Published OnlineFirst October 19, 2012; DOI: 10.1158/1940-6207.CAPR-12-0290

Cancer Prevention Research

Research Article

See related commentary by Gomperts et al., p. 4

Characterizing the Molecular Spatial and Temporal Field of Injury in Early-Stage Smoker Non–Small Cell Lung Cancer Patients after Definitive Surgery by Expression Profiling

Humam Kadara[1], Li Shen[2], Junya Fujimoto[1], Pierre Saintigny[1], Chi-Wan Chow[1], Wenhua Lang[1], Zuoming Chu[1], Melinda Garcia[1], Mohamed Kabbout[1], You-Hong Fan[1], Carmen Behrens[1], Diane A. Liu[4], Li Mao[5], J. Jack Lee[4], Kathryn A. Gold[1], Jing Wang[2], Kevin R. Coombes[2], Edward S. Kim[1], Waun Ki Hong[1], and Ignacio I. Wistuba[1,3]

Abstract

Gene expression alterations in response to cigarette smoke have been characterized in normal-appearing bronchial epithelium of healthy smokers, and it has been suggested that adjacent histologically normal tissue displays tumor-associated molecular abnormalities. We sought to delineate the spatial and temporal molecular lung field of injury in smoker patients with early-stage non–small cell lung cancer (NSCLC; $n = 19$) who were accrued into a surveillance clinical trial for annual follow-up and bronchoscopies within 1 year after definitive surgery. Bronchial brushings and biopsies were obtained from six different sites in the lung at the time of inclusion in the study and at 12, 24, and 36 months after the first time point. Affymetrix Human Gene 1.0 ST arrays were used for whole-transcript expression profiling of airways ($n = 391$). Microarray analysis identified gene features ($n = 1{,}165$) that were nonuniform by site and differentially expressed between airways adjacent to tumors relative to more distant samples as well as those ($n = 1{,}395$) that were significantly altered with time up to 3 years. In addition, gene interaction networks mediated by phosphoinositide 3-kinase (PI3K) and extracellular signal–regulated kinase (ERK)1/2 were modulated in adjacent compared with contralateral airways and the latter network with time. Furthermore, phosphorylated AKT and ERK1/2 immunohistochemical expression were significantly increased with time (nuclear pAKT, $P = 0.03$; cytoplasmic pAKT, $P < 0.0001$; pERK1/2, $P = 0.02$) and elevated in adjacent compared with more distant airways (nuclear pAKT, $P = 0.04$; pERK1/2, $P = 0.03$). This study highlights spatial and temporal cancer–associated expression alterations in the molecular field of injury of patients with early-stage NSCLCs after definitive surgery that warrant further validation in independent studies. *Cancer Prev Res; 6(1); 8–17. ©2012 AACR.*

Introduction

Lung cancer, of which non–small cell lung cancer (NSCLC) comprises the majority, is the leading cause of cancer-related deaths in the United States and worldwide (1, 2). The high mortality of this disease is, in part, due to the late diagnosis of the majority of lung cancers after regional or distant spread of the malignancy (3). Recent data from the National Lung Screening Trial (4), indicating that screening increases early detection rates, are expected to augment the number of early-stage NSCLCs detected warranting the need for better clinical management of this growing subpopulation. Besides adjuvant therapy, there are no effective chemoprevention strategies for patients with early-stage NSCLCs (5) who comprise approximately 50% of all diagnosed cases and have a relatively high rate of relapse (3). Improved clinical management of early-stage NSCLCs is tightly linked to the identification of new effective early biomarkers that can guide potential chemoprevention strategies.

Most diagnosed NSCLC cases (85%) are attributable to cigarette smoking (6–8). Auerbach and colleagues found that cigarette smoke induces extensive histologic changes in the bronchial epithelia in the lungs of smokers and that premalignant lesions are widespread and multifocal throughout the respiratory epithelium, suggestive of a field effect (9). Many molecular abnormalities, such as LOH

Authors' Affiliations: Departments of [1]Thoracic/Head and Neck Medical Oncology, [2]Bioinformatics, [3]Pathology, and [4]Biostatistics, The University of Texas MD Anderson Cancer Center, Houston, Texas; and [5]School of Dentistry, The University of Maryland, Baltimore, Baltimore, Maryland

Note: Supplementary data for this article are available at Cancer Prevention Research Online (http://cancerprevres.aacrjournals.org/).

Current address for E.S. Kim: Department of Solid Tumor Oncology and Investigational Therapeutics, Levine Cancer Institute, Carolinas Healthcare System, Charlotte, North Carolina.

Corresponding Author: Ignacio I. Wistuba, Departments of Thoracic/Head and Neck Medical Oncology and Pathology, the University of Texas MD Anderson Cancer Center, Houston, TX 77030. Phone: 713-563-9184; Fax: 713-730-0309; E-mail: iiwistuba@mdanderson.org

doi: 10.1158/1940-6207.CAPR-12-0290

Published OnlineFirst October 19, 2012; DOI: 10.1158/1940-6207.CAPR-12-0290

(10–12), mutations in *TP53* (13), methylation of *p16* tumor suppressor, death-associated protein kinase (*DAPK*), and retinoic acid receptor 2 beta (*RAR-β2*), were detected in bronchial epithelia of cancer-free former smokers (14–18), some of which persist for many years after smoking cessation (15). More recently, global mRNA and miRNA expression profiles have been described in the normal-appearing bronchial epithelium of healthy smokers (19, 20) that are different from those in nonsmokers. Moreover, expression and pathway signatures have been derived from normal bronchial epithelium of smokers that exhibited diagnostic properties (21, 22). Molecular changes involving LOH of chromosomal regions 3p (*DDUT* and *FHIT* genes), 9p (*CDKN2A*), genomic instability (increased microsatellite repeats), and *p16* methylation have been shown in histologically normal epithelium in patients with squamous cell carcinoma and in the sequence of pathogenesis of the disease (14, 23, 24). Moreover, Nelson and colleagues showed that *KRAS* is also mutated in histologically normal lung tissue adjacent to lung tumors (25). Furthermore, Tang and colleagues found higher rates of *EGFR* mutations in adjacent normal bronchial epithelia (NBE; refs. 26, 27) suggestive of a potential localized field effect.

It is plausible to assume that understanding early molecular aberrations in histologically normal smoke-damaged airway epithelium of early-stage patients would serve as a critical first step toward identification of biomarkers that can guide lung cancer prevention strategies. However, the global molecular airway field of injury in patients with early-stage NSCLCs, in particular after definitive surgery, is unknown. In this study, we used transcript-level expression profiling coupled with gene interaction network analysis and immunohistochemical (IHC) analysis to characterize, in-depth, site- and time-dependent global molecular alterations in airways of smoker patients with early-stage NSCLCs.

