[CANCER RESEARCH 61, 6640–6648, September 15, 2001]

Advances in Brief

Characterization of the 13q14 Tumor Suppressor Locus in CLL: Identification of *ALT1*, an Alternative Splice Variant of the *LEU2* Gene[1]

Florencia Bullrich, Hidenori Fujii,[2] George Calin, Hideaki Mabuchi,[3] Massimo Negrini,[4] Yuri Pekarsky, Laura Rassenti, Hansjuerg Alder, John C. Reed, Michael J. Keating, Thomas J. Kipps, and Carlo M. Croce[5]

Department of Microbiology and Immunology, Kimmel Cancer Center, Thomas Jefferson University, Philadelphia, Pennsylvania 19107 [F. B., H. F., G. C., H. M., M. N., Y. P., H. A., C. M. C.]; The Burnham Institute Cancer Research Center, La Jolla, California 92037 [J. C. R.]; Section of Leukemia, Division of Medicine, The University of Texas M. D. Anderson Cancer Center, Houston, Texas 77030 [M. J. K.]; Department of Medicine, University of California San Diego, La Jolla, California 92093 [L. R., T. J. K.]; and The CLL Research Consortium [L. R., J. C. R., M. J. K., T. J. K., C. M. C.]

Abstract

Chromosome 13q14 deletions constitute the most common genetic abnormality in chronic lymphocytic leukemia (CLL). To identify the putative tumor suppressor gene targeted by 13q14 genomic loss, we completely sequenced and characterized a segment of 790 kb at 13q14 spanning the minimal region of loss in CLL. Transcribed sequences in the region were identified through database homology searches and exon-prediction analysis. Two-hundred kb at the centromeric end of the sequence contain five CpG islands, three previously identified genes *LEU5/RFP2*, *LEU2*, and *LEU1*, seven of seven EST clusters composed of >10 ESTs, and a large number of predicted exons. Homology searches against the mouse EST database have allowed us to identify a highly conserved alternative first exon of the *LEU2* gene, giving rise to a novel transcript, *ALT1* (GenBank accession no. AF380424), which originates within a G+C region in the vicinity of the *D13S272* marker. Two novel 3′ exons of *LEU2* were also identified and are present in both *LEU2* and *ALT1* transcripts. However, we have not identified any mutations in leukemia cases, or alterations in expression of mRNAs in the region, that might directly implicate these mRNAs in the pathology of CLL. The centromeric end of the sequence, where all reported genes are located, contains twice the expected amount of ALU repeats, whereas the telomeric end is LINE1 rich and contains four LINE1 elements longer than 4 kb, including two full-length LINE1 sequences. This feature of the sequence may favor the occurrence of chromosomal rearrangements and may confer instability to the region, resulting in deletions that may inactivate an as yet unidentified tumor suppressor.

Introduction

CLL[6] is the most common form of adult leukemia in the Western hemisphere. CLL is a largely incurable malignancy of mature B lymphocytes. In persons >70 years, CLL has an incidence of 10 of 100,000, making it 10 times more common in the older population than in people <50 years (1). In contrast to the majority of leukemias and lymphomas, the molecular etiology of CLL remains largely unknown. Indeed, CLL, remains a challenge for physicians and scientists alike. Translocations such as t(11;14)(q13;q32) and t(14;18)(q32;21) involving the immunoglobulin locus on chromosome 14 are characteristic of B-cell neoplasms, such as mantle cell lymphoma and follicular lymphoma. These translocations, which were shown to result in overexpression of cellular proto-oncogenes that either promote cell division (*BCL1/Cyclin D1* at 11q13) or inhibit cell death (*BCL2* at 18q21; Refs. 2–5), are rare in CLL. In fact, a minority of B-CLL has been reported to carry rearrangements at 14q32, but most cases are now thought to represent mantle cell lymphoma (6), and although *BCL2* is overexpressed in >85% of B-CLLs, *BCL2* gene rearrangements in B-CLL are infrequent events (6).

The use of fluorescence *in situ* hybridization-based interphase analysis, comparative genomic hybridization, and microsatellite screening has revealed that hemizygous and homozygous loss at 13q14 occurs in more than half of CLL and constitutes the most common chromosomal abnormality in this leukemia. 13q14 deletions also occur in ~50% of mantle cell lymphomas and in 16–40% of multiple myeloma (6–11), suggesting that one or more tumor suppressor genes at 13q14 are involved in the pathogenesis of this clinically and pathologically diverse group of tumors.

Loss of heterozygosity studies were used to characterize the 13q14 CLL tumor suppressor locus, to define the minimal region of loss, and with the ultimate aim of identifying the critical gene. BAC, PAC, and cosmid contigs of the region have been developed (9, 12–20), and two core regions of loss have emerged from a series of studies aimed at identifying the critical gene. Here, we report the complete sequence and characterization of genes and expressed sequences within a region of 790 kb between markers *D13S1150* and *D13S25* and encompassing the entire 13q14 CLL tumor suppressor locus.

Materials and Methods

Patient Samples and Cell Lines. Normal/tumor paired CLL samples 23–120 were obtained from patients diagnosed with B-CLL and enrolled in Eastern Cooperative Oncology Group protocols and have been described previously (14). Other patient samples were obtained after informed consent from patients diagnosed with CLL at one of the CLL Research Consortium institutions. Briefly, peripheral blood was obtained from CLL patients, and mononuclear cells were isolated through Ficoll-Hypaque gradient centrifugation (Amersham Pharmacia Biotech, Piscataway, NJ). Patient and cell line DNA was isolated by proteinase K digestion and phenol extraction. Total RNA was isolated with TRIzol buffer following the manufacturer's instructions (Invitrogen, Life Technologies, Inc., Carlsbad, CA). Poly(A)+ RNA was isolated from selected cases with the Fast Track 2 kit (Invitrogen, Life Technologies, Inc.).

BAC Contig Construction. Eight BAC clones from 13q14, spanning the region between *D13S1150* and *D13S25*, were isolated from a commercial human BAC library (Research Genetics, Huntsville, AL) by PCR screening with oligonucleotides specific for microsatellite markers within the region (14). BAC DNA was isolated with a Qiagen Maxi kit (Qiagen, Valencia, CA), and BAC ends were sequenced with vector-specific primers by dyedeoxy-

Received 5/16/01; accepted 7/26/01.

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

[1] Supported by Program Project Grant P01CA76259, P01CA81534, and P30CA56036 from the National Cancer Institute and by a Kimmel Scholar award to F. B. and Y. P.

[2] Present address: Fukui Red Cross Hospital, Fuikui 918-8501, Japan.

[3] Present address: Department of General Gastroenterological Surgery, Osaka Medical College, Osaka 569-8686, Japan.

[4] Present address: Department of Experimental and Diagnostic Medicine, University of Ferrara, 44100 Ferrara, Italy.

[5] To whom requests for reprints should be addressed, at Department of Microbiology and Immunology, Kimmel Cancer Center, Thomas Jefferson University, Philadelphia, PA 19107. Phone: (215) 503-4645; Fax: (215) 923-3528.

[6] The abbreviations used are: CLL, chronic lymphocytic leukemia; BAC, bacterial artificial chromosome; PAC, P1-derived artificial chromosome; EST, expressed sequence tag; RT-PCR, reverse transcriptase-PCR; RACE, rapid amplification of cDNA ends; LINE, long interspersed element; ORF, open reading frame; dbEST, EST database.

