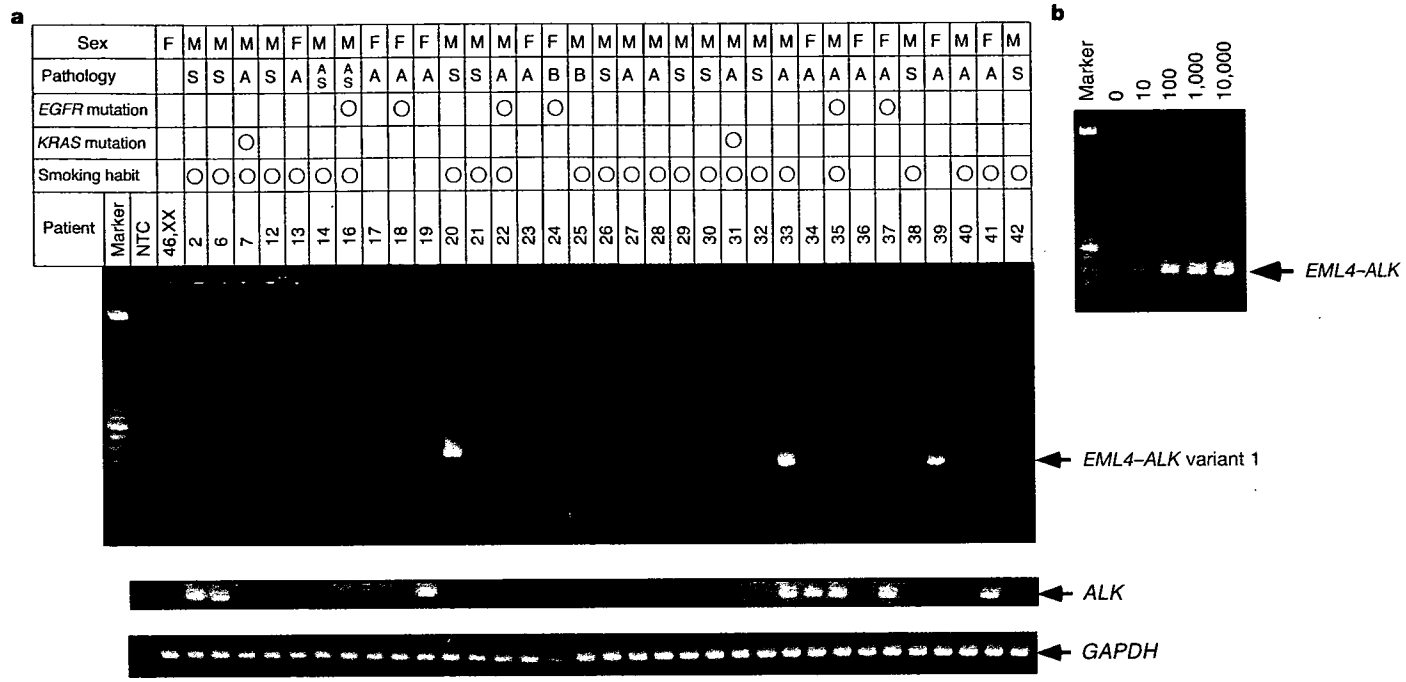

Figure 3 | Screening of NSCLC specimens for *EML4–ALK* variant 1 mRNA. **a**, A consecutive panel of NSCLC specimens ($n = 33$) was subjected to RT–PCR with the Fusion-RT primers (Supplementary Fig. 1). Patients positive for the *EML4–ALK* (variant 1) PCR product are shown in red. Peripheral blood mononuclear cells from a healthy female (46,XX) were also examined as a negative control. RT–PCR for wild-type *ALK* mRNA (with a primer set corresponding to the extracellular domain of ALK) and for glyceraldehyde-3-phosphate dehydrogenase (*GAPDH*) mRNA is also shown. Patient characteristics (sex, pathological classification of NSCLC, the presence of *EGFR* or *KRAS* mutations, and smoking habit) are indicated at the top. A, adenocarcinoma; AS, adenosquamous carcinoma; B, bronchiolo-alveolar carcinoma; F, female; M, male; NTC, no-template control; S, squamous carcinoma. Marker, 50-bp DNA ladder. **b**, Sputum (1 ml) was mixed with 0, 10, 100, 1,000 or 10,000 BA/F3 cells expressing *EML4–ALK* (variant 1) and was then subjected to RT–PCR with the Fusion-RT primer set for detection of *EML4–ALK* mRNA. Marker, 50-bp DNA ladder.

ALK, as variant 2. Multiple variants of the TRK-fused gene (*TFG*)–*ALK* fusion gene associated with anaplastic large-cell lymphoma have also been identified[19].

Given the head-to-head orientation of *EML4* and *ALK* on chromosome 2, RT–PCR with the primers used in our study would not be expected to yield specific products in normal tissues or in any cancers that do not harbour the fusion gene. RT–PCR for *EML4–ALK* mRNA may thus provide a highly sensitive means for detection of lung cancer (with the corresponding chromosomal rearrangement). Although cytological examination of sputum is a reliable method for detection of lung cancer, it is usually only effective at advanced stages. Detection of *EGFR* mutations in sputum is also problematic because NSCLC cells often constitute only a small fraction of cells within the specimen. In contrast, RT–PCR would be expected to detect the presence of only a few cells harbouring the *EML4–ALK* fusion gene among the tens of thousands of non-cancerous cells in sputum. To examine this issue, we established mouse BA/F3 cells[20] that express Flag-tagged EML4–ALK, mixed various numbers of these cells with control sputum, and subjected the mixtures to RT–PCR for detection of *EML4–ALK* mRNA. The fusion mRNA was detected in sputum containing as few as ten BA/F3 cells per ml (Fig. 3b). Early diagnosis of NSCLC (of the *EML4–ALK*$^+$ subtype) by RT–PCR may thus be realistic, as is clinical detection of *Mycobacterium tuberculosis* in sputum by PCR[21]. It is also possible that *EML4–ALK* mRNA would be detected by RT–PCR using pleural effusion, bronchoalveolar lavage, lung biopsy or peripheral blood specimens of patients with lung cancer.

EML4–ALK as a potential therapeutic target

Several small compounds have recently been shown to inhibit the kinase activity of ALK and to suppress the growth of cells expressing NPM–ALK[17,22,23]. To investigate whether use of such ALK inhibitors might be an effective treatment for EML4–ALK$^+$ NSCLC, we expressed Flag-tagged ALK, EML4–ALK or EML4–ALK(K589M) in BA/F3 cells, which are dependent on interleukin-3 (IL-3) for growth. Whereas all transfected cells grew exponentially in the presence of IL-3, only those expressing EML4–ALK proliferated at a similar rate in the absence of IL-3 (Fig. 4a), again confirming the kinase-dependent oncogenic activity of EML4–ALK.

In the presence of IL-3, BA/F3 cells proliferate in a manner dependent on the activity of the tyrosine kinase Janus kinase 2 (JAK2, ref. 24). Addition of a chemical inhibitor (WHI-P154)[22] of ALK to the culture medium affected the IL-3-dependent growth of BA/F3 cells only slightly at concentrations $\geq 5\ \mu M$ (Fig. 4b). Given that WHI-P154 was originally developed as a specific inhibitor of JAK3, it might be expected to show a weak cross-reactivity with JAK2, possibly accounting for the small effect on the JAK2-dependent growth of BA/F3 cells. In contrast, WHI-P154 markedly inhibited the growth of BA/F3 cells expressing EML4–ALK, which do not require IL-3 for proliferation (Fig. 4b). At a concentration of $10\ \mu M$, WHI-P154 rapidly induced the death of these cells in the absence of IL-3. Consistent with these observations, immunoblot analysis revealed that WHI-P154 inhibited the tyrosine phosphorylation of EML4–ALK (on the residue corresponding to Tyr 1604 of wild-type ALK) in a concentration-dependent manner in the transfected BA/F3 cells (Fig. 4c).

Discussion

Using retrovirus-mediated expression screening, we have identified an oncogene, *EML4–ALK*, in a specimen of NSCLC. The 75 NSCLC patients examined in the present study (5 of whom were positive for *EML4–ALK*) were all Japanese. Given that the association of *EGFR* mutations with lung cancer is most prominent in Asian populations[5], it will be important to examine the association of *EML4–ALK* with lung cancer in other ethnic groups. Our data obtained with WHI-P154 suggest that inhibition of the tyrosine kinase activity of

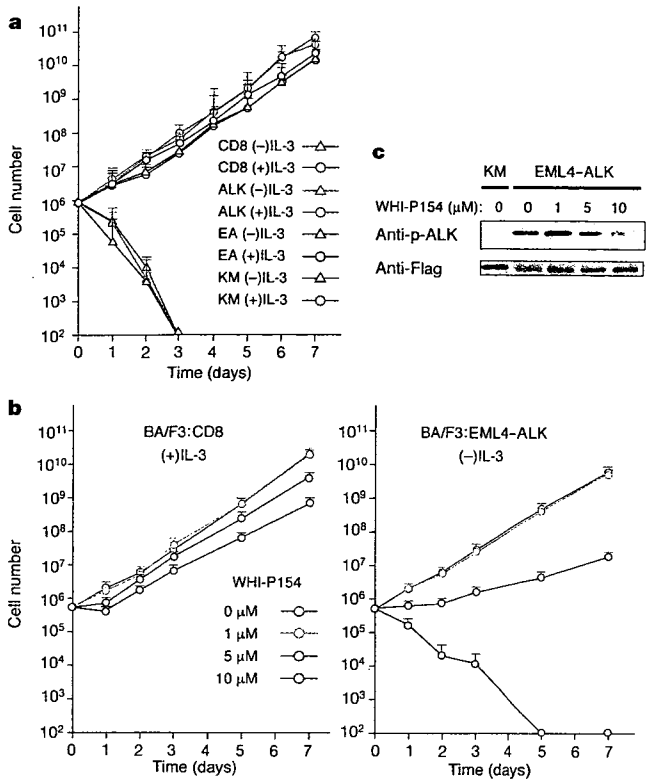

Figure 4 | Inhibition of the growth of BA/F3 cells expressing EML4–ALK variant 1 by a chemical inhibitor of ALK. a, Mouse BA/F3 cells expressing CD8 either alone or together with wild-type ALK, EML4–ALK (EA) or EML4–ALK(K589M) (KM) were cultured in the absence (−) or presence (+) of IL-3 (1 ng ml^{-1}). Cell number was determined at the indicated times. Data are means plus s.d. of values from three separate experiments. **b**, BA/F3 cells expressing CD8 alone were cultured with IL-3 and 0, 1, 5 or 10 μM WHI-P154 (left panel), or those expressing CD8 and EML4–ALK were incubated without IL-3 but with 0, 1, 5 or 10 μM WHI-P154 (right panel). Cell number was determined at the indicated times. Data are means plus s.d. of values from three separate experiments. **c**, BA/F3 cells expressing Flag-tagged EML4–ALK or EML4–ALK(K589M) were incubated with the indicated concentrations of WHI-P154 for 3.5 h, after which total cell lysates (25 μg of protein per lane) were subjected to immunoblot analysis with antibodies to tyrosine-phosphorylated ALK (p-ALK) or to Flag.

EML4–ALK may induce cell death in tumours expressing this fusion protein. Given the lack of apparent phenotypes in *Alk* knockout mice[25], suppression of EML4–ALK function with ALK inhibitors might be expected to be free of severe side effects in NSCLC patients. Furthermore, given that the population of patients who harbour *EGFR* mutations is distinct from that which harbours the *EML4–ALK* fusion gene, ALK inhibitors may provide a means to control NSCLC in the latter population of patients, for whom effective treatments are rarely available.

METHODS SUMMARY

A recombinant retroviral cDNA expression library was constructed as described previously[7–10] from a lung cancer specimen, and was used to infect mouse 3T3 fibroblasts. Transformed foci isolated from the cells after 2 weeks of culture were subjected to extraction of genomic DNA and amplification of retroviral insert cDNAs by PCR. The *EML4–ALK* fusion cDNA was detected by RT–PCR analysis of total RNA from clinical specimens with the Fusion-RT-S (5′-GTGCAGTGTTTAGCATTCTTGGGG-3′) and Fusion-RT-AS (5′-TCTTGCCAGCAAAGCAGTAGTTGG-3′) primers. The *EML4–ALK* variant 1 gene was detected by PCR with genomic DNA of clinical specimens and the Fusion-genome-S (5′-CCACACCTGGGAAAGGACCTAAAG-3′) and Fusion-genome-AS (5′-AGCTTGCTCAGCTTGTACTCAGGG-3′) primers. Expression plasmids for Flag-epitope-tagged ALK, EML4 and EML4–ALK were generated with the retroviral vector pMXS[26]. The kinase-inactive mutant (K589M) of EML4–ALK was constructed by site-directed mutagenesis. Internal deletion mutants of EML4–ALK were also generated by mutagenesis. The expression plasmids were introduced into 3T3 cells by the calcium phosphate method, and the cells were then either cultured for 21 days or injected subcutaneously into nude mice. For analysis of EML4–ALK dimerization, expression vectors for Flag- or Myc-epitope-tagged EML4–ALK or its mutants were introduced into HEK 293 cells, cell lysates were subjected to immunoprecipitation with anti-Flag or anti-Myc, and the resulting precipitates were subjected to immunoblot analysis with the same antibodies. Anti-Flag immunoprecipitates were also subjected to an *in vitro* kinase assay with the synthetic YFF peptide[18]. For BA/F3 experiments, cDNAs for ALK, EML4–ALK or EML4–ALK(K589M) were inserted into the plasmid pMX-iresCD8 (ref. 27) to confer simultaneous expression of the protein of interest and mouse CD8. BA/F3 cells were infected with recombinant retroviruses generated from each plasmid, and the resulting CD8$^+$ cells were purified and incubated with various concentrations of WHI-P154 (EMD Biosciences).

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 15 February; accepted 17 May 2007.
Published online 11 July 2007.

