Mutations of the Epidermal Growth Factor Receptor Gene in Lung Cancer: Biological and Clinical Implications

Takayuki Kosaka,1,3 Yasushi Yatabe,2 Hideki Endoh,1,3 Hiroyuki Kuwano,3 Takashi Takahashi,4 and Tetsuya Mitsudomi,2

1Departments of Thoracic Surgery and 2Pathology and Molecular Diagnostics, Aichi Cancer Center Hospital, Nagoya, Japan; 3Department of Surgery I, Gifu University School of Medicine, Gifu, Japan; and 4Division of Molecular Oncology, Aichi Cancer Center Research Institute, Nagoya, Japan

INTRODUCTION
Non–small-cell lung cancer (NSCLC) frequently overexpresses receptors of the erbB family including the epidermal growth factor receptor (EGFR) encoded by erbB-1 (HER1; ref. 1, 2). The EGFR is a 170 kilodaltons receptor tyrosine kinases (TK) that dimerizes and phosphorylates several tyrosine residues after binding of several specific ligands (1). These phosphorylated tyrosines serve as the binding sites for several signal transducers that initiate multiple signaling pathways resulting in cell proliferation, migration, and metastasis, evasion from apoptosis, or angiogenesis, all of which are associated with cancer phenotypes (1). Downstream pathways include ras-raf-MEK-ERK (raf-mitogen-activated protein kinase–extracellular signal-regulated kinase), phosphatidylinositol-3 kinase-AKT and PAK-JNK-JNKK-JNK (p21-activated protein kinase–c-Jun NH2 terminal signal-regulated kinase), phosphatidylinositol-3 kinase-AKT, respectively. EGFR mutations were more frequent in well to moderately differentiated NSCLC to correlate them with clinical and pathologic features. EGFR mutations were present in 111 patients (40%). Fifty-two were in-frame deletions around codons 746–750 in exon 19, 54 were point mutations including 49 at codon 858 in exon 21 and 4 at codon 719 in exon 18, and 5 were duplications/insertions mainly in exon 20. They were significantly more frequent in female (P < 0.001), adenocarcinomas (P = 0.0013), and in never-smokers (P < 0.001). Multivariate analysis suggested EGFR mutations were independently associated with adenocarcinoma histology (P = 0.0012) and smoking status (P < 0.001), but not with female gender (P = 0.9917). In adenocarcinomas, EGFR mutations were more frequent in well to moderately differentiated tumors (P < 0.001) but were independent of patient age, disease stages, and patient survival. KRAS and TP53 mutations were present in 13 and 41%, respectively. EGFR mutations never occurred in tumors with KRAS mutations, whereas EGFR mutations were independent of TP53 mutations. EGFR mutations define a distinct subset of pulmonary adenocarcinoma without KRAS mutations, which is not caused by tobacco carcinogens.

MATERIALS AND METHODS
Patients. Primary tumor samples were obtained from 277 unselected patients with lung cancer who underwent potentially curative pulmonary resection at the Department of Thoracic Surgery, Aichi Cancer Center Hospital from May, 2000 through November, 2000 and from January, 2001 through December, 2002, after obtaining appropriate approval from the institutional review and patients’ written informed consent. These cases corresponded to 82% of all consecutive cases. Inclusion of the cases into this study was dependent on availability of frozen tumor material. About 20 cases were excluded because tumor cells were too few to sufficiently extract tumor DNA because of inflammation and/or necrosis. There were 159 males and 118 females with an age at diagnosis ranging from 26 to 89 (median 64) years. One hundred fifty-nine patients had stage I disease, 39 had stage II, 74 had stage III and 5 had stage IV diseases. There were 224 adenocarcinomas, 35 squamous cell carcinomas, 9 large cell carcinomas, 5 adenosquamous carcinomas, 3 small cell carcinomas, and 1 carcinoid. There were 115 never-smokers and 162 ever-smokers including current and former smokers. Smoking history was obtained by interviewing each patient at admission or first outpatient visit.

