Localization of preferential sites of rearrangement within the BCR gene in Philadelphia chromosome-positive acute lymphoblastic leukemia

CHRISTOPHER T. DENNY*, NEIL P. SHAH†, SHEILA OGDEN*, CHERYL WILLMAN‡, THOMAS McCONNELL‡, WILLIAM CRIST§¶, ANDREW CARROLL¶¶, and OWEN N. WITTE†***

*Department of Pediatrics, Gwynne Hazen Cherry Memorial Laboratories, Division of Hematology/Oncology, University of California, Los Angeles, School of Medicine, Los Angeles, CA 90024; †Department of Microbiology and **Howard Hughes Medical Institute, University of California, Los Angeles, 405 Hilgard Avenue, Los Angeles, CA 90024; ‡Department of Cell Biology and Pathology, University of New Mexico Cancer Center, 900 Camino de Salud, N.E., Albuquerque, NM 87131; §Saint Jude Children’s Research Hospital, University of Tennessee Memphis College of Medicine, Memphis, TN 38101; ¶University of Alabama at Birmingham, Laboratory of Medical Genetics, Birmingham, AL 35233; and ‡The Pediatric Oncology Group, Saint Louis, MO 63108

Communicated by Michael Potter, January 30, 1989 (received for review October 28, 1988)

ABSTRACT The Philadelphia chromosome associated with acute lymphoblastic leukemia (ALL) has been linked to a hybrid BCR/ABL protein product that differs from that found in chronic myelogenous leukemia. This implies that the molecular structures of the two chromosomal translocations also differ. Localization of translocation breakpoints in Philadelphia chromosome-positive ALL has been impeded due to the only partial characterization of the BCR locus. We have isolated the entire 130-kilobase BCR genomic locus from a human cosmid library. A series of five single-copy genomic probes from the 70-kilobase first intron of BCR were used to localize rearrangements in 8 of 10 Philadelphia chromosome-positive ALLs. We have demonstrated that these breakpoints are all located at the 3' end of the intron around an unusual restriction fragment length polymorphism caused by deletion of a 1-kilobase fragment containing Alu family reiterated sequences. This clustering is unexpected in light of previous theories of rearrangement in Philadelphia chromosome-positive chronic myelogenous leukemia that would have predicted a random dispersion of breakpoints in the first intron in Philadelphia chromosome-positive ALL. The proximity of the translocation breakpoints to this constitutive deletion may indicate shared mechanisms of rearrangement or that such polymorphisms mark areas of the genome prone to recombination.

Acute lymphoblastic leukemia (ALL) is the most common pediatric malignancy afflicting 1 in 30,000 children annually (1). Though therapeutic advances have resulted in long-term disease-free survival rates of 60% (2), those tumors that contain a karyotypic abnormality have a significantly poorer prognosis (3). This would suggest that neoplastic mechanisms may be at work as a consequence of chromosomal rearrangement that make these tumors more resistant to treatment. The most common karyotypic abnormality seen in ALL is the Philadelphia chromosome (Ph). It is this same chromosomal translocation that is characteristically linked to >90% of patients with chronic myelogenous leukemia (CML) (4). In ALL, however, the Ph is less frequently observed and is present in ~10% of patients (5). Karyotypic analysis has revealed that this marker chromosome is a result of a reciprocal translocation between chromosome 9 band q34 and chromosome 22 band q11.2 (t(9;22)(q34;q11.2) (6). Though these rearrangements in ALL and CML are karyotypically indistinguishable, their molecular structures differ. In both instances this translocation results in the juxtaposition of upstream regions of the BCR gene located on chromosome 22 to a distal portion of the ABL protooncogene on chromosome 9 (7–10). This results in the formation of both hybrid BCR/ABL transcripts and proteins (11, 12). In CML, this rearrangement has been mapped to within a 5- to 6-kilobase pair (kb) region of the BCR gene and results in a 210- to 216-kDa hybrid protein (P210) (13–18).

