肺癌による一側肺全摘後の予測呼吸機能に関する検討

安川元章[1]・中川勝裕[1]・阪口全宏[1]・
岩崎輝夫[1]・佐々木尚子[1]

要旨——目的．肺癌による一側肺全摘例につき，術前の呼吸器疾患合併例と非合併例にわけ，それぞれの予測呼吸機能と実測値との乖離につき検討する．方法．1990年から2001年までの当院での肺癌全摘例133例中，術後合併症を併発することなしに術後3ヶ月から6ヶ月の間に呼吸機能検査を施行できた33例を対象とした．術前のスパイロメトリーと肺血流シンチグラムから術後の努力肺活量（FVC），1秒量（$FEV_{1.0}$）の予測値を算出した．術後のFVC，$FEV_{1.0}$はスパイロメトリーで実測した．呼吸器疾患合併群21例と非合併群12例にわけ，予測残存率と実測残存率をpaired *t* testにて検定した．結果．呼吸器疾患非合併群ではFVCと$FEV_{1.0}$において，予測残存率と実測残存率に有意差を認めなかった．一方，呼吸器疾患合併例では予測残存率と実測残存率はFVCで$p<0.001$，$FEV_{1.0}$で$p=0.0486$で，両項目とも実測残存率は予測残存率よりも有意に低値であった．結論．呼吸器疾患合併症例では，実測残存率は予測残存率よりも有意に低値であったことより，術後残存呼吸機能は予測値よりも低下しうると考えられ，注意が必要である．
（肺癌．2004;44:683-687）

索引用語——肺癌，予測呼吸機能，肺血流シンチグラム

Prediction of Postoperative Pulmonary Function After Pneumonectomy for Lung Cancer

Motoaki Yasukawa[1]; *Katsuhiro Nakagawa*[1]; *Masahiro Sakaguchi*[1];
Teruo Iwasaki[1]; *Naoko Sasaki*[1]

ABSTRACT —— ***Objective.*** Patients undergone pneumonectomy for lung cancer were divided into two groups according to the presence or absence of preoperative respiratory complications, and discordances between predicted respiratory functions and actual values are examined for each group. ***Methods.*** Of the 133 cases with pneumonectomy for lung cancer carried out at our hospital in the period from 1990 to 2001,33 were selected as study subjects, among which no incidence of postoperative complications was observed and in whom respiratory function tests could be performed 3 to 6 months after the surgical treatment. Prior to the operation, predicted values of forced vital capacity (FVC) and forced expiratory volume in 1 second ($FEV_{1.0}$) were calculated using blood flow scintigrams, and actual values of FVC and $FEV_{1.0}$ were determined postoperatively using spirometry. The subjects were divided into two groups with or without preoperative respiratory disease complications, and predicted and actual functional persistence rates were compared by Student's paired t-test for each group. ***Results.*** There were 21 cases of preoperative respiratory disease complications. In the group without any respiratory disease complications, there were no statistically significant differences in either FVC and $FEV_{1.0}$ between predicted and actual values. However, in the respiratory disease complication group, there were statistically significant differences between predicted and actual values in FVC ($p<0.001$) and $FEV_{1.0}$ ($p=$

[1]大阪府立呼吸器・アレルギー医療センター呼吸器外科．
別刷請求先：安川元章，大阪府立呼吸器・アレルギー医療センター呼吸器外科，〒583-8588 大阪府羽曳野市はびきの3-7-1，大阪府立呼吸器・アレルギー医療センター．
[1]Department of Thoracic Surgery, Osaka Prefectural Medical Center for Respiratory and Allergic Diseases, Osaka, Japan.
Reprints: Motoaki Yasukawa, Department of Thoracic Surgery, Osaka Prefectural Medical Center for Respiratory and Allergic Diseases, 3-7-1 Habikino, Habikino-shi, Osaka 583-8588, Japan.
Received April 5, 2004; accepted August 10, 2004.

0.0486), the actual functional persistence rates being significantly lower than the predicted values for both parameters. ***Conclusion.*** In cases of preoperative respiratory disease complications, actual postoperative respiratory functions might decline to a lower than predicted level, which requires cautions. (*JJLC.* 2004;44:683-687)

KEY WORDS —— Lung cancer, Predicted postoperative pulmonary function, Lung perfusion scanning

はじめに

肺癌手術の機能的適応を決定する手段として肺血流シンチグラムによる残存呼吸機能予測が多く用いられている．この方法による予測呼吸機能と術後実測値は強く相関するとされている.[1-5] しかし術後に，予測呼吸機能よりも実測値が低下する症例に遭遇する機会を経験することもある．今回我々は合併肺疾患の有無がこの乖離の原因と考え，肺癌全摘例につき検討したので報告する．

なお，呼吸機能の検討項目は術前評価に最も広範にわたり用いられている努力肺活量(FVC)，1秒量($FEV_{1.0}$)とし，術前の実測値に対する術後予測値，実測値のそれぞれの割合で検討した．

対象と方法

当院における機能面での肺全摘基準は予測残存 $FEV_{1.0}$ 800 ml 以上，術前の Eastern Cooperative Oncology Group (ECOG) performance status が0もしくは1としている．これを満たした肺全摘例は1990年から2001年までに133例であった．うち2例が術後肺炎で術死した．また治療を要した術後合併症は67例（不整脈32例，嗄声24例，肺炎16例，心不全5例，無気肺1例，その他9例）であった．今回の検討目的は，呼吸器疾患を有している症例における術後予測値の正確さを検討することにあるので，肺全摘例133例中，術後合併症を併発することなしに術後3ヶ月から6ヶ月の間に呼吸機能検査を施行でき，また切除肺と残存肺がほぼ同等の機能を有していると仮定するために術前の胸部CTで背景肺に左右差がないと判断された33症例とした．切除肺における肺病変は摘出肺を浸漬保存後，HE染色にて切除標本を作成し，腫瘍病変部以外を背景肺組織とし，合併肺疾患の有無を検討した．呼吸器疾患合併症例（C群）は21例で，内訳は肺気腫12例，閉塞性肺炎8例（すべて同側），肺線維症2例，珪肺1例，陳旧性肺結核1例であった（のべ数）．呼吸器疾患非合併例(N群)の12例は切除肺に有意の病理学的異常所見を認めなかった．手術前3週間以内にスパイロメトリーでFVC，$FEV_{1.0}$を測定した．また^{99m}Tc-MAA肺血流シンチグラムにて肺血流左右分布を測定し，残存予測FVCおよび$FEV_{1.0}$をこの分布比より算出し，術後予測残存率(＝術後予測値/術前実測値)として示した．また術後3ヶ月から6ヶ月目にスパイロメトリーでFVCと$FEV_{1.0}$を実測し，術前値を基準とし術後実測残存率(＝術後実測値/術前実測値)を算出した．paired t testにより両群間の比較を行った．

Table 1. Patient Characteristics

	No respiratory disease	Respiratory disease	P values
No	12	21	
Age	56.7 ± 11.2	60.7 ± 4.1	0.372
Tumor size [mm]	41.3 ± 25.2	57.0 ± 15.4	0.149
Sex (M/F)	12/0	20/1	0.773
Brinkman index	923 ± 421	1172 ± 656	0.209
Side (Lt/Rt)	8/4	15/6	0.914
His (Ad/Sq)	5/7	8/13	0.866
Stage ⅠA	0	0	0.310
ⅠB	2	1	
ⅡA	0	0	
ⅡB	4	2	
ⅢA	3	14	
ⅢB	3	4	
FVC [l]	3.79 ± 0.74	3.24 ± 0.71	0.038
%FVC [%]	108.8 ± 16.8	96.1 ± 16.1	0.040
$FEV_{1.0}$ [l]	2.80 ± 0.51	2.23 ± 0.47	0.002
$FEV_{1.0}$% [%]	75.6 ± 8.5	70.0 ± 9.2	0.097

