BRG1 may compose a large protein complex, which may play a crucial role in the expression of growth- or differentiation-associated genes. Altered expression of any component may contribute to cancer development and progression.

Our study demonstrates that class II HDAC genes might play a crucial role in cancer development/progression. Further studies on the relation of altered expression of class II HDAC genes with expression profiles and epigenetic alterations may clarify the biologic and pathogenetic function of class II HDAC genes. HDAC inhibitors are promising cancer therapeutic reagents. Some therapeutic compounds (FK228[24] and MS-27-275[25]) have demonstrated inhibitory activity, specifically against class I HDACs. These HDAC inhibitors with class I specificity might be desirable as they exhibit anticancer effects.

REFERENCES

1. Osada H, Takahashi T. Genetic alterations of multiple tumor suppressors and oncogenes in the carcinogenesis and progression of lung cancer. Oncogene 2002;21:7421–34.
2. Sekido Y, Fong KM, Minna JD. Molecular genetics of lung cancer. Annu Rev Med 2003;54:73–87.
3. Rountree MR, Bachman KE, Herman JG, Baylin SB. DNA methylation, chromatin inheritance, and cancer. Oncogene 2001;20:3156–65.
4. Jenuwein T, Allis CD. Translating the histone code. Science 2001; 293:1074–80.
5. Usadel H, Brabender J, Danenberg KD, Jeronimo C, Harden S, Engles J, Danenberg PV, Yang S, Sidransky D. Quantitative adenomatous polyposis coli promoter methylation analysis in tumor tissue, serum, and plasma DNA of patients with lung cancer. Cancer Res 2002;62: 371–5.
6. Tsou JA, Hagen JA, Carpenter CL, Laird-Offringa IA. DNA methylation analysis: a powerful new tool for lung cancer diagnosis. Oncogene 2002;21:5450–61.
7. Zochbauer-Muller S, Fong KM, Virmani AK, Geradts J, Gazdar AF, Minna JD. Aberrant promoter methylation of multiple genes in non-small cell lung cancers. Cancer Res 2001;61:249–55.
8. Baylin SB, Esteller M, Rountree MR, Bachman KE, Schuebel K, Herman JG. Aberrant patterns of DNA methylation, chromatin formation and gene expression in cancer. Hum Mol Genet 2001;10:687–92.
9. Marks P, Rifkind RA, Richon VM, Breslow R, Miller T, Kelly WK. Histone deacetylases and cancer: causes and therapies. Nat Rev Cancer 2001;1:194–202.
10. de Ruijter AJ, van Gennip AH, Caron HN, Kemp S, van Kuilenburg AB. Histone deacetylases (HDACs): characterization of the classical HDAC family. Biochem J 2003;370:737–49.
11. Verdin E, Dequiedt F, Kasler HG. Class II histone deacetylases: versatile regulators. Trends Genet 2003;19:286–93.
12. Osada H, Tatematsu Y, Masuda A, Saito T, Sugiyama M, Yanagisawa K, Takahashi T. Heterogeneous transforming growth factor (TGF)-β unresponsiveness and loss of TGF-beta receptor type II expression caused by histone deacetylation in lung cancer cell lines. Cancer Res 2001;61:8331–9.
13. Zhang CL, McKinsey TA, Chang S, Antos CL, Hill JA, Olson EN. Class II histone deacetylases act as signal-responsive repressors of cardiac hypertrophy. Cell 2002;110:479–88.
14. Czubryt MP, McAnally J, Fishman GI, Olson EN. Regulation of peroxisome proliferator-activated receptor γ coactivator 1 α (PGC-1α) and mitochondrial function by MEF2 and HDAC5. Proc Natl Acad Sci USA 2003;100:1711–6.
15. Youn HD, Liu JO. Cabin1 represses MEF2-dependent Nur77 expression and T cell apoptosis by controlling association of histone deacetylases and acetylases with MEF2. Immunity 2000;13:85–94.
16. Lemercier C, Brocard MP, Puvion-Dutilleul F, Kao HY, Albagli O, Khochbin S. Class II histone deacetylases are directly recruited by BCL6 transcriptional repressor. J Biol Chem 2002;277:22045–52.
17. Jepsen K, Rosenfeld MG. Biological roles and mechanistic actions of co-repressor complexes. J Cell Sci 2002;115:689–98.
18. Chan JK, Sun L, Yang XJ, Zhu G, Wu Z. Functional characterization of an amino-terminal region of HDAC4 that possesses MEF2 binding and transcriptional repressive activity. J Biol Chem 2003;278:23515–21.
19. Grozinger CM, Schreiber SL. Regulation of histone deacetylase 4 and 5 and transcriptional activity by 14-3-3-dependent cellular localization. Proc Natl Acad Sci USA 2000;97:7835–40.
20. Zhang CL, McKinsey TA, Olson EN. Association of class II histone deacetylases with heterochromatin protein 1: potential role for histone methylation in control of muscle differentiation. Mol Cell Biol 2002; 22:7302–12.
21. Kirschmann DA, Lininger RA, Gardner LM, Seftor EA, Odero VA, Ainsztein AM, Earnshaw WC, Wallrath LL, Hendrix MJ. Down-regulation of HP1Hsα expression is associated with the metastatic phenotype in breast cancer. Cancer Res 2000;60:3359–63.
22. Reisman DN, Sciarrotta J, Wang W, Funkhouser WK, Weissman BE. Loss of BRG1/BRM in human lung cancer cell lines and primary lung cancers: correlation with poor prognosis. Cancer Res 2003;63:560–6.
23. Nielsen AL, Sanchez C, Ichinose H, Cervino M, Lerouge T, Chambon P, Losson R. Selective interaction between the chromatin-remodeling factor BRG1 and the heterochromatin-associated protein HP1α. EMBO J 2002;21:5797–806.
24. Furumai R, Matsuyama A, Kobashi N, Lee KH, Nishiyama M, Nakajima H, Tanaka A, Komatsu Y, Nishino N, Yoshida M, Horinouchi S. FK228 (depsipeptide) as a natural prodrug that inhibits class I histone deacetylases. Cancer Res 2002;62:4916–21.
25. Hu E, Dul E, Sung CM, Chen Z, Kirkpatrick R, Zhang GF, Johanson K, Liu R, Lago A, Hofmann G, Macarron R, de los Frailes M, et al. Identification of novel isoform-selective inhibitors within class I histone deacetylases. J Pharmacol Exp Ther 2003;307:720–8.

[CANCER RESEARCH 64, 8919–8923, December 15, 2004]

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[1,2]

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

ABSTRACT

Recently it has been reported that mutations in the tyrosine kinase domain of the *epidermal growth factor receptor* (*EGFR*) gene occur in a subset of patients with lung cancer showing a dramatic response to EGFR tyrosine kinase inhibitors. To gain further insights in the role of EGFR in lung carcinogenesis, we sequenced exons 18–21 of the tyrosine kinase domain using total RNA extracted from unselected 277 patients with lung cancer who underwent surgical resection and correlated the results 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, or 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.

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 kinase-extracellular signal-regulated kinase), phosphatidylinositol-3 kinase-AKT and PAK-JNKK-JNK (p21-activated protein kinase–c-Jun NH_2 terminal kinase kinase–c-Jun NH_2 terminal kinase; ref. 1). Gefitinib is an orally administered small molecule that specifically inhibits EGFR tyrosine phosphorylation (3). Clinical trials revealed that there was a significant variability in response to gefitinib. Good clinical response has been observed most frequently in women, nonsmokers, patients with adenocarcinomas, and Japanese patients (4, 5). However, it has not been possible to predict gefitinib sensitivity by levels of EGFR overexpression as determined by immunohistochemistry (6) or immunoblotting (7). The factor(s) that determine gefitinib sensitivity has long been an enigma. It has been reported recently that activating mutations of EGFR are present in a subset of pulmonary adenocarcinomas and that tumors with *EGFR* mutations are highly sensitive to gefitinib (8, 9). Furthermore, the incidence of *EGFR* mutations is higher in Japanese than in Caucasian patients (8). In this study, we searched for *EGFR* mutations in a large cohort of unselected Japanese NSCLC to correlate them with clinical and pathologic features including *KRAS* or *TP53* mutations.

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 RNA 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 RNAeasy 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′-AGCTTGTGGAGCCTCTTACACC-3′) and R1 (5′-TAAAATTGATTCCAATGCCATCC-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 NM005228). 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).

Statistical Analysis. For comparisons of proportions, the χ^2 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

Received 8/5/04; revised 9/8/04; accepted 10/19/04.

Grant support: in part by Grant-in-Aid (16591424) from the Ministry of Education, Culture, Sports, Science and Technology of Japan.

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: T. Kosaka and Y. Yatabe contributed equally to the present study.

Requests for reprints: Tetsuya Mitsudomi, Department of Thoracic Surgery, Aichi Cancer Center Hospital, 1-1 Kanokoden, Chikusa-ku, Nagoya 464-8681, Japan. Phone: 81-52-762-6111; Fax: 81-52-764-2963; E-mail: mitsudom@aichi-ml.jp.

analyzed by the log-rank test. The two-sided significance level was set at $P < 0.05$. To identify which independent factors jointly had a significant influence on the incidence of EGFR mutations, the logistic regression modeling technique was used. We did all analyses using a StatView (version 5, SAS Institute Inc., Cary, NC) software on a Macintosh computer.

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 T to a 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, S768I, R776C, V769L), 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 0.996, $P = 0.9917$).

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 to moderately 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 both 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.[5] However, *EGFR* mutations also might have prognostic impact on patients with pulmonary adenocarcinoma, even when the patients were not exposed to gefitinib because EGFR

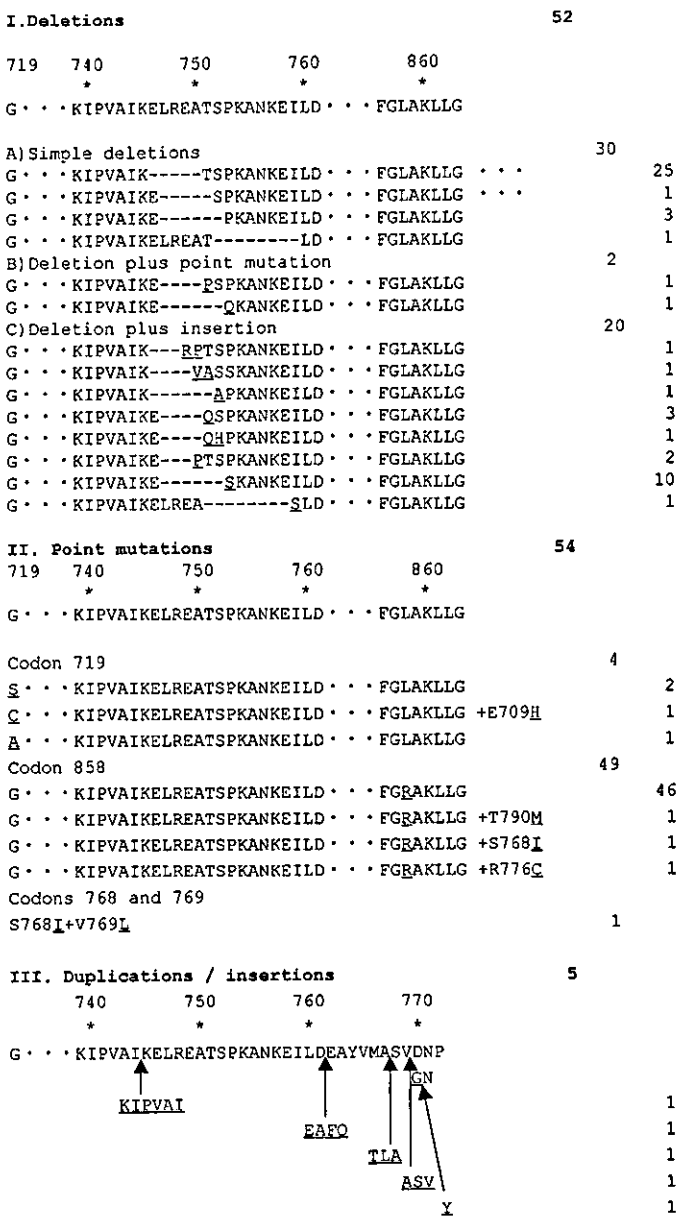

Fig. 1. Analysis of 111 *EGFR* mutations in the TK domain of the *EGFR* gene found in unselected cases with lung cancer.

[5] T. Mitsudomi, T. Kosaka, H. Endoh, Y. Horio, T. Hida, S. Mori, S. Hatooka, M. Shinoda, T. Takahashi, Y. Yatabe, submitted for publication.

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

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

There were five BACs in our cohort, of which three harbored *EGFR* mutations.

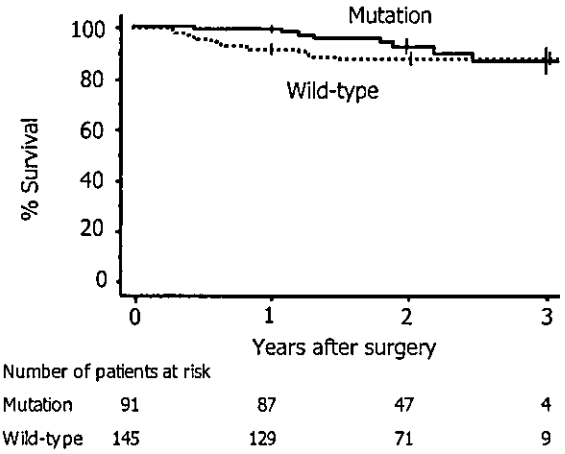

Fig. 2. Effect of *EGFR* mutations on survival of patients with adenocarcinoma calculated from the day of surgery. Patients later treated with gefitinib and those whose surgery was done for recurrent or second primary cancers were excluded.

