Genomics **70,** 327–334 (2000)
doi:10.1006/geno.2000.6386, available online at http://www.idealibrary.com on IDEAL®

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 14

Bagrat Kapanadze,*,†,[1] Natalia Makeeva,*,†,[1] Martin Corcoran,‡ Niclas Jareborg,§ Marianne Hammarsund,* Ancha Baranova,† Eugene Zabarovsky,¶ Olga Vorontsova,¶ Mats Merup,‖ Gösta Gahrton,‖ Monika Jansson,‖ Nick Yankovsky,† Stefan Einhorn,* David Oscier,‡ Dan Grandér,*,[2] and Olle Sangfelt*

**Radiumhemmets Research Laboratory, CCK, R8:03, S-171 76, Karolinska Hospital, Stockholm, Sweden; ‡Molecular Biology Laboratory, Royal Bournemouth Hospital, Bournemouth, BH7 7DW United Kingdom; †Genome Analysis Laboratory, Institute of General Genetics, Russian Academy of Science, Moscow 117 809, Russia; §Center for Genomics Research, and ¶Microbiology and Tumor Biology Center and Center for Genomics Research, Karolinska Institute, S-171 77, Stockholm, Sweden; and ‖Department of Hematology and Medicine, Huddinge Hospital, 141 86, Huddinge, Sweden*

Received July 18, 2000; accepted September 14, 2000

Previous studies have indicated the presence of a putative tumor suppressor gene on human chromosome 13q14, commonly deleted in patients with B-cell chronic lymphocytic leukemia (B-CLL). We have recently identified a minimally deleted region encompassing parts of two adjacent genes, termed LEU1 and LEU2 (leukemia-associated genes 1 and 2), and several additional transcripts. In addition, 50 kb centromeric to this region we have identified another gene, LEU5/RFP2. To elucidate further the complex genomic organization of this region, we have identified, mapped, and sequenced the homologous region in the mouse. Fluorescence *in situ* hybridization analysis demonstrated that the region maps to mouse chromosome 14. The overall organization and gene order in this region were found to be highly conserved in the mouse. Sequence comparison between the human deletion hotspot region and its homologous mouse region revealed a high degree of sequence conservation with an overall score of 74%. However, our data also show that in terms of transcribed sequences, only two of those, human LEU2 and LEU5/RFP2, are clearly conserved, strengthening the case for these genes as putative candidate B-CLL tumor suppressor genes.

INTRODUCTION

B-cell chronic lymphocytic leukemia (B-CLL) is the most frequent type of leukemia in adults of the western countries, but the molecular pathogenesis of B-CLL remains largely unknown (O'Brien *et al.*, 1995). Cytogenetic analysis has shown recurring abnormalities of which the most common are trisomy 12 and structural chromosome abberations involving chromosome 13q (Juliusson and Gahrton, 1993; Oscier, 1994). We have recently defined a minimally deleted region of 130 kb centromeric to the marker D13S272 on chromosome 13q (Corcoran *et al.*, 1998), and in an extension of these studies, we identified two neighbouring genes, termed LEU1 and LEU2, covering a major deletion hotspot of no more than 10 kb in close proximity to the marker D13S319 (Liu *et al.*, 1997). Despite extensive mutational analysis of both LEU1 and LEU2, no intragenic mutations could be detected in hemizygously deleted patients in more than 170 B-CLL samples (Liu *et al.*, 1997). Subsequent genomic sequence analysis of this hotspot deletion region identified several additional ESTs, two of which were termed LEU3 and LEU4 (Liu *et al.*, 1997). In addition, 50–60 kb centromeric to this area we previously cloned another gene, termed LEU5 (Liu *et al.*, 1997; Kapanadze *et al.*, 1998). This gene, now denoted RFP2, contains an open reading frame (ORF) of 407 amino acids that shares strong homology with some proteins taking part in early embryogenesis and malignant transformation (Kapanadze *et al.*, 1998). Moreover, we have also identified and isolated a number of CpG islands as markers for the probable location of genes in the area (Corcoran *et al.*, 1998).

Cross-species DNA sequence comparisons are being used increasingly to address structural and evolutionary questions, but also to identify sequences likely to encode important biological functions. Several studies

[1] These authors contributed equally to the work.

[2] To whom correspondence should be addressed. Telephone: (46) (8) 5177 6262. Fax: (46) (8) 339031. E-mail: dan.grander@cck.ki.se.

0888-7543/00 $35.00

that have compared large segments of human and rodent sequences indicate gene-coding and regulatory domains as generally well conserved, in contrast to noncoding (intronic and intergenic) domains (Hood *et al.*, 1993; Koop and Hood, 1994; Koop, 1995; Hardison *et al.*, 1997; Oeltjen *et al.*, 1997). Moreover, some of the noncoding sequence conservation could be attributable to constraints by splicing, chromatin condensation, matrix association, and replication origins (Koop, 1995; Oeltjen *et al.*, 1997).

Because critical genes, including those involved in malignant transformation, tend to be highly conserved between mouse and human, direct cross-species sequence analysis could potentially aid in the identification of putative candidate tumor suppressor genes. To elucidate further the extremely EST/gene-rich region on human chromsosome 13q thought to harbor a tumor suppressor gene involved in CLL and other malignancies, we decided to compare directly the human major deletion hotspot and its flanking regions with that of the mouse. In addition, these data may enable the elucidation of conserved intronic/intergenic regions of possible importance in the transcriptional regulation of these genes.

