

Combined oligonucleotide microarray-bioinformatics and constructed membrane arrays to analyze the biological pathways in the carcinogenesis of human lung adenocarcinoma

MEI-YIN CHANG¹, YA-PING YU⁶, JONG-RUNG TSAI²,
CHAU-CHYUN SHEU², INN-WEN CHONG², SHIU-RU LIN³⁻⁵

¹Graduate Institute of Medicine, College of Medicine; ²Division of Pulmonary and Critical Care Medicine, Kaohsiung Medical University Hospital; ³Graduate Institute of Medical Genetics, College of Medicine; ⁴BioMedi Innovation Incubation Center; ⁵Department of Laboratory Medicine, Kaohsiung Medical University Hospital, Kaohsiung Medical University; ⁶Gene Target Technology Co., Ltd., No. 332, Mingcheng 2nd Road, Zuoying District, Kaohsiung 807, Taiwan, ROC

Received January 24, 2007; Accepted March 22, 2007

Abstract. The present study systematically explores the biological pathways and altered expression of genes speculatively participating in lung carcinogenesis by using oligonucleotide microarray-bioinformatic analysis methods. The results revealed that 1,396 genes were up-regulated and 1,965 were down-regulated in lung adenocarcinoma carcinogenesis. Gene ontology and relevant bioinformatics tools indicated that the functional category to which the most frequently differentially expressed genes were classified, was to the cytokine-cytokine receptor interaction pathway, focal adhesion pathway and the mitogen-activated protein kinase signaling pathway. Furthermore, we constructed a membrane array, consisting of 51 up-regulated genes in lung adenocarcinoma, in order to verify the biological pathways involved in the carcinogenesis of lung cancer. The analysis of 45 lung adenocarcinoma tissue specimens demonstrated that the genes involved in these three biological pathways had high

rates of overexpression. Out of the 51 genes, 17 genes were demonstrated to be overexpressed in all 45 lung adenocarcinoma tissues compared to the paired normal lung tissues. These findings could have implications in understanding the process of lung adenocarcinoma carcinogenesis. Moreover, our developed membrane arrays could be a potentially feasible and promising tool in clinical practice for analyzing the molecular mechanisms of lung adenocarcinoma carcinogenesis.

Introduction

Lung cancer is the major cause of cancer related mortality worldwide, and is also the most frequent cause of cancer death in Taiwan, with over 7,000 new cases per year (<http://www.doh.gov.tw/statistic/index.htm>; accessed in April 2006). In Taiwan, the disease results in an annual mortality rate of 31.58 per 100,000 inhabitants. Even with the innovations in diagnostic testing, surgical techniques, and the development of new chemotherapeutic agents, the outcome of patients with lung cancer has improved only modestly over the past few decades, and the overall 5-year survival rate has remained at ~10-15% (1,2). Non-small cell lung cancer (NSCLC) is the predominant form of lung cancer and consists of two major histological subtypes: Squamous cell carcinoma and adenocarcinoma. Throughout the past 3 decades the incidence of squamous cell carcinoma has decreased while the incidence of adenocarcinoma has increased (3), and in recent years, adenocarcinoma has become the most common type.

Lung adenocarcinoma is the most common histological type and accounts for approximately 40-45% of all lung cancer cases. Although the clinical courses of these tumors are similar, adenocarcinomas are characterized by peripheral location in the lung and often have activating mutations in the *K-ras* oncogene (4,5). In contrast, squamous cell carcinomas are usually centrally located and more frequently

Correspondence to: Dr Shiu-Ru Lin, MedicoGenomic Research Center, Kaohsiung Medical University, 100 Shin-Chuan 1st Road, Kaohsiung 807, Taiwan, ROC
E-mail: srlin@ms2.hinet.net

Abbreviations: DAVID, Database for Annotation, Visualization, and Integrated Discovery; GO, gene ontology; KEGG, Kyoto Encyclopedia of Genes and Genomes; MAPK, mitogen-activated protein kinase

Key words: lung adenocarcinoma; cytokine-cytokine receptor interaction pathway, focal adhesion pathway, mitogen-activated protein kinase signaling pathway, bioinformatics, gene ontology

carry *p53* gene mutations (6). Although many molecular abnormalities are common to both NSCLC and small cell lung cancer (SCLC), there are differences between these histological types, even within the NSCLC histological subtypes (7). The large-cell carcinoma gene set includes a preponderance of genes associated with cellular proliferation and cell-cycle regulation, whereas adenocarcinoma genes are associated with differentiation and signal transduction (1). Moreover, adenocarcinoma of the lung has been reported to be more strongly associated with tobacco smoke exposure than squamous cell and small cell carcinoma (8,9). Understanding the genes specifically expressed in lung adenocarcinoma can provide novel markers for disease detection, risk assessment as well as potential markers for the rational therapy of this disease. There are several molecular mechanisms involved in the carcinogenesis of lung cancer through which respiratory epithelial cells become preneoplastic and then invasive cancer. Though many molecular changes associated with lung cancer have been investigated (10,11), the molecular characteristics of lung cancer, with particular reference to lung adenocarcinoma, analyzed by combined functional genomics-bioinformatics have been rarely explored.

The objective of this study was to analyze the differentially expressed genes between lung adenocarcinoma and normal pulmonary tissues, initially screened by microarray assays, as well as to determine the molecular pathways involved in lung adenocarcinoma carcinogenesis with relevant bioinformatics tools. The subsequent validation of these molecular pathway-related genes was confirmed in additional lung adenocarcinoma tissue samples using the similar methodology of constructed membrane arrays published recently (12). Consequently, we could elucidate the role and significance of each identified molecular pathway underlying the development of lung cancer, and all these approaches can hopefully identify potential diagnostic/therapeutic markers, translating into clinical benefit for lung cancer patients.

Materials and methods

Tumor samples collection. Lung adenocarcinoma and corresponding normal tissues were collected from non-selected 45 patients who underwent surgical resection at the Division of Chest Surgery, Department of Surgery, Kaohsiung Medical University Hospital. Of all these patients, there were eight males and seven females (mean age was 57.1 years, ranging from 38 to 80 years). Half of the tissue samples were immediately placed in liquid nitrogen for shipment to the laboratory, and then stored in -80°C freezers until RNA isolation. The other half of the tissue samples were sent unfrozen for routine histopathological diagnosis. Written informed consent was obtained from all of the subjects and/or guardians for the use of their tissue samples. The institutional review board of the Kaohsiung Medical University Hospital also approved the sample acquisition and its subsequent use. Complete staging procedures including chest radiography, bronchoscopy, brain and thoracic computed tomography, sonography, and bone scintigraphy were carried out to precisely determine the characteristics of the primary tumor (T), nodal involvement (N), and metastasis (M) in the

NSCLC patients according to the TNM International Staging System for Lung cancer (41). Of the 45 lung adenocarcinoma patients, 18 were subsequently staged as stage I; 9 as stage II; 6 as stage III; and 12 as stage IV.

mRNA isolation and cDNA synthesis. Total RNA was extracted from the tumor and paired normal tissues of lung adenocarcinoma patients with IsogenTM (Nippon Gene Co., Ltd., Toyama, Japan) by following the modified acid-guanidine thiocyanate and phenol/chloroform extraction method (42). All the RNA samples were verified to be in good quality for three clear bands corresponding to the 28S, 18S and 5S ribosomes clearly present in an ethidium bromide-stained gel. Poly(A)⁺-enriched RNA was then purified from total RNA by using Dynabeads[®] mRNA DirectTM kit (DynaL A.S., Oslo, Norway).

