Differences in KRAS Mutation Spectrum in Lung Cancer Cases between African Americans and Caucasians after Occupational or Environmental Exposure to Known Carcinogens

Jay D. Hunt, Anna Strimas, Julie E. Martin, Marilyn Eyer, Monica Haddican, Brian G. Luckett, Bernardo Ruiz, T. William Axelrad, Wayne L. Backes, and Elizabeth T. H. Fontham


Abstract

Elevated mortality rates of lung cancer in the Mississippi River corridor in Louisiana have been clearly documented for the past half-century and rank among the highest in the nation. A population-based case-control study of lung cancer termed Lower Mississippi River Interagency Cancer Study was conducted in southern Louisiana. Lung tumor specimens were collected, isolated by laser capture microdissection, subjected to PCR to amplify KRAS, and sequenced to confirm mutation status and specificity. Of the 116 lung tumors analyzed to date, 32 (27.6%) contained mutations in either codon 12 or 13 of KRAS. This frequency is comparable to that reported in the literature; however, the mutation spectrum was strikingly different. Of the 32 mutations observed, 21 (65.6%) resulted in the inappropriate insertion of cysteine, 6 (18.8%) resulted in the insertion of serine, and 1 (3.1%) each resulted in the insertion of aspartate and alanine. These data indicate that an abnormally high proportion of cysteine (P = 0.010) and serine (P = 0.002) mutations was observed in our sample group versus lung cancers reported in the literature. KRAS mutations were more common in African Americans with an odds ratio of 2.4 (P = 0.048), as were serine mutations, although the latter did not reach statistical significance (odds ratio, 2.6; P = 0.373). No association was found between the observed mutation spectrum and known lung cancer risk factors.

Introduction

The 11-parish region from Baton Rouge to the mouth of the Mississippi River (including Ascension, East Baton Rouge, Iberville, Jefferson, Orleans, Plaquemines, St. Bernard, St. Charles, St. John, St. James, and West Baton Rouge parishes) has been termed “cancer corridor” because of the elevated mortality rates of cancer in the region. The Louisiana Tumor Registry database revealed elevated incidence rates between 1993 and 1997 of lung cancer, but not breast, prostate, or colorectal cancers (1). The average annual age-adjusted (US1970 census) incidence rates per 100,000 for cancers of the lung and bronchus in Louisiana between 1993 and 1997 were 97.1 for white males (versus 72.4 for US SEER) and 129.4 for black males (versus 110.5 for US SEER; Ref. 2). These elevated incidence rates resulted in increased average annual age-adjusted mortality rates per 100,000 from cancers of the lung and bronchus of 83.3 for white males (versus 67.9 for US SEER) and 116.3 for black males (versus 96.7 for US SEER).

A case-control study of lung cancer was initiated to address many of the key issues related to cancer risk in this area. The LMRICS is a large population-based case-control study that assesses lung cancer risk associated with environmental exposures in this highly industrialized corridor. Table 1 summarizes the industrial releases from 1999 within the study area that may place surrounding communities at increased risk of lung cancer. The table does not include chemicals added in 1990, 1991, 1994, or 1995 or chemicals deleted in any year. The Environmental Protection Agency determines which substances are reportable using criteria based on the acute toxicity of the substance, the potential health effects from long-term exposure to the substance, and the environmental impact of the substance (3–5). Any substance released in the area that has been linked by epidemiological evidence and/or animal studies to the risk of developing lung cancer is included in Table 1. A rapid case-ascertainment system based within the Louisiana Tumor Registry was used by LMRICS to ensure prompt identification of newly diagnosed cases for interview and collection of biological specimens. Controls, frequency-matched to cases by age, race, and sex, were chosen by sampling from the combined Louisiana Driver’s License/Identification files of the Louisiana Department of Public Safety and Corrections.

