The EMBO Journal vol.9 no.2 pp.415-424, 1990

The *tal* gene undergoes chromosome translocation in T cell leukemia and potentially encodes a helix-loop-helix protein

Qi Chen, Jiin-Tsuey Cheng, Li-Huei Tsai, Nancy Schneider, George Buchanan[1], Andrew Carroll[1,2], William Crist[1,3], Bradford Ozanne, Michael J.Siciliano[4] and Richard Baer

Departments of Microbiology, Pathology and Pediatrics, University of Texas Southwestern Medical Center, Dallas, TX 75235, [1]Pediatric Oncology Group, St. Louis, MO 63108, [2]Laboratory of Medical Genetics, University of Alabama, Birmingham, AL 35294, [3]St Jude Children's Research Hospital and University of Tennessee, Memphis College of Medicine, Memphis, TE 38101 and [4]Department of Molecular Genetics, University of Texas M.D.Anderson Hospital Cancer Center, Houston, TX 77030, USA

Communicated by T.H.Rabbitts

We have analyzed t(1;14)(p32;q11) chromosome translocations from two patients with T cell acute lymphocytic leukemia. The chromosome 1 breakpoints of these patients lie within a kilobasepair of each other, and thus define a genetic locus (designated *tal*) involved in T cell oncogenesis. Moreover, we have identified sequences within *tal* that potentially encode an amphipathic helix-loop-helix motif, a DNA-binding domain found in a variety of proteins that control cell growth and differentiation. The homology domain of *tal* is especially related to that of *lyl*-1, a gene on chromosome 19 that has also been implicated in T cell oncogenesis. Hence, *tal* and *lyl*-1 encode a distinct family of helix-loop-helix proteins involved in the malignant development of lymphocytes.

Key words: chromosome translocation/helix-loop-helix motif/*tal* gene/T cell acute lymphocytic leukemia

Introduction

Cytogenetic abnormalities involving chromosome band 14q11 are often observed in the malignant cells of patients with T cell tumors. Three such karyotypic defects are recurrently found in T cell acute lymphocytic leukemia (T-ALL): the chromosome translocations t(8;14)(q24;q11); t(10,14)(q24;q11); and t(11,14)(p13;q11) (Williams *et al.*, 1984; Dubé *et al.*, 1986; Raimondi *et al.*, 1988). Their unique association with T-ALL suggests that these translocations actively promote leukemic development. Molecular studies show that the chromosome 14 breakpoints occur within the T cell receptor (TCR) α/δ chain gene at band 14q11 (Finger *et al.*, 1986; McKeithan *et al.*, 1986; Shima *et al.*, 1986; Kagan *et al.*, 1987; Boehm *et al.*, 1988b). Thus, expression of gene sequences from the other participating chromosomes could be altered upon juxtaposition with the TCR locus, and indeed this may be the mechanism by which these translocations influence neoplastic development. In turn, this hypothesis implies the existence of proto-oncogenes near the breakpoints of the other participating chromosomes. Interestingly, studies of the t(8;14)(q24;q11) translocation have borne out this prediction by localizing the chromosome 8 breakpoints to the vicinity of c-*myc*—a proto-oncogene frequently implicated in lymphoid tumorigenesis (Finger *et al.*, 1986; McKeithan *et al.*, 1986; Shima *et al.*, 1986).

Recent studies indicate that the t(1;14)(p32;q11) chromosome translocation may also be non-randomly associated with T-ALL. Lampert *et al.* (1988) observed t(1;14)(p32;q11) in two of ten patients with T-ALL, an isolated case was reported by Mathieu-Mahul *et al.* (1986), and five cases were uncovered in a cytogenetic study of 251 T-ALL patients (A.Carroll and W.Crist, unpublished data). Thus, the t(1;14)(p32;q11) translocation is a rare but none the less recurrent karyotypic marker of all T-ALL. A grossly similar cytogenetic defect, t(1;14)(p33;q11), has been observed in a stem cell leukemia with both lymphoid and myeloid manifestations (Hershfield *et al.*, 1984); it remains to be seen whether this is related at the molecular level to the t(1;14)(p32;q11) translocation associated with typical T-ALL.

Here we report the molecular characterization of the t(1;14)(p32;q11) chromosome translocation. Analysis of the translocation junctions from two T-ALL patients reveals that

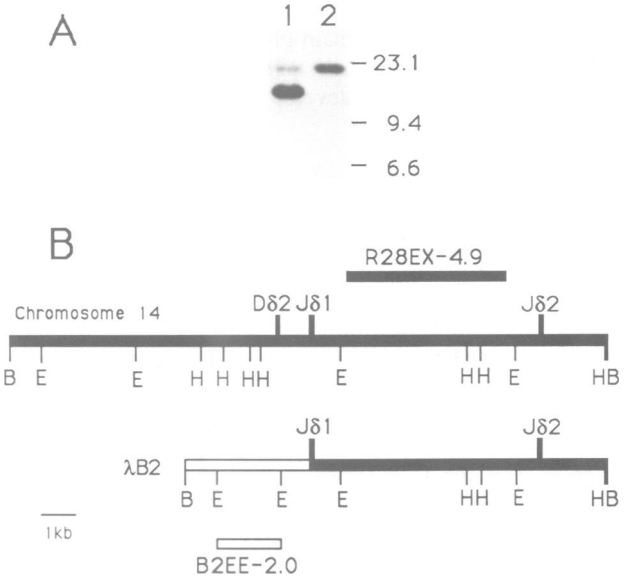

Fig. 1. Isolation of the t(1;14)(p32;q11) translocation junction from patient 4. (**A**) Southern analysis of *Bam*HI-digested DNAs hybridized with the TCR Jα probe R28EX-4.9. **Lane 1**, patient 4 leukemic DNA; **lane 2**, patient 4 remission DNA. Sizes of *Hin*dIII-digested λ DNA fragments are indicated (in kb) to the right of the autoradiogram. (**B**) Recombinant clone λB2 was isolated from a phage library of patient 4 leukemia DNA with probe R28EX-4.9. The map of λB2 is compared to that of the Dδ-Jδ region of a germline TCR α/δ chain gene. Closed boxes represent chromosome 14 sequences and open boxes chromosome 1 sequences. Positions of the DNA probes R28EX-4.9 and B2EE-2.0 are indicated. Restriction sites: B, *Bam*HI; E, *Eco*RI; H, *Hin*dIII.

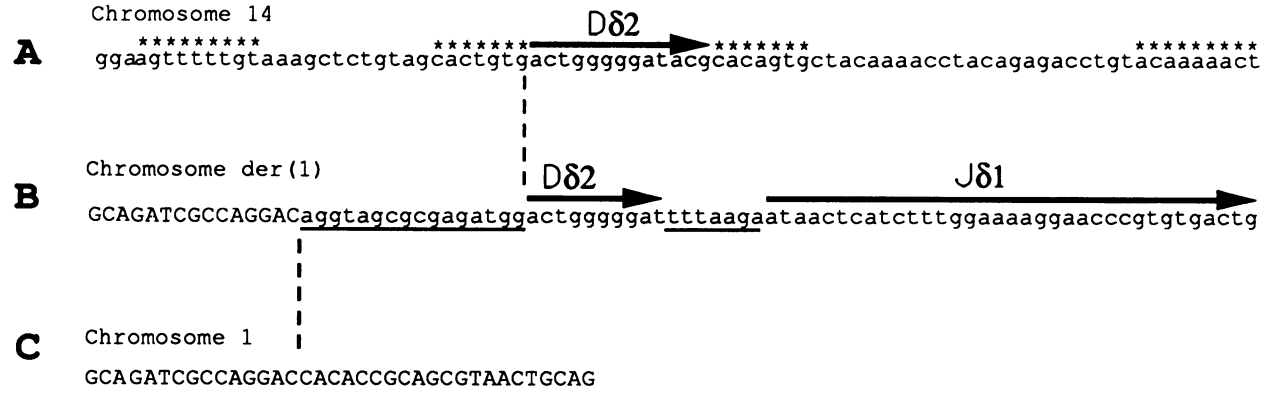

Fig. 2. The t(1;14)(p32;q11) translocation junction of patient 4. Comparison of (**A**) chromosome 14 sequences around the Dδ2 gene segment (Boehm *et al.*, 1988b), (**B**) chromosome der(1) sequences encompassing the t(1;14)(p32;q11) junction and (**C**) chromosome 1 sequences from the *tal* locus. Nucleotides in upper-case letters are derived from chromosome 1 and those in lower-case letters are derived from chromosome 14 (excepting nucleotides generated by N-region insertion, which are underlined). Coding sequences of the TCR Dδ2 and Jδ1 gene segments are overlined with arrows. The conserved heptamer and nonamer elements of the Dδ2 recombination signals are marked with asterisks. Dashed vertical lines indicate the translocation breakpoints on chromosomes 1 and 14.

