A prospective PCR–based screening for the EML4-ALK oncogene in non-small cell lung cancer

Running title: PCR–based screening for EML4-ALK

Manabu Soda,1 Kazutoshi Isobe,2 Akira Inoue,3 Makoto Maemondo,4 Satoshi Oizumi,5 Yuka Fujita,6 Akihiko Gemma,7 Yoshihiro Yamashita,1 Toshioide Ueno,1 Kengo Takeuchi,8 Young Lim Choi,1,9 Hitoshi Miyazawa,10 Tomoaki Tanaka,10 Koichi Hagiwara,10 and Hiroyuki Mano,1,9,11 for
the North-East Japan Study Group and the ALK Lung Cancer Study Group

1Division of Functional Genomics, Jichi Medical University, Tochigi 329-0498, Japan;
2Department of Respiratory Medicine, Toho University Omori Medical Center, Tokyo 143-8541, Japan; 3Tohoku University Hospital, Miyagi 980-8574, Japan; 4Miyagi Cancer Center, Miyagi 981-1293, Japan; 5First Department of Medicine, Hokkaido University School of Medicine, Hokkaido 060-8638, Japan; 6Asahikawa Medical Center, Hokkaido 070-8644, Japan; 7Nippon Medical School Hospital, Tokyo 113-8603, Japan; 8Pathology Project for Molecular Targets, The Cancer Institute, Tokyo 135-8550, Japan; 9Department of Medical Genomics, Graduate School of Medicine, University of Tokyo, Tokyo 113-0033, Japan; 10Saitama Medical University Hospital, Saitama 350-0495, Japan; 11CREST, Japan Science and Technology Agency, Saitama 332-0012, Japan
Key words: EML4-ALK, RT-PCR, molecular diagnosis, lung carcinoma

Corresponding author: Hiroyuki Mano, Division of Functional Genomics, Jichi Medical University, 3311-1 Yakushiji, Shimotsukeshi, Tochigi 329-0498, Japan. Tel.: +81-285-58-7449. Fax: +81-285-44-7322. E-mail: hmano@jichi.ac.jp

Conflict of Interest Statement: H.M. is an employee of CureGene Co., Ltd., and receives research grants from Astellas Pharma Inc. and Illumina Inc. H.M. also owns a patent for EML4-ALK oncogene, and serves as a member of scientific advisory board to Astellas Pharma Inc., Chugai Pharmaceutical and Daiichi Sankyo Co., Ltd.

Word count (excluding references): 4567
Total number of tables and figures: 4
Translational Relevance

The recent approval of an ALK inhibitor by the U.S. Food and Drug Administration has rendered urgent the development of a diagnostic scheme for tumors harboring ALK fusion genes. Whereas FISH is effective for analysis of formalin-fixed, paraffin-embedded (FFPE) tissue, how to test other types of specimen remains unsettled. We conducted a prospective, nationwide screening for EML4-ALK– or KIF5B-ALK–positive lung carcinomas in Japan with the use of a newly developed multiplex reverse transcription (RT)-PCR system. Various subtypes of EML4-ALK cDNA were identified in 36 of 808 specimens with adequate RNA quality. The RT-PCR results were concordant with those of immunohistochemistry, and EML4-ALK PCR products were detected in independent specimens from the same individuals. As far as we are aware, our study represents the first prospective, RT-PCR–based screening for EML4-ALK, and it demonstrates that multiplex RT-PCR is reliable for detection of the fusion gene in non-FFPE specimens.
Abstract

**Purpose:** *EML4-ALK* is a lung cancer oncogene, and ALK inhibitors show marked therapeutic efficacy for tumors harboring this fusion gene. It remains unsettled, however, how the fusion gene should be detected in specimens other than formalin–fixed, paraffin–embedded tissue. We here tested whether reverse transcription (RT) and PCR–based detection of *EML4-ALK* is a sensitive and reliable approach.

**Experimental Design:** We developed a multiplex RT-PCR system to capture *ALK* fusion transcripts and applied this technique to our prospective, nationwide cohort of non-small cell lung cancer in Japan.

**Results:** During February to December 2009, we collected 916 specimens from 853 patients, quality filtering of which yielded 808 specimens of primary non-small cell lung cancer from 754 individuals. Screening for *EML4-ALK* and *KIF5B-ALK* with our RT-PCR system identified *EML4-ALK* transcripts in 36 samples (4.46%) from 32 individuals (4.24%). The RT-PCR products were detected in specimens including bronchial washing fluid (*n* = 11), tumor biopsy (*n* = 8), resected tumor (*n* = 7), pleural effusion (*n* = 5), sputum (*n* = 4), and metastatic lymph node (*n* = 1). The results of RT-PCR were concordant with those of sensitive immunohistochemistry with ALK antibodies.

