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.

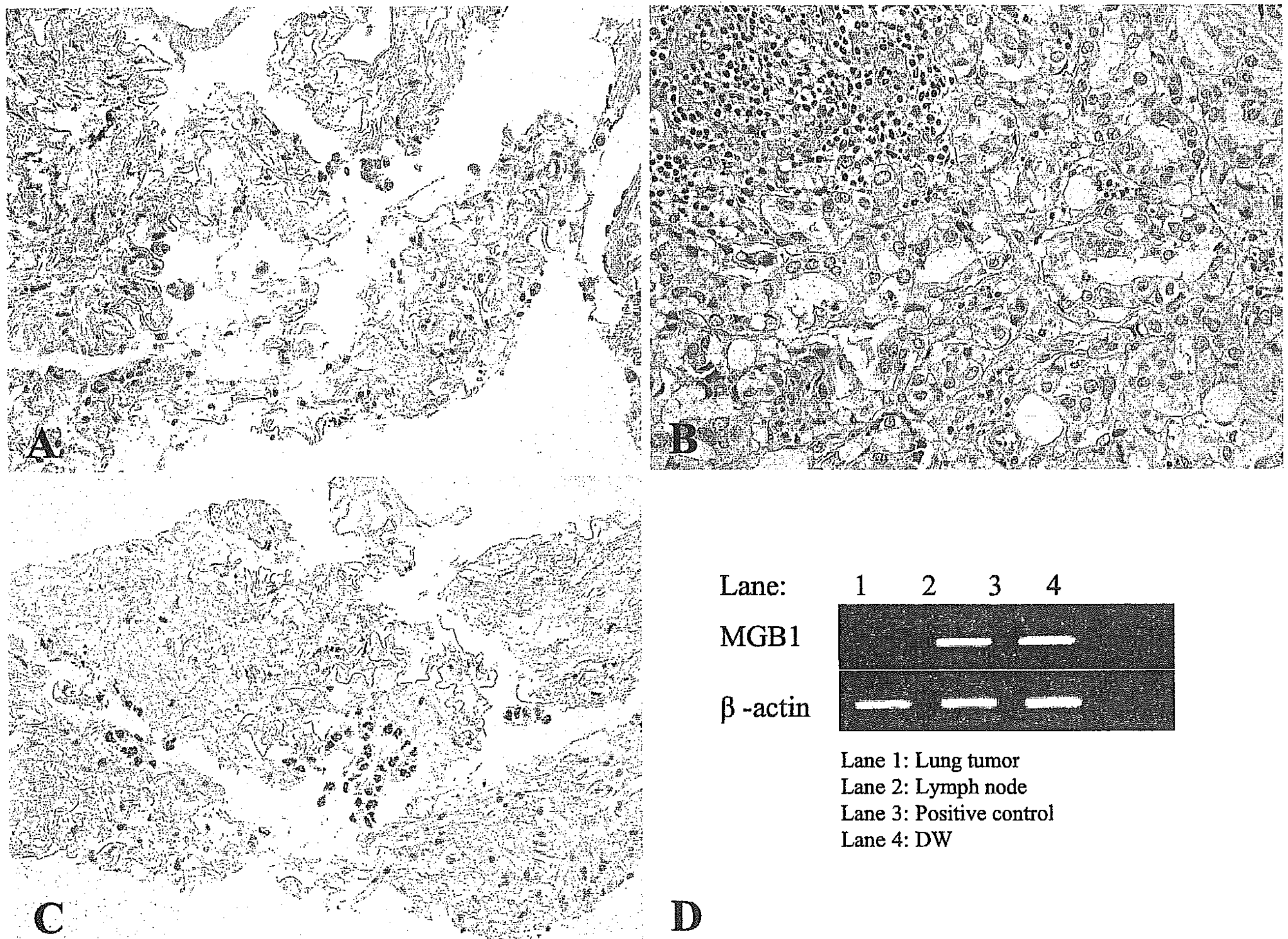


Figure 2. Practical application of MGB1 to an 80-year-old woman, who had a history of breast cancer 13 years previously. A lung tumor (**A**) and a lymph node of her left neck (**B**) were biopsied. In the lung tumor specimen, a few degenerated atypical cells are seen in the background fibrosis. Although 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 $MGB1^+$, $TTF\text{-}1^+$ 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.

 © Oxford University Press 2004; all rights reserved.

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.

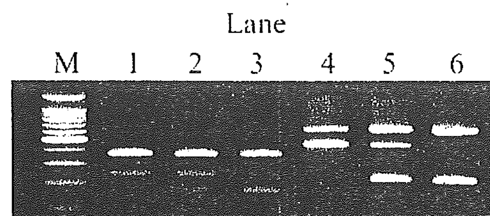


Fig. 1. Representative results for the *APE1* Asp148Glu (lanes 1–3) and *XRCC1* Arg399Gln (lanes 4–6) polymorphisms by the PCR-CTPP method. DNA fragments stained with ethidium bromide are shown. Lane M, markers; lane 1, Asp/Asp; lane 2, Asp/Glu; lane 3, Glu/Glu; lane 4, Arg/Arg; lane 5, Arg/Gln; lane 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 148*Glu/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 noncancer 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

VOLUME 22 · NUMBER 5 · MARCH 1 2004

JOURNAL OF CLINICAL ONCOLOGY

ORIGINAL REPORT

Prognostic Model of Pulmonary Adenocarcinoma by Expression Profiling of Eight Genes As Determined by Quantitative Real-Time Reverse Transcriptase Polymerase Chain Reaction

Hideki Endoh, Shuta Tomida, Yasushi Yatabe, Hiroyuki Konishi, Hirotaka Osada, Kohei Tajima, Hiroyuki Kuwano, Takashi Takahashi, and Tetsuya Mitsudomi

From the Department of Thoracic Surgery, Department of Pathology and Molecular Diagnostics, Aichi Cancer Center Hospital; Division of Molecular Oncology, Aichi Cancer Center Research Institute, Aichi Cancer Center, Nagoya, Japan; and Department of Surgery I, Gunma University, School of Medicine, Maebashi, Japan.

Submitted April 14, 2003; accepted October 3, 2003.

Supported in part by the Aichi Cancer Research Foundation.

Authors' disclosures of potential conflicts of interest are found at the end of this article.

Address reprint requests to Tetsuya Mitsudomi, MD, PhD, Department of Thoracic Surgery, Aichi Cancer Center, Hospital, 1-1 Kanokoden, Chikusa-ku, Nagoya 464-8681, Japan; e-mail: mitsudom@aichi-cc.jp.

© 2004 by American Society of Clinical Oncology

0732-183X/04/2205-811/$20.00

DOI: 10.1200/JCO.2004.04.109

ABSTRACT

Purpose

Recently, several expression-profiling experiments have shown that adenocarcinoma can be classified into subgroups that also reflect patient survival. In this study, we examined the expression patterns of 44 genes selected by these studies to test whether their expression patterns were relevant to prognosis in our cohort as well, and to create a prognostic model applicable to clinical practice.

Patients and Methods

Expression levels were determined in 85 adenocarcinoma patients by quantitative reverse transcriptase polymerase chain reaction. Cluster analysis was performed, and a prognostic model was created by the proportional hazards model using a stepwise method.

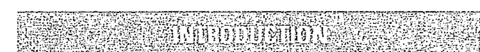
Results

Hierarchical clustering divided the cases into three major groups, and group B, comprising 21 cases, had significantly poor survival ($P = .0297$). Next, we tried to identify a smaller number of genes of particular predictive value, and eight genes (*PTK7, CIT, SCNN1A, PGES, ERO1L, ZWINT*, and two *ESTs*) were selected. We then calculated a risk index that was defined as a linear combination of gene expression values weighted by their estimated regression coefficients. The risk index was a significant independent prognostic factor ($P = .0021$) by multivariate analysis. Furthermore, the robustness of this model was confirmed using an independent set of 21 patients ($P = .0085$).

