The EMBO Journal vol.9 no.10 pp.3343-3351, 1990

Site-specific recombination of the *tal*-1 gene is a common occurrence in human T cell leukemia

Lamorna Brown, Jiin-Tsuey Cheng, Qi Chen, Michael J.Siciliano[2], William Crist[1,3], George Buchanan[3] and Richard Baer

Departments of Microbiology and Pediatrics, University of Texas Southwestern Medical Center, 5323 Harry Hines Boulevard, Dallas, TX 75235, [1]St Jude Children's Research Hospital and University of Tennessee, Memphis College of Medicine, Memphis, TN 38101, [2]Department of Molecular Genetics, University of Texas M.D. Anderson Hospital Cancer Center, Houston, TX 77030 and [3]Pediatric Oncology Group, St Louis, MO 63108, USA

Communicated by T.H.Rabbitts

The *tal*-1 gene is altered as a consequence of the t(1;14)(p32;q11) chromosome translocation observed in 3% of patients with T cell acute lymphoblastic leukemia (T-ALL). *tal*-1 encodes a helix-loop-helix (HLH) domain, a DNA binding and dimerization motif found in a number of proteins involved in cell growth and differentiation. We now report that an additional 25% of T-ALL patients bear *tal*-1 gene rearrangements that are not detected by karyotype analysis. These rearrangements result from a precise 90 kb deletion (designated tal^d) that arises independently in different patients by site-specific DNA recombination. Since the deletion junctions resemble the coding joints of assembled immunoglobulin genes, tal^d rearrangements are likely to be mediated by aberrant activity of the immunoglobulin recombinase. Moreover, t(1;14)(p32;q11) translocations and tal^d rearrangements disrupt the coding potential of *tal*-1 in an equivalent manner, and thereby generate a common genetic lesion shared by a significant proportion of T-ALL patients.

Key words: gene rearrangement/human/site-specific recombination/*tal*-1/T cell leukemia

Introduction

Certain tumors are characterized by unique chromosome abnormalities that are present in a significant proportion of affected patients (Yunis, 1983). Much evidence supports the view that these abnormalities represent genetic events that promote development of the associated neoplasms (Mitelman, 1987). For example, the t(8;14)(q24;q32) translocation is observed in the malignant cells of nearly 90% of patients with Burkitt's lymphoma; at the molecular level, this defect represents the activation of the c-*myc* proto-oncogene by transposition from its normal location on chromosome 8 into the immunoglobulin heavy chain locus on chromosome 14 (reviewed by Leder *et al.*, 1983; Rabbitts, 1985).

By contrast, cytogenetic studies have not uncovered a major karyotypic abnormality associated with T cell acute lymphoblastic leukemia (T-ALL). Instead, a number of minor defects have been reported, many of which feature cytogenetic breakage within the T cell receptor (TCR) genes on chromosomes 7 and 14 (for recent reviews see Boehm and Rabbitts, 1989; Tycko and Sklar, 1990). Individually, these chromosome abnormalities are present in relatively small proportions of T-ALLs; for example, the most prevalent, the t(11;14)(p13;q11) translocation, is only observed in ~7% of T-ALL patients. Nevertheless, each has been implicated as a contributing factor in leukemogenesis on the basis of its recurrence in unrelated patients and its unique association with T-ALL.

The *tal*-1 gene was identified upon analysis of t(1;14)(p32;q11), a chromosome translocation observed in only 3% of T-ALL patients (Begley *et al.*, 1989a; Finger *et al.*, 1989; Bernard *et al.*, 1990; Chen *et al.*, 1990). As a consequence of the translocation, *tal*-1 is transposed from its normal location on chromosome 1 into the TCR α/δ chain locus on chromosome 14. The amino acid translation of *tal*-1 includes sequence homologous to the helix-loop-helix (HLH) motif (Begley *et al.*, 1989; Chen *et al.*, 1990), a DNA binding domain present in a family of proteins involved in the control of cell growth and differentiation (Villares and Cabrera, 1987; Murre *et al.*, 1989a). The homology domains of several HLH proteins have been shown to specifically bind core sequences of transcription enhancer elements, including those associated with the immunoglobulin κ chain locus and the muscle creatine kinase gene (Lassar *et al.*, 1989; Murre *et al.*, 1989a,b; Henthorn *et al.*, 1990). Thus, proteins of the HLH family may serve as transcriptional regulatory factors, a function that accords well with their profound influence over both normal and malignant development.

Several genes that encode HLH proteins have been implicated in human leukemia, including the c-*myc*, *lyl*-1 and *E2A* proto-oncogenes (Mellentin 1989a,b; Kamps *et al.*, 1990; Nourse *et al.*, 1990). Therefore, it is tempting to propose that the t(1;14)(p32;q11) translocations alter *tal*-1 gene expression in a manner that promotes T-ALL formation. However, t(1;14)(p32;q11) is a rare marker of T-ALL, detectable in only 3% of patients with the disease (Carroll *et al.*, 1990). Nevertheless, in this report we demonstrate rearrangements of the *tal*-1 gene in 25% of T-ALL patients, including those without apparent cytogenetic aberrations of chromosome 1. Surprisingly, the rearrangements observed in different patients are identical, i.e. they all arose from a precise 90 kb deletion that disrupts the coding region of *tal*-1 in a manner analogous to the t(1;14)(p32;q11) translocation. The extraordinary precision of these deletions (designated tal^d) suggests that they are mediated by a site-specific DNA recombinase. Moreover, analysis of the deletion junctions indicates that tal^d rearrangement is engendered by aberrant activity of the same recombinase that controls immunoglobulin and TCR gene assembly.

Results

The tal-1 gene is altered in a high proportion of T-ALL patients

The *tal*-1 gene is comprised of multiple exons with amino acid coding potential (Figure 1A). Two alternatively spliced mRNA species have been identified by cDNA analysis, one of which includes exons Ia-II-III (type A) and the other exons Ib-II-III (type B). In three out of four patients with t(1;14)(p32;q11) translocations, the chromosome 1 breakage occurred within a 1 kb region located just downstream of exon Ia (Figure 1A). Thus, as a consequence of t(1;14)(p32;q11), the structure of *tal*-1 is often disrupted in a manner that precludes production of type A mRNA from the translocated allele.

Since t(1;14)(p32;q11) is only observed in 3% of T-ALL patients, we sought to determine whether *tal*-1 is also altered in patients without obvious karyotypic aberrations of chromosome 1. Therefore, we examined leukemic DNAs derived from either established T-ALL cell lines or fresh T-ALL specimens. Figure 1 shows Southern analyses of *Bam*HI digested DNAs hybridized with B2EE-2.0, a probe representing sequences from the translocation breakpoint region of *tal*-1 (Figure 1A). DNAs derived from eight T-ALL cell lines were examined and, as illustrated in Figure 1B, four of these displayed a rearranged 4.5 kb *Bam*HI fragment in addition to the normal 5.0 kb fragment. The *tal*-1 DNA rearrangement was also observed in fresh leukemic cells from four out of eight T-ALL patients (Figure 1C).