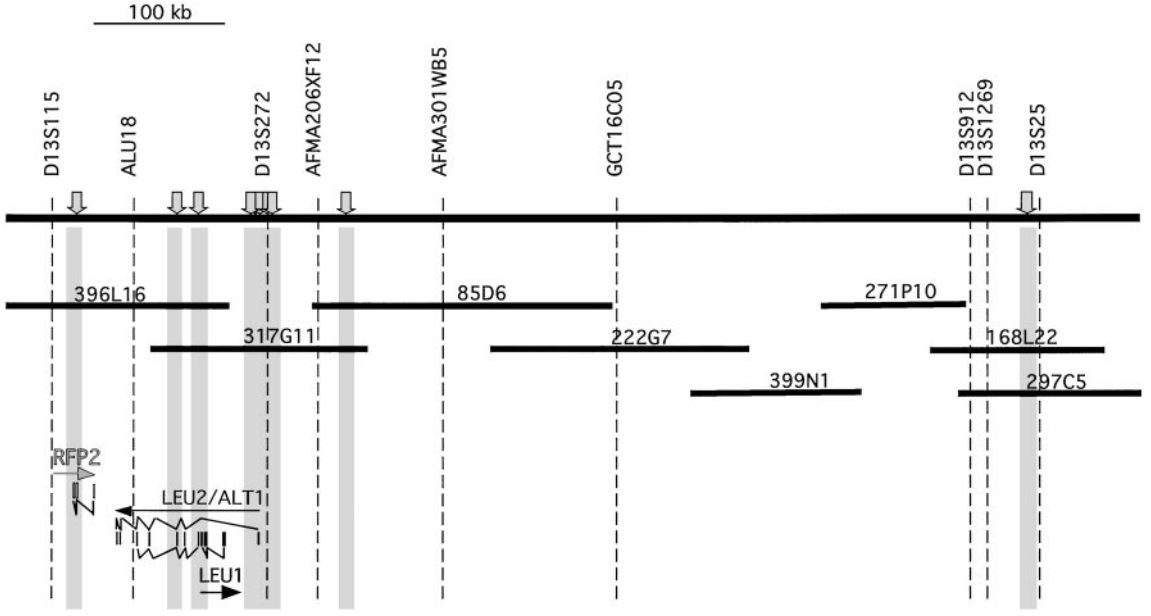

Fig. 1. BAC contig spanning the 13q14 locus. *RFP2/LEU5*, *LEU2/ALT1*, and *LEU1* exons are represented as *shaded boxes*. The direction of transcription is marked by *horizontal arrows* for each gene. *Thick vertical arrows* and *gray columns* under the *thick arrows* mark the position of CpG islands identified by GrailEXP.

terminator chemistry on a Perkin-Elmer/Cetus DNA Thermal Cycler 9600 and the Applied Biosystems Model 377 DNA sequencing systems (Perkin-Elmer/Cetus, Applied Biosystems, Foster City, CA). Gaps in the contig were closed using BAC-end-specific oligonucleotide pairs to screen the BAC library and isolate overlapping clones.

Shotgun Sequencing. BAC DNA was isolated as described above. For each clone, 30-ml aliquots containing 30 mg of DNA were sonicated at the lowest energy setting for 5–10 s at 0°C with the 3-mm probe. Aliquots were electrophoresed on 0.8% agarose gels, and preparations with peak fragment size between 2 and 3 kb were selected. DNA aliquots of 5 ml were prepared and digested for 10 min with *BAL-31* restriction enzyme (New England Biolabs, Beverly, MA), extracted with phenol, precipitated, dissolved, and size fractionated on 0.8% low melting temperature agarose gels. DNA ranging in size between 1.5 and 2 kb was cut from the gel, purified with a GeneClean III kit (BIO 101), and dissolved in 10 μl of distilled water. Plasmid libraries from the fragments were prepared as described elsewhere (21). Recombinant plasmids were isolated from each BAC, and DNA from the plasmids was prepared with the BioRobot 9600 (Qiagen). Plasmid DNA concentration was measured with a FluorImager SI analyzer. The presence of insert in recombinant plasmids was confirmed by *Eco*RI and *Hin*dIII digestion followed by agarose gel electrophoresis. Plasmids were sequenced with vector-specific oligonucleotide primers on an Applied Biosystems automated sequencer as described above.

Contig Assembly and Sequence Analysis. Sequence data were edited, assembled, and analyzed with the Sequencher 3.0 software (Genecodes, Ann Arbor, MI). Primer walking and PCR amplification and sequencing were used to close gaps or to extend the sequence read of a given subclone. Each clone was sequenced at least twice in one direction and once in the opposite direction. Sequences were proofread, and final sequences for each BAC clone were assembled into the contig using both BLAST (22)[7] and Sequencher. The final sequence of 790716 bp is accessible through the GenBank database[8] with accession no. AF380421.

Identification Expressed Sequence Tags and Exons within the Region. Repeat sequences were identified and filtered using the Repeat Masker database available through the University of Washington repeat Masker Server[9] (23). CpG islands were located with the GrailEXP software accessed through the Oak Ridge National Laboratory Computational Biology Resources Web site.[10] ESTs within the region were identified through BLAST-based homology searches. Overlapping 25–200 kb sequence blocks from our contig were used to query the human and mouse dbESTs through the National Center for Biotechnology Information network service.[11] Putative exons were identified using GRAIL (24) and FGENESH (25), both of which are accessible through the Baylor College of Medicine Human Genome Sequencing Center,[12] and GENESCAN[13] (26). Single copy sequences and putative exons showing high homology (>90%) to clusters of more than two ESTs and/or significant homology to mouse ESTs were additionally tested for expression by Northern blotting and RT-PCR. Fig. 2 was prepared using the PipMaker (27) Web server.[14] Fig. 3 was prepared using GeneQuest and GenVision Beta software (DNAstar Corp., Madison, WI).

RT-PCR. Patient and cell line cDNAs were synthesized from 2 μg of total or 0.2 μg of poly(A)+ RNA with the Superscript II polymerase (Invitrogen, Life Technologies, Inc.) enzyme primed with 100 pmol of oligodeoxythymidylic acid and/or random primers (Invitrogen, Life Technologies, Inc.) following the manufacturer's instructions. cDNA (1–3 μl) was used as template for RT-PCR reactions with the Advantage II RT-PCR kit (CLONTECH, Palo Alto, CA) under standard conditions as described above for genomic PCR. When necessary to avoid amplification from DNA contaminating the RNA sample, RNA was treated with amplification grade DNaseI (Invitrogen, Life Technologies, Inc.) before cDNA synthesis. Parallel reactions were carried out with cDNA from each sample prepared with and without reverse transcriptase enzyme to check for DNA-derived amplification products.

cDNA Library Screening. cDNA libraries from fetal liver, fetal brain, peripheral blood lymphocytes, kidney, spleen, and testis were purchased from CLONTECH. One million to 3×10^6 plaque from each library were screened with EST or exon-derived probes following the manufacturer's instructions, and isolated clones were analyzed by restriction endonuclease digestion and sequencing.

RACE. The 5′ and 3′ ends of mRNAs were isolated from double-stranded cDNA synthesized from bone marrow, fetal liver, lymph node, and peripheral blood leukocyte RNA (CLONTECH) with oligodeoxythymidylate primers following Marathon RACE amplification kit (CLONTECH) procedures. The cDNA was ligated to Marathon adapters, and RACE products were generated by long PCR with a gene-specific primer and the AP1 primer for 25 cycles. A nested primer was used in conjunction with AP2 for the second PCR. RACE products were separated on 1.5% agarose gels, and bands were cut out of the

[7] Internet address: http://www.ncbi.nlm.nih.gov/BLAST/.
[8] Internet address: http://www.ncbi.nlm.nih.gov/GenBank/GenBankSearch.html.
[9] Internet address: http://repeatmasker.genome.washington.edu/.
[10] Internet address: http://grail.lsd.ornl.gov/grailexp/.
[11] Internet address: http://www.ncbi.nlm.nih.gov.
[12] Internet address: http://searchlauncher.bcm.tmc.edu:9331/.
[13] Internet address: http://genes.mit.edu/GENESCAN.html/.
[14] Internet address: http://bio.cse.psu.edu/pipmaker/.

Table 1 *Repeat sequences within the 790-kb 13q14 locus*[a]

Repeat type	Observed genome wide (%)	790 kb (%)	Centromeric end (178 kb, %)	Telomeric end (612 kb, %)
SINEs[b]	13.14	11.47	23.17	8.09
ALU	10.60	8.85	21.79	5.10
MIR[b]	2.20	2.63	1.38	2.99
LINEs	20.42	20.41	13.31	22.52
LINE1	16.89	16.30	9.33	18.34
LINE2	3.22	3.58	3.37	3.69
LINE3	0.31	0.53	0.50	0.48
LTR elements	8.29	5.39	3.91	5.77
DNA elements	2.84	3.44	2.99	3.55
Total interspersed	44.83	40.71	43.27	39.92

[a] Expected percentage values are from Ref. 28.
[b] SINE, short interspersed nuclear elements; MIR, mammalian-wide interspersed repeat.

gel, purified, and sequenced. Primers for 5′ RACE experiments for *ALT1* were as follows: R1 5′ AAG AAG TGA GGC TGT TCT CCA G 3′, R2 5′ ATT GGT TCT CTG CTA CAG GTC AA 3′, R3 5′ AGC TGT GCC TCA GCC AGG TGT 3′, R4 5′ CAT CGC CGG GGA CCA GGG AA 3′, R5 5′ CTC GAT TTT TTG TGC AGT TTC AGC 3′. R1 and R2 are located on exon 2 of the *LEU2* gene. R3, 4, and 5 are located on exon 1a of *LEU2*. All combinations of primers were used that would allow for a nested RACE reaction.

