Oncogene (2000) 19, 4446–4450

www.nature.com/onc

SHORT REPORT

AF15q14, a novel partner gene fused to the *MLL* gene in an acute myeloid leukaemia with a t(11;15)(q23;q14)

S Hayette[1,2], I Tigaud[1], A Vanier[1], S Martel[2], L Corbo[2], C Charrin[1], E Beillard[3], G Deleage[4], JP Magaud[1,2] and R Rimokh*[,2]

[1]*Laboratoire de Cytogénétique Moléculaire, Hôpital Edouard Herriot, 69437 Lyon, France;* [2]*Unité INSERM U453, Centre Léon Bérard, 69373 Lyon cedex 08, France;* [3]*Laboratoire de Biochimie et Biologie Moléculaire, Faculté de Médecine Nord, 13916 Marseille Cedex 20, France;* [4]*UMR 5086, Pôle Rhône-Alpin de BioInformatique, IBCP, 69367 Lyon cedex 07, France*

In haematopoietic malignancies the *MLL* gene, located on chromosome 11q23, is frequently disrupted by chromosome rearrangement, generally resulting in fusion to various partner genes. We have previously reported a t(11;15)(q23;q14) in a case of acute myeloblastic leukaemia. Here, we report the cloning of a novel *MLL* partner, *AF15q14*, at chromosome 15q14. In this translocation, the breakpoint occurred in exon 8 of *MLL* and exon 10 of *AF15q14*. The normal *AF15q14* transcripts of approximately 8.5 kb in size, are expressed in different tumoral cell lines, in a variety of normal tissues, and in all the foetal tissues tested. Sequencing of *AF15q14* cDNA revealed a putative open reading frame of 1833 amino acids that had no homology with any other known protein. The C-terminal end of the putative AF15q14 contained a bipartite nuclear localization site. The translocation t(11;15) preserved the open reading frame between *MLL* and the 3′ end of *AF15q14*. The contribution of AF15q14 to the fusion protein was only 85 amino acids. Immunofluorescence staining experiments with expression vectors encoding these 85 amino acids confirmed the functionality of the predicted nuclear localization site. *Oncogene* (2000) **19,** 4446–4450.

Keywords: t(11;15)(q23;q14); 11q23; MLL; AF15q14; leukaemogenesis

The *MLL* gene (also known as *ALL1* and *HTRX*), located on chromosome 11q23, is found rearranged in approximately 10% of acute lymphoblastic leukaemias, and in more than 5% of acute myeloid leukaemias (AML), and is generally associated with a very poor prognosis. This gene encodes a large protein (3968 amino acids) which contains several features of transcription factors, AT-hooks, methyl-transferase domain, zinc fingers, a nuclear localization site and a highly conserved SET domain. The *MLL* gene could be the human homologue of *Drosophila trithorax* (reviewed in Dimartino and Cleary, 1999). Defects in yolk-sac haematopoiesis in *MLL*-null mouse embryos (Yu *et al*., 1995), and the differentiation block in *MLL*-null embryonic stem cells (Fidanza *et al*., 1996), suggest that *MLL* plays an important role in haematopoiesis through its ability to regulate the expression of *Hox* genes. Most rearrangements that disrupt the *MLL* gene, e.g. reciprocal translocation and self-fusion, occur within a clustered region of 8.3 kb (Gu *et al*., 1994). Translocation disrupts the *MLL* gene between the AT-hooks and the zinc finger domains to produce a fused gene encoding a chimeric protein. Molecular and cytogenetic evidences indicate that the crucial chimeric product is the 5′*MLL*3′-partner fusion transcript in *MLL* associated leukaemias (Rowley, 1992). To date, the cloning of the fusion partners has led to the discovery of 20 different genes encoding various kinds of proteins with no common features. The apparently-insignificant role of the partner genes in leukaemogenesis is suggested by the tandem duplication of the *MLL* gene, whose effects are similar to those of translocations, and by a recent study which showed that truncated *MLL*, when fused to the bacterial *Lac Z* gene, is sufficient for tumorigenesis (Dobson *et al*., 2000). However, the contribution of the partner gene to leukaemogenesis could still be critical, in view of a mouse knock-in study which has shown that *MLL-AF9* chimeras develop AML, whereas *MLL-myc* chimeras do not (Corral *et al*., 1996). Moreover, the immortalization of murine haematopoietic progenitors by the retrovirus-mediated gene transfer of *MLL-ENL* requires the integrity of the transactivation activity of *ENL* (Schreiner *et al*., 1999). Thus it would seem that the other partners of *MLL* have to be identified before one can attempt to find out whether it is the functional gain or the production of a dominant-negative protein that dictates the tumoral phenotype.