Conclusion

By analyzing a reasonably small number of genes, patients with adenocarcinoma could be stratified according to their prognosis. The prognostic model could be applicable to future decisions concerning treatment.

J Clin Oncol 22:811-819. © 2004 by American Society of Clinical Oncology



INTRODUCTION

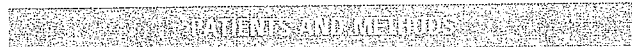
In Japan, as in many Western countries, lung cancer is the leading cause of cancer-related death, claiming more than 50,000 lives annually, and the situation is worsening [1]. Approximately 30% of patients with non–small-cell lung carcinoma (NSCLC) have localized disease, and successful surgical management with long-term disease control is generally restricted to this group of early-stage patients. NSCLC is histopathologically and clinically distinct from small-cell lung carcinoma (SCLC), and is further subdivided into adenocarcinoma, squamous cell carcinoma, and large-cell carcinoma [2]. Although these types share common characteristics, they are thought to develop from at least partially different sets of genetic alterations [3].

Adenocarcinoma is currently the most predominant histological subtype of NSCLC in Japan as well as in the United States [4]. Although morphological features and clinical stage based on the tumor-node-metastasis system can roughly stratify patients for prognosis, it is often difficult to predict either which surgically managed patients are at risk for early relapse or which rare advanced-stage patients

may experience prolonged survival [2]. To guide clinical decisions on the optimum treatment regimen, there is clearly a need to accurately identify patients at high risk for recurrent or metastatic disease. Therefore, search of the genetic lesions identified by recent advances in cancer molecular biology for those relevant to predicting patient prognosis is considered to be of great importance. Many molecular markers that predict patient survival independent of the tumor-node-metastasis staging system have been reported [5]. These include oncogenes (*K-ras*, *Bcl2*, Her2/*neu*, *EGFR*), tumor suppressor genes (*p53*, *RB*, *p16*, *p27*), cell cycle modulators (cyclins), molecules related to tumor invasion and metastasis (CD44, cathepsin B, matrix metalloproteinase), telomerase, molecules involved in tumor angiogenesis (vascular endothelial growth factor, vascular endothelial growth factor receptor) and cyclo-oxygenase 2 [5]. However, for the moment there is no single biomarker available that can be routinely used for prediction of prognosis of NSCLC. This may be quite reasonable considering that cancer is a complex multigene disease.

Recently, cDNA microarray technologies that simultaneously analyze the expression tens of thousands of genes have been used to correlate gene-expression patterns in individuals with various clinical parameters, including morphologic features and tumor behavior. In 2001, groups from Stanford [6] and Boston [7] applied expression profiling technologies to lung cancer, and they both concluded that (1) clusters defined by gene expression patterns recapitulate morphological classification of the tumors into squamous, small-cell, large-cell, and adenocarcinoma, and that (2) adenocarcinoma can be classified into subgroups that reflect patient survival. However, these two groups selected quite different sets of genes that influenced patient survival [6,7]. Similar subsequent studies also identified sets of genes of prognostic interest; but again, the selected sets of genes were quite different among reports [8-10].

In this study, we used real-time reverse transcriptase polymerase chain reaction (RT-PCR) to examine expression of 44 genes reported previously [6,7] as putative prognostic markers in our cohort of pulmonary adenocarcinoma patients. Our objectives were (1) to confirm whether expression patterns of these 44 genes were relevant to prognosis in our cohort of patients, and (2) to create a prognostic model that could reasonably be used in routine clinical practice by further selecting a smaller number of genes.



Cell Lines and Patients

RNAs derived from five SCLC and 15 NSCLC cell lines were used for optimization of PCR conditions. Details of their derivation and culture conditions have been described [11,12]. Adenocarcinoma samples were obtained from 85 consecutive patients who underwent pulmonary resection at the Aichi Cancer Center Hospital (Nagoya, Japan) from December 1995 through May 1998, after obtaining approval from the institutional review board, and patients' written informed consent. The patients were 44 males and 41 females, with age at diagnosis ranging from 32 to 84 years (median age, 62 years). Forty-eight patients had stage I disease, six had stage II, 30 had stage III, and one patient had stage IV disease. Twenty-four patients had poorly differentiated; 47, moderately differentiated; and 14, well-differentiated adenocarcinoma. Thirty-eight patients were smokers, with a median Brinkman index (number of cigarettes per day × years) of 855; the remaining 47 were never smokers.

For validation of our prognostic model, we used an independent set of 21 patients with pulmonary adenocarcinoma who had undergone pulmonary resection from March 1994 through November 1995. These patients were 12 males and nine females, with an age at diagnosis ranging from 43 to 78 years (median age, 60 years), and 13 patients were smokers with a median Brinkman index of 750. Eight patients had stage I disease, one had stage II, 10 had stage III, and two had stage IV disease. Nineteen patients had moderately differentiated, and two had well-differentiated adenocarcinoma.

All patients underwent potentially curative resection (84 of 85 and 21 of 21 underwent formal pulmonary lobectomy or more, with systematic ipsilateral mediastinal lymph node dissection). One remaining patient underwent partial resection due to poor pulmonary reserve. Stages were determined after pathologic evaluation of resected specimens according to the International System for Staging Lung Cancer, revised in 1997 [13].

Tumor samples were processed immediately after surgical removal. Tissue specimens were grossly examined by a surgical pathologist (Y.Y.), and a piece of the tumor tissue was carefully obtained so as to maximize tumor content. A half of the piece was snap frozen in liquid nitrogen, followed by storage at −80°C until use. The other half was fixed with prechilled acetone, and embedded in paraffin for evaluation of tumor contents and scored in one of four classes (eg, < 25%, 25% to 50%, 50% to 75%, and > 75%). Total RNA was isolated using the acid guanidinium isothiocyanate/cesium chloride procedure [14]. All samples used in this study were analyzed by RT-PCR of glyceraldehyde-3-phosphate dehydrogenase (*GAPDH*) gene (539 base pair [bp]) to ensure the integrity of RNA using the following primers: 5′-GTCAACGGATTTGGTCG-TATT-3′ and 5′-AGTCTTCTGGGTGGCAGTGAT-3′.

Genes Examined in the Present Study

Initially, we tried to examine 48 genes that had been previously reported to be relevant to subgrouping of adenocarcinoma for 28 genes [6] and 20 genes [7] (Table 1). PCR primers were designed to amplify 3′ untranslated regions of the genes to yield small fragments (range, 91 to 333 bp; mean, 170 bp). Primer sequences are available on request. Most of the primers were purchased as Human Genepair (Research Genetics, Huntsville, AL), and others were synthesized according to the database at UniSTS (National Center for Biotechnology Information, Bethesda, MD). Primers for five genes were designed using Primer Express Software version 1.5 (Applied Biosystems, Foster City, CA).

