Neoplasia • Vol. 7, No. 12, December 2005, pp. 1073–1080 1073

www.neoplasia.com

BJ-TSA-9, a Novel Human Tumor–Specific Gene, Has Potential as a Biomarker of Lung Cancer[1]

Yunyan Li[2], Xueyuan Dong[2], Yanhui Yin, Yanrong Su, Qingwen Xu, Yuxia Zhang, Xuewen Pang, Yu Zhang and Weifeng Chen

Immunology Department, Peking University Health Science Center, Beijing, China

Abstract

Using bioinformatics, we have identified a novel tumor-specific gene *BJ-TSA-9*, which has been validated by Northern blot analysis and reverse transcription–polymerase chain reaction (RT-PCR). *BJ-TSA-9* mRNA was expressed in 52.5% (21 of 40) of human lung cancer tissues and was especially higher in lung adenocarcinoma (68.8%). To explore the potential application of *BJ-TSA-9* for the detection of circulating cancer cells in lung cancer patients, nested RT-PCR was performed. The overall positive detection rate was 34.3% (24 of 70) in peripheral blood mononuclear cells (PBMCs) of patients with various types of lung cancers and was 53.6% (15 of 28) in PBMCs of lung adenocarcinoma patients. In combination with the detection of two known marker genes *SCC* and *LUNX*, the detection rate was increased to 81.4%. A follow-up study was performed in 37 patients after surgical removal of tumor mass. Among nine patients with persistent detection of two to three tumor marker transcripts in PBMCs, six patients had recurrence/metastasis. In contrast, 28 patients with transient detection of one tumor marker or without detection of any tumor marker were all in remission. Thus, *BJ-TSA-9* may serve as a marker for lung cancer diagnosis and as a marker, in combination with two other tumor markers, for the prediction of the recurrence and prognosis of lung cancer patients.

Neoplasia (2005) **7,** 1073–1080

Keywords: Biomarker, *BJ-TSA-9*, lung cancer, micrometastasis, tumor-specific gene.

Introduction

Lung cancer is the leading cause of cancer-related deaths worldwide. In all developed countries and many developing countries, the 5-year survival rates for lung cancer are less than 15% and 5%, respectively [1]. Despite some advances in early detection, more radical surgical treatments, and multimodal therapeutic strategies over the last decade, the prognosis of patients with lung cancer remains relatively poor. Nearly 50% of early-stage lung cancer patients will relapse or develop distant metastases within 5 years after surgical removal of tumor mass. This fact suggests the existence of occult metastatic cells, which are undetectable by current methods [2]. The spread of cancer cells by the hematogenous route is important in lung cancer metastasis. Thus, development of sensitive and specific detection methods for circulating cancer cells in the peripheral blood may have important diagnostic, prognostic, and therapeutic implications [3].

Conventional immunohistochemistry and pathology are relatively insensitive to the detection of occult cancer cells in peripheral blood samples. Several research groups, including our group, have reported the use of reverse transcription–polymerase chain reaction (RT-PCR) to detect disseminated tumor cells in the peripheral blood [4–7]. This assay, which is currently believed to be the most sensitive means for identifying circulating tumor cells [8,9] and has been applied in the diagnosis of malignant melanomas [10–12], colon cancers [13], breast cancers [14–16], neuroblastomas [17], thyroid cancers [18], and hepatocellular carcinomas [6,19,20], is able to detect a single tumor cell present in the peripheral blood. As for lung cancer, several marker genes, such as *preproGRP*, *cytokeratin*, and carcinoembryonic antigen (*CEA*), have been employed to detect tumor cells dispersed in the circulation [21–23]. However, most of these known transcripts displayed significant levels of illegitimate transcription in peripheral blood samples taken from healthy volunteers [24–26]. It is hence desirable to have markers expressed in embryonic or cancerous tissues, but not in adult normal human tissues.

We have recently identified a novel tumor-specific gene *BJ-TSA-9*, which is highly expressed in lung cancer tissues. By nested RT-PCR, *BJ-TSA-9* was detected in peripheral blood mononuclear cell (PBMC) samples from lung cancer patients, indicating the presence of circulating cancer cells. We have also chosen two known tumor marker genes, squamous cell carcinoma antigen (*SCC*) and lung-specific X protein (*LUNX*), for detecting circulating lung cancer cells. *SCC* is a purified subfraction of tumor antigen TA-4 obtained from squamous cell

Abbreviations: *CEA*, carcinoembryonic antigen; *G3PDH*, glyceraldehyde-3-phosphate dehydrogenase; HERV, human endogenous retrovirus; ORF, open reading frame; PBMC, peripheral blood mononuclear cell; *SCC*, squamous cell carcinoma antigen

Address all correspondence to: Weifeng Chen, Immunology Department, Peking University Health Science Center, Beijing 100083, China. E-mail: wfchen@public.bta.net.cn

[1]The work was supported by grants from the James R. Kerr Program, the Ludwig Institute for Cancer Research (KSP003), and the China National 863 High Technology Program (2001AA217151).

[2]Yunyan Li and Xueyuan Dong contributed equally to this work.

Received 17 June 2005; Revised 28 August 2005; Accepted 29 August 2005.

DOI 10.1593/neo.05406

cancer tissues of the uterine cervix [27], which is also expressed at high concentrations in various neoplasias derived from squamous cells [28] and has been applied as a tumor marker of lung cancer [29]. *LUNX* is a lung-specific gene whose mRNA expression is strictly limited to lung tissues and has been reported to be a sensitive marker for lung cancer [30,31]. With the combination of *BJ-TSA-9* with *SCC* and *LUNX*, the detection rate of circulating tumor cells was markedly increased, and such detection was correlated with the clinical outcome of lung cancer patients. Thus, *BJ-TSA-9* appears to be a promising tumor biomarker of lung cancer.

