From bloodjournal.hematologylibrary.org by guest on May 30, 2013. For personal use only.

blood

2001 97: 2098-2104
doi:10.1182/blood.V97.7.2098

Nucleotide sequence, transcription map, and mutation analysis of the 13q14 chromosomal region deleted in B-cell chronic lymphocytic leukemia

Anna Migliazza, Francesc Bosch, Hirokazu Komatsu, Eftihia Cayanis, Stefano Martinotti, Elena Toniato, Ernesto Guccione, Xiaoyan Qu, Minchen Chien, V. V. V. Murty, Gianluca Gaidano, Giorgio Inghirami, Peisen Zhang, Stuart Fischer, Sergey M. Kalachikov, James Russo, Isidore Edelman, Argiris Efstratiadis and Riccardo Dalla-Favera

Updated information and services can be found at:
http://bloodjournal.hematologylibrary.org/content/97/7/2098.full.html

Articles on similar topics can be found in the following Blood collections
Neoplasia (4217 articles)

Information about reproducing this article in parts or in its entirety may be found online at:
http://bloodjournal.hematologylibrary.org/site/misc/rights.xhtml#repub_requests

Information about ordering reprints may be found online at:
http://bloodjournal.hematologylibrary.org/site/misc/rights.xhtml#reprints

Information about subscriptions and ASH membership may be found online at:
http://bloodjournal.hematologylibrary.org/site/subscriptions/index.xhtml

Blood (print ISSN 0006-4971, online ISSN 1528-0020), is published weekly by the American Society of Hematology, 2021 L St, NW, Suite 900, Washington DC 20036.
Copyright 2011 by The American Society of Hematology; all rights reserved.

From bloodjournal.hematologylibrary.org by guest on May 30, 2013. For personal use only.

Nucleotide sequence, transcription map, and mutation analysis of the 13q14 chromosomal region deleted in B-cell chronic lymphocytic leukemia

Anna Migliazza, Francesc Bosch, Hirokazu Komatsu, Eftihia Cayanis, Stefano Martinotti, Elena Toniato, Ernesto Guccione, Xiaoyan Qu, Minchen Chien, V. V. V. Murty, Gianluca Gaidano, Giorgio Inghirami, Peisen Zhang, Stuart Fischer, Sergey M. Kalachikov, James Russo, Isidore Edelman, Argiris Efstratiadis, and Riccardo Dalla-Favera

Deletions of the 13q14 chromosome region are associated with B-cell chronic lymphocytic leukemia (B-CLL) and several other types of cancer, suggesting the presence of a tumor suppressor gene. In previous studies the minimal region of deletion (MDR) was mapped to a less than 300-kilobase (kb) interval bordered by the markers 173a12-82 and 138G4/1.3R. For the identification of the putative tumor suppressor gene, the entire MDR (approximately 347 kb) has been sequenced, and transcribed regions have been identified by exon trapping, EST-based full-length complementary DNA cloning, database homology searches, and computer-assisted gene prediction analyses. The MDR contains 2 pseudogenes and 3 transcribed genes: *CAR*, encoding a putative RING-finger containing protein; *1B4/Leu2*, generating noncoding transcripts; and *EST70/Leu1*, probably representing another noncoding gene (longest open reading frame of 78 codons). These genes have been sequenced in 20 B-CLL cases with 13q14 hemizygous deletion, and no mutations were found. Moreover, no somatic variants were found in the entire MDR analyzed for nucleotide substitutions by a combination of direct sequencing and fluorescence-assisted mismatch analysis in 5 B-CLL cases displaying 13q14-monoallelic deletion. The nondeleted allele of the *CAR* and *EST70/Leu1* genes was expressed in B-CLL specimens, including those with monoallelic loss, whereas no expression of *1B4/Leu2* was detectable in B-CLL, regardless of the 13q14 status. These results indicate that allelic loss and mutation of a gene within the MDR is an unlikely pathogenetic mechanism for B-CLL. However, haplo-insufficiency of one of the identified genes may contribute to tumorigenesis. (Blood. 2001; 97:2098-2104)

