

長崎大学

宮崎 泰司 業績リスト

発表者名	論文タイトル名	発表誌名	巻号	ページ	出版年
Tominaga-Sato S, Tsushima H, Ando K, Itonaga H, Imaizumi Y, Imanichi D, Iwanaga M, Taguchi J, Fukushima T, Yoshida S, Hata T, Moriuchi Y, Kuriyama K, Mano H, Tomonaga M, Miyazaki Y	Expression of myeloperoxidase and gene mutations in AML patients with normal karyotype: double CEBPA mutations are associated with high percentage of MPO positivity in leukemic blasts	Int J Hematol	94	81-89	2011
Usui N, Takeshita A, Nakaseko C, Dobashi N, Fujita H, Kiyoi H, Kobayashi Y, Sakura T, Yahagi Y, Shigeno K, Ohwada C, Miyazaki Y, Ohtake S, Miyawaki S, Naoe T, Ohnishi K for the Japan Adult Leukemia Study Group	Phase I trial of gemtuzumab ozogamicin in intensive combination chemotherapy for relapsed or refractory adult acute myeloid leukemia (AML): Japan Adult Leukemia Study Group (JALSG)-AML206 study	Cancer Sci	102	1358-1365	2011
Miyawaki S, Ohtake S, Fujisawa S, Kiyoi H, Shinagawa K, Usui N, Sakura T, Miyamura K, Nakaseko C, Miyazaki Y, Fujieda A, Nagai T, Yamane T, Taniwaki M, Takahashi M, Yagasaki F, Kimura Y, Asou N, Sakamaki H, Handa H, Honda S, Ohnishi K, Naoe T, Ohno R	A randomized comparison of 4 courses of standard-dose multiagent chemotherapy versus 3 courses of high-dose cytarabine alone in postremission therapy for acute myeloid leukemia in adults: the JALSG AML201 Study	blood	117	2366-2372	2011
Ohtake S, Miyawaki S, Fujita H, Kiyoi H, Shinagawa K, Usui N, Okumura H, Miyamura K, Nakaseko C, Miyazaki Y, Fujieda A, Nagai T, Yamane T, Taniwaki M, Takahashi M, Yagasaki F, Kimura Y, Asou N, Sakamaki H, Handa H, Honda S, Ohnishi K, Naoe T, Ohno R	Randomized study of induction therapy comparing standard-dose idarubicin with high-dose daunorubicin in adult patients with previously untreated acute myeloid leukemia: the JALSG AML201 Study	blood	117	2358-2365	2011
Kako S, Morita S, Sakamaki H, Ogawa H, Fukuda T, Takahashi S, Kanamori H, Onizuka M, Iwato K, Suzuki R, Atsuta Y, Kyo T, Sakura T, Jinnai I, Takeuchi J, Miyazaki Y, Miyawaki S, Ohnishi K, Naoe T, Kaneda Y	A decision analysis of allogeneic hematopoietic stem cell transplantation in adult patients with Philadelphia chromosome-negative acute lymphoblastic leukemia in first remission who have an HLA-matched sibling donor	Leukemia	25	259-265	2011
Ando K, Miyazaki Y, Sawayama Y, Tominaga S, Matsuo E, Yamasaki R, Inoue Y, Iwanaga M, Imanishi D, Tsushima H, Fukushima T, Imaizumi Y, Taguchi J, Yoshida S, Hata T, Tomonaga M	High expression of 67-kDa laminin receptor relates to the proliferation of leukemia cells and increases expression of GM-CSF receptor	Experimental Hematology	39	179-186	2011

Iwanaga M, Hsu WL, Soda M, Takasaki Y, Tawara M, Joh T, Amenomori T, Yamamura M, Yoshida Y, Koba T, Miyazaki Y, Matsuo T, Preston D L, Suyama A, Kodama K, Tomonaga M	Risk of Myelodysplastic Syndromes in People Exposed to Ionizing Radiation: A Retrospective Cohort Study of Nagasaki Atomic Bomb Survivors	JOURNAL OF CLINICAL ONCOLOGY	29	428-434	2011
Mizuta S, Matsuo K, Yagasaki F, Yujiri T, Hatta Y, Kimura Y, Ueda Y, Kanamori H, Usui N, Akiyama H, Miyazaki Y, Ohtake S, Atsuta Y, Sakamaki H, Kawa K, Morishima Y, Ohnishi K, Naoe T, Ohno R	Pre-transplant imatinib-based therapy improves the outcome of allogeneic hematopoietic stem cell transplantation for BCR-ABL-positive acute lymphoblastic leukemia	Leukemia	25	41-47	2011

High-throughput resequencing of target-captured cDNA in cancer cells

Toshihide Ueno,^{1,5} Yoshihiro Yamashita,^{1,5} Manabu Soda,¹ Kazutaka Fukumura,² Mizuo Ando,² Azusa Yamato,¹ Masahito Kawazu,² Young Lim Choi^{1,2} and Hiroyuki Mano^{1,2,3,4}

¹Division of Functional Genomics, Jichi Medical University, Tochigi; ²Department of Medical Genomics, Graduate School of Medicine, University of Tokyo, Tokyo; ³CREST Japan Science and Technology Agency, Saitama, Japan

(Received May 30, 2011/Revised September 7, 2011/Accepted September 14, 2011/Accepted manuscript online September 20, 2011/Article first published online October 13, 2011)

The recent advent of whole exon (exome)-capture technology, coupled with second-generation sequencers, has made it possible to readily detect genomic alterations that affect encoded proteins in cancer cells. Such target resequencing of the cancer genome, however, fails to detect most clinically-relevant gene fusions, given that such oncogenic fusion genes are often generated through intron-to-intron ligation. To develop a resequencing platform that simultaneously captures point mutations, insertions-deletions (indels), and gene fusions in the cancer genome, we chose cDNA as the input for target capture and extensive resequencing, and we describe the versatility of such a cDNA-capture system. As a test case, we constructed a custom target-capture system for 913 cancer-related genes, and we purified cDNA fragments for the target gene set from five cell lines of CML. Our target gene set included Abelson murine leukemia viral oncogene homolog 1 (*ABL1*), but it did not include breakpoint cluster region (*BCR*); however, the sequence output faithfully detected reads spanning the fusion points of these two genes in all cell lines, confirming the ability of cDNA capture to detect gene fusions. Furthermore, computational analysis of the sequence dataset successfully identified non-synonymous mutations and indels, including those of tumor protein p53 (*TP53*). Our data might thus support the feasibility of a cDNA-capture system coupled with massively parallel sequencing as a simple platform for the detection of a variety of anomalies in protein-coding genes among hundreds of cancer specimens. (*Cancer Sci* 2012; 103: 131–135)

Cancer is thought to result from various alterations of the genome, including point mutations, insertions-deletions (indels), and genomic rearrangements.⁽¹⁾ Whereas comprehensive sequencing of the cancer genome, or “cancer genome resequencing”, is a promising approach to the identification of such anomalies, and to provide a basis for the development of effective treatment strategies for cancer, determination of the nucleotide sequence of the entire human genome with conventional Sanger sequencers remains a highly demanding task. However, the recent advent of massively parallel sequencing systems, or second-generation sequencers, has rendered such projects manageable in private laboratories⁽²⁾ and triggered the formation of large-scale consortia, such as The Cancer Genome Atlas and International Cancer Genome Consortium,⁽³⁾ to undertake cancer genome resequencing for hundreds of specimens. Cancer genome resequencing with massively parallel sequencers has already provided a wealth of information on genome-wide mutation status for melanoma,⁽⁴⁾ acute myeloid leukemia,⁽⁵⁾ hepatocellular carcinoma,⁽⁶⁾ and other cancers.

Even with the current massively parallel sequencers, however, the determination and compilation of the full genome sequence for a given sample might still take almost 1 month. Comparison of the cancer genome among many specimens thus remains time-consuming and labor intensive. Anomalies in protein-coding genes likely play a major role in carcinogenesis. Given that

exonic regions occupy only ~1.3% of the human genome, sequencing such targeted regions would be expected to markedly facilitate the discovery of proteins that are activated or inactivated specifically in cancer cells. Indeed, target-capture strategies, coupled with massively parallel sequencers, have revealed important genetic changes in cancer,⁽⁷⁾ as well as in hereditary disorders.^(8,9)

One important drawback of such target-capture approaches, however, is their inability to detect gene fusions. Most cancer-associated gene fusion events occur within introns (resulting in exon-to-exon ligation in the corresponding mRNA), and exon capture does not reveal breakage and ligation of intronic regions. Recurrent gene fusions were once thought to be rare in epithelial tumors compared with hematologic malignancies and sarcomas;⁽¹⁰⁾ however, our recent discovery of the echinoderm microtubule associated protein like-4 (*EML4*)-anaplastic lymphoma kinase (*ALK*) fusion gene in lung cancer and the discovery by others of rearrangements in loci for the v-ets avian erythroblastosis virus E26 oncogene homolog (*ETS*) family of transcription factors in prostate cancer have led to a revision of this notion.^(11,12) It would thus be desirable to develop a resequencing platform that is able to capture, within a reasonable timeframe, all gene fusions, point mutations, and indels in the cancer genome. In pursuit of this goal, we have now examined the efficacy of high-throughput sequencing of captured cDNA for the identification of such cancer genome anomalies.

Materials and Methods

Cell lines. Cell lines established from the blast crisis stage of CML, including MEG-01s, KCL-22-SR, K562, NCO2, and KU812,^(13,14) were obtained from the Japanese Collection of Research Bioresources (Osaka, Japan) and were maintained in RPMI-1640 medium (Invitrogen, Carlsbad, CA, USA) supplemented with 10% FBS (Invitrogen). Total RNA was isolated from each cell line with the use of an RNeasy mini kit (Qiagen, Valencia, CA, USA) and was subjected to cDNA synthesis with an oligo(dT) primer.

Gene expression profiling. The cDNA prepared from total poly(A)-RNA of KCL-22-SR cells was subjected to hybridization with the HGU95Av2 microarray (Affymetrix, Santa Clara, CA, USA), as described previously.⁽¹⁵⁾ The expression intensity of each test gene on the array was normalized by the 50th percentile value.

cDNA-capture methods. RNA probes of 120 bases were designed to cover (with a 60-base overlap) cDNA of 913 human protein-coding genes (Table S1), and were synthesized by Agilent Technologies (Santa Clara, CA, USA). During the design of the probes, the Repeat Masker dataset (<http://www.repeatmasker.org>) was used to remove probes corresponding to

⁴To whom correspondence should be addressed. E-mail: hmano@jichi.ac.jp

⁵These authors contributed equally to this work.

repetitive sequences in the human genome. Hybridization of DNA fragments to the RNA probes was performed according to the protocols recommended for the SureSelect Target Enrichment system (Agilent). We also used the SureSelect Human X Chromosome Demo kit (Agilent) to examine purification efficiency. Purified DNA fragments were then subjected to sequencing with a Genome Analyzer IIx (GAIIx; Illumina, San Diego, CA, USA) for 76 bases from both ends by the paired-end sequencing system.

Computational pipeline. Raw read data were quality filtered on the basis of the presence of the Illumina adaptor sequences and a *Q*-value of ≥ 20 . The resulting read sequences were then subjected to an in-house computational pipeline to identify various mutations (Fig. S1). In brief, read sequences were matched with the Bowtie algorithm⁽¹⁶⁾ to the cDNA sequences of the 913 genes used to construct our custom-made SureSelect system. The matched reads were then examined for the presence of non-synonymous mutations and single nucleotide polymorphisms (SNP) deposited in dbSNP (build 132, <http://www.ncbi.nlm.nih.gov/projects/SNP/index.html>). The remaining reads were further matched to the cDNA sequences with Burrows-Wheeler Aligner (BWA) and Basic Local Alignment Search Tool (BLAST) algorithms to search for indels and multiple mutations.^(17,18) Candidates for non-synonymous mutations were identified only when $\geq 20\%$ of reads correspond to the mutations at positions with ≥ 50 coverage.

For the selection of reads corresponding to possible fusion cDNA, nucleotide sequences of 20 bp were obtained from both ends of each read and were separately matched to RefSeq mRNA (<http://www.ncbi.nlm.nih.gov>), KnownGeneMrna,⁽¹⁹⁾ and the human genome sequence (GRCh37, <http://www.ncbi.nlm.nih.gov/projects/genome/assembly/grc/human/data/?build=37>). Reads were considered to be derived from fusion genes if the ends of a given read matched to different genes within the 913-gene group, or one end matched to a single gene within the 913-gene group and the other end matched to a sequence in RefSeq, KnownGeneMrna, or the human genome sequence that did not correspond to the 913 genes. Candidates for fusion genes were identified only when four or more reads were mapped to possible fusion points.

RT-PCR. To confirm the presence of an alternatively-spliced mixed-lineage leukemia (*MLL*) mRNA, we subjected oligo(dT)-primed cDNA of KU812 cells to PCR with the combination of the F-1 primer (5'-ACCTCGTGGGAGACCTAGAAGTGG-3') and the R primer (5'-AGTCATTGGAAGCTTGTCTGCCTG-3'), or with the combination of the F-2 primer (5'-CCTGTGGGTA-GGGTTTCCAAAGAG-3') and the R primer.

Results

Efficiency of cDNA-capture sequencing. Paired-end sequencing of target-captured cDNA was briefly described in a previous study;⁽²⁰⁾ however, how the efficiency of target purification with cDNA compares with that with genomic DNA remains unclear. We therefore attempted to optimize the conditions for cDNA purification with the SureSelect system. Oligo(dT)-primed cDNA of KCL-22-SR cells were fragmented to a mean size of 500 or 200 bp and then subjected to purification with the use of the SureSelect Human X Chromosome Demo kit, which is designed to capture genomic sequences derived from the human X chromosome. Genomic DNA of KCL-22-SR cells was similarly processed and hybridized with the X Chromosome Demo kit. The purified fragments at either 4 or 8 pM were then sequenced by the GAIIx system.