In the present work, we report the cloning of *AF15q14*, the fusion partner of the *MLL* gene in t(11;15)(q23;q14). We previously reported the case of a 48-year-old man with no previous history of toxic exposure, suffering from AML-M4, which was cytogenetically characterized as 46,XY,-3, t(11;15)(q23;q14), +mar (Archimbaud *et al*., 1998; Harrison *et al*., 1998). The bone marrow was hypercellular, with 80% blast cells. The patient was treated by intensive chemotherapy, and died four months after diagnosis. *MLL* involvement in the reciprocal translocation t(11;15)(q23;q14) was demonstrated by Southern blot analysis of the leukaemia-cell DNA. Rearranged DNA fragments of the der(11) and der(15) chromosomes were observed with MLLa probe encompassing the breakpoint cluster region (Figure 1a). Restriction mapping by Southern blot analysis, performed with 5′MLL and 3′MLL probes (data not shown), allowed us to locate the breakpoint between

*Correspondence: R Rimokh
Received 11 April 2000; revised 29 June 2000; accepted 6 July 2000

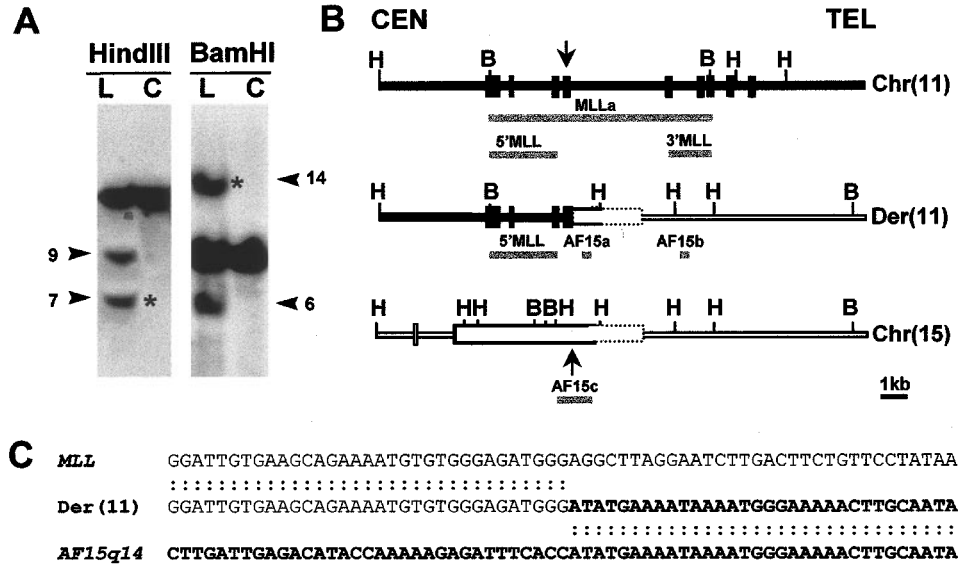


Figure 1 The t(11;15)(q23;q14) translocation involves the *MLL* and *AF15q14* genes. (**a**) Southern blot analysis performed with 10 μg of genomic DNA extracted from leukaemic cells with the t(11;15) translocation (L) and human placenta (C), which were digested with the indicated enzymes. The filter was assayed for hybridization to the *MLL* cDNA probe (MLLa), nucleotides 3750–4614 (Tkachuk *et al*., 1992). Numbers indicate the size (kbp). The rearranged fragments are indicated by arrows, and the der(11) fragments by stars. (**b**) In order to characterize the genomic sequences fused to *MLL*, two genomic libraries in the EMBL3 λ phage vector (Stratagene) were constructed with leukaemic cells and human placental DNA. The leukaemic library was screened with the 5′MLL and 3′MLL probes, which surround the *MLL* breakpoint. These probes were PCR-amplified fragments of nucleotides, 3750–4250 and 4398–4615 respectively. After restriction mapping, recombinant phages containing inserts representative of the der(11) chromosome were isolated. Repeat-free AF15a (198 bp) and AF15b (230 bp) probes derived from the der(11) chromosome were used to screen the placental genomic DNA library, which allowed us to clone recombinant phages corresponding to the normal chromosome 15 counterpart. Thin black lines: chromosome 11 DNA region; thin white lines: chromosome 15 DNA region. Exons are indicated by boxes (unknown 3′end of *AF15q14* exon is indicated by dotted line). The probes are indicated by shaded lines. The breakpoints between the two chromosomes are shown by arrows. B, *Bam*HI; H, *Hind*III. (**c**) Subcloning into Bluescript SK(−) vector (Stratagene), and an analysis of the sequence of the junction in the der(11) chromosome, showed that the breakpoint is in exon 8 of the *MLL* gene (Gu *et al*., 1992). Chromosome 11 and 15 sequences are in normal and bold type respectively