**Conclusions:** Multiplex RT-PCR was confirmed to be a reliable technique for detection of *ALK* fusion transcripts. We propose that diagnostic tools for *EML4-ALK* should be selected in a manner dependent on the available specimen types. FISH and sensitive immunohistochemistry should be applied to formalin–fixed, paraffin–embedded tissue, but multiplex RT-PCR is appropriate for other specimen types.
Introduction

An oncogenic fusion between the echinoderm microtubule-associated protein–like 4 gene (*EML4*) and the anaplastic lymphoma kinase gene (*ALK*) was discovered by functional screening with a non-small cell lung cancer (NSCLC) specimen (1). *EML4* and *ALK* are located within a short distance (~12 Mbp) of each other on the short arm of human chromosome 2, and a small inversion involving the two loci is responsible for generation of the *EML4-ALK* fusion in lung cancer. The *EML4-ALK* tyrosine kinase undergoes constitutive dimerization through a coiled-coil domain within *EML4*, resulting in kinase activation and conferring potent transforming ability (2, 3). Transgenic mice expressing *EML4-ALK* in lung alveolar cells develop multiple adenocarcinoma nodules soon after birth, but treatment with an ALK inhibitor results in the rapid clearance of such nodules, confirming the addiction of EML4-ALK–positive tumors to the kinase activity of the fusion protein (4). The therapeutic efficacy of ALK inhibitors has been confirmed in other transgenic mice expressing EML4-ALK (5).

Several ALK inhibitors have already entered clinical trials or are under preclinical development (6-10). Marked therapeutic efficacy of one such compound, crizotinib, has been described in patients with NSCLC positive for *EML4-ALK*, with an overall response rate of 57% (7), and crizotinib was recently approved as a therapeutic drug by the U.S. Food and Drug Administration within a remarkably short period after target discovery (3, 11).

The failure of crizotinib treatment in individuals without oncogenic ALK fusions (12) and an adverse effect of treatment with gefitinib on the prognosis of NSCLC patients who do not harbor mutations of the epidermal growth factor receptor gene (*EGFR*) (13) both suggest that ALK inhibitors should be administered only to patients positive for oncogenic ALK proteins. FISH–based detection of *ALK* rearrangements has proved to be of diagnostic utility in the trials with crizotinib (7). Furthermore, detection of ALK proteins by sensitive immunohistochemistry (IHC) has been described (14, 15), and one such IHC screening approach resulted in the identification of another oncogenic ALK fusion, KIF5B-ALK (14). However, a substantial proportion of patients attending clinics are diagnosed with lung cancer on the basis of...
pathological analysis of bronchial lavage fluid, pleural effusion, or sputum. Given that these specimens are not always suitable for the preparation of formalin–fixed, paraffin–embedded (FFPE) tissue required for FISH or IHC, individuals who are diagnosed solely by analysis of such specimens cannot receive EML4-ALK tests. To allow the sensitive detection of EML4-ALK and KIF5B-ALK in such specimens, we have now developed a multiplex reverse transcription (RT)-PCR system that captures the two ALK fusions, and we have tested its reliability as a diagnostic tool in our large-scale prospective cohort.

Materials and Methods

Prospective collection of NSCLC specimens

During February to December of 2009, we collected a total of 916 lung cancer specimens from 853 independent patients through our multicenter, nationwide networks in Japan. All specimens but resected tumors were mixed with RLT buffer (Qiagen, Valencia, CA) immediately after sampling, a step that markedly inhibits RNA degradation for up to 3 days at room temperature (data not shown). Resected tumor samples were snap frozen and stored at –80°C until extraction of RNA and DNA. Portions of the samples were sent to Jichi Medical University for multiplex RT-PCR analysis of EML4-ALK and KIF5B-ALK fusions and to Saitama Medical University for peptide nucleic acid–locked nucleic acid (PNA-LNA) PCR clamp analysis of EGFR mutations (16). All specimens were confirmed by pathological analysis to contain malignant cells. More than half of the specimens were collected through the North-East Japan Study Group network according to the NEJ004 protocol. The study was approved by the institutional review board of each participating center, and written informed consent was obtained from each study subject. All statistical analysis was performed with two-sided tests, and a P value of <0.05 was considered statistically significant.

Clinicopathologic features of EML4-ALK–positive NSCLC

The clinicopathologic features of patients with EML4-ALK–positive or –negative tumors in our cohort are summarized in Table 1 and Supplementary Table S1. Consistent with previous
observations, \textit{EML4-ALK}–positive patients were significantly younger than those without \textit{EML4-ALK} \((P < 0.001, \text{Student’s } t \text{ test})\) and were enriched in never- or light smokers \((P < 0.001, \text{Fisher’s exact test})\). Our data also indicated that \textit{EML4-ALK}–positive tumors are more likely to occur in women than in men \((P < 0.001, \text{Fisher’s exact test})\). In the present cohort, \textit{EML4-ALK} was detected only in lung adenocarcinoma \((P < 0.001, \text{Fisher’s exact test})\), for which the fusion-positive rate was 6.11%.

A total of 718 specimens was screened for \textit{EGFR} mutations, with such mutations being detected in 171 cases (23.8%). Whereas most \textit{EML4-ALK}–positive tumors did not harbor \textit{EGFR} mutations \((P = 0.002, \text{Fisher’s exact test})\), we did detect one tumor doubly positive in this regard. \textit{EML4-ALK} and \textit{EGFR} mutations are largely mutually exclusive (17, 18), but, importantly, such exclusiveness may not be absolute (19). Given that the presence of \textit{EML4-ALK} and \textit{EGFR} mutations in our doubly positive patient was examined with cells isolated from bronchial washing fluid, which was the only available specimen for molecular analysis in this individual, we were not able to determine whether there was a genuinely double-positive tumor in the lung or there were multiple independent tumors each positive for \textit{EML4-ALK} or mutated \textit{EGFR}.

