[CANCER RESEARCH 53, 3853–3856, September 1, 1993]

Advances in Brief

Heterogeneity in MLL/AF-4 Fusion Messenger RNA Detected by the Polymerase Chain Reaction in t(4;11) Acute Leukemia[1]

Joanne M. Hilden,[2] Chien-Shing Chen, Roderick Moore, Joy Frestedt, and John H. Kersey

Departments of Laboratory Medicine/Pathology, Pediatric Hematology/Oncology, and Bone Marrow Transplantation, University of Minnesota, Minneapolis, Minnesota 55455

Abstract

We have designed a single polymerase chain reaction (PCR) primer pair that detects the *MLL/AF-4* fusion mRNA encoded by the derivative 11 chromosome from t(4;11)(q21;q23) leukemia cells using the reverse transcriptase PCR technique. PCR amplification was possible in seven of seven cells studied. Sequencing of the amplified products showed three different breakpoints on 11q23 and three on 4q21, resulting in six unique fusion sequences. All fusion sequences maintained an open reading frame. The areas of the *MLL* and *AF-4* genes that are conserved in all derivative 11 fusion RNAs and therefore likely to contribute to the function of the oncogenic fusion protein are centromeric regions of *MLL* through exon 6 (retaining the AT hook motif) and telomeric regions of *AF-4* beginning at codon 491 (containing nuclear localization and GTP-binding motifs). A single primer pair was able to detect the derivative 11 fusion transcript in seven of seven cases of t(4;11) acute leukemia tested. Given the variability shown in specific fusion sequences, studies correlating differential exon usage with clinical parameters will require different fusion-specific oligonucleotides or PCR primer pairs.

Introduction

A reciprocal translocation involving chromosomes 4 and 11 results in a characteristic form of acute leukemia. Leukemia with the t(4;11) is the most common form of infant ALL[3] and is characterized by a mixed lymphoid/myeloid phenotype and an extremely poor prognosis with conventional chemotherapy (1, 2). The two genes involved in the translocation, located within chromosome bands 11q23 and 4q21, have been cloned and are expressed in a variety of cell lines and tissues (3–10). The 11q23 gene (known as *MLL*, *HRX*, or *ALL-1*) is homologous to the *Drosophila trithorax* gene and has been shown to contain two different DNA-binding motifs, a centromeric AT-hook motif and a zinc-finger region located in the middle of the gene (3–7). The 4q21 gene (known as *AF-4* or *FEL*) encodes a serine/proline-rich protein (such regions have been associated with transcriptional activation) and contains possible nuclear localization and GTP-binding motifs in the telomeric region (9–11). Close breakpoint proximity on 11q23 has been shown in many studies (2, 4, 12–14); breakpoints on *AF-4* are less tightly clustered (8). In this study we have designed a single set of primers which will detect the derivative 11 fusion mRNA, using the RT-PCR technique, from seven t(4;11) leukemias, to specifically analyze breakpoint sequences. We have generated PCR products in all seven cases using an upstream primer in exon 6, just centromeric to *MLL* exon 7, and a downstream primer telomeric to *AF-4* codon 391. We have shown six unique fusion sequences, documenting variable *MLL* and *AF-4* exon usage in the der(11) fusion transcripts in t(4;11) acute leukemia.

Materials and Methods

Patient Samples and Cell Lines. The RS4;11 cell line, established in our laboratory (15), is available from American Type Culture Collection (Rockville, MD). MV4;11 and B1 are established t(4;11) cell lines described in detail elsewhere (16, 17). The AN4;11 cell line was established from an 8-month-old infant with t(4;11) ALL as described (8). Leukemia cells were obtained from three other patients (S. W. and M. P. were infants; A. B. was 14 years old) with t(4;11) ALL and cryopreserved before their use in these studies. In patient S. W., 1×10^8 leukemia cells were injected into the peritoneum of SCID mice and harvested 2 months later for cell culture; mRNA was extracted at this point. Additional cell lines obtained from American Type Culture Collection for use as negative controls were THP-1, a human monocytic cell line, and Nalm 6, a human pre-B cell line.

