Laser Capture Microdissection and Microarray Expression Analysis of Lung Adenocarcinoma Reveals Tobacco Smoking- and Prognosis-related Molecular Profiles


Laboratory of Human Carcinogenesis [K. M., E. D. B., A. I. R., P. H., C. C. H.], Biometric Research Branch, Division of Cancer Treatment and Diagnosis [R. S.] and Laboratory of Pathology [L. C., I. A. I.], National Cancer Institute, and Genetics of Development and Disease Branch, National Institute of Diabetes and Digestive and Kidney Diseases [H. M.], NIH, Bethesda, Maryland 20892; EMMES Corporation, Rockville, Maryland 20850 [A. C. P.]; Department of Pathology and Program in Oncology, University of Maryland, Baltimore, Maryland 21201 [R. T. J.]; and Laboratory of Molecular Medicine, Human Genome Center, Institute of Medical Science, University of Tokyo, Tokyo 108-8639, Japan [To. K., Ta. K., Y. N.]

ABSTRACT

Recent expression profile analyses revealed that lung adenocarcinomas can be divided into several subgroups with diverse pathological features. Because cellular heterogeneity of tumours can confound these analyses, we used laser capture microdissection and microarray expression analysis to characterize the molecular profiles of lung adenocarcinomas. We found 45 genes delineating smokers and nonsmokers that were located at chromosomal loci frequently altered in non-small cell lung cancers, and 27 genes, which were differentially expressed between survivors and nonsurvivors 5 years after surgery. These results are consistent with the hypothesis that the abnormal expression of genes involved in maintaining the mitotic spindle checkpoint and genomic stability, e.g., hBUB3, hZW10, and APC2, contribute to the molecular pathogenesis and tumor progression of tobacco smoke-induced adenocarcinoma of the lung.

INTRODUCTION

Lung cancer is the leading cause of male and female cancer deaths in the United States. Lung carcinoma is classified usually as small-cell lung carcinoma or NSCLC (adenocarcinoma, squamous cell carcinoma, and large cell carcinoma). In particular, adenocarcinoma is the most common type of lung cancer in women and nonsmokers, and it is increasingly associated with tobacco smoking as well. Since the 1950s, steady rises in the incidence of adenocarcinoma of the lung have been observed in many developed countries with ethnically diverse populations (1). Recent reports (2, 3) used the gene expression profiling to divide lung adenocarcinomas into several subgroups and to discriminate primary cancers from metastases of extrapulmonary origin. Lung adenocarcinomas, when compared with squamous cell lung carcinomas or small cell lung carcinomas, show striking differences in the expression patterns (2, 3). In the analysis of clinical tumor specimens in these and other studies, a significant confounder is the cellular heterogeneity of normal and diseased tissue. To overcome this problem, LCM was developed to analyze clinical samples (4 –7). We used laser capture microdissection and microarray expression analysis to identify the genes differentially expressed in lung adenocarcinoma associated with the following clinical phenotypic subgroups: prognosis, smoking, and gender, and to generate the hypotheses concerning the molecular pathogenesis and tumor progression of lung carcinoma.

MATERIALS AND METHODS

Patients and Surgically Resected Frozen Tissues. The study was approved by the Institutional Review Boards of the University of Maryland, Baltimore, and the National Cancer Institute. Three pathologists diagnosed the tumor tissues with adenocarcinomas (8). We chose lung adenocarcinoma cases with tumor stages I or II by Tumor-Node-Metastasis classification to minimize any secondary or tertiary effects related to tumor stages III or IV. In Table 1, we summarized the characteristics of patients. Among the 19 patients with adenocarcinoma, 6 survived 5 years or more after surgery, 12 were nonsurvivors that died of a recurrence of lung cancer, and we could not find the exact cause of death for 1 patient. After the classical criteria, patients were defined as nonsmokers if they smoked <100 cigarettes in their life time. Five were nonsmokers, and 14 were smokers, and the pack-years among 14 smokers were 20–125 (mean ± SD; 48 ± 27). Ten were males and 9 were females.

RNA Extraction, LCM, and T7 Amplification. From each flash-frozen tumor sample, we prepared 8-μm thick frozen sections. Total RNA was extracted from one section, and the quality was evaluated (Fig. 1). Serial frozen sections (~18/case) were used for the subsequent analysis. We used the PiscCell II LCM System from Acturus Engineering (Mountain View, CA) for laser capture and followed the manufacturer’s protocol with several modifications. The sections were immersed in the relevant fixatives or staining solutions (70%, 95%, or 100% ethanol, hematoxylin, eosin, or RNase-free water) for ~10 s each, followed by dehydration with xylene for 1 min. We recorded the pathological images for each case (Fig. 1, C–M). We repeated reverse transcription and T7 amplification twice to get two-round aRNA as probes for microarray analysis as reported previously (5).

