The Impact of Sex and Smoking Status on the Mutational Spectrum of Epidermal Growth Factor Receptor Gene in Non–small Cell Lung Cancer

Shinichi Toyooka,[1] Keitaro Matsuo,[3] Hisayuki Shigematsu,[1] Takayuki Kosaka,[4] Masaki Tokumo,[1] Yasushi Yatabe,[5] Syuji Ichihara,[1] Michio Inukai,[1] Hiroshi Suehisa,[1] Junichi Soh,[1] Katsuyuki Kiura,[2] Kwun M. Fong,[6] Huei Lee,[7] Ignacio I. Wistuba,[8] Adi F. Gazdar,[9,10] Tetsuya Mitsudomi,[4] and Hiroshi Date[1]

Abstract

Purpose: Mutation of epidermal growth factor receptor (EGFR) gene has been reported to be present in non–small cell lung cancer (NSCLC) and significantly associated with female sex and never-smoking status. In this study, we extensively investigated the impact of sex and smoking on the *EGFR* mutation.

Experimental Design: We examined *EGFR* exons 18 to 21 status in 1,467 NSCLC patients by direct sequencing to study the impact of sex and smoking status on the *EGFR* mutational spectrum.

Results: Among 1,467 patients, 197 mutations were found at exon 19, 176 at exon 21, 21 at exon 18, and 24 at exon 20. To examine the independent effect of sex and smoking, the mutational status of each exon was compared between smokers and never smokers in each sex and between males and females stratified by smoking status. In females, exon 19 ($P = 0.001$) and exon 21 ($P < 0.001$) mutations were significantly less frequent in ever smokers compared with never smokers. In males, exon 19 ($P < 0.001$), exon 21 ($P < 0.001$), and exon 18 ($P = 0.003$) mutations were significantly less frequent in ever smokers compared with never smokers. In analysis stratified by smoking, there was no difference in sex among never smokers. However, exon 19 mutations were significantly less frequent in males compared with females among ever smokers ($P = 0.003$). In addition, the interactive effect of male sex and ever smoking status significantly decreased the frequency of exon 19 mutations ($P = 0.047$) when female never smoker was set as a reference.

Conclusion: Both sex and smoking status could influence the *EGFR* mutational spectrum. Our findings suggest that individual *EGFR* exons may have differing susceptibilities for mutagenesis.

Authors' Affiliations: Departments of [1]Cancer and Thoracic Surgery and [2]Hematology, Oncology and Respiratory Medicine, Graduate School of Medicine, Dentistry and Pharmaceutical Sciences, Okayama University, Okayama, Japan; Divisions of [3]Epidemiology and Prevention, [4]Thoracic Surgery, and [5]Pathology, Aichi Cancer Center Research Institute, Nagoya, Japan; [6]The Prince Charles Hospital, Brisbane, Australia; [7]Institute of Medical and Molecular Toxicology, Chung Shan Medical University, Taichung, Taiwan; [8]Department of Pathology, University of Texas M. D. Anderson Cancer Center, Houston, Texas; and [9]Hamon Center for Therapeutic Oncology Research and [10]Department of Pathology, University of Texas Southwestern Medical Center, Dallas, Texas

Received 1/26/07; revised 6/3/07; accepted 7/12/07.

Grant support: Grant-in-Aid for Scientific Research from the Ministry of Education, Science, Sports, Culture and Technology of Japan (18790993 for S. Toyooka), AstraZeneca Research Grant (S. Toyooka), and University of Texas Specialized Program for Research Excellence in Lung Cancer, National Cancer Institute, Bethesda, MD (for A.F. Gazdar and I.I. Wistuba).

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.

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

Requests for reprints: Shinichi Toyooka, Department of Cancer and Thoracic Surgery, Graduate School of Medicine, Dentistry and Pharmaceutical Sciences, Okayama University, 2-5-1 Shikata-cho, Okayama 700-8558, Japan. Phone: 81-86-235-7265; Fax: 81-86-235-7269; E-mail: toyooka@md.okayama-u.ac.jp.