MATERIALS AND METHODS

General methods. All basic molecular biology and microbiology procedures (isolation of DNA, growing of bacteria, cloning, etc.) were performed according to standard procedures (Sambrook *et al.*, 1989). A human genomic CpG island, 9E4.3, previously cloned and isolated by us (Corcoran *et al.*, 1998), was subcloned and partially sequenced during this analysis (Corcoran *et al.*, manuscript in preparation).

PAC library screening, subcloning, and creation of sublibraries. Mouse PAC library RPCI21 was screened with the following cDNA and genomic probes: mouse LEU2 cDNA (GenBank Accession No. AA171261), a cDNA fragment of the human LEU5/RFP2 coding exon (Kapanadze *et al.*, 1998), and the human genomic fragment 9E4.3 (Corcoran *et al.*, 1998). Positive PAC clones were purified according to standard protocols, cleaved with *Eco*RI and *Hin*dIII, and analyzed by conventional hybridization techniques with the same probes used for library screening. In addition, these PACs were also hybridized with human LEU1 and LEU2 cDNA probes (Liu *et al.*, 1997) and a human LEU4 EST (GenBank Accession No. AA281817). Based on these hybridizations, the following PACs were chosen for further analysis: 454-C18, 602-N13, and 588-K23. The appropriate enzymes and a combination of the two, *Eco*RI, *Pst*I, and *Eco*RI/*Pst*I, were used for subcloning fragments into pBS II plasmid vectors. Sublibraries (grids of >400 clones) were created from each enzyme combination, as previously described (Kapanadze *et al.*, 1996), and further screened with the above-mentioned probes.

PCR and sequencing. Human and mouse RNA (mRNA and total RNA) was purchased directly from Clontech (Palo Alto, CA), and cDNA was prepared with first-strand beads (Ready-To-Go), according to the manufacturer's protocol (Amersham Pharmacia Biotech Inc., Uppsala, Sweden). LEU5/RFP2 and LEU2 were amplified with the following primers: RFP2F, 5′-CAGCTCCATTCAAGTGTCCTACATGCC-3′, and RFP2R, 5′-GACCATTCTAGGAAATGGTAGGACC-3′; and LEU2F, 5′-GTTCCCTGGTCCCCGATGTTGG-3′, and LEU2R, 5′-CCTTAATCGCCTATCATATTCCAATG-3′. PAC and mouse genomic DNA was used to amplify a 3.0-kb PCR product with the primers L2, 5′-GTGCCGGGAGAACTATATCAGCACC-3′, and LEU2A/9, 5′-CTAGGATCACGCCGGAAAATCCC-3′, using the Advantage-GC 2 PCR kit (Clontech) with cycling parameters of 95°C for 3 min, 1 cycle; 95°C for 25 s, 65°C for 20 s, 68°C for 3 min, 35 cycles; 68°C for 7 min, 1 cycle. All PCRs were performed in a GeneAmp 2400 thermal cycler (Perkin–Elmer Biosystems, Palo Alto, CA). Cycling parameters for the amplification of cDNA were 95°C for 30 s, 1 cycle; 94°C for 10 s, 65°C for 20 s, 68°C for 1 min, 5 cycles; 94°C for 10 s, 63°C for 20 s, 68°C for 1 min, 25 cycles; 68°C for 7 min, 1 cycle. Amplified PCR products were gel-purified and subcloned into pGEM-T using an Easy Vector System (Promega, Soton, UK) and into pCR.2.1 using TOPA-TA cloning kits (Invitrogen, Carlsbad, CA), according to the manufacturer's protocols. Bidirectional sequencing of all plasmids and PCR clones with vector- or gene-specific primers was performed by fluorescence tagged chain termination (Big Dye Terminator, Perkin–Elmer–ABI) followed by fractionation on an Applied Biosystems 310 automated DNA sequencer (PE Biosystems).

Expression analysis. A mouse multiple tissue Northern (MTN) blot (Clontech) was hybridized with a mouse LEU2 cDNA probe (GenBank Accesion No. AI465804), and a mouse PCR product was amplified with the primers MLEU4F, 5′-GACTCAGGCTTCCGGAGACCC-3′, and MLEU4R, 5′-GATTACATGCAGAATAGAAATGAGTCG-3′, corresponding to the genomic area covered by human LEU4 ESTs (Accession No. AA281817).

Fluorescence in situ hybridization (FISH). Fibroblasts from a normal C57 black mouse were cultured for 6 days. Colchicine (0.005 μg/ml) was added in passage 2, and after 2 h the cells were trypsinized for 1 min, treated with 0.075 M KCl for 10 min, and fixed in methanol:acetate acid, 3:1. Chromosome preparation, labeling of the probe PAC 602-N13, and hybridization were performed as previously described (Wu *et al.*, 1999).

Sequence analysis and computer programs. Sequence compilation was performed using the Sequencher DNA analysis package (Genecodes Corp., Ann Arbor, MI). Sequences were aligned with the Bestfit program (Wisconsin Package Version 10.0, Genetics Computer Group (GCG), Madison, WI), using default parameters. The degree of similarity for consecutive 50-bp sections of the resulting alignment was plotted against the position in the alignment. The Alfresco program (Jareborg and Durbin, 2000) was used to obtain a graphical representation of the human–mouse comparison and to identify positions of conserved blocks, repetitive elements, and predicted CpG islands. Alfresco uses BLASTN (Altschul *et al.*, 1997) filtered by MSPcrunch (Sonnhammer and Durbin, 1994) and the DBA (DNA Block Aligner) program (Jareborg *et al.*, 1999) to identify conserved blocks of high similarity. RepeatMasker (A. F. A. Smith and P. Green, http://ftp.genome.washington.edu/RM/RepeatMasker.html) was used to identify repetitive elements. CpG islands were predicted using the CPG program as previously described (Jareborg *et al.*, 1999). Exon–intron boundaries, CpG islands, and putative gene elements were also analyzed using the NIX program package, UK HGMP Resource Centre (http://menu.hgmp.mrc.ac.uk/menu-bin/Nix/Nix.pl).