The X chromosome-mapped cDNA reads occupied 62.1%, 81.6%, 62.4%, and 82.2% of quality filter-passed reads for the experiments with 4 pM of 500-bp fragments, 4 pM of 200-bp fragments, 8 pM of 500-bp fragments, and 8 pM of 200-bp frag-

ments, respectively (Fig. 1). Thus, these results suggested that the shorter cDNA fragments were captured more efficiently than the longer ones. Furthermore, the purification efficiency for genomic DNA fragments was not higher than that for cDNA, irrespective of DNA concentration and fragmentation size (Fig. 1), supporting the feasibility of cDNA-capture approaches.

The ability to detect breakpoint cluster region (*BCR*)-Abelson murine leukemia viral oncogene homolog 1 (*ABL1*) fusion reads was reduced for the cDNA sheared to ~ 200 bp compared with that for those of ~ 500 bp (see below). The former cDNA detected 83.7% or 76% of the fusion reads detected by the latter cDNA at input concentrations of 4 and 8 pM, respectively. This result is in line with our computational bootstrap trial ($n = 10\,000$) showing that the number of randomly-fragmented, 200-bp reads encompassing the *BCR-ABL1* fusion point is ~ 2.5 times higher than that of 500-bp reads (data not shown). However, given that the total number of high-quality reads was much higher in the data for the 200-bp cDNA than in those for the 500-bp cDNA (Fig. 1), we chose to use 8 pM of cDNA with a mean size of 200 bp for further experiments.

Custom cDNA-capture system. We also tested whether extensive sequencing of cDNA generated from total poly(A)-RNA (unselected cDNA) might serve to identify gene fusions, point mutations, and indels. For this purpose, unselected cDNA were prepared from KCL22-SR cells, and subjected to GAIIx sequencing, yielding 34.1 million reads, which mapped to 36 128 RefSeq entries (data not shown). The distribution of read number per transcript in the data is shown in Figure 2a. Among the 36 128 entries, only 200 (0.55%) accounted for $\sim 20\%$ of total reads, and 4.55% accounted for $\sim 50\%$ of reads. Thus, as expected, resequencing data for unselected cDNA consist mostly of reads corresponding to a limited number of highly-abundant transcripts.

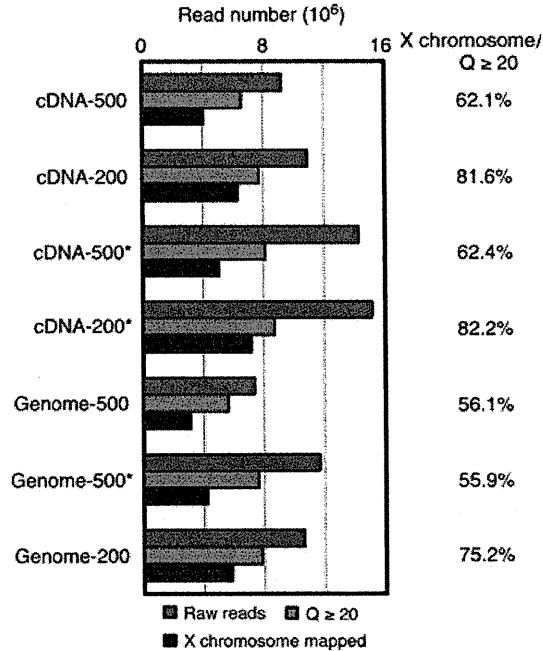


Fig. 1. Comparison of capture efficiency between cDNA and genomic DNA. Genomic DNA or cDNA of KCL-22-SR cells was fragmented to a mean size of 200 or 500 bp, and then subjected to purification with the SureSelect Human X Chromosome Demo kit, followed by GAIIx sequencing at a concentration of 4 or 8 pM (the latter indicated by an asterisk). Numbers of raw reads, reads with a *Q*-value of ≥ 20 ($Q \geq 20$), and reads mapped to the human X chromosome are shown for each experiment. Percentage of X chromosome-mapped reads among the reads with a *Q*-value of ≥ 20 is shown on the right.

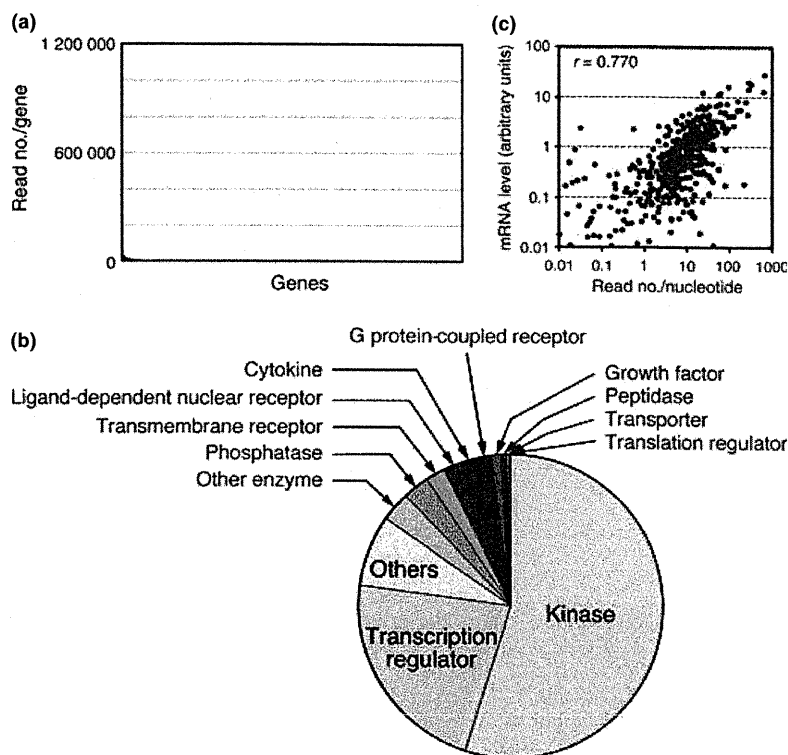


Fig. 2. Capture of a selected set of cDNA. (a) Read number for each gene was calculated from the sequencing data for the unselected cDNA of KCL-22-SR cells. Genes were sorted according to their read number. A small number of genes accounted for most of the sequence reads. (b) Functional annotation for the encoded proteins of our target cDNA ($n = 913$). (c) Read number per nucleotide for each captured cDNA in KCL-22-SR cells is compared with the expression intensity (arbitrary units) of the same cDNA examined with an HGU95Av2 microarray. Pearson's correlation coefficient (r) for the comparison is also demonstrated.

We therefore attempted to construct a custom SureSelect system to capture cDNA for cancer-related genes. For this purpose, we selected 913 genes that yielded 56 892 hybridization probes corresponding to ~ 3.77 Mbp of total capture capacity. The target genes encoded human protein kinases (all members in the human genome), transcription regulators, phosphatases, and other proteins (Fig. 2b; Table S1).

To compare the information provided by the sequence data from unselected and captured cDNA, we purified target cDNA from KCL-22-SR cells with the use of our custom SureSelect system, and determined their nucleotide sequences with GAIIX. A comparable amount of filter-passed reads (39.2 million) to that of unselected cDNA were thus obtained. We found that 88% of the captured cDNA were mapped to the target genes in our SureSelect system, while only 6.6% of the unselected cDNA were mapped to the 913 targets (data not shown). The read number obtained for each gene in the captured cDNA dataset is shown in Figure S2, with the distribution being markedly different from that obtained by sequencing of the unselected cDNA (Fig. 2a). As expected, the read number per nucleotide in each cDNA for the captured dataset was highly correlated to the expression intensity of the same gene quantified with the HGU95Av2 GeneChip expression array (Pearson's correlation coefficient = 0.770, $P < 2.2 \times 10^{-16}$) (Fig. 2c).

We further isolated target cDNA from other CML cell lines, including K562, KU812, MEG-01s, and NCO2, and the purified cDNA fragments were subjected to GAIIX sequencing. As in the case for KCL-22-SR, 86–88% of the obtained reads were successfully mapped to the target cDNA in each cell line (Table S2).

Screening of fusion cDNA. Our target set of 913 genes did not include *BCR*, but it did contain *ABL1*. Thus, if we were able to isolate sequence reads encompassing the fusion point of *BCR-ABL1*, cDNA-capture approaches for a given gene set would likely be able to detect gene fusions to unknown partners. In fact, we detected 45 sequence reads for KCL-22-SR cells that covered the *BCR-ABL1* fusion point (Fig. 3a). Likewise, the sequence datasets for K562, KU812, MEG-01s, and NCO2 cells

contained 53, 8, 11, and 10 such fusion reads, respectively (data not shown). Furthermore, our sequence data faithfully recapitulated two variants of *BCR-ABL1* cDNA in these cell lines; a fusion variant between exon 13 of *BCR* and exon 2 of *ABL1* was detected in KCL-22-SR, MEG-01s, and NCO2 cells, whereas a fusion variant between exon 14 of *BCR* and exon 2 of *ABL1* was detected in K562 and KU812 cells.⁽¹⁴⁾

In addition to *BCR-ABL1*, we identified 72 independent candidates for fusion cDNA (including fusions to non-coding RNA) from the CML cell lines. Surprisingly, however, the screening of fusion genes among the unselected cDNA of KCL-22-SR with our rather non-stringent threshold (≥ 4 reads mapped to a candidate fusion point) failed to isolate *BCR-ABL1* cDNA. We could not even detect any fusion candidates (involving one of our target genes in either or both ends of fusion events) from this dataset, while a total of nine candidates (including *BCR-ABL1*) were isolated from the captured cDNA of the same cell line.

Our Bowtie mapping of both ends of each read to human mRNA or genome databases (Fig. S1) resulted in the detection of not only *BCR-ABL1* fusions, but also a large number of alternatively-spliced messages. From the captured cDNA of KCL-22-SR, for instance, we could detect 79 alternatively-spliced transcripts for 72 independent genes (data not shown). In contrast, from the unselected cDNA of the same cell line, only three independent, alternatively-spliced transcripts were identified among three genes within the 913 targets.

One such example of alternatively-spliced message was *MLL* (ensemble accession no.: ENST00000389506) in KU812, MEG-01s, and K562 cells. In addition to a set of reads that completely matched exon 3 of *MLL*, we obtained reads that lacked an internal 2193-bp sequence in exon 3 (Fig. 3b). Such in-frame truncation would be expected to generate an *MLL* protein lacking amino acids 276–1006 of the wild-type protein. To confirm the presence of such transcripts, we performed RT-PCR analysis with total RNA from KU812 cells, and PCR primers designed as in Figure 3b. The combination of the F-1 and R primers would be expected to yield both the wild-type (2536 bp) and truncated

However, for experiments based on capture of genomic DNA, sequencing a paired normal specimen allows the efficient subtraction of rare SNP not present in the current databases from the dataset of cancer tissue. This is not always the case, however, for the cDNA-capture approach, given that gene expression profiles differ markedly among samples (even among those obtained from the same individual). Genes with sequence alterations in the cancer specimen might not be expressed in a given normal specimen, and it is not possible to readily determine whether such alterations are germ-line polymorphisms, while algorithms to predict the effect on protein functions for a given amino acid change are currently available⁽²³⁾ and synonymous-to-non-synonymous ratio of nucleotide alterations for a given gene/dataset might provide clues as to how such changes are selected in tumor cells.⁽²⁴⁾

In addition, the cDNA-capture system cannot obtain a sufficient number of reads for genes expressed at a low level, and the overall sensitivity of cDNA capture is dependent on the total read number provided by sequencers. We are able to run only two samples per flow cell of the GAIIX system, whereas up to eight samples can be run in a single flow cell for whole exome sequencing of human genomic DNA.

Despite such limitations, our study shows that cDNA capture is an efficient process, and extensive sequencing of such purified

cDNA is a straightforward approach to interrogate the target cDNA for various genetic changes in a single platform. Large-scale resequencing of hundreds of cancer specimens might thus become within the scope of private laboratories with the adoption of the cDNA-capture approach.

Acknowledgments

This study was supported in part by grants for Research on Human Genome Tailor-Made and for Third-Term Comprehensive Control Research for Cancer from the Ministry of Health, Labor, and Welfare of Japan; Grants-in-Aid for Scientific Research (B) and for Young Scientists (A) from the Ministry of Education, Culture, Sports, Science, and Technology of Japan; and by grants from the Japan Society for the Promotion of Science, Takeda Science Foundation, the Naito Foundation, Sankyo Foundation of Life Science, The Sagawa Foundation for Promotion of Cancer Research, the Yasuda Medical Foundation, the Mitsubishi Foundation, and Kobayashi Foundation for Cancer Research.

Disclosure Statement

K. Fukumura, M. Ando, M. Kawazu, Y.L. Choi and H. Mano belong to the Department of Medical Genomics, Graduate School of Medicine, University of Tokyo, which receives research funding from Illumina Inc.