No: Number of cases; Lt: Left; Rt: Right; Ad: Adenocarcinoma; Sq: Squamous cell carcinoma

結　果

C群とN群に年齢，性別，組織型，進行度，左右別，腫瘍径，Brinkman indexに有意差は認めなかった．術前FVCおよび$FEV_{1.0}$においてN群が有意に良好であった(Table 1)．

C群とN群各群別に術前後のFVC，$FEV_{1.0}$を比較検討した．N群のFVCで術後実測残存率(Y)は術後予測残存率(X)との間に$Y=0.9225X-0.0007$ ($n=12$, $\gamma=0.9954$, $p<0.001$)の相関を認めた．同様に$FEV_{1.0}$でも$Y=0.9972X-0.0005$ ($n=12$, $\gamma=0.9987$, $p<0.001$)の相関関係を認めた（Figure 1)．

C群のFVCで術後実測残存率（y）と術後予測残存率(x)との間には$y=0.8622x+0.0113$ ($n=21$, $\gamma=0.9881$, $p<0.001$)；$FEV_{1.0}$に関しては$y=0.8998x+0.0201$（$n=21$, $\gamma=0.9693$, $p<0.001$）の相関関係を認めた（Figure 2)．呼吸器疾患の有無を問わず，予測残存率と実測残存

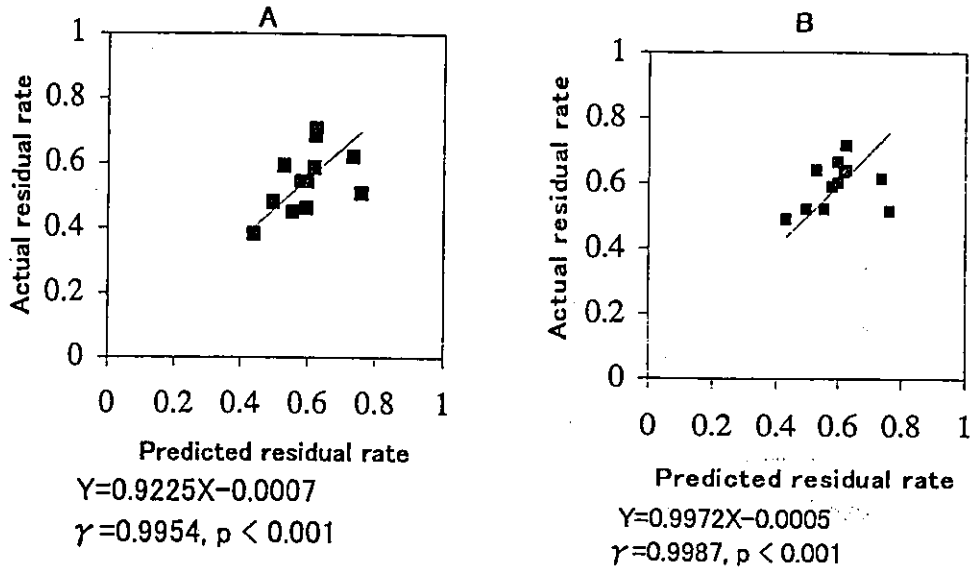


Figure 1. Correlation between predicted and actual postoperative residual rate of FVC and $FEV_{1.0}$ in patients without pulmonary parenchymal disease. A: FVC, B: $FEV_{1.0}$

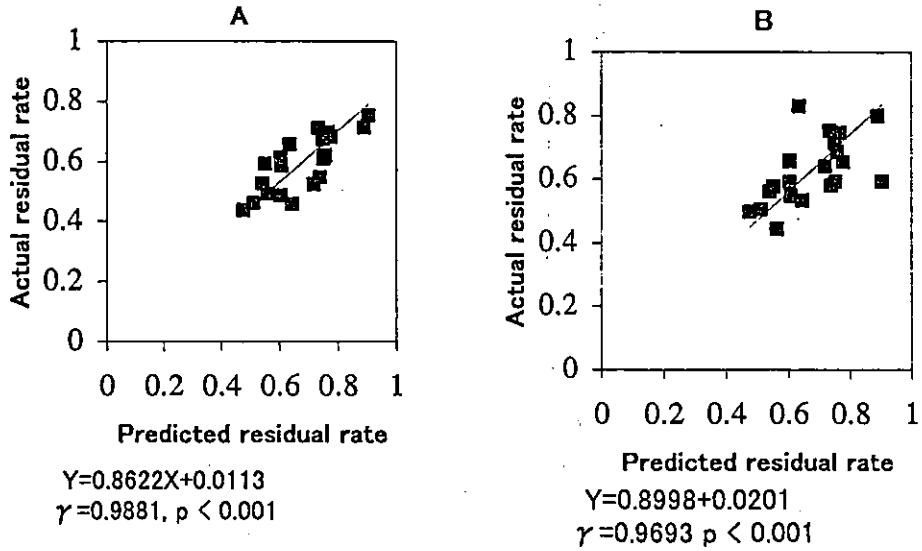
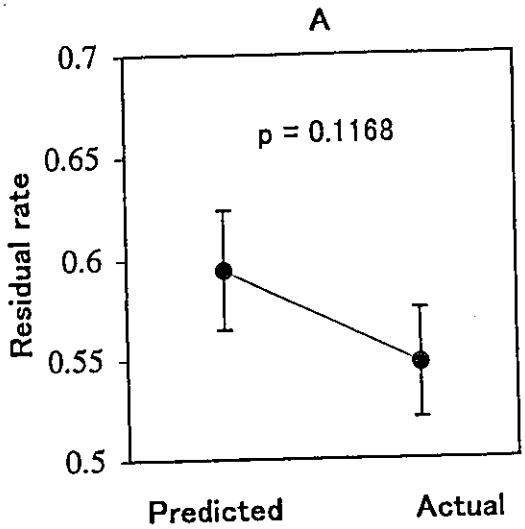


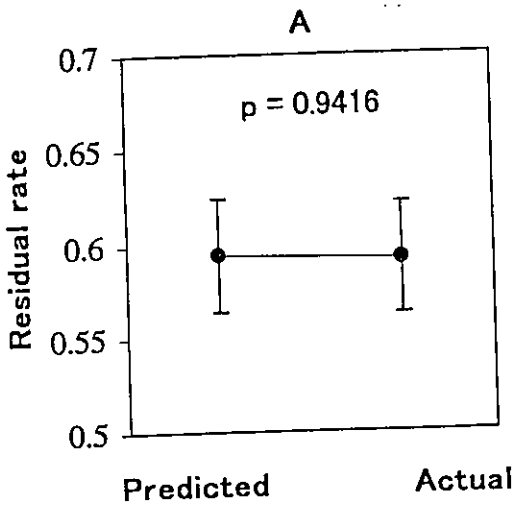
Figure 2. Correlation between predicted and actual postoperative residual rate of FVC and $FEV_{1.0}$ in patients with pulmonary parenchymal disease. A: FVC, B: $FEV_{1.0}$

率は非常に強い相関関係を認めた.

続いて，相関は認めるものの術後予測残存率と術後実測残存率に差があるのかを検討してみた(paired t test). N群のFVCで予測残存率は0.5946±0.5380，実測残存率は0.5486±0.4986でp値は0.1168，$FEV_{1.0}$で予測残存率は0.5946±0.5380，実測残存率は0.5925±0.5372でp値は0.9416とそれぞれ有意差を認めなかった（Figure 3 A，4A).

C群のFVCでは予測残存率は0.6735±0.3828，実測残存率は0.5920±0.3341で両値の間のp値は<0.001，$FEV_{1.0}$では予測残存率は0.6735±0.3828，実測残存率は0.6261±0.3554でp値は0.0486とそれぞれ実測残存率が予測残存率より有意に低値を示した(Figure 3B，4B). 以上からN群では予測残存率と実測残存率に差を認めないが，C群では実測残存率は予測残存率より有意に低値を示した.