RESULTS

Structure, Molecular Cloning, and FISH Analysis of Mouse Genomic DNA

To obtain mouse genomic DNA homologous to the region of interest on chromosome 13q14 (Corcoran *et al.*, 1998), a mouse genomic PAC library was screened with a series of probes as described under Materials and Methods. A total of 22 PACs were found to be positive for at least one of these probes, and 3 of these PACs (454-C18, 602-N13, and 588-K23) were selected for further analysis. The results from Southern hybridizations indicate that all 3 PACs represent the wild-type DNA structure around this locus, as all informative probes identified identical restriction fragments (data not shown). Interestingly, no LEU1 fragment

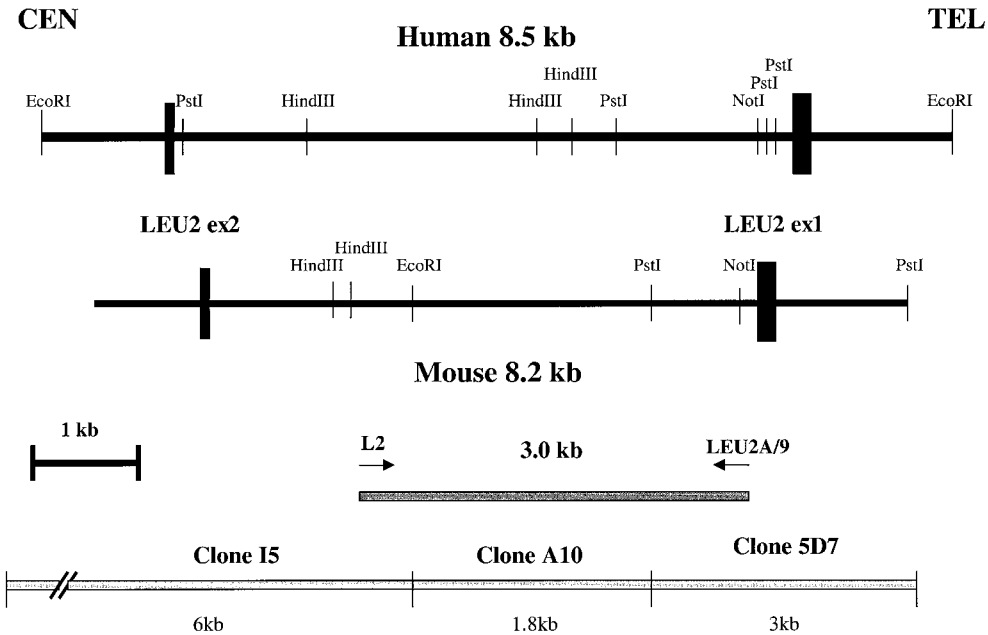

FIG. 1. Diagrammatic map of a deletion hotspot region in human and its homologous mouse region. The two filled rectangles in each map represent the location of human and mouse LEU2 exons 1 and 2. Restriction sites of *Eco*RI, *Pst*I, *Not*I, and *Hin*dIII are displayed. The isolated clones I5, A10, and 5D7 used to assemble the contiguous map of the designated mouse region are shown at the bottom.

could be detected in any of these PACs, even under less stringent hybridization conditions. Similarly, no "positive" clones were detected following hybridization of the mouse genomic PAC library with a human LEU1 cDNA probe. To determine the chromosomal location of the homologous mouse genomic region, metaphase FISH analysis was performed with the mouse PAC 602-N13 as a probe. The region was localized to mouse chromosome 14 D1-3 (data not shown).

Sequence and Computer Analyses

Sublibraries from the selected PACs were used to isolate mouse genomic DNA for sequencing of the region homologous to the human deletion hotspot. Following hybridization with a mouse LEU2 cDNA probe, two positive clones, I5 (a 6-kb *Pst*I/*Eco*RI clone) and 5D7 (a 3.0-kb *Pst*I clone) were isolated and used for sequencing (Fig. 1). A third intervening fragment, A10 (a 1.8-kb *Pst*I clone), was subsequently identified through hybridization with a 3.0-kb PCR product overlapping all three clones, I5, 5D7, and A10 (Fig. 1). Sequence analysis of the I5 clone identified sequences homologous to the entire exon 2 and a portion of introns 1 and 2 of the human LEU2 gene. Sequencing of the A10 clone and the 5D7 clone revealed the remaining part of LEU2 intron 1 as well as the entire exon 1 and its upstream 5′ region (Fig. 1). Complete bidirectional sequencing of multiple independent PCRs performed with primers localized in different clones, including the 3.0-kb PCR product, compiled a 8170-bp mouse genomic consensus sequence. A restriction map of the previously published human sequence (Liu *et al.*, 1997) and its homologous mouse consensus sequence is shown in Fig. 1. Mouse genomic clones corresponding to parts of LEU5/RFP2 and the CpG island 9E4.3, flanking the deletion hotspot, were isolated and subjected to sequencing as described above. Based on the hybridization results from the screening of the mouse PAC library with these flanking probes, we estimate the distance (Kapanadze *et al.*, 1996) encompassed by the homologous LEU5/RFP2-LEU2 and LEU2-9E4.3 regions in the mouse to 40 and 35 kb, respectively, which are slightly shorter than the predicted human genomic distance (Corcoran *et al.*, 1998; and data not shown).