References

- Stratton MR, Campbell PJ, Futreal PA. The cancer genome. *Nature* 2009; **458**: 719–24.
- Mardis ER. A decade's perspective on DNA sequencing technology. *Nature* 2011; **470**: 198–203.
- Ledford H. Big science: the cancer genome challenge. *Nature* 2010; **464**: 972–4.
- Pleasant ED, Cheetham RK, Stephens PJ *et al*. A comprehensive catalogue of somatic mutations from a human cancer genome. *Nature* 2010; **463**: 191–6.
- Ley TJ, Mardis ER, Ding L *et al*. DNA sequencing of a cytogenetically normal acute myeloid leukaemia genome. *Nature* 2008; **456**: 66–72.
- Totoki Y, Tatsuno K, Yamamoto S *et al*. High-resolution characterization of a hepatocellular carcinoma genome. *Nat Genet* 2011; **43**: 464–9.
- Wei X, Walia V, Lin JC *et al*. Exome sequencing identifies GRIN2A as frequently mutated in melanoma. *Nat Genet* 2011; **43**: 442–6.
- Otto EA, Hurd TW, Airik R *et al*. Candidate exome capture identifies mutation of SDCCAG8 as the cause of a retinal-renal ciliopathy. *Nat Genet* 2010; **42**: 840–50.
- Bilguvar K, Ozturk AK, Louvi A *et al*. Whole-exome sequencing identifies recessive WDR62 mutations in severe brain malformations. *Nature* 2010; **467**: 207–10.
- Mitelman F. Recurrent chromosome aberrations in cancer. *Mutat Res* 2000; **462**: 247–53.
- Soda M, Choi YL, Enomoto M *et al*. Identification of the transforming *EMLA-ALK* fusion gene in non-small-cell lung cancer. *Nature* 2007; **448**: 561–6.
- Tomkins SA, Rhodes DR, Perner S *et al*. Recurrent fusion of *TPMRSS2* and *ETS* transcription factor genes in prostate cancer. *Science* 2005; **310**: 644–8.
- Ohmine K, Nagai T, Tarumoto T *et al*. Analysis of gene expression profiles in an imatinib-resistant cell line, KCL22/SR. *Stem Cells* 2003; **21**: 315–21.
- Drexler HG, MacLeod RA, Uphoff CC. Leukemia cell lines: *in vitro* models for the study of Philadelphia chromosome-positive leukemia. *Leuk Res* 1999; **23**: 207–15.
- Choi YL, Tsukasaki K, O'Neill MC *et al*. A genomic analysis of adult T-cell leukemia. *Oncogene* 2007; **26**: 1245–55.
- Langmead B, Trapnell C, Pop M, Salzberg SL. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol* 2009; **10**: R25.
- Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. *J Mol Biol* 1990; **215**: 403–10.
- Li H, Durbin R. Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics* 2009; **25**: 1754–60.
- Fujita PA, Rhead B, Zweig AS *et al*. The UCSC Genome Browser database: update 2011. *Nucleic Acids Res* 2011; **39**: D876–82.
- Levin JZ, Berger MF, Adiconis X *et al*. Targeted next-generation sequencing of a cancer transcriptome enhances detection of sequence variants and novel fusion transcripts. *Genome Biol* 2009; **10**: R115.
- Calabretta B, Perrotti D. The biology of CML blast crisis. *Blood* 2004; **103**: 4010–22.
- Shen P, Wang W, Krishnakumar S *et al*. High-quality DNA sequence capture of 524 disease candidate genes. *Proc Natl Acad Sci USA* 2011; **108**: 6549–54.
- Adzhubei IA, Schmidt S, Peshkin L *et al*. A method and server for predicting damaging missense mutations. *Nat Methods* 2010; **7**: 248–9.
- Babenko VN, Basu MK, Kondrashov FA, Rogozin IB, Koonin EV. Signs of positive selection of somatic mutations in human cancers detected by EST sequence analysis. *BMC Cancer* 2006; **6**: 36.

Supporting Information

Additional Supporting Information may be found in the online version of this article:

Fig. S1. Algorithm of the computational pipeline.

Fig. S2. Read number distribution of all poly(A)-RNA data.

Fig. S3. Numbers of 1-, 2-, or 3-bp indels for the entire dataset.

Fig. S4. A CAG-deletion in the *TP53* message in MEG-01s cells.

Table S1. Gene list for the custom cDNA-capture system.

Table S2. Purification of the target cDNA in CML cell lines.

Table S3. *TP53* mutation status in CML cell lines.

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.

KLC1-ALK: A Novel Fusion in Lung Cancer Identified Using a Formalin-Fixed Paraffin-Embedded Tissue Only

Yuki Togashi^{1,2}, Manabu Soda³, Seiji Sakata¹, Emiko Sugawara^{1,4}, Satoko Hatano^{1,2}, Reimi Asaka^{1,2}, Takashi Nakajima⁵, Hiroyuki Mano^{3,6}, Kengo Takeuchi^{1,2*}

1 Pathology Project for Molecular Targets, The Cancer Institute, Japanese Foundation for Cancer Research, Tokyo, Japan, **2** Division of Pathology, The Cancer Institute, Japanese Foundation for Cancer Research, Tokyo, Japan, **3** Division of Functional Genomics, Jichi Medical University, Tochigi, Japan, **4** Department of Comprehensive Pathology, Graduate School, Tokyo Medical and Dental University, Tokyo, Japan, **5** Division of Diagnostic Pathology, Shizuoka Cancer Center, Nagaizumi, Shizuoka, Japan, **6** Department of Medical Genomics, Graduate School of Medicine, University of Tokyo, Tokyo, Japan

Abstract

The promising results of anaplastic lymphoma kinase (ALK) inhibitors have changed the significance of ALK fusions in several types of cancer. These fusions are no longer mere research targets or diagnostic markers, but they are now directly linked to the therapeutic benefit of patients. However, most available tumor tissues in clinical settings are formalin-fixed and paraffin-embedded (FFPE), and this significantly limits detailed genetic studies in many clinical cases. Although recent technical improvements have allowed the analysis of some known mutations in FFPE tissues, identifying unknown fusion genes by using only FFPE tissues remains difficult. We developed a 5'-rapid amplification of cDNA ends-based system optimized for FFPE tissues and evaluated this system on a lung cancer tissue with ALK rearrangement and without the 2 known ALK fusions EML4-ALK and KIF5B-ALK. With this system, we successfully identified a novel ALK fusion, KLC1-ALK. The result was confirmed by reverse transcription-polymerase chain reaction and fluorescence *in situ* hybridization. Then, we synthesized the putative full-length cDNA of KLC1-ALK and demonstrated the transforming potential of the fusion kinase with assays using mouse 3T3 cells. To the best of our knowledge, KLC1-ALK is the first novel oncogenic fusion identified using only FFPE tissues. This finding will broaden the potential value of archival FFPE tissues and provide further biological and clinical insights into ALK-positive lung cancer.

Citation: Togashi Y, Soda M, Sakata S, Sugawara E, Hatano S, et al. (2012) KLC1-ALK: A Novel Fusion in Lung Cancer Identified Using a Formalin-Fixed Paraffin-Embedded Tissue Only. PLoS ONE 7(2): e31323. doi:10.1371/journal.pone.0031323

Editor: Anthony W.I. Lo, The Chinese University of Hong Kong, Hong Kong

Received: October 17, 2011; **Accepted:** January 5, 2012; **Published:** February 8, 2012

Copyright: © 2012 Togashi et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: This work was supported in part by Grants-in-Aid for Scientific Research from the Ministry of Education, Culture, Sports, Science, and Technology of Japan as well as by grants from the Japan Society for the Promotion of Science; the Ministry of Health, Labor, and Welfare of Japan; the Vehicle Racing Commemorative Foundation of Japan; the Princess Takamatsu Cancer Research Fund; and the Uehara Memorial Foundation. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: kentakeuchi-tyk@umin.net

Introduction

Anaplastic lymphoma kinase (ALK) is a receptor tyrosine kinase that was discovered in anaplastic large-cell lymphoma (ALCL) in the form of a fusion protein, NPM-ALK [1,2]. The formation of a fusion protein with a partner through chromosomal translocations is the most common mechanism of ALK overexpression and ALK kinase domain activation. Recent promising results of clinical trials with an ALK inhibitor, crizotinib, have changed the significance of ALK fusions in lung cancer [3,4,5,6], inflammatory myofibroblastic tumors (IMTs) [7], and ALCL [8]. ALK fusions are no longer mere research targets or diagnostic markers and are now directly linked to the therapeutic benefit of patients.

In lung cancer, 3 fusion partners of ALK have been reported—EML4, TFG, and KIF5B—although the presence of TFG-ALK in lung cancer has not yet been proven with histopathological evidence [9,10,11]. In addition to lung cancer, ALK has further been found to generate fusions in ALCL (fused to NPM, TPM3, TPM4, ATIC, TFG, CLTC, MSN, MYH9, or ALO17) [1,2,12,13,14,15,16,17,18,19], IMT (TPM3, TPM4, CLTC, CARS, RANBP2, ATIC, or SEC31A) [19,20,21,22,23,24], ALK-positive large B-cell lymphoma (CLTC, NPM, SEC31A,

or SQSTM1) [25,26,27,28], and renal cancer (VCL, TPM3 or EML4) (Table 1) [29,30]. In addition to TFG-ALK in lung cancer, some ALK fusions have been reported without histopathological evidence: TPM4-ALK in esophageal squamous cell carcinoma [31,32] and EML4-ALK in colon and breast carcinomas [33].

Anti-ALK immunohistochemistry played an important role in identifying these ALK fusion partners. Several ALK fusions exhibit a characteristic staining pattern in anti-ALK immunohistochemistry because the subcellular localization of ALK fusion proteins depends on the fusion partner. For example, NPM-ALK, which is the most common fusion in ALK-positive ALCL (85%), exhibits a nuclear and cytoplasmic staining pattern because the heterodimer of NPM and NPM-ALK localizes in the nucleus and the homodimer of NPM-ALK in the cytoplasm; CLTC-ALK exhibits a cytoplasmic granular pattern because it localizes in the small vesicles. If a tumor exhibits an unrecognized anti-ALK staining pattern, the patient may have a novel fusion partner. In addition to the difference in subcellular localization, the difference in staining intensity is a key to identifying novel partners. EML4-ALK is hardly stained by conventional anti-ALK immunohistochemistry [11,34]. To overcome this limitation, we developed the intercalated antibody-enhanced polymer (iAEP) method, which moderately increases

Table 1. ALK fusion partners.

Reported year	Partner	Locus	ALK+ALCL	ALK+LBCL	IMT	NSCLC	RCC
1994	NPM	5q35.1	+	+			
1999	TPM3	1p23	+		+		+
1999	TFG	3q12.2	+			+	
2000	ATIC	2q35	+		+		
2000	TPM4	19p13	+		+		
2001	CLTC	17q23	+	+	+		
2001	MSN	Xp11.1	+				
2002	ALO17	17q25.3	+				
2003	MYH9	22q13.1	+				
2003	RANBP2	2q13			+		
2003	CARS	11p15			+		
2006	SEC31A	4q41		+	+		
2007	EML4	2p21					+
2009	KIF5B	10p11.22				+	
2011	SQSTM1	5q35.3		+			
2011	PPFIBP1	12p11			+		
2011	VCL	10q22.2					+
Present study	KLC1	14q32.1				+	

*Histopathological evidence is lacking. Abbreviations: ALCL, anaplastic large cell lymphoma; LBCL, large B-cell lymphoma; IMT, inflammatory myofibroblastic tumor; NSCLC, non-small cell lung carcinoma; RCC, renal cell carcinoma.
doi:10.1371/journal.pone.0031323.t001

sensitivity in the immunohistochemical detection system, and EML4-ALK was consistently stained with this method [11]. This indicated that a tumor that is positively immunostained for ALK only by a sensitive immunohistochemistry method but not by conventional methods may harbor a novel ALK fusion. Based on this hypothesis, we successfully identified PPFIBP1-ALK in 2 IMT cases that were positive in anti-ALK immunohistochemistry only when stained by the iAEP method [35].

Anti-ALK immunohistochemistry may thus be useful to detect candidate tumors for a novel ALK fusion. However, to identify the fusion partner, other molecular techniques are usually required such as 5'-rapid amplification of cDNA ends (5'-RACE) or inverse reverse-transcription polymerase chain reaction (RT-PCR). To the best of our knowledge, no novel oncogenic fusions have been discovered using formalin-fixed paraffin-embedded (FFPE) tissues only because nucleic acids extracted from FFPE tissues are severely degraded during the fixation process. In the present study, we developed a 5'-RACE method optimized for *ALK* fusion partner detection that was applicable to FFPE tissues and identified a novel fusion, kinesin light chain 1 (KLC1)-ALK, in lung cancer by using only an FFPE tissue.

Methods

Materials

A FFPE tissue block of pulmonary adenocarcinoma in situ, nonmucinous (formerly called bronchioloalveolar carcinoma) [36], which was excised from a 47-year-old female patient was used [37]. This carcinoma was negative for EML4-ALK and KIF5B-ALK, although the presence of *ALK* rearrangement was confirmed by anti-ALK iAEP immunohistochemistry and a split fluorescence in situ hybridization (FISH) assay for ALK (hereafter referred to as the unknown ALK fusion-positive case) (Figure 1) [37]. Two FFPE tissue blocks of ALK-positive tumor cases were also employed, for which

the presence of EML4-ALK or KIF5B-ALK had already been confirmed. Total RNA was extracted from each FFPE tissue with the use of the RecoverAll™ Total Nucleic Acid Isolation Kit for FFPE (Applied Biosystems Japan, Tokyo, Japan). The ages of the 3 FFPE blocks used (time from FFPE tissue production to RNA extraction) were 65, 40, and 51 months for the unknown ALK fusion-positive case, EML4-ALK, and KIF5B-ALK, respectively. Written informed consent was obtained from each patient. The study was approved by the institutional review board of the Shizuoka Cancer Center (approval ID 22-J132-22-1) and the Japanese Foundation for Cancer Research (approval ID 2010-1011).