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 1965 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

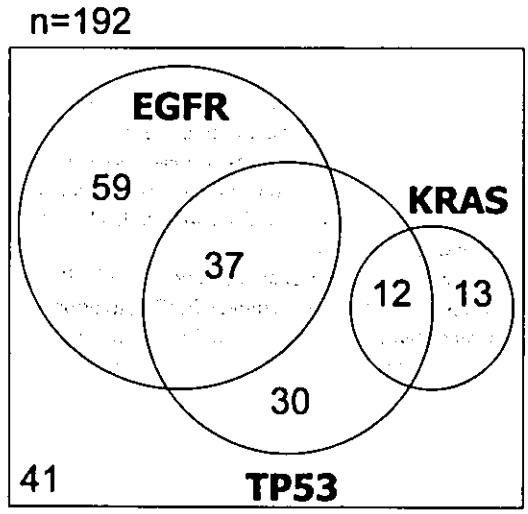

Fig. 3. The Venn diagram illustrating relationship among *EGFR* mutations, *KRAS* mutations and *TP53* mutations in patients with adenocarcinoma ($n = 192$). Diameters of each circle are roughly proportional to the number of mutations.

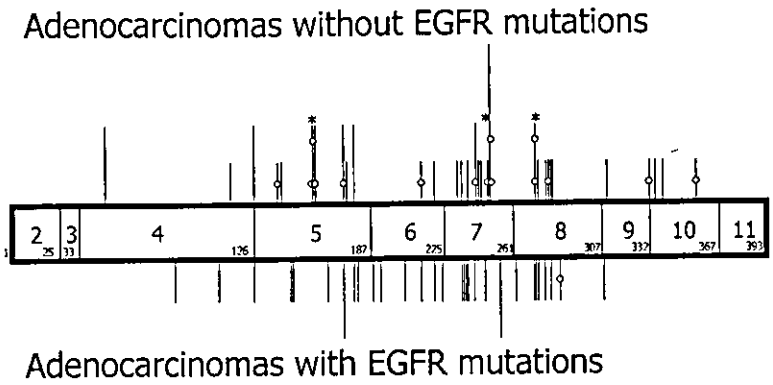

Fig. 4. Distribution of *TP53* gene mutations in adenocarcinomas without *EGFR* mutations (n = 42) or with *EGFR* mutations (n = 32). Numbers below show codons of exon boundaries. *Asterisks* show codons 157, 248, and 273, where strong and selective benzo(*a*)pyrene diol epoxide adduct formation is reported 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 are not possible at this time, but it is natural to assume that *EGFR* mutations are caused by carcinogen(s) other than those contained in tobacco smoke. In Taiwan, human papilloma virus type, 16 of 18 infections (20) or cooking oil fume (21) have been investigated as a cause of lung cancer occurring in nonsmoking women. These observations might be relevant with preferential *EGFR* mutations in nonsmoking women. Nevertheless, *EGFR* mutations should provide a clue for pathogenesis of adenocarcinoma occurring in nonsmokers and should ultimately lead to discovery of effective prevention.

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]

[6] Y. Yatabe, T. Kosaka, T. Takahashi, T. Mitsadomi, submitted for publication.

EGFR mutations were not associated with stage of disease, suggesting that EGFR mutations occurs 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

1. Arteaga CL. Overview of epidermal growth factor receptor biology and its role as a therapeutic target in human neoplasia. Semin Oncol 2002;29:3–9.
2. Bunn PA Jr, Franklin W. Epidermal growth factor receptor expression, signal pathway, and inhibitors in non-small cell lung cancer. Semin Oncol 2002;29:38–44.
3. Ciardiello F, Tortora G. A novel approach in the treatment of cancer: targeting the epidermal growth factor receptor. Clin Cancer Res 2001;7:2958–70.
4. Kris MG, Natale RB, Herbst RS, et al. Efficacy of gefitinib, an inhibitor of the epidermal growth factor receptor tyrosine kinase, in symptomatic patients with non-small cell lung cancer: a randomized trial. JAMA 2003;290:2149–58.
5. Fukuoka M, Yano S, Giaccone G, et al. Multi-institutional randomized phase II trial of gefitinib for previously treated patients with advanced non-small-cell lung cancer. J Clin Oncol 2003;21:2237–46.
6. Cappuzzo F, Gregorc V, Rossi E, et al. Gefitinib in pretreated non-small-cell lung cancer (NSCLC): analysis of efficacy and correlation with HER2 and epidermal growth factor receptor expression in locally advanced or metastatic NSCLC. J Clin Oncol 2003;21:2658–63.
7. Suzuki T, Nakagawa T, Endo H, et al. The sensitivity of lung cancer cell lines to the EGFR-selective tyrosine kinase inhibitor ZD1839 ("Iressa") is not related to the expression of EGFR or HER-2 or to K-ras gene status. Lung Cancer 2003;42:35–41.
8. Paez JG, Janne PA, Lee JC, et al. EGFR mutations in lung cancer: correlation with clinical response to gefitinib therapy. Science (Wash D C) 2004;304:1497–500.
9. 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.
10. Yatabe Y, Koga T, Mitsudomi T, Takahashi T. CK20 expression, CDX2 expression, K-ras mutation, and goblet cell morphology in a subset of lung adenocarcinomas. J Pathol 2004;203:645–52.
11. Yatabe Y, Mitsudomi T, Takahashi T. Maspin expression in normal lung and non-small-cell lung cancers: cellular property-associated expression under the control of promoter DNA methylation. Oncogene 2004;23:4041–9.
12. Travis WD, Colby TV, Sobin SH, Corrin B, Shimosato Y, Brambilla E. Histological typing of lung and pleural tumors, 3rd edition. New York: Springer-Verlag New York Inc.; 1999.
13. Denissenko MF, Pao A, Tang M, Pfeifer GP. Preferential formation of Benzo[a]pyrene adducts at lung cancer mutational hotspots in p53. Science (Wash D C) 1996;274:430–2.
14. Meuth M. The structure of mutation in mammalian cells. Biochim Biophys Acta 1990;1032:1–17.
15. Watanabe T, Hirono T, Koike T, et al. Registration of resected lung cancer in niigata prefecture. Jpn J Thorac Cardiovasc Surg 2004;52:225–30.
16. Minna JD, Gazdar AF, Sprang SR, Herz J. Cancer. A bull's eye for targeted lung cancer therapy. Science (Wash D C) 2004;304:1458–61.
17. Suzuki H, Takahashi T, Kuroishi T, Suyama M, Ariyoshi Y, Takahashi T. p53 mutations in non-small cell lung cancer in Japan: association between mutations and smoking. Cancer Res 1992;52:734–6.
18. Ahrendt SA, Decker PA, Alawi EA, et al. Cigarette smoking is strongly associated with mutation of the K-ras gene in patients with primary adenocarcinoma of the lung. Cancer (Phila) 2001;92:1525–30.
19. Mitsudomi T, Oyama T, Nishida K, et al. Loss of heterozygosity at 3p in non-small cell lung cancer and its prognostic implication. Clin Cancer Res 1996;2:1185–9.
20. Cheng YW, Chiou HL, Sheu GT, et al. The association of human papillomavirus 16/18 infection with lung cancer among nonsmoking Taiwanese women. Cancer Res 2001;61:2799–803.
21. Ko YC, Cheng LS, Lee CH, et al. Chinese food cooking and lung cancer in women nonsmokers. Am J Epidemiol 2000;151:140–7.
22. Toyooka S, Pass HI, Shivapurkar N, et al. Aberrant methylation and simian virus 40 tag sequences in malignant mesothelioma. Cancer Res 2001;61:5727–30.
23. Minami Y, Tateno H. Associations between cigarette smoking and the risk of four leading cancers in Miyagi Prefecture, Japan: a multi-site case-control study. Cancer Sci 2003;94:540–7.
24. Kobrinsky NL, Klug MG, Hokanson PJ, Sjolander DE, Burd L. Impact of smoking on cancer stage at diagnosis. J Clin Oncol 2003;21:907–13.
25. Rodenhuis S, Slebos RJ. Clinical significance of ras oncogene activation in human lung cancer. Cancer Res 1992;52:2665s–9s.
26. Otterson GA, Kratzke RA, Coxon A, Kim WK, Kaye FJ. Absence of p16INK4 protein is restricted to the subset of lung cancer lines that retains wild type rb. Oncogene 1994;9:3375–8.
27. Miller VA, Kris MG, Shah N, et al. Bronchioloalveolar pathologic subtype and smoking history predict sensitivity to gefitinib in advanced non-small-cell lung cancer. J Clin Oncol 2004;22:1103–9.
28. Yatabe Y. Role of expression of thyroid transcription factor-1 in pulmonary adenocarcinoma. In: Hyat MA, editor, Immunohistochemistry and in situ hybridization of human carcinomas. New York: Elsevier Science/Academic Press; 2004. p. 169–79.
29. Yatabe Y, Mitsudomi T, Takahashi T. TTF-1 expression in pulmonary adenocarcinomas. Am J Surg Pathol 2002;26:767–73.

Journal of Molecular Diagnostics, Vol. 6, No. 2, May 2004

Identification of MGB1 as a Marker in the Differential Diagnosis of Lung Tumors in Patients with a History of Breast Cancer by Analysis of Publicly Available SAGE Data

Takaomi Koga,* Yoshitsugu Horio,† Tetsuya Mitsudomi,‡ Takashi Takahashi,§ and Yasushi Yatabe*

From the Departments of Pathology and Molecular Diagnostics, Internal Medicine,† and Thoracic Surgery,‡ Aichi Cancer Center Hospital, Nagoya; and the Division of Molecular Oncology,§ Aichi Cancer Center Research Institute, Nagoya, Japan*

The risk of developing second primary cancers is increased in patients with breast cancer. The lung is one of the major target organs, and therefore a differential diagnosis between primary and metastatic cancers is required for the treatment of lung tumors in patients with a history of breast cancer. However, biopsy specimens frequently result in small, fragmented tissues containing only a few, degenerated cancer cells. We attempted to find a useful marker for differential diagnosis, using the online SAGE database. We selected three molecules, small breast epithelial mucin (SBEM), prostate epithelium-specific Ets transcription factor (PDEF), and mammaglobin (MGB1), as potential markers for breast cancer. SBEM and PDEF proved of no use for practical differential diagnosis because they are expressed in the normal bronchus. In contrast, expression of MGB1 was detected in all 22 primary breast cancers, but not in 22 normal lung tissues. Furthermore, all 12 metastatic breast cancers examined demonstrated positive MGB1 transcripts, whereas one of 48 primary lung adenocarcinomas expressed MGB1. This suggests that MGB1 can serve as a differential molecular marker. In practice, prospective examination, using the nine cases with a history of breast cancer, confirmed the usefulness of MGB1 in differential diagnosis. ***(J Mol Diagn 2004, 6:90–95)***

The lung is a major target of hematogeneous metastases from a variety of cancers. Thus, a diagnosis differentiating between primary and metastatic cancers is always required in clinical practice. In our institute, on average 250 lung biopsies are performed every year, and about two thirds of the tumors are diagnosed as malignant. Metastatic cancers make up 10% to 20% of these. Although this incidence may not be very high, a differential diagnosis of the metastatic cancer is important to determine the therapeutic strategy. For example, in the case of a small solitary lung tumor without any lymphadenopathy, the patient may be treated with chemotherapy or may undergo partial resection of the lung when the lung tumor is diagnosed as a metastatic breast cancer. On the other hand, standard lobectomy may be the treatment of choice when the diagnosis is of a primary non-small cell lung cancer.

The risk of developing second primary cancers is increased in patients with breast cancer, and the lung is one of the major sites involved.[1–3] Some articles have described an increased risk of primary lung cancers in association with radiation therapy following mastectomy.[4,5] Furthermore, the long latent period before identification of metastasis makes a differential diagnosis challenging. Indeed, a latent period of more than 10 years is not rare in patients with breast cancer. Histologically, a differential diagnosis between metastatic breast cancer and primary lung adenocarcinoma is difficult. Cytoplasm with secretory feature and stromal fibrosis were frequently observed in both adenocarcinomas. Moreover, metastatic breast cancer can grow along with the alveolar septa, in a similar manner to bronchioloalveolar carcinomas.[6] Difficulty is also caused by the need to carry out a differential diagnosis on biopsy specimens. Often, lung biopsies produce small amounts of fragmented tissue, which contain only a few degenerated cancer cells. Therefore, the differential diagnosis has to draw on auxiliary analysis, such as immunohistochemistry.

There are a limited number of immunohistochemical markers to identify breast cancers. Gross cystic disease fluid protein-15 (GCDFP-15) is one such marker; it is positive in only a few normal breast epithelia, but frequently expressed in breast carcinomas showing apo-

This study is supported partly by Grant-in-Aid for Encouragement of Young Scientists (B) and by Grant-in-Aid for Scientific Research (C, 14571294) from the Ministry of Education, Science, Sports and Culture, Japan.

Accepted for publication January 27, 2004.

