Am J Cancer Res 2014;4(3):256-269
www.ajcr.us /ISSN:2156-6976/ajcr0000530

Original Article

Protein signature for non-small cell lung cancer prognosis

Wei Liu[1], Yong Wu[1], Libo Wang[3], Ling Gao[4], Yingping Wang[5], Xiaoliang Liu[4], Kai Zhang[2], Jena Song[2], Hongxia Wang[2], Thomas A Bayer[2], Laurel Glaser[2], Yezhou Sun[6], Weijia Zhang[6], Michael Cutaia[7], David Y Zhang[2], Fei Ye[2]

Departments of [1]Thoracic Surgery, [3]Endoscopic Center, [4]Cancer Center, [5]Pathology, The First Hospital of Jilin University, Changchun, Jilin 130021, China; Departments of [2]Pathology, [6]Medicine, Mount Sinai School of Medicine, New York, NY 10029, USA; [7]Department of Medicine, SUNY/Downstate Health Sciences Center, New York Harbor Health Care System, Brooklyn, NY 11203, USA

Received April 14, 2014; Accepted April 23, 2014; Epub May 26, 2014; Published June 1, 2014

Abstract: Background: Current histopathological classification and TNM staging have limited accuracy in predicting survival and stratifying patients for appropriate treatment. The goal of the study is to determine whether the expression pattern of functionally important regulatory proteins can add additional values for more accurate classification and prognostication of non-small lung cancer (NSCLC). Methods: The expression of 108 proteins and phosphoproteins in 30 paired NSCLC samples were assessed using Protein Pathway Array (PPA). The differentially expressed proteins were further confirmed using a tissue microarray (TMA) containing 94 NSCLC samples and were correlated with clinical data and survival. Results: Twelve of 108 proteins (p-CREB(Ser133), p-ERK1/2(Thr202/Tyr204), Cyclin B1, p-PDK1(Ser241), CDK4, CDK2, HSP90, CDC2p34, β-catenin, EGFR, XIAP and PCNA) were selected to build the predictor to classify normal and tumor samples with 97% accuracy. Five proteins (CDC2p34, HSP90, XIAP, CDK4 and CREB) were confirmed to be differentially expressed between NSCLC (n=94) and benign lung tumor (n=19). Over-expression of CDK4 and HSP90 in tumors correlated with a favorable overall survival in all NSCLC patients and the over-expression of p-CREB(Ser133) and CREB in NSCLC correlated with a favorable survival in smokers and those with squamous cell carcinoma, respectively. Finally, the four proteins (CDK4, HSP90, p-CREB and CREB) were used to calculate the risk score of each individual patient with NSCLC to predict survival. Conclusion: In summary, our data demonstrated a broad disturbance of functionally important regulatory proteins in NSCLC and some of these can be selected as clinically useful biomarkers for diagnosis, classification and prognosis.

Keywords: Lung cancer, biology marker, survival analysis

Background

Lung cancer is the leading cause of cancer-related mortality in both men and women worldwide, and is expected to account for 26% of all female cancer deaths and 28% of all male cancer deaths in the United States in 2013 [1]. Lung cancer can be classified into 4 major histological types, including adenocarcinoma, squamous carcinoma, large cell carcinoma, and small cell carcinoma. The first 3 types are collectively called Non-Small Cell Lung Cancer (NSCLC), which accounts for about 85% of lung cancer. Despite significant advancement in diagnostic tests, surgical techniques and therapeutic agents, the overall 5-year survival is only 12-16% [2, 3]. The mortality rate is high because approximately two thirds of patients are diagnosed at an advanced stage, for which curative treatment is not available. Although patients diagnosed with early-stage NSCLC have an overall 5-year survival rate of 63%, nearly 35% will relapse after surgical resection [4, 5]. Therefore, identification of biomarkers for early diagnosis, accurate prognosis, and reliable prediction of treatment response has become an urgent task.

In the past years, the molecular mechanisms responsible for lung cancer development have been extensively studied [6-8]. However, the biomarkers for early diagnosis, therapeutic

response prediction and survival prognosis are still limited due to lack of understanding about the multifactorial process of lung carcinogenesis and the heterogeneous nature of the disease. The most important clinical prognostic marker in NSCLC is stage [9]. Molecular prognostic markers that have been studied in NSCLC include mutations in EGFR, k-ras, and p53 [10, 11] and mRNA expression levels of ERCC1, Bcl-2, and c-erbB-2 [12, 13]. Also, the extensive genomic studies of lung cancer demonstrated the ability of mRNA-based gene expression profiles to predict survival in NSCLC [14, 15]. To date, however, very few prognostic protein markers have been accepted for routine clinical use.

Unlike genomic studies, where individual changes may not have functional significance, protein expression is closely aligned with cellular function and activity. The proteomic profiling of functionally important regulatory proteins in cancer cells may shed light on the molecular mechanisms of cancer development and metastasis. We recently developed a powerful Protein Pathway Array (PPA) analysis [16-19] that allows for the identification of important, but low abundance, proteins and phosphoproteins in NSCLC. The proteins included in the PPA analysis are important in carcinogenesis, including cell cycle regulation and proliferation, adhesion, migration, invasion, metastasis and angiogenesis [16-19].

In this study, the PPA analysis was used to evaluate 108 functionally important proteins and phosphoproteins and identified a robust set of proteins that are differentially expressed in NSCLC. The changes of these proteins were further confirmed in an independent NSCLC cohort using Tissue Microarray (TMA) and demonstrated the correlations of these proteins with clinical outcomes. These proteins may serve as diagnostic and prognostic markers for NSCLC.

Materials and methods

Tissue specimens and patient characteristics

For Protein Pathway Array analysis (training cohort), a total of 30 fresh frozen tumor specimens along with 30 adjacent normal lung tissues (at least 1 cm away from tumor tissue) (**Table 1**) were obtained during surgical resection from patients with primary non-small cell lung cancer at The First Hospital of Jilin University, Changchun, China, between May 2008 and May 2010. These 30 pairs of NSCLC specimens were used to identify differentially expressed proteins between NSCLC and normal lung tissues. For Tissue Microarray Analysis (TMA) (validation cohort), 94 formalin-fixed paraffin-embedded NSCLC specimens (**Table 1**) as well as 19 benign lung tumors were retrieved from the archives of the Department of Pathology at the First Hospital of Jilin University between January 2006 and December 2009. The tissue samples were from 94 NSCLC patients who did not receive preoperative chemotherapy. Data concerning clinicopathological characteristics, including age, gender, smoking history (smokers or nonsmokers), pathologic TNM stage and histologic subtypes (adenocarcinoma or squamous-cell carcinoma) were obtained via medical record review. Patients who smoked at least 20 packs/year [20] were defined to be smokers. The pathological TNM stage was determined according to the International Union Against Cancer classification scheme [21]. This study was reviewed and approved by the Institutional Ethical Review Boards of the First Hospital of Jilin University, and written informed consent for research use of specimens was obtained from all patients.

Protein pathway array analysis

Total proteins were extracted from 30 paired fresh frozen NSCLC tumor tissues and the surrounding non-tumor tissues using 1×cell lysis buffer (Cell Signaling Technology, Danvers, MA) containing 20 mM Tris-HCL (pH 7.5), 150 mM NaCl, 1 mM Na_2EDTA, 1 mM EGTA, 1% Triton X-100, 2.5 mM sodium pyrophosphate, 1 mM β-glycerophosphate, 1 mM Na_3VO_4, and 1 μg/ml leupeptin in the presence of 1×proteinase inhibitor cocktail (Roche Applied Science, Indianapolis, IN) and 1×phosphatase inhibitor cocktail (Roche Applied Science, Indianapolis, IN). The lysate was sonicated three times for 15 seconds each time, and then centrifuged at 14,000 rpm for 30 minutes at 4°C. The protein concentration was determined with the BCA Protein Assay kit (PIERCE, Rockford, IL). Three hundred μg of lysated protein was loaded in one well across the entire width of 10% SDS polyacrylamide and separated by electrophoresis as described previously [16-19]. After electrophoresis, the proteins were transferred to a

Table 1. Characteristics of NSCLC patients

Clinicopathological Characteristics	Patient number (%)	
	PPA (n=30)	TMA (n=94)
Age		
≤60 yr	21 (70)	46 (49)
>60 yr	9 (30)	48 (51)
Sex		
Male	19 (63)	70 (74)
Female	11 (37)	24 (26)
Smoking		
Yes	18 (60)	61 (65)
No	12 (40)	33 (35)
Pathological TNM stage		
Stage I	13 (43)	42 (45)
Stage II	8 (27)	26 (28)
Stage III	8 (27)	26 (28)
Stage IV	1 (3)	0 (0)
Tumor		
T1	5 (17)	3 (3)
T2	19 (63)	67 (71)
T3	0 (0)	17 (18)
T4	6 (20)	7 (7)
Node status		
N0	16 (53)	50 (53)
N1	11 (37)	29 (31)
N2	3 (10)	15 (16)
Distant Metastasis		
M0	29 (97)	94 (100)
M1	1 (3)	0 (0)
Histologic type		
Squamous cell carcinoma	18 (60)	57 (61)
Adenocarcinoma	12 (40)	37 (39)