Molecular Analysis of Lung Cancer Specimens. Tumor samples were obtained at the time of surgery, rapidly frozen in liquid nitrogen, and stored at −80°C. Frozen tissue of the tumor specimens were grossly dissected to enrich as much tumor cells as possible by a surgical pathologist (Y. Y.). We isolated total RNA using the RNeasy kit (Qiagen, Valencia, CA). The first four exons (exons 18–21) of the seven exons (exons 18–24) that code for TK domain of the EGFR gene that includes all of the mutations reported thus far (8, 9) were amplified with primers F1 (5′-AGCTTGGTGGAGGCTTTACACC-3′) and R1 (5′-TAAAAATGATCAACCTGCTCC-3′), in a one-step reverse transcription-PCR setup with Qiagen OneStep reverse transcription-PCR kit (Qiagen, Valencia, CA). The cDNA sequence of EGFR gene was obtained from GenBank (accession number NM005258). Reverse transcription-PCR conditions were available after request. Reverse transcription-PCR products were diluted and cycle-sequenced with the Big Dye Terminator v3.1/1.1 cycle sequencing kit (Applied Biosystems, Foster City, CA). Sequencing reactions were electrophoresed on an ABI PRISM 3100 (Applied Biosystems). Both the forward and reverse sequences obtained were analyzed by BLAST and chromatograms by manual review.

KRAS and TP53 Gene Analysis. We had previously examined the same cohort for KRAS mutations and TP53 mutations (10, 11). Briefly, TP53 gene (exon 4 through 10) and KRAS gene (exons 1 and 2) were amplified and directly sequenced with ABI PRISM 310 Genetic Analyzer (Applied Biosystems). The forward and reverse sequences obtained were analyzed by BLAST and chromatograms by manual review.

KRAS and TP53 Gene Analysis. We had previously examined the same cohort for KRAS mutations and TP53 mutations (10, 11). Briefly, TP53 gene (exon 4 through 10) and KRAS gene (exons 1 and 2) were amplified and directly sequenced with ABI PRISM 310 Genetic Analyzer (Applied Biosystems). The forward and reverse sequences obtained were analyzed by BLAST and chromatograms by manual review.

Statistical Analysis. For comparisons of proportions, the χ2 test or Fisher’s exact test were used. The Kaplan-Meier method was used to estimate the probability of survival as a function of time, and survival differences were
Forty-six of the 54 point mutations were from a Tt to an G transversion. Fig. 1. It is noted that, in all cases, such alterations were in-frame. Details of the resulting changes in EGFR protein as a consequence of these mutations are illustrated in Fig. 1. In 14 tumors, corresponding cDNA from normal lung tissue far from the tumors was also sequenced, which confirmed that all these mutations were somatic. Of the mutations found, the four major classes of mutations (i.e., deletions, L858R, mutations at codon 719, duplications/insertions) never occurred simultaneously. Furthermore, it is of note that only mutant sequences were present in chromatograms in 19 of 52 deletions, 13 of 54 point mutations, and 5 duplication/insertion mutations. In 14 tumors, rare mutations (all 5 insertions, E709H, T790M, R776C, S768I, L858R) were resequenced and confirmed that these mutations were actually present. In summary, 52 of the 111 (47%) EGFR mutations were deletions around codons 746–750 and 49 (44%) were L858R, altogether accounting for 91% of all of the EGFR mutations found. The four major classes of mutations (i.e., deletions, L858R, mutations at codon 719, duplications/insertions) never occurred simultaneously. Furthermore, it is of note that only mutant sequences were present in chromatograms in 19 of 52 deletions, 13 of 46 in L858R, and 1 of 4 codon 719 mutations.}

**RESULTS**

**EGFR Mutations in Unselected Lung Cancer Specimens.** Of 277 unselected patients who underwent surgical resection of their tumors, we found that 111 patients (40%) had mutations in exons 18–21 of the EGFR gene. There were 52 deletion mutations, 54 point mutations, and 5 duplication/insertion mutations. In 14 tumors, corresponding cDNA from normal lung tissue far from the tumors was also sequenced, which confirmed that all these mutations were somatic. Details of the resulting changes in EGFR protein as a consequence of these mutations are illustrated in Fig. 1.