doi:10.1158/1078-0432.CCR-07-0216

The epidermal growth factor receptor (EGFR) is a receptor tyrosine kinase (TK) that is frequently overexpressed in malignant tumors, and its signal transduction cascade leads to a multitude of effects, including cell proliferation, cell differentiation, angiogenesis, metastasis, and antiapoptosis (1, 2). Recent findings have shown that *EGFR* mutations are exclusively present in lung cancer and showed significant association with sensitivity of EGFR TK inhibitor (EGFR-TKI), including gefitinib and erlotinib (3–6). Several clinicopathologic factors were identified to be related to the frequency of *EGFR* mutations, including adenocarcinoma histology, female sex, and never-smoking status and East Asian ethnicity (6). Most mutations involved nucleotides of exons 18 to 21 in the TK domain, especially deletions in exon 19 and the L858R point mutation in exon 21. We previously reported that exon 19 deletions and exon 21 point mutations were frequent in never smoker compared with ever smoker and in female sex compared with male sex, but there was no bias for sex and smoking status in exon 20 insertion mutation (7). More recently, we found the lack of sex effect for the EGFR-TKI resistance related *EGFR* T790M mutation located in exon 20 in a minor clone of the tumor (8). These findings suggest the sex, smoking status, or their interactions influence the *EGFR*

Table 1. Patient characteristics

Variables	Japan	United States	Taiwan	Australia
Age				
Median (range)	65 (26-91)	67 (40-90)	65 (26-87)	63 (38-83)
Sex				
Female	368	74	73	36
Male	694	86	71	65
Smoking status				
Never smoker	370	40	105	25
Ever smoker	692	120	39	76
Histology				
Adenocarcinoma	773	97	94	47
Non-adenocarcinoma	289	63	50	54

mutational spectrum in an exon-specific manner in non–small cell lung cancer (NSCLC) cases. In addition, the clinical impact of EGFR-TKIs on patients with various mutations was different according to the type of *EGFR* mutant, indicating that biological significance of mutations is not equal according to different mutant types (6, 9–11). These observations indicate that a better understanding of factors influencing the *EGFR* mutational spectrum may be important to understand the tumor pathogenesis and response for treatment of NSCLC.

We have independently analyzed and reported the relationship between *EGFR* mutations and clinicopathologic factors, but a large-scale study is necessary to glean novel and important features of *EGFR* mutations in NSCLC. Because no single study can analyze more than a modest number of mutations, we combined our data to build up database for analysis. In this study, we examined *EGFR* exons 18 to 21 mutation status in 1,467 NSCLC patients to investigate the impact of sex factor, smoking factor, or interacted factor of sex and smoking status on the *EGFR* mutational spectrum in NSCLC.

Table 2. The *EGFR* mutational spectrum and the effect of interaction between sex and smoking

A. Total

EGFR status	Female				
	Never smoker		Ever smoker		
	n (%)	OR (95% CI)	*n* (%)	OR (95% CI)	*P*
All cases	420		140		
Wild type	181 (60.9)		116 (39.1)		
Exon 19	119 (89.5)	1.00 (reference)	14 (10.5)	0.34 (0.18-0.65)	0.001
Exon 21	100 (94.3)	1.00 (reference)	6 (5.7)	0.18 (0.07-0.44)	<0.001
Exon 18	8 (80.0)	1.00 (reference)	2 (20.0)	0.50 (0.10-2.54)	0.399
Exon 20	12 (85.7)	1.00 (reference)	2 (14.3)	0.48 (0.10-2.36)	0.366
B. East Asia					
All cases	378		70		
Wild type	155 (74.5)		53 (25.5)		
Exon 19	113 (92.6)	1.00 (reference)	9 (7.4)	0.30 (0.14-0.66)	0.003
Exon 21	93 (94.9)	1.00 (reference)	5 (5.1)	0.21 (0.08-0.56)	0.002
Exon 18	6 (75.0)	1.00 (reference)	2 (25.0)	1.18 (0.22-6.33)	0.855
Exon 20	11 (91.7)	1.00 (reference)	1 (8.3)	0.38 (0.04-3.11)	0.370
C. United States and Australia					
All cases	42		70		
Wild type	26 (29.2)		63 (70.8)		
Exon 19	6 (54.6)	1.00 (reference)	5 (45.4)	0.37 (0.10-1.41)	0.146
Exon 21	7 (88.9)	1.00 (reference)	1 (11.1)	0.07 (0.01-0.60)	0.016
Exon 18	2 (100)	1.00 (reference)	0 (0)	NA	
Exon 20	1 (50.0)	1.00 (reference)	1 (50.0)	0.30 (0.02-5.37)	0.415

NOTE: For A, model included age, region (East Asia or United States and Australia), institution, histology (adenocarcinoma or non-adenocarcinoma) and sex, smoking, and interaction term between sex and smoking. For B and C, model included age, institution, histology (adenocarcinoma or non-adenocarcinoma), and sex, smoking, and interaction term between sex and smoking. Comparisons were made for each exon mutated with reference to those without mutation. Interaction OR indicates how many times if being male ever smoker impacts on having mutation compared with female never smoker.
Abbreviation: NA, not available.

