Prognosis of Non–Small Cell Lung Cancer Patients by Detecting Circulating Cancer Cells in the Peripheral Blood with Multiple Marker Genes

Yuh-Pyng Sher,[1,3] Jin-Yuan Shih,[2] Pan-Chyr Yang,[1,2,4] Steve R. Roffler,[1] Yi-Wen Chu,[4] Cheng-Wen Wu,[4] Chia-Li Yu,[3] and Konan Peck[1]

[1]Institute of Biomedical Sciences, Academia Sinica; [2]Department of Internal Medicine, National Taiwan University Hospital and [3]Institute of Molecular Medicine, National Taiwan University; and [4]National Health Research Institutes, Taipei, Taiwan, Republic of China

ABSTRACT

Purpose: **Current lung cancer staging and prognosis methods are based on imaging methods, which may not be sensitive enough for early and accurate detection of metastasis. This study aims to validate the use of a panel of markers for circulating cancer cell detection to improve the accuracy of cancer staging, prognosis, and as a rapid assessment of therapeutic response.**

Experimental Design: **We analyzed the National Cancer Institute-Cancer Genome Anatomy Project database to identify potential marker genes for the detection of circulating cancer cells in peripheral blood. Nested real-time quantitative PCR and a scoring method using cancer cell load L_c were employed to correlate the amount of circulating cancer cells with clinical outcomes in 54 non–small cell lung cancer (NSCLC) patients. The Kaplan-Meier method was employed for analysis of prognostic variables.**

Results: **A panel of four marker genes was identified and experimentally validated. With these marker genes, we achieved an overall positive detection rate of 72% for circulating cancer cells in the peripheral blood of NSCLC patients. Patients who had higher L_c values had worse outcomes and shorter survival times. Patients with poor therapeutic response were revealed by positive detection of circulating cancer cells after therapy. The results correlated well with the patients' survival time.**

Conclusion: **Circulating cancer cell detection by a panel of markers and the L_c scoring method can supplement the current tumor, node, metastasis staging method for improved prognosis and for rapid assessment of therapeutic response. Together, they may facilitate the design of better therapeutic strategies for the treatment of NSCLC patients.**

INTRODUCTION

Lung cancer is the leading cause of cancer-related death and non–small cell lung cancer (NSCLC) accounts for ~80% of the cases. Attempts to use serum protein markers for the early diagnosis of lung cancer have not yielded satisfactory results for routine screening, and newly developed early diagnostic methods using serum DNA as a diagnostic marker await further validation (1, 2). Current therapeutic measures remain unable to lower the mortality rate of late-stage lung cancer patients. Surgical resection is still the best cure for the early-stage patients. The tumor, node, metastasis (TNM) classification has been used for cancer staging and prognosis for decades. A large portion of early-stage patients, defined by the current staging system and available imaging modalities, still develop distant metastases although they received surgical removal of the tumor mass. The inability to detect disseminated tumor cells with the current imaging techniques is a major obstacle to accurate cancer staging.

Several research groups have reported the presence of epithelial cancer cells in the bone marrow (3, 4) and in the peripheral blood (5, 6) of patients with carcinoma. In contrast to bone marrow aspirates, peripheral blood samples can be obtained routinely and more readily. Techniques such as immunocytology, flow cytometry (7), and PCR (8–11) have been employed to detect disseminated tumor cells in peripheral blood. The positive detection rate of circulating cancer cells in lung cancer patients was only about 40% when a single marker gene (cytokeratin 19, *CK19*) was employed (9). To overcome this problem, a multimarker reverse transcription-PCR assay should be beneficial as shown in the diagnosis of melanoma and breast cancer (12, 13).

In this report, we identify a panel of markers for the detection of circulating cancer cells in NSCLC patients by *in silico* analysis of the National Cancer Institute-Cancer Genome Anatomy Project database (http://cgap.nci.nih.gov/). The load of cancer cells in the circulation was correlated with the clinical outcomes of the patients. An accurate cancer staging method is prerequisite for prescribing optimal therapeutic courses for early- and late-stage lung cancer patients. Detection of the presence of circulating cancer cells in conjunction with the conventional TNM classification may better define the stage and prognosis of cancer patients. This method provides a useful tool to optimize NSCLC treatment and management.

MATERIALS AND METHODS

Patients and Specimens. Peripheral blood samples were obtained with informed consent from 54 patients with histologically documented NSCLC in the National Taiwan University Hospital. Staging procedures included chest radiography, bronchoscopy, brain and thoracic computed tomography, sonography, and bone scintigraphy. The follow-up period of the lung cancer

Received 7/2/04; revised 8/9/04; accepted 9/7/04.
Grant support: National Health Research Institutes, Taipei, Taiwan, Republic of China Metastasis Program Project grant NHRI-CN-PL-9101P.
The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.
Requests for reprints: Konan Peck, Institute of Biomedical Sciences, Academia Sinica, 128, Section 2, Academia Road, Taipei, Taiwan 115, Republic of China. Phone: 886-2-2652-3072; Fax: 886-2-2785-8594; E-mail: konan@ibms.sinica.edu.tw.

patients was up to 85 months for stage I patients who were still alive in May 2003. The normal control population consisted of 24 healthy volunteers with no history or present diagnosis of malignancy. Among the 54 patients, 32 were men and 22 were women with median age of 65 years (SD = 11, range = 28-81). In the healthy control group, 16 were men and 8 were women with median age of 57.5 years (SD = 15.1, range = 27-78). Adenocarcinoma lung cancer cell lines (A549, CL3, H928, CL1-0, CL1-5, CRL-5865, CRL-5806, and CRL-5807) and squamous carcinoma lung cancer cell lines (NCI-H520, H2981, CRL-5802, and HTB-54) were used to validate the candidate markers and for analyzing the correlation of L_c value and cancer cell number.

