Proteomic Analysis of Cytokeratin Isoforms Uncovers Association with Survival in Lung Adenocarcinoma[1]

Tarek G. Gharib, Guoan Chen*, Hong Wang†, Chiang-Ching Huang‡, Michael S. Prescott*, Kerby Shedden§, David E. Misek†, Dafydd G. Thomas¶, Thomas J. Giordano¶, Jeremy M.G. Taylor‡, Sharon Kardia‡, John Yee*, Mark B. Orringer*, Samir Hanash† and David G. Beer**

Departments of *Surgery; †Pediatrics; ‡Biostatistics; §Statistics; ¶Pathology, University of Michigan, Ann Arbor, MI 48109, USA

Abstract

Cytokeratins (CK) are intermediate filaments whose expression is often altered in epithelial cancer. Systematic identification of lung adenocarcinoma proteins using two-dimensional polyacrylamide gel electrophoresis and mass spectrometry has uncovered numerous CK isoforms. In this study, 93 lung adenocarcinomas (64 stage I and 29 stage III) and 10 uninvolved lung samples were quantitatively examined for protein expression. Fourteen of 21 isoforms of CK 7, 8, 18, and 19 occurred at significantly higher levels ($P<.05$) in tumors compared to uninvolved adjacent tissue. Specific isoforms of the four types of CK identified correlated with either clinical outcome or individual clinical–pathological parameters. All five of the CK7 isoforms associated with patient survival represented cleavage products. Two of five CK7 isoforms (nos. 2165 and 2091), one of eight CK8 isoforms (no. 439), and one of three CK19 isoforms (no. 1955) were associated with survival and significantly correlated to their mRNA levels, suggesting that transcription underlies overexpression of these CK isoforms. Our data indicate substantial heterogeneity among CK in lung adenocarcinomas resulting from posttranslational modifications, some of which correlated with patient survival and other clinical parameters. Therefore, specific isoforms of individual CK may have utility as diagnostic or predictive markers in lung adenocarcinomas.

Neoplasia (2002) **4,** 440–448 doi:10.1038/sj.neo.7900257

Keywords: lung adenocarcinoma, isoforms, 2D PAGE, mass spectrometry, microarrays.

Introduction

Lung cancer is the leading cause of cancer mortality among men and women in the US, accounting for 158,900 deaths in 1999 [1]. An increase in the percentage of adenocarcinomas has been recently observed, making adenocarcinomas the most common histologic type of non small cell lung cancer. In contrast to patients with advanced stages of lung adenocarcinomas, those presenting with stage I disease have overall 5-year survival rates of 55% to 63% [2]. Biomarkers that serve to correctly stratify patients according to risk or to serve as an early detection system may impact survival or help direct therapy. To correctly identify these high-risk patients, identification of biomarkers determined using genomic or proteomic technologies may have great utility.

The cytokeratins (CK) belong to a large family of intermediate filaments that are primarily expressed in epithelial cells and whose members are expressed in various combinations in both normal and neoplastic epithelium (reviewed in Moll et al. [3]). They are abundant proteins involved in the structural organization of cells and the specific forms CK7, CK8, CK18, and CK19 have been reported in lung adenocarcinomas [4–7] and in other cancer types [8,9]. These proteins are often highly expressed by lung tumors and also specifically modified with proteolytically altered forms detected in the serum of lung cancer patients [7,8]. Little is understood, however, regarding the expression of the individual isoforms of these proteins or their relationship to specific clinical or pathological features of these tumors.

In the present study, we utilized two-dimensional polyacrylamide gel electrophoresis (2D PAGE) and mass spectrometry (MS) to examine CK expression profiles of 93 lung adenocarcinomas and 10 uninvolved lung samples. We have examined within this cohort of lung adenocarcinomas both early stage I and more aggressive stage III tumors to help identify specific CK isoforms that are associated with poor clinical outcome. Correlation between protein levels of the CK isoforms and the CK gene mRNA expression was examined to help delineate potential mechanisms underlying overexpression of these proteins in lung tumors.

Materials and Methods

Sample Acquisition and Preparation

Sequential patients seen between May 1991 and July 2000 by the General Thoracic Surgery at the University of

Abbreviations: CK, cytokeratins; MS, mass spectrometry; 2D PAGE, two-dimensional polyacrylamide gel electrophoresis

Address all correspondence to: David G. Beer, General Thoracic Surgery, MSRB II, B560, Box 0686, University of Michigan Medical School, Ann Arbor, MI 48109-086, USA. E-mail: dgbeer@umich.edu

[1]This work was supported by NCI grant U19 CA-85953.
Received 8 April 2002; Accepted 14 May 2002.

Michigan Hospital for resection of stages I and III lung adenocarcinoma were evaluated for inclusion in this study. Consent was received from all patients and the project approved by the Institutional Review Board. Patient medical records were reviewed and patient identifiers coded to protect confidentiality. Tumor tissues and adjacent non-neoplastic lung tissues were acquired immediately after resection and carried to the laboratory in Dulbecco's modified Eagle's medium (Life Technologies, Gaithersburg, PA) on ice. This included 64 stage I, 29 stage III lung adenocarcinomas, as well as 10 non-neoplastic lung samples (Table 1). A portion of each tumor and/or lung tissue was embedded in OCT (Miles Scientific, Naperville, IL), frozen in isopentane cooled with liquid nitrogen for cryostat sectioning, and then stored at -80°C. Hematoxylin- and eosin-stained cryostat sections (5 μm), prepared from tumor pieces to be utilized for mRNA isolation, were evaluated by a study pathologist, and compared to H&E sections made from paraffin blocks of the same tumor. Specimens were excluded based on unclear or mixed histology (e.g., adenosquamous), tumor cellularity less than 70%, potential metastatic origin as indicated by previous tumor history, extensive lymphocytic infiltration, extensive fibrosis, prior chemotherapy, or radiotherapy. Tumors were histopathologically divided into two categories: bronchial-derived, if they exhibited invasive features with architectural destruction, and bronchioloalveolar, if they exhibited preservation of the lung architecture.

