AF5q31, a newly identified *AF4*-related gene, is fused to *MLL* in infant acute lymphoblastic leukemia with ins(5;11)(q31;q13q23)

Tomohiko Taki*, Hirotsugu Kano*, Masafumi Taniwaki†, Masahiro Sako‡, Masayoshi Yanagisawa*, and Yasuhide Hayashi*§

*Department of Pediatrics, Faculty of Medicine, University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8655, Japan; †3rd Department of Internal Medicine, Kyoto Prefectural University of Medicine, Kawaramachi-Hirokoji, Kamigyo-ku, Kyoto 602-8566, Japan; and ‡Department of Pediatrics, Osaka City General Hospital, 2-13-22 Miyakojimahon-dori, Miyakojima-ku, Osaka 534-0021, Japan

Edited by Janet D. Rowley, The University of Chicago Medical Center, Chicago, IL, and approved October 12, 1999 (received for review August 10, 1999)

Infant acute lymphoblastic leukemia (ALL) with *MLL* gene rearrangements is characterized by early pre-B phenotype (CD10$^-$/CD19$^+$) and poor treatment outcome. The t(4;11), creating *MLL-AF4* chimeric transcripts, is the predominant 11q23 chromosome translocation in infant ALL and is associated with extremely poor prognosis as compared with other 11q23 translocations. We analyzed an infant early preB ALL with ins(5;11)(q31;q13q23) and identified the *AF5q31* gene on chromosome 5q31 as a fusion partner of the *MLL* gene. The *AF5q31* gene, which encoded a protein of 1,163 aa, was located in the vicinity of the cytokine cluster region of chromosome 5q31 and contained at least 16 exons. The *AF5q31* gene was expressed in fetal heart, lung, and brain at relatively high levels and fetal liver at a low level, but the expression in these tissues decreased in adults. The AF5q31 protein was homologous to AF4-related proteins, including AF4, LAF4, and FMR2. The AF5q31 and AF4 proteins had three homologous regions, including the transactivation domain of AF4, and the breakpoint of AF5q31 was located within the region homologous to the transactivation domain of AF4. Furthermore, the clinical features of this patient with the *MLL-AF5q31* fusion transcript, characterized by the early pre-B phenotype (CD10$^-$/CD19$^+$) and poor outcome, were similar to those of patients having *MLL-AF4* chimeric transcripts. These findings suggest that AF5q31 and AF4 might define a new family particularly involved in the pathogenesis of 11q23-associated-ALL.

The *MLL* gene (also called *ALL-1*, *HRX*, and *HTRX-1*) has been identified in 11q23 translocations (1–4) and is considered to be a transcriptional maintenance factor (5). At least 30 chromosomal regions for partners of 11q23 have been observed, such as t(4;11), t(9;11), and t(11;19) (6). *MLL* gene rearrangement, which is found in the majority of infant (7, 8) and therapy-related leukemias (9, 10) is strongly associated with poor outcome in infants with acute lymphoblastic leukemia (ALL) compared with older children and adults with ALL, or with acute myeloid leukemia (AML). Of these, t(4;11)(q21;q23) creating a *MLL-AF4* fusion transcript was reported to be the most important prognostic factor in infant ALL (11). Currently almost 20 partner genes for *MLL* have been cloned from leukemia cells with various types of reciprocal 11q23 translocations (12). The functions of some genes have been revealed, including those for a Ras-binding protein (AF6) (13), an RNA polymerase II elongation factor (ELL/MEN) (14), transcriptional coactivator/histone acetyltransferase (CBP and p300) (15–17), and ABL and eps8-binding protein (ABI-1) (18). It has been shown both by "knock-in" mice and by retroviral transformation of murine bone marrow that *MLL* and the partner gene both are required for leukemogenesis (19, 20).

Hitherto, only a few genes have been identified as fusion partners of the *MLL* gene from ALL with 11q23 translocations (12). In the present study, we analyzed an infant ALL with ins(5;11)(q31;q13q23) and identified a novel *AF4*-related gene as a fusion partner of the *MLL* gene.

Materials and Methods

Patient. A 4-month-old girl was diagnosed as having ALL. Leukemic cells expressed CD15, CD19, and HLA-DR but not CD10, and were cytogenetically characterized as ins(5;11)(q31;q13q23), i(17q). She achieved a complete remission by intensive chemotherapy, but relapsed three times and died 20 months after diagnosis.

Southern Blot Analysis. High molecular weight DNA was extracted from bone marrow cells from the patient by proteinase K digestion and phenol/chloroform extraction. Ten micrograms of DNA was digested with appropriate restriction enzymes, subjected to electrophoresis on 0.8% agarose gels, transferred to charged nylon filters (Amersham Pharmacia), and hybridized to DNA probes labeled by the random hexamer method (15). A 0.9-kb *Bam*HI fragment (designated probe x) derived from *MLL* cDNA was used as a probe (21).

Preparation of mRNA and cDNA Libraries. Poly(A)$^+$ RNA from frozen cells was extracted with a Fast Track mRNA Isolation Kit (Invitrogen). A cDNA library was constructed with poly(A)$^+$ mRNA from patient cells and the BALM14 cell line following established procedures (18). Briefly, random hexanucleotide-primed synthesized cDNAs were ligated with *Eco*RI adaptors and cloned into the *Eco*RI-digested λgt10 cloning vector (Promega). After packaging with commercial packaging kits (Epicentre Technologies, Madison, WI), phage plaques were screened with probes labeled by using a random primer synthesis kit (Stratagene). The probe x was used for screening the patient cDNA library, and the 114-bp *AF5q31* cDNA probe derived from the *MLL-AF5q31* chimeric clone was used for screening the BALM14 and human placenta cDNA libraries.

Reverse Transcription–PCR and Genomic PCR. Total cellular RNA was extracted from bone marrow cells of the patient by the acid guanidine isothiocyanate-phenol-chloroform method (22). Four micrograms of total RNA was reverse transcribed to cDNA in a total volume of 20 μl with random hexamers and 20 units of reverse transcriptase (avian myeloblastosis virus) (Boehringer Mannheim). One-twentieth of the cDNA was amplified by PCR in a total volume of 100 μl with 50 mM KCl, 1.5 mM $MgCl_2$, 10 mM Tris·HCl (pH 9.0 at room temperature), 25 pmol of each primer, 75 μM of each dNTP, and 1 unit of *Taq* polymerase (Applied Biosystems). After 35 rounds of PCR (30 sec at 94°C,

MEDICAL SCIENCES

This paper was submitted directly (Track II) to the PNAS office.

Abbreviations: ALL, acute lymphoblastic leukemia; AML, acute myeloid leukemia.

