Research Article

Cancer Epidemiology, Biomarkers & Prevention

Collagen XXIII: A Potential Biomarker for the Detection of Primary and Recurrent Non–Small Cell Lung Cancer

Kristin A. Spivey[1], Jacqueline Banyard[1], Luisa M. Solis[5], Ignacio I. Wistuba[5,6], Justine A. Barletta[4], Leena Gandhi[3], Henry A. Feldman[2], Scott J. Rodig[4], Lucian R. Chirieac[4], and Bruce R. Zetter[1]

Abstract

Background: Collagen XXIII is a transmembrane collagen previously shown to be upregulated in metastatic prostate cancer. The purpose of this study was to determine the protein expression of collagen XXIII in tumor tissues from a variety of cancers and to assess the utility of collagen XXIII as a biomarker for non–small cell lung cancer (NSCLC).

Methods: A multicancer tissue microarray was used for the immunohistochemical examination of collagen XXIII protein expression in a variety of cancers. Subsequently, collagen XXIII expression was analyzed in three separate cohorts using tissue microarrays with representative tumor and control lung tissues from NSCLC patients. In addition, NSCLC patient urine samples were analyzed for the presence of collagen XXIII through Western blot.

Results: Collagen XXIII was present in tissue samples from a variety of cancers. Within lung cancer tissues, collagen XXIII staining was enriched in NSCLC subtypes. Collagen XXIII was present in 294 of 333 (88%) lung adenocarcinomas and 97 of 133 (73%) squamous cell carcinomas. In urine, collagen XXIII was present in 23 of 29 (79%) NSCLC patient samples but only in 15 of 54 (28%) control samples. High collagen XXIII staining intensity correlated with shorter recurrence-free survival in NSCLC patients.

Conclusions: We show the capability of collagen XXIII as a tissue and urinary biomarker for NSCLC, in which positivity in tissue or urine significantly correlates with the presence of NSCLC and high staining intensity is a significant recurrence predictor.

Impact: Inclusion of collagen XXIII in a tissue- or urine-based cancer biomarker panel could inform NSCLC patient treatment decisions. *Cancer Epidemiol Biomarkers Prev; 19(5); 1362–72. ©2010 AACR.*

Introduction

Lung cancer is the leading cause of cancer-associated mortality, accounting for over 25% of cancer-related deaths and ~15% of new cancer cases annually (1). Lung cancer is divided into small cell lung cancer (SCLC) and non-SCLC (NSCLC), with prevalence rates of 14% and 85%, respectively. NSCLC is classified into three types: adenocarcinoma, squamous cell carcinoma (SqCC), and large cell carcinoma, based on histologic and immunophenotypic characteristics (2). Cigarette smoking, measured by quantity and frequency of consumption, is still the leading risk factor for lung cancer, accounting for 75% to 87% of lung cancer cases, although the proportion of never smoker–related lung cancer is increasing (1, 3). Lung cancer patients have an overall 5-year survival rate of 15%, with only 16% of patients diagnosed with early-stage disease that is still treatable by curative therapy and 5-year survival rates of 50% (1). Current screening tools such as chest X-rays, sputum cytologic examination, or computed tomography scans have yet to show an effect on mortality, although computed tomography scans are capable of detecting early-stage cancers (4, 5).

Recent technological advances have facilitated the identification of genetic, epigenetic, and proteomic alterations in a variety of cancers. For example, treatment with tyrosine kinase inhibitors is informed by genetic and proteomic changes in epidermal growth factor receptor and v-Ki-ras2 Kirsten rat sarcoma viral oncogene homologue (*KRAS*; refs. 6-8) in NSCLC patients. Histological and immunohistochemical (IHC) analysis of biopsied tissue is the current standard for diagnosing

Authors' Affiliations: [1]Vascular Biology Program, Department of Surgery and [2]Clinical Research Program, Children's Hospital; [3]Thoracic Oncology Program, Dana-Farber Cancer Institute; and [4]Department of Pathology, Brigham and Women's Hospital, Boston, Massachusetts; and Departments of [5]Pathology and [6]Thoracic, Head and Neck Medical Oncology, M.D. Anderson Cancer Center, Houston, Texas

Note: Supplementary data for this article are available at Cancer Epidemiology Biomarkers and Prevention Online (http://cebp.aacrjournals.org/).

Corresponding Author: Bruce Zetter, Charles Nowiszewski Professor of Cancer Biology in the Department of Surgery, Karp Family Research Laboratories, Room 11.125, One Blackfan Circle, Boston, MA 02115. Phone: 617-919-2320; Fax: 1-617-730-0268. E-mail: Bruce.Zetter@childrens.harvard.edu

doi: 10.1158/1055-9965.EPI-09-1095

and classifying tumors, but the invasiveness and associated morbidity of the technique makes it impractical for patient screening and monitoring. In contrast, biological fluids such as urine and serum are an accessible source of disease markers and can be easily and repeatedly obtained from patients. There is also a renewed interest in analyzing breath condensate for lung cancer detection (9, 10), although standardized protocols are still being developed. Identification of biomarkers capable of detecting primary lung cancers or recurrent tumors is imperative for improved patient outcome.

Collagen XXIII, a type II transmembrane collagen, was originally discovered in our laboratory using differential display analysis to identify transcripts that were upregulated in highly metastatic prostate cancer (11). It was later validated as a tissue marker for disease recurrence after prostatectomy (12). In this study, we identify additional cancers that overexpress collagen XXIII, with a focus on lung cancer due to the scarcity of validated biomarkers for this disease. We show that collagen XXIII positivity significantly correlates with the presence of NSCLC in biopsied lung tissue and in patient urine samples and, therefore, is a potential noninvasive biomarker for NSCLC. Additionally, elevated collagen XXIII staining intensity correlates with shorter time to recurrence in NSCLC patients, indicating its utility as a biomarker for recurrent disease.

Materials and Methods

Cell culture, transfection, and lysate preparation

H460 and HEK293 cells were obtained from the American Type Culture Collection and were maintained in their recommended media (RPMI or DMEM, respectively) supplemented with 10% fetal bovine serum and 1% penicillin-streptomycin-glutamine (Invitrogen). For overexpression studies, HEK293 cells were transfected with either an empty PCDNA3 vector (Invitrogen), PCDNA3-human collagen XXIII (11), PCDNA3-human collagen XIII (12), or pCMV6-XL4-human collagen XXV transcript variant 1 (Origene) using the Fugene 6 reagent (Roche Diagnostics). For knockdown studies, H460 cells were transfected with an empty pRS vector (Origene) or pRS encoding short hairpin RNA against collagen XXIII (pRS-shc23; Origene) using Lipofectamine 2000 (Invitrogen). After selection in puromycin, clonal cell lines were established. Whole-cell lysates were collected using Triton Lysis Buffer (Boston BioProducts) and supplemented with compete mini protease inhibitor cocktail tablets (Roche Diagnostics). Lysate protein concentration was determined using the BCA assay (Pierce Biotechnology).