Oligonucleotide microarray. The oligonucleotide array contains 22,500 elements designed for expression profiling (Human 1A V2, Agilent Technologies, Palo Alto, CA, USA), for which over 18,000 well-characterized, full-length human genes were defined. First-strand cDNA targets for hybridization were made by reverse transcription of the mRNA isolated from both the tumor and paired normal tissues from the same case by using SuperScript II RT (Gibco-BRL, Gaithersburg, MD, USA) in the presence of either Cy3- or Cy5-labeled dUTP (Amersham Pharmacia Biotech, Piscataway, NJ, USA). The targets were dried to 18 μl by a SpeedVacTM concentrator (Thermo Electron Corp., Waltham, MA, USA), and 3.6 μl 20X SSC, 1.8 μl 10 mg/ml poly-A and 0.54 μl 10% SDS were added. Then, the mixture was heated to 100°C for 2 min proceeding to hybridization reaction on Human 1A Oligo Microarray V2 array slides (Agilent Technologies) in an incubator at 63°C for 12-16 h. After being sequentially washed with 1X SSC, 0.2X SSC and 0.05X SSC, hybridized microarray slides were scanned and fluorescence signals were detected by using an Axon GenePix 4000A dual-color confocal laser scanner (Axon Instruments, Union, CA, USA). Subsequent quantification analysis was performed using the commercial software GenePix Pro 3.0TM (Axon Instruments). The acceptance criterion for a gene's signal was a signal-to-noise ratio of ≥ 2 . If either the Cy3 or Cy5 signal of a specific spot passed the criterion, the flag of its ratio was counted to be 'TRUE'. The elements with the 'TRUE' flag were analyzed with GeneSpring GX7 (Silicon Genetics, Redwood, CA, USA). The differentially expressed elements were figured out by the two-sided statistical tolerance interval (95%).

Functional genomics analysis. The functional genomics analysis was performed according to our recent study (13). Briefly, a program package, MatchMiner (<http://discover.nci.nih.gov/matchminer/index.jsp>), was used to scan human genome organization (HUGO) names for significantly expressed genes (≥ 2 -fold increase in intensity) initially screened by cDNA microarray analysis. Then, the HUGO names were input into GoMiner (<http://discover.nci.nih.gov/gominer>), which leveraged the Gene Ontology (GO) in order to identify the biological processes, functions and components represented in the gene profile. We also used the

Table I. The distribution of differentially expressed genes after mapping the pathway diagrams by linking to the DAVID and KEGG databases.

Pathway Term		Differential expression genes / category genes (%)
Cytokine-Cytokine Receptor Interaction (Homo sapiens)	██████████	34
Focal adhesion (Homo sapiens)	██████████	28
MAPK Signaling Pathway (Homo sapiens)	██████	15
Regulation of actin cytoskeleton (Homo sapiens)	██████	12
Hematopoietic cell lineage (Homo sapiens)	█	3
ECM-receptor interaction (Homo sapiens)	█	2
Complement and coagulation cascades (Homo sapiens)	█	2

Database for Annotation, Visualization, and Integrated Discovery (DAVID) (<http://apps1.niaid.nih.gov/david/>) and the Kyoto Encyclopedia of Genes and Genomes (KEGG) database (<http://www.genome.ad.jp/kegg>). For this analysis, microarray expression data was loaded by MatchMiner and analyses for each GO biological process, cellular component, and molecular function ontology were carried out. In addition, DAVID tools were used to analyze the roles of these genes in metabolic pathways and to show the biological relationships between genes or gene-products.

Design of oligonucleotide probes and preparation of oligonucleotide membrane arrays. The procedure for the design and preparation of the membrane array was carried out according to our previously described method (12). Visual OMP 3 (Oligonucleotide Modeling Platform, DNA Software, Ann Arbor, MN, USA) (14,15) was used to design probes for each of the gene targets. The probe selection criteria included strong mismatch discrimination, minimal or no secondary structure, the signal strength at the assay temperature, and the lack of cross-hybridization. Oligonucleotide probes are synthesized according to well designed sequences, purified and controlled before being grafted onto the substrates. The newly synthesized oligonucleotide fragments were dissolved in di-water to a concentration of 20 mM and then applied to a AD3050™ Dispensing System (BioDot Inc., Irvine, CA, USA), which sequentially blotted the 51 target DNAs (Table II), 1 housekeeping gene (β -actin), and 1 nonmammalian plant gene (50 nl per spot and 1.5 mm between spots) on a Nytran® SuperCharge nylon membrane (Fig. 3A) (Schleicher and Schuell, Dassel, Germany) and then cross-linked to the membrane using a UV Stratallinker 1800 (Stratagene, La Jolla, CA, USA) in triplicate. Each spot consists of 20 ng PCR-amplified DNA derived from sequence-verified cDNA clones. DMSO was also dispensed onto the membrane as a blank control (Table II).

Preparation of digoxigenin (DIG)-labeled cDNA targets and hybridization. First-strand cDNA targets for hybridization were made by reverse transcription of the mRNA from the tumor and corresponding normal tissues of lung adenocarcinoma patients in the presence of DIG-labeled UTP (Roche Diagnostics GmbH, Penzberg, Germany) using

SuperScript II reverse transcriptase (Gibco-BRL). The membrane array needs prehybridization at 42°C for 2 h in a hybridization oven (Autoblot Bellco, Vineland, NJ, USA), and subsequent blocking before hybridization. The lifts were covered with the ExpressHyb Hybridization Solution (BD Biosciences, Palo Alto, CA, USA) containing DIG-11-UTP-labeled cDNA probes, and then incubated with an anti-DIG alkaline phosphatase conjugated antibody (Roche Diagnostics) at a dilution of 1:2500 (75 mU/ml) in 1 x blocking buffer (100 mM Maleic acid, 150 mM NaCl, pH 7.5). We incubated the arrays for hybridization at 42°C overnight in a humid shaking chamber. After washing each one three times for 10 min in 20 ml washing buffer, the arrays were then exposed to light that excites the light. For signal detection, the membranes we incubated for 15 min without shaking in a chromogen solution containing nitroblue-tetrazolium and 5-bromo-4-chloro-3-indolyl-phosphate (NBT/BCIP).

Analysis of spot intensity ratio for membrane array. The hybridized membrane-arrays were then scanned using an Epson Perfection 1670 flat bed scanner (Seiko Epson Corp., Nagano-ken, Japan). Subsequent quantification analysis of each spot's intensity was carried out using commercial AlphaEase®FC software (Alpha Innotech Corp., San Leandro, CA, USA). A deformable template extracted the gene spots and quantified their expression levels by the integrated intensity of the spot after background subtraction. The fold ratio for each gene was calculated as follows: Spot intensity ratio = (mean intensity of target gene)/(mean intensity of corresponding β -actin). Spots consistently differing by a factor of two or more were regarded as differentially expressed. The membrane arrays could be stored in tap water for several months without a loss of color.

Results

Gene expression analyses. In order to search for new target genes that were differentially overexpressed in lung adenocarcinoma and thought to be involved in pulmonary carcinogenesis, cDNA microarray analysis was used. Three pairs of early stage lung adenocarcinoma and normal lung tissues were used for microarray analysis. Raw images of the scanned slides for adenocarcinoma versus the normal tissues

Table II. Oligonucleotide sequence for membrane array.