© 2001 by The American Society of Hematology

Introduction

B-cell chronic lymphocytic leukemia (B-CLL) represents the most common leukemia in the Western countries with an estimated incidence of 1 per 100 000 per year. The disease is characterized by the monoclonal expansion of B lymphocytes expressing the CD5 marker and exhibiting a long life span, possibly because of a perturbed apoptotic program.[1]

Current knowledge of the molecular pathogenesis of B-CLL is limited because no specific genetic alteration has yet been associated with this disease. In particular, B-CLL is not associated with reciprocal balanced chromosomal translocations.[2] Accordingly, none of the proto-oncogenes involved in chromosomal translocations in other mature B cell malignancies, including *cyclin D1*, *BCL-2*, *BCL-6*, *PAX-5*, and *c-MYC*, are primarily altered in B-CLL.[3,4] Although in a fraction of B-CLL cases, inactivation of the tumor suppressor gene *p53* (on chromosome 17p13) and deletions or mutations of the *ATM* gene (on chromosome 11q22-23) have been reported,[5-9] such lesions were observed in late stages of the disease and may not represent primary tumorigenic events.

Chromosome 13q14 deletions (approximately 50% of the patients)[2,10,11] are the commonest chromosomal alterations associated with B-CLL, followed by structural aberrations of chromosome 11q (19% of the cases) and trisomy of chromosome 12 (15%).[2] Moreover, chromosome 13 deletions can be the only cytogenetically detectable abnormality, suggesting an early role in B-CLL pathogenesis. Together, these findings point to the presence in this region of a tumor suppressor gene whose loss or inactivation may be crucial for the leukemogenesis.

Deletions of chromosome 13q14 have also been reported in a variety of human tumors, including various other types of lymphoid tumors[12-17] and myeloid leukemias,[18] as well as prostate,[19,20] head and neck,[21] and non-small-cell lung cancers.[22] Although the deleted region of chromosome 13q14 has not yet been defined precisely in most of these neoplasms, these observations suggest that a common tumor suppressor gene may reside on this chromosomal segment.

To identify the B-CLL-associated tumor suppressor gene, we previously constructed a high-density contiguous cosmid-based physical map encompassing the deleted interval.[10] This map allowed us to demonstrate that somatic loss at 13q14 occurs in 54% of the patients[10] (and unpublished results) and that the loss is monoallelic in 81% of such cases and biallelic in the remaining 19%. A minimal deleted region (MDR) in B-CLL, spanning less than 300 kilobase (kb) and representing the site of the B-CLL-associated putative tumor suppressor gene, was assigned on our

From the Institute of Cancer Genetics, the Departments of Pathology and Genetics & Development, and the Genome Center, Columbia University, New York; the Department of Medical Sciences, A. Avogadro University of Eastern Piedmont, Novara, Italy; and the Department of Pathology, New York University Medical Center, New York.

Submitted June 29, 2000; accepted November 21, 2000.

Supported by grant 985111 from Fondo de Investigationes Sanitarias (FIS) (F.B.).

Reprints: Riccardo Dalla-Favera, Institute of Cancer Genetics, Columbia University, 1150 St Nicholas Ave, New York, NY 10032; e-mail: rd10@columbia.edu.

The publication costs of this article were defrayed in part by page charge payment. Therefore, and solely to indicate this fact, this article is hereby marked "advertisement" in accordance with 18 U.S.C. section 1734.

© 2001 by The American Society of Hematology

From bloodjournal.hematologylibrary.org by guest on May 30, 2013. For personal use only.

physical map in the interval between markers 173a12-82 and 138G4/1.3R.[10] This interval has also been confirmed by other studies.[13] In fact, the MDR exhibits a partial overlap with the regions of deletion as assigned in additional reports,[23-26] and it includes a locus of 10 kb that was previously defined as a minimal consensus for 13q14 deletion in B-CLL, based on the pattern of loss in 2 B-CLL cases and one cell line.[27]

Here, we present the complete annotated sequence of approximately 347 kb encompassing the MDR and describe a complete characterization of the genes contained in this interval, an analysis of their expression status in B-CLL samples, and the screening of the entire MDR for mutations.