Materials and Methods

NSCLC samples. Surgically resected specimens of 1,467 NSCLC patients were collected from 1,062 patients in Japan (415 from Aichi Cancer Center Hospital, 413 from Okayama University Hospital, and 234 from Chiba University Hospital), 160 patients in the United States (M.D. Anderson Cancer Center, Houston, TX), 144 patients from Taiwan (Veterans General Hospital, Taichung, Taiwan), and 101 patients from Australia (Prince Charles Hospital, Brisbane, Australia). Regarding ethnicity, all cases of Japan and Taiwan were East Asian, and the majority of cases from the United States and Australia were Caucasian (6). The details of patient characteristics are shown in Table 1. A total of 1,467 patients consisted of 916 males and 551 females, 927 ever smokers and 540 never smokers, and 1,011 adenocarcinoma, 377 squamous cell carcinoma, 41 large cell carcinoma, 23 adenosquamous carcinoma, and 15 other types of NSCLC. Never smokers were defined as those with lifetime exposure of 100 cigarettes or less. Ever smokers were defined as current or former smoker with lifetime exposure of more than 100 cigarettes. Specimens of 1,437 patients were from frozen samples, and those of 30 patients were from paraffin-embedded specimens (6, 12, 13). All cases from the United States and Australia, 93 cases from Taiwan, and 666 cases from Japan were previously reported (6, 12, 13). In this study, 137 cases from the Aichi Cancer Center Hospital, 259 cases from the Okayama University Hospital, 51 patients from the Veterans General Hospital, Taiwan, were added to the previous study. Corresponding nonmalignant lung tissues were examined in 921 patients. Institutional Review Board permission and informed consent were obtained at each collection site.

DNA extraction and sequencing analysis. Genomic DNAs were isolated by digestion with proteinase K followed by phenol-chloroform (1:1) extraction and ethanol precipitation from frozen specimen (14) and by DEXPAT (TaKaRa) from paraffin-embedded tissues following the manufacturer's instructions. For specimens analyzed at the Aichi Cancer Center, the total RNAs were isolated from frozen tissues of the tumor specimens using the RNAeasy kit (Qiagen). DNA-based direct sequencing for exons 18 to 21 of *EGFR* was done using same PCR primers and conditions at the University of Texas Southwestern Medical Center and Okayama University, except paraffin samples (6, 13). The RNA-based analysis using one-step reverse transcription-PCR for exons 18 to 21 region using Qiagen OneStep reverse transcription-PCR kit (Qiagen) was done at the Aichi Cancer Center (12).

Statistical analyses. All the statistical analyses were conducted by STATA 9.2 (Stata Corp.). A χ^2 test was applied to test the frequencies if appropriate. A multivariable logistic regression model was applied to estimate odds ratios (OR) and their confidence intervals (CI). Age, sex, histology (adenocarcinoma or non-adenocarcinoma) region (East Asia or United States and Australia), and institutions were adjusted in the model. A model evaluated sex (male versus female) and smoking status (ever versus never) as binary variable, and a model including the interaction between them was also evaluated. We regarded P values <0.05 as statistically significant.

Results

Detection of EGFR mutations in NSCLC patients. We examined the *EGFR* mutation status in 1,469 lesions from

Table 2. The *EGFR* mutational spectrum and the effect of interaction between sex and smoking (Cont'd)

A. Total

Male						Interaction OR (95% CI)	*P*
Never smoker			Ever smoker				
n (%)	OR (95% CI)	*P*	*n* (%)	OR (95% CI)	*P*		
130			788				
73 (9.6)			690 (90.4)			—	
29 (45.3)	0.84 (0.49-1.42)	0.505	35 (54.7)	0.12 (0.08-0.18)	<0.001	0.41 (0.17-0.99)	0.047
23 (32.9)	0.79 (0.45-1.39)	0.408	47 (67.1)	0.18 (0.12-0.26)	<0.001	1.23 (0.42-3.60)	0.706
4 (36.4)	1.58 (0.43-5.81)	0.487	7 (63.6)	0.25 (0.08-0.74)	0.012	0.32 (0.04-2.62)	0.287
1 (10.0)	0.23 (0.03-1.83)	0.164	9 (90.0)	0.31 (0.13-0.77)	0.012	2.86 (0.20-40.4)	0.437