The t(9;22) translocation in Ph+ ALL has been shown to produce a shorter BCR/ABL fusion protein of 185 kDa (P185) (19–22). Further analysis has revealed that the mRNAs coding for the P210 and P185 proteins differ only in the presence or absence of specific mid-BCR exons. Though their 5' ends were identical, less of the BCR gene was incorporated into the P185 mRNA. This implied that BCR rearrangement in Ph+ ALL occurred upstream of the CML breakpoint cluster region. The precise locations of these BCR breakpoints have been determined in only a few cases of Ph+ ALL and have been shown to occur within the first intron of BCR (23, 24). Breakpoint locations of the majority of Ph+ ALL have eluded molecular definition. Since only portions of the large first intron have been isolated, it has been postulated that these molecularly undetectable breakpoints occurred elsewhere within this intron.

To test this hypothesis, we have cloned the entire BCR gene from a human cosmid library. We have found that the first intron is ~70 kb long. By applying a series of single-copy molecular probes derived from this intron we have localized the t(9;22) translocation breakpoints in seven of nine Ph+ ALL samples and in one patient with acute undifferentiated leukemia. We demonstrate that these breakpoints are not randomly distributed throughout this intron but are more frequently found within a 20-kb region.

MATERIALS AND METHODS

Cosmid Library Creation. High molecular weight genomic DNA was prepared from washed TC-32 cells by a modified method of Maniatis et al. (25). The DNA was then partially digested with Sau3A (Boehringer Mannheim) and size-fractionated by ultracentrifugation over 10–40% sucrose gradients. The fractions containing fragments of 35–50 kb were pooled and ligated to the cosmid vector pWE-15 (Stratagene) that had been digested with BamHI and calf intestine alkaline phosphatase (Boehringer Mannheim). Ligated constructs were packaged into λ phage by using Gigapack Gold packaging extract (Stratagene). The RecA− Escherichia coli host 490A− was then infected with recombinant phage.

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

Abbreviations: ALL, acute lymphoblastic leukemia; CML, chronic myelogenous leukemia; Ph, Philadelphia chromosome; RFLP, restriction fragment length polymorphism.
Replica plating and screening of cosmid libraries was performed as described (26).

Riboprobe Screening of Cosmid Libraries. 32P-labeled cosmid end probes were made according to Wahl et al. (27). Cosmid DNAs were first digested to completion with Rsa I (Stratagene). Radiolabeled RNAs were then synthesized from T3 or T7 RNA promoters by using reagents and instructions prepared by Stratagene Cloning Systems.

Hybridizations using riboprobes were performed as follows. Cosmid filter blots (Biortrans; ICN) were prehybridized for 2 hr at 50°C in a solution of 50% (vol/vol) deionized formamide, 6 x SSC (1 x SSC is 0.15 M NaCl/0.015 M sodium citrate, pH 7.0), 5 x Denhardt’s solution (1 x Denhardt’s solution is 0.02% bovine serum albumin/0.02% polyvinylpyrrolidine/0.02% Ficoll), denatured herring sperm DNA (200 

Southern Analyses. Genomic DNAs for Southern experiments were prepared in similar fashion to that used for cosmid library generation or by guanidinium isothiocyanate lysis (28). Endonuclease digestions were performed according to manufacturer’s recommendations (Boehringer Mannheim). Digested DNAs were size-fractionated on 0.8% agarose gels (Seakem LE, FMC) and transferred to nitrocellulose membranes [Nitro Plus (200, Micron Separations (Westboro, MA)] (29). Southern blots were hybridized to 32P-labeled nick-translated genomic probes and subsequently washed as described (30).

Nucleotide Sequence Procedure. Dideoxynucleotide sequencing was performed using reagents and instructions prepared by United States Biochemical.

RESULTS

The First Intron of BCR Is 70 kb Long. A cosmid library was made from the neuroepithelioma cell line TC-32. The complete karyotype of this cell line has been described: +5, +10, t(iq1)(t11;22)(q24:q12) (31). This cell line contains a translocation involving chromosomes 11 and 22 in addition to its normal chromosome 22. However, this particular tumor-associated rearrangement has been shown to lie telomeric to BCR and, therefore, would not affect our experiments (32).