Materials and methods

DNA sequencing

Sixteen overlapping cosmid clones covering the MDR were described previously.[10] The nucleotide sequence of these cosmids was obtained by the shotgun DNA sequencing method.[28] Briefly, cosmid DNA, purified as described,[10] was sonicated or nebulized, and fragments between 1 and 4 kb were selected following agarose gel electrophoresis. The isolated DNA fragments were ligated into the dephosphorylated *Hinc*II or *Sma*I sites of either M13mp18 or pUC19 and electroporated into DH5α (Life Technologies, Rockville, MD) or XL1Blue (Stratagene, La Jolla, CA) *Escherichia coli* cells. Plasmid DNA was obtained by using Qiagen Turbo kits on the BioRobot 9600 (Qiagen, Valencia, CA), and single-strand M13 DNA was isolated as described.[29] Unidirectional sequencing with an M13 forward primer or bidirectional sequencing with both forward and reverse M13 primers was performed with either FS dye terminator or big dye terminator chemistry (PE Applied Biosystems, Foster City, CA). The products were analyzed, using ABI 377 sequencers. DNA sequences were filtered for vector and *E coli* sequences, assembled, and edited with the Phred, Cross_Match, Phrap, and Consed software package[30-32] (available at http://bozeman.mbt.washington.edu) on a Sun SPARC20 workstation. Sequence contigs were bridged by "walking," using custom-generated primers (Life Technologies). Overall, 5.3-fold sequence coverage was achieved. A unique contig of 347 503 base pair (bp) encompassing the region of deletion in B-CLL was then compiled and used as the reference sequence. Homology searches on public databases were performed with the use of BLASTN and TBLASTX,[33] after filtering out repeated and vector sequences with RepeatMasker, part of the Phred package.

Computational sequence annotation

The entire consensus sequence of 347 503 bp was annotated with homologies to sequences in the public databases, using BLASTN and TBLASTX. In addition, computer-predicted exons were also identified, using the GENSCAN program[34] (see the GENSCAN web server at Massachusetts Institute of Technology at http://ccr-081.mit.edu/GENSCAN.html). A graphical output of these annotations is available at our B-CLL homepage (http//:genome 2.cpmc.columbia.edu/~bcll/).

Isolation and sequencing of complementary DNA clones

Standard hybridization screenings using gene-specific probes were performed on complementary (cDNA) libraries generated from various RNAs derived from spleen (Clontech, Palo Alto, CA), placenta (ATCC, Rockville, MD), tonsil (ATCC), the B-cell lymphoma cell line BJAB (in the λZipLox vector [Life Technologies], using the SuperscriptII lambda system for cDNA synthesis and cloning [Life Technologies]), and the B-CLL cell line MO1043[35] (in the λgt11 system [Promega, Madison, WI]). The cDNA clone inserts were amplified by polymerase chain reaction (PCR), using vector-derived primers, gel purified, and extracted with GFX columns (Amersham Pharmacia Biotech, Piscataway, NJ) and then fully sequenced with vector-derived or custom-generated oligonucleotides (Life Technologies).

Rapid amplification of cDNA ends

Full-length cloning of the identified genes was completed, using the Rapid amplification of cDNA ends (RACE) Marathon-Ready cDNA kit (Clontech, Palo Alto, CA). Gene-specific primer pairs for both 3′ and 5′ extension starting from the most 3′ and 5′ cloned sequences were used. The cDNAs used were derived from spleen, lymph node, and testis.