Antibody validation

A new human collagen XXIII mouse monoclonal antibody was made against the cleaved portion of human collagen XXIII by R&D Systems. The specificity of this antibody, clone 468642, for endogenously expressed collagen XXIII was examined through Western blotting. Twenty micrograms of whole-cell lysates from H460 cells expressing pRS or pRS-shc23 were resolved on a 7.5% Tris-HCl gel and transferred to a 0.45 μmol/L polyvinylidene difluoride membrane; (Millipore). Following transfer, membranes were immunoblotted with the 468642 antibody, collagen XXV (R&D Systems), or actin (Chemicon) antibodies and developed using the ECL Plus chemiluminescence kit (Amersham). Western blot analysis revealed weak cross-reactivity between the 468642 antibody and overexpressed collagen XXV,[7] but no detection of endogenous collagen XXV (Supplementary Fig. S1A). The range of reactivity of the 468642 antibody was determined by immunoblotting membranes containing 0 to 5 ng of recombinant collagen XXIII protein (R&D Systems; Supplementary Fig. S1B).

Immunohistochemical and Western blotting peptide competition experiments were used to confirm the specificity of a newly glutathione *S*-transferase (GST)-collagen XXIII peptide purified batch of rabbit polyclonal collagen XXIII antibody (6.1B) as previously described (Supplementary Fig. S2A and B); ref. 12). Briefly, the antibody was preincubated with GST or a GST-collagen XXIII peptide at a ratio of 1:4. Sections of lung adenocarcinoma tumor tissue or membranes containing HEK293 lysates overexpressing collagen XXIII were stained with the antibody alone or GST-collagen XXIII peptide competed antibody. Antibody cross-reactivity was determined through Western blotting (Supplementary Fig. S2C). Briefly, whole-cell lysates from HEK293 cells transfected with PCDNA3, PCDNA3-human collagen XXIII, PCDNA3-human collagen XIII, or pCMV6-XL4-human collagen XXV were resolved on a 7.5% Tris-HCl gel and transferred to a polyvinylidene difluoride membrane. Membranes were immunoblotted with the 6.1B antibody or rabbit polyclonal collagen XXV antibody raised against the third noncollagenous domain (NC3; ref. 13) and developed using the ECL Plus chemiluminescence kit (Amersham). Weak cross-reactivity was observed between the polyclonal collagen XXIII 6.1B antibody and overexpressed collagen XXV.

Tissue microarray construction

To analyze collagen XXIII expression in NSCLC, immunohistochemical staining was done on three separate tissue microarrays (TMA) representing different patient cohorts.

Brigham & Women's TMA. The TMA was constructed as previously described (14). In brief, the TMA was constructed from representative tissue blocks from formalin-fixed, paraffin-embedded archival tissue specimens. Three areas of tumor and three adjacent areas of normal lung parenchyma from 86 patients with lung adenocarcinoma were sampled from each case in 0.6-mm-diameter cores and transferred into paraffin recipient blocks using a manual arrayer (Beecher Instruments, Inc.). The three

[7] Unpublished data.

Table 1. Collagen XXIII immunohistochemical analysis of multiple cancer tissues

Organ	Pathology	*n*	Collagen XXIII staining*†
Bladder	Transitional cell carcinoma	18	+
Breast	Infiltrating lobular carcinoma	8	+
Breast	Nonspecific infiltrating duct carcinoma	16	+
Cerebrum	Astrocytoma	16	+
Colon	Adenocarcinoma	18	+
Fatty tissue	Liposarcoma	10	–
Fibrous tissue	Fibrosarcoma	5	–
Head and neck	SqCC	16	+
Kidney	Clear-cell carcinoma	10	+
Kidney	Granular cell carcinoma	10	+
Liver	Hepatocellular carcinoma	19	+
Lung	NSCLC	21	+
Lung	SCLC	7	–
Lymph node	Hodgkin's disease	9	–
Ovary	Serous papillary carcinoma	11	+
Pancreas	Adenocarcinoma	17	+
Prostate	Adenocarcinoma	16	+
Skin	Malignant melanoma	18	+
Skin	SqCC	7	+
Stomach	Adenocarcinoma	19	+
Testis	Seminoma	15	+
Thyroid	Follicular carcinoma	9	+
Thyroid	Papillary carcinoma	7	+
Uterus	Adenocarcinoma	18	+

*Average collagen XXIII intensity values <0.5 (–) or >5 (+) on a 0- to 3-point staining intensity scoring scale are reported for each cancer type.
†IHC was done using the 468642 antibody.

cored areas of tumor on each donor block were randomly selected from different parts of the tumor tissue based on a histologic characterization of the H&E-stained slide.

M.D. Anderson Cancer Center TMA. Archived formalin-fixed, paraffin embedded material was obtained from the Lung Cancer specialized Program of Research Excellence Tissue Bank at The University of Texas M.D. Anderson Cancer Center. Use of these tissues has been approved by M.D. Anderson Cancer Center Institutional Review Board. We selected 225 NSCLC tumor tissues from surgical resected lung cancer specimens (138 adenocarcinomas and 87 SqCCs). Tumor tissues were histologically analyzed and classified using the 2004 WHO classification system (2). Three areas of tumor tissue, representing the center, intermediate, and peripheral zone of the tumor, were sampled in 1-mm-diameter cores and included on a TMA. Detailed clinicopathologic information including demographics, smoking status (never, former, and current smokers), and clinicopathologic tumor-node-metastasis stage was available for most cases.

U.S. Biomax TMA. Multicancer (MC5001) and lung cancer (LC801, LC810, LC992, BS04071, and BS04081) TMAs were purchased from U.S. Biomax. The TMAs were composed of one to three representative formalin-fixed, paraffin-embedded tissue specimens per patient. The lung cancer TMAs included 155 tissue samples from lung adenocarcinoma and squamous cell patients and 38 normal control tissues.

Immunohistochemical staining and evaluation

Tissue sections were deparaffinized in xylene and rehydrated through a series of alcohols. Sections were treated with Antigen Retrieval Solution (Vector Laboratories), heated using a pressure cooker (Biocare Medical; peak temp 125°C for 30 seconds, followed by 90°C for 10 s), and cooled for 20 minutes. After washing in water and TBS-T, slides were placed in a humidified chamber and incubated with a peroxidase-blocking reagent (DAKO) for 5 minutes. Immunohistochemistry was done by blocking slides in 20% goat serum (DAKO) before incubating slides with the rabbit polyclonal collagen XXIII 6.1B antibody, followed by a biotinylated goat-anti-rabbit secondary antibody (DAKO). Extensive TBS washes were done between steps. The chromogen substrate, 3,3′-diaminobenzidine (DAKO), was added for 5 minutes to visualize bound antibodies and Mayer's hematoxylin

was used to visualize cell nuclei. This same protocol was used for the mouse monoclonal 468642 antibody with the following alterations: EDTA (pH 6.0; Zymed) was used for antigen retrieval and Mouse Envision (DAKO) was used for the secondary antibody. Same-species IgG serum was used as a negative control. The tissue core staining score was calculated using established pathology methods, using a range of 0 to 3 or 0 to 300.

The Brigham & Women's Hospital TMA was immunostained using the collagen XXIII 6.1B antibody. Staining intensity was scored by two pathologists (JAB and LRC) on a four-point scale: absent (0), mild (1+), moderate (2+) or intense (3+) staining (Supplementary Fig. S3).

