厚生労働科学研究費補助金（創薬基盤推進研究事業）
分担研究報告書
「肺がん検体における大規模 FISH スクリーニング」に関する研究

分担研究者：　竹内　賢吾　　がん研究会がん研究所・主任研究者

研究要旨：フォルマリン固定パラフィン包埋（FFPE）病理検体のみによる新規 ALK 融合同定法の開発を試みた。FFPE 用に特化した 5’-RACE を設計し、抗 ALK iAEP 免疫染色陽性、FISH 法で *ALK* split assay 陽性、*EML4-ALK*、*KIF5B-ALK* fusion assay が陰性であった一症例の FFPE 検体から合成された cDNA に対して施行したところ新規 KLC1-ALK 融合を得た。得られた配列から想定される全長 ALK 融合 cDNA を合成、3T3 細胞にトランスフェクトし、フォーカスフォーメーションアッセイおよびヌードマウスを用いた試験により、KLC1-ALK の造腫瘍能を証明した。

A　研究目的

主任研究者らは 62 才男性肺腺がん患者外科切除検体より cDNA 発現レトロウィルスライブラリーを構築し、マウス 3T3 細胞を用いて形質転換フォーカスのスクリーニングを行った結果、新規がん遺伝子 EML4-ALK を発見することに成功した（Nature 448: 561-566）。これを受け、病理診断医である分担研究者らは、ALK 陽性肺がんの簡便な診断法として multiplex RT-PCR 法（*Clin Cancer Res* 2008;14:6618-6624）、高感度免疫染色法 intercalated antibody-enhanced polymer (iAEP)法（*Clin Cancer Res* 2009;15:3143-3149）および FISH 法を開発した。

これらの診断法が実施された検体のうち、未知の ALK 融合の存在が想定されたにもかかわらず凍結検体が保存されていなかった症例が存在した。本分担研究の今年度の目的は、フォルマリン固定パラフィン包埋（FFPE）病理検体のみによる新規 ALK 融合同定法の開発にある。

B　研究方法

抗 ALK iAEP 免疫染色陽性、FISH 法で *ALK* split assay 陽性、*EML4-ALK*、*KIF5B-ALK* fusion assay がそれぞれ陰性であった一症例、EML4-ALK、KIF5B-ALK 陽性それぞれ一症例ずつの合計 3 つの肺がん FFPE ブロックを用いた。型どおり RNA を FFPE から抽出し合成した cDNA に対し、FFPE 用に特化・設計した 5’-RACE を施行した。得られた配列から想定される全長 ALK 融合 cDNA を合成、3T3 細胞にトランスフェクトし、フォーカスフォーメーションアッセイおよびヌードマウスを用いた造腫瘍能の証明をおこなった。

（倫理面への配慮）

検体収集に関しては自治医科大学の生命倫理委員会認可を受けた事業として開始し、連結可能匿名化のもとで研究を行った。

C　研究結果

ALK の融合点は常に exon 20 の先端であることが経験的にわかっている。そこで、RACE に用いるプライマー群を *ALK* exon 20 の先端に集中させることにより、FFPE ブロックから合成される短い cDNA にも対応できるようにした。

EML4-ALK、KIF5B-ALK 陽性症例の FFPE ブロックから合成された cDNA に対して本法を施行し、産物のシークエンス解析をおこなったところ、それぞれの融合ポイントを含む配列が得られた。そこで、抗 ALK iAEP 免疫染色陽性、FISH 法で *ALK* split assay 陽性、*EML4-ALK*、*KIF5B-ALK* fusion assay がそれぞれ陰性であった一症例に対し本法を施行したところ、*KLC1* exon 9 が *ALK* exon 20 に融合した配列が得られた。

次に、*KLC1-ALK* 融合に特異的な RT-PCR を施行した。その結果、特異的増幅産物を得て、mRNA レベルでの *KLC1-ALK* 発現を証明した。また、BAC clone をもちいて *KLC1* 遺伝子に対するプローブを作成、*KLC1-ALK* fusion FISH assay を FFPE 組織切片に対して施行したところ、融合シグナルが複数の癌細胞に観察され、ゲノムレベルでも *KLC1-ALK* の存在を証明した。

造腫瘍能の証明のために、想定される *KLC1-ALK* 全長 cDNA を人工的に合成し、マウス 3T3 細胞にトランスフェクトした。培養系では、明らかな形質転換フォーカスを形成し、マウスに皮下注射したものは肉眼的に確認できる腫瘤を形成したことにより、KLC1-ALK の造腫瘍能が証明された。

D&E. 考察及び結論

肺がんにおける新規ALK融合KLC1-ALKを同定し、その造腫瘍能を示した。本症例は病理組織学的に細気管支肺胞上皮がんであり、EML4-ALKの存在が報告されていない病理型である。パートナーにより組織型が異なる可能性を提示した事例といえるが、その証明には症例の蓄積が必要である。