Identification of Candidate Marker Genes. To take advantage of the vast information of the expressed sequence tags databases generated with cancer cell lines, we used the cDNA Digital Gene Expression Displayer developed by the Cancer Genome Anatomy Project (14) to identify genes that were differentially expressed between lung cancer cells and leukocytes. The Digital Gene Expression Displayer program identified differentially expressed genes among 47,036 sequences in five lung cancer cDNA libraries and 21,460 sequences in six leukocyte cDNA libraries with a P filter set at 0.01. The differentially expressed genes were ranked by sequence odds ratio. The genes with the highest sequence odds ratios were selected as candidate marker genes for quantitative PCR (qPCR) assay.

Sample Collection and RNA Preparation. The blood sample collection and RNA preparation methods were the same as described by Peck et al. (9). Briefly, two samples were collected from each subject with Vacutainers (Becton Dickinson, Rutherford, NJ). The first tube with 1 to 2 mL of peripheral blood was discarded and only the second tube with 3 to 4 mL of blood was assayed to avoid epithelial cell contamination by the needle when it pierced through the skin. Total cellular RNA was extracted with the QIAamp RNA Blood Mini kit (Qiagen, Hiden, Germany) within 2 hours after the blood samples were collected.

Nested Reverse Transcription-PCR Assay. To detect a few cancer cells in the circulation, a highly sensitive nested reverse transcription-PCR is necessary. cDNA was derived from 1 to 2 μg of total RNA by random primed reverse transcription, and nested PCR was used to amplify the candidate marker gene transcripts for detection. The primer sets are listed in Table 1. The first round (outer) of the nested PCR was done using 1 μL of 20-fold diluted cDNA with a PCR mixture containing 0.1 μmol/L outer primer pair, 0.2 mmol/L deoxynucleotide triphosphate, 50 mmol/L Tris-HCl (pH 8.3), 10 mmol/L KCl, 5 mmol/L $(NH_4)_2SO_4$, 2 mmol/L $MgCl_2$, and 0.75 units of FastStart Taq DNA polymerase (Roche, Mannheim, Germany) in a total volume of 12.5 μL. The PCR conditions were one cycle at 94°C for 7 minutes followed by 25 cycles at 94°C for 50 seconds, 60°C for 50 seconds, 72°C for 35 seconds, and a final extension at 72°C for 10 minutes.

Table 1 DNA sequences of the PCR primer pairs for detecting the marker genes

Marker gene		Primer sequence (5′ to 3′)	Amplicon size (bp)
A	Outer_F	aagctaaccatgcagaacctcaacgaccgc	1,070
	Outer_R	ttattggcaggtcaggagaagagcc	
	Inner_F	tcccgcgactacagccactactacacgacc	745
	Inner_R	cgcgacttgatgtccatgagccgctggtac	
B	Outer_F	caatcggacttattcacgca	557
	Outer_R	gtgcttaagtgcctgggtgt	
	Inner_F	tgtgcagagaattcaccgag	200
	Inner_R	gcgtgtctgcagaacagaag	
C	Outer_F	gcattggagatcagtgggat	535
	Outer_R	tatccaaggtttctgggtgg	
	Inner_F	atgatgaggtgcacgtgtgt	219
	Inner_R	ctctgaatcctggcattggt	
D	Outer_F	atgtgagcgtgtactgctgg	446
	Outer_R	ggacagagaacagagccagg	
	Inner_F	atgtgagcgtgtactgctgg	237
	Inner_R	acgtctgccttgtcctcagt	

For the second round (inner) of the nested PCR amplification, the reaction mixture contained 2 μL of the first round PCR product, 0.25 μmol/L inner primers, and SYBR Green PCR master mix (Applied Biosystems, Foster City, CA) in a total volume of 20 μL. The real-time qPCR assays were done with an ABI prism 7000 SDS (Applied Biosystems) instrument. The real-time qPCR condition was 95°C for 10 minutes followed by 40 cycles at 95°C for 15 seconds, 60°C for 25 seconds, and 72°C for 35 seconds.