We also attempted to examine the mutation status of \textit{KRAS} among our 32 cases positive for \textit{EML4-ALK}. We were able to sequence \textit{KRAS} cDNAs for 26 of these patients, none of whom showed \textit{KRAS} alterations (data not shown), confirming the mutual exclusivity of \textit{EML4-ALK} and \textit{KRAS} mutations (17, 20, 21).

\textbf{Quality assessment of samples}

Complementary DNA prepared from the specimens was first subjected to RT-PCR analysis with primers (5’-CTGTGGAGGCTGAGCTGGATC-3’ and 5’-TCATCAAAGCTCCACGGTG-3’) specific for the human ribonuclease P (RNase P) gene (GenBank accession number NM_005837). Given that we previously showed that the abundance of RNase P mRNA is similar to that of \textit{EML4-ALK} mRNA in NSCLC (data not shown), we used the successful amplification of RT-PCR products for RNase P as a threshold for selection of specimens for further analysis. Exclusion of small cell lung cancer specimens and filtering on the basis of RNase P mRNA abundance resulted
in the isolation of 808 specimens of primary NSCLC obtained from 754 individuals.

As shown in Supplementary Figure S1, bronchial washing fluid, including bronchoalveolar lavage fluid and washing fluid for the brush, needle, forceps, and other implements used in bronchoscopy, constituted 66.3% of the 808 eligible samples, with the remaining specimens including pleural effusion (12.8%); surgically resected tumor (7.05%); sputum (4.33%); tumor biopsy tissue including that obtained by transbronchial lung biopsy and transbronchial needle aspiration (3.71%); peripheral blood (3.71%); cardiac effusion, spinal fluid, or ascites (1.36%); and metastatic lesions of NSCLC (0.74%).

**Multiplex RT-PCR analysis of *EML4-ALK* and *KIF5B-ALK***

Each specimen (with the exception of resected tumors) was mixed with an equal volume of RLT buffer at the institute at which it was harvested. The resulting mixture was sent to Jichi Medical University, where DNA and RNA were extracted with the use of an automated BioRobot EZ1 workstation (Qiagen). The isolated RNA was subjected to RT with a ReverTra Ace qPCR RT kit (Toyobo, Osaka, Japan), and the resulting cDNA was subjected to PCR for 50 cycles of incubation at 94°C for 15 s, 60°C for 30 s, and 72°C for 1 min with AmpliTaq Gold DNA polymerase (Applied Biosystems, Foster City, CA) and with 2 μM of each of the following primers: F-1, 5’-GCTTTCCCCGCAAGATGGACGG-3’; F-2, 5’-TACCAGTGTGCTTCAATTGCAGG-3’; F-3, 5’-GTGCAGTGTTTAGCATTCTTGGG-3’; F-4, 5’-AGCTACATCACACACCTTGACTGG-3’; F-5, 5’-TCAAGCACATCTCAAGAGCAAGTG-3’; F-6, 5’-ATCCTGCGGAACACTATTCAGTGG-3’; F-7, 5’-GACAGTTGGAGGAATCTGTCGATG-3’; F-8, 5’-CAGCTGAGAGAGTGAAAGCTTTGG-3’; and R-1, 5’-TCTTGCCAGCAAAGCAGTAGTTGG-3’. All PCR products were subjected to Sanger sequencing to confirm the presence of *EML4-ALK* or *KIF5B-ALK* cDNA.
Results

Multiplex RT-PCR system

In addition to the original EML4-ALK fusion cDNA in which exon 13 of EML4 is fused to exon 20 of ALK in an in-frame manner (designated the E13;A20 variant by analogy with karyotype nomenclature; see http://atlasgeneticsoncology.org/Tumors/inv2p21p23NSCCLungID5667.html), 14 different variants of EML4-ALK have been described (1, 14, 21-27). Seven exons of EML4 are theoretically capable of in-frame fusion with exon 20 of ALK (Fig. 1A), and all but the E1;A20 variant would be expected to produce an oncogenic EML4-ALK protein, given that the coiled-coil domain encoded by exon 2 is required for constitutive dimerization of EML4-ALK. In addition, six different exons of KIF5B are theoretically capable of in-frame fusion with exon 20 of ALK (Fig. 1A).

To detect any such EML4-ALK or KIF5B-ALK fusion mRNAs, we developed a multiplex RT-PCR system. We had previously screened our archive of frozen tumors by RT-PCR analysis with two forward primers targeted to EML4 and one reverse primer targeted to ALK (24), but such PCR conditions resulted in the amplification of products as large as ~1300 bp for some variants. In this prospective study, we were faced with the analysis of a large number of samples with different levels of RNA quality. If the size of PCR products varied substantially among different EML4-ALK or KIF5B-ALK variants, some variants with large PCR products might not be amplified efficiently from specimens with low RNA quality. To be able to diagnose all possible fusions even with such samples, we therefore designed four forward primers for each of EML4 and KIF5B so that the size variation among all possible RT-PCR products is minimal (Fig. 1A). This new multiplex system faithfully detected all known fusion variants from EML4-ALK–positive specimens in our previous archive of NSCLC (data not shown).