Data deposition: The sequence reported in this paper has been deposited in the GenBank database (accession no. AF197927).

§To whom reprint requests should be addressed. E-mail: hayashi@ped.h.u-tokyo.ac.jp.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

30 sec at 55°C, and 1 min at 72°C), 5 μl of PCR product was electrophoresed in a 3% agarose gel. The primers used were as follows: MLL-7S, 5′-TCCTCAGCACTCTCTCCAAT-3′; MLL-11A, 5′-TTTGCCTGGAGTTGTGGATC-3′; 5–1A, 5′-CCATCACTGTCTTCACTGCT-3′; and 5–13S, 5′-ACACCATGCAAAACAGAACCT-3′.

One hundred nanograms of genomic DNA derived from the patient's leukemic cells was amplified under the same condition as reverse transcription–PCR. The primers used were MLL-7S and 5–21A, 5′-TGCCAAATCTAAATGACCTGG-3′.

Nucleotide Sequencing. PCR products were cloned into the TA cloning vector (Invitrogen). Nucleotide sequences of phage clones and PCR products were determined by the fluorometric method (18) (Dye Terminator Cycle Sequencing Kit, Applied Biosystems).

Northern Blot Analysis. Multiple human tissue Northern blots (CLONTECH) were hybridized with ^{32}P-labeled 0.6-kb *AF5q31* cDNA probe, which covered nucleotides 1,135 to 1,735, and 0.6-kb *AF4* cDNA probe, which covered nucleotides 476 to 1,117 (GenBank accession number L13773), respectively.

Fluorescence *in Situ* Hybridization (FISH) Analysis. Chromosomal mapping of the genomic clone λH17–6 was performed by the FISH method (23). The phage clone was labeled by the standard nick-translation method using biotin-16-dUTP (Boehringer Mannheim). Chromosomal *in situ* suppression hybridization, washing, and signal detection procedures were performed as described (23). Preparations were analyzed under a conventional fluorescence microscope (BX40-RF; Olympus, Tokyo), and images were captured with a charge-coupled device camera (SenSys0400-G1, Photometrics, Tucson, AZ). Each chromosomal band was identified based on 4′,6-diaminido-2-phenylindole dihydrochloride (DAPI) staining properties.

Results

Isolation of the *MLL* Fusion cDNAs in ins(5;11)(q31;q13q23). Southern blot analysis of DNA prepared from the leukemic cells of the patient using probe x revealed a chromosomal breakpoint within the breakpoint cluster region of the *MLL* gene at 11q23 (Fig. 1), which spanned exons 8–14 in the *MLL* gene (nomenclature according to ref. 24). To isolate fusion transcripts of *MLL*, we prepared a cDNA library from mRNA of the patient's leukemic cells. Four cDNA clones were isolated by screening with probe x, and one (clone 17–6) of them was found to represent a fusion transcript of *MLL*. Clone 17–6, 534 bp in size, contained a 420-bp sequence corresponding to exons 8–10 in the *MLL* gene at the 5′ region, and the remaining 114-bp sequence did not match the *MLL* gene or the partner genes of *MLL* previously cloned (Fig. 2*a*). We could not isolate any 3′-*MLL* fusion clones.

Isolation of the *AF5q31* Gene Encoding a Protein That Is Similar to AF4 Protein. The 114-bp sequence identified in the chimeric clone was used as a probe to screen cDNA libraries from the BALM14 cell line and human placenta. We isolated two overlapping clones that spanned the complete coding region and encoded a protein of 1,163 aa with a predicted molecular mass of 127,457 Da (Fig. 2 *b* and *c*). Sequence comparisons of the predicted AF5q31 protein using the BLAST file showed partial similarity to the AF4 protein, which is a fusion partner of the *MLL* gene in t(4;11)(q21;q23) (Fig. 3*a*).

Detection of the *MLL-AF5q31* Fusion Transcripts and the Genomic Junction of the Breakpoint. Using a sense primer from *MLL* exon 10 (MLL-7S) and an antisense primer from *AF5q31* (5–1A), we obtained a PCR product of 214 bp from the patient (Fig. 4). However, reciprocal PCR products of *AF5q31-MLL* fusion transcripts were not generated by reverse transcription–PCR (Fig. 4). Furthermore, we cloned a genomic junction of the breakpoint by genomic PCR followed by sequencing (Fig. 2*d*).

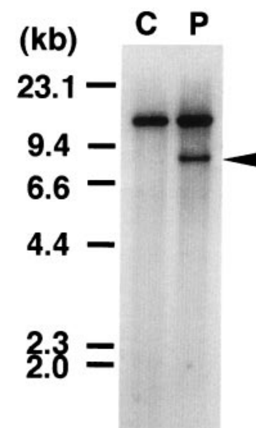

Fig. 1. Southern blot of DNA digested with *Hin*dIII and probed with the 0.9-kb fragment of the *MLL* gene. C, normal peripheral lymphocytes. P, leukemic cells from the patient. The patient exhibited a rearranged band (arrow) with this probe.

Expression of the *AF5q31* Gene Compared with *AF4* Gene in Normal Tissues. To examine for expression of the *AF5q31* gene, we performed Northern blot analysis on poly(A)$^+$ RNA from various human tissues. Expression of the 9.5-kb transcript was detected in adult heart, placenta, skeletal muscle, and pancreas, and fetal heart, lung, and brain at relatively high levels, and in adult brain and fetal liver at low levels (Fig. 5). We also performed Northern blot analysis on the same blots using *AF4* cDNA probe. Expression of the 9.5-kb transcript was detected in adult heart, placenta, skeletal muscle, and pancreas, and fetal lung, liver, and brain at relatively high levels, and in fetal heart at a low level (Fig. 5).

Chromosomal Assignment of the *AF5q31* Gene. To assign the chromosomal location for the *AF5q31* gene, we obtained a phage clone (λH17–6) after screening of a genomic library from human placental DNA using an *AF5q31* cDNA probe. The phage clone showed specific signals at band 5q31.1 in all 25 metaphase cells tested (Fig. 6).

***AF5q31* Gene Is Located in the Vicinity of the Cytokine Cluster Region of Chromosome 5q31.** By BLAST search of GenBank, we found that the sequence of a P1 clone from chromosome 5q31 partially matched the sequence of *AF5q31* cDNA. Sequence comparison between *AF5q31* and this P1 clone revealed that the P1 clone contained the 3′ region of the breakpoint in the *AF5q31* gene and that the *AF5q31* gene contained at least 16 exons. Furthermore, six overlapping P1 clones covered a genomic region of about 380 kb, containing the *GDF-9*, *KIF3A*, *IL-4*, *IL-13*, *RAD50*, and *IL-5* genes on chromosome 5q31 (Fig. 2*e*).