Relative Quantification by Real-Time RT-PCR

First-strand cDNAs were synthesized from total RNA using Superscript II (Invitrogen, Carlsbad, CA) and random hexamer primers (Roche Applied Science, Alameda, CA). Real-time quantitative PCR amplifications were performed by SYBR Green assay in an ABI PRISM 7900-HT (Applied Biosystems). The reactions were carried out in a 96-well plate in 25-μL reactions containing 2 × SYBR Green Master Mix (Applied Biosystems), 200 to 250 nmol/L each, forward and reverse primer, and a cDNA template

Table 1. 48 Genes Examined in This Study

No.	Gene Name	Symbol	Accession Number	uniSTS Code	Location	Size (base pairs)
1	Intercellular adhesion molecule 1 (CD54), human rhinovirus receptor	*ICAM1*	J03132	WI-7031	19p13.3-13.2	224
2	Protein tyrosine kinase 7	*PTK7*	U40271	WI-8716	6p21.1-12.2	108
3	Carcinoembryonic antigen-related cell adhesion molecule 1 (biliary glycoprotein)	*CEACAM1*	S71326	WI-17429	19q13.2	104
4	Dipeptidylpeptidase IV (CD26, adenosine deaminase complexing protein 2)	*DPP4*	X60708	WI-7620	2q24.3	332
5	Collagen, type IX, alpha 2	*COL9A2*	AF019406	SGC30746	1p33-32	127
6	Thyroid transcription factor 1	*TITF1*	X82850	SGC35528	14q13	204
7	Epididymis-specific, whey-acidic protein type, four-disulfide core domain 2, HE4	*WFDC2*	X63187	SGC30446	20q12-13.2	150
8	Citron (rho-interacting, serine/threonine kinase 21)	*CIT*	AB023166	stSG3138	12q24	142
9	Hepsin (transmembrane protease, serine 1)	*HPN*	M18930	WI-7579	19q11-13.2	276
10	Ornithine decarboxylase 1	*ODC1*	XM002679	stSG1950	2p25	165
11	Tumor suppressor deleted in oral cancer-related 1	*DOC-1R*	G24795	SHGC34148	11q13	100
12	Cartilage paired-class homeoprotein 1	*CART1*	U31986	sts-N20106	12q21.3-22	137
13	Sodium channel, nonvoltage-gated 1 alpha	*SCNN1A*	XM033306	sts-X76180	12p13	170
14	Solute carrier family 2 (facilitated glucose transporter), member 1	*SLC2A1*	NM006516	WI-15743	1p35-31.3	150
15	Ataxia-telangiectasia group D-associated protein (tripartite motif-containing 29)	*TRIM29*	NM012101	WI-7302	11q22-23	333
16	*KIAA0101*	*KIAA0101*	NM014736	D14657	15q11.2	184
17	Prostaglandin E synthase	*PGES*	AF027740		9q34.3	151
18	Cathepsin L	*CTSL*	X12451	WI-7541	9q21-22	229
19	ESTs Hs.11607 *AA443569* (*FLJ32205*)	*EST AA443569*	AK056767	stSG42238	7p22.3	154
20	Dickkopf (Xenopus laevis) homolog 1	*DKK1*	AF127563		10q11.2	121
21	ESTs, highly similar to LB4D human NADP-dependent Leukotriene B4 12-hydroxydehydrogenase	*EST-LB4D*	BI254118		9p32	151
22	Vascular endothelial growth factor C	*VEGFC*	NM005429	stSG2713	4q34.1-34.3	130
23	ERO1L (S. cerevisiae)-like	*ERO1L*	AF081886	G32587	14q22.1	108
24	v-erb-b2 avian erythroblastic leukemia viral oncogene homolog 2	*ERBB2*	M11730	GDB:181407	17q21.1	148
25	Similar to phosphatidylcholine transfer protein 2 serologically defined colon cancer antigen 28	*SDCCAG28*	AF151810	stSG4384	11q13	125
26	EST Hs.98803 *AA434256*	*EST AA434256*	AA434256	stSG47892	9	120
27	Islet cell autoantigen 1 (69kD) autoantigen p69	*ICA1*	U38260	WI-7176	7p22	128
28	ESTs Hs.102406 *AA468094*	*EST AA468094*	AA468094		2	121
29	Kallikrein 11	*KLK11*	AF164623	SHGC57422	19q13.3-13.4	199
30	Achaete-scute complex homolog-like 1	*ASCL1*	L08424	WI-9226	12q22-23	111
31	Carboxypeptidase E	*CPE*	X51405	WI-7540	4q32.3	325
32	Calcitonin/carcitonin-related polypeptide, alpha	*CALCA*	X00356	WI-6982	11p15.2-15.1	274
33	Tubulin, beta polypeptide	*TUBB*	X79535	WI-7931	6p25	139
34	Tumor rejection antigen(gp96) 1	*TRA1*	X15187	R99860	12q24.2-24.3	141
35	X-box binding protein 1	*XBP1*	NM005080	WI-8513	22q12.1	344
36	Dual specificity phosphatase 4 (MAP kinase phosphatase 2)	*DUSP4*	U21108	U21108	8p12-11	207
37	Platelet-activating factor acetylhydrolase	*PAFAH1B3*	NM002573	NIB1825	19q13.1	267
38	Trefoil factor 3 (intestinal), TFF3	*(H)ITF*	L08044	WI-7267	21q22.3	125
39	Proprotein convertase subtilisin/kexin type 1	*PCSK1*	X64810	RH92225	5q15-21	143
40	Nuclear transcription factor Y, beta	*NFYB*	NM006166	RH71289	12q22-23	130
41	ZW10 interactor	*ZWINT*	AF067656	sts-H99221	10q21-22	126
42	Oncogene Ret/Ptc2	*Ret/Ptc2*	L03357	RH66458	10q11.2	110
43	Trinucleotide repeat containing 9	*TNRC9*	U80736		16q12.2	91
44	*KIAA1025*	*KIAA1025*	AB028948	stSG1522	12q24.22	273
45	*KIAA1128*	*KIAA1128*	AF241785	WI-8523	10q23.2	110
46	Dopa decarboxylase	*DDC*	M76180	M76180	7p11	216
47	Hepatocyte nuclear factor 3, alpha	*HNF3A*	U39840	stSG43208	14q12-13	125
48	*KIAA0767*	*KIAA0767*	BC025418	A005V42	22q13.31	121

NOTE. No. 1 through 28 are derived from Garber et al [6]; 29 through 48, from Bhattacharjee et al [7]; and 28, 46, 47, and 48 were excluded in the analysis (see Results).

corresponding to 20 ng total RNA. SYBR Green PCR conditions were 50°C for 2 minutes, 95°C for 10 minutes, followed by 95°C for 50 seconds, 57°C or 60°C for 50 seconds (except Ret/Ptc2, 68°C), and 72°C for 1 minute for 40 to 50 cycles. In the SYBR Green Master Mix, there is an internal passive dye, ROX, in addition to the SYBR Green dye. The increase in the fluorescence of SYBR Green against that of ROX was measured at the end of each cycle. In each 96-well reaction plate, six standard samples, diluted up to 1/1000 of cDNA of any lung cancer cell line selected in preliminary experiments of each gene, were run with unknown tumor samples. Finally, relative quantitative values of each sample were compared with those of their 18S ribosomal RNA (rRNA) genes (186 bp), since expression of 18S rRNA was more consistent than expression of β-actin (275 bp) or *GAPDH* (225 bp) among 20 cell lines in our preliminary experiments (ie, standard deviations of these housekeeping gene expressions were 0.87 for β-actin, 0.56 for *GAPDH*, and 0.25 for 18S rRNA when expression levels in the ACC-LC-319 cell line were set to 1.0).

Hierarchical Clustering

For using the cluster analysis program, we performed a logarithmic ($\log_2$) transformation of the data to stabilize the variance, and the gene expression profile of each tumor was normalized to the median gene expression level for the entire sample set. Average linkage hierarchical clustering was performed using Cluster and TreeView software (http://rana.lbl.gov/EisenSoftware.htm) [15].