To examine the sensitivity of our RT-PCR system, we mixed EML4-ALK–expressing BA/F3 cells (0 to 1 × 10^6) with EML4-ALK–negative cells (1 × 10^6) and then subjected them to RT-PCR analysis. A fusion cDNA was readily identified even with 10 positive cells (0.001%) among 1 × 10^6 negative cells (Fig. 1B), demonstrating the high sensitivity of the RT-PCR system.
To confirm the potential of our RT-PCR–based system, we compared it with a sensitive IHC approach and with FISH for the diagnosis of our archive of surgically resected and freshly frozen tumors with high RNA quality. Fifteen NSCLC specimens that previously stained positive by our sensitive IHC approach, which is based on an intercalated antibody–enhanced polymer (iAEP) method (14), were analyzed by RT-PCR and FISH together with 96 iAEP-negative specimens in a blinded manner. RT-PCR analysis of all these specimens \( n = 111 \) yielded a diagnosis identical to that obtained with the iAEP method \( P = 7.3 \times 10^{-19} \), Fisher’s exact test \( (\text{data not shown}) \). Analysis of the same sample set by a split FISH assay with Vysis probes \( \text{(Abbott Laboratories, Abbott Park, IL)} \) revealed that all of the iAEP-positive cases showed a rearranged \( ALK \) locus, whereas one iAEP-negative sample gave a discordant result (negative by iAEP and RT-PCR but positive by FISH) \( (\text{Supplementary Fig. S2}) \). The reason for this discrepant result remains unclear, but the multiple signals obtained with the 3’-\( ALK \) probe in the FISH analysis are indicative of amplification of the \( ALK \) gene or its adjacent region. Despite this discrepancy, the RT-PCR and iAEP data were highly concordant with the FISH results \( P = 1.2 \times 10^{-17} \), Fisher’s exact test \( ). \) Compared with the iAEP method, therefore, both the sensitivity and specificity of our RT-PCR system were 100%. In comparison with the Vysis FISH, the sensitivity and specificity of RT-PCR were 93.8% and 100%, respectively.

**Detection of \( EML4-ALK \)**

Screening of the 808 eligible specimens with our multiplex RT-PCR system identified positive products in 36 samples \( (4.46\%) \) obtained from 32 different individuals \( (4.24\%) \) \( (\text{Table 1, Fig. 2A}) \). Nucleotide sequencing of each PCR product identified 19 cases positive for the E13;A20 variant, 10 cases for E6a/b;A20, and a single case each for E18;A20, E20;A20, and a novel variant. \( EML4-ALK \) was detected in a wide range of specimens including bronchial washing fluid \( (n = 11) \), tumor biopsy \( (n = 8) \), resected tumor \( (n = 7) \), pleural effusion \( (n = 5) \), sputum \( (n = 4) \), and metastatic lymph node \( (n = 1) \). We did not detect any \( KIF5B-ALK \) cDNAs, confirming the rarity of this fusion gene.

Importantly, an E13;A20 product was consistently identified in both of the sputa obtained
at different time points from patient J-#1. Likewise, an E13;A20 product was detected in both the
tumor biopsy and sputum from patient J-#53 as well as in the pleural effusion and two resected
tumor specimens from patient J-#330, supporting the reliability of our RT-PCR approach.

Sequence determination for the RT-PCR product from patient J-#189 revealed that exon 14 of \textit{EML4} was fused to exon 20 of \textit{ALK} with an intervening sequence. Genomic PCR analysis
of the J-#189 specimen with a forward primer targeted to exon 14 of \textit{EML4} and a reverse primer
targeted to exon 20 of \textit{ALK} yielded a specific product, nucleotide sequencing of which revealed
that a position 453 bp downstream of \textit{EML4} exon 14 was ligated to a position 56 bp upstream of
\textit{ALK} exon 20 (Fig. 2B). In the transcript of this fusion gene, exon 14 of \textit{EML4} is thus spliced to a
TT sequence that is located within \textit{EML4} intron 14 and which is directly ligated to intron 19 of
\textit{ALK}. This splicing event results in an in-frame fusion between the mRNA sequences derived
from \textit{EML4} and \textit{ALK}. Furthermore, a full-length cDNA for this variant, here designated
E14::ins2;ins56A20, was isolated by RT-PCR analysis (Supplementary Fig. S3), and the potent
transforming ability of the encoded protein was confirmed with an in vitro focus formation assay
(Supplementary Fig. S4).

\textbf{Comparison between multiplex RT-PCR and sensitive IHC}

Finally, we applied the iAEP method to the \textit{EML4-ALK}–positive cases for which FFPE
specimens were also available \((n = 15)\). All but two cases (J-#393 and J-#927) manifested clear
immunoreactivity with antibodies to \textit{ALK} (Table 1). FISH analysis of these two specimens also
failed to detect the \textit{EML4-ALK} rearrangements (Fig. 3). Given that genomic DNA was not
available for the tumor of patient J-#393, we were not able to determine whether the PCR result
was a false positive. For J-#927, however, PCR analysis of genomic DNA with a forward primer
targeted to \textit{EML4} exon 6 and a reverse primer to \textit{ALK} exon 20 resulted in the amplification of an
~8.8-kbp genomic fragment, nucleotide sequencing of which revealed a fusion event between
intron 6 of \textit{EML4} and intron 19 of \textit{ALK} (Supplementary Fig. S5). Isolation of the genomic fusion
point thus indicates that J-#927 indeed harbors an \textit{EML4-ALK}–positive tumor.
Discussion