Materials and Methods

Tumor Specimens

All samples from cancer patients were collected with written consents and were approved by the Hospital Ethics Review Committee. Clinical diagnosis was confirmed by pathological examination. Tumor stages and grades were classified according to the fifth edition of the Tumor–Node–Metastasis classification of the International Union Against Cancer. Samples of human lung cancer tissues and paired noncancerous tissues (5 cm away from the tumor) were obtained during surgical resection from the Peking University First Hospital. The resected tissue samples were immediately cut into small pieces and snap frozen in liquid nitrogen until use.

Blood Samples

Venous blood samples from lung cancer patients ($n = 70$) were collected preoperatively by standard transcutaneous needle venipuncture and placed into citrate sodium–containing tubes. Blood was collected into two tubes, with 1 ml in the first tube and 5 ml in the second tube. The blood in the first tube, which may have been contaminated with epithelial cells picked up by the needle when it pierced the skin, was discarded, whereas the blood in the second tube was used for RNA preparation. The blood samples were also collected from healthy donors ($n = 20$). The blood was placed on a Ficoll-Hypaque layer (Gibco BRL, Grand Island, NY), and the PBMCs were collected after density gradient centrifugation. The collected cells were washed twice with sterile phosphate buffer solution. Cell pellets were then snap frozen and stored at −80°C until RNA extraction.

Cell Lines

BJ-TSA-9 transcript was detected by RT-PCR in the following tumor cell lines: A549 and PAa (lung adenocarcinoma); Bel-7402, Bel-7405, HLE, and HepG2 (hepatocellular carcinoma); Jurkat (leukemia); and Mel-kenny (melanoma). These cell lines were maintained in DMEM (Gibco BRL) medium containing 10% fetal calf serum (Hyclone, Logan, UT).

cDNA samples from eight tumor cell lines of melanoma, lung, breast, pancreas, colon, prostate, and ovary cell lines were purchased from Clontech Laboratories, Inc. (Palo Alto, CA) and enrolled in this study.

RNA Preparation and cDNA Synthesis

Total RNA was extracted from cancer tissues and cell lines using Trizol (Gibco BRL) reagent, according to the protocol provided by the manufacturer. Total RNA was checked by agarose gel electrophoresis for ribosomal RNA integrity, and RNA quantity was measured by an ultraviolet spectrometer. Total RNA (2 μg) was primed with an oligo-$(dT)_{15}$ oligonucleotide (Promega, Madison, WI) and reverse-transcribed into cDNA using advantage reverse transcriptase (Clontech Laboratories, Inc.), according to the protocol provided by the manufacturer.

mRNA expression pattern in human normal tissues was assessed by RT-PCR, using a panel of commercially available cDNA (Clontech Laboratories, Inc.). The panel is composed of cDNA generated from RNA isolated from 16 human normal tissues: brain, heart, kidney, liver, lung, pancreas, placenta, skeletal muscle, colon, ovary, peripheral blood leukocyte, prostate, small intestine, spleen, testis, and thymus.

Northern Blot Analysis

Northern blot analysis was performed with total RNA samples extracted from lung cancer tissues and paired noncancerous tissues. RNA integrity was examined by electrophoresis in formalin/4-morpholine propane sulfuric acid gels. Twenty micrograms of total RNA per lane was first separated by electrophoresis in 1.2% agarose containing 3% formaldehyde and then blotted onto Hybond-C nylon membranes (Amersham, Piscataway, NJ). The membrane was cross-linked by ultraviolet radiation. DIG-labeled probes spanned the full length of the open reading frame (ORF) sequence (1101 bp) of *BJ-TSA-9*. After prehybridization, the membrane was hybridized overnight to specific DIG-labeled probes at 42°C in hybridization solution (50% formamide, 5× SSC, 0.1% *N*-lauroylsarcosine, 0.02% SDS, 2% blocking reagent, and 100 μg/ml sheared salmon sperm DNA). The hybridized filter was then sequentially washed at room temperature for 30 minutes in 0.1% SDS at 2× SDS, and at 68°C for 30 minutes in 0.1% SDS and 0.1× SDS. After stringent washes, the corresponding mRNA in Northern blot analysis was detected by chemiluminescence using CSPD ready-to-use reagent (Roche, Indianapolis, IN).

RT-PCR Analysis

Sequences of primers for target genes are listed in Table 1. Sample cDNA (2.5 μl) was 10-fold diluted with a PCR mixture containing 0.1 μmol per primer pair, 0.2 mM deoxynucleotide triphosphate, 50 mM Tris–HCl (pH 8.3), 10 mM KCl, 5 mM$(NH_4)_2SO_4$, 2 mM $MgCl_2$, and 0.75 U of Taq DNA polymerase (Tianwei, Beijing, China) in a total volume of 25 μl. PCR was performed using GeneAmp PCR system 9700 (Applied Biosystems, Foster City, CA). For cancer tissues and cancer cell lines, RT-PCR of *BJ-TSA-9* was performed with primers B2 and B4. The parameters were as follows: 35 cycles at 94°C for 20 seconds, at 64°C for 30 seconds, and at 72°C for 60 seconds, followed by 7 minutes at 72°C.

For peripheral blood samples, the primers of all these three genes for PCR amplification are listed in Table 1. Taking 2.5 μl of 10-fold–diluted cDNA as template for the

Table 1. Sequences of PCR Primers for Detecting the Three Marker Genes.