***In Silico* Cloning of Mouse mRNAs.** Mouse ESTs with similarity to *LEU2* and *ALT1* were downloaded from public databases, imported onto Sequencher 4.5 software (Genecodes), assembled, and additionally analyzed to determine alternatively spliced forms and putative ORFs. BLAST was used to evaluate overall similarity between the human and mouse genes.

Expression Analysis: Northern Blotting and Semiquantitative PCR. Multiple tissue Northern blots containing RNA derived from a variety of human tissues, including immune and fetal tissues, were purchased from CLONTECH. In addition, Northern blots were prepared with 10–20 μg of total RNA from CLL samples, normal tissues, and cell lines. RNA was electrophoresed on 0.8% agarose gels, transferred onto HybondN+ nylon membranes (Amersham Pharmacia Biotech), and hybridized with radioactively labeled probes following standard procedures. Relative amounts of RNA on each lane were determined by hybridization with a G3PDH control probe (CLONTECH). Experimental RT-PCRs were also performed with 22–25 cycles of amplification. After Southern blotting and hybridization with product-specific probes, the intensity of the experimental amplification product was compared with that of the control product for each sample.

Mutational Analysis. Genomic DNA (50–100 ng) was used as the template for PCR amplification. Primers were designed from intron sequences flanking exons. PCR reactions were carried out with 10 pmol of each specific primer and 10 nM each deoxynucleotide triphosphate in a total volume of 25 ml under standard conditions. Reactions consisted of an initial cycle of denaturation at 94°C for 5 min followed by 35 cycles of amplification as follows: 94°C for 30, 55°C-65°C for 30, 72°C for 30, and a final extension at 72°C for 5 min. Primers were as follows: for *LEU5* forward Leu5gen_F 5′ GTT CAC CAG GTC TAA ACA GC 3′ and Leu5gen_R 5′ GAA AGT GGA TGG GAA CAT ACG 3′; for *CLLD4* CLL4F1 5′ GGT TGA TTT CCT AGG TGT GGC TGA 3′ (forward), CLL4R2 5′ TTT GGT CAT TCT ACC CAA CTC CAC 3′ (reverse); for *CLLD3* ORFc_F_cp5 5′ ACT AGA GCT GAG CTG GGA CGA 3′, ORFc_R_cp5 5′ TTG GAA GGA AAC CAG GAG CGA 3′, ORF3′_F_cp5 5′ TGC CAC AGC ACT CCT AAG AC 3′, ORF3′_R_cp5 5′ TTG CCT CGT CGA ACT TTC AGC 3′, ORF5′_F_cp5 5′ AGG GCA GAG GTG TGA GGG

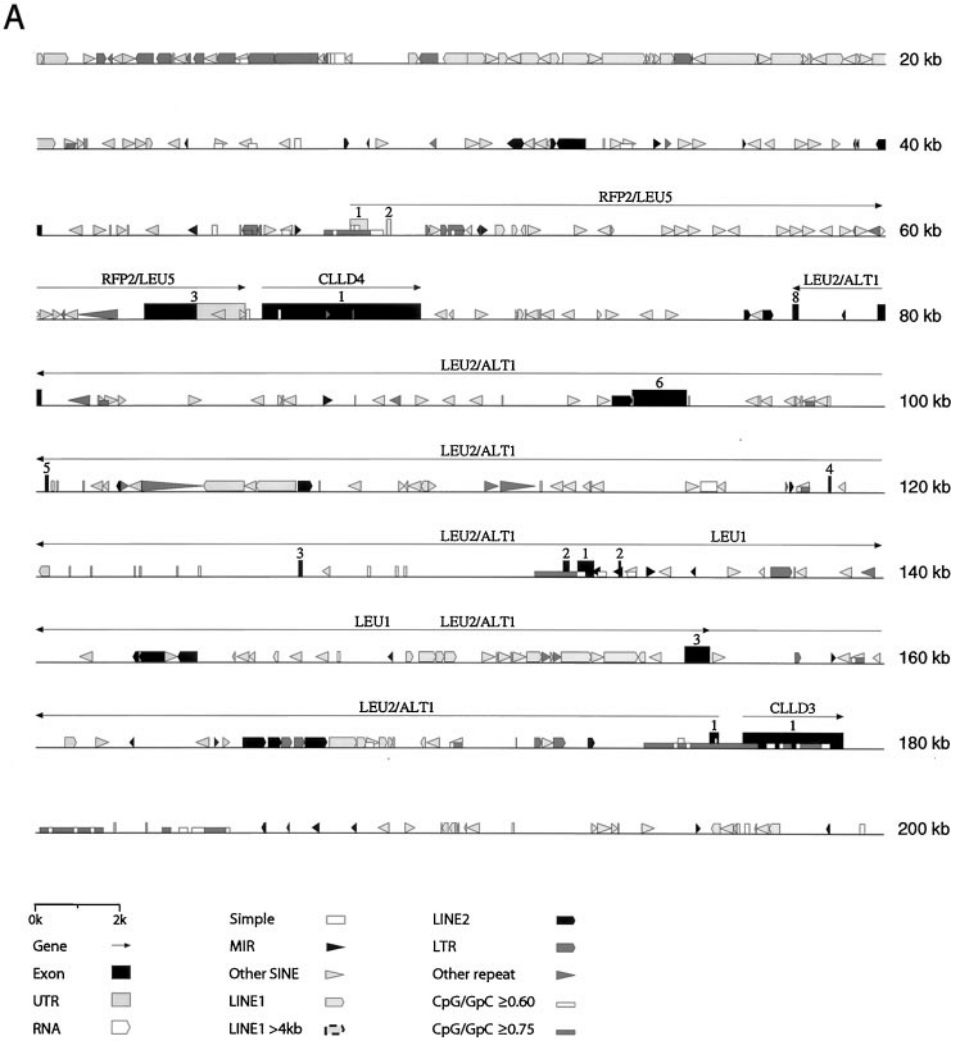

Fig. 2. Repeat sequences within the 13q14 locus. *A*, 0–200 kb; *B*, 200–790 kb. LINE1 sequences longer than 4 kb are highlighted by a *thicker line* under the corresponding *box*.

Fig. 2. Continued.

AT 3′, ORF5′_R_cp5 5′ CCG GTA GGG AGC TCA GTT TC 3′; for *LEU2/ALT1* exon 6 forward 01E6GSP2 5′ ATG CCT GAA GCG TTT CCT 3′, and reverse 02E6GSP2 5′ CTG CAC TCG CAA GTT ATC 3′; for *LEU2/ALT1* exon 7 forward 01E7GSP2 5′ GTC AAG AAT TCA ACC AGC 3′, and reverse 02E7GSP2 5′ TTA CAA GCA TGC AGG CCA 5′. PCR products were separated by gel electrophoresis in 1.5–2% agarose. Reactions were either purified directly or bands were cut out of the gel and purified with the QIAquick PCR purification kit (Qiagen). Exon 1 of *ALT1* was amplified using the Fail Safe PCR system (Epicentre Technologies, Madison, WI) using primers 01E1GSP2 (forward) 5′ CTT TTC TCT CCC AGG AGC 3′ and 04E1GSP2 (reverse) 5′ GCA CCG GCC ATA CGA CGA 3′, with buffer H and 60°C annealing temperature.

Genomic and RT-PCR products were purified directly or after cutting of the band from an agarose gel using the QIAquick purification kit (Qiagen). Purified products were directly sequenced as described above. Sequences were aligned using the Sequencher 3.0 program (Genecodes). Any alterations found were verified by reamplification and sequencing of the product from the tumor and from the paired normal sample if the latter was available.

***In Vitro* Translation.** PCR amplification with adapter primers was used to clone all three ORFs of *ALT1* into the TA cloning vector (Invitrogen). Transformants were selected and sequenced to verify the absence of mutations in the insert. Inserts were then subcloned under the control of both the SP6 and T7 promoters in the pGEM T Easy vector (Promega, Madison, WI). *In vitro* translation experiments were performed in rabbit reticulocyte lysates with the TNT-coupled transcription/translation system (Promega) following the manufacturer's instructions and using C14 leucine as a label. *FHIT* and *TCL1* expression vectors (28, 29) were used as positive controls.