B. East Asia

104			662				
52 (8.4)			566 (91.6)			—	
27 (44.3)	0.90 (0.51-1.58)	0.712	34 (55.7)	0.12 (0.08-0.199)	<0.001	0.45 (0.17-1.20)	0.109
22 (32.4)	0.88 (0.48-1.59)	0.661	46 (67.7)	0.19 (0.12-0.28)	<0.001	1.02 (0.32-3.22)	0.975
2 (22.2)	1.28 (0.23-6.97)	0.775	7 (77.8)	0.36 (0.11-1.15)	0.085	0.24 (0.02-2.54)	0.237
1 (10.0)	0.27 (0.03-2.24)	0.227	9 (90.0)	0.37 (0.15-0.93)	0.034	3.51 (0.18-68.7)	0.409

C. United States and Australia

26			126				
21 (14.5)			124 (85.5)			—	
2 (66.7)	0.47 (0.08-2.74)	0.403	1 (33.3)	0.05 (0.01-0.44)	0.007	0.28 (0.02-4.61)	0.372
1 (50.0)	0.22 (0.02-2.03)	0.179	1 (50.0)	0.05 (0.01-0.41)	0.006	3.23 (0.09-117.4)	0.522
2 (100)	1.47 (0.17-12.7)	0.726	0 (0)	NA		NA	
0 (0)	NA		0 (0)	NA		NA	

Table 3. The impact of smoking or sex status on the *EGFR* mutation spectrum

A. The impact of smoking status according to sex status

EGFR status	Never smoker		Ever smoker		
	n (%)	OR (95% CI)	*n* (%)	OR (95% CI)	*P*
Female					
All cases	420		140		
Wild type	181 (61.6)		116 (38.4)		
Exon 19	119 (89.5)	1.00 (reference)	14 (10.5)	0.31 (0.16-0.60)	0.001
Exon 21	100 (94.4)	1.00 (reference)	6 (5.6)	0.16 (0.06-0.38)	<0.001
Exon 18	8 (81.8)	1.00 (reference)	2 (18.2)	0.66 (0.12-3.65)	0.630
Exon 20	12 (85.7)	1.00 (reference)	2 (14.3)	0.34 (0.06-1.76)	0.197
Male					
All cases	130		788		
Wild type	73 (9.8)		690 (90.2)		
Exon 19	29 (45.3)	1.00 (reference)	35 (54.7)	0.14 (0.08-0.25)	<0.001
Exon 21	23 (32.9)	1.00 (reference)	47 (67.1)	0.21 (0.11-0.40)	<0.001
Exon 18	4 (41.7)	1.00 (reference)	7 (58.3)	0.13 (0.03-0.51)	0.003
Exon 20	1 (10.0)	1.00 (reference)	9 (90.0)	1.54 (0.18-13.2)	0.693

B. The impact of sex status according to smoking status

EGFR status	Female		Male		
	n (%)	OR (95% CI)	*n* (%)	OR (95% CI)	*P*
Never smoker					
All cases	420		130		
Wild type	181 (71.3)		73 (28.7)		
Exon 19	119 (80.4)	1.00 (reference)	29 (19.6)	0.85 (0.50-1.45)	0.553
Exon 21	100 (81.3)	1.00 (reference)	23 (18.7)	0.76 (0.43-1.34)	0.344
Exon 18	8 (66.7)	1.00 (reference)	4 (33.3)	1.28 (0.35-4.77)	0.710
Exon 20	12 (92.3)	1.00 (reference)	1 (7.7)	0.21 (0.03-1.71)	0.145
Ever smoker					
All cases	140		788		
Wild type	116 (38.4)		690 (90.2)		
Exon 19	14 (28.6)	1.00 (reference)	35 (71.4)	0.34 (0.16-0.70)	0.003
Exon 21	6 (11.3)	1.00 (reference)	47 (88.7)	0.88 (0.35-2.22)	0.785
Exon 18	2 (22.2)	1.00 (reference)	7 (77.8)	0.28 (0.05-1.50)	0.138
Exon 20	2 (18.2)	1.00 (reference)	9 (81.8)	0.62 (0.12-3.20)	0.567

NOTE: Model included age, area (East Asia or United States and Australia), histology (adenocarcinoma or non-adenocarcinoma), and institution.