The WHO classification of lung adenocarcinomas includes acinar, papillary, bronchioloalveolar, and solid, indicating the large heterogeneity recognized within these tumors and there is also an overlap that exists between WHO-recognized classes [10]. Because the bronchioloalveolar carcinomas have distinctive clinicopathologic features, it is justifiable to separate them. We separated the tumors into two broad histopathological classifications, bronchial-derived and bronchioloalveolar. H&E–stained paraffin sections were utilized to examine lymphocytic response of tumors. Those demonstrating lymphocytic accumulation either within or surrounding the tumors were classified as having a positive lymphocytic response. Tumors showing only mild or no lymphocytes were deemed negative and those with moderate or extensive lymphocytes were deemed positive.

2D PAGE

Protein samples were solubilized in standard lysis buffer (9.5 M urea, 20 μl of nonionic detergent [Nonidet P-40], 20 μl of ampholines [pH 3.5–10; Pharmacia/LKB, Piscataway, NJ], 20 μl of 2-mercaptoethanol, and 0.89 M phenylmethylsulfonyl fluoride per milliliter of deionized water). Sample volumes of between 15 and 30 μl were immediately applied to isoelectric focusing gels containing 50 μl of ampholytes per milliliter (pH 3.5–10). First dimensional separation was performed using 700 V for 16 hours, and then 1000 V for 2 hours at room temperature. Second dimensional separation was accomplished using an 18$\times$ 18 cm gel containing an acrylamide gradient of 11.4 to 14 g/dl. Samples were run in 20 gel batches. After separation, the protein spots were visualized by a photochemical silver-based staining technique.

Detection and Quantification

Following 2D separation, each gel was scanned using a Kodak CCD camera. A 1024$\times$1024 pixel format was used,

Table 1. Clinical–Pathological Characteristics of 93 Lung Adenocarcinoma Patients.

Variable	*n*
Age	
<65	49
>65	44
Gender	
Female	53
Male	40
Race	
White	76
Nonwhite (Asian, Black)	8
Unknown	9
Smoking	
Nonsmoking	10
Smoking	79
Unknown	4
Stage	
I	64
III	29
T status	
T_1	49
T_{2-4}	44
N status	
N_0	68
N_{1-2}	25
Classification	
Bronchoalveolar	14
Bronchial-derived	76
Mixed	3
Differentiation	
Well	22
Moderate	47
Poor	23
Lymphocytic response	
Yes	41
No	52
Angiolymphatic invasion	
Yes	16
No	77
Tumor location	
Left lobe	31
Right lobe	62
p53 nuclear accumulation	
Positive	19
Negative	69
K-ras mutational status	
Positive	34
Negative	40

yielding pixel widths of 163 μm where each pixel had 256 possible gray scale values (optical density). Spot detection was accomplished by Bio Image Visage System software (Bioimage, Ann Arbor, MI). Each gel generated 1600 to 2200 detectable spots, with 820 well-defined spots in most samples, chosen to obtain quantitative measurements. The integrated intensity, which is the value of each spot, was calculated as the measured optical density units$\times$square millimeters. Then spots from each gel were matched to the 820 spots on a "master" gel [11] to allow for the identification of identical polypeptides between each gel. A total of 250 spots were chosen as ubiquitously expressed reference spots to allow adjustment for variations in protein loading and gel staining. Each of the 820 spots was then mathematically adjusted in relation to the reference spots [12].

MS

Protein spots identified for analysis were extracted from preparative 2D gels of extracts from A549 lung adenocarcinoma cell lysates (obtained from ATCC) or lung adenocarcinoma tissue lysates (obtained from patient). The conditions were identical to the analytical 2D gel, except there was a 30% greater protein loading, and were silver-stained by successive incubations in 0.02% sodium thiosulfate for 2 minutes, 0.1% silver nitrate for 40 minutes, and then 0.014% formaldehyde plus 2% sodium carbonate for 10 minutes. Identification of proteins was performed by trypsin digestion followed by either matrix-assisted laser desorption ionization time-of-flight mass spectrometry (MALDI-TOF MS) in the Perseptive Voyager Biospectrometry Workstation (PerSeptive Biosystem, Framingham, MA), or nanoflow capillary liquid chromatography coupled with electrospray tandem mass spectrometry (ESI MS/MS) in a Q-TOF *micro* (Micromass, Manchester, UK). MALDI-TOF MS gave a "fingerprint" for each spot based on the molecular mass of trypsin-digested products and compared the resulted mass with known trypsin digest databases using the MSFit database searching (http://prospector.ucsf.edu/ucsfhtml3.2/msfit.htm). MS/MS spectra produced by ESI MS/MS were automatically processed and searched against a nonredundant database using ProteinLynx Global SERVER (www. micromass.co.uk).

RNA Isolation for Microarrays

RNA was isolated from 75 of the tumors and the 10 normal samples used in the protein studies. Two contiguous 2-mm^3 samples were removed for RNA and protein isolation, respectively. Total cellular RNA was isolated using Trisol reagent (Life Technologies) and subjected to further purification using RNeasy columns (Qiagen, Valencia, CA). Five micrograms of total RNA was used as template. All protocols used for mRNA reverse transcriptase, second strand synthesis, production of cDNA and RNA amplification, hybridization, and washing conditions for the 6800 gene HUGeneFL oligonucleotide arrays are as provided by the manufacturer (Affymetrix, Santa Clara, CA). More detailed information is provided in Giordano et al. [13].

Mutational Analysis of K-ras

Genomic DNA was isolated from each tumor sample and 50 ng was subjected to PCR amplification using the primers that encompass codons 12 and 13 of the K-*ras* gene. The sequences of forward and reverse primers are 5′ TATAAGGCCTGCTGAAAAT 3′ and 5′ CCTGCACCAGTAATATGC 3′, respectively. Two nanograms of purified PCR products containing the exon 1 of the K-*ras* gene was then subjected to thermal cycle sequencing with an internal nested primer (5′ AGGCCTGCTGAAAATGACT 3′) and resolved in 8% urea PAGE gels, dried, and exposed to Phosphor-Image screens and visualized using a Phosphor-Image scanner (Molecular Dynamics, Sunnyvale, CA). The mutations were determined by comparing each tumor DNA sequences of K-*ras* 12th and 13th codons to its wild-type sequence GGTGGC.