The tal^d rearrangement: a common alteration of the tal-1 gene in ~25% of T-ALL patients

It is intriguing that all eight patients with *tal*-1 gene alterations had rearranged *Bam*HI fragments with identical electrophoretic mobilities (Figure 1). This phenomenon was investigated further by Southern analyses of T-ALL DNAs digested with six additional restriction endonucleases. Figure 2A shows the hybridization pattern obtained for DNA from the T-ALL cell line RPMI8402; in each digest, the B2EE-2.0 probe detected an equimolar ratio of the normal DNA fragment and a rearranged DNA fragment, indicating that one allele of *tal*-1 had undergone a structural rearrangement in RPMI8402 cells. Figure 2B shows the Southern analysis of DNA from CCRF-CEM, a leukemic cell line derived from an unrelated T-ALL patient; again, the B2EE-2.0 probe detects DNA rearrangement of one of the two alleles of *tal*-1. Surprisingly, however, the rearranged DNA fragment in each restriction digest of CCRF-CEM DNA is similar in size to that observed in RPMI8402 DNA, indicating that both cell lines bear an identical rearrangement of the *tal*-1 locus. Furthermore, an identical pattern of rearranged DNA fragments was obtained upon analysis of each of the other T-ALL samples with *tal*-1 gene alterations (data not shown). Therefore, a high proportion of T-ALL patients exhibit a common rearrangement of the *tal*-1 locus (designated tal^d) that, at least at the level of Southern analysis, appears to be the same in each patient.

To obtain a more accurate estimate of the frequency of *tal*-1 gene alteration, we examined 50 random T-ALL patients from the Pediatric Oncology Group, 13 of which (26%) harbored tal^d rearrangement (W.Crist and R.Baer, unpublished results). Thus, alteration of the *tal*-1 gene, either by tal^d rearrangement or t(1;14)(p32;q11) translocation, appears to be a prominent feature of T-ALL.

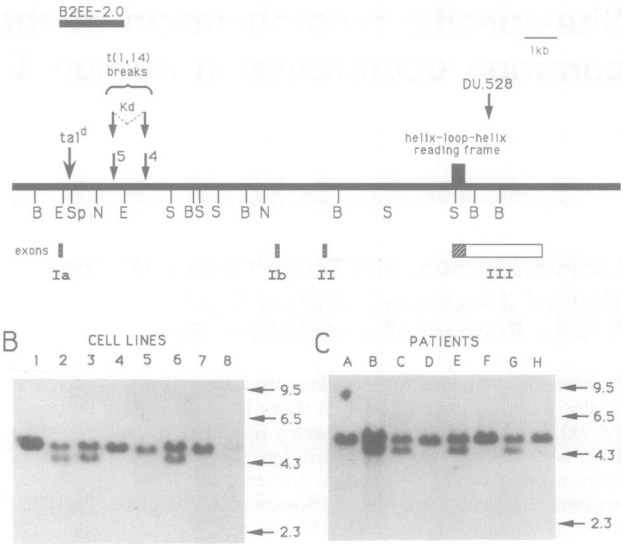

Fig. 1. Rearrangement of the *tal*-1 gene in T-ALL cells. (**A**) A restriction map of the *tal*-1 gene in its normal configuration. The small arrows designate the site of chromosome breakage due to t(1;14)(p32;q11) translocations from patients 4 and 5 (Chen *et al.*, 1990), DU.528 (Begley *et al.*, 1989b; Finger *et al.*, 1989) and Kd (Bernard *et al.*, 1990); the translocation in patient Kd was accompanied by loss of sequences between the two arrows (Bernard *et al.*, 1990). The large arrow indicates the downstream endpoint of the tal^d deletion. The exons of *tal*-1 were localized by comparing sequences of *tal*-1 cDNA clones (Begley *et al.*, 1989b; Chen *et al.*, 1990) and genomic DNA (Q.Chen and R.Baer, unpublished results); the stippled regions of the *tal*-1 exons denote coding sequences and the open region denotes 3′ non-coding sequences. Restriction sites: B, *Bam*HI; E, *Eco*RI; N, *Not*I; S, *Sac*I; Sp, *Sph*I (*Sph*I sites are not complete). (**B**) Rearrangement of the *tal*-1 gene in T-ALL cell lines. A Southern filter of *Bam*HI digested DNAs was hybridized with *tal*-1 probe B2EE-2.0. The DNAs were derived from T-ALL cell lines Jurkat (lane 1), RPMI8402 (2), CCRF-CEM (3), MOLT-3 (4), MOLT-13 (5), MOLT-16 (6), PEER (7) and CCRF-HSB-2 (8). The sizes of *Hin*dIII λ DNA marker fragments are indicated in kb. (**C**) Rearrangement of the *tal*-1 gene in primary T-ALL cells. A Southern filter of *Bam*HI digested DNAs was hybridized with B2EE-2.0. The DNAs were derived from peripheral blood obtained from T-ALL patients before treatment.

The tal^d rearrangement is tumor-specific

Comparative Southern analyses of normal and leukemic cells from the same patient indicate that tal^d rearrangement is tumor-specific (Figure 3). For example, tal^d is apparent in T lymphoblastoid lines from two T-ALL patients (RPMI8402 and CCRF-HSB-2), but not in non-leukemic B cell lines derived from the same patients (RPMI8392 and CCRF-SB, respectively) (Hayata *et al.*, 1975). Similar results were obtained upon analysis of fresh specimens from T-ALL patients. Three such patients were treated at the Children's Medical Center of Dallas, where blood specimens were obtained before chemotherapy and after complete remission; in each case the leukemic sample bore the tal^d rearrangement and the remission sample did not (Figure 3). Therefore, tal^d is not a genetic polymorphism of the *tal*-1 gene, but instead represents an acquired alteration that appears to be restricted to the leukemic cells of T-ALL patients.

The tal^d rearrangement is generated by local DNA recombination

To investigate the nature of the tal^d rearrangement, a bacteriophage λ library of genomic DNA from RPMI8402

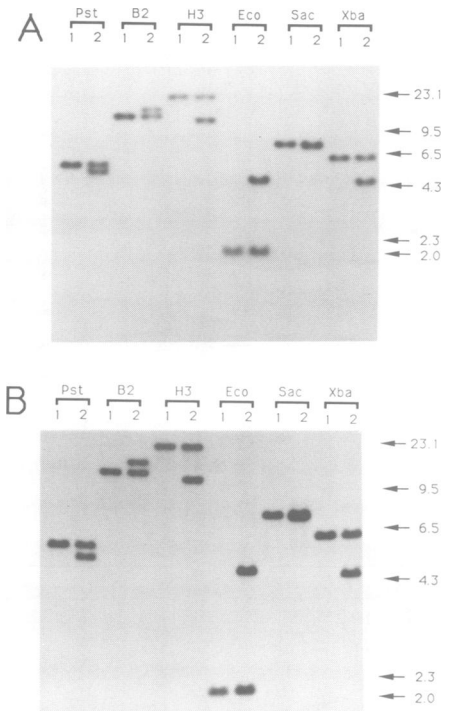

Fig. 2. *tal*-1 gene rearrangements in cell lines from unrelated T-ALL patients are identical. (**A**) Southern analysis of genomic DNAs digested with any of six different restriction endonucleases and hybridized with the B2EE-2.0 probe. Genomic DNAs were derived from the non-leukemic B cell line RPMI8392 (lanes 1) and T-ALL cell line RPMI8402 (lanes 2). (**B**) Southern analysis of DNAs hybridized with the B2EE-2.0 probe. DNAs were derived from non-leukemic control cells (lanes 1) and the T-ALL cell line CCRF-CEM (lanes 2).