Materials and Methods

Patient population and airway epithelial cell collection

Early-stage (I/II), current or former smoker patients with NSCLCs with at least a 10-pack-year smoking history and without evidence of disease after definitive surgery were recruited into the Vanguard phase II surveillance clinical trial (clinical trial number NCT00352391) within 1 year from time of surgery. Patients were accrued between 2004 and 2008 and underwent frequent testing including chest X-rays, computed tomographic (CT) scans, laboratory work, serologies, flexible bronchoscopies, and airway biopsy collections within 1 year from surgery (average, 6 months; range, 1–12; first time point), and at months 12, 24, and 36 following the first time point. Bronchoscopies were conducted using white light or both white light and autofluorescence modalities. Biopsies were obtained from all potential anatomic locations and time points per patient. In total, there were 324 evaluable airway biopsies. Histologic assessment was conducted to determine whether malignant changes will occur during the time period. Once patients have completed 3 years of testing, they were followed until the study is completed. Patients were composed of former ($n = 16$) and current ($n = 3$) smokers. One of the 3 current smoker patients quit smoking 6 months before the 24-month time point. The clinicopathologic variables of patients in this study are summarized in Table 1. The study was approved by the Institutional Review Boards, and all participants provided written informed consents.

Bronchial airway epithelial cells were obtained from 5 to 6 different sites (main carina, 4 airways from 4 lobes, and the bronchial stump or main stem bronchus adjacent to the originally resected tumor and lobe; Fig. 1) at each time point using an Olympus fiberoptic bronchoscope (Olympus America Inc.) and cytobrushes (Cellebrity Endoscopic Cytobrush, Boston Scientific). Patients ($n = 19$) with samples/specimens available for analysis that were obtained serially up to either 24 or 36 months and from at least 5 different sites at each time point ($n = 391$ airway samples) were selected for the study. Epithelial cell content was confirmed by cytokeratin staining which yielded a 90% epithelial cell content. Brushes were immediately placed in serum-free RPMI medium on ice, vortexed gently to disperse epithelia into the media, and then removed. Samples were then immediately centrifuged, and cell pellets were resuspended in 1 mL of PBS. About 500 μL of the sample was then again centrifuged, and the pellet resuspended in 500 μL β-mercaptoethanol containing RLT buffer (Qiagen), homogenized, and stored in −80°C until further processing. Total RNA was isolated using the RNeasy Mini Kit according to the manufacturer's instructions (Qiagen).

RNA processing for microarrays

Total RNA samples were preprocessed for subsequent hybridization to expression arrays using the WT-Ovation and Encore Biotin Module from NuGEN Technologies Inc. (San Carlos, CA) according to the manufacturer's instructions. Briefly, the WT-Ovation Pico RNA amplification system was used to generate amplified cDNA using 5 ng of starting RNA material. After formation of double stranded cDNA, DNA was amplified using the SPIA Amplification Method, a linear isothermal DNA amplification process developed by the vendor (NuGEN Technologies). The WT-Ovation Exon Module (NuGEN Technologies) was then used for generation of amplified sense strand cDNA (ST-cDNA) that is suitable for subsequent array analysis with the Affymetrix Human Gene 1.0 ST array platform (Affymetrix, Santa Clara, CA). Fragmented and biotin-labeled cDNA was then generated using the Encore Biotin Module (NuGEN Technologies) using 5 μg of amplified cDNA. Quality and size distribution of unfragmented SPIA-amplified cDNA and subsequent fragmented labeled cDNA were assessed by loading samples on an RNA 6000 Nano LabChip (Agilent) and analysis with Agilent Bioanalyzer 2000 (Agilent). No differences in quality were noted on the basis of the duration of sample storage.

Generation of microarray raw data and analysis

Fragmented and labeled cDNA (2.5 μg) were hybridized onto Human Gene 1.0 ST arrays according to the

Published OnlineFirst October 19, 2012; DOI: 10.1158/1940-6207.CAPR-12-0290

Table 1. Clinicopathologic features of patients with NSCLCs included in the study

Patient	Histology	Anatomic site[a]	Stage	Recurrence	Adjuvant chemotherapy	Gender/ age, y	Smoking status[b]	Pack-years	Years quit smoking[c]	Months to inclusion from surgery[d]
1	ADC	LUL	IIA	Yes	No	Male/81	Former	100	12.0	1
2	ADC	RLL	IB	No	No	Male/64	Former	27	19.1	5
3	ADC	RUL	IB	Yes	Yes	Male/58	Former	12	25.5	7
4	SCC	RLL	IB	Yes	Yes	Female/62	Former	60	7.6	7
5	ADC	RLL	IA	No	No	Female/60	Former	31	11.8	6
6	ADC	LLL	IA	Yes	Yes	Male/62	Former	40	5.8	5
7	ADC	RUL	IB	Yes	Yes	Male/61	Current	70	NA	10
8	SCC	LUL	IB	No	No	Male/64	Current	100	NA	5
9	ADC	RUL	IA	No	No	Male/69	Former	12	36.5	8
10	ADC	RUL	IB	No	No	Female/63	Former	18	8.6	5
11	ADC	RUL	IB	Yes	No	Female/53	Former	42	5.8	3
12	SCC	RUL	IA	No	No	Male/62	Former	68	0.4	5
13	SCC	LUL	IA	No	No	Male/70	Former	96	1.4	7
14	ADC	LUL	IA	Yes	Yes	Female/45	Former	12.5	0.9	12
15	ADC	RLL	IA	No	No	Female/57	Current	80	NA	4
16	ADC	LLL	IA	No	No	Male/65	Former	48	21.0	7
17	ADC	LLL	IA	No	No	Male/64	Former	92	5.1	7
18	ADC	RML	IB	No	No	Male/71	Former	102	2.3	2
19	ADC	LUL	IA	No	No	Female/66	Former	50	18.5	5

Abbreviations: ADC, adenocarcinoma; LLL, left lower lobe; LUL, left upper lobe; RLL, right lower lobe; RML, right middle lobe; RUL, right upper lobe; SCC, squamous cell carcinoma.

[a]Location of primary tumor in the lung.

[b]Smoking status at time of inclusion into the study (patient 7 quit smoking during the course of the study).

[c]Years from smoking cessation to time of inclusion into the study.