Data Analysis

The χ^2 test, Student's *t* test, and Spearman rank correlation coefficient were used to compare the results. The Kaplan-Meier method was employed to estimate the probability of survival as a function of time, and survival differences were analyzed by the log-rank test. To identify which independent factors jointly had a significant influence on the overall survival, Cox proportional hazards modeling was applied. The two-sided significance level was set at $P < .05$. All analyses were performed using StatView software (version 5; SAS Institute Inc, Cary, NC [SAS/STAT User's Guide, Version 6; SAS Institute, 1990]).

RESULTS

Relative Quantification by Real-Time RT-PCR

We have preliminarily examined reliability of this assay, including reproducibility using cell line samples. Four of the 48 genes (*EST AA468094, DDC* [dopa decarboxylase], *HNF3A* [hepatocyte nuclear factor 3α], and *KIAA0767*) did not give consistent bands on agarose gels or identical melting curve of PCR products, and thus, subsequent analyses were performed using the remaining 44 genes. We next asked whether there was a good correlation between real-time RT-PCR assay and immunohistochemistry (IHC). Since we previously examined overlapping cohorts of tumors by IHC of thyroid transcription factor 1 (*TITF1*) [16], we compared the results obtained by RT-PCR with IHC. *TITF1* expressions were highly in agreement with mRNA quantities of our samples ($P < .0001$, Spearman rank correlation coefficient).

Hierarchical Clustering

Unsupervised hierarchical clustering based on $\log_2$ transformation of relative expression values of 44 genes classified 85 adenocarcinoma samples into a cluster tree with three major subgroups (Fig 1). These three clusters were independent of pathological stage ($P = .1732$, χ^2 test), differentiation ($P = .5498$), or degree of tumor content ($P = .4095$). Kaplan-Meier plots of one of the subgroups (group B) showed a statistically significant difference in overall survival as compared with the other two subgroups ($P = .0297$, log-rank test; Fig 2). Group B included preferentially more males ($P = .0020$, χ^2 test), more elderly patients ($P = .0277$), more smokers ($P = .0197$), and more tumors larger than 30 mm ($P = .0019$; Table 2). Tumors in group B had lower levels of expression of *DPP4* (dipeptidylpeptidase IV), *WFDC2* (whey-acidic protein four-disulfide core domain 2), *TITF1*, *CIT* (citron), *HPN* (hepsin), *PTK7* (protein tyrosine kinase 7), *ICAM1* (intercellular adhesion molecule 1), *DOC-1R* (deleted in oral cancer-related 1), *CEACAM1* (carcinoembryonic antigen-related cell adhesion molecule 1), *COL9A2* (collagen type IX alpha 2), *SDCCAG28* (serologically defined colon cancer antigen 28), and two *EST* genes. On the other hand, group B tumors had higher expression of *DUSP4* (dual specificity phosphatase 4), *TRIM29* (tripartite motif-containing 29), and *SLC2A1* (solute carrier family 2 member 1; Table 3). Expression of neuroendocrine markers such as *KLK11* (kallikrein 11) and *ASCL1* (ASH1, achaete-scute homolog 1) were highest in tumors of group C. Patients of this group did not, however, have poor prognosis.

Identification of a Smaller Number of Survival-Related Genes and Creation of a Prognostic Model

Although it is of interest that the above unsupervised cluster analysis of 44 genes was able to identify patients at higher risk, examination of 44 genes is too laborious in clinical practice. Furthermore, the hierarchical clustering method can only be applicable to a retrospective analysis of a cohort of patients and cannot be used to predict clinical outcome for any future patients. Therefore, we tried to identify a smaller number of genes relevant to patient prognosis, and to create a prognostic model that could be applied prospectively. For the selection of genes, a stepwise multivariate Cox proportional hazards model was used, and eight genes were selected as significant independent prognostic factors (Table 4) when the cutoff *P* value was set at .1. The genes were *PTK7*, *CIT*, *SCNN1A* (sodium channel, nonvoltage-gated 1 alpha), *PGES* (prostaglandin E synthase), *ERO1L* (*ERO1*-like), *ZWINT* (ZW10 interactor), and two *EST*s. Coefficients for *PTK7*, *CIT*, *ERO1L*, and *EST AA434256* were negative, while those for the other four genes were positive, suggesting that high expression of the former four genes was associated with good prognosis, and that high expression of the latter four genes was associated with poor prognosis. We then calculated risk indices (RIs) that were defined as a linear combination of gene expression values weighed by their estimated regression coefficients. When the cutoff of the RI was set at the 50th percentile,

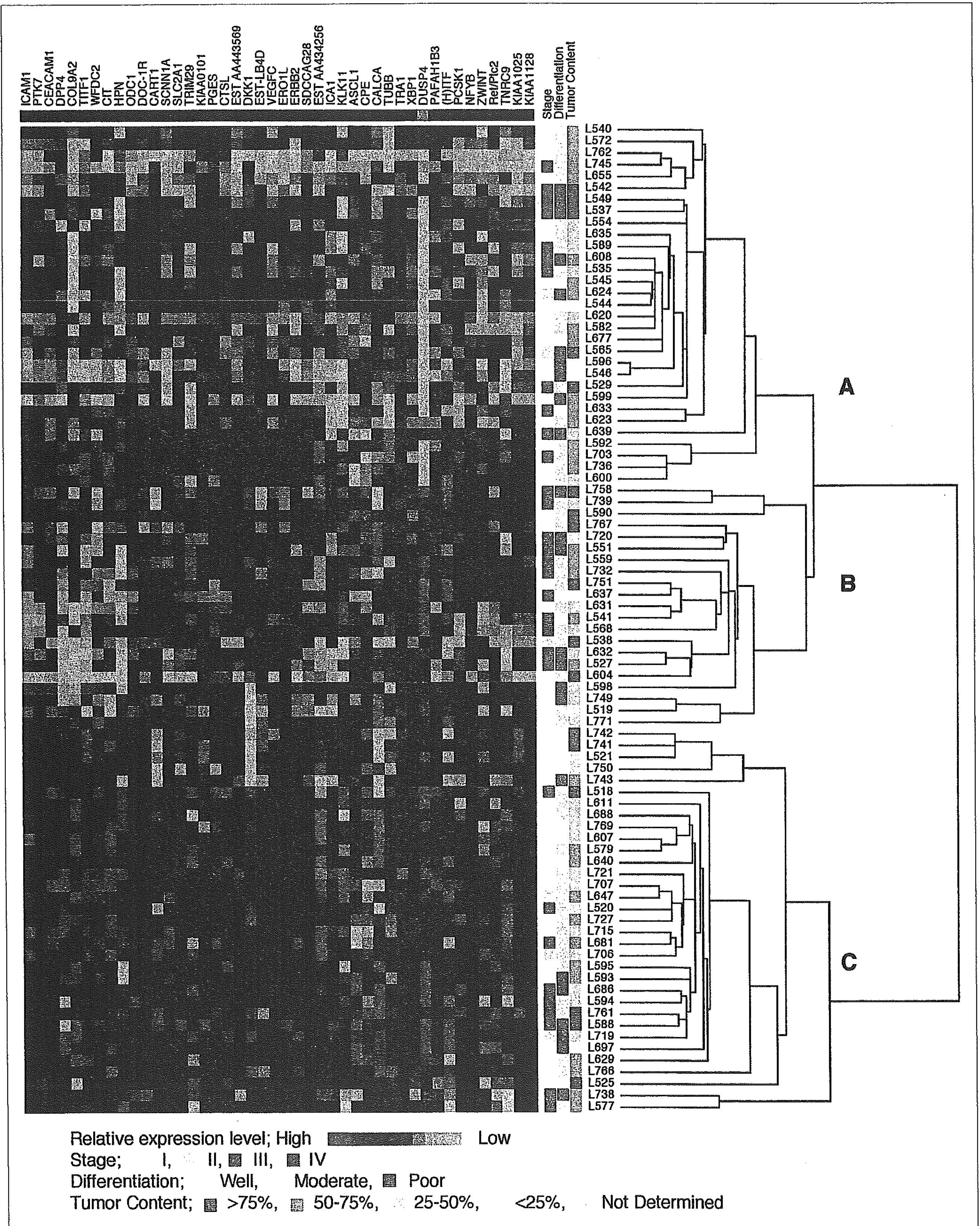


Fig 1. Gene expression profile. Eighty-five patients with clinical outcome data were clustered hierarchically based on expression of 44 genes. Cluster tree of individual patient samples and overall pattern of median-centered gene expression data are shown.