Discussion

In the present study, we isolated a novel fusion partner of the *MLL* gene, *AF5q31*, in an infant ALL with ins(5;11)(q31;q13q23). In the patient's leukemic cells, only an *MLL-AF5q31* fusion transcript was detected, but not an *AF5q31-MLL* transcript. Although we could not exclude the possibility of the presence of any other 3′-*MLL* fusion transcripts, the 5′-*MLL-AF5q31*-3′ transcript is thought to be critical in leukemogenesis, as described (19, 20).

a

MLL exon 10 ▼ AF5q31

```
CACAGGATCAGAGTGGACTTTAAGGAGTCTCAGCAGTCCAATTTTGGC
H  R  I  R  V  D  F  K  E  S  Q  Q  S  N  F  G
```

b

AF5q31 — ATG ▼ TGA

B16

P2-1

1kb

c

```
   1  AGATTCGCGGCCGCGTCGACCGCCAGCGGACGGGGCGGGGATCCTGGTGCCAGGTCTTCGTTAACCCGCCCCCGCCTGCCTCAGTCAGCTGTCCCAGCCCGGGCCCGACCGTCATGGAGC
 121  GGTGCGGGGCCTACGCCTGTTCCGTGTGAGGCGCATCGCTGCCGCCGCCGCCGACGCCGCCTTCGTTTGTCGAGGCCTCCGGAACCCCCGCCGCCTCCCGGAGCCCCTCGCCCTTGGCAG
 241  CCTGTCACCGCCCCCTGGGCGGGCTCGAATGCGCGTCCGGTGAAGGTGCAGGCCCGGCGCCGCCACTGCCGCAGCCAGGAGATGGTTCGGGCCTAGCGGAGCCGGGACTGGAGCAACATG
   1                                                                                                                          M
 361  AACCGTGAAGACCGGAATGTGCTGCGTATGAAAGAACGGGAAAGGCGGAATCAGGAAATTCAGCAGGGCGAAGACGCCTTCCCACCTAGCTCTCCTCTCTTTGCAGAGCCATACAAAGTT
   2  N R E D R N V L R M K E R E R R N Q E I Q Q G E D A F P P S S P L F A E P Y K V
 481  ACTAGCAAAGAAGATAAGTTATCAAGTCGTATTCAGAGTATGCTTGGAAACTACGATGAAATGAAGGATTTCATAGGAGACAGATCTATACCAAAGCTTGTTGCAATTCCCAAGCCTACA
  42  T S K E D K L S S R I Q S M L G N Y D E M K D F I G D R S I P K L V A I P K P T
 601  GTACCACCATCAGCAGATGAAAAATCTAACCCAAATTTCTTTGAACAGAGACATGGAGGCTCTCATCAGAGTAGCAAATGGACTCCAGTAGGACCCGCACCCAGCACTTCTCAGTCTCAG
  82  V P P S A D E K S N P N F F E Q R H G G S H Q S S K W T P V G P A P S T S Q S Q
 721  AAACGGTCCTCAGGCTTACAGAGTGGACATAGTAGCCAGCGGACCAGCGCAGGTAGCAGTAGTGGCACTAACAGTAGTGGTCAGAGGCACGACCGTGAGTCATATAACAATAGTGGAGC
 122  K R S S G L Q S G H S S Q R T S A G S S S G T N S S G Q R H D R E S Y N N S G S
 841  AGTAGCCGGAAAAAAGGCCAGCATGGATCAGAACACTCCAAATCACGTTCTTCCAGCCCTGGAAAACCCCAGGCTGTTTCTTCATTAAACTCTAGTCATTCCAGGTCTCATGGGAATGAT
 162  S S R K K G Q H G S E H S K S R S S S P G K P Q A V S S L N S S H S R S H G N D
 961  CACCATAGCAAGGAACATCAACGCTCCAAATCACCTCGGGACCCTGATGCAAACTGGGATTCTCCTTCCCGTGTACCTTTTTCAAGTGGGCAGCACTCAACTCAATCTTTCCCACCCTCA
 202  H H S K E H Q R S K S P R D P D A N W D S P S R V P F S S G Q H S T Q S F P P S
1081  TTGATGTCAAAGTCCAATTCAATGTTACAGAAACCCACTGCCTATGTGCGGCCCATGGACGGACAGGAGTCCATGGAACCAAAGCTGTCCTCTGAGCACTACAGCAGCCAATCCCATGGC
 242  L M S K S N S M L Q K P T A Y V R P M D G Q E S M E P K L S S E H Y S S Q S H G
1201  AACAGCATGACTGAGCTGAAGCCCAGCAGCAAAGCACATCTCACCAAGCTGAAAATACCTTCCCAACCACTGGATGCATCAGCTTCTGGTGATGTGAGCTGTGTGGATGAAATCCTAAAA
 282  N S M T E L K P S S K A H L T K L K I P S Q P L D A S A S G D V S C V D E I L K
                                                                                                   ▼
1321  GAGATGACGCATTCATGGCCTCCCCCTCTAACGGCTATTCATACACCATGCAAAACAGAACCTTCCAAATTTCCTTTTCCAACTAAGGAGTCTCAGCAGTCCAATTTTGGCACTGGAGAA
 322  E M T H S W P P P L T A I H T P C K T E P S K F P F P T K E S Q Q S N F G T G E
1441  CAAAAAAGATATAATCCTTCTAAAACTTCAAATGGGCACCAGTCTAAATCTATGTTAAAAGATGACTTAAAACTAAGCAGCAGTGAAGACAGTGATGGGGAACAGGATTGTGATAAGACA
 362  Q K R Y N P S K T S N G H Q S K S M L K D D L K L S S S E D S D G E Q D C D K T
1561  ATGCCGAGGAGTACACCAGGAAGTAACTCTGAACCTTCACACCATAATAGTGAAGGAGCAGATAACTCCAGGGATGATTCTAGTAGCCACAGTGGATCTGAAAGCAGCTCTGGATCTGAC
 402  M P R S T P G S N S E P S H H N S E G A D N S R D D S S S H S G S E S S S G S D
1681  TCAGAGAGTGAAAGTAGTTCCAGTGACAGTGAGGCAAATGAGCCATCCCAGAGTGCATCTCCCGAGCCTGAACCCCCGCCAACAAACAAATGGCAACTTGATAATTGGCTGAATAAAGTG
 442  S E S E S S S S D S E A N E P S Q S A S P E P E P P P T N K W Q L D N W L N K V
1801  AACCCACATAAAGTGTCACCCGCCTCTTCAGTGGACAGTAACATCCCATCATCTCAAGGCTACAAAAAGGAAGGCCGAGAGCAGGGCACTGGGAATAGCTACACTGATACAAGTGGACCT