Note: PPA, Protein Pathway Array; TMA, Tissue Microarray.

nitrocellulose membrane (Bio-rad, Hercules, CA) which was then blocked for 1 hour with blocking buffer including either 5% milk or 3% BSA in 1×TBST containing 20 mM Tris-HCl (pH 7.5), 100 mM NaCl and 0.1% Tween-20. Next, the membrane was clamped on a Western blotting manifold (Mini-PROTEAN II Multiscreen apparatus, Bio-Rad, Hercules, CA) that isolates 20 channels across the membrane. The multiplex immunoblot was performed using a total of 108 protein-specific or phosphorylation site-specific antibodies (An additional table file shows this in more detail [see Table S1]). Three sets of antibodies (a total of 36 protein-specific or phosphorylation site-specific antibodies per set) were individually used for each membrane and all antibodies (from various companies) were validated independently before inclusion in PPA. For the first set of 36 primary antibodies, a mixture of two antibodies in the blocking buffer were added to each channel and then incubated at 4°C overnight. The membrane was then washed with 1×TBS and 1×TBST, and was further incubated with secondary anti-rabbit or anti-mouse antibody conjugated with horseradish peroxidase (Bio-rad, Hercules, CA) for 1 hour at room temperature. The membrane was developed with chemiluminescence substrate (Immun-Star™ HRP Peroxide Buffer/ Immun-Star™ HRP Luminol Enhancer, Bio-rad, Hercules, CA), and chemiluminescent signals were captured using the ChemiDoc XRS System (Bio-rad, Hercules, CA). The same membrane was then treated with stripping buffer (Restore™ Western blot stripping buffer, Thermo Scientific, Rockford, IL) and used to detect a second set of 36 primary antibodies as described above.

For PPA data analysis, the correct band of each protein was identified visually and the signal intensity of each protein was determined by densitometric scanning (Quantity One software package, Bio-rad). The background was locally subtracted from raw protein signal and the background-subtracted intensity was normalized by a "global median subtraction" normalization method to reduce the variations arising from different runs (such as transferring and blotting efficiency, total protein loading amount, and exposure density). In detail, each protein signal intensity was divided by the total intensities of all proteins in the same blot membrane, and then multiplied by average intensities of each protein in all samples. The normalized data were used in subsequent statistical analysis.

Tissue microarray and immunohistochemical staining

H&E stained slides of all cases were reviewed by a pathologist (Y. Wang) and 94 NSCLC FFPE blocks along with 19 benign tumor FFPE blocks were selected. For each tumor, an area with minimal necrosis was marked and two representative tissue cores (1.0 mm diameter each) were punched from each block. The TMA was constructed and stained at Shanghai Biochip

Company (Shanghai, China). Immunohistochemical staining was performed using a standard two-step indirect immunohistochemistry protocol [22]. Antibodies for p-CREB(Ser133) (1:50 dilution) and CREB (1:150 dilution) were purchased from Cell Signaling Technology (Danvers, MA). Antibodies for CDK4 (1:800 dilution), XIAP (1:400 dilution), CDC2p34 (1:1000 dilution), and HSP90 (1:1600 dilution) were obtained from Santa Cruz Biotechnology (Santa Cruz, CA). The immunohistochemical staining pattern of each antibody were reviewed by D. Zhang and a final histochemical score (H-score) was calculated as the product of staining intensities (0-3) and percentage of stained cells (0-100%) [23, 24]. In this study, high expression for CDC2p34, HSP90, XIAP, and CDK4 was defined when the H-score was greater than 70, 70, 100, and 90, respectively. However, low expression for p-CREB(Ser133) or CREB was defined when H-score was less than 20.

Statistical analysis

Paired t-test and Significant Analysis of Microarray (SAM) tool (http://www-stat.srd.edu/~tibs/SAM/) were used to identify the proteins differentially expressed between normal and tumor tissues. The leave-one-out cross-validation approach (LOOCV), including k-nearest neighbor (k=3) and support vector machine (SVM), was used to select proteins that can accurately classify tumor and benign tissues (http://linus.nci.nih.gov/BRB-ArrayTools.html) [25]. Chi-square test and Fisher's exact test were used to determine the association of expression of proteins with clinical parameters such as age, sex, smoking history, tumor stage, tumor size, and histology subtype. The Kaplan-Meier and the Cox proportional hazard regression were used for survival analysis. All statistical analyses, except SAM and LOOCV, were performed using SPSS version 17.0. Statistical significance was determined based on a two-sided significance level of 0.05.

Results

Clinicopathological characteristics of the patient cohorts

Two cohorts of lung cancer patients were included in this study: the Protein Pathway Array (PPA) cohort (training set) and the Tissue Microarray (TMA) cohort (validation set). The demographic and clinicopathological characteristics were summarized in **Table 1**. The median age of the patients in the PPA cohort was 58 years (ranged from 23 to 76) and 63% of the patients were men; while the median age of the patients in the TMA cohort was 60 years (ranged from 35 to 77) and 74% of the patients were men. The median follow-up for the TMA cohort was 51 months (ranged from 7 to 122 months, started from the day of diagnosis). The overall survival rate during follow-up period was 40.4% (38/94) and the 1-year, 3-year, and 5-year survival rates were 90%, 60%, and 44%, respectively.

Identification of differentially expressed proteins and phosphoproteins between tumor and normal tissues

Thirty pairs of NSCLC and surrounding normal tissues were initially used to screen the differentially expressed proteins using PPA (An additional figure file shows this in more detail [see Figure S1]). Among 108 proteins and phosphoproteins tested, 51 were detected in either tumors or normal tissues (An additional table file shows this in more detail [see Table S1]). Of these proteins, 21 showed significant differences between tumors and corresponding normal tissues using paired t-test ($p<0.05$) (**Figure 1**) and SAM ($q<0.05$). Fourteen proteins and phosphoproteins were predominantly over-expressed in the tumor tissues, including CDC2p34, HSP90, PCNA, p-PDK1(Ser241), XIAP, β-catenin, EGFR, CDK2, CDK4, p-CDC2p34(Tyr15), p53, 14-3-3-β, Notch4, and p-p38(Thr180/Tyr182) (ranked by *p*-value; from lowest to highest), and 7 were predominantly down-regulated in tumors, including Cyclin B1, p-CREB(Ser133), p-ERK1/2(Thr202/Tyr204), ERK1/2, NFKB p50, CDC-25C, and BCL-6 (ranked by *p*-value; from lowest to highest) (**Figure 1**). The ratios of the average expression level of each protein between tumors and normal tissues ranged from 9.01 fold (p-CDC2p34(Tyr15)) to 1.50 fold (p-p38(Thr180/Tyr182)) (**Figure 1**). The percentage of the proteins with increased or decreased expression (>1.5 fold) among 30 NSCLC tumors was ranged from 17.2 to 89.7% and 20.7 to 65.5%, respectively. The top five proteins with increased expression in 30 tumors were CDC2p34 (89.7%), HSP90 (86.2%), p-PDK1(Ser241) (75.9%), β-catenin (72.4%), and PCNA (69%). The top 5 proteins with decreased expression in 30 tumors were Cyclin B1 (72.4%), p-ER-

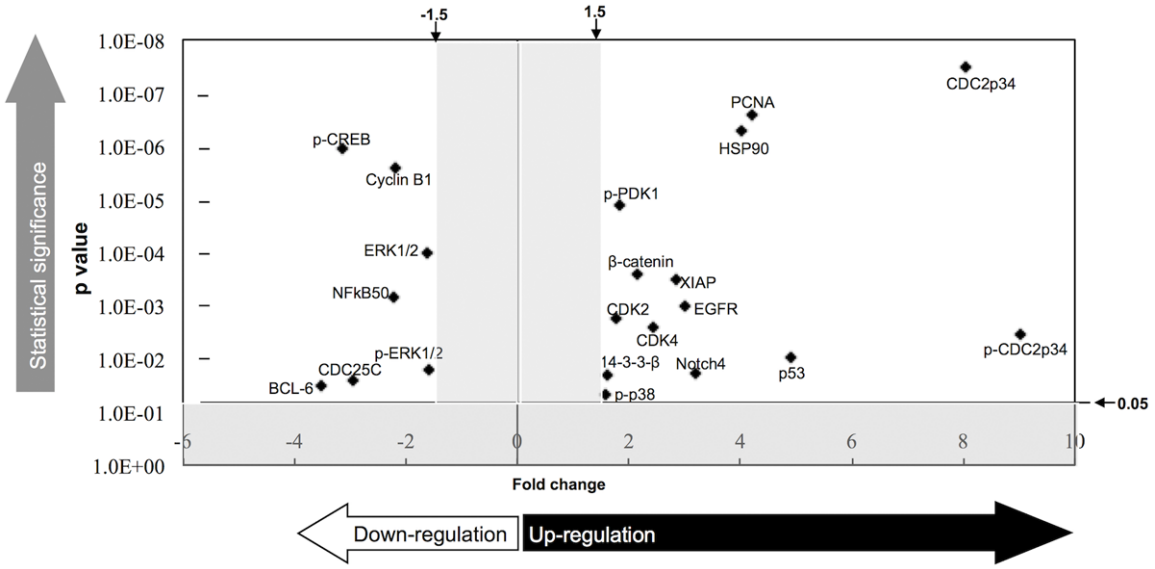

Figure 1. The differentially expressed proteins and phosphoproteins. A total of 21 proteins (14 up-regulated and 7 down-regulated) was considered significant in their expression between tumor and normal tissues. The expression difference for each protein was significant when the average fold difference was >1.5 folds between tumors and normal tissues and statistical significance was $p<0.05$.