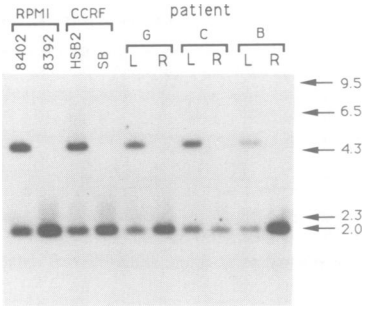

Fig. 3. *tal*^{d} rearrangements are tumor-specific. Southern hybridization analysis of *Eco*RI digested DNAs hybridized with the B2EE-2.0 *tal*-1 probe. *tal*^{d} gene rearrangements are seen in DNAs of T-ALL cell lines (RPMI8402 and CCRF-HSB-2) but not non-leukemic B cell lines (RPMI8392 and CCRF-SB respectively) from the same patients. Similarly, *tal*^{d} rearrangements can be detected in DNA from peripheral blood obtained from T-ALL patients before treatment (lanes L) but not after remission induction (lanes R).

cells was screened with B2EE-2.0, and several clones with inserts spanning the rearrangement were obtained (e.g. λBL1 and λBL3; Figure 4). A restriction map encompassing the *tal*^{d} rearrangement was compiled by analysis of these clones, and in Figure 4 this map is compared to that of the normal *tal*-1 locus. As illustrated, the maps diverge within a 0.25 kb *Eco*RI−*Sph*I fragment of the normal *tal*-1 locus, at a position ~1 kb upstream of the t(1;14)(p32;q11) breakpoint region. As a consequence of the rearrangement, novel DNA sequences are juxtaposed with the *tal*-1 locus

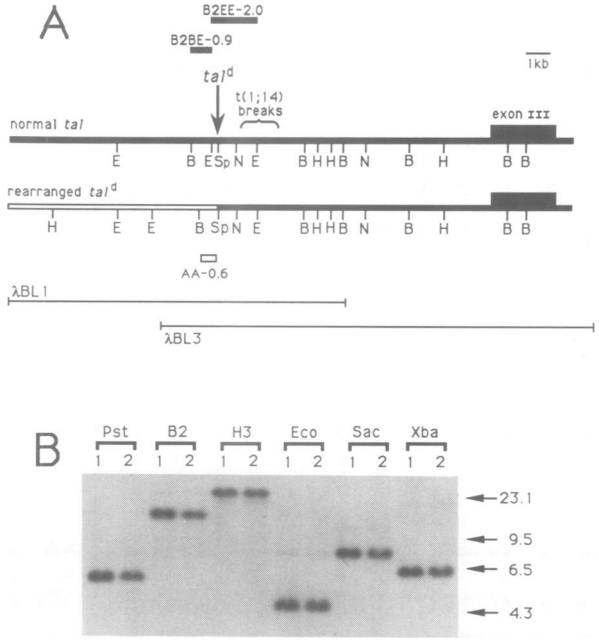

Fig. 4. The structure of the *tal*^{d} rearrangement. (**A**) The λBL1 and λBL3 clones were isolated by screening a library of genomic DNA from the *tal*^{d} positive RPMI8402 cell line with probe B2EE-2.0. The restriction map of the rearranged *tal*^{d} allele was compiled from those of λBL1 and λBL3, and is compared to that of the normal *tal*-1 locus. The downstream endpoint of the *tal*^{d} deletion is denoted by a large arrow. The B2BE-9.0 and B2EE-2.0 probes are derived from normal *tal*-1 sequences, and the AA-0.6 probe is derived from novel sequences engendered by the *tal*^{d} rearrangement. The major breakpoint region of t(1;14)(p32;q11) is bracketed and the position of the HLH-encoding exon III is indicated. Restriction sites are marked as in Figure 1A. (**B**) A Southern filter identical to that shown in Figure 2A was hybridized with the B2BE-9.0 probe. Genomic DNAs were derived from the non-leukemic B cell line RPMI8392 (lanes 1) and the T-ALL cell line RPMI8402 (lanes 2).

(Figure 4). These sequences are likely to be derived from the same chromosome as *tal*-1 since RPMI8402 cells do not have karyotypic defects involving the short arm of chromosome 1 (Le Beau *et al.*, 1986). To study this issue, a 0.6 kb DNA fragment (AA-0.6) derived from the novel sequences was used as a probe in Southern hybridizations with DNAs from a panel of human−hamster somatic cell hybrids with randomly segregated human chromosomes (Thompson *et al.*, 1987). The hybridization of AA-0.6 was perfectly concordant with chromosome 1 and randomly associated (18−65% discordance) with every other human chromosome. The regional localization of AA-0.6 was determined by analysis with a panel of hybrids containing broken derivatives of chromosome 1 (Stallings *et al.*, 1988). Since the AA-0.6 probe showed low discordance with short arm markers (17−24%) and high discordance with long arm markers (44−75%), it is likely to be derived from the short arm of chromosome 1. Moreover, the same hybrid panel had been analyzed previously with the B2EE-2.0 probe (Chen *et al.*, 1990), and comparison of the data reveals perfect concordance with AA-0.6. This implies close linkage between the AA-0.6 and B2EE-2.0 sequences. Hence, the *tal*^{d} rearrangement represents recombination of local DNA sequences on the short arm of chromosome 1.

The origin of AA-0.6 was further analyzed by isolating corresponding unrearranged genomic DNA from SUP-T1, a cell line that does not bear the *tal*^{d} rearrangement

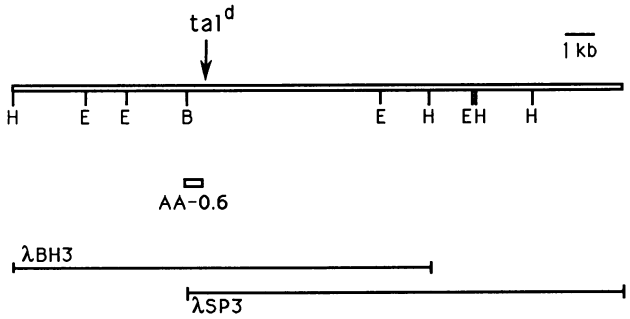

Fig. 5. The normal structure of the AA-0.6 locus. The λBH3 and λSP3 clones were obtained by screening a library of genomic DNA from the *tal*^{d} negative cell line SUP-T1 with probe AA-0.6. The restriction map of the unrearranged AA-0.6 locus was compiled from those of λBH3 and λSP3. The upstream endpoint of the *tal*^{d} deletion is denoted with an arrow. Restriction sites are marked as in Figure 1A.

(unpublished data). Hence, a λ library of SUP-T1 DNA was screened with AA-0.6, and two hybridizing clones were identified (λSP3 and λBH3). Restriction analysis of these clones generated a 22 kb map of the normal genomic DNA encompassing AA-0.6 (Figure 5).