Although cancer of the lung is the most common cause of
death from neoplasia in both men and women in the US, the complex genetic changes have not been well characterized. One common genetic aberration that occurs in lung cancer is conversion of the proto-oncogene KRAS to its activated oncogenic form through mutations in codons 12, 13, or 61 (reviewed in Ref. 6). Typically, mutations in codons 12 or 13 are the predominant abnormality of KRAS observed in lung cancers, and mutations in codon 61 are so infrequent that they are ignored (6). Many investigators have evaluated mutations in KRAS as a prognostic indicator of progression or survival. We and others (6). Typically, mutations in codons 12 or 13 are the predominant abnormality of KRAS observed in lung cancers, and mutations in codon 61 are so infrequent that they are ignored (6). Many investigators have evaluated mutations in KRAS as a prognostic indicator of progression or survival. We and others have demonstrated that the specific presence of a KRAS mutation is not a significant biomarker of prognosis (7–11), whereas other studies demonstrate that KRAS mutation frequency does predict survival (12–21). As a general rule, studies done in the US have generally demonstrated that KRAS mutation frequency is not a prognostic marker, whereas studies in Europe and in Asia have demonstrated that mutation frequency is prognostic. In 1997, we demonstrated that the mutation spectrum in KRAS differed between American and European studies and showed that the specific amino acid substitution in mutated KRAS was an important prognostic biomarker of progression and survival and may explain the lack of correlation between studies concerning KRAS as a biomarker in lung cancer (7). This study demonstrated the importance of determining the KRAS mutation spectrum in one’s study population. There are no reports that specifically address the association between KRAS mutation and the African-American race.

The tumor suppressor gene p53 has also been the focus of much attention in a variety of human cancers (22). This gene encodes a M, 53,000 nuclear phosphoprotein that is important in the control of the cell cycle and has been referred to as “the guardian of the human genome” (23–25). Defects in the p53 gene have been found to be the most common genetic alteration in a variety of human cancers (26) and have been extensively documented in lung cancer (27–32). In lung cancer, mutations of the p53 gene have been reported to be present in 35–70% of cases (33, 34). Mutations in p53 appear to be an early event in the formation of lung tumors because they have been observed in early bronchial neoplasia (35) and can be detected in precursor lesions such as dysplasia (36). These changes, which occur almost exclusively within the highly conserved domains of the gene (exons 5–8), appear to have prognostic value; several reports indicate that patients with tumors containing mutations in these regions have a significantly worse prognosis than those without such alterations (37–40). As with KRAS, there is a paucity of data on the association of p53 mutation in lung cancer for African Americans.

To determine whether the KRAS mutation spectrum was associated with social, occupational, or environmental risk factors for lung cancer in the LMRICS study area, we isolated available specimens from patients in the 11-parish area. We compared KRAS mutation frequency and spectrum with self-reported race and with environmental or occupational exposure to cigarette smoke, asbestos, cotton fibers or dust, other textile fibers or dust, fiberglass, glass wool, silica dust (including sand and concrete dust), metal oxide dust, coal dust, coal tar, paints, lacquers, stains, inks, hair dyes, hair tints, cooking fumes, cooking oils, or pesticides or with employment in a refinery or a chemical plant.

The carcinogenicity of many compounds including PAHs is mediated through metabolically formed epoxides. Many studies have indicated that the conversion of PAHs to dihydrodiol epoxides is critical for the formation of the ultimate carcinogen (reviewed in Ref. 41). 7β,8α-Dihydroxy-9α,10α-epoxy-7,8,9,10-tetrahydrobenzo(a)pyrene is a mutagenic and carcinogenic metabolite of benzo(a)pyrene found in tobacco smoke (42–44). Of the many metabolites of benzo(a)pyrene, studies indicate that the most potent carcinogenic metabolite of benzo(a)pyrene is the (+) enantiomer benzo(a)pyrene-7,8-diol-9,10-epoxide-1 (45). The (+) enantiomer binds preferentially to the 2-amino group of guanine in double-stranded DNA and binds to a 20-fold greater extent than does the (−) enantiomer (46). When benzo(a)pyrene-7,8-diol-9,10-epoxide-1 was injected into newborn mice, lung adenomas and adenocarcinomas, as well as lymphomas, and liver tumors developed in a dose-dependent fashion (45). The P450 family of mixed function oxidases is important in the activation of a number of chemical carcinogens to their ultimate carcinogenic metabolites (41), including 2-acetylaminofluorene, other arylamines, aflatoxin B, nitrosamines, and PAHs. P450 enzymes predominantly form the more carcinogenic (+) enantiomers of benzo(a)pyrene-7,8-diol-9,10-epoxides from benzo(a)pyrene (45). The effects of particular agents that modulate the P450 enzymes in vivo on PAH carcinogenesis are poorly understood, but it is assumed that ingestion of drugs, the smoking of cigarettes, exposure to halogenated hydrocarbons, and diet can influence the metabolism and therefore the carcinogenicity of PAHs. Variability in enzyme activity has been associated with genotype at several polymorphic loci and has been associated with increased risk of cancer. One cytochrome P450A1 (or CYP1A1) has been extensively studied because of its association with cigarette smoking (47–53). The PAHs of cigarette smoke will induce expression of CYP1A1. Genetic polymorphisms of CYP1A1 combined with a genetic deficiency in glutathione S-transferase, which detoxifies the electrophilic metabolites of PAHs, are associated with increased risk of cigarette smoking.