K1/2(Thr202/Tyr204) (65.6%), p-CREB(Ser133) (58.6%), ERK1/2 (55.2%), and NFKB p50 (48.3%).

To obtain an optimal number of proteins for the accurate classification of tumor and normal, a supervised leave-one-out cross-validation using two class prediction models including a support vector machine (SVM) and 3-nearest neighbor (3NN) algorithms was performed with significance level between the two classes at $p<0.001$. Twelve proteins (p-CREB(Ser133), p-ERK1/2(Thr202/Tyr204), Cyclin B1, p-PDK-1(Ser241), CDK4, CDK2, HSP90, CDC2p34, β-catenin, EGFR, XIAP, and PCNA) were chosen to build SVM and 3NN predictor models with a cross-validation accuracy of 97%. The results showed that these predictor proteins including 3 down-regulated proteins (p-CREB(Ser133), p-ERK1/2(Thr202/Tyr204) and Cyclin B1) and 9 up-regulated proteins (p-PDK1(Ser241), CDK2, CDK4, HSP90, CDC2p34, β-catenin, EGFR, XIAP, and PCNA) in tumors can successfully distinguish between tumor and normal tissues.

Verification of the significant predictor protein expression in NSCLC using tissue microarray

To validate the PPA results, we selected 5 proteins (out of 12) to assess their expression patterns on a TMA. These proteins were selected based on the following criteria: 1) top ranked *p*-values based on t-test, 2) more than two-fold change in the expression level between cancer and normal tissues based on PPA, 3) involved in different, but important, regulatory pathways. CDC2p34 (p=2.43E-09) was increased 8.04 fold in cancer and is involved in cell progression from G2 to M phase. HSP90 (p=4.15E-08) was increased 4.02 fold in cancer and is involved in stress-related signaling pathway. Phosphorylated CREB (p=7.04E-06) was decreased 2.91 fold in cancer and is involved in CREB pathway. In addition, non-phosphorylated CREB was also assessed. XIAP (p=1.41E-04) was increased 2.83 fold in cancer and inhibits apoptosis. CDK4 (p=1.83E-03) was increased 2.45 fold in cancer and promotes G_1-S phase progression. Immunohistochemical staining was performed on 94 NSCLC samples and scored based on the percentage and intensity of immunopositivity. **Figure 2** showed that CDC2p34, HSP90, XIAP, and CDK4 were localized to the cytoplasm, and p-CREB and CREB to the nucleus. H-scores (mean score+1SD) were determined for CDC-2p34 (117+78), HSP90 (108+64), XIAP (115+55), CDK4 (119+70), p-CREB(Ser133) (76+92) and CREB (56+71). The percentage of cancer samples with high expression (CDC2-p34=68.1%, HSP90=84.0%, XIAP=41.5%, and

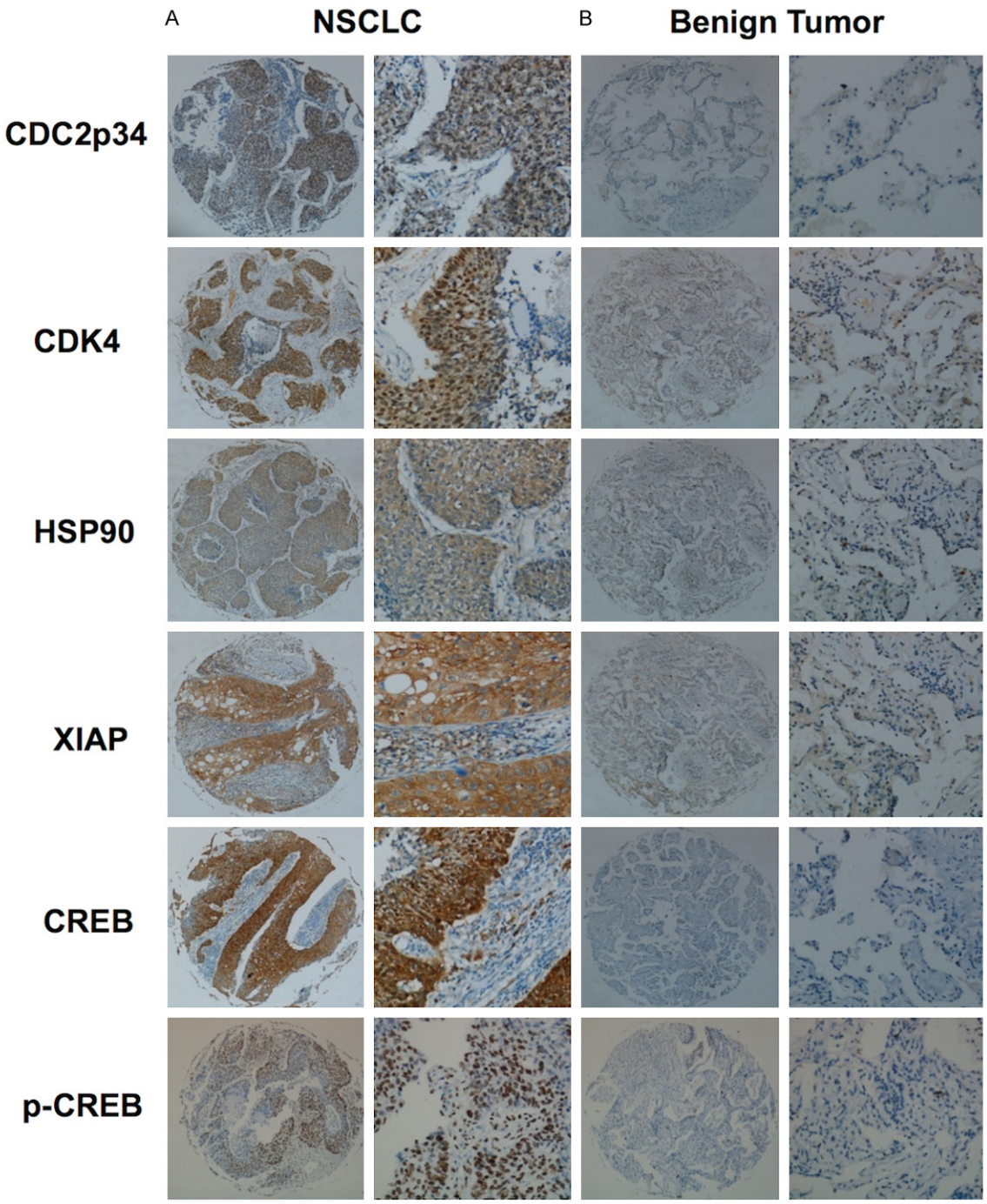

Figure 2. Representative images of immunohistochemical staining of TMA. A. NSCLC (Squamous cell carcinoma) (n=94). B. Benign lung tumor (hamartoma) (n=19). IHC included CDC2p34, CDK4, HSP90, XIAP and CREB.

CDK4=53.2%) was similar to that observed in PPA (CDC2p34=89.7%, HSP90=86.2%, XIAP=51.7%, and CDK4=55.2%) (An additional figure file shows this in more detail [see Figure S2]). The percentage of cancer samples with low expression of p-CREB(Ser133) in TMA (45.7%) was compatible with those by PPA (58.6%) (An additional figure file shows this in more detail [see Figure S2]). These data suggested an accurate quantification of protein expression by PPA, and its consistency with the TMA results.

Table 2. Correlation between clinicopathologic data and the level of protein expression

Clinical Data		CDC2 p34	HSP90	p-CREB	CREB	XIAP	CDK4
Gender	Male Female	p=0.090	p=0.592	p=0.642	p=0.904	p=0.156	p=0.001**
Age	≥60 y <60 y	p=0.485	p=0.858	p=0.439	p=0.650	p=0.728	p=0.464
Smoking	Yes No	p=0.253	p=0.875	p=0.695	p=0.038*	p=0.458	p=0.016*
Histology	Squamous-Cell Carcinoma Adenocarcinoma	p=0.001**	p=0.379	p=0.157	p=0.033*	p=0.020*	p=0.014*
Stage	Stage I-II Stage III	p=0.256	p=0.592	p=0.330	p=0.040*	p=0.921	p=0.191
T	T1-2 T3-4	p=0.738	p=0.912	p=0.151	p=0.904	p=0.327	p=0.717
N	N0 N1-2	p=0.050*	p=0.520	p=0.891	p=0.094	p=0.233	p=0.717

Note: Chi-square test was performed to determine the association of expression of each protein with the clinical parameters (*p<0.05, **p<0.01).