The U.S. Biomax TMAs (LC801, LC992, LC5001, LC810, BS04071, and BS04081) were immunostained using the 468642 antibody. Staining intensity was scored by a pathologist (SJR) on a four point scale: absent (0), mild (1+), moderate (2+), or intense (3+) staining. For the multicancer TMA, average collagen XXIII intensity values <0.5 (−) or >0.5 (+) are reported for each cancer type (Table 1).

The M.D. Anderson Cancer Center TMA was immunostained using the 468642 antibody. For each tissue core, a pathologist (LMS) determined the percentage of tumor cells (0-100%) with absent (0), mild (1+), moderate (2+), or intense (3+) cytoplasmic staining. The following equation was used to calculate a representative score per tissue core: staining score (0-300) = 0 (% cells with absent staining) + 1 (% cells with mild staining) + 2 (% cells with moderate staining) + 3 (% cells with intense staining). Core scores greater than zero were considered positive in this study.

Individual tissue images were captured using the Spot Insight imaging system and white balanced using Adobe Photoshop CS2.

Patients were excluded if they lacked evaluable tumor tissue samples on the TMA, were previously diagnosed with cancer within 5 years of the surgery date, did not have adenocarcinoma or SqCC, or received neoadjuvant therapy.

Patient urine samples

NSCLC patient urine samples. Subjects were identified from the Thoracic Oncology Clinic at Dana-Farber Cancer Institute and consented on an Institutional Review Board–approved protocol for the collection of clinical specimens, including urine in patients with known or suspected thoracic malignancies. Specimens were frozen on dry ice within 1 hour of collection and stored at −80C until processing for Western blots. The age of the patients ranged from 36 to 76 years.

Control urine samples. Normal urine samples were collected under an approved Institutional Review Board protocol issued by the Children's Hospital Boston. Control samples included disease-free individuals and several nonmalignant urological conditions, such as vasectomy and infertility. The age of the control urine patients ranged from 20 to 79 years.

Urine samples were thawed on ice and diluted in one volume of 10 mmol/L ammonium bicarbonate buffer (pH 8) supplemented with Halt Protease Inhibitor +EDTA (Pierce Biotechnology). Samples were concentrated to 1 mL by centrifugation, frozen at −80°C, lyophilized for 2 days, and then resuspended in one volume of 10 mmol/L ammonium bicarbonate (pH 8) and protease inhibitors. Protein concentration was determined (Bio-Rad Protein Assay) and samples were denatured in Laemmli's sample buffer by boiling for

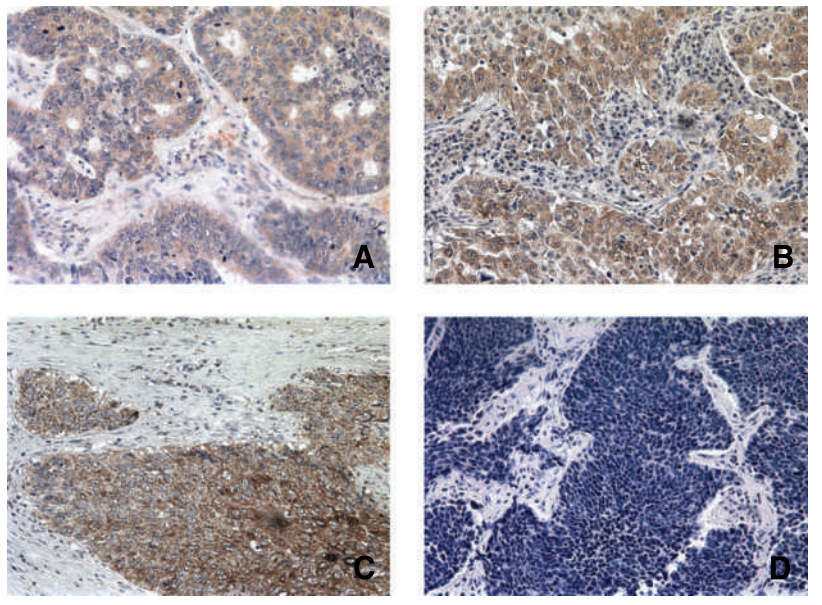

Figure 1. Immunohistochemical analysis of collagen XXIII in lung cancer subtypes. Representative images showing collagen XXIII staining in the NSCLC subtypes: (A) lung adenocarcinoma, (B) large cell carcinoma, and (C) SqCC. D, collagen XXIII staining is absent in SCLC tissues. Staining was done using the 468642 antibody and images were acquired using a ×20 objective lens.

5 minutes. Fifty micrograms of each urine sample and 1 ng of recombinant collagen XXIII protein were resolved on each gel. Following transfer, membranes were immunoblotted with the 468642 antibody and developed using the ECL Plus chemiluminescence kit (Amersham). Membrane images with similar signal intensities for the 1-ng recombinant collagen XXIII lane were chosen for subsequent scoring. Patients were scored as positive for collagen XXIII if a detectable band was present in a lane containing their urine sample. The representative image in the main text was cropped in Adobe Photoshop CS but is otherwise unaltered.

Statistical analysis

Mixed model linear regression, with a random effect to adjust for multiple cores per patient, was used to model the relationship between collagen XXIII expression and patient clinical demographic variables. These variables, examined individually, included age, gender, differentiation status (differentiation), tumor status (T), tumor size (cm), nodal status (N), pathologic stage, smoking status (never, former, or current), and tobacco story (never or 'former or current'). The coefficient, β, 95% confidence intervals, and P value, if significant, are reported for each model. The χ^2 test or Fisher's exact test were used to analyze the relationship between urinary or tissue collagen XXIII positivity and NSCLC.

To analyze the specificity and sensitivity of collagen XXIII positivity or collagen XXIII staining intensity (low or high) as an indicator of NSCLC, we generated receiver operating characteristic (ROC) curves for the Brigham & Women's Hospital and the U.S. Biomax cohorts individually and collectively. The M.D. Anderson

Table 2. Clinical demographics of patient groups from three NSCLC patient cohorts

		Brigham & Women's Hospital*		U.S. Biomax†		M.D. Anderson Cancer Center†	
No. of patients		86		155		225	
Age	Median (range)	68	(34-84)	56	(30-78)	67	(33-90)
Diagnosis	Adenocarcinoma	86	100%	109	70%	138	61%
	SqCC	0	0%	46	30%	87	39%
Gender	Female	56	65%	51	33%	115	51%
	Male	30	35%	104	67%	110	49%
Differentiation‡	Well (1)	10	11%	23	15%	49	22%
	Moderate (2)	41	48%	62	40%	112	50%
	Poor (3)	35	41%	66	43%	64	28%
Tumor status (T)	T1	32	37%	§	§	84	37%
	T2	40	47%	§	§	118	52%
	T3	3	3%	§	§	10	4%
	T4	11	13%	§	§	13	6%
Nodal status (N)‖	N0	55	64%	§	§	155	69%
	N1	13	15%	§	§	45	20%
	N2	9	10%	§	§	25	11%
Distant metastasis	Absent/undetected	77	90%	§	§	217	96%
	Present	9	10%	§	§	8	4%
Smoking status‖	Current	42	49%	§	§	78	35%
	Former	23	27%	§	§	113	50%
	Never	10	12%	§	§	34	15%
Pathologic stage¶	IA & IB	§	§	§	§	139	62%
	IIA & IIB	§	§	§	§	46	20%
	IIIA & IIIB	§	§	§	§	31	14%
	IV	§	§	§	§	8	4%

*IHC was done using the 6.1B antibody.
†IHC was done using the 468642 antibody.
‡Missing individual patient information for U.S. Biomax TMAs.
§Data not available for any patients.
‖Missing individual patient information for Brigham and Women's Hospital TMA.
¶Missing individual patient information for M.D. Anderson Cancer Center TMA.