Gene	Sequence of Primers	Product (bp)
BJ-TSA-9	Outer: B1, 5′-CACTTCTTGGAGGTGCCCTGCACG-3′; B2, 5′-TAGAGGCAGCCAACAAGCGTG-3′	1421
	Inner: B3, 5′-CACTTTGGTTCCTGATGGCTT-3′; B4, 5′-CTCCCAAGCCCAGCTTTCCAAAGG-3′	564
SCC	Outer: S1, 5′-GCCCACCTCTGCTTCCTCTA-3′; S2, 5′-GCTTCTGCTCCCTCCTCTGT-3′	1083
	Inner: S3, 5′-GCAAATGCTCCAGAAGAAAG-3′; S4, 5′-CGAGGCAAAATGAAAAGATG-3′	261
LUNX	Outer: L1, 5′-GGGCCTCATTGTCTTCTACGGG-3′; L2, 5′-GAAGGGTGCAGTCACCAAGGAC-3′	534
	Inner: L3, 5′-CTCATTGTCTTCTACGGCTGTTAG-3′; L4, 5′-CTTTATGCCGAGAGGGATGGT-3′	396

first round, PCR conditions were as follows: 35 cycles at 94°C for 20 seconds, at 60°C for 30 seconds, and at 72°C for 90 seconds, followed by 7 minutes at 72°C. Then 2 μl of the 1:100–diluted first-round PCR product was used as template for the second round of the PCR reaction. The second round was cycled at 94°C for 30 seconds, 58°C for 20 seconds, and then 72°C for 50 seconds, for 28 cycles.

The risk of genomic DNA yielding a product was prevented by designing primers that bridged intronic sequences. Glyceraldehyde-3-phosphate dehydrogenase (*G3PDH*) gene was used as an endogenous control. *G3PDH* primer sequences were 5′-ACC ACA GTC CAT GCC ATC AC-3′ and 5′-TCC ACC ACC CTG TTG CTG TA-3′. Reactions in which pure water replaced RNA were used as RT-PCR–negative controls. Visualization of target bands on a 0.8% agarose gel with ethidium bromide staining was performed to determine the expression of mRNA transcripts.

Follow-up

Follow-up survey was obtained from 37 lung cancer patients every 3 to 6 months after operation. All patients were checked by chest X-ray, computed tomography, and others to confirm if there was any metastasis and/or recurrence. The average follow-up period was 24 months (range 19–30 months).

Statistical Analysis

Statistical significance was determined by SPSS 11.0 for windows. Chi-square test was performed for multiple comparisons among individual samples, and significant difference was defined as $P < .05$.

Results

Prediction of a Putative Tumor-Specific Gene from the Analysis of Human Chromosome 8q24

The DNA sequence of the q24 region of human chromosome 8 was obtained from the Human Genome Resource. Unknown coding regions in the DNA sequence were obtained by analysis of Genemark, an online tool used to predict coding regions in a DNA sequence. All potential coding regions were analyzed by BLAST programs for human expressed sequence tag (EST), Unigene, and virtual Northern blot analysis in the National Center for Biological Information server. In our analysis, when all potential coding regions located in chromosome 8q24 had been obtained, four criteria were used to screen the candidates for tumor-specific genes: first, there must be EST fragments representing candidate coding regions, and these candidate coding regions must be transcribed *in vivo*; second, these EST fragments must be cloned from libraries of either the tumor or the testis, implying that these EST fragments may be inside tumor-specific gene sequences; third, if there exist SAGE tags for sequences of the candidate coding regions, only the tags absolutely corresponding to the candidate gene should be selected because there are many SAGE tags that are homologous to the candidate genes but represent other genes; fourth, the candidate coding regions must represent the genes that have not yet been recorded as tumor-associated genes. When these four criteria are applied, a potential coding region is found and considered to be the most promising candidate for tumor-specific gene. In the Unigene database, there were 18 ESTs representing the coding region, all of which were derived from the libraries of different histologic types of cancers, including lung cancer, colon carcinoma, and pancreatic carcinoma. But there were no ESTs found in the libraries of normal colon or pancreas. In the virtual Northern blot analysis, one SAGE tag (AGCAGAACCG) representing the coding region was identified. We found that five SAGE libraries in the database contained at least one copy of this tag, including pancreatic cancer, ovary cancer, mesothelioma, and breast cancer cell line MCF7 SAGE libraries. We have proposed this coding region to be a potential tumor-specific gene and named it as *BJ-TSA-9* (GenBank accession no. AF497803).

Full-Length cDNA Cloning of BJ-TSA-9 *and Its Sequence Analysis*

Based on the potential coding region obtained from genomic DNA sequence analysis, *BJ-TSA-9* was assembled and extended by searching the human EST database. Gene-specific primers were designed, and the PCR products were sequenced to confirm the extended sequence. The sequence of *BJ-TSA-9* could not be further extended by RACE PCR. The 5′ untranslated sequence of *BJ-TSA-9* contained a stop codon in each of the three possible translational frames. Exon–intron boundaries were identified by comparing the genomic sequence with the full-length cDNA sequence. All sequences at exon–intron junctions are consistent with the AG–GT rule. Gene structure is shown in Figure 1, which comprises five exons with a full length of 1936 bp. Its ORF of 1101 bp encompasses parts of exons 1 and 5 and the full length of exons 2, 3, and 4. This gene encodes a protein of 367-amino-acid residues. The mRNA of *BJ-TSA-9* is in good agreement with the size observed in Northern blot analysis (Figure 2).