In order to determine whether the changes of these proteins are specific for NSCLC, 19 benign lung tumors, including inflammatory psuedotumor, hamartoma, lymphangiomyoma and sclerosing hemangioma, were included in the TMA. Five antibodies (CDC2p34, HSP90, XIAP, CDK4 and CREB) did not stain or weakly stained in benign tumors (**Figure 2B**) and the differences between NSCLCs and benign tumors were significant (p<0.05) (Table S2). However, there was a higher positive rate for p-CREB(Ser133) in benign tumors with p=0.166 between NSCLC and benign tumors (An additional table file shows this in more detail [see Table S2]). The results suggest that these protein, except for p-CREB(Ser133), are relatively specific for NSCLC (An additional table file shows this in more detail [see Table S2]).

Correlation between the protein expression status and clinicopathological variables and overall survivals

The potential correlation between the immunophenotypes and clinical parameters including gender, age, smoking history, histology subtypes, pathological stages, tumor size (T), and node status (N) was assessed (**Table 2**). The high expression of CDK4 was significantly correlated with male gender (p=0.001). Both CREB and CDK4 expressions were increased in NSCLC with a smoking history (p=0.038 and p=0.016, respectively). The high expressions of CDC2p34, CREB, CDK4, and XIAP were more frequently observed in squamous cell carcinoma than in adenocarcinoma (p=0.001, p=0.033, p=0.014, and p=0.02, respectively). The high expression of CREB showed a stronger association with early stage tumor (I-II) than with advanced stage tumor (III) (p=0.040), and the high expression of CDC2p34 showed a stronger association with node-positive tumors than in node-negative tumors (p=0.05). No significant correlations were observed between the protein expression and age and tumor size.

The relationship between the protein expression and overall survival of 94 NSCLC patients was examined. For this analysis, the expression levels of each protein, including CDC2p34, HSP90, XIAP, CDK4, p-CREB(Ser133), and CREB, were divided into two groups, i.e. low and high expression based on the H-scores (see Materials and Methods section). Among these 6 proteins, only CDK4 and HSP90 showed significant correlation with overall survival rate in NSCLC patients: high expression of CDK4 and HSP90 in NSCLC correlated with a favorable overall survival (log-rank p=0.007 and p=0.040, respectively) (**Figure 3A** and **3B**). Based on the univariate Cox regression analysis, hazard ratio for CDK4 was 0.488 (95% CI=0.286~0.834, p=0.009), and for HSP90 was 0.506 (95% CI=0.260~0.984, p=0.045).

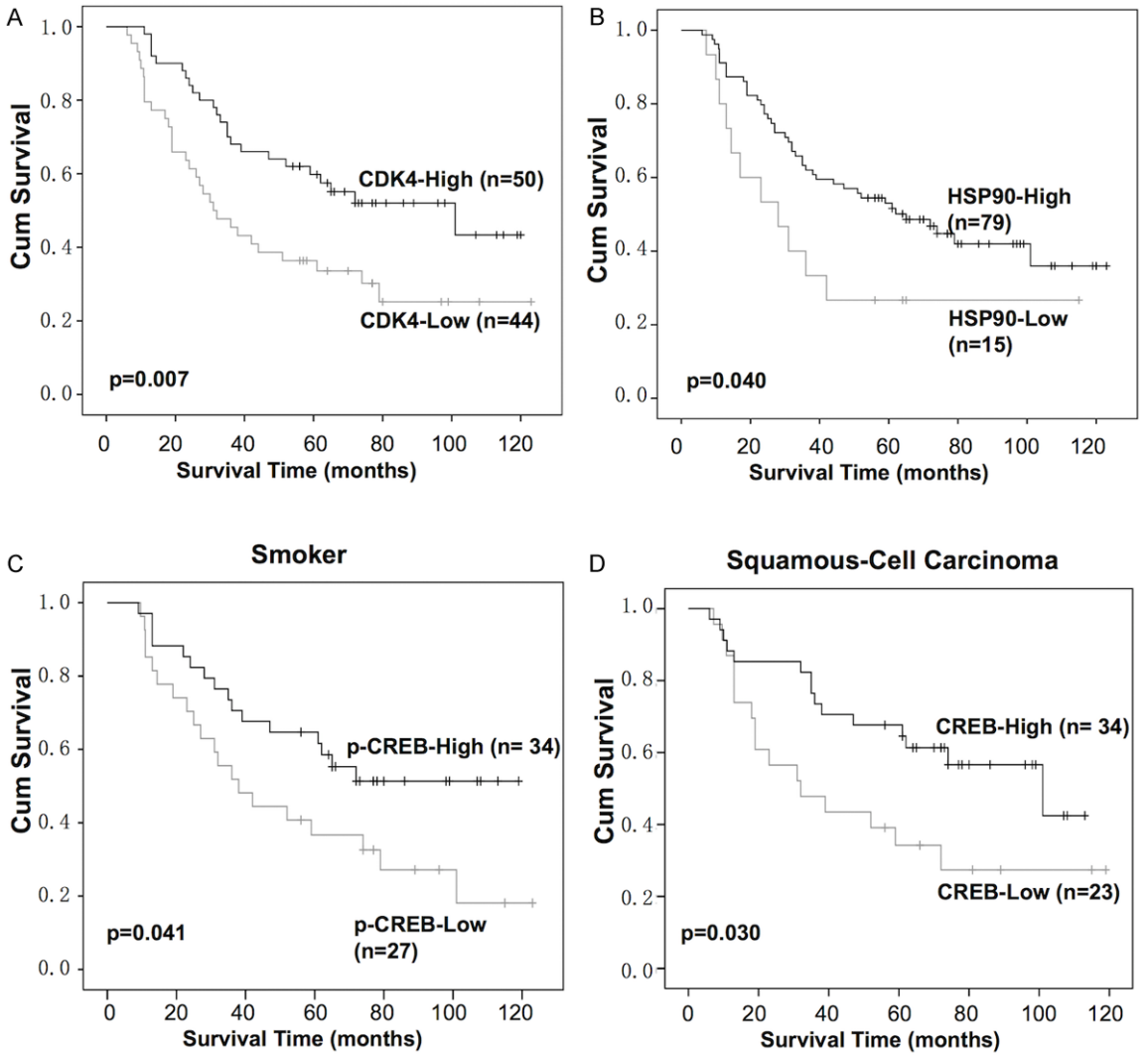

Figure 3. Kaplan-Meier survival analysis of patients with NSCLC based on the four protein expression levels. The overall survival of the patients with NSCLC (n=94) was better in the group with high level expression of CDK4 (A) or HSP90 (B). (C) Increased p-CREB in patients with smoking history (n=61) associated a favorable prognosis. (D) Increased p-CREB in patients with squamous cell carcinoma (n=57) associated a favorable prognosis. *p* values were determined by log-rank test.

Because many clinical variables, such as gender, age, smoking history, histology subtypes, stage, tumor size, and lymph node metastasis may affect patients' survival, multivariate analyses using a Cox regression model were performed for CDK4 and HSP90. After adjusting for the above clinical variables, the hazard ratios of CDK4 or HSP90 expression remains statistically significant (p=0.028 and p=0.031, respectively) (**Table 3**), suggesting that CDK4 and HSP90 were independent prognostic factors. No overall correlation between other proteins (CDC2p34, XIAP, p-CREB(Ser133) and CREB) with the risk of death was observed (data not shown). However, high expression of p-CREB(Ser133) was associated with a favorable prognosis in those with smoking history (p=0.041, **Figure 3C**) and high expression of CREB was associated with a favorable prognosis in those with squamous cell carcinoma (p=0.03, **Figure 3D**).

In order to improve the prognostic capability, a patient's risk score was calculated as the product of the H-score of each protein and its corresponding coefficient (**Figure 4A**). The coefficient of each protein (CDK4=-0.717, HSP90=-0.681, p-CREB=-0.356 and CREB=-0.318) was