[d]Months elapsed from surgery to time of inclusion into the study/first bronchoscopy time point.

manufacturer's instructions (Affymetrix). Hybridization cocktails containing samples, control oligonucleotide and eukaryotic hybridization controls in addition to hybridization mixes, dimethyl sulfoxide and nuclease-free water were heat denatured at 99°C for 5 minutes, cooled to 45°C for 5 minutes, and finally centrifuged at maximum speed for 1 minute. After injecting 80 μL of the hybridization cocktails, arrays were incubated for 17 hours in a hybridization oven set to a temperature of 45°C with 60 rpm rotation. Arrays were washed, stained, and processed using Affymetrix GeneChip Fluidics Station 450 systems after which they were imaged using Affymetrix GeneChip Scanner 3000 7G for subsequent generation of raw data (*CEL files). Raw data were quantified using Robust Multichip Array (RMA) background correction, quantile normalization, and RMA probe-level models (RmaPlm) and summarization methods. MIAME compliant metadata, normalized expression values, and 391 CEL files were submitted to the Gene Expression Omnibus (GEO; samples GSM992943-GSM993345, series GSE40407). After data preprocessing and normalization, a $\log_2$ transformation was applied. Pathways and gene interaction network analyses were conducted using the commercially available software Ingenuity Pathways Analysis (IPA, Redwood City, CA). All details of the microarray analysis including pairwise analysis of adjacent and contralateral airways for patient clustering are included in the Supplementary Information and in the 4 Supplementary Sweave Reports accompanying the manuscript.

IHC analysis of airway biopsies

Immunohistochemistry was done on histologic sections of 4-μm formalin-fixed, paraffin-embedded tissue samples prepared by a tissue arrayer as described previously (28). IHC analysis was conducted using purified rabbit polyclonal primary antibodies raised against phospho-AKT(threonine308) (1:200 dilution, clone C31E5E, catalog number 2965) and phospho-ERK1/2(Thr202/Tyr204) (1:400 dilution, clone D13.14.4E, catalog number 4370; Cell Signaling Technology). Antigen retrieval was conducted using the Dako Target Retrieval System at a pH of 6 (Dako). Intrinsic peroxidase activity was blocked by 3% methanol and hydrogen peroxide for 15 minutes and serum-free protein block (Dako) was used for 7 minutes for blocking nonspecific antibody binding. Slides were then incubated with the antibodies against phospho-AKT and phospho-ERK1/2 at room temperature for 90 and 120 minutes, respectively. After 3 washes in TBS, slides were incubated for 30 minutes with Dako Envision+ Dual Link at room temperature.

Published OnlineFirst October 19, 2012; DOI: 10.1158/1940-6207.CAPR-12-0290

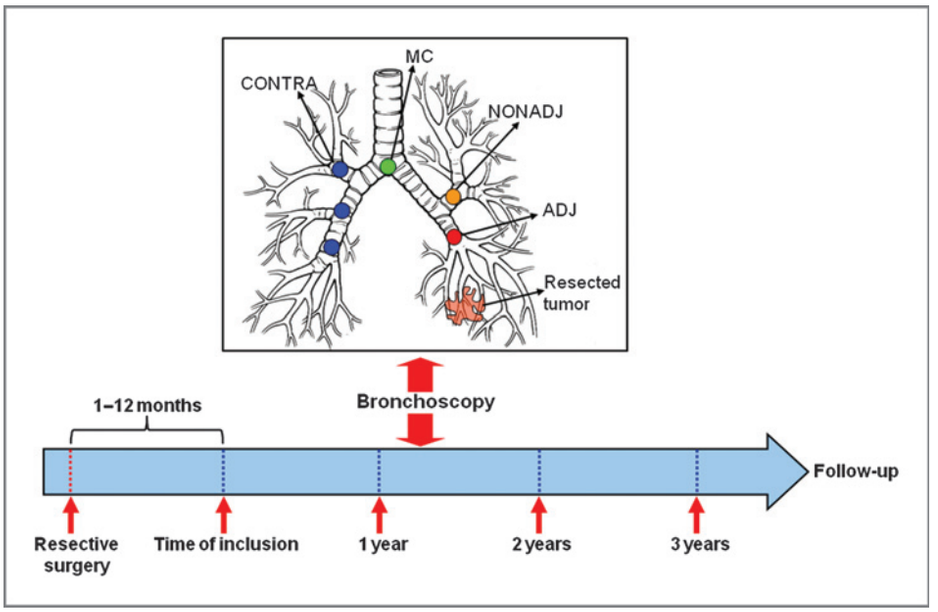

Figure 1. Spatial and temporal molecular field of injury in patients with early-stage NSCLCs after definitive surgery. Schematic depicting the site-dependent (top) and time-dependent (bottom) collection of airway epithelia brushings by bronchoscopy. Smoker patients with early-stage NSCLCs were enrolled into a surveillance clinical trial for annual follow-up and bronchoscopies within 1 year after definitive surgery. Bronchial airway epithelial cells (brushings and biopsies) were obtained from 5 to 6 different sites composed of the main carina (MC), 4 airways from 4 lobes ipsilateral (NON-ADJ), or contralateral (CONTRA) to the originally resected tumor and of the bronchial stump or main stem bronchus adjacent to the tumor and lobe. All site-different airway epithelia were collected at the time of inclusion in the study and at 12, 24, and 36 months following the starting time point (391 airways from 19 patients with NSCLCs).

Following 3 additional washes, slides were incubated with Dako chromogen substrate for 5 minutes and were counterstained with hematoxylin for another 5 minutes. Formalin-fixed, paraffin-embedded pellets from lung cancer cell lines displaying positive phospho-AKT and phospho-ERK1/2 expression by Western blot analysis were used as a positive control, whereas samples and whole-section tissue specimens processed similarly, except for the omission of the primary antibodies used as negative controls. The intensity and extent of cytoplasmic and nuclear phospho-AKT and phospho-ERK1/2 immunostaining were evaluated using a light microscope (magnification, $\times 20$) independently by 2 pathologists (J. Fujimoto and I.I. Wistuba) who were blinded to the identity of the samples. Cytoplasmic expression was quantified using a 4-value intensity score (0, none; 1, weak; 2, moderate; and 3, strong) and the percentage (0%–100%) of the extent of reactivity. A final cytoplasmic expression score was obtained by multiplying the intensity and reactivity extension values (range, 0–300). Nuclear expression score was quantified using the percentage of extent of reactivity (range, 0–100).

Summary statistics, including frequency tabulation, means, SDs, median, and range, were given to describe subject characteristics and IHC protein expression. Repeated measures analysis was conducted to assess the differential effect of sites on phosphorylated AKT and extracellular signal–regulated kinase (ERK)1/2 expression variation with time. Mixed-effects models were generated to assess significance of site, time, and the interaction of both factors to the expression of both proteins. All statistical tests were 2-sided and $P \leq 0.05$ was considered to be statistically significant. Statistical analysis was conducted with standard statistical software, including SAS Release 8.1 (SAS Institute) and S-Plus 2000 (Mathsoft Inc.).