Semiquantification of the Nested PCR Results. By using a real-time qPCR instrument, the threshold cycle (C_T), the fractional cycle number at which the SYBR Green I fluorescence exceeded a set level above baseline, was determined. We used GAPDH mRNA as an internal control. The relative amount of mRNA, normalized against the GAPDH mRNA, was expressed as $\Delta C_T = C_T^{(\text{GAPDH})} - C_T^{(\text{marker gene})}$. If the fluorescence signal was undetected after 40 cycles, the C_T value was given the maximum cycle number of 40 for analysis convenience. The differential expression ratio of a candidate marker gene, Q, for patients versus normal controls was calculated by $Q = 2^{\Delta C_T - \text{mean of } \Delta C_T \text{ in normal}}$. To estimate the number of circulating cancer cells, we normalized the differential expression ratio of each marker gene to take into account the different amount of the marker gene transcripts. We then summed up the expression ratios of the marker genes to estimate the cancer cell load in the circulation in a semiquantitative way. The formula for normalizing the expression ratio, E_{ij}, of a marker gene is $E_{ij} = \frac{Q_{ij} - \bar{Q}_j}{\sigma_j}$, where i is the patient index, j is the marker gene index, Q_{ij} is the differential expression ratio of marker gene j in patient i, $\bar{Q}_j$ is the mean and σ_j is the SD of the expression ratios of the 54 patients for marker gene j. The load of cancer cells in the circulation of a patient is defined as $L_c = \sum_{j=1}^{n} E_j$ where n is the number of marker genes. In this study, the value of L_c ranged between −2.2 and 8.4.

Statistical Analysis. Fisher's exact test and Student's t test were used to compare the clinicopathologic characteristics of patients with low and high L_c values. All statistical tests were two sided. Survival time of the patients was calculated from the day of specimen collection. Survival curves were obtained by the Kaplan-Meier method. The difference of survival times between two groups was analyzed with the log-rank test. $P < 0.05$ was considered statistically significant.

RESULTS

Marker Genes for Detecting Circulating NSCLC cells. The *in silico* Digital Gene Expression Displayer program search

of the National Cancer Institute-Cancer Genome Anatomy Project database yielded 85 overexpressed genes with a sequence odds ratio >16 between the lung cancer cDNA libraries and the leukocyte cDNA libraries. These candidate genes were further verified by nested real-time qPCR. All of the cancer cell lines listed in the MATERIALS AND METHODS and pooled peripheral blood mononuclear cells from 12 healthy controls were used as samples in the first round verification. Fifty-nine candidate marker genes showed >2-fold differential expression ratios, but only 19 genes had differential expression ratios >100,000. Marker genes with large differential expression ratios are required to detect rare circulating cancer cells in blood samples containing millions of peripheral blood mononuclear cells. By using real-time qPCR to detect the presence of 19 candidate genes in the clinical specimens of 54 NSCLC patients and 24 normal controls, four marker genes including *CK19* were identified to show positive detection in at least two NSCLC patients. The four marker genes are listed in Table 2. The first three marker genes had negligible expression in the blood samples of normal controls by the real-time nested qPCR assay and are labeled as genes A to C. The fourth gene, labeled as gene D, had high expression in the cancer patient group versus the normal control group, but had residual expression in normal peripheral blood mononuclear cells. Therefore, we set the positive detection threshold at 99% confidence interval in the normal control group to avoid false positive detection.

Enhancement of Positive Detection Rate with Multiple Genes. As shown in Fig. 1*A*, the positive detection rate of circulating cancer cells in NSCLC patients was 41%, 11%, 39%, and 11% for genes A to D, respectively. The fraction of positives among all the patients are indicated in the columns.

Increasing the number of marker genes raised the positive detection rate for NSCLC patients to 72% (39 of 54) compared with 41% (22 of 54) positive detection rate using the *CK19* marker gene alone (Fig. 1*B*). These results prove that using multiple markers indeed improves the positive detection rate. The B marker gene was not detected in squamous cell lung cancer patients, but the other three genes were detected in patients with adenocarcinoma or squamous carcinoma. Further analysis of these marker genes did not show statistically significant results for lung cancer subtyping.

Circulating Cancer Cell Load and Patient Outcome. To investigate the correlation between the number of circulating cancer cells and patient outcome, we used cancer cell load L_c (see MATERIALS AND METHODS) as a measure of the number of cancer cells in circulation. By definition, L_c increases with the number of circulating cells and/or the number of detected marker genes in the circulating cancer cells. We spiked different numbers of lung cancer cells (CL1-0) into peripheral blood mononuclear cells to determine the correlation between the number of circulating cancer cells and L_c values (Fig. 2*A*). The results show a good correlation ($r = 0.99$) between cancer cell number and L_c value.

To determine whether L_c can be used as an indicator of patient outcome, we analyzed the L_c value versus the survival time of stage I to IIIa patients who received surgical resection. We found that patients who survived for more than 5 years had low L_c values (<1; data not shown). Therefore, we set $L_c = 1$ as the threshold for the subsequent prognosis studies in the 54 lung cancer patients. We divided the patients at different stages into two groups ($L_c \geq 1$ or $L_c < 1$) and examined their survival time. As expected, late-stage patients had shorter survival times than did early-stage patients (Fig. 2*B*, *inset*). The cancer load (L_c) study further distinguished that for patients of the same stage, those with low L_c had longer survival time than those with high L_c value (Fig. 2*B*). For the operable early-stage patients, the survival time is highly correlated ($P = 0.002$) with L_c value. We then used the L_c value to analyze the prognosis of late-stage patients (stages IIIb and IV) with Kaplan-Meier survival plots. The survival time difference was more significant by using L_c as a classification parameter (Fig. 3*B*, $P = 0.006$) than by using the traditional TNM classification (Fig. 3*A*, $P = 0.03$). The number of patients in each group is shown in parentheses in the figures. We further used the L_c parameter to divide the patients of the same stage into low ($L_c < 1$) and high ($L_c \geq 1$) value groups and examined the survival plots. The survival time of stage IIIb patients with low and high L_c value was significantly different