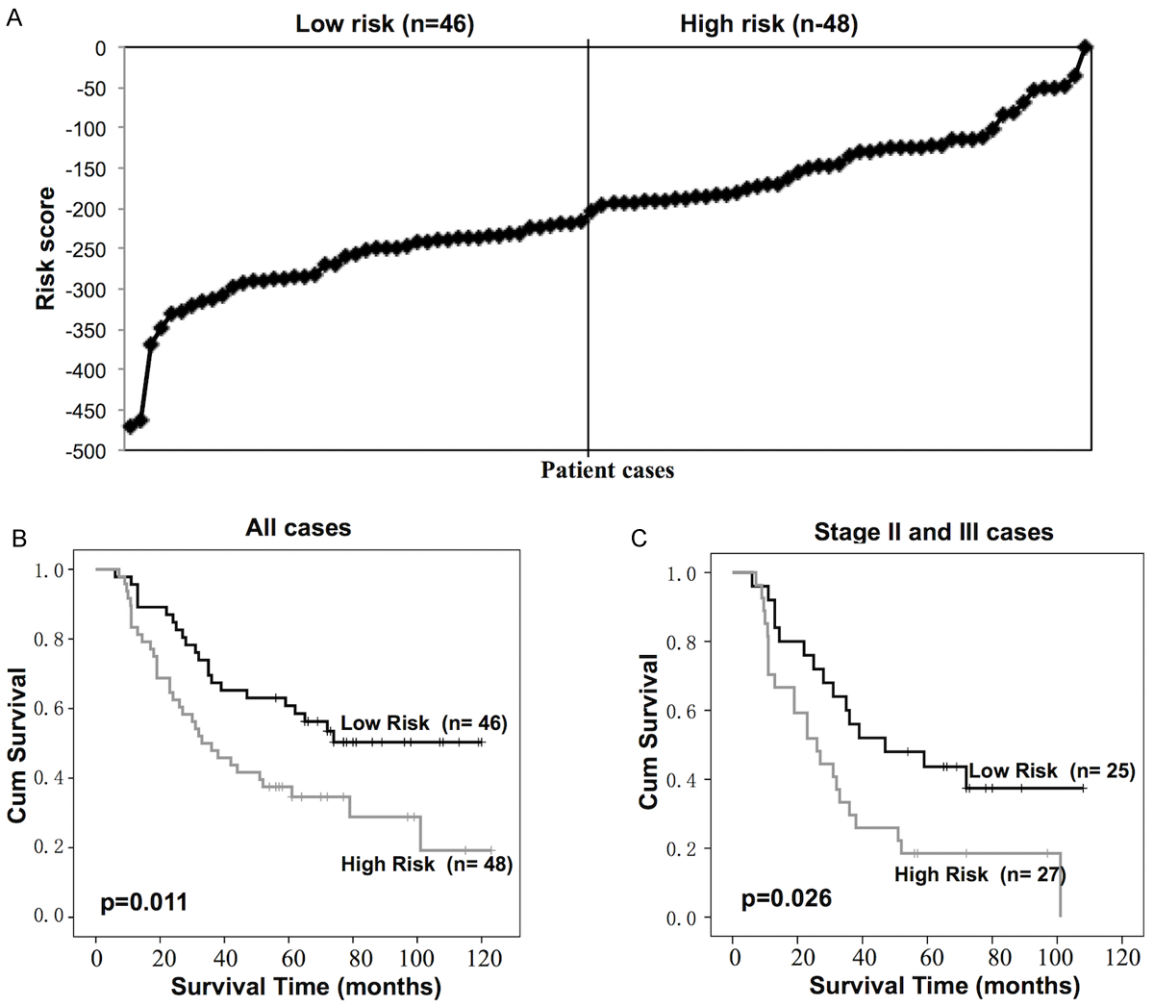

Figure 4. Kaplan-Meier survival analysis of patients with NSCLC based on the risk scores. A. The cases (n=94) were ranked according to the risk scores calculated based on the expression level of CDK4, HSP90, p-CREB and CREB in each sample. The line divided the cases into low and high risk groups. B. The overall survival of all NSCLC patients (n=94) was different in high and low risk groups as shown in A. C. The overall survival of the patients with stage II-III NSCLC (n=52) was different based on their risk scores. *p* values were determined by log-rank test.

determined based on the univariate Cox regression analysis. These 4 proteins were chosen because they were significantly correlated with patients' survival either in all or in a subset of NSCLC patients. To avoid the effect of extreme values, the 50th percentile (median) was chosen to separate the patients into two groups: high-risk score group (-202.47 to 0) and low risk score group (-471.63 to -202.48) (**Figure 4A**). The results showed that the patients with high-risk scores had a shorter median overall survival than those with low-risk scores (35 months vs. 66 months, p=0.011 by log-rank test) (**Figure 4B**). In order to eliminate the confounding effect of stages, the survival rate of patients with stage II-III cancer was analyzed using risk scores (**Figure 4C**). Our results showed that the high risk group (risk score -191.69 to -35.04) had an unfavorable prognosis as compared with low risk group (risk score -463.85 to -192.13) (**Figure 4C**). These data suggest that a risk score based on expression level CDK4, HSP90, p-CREB, and CREB can be used to stratify patients with NSCLC.

Discussion

Major efforts have been made to develop molecular signature-based tests to complement the traditional histopathological diagnosis and prognosis in lung cancer and also to understand the biology of lung cancer at a molecular level. Currently, the majority of

Table 3. Hazard ratios of overall survival based on multivariate Cox regression analysis

Clinical Data	*p* value	HR	95.0% CI for HR	
			Lower	Upper
Age	0.304	0.730	0.401	1.329
Gender	0.940	0.971	0.456	2.070
Histology	0.969	1.013	0.532	1.927
Stage	0.022*	1.618	1.071	2.443
T	0.861	1.069	0.510	2.240
N	0.036*	2.216	1.053	4.666
Smoking History	0.127	0.559	0.265	1.180
HSP90	0.031*	0.452	0.219	0.932
CDK4	0.028*	0.477	0.246	0.925

Note: multivariate Cox regression analysis was performed against each of the variables. HR: hazard ratio. 95% CI: 95% confidence interval. *p<0.05.

molecular signatures for lung cancer have been derived from gene expression microarray [6, 26-28]. However, we take a different approach to identification of protein signatures with a focus on regulatory proteins using novel Protein Pathway Array technology [16-19]. Our hypothesis is that the changes at the genetic and epigenetic levels in cancer cells will affect the expression and activation of regulatory proteins and signal transduction pathways, which ultimately control the cancer cell proliferation, invasion, and metastasis. Therefore, the pattern of regulatory protein expression and phosphorylation can serve as protein signatures to predict different phenotypes of lung cancer.

In this study, we assessed the expression of 108 proteins and phosphoproteins in NSCLC and compared with those expressed in adjacent normal tissues using PPA. Among these, 51 (47%) were detected, 21 (19%) were differentially expressed and 12 (11%) could distinguish tumors from normal tissues (see Table S1 and **Figure 1**). These differentially-expressed proteins in NSCLC are important in diverse cellular processes, including cell cycle regulation, cell proliferation, DNA replication and repair, cell death. Specifically, 7 were related to signaling pathways (p-PDK1, p-ERK1/2, ERK1/2, β-catenin, HSP90, p-p38, Notch4), 8 were involved in controlling cell cycle (CDC2p34, p-CDC2p34, Cyclin B1, PCNA, CDK2, CDK4, 14-3-3β and CDC25C), one was a membrane receptor (EGFR), 3 were related apoptosis (XIAP, BCL-6 and p53) and 2 were transcript factors (p-CREB and NFKB p50). Based on Ingenuity Pathway Analysis (IPA) (www.Ingenuity.com) of our data, which utilizes a knowledge-based database, the top 6 canonical pathways dysregulated in NSCLC are ATM signaling (p=3.981E-11), PI3K/AKT signaling (p=6.31E-9), p53 signaling (p=1.122E-7), PTEN signaling (p=2.138E-4), ERK/MAPK signaling (p=1.585E-3), and EGF signaling (p=1.262E-3). These results suggest that the dysregulation of these proteins in NSCLC lead to increased proliferation and survival as well as reduced ability of cell death, a hallmark of malignancy.

The dysregulation of some of these proteins, including EGFR signaling pathway (EGFR and ERK1/2) [29, 30], proliferation marker PCNA [31], β-catenin [32], tumor surpressor p53 [33], cell cycle protein Cyclin B1 [34], CDK4 [35] and cyclin dependent kinases CDC2p34 [36], anti-apoptosis protein XIAP [37], and transcription factor phosphorylated-cyclic AMP response element-binding protein (p-CREB) [38] was reported previously in NSCLC. However, the dysregulated expression of heat shock protein 90 (HSP90) and protein kinase p-PDK1 in NSCLC was discovered for the first time by our PPA study in human NSCLC tissues. The differential expression of CDC2p34, HSP90, p-CREB, XIAP, and CDK4 in NSCLC was further confirmed by TMA assay and the percentage of these proteins expressed in NSCLC was very similar between PPA and TMA (An additional table file shows this in more detail [see Figure S2]). Interestingly, we observed that dysregulation of any given protein does not occur in all NSCLC. For example, about 89.7% of NSCLC exhibited an increased expression of CDC2p34 and 58.6% of NSCLC had a decreased expression of p-CREB (An additional table file shows this in more detail [see Figure S2]). These results suggest that a great heterogeneity exists in NSCLC and no single pathway can account for the cause of NSCLC.