Tumor samples and cell lines

The B-CLL specimens used in this study were diagnosed according to the criteria of the Fifth International Workshop on CLL and the National Cancer Institute-sponsored Working Group[36,37] and were characterized by fractionation of peripheral blood from B-CLL patients on a Ficoll-Hypaque gradient (Amersham Pharmacia Biotech) and flow cytometric analysis. The majority of the cells (> 64%; average, 88%) expressed the CD19 and CD20 markers. These cells also expressed the CD5 and CD23 markers as well as low-density monoclonal surface immunoglobulin of the immunoglobulin M (IgM) or IgD (κ or λ) isotypes. The fraction of CD5-positive B cells was between 64% and 96%, with a median of 84%. Normal granulocytes in the postgradient pellet from the same patients were also collected. Various control B-cell lines were used in this study: the B-cell lymphoma-derived BJAB, the cord blood-derived CB33, the lymphoblastoid cell line EREB-E2,[38] the B-CLL-derived cell lines MEC1 and MEC2,[39] MO1043,[35] Handman (G. Inghirami, unpublished), CLL183E95, and WaCaCD5$^+$.[40] Cellular DNA from both tumor samples and cell lines was extracted by cell lysis, proteinase K digestion, "salting out" extraction, and ethanol precipitation, as described.[41]

Mutational analysis of the B-CLL candidate genes

Somatic mutations were assessed on 20 B-CLL patient-derived DNAs displaying hemizygous deletion at 13q14,[10] as follows: Exons and flanking intron sequences were amplified by standard PCR reactions in a final volume of 25 μL starting from 100 ng of tumor genomic DNA. An aliquot of the PCR product (2 μL) was examined on an agarose gel and re-amplified, if the total product yield was less than 20 ng. The PCR products were then subjected to direct sequencing (with the same primers used for the amplification reaction, except where indicated), after enzymatic degradation of primers and dNTPs with *E coli* exonuclease I and shrimp alkaline phosphatase (United States Biochemical, Cleveland, OH) as described.[42] The sequences of the primers used for the mutational analysis are available at the B-CLL homepage. Each sequence variant detected in a tumor sample was ruled out as a polymorphism by sequencing the same segment on normal case-matched DNA.

Mutational analysis of the B-CLL MDR

A genomic interval of 317 kb encompassing the B-CLL MDR was analyzed by a combination of fluorescence-assisted mismatch analysis (FAMA),[43,44] and direct sequencing in 5 B-CLL patients (B-CLL Nos. 19, 22, 31, 33, and 85) carrying a hemizygous deletion at 13q14. The deletion status of B-CLL Nos. 19, 31, and 33 was described previously in detail[10]; B-CLL Nos. 22 and 85 showed a deletion throughout the 13q14 locus (data not shown). Because these B-CLL samples exhibited, on average, 23% contamination with normal lymphocytes (range, 25% to 19%), the allelic frequency (percentage of tumor-derived allele in the sample) was approximately 62%. The sensitivity of the FAMA method was tested with known mutations and detected a base change even when the frequency of the mutated allele was 15% (data not shown). FAMA was performed as described.[43] Briefly, overlapping PCR amplicons of 900 to 1300 bp were generated with chromosome 13-specific primers containing a "universal" tail (see below), starting from 100 ng of tumor DNA under standard PCR conditions (final volume, 12-50 μL). These products were then labeled by PCR (final volume, 50 μL), using universal 5′-fluorescent primers (6-FAM-5′HUP: GGACCGTTAGTAGTCGAC and HEX-3′HUP: GGTCGGATAGCTAGTCGT) that anneal to the chromosome 13-specific primer "tail." After chemical modification and cleavage, the products were electrophoresed on a 4% polyacrylamide denaturing gel, using an ABI377 sequencer together with standard size markers (TAMRA2500). The results were

From bloodjournal.hematologylibrary.org by guest on May 30, 2013. For personal use only.

analyzed with GeneScan software (PE Applied Biosystems). For direct sequencing, the products were amplified under standard PCR conditions in a final volume of 50 μL and were purified by enzymatic degradation of primers and dNTPs as described above. All sequence variants detected by FAMA or by direct sequencing were confirmed on normal DNA from the same patient displaying such nucleotide substitutions. The 13q14 region analyzed for mutations in B-CLL patients was 317 kb (between position 1-316811 of the MDR sequence). The sequences of the primers used for FAMA and direct sequencing are available at the B-CLL homepage.