RNA Preparation and RT-PCR Amplification of *MLL/AF-4* Junctions. Polyadenylated RNA was isolated from cell lines using the FastTrack mRNA isolation kit (Invitrogen, San Diego, CA). For patient samples with limited cell number, the MicroFast Track (Invitrogen) system was used, with 2 million to 4 million cells/reaction tube. *In vitro* reverse transcription of mRNA to cDNA was performed with random hexamer priming using Avian Myeloblastosis virus reverse transcriptase in a 20-μl reaction mixture using a commercial kit (Invitrogen). Two μl of this reaction was diluted to 50 μl with a PCR mixture containing 2.5 units of Taq DNA polymerase, 10 mM Tris HCl, 50 nM KCl, 4.0 mM $MgCl_2$, 200 μM deoxynucleotide triphosphate, and 20 pmol of primers. We used the sequence of a der(11) cDNA clone from a RS4;11 cDNA library (9) to design PCR primers around the RS4;11 breakpoint. In addition, we also utilized the published *ALL-1* sequence (5). The primers used were: M1 (which primes in *MLL* exon 6), 5′-GAGCAGAGCAAACAGAAAA-3′, as the forward primer on the *MLL* gene; A1 (priming at *AF-4* codon 387), 5′-GGAAAGGAAACTTGGATGG-3′; and A2, (priming at *AF-4* codon 520), 5′-TTGGTCAGCCAGTTGTCCA-3′ as the reverse primers on the *AF-4* gene. For clarity, codon numbering for both the *AF-4* (10) and *MLL/HRX* (5) genes maintains that used in previous reports. After initial denaturation at 94°C for 7 min, denaturation, annealing, and extension were performed on an automated heat-block (DNA Thermal Cycler, Perkin Elmer-Cetus) at 94°C for 1 min, 55°C for 45 s, and 72°C for 1 min, respectively, for a total of 35 cycles. Nine μl of PCR mixture were size-fractionated on a 1% agarose gel and stained with ethidium bromide. Positive and negative controls were carried through all steps. Negative results from the control cell lines were verified by repeat testing with each primer set and the demonstration of integrity of isolated mRNA with amplification of the β-actin gene. Amplification of β-actin mRNA was accomplished with the following primers: forward 5′-CCTTCCTGGGCATGGAGTCCTG-3′; and reverse 5′-GGAGCAATGATCTTGATCTTC-3′.

Cloning, Southern Analysis, and Sequencing of PCR Products. PCR products were cloned into the pCR II sequencing vector (TA cloning; Invitrogen), which uses an AT overhang cloning site flanked by *Eco*RI restriction sites. The products were excised with *Eco*RI, size-fractionated on agarose gels, and transferred to Hybond-N nylon membranes (Amersham, Arlington Heights, IL). The membranes were probed with a 19-base pair oligonucleotide, p363, 5′-TGACCCATTCATGGCCTCCT-3′, which is the cDNA sequence beginning at *AF-4* codon 363, immediately 3′ to an exon telomeric to the RS4;11 breakpoint (8). The probe was end-labeled with Digoxigenin-NHS ester (Boehringer Mannheim, Indianapolis, IN) and hybridized according to the manufacturer's directions. Alternatively, the membranes were probed with

Received 5/13/93; accepted 7/20/93.

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

[1] This research was supported in part by a grant from the NIH and an Outstanding Investigator Grant Award (R35 CA49721) from the National Cancer Institute to J. H. K., and a Childrens Cancer Research Fund grant and an American Cancer Society Clinical Oncology fellowship (93–111–2) to J. M. H.

[2] To whom requests for reprints should be addressed, at Harvard Street at East River Road, Box 484 UMHC, Department of Pediatrics, University of Minnesota, Minneapolis, MN 55455.

[3] The abbreviations used are: ALL, acute lymphoblastic leukemia; RT, reverse transcriptase; PCR, polymerase chain reaction; cDNA, complementary DNA; Der, derivative chromosome.

pc606, which contains 205 base pairs of *MLL* cDNA fused to 1184 base pairs of *AF-4* cDNA. This probe was ^{32}P-labeled by random priming (US Biochemical, Cleveland, OH). Positive clones were sequenced using the T7 and SP6 promotor sequences in the pCR II vector as primers, by the dideoxy chain termination method. Both strands of two clones were sequenced from each cell line or patient sample.