 482  N P H K V S P A S S V D S N I P S S Q G Y K K E G R E Q G T G N S Y T D T S G P
1921  AAAGAAACGAGTTCCGCTACTCCGGGACGAGACTCCAAAACCATCCAAAAGGGATCAGAAAGTGGGCGTGGGAGGCAGAAATCTCCTGCACAGAGTGACAGCACAACACAGAGAAGAACT
 522  K E T S S A T P G R D S K T I Q K G S E S G R G R Q K S P A Q S D S T T Q R R T
2041  GTAGGCAAAAAACAACCCAAAAAGGCTGAGAAGGCAGCTGCTGAAGAGCCTCGTGGAGGCCTGAAGATAGAAAGTGAAACCCCTGTAGACTTGGCTAGCAGCATGCCCTCCAGCAGACAC
 562  V G K K Q P K K A E K A A A E E P R G G L K I E S E T P V D L A S S M P S S R H
2161  AAAGCAGCCACCAAAGGCTCAAGGAAACCCAATATAAAGAAGGAGTCTAAGTCTTCCCCTCGACCTACAGCAGAGAAAAAGAAATATAAGTCAACAAGTAAATCTTCCCAGAAATCAAGG
 602  K A A T K G S R K P N I K K E S K S S P R P T A E K K K Y K S T S K S S Q K S R
2281  GAAATCATAGAAACAGATACCTCATCCTCAGATTCAGATGAAAGTGAGAGCCTTCCTCCTTCCTCACAAACTCCTAAGTACCCCGAGAGCAATAGGACTCCTGTTAAACCCTCCTCAGTG
 642  E I I E T D T S S S D S D E S E S L P P S S Q T P K Y P E S N R T P V K P S S V
2401  GAGGAAGAAGATAGCTTTTTTCGGCAACGAATGTTCTCTCCTATGGAAGAGAAGGAACTTCTTTCACCCCTCAGTGAGCCTGATGACAGGTACCCACTTATTGTGAAGATTGACCTGAAT
 682  E E E D S F F R Q R M F S P M E E K E L L S P L S E P D D R Y P L I V K I D L N
2521  CTTTTGACTAGAATACCAGGAAAGCCTTACAAAGAAACAGAGCCGCCCAAGGGGGAAAAGAAAAATGTGCCAGAAAAGCACACGAGAGAGGCTCAGAAACAAGCCTCAGAAAAAGTTTCC
 722  L L T R I P G K P Y K E T E P P K G E K K N V P E K H T R E A Q K Q A S E K V S
2641  AACAAAGGCAAGAGGAAGCATAAGAATGAAGATGATAACCGAGCCAGTGAGAGCAAGAAACCCAAAACGGAGGACAAGAATTCAGCAGGCCATAAGCCATCCAGCAACAGAGAGTCATCT
 762  N K G K R K H K N E D D N R A S E S K K P K T E D K N S A G H K P S S N R E S S
2761  AAGCAGAGTGCTGCAAAAGAAAAGGATTTGTTGCCTTCTCCCGCTGGGCCTGTTCCTTCAAAAGATCCAAAAACAGAGCATGGCTCTCGGAAGAGGACTATTAGTCAGTCTTCTTCCTTA
 802  K Q S A A K E K D L L P S P A G P V P S K D P K T E H G S R K R T I S Q S S S L
2881  AAGTCAAGCAGTAACAGCAACAAGGAGACGAGTGGCAGCAGCAAAAACAGTTCCTCCACATCAAAGCAGAAGAAGACCGAAGGGAAGACTTCCAGTAGCTCCAAGGAGGTTAAGGAAAAG
 842  K S S S N S N K E T S G S S K N S S S T S K Q K K T E G K T S S S S K E V K E K
3001  GCTCCAAGTAGCTCCTCTAACTGTCCTCCATCTGCACCAACTCTTGATTCTTCTAAGCCTCGGAGAACAAAGCTTGTCTTTGATGACAGAAATTATTCAGCAGACCATTATTTACAAGAA
 882  A P S S S S N C P P S A P T L D S S K P R R T K L V F D D R N Y S A D H Y L Q E
3121  GCAAAAAAGCTAAAGCACAATGCAGATGCATTGTCTGATAGGTTTGAGAAAGCTGTATACTATCTTGATGCTGTGGTATCTTTCATTGAATGTGGGAATGCATTAGAGAAGAATGCTCAG
 922  A K K L K H N A D A L S D R F E K A V Y Y L D A V V S F I E C G N A L E K N A Q
3241  GAATCCAAATCCCCATTCCCTATGTATTCAGAGACGGTGGATCTCATCAAATACACTATGAAGCTAAAGAATTACTTGGCACCAGATGCTACAGCTGCAGATAAACGACTCACAGTACTT
 962  E S K S P F P M Y S E T V D L I K Y T M K L K N Y L A P D A T A A D K R L T V L
3361  TGCCTGCGATGCGAGTCTTTGCTGTACCTGAGGCTGTTCAAACTGAAGAAGGAAAATGCTCTGAAGTACTCAAAGACACTGACAGAGCACCTGAAGAATTCTTATAATAATTCTCAAGCA
1002  C L R C E S L L Y L R L F K L K K E N A L K Y S K T L T E H L K N S Y N N S Q A
3481  CCATCGCCTGGCTTGGGAAGCAAAGCTGTGGGGATGCCTTCCCCTGTTTCTCCAAAGCTGTCACCAGGCAATTCAGGAAATTATTCATCTGGGGCCAGTAGTGCTTCTGCAAGTGGTTCT
1042  P S P G L G S K A V G M P S P V S P K L S P G N S G N Y S S G A S S A S A S G S
3601  TCAGTGACCATTCCACAGAAGATCCACCAGATGGCAGCCAGCTATGTTCAGGTCACATCCAACTTCCTCTATGCCACCGAAATTTGGGACCAAGCTGAACAGCTTTCCAAAGAGCAAAAA
1082  S V T I P Q K I H Q M A A S Y V Q V T S N F L Y A T E I W D Q A E Q L S K E Q K
3721  GAATTCTTTGCTGAACTGGATAAAGTAATGGGCCCTCTCATCTTTAATGCAAGCATCATGACAGATCTAGTTCGTTATACCCGGCAGGGACTGCACTGGCTTCGCCAGGATGCCAAGTTG
1122  E F F A E L D K V M G P L I F N A S I M T D L V R Y T R Q G L H W L R Q D A K L
3841  ATATCTTGAACTGAACACATTCTCGTTGCCTCTGATTTTCTCCACAACACTGTGTCACATCACGAAGGAAAACTGCCATAACATACCACCTAGTCGACACTAAGAATGAGGAATAGTTTT
1162  I S *
3961  CTCCTCGTTGGTTCATGTGTTGTTGTTTTGATAATCCAAAGCGATCATGTCAGTTGGCCCTTTAATATTTCCAATGTGAAAGATTATTTAAATGCTTTTAAATCTGCAGCACATTGATA
4081  AGATGGTTTCCGTGAGCTATGATAAGATTGAAATTCCAGTTGCAATTCATAACTAATGAATTGAAATTTTATTTATTTAATTTTTTTAAGTTCCCAGATTGGGGCTAGTTTAAGACTTA
4201  GTTATTTCTGCCTATAAATTTACTCTGTTGAATATTTAGCATAAATTTAATGTAG
```