the chromosome 14 breakpoints occur within the TCR δ chain gene. Also, the chromosome 1 breakpoints from both patients lie just 1 kb apart, suggesting that this region of chromosome 1 harbors a genetic locus (designated *tal*, for T-cell acute leukemia) involved in leukemogenesis. Moreover, we have identified sequences within *tal* that encode a protein motif homologous to the helix−loop−helix domain found in a number of proteins involved in the control of cell growth and differentiation (Villares and Cabrera, 1987; Murre *et al.*, 1989a) including MyoD1, myogenin, Myf-5, the *Drosophila achaete-scute, twist, daughterless* and *enhancer of split* proteins, the E12/E47 immunoglobulin enhancer-binding proteins and products of the *myc* and *lyl*-1 oncogenes. Since this domain plays a demonstrated role in DNA binding (Murre *et al.*, 1989a), the *tal* gene product may influence leukemic development by direct interaction with DNA.

Results and Discussion

The junction of the t(1;14)(p32;q11) chromosome translocation from a T-ALL patient

Cytogenetic analysis identified a t(1;14)(p32;q11) chromosome translocation in the leukemic cells of T-ALL patient 4 (see Materials and methods). To determine whether the chromosome 14 breakpoint occurred within the TCR α/δ chain gene, we examined patient DNA by Southern hybridization with a genomic DNA probe (R28EX-4.9) derived from the TCR Jδ region. In addition to the germline 17.5 kb *Bam*HI fragment, this probe detected a rearranged 12.2 kb fragment in leukemic DNA (Figure 1A, lane 1) that was clearly absent in DNA obtained from the patient after remission (Figure 1A, lane 2). To study further the tumor-specific δ gene rearrangement from this patient, we constructed a lambda phage library of *Bam*HI-digested leukemic DNA. Screening with the R28EX-4.9 probe identified recombinant clones containing the rearranged 12.2 kb fragment, and analysis of one such clone (λB2) revealed that its restriction map diverges from that of the germline δ chain locus in the vicinity of the Jδ1 gene segment (Figure 1B).

If this DNA rearrangement represents a junction of the t(1;14)(p32;q11) translocation, then sequences upstream of the divergence point should be derived from chromosome 1. Therefore a 2.0 kb *Eco*RI fragment from this region was isolated and subcloned into a plasmid vector. This clone (B2EE-2.0) was then used as a probe in Southern filter hybridizations with DNAs extracted from a panel of 17 human/hamster somatic cell hybrids with randomly segregated human chromosomes (Thompson *et al.*, 1987; Stallings *et al.*, 1988). The hybridization of B2EE-2.0 was perfectly concordant with chromosome 1 and randomly associated (18−65% discordancy) with every other human chromosome (Table I, Materials and methods). Furthermore, the pattern of hybridization with a panel of hybrids containing broken derivatives of chromosome 1 (Thompson *et al.*, 1987) provided a regional localization of B2EE-2.0 to the short arm of chromosome 1 (see Materials and methods). Thus the contiguity of sequences from chromosome 14 (i.e. TCR δ gene) and chromosome 1 (i.e. B2EE-2.0) in the 12.2 kb *Bam*HI fragment of λB2 demonstrates that the intervening DNA rearrangement represents the junction of t(1;14)(p32;q11).

The t(1;14)(p32;q11) translocation generates two abnormal chromosomes, designated der(1) and der(14); these contain the centromeres derived from chromosomes 1 and 14, respectively. From the known orientation of the TCR α/δ chain gene on chromosome 14 we can deduce that phage clone λB2 represents the t(1;14)(p32;q11) junction on chromosome der(1). To evaluate further the nature of the der(1) junction, the divergence point of λB2 was subjected to nucleotide sequence analysis. As illustrated in Figure 2(B), the TCR Dδ2 and Jδ1 segments of chromosome der(1) are juxtaposed as a consequence of TCR gene recombination. In addition, however, sequences of unknown origin are found upstream of the Dδ2 segment (Figure 2B). If these novel sequences are indeed derived from chromosome 1, then the 14q11 breakpoint of t(1;14)(p32;q11) occurs immediately upstream of the recombined Dδ2−Jδ1 segments.

The t(1;14)(p32;q11) junction from a second T-ALL patient

A t(1;14)(p32;q11) chromosome translocation was also observed in the leukemic cells of T-ALL patient 5 (see

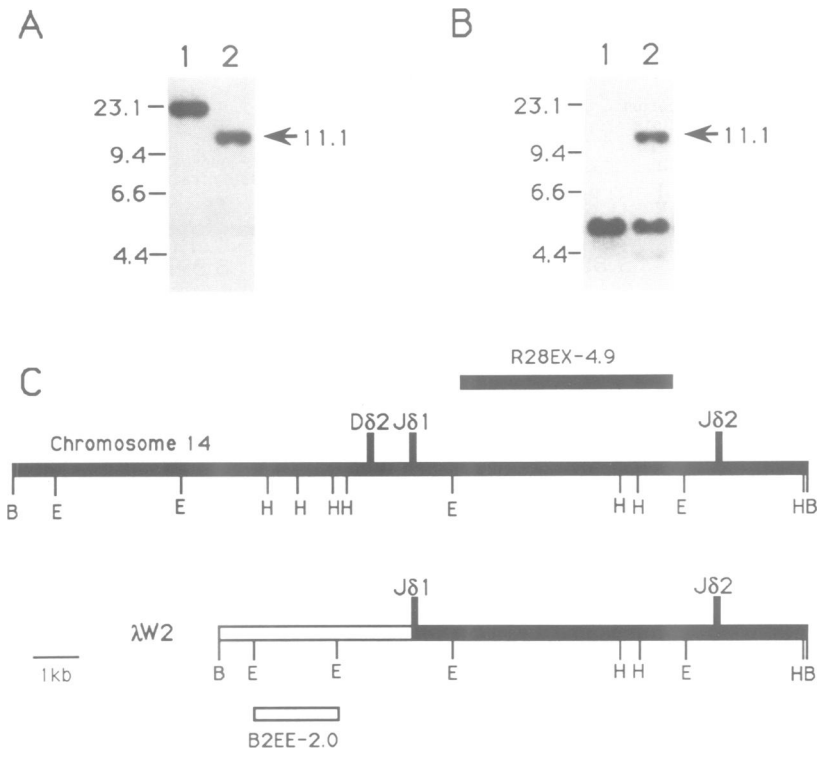

Fig. 3. Isolation of the t(1;14)(p32;q11) junction from patient 5. Southern analysis of *Bam*HI-digested DNAs hybridized with probe R28EX-4.9 (**A**) or probe B2EE-2.0 (**B**). **Lanes 1**, control DNA from an unrelated Wilm's tumor specimen; **lanes 2**, patient 5 leukemic DNA. Sizes of *Hin*dIII-digested λ DNA fragments are indicated (in kb) to the left of each autoradiogram. The rearranged 11.1 kb *Bam*HI fragment is marked with arrows. (**C**) Recombinant clone λW2 was isolated from a phage library of patient 5 leukemic DNA with probe B2EE-2.0. The λW2 map is compared to that of the Dδ–Jδ region of a germline TCR α/δ chain gene. Closed boxes represent chromosome 14 sequences and open boxes chromosome 1 sequences. Positions of the DNA probes R28EX-4.9 and B2EE-2.0 are indicated. Restriction sites: B, *Bam*HI; E, *Eco*RI; H, *Hin*dIII.