It is of clinical significance that some of the proteins identified in our study are relatively specific for malignant tumors (An additional table file shows this in more detail [see Table S2]), suggesting that the mechanisms of development of malignant and benign tumors are different. Our TMA results showed that 4 of 6 proteins including CDC2p34, HSP90, XIAP and CDK4 were over-expressed predominantly in

NSCLC. These proteins may be used to differentiate malignant lung cancers from benign processes, which sometimes can be very challenging to distinguish at a histological level. These proteins are also associated with several malignant behaviors including invasion (over-expression of CREB in early stage tumors) and lymph node metastasis (over-expression of CDC2p34 in lymph node positive tumor cells), as well as squamous differentiation (over-expression of CDC2p34, CREB, XIAP and CDK4 in squamous carcinoma cell) (**Table 2**). CREB is a member of a leucine zipper class of transcription factors that binds to cAMP-response elements (CREs) found within the promoter and enhancer regions of hundreds of genes. CREB regulates the expression of genes that suppress apoptosis, induce cell proliferation, and mediate inflammation and tumor metastasis. The roles of the protein in tumor development and progression have been previously reported [39-41]. For example, the expression of CREB in breast cells increases during their development and the CREB upregulation plays an important role in multiple steps of bone metastasis [41]. Furthermore, CREB over-expression in bone marrow was associated with an unfavorable prognosis in AML patients compared to those without over-expression [40]. More interestingly, the expression levels of CREB and p-CREB were significantly higher in the lung squamous cell carcinoma than in the adenocarcinoma [38], which is consistent with our finding (**Table 2**). CDC2p34 (cell division control protein 2 homolog), also known as CDK1 (cyclin dependent kinase 1), is a cyclin-dependent serine/threonine kinase that phosphorylates many proteins. CDC2p34 is a key player in regulating cell cycle progression through the G_1-S and G_2-M checkpoints. Increased expression of CDC2p34 has been associated with leiomyosarcoma of the uterus [42], neoplastic squamous epithelium of the cervix [43], and hepatocellular carcinoma [44]. Recent studies indicate that over-expression of CDC2p34 is associated with lymph node metastasis of breast [45], colon [46] and endometrial cancers [47]. It is interesting to note that both CREB and CDK4 are associated with smokers, suggesting the activation of these pathways by carcinogens in cigarettes (**Table 2**). Furthermore, unfavorable prognosis was observed in smokers with lower level expression of p-CREB and in patients with lower level expression of CREB in squamous carcinoma (**Figure 3C** and **3D**).

It is well known that the prognosis varies widely in NSCLC even among patients with similar clinical and pathological features, suggesting that current staging systems for lung cancer that are based on clinical and pathological findings may have reached their limit of usefulness in predicting outcomes. Therefore, addition of molecular biomarkers may improve the accuracy of predicting patient survival. In this study, we demonstrated that the expression level of HSP90 and CDK4 are independent markers, together with TNM stage and lymph node status, to predict the overall survival of NSCLC patients (**Table 3**). HSP90 (heat shock protein 90) is a molecular chaperone and is significantly increased in cells after stress, such as heat. It is involved in cell signaling and tumor growth by stabilizing some important signaling proteins including EGFR, VEGF, AKT, and mutant form of p53 [48]. HSP90 is overexpressed in many tumor types, including breast cancer, gastrointestinal stromal tumor (GIST), and esophageal carcinoma, making it an attractive therapeutic target [48]. CDK4 is a cyclin-dependent serine/threonine-specific protein kinase and is critical in a cell cycle G1 phase progression via phosphorylation of retinoblastoma gene product (Rb). Previous reports demonstrated that the expression of HSP90 and CDK4 is associated with prognosis of many cancers. Increased HSP90 expression indicated a worse prognosis for patients with breast cancer [49], whereas loss of HSP90 expression predicted a poor clinical outcome in patients with bladder carcinoma [50]. Moreover, in GIST HSP90 over-expression significantly correlated with poor clinical outcome [51]. Studies also demonstrated that CDK4 over-expression was associated with worse survival in patients with glioblastoma multiforme [52] and colorectal carcinoma [53] and lung cancer [35]. However, to our knowledge, this is the first study demonstrating that HSP90 is dysregulated in NSCLC tissues and correlate significantly with patient survival. It is interesting to note that over-expression of HSP90 and CDK4 predicts different clinical outcomes (i.e. favorable or unfavorable), depending on the types and histological origin of the cancers. These data suggest that HSP90 and CDK4 may play different roles in different cancers. In order to predict the risk of each patient, we further developed prognostic risk scores based on the combined expression level of CDK4, HSP90, p-CREB, and CREB. The lower

risk scores predict favorable prognosis for all stages of the patients. It is our hope that the risk score can be used for more accurate patient stratification for surgical and chemotherapeutic treatment in combination with TNM stage.

Conclusion

In conclusion, our data showed that there is a broad dysregulation of regulatory proteins in NSCLC, suggesting the important roles of these proteins in carcinogenesis. The altered expression of some of the proteins correlated with tumor stage and lymph node metastasis, while other proteins correlated with overall survival, indicating that different sets of regulatory proteins associate with different tumor behaviors and clinical outcomes. Future study will be focused on understanding the roles of these proteins in controlling tumor behaviors and confirming the ability of these proteins to classify lung cancer and predict survival in a different cohort of patients.

Acknowledgements

The study was partially supported by funding from The First Hospital of Jilin University, Changchun, China; and Science and Technology Administration of Jilin Province, Changchun, China (#20071133 and #20110734). The authors wish to thank Drs. Yuan Hu and Wei Li for their excellent technical support in statistical analysis and Dr. Peng Lee for valuable comments on this manuscript.

Disclosure of conflict of interest

There is no potential competing interest.

Address correspondence to: Dr. Fei Ye or Dr. David Y Zhang, Department of Pathology, Mount Sinai School of Medicine, New York, NY 10029, USA. Tel: 212-659-8137; Fax: 212-659-8230; E-mail: fei.ye@mssm.edu (FY); Tel: 212-659-8173; Fax: 212-659-8230; E-mail: david.zhang@mssm.edu (DYZ)

References

[1] Siegel R, Naishadham D, Jemal A. Cancer statistics, 2013. CA Cancer J Clin 2013; 63: 11-30.

[2] Fry WA, Phillips JL, Menck HR. Ten-year survey of lung cancer treatment and survival in hospitals in the United States: a national cancer data base report. Cancer 1999; 86: 1867-1876.

[3] Jemal A, Siegel R, Ward E, Hao Y, Xu J, Thun MJ. Cancer statistics, 2009. CA Cancer J Clin 2009; 59: 225-249.

[4] Hoffman PC, Mauer AM, Vokes EE. Lung cancer. Lancet 2000; 355: 479-485.

[5] Martins SJ, Pereira JR. Clinical factors and prognosis in non-small cell lung cancer. Am J Clin Oncol 1999; 22: 453-457.

[6] Santos ES, Blaya M, Raez LE. Gene expression profiling and non-small-cell lung cancer: where are we now? Clin Lung Cancer 2009; 10: 168-173.

[7] Zhu CQ, Ding K, Strumpf D, Weir BA, Meyerson M, Pennell N, Thomas RK, Naoki K, Ladd-Acosta C, Liu N, Pintilie M, Der S, Seymour L, Jurisica I, Shepherd FA, Tsao MS. Prognostic and predictive gene signature for adjuvant chemotherapy in resected non-small-cell lung cancer. J Clin Oncol 2010; 28: 4417-4424.

[8] Chen HY, Yu SL, Chen CH, Chang GC, Chen CY, Yuan A, Cheng CL, Wang CH, Terng HJ, Kao SF, Chan WK, Li HN, Liu CC, Singh S, Chen WJ, Chen JJ, Subramanian J, Simon R. Gene expression-based prognostic signatures in lung cancer: ready for clinical use? J Natl Cancer Inst 2010; 102: 464-474.

[9] Subramanian J, Simon R. Gene expression-based prognostic signatures in lung cancer: ready for clinical use? J Natl Cancer Inst 2010, 102: 464-474.

[10] Harpole DH Jr, Herndon JE 2nd, Wolfe WG, Iglehart JD, Marks JR. A prognostic model of recurrence and death in stage I non-small cell lung cancer utilizing presentation, histopathology, and oncoprotein expression. Cancer Res 1995; 55: 51-56.

[11] Horio Y, Takahashi T, Kuroishi T, Hibi K, Suyama M, Niimi T, Shimokata K, Yamakawa K, Nakamura Y, Ueda R. Prognostic significance of p53 mutations and 3p deletions in primary resected non-small cell lung cancer. Cancer Res 1993; 53: 1-4.

[12] Olaussen KA, Dunant A, Fouret P, Brambilla E, Andre F, Haddad V, Taranchon E, Filipits M, Pirker R, Popper HH, Stahel R, Sabatier L, Pignon JP, Tursz T, Le Chevalier T, Soria JC; IALT Bio Investigators. DNA repair by ERCC1 in non-small-cell lung cancer and cisplatin-based adjuvant chemotherapy. N Engl J Med 2006; 355: 983-991.

[13] Brundage MD, Davies D, Mackillop WJ. Prognostic factors in non-small cell lung cancer: a decade of progress. Chest 2002; 122: 1037-1057.