We have performed a large-scale prospective screening for EML4-ALK with an RT-PCR–based approach. Whereas RNA extraction and cDNA synthesis add extra labor to the diagnostic procedure, certain introns of EML4 are too large (intron 6 spans >16 kbp, for instance) for reliable amplification by genomic PCR. We therefore adopted RT-PCR as the method for our prospective screening. Specific PCR products were successfully isolated from different types of specimen, even from sputum (J-#1, J-#53, J-#215) and washing fluid of a tumor biopsy needle (J-#530). Multiple positive results obtained with different specimens of the same individuals further reinforce the reliability of our multiplex RT-PCR system as a diagnostic tool for EML4-ALK–positive tumors. Importantly, a subset of EML4-ALK–positive individuals diagnosed in the present study entered a clinical trial for crizotinib, and the response rate of the evaluable patients (n = 9) was 100% with this drug, again verifying the accuracy of our RT-PCR–based diagnosis.

The frequency of EML4-ALK in our cohort was 4.24% for all NSCLC cases and 6.11% for lung adenocarcinoma, values similar to those obtained in previous studies (20, 21). However, our prevalence data might be overestimates because the knowledge of mutual exclusiveness for EML4-ALK and EGFR mutations may have affected patient selection for our specimen collection. Indeed, EGFR mutation frequency among our cohort (23.8%) is slightly lower than that (30.9%) determined in a previous large-scale screening in Japan (28).

The clinicopathologic features of patients with EML4-ALK–positive tumors determined in the present study are also in agreement with those previously described, with a bias toward a young age, adenocarcinoma histology, and never or light smoking. Whereas a previous large-scale screening for EML4-ALK based on FISH did not detect a sex preference for the fusion gene (7), our cohort revealed a significant female preference. Such a sex difference was evident even among individuals below 40 years of age (P = 0.03, Fisher’s exact test) and among those with an adenocarcinoma histology (P = 0.005, Fisher’s exact test). Further large-scale studies are warranted to determine whether this uneven sex distribution of EML4-ALK is related to particular
clinicopathologic features or ethnic groups.

Given that *EML4-ALK* and *EGFR* mutations are almost mutually exclusive and that the fusion gene is enriched in lung adenocarcinoma with an early onset, it should prove to be clinically beneficial to pay special attention to such subsets of patients. Indeed, *EML4-ALK* was detected in 27.7% of *EGFR* mutation–negative adenocarcinomas in individuals of <50 years of age and in 50.0% of those in individuals of <40 years of age in our cohort. Given the marked efficacy of ALK inhibitors in patients with *EML4-ALK*–positive NSCLC (7), however, physicians should not dismiss the diagnosis in other subsets of patients. For example, *EML4-ALK* was even detected in an 80-year-old woman and in another woman with an intense smoking history (82 pack-years) (Table 1).

Multiplex RT-PCR has both advantages and disadvantages compared with other techniques. Importantly, the accuracy of RT-PCR–based diagnosis depends markedly on the RNA quality of specimens. In our cohort, for instance, 71 (7.75%) of the initial 916 specimens were excluded from *EML4-ALK* screening because of a failure to obtain PCR products for RNase P (the other 37 samples were excluded because they were not NSCLC). Low RNA quality thus clearly hampers reliable RT-PCR–based diagnosis.

Also, as expected, there was a large variation in the PCR cycle number required for successful amplification among specimens. In our cohort, 50 cycles of PCR allowed detection of PCR products for all positive cases, but such extensive amplification may also generate nonspecific products (as shown in Fig. 2A). Further optimization of primer sequences or combinations may minimize the generation of such by-products. Furthermore, whereas our system should be able to capture all in-frame fusions of *ALK* to *EML4* or *KIF5B*, it is not capable in its present form of detecting *ALK* fusions to other partners, such as *KLC1-ALK*, which was recently shown to be present infrequently in NSCLC (29).

On the other hand, RT-PCR can be readily applied to specimens such as sputum, bronchial washing fluid, or pleural effusion that may not be suitable for preparation of FFPE samples. Whereas the latter two specimen types can be used for the preparation of cell blocks suitable for
analysis by FISH or IHC, this procedure may not be as widely adopted in the clinic as is FISH or IHC. More importantly, it is difficult to generate cell blocks or FFPE samples from sputum. Our current prospective screening identified four EML4-ALK–positive sputa out of 35 samples (Table 1, Supplementary Fig. S1), showing that sputum is a suitable specimen for RT-PCR analysis. Indeed, sputum was the only available specimen from patient J-#215 both for the diagnosis of NSCLC and for the detection of EML4-ALK. If RT-PCR had not been applied to this patient’s sputum, we would not have been able to identify her tumor as positive for EML4-ALK, and she would not have had the chance to receive treatment with an ALK inhibitor in Japan.

Furthermore, PCR-based detection of EML4-ALK should have a higher analytic sensitivity compared with IHC or FISH (Fig. 1B). Even with sputum obtained from a patient with chronic bronchitis, RT-PCR was able to readily detect EML4-ALK at a concentration of 10 positive cells/mL (1). Thus, provided that RNA is not substantially degraded, RT-PCR–based diagnosis is expected to have a strong advantage with regard to the detection of low numbers of EML4-ALK–positive cells.