Address reprint requests to Yasushi Yatabe, M.D., Department of Pathology and Molecular Diagnostics, Aichi Cancer Center Hospital, Kanokoden, Chikusa-ku, Nagoya 464-8681, Japan. E-mail: yyatabe@aichi-cc.jp.

crine features. Estrogen receptor (ER) is also commonly used for differential diagnosis. The combined application of GCDFP-15 and ER are very helpful for differential diagnosis, but each marker is not complete. About 50% to 60% of breast cancers express one or both, but the remaining tumors are negative for both markers. Indeed, Perry et al[7] reported that GCDFP-15 and ER are specific, but insensitive, for breast origin through the differential diagnosis of 68 metastatic adenocarcinomas to the brain. Currently, extensive human genome data have been accumulated, and the data contain much information that can be directly used in clinical practice. In the present study, we searched for molecules of potential breast-specific expression using the online database of serial analysis of gene expression (SAGE) from the National Center for Biotechnology Information (NCBI). Our results demonstrated that one of the molecules examined was specifically expressed in breast cancers, indicating that the molecule can serve as a differential marker.

Materials and Methods

Patients

Using a database of the Department of Pathology and Molecular Diagnostics of Aichi Cancer Center Hospital (Nagoya, Japan) , we first analyzed the incidence of lung biopsies that required a differential diagnosis of primary or metastatic lung cancer. For RT-PCR studies, 70 primary lung cancers, 51 metastatic lung cancers, and 22 normal lung tissue samples, as well as 22 invasive breast cancers were analyzed. All these tissues were obtained immediately after surgery, snap-frozen, and stored at −80°C until use. In addition, for prospective analysis, nine touch-imprint slides were prepared from fine-needle biopsy specimens.

Reverse Transcription and PCR (RT-PCR) Analysis

Total RNA was extracted using a standard acid guanidinium thiocyanate-phenol-chloroform method,[8] and was digested with DNase I, followed by conversion to cDNA with Superscript II reverse transcriptase (Invitrogen, Carlsbad, CA) and random hexamers. After confirmation of RNA integrity and negative contamination of genomic DNA by RT-PCR for β-actin (275 and 369 bp products for cDNA and genomic DNA, respectively), cDNA was subjected to PCR analysis. Gene-specific amplification was performed using Amplitaq Gold (Applied Biosystems, Foster City, CA). The nucleotide sequences of MGB1, SBEM and PDEF were obtained from GenBank (NM 002411, NM 058173, and NM 012391, respectively) and primers were designed to span an exon junction as follows: MGB1, 5′-ACCATGAAGTTGCTGATGGTC-3′ and 5′-AAACACCTCAACATTGCTCAGA-3′; SBEM, 5′-GTATCCAGCTACTGGTCCTGCT-3′ and 5′-CAATTGCAGAAGACTCAAGCTG-3′; PDEF, 5′-CGGTCATTGACAGCCAAG-3′ and 5′-AGGAGCCACTTCTGCACATT-3′.

Products were analyzed by electrophoresis on 2.5% high-resolution gels (NuSieve GTG agarose, BioWhittaker Molecular Applications, Rockland, ME). Some products were directly sequenced using an ABI 310 Genetic Analyzer and BigDye Primer Cycle Sequencing Kits (ABI, Foster City, CA), to confirm the amplified sequences. To determine the precise location of the mRNA expression of the gene of interest, parts of the tissues were isolated with a laser-captured microdissection system (Arcturus, Mountain View, CA). Extraction of RNA and RT-PCR were carried out as described above.

Results

Incidence of Lung Biopsies Requiring Differential Diagnosis

For the years 2000–2002, 834 lung biopsies were submitted to our department; 564 were diagnosed as malignancies or suspicious malignancies. A differential diagnosis between primary and metastatic cancers was required for 174 (21%) of the specimens submitted, because these patients had histories of cancer before the identification of lung tumors. The primary sites of the previous cancers included lung (26%), head and neck (22%), breast (13%), stomach (13%), colon (7%), and others (19%). We have reported on the differential diagnosis of metachronous[9] and synchronous lung cancers,[10] and, therefore, this study focused on the differential diagnosis of primary lung cancer and metastatic breast cancer.

Searching for Sensitive Markers of Metastatic Breast Cancer

To obtain markers that are sensitive and specific to breast cancers, we searched the online SAGE database of the NCBI (http://www.ncbi.nlm.nih.gov/SAGE/). Three candidates, mammaglobin (MGB1, Hs.46452), small breast epithelial mucin (SBEM; Hs.348419), and prostate epithelium-specific Ets transcription factor (PDEF; Hs.79414), were selected. Although only a few articles have described these molecules, they support breast-specific expression.[11–15] Detailed results from the SAGE database and the literature are summarized in Table 1. We next examined whether expression of these molecules could be used as practical distinguishing markers between primary lung cancers and metastatic breast cancers. In the 22 primary breast cancers, MGB1 was expressed in all 22, SBEM in 20, and PDEF in 20. In contrast to the results from the SAGE database, SBEM and PDEF transcripts were detected in 21 and 20 of 22 normal lung tissues, respectively, whereas MGB1 was not expressed in any (Table 1). SBEM has also been reported to be expressed in salivary glands,[14] which histologically resemble bronchial glands. We therefore determined the precise location of expressions in isolated bronchial glands, bronchial surface epithelium, and peripheral lung tissue, using a laser-capture microdis-

Table 1. Expression of MGB1, SBEM, and PDEF in Breast Cancers and in the Normal Lung

	MGB1	SBEM	PDEF
SAGE database*			
Normal tissues	9:2	376:0	4:0
Cancer tissues	17:4	32:0	126:0
Reported description			
Tissue-specific expression	Yes	Yes	Yes
Expressing organ(s)	Breast†	Breast and salivary glands	Breast and prostate
Expression in breast cancers	Yes	Yes	Yes
References	11, 16	14	12, 13
RT-PCR			
Primary breast cancer (n = 22)‡	22	20	20
Normal lung (n = 22)	0	21	20

*Ratio of total breast counts to lung counts; libraries used were GSM692, 677, 691, 760, and 780 for normal breast tissue; GSM 670, 671, 672, 673, and 694 for breast cancer tissue, and GSM762 for normal lung tissue.
†One article reported low-level of expression in gynecological malignancies, using nested RT-PCR.
‡All are invasive ductal carcinoma.

section device. SBEM expression was detected both in bronchial glands and bronchial surface epithelium, but not in parenchyma without bronchioles (Figure 1). The same experiment was carried out on PDEF, which was also expressed in bronchial cells.

Expression of MGB1 in Primary and Metastatic Cancers of the Lung

The lack of expression of MGB1 in the normal lung prompted us to examine whether MGB1 could be used in differential diagnosis between primary lung cancers and metastatic breast cancers. All of the 14 metastatic breast cancers of the lung were confirmed to express MGB1, whereas only seven of 70 primary lung cancers (10%) were positive (Table 2). No MGB1 expression could be detected in the 47 cases of adenocarcinoma of the lung

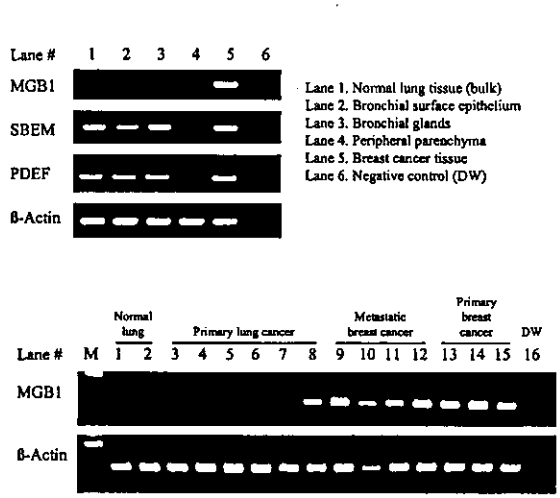

Figure 1. Expression analysis of potential breast-specific molecules (**A**) reveals that bulk tissue of normal lung (**lane 1**) expressed SBEM and PDEF, but not MGB1. Detailed examination using laser capture microdissection (**lanes 2 to 4**) demonstrated that bronchial surface epithelium and bronchial gland cells, but not the peripheral lung, were the source of the expression. In tumors (**B**), MGB1 expression was specific to breast cancers (**B, lanes 9 to 15**), with the notable exception of small-cell lung carcinomas (**B, lane 8**). DW indicates distilled water.

except for one, which interestingly metastasized to the breast 1 year after the lung operation. Detailed immunohistochemical analysis of this case revealed that both lung and breast tumors were positive for thyroid transcription factor-1 (TTF-1) and surfactant apoprotein A, and negative for estrogen and progesterone receptors. The morphological features were that of an ordinary lung adenocarcinoma. Five of the remaining six tumors with unexpected MGB1 expression were three small cell lung cancers, two large cell neuroendocrine carcinomas and a squamous cell carcinoma expressing neuroendocrine markers (CD56 and synaptophysin) in parts. Conversely, for high-grade neuroendocrine tumors of the lung, half the small cell lung cancers and large cell neuroendocrine carcinomas were positive for MGB1 expression.

We also examined the expression status of MGB1 in various other metastatic cancers. All 15 colon cancers and 12 sarcomas were negative for MGB1 expression, whereas three of four salivary gland carcinomas (adenoid cystic carcinomas) and one each of metastatic esophageal and endometrial cancers showed MGB1 expression. All of the PCR products of MGB1 in the non-breast metastatic cancer cases were confirmed by direct sequencing of the products. The other metastatic cancers, including one each of thyroid, tongue, gastric, pancreas and uterine cervix cancers, were negative for MGB1 expression. Because MGB1 transcripts were detected in the metastatic tumors, 17 primary esophageal cancers, and 10 primary salivary tumors were examined: the esophageal cancers were all negative for MGB1, and six of the 10 salivary gland tumors expressed MGB1.

Practical Application of MGB1 in Differential Diagnosis

These results indicated that MGB1 could serve as a marker of breast cancers, and thus we prospectively evaluated MGB1 expression in touch-imprint specimens of fine-needle biopsies from nine lung tumors with a breast cancer history. In six of the nine, MGB1 expression was detected, and detailed immunohistochemical results

Table 2. Expression of MGB1 in Breast and Lung Tumors

	n	Positive cases	Frequency
Cancers metastatic to lung (total)	51	16	31%
Metastatic breast cancer	12	12	100%
Metastatic colon cancer	15	0	0%
Metastatic sarcoma	12	0	0%
Metastatic salivary gland cancer	4	3	75%
Metastatic cancer, others	8	1*	13%
Primary lung cancers (total)	70	7	11%
Adenocarcinoma	48	1†	2%
Squamous cell carcinoma	12	1	8%
High-grade neuroendocrine tumor	10	5‡	50%
Low-grade neuroendocrine tumor§	3	0	0%

*A metastatic endometroid cancer.
†A primary lung adenocarcinoma that metastasized to the breast.
‡Positive cases were 2 of 4 large cell neuroendocrine carcinomas and 3 of 6 small cell carcinomas.
§All these were carcinoid tumors.

of these cases were summarized in Table 3. Among these, a representative case (Case 2) is presented below.

Case 2

A small lung nodule and right cervical lymphadenopathy were found in an 80-year-old woman, and both lesions were biopsied. She had undergone mastectomy for breast cancer 13 years before the biopsy. In the fine-needle lung biopsy, atypical carcinoma cells were identified (Figure 2). However, the tumor cells were so few that a diagnosis differentiating between metastatic breast cancer and primary lung cancer was difficult. The lymph node biopsied was replaced by an infiltration of metastatic cancer, which histologically resembled breast cancer cells. Immunohistochemical and transcript profiles were as follows: TTF-1$^-$, surfactant apoprotein A$^-$ and ER$^+$ in the breast cancer; TTF-1$^+$, surfactant apoprotein A$^+$, ER$^-$ and MGB1$^-$ in the lung tumor, and TTF-1$^-$, surfactant apoprotein A$^-$, ER$^+$, MGB1$^+$ in metastatic cancer cells of the lymph node. This suggested that the lung tumor was a primary lung adenocarcinoma, and the tumor cells in the lymph node were metastases from the breast cancer.

Discussion

We describe here a novel approach to the differential diagnosis of lung tumors in patients with a history of breast cancer. Immunohistochemical analysis is easier and more practical than nucleic acid-based assays; however, accumulated human genome data can be used for the nucleic acid-based assays. In this study, three candidate molecules, MGB1, SBEM, and PDEF, were selected from the public database as highly expressed in breast cancers but not expressed or very low in normal lung. Only a few articles have described these molecules, but they support breast-specific expression.[11–15] However, both SBEM and PDEF could be detected in normal lung. Such a discrepancy might result from our highly sensitive RT-PCR analysis, which we used to obviate the possibility of pseudo-tissue-specific expression due to a low level of expression. Indeed, expression of SBEM and PDEF were detected in the bronchus, which was a minor component in the lung parenchyma.

In contrast, MGB1 exhibited a very restricted expression pattern in the breast and salivary gland. Gruenewald et al[11] reported that MGB1 is also expressed in normal and cancerous tissues of the ovary, endometrium, and uterine cervix, as well as in normal breast and breast

Table 3. Summary of Prospective Analysis, Using Biopsy Specimens

Case	Years after breast cancer	Lung tumor	MGB1	ER	TTF-1	Surfactant	Evaluation
1	2	Solitary, lymphadenopathy	+	−	−	−	Metastatic breast cancer
2	13	Solitary, lymphadenopathy	−	−	+	+	Primary lung cancer
3	20	Solitary, 20 mm	+	+	−	−	Metastatic breast cancer
4	11	Solitary, 18 mm	+	+	−	−	Metastatic breast cancer
5	0	Multiple*	−	−	+	+	Primary lung cancer
6	12	Solitary, 37 mm	+	−	−	−	Metastatic breast cancer
7	2	Solitary, 25 mm	+	−	−	−	Metastatic breast cancer
8	2	Solitary, 20 mm	−	−	+	−	Primary lung cancer
9	12	Solitary, 15 mm	+	+	−	−	Metastatic breast cancer

*Simultaneous presentation of multiple lung nodules and a breast tumor.