RNA Reference Pool. As an experimental control, we used the BEAS-2B cell line that was isolated from normal human bronchial epithelium obtained from an autopsy of a noncancerous individual, which was immortalized with an adenovirus 12-SV40 virus hybrid (9).

Microarray and Acquisition of Data. We used cDNA microarray slides generated by Amersham Pharmacia Biotech (Piscataway, NJ) and fabricated in Laboratory of Molecular Medicine, University of Tokyo. Four sets of array slides contained 18,432 cDNAs selected from the UniGene database with 15,737 unique genes. The cDNAs were amplified by reverse transcription-PCR without any repetitive or poly(A) sequences. The PCR products were spotted in duplicate using an Array Spotter Generation III (Amersham). Each slide contained 48 housekeeping genes. We labeled 2.5 μg of two-round aRNA from LCM-captured tissues or BEAS-2B by reverse transcription with Cy3-dCTP or Cy5-dCTP, respectively. Probes were hybridized to the microarrays in the Automated Slide Processor (Amersham), and signal intensities were quantified using ArrayVision software (Imaging Research, Inc., St. Catharines, Ontario, Canada). Four sets of signal intensities (Cy3 or Cy5, left panel or right panel) from duplicate spots were transformed to log2 scale, and the duplicate log-ratios were averaged. We excluded clones for which the Cy3-Cy5 ratios for duplicate spots differed by more than 2-fold. The log-ratios of nonexcluded clones were normalized by a median centering the 48 housekeeping genes on each array. We selected 6,216 clones with...
consistent duplicate measurements on at least 15 of the 19 arrays for additional analysis. The analyses were performed using BRB-ArrayTools.5

Hierarchical Clustering. We selected 2270 clones with consistent log-ratios present in at least 15 of the 19 samples, and which showed more variant expression patterns among the 19 samples. Then, we performed averaged linkage hierarchical clustering for 19 samples and 2270 clones with centered correlation as the distance metric (10).

Genes That Distinguish Phenotype Classes. Among 19 cancer cases, to identify genes that discriminate between two phenotype subclasses (e.g., smoker versus nonsmoker), the F test with 0.0025 as a nominal significance level (P) was computed on log2-expression ratios of each clone. Next, phenotype labels were randomly permuted among the 19 samples, the F statistic for each gene in the permuted data set was recomputed, and the number of clones significant at the nominal 0.0025 level was counted. This process was repeated 2000 times, and the proportion of the 2000 replications with at least as many genes significant at the 0.0025 level, provided a global test of the null hypothesis that the expression profile is unrelated to phenotype class.

RESULTS

Following the strategy described in Fig. 1A, we recorded the tumor pathology (Fig. 1C), and the number of sections and spots used for LCM, and calculated the yield of aRNA after two rounds of amplification (Fig. 1B). Within the context of hierarchical clustering among 19 samples and 2270 clones (Fig. 2), 4 of the adenocarcinomas (cases 3, 8, 9, and 15) that clustered separately were diffusely invasive types. In analyzing genetic alterations or expression profiles in cancer cells, the contamination of normal epithelia or stromal cells may confound the analysis, especially when using tissues such as lung, brain, kidney, ovary, mammary gland, or prostate (5, 7, 11, 12), which are unlike the more homogenous fluorescence-activated cell sorted blood cells (13). The use of LCM should improve the sample preparation for microarray expression analysis (4). One disadvantage of microdissection is that RNA amplification is essential for the preparation of an adequate amount of RNA for probe generation. The reproducibility and usefulness of T7 amplification in gene expression profiling has been reported previously (14, 15). In our hands, the gene expression profile from two independent two-rounds of amplification of BEAS-2B RNA was highly correlated (r2 = 0.91; data not shown).