Immunohistochemical Staining

Diagnoses for all primary lung adenocarcinomas used in this study were confirmed by a board-certified pathologist. A tissue microarray block was assembled based on the best morphological areas of the tumors used in this study according to the method of Kononen et al. [14]. Deparaffinized sections of the pulmonary adenocarcinoma tissue microarray were microwave-pretreated in citric acid to retrieve antigenicity. The sections were incubated with 1% hydrogen peroxide for 60 minutes to inhibit endogenous peroxidase activity at room temperature. Following blocking to reduce nonspecific binding, the sections were incubated with primary antibodies overnight at 4°C. Antibodies included anti-p53 (M-7001; 1:500), anti-CK19 (M0772; 1:500), and anti-CK7 (OV-TL 12/30; 1:500) from DAKO (Carpenteria, CA); anti-CK7 (K72.7; 1:500; Neomarkers, Fremont, CA); and anti-CK18 (DC-10; 1:500) and anti-CK8 (TS1; 1:500) from Novo Castra Laboratories (Newcastle, UK). The immuno-complex was visualized by the immunoglobulin enzyme bridge technique using a Vector ABC peroxidase kit (Vector Laboratories, Burlingame, CA). The enzyme substrate was 3,3′ diaminobenzidine tetrachloride, resulting in a brown reactant. The sections were lightly counterstained with hematoxylin.

2D Western Blotting

Protein extracts of A549 lung adenocarcinoma cells were run on 2D gels using the identical conditions as used for the analytical 2D gels. The separated proteins were transferred onto polyvinylidene fluoride membranes and incubated for 2 hours at room temperature with a blocking buffer consisting of TBST (Tris-buffered saline, 0.01% Tween 20) and 5% nonfat dry milk. Individual membranes were washed and incubated with monoclonal antibodies listed above for immunohistochemistry and used at 1:500 dilution. After additional washes with TBST, the membranes were incubated with a secondary antibody conjugated with horseradish peroxidase (HRP) at a 1:5000 dilution for 1 hour, further washed, and then incubated for 1 minute with enhanced chemiluminescence (ECL) (Pierce, Rockford, IL).

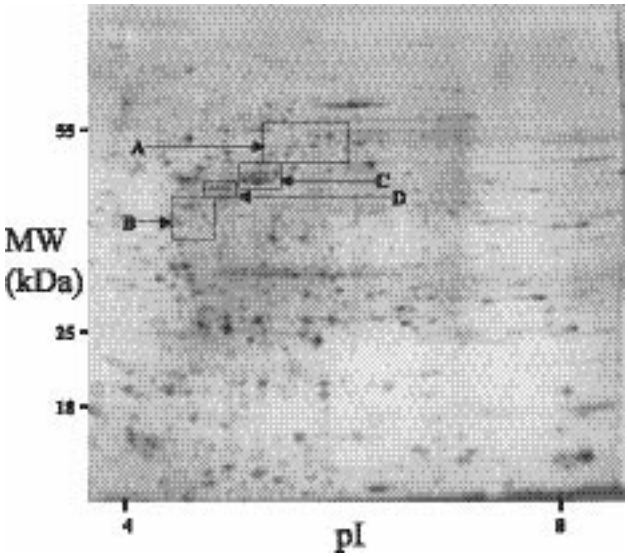

Figure 1. *2D PAGE gel image of a stage I lung adenocarcinoma with proteins separated by molecular mass and isoelectric point. (A) Location of CK8 spots. (B) Location of CK7 spots. (C) Location of CK18 spots. (D) Location of CK19 spots.*

Statistical Analysis

A Student's *t*-test was used for comparing protein expression in lung adenocarcinoma *versus* uninvolved (normal) lung. An *F*-test was used to compare between values of other clinical–pathological variables by constructing background-adjusted protein spot size quantification. The transform $x \rightarrow \log(1+x)$ was applied to all protein expression measurements. A probability (P) level $<.05$ (two-sided) was considered statistically significant. Potential associations between mRNA and protein expression were made by calculating Spearman correlation coefficients for individual genes and gene products within the same samples.

Survival time was defined as the interval in months between the day of operation for the lung adenocarcinoma and the date of lung cancer–related death or last follow-up. A total of 682 protein spots were included in the survival analysis after the elimination of spots that were absent from at least 60% of all the gels. Univariate Cox proportional hazards regression methods were used to identify proteins associated with variability in survival times.

Results

CK Isoforms Show Altered Expression in Lung Adenocarcinoma

Systematic identification of proteins detected in lung adenocarcinoma tumor and cell lines has uncovered to date over 300 differentially expressed proteins expressed in our lung cancer protein database [15]. Four CK types (CK7, CK8, CK18, and CK19) have been identified (Figure 1), each of which demonstrated multiple protein isoforms in tumor samples (Table 2). All four CKs had at least one or more isoforms that showed significant increases ($P<.05$) in lung adenocarcinomas compared to normal lung. For CK19, two isoforms were overexpressed

Table 2. CK Expression in Lung Adenocarcinoma and Their Frequency of Expression in Tumor and Normal Lung Samples.