Bestfit alignments of these sequences revealed extensive similarities between the human sequence and its homologous mouse sequence. The highest degree of sequence similarity (>95% identity) was found in the telomeric 9E4.3 region with an almost perfect sequence conservation of the predicted CpG island in that region (GC content in mouse, 69%; GC content in human, 72%). Alignment of the centromeric region encoding parts of the human LEU5/RFP2 gene was also found to be highly conserved (89% identity) (Fig. 2; and data not shown). The mouse sequence obtained from this region also exhibited >98% sequence identity with the mouse ESTs corresponding to human LEU5/RFP2, present in the public EST databases (GenBank Accession Nos. AI596473 and AA499774). The complete determination of a mouse consensus sequence homologous to the human deletion hotspot region enabled us to make detailed comparisons of this highly EST-rich locus, using the Alfresco program. A map of the result from pairwise comparisons between the consensus sequences is shown in Fig. 2. The degree of conservation of various parts between human and mouse is displayed both as BLAST alignments and as DBA alignments, which

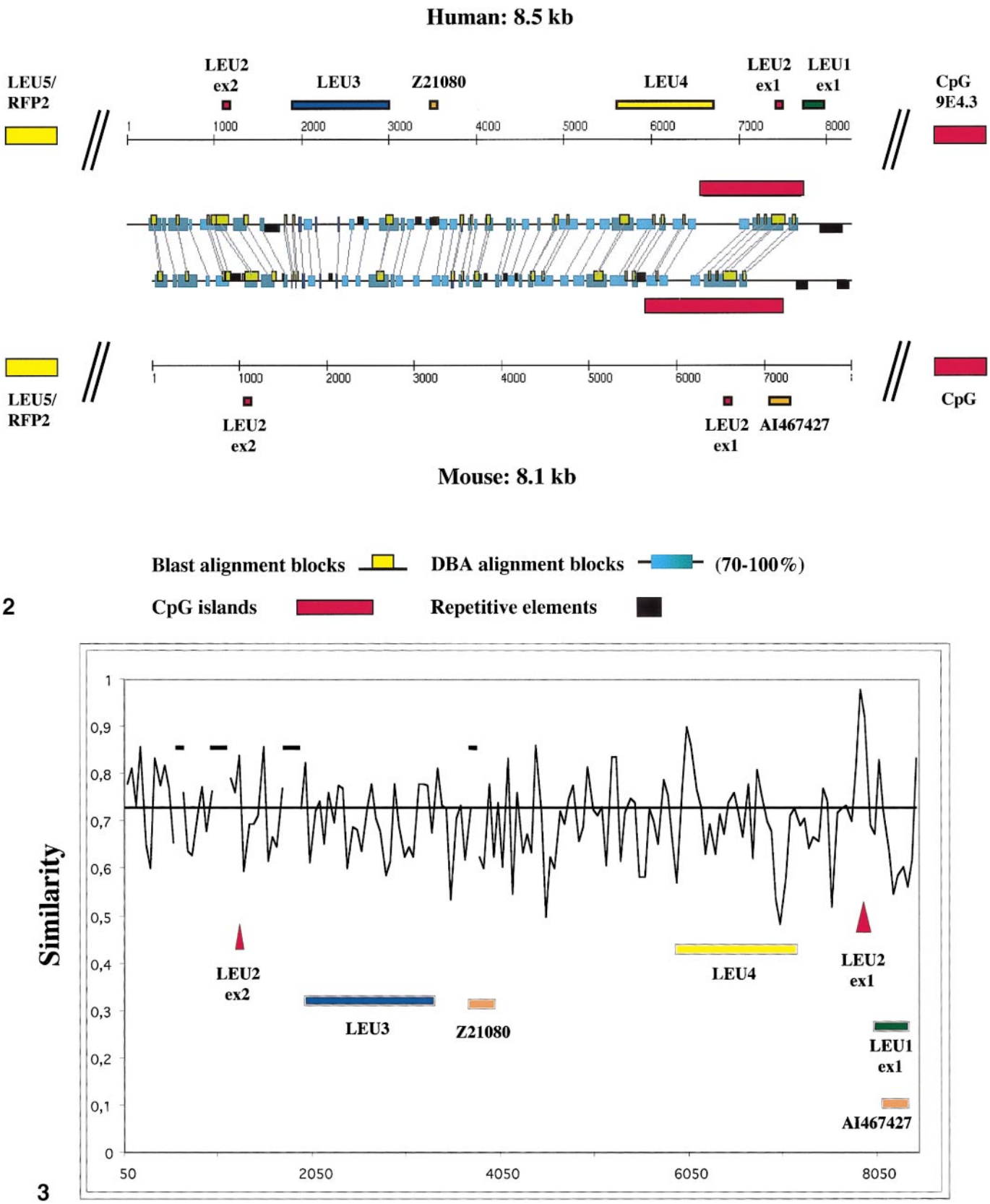

FIG. 2. The Alfresco program (Jareborg and Durbin, 2000) was used to obtain a graphical representation of the human and mouse sequence comparison. Gray lines connect BLAST and DBA alignments between human and mouse sequences. Repetitive elements are indicated by smaller boxes in black. Predicted CpG islands are shown as red boxes one level away from the line representing the sequence. The locations of various exons and ESTs in the human and mouse sequences are indicated by colored rectangles at the top and bottom of the figure, respectively. The homologous flanking regions, LEU5/RFP2 and 9E4.3, are also indicated.