The tal^d *rearrangement represents an ~90 kb DNA deletion*

The *tal*^{d} rearrangement might conceivably arise by any of a number of distinct processes, including local DNA inversion, duplication, insertion or deletion. To evaluate these possibilities, we hybridized both normal and leukemic DNAs with B2BE-0.9, a probe located immediately upstream of B2EE-2.0 in normal DNA, but on the opposite flank of the rearrangement site (see Figure 4A). If *tal*^{d} rearrangement occurs without loss of genetic material, then Southern analyses with B2BE-0.9 should reveal rearranged DNA fragments upon digestion of RPMI8402 DNA with restriction enzymes that recognize sites flanking both B2BE-0.9 and B2EE-2.0 (e.g. *Bam*HI, *Hin*dIII, *Bgl*II, *Pst*I). Nevertheless, as shown in Figure 4B, only normal DNA fragments are detected with B2BE-0.9. Hence, one allele of B2BE-0.9 has been lost from the genome of RPMI8402. B2BE-0.9 exhibits the same pattern of hybridization with genomic DNA from each of the other T-ALL samples that bears *tal*^{d} (data not shown). These results are consistent with a model in which *tal*^{d} is generated by local DNA deletion, with concomitant loss of sequences between the two recombining elements (represented by AA-0.6 and B2EE-2.0).

Although AA-0.6 and B2EE-2.0 originate from the same region of chromosome 1, the restriction maps of normal genomic DNA around these markers do not overlap. Indeed, direct comparison of these maps indicates that AA-0.6 and B2EE-2.0 are separated by at least 35 kb (Figures 1A and 5). Consequently, if *tal*^{d} is generated by simple deletion, then the segment of DNA deleted is likely to be substantial. To evaluate the linkage between AA-0.6 and B2EE-2.0, *Not*I-digested DNAs from RPMI8392 and RPMI8402 cells were fractionated by transverse alternating field electrophoresis, and analyzed by Southern hybridization (Figure 6A). Interestingly, the AA-0.6 and B2EE-2.0 probes co-hybridized with *Not*I fragments of 220 and 190 kb in DNA from RPMI8392 (Figure 6A, lanes 1 and 3), a B-lymphoblastoid line that does not bear the *tal*^{d} rearrangement. Southern analyses of *Not*I-digested DNAs from other

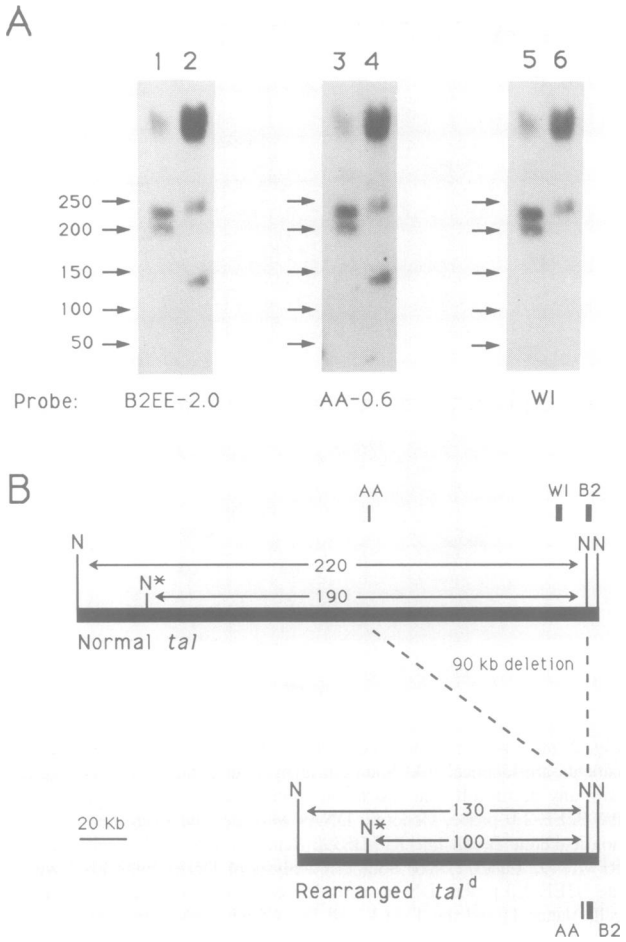

Fig. 6. *tal*^{d} rearrangement generates a 90 kb deletion. (**A**) *Not*I digested genomic DNAs from the non-leukemic B cell line RPMI8392 (lanes 1, 3 and 5) and the T-ALL cell line RPMI8402 (lanes 2, 4 and 6) were fractionated by pulsed-field gel electrophoresis and transferred to a membrane filter. The filter was hybridized, stripped of radioactivity and rehybridized successively with probes B2EE-2.0 (Figure 4A), AA-0.6 (Figure 4A) and W1 (Figure 6B). Size markers are concatemers of λ DNA spaced at ~50 kb intervals. (**B**) The *Not*I restriction map of the normal and rearranged alleles of *tal*-1. The positions of the AA-0.6 (AA), WI and B2EE-2.0 (B2) probes are indicated. The WI probe represents sequences located 12 kb upstream of B2EE-2.0 in the normal *tal*-1 allele (unpublished data). The sizes of *Not*I restriction fragments that co-hybridized with AA-0.6 and B2EE-2.0 are indicated in kb. The asterisk denotes the *Not*I site that exhibits variable resistance to *Not*I digestion, presumably due to cytosine methylation.

cell lines without *tal*^{d} exhibit a variable pattern in which the AA-0.6 and B2EE-2.0 probes co-hybridize with either two fragments of 220 and 190 kb or a single fragment of 220 kb (data not shown). Therefore, in normal DNA, the AA-0.6 and B2EE-2.0 markers reside within a 190 kb *Not*I fragment; the 220 kb species is likely to arise in certain cell lines as a consequence of partial cytosine methylation at one of the flanking *Not*I sites. Restriction analysis of cloned DNA encompassing the *tal*-1 locus identified several *Not*I sites, one of which lies within B2EE-2.0 (Figure 1A). Hence, the B2EE-2.0 probe overlaps neighboring *Not*I fragments of 190 kb and 5.3 kb in normal DNA. *tal*-1 gene probes located downstream of the 5.2 kb *Not*I fragment hybridize to a *Not*I fragment of >500 kb (data not shown); therefore, the 220 kb species detected with B2EE-2.0 is likely to be