Spot Number	Protein Name	Estimated Molecular Mass	Estimated pI	Univariate Cox Model *P* Value	Coefficient β	Normal ($n=10$) Mean ± SD	Tumor ($n=93$) Mean ± SD	*P* (*t*-test)	Fold Change (Tumor/Normal)	Tumor Frequency* (%)	Normal Frequency† (%)
691	CK7	41.7	4.9	.025	0.406	0.049 ± 0.066	0.231 ± 0.182	<.0001	4.714	58.1	0.0
871	CK7	41.7	4.7	.005	−0.659	0.125 ± 0.167	0.159 ± 0.177	.5566	–	–	–
1968	CK7	41.0	4.6	.014	0.516	0.303 ± 0.104	0.493 ± 0.282	.0002	1.627	39.80	10.00
2091	CK7	40.4	4.6	.029	0.356	0.325 ± 0.355	0.246 ± 0.211	.3227	–	–	–
2165	CK7	36.7	4.7	.019	−0.487	0.136 ± 0.049	0.119 ± 0.099	.3845	–	–	–
352	CK8	54.3	5.6	.442	–	0.050 ± 0.023	0.070 ± 0.081	.0768	–	–	–
436	CK8	50.6	5.2	.101	–	0.313 ± 0.168	0.237 ± 0.123	.1939	–	–	–
439	CK8	50.7	5.4	.033	0.382	0.165 ± 0.080	0.343 ± 0.132	<.0001	2.078	49.5	0.0
441	CK8	49.7	5.2	.681	–	0.048 ± 0.062	0.103 ± 0.057	.0219	2.158	10.8	0.0
443	CK8	53.1	5.5	.481	–	0.029 ± 0.060	0.042 ± 0.066	.5254	–	–	–
444	CK8	52.7	5.4	.594	–	0.008 ± 0.026	0.039 ± 0.058	.0062	4.796	23.7	10.0
446	CK8	53.1	5.6	.990	–	0.019 ± 0.030	0.103 ± 0.103	<.0001	5.453	50.5	0.0
451	CK8	54.1	5.7	.814	–	0.013 ± 0.028	0.027 ± 0.052	.1783	–	–	–
523	CK18	45.4	5.1	.393	–	0.126 ± 0.074	0.304 ± 0.191	<.0001	2.413	50.5	0.0
527	CK18	45.3	5.2	.221	–	0.044 ± 0.060	0.103 ± 0.102	.0186	2.331	15.1	0.0
529	CK18	46.0	5.4	.108	–	0.121 ± 0.097	0.388 ± 0.191	<.0001	3.219	60.2	0.0
2324	CK18	46.0	5.2	.642	–	0.000 ± 0.000	0.117 ± 0.168	<.0001	N/A	36.6	0.0
2381	CK18	46.0	5.3	.084	–	0.000 ± 0.000	0.124 ± 0.111	<.0001	N/A	49.5	0.0
609	CK19	43.6	4.7	.724	–	0.057 ± 0.081	0.132 ± 0.133	.0215	2.305	26.9	0.0
1955	CK19	43.6	4.6	.026	0.433	0.010 ± 0.032	0.163 ± 0.142	<.0001	16.171	67.7	10.0
608	CK19	43.7	4.8	.567	–	0.987 ± 0.512	0.512 ± 0.348	.0170	0.519	50.5	20.0

Normal average + 2 SD was used as cut-off value for proteins that were overexpressed in tumor samples. Normal average − 2 SD was used as cut-off value for proteins that were underexpressed in tumor samples.

(–) Placed where value is insignificant due to univariate Cox model *P* value >.05 or *t*-test *P* value >.05.

N/A represents a fold change that cannot be expressed due to normal mean value being zero.

*Frequency of expression in tumor samples is percent of samples greater than cut-off value.

†Frequency of expression in normal samples is percent of samples greater than cut-off value.

Table 3. Relationship Between CKs and Different Clinical–Pathological Variables.

Spot Number	Protein Name	Expression in Stage III Samples (Versus Stage I) (n=29 vs 64)	Expression Bronchial-Derived Samples (Versus Bronchioloalveolar) (n=73 vs 14)	Expression in Poorly Differentiated Samples (Versus Well, Moderate) (n=23 vs 22,47)	Expression in p53+ (Versus p53−) (n=19 vs 69)	Expression in K-*ras*+ (Versus K-*ras*−) (n=34 vs 40)	Expression in Samples with Presence of Lymphocytic Response (n=39 vs 51)
691	CK7		down (P=.021)				down (P=.011)
352	CK8			up (P=.027)			
436	CK8	up (P=.013)					down (P=.012)
451	CK8					down (P=.038)	
2381	CK18						up (P=.032)
608	CK19				up (P=.027)	up (P=.028)	
609	CK19			down (P=.003)			

Other CK spots that did not significantly change in expression: 871, 1968, 2091, 2165 (CK7); 439, 441, 443, 444, 446 (CK8); 527, 529, 2324 (CK18); 1955 (CK19).

and one isoform (no. 608) was underexpressed in lung adenocarcinoma.

CK Isoforms Associated with Patient Survival and Other Clinical–Pathologic Variables

Univariate Cox proportional hazards regression analysis revealed that one CK8, one CK19, and all five CK7 isoforms identified by MS were significantly associated (P<.05) with patient survival. Interestingly, the five CK7 isoforms had estimated molecular mass ranging from 36.7 to 41.7 kDa, and estimated p*I*s ranging from 4.6 to 4.9 that are quite different from the reported native CK7 molecular mass of 51.3 kDa and p*I* of 5.5. The estimated p*I* and molecular mass for CK 8, 18, and 19 isoforms were relatively similar to their reported standards [3].

Of the 21 CK isoforms examined, only a subset showed a significant correlation to other clinical–pathological variables (Table 3). Relationships between these variables may provide some insight into the role of these CKs in lung cancer. *F*-tests showed that some CK isoforms such as CK7 (no. 691), which had a relationship to survival, were decreased in bronchial-derived tumors and those with a positive lymphocytic response. None of the other CK7 isoforms showed a relationship to any clinical–pathological variables. However, three CK8 isoforms showed relationships to differentiation, positive lymphocytic response, or K-*ras* mutation status. One CK18 isoform was also increased in tumors with a positive lymphocytic response, and of the three CK19 isoforms, one was increased in tumors with *p53* and K-*ras* mutations, and one was decreased in poorly differentiated tumors.