Results

Detailed site- and time-dependent airway sampling of the field of injury in early-stage NSCLC patients after definitive surgery

Expression patterns molecularly exemplifying the impact of smoking on the airway epithelium of cancer-free individuals have been characterized (8, 19, 29). Moreover, molecular abnormalities typically found in lung tumors have been detected in normal resected margins suggestive of a field effect (15, 25–27, 30, 31). However, the biologic nature of the field of injury in particular after complete removal of the tumor in patients with early-stage NSCLCs, who are increasing in number and for whom there are no chemoprevention strategies, are yet unknown. Smoker patients with early-stage NSCLCs were recruited on a prospective phase II surveillance clinical trial that included frequent CT scans, serologies, and annual bronchoscopies in which airway brushings and biopsies were obtained within 1 year following tumor resection and at 12, 24, and 36 months following the first time point (Fig. 1). The first time point bronchoscopies were all conducted within 1 year of definitive surgery (average, 6 months; range, 1–12). Nineteen patients were selected for the study (Table 1) on the basis of airway sampling of at least 5 different sites per time point

Published OnlineFirst October 19, 2012; DOI: 10.1158/1940-6207.CAPR-12-0290

and continuously up to 24 or 36 months giving rise to 391 airway samples for transcript-level expressing profiling. The patients were accrued between 2004 and 2008 and were composed of former ($n = 16$) and current ($n = 3$) smokers. Brushings and biopsies were obtained from the main carina, airways relatively adjacent to the previously resected tumor and from airways more distant from the tumor in the ipsilateral and contralateral lung (Fig. 1).

Following normalization of the raw expression data, a mixed-effects model was applied to characterize the expression pattern of each gene that incorporated fixed-effects such as the site from the tumor and time after surgery of the collected airway samples (Supplementary Information). Histogram *P* value distribution plots after fitting beta-uniform mixture (BUM) models (32) for derived *P* values on the basis of the fixed-effects suggested that both site and time of the airway samples influenced gene expression modulation (Supplementary Fig. S1A and S1B).

Site-dependent differential expression patterns in airway epithelia of early-stage NSCLC patients

We first sought to determine whether gene expression profiles are differentially expressed in airways by site from the tumor including those relatively adjacent to the resected tumors compared with more distant airways. On the basis of the generated BUM models and *P* value distributions, genes differentially expressed by site were selected on the basis of a 1% false discovery rate (FDR). We identified 1,165 gene features that were statistically significantly differentially expressed by site (Supplementary Table S1 and Supplementary Sweave Report S2). Two-dimensional hierarchical clustering showed that the airway samples were divided into 2 main branches or clusters (Fig. 2A) on the basis of expression of the genes. Moreover, the left cluster in the indicated heatmap's dendogram contained a statistically significantly higher number of adjacent airway samples than the right branch which contained a significantly higher proportion of main carinas and contralateral to the tumor airways ($P = 0.0027$ of the Fisher exact test for count data). In addition, the 2-dimensional clustering revealed 8 different gene expression patterns which are indicated by the color bar and code along the left side of the heatmap (Fig. 2A). Some of the different gene clusters or classes were associated with a specific group of airway samples. Notably, a cluster of 263 genes (cluster 1, Supplementary Table S1), indicated by the dark green color and asterisk on the heatmap, was found to have highest average expression in adjacent airways (Fig. 2A and Supplementary Table S2). In contrast, another cluster of 348 genes (indicated by magenta color) appeared to be highest in expression in main carinas (Fig. 2A).

We then determined to functionally analyze the cluster of genes ($n = 263$) that was found to exhibit the highest average expression in adjacent airways between adjacent and contralateral airways. Functional pathways analysis using IPA depicted several significantly modulated

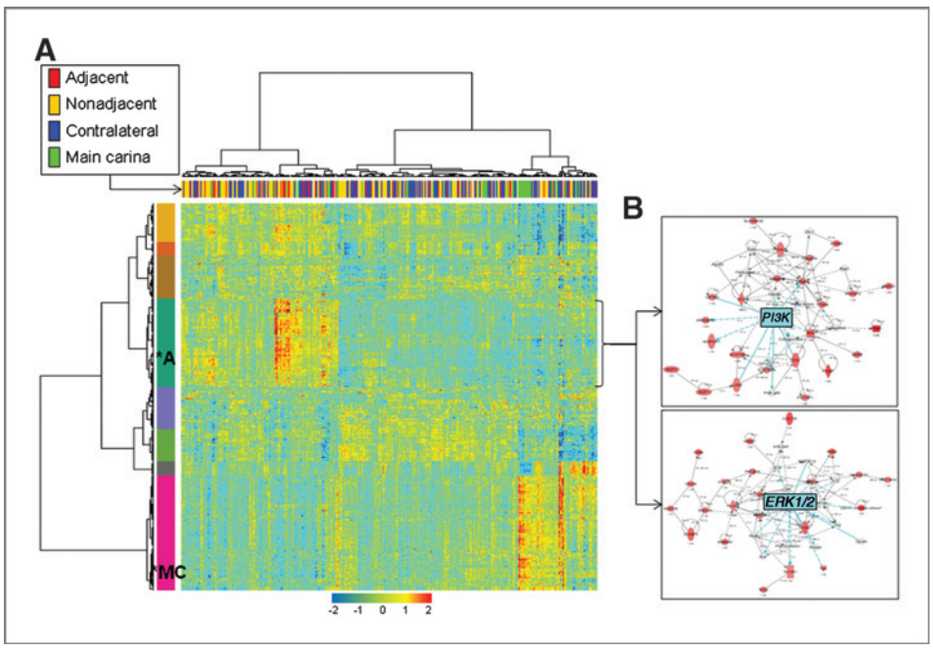

Figure 2. Site-dependent airway epithelia differential gene expression patterns. A, heatmap depicting 2-dimensional clustering of airway samples ($n = 391$) and genes ($n = 1,165$) that were determined to be differentially expressed by site in the mixed-effects model on the basis of a 1% FDR cutoff. The identified 8 gene clusters are labeled with different colors with the green cluster of genes ($n = 263$, *A) exhibiting highest average expression in adjacent airways and the magenta cluster ($n = 348$, *MC) having highest expression in main carinas. B, gene interaction analysis by IPA depicting networks with significant scores that indicate the likelihood of genes in a network being found together than due to chance. The depicted networks were found to be mediated by *PI3K* (top) and *ERK1/2* (bottom) with both kinases themselves not modulated in expression. Gene expression variation based on the statistical cutoff described above is depicted by color in the network (red, upregulated; green, downregulated).