[14] Chen G, Gharib TG, Wang H, Huang CC, Kuick R, Thomas DG, Shedden KA, Misek DE, Taylor JM, Giordano TJ, Kardia SL, Iannettoni MD, Yee J, Hogg PJ, Orringer MB, Hanash SM, Beer DG. Protein profiles associated with survival in lung

adenocarcinoma. Proc Natl Acad Sci U S A 2003; 100: 13537-13542.
[15] Beer DG, Kardia SL, Huang CC, Giordano TJ, Levin AM, Misek DE, Lin L, Chen G, Gharib TG, Thomas DG, Lizyness ML, Kuick R, Hayasaka S, Taylor JM, Iannettoni MD, Orringer MB, Hanash S. Gene-expression profiles predict survival of patients with lung adenocarcinoma. Nat Med 2002; 8: 816-824.
[16] Ye F, Che Y, McMillen E, Gorski J, Brodman D, Saw D, Jiang B, Zhang DY. The effect of Scutellaria baicalensis on the signaling network in hepatocellular carcinoma cells. Nutr Cancer 2009; 61: 530-537.
[17] Li W, Ye F, Wang D, Sun X, Tong W, Lian G, Jiang J, Suo J, Zhang DY. Protein predictive signatures for lymph node metastasis of gastric cancer. Int J Cancer 2013; 132: 1851-1859.
[18] Che Y, Ye F, Xu R, Qing H, Wang X, Yin F, Cui M, Burstein D, Jiang B, Zhang DY. Co-expression of XIAP and cyclin D1 complex correlates with a poor prognosis in patients with hepatocellular carcinoma. Am J Pathol 2012; 180: 1798-1807.
[19] Wang D, Ye F, Sun Y, Li W, Liu H, Jiang J, Zhang Y, Liu C, Tong W, Gao L, Sun Y, Zhang W, Seetoe T, Lee P, Suo J, Zhang DY. Protein signatures for classification and prognosis of gastric cancer a signaling pathway-based approach. Am J Pathol 2011; 179: 1657-1666.
[20] Haiman CA, Stram DO, Wilkens LR, Pike MC, Kolonel LN, Henderson BE, Le Marchand L. Ethnic and racial differences in the smoking-related risk of lung cancer. N Engl J Med 2006; 354: 333-342.
[21] Mountain CF. Revisions in the International System for Staging Lung Cancer. Chest 1997; 111: 1710-1717.
[22] Zhu XD, Zhang JB, Zhuang PY, Zhu HG, Zhang W, Xiong YQ, Wu WZ, Wang L, Tang ZY, Sun HC. High expression of macrophage colony-stimulating factor in peritumoral liver tissue is associated with poor survival after curative resection of hepatocellular carcinoma. J Clin Oncol 2008; 26: 2707-2716.
[23] Turner BC, Krajewski S, Krajewska M, Takayama S, Gumbs AA, Carter D, Rebbeck TR, Haffty BG, Reed JC. BAG-1: a novel biomarker predicting long-term survival in early-stage breast cancer. J Clin Oncol 2001; 19: 992-1000.
[24] Rakha EA, Abd El Rehim D, Pinder SE, Lewis SA, Ellis IO. E-cadherin expression in invasive non-lobular carcinoma of the breast and its prognostic significance. Histopathology 2005; 46: 685-693.
[25] Bloom GC, Eschrich S, Zhou JX, Coppola D, Yeatman TJ. Elucidation of a protein signature discriminating six common types of adenocarcinoma. Int J Cancer 2007; 120: 769-775.
[26] Zhu CQ, Strumpf D, Li CY, Li Q, Liu N, Der S, Shepherd FA, Tsao MS, Jurisica I. Prognostic gene expression signature for squamous cell carcinoma of lung. Clin Cancer Res 2010; 16: 5038-5047.
[27] Petrosyan F, Daw H, Haddad A, Spiro T, Sood R. Gene Expression Profiling for Early-Stage NSCLC. Am J Clin Oncol 2013; [Epub ahead of print].
[28] Yao J, Zhao Q, Yuan Y, Zhang L, Liu X, Yung WK, Weinstein JN. Identification of common prognostic gene expression signatures with biological meanings from microarray gene expression datasets. PLoS One 2012; 7: e45894.
[29] Kotsakis A, Georgoulias V. Targeting epidermal growth factor receptor in the treatment of non-small-cell lung cancer. Expert Opin Pharmacother 2010; 11: 2363-2389.
[30] Vicent S, Lopez-Picazo JM, Toledo G, Lozano MD, Torre W, Garcia-Corchon C, Quero C, Soria JC, Martin-Algarra S, Manzano RG, Montuenga LM. ERK1/2 is activated in non-small-cell lung cancer and associated with advanced tumours. Br J Cancer 2004; 90: 1047-1052.
[31] Mattern J, Volm M. Imbalance of cell proliferation and apoptosis during progression of lung carcinomas. Anticancer Res 2004; 24: 4243-4246.
[32] Tseng RC, Lee SH, Hsu HS, Chen BH, Tsai WC, Tzao C, Wang YC. SLIT2 attenuation during lung cancer progression deregulates beta-catenin and E-cadherin and associates with poor prognosis. Cancer Res 2010; 70: 543-551.
[33] Huang CL, Yokomise H, Miyatake A. Clinical significance of the p53 pathway and associated gene therapy in non-small cell lung cancers. Future Oncol 2007; 3: 83-93.
[34] Yoshida T, Tanaka S, Mogi A, Shitara Y, Kuwano H. The clinical significance of Cyclin B1 and Wee1 expression in non-small-cell lung cancer. Ann Oncol 2004; 15: 252-256.
[35] Wu A, Wu B, Guo J, Luo W, Wu D, Yang H, Zhen Y, Yu X, Wang H, Zhou Y, Liu Z, Fang W, Yang Z. Elevated expression of CDK4 in lung cancer. J Transl Med 2011; 9: 38.
[36] Zhang C, Elkahloun AG, Robertson M, Gills JJ, Tsurutani J, Shih JH, Fukuoka J, Hollander MC, Harris CC, Travis WD, Jen J, Dennis PA. Loss of cytoplasmic CDK1 predicts poor survival in human lung cancer and confers chemotherapeutic resistance. PLoS One 2011; 6: e23849.
[37] Krepela E, Dankova P, Moravcikova E, Krepelova A, Prochazka J, Cermak J, Schutzner J, Zatloukal P, Benkova K. Increased expression of inhibitor of apoptosis proteins, survivin and XIAP, in non-small cell lung carcinoma. Int J Oncol 2009; 35: 1449-1462.

[38] Seo HS, Liu DD, Bekele BN, Kim MK, Pisters K, Lippman SM, Wistuba II, Koo JS. Cyclic AMP response element-binding protein overexpression: a feature associated with negative prognosis in never smokers with non-small cell lung cancer. Cancer Res 2008; 68: 6065-6073.

[39] Sofi M, Young MJ, Papamakarios T, Simpson ER, Clyne CD. Role of CRE-binding protein (CREB) in aromatase expression in breast adipose. Breast Cancer Res Treat 2003; 79: 399-407.

[40] Shankar DB, Cheng JC, Kinjo K, Federman N, Moore TB, Gill A, Rao NP, Landaw EM, Sakamoto KM. The role of CREB as a proto-oncogene in hematopoiesis and in acute myeloid leukemia. Cancer Cell 2005; 7: 351-362.

[41] Son J, Lee JH, Kim HN, Ha H, Lee ZH. cAMP-response-element-binding protein positively regulates breast cancer metastasis and subsequent bone destruction. Biochem Biophys Res Commun 2010; 398: 309-314.

[42] Zhai YL, Nikaido T, Shiozawa T, Orii A, Fujii S. Expression of cyclins and cyclin-dependent kinases in smooth muscle tumors of the uterus. Int J Cancer 1999; 84: 244-250.

[43] Kanai M, Shiozawa T, Xin L, Nikaido T, Fujii S. Immunohistochemical detection of sex steroid receptors, cyclins, and cyclin-dependent kinases in the normal and neoplastic squamous epithelia of the uterine cervix. Cancer 1998; 82: 1709-1719.

[44] Ito Y, Takeda T, Sakon M, Monden M, Tsujimoto M, Matsuura N. Expression and prognostic role of cyclin-dependent kinase 1 (cdc2) in hepatocellular carcinoma. Oncology 2000; 59: 68-74.

[45] Ohta T, Fukuda M, Arima K, Kawamoto H, Hashizume R, Arimura T, Yamaguchi S. Analysis of Cdc2 and Cyclin D1 Expression in Breast Cancer by Immunoblotting. Breast Cancer 1997; 4: 17-24.

[46] Nozoe T, Honda M, Inutsuka S, Korenaga D. P34cdc2 Expression is an Independent Indicator for Lymph Node Metastasis in Colorectal Carcinoma. J Cancer Res Clin Oncol 2003; 129: 498-502.

[47] Bidus MA, Risinger JI, Chandramouli GV, Dainty LA, Litzi TJ, Berchuck A, Barrett JC, Maxwell GL. Prediction of lymph node metastasis in patients with endometrioid endometrial cancer using expression microarray. Clin Cancer Res 2006; 12: 83-88.

[48] Ekman S, Bergqvist M, Tell R, Bergstrom S, Lennartsson J. Hsp90 as a therapeutic target in patients with oesophageal carcinoma. Expert Opin Ther Targets 2010; 14: 317-328.

[49] Pick E, Kluger Y, Giltnane JM, Moeder C, Camp RL, Rimm DL, Kluger HM. High HSP90 expression is associated with decreased survival in breast cancer. Cancer Res 2007; 67: 2932-2937.