2D Western Blot Analysis of CKs

We determined the patterns of reactivity of CKs with 2D Western blot analysis (Figure 2) of the regions of 2D gels demonstrating the immunoreactivity shown in Figure 1, *boxes A–D*. Outlined box B indicated the region from which the five overexpressed CK7 proteins identified by MS were located. The CK8 antibody identified at least 15 separate immunoreactive proteins (Figure 2*B*) including the eight isoforms that were identified by MS (Figure 1, *box A*), and for which quantitative analysis was performed (Figure 2*A*). Two different monoclonal antibodies to CK7 representing two different clones were examined (Figure 2, *C* and *D*). Both antibodies detected some isoforms in common as well as distinct CK7 isoforms. The location of the immunoreactive CK7 isoforms is very close to the location to the CK8 fragments (Figure 2*B*), and three of the isoforms (436, 444, 446) appear to overlap between CK7 and CK8 fragments. All

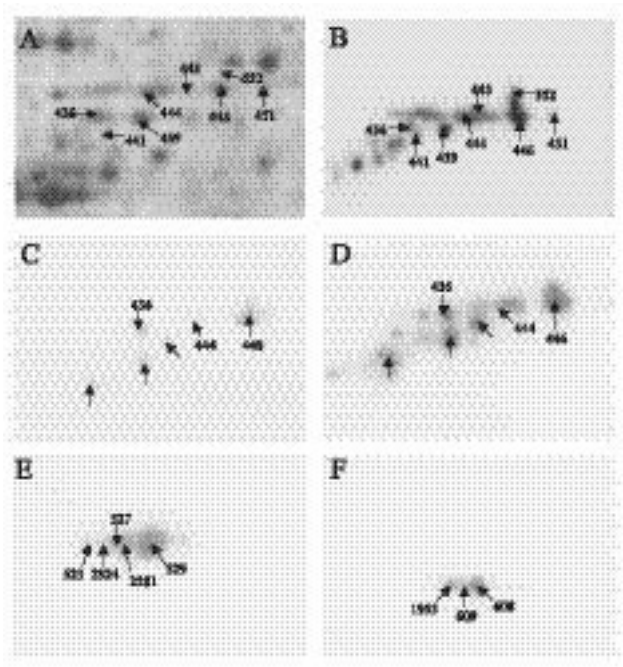

Figure 2. *(A) 2D gel digital image of the silver-stained region shown in Figure 1, Box A, containing eight CK8 spots expressed in a primary lung adenocarcinoma. (B) 2D Western blot analysis using anti-CK8 monoclonal antibody clone TS1 showing the eight isoforms quantified of the approximately 20 that are immunoreactive in A549 cells. The region is shown in Figure 1, Box A. (C) 2D Western blot analysis using anti-CK7 monoclonal antibody clone OV-TL 12/30. The region is shown in Figure 1, Box A. (D) 2D Western blot analysis using anti-CK7 monoclonal antibody clone K72.7. Three of the spots (nos. 436, 444, and 446) appear to be at very similar location to CK8 spots and are detected with both CK7 antibodies. None of the five CK7 isoforms detected by mass spectrometry reacted with either antibody. The unnumbered spots, although detected in the A549, were not abundant enough to be quantified in the lung and tumor samples. The region is shown in Figure 1, Box A. (E) 2D Western blot analysis using anti-CK18 monoclonal antibody clone DC-10. The region is shown in Figure 1, Box C. Five individual isoforms were quantified. (F) 2D Western blot analysis using anti-CK19 monoclonal antibody clone M0772. The region is shown in Figure 1, Box D. Three CK19 isoforms were quantified.*

Table 4. Correlation Coefficients Between CK Protein and mRNA Values.

Spot Number	Gene Name	Correlation Coefficient (r)*
2165	CK7	0.35
2091	CK7	0.29
1968	CK7	0.14
691	CK7	0.10
871	CK7	−0.18
439	CK8	0.30
441	CK8	0.18
436	CK8	0.11
443	CK8	0.09
352	CK8	0.04
446	CK8	0.03
451	CK8	−0.01
444	CK8	−0.13
523	CK18	0.33
2324	CK18	0.26
2381	CK18	0.20
529	CK18	0.13
527	CK18	0.04
1955	CK19	0.39
609	CK19	0.15
608	CK19	0.04

*$P<.05$ if $r>0.25$.

other fragments are clearly distinct from each other, although their location is similar. The redundancy of the immunostaining may be the result of the extensive sequence similarity of CK7 and CK8 [16,17].

The immunoreactive CK7 proteins have the same p*I* and molecular mass predicted from their sequence; however, neither of the two CK7 antibodies directed against distinct clones reacted with the five isoforms identified by MS, suggesting that these smaller CK7 proteins represented cleavage products that lacked the epitopes recognized by these monoclonal antibodies. Quantitative values for the immunoreactive CK7 isoforms could not be obtained from the silver-stained gels due to low abundance, although these forms are relatively abundant in the A549 lung cell line. All five CK7 isoforms were significantly associated with survival ($P<.05$), significantly increased in the adenocarcinomas, and frequently expressed (58.1% *vs* 39.8%) relative to normal lung (Table 2). None of the immunoreactive CK7/CK8 isoforms showed a significant relationship with survival; however, spot nos. 444 and 446 were significantly increased in tumors relative to normal lung. In addition, spot no. 436 showed a significant increased expression in stage III relative to stage I tumors, and significant decreased expression in tumors showing lymphocytic response.

Five CK18 and three CK19 isoforms were detected using MS and 2D Western blots (Figure 2, *E* and *F*) and their protein abundance quantitatively analyzed (Table 2). Unlike CK7 and CK8, all CK18 and CK19 isoforms demonstrated similar molecular mass but varied in p*I*. Interestingly, all five CK18 isoforms were significantly increased in the adenocarcinomas and detected frequently in these tumors relative to normal lung (15–60%), but an association with survival was not observed. The isoform nos. 529 and 523 are the most abundant in both normal lung and lung adenocarcinomas. All three CK19 isoforms were significantly increased in the adenocarcinoma but only one isoform (no. 1955), the most positively charged isoform, was associated with survival and demonstrated the greatest differential expression between normal lung and the adenocarcinomas. This isoform was significantly associated with the CK19 mRNA levels, suggesting that the level of transcription in part can influence its abundance.

CK mRNA Levels and Correlation to Other Genes

Because the same lung adenocarcinomas and non-neoplastic lung samples that were quantitatively analyzed for protein expression were examined at the level of mRNA abundance using oligonucleotide arrays, correlation coefficients between these CK proteins and the levels of their corresponding mRNA could be determined.