Table 2 List of a panel of four marker genes

Gene	Title	Genbank accession no.
A	*Homo sapiens* keratin 19 (*KRT19*)	NM_002276
B	Ubiquitin thiolesterase	NM_004181
C	Highly similar to HSFIB1 for fibronectin	NM_054034
D	Tripartite motif-containing 28 (*TRIM28*)	NM_005762

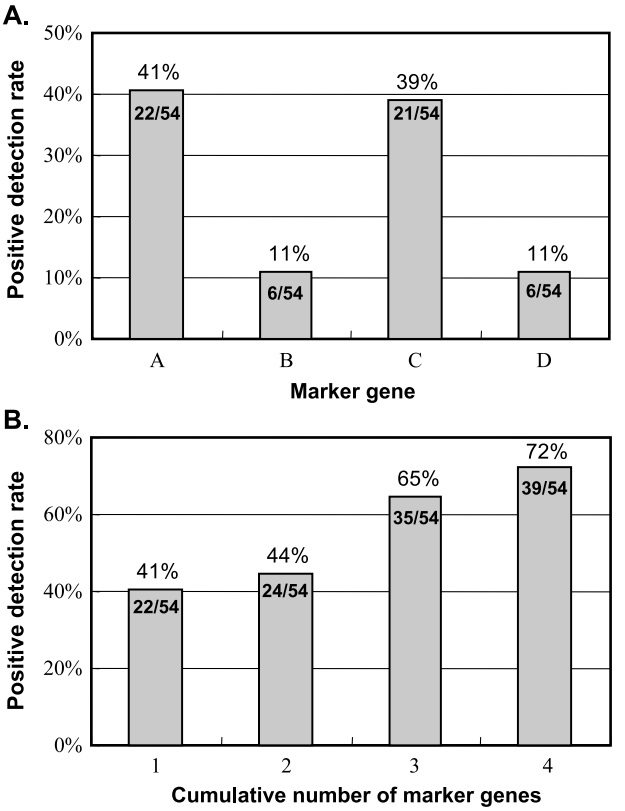

Fig. 1 Analysis of positive detection rates with the multimarker gene panel. *A,* positive detection rate for the individual four marker genes. *B,* positive detection rate increases with the number of marker genes used in the assay.

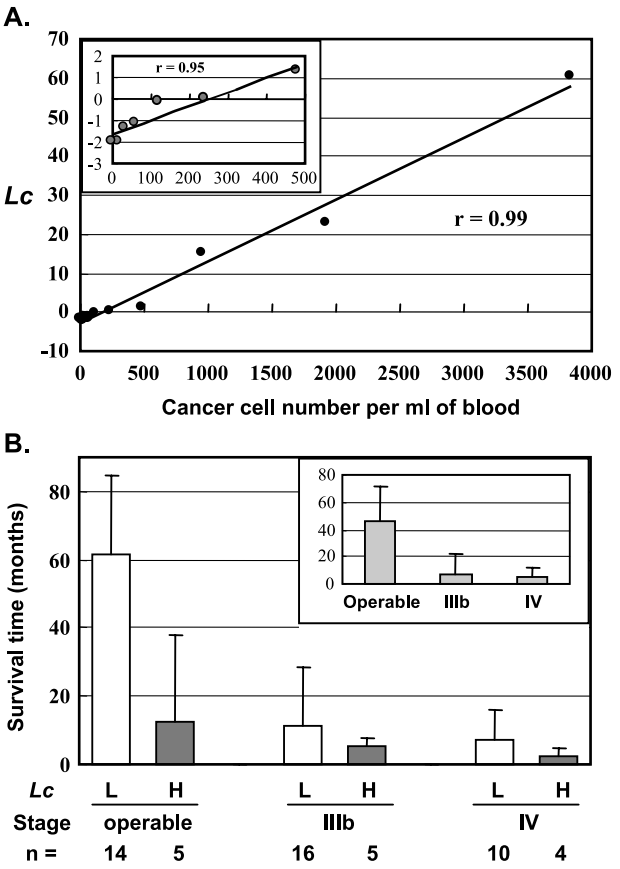

Fig. 2 L_c value characterization. *A,* plot to show that L_c value is linearly correlated with the number of circulating cancer cells. *Inset,* L_c value versus the number of cancer cells between 0 and 480 cells/mL of blood. *B,* median survival time for patients of various TNM stages and with high (≥1) or low (<1) L_c values. *Inset,* survival time of patients classified by the TNM staging method. *Bars,* SD.

(Fig. 3*C*, $P = 0.026$). The same analysis on stage IV patients did not achieve statistical significance (Fig. 3*D*, $P = 0.09$).

No statistically significant association was found between L_c values and clinicopathologic variables such as age, smoking, histologic type, and stage, with the exception of gender (Table 3). A search of the available literature reports on the four marker genes did not offer a plausible explanation for this gender bias. An investigation on this finding will be conducted in a separate study.