Published OnlineFirst October 19, 2012; DOI: 10.1158/1940-6207.CAPR-12-0290

pathways and molecular functions in particular, T cell ($P = 1.2 \times 10^{-9}$), chemokine C-C motif receptor 5 (CCR5; $P = 1.5 \times 10^{-9}$), and phospholipase C signaling ($P = 3.6 \times 10^{-9}$). In addition, topological gene interaction network analysis highlighted functionally modulated and upregulated gene networks mediated by phosphoinositide 3-kinase (*PI3K*) and *ERK* in the adjacent airways (Fig. 2B). It is worthwhile to note that we also observed increased modulation of *ERK*/mitogen-activated protein kinase (*MAPK*)-mediated gene interaction network using a different analytic method where we compared expression profiles between adjacent and contralateral airways between patients in a pairwise fashion (data not shown, Supplementary Sweave Report S3). These findings suggest that airway site-dependent differential gene expression profiles in patients with early-stage NSCLCs exhibit increased molecular features and gene interaction networks typically associated with cancers.

We then determined to analyze gene expression profiles while excluding main carinas because their epithelial anatomy is suggested to be different from that of other airways (33) and thus may confound site-dependent observations. Following exclusion of main carinas, we found a reduced number of genes ($n = 136$) that were significantly modulated by site in the field of injury (Supplementary Sweave Report S4). Two-dimensional hierarchical clustering showed that airway samples were divided into 2 main branches or clusters (Supplementary Fig. S2) with significantly more adjacent airways in the right branch ($P = 0.0002$ of the Fisher exact t test). Moreover, the differentially expressed genes (Supplementary Table S4) were composed of 2 main subgroups with one cluster (top cluster) of 113 genes exhibiting highest average expression in adjacent airways. It is important to note that when we cross-compared gene clusters that we had found to exhibit the highest average expression in adjacent airways when including (263-gene cluster) or excluding (113-gene cluster) main carinas, we found a highly significant overlap ($P = 2.46 \times 10^{-191}$) in the number of genes ($n = 96$). Moreover, the site-dependent genes identified after exclusion of main carinas were also topologically organized following functional pathways analysis into interaction networks involving PI3K and ERK.

Gene expression profiles in the lung airway epithelia of early-stage NSCLC patients are modulated with time following definitive surgery

We then determined to identify genes that were differentially expressed with time. On the basis of the generated BUM models and P value distributions, time-dependent differentially expressed genes were identified and selected on the basis of a 5% FDR cutoff ($n = 1{,}395$; Supplementary Table S3 and Supplementary Sweave Report S2). Hierarchical clustering of samples indicated that the dendogram's main branches were statistically significantly unbalanced with respect to time; the main left branch comprised a higher number of 24- and 36-month time points than the right cluster ($P = 4.15 \times 10^{-7}$ of the Fisher exact test for count data). Two-dimensional clustering of both genes and samples revealed 2 main classes of genes: those that displayed increased (upper cluster) and decreased (lower cluster) expression with time (Fig. 3A). Functional pathways and gene interaction network analysis of genes differentially expressed between 36 months and the first time point revealed statistically significantly modulated pathways, in particular protein ubiquitination (5.3×10^{-5}), glutathione metabolism (6.0×10^{-5}), mitochondrial dysfunction (1.4×10^{-4}), and oxidative phosphorylation (2.9×10^{-3}) as well as eukaryotic initiation factor 2 (eIF2) signaling (2.6×10^{-3}). In addition, network analysis highlighted functionally modulated and elevated gene interaction

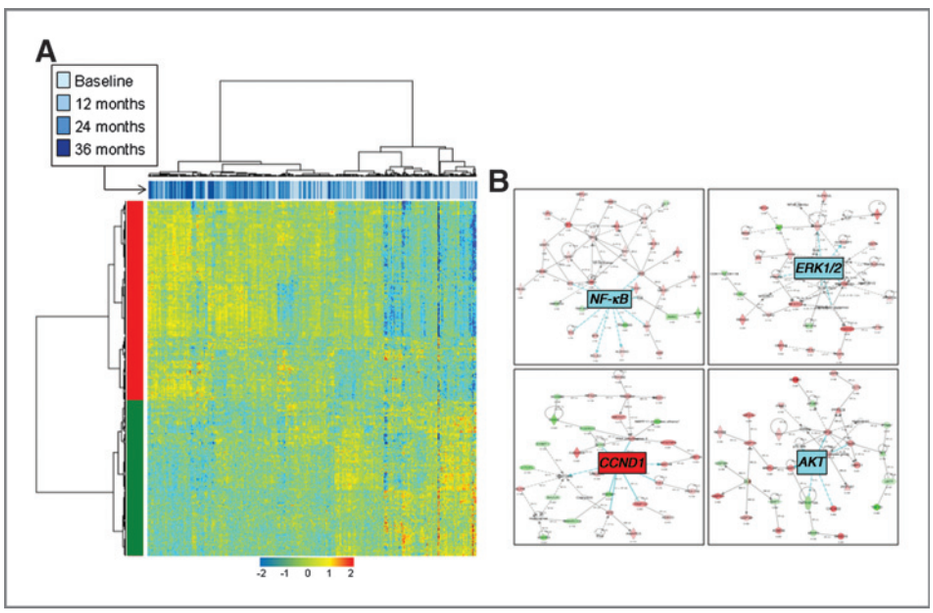

Figure 3. Temporal modulation of the molecular field of injury after definitive surgery in patients with early-stage NSCLCs. A, heatmap depicting 2-dimensional clustering of airway samples ($n = 391$) and genes ($n = 1{,}395$) that were determined to be differentially expressed by time in the mixed-effects model on the basis of a 5% FDR cutoff. B, gene interaction analysis, similar to that in Fig. 2, by IPA depicting networks with increased likelihood of genes being found together than due to chance and mediated by *NF-κB*, *ERK1/2*, *AKT*, and *CCND1*. *CCND1* itself was upregulated at the expression level. Gene expression variation based on the statistical cutoff is depicted by color in the network (red, upregulated; green, downregulated).

Published OnlineFirst October 19, 2012; DOI: 10.1158/1940-6207.CAPR-12-0290

networks with time in particular those mediated by *NF-κB, ERK, AKT*, and cyclin-B1 (*CCNB1*; Fig. 3B).

We then sought to determine the relationship of genes that were significantly modulated by site and time in the molecular field of injury. A smooth scatter plot of transformed *P* values indicated that site- and time-dependent expression modulations were largely independent (Supplementary Fig. S3 and Supplementary Sweave Report S2). We then cross-compared the site-dependent ($n = 1{,}165$) and time-dependent ($n = 1{,}395$) profiles that we had noted in the molecular field of injury. Using hypergeometric tests for overlapping probability, we found no significant overlap between genes we had determined to be differentially expressed by site and time in the molecular field of injury ($P = 0.865$; Supplementary Sweave Report S4).