identifies colinear blocks of high similarity in two DNA sequences. As can be seen in Fig. 2, there is an even distribution of conserved parts throughout this locus, with distinct BLAST hits at the position for LEU2 exon 1 and exon 2. In addition, there are a number of BLAST hits of various lengths in the first intron and in the beginning of intron 2 of LEU2. DBA identified conserved blocks that cover large parts of the region, all of which showed similarities above 70%. Some of these blocks coincide partly with the sequence from the two human ESTs, previously termed LEU3 and LEU4 (Liu *et al.*, 1997). Interestingly, the first exon of the previously described human LEU1 gene is clearly not conserved in the mouse (Fig. 2). We also analyzed the presence of the predicted human CpG island in the mouse consensus sequence. As can be seen in Fig. 2, as in human, a CpG island was predicted also in mouse, with a size of 1505 bp and a GC content of 64.1%. The human CpG island is slightly shorter (1116 bp, 62.1% GC content), and the actual nucleotide sequence is not conserved in parts that extend into the first intron of LEU2. However, in both species, the CpG islands overlap the highly conserved first exon 1 of the LEU2 gene (Fig. 2).

Repetitive elements in both the mouse and the human sequences were localized and identified using RepeatMasker (Table 1 and Fig. 2). Only a small percentage of the sequence in both species was identified as repetitive sequences, with a total of 564 bp (6.9%) being repetitive in mouse, compared to 716 bp (9%) in human. Except for a MIR repeat located slightly telomeric of LEU2 exon 1 in both species, there was no clear conservation in the location of any of the other repetitive elements found in the area (Fig. 2). The numbers of elements and the percentage of sequence occupied by the different classes of repetitive elements are summarized in Table 1.

Bestfit alignment over the entire length of the human deletion hotspot region with its corresponding mouse loci resulted in an overall similarity score of 74% identity, with individual homologous areas yielding similarity values ranging from 48 to 98% (Fig. 3). Peaks of high sequence identity (>80%) were found in LEU2 exon 1 (98%) and exon 2 (84%) and in sequence segments close to the beginning of these exons (Fig. 3). High sequence identity was also observed in some parts of the LEU4 segment (Fig. 3). Areas of low conservation (<50%) were also detected at various positions in the sequences, such as in parts of LEU4 and across the entire first exon of LEU1 (Fig. 3).

TABLE 1

Distribution of Repetitive Elements

	Number of elements (human)/(percentage of human sequence)	Number of elements (mouse)/(percentage of mouse sequence)
Alus	1/(2%)	—
MIRs	2/(5%)	1/(1.2%)
B1s	—	2/(2.5%)
B2–B4	—	1/(0.8%)
Low complexity	2/(1.4%)	4/(1.6%)
Simple repeat	1/60 bp	2/62 bp

Note. The distribution of human and mouse repetitive elements was determined by RepeatMasker. All low-complexity repeats in both human and mouse were AT-rich. The simple repetitive sequence in human was a $(CAAA)_n$ repeat, and those in mouse were $(GAAA)_n$ and $(CA)_n$ repeats.

Expression Analysis

Previous expression studies of the human LEU2 gene have shown that it is expressed in various human normal tissues as two major transcripts of 1.4 and 1.8 kb (Liu *et al.*, 1997). Northern analysis of a panel of mouse embryonic tissues reveals that the mouse LEU2 gene is also expressed as two major transcripts of 1.2 and 1.4 kb (Fig. 4). Expression of mouse LEU2 was also analyzed by RT-PCR, confirming its expression in a variety of mouse tissues (data not shown). Previous Northern hybridizations with a human LEU4 EST detected two weak bands corresponding to 8- and 5.5-kb transcripts in skeletal muscle and heart (Liu *et al.*, 1997). In the current study, Northern hybridization using a mouse PCR product corresponding to the genomic segment covering the human LEU4 sequence gives rise to weak bands of approximately 2 and 5 kb in various mouse tissues (data not shown).

DISCUSSION

Deletion of 13q14.3 telomeric to the retinoblastoma gene is the most frequent genetic abnormality in CLL, with heterozygous loss of this region being found in 30–40% of cases and homozygous loss in 10–20% of cases (Chapman *et al.*, 1994; Gardiner *et al.*, 1997), suggesting the presence of a tumor suppressor gene in this area. Recently it has also become apparent that loss in the same region of 13q is frequent not only in other hematological malignancies such as myeloma (Chang *et al.*, 1999), diffuse large cell lymphoma (Cuneo *et al.*, 1999; Wada *et al.*, 1999), and myelofibrosis (Tanaka *et al.*, 1999), but also in solid tumors including head and neck tumors and prostate carcinoma (Mae-

FIG. 3. Sequence similarity index based on the aligned human and mouse sequences. Bestfit alignment was used to determine the consensus sequence between the two species. Analysis of 50-bp intervals of the consensus sequence for sequence similarity was performed, and values were plotted against the *Y* axis (a similarity index of 1 corresponds to 100% sequence identity). Small horizontal bars (▬) reflects major gaps (>50 bp) in similarity between the two species. The large horizontal line shows the mean similarity (74%) over the whole region. Red triangles represent the mouse and human LEU2 exons. Colored rectangles representing other exons/ESTs are also shown.