Table 1 Total reported releases (in pounds) in 1999 for 1988 core chemicals in the LMRICS areaa

<table>
<thead>
<tr>
<th>Parish</th>
<th>Air</th>
<th>Water</th>
<th>Underground</th>
<th>Land</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ascension</td>
<td>4,151,762</td>
<td>100,093</td>
<td>2,233,128</td>
<td>10,365</td>
<td>6,784,708</td>
</tr>
<tr>
<td>East Baton Rouge</td>
<td>4,264,741</td>
<td>216,655</td>
<td>367,518</td>
<td>5,804,365</td>
<td>12,193,374</td>
</tr>
<tr>
<td>Iberville</td>
<td>2,207,929</td>
<td>7,312</td>
<td>2</td>
<td>336,822</td>
<td>2,585,813</td>
</tr>
<tr>
<td>Jefferson</td>
<td>592,861</td>
<td>8,133</td>
<td>12,337,645</td>
<td>5</td>
<td>12,967,471</td>
</tr>
<tr>
<td>Orleans</td>
<td>183,553</td>
<td>2,010</td>
<td>1</td>
<td>207,563</td>
<td></td>
</tr>
<tr>
<td>Plaquemines</td>
<td>183,144</td>
<td>4,999</td>
<td>134,283</td>
<td>431,192</td>
<td></td>
</tr>
<tr>
<td>St. Bernard</td>
<td>518,135</td>
<td>698</td>
<td>16,499</td>
<td>539,911</td>
<td></td>
</tr>
<tr>
<td>St. Charles</td>
<td>2,667,911</td>
<td>39,932</td>
<td>7,645,753</td>
<td>10,824,433</td>
<td></td>
</tr>
<tr>
<td>St. James</td>
<td>328,925</td>
<td>3,160</td>
<td>45,069</td>
<td>382,900</td>
<td></td>
</tr>
<tr>
<td>St. John</td>
<td>618,254</td>
<td>373</td>
<td>41,217</td>
<td>679,518</td>
<td></td>
</tr>
<tr>
<td>West Baton Rouge</td>
<td>476,859</td>
<td>1,362</td>
<td>1</td>
<td>588,185</td>
<td></td>
</tr>
</tbody>
</table>

a Summarized from the Toxics Release Inventory Explorer, data source: 1999 data update as of August 1, 2001 (http://www.epa.gov/triexplorer/reports.htm).
induced lung cancer (47, 51, 54–56). GST, which conjugates xenobiotics with glutathione, comprises four known classes: α (A); μ (M); π (P); and θ (T). Activated forms of benzo-(a)pyrene are substrates for GSTM1 (GST-μ). An allelic variant of GSTM1 (GSTM1-null) contains essentially a complete deletion of the gene, resulting in no production of protein (GSTM1-null genotype in homozygous individuals, ~56% of the population). Individuals who are GSTM1-null and have been exposed to chemical carcinogens show increased risk for several types of cancer (57–59). Hayashi et al. (60) reported a relative risk for lung cancer of 5.8 for individuals who were homozygous for the rare allele of CYP1A1 and for GSTM1-null. An OR of 3.0 was reported for squamous cell carcinoma of the lung in individuals who were GSTM1-null and carried the MspI polymorphism in CYP1A1 (61). In this study, phenotypes were determined for the GSTM1-null genotype, the CYP1A1 MspI polymorphism, and the CYP1A1 exon 7 polymorphism, and these were compared with the frequency and spectrum of KRAS gene mutation in LMRICS tumor samples.

Materials and Methods
Selection of Cases and Controls. LMRICS is a population-based case-control study to assess lung cancer risk associated with occupational or environmental exposures. Eligible cases aged 20–74 years must have resided in 1 of the 11 Louisiana parishes of the LMRICS area at the time of diagnosis with microscopically confirmed primary carcinoma of the lung (International Classification of Diseases ICD-9, 162.2–162.9), diagnosed before death; no history of a previous cancer (except basal or squamous carcinoma of the skin) was allowed. Determination of race was based on self-report, obtained first from medical records and confirmed by the respondent at interview. No exclusions were made based on race or ethnicity; however, most study subjects were either African American or non-Hispanic Caucasian (Table 2). The LMRICS project interviewed and collected blood or buccal cells from newly diagnosed lung cancer patients and from population-based controls frequency-matched to cases by age, race, and sex. Available tumor tissue blocks were obtained from participating hospitals. The Louisiana State University Health Sciences Center Institutional Review Board approved this study. All study volunteers gave their informed consent before inclusion in this study.