Modified 5'-RACE for ALK fusions applicable to FFPE tissues

5'-RACE was performed with the SMARTer RACE cDNA Amplification Kit (Clontech) according to the manufacturer's instruction with minor modifications. In brief, instead of the primers included in the kit, ALK-3242R (5'-CTCAGCTTG-TACTCAGGGC-3') was used for cDNA synthesis. The cDNA was subjected to 5'-RACE PCR using PrimeSTAR HS DNA Polymerase (TaKaRa) and the following primers: Universal Primer A Mix of the kit and ALK-3206R (5'-ATGGCTTG-CAGCTCCTGGTGCTT-3'). The PCR condition consisted of 5 cycles at 94°C for 30 s and 72°C for 3 min; 5 cycles at 94°C for 30 s, 70°C for 30 s, and 72°C for 3 min; and 30 cycles at 94°C for 30 s, 68°C for 30 s, and 72°C for 3 min.

FISH

FISH analysis of fusion genes was performed with DNA probes for KLC1 and ALK. Unstained sections (4-μm thick) were subjected to hybridization with an ALK-split probe set (Dako, Tokyo, Japan) or with bacterial artificial chromosome (BAC) clone-derived probes for ALK (RP11-984I21 and RP11-62B19)

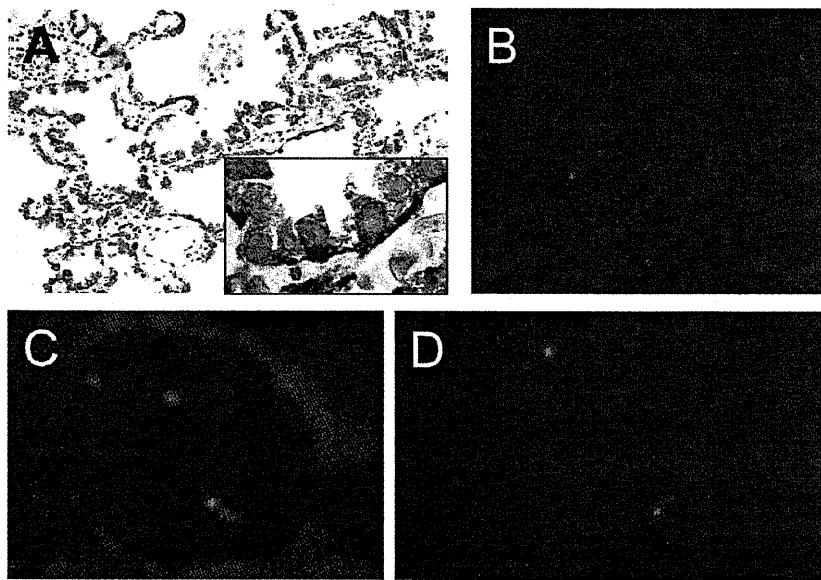


Figure 1. ALK-rearranged lung adenocarcinoma without EML4-ALK and KIF5B-ALK. Panel A shows the results of anti-ALK immunohistochemistry with the iAEP method on pulmonary adenocarcinoma in situ, nonmucinous. The staining pattern was diffusely cytoplasmic. The basal side of tumor cells was more strongly stained, indicating an uneven subcellular localization of KLC1-ALK protein. FISH analyses revealed that this case was positive in the split assay for ALK (Panel B: individual 5'- and 3'-signals are observed) and negative in EML4-ALK and KIF5B-ALK fusion assays (Panel C: EML4, red; ALK, green; Panel D: KIF5B, green; ALK, red). doi:10.1371/journal.pone.0031323.g001

and KLC1 (RP11-186F6). Hybridized slides were then stained with DAPI and examined using a BX51 fluorescence microscope (Olympus, Tokyo, Japan).

Synthesis of the putative cDNA of *KLC1-ALK*

Two independent PCRs were performed using cDNA synthesized from a tumor tissue expressing KIF5B-ALK with the following primer sets: KLC1-NheI-M (5'-GCGCTAGCGAATGTATGAC-AACATGTCCAC-3') and KLC1-bpR (5'-GTGCTTCCGGCCGTACACATCTACAGAACCAACTC-3'), and ALK-bpF (5'-GGAGTTTGGTTCTGTAGATGTGTACCGCCGGAAGC-3') and ALK-EcoRI (5'-GATAGAATTCTCAGGGCCCAGGCT-3'). Then, the second PCR was performed using a 1/100 dilution of a mixture of the first PCR products as a template with the KLC1-NheI-M and ALK-EcoRI primers (Figure 2).

Transformation assay for *KLC1-ALK*

Analysis of the transforming activity of kinase fusions was performed as described previously [9,38,39]. A pMXS-based expression plasmid for each fusion was used to generate recombinant ecotropic retroviruses [40], which were then used individually to infect mouse 3T3 fibroblasts. The formation of transformed foci was evaluated after culturing the cells for 4 days. The same set of 3T3 cells was injected subcutaneously into nu/nu mice, and tumor formation was examined after 14 days. The animal experiments were approved by the animal ethics committee of Jichi Medical University (approval ID 1135).

Results

Identification of *KLC1-ALK* as a novel ALK fusion gene

Our modified 5'-RACE faithfully isolated cDNA fragments for *EML4-ALK* or *KIF5B-ALK* from known ALK-positive tumors

(Supplementary Figure S1A and B). We then attempted to isolate cDNA fragments encompassing the fusion points from the unknown ALK fusion-positive case. Nucleotide sequencing of such 5'-RACE products revealed that 2 of 10 clones contained the 3'-terminus of exon 9 of *KLC1* (ENST00000348520) fused to the first nucleotide of exon 20 of *ALK* (ENST00000389048), indicating the presence of a novel fusion between *KLC1* and *ALK*. As this rearrangement constituted an in-frame fusion between the 2 genes, the full-length *KLC1-ALK* cDNA probably produces a protein of 984 amino acids containing an amino-terminal two-thirds of KLC1 and an intracellular region of ALK (Figure 3A). RT-PCR-mediated isolation of a fusion point successfully confirmed the in-frame fusion between the 2 messages (Figure 3A and B). Further, to confirm the genomic rearrangement responsible for the fusion, a fusion FISH assay was performed (Figure 3C). These results were consistent with the presence of t(2;14)(p23;q32.3), leading to the generation of *KLC1-ALK*.

Transforming potential of *KLC1-ALK*

The putative full-length cDNA of *KLC1-ALK* was synthesized from the frozen tissue with KIF5B-ALK fusion expression (Figure 2, Supplementary Figure S2), and was used to generate a recombinant retrovirus expressing the fusion protein with an amino-terminal FLAG epitope tag. Infection of 3T3 cells with the virus expressing KLC1-ALK readily produced multiple transformed foci in culture and subcutaneous tumors in a nude mouse tumorigenicity assay (Figure 4), confirming the potent transforming ability of KLC1-ALK.

Discussion

Here, by analyzing the FFPE tissues only, we successfully discovered a novel ALK fusion, KLC1-ALK. While snap-frozen materials sampled from biopsied or surgically removed specimens

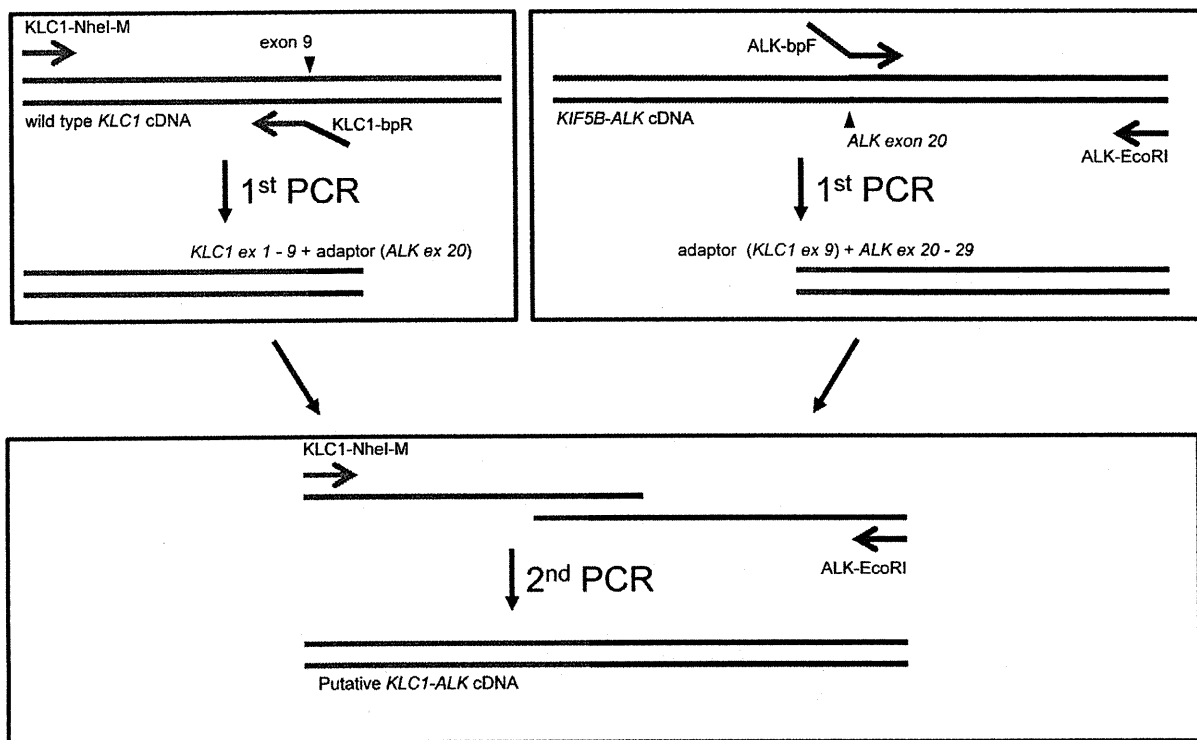


Figure 2. Synthesis of the putative KLC1-ALK full-length cDNA. Two first-round PCRs were performed separately using cDNA synthesized from a tumor tissue expressing KIF5B-ALK with the following primer sets: KLC1-NheI-M and KLC1-bpR, and ALK-bpF and ALK-EcoRI. KLC1-bpR and ALK-bpF had sequences downstream of the ALK break point (exon 20) and upstream of the KLC1 break point (exon 9) as adaptor sequences, respectively. Then, the second PCR was performed using a 1/100 dilution of the mixture of the first PCR products as a template with primers KLC1-NheI-M and ALK-EcoRI. The first PCR products were annealed, extended with each other, and then amplified with the primers.
doi:10.1371/journal.pone.0031323.g002

can be used for various types of molecular analyses, they are not routinely sampled in most clinical settings. In contrast, FFPE specimens are usually produced, and histopathology diagnostic archives are an extremely large resource of FFPE tissues in ordinary diagnostic pathology laboratories. However, DNA and RNA extracted from FFPE tissues are severely degraded during formalin fixation and are usually not suitable for assays that need long DNA/RNA of high quality. Recent technical advances have allowed some analyses for known point mutations and known fusion genes, but it is still difficult to identify an unrecognized gene aberration using only an FFPE tissue.

In most *ALK* fusions, the break point of *ALK* is located within intron 19, and the fusion point in mRNA is typically the first nucleotide of exon 20. Therefore, if the primers for 5'-RACE are placed immediately downstream of the first nucleotide of *ALK* exon 20, such 5'-RACE may successfully isolate PCR products containing the partner gene sequence even using FFPE tissues. Based on this hypothesis, we established a 5'-RACE system for *ALK* fusions optimized for FFPE tissues. With this system, we identified a novel *ALK* fusion, *KLC1-ALK*. To the best of our knowledge, this is the first novel oncogenic fusion identified using only an FFPE tissue.

Caution, however, is needed. In some rare cases with *ALK* fusion, the break point of *ALK* fusion mRNA may not be at the 5'-end of exon 20. For example, in variant 4 of *EML4-ALK*, exon 14 of *EML4* is fused to an unknown sequence of 11 bp, which in turn is connected to nucleotide 50 of *ALK* exon 20 (E14;in-

s11;del49A20) [38]. Our 5'-RACE system would not work on such a case because the reverse primer ALK-3206R corresponds to nucleotides 12–34 of *ALK* exon 20. Therefore, if our modified 5'-RACE fails to isolate fusion cDNAs from cases with a confirmed *ALK* rearrangement, other primer settings may be attempted.

Kinesin is a heterotetramer of 2 kinesin heavy chains and 2 kinesin light chains, and it moves on the microtubules towards their plus ends carrying various cargos. The heavy chains harbor the motor activity, whereas the light chains play roles in cargo binding and in modulating the activity and subcellular localization of the heavy chains. KLC1 binds to the kinesin heavy chains with an N-terminal domain and to various cargos via the tetratricopeptide repeat domains [41,42]. Of the 3 histopathologically confirmed *ALK* fusion partners in lung cancer, *EML4* colocalizes with microtubules and may contribute to the stabilization of microtubules [43], *KIF5B* moves on the microtubules as a kinesin heavy chain [44], and *KLC1* binds to kinesin heavy chains as a kinesin light chain. Therefore, it is interesting that all the 3 *ALK* fusions in lung cancer are likely to colocalize with microtubules.