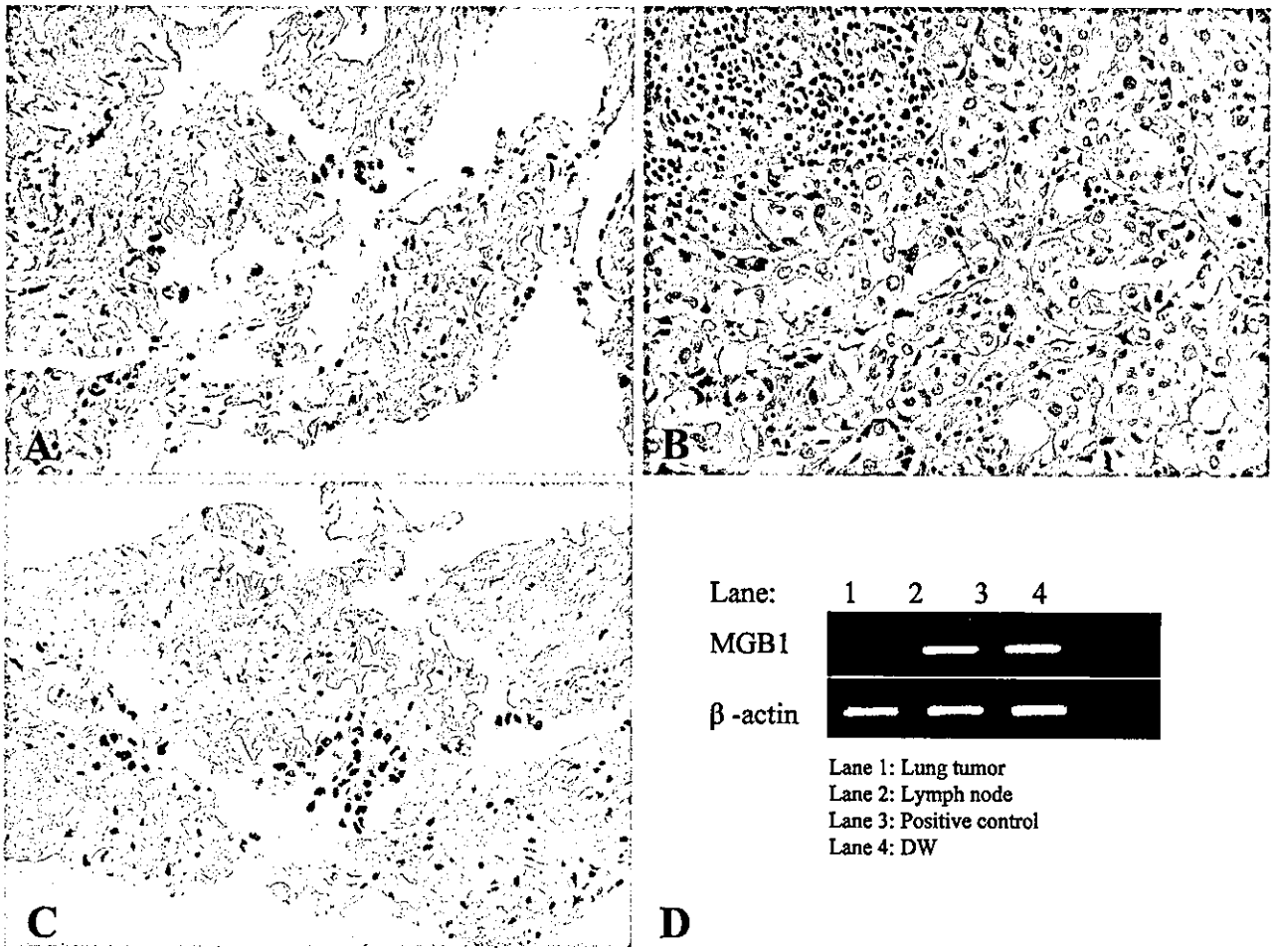

Figure 2. Practical application of MGB1 to an 80-year-old woman, who had a history of breast cancer 13 years previously. A lung tumor (**A**) and a lymph node of her left neck (**B**) were biopsied. In the lung tumor specimen, a few degenerated atypical cells are seen in the background fibrosis. Although the atypical cells are suggestive of an adenocarcinoma, it is impossible to determine whether the adenocarcinoma is primary or metastatic simply from HE sections (**A**). Positive staining of TTF-1 (**C**) and the absence of a transcript for MGB1 (**D**) suggests that this lung tumor is a primary pulmonary adenocarcinoma. By contrast, the metastatic cancer in the lymph node specimen is negative for TTF-1 and positive for MGB1 (**D**), suggesting metastatic breast cancer.

cancer tissues. Indeed, a metastatic endometrial cancer in our series expressed this molecule. MGB1 is a member of the uteroglobulin gene family, and is characterized as being a small secretory protein with glycosylation sites. It is regulated by steroid hormones, including estrogen and androgen.[16] It is of note that all members of this gene family show tissue-specific expression: Clara cell antigen (secretoglobin, family 1A, member 1) in the lung and kidney;[17] prostatein-like lipophilin A (secretoglobin, family 1D, member 1) in the prostate and tears;[18] and uteroglobin-related protein 1 (secretoglobin, family 3A, member 2) in the lung.[19] Furthermore, all but uteroglobin-related protein 1 are localized in tandem in chromosomal region 11q12.2. In addition, their physiological functions remain unclear.

The present study demonstrated that MGB1 is a sensitive and specific marker to identify metastatic breast cancers in the differential diagnosis. However, there was a notable exception of unexpected MGB1 expression in small cell carcinomas. Interestingly, small cell carcinomas unexpectedly expressed TTF-1, which regulates functional pulmonary molecules, such as surfactant apoprotein, and is expressed commonly in terminal airway unit cells and their cancers. Recent research suggests that stem cells can express a broad range of genes.[20,21] Morphologically, small cell carcinomas appear as very primitive or undifferentiated cells, and thus the unexpected co-expression of MGB1 and TTF-1 may have some association with the multi-lineage gene expression of stem cells. Indeed, ectopic expression of c-Kit, stem cell factor,[22,23] and some cancer testis antigens[24] are more common in small cell lung carcinomas than in non-small cell lung cancers.

In contrast to morphological analysis, RNA just for RT-PCR is relatively tolerant of degeneration, and RT-PCR analysis of MGB1 can be applied to a small number of tumor cells, even in a biopsy. Application of the assay to paraffin-embedded tissues may be more practical. However, a preliminary study using paraffin-embedded tissues resulted in successful detection for MGB1 amplification in only half of the positive controls studied (data not shown). RT-PCR of shorter sequences and RT-PCR followed by RNA amplification may lead to more consistent detection. Alternative material for the assay is a touch-imprint specimen of the biopsy, as described. These specimens provide high quality and sufficient amounts of RNA to perform RT-PCR. Moreover, tumor

cells were significantly enriched by the procedure.[25] A similar approach using SAGE and cDNA-microarray databases may allow the identification of molecules specific for other organs or cancers that are applicable for tumors of unknown origin.

In interpreting the results of MGB1 expression, attention should be paid to two factors. First, high-grade neuroendocrine cancers may show positive transcripts for MGB1. In this case, TTF-1, which is commonly used for the identification of pulmonary adenocarcinomas, may also be positive. When tumor cells show poorly differentiated or undifferentiated morphology, and/or atypical gene expression, such as the MGB1$^+$, TTF-1$^+$ phenotype, the possibility of high-grade neuroendocrine tumors should be excluded. Second, metastatic salivary gland cancers and gynecologic malignancies are possible tumors that may express MGB1. Like lung cancers, ovarian cancers and endometrial cancers preferentially occur as second primary neoplasms.[3] However, gynecological tumors rarely metastasize to the lung as solitary tumors without peritoneal involvement.

Although a validation study based on a larger series of various cancers is needed, the strategy described in this study provides a useful tool to develop methods for the differential diagnosis of primary and metastatic tumors. It appears to be more important to develop novel methods for the differential diagnosis of squamous cell carcinomas in those patients with head and neck cancers, because differential diagnosis is required second most frequently for these cases. Squamous cell carcinomas of the head and neck often arise in a multi-focal fashion, which is explained by the field cancerization theory, and squamous cell carcinomas of the head and neck and of the lung are morphologically indistinguishable. It is therefore of urgent need to develop markers distinguishing squamous cell carcinomas arising from the two organs.

Acknowledgments

We thank K. Hayashi for excellent technical assistance and M. Mabuchi and H. Ishida for management of paraffin blocks and preparations of slides.

References

1. Levi F, Te VC, Randimbison L, La Vecchia C: Cancer risk in women with previous breast cancer. Ann Oncol 2003, 14:71–73
2. Prochazka M, Granath F, Ekbom A, Shields PG, Hall P: Lung cancer risks in women with previous breast cancer. Eur J Cancer 2002, 38:1520–1525
3. Volk N, Pompe-Kirn V: Second primary cancers in breast cancer patients in Slovenia. Cancer Causes Control 1997, 8:764–770
4. Neugut AI, Robinson E, Lee WC, Murray T, Karwoski K, Kutcher GJ: Lung cancer after radiation therapy for breast cancer. Cancer 1993, 71:3054–3057
5. Inskip PD, Stovall M, Flannery JT: Lung cancer risk and radiation dose among women treated for breast cancer. J Natl Cancer Inst 1994, 86:983–988
6. Dail DH: Uncommon tumors. Pulmonary Pathology: Tumors. Edited by Dail DH, Hammer SP, Colby TV. New York, Springer-Verlag, 1995, pp 182–184
7. Perry A, Parisi JE, Kurtin PJ: Metastatic adenocarcinoma to the brain: an immunohistochemical approach. Hum Pathol 1997, 28:938–943
8. Chomczynski P, Sacchi N: Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. Anal Biochem 1987, 162:156–159
9. Mitsudomi T, Yatabe Y, Koshikawa T, Hatooka S, Shinoda M, Suyama M, Sugiura T, Ogawa M, Takahashi T: Mutations of the P53 tumor suppressor gene as clonal marker for multiple primary lung cancers. J Thorac Cardiovasc Surg 1997, 114:354–360
10. Shimizu S, Yatabe Y, Koshikawa T, Haruki N, Hatooka S, Shinoda M, Suyama M, Ogawa M, Hamajima N, Ueda R, Takahashi T, Mitsudomi T: High frequency of clonally related tumors in cases of multiple synchronous lung cancers as revealed by molecular diagnosis. Clin Cancer Res 2000, 6:3994–3999
11. Grunewald K, Haun M, Fiegl M, Urbanek M, Muller-Holzner E, Massoner A, Riha K, Propst A, Marth C, Gastl G: Mammaglobin expression in gynecologic malignancies and malignant effusions detected by nested reverse transcriptase-polymerase chain reaction. Lab Invest 2002, 82:1147–1153
12. Ghadersohi A, Sood AK: Prostate epithelium-derived Ets transcription factor mRNA is overexpressed in human breast tumors and is a candidate breast tumor marker and a breast tumor antigen. Clin Cancer Res 2001, 7:2731–2738
13. Oettgen P, Finger E, Sun Z, Akbarali Y, Thamrongsak U, Boltax J, Grall F, Dube A, Weiss A, Brown L, Quinn G, Kas K, Endress G, Kunsch C, Libermann TA: PDEF, a novel prostate epithelium-specific ets transcription factor, interacts with the androgen receptor and activates prostate-specific antigen gene expression. J Biol Chem 2000, 275:1216–1225
14. Miksicek RJ, Myal Y, Watson PH, Walker C, Murphy LC, Leygue E: Identification of a novel breast- and salivary gland-specific, mucin-like gene strongly expressed in normal and tumor human mammary epithelium. Cancer Res 2002, 62:2736–2740
15. Watson MA, Fleming TP: Mammaglobin, a mammary-specific member of the uteroglobin gene family, is overexpressed in human breast cancer. Cancer Res 1996, 56:860–865
16. Watson MA, Darrow C, Zimonjic DB, Popescu NC, Fleming TP: Structure and transcriptional regulation of the human mammaglobin gene, a breast cancer associated member of the uteroglobin gene family localized to chromosome 11q13. Oncogene 1998, 16:817–824
17. Singh G, Katyal SL: Clara cell proteins. Ann NY Acad Sci 2000, 923:43–58
18. Lehrer RI, Xu G, Abduragimov A, Dinh NN, Qu XD, Martin D, Glasgow BJ: Lipophilin, a novel heterodimeric protein of human tears. FEBS Lett 1998, 432:163–167
19. Niimi T, Keck-Waggoner CL, Popescu NC, Zhou Y, Levitt RC, Kimura S: UGRP1, a uteroglobin/Clara cell secretory protein-related protein, is a novel lung-enriched downstream target gene for the T/EBP/NKX2.1 homeodomain transcription factor. Mol Endocrinol 2001, 15:2021–2036
20. Akashi K, He X, Chen J, Iwasaki H, Niu C, Steenhard B, Zhang J, Haug J, Li L: Transcriptional accessibility for genes of multiple tissues and hematopoietic lineages is hierarchically controlled during early hematopoiesis. Blood 2003, 101:383–389
21. Shamblott MJ, Axelman J, Littlefield JW, Blumenthal PD, Huggins GR, Cui Y, Cheng L, Gearhart JD: Human embryonic germ cell derivatives express a broad range of developmentally distinct markers and proliferate extensively in vitro. Proc Natl Acad Sci USA 2001, 98:113–118
22. Hibi K, Takahashi T, Sekido Y, Ueda R, Hida T, Ariyoshi Y, Takagi H: Coexpression of the stem cell factor and the c-kit genes in small-cell lung cancer. Oncogene 1991, 6:2291–2296
23. Sekido Y, Obata Y, Ueda R, Hida T, Suyama M, Shimokata K, Ariyoshi Y, Takahashi T: Preferential expression of c-kit protooncogene transcripts in small cell lung cancer. Cancer Res 1991, 51:2416–2419
24. Sugita M, Geraci M, Gao B, Powell RL, Hirsch FR, Johnson G, Lapadat R, Gabrielson E, Bremnes R, Bunn PA, Franklin WA: Combined use of oligonucleotide and tissue microarrays identifies cancer/testis antigens as biomarkers in lung carcinoma. Cancer Res 2002, 62:3971–3979
25. Maitra A, Wistuba II, Virmani AK, Sakaguchi M, Park I, Stucky A, Milchgrub S, Gibbons D, Minna JD, Gazdar AF: Enrichment of epithelial cells for molecular studies. Nat Med 1999, 5:459–463