Materials and methods). Southern hybridization with the TCR Jδ gene probe (R28EX-4.9) identified a rearranged 11.1 kb *Bam*HI fragment in leukemic DNA from this patient (Figure 3A); hybridization with the chromosome 1 probe B2EE-2.0 also detected a rearranged *Bam*HI fragment of 11.1 kb (Figure 3B). Therefore a lambda phage library of *Bam*HI-digested leukemic DNA from patient 5 was constructed, and screening with the B2EE-2.0 probe yielded recombinant phage containing the rearranged 11.1 kb *Bam*HI fragment. Simultaneous screening with the chromosome 14 probe (R28EX-4.9) revealed that these clones also annealed to DNA from the TCR Jδ region. Hence, the rearranged 11.1 kb *Bam*HI fragment contains sequences derived from both 1p32 and 14q11, and therefore is likely to constitute the der(1) junction of the t(1;14)(p32;q11) chromosome translocation. This was confirmed by restriction mapping of one of these clones (λW2), which shows it to be comprised of chromosome 1 sequences juxtaposed with sequences from the TCR α/δ chain locus of chromosome 14 (Figure 3C). Nucleotide sequence analysis of λW2 reveals DNA of unknown origin joined to coding sequences of the TCR Jδ1 gene segment (Figure 4), suggesting that the 14q11 breakpoint in patient 5 occurs within Jδ1.

The tal locus; a breakpoint cluster region on chromosome 1

Restriction mapping of the t(1;14)(p32;q11) junctions from patients 4 and 5 revealed a common breakpoint region on chromosome 1 that can be identified with the DNA probe B2EE-2.0 (Figures 1B and 3C); this region (designated *tal*)

Fig. 4. The t(1;14)(p32;q11) junction of patient 5. Comparison of (**A**) chromosome 14 sequences around the Jδ1 gene segment Boehm *et al.*, 1988a), (**B**) chromosome der(1) sequences encompassing the t(1;14)(p32;q11) junction and (**C**) chromosome 1 sequences from the *tal* locus. Nucleotides in upper-case letters are derived from chromosome 1 and those in lower-case letters are derived from chromosome 14 (excepting nucleotides generated by N-region insertion, which are underlined). Coding sequences of the TCR Jδ1 segment are overlined with an arrow. Dashed vertical lines indicate the translocation breakpoints on chromosome 1 and 14.

was further characterized by isolating corresponding unrearranged genomic DNA sequences. The λW3 recombinant phage clone was isolated upon screening the patient 5 leukemic DNA library with B2EE-2.0; λW3 harbors a germline 5.0 kb *Bam*HI fragment (Figures 3B and 5) that is presumably derived from the normal chromosome 1 in patient 5 leukemic cells. The λSU2 clone was isolated by B2EE-2.0 screening of a library of genomic DNA from SUP-T1 (Baer *et al.*, 1985), a T cell lymphoma line that

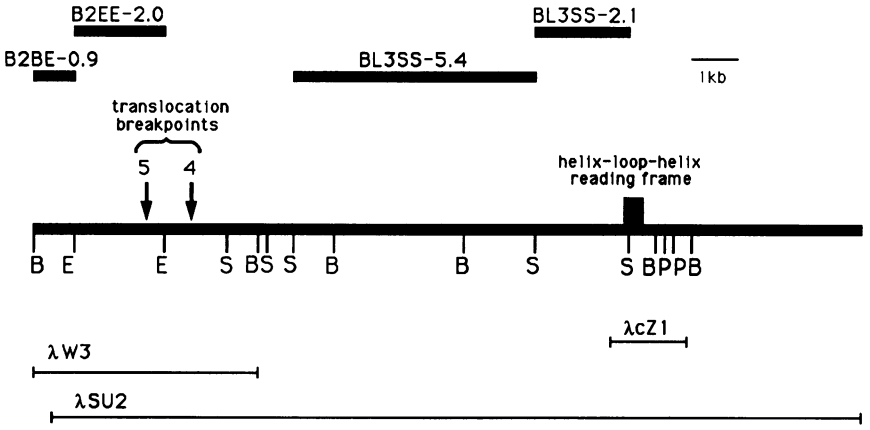

Fig. 5. Map of the *tal* locus on chromosome 1. The λW3 and λSU2 clones were isolated by screening libraries of genomic DNA with the B2EE-2.0 probe. Analyses of λW3 and λSU2 provide an 18.6 kb restriction map of the *tal* locus in its normal configuration. The t(1;14)(p32;q11) translocation breakpoints from patients 4 and 5 are marked with arrows. DNA probes are indicated above the map. The co-linearity between the 1.8 kb cDNA clone λcZ1 and the *tal* locus is illustrated, as well as the position of the helix−loop−helix reading frame. Restriction sites: B, *Bam*HI; E, *Eco*RI; P, *Pst*I; S, *Sac*I (*Pst*I sites are not complete).

bears only the normal homologs of chromosome 1 (Hecht *et al.*, 1984). Alignment of the overlapping λW3 and λSU2 clones provides an 18.6 kb restriction map of the *tal* locus (Figure 5); extensive Southern analyses of non-leukemic DNAs from a variety of human sources confirm that this map represents the normal structure of *tal* (data not shown).

Recently, two groups have analyzed the t(1;14)(p33;q11) translocation from DU.528 (Begley *et al.*, 1989; Finger *et al.*, 1989), a stem cell leukemia line with lymphoid and myeloid characteristics (Hershfield *et al.*, 1984; Kurtzberg *et al.*, 1985). It would be intriguing to determine whether this translocation is functionally equivalent to the t(1;14)-(p32;q11) defect observed in typical T-ALL. Interestingly, the chromosome 14 breakpoint in DU.528 also occurs within the TCR δ chain gene (Begley *et al.*, 1989; Finger *et al.*, 1989). However, available restriction data for the chromosome 1 breakpoint are not compatible with the map of the *tal* locus, suggesting that the DU.528 breakpoint occurs at a distal site within *tal* or involves a distinct genetic entity on chromosome 1.

As illustrated in Figure 5, the t(1;14)(p32;q11) breakpoints of patients 4 and 5 lie just 1 kb apart within the *tal* locus. Nucleotide sequences of the germline *tal* locus were determined, and in Figures 2(C) and 4(C) these are compared with the corresponding t(1;14)(p32;q11) junctions from patients 4 and 5 respectively. These comparisons reveal the presence of short sequences at the translocation junctions (16 residues in patient 4, Figure 2B; 5 residues in patient 5, Figure 4B) that are derived from neither chromosome 1 or 14. These are reminiscent of the N-region insertions generated during normal TCR gene assembly (Kronenberg *et al.*, 1986), suggesting that the translocations were engendered by aberrant activity of the TCR gene recombinase.

RNA transcription of the tal locus

In order to identify transcribed sequences within the *tal* locus, a collection of genomic probes (B2BE-0.9, B2EE-2.0, BL3SS-5.4, and BL3SS-2.1; Figure 5) was used to screen a cDNA library prepared from poly(A)-selected RNA of the T-ALL cell line CCRF-CEM (Foley *et al.*, 1965). A single hybridizing clone was identified (λcZ1) and restriction mapping of the 1.8 kb cDNA insert (Figure 7A) revealed

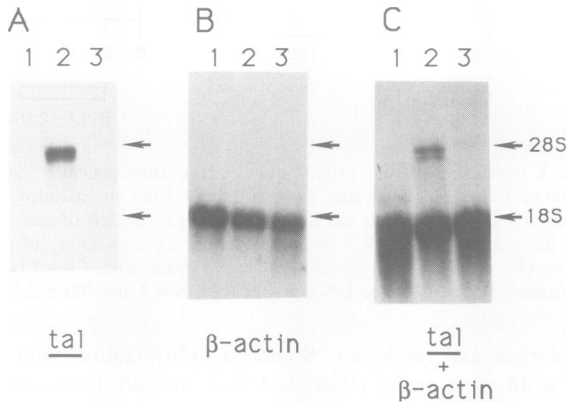

Fig. 6. RNA transcription of the *tal* locus. Northern analyses of poly(A)$^+$RNA from the pre-B leukemia cell line SMS-SB (**lanes 1**), the T-ALL line Jurkat (**lanes 2**) and the Burkitt's lymphoma line Ly65 (**lanes 3**). (**A**) Northern filter hybridized with the *tal* cDNA probe λcZ1. (**B**) The same filter was stripped of λcZ1 probe, and rehybridized with a β-actin cDNA probe. (**C**) A second filter was hybridized with λcZ1, and (without removal of the λcZ1 probe) rehybridized with the β-actin cDNA probe. Mobilities of the 28S and 18S rRNA markers are indicated to the right of each autoradiogram.

it to be co-linear with genomic DNA located ~11 kb from the translocation breakpoints (see Figure 5). RNA transcription of *tal* was evaluated by using the λcZ1 insert DNA as a probe in Northern hybridization experiments. Unfortunately, RNA preparations from leukemic cells of patients 4 and 5 were not available, thereby precluding analysis of transcription from translocated alleles of the *tal* locus. Regardless, the λcZ1 probe detected two closely migrating species in poly(A)$^+$ RNA from the T-ALL cell line Jurkat; these transcripts were not observed, however, in RNA prepared from the B cell tumor lines SMS-SB and Ly65 (Figure 6A). The specific expression of *tal* sequences in Jurkat cells was confirmed in a control experiment that identified similar levels of β-actin gene transcripts in the three cell lines (Figure 6B). The *tal* transcripts migrate as ~5.0 and ~4.8 kb species which are better resolved in the Northern experiment presented in Figure 6(C). Preliminary analyses of *tal* transcription in a larger panel of lymphoid tumors suggest expression of the 5.0/4.8 kb species in some, but not all, T-ALL cell lines.