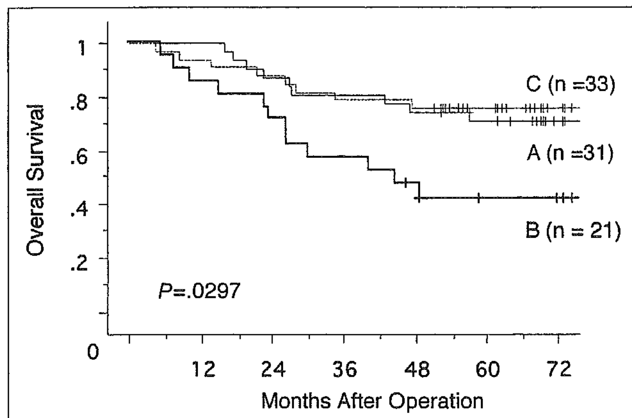


Fig 2. Kaplan-Meier curve divided by hierarchical clustering using all 44 genes and 85 cases.

Kaplan-Meier analysis and a log-rank test showed that there was a large difference in survival between high- and low-risk patients ($P < .0001$, log-rank test; Fig 3A). There was a significant association between RIs and tumor stage or tumor differentiation. Seventy percent of patients with stage III and IV disease, and 40% of patients with stage I and II disease belonged to the high risk group ($P = .0082$, χ^2 test). In addition, the high-risk group contained 57% of patients with poorly to moderately differentiated carcinoma, and only 15% of those with well-differentiated carcinoma ($P = .0052$). The prognostic stratification was even more prominent when patients were divided into four groups by setting the cutoffs at the 25th, 50th, and 75th percentiles (Fig 3B).

We also created a model including conventional prognostic factors (sex, age, differentiation, and stage) as well as RI. The Cox proportional hazards model selected pathological stage (hazard ratio [HR], 2.915; $P = .0069$) and RI (HR, 5.017; $P = .0021$) as independent prognostic factors (Table 5).

Since these eight genes were selected and regression coefficients were calculated to explain prognosis of these learning set of 85 patients, it was important to validate this approach by applying the model to an independent set of patients. Therefore, quantitative RT-PCR was performed for these eight genes, and RI was calculated in a test set of 21 additional adenocarcinoma patients who had undergone surgery from 1994 through 1995—a period just before the learning set of cases was collected. The Kaplan-Meier curves stratified by RI are shown in Figure 4. The difference in overall survival between patients with low RI and high RI was also significant ($P = .0273$, log-rank test), validating this model. As with the learning set, multivariate analysis with the test cases also identified RI (HR, 6.562; $P = .0085$) and pathological stage (HR, 14.819; $P = .0038$) as independent prognostic factors (Table 6).

Table 2. Relationship Between Hierarchical Cluster and Clinicopathologic Variables (χ^2 test)

	Group B (n = 21)		Groups A and C (n = 64)		
Variable	No. of Patients	%	No. of Patients	%	P
Sex					
Male	17	81	27	42	.0020
Female	4	19	37	58	
Age, years					
≥ 62	15	71	28	44	.0277
62 >	6	29	36	56	
Smoking habit					
Smoker	14	67	24	38	.0197
Never	7	33	40	62	
CEA, ng/mL					
≥ 5	10	48	19	30	.1326
5 >	11	52	45	70	
Differentiation					
Poor	7	33	17	27	.5498
Well + moderate	14	67	47	73	
Tumor size, mm					
> 30	15	71	21	33	.0019
≤ 30	6	29	43	67	
Pleural invasion					
+	8	38	25	39	.9371
−	13	62	39	61	
LN metastasis					
N1-3	8	38	22	34	.7569
N0	13	62	42	66	
p-Stage					
III + IV	10	48	20	31	.1732
I + II	11	52	44	69	
Tumor content, %					
< 25	1	5	4	6	.4095
25-50	7	37	18	30	
50-75	5	26	28	46	
> 75	6	32	11	18	

Abbreviations: CEA, carcinoembryonic antigen in serum; LN, lymph node.

DISCUSSION

cDNA microarray technology is promising in that it can analyze thousands of unselected genes simultaneously, and thus, it gives a comprehensive view of gene expression characteristics of tumors in each patient. However, microarray technology is still developing, and it has potential technical variances that may compromise the reproducibility of results. These technical variances may be derived from variation in printing or processing of chips, hybridization or scanning, sample preparation, or probes [17]. In addition, cDNA chips are still very expensive for routine clinical use. As a complementary approach, we determined the expression levels of each gene by quantitative RT-PCR, which can provide more accurate and reproducible RNA quantification and requires smaller quantities of tumor tissue [18-20].

In the present study, we did not use microdissected tissue but used bulk of cancer tissue. This may raise argument that contamination of normal stromal cells may dilute molecular characteristic of tumors of interest. However, we have shown that the degree of tumor content did not signif-

Table 3. Characterization of Group B Divided by Hierarchical Cluster (t test in $\log_2$ transformation; median centered data)

Variable	Group B	Groups A and C	P
Genes with decreased expression in group B			
DPP4	−2.006	0.433	< .0001
WFDC2	−1.27	0.283	< .0001
TITF1	−1.472	0.238	.0001
CIT	−0.876	0.754	.0002
HPN	−1.726	0.093	.0003
EST AA434256	−0.821	0.654	.0008
PTK7	−0.563	0.106	.0026
ICAM1	−0.722	0.104	.0029
DOC-1R	−0.51	0.07	.0079
CEACAM1	−0.4	0.274	.0108
EST-LB4D	−0.437	0.088	.0225
COL9A2	−1.272	0.072	.0239
SDCCAG28	−0.47	−0.051	.0261
Genes with increased expression in group B			
DUSP4	1.587	−2.196	.0007
TRIM29	0.981	−0.231	.0046
SLC2A1	0.556	0.002	.0438

Abbreviation: $\log_2$, logarithm.

icantly affect clustering. In a clinical setting, it is too laborious to perform microdissection for each patient. Furthermore, gene-expression signature might arise from both malignant and stromal elements in primary tumors, as suggested by Ramaswamy et al [10].