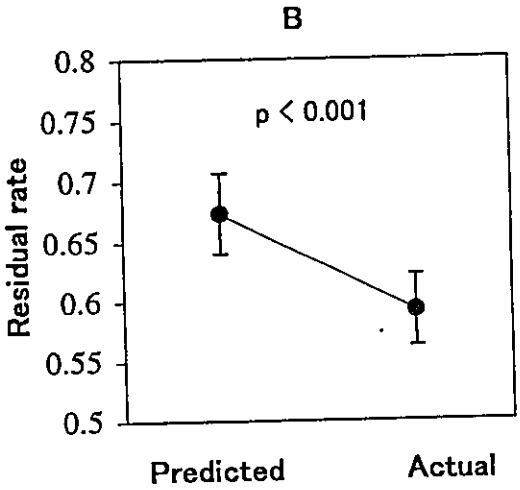


Figure 3. Student's *t* test result of predicted and actual postoperative residual rate of FVC in patients without and with pulmonary parenchymal disease. A: no respiratory disease, B: respiratory disease

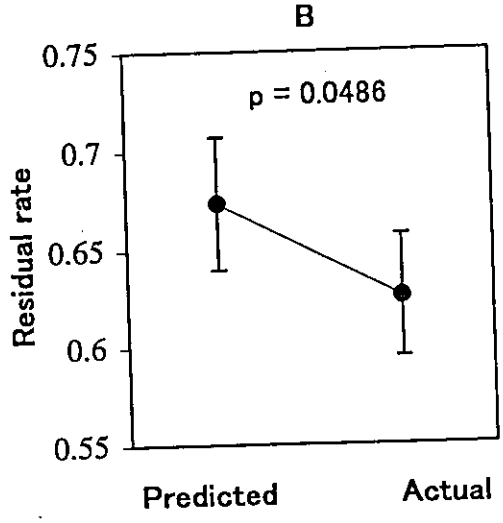


Figure 4. Student's *t* test result of predicted and actual postoperative residual rate of $FEV_{1.0}$ in patients without and with pulmonary parenchymal disease. A: no respiratory disease, B: respiratory disease

考　察

高齢化社会に伴い呼吸器疾患，特に肺実質病変を伴う肺癌症例に遭遇する機会が増えてきている．肺全摘術は侵襲も大きく，より厳密な術後呼吸機能予測が必要であると考える．呼吸器疾患合併の有無を考慮することにより，肺全摘後の残存呼吸機能がより正確に予測可能ではないかと考え，本検討を行った．

肺切除後の残存呼吸機能の予測は肺血流シンチグラム，quantitative CT などの方法が報告されており，いずれも肺切除後の残存呼吸機能の予測値と術後実測値は良好な相関関係を示すという報告[1-7]がなされている．肺血流シンチグラムを用いた術後予測呼吸機能として，中原ら[12]は 13 例の肺全摘例で，FVC は $\gamma=0.755$，$FEV_{1.0}$ は $\gamma=0.723$，Olsen ら[9]は 13 例の肺全摘例で，FVC は $\gamma=0.705$，$FEV_{1.0}$ は $\gamma=0.722$ との相関関係を報告している．諸家[1-6]の報告も同様であった．本検討において呼吸器疾患非合併群（N 群）では，FVC で $\gamma=0.9954$，$FEV_{1.0}$ で $\gamma=0.9987$，呼吸器疾患合併群（C 群）では FVC で $\gamma=0.9881$，

$FEV_{1.0}$ で $\gamma=0.9693$ と，これまでの報告と比較し良好な相関関係であった．従来の報告と比較し術後残存予測呼吸機能と実測呼吸機能により良好な相関関係が得られたのは，術後合併症例を除外したことの他に，呼吸器疾患合併例群と非合併例群にわけて検討したことが一因と考えられた．実際，呼吸器疾患の合併の有無にかかわらず，33 例全例にて予測値と実測値の相関関係を検討すると，FVC で $\gamma=0.7004$，$FEV_{1.0}$ で $\gamma=0.5271$ となった．

予測残存率と実測残存率に相関関係を認めたが，C 群（21 例）において FVC で 17 例，$FEV_{1.0}$ で 16 例が予測残存率より実測残存率が低値であった．さらに予測残存率と実測残存率の差を検定すると，N 群では FVC と $FEV_{1.0}$ の両項目において，予測残存率と実測残存率に有意差を認めなかったが，C 群では実測残存率は予測残存率より有意（FVC で $p<0.01$，$FEV_{1.0}$ で $p=0.0486$）の低下を認めた．

Olsen ら[9] は FVC，$FEV_{1.0}$ の両因子で実測値と予測値の間に 10% 以上の差を認めたものは 13 例中 10 例も存在したと報告している．我々の結果では FVC，$FEV_{1.0}$ の両者ともに 10% 以上の差が認められたのは 33 例中 10 例であり，いずれも実測値が予測値より低値を示した．うち C 群が 8 例であった．8 例の内訳は肺気腫が 6 例，閉塞性肺炎が 2 例，肺線維症が 1 例，肺結核が 1 例であった．

野々山ら[13] は予測値と実測値の差が 10% 以上生じる症例は，1）レントゲン上 2 葉以上にわたる無気肺所見，2）肺血流シンチグラムにて残存肺血流比が 75% 以上を認めるもの，の 2 つを挙げ，換気血流の不均等が大きくなることによるとしている．自験例で 10% 以上の差を認めた 10 例のうち，8 例で残存肺血流比が 70% 以上で，5 例が 75% 以上であった．このように FVC，$FEV_{1.0}$ の両者または，いずれかに 10% 以上の差を認めたものは，残存肺血流比が 70% 以上の症例が多かった．我々は換気シンチグラムを施行していないが，換気血流のミスマッチが生じた症例では予測呼吸機能より実測値は低下すると考えられた．

術後残存呼吸機能は，呼吸筋の損傷や術後の縦隔偏位の程度などを考慮する必要があるが，今回の結果から呼吸器疾患合併症例では術前の予測残存機能と術後実測呼吸機能に乖離が認められることを念頭に置いた手術適応の決定が必要であると考える．

呼吸器疾患合併症例での肺全摘術手術適応を見直すべきかについては，今回の検討症例はわずか 21 例の検討であり，今後症例を重ねて検討する必要がある．

REFERENCES

1. 良河光一．肺シンチグラフィーによる肺切除術後残存肺機能の予測に関する検討．日胸外会誌．1988;36:2437-2446.
2. Kristersson S, Lindell SE, Svanberg L. Prediction of pulmonary function loss due to pneumonectomy using 133 Xe-radiospirometry. *Chest*. 1972;62:694-698.
3. Ali MK, Mountain CF, Miller JM, et al. Regional pulmonary function before and after pneumonectomy using 133 xenon. *Chest*. 1975;68:288-296.
4. 小西　洋．肺癌患者における定量的肺シンチグラフィーによる術後呼吸機能の予測．日胸外会誌．1982;30:1784-1795.
5. Chen CY, Kao CH, Hsu NY, et al. Prediction probability of pneumonectomy for lung cancer using Tc-99 m MAA perfusion lung imaging. *Clin Nucl Med*. 1994;19:1094-1097.
6. Chris TB, Claudius G, Hermann E, et al. Prediction of functional reserves after lung resection: comparison between quantitative computed tomography, scintigraphy and anatomy. *Respiration*. 2002;69:482-489.
7. Wu MT, Chang JM, Chiang AA, et al. Use of quantitative CT to predict postoperative lung function in patients with lung cancer. *Radiology*. 1994;191:257-262.
8. 野々山明，斉藤幸人，大本一夫，他．高齢者（70 歳以上）の肺癌の肺剔徐術例の検討．日胸外会誌．1988;36:1096-1105.
9. Olsen GN, Block AJ, Tobias JA, et al. Prediction of postpneumonectomy pulmonary function using quantitative macroaggregate lung scanning. *Chest*. 1974;66:13-16.
10. Kristersson S, Arborelius M Jr, Jungquist G, et al. Prediction of ventilatory capacity after lobectomy. *Scand J Respir Dis*. 1973;54:315-325.
11. Wemly JA, DeMeester TR, Kirchner PT, et al. Clinical value of quantitative ventilation-perfusion lung scans in the surgical management of bronchogenic carcinoma. *J Thorac Cardiovasc Surg*. 1980;80:535-543.
12. 中原数也，藤本祐三郎，池田正人，他．手術術式別にみた原発性肺癌手術後予測呼吸機能と実測値の相関性について．胸部外科．1983;36:863-866.
13. 野々山明，斉藤幸人，田中一穂，他．肺血流シンチグラムを用いた肺剔除術例の術後肺機能の予測値についての検討．日胸外会誌．1988;36:65-72.