Ideally, every NSCLC case should be examined for the presence of EML4-ALK, with a sensitive and accurate diagnostic strategy for the oncogenic fusion being essential for the adoption of ALK inhibitors in the clinic. Given the reliable detection of EML4-ALK mRNA by multiplex RT-PCR demonstrated in the present study, we propose the following scheme for the comprehensive diagnosis of EML4-ALK–positive NSCLC. For sputum, bronchial lavage fluid, pleural effusion, or other specimens that may not be suitable for the preparation of FFPE tissue, multiplex RT-PCR should be applied to detect ALK fusion mRNAs. In contrast, given that FFPE specimens usually have fragmented RNA, they should be subjected to FISH and to sensitive IHC analysis such as that described previously (14, 15). Furthermore, FISH or IHC can be applied to cell blocks prepared from some non-FFPE specimens. No single technique is therefore able to detect EML4-ALK in all types of specimen, and appropriate tests should be chosen on the basis of the specimens available for a given patient.
Notes

Supplementary data for this article are available at Clinical Cancer Research Online (http://clincancerres.aacrjournals.org/). The nucleotide sequence of the novel $EML4-ALK$ variant cDNA from patient J-#189 has been deposited in the DDBJ/EMBL/GenBank databases under the accession number AB663645.

Grant Support

This study was supported in part by a grant for Research on Human Genome Tailor-made from the Ministry of Health, Labor, and Welfare of Japan, by Grants-in-Aid for Scientific Research from the Ministry of Education, Culture, Sports, Science, and Technology of Japan, and by grants from the Japan Society for the Promotion of Science, from Takeda Science Foundation, from Mochida Memorial Foundation for Medical and Pharmaceutical Research, from The Mitsubishi Foundation, and from The Sagawa Foundation for Promotion of Cancer Research.
References


cancers are characterized by rare other mutations, a TTF-1 cell lineage, an acinar histology, and young onset. Mod Pathol 2009;22:508-15.


Figure Legends

Figure 1. Multiplex RT-PCR system for detection of EML4-ALK and KIF5B-ALK.

(A) Schematic representation of the structure of EML4, KIF5B, and ALK proteins. The positions of exons (E for EML4 and K for KIF5B) theoretically capable of fusing in-frame to exon 20 (A20) of ALK are indicated by arrows. The positions of eight forward primers (F-1 to F-8) and one reverse primer (R-1) for PCR are also indicated below the corresponding proteins. EML4 contains a coiled-coil domain (CC), a hydrophobic EMAP-like protein domain (HELP), and WD repeats (WD). KIF5B consists of an amino-terminal ATP-dependent motor domain, a neck region, and a stalk region. (B) Various numbers (0 to 1 × 10⁶) of EML4-ALK (E13;A20)–positive BA/F3 cells (1) were mixed with a fixed number (1 × 10⁶) of EML4-ALK–negative BA/F3 cells, and each mixture was analyzed with our multiplex RT-PCR system. A cDNA for mouse Gapdh was also amplified by PCR as an internal control with the primers 5’-TGTGTCCGTCGTGGATCTGA-3’ and 5’-CCTGCTTCACCACCTTCTTGA-3’.

Figure 2. Multiplex RT-PCR detection of EML4-ALK–positive NSCLC.

(A) RT-PCR products for each of the EML4-ALK variants identified in our cohort were separated by agarose gel electrophoresis. RT-PCR products spanning the EML4-ALK fusion points are indicated by arrows; the asterisk indicates a nonspecific product. A NSCLC cell line, H2228, harboring the E6a/b;A20 variant of EML4-ALK was used as a positive control for the PCR reaction. Size markers include a 50-bp DNA ladder (Invitrogen). NTC, no-template control. (B) Genomic structure of the fusion point for a novel variant of EML4-ALK. Nucleotide sequencing of the genomic PCR and RT-PCR products from patient J-#189 revealed that exon 14 of EML4 (blue) was spliced to a TT sequence adjacent to the genomic ligation point, with transcription continuing in an in-frame manner into intron 19 and exon 20 of ALK (red).

Figure 3. Specimens positive for EML4-ALK by RT-PCR but negative by iAEP-based IHC and by FISH.
Sections of tumor biopsy specimens for J-#393 tumor (A) and J-#927 (B) were stained with hematoxylin-eosin (HE), subjected to IHC analysis by the iAEP method, and examined by split or fusion FISH. The color of fluorescence for the probes in each hybridization is indicated below the FISH images. Nuclei are stained blue with 4’,6-diamidino-2-phenylindole.
Figure 1B

EML4-ALK-positive cells among 10^6 negative cells

No template

EML4-ALK

Gapdh

0 1 10 10^2 10^3 10^4 10^5 10^6
**Figure 2**

**A**

![Image of gel electrophoresis showing size markers and specific samples like NTC, J-#817 (E13;A20), J-#4 (E20;A20), J-#646 (E6a/b;A20), J-#848 (E18;A20), J-#189 (E14::ins2;ins56A20), and H2228.]

**B**

![Diagram of genomic region showing EML4 and ALK genes with specific sequences and distances.]

Size markers:
- NTC
- J-#817 (E13;A20)
- J-#4 (E20;A20)
- J-#646 (E6a/b;A20)
- J-#848 (E18;A20)
- J-#189 (E14::ins2;ins56A20)
- H2228

Size markers.
Figure 3A

HE

iAEP

Fusion FISH

EML4     ALK
Figure 3B

HE

iAEP

Split FISH

Fusion FISH

5′-ALK  3′-ALK

EML4  ALK
Table 1. Characteristics of subjects positive for EML4-ALK by the RT-PCR diagnostic system.