Expression analysis of the B-CLL candidate genes

Total RNA was isolated from tumor samples and cell lines by guanidine isothiocyanate extraction.[45] Poly(A)$^+$ isolation was then accomplished with the PolyATtract system (Promega). Northern analysis was performed by using 25 μg total RNA or 5 μg poly(A)$^+$ RNA. Northern blot filters (including the "Immune System" and "Multiple Tissue") obtained from Clontech (Palo Alto, CA) were also used for analysis. In addition, expression profiles were determined by reverse transcriptase (RT)-PCR assays, using gene- or exon-specific primers. Briefly, cDNAs were generated in a final volume of 50 μL, starting from 2 μg of total RNA that had been treated with RQ1 DNase (Promega, Madison, WI) using SuperscriptII RNaseH$^-$ RT (Life Technologies). A control reaction lacking the RT enzyme was always performed in parallel to exclude the presence of contaminant DNA. For amplification, 2 μL of the cDNA were used per PCR reaction with gene-specific primers in a final volume of 20 μL (denaturation at 94°C for 10 seconds; annealing at 56°C to 60°C for 30 seconds; and extension at 72°C for 30 seconds for 30 cycles).

Results

Complete sequence analysis of the MDR

A total of 16 overlapping cosmids,[10] spanning the interval bordered by markers WI-10171 and GCT16C05 covering the B-CLL MDR were sequenced to completion. A continuous sequence of 347 503 bp (available at the B-CLL homepage; GenBank accession No. AF272953) was generated. This was achieved by first generating 6 sequence contigs assembled by Phred/Phrap from cosmid shotgun sequences (ranging in size from 10 047-120 816 bp), which were then bridged with sequences derived from PCR products generated on genomic DNA (ranging in size from 150-800 bp). Overall, the calculated sequence coverage was 5.3-fold redundant. In addition, approximately 250 kb of sequence (available on request) was also generated, which covers the region between the MDR telomeric boundary and the marker D13S25 (Figure 1).

The average GC content of the MDR region is 39.3%. Two CpG islands (at position 84091-85217 and 126795-129244) were identified, at least one of which is associated with the *EST70/Leu1* and *1B4/Leu2* genes. Repetitive elements comprised 40.2% of the complete sequence, mainly Alu and LINE1 repeats (12.6% and 14.3%, respectively; see the B-CLL homepage for a summary table and full description of repetitive elements).

A total of 200 nucleotide polymorphisms (SNPs) (average SNP frequency, 1/1400 bp) were also identified in the MDR region by comparing the sequences among 5 B-CLL patients. All of the identified SNPs are reported at the B-CLL homepage, and they have also been deposited in the GenBank SNP database at http://www.ncbi.nlm.nih.gov/SNP/.

Identification of transcribed sequences

Genes located within the B-CLL MDR were identified by homology searches of public databases, analysis of computer-predicted exons, and exon trapping reported previously.[10]