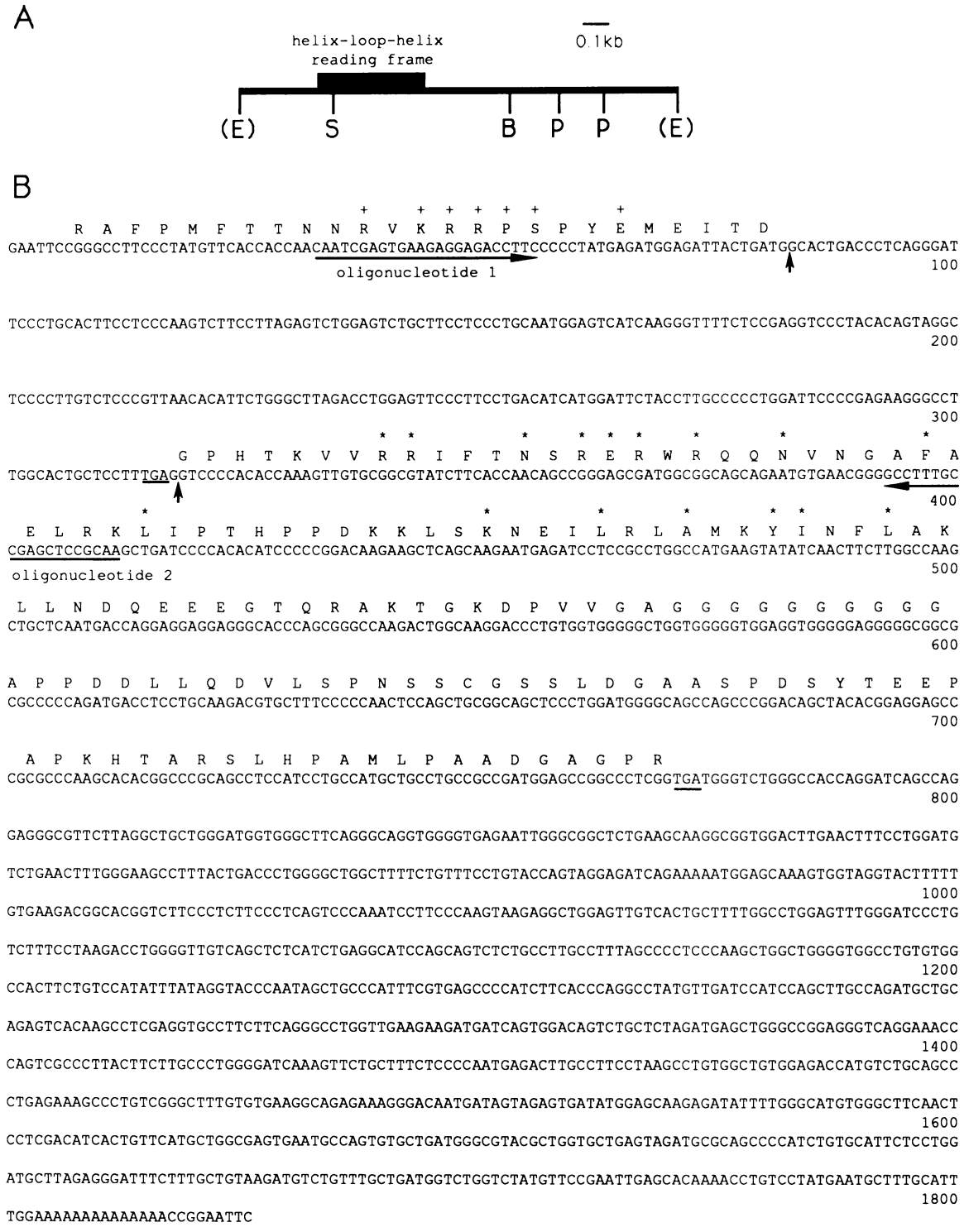

Fig. 7. Identification of the helix – loop – helix motif encoded by *tal*. (**A**) Restriction map of λcZ1, a cDNA clone of a *tal* RNA transcript from CCRF-CEM cells. The location of the helix – loop – helix reading frame is indicated by a closed box. Restriction sites: B, *Bam*HI; E, *Eco*RI; P, *Pst*I; S, *Sac*I. *Eco*RI sites in parentheses were generated during cDNA construction. (**B**) The complete nucleotide sequence of the *tal* cDNA clone λcZ1. The conceptual reading frame that includes the helix – loop – helix motif is specified by nt 318 – 770, and the TGA termination codons that delimit the reading frame are underlined. The 16 consensus amino acid of the helix – loop – helix motif, as defined by Mellentin *et al.* (1989), are marked with asterisks. The conceptual reading frame encoded by nt 7 – 82 is homologous to the amino acid sequence just preceding the *lyl*-1 helix – loop – helix domain; the seven residues marked by asterisks in this frame are identical in *tal* and *lyl*-1. The vertical arrows indicate RNA splice sites identified by experiments described in Figure 9. The positions of synthetic oligonucleotide primers are illustrated; the oligonucleotide 1 sequence is as shown, whereas the oligonucleotide 2 sequence is the reverse complement of that shown (see Materials and methods).

tal sequences encode an amphipathic helix – loop – helix motif

The 1.8 kb sequence of the λcZ1 cDNA clone is presented in Figure 7(B). In order to identify amino acid coding sequences of *tal*, potential reading frames were compared with the Protein Identification Resource (National Biomedical Research Foundation) using a computer search algorithm (Lipman and Pearson, 1985). One such reading frame of

```
                          amphipathic                               amphipathic
                            helix I                                   helix II
                        ===============                            =============
             **   * *** *    *  *  *                            *   *  *  **  *
tal         VRRIFTNSRERWRQQNVNGAFAELRKLIP----THPP----------DKKLSKNEILRLAMKYINFLAKLLN
lyl-1       ARRVFTNSRERWRQQNVNGAFAELRKLLP----THPP----------DRKLSKNEVLRLAMKYIGFLVRLLR
E12         ERRVANNARERLRVRDINEAFKELGRMCQ----LHLN---------SEKPQTKLLILHQAVSVILNLEQQVR
E47         ERRVANNARERVRVRDINEAFRELGRMCQ----MHLK---------SDKAQTKLLILQQAVQVILGLEQQVR
N-myc       ERRRNHNILERQRRNDLRSSFLTLRDHVP----ELVK----------NEKAAKVVILKKATEYVHSLQAEEH
c-myc       VKRRTHNVLERQRRNELKRSFFALRDQIP----ELEN----------NEKAPKVVILKKATAYILSVQAEEQ
L-myc       TKRKNHNFLERKRRNDLRSRFLALRDQVP----TLAS----------CSKAPKVVILSKALEYLQALVGAEK
MyoD1       DRRKAATNRERRRLSKVNEAFETLKRCTS----SNP-----------NQRLPKVEILRNAIRYIEGLQALLR
myogenin    DRRRAATLREKRRLKKVNEAFEALKRSTL----JNP-----------NQRLPKVEILRSAIQYIERLQALLS
Myf-5       DRRKAATMRERRRLKKVNQAFETLKRCTT----TNP-----------NQRLPKVEILRNAIRYIESLQELLR
da          ERRQANNARERIRIRDINEALKELGRMCM----THLK---------SDKPQTKLGILNMAVEVIMTLEQQVR
twist       NQRVMANVRERQRTQSLNDAFKSLQQIIP----TLPS-----------DKLSKIQTLKLATRYIDFLCRMLS
AS-C T3     ARR---NARERNRVKQVNNGFVNLRQHLPQTVVNSLSNG----GRGSSKKLSKVDTLRIAVEYIRGLQDMLD
AS-C T4     QRR---NARERNRVKQVNNSFARLRQHIPQSIITDLTKG--G-GRGPHKKISKVDTLRIAVEIYRSLQDLVD
AS-C T5     IRR---NARERNRVKQVNNGFSQLRQHIPAAVLADLSNGRRGIGPGANKKLSKVSTLKMAVEIYRRLQKVLH
AS-C T8     ARR---NARERNRVKQVNNGFALLREKIPEEVSEAFEAQ--GAGRGASKKLSKVETLRMAVEYIRSLEKLLG
```