d

```
#11q23      gtaaaggtgttcagtgatcataaagtatattgagtgtcaaagactttaaataaagaaaatgctactaccaaaggtgttgaaagaggaaatcagcac
            ||||||||||||||||||||||||||||||||||||||||||||||
MLL-AF5q31  gtaaaggtgttcagtgatcataaagtatattgagtgtcaaagactttgtaccctttagctgttatccctctttccctctaccttcctagccttaag
                                                          ||||||||||||||||||||||||||||||||||||||||||||||||||||
#5q31       tcagtagttttgagtttatttataaatcatcaactcaaaaagaaactgtaccctttagctgttatccctctttccctctaccttcctagccttaag
```

e

centromere ← → telomere

AF5q31 GDF-9 KIAA0202 KIF3A IL4 IL13 Rad50 IL-5

λH17-6 1308e5 876h9 96c7 50kb

1076B9 565a12 282d12

Fig. 2. (*a*) Partial sequence of the junction of the *MLL-AF5q31* chimeric transcript. The arrow indicates the fusion point. (*b*) Cloning of *AF5q31* cDNA clones. (*c*) Sequencing of the *AF5q31* cDNA. * indicates a termination codon. The arrow indicates the breakpoint in the patient. (*d*) Sequence of the ins(5;11)(q31;q13q23) breakpoint region in the patient. (*e*) Physical map of the chromosome 5q31-cytokine cluster region and P1 clones.

a

```
AF5q31      1      MNREDRNVLRMKERERRNQEIQQGEDAFPPSSPLFAEPYKVTSKEDKLSSRIQSMLGNYDEMKDF
AF4 human   1   MAAQSSLYNDDRNLLRIREKERRNQEAHQEKEAFPEKIPLFGEPYK-TAKGDELSSRIQNMLGNYEEVKEF
Af4 mouse   1   MAAHSSLYNEDRNLLRIREKERRNQEAHQEKEAFPEKAPLFPEPYK-TAKGDELSSRIQTMLGDYEEMKEF
LAF4       31   DLESLCVYEPDRNALRRKERERRNQETQQDDGTFNSSYSLFSEPYK-TNKGDELSNRIQNTLGNYDEMKDF
FMR2       11   DLEQQCHYEQDRSALKKREWERRNQEVQQEDDLFSSGFDLFGEPYK-TNKGDAIANRVQNTLGNYDEMKNL

AF5q31    232   QHSTQSFPP-SLMSKSNSMLQKPTAYVRPMDGQ-----ESMEPKLSSE-----HYSSQSHGNSMTELKPSS
AF4 human 268   QPPSQTFPPPSLPSKSVAMQQKPTAYVRPMDGQDQAPSESPELKPLPE-----DYRQQTFEK--TDLKVPA
Af4 mouse 254   QPPCQTFPPPPLPSKSAAMQQKPTAYVRPMDGQDQAPSESPELKLPLE-----DYGQQSFEK--PDLKVPA
LAF4      228   GHCVQNFPP-SLASKPSLVQQKPTAYVRPMDGQDQAPDESPKLKSSSE--TSVHCTSYRGVPASKPEPARA
FMR2      260   GLSVQNFPP-GLYCKTSMGQQKPTAYVRPMDGQDQAPDISPTLKPSIEFENSFGNLSFGTLLDGKPSAASS

AF5q31    292   KAHLTKLKIPSQPLDASASGDVSCVDEILKEMTHSWPPPLTAIHTPCKTEPSKFPFPTKESQQSNFGTGEQ
AF4 human 332   KAKLTKLKMPSQSVEQTYSNEVHCVEEILKEMTHSWPPPLTAIHTPSTAEPSKFPFPTKDSQHVSSVTQNQ
Af4 mouse 318   KAKLTRLRMPSQSVEQPYSNEVHCVEEILKEMTHSWPPPLTAIHTPSTAEPSRFPFPTKDLLHVSPATQSQ
LAF4      296   KAKLSKFSIPKQGEESRSGETNSCVEEIIREMT--SLPPLSAIQAPGKVEPTKFPFPNKDSQLVSSGHNNP
FMR2      330   KTKLPKFTIL-QTSEVSLPSDPSCVEEILREH----------------------------------SQHLTP

AF5q31    363   KRYN-PSKT-SNGHQSKS-MLKDDLKLSSSEDSDGEQDCDKTMPRSTPGSNSEPSHHNSEGADNSRDDSSS
AF4 human 403   KQYDTSSKTHSNSQQGTSSMLEDDLQLSDSEDSDSEQTPEKPPSSSAPPSAPQ-SLPEPVASAHSSSAESE
Af4 mouse 389   KQYDTPSKTHPNPQQGTSSMLEDDLQLSDSEDSDTEQATEKPPSPPAPPSAPQ-TLPEPVASAHSSSGESE
LAF4      365   KKGDAEPESPDNGTSNTS-MLEDDLKLSSDEEENE---QQAAQ--RTALRALSDSAVVQQPNCRTSVPSSK
FMR2      366   GFTLQKWNDPTTRASTKK--MLEDDLKLSSDEDDLEPVKTLTTQCTATELYQAVEKAKPRNNPVNPPLATPQ

AF5q31    431   HSGSE-----SSSGSDSESESSSSDSEA---NEPSQSAS------PEPEPPPTNKWQLDNWLNKVN
AF4 human 473   STSD------SDSSSDSESESSSSDSEE---NEPLETPA------PEPEPPTTNKWQLDNWLTKVS
Af4 mouse 458   SSE-------SDSSSDSESESSSSDSEEEEENEPLETRA------PEPEPPTTNKWQLDNWLTKVN
LAF4      430   GSSSSSSSGTSSSSSDSESSS-GSDSETESSSSESEGSKPPHFSSPEAEPASSNKWQLDKWLNKVN
FMR2      435   PPPAVQASGGSGSSSESESSS-ESDSDTESSTTDSESNEAPRVATPEPEPPSTNKWQLDKWLNKVT

AF5q31    905   KLVFDDRNYSADHYLQ----EAKKLKHNADALSDRFEKAVYYLDAVVSFIECGNALEKNAQESKSPF-PMY
AF4 human 952   --PGKPQVKFDKQQADLHMREAKKMKQKAELMTDRVGKAFKYLEAVLSFIECGIATESESQSSKSAY-SVY
Af4 mouse 956   AKPGKPQVKSDRQQADFHMKEAKKLKCKAETMVDKAGKAFKYLEAVLSFIECGMASESES-SAKSAY-AVY
LAF4      985   KLVFDDMPRSADYFMQ----EAKRMKHKADAMVEKFGKALNYAEAALSFIECGNAMEQGPMESKSPYYLMY
FMR2     1023   KLTFDDSVHNADYYMQ----EAKKLKHKADALFEKFGKAVNYADAALSFTECGNAMERDPLEAKSPY-TMY

AF5q31    971   SETVDLIKYTMKLKNYLARDATAADKRLTVLCLRCESLLYLRLFKLKKENALKYSKTLTEHLK-NSYNNSQ
AF4 human 1020  SETVDLIKFIMSLKSFSDATAPTQEKIFAVLCMRCQSILNMAMFRCKKDIAIKYSRTLNKHFE-SSSKVAQ
Af4 mouse 1025  SETIDLIRYVMSLKCFSDNTMPAQEKIFAVLCLRCQSLLNMAMFRCKKDTVMKYSRTLSEHFK-STSKVAQ
LAF4     1052   SETVELIRYAMRLKTHSGPNATPEDKQLAALCYRCLALLYWRMFRLKRDHAVKYSKALIDYFK-NSSKAAQ
FMR2     1089   SETVELLRYAMRLKNFASPLASDGDKKLAVLCYRCLSLLYLRMFKLKKDHAMKYSRSLMEYFKQNASKVAQ

AF5q31   1041   APSPGLGS-KAVGMPSPVSPKLSPGNSGNYSSGASSASASGSS---VTIPQKIHQMAASYVQVTSNFLYAT
AF4 human 1090  APSPCIAS---TGTPSPLSPMPSPASSVGSQSSAGS-VGSSGVAATISTPVTIQNMTSSYVTITSHVLTAF
Af4 mouse 1095  APSPCTARS--TGVPSPLSPMPSPASSVGSQSSAGSSMGSVGVTATVSTPVSIQNMTSSYVTITSHVLTAF
LAF4     1122   APSPWGASGKSTGTPSPISPNPFPGSSVGSQGSLSNASALSPSTI-VSIPQRIHQMAANHVSITNSILHSY
FMR2     1160   IPSPWVSNGKNTPSPVSLNNVSPINAMGNCNNGP------------VTIPQRIHHMAASHVNITSNVLRGY

AF5q31   1108   EISDQAEQLSKEQKEFFAELDKVMGPLIFNASIMTDLVRYTRQGLHWLRQDAKLIS
AF4 human 1157  DLWEQAEALTRKNKEFFARLSTNVCTLALNSSL-VDLVHYTRQGFQQLQELTKTP
Af4 mouse 1164  SLWEQAEALTRKNKEFFAQLSTKVRVLALNSSL-VDLVHYTRQGLQRLKQSPKGP
LAF4     1192   DYWEMADNLAKENREFFNDLDLLMGPVTLHSS-MEHLVQYSQQGLHWLRNSAHLS
FMR2     1219   EHSDMADKLTRENKEFFGDLDTLMGPLTQHSS-MTNLVRYVRQGLCWLRIDAHLL
```