The most frequent *ALK* fusion in lung cancer is *EML4-ALK* (4–7%) [9,38], and the second is *KIF5B-ALK* (0.5%) [11]. One case with *TFG-ALK* is reported [10]. *KLC1-ALK* may be rare but exists in lung adenocarcinoma, and the patients with this fusion are highly likely to benefit from *ALK* inhibitor therapy as do patients with other *ALK* fusions. The incidence may be low, but the significance of this fusion is very high from the perspective

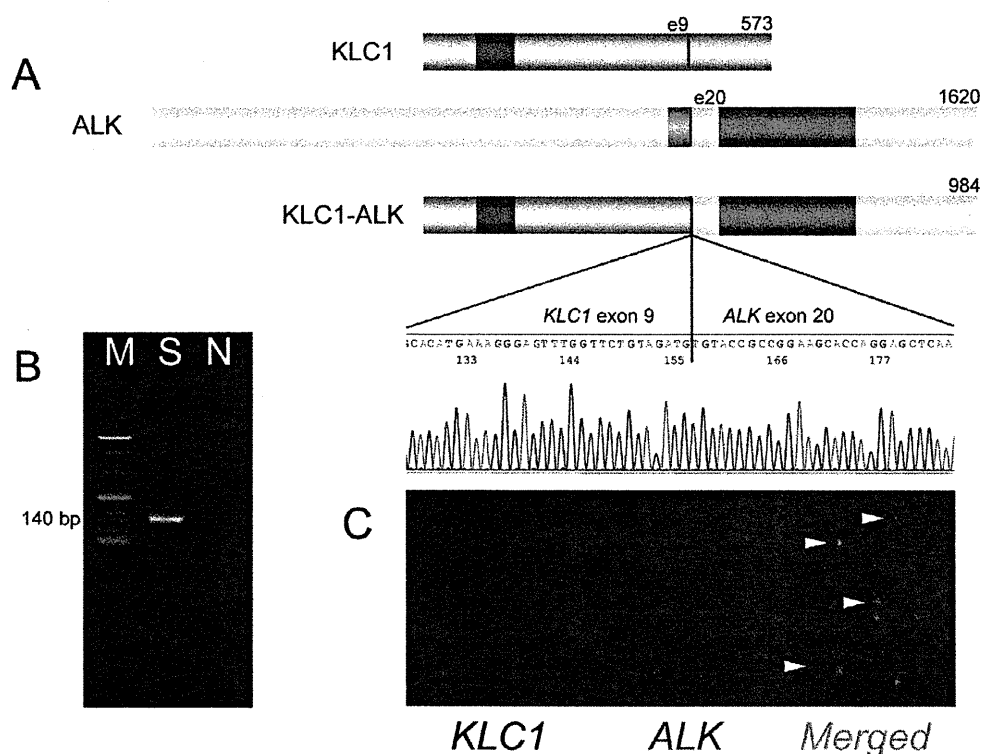


Figure 3. Identification of KLC1-ALK. Panel A shows the schematic structure of KLC1, ALK, and KLC1-ALK proteins and the cDNA sequence around the fusion point. Dark blue, orange, and red parts represent coiled-coil, transmembrane, and kinase domains, respectively. The break point exons and the number of amino acids are indicated. KLC1-ALK-specific RT-PCR using RNA extracted from the FFPE tissue of the unknown ALK fusion-positive case amplified a fragment of the expected product size (140 bp, Panel B) with the consistent fusion sequence (Panel A). A fusion FISH assay for KLC1-ALK revealed a fusion signal (yellow) in multiple tumor cells (Panel C). M, marker (100-bp ladder); S, sample (the unknown ALK fusion-positive case); N, no template control.

doi:10.1371/journal.pone.0031323.g003

of a tailor-made therapeutic option for the patient. Another important point is that KLC1-ALK was found in adenocarcinoma in situ, nonmucinous (formerly called bronchioloalveolar carcinoma, BAC). BAC is recognized to rarely harbor ALK fusions, although a small number of BAC cases has been examined for ALK fusion compared with invasive adenocarcinoma. It would be

interesting from a pathobiological perspective to examine a large-scale cohort of BAC and other premalignant conditions for ALK fusion.

There are 3 methods for the detection of ALK fusions: RT-PCR, ALK split FISH, and high-sensitivity anti-ALK immunohistochemistry. For RT-PCR, the 5' partner gene must be known. Our findings in this study identified one more partner gene that should be targeted in ALK-fusion detection using RT-PCR in lung cancer. The other 2 methods can detect all ALK fusions regardless of fusion partner and, therefore, are suitable for ALK-fusion screening. In other words, these 2 methods cannot identify the fusion partner and need to be succeeded by partner-specific RT-PCR and/or fusion FISH for this purpose. If it is revealed that the partner gene in the tested case is unknown, a novel partner gene is highly likely to be discovered, as was shown in the present study. In fact, using high-sensitivity anti-ALK immunohistochemistry (iAEP method) as screening, we have identified several novel ALK fusions in various types of cancers including lung adenocarcinoma [11], lymphoma [28], sarcoma [35], and renal cell carcinoma [30].

Many efficient tools have been established for the detection of ALK fusion-positive cases using FFPE tissues, including anti-ALK immunohistochemistry and FISH. Our findings will further expand the potential value of archival FFPE tissues and provide further biological and clinical insights into ALK-positive cancers in the forthcoming era of ALK inhibitor therapy.

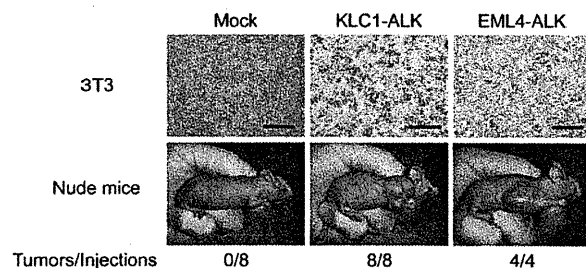


Figure 4. Transforming potential of KLC1-ALK. Upper panels: Mouse 3T3 fibroblasts were infected with retroviruses encoding KLC1-ALK or EML4-ALK or with the corresponding empty virus (Mock). The cells were photographed after 4 days of culture. Scale bar, 1 mm. Lower panels: Nude mice were injected subcutaneously with the corresponding 3T3 cells, and tumor formation was examined after 14 days. The number of tumors formed per injections is indicated at the bottom.

doi:10.1371/journal.pone.0031323.g004

Supporting Information

Figure S1 5'-RACE products using FFPE tissues. Our modified 5'-RACE faithfully isolated cDNA fragments for *EML4-ALK* (A) or *KIF5B-ALK* (B) from known ALK-positive tumors. (TIF)

Figure S2 Putative cDNA sequence of KLC1-ALK. The putative full-length cDNA of *KLC1-ALK* was synthesized from the frozen tissue with KIF5B-ALK fusion expression. (PDF)

References

- Morris SW, Kirstein MN, Valentine MB, Dittner KG, Shapiro DN, et al. (1994) Fusion of a kinase gene, ALK, to a nucleolar protein gene, NPM, in non-Hodgkin's lymphoma. *Science* 263: 1281–1284.
- Shiota M, Fujimoto J, Semba T, Satoh H, Yamamoto T, et al. (1994) Hyperphosphorylation of a novel 80 kDa protein-tyrosine kinase similar to Ltk in a human Ki-1 lymphoma cell line, AMS3. *Oncogene* 9: 1567–1574.
- Kwak EL, Bang YJ, Camidge DR, Shaw AT, Solomon B, et al. (2010) Anaplastic lymphoma kinase inhibition in non-small-cell lung cancer. *N Engl J Med* 363: 1693–1703.
- Kimura H, Nakajima T, Takeuchi K, Soda M, Mano H, et al. (2011) ALK fusion gene positive lung cancer and 3 cases treated with an inhibitor for ALK kinase activity. *Lung Cancer*.
- Kijima T, Takeuchi K, Tetsumoto S, Shimada K, Takahashi R, et al. (2011) Favorable response to crizotinib in three patients with echinoderm microtubule-associated protein-like 4-anaplastic lymphoma kinase fusion-type oncogene-positive non-small cell lung cancer. *Cancer Sci* 102: 1602–1604.
- Nakajima T, Kimura H, Takeuchi K, Soda M, Mano H, et al. (2010) Treatment of Lung Cancer with an ALK Inhibitor After EML4-ALK Fusion Gene Detection Using Endobronchial Ultrasound-Guided Transbronchial Needle Aspiration. *J Thorac Oncol* 5: 2041–2043.
- Butrynski JE, D'Adamo DR, Hornick JL, Dal Cin P, Antonescu CR, et al. (2010) Crizotinib in ALK-rearranged inflammatory myofibroblastic tumor. *N Engl J Med* 363: 1727–1733.
- Gambacorti-Passerini C, Messa C, Pogliani EM (2011) Crizotinib in anaplastic large-cell lymphoma. *N Engl J Med* 364: 775–776.
- Soda M, Choi YL, Enomoto M, Takada S, Yamashita Y, et al. (2007) Identification of the transforming EML4-ALK fusion gene in non-small-cell lung cancer. *Nature* 448: 561–566.
- Rikova K, Guo A, Zeng Q, Possemato A, Yu J, et al. (2007) Global survey of phosphotyrosine signaling identifies oncogenic kinases in lung cancer. *Cell* 131: 1190–1203.
- Takeuchi K, Choi YL, Togashi Y, Soda M, Hatano S, et al. (2009) KIF5B-ALK, a novel fusion oncogene identified by an immunohistochemistry-based diagnostic system for ALK-positive lung cancer. *Clin Cancer Res* 15: 3143–3149.
- Lamant L, Dastugue N, Pulford K, Delsol G, Mariame B (1999) A new fusion gene TPM3-ALK in anaplastic large cell lymphoma created by a (1;2)(q25;p23) translocation. *Blood* 93: 3088–3095.
- Meech SJ, McGavran L, Odom LF, Liang X, Meltesen L, et al. (2001) Unusual childhood extramedullary hematologic malignancy with natural killer cell properties that contains tropomyosin 4-anaplastic lymphoma kinase gene fusion. *Blood* 98: 1209–1216.
- Colleoni GW, Bridge JA, Garicochea B, Liu J, Filippa DA, et al. (2000) ATIC-ALK: A novel variant ALK gene fusion in anaplastic large cell lymphoma resulting from the recurrent cryptic chromosomal inversion, inv(2)(p23q35). *Am J Pathol* 156: 781–789.
- Hernandez L, Pinyol M, Hernandez S, Bea S, Pulford K, et al. (1999) TRK-fused gene (TFG) is a new partner of ALK in anaplastic large cell lymphoma producing two structurally different TFG-ALK translocations. *Blood* 94: 3265–3268.
- Touriol C, Greenland C, Lamant L, Pulford K, Bernard F, et al. (2000) Further demonstration of the diversity of chromosomal changes involving 2p23 in ALK-positive lymphoma: 2 cases expressing ALK kinase fused to CLTCL (clathrin chain polypeptide-like). *Blood* 95: 3204–3207.
- Tort F, Pinyol M, Pulford K, Roncador G, Hernandez L, et al. (2001) Molecular characterization of a new ALK translocation involving moesin (MSN-ALK) in anaplastic large cell lymphoma. *Lab Invest* 81: 419–426.
- Lamant L, Gascoyne RD, Duplantier MM, Armstrong F, Raghab A, et al. (2003) Non-muscle myosin heavy chain (MYH9): a new partner fused to ALK in anaplastic large cell lymphoma. *Genes Chromosomes Cancer* 37: 427–432.
- Cools J, Wlodarska I, Somers R, Mentens N, Pedutoir F, et al. (2002) Identification of novel fusion partners of ALK, the anaplastic lymphoma kinase, in anaplastic large-cell lymphoma and inflammatory myofibroblastic tumor. *Genes Chromosomes Cancer* 34: 354–362.
- Lawrence B, Perez-Atayde A, Hibbard MK, Rubin BP, Dal Cin P, et al. (2000) TPM3-ALK and TPM4-ALK oncogenes in inflammatory myofibroblastic tumors. *Am J Pathol* 157: 377–384.
- Bridge JA, Kanamori M, Ma Z, Pickering D, Hill DA, et al. (2001) Fusion of the ALK gene to the clathrin heavy chain gene, CLTC, in inflammatory myofibroblastic tumor. *Am J Pathol* 159: 411–415.
- Ma Z, Hill DA, Collins MH, Morris SW, Sumegi J, et al. (2003) Fusion of ALK to the Ran-binding protein 2 (RANBP2) gene in inflammatory myofibroblastic tumor. *Genes Chromosomes Cancer* 37: 98–105.
- Debiec-Rychter M, Marynen P, Hagemeyer A, Pauwels P (2003) ALK-ATIC fusion in urinary bladder inflammatory myofibroblastic tumor. *Genes Chromosomes Cancer* 38: 187–190.
- Panagopoulos I, Nilsson T, Domanski HA, Isaksson M, Lindblom P, et al. (2006) Fusion of the SEC31L1 and ALK genes in an inflammatory myofibroblastic tumor. *Int J Cancer* 118: 1181–1186.
- Delsol G, Lamant L, Mariame B, Pulford K, Dastugue N, et al. (1997) A new subtype of large B-cell lymphoma expressing the ALK kinase and lacking the 2;5 translocation. *Blood* 89: 1483–1490.
- Gascoyne RD, Lamant L, Martin-Subero JI, Lestou VS, Harris NL, et al. (2003) ALK-positive diffuse large B-cell lymphoma is associated with Clathrin-ALK rearrangements: report of 6 cases. *Blood* 102: 2568–2573.
- Van Roosbroeck K, Cools J, Dierickx D, Thomas J, Vandenberghe P, et al. (2010) ALK-positive large B-cell lymphomas with cryptic SEC31A-ALK and NPM1-ALK fusions. *Haematologica* 95: 509–513.
- Takeuchi K, Soda M, Togashi Y, Ota Y, Sekiguchi Y, et al. (2010) Identification of a novel fusion, SQSTM1-ALK, in ALK-positive large B-cell lymphoma. *Haematologica*.
- Debelenko LV, Raimondi SC, Daw N, Shivakumar BR, Huang D, et al. (2011) Renal cell carcinoma with novel VCL-ALK fusion: new representative of ALK-associated tumor spectrum. *Mod Pathol* 24: 430–442.
- Sugawara E, Togashi Y, Kuroda N, Sakata S, Hatano S, et al. (in press) Identification of ALK Fusions in Renal Cancer: Large-Scale Immunohistochemical Screening by intercalated Antibody-Enhanced Polymer Method. *Cancer*.
- Du XL, Hu H, Lin DC, Xia SH, Shen XM, et al. (2007) Proteomic profiling of proteins dysregulated in Chinese esophageal squamous cell carcinoma. *J Mol Med* 85: 863–875.
- Jazji FR, Najafi Z, Malekzadeh R, Conrads TP, Ziaee AA, et al. (2006) Identification of squamous cell carcinoma associated proteins by proteomics and loss of beta tropomyosin expression in esophageal cancer. *World J Gastroenterol* 12: 7104–7112.
- Lin E, Li L, Guan Y, Soriano R, Rivers CS, et al. (2009) Exon array profiling detects EML4-ALK fusion in breast, colorectal, and non-small cell lung cancers. *Mol Cancer Res* 7: 1466–1476.
- Martelli MP, Sozzi G, Hernandez L, Pettrossi V, Navarro A, et al. (2009) EML4-ALK rearrangement in non-small cell lung cancer and non-tumor lung tissues. *Am J Pathol* 174: 661–670.
- Takeuchi K, Soda M, Togashi Y, Sugawara E, Hatano S, et al. (2011) Pulmonary inflammatory myofibroblastic tumor expressing a novel fusion, PPFBP1-ALK: reappraisal of anti-ALK immunohistochemistry as a tool for novel ALK fusion identification. *Clin Cancer Res* 17: 3341–3348.
- Travis WD, Brambilla E, Noguchi M, Nicholson AG, Geisinger KR, et al. (2011) International association for the study of lung cancer/american thoracic society/european respiratory society international multidisciplinary classification of lung adenocarcinoma. *J Thorac Oncol* 6: 244–285.
- Yamamoto M, Takeuchi K, Shimoji M, Maniwa T, Isaka M, et al. (in press) Small non-mucinous bronchioloalveolar carcinoma with ALK immunoreactivity: A novel ALK fusion gene? *Cancer Sci*.
- Takeuchi K, Choi YL, Soda M, Inamura K, Togashi Y, et al. (2008) Multiplex reverse transcription-PCR screening for EML4-ALK fusion transcripts. *Clin Cancer Res* 14: 6618–6624.
- Choi YL, Takeuchi K, Soda M, Inamura K, Togashi Y, et al. (2008) Identification of novel isoforms of the EML4-ALK transforming gene in non-small cell lung cancer. *Cancer Res* 68: 4971–4976.
- Onishi M, Kinoshita S, Morikawa Y, Shibuya A, Phillips J, et al. (1996) Applications of retrovirus-mediated expression cloning. *Exp Hematol* 24: 324–329.
- Stenoien DL, Brady ST (1997) Immunohistochemical analysis of kinesin light chain function. *Mol Biol Cell* 8: 675–689.