A positive correlation was observed between the CK protein expression values for two CK7 isoforms (nos. 165 and 2091) and CK7 mRNA levels (Table 4). One CK8 isoform (no. 439), two CK18 isoforms (nos. 523 and 2324), and one CK19 isoform (no. 1955) were also significantly associated with mRNA levels of corresponding genes. Importantly, univariate Cox proportional hazards regression analysis revealed that CK7, CK8, CK18, and CK19 mRNA showed a significant relationship to survival and all four were significantly increased at the mRNA level in the adenocarcinomas relative to normal lung (Table 5). Further, correlation analysis showed that the mRNA levels for all four CK genes were significantly correlated ($P<.05$) to one another (Table 6), suggesting that these increases may reflect a common mechanism. Further analysis of 4966-expressed genes examined in these samples revealed two genes encoding liver-specific bHLH-Zip transcription factor and smooth and nonmuscle myosin light chain polypeptide to be significantly correlated to all four CKs.

Cellular Localization of Candidate Proteins

Utilizing the same antibodies used for 2D Western blot analysis (Figure 2), the expression of the four CK forms was confirmed as present specifically in lung adenocarcinoma by immunohistochemistry of tumor tissue arrays (Figure 3).

Table 5. CK mRNA Expression.

Gene Name	P (Cox Model)	β Coefficient	Normal Mean ($n=10$)	Tumor Mean ($n=86$)	Fold Change (Tumor/Normal)	P (t-test)
CK7	.0004	0.0003	951	2146	2.26	<.0001
CK8	.0934	0.0001	1368	3436	2.51	<.0001
CK18	.0006	0.0003	1310	3826	2.92	<.0001
CK19	.0009	0.0003	952	2522	2.65	<.0001

Table 6. Correlation Coefficients between CKs and Other mRNA.

Gene Name	Correlation Coefficient (r)*			
	CK7	CK8	CK18	CK19
CK7	1.00	0.53	0.62	0.43
CK8	–	1.00	0.75	0.52
CK18	–	–	1.00	0.40
CK19	–	–	–	1.00
LISCH7	0.41	0.52	0.43	0.40
MYL6	0.43	0.61	0.41	0.44

*$P < .05$ if $r > 0.25$.

These arrays contained the same tumors examined in this study at both the protein and mRNA levels. A very similar pattern and cytoplasmic localization of all four of these CK proteins were observed in the lung adenocarcinomas, with much lower levels detected in normal lung. Interestingly, all four CKs showed relatively abundant expression in the epithelial counterpart of the fetal lung, consistent with previous studies [18].

Discussion

The expression of CKs 7, 8, 18, and 19 in lung adenocarcinoma includes multiple isoforms for each type and suggests that regulation may occur at a number of different levels. Several mechanisms for the regulation of CKs have been proposed. These intermediate filaments appear to be assembled and disassembled through rapid phosphorylation and dephosphorylation reactions on specific serine and threonine residues [19,20]. Phosphorylation of CKs induced by cyclic AMP stimulation [21] can be accelerated by phosphatase inhibition, and can result in the disappearance of the intermediate filament cytoskeleton [22]. Caspase-mediated cleavage leading to cytoskeleton disassembly, formation of pleomorphic cytoplasmic inclusions, and stable CK fragments [23,24] has been reported. This early apoptosis-related process has been shown to expose a neoepitope of CK18 that is not detectable in nonapoptotic epithelial cells [25].

Bergqvist et al. [5] have reported an increase in both CK8 and CK18 in the serum of non small cell lung cancer patients. The authors attributed this to increased tumor cell turnover resulting in an increased level of these CK fragments in the circulation. Alberti et al. [26] also reported massive fragmentation of CKs induced by calcium-dependent proteases. This rearrangement of CKs has been hypothesized to play a direct role in carcinogenesis by triggering the reorganization of chromatin [27], and providing a growth advantage to the transformed cell. Increased vascularization, a characteristic of growing tumors, and the altered permeability of these new blood vessels were suggested as responsible for the increase in CK fragments in the serum. It has been also postulated that endogenously produced cytokines may lead to reconstruction of the cytoskeleton in epithelial cells, thus increasing the CKs [5]. Increased CK8 expression was reported in breast carcinoma cell lines, where it functions as a plasminogen receptor, activating plasmin, which is important for tumor invasion and cellular migration [28]. CK8 was reported to contain antigenic epitopes for antigen CA19-9 [7], which serves as a ligand for endothelial cell leukocyte molecule-1 (ELAM-1). ELAM-1 mediates cell–cell interactions between platelets and endothelial cells with neutrophils, monocytes, and cancer cells [29,30]. These studies suggest that increased expression of the CKs may reflect changes in the organization of cellular architecture during tumor development and progression. It is likely that the variable expression of the specific CK isoforms examined in this study also reflects the regulatory and apoptotic events occurring in lung adenocarcinomas.

Importantly, we not only observed a relationship among the mRNA overexpression of the four CKs examined in this study (CKs 7, 8, 18, and 19) and patient outcome in the lung adenocarcinomas, but also significant overexpression of many specific isoforms for these proteins. For example, all five isoforms of CK7 were associated with patient survival. CK7 has an apparent molecular mass range of 49 to 54.3