Expression of BJ-TSA-9 *in Normal and Cancer Tissues*

For the detection of *BJ-TSA-9* expression, 16 different normal tissues, including the brain, heart, kidney, liver, lung,

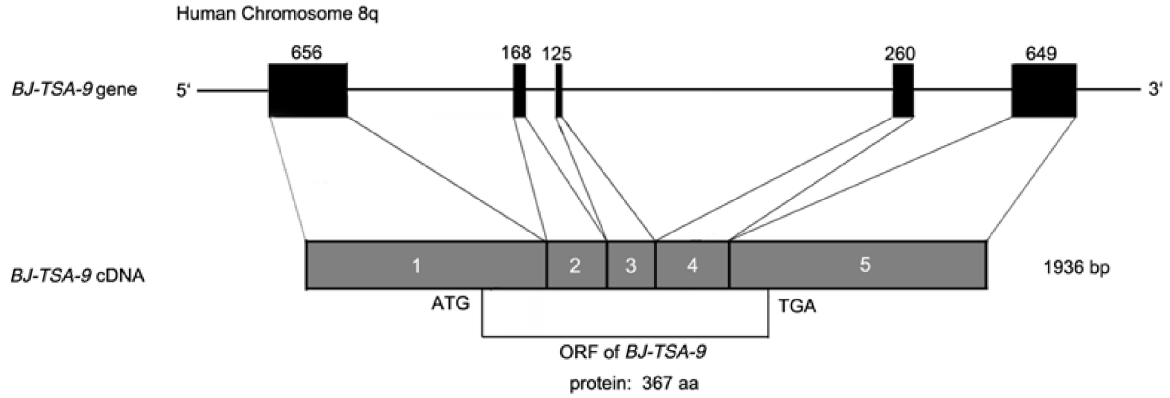

Figure 1. *Genomic structure of BJ-TSA-9 gene and its derived cDNA. BJ-TSA-9 comprises five exons. Exons are depicted schematically as boxes, and introns are depicted as lines with relative sizes. The number in the boxes indicates the exon number. The full-length cDNA and the ORF of BJ-TSA-9 are shown beneath the genomic structure.*

pancreas, placenta, skeletal muscle, colon, peripheral blood leukocyte, prostate, small intestine, spleen, testis, and thymus, were tested by RT-PCR. *BJ-TSA-9* mRNA transcript was undetectable in all these normal tissues tested (Figure 3*A*).

The mRNA expression of *BJ-TSA-9* was analyzed in the samples of lung cancer tissues and other cancer tissues. The *BJ-TSA-9* mRNA transcript was expressed in 21 of 40 lung cancer samples (52.5%) (Figure 3*B*), 4 of 12 colon cancer samples (33.3%), and 3 of 10 gastric cancer tissues (30.0%) (data not shown), but was not detected in surrounding noncancerous tissues. We further analyzed the frequency of *BJ-TSA-9* expression in different pathological types of lung cancer, and the expression rate was much higher in adenocarcinoma (68.8%) than in other types of lung cancer (squamous epithelial carcinoma, 47.1%; large cell lung cancer, 33.3%; small cell lung cancer, 25%). In the assessment for the expression of *BJ-TSA-9* mRNA in eight cancer cell lines available in our laboratory, including the lung adenocarcinoma (A549 and PAa), hepatocellular carcinoma (Bel-7402, Bel-7405, HLE, and HepG2), leukemia (Jurkat), and melanoma (Mel-kenny) cell lines, the *BJ-TSA-9* mRNA was detected in lung adenocarcinoma (PAa) only. There was no positive *BJ-TSA-9* signal detected in the cDNA samples of eight tumor cell lines (melanoma, lung, breast, pancreas, colon, prostate, and ovary cell lines) purchased from Clontech Laboratories, Inc. (data not shown).

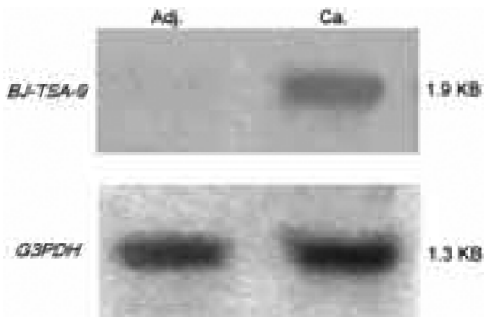

Figure 2. *Northern blot analysis of BJ-TSA-9 expression in lung cancer tissues and paired noncancerous tissues. G3PDH was used to monitor the quality of the RNA samples. Ca., cancerous tissues; Adj., adjacent noncancerous tissues.*

Detection of Disseminated Tumor Cells in the Peripheral Blood of Lung Cancer Patients

Due to the high detection rate of *BJ-TSA-9* in lung cancer tissues, we wondered whether it could be a biomarker of lung cancer metastasis. Then, nested RT-PCR was performed to detect disseminated tumor cells in the peripheral blood of lung cancer patients. The overall positive detection rate of *BJ-TSA-9* transcript in PBMCs was 34.3% (24 of 70). Taking into account the positive detection rate (52.5%) of *BJ-TSA-9* mRNA in lung cancer, the relative positive frequency of PBMC samples should be corrected to 65.3%, as only the lung cancer that expresses *BJ-TSA-9* mRNA can be detected in PBMCs. *BJ-TSA-9* transcript was undetectable in the PBMC samples of 20 healthy controls.