また、非浸潤がんにALK融合が発見されたことからは、ALK融合が発がんの極めて早期に生じていることを改めて示した。

FFPE検体のみを用いて新規融合がん遺伝子を同定したという初めての事例であり、世界中で豊富に蓄積されているFFPE検体の潜在的価値を増加させた研究ともいえよう。

F. 健康危険情報

なし

G. 研究発表

1. 論文発表

1. Takeuchi K, Soda M, Togashi Y, Suzuki R, Sakata S, Hatano S, Asaka R, Hamanaka W, Ninomiya H, Uehara H, Lim Choi Y, Satoh Y, Okumura S, Nakagawa K, Mano H, Ishikawa Y. RET, ROS1 and ALK fusions in lung cancer. ***Nat Med***. Epub ahead of print, 2012.
2. Yamamoto M, Takeuchi K, Shimoji M, Maniwa T, Isaka M, Nakagawa K, Ohde Y, Kondo H, Nakajima T. Small non-mucinous bronchioloalveolar carcinoma with anaplastic lymphoma kinase immunoreactivity: A novel ALK fusion gene? ***Cancer Sci***. 2012;103:390-392.
3. Sugawara E, Togashi Y, Kuroda N, Sakata S, Hatano S, Asaka R, Yuasa T, Yonese J, Kitagawa M, Mano H, Ishikawa Y, Takeuchi K. Identification of anaplastic lymphoma kinase fusions in renal cancer: Large-scale immunohistochemical screening by the intercalated antibody-enhanced polymer method. ***Cancer***. 2012.
4. Kimura H, Nakajima T, Takeuchi K, Soda M, Mano H, Iizasa T, Matsui Y, Yoshino M, Shingyoji M, Itakura M, Itami M, Ikebe D, Yokoi S, Kageyama H, Ohira M, Nakagawara A. ALK fusion gene positive lung cancer and 3 cases treated with an inhibitor for ALK kinase activity. ***Lung Cancer***. 2012;75:66-72.
5. Togashi Y, Soda M, Sakata S, Sugawara E, Hatano S, Asaka R, Nakajima T, Mano H, Takeuchi K. KLC1-ALK: A novel fusion in lung cancer identified using a formalin-fixed paraffin-embedded tissue only. ***PLoS One***. 2012;7:e31323.
6. Kijima T, Takeuchi K, Tetsumoto S, Shimada K, Takahashi R, Hirata H, Nagatomo I, Hoshino S, Takeda Y, Kida H, Goya S, Tachibana I, Kawase I. 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***. 2011;102:1602-1604.
7. Takeuchi K, Soda M, Togashi Y, Sugawara E, Hatano S, Asaka R, Okumura S, Nakagawa K, Mano H, Ishikawa Y. Pulmonary inflammatory myofibroblastic tumor expressing a novel fusion, PPFIBP1-ALK: reappraisal of anti-ALK immunohistochemistry as a tool for novel ALK fusion identification. ***Clin Cancer Res***. 2011;17:3341-3348.

H. 知的財産権の出願・登録状況

該当せず

自治医科大学　　間野博行　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
Ueno T, Yamashita Y, Soda M, Fukumura K, Ando M, Yamato A, Kawazu M, Choi YL & Mano H	High-throughput resequencing of target-captured cDNA in cancer cells	Cancer Sci	103	131-135	2012
Togashi Y, Soda M, Sakata S, Sugawara E, Hatano S, Asaka R, Nakajima T, Mano H & Takeuchi K	KLC1-ALK: a novel fusion in lung cancer identified using a formalin-fixed paraffin-embedded tissue only	PLoS ONE	7	e31323	2012
Takeuchi K, Soda M, Togashi Y, Suzuki R, Sakata S, Hatano S, Asaka R, Hamanaka W, Ninomiya H, Uehara H, Choi YL, Satoh Y, Okumura S, Nakagawa K, Mano H & Ishikawa Y	RET, ROS1, and ALK fusions in lung cancer	Nat Med		in press	2012
Sugawara E, Togashi Y, Kuroda N, Sakata S, Hatano S, Asaka R, Yuasa T, Yonese J, Kitagawa M, Mano H, Ishikawa Y & Takeuchi K	Identification of anaplastic lymphoma kinase fusions in renal cancer: Large-scale immunohistochemical screening by the intercalated antibody-enhanced polymer method	Cancer		in press	2012
Ng KP, Hillmer AM, Chuah CTH, Juan WC, Ko T-K, Teo ASM, Ariyaratne PN, Takahashi N, Sawada K, Fei Y, Soh S, Lee WH, Huang JWJ, Allen Jr. JC, Woo XY, Nagarajan N, Kumar V, Thalamuthu A, Poh WT, Ang AL, Mya HT, How GF, Yang LY, Koh LP, Chowbay B, Chang C-T, Nadarajan VS, Chung WJ, Than H, Lim LC, Goh YT, Zhang S, Poh D, Tan P, Seet J-E, Ang M-K, Chau N-M, Ng Q-S, Tan DSW, Soda M, Isobe K, Nothen MM, Wong TY, Shahab A, Ruan X, Cacheux-Rataboul V, Sung W-K, Tan EH, Yatabe Y, Mano H, Soo RA, Chin TM, Lim W-T, Ruan Y & Ong ST	A common BCL2L11 deletion polymorphism mediates intrinsic resistance and inferior responses to tyrosine kinase inhibitors in cancer	Nat Med		in press	2012
Kimura H, Nakajima T, Takeuchi K, Soda M, Mano H, Iizasa T, Matsui Y, Yoshino M, Shingyoji M, Itakura M, Itami M, Ikebe D, Yokoi S, Kageyama H, Ohira M & Nakagawara A	ALK fusion gene positive lung cancer and 3 cases treated with an inhibitor for ALK kinase activity	Lung Cancer	75	66-72	2012
Choi YL, Soda M, Ueno T, Hamada T, Haruta H, Yamato A, Fukumura K, Ando M, Kawazu M, Yamashita Y & Mano H	Oncogenic MAP2K1 mutations in human epithelial tumors	Carcinogenesis		in press	2012