1. Jemal, A. *et al.* Cancer statistics, 2006. *CA Cancer J. Clin.* 56, 106–130 (2006).
2. Schiller, J. H. *et al.* Comparison of four chemotherapy regimens for advanced non-small-cell lung cancer. *N. Engl. J. Med.* 346, 92–98 (2002).
3. Lynch, T. J. *et al.* Activating mutations in the epidermal growth factor receptor underlying responsiveness of non-small-cell lung cancer to gefitinib. *N. Engl. J. Med.* 350, 2129–2139 (2004).
4. Paez, J. G. *et al.* EGFR mutations in lung cancer: correlation with clinical response to gefitinib therapy. *Science* 304, 1497–1500 (2004).
5. Druker, B. J. *et al.* Efficacy and safety of a specific inhibitor of the BCR-ABL tyrosine kinase in chronic myeloid leukemia. *N. Engl. J. Med.* 344, 1031–1037 (2001).
6. Pao, W. *et al.* EGF receptor gene mutations are common in lung cancers from "never smokers" and are associated with sensitivity of tumors to gefitinib and erlotinib. *Proc. Natl Acad. Sci. USA* 101, 13306–13311 (2004).
7. Yoshizuka, N. *et al.* An alternative transcript derived from the *Trio* locus encodes a guanosine nucleotide exchange factor with mouse cell-transforming potential. *J. Biol. Chem.* 279, 43998–44004 (2004).
8. Hatanaka, H. *et al.* Transforming activity of purinergic receptor P2Y, G-protein coupled, 2 revealed by retroviral expression screening. *Biochem. Biophys. Res. Commun.* 356, 723–726 (2007).
9. Choi, Y. L. *et al.* Identification of a constitutively active mutant of JAK3 by retroviral expression screening. *Leuk. Res.* 31, 203–209 (2007).
10. Fujiwara, S. *et al.* Transforming activity of the lymphotoxin-β receptor revealed by expression screening. *Biochem. Biophys. Res. Commun.* 338, 1256–1262 (2005).
11. Pollmann, M. *et al.* Human EML4, a novel member of the EMAP family, is essential for microtubule formation. *Exp. Cell Res.* 312, 3241–3251 (2006).
12. Eichenmuller, B., Everley, P., Palange, J., Lepley, D. & Suprenant, K. A. The human EMAP-like protein-70 (ELP70) is a microtubule destabilizer that localizes to the mitotic apparatus. *J. Biol. Chem.* 277, 1301–1309 (2002).
13. Smith, T. F., Gaitatzes, C., Saxena, K. & Neer, E. J. The WD repeat: a common architecture for diverse functions. *Trends Biochem. Sci.* 24, 181–185 (1999).
14. Morris, S. W. *et al.* Fusion of a kinase gene, ALK, to a nucleolar protein gene, NPM, in non-Hodgkin's lymphoma. *Science* 263, 1281–1284 (1994).
15. Shiota, M. *et al.* 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 (1994).
16. Pulford, K., Morris, S. W. & Turturro, F. Anaplastic lymphoma kinase proteins in growth control and cancer. *J. Cell. Physiol.* 199, 330–358 (2004).
17. Galkin, A. V. *et al.* Identification of NVP-TAE684, a potent, selective, and efficacious inhibitor of NPM-ALK. *Proc. Natl Acad. Sci. USA* 104, 270–275 (2007).
18. Donella-Deana, A. *et al.* Unique substrate specificity of anaplastic lymphoma kinase (ALK): development of phosphoacceptor peptides for the assay of ALK activity. *Biochemistry* 44, 8533–8542 (2005).
19. Hernandez, L. *et al.* Diversity of genomic breakpoints in *TFG-ALK* translocations in anaplastic large cell lymphomas: identification of a new TFG-ALK$_{XL}$ chimeric gene with transforming activity. *Am. J. Pathol.* 160, 1487–1494 (2002).
20. Palacious, R. & Steinmetz, M. IL-3 dependent mouse clones that express B-220 surface antigen, contain Ig genes in germ-line configuration, and generate B lymphocytes *in vivo*. *Cell* 41, 727–734 (1985).
21. Kaul, K. L. Molecular detection of *Mycobacterium tuberculosis*: impact on patient care. *Clin. Chem.* 47, 1553–1558 (2001).
22. Marzec, M. *et al.* Inhibition of ALK enzymatic activity in T-cell lymphoma cells induces apoptosis and suppresses proliferation and STAT3 phosphorylation independently of Jak3. *Lab. Invest.* 85, 1544–1554 (2005).
23. Li, R. *et al.* Design and synthesis of 5-aryl-pyridone-carboxamides as inhibitors of anaplastic lymphoma kinase. *J. Med. Chem.* 49, 1006–1015 (2006).
24. Watanabe, S., Itoh, T. & Arai, K. JAK2 is essential for activation of c-*fos* and c-*myc* promoters and cell proliferation through the human granulocyte-macrophage colony-stimulating factor receptor in BA/F3 cells. *J. Biol. Chem.* 271, 12681–12686 (1996).

25. Duyster, J., Bai, R. Y. & Morris, S. W. Translocations involving anaplastic lymphoma kinase (ALK). *Oncogene* 20, 5623–5637 (2001).
26. Onishi, M. *et al.* Applications of retrovirus-mediated expression cloning. *Exp. Hematol.* 24, 324–329 (1996).
27. Yamashita, Y. *et al.* Sak serine/threonine kinase acts as an effector of Tec tyrosine kinase. *J. Biol. Chem.* 276, 39012–39020 (2001).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank R. Moriuchi for suggestions.

Author Contributions M.S. and Y.L.C. contributed equally to this work. M.S., S.-i.F., H.W. and H.H. constructed the cDNA library and screened for transforming genes. Y.L.C. sequenced the *EML4–ALK* cDNA and conducted the experiments with BA/F3 cells. Y.Y. and S.T. searched for *EGFR* and *KRAS* mutations. M.E., S.I., K.K., M.B., S.O., S.T., Y.I. and H.A. performed RT–PCR for *EML4–ALK* transcripts in cancer specimens. T.N., Y. Sohara, Y. Sugiyama and H.M. designed the overall project, and H.M. wrote the manuscript. All authors discussed the results and commented on the manuscript.

Author Information The nucleotide sequences of *EML4–ALK* variant 1 and variant 2 cDNA have been deposited in DDBJ, EMBL and GenBank under the accession numbers AB274722 and AB275889, respectively. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to H.M. (hmano@jichi.ac.jp).

Analysis of chromosome copy number in leukemic cells by different microarray platforms

Leukemia (2007) **21**, 1333–1337. doi:10.1038/sj.leu.2404636; published online 15 March 2007

Changes in the copy number of chromosomes, or copy number alterations (CNAs), are frequently apparent in the genome of cancer cells and may result in the amplification of oncogenes or deletion of tumor suppressor genes.[1] Such CNAs range in size from an entire chromosome to several kilobase pairs, with many of the smaller changes being undetectable by conventional bacterial artificial chromosome-based comparative genomic hybridization (CGH), which has a resolution of several hundred kilobase pairs.[2] This limitation has recently been overcome by the adaptation of high-density oligonucleotide microarrays that were originally developed for typing of single nucleotide polymorphisms (SNPs) to the evaluation of CNAs.[3] Sophisticated software for such analysis, including dChip (http://biosun1.harvard.edu/complab/dchip) and CNAG (http://www.genome.umin.jp/CNAG.html), is now available online.

Although high-density SNP-typing arrays allow determination of CNAs at a resolution of <100 kbp, it is not a simple task to link such data to changes in the number of the corresponding genes. The HGU133 Plus 2.0 (HGU133P2) microarray manufactured by Affymetrix (USA) was designed to quantitate the abundance of >47 000 human transcripts and has been widely used for gene expression profiling. Most of the probe sequences on the array are targeted to exons corresponding to the 3′ untranslated region of each transcript. Although, as far as we are aware, there have been no reports on the use of this microarray for CNA assessment or CGH, we reasoned that HGU133P2 may be a useful platform for evaluation of CNAs because (1) cDNA sequences are one of the most updated and comprehensive information for human genome, and (2) copy number values could be calculated directly for each gene. Furthermore, it would be straightforward to compare the copy number and expression level of a given gene through separate hybridization of genomic DNA and mRNA to the array. For this purpose, we modified a protocol for labeling and hybridizing genomic DNA, provided by Affymetrix (Supplementary Information).

To examine the fidelity of CNA analysis with the HGU133P2 array, we first isolated genomic DNA from mononuclear cells (MNCs) of peripheral blood from a healthy male volunteer (karyotype: 46,XY) and a healthy female volunteer (karyotype: 46,XX). Genomic DNA was also isolated from a human leukemia cell line (KCL22) with trisomy 8 plus t(9;22) and from transformed human B cells with a karyotype of 48,XXXX or of 47,XY,+21 (both from Coriell Cell Repositories, http://ccr.coriell.org). The DNA was fragmented by treatment with DNase I, end-labeled with biotin-conjugated deoxynucleotides and subjected to hybridization with the HGU133P2 array. We first chose 26 354 probe sets that gave reliable signals — those that received the 'Present' call from the GeneChip operating software (GCOS, Affymetrix) and had a signal intensity of ⩾1.0 — in the control experiment with the normal male genome. This group contained 960 probe sets that mapped to the X chromosome. The signal intensities of these X chromosome-specific probe sets in the experiment with the normal female genome (46,XX) were then normalized by those of the corresponding probe sets in the data for 46,XY. The mean±s.d. value of these XX/XY ratios for the X chromosome-specific probe sets was 2.07±1.19 (Figure 1a) and thus matched the predicted value of 2. The intensities of these probe sets for B cells with the

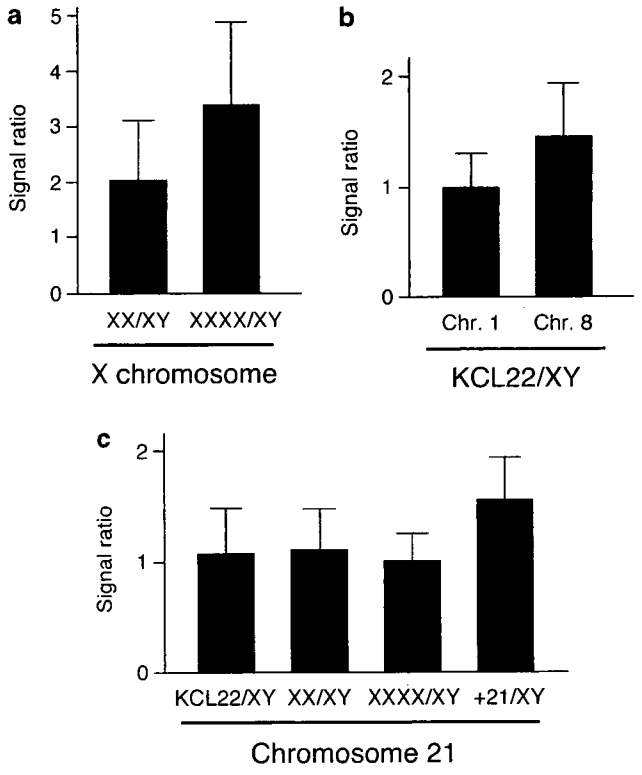

Figure 1 Comparison of chromosome-specific signal intensities obtained by HGU133P2 analysis of genomic DNA isolated from various samples. (**a**) Comparison of signal intensities for the X chromosome. Genomic DNA was purified from MNCs of peripheral blood from a male (XY) or female (XX) volunteer as well as from transformed B cells with a karyotype of 48,XXXX. The DNA was subjected to whole-genome amplification with the REPLI-g kit (Qiagen, Valencia, CA, USA), and the amplification products (100 μg) were digested, labeled with biotin and subjected to hybridization with an HGU133P2 microarray. The signal intensities of the probe sets specific for the X chromosome for the female volunteer or the B cells were normalized by those of the corresponding probe sets for the male volunteer, and the mean+s.d. values were determined. The detailed protocol for CGH analysis with HGU133P2 is described in Supplementary Information. (**b**) Comparison of signal intensities for chromosome 1 or 8. Genomic DNA of KCL22 cells was subjected to hybridization with HGU133P2. The signal intensities of the probe sets specific for chromosome (Chr) 1 or 8 were normalized by those of the corresponding probe sets for the male volunteer, and the mean+s.d. values were determined. (**c**) Comparison of signal intensities for chromosome 21. Genomic DNA of transformed B cells with a karyotype of 47,XY,+21 (+21) was subjected to hybridization with HGU133P2. The signal intensities of the chromosome 21-specific probe sets for these cells, as well as for KCL22, the female volunteer and the B cells with the karyotype 48,XXXX, were normalized by those of the corresponding probe sets for the male volunteer, and the mean+s.d. values were determined.

karyotype 48,XXXX were also normalized by those for 46,XY, giving a mean ± s.d. value for the XXXX/XY ratios of 3.34 ± 1.52.

To examine further the linearity of the signal intensity data, we compared the hybridization signals of the probe sets that mapped to chromosome 1 or 8 between the KCL22 cell line and the male control (Figure 1b). The mean ± s.d. of the signal intensity ratios (KCL22/XY) for the probe sets on chromosome 8 (1.47 ± 0.516) was well matched to the predicted value of 1.5, whereas that for the probe sets on chromosome 1 (with no abnormality in KCL22) was close to 1.0 (1.01 ± 0.360).

We next compared the signal intensities of the probe sets that mapped to chromosome 21 between the B cells with the karyotype 47,XY, +21, which contain three copies of chromosome 21, and the other samples, all of which contain two copies of this chromosome (Figure 1c). The mean ± s.d. of the signal intensities relative to the male control for the cells with trisomy 21 (1.52 ± 0.40) was significantly greater ($P<1.32\times10^{-29}$, Student's *t*-test) than that for the female control (1.13 ± 0.36), for the B cells with 48,XXXX (1.00 ± 0.28) or for KCL22 (1.11 ± 0.41); no significant difference was apparent among the values for the female control, for 48,XXXX and for KCL22.