Results

General Sequence Landmarks. The sequenced fragment extends for 790 kb and covers the region between *D13S1150* on the centromeric end and *D13S25* on the telomeric end (Fig. 1). The 790 kb overlap partially with public database sequences (accession no. AL157367), as well as with published sequences (30, 31). The sequence has an overall G+C content of 39.57%. This is below the genome-wide average of 41% and represents a significant departure from the expected SD for a sequence of this size (approximately = –ˆ[(41)(59)/792,000)]% or ± 0.054% (26). Indeed, this sequence is

part of a larger (40 Mb) region of chromosome 13 extending from 13q14 to 13q22, between sequence-tagged site markers A005X38 and stsG30423, which has been reported as having only 36% G+C content (26). Regions of low G+C content have been observed to contain a relatively low content of genes (32, 33).

Overall, the sequence has the expected repeat content for short interspersed nuclear elements, LINEs, long terminal repeats, and DNA elements (Table 1 and Fig. 2) with a slightly lower content of ALU sequences than expected. However, when the most centromeric ~200 kb of the sequence is analyzed separately from the telomeric 600 kb, the concentration of ALU sequences in the centromeric end is twice as high as expected, and the LINE1 content is 45% below the expected proportion. The opposite holds true for the telomeric end, where LINE1 sequences are more common and longer than might be expected. The sequence includes four LINE1 elements of >4 kb in length (Fig. 2), two of which are full length and contain complete copies of both ORF1 and ORF2 (34, 35). The average length of inactive L1 elements in the genome is 900 bp, and only 1% of L1s are thought to be full length (32, 34, 35). We identified six CpG islands within the sequence (Fig. 2), two of which were reported previously (30). Three of the islands are located at the 5′ end of transcribed sequences (Fig. 2).

Identification of Genes and Transcribed Sequences within the Locus. Expressed sequences were identified by querying the human dbEST with overlapping fragments of our sequence ranging from 25 to 200 kb in length. GENESCAN, FGENESH, and GRAIL were used to identify potential genes by computer-aided gene prediction. Fig. 3 presents the annotated sequence with known genes, EST clusters, and predicted genes and exons. Five pseudogenes were identified within the sequence, including three previously reported ones (30). Four of the pseudogenes were derived from ribosomal proteins, RPL12 (accession no. P30050), RPL18 (accession no. AAA16329), RPL18a (accession no. NM 00980.1), and RPL34 (accession no. P49207), and one pseudogene was derived from the HSP70-interacting protein gene (accession no. AAH01236). Only seven UniGene clusters[15] were found that were formed of >10 ESTs; all were located within the centromeric 1/3 of the sequence. Five UniGene clusters with >10 ESTs correspond to known genes *RFP2/LEU5* (clusters HS.151428, Hs.246191, and Hs.114659), *LEU2* (Hs.43628), and *LEU1* (Hs.20149), and one cluster corresponds to hereby reported additional exons 6 and 7 of *LEU2* (Hs.192062). Sixteen UniGene clusters were found which were formed of <10 ESTs (Fig. 3). Only 1, Hs.246191, from *RPF2/LEU5*, had intervening genomic sequence suggesting the presence of an intron. In addition, 10 clusters (cluster A and C–K in Fig. 3) were created by overlap of ≥2 ESTs from the database but do not correspond to UniGene clusters. Table 2 gives the accession number for 1 EST in cluster A and C–K. Finally, BLAST analysis identified 58 single ESTs (data not shown) with >90% homology to our sequence. As shown in Fig. 3, FGENESH predicted nine genes, and GENESCAN predicted 17 genes within the 790-kb sequence. Although both programs predicted the presence of the only coding exon (exon 3) of *RFP2/LEU5* (genes B and 2 in Fig. 3), neither predicted the correct structure of the gene. Similarly, for *LEU2* and *LEU1*, although some predicted exons overlap with actual exons of the genes, and although the predicted direction of transcription is accurate, neither program predicted the correct gene structure. Both programs predicted exons overlapping partially with *CLLD3* and *CLLD4*, two transcripts in the region (see below and Fig. 3). To identify additional genes within the sequence, we performed BLAST searches by querying the mouse dbEST with our sequence. This search allowed us to identify a novel alternative first exon of *LEU2* which is highly conserved in mouse and is located within a C+G-rich region ~176 kb from the centromeric end of the sequence (Figs. 1–3). Neither GENESCAN nor FGENESH predicted this exon, nor were there any human ESTs homologous to this region.

Characterization of Genes within the Region. The three known genes in the locus, *RFP2/LEU5*, *LEU2*, and *LEU1*, have been extensively described, and mutation studies have not revealed the presence of alterations in any of these genes regardless of the presence or absence of 13q14 deletions in the leukemias studied (20, 30). UniGene clusters that did not overlap with known genes were tested for expression by RT-PCR amplification with specific primers on DNase-treated cDNA. Six clusters showed amplification products and were tested additionally by Northern blotting. Three of these (Hs.42321, Hs.5244, and Hs.182953) corresponding to *CLLD4* (for chronic lymphocytic leukemia deleted 4) and *CLLD3* (for chronic lymphocytic leukemia deleted 3), respectively, in Fig. 3 showed hybridization to Northern blots and overlapped with predicted exons from both FGENSH and GENESCAN. Therefore, cDNA libraries were screened to identify any additional exons and to clone the full-length gene. In both cases, the cDNA clones isolated were colinear with the genomic DNA sequence. *CLLD3* cDNA clones contained three putative ORFs of 88, 138, and 151 amino acids (in order, 5′ to 3′) with no significant homology to known proteins. *CLLD4* cDNA clones contain a 138 amino acid ORF, a portion of which shows homology to potassium channel proteins. Interestingly, a rat EST and a mouse EST (BF543451 and BB60593, respectively) with homology to potassium channel proteins showed >80% similarity to our sequence. Because of these features, mutation analysis of all putative ORFs of *CLLD3* and *CLLD4*, as well as for the LEU5 ORF was undertaken, but no alterations were found in any of the >70 leukemias analyzed. Because of the low levels of expression and the lack of introns in the *CLLD4* and *CLLD3* transcripts, it is possible that they represent basal transcription products.

Sequence Comparison to Mouse ESTs. As mentioned above, a query against the mouse dbEST identified mouse ESTs homologous to a G+C-rich segment of our sequence. The human segment was amplified by PCR, and the product was used as a probe on Northern blots containing RNAs from different human tissues. The probe hybridized to a ubiquitously expressed 1.8-kb mRNA, suggesting that it corresponded to a functional gene (Fig. 4*A*). cDNA libraries from spleen, fetal liver, and kidney were screened to isolate the full-length cDNA. Several clones were identified and characterized. Three clones were full length and contained, at the 5′ end, sequences derived from the human fragment conserved in mouse as well as exons 2–5 of *LEU2* and two additional novel 3′ exons (see cluster Hs.192062). In this novel transcript, exon 5 of *LEU2* was spliced at an alternative splice donor site 118 bp from its 5′ end (Fig. 5*B*). These results suggest that the 1.8-kb mRNA is derived from utilization of a transcription start site located upstream of the previously reported first exon of *LEU2*. The structure of this transcript, which we have named *ALT1* (accession no. AF380424), is shown in Fig. 5*A* compared with that of *LEU2*. Both *ALT1* and *LEU2* are highly conserved in mouse with *LEU2* showing 76% identity in exons 1, 2, and 3 with human *LEU2* at the DNA level. *ALT1* exon 1 is very highly conserved, with 95 bp at the 5′ end of the mRNA showing 91% identity and the remainder of exons 1, 2, and 3 showing again 76% identity at the DNA level (Fig. 5*A*). The proposed mouse sequences for both Alt1 (accession no. AF380423) and Leu2 (accession no. AF380425) derive from EST sequences downloaded from mouse EST UniGene cluster Mm.25679 and assembled using Sequencher 4.5. Complete sequencing of the human *ALT1* mRNA revealed the presence of the previously reported 55 amino acid ORF derived from exons 4 and 5 of *LEU2*,

[15] Internet address: http://www.ncbi.nlm.nih.gov/UniGene/.