Results

PCR Primers to Amplify the *MLL/AF-4* Fusion RNA. Fig. 1 shows the heterogeneity in PCR product size using primer pair M1/A2, which is the pair that amplifies a product from all seven t(4;11) leukemias tested. Identically sized fragments were obtained from RS4;11 and patient A. B. (approximately 700 base pairs) and from AN4;11 and patient S.W . (approximately 600 base pairs); B1 yields a fragment slightly smaller than RS4;11 and A.B. (Fig. 1*A*). The cell line MV4;11 yielded the smallest product, approximately 550 base pairs, and patient M. P. yielded the largest, approximately 850 base pairs. All cells except B1 yielded a product with the primer pair M1/A1 (which moves the *AF-4* priming site 400 base pairs in a centromeric direction; data not shown). PCR product size was not predictive of fusion sequence similarity (see below). The cell lines THP-1 (data not shown) and Nalm-6 (Fig. 1*A*) were consistently negative with both primer pairs but yielded the appropriate-sized fragment with β-actin primers (Fig. 1*B*).

Sequencing of *MLL/AF-4* Fusions and Characterization of Fusion mRNA Diversity. All PCR products yielded clones which were positive by Southern analysis using probes as described in "Materials and Methods." Sequencing of the PCR products from the seven t(4;11) leukemias yielded six different fusion junction patterns (Fig. 2). AN4;11 and S. W. yielded identical fusions, with *MLL* codon 1362 fused in frame to *AF-4* codon 347. RS4;11 and A. B. showed fusions that joined the same *MLL* and *AF-4* exons, but that differed by one codon on *AF-4* (a CAG trinucleotide, resulting in the absence of a glutamine residue in A. B.), with *MLL* codon 1406 fused in frame to *AF-4* codon 347 in RS4;11 but to *AF-4* codon 348 in A. B. This breakpoint occurs at the beginning of a 45-base pair exon on *AF-4*, which starts at codon 347 with a CAG trinucleotide, and is flanked on the 5′ end by an intron ending with a CAG trinucleotide (8). Thus this fusion seems to be the result of the alternate use of two adjacent CAG trinucleotides as a splice site, as previously noted by Domer *et al.* (9).

Despite very similar PCR fragment sizes, RS4;11 and B1 yield a very different fusion sequence. The *MLL* breakpoints are separated by 129 nucleotides, and the *AF-4* breakpoints are separated by 132 nucleotides. In B1, this results in the fusion, in frame, of *MLL* codon 1444 (which is the 3′ end of *MLL* exon eight) to *AF-4* codon 391. The breakpoints of M. P. are at *MLL* codon 1444 and *AF-4* codon 348. MV4;11 showed yet another fusion pattern, with *MLL* codon 1362 fused in frame to *AF-4* codon 362.

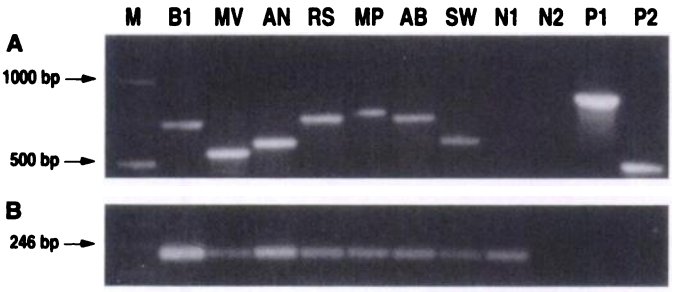

Fig. 1. PCR amplification of *ALL/AF-4* mRNA. Amplified fragments were separated by electrophoresis on 1% agarose gels. *A*, primer pair M1/A2. *B*, β-actin primers. *M*, size markers. *RS*, cell line RS4;11. *B1*, cell line B1. *MV*, cell line MV4;11. AN, cell line AN4;11. Patients were M.P., A. B., and S. W. *N1*, cell line Nalm 6. *N2*, no template. *P1* and *P2* are positive controls for the reverse transcription and PCR, respectively.