<table>
<thead>
<tr>
<th>Identification number</th>
<th>Sex</th>
<th>Age (years)</th>
<th>Pathological classification</th>
<th>Specimen type</th>
<th>EML4-ALK variant</th>
<th>Smoking history (pack-years)</th>
<th>TNM classification</th>
<th>Clinical stage</th>
<th>iAEP</th>
<th>EGFR mutation</th>
<th>KRAS mutation</th>
</tr>
</thead>
<tbody>
<tr>
<td>J-#1</td>
<td>M</td>
<td>27</td>
<td>Adenocarcinoma</td>
<td>Sputum (two different time points)</td>
<td>E13;A20</td>
<td>0</td>
<td>cT4N3M1</td>
<td>4</td>
<td>+</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>J-#4</td>
<td>F</td>
<td>39</td>
<td>Adenocarcinoma</td>
<td>Metastatic lymph node</td>
<td>E20;A20</td>
<td>NA</td>
<td>cTxN3M1</td>
<td>4</td>
<td>+</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>J-#7</td>
<td>M</td>
<td>74</td>
<td>Adenocarcinoma</td>
<td>Bronchial washing fluid</td>
<td>E13;A20</td>
<td>50</td>
<td>cT3N3M1</td>
<td>4</td>
<td>ND</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>J-#12</td>
<td>F</td>
<td>36</td>
<td>Adenocarcinoma</td>
<td>Resected tumor</td>
<td>E13;A20</td>
<td>0</td>
<td>cT1N0M0</td>
<td>1A</td>
<td>+</td>
<td>–</td>
<td>ND</td>
</tr>
<tr>
<td>J-#43</td>
<td>M</td>
<td>48</td>
<td>Adenocarcinoma</td>
<td>Tumor biopsy/sputum</td>
<td>E13;A20</td>
<td>0</td>
<td>cT3N2M1</td>
<td>4</td>
<td>+</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>J-#88</td>
<td>F</td>
<td>37</td>
<td>Adenocarcinoma</td>
<td>Pleural effusion</td>
<td>E13;A20</td>
<td>0</td>
<td>cT4N3M1</td>
<td>4</td>
<td>ND</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>J-#127</td>
<td>F</td>
<td>49</td>
<td>Adenocarcinoma</td>
<td>Tumor biopsy</td>
<td>E6a,b,A20</td>
<td>0.9</td>
<td>cT1N2M1</td>
<td>4</td>
<td>+</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>J-#189</td>
<td>F</td>
<td>37</td>
<td>Adenocarcinoma</td>
<td>Resected tumor</td>
<td>E14::ins2,ins56A20</td>
<td>0</td>
<td>cT2N1M1</td>
<td>4</td>
<td>+</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>J-#210</td>
<td>F</td>
<td>37</td>
<td>Adenocarcinoma</td>
<td>Resected tumor</td>
<td>E13;A20</td>
<td>0</td>
<td>cT4N2M1</td>
<td>4</td>
<td>ND</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>J-#215</td>
<td>F</td>
<td>61</td>
<td>Adenocarcinoma</td>
<td>Sputum</td>
<td>E13;A20</td>
<td>82</td>
<td>cT4N2M1</td>
<td>4</td>
<td>ND</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>J-#330</td>
<td>M</td>
<td>72</td>
<td>Adenocarcinoma</td>
<td>Pleural effusion/resected tumor (two different regions)</td>
<td>E13;A20</td>
<td>0</td>
<td>cT4N1M1</td>
<td>4</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>J-#350</td>
<td>F</td>
<td>53</td>
<td>Adenocarcinoma</td>
<td>Pleural effusion</td>
<td>E13;A20</td>
<td>0</td>
<td>cT4N2M0</td>
<td>3B</td>
<td>ND</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>J-#378</td>
<td>F</td>
<td>78</td>
<td>Adenocarcinoma</td>
<td>Resected tumor</td>
<td>E13;A20</td>
<td>0</td>
<td>cT1N0M0</td>
<td>1A</td>
<td>ND</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>J-#385</td>
<td>F</td>
<td>80</td>
<td>Adenocarcinoma</td>
<td>Pleural effusion</td>
<td>E6a,b,A20</td>
<td>0</td>
<td>cT4N3M1</td>
<td>4</td>
<td>ND</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>J-#391</td>
<td>F</td>
<td>55</td>
<td>Adenocarcinoma</td>
<td>Tumor biopsy</td>
<td>E13;A20</td>
<td>16.5</td>
<td>cT2N2M1</td>
<td>4</td>
<td>+</td>
<td>–</td>
<td>ND</td>
</tr>
<tr>
<td>J-#392</td>
<td>F</td>
<td>38</td>
<td>Adenocarcinoma</td>
<td>Tumor biopsy</td>
<td>E13;A20</td>
<td>34</td>
<td>cT4N2M0</td>
<td>3B</td>
<td>+</td>
<td>–</td>
<td>ND</td>
</tr>
<tr>
<td>J-#393</td>
<td>F</td>
<td>42</td>
<td>Adenocarcinoma</td>
<td>Tumor biopsy</td>
<td>E13;A20</td>
<td>0</td>
<td>cT4N3M1</td>
<td>4</td>
<td>–</td>
<td>–</td>
<td>ND</td>
</tr>
<tr>
<td>J-#409</td>
<td>F</td>
<td>35</td>
<td>Adenocarcinoma</td>
<td>Tumor biopsy</td>
<td>E13;A20</td>
<td>0</td>
<td>cT4N0M0</td>
<td>3B</td>