Butler MO, Imataki O, Yamashita Y, Tanaka M, Ansen S, Berezovskaya A, Metzler G, Milstein MI, Mooney MM, Murray AP, Mano H, Nadler LM & Hirano N	Ex vivo expansion of human CD8 T cells using autologous CD4 T cell help	PLoS One	7	e30229	2012
Tominaga-Sato S, Tsushima H, Ando K, Itonaga H, Imaizumi Y, Imanishi 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
Takeuchi K, Soda M, Togashi Y, Sugawara E, Hatano S, Asaka R, Okumura S, Nakagawa K, Mano H & Ishikawa Y	Pulmonary inflammatory myofibroblastic tumor expressing a novel fusion, PPFIBP1-ALK: reappraisal of anti-ALK immunohistochemistry as a tool for novel ALK fusion identification	Clin Cancer Res	17	3341-3348	2011
Takeuchi K, Soda M, Togashi Y, Ota Y, Sekiguchi Y, Hatano S, Asaka R, Noguchi M & Mano H	Identification of a novel fusion, SQSTM1-ALK, in ALK-positive large B-cell lymphoma	Haematologica	96	464-467	2011
Suzuki HI, Arase M, Matsuyama H, Choi YL, Ueno T, Mano H, Sugimoto K & Miyazono K	MCPIP1 ribonuclease antagonizes dicer and terminates microRNA biogenesis through precursor microRNA degradation	Mol Cell	44	424-436	2011
Sasaki D, Imaizumi Y, Hasegawa H, Osaka A, Tsukasaki K, Choi YL, Mano H, Marquez VE, Hayashi T, Yanagihara K, Moriwaki Y, Miyazaki Y, Kamihira S & Yamada Y	Overexpression of Enhancer of zeste homolog 2 with trimethylation of lysine 27 on histone H3 in adult T-cell leukemia/lymphoma as a target for epigenetic therapy	Haematologica	96	712-719	2011
Matsuyama H, Suzuki HI, Nishimori H, Noguchi M, Yao T, Komatsu N, Mano H, Sugimoto K & Miyazono K	miR-135b mediates NPM-ALK-driven oncogenicity and renders IL-17-producing immunophenotype to anaplastic large cell lymphoma	Blood	118	6881-6892	2011
Iida A, Shinoe T, Baba Y, Mano H & Watanabe S	Dicer plays essential roles for retinal development by regulation of survival and differentiation	Invest Ophthalmol Vis Sci	52	3008-3017	2011

自治医科大学　　杉山幸比古　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
Mizushina Y, Bando M, Hosono T, Mato N, Nakaya T, Ishii Y, Yamasawa H, Sugiyama Y	Clinical features of lymphangioleiomyomatosis complicated by renal angiomyolipomas	Intern Med	50	285-289	2011
Taniguchi H, Kondoh Y, Ebina M, Azuma A, Ogura T, Taguchi Y, Suga M, Takahashi H, Nakata K, Sato A, Sugiyama Y, Kodoh S, Nukiwa T, Pirfenidone Clinical Study Group in Japan	The clinical significance of 5% change in vital capacity in patients with idiopathic pulmonary fibrosis: extended analysis of the pirfenidone trial	Respir Res	12	93	2011
Bando M, Miyazawa T, Sugiyama Y, Shinohara H, Owada T, Terakado M	An epidemiological study of the effects of statin use on airflow limitation in patients with chronic obstructive pulmonary disease	Respirology		Epub ahead of print	2011
Uto T, Bando M, Yamauchi H, Nakayama M., Ohata, M, Mato N, Nakaya T, Yamasawa H, Kawai T, Sugiyama, Y	Primary cardiac angiosarcoma of the right auricle with difficult-to-treat bilateral pleural effusion	Intern Med	50	2371-2374	2011

東京大学　　崔　永林　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
Ueno T, Yamashita Y, Soda M, Fukumura K, Ando M, Yamato A, Kawazu M, Choi YL & Mano H	High-throughput resequencing of target-captured cDNA in cancer cells	Cancer Sci	103	131-135	2012
Takeuchi K, Soda M, Togashi Y, Suzuki R, Sakata S, Hatano S, Asaka R, Hamanaka W, Ninomiya H, Uehara H, Choi YL, Satoh Y, Okumura S, Nakagawa K, Mano H & Ishikawa Y	RET, ROS1, and ALK fusions in lung cancer	Nat Med	in press		2012
Choi YL, Soda M, Ueno T, Hamada T, Haruta H, Yamato A, Fukumura K, Ando M, Kawazu M, Yamashita Y & Mano H	Oncogenic MAP2K1 mutations in human epithelial tumors	Carcinogenesis	in press		2012
Suzuki HI, Arase M, Matsuyama H, Choi YL, Ueno T, Mano H, Sugimoto K & Miyazono K	MCPIP1 ribonuclease antagonizes dicer and terminates microRNA biogenesis through precursor microRNA degradation	Mol Cell	44	424-436	2011
Sasaki D, Imaizumi Y, Hasegawa H, Osaka A, Tsukasaki K, Choi YL, Mano H, Marquez VE, Hayashi T, Yanagihara K, Moriwaki Y, Miyazaki Y, Kamihira S & Yamada Y	Overexpression of Enhancer of zeste homolog 2 with trimethylation of lysine 27 on histone H3 in adult T-cell leukemia/lymphoma as a target for epigenetic therapy	Haematologica	96	712-719	2011

東京大学　　鯉沼代造　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
Horiguchi K, Sakamoto K, Koinuma D, Semba K, Inoue A, Inoue S, Fujii H, Yamaguchi A, Miyazawa K, Miyazono K & Saitoh M.	TGF-beta drives epithelial-mesenchymal transition through deltaEF1-mediated downregulation of ESRP	Oncogene	in press		2011
Kawata M, Koinuma D, Ogami T, Umezawa K, Iwata C, Watabe T & Miyazono K.	TGF-beta-induced epithelial-mesenchymal transition of A549 lung adenocarcinoma cells is enhanced by proinflammatory cytokines derived from RAW 264.7 macrophage cells	J Biochem	151	205-206	2012
Miyazono K, Ehata S & Koinuma D.	Tumor-promoting functions of transforming growth factor-beta in progression of cancer	Ups J Med Sci	in press		2011
Koinuma D, Shinozaki M, Nagano Y, Ikushima H, Horiguchi K, Goto K, Chano T, Saitoh M, Imamura T, Miyazono K & Miyazawa K.	RB1CC1 Protein Positively Regulates Transforming Growth Factor-beta Signaling through the Modulation of Arkadia E3 Ubiquitin Ligase Activity	J Biol Chem	286	32502-32512	2011
Mizutani A, Koinuma D, Tsutsumi S, Kamimura N, Morikawa M, Suzuki HI, Imamura T, Miyazono K & Aburatani H.	Cell type-specific target selection by combinatorial binding of Smad2/3 proteins and hepatocyte nuclear factor 4alpha in HepG2 cells	J Biol Chem	286	29848-29860	2011
Morikawa M, Koinuma D, Tsutsumi S, Vasilaki E, Kanki Y, Heldin CH, Aburatani H & Miyazono K.	ChIP-seq reveals cell type-specific binding patterns of BMP-specific Smads and a novel binding motif	Nucleic Acids Res	39	8712-8727	2011