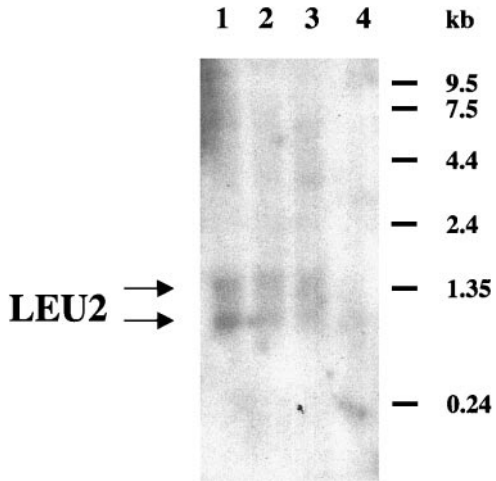

FIG. 4. Multiple tissue Northern blot of mouse LEU2. A mouse MTN blot of poly(A)$^+$ RNA derived from mouse embryo tissues was hybridized with a LEU2 probe as described under Materials and Methods. Lane **1,** 17-day embryo; lane **2,** 15-day embryo; lane **3,** 11-day embryo; and lane **4,** 7-day embryo. The position of two major transcripts of 1.4 and 1.2 kb is shown.

stro *et al.,* 1996; Gupta *et al.,* 1999; Hyytinen *et al.,* 1999).

In previous studies we defined a major deletion hotspot of less than 10 kb close to the marker D13S319, which was found to be highly EST-rich. Within this region of minimal loss, we identified exons of two separately spliced genes, termed LEU1 and LEU2 (Liu *et al.,* 1997). Additional transcripts termed LEU3 and LEU4 were localized to the first intron of LEU2, and the gene LEU5/RFP2 was found within 50 kb of the minimal loss region (Liu *et al.,* 1997; Corcoran *et al.,* 1998). Although LEU1, LEU2, and LEU5/RFP2 may be considered strong candidate genes due to the genomic location of these transcripts, neither we nor others have found mutations in the retained allele of these genes in patients with heterozygous loss in the region (Liu *et al.,* 1997). In addition to these published genes, a large series of EST transcripts has been mapped within and close to the identified critical region by us and other groups. The identification of important genes from these candidate transcripts has to date been hampered by their lack of sequence identity or homology to previously described genes. To overcome this problem, we have used a strategy of direct comparison of the human genomic sequence with that of the mouse to identify the location of potentially important sequences. DNA sequence homology between important mouse and human genes can be high, coding sequences of the p53 gene, for instance, share 83% homology at the DNA level, while other important genes such as RB1 (86%), p16 (82%), and ATM (86%) indicate that the identification of such regions of high sequence homology can provide a way of identifying important genes. To date, some comparative studies of homologous human and rodent genes have been performed on gene loci, gene clusters, and chromosomal regions (Ansari-Lari *et al.,* 1998; Ling *et al.,* 1999; Lund *et al.,* 2000). The results from these studies demonstrate that the number, order, and orientation of most genes are conserved between the two species. Importantly, it has recently been demonstrated that comparative sequence analysis between different species can be even more effective than the tested computer algorithms at identifying genes from genomic DNA (Lund *et al.,* 2000).

Thus, to obtain a better understanding of the complex region on human 13q14, and to make an assessment of the conservation of the human deletion hotspot with its syntenic region in mouse, we cloned, mapped, and sequenced the homologous mouse region. A number of results clearly demonstrate that the human 13q14 loci and the syntenic mouse region are highly conserved. First, interphase FISH analysis of the location of this region in the mouse indicates that the region shares a similar cytogenetic location on mouse chromosome 14 with the RB1 and Esterase-D genes, which are likewise located in the human 13q14 region. Furthermore, the overall structure of the region with regard to the location and distance between RFP2, LEU2, and 9E4.3 is similar in both species.

The results from direct sequence comparisons clearly demonstrate that these areas are also highly conserved at the nucleotide level and have preserved their CpG islands (Fig. 2). As already mentioned, the human 13q14 hotspot region is highly EST-rich with several ESTs localized in the first intron of LEU2 (Figs. 2 and 3). An interesting finding is the complete lack of mouse ESTs corresponding to the huge number of human LEU1 ESTs and the weak sequence conservation observed in this segment of the region (Figs. 2 and 3). There is still a possibility that the LEU1 gene has a different location in the mouse. However, hybridizations of the mouse genomic PAC library and mouse genomic DNA, with a human LEU1 cDNA probe, failed to detect any positive clones, which is in sharp contrast to the other probes from the region described above. Moreover, we also sequenced another 2-kb mouse genomic segment telomeric to the expected location of exon 1 of LEU1, without finding any sequence corresponding to LEU1 sequences. Instead, we found one unrelated mouse EST that maps around the expected location of LEU1 exon 1 (Figs. 2 and 3).

The only ESTs conserved between human and mouse in the critical loss region belong to LEU2. Within the LEU2 gene, exons 1 and 2 readily align across the entire corresponding exon in the other species, indicative of the importance of these exons in the function of the LEU2 gene. Moreover, we have recently identified an alternative first 5′ exon of human LEU2 in the 9E4.3 region (Corcoran *et al.,* manuscript in preparation). In the mouse, several ESTs with this alternative first exon also exist (i.e., GenBank Accession No. AI047357), located in the highly homologous CpG island corresponding to the human 9E4.3 region (data not shown).

The human LEU3 segment that is localized just out-

side the deletion hotspot border (Liu *et al.*, 1997) is in no place as conserved as the LEU2 exons, but does show a few small blocks of fairly high conservation between the species (Figs. 2 and 3). However, to date there are no homologous mouse ESTs corresponding to human LEU3. Together, these findings suggest that the LEU3 segment is not a coding part of a functional gene, as coding sequences are usually conserved to a much higher degree (Ling *et al.*, 1999).

The human LEU4 segment was found to be highly conserved at the nucleotide level in the centromeric part, whereas the telomeric part shows a very weak degree of conservation (Fig. 3). Although there are no corresponding mouse LEU4 ESTs described, the fact that LEU4 is partially conserved, and possibly expressed in both species (data not shown), might indicate that LEU4 is a true gene and not simply derived from intronic sequences. However, another possibility could be that this genomic area has some conserved regulatory role. In line with this hypothesis, we found that several transcription factor-binding sites for factors involved in lymphocytic development (i.e., Ikaros 2 and NFAT) were present in the highly homologous part of LEU4 in both species (data not shown).