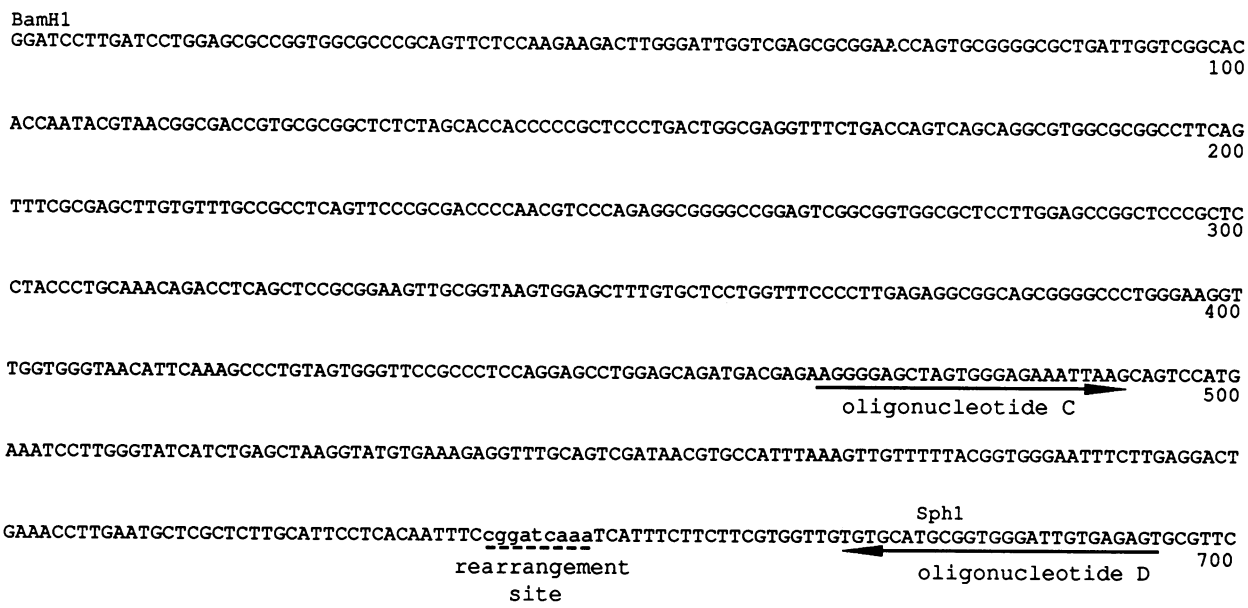

Fig. 7. Nucleotide sequence encompassing the *tal*^{d} rearrangement of RPMI8402. The sequence includes the 0.7 kb *Bam*HI−*Sph*I fragment of the *tal*^{d} allele of RPMI8402 cells. The rearrangement site was determined by comparative analysis of normal and rearranged *tal*-1 sequences (see Figure 8). The positions of synthetic oligonucleotide primers are indicated; the oligonucleotide C sequence is as shown, whereas the oligonucleotide D sequence is the reverse complement of that shown.

generated due to partial methylation of the upstream (rather than the downstream) *Not*I site of the 190 kb fragment, as shown schematically in Figure 6B.

The relative position of AA-0.6 within the 190 kb fragment can be deduced by Southern analysis of *Not*I-digested DNA from cells bearing *tal*^{d}. For example, in RPMI8402 cells, the *tal*^{d} rearrangement generates a 130 kb *Not*I fragment that co-hybridizes with the AA-0.6 and B2EE-2.0 probes (Figure 6A, lanes 2 and 4). In other *tal*^{d} positive cell lines, these probes co-hybridize with either a single rearranged 130 kb fragment (e.g. CCRF-CEM) or with two rearranged fragments of 130 and 100 kb (e.g. CCRF-HSB-2) (data not shown). Again, these patterns are compatible with variable methylation at the distal *Not*I site of the smaller fragment (Figure 6B). AA-0.6 must be located near one end of this fragment since, as a result of *tal*^{d} rearrangement, AA-0.6 is closely juxtaposed with B2EE-2.0 sequences that contain a *Not*I recognition site. As depicted schematically in Figure 6B, this in turn localizes AA-0.6 to a position ~90 kb upstream of B2EE-2.0 in normal genomic DNA. It is noteworthy that *tal*-1 DNA probes located between AA-0.6 and B2EE-2.0 (e.g. WI; Figure 6B) hybridize with the normal 220/190 kb *Not*I fragment(s) but not with the rearranged 130/100 kb species (Figure 6A, lanes 5 and 6). This provides further support for a deletional model in which *tal*^{d} arises by site-specific recombination between AA-0.6 and B2EE-2.0, accompanied by loss of the 90 kb of intervening sequence.

The tal^{d} *rearrangement is site-specific*

In view of its substantial size, it is surprising that the *tal*^{d} deletion is indistinguishable—at least at the level of Southern analysis—in different T-ALL patients. As shown in Figure 4, the deletion junction from RPMI8402 cells can be localized to a 0.7 kb *Bam*HI−*Sph*I fragment, the sequence of which is presented in Figure 7. We also determined the nucleotide sequences of corresponding germline DNA in the vicinity of AA-0.6 and B2EE-2.0, and in Figure 8 these sequences are aligned so as to illustrate the deletion junction. Interestingly, the RPMI8402 junction contains a stretch of nine nucleotides which are not derived from germline sequences in the region of either AA-0.6 or B2EE-2.0.

Deletion junctions of *tal*^{d} were isolated from additional patients by the polymerase chain reaction. Hence, oligonucleotide primers that flank the RPMI8402 junction (i.e. oligos C and D; Figure 8) were used to amplify genomic DNAs from various sources, including leukemic cells from 21 T-ALL patients with *tal*^{d}. In this manner, a discrete amplification product of ~220 bp was generated from each of the *tal*^{d} positive DNAs, but not from genomic DNAs without *tal*^{d}. Nucleotide sequence analyses confirmed that the amplified product from each patient represents the deletion junction of *tal*^{d}; thus, as shown in Figure 9, each product is comprised of AA-0.6 sequences juxtaposed with B2EE-2.0 sequences in a fashion similar to that of the RPMI8402 junction. Nevertheless, the deletion junction from each patient is unique due to sequence variation at the recombination site (Figure 9). As discussed below, the junctional diversity generated by the *tal*^{d} rearrangement is reminiscent of that engendered during site-specific rearrangement of the immunoglobulin and TCR genes.

Discussion

Site-specific DNA recombination as a means of oncogene alteration

The structural defects that activate proto-oncogenes are of various types, including point mutation, retroviral insertion, gene amplification and chromosome translocation (Klein and Klein, 1985). The *tal*-1 gene was originally implicated in oncogenesis in studies of t(1;14)(p32;q11), a chromosome translocation uniquely associated with T-ALL (Begley *et al.*, 1989a; Finger *et al.*, 1989; Bernard *et al.*, 1990; Chen *et al.*, 1990). As a consequence of the translocation, *tal*-1 is transposed from its normal position on chromosome 1 into the TCR α/δ chain locus on chromosome 14, whereupon

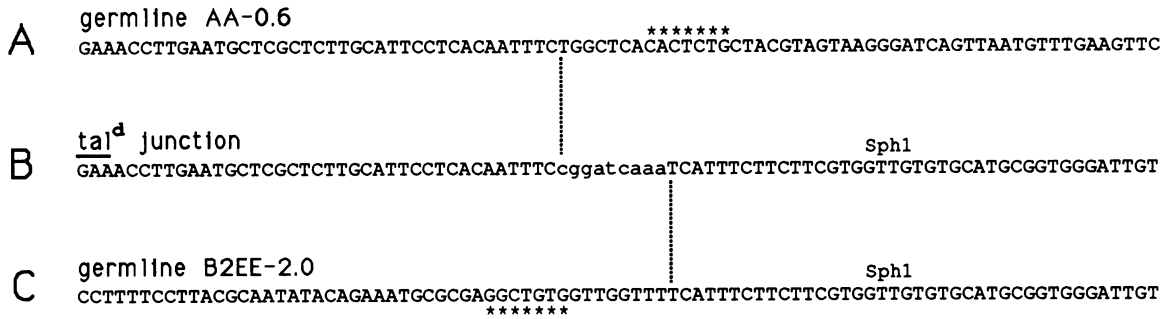

Fig. 8. The deletion junction of *tal*^{d} from RPMI8402 cells. The *tal*^{d} deletion junction of RPMI8402 (**B**) is identified by comparison with germline sequences of the AA-0.6 region (**A**) and the B2EE-2.0 region (**C**). The nine nucleotide residues at the junction (in lower case letters) are not derived from germline sequences of AA-0.6 or B2EE-2.0. Heptamer sequences of the putative recombination signals are marked with asterisks.