Carcinogenesis vol.25 no.8 pp.1395–1401, 2004
doi:10.1093/carcin/bgh153

Gene–environment interactions between the smoking habit and polymorphisms in the DNA repair genes, *APE1* Asp148Glu and *XRCC1* Arg399Gln, in Japanese lung cancer risk

Hidemi Ito[1,2,8], Keitaro Matsuo[1], Nobuyuki Hamajima[3], Tetsuya Mitsudomi[4], Takahiko Sugiura[5], Toshiko Saito[1], Tetsuo Yasue[6], Kyoung-Mu Lee[7], Daehee Kang[7], Keun-Young Yoo[7], Shigeki Sato[2], Ryuzo Ueda[2] and Kazuo Tajima[1]

[1]Division of Epidemiology and Prevention, Aichi Cancer Center Research Institute, Nagoya Aichi, [2]Department of Internal Medicine and Molecular Science, Nagoya City University, Graduate School of Medical Sciences, Nagoya Aichi, [3]Department of Preventive Medicine, Biostatistics and Medical Decision Making, Nagoya Graduated School of Medicine, Nagoya Aichi, [4]Department of Thoracic Surgery, Aichi Cancer Center Hospital, Nagoya Aichi, [5]Department of Thoracic Oncology, Aichi Cancer Center Hospital, Nagoya Aichi, [6]Division of Clinical Laboratories, Aichi Cancer Center Hospital, Nagoya Aichi, Japan and [7]Department of Preventive Medicine, General Surgery and Cancer Research Institute, Seoul National University College of Medicine, Seoul, South Korea

[8]To whom correspondence should be addressed
Email: hidemi@aichi-cc.jp

APE1 (apurinic/apyrimidinic endonuclease 1) and XRCC1 (X-ray cross-complementing group 1) are DNA repair proteins that play important roles in the base excision repair (BER) pathway. Polymorphisms in their encoding genes are associated with altered DNA repair capacity and thus may impact on cancer risk. In the present case-control study with 178 Japanese incident lung cancer cases and 449 age- and sex-matched controls, we investigated the gene–environment interaction among *APE1* Asp148Glu, *XRCC1* Arg399Gln and smoking habit in lung cancer risk. The results were analyzed by using conditional logistic regression models, adjusted for age, sex and smoking status. The adjusted odds ratio for the current smokers with *APE1* 148Asp/Asp, Asp/Glu and Glu/Glu genotype as compared with the never smokers with the Asp/Asp genotype were 3.01 (95% CI 1.39–6.51, $P = 0.005$), 2.73 (95% CI 1.29–5.77, $P = 0.008$) and 7.33 (95% CI 2.93–18.3, $P < 0.001$), respectively. The gene–environment interaction between current smoking and *APE1* 148Glu/Glu genotype was statistically significant (OR 3.59, 95% CI 1.28–10.1, $P = 0.015$). When *APE1* Asp148Glu and *XRCC1* Arg399Gln polymorphisms were evaluated together, the adjusted odds ratios for the current smokers with 0–1, 2 and 3–4 of *APE1* 148Glu or *XRCC1* 399Gln alleles as compared with never smokers with the 0 of these alleles were 2.96 (95% CI 1.57–5.58, $P = 0.001$), 3.86 (95% CI 1.85–8.05, $P < 0.001$) and 6.01 (95% CI 2.25–16.1, $P < 0.001$), respectively. The gene–environment interaction between current smoking and three or more *APE1* 148Glu or *XRCC1* 399Gln alleles was statistically significant (OR 2.44, 95% CI 1.00–9.22, $P = 0.049$). The OR for the gene–environment interaction of Glu/Glu genotype of *APE1* codon 148 with heavy smoking was 1.04 (95% CI 0.38–2.90, $P = 0.936$) and that with light smoking was 2.67 (95% CI 1.00–7.68, $P = 0.049$). These results suggest that *APE1* Asp148Glu and *XRCC1* Arg399Gln polymorphisms might modify the risk of lung cancer attributable to cigarette smoking exposure.

Introduction

Cigarette smoke contains large quantities of carcinogens, including polycyclic aromatic hydrocarbons, such as benzo[*a*]-pyrene, which damage DNA by covalent binding or oxidation, following activation *in vivo* into benzo[*a*]pyrene-diol epoxide (1). Although extensive prospective epidemiologic data have clearly established cigarette smoking as the major cause of lung cancer (2), only a fraction of cigarette smokers develop smoking-related lung cancer (3). This variation has been suggested to be due, in part, to genetically determined variation in carcinogen metabolism (4) and/or in the capacity of DNA repair (5–13), which is essential in protecting the genome of cells.

The apurinic/apyrimidinic (AP) endonuclease (APE1) and DNA repair enzyme X-ray repair cross-complementing group 1 (XRCC1) coordinate (14) and play a central role in the DNA base excision repair (BER) pathway, which operates on small lesions such as oxidized or reduced bases, fragmented or non-bulky adducts, or those produced by methylating agents (15). APE1, the rate-limiting enzyme in the BER pathway (16), assembles pol β onto AP sites and allows pol β and ligase III to engage in DNA repair (17). Although the *APE1* Asp148Glu polymorphism does not result in reduced endonuclease activity (18), the Glu allele may have higher sensitivity to ionizing radiation (19). Only one result has so far been reported about its relevance to lung cancer risk and this showed no significant association (20). XRCC1 interacts with ligase III, DNA polymerase β and poly (ADP-ribose) polymerase (PARP) in the C-terminal, N-terminal and central regions of XRCC1, respectively. Contradictory results have been reported about the association of the *XRCC1* Arg399Gln polymorphism with either DNA repair capacity (21,22) or the risk of lung cancer (20,23–28). The Gln allele of this polymorphism is associated with increased levels of DNA damage that may be due to reduced DNA repair, as reflected in a higher level of DNA adducts (21,29), glycophorin A variants (21,30) and bleomycin sensitivity (31) as well as chromatid exchange frequencies (22). However, other authors found no association between this polymorphism and DNA adducts (30,32). A joint effect of *APE1* Asp148Glu and *XRCC1* Arg399Gln allele genotypes has been reported regarding elevation of sensitivity to ionizing radiation (19).

Cigarette smoking may induce DNA damage (33) and individuals with a reduced capacity of DNA repair would be expected to have more carcinogen–DNA adducts in their tissue (13). Indeed, lung cancer patients may have a lower capacity of DNA repair when compared with healthy subjects and this

Abbreviations: AP, apurinic/apyrimidinic; APE1, apurinic/apyrimidinic endonuclease 1; BER, base excision repair; XRCC1, X-ray cross-complementing group 1.

may modulate the risk of lung cancer associated with smoking (11,12,34). Polymorphisms of DNA repair genes that impair their function should theoretically predispose an individual to development of tobacco-related cancers such as those in the lung (34). Therefore, we conducted the present hospital-based case-control study to test this biological hypothesis by evaluating the relationship between polymorphisms of two DNA repair genes, *APE1* and *XRCC1*, smoking and the risk of lung cancer.

Materials and methods

Study subjects

Cases and controls were first-visit outpatients at Aichi Cancer Center Hospital (ACCH) who were enrolled in the Hospital-based Epidemiologic Research Program at Aichi Cancer Center (HERPACC) (35,36). All subjects gave written informed consent to participate in the study, completed a self-administered questionnaire and provided peripheral blood. Cases of lung cancer newly diagnosed on the basis of pathologic examination at ACCH from November 2000 to April 2002 were deemed eligible, a total of 178 cases. Controls were selected by random sampling from 2158 cancer-free individuals without a past history of cancer, who visited ACCH and provided peripheral blood between November 2000 and October 2001. They were confirmed not to have cancer by the hospital-based cancer registry system by the end of 2002 and were frequency-matched to cases by sex and age group. Consequently, 178 cases and 449 controls were selected for the study.

The Institutional Ethical Review Board of Aichi Cancer Center approved this study before it was commenced (approved number 41-2).

Genotyping procedure

DNA was extracted from the buffy coat fraction using QIAamp blood mini kit (Qiagen, Valencia, CA) and genotyping for *APE1* Glu148Asp and *XRCC1* Arg399Gln polymorphisms was performed by a PCR–CTPP (PCR with confronting two-pair primers) method (37). For the *APE1* Glu148Asp (2197 T to G) polymorphism, extracted DNA was amplified with the four primers by 'Ampli Taq Gold' (Perkin-Elmer, Foster City, CA); F1, 5′-CCT ACG GCA TAG GTG AGA CC-3′; R1, 5′-TCC TGA TCA TGC TCC TCC-3′; F2, 5′-TCT GTT TCA TTT CTA TAG GCG AT-3′; and R2, 5′-GTC AAT TTC TTC ATG TGC CA-3′. PCR conditions were 1-min denaturation at 95°C followed by 30 cycles of 95°C for 1 min, 60°C for 1 min, and 72°C for 1 min, with a 5-min extension at 72°C. Primer pairs F1 and R1 for the *G* allele (148Glu), F2 and R2 for the *T* allele (148Asp) produced allele-specific bands of 167- and 236-bp, respectively, as well as a 360-bp common band. For *XRCC1* Arg399Gln (28152 G to A), extracted DNA was amplified with the four primers by 'Ampli Taq Gold' (Perkin-Elmer, Foster City, CA); F1, 5′-TCC CTG CGC CGC TGC AGT TTC T-3′; R1, 5′-TGG CGT GTG AGG CCT TAC CTC C-3′; F2, 5′-TCG GCG GCT GCC CTC CCA-3′; and R2, 5′-AGC CCT CTG TGA CCT CCC AGG C-3′. PCR conditions were 1-min denaturation at 94°C followed by 30 cycles of 94°C for 1 min, 59°C for 1 min, and 72°C for 1 min, with a 10-min extension at 72°C. Primer pairs F1 and R1 for the *G* allele (399Arg) and F2 and R2 for the *A* allele (399Gln) produced allele-specific bands of 447- and 222-bp, respectively, as well as a 630-bp common band. Both of *APE1* and *XRCC1* genotyping were confirmed by PCR–RFLP with *Baf*I (19) and *Msp*I digestion (33), respectively.

Statistical analysis

All statistical analyses were performed with STATA Version 8 software (STATA, College Station, TX). The observed genotype frequencies for controls were compared with those calculated from Hardy–Weinberg disequilibrium theory. The odds ratios (ORs) and their 95% confidence intervals (CIs) were calculated by conditional logistic regression analysis with adjustment for age, sex and smoking status. Smoking status at interview was classified into three categories: current smokers (individuals who either were currently smoking or had quit smoking within the previous 1 year); never smokers [those who smoked <100 cigarettes in their lifetime (before diagnosis for cases)]; former smokers (those who had quit smoking 1 year and more previously). For analysis of combined effect of *APE1* and *XRCC1* genotypes, three categories, 0–1, 2 and 3–4 were defined according to the number of rare allele (APE 148Glu and XRCC1 399Gln) in *APE1* and *XRCC1* genotypes. Gene–environment interactions were estimated by the logistic regression model (38), which included an interaction term as well as a variables for exposure (smoking), genotypes (*APE1* or *XRCC1*) or haplotype (number of rare alleles) and potential confounders (age and sex).

Results

As shown in Table I, the analysis included 178 cases and 449 controls. Because of frequency-matching by sex and age strata, there were no significant differences in the sex distribution and the mean age between the cases (male, 70.2%, mean ± SD, 62.9 ± 9.1, range 36–79 years) and the controls (male, 69.9% mean ± SD, 62.6 ± 9.1, range 35–79 years) ($P = 0.943$ for sex distribution, two-sided χ^2-test and $P = 0.677$ for mean age, *t*-test). Histological subtypes of the cases were: adenocarcinoma, 62.4% ($n = 111$); squamous cell carcinoma, 19.7% ($n = 35$); small cell carcinoma, 12.4% ($n = 22$) and others, 5.5% ($n = 10$). There were more current smokers in the cases (50.6%) than in the controls (29.2%) ($P < 0.001$). In addition, there were more heavy smokers (41 or more of pack-years of smoking) in the cases (69.6%) than in the smoker controls (33.7%) among smokers ($P < 0.001$).

