cancer Res 2007; 67: 7972-7976.

51) Dragnev K, M You, Y Wang, R Lubet. Lung cancer chemoprevention: difficulties, promise and potential agents? Expert Opin Investig Drugs 2013; 22: 35-47.

52) Khuri FR. The dawn of a revolution in personalized lung cancer prevention. Cancer Prev Res (Phila) 2011; 4: 949-953.

53) Szabo E. Altered histology provides a positive clinical signal in the bronchial epithelium. Cancer Prev Res (Phila) 2011; 4: 775-778.

54) Pehlivan Y, IH Turkbeyler, O Balakan, A Sevinc, M Yilmaz, K Bakir, AM Onat. Possible anti-metastatic effect of iloprost in a patient with systemic sclerosis with lung cancer: a case study. Rheumatol Int 2012; 32: 1437-1441.

55) Tennis MA, M Van Scoyk, LE Heasley, K Vandervest, M Weiser-Evans, S Freeman, RL Keith, P Simpson, RA Nemenoff, RA Winn. Prostacyclin inhibits non-small cell lung cancer growth by a frizzled 9-dependent pathway that is blocked by secreted frizzled-related protein 1. Neoplasia 2010; 12: 244-253.

56) Nagathihalli NS, PP Massion, AL Gonzalez, P Lu, PK Datta. Smoking induces epithelial-to-Mesenchymal Transition in Non-Small Cell Lung Cancer through HDAC-Mediated Downregulation of E-Cadherin. Mol Cancer Ther 2012; 11: 2362-2372.

57) Suda K, K Tomizawa, M Fujii, H Murakami, H Osada, Y Maehara, Y Yatabe, Y Sekido, T Mitsudomi. Epithelial to mesenchymal transition in an epidermal growth factor receptor-mutant lung cancer cell line with acquired resistance to erlotinib. J Thorac Oncol 2011; 6: 1152-1161.

58) Kakihana M, T Ohira, D Chan, RB Webster, H Kato, HA Drabkin, RM Gemmill. Induction of E-cadherin in lung cancer and interaction with growth suppression by histone deacetylase inhibition. J Thorac Oncol 2009; 4: 1455-1465.

59) Niesen MI, G Blanck. Rescue of major histocompatibility-DR surface expression in retinoblastoma-defective, non-small cell lung carcinoma cells by the MS-275 histone deacetylase inhibitor. Biol Pharm Bull 2009; 32: 480-482.

60) Hurtubise A, ML Bernstein, RL Momparler. Preclinical evaluation of the antineoplastic action of 5-aza-2'-deoxycytidine and different histone deacetylase inhibitors on human Ewing's sarcoma cells. Cancer Cell Int 2008; 8: 16.