intron 7 and intron 8 within the common breakpoint cluster region of *MLL* (Figure 1b).

The t(11;15) translocation has already been described several times (Hernandez *et al*., 1995; Hunger *et al*., 1993; Rubnitz *et al*., 1994), and it was the detection of a t(11;15) in a patient with AML, involving the *MLL* locus, that prompted us to study the resulting molecular rearrangement. Unfortunately, no mRNA from this patient was available so we used genomic phage libraries prepared from leukaemic cells and placental tissue to isolate the der(11) chromosome and the normal counterpart on chromosome 15 (Figure 1b). Southern blot analysis of the leukaemic cells and control DNA, performed with AF15a, AF15b and AF15c probes, revealed the expected germinal and rearranged bands (data not shown). Hybridization onto DNA from interspecies somatic hybrid (Clontech) showed that these probes were derived from chromosome 15 (data not shown). Fluorescence *in situ* hybridization to normal human metaphases, using an artificial bacterial chromosome isolated by PCR screening with *AF15q14* primers (BAC No. B590H1, CEPH, Paris, France), confirmed the location of these probes on chromosome 15 at band 15q14 (data not shown). Nucleotide sequence analysis revealed that the breakpoint on the der(11) chromosome was located in exon 8 of *MLL* (Figure 1c). Unlike other *MLL* rearrangements previously studied, no extra nucleotides or heptamer-like signals were observed at the junction point, which probably rules out any involvement of VDJ recombinase. It has been shown that large numbers of ALU sequences are present at the *MLL* breakpoint (Marschalek *et al*., 1997), which may render this region particularly prone to recombination events. An ALU sequence was found on chromosome 15 660 bp downstream from the breakpoint, which could thus be predisposed to translocation events.

In order to identify a transcriptional unit on 15q14, repeat-free genomic DNA fragments were used as probes against Northern blot containing poly(A)$^+$ RNA from a variety of cell types. The AF15c probe hybridized to a transcript of about 8.5 kb which is relatively abundant in HeLa cells (Figure 2a). In adult tissue, *AF15q14* mRNA was found to be almost wholly restricted to the thymus, testis and bone marrow, whereas it seemed to be ubiquitous in the foetal tissues tested (Figure 2b). In line with the nomenclature for *MLL* fusion partner genes, we designated this gene *AF15q14* (*ALL-1* fused gene from chromosome 15q14). In order to characterize the *AF15q14* transcript, a HeLa cDNA library was screened with the AF15c probe. Among the nine cDNAs isolated, the longest ones (1.2 kbp, 0.8 kbp and 0.6 kbp) were sequenced. Most of the 5′ and 3′ *AF15q14* sequences were found by PCR screening of a testis cDNA library and/or by the rapid amplification of 5′/3′cDNA ends obtained from HeLa-cell poly(A)$^+$ RNA. Sequencing of the genomic counterpart confirmed the nucleotide sequences of the PCR-isolated fragments. Subsequent walking and compilation of sequence data from the different overlapping cDNA clones allowed us to sequence 5925 bp of the *AF15q14* transcript. The

AF15q14 sequence has been deposited at the Data Library of the National Center for Biotechnology Information (accession No. AF248041). Genomic sequencing showed that the *AF15q14* gene spanned a region of more than 35 kb, and contained at least 10 exons (Figure 3a), with typical donor and acceptor sites at the intron-exon junctions (data not shown).