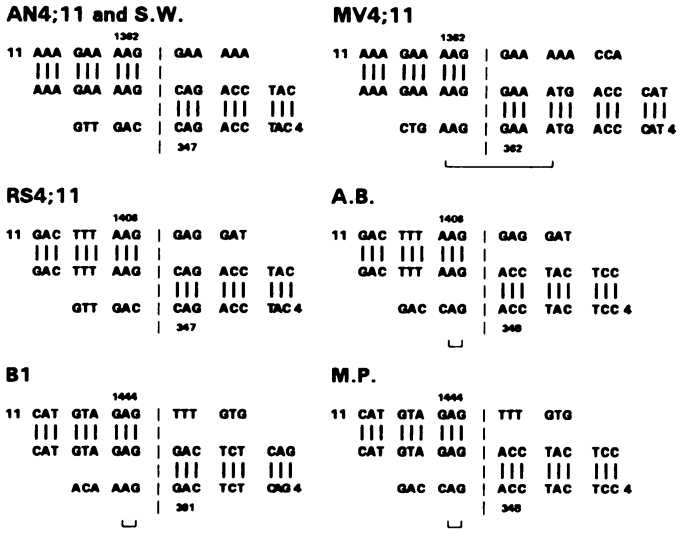

Fig. 2. Nucleotide sequences at junctions of *MLL* and *AF-4* fusion transcripts in the indicated t(4;11) cell line or patient sample. In each set, the *top row* is published *MLL* sequence, the *bottom row* is published *AF-4* sequence, and the *center row* is our der(11) fusion sequence. The codon numbers are taken from Ref. 5 for *MLL* and Ref. 10 for *AF-4*. *Brackets*, areas of sequence identity between *MLL* and *AF-4* at the fusion point.

These results are consistent with three different breakpoints on *MLL* (between exons 6 and 7, 7 and 8, and 8 and 9) as reported (4) and identifies three breakpoints on *AF-4* (at codons 347, 362, and 391) (Fig. 3). As shown schematically in Fig. 3, all cases tested retain on the derivative 11 fusion transcript the *MLL* gene through exon 6 and the *AF-4* gene telomeric to codon 391.

Discussion

In the present study, we have designed one pair of PCR primers which will detect the fusion region of the derivative 11 transcript in seven of seven t(4;11) acute leukemias and sequenced this region. Several lines of evidence suggest that it is the derivative 11 product that is etiologically important in 11q23 acute leukemia. Studies of complex translocations involving 11q23 show consistent conservation of the der(11) chromosome (18). Molecular studies have shown deletions of telomeric chromosome 11 DNA (13). We therefore sought to design a strategy to determine, using the RT-PCR technique with sequencing of the amplified products, those portions of the *MLL* and *AF-4* genes which are conserved and thus likely to be etiologically significant, in the der(11) transcript in different cases of t(4;11) acute leukemia.

Studies of the *MLL* and *AF-4* genes have shown that both are positioned with the 5′ end toward the centromere and that the der(11) transcript in t(4;11) leukemia fuses centromeric *MLL* to telomeric *AF-4* (4, 5, 9, 11). The *MLL* gene has many small (74–147-base pair) exons; several investigators have shown by Southern and sequence analysis that 11q23 breakpoints cluster between exons 6 and 7, 7 and 8, and 8 and 9 (4, 5, 12). Our sequence analysis corroborates exons 6 through 9 as the *MLL* fusion point in t(4;11) leukemia. In our previous work at the genomic level, breakpoints on *AF-4* were less clustered than those on *MLL*, and we identified an exon of 45 base pairs (codons 347–361) at the cDNA junction of the der(4) fusion transcript in the RS4;11 cell line (8). In the present study, we showed five of seven *AF-4* cases (including all three infants tested) fused *MLL* to one *AF-4* position, at the centromeric end of this 45-base pair exon. One of the other 2 breakpoints on *AF-4* is at the telomeric end of this exon (at codon 362). A cDNA sequence of a t(4;11) der(11) fusion reported from a patient with ALL (10) is in agreement with these results, with the *MLL* breakpoint between exons 7 and 8 and the *AF-4* break at