We then screened for CNAs in leukemic blasts by using HGU133P2 platform and compared the results to the data obtained with a commercially available oligonucleotide-based CGH microarray (Human Genome 44A (CGH44A); Agilent, USA), which harbors >40 000 60-nucleotide oligomers that cover the human genome at a mean resolution of ~35 kbp. The CD34-positive progenitor (leukemic) fraction and the CD34-negative differentiated (control) fraction were isolated from the bone marrow of 23 patients with leukemia (Supplementary Table 1). We were also able to isolate MNCs from the bone marrow of four leukemia patients in complete remission (patient ID nos. 4, 6, 10 and 13), and we thus compared for these individuals the $CD34^+$ fraction obtained before chemotherapy with the $CD34^-$ fraction obtained during complete remission. Although CNA analysis with the HGU133P2 platform was conducted for all subjects, experiments with the CGH44A platform were carried out for 16 patients (ID nos. 4–49).

For the analysis with HGU133P2, genomic DNA isolated from the leukemic and control fractions of each patient was subjected to hybridization with the array separately, and the leukemic fraction/control fraction ratio of signal intensity was calculated for each probe set. For analysis with CGH44A, genomic DNA of the leukemic and control fractions of each patient was labeled with Cy5 and Cy3, respectively, and was mixed before hybridization with the array. To confirm the CGH44A data, we performed a dye-swap experiment for each patient. The influence of outlier probes in the CGH analysis was minimized by transformation of the intensity of each signal with a moving window of 1-Mbp width with the use of CGH Analytics 3.1.8 software (Agilent). Furthermore, gain or loss of chromosome number at given loci was considered significant only if the *z*-score[4] for the corresponding position was $\geqslant 1.0$. We considered CNAs identified by CGH reliable only if the original and dye-swap experiments revealed opposite changes each with a *z*-score of $\geqslant 1.0$.

The resulting CNA data were highly similar between the HGU133P2 and CGH44A experiments. Nine (~39%) out of the 23 leukemia patients showed a constant copy number of 2 for all chromosomes (data not shown). These results were also confirmed for patient nos. 4, 5 and 6 by hybridization of the corresponding genomic DNA to the Affymetrix 100K SNP-typing array (data not shown). The remaining patients manifested CNAs of various sizes. For instance, both the HGU133P2 and CGH44A analyses revealed a gain of the entire long arm of chromosome 1 and a loss of a short segment of 9q33.1 in patient no. 29 (Figure 2a). The CNAs identified in patient no. 29 by both arrays were not detected by clinical karyotyping (the blasts of the patient were designated 46,XX).

For patient no. 43, both HGU133P2 and CGH44A analyses revealed deletions of 5q, 7q, 9q and 16, a gain of 19q and the same pattern of complex copy number changes at 8p and chromosome 17 (Figure 2b). Again, the clinical karyotyping for this patient (46,XY) failed to detect these various changes.

To confirm the CNAs identified by both HGU133P2 and CGH44A, we measured chromosome copy number in the $CD34^+$ fractions of the leukemia patients with the use of quantitative, real-time polymerase chain reaction (PCR) analysis. The chromosome copy numbers determined by PCR were highly similar to those determined with the arrays for chromosomes 1 and 9q in patient no. 29 and for 8p in patient no. 43 (Figure 3a).

We also prepared single-stranded cDNA from the $CD34^+$ fractions of all subjects except patient no. 72 (RNA with a sufficient quality could not be obtained from this patient), and subjected it to hybridization with HGU133P2 in order to quantitate the amount of each transcript targeted by the array. Given that portions of chromosome 7q were frequently deleted in the leukemia patients, we compared the expression levels of the genes in the deleted regions between the individuals with or without 7q− (Figure 3b). The expression levels of the genes on 7q were significantly lower in the patients with 7q− (patient nos. 28, 43, 44, 60, 61, 70 and 73) than in those without it (18.1 ± 79.5 versus 23.3 ± 87.6 U, mean ± s.d.; $P=0.014$, *t*-test).

Additionally, we also examined whether an increase in chromosome copy number affects the expression level of the genes mapped to the chromosomes. As shown in Figure 3c, +8q was observed commonly in five patients (ID nos. 43, 59, 70, 71 and 72). Quantitation of mRNA for the genes mapped to the 8q region has revealed that the expression level of such genes was higher in the individuals with +8q than those without it (42.8 ± 147.3 versus 34.7 ± 114.9; $P=0.022$, *t*-test).

These data indicate that chromosome copy number affects the corresponding level of gene expression; however, given the relatively large s.d. in each data set, other factors (epigenetic changes or interactions with transcriptional factors, for instance) likely have a substantial effect on transcriptional activities. Consistent with this notion, the amount of genomic DNA was not significantly correlated with that of mRNA for the entire collection of probe sets on HGU133P2 (Pearson's correlation coefficient (r) $=6.48\times10^{-4}$, $P=0.881$).

The fluorescence signal intensity for a given gene (probe set) represented on the HGU133P2 array is calculated from those of 11 probes designed for each gene. Given that the physical distance between some of these probes is >10 kbp in the genome, small amplifications or deletions of the genome may be undetectable with HGU133P2. Nevertheless, our data have shown that HGU133P2 is potentially useful for quantitation of chromosome copy number in a gene-oriented manner. Both HGU133P2 and CGH44A array systems have a high resolution, revealing, for instance, a homozygous deletion of a small region of chromosome 9 containing *CDKN2A* and *CDKN2B* in the genome of patient no. 9 (data not shown). The availability of different platforms for CNA assessment will provide insight into different aspects of genomic structural anomalies associated with hematologic malignancies.

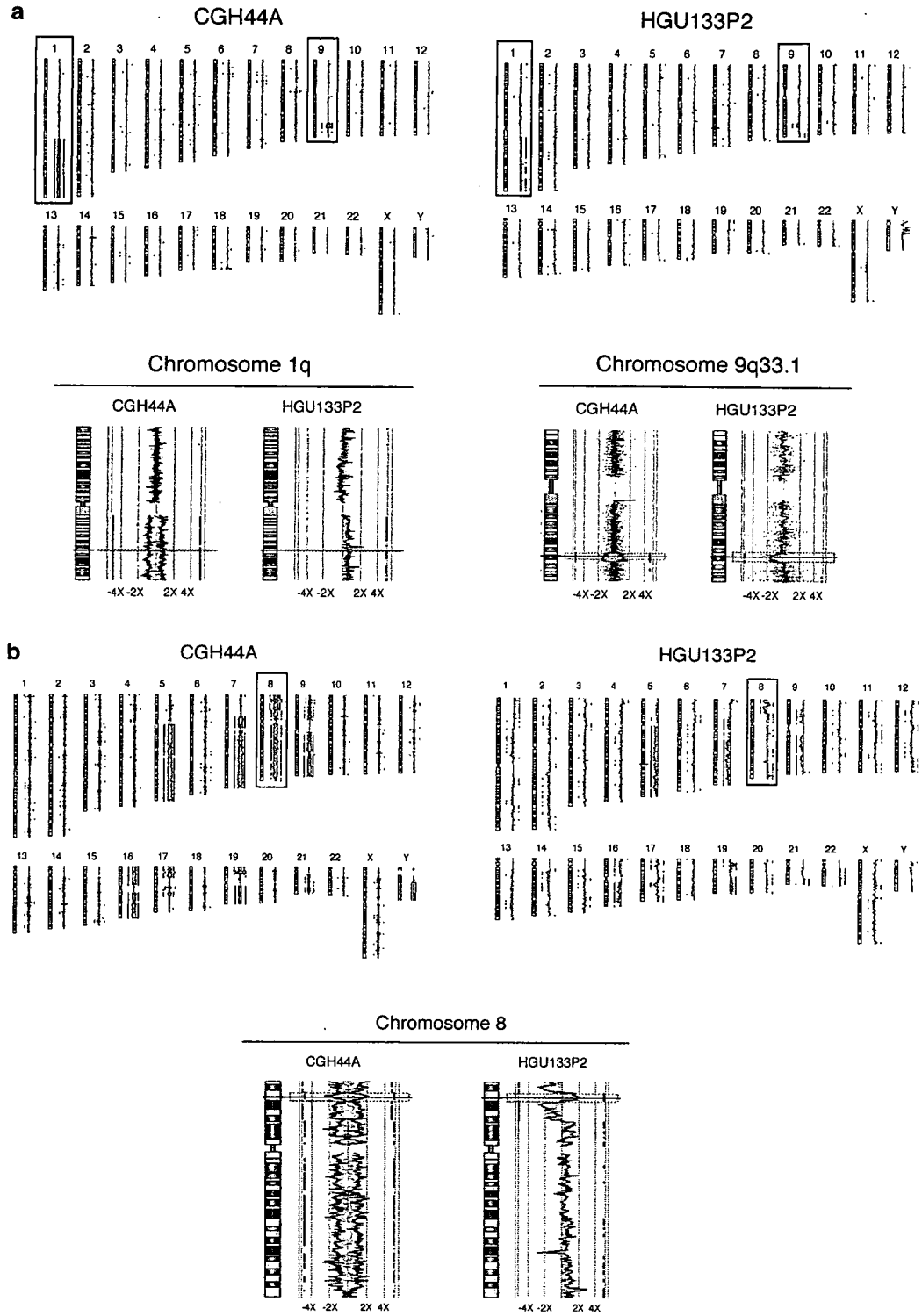

Figure 2 Evaluation of CNAs in leukemia patients with the use of HGU133P2 and CGH44A microarrays. (**a**) Assessment of patient no. 29. $CD34^+$ and $CD34^-$ fractions were isolated from bone marrow with the use of a magnetic bead-conjugated monoclonal antibody to CD34 (CD34 MicroBeads) and a MIDI-MACS cell separation column (both from Miltenyi Biotec, Auburn, CA, USA). For CGH44A analysis, genomic DNA was isolated from $CD34^+$ and $CD34^-$ fractions and labeled with Cy5 and Cy3, respectively; the labeled samples were then mixed and subjected to hybridization with the array.[5] A dye-swap experiment was also performed, and both the original (blue lines) and dye-swap (red lines) data are shown aligned with human chromosome figures by CGH Analytics 3.1.8 (upper panels). Regions with a *z*-score $\geqslant 1.0$ are indicated by side bars and shading (in corresponding colors). The data for chromosome 1 and 9 are shown magnified in the lower panels. For HGU133P2 analysis, genomic DNA isolated from $CD34^+$ and $CD34^-$ fractions was subjected to hybridization with the array separately, and the signal intensities for the former fraction were normalized by those for the latter. Regions with a *z*-score $\geqslant 1.0$ are shown by side bars and blue shading. All data obtained with the CGH44A and HGU133P2 arrays are available at the Gene Expression Omnibus (GEO) web site (http://www.ncbi.nlm.nih.gov/geo) under the accession numbers GSE4659 and GSE4602, respectively. (**b**) Assessment of patient no. 43. Signal intensity data for patient no. 43 were obtained as in (**a**), and the results for chromosome 8 are shown magnified in the lower panel.