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.

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.

PATIENTS AND METHODS

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′-GTCAACGGATTTGGTCGTATT-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,

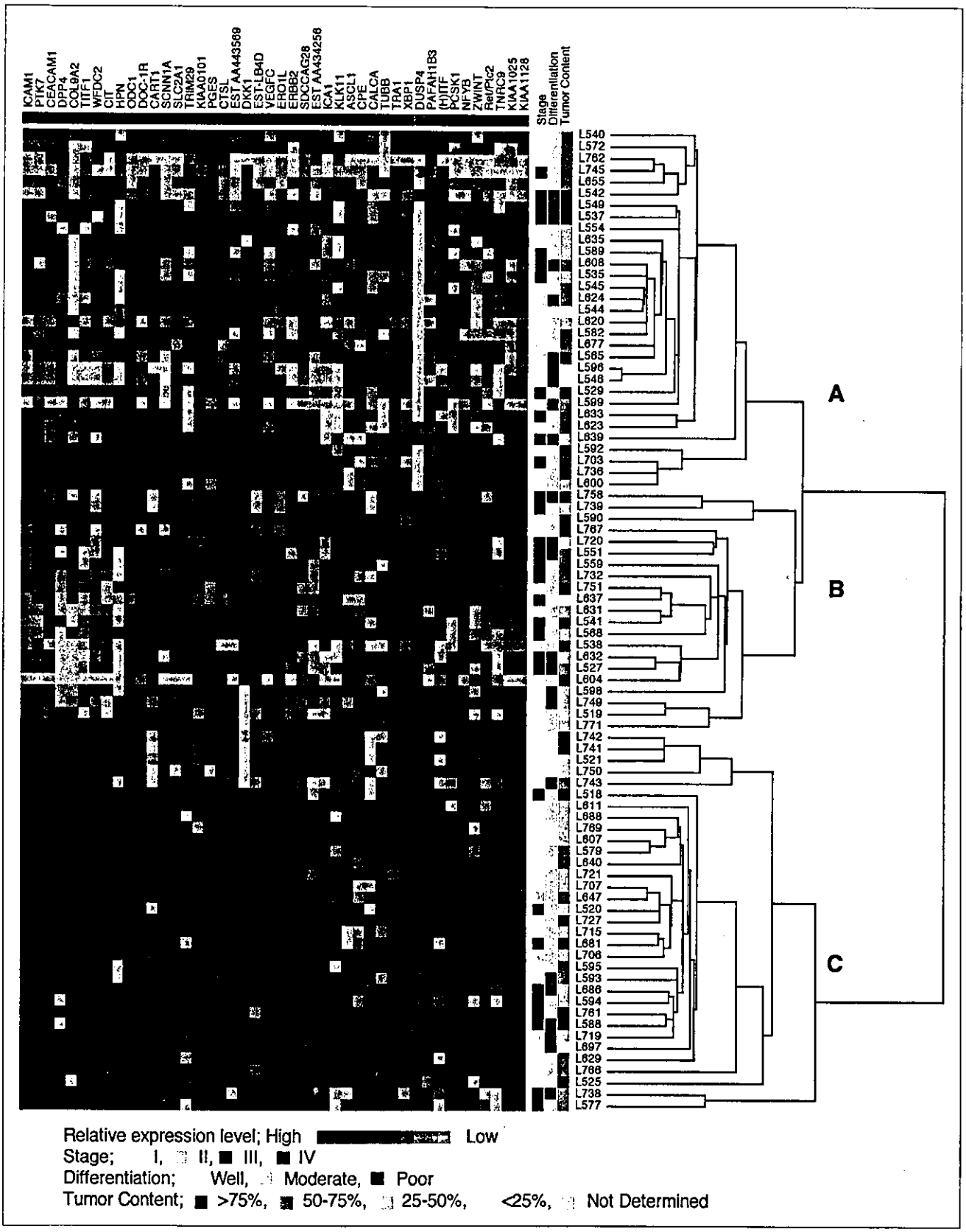


Fig 1. Gene expression profile. Eighty-five patients with clinical outcome data were clustered hierarchically based on expression of 44 genes. Cluster tree of individual patient samples and overall pattern of median-centered gene expression data are shown.

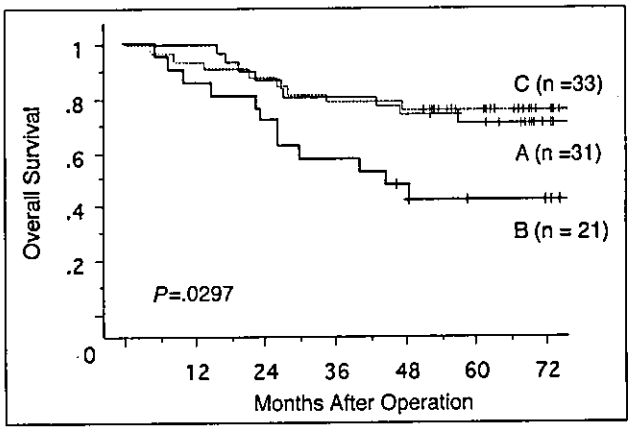


Fig 2. Kaplan-Meier curve divided by hierarchical clustering using all 44 genes and 85 cases.

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)

Variable	Group B (n = 21) No. of Patients	%	Groups A and C (n = 64) 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.

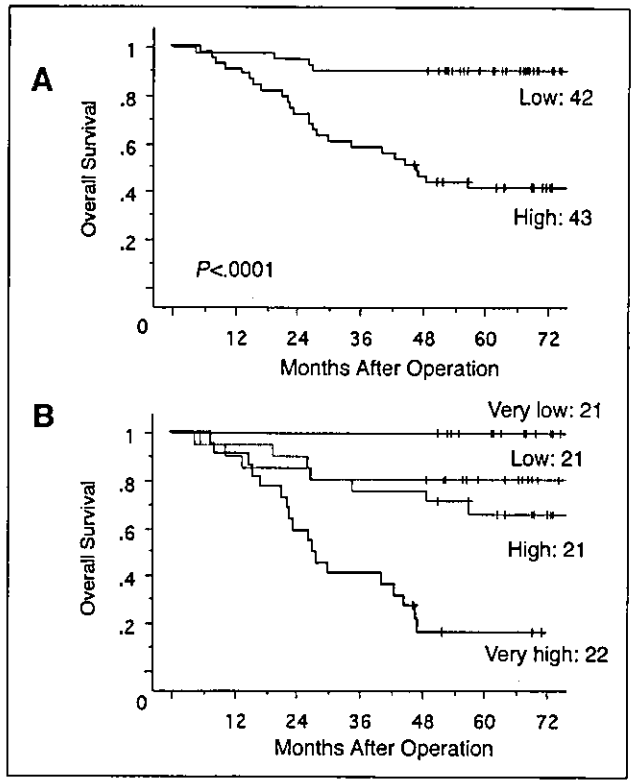


Fig 3. Relationship between survival in the 85 cases and their risk assignments based on the eight-gene risk index. The cutoff of the risk index was set at the 50th percentile (A) or the 25th, 50th, and 75th percentiles (B).