1,467 NSCLC patients and found 418 mutations in 407 patients. Seven patients had double mutations, and one patient had a triple mutation in a tumor. One patient had three tumors, and each of them harbored a mutation [one exon 19 deletion and two exon 21 mutations (L858R mutations)]. These multiple mutant cases were regarded as each single mutation; thus, our cohort was assumed to have 1,478 cases. There were 197 mutations in exon 19, 176 in exon 21, 21 in exon 18, and 24 in exon 20. Mutations of exon 19 consisted of 196 in-frame deletions involving three to eight codons around codons 747 to 749 (Leu-Arg-Glu sequence) and one point mutation (D761Y; ref. 15). Mutation of exon 20 consisted of 22 insertion and two point mutations (T790M; ref. 16). Other mutations in each exon were point mutations. *EGFR* mutations were present in 328 (30.9%) patients (336 mutations) of 1,062 patients from Japan, 15 (9.4%) patients (15 mutations) of 160 patients from the United States, 52 (36.1%) patients (52 mutations) of 144 patients from Taiwan, and 12 (11.9%) patients (15 mutations) of 101 patients from Australia. Because *EGFR* mutation was predominantly present in adenocarcinoma and experimental data also showed a close relationship between *EGFR* mutations and adenocarcinoma histology, the frequency of *EGFR* mutations in adenocarcinoma was also examined (6, 17). *EGFR* mutations were present in 317 (41.0%) patients (324 mutations) of 773 patients in Japan, 15 (15.5%) patients (15 mutations) of 97 patients in the United States, 50 (53.2%) patients (50 mutations) of 94 patients in Taiwan, and 9 (19.1%) patients (11 mutations) of 47 in Australia. These results showed that there was a similar tendency in the prevalence of *EGFR* mutations in Japan and Taiwan (East Asia) and in the United States and Australia. The *EGFR* mutation was not found in 921 nonmalignant samples examined (6, 12, 13).

The EGFR mutational spectrum according to sex and smoking status. We divided our cases into four groups based on sex and smoking status: 420 female never-smoker cases, 140 female ever-smoker cases, 130 male never-smoker cases, and 788 male ever-smoker cases. In females, never smokers had 119 mutations in exon 19, 100 in exon 21, 8 in exon 18, and 12 in exon 20. Female ever smokers had 14 mutations in exon 19, 6 in exon 21, 2 in exon 18, and 2 in exon 20. In males, never smokers had 29 mutations in exon 19, 23 in exon 21, 4 in exon 18, and 1 in exon 20. Male ever smokers had 35 mutations in exon 19, 47 in exon 21, 7 in exon 18, and 9 in exon 20.

The *EGFR* mutation spectrum of the total group, East Asia, and the United States and Australia cases stratified by histologic subtypes are shown in Table 2 and Supplementary Material 1.

The impact of smoking status on the mutation spectrum according to sex status. The impact of smoking status on the *EGFR* mutation spectrum was examined for each sex. Never-smoker female or male status was set as a reference. In females, exon 19 and exon 21 mutations were significantly less frequent in cases from ever smokers compared with never smokers (exon 19: OR, 0.31; 95% CI, 0.16-0.60; $P = 0.001$; exon 21: OR, 0.16; 95% CI, 0.06-0.38; $P < 0.001$), and there were no significant differences between exons 18 and 20. In males, exons 19, 21, and 18 mutations were significantly less frequent in cases from ever smokers compared with never smokers (exon 19: OR, 0.14; 95% CI, 0.08-0.25; $P < 0.001$; exon 21: OR, 0.21; 95% CI, 0.11-0.40; $P < 0.001$; exon 18: OR, 0.13; 95% CI, 0.03-0.51; $P = 0.003$; Table 3A).