The cosmid vector pWE-15 was chosen as the optimal cloning vehicle (27). This 8.8-kb plasmid contains the phage cos site, plasmid origin, and bacterial resistance genes necessary for successful packaging into lambda phage and propagation in bacterial hosts. In addition T3 and T7 RNA promoters flank the BamHI cloning site. Utilizing these promoters, 32P-labeled cosmid end probes can easily be generated and applied to “chromosome walking” strategies.

An unamplified library of 7 x 105 colonies was screened with two probes specific for the first and second exons of BCR. The results from this and subsequent screens are shown schematically in Fig. 1. The upstream genomic fragment derived from exon 1 hybridized to two cosmid clones: P1-11 and P1-9. Library screens with the exon 2 probe yielded a cluster of four overlapping clones: M1-4, M1-5, M1-11, and M1-14. To cross the gap between these two cosmid clusters a 32P-labeled riboprobe was generated from the 3' end of clone P1-9 utilizing the T3 RNA promoter. Screening with this probe resulted in four additional cosmid clones: P2-14, P2-21, P2-22, and P2-23. Restriction endonuclease mapping and hybridization studies using cosmid end probes showed that overlap of all cosmid clones had been achieved.

Two cosmid clones P1-9 and P2-23 together spanned the entire intron and were selected for further restriction endonuclease mapping. EcoRI fragments were subcloned into the plasmid vectors pGEM-4 or Bluescript II and mapped in detail. Subclone maps were unambiguously placed in the intron by comparisons with cosmid digests as well as hybridizations using subclone end fragments to genomic and cosmid Southern blots (Fig. 2).

The precise locations of most of the exons in BCR have not been determined. Intron–exon borders have been mapped in only the first exon and the five small exons surrounding the CML breakpoint cluster region (14, 23). To determine whether any exonic sequences were present within this region hybridization studies were performed. A cDNA fragment containing the first and second exons of BCR was used to probe restriction endonuclease-digested cosmid DNAs. Hybridization was only seen in those cosmid fragments to which exons 1 and 2 had been mapped, supporting the conclusion that no exonic sequences exist in this intervening

![Fig. 1. BCR genomic locus. The overall genomic structure of the BCR locus is schematically illustrated. Only cosmid clones that define the first intron are displayed (see text). The remaining clones, that describe the 3’ end of the BCR gene, are not shown. Exon locations and 3’ EcoRI restriction sites have been determined by Hermans et al. (23). Solid boxes denote BCR exons that have been mapped; stippled boxes indicate restriction fragments containing less precisely defined exons. Vertical lines indicate locations of EcoRI restriction sites. Locations of the initial exon 1 and exon 2 DNA probes used in cosmid selection are denoted by solid triangles. The exon 1 probe was a 1-kb BamHI genomic fragment containing the 3’ 800 base pairs (bp) of exon 1 and 200 bp of intronic sequence (see Fig. 2). The exon 2 probe was a 1.1-kb Bgl II BCR cDNA fragment derived from Bgl II sites 1230 bp and 2364 bp on the 3’ side of the coding ATG. The open triangle illustrates the location of the synthesized riboprobe used to cross the intron (see text).](image-url)
region (data not shown). The size of this first intron, therefore, is 70 kb.

**Breakpoints Are Localized to the First Intron of BCR.** Five single-copy genomic probes were generated from the first intron of BCR. Regions rich in reiterated sequence were identified by hybridizing subclone plasmid Southern blots with 32P-labeled total human genomic DNA. Those fragments that were not reiterated and would detect rearrangement over a large distance were isolated by preparative agarose or polyacrylamide electrophoresis. These candidate probes were then 32P-labeled and hybridized to normal human genomic DNAs. Those probes that hybridized to a single band with minimum background were used to detect the t(9;22) translocation in Ph+ ALL.