Increased expression of phosphorylated AKT and ERK1/2 in airway epithelia by site from the tumor and with time following surgery

Our findings on the modulation of PI3K- and ERK1/2-mediated networks by site and time after surgery prompted us to examine the IHC expression of surrogate markers of both signaling cascades in corresponding formalin-fixed, paraffin-embedded airway biopsies. We sought to examine expression of phosphorylated AKT at Threonine308 because phosphorylation of this amino acid is well known to occur through phosphoinositide-dependent kinase 1 (PDPK1) following PI3K activation (34). We assessed by immunohistochemistry the IHC expression of phospho-AKT (Thr308) and phospho-ERK1/2(Thr202/Tyr204) in available and evaluable histologically NBE biopsies ($n = 324$) corresponding to the brushings analyzed by expression profiling. Immunoreactivity of phospho-AKT (minimum, 0; maximum, 300) was variable as depicted in the representative photomicrographs in Fig. 4A and was detected in both the cytoplasm and nucleus of NBE (Fig. 4A). IHC analysis showed that cytoplasmic ($P < 0.0001$) and nuclear ($P = 0.01$) phospho-AKT statistically significantly increased with time up to 3 years in all NBE (Fig. 4B) with highest expression at the 36-month time point. Nuclear phosphorylated AKT was also statistically significantly increased in adjacent NBE compared with airways more distant from tumors in the mixed-effects model ($P = 0.04$; Fig. 4B).

Immunoreactivity of phospho-ERK1/2 was also variable (minimum, 0; maximum, 209) and mainly cytoplasmic (Fig. 4C). IHC analysis showed that phosphorylated ERK1/2 was statistically significantly elevated in adjacent NBE ($P = 0.03$) and significantly increased with time up to 3 years in all airways when averaged together ($P = 0.02$; Fig. 4D) in the model. Notably, there was a significant interaction ($P = 0.019$) between the site of NBE and the time of sampling, as phospho-ERK1/2 expression was significantly increased with time in adjacent NBE but not in contralateral airways and main carinas in the model (Fig. 4D) with highest

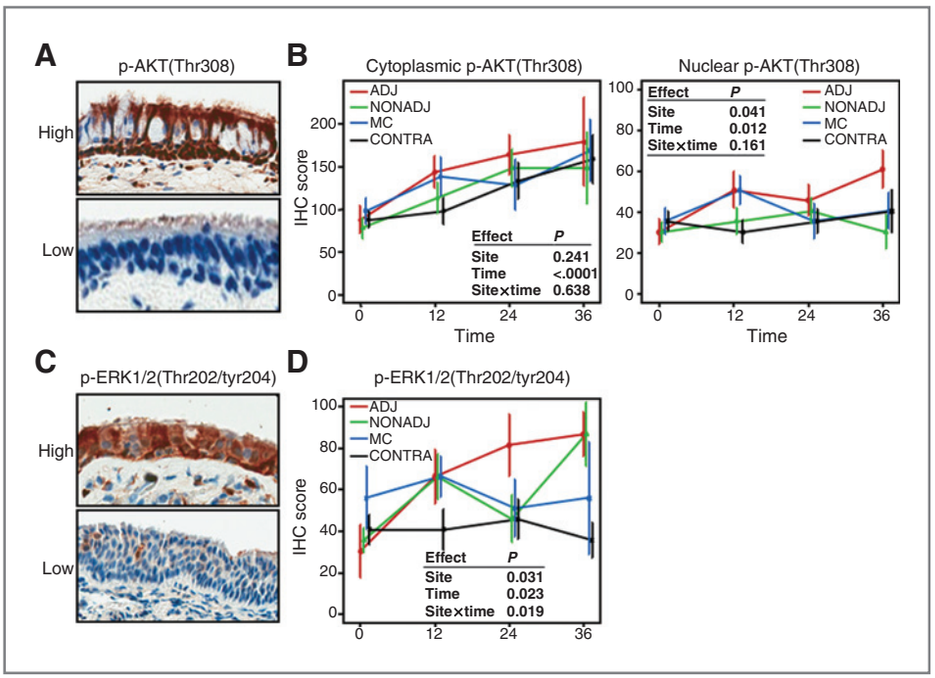

Figure 4. Site- and time-dependent IHC expression of phosphorylated AKT and ERK1/2 in the airway field of injury. A, representative photomicrographs (magnification, ×20) depicting strong (top) and weak (bottom) phospho-AKT(Thr308) immunostaining. B, IHC scores of cytoplasmic (left) and nuclear phospho-AKT (middle) were assessed for statistically significant differences by site and time in a mixed-effect model and plotted in main carinas (MC) and in adjacent (ADJ), nonadjacent (ipsilateral, NON-ADJ), and contralateral (CONTRA) airways with time. C, representative photomicrographs (magnification, ×20) depicting strong (top) and weak (bottom) phospho-ERK1/2(Thr202/Tyr204) immunostaining. D, IHC scores of phosphorylated ERK1/2 levels were assessed for statistically significant differences by site and time in a mixed-effect model and repeated measures analysis, site × time, term for interaction between site and time factors. Error bars represent SEM.

Published OnlineFirst October 19, 2012; DOI: 10.1158/1940-6207.CAPR-12-0290

expression in adjacent airways and nonadjacent (ipsilateral, green plot) airways observed at the latest time point. Similar data were obtained when we excluded main carinas in the mixed-effects model (data not shown). Moreover, we noted similar findings when all airway samples were analyzed irrespective of the presence of preneoplastic lesions (e.g., dysplasias).

These data showed that like differential gene expression profiles within the lung airway field of injury, canonical activated oncogenes are modulated by site from the resected tumor and time following definitive surgery in patients with early-stage NSCLCs.

Discussion

In this report, we characterized differential expression profiles and protein expression within the lung airway field of injury of patients with early-stage NSCLCs by site from the tumor and time in years following surgery. We showed, and to our knowledge for the first time, that gene expression profiles in histologically normal airways of patients with early-stage NSCLCs are nonuniform by site and are modulated with time up to 3 years following surgery. Moreover, functional analysis of the expression profiles showed that canonical expression patterns and protein kinase activation, typical of tumors, are increased in airway sites adjacent to tumors as well as remain or are further modulated in the lung airway field of injury for three years after definitive surgery. In particular, phosphorylated ERK/MAPK and AKT expression were upregulated in NBE with time and by site from tumors. In light of the growing subpopulation of early-stage NSCLCs, our findings are, in part, proof-of-principle and raise the intriguing possibility of the importance of intense surveillance and molecular characterization of the remaining smoking-injured airway epithelia and its potential integration in the future into clinical practice and management of early-stage disease. However, it is noteworthy that our study's patient cohort, despite its uniqueness in which expression profiling was serially conducted on airways from multiple sites collected during bronchoscopies for 36 months following surgery, is of limited size. Moreover, the reported findings warrant the need for validation or confirmation in independent larger sets.