The same analyses were done among histologic subtypes. Restricted to adenocarcinomas, we had similar results to those of total cases (analyses of histologic subtypes were adjusted in total cases as described). In females, exons 19 and 21 mutations were significantly less frequent in cases from ever smokers compared with never smokers (exon 19: OR, 0.33; 95% CI, 0.17-0.65; $P = 0.001$; exon 21: OR, 0.14; 95% CI, 0.05-0.37; $P < 0.001$), and there were no significant differences between exons 18 and 20. In males, exons 19, 21, and 18 mutations were significantly less frequent in cases from ever smokers compared with never smokers (exon 19: OR, 0.13; 95% CI, 0.07-0.24; $P < 0.001$; exon 21: OR, 0.21; 95% CI, 0.11-0.39; $P < 0.001$; exon 18: OR, 0.16; 95% CI, 0.03-0.96; $P = 0.045$; Supplementary Material 2). In non-adenocarcinoma, whereas the number of cases were limited, exon 18 mutations were significantly less frequent in cases from ever smokers compared with never smokers in males (OR, 0.10; 95% CI, 0.01-0.82; $P = 0.032$; Supplementary Material 2).

The impact of sex difference on the mutational spectrum according to smoking status. The impact of sex status on the *EGFR* mutational spectrum was examined in ever smoker and never smoker cases separately. Female never-smoker or ever-smoker status was set as a reference. In never smokers, there was no difference between female and male cases on each exon. In ever smokers, exon 19 mutations were significantly less frequent in male compared with female cases (OR, 0.34; 95% CI, 0.16-0.70; $P = 0.003$; Table 3B). Restricted to adenocarcinomas, similar results were obtained; in ever smokers, exon 19 mutations were significantly less frequent in males compared with females (OR, 0.32; 95% CI, 0.15-0.67; $P = 0.003$; Supplementary Material 2). In non-adenocarcinoma, the number of cases was too limited to analyze.

The impact of interaction between sex and smoking status on the mutational spectrum. We also examined the impact of interaction between sex and smoking status on each exon when female never smoker was set as a reference. Regarding exon 19 mutations, the effect of male sex and smoking status was significantly decreased compared with an expectation from the independent effect of male sex and smoking status (Interaction OR, 0.41; 95% CI, 0.17-0.99; $P = 0.047$; Table 2A). On the other hand, there was no significant interactive effect of smoking and sex on exon 21, 18, and 20 mutations. Stratified by histologic subtypes, adenocarcinoma cases showed similar results to total cases that the effect of male sex and smoking status was significantly decreased in exon 19, as compared with an expectation from independent effect of male sex and smoking status (Interaction OR, 0.36; 95% CI, 0.15-0.89; $P = 0.027$; Supplementary Material 1). There was no significant interactive effect on exons 21, 18, and 20 mutations. In non-adenocarcinoma cases, the number of cases was too small to analyze the impact of interaction between sex and smoking status on each exon.

The impact of smoking and sex differences on the mutational types. We analyzed the relationship between sex and smoking status and types of mutation consisting of deletions, insertions, and point mutation. As shown above, cases with deletion were only present in exon 19; cases with insertion were only present in exon 20; and pointed mutation cases consisted of all cases of exon 18, one of exon 19, one of exon 20, and all of exon 21. Thus, the impact of smoking and sex differences on the mutational types were similar to that on the mutational spectrum (Supplementary Material 3). Indeed, the effect of male sex and smoking status in total cases tended to be decreased in deletion type compared with an expectation from independent effect of male sex and smoking status (Interaction OR, 0.44; 95% CI, 0.18-1.08; $P = 0.074$). Stratified by histologic subtypes, adenocarcinoma cases showed similar results to total cases that the effect of male sex and smoking status was significantly decreased in exon 19, as compared with an expectation from the independent effect of male sex and smoking status (Interaction OR, 0.39; 95% CI, 0.16-0.97; $P = 0.043$). There was no significant interactive effect on point mutation and insertion.

Discussion

Previous studies examined the effects of sex or smoking status among *EGFR* mutant cases and found that the *EGFR* L858R mutation in exon 21 was more frequent in females and never smokers as compared with exon 19 deletions (13, 18). By contrast, in this study, NSCLC cases with *EGFR* wild type were considered as a reference, and the impact of sex and smoking status on the *EGFR* mutation spectrum was evaluated to identify the different effects of sex and smoking status on mutation of each exon. Furthermore, the discovery of an interaction of sex and smoking on exon 19 mutation (or deletion mutation) implies that the pathogenesis of exon 19 mutations (which mainly consist of deletions) may be different from mutation in other exons (mainly point mutations). Nonetheless, the mechanisms of how deletion or point mutation occurs in genomic DNA are unknown.