In light of these comparative analyses, taking into account the low degree of conservation between several human ESTs and the corresponding mouse genomic sequence and the lack of corresponding mouse ESTs (Fig. 3), at least some of the human ESTs in the deletion hotspot area are most likely artifactual. Others could represent antisense transcripts (Nellen and Lichtenstein, 1993) or parts of unprocessed transcripts or differentially spliced exons of the genes in the area, but RT-PCR experiments do not support the latter possibility (data not shown).

In addition to sequence conservation in segments covered by ESTs, a number of BLAST and DBA block alignments were observed flanking these segments (Fig. 2). These blocks of high sequence identity may be conserved because of functional constraints and may, for example, have a role in transcriptional regulation. One possibility may be that the overall high sequence conservation (74%) of this whole area indicates that it functions as an enhancer region for an adjacent gene.

In summary, this investigation defines a syntenic region on mouse chromosome 14 to the deletion hotspot region on human chromosome 13q14, where the location of a tumor suppressor gene, of importance in several malignancies, has been hypothesized. This syntenic region was found to be highly conserved in overall structure, gene order, and content and at the nucleotide level. Several other important conclusions can be drawn from the data obtained. First, one of our previous candidate tumor suppressor genes in the region, LEU1, is not conserved in the mouse, clearly diminishing the possibility that this gene is the CLL tumor suppressor gene. Furthermore, several of the previously described human ESTs are found to lack corresponding mouse ESTs, and only small parts of them show the degree of conservation expected in conserved coding exons. These results should help clarify the identity of the important transcripts located within this region and help us discount other identified transcripts on the basis of lack of sequence conservation, indicating that they are not critical genes. The only transcribed sequences in the region that clearly seem to be conserved to a degree compatible with an important conserved function are LEU2, which resides in the deletion hotspot region, and LEU5/RFP2, which is adjacent to but clearly outside of the critical region, indicating that LEU2 remains on the basis of its genomic location a strong candidate as the critical tumor suppressor gene involved in CLL.

ACKNOWLEDGMENTS

We thank Ms. Birgitta Stellan for skillful assistance. This study was supported by grants from Kay Kendall Leukemia Research Fund UK, the King Gustav V Jubilee Fund, the Swedish Cancer Foundation, and the Cancer Society of Sweden.

REFERENCES

Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D. J. (1997). Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. *Nucleic Acids Res.* **25:** 3389–3402.

Ansari-Lari, M. A., Oeltjen, J. C., Schwartz, S., Zhang, Z., Muzny, D. M., Lu, J., Gorrell, J. H., Chinault, A. C., Belmont, J. W., Miller, W., and Gibbs, R. A. (1998). Comparative sequence analysis of a gene-rich cluster at human chromosome 12p13 and its syntenic region in mouse chromosome 6. *Genome Res.* **8:** 29–40.

Chang, H., Bouman, D., Boerkoel, C. F., Stewart, A. K., and Squire, J. A. (1999). Frequent monoallelic loss of D13S319 in multiple myeloma patients shown by interphase fluorescence in situ hybridization. *Leukemia* **13:** 105–109.

Chapman, R. M., Corcoran, M. M., Gardiner, A., Hawtorn, L. A., Cowell, J. K., and Oscier, D. G. (1994). Frequent homozygous deletions of the D13S25 locus in chromosome region 13q14 defines the location of a gene critical in leukemogenesis in chronic B-cell lymphocytic leukemia. *Oncogene* **9:** 1289–1293.

Corcoran, M. M., Rasool, O., Liu, Y., Iyengar, A., Grandér, D., Ibbotson, R. E., Merup, M., Wu, X., Brodjansky, 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. (1998). Detailed molecular delineation of 13q14.3 loss in B-cell chronic lymphocytic leukemia. *Blood* **91:** 1382–1390.

Cuneo, A., Bigoni, R., Rigolin, G. M., Roberti, M. G., Bardi, A., Campioni, D., Minotto, C., Agostini, P., Milani, R., Bullrich, F., Negrini, M., Croce, C., and Castoldi, G. (1999). 13q14 deletion in non-Hodgkin's lymphoma: Correlation with clinicopathologic features. *Haematologica* **84:** 589–593.

Gardiner, A. C., Corcoran, M. M., and Oscier, D. G. (1997). Cytogenetic, fluorescence in situ hybridisation, and clinical evaluation of translocations with concomitant deletion at 13q14 in chronic lymphocytic leukaemia. *Genes Chromosomes Cancer* **20:** 73–81.

Gupta, V. K., Schmidt, A. P., Pashia, M. E., Sunwoo, J. B., and Scholnick, S. B. (1999). Multiple regions of deletion on chromosome arm 13q in head-and-neck squamous-cell carcinoma. *Int. J. Cancer* **84:** 453–457.

Hardison, R. C., Oeltjen, J., and Miller, W. (1997). Long human–mouse sequence alignments reveal novel regulatory elements: A reason to sequence the mouse genome. *Genome Res.* **7:** 959–966.

Hood, L., Koop, B. F., Rowen, L., and Wang, K. (1993). Human and mouse T-cell-receptor loci: The importance of comparative large-scale DNA sequence analysis. *Cold Spring Harbor Symp. Quant. Biol.* **58:** 339–348.