b

Fig. 3. (*a*) Comparison of predicted coding sequence of AF5q31 with other related sequences. (*b*) Schematic representation of homology between AF5q31 and AF4 proteins. The percentage of amino acid homology between corresponding regions of AF5q31 and AF4 is indicated. Arrowheads indicate the fusion points with *MLL*.

The AF5q31 protein is homologous to the AF4 protein (Fig. 3 *a* and *b*). The AF5q31 and AF4 proteins have three homologous regions, including the transactivation domain of AF4 (25, 26) (Fig. 3*b*). Although the functions of the other two regions

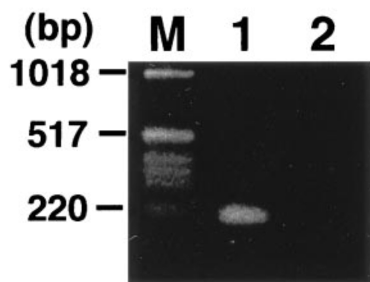

Fig. 4. Detection of the *MLL-AF5q31* chimeric transcripts by reverse transcription–PCR. Primers used were MLL-7S and 5–1A (lane 1), and 5–13S and MLL-11A (lane 2), respectively. M, size marker.

remain unknown, these homologous regions may include potential common functional domains of AF5q31 and AF4. The *AF4* gene was cloned from ALL with t(4;11)(q21;q23) (3). Patients with the t(4;11) are characterized by very young age, hyperleukocytosis, early pre-B phenotype ($CD10^{-}/CD19^{+}$), and poor treatment outcome for infants and patients aged >10 years (27). The t(4;11) is the predominant 11q23 chromosome translocation in infant ALL (7, 8). As compared with other *MLL* chimeric transcripts in ALL and AML, *MLL-AF4* chimeric transcripts in infant ALL have been reported to be associated with extremely poor prognosis, regardless of advances in therapy for childhood leukemia (11). The clinical features of this patient with *MLL-AF5q31* fusion transcript, characterized by early pre-B phenotype ($CD10^{-}/CD19^{+}$) and poor outcome, were similar to those of the patients having *MLL-AF4* chimeric transcripts. t(5;11)(q31;q23) has been described in only a few patients with *de novo* ALL (28, 29) and therapy-related AML (30). Although it remains unknown whether MLL-AF5q31 is associated with only ALL or not, these findings suggest that AF5q31 and AF4 might define a new family particularly involved in the pathogenesis of 11q23-associated-ALL. Further accumulation of these patients with this *MLL-AF5q31* transcript may clarify the association between t(5;11)-ALL and t(4;11)-ALL.