Fig. 8. Amino acid similarities between *tal* and other helix–loop–helix proteins. The amino acid residues of *tal* and other known proteins are aligned to illustrate homologies between their helix–loop–helix domains. The 16 consensus amino acids shared by most of the proteins are marked with asterisks (Mellentin *et al.*, 1989). Those residues that form the proposed amphipathic helices are overlined with bars (Murre *et al.*, 1989a).

150 amino acids displayed remarkable homology to a family of proteins involved in the control of cell growth and differentiation (nt 318–770; Figure 7B). The region of homology is restricted to an ~60 amino acid domain that has the potential to form two amphipathic helices separated by an intervening loop (Murre *et al.*, 1989a). The 'helix–loop–helix' homology domain has been identified in oncoproteins encoded by the *myc* gene family (Villares and Cabrera, 1987) and the recently described *lyl*-1 gene (Mellentin *et al.*, 1989). The MyoD1, myogenin and Myf-5 proteins that mediate myoblast conversion also share this motif (Davis *et al.*, 1987; Wright *et al.*, 1989; Braun *et al.*, 1989; Edmondson and Olson, 1989), as do the immunoglobulin enhancer-binding proteins E12 and E47 (Murre *et al.*, 1989a). Helix–loop–helix domains are also encoded by several loci involved in *Drosophila* development, including four genes of the *achaete-scute* complex required for neurogenesis (Villares and Cabrera, 1987; Alonso and Cabrera, 1988); three transcripts of the *enhancer of split* gene complex also involved in neurogenesis (Klaembt *et al.*, 1989); *daughterless*, a locus essential for sex determination and neural differentiation (Caudy *et al.*, 1988); and *twist*, which controls germ layer formation during embryogenesis (Thisse *et al.*, 1988). In Figure 8, the amino acid homology domains of known helix–loop–helix proteins are compared; the sequence identities between any of these 15 proteins and the corresponding helix–loop–helix domain encoded by *tal* range 29–84%. Significantly, the *tal* motif contains each of the 16 consensus residues (marked by asterisks in Figure 8) characteristic of the helix–loop–helix domain as defined by Mellentin *et al.* (1989).

The helix–loop–helix domain of *tal* is comprised of the amino-terminal 60 residues of its 150 amino acid reading frame in λcZ1 (Figure 7B). However, the λcZ1 cDNA is likely to have been derived from a non-functional RNA transcript since the helix–loop–helix domain is immediately preceded by an in-phase termination codon (underlined in Figure 7B). The non-functional transcript may have arisen by either incomplete or improper splicing of *tal* mRNA. This is consistent with Northern analyses which detect little, if any, 5.0/4.8 kb *tal* transcripts in CCRF-CEM cells (unpublished data). Nucleotide sequence comparison of genomic *tal* DNA (Figure 9A) and λcZ1 (Figure 9B) reveals the presence of an RNA splice acceptor site immediately upstream of the genomic sequence that encodes the helix–loop–helix domain (Mount, 1982). Splicing at this position would be required for expression of the helix–loop–helix domain since the genomic sequence is also preceded by an in-phase termination codon (underlined in Figure 9A). Clearly, the λcZ1 transcript had undergone splicing at this site (Figure 9B), but it had done so in a manner that generated a non-functional message.

As discussed below, the helix–loop–helix domain of *tal* is most closely related to that encoded by the *lyl*-1 gene (Mellentin *et al.*, 1989). Thus, it is noteworthy that sequences of λcZ1 (nt 37–66; Figure 7B) potentially encode a ten amino acid stretch with strong homology to the region immediately preceding the helix–loop–helix domain of *lyl*-1 (i.e. seven of ten residues are identical; these are marked with the + symbol in Figure 7B). Hence, these may correspond to the upstream sequences that are normally spliced to the helix–loop–helix coding sequences in functional *tal* mRNA. To investigate this possibility we prepared synthetic oligonucleotides that flank the potential splice site—i.e. a 24mer complementary to nts 33–56 (oligonucleotide 1; Figure 7B) and a 20mer complementary to nts 393–412 (oligonucleotide 2; Figure 7B). These were used as primers in a polymerase chain reaction to amplify sequences of poly(A)$^+$ RNA from Jurkat, a T-ALL cell line that expresses detectable levels of the 5.0/4.8 kb *tal* mRNA (Figure 6). This reaction generated a single amplification product of 144 bp, the nucleotide sequence of which is illustrated in Figure 9(C). As shown, the upstream *tal* sequences that encode the short amino acid stretch with *lyl*-1 homology are spliced in-frame with *tal* sequences encoding the helix–loop–helix domain (Figure 9B). Hence, this is likely to represent the sequence of properly spliced, functional *tal* mRNA.