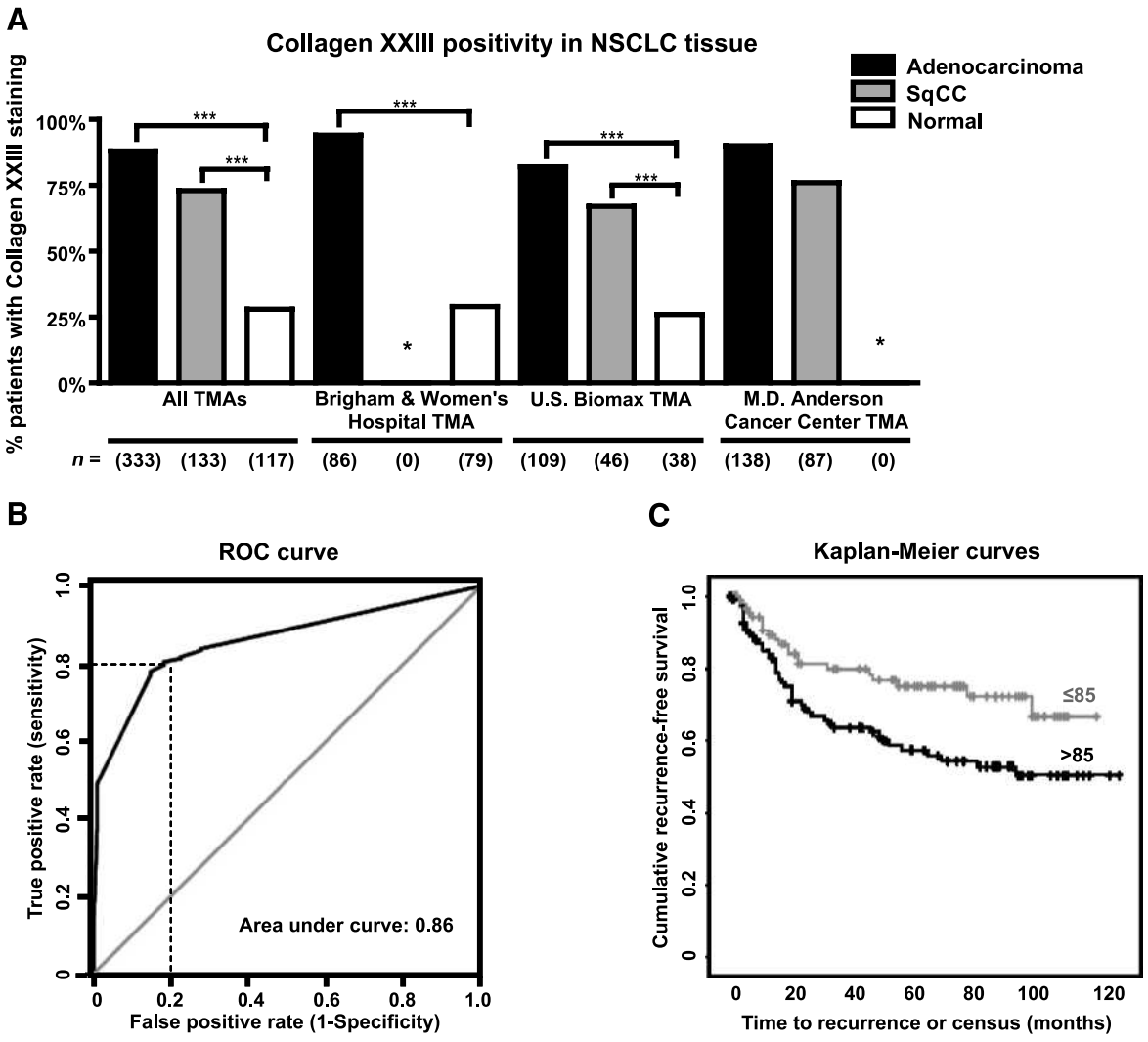

Figure 2. Collagen XXIII positivity in tissue samples correlates with the presence of adenocarcinoma and SqCC. A, bar graph showing percentage of patient tumor or normal tissue cores expressing collagen XXIII across all NSCLC TMAs and each individual TMA. *, unavailable data. Patients lacking detectable collagen XXIII staining in all of their representative tissue cores were considered negative for collagen XXIII protein expression (Supplementary Table S1). χ^2 analysis of a 3 by 2 contingency table for the combined TMA data were highly significant ($P < 0.001$). A two-tailed P value was calculated for each comparison between NSCLC subtype and normal lung tissue using Fisher's Exact Test. Highly significant differences ($P < 0.001$) are indicated (***). For immunohistochemical analysis, the 468642 antibody was used on the U.S. Biomax and M.D. Anderson Cancer Center TMAs, whereas the 6.1B antibody was used on the Brigham and Women's Hospital TMA. B, ROC curve plotting the true positive rate against the false positive rate for the different possible collagen XXIII staining intensity thresholds (tabulated data in Supplementary Table S2); area under the curve, 0.860. C, representative Kaplan Meier curves showing reduced recurrence-free survival in patients with high (>85) collagen XXIII staining intensity ($P = 0.013$). Data for all examined threshold values is included in Supplementary Table S4.

cohort was not used for this analysis because it lacked normal tissue samples. Threshold values of collagen XXIII staining intensity were evaluated in subsequent statistical analysis if the specificity at that threshold value was 80%. For the M.D. Anderson Cancer Center cohort, multiple threshold values were evaluated for their ability to predict patient recurrence or survival.

Fisher's Exact Test was done to examine the relationship between collagen XXIII positivity (present or absent) or staining intensity (low or high) and recurrence. To assess the association of collagen XXIII positivity with the duration of survival and time to recurrence, we compared the Kaplan-Meier curves for collagen XXIII–positive and collagen XXII–negative tissues by the log-rank test. This analysis was repeated using different staining intensity thresholds. To corroborate the result, we conducted Cox Proportional Hazards regression analysis using high or low staining intensity as predictor variable, first alone and then controlled for stage and age. Only patients with recurrence or census within 120

months of diagnosis were included in the survival analysis. The analyses were done by SPSS 15.0 for Windows.