がん研究会　　竹内賢吾　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
Takeuchi K, Soda M, Togashi Y, Suzuki R, Sakata S, Hatano S, Asaka R, Hamanaka W, Ninomiya H, Uehara H, Lim Choi Y, Satoh Y, Okumura S, Nakagawa K, Mano H, Ishikawa Y.	RET, ROS1 and ALK fusions in lung cancer.	Nat Med.			on line
Yamamoto M, Takeuchi K, Shimoji M, Maniwa T, Isaka M, Nakagawa K, Ohde Y, Kondo H, Nakajima T.	Small non-mucinous bronchioloalveolar carcinoma with anaplastic lymphoma kinase immunoreactivity: A novel ALK fusion gene?	Cancer Sci	103	390-392	2012
Sugawara E, Togashi Y, Kuroda N, Sakata S, Hatano S, Asaka R, Yuasa T, Yonese J, Kitagawa M, Mano H, Ishikawa Y, Takeuchi K.	Identification of anaplastic lymphoma kinase fusions in renal cancer: Large-scale immunohistochemical screening by the intercalated antibody-enhanced polymer method.	Cancer			on line
Kimura H, Nakajima T, Takeuchi K, Soda M, Mano H, Iizasa T, Matsui Y, Yoshino M, Shingyoji M, Itakura M, Itami M, Ikebe D, Yokoi S, Kageyama H, Ohira M, Nakagawara A.	ALK fusion gene positive lung cancer and 3 cases treated with an inhibitor for ALK kinase activity.	Lung Cancer	75	66-72	2012
Togashi Y, Soda M, Sakata S, Sugawara E, Hatano S, Asaka R, Nakajima T, Mano H, Takeuchi K.	KLC1-ALK: A novel fusion in lung cancer identified using a formalin-fixed paraffin-embedded tissue only.	PLoS One	7	e31323	2012
Kijima T, Takeuchi K, Tetsumoto S, Shimada K, Takahashi R, Hirata H, Nagatomo I, Hoshino S, Takeda Y, Kida H, Goya S, Tachibana I, Kawase I.	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	2011
Takeuchi K, Soda M, Togashi Y, Sugawara E, Hatano S, Asaka R, Okumura S, Nakagawa K, Mano H, Ishikawa Y.	Pulmonary inflammatory myofibroblastic tumor expressing a novel fusion, PPFIBP1-ALK: reappraisal of anti-ALK immunohistochemistry as a tool for novel ALK fusion identification.	Clin Cancer Res.	17	3341-3348	2011

Cancer Science
The official journal of the Japanese Cancer Association JCA

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]

[1]Division of Functional Genomics, Jichi Medical University, Tochigi; [2]Department of Medical Genomics, Graduate School of Medicine, University of Tokyo, Tokyo; [3]CREST 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

[4]To whom correspondence should be addressed. E-mail: hmano@jichi.ac.jp
[5]These authors contributed equally to this work.

doi: 10.1111/j.1349-7006.2011.02105.x

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 ≥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′-CCTGTGGGTAGGGTTTCCAAAGAG-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 fragments, 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 ~20% of total reads, and 4.55% accounted for ~50% of reads. Thus, as expected, resequencing data for unselected cDNA consist mostly of reads corresponding to a limited number of highly-abundant transcripts.

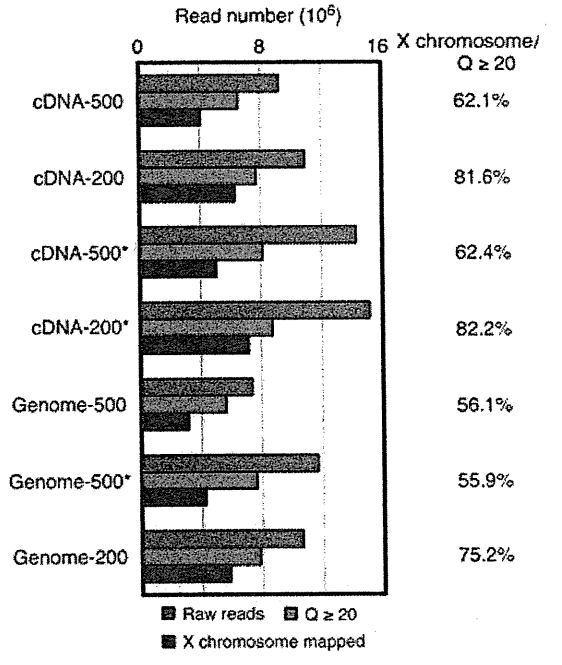

Fig. 1. Comparison of capture efficiency between cDNA and genomic DNA. Genomic DNA or cDNA of KCL-22-SR cells was fragmented to a mean size of 200 or 500 bp, and then subjected to purification with the SureSelect Human X Chromosome Demo kit, followed by GAIIx sequencing at a concentration of 4 or 8 pM (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.