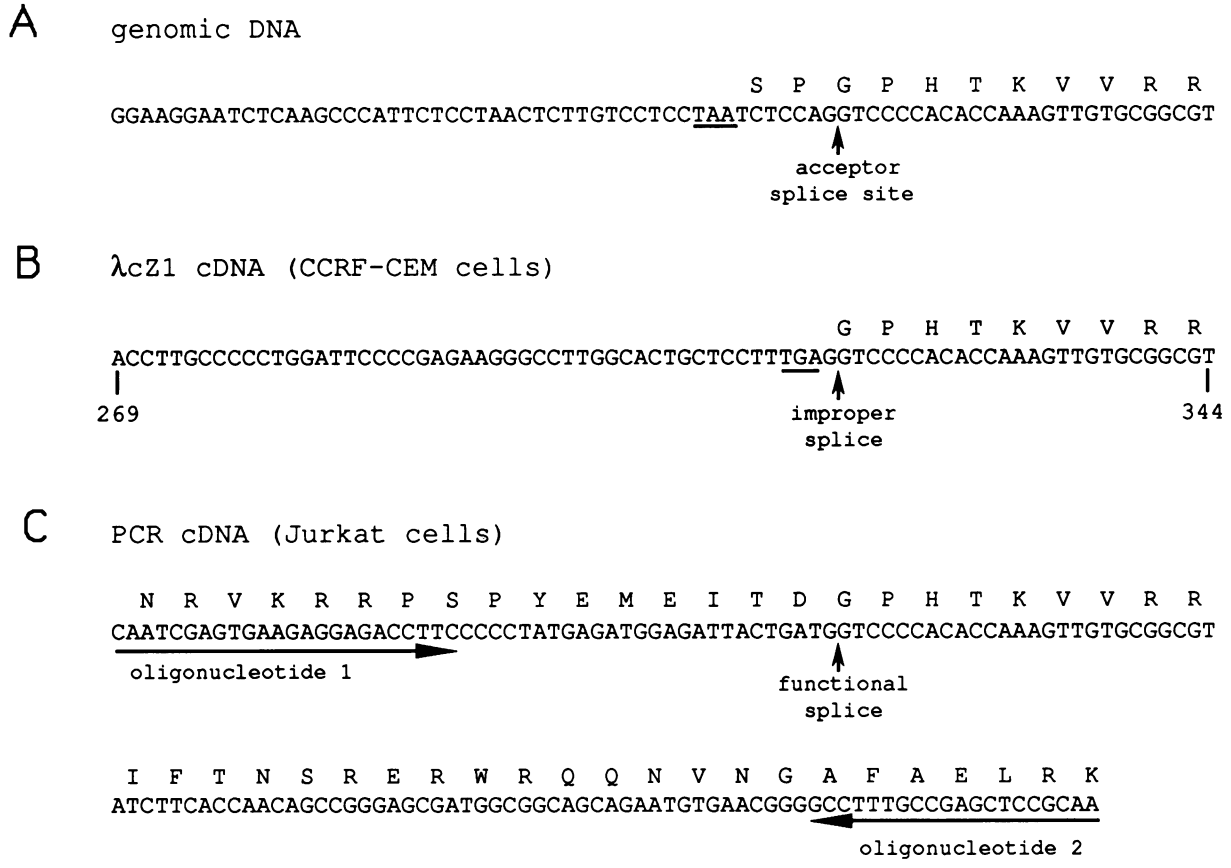

Fig. 9. RNA splicing of *tal* gene transcripts. (**A**) Nucleotide sequence of *tal* genomic DNA at the 5′ exon–intron border of the helix–loop–helix reading frame. The sequence was derived from λSU2, the genomic DNA clone illustrated in Figure 5. The in-phase TAA termination codon is underlined, and the RNA acceptor splice site is indicated by a vertical arrow (Mount, 1982). (**B**) Nucleotide residues 269–344 of λcZ1, the cDNA clone derived from poly(A)$^+$ RNA of CCRF-CEM cells (Figure 7). The in-phase TGA termination codon is underlined, and the position of the improper RNA splice is indicated. (**C**) Nucleotide sequence of a cDNA clone derived from poly(A)$^+$ RNA of Jurkat cells. The 148 bp cDNA fragment was generated by polymerase chain reaction using oligonucleotides 1 and 2 (Figure 7) as primers. The position of the functional RNA splice is indicated.

tal and lyl-1: a family of helix–loop–helix proteins involved in T cell oncogenesis

The helix–loop–helix proteins all share a moderate level of primary sequence homology that is likely to be sufficient for formation of the domain's proposed tertiary structure (Murre *et al.*, 1989a). However, those proteins that were identified on the basis of similar functional properties exhibit far more dramatic similarities in their helix–loop–helix domains. For example, the Myc oncoproteins (c-Myc, N-Myc, and L-Myc) share 53–67% amino acid identity in this region (Alt *et al.*, 1986). Striking sequence identities are also apparent amongst the helix–loop–helix domains of the myogenesis regulatory proteins MyoD1, myogenin, and Myf-5 (77–88% identity), and the products of four genes of the *achaete-scute* complex required for neural development (60–70% identity). These similarities contrast with the modest sequence homologies shared by unrelated helix–loop–helix proteins; e.g. the corresponding domains in c-Myc and MyoD1 exhibit just 29% sequence identity (Davis *et al.*, 1987). It is noteworthy, therefore, that the helix–loop–helix domain of *tal* is especially related to that encoded by *lyl*-1, a gene on chromosome 19 that was identified on the basis of chromosomal rearrangement in a human T cell leukemia (Mellentin *et al.*, 1989). The amino acid alignment in Figure 10 illustrates the 84% level of sequence identity between the homology domains encoded by *tal* and *lyl*-1. Hence, these genes specify a family of helix–loop–helix proteins that can potentially contribute to the formation of human T-ALL.

The normal functions of the *tal* and *lyl*-1 gene products are not known. However, in view of the proposed regulatory interactions between genes encoding helix–loop–helix proteins responsible for neurogenesis (Cabrera *et al.*, 1987) and myogenesis (Pinney *et al.*, 1988; Thayer *et al.*, 1989; Wright *et al.*, 1989), it is conceivable that *tal* and *lyl*-1 are co-ordinately involved in the control of lymphoid development. Moreover, their gene products may function by direct interaction with DNA. Murre *et al.* (1989a) showed that sequence-specific binding of E47 to the immunoglobulin κ-chain enhancer is mediated by its helix–loop–helix domain. They further argued that other polypeptides bearing this motif also function as sequence-specific DNA-binding proteins. Indeed, recent studies reveal that the helix–loop–helix domain is required for specific binding of MyoD1 to the muscle creatine kinase enhancer (Lassar *et al.*, 1989). Likewise, the normal and malignant activities of *tal* may also be mediated by specific recognition of transcriptional regulatory sequences.

The E47 protein binds the immunoglobulin κ-chain enhancer as a dimer (Murre *et al.*, 1989a). Since the

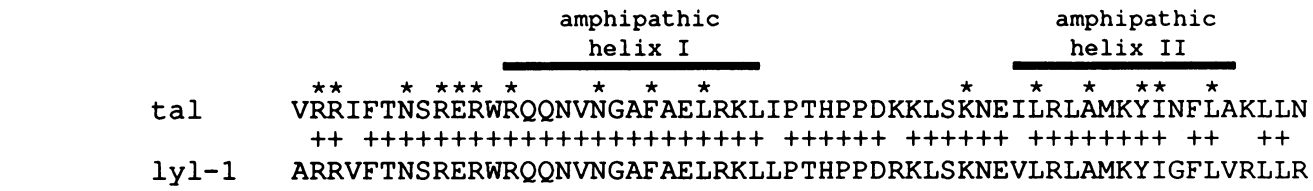

Fig. 10. Alignment of the amphipathic helix−loop−helix domains of *tal* and *lyl*-1. The 16 consensus amino acids residues of the helix−loop−helix motif are marked with asterisks (Mellentin *et al.*, 1989), and the residues that form the proposed helices are overlined with bars (Murre *et al.*, 1989a). The *lyl*-1 sequence was determined by Mellentin *et al.* (1989).

Table I. Concordancy analysis of each human chromosome with B2EE-2.0 in the 17 hybrids of the hybrid clone human mapping panel

Chrom	No. of hybrids[a]				% disc[b]	Chrom	No. of hybrids				% disc
	+/+	+/−	−/+	−/−			+/+	+/−	−/+	−/−	
1	5	0	0	12	0	13	2	6	3	6	53
2	1	4	4	8	47	14	5	4	0	8	24
3	3	3	2	9	29	15	2	2	3	10	29
4	3	4	2	8	35	16	3	4	2	8	35
5	4	6	1	6	41	17	0	2	5	10	41
6	4	3	1	9	24	18	3	5	2	7	41
7	4	2	1	10	18	19	4	8	1	4	53
8	3	9	2	3	65	20	3	3	2	9	29
9	1	6	4	6	59	21	2	8	3	4	65
10	1	5	4	7	53	22	4	6	1	6	41
11	3	3	2	9	29	X	3	3	2	9	29
12	3	5	2	7	41						

[a]+/+ have the chromosome and B2EE, +/− have the chromosome but not B2EE, −/+ do not have the chromosome but have B2EE, −/− have neither the chromosome or B2EE.
[b]% discordant = no. +/− and no. −/+ divided by 17 × 100.

helix−loop−helix domain is necessary and sufficient for dimerization, formation of the E47 homodimer may be facilitated by hydrophobic interactions between the amphipathic helices. Curiously, heterodimers of E47 and other helix−loop−helix proteins—including some that are not expressed in lymphocytes (e.g. MyoD1 and AS-C T3)—not only bind the κ-chain enhancer but do so with higher affinity than the E47 homodimer itself (Murre *et al.*, 1989b). Similarly, some of the functions of the *tal* gene product may be expressed upon heterodimer formation with other helix−loop−helix proteins present in lymphocytes, including E47 or the c-*myc* or *lyl*-1 oncoproteins.

Materials and methods

The patients

Patient 4 was a 7 year old male with T-ALL. At presentation the white blood cell count was 62 700/mm^3 with 51% lymphoblasts. Cytogenetic analysis of bone marrow lymphoblasts revealed an abnormal karyotype: 46,XY,t(1;14)(p32;q11), del(6)(q15q23). The patient entered complete remission following intensive chemotherapy. Patient materials analyzed in this report include a leukemic sample of peripheral blood obtained just prior to treatment, and a remission sample of peripheral blood obtained 4 months later.

Patient 5 was an 8 year old male with T-ALL. The white blood cell count at presentation was 112 400/mm^3 with 41% lymphoblasts. Cytogenetic analysis of lymphoblasts from pleural fluid revealed an abnormal karyotype: 46,XY,t(1;14)(p32;q11). The leukemic sample analyzed in this report comprises cells obtained from the pleural fluid at presentation.