Results

Immunohistochemical analysis of collagen XXIII staining in multiple cancer tissues

We sought to determine whether collagen XXIII protein expression is upregulated in other human cancers in addition to prostate cancer. Immunohistochemical analysis of TMAs allowed for high-throughput examination of collagen XXIII expression across multiple patients and cancer types. In addition to prostate adenocarcinoma, collagen XXIII was detected in a variety of cancers, including cancer of the bladder, breast, kidney, lung, liver, skin, stomach, testis, thyroid, and uterus (Supplementary Figs. S4 and S5; Table 1). The absence of collagen XXIII in fibrosarcoma, liposarcoma, or Hodgkin's disease tissue samples suggests an expression bias for carcinomas. In all tissues examined, collagen XXIII staining was predominantly cytoplasmic.

For this study, a new monoclonal collagen XXIII antibody, clone 468642 from R&D Systems, was used. Its specificity for endogenous collagen XXIII was determined through Western blot analysis of collagen XXIII and collagen XXV in H460 cells with or without short hairpin RNA against collagen XXIII (Supplementary Fig. S1). The 468642 antibody recognized the full-length and cleaved forms of collagen XXIII, running at 75 kDa and 60 kDa, respectively, in parental and control H460 lysates that were absent in the knockdown cell lysate (11). In contrast, endogenous collagen XXV (85 kDa) was not recognized by the 468642 antibody.

A newly purified batch of the collagen XXIII 6.1B polyclonal antibody (12) was also screened for specificity as previously described. Collagen XXIII-GST peptides successfully blocked collagen XXIII antibody reactivity through Western blot analysis and IHC (Supplementary Fig. S2).

Collagen XXIII in lung cancer tissue

In screening multiple cancer tissues for collagen XXIII expression, we observed that collagen XXIII expression was elevated in a large subset of NSCLC tissues. Specifically, collagen XXIII staining was present in many of the large cell carcinoma, SqCC, and adenocarcinoma tissue samples examined (Fig. 1). In most cases, collagen XXIII expression was limited to the epithelial cells comprising the tumor, but absent or weakly present in the tumor stroma. The scarcity of large cell lung carcinoma tissue arrays and the lack of collagen XXIII expression in six of seven SCLC tissue samples led us to study the association between collagen XXIII expression in the two most common types of NSCLC, SqCC, and adenocarcinoma.

Collagen XXIII positivity in tissue samples from NSCLC patients

To analyze collagen XXIII expression in NSCLC, immunohistochemical staining was done on three separate TMAs representing different patient cohorts and scored as described in Materials and Methods section (Table 2). Initially, we examined collagen XXIII positivity in NSCLC and normal lung tissues. Because some variability in

Table 3. Clinical demographics of patient groups from three NSCLC TMAs*

		Brigham & Women's Hospital†	U.S. Biomax‡,§	M.D. Anderson Cancer Center‡,§
Age	Years	−0.02 (−0.03, 0.00)‖	0.00 (−0.16, 0.02)	−0.86 (−1.59, −0.12)‖
Gender	Male (0) Female (1)	−0.14 (−0.50, 0.22)	0.18 (−0.13, 0.49)	−11.14 (−27.06, 4.78)
Differentiation	Well (1), moderate (2), poor (3)	0.37 (0.12, 0.61)¶	0.05 (−0.15, 0.24)	−1.45 (−12.79, 9.88)
Tumor status (T)	T1-4 (1-4)	0.21 (0.04, 0.39)‖	—	−6.43 (−16.64, 3.76)
Tumor size	cm	0.20 (0.10, 0.30)¶	—	−2.78 (−6.53, 0.97)
Nodal status	N0 (0), N1 (1), N2 (2)	0.24 (−0.17, 0.64)	—	−0.55 (−12.22, 11.13)
Pathologic stage**	I-IV (1-4)	—	—	−1.15 (−10.46, 8.16)
Smoking status††	Never (0), former (1), current (2)	0.13 (−0.12, 0.37)	—	7.30 (−4.44, 19.03)
Tobacco story††	Never (0), former, or current (1)	0.29 (−0.22, 0.81)	—	12.87 (−9.32, 35.06)

*Coefficient from linear regression, β (linear mixed models with patient ID; 95% confidence intervals). **$P < 0.01$
†IHC was done using the 6.1B antibody.
‡Combined NSCLC patients.
§IHC was done using the 468642 antibody.
‖$P < 0.05$.
¶$P < 0.001$.
**Missing individual patient data from M.D. Anderson Cancer Center TMA.
††Missing individual patient data from Brigham Women's & Hospital TMA.

replicate tissue core staining was observed, patients without staining in any of their replicate tissue cores were considered negative for collagen XXIII protein expression. Collagen XXIII was present in 88% of adenocarcinoma and 73% of SqCC patient tissues, but only in 28% of normal lung tissues. Importantly, collagen XXIII staining intensity in normal lung tissue was always mild or absent. Analysis of a contingency table containing either the combined TMA data or individual TMA data revealed that collagen XXIII positivity was significantly correlated with the presence of NSCLC ($P < 0.001$;Supplementary Table S1; Fig. 2A). Subsequent analysis of the combined TMA data using the Fisher's exact test confirmed the association between collagen XXIII positivity and both SqCC ($P < 0.001$) and adenocarcinoma ($P < 0.001$). Highly significant relationships between lung adenocarcinoma and collagen XXIII presence were corroborated in the Brigham & Women's Hospital ($P < 0.001$) and U.S. Biomax ($P < 0.001$) TMAs. Furthermore, collagen XXIII positivity also correlates with presence of SqCC in the U.S. Biomax TMA ($P < 0.001$). Collagen XXIII tissue positivity displayed sensitivities of 88% in adenocarcinoma and 73% in SqCC patients and specificity of 72%.

ROC analysis

ROC analysis was used to evaluate how distinct cutoff values serve to discriminate between cancer and normal tissue in the Brigham and Women's Hospital and U.S. Biomax cohorts, both of which included both cancer and normal tissue cores. We found that an average staining intensity cutoff of 0.5 on a three-point scale gave 88% sensitivity and 82.7% specificity in the Brigham and Women's Hospital cohort and 78% sensitivity and 75% specificity in the U.S. Biomax cohort. When the cohorts were analyzed together, a staining intensity cutoff of 0.5 exhibited 80% specificity and 81% sensitivity. The area under the ROC curve for both cohorts combined was 0.86. These results suggest that establishment of a collagen XXIII staining intensity threshold value can better discriminate between cancer and normal tissue compared with staining positivity (Supplementary Table S2; Fig. 2B).