We first analyzed our expression data using unsupervised hierarchical clustering. Hierarchical clustering separates samples into subgroups of related expression patterns in an unbiased manner. Although the number of genes we examined was small, we were able to cluster adenocarcinoma into three groups, confirming the heterogeneity of this type of NSCLC as has been suggested by other researchers [6-8]. Patients in group B by our expression profiling had a significantly poor prognosis. Group B had more male elderly patients with smoking habits and larger tumors. However, there was no statistical difference in terms of pleural invasion or lymph node metastasis. Tumors in group B were characterized by lower expression of adhesion molecules, such as *DPP4*, *ICAM1*, and *CEACAM1*, and growth suppressors such as *HPN* [21] and *DOC-1R* [22]. *ICAM1* is involved in intracellular signaling in a variety of physiological and pathological processes, including metastasis and tumor growth [23]. *WFDC2*, *TITF1*, and *CIT* were also low in expression in group B. *WFDC2* (*HE4*), a protease inhibitor, is expressed in pulmonary epithelium and may be part of the host defense shield of the airways [24]. It was suggested to be a growth inhibitor [25], as similarly, its family member gene *WFDC1* is a candidate tumor suppressor gene. *TITF1* is implicated in the regulation of surfactant

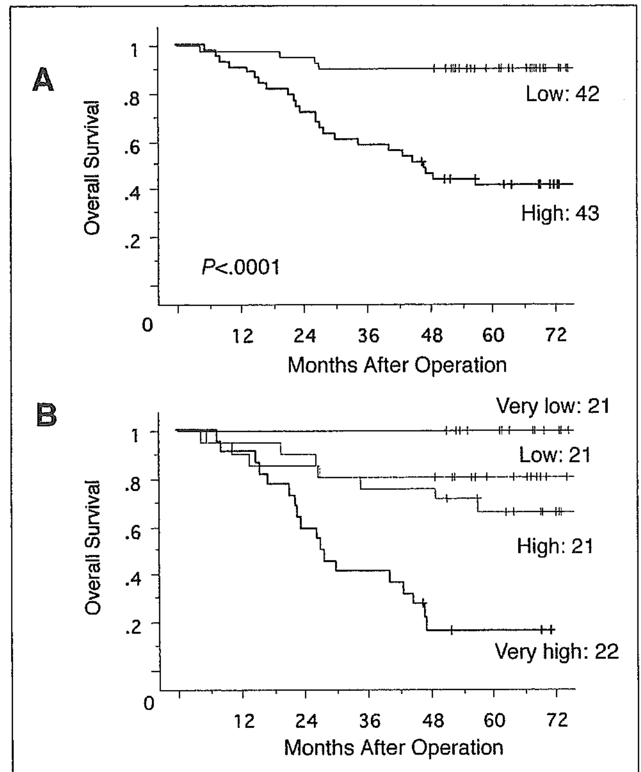


Fig 3. Relationship between survival in the 85 cases and their risk assignments based on the eight-gene risk index. The cutoff of the risk index was set at the 50th percentile (A) or the 25th, 50th, and 75th percentiles (B).

Table 4. Eight Genes Selected by Stepwise Multivariate Analysis of Overall Survival Using the Cox Proportional Hazards Model in 85 Cases

Variable	HR	95% CI	P
PTK7	0.500	0.244 to 1.024	.0582
CIT	0.611	0.428 to 0.870	.0064
SCNNIA	1.744	1.185 to 2.566	.0048
PGES	2.228	1.299 to 3.822	.0036
EST-LB4D	1.962	1.219 to 3.158	.0055
ERO1L	0.289	0.157 to 0.531	< .0001
EST AA434256	0.481	0.342 to 0.677	< .0001
ZWINT	2.398	1.586 to 3.626	< .0001

Abbreviation: HR, hazard ratio.

Table 5. Multivariate Analysis of Overall Survival Using the Cox Proportional Hazards Model in 85 Cases

Variable	Category	HR	95% CI	P
Sex	Male/female	1.451	0.662 to 3.179	.3525
Age, years	≥ 62/< 62	1.102	0.513 to 2.367	.8033
Differentiation	Well + moderate/poor	1.311	0.581 to 2.957	.5143
p-Stage	III+IV/I+II	2.915	1.341 to 6.340	.0069
Risk index	High/low*	5.017	1.795 to 14.018	.0021

Abbreviation: HR, hazard ratio.
*Divided by median value.

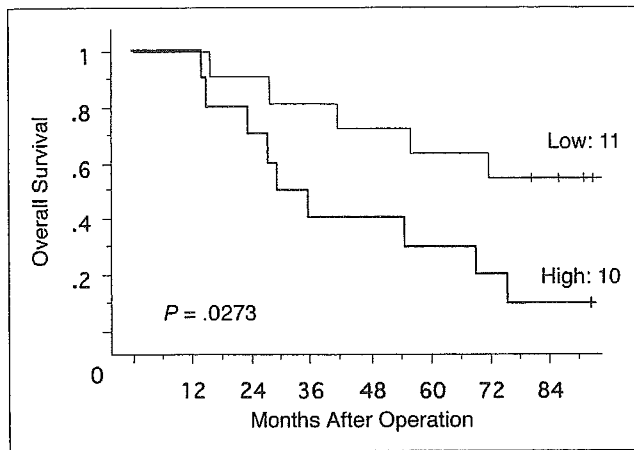


Fig 4. Relationship between survival in the 21 test cases and their risk assignments based on the eight-gene risk index.

gene expression, and *TITF1* expression is initiated at a very early stage of lung morphogenesis [26]. CIT is a serine/threonine kinase and rho effector suggested as a crucial regulator in cytokinesis [27]. *SLC2A1* is also known as glucose transporter (*GLUT1*), and it was highly expressed in group B. Increased *SLC2A1* expression is observed under conditions that induce greater dependence on glycolysis as an energy source, such as ischemia or hypoxia [28-30]. These data suggest that overexpression of *SLC2A1* may play an important role in the survival tumor cells, especially in large tumors. In the literature, decreased expression of *ICAM1* [31] and *TITF1* [16], and increased expression of *SLC2A1* [32], have been previously reported to be poor prognostic factors in NSCLC.

For several genes, the prognostic impact appeared to be somewhat at odds with their function. Although protein tyrosine kinases (PTKs) are regarded as oncogenes, *PTK7* was more strongly expressed in tumors with favorable prognosis, in agreement with the Stanford group [6]. It is also reported to be downregulated during melanoma development [33]. *DUSP4* (*MKP-2*) is one of the dual specificity phosphatases, which are also known as MKPs (mitogen-activated protein kinase phosphatases). *DUSP4* inactivates MAPKs (extracellular signal-regulated kinase and c-Jun *N*-terminal kinase) [34] and is thought to be a negative growth regulator. However, high expression of *DUSP4* was associated with poor prognosis in the present study, in agreement with the Boston group [7] and with the Stanford group [6]. Indeed, *DUSP4* was recently reported to be a possible component of a novel, transforming pathway [35], and to be highly expressed in hepatoma [36] and pancreatic tumor cell lines [37]. Moreover, it is reported that expression of *MKP-1*, highly homologous to *DUSP4*, is associated with a shorter progression-free survival in ovarian tumors [38]. Although the Boston group reported that adenocarcinoma with overexpression of neuroendocrine markers had the worst prognosis [7], our group C, which had highest expression of neuroendocrine markers, did not have poor prognosis.

Table 6. Multivariate Analysis of Overall Survival Using the Cox Proportional Hazards Model in 21 Test Cases

Variable*	Category	HR	95% CI	*P* Value
Sex	Female/male	2.083	0.621 to 6.981	.2344
Age, years	≥ 60/< 60	5.385	1.126 to 25.744	.0349
p-Stage	III+IV/I+II	14.819	2.387 to 91.994	.0038
Risk index	High/low†	6.562	1.617 to 26.625	.0085

Abbreviation: HR, hazard ratio.
*There was no poorly differentiated adenocarcinoma in the 21 cases.
†Divided by median value.