Molecular studies

DNA extracted from patient specimens was analyzed by Southern hybridization with radiolabeled DNA probes (Southern, 1975; Feinberg and Vogelstein, 1983). Genomic DNA libraries of *Bam*HI-digested patient DNAs were constructed in phage vector λ2001 (Karn *et al.*, 1984). A cDNA library of poly(A)-selected RNA from CCRF-CEM cells (Foley *et al.*, 1965) prepared in the phage vector λZAP II (Short *et al.*, 1988) was obtained from Clontech Laboratories, Inc. (Palo Alto, CA). These libraries were screened by the method of Benton and Davis (1977), and restriction fragments of recombinant lambda phage DNA were subcloned into plasmid and M13 phage vectors (Yanisch-Perron *et al.*, 1985; Pridmore, 1987; Short *et al.*, 1988). Nucleotide sequence analyses were performed on M13 single-stranded templates by the chain-terminator method (Sanger *et al.*, 1980).

Somatic cell hybrid analyses

The B2EE-2.0 clone was used as a probe in Southern filter hybridizations with *Eco*RI-digested DNAs extracted from a panel of 17 human/hamster somatic cell hybrids with randomly segregated human chromosomes. The B2EE-2.0 probe hybridized with both human and hamster DNAs, but the resolvable difference in fragment size (human, 2.0 kb; hamster, 7.3 kb) allowed assessment of the presence or absence of human sequence among hybrids of the panel. The hybridization of B2EE-2.0 was perfectly concordant with chromosome 1 and randomly associated (18−65% discordancy) with every other human chromosome (Table I). Many of the ~200 independent human−hamster hybrids generated and analyzed in the laboratory contain broken human chromosomes (Thompson *et al.*, 1987). These were screened for hybrids in which chromosome 1 was disrupted as indicated by the presence of one or more human chromosome 1 markers in the absence of others. A panel of 29 such hybrids was identified and screened for the presence or absence of B2EE-2.0 (as above) in addition to five chromosome 1 markers representative of regions on the p and q arms—PGD at p36, AK2 at p34, PGM1 at p22, AT3 at q23, and PEPC at q24 or q42 (assignments from HGM9, Morton and Bruns, 1988). The lowest levels of discordance were between B2EE-2.0 and the p-arm markers—24% with PGD, 24% with AK2, and 17% with PGM1. Discordancy between B2EE-2.0 and the q-arm markers was high (44% for AT3 and 75% for PEPC). The data indicate a chromosome 1p location for B2EE-2.0.

RNA analysis

RNA was prepared from cell lines derived from human lymphoid tumors, including the pre-B cell line SMS-SB (Smith *et al.*, 1981), the T-ALL line, Jurkat (Schneider *et al.*, 1977) and the Burkitt's lymphoma line Ly65 (Taub *et al.*, 1984). RNA was isolated either by guanidinium−cesium chloride (Ullrich *et al.*, 1977) or acid guanidinium thiocyanate phenol−chloroform extraction (Chomcyznski and Sacchi, 1987). Poly(A)$^+$ RNA was selected

by oligo(dT)–cellulose chromatography (Aviv and Leder, 1972). RNA was fractionated by electrophoresis on a 1% agarose gel after denaturation by glyoxal (McMaster and Carmichael, 1977) and transferred to a nylon membrane (Zeta-Probe, Biorad) for hybridization (Thomas, 1980). The relative electrophoretic mobilities of 28S and 18S rRNAs were determined by fractionation of total cellular RNA in adjacent gel lanes, and methylene blue staining of the rRNA species after transfer to a nylon membrane (Maniatis *et al.*, 1982). The chicken cDNA clone pA1 was used to detect β-actin gene transcripts (Cleveland *et al.*, 1980).

Polymerase chain reaction

Amplification of poly(A)$^+$ RNA sequences from Jurkat cells was conducted by polymerase chain reaction (Saiki *et al.*, 1988) with a pair of oligonucleotide primers complementary to *tal* sequences (Figure 7B). The sequences of oligonucleotides 1 and 2 are, respectively, CAATCGAGTGA-AGAGGAGACCTTC and TTGCGGAGCTCGGCAAAGGC. Each reaction consisted of 50 mM KCl, 10 mM Tris–HCl (pH 8.3), 1.5 mM $MgCl_2$, 0.01% gelatin, 1 μM each oligonucleotide primer, 125 μM each deoxynucleotide triphosphate (dNTP), 40 U human placental ribonuclease inhibitor (Promega), 1 μg poly(A)$^+$ RNA, and 7.5 U avian myeloblastosis virus reverse transcriptase (Promega). Reaction volumes were 100 μl. The reactions were incubated at 42°C for 2 h to allow synthesis of complementary DNA. After addition of 0.25 U of *Taq* DNA polymerase (Perkin-Elmer Cetus) each reaction was subjected to 50 cycles of denaturation (1 min at 94°C), primer annealing (3 min at 55°C) and extension (3 min at 72°C). Temperature cycling was carried out in an automated heating/cooling block (Gene Machine, USA/Scientific Plastics). The reactions were fractionated by electrophoresis on a 10% polyacrylamide gel, and the 148 bp amplification product was visualized by ethidium bromide staining. After elution from the gel, the amplification product was phosphorylated with polynucleotide kinase (New England Biolabs) and cloned into the *Sma*I site of M13mp18 for nucleotide sequence analysis. Multiple independent M13 clones of the 148 bp amplification product were analyzed in both orientations, all of which exhibited the nucleotide sequence presented in Figure 9(C).

Acknowledgements

We thank Cary Ying-Chuan Yang and Julia Tsan for technical assistance, Cindy Baselski for preparing the manuscript, Michael Link (Stanford) for immunophenotyping the leukemic samples, and Donald Capra and Philip Tucker for use of computer facilities. We are also grateful to Joseph Sambrook and Eric Humphries for their comments on the manuscript and to Terrence Rabbitts and Woodring Wright for their advice. This work was supported by the National Cancer Institute (CA46593) and the American Cancer Society (CH-414). G.B., A.C. and W.C. are supported in part by National Institutes of Health grants CA21765 (CORE), CA-33625 and CA-30969, and by the American Lebanese Syrian Associated Charities (W.C. only). R.B. is a member of the Simmons Arthritis Research Center. Q.C., J.-T.C. and L.-H.T. are students in the Microbiology Graduate Program at UT Southwestern Graduate School.

References

Alonso,M.C. and Cabrera,C.V. (1988) *EMBO J.*, **7**, 2585–2591.

Alt,F.W., DePinho,R.A., Zimmerman,K., Legouy,E., Hutton,K., Ferrier,P., Tesfaye,A., Yancopoulos,G.D. and Nisen,P. (1986) *Cold Spring Harbor Symp. Quant. Biol.*, **51**, 931–941.

Aviv,H. and Leder,P. (1972) *Proc. Natl. Acad. Sci. USA*, **69**, 1408–1412.

Baer,R., Chen,K.-C., Smith,S.D. and Rabbitts,T.H. (1985) *Cell*, **43**, 705–713.

Baer,R., Boehm,T., Yssel,H., Spits,H. and Rabbitts,T.H. (1988) *EMBO J.*, **7**, 1661–1668.

Begley,C.G., Aplan,P.D., Davey,M.P., Nakahara,K., Tchorz,K., Kurtzberg,J., Hershfield,M.S., Haynes,B.F., Cohen,D.I., Waldmann,T.A. and Kirsch,I.R. (1989) *Proc. Natl. Acad. Sci. USA*, **86**, 2031–2035.

Benton,W.D. and Davis,R.W. (1977) *Science*, **196**, 180–182.

Boehm,T., Baer,R., Lavenir,I., Forster,A., Waters,J.J., Nacheva,E. and Rabbitts,T.H. (1988a) *EMBO J.*, **7**, 385–394.

Boehm,T., Buluwela,L., Williams,D., White,L. and Rabbitts,T.H. (1988b) *EMBO J.*, **7**, 2011–2017.

Braun,T., Buschhausen-Denker,G., Bober,E., Tannich,E. and Arnold,H.H. (1989) *EMBO J.*, **8**, 701–709.