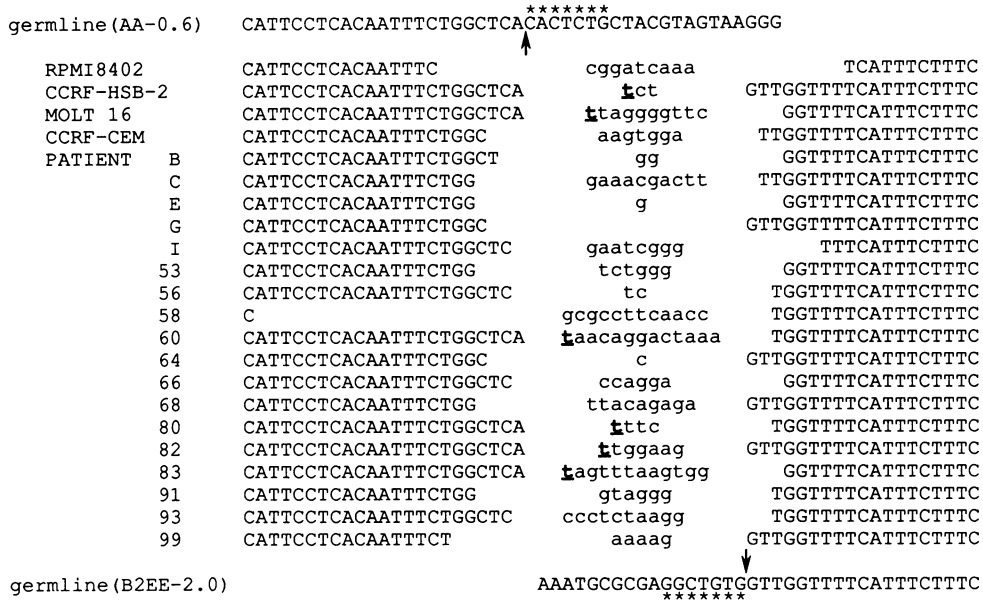

Fig. 9. *tal*^{d} deletion junctions resemble the coding joints of assembled immunoglobulin genes. The *tal*^{d} junctions of three additional *tal*^{d} positive cell lines and 18 *tal*^{d} positive primary T-ALL specimens were isolated by PCR amplification using oligonucleotides C and D (Figure 7). The *tal*^{d} junctions are aligned with germline sequences from AA-0.6 and B2EE-2.0. Heptamer sequences of the putative recombination signals within AA-0.6 and B2EE-2.0 are marked with asterisks, and the proposed sites of recombination are denoted with arrows. Junctional nucleotides in lower case letters are not derived from the germline sequences of AA-0.6 or B2EE-2.0. The underlined thymidine residues at the AA-0.6 junctions of CCRF-HSB-2, MOLT16, patients 60, 80, 82 and 83 are proposed to be P nucleotides (Lafaille *et al.*, 1990).

its expression is presumably altered in a manner that promotes leukemogenesis. Since t(1;14)(p32;q11) is observed in only 3% of T-ALL patients (Carroll *et al.*, 1990), malignant activation of *tal*-1 by chromosome translocation is clearly an uncommon pathway towards the formation of T-ALL. Here we report that ~25% of T-ALL patients harbor a tumor-specific deletion of ~90 kb from one allele of the *tal*-1 locus. Since the deletion (designated *tal*^{d}) is found in tumors without detectable cytogenetic lesions of chromosome 1, it is apparently beyond the current resolution of karyotype analysis.

Structural deletions of the *rb*-1 locus are routinely observed in the malignant tissues of patients with retinoblastoma (Dryja *et al.*, 1986). Nevertheless, these deletions are of variable length and probably serve to inactivate expression of *rb*-1, a gene that promotes malignancy in a recessive fashion by loss of genetic function. In contrast, the *tal*^{d} deletions exhibit nearly identical endpoints in different T-ALL patients. Hence, *tal*^{d} rearrangement is a distinct mechanism of oncogene alteration in which a common genetic lesion is generated independently in different patients by site-specific DNA recombination.

The role of the tal-1 gene in T cell acute leukemia

In this study, *tal*^{d} rearrangement was observed in at least 25% of T-ALL samples, including leukemic specimens obtained directly from T-ALL patients and leukemic cell lines established from these patients. We have not observed similar lesions in other hematopoietic tumors, including pre-B-ALL (28 cases analyzed to date) and the other forms of T cell neoplasia (15 cases). Recently, two groups have shown that *tal*-1 is altered by t(1;14)(p32;q11) translocation in DU.528 (Begley *et al.*, 1989b; Finger *et al.*, 1989), a cell line derived from a rare stem cell leukemia (Kurtzberg *et al.*, 1985). It may be significant, however, that the DU.528 patient presented with T-ALL at the initial diagnosis (Hershfield *et al.*, 1984). Thus, alterations of the *tal*-1 gene, either by *tal*^{d} rearrangement or t(1;14)(p32;q11) translocation, appears to be predominantly, if not exclusively, associated with T-ALL.

Although the initiation sites of *tal*-1 gene transcription are presently unknown, two distinct mRNA species have been identified, both of which encode the HLH domain (Begley *et al.*, 1989b; Chen *et al.*, 1990). These are clearly generated by alternative RNA splicing; hence, type A mRNA includes the three coding exons provisionally designated Ia, II and III, while type B mRNA includes exons Ib, II and III (see Figure 1A). Three of the four t(1;14)(p32;q11) translocations analyzed to date feature breakage within a 1 kb region of the *tal*-1 locus. It is noteworthy that the translocation breakpoint region and the site of tal^d rearrangement both fall between exons Ia and Ib. Thus, tal^d deletion and t(1;14)(p32;q11) translocation are structurally equivalent alterations of *tal*-1 in that each removes exon Ia from the remainder of the locus and thereby precludes the production of type A mRNA. The effect of these lesions on expression of type B mRNA cannot be evaluated until its transcription start site is defined. Nevertheless, the truncation of *tal*-1 observed in T-ALL resembles the structural alterations sustained by genes that encode other HLH proteins implicated in human leukemogenesis. For example, it has been proposed that the malignant potential of c-*myc* and *lyl*-1 can also be unleashed by chromosome translocations that remove or mutate the 5′ coding exons of these genes (Rabbitts *et al.*, 1984; Taub *et al.*, 1984; Pelicci *et al.*, 1986; Hann *et al.*, 1988; Mellentin *et al.*, 1989a).