Pathway	Gene name	Sequence of oligonucleotide probe	Length
Cytokine-cytokine receptor interaction	Growth hormone 2/GH2	TCTTGAGCAGTGCCTCATCGTTGTGCGATTTTGTGTCAAACCTGCTGTAG	50
	Epidermal growth factor (beta-urogastrone)/EGF	ATTCATTAACATCTTCACAGTACTTCCGGTCTCGACTTAGGGCGTATCCC	50
	Fms-related tyrosine kinase 1 (vascular endothelial growth factor/vascular permeability factor receptor)/FLT1	GGGCTCCTTGTAGAAACCGTCAGAATCCTCCTCTTCTCAACATCACTCA	50
	Erythropoietin receptor/EPOR	CGCCATCCCTGTTCCATAAGTCTTGAGTCTGCACTGGTTCTCTGAGTCAT	50
	Interferon epsilon 1/IFNT1	ACACAAAGAACAGACATCGGCTGATTTCTACTTGGACAATGGCCCAGGCA	50
	Interferon, alpha 16/IFNA16	CCCAGAGAACAGATGGATTTGTAGCTGAGCACCAGCAGGCCATCAGTAA	50
	Interleukin 11/IL11	TGGTGAACACAGCCATAATGAACCATCTTGACCTTGCAGCTTTGTCCCAT	50
	Interleukin 20 receptor, alpha/IL20RA	CCAGGTCACTAGGCAGTAAAACCTCCTCTTGTAATGAGCGAGGCGAGT	50
	Interleukin 22 receptor, alpha 1/IL20RA1	TTGCCACCCAGTCCCTCTCTCCGTACGTCTTATACTCGATGCTGTAGACC	50
	Interleukin 28 receptor, alpha (interferon, lambda receptor)/IL28RA	GAATGGGACTTCTAAGGACCACAACGGTCAACCAAGAGTCTAACGTGCCA	50
	Leukemia inhibitory factor (cholinergic differentiation factor)/LIF	CGCTTGTGTCGGGAGCCATCATCATATTTTGTGAGTAGGGACCAGGGAA	50
	Transcriptional intermediary factor 1/TIF1	AGTAGAGGATGTGCTGTTGGTAGTTGATGGGGTCCCTGTCCACTGCTGA	50
	Thyroid peroxidase/TPO	CATTCTCCACGCTCTCTGGGAAGCCACACTTGTGCTCTTGGAGAAAGGTT	50
	Bone morphogenetic protein 7 (osteogenic protein 1)/BMP7	TAGGCGCTCATAATTACCTCTGGAACGAGTCCGTGCATGGCTGAGACTT	50
	Chemokine (C-C motif) ligand 16/CCL16	CGAGGCAGTACAGCTTCAGGGAGAGCCGAATGAAGATGTTGTCTGTTGCT	50
	Chemokine (C-C motif) ligand 20/CCL20	TCCATGATGTGCAAGTGAAACCTCCAACCCAGCAAGGTTCTTTCTGTTC	50
	Chemokine (C-C motif) ligand 26/CCL26	GCTTGTGGCTGTATTGGAAGCAGCAGGTCTTGATATGCACTCCACCGT	50
	Chemokine (C-X-C motif) ligand 1 (melanoma growth stimulating activity, alpha)/CXCL1	CAGCATCTTTTCGATGATTTTCTTAACCTATGGGGGATGCAGGATTGAGGC	50
	Chemokine (C-X-C motif) ligand 1 (melanoma growth stimulating activity, alpha)/CXCL1	CAGCATCTTTTCGATGATTTTCTTAACCTATGGGGGATGCAGGATTGAGGC	50
	Chemokine (C-X-C motif) ligand 5/CXCL5	ACCTTGAGCACTGTGGCCTATGGCGAACACTTGCAGATTACTGATCAT	50
	Colony stimulating factor 1 (macrophage)/CSF1	TGGAGAGGTGTCTCATAGAAAGTTCCGACGCAGGCCTTGTGCTGCTCTTC	50
	Ectodysplasin A/EDA	CAAACATGAAGGAGTGGTTCTTGTGTTTCTGCTACAACCTGCGGACTGT	50
	Interleukin 17 receptor B/IL17RB	GTGCTGTGTTGGATAAGAGCCATGTATCTGTTTCCAGGGGAGTGGTTGT	50
Inhibin, beta E/INHBE	TTGACCACATTGCCATTATGATCCAGGTAGAGGAGAGAGAGGGGCCTTCG	50	
Tumor necrosis factor receptor superfamily, member 21/SF21	AGGTGGGCAAGTGCATTCTCGGTCAGTCAAGGCAGCACAAGGTAATTCT	50	
MAPK signaling pathway	Dual specificity phosphatase 16/DUSP16	AAGCTGTCATTACAGGCACACGCAGGAAATGAGACTCGGGGATAAAGTC	50
	Dual specificity phosphatase 6/DUSP6	TAAAGTGAGCCCATGATTTGGTGTCTTTCCAGATTATCCCCTTTGCGAC	50
	Fibroblast growth factor receptor 2/FGFR2	CGGCGATATTTTGTCTGATGTAGGTATGAGGCTGGATCTTTTGGTGAGGT	50
	Fibroblast growth factor receptor 3 (achondroplasia, thanatophoric dwarfism)/FGFR3	TGATGTTTTTGTGTTTCCCGATCATCTTCATCATCTCCATCTCAGACACC	50
	Mitogen-activated protein kinase 13/MAPK13	CTCCTTGATAGATGTGCTGCTTCCATTCATCCACTGTGAGTTTCTCGTGT	50
	Mitogen-activated protein kinase 4/MAPK4	CACCAAACCTCACCTTAGCCTACACTATGAGGGGCATTGCTGCTTTGAGTC	50
Mitogen-activated protein kinase 8 interacting protein 2/MAPK8IP2	GAGAGAGAAAAGGGAGACTGGGAAGGGAAGGAGGAGGACAACACAGAGAA	50	

Table II. Continued.