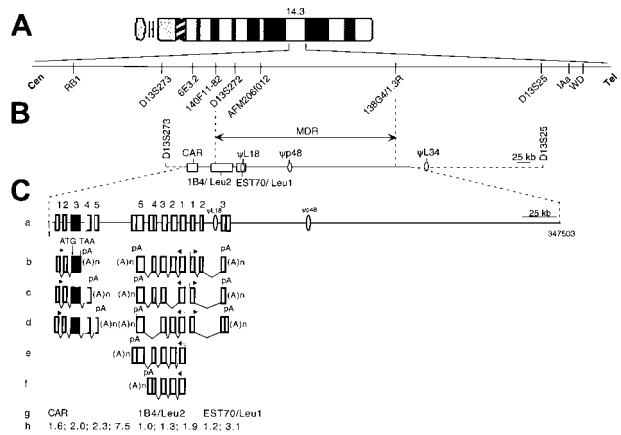

Figure 1. Schematic representation of the 13q14 chromosomal region deleted in B-CLL. (A) The chromosome 13 ideogram is shown with approximate location of reference markers used in previous studies for deletion analysis in B-CLL samples. Markers 140F11-82 and 138G4/1.3R border the MDR region of approximately 300 kb as defined previously,[10] located between the retinoblastoma 1 gene locus (RB1) and the Wilson disease (WD) locus. (B) Distribution of genes in an approximate 600-kb region (between markers D13S273 and D13S25), spanning the MDR (indicated by an arrow): 3 transcribed sequences (*CAR*, *1B4/Leu2*, and *EST70/Leu1*) and 3 pseudogenes (ΨL18, Ψp48, and ΨL34) were identified. The continuous line corresponds to the fully sequenced region of 347 503 bp; the dashed line indicates a region partially sequenced. (C) Row a. The region of approximate 347 kb completely sequenced and including the MDR is shown with various gene annotations: the gene exons (rectangles: empty, noncoding; filled, coding) and their relative positions along the MDR are indicated. The 2 pseudogenes located within the MDR are also shown. Rows b-c. Diagram of various mRNA splicing variants identified. The transcriptional start site as well as the transcriptional orientation for each candidate gene is indicated by an arrow. pA indicates the presence of a typical polyA [(A)n] addiction signal within the gene sequences. Row g. Gene names. Row h. Transcripts sizes of the detected mRNA for each gene.

Database screening. The entire MDR consensus sequence of 347 503 bp and the sequence of 250 kb covering the interval between the MDR and the D13S25 marker were used as queries against the GenBank Expressed Sequence Tag database (dbEST). A total of 244 hits were detected, representing 74 unique cDNA clusters (28 within the MDR). These dbEST hits are summarized in a table at the B-CLL homepage and are also shown as part of a graphical map of the region. Of these clusters, only 7 yielded a positive signal in Northern analysis, whereas the remaining EST clusters gave positive signals only when analyzed by RT-PCR in a variety of B-CLL and lymphoid-cell line-derived cDNAs (data not shown), suggesting low levels of expression in the examined tissues. Our attempts to extend these cDNAs by library screening and RACE experiments were negative. Accordingly, these ESTs were considered as products of basal transcription and were not pursued further.

Strong homologies to 2 known genes were found in the region; the MDR segment from 102 255 to 102 851 exhibited 83% homology to the human ribosomal protein *L18* gene (GenBank accession No. L11566), whereas the homology between the MDR segment between 174 971 and 176 557 and the progesterone-receptor associated protein p48/HSP70-interacting protein, also named putative tumor suppressor gene *ST13* (U28 918 and U17 714), was 96% (Figure 1). In addition, a sequence with 99% homology with the human ribosomal protein *L34* gene (L38 941) was observed in the region telomeric to the MDR. Several observations suggest that these sequences represent processed pseudogenes on chromosome 13 because they lack introns, harbor numerous stop codons interrupting the open reading frame of the homologous genes, and lack detectable expression (at least in B

From bloodjournal.hematologylibrary.org by guest on May 30, 2013. For personal use only.

cells) when analyzed by RT-PCR with the use of primers specific for these loci.

Computer-predicted genes. GENSCAN[34] predicts the presence of probable genes and/or exons, based on genomic sequence composition and gene structure features. A total of 22 exons predicted by GENSCAN with medium to very high probability ($P > .5$) was analyzed for expression. The results of this analysis, including exon location within the framework sequence, are summarized at the B-CLL homepage. Weak RT-PCR signals were detected for 12 exons, but in no case could these exons be extended further by cDNA library screening or RACE experiments. Only one predicted exon, overlapping with known EST sequences, was shown to recognize several transcripts by Northern analysis (see below).

Characterization of B-CLL candidate genes

A total of 9 transcripts, corresponding to the 3 genes described below, was identified within the approximate 347-kb interval (Figure 1B and Figure 2) and were fully cloned and characterized.

CAR. This gene (GenBank accession No. AF279659), named for "CLL-Associated RING finger" (and previously named *Leu5*[46]), was identified both by homology searches as well as by GENSCAN prediction, and it includes the WI-10171 EST cluster that we described previously.[10] The *CAR* gene includes at least 5 exons distributed over 21 kb of genomic DNA (bp 207-21543 of the approximate 347-kb MDR consensus sequence, Figure 1B). Different polyadenylation (see the EST hits summary table at the B-CLL homepage), and 3′ exon usage generate 4 transcripts of 1.6, 2.0, 2.3, and 7.5 kb sharing the same coding potential (Figure 1C).