Acknowledgments

We thank Mr. Motoyoshi Iwakoshi, Ms. Keiko Shiozawa, Ms. Tomoyo Kakita, and Ms. Kimie Nomura for their technical assistance, and Ms. Sayuri Sengoku for providing administrative assistance.

Author Contributions

Conceived and designed the experiments: KT HM. Performed the experiments: YT MS SS ES SH RA. Analyzed the data: YT MS HM KT. Contributed reagents/materials/analysis tools: RA TN. Wrote the paper: KT HM.

42. Rahman A, Kamal A, Roberts EA, Goldstein LS (1999) Defective kinesin heavy chain behavior in mouse kinesin light chain mutants. *J Cell Biol* 146: 1277–1288.
43. Houtman SH, Rutteman M, De Zeeuw CI, French PJ (2007) Echinoderm microtubule-associated protein like protein 4, a member of the echinoderm microtubule-associated protein family, stabilizes microtubules. *Neuroscience* 144: 1373–1382.
44. Sablin EP (2000) Kinesins and microtubules: their structures and motor mechanisms. *Curr Opin Cell Biol* 12: 35–41.

RET, ROS1 and ALK fusions in lung cancer

Kengo Takeuchi^{1,2}, Manabu Soda³, Yuki Togashi^{1,2}, Ritsuro Suzuki⁴, Seiji Sakata¹, Satoko Hatano¹, Reimi Asaka^{1,2}, Wakako Hamanaka², Hironori Ninomiya², Hirofumi Uehara⁵, Young Lim Choi⁶, Yukitoshi Satoh^{5,7}, Sakae Okumura⁵, Ken Nakagawa⁵, Hiroyuki Mano^{3,6} & Yuichi Ishikawa²

Through an integrated molecular- and histopathology-based screening system, we performed a screening for fusions of anaplastic lymphoma kinase (ALK) and c-ros oncogene 1, receptor tyrosine kinase (ROS1) in 1,529 lung cancers and identified 44 ALK-fusion-positive and 13 ROS1-fusion-positive adenocarcinomas, including for unidentified fusion partners for ROS1. In addition, we discovered previously unidentified kinase fusions that may be promising for molecular-targeted therapy, kinesin family member 5B (KIF5B)-ret proto-oncogene (RET) and coiled-coil domain containing 6 (CCDC6)-RET, in 14 adenocarcinomas. A multivariate analysis of 1,116 adenocarcinomas containing these 71 kinase-fusion-positive adenocarcinomas identified four independent factors that are indicators of poor prognosis: age ≥ 50 years, male sex, high pathological stage and negative kinase-fusion status.

Echinoderm microtubule associated protein like 4 (EML4)-ALK was the first targetable fusion oncokine to be identified in non-small cell lung cancer (NSCLC)¹. This fusion is found in approximately 4–6% of lung adenocarcinomas^{2,3}. ROS1 is another receptor tyrosine kinase that forms fusions in NSCLC⁴. Solute carrier family 34 (sodium phosphate), member 2 (SLC34A2)-ROS1 and CD74 molecule, major histocompatibility complex, class II invariant chain (CD74)-ROS1 were identified in 1 out of 41 NSCLC cell lines and 1 out of 150 lung cancer samples, respectively⁴. However, the oncogenic ability of these ROS1 fusion proteins and the incidence of ROS1 fusions in lung cancers are still unclear.

We screened for known and unknown kinase fusions in lung cancers using a histopathology-based system with tissue microarrays of 1,528 surgically removed tissues (Supplementary Methods and Supplementary Appendix). Immunohistochemistry of antibodies to ALK using the intercalated antibody-enhanced polymer method^{2,3,5–7} detected 45 tumors with ALK kinase domain expression (Supplementary Fig. 1). In 44 adenocarcinomas, multiplex RT-PCR^{2,3}

identified 41 *EML4*-ALK-positive and 3 *KIF5B*-ALK-positive adenocarcinomas, including a previously unidentified *KIF5B*-ALK fusion variant, K17;A20 (Supplementary Table 1). Further, we used fluorescence *in situ* hybridization (FISH) for split and fusion assays to confirm the presence of ALK fusions^{2,3,8}. The FISH results for the ALK split assay, the *EML4*-ALK fusion assay and the *KIF5B*-ALK fusion assay in the 44 adenocarcinomas were all consistent with the presence of the corresponding fusion gene (Supplementary Figs. 2 and 3). The remaining tumor that was positive for antibodies to ALK as determined by immunohistochemistry (a large-cell neuroendocrine carcinoma) was negative in the FISH assays and expressed wild-type ALK. ALK fusions existed in 3.0% (44 out of 1,485) of the NSCLCs and 3.9% (44 out of 1,121) of the adenocarcinomas. We included 20 previously reported ALK-fusion-positive and 304 ALK-fusion-negative tumors, all of which were screened with multiplex RT-PCR. Because specimens of these 324 patients were collected consecutively during the period of tissue collection, they served as positive and negative controls, respectively^{1–3,8,9}. The immunohistochemistry results using the intercalated antibody-enhanced polymer method were complete matches in the 20 fusion-positive and the 304 fusion-negative tumors.

We used split FISH assays for the screening for ROS1 gene rearrangement (Fig. 1). In 11 of the 13 ROS1 split FISH-positive tumors (Fig. 1a), 5' rapid amplification of complementary DNA ends (5' RACE) identified two known and three unknown fusion partners for ROS1: *TPM3*, *SDC4*, *SLC34A2*, *CD74* and *EZR* (Fig. 1b and Supplementary Table 1); RT-PCR confirmed this finding (Fig. 1c). In a 5'-RACE-negative tumor (ROS#12) (again, where split FISH is used to detect candidate fusion genes of interest by the presence of rearrangements and RACE is used for the identification of fusion partners), each fusion-specific RT-PCR (using a common reverse primer) amplified the same band, which contained an *LRIG3* sequence. This tumor was proven fusion-positive in RT-PCR specific to *LRIG3*-ROS1, an unidentified fusion. Fusion FISH results confirmed that all 12 cases harbored the corresponding fusion (Fig. 1a). All fusion FISH assays for these six ROS1 fusions were negative for the tumor ROS#13 (the frozen material had been consumed), indicating an unknown fusion partner for ROS1. ROS1 split FISH screening failed for nine NSCLCs, including five adenocarcinomas. We identified ROS1 fusions in 0.9% (13 out of 1,476) of the NSCLCs and 1.2% (13 out of 1,116) of the adenocarcinomas.

We performed *KIF5B* split FISH to discover new fusion kinases, as we previously identified *KIF5B*-ALK fusions in lung cancer³. As such, we hypothesized that *KIF5B* might be rearranged in lung cancer. In 24 *KIF5B* split FISH-positive tumors, 3' RACE identified an in-frame fusion between *KIF5B* exon 23 and *RET* exon 12

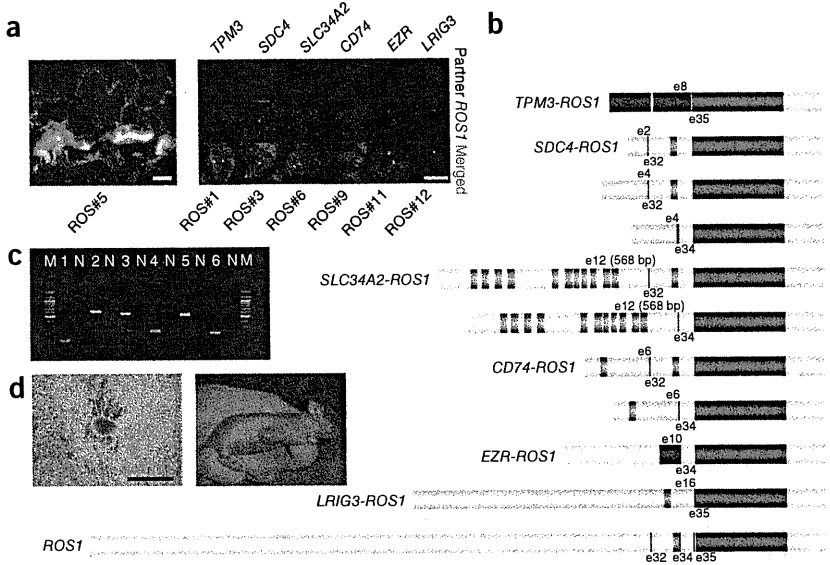
¹Pathology Project for Molecular Targets, the Cancer Institute, Japanese Foundation for Cancer Research, Tokyo, Japan. ²Division of Pathology, the Cancer Institute, Japanese Foundation for Cancer Research, Tokyo, Japan. ³Division of Functional Genomics, Jichi Medical University, Tochigi, Japan. ⁴Department of Hematopoietic Stem Cell Transplantation Data Management and Biostatistics, Nagoya University Graduate School of Medicine, Nagoya, Japan. ⁵Department of Thoracic Surgical Oncology, Thoracic Center, the Cancer Institute Hospital, Japanese Foundation for Cancer Research, Tokyo, Japan. ⁶Department of Medical Genomics, Graduate School of Medicine, University of Tokyo, Tokyo, Japan. ⁷Present address: Department of Thoracic Surgery, Kitasato University School of Medicine, Kanagawa, Japan. Correspondence should be addressed to K.T. (kentakuchi-ty@umin.net).

Received 28 September 2011; accepted 3 January 2012; published online 12 February 2012; doi:10.1038/nm.2658

BRIEF COMMUNICATIONS

Figure 1 Identification of ROS1 fusions.

(a) *ROS1* split (left) and fusion (right) FISH assay data (scale bars, 20 μ m). In the split assay, multiple tumor cells harbored individual 3' side signals (green), indicating the presence of a *ROS1* rearrangement. In the fusion assay, a fusion signal (yellow) was observed in the representative tumor cell of each subject, which is consistent with the presence of t(1;6)(q21.2;q22) for *TPM3-ROS1*, t(6;20)(q22;q12) for *SDC4-ROS1*, t(4;6)(q15.2;q22) for *SLC34A2-ROS1*, t(5;6)(q32;q22) for *CD74-ROS1*, inv(6)(q22q25.3) for *EZR-ROS1* or t(6;12)(q22;q14.1) for *LRIG3-ROS1*. (b) The break points of *ROS1* are exons 32, 34 and 35. All of the break points allow the resulting fusion to harbor the kinase domain of *ROS1* (red), and the exon 32 break point allows the resulting fusion to harbor the transmembrane domain of *ROS1* (orange). In the fusion partners, dark blue and orange represent coiled-coil and transmembrane domains, respectively. Coiled-coil domains may contribute to homodimerization, but only *TPM3* and *EZR* contained these domains. In contrast to *ALK* and *RET* fusions, the role of the fusion partner's coiled-coil domain is unknown in *ROS1* fusions. (c) Results for fusion-specific RT-PCR for tumors ROS#1 (lane 1, *TPM3-ROS1*, T8;R35, predicted product size of 119 bp), ROS#3 (lane 2, *SDC4-ROS1*, S2;R32, 596 bp), ROS#6 (lane 3, *SLC34A2-ROS1*, S13del2046;R32 and S13del2046;R34, 544 bp and 235 bp, respectively), ROS#8 (lane 4, *CD74-ROS1*, C6;R34, 230 bp), ROS#10 (lane 5, *EZR-ROS1*, E10;R34, 527 bp), and ROS#12 (lane 6, *LRIG3-ROS1*, L16;R35, 218 bp). M and N represent the size marker (100-bp ladder) and the non-template control, respectively. (d) The transforming potential of the *ROS1* fusion. Mouse 3T3 fibroblasts infected with a retrovirus encoding *SDC4-ROS1* derived from tumor ROS#4 formed multiple foci (scale bar, 1 mm). All of the four nude mice injected with the corresponding 3T3 cells developed a subcutaneous tumor (right).