<td>+</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>J-#422</td>
<td>M</td>
<td>69</td>
<td>Adenocarcinoma</td>
<td>Tumor biopsy</td>
<td>E6a,b,A20</td>
<td>0</td>
<td>cT2N2M0</td>
<td>3A</td>
<td>ND</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>J-#450</td>
<td>F</td>
<td>30</td>
<td>Adenocarcinoma</td>
<td>Bronchial washing fluid</td>
<td>E6a,b,A20</td>
<td>0</td>
<td>cT4N2M1</td>
<td>4</td>
<td>+</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>J-#530</td>
<td>F</td>
<td>35</td>
<td>Adenocarcinoma</td>
<td>Bronchial washing fluid</td>
<td>E13;A20</td>
<td>0</td>
<td>cT1N0M0</td>
<td>1A</td>
<td>ND</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>J-#646</td>
<td>F</td>
<td>36</td>
<td>Adenocarcinoma</td>
<td>Bronchial washing fluid</td>
<td>E6a,b,A20</td>
<td>0</td>
<td>cT2N3M0</td>
<td>3B</td>
<td>ND</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>J-#657</td>
<td>F</td>
<td>62</td>
<td>Adenocarcinoma</td>
<td>Bronchial washing fluid</td>
<td>E13;A20</td>
<td>15</td>
<td>cT4N2M1</td>
<td>3B</td>
<td>ND</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>J-#759</td>
<td>F</td>
<td>32</td>
<td>Adenocarcinoma</td>
<td>Resected tumor</td>
<td>E13;A20</td>
<td>12</td>
<td>cT1N0M0</td>
<td>1A</td>
<td>ND</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>J-#771</td>
<td>M</td>
<td>32</td>
<td>Adenocarcinoma</td>
<td>Tumor biopsy</td>
<td>E6a,b,A20</td>
<td>15</td>
<td>cT3N3M1</td>
<td>3B</td>
<td>ND</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>J-#817</td>
<td>M</td>
<td>33</td>
<td>Adenocarcinoma</td>
<td>Pleural effusion</td>
<td>E13;A20</td>
<td>0</td>
<td>cT2N1M1</td>
<td>4</td>
<td>ND</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>J-#848</td>
<td>M</td>
<td>57</td>
<td>Adenocarcinoma</td>
<td>Bronchial washing fluid</td>
<td>E18;E20</td>
<td>0</td>
<td>cT4N2M0</td>
<td>3B</td>
<td>ND</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>J-#887</td>
<td>F</td>
<td>32</td>
<td>Adenocarcinoma</td>
<td>Bronchial washing fluid</td>
<td>E6a,b,A20</td>
<td>0</td>
<td>cT3N3M1</td>
<td>4</td>
<td>ND</td>
<td>–</td>
<td>ND</td>
</tr>
<tr>
<td>J-#927</td>
<td>M</td>
<td>36</td>
<td>Adenocarcinoma</td>
<td>Bronchial washing fluid</td>
<td>E6a,b,A20</td>
<td>30</td>
<td>cT4N3M1</td>
<td>4</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>J-#928</td>
<td>F</td>
<td>71</td>
<td>Adenocarcinoma</td>
<td>Bronchial washing fluid</td>
<td>E6a,b,A20</td>
<td>0</td>
<td>cT4N3M1</td>
<td>4</td>
<td>ND</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>J-#996</td>
<td>M</td>
<td>52</td>
<td>Adenocarcinoma</td>
<td>Bronchial washing fluid</td>
<td>E6a,b,A20</td>
<td>0</td>
<td>cT3N3M0</td>
<td>3B</td>
<td>ND</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>J-#1001</td>
<td>F</td>
<td>32</td>
<td>Adenocinoma</td>
<td>Bronchial washing fluid</td>
<td>E13;A20</td>
<td>6.5</td>
<td>cT2N2M1</td>
<td>4</td>
<td>+</td>
<td>–</td>
<td>–</td>
</tr>
</tbody>
</table>

M, male; F, female; ND, not determined; NA, not available.
A prospective PCR-based screening for the EML4-ALK oncogene in non-small cell lung cancer

Manabu Soda, Kazutoshi Isobe, Akira Inoue, et al.

Clin Cancer Res  Published OnlineFirst August 20, 2012.

Updated version  Access the most recent version of this article at: doi:10.1158/1078-0432.CCR-11-2947

Supplementary Material  Access the most recent supplemental material at: http://clincancerres.aacrjournals.org/content/suppl/2012/08/20/1078-0432.CCR-11-2947.DC1

Author Manuscript  Author manuscripts have been peer reviewed and accepted for publication but have not yet been edited.

E-mail alerts  Sign up to receive free email-alerts related to this article or journal.

Reprints and Subscriptions  To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org.

Permissions  To request permission to re-use all or part of this article, contact the AACR Publications Department at permissions@aacr.org.