There is a large body of evidence that patients who have survived an upper aerodigestive cancer comprise a high-risk population that may be targeted for early detection and chemoprevention efforts (5, 31). Currently, there are no established adjuvant treatments in the tertiary prevention setting of patients with early-stage NSCLCs. It has been suggested that failures in advances of chemoprevention are, in part, due to the lack of clear and specific molecular targets (5, 35). Our extensive profiling of the molecular field of injury in patients with NSCLCs identified cancer-associated pathways (PI3K and ERK) aberrantly regulated in NBE of patients with NSCLCs after tumor removal. In this context, it is plausible to suggest that a thorough characterization of the molecular field of injury in patients with early-stage NSCLCs can aid in identification of aberrantly expressed pathways, for example, PI3K and ERK, which could potentially serve as suitable targets for chemoprevention. However, it cannot be neglected that the alternative hypothesis can counter argue that activation of PI3K and ERK/MAPK pathways may be beneficial for chemoprevention as markers of both pathways (phosphorylated AKT and ERK) increased following surgical tumor resection. Our suggestion that such pathways may serve as chemoprevention targets should be interpreted cautiously and is presented owing to the known promalignant function of the pathways and gene networks highlighted in our analysis (36).

The first time point brushings in this study were obtained within 1 year of patient accrual and not at the time when the tumor was still *in situ*. The variable starting time point from surgery among patients is a limiting factor in our analysis as it is plausible to assume that the molecular field of injury may vary upon tumor removal. We were not able to avoid this caveat given the difficulty of accrual of patients following surgical tumor resection to obtain bronchial brushings at 6 different anatomical sites in the bronchial tree annually up to 3 years. However, it is important to mention that the time-dependent gene expression profiles we had identified, despite not incorporating the molecular field effect at the time the tumor was present, showed gradual changes in expression with time. This effect was also evident when analyzing the IHC expression of phosphorylated AKT and ERK in corresponding bronchial biopsies with highest expression at 3 years. Our findings warrant future larger studies in which the molecular field of injury at the time the tumor is still *in situ* can also be serially monitored for several years.

The patient population we had studied was composed of patients with early-stage NSCLCs, which is in contrast to earlier transcriptomic studies of the molecular field of injury that focused on phenotypically normal smokers and nonsmokers (19, 21, 22). It is still not clear whether the differences in expression described in this study reflect an already present gradient field of injury that may have contributed to tumor development in light of the differential cancer-associated pathways identified or one that arises due to the molecular impact of the tumor on the adjacent field. It is important to note that in this study, the spatial and temporal molecular field of injury in patients with lung cancer was profiled prospectively starting within 1 year following definitive surgery. Thus, the above speculation may be addressed by a similar thorough spatial and temporal characterization of the molecular field of injury before and after surgery in early-stage patients. In addition, we did not have access to similar type of brushings from cancer-free individuals such as high-risk smokers. Similar analysis of the molecular field of injury in cancer-free individuals will shed light on the nature of site- and time-dependent expression patterns in the field of injury and whether such changes in patients with cancer reflect recovery from surgery (temporal profiles) or are a cause or consequence of tumor development in the adjacent field (spatial profiles).

In conclusion, our unique study identified gene expression profiles, functional gene networks, and activated levels

Published OnlineFirst October 19, 2012; DOI: 10.1158/1940-6207.CAPR-12-0290

of oncogenic protein kinases within the field of injury of patients with early-stage NSCLCs that are modulated or increased in airways spatially from the tumor and temporally following surgery. Moreover, the herein previously uncharacterized airway cancer–associated expression and protein kinase alterations harbor potential valuable targets for chemoprevention and warrant confirmation and further studies in larger independent cohorts.

Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

Authors' Contributions

Conception and design: H. Kadara, L. Mao, E.S. Kim, I.I. Wistuba
Development of methodology: H. Kadara, L. Shen, L. Mao, I.I. Wistuba
Acquisition of data (provided animals, acquired and managed patients, provided facilities, etc.): H. Kadara, J. Fujimoto, P. Saintigny, M. Kabbout, Y.-H. Fan, C. Behrens, K.A. Gold
Analysis and interpretation of data (e.g., statistical analysis, biostatistics, computational analysis): H. Kadara, L. Shen, J. Fujimoto, M. Kabbout, D.A. Liu, J.J. Lee, J. Wang, K.R. Coombes, E.S. Kim, I.I. Wistuba
Writing, review, and/or revision of the manuscript: H. Kadara, J. Fujimoto, P. Saintigny, J.J. Lee, K.A. Gold, K.R. Coombes, W.K. Hong, I.I. Wistuba
Administrative, technical, or material support (i.e., reporting or organizing data, constructing databases): J. Fujimoto, C.-W. Chow, W. Lang, Z. Chu, E.S. Kim, I.I. Wistuba
Study supervision: H. Kadara, L. Mao, E.S. Kim, W.K. Hong, I.I. Wistuba

Grant Support

This study was funded in part by Department of Defense (DoD) grants W81XWH-04-1-0142 (to W.K. Hong and I.I. Wistuba) and W81XWH-10-1-1007 (to H. Kadara and I.I. Wistuba) and by Cancer Center Support Grant CA16672 (MD Anderson Cancer Center microarray core facility).

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 July 3, 2012; revised September 27, 2012; accepted September 28, 2012; published OnlineFirst October 19, 2012.