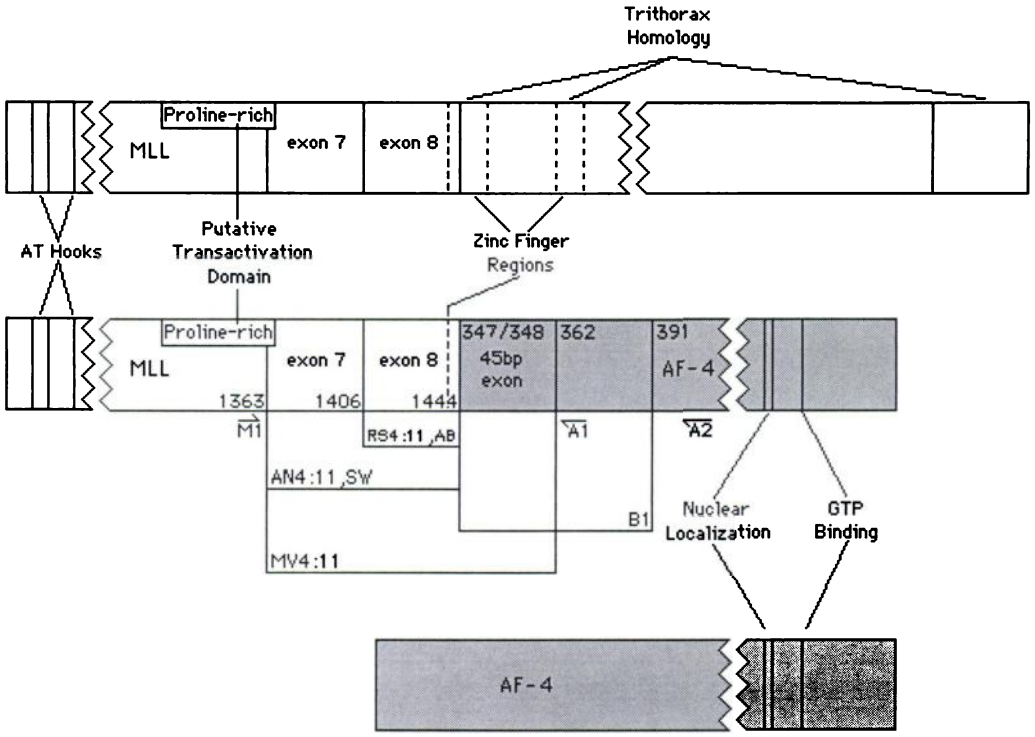

Fig. 3. Schematic representation of *MLL* and *AF-4* exons and functional domains and their location on the nonrearranged genes and the der(11) fusion genes resulting from the 7 different fusion patterns. The regions of *MLL* and *AF-4* deleted in each fusion transcript was determined by PCR product sequencing. Patient M. P. is shown as the prototype fusion of *MLL* exon 8 to *AF-4* codon 348. *Brackets*, deletions of regions of *MLL* and *AF-4* beyond that deleted in patient M. P.. The codon numbers of breakpoints and exon boundaries are indicated. *Interrupted lines*, region of zinc finger homology. *M1*, forward primer on *MLL*. *A1* and *A2*, reverse primers on *AF-4*.

codon 362. At the mRNA level, *AF-4* breakpoints are thus limited in their distribution, probably reflecting the small size of the exons involved relative to the large genomic region in which the breakpoints are distributed. Our data show that *MLL* exons 7 and 8, and the *AF-4* exons in codons 347–390, are inconsistently retained or lost in the t(4;11) der(11) fusion transcript and thus are unlikely to contribute to the function of the oncogenic fusion protein.

Comparison of our fusion sequences to published reports of *MLL* and *AF-4* regions of homology to known functional domains shows which of these are conserved in all of our cases (Fig. 3). The *MLL* gene is known to contain a centromeric AT-hook motif (5) (this domain binds to AT-rich regions in the minor groove of DNA); this is conserved in all 11q23 translocations since it is very close to the centromeric end of the gene. *MLL* also has three regions of homology to the *Drosophila trithorax* gene, with one at the 3′ end, and the other two, with homology to the zinc finger domains, beginning at the end of exon eight (4, 5). The zinc finger region is lost in five of seven of our cases, while the other two (M. P. and B1) retain 11 amino acids of the first zinc finger, keeping only two of the four cysteine residues needed to fold into the finger-like conformation. This would alter the DNA-binding function of the protein, however, as the first two zinc-fingers have a critical role in the specificity of the protein-nucleic acid interaction (19). Thus the zinc finger regions does not appear necessary for malignant transformation. Centromeric to the zinc fingers, Gu *et al.* (4) have reported a proline rich (27%) region, which could serve as a transcriptional activation domain. An elegant study of the proline rich (25%) region in the CTF/NF-1 family of CCAAT box binding proteins has shown that this region retains transcriptional activation activity even when separated from its DNA binding motif (20). In all cases we studied, the proline-rich region in *MLL* is retained on der(11) and thus separated from the DNA-binding activity of the zinc finger regions. On *AF-4*, areas of homology to known functional motifs are two telomeric areas retained in all fusion sequences, a nuclear localization consensus, and a possible GTP-binding motif, suggesting that this gene functions as a transcription factor (9, 11). Having determined the areas of each gene consistently retained on the der(11) fusion gene and thus potentially contributing to the function of the oncogenic fusion protein, these studies help to pinpoint the areas warranting further study using tools such as *in vitro* expression systems.