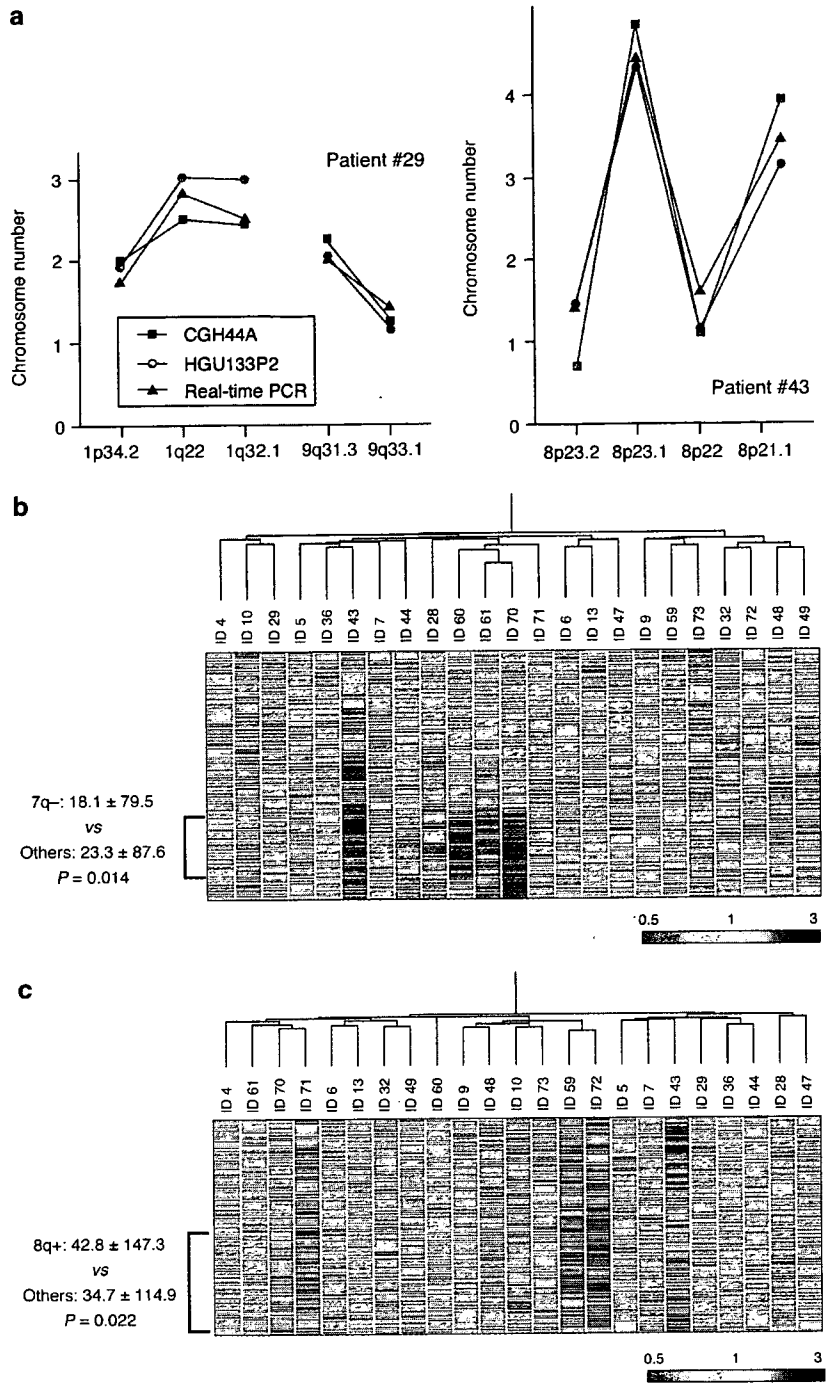

Figure 3 Confirmation of microarray data by quantitative PCR and the effect of DNA copy number on gene expression level. (**a**) Confirmation of microarray data by PCR. The amount of DNA at the indicated chromosomal positions for the $CD34^+$ fractions of patient nos 29 and 43 was determined by real-time PCR and expressed relative to that of DNA corresponding to the glyceraldehyde-3-phosphate dehydrogenase gene. The data are compared with chromosome number determined by CGH44A and HGU133P2 analyses. The sequences of the PCR primers are provided in Supplementary Information. (**b**) Effect of DNA copy number on gene expression level. The leukemic fraction/control fraction DNA ratio was calculated for each probe set on chromosome 7, and the study samples were subjected to hierarchical clustering analysis based on such raw DNA ratio profiles. Each column corresponds to a separate sample, and each row to a probe set whose leukemic fraction/control fraction DNA ratio is color-coded according to the indicated scale. Probe sets are ordered by their physical position in chromosome 7 from top to bottom. mRNA was isolated from the $CD34^+$ fractions of 22 leukemia patients, and the abundance of specific transcripts was quantitated by HGU133P2 array analysis as described previously.[6] The expression level of the genes on a region of the long arm of chromosome 7 was compared between individuals with or without the deletion of the region. The ID numbers of the patients with a loss in chromosome copy number in the region are shown in green. All array data for mRNA quantitation are available at the GEO web site under the accession number GSE4608. (**c**) The leukemic fraction/control fraction DNA ratio was calculated for each probe set on chromosome 8, and the subject tree was constructed as in (**b**). Probe sets are ordered by their physical position in chromosome 8 from top to bottom. The expression level of the genes on the long arm of chromosome 8 was compared between individuals with or without 8q+. The ID numbers of the patients with a gain in chromosome copy number in the region are shown in red.

Acknowledgements

We thank Dr Raji Pillai and her colleagues at Affymetrix for technical suggestions and for sharing data before publication. This work was supported in part by grants for Research on Human Genome and Tissue Engineering and for Third-Term Comprehensive Control Research for Cancer from the Ministry of Health, Labor, and Welfare of Japan, as well as by a Grant-in-Aid for Scientific Research on Priority Areas 'Applied Genomics' from the Ministry of Education, Culture, Sports, Science, and Technology of Japan.

Y Yamashita[1], K Minoura[2], T Taya[2], S-i Fujiwara[1], K Kurashina[1], H Watanabe[1], YL Choi[1], M Soda[1], H Hatanaka[1], M Enomoto[1], S Takada[1] and H Mano[1,3]

[1]*Division of Functional Genomics, Center for Molecular Medicine, Jichi Medical University, Tochigi, Japan;*
[2]*Bioapplication Group, Yokogawa Analytical Systems Inc., Tokyo, Japan and*
[3]*CREST, Japan Science and Technology Agency, Saitama, Japan*
E-mail: hmano@jichi.ac.jp

References

1 Hanahan D, Weinberg RA. The hallmarks of cancer. *Cell* 2000; **100**: 57–70.
2 Lockwood WW, Chari R, Chi B, Lam WL. Recent advances in array comparative genomic hybridization technologies and their applications in human genetics. *Eur J Hum Genet* 2006; **14**: 139–148.
3 Lieberfarb ME, Lin M, Lechpammer M, Li C, Tanenbaum DM, Febbo PG *et al.* Genome-wide loss of heterozygosity analysis from laser capture microdissected prostate cancer using single nucleotide polymorphic allele (SNP) arrays and a novel bioinformatics platform dChipSNP. *Cancer Res* 2003; **63**: 4781–4785.
4 Kincaid R, Ben-Dor A, Yakhini Z. Exploratory visualization of array-based comparative genomic hybridization. *Inf Vis* 2005; **4**: 176–190.
5 Barrett MT, Scheffer A, Ben-Dor A, Sampas N, Lipson D, Kincaid R *et al.* Comparative genomic hybridization using oligonucleotide microarrays and total genomic DNA. *Proc Natl Acad Sci USA* 2004; **101**: 17765–17770.
6 Tsutsumi C, Ueda M, Miyazaki Y, Yamashita Y, Choi YL, Ota J *et al.* DNA microarray analysis of dysplastic morphology associated with acute myeloid leukemia. *Exp Hematol* 2004; **32**: 828–835.

Supplementary Information accompanies the paper on the Leukemia website (http://www.nature.com/leu)

Expression of the *JAK2* V617F mutation is not found in *de novo* AML and MDS but is detected in MDS-derived leukemia of megakaryoblastic nature

Leukemia (2007) **21**, 1337–1338. doi:10.1038/sj.leu.2404626; published online 8 March 2007

A somatic point mutation in the tyrosine kinase *JAK2* (V617F) is detected in most patients with polycythenia vera (PV) and in approximately half of those with essential thrombocythemia (ET), as well as in other myeloid disorders such as *de novo* acute myeloid leukemia (AML; ~6%), myelodysplastic syndrome (MDS; ~4%), chronic myelomonocytic leukemia (CMML; ~20%) and atypical myeloproliferative disease (MPD). In contrast, *JAK2* V617F was not detected in the patients with acute or chronic lymphoid malignancies.[1] In these studies, most of these mutations were heterozygous mutation and homozygous mutation was only detected in 30% of patients with *JAK2* mutation using DNA samples.[2] Thus, the prevalence of the expression of *JAK2* V617F in myeloid malignancies remains unclear.

In the present study, we examined the expression of *JAK2* V617F in 241 Japanese patients with myeloid malignancies. The morphologic subtypes of AML and MDS were initially classified according to the French–American–British (FAB) criteria. A diagnosis of transformation from MDS to AML (MDS-AML) was made when the bone marrow (BM) consisted of more than 30% blasts. Leukemic cells from BM were collected after informed consent was obtained. Mononuclear cells were isolated using the Ficoll–Hypaque density gradient centrifugation method, followed by total cellular RNA extraction. Expression of the *JAK2* gene was determined by reverse transcription-PCR, as described previously.[3] Mutation analysis of V617F was then performed by direct sequencing of these PCR products.

A homozygous *JAK2* V617F mutation was detected in the HEL cell line, which was established from a patient with erythroleukemia (AML M6 as defined by the FAB classification) and in which the JAK2/STAT5 signal transduction pathway is constitutively activated.[4] Among *de novo* AML, only heterozygous *JAK2* V617F mutation was detected in the patients with AML M6 (1 of 53 patients; 2%) and AML M7 (2 of 11 cases; 18%).[5,6] None of the 198 patients with *de novo* AML tested showed expression of *JAK2* V617F, including nine cases with M6 and eight cases with M7 (Table 1). These observations suggest that the prevalence of the expression of *JAK2* V617F in *de novo* AML appears to be less frequent than the reported detection rate of the genetic mutation using DNA samples.

In contrast to *de novo* AML, expression of the *JAK2* V617F mutation was found in 2 of 18 cases (11%) of MDS-AML but not

Table 1 Prevalence of the expression of *JAK2* V617F in myeloid malignancies

Disease		*Tested samples*	*JAK2 V617F*	*% Mutation*
De novo AML	M0	5	0	0
	M1	21	0	0
	M2	90	0	0
	M3	12	0	0
	M4	32	0	0
	M5	21	0	0
	M6	9	0	0
	M7	8	0	0
MDS	RA	8	0	0
	RAEB	17	0	0
MDS-AML		18	2	11

Abbreviations: AML, acute myeloid leukemia; MDS, myelodysplastic syndrome.

Table 2 Clinical and laboratory features of MDS-AML showing expression of *JAK2* V617F

Patient	*Age/sex*	*WBC* ($\times 10^9$/l)	*Hb* (g/dl)	*Platelets* ($\times 10^9$/l)	*% blasts in BM*	*Myelofibrosis*	*Surface marker*[a]	*Karyotype*	*Splenomegaly*
1396	51/M	1.6	5.6	8.0	85	—	CD34,CD33,CD13,CD41,CD7	46,XY	—
2704	52/M	12.6	7.9	115.0	75	—	CD34,CD41	46,XY	—

Abbreviations: BM, bone marrow; Hg, hemoglobin; M, male; WBC, white blood cells; —, not found.
[a]Positive antigens on the blasts.

in eight cases with refractory anemia (RA) or 17 cases with refractory anemia with excess blasts (RAEB) (Table 1). In both cases with the mutation, most of fluorescent intensity at codon 617 was V617F and the wild type was only faintly visible, suggesting that most of the leukemic cells expressed V617F. A previous study found a high occurrence of the *JAK2* V617F mutation in RA with ringed sideroblasts associated with marked thrombocytosis (71%),[7] and this mutation has also been detected in MDS with myelofibrosis (33%).[8] In our two cases with the mutation, neither thrombocytosis nor myelofibrosis was detected (Table 2). Interestingly, the blastic cells of these two cases were small with pale agranular cytoplasms and cytoplasmic blebs, and expressed CD41 antigen, suggesting that blastic cells have a megakaryoblastic nature. The progenitor cells of MDS are believed to undergo a multistep process during transformation into overt acute myelogenous leukemia. Several genetic abnormalities have been detected in both MDS and MSD-AML patients, which have suggested that genetic alterations are closely associated with disease progression of MDS.[3] Therefore, MDS serves as a useful model for studying the abnormal genetic events that occur in leukemogenesis. We detected expression of *JAK2* V617F in MDS-AML but not in RA or RAEB, suggesting that expression of *JAK2* V617F may be one of the genetic factors involved in the progression of MDS to AML with a megakaryoblastic nature.

In conclusion, detection of the expression of the *JAK2* gene mutation in MDS-AML patients may enhance both the management of these patients and the application of adequate therapeutic strategies such as tyrosine kinase inhibitors.

Acknowledgements

We thank Dr K Kita (Tokura Hospital, Kyoto, Japan) for providing the clinical data. This work was supported by research grants from the Mie Medical Research Fund (KN, 2004) Japan Leukemia Research Fund (KN, 2004), and a Grant-in-Aid for Scientific Research (C, KAKENHI: KN, 17590992) from the Japan Society for the Promotion of Science (JSPS).

K Nishii, R Nanbu, F Lorenzo V, F Monma, K Kato, H Ryuu and N Katayama
Division of Hematology and Oncology, Mie University School of Medicine, Tsu, Mie, Japan
E-mail: kaz@clin.medic.mie-u.ac.jp

References

1 Tefferi A, Pardanani A. Mutation screening for JAK2V617F: when to order the test and how to interpret the results. *Leuk Res* 2006; **30**: 739–744.
2 James C, Ugo V, Le Couedic JP, Staerk J, Delhommeau F, Lacout C *et al.* A unique clonal JAK2 mutation leading to constitutive signaling causes polycythaemia vera. *Nature* 2005; **434**: 1144–1148.
3 Lorenzo F, Nishii K, Monma F, Kuwagata S, Usui E, Shiku H. Mutational analysis of the KIT gene in myelodysplastic syndrome (MDS) and MDS-derived leukemia. *Leuk Res* 2006; **30**: 1235–1239.
4 Papayannopoulou T, Yokochi T, Nakamoto B, Martin P. The surface antigen profile of HEL cells. *Prog Clin Biol Res* 1983; **134**: 277–292.
5 Jelinek J, Oki Y, Gharibyan V, Bueso-Ramos C, Prchal JT, Verstovsek S *et al.* JAK2 mutation 1849G>T is rare in acute leukemias but can be found in CMML, Philadelphia chromosome-negative CML, and megakaryocytic leukemia. *Blood* 2005; **106**: 3370–3373.
6 Frohling S, Lipka DB, Kayser S, Scholl C, Schlenk RF, Dohner H *et al.* Rare occurrence of the JAK2 V617F mutation in AML subtypes M5, M6, and M7. *Blood* 2006; **107**: 1242–1243.
7 Renneville A, Quesnel B, Charpentier A, Terriou L, Crinquette A, Lai JL *et al.* High occurrence of JAK2 V617 mutation in refractory anemia with ringed sideroblasts associated with marked thrombocytosis. *Leukemia* 2006; **20**: 2067–2070.
8 Ohyashiki K, Aota Y, Akahane D, Gotoh A, Miyazawa K, Kimura Y *et al.* The JAK2 V617F tyrosine kinase mutation in myelodysplastic syndromes (MDS) developing myelofibrosis indicates the myeloproliferative nature in a subset of MDS patients. *Leukemia* 2005; **19**: 2359–2360.