References

1. Jemal A, Bray F, Center MM, Ferlay J, Ward E, Forman D. Global cancer statistics. CA Cancer J Clin 2011;61:69–90.
2. Siegel R, Naishadham D, Jemal A. Cancer statistics, 2012. CA Cancer J Clin 2012;62:10–29.
3. Herbst RS, Heymach JV, Lippman SM. Lung Cancer. N Engl J Med 2008;359:1367–80.
4. Aberle DR, Adams AM, Berg CD, Black WC, Clapp JD, Fagerstrom RM, et al. Reduced lung-cancer mortality with low-dose computed tomographic screening. N Engl J Med 2011;365:395–409.
5. Gold KA, Kim ES, Lee JJ, Wistuba II, Farhangfar CJ, Hong WK. The BATTLE to personalize lung cancer prevention through reverse migration. Cancer Prev Res 2011;4:962–72.
6. Goldstraw P, Ball D, Jett JR, Le Chevalier T, Lim E, Nicholson AG, et al. Non-small-cell lung cancer. Lancet 2011;378:1727–40.
7. Gazdar AF, Thun MJ. Lung cancer, smoke exposure, and sex. J Clin Oncol 2007;25:469–71.
8. Steiling K, Ryan J, Brody JS, Spira A. The field of tissue injury in the lung and airway. Cancer Prev Res 2008;1:396–403.
9. Auerbach O, Stout AP, Hammond EC, Garfinkel L. Changes in bronchial epithelium in relation to cigarette smoking and in relation to lung cancer. N Engl J Med 1961;265:253–67.
10. Mao L, Lee JS, Kurie JM, Fan YH, Lippman SM, Lee JJ, et al. Clonal genetic alterations in the lungs of current and former smokers. J Natl Cancer Inst 1997;89:857–62.
11. Powell CA, Klares S, O'Connor G, Brody JS. Loss of heterozygosity in epithelial cells obtained by bronchial brushing: clinical utility in lung cancer. Clin Cancer Res 1999;5:2025–34.
12. Wistuba II, Lam S, Behrens C, Virmani AK, Fong KM, LeRiche J, et al. Molecular damage in the bronchial epithelium of current and former smokers. J Natl Cancer Inst 1997;89:1366–73.
13. Franklin WA, Gazdar AF, Haney J, Wistuba II, La Rosa FG, Kennedy T, et al. Widely dispersed p53 mutation in respiratory epithelium. A novel mechanism for field carcinogenesis. J Clin Invest 1997;100:2133–7.
14. Belinsky SA, Nikula KJ, Palmisano WA, Michels R, Saccomanno G, Gabrielson E, et al. Aberrant methylation of p16(INK4a) is an early event in lung cancer and a potential biomarker for early diagnosis. Proc Natl Acad Sci U S A 1998;95:11891–6.
15. Belinsky SA, Palmisano WA, Gilliland FD, Crooks LA, Divine KK, Winters SA, et al. Aberrant promoter methylation in bronchial epithelium and sputum from current and former smokers. Cancer Res 2002;62:2370–7.
16. Heller G, Zielinski CC, Zochbauer-Muller S. Lung cancer: from single-gene methylation to methylome profiling. Cancer Metastasis Rev 2010;29:95–107.
17. Soria JC, Rodriguez M, Liu DD, Lee JJ, Hong WK, Mao L. Aberrant promoter methylation of multiple genes in bronchial brush samples from former cigarette smokers. Cancer Res 2002;62:351–5.
18. Zochbauer-Muller S, Lam S, Toyooka S, Virmani AK, Toyooka KO, Seidl S, et al. Aberrant methylation of multiple genes in the upper aerodigestive tract epithelium of heavy smokers. Int J Cancer 2003;107:612–6.
19. Spira A, Beane J, Shah V, Liu G, Schembri F, Yang X, et al. Effects of cigarette smoke on the human airway epithelial cell transcriptome. Proc Natl Acad Sci U S A 2004;101:10143–8.
20. Schembri F, Sridhar S, Perdomo C, Gustafson AM, Zhang X, Ergun A, et al. MicroRNAs as modulators of smoking-induced gene expression changes in human airway epithelium. Proc Natl Acad Sci U S A 2009;106:2319–24.
21. Spira A, Beane JE, Shah V, Steiling K, Liu G, Schembri F, et al. Airway epithelial gene expression in the diagnostic evaluation of smokers with suspect lung cancer. Nat Med 2007;13:361–6.
22. Gustafson AM, Soldi R, Anderlind C, Scholand MB, Qian J, Zhang X, et al. Airway PI3K pathway activation is an early and reversible event in lung cancer development. Sci Transl Med 2010;2:26ra5.
23. Wistuba II, Behrens C, Virmani AK, Mele G, Milchgrub S, Girard L, et al. High resolution chromosome 3p allelotyping of human lung cancer and preneoplastic/preinvasive bronchial epithelium reveals multiple, discontinuous sites of 3p allele loss and three regions of frequent breakpoints. Cancer Res 2000;60:1949–60.
24. Wistuba II, Behrens C, Virmani AK, Milchgrub S, Syed S, Lam S, et al. Allelic losses at chromosome 8p21-23 are early and frequent events in the pathogenesis of lung cancer. Cancer Res 1999;59:1973–9.
25. Nelson MA, Wymer J, Clements N Jr. Detection of K-ras gene mutations in non-neoplastic lung tissue and lung cancers. Cancer Lett 1996;103:115–21.
26. Tang X, Shigematsu H, Bekele BN, Roth JA, Minna JD, Hong WK, et al. EGFR tyrosine kinase domain mutations are detected in histologically normal respiratory epithelium in lung cancer patients. Cancer Res 2005;65:7568–72.
27. Tang X, Varella-Garcia M, Xavier AC, Massarelli E, Ozburn N, Moran C, et al. Epidermal growth factor receptor abnormalities in the pathogenesis and progression of lung adenocarcinomas. Cancer Prev Res 2008;1:192–200.
28. Sun M, Behrens C, Feng L, Ozburn N, Tang X, Yin G, et al. HER family receptor abnormalities in lung cancer brain metastases and corresponding primary tumors. Clin Cancer Res 2009;15:4829–37.
29. Beane J, Sebastiani P, Liu G, Brody JS, Lenburg ME, Spira A. Reversible and permanent effects of tobacco smoke exposure on airway epithelial gene expression. Genome Biol 2007;8:R201.

Published OnlineFirst October 19, 2012; DOI: 10.1158/1940-6207.CAPR-12-0290

30. Wistuba II, Behrens C, Milchgrub S, Bryant D, Hung J, Minna JD, et al. Sequential molecular abnormalities are involved in the multistage development of squamous cell lung carcinoma. Oncogene 1999;18:643–50.
31. Wistuba II, Gazdar AF. Lung cancer preneoplasia. Annu Rev Pathol 2006;1:331–48.
32. Pounds S, Morris SW. Estimating the occurrence of false positives and false negatives in microarray studies by approximating and partitioning the empirical distribution of p-values. Bioinformatics 2003;19:1236–42.
33. Lips KS, Volk C, Schmitt BM, Pfeil U, Arndt P, Miska D, et al. Polyspecific cation transporters mediate luminal release of acetylcholine from bronchial epithelium. Am J Respir Cell Mol Biol 2005;33:79–88.
34. Engelman JA. Targeting PI3K signalling in cancer: opportunities, challenges and limitations. Nat Rev Cancer 2009;9:550–62.
35. Khuri FR. The dawn of a revolution in personalized lung cancer prevention. Cancer Prev Res 2011;4:949–53.
36. Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. Cell 2011;144:646–74.

Published OnlineFirst October 19, 2012; DOI: 10.1158/1940-6207.CAPR-12-0290

Cancer Prevention Research

Characterizing the Molecular Spatial and Temporal Field of Injury in Early-Stage Smoker Non –Small Cell Lung Cancer Patients after Definitive Surgery by Expression Profiling

Humam Kadara, Li Shen, Junya Fujimoto, et al.

Cancer Prev Res 2013;6:8-17. Published OnlineFirst October 19, 2012.