doi: 10.1111/j.1349-7006.2011.02105.x

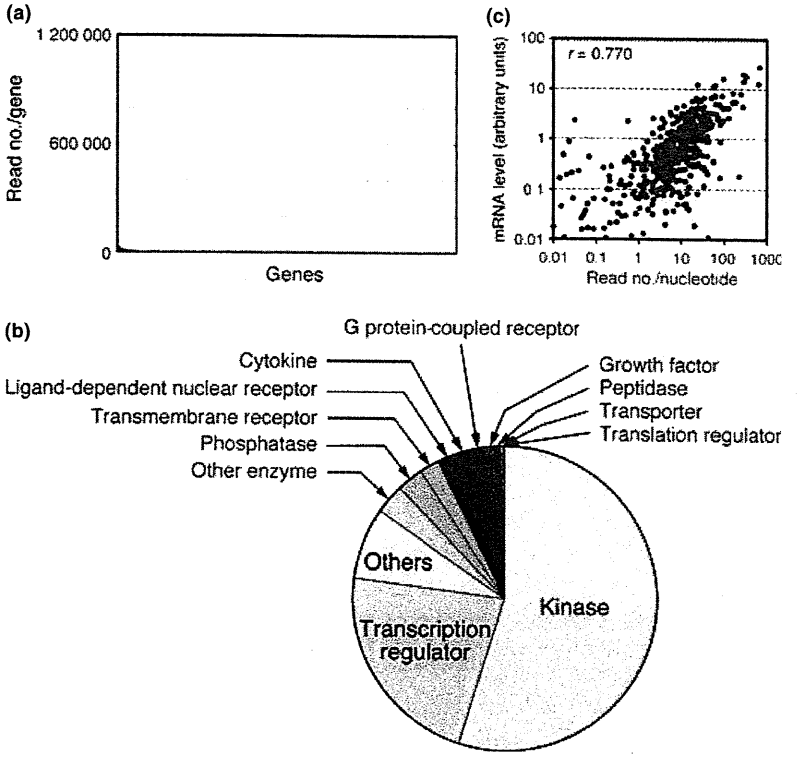

Fig. 2. Capture of a selected set of cDNA. (a) Read number for each gene was calculated from the sequencing data for the unselected cDNA of KCL-22-SR cells. Genes were sorted according to their read number. A small number of genes accounted for most of the sequence reads. (b) Functional annotation for 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

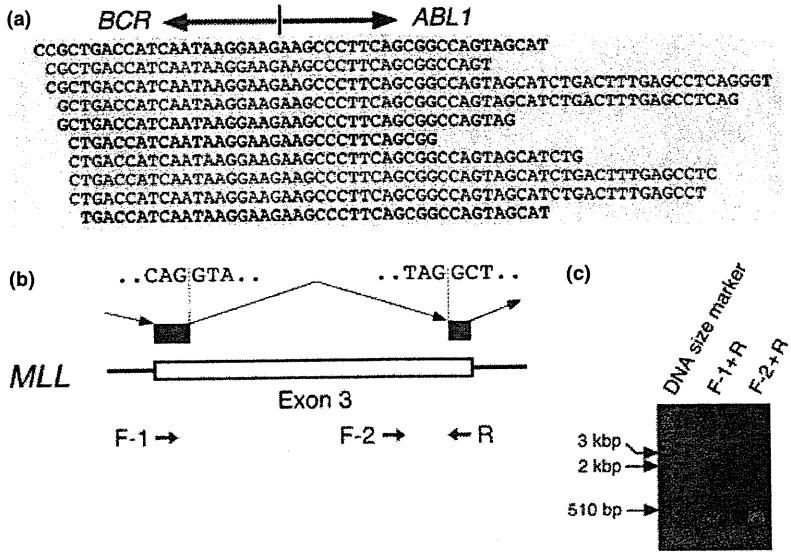

Fig. 3. Detection of gene fusions and alternative mRNA splicing in CML cells. (a) Our computational pipeline yielded 45 reads for KCL-22-SR cells that encompassed the fusion point of breakpoint cluster region (*BCR*)-Abelson murine leukemia viral oncogene homolog 1 (*ABL1*) cDNA, some of which are shown aligned. Reads in the sense or antisense strand are designated in black and blue letters, respectively, and the *BCR* and *ABL1* portions of the sequences are shaded differentially. (b) Some of the reads that mapped to exon 3 of mixed-lineage leukemia (*MLL*) skipped a 2193-bp region within this exon. Nucleotide sequences of the cryptic splicing sites are shown, as are the positions of PCR primers used to confirm the alternative splicing. (c) Gel electrophoresis of the RT–PCR products obtained with total RNA isolated from KU812 cells and with either the F-1 and R primer pair or the F-2 and R primer pair. A 1-kb ladder of DNA size markers was also included.

(343 bp) products, whereas that of the F-2 and R primers would yield only the wild-type product of 339 bp. Gel electrophoresis of the RT–PCR products confirmed the presence of the truncated mRNA (Fig. 3c). Given that the donor and acceptor sites for this alternative splicing harbor the consensus sequences for mRNA splicing (Fig. 3b), some CML cells likely make use of such cryptic splicing sites after *MLL* transcription.

Other variants. From the captured cDNA for KCL-22-SR, NCO2, MEG-01s, K562, and KU812 cells, we detected 156, 18, 28, 23, and 21 non-synonymous mutations among the 913 target genes, respectively. An analysis of the unselected cDNA from KCL-22-SR, however, identified only 19 mutations within the target genes, 16 of which were discovered in the captured cDNA as well. Comparison of the read sequences from the unselected KCL-22-SR cDNA to all RefSeq exonic sequences discovered a total of 597 non-synonymous mutations.

Furthermore, 19, eight, four, 11, and two indels were detected with the captured cDNA of KCL-22-SR, NCO2, MEG-01s, K562, and KU812, respectively. Most of the detected indels were only 1 bp in length, whereas the others were either 2 or 3 bp (Fig. S3). Detailed analysis of these nucleotide changes will be described elsewhere (Toshihide Ueno and Yoshihiro Yamashita, personal communication).