All of the 52 deletion mutations occurred around codons 746–750 in exon 19. About half (25 of 52) of deletion mutations were simple deletions of five amino acid residues ELREA from codon 746 to 750. However, 22 deletions were coupled with point mutations or insertions, yielding various changes in amino acid sequences as shown in Fig. 1. It is noted that, in all cases, such alterations were in-frame. Forty-six of the 54 point mutations were from a Tt to an G transversion at the second nucleotide of codon 858 in exon 21 resulting in substitution of leucine with arginine residue. Four of the point mutations occurred at codon 719 in exon 18. We noted that one tumor with a mutation at codon 719 and three tumors with mutations at codon 858 had another mutation occurring at codons 709, 768, 776, and 790, respectively. For rare mutations (all 5 insertions, E709H, T790M, R776C, S768I, L858R), we resequenced and confirmed that these mutations were actually present. In summary, 52 of the 111 (47%) EGFR mutations were deletions around codons 746–750 and 49 (44%) were L858R, altogether accounting for 91% of all of the EGFR mutations found. The four major classes of mutations (i.e., deletions, L858R, mutations at codon 719, duplications/insertions) never occurred simultaneously. Furthermore, it is of note that only mutant sequences were present in chromatograms in 19 of 52 deletions, 13 of 46 in L858R, and 1 of 4 codon 719 mutations.

**Relationship between EGFR Mutations and Clinical-Pathologic Features.** EGFR mutations were significantly more frequent in females (59%) than males (26%; P < 0.001), in never-smokers (66%) than ever-smokers (22%; P < 0.001), and in patients with adenocarcinomas (49%) than in those with nonadenocarcinomas (2%; P < 0.001). There was only one patient with an EGFR mutation of 53 nonadenocarcinoma patients. This patient was a 61-year-old male with adenosquamous carcinoma. Because female patients tended to be never-smokers and were likely to have adenocarcinoma, we did logistic regression analysis to determine which of these three variables independently contributed to the EGFR mutations. The result suggested that smoking status and adenocarcinoma histology independently affected EGFR mutations whereas female gender did not (smoking status, odds ratio 3.949, P < 0.001; histologic type, odds ratio 27.486, P = 0.0013; gender odds ratio 0.9917).

**Further Analysis of Patients with Adenocarcinoma.** EGFR mutations were found almost exclusively in adenocarcinomas with only one exception; hence, we did more detailed analysis limited to this subset of patients (Table 1). EGFR mutations were also significantly frequent in female, nonsmoking patients. When we divided ever-smokers into 3 categories depending on smoke exposure, there was a trend that the higher the exposure, the lower the incidence of EGFR mutations. EGFR mutations were significantly more frequent in well differentiated adenocarcinomas (58%) than in poorly differentiated adenocarcinomas (30%; P < 0.001). There were five bronchioloalveolar cell carcinomas (BAC) in our cohort, of which three harbored EGFR mutations (60%), according to the World Health Organization classification of lung cancers (which states that BAC is a true noninvasive cancer without stromal or pleural invasion; ref. 12). It seemed that EGFR mutations were associated neither with age of the patients nor with stage of diseases. There was no difference in incidence of EGFR mutations between sexes in patients of age 50 (average age of menopause in Japan) or younger, although the number of patients of this age group was small (2 of 7 males, 2 of 7 females).

Our preliminary study indicated that patients with EGFR mutations survived for a longer period after gefitinib treatment than those without EGFR mutations. However, EGFR mutations also might have prognostic impact on patients with pulmonary adenocarcinoma, even when the patients were not exposed to gefitinib because EGFR differentiation.