Figure 1 shows representative results of genotyping of *APE1* Asp148Glu and *XRCC1* Arg399Gln genotypes by the

Table I. Characteristics of cases and controls

Characteristic	Cases (%) ($n = 178$)	Controls (%) ($n = 449$)	*P* value
Sex			
Male	125 (70.2)	314 (69.9)	0.943[a]
Female	53 (29.8)	135 (30.1)	
Age at diagnosis			
−39	2 (1.1)	5 (1.1)	1.000[a]
40–49	12 (6.8)	31 (6.9)	
50–59	49 (27.5)	124 (27.6)	
60–69	70 (39.3)	176 (39.2)	
70+	45 (25.3)	113 (25.2)	
Mean age ± SD	62.9 ± 9.1	62.6 ± 9.1	0.677[b]
Histology			
Adenocarcinoma	111 (62.4)	-	-
Squamous cell carcinoma	35 (19.7)	-	
Small cell carcinoma	22 (12.4)	-	
Others	10 (5.5)	-	
Smoking status			
Never smoker	53 (29.8)	182 (40.5)	<0.001[a]
Former smoker	35 (19.7)	136 (30.3)	
Current smoker	90 (50.6)	131 (29.2)	
Pack-years of smoking (Among smokers)			
1–40	35 (28.0)	177 (66.3)	<0.001[a]
41+	87 (69.6)	90 (33.7)	
Unknown	3 (2.4)	0	

[a]Two-sided χ^2 test.
[b]*t*-test.

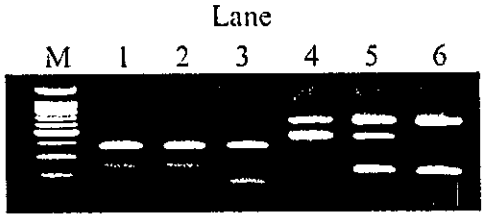

Fig. 1. Representative results for the *APE1* Asp148Glu (lanes 1–3) and *XRCC1* Arg399Gln (lanes 4–6) polymorphisms by the PCR–CTPP method. DNA fragments stained with ethidium bromide are shown. Lane M, markers; lane 1, Asp/Asp; lane 2, Asp/Glu; lane 3, Glu/Glu; lane 4, Arg/Arg; lane 5, Arg/Gln; lane 6, Gln/Gln.

PCR-CTPP methods. The genotyping results were completely in accordance with those generated by the PCR-RFLP method. The *APE1* 148Glu allele frequencies for controls and cases were 0.39 and 0.41 and the genotype distribution among controls was in accordance with the Hardy-Weinberg equilibrium law ($P = 0.250$, χ^2 test): Asp/Asp, 35.4%; Asp/Glu, 50.3%; Glu/Glu, 14.3%. The distribution among cases was: Asp/Asp, 34.8%; Asp/Glu, 47.2%; Glu/Glu, 18.0%. The *XRCC1* 399Gln allele frequencies for controls and cases were 0.25 and 0.26 and the genotype distribution among controls was again in accordance with the Hardy-Weinberg equilibrium law: Arg/Arg, 56.5%; Arg/Gln, 37.7%; Gln/Gln, 5.8% ($P = 0.750$, χ^2 test). The distribution among cases was: Arg/Arg, 55.0%; Arg/Gln, 37.1%; Gln/Gln, 7.9%. Thirty-two cases (18%) and 64 controls (14.3%) were homozygous for the codon 148 polymorphism (148 Glu/Glu) of *APE1* ($P = 0.507$, χ^2 test), and 14 cases (7.9%) and 26 controls (5.8%) were homozygous for codon 399 polymorphism (399 Gln/Gln) of *XRCC1* ($P = 0.635$, χ^2 test) (Table II).

The crude ORs for *APE1* Asp/Glu and Glu/Glu as compared with Asp/Asp genotype were not statistically significant and risks were virtually unchanged after adjustment for age, sex, smoking status and pack-years of smoking (adjusted OR = 0.99, 95% CI 0.66–1.49 for Asp/Glu and 1.29 0.75–2.21 for Glu/Glu, respectively; Table II). Likewise, the ORs for *XRCC1* Arg/Gln and Gln/Gln compared with Arg/Arg genotype were not statistically significant (adjusted OR = 1.02 for Arg/Gln, 0.69–1.50 and 1.36, 0.65–2.79 for Gln/Gln, respectively; Table II). When we examined the combined *APE1* and *XRCC1* genotypes, similar results were obtained [adjusted OR = 1.18, 0.79–1.75 for the subjects with 2 rare alleles (*APE1* 148Glu and *XRCC1* 399Gln) and 1.22, 0.63–2.37 for the subjects with 3–4 rare alleles, respectively; Table II].

The adjusted ORs for joint effect of environmental factor (smoking habit) and *APE1* codon 148 and *XRCC1* codon 399 genotype are shown in Table III. For *APE1* codon 148 genotype, the impact of the Glu/Glu genotype in current smokers appeared higher than that of Asp/Asp and Asp/Glu genotypes. ORs of current smokers with Asp/Asp, Asp/Glu and Glu/Glu genotypes were 3.01 (95% CI 1.39–6.51, $P = 0.005$), 2.73 (95% CI 1.29–5.77, $P = 0.008$) and 7.33 (95% CI 2.93–18.3, $P < 0.001$), respectively, when compared with never smokers

Table II. *APE1* codon 148 genotype and *XRCC1* codon 399 genotype frequencies and odds ratios in lung cancer patients and controls

Genotype	Case (%)	Control (%)	Crude OR (95% CI)	Adjusted OR[a] (95% CI)
APE1				
Asp/Asp	62 (34.8)	159 (35.4)	1.00 (reference)	1.00 (reference)
Asp/Glu	84 (47.2)	226 (50.3)	0.95 (0.65–1.40)	0.99 (0.66–1.49)
Glu/Glu	32 (18.0)	64 (14.3)	1.29 (0.77–2.15)	1.29 (0.75–2.21)
XRCC1				
Arg/Arg	98 (55.0)	253 (56.5)	1.00 (reference)	1.00 (reference)
Arg/Gln	66 (37.1)	169 (37.7)	1.01 (0.70–1.45)	1.02 (0.69–1.50)
Gln/Gln	14 (7.9)	26 (5.8)	1.39 (0.70–2.76)	1.35 (0.65–2.79)
APE1 and *XRCC1*				
0–1[b]	101 (56.7)	270 (60.3)	1.00 (reference)	1.00 (reference)
2[b]	59 (33.2)	144 (32.1)	1.09 (0.75–1.59)	1.18 (0.79–1.75)
3–4[b]	18 (10.1)	34 (7.6)	1.41 (0.76–2.63)	1.22 (0.63–2.37)

[a]ORs are adjusted for age, sex, smoking status and pack-years of smoking.
[b]0–1, 2 and 3–4 were defined according to the number of rare alleles (*APE1* 148Glu and *XRCC1* 399Gln) in the *APE1* and *XRCC1* genotypes.

Table III. Adjusted odds ratios and 95% CI for the joint effect of smoking habit and polymorphisms of APE1 Aps148Glu and XRCC1 Arg399Gln, and combined these genotypes

	Cases	Controls	OR[a] (95%CI)		
	Never/former/current	Never/former/current	Never	Former	Current
APE1					
Asp/Asp	19/11/32	62/47/50	1.00 (reference)	1.09 (0.44–2.71)	3.01 (1.39–6.51)*
Asp/Glu	29/19/36	91/70/65	1.11 (0.57–2.16)	1.24 (0.55–2.80)	2.73 (1.29–5.77)*
Glu/Glu	5/5/22	29/19/16	0.58 (0.20–1.74)	1.17 (0.36–3.80)	7.33 (2.93–18.3)*
XRCC1					
Arg/Arg	28/19/51	97/79/77	1.00 (reference)	1.08 (0.52–2.27)	3.39 (1.76–6.50)*
Arg/Gln	19/15/32	77/49/43	0.83 (0.43–1.61)	1.51 (0.67–3.39)	3.67 (1.80–7.46)*
Gln/Gln	6/1/7	8/7/11	2.80 (0.90–8.77)	0.66 (0.07–5.79)	3.15 (1.04–5.79)*
APE1 and *XRCC1*					
0–1[b]	31/21/49	106/81/83	1.00 (reference)	1.21 (0.59–2.46)	2.96 (1.57–5.58)*
2[b]	20/11/28	60/47/37	1.20 (0.63–2.31)	1.11 (0.47–2.63)	3.86 (1.85–8.05)*
3–4[b]	2/3/13	16/7/11	0.43 (0.09–2.00)	2.08 (0.47–9.15)	6.01 (2.25–16.1)*
All	53/35/90	182/136/131	1.00 (reference)	1.17 (0.65–2.13)	3.36 (1.97–5.73)*

[a]ORs are adjusted for age and sex.
[b]0–1, 2 and 3–4 were defined according to the number of rare alleles (*APE1* 148Glu and *XRCC1* 399Gln) in the *APE1* and *XRCC1* genotypes.
*$P < 0.05$.

Table IV. Adjusted odds ratios and 95% CI for the joint effect of smoking habit and polymorphisms of APE1 Aps148Glu and XRCC1 Arg399Gln, and combined these genotypes

	Cases (n = 178)	Control (n = 449)	OR[a] (95% CI)		
	Never/light/heavy	Never/light/heavy	Never	Light[b]	Heavy[b]
APE1					
Asp/Asp	19/14/29	62/61/36	1.00 (reference)	1.24 (0.52–2.94)	5.56 (2.35–13.2)*
Asp/Glu	29/12/41	91/96/39	1.10 (0.57–2.15)	0.68 (0.29–1.63)	7.24 (3.16–16.6)*
Glu/Glu	5/9/17	29/2015	0.56 (0.19–1.66)	2.62 (0.92–7.48)	8.38 (3.09–22.7)*
XRCC1					
Arg/Arg	28/19/49	97/114/42	1.00 (reference)	0.92 (0.45–1.90)	8.07 (3.82–17.0)*
Arg/Gln	19/16/30	77/52/40	0.86 (0.44–1.67)	1.69 (0.77–3.06)	5.41 (2.46–11.9)*
Gln/Gln	6/0/8	8/11/7	2.72 (0.86–8.57)	NA	7.55 (2.28–24.9)*
APE1 and *XRCC1*					
0–1[b]	31/21/49	106/91/83	1.00 (reference)	1.00 (0.50–2.1)	6.81 (3.26–14.2)*
2[b]	20/11/28	60/47/37	1.20 (0.62–2.31)	1.16 (0.49–2.73)	6.88 (3.08–15.4)*
3–4[b]	2/3/13	16/7/11	0.42 (0.09–1.93)	2.76 (0.60–12.8)	7.65 (2.75–21.3)*
All	53/35/87	182/177/90	1.00 (reference)	1.07 (0.59–1.92)	6.65 (3.55–12.4)*

[a]ORs are adjusted for age and sex.
[b]Light and heavy were defined according to pack-years of smoking (Light; pack-years ≤40, Heavy; pack-year >40).
[c]0–1, 2 and 3–4 were defined according to the number of rare alleles (*APE1* 148Glu and *XRCC1* 399Gln) in the *APE1* and *XRCC1* genotypes.
*$P < 0.05$.

with Asp/Asp genotypes. The OR for the gene–environment interaction between Glu/Glu genotype of *APE1* codon 148 and current smoking in lung cancer risk was 3.59 (95% CI 1.28–10.12, $P = 0.015$). For *XRCC1* codon 399 genotype, on the other hand, OR for current smokers with Gln/Gln genotype (3.15, 95% CI 1.04–5.79, $P < 0.042$) was similar to that with Arg/Arg and Arg/Gln (OR = 3.39, 95% CI 1.76–6.50, $P < 0.001$ and OR = 3.67, 95% CI 1.80–7.46, $P < 0.001$, respectively), and this gene–environment interaction between Gln/Gln genotype and current smoking in lung cancer risk was not statistically significant (OR, 0.95, 95% CI, 0.46–1.96, $P = 0.893$). Joint effects of combined *APE1* and *XRCC1* genotypes and the smoking habit are also shown in Table III. The impact of 3–4 rare alleles among current smokers appeared substantially higher than that of 2 and fewer rare alleles. Adjusted ORs were 2.96 (95% CI 1.57–5.58, $P = 0.001$) for current smokers with 0–1 rare alleles, 3.86 (95% CI 1.85–8.05, $P < 0.001$) for those with 2 rare alleles and 6.01 (95% CI 2.25–16.1, $P < 0.001$) among those with 3–4 rare alleles. The gene–environment interaction between 3 and 4 rare alleles and current smoking was statistically significant (adjusted OR 2.44, 95% CI 1.02–9.22, $P = 0.049$).

The adjusted ORs for joint effect of tobacco exposure (pack-years of smoking) and *APE1* codon 148 and *XRCC1* codon 399 genotype are shown in Table IV. The impact of the *APE1* 148Glu/Glu genotype in light smokers appeared substantially higher than that in heavy smokers. The ORs of heavy smokers with *APE1* 148Asp/Asp, Asp/Glu and Glu/Glu genotypes were 5.56 (95% CI 2.35–13.2, $P < 0.001$), 7.24 (95% CI 3.16–16.6, $P < 0.001$) and 8.38 (95% CI 3.09–22.7, $P < 0.001$), respectively, when compared with never smokers with Asp/Asp genotypes. And the OR for the gene–environment interaction between Glu/Glu genotype of *APE1* codon 148 and heavy smoking in lung cancer risk was 1.04 (95% CI 0.38–2.90, $P = 0.936$). On the other hand, the ORs for light smokers with Asp/Asp, Asp/Glu and Glu/Glu genotypes were 1.24 (95% CI 0.52–2.94, $P = 0.623$), 0.68 (95% CI 0.29–1.63, $P = 0.389$) and 2.62 (95% CI 3.09–22.7, $P = 0.049$), respectively, when compared with never smokers with Asp/Asp genotypes, and the OR for the gene–environment interaction between Glu/Glu genotype of *APE1* codon 148 and light smoking in lung cancer risk was 2.67 (95% CI 1.00–7.68, $P = 0.049$). For *XRCC1* Arg399Gln polymorphism, comparison of joint effects with light smoking could not be estimated due to small sample size in this category and the differences of the ORs among heavy smokers between the polymorphisms was not observed (8.07, 95% CI 3.82–17.0, $P < 0.001$ for Arg/Arg, 5.41, 95% CI 2.46–11.9, $P < 0.001$ for Arg/Gln and 7.55, 95% CI 2.28–24.9, $P = 0.001$ for Gln/Gln, respectively). For combined *APE1* and *XRCC1* genotypes and the smoking exposure, the impact of 3–4 rare alleles among light smokers also appeared higher than among heavy smokers. Gene–environment interactions with the 3–4 rare alleles regarding lung cancer risk were 2.20 (95% CI 0.45–10.8, $P = 0.329$) for light smoking and 1.43 (95% CI 0.37–5.48, $P = 0.602$) for heavy smoking.