Several of the genes with lower expression among smokers are located in the chromosomal regions where genomic imbalance in NSCLC has been observed previously with high frequency. In Table 2, we summarized the reported data of genomic imbalance including homozygous deletion, which was analyzed with comparative genomic hybridization, LOH analysis, and other methods. For example, chromosome 3p21.3 is a well-known region for a frequent homozygous deletion in lung cancers. Lerman and Minna (16) analyzed 25 genes localized in the 630-kb homozygous deletion region on 3p21.3. Using the F test for smoking, we found two genes, 101F6 and CACT, located on 3p21 region. In addition to 3p21.3, 11q23–24, where NCAM1 is also located, is known for a high frequency of LOH, and putative tumor suppressor genes may be found in this region (17). In addition, a high frequency of LOH in lung cancers has been reported in chromosomes 19p12–13, 19q13.3, and 4q, where C19ORF3, EDG4, SIGLEC5, and RRH are located (18, 19). Mori et al. (20) showed recently that microsatellite instability was observed frequently in the coding region of PA2G4 on chromosome 12q13. The inactivation of these genes may be related to tobacco carcinogenesis. The high expression of RAB4, DJ1, MCT, and ribosomal protein L22 (RPL22) also may be related to the tobacco smoke-induced adenocarcinoma. RAB4 gene shares biochemical properties with the Ras gene super-family and encodes GTP-binding proteins. DJ1 cooperates with H-RAS in transforming NIH3T3 cells (21). The MCT protein is a novel candidate oncogene sharing a homology with cyclin H, which increases G1 cyclin/cyclin-dependent kinase activity (22). DNA amplification of chromosome 3q26.1-q26.3 is detected frequently in lung squamous cell carcinomas, where RPL22 is also located (23).

DISCUSSION

Analysis of clinical samples is difficult because of the heterogeneity of cellular components in the tissues and their diverse pathological features. In analyzing genetic alterations or expression profiles in cancer cells, the contamination of normal epithelia or stromal cells may confound the analysis, especially when using tissues such as lung, brain, kidney, ovary, mammary gland, or prostate (5, 7, 11, 12), which are unlike the more homogenous fluorescence-activated cell sorted blood cells (13). The use of LCM should improve the sample preparation for microarray expression analysis (4). One disadvantage of microdissection is that RNA amplification is essential for the preparation of an adequate amount of RNA for probe generation. The reproducibility and usefulness of T7 amplification in gene expression profiling has been reported previously (14, 15). In our hands, the gene expression profile from two independent two-rounds of amplification of BEAS-2B RNA was highly correlated (r2 = 0.91; data not shown).

Several of the genes with lower expression among smokers are located in the chromosomal regions where genomic imbalance in NSCLC has been observed previously with high frequency. In Table 2, we summarized the reported data of genomic imbalance including homozygous deletion, which was analyzed with comparative genomic hybridization, LOH analysis, and other methods. For example, chromosome 3p21.3 is a well-known region for a frequent homozygous deletion in lung cancers. Lerman and Minna (16) analyzed 25 genes localized in the 630-kb homozygous deletion region on 3p21.3. Using the F test for smoking, we found two genes, 101F6 and CACT, located on 3p21 region. In addition to 3p21.3, 11q23–24, where NCAM1 is also located, is known for a high frequency of LOH, and putative tumor suppressor genes may be found in this region (17). In addition, a high frequency of LOH in lung cancers has been reported in chromosomes 19p12–13, 19q13.3, and 4q, where C19ORF3, EDG4, SIGLEC5, and RRH are located (18, 19). Mori et al. (20) showed recently that microsatellite instability was observed frequently in the coding region of PA2G4 on chromosome 12q13. The inactivation of these genes may be related to tobacco carcinogenesis. The high expression of RAB4, DJ1, MCT, and ribosomal protein L22 (RPL22) also may be related to the tobacco smoke-induced adenocarcinoma. RAB4 gene shares biochemical properties with the Ras gene super-family and encodes GTP-binding proteins. DJ1 cooperates with H-RAS in transforming NIH3T3 cells (21). The MCT protein is a novel candidate oncogene sharing a homology with cyclin H, which increases G1 cyclin/cyclin-dependent kinase activity (22). DNA amplification of chromosome 3q26.1-q26.3 is detected frequently in lung squamous cell carcinomas, where RPL22 is also located (23).