Hyytinen, E. R., Frierson, H. F., Jr., Boyd, J. C., Chung, L. W., and Dong, J. T. (1999). Three distinct regions of allelic loss at 13q14, 13q21–22, and 13q33 in prostate cancer. *Genes Chromosomes Cancer* **25:** 108–114.

Jareborg, N., and Durbin, R. (2000). Alfresco—A workbench for comparative genomic sequence analysis. *Genome Res.* **10:** 1148–1157.

Jareborg, N., Birney, E., and Durbin, R. (1999). Comparative analysis of noncoding regions of 77 orthologous mouse and human gene pairs. *Genome Res.* **9:** 815–824.

Juliusson, G., and Gahrton, G. (1993). Cytogenetics in CLL and related disorders. *Bailliers Clin. Haematol.* **6:** 821–848.

Kapanadze, B., Brodianskii, V. M., Baranova, A. V., Sevetíanov, S., Fedorova, N. D., Kurskov, M. M., Kostina, M. A., Mironov, A. A., Sineokii, S. P., Zakharév, V. M., Grafodatskii, A. S., Modianov, N. N., and Iankovskii, N. K. (1996). *Genetika* **32:** 331–340. [In Russian]

Kapanadze, B., Kashuba, V., Baranova, A., Rasool, O., Van Everdink, W., Lie, Y., Syomov, A., Corcoran, M. 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., Grandér, D., Buys, C., Oscier, D. G., Zabarovsky, E. R., Einhorn, S., and Yankovsky, N. (1998). 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 tumor suppressor gene, Leu5. *FEBS Lett.* **426:** 266–270.

Koop, B. F. (1995). Human and rodent DNA sequence comparisons: A mosaic model of genomic evolution. *Trends Genet.* **11:** 367–371.

Koop, B. F., and Hood, L. (1994). Striking sequence similarity over almost 100 kilobase of human and mouse T-cell receptor DNA. *Nat. Genet.* **7:** 48–53.

Ling, V., Wu, P. W., Finnerty, H. F., Sharpe, A. H., Gray, G. S., and Collins, M. (1999). Complete sequence determination of the mouse and human CTLA4 gene loci: Cross-species DNA sequence similarity beyond exon borders. *Genomics* **60:** 341–355.

Liu, Y., Corcoran, M. M., Rasool, O., Ivanova, G., Ibbotson, R. E., Grandér, D., Iyengar, A., Baranova, A., Kashuba, V., Merup, M., Wu, X., Gardiner, A., Mullenbach, R., Poltaraus, A., Hultström, A. L., Juliusson, G., Chapman, R. M., Tiller, M., Cotter, F., Gahrton, G., Yankovsky, N., Zabarovsky, E., Einhorn, S., and Oscier, D. G. (1997). 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.

Lund, J., Chen, F., Hua, A., Roe, B., Budarf, M., Emanuel, B. S., and Reeves, R. H. (2000). Comparative sequence analysis of 634 kb of the mouse chromosome 16 region of conserved synteny with the human velocardiofacial syndrome region on chromosome 22q11.2. *Genomics* **63:** 374–383, doi:10.1006/geno.1999.6044.

Maestro, R., Piccinin, S., Doglioni, C., Gasparotto, D., Vukosavljevic, T., Sulfaro, S., Barzan, L., and Boiocchi, M. (1996). Chromosome 13q deletion mapping in head and neck squamous cell carcinomas: Identification of two distinct regions of preferential loss. *Cancer Res.* **56:** 1146–1150.

Nellen, W., and Lichtenstein, C. (1993). What makes an mRNA anti-sense-itive? *Trends Biochem. Sci.* **18:** 419–423.

O'Brien, S., del Giglio, A., and Keating, M. (1995). Advances in the biology and treatment of B-cell chronic lymphocytic leukemia. *Blood* **85:** 307–318.

Oeltjen, J. C., Malley, T. M., Muzny, D. M., Miller, W., Gibbs, R. A., and Belmont, J. W. (1997). Large scale comparative sequence analysis of the human and murine Bruton's tyrosine kinase loci reveals conserved regulatory domains. *Genome Res.* **7:** 315–329.

Oscier, D. G. (1994). Cytogenetic and molecular abnormalities in chronic lymphocytic leukaemia. *Blood* **8:** 88–97.

Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989). "Molecular Cloning: A Laboratory Manual," Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Sonnhammer, E. L., and Durbin, R. (1994). A workbench for large scale sequence homology analysis. *Comput. Appl. Biosci.* **10:** 301–307.

Sonnhammer, E. L., and Durbin, R. (1995). A dot matrix program with dynamic threshold control suited for genomic DNA and protein sequence analysis. *Gene* **167:** 1–10.

Tanaka, K., Arif, M., Eguchi, M., Guo, S. X., Hayashi, Y., Asaoku, H., Kyo, T., Dohy, H., and Kamada, N. (1999). Frequent allelic loss of the RB, D13S319 and D13S25 locus in myeloid malignancies with deletion/translocation at 13q14 of chromosome 13, but not in lymphoid malignancies. *Leukemia* **13:** 1367–1373.

Wada, M., Okamura, T., Okada, M., Teramura, M., Masuda, M., Motoji, T., and Mizoguchi, H. (1999). Frequent chromosome arm 13q deletion in aggressive non-Hodgkin's lymphoma. *Leukemia* **13:** 792–798.

Wu, X., Ivanova, G., Merup, M., Jansson, M., Stellan, B., Grandér, D., Zabarovsky, E., Gahrton, G., and Einhorn, S. (1999). Molecular analysis of the human chromosome 5q13.3 region in patients with hairy cell leukemia and identification of tumor suppressor gene candidates. *Genomics* **60:** 161–171.