Assessment of Therapy Efficacy. We further examined whether cancer cell load could be employed to monitor the efficacy of therapy. Fig. 4 shows the assessment of six patients before and after chemotherapy. The stage I, II, and IIIa patients received surgical resection, whereas the stage IIIb and IV patients received chemotherapy. "d−" indicates that the samples were collected before therapy, "d*n*" indicates that the samples were collected on the *n*th day after therapy, and "*n*/" indicates the course number of chemotherapy. The stage I patient had negative L_c values before and after treatment and was still alive on the last follow-up date (72.3 months). The L_c values decreased for the stage II and IIIb patients after therapy. The stage IIIa and IV patients had higher L_c values after the last course of treatment and had short survival times. These results suggest that measurement of the cancer cell load can be used to reveal whether therapy is efficacious.

DISCUSSION

In this study, we employed *in silico* analysis to identify a panel of marker genes for NSCLC prognosis and assessment of therapy efficacy. Based on our experimental results, the National Cancer Institute-Cancer Genome Anatomy Project database and the Digital Gene Expression Displayer program are useful tools for identifying differentially expressed genes between two pools of samples provided that a sufficient number of expressed sequence tags libraries for the tissue of interest are archived in the database. The differentially expressed genes can be further developed into marker genes for diagnostic or prognostic purposes by experimental verification procedures such as real-time qPCR.

NSCLC is heterogeneous with respect to histology and biological characteristics (15). Individual NSCLC cells within a tumor and in different patients' tumors express different amounts of marker gene transcripts. The heterogeneity of marker gene expression levels in NSCLC cells limits the reliability of an assay method with a single-marker detection scheme. Several literature reports have described PCR methods for the detection of tumor cells dispersed in the circulation (8–11). However, not one tumor marker is consistently and specifically expressed in all of the primary tumors of a particular malignancy (16, 17). Literature reports have also shown that a panel of marker genes provides a more reliable and informative approach than a single-marker assay for the detection of melanoma and breast cancer cells in blood (12, 13). Such assays for lung cancer have been limited by the availability of molecular markers (18). To overcome this obstacle, we developed an effective method to identify marker genes for NSCLC and showed the use of a multimarker reverse transcription-qPCR assay for detecting metastatic lung cancer cells in blood with a good detection rate.

Table 3 Clinicopathologic characteristics and their correlation with L_c value of NSCLC patients

Characteristic	L_c		*P*
	Low	High	
Age (y), mean ± SD	63.6 ± 10.5	61.3 ± 13.0	0.549*
Gender, no. patients			
Male	20	12	0.03
Female	20	2	
Smoking, no. patients			
No	24	7	0.546
Yes	16	7	
Histology, no. patients			
Adenocarcinoma	25	10	0.442
Squamous cell carcinoma	12	2	
Poorly differentiated	3	2	
Stage†			
I-IIIa	14	5	1.0
IIIb-IV	26	9	

*Derived with Student's *t* test; other *P*s were derived with Fisher's exact test. All statistical tests were two-sided.

†Tumor stage was classified according to the International System for Staging Lung Cancer.

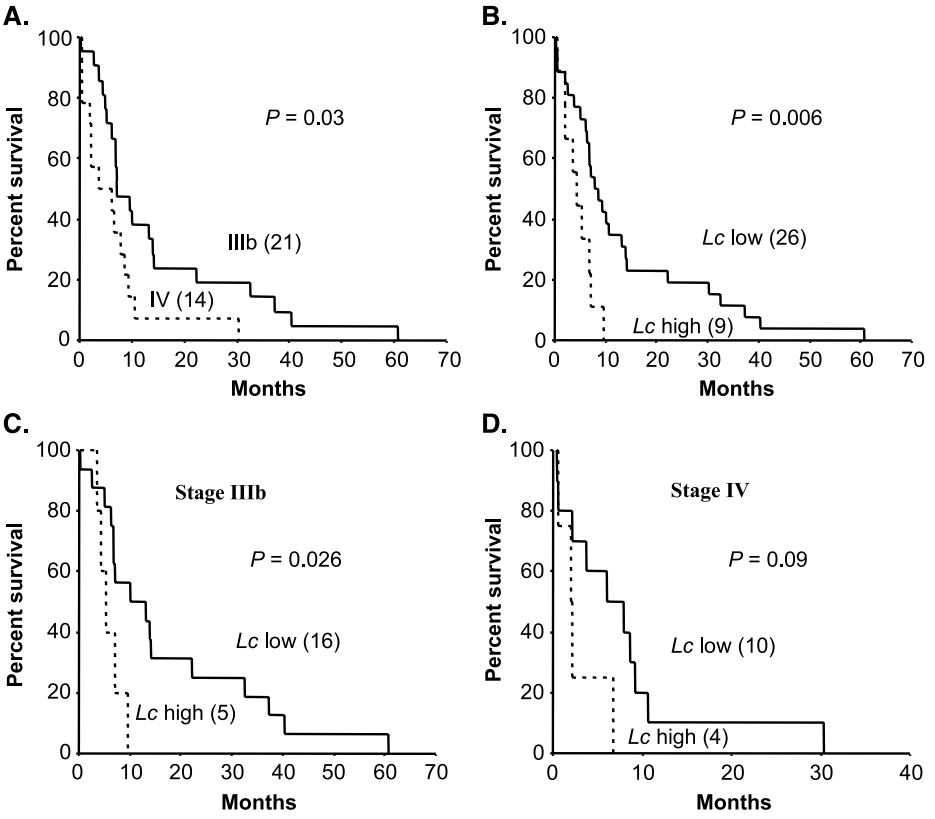

Fig. 3 Survival analysis of late stage patients with high (≥1) or low (<1) L_c values. *A,* Kaplan-Meier survival plots of late TNM stage patients (IIIb and IV). *B,* Kaplan-Meier survival plots for the same late stage patients grouped by the L_c value. *C,* Kaplan-Meier survival plots for the stage IIIb patients grouped by the L_c value. *D,* Kaplan-Meier survival plots for the stage IV patients grouped by the L_c value.