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

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
SCNN1A	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.

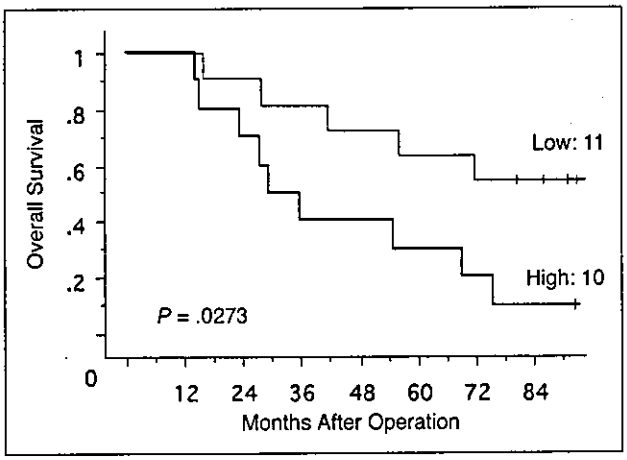


Fig 4. Relationship between survival in the 21 test cases and their risk assignments based on the eight-gene risk index.

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.

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.

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.

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 tracheoesophageal 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

Journal of Pathology
J Pathol 2004; **203**: 645–652
Published online in Wiley InterScience (www.interscience.wiley.com). **DOI:** 10.1002/path.1566

Original Paper

CK20 expression, CDX2 expression, *K-ras* mutation, and goblet cell morphology in a subset of lung adenocarcinomas

Yasushi Yatabe,[1]* Takaomi Koga,[1] Tetsuya Mitsudomi[2] and Takashi Takahashi[3]
[1]*Department of Pathology and Molecular Diagnostics, Aichi Cancer Center Hospital, Nagoya, Japan*
[2]*Department of Thoracic Surgery, Aichi Cancer Center Hospital, Nagoya, Japan*
[3]*Division of Molecular Oncology, Aichi Cancer Center Research Institute, Nagoya, Japan*

*Correspondence to: Yasushi Yatabe, MD, Department of Pathology and Molecular Diagnostics, Aichi Cancer Center, Kanokoden, Chikusa-ku, Nagoya 464-8681, Japan.
E-mail: yyatabe@aichi-cc.jp

Received: 18 December 2003
Revised: 9 February 2004
Accepted: 27 February 2004

Abstract

There are data in the literature that suggest a close relationship between the expression of CK20 and CDX2, *K-ras* mutations, and goblet cell morphology. The present study has examined these factors in a cohort of 264 non-small cell lung cancers. Thirteen of 212 adenocarcinomas expressed CK20; 29 expressed CDX2; *K-ras* mutation was identified in 28; and goblet cell features were present in 19. These four factors correlated with each other in a complex way and therefore a logistic regression model was constructed. Significant correlations were found between CK20 and CDX2 expression, and between *K-ras* mutation and goblet cell morphology, and there was a marginal correlation between CDX2 immunoreactivity and goblet cell morphology. These four features have also been commonly detected in colorectal, pancreato-biliary, and ovarian mucinous carcinomas, suggesting that these adenocarcinomas may be prototypical, independent of the organ of origin. Furthermore, as high and uniform expression of CDX2 was characteristic of metastatic colorectal cancer, weak and/or focal CDX2 expression should alert surgical pathologists to the possibility of primary lung adenocarcinoma, especially in the presence of goblet cell morphology. However, some lung adenocarcinomas may express CDX2 strongly: in this case, CK20 also tends to be positive.

Keywords: CK20; CDX2; *K-ras*; goblet cells; mucinous carcinoma; non-small cell carcinoma

Introduction

Lung cancers, especially adenocarcinomas, are characterized by a high degree of morphological heterogeneity, which implies both intra- and inter-tumoural diversity. Morphologically, lung adenocarcinomas show a wide range of cellular features and are subdivided into a number of categories including type II pneumocyte type, Clara cell type, bronchial surface epithelium type, and bronchial gland type, according to the scheme of cellular classification of lung adenocarcinomas proposed by Shimosato [1]. Goblet cell adenocarcinoma is one such subtype, about which interesting data have been reported. Several types of adenocarcinoma in the WHO classification, including mucinous type of bronchioloalveolar carcinoma, mucinous (colloid) carcinoma, mucinous cystadenocarcinoma, and adenocarcinoma, mixed mucinous type, belong to this category.

Close correlations have been suggested between the occurrence of mucinous bronchioloalveolar carcinomas (BACs) and mutations in the *K-ras* proto-oncogene [2,3]. Tsuchiya *et al* reported that goblet cell adenocarcinomas exhibited a very high rate of *K-ras* mutation in comparison with other types of adenocarcinoma [3]. Marchetti *et al* found *K-ras* mutations in ten out of ten mucinous BACs, but only in 34 of 98 non-mucinous adenocarcinomas [4].

Another interesting finding in mucinous carcinomas is differential expression of cytokeratins. Clinically, the expression of cytokeratin 20 (CK20), as well as CK7 and thyroid transcription factor-1 (TTF-1), has been used to discern primary from metastatic adenocarcinomas in the lung [5–10]. CK20 expression is restricted to a limited group of adenocarcinomas: it is detectable in almost all colon cancers, about half of gastric and pancreato-biliary carcinomas, and about 30% of transitional cell carcinomas. CK20 expression is rare in lung cancer, with less than 8% of lung adenocarcinomas positive [5,9,10]. However, some have reported that the expression in lung tumours is linked to a particular histological subtype. Shah *et al* [11] reported that 17 of 19 mucinous BACs showed immunoreactivity for CK20, whereas only ten of 80 adenocarcinomas NOS and four of 14 non-mucinous BACs were positive. Goldstein and Thomas

[12] also reported differential immunoreactivity for CK20 between mucinous and non-mucinous BACs; seven of 14 mucinous BACs were positive for CK20, in contrast to only one of 26 non-mucinous BACs.

CDX2 is a homeobox gene related to the *Drosophila Caudal* gene and encodes a transcription factor that plays an important role in pattern formation in the developing embryo and induction of intestine-specific genes [13]. Transfection of *CDX2* into an undifferentiated intestinal cell line resulted in arrest of proliferation, followed by morphological and phenotypic differentiation to mature intestinal epithelial-like cells, suggesting that the expression of this gene affects the proliferation and differentiation of intestinal cells [14]. Indeed, *CDX2* (+/−) heterozygous mice develop colonic hamartomas and polyps, whereas *CDX2* (−/−) homozygous mice are embryonic lethal. In adult human tissue, CDX2 is specifically expressed in the intestinal epithelium and most colorectal cancers are positive for CDX2. In addition, 50% of gastric cancers express CDX2 in association with their intestinal phenotype and positivity for CDX2 was observed in some pancreato-biliary carcinomas [15,16]. Whereas expression in organs derived from the mesodermal component of the digestive tract during development is understandable, ovarian mucinous carcinomas, which are positive for both *K-ras* mutations and CK20 expression, are also positive for CDX2.

These previous studies reported various pieces of information regarding the relationships between *K-ras* mutation, CK20 expression, goblet cell morphology, and CDX2 expression in lung cancer and in carcinomas in other sites. In this study, using a consecutive series of 264 non-small cell lung cancers, we attempted to organize and integrate the data on lung cancers and compared these with carcinomas at other sites. The results suggest the existence of an adenocarcinoma prototype, in which a subset of lung cancers may be included.