DNA Isolation, PCR Amplification, and DNA Sequencing. DNA was isolated from microdissected tumor specimens using Puregene DNA isolation protocol (Gentra Systems, Inc., Minneapolis, MN) as described previously (7). To amplify the small quantities of DNA, a nested PCR reaction was used to amplify the first exon of KRAS. The 25-μl outer PCR reaction was conducted using primers XKF (5′-GUAA-CUG-GUG-UAU-UUG-AUA-3′) and XKR (5′-GGU-CAG-AGA-AAC-CUU-UAA-CUG-UAA-C-3′) by an initial exposure to 95°C for 5 min followed by 36 cycles of 94°C for 30 s, 55°C for 1 min, and 72°C for 1 min. A final incubation at 72°C for 10 min was conducted before termination of the PCR. Water was substituted for DNA template as a negative control to determine if carry-over contamination was present. DNA isolated from the cell line Calu-1, which has a mutation in KRAS, was used as a positive control. After the inner PCR reaction, 5 μl of the reaction were treated with shrimp alkaline phosphatase and manually sequenced as described previously (7). In many instances, both a normal band (from DNA from normal cells contaminating the tumor, the normal allele of a tumor containing a heterozygous mutated KRAS, or both) and a mutated band are observed. To ensure that the data are correct, two individuals (J. D. H. and A. S. or J. E. M.) independently scored KRAS mutation status and sequence before a sample was termed mutated. The sequencing of KRAS was done in the reverse direction (3′ to 5′) because the mutations at codon 12 or 13 are too close to the 5′ end of the gene to obtain reasonable accuracy if the reaction is primed from the 5′ end of the PCR-amplified DNA. Some mutations were confirmed using Denaturing high-performance liquid chromatography (Transgenomic, Omaha, NE).

Examination of Polymorphisms in Drug-metabolizing Enzymes. DNA was isolated for PCR analysis using the Bio-Rad (Hercules, CA) InstaGene Whole Blood Kit according to the manufacturer’s instructions. The CYP1A1 mutation found in the 3′-flanking region was detected by PCR and RFLP analysis using the MspI restriction enzyme (62, 63). The DNA fragment was amplified using the following primers: 5′-CAG-TGA-AGA-GTG-CTG-TGA-GCC-GCT-3′ and 5′-TAG-GAG-TCT-TGC-AGG-TAT-3′. After amplification, the PCR product was subjected to restriction digestion using MspI and separated by agarose gel electrophoresis. The wild-type allele produced a 340-bp band, whereas the variant produced bands at 200 and 140-bp. Polymorphisms in the coding region of the CYP1A1 gene were detected by allele-specific amplification as described previously (63). DNA samples were subjected to PCR ampli-

<table>
<thead>
<tr>
<th>Table 2</th>
<th>Descriptive statistics for cases from whom lung tumors were obtained</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ethnicity (self-reported)</td>
<td>Frequency</td>
</tr>
<tr>
<td>Non-Hispanic Caucasian</td>
<td>51</td>
</tr>
<tr>
<td>African American</td>
<td>60</td>
</tr>
<tr>
<td>Native American</td>
<td>2</td>
</tr>
<tr>
<td>Hispanic</td>
<td>1</td>
</tr>
<tr>
<td>Other</td>
<td>2</td>
</tr>
<tr>
<td>Sex</td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>37</td>
</tr>
<tr>
<td>Male</td>
<td>79</td>
</tr>
<tr>
<td>Age at diagnosis (yrs)</td>
<td>59.7±</td>
</tr>
<tr>
<td>Smoking history</td>
<td>60.6 pack-years±</td>
</tr>
</tbody>
</table>

* Mean.
* SD.
fication using the common primer 5'-GAA-CTG-CCA-CTTCAG-CTG-TCT-3' and either of the following primers: 5'-AAG-ACC-TCC-CAG-CGG-GCA-AT-3' (CA4) or 5'-AAG-ACC-TCC-CAG-CGG-GAA-C-3' (CA5). Wild-type samples produced a band with the common/CA4 primers, whereas the variant samples produced bands with the common/CA5 primers. The GSTM1 genotype was detected after PCR amplification using primers for the GSTM1 gene (64) and the globin gene. The GSTM1 primers were 5'-CTG-CCC-TAC-ATG-ATT-GAT-GGG-3' and 5'-CTG-GAT-TGT-AGA-AGA-TGC-3'. The wild-type samples produced a band at 300 bp. In the variant samples, the GSTM1 gene was absent, and no band was observed. A portion of the globin gene was amplified as a positive control, producing a 200-bp fragment. The following