Analysis of *AF15q14* cDNA showed an open reading frame of 5499 bp encoding a protein of 1833 amino acids, with a predicted molecular mass of 206 kDa. This open reading frame contained an initiation sequence consensus ATGG (Kozak, 1986) at nucleotide 146 in exon 2, and a stop codon at nucleotide 5645 in exon 10. We observed an imbalance in its nucleotide composition, with 63% of A/T. Sequence comparison of the *AF15q14* cDNA and the deduced protein with those found in databases did not reveal any significant homologies. Computerized analysis of the AF15q14 revealed two closely-related motifs of 105 and 103 amino acids coded by exon 10 (Figure 3b). These

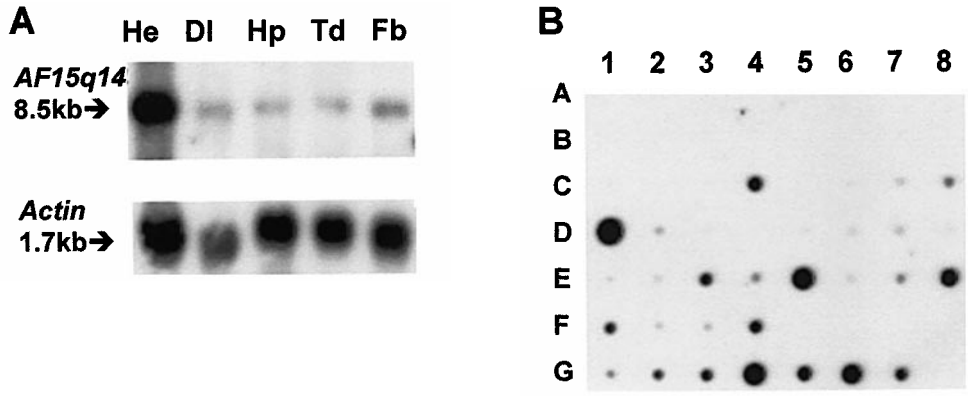
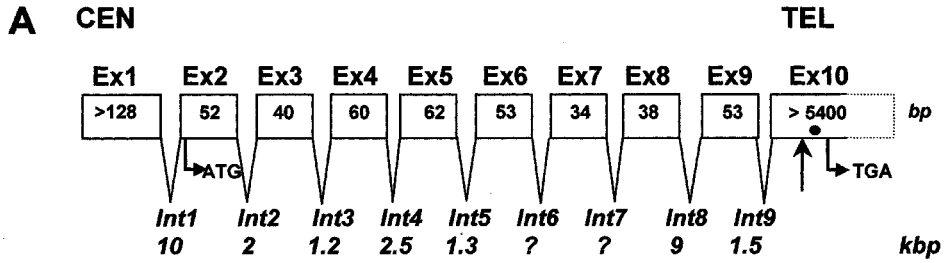


Figure 2 Expression analysis of *AF15q14*. (**a**) Northern blot analysis performed with 5 μg of poly(A)$^+$ RNA extracted from different cell types with a Fast Track mRNA isolation Kit (Invitrogen), and hybridized with the AF15c probe (Figure 1b). In order to assess the integrity of mRNA, this Northern blot was then hybridized with a β-actin probe. *AF15q14* transcripts were about 8.5 kb. Human tumoral cell lines: He, HeLa (cervix); Dl, Del (lymphoid); Hp, HepG2 (liver); Td, T47D (breast) and human fibroblasts (Fb). (**b**) Commercial dot blot (Master blot, Clontech) was hybridized to the AF15c probe. *AF15q14* was found to be highly expressed in the adult testis (D1), thymus (E5) and bone marrow (E8), as well as in the foetal liver (G4) and thymus (G6). It was relatively strongly expressed in the adult colon (C4), small intestine (E3), appendix (F1), placenta (F4), and in all the remaining foetal tissues (G1 to G7)



B

```
859  DKTIVFSEDDKNDMDITKSYTIEINHRPLLEKRDCHLVPLAGTSETILYTCG  910
     ********..*******.*..**..  **  *  .***.*.****.**
1073 DKTIVFSENHKNDMDITQSCMVEIDNESALE--DKEDFHLAGASKTILYSCG 1122

911  QDDMEITRSHTTALECKTVSPDEITTRPMDKTVVFVDNHVELEMTESHTVFID 963
     ******************. *.**..*******.*.**. .**.*.*******
1123 QDDMEITRSHTTALECKTLLPNEIAIRPMDKTVLFTDNYSDLEVTDSHTVFID 1175
```