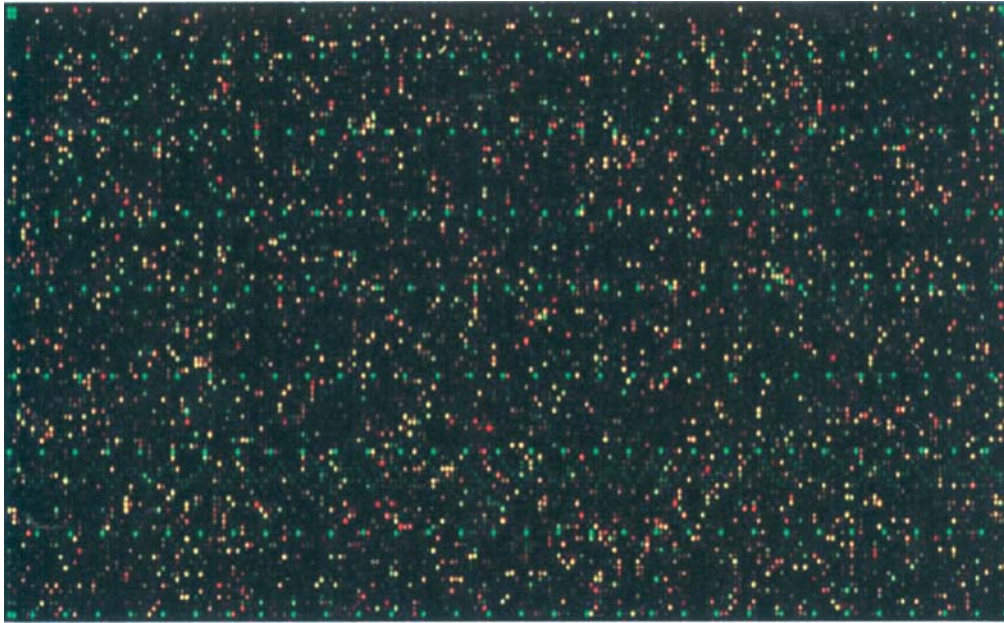
Pathway	Gene name	Sequence of oligonucleotide probe	Length
	Mitogen-activated protein kinase 13/ MAP3K13	ATCCATGCGACCGTGCCAGCAAATGACATCTTGGTACTTTTGTCACTGAG	50
	Mitogen-activated protein kinase 5/ MAP3K5	TGGTTGCTCAAGTCCCGACCTGCGTAGACTATCCATAAGTGCCTTTTCC	50
	Myelodysplasia syndrome 1/MDS1	TCTCTCAATCCACACTCGCTATCTCTCCAGCATTGTGAGTTTGGACACCT	50
	Phospholipase A2, group IB (pancreas/ PLA2G1B	GATTGCCGAGCCAGAGCAGGATGAATAGGTGTGGGTGTACGGGTTGT	50
	Phospholipase A2, group III/PLA2G3	TGTGGAATCGGTAGTTTCGGATGCCATAGTTGTACTGCAAGGGTGAGATG	50
	Phospholipase A2, group X/PLA2G10	GCGCCCAAAGGGAACAGAAGACTTGGAGGAGGGGTAGTCCGTAATTTGCT	50
	Phospholipase A2, group X1IA/ PLA2G12	ACGTGGGAAAGGCTTAGATCCGTCCTACTGCATTTATACTGGCAGAGACCGT	50
	Protein phosphatase 1B (formerly 2C), magnesium-dependent, beta isoform/ PPM1B	ATCAGGCATTCTTCCTCGCCAGACTTCTCCATAATCTCTTCAACCCGTG	50
	RAS guanyl releasing protein 1 (calcium and DAG-regulated)/RASGRP1	TTTCCAGATTTCTGTGATCCCACTAGCCCCATGTTGCTCAACACAAATAC	50
Focal adhesion	Baculoviral IAP repeat-containing 4/BIRC4	TTAGCCCGTAATCAAAGGGACAGGGGAAAAAACCAAATGTAGCAGGAAAG	50
	Calpain 9/CAPN9	CAGGTGTTCTGCTTTTGTGAGGGCACTCATAAATGGCATAGCCGATTGTCA	50
	Collagen, type XI, alpha 1/COL11A1	ACCTTTAGATCCCTTGAGACCTCTGACACCATCTGCTCCCTTTACTCCCC	50
	Integrin, alpha 11/ITGA11	TGCAGGTTTGTGACTGCGAGATATTTAGGACCGTGTGTAGGCGTTCTC	50
	Integrin, beta 4/ITGB4	GGTCAAGTTTGAAGAAGTGTGGTCCATGTGGGTGCTAAGGGTTCCGCT	50
	Laminin, beta 3/LAMB3	TCTCCACCATCTTTGTCTGTCAAGTGAAGTGTCCATTGGCTCAGGCTC	50
	Fyn-related kinase/FRK	AAACTGACCAGATCCCAATCGCTTCAGAAGCTGTATGGAGTTGCGGTCTA	50
	Protein kinase STYK1/STYK1	TATTGATGGAATTGGATTCGCCCTAAGAAATCTGTACCTCATGGAGCCC	50
	Secreted phosphoprotein 1/SPP1	ACGGCTGACTTTGGAAAGTCTGACTATCAATCACATCGGAATGCTCAT	50
	v-erb-b2 erythroblastic leukemia viral oncogene homolog 2, neuro/glioblastoma derived oncogene homolog (avian)/ERBB2	CCCCAAAAGTCATCAGCTCCACACAGTCACACCATAACTCCACACATCA	50
	v-src sarcoma (Schmidt-Ruppin A-2) viral oncogene homolog (avian)/SRC	TGTACTCCGTGACGATGTAAATGGGCTCCTCTGAAACCACAGCATAACAAC	50
Negative control	Desulfotalea psychrophila LSv54, complete genome	TCCCATGCAGTACCATCGGCGCTGAAGAGCTTAACTTCCGTGTTT	45
Positive control	Actin, beta/ACTB	AACATAATCTGAGTCATCTTCTCTCTGTTGGCCTTGGGGTTCAGGGGGGC	50

are shown in Fig. 1A. All of the cDNAs printed on the arrays have been sequence-verified. Genes exhibiting greater intensities in the adenocarcinoma tissues than in the corresponding normal tissues by at least 2-fold were considered significantly overexpressed. Scatter plots of spot intensities of mRNA hybridized to the Human 1A V2 array for adenocarcinoma versus normal tissues are shown in Fig. 1B. Part of the spot intensities of the genes were located on the X=Y line of each scatter plot, except the spots of the differentially expressed genes. This X=Y line can be the indicator for the quality control of the microarray technique. The results showed that there were 1396 up-regulated genes and 1965 down-regulated genes in lung adenocarcinoma. Subsequently, the biological process and molecular function of these identified genes in lung adenocarcinoma tumorigenesis were categorized through the cancer genome anatomy project (CGAP) GO browser. These identified genes were later found to play certain crucial roles in lung cancer carcinogenesis when the

biological functions of these genes in cancer were explored. *GoMiner displays for microarray gene-expression data of functional categories.* GoMiner classified the genes into three functional categories, i.e. biological process, cellular component and molecular function (Fig. 2). The relative enrichment, the parameter for interpretation within the parenthesis following each category, was indicated by blue numbers for total flagged genes and by red and green numbers for over- and underexpressed genes, respectively. The last blue number for each category was a two-sided P-value from Fisher's exact test. There were 1653, 1644, and 1969 genes categorized into the biological process, cellular component, and molecular function ontology, respectively.

Differentially expressed genes in biological pathways. Through the DAVID and KEGG database analyses, the differentially overexpressed genes were categorized into 7 distinct pathways. We calculated the ratio of the differentially

A



B

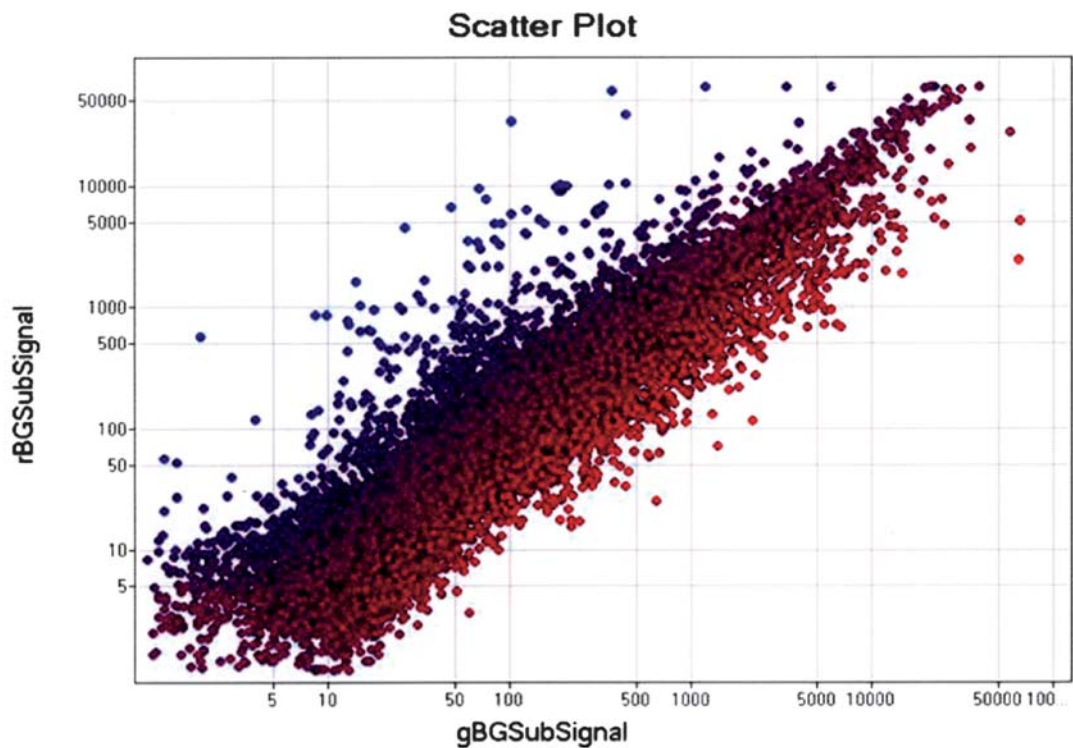


Figure 1. (A) Gene expression analysis of oligonucleotide microarray. Competitive cDNA microarray images of human lung adenocarcinoma and corresponding normal tissue. For a same-versus-same comparison, 2.5 μg Cy5-labeled mRNAs (red) from adenocarcinoma tissues and 2.5 μg Cy3-labeled mRNA (green) from normal tissues were hybridized to TGS-4K array glass slides. The hybridized slides were scanned by the GenTec UC4 microarray scanner. Raw data was collected and calibrated. (B) Scatter plot of oligonucleotide microarray. The intensities of TF(-) (labeled with Cy5-532 nm) were calibrated with the Normalization Factor (Formula 1-3), and the TF(+) intensities (labeled with Cy3-635 nm) shown are the original values. The genes with a ratio of (Cy-5: Cy-3) ≥ 2 or ≤ 0.5 were considered to be differentially repressed. (Signal: mean signal intensity of Cy3 channel - mean background light intensity of Cy3 channel).