<table>
<thead>
<tr>
<th>Study</th>
<th>No. mutated (%)</th>
<th>AA</th>
<th>Count</th>
<th>% of Total</th>
<th>% of Mutated</th>
</tr>
</thead>
<tbody>
<tr>
<td>LMRICS lung cancer (n = 116)</td>
<td>32 (27.6%)</td>
<td>Cys</td>
<td>21</td>
<td>18.1</td>
<td>65.6</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Ser</td>
<td>6</td>
<td>5.2</td>
<td>18.8</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Val</td>
<td>3</td>
<td>2.6</td>
<td>9.4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Asp</td>
<td>1</td>
<td>0.9</td>
<td>3.1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Ala</td>
<td>1</td>
<td>0.9</td>
<td>3.1</td>
</tr>
<tr>
<td>Lung cancer (other studies) (n = 1396)</td>
<td>336 (24.1%)</td>
<td>Cys</td>
<td>141</td>
<td>10.1</td>
<td>42.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Asp</td>
<td>80</td>
<td>5.7</td>
<td>23.8</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Val</td>
<td>71</td>
<td>5.1</td>
<td>21.1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Ala</td>
<td>17</td>
<td>1.2</td>
<td>5.1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Ser</td>
<td>12</td>
<td>0.9</td>
<td>3.6</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Arg</td>
<td>13</td>
<td>0.9</td>
<td>3.9</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Phe</td>
<td>2</td>
<td>0.1</td>
<td>0.6</td>
</tr>
<tr>
<td>Siegfried et al. (Ref. 7) (n = 181)</td>
<td>57 (31.5%)</td>
<td>Cys</td>
<td>28</td>
<td>15.5</td>
<td>49.1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Val</td>
<td>12</td>
<td>6.6</td>
<td>21.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Asp</td>
<td>8</td>
<td>4.4</td>
<td>14.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Ala</td>
<td>5</td>
<td>2.8</td>
<td>8.8</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Arg</td>
<td>3</td>
<td>1.6</td>
<td>5.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Phe</td>
<td>1</td>
<td>0.6</td>
<td>1.8</td>
</tr>
<tr>
<td>Significantly fewer serine substitutions</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(P = 0.002, Fisher’s exact test)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rodenhuis and Slebos (Ref. 12) (n = 280)</td>
<td>45 (16.1%)</td>
<td>Cys</td>
<td>25</td>
<td>8.9</td>
<td>55.6</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Asp</td>
<td>10</td>
<td>3.6</td>
<td>22.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Val</td>
<td>7</td>
<td>2.5</td>
<td>15.6</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Ala</td>
<td>2</td>
<td>0.7</td>
<td>4.4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Arg</td>
<td>1</td>
<td>0.4</td>
<td>2.2</td>
</tr>
<tr>
<td>Significantly fewer serine substitutions</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(P = 0.004, Fisher’s exact test)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Silini et al. (Ref. 13) (n = 109)</td>
<td>32 (29.4%)</td>
<td>Cys</td>
<td>15</td>
<td>13.8</td>
<td>46.9</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Val</td>
<td>7</td>
<td>6.4</td>
<td>21.9</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Ala</td>
<td>5</td>
<td>4.6</td>
<td>15.6</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Asp</td>
<td>5</td>
<td>4.6</td>
<td>15.6</td>
</tr>
<tr>
<td>Significantly fewer serine substitutions</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(P = 0.024, Fisher’s exact test)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Keohavong et al. (Ref. 8) (n = 173)</td>
<td>42 (24.3%)</td>
<td>Cys</td>
<td>13</td>
<td>7.5</td>
<td>31.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Asp</td>
<td>12</td>
<td>6.9</td>
<td>28.6</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Val</td>
<td>12</td>
<td>6.9</td>
<td>28.6</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Arg</td>
<td>2</td>
<td>1.2</td>
<td>4.8</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Ser</td>
<td>2</td>
<td>1.2</td>
<td>4.8</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Phe</td>
<td>1</td>
<td>0.6</td>
<td>2.4</td>
</tr>
<tr>
<td>No statistical difference in number of</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>serine substitutions (P = 0.07, Fisher’s</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>exact test)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rosell et al. (Ref. 20) (n = 275)</td>
<td>52 (18.9%)</td>
<td>Val</td>
<td>14</td>
<td>5.1</td>
<td>26.9</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Asp</td>
<td>13</td>
<td>4.7</td>
<td>25.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Cys</td>
<td>10</td>
<td>3.6</td>
<td>19.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Ser</td>
<td>8</td>
<td>2.9</td>
<td>15.4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Arg</td>
<td>6</td>
<td>2.2</td>
<td>11.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Ala</td>
<td>1</td>
<td>0.4</td>
<td>1.9</td>
</tr>
<tr>
<td>No statistical difference in number of</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>serine substitutions (P = 0.77, Fisher’s</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>exact test)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other cancers (non-lung) (n = 321)</td>
<td>163 (50.8%)</td>
<td>Val</td>
<td>39</td>
<td>12.1</td>
<td>23.9</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Asp</td>
<td>33</td>
<td>10.3</td>
<td>20.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Ser</td>
<td>22</td>
<td>6.9</td>
<td>13.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Cys</td>
<td>21</td>
<td>6.5</td>
<td>12.9</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Ala</td>
<td>18</td>
<td>5.6</td>
<td>11.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Arg</td>
<td>2</td>
<td>0.6</td>
<td>1.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Unknown/other</td>
<td>29</td>
<td>9.0</td>
<td>17.8</td>
</tr>
</tbody>
</table>