The influence of sex and smoking status on lung cancer has been an issue of interest. Zang and Wynder have reported that the ORs for major lung cancer types at every level of exposure to tobacco smoke were higher in women than in men likely due to the higher susceptibility of women to tobacco smoking (19), and a recent review supported this observation (20). Previous molecular data indicate that cancers arising in women smokers harbored significantly more tobacco-related p53 mutations, G:C to T:A transversions, than those in male smokers (20). Furthermore, DNA adduct levels of benzo(*a*)pyrene diol epoxide (a tobacco carcinogen) in females with lung cancer were markedly greater than those arising in males, suggesting sex differences in susceptibility to DNA damage derived from environmental carcinogen exposure (21). Of note, there was no

specific association between smoking status and the type of point mutation (transversion and transition) in *EGFR* mainly because that majority of point mutation is G to T transversion at the second letter of codon 858.

The biological differences in *EGFR* mutations have also been identified in experimental models. The *in vitro* transforming activity of mutants L747-749delA750P (exon 19 deletion) and D770-771insNPG (exon 20 insertion) were higher than that of L858R (exon 21 mutation) and G719S (exon 18 mutation) mutations (9). In a transgenic mouse model, L858R mutants were associated with a more aggressive adenocarcinoma compared with exon 19 deletion mutants (delL747-S752; ref. 17).

Clinically, individual mutant types also show different outcomes following EGFR-TKI treatment. The exon 20 insertion mutant cells revealed resistance to gefitinib or erlotinib, and those of exon 18 mutant revealed intermediate sensitivity (lower than exon 19 or exon 21 mutant cells) *in vitro* (9). In patients treated with EGFR-TKIs, survival time of exon 19 mutant cases was longer than that of exon 21 mutation (10, 11, 22). These findings strengthen the role and importance of *EGFR* mutational types in the pathogenesis and treatment response of NSCLC.

In this study, we analyzed cases from four countries, Japan, the United States, Taiwan, and Australia. Our previous study showed the presence of the ethnic differences in the frequency of *EGFR* mutations (6). Besides ethnicity, there is a possibility that other factors, including differences in lifestyle, may affect the *EGFR* mutational spectrum. Thus, care is required to account for possible differences found when analyzing such multinational cohorts. Notwithstanding, much useful information may be obtained from a large-scale international database, and our study will encourage the establishment of the worldwide database for the *EGFR* mutation for common scientific use. Indeed, a recent study conducted by co-authors (K.T., T.M., Y.Y., and T.K.) using a case-control design in a single institution indicated that the impacts of smoking and sex on the risk of NSCLC might differ between *EGFR* mutant and *EGFR* wild-type tumor status (23). The results in the study basically support our finding that sex, smoking, and their interaction differentially affected the site of somatically acquired *EGFR* mutations. However, the reasons for our findings, especially the impact of sex difference on mutation spectrum, are unclear. Further study is warranted from various viewpoints including the effect of sex hormones.

In summary, we found that sex, smoking status, and the interactive effect of sex and smoking status influence the *EGFR* mutational spectrum, suggesting that individual exons of *EGFR* may have differing susceptibilities for mutagenesis. In addition, establishment of a large-scale database for *EGFR* mutations may be a useful resource to study the role of *EGFR* mutations for understanding the pathogenesis and novel therapeutic strategies for treating and preventing lung cancer.

Acknowledgments

We thank Makiko Tabata, Cancer and Thoracic Surgery, and Yukinari Isomoto, Central Research Laboratory, Graduate School of Medicine and Dentistry, Okayama University, for excellent technical support. We thank Dr. Makoto Suzuki, Department of Thoracic Surgery, Chiba University, Chiba, Japan, for providing some specimens.