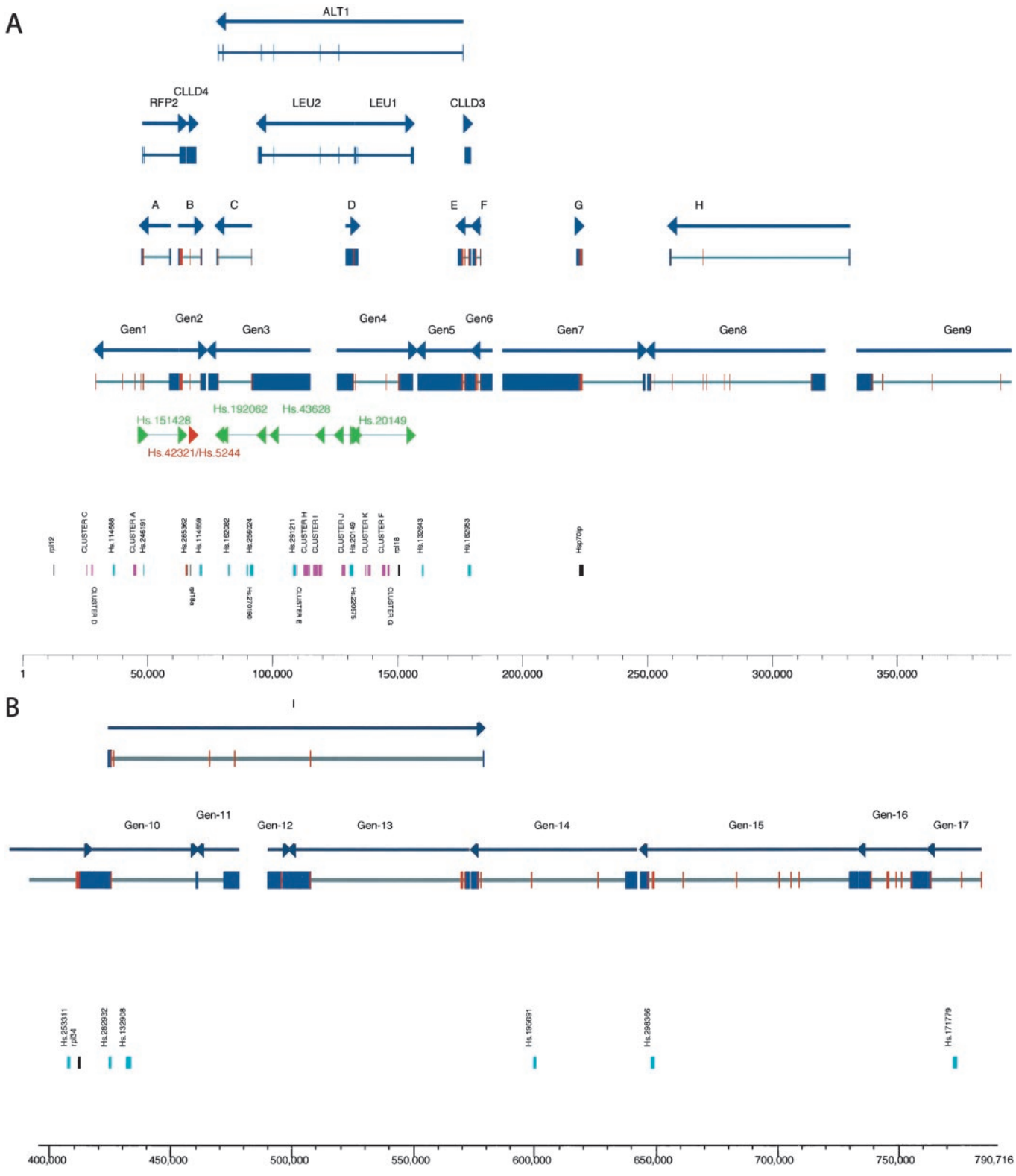

Fig. 3. Sequence annotation of the 13q14 tumor suppressor locus (base 1 is at the centromeric end of the sequence). *A*, centromeric end, 390 kb. *B*, telomeric end, 400 kb. The direction of *arrows* indicates the direction of transcription. *Blue arrows* span the genomic locus of each gene from 5′ UTR to 3′ UTR. *Dark blue boxes*, noncoding GENESCAN and FGENESH predicted exons; *red boxes*, coding GENESCAN and FGENESH predicted exons; *green arrows*, UniGene clusters of >10 ESTs that correspond to known genes in the region. *Red arrow*, UniGene cluster of >10 ESTs not deriving from a known gene; *light-blue boxes*, UniGene clusters of <10 ESTs; *pink boxes*, novel clusters constructed on the basis of the overlap of ≥2 ESTs (see Table 2); *black boxes*, pseudogenes.

which shows neither homology to known proteins nor conservation in the mouse mRNAs. We identified an ORF at the 5′ end of the cDNA clones, suggesting that perhaps this ORF represented the bonafide translation unit. Five prime RACE experiments using bone marrow and testis cDNA were performed which extended the sequence ~150 nucleotides 5′ of the cloned cDNA sequence. The RACE sequence was colinear with the genomic DNA immediately 5′ of the cDNA and was thus derived from the same exon. No additional ORFs became

Table 2 *Novel clusters (see Fig. 3) and EST accession no. at 13q14*

Cluster name	Accession no.
Cluster A	H96691
Cluster C	BE241584
Cluster D	AA324052
Cluster E	BF350686
Cluster F	BE066024
Cluster G	AW963844
Cluster H	AA991380
Cluster I	AV761241
Cluster J	AV712385
Cluster K	AW970604

apparent when the full-length cDNA was analyzed. *In vitro* translation experiments with both *LEU2* and *ALT1* were negative, suggesting that neither of these two mRNAs codes for a functional protein product. To determine whether exons 6 and 7 were expressed exclusively as part of *ALT1*, we screened the same cDNA libraries as above with a *LEU2* exon 1 probe. One clone isolated from a peripheral blood leukocyte library contained *LEU2* exon 1, as well as exons 2–7, indicating that exon 1a is the only exon specific to *ALT1*. In addition to this transcript, several other transcripts with polyadenylation tails at exon 4 as well as a clone lacking exon 2 and amplification products lacking exon 6, were isolated from the libraries, indicating that *LEU2*/*ALT1* has several splice variants originating at either of the transcription start points. No mutations were found at the DNA level in 100 tumors analyzed with and without 13q14 loss.

As reported previously (36), regions of significant homology with mouse ESTs were found only for the *RFP2/LEU5* gene and for *LEU2*/*ALT1* only, suggesting that these are the only conserved sequences within the 790 kb reported in this work.

Expression Analysis. Northern blot analysis on normal human tissues revealed that the *LEU1* message is detectable in most tissues but not in peripheral blood lymphocytes, whereas *LEU2* shows very low expression levels in most tissues, as well as higher expression in lymphoid tissues, such as thymus, bone marrow, fetal liver, and peripheral blood lymphocytes, and *ALT1* is ubiquitously expressed (Fig. 4). Because an expressed pseudogene of *LEU2* is located on chromosome 1 (37), the hybridization signal seen on Northern blots could be attributable to the pseudogene mRNA, which is apparently expressed at similar levels to the *13q14* gene. Indeed, our Southern analysis with an *ALT1*-specific probe suggests that the first exon of *ALT1* is also duplicated, although the chromosomal location of the duplicated segment is not known. Our expression analysis of *LEU1*, *LEU2*, and *ALT1* in CLL cases and lymphoid tumor cell lines suggests that the expression levels of these genes cannot be directly correlated with the presence or absence of 13q14 loss in the tumors.

Discussion

We report the complete sequence of a 790-kb region of 13q14, between *D13S1150* and *D13S25*, thought to harbor a tumor suppressor gene playing a crucial role in CLL pathogenesis. The sequence partially overlaps with public database releases as of the genome draft as well as with sequence segments published previously by us *et al.* (30, 31). Our analysis of general sequence landmarks has revealed a marked inequality in the distribution of repeat elements between the centromeric and telomeric ends of the sequence. In fact, ALU sequences are twice as common as expected on the most centromeric ~200 kb, whereas LINE1 repeats are half as common as expected (Table 1). On the telomeric end of the sequence, ALUs are half as common as expected, whereas the proportion of LINE1 repeats is higher than expected. In particular, LINE1 sequences at the telomeric end are longer and more common than expected. One of two full-length LINE1 sequences (at 231 kb in Fig. 2) belongs to the PA2 family, which is the youngest inactive LINE1 family (34, 35), suggesting that the region contained active transposons relatively recently in terms of the evolution of the genome. Within the FRA3B fragile site, the most common constitutive fragile site in the human genome, there is significant variation in the distribution of repeat elements along the length of the fragile site, with the highest concentration (>21%) occurring within intron 4 where most cancer breaks occur (38). In addition, within the FRA3B locus, there are three LINE1 elements longer than 4 kb but no full-length elements.[16] For comparison, within a 1.3-Mb region of chromosome 22 (1,314,740 bp from accession no. NT_011523), which was selected at random for analysis, no LINE1 elements of >4 kb were found. The association between LINE1 and other transposable elements with chromosomal breaks found in cancer is well documented and has been postulated to mediate breakage and repair of chromosome ends within the *FHIT* locus, which spans the *FRA3B* fragile site and is the target of numerous deletions in cancer (38). The human genome has been found to have wide variations in repeat content, such as the reported 525 kb of chromosome X with an 89% repeat content (32). The human genome also contains a region larger than 100 kb with ≥56% ALU (32). The variation in the frequency and distribution of repeat frequencies described for this 790-kb segment of 13q14 is not as extreme as the ones mentioned above. However, the presence of relatively long LINE1 sequences and the concentration of ALUs within the 200 kb at the centromeric end may compromise the stability of the region and could provide the means for inter- or intra-chromosomal recombination that might cause the deletions observed in CLL and other tumors.