We further analyzed the association of clinicopathological variables with *BJ-TSA-9* transcript expression in the

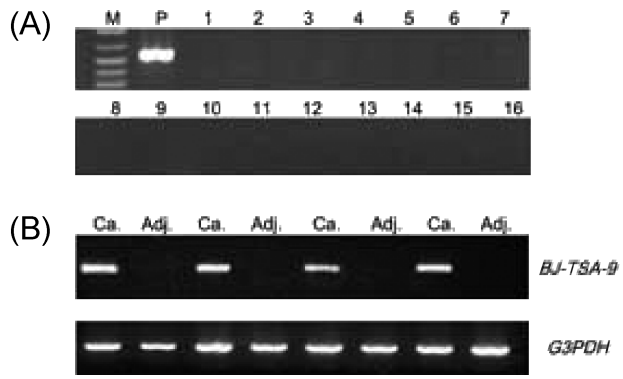

Figure 3. *The expression of the BJ-TSA-9 mRNA transcript in normal tissues and lung cancer samples. The assay was performed by RT-PCR. (A) BJ-TSA-9 mRNA was not expressed in any normal tissue. M, molecular maker; P, PAa (lung adenocarcinoma cell line); (1) brain; (2) heart; (3) kidney; (4) liver; (5) lung; (6) pancreas; (7) skeletal muscle; (8) placenta; (9) ovary; (10) thymus; (11) prostate; (12) spleen; (13) testis; (14) small intestine; (15) colon; and (16) leukocytes. (B) Expression of BJ-TSA-9 mRNA in representative lung cancer tissues and paired adjacent noncancerous tissues.* G3PDH *was used to monitor the quality of the RNA sample. Ca., cancerous tissues; Adj., adjacent noncancerous tissues.*

Table 2. Clinicopathological Characteristics and *BJ-TSA-9* Expression in Circulating Cancer Cells of 70 Lung Cancer Patients.

Factors Evaluated	*n*	Positive (*n*)	Positive Detection Rate
Pathology			
Squ	30	7	7/30 (23.3%)
Ade	28	15	15/28 (53.6%)
LCLC	5	1	1/5 (20.0%)
SCLC	7	1	1/7 (14.3%)
Clinical stages			
I and II	37	9	9/37 (24.3%)
III	23	9	9/23 (39.1%)
IV	10	6	6/10 (60.0%)
Tumor differentiation			
Good	19	6	6/19 (31.6%)
Moderate	28	11	11/28 (39.3%)
Poor	23	7	7/23 (30.4%)
Sex			
Male	14	5	5/14 (35.7%)
Female	56	19	19/56 (33.9%)
Total	70	24	24/70 (34.3%)

Squ, squamous epithelial carcinoma; *Ade*, adenocarcinoma; *LCLC*, large cell lung cancer; *SCLC*, small cell lung cancer.

peripheral blood of lung cancer patients. *BJ-TSA-9* showed different detection rates for different histologic types of lung cancer; the positive detection rate was relatively high in adenocarcinomas (53.6%; 15 of 28) and low in squamous epithelial carcinoma (23.3%; 7 of 30), large cell lung cancer (20.0%; 1 of 5), and small cell lung cancer (14.3%; 1 of 7) (Table 2). Of note, the positive detection rate of *BJ-TSA-9* correlated with the progression of clinical stages, being 24.3% for stages I and II, 39.1% for stage III, and 60.0% for stage IV. There was no statistically significant association found between *BJ-TSA-9* expression and gender, as well as tumor differentiation states.

Detection of Circulating Cancer Cells with Multiple Genes

In the attempt to increase the positive detection rate of circulating cancer cells in lung cancer patients, two additional known biomarkers of lung cancer (*SCC* and *LUNX*), together with *BJ-TSA-9*, were also detected in the PBMC samples of lung cancer patients (Figure 4). Among 70 lung cancer patients, the positive detection rate of *SCC* and *LUNX* in PBMCs was 32.9% and 40.0%, respectively. With the combination of transcripts of the two tumor markers (*BJ-TSA-9* + *LUNX*), the detection rate of circulating tumor cells increased to 62.9%; however, with the combination of three tumor markers (*BJ-TSA-9* + *LUNX* + *SCC*), the detection rate of circulating tumor cells increased to 81.4% (Figure 5). Therefore, the positive detection rate was dramatically increased with the combination of the three tumor marker gene transcripts for detection. In the control assays, all these three gene transcripts were undetectable in the PBMC samples from healthy controls.

Follow-up

To investigate the relationship between the detection of circulating tumor cells and the clinical outcome of lung cancer patients, a follow-up study was performed for an average of 24 months (range 19–30 months) in 37 patients after surgical removal of the tumor mass. Four patients with persistent detection of the three tumor marker transcripts (*BJ-TSA-9*, *LUNX*, and *SCC*) in PBMCs had recurrence and/or metastasis. Among these four patients, one died during this period. In another group of five patients with persistent detection of two markers, two had recurrence and/or metastasis. In contrast, 21 patients with transient detection of one tumor marker and seven patients without detection of any tumor marker all went into remission. The follow-up range and clinical data of all 37 patients are shown in Table 3.

Discussion

The unveiling of the sequence of human genome makes bioinformatics an important tool for discovering novel genes. It is well known that oncogenic retroviruses can cause cancers in various species. Human endogenous retroviruses (HERVs) are remnants of ancient retroviral infections that became fixed in the germ line DNA millions of years ago.