Figure 3 Schematic representation of the *AF15q14* locus and the repeat motif encoded by exon 10 of *AF15q14*. (**a**) Human *AF15q14* cDNA was analysed by a combination of rapid amplification of 5′/3′cDNA ends from HeLa cells poly(A)$^+$ RNA (5′/3′ RACE kit, Roche), from testis cDNA (Marathon-Ready® cDNA, Clontech) and by the screening of a human HeLa cDNA library (Clontech) with the AF15c probe. A control with a tube containing RNA, but without reverse transcriptase, was performed in order to rule out possible genomic amplification. PCR products were subcloned into the TA cloning vector (InVitrogen). Genomic DNA fragments from normal chromosome 15 were obtained either by placenta-library screening and analysis of the phages subcloned into Bluescript SK(−) vector (Stratagene) or by direct sequencing of PCR-amplified genomic DNA fragments (Expand® long Template PCR System, Roche). Sequence analysis of cDNA from both strands and their genomic counterparts was performed by the dideoxynucleotide method. Comparison of the human cDNA and genomic DNA *AF15q14* sequences showed that the *AF15q14* gene extended for more than 35 kb, and contained at least 10 exons. Analysis of overlapping cDNA clones allowed us to sequence 5925 bp of the *AF15q14* transcript (GenBank accession No. AF248041). The breakpoint on the *AF15q14* gene is indicated by a vertical arrow. The sequence encoding the NLS-BP is shown by a black dot. Ex: exon (unknown 3′end of *AF15q14* exon 10 is indicated by dotted line); Int: intron; CEN: centromere; TEL: telomere. (**b**) Analysis of the predictive AF15q14 protein (1833 amino acids) with the Antheprot program (Geourjon and Deleage, 1995) revealed a repeat motif of about 105 amino acids between residues 859 and 1175. These motifs, though sharing no particular specificity, had about 73% of amino-acids identity

motifs, which are probably due to a functionally-important duplicated region, do not themselves have any currently known functional features. A database search for functional domains revealed only the presence of a predicted Bipartite Nuclear Localization Site (NLS-BP) at the C-terminal end of the putative AF15q14 protein (amino acids 1761–1778).

As with all the other 11q23 translocations so far investigated, the t(11;15)(q23;q14) breakpoint resulted in frame fusion between exon 8 (nt 4253) of *MLL* and exon 10 (nt 5389) of *AF15q14* cDNA. The fused transcript *MLL-AF15q14* could encode a putative protein of 1503 amino acids, composed of the N-terminal part of MLL (1418 amino acids) and the C-terminal part of AF15q14 (85 amino acids), which includes the NLS-BP. As no mRNA from leukaemic cells was available to us, this chimeric transcript remains to be studied in other t(11;15) translocations. In an attempt to assess the functionality of the NLS-BP, expression vectors encoding the C-terminal part of AF15q14 (85 amino acids) were transfected into COS-7 cells. Nuclear immunofluorescence was observed in COS-7 cells transfected with the pSG5flag*AF15*-85aa expression vector (Figure 4). The wild-type pSG5-flag vector, used as a control, did not show any signal (data not shown). Moreover an expression vector encoding these 85 amino acids, fused to Green Fluorescent Protein (pGFP expression vector, Clontech) and transfected into COS-7 cells, showed a similar nuclear localization pattern (data not shown). Our results indicate that AF15q14 protein may be a nuclear protein. Similarly to other MLL-partners such as AF3p21, AF4, AF5α, AF9, ENL and ELL, the NLS of AF15q14 is fused to MLL. All the MLL fusion proteins studied so far have been found to localize in the nucleus and although some partners of MLL such as AF1p/eps15 and AF6 localize in the cytoplasm, they show a nuclear localization when they are fused to MLL (Rogaia *et al*., 1997; Joh *et al*., 1997).

Further functional analysis of translocation partner genes, as well as *MLL*, should provide new insights into the leukaemogenesis of 11q23-associated leukaemia. Molecular analysis of t(11;15)(q23;q14), of which only a few cases have so far been reported, allowed us to isolate a new gene, *AF15q14*, fused to *MLL*. It will be interesting to see if this gene is found to be involved in other cases of t(11;15)(q23;q14).