One of the most frequent genetic changes in the blast crisis of CML is point mutation or loss (or both) of *TP53*.[21] Indeed, our sequence data for this gene revealed non-synonymous point mutations in NCO2 and KU812 cells, a 1-bp insertion in K562 cells, a 1-bp deletion in KCL-22-SR cells, and a 3-bp deletion in MEG-01s cells (Fig. 4; Fig. S4; Table S3), all of which were confirmed by Sanger sequencing (data not shown). In NCO2 cells, for instance, 100% of *TP53* reads harbored a G-to-C substitution at nucleotide position 993 of *TP53* mRNA (GenBank accession no.: NM_000546), resulting in a glycine-to-arginine amino acid change (Fig. 4a). The data were also indicative of loss of heterozygosity for *TP53* in NCO2 cells. Similarly, 75% or 78% of *TP53* reads contained a C insertion or a CAC deletion in K562 (Fig. 4b) or MEG-01s (Fig. S4) cells, respectively.

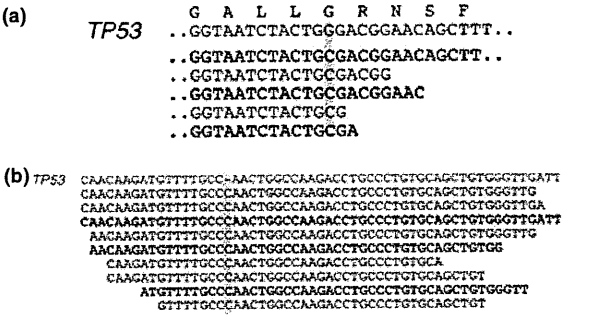

Fig. 4. Anomalies in *TP53* in CML cell lines. (a) Read sequences for NCO2 cells are shown aligned with the reference nucleotide and predicted amino acid sequences (red letters) for *TP53*, revealing a G-to-C substitution in all the reads. Sense or antisense strands are denoted in black and blue letters, respectively. (b) Alignment of the read sequences for K562 cells with the cDNA sequence of *TP53* as in (a), revealing a C insertion.

Discussion

We have shown that a cDNA-capture system, coupled with massively parallel sequencing, is a feasible and relatively simple approach to the simultaneous detection of point mutations, indels, and gene fusions in target cDNA. There are, however, both advantages and disadvantages of this technique compared with the conventional exon-capture system for genomic DNA.

The ability to detect gene fusions, in addition to other mutations with a single sequencing reaction, is one of the most important benefits of the cDNA-capture approach. Furthermore, the efficiency of exon capture with genomic DNA is dependent on the sequence context of each exon. The mean exon size for the human genome is only <200 bp, and the efficiency of exon purification is markedly affected by GC content and sequence complexity.[22] In contrast, even exons with a high GC content might be well isolated by the cDNA-capture system if adjacent exons have a normal GC content and are efficiently targeted by hybridization probes.

Levin *et al.*[20] conducted deep sequencing of captured cDNA for K562 cells, and identified five candidates for fusion genes in addition to *BCR–ABL1*. However, we could not detect any of the five candidates through our analysis with K562, probably because our 913 target genes did not contain those involved in the gene fusions in their report, other than nascent polypeptide-associated complex alpha subunit (*NACA*). While Levin *et al.* discovered primase, DNA, polypeptide 1 (*PRIM1*)–*NACA* fusion transcripts, the low expression level of *PRIM1–NACA* in K562 (only 2.5% of that of *BCR–ABL1* in their dataset)[20] might account for the failure in our analysis.

doi: 10.1111/j.1349-7006.2011.02105.x

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

1 Stratton MR, Campbell PJ, Futreal PA. The cancer genome. *Nature* 2009; **458**: 719–24.
2 Mardis ER. A decade's perspective on DNA sequencing technology. *Nature* 2011; **470**: 198–203.
3 Ledford H. Big science: the cancer genome challenge. *Nature* 2010; **464**: 972–4.
4 Pleasance ED, Cheetham RK, Stephens PJ *et al.* A comprehensive catalogue of somatic mutations from a human cancer genome. *Nature* 2010; **463**: 191–6.
5 Ley TJ, Mardis ER, Ding L *et al.* DNA sequencing of a cytogenetically normal acute myeloid leukaemia genome. *Nature* 2008; **456**: 66–72.
6 Totoki Y, Tatsuno K, Yamamoto S *et al.* High-resolution characterization of a hepatocellular carcinoma genome. *Nat Genet* 2011; **43**: 464–9.
7 Wei X, Walia V, Lin JC *et al.* Exome sequencing identifies GRIN2A as frequently mutated in melanoma. *Nat Genet* 2011; **43**: 442–6.
8 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.
9 Bilguvar K, Ozturk AK, Louvi A *et al.* Whole-exome sequencing identifies recessive WDR62 mutations in severe brain malformations. *Nature* 2010; **467**: 207–10.
10 Mitelman F. Recurrent chromosome aberrations in cancer. *Mutat Res* 2000; **462**: 247–53.
11 Soda M, Choi YL, Enomoto M *et al.* Identification of the transforming *EML4–ALK* fusion gene in non-small-cell lung cancer. *Nature* 2007; **448**: 561–6.
12 Tomlins SA, Rhodes DR, Perner S *et al.* Recurrent fusion of TMPRSS2 and ETS transcription factor genes in prostate cancer. *Science* 2005; **310**: 644–8.
13 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.
14 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.
15 Choi YL, Tsukasaki K, O'Neill MC *et al.* A genomic analysis of adult T-cell leukemia. *Oncogene* 2007; **26**: 1245–55.
16 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.
17 Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. *J Mol Biol* 1990; **215**: 403–10.
18 Li H, Durbin R. Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics* 2009; **25**: 1754–60.
19 Fujita PA, Rhead B, Zweig AS *et al.* The UCSC Genome Browser database: update 2011. *Nucleic Acids Res* 2011; **39**: D876–82.
20 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.
21 Calabretta B, Perrotti D. The biology of CML blast crisis. *Blood* 2004; **103**: 4010–22.
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.
23 Adzhubei IA, Schmidt S, Peshkin L *et al.* A method and server for predicting damaging missense mutations. *Nat Methods* 2010; **7**: 248–9.
24 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.