expressed genes to the total number of genes on the microarray for each pathway, and found that the ranking of the pathways by the ratio was consistent with the output from DAVID (Table I). Of the 7 pathways, the cytokine-cytokine receptor interaction pathway, focal adhesion pathway and the

mitogen-activated protein kinase (MAPK) signaling pathway consisted of three top numbers of genes. There were 24, 13, and 18 up-regulated genes as well as 60, 62, and 48 down-regulated genes in each pathway (data not shown). Finally, fifty one up-regulated genes were identified and selected as

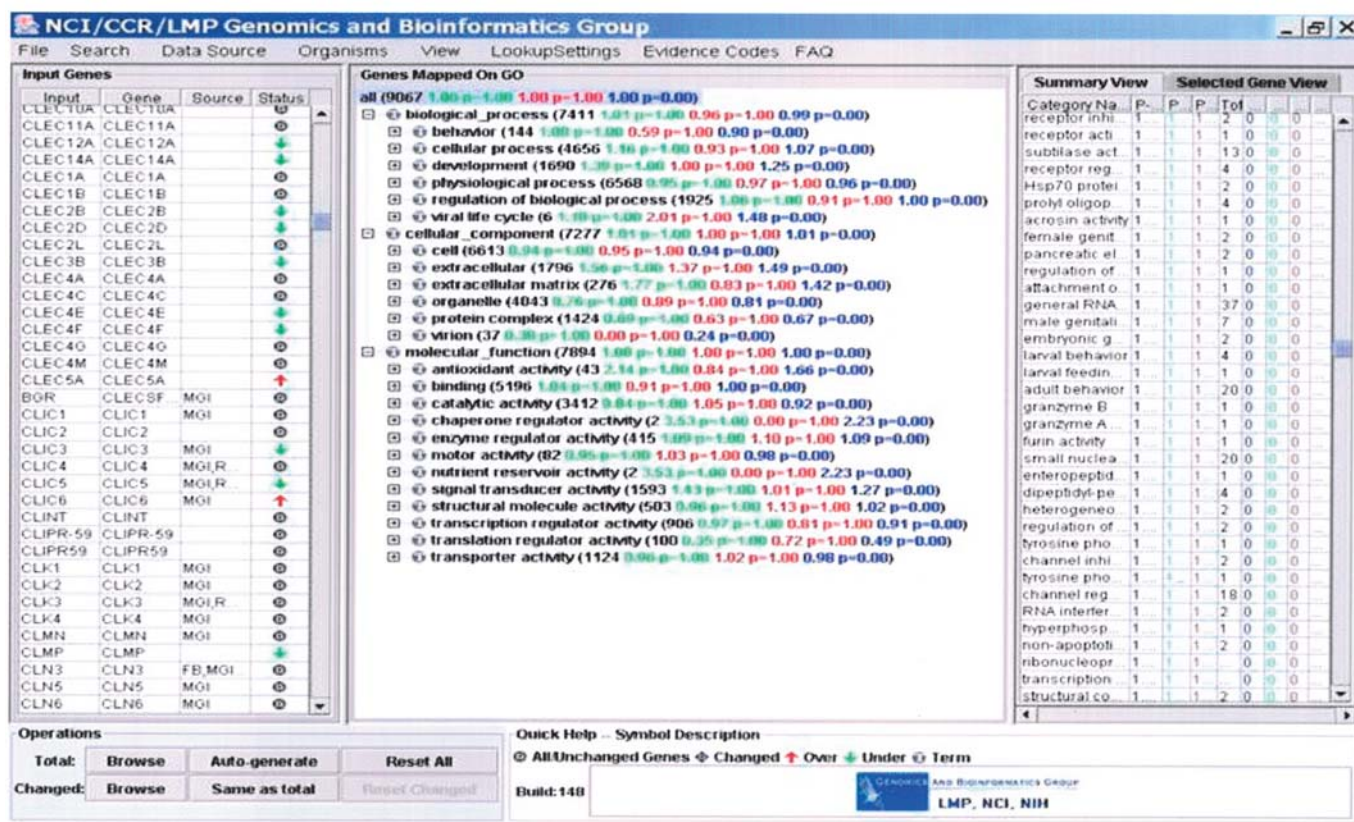


Figure 2. GoMiner displays the gene expression data from microarray analysis in the form of a tree-like structure. (Left panel) The tree-like display shows underexpressed genes (green down-arrows), overexpressed genes (red up-arrows), and constant genes (gray circles) within the framework of the GO hierarchy. (Right panel) Black number, total genes in the category; green number, relative enrichment of underexpressed genes in the category; red number, relative enrichment of overexpressed genes in the category; blue number, relative enrichment of all changed genes in the category; P-value, Fisher's exact P-value for all changed genes in the category.

candidate genes for the membrane arrays construction after the deletion of repetition. The gene expressions involved in these three pathways could play some significant roles in lung adenocarcinoma tumorigenesis.

Gene validation by membrane array analysis. Fig. 3B and C shows the representative results of membrane array hybridization for a normal lung tissue and lung adenocarcinoma tissue, respectively. A significant overexpression of the candidate genes was observed in the lung adenocarcinoma tissue as opposed to the corresponding normal lung tissue. Of the 51 candidate genes, only 3 genes, including the growth factor 2 (GH2), epidermal growth factor (EGF), and the bone morphogenetic protein 7 (BMP7), were determined to be overexpressed in less than half of all 45 lung adenocarcinoma tissues compared to the paired normal counterparts. There were 17 genes, including the erythropoietin receptor (EPOR), interleukin 11 (IL11), the leukemia inhibitory factor (LIF), chemokine (C-X-C motif) ligand 1 (CXCL1), inhibin beta E (INHBE), tumor necrosis factor receptor superfamily member 21 (SF21), collagen type XI alpha 1 (COL11A1), protein kinase STYK1 (STYK1), v-erb-b2 erythroblastic leukemia viral oncogene homolog 2 (ERBB2), protein inhibitor of activated STAT 3 (PIAS3), dual specificity phosphatase 6 (DUSP6), fibroblast growth factor receptor 2 (FGFR2), MAPK13, MAPK4, mitogen-activated protein kinase 8

interacting protein 2 (MAPK8IP2), the phospholipase A2 group III (PLA2G3), and the protein phosphatase 1B magnesium-dependent beta isoform (PPM1B) that were overexpressed in all 45 lung adenocarcinoma tissues compared to their normal counterparts.

Discussion

The high mortality of lung cancer is largely attributable to difficulties in early diagnosis and the lack of effective therapeutic methods. In order to contribute to the improvement of survival rates in lung cancer patients, a comprehensive analysis of the molecular signature in biological processes is mandatory for identifying better tumor markers for diagnosis and new molecular targets for drug development. Microarray-based analysis is a promising new technology that allows the rapid detection of gene expression profiles of large numbers of genes simultaneously (14,15). Nowadays, adequate bioinformatics approaches are developed at a rapid pace to facilitate the information management. Although many studies have explored genes differentially expressed between tumor tissues and normal lung tissue in lung adenocarcinoma patients (1,10,11,16-21), the combined microarray-bioinformatics which can provide insight into the complex networks of molecular pathways underlying the development of lung adenocarcinoma, have not been revealed.