Table 1 Characteristics of patients

<table>
<thead>
<tr>
<th>ID</th>
<th>Age</th>
<th>Gender</th>
<th>Prognosis</th>
<th>Smoking</th>
<th>Pack-years</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>75</td>
<td>M</td>
<td>D9</td>
<td>N</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>54</td>
<td>F</td>
<td>A</td>
<td>Y</td>
<td>40</td>
</tr>
<tr>
<td>3</td>
<td>48</td>
<td>M</td>
<td>A</td>
<td>Y</td>
<td>50</td>
</tr>
<tr>
<td>4</td>
<td>66</td>
<td>F</td>
<td>D</td>
<td>Y</td>
<td>47</td>
</tr>
<tr>
<td>5</td>
<td>74</td>
<td>M</td>
<td>D</td>
<td>Y</td>
<td>60</td>
</tr>
<tr>
<td>6</td>
<td>59</td>
<td>M</td>
<td>D</td>
<td>Y</td>
<td>125</td>
</tr>
<tr>
<td>7</td>
<td>72</td>
<td>F</td>
<td>D</td>
<td>Y</td>
<td>60</td>
</tr>
<tr>
<td>8</td>
<td>53</td>
<td>F</td>
<td>D</td>
<td>Y</td>
<td>20</td>
</tr>
<tr>
<td>9</td>
<td>55</td>
<td>M</td>
<td>A</td>
<td>Y</td>
<td>80</td>
</tr>
<tr>
<td>10</td>
<td>66</td>
<td>F</td>
<td>D</td>
<td>Y</td>
<td>40</td>
</tr>
<tr>
<td>11</td>
<td>49</td>
<td>M</td>
<td>D</td>
<td>Y</td>
<td>25</td>
</tr>
<tr>
<td>12</td>
<td>73</td>
<td>M</td>
<td>D</td>
<td>N</td>
<td>0</td>
</tr>
<tr>
<td>13</td>
<td>78</td>
<td>F</td>
<td>D</td>
<td>Y</td>
<td>60</td>
</tr>
<tr>
<td>14</td>
<td>61</td>
<td>M</td>
<td>D</td>
<td>Y</td>
<td>-46</td>
</tr>
<tr>
<td>15</td>
<td>70</td>
<td>M</td>
<td>Unknown</td>
<td>Y</td>
<td>60</td>
</tr>
<tr>
<td>16</td>
<td>57</td>
<td>F</td>
<td>A</td>
<td>N</td>
<td>0</td>
</tr>
<tr>
<td>17</td>
<td>38</td>
<td>F</td>
<td>A</td>
<td>N</td>
<td>0</td>
</tr>
<tr>
<td>18</td>
<td>63</td>
<td>F</td>
<td>A</td>
<td>N</td>
<td>0</td>
</tr>
<tr>
<td>19</td>
<td>63</td>
<td>M</td>
<td>D</td>
<td>Y</td>
<td>30</td>
</tr>
</tbody>
</table>

* The status 5 years after surgery.
* D, dead; A, alive; N, no; Y, yes.

Fig. 1. Experimental course of LCM and T7 amplification. A, flow chart of tissue preparation, LCM, and two-round T7 amplification. ss cDNA, single-strand cDNA; ds cDNA, double-strand cDNA. B, the summary of LCM and two-round aRNA yields. C–M, pathological images of LCM. Left panels, before LCM; middle panels, after LCM; and right panels, the tissues captured on LCM caps. C, case 1; D, case 2; E, case 3; F, case 5; G, case 7; H, case 8; I, case 11; J, case 12; K, case 15; L, case 18; and M, case 19.

Fig. 2. A hierarchical clustering and gene expression cluster view of 19 samples and 2270 cDNA clones. Green, transcript levels below the median; black, equal to the median; red, greater than median; and gray, missing data for filtering.
Decreased expression of glioma tumor suppressor candidate region gene 2 (GLTSCR2) and high expression of candidate oncogene EDG4 among nonsmokers may contribute to the carcinogenesis of nonsmokers. The mitochondrial enzyme TK2 is also higher expressed in adenocarcinomas in nonsmokers than in smokers. TK2 is active throughout the cell cycle and is most likely correlated with mitochondrial content in the tumors. TK2 contributes to the metabolic activation of 2',2'-difluorodeoxycytidine (gemcitabine), which has significant activity against non-small cell cancers (24). These data suggest the hypothesis that gemcitabine may be more active in the TK2 higher-expressing adenocarcinoma in nonsmokers when compared with smokers.
Among the 21 genes on chromosome 3p21.3 listed in Table 3, 16 genes including PL6, 101F6, HYAL2, FUS2, and RBM6, which are located in 630 kb with a high frequency of homozygous deletions (16), showed lower expression in smokers. This result supported our hypothesis that genes with lower expression among smokers are located in the chromosomal regions with a high frequency of LOH, including 3p21.3.