The positive detection rate was significantly improved by using a panel of four marker genes as compared with using a single *CK19* marker gene, which is a characteristic marker to epithelial cells. Other than *CK19*, the roles of the other three marker genes in cell differentiation, carcinogenesis, tumor invasion, and metastasis are not yet certain. Literature reports characterizing the marker genes and their association with cancers are available for ubiquitin thiolesterase and fibronectin but not TRIM28.

Ubiquitin thiolesterase, also called protein gene product 9.5 (PGP9.5), is a neurospecific peptide that functions to remove ubiquitin from ubiquitinated proteins and prevents them from targeted degradation in proteasomes. PGP9.5 can contribute to the uncontrolled growth of somatic cells (19). Recent studies have identified PGP9.5 as a marker for invasive colorectal cancer (20) and as a prognostic factor in pancreatic cancer (21) and NSCLC (22). In our study, the *PGP9.5* gene was preferentially expressed in stage IIIb and IV NSCLC (17%) compared with early-stage (I, II, and IIIa) patients. This result is similar to previous findings (22).

Fibronectin, an extracellular matrix protein, is produced by a wide variety of cells such as fibroblasts (23), endothelial cells (24), epithelial cells (25), macrophages (26), and some carcinoma cells (27, 28). Fibronectin plays several important roles in cell adhesion, morphology, migration, oncogenic transformation, neovascularization, tumor invasion, and metastasis (29, 30). The intracellular expression of fibronectin has been reported to be stronger in the invading parts of most papillary thyroid carcinomas tumors compared with the central tumor regions (31). Up-regulation of fibronectin has also been reported in several other human cancers and its overexpression may be tumor promoting in hepatoblastoma (32) and ovarian cancer (33).

Using the four marker genes, the positive detection rates of patients with circulating cancer cells were 67% (4 of 6) for stage I, 100% (4 of 4) for stage II, 67% (6 of 9) for stage IIIa, 67% (14 of 21) for stage IIIb, and 79% (11 of 14) for stage IV. The detection rate indicates the sensitivity of using the four markers for detecting the presence of circulating NSCLC cells. The detection rate, however, does not provide information on the number of circulating cancer cells and the prognosis, which are better measured and assessed by the cancer cell load L_c. Our data showed that 16.7% of stage I, 50% of stage II, 22% of stage IIIa, 24% of stage IIIb, and 28.6% of stage IV patients had $L_c > 1$.

The percentage of patients with high L_c value is greater for the late-stage patients than for the early-stage patients except for the stage II patients. We looked into this issue and found that the markers have different detection rates for different histologic types of NSCLC. The markers are more sensitive for detecting squamous carcinoma (85.7%; 12 of 14) compared with adenocarcinomas (68.6%; 24 of 35) and others (60%, 3 of 5). An investigation on the histologic types of the NSCLC patients revealed that the available stage II patients were composed of three (75%) squamous carcinoma and one (25%) adenocarcinoma patients, whereas the other stage patients were composed of 11% to 33% squamous carcinoma, 50% to 71% adenocarcinoma, and 5% to 22% poorly differentiated cell type patients. The higher detection rate of the stage II patients can therefore be attributed to the limited number of clinical samples and their histologic composition.

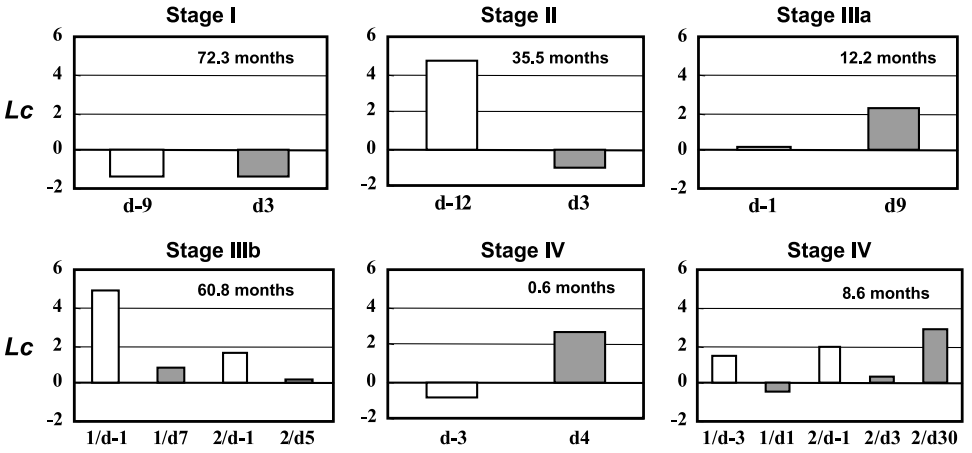

Fig. 4 Assessment of therapy efficacy with circulating cancer cell detection for six different NSCLC patients. Median survival time of each patient is indicated in the figure panels. The designations of the treatment and sampling day are described in the text.