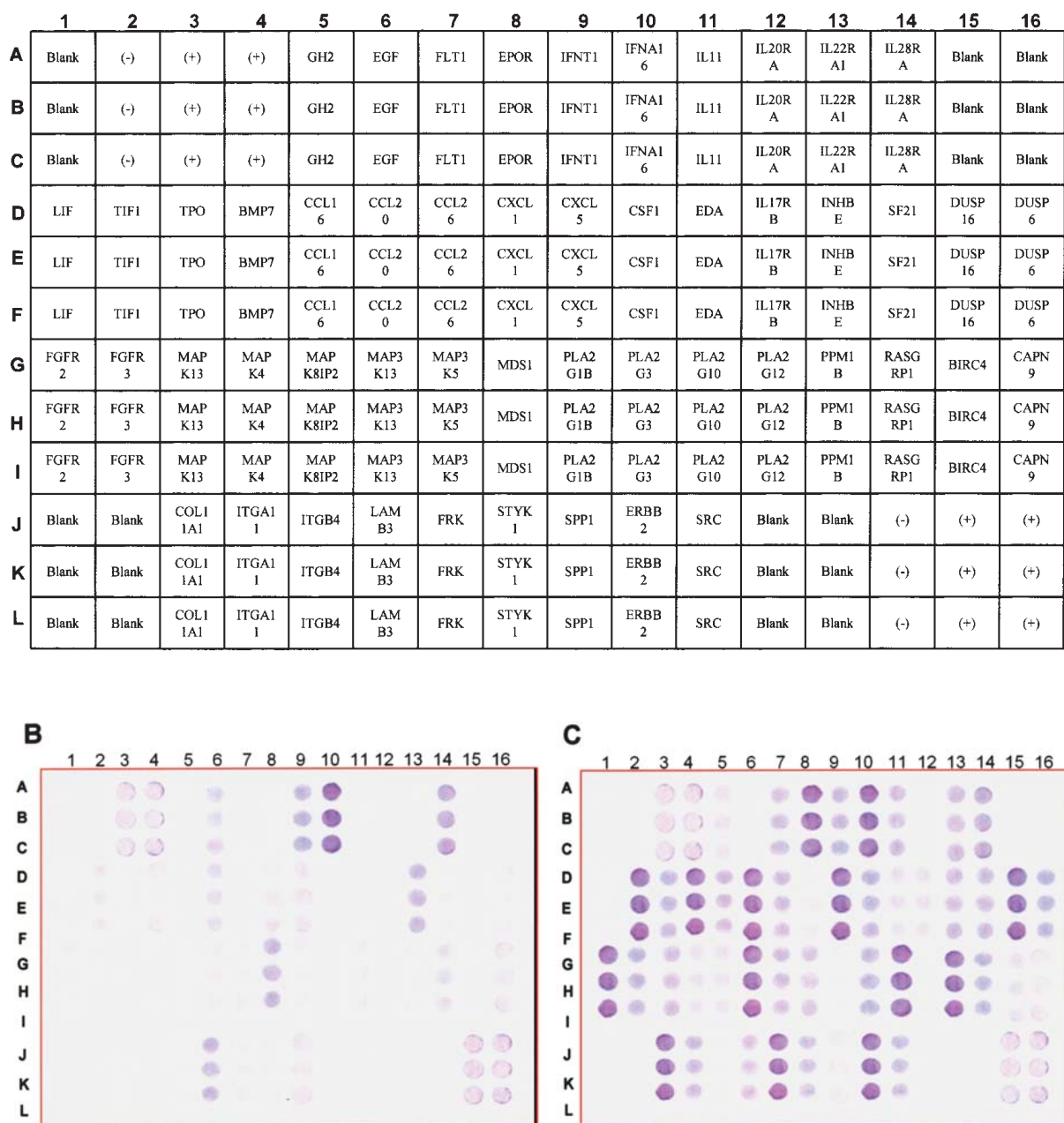


Figure 3. Membrane array analysis of the 51 up-regulated gene expression profile. (A) β -actin was used as an internal control, while DMSO was used as a blank control. Oligonucleotide fragments were blotted on membranes in triplicate. Each spot's intensity was normalized to β -actin, and the corrected expression level exhibiting at least twice the level of the corresponding normal control was determined as overexpression. (See the definition of abbreviations in Table II). (B) Result of membrane array of normal lung tissue gene expression. (C) Result of membrane array of lung adenocarcinoma gene expression.

In the present study, using a combined oligonucleotide microarray-bioinformatics method, we demonstrated 3,361 differentially expressed genes, and three potentially significant molecular pathways participating in lung adenocarcinoma carcinogenesis. Furthermore, our constructed membrane arrays with 51 related genes verified the similar gene expression patterns in additional lung adenocarcinoma tissues.

We identified 1,396 up-regulated genes and 1,965 down-regulated genes in the present investigation using the oligonucleotide microarray analysis. Some of our results were the same as those of other lung adenocarcinoma reports. Using the same Agilent's Human 1A Oligo Microarray, Li *et al* identified 275 genes that showed increased mRNA levels in 6

lung adenocarcinoma cell lines (22). Thirty-four genes had the same increased expression pattern which was consistent with our results. Among these 34 genes, COL11A1, SPP1 and ERBB2 were the up-regulated genes that involve the focal adhesion pathway. Although there were some differences between the cell lines and lung adenocarcinoma specimens, it is meaningful to use our native lung adenocarcinoma to detect our native lung adenocarcinoma carcinogenesis. Examining the early stage (stage I) NSCLC by high-resolution comparative genomic hybridization, 18 genes were found to have distinct differences in primary lung adenocarcinoma compared to their paired normal lung tissue (23). The subsequent fluorescence *in situ* hybridization analysis of these 18 genes found that there was a strong correlation with

cDNA microarray comparative genomic hybridization. Comparing our results with these 18 overexpressed genes, 6 genes had the same expression pattern, especially the tumor protein p73-like (TP73L), the small proline-rich protein (SPRP) and the SRY (sex determining region Y)-box 2 (SOX2). These were the up-regulated genes. The function of these 3 genes can enhance the initiation and progression of lung tumorigenesis and can be developed as new biomarkers for the early detection and classification of lung cancer.

In order to characterize the differentially expressed genes involved in the various biological processes in lung adenocarcinoma carcinogenesis, the pathways with the most frequently participating genes were categorized into the cytokine-cytokine receptor interaction pathway, the focal adhesion pathway and the MAPK signaling pathway. The three pathways are prominently correlated to cell signaling transduction, and are therefore thought to play a direct role in the early stages of tumor cell development and growth (24-28). Pathological angiogenesis is a hallmark of cancer and enhanced tumor cell growth and migration (29). The vascular endothelial growth factor (VEGFs)/VEGF receptor (VEGFRs) system also plays an important role in tumor growth and metastasis (28). Coagulation can facilitate tumor cells spreading into the pulmonary vasculature during early metastatic colony formation (30). In our study, several platelet-derived growth factor (PDGF) family genes, including PDGF, VEGFR, and the basic fibroblast growth factor (bFGF), had significant differential expressions in addition to the hematopoietin genes. The activation of the cytokine-cytokine receptor interaction pathway enhances the angiogenesis of lung adenocarcinoma and promotes cancer cell invasion. These angiogenesis markers had a great potential for predicting the prognosis of lung cancer (31).