Using Cox proportional hazards model, we selected eight genes that would jointly predict patient prognosis. These were *PTK7*, *CIT*, *SCNN1A*, *PGES*, *ERO1L*, *ZWINT*, and two ESTs, and half of these genes also characterized hierarchical cluster group B. *PGES* has been shown to stimulate transcription, influence mitogenesis of normal human bone cells, and to promote growth and metastasis of tumors [39]. PGE_2 is synthesized from PGH_2, whose production is catalyzed by cyclooxygenase (*COX*), and inducible *PGES* is specifically overexpressed in NSCLC [40]. Overexpression of inducible form of *COX* (*COX-2*) is shown to be associated with poor prognosis in NSCLC patients [41]. *ZWINT*, human ZW10 interacting protein-1, plays an important role in normal centromere function [42]. *SCNN1A* is related to hypertension [43], and *ERO1L* is thought to be involved in oxidative protein folding in the endoplasmic reticulum [44]. Although the mechanisms by which some of these genes affect patient prognosis are not very clear, it seemed that whole set of expression pattern of these genes might have the important information for cancer prognosis.

In conclusion, we were able to select eight genes that were more relevant to patients' survival from expression data obtained by real-time PCR. Both unsupervised clustering and a supervised method were useful to predict survival of patients with pulmonary adenocarcinoma. It is of special clinical interest that the risk index calculated from expression of a reasonably small number of genes may be useful in routine clinical practice.

Acknowledgment

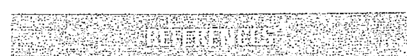
We thank Dr Hamajima, professor of Nagoya University, for help in statistical approaches such as Cox proportional hazards regression modeling.

Appendix

The appendix is included in the full-text version of this article, available on-line at www.jco.org. It is not included in the PDF (via Adobe® Acrobat Reader®) version.

Authors' Disclosures of Potential Conflicts of Interest

The authors indicated no potential conflicts of interest.



1. Kuroishi T, Hirose K, Tajima K, et al: Cancer mortality in Japan, in Tominaga S, Oshima A (eds): Cancer mortality and morbidity statistics. Gann Monograph Cancer Research, No. 47, Tokyo, Japan, Japan Scientific Societies Press, 1999, pp 1-38

2. Vaporciyan AA, Nesbitt JC, Lee JS, et al: Cancer of the lung, in Bast RC, Kufe DW, Pollock RE, et al (eds): Cancer Medicine (ed 5). Hamilton, Canada, B.C. Decker Inc, 2000, pp 1227-1292

3. Balsara BR, Testa JR: Chromosomal imbalance in human lung cancer. Oncogene 21: 6877-6883, 2002

4. Fry WA, Phillips JL, Menck HR: Ten-year survey of lung cancer treatment and survival in hospitals in the United States: A national cancer data base report. Cancer 86:1867-1876, 1999

5. Brundage MD, Davies D, Mackillop WJ: Prognostic factors in non-small cell lung cancer: A decade of progress. Chest 122:1037-1057, 2002

6. Garber ME, Troyanskaya OG, Schluens K, et al: Diversity of gene expression in adenocarcinoma of the lung. Proc Natl Acad Sci U S A 98:13784-13789, 2001

7. Bhattacharjee A, Richards WG, Staunton J, et al: Classification of human lung carcinomas by mRNA expression profiling reveals distinct adenocarcinoma subclasses. Proc Natl Acad Sci U S A 98:13790-13795, 2001

8. Beer DG, Kardia SL, Huang CC, et al: Gene-expression profiles predict survival of patients with lung adenocarcinoma. Nat Med 8:816-824, 2002

9. Wigle DA, Jurisica I, Radulovich N, et al: Molecular profiling of non-small cell lung cancer and correlation with disease-free survival. Cancer Res 62:3005-3008, 2002

10. Ramaswamy S, Ross KN, Lander ES, et al: A molecular signature of metastasis in primary solid tumors. Nat Genet 33:49-54, 2003

11. Takahashi T, Ueda R, Song X, et al: Two novel cell surface antigens on small cell lung carcinoma defined by mouse monoclonal antibodies NE-25 and PE-35. Cancer Res 46:4770-4775, 1986

12. Hida T, Ariyoshi Y, Kuwabara M, et al: Glutathione S-transferase pi levels in a panel of lung cancer cell lines and its relation to chemoradiosensitivity. Jpn J Clin Oncol 23:14-19, 1993

13. Mountain CF: Revisions in the international system for staging lung cancer. Chest 111:1710-1717, 1997

14. Sambrook J, Fritsch EF, Maniatis T: Isolation of total RNA by extraction with strong denaturants, in Molecular Cloning: A Laboratory Manual, New York, NY, Cold Spring Harbor Laboratory Press, 1989, pp 7.18-7.25

15. Eisen MB, Spellman PT, Brown PO, et al: Cluster analysis and display of genome-wide expression patterns. Proc Natl Acad Sci U S A 95:14863-14868, 1998

16. Yatabe Y, Mitsudomi T, Takahashi T: TTF-1 expression in pulmonary adenocarcinomas. Am J Surg Pathol 26:767-773, 2002

17. Miller LD, Long PM, Wong L, et al: Optimal gene expression analysis by microarrays. Cancer Cell 2:353-361, 2002

18. Dhar AK, Roux MM, Klimpel KR: Detection and quantification of infectious hypodermal and hematopoietic necrosis virus and white spot virus in shrimp using real-time quantitative PCR and SYBR Green chemistry. J Clin Microbiol 39:2835-2845, 2001

19. Nazarenko I, Lowe B, Darfler M, et al: Multiplex quantitative PCR using self-quenched primers labeled with a single fluorophore. Nucleic Acids Res 30:e37, 2002

20. Riou P, Saffroy R, Comoy J, et al: Investigation in liver tissues and cell lines of the transcription of 13 genes mapping to the 16q24 region that are frequently deleted in hepatocellular carcinoma. Clin Cancer Res 8:3178-8316, 2002

21. Srikantan V, Valladares M, Rhim JS, et al: HEPSIN inhibits cell growth/invasion in prostate cancer cells. Cancer Res 62:6812-6816, 2002

22. Zhang X, Tsao H, Tsuji T, et al: Identification and mutation analysis of DOC-1R, a DOC-1 growth suppressor-related gene. Biochem Biophys Res Commun 255:59-63, 1999

23. Yasuda M, Tanaka Y, Tamura M, et al: Stimulation of beta1 integrin down-regulates ICAM-1 expression and ICAM-1-dependent adhesion of lung cancer cells through focal adhesion kinase. Cancer Res 61:2022-2030, 2001

24. Bingle L, Singleton V, Bingle CD: The putative ovarian tumour marker gene HE4 (WFDC2), is expressed in normal tissues and undergoes complex alternative splicing to yield multiple protein isoforms. Oncogene 21:2768-2773, 2002

25. Larsen M, Ressler SJ, Lu B, et al: Molecular cloning and expression of ps20 growth inhibitor: A novel WAP-type "four-disulfide core" domain protein expressed in smooth muscle. J Biol Chem 273:4574-4584, 1998

26. Minoo P, Su G, Drum H: Defects in traceoesophageal and lung morphogenesis in Nkx2.1(-/-) mouse embryos. Dev Biol 209:60-71, 1999

27. Di Cunto F, Calautti E, Hsiao J, et al: Citron rho-interacting kinase, a novel tissue-specific ser/thr kinase encompassing the Rho-Rac-binding protein Citron. J Biol Chem 273:29706-29711, 1998

28. Kurata T, Oguri T, Isobe T, et al: Differential expression of facilitative glucose transporter (GLUT) genes in primary lung cancers and their liver metastases. Jpn J Cancer Res 90:1238-1243, 1999

29. Merrall NW, Plevin R, Gould GW: Growth factors, mitogens, oncogenes and the regulation of glucose transport. Cell Signal 5:667-675, 1993