The data shown in Figs. 2*B* and 3 indicate that patients of the same stage who had higher L_c had worse outcomes. The results indicate that L_c measurement is a supplementary tool to the traditional TNM staging method to better predict the outcome of cancer patients. In traditional metastasis models, metastatic cells are rare and arise during late stages of tumor progression (34). These models have been challenged by recent studies, which reported that most cancer cells in a primary tumor have a "metastatic phenotype" (35, 36). TNM classification is based on pathologic observations to determine the cancer progression stage by a set of guidelines using tumors size, metastases in the surgically removed regional lymph nodes, and the presence of distant metastases as parameters. Although the presence of metastases in the axillary lymph node predicts the development of distant metastases in breast carcinoma patients, 20% to 30% of the patients who were free of axillary lymph node metastases also developed disease at distant sites (37). This observation indicates that breast tumor cells can bypass the lymph nodes and disseminate directly through the blood to distant organs. Our observation of the presence of circulating cancer cells in the stage I NSCLC cancer patients also concur with the above findings in breast cancer patients.

Our data show that detection of circulating cancer cells is a valid supplement to the TNM method for better cancer staging. The two methods combined together provide better information for designing lung cancer treatment strategies. In this study, the 5-year survival rate of early-stage (I-IIIa) NSCLC patients was 30% to 50% after surgical resection. A more precise staging method that includes detection of circulating cancer cells will aid in deciding whether adjuvant therapeutic regiment in addition to tumor resection are beneficial to the patients. Early-stage patients who have circulating cancer cells in their peripheral blood may actually have later-stage disease. An additional C factor for detection of circulating cancer cells or disseminated tumor cells should make the TNM staging method more precise for prognosis and therapeutic purposes. On the other hand, although it is statistically more significant to use L_c for classifying late-stage patients compared with TNM staging ($P = 0.006$ versus $P = 0.03$ in Fig. 3*B* and *A*), the available data do not warrant using L_c alone for cancer staging. The detection of circulating cancer cells is more appropriate as a supplementary factor to the TNM staging.

At present, a reliable serologic biomarker assay for assessing the treatment response of NSCLC patients is not available. Two courses of chemotherapy are traditionally given before imaging is done to evaluate the treatment response of NSCLC patients. We showed in a limited number of patients that the L_c could be used for monitoring therapeutic response. Because different cancer cell types have different levels of marker gene expression, patients with higher L_c values do not necessarily have more circulating cancer cells in their peripheral blood than do patients with lower L_c values. Nevertheless, this semiquantitative approach is useful for measuring the relative cancer cell load in a patient's peripheral blood to monitor the effectiveness of treatment. This method highlights an alternative approach to rapidly assess the treatment response of NSCLC patients. This method may therefore help to design more comprehensive and reasonable therapeutic regimens at earlier dates for NSCLC patients.

REFERENCES

1. Bunn PA. Early detection of lung cancer using serum RNA or DNA markers: ready for "prime time" or for validation? J Clin Oncol 2003;21: 3891–3.

2. Sozzi G, Conte D, Leon M, et al. Quantification of free circulating DNA as a diagnostic marker in lung cancer. J Clin Oncol 2003;21:3902–8.

3. Forus A, Hoifodt HK, Overli GE, et al. Sensitive fluorescent *in situ* hybridisation method for the characterisation of breast cancer cells in bone marrow aspirates. Mol Pathol 1999;52:68–72.

4. Muller P, Weckermann D, Riethmuller G, et al. Detection of genetic alterations in micrometastatic cells in bone marrow of cancer patients by fluorescence *in situ* hybridization. Cancer Genet Cytogenet 1996; 88:8–16.

5. Denis MG, Lipart C, Leborgne J, et al. Detection of disseminated tumor cells in peripheral blood of colorectal cancer patients. Int J Cancer 1997;74:540–4.

6. Hardingham JE, Hewett PJ, Sage RE, et al. Molecular detection of blood-borne epithelial cells in colorectal cancer patients and in patients with benign bowel disease. Int J Cancer 2000;89:8–13.

7. Gross H, Verwer B, Houck D, et al. Model study detecting breast cancer cells in peripheral blood mononuclear cells at frequencies as low as 10^{-7}. Proc Natl Acad Sci U S A 1995;92:537–41.

8. Mitas M, Cole DJ, Hoover L, et al. Real-time reverse transcription-PCR detects KS1/4 mRNA in mediastinal lymph nodes from patients with non-small cell lung cancer. Clin Chem 2003;49:312–5.

9. Peck K, Sher YP, Shih JY, et al. Detection and quantitation of circulating cancer cells in the peripheral blood of lung cancer patients. Cancer Res 1998;58:2761–5.

10. Yamashita J, Matsuo A, Kurusu Y, Saishoji T, Hayashi N, Ogawa M. Preoperative evidence of circulating tumor cells by means of reverse transcriptase- PCR for carcinoembryonic antigen messenger RNA is an independent predictor of survival in non-small cell lung cancer: a prospective study. J Thorac Cardiovasc Surg 2002;124:299–305.