References

1. Rusch V, Baselga J, Cordon-Cardo C, et al. Differential expression of the epidermal growth factor receptor and its ligands in primary non – small cell lung cancers and adjacent benign lung. Cancer Res 1993;53:2379–85.
2. Ciardiello F, De Vita F, Orditura M, Tortora G. The role of EGFR inhibitors in nonsmall cell lung cancer. Curr Opin Oncol 2004;16:130–5.
3. Lynch TJ, Bell DW, Sordella R, et al. Activating mutations in the epidermal growth factor receptor underlying responsiveness of non – small-cell lung cancer to gefitinib. N Engl J Med 2004;350:2129–39.
4. Paez JG, Janne PA, Lee JC, et al. EGFR mutations in lung cancer: correlation with clinical response to gefitinib therapy. Science 2004;304:1497–500.
5. Pao W, Miller V, Zakowski M, et al. EGF receptor gene mutations are common in lung cancers from "never smokers" and are associated with sensitivity of tumors to gefitinib and erlotinib. Proc Natl Acad Sci U S A 2004;101:13306–11.
6. Shigematsu H, Lin L, Takahashi T, et al. Clinical and biological features associated with epidermal growth factor receptor gene mutations in lung cancers. J Natl Cancer Inst 2005;97:339–46.
7. Shigematsu H, Toyooka S, Suzuki M. The need for an individual approach to lung cancer treatment. PLoS Med 2006;3:560–1.
8. Inukai M, Toyooka S, Ito S, et al. Presence of epidermal growth factor receptor gene T790M mutation as a minor clone in non – small cell lung cancer. Cancer Res 2006;66:7854–8.
9. Greulich H, Chen TH, Feng W, et al. Oncogenic transformation by inhibitor-sensitive and -resistant EGFR mutants. PLoS Med 2005;2:e313.
10. Jackman DM, Yeap BY, Sequist LV, et al. Exon 19 deletion mutations of epidermal growth factor receptor are associated with prolonged survival in non – small cell lung cancer patients treated with gefitinib or erlotinib. Clin Cancer Res 2006;12:3908–14.
11. Riely GJ, Pao W, Pham D, et al. Clinical course of patients with non – small cell lung cancer and epidermal growth factor receptor exon 19 and exon 21 mutations treated with gefitinib or erlotinib. Clin Cancer Res 2006;12:839–44.
12. Kosaka T, Yatabe Y, Endoh H, Kuwano H, Takahashi T, Mitsudomi T. Mutations of the epidermal growth factor receptor gene in lung cancer: biological and clinical implications. Cancer Res 2004;64:8919–23.
13. Tokumo M, Toyooka S, Kiura K, et al. The relationship between epidermal growth factor receptor mutations and clinicopathologic features in non – small cell lung cancers. Clin Cancer Res 2005;11:1167–73.
14. Herrmann BG, Frischauf AM. Isolation of genomic DNA. Methods Enzymol 1987;152:180–3.
15. Tokumo M, Toyooka S, Ichihara S, et al. Double mutation and gene copy number of EGFR in gefitinib refractory non – small-cell lung cancer. Lung Cancer 2006;53:117–21.
16. Toyooka S, Kiura K, Mitsudomi T. EGFR mutation and response of lung cancer to gefitinib. N Engl J Med 2005;352:2136.
17. Politi K, Zakowski MF, Fan PD, Schonfeld EA, Pao W, Varmus HE. Lung adenocarcinomas induced in mice by mutant EGF receptors found in human lung cancers respond to a tyrosine kinase inhibitor or to down-regulation of the receptors. Genes Dev 2006;20:1496–510.
18. Sugio K, Uramoto H, Ono K, et al. Mutations within the tyrosine kinase domain of EGFR gene specifically occur in lung adenocarcinoma patients with a low exposure of tobacco smoking. Br J Cancer 2006;94:896–903.
19. Zang EA, Wynder EL. Differences in lung cancer risk between men and women: examination of the evidence. J Natl Cancer Inst 1996;88:183–92.
20. Toyooka S, Tsuda T, Gazdar AF. The TP53 gene, tobacco exposure, and lung cancer. Hum Mutat 2003;21:229–39.
21. Cheng YW, Hsieh LL, Lin PP, et al. Gender difference in DNA adduct levels among nonsmoking lung cancer patients. Environ Mol Mutagen 2001;37:304–10.
22. Mitsudomi T, Kosaka T, Endoh H, et al. Mutations of the epidermal growth factor receptor gene predict prolonged survival after gefitinib treatment in patients with non – small-cell lung cancer with postoperative recurrence. J Clin Oncol 2005;23:2513–20.
23. Matsuo K, Ito H, Yatabe Y, et al. Risk factors differ for non – small-cell lung cancers with and without EGFR mutation: assessment of smoking and sex by a case-control study in Japanese. Cancer Sci 2007;98:96–101.