Correlations with clinical and pathologic variables

Due to the variability of collagen XXIII staining intensity among the patients, we hypothesized that collagen XXIII might also be a useful predictor for disease stage or might correlate with patient clinical variables. Mixed-model linear regression, with an adjustment for multiple cores per patient, was used to model the relationship between collagen XXIII staining intensity on a continuous scale and patient variables. Significant correlations with differentiation status, tumor size (cm), tumor status (T), and age were observed in at least one TMA, but not with gender, nodal status, or smoking status (Table 3). For example, in the Brigham and Women's Hospital cohort, which was limited to adenocarcinoma patients, the collagen XXIII staining score increased by ~0.2 for each unit change in tumor size [cm; 0.20 (0.10, 0.30) $P < 0.001$] and tumor status [0.21 (0.04, 0.39) $P < 0.05$]. Increased collagen XXIII staining scores were also observed for each unit change in differentiation status in the Brigham and Women's Hospital cohort [0.37 (0.12, 0.61) $P < 0.01$] and when analyzing only adenocarcinoma patients in the M.D. Anderson Cancer Center cohort [14.84 (0.75, 28.95) $P < 0.05$; Supplementary Table S3]. We did not observe a significant association between collagen XXIII staining and differentiation status or tumor status in the U.S. Biomax or M.D. Anderson Cancer Center cohorts when all NSCLC patients were analyzed together. The collagen XXIII staining score decreased by 0.02 and 0.85 for each unit change in age when analyzing the Brigham and Women's Hospital [−0.02 (−0.03, −0.003) $P < 0.05$] and M.D. Anderson Cancer Center [−0.86 (−1.59, −0.12) $P < 0.05$] TMAs, respectively, suggesting that collagen XXIII expression is higher in younger patients with NSCLC in these cohorts. Stratification of patients by histologic type yielded similar association trends in most cases (Supplementary Table S3).

Using Fisher's exact test, we observed a significant correlation between collagen XXIII positivity and recurrence ($P = 0.022$). When examined as a continuous variable, collagen XXIII staining was not a significant predictor of recurrence, suggesting that collagen XXIII staining should be evaluated as a binary variable. Subsequently, Kaplan Meier curves were constructed to determine whether collagen XXIII positivity is predictive of time to recurrence in the M.D. Anderson Cancer Center TMA and time to survival in the M.D. Anderson Cancer Center and Brigham and Women's Hospital TMAs. Positive staining for collagen XXIII was not a significant predictor of time to recurrence (log-rank test, $P = 0.08$). Based on the ROC analysis, we examined next whether the use of various collagen XXIII staining threshold values could better predict recurrence-free survival in this cohort. Kaplan Meier curves were generated to compare time to recurrence in NSCLC patients using different staining intensity thresholds between 0 and 100 on a 300-point scale and were compared using the log-rank test. Using threshold values between 40 and 85, collagen XXIII staining intensity significantly correlated with shorter time to recurrence (log-rank test, $P \leq 0.05$; Supplementary Table S4). The most significant association between collagen XXIII and time to recurrence was observed when using a staining score of 85 as the threshold (log-rank test, $P = 0.013$; Fig. 2C). Collagen XXIII staining positivity and staining intensity were not significant predictors of time to survival in either cohort when analyzing all patients together or when stratifying patients by NSCLC subtype. Our results were corroborated using a Cox regression model in which collagen XXIII staining intensity, using a threshold value of 85, was a significant predictor of recurrence in NSCLC patients ($P = 0.015$). Overall, collagen XXIII correlated strongly with NSCLC occurrence and was a significant predictor of recurrence when analyzed as a binary indicator.

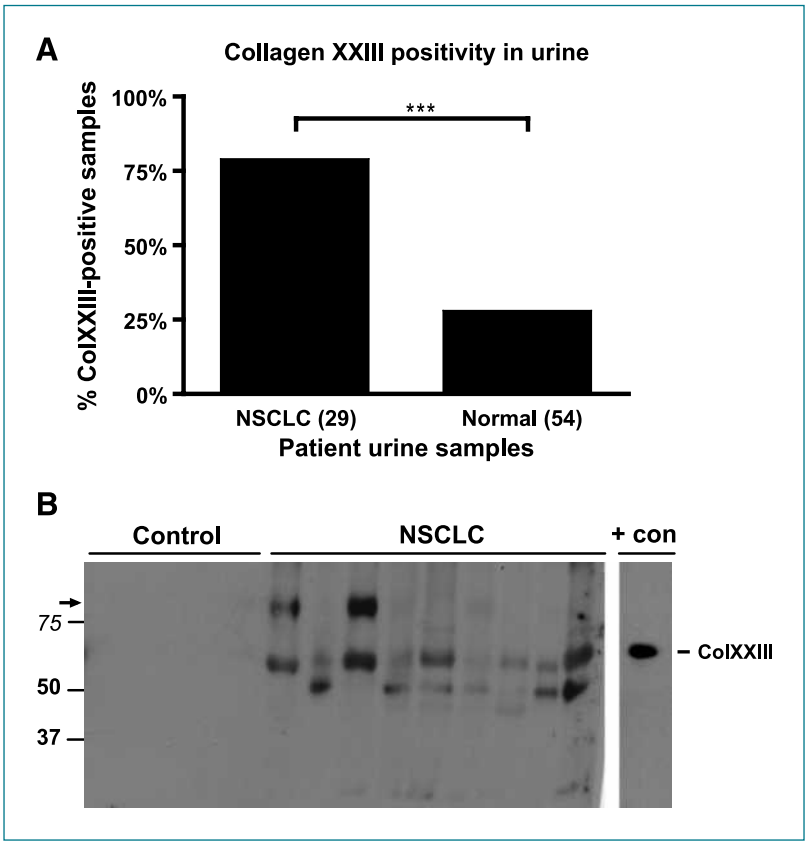

Figure 3. Detection of urinary collagen XXIII in human urine samples. A, collagen XXIII is detected in urine from 23 of 29 (79%) NSCLC patients but only in 15 of 54 (28%) control urine samples (Fisher's Exact test; ***, $P < 0.0001$). B, representative Western blot showing collagen XXIII positivity in a subset of urine samples using the 468642 antibody. Recombinant collagen XXIII corresponding to the cleaved ectodomain was used as a positive control (+ con) and is the predominant form of urinary collagen XXIII. Arrow, potential nonreduced protein complexes.

Collagen XXIII positivity in urine samples from NSCLC patients

Previous studies by Banyard et al. (12) have shown that the cleaved form of collagen XXIII is detectable in urine samples from prostate cancer patients and that levels decrease after prostatectomy. To determine whether collagen XXIII was also present in urine from NSCLC patients, Western blot analysis was done on 50 μg of processed patient urine. Urinary collagen XXIII was observed as two bands with approximate molecular weights of 60,000 and 50,000, corresponding to the furin-cleaved ectodomain of collagen XXIII and a lower molecular mass collagen XXIII fragment. Urinary collagen XXIII was detected in 23 of 29 NSCLC (79%) patients but only in 15 of 54 control (28%) patients ($P < 0.001$, Fisher's exact test; Supplementary Fig. S6; Fig. 3A and B) with a sensitivity of 79% and a specificity of 72%. These results revealed a significant association between urinary collagen XXIII positivity and presence of NSCLC.