Genomic DNA from six Ph+ ALL tumor samples, three tumor-derived cell lines, and one tumor sample from a patient with Ph+ acute undifferentiated leukemia were assembled. The clinicopathologic features of this group are enumerated in Table 1. With the exception of the acute undifferentiated leukemia sample, all specimens were consistent with the diagnosis of ALL of pre-B-cell lineage.

The bank of single-copy genomic probes from the first intron of BCR were applied to restriction endonuclease digested-Ph+ ALL DNAs to localize t(9;22) translocation breakpoints. Rearrangement was detected in 8 of the 10 samples using these probes. Representative examples of these experiments are displayed in Fig. 3. All apparent rearrangements were confirmed and further localized by at least one other restriction enzyme digest. No apparent rearrangement was seen with any of the first intron probes in DNAs from Ph− ALL or Ph+ CML samples. Conversely, a probe derived from the CML breakpoint cluster region did detect rearrangement in this CML sample but failed to show any genomic structural abnormality in any of the Ph+ ALL samples (data not shown). The individual breakpoint locations from each of the 8 samples are shown schematically in Fig. 2.

**BCR Breakpoints Are Found Near a RFLP.** Breakpoint mapping studies demonstrated that all BCR rearrangements were localized to the 3' end of the first intron. At the center of this breakpoint region an unusual RFLP was found. Seven of the eight detected Ph+ samples had breakpoints that localized to within 10 kb of this polymorphism. Restriction endonuclease mapping demonstrated that it resulted from a 1-kb internal deletion. The size and location of this deletion was the same in cosmid clones as well as in genomic maps from both tumor and normal control DNAs (data not shown). This is most consistent with the deletion being a stable RFLP of the human genome rather than a cloning artifact. Analysis of both normal and tumor cell DNAs showed that the deleted allele occurred in approximately one in three haploid genomes (data not shown).

Since the cell line TC-32 contained both alleles of this polymorphism, we were able to isolate both forms from previously defined cosmid clones. Fragments with and without the internal deletion were subcloned, further mapped, and sequenced (Fig. 4). No obvious secondary structures were found near the breakpoints of this 1-kb internal deletion that could account for its formation. Examination of this deletion revealed the presence of Alu reiterated sequences contained near its 3' end. These sequences had all the structural features

---

**Table 1. Marker studies**

<table>
<thead>
<tr>
<th>Sample</th>
<th>Disease</th>
<th>Age, years</th>
<th>CALLA</th>
<th>TdT</th>
<th>B1</th>
<th>Clg</th>
<th>Ref.</th>
</tr>
</thead>
<tbody>
<tr>
<td>SUP-B13</td>
<td>ALL</td>
<td>8</td>
<td>+</td>
<td>+</td>
<td>NA</td>
<td>+</td>
<td>*</td>
</tr>
<tr>
<td>UC-6</td>
<td>ALL</td>
<td>4</td>
<td>+</td>
<td>+</td>
<td>NA</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>UC-7</td>
<td>AUL</td>
<td>54</td>
<td>NA</td>
<td>−</td>
<td>NA</td>
<td>NA</td>
<td>33</td>
</tr>
<tr>
<td>ALL-1</td>
<td>ALL</td>
<td>7</td>
<td>+</td>
<td>+</td>
<td>−</td>
<td>−</td>
<td>34</td>
</tr>
<tr>
<td>SSPB</td>
<td>ALL</td>
<td>13</td>
<td>+</td>
<td>NA</td>
<td>+</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>SWG-1</td>
<td>ALL</td>
<td>40</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>−</td>
<td></td>
</tr>
<tr>
<td>SWG-2</td>
<td>ALL</td>
<td>5</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>NA</td>
<td></td>
</tr>
<tr>
<td>SWG-3</td>
<td>ALL</td>
<td>55</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>−</td>
<td></td>
</tr>
<tr>
<td>UC-8</td>
<td>ALL</td>
<td>1</td>
<td>+</td>
<td>NA</td>
<td>NA</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>SWG-4</td>
<td>ALL</td>
<td>14</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td></td>
</tr>
</tbody>
</table>