Materials and methods

Patients

A series of 264 consecutive, non-small cell carcinomas presenting between September 2000 and December 2002 at the Department of Pathology and Molecular Diagnostics, Aichi Cancer Center, Nagoya, Japan were used for the present study, after obtaining the approval of the institutional review board and patients' written informed consent. To exclude the possibility that the lung tumours were metastatic, all of the patients who underwent surgical resection were routinely examined by abdominal computed tomography and further analysis was performed when abnormal findings were obtained. Twenty-one metastatic colorectal carcinomas were used as controls. In addition, a spectrum of 32 neuroendocrine tumours of the lung was examined, comprising six typical carcinoid tumours, two atypical carcinoid tumours, eight large cell neuroendocrine carcinomas, and 16 small cell carcinomas. Clinical information was collected from the database and pathological staging was determined according to the *AJCC Cancer Staging Manual* [17].

Tissue microarrays

Four spots were selected per tumour and tissue microarrays were constructed using an MTA-1 manual tissue arrayer (Beecher Instruments, Inc, Silver Spring, MD, USA). Briefly, selected spots of the donor paraffin wax block were punched with a 0.6-mm-diameter coring needle, and transferred and arrayed in the recipient block using the arrayer. Serial 4-μm-thick sections on coated slide glasses were prepared for immunohistochemical analysis, as described previously [18,19].

Immunohistochemistry

Immunohistochemical examination proceeded according to the standard avidin–biotin–peroxidase complex method using monoclonal antibodies against CDX2 (CDX2-88; Biogenex, San Ramon, CA, USA), CK20 (Ks 20.8; DAKO, Copenhagen, Denmark), and TTF-1 (8G7G3, DAKO). Antigens were retrieved by autoclave for CDX2 and TTF-1, or by trypsin treatment for CK20. The tissue array was used for screening and all of the positive cases or cases with suspicious positive reactions were examined using regular large sections. Absence of false-negative cases in the tissue microarray was confirmed similarly with regular large sections, using 20 each of the cases evaluated as negative. Selected large sections were evaluated semiquantitatively with the following criteria. More than a moderate intensity of signal was considered as positive and the proportion of positive cells was scored as 0, negative; 1, less than 25% of positive tumour cells; 2, 26–50%; 3, 51–75%; and 4, 76% or more.

Relative quantification by real-time RT-PCR

First-strand cDNAs were synthesized from DNase I-treated total RNA extracts, using Superscript II (Invitrogen, Carlsbad, CA, USA) and random hexamer primers (Roche Applied Science, Alameda, CA, USA). Real-time quantitative PCR amplifications were performed using a Smart Cycler system (SC-100, Cepheid, Sunnyvale, CA, USA). The reactions were performed using QuantiTect SYBR Green PCR kits (Qiagen, Valencia, CA, USA). The primer sequences for CDX2 were forward, 5′-CCGAACAGGGACTTGTTTAGAG- 3′ and reverse, 5′-CTCTGGCTTGGATGTTACACAG-3′. In each reaction, standard samples diluted up to 1/1000 of cDNA from a colon cancer cell line, SW480, were run with unknown tumour samples. Quantitative values of each sample relative to HT-29, which is known to have minimal expression, were compared with those of β-actin.

Mutation status of K-ras

Frozen tissues from the tumour specimens were dissected macroscopically to enrich tumour cells in the extracted tissues, followed by extraction of total RNA with the RNAeasy kit (Qiagen). Using a standard RT-PCR procedure, *K-ras* was amplified using the following primers: forward, 5′-GGCCTGCTGAAAATGACTGA- 3′ and 5′-TCTTGCTAAGTCCTGAGCCTGTT-3′. The products were directly sequenced using an ABI PRISM 310 Genetic Analyzer (Applied Biosystems, Foster City, CA, USA).

Statistical analysis

The χ^2 test for independence and the unpaired t-test were used to compare the gene expression data. Logistic regression models were constructed to analyse complex relationships using SYSTAT software (SYSTAT Software Inc, Richmond, CA, USA). $p < 0.05$ was considered statistically significant.

Results

K-ras mutation and expression of CDX2 and CK20 in non-small cell lung cancer

Careful screening using tissue microarray and validation with regular whole sections revealed that 13 and 32 of 264 non-small cell lung cancers were positive for CK20 and CDX2 expression, respectively. The distribution of expression among the histological subtypes is listed in Table 1. All of the tumours expressing CK20 and CDX2 were adenocarcinomas except for three CDX2-positive squamous carcinomas. The intensity of the positive reactions in the primary lung adenocarcinomas was variable, whereas 21 metastatic cancers of colorectal origin were uniformly and intensely positive. The distribution of reactions and representative pictures are shown in Figures 1 and 2. The mean positive score for lung adenocarcinomas was significantly lower than that of metastatic colorectal cancers, and most of the reactions were regional. However, in a limited number of the lung adenocarcinomas, uniform positive reactions appeared to mimic those of metastatic colon cancers. Examination of clinical charts revealed no history of advanced gastrointestinal carcinoma prior to undergoing surgery for lung cancer in the patients with tumours positive for these markers.

K-ras mutations were observed in 30 of 251 tumours. The majority of the mutations were at codon 12 (26 of 30) and the others included one at codon 13 and three at codon 61. The significant association of *K-ras* mutations with adenocarcinomas (28 of 30 mutations found) is consistent with previous reports [20,21].

Table 1. Expression status in non-small cell lung cancers and metastatic colon cancers

	CK20 expression	CDX2 expression	*K-ras* mutations
Adenocarcinomas	13/212	29/212	28/203
Squamous cell carcinomas	0/41	3/41	1/39
Adenosquamous carcinomas	0/5	0/5	1/5
Large cell carcinomas	0/6	0/6	0/4
Metastatic colon cancers	20/21	21/21	8/19

Correlation of the expression of CK20, CDX2, and mutated *K-ras* with goblet cell morphology

Because *K-ras* mutations and CK20 expression have been reported in association with goblet cell morphology, we examined the relationship between these features in lung adenocarcinomas. Goblet cell adenocarcinomas were observed in 19 of the 212 tumours. This was associated with a higher frequency of *K-ras* mutations ($p < 0.05$) and CDX2 expression ($p < 0.05$). Similarly, a statistically significant relationship was found between CDX2 and CK20 expression ($p < 0.01$), and *K-ras* tended to be mutated more frequently in CDX2-expressing adenocarcinomas ($p = 0.06$). The clinicopathological features of the four groups are summarized in Table 2. The relationships were associated with each other and we therefore constructed logistic regression models to analyse these complex relationships, together with various clinicopathological features (Table 3). CK20 expression was associated only with CDX2 expression and *K-ras* mutations only with goblet cell morphology, respectively. CDX2 expression was significantly associated with CK20 expression, male sex, a tumour diameter of more than 30 mm, and lack of TTF-1 expression; there was a marginal association with goblet cell morphology. Goblet cell morphology was related significantly to lack of TTF-1 expression and marginally with *K-ras* mutation. These results suggested three kinds of relationship: first, a close association between CK20 and CDX2; second, correlations between *K-ras* mutations and goblet cell morphology; and third, a weak association between CDX2 expression and goblet cell morphology.

These findings raised the question of whether a small proportion of adenocarcinomas were individually positive for all of these markers, or whether a subset of adenocarcinomas specifically expressed some of the markers. The former possibility suggests the presence of a distinct form of adenocarcinoma, whereas in the latter case, expression represents a vague tendency for a certain differentiation feature. The results revealed that tumours expressing just one or two markers were most frequent and that only two expressed all markers.