Several groups have used a variety of methods to identify genes within this 13q14 locus, ranging from exon trapping and cDNA selection to computer-assisted gene identification. Thus far, three genes have been identified between *D13S273* and *D13S25*: RFP2/*LEU5* (17) is homologous to zinc-finger proteins involved in cancer but shows no alterations in B-CLL (this study and Ref. 30). *LEU1* and *LEU2*/*ALT1* (Ref. 20 and this study) represent two transcriptional units located within a homozygous deletion in CLL. The genes are transcribed in opposite orientation and are separated by only 200 bp. *LEU1* is transcribed into a 1.1-kb mRNA, whereas *LEU2* shows a complex transcription pattern with major transcripts of 1.4 and 1.8 kb. Sequence homology searches against the human and mouse dbESTs, and the use of exon prediction software has allowed us to identify three novel exons of the *LEU2* gene, including a highly conserved G+C-rich alternative first exon. The hereby-reported mutational analysis of these novel exons, as well as previously reported analyses (20, 30) of known exons, have not revealed the presence of mutations in CLL cases. Expression analysis of these genes has also not revealed a clear correlation between level of expression and 13q14 loss (Refs. 30 and 37 and this study). The function of both *LEU1* and *LEU2*/*ALT1* remains unknown. Because of the lack of experimental evidence indicating that these mRNAs code for a functional protein, it is thought that they function as noncoding RNAs. However, although *LEU2* is conserved in mouse, neither the *LEU2* nor the *LEU1* nucleotide sequences share homology with other noncoding RNAs. Therefore, their role, if any, in the pathogenesis of CLL remains unknown.

Our finding of a dysequilibrium in the quality and quantity of transposable elements within the 13q14 locus, combined with the apparent paucity of genes in the region, could suggest that 13q14 deletions may not have a pathogenetic role in CLL. However, the presence of homozygous as well as heterozygous deletions and the high frequency of 13q14 loss in a tumor characterized by its relative genomic stability, as well as the clonal character of deletions, indicate

[16] F. Bullrich, unpublished observations.

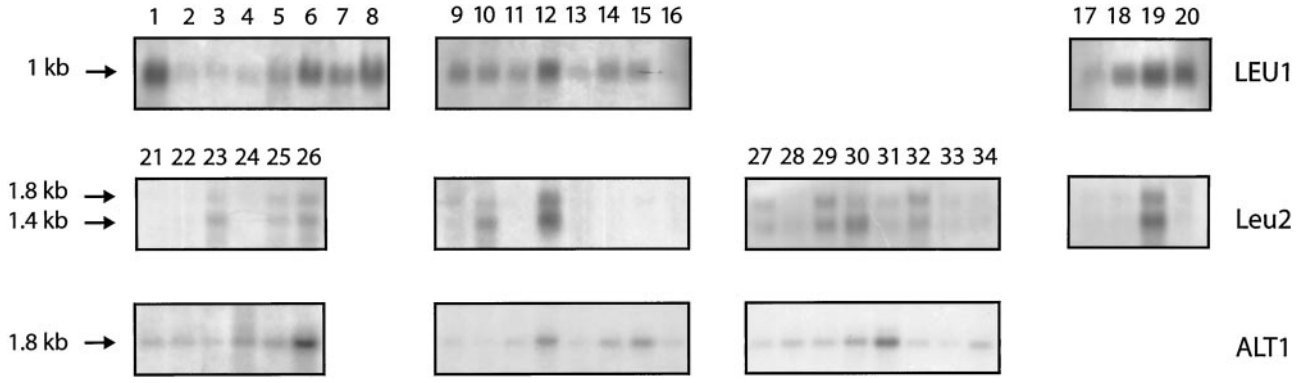

Fig. 4. Expression of *LEU1* (*top row*), *LEU2* (*middle row*), and *ALT1* (*bottom row*) on human tissues. *Lane 1*, heart; *Lane 2*, brain; *Lane 3*, placenta; *Lane 4*, lung; *Lane 5*, liver; *Lane 6*, skeletal muscle; *Lane 7*, kidney; *Lane 8*, pancreas; *Lane 9*, spleen; *Lane 10*, thymus; *Lane 11*, prostate; *Lane 12*, testis; *Lane 13*, ovary; *Lane 14*, small intestine; *Lane 15*, colon; *Lane 16*, peripheral blood leukocytes; *Lane 17*, fetal brain; *Lane 18*, fetal lung; *Lane 19*, fetal liver; *Lane 20*, fetal kidney; *Lane 21*, spleen; *Lane 22*, lymph node; *Lane 23*, thymus; *Lane 24*, peripheral blood leukocytes; *Lane 25*, bone marrow; *Lane 26*, fetal liver; *Lane 27*, promyelocytic leukemia cell line HL-60; *Lane 28*, cervical cancer cell line HeLa S3; *Lane 29*, chronic myelogenous leukemia cell line K562; *Lane 30*, acute lymphocytic leukemia cell line MOLT-4; *Lane 31*, Burkitt's lymphoma cell line Raji; *Lane 32*, colorectal adenocarcinoma cell line SW480; *Lane 33*, lung carcinoma cell line A549; *Lane 34*, melanoma cell line G-361.

that 13q14 deletions are indeed of pathogenetic significance. It is possible that the 13q14 tumor suppressor is normally expressed during a small window of B-cell development and that it is therefore not represented in current versions of the public databases. In addition, the pattern of allelic loss at 13q14 in CLL, with interrupted deletions (9, 14–16) and more than one core region of loss, may be attributable to genetic heterogeneity within the tumors or to the characteristics of the tumor suppressor gene itself. Most deletions described in B-CLL are relatively large and extend for >500 kb from the vicinity of *D13S272* to *D13S25*. A large gene with relatively few exons separated by large introns could produce a pattern of loss similar to that observed in CLL. Definition of the region of loss in other malignancies, particu-