Discussion

Collagen XXIII was recently proposed as a biomarker of prostate cancer progression and metastasis and is detectable in prostate cancer patient tissue and urine (12). We hypothesized that collagen XXIII expression was not limited to prostate cancer and that it may be a potential biomarker for other cancers. Collagen XXIII was expressed in several carcinoma tissues examined using a multicancer TMA, but was not readily detectable in sarcomas. In normal tissue, collagen XXIII expression is restricted. It is present in extracts from the brain, lung, kidney, cornea, and other epithelial structures but not in the tissue stroma (15). In normal lung tissue, collagen XXIII is rarely detected in the cells comprising the alveolar sacs but can be detected in the bronchiolar epithelial cells.[8] These observations suggest that collagen XXIII expression in carcinomas can occur in multiple ways: either collagen XXIII is already present in neoplastic precursor cells or collagen XXIII expression is activated in response to cellular alterations that occur during tumor initiation and progression. Collagen XXIII was detected in the majority of NSCLC tissues but was not detected in the majority of SCLC tissues sampled. Differences in the cellular origin and molecular pathogenesis of these lung cancers could potentially explain this disparity (16). Expression of the pro-protein convertase furin, which cleaves collagen XXIII (11, 17), is also increased in NSCLC compared with SCLC (18).

Traditionally, lung carcinomas are diagnosed using a multimodality approach including imaging (19), fluorescence

[8] Unpublished data.

bronchoscopy (20), or radiology and histopathologic examination of the tumor. We examined the utility of collagen XXIII presence in tumor tissue as a surrogate marker for NSCLC and we showed a significant correlation between collagen XXIII positivity and the presence of NSCLC in three different patient cohorts. Collagen XXIII does not distinguish between different classes of NSCLC, as positive staining was observed in 88% of lung adenocarcinomas and 73% of SqCCs. Markers such as thyroid transcription factor (TTF-1), cytokeratin 5/6 (CK5/6), and p63 (21-23) have been used previously to distinguish between SqCC and adenocarcinoma. Instead, we evaluated the potential of collagen XXIII as a tissue and circulating molecular correlate of NSCLC, as has been reported for cytokeratin-19, carcinoembryonic antigen, c-met, and heterogeneous nuclear ribonucleoprotein, which are overexpressed in NSCLC tissues, respectively, and are detected also in peripheral blood (24). By setting a threshold value for collagen XXIII staining intensity, we were able to distinguish between NSCLC and normal tissue with a sensitivity of 80% and a specificity of 81%.

The need for predictive and prognostic biomarkers for NSCLC is apparent from the low 5-year survival rate in NSCLC. Although pathologic staging and histologic classification are important determinants of treatment strategy, concurrent analysis of predictive biomarkers such as epidermal growth factor receptor, KRAS and DNA excision repair protein ERCC-1 (25-27), and prognostic biomarkers such as matrix metalloproteinase (MMP)-2 (28) and miR-34a (29) can be used to further guide treatment decisions in NSCLC patients. We therefore examined collagen XXIII staining intensity as an indicator of disease recurrence, as observed previously with collagen XXIII in prostate cancer (12) and with collagen XVIII in NSCLC (30). We find that collagen XXIII is an informative biomarker for NSCLC disease status and propensity for recurrence. High collagen XXIII staining intensity correlates with shorter recurrence-free survival times in NSCLC patients. Patients with high collagen XXIII staining may consequently benefit from more aggressive treatment strategies such as chemotherapy after surgery—even for patients with early-stage disease. Concurrent examination of existing biomarkers in combination with collagen XXIII could aid in the implementation of improved treatment strategies.

Although molecular changes in tumor tissue can provide important information about disease progression and outcome, access to such material is limited, as tumor resections and tissue biopsies are highly invasive procedures with an inherent risk of pneumothorax. Although examination of tissue biomarkers can initially assist with NSCLC diagnosis and guide treatment decisions, noninvasive procedures are essential for screening and for posttreatment monitoring. Biomarkers present in biological fluids, such as urine and serum, are easy to access noninvasively and measure repeatedly. Urinary biomarkers have additional advantages. First, urine collection does not require trained personnel or invasive sample collection, which represents a potential barrier to patient compliance—especially during extended patient monitoring. Second, the serum proteome is quite complex and includes abundant proteins, such as albumin, which may impair the detection of concentration changes in less abundant proteins. As detection methods have improved, larger and more highly charged proteins are increasingly found in urine samples. For example, urinary MMP-9 (92 kDa) and MMP-2 (72 kDa) are useful for distinguishing patients with localized cancer from patients without cancer (31). In the same report, Moses et al. (31) showed that urinary MMP activity did not correlate with creatinine levels, suggesting that large proteins in the urine are not always attributable to kidney dysfunction.

We have presented evidence that collagen XXIII is present in urine samples from NSCLC patients and can discriminate between the majority of control and NSCLC patients with sensitivity of 79% and specificity of 72%. For comparison, a panel of the four proteins in serum, carcinoembryonic antigen, retinol-binding protein, α1-antitrypsin, and SqCC antigen, can distinguish between the majority of control and lung cancer patients with sensitivities ranging from 77% to 89% and specificities ranging from 75% to 84% or 75.4% depending on the patient cohort (32). Other fluid biomarkers indicative of NSCLC risk or disease have also been reported. Increased levels of urinary 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone metabolites are found in smokers with increased risk of lung cancer (33), whereas dihydrodiol dehydrogenase levels are increased in tissue and serum from NSCLC patients and secreted from cancer cell lines (34, 35). Due to the presence of urinary collagen XXIII in other cancers, such as prostate cancer, collagen XXIII could not be used singularly to identify NSCLC patients. Instead, it could be used as part of a biomarker panel for noninvasive detection of primary cancer or monitoring of previously treated NSCLC patients for recurrence—supplementing more expensive imaging modalities and potentially improving the sensitivity and specificity of currently used fluid biomarkers. Notably, the fluid biomarkers, carcinoembryonic antigen, CA-125, and tissue polypeptide–specific antigen, are used currently to detect disease recurrence in a subset of NSCLC patients despite their lack of specificity for lung cancer (36).

Collagen XXIII transcriptional and translational regulation remains incompletely understood, as only the posttranslational cleavage of collagen XXIII by furin has been characterized. Furin-mediated cleavage of collagen XXIII is the presumed mechanism by which collagen XXIII is released into biological fluids such as urine. In our study, collagen XXIII is expressed in ~80% of tissue and urine samples from patients with NSCLC, which is similar to the proportion of furin in lung SqCC tissue (37). These data suggest that increased furin production may be responsible for increased extracellular presence of collagen XXIII.

Disclosure of Potential Conflicts of Interest

B. Zetter is a consultant and has equity in Predictive Biosciences, which has licensed the rights to Collagen XXIII when used as a cancer biomarker.

Acknowledgments

We thank Jeanette Muniz for her help processing the urine samples, Dr. Tadafumi Hashimoto and Dr. Takeshi Iwatsubo for collagen XXV antibody.

Grant Support

Grants from the Department of Defense grant W81XWH-05-2-0027 (I.I. Wistuba); the Specialized Program of Research Excellence in Lung Cancer Grant P50CA70907 (I.I. Wistuba); and NIH grant no. CA90578, CA074386, and CA092824 (L.R. Chirieac).

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Received 10/19/2009; revised 02/24/2010; accepted 03/03/2010; published online 05/06/2010.