Discussion

Polymorphisms altering DNA repair capacity may lead to synergistic effects with tobacco carcinogen-induced lung cancer risk. Based on this hypothesis, we examined the relationships between polymorphisms of two DNA repair genes, *APE1* Asp148Glu and *XRCC1* Arg399Gln, smoking and the risk of lung cancer. We found a statistically significant interaction of current/light smoking with *APE1* Asp148Glu polymorphism but not with *XRCC1* Arg399Gln. Moreover, we found the combination of these polymorphisms to have a statistically significant joint effect with current smoking. In contrast, we did not find significant associations with *APE1* Asp148Glu polymorphism alone as well as *XRCC1* Arg399Gln polymorphism regarding the risk of lung cancer. To the best of our knowledge, this is the first report showing a gene–environment interaction between the *APE1* and *XRCC1* genotypes and cigarette smoking with regard to lung cancer risk.

In this study, the allele frequency of *APE1* codon 148 Glu (0.39) was consistent with the previous studies (18,19). However, Misra *et al.* (20) reported Glu allele frequency of 0.52 for *APE1* codon 148 among male smokers. That of *XRCC1* 399Gln (0.26) was much lower, as well as in Koreans (0.22)

(39), in Chinese (0.25–0.27) (23,39) and in Taiwanese (0.26) (21) than those in Caucasians (0.34–0.38) (21,25,29). The differences in allele frequencies detected among these studies might be due to ethnic variation, heterogeneity of study populations and different sample sizes.

The polymorphism of *APE1* Asp148Glu has so far only been looked at regarding lung cancer risk among male smokers and a lack of any link has been reported for Caucasians (20). Regarding biological significance, the Glu allele of this polymorphism appears to be associated with hypersensitivity to ionizing radiation (19). Another study found a possible effect on endonuclease and DNA binding ability for *APE1* codon 148 Glu allele (18). Although the authors did not observe APE1 148 Glu protein defective in endonuclease and DNA binding activity, their results suggested a reduced ability to communicate with other BER proteins giving rise to reduced BER efficacy. The available evidence is thus basically accordant with our observations.

The *XRCC1* Arg399Gln polymorphism occurs within the BRCA C-terminal domain, which interacts with PARP (40). Considering the important roles of BRCA1 and PARP in DNA repair, the XRCC1 399Gln may have functional significance. This polymorphism has been reported to be linked with a higher level of DNA adducts (21,29), glycophorin A variants (21,30), bleomycin sensitivity (31) and chromatid exchange frequencies (22). Based on potential biological significance, this polymorphism has been evaluated epidemiologically in many cancers. However, no association was found with esophageal cancer (32,41), bladder cancer (32) or malignant lymphoma (42). On the other hand, possible associations have been reported for pancreatic cancer (43), prostate cancer (44), breast cancer (45) and gastric cancer (46). Regarding lung cancer, two studies in Caucasians (22,24), an American-African study (26) and a Korean study (27) demonstrated significant association with the *XRCC1* Arg399Gln polymorphism. A Caucasian study (26), Chinese studies (23,39) and our study failed to demonstrate significant association, although the trend was positive, in line with the significant associations (22,24). Although the reason for inconsistency across various types of cancer is unclear, one may say that the effect of this polymorphism on lung cancer is consistent across studies and is accordant with biological mechanisms. For different polymorphism of *XRCC1*, Arg194Trp, inconsistent results have been reported regarding lung cancer risk and gene–environment interaction with smoking, alcohol and serum antioxidants (23,39,47).

Although the result for joint effect of polymorphisms of *APE1* Asp148Glu and *XRCC1* Arg399Gln on lung cancer risk was not significant, it suggested that the individuals with 3–4 rare alleles are at increased risk of lung cancer. This result is in agreement with the evidence that APE1 and XRCC1 coordinate to the BER pathway (14) and that the joint effect of two genotypes yields higher sensitivity to ionizing radiation (19).

The risk for lung cancer among smokers is thought to increase with cumulative tobacco exposure (48), and genetic susceptibility to lung cancer may depend on the level of exposure to tobacco smoke (49,50). Therefore, we examined further gene–environment interaction between tobacco smoke exposure (pack-years of smoking) and the polymorphisms of *APE1* and *XRCC1*. When the subjects were divided into two groups according to cumulative cigarette consumption ($\leq$40 pack-years and $>$40 pack-years of smoking), we found that light smoking had a statistically significant interaction with the Glu/Glu genotype of *APE1* codon 148 regarding risk of lung cancer, while heavy smoking did not. Although comparison of joint effects of the light and heavy smoking with *XRCC1* Arg399Gln polymorphism could not be estimated due to small sample size in this study, such epidemiological comparisons have been recently conducted (26–28). In these studies, results among Caucasians, African-Americans and Koreans have consistently demonstrated that the *XRCC1* Gln allele may confer higher risk in light smokers. One study with healthy Italian subjects had significantly higher DNA adduct levels in lymphocytes with Gln/Gln genotypes only in never smokers, but not ever smokers (32). The exact mechanism of how cigarette smoking changes the DNA repair capacity by each genotype of these DNA repair polymorphisms is unknown. One possible mechanism is that, at high levels of exposure, the DNA repair capacity is saturated even in individuals having higher repair capacity (49,50).

Although we have evaluated the feasibility of using non-cancer outpatients participating in HERPACC program as controls and found them to reflect the general population in Japan (51), the present investigation, as a hospital-based case-control study, has several limitations. Our results may be biased by the relatively small number of subjects in the various subgroups and therefore need to be duplicated by others. Further studies with a larger sample and more complete measures of tobacco exposure are needed to clarify the gene–environment (smoking) interaction. Furthermore, we only evaluated the specific distribution of *APE1* Asp148Glu and *XRCC1* Arg399Gln polymorphisms but not the other polymorphisms of these two genes, including the *APE1* Gln51His, Ile64Val and Gly241Arg and the *XRCC1* Arg194Trp, because of relatively low allele frequencies for these polymorphisms and limited information regarding their functional significance (18,21,22,25,26). Cigarette smoke is a complex mixture of substances, and APE1 and XRCC1 contribute partially to BER. It is possible that polymorphisms of other genes not evaluated in this study could play a role in lung cancer risk, but evaluation of more polymorphisms would require larger sample sizes. Although the exact biological mechanisms for the gene–environment (smoking) interaction related to the APE1 and XRCC1 phenotypes as consequences of these polymorphisms could not be clarified, this study did provide important additional evidence of gene–environment interactions between *APE1* and *XRCC1* polymorphisms and smoking.

In conclusion, *APE1* Asp148Glu and *XRCC1* Arg399Gln polymorphisms appear to play an important role in modifying the direction and magnitude of the association between cigarette smoking exposure and lung cancer risk.

Acknowledgements

The authors are grateful to Ms Naomi Takeuchi, Ms Keiko Asai and Ms Hiroko Fujikura for their technical assistance. This work was supported in part by a Grant-in-Aid for Scientific Research on Priority Areas of Cancer from the Japanese Ministry of Education, Science, Culture, Sports, Science and Technology.