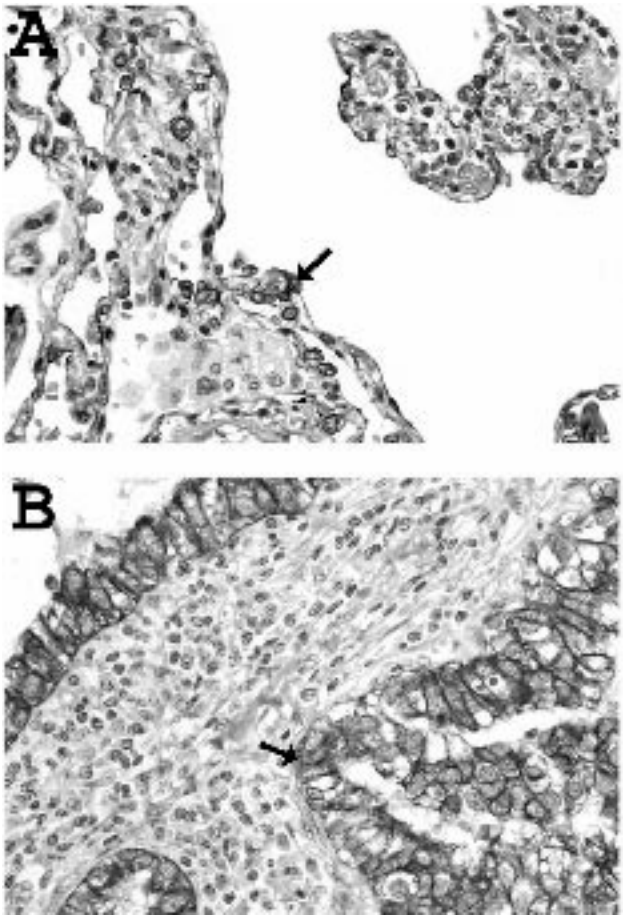

Figure 3. *(A) Immunohistochemical analysis of CK18 expression in normal non-neoplastic lung parenchymal tissue. Staining of individual epithelial cells is observed but not stromal inflammatory or supporting cell elements. Normal bronchiole epithelium also demonstrated expression of CK18 (not shown). Original magnification, ×400. (B) Immunohistochemical analysis of CK18 expression in a lung adenocarcinoma showing abundant cytoplasmic and membrane staining in all tumor cells (arrow) but not stromal cells. Original magnification, ×400.*

kDa and a p*I* range of 5.2 to 5.5, yet all five forms identified by MS were smaller and not detected by 2D Western blot analysis using two different anti-CK7 antibodies. These results suggest that these smaller CK7 fragments may be missing the epitopes recognized by the CK7 antibodies and may represent proteolyically modified forms. Given the potential importance of these smaller CK7 isoforms by their strong association with patient survival, determining their structure and possibly directing antibody production against these specific isoforms may be very clinically useful.

Eight CK8 of nearly 15 isoforms were quantified and four found to be significantly overexpressed in lung adenocarcinomas. One isoform (no. 439) was associated with an unfavorable prognosis and was correlated with CK8 mRNA levels, suggesting that the abundance of this isoform may be regulated at the level of transcription. Comparison of the immunoreactive proteins between CK7 and CK8 revealed a very similar pattern of fragments with three fragments appearing to overlap. This result is not unexpected given the extensive protein sequence similarity between both CKs [16,17]; however, our computer-based matching analyses reveal many of the isoforms to be unique in their molecular mass and p*I* values, suggesting that modifications may differ. Further, the antibodies utilized for Western blots reportedly do not cross-react.

All five CK18 isoforms are overexpressed in lung adenocarcinoma but the native form (529) had the highest frequency of expression. Given the similar molecular mass but differing p*I* of the CK18 isoforms, modification by phosphorylation may underlie these different charged isoforms. The isoforms (nos. 523 and 2324) were significantly correlated with CK18 mRNA, also suggesting that their abundance may relate to transcriptional regulation.

CK19 demonstrated three isoforms of similar molecular mass yet different p*I*s indicative of potential modification by phosphorylation. CK19 has been documented as a useful tumor marker in lung cancer [31,32]. The most abundant CK19 isoform in both normal and neoplastic lung was no. 608, which was significantly decreased in lung adenocarcinomas. The two acidic (nos. 609 and 1955), and potentially phosphorylated, isoforms were significantly overexpressed in the adenocarcinomas. Interestingly, the most phosphorylated isoform (no. 1955) was highly elevated (16-fold) in the lung tumors and was associated with unfavorable prognosis.

Correlation analysis of CK mRNA expression revealed a significant association to the expression of all four CKs to one another. This suggests that the mechanisms underlying the increased expression of these proteins in lung tumors may be similar, possibly reflecting the events associated with adenocarcinoma development and progression. Of interest, correlation analysis also revealed two other genes that also were significantly correlated in their expression to all four CK genes. This included liver-specific bHLH-Zip transcription factor and smooth, and nonmuscle, myosin light polypeptide 6 genes. The potential for these genes to be involved in CK regulation is to be determined, but may provide an approach to determine whether a common pathway is involved in the regulation of these proteins.

Acknowledgements

The authors acknowledge the contributions from Lin Lin, Charles Miller, Melissa Krause, David Wood, Robert Hinderer, Rork Kuick, and Eric Puravs.

References

[1] Landis SH, Murray T, Bolden S, and Wingo PA (1999). Cancer statistics. *CA Cancer J Clin* **49**, 6–7.

[2] Fry WA, Phillips JL, and Menck HR (1999). Ten-year survey of lung cancer treatment and survival in hospitals in the United States: a national cancer data base report. *Cancer* **86**, 1867–76.

[3] Moll R, Franke WW, and Schiller DL (1982). The catalog of human cytokeratins: patterns of expression in normal epithelia, tumors and cultured cells. *Cell* **31**, 11–24.

[4] Stigbrand T (2001). The versatility of cytokeratins as tumor markers. *Tumour Biol* **22**, 1–3.

[5] Bergqvist M, Brattstrom D, Hesselius P, Wiklund B, Silen A, Wagenus G, and Brodin O. Cytokeratin 8 and 18 fragments measured in serum and their relation to survival in patients with non-small cell lung cancer. *Anticancer Res* **19**, 1833–36.

[6] Fernandez-Madrid F, VandeVord PJ, Yang X, Karvonen RL, Simpson PM, Kraut MJ, Granda JL, and Tomkiel JE (1999). Antinuclear antibodies as potential markers of lung cancer. *Clin Cancer Res* **5**, 1393–400.

[7] Fujita J, Dobashi N, Ohtsuki Y, Ueda Y, Bandoh S, Yamadori I, and Takahara J (1999). Detection of large molecular weight cytokeratin 8 as carrier protein of CA19-9 in non-small-cell lung cancer cell lines. *Br J Cancer* **81**, 769–73.

[8] Nagai T, Murota M, Nishioka M, Fujita J, Ohtsuki Y, Dohmoto K, Hojo S, Dobashi N, and Takara J (2001). Elevation of cytokeratin 19 fragment in serum in patients with hepatoma: its clinical significance. *Eur J Gastroenterol Hepatol* **13**, 157–61.

[9] Wells MJ, Hatton MW, Hewlett B, Podor TJ, Sheffield WP, and Blajchman MA (1997). Cytokeratin 18 is expressed on the hepatocyte plasma membrane surface and interacts with thrombin–antithrombin complexes. *J Biol Chem* **272**, 28574–81.