CDX2 mRNA expression in lung cancers

In contrast to previous studies, we observed CDX2 expression in lung tumours. To confirm this,

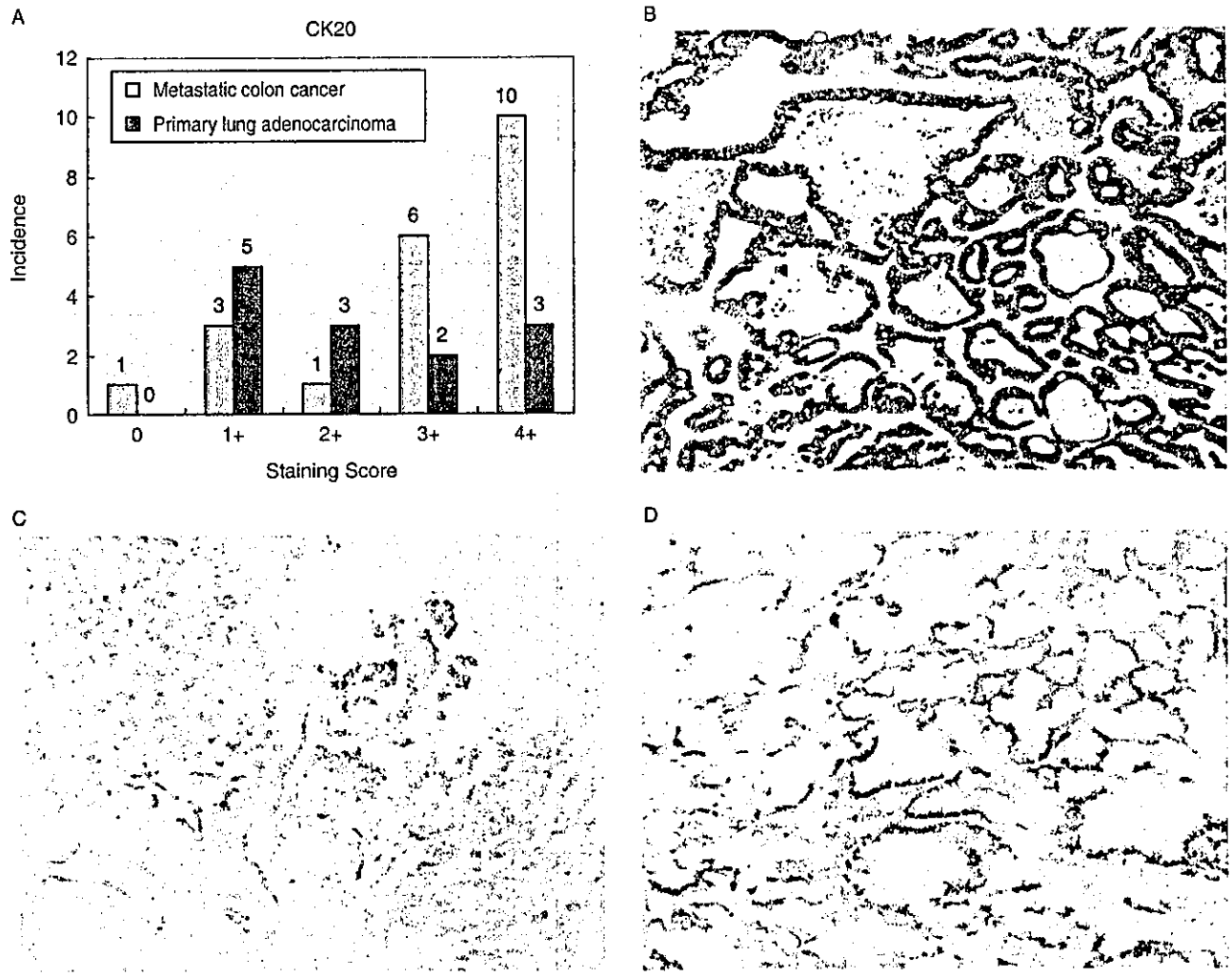


Figure 1. CK20 expression in metastatic colon cancers and lung adenocarcinomas. The histogram (A) indicates the distribution of CK20-postive scores between the metastatic colon cancers and lung adenocarcinomas examined: mean scores are 3.0 and 2.2, respectively. The numbers on the tops of the bars indicate the frequency of each staining category. Representative sections are shown for a score of 4+ in a metastatic colon cancer (B), a score of 2+ in a lung adenocarcinoma (C), and a score of 4+ in a lung adenocarcinoma (D)

expression relative to that of a colon cancer cell line, HT29, was examined using real-time RT-PCR in three normal lungs, three metastatic colon cancers, five CDX2-positive lung adenocarcinomas, and five CDX2-negative lung adenocarcinomas (Figure 3). No or negligible transcript levels were detected in normal lung and CDX2-negative lung adenocarcinomas, whereas CDX2-positive lung adenocarcinomas and colorectal cancers showed, respectively, 10.7 and 36.6 times as much transcript relative to HT29 cells. In the lung adenocarcinomas with positive CDX2 expression, relative transcript values varied from 0.1 to 49.3, which correlated well with the staining scores.

CDX2 expression in neuroendocrine carcinomas

We noted that, in addition to lung adenocarcinomas, some of the neuroendocrine tumours expressed CDX2. This was observed in six of 24 high-grade neuroendocrine tumours (25%), including large cell neuroendocrine carcinomas and small cell lung cancers (Figure 4). Although the expression in about half of the positive cases was restricted to a portion of the tumour (less than 25% of the tumour area), the other half demonstrated relatively uniform expression. The distribution of the staining scores and representative scoring 3+ is shown in Figure 4.

Discussion

To date, pieces of information in the literature have indicated a close relationship between the expression of CK20 and CDX2, *K-ras* mutations, and goblet cell morphology. CK20 is expressed commonly in colorectal adenocarcinomas, ovarian mucinous carcinomas, transitional cell carcinomas, gastric adenocarcinomas, and pancreatic carcinomas [5,9,10]. CDX2 is expressed in colorectal adenocarcinomas, ovarian mucinous carcinomas, gastric adenocarcinomas, pancreatic ductal carcinomas, and biliary duct carcinomas [15,16,22]. *K-ras* is frequently mutated in pancreatic carcinomas, colorectal carcinomas, ovarian mucinous