Quantification of *ex vivo* generated dendritic cells (DC) and leukemia-derived DC contributes to estimate the quality of DC, to detect optimal DC-generating methods or to optimize DC-mediated T-cell-activation-procedures *ex vivo* or *in vivo*

Leukemia (2007) **21**, 1338–1341. doi:10.1038/sj.leu.2404639; published online 22 March 2007

There is a need for less intensive (post-remission) immunotherapies to maintain stable remissions in AML and at least stable diseases in MDS before or after stem cell transplantation (SCT). The significance of T cells to mediate cellular anti leukemic reactions has been demonstrated by donor-lymphocyte-infusion (DLI) therapy of relapsed AML,[1] although not all of the patients treated respond to this therapy. Professional antigen-presenting cells like dendritic cells (DC) could be used to overcome this therapeutic resistance and (re)-activate anti-leukemia-directed allogeneic or autologous T cells.[2] Leukemic

Available online at www.sciencedirect.com
SCIENCE DIRECT®
Leukemia Research 29 (2005) 943–949
Leukemia Research
www.elsevier.com/locate/leukres

Retroviral expression screening of oncogenes in natural killer cell leukemia

Young Lim Choi [a, b], Ryozo Moriuchi [c], Mitsujiro Osawa [d], Atsushi Iwama [d], Hideki Makishima [e], Tomoaki Wada [a], Hiroyuki Kisanuki [a], Ruri Kaneda [a], Jun Ota [a, f], Koji Koinuma [a], Madoka Ishikawa [a], Shuji Takada [a], Yoshihiro Yamashita [a], Kazuo Oshimi [b], Hiroyuki Mano [a, f, *]

[a] *Division of Functional Genomics, Jichi Medical School, 3311-1 Yakushiji, Kawachigun, Tochigi 329-0498, Japan*
[b] *Division of Hematology, Department of Medicine, Juntendo University School of Medicine, Tokyo, Japan*
[c] *Department of Molecular Microbiology and Immunology, Nagasaki University Graduate School of Medicine, Nagasaki, Japan*
[d] *Center for Experimental Medicine, Institute of Medical Science, University of Tokyo, Tokyo, Japan*
[e] *Second Department of Internal Medicine, Shinshu University School of Medicine, Nagano, Japan*
[f] *CREST, Japan Science and Technology Agency, Saitama, Japan*

Received 17 October 2004; accepted 22 January 2005
Available online 24 February 2005

Abstract

Aggressive natural killer cell leukemia (ANKL) is an intractable malignancy that is characterized by the outgrowth of NK cells. To identify transforming genes in ANKL, we constructed a retroviral cDNA expression library from an ANKL cell line KHYG-1. Infection of 3T3 cells with recombinant retroviruses yielded 33 transformed foci. Nucleotide sequencing of the DNA inserts recovered from these foci revealed that 31 of them encoded KRAS2 with a glycine-to-alanine mutation at codon 12. Mutation-specific PCR analysis indicated that the *KRAS* mutation was present only in KHYG-1 cells, not in another ANKL cell line or in clinical specimens ($n = 8$).

Keywords: Aggressive NK cell leukemia; cDNA expression library; Retrovirus; KRAS2 oncogene

1. Introduction

Outgrowth of $CD3^{-}CD16/CD56^{+}$ natural killer (NK) cells in peripheral blood is diagnosed as either chronic NK lymphocytosis (CNKL) or aggressive NK cell leukemia (ANKL) [1,2]. Whereas the former condition has an indolent clinical course with few symptoms, the latter is characterized by chemoresistance and multiorgan failure and has a poor outcome.

The Epstein–Barr virus (EBV) genome is frequently present episomally in ANKL cells [3], suggesting a role for EBV in disease pathogenesis. However, little is known of how infection with EBV might trigger clonal growth of NK cells. Inactivation of tumor suppressor genes has been associated with NK cell neoplasia. For instance, a homozygous deletion of the genes for p16INK4A, p15INK4B, or p14ARF has been detected [4]. Additionally, inactivating mutations of the FAS gene have been found in nasal NK/T cell lymphoma [5].

A few studies have identified a potential contribution of oncogenes to NK cell malignancy. Mutations that affect codons 13 or 22 of *KRAS2* were found in NK/T cell lymphoma [6] but not in ANKL [7]. Furthermore, although mutations in *KIT* were associated with NK/T cell lymphoma, transforming activity of the mutant KIT proteins was not demonstrated [8]. A role for oncogenes in ANKL has not been identified to date.

Functional screening based on transforming ability is one potential approach to the efficient isolation of tumor-promoting genes in ANKL. Focus formation assays with

* Corresponding author. Tel.: +81 285 58 7449; fax: +81 285 44 7322.
E-mail address: hmano@jichi.ac.jp (H. Mano).

0145-2126/$ – see front matter © 2005 Elsevier Ltd. All rights reserved.
doi:10.1016/j.leukres.2005.02.002

mouse 3T3 fibroblasts have indeed proved successful for the identification of oncogenes in human cancer [9]. In such screening assays, genomic DNA isolated from cancer specimens is used to transfect 3T3 cells and the formation of transformed cell foci is then evaluated. Expression of the exogenous genes in such experiments is driven by their own promoters or enhancers, however, so that oncogenes can exert transforming effects in 3T3 cells only if their regulatory regions are active in fibroblasts. Given the distinct developmental origins of NK cells and fibroblasts, expression of oncogenes associated with ANKL in 3T3 cells under these conditions is not guaranteed.

This problem might be expected to be overcome by the expression of test cDNAs under the control of an ectopic promoter in 3T3 cells. We have therefore constructed a retroviral cDNA expression library from the ANKL cell line KHYG-1 [10] and used this library to infect 3T3 cells. In preparation of the cDNA library, we took advantage of a polymerase chain reaction (PCR)-based system that preferentially amplifies full-length cDNAs. The resulting library was found to have sufficient complexity and to contain a high percentage of full-length cDNAs. Focus formation assays with 3T3 cells resulted in the identification of *KRAS2* as a transforming gene in KHYG-1 cells.

2. Materials and methods

2.1. Cell culture and clinical samples

KHYG-1 and NKL cells [11] were kindly provided by M. Yagita and Y. Yokota, respectively, and were cultured in RPMI 1640 medium (Invitrogen, Carlsbad, CA) supplemented with 10% fetal bovine serum (Invitrogen) and human interleukin-2 (20 U/mL) (Roche, St. Louis, MO). The BOSC23 packaging cell line for ecotropic retroviruses [12] and mouse 3T3 fibroblasts were maintained in Dulbecco's modified Eagle's medium (DMEM)–F12 (Invitrogen) supplemented with 10% fetal bovine serum and 2 mM L-glutamine.

Mononulear cells were isolated by Ficoll–Hypaque density gradient centrifugation from peripheral blood of the subjects with informed consent. The cells were incubated with anti-CD3 MicroBeads (Miltenyi Biotec, Auburn, CA), and loaded onto MIDI-MACS magnetic cell separation columns (Miltenyi Biotec) to remove $CD3^+$ cells. The flow-through was then mixed with anti-CD56 MicroBeads (Miltenyi Biotec), and was subjected to a MINI-MACS column for the "positive selection" of $CD56^+$ cells. Cells bound specifically to the column was then eluted according to the manufacturer's instructions.

2.2. Construction of a retrovirus library

Total RNA was extracted from KHYG-1 cells with the use of an RNeasy Mini column and RNase-free DNase (Qiagen, Valencia, CA), and first-strand cDNA was synthesized from the RNA with PowerScript reverse transcriptase, a SMART IIA oligonucleotide, and CDS primer IIA (Clontech, Palo Alto, CA). The resulting cDNA molecules were then amplified for 12 cycles with 5′-PCR primer IIA and a SMART PCR cDNA synthesis kit (Clontech), with the exception that LA Taq polymerase (Takara Bio, Shiga, Japan) was substituted for the Advantage 2 DNA polymerase provided with the kit. The PCR products were treated with proteinase K, rendered blunt-ended with T4 DNA polymerase, and ligated to a *Bst*XI adapter (Invitrogen). Unbound adapters were removed with a cDNA size fractionation column (Invitrogen), and the modified cDNAs were ligated into the pMX retroviral plasmid (kindly provided by T. Kitamura) [13] that had been digested with *Bst*XI. The pMX-cDNA plasmids were introduced into ElecroMax DH10B cells (Invitrogen) by electroporation.

2.3. Focus formation assay

BOSC23 cells (1.8×10^6) were seeded onto 6-cm culture plates, cultured for 1 day, and then transfected with a mixture comprising 2 μg of retroviral plasmids, 0.5 μg of pGP plasmid (Takara Bio), 0.5 μg of pE-eco plasmid (Takara Bio), and 18 μL of Lipofectamine reagent (Invitrogen). Two days after transfection, polybrene (Sigma, St. Louis, MO) was added at a concentration of 4 μg/mL to the culture supernatant, which was then used to infect 3T3 cells for 48 h. For the focus formation assay, the culture medium of 3T3 cells was changed to DMEM–high glucose (Invitrogen) supplemented with 5% calf serum and 2 mM L-glutamine. Transformed foci were isolated after 3 weeks of culture.

2.4. Recovery of cDNAs from 3T3 cells

Each 3T3 cell clone was harvested with a cloning syringe and cultured independently in a 10-cm culture plate. Genomic DNA was subsequently extracted from the cells and subjected to PCR with 5′-PCR primer IIA and LA Taq polymerase for 50 cycles of 98 °C for 20 s and 68 °C for 6 min. Amplified genomic fragments were purified by gel electrophoresis and ligated into the pT7Blue-2 vector (EMD Biosciences, San Diego, CA) for nucleotide sequencing.

2.5. Mutation-specific PCR for KRAS2

Detection of $KRAS2^{G12A}$ cDNA was performed as described previously [14] but with minor modifications. In brief, a 5′-region of *KRAS2* cDNA was amplified from oligo(dT)–primed cDNA by PCR with 5′-RAS primer (5′-ACTGAATATAAACTTGTGGTAGTTGGACCT-3′; the underlined cytosine was incorporated to generate a *Bst*NI site) and 3′-RAS primer A (5′-CTGTGTCGAGAATATCCAAGAGACA-3′). The PCR product was subjected to digestion with *Bst*NI (New England Biolabs, Beverly, MA) and then to a second PCR with 5′-RAS primer and 3′-RAS primer B (5′-CTGTGTCGAGAATCCAGGAGACA-3′; the under-

lined guanine was incorporated to generate a *Bst*NI site). The second PCR product was then also subjected to digestion with *Bst*NI, and the resulting DNA fragments were separated by agarose gel electrophoresis.

3. Results

3.1. Construction of a full-length cDNA expression library for KHYG-1 cells

Full-length cDNAs were selectively amplified from mRNA of KHYG-1 cells and ligated into the retroviral vector pMX. We obtained a total of 5.61×10^6 colony-forming units (cfu) of independent plasmid clones. To evaluate the quality of the library, we randomly selected 40 clones and examined the incorporated cDNAs. Thirty-nine of the 40 clones contained inserts with an average size of 2.03 kbp. The cDNA inserts from 20 out of these 39 clones were sequenced from both ends, and the determined sequences were used to screen, with the BLAT search program [15], the nucleotide sequence database assembled as of July 2003 by the Genome Bioinformatics Group of the University of California at Santa Cruz (http://genome.ucsc.edu/). Both ends of 14 of the 20 cDNAs could be matched to the mRNA sequences of known genes, and 13 of these cDNAs included complete open reading frames (data not shown). We therefore concluded that the retroviral cDNA expression library was of sufficient complexity and sufficiently enriched in full-length cDNAs for the present study.

3.2. Identification of $KRAS2^{G12A}$ in KHYG-1 cells

We generated a recombinant ecotropic retrovirus library by introducing 7.1×10^5 cfu of the generated plasmids into a packaging cell line. This library was then used to infect mouse 3T3 fibroblasts. After culture of the infected cells for 3 weeks, we detected 33 transformed foci (Fig. 1). Each focus was isolated, expanded independently, and subjected to extraction of genomic DNA for the recovery of retroviral inserts by PCR with the primer used originally to amplify the cDNAs during construction of the library. In most instances, a single major DNA fragment was recovered from each genome (Fig. 2A), suggestive of original infection of a single 3T3 cell with a single retrovirus.

The recovered cDNA fragments were sequenced from both ends for all 33 clones. Screening of the human genome sequence database with the insert sequences revealed that those from 31 of the 33 clones (#1–#29, #31, #33) matched, with >98% identity, the sequence of human *KRAS2* (GenBank accession number, NM_004985). The genome of 3T3 clone ID #30 yielded two PCR fragments (Fig. 2A); the larger (~1.4 kbp) and the smaller (~0.9 kbp) fragments were revealed to be derived from β-actin (*ACTB*; GenBank accession number, NM_001101) and profilin 1 (*PFN1*; GenBank accession number, NM_005022) genes, respectively. The final 3T3 clone (#32) yielded a major PCR fragment corresponding to the gene for isocitrate dehydrogenase 3 (NAD^+) β (*IDH3B*; GenBank accession number, NM_006899).

KRAS2 belongs to the *RAS* gene family and is involved in a wide variety of human cancers [16]. Given that point

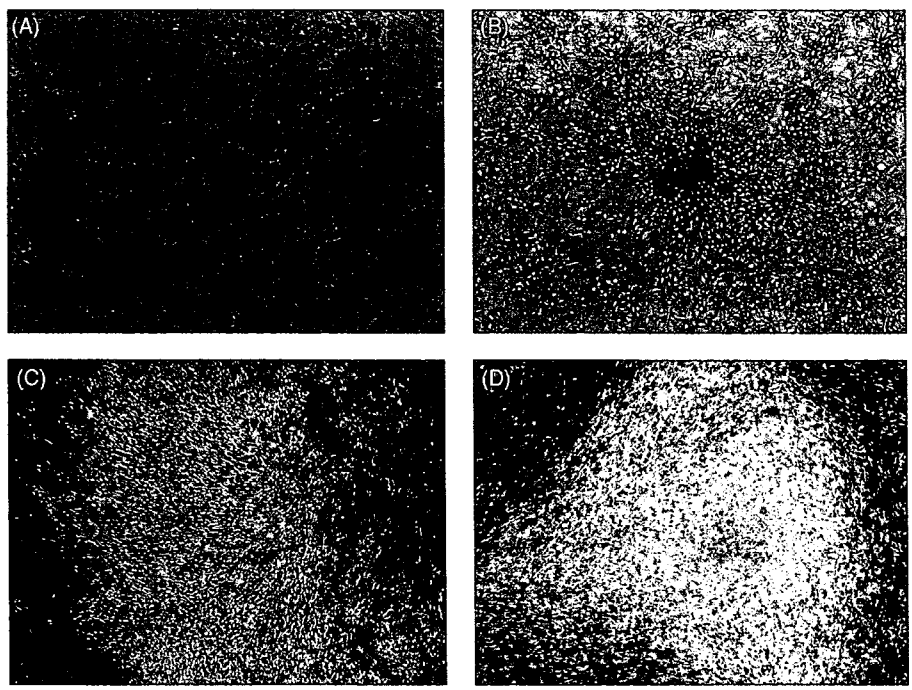

Fig. 1. Focus formation assay with a retroviral library derived from KHYG-1 cells. Mouse 3T3 cells were infected with the empty virus (A), a retrovirus expressing v-Ras as a positive control (B), or retroviruses from the KHYG-1 cell library (C and D). The cultures were photographed 3 weeks after infection. Scale bar, 100 μm.