Among the 27 genes that were differentially expressed between smokers and nonsmokers, high expression in smokers and low expression in nonsmokers included HYA22 protein, SH3 protein, 101F6, FUS2, and RBM6, which are located in 630 kb with a high frequency of homozygous deletions (16), showed lower expression in smokers. This result supported our hypothesis that genes with lower expression among smokers are located in the chromosomal regions with a high frequency of LOH, including 3p21.3.
mitotic spindle checkpoint. Whereas genes are involved in the metaphase-anaphase transition and the 5-year survivors or more and nonsurvivors (Table 4), several of the – is activated by were expressed higher in cases with a better prognosis. The APC/C proteasome 26S subunit, ATPase hBUB3 higher level in cases with a poor prognosis, (26). In APC2 mutants arrest in metaphase (27). in which inhibitor, which are subsequently degraded by the 26S proteasome, lytic cleavage of cyclins and anaphase inhibitors is necessary for unit E3 ubiquitin ligase that targets proteins of which the proteolytic cleavage of cyclins and anaphase inhibitors is necessary for mitotic checkpoint complex that inhibits APC/C (28), a multisubunit E3 ubiquitin ligase in which APC/C is activated by Cdc20 and Cdh1, and promotes metaphase-anaphase transition by ubiquitinating cyclins and the anaphase inhibitor, which are subsequently degraded by the 26S proteasome, in which PSMC1 and PSMC5 gene products are subunits (25). APC2 is an evolutionarily conserved component of the APC/C (26). In Saccharomyces cerevisiae, temperature-sensitive APC2 mutants arrest in metaphase (27). hBUB3 is a component of the mitotic checkpoint complex that inhibits APC/C (28), a multisubunit E3 ubiquitin ligase that targets proteins of which the proteolytic cleavage of cyclins and anaphase inhibitors is necessary for sister chromatid separation and the exit from mitosis (29, 30). BUB3 is evolutionarily conserved in eukaryotes (31–33). Mutation of BUB3 inactivates the mitotic spindle cell checkpoint in both the budding yeast S. cerevisiae (34) and the mouse (35). Murine BUB3 gene knockout embryos accumulate mitotic errors including lagging chromosomes and micronuclei by 4.5 days and do not survive past 7.5 days (35). ZW10 recruits dynactin and dynein to the centromeric kinetochore (36), a multiregion DNA complex responsible for attachment to spindle microtubules, prometaphase chromosome congression, anaphase initiation, and poleward movement. Mutation in the Drosophila ZW10 gene results in both lagging chromosomes and aneuploidy (36). The positive correlation between poor prognosis and the reduced expression of hBUB3 and hZW10 is consistent with the hypothesis that a defective mitotic spindle checkpoint, and the resultant abnormal segregation and increasing aneuploidy would enhance tumor progression. Bhattacharjee et al. (3) also have found that increased expression of hBUB3 is associated with a better prognosis in lung adenocarcinoma cases.

Our study and others (2, 3) generate hypotheses for both clinical and laboratory investigations. The combined use of LCM with microarray expression analysis should decrease the confounder of tissue heterogeneity. Laboratory studies of differentially expressed genes associated with prognosis may identify genes and their function involved in metastasis.

ACKNOWLEDGMENTS

We thank Drs. Hiroshi Murakami, Ryuchi Wada, Atul Kumar, Ivan Rosas, Masayuki Shiseki, Remy Pedeux, Shu Okamura, Stefan Ambs, Hasan Seker, Xin Wang, and Makoto Nagashima for technical advice and suggestions, Dr. John Gillespie for helping with the pathological diagnosis, Dr. Joanna Shih for helpful discussion on the statistical analysis, the microarray group of Dr. Nakamura’s laboratory for the fabrication of microarrays, Mary G. McMenamin for cell culture, Audrey Salabes for collecting the clinical information, and Dorothea Dudek for editorial assistance.

3249
Laser Capture Microdissection and Microarray Expression Analysis of Lung Adenocarcinoma Reveals Tobacco Smoking- and Prognosis-related Molecular Profiles


Updated version
Access the most recent version of this article at:
http://cancerres.aacrjournals.org/content/62/11/3244

Cited articles
This article cites 46 articles, 28 of which you can access for free at:
http://cancerres.aacrjournals.org/content/62/11/3244.full.html#ref-list-1

Citing articles
This article has been cited by 25 HighWire-hosted articles. Access the articles at:
/content/62/11/3244.full.html#related-urls

E-mail alerts
Sign up to receive free email-alerts related to this article or journal.

Reprints and Subscriptions
To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org.

Permissions
To request permission to re-use all or part of this article, contact the AACR Publications Department at permissions@aacr.org.