[50] Lebret T, Watson RW, Molinie V, O'Neill A, Gabriel C, Fitzpatrick JM, Botto H. Heat shock proteins HSP27, HSP60, HSP70, and HSP90: expression in bladder carcinoma. Cancer 2003; 98: 970-977.

[51] Li CF, Huang WW, Wu JM, Yu SC, Hu TH, Uen YH, Tian YF, Lin CN, Lu D, Fang FM, Huang HY. Heat shock protein 90 overexpression independently predicts inferior disease-free survival with differential expression of the alpha and beta isoforms in gastrointestinal stromal tumors. Clin Cancer Res 2008; 14: 7822-7831.

[52] Ruano Y, Ribalta T, de Lope AR, Campos-Martin Y, Fiano C, Perez-Magan E, Hernandez-Moneo JL, Mollejo M, Melendez B. Worse outcome in primary glioblastoma multiforme with concurrent epidermal growth factor receptor and p53 alteration. Am J Clin Pathol 2009; 131: 257-263.

[53] Zhao P, Hu YC, Talbot IC. Expressing patterns of p16 and CDK4 correlated to prognosis in colorectal carcinoma. World J Gastroenterol 2003; 9: 2202-2206.

Table S1. List of antibodies included in the Protein Pathway Array (Underlines indicate a detectable expression in NSCLC tissues)

Angiogenesis: VEGFR, TGF-β, TNF-α, HIF-1α, HIF-2α, HIF-3α
Apoptosis/Autophagy: Bax, BCL-2, BCL-6, cleaved caspase 3, XIAP, NFkB65, NFkB52, NFkB50, Survivin, Cytochrome C
Signal transduction: ERK1/2, p-ERK1/2(Thr202/Tyr204), Akt, p-Akt(Ser473), p-HGFR(Y1234/Y1235), p-p38(Thr180/Tyr182), p-JNK(Thr183/Tyr185), FGFR, p-FGFR (Tyr653/654), p-VEGFR(Tyr951), cPKCα, p-PKCα(Ser657), p-PKCα/β(Thr638/641), p-PKCδ (Thr505), p-PTEN(Ser380), EGFR, p-EGFR(Tyr1068), p-EGFR(Tyr1148), Her2, p-Her2 (Tyr1221/1222), p-PDK1(Ser241), p-mTor(Ser2448), p-IKB(Ser32), p-c-Kit(Tyr719), β-catenin, p-β-catenin(Ser33/37/Thr41), p-Stat3(Ser727), p-Stat5(Tyr694), p-Smad(Ser463/465), p-p70 S6 Kinase(Thr389), p-eIF4B(Ser422), p-ERK5(Thr218/Tyr220), p-p90RSK(Ser380), p-FLT3(Tyr 591), eIF4B, p-Survivin(Thr34), p-HGFR (Y1003), p-IGFR(Tyr1131/Tyr1146), p-c-Jun(Ser73), Notch4, Notch1
Oncoproteins/Suppressors: p21, p27, p16, p14, PTEN, p53, p-p53(Ser392), MDM2, c-Kit, WT1, SRC-1
Cell cycle: CDK2, CDK4, CDK6, CDC2p34, CDC25B, CDC25C, Cyclin B1, Cyclin D1, Cyclin E, Cdc42, Chk1, BRCA1, 14-3-3β, PCNA, Ki-67, Wee 1, p-RB(Ser780), p-RB(Ser807/811), p-CDC2(Tyr15)
Invasion/metastasis: COX-2, Osteopontin, HSP90, N-cadherin, E-cadherin
Transcription factor: c-Jun, ETS1, c-MYC, E2F-1, TERT, Gata1, p-CREB(Ser133), Trap, p300
Others: Vimentin, Mesothelin, Calretinin

Note: p-represents phosphorylation.

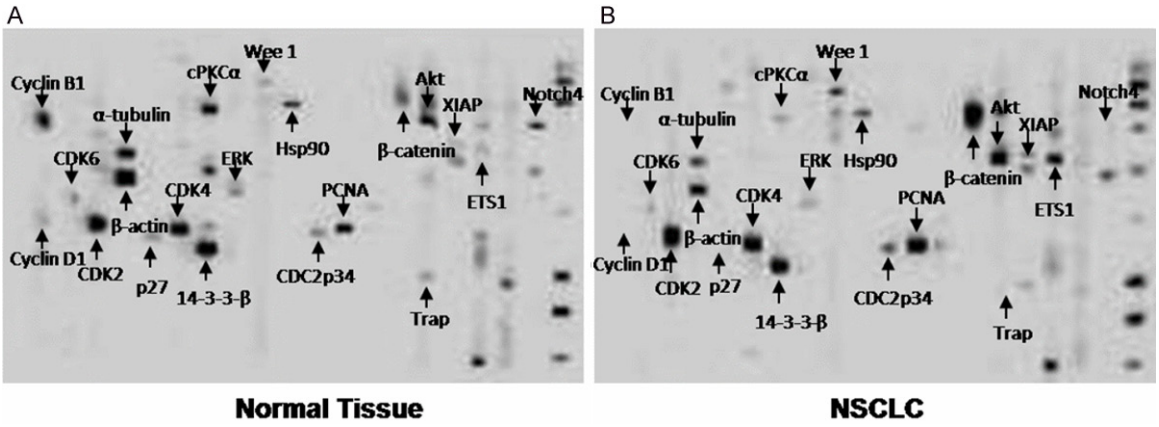

Figure S1. Representative images of Protein Pathway Array showing the expression and phosphorylation of different proteins. (A) Normal tissue and (B) NSCLC tissue.

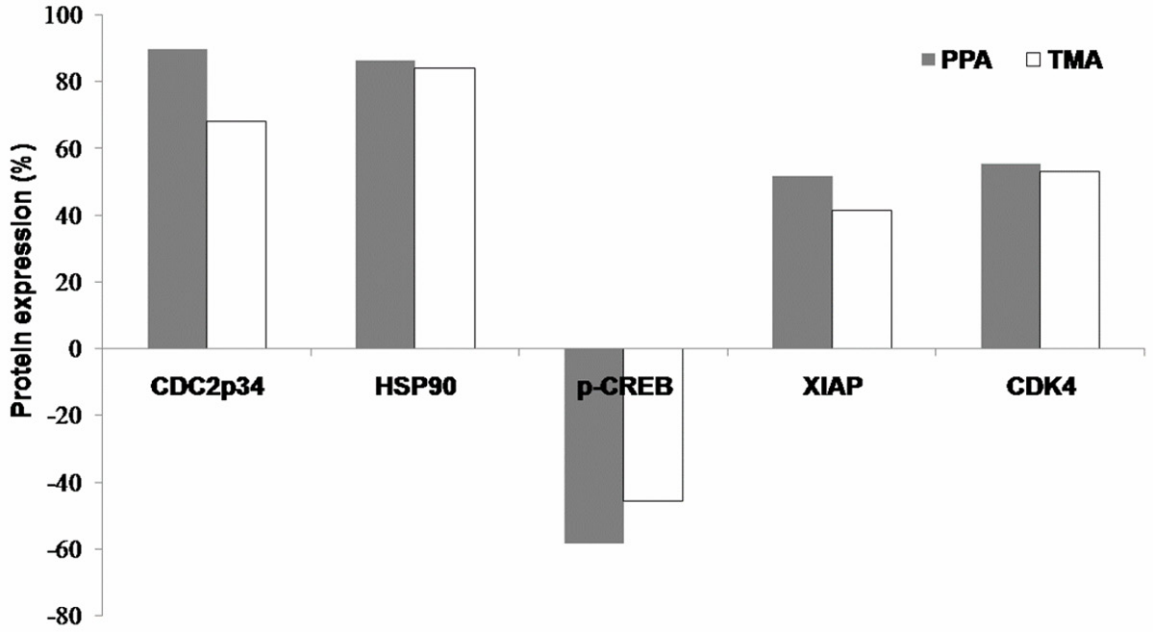

Figure S2. Comparison of percentages of protein expression in NSCLC as determined by PPA and TMA. The percentages of CDC2p34, HSP90, p-CREB(Ser133), XIAP and CDK4 in NSCLCs were determined and overall, the expression percentages of these proteins between 2 platforms (PPA and TMA) were very similar.

Table S2. The comparison of protein expression between NSCLC and benign tumors

Tumor Types	CDC2 p34			HSP90			p-CREB		
	+	-	*p* value	+	-	*p* value	+	-	*p* value
NSCLC	64	30	p=0.000**	79	15	p=0.000**	51	43	p=0.166
Benign Tumor	1	18		3	16		7	12	
Tumor Types	CREB			XIAP			CDK4		
	+	-	*p* value	+	-	*p* value	+	-	*p* value
NSCLC	46	48	p=0.026*	39	55	p=0.008**	50	44	p=0.003**
Benign Tumor	4	15		2	17		3	16	

Note: Fisher's exact test was performed to determine the significant difference of each protein between NSCLC and benign tumors (*p<0.05, **p<0.01).