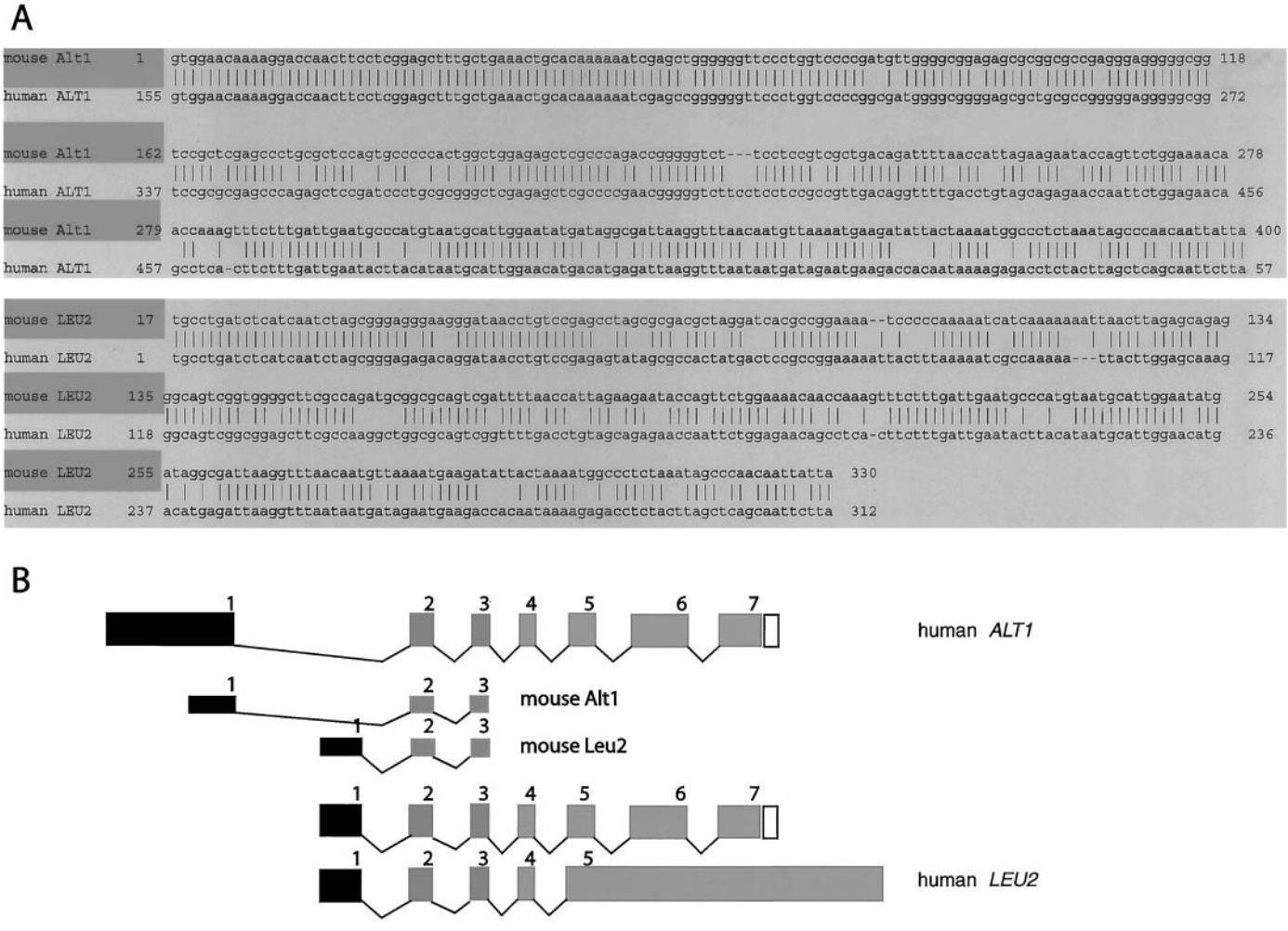

Fig. 5. Sequence homology (*A*) and exon structure (*B*) of human and mouse *LEU2* and *ALT1*. Conserved exons are depicted as *shaded boxes*. *Black*, >90% identity; *dark gray*, 76% identity; *light gray*, no sequences in the mouse dbEST.

larly in multiple myeloma, as well as the availability of the DNA sequence of the region involved, will undoubtedly aid in the identification of the critical gene.

Acknowledgments

We thank Shashi Rattan and Masaioshi Shimizu for expert technical assistance and Dr. T. Burland of DNAstar Corp. for his valuable input regarding sequence annotation with GenVision and GeneQuest.