The RT-PCR method for detection of the t(4;11) fusion mRNA can now be applied to clinical cases. To screen infant ALL or to detect minimal residual disease in treated patients, a single primer pair, M1/A2, will detect the der(11) transcript in all cases we have tested so far. To look for a relationship between differential exon use and disease parameters, it will be necessary to detect specific fusion site sequences. This will require a set of primers or a fusion-specific oligonucleotide for each fusion pattern, since we have demonstrated that this cannot be extrapolated from PCR product size on agarose gels. From such studies we will learn more about the function of the oncogenic fusion protein and its clinical relevance, and we hope to have an impact on the therapy of this particularly refractory leukemia.

Acknowledgments

We wish to thank Bruce Bejcek, Ph.D., for helpful suggestions and Wendy Hudson for assistance in preparation of the figures. We also thank Dr. Martha Greenwood and Dr. Gregory Reaman for referral of samples from patients.

References

1. Arthur, D. C., Bloomfield, C. D., Lindquist, L. L., and Nesbit, M. E. Translocation 4;11 in acute leukemia. Clinical characteristics and prognostic significance. Blood, *59:* 96–99, 1982.
2. Chen, C-S., Sorenson, P. H. B., Domer, P. H., Reaman, G. H., Korsmeyer, S. J., Heerema, N. A., Hammond, G. D., and Kersey, J. H. Molecular rearrangements on chromosome 11q23 predominate in infant acute lymphoblastic leukemia and are associated with specific biologic variables and poor outcome. Blood, *81:* 2386–2393, 1993.
3. Ziemin-van der Poel, S., McCabe, N. R., Gill, H. J., Espinosa, R., Patel, Y., Harden, A., Rubinelli, P., Smith, S. D., LeBeau, M. M., Rowley, J. D., and Diaz, M. O. Identification of a gene, *MLL*, that spans the breakpoint in 11q23 translocations

associated with human leukemias. Proc. Natl. Acad. Sci. USA, *88:* 10735–10739, 1991.