The mechanism of tal^d rearrangement

A remarkable feature of the tal^d rearrangement is its apparent site-specificity, especially in view of the substantial size of the deletion that it engenders. Two interpretations of this phenomenon seem plausible. The first posits that the *tal*-1 gene undergoes rare DNA rearrangements of a random nature, but only those that resemble tal^d are capable of eliciting its malignant potential. However, this hypothesis does not adequately explain the role of t(1;14)(p32;q11), which clearly generates chromosome 1 breakpoints distinct from the site of tal^d deletion. A second interpretation assumes that *tal*-1 can be malignantly activated by a variety of distinct structural alterations (e.g. those that eliminate exon Ia); however, tal^d emerges in T-ALL patients because sequences near the deletion endpoints are especially prone to recombination. For example, these sequences might resemble the recognition signals of a site-specific recombinase. Such a resemblance could be fortuitous, without normal physiological significance. Nevertheless, when combined with the forces of tumor selection engendered by *tal*-1 gene alteration, even a low level of enzymatically driven recombination should be sufficient to account for the prevalence of tal^d rearrangements in T-ALL patients.

The only site-specific DNA rearrangements observed in vertebrates are those involved in the assembly of the immunoglobulin (Ig) and TCR genes during lymphoid development (reviewed by Tonegawa, 1983; Blackwell and Alt, 1989; Lewis and Gellert, 1989). Since tal^d rearrangements arise in T-lineage cells, these may also be mediated by the same recombination system. Rearrangements within the Ig/TCR loci are directed by signals that flank the rearranging gene segments and presumably serve as recognition sites for the Ig/TCR recombinase. These signals are comprised of a conserved heptamer element that is separated from a conserved nonamer by either 12 or 23 bp of relatively unconserved sequence. As illustrated in Figure 9, sequences bearing resemblance to the consensus heptamer of Ig/TCR recombination signals (CACAGTG) can be found at appropriate positions within germline DNA of both AA-0.6 (CACTCTG; marked by asterisks in Figure 8) and B2EE-2.0 (CACAGCC; marked by asterisks in reverse complement). These heptamers should be relatively inefficient at directing recombination since they are not associated with conserved nonamer elements and they bear sequence deviations from the consensus heptamer that are likely to reduce, but not eliminate, the rate of recombination (Hesse *et al.*, 1989).

The role of the Ig/TCR recombinase in tal^d rearrangement can be further evaluated by examination of tal^d deletion junctions from T-ALL patients. During normal Ig/TCR gene rearrangement, two reciprocal products of recombination are generated: a 'coding joint', which constitutes the fusion of two gene segments of the rearranging locus, and a 'signal joint', comprised of the two recombination signals that had previously flanked the rearranged gene segments. Signal joints are usually formed in a conservative fashion without the loss or gain of nucleotides at the recombination junction. In contrast, coding joints are diversified as a result of both the random trimming and random addition of nucleotides at the junction. Interestingly, the tal^d deletion junctions bear a striking resemblance to the coding joints of assembled Ig/TCR genes. For example, if cleavage within germline AA-0.6 occurs adjacent to the proposed heptamer (see Figure 9), then a variable trimming of nucleotides (0−22 residues) clearly takes place before religation of the AA-0.6 end to form tal^d; exonucleolytic trimming of the B2EE-2.0 sequence (0−8 resiudes) is also evident. Moreover, 21 of the 22 tal^d junctions bear random nucleotides (1−13 residues) that are not derived from germline sequences, and thus may have been generated in a manner similar to the N-region nucleotides of Ig/TCR coding joints (Alt and Baltimore, 1982). Recently, Lafaille *et al.* (1989) showed that some Ig/TCR coding joints acquire nonrandom insertions of defined mono- and dinucleotides (designated P nucleotides). Although the complete rules for their identification are complex, P nucleotides are only found appended to coding ends that have not suffered exonucleolytic trimming. Notably, a thymidine residue (underlined in Figure 9) that fulfils the criteria of P nucleotides can be seen in each of the six tal^d junctions that bear an untrimmed AA-0.6 sequence (CCRF-HSB-2, MOLT-16, and patients 60, 80, 82 and 83).

The strong resemblance between tal^d junctions and the coding joints of assembled Ig/TCR genes implies that tal^d deletions are mediated by the Ig/TCR recombinase. Aberrant activity of the recombinase has also been implicated in the formation of chromosome translocations involving the Ig/TCR loci (Boehm and Rabbitts, 1989; Tycko and Sklar, 1990). Nevertheless, at least one of the two recombining elements responsible for these translocations corresponds to an Ig/TCR sequence that normally serves as a recombination signal. In contrast, both recombining elements (i.e. the AA-0.6 and B2EE-2.0 heptamers) involved in tal^d rearrangement are unnatural substrates for the recombinase; moreover, these are probably poor substrates as well, due to the deviations from the consensus heptamer sequence and the absence of an associated nonamer element (Hesse *et al.*, 1989). In view of the likely inefficiency of tal^d recombination, the recurrence of tal^d in T-ALL patients is all the

more remarkable and argues strongly that alteration of the *tal*-1 gene is a critical factor in T cell leukemogenesis.

Materials and methods

Tumor specimens and cell lines

Leukemic specimens were provided by the Pediatric Oncology Group (St Louis, MO); these represent peripheral blood or bone marrow aspirates obtained from T-ALL patients prior to treatment. Cell lines were either obtained from the American Type Culture Collection (Rockville, MD) or they were kindly provided by Drs Manuel Diaz (University of Chicago), Michael Krangel (Dana Farber Cancer Institute) or Peter Lipsky (U.T. Southwestern).

DNA analysis and cloning

DNA extracted from patient specimens and cell lines was analyzed by Southern hybridization with radiolabeled DNA probes (Southern, 1975; Feinberg and Vogelstein, 1983). λ phage libraries of genomic DNA from the SUP-T1 (Baer *et al.*, 1985) and RPMI8402 cell lines (Baer *et al.*, 1988) were screened by the method of Benton and Davis (1977), and restriction fragments of recombinant phage DNA were subcloned into plasmid and M13 phage vectors (Yanisch-Perron *et al.*, 1985; Pridmore, 1987). Nucleotide sequence analyses were performed on M13 single-stranded templates by the chain termination method (Sanger *et al.*, 1980). Chromosomal localization of the AA-0.6 DNA fragment by somatic cell hybrid analysis was conducted exactly as described previously for the B2EE-2.0 fragment (Chen *et al.*, 1990). Pulsed-field gel electrophoresis (Schwartz and Cantor, 1984) was conducted on a transverse alternating field (Gardiner *et al.*, 1986) apparatus obtained from Beckman Instruments (GeneLine). DNA from tissue culture cells was prepared in agarose blocks for restriction endonuclease digestion (Smith *et al.*, 1988). Electrophoresis was carried out in a TAFE buffer (10 mM Tris−acetate, pH 7.0, 0.4 mM EDTA) at constant temperature (12°C) and amperage (150 mA) using 10 s pulses for 20 h.