OPEN ACCESS Freely available online

PLoS one

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

Yuki Togashi[1,2], Manabu Soda[3], Seiji Sakata[1], Emiko Sugawara[1,4], Satoko Hatano[1,2], Reimi Asaka[1,2], Takashi Nakajima[5], 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

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-tky@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)

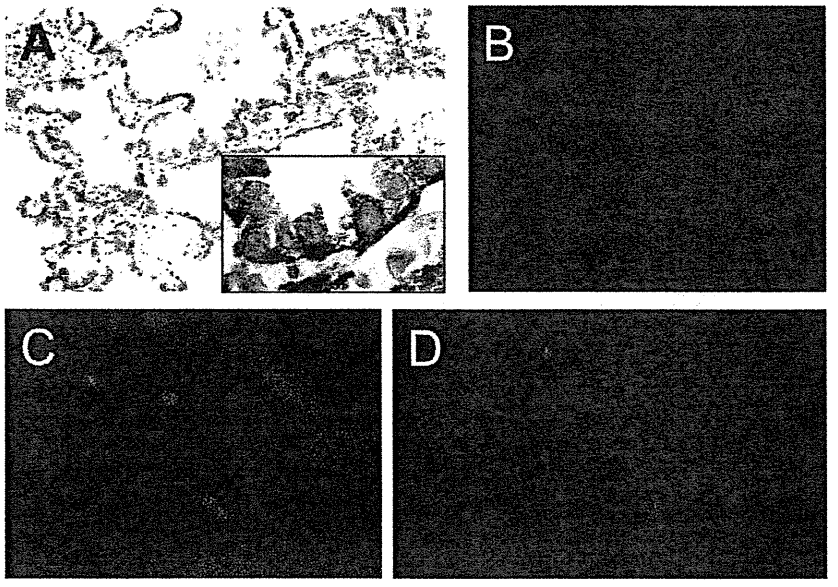

Figure 1. ALK-rearranged lung adenocarcinoma without EML4-ALK and KIF5B-ALK. Panel A shows the results of anti-ALK immunohistochemistry with the iAEP method on pulmonary adenocarcinoma in situ, nonmucinous. The staining pattern was diffusely cytoplasmic. The basal side of tumor cells was more strongly 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′-GCGCTAGCGAATGTATGACAACATGTCCAC-3′) and KLC1-bpR (5′-GTGCTTCCGGCGGTACACATCTACAGAACCAAACTC-3′), and ALK-bpF (5′-GGGAGTTTGGTTCTGTAGATGTGTACCGCCGGAAGC-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

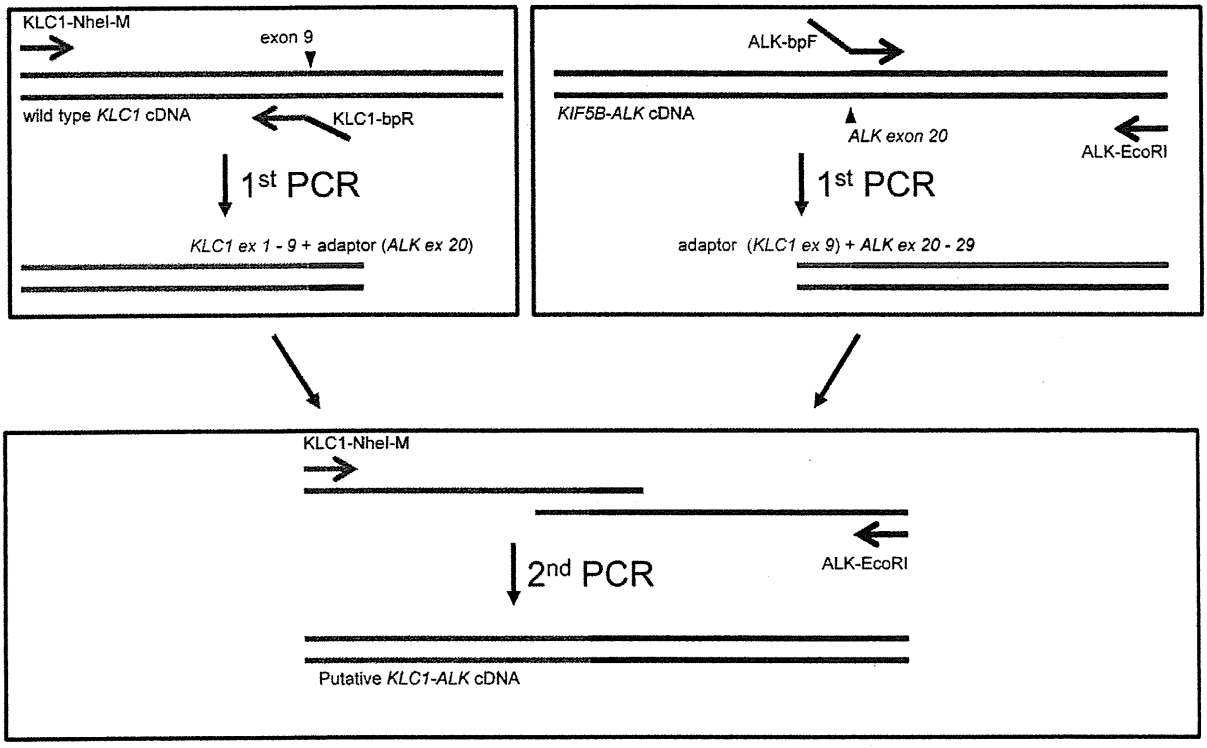

Figure 2. Synthesis of the putative KLC1-ALK full-length cDNA. Two first-round PCRs were performed separately using cDNA synthesized from a tumor tissue expressing KIF5B-ALK with the following primer sets: KLC1-NheI-M and KLC1-bpR, and ALK-bpF and ALK-EcoRI. KLC1-bpR and ALK-bpF had sequences downstream of the ALK break point (exon 20) and upstream of the KLC1 break point (exon 9) as adopter 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;insl1;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