Table 3: Comparisons of KRAS mutation spectrum in LMRICS samples

Significantly fewer serine substitutions (P = 0.002, Fisher’s exact test)
primers were used: 5'-GAA-GAG-CCA-AGG-ACA-GGT-AC-3' and 5'-GGT-GTC-TGT-TTG-AGG-TTG-CT-3'.

Statistics. Poole t tests for independent samples with equal variances, Pearson χ² analyses, and Fisher's exact tests were conducted using StatView 5.0.1 (SAS Institute, Inc., Cary, NC). Satterthwaite t tests for unequal variances were conducted using The SAS System (SAS Institute, Inc.). The Satterthwaite test was used when the folded F-test for variances failed to reject the null hypothesis at α = 0.05. In all tests, the threshold for significance was set at α = 0.05.

Results

The KRAS Mutation Spectrum Is Unusual in LMRICS Cases. From the 116 lung tumor specimens, the KRAS mutation spectrum was determined by a nested PCR reaction followed by direct DNA sequencing. As seen in Table 3, we demonstrated that 32 of 116 (27.6%) specimens contained mutations in KRAS, which is comparable with 336 of 1396 (24.1%; P = 0.396, χ² goodness-of-fit analysis with 1 DF) reported in the literature (7, 8, 11–14, 16, 18, 20, 21, 65–69). Of the 32 that were mutated, 31 of 32 involved codon 12, and only 1 of 32 involved codon 13. The KRAS mutation frequency was not associated with age at diagnosis, smoking history, or sex (Table 4). However, we found profound dissimilarity in the KRAS mutation spectrum between the LMRICS lung cancer specimens and the spectrum reported in the literature. The most common mutation observed in KRAS for lung cancer specimens is a G→T transversion in codon 12 (GGT→TGT) resulting in an inappropriate insertion of cysteine for glycine (6). We found a disproportionately high number of cysteine substitutions (Table 3; P = 0.010, χ² goodness-of-fit analysis with 1 DF) in the LMRICS sample set. Even more importantly, we observed a profoundly high number of inappropriate serine substitutions (Table 3; P = 0.002, two-sided Fisher's exact test), which result from G→A transitions in codon 12 (GGT→AGT). A comparison was made between the present study and previously reported individual studies (Table 3). Of the 15 reported studies, only 5 had sample sizes providing sufficient statistical power to conduct analyses (7, 8, 12, 13, 20). We found that three of the studies reported no serine substitutions in KRAS (7, 12, 13) in a combined 470 cases. The other two studies (8, 20) reported serine substitutions in KRAS in 2 of 173 (1.2%) and 8 of 275 (2.9%) cases, which was similar to the numbers observed in the present study [6 of 116 (5.2%); P = 0.07 and 0.77, respectively, using Fisher’s exact test].