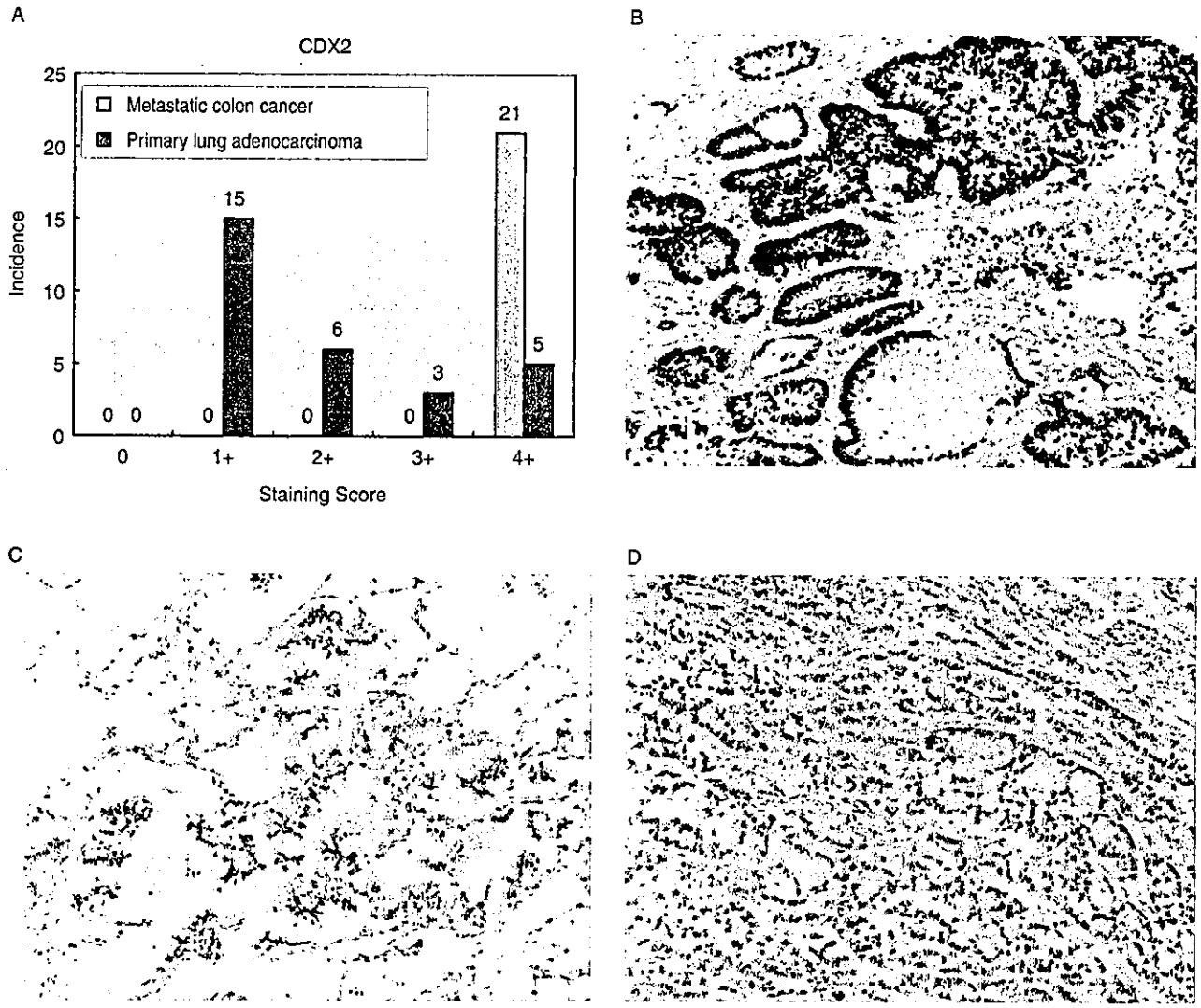


Figure 2. CDX2 expression in metastatic colon cancers and lung adenocarcinomas. The histogram (A) indicates the distribution of CDX2-postive scores between metastatic colon cancers and lung adenocarcinomas: mean scores are 4.0 and 1.9, respectively. The numbers on the tops of the bars indicate the frequency of each staining category. Representative figures are shown for a score of 4+ in a metastatic colon cancer (B), a score of 2+ in a lung adenocarcinoma (C), and a score of 4+ in a lung adenocarcinoma (D)

carcinomas, and endometrial carcinomas. Therefore, colorectal carcinomas, ovarian mucinous carcinomas, pancreatic carcinomas, and some endometrial carcinomas share the four properties examined in the present study and these tumours make up a significant proportion of human adenocarcinomas. These findings suggested that overlap of these four features is rather common and a cancer with these features may represent one of the adenocarcinoma prototypes, which is independent of the organ of origin. Characteristics of this prototype may be cell type of 'intestinal' goblet cells. Ichikawa *et al* reported that intestinal-type mucinous ovarian, but not endocervical, tumours exhibit *K-ras* mutation, suggesting a specific association of this mutation with 'intestinal' goblet cell-type adenocarcinomas [23]. Indeed, although endocervical adenocarcinomas show goblet cell morphology, *K-ras* mutation and expression of CK20 and CDX2 are rarely detected. This kind of alteration that is specific for cellular context and/or cell type has been documented. Some skin tumours show a preference for certain genes to be mutated, ie PTCH in basal cell carcinomas [24] and CYLD in cylindroma [25]. It is of note that despite induction throughout the epidermis and skin appendages using a keratin-5 promoter, mice transgenic for *H-ras, GLI2*, and *β-catenin* developed different types of tumour: squamous cell carcinoma [26], basal cell carcinoma [27], and hair-follicle tumour [28], respectively. Another example is mice transgenic for *K-ras*. Even though the mice expressed $K\text{-}ras^{V12}$ throughout the body, hyperproliferative properties and consequent progression to cancer were induced only in the bronchioloalveolar cells in the lung. These findings suggest that the molecular mechanism of carcinogenesis is different among cell types and/or cellular contexts, and similar cell types and/or contexts may share mechanisms that result in a similar genotype and phenotype.

To our knowledge, there are only three reports concerning CDX2 expression in the differential diagnosis

Table 2. Patient characteristics

	CDX2 expression	CK20 expression	*K-ras* mutation	Goblet cell morphology
Number of cases	29	13	28	19
Histological subtypes in the WHO classification				
Mixed	16	6	15	16
Acinar	8	4	5	3
Papillary	4	3	6	0
Solid with mucin	1	0	2	0
TNM stage				
pT(1/2/3/4)	9/12/6/2	9/4/0/0	11/13/2/2	7/7/0/5
pN(0/1/2/3)	20/2/7/0	10/0/2/1	18/0/9/1	16/1/2/0
pM(0/1)	28/1	13/0	28/0	19/0
Male gender	22	7	21	5
More than median age	14	8	10	11
Smoker	21	8	21	7
More than 30 mm of tumour size	19	3	12	7
Invasion beyond visceral pleura	13	2	7	4
Nodal metastasis	9	2	9	3
TTF-1 expression	11	8	17*	7

Each number indicates the frequency in each category.
* Two cases were not examined.

Table 3. Four logistic regression models for the analysis of relationships

	Dependent variables			
Independent variables	**CK20 expression**	**CDX2 expression**	***K-ras* mutations**	**Goblet cell morphology**
CK20 expression	—	<u>< 0.01</u>	0.11	0.16
CDX2 expression	<u>< 0.01</u>	—	0.43	0.25
K-ras mutations	0.30	0.49	—	*0.06*
Goblet cell morphology	0.43	*0.05*	<u>0.02</u>	—
More than median age	0.80	0.73	0.10	0.23
Male gender	0.20	<u>< 0.01</u>	0.20	0.12
Smoker	0.49	0.52	0.16	0.38
Tumour diameter > 30 mm	0.40	<u>0.03</u>	0.74	0.80
Invasion to visceral pleura	0.22	0.27	0.52	0.19
Pathological stage I (early stage)	0.73	0.86	0.67	0.33
Nodal metastasis	0.78	0.40	0.39	0.10
TTF-1 expression	0.53	<u>0.02</u>	0.85	<u>< 0.01</u>

Each value indicates a p value for the comparison of independent variables (rows) with individual models (columns). For example, the CK20 column indicates that although CD20 expression is potentially influenced by various factors listed in rows (independent variables), the logistic regression model for CK20 expression (second column) reveals a statistically significant association with CDX2 expression ($p = 0.01$) but not the other factors.
Underlines and italics represent statistically significant and marginal correlations, respectively.

between primary lung adenocarcinoma and metastatic cancers of colorectal origin [15,16,22]. The present study is in agreement with those in terms of the marked difference in CDX2 expression between primary and metastatic adenocarcinomas; high and uniform expression was characteristic of colorectal cancers. However, as shown with real-time PCR, expression is not completely absent from primary lung adenocarcinomas. Indeed, one of the previous three reports documented one case each of 1+ and 2+ CDX2 staining among 33 primary lung adenocarcinomas examined [15]. Furthermore, other non-gastrointestinal tumours were also positive for CDX2. Franchi *et al* demonstrated uniform expression in most ethmoid sinus adenocarcinomas, which were also positive for CK20 and goblet cell morphology [29]. It is therefore reasonable to suggest that CDX2 can be expressed in this specific subset of lung adenocarcinomas, as a reflection of the characteristics of the cell- or cancer-type. This also alerts surgical pathologists to a pitfall in differential diagnosis using CDX2 immunohistochemistry.

It is of note that high-grade neuroendocrine carcinomas also expressed CDX2. It is well known that *K-ras* is rarely mutated in small cell lung cancers [21] and that high-grade neuroendocrine tumours are negative for CK20 [9]. Thus, CDX2 is not expressed in association with a certain adenocarcinoma prototype. Rather, this finding may represent aberrant expression of lineage-specific molecules in pulmonary high-grade neuroendocrine carcinomas. We have reported