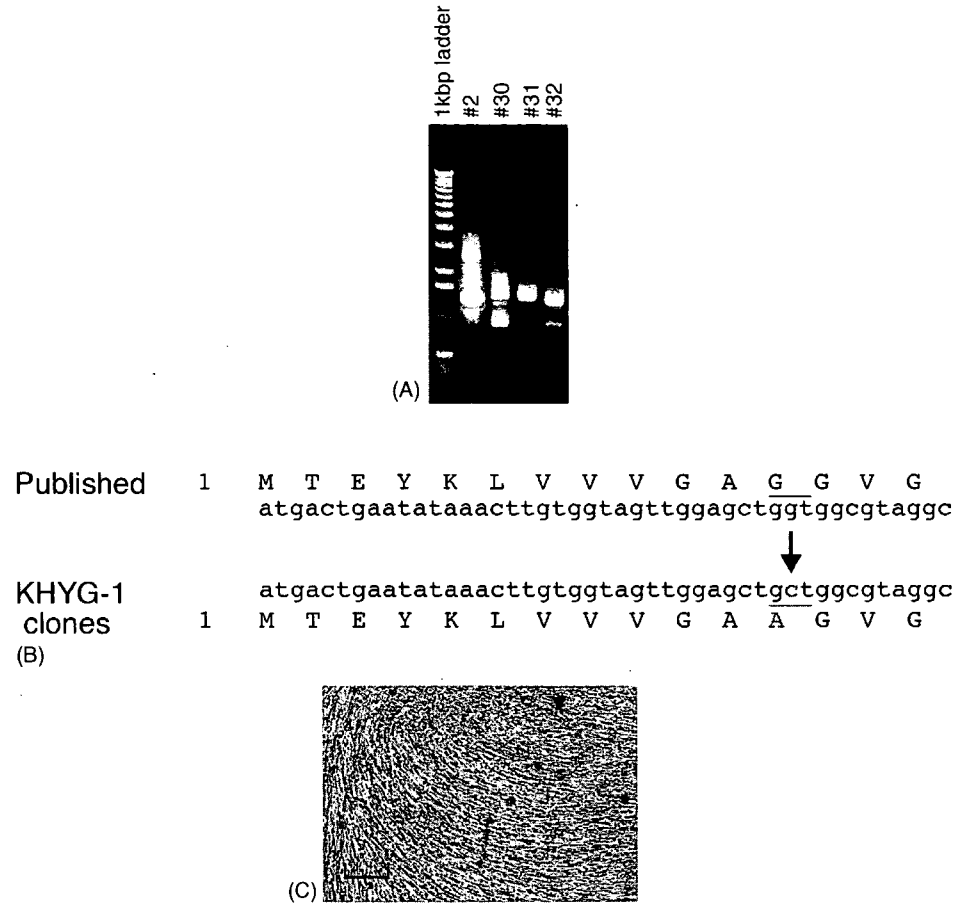

Fig. 2. Identification of a *KRAS* mutant gene with transforming activity: (A) Genomic DNA isolated from transformed 3T3 cell foci (clones #2, #30, #31, and #32) was subjected to PCR for amplification of the DNA inserts. The left lane contains DNA size markers (1-kbp DNA ladder; Invitrogen); (B) The nucleotide sequences of the DNA inserts derived from 31 of the 33 transformed foci matched that of *KRAS* with a single nucleotide substitution (G to C) in the second position of codon 12 that results in a change in the encoded amino acid from glycine to alanine; (C) A recombinant retrovirus encoding KRAS2^{G12A} was used to infect 3T3 cells. The cells were photographed after culture for 2 weeks. Scale bar, 50 μm.

mutations in *KRAS2* confer oncogenic activity, we compared the nucleotide sequences of the *KRAS2* cDNAs derived from KHYG-1 cells with the published sequence of the wild-type gene. Although oncogenic mutations have previously been shown to affect codons 12, 13, and 59 of *KRAS2*, all of the KHYG-1 cell cDNAs harbored the same mutation: the GGT sequence of codon 12 was changed to GCT, resulting in the substitution of an alanine residue for the glycine normally present at this position (Fig. 2B). To verify the transforming ability of *KRAS2*^{G12A}, we inserted the corresponding cDNA into the pMX retroviral vector and generated recombinant retroviruses for infection of 3T3 cells. The recipient 3T3 cells indeed underwent transformation (Fig. 2C), confirming that *KRAS2*^{G12A} possesses oncogenic activity.

3.3. Screening for *KRAS2*^{G12A} in NK cell leukemia

To determine whether *KRAS2*^{G12A} is frequently associated with NK cell leukemia, we applied a slightly modified version of a rapid screening method previously described by Kahn et al. [14]. *KRAS2* cDNA was first amplified by PCR with 5′-RAS primer and 3′-RAS primer A (Fig. 3A). The 3′ end of 5′-RAS primer corresponds to codon 11 of *KRAS2* but contains a cytosine substitution in the first position of codon 11, which generates a *Bst*NI recognition site [CC(T/A)GG] that includes the first and second positions of codon 12. The presence of a point mutation at the first or second position of codon 12 would therefore prevent digestion of the PCR product by *Bst*NI.

After *Bst*NI digestion, the PCR product was subjected to a second PCR with 5′-RAS primer and 3′-RAS primer B. Given that *Bst*NI digestion removes the binding site for 5′-RAS primer, only *KRAS2* cDNA with a mutation at the first or second position of codon 12 should yield a second PCR product. Even if *Bst*NI digestion of the initial PCR product was not complete and the second PCR amplified a trace amount of wild-type *KRAS2* cDNA, a second *Bst*NI digestion further discriminates between the wild-type and mutant genes. The 3′-RAS primer B thus contains a guanine substitution that generates a *Bst*NI site. The second PCR product of wild-type *KRAS2* cDNA would thus contain two *Bst*NI sites, whereas that of mutant *KRAS2* contains only one.

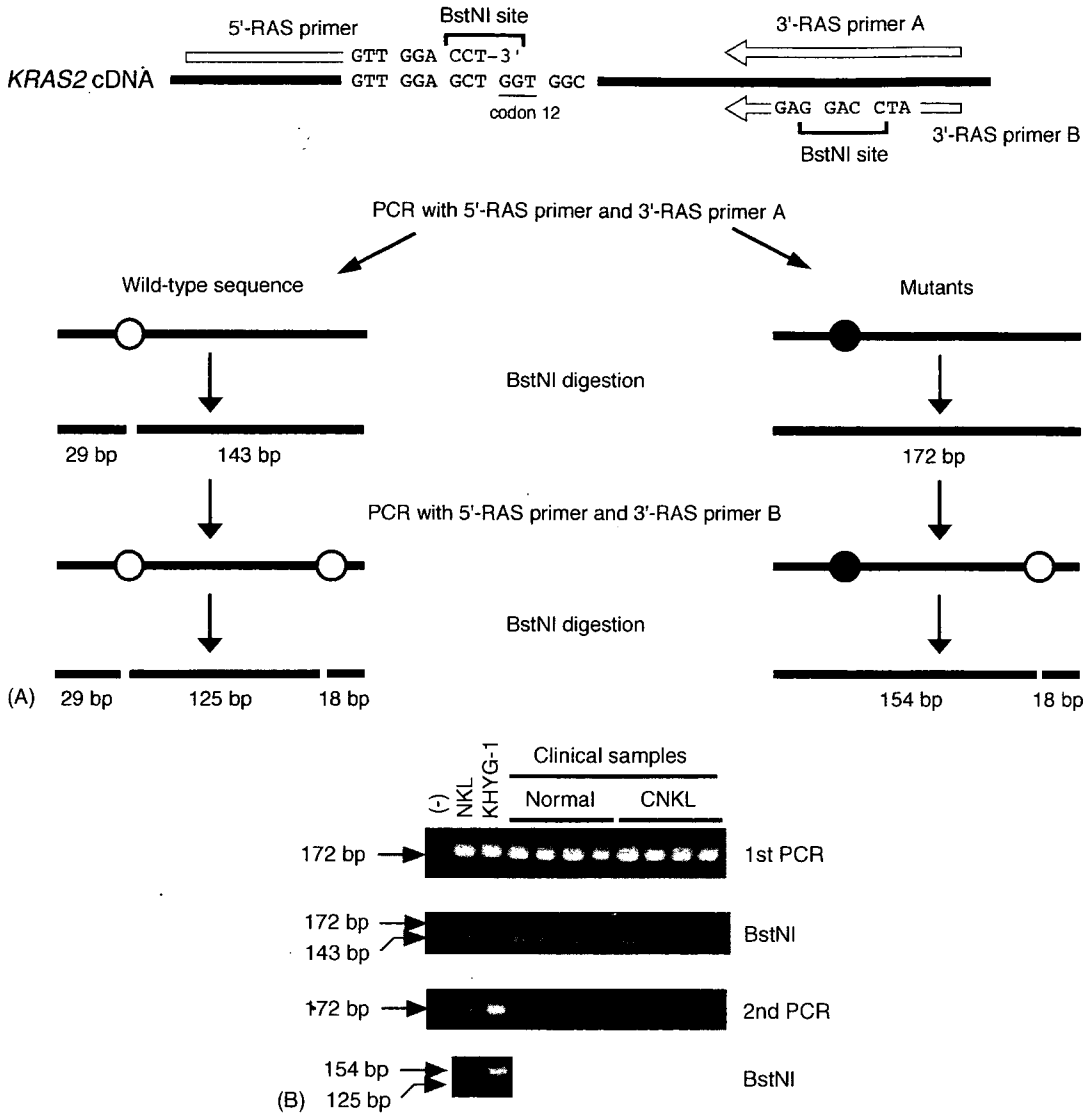

Fig. 3. Mutation-specific PCR analysis of NK cell leukemia cell lines and clinical specimens: (A) *KRAS2* cDNA was amplified with 5′-RAS primer and 3′-RAS primer A. The PCR product was subjected to digestion with *Bst*NI and then to a second PCR with 5′-RAS primer and 3′-RAS primer B. The second PCR product was also subjected to digestion with *Bst*NI. The nucleotides shown in red were incorporated into the primers to generate a *Bst*NI site. Open circles indicate *Bst*NI sites; closed circles indicate corresponding mutant sequences that are not susceptible to *Bst*NI; (B) cDNA isolated from the NKL and KHYG-1 cell lines as well as from $CD3^-CD56^+$ NK cell fractions derived from healthy volunteers (Normal) or individuals with CNKL was subjected to mutation-specific PCR analysis. A reaction without input DNA was also performed as a negative control (–). The size of DNA fragments is indicated on the left.

With this approach, we analyzed cDNA prepared from two ANKL cell lines (NKL and KHYG-1) and from $CD3^-CD56^+$ NK cell fractions purified from the peripheral blood of healthy individuals ($n=4$) and patients with CNKL ($n=4$). The first PCR step yielded a single DNA fragment of 172 bp from all samples. Furthermore, only the PCR product from KHYG-1 cells was refractory to *Bst*NI digestion, indicating that only KHYG-1 cells harbor a codon 12 mutation of *KRAS2*. The presence of the 143-bp band may indicate that KHYG-1 cells are heterozygous for the *KRAS2* mutation. The second PCR generated a 172-bp DNA fragment only with the NKL and KHYG-1 cell samples. Whereas this fragment derived from NKL cells was completely digested by *Bst*NI to generate a 125-bp band, *Bst*NI digestion of the fragment derived from KHYG-1 cells generated a band of 154 bp. Of all the samples analyzed, therefore, mutation of the first or second position of codon 12 of *KRAS2* was detected only in KHYG-1 cells.

4. Discussion

We have constructed a retroviral cDNA expression library for an ANKL cell line. Given that >97% (39/40) of the viral plasmids contained cDNA inserts and that the overall clone number was $>5 \times 10^6$, our library likely represents most of

the transcriptome of KHYG-1 cells. The high probability that the incorporated cDNAs are full length is also an important advantage for functional screening.

In our screening, most of the 3T3 transformants were found to have incorporated a single cDNA fragment corresponding to *KRAS2*^{G12A}, with only two transformants found to contain other cDNAs. One of these two cDNA inserts was derived from the gene for PFN1, a protein that binds to unpolymerized actin [17]. Homozygous deletion of *Pfn1* results in embryonic death in mice, with the encoded protein apparently being indispensable for cell growth or differentiation during embryonic development [18]. The other cDNA insert isolated from 3T3 transformants contained the entire open reading frame for IDH3B, which catalyzes the oxidative decarboxylation of isocitrate and is a key enzyme in the tricarboxylic acid cycle [19]. Neither *PFN1* nor *IDH3B* has previously been shown to possess oncogenic activity. It is currently under examination whether a long terminal repeat (LTR)-driven overexpression of *PFN1* or *IDH3B* leads focus formation in 3T3 cells.

Comparative genomic hybridization analysis identified a wide variety of genetic alterations at a high frequency in ANKL cells [20], suggesting that leukemogenesis in ANKL is associated with multiple steps of oncogene activation. An analysis of patients with NK cell neoplasia failed to detect any changes in the genes for members of the RAS and MYC families of proteins [7], however. This previous study did find a frequent increase in the abundance of the cell cycle regulator MDM2.