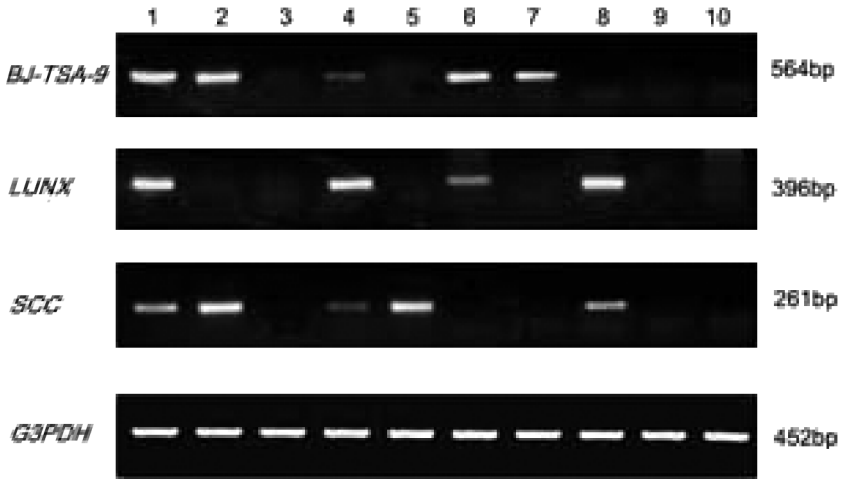

Figure 4. *Detection of marker genes in circulating cancer cells of lung cancer patients. The assay was performed by nested RT-PCR. Lanes 1–8: peripheral blood samples of eight lung cancer patients; lanes 9–10: peripheral blood samples of two healthy volunteers. G3PDH was used to monitor the quality of the RNA samples.*

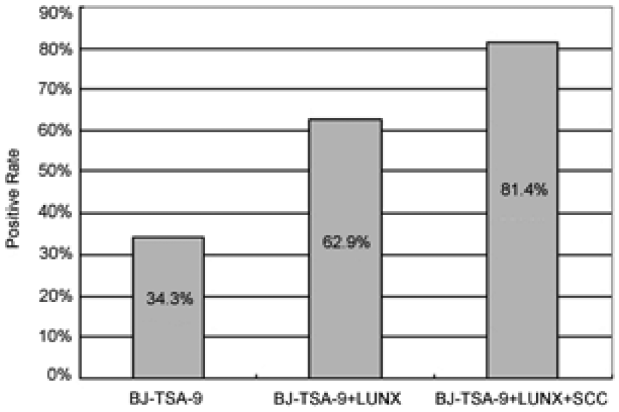

Figure 5. *Analysis of positive detection rates using different numbers of marker genes. The circulating lung cancer cells were detected by one, two, or three tumor markers using nested RT-PCR. The positive detection rates were correlated with the increase in the number of marker genes assessed.*

HERVs can potentially affect the expression of adjacent genes by insertion, recombination, and transcriptional/translational effects mediated by control elements in long terminal repeats. As there are some integration sites of HERV in the region of chromosome 8q24, we assumed that the integration of HERV might activate some tumor-specific genes, so that we have analyzed all the potential coding regions located in chromosome 8q24, including known and unknown genes. We identified a novel tumor-specific gene, named *BJ-TSA-9*, whose mRNA expression is strictly limited to cancer tissues and is not expressed in any normal tissue. For its specific expression pattern, *BJ-TSA-9* may have potential as a biomarker candidate for cancer.

Lung cancer is a very common and aggressive malignancy. Difficulty in early diagnosis and its propensity to metastasize or relapse result in the poor 5-year survival rate of lung cancer patients. One of the major causes of recurrence is metastasis of disseminated tumor cells through the blood circulation. The number of circulating tumor cells is so small such that they cannot be detected by conventional diagnostic methods, such as imaging studies and assays for serum marker detection. Recently, RT-PCR amplification of target DNA sequences in the peripheral blood has been increasing applied for the detection of micrometastases of cancer cells. This technology may have three advantages for cancer patients [10]: 1) detection of occult tumor cells at a very early stage [14]; 2) more efficient analysis of ways to monitor the efficacy of therapy [24]; and 3) prediction of the kinds of patients who may have favorable outcomes following removal/treatment of the primary lesion [32]. In this study, we have applied it in the detection of micrometastasis in lung cancer patients during the postoperative period.

CEA, *cytokeratin*, and *preproGRP* are the markers most frequently used to detect circulating tumor cells in the peripheral blood of lung cancer patients. However, the high false-positive detection rate in healthy controls is not satisfactory [25,33]. Aberrant transcription or contamination makes it unreliable for these markers to specifically detect disseminated cancer cells in the blood of lung cancer patients [24–26]. Thus, molecular markers with strict expression in cancers are preferred for the detection of disseminated cancer cells. *BJ-TSA-9* is a tumor-specific gene and is highly expressed in lung cancer. We herein applied *BJ-TSA-9* to detect circulating cancer cells by nested RT-PCR in lung cancer patients. The positive detection rate of circulating cancer cells was 34.3% (24 of 70) and 53.6% (15 of 28) in the peripheral blood of patients with various types of lung cancers and lung adenocarcinomas, respectively. As the *BJ-TSA-9* transcript was undetectable in the peripheral blood of 20 healthy donors and because there was no nonregulated transcription, *BJ-TSA-9* seemed fit as a biomarker of circulating lung cancer cells.

Further analysis on the correlation of clinicopathological variables with *BJ-TSA-9* transcript expression in the peripheral blood of lung cancer patients showed that *BJ-TSA-9* was more sensitive to detecting adenocarcinoma (53.6% positive) than other types (14.3–23.3% positive) of lung cancer. Early hematogenous metastasis is a prominent characteristic of lung adenocarcinoma, and its prognosis is poor. Therefore, the application of *BJ-TSA-9* is of significant value in the early detection of micrometastasis for lung adenocarcinoma patients.

Table 3. Follow-up Study in 37 Lung Cancer Patients: PCR Results of Tumor Marker Genes and Clinical Progression.