30. Clavo A, Brown R, Wahl R: 2-fluoro-2-deoxy-D-glucose (FDG) uptake into human cancer cell lines is increased by hypoxia. J Nucl Med 36:1625-1632, 1995

31. Passlick B, Izbicki JR, Simmel S, et al: Expression of major histocompatibility class I and class II antigens and intracellular adhesion molecule-1 on operable non-small cell lung carcinomas: Frequency and prognostic significance. Eur J Cancer 30:376-381, 1994

32. Younes M, Brown RW, Stephenson M, et al: Overexpression of Glut1 and Glut3 in stage I nonsmall cell lung carcinoma is associated with poor survival. Cancer 80:1046-1051, 1997

33. Eastey DJ, Mitchell PJ, Patel K, et al: Loss of expression of receptor tyrosine kinase family genes PTK7 and SEK in metastatic melanoma. Int J Cancer 71:1061-1065, 1997

34. Hirsch DD, Stork PJ: Mitogen-activated protein kinase phosphatases inactivate stress-activated protein kinase pathways in vivo. J Biol Chem 272:4568-4575, 1997

35. Fu SL, Waha A, Vogt PK: Identification and characterization of genes upregulated in cells transformed by v-Jun. Oncogene 19:3537-3545, 2000

36. Yokoyama A, Karasaki H, Urushibara N, et al: The characteristic gene expressions of MAPK phosphatases 1 and 2 in hepatocarcinogenesis, rat ascites hepatoma cells, and regenerating rat liver. Biochem Biophys Res Commun 239:746-751, 1997

37. Yip-Schneider MT, Lin A, Marshall MS: Pancreatic tumor cells with mutant K-ras suppress ERK activity by MEK-dependent induction of MAP kinase phosphatase-2. Biochem Biophys Res Commun 280:992-997, 2001

38. Denkert C, Schmitt WD, Berger S, et al: Expression of mitogen-activated protein kinase phosphatase-1 (MKP-1) in primary human ovarian carcinoma. Int J Cancer 102:507-513, 2002

39. Sales KJ, Katz AA, Howard B, et al: Cyclooxygenase-1 is up-regulated in cervical carcinomas: Autocrine/paracrine regulation of cyclooxygenase-2, prostaglandin E receptors, and angiogenic factors by cyclooxygenase-1. Cancer Res 62:424-432, 2002

40. Yoshimatsu K, Altorki NK, Golijanin D, et al: Inducible prostaglandin E synthase is overexpressed in non-small cell lung cancer. Clin Cancer Res 7:2669-2674, 2001

41. Achiwa H, Yatabe Y, Hida T, et al: Prognostic significance of elevated cyclooxygenase 2 expression in primary, resected lung adenocarcinomas. Clin Cancer Res 5:1001-1005, 1999

42. Starr DA, Saffery R, Li Z, et al: Hzwint-1, a novel human kinetochore component that interacts with HZW10. J Cell Sci 113:1939-1950, 2000

43. Iwai N, Baba S, Mannami T, et al: Association of a sodium channel alpha subunit promoter variant with blood pressure. J Am Soc Nephrol 13:80-85, 2002

44. Cabibbo A, Pagani M, Fabbri M, et al: ERO1-L, a human protein that favors disulfide bond formation in the endoplasmic reticulum. J Biol Chem 275:4827-4833, 2000

Cancer Statistics Digest

Mortality trend of uterine cancer in Japan: 1960–2000

Time trends in mortality rates and number of deaths of overall uterine cancer are presented from 1960 to 2000 (Fig. 1). The age-standardized uterine cancer mortality rate (ASR) decreased up to the mid-1990s. The tendency of decline in the ASR seems to have stopped in recent years. The decrease after 1960 onwards was observed in all age groups except for females aged 85+ years. Among females aged 25–59 years no further decline, but a consecutive increase was observed. Years when the trends in mortality switched from decrease to increase appeared earlier among younger females than among older females.

The total number of uterine cancer deaths (Fig. 1) has been gradually decreasing for three decades since 1960: 7068 in 1960 to 4444 in 1993 (for age 0+ years). Since 1993, the total number of uterine cancer deaths has shown a constant increase and reached 5202 in 2000.

The number of deaths of the cancer of the cervix, corpus and unspecified is presented in Fig. 2 (left). 'Unspecified' uterine cancer had been decreasing since 1960. Since 1988, the percentage with a specific cancer site identified, i.e. cervix or corpus, was over 50% of overall uterine cancer. In 2000, the number of deaths from cervical cancer was approximately twice that of corpus cancer, although the proportion of 'unspecified' cancer of the uterus was still 32%. The percentages of specific sites compared to all uterine cancer by 10-year age groups in 2000 are presented (Fig. 2, right). Although the percentages of unspecified uterine cancer varied by age group, cervical cancer tended to be more frequent among younger females compared to older females.

The proportion of uterine cancer relative to total cancer deaths has declined from about 17% in 1960 to 4% in 2000 (Fig. 3). Note that ICD code of uterine cancer was 179–182 for ICD-7 in 1960–1967, 179–182 for ICD-8 in 1968–1978, 179–182 for ICD-9 in 1979–1994, and C53-C55 for ICD-10 in 1995–2000.

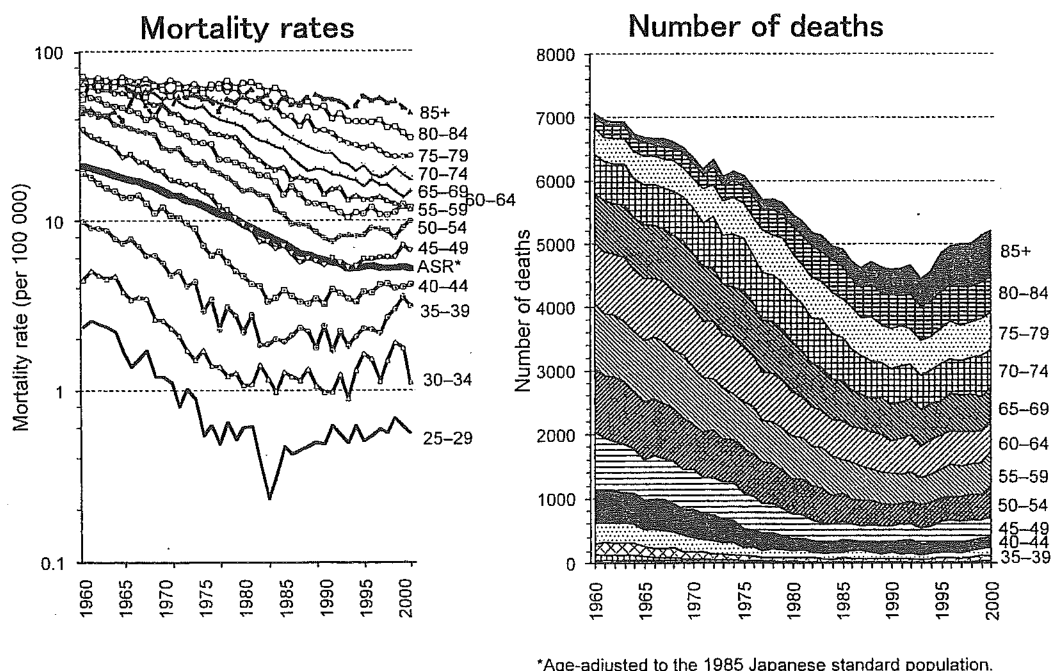


Figure 1. Time trends of uterine cancer mortality rates and number of deaths.

© 2004 Foundation for Promotion of Cancer Researcl