Cabrera,C.V., Martinez-Arias,A. and Bate,M. (1987) *Cell*, **50**, 425–433.

Caudy,M., Vässin,H., Brand,M., Tuma,R., Jan,L.Y. and Jan,Y.N. (1988) *Cell*, **55**, 1061–1067.

Chomcyznski,P. and Sacchi,N. (1987) *Anal. Biochem.*, **162**, 156–159.

Cleveland,D.W., Lopata,M.A., MacDonald,R.J., Cowan,N.J., Rutter,W.J. and Kirschner,M.W. (1980) *Cell*, **20**, 95–105.

Davis,R.L., Weintraub,H. and Lassar,A.B. (1987) *Cell*, **51**, 987–1000.

Dubé,I.D., Raimondi,S.C., Pi,D. and Kalousek,D.K. (1986) *Blood*, **67**, 1181–1184.

Edmondson,D.G. and Olson,E.N. (1989) *Genes Dev.*, **3**, 628–640.

Feinberg,A.P. and Vogelstein,B. (1983) *Anal. Biochem.*, **132**, 6–13.

Finger,L.R., Harvey,R.C., Moore,R.C.A., Showe,L.C. and Croce,C.M. (1986) *Science*, **234**, 982–985.

Finger,L.R., Kagan,J., Christopher,G., Kurtzberg,J., Hershfield,M.S., Nowell,P.C. and Croce,C.M. (1989) *Proc. Natl. Acad. Sci. USA*, **86**, 5039–5043.

Foley,G.E., Lazarus,H., Farber,S., Uzman,B.G., Boone,B.A. and McCarthy,R.E. (1965) *Cancer*, **18**, 522–529.

Hecht,F., Morgan,R., Hecht,B.K.-M. and Smith,S.D. (1984) *Science*, **226**, 1445–1447.

Hershfield,M.S., Kurtzberg,J., Harden,E., Moore,J.O., Whang-Peng,J. and Haynes,B.F. (1987) *Proc. Natl. Acad. Sci. USA*, **81**, 253–257.

Kagan,J., Finan,J., Letofsky,J., Besa,E.C., Nowell,P.C. and Croce,C.M. (1987) *Proc. Natl. Acad. Sci. USA*, **84**, 4543–4546.

Karn,J., Matthes,H.W.D., Gait,M.J. and Brenner,S. (1984) *Gene*, **32**, 217–224.

Klaembt,C., Knust,E., Tietze,K. and Campos-Ortega,J.A. (1989) *EMBO J.*, **8**, 203–211.

Kronenberg,M., Siu,G., Hood,L. and Shastri,N. (1986) *Annu. Rev. Immunol.*, **4**, 529–591.

Kurtzberg,J., Bigner,S.H. and Hershfield,M.S. (1985) *J. Exp. Med.*, **162**, 1561–1578.

Lampert,F., Harbott,J., Ritterbach,J., Ludwig,W.-D., Fonatsch,C., Schwamborn,D., Stier,B., Gnekow,A., Gerein,V., Stollmann,B., Jobke,A. and Janka-Schaub,G. (1988) *Blut*, **56**, 117–123.

Lassar,A.B., Buskin,J.N., Lockshon,D., Davis,R.L., Apone,S., Hauschka,S.D. and Weintraub,H. (1989) *Cell*, **58**, 823–831.

Lipman,D.J. and Pearson,W.R. (1985) *Science*, **227**, 1435–1441.

Maniatis,T., Fritsch,E.F. and Sambrook,J. (1982) *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Matthieu-Mahul,D., Bernheim,A., Sigaux,F., Daniel,M.-T., Larsen,C.-J. and Berger,R. (1986) *C.R. Acad.Sci. Paris*, **302**, 525–528.

McKeithan,T.W., Shima,E.A., Le Beau,M.M., Minowada,J., Rowley,J.D. and Diaz,M.O. (1986) *Proc. Natl. Acad. Sci. USA*, **83**, 6636–6640.

McMaster,G.K. and Carmichael,G.G. (1977) *Proc. Natl. Acad. Sci. USA*, **74**, 4835–4838.

Mellentin,J.D., Smith,S.D. and Cleary,M.L. (1989) *Cell*, **58**, 77–83.

Morton,N.E. and Bruns,G.A. (1987) *Cytogenet. Cell Genet.*, **46**, 102–130.

Mount,S.M. (1982) *Nucleic Aids Res.*, **10**, 459–472.

Murre,C., McCaw,P.S. and Baltimore,D. (1989a) *Cell*, **56**, 777–783.

Murre,C., McCaw,P.S., Vaessin,H., Caudy,M., Jan,L.Y., Jan,Y.N., Cabrera,C.V., Buskin,J.N., Hauschka,S.D., Lassar,A.B., Weintraub,H. and Baltimore,D. (1989b) *Cell*, **58**, 537–544.

Pinney,D.F., Pearson-White,S.H., Konieczny,S.F., Latham,K.E. and Emerson,C.P.,Jr (1988) *Cell*, **53**, 781–793.

Pridmore,R.D. (1987) *Gene*, **56**, 309–312.

Raimondi,S.C., Behm,F.G., Roberson,P.K., Pui,C.-H., Rivera,G.K., Murphy,S.B. and Williams,D.L. (1988) *Blood*, **72**, 1560–1566.

Saiki,R.K., Gelfand,D.H., Stoffel,S., Scharf,S.J., Higuchi,R., Horn,G.T., Mullis,K.B. and Erlich,H.A. (1988) *Science*, **239**, 487–491.

Sanger,F., Coulson,A.R., Barrell,B.G., Smith,A.J.H. and Roe,B.. (1980) *J. Mol. Biol.*, **143**, 161–178.

Schneider,U., Schwenk,H.-U. and Bornkamm,G. (1977) *Int. J. Cancer*, **19**, 621–626.

Shima,E.A., Le Beau,M.M., McKeithan,T.W., Minowada,J., Showe,L.C., Mak,T.W., Minden,M.D., Rowley,J.D. and Diaz,M.O. (1986) *Proc. Natl. Acad. Sci. USA*, **83**, 3439–3443.

Short,J.M., Fernandez,J.M., Sorge,J.A. and Huse,W.D. (1988) *Nucleic Acids Res.*, **16**, 7583–7600.

Southern,E. (1975) *J. Mol. Biol.*, **98**, 503–517.

Smith,R.G., Dev,V.G. and Shannon,W.A. (1981) *J. Immunol.*, **126**, 596–602.

Stallings,R.L., Olson,E., Straus,A.W., Thompson,L.H., Bachinski,L.L. and Siciliano,M.J. (1988) *Am. J. Hum. Genet.*, **43**, 144–151.

Taub,R., Moulding,C., Battey,J., Murphy,W., Vasicek,T., Lenoir,G.M. and Leder,P. (1984) *Cell*, **36**, 339–348.

Thayer,M.J., Tapscott,S.J., Davis,R.L., Wright,W.E., Lassar,A.B. and Weintraub,H. (1989) *Cell*, **58**, 241–248.

Thisse,B., Stoetzel,C., Gorostiza-Thisse,C. and Perrin-Schmitt,F. (1988) *EMBO J.*, **7**, 2175–2183.

Thomas,P.S. (1980) *Proc. Natl. Acad. Sci. USA*, **77**, 5201–5205.
Thompson,L.E., Carrano,A.V., Sato,K., Salazar,E.P., White,B.F., Stewart,S.A., Minkler,J.L. and Siciliano,M.J. (1987) *Somat. Cell Mol. Genet.*, **13**, 539–551.
Ullrich,A., Shine,J., Chirgwin,J., Pictet,R., Tischer,E., Rutter,W.J. and Goodman,H.M. (1977) *Science*, **196**, 1313–1319.
Villares,R. and Cabrera,C.V. (1987) *Cell*, **50**, 415–424.
Williams,D.L., Look,A.T., Melvin,S.L., Roberson,P.K., Dahl,G., Flake,T. and Stass,S. (1984) *Cell*, **36**, 101–109.
Wright,W.E., Sassoon,D.A. and Lin,V.K. (1989) *Cell*, **56**, 607–617.
Yanisch-Perron,C., Vieira,J. and Messing,J. (1985) *Gene*, **33**, 103–119.

Received on September 12, 1989; revised on November 14, 1989