---

mutations defined subsets of pulmonary adenocarcinoma with distinct features as described. Therefore, we did survival analysis in patients excluding those who were treated with gefitinib when they had recurrent diseases. The Kaplan-Meier curve (Fig. 2) indicated that EGFR mutations did not affect prognosis of the patients ($P = 0.9933$), although the follow up period was relatively short (median follow up, 788 days).

**KRAS** and **TP53** Gene Mutational Analysis. Of 224 patients with adenocarcinoma, KRAS and TP53 data were available for 196 and 192 patients, respectively. KRAS mutations were present in 26 of 196 patients (13%; 22 at codon 12, 1 at codon 13, and 3 at codon 61). TP53 mutations were present in 79 of 192 (41%). KRAS and TP53 mutations were significantly more frequent in ever-smokers, respectively [20% versus 6% for KRAS ($P = 0.0054$) and 54% versus 30% for TP53 ($P < 0.001$)]. Interestingly, EGFR mutations were never found in tumors with KRAS mutations, showing a mutually exclusive relationship. By contrast, EGFR mutations and TP53 mutations seemed to occur independently. Figure 3 shows the relationship among the three mutations by a Venn diagram in 192 patients in whom information about the status of these three genes was available.

TP53 mutations seemed more widely distributed in tumors without EGFR mutations (Fig. 4). Of seven mutations either at codon 157, 248, or 273 in which strong and selective adduct formation of benzo(a)pyrene diol epoxide, one of the major tobacco carcinogens, occurs (13), six were in tumors without EGFR mutations (Fig. 3). Furthermore, of 16 mutations caused by a G to a T transversions characteristic of mutations caused by aromatic polycyclic hydrocarbons (14), 15 were in tumors without EGFR mutations (Fig. 3).

**DISCUSSION**

Adenocarcinoma is the most predominant histologic subtype, and its incidence is increasing in Japan. Registration of resected lung cancer in Niigata prefecture, Japan, revealed that the incidence of adenocarcinoma is 71% of 1211 patients operated on from 2001 to 2002 (15). In our institution, adenocarcinoma accounted for 54% of 975 patients who were operated on from 1996 through 1995, 69% of 522 from 1996 through 2000, and 76% of 407 from 2001 through 2003. Considerable evidence indicates that the EGFR pathway also plays an important role in both the pathogenesis and the progression of lung cancer (1).

We found that 40% of 277 unselected patients with lung cancer carried mutations in the TK domain of the EGFR gene. More than 90% of the mutations were either deletions around codons 746–750 in

---

**Table 1** Relationship between EGFR mutations and clinical and pathologic features in a subset of patients with adenocarcinoma

<table>
<thead>
<tr>
<th>Variables</th>
<th>Category</th>
<th>EGFR</th>
<th>Wild-type</th>
<th>$P$</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td></td>
<td>110 (49)</td>
<td>114</td>
<td></td>
</tr>
<tr>
<td>Gender</td>
<td>Male</td>
<td>40 (36)</td>
<td>71</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>Female</td>
<td>70 (62)</td>
<td>43</td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>$\leq$64</td>
<td>51 (46)</td>
<td>60</td>
<td>0.3481</td>
</tr>
<tr>
<td></td>
<td>$&gt;64$</td>
<td>59 (52)</td>
<td>54</td>
<td></td>
</tr>
<tr>
<td>Smoking status</td>
<td>Never-smoker</td>
<td>76 (68)</td>
<td>36</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>Ever-smoker</td>
<td>34 (31)</td>
<td>78</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Pack years $&lt;$20</td>
<td>11 (55)</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td></td>
<td>20–50</td>
<td>15 (27)</td>
<td>40</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$&gt;$20</td>
<td>8 (22)</td>
<td>29</td>
<td></td>
</tr>
<tr>
<td>Differentiation</td>
<td>Well to moderately differentiated</td>
<td>89 (58)</td>
<td>65</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>Poorly differentiated</td>
<td>21 (30)</td>
<td>49</td>
<td></td>
</tr>
<tr>
<td>Stage</td>
<td>IA and IB</td>
<td>69 (50)</td>
<td>70</td>
<td>0.8383</td>
</tr>
<tr>
<td></td>
<td>IIA through IIV</td>
<td>41 (48)</td>
<td>44</td>
<td></td>
</tr>
<tr>
<td>Survival</td>
<td>3-year survival rate</td>
<td>86%</td>
<td>91%</td>
<td>0.9933</td>
</tr>
<tr>
<td>KRAS mutation</td>
<td>Mutated</td>
<td>0 (0)</td>
<td>26</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>Wild-type</td>
<td>97 (57)</td>
<td>73</td>
<td></td>
</tr>
<tr>
<td>TP53 mutation</td>
<td>Mutated</td>
<td>37 (47)</td>
<td>42</td>
<td>0.4634</td>
</tr>
<tr>
<td></td>
<td>Wild-type</td>
<td>59 (52)</td>
<td>54</td>
<td></td>
</tr>
</tbody>
</table>