In contrast, we have detected a transforming *KRAS2* mutant gene in an ANKL cell line. Given that the mutation in codon 12 of this gene was detected by two different approaches (retroviral screening of PCR-amplified cDNAs and mutation-specific PCR), we conclude that it was not an artifact of PCR. KRAS2 is a GTP-binding protein with a relative molecular mass of ~21 kDa. Together with HRAS and NRAS, it plays an important role in cell growth and differentiation. Many mitogenic signals promote the loading of KRAS2 with GTP, which in turn triggers various downstream signaling events including activation of the mitogen-activated protein kinase (MAPK) pathway.

Activating mutations of *KRAS2* have been identified in a wide range of human cancers. Mutations of codon 12, for example, are associated with acute lymphoblastic leukemia [21], lung cancer [22], and pancreatic cancer [23]. No such mutations have previously been detected in association with ANKL, however. Although we have now identified a *KRAS2* mutation affecting codon 12 in the ANKL cell line KHYG-1, we did not detect this mutation in another ANKL cell line (NKL) or in $CD3^{-}CD56^{+}$ NK cell fractions isolated from healthy volunteers or from individuals with CNKL. Mutation of *KRAS2*, at least of codon 12 of this gene, might therefore be an infrequent event in NK cell neoplasia. Indeed, it remains possible that the detected *KRAS2* mutation is specific to the KHYG-1 cell line. Nevertheless, our identification of an activating *KRAS2* mutation in KHYG-1 cells might provide insight into the role of the RAS-MAPK signaling pathway in ANKL carcinogenesis. Furthermore, given the high quality of our retroviral expression library, the strategy adopted in the present study also might prove useful for the functional screening of genes associated with various clinical characteristics of ANKL, such as chemoresistance.

Acknowledgments

This work was supported in part by grants for Research on Human Genome and Tissue Engineering and for Third-Term Comprehensive Control Research for Cancer from the Ministry of Health, Labor, and Welfare of Japan, as well as by grants from Research Foundation for Community Medicine of Japan, Sankyo Foundation of Life Science, Takeda Science Foundation, and Mitsubishi Pharma Research Foundation. Y.-L.C conducted most of the experiments. R.M., M.O., A.I. and T.W. helped to establish a retroviral expression library. Hideki Makishima, J.O. and K.O. collected the ANKL specimens, and conducted the mutation-specific PCR method. H.K., R.K., K.K., M.I., S.T. and Y.Y. helped the 3T3 focus formation screening, and provided suggestions on molecular biology. Hiroyuki Mano designed this project with Y.-L.C., and was responsible for all aspects of this project.

References

[1] Harris NL, Jaffe ES, Diebold J, Flandrin G, Muller-Hermelink HK, Vardiman J, et al. World Health Organization classification of neoplastic diseases of the hematopoietic and lymphoid tissues: report of the Clinical Advisory Committee meeting, Airlie House, Virginia, November. J Clin Oncol 1999;17:3835–49.

[2] Oshimi K. Leukemia and lymphoma of natural killer lineage cells. Int J Hematol 2003;78:18–23.

[3] Gelb AB, van de Rijn M, Regula Jr DP, Cornbleet JP, Kamel OW, Horoupian DS, et al. Epstein–Barr virus-associated natural killer-large granular lymphocyte leukemia. Hum Pathol 1994;25:953–60.

[4] Sakajiri S, Kawamata N, Egashira M, Mori K, Oshimi K. Molecular analysis of tumor suppressor genes, Rb, p53, p16INK4A, p15INK4B and p14ARF in natural killer cell neoplasms. Jpn J Cancer Res 2001;92:1048–56.

[5] Takakuwa T, Dong Z, Nakatsuka S, Kojya S, Harabuchi Y, Yang WI, et al. Frequent mutations of Fas gene in nasal NK/T cell lymphoma. Oncogene 2002;21:4702–5.

[6] Hoshida Y, Hongyo T, Nakatsuka S, Nishiu M, Takakuwa T, Tomita Y, et al. Gene mutations in lymphoproliferative disorders of T and NK/T cell phenotypes developing in renal transplant patients. Lab Invest 2002;82:257–64.

[7] Sugimoto KJ, Kawamata N, Sakajiri S, Oshimi K. Molecular analysis of oncogenes, ras family genes (N-ras, K-ras, H-ras), myc family genes (c-myc, N-myc) and mdm2 in natural killer cell neoplasms. Jpn J Cancer Res 2002;93:1270–7.

[8] Hongyo T, Li T, Syaifudin M, Baskar R, Ikeda H, Kanakura Y, et al. Specific c-kit mutations in sinonasal natural killer/T-cell lymphoma in China and Japan. Cancer Res 2000;60:2345–7.

[9] Aaronson SA. Growth factors and cancer. Science 1991;254:1146–53.

[10] Yagita M, Huang CL, Umehara H, Matsuo Y, Tabata R, Miyake M, et al. A novel natural killer cell line (KHYG-1) from a patient with aggressive natural killer cell leukemia carrying a p53 point mutation. Leukemia 2000;14:922–30.
[11] Robertson MJ, Cochran KJ, Cameron C, Le JM, Tantravahi R, Ritz J, et al. Characterization of a cell line, NKL, derived from an aggressive human natural killer cell leukemia. Exp Hematol 1996;24:406–15.
[12] Pear WS, Nolan GP, Scott ML, Baltimore D. Production of high-titer helper-free retroviruses by transient transfection. Proc Natl Acad Sci USA 1993;90:8392–6.
[13] Onishi M, Kinoshita S, Morikawa Y, Shibuya A, Phillips J, Lanier LL, et al. Applications of retrovirus-mediated expression cloning. Exp Hematol 1996;24:324–9.
[14] Kahn SM, Jiang W, Culbertson TA, Weinstein IB, Williams GM, Tomita N, et al. Rapid and sensitive nonradioactive detection of mutant K-ras genes via 'enriched' PCR amplification. Oncogene 1991;6:1079–83.
[15] Kent WJ. BLAT—the BLAST-like alignment tool. Genome Res 2002;12:656–64.
[16] Ayllon V, Rebollo A. Ras-induced cellular events. Mol Membr Biol 2000;17:65–73.
[17] Goldschmidt-Clermont PJ, Janmey PA, Profilin. A weak CAP for actin and RAS. Cell 1991;66:419–21.
[18] Witke W, Sutherland JD, Sharpe A, Arai M, Kwiatkowski DJ. Profilin I is essential for cell survival and cell division in early mouse development. Proc Natl Acad Sci USA 2001;98:3832–6.
[19] Kim YO, Park SH, Kang YJ, Koh HJ, Kim SH, Park SY, et al. Assignment of mitochondrial NAD(+)-specific isocitrate dehydrogenase beta subunit gene (IDH3B) to human chromosome band 20p13 by in situ hybridization and radiation hybrid mapping. Cytogenet Cell Genet 1999;86:240–1.
[20] Siu LL, Wong KF, Chan JK, Kwong YL. Comparative genomic hybridization analysis of natural killer cell lymphoma/leukaemia. Recognition of consistent patterns of genetic alterations. Am J Pathol 1999;155:1419–25.
[21] Perentesis JP, Bhatia S, Boyle E, Shao Y, Shu XO, Steinbuch M, et al. RAS oncogene mutations and outcome of therapy for childhood acute lymphoblastic leukemia. Leukemia 2004;18:685–92.
[22] Santos E, Martin-Zanca D, Reddy EP, Pierotti MA, Della Porta G, Barbacid M. Malignant activation of a K-ras oncogene in lung carcinoma but not in normal tissue of the same patient. Science 1984;223:661–4.
[23] Motojima K, Urano T, Nagata Y, Shiku H, Tsurifune T, Kanematsu T. Detection of point mutations in the Kirsten-ras oncogene provides evidence for the multicentricity of pancreatic carcinoma. Ann Surg 1993;217:138–43.

Available online at www.sciencedirect.com
ScienceDirect
Methods 43 (2007) 118–122
METHODS
www.elsevier.com/locate/ymeth

mRAP, a sensitive method for determination of microRNA expression profiles

Hiroyuki Mano [a,b,*], Shuji Takada [a]

[a] *Division of Functional Genomics, Jichi Medical University, 3311-1 Yakushiji, Shimotsukeshi, Tochigi 329-0498, Japan*
[b] *CREST, Japan Science and Technology Agency, Saitama 332-0012, Japan*

Accepted 13 April 2007

Abstract

MicroRNAs (miRNAs) are noncoding RNA molecules of 21–24 nucleotides that regulate the expression of target genes in a posttranscriptional manner. Although evidence indicates that miRNAs play essential roles in embryogenesis, cell differentiation, and pathogenesis of human diseases, extensive miRNA profiling in cells or tissues has been hampered by the lack of sensitive cloning methods. Here we describe a highly efficient profiling strategy, termed miRNA amplification profiling (mRAP), that relies on the use of a long, optimized 5′ adaptor, the SMART (switching mechanism at the 5′ end of RNA templates of reverse transcriptase) method, the polymerase chain reaction, and cDNA concatamerization after BanI digestion. This approach is highly sensitive, readily allowing the isolation of $>1 \times 10^4$ independent miRNA-derived cDNAs from $\leqslant 1 \times 10^4$ cells. The mRAP method thus makes it possible to analyze miRNA expression profiles for small quantities of tissue or cells such as fresh clinical specimens.

Keywords: MicroRNA; SMART; Cloning; cDNA library

1. Introduction

MicroRNAs (miRNAs) are short noncoding RNAs that bind target mRNAs by incomplete base-pairing to their 3′ untranslated regions. Mainly through translational repression, miRNAs regulate the expression of target genes and thereby have a plethora of effects on cellular functions [1,2]. Given that changes in miRNA expression have been implicated in tumorigenesis [3,4], it is important to characterize miRNA expression profiles in fresh cancer specimens.

Isolation of miRNA clones has not been a simple task with limited amounts of tissue or small numbers of cells, however. Given that conventional miRNA isolation procedures require $\geqslant 100$ μg of total RNA as starting material [5], several strategies have been employed to overcome this obstacle with the use of stem-loop reverse transcription-polymerase chain reaction (RT-PCR) analysis [6], the SMART (switching mechanism at the 5′ end of RNA templates of reverse transcriptase) method [7], serial analysis of gene expression (SAGE) [8], and bead-based flow cytometry [9], for example.

We have attempted to isolate cDNAs for miRNAs from fresh clinical specimens. A simple application of the SMART method to cDNA cloning did not allow an efficiency sufficient for the isolation of miRNA-derived clones, especially from small numbers of cells. We therefore developed a highly sensitive miRNA cloning strategy, mRAP (miRNA amplification profiling), that relies on a long, optimized 5′ adaptor and the SMART method. This approach readily allows the isolation of $>1 \times 10^4$ miRNA-derived clones even from $\leqslant 1 \times 10^4$ cells, making it possible to determine miRNA expression profiles for small quantities of tissue or cells such as fresh clinical specimens.

* Corresponding author. Address: Division of Functional Genomics, Jichi Medical University, 3311-1 Yakushiji, Shimotsukeshi, Tochigi 329-0498, Japan. Fax: +81 285 44 7322.
E-mail address: hmano@jichi.ac.jp (H. Mano).

1046-2023/$ - see front matter © 2007 Elsevier Inc. All rights reserved.
doi:10.1016/j.ymeth.2007.04.006

2. Isolation and fractionation of small RNA molecules

A small-RNA fraction is isolated from cells or tissue, the RNA molecules are separated by denaturing polyacrylamide gel electrophoresis (PAGE), and the region of the gel containing RNA molecules of 19–24 nucleotides (nt) is excised. The portion of the gel containing the sample RNA should not be stained with a dye or exposed to ultraviolet light.

1. Purify small RNA molecules with the use of a *mir*Vana miRNA isolation kit (Ambion, Austin, TX) and the small-RNA fraction option. Other protocols or kits suitable for the enrichment of small RNAs may be used.
2. Prepare the sample mixture: 5 μl of isolated RNA (≤5 μg) and 5 μl of 2× Dye (*mir*Vana).
3. Prepare the size-standard mixture: 0.25 μl of size standards (synthesized RNA oligonucleotides of 19, 24, and 33 nt at a concentration of 0.1 mg/ml each), 4.75 μl of water, and 5 μl of 2× Dye (*mir*Vana).
4. Incubate each mixture at 90 °C for 20 s and then place on ice.
5. Subject the sample and size-standard mixtures to PAGE on a 15% gel (SequaGel Sequencing System; National Diagnostics, Atlanta, GA).
6. Cut out and stain only the marker lane with SYBR Green II (Takara Bio, Shiga, Japan) or with ethidium bromide for 5–10 min and then photograph the gel aligned with a ruler.
7. Excise the portion of the sample lane containing RNA molecules of 19–24 nt.
8. Chop the isolated gel into small fragments and then transfer them to a microcentrifuge tube.
9. Add 125 μl of water to the tube and maintain it at 4 °C overnight with gentle agitation.
10. Separate the gel pieces by brief centrifugation, and transfer 100 μl of the supernatant containing the small RNA molecules to another microcentrifuge tube.
11. Add 250 μl of 2-butanol to the RNA, invert the tube several times, and centrifuge the mixture briefly.
12. Check the volume of the bottom, water phase (should be <10 μl) and discard the upper, 2-butanol phase.
13. Subject the water phase to chloroform extraction followed by ethanol precipitation with 1 μl of glycogen (Roche Diagnostics, Mannheim, Germany; 20 mg/ml stock) as a carrier.
14. Dissolve the purified RNA molecules in 8.75 μl of water.