Metastasis and/or Relapse	Follow-up (months)	A/D	Site of Metastasis	PCR Result of Three Marker Genes (Case/Pathological Type)							
				$L^+B^+S^+$	L^+S^+	L^+B^+	S^+B^+	L^+	S^+	B^+	$L^-B^-S^-$
No	19–20	A				1/Squ		1/Ade		1/Ade	1/Ade
	21–22	A						1/Squ	4/Squ	2/Ade	3/Squ
								2/Ade		1/SCLC	
	23–24	A						1/Squ	1/Squ	1/Squ	1/Ade
										2/Ade	
	25–26	A					1/Squ				1/Squ
	27–28	A					1/Squ	1/Ade	2/Ade		
	29–30	A							1/Squ		1/Squ
Yes	23	A	Relapse		1/Squ						
	24	A	Liver			1/SCLC					
	21	A	Liver	1/Ade							
	24	A	Liver	1/Squ							
	25	A	Bone	1/Ade							
	26	D	Bone	1/Ade							

Squ, squamous epithelial carcinoma; *Ade*, adenocarcinoma; *SCLC*, small cell lung cancer; *B*, *BJ-TSA-9*; *S*, *SCC*; *L*, *LUNX*; *A*, alive; *D*, dead.

To increase the detection rate, we have chosen two additional known genes, *SCC* and *LUNX*, together with *BJ-TSA-9*, to detect circulating lung cancer cells. As shown in Figure 5, the detection rate of circulating lung cancer cells increased to 62.9% using two tumor markers, and to 81.4% using three tumor markers. With the combination of the three tumor markers, the detection of circulating lung adenocarcinoma cells also increased to 71.4%. Therefore, detection by nested RT-PCR assay of the presence of circulating cancer cells through the application of the three tumor markers, in combination with conventional Tumor–Node–Metastasis classification, may lead to early diagnosis and corresponding therapeutic modalities in patients with lung cancer, especially in patients with adenocarcinoma.

Moreover, the prognosis of lung cancer patients may be precisely predicted by the detection of circulating tumor cells using three the tumor markers described. The prognosis was extremely poor in patients with three tumor markers detected in the blood, was relatively poor in patients with two tumor markers detected in the blood, and could not be determined in patients with only one tumor marker detected in the blood. In our 24-month follow-up survey, there was no recurrence and/or metastasis in the lung cancer patients with one tumor marker detected in the blood or in those without tumor markers. To confirm the validity of our study, more patients and long-term follow-up surveys should be carried out.

In conclusion, *BJ-TSA-9* is a novel tumor-specific gene that is highly expressed in human lung adenocarcinoma. Detection of *BJ-TSA-9*, *SCC*, and *LUNX* transcripts in the peripheral blood by nested RT-PCR is a sensitive tool for monitoring disseminated cancer cells, and such detection may lead to early diagnosis of the recurrence and/or metastasis of lung cancer.

References

[1] Parkin DM (2001). Global cancer statistics in the year 2000. *Lancet Oncol* **2,** 533–543.

[2] Hosch SB, Scheunemann P, and Izbicki JR (2001). Minimal residual disease in non–small-cell lung cancer. *Semin Surg Oncol* **20,** 278–281.

[3] Matsunaga H, Hangai N, Aso Y, Okano K, Kawamura M, Kobayashi K, Kambara H, Hoger JH, and Mitsuhashi M (2002). Application of differential display to identify genes for lung cancer detection in peripheral blood. *Int J Cancer* **100,** 592–599.

[4] Castells A, Boix L, Bessa X, Gargallo L, and Pique JM (1998). Detection of colonic cells in peripheral blood of colorectal cancer patients by means of reverse transcriptase and polymerase chain reaction. *Br J Cancer* **78,** 1368–1372.

[5] Mori M, Mimori K, Ueo H, Karimine N, Barnard GF, Sugimachi K, and Akiyoshi T (1996). Molecular detection of circulating solid carcinoma cells in the peripheral blood: the concept of early systemic disease. *Int J Cancer* **68,** 739–743.

[6] Mou DC, Cai SL, Peng JR, Wang Y, Chen HS, Pang XW, Leng XS, and Chen WF (2002). Evaluation of MAGE-1 and MAGE-3 as tumour-specific markers to detect blood dissemination of hepatocellular carcinoma cells. *Br J Cancer* **86,** 110–116.

[7] Sher YP, Shih JY, Yang PC, Roffler SR, Chu YW, Wu CW, Yu CL, and Peck K (2005). Prognosis of non–small cell lung cancer patients by detecting circulating cancer cells in the peripheral blood with multiple marker genes. *Clin Cancer Res* **11,** 173–179.

[8] Burchill SA and Selby PJ (2000). Molecular detection of low-level disease in patients with cancer. *J Pathol* **190,** 6–14.

[9] Ghossein RA, Carusone L, and Bhattacharya S (1999). Review: polymerase chain reaction detection of micrometastases and circulating tumor cells: application to melanoma, prostate, and thyroid carcinomas. *Diagn Mol Pathol* **8,** 165–175.

[10] Alao JP, Mohammed MQ, Slade MJ, and Retsas S (1999). Detection of tyrosinase mRNA by RT-PCR in the peripheral blood of patients with advanced metastatic melanoma. *Melanoma Res* **9,** 395–399.

[11] Kulik J, Nowecki ZI, Rutkowski P, Ruka W, Rochowska M, Skurzak H, and Siedlecki JA (2001). Detection of circulating melanoma cells in peripheral blood by a two-marker RT-PCR assay. *Melanoma Res* **11,** 65–73.

[12] Smith B, Selby P, Southgate J, Pittman K, Bradley C, and Blair GE (1991). Detection of melanoma cells in peripheral blood by means of reverse transcriptase and polymerase chain reaction. *Lancet* **338,** 1227–1229.

[13] Denis MG, Lipart C, Leborgne J, LeHur PA, Galmiche JP, Denis M, Ruud E, Truchaud A, and Lustenberger P (1997). Detection of disseminated tumor cells in peripheral blood of colorectal cancer patients. *Int J Cancer* **74,** 540–544.