There were five BACs in our cohort, of which three harbored EGFR mutations.
EGFR MUTATIONS IN LUNG CANCER

Adenocarcinomas with EGFR mutations

Adenocarcinomas without EGFR mutations

Fig. 4. Distribution of *TP53* gene mutations in adenocarcinomas without *EGFR* mutations (n = 42) or with *EGFR* mutations (n = 32). Numbers below show codons of exon boundaries. Asterisks show codons 157, 248, and 273, where strong and selective benzo(e)pyrene diol epoxide adduct formation is reported to occur (13). White circles indicate where *TP53* mutations were caused by a G to a T transversion.

exon 19 or L858R in exon 21, which all flank the ATP-binding pocket that is important for TK activity (8, 9). We also noted that in about 30% of the cases with *EGFR* mutations, only bands derived from mutant allele were detected on chromatogram. This is somewhat puzzling considering the heterozygous nature of the *EGFR* mutations reported thus far (8, 9) and the presence of stromal cells in resected tumor specimens. This finding may suggest that loss of wild-type alleles or amplification of mutant alleles accompanied with mutations in these cases, as indicated by Minna et al. (16).

*EGFR* mutations were almost exclusively present in adenocarcinoma. Mutations were more prevalent in females and nonsmokers, confirming and extending the results of previous reports (8, 9). It is noteworthy that these characteristics and Japanese ethnicity are all predictors of gefitinib sensitivity at least by univariate analysis (4, 5). Multivariate analysis suggested that nonsmoking status and adenocarcinoma histology independently contributed to *EGFR* mutations but female gender did not. The fact that premenopausal women did not show higher incidence of *EGFR* mutations further suggested that apparent difference between female and male was caused by a difference in lifestyle including smoking habit rather than involvement of sexual environment.

Previously described genetic alterations in lung cancer are almost always more frequent in smokers than nonsmokers. For example, mutations of the *TP53* gene (17), *KRAS* genes (18), or deletion of the short arm of chromosome 3 (19) are known to be more frequent in smokers, as was the case in the present study for the first two. A plausible explanation for the reason why *EGFR* mutations are associated with nonsmoking status and adenocarcinoma histology independently contributed to *EGFR* mutations but female gender did not. The fact that premenopausal women did not show higher incidence of *EGFR* mutations further suggested that apparent difference between female and male was caused by a difference in lifestyle including smoking habit rather than involvement of sexual environment.