11. Keilholz U, Goldin-Lang P, Bechrakis NE, et al. Quantitative detection of circulating tumor cells in cutaneous and ocular melanoma and quality assessment by real-time reverse transcriptase-polymerase chain reaction. Clin Cancer Res 2004;10:1605–12.

12. Hoon DS, Wang Y, Dale PS, et al. Detection of occult melanoma cells in blood with a multiple-marker polymerase chain reaction assay. J Clin Oncol 1995;13:2109–16.

13. Taback B, Chan AD, Kuo CT, et al. Detection of occult metastatic breast cancer cells in blood by a multimolecular marker assay correlation with clinical stage of disease. Cancer Res 2001;61:8845–50.

14. Strausberg RL. The Cancer Genome Anatomy Project: new resources for reading the molecular signatures of cancer. J Pathol 2001;195:31–40.

15. Weynants P, Humblet Y, Canon JL, Symann M. Biology of small cell lung cancer: an overview. Eur Respir J 1990;3:699–714.

16. Bostick PJ, Chatterjee S, Chi DD, et al. Limitations of specific reverse-transcriptase polymerase chain reaction markers in the detection of metastases in the lymph nodes and blood of breast cancer patients. J Clin Oncol 1998;16:2632–40.

17. Sarantou T, Chi DD, Garrison DA, et al. Melanoma-associated antigens as messenger RNA detection markers for melanoma. Cancer Res 1997;57:1371–6.

18. Iwao K, Watanabe T, Fujiwara Y, et al. Isolation of a novel human lung-specific gene, LUNX, a potential molecular marker for detection of micrometastasis in non-small cell lung cancer. Int J Cancer 2001;91: 433–7.

19. Spataro V, Norbury C, Harris AL. The ubiquitin-proteasome pathway in cancer. Br J Cancer 1998;77:448–55.

20. Yamazaki T, Hibi K, Takase T, et al. PGP9.5 as a marker for invasive colorectal cancer. Clin Cancer Res 2002;8:192–5.

21. Tezel E, Hibi K, Nagasaka T, et al. PGP9.5 as a prognostic factor in pancreatic cancer. Clin Cancer Res 2000;6:4764–7.

22. Sasaki H, Yukiue H, Moriyama S, et al. Expression of the protein gene product 9.5, PGP9.5, is correlated with t-status in non-small cell lung cancer. Jpn J Clin Oncol 2001;31:532–5.

23. Yamada KM, Weston JA. Isolation of a major cell surface glycoprotein from fibroblasts. Proc Natl Acad Sci U S A 1974;71: 3492–6.

24. Jaffe EA, Mosher DF. Synthesis of fibronectin by cultured human endothelial cells. J Exp Med 1979;147:1779–91.

25. Chen LB, Maitland N, Gallimore PH, McDougall JK. Detection of the large external transformation-sensitive protein on some epithelial cells. Exp Cell Res 1977;106:39–46.

26. Villiger B, Kelly DG, Kuhn C, McDonald JA. Human alveolar macrophage fibronectin: synthesis and ultrastructural localisation. Clin Res 1980;28:745A.

27. Olden K, Yamada KM. Mechanism of the decrease in the major cell surface protein of chicken embryo fibroblasts after transformation. Cell 1997;11:957–69.

28. Smith HS, Riggs JL, Mosesson NW. Production of fibronectin by human epithelial cells in culture. Cancer Res 1979;39:4138–44.

29. D'ardenne AJ, McGee JO'D. Fibronectin in disease. J Pathol 1984; 12:235–351.

30. Akiyama SK, Olden K, Yamada KM. Fibronectin and integrins in invasion and metastasis. Cancer Metastasis Rev 1995;14:173–89.

31. Ryu S, Jimi S, Eura Y, Kato T, Takebayashi S. Strong intracellular and negative peripheral expression of fibronectin in tumor cells contribute to invasion and metastasis in papillary thyroid carcinoma. Cancer Lett 1999;146:103–9.

32. Takayasu H, Horie H, Hiyama E, et al. Frequent deletions and mutations of the β-catenin are associated with overexpression of cyclin D1 and fibronectin and poorly differentiated histology in childhood hepatoblastoma. Clin Cancer Res 2001;7:901–8.

33. Tapper J, Kettunen E, El-Rifai W, Seppälä M, Andersson LC, Knuutila S. Changes in gene expression during progression of ovarian carcinoma. Cancer Genet Cytogenet 2001;128:1–6.

34. Fidler IJ. The pathogenesis of cancer metastasis: the "seed and soil" hypothesis revisited. Nat Rev Cancer 2003;3:453–8.

35. Van't Veer LJ, Dai H, van de Vijver MJ, et al. Gene expression profiling predicts clinical outcome of breast cancer. Nature 2002;415: 530–6.

36. Bernards R, Weinberg RA. A progression puzzle. Nature 2002;418: 823–4.

37. Braun S, Pantel K, Muller P, et al. Cytokeratin-positive bone marrow micrometastases and survival of breast cancer patients with stage I-III disease. N Engl J Med 2000;342:525–33.