Interferon alpha 16 (IFNA16), an important gene of the cytokine-cytokine receptor interaction pathway, had increased its expression ratio parallel to the advanced stages of lung cancer in our subsequent membrane array analysis (stage I: 33%, stage II: 33%, stage III: 50%, stage IV: 50%). IFN-alpha has emerged as an important regulator of cancer cell growth and differentiation and affected cellular communication and signal transduction. The protection of tumor cells from apoptosis induced by IFN-alpha is dependent on the EGF-mediated Ras/extracellular signal regulated kinase (Erk) signaling (32). Activated IFNA16 can facilitate invasion and metastasis in the advanced stages of lung adenocarcinoma and can be an excellent index for the molecular staging of lung cancer.

The focal adhesion pathway can modulate cell motility, cell proliferation, the cell cycle and cell survival (33-35). The focal adhesion kinase (FAK), a non-receptor tyrosine kinase that is enriched in focal adhesions, is an important mediator of integrin signaling. The adhesion and spreading of cells on a variety of extracellular matrix (ECM) proteins leads to an increase in tyrosine phosphorylation and the activation of FAK. The binding of FAK with Src also links FAK to the adaptor protein, Grb2, and then to the Ras signaling pathway. The downstream target of the Ras signaling pathway is the MAPK pathway, which is the third important pathway in our present study. The overexpression and phosphorylation of the focal adhesion kinase correlates with increased motility,

invasion and alteration in the cytoskeleton of tumor cells. However, in our report, FAK did not have a significant overexpression (T/N ratio: 1.32) although some upstream genes of FAK, such as the protein kinase C, Src, were overexpressed. The reasons for the underexpression of FAK could be the early stage lung adenocarcinoma samples used for the microarray analysis, which have certain factors disrupting the interaction of the upstream genes and FAK. The exact mechanisms of the focal adhesion pathway involving lung adenocarcinoma carcinogenesis need further investigation.

The Baculoviral IAP repeat-containing protein gene (BIRC4), a gene of the focal adhesion pathway, can inhibit caspases and suppress apoptosis. It had an increased expression ratio in the advanced stages of lung adenocarcinoma by membrane array analysis (stage I: 17%, stage II: 33%, stage III: 50%, stage IV: 75%). These results are consistent with the conclusion that the increased expression of BIRC4 contributes to the anoikis resistance of the circulating metastatic human prostate carcinoma cells (36). The down-regulation of the BIRC4 gene can induce apoptosis. The inhibition of the BIRC4 gene expression can also sensitize NSCLC to the cytotoxic effects of doxorubicin, Taxol, vinorelbine, and etoposide (37). BIRC4 could be a valuable target for cancer gene-therapy in human NSCLC, as well as a prognostic marker.

The MAPK pathway can elucidate signal transduction and was the downstream pathway of the cytokine-cytokine receptor interaction and focal adhesion pathways. Nicotine is one of the primary constituents in cigarette smoke. Several reports have shown that nicotine is responsible for airway irritation evoked by cigarette smoke (38,39). In our previous report, we demonstrated that the MAPK pathway is significantly activated in human bronchial epithelial cells by nicotine (Tsai *et al*) (21). Epidemiological evidence has revealed a statistically significant association between exposure to environmental tobacco smoke and lung cancer risk (8). Of the three major lung cancer subtypes, adenocarcinoma of the lung is strongly associated with cigarette smoking (9). Nicotine can play a critical role in the pathogenesis of lung adenocarcinoma carcinogenesis through the MAPK pathway activation. With the MAPK pathway antagonists inhibiting the production of IL-8, a well-known proinflammatory mediator could be correlated with tumor angiogenesis and progression (40), and could also have benefits in the treatment of lung cancer.

There are some limitations in this study. Three pairs of early stage lung adenocarcinoma samples were used for the microarray analysis and this limited the discovery of gene involvement in the advanced stages of lung cancer. The values of these 51 genes were also limited to the 15 cases of lung adenocarcinoma. More lung adenocarcinoma specimens are needed in order to prove the exact carcinogenesis pathway of lung adenocarcinoma.

In conclusion, we used microarray analysis with bioinformatics to demonstrate the signaling pathways of lung adenocarcinoma carcinogenesis. The analyses of the GO annotation tool (GOAT) revealed that of these 3,361 genes, 1,162 (34%) and 415 (12%) were involved in functions associated with cell binding and signal transduction, respectively. The three most-abundant classes of differentially

expressed genes were related to cytokine-cytokine receptor interaction, focal adhesion, and the MAPK signaling pathway. The consideration of the MAPK pathway as the core pathway could give profound insight to bundles of genes contributing to coordinate and integral functions in lung carcinogenesis. Our findings have shed some light on the mystery of the mechanisms underlying the carcinogenesis of lung adenocarcinoma in modern lung cancer research, and could also be helpful in seeking potential diagnostic/therapeutic markers for patients with early lung cancer.

Acknowledgements

The authors wish to express their thanks to the generous research support provided by Drs TF Chen and OL Hsu (the founders of Sunrider International), through the Give2Asia Foundation, to the Kaohsiung Medical University Education and Development Fund.