We were able to confirm higher incidence of *EGFR* mutations in Japanese patients. Lynch et al. found *EGFR* mutation in 2 of 25 unselected United States patients (9), and Paez et al. (8) did so in 1 of 61 United States patients and 15 of 58 Japanese patients. The reason for this marked difference between Japanese and United States patients is not very clear. However, difference in incidence of nonsmoking patients between Japanese and American female patients with lung cancer may partly account for this. In our cohort, 83% of female patients and 10% of male patients were never-smokers. This trend is common in Japan. For example, Toyooka et al. (22) and Minami et al. (23) reported that the proportion of never-smoking women in lung cancer patients is 96% and 75%, respectively. This makes quite a contrast with the fact that only 15% of 706 United States female and 6% of 1,347 male patients with lung cancer are never-smokers (24).

We found that *EGFR* mutations and *KRAS* mutations known to play an important role in pathogenesis of adenocarcinoma of the lung (25) were strictly mutually exclusive, reminding us of a similar exclusionary relationship between retinoblastoma and p16 inactivation in lung cancer (26). This finding may be explained by the fact that the KRAS-mitogen-activated protein kinase pathway is one of the downstream signaling pathways of EGFR (1). Because it has been shown that L858R and delL747-P753ins S are activating mutations that result in markedly increased phosphorylation of EGFR when EGF was added (8, 9), tumors with *KRAS* mutations that already have activated further downstream effectors do not need to have *EGFR* mutations. The high incidence of *EGFR* mutations in lung adenocarcinomas may explain why *KRAS* mutations are lower in Japanese than in Caucasian patients. In the present study, *KRAS* mutations were found in 13% of adenocarcinomas, whereas they were present in 33% of Dutch cases (25). This may be also at least partially attributable to the difference in smoking status, because *KRAS* mutations were more frequent in smokers as reported previously (18). In contrast, the incidence of *TP53* mutations was not associated with *EGFR* mutations, although *TP53* mutations also occurred more frequently in smokers (17). However, *TP53* mutations in tumors without *EGFR* mutations showed characteristics of mutations caused by tobacco carcinogens in terms of sites or base substitution patterns (13, 14).

We also noted that well to moderately differentiated adenocarcinomas had a significantly higher incidence of *EGFR* mutations than poorly differentiated ones. This observation might be relevant to the fact that adenocarcinomas showing BAC feature show higher sensitivity to gefitinib (27). However, when we used the strict criteria as stated by the World Health Organization Classification of lung tumors (12), our cohort included only five BAC, of which three had *EGFR* mutations. Unfortunately, these strict criteria are not applied by many pathologists, leading to considerable confusion between BAC and adenocarcinoma with BAC features in the literature. Alternatively, we proposed terminal respiratory unit type adenocarcinoma that is characterized by morphological resemblance to type II pneumocytes, Clara cells, and/or bronchioles as well as expression of thyroid transcription factor-1 and surfactant proprotein B (refs. 28, 29). In the World Health Organization classification, most nonmucinous bronchioloalveolar, mixed bronchioloalveolar and acinar subtypes, and some papillary subtypes belong to the terminal respiratory unit type adenocarcinoma (28, 29). We found that most adenocarcinoma with *EGFR* mutations were categorized into terminal respiratory unit type adenocarcinoma.6

EGFR mutations were not associated with stage of disease, suggesting that EGFR mutations occur relatively early in clinical course and are associated with pathogenesis of adenocarcinoma rather than progression.

In conclusion, we found a high incidence of EGFR mutations in Japanese patients with pulmonary adenocarcinoma, especially in those who never smoked. EGFR mutations were never present in tumors with KRAS mutations, indicating possibilities of genotype-oriented approach for pulmonary adenocarcinoma.

ACKNOWLEDGMENTS

The authors thank Kaori Hayashi-Hirano for excellent technical assistance in molecular analysis of tumors and Ryuzo Ohno, President of Aichi Cancer Center for special encouragement and support.

REFERENCES

Mutations of the *Epidermal Growth Factor Receptor* Gene in Lung Cancer: Biological and Clinical Implications

Takayuki Kosaka, Yasushi Yatabe, Hideki Endoh, et al.

*Cancer Res* 2004;64:8919-8923.