3. Synthesis of cDNAs, PCR, and concatamer formation

The mRAP procedure is depicted schematically in Fig. 1. The miRNAs isolated by gel electrophoresis are ligated at their 3′ ends to a 3′ adaptor with the use of RNA ligase. Complementary DNAs corresponding to the miRNAs are then synthesized with the use of reverse transcriptase and an RT primer complementary to the 3′ adaptor. Given that some reverse transcriptases possess terminal deoxynucleotidyl transferase activity, the synthesized cDNA strands frequently have a small poly(C) overhang at their 3′ ends. After annealing of a long 5′ adaptor to such poly(C) overhangs, PCR is used to amplify the miRNA-derived cDNAs. As shown in Fig. 1, there are usually three main types of PCR product of different sizes (~70, ~90, and ~120 bp). The ~90-bp products are the ones that contain the cDNAs; the smaller and larger products comprise dimers and trimers of the 5′ adaptor and 3′ adaptor (without cDNA). The desired products are purified, digested with the restriction endonuclease BanI (the target sites of the enzyme are incorporated into the PCR primers), and ligated to generate cDNA concatamers.

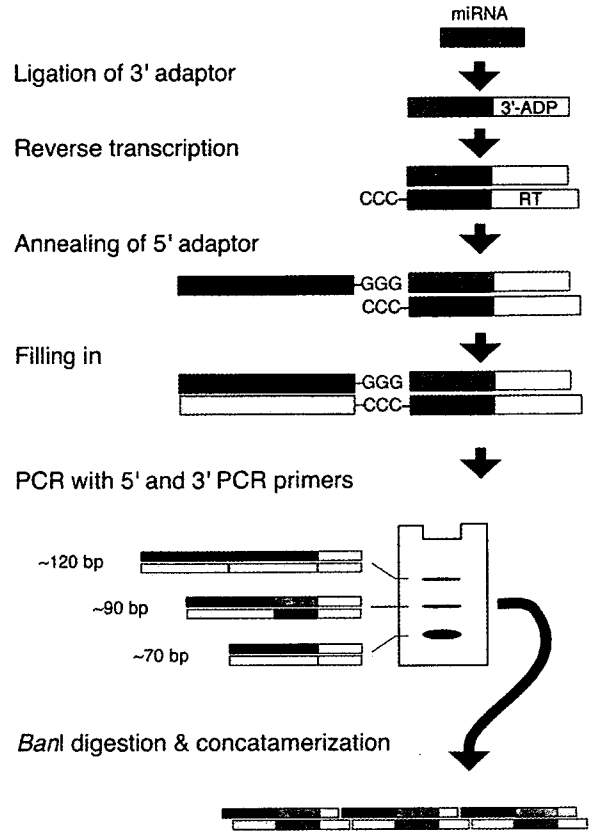

Fig. 1. The mRAP protocol. Isolated small RNA molecules are ligated to the 3′ adaptor (3′-ADP) and subjected to reverse transcription with the RT primer. After annealing of the 5′ adaptor (5′-ADP) to the poly(C) overhang at the 3′ end of the synthesized cDNAs, the latter are subjected to PCR with the 5′ and 3′ PCR primers. Of the three major sizes of amplicon generated, only the middle one includes cDNAs derived from miRNAs. These cDNAs are isolated, digested with BanI, and self-ligated to yield concatamers. Reproduced from [10] with permission of Oxford University Press.

3.1. Required oligonucleotides

3′ adaptor: 5′-(Pu)uuAACCGCGAATTCCAG(idT)-3′ (lowercase letters indicate RNA, uppercase letters

indicate DNA, Pu denotes 5′-phosphorylated uridine, and idT represents 3′-inverted deoxythymidine [Dharmacon, Chicago, IL])
5′ adaptor: 5′-GACCACGCGTATCGGGCACCACG TATGCTATCGATCGTGAGATGGG-3′
RT primer: 5′-GACTAGCTGGAATTCGCGGTTAA A-3′
5′ PCR primer: 5′-GCGTATCGGGCACCACGTAT GC-3′
3′ PCR primer: 5′-GACTAGCTTGGTGCCGAATTC GCGGTTAAA-3′

3.2. Synthesis of cDNAs from miRNAs

1. Mix the small RNAs (8.75 μl) with 1 μl of 10× NEB buffer 3 (New England Biolabs, Ipswich, MA) and 0.25 μl of calf intestinal alkaline phosphatase (New England Biolabs) in a microcentrifuge tube, and incubate the mixture at 50 °C for 30 min.
2. Subject the mixture to phenol–chloroform extraction and chloroform extraction followed by ethanol precipitation in the presence of 1 μl of glycogen (20 mg/ml stock) and 0.5 μl of 100 μM 3′ adaptor.
3. Isolate the RNA by centrifugation and allow it to dry before dissolving it in 3 μl of water.
4. To the dissolved RNA, add the following: 1 μl of 10× Ligation Buffer (New England Biolabs), 1 μl of acetylated bovine serum albumin (Invitrogen, Carlsbad, CA; 1 mg/ml stock), 1 μl of 1 mM ATP, and 3 μl of 50% dimethyl sulfoxide. Incubate the mixture at 90 °C for 30 s and then place it on ice for 20 s.
5. Add 1 μl of T4 RNA ligase (New England Biolabs), and incubate the mixture at 37 °C for 1 h. Perform phenol–chloroform extraction, chloroform extraction, and ethanol precipitation. Isolate the RNA by centrifugation and dissolve it in 4 μl of water.
6. Add 0.5 μl of 100 μM RT primer and 0.5 μl of 100 μM 5′ adaptor. Incubate the mixture at 70 °C for 2 min, and then place it on ice.
7. Add the following: 2 μl of 5× RT buffer (Clontech, Mountain View, CA), 1 μl of 0.1 M dithiothreitol, 1 μl of a mixture of the four deoxynucleoside triphosphates (dNTPs) each at 2.5 mM, and 1 μl of PowerScript reverse transcriptase (Clontech). Incubate the resulting mixture at 42 °C for 1 h.
8. Add 40 μl of water, and incubate the mixture at 72 °C for 7 min. Add 160 μl of water.

3.3. PCR amplification of miRNA-derived cDNAs

1. Mix the following: 49 μl of cDNA solution, 1706 μl of water , 245 μl of 10× AmpliTaq buffer (Applied Biosystems, Foster City, CA), 245 μl of a mixture of the four dNTPs each at 2 mM, 196 μl of a mixture of the 5′ PCR primer and 3′ PCR primer each at 10 μM, and 9.8 μl of AmpliTaq Gold DNA polymerase (Applied Biosystems).
2. Transfer 50 μl of the reaction mixture to each of 48 PCR tubes (0.5-ml scale).
3. Perform PCR with an initial incubation at 95 °C for 4.5 min; 32 cycles of 95 °C for 30 s and 65 °C for 30 s; and a final incubation at 72 °C for 5 min.
4. Subject each reaction mixture to sodium acetate–ethanol precipitation as follows: add 1/10th volume of 3 M NaOAc (pH 5.5), mix, add two volumes of 100% ethanol, mix, and chill at −70 °C for 15 min. Centrifuge the tubes, and dissolve the final precipitates in water.

3.4. Concatamer formation

1. Subject the PCR products and size standards to PAGE on a 10% gel under nondenaturing conditions.
2. Stain the gel with ethidium bromide, and excise the portion of the gel containing DNA molecules of 90–95 bp.
3. Chop the gel portion into small fragments, transfer the fragments to a microcentrifuge tube, and add 200 μl of 0.3 M NaCl. Incubate the mixture at 37 °C for $\geqslant$8 h.
4. Briefly centrifuge the tube and harvest the supernatant containing the cDNAs.
5. Subject the cDNA preparation to ethanol precipitation (in the presence of glycogen). Isolate the DNA by centrifugation and dissolve it in 43 μl of water.
6. Add 5 μl of 10× K buffer (New England Biolabs) and 2 μl of BanI (New England Biolabs; 20 U/μl stock). Incubate the mixture at 37 °C for 2 h or overnight.
7. Purify the DNA with ProbeQuant G50 (GE Healthcare Bio-Sciences, Uppsala, Sweden) and then subject it to phenol-chloroform extraction, chloroform extraction, and ethanol precipitation (in the presence of glycogen). Isolate the DNA by centrifugation and dissolve it in 2 μl of water.
8. Add 2 μl of Ligation High solution (Toyobo, Tokyo, Japan), and incubate the resulting mixture at 16 °C for 4 h or overnight.
9. Twenty minutes before the end of the ligation incubation, prepare the following mixture (Taq mix): 76 μl of water, 10 μl of 10 × NH_4 buffer (Bioline, London, UK), 3 μl of 50 mM $MgCl_2$, 10 μl of a mixture of the four dNTPs each at 2 mM, and 1 μl of BioTaq DNA polymerase (Bioline).
10. Incubate the Taq mix at 95 °C for 10 min and then maintain it at 72 °C.
11. Add the ligation mixture to the Taq mix and incubate at 72 °C for an additional 30 min.
12. Subject the mixture to ammonium acetate–ethanol precipitation as follows: add 1/4th volume of 10 M ammonium acetate (i.e., 2 M final), mix, add two volumes of 100% ethanol, mix, and chill at −70 °C for 15 min. Centrifuge the tubes, and dissolve the final precipitates in water.

13. Fractionate the DNA molecules by PAGE on a 10% gel under nondenaturing conditions.
14. Excise the portion of the gel containing DNA molecules of 500–2000 bp.
15. Chop the excised region of the gel into small fragments and transfer them to a D-tube (EMD Biosciences, San Diego, CA). Subject the fragments to electrophoretic elution at 100 V for 4 h in Tris–borate–EDTA buffer.
16. Harvest the solution and subject it to sodium acetate–ethanol precipitation in the presence of glycogen.
17. Isolate the cDNA concatamers by centrifugation and dissolve them in water. Ligate the concatamers into a TA cloning vector such as pGEM-Teasy (Promega, Madison, WI) for nucleotide sequencing.

4. Sequencing of mRAP products

We usually first assess the quality of the mRAP plasmid library by sequencing the inserts of ~100 randomly isolated plasmids. The plasmid inserts typically contain two to six cDNAs (Fig. 2). BLAST searching of the insert cDNAs against the genome sequence of interest has revealed that about one-third of the cDNAs correspond to miRNAs, one-third to ribosomal or transfer RNA, and the remaining one-third to other sequences [10]. In our experience, a high proportion of "other sequences" indicates poor quality of the initial RNA preparation. Abundant miRNA sequences can be obtained even from $\leqslant 1 \times 10^4$ cells, provided that the isolated RNA is in good condition.

Our initial trial with mRAP to characterize the entire miRNA repertoire of the mouse resulted in the identification of a total of 77,736 small RNA reads [10]. Even at this scale, many miRNAs (both known and predicted) were present in only one read per tissue, indicating that our trial did not exhaust the miRNA repertoire. It might therefore be of interest to couple mRAP to recently developed high-throughput sequencing technologies. Pyrosequencing or other recent methods are capable of simultaneously determining short nucleotide sequences for 1×10^5–1×10^7 clones [11]. Sequencing of mRAP products on such a scale should allow characterization of a complete body map of miRNA profiles for any organism.

5. Troubleshooting

5.1. Three bands are not observed in the gel after RT-PCR at step 3.4.2

If small amounts of starting material, such as fresh clinical specimens, are used, distinct bands around 90–95 bp may not be observed at step 3.4.2; instead a smear may appear. Since the cDNA/adaptor products should exist in this smear (at 90–95 bp), cut the corresponding region according to the size marker. Since this process may result in cutting out a region of an incorrect size, also individually cut out the fractions of the gel above and below this region. Treat these three fractions in parallel in succeeding steps described above. Sequencing reveals which fraction contains the targets.

5.2. Most of the sequences obtained are derived from the adaptors used in mRAP

The reason for this is that ~70 bp and/or ~120 bp bands are contaminated at step 3.4.2. Contamination of ~70 bp or ~120 bp bands leads to obtaining the sequences for "5′ adaptor-3′ adaptor" or "5′ adaptor-5′ adaptor-3′ adaptor", respectively. Since abundances of the adaptor-derived bands (~70 and ~120 bp) are much greater than that of the desired adaptor/cDNA band (90–95 bp), cut out only the gel at 90–95 bp without even touching the ~70 and ~120 bp bands. An extended time of electrophoresis facilitates separating adaptors and cDNA/adaptor bands. For this purpose electrophoresis can run until BPB dye is about to run out the bottom edge of the gel.

6. Concluding remarks

Given the emerging roles of miRNA in animal development or pathogenesis of human disorders, it is mandatory

GCACCACGTATGCTATCGATCCTGAGATGGGTGTTGCACTTGTCCCGGCCTGTTTTAACC
GCGAATTCGGCACCACGTATGCTATCGATCGTGAGATGGGGTAGTGTTTCCTACTTTATG
GATTTAACCGCGAATTCGGCACCACGTATGCTATCGATCGTGAGATGGGCGGGCGGCGGC
GGTCGGCGGGCTTTAACCGCGAATTCGGCACCACGTATGCTATCGATCGTGAGATGGGGA
AGCGGGTTTTAACCGCGAATTCGGCACCACGTATGCTATCGATCGTGAGATGGGGTAGTG
TTTCCTACTTTATGGATTTAACCGCGAATTCGGCACCACGTATGCTATCGATCGTGAGAT
GGGGTAGTGTTTCCTACTTTATGGAAATTTAACCGCGAATTCGGCACCACGTATGCTATC
GATCGTGAGATGGGGGGCTGGGGCGCGAAGCGGGGCTTTAACCGCGAATTCGGCACCACG
TATGCTATCGATCGTGAGATGGGGGGTCGGGGCGGCGGCGGCGGTTTAACCGCGAAT
TCGGCACCACGTATGCTATCGATCGTGAGATGGGTGTGCAAATCTATGCAAAACTG

Fig. 2. Example of a plasmid insert of an mRAP library. The nucleotide sequence of one insert in a library prepared from human Jurkat cells is shown. Black letters indicate the sequences of adaptors, whereas red or blue letters indicate the cDNA sequences for miRNAs or for other sequences (ribosomal RNA or genomic DNA sequences), respectively. BanI sites are underlined.