[10] Colby TV, Koss MN, and Travis WD (1995). AFIP Atlas of Tumor Pathology: Tumors of the Lower Respiratory Tract. Armed Forces Institute of Pathology, Washington, DC.

[11] Kuick RD, Skolnick MM, Hanash SM, and Neel JV (1991). A two-dimensional electrophoresis-related laboratory information processing system: spot matching. *Electrophoresis* **12**, 736–46.

[12] Kuick RD, Hanash SM, Chu EHY, and Strahler JR (1987). A comparison of some adjustment techniques for use with quantitative spot data from two-dimensional gels. *Electrophoresis* **8**, 199–204.

[13] Giordano TJ, Shedden KA, Schwartz DR, Kuick R, Taylor JMG, Lee N, Misek DE, Greenson JK, Kardia SLR, Beer DG, Rennert G, Cho KR, Gruber SB, Fearon ER, and Hanash S (2001). Organ-specific molecular classification of primary lung, colon, and ovarian adenocarcinomas using gene expression profiles. *Am J Pathol* **159**, 1231–38.

[14] Kononen J, Bubendorf L, Kallioniemi A, Barlund M, Schraml P, Leighton S, Torhorst J, Mihatsch MJ, Sauter G, and Kallioniemi OP (1998). Tissue microarrays for high-throughput molecular profiling of tumor specimens. *Nat Med* **4**, 844–47.

[15] Oh JMC, Brichory F, Puravs E, Kuick R, Wood C, Rouillard JM, Tra J, Kardia S, Beer D, and Hanash S (2001). A database of protein expression in lung cancer. *Proteomics* **1**, 1303–19.

[16] Glass C, and Fuchs E (1988). Isolation, sequence, and differential expression of a human K7 gene in simple epithelial cells. *J Cell Biol* **107**, 1337–50.

[17] Leube RE, Bosch FX, Romano V, Zimbelmann R, Hofler H, and Franke WW (1986). Cytokeratin expression in simple epithelia: III. Detection of mRNAs encoding human cytokeratins nos. 8 and 18 in normal and tumor cells by hybridization with cDNA sequences *in vitro* and *in situ*. *Differentiation* **33**, 69–85.

[18] Broers JL, de Leij L, Rot MK, ter Haar A, Lane EB, Leigh IM, Wagnaar SS, Vooijs GP, and Ramaekers FC (1989). Expression of intermediate filament proteins in fetal and adult human lung tissues. *Differentiation* **40**, 119–28.

[19] Nakamura Y, Takeda M, and Nishimura T (1996). Dynamics of bovine glial fibrillary acidic protein phosphorylation. *Neurosci Lett* **205**, 91–94.

[20] Giasson BI, and Mushynski WE (1998). Intermediate filament disassembly in cultured dorsal root ganglion neurons is associated with amino-terminal head domain phosphorylation of specific subunits. *J Neurochem* **70**, 1869–75.

[21] Escribano J, and Rozengurt E (1988). Cyclic AMP increasing agents rapidly stimulate vimentin phosphorylation in quiescent cultures of Swiss 3T3 cells. *J Cell Physiol* **137**, 223–34.

[22] Eriksson JE, Brautigan DL, Vallee R, Olmsted J, Fujiki H, and Goldman RD (1992). Cytoskeletal integrity in interphase cells requires protein phosphatase activity. *Proc Natl Acad Sci USA* **89**, 11093–97.

[23] Ku NO, and Omary MB (1994). Identification of the major physiologic phosphorylation site of human cytokeratin 18: potential kinases and a role in filament organization. *J Cell Biol* **127**, 161–71.

[24] MacFarlane M, Merrison W, Dinsdale D, and Cohen GM (2000). Active caspases and cleaved cytokeratins are sequestered into cytoplasmic inclusions in TRAIL-induced apoptosis. *J Cell Biol* **148**, 1239–54.

[25] Leers MP, Kolgen W, Bjorklund P, Ramaekers FC, Bjorklund B, Nap M, Jornvall H, and Schutte B (1999). Immunocytochemical detection and mapping of a cytokeratin 18 neo-epitope exposed during early apoptosis. *J Pathol* **187**, 567–72.

[26] Alberti I, Barboro P, Barbesino M, Sanna P, Pisciotta L, Parodi S, Nicolo G, Boccardo F, Galli S, Patrone E, and Balbi C (2000). Changes in the expression of cytokeratins and nuclear matrix proteins are correlated with the level of differentiation in human prostate cancer. *J Cell Biochem* **79**, 471–85.

[27] Spencer AV, Coutts AS, Samuel AK, Murphy LC, and Davie JR (1998). Estrogen regulates the association of intermediate filament proteins with nuclear DNA in human breast cancer cells. *J Biol Chem* **273**, 29093–97.

[28] Hembrough TA, Kralovich KR, Li L, and Gonias SL (1996). Cytokeratin 8 released by breast carcinoma cells *in vitro* binds plasminogen and tissue-type plasminogen activator and promotes plasminogen activation. *Biochem J* **317**, 763–69.

[29] Takada A, Ohmori K, Takahashi N, Tsuyuoka K, Yago A, Zenita K, Hasegawa A, and Kannagi R (1991). Adhesion of human cancer cells to vascular endothelium mediated by a carbohydrate antigen, sialyl Lewis A. *Biochem Biophys Res Commun* **179**, 713–19.

[30] Takada A, Ohmori K, Yoneda T, Tsuyuoka K, Hasegawa A, Kiso M, and Kannagi R (1993). Contribution of carbohydrate antigens sialyl Lewis A and sialyl Lewis X to adhesion of human cancer cells to vascular endothelium. *Cancer Res* **53**, 354–61.

[31] Pujol JL, Grenier J, Daures JP, Daver A, Pujol H, and Michel FB (1993). Serum fragment of cytokeratin subunit 19 measured by CYFRA 21-1 immunoradiometric assay as a marker of lung cancer. *Cancer Res* **53**, 61–66.

[32] Pujol JL, Gernier J, Parrat E, Lehmann M, Lafontaine T, and Quanyin X (1996). Cytokeratins as serum markers in lung cancer: a comparison of CYFRA-21-1 and TPS. *Am J Respir Crit Care Med* **154**, 725–33.