AUL, acute undifferentiated leukemia. Marker studies are scored as follows: +, >80% reactivity; −, <20% reactivity; NA, data was not available. Immunophenotypic marker abbreviations are as follows: CALLA, common ALL antigen; TdT, terminal deoxynucleotide transferase; Clg, cytoplasmic μ immunoglobulin. Select references for each patient are given when the data was prepared by others. *, Data for SUP-B13 and SSPB were supplied by Stephen D. Smith (Stanford University).
characteristic of Alu family repeats and had a 90% sequence similarity to Alu consensus sequences (35).

**DISCUSSION**

We have defined the 70-kb first intron of the *BCR* gene in an effort to localize the translocation breakpoints in Ph⁺ ALL. Utilizing five single-copy genomic probes derived from this region, we have localized the t(9;22) breakpoints in 8 of 10 Ph⁺ samples. *BCR* rearrangement in these 8 patients occurred exclusively within the first intron. This is in contrast to previous reports that approximately half the time rearrangement in Ph⁺ ALL involved the CML breakpoint cluster region (34, 36–38). There is no straightforward explanation

**Fig. 3.** Southern analyses demonstrating rearrangement within the first intron of *BCR* in Ph⁺ samples. (A) Samples digested with HindIII and probed with Ph-10 detected rearrangement in SWG-1 and SWG-2. (B) Probe Ph-15 detected an additional recombinant band in EcoRI-digested DNA of patient sample SWG-3. (C) Sample DNAs cut with HindIII and hybridized to probe Ph-15 detected both chromosomal rearrangement and an RFLP. Rearrangement caused by the Ph translocation is seen in samples UC-6, UC-7, SUP-B13, and SSFB. The presence of the two alleles of the 1-kb deletion RFLP is seen in one Ph⁻ ALL sample and in a normal human sperm DNA control (HS). Sizes of germ-line bands (in kb) are enumerated to the right of each panel. Rearranged bands are demarcated by arrowheads to the left of each lane. Specific probes correspond to genomic fragments from the first intron of *BCR* as shown in Fig. 2.

**Fig. 4.** Sequence analysis of the 1-kb deletion RFLP. (A) Breakpoints of the 1-kb deletion were determined by sequencing both alleles isolated from the TC-32 cell line and are bracketed. Alu repetitive sequences present within the deleted segment are boxed. (B) The 1-kb RFLP showing detailed restriction map and sequencing strategy. Bold vertical lines indicate breakpoint locations. Restriction enzyme abbreviations are as follows: E, EcoRI; Bg, Bgl II; S, Sal I; X, Xba I; Bal, Bal I; P, Psi I; K, Kpn I.

for this discrepancy. Clinicopathologic data from our patient population did not differ significantly from those previously described. Further study is required before this disparity can be resolved.

Two of the 10 samples in our study failed to demonstrate rearrangement with any of the five first-intron probes on multiple digests. These tumor specimens also failed to show *BCR* rearrangement anywhere in the 50 kb on the 3' side of the first intron including the CML breakpoint cluster region (data not shown). In these cases rearrangement may have occurred within the first intron that would not be detected with the present set of probes. If during translocation a deletion of *BCR* probe sequences occurred then rearrangement detection with that probe would not be possible. Deletions of this nature have been found in >10% of t(9;22) translocations in CML (39). Alternatively, the translocations in these tumor samples may involve an entirely different *BCR* locus. Croce et al. (40) have reported three other closely related *BCR* genes that are located on chromosome 22 within band q11.2. Rearrangement to any of these genes would be karyotypically indistinguishable from the Ph chromosome in ALL. Since these genes diverge at their 5' ends, probes derived from the first intron of *BCR* would not be expected to strongly cross react with these other loci. On this basis rearrangement with one of these other genes would escape detection. The functional roles of these alternate *BCR* loci as well as their eligibility as possible partners in t(9;22) translocations remains to be clarified.