Hitherto, two *AF4*-related genes had been identified, *LAF4* (25) and *FMR2* (31, 32), and sequencing analysis revealed that both LAF4 and FMR2 are also homologous to AF5q31 (Fig. 3*a*). LAF4, isolated as a lymphoid-restricted nuclear protein from a subtracted cDNA library, was expressed at high levels in lymphoid tissues and at lower levels in brain and lung (25). In human and mouse lymphoid cell lines, LAF4 expression was highest in pre-B cells, intermediate in mature B cells, and absent in plasma

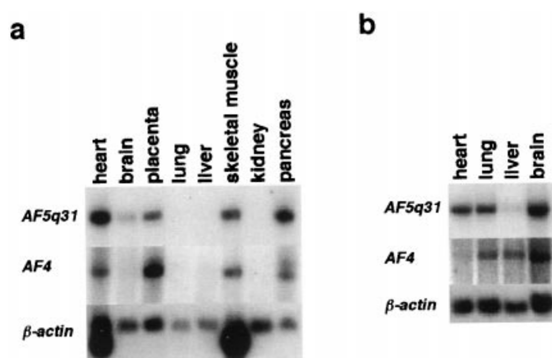

Fig. 5. Northern blot analysis of RNAs from adult (*a*) and fetal (*b*) human tissues. U*AF5q31* and *AF4* cDNA fragments were used as probes for the Northern blots in the upper and middle figures, respectively. Membranes were rehybridized to the β-actin probe for the lower figures. The organs from which tissues were analyzed are indicated on top of each lane.

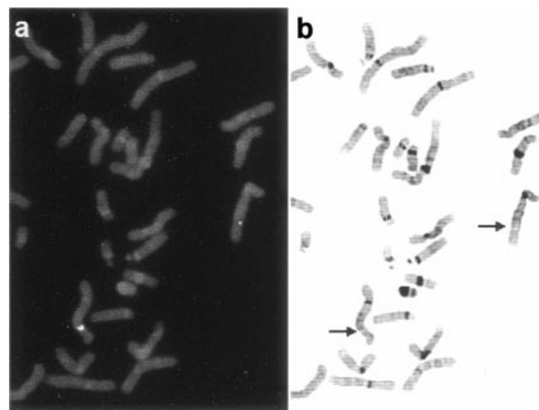

Fig. 6. Chromosomal localization of the *AF5q31* gene. (*a*) A fluorescence *in situ* hybridization pattern obtained with the phage clone λH17–6 as a probe. (*b*) The same metaphase spread stained with 4′,6-diaminido-2-phenylindole dihydrochloride (DAPI). A DAPI image was inverted and enhanced in terms of band image contrast. The *AF5q31* gene was assigned to band 5q31.1 as indicated by arrows.

cells, thus suggesting that LAF4 plays a potential regulatory role in early lymphoid development. *FMR2* originally was identified from chromosome Xq28 as a gene associated with FRAXE mental retardation (31, 32). *FMR2* encodes a nuclear protein of 1,311 aa with putative nuclear transcription transactivation potential. Expression of the *FMR2* gene was found in adult brain and placenta, and fetal brain, lung, and kidney (33). The expression pattern of the *FMR2* gene in fetal tissues was similar to that of the *AF5q31* gene, suggesting that both AF5q31 and FMR2 may play common roles in fetal development. Particularly, expression of the *AF5q31* gene was decreased in adult lung, liver, and brain compared with that in fetal tissues, suggesting that AF5q31 may play critical roles in the fetal development of these tissues.

The 5q31 chromosomal region is known to be a very critical region in which many cytokine genes are clustered (34, 35). This region also has been suggested to be associated with interstitial deletion of 5q observed in AML and myelodysplastic syndrome (MDS), suggesting that a tumor suppressor gene for AML and MDS may be present in this region (34, 36). Therefore, this region has been under investigation for a long time by many researchers, and large amounts of sequence data from the region have been stored in databases. Interestingly, it was predicted that another *AF4*-related gene was present in the chromosome 5q31 region based on a homology search of databases by Frestedt *et al.* (37). Gecz *et al.* (33) showed that *FMR2* is homologous to two expressed sequence tags (W26686 and AA025630) mapped to chromosome 5q31. *AF5q31* is likely to be the gene predicted by those authors. However, the location of the *AF5q31* gene is outside of the commonly deleted region previously reported in AML and MDS (36).

A few papers about the function of MLL fusion proteins have been published (19, 20, 38–41). However, the function of the MLL-AF4 fusion protein has not been analyzed yet. Functional analyses of the AF5q31 and MLL-AF5q31 fusion proteins may provide new insights into the function of the MLL-AF4 fusion protein and also the leukemogenesis of 11q23-associated-ALL.

We thank M. Seto, Aichi Cancer Center Research Institute, for providing the *MLL* cDNA probe (probe x). We express appreciation to S. Sohma and H. Soga for technical assistance. This work was supported by a Grant-in-Aid for Cancer Research from the Ministry of Health and Welfare of Japan, a Grant-in-Aid for Scientific Research on Priority Areas, and Grant-in-Aid for Scientific Research (B) and (C) from the Ministry of Education, Science, Sports, and Culture of Japan.