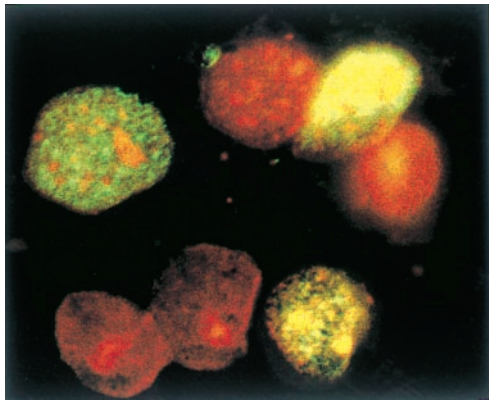


Figure 4 Nuclear localization of the C-terminal portion of the AF15q14 protein. Immunofluorescence staining of Cos-7 cells was carried out after transient transfection, using Fugene™ (Roche), either by the wild-type pSG5-flag vector (Rouault *et al*., 1998), used as a control (data not shown), or by the pSG5-flag expression vector which encodes the last 85 amino acids of AF15q14. The anti-Flag M2 monoclonal antibody (Stratagene) was used to detect the flagAF15-85aa fusion protein (green). Nuclear staining (red) was performed using propidium iodide

Acknowledgements
The authors wish to thank Bénédicte Santa-Lucia and Annick Bertholin for their technical assistance. This work was supported by grants from INSERM, the Association pour la Recherche contre le Cancer and the Ligue Nationale contre le Cancer (Comités du Rhône et de la Saône et Loire).

References

Archimbaud E, Charrin C, Magaud JP, Campos L, Thomas X, Fiere D and Rimokh R. (1998). *Leukemia,* **12,** 25–33.
Corral J, Lavenir I, Impey H, Warren AJ, Forster A, Larson TA, Bel S, McKenzie AN, King G and Rabbitts TH. (1996). *Cell,* **85,** 853–861.
Dimartino JF and Cleary ML. (1999). *Br. J. Haematol.,* **106,** 614–626.
Dobson CL, Warren AJ, Pannell R, Forster A and Rabbitts TH. (2000). *EMBO J.,* **19,** 843–851.
Fidanza V, Melotti P, Yano T, Nakamura T, Bradley A, Canaani E, Calabretta B and Croce CM. (1996). *Cancer Res.,* **56,** 1179–1183.
Geourjon C and Deleage G. (1995). *J. Mol. Graph.,* **13,** 209–212, 199–200.
Gu Y, Alder H, Nakamura T, Schichman SA, Prasad R, Canaani O, Saito H, Croce CM and Canaani E. (1994). *Cancer Res.,* **54,** 2326–2330.
Gu Y, Nakamura T, Alder H, Prasad R, Canaani O, Cimino G, Croce CM and Canaani E. (1992). *Cell,* **71,** 701–708.
Harrison CJ, Cuneo A, Clark R, Johansson B, Lafage-Pochitaloff M, Mugneret F, Moorman AV and Secker-Walker LM. (1998). *Leukemia,* **12,** 811–822.
Hernandez JM, Mecucci C, Beverloo HB, Selleri L, Wlodarska I, Stul M, Michaux L, Verhoef G, Van Orshoven A, Cassiman JJ, Evans GA, Hagemeijer A and Van Den Berghe H. (1995). *Leukemia,* **9,** 1162–1166.
Hunger SP, Tkachuk DC, Amylon MD, Link MP, Carroll AJ, Welborn JL, Willman CL and Cleary ML. (1993). *Blood,* **81,** 3197–3203.
Joh T, Yamamoto K, Kagami Y, Kakuda H, Sato T, Yamamoto T, Takahashi T, Ueda R, Kaibuchi K and Seto M. (1997). *Oncogene,* **15,** 1681–1687.
Kozak M. (1986). *Cell,* **44,** 283–292.
Marschalek R, Nilson I, Lochner K, Greim R, Siegler G, Greil J, Beck JD and Fey GH. (1997). *Leuk. Lymphoma.,* **27,** 417–428.
Rogaia D, Grignani F, Carbone R, Riganelli D, LoCoco F, Nakamura T, Croce CM, Di Fiore PP and Pelicci PG. (1997). *Cancer Res.,* **57,** 799–802.
Rouault JP, Prevot D, Berthet C, Birot AM, Billaud M, Magaud JP and Corbo L. (1998). *J. Biol. Chem.,* **273,** 22563–22569.
Rowley JD. (1992). *Genes Chromosomes Cancer,* **5,** 264–266.

Rubnitz JE, Link MP, Shuster JJ, Carroll AJ, Hakami N, Frankel LS, Pullen DJ and Cleary ML. (1994). *Blood,* **84,** 570 – 573.

Schreiner SA, Garcia-Cuellar MP, Fey GH and Slany RK. (1999). *Leukemia,* **13,** 1525 – 1533.

Tkachuk DC, Kohler S and Cleary ML. (1992). *Cell,* **71,** 691 – 700.

Yu BD, Hess JL, Horning SE, Brown GA and Korsmeyer SJ. (1995). *Nature,* **378,** 505 – 508.