References

1. Freeman, A. S., and Nadler, L. M. Malignancies of lymphoid cells. *In:* A. S. Fauci, E. Braunwald, K. J. Isselbacher, J. D. Wilson, J. B. Martin, D. L. Kasper, S. L. Hauser, and D. L. Longo (eds.), Harrisson's Principles of Internal Medicine, Ed. 14, pp. 695–700. New York: McGraw-Hill, Inc., 1998.
2. Tsujimoto, Y., Yunis, J., Onorato-Showe, L., Nowell, P. C., and Croce, C. M. Molecular cloning of the chromosomal breakpoint of B cell lymphomas and leukemias with the t(11;14) chromosome translocation. Science (Wash. DC), *224:* 1403–1406, 1984.
3. Tsujimoto, Y., Jaffe, E., Cossman, J., and Croce, C. M. Clustering of breakpoints on chromosome 11 in human B-cell neoplasms with the t(11;14) chromosome translocation. Nature (Lond.), *315:* 340–343, 1985.
4. Tsujimoto, Y., Finger, L. R., Yunis, J., Nowell, P. C., and Croce, C. M. Cloning of the chromosome breakpoint of neoplastic B-cells with the t(14;18) chromosome translocation. Science (Wash. DC), *226:* 1097–1099, 1984.
5. Tsujimoto, Y., Cossman, J., Jaffe, E., and Croce, C. M. Involvement of the *bcl-2* gene in human follicular lymphoma. Science (Wash. DC), *228:* 1440–1443, 1985.
6. Dohner, H., Stilgenbauer, S., Dohner, K., Bentz, M., and Lichter, P. Chromosome aberrations in B-cell chronic lymphocytic leukemia: reassessment based on molecular cytogenetic analysis. J. Mol. Med., *77:* 266–281, 1999.
7. Bigoni, R., Cuneo, A., Roberti, M. G., Bardi, A., Rigolin, G. M., Piva, N., Scapoli, G., Spanedda, R., Negrini, M., Bullrich, F., Veronese, M. L., Croce, C. M., and Castoldi, G. Chromosome aberrations in atypical chronic lymphocytic leukemia: a cytogenetic and interphase cytogenetic study. Leukemia, *11:* 1933–1940, 1997.
8. Stilgenbauer, S., Nickolenko, J., Wilhelm, J., Wolf, S., Weitz, S., Dohner, K., Boehm, T., Dohner, H., and Lichter, P. Expressed sequences as candidates for a novel tumor suppressor gene at band 13q14 in B-cell chronic lymphocytic leukemia and mantle cell lymphoma. Oncogene, *16:* 1891–1897, 1998.
9. Cuneo, A., Bigoni, R., Rigolin, G. M., Roberti, M. G., Bardi, A., Piva, N., Milani, R., Bullrich, F., Veronese, M. L., Croce, C., Birg, F., Dohner, H., Hagemeijer, A., and Castoldi, G. Cytogenetic profile of follicle mantle lineage: correlation with clinicobiologic features. Blood, *93:* 1372–1380, 1999.
10. Zojer, N., Königsberg, R., Ackermann, J., Fritz, E., Dallinger, S., Krömer, E., Kaufmann, H., Riedl, L., Gisslinger, H., Schreiber, S., Heinz, R., Huber, H., and Drach, J. Deletion of 13q remains an independent adverse prognostic variable in multiple myeloma despite its frequent detection by interphase fluorescence *in situ* hybridization. Blood, *95:* 1925–1930, 2000.
11. Desikan, R., Barlogie, B., Sawyer, J., Ayers, D., Tricot, G., Badros, A., Zangari, M., Munshi, N. C., Anaissie, E., Spoon, D., Siegel, D., Jagannath, S., Vesole, D., Epstein, J., Shaughnessy, J., Fassas, A., Lim, S., Roberson, P., and Crowley, J. Results of high-dose therapy for 1000 patients with multiple myeloma: durable complete remissions and superior survival in the absence of chromosome 13 abnormalities. Blood, *95:* 4008–4010, 2000.
12. Liu, Y., Hermanson, M., Grander, D., Merup, M., Wu, X., Heyman, M., Rasool, O., Juliusson, G., Gahrton, G., Detlofsson, R., Nikiforova, N., Buys, C., Söderhäll, S., Yankovsky, N., Zabarovsky, E., and Einhorn, S. 13q deletions in lymphoid malignancies. Blood, *86:* 1911–1915, 1995.
13. Stilgenbauer, S., Leupolt, E., Ohl, S., Weiss, G., Schroder, M., Fischer, K., Bentz, M., Lichter, P., and Dohner, H. Heterogeneity of deletions involving RB-1 and the D13S25 locus in B-cell chronic lymphocytic leukemia revealed by fluorescence *in situ* hybridization. Cancer Res., *55:* 3475–3477, 1995.
14. Bullrich, F., Veronese, M. L., Kitada, S., Jurlander, J., Caligiuri, M. A., Reed, J. C., and Croce, C. M. Minimal region of loss at 13q14 in B-cell chronic lymphocytic leukemia. Blood, *88:* 3109–3115, 1996.
15. Kalachikov, S., Migliazza, A., Cayanis, E., Fracchiolla, N. S., Bonaldo, M. F., Lawton, L., Jelenc, P., Ye, X., Qu, X., Chien, M., Hauptschein, R., Gaidano, G., Vitolo, U., Saglio, G., Resegotti, L., Brodjansky, V., Yankovsky, N., Zhang, P., Soares, M. B., Russo, J., Edelman, I. S., Efstratiadis, A., Dalla-Favera, R., and Fischer, S. G. Cloning and gene mapping of the chromosome 13q14 region deleted in chronic lymphocytic leukemia. Genomics, *42:* 369–377, 1997.
16. Bouyge-Moreau, I., Rondeau, G., Avet-Loiseau, H., Andre, M. T., Bezieau, S., Cherel, M., Saleun, S., Cadoret, E., Shaikh, T., De Angelis, M. M., Arcot, S., Batzer, M., Moisan, J. P., and Devilder, M. C. Construction of a 780-kb PAC, BAC, and cosmid contig encompassing the minimal critical deletion involved in B cell chronic lymphocytic leukemia at 13q14.3. Genomics, *46:* 183–190, 1997.
17. Kapanadze, B., Kashuba, V., Baranova, A., Rasool, O., van Everdink, W., Liu, Y., Syomov, A., Corcoran, M., Poltaraus, A., Brodyansky, V., Syomova, N., Kazakov, A., Ibbotson, R., van den Berg, A., Gizatullin, R., Fedorova, L., Sulimova, G., Zelenin, A., Deaven, L., Lehrach, H., Grander, D., Buys, C., Oscier, D., Zabarovsky, E. R., Einhorn, S., and Yankovsky, N. A cosmid and cDNA fine physical map of a human chromosome 13q14 region frequently lost in B-cell chronic lymphocytic leukemia and identification of a new putative tumor suppressor gene, *Leu5*. FEBS Lett., *426:* 266–270, 1998.
18. Bezieau, S., Devilder, M. C., Rondeau, G., Cadoret, E., Moisan, J. P., and Moreau, I. Assignment of 48 ESTs to chromosome 13 band q14.3 and expression pattern for ESTs located in the core region deleted in B-CLL. Genomics, *52:* 369–373, 1998.
19. Corcoran, M. M., Rasool, O., Liu, Y., Iyengar, A., Grander, D., Ibbotson, R. E., Merup, M., Wu, X., Brodyansky, V., Gardiner, A. C., Juliusson, G., Chapman, R. M., Ivanova, G., Tiller, M., Gahrton, G., Yankovsky, N., Zabarovsky, E., Oscier, D. G., and Einhorn, S. Detailed molecular delineation of 13q14.3 loss in B-cell chronic lymphocytic leukemia. Blood, *91:* 1382–1390, 1998.
20. Liu, Y., Corcoran, M., Rasool, O., Ivanova, G., Ibbotson, R., Grander, D., Iyengar, A., Baranova, A., Kashuba, V., Merup, M., Wu, X., Gardiner, A., Mullenbach, R., Poltaraus, A., Hultstrom, A. L., Juliusson, G., Chapman, R., Tiller, M., Cotter, F., Gahrton, G., Yankovsky, N., Zabarovsky, E., Einhorn, S., and Oscier, D. Cloning of two candidate tumor suppressor genes within a 10 kb region on chromosome 13q14, frequently deleted in chronic lymphocytic leukemia. Oncogene, *15:* 2463–2473, 1997.
21. Fleischmann, R. D., Adams, M. D., White, O., Clayton, R. A., Kirkness, E. F., Kerlavage, A. R., Bult, C. J., Tomb, J. F., Dougherty, B. A., Merrick, J. M., *et al.* Whole-genome random sequencing and assembly of *Haemophilus influenzae* Rd. Science (Wash. DC), *269:* 496–512, 1995.
22. Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. Basic local alignment search tool. J. Mol. Biol., *215:* 403–410, 1990.
23. Smit, A. F. The origin of interspersed repeats in the human genome. Curr. Opin. Genet. Dev., *6:* 743–748, 1996.
24. Xu, Y., Einstein, J. R., Mural, R. J., Shah, M., and Uberbacher, E. C.(eds.). An Improved System for Exon Recognition and Gene Modeling in Human DNA Sequences, pp. 376–384. Menlo Park, CA: Press, A. AAAI, 1994.
25. Solovyev, V. V., Salamov, A. A., and Lawrence, C. B. Predicting internal exons by oligonucleotide composition and discriminant analysis of spliceable open reading frames. Nucleic Acids Res., *22:* 5156–5163, 1994.
26. Burge, C., and Karlin, S. Prediction of complete gene structures in human genomic DNA. J. Mol. Biol., *268:* 78–94, 1997.
27. Schwartz, S., Zhang, Z., Frazer, K. A., Smit, A., Riemer, C., Bouck, J., Gibbs, R., Hardison, R., and Miller, W. PipMaker—A Web Server for Aligning Two Genomic DNA Sequences. Genome Res., *10:* 577–586, 2000.
28. Pekarsky, Y., Koval, A., Hallas, C., Bichi, R., Tresini, M., Malstrom, S., Russo, G., Tsichlis, P., and Croce, C. M. Tcl1 enhances Akt kinase activity and mediates its nuclear translocation. Proc. Natl. Acad. Sci. USA, *97:* 3028–3033, 2000.
29. Barnes, L. D., Garrison, P. N., Siprashvili, Z., Guranowski, A., Robinson, A. K., Ingram, S. W., Croce, C. M., Ohta, M., and Huebner, K. Fhit, a putative tumor suppressor in humans, is a dinucleoside 5′,5′-P1, P3-triphosphate hydrolase. Biochemistry, *35:* 11529–11535, 1996.
30. Migliazza, A., Bosch, F., Komatsu, H., Cayanis, E., Martinotti, S., Toniato, E., Guccione, E., Qu, X., Chien, M., Murty, V. V., Gaidano, G., Inghirami, G., Zhang, P., Fischer, S., Kalachikov, S. M., Russo, J., Edelman, I., Efstratiadis, A., and Dalla-Favera, R. Nucleotide sequence, transcription map, and mutation analysis of the 13q14 chromosomal region deleted in B-cell chronic lymphocytic leukemia. Blood, *97:* 2098–2104, 2001.
31. Mabuchi, H., Fujii, H., Calin, G., Alder, H., Negrini, M., Rassenti, L., Kipps, T. J., Bullrich, F., and Croce, C. M. Cloning and characterization of CLLD6, CLLD7, and CLLD8, novel candidate genes for leukemogenesis at chromosome 13q14, a region commonly deleted in B-cell chronic lymphocytic leukemia. Cancer Res., *61:* 2870–2877, 2001.
32. International Human Genome Consortium. Initial sequencing and analysis of the human genome. Nature (Lond.), *409:* 860–921, 2001.
33. Venter, J. C., Adams, M. D., Myers, E. W., Li, P. W., Mural, R. J., Sutton, G. G., Smith, H. O., Yandell, M., Evans, C. A., Holt, R. A. *et al.* The sequence of the human genome. Science (Wash. DC), *291:* 1304–1351, 2001.
34. Sassaman, D. M., Dombroski, B. A., Moran, J. V., Kimberland, M. L., Naas, T. P., DeBerardinis, R. J., Gabriel, A., Swergold, G. D., and Kazazian, H. H., Jr. Many human L1 elements are capable of retrotransposition. Nat. Genet., *16:* 37–43, 1997.
35. Luning Prak, E .T., and Kazazian, H. H. Mobile elements and the human genome. Nat. Rev. Genet., *1:* 134–144, 2000.
36. Kapanadze, B., Makeeva, N., Corcoran, M., Jareborg, N., Hammarsund, M., Baranova, A., Zabarovsky, E., Vorontsova, O., Merup, M., Gahrton, G., Jansson, M., Yankovsky, N., Einhorn, S., Oscier, D., Grandér, D., and Sangfelt, O. Comparative sequence analysis of a region on human chromosome 13q14, frequently deleted in b-cell chronic lymphocytic leukemia, and its homologous region on mouse chromosome. Genomics, *70:* 327–334, 2000.
37. Mertens, D., Wolf, S., Bullinger, L., Ohl, S., Schaffner, C., Dohner, H., Stilgenbauer, S., and Lichter, P. BCMSUN, a candidate gene for B-cell chronic lymphocytic leukemia and mantle-cell lymphoma, has an independently expressed homolog on 1p22–p31, BCMSUN-like. Int. J. Cancer, *88:* 692–697, 2000.
38. Mimori, K, Druck, T., Inoue, H., Alder, H., Berk, L., Mori, M., Huebner, K., and Croce, C. M. Cancer-specific chromosome alterations in the constitutive fragile region FRA3B. Proc. Natl. Acad. Sci. USA, *96:* 7456–7461, 1999.