Overall mutation frequency in 115 of the cases was not associated with self-reported environmental or occupational exposure to asbestos, cotton fibers or dust, other textile fibers or dust, fiberglass, glass wool, silica dust (including sand and concrete dust), metal oxide dust, coal dust, coal tar, paints, lacquers, stains, inks, hair dyes, hair tints, cooking fumes, cooking oils, or pesticides or with employment in a refinery or a chemical plant (no exposure data were available for one case). Likewise, there was no association of overall KRAS mutation frequency or with the KRAS mutation spectrum with the GSTM1-null genotype, the CYP1A1 MspI polymorphism, or the CYP1A1 exon 7 polymorphism; however, these analyses had limited statistical power.

There was an association of KRAS mutation frequency with race (Table 5). African Americans had significantly greater frequency of KRAS mutation than did Caucasians (OR, 2.4; P = 0.048, χ² goodness-of-fit analysis with 1 DF). Although the rare serine mutations were more frequent in African Americans than in Caucasians, the frequency was not statistically significantly higher (Table 6; P = 0.274, χ² goodness-of-fit analysis with 2 DF). In comparing the association of race with serine substitutions versus all other amino acid substitutions, a nonsignificantly elevated OR of 2.6 (P = 0.373, one-sided Fisher's exact test) was found for African Americans as compared with Caucasians.

Discussion

Lung cancer is the most common cancer among residents of Louisiana, accounting for 19% of the total new cases during the 5-year period of 1988–1992 (2). Although the incidence of lung cancer in Louisiana is elevated, the overall KRAS mutation frequency found in the LMRICS tumor set was comparable with the frequency observed worldwide as reported in the literature. Occupational or environmental risk factors for lung cancer found in the highly industrialized Mississippi River corridor do not lead to a greater overall rate of KRAS mutations. However, we observed uncharacteristically high levels of cysteine and serine amino acid substitutions in our sample set as compared with those observed by us (7) and others (8, 11–14, 16, 18, 20, 21, 65–69) in lung cancers from patients outside of the LMRICS study area (the samples used in the present study are not the same samples previously used by us in Ref. 7).

Cysteine missense substitutions result from G→T transversions at the first base in either codon 12 or 13; the wild-type sequence is GGT GGC (glycine-glycine). Typically, G→T

### Table 4 KRAS mutation frequency by age at diagnosis, smoking history, and sex

<table>
<thead>
<tr>
<th>A. Variable</th>
<th>Mean</th>
<th>SD</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age at diagnosis</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No KRAS mutation</td>
<td>60.6 yrs</td>
<td>9.8</td>
<td>0.123³</td>
</tr>
<tr>
<td>KRAS mutation</td>
<td>57.5 yrs</td>
<td>8.6</td>
<td></td>
</tr>
<tr>
<td>Smoking history</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No KRAS mutation</td>
<td>64.2 pack-years</td>
<td>44.5</td>
<td>0.084⁴</td>
</tr>
<tr>
<td>KRAS mutation</td>
<td>51.3 pack-years</td>
<td>31.3</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>B. Sex</th>
<th></th>
<th></th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>No KRAS mutation</td>
<td>56</td>
<td>28</td>
<td>0.591τ</td>
</tr>
<tr>
<td>KRAS mutation</td>
<td>23</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

a Pooled t test for equal variances.

b Satterthwaite t test for unequal variances (folded F-test for variances, P = 0.0298).

c χ² goodness-of-fit analysis with 1 DF.

### Table 5 KRAS mutations are more common in African Americans

<table>
<thead>
<tr>
<th>Race</th>
<th>KRAS mutation</th>
<th>No KRAS mutation</th>
</tr>
</thead>
<tbody>
<tr>
<td>African American</td>
<td>22 (37%)</td>
<td>38 (63%)</td>
</tr>
<tr>
<td>Caucasian</td>
<td>10 (20%)</td>
<td>41 (80%)</td>
</tr>
</tbody>
</table>

### Table 6 Rare serine mutations are more common in African Americans

<table>
<thead>
<tr>
<th>Race</th>
<th>Amino acid substitution</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Cysteine</td>
</tr>
<tr>
<td>African American</td>
<td>15 (68.2%)</td>
</tr>
<tr>
<td>Caucasian</td>
<td>6 (60.0%)</td>
</tr>
</tbody>
</table>