4. Gu, Y., Nakamura, T., Alder, H., Prasad, R., Canaani, O., Cimino, G., Croce, C. M., and Canaani, E. The t(4;11) chromosome translocation of human acute leukemias fuses the *ALL-1* gene, related to *Drosophila trithorax*, to the *AF-4* gene. Cell, *71:* 701–708, 1992.
5. Tkachuk, D. C., Kohler, S., and Cleary, M. L. Involvement of a homolog of *Drosophila trithorax* by 11q23 chromosomal translocations in acute leukemias. Cell, *71:* 691–700, 1992.
6. Djabali, M., Selleri, L., Parry, P., Bower, M., Young, B. D., and Evans, G. A. A trithorax-like gene is interrupted by chromosome 11q23 translocations in acute leukemias. Nature Genet., *2:* 113–118, 1992.
7. 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., and Diaz, M. O. Cloning of cDNAs of the *MLL* gene that detect DNA rearrangements and altered RNA transcripts in human leukemic cells with 11q23 translocations. Proc. Natl. Acad. Sci. USA, *89:* 11794–11798, 1992.
8. Chen, C-S., Hilden, J. M., Frestedt, J., Domer, P. H., Moore, R., Korsmeyer, S. J., and Kersey, J. H. The chromosome 4q21 gene (*AF-4/FEL*) is widely expressed in normal tissues and shows breakpoint diversity in t(4;11)(q21↖3) acute leukemia. Blood, *82:* 1080–1085, 1993.
9. Domer, P. H., Fakharzadeh, S. S., Chen, C-S., Jockel, J., Johansen, L., Silverman, G. A., Kersey, J. H., and Korsmeyer, S. J. The acute mixed lineage leukemia associated with t(4;11)(q21:q23) produces an *MLL-AF-4* fusion transcript. Proc. Natl. Acad. Sci. USA, in press, 1993.
10. Morrissey, J., Tkachuk, D. C., Milatovich, A., Francke, U., Link, M., and Cleary, M. A serine/proline-rich protein is fused to *HRX* in t(4;11) acute leukemias. Blood, *81:* 1124–1131, 1993.
11. Nakamura, T., Alder, H., Gu, Y., Prasad, R., Canaani, O., Kamada, N., Gale, R. P., Lange, B., Crist, W. M., Nowell, P. C., Croce, C. M., and Canaani, E. Genes on chromosomes 4, 9, and 19 involved in 11q23 abnormalities in acute leukemia share sequence homology and/or common motifs. Proc. Natl. Acad. Sci. USA, *90:* 4631–4635, 1993.
12. Chen, C-S., Medberry, P. S., Arthur, D. C., and Kersey, J. H. Breakpoint clustering in t(4;11)(q21↖3) acute leukemia. Blood, *78:* 2498–2504, 1991.
13. Rowley, J. R., Diaz, M. O., Espinosa, R., Patel, Y. D., van Melle, E., Ziemin, S., Taillon-Miller, P., Lichter, P., Evans, G. A., Kersey, J. H., Ward, D. C., Domer, P. H., and LeBeau, M. M. Mapping chromosome band 11q23 in human acute leukemia with biotinylated probes: identification of 11q23 translocation breakpoints with a yeast artificial chromosome. Proc. Natl. Acad. Sci. USA, *87:* 9358–9362, 1990.
14. Cimino, G., Nakamura, T., Canaani, O., Prasad, R., Crist, W. M., Carroll, A. J., Baer, M., Bloomfield, C. D., Nowell, P. C., Croce, C. M., and Canaani, E. An altered 11-kilobase transcript in leukemic cell lines with the t(4;11)(q21↖3) chromosome translocation. Cancer Res., *52:* 3811–3813, 1992.
15. Stong, R. C., Korsmeyer, S. J., Parkin, J. L., Arthur, D. C., and Kersey, J. H. Human acute leukemia cell line with the t(4;11) chromosomal rearrangement exhibits B lineage and monocytic characteristics. Blood, *65:* 21–31, 1985.
16. Lange, B., Valtieri, M., Santoli, D., Caracciolo, D., Mavilio, F., Gemperlein, I., Griffin, C., Emanuel, B., Finan, J., Nowell, P., and Robera, G. Growth factor requirements of childhood acute leukemia: establishment of GM-CSF dependent cell lines. Blood, *70:* 192–198, 1987.
17. Cohen, A., Grunberger, T., Vanek, W., Dube, I. D., Doherty, P. J., Letarte, M., Roifman, C., and Freeman, M. H. Constitutive expression and role in growth regulation of interleukin-1 and multiple cytokine receptors in a biphenotypic leukemic cell line. Blood, *78:* 94–102, 1991.
18. Rowley, J. D. The der(11) chromosome contains the critical breakpoint junction in 4;11, 9;11, and 11;19 translocations in acute leukemia. Genes Chromos. Cancer, *5:* 264–266, 1992.
19. Latchman, D. S. Review article: eukaryotic transcription factors. Biochem. J., *270:* 281–289, 1990.
20. Mermod, N., O'Neill, E. A., Kelly, T. J., and Tijan, R. The proline-rich transcriptional activator of CTF/NF-1 is distinct from the replication and DNA binding domain. Cell, *58:* 741–753, 1989.

Heterogeneity in MLL/AF-4 Fusion Messenger RNA Detected by the Polymerase Chain Reaction in t(4;11) Acute Leukemia

Joanne M. Hilden, Chien-Shing Chen, Roderick Moore, et al.

Cancer Res 1993;53:3853-3856.