References

1. Hecht,S.S., (1999) Tobacco smoke carcinogens and lung cancer. *J. Natl Cancer Inst.*, 91, 1194–1210.

2. Blot,D. and Fraumeni,J.J. (1996) Cancers of the lung and pleura. In Schottenfeld,D. and Fraumeni,J.F.,Jr (eds) *Cancer Epidemiology and Prevention*. Oxford University Press, New York, pp. 635–637.
3. Mattson,M.E., Pollack,E.S. and Cullen,J.W. (1987) What are the odds that smoking will kill you? *Am. J. Public Health*, **77**, 425–431.
4. Caporaso,N., Landi,M.T. and Vineis,P. (1991) Relevance of metabolic polymorphisms to human carcinogenesis: evaluation of epidemiologic evidence. *Pharmacogenetics*, **1**, 4–19.
5. Goode,E.L., Ulrich,C.M. and Potter,J.D. (2002) Polymorphisms in DNA repair genes and associations with cancer risk. *Cancer Epidemiol. Biomarkers Prev.*, **11**, 1513–1530.
6. Sellers,T.A., Potter,J.D., Bailey-Wilson,J.E., Rich,S.S., Rothschild,H. and Elston,R.C. (1992) Lung cancer detection and prevention: evidence for an interaction between smoking and genetic predisposition. *Cancer Res.*, **52**, 2694s–2697s.
7. Raunio,H., Husgafvel-Pursiainen,K., Anttila,S., Hietanen,E., Hirvonen,A. and Pelkonen,O. (1995) Diagnosis of polymorphisms in carcinogen-activating and inactivating enzymes and cancer susceptibility—a review. *Gene*, **159**, 113–121.
8. Perera,F.P. (1996) Molecular epidemiology: insights into cancer susceptibility, risk assessment and prevention. *J. Natl Cancer Inst.*, **88**, 496–509.
9. Li,D., Wang,M., Cheng,L., Spitz,M.R., Hittelman,W.N. and Wei,Q. (1996) *In vitro* induction of benzo(a)pyrene diol epoxide-DNA adducts in peripheral lymphocytes as a susceptibility marker for human lung cancer. *Cancer Res.*, **56**, 3638–3641.
10. Li,D., Firozi,P.F., Wang,L.E., Bosken,C.H., Spitz,M.R., Hong,W.K. and Wei,Q. (2001) Sensitivity to DNA damage induced by benzo(a)pyrene diol epoxide and risk of lung cancer: a case-control analysis. *Cancer Res.*, **61**, 1445–1450.
11. Spitz,M.R., Wu,X., Wang,Y., Wang,L.E., Shete,S., Amos,C.I., Guo,Z., Lei,L., Mohrenweiser,H. and Wei,Q. (2001) Modulation of nucleotide excision repair capacity by XPD polymorphisms in lung cancer patients. *Cancer Res.*, **61**, 1354–1357.
12. Wei,Q., Cheng,L., Hong,W.K. and Spitz,M.R. (1996) Reduced DNA repair capacity in lung cancer patients. *Cancer Res.*, **56**, 4103–4107.
13. Wei,Q. and Spitz,M.R. (1997) The role of DNA repair capacity in susceptibility to lung cancer: a review. *Cancer Metastasis Rev.*, **16**, 295–307.
14. Vidal,A.E., Boiteux,S., Hickson,I.D. and Radicella,J.P. (2001) XRCC1 coordinates the initial and late stages of DNA abasic site repair through protein-protein interactions. *EMBO J.*, **20**, 6530–6539.
15. Lu,A.L., Li,X., Gu,Y., Wright,P.M. and Chang,D.Y. (2001) Repair of oxidative DNA damage: mechanisms and functions. *Cell Biochem. Biophys.*, **35**, 141–170.
16. Ramana,C.V., Boldogh,I., Izumi,T. and Mitra,S. (1998) Activation of apurinic/apyrimidinic endonuclease in human cells by reactive oxygen species and its correlation with their adaptive response to genotoxicity of free radicals. *Proc. Natl Acad. Sci. USA*, **95**, 5061–5066.
17. Bennett,R.A., Wilson,D.M.,3rd, Wong,D. and Demple,B. (1997) Interaction of human apurinic endonuclease and DNA polymerase beta in the base excision repair pathway. *Proc. Natl Acad. Sci. USA*, **94**, 7166–7169.
18. Hadi,M.Z., Coleman,M.A., Fidelis,K., Mohrenweiser,H.W. and Wilson,I.D. (2000) Functional characterization of Ape1 variants identified in the human population. *Nucleic Acids Res.*, **28**, 3871–3879.
19. Hu,J.J., Smith,T.R., Miller,M.S., Mohrenweiser,H.W., Golden,A. and Case,L.D. (2001) Amino acid substitution variants of APE1 and XRCC1 genes associated with ionizing radiation sensitivity. *Carcinogenesis*, **22**, 917–922.
20. Misra,R.R., Ratnasinghe,D., Tangrea,J.A., Virtamo,J., Andersen,M.R., Barrett,M., Taylor,P.R. and Albanes,D. (2003) Polymorphisms in the DNA repair genes XPD, XRCC1, XRCC3 and APE/ref-1 and the risk of lung cancer among male smokers in Finland. *Cancer Lett.*, **191**, 171–178.
21. Lunn,R.M., Langlois,R.G., Hsieh,L.L., Thompson,C.L. and Bell,D.A. (1999) XRCC1 polymorphisms: effects on aflatoxin B1-DNA adducts and glycophorin A variant frequency. *Cancer Res.*, **59**, 2557–2561.
22. Abdel-Rahman,S.Z. and El-Zein,R.A. (2000) The 399Gln polymorphism in the DNA repair gene XRCC1 modulates the genotoxic response induced in human lymphocytes by the tobacco-specific nitrosamine NNK. *Cancer Lett.*, **159**, 63–71.
23. Ratnasinghe,D., Yao,S.X., Tangrea,J.A., Qiao,Y.L., Andersen,M.R., Barrett,M.J., Giffen,C.A., Erozan,Y., Tockman,M.S. and Taylor,P.R. (2001) Polymorphisms of the DNA repair gene XRCC1 and lung cancer risk. *Cancer Epidemiol. Biomarkers Prev.*, **10**, 119–123.
24. Divine,K.K., Gilliland,F.D., Crowell,R.E., Stidley,C.A., Bocklage,T.J., Cook,D.L. and Belinsky,S.A. (2001) The XRCC1 399 glutamine allele is a risk factor for adenocarcinoma of the lung. *Mutat. Res.*, **461**, 273–278.
25. Butkiewicz,D., Rusin,M., Enewold,L., Shields,P.G., Chorazy,M. and Harris,C.C. (2001) Genetic polymorphisms in DNA repair genes and risk of lung cancer. *Carcinogenesis*, **22**, 593–597.
26. David-Beabes,G.L. and London,S.J. (2001) Genetic polymorphism of XRCC1 and lung cancer risk among African-Americans and Caucasians. *Lung Cancer*, **34**, 333–339.
27. Park,J.Y., Lee,S.Y., Jeon,H.S., Bae,N.C., Chae,S.C., Joo,S., Kim,C.H., Park,J.H., Kam,S., Kim,I.S. and Jung,T.H. (2002) Polymorphism of the DNA repair gene XRCC1 and risk of primary lung cancer. *Cancer Epidemiol. Biomarkers Prev.*, **11**, 23–27.
28. Zhou,W., Liu,G., Miller,D.P., Thurston,S.W., Xu,L.L., Wain,J.C., Lynch,T.J., Su,L. and Christiani,D.C. (2003) Polymorphisms in the DNA repair genes XRCC1 and ERCC2, smoking and lung cancer risk. *Cancer Epidemiol. Biomarkers Prev.*, **12**, 359–365.
29. Matullo,G., Palli,D., Peluso,M. *et al.* (2001) XRCC1, XRCC3, XPD gene polymorphisms, smoking and (32)P-DNA adducts in a sample of healthy subjects. *Carcinogenesis*, **22**, 1437–1445.
30. Palli,D., Russo,A., Masala,G., Saieva,C., Guarrera,S., Carturan,S., Munnia,A., Matullo,G. and Peluso,M. (2001) DNA adduct levels and DNA repair polymorphisms in traffic-exposed workers and a general population sample. *Int. J. Cancer*, **94**, 121–127.
31. Wang,Y., Spitz,M.R., Zhu,Y., Dong,Q., Shete,S. and Wu,X. (2003) From genotype to phenotype: correlating XRCC1 polymorphisms with mutagen sensitivity. *DNA Rep.*, **2**, 901–908.
32. Matullo,G., Guarrera,S., Carturan,S., Peluso,M., Malaveille,C., Davico,L., Piazza,A. and Vineis,P. (2001) DNA repair gene polymorphisms, bulky DNA adducts in white blood cells and bladder cancer in a case-control study. *Int. J. Cancer*, **92**, 562–567.
33. Duell,E.J., Wiencke,J.K., Cheng,T.J., Varkonyi,A., Zuo,Z.F., Ashok,T.D., Mark,E.J., Wain,J.C., Christiani,D.C. and Kelsey,K.T. (2000) Polymorphisms in the DNA repair genes XRCC1 and ERCC2 and biomarkers of DNA damage in human blood mononuclear cells. *Carcinogenesis*, **21**, 965–971.
34. Shen,H., Spitz,M.R., Qiao,Y., Guo,Z., Wang,L.E., Bosken,C.H., Amos,C.I. and Wei,Q. (2003) Smoking, DNA repair capacity and risk of nonsmall cell lung cancer. *Int. J. Cancer*, **107**, 84–88.
35. Tajima,K., Hirose,K., Inoue,M., Takezaki,T., Hamajima,N. and Kuroishi,T. (2000) A model of practical cancer prevention for out-patients visiting a hospital: the hospital-based epidemiologic research program at Aichi Cancer Center (HERPACC). *Asian Pac. J. Cancer Prev.*, **1**, 35–47.
36. Hamajima,N., Matsuo,K., Saito,T., Hirose,K., Inoue,M., Takezaki,T., Kuroishi,T. and Tajima,K. (2001) Gene-environment interactions and polymorphism studies of cancer risk in the hospital-based epidemiologic research program at Aichi Cancer Center II (HERPACC-II). *Asian Pac. J. Cancer Prev.*, **2**, 99–107.
37. Hamajima,N., Saito,T., Matsuo,K., Kozaki,K., Takahashi,T. and Tajima,K. (2000) Polymerase chain reaction with confronting two-pair primers for polymorphism genotyping. *Jpn. J. Cancer Res.*, **91**, 865–868.
38. Khoury,M.J. and Flanders,W.D. (1996) Nontraditional epidemiologic approaches in the analysis of gene-environment interaction: case-control studies with no controls! *Am. J. Epidemiol.*, **144**, 207–213.
39. Chen,S., Tang,D., Xue,K., Xu,L., Ma,G., Hsu,Y. and Cho,S.S. (2002) DNA repair gene XRCC1 and XPD polymorphisms and risk of lung cancer in a Chinese population. *Carcinogenesis*, **23**, 1321–1325.
40. Shen,M.R., Jones,I.M. and Mohrenweiser,H. (1998) Nonconservative amino acid substitution variants exist at polymorphic frequency in DNA repair genes in healthy humans. *Cancer Res.*, **58**, 604–608.
41. Xing,D., Qi,J., Miao,X., Lu,W., Tan,W. and Lin,D. (2002) Polymorphisms of DNA repair genes XRCC1 and XPD and their associations with risk of esophageal squamous cell carcinoma in a Chinese population. *Int. J. Cancer*, **100**, 600–605.
42. Matsuo,K., Hamajima,N., Suzuki,R., Andoh,M., Nakamura,S., Seto,M., Morhisima,Y. and Tajima,K. (2004) Lack of association between DNA base excision repair gene *XRCC1* Arg399Gln polymorphism and risk of malignant lymphoma. *Cancer Genet. Cytogenet.*, **149**, 77–80.
43. Duell,E.J., Holly,E.A., Bracci,P.M., Wiencke,J.K. and Kelsey,K.T. (2002) A population-based study of the Arg399Gln polymorphism in X-ray repair cross-complementing group 1 (XRCC1) and risk of pancreatic adenocarcinoma. *Cancer Res.*, **62**, 4630–4606.
44. van Gils,C.H., Bostick,R.M., Stern,M.C. and Taylor,J.A. (2002) Differences in base excision repair capacity may modulate the effect of dietary antioxidant intake on prostate cancer risk: an example of polymorphisms in the XRCC1 gene. *Cancer Epidemiol. Biomarkers Prev.*, **11**, 1279–1284.
45. Duell,E.J., Millikan,R.C., Pittman,G.S., Winkel,S., Lunn,R.M., Tse,C.K., Eaton,A., Mohrenweiser,H.W., Newman,B. and Bell,D.A. (2001)

Polymorphisms in the DNA repair gene XRCC1 and breast cancer. *Cancer Epidemiol. Biomarkers Prev.*, **10**, 217–222.
46. Shen,H., Xu,Y., Qian,Y., Yu,R., Qin,Y., Zhou,L., Wang,X., Spitz,M.R. and Wei,Q. (2000) Polymorphisms of the DNA repair gene XRCC1 and risk of gastric cancer in a Chinese population. *Int. J. Cancer*, **88**, 601–606.
47. Ratnasinghe,D.L., Yao,S.X., Forman,M., Qiao,Y.L. Andersen,M.R., Giffen,C.A., Erozan,Y., Tockman,M.S. and Taylor,P.R. (2003) Gene-environment interactions between the codon 194 polymorphism of XRCC1 and antioxidants influence lung cancer risk. *Anticancer Res.*, **23**, 627–632.
48. Ruano-Ravina,A., Figueiras,A., Montes-Martinez,A. and Barros-Dios,J.M. (2003) Dose-response relationship between tobacco and lung cancer: new findings. *Eur. J. Cancer Prev.*, **12**, 257–263.
49. Nakachi,K., Imai,K., Hayashi,S., Watanabe,J. and Kawajiri,K. (1991) Genetic susceptibility to squamous cell carcinoma of the lung in relation to cigarette smoking dose. *Cancer Res.*, **51**, 5177–5180.
50. Vineis,P. (1997) Molecular epidemiology: low-dose carcinogens and genetic susceptibility. *Int. J. Cancer*, **71**, 1–3.
51. Inoue,M., Tajima,K., Hirose,K., Hamajima,N., Takezaki,T., Kuroishi,T. and Tominaga,S. (1997) Epidemiological features of first-visit outpatients in Japan: comparison with general population and variation by sex, age and season. *J. Clin. Epidemiol.*, **50**, 69–77.

Received November 1, 2003; revised March 10, 2004; accepted March 18, 2004

Progression of Focal Pure Ground-Glass Opacity Detected by Low-Dose Helical Computed Tomography Screening for Lung Cancer

Ryutaro Kakinuma, MD, Hironobu Ohmatsu, MD, Masahiro Kaneko, MD, Masahiko Kusumoto, MD, Junji Yoshida, MD, Kanji Nagai, MD, Yutaka Nishiwaki, MD, Toshiaki Kobayashi, MD, Ryosuke Tsuchiya, MD, Hiroyuki Nishiyama, MD, Eisuke Matsui, MD, Kenji Eguchi, MD, and Noriyuki Moriyama, MD

Objective: To clarify the progression of focal pure ground-glass opacity (pGGO) detected by low-dose helical computed tomography (CT) screening for lung cancer.

Methods: A total of 15,938 low-dose helical CT examinations were performed in 2052 participants in the screening project, and 1566 of them were judged to have yielded abnormal findings requiring further examination. Patients with peripheral nodules exhibiting pGGO at the time of the first thin-section CT examination and confirmed histologically by thin-section CT after follow-up of more than 6 months were enrolled in the current study. Progression was classified based on the follow-up thin-section CT findings.

Results: The progression of the 8 cases was classified into 3 types: increasing size (n = 5: bronchioloalveolar carcinoma [BAC]), decreasing size and the appearance of a solid component (n = 2: BAC, n = 1; adenocarcinoma with mixed subtype [Ad], n = 1), and stable size and increasing density (n = 1: BAC). In addition, the decreasing size group was further divided into 2 subtypes: a rapid-decreasing type (Ad: n = 1) and a slow-decreasing type (BAC: n = 1). The mean period between the first thin-section CT and surgery was 18 months (range: 7–38 months). All but one of the follow-up cases of lung cancer were noninvasive whereas the remaining GGO with a solid component was minimally invasive.

Conclusions: The pGGOs of lung cancer nodules do not only increase in size or density, but may also decrease rapidly or slowly with the appearance of solid components. Close follow-up until the appearance of a solid component may be a valid option for the management of pGGO.

Key Words: ground-glass opacity, low-dose helical computed tomography screening, lung cancer

(*J Comput Assist Tomogr* 2004;28:17–23)

Focal pure ground-glass opacities (pGGOs), or nodules of the lungs, has become a major concern as low-dose helical computed tomography (CT) screening for lung cancer becomes more widely available, not only in the field of diagnostic imaging,[1–5] but also in the field of limited surgery.[6–10] GGO is a finding on thin-section CT images of the lung which has been described as a hazy, increased attenuation of the lung tissue with preservation of the bronchial and vascular margins. GGO is usually a nonspecific finding that is found in many types of pulmonary disease.[11] However, some investigators have recently reported that most localized pGGOs or focal GGOs are malignant.[1,2,5] Although a few reports have described the evolution of lung cancer using conventional chest CT,[12–14] thin-section CT[15–17] and low-dose screening CT,[18,19] the natural history of peripheral lung cancers that exhibit as pGGO on thin-section CT images detected using low-dose helical CT screening is still unclear.

The purpose of this retrospective study was to clarify the progression of pGGOs, which were not visible on chest radiographs, detected by low-dose helical CT screening examinations performed every 6 months. We evaluated the progression of pGGOs based on the thin-section CT findings obtained during the follow-up after the first thin-section CT.

PATIENTS AND METHODS

Subjects

Between September 1993 and January 2003, low-dose helical CT screening was conducted semiannually in Tokyo by

From the Division of Thoracic Oncology (Drs Kakinuma, Ohmatsu, Yoshida, Nagai, and Nishiwaki), National Cancer Center Hospital East, Tsukiji, Chiba Japan; the Divisions of Endoscopy (Drs Kaneko and Kobayashi), Diagnostic Radiology (Drs Kusumoto and Moriyama), and Thoracic Surgery (Dr Tsuchiya), National Cancer Center Hospital; the Division of Thoracic Surgery (Dr Nishiyama), Social Health Insurance Medical Center, Okubo, Japan; (Dr Matsui); the Anti–Lung Cancer Association, Ichigaya, Japan; and the Division of Internal Medicine (Dr Eguchi), School of Medicine, Tokai University, Isehara, Japan.

This study was supported in part by a Grant-in-Aid for Cancer Research (13-8) from the Ministry of Health, Labor, and Welfare of Japan and by a Grant-in-Aid from the Second-Term Comprehensive 10-Year Strategy for Cancer Control.

Reprints: Ryutaro Kakinuma, MD, National Cancer Center Hospital East, 6-5-1 Kashiwa-no-ha, Kashiwa, Chiba 277-8577, Japan (e-mail: rkaki@east.ncc.go.jp).