P = 0.274 (χ² goodness-of-fit analysis with 2 DF).
transversions in lung cancer are associated with the PAH benz(a)pyrene present in tobacco smoke, the predominant risk factor for lung carcinogenesis. However, in other tumor types (e.g., pancreatic tumors, angiosarcoma, extrahepatic tumors of the bile system, biliary tract tumors, colorectal adenomas and carcinomas, and colorectal metastases) containing KRAS mutations in which tobacco smoke is a less significant risk factor, G→A transitions are seen at a much higher rate (19, 70–74). G→A transitions in the first base of either codon 12 or 13 will result in the inappropriate substitution of serine for glycine. When the serine mutation frequency of the LMRICS lung cancer specimens was compared with the serine mutation frequency reported in tumor types other than lung cancer (19, 70–74), there was no difference in serine mutation frequency (P = 0.501, χ² goodness-of-fit analysis with 1 DF). This may indicate that a mutagen other than tobacco smoke may be responsible for the elevated number of observed serine missense substitutions in the LMRICS lung cancer samples. However, it is important to note that the increase in serine substitutions did not simply result from a disproportionate number of G→A transitions. As noted previously, G→A transitions in the first base of either codon 12 or 13 will result in serine substitutions; however, a G→A transition in the second position of either codon will result in the inappropriate substitution of aspartate for glycine. When the number of G→A mutations in the LMRICS sample set (7 of 32 or 21.9%) was compared with G→A frequency in the 1396 lung cancers from the literature (92 of 336 or 27.4%), no difference was observed (P = 0.502, χ² goodness-of-fit analysis with 1 DF). Indeed, the number of serine substitutions is appreciably higher in the LMRICS cases, and the number of aspartate substitutions is appropiately lower, suggesting that a specific carcinogen or risk factor is associated with these substitutions. This may have adverse consequences for the patients with serine substitutions because we have previously demonstrated that the specificity of the KRAS substitution has particular bearing on the prognosis of the lung cancer patient (7). We demonstrated that patients with lung cancers that contained KRAS substitutions of hydrophilic charged amino acids aspartate and arginine had significantly reduced overall survival versus those patients whose tumors contained wild-type glycine or hydrophobic amino acid substitutions. Serine is hydrophilic, and although we have not determined the survival of patients who have this substitution, we speculate that serine substitutions may be associated with a poorer survival. This will be addressed in a future study on survival in the LMRICS group.

Nationally, lung cancer rates are elevated in African Americans (75–77). In Louisiana, the incidence rate of lung cancer for both African American and Caucasian individuals is elevated over national rates for both racial groups (1, 2); however, the rate for African-American-Louisianians is dramatically higher than national levels (78). The use of race as a categorical identifier in medicine has recently been criticized (79). Although it seems unlikely that there are inherent biological differences between different ethnic groups, we have noted a profound difference in the KRAS mutation frequency between African Americans and Caucasians; conversely, the mutation spectrum between African Americans and Caucasians is not significantly different. Individuals of both racial groups had a higher prevalence of cysteine substitutions, although African Americans were much more likely to have very rare serine substitutions. If there are no major biological differences in susceptibility between African Americans and Caucasians, then differences in exposure are likely to account for the observed differences in KRAS mutation frequency. The primary noncigarette-related respiratory exposure to PAHs is mainly from tobacco smoke and urban air (80). Other important lung cancer carcinogens found in tobacco smoke are N-nitrosamines (81). More African Americans smoke than the general population in the US, although African Americans smoke significantly fewer cigarettes than do Caucasians (75). Therefore, it seems unlikely that tobacco consumption alone accounts for the differences in lung cancer rates and KRAS mutation frequency between African Americans and Caucasians. However, approximately 75–90% of African-American smokers prefer menthol cigarettes compared with 20–30% of Caucasian smokers (75). It is known that combustion of menthol produces benz(a)pyrene, which may account for the higher KRAS mutation frequency observed in African Americans. The elevated number of serine substitutions observed in African Americans is unexplained. Given that large numbers of serine substitutions have not been observed in previous studies, it appears that a specific risk factor common to the African-American community in the LMRICS study area may be responsible.

References


Differences in KRAS Mutation Spectrum in Lung Cancer Cases between African Americans and Caucasians after Occupational or Environmental Exposure to Known Carcinogens

Jay D. Hunt, Anna Strimas, Julie E. Martin, et al.


Updated version Access the most recent version of this article at: http://cebp.aacrjournals.org/content/11/11/1405

Cited articles This article cites 72 articles, 25 of which you can access for free at: http://cebp.aacrjournals.org/content/11/11/1405.full#ref-list-1

Citing articles This article has been cited by 6 HighWire-hosted articles. Access the articles at: http://cebp.aacrjournals.org/content/11/11/1405.full#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.