Polymerase chain reaction

Amplification of tal^d deletion junctions was conducted by the polymerase chain reaction (Saiki *et al.*, 1988) using oligonucleotide primers C (AGGGGAGCTCGTGGGAGAAATTAAG) and D (TCACAATCC-CACCGCATGCACA). The reaction conditions have been described (Cheng *et al.*, 1990). The amplification products were fractionated by electrophoresis on 10% polyacrylamide gels and visualized by ethidium bromide staining. After elution from the gel, the amplification products were phosphorylated with polynucleotide kinase (New England Biolabs) and cloned into the *Sma*I site of M13mp18 for nucleotide sequence analysis.

Acknowledgements

We thank Cary Ying-Chuan Yang and Julia Tsan for technical assistance, Cindy Baselski for preparing the manuscript, and Michael Link (Stanford) for immunophenotyping the leukemic samples. We are also grateful to Donald Capra, Eric Humphries, Joseph Sambrook, Philip Tucker and Jonathan Uhr for their comments on the manuscript, and Thomas Boehm, Andrew Carroll, Terrence Rabbitts and Graham Smith for their advice. This work was supported by the National Cancer Institute (CA46593 and CA47975) and the American Cancer Society (CH-414A). G.B. and W.C. are supported in part by National Institutes of Health grants CA-21765 (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 and a recipient of a Junior Faculty Research Award from the American Cancer Society. L.B., J.-T.C. and Q.C. are students in the Microbiology Program of U.T. Southwestern Graduate School.

References

Alt,F.W. and Baltimore,D. (1982) *Proc. Natl. Acad. Sci. USA*, **79**, 4118−4122.

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. (1989a) *Proc. Natl. Acad. Sci. USA*, **86**, 2031−2035.

Begley,C.G., Aplan,P.D., Denning,S.M., Haynes,B.F., Waldmann,T.A. and Kirsch,I.R. (1989b) *Proc. Natl. Acad. Sci. USA*, **86**, 10128−10132.

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

Bernard,O., Guglielmi,P., Jonveaux,P., Cherif,D., Gisselbrecht,S., Mauchauffe,M., Berger,R., Larsen,C.-J. and Mathieu-Mahul,D. (1990) *Genes, Chromosomes Cancer*, **1**, 194−208.

Blackwell,T.K. and Alt,F.W. (1989) *J. Biol. Chem.*, **264**, 10327−10330.

Boehm,T. and Rabbitts,T.H. (1989) *FASEB J.*, **3**, 2344−2359.

Carroll,A., Crist,W., Link,M.P., Amylon,M.D., Pullen,D.J., Ragab,A.H., Buchanan,G.R., Wimmer,R.S. and Vietti,T.J. (1990) *Blood*, in press.

Chen,Q., Cheng,J.-T., Tsai,L.-H., Schneider,N., Buchanan,G., Carroll,A., Crist,W., Ozanne,B., Siciliano,M.J. and Baer,R. (1990) *EMBO J.*, **9**, 415−424.

Cheng,J.-T., Yang,C.Y.-C., Hernandez,J., Embrey,J. and Baer,R. (1990) *J. Exp. Med.*, **171**, 489−501.

Dryja,T.P., Rapaport,J.M., Joyce,J.M. and Peterson,R.A. (1986) *Proc. Natl. Acad. Sci. USA*, **83**, 7391−7394.

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

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.

Gardiner,K., Laas,W. and Patterson,D. (1986) *Somat. Cell Mol. Genet.*, **12**, 185−195.

Hann,S.R., King,M.W., Bentley,D.L., Anderson,C.W. and Eisenman,R.N. (1988) *Cell*, **52**, 185−195.

Hayata,I., Oshimura,M., Minowada,J. and Sandberg,A.A. (1975) *In Vitro*, **11**, 361−368.

Henthorn,P., Kiledjian,M. and Kadesch,T. (1990) *Science*, **247**, 467−470.

Hershfield,M.S., Kurtzberg,J., Harden,E., Moore,J.O., Whang-Peng,J. and Haynes,B.F. (1984) *Proc. Natl. Acad. Sci. USA*, **81**, 253−257.

Hesse,J.E., Lieber,M.R., Mizuuchi,K. and Gellert,M. (1989) *Genes Dev.*, **3**, 1053−1061.

Kamps,M.P., Murre,C., Sun,X.-H. and Baltimore,D. (1990) *Cell*, **60**, 547−555.

Klein,G. and Klein,E. (1985) *Nature*, **315**, 190−195.

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

Lafaille,J.J., DeCloux,A., Bonneville,M., Takagaki,Y. and Tonegawa,S. (1989) *Cell*, **59**, 859−870.

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

Le Beau,M.M., McKeithan,T.W., Shima,E.A., Goldman-Leikin,R.E., Chan,S.J., Bell,G.I., Rowley,J.D. and Diaz,M.O. (1986) *Proc. Natl. Acad. Sci. USA*, **83**, 9744−9748.

Leder,P., Battey,J., Lenoir,G., Moulding,C., Murphy,W., Potter,H., Stewart,T. and Taub,R. (1983) *Science*, **222**, 765−771.

Lewis,S. and Gellert,M. (1989) *Cell*, **59**, 585−588.

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

Mellentin,J.D., Murre,C., Donlon,T.A., McCaw,P.S., Smith,S.D., Carroll,A.J., McDonald,M.E., Baltimore,D. and Cleary,M.L. (1989b) *Science*, **246**, 379−382.

Mitelman,F. (1987) *Cancer Cytogenetics*. A.R.Liss, New York.

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.

Nourse,J., Mellentin,J.D., Galili,N., Wilkinson,J., Stanbridge,E., Smith,S.D. and Cleary,M.L. (1990) *Cell*, **60**, 535−545.

Pellicci,P.-G., Knowles,D.M., II, Magrath,I. and Dalla-Favera,R. (1986) *Proc. Natl. Acad. Sci. USA*, **83**, 2984−2988.

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

Rabbitts,T.H. (1985) *Trends Genet.*, **1**, 327−331.

Rabbitts,T.H., Forster,A., Hamlyn,P. and Baer,R. (1984) *Nature*, **309**, 592−597.

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.A. (1980) *J. Mol. Biol.*, **143**, 161−178.

Schwartz,D.C. and Cantor,C.R. (1984) *Cell*, **37**, 67−75.

Smith,C.L., Kico,S.R. and Cantor,C.R. (1988) In Davies,K. (ed.), *Genome Analysis: A Practical Approach*. IRL Press, Oxford, pp. 41−47.

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

Stallings,R.L., Olson,E., Strauss,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.

Thompson,L.H., 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.
Tonegawa,S. (1983) *Nature*, **302**, 575–581.
Tycko,B. and Sklar,J. (1990) *Cancer Cells*, **2**, 1–8.
Villares,R. and Cabrera,C.V. (1987) *Cell*, **50**, 415–424.
Yanisch-Perron,C., Vieiera,J. and Messing,J. (1985) *Gene*, **33**, 103–119.
Yunis,J.J. (1983) *Science*, **221**, 227–236.

Received on May 16, 1990; revised on July 2, 1990