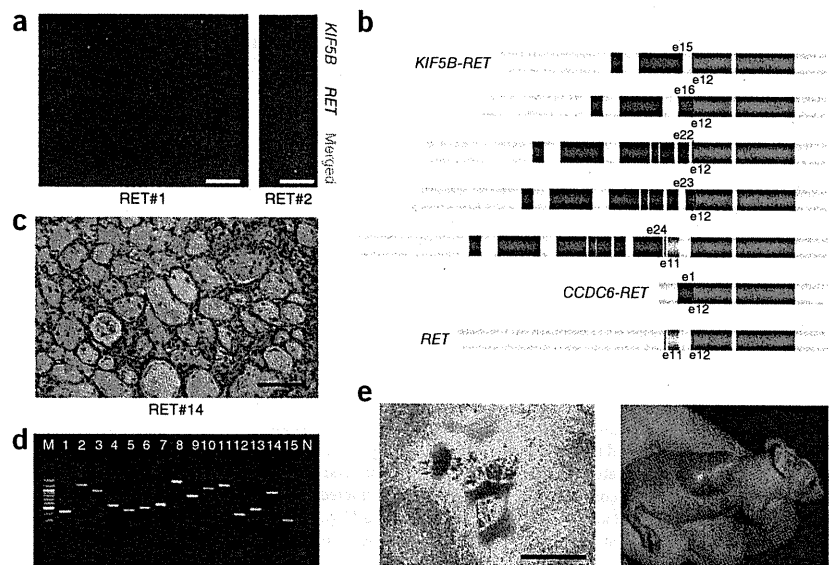
(tumor RET#11). *RET* split FISH on the tissue arrays identified 22 fusion-positive tumors in 1,528 lung cancers (Fig. 2a), from which a multiplex RT-PCR system that captures all possible *KIF5B-RET* fusions detected 12 fusion-positive tumors: eight tumors with the fusion of *KIF5B* exon 15 and *RET* exon 12 (K15;R12) and one tumor each with the K16;R12, K22;R12, K23;R12 and K24;R11 fusions (Fig. 2b and Supplementary Table 1). The *KIF5B-RET* fusion FISH results were consistent with the presence of inv(10)(p11.22q11.2) in all 12 of these tumors (Fig. 2a).

In a routine histopathological diagnosis, we encountered an adenocarcinoma that showed a mucinous cribriform pattern (Fig. 2c) that was previously reported as a histopathological marker for the presence of *EML4-ALK* (Supplementary Fig. 4)⁹⁻¹¹. Notably, this adenocarcinoma (tumor RET#14) was negative for *ALK* fusion and was positive for *CCDC6-RET*, as determined by FISH and inverse RT-PCR; the latter fusion gene was first described in thyroid cancer¹². RT-PCR identified another tumor positive for the *CCDC6-RET* fusion (RET#13) in the remaining 10 tumors. The 14 *RET*-positive tumors (out of the total 1,528 tumors tested, with one additional tumor (RET#14) found through routine pathology diagnostic service) were also positive in the revised multiplex RT-PCR that captured *EML4-ALK*, *KIF5B-ALK*, *KIF5B-RET* and *CCDC6-RET* simultaneously (Fig. 2d). The *RET* kinase domain expression using real-time RT-PCR was weak or undetectable for the remaining nine tumors determined to be positive in the *RET* split FISH screening. Perhaps the genomic rearrangement occurred downstream of the *RET* break points. *RET* split FISH screening failed in three NSCLCs, including two adenocarcinomas. RET#14 was the index case found in routine pathology diagnostic service but not in the 1,528 cohort. *RET* fusions existed in 0.9% (13 out of 1,482) of the NSCLCs and 1.2% (13 out of 1,119) of the adenocarcinomas. The 14 *RET* fusion-positive subjects did not receive vandetanib.

We concluded that the rearrangements described above are somatic without using any matched normal tissues. Our histopathology-based screening method preserves the samples' histological architecture. This allows observers to confirm that internal non-tumor cells, for example, epithelial cells, inflammatory cells or fibroblasts, are negative in a test of interest.

All 71 kinase-fusion-positive (44 *ALK*, 13 *ROS1* and 14 *RET* fusions) lung cancers were exclusively adenocarcinomas (6% of all adenocarcinomas in the present study), were positive for antibodies to TTF1, which is regarded as a marker for lung adenocarcinoma, as determined by immunohistochemistry (excluding two *ALK*-positive tumors) and were negative for *EGFR* and *KRAS* mutations. Thirteen of the 44 *ALK*-positive tumors (30%) were weakly positive for p63 expression (were weakly positive for a squamous cell carcinoma marker, p63) (Supplementary Table 1). Thirty-three tumors showed a mucinous cribriform pattern in at least 5% of their area; 22 tumors had this pattern in >25% of their area (Fig. 2c, Supplementary Table 1 and Supplementary Fig. 4). The frequency of mucinous cribriform carcinoma was significantly higher in the kinase-fusion-positive group of tumors than in the 77 fusion-negative adenocarcinomas (22 out of 71 compared to 7 out of 77, respectively; $P = 0.00088$). Notably, we observed this pattern preferentially in *EML4-ALK*-positive tumors (70%, 29 out of 41); all three *CD74-ROS1*-positive tumors also showed this pattern. Recognizing this pattern in routine pathology diagnoses led to the identification of the *CCDC6-RET* fusion (tumor RET#14). In organs other than the lung, secretory breast carcinoma, which is characterized by a cribriform pattern with abundant secretory material, harbors the ets variant 6 (ETV6)-neurotrophic tyrosine kinase, receptor, type 3 (NTRK3) fusion (ref. 13). We identified an *ALK*-fusion-positive renal cell carcinoma that showed a mucinous cribriform pattern⁷. This pattern may be linked to the presence of particular kinase fusions¹⁰, and this possibility warrants further study.

Figure 2 Discovery of RET fusions. (a) *RET* split (left) and fusion (right) FISH assay data (scale bars, 20 μ m). In the split assay, multiple tumor cells harbored individual 3' side signals (green), indicating the presence of *RET* rearrangement. In the fusion assay, a fusion signal (yellow) was observed in the representative tumor cell of subject RET#2, which is consistent with the presence of *inv*(10)(p11.22q11.2). (b) The break points of *RET* are exons 11 and 12. Both of the break points allow the resulting fusion to harbor the kinase domain of *RET* (red), and the exon 11 break point allows the resulting fusion to harbor the transmembrane domain of *RET* (orange). In the fusion partners, dark blue represents a coiled-coil domain, which probably contributes to the homodimerization of the fusion. Only the longer isoforms of *RET* and the *RET* fusions are shown. (c) Subject RET#14 showed the representative histopathology of mucinous cribriform carcinoma (scale bar, 100 μ m). (d) The results for fusion-specific RT-PCR for subjects ALK#10 (lane 1, EML4-ALK, E13;A20, predicted product size of 432 bp), ALK#16 (lane 2, EML4-ALK, E20;A20, 1185 bp), ALK#26 (lane 3, EML4-ALK, E6;A20, 913 bp), ALK#38 (lane 4, EML4-ALK, E14;ins11del49A20, 546 bp), ALK#39 (lane 5, EML4-ALK, E2;A20, 454 bp), ALK#40 (lane 6, EML4-ALK, E13;ins69A20, 501 bp), ALK#41 (lane 7, EML4-ALK, E14;del14A20, 570 bp), ALK#42 (lane 8, KIF5B-ALK, K17;A20, 1,483 bp), ALK#44 (lane 9, KIF5B-ALK, K24;A20, 814 bp), RET#6 (lane 10, KIF5B-RET, K15;R12, 1,104 bp), RET#9 (lane 11, KIF5B-RET, K16;R12, 1,293 bp), RET#10 (lane 12, KIF5B-RET, K22;R12, 420 bp), RET#11 (lane 13, KIF5B-RET, K23;R12, 525 bp), RET#12 (lane 14, KIF5B-RET, K24;R11, 999 bp) and RET#13 (lane 15, CCDC6-RET, C1;R12, 352 bp). M and N represent the size marker (100-bp ladder) and non-template control, respectively. (e) The transforming potential of the KIF5B-RET fusion. Mouse 3T3 fibroblasts infected with a retrovirus encoding K15;R12L derived from tumor RET#7 formed multiple foci (scale bar, 1 mm). All of the four nude mice injected with the corresponding 3T3 cells developed a subcutaneous tumor (right).



Supplementary Tables 1–4 summarize the clinicopathological features of the subjects. Briefly, young age, low smoking index and small tumor size characterized the kinase-fusion–positive group of subjects (**Supplementary Table 2**). A multivariate analysis of the adenocarcinomas revealed four independent factors that were indicators of poor prognosis: age ≥ 50 years, male sex, high pathological stage and negative kinase-fusion status (**Supplementary Table 3**). There was no significant difference in overall survival between the kinase-positive and epidermal growth factor receptor (EGFR)-mutant groups ($P = 0.32$). **Supplementary Table 4** shows the clinicopathological features of the subjects stratified by each fusion.

The transforming ability of CCDC6-RET and all of the ALK fusions, excluding K17;A20, was shown previously^{1–3,8,12}. 3T3 cells infected with a virus expressing K17;A20, tropomyosin 3 (TPM3)-ROS1, syndecan 4 (SDC4)-ROS1, SLC34A2-ROS1, CD74-ROS1, ezrin (EZR)-ROS1, leucine-rich repeats and immunoglobulin-like domains 3 (LRIG3) (transcript variant 2)-ROS1 or KIF5B-RET (with both the longer (RET51) and shorter (RET9) RET isoforms) led to multiple transformed foci formation in culture and in subcutaneous tumors in a nude mouse tumorigenicity assay (Figs. 1d, 2e and **Supplementary Fig. 5**).

To test whether vandetanib, an inhibitor of vascular endothelial growth factor receptor (VEGFR-2), VEGFR-3, EGFR and RET¹⁴, might be effective for the treatment of RET-fusion–positive tumors, we induced Flag-tagged EML4-ALK (E13;A20) or KIF5B-RET (K15;R12L and K15;R12S) in Ba/F3 cells, which are dependent on interleukin-3 (IL-3) for growth. All transfected cells, including those without any kinase fusion, proliferated in the presence of IL-3, but only cells expressing E13;A20 or K15;R12L grew in the absence of IL-3 (**Supplementary Fig. 6a**). In the absence of IL-3, vandetanib inhibited the proliferation of cells expressing K15;R12L (**Supplementary Fig. 6c**)

but not the proliferation of cells expressing E13;A20 (**Supplementary Fig. 6d**). Crizotinib was not effective in inhibiting the proliferation of Ba/F3 cells expressing K15;R12L (**Supplementary Fig. 7**).

In 1985, a 3T3 assay identified *RET* as a rearranged transforming gene¹⁵. RET fusions have been identified exclusively in papillary thyroid carcinoma and are more frequently observed in radiation-associated thyroid cancers (for example, in survivors of the Chernobyl accident¹⁶, atomic bomb survivors¹⁷ and post-radiation therapy patients¹⁸). Therefore, a retrospective comparison of RET fusions in individuals with lung cancer with and without a history of radiation exposure warrants further study. If a positive association is found between RET fusion and radiation exposure in these studies, it might be desirable for individuals with internal or therapeutic exposure to irradiation (for example, those individuals involved in the Fukushima accident) to be monitored prospectively for lung cancer as well as thyroid cancer.

In Japan, more than 40% of lung adenocarcinomas in younger individuals harbor EGFR mutations¹⁹. In this study, 16% (17 out of 107) of younger individuals (≤ 50 years of age) with adenocarcinoma harbored a kinase fusion. Collectively, as long as molecular target diagnoses are properly performed, $>50\%$ of the individuals with lung adenocarcinoma in this generation may benefit from treatment with corresponding kinase inhibitors. Integrated pathology-based screening techniques can also be used for the selection of individuals to receive this treatment²⁰. The results of our study will facilitate the development of a molecular classification of lung adenocarcinomas that is closely related to both the pathogenesis and the treatment of disease. This study was approved by the Institutional Review Board of the Cancer Institute Hospital, and all subjects provided informed consent.

BRIEF COMMUNICATIONS

METHODS

Methods and any associated references are available in the online version of the paper at <http://www.nature.com/naturemedicine/>.

Note: Supplementary information is available on the Nature Medicine website.

ACKNOWLEDGMENTS

We thank M. Iwakoshi, K. Shiozawa, T. Kakita, H. Nagano and K. Nomura for their technical assistance and S. Sengoku for providing administrative assistance. This work was supported in part by Grants-in-Aid for Scientific Research from the Ministry of Education, Culture, Sports, Science and Technology of Japan, as well as by grants from the Japan Society for the Promotion of Science; the Ministry of Health, Labor and Welfare of Japan; the Vehicle Racing Commemorative Foundation of Japan; the Princess Takamatsu Cancer Research Fund; and the Uehara Memorial Foundation.