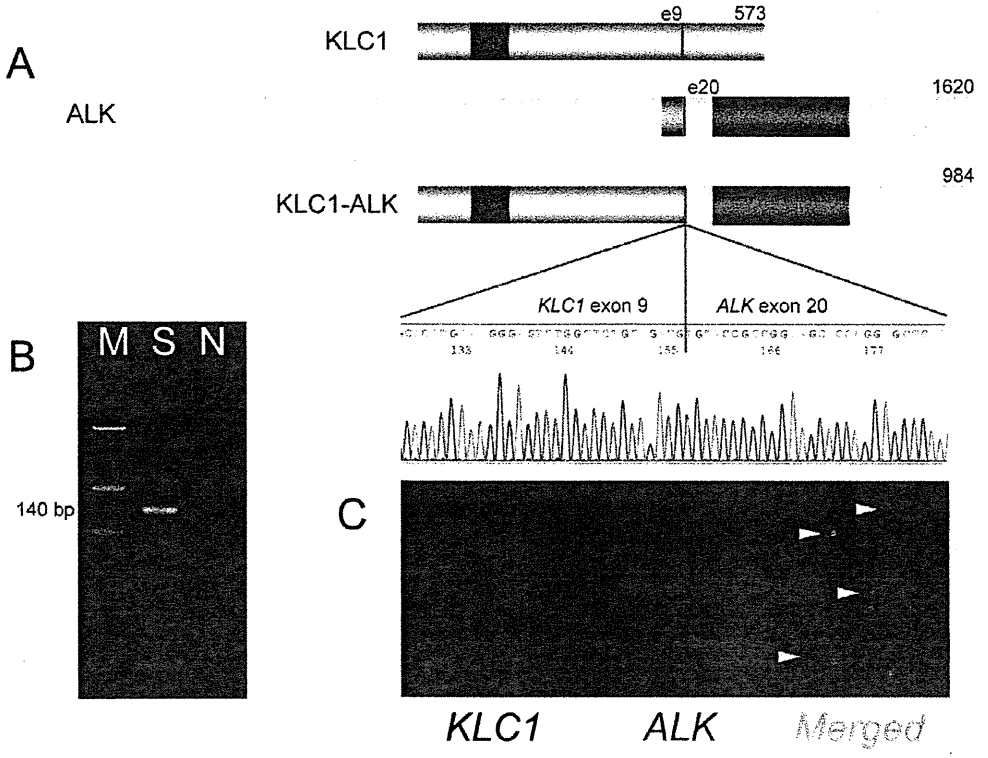

Figure 3. Identification of KLC1-ALK. Panel A shows the schematic structure of KLC1, ALK, and KLC1-ALK proteins and the cDNA sequence around the fusion point. Dark blue, orange, and red parts represent coiled-coil, transmembrane, and kinase domains, respectively. The break point exons and the number 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.

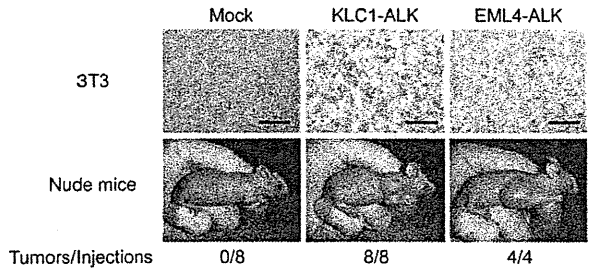

Figure 4. Transforming potential of KLC1-ALK. Upper panels: Mouse 3T3 fibroblasts were infected with retroviruses encoding KLC1-ALK or EML4-ALK or with the corresponding empty virus (Mock). The cells were photographed after 4 days of culture. Scale bar, 1 mm. Lower panels: Nude mice were injected subcutaneously 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

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.

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)

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.

References

1. Morris SW, Kirstein MN, Valentine MB, Dittmer 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.
2. 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.
3. 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.
4. 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.
5. 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.
6. 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.
7. 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.
8. Gambacorti-Passerini C, Messa C, Pogliani EM (2011) Crizotinib in anaplastic large-cell lymphoma. N Engl J Med 364: 775–776.
9. 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.
10. 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.
11. Takeuchi K, Choi YL, Togashi Y, Soda M, Hatano S, et al. (2009) KIF5B-ALK, a novel fusion oncokinase identified by an immunohistochemistry-based diagnostic system for ALK-positive lung cancer. Clin Cancer Res 15: 3143–3149.
12. 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.
13. 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.
14. 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.
15. 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.
16. 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.
17. 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.
18. 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.
19. Cools J, Wlodarska I, Somers R, Mentens N, Pedeutour 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.
20. 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.
21. 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.
22. 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.
23. Debiec-Rychter M, Marynen P, Hagemeijer A, Pauwels P (2003) ALK-ATIC fusion in urinary bladder inflammatory myofibroblastic tumor. Genes Chromosomes Cancer 38: 187–190.
24. 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.
25. 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.
26. 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.
27. 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.
28. 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.
29. 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.
30. 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.
31. Du XL, Hu H, Lin DC, Xia SH, Shen XM, et al. (2007) Proteomic profiling of proteins dysregulted in Chinese esophageal squamous cell carcinoma. J Mol Med 85: 863–875.
32. Jazii 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.
33. 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.
34. Martelli MP, Sozzi G, Hernandez L, Pettirossi 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.
35. Takeuchi K, Soda M, Togashi Y, Sugawara E, Hatano S, et al. (2011) Pulmonary inflammatory myofibroblastic tumor expressing a novel fusion, PPFIBP1-ALK: reappraisal of anti-ALK immunohistochemistry as a tool for novel ALK fusion identification. Clin Cancer Res 17: 3341–3348.
36. 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.
37. 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.
38. 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.
39. 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.
40. Onishi M, Kinoshita S, Morikawa Y, Shibuya A, Phillips J, et al. (1996) Applications of retrovirus-mediated expression cloning. Exp Hematol 24: 324–329.
41. Stenoien DL, Brady ST (1997) Immunochemical analysis of kinesin light chain function. Mol Biol Cell 8: 675–689.

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.