References

1. Cancer Facts & Figures. Atlanta: American Cancer Society; 2009.
2. World Health Organization classification of tumours. In: Travis WD, Brambilla E, Muller-Hemelink HK, Harris CC, editors. Pathology and Genetics of Tumous of the Lung, Pleura, Thymus and Heart. Lyon: IARC Press; 2004.
3. Sun S, Schiller JH, Gazdar AF. Lung cancer in never smokers-a different disease. Nat Rev Cancer 2007;7:778–90.
4. Humphrey LL, Teutsch S, Johnson M. Lung cancer screening with sputum cytologic examination, chest radiography, and computed tomography: an update for the U.S. Preventive Services Task Force. Ann Intern Med 2004;140:740–53.
5. Alberts WM. Diagnosis and management of lung cancer executive summary: ACCP evidence-based clinical practice guidelines (2nd Edition). Chest 2007;132:1–19S.
6. Sharma SV, Bell DW, Settleman J, Haber DA. Epidermal growth factor receptor mutations in lung cancer. Nat Rev Cancer 2007;7:169–81.
7. Riely GJ, Marks J, Pao W. KRAS mutations in non-small cell lung cancer. Proc Am Thorac Soc 2009;6:201–5.
8. Sawyers CL. The cancer biomarker problem. Nature 2008;452:548–52.
9. Chan HP, Lewis C, Thomas PS. Exhaled breath analysis: novel approach for early detection of lung cancer. Lung Cancer 2009;63:164–8.
10. Phillips M. Volatile organic compounds in breath as markers of lung cancer: a cross-sectional study. Lancet 1999;353:1930–3.
11. Banyard J, Bao L, Zetter BR. Type XXIII collagen, a new transmembrane collagen identified in metastatic tumor cells. J Biol Chem 2003;278:20989–94.
12. Banyard J, Bao L, Hofer MD, et al. Collagen XXIII expression is associated with prostate cancer recurrence and distant metastases. Clin Cancer Res 2007;13:2634–42.
13. Hashimoto T, Wakabayashi T, Watanabe A, et al. CLAC: a novel Alzheimer amyloid plaque component derived from a transmembrane precursor, CLAC-P/collagen type XXV. EMBO J 2002;21:1524–34.
14. Barletta JA, Perner S, Iafrate AJ, et al. Clinical significance of TTF-1 protein expression and TTF-1 Gene amplification in lung adenocarcinoma. J Cell Mol Med 2008;13:1977–86.
15. Koch M, Veit G, Stickler S, et al. Expression of type XXIII collagen mRNA and protein. J Biol Chem 2006;281:21546–57.
16. Wistuba II, Gazdar AF. Lung cancer preneoplasia. Annu Rev Pathol 2006;1:331–48.
17. Veit G, Zimina EP, Franzke CW, et al. Shedding of collagen XXIII is mediated by furin and depends on the plasma membrane microenvironment. J Biol Chem 2007;282:27424–35.
18. Mbikay M, Sirois F, Yao J, Seidah NG, Chretien M. Comparative analysis of expression of the proprotein convertases furin, PACE4, PC1 and PC2 in human lung tumours. Br J Cancer 1997;75:1509–14.
19. Mulshine JL. Current issues in lung cancer screening. Oncology (Williston Park) 2005;19:1724–30; discussion 30–1.
20. Lam S, MacAulay C, leRiche JC, Palcic B. Detection and localization of early lung cancer by fluorescence bronchoscopy. Cancer 2000;89:2468–73.
21. Ring BZ, Seitz RS, Beck RA, et al. A novel five-antibody immunohistochemical test for subclassification of lung carcinoma. Mod Pathol 2009;22:1032–43.
22. Kaufmann O, Fietze E, Mengs J, Dietel M. Value of p63 and cytokeratin 5/6 as immunohistochemical markers for the differential diagnosis of poorly differentiated and undifferentiated carcinomas. Am J Clin Pathol 2001;116:823–30.
23. Khayyata S, Yun S, Pasha T, et al. Value of P63 and CK5/6 in distinguishing squamous cell carcinoma from adenocarcinoma in lung fine-needle aspiration specimens. Diagn Cytopathol 2009;37:178–83.
24. Sheu CC, Chang MY, Chang HC, et al. Combined detection of CEA, CK-19 and c-met mRNAs in peripheral blood: a highly sensitive panel for potential molecular diagnosis of non-small cell lung cancer. Oncology 2006;70:203–11.
25. Subramanian J, Govindan R. Molecular genetics of lung cancer in people who have never smoked. Lancet Oncol 2008;9:676–82.
26. Shigematsu H, Gazdar AF. Somatic mutations of epidermal growth factor receptor signaling pathway in lung cancers. Int J Cancer 2006;118:257–62.
27. Olaussen KA, Mountzios G, Soria JC. ERCC1 as a risk stratifier in platinum-based chemotherapy for nonsmall-cell lung cancer. Curr Opin Pulm Med 2007;13:284–9.
28. Leinonen T, Pirinen R, Bohm J, Johansson R, Kosma VM. Increased expression of matrix metalloproteinase-2 (MMP-2) predicts tumour recurrence and unfavourable outcome in non-small cell lung cancer. Histol Histopathol 2008;23:693–700.
29. Gallardo E, Navarro A, Vinolas N, et al. miR-34a as a prognostic marker of relapse in surgically resected non-small-cell lung cancer. Carcinogenesis 2009;30:1903–9.
30. Iizasa T, Chang H, Suzuki M, et al. Overexpression of collagen XVIII is associated with poor outcome and elevated levels of circulating serum endostatin in non-small cell lung cancer. Clin Cancer Res 2004;10:5361–6.
31. Moses MA, Wiederschain D, Loughlin KR, Zurakowski D, Lamb CC, Freeman MR. Increased incidence of matrix metalloproteinases in urine of cancer patients. Cancer Res 1998;58:1395–9.
32. Patz EF, Jr., Campa MJ, Gottlin EB, Kusmartseva I, Guan XR, Herndon JE II. Panel of serum biomarkers for the diagnosis of lung cancer. J Clin Oncol 2007;25:5578–83.
33. Yuan JM, Koh WP, Murphy SE, et al. Urinary levels of tobacco-specific nitrosamine metabolites in relation to lung cancer development in two prospective cohorts of cigarette smokers. Cancer Res 2009;69:2990–5.
34. Huang LJ, Chen SX, Huang Y, et al. Proteomics-based identification of secreted protein dihydrodiol dehydrogenase as a novel serum markers of non-small cell lung cancer. Lung Cancer 2006;54:87–94.
35. Huang LJ, Chen SX, Luo WJ, Jiang HH, Zhang PF, Yi H. Proteomic analysis of secreted proteins of non-small cell lung cancer. Ai Zheng 2006;25:1361–7.
36. Cho WC. Potentially useful biomarkers for the diagnosis, treatment and prognosis of lung cancer. Biomed Pharmacother 2007;61:515–9.
37. Lopez de Cicco R, Bassi DE, Page R, Klein-Szanto AJ. Furin expression in squamous cell carcinomas of the oral cavity and other sites evaluated by tissue microarray technology. Acta Odontol Latinoam 2002;15:29–37.

Collagen XXIII: A Potential Biomarker for the Detection of Primary and Recurrent Non–Small Cell Lung Cancer

Kristin A. Spivey, Jacqueline Banyard, Luisa M. Solis, et al.

Cancer Epidemiol Biomarkers Prev 2010;19:1362-1372.