[14] Berois N, Varangot M, Aizen B, Estrugo R, Zarantonelli L, Fernandez P, Krygier G, Simonet F, Barrios E, Muse I, et al. (2000). Molecular detection of cancer cells in bone marrow and peripheral blood of patients with operable breast cancer. Comparison of CK19, MUC1 and CEA using RT-PCR. *Eur J Cancer* **36,** 717–723.

[15] Johnson PW, Burchill SA, and Selby PJ (1995). The molecular detection of circulating tumour cells. *Br J Cancer* **72,** 268–276.

[16] Kahn HJ, Yang LY, Blondal J, Lickley L, Holloway C, Hanna W, Narod S, McCready DR, Seth A, and Marks A (2000). RT-PCR amplification of CK19 mRNA in the blood of breast cancer patients: correlation with established prognostic parameters. *Breast Cancer Res Treat* **60,** 143–151.

[17] Mattano LA Jr, Moss TJ, and Emerson SG (1992). Sensitive detection of rare circulating neuroblastoma cells by the reverse transcriptase–polymerase chain reaction. *Cancer Res* **52,** 4701–4705.

[18] Ditkoff BA, Marvin MR, Yemul S, Shi YJ, Chabot J, Feind C, and Lo Gerfo PL (1996). Detection of circulating thyroid cells in peripheral blood. *Surgery* **120,** 959–964 (discussion 964–955).

[19] Paterlini-Brechot P, Vona G, and Brechot C (2000). Circulating tumorous cells in patients with hepatocellular carcinoma. Clinical impact and future directions. *Semin Cancer Biol* **10,** 241–249.

[20] Wong IH, Lau WY, Leung T, and Johnson PJ (2000). Quantitative comparison of alpha-fetoprotein and albumin mRNA levels in hepatocellular carcinoma/adenoma, non-tumor liver and blood: implications in cancer detection and monitoring. *Cancer Lett* **156,** 141–149.

[21] Castaldo G, Tomaiuolo R, Sanduzzi A, Bocchino ML, Ponticiello A, Barra E, Vitale D, Bariffi F, Sacchetti L, and Salvatore F (1997). Lung cancer metastatic cells detected in blood by reverse transcriptase–polymerase chain reaction and dot-blot analysis. *J Clin Oncol* **15,** 3388–3393.

[22] Peck K, Sher YP, Shih JY, Roffler SR, Wu CW, and Yang PC (1998). Detection and quantitation of circulating cancer cells in the peripheral blood of lung cancer patients. *Cancer Res* **58,** 2761–2765.

[23] Salerno CT, Frizelle S, Niehans GA, Ho SB, Jakkula M, Kratzke RA, and Maddaus MA (1998). Detection of occult micrometastases in non–small cell lung carcinoma by reverse transcriptase–polymerase chain reaction. *Chest* **113,** 1526–1532.

[24] Bostick PJ, Chatterjee S, Chi DD, Huynh KT, Giuliano AE, Cote R, and Hoon DS (1998). 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* **16,** 2632–2640.

[25] Dingemans AM, Brakenhoff RH, Postmus PE, and Giaccone G (1997). Detection of cytokeratin-19 transcripts by reverse transcriptase–polymerase chain reaction in lung cancer cell lines and blood of lung cancer patients. *Lab Invest* **77,** 213–220.

[26] Zippelius A, Kufer P, Honold G, Kollermann MW, Oberneder R, Schlimok G, Riethmuller G, and Pantel K (1997). Limitations of reverse-transcriptase polymerase chain reaction analyses for detection of micrometastatic epithelial cancer cells in bone marrow. *J Clin Oncol* **15,** 2701–2708.

[27] Kato H and Torigoe T (1977). Radioimmunoassay for tumor antigen of human cervical squamous cell carcinoma. *Cancer* **40,** 1621–1628.

[28] Mino N, Iio A, and Hamamoto K (1988). Availability of tumor-antigen 4 as a marker of squamous cell carcinoma of the lung and other organs. *Cancer* **62,** 730–734.

[29] Kano M, Shimada Y, Kaganoi J, Sakurai T, Li Z, Sato F, Watanabe G, and Imamura M (2000). Detection of lymph node metastasis of oesophageal cancer by RT-nested PCR for *SCC* antigen gene mRNA. *Br J Cancer* **82,** 429–435.

[30] Iwao K, Watanabe T, Fujiwara Y, Takami K, Kodama K, Higashiyama M,

Yokouchi H, Ozaki K, Monden M, and Tanigami A (2001). 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* **91,** 433–437.

[31] Mitas M, Hoover L, Silvestri G, Reed C, Green M, Turrisi AT, Sherman C, Mikhitarian K, Cole DJ, Block MI, et al. (2003). Lunx is a superior molecular marker for detection of non–small cell lung cancer in peripheral blood [corrected]. *J Mol Diagn* **5,** 237–242.

[32] Shin JH, Chung J, Kim HO, Kim YH, Hur YM, Rhim JH, Chung HK, Park SC, Park JG, and Yang HK (2002). Detection of cancer cells in peripheral blood of stomach cancer patients using RT-PCR amplification of tumour-specific mRNAs. *Aliment Pharmacol Ther* **16 Suppl 2,** 137–144.

[33] Ko Y, Grunewald E, Totzke G, Klinz M, Fronhoffs S, Gouni-Berthold I, Sachinidis A, and Vetter H (2000). High percentage of false-positive results of cytokeratin 19 RT-PCR in blood: a model for the analysis of illegitimate gene expression. *Oncology* **59,** 81–88.