1. McCabe, N. R., Burnett, R. C., Gill, H. J., Thirman, M. J., Mbangkollo, D., Kipiniak, M., van Melle, E., Ziemin-van der Poel, S., Rowley, J. D. & Diaz, M. O. (1992) *Proc. Natl. Acad. Sci. USA* **89,** 11794–11798.
2. Tkachuk, D. C., Kohler, S. & Cleary, M. L. (1992) *Cell* **71,** 691–700.
3. Gu, Y., Nakamura, H., Alder, H., Prasad, R., Canaani, O., Cimino, G., Croce, C. M. & Canaani, E. (1992) *Cell* **71,** 701–708.
4. Djabali, M., Selleri, L., Parry, P., Bower, M., Young, B. D. & Evans, G. A. (1992) *Nat. Genet.* **2,** 113–118.
5. Yu, B. D., Hanson, R. D., Hess, J. L., Horning, S. E. & Korsmeyer, S. J. (1998) *Proc. Natl. Acad. Sci. USA* **95,** 10632–10636.
6. Rowley, J. D. (1998) *Annu. Rev. Genet.* **32,** 495–519.
7. Rubnitz, J. E., Link, M. P., Shuster, J. J., Carroll, A. J., Hakami, N., Frankel, L. S., Pullen, D. J. & Cleary, M. L. (1994) *Blood* **84,** 570–573.
8. Taki, T., Ida, K., Bessho, F., Hanada, R., Kikuchi, A., Yamamoto, K., Sako, M., Tsuchida, M., Seto, M., Ueda, R. & Hayashi, Y. (1996) *Leukemia* **10,** 1303–1307.
9. Super, H. J., McCabe, N. R., Thirman, M. J., Le Beau, M. M., Pedersen-Bjergaard, J., Philip, P., Diaz, M. O. & Rowley, J. D. (1993) *Blood* **82,** 3705–3711.
10. Hunger, S. P., Tkachuk, D. C., Amylon, M. D., Link, M. P., Carroll, A. J., Welborn, J. L., Willman, C. L. & Cleary, M. L. (1993) *Blood* **81,** 3197–3203.
11. Heerema, N. A., Sather, H. N., Ge, J., Arthur, D. C., Hilden, J. M., Trigg, M. E. & Reaman, G. H. (1999) *Leukemia* **13,** 679–686.
12. Rubnitz, J. E., Behm F. G. & Downing, J. R. (1996) *Leukemia* **10,** 74–82.
13. Yamamoto, T., Harada, N., Kawano, Y., Taya, S. & Kaibuchi, K. (1999) *Biochem. Biophys. Res. Commun.* **259,** 103–107.
14. Shilatifard, A., Lane, W. S., Jackson, K. W., Conaway, R. C. & Conaway, J. W. (1996) *Science* **271,** 1873–1876.
15. Taki, T., Sako, M., Tsuchida, M. & Hayashi, Y. (1997) *Blood* **89,** 3945–3950.
16. Sobulo, O. M., Borrow, J., Tomek, R., Reshmi, S., Harden, A., Schlegelberger, B., Housman, D., Doggett, N. A., Rowley, J. D. & Zeleznik-Le, N. J. (1997) *Proc. Natl. Acad. Sci. USA* **94,** 8732–8737.
17. Ida, K., Kitabayashi, I., Taki, T., Taniwaki, M., Noro, K., Yamamoto, M., Ohki, M. & Hayashi, Y. (1997) *Blood* **90,** 4699–4704.
18. Taki, T., Shibuya, N., Taniwaki, M., Hanada, R., Morishita, K., Bessho, F., Yanagisawa, M. & Hayashi, Y. (1998) *Blood* **92,** 1125–1130.
19. Corral, J., Lavenir, I., Impey, H., Warren, A. J., Foster, A., Larson, T. A., Bell, S., McKerzie, A. N. J., King, G. & Rabbitts, T. H. (1996) *Cell* **85,** 853–861.
20. Lavau, C., Szilvassy, S. J., Slany, R. & Cleary, M. L. (1997) *EMBO J.* **16,** 4226–4237.
21. Yamamoto, K., Seto, M., Komatsu, H., Iida, S., Akao, Y., Kojima, S., Kodera, Y., Nakazawa, S., Ariyoshi, Y., Takahashi, T. & Ueda, R. (1993) *Oncogene* **8,** 2617–2625.
22. Ida, K., Taki, T., Bessho, F., Kobayashi, M., Taira, F., Hanada, R., Yamamoto, K., Okimoto, Y., Seto, M., Ueda, R. & Hayashi, Y. (1997) *Med. Pediatr. Oncol.* **28,** 325–332.
23. Taniwaki, M., Nishida, K., Ueda, Y., Misawa, S., Nagai, M., Tagawa, S., Yamagami, T., Sugiyama, H., Abe, M., Fukuhara, S. & Kashima, K. (1995) *Blood* **85,** 3223–3228.
24. Nilson, I., Lochner, K., Siegler, G., Greil, J., Beck, J. D., Fey, G. H. & Marschalek, R. (1996) *Br. J. Haematol.* **93,** 966–972.
25. Ma, C. & Staudt, L. M. (1996) *Blood* **87,** 734–745.
26. Li, Q., Frestedt, J. L. & Kersey, J. H. (1998) *Blood* **92,** 3841–3847.
27. Pui, C. H., Frankel, L. S., Carroll, A. J., Raimondi, S. C., Shuster, J. J., Head, D. R., Crist, W. M., Land, V. J., Pullen, D. J., Steuber, C. P., *et al*. (1991) *Blood* **77,** 440–447.
28. Kearney, L., Bower, M., Gibbons, B., Das, S., Chaplin, T., Nacheva, E., Chessells, J. M., Reeves, B., Riley, J. H., Lister, T. A. & Young, B. D. (1992) *Blood* **80,** 1659–1665.
29. Rubnitz, J. E., Link, M. P., Shuster, J. J., Carroll, A. J., Hakami, N., Frankel, L. S., Pullen, D. J. & Cleary, M. L. (1994) *Blood* **84,** 570–573.
30. Hoyle, C. F., de Bastos, M., Wheatley, K., Sherrington, P. D., Fischer, P. J., Rees, J. K., Gray, R. & Hayhoe, F. G. (1989) *Br. J. Haematol.* **72,** 45–53.
31. Gecz, J., Gedeon, A. K., Sutherland, G. R. & Mulley, J. C. (1996) *Nat. Genet.* **13,** 105–108.
32. Gu, Y., Shen, Y., Gibbs, R. A. & Nelson, D. L. (1996) *Nat. Genet.* **13,** 109–113.
33. Gecz, J., Bielby, S., Sutherland, G. R. & Mulley, J. C. (1997) *Genomics* **44,** 201–213.
34. Le Beau, M. M., Lemons, R. S., Espinosa, R., 3rd, Larson, R. A., Arai, N. & Rowley, J. D. (1989) *Blood* **73,** 647–650.
35. Marsh, D. G., Neely, J. D., Breazeale, D. R., Ghosh, B., Freidhoff, L. R., Ehrlich-Kautzky, E., Schou, C., Krishnaswamy, G. & Beaty, T. H. (1994) *Science* **264,** 1152–1156.
36. Fairman, J., Chumakov, I., Chinault, A. C., Nowell, P. C. & Nagarajan, L. (1995) *Proc. Natl. Acad. Sci. USA* **92,** 7406–7410.
37. Frestedt, J. L., Hilden, J. M. & Kersey, J. H. (1996) *DNA Cell Biol.* **15,** 669–678.
38. Slany, R. K., Lavau, C. & Cleary, M. L. (1998) *Mol. Cell. Biol.* **18,** 122–129.
39. Joh, T., Hosokawa, Y., Suzuki, R., Takahashi, T. & Seto, M. (1999) *Oncogene* **18,** 1125–1130.
40. Maki, K., Mitani, K., Yamagata, T., Kurokawa, M., Kanda, Y., Yazaki, Y. & Hirai, H. (1999) *Blood* **93,** 3216–3224.
41. Dobson, C. L., Warren, A. J., Pannell, R., Forster, A., Lavenir, I., Corral, J., Smith, A. J. & Rabbitts, T. H. (1999) *EMBO J.* **18,** 3564–3574.