References

- Borcuk AC, Gorenstein L, Walter KL, Assaad AA, Wang L and Powell CA: Non-small-cell lung cancer molecular signatures recapitulate lung developmental pathways. *Am J Pathol* 163: 1949-1960, 2003.
- Yankelevitz DF, Reeve AP, Kostis WJ, Zhao B and Henschke CI: Small pulmonary nodules: volumetrically determined growth rates based on CT evaluation. *Radiology* 217: 251-256, 2000.
- Devesa SS, Bray F, Vizcaino AP and Parkin DM: International lung cancer trends by histologic type: male:female differences diminishing and adenocarcinoma rates rising. *Int J Cancer* 117: 294-299, 2005.
- Graziano SL, Gamble GP, Newman NB, Abbott LZ, Rooney M, Mookherjee S, Lamb ML, Kohman LJ and Poesz BJ: Prognostic significance of K-ras codon 12 mutations in patients with resected stage I and II non-small-cell lung cancer. *J Clin Oncol* 17: 668-675, 1999.
- Sugio K, Kishimoto Y, Virmani AK, Hung JY and Gazdar AF: K-ras mutations are a relatively late event in the pathogenesis of lung carcinomas. *Cancer Res* 54: 5811-5815, 1994.
- Niklinska W, Chyczewski L, Laudanski J, Sawicki B and Niklinski J: Detection of P53 abnormalities in non-small cell lung cancer by yeast functional assay. *Folia Histochem Cytobiol* 39: 147-148, 2001.
- McDoniels-Silvers AL, Stoner GD, Lubet RA and You M: Differential expression of critical cellular genes in human lung adenocarcinomas and squamous cell carcinomas in comparison to normal lung tissues. *Neoplasia* 4: 141-150, 2002.
- Sasco AJ, Secretan MB and Straif K: Tobacco smoking and cancer: a brief review of recent epidemiological evidence. *Lung Cancer* 45: S3-S9, 2004.
- Yang P, Cerhan JR, Vierkant RA, Olson JE, Vachon CM, Limburg PJ, Parker AS, Anderson KE and Sellers TA: Adenocarcinoma of the lung is strongly associated with cigarette smoking: further evidence from a prospective study of women. *Am J Epidemiol* 156: 1114-1122, 2002.
- Cheng S, Gao Y, Dong X, Lu Y, An Q, Tong T and Wang Y: Molecular and cytogenetic alterations in early stage of carcinogenesis of human lung. *Cancer Lett* 162 Suppl: S5-S10, 2001.
- Forgacs E, Zochbauer-Muller S, Olah E and Minna JD: Molecular genetic abnormalities in the pathogenesis of human lung cancer. *Pathol Oncol Res* 7: 6-13, 2001.
- Wang JY, Yeh CS, Chen YF, Wu CH, Hsieh JS, Huang TJ, Huang SY and Lin SR: Development and evaluation of a colorimetric membrane-array method for the detection of circulating tumor cells in the peripheral blood of Taiwanese patients with colorectal cancer. *Int J Mol Med* 17: 737-747, 2006.
- Yeh CS, Wang JY, Cheng TL, Juan CH, Wu CH and Lin SR: Fatty acid metabolism pathway play an important role in carcinogenesis of human colorectal cancers by Microarray-Bioinformatics analysis. *Cancer Lett* 233: 297-308, 2006.
- Kurian KM, Watson CJ and Wyllie AH: DNA chip technology. *J Pathol* 187: 267-271, 1999.
- Lockhart DJ, Dong H, Byrne MC, Follettie MT, Gallo MV, Chee MS, Mittmann M, Wang C, Kobayashi M, Horton H and Brown EL: Expression monitoring by hybridization to high-density oligonucleotide arrays. *Nat Biotechnol* 14: 1675-1780, 1996.
- Bhattacharjee A, Richards WG, Staunton J, Li C, Monti S, Vasa P, Ladd C, Beheshti J, Bueno R, Gillette M, Loda M, Webe G, Mark EJ, Lander ES, Wong W, Johnson BE, Golub TR, Sugarbaker DJ and Meyerson M: Classification of human lung carcinomas by mRNA expression profiling reveals distinct adenocarcinoma subclasses. *Proc Natl Acad Sci USA* 98: 13790-13795, 2001.
- Garber ME, Troyanskaya OG, Schluens K, Petersen S, Thaesler Z, Pacyna-Gengelbach M, van de Rijn M, Rosen GD, Perou CM, Whyte RI, Altman RB, Brown PO, Botstein D and Petersen I: Diversity of gene expression in adenocarcinoma of the lung. *Proc Natl Acad Sci USA* 98: 13784-13789, 2001.
- Nakamura N, Kobayashi K, Nakamoto M, Kohno T, Sasaki H, Matsuno Y and Yokota J: Identification of tumor markers and differentiation markers for molecular diagnosis of lung adenocarcinoma. *Oncogene* 30: 4245-4255, 2006.
- Shibata T, Uryu S, Kokubu A, Hosoda F, Ohki M, Sakiyama T, Matsuno Y, Tsuchiya R, Kanai Y, Kondo T, Imoto I, Inazawa J and Hirohashi S: Genetic classification of lung adenocarcinoma based on array-based comparative genomic hybridization analysis: its association with clinicopathologic features. *Clin Cancer Res* 11: 6177-6185, 2005.
- Wikman H, Kettunen E, Seppanen JK, Karjalainen A, Hollmen J, Anttila S and Knuutila S: Identification of differentially expressed genes in pulmonary adenocarcinoma by using cDNA array. *Oncogene* 21: 5804-5813, 2002.
- Sheu CC, Chang MY, Chang HC, Tsai JR, Lin SR, Chang SJ, Hwang JJ, Huang MS and Chong IW: Combined detection of CEA, CK-19 and c-met mRNAs in peripheral blood: a highly sensitive panel for potential molecular diagnosis of non-small cell lung cancer. *Oncology* 3: 203-211, 2006.
- Li R, Wang H, Bekele BN, Yin Z, Caraway NP, Katz RL, Stass SA and Jiang F: Identification of putative oncogenes in lung adenocarcinoma by a comprehensive functional genomic approach. *Oncogene* 25: 2628-2635, 2006.
- Jiang F, Yin Z, Caraway NP, Li R and Katz RL: Genomic profiles in stage I primary non small cell lung cancer using comparative genomic hybridization analysis of cDNA microarrays. *Neoplasia* 6: 623-635, 2004.
- Michiel DF and Oppenheim JJ: Cytokines as positive and negative regulators of tumor promotion and progression. *Semin Cancer Biol* 3: 3-15, 1992.
- Mukohara T, Kudoh S, Yamauchi S, Kimura T, Yoshimura N, Kanazawa H, Hirata K, Wanibuchi H, Fukushima S, Inoue K and Yoshikawa J: Expression of epidermal growth factor receptor (EGFR) and downstream-activated peptides in surgically excised non-small-cell lung cancer (NSCLC). *Lung Cancer* 41: 123-130, 2003.
- Murakami T, Cardones AR and Hwang ST: Chemokine receptors and melanoma metastasis. *J Dermatol Sci* 36: 71-78, 2004.
- Sunaga N, Miyajima K, Suzuki M, Sato M, White MA, Ramirez RD, Shay JW, Gazdar AF and Minna JD: Different roles for caveolin-1 in the development of non-small cell lung cancer versus small cell lung cancer. *Cancer Res* 64: 4277-4285, 2004.
- Tanno S, Ohsaki Y, Nakanishi K, Toyoshima E and Kikuchi K: Human small cell lung cancer cells express functional VEGF receptors, VEGFR-2 and VEGFR-3. *Lung Cancer* 46: 11-19, 2004.
- Carmeliet P and Jain RK: Angiogenesis in cancer and other diseases. *Nature* 407: 249-257, 2000.
- Im JH, Fu W, Wang H, Bhatia SK, Hammer DA, Kowalska MA and Muschel RJ: Coagulation facilitates tumor cell spreading in the pulmonary vasculature during early metastatic colony formation. *Cancer Res* 64: 8613-8619, 2004.
- Singhal S, Vachani A, Antin-Ozerkis D, Kaiser LR and Albelda SM: Prognostic implications of cell cycle, apoptosis, and angiogenesis biomarkers in non-small cell lung cancer (Review). *Clin Cancer Res* 11: 3974-3986, 2005.
- Caraglia M, Vitale G, Marra M, Budillon A, Tagliaferri P and Abbruzzese A: Alpha-interferon and its effects on signalling pathways within cells. *Curr Protein Pept Sci* 5: 475-485, 2004.
- Mitra SK, Hanson DA and Schlaepfer DD: Focal adhesion kinase: in command and control of cell motility. *Nat Rev Mol Cell Biol* 6: 56-68, 2005.

34. Schlaepfer DD and Mitra SKL: Multiple connections link FAK to cell motility and invasion. *Curr Opin Genet Dev* 14: 92-101, 2004.
35. Turner CE: Paxillin and focal adhesion signalling. *Nat Cell Biol* 2: E231-E236, 2000.
36. Berezovskaya O, Schimmer AD, Glinskii AB, Pinilla C, Hoffman RM, Reed JC and Glinsky GV: Increased expression of apoptosis inhibitor protein XIAP contributes to anoikis resistance of circulating human prostate cancer metastasis precursor cells. *Cancer Res* 65: 2378-2386, 2005.
37. Hu Y, Cherton-Horvat G, Dragowska V, Baird S, Korneluk RG, Durkin JP, Mayer LD and LaCasse EC: Antisense oligonucleotides targeting XIAP induce apoptosis and enhance chemotherapeutic activity against human lung cancer cells *in vitro* and *in vivo*. *Clin Cancer Res* 9: 2826-2836, 2003.
38. Lee LY, Gerhardstein DC, Wang AL and Burki NK: Nicotine is responsible for airway irritation evoked by cigarette smoke inhalation in men. *J Appl Physiol* 75: 1955-1961, 1993.
39. Totti N III, McCusker KT, Campbell EJ, Griffin GL and Senior RM: Nicotine is chemotactic for neutrophils and enhances neutrophil responsiveness to chemotactic peptides. *Science* 223: 169-171, 1984.
40. Yuan A, Yang PC, Yu CJ, Chen WJ, Lin FY, Kuo SH and Luh KT: Interleukin-8 messenger ribonucleic acid expression correlates with tumor progression, tumor angiogenesis, patient survival, and timing of relapse in non-small-cell lung cancer. *Am J Respir Crit Care Med* 162: 1957-1963, 2000.
41. Mountain CF: Revisions in the International system for staging lung cancer. *Chest* 111: 1710-1717, 1997.
42. Chomczynski and Sacchi N: Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. *Anal Biochem* 162: 156-159, 1987.