*BCR* breakpoints in Ph⁺ ALL are not randomly distributed throughout the first intron but instead fall within a 20-kb region at the 3' end. This contrasts with *BCR* breakpoints in CML that are more closely localized to the 5- to 6-kb breakpoint cluster region. Within the breakpoint cluster region are five short exons separated by introns no more than 2 kb. It was initially believed that the limits of the breakpoint cluster region were defined by the suitability of specific introns to serve as recombination targets in CML (14). This suitability was dictated by the reading frame of the 5' exon since only those exons in frame with the second exon of *ABL* would be capable of producing functional BCR/ABL hybrid products. Rearrangement outside of the breakpoint cluster region would lead to biologically inactive products and, therefore, not be seen with CML.

Different mechanisms must be invoked to account for the grouping of *BCR* breakpoints in Ph⁺ ALL and suggest that the site specificity in the CML breakpoint cluster region may not be entirely biologically driven. Rearrangement anywhere within the first intron would result in the same BCR/ABL mRNA and protein. Nevertheless, a random dispersion of breakpoints is not seen. Site selection could be caused by the
presence of regulatory elements in the 5' portion of the first intron necessary for efficient transcription. Alternatively, clustering of BCR breakpoints could reflect an increased accessibility to rearrangement of a part of the BCR first intron. The presence of an atypical RFLP centered in the Ph+ ALL BCR breakpoint region suggests that this portion of the first intron may be predisposed to rearrangement. In this case polymorphism is generated by a 1-kb deletion that is present in approximately one-third of haploid genomes (ref. 24 and data not shown). Though the majority of genomic polymorphisms are caused by either point mutation of a restriction endonuclease site or the presence of a variable number of tandem repeated sequences. Deletion/insertion RFLPs have been infrequently described (41–43). Such an RFLP has been demonstrated to be near the breakpoint of an X-Y translocation resulting in a human XXY male (44). For Ph+ ALL it is unclear from our data whether the presence or absence of this deletion results in a direct increased susceptibility toward rearrangement.

The mechanisms underlying the formation of this RFLP or that mediate the t(9;22) chromosomal translocation are not understood. Both, however, have been linked to the presence of Alu family dispersed repeats. Sequence analysis of the 1-kb RFLP deletion associated with BCR rearrangement in Ph+ ALL demonstrated that it contains an Alu repeat. Alu family repeats have also been noted to be in close proximity to both der9 and der22 breakpoints in specific instances of Ph+ ALL (14, 45) and ALL (46). The association of Alu repeats to chromosomal rearrangement is not limited to the t(9;22) translocation. Recombinations involving Alu sequences have been found resulting in alternation of the low density lipoprotein receptor in familial hypercholesterolemia (47, 48). Alu repeats have been physically linked to mutations of adenine deaminase (49) and β-hexosaminidase genes (50). They have also been implicated in chemically induced rearrangements of the MET protooncogene (51). In spite of this increasing association of Alu sequences and genomic recombination, it is likely that other factors must come into play before rearrangement occurs. Tertiary genomic structure and transcriptional activity may exert important influences on the location and frequency of certain recombinant events. Complete characterization of the BCR gene will allow Ph+ ALL to serve as a model system to further delineate these factors.

We gratefully acknowledge the following people for their important contributions to this work: C. Patrick Reynolds for supplying the neuroepithelioma cell line TC-32; Eric Lai for assistance in cosmid library generation; and Stephen D. Smith for the contribution of Ph+ samples SUP-B13 and SSPB. W.C. and A.C. were supported by the American Lebanese/Syrian Associated Charities and Grants CA 20180, CA 21765, and CA 31566 from the National Institutes of Health. Work in the laboratory of O.N.W. was supported by Grants CA 27507, CA 32737, and CA 40957 from the National Institutes of Health.

12. Grosveld, G., Verwoerd, T., van Aghoven, T., de Klein, A., Ramachana-