AUTHOR CONTRIBUTIONS

K.T. conceived of and led the entire project, designed the FISH probes, screened samples using FISH and immunohistochemistry, performed histopathological analyses, generated figures and tables and wrote the manuscript. M.S. performed functional analyses and generated the figures. Y.T. performed inverse RT-PCR and RACE experiments and their corresponding analyses. R.S. conducted statistical analyses. S.S. performed FISH and histopathological analyses. S.H. processed and analyzed the tissue microarrays and FISH screening and generated figures. R.A. processed the FISH probe library. W.H. made and analyzed the database and processed tissue microarrays. H.N., H.U., Y.S., S.O. and K.N. collected specimens and clinical information and were involved in planning the project. Y.L.C. conducted functional analyses. H.M. supervised the functional analyses and planned the project. Y.I. performed histopathological analyses and

collected specimens. All authors participated in the discussion and interpretation of the data and the results.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

Published online at <http://www.nature.com/naturemedicine/>.

Reprints and permissions information is available online at <http://www.nature.com/reprints/index.html>.

1. Soda, M. *et al. Nature* **448**, 561–566 (2007).
2. Takeuchi, K. *et al. Clin. Cancer Res.* **14**, 6618–6624 (2008).
3. Takeuchi, K. *et al. Clin. Cancer Res.* **15**, 3143–3149 (2009).
4. Rikova, K. *et al. Cell* **131**, 1190–1203 (2007).
5. Takeuchi, K. *et al. Haematologica* **96**, 464–467 (2011).
6. Takeuchi, K. *et al. Clin. Cancer Res.* **17**, 3341–3348 (2011).
7. Sugawara, E. *et al. Cancer* published online, doi:10.1002/cncr.27391 (17 January 2012).
8. Choi, Y.L. *et al. Cancer Res.* **68**, 4971–4976 (2008).
9. Inamura, K. *et al. J. Thorac. Oncol.* **3**, 13–17 (2008).
10. Takeuchi, K. *Pathol. and Clin. Med.* **28**, 139–144 (2010).
11. Jokoji, R. *et al. J. Clin. Pathol.* **63**, 1066–1070 (2010).
12. Grieco, M. *et al. Cell* **60**, 557–563 (1990).
13. Tognon, C. *et al. Cancer Cell* **2**, 367–376 (2002).
14. Flanagan, J., Deshpande, H. & Gettinger, S. *Biologics* **4**, 237–243 (2010).
15. Takahashi, M., Ritz, J. & Cooper, G.M. *Cell* **42**, 581–588 (1985).
16. Ito, T. *et al. Lancet* **344**, 259 (1994).
17. Hamatani, K. *et al. Cancer Res.* **68**, 7176–7182 (2008).
18. Bounacer, A. *et al. Oncogene* **15**, 1263–1273 (1997).
19. Kosaka, T. *et al. Cancer Res.* **64**, 8919–8923 (2004).
20. Han, B. *et al. Cancer Res.* **68**, 7629–7637 (2008).

Identification of Anaplastic Lymphoma Kinase Fusions in Renal Cancer

Large-Scale Immunohistochemical Screening by the Intercalated Antibody-Enhanced Polymer Method

Emiko Sugawara, MD^{1,2}; Yuki Togashi, MS^{1,3}; Naoto Kuroda, MD⁴; Seiji Sakata, MD, PhD¹; Satoko Hatano, BS^{1,3}; Reimi Asaka, BS^{1,3}; Takeshi Yuasa, MD, PhD⁵; Junji Yonese, MD, PhD⁵; Masanobu Kitagawa, MD, PhD²; Hiroyuki Mano, MD, PhD^{6,7}; Yuichi Ishikawa, MD, PhD³; and Kengo Takeuchi, MD, PhD^{1,3}

BACKGROUND: Several promising molecular-targeted drugs are used for advanced renal cancers. However, complete remission is rarely achieved, because none of the drugs targets a key molecule that is specific to the cancer, or is associated with “oncogene addiction” (dependence on one or a few oncogenes for cell survival) of renal cancer. Recently, an anaplastic lymphoma kinase (ALK) fusion, *vinculin-ALK*, has been reported in pediatric renal cell carcinoma (RCC) cases who have a history of sickle cell trait. In this context, ALK inhibitor therapy would constitute a therapeutic advance, as has previously been demonstrated with lung cancer, inflammatory myofibroblastic tumors, and anaplastic large cell lymphomas. **METHODS:** Anti-ALK immunohistochemistry was used to screen 355 tumor tissues, using the intercalated antibody-enhanced polymer (iAEP) method. The cohort consisted of 255 clear cell RCCs, 32 papillary RCCs, 34 chromophobe RCCs, 6 collecting duct carcinomas, 10 unclassified RCCs, 6 sarcomatoid RCCs, and 12 other tumors. **RESULTS:** Two patients (36- and 53-year-old females) were positive for ALK as determined by iAEP immunohistochemistry. Using 5'-rapid amplification of complementary DNA ends, we detected *TPM3-ALK* and *EML4-ALK* in these tumors. The results of this study were confirmed by fluorescence in situ hybridization assays. The 2 ALK-positive RCCs were unclassified (mixed features of papillary, mucinous cribriform, and solid patterns with rhabdoid cells) and papillary subtype. They comprised 2.3% of non-clear cell RCCs (2 of 88) and 3.7% of non-clear cell and nonchromophobe RCCs (2 of 54). **CONCLUSIONS:** The results of this study indicate that ALK fusions also exist in adult RCC cases without uncommon backgrounds. These findings confirm the potential of ALK inhibitor therapy for selected cases of RCC. *Cancer* 2012;000:000-000. © 2012 American Cancer Society.

KEYWORDS: anaplastic lymphoma kinase, molecular-targeted therapy, renal cell carcinoma, immunohistochemistry, intercalated antibody-enhanced polymer.

INTRODUCTION

Renal cancer is one of the major cancers. The incidence and mortality of cases are estimated at 273,518 and 116,368 in the world; 14,963 and 6957 in Japan; and 56,678 and 13,711 in the United States.¹ The 5-year survival rate of patients with localized disease is relatively good: 65% to 93% and 47% to 77% for stages 1 and 2, respectively.² For advanced renal cancers (34%-80% and 2%-20% 5-year survival rates in stages 3 and 4, respectively),² several molecular-targeted drugs have been recently approved by the US Food and Drug Administration. These drugs, which include sunitinib, sorafenib, temsirolimus, everolimus, bevacizumab, pazopanib, and axitinib, are promising. However, none of them targets a key molecule that is specific to the cancer, or is associated with “oncogene addiction” of renal cancer, namely, the dependence on one or a few oncogenes for maintenance of the malignant phenotype and cell survival.

Anaplastic lymphoma kinase (ALK) fusion is a potential vulnerability, an “Achilles’ heel”, of many types of human cancer, including lymphoma,^{3,4} sarcoma,⁵ and carcinoma.^{6,7} Experimentally, lung adenocarcinomas developed in *EML4-ALK* (fusion of ALK with echinoderm microtubule-associated protein like 4) transgenic mice were successfully treated with an ALK inhibitor.⁸ The ALK inhibitor crizotinib has recently been used in patients with lung cancer, inflammatory myofibroblastic tumors (IMTs), or anaplastic large cell lymphomas (ALCLs), which harbor various ALK fusions. The compound showed an 81% response rate in ALK-positive lung cancers defined by at least 2 diagnostic methods,^{9,10} and a

Corresponding author: Kengo Takeuchi, MD, PhD, Pathology Project for Molecular Targets, The Cancer Institute, Japanese Foundation for Cancer Research, 3-8-31 Ariake, Koto, Tokyo 135-8550, Japan; Fax: (81) 3-3570-0230; kentakeuchi-tyk@umin.net

¹Pathology Project for Molecular Targets, The Cancer Institute, Japanese Foundation for Cancer Research, Tokyo, Japan; ²Department of Comprehensive Pathology, Graduate School, Tokyo Medical and Dental University, Tokyo, Japan; ³Division of Pathology, The Cancer Institute, Japanese Foundation for Cancer Research, Tokyo, Japan; ⁴Department of Diagnostic Pathology, Kochi Red Cross Hospital, Kochi City, Kochi, Japan; ⁵Department of Urology, The Cancer Institute Hospital, Japanese Foundation for Cancer Research, Tokyo, Japan; ⁶Division of Functional Genomics, Jichi Medical University, Tochigi, Japan; ⁷Department of Medical Genomics, Graduate School of Medicine, University of Tokyo, Tokyo, Japan

We thank Tomoyo Kakita, Keiko Shiozawa, and Motoyoshi Iwakoshi for their technical assistance, and Sayuri Sengoku for providing administrative assistance.

DOI: 10.1002/cncr.27391, **Received:** October 1, 2011; **Revised:** October 31, 2011; **Accepted:** November 10, 2011, **Published online** in Wiley Online Library (wileyonlinelibrary.com)

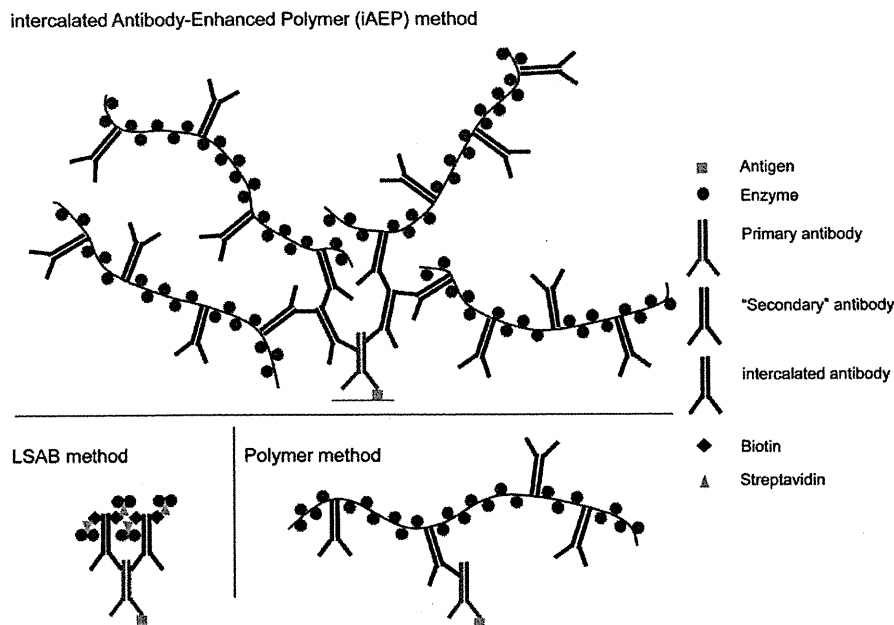


Figure 1. Schematic of intercalated antibody-enhanced polymer (iAEP) method is shown. The labeled streptavidin biotin (LSAB) and polymer methods are common conventional immunohistochemistry methods. In the iAEP method, a step of “intercalated antibody” is added between those of the primary antibody and polymer reagent. Thus, the iAEP method has an additional step compared with the polymer method, but the same number of steps as the LSAB method. There are generally 2 ways to raise the sensitivity of immunohistochemistry. The first is to raise the sensitivity of the antigen-antibody reaction, by increasing the concentration of the primary antibody, using a more sensitive antibody, antigen-retrieval technique, and so forth. The second is to raise the sensitivity of the detection system for the antigen-antibody immune complex. These 2 techniques may appear to generate the same result; however, in principle, they are totally different. The staining results are more likely to differ, especially when the antigen density is very low, such as for EML4-ALK (fusion of echinoderm microtubule-associated protein like 4 with anaplastic lymphoma kinase) or PPFIBP1-ALK (fusion of PTPRF interacting protein binding protein 1 with ALK).^{13,24} In such a setting, the latter technique is more advantageous. The staining intensity depends on the density of enzyme in the antigen site. However sensitive a primary antibody is, the antigen-antibody complex cannot exceed the number of antigens. In contrast, it is easy to increase the enzyme density per antigen-antibody complex with use of the latter technique, which includes the iAEP method.

strong response in IMT for several months.¹¹ Two patients with ALCL who were receiving crizotinib achieved complete remission.¹² These findings indicate that ALK fusion addition is one of the most promising targets in cancer therapy.

To ensure that such molecular-targeted therapy is effective and less toxic, accurate screening methods to detect ALK fusions are crucial. However, although immunohistochemistry has been a gold standard for the detection of ALK fusions in ALCL and IMT,^{13,14} conventional anti-ALK immunohistochemistry is not sensitive enough to detect EML4-ALK, which was first described in lung cancer in 2007.^{6,7} To overcome this, we developed a sensitive immunohistochemical tool, the intercalated antibody-enhanced polymer (iAEP) method (Fig. 1).¹³ Combined with a conventional anti-ALK mouse monoclonal antibody 5A4, the iAEP method efficiently and consistently detected EML4-ALK in paraffin-embedded sections. In various studies on ALK-positive lung cancer,

anti-ALK immunohistochemistry by iAEP or essentially equivalent methods was used to examine surgically resected specimens,^{13,15-19} transbronchial lung biopsy specimens,²⁰ and endobronchial ultrasound-guided transbronchial needle aspiration specimens.^{17,21,22} More importantly, some of the patients screened by anti-ALK iAEP immunohistochemical analysis received crizotinib therapy and showed a good response.^{16,17,22} Novel ALK fusions, including v6 and v7 of EML4-ALK,¹³ kinesin family member 5B (KIF5B)-ALK,¹³ sequestosome 1 (SQSTM1)-ALK,²³ and PTPRF interacting protein, binding protein 1 (PPFIBP1)-ALK²⁴ have been identified using anti-ALK iAEP immunohistochemical analysis. Thus, anti-ALK iAEP immunohistochemistry constitutes a powerful tool for clinical and also research purposes.

The development of anti-ALK antibodies has facilitated the investigation of many types and cases of cancer, including lung cancer.²⁵⁻²⁷ Since 1994, ALK-positive tumors have been identified exclusively in lymphoma