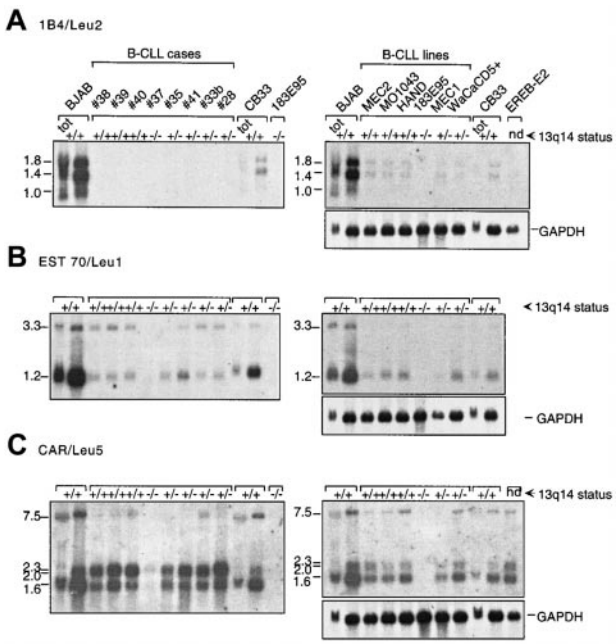

Figure 2. Expression analysis of the candidate genes in B-CLL. The 13q14 status of each sample is indicated: $^{+/+}$, no loss; $^{+/-}$, monoallelic loss; $^{-/-}$, biallelic loss; nd, not determined. PolyA$^+$ (or total, tot) RNA derived from B-CLL samples (indicated by the corresponding number) or cell lines was sequentially hybridized with A, *1B4/Leu2*-; B, *EST70/Leu1*-; and C, *CAR*-specific probes (in this order). The cell lines used include 6 B-CLL-derived cell lines,[35,39,40] one B-cell lymphoma-derived line (BJAB), one cord blood-derived line (CB33), and one lymphoblastoid cell line (EREB-E2[38]). The transcript sizes (kb) are indicated. Exposure times are 6 days for the *1B4/Leu2* gene and 2 days for the *EST70/Leu1* and *CAR* genes. The normal retained allele of genes *EST70/Leu1* and *CAR* is expressed in the deleted samples. On the contrary, no expression of the *1B4/Leu2* gene is detected in B-CLL tumors, whereas traces of expression were detected in B-CLL-derived cell lines.

Expression of these messenger RNAs (mRNAs) was detected in all tissues tested. Exons 1 to 3, which recognize all 4 messages and correspond to the 1.6-kb messages, and possibly 2.0- and 2.3-kb messages, were found to be linked in dbEST- or library-derived cDNA clones. Exons 4 and 5 detected the 7.5-kb message only.

The predicted protein product of the *CAR* gene is encoded entirely by exon 3 and consists of 407 amino acids (aa). It belongs to the RING-finger (aa 10-57) B-box (aa 89-131) coiled-coil domain-containing subfamily of RING-finger proteins.[47] It is notable that a region included in the 3′ untranslated sequence potentially encodes a 213-aa-long protein containing a BTB/POZ domain (aa 5-106) found in a variety of potassium channel proteins.

1B4/Leu2. The full-length cDNA cloning of this gene (GenBank accession No. AF279658) revealed that it is composed of at least 5 exons spanning approximately 40 kb of DNA (bp 46 433-84 921 of the approximate 347-kb MDR consensus sequence; Figure 1). Two transcripts of 1.3 and 1.9 kb (previously described partially by our group and also by others[10,27]) are generated by differential polyadenylation of exon 5. Expression was detected in all of the lymphoid tissues examined as well as in fetal liver, bone marrow, and thymus (Figure 2 and data not shown). Various alternatively spliced RNA species were also identified in spleen or lymph node (Figure 1C). A minor transcript of 1.0 kb was also detected occasionally, deriving from polyadenylation after the exon 4 sequences (Figure 1C and Figure 2). We note that several additional exons downstream of exon 5 (ie, extending the gene size) were found by 3′ RACE experiments. However, such sequences failed to generate a signal in Northern blot experiments. Therefore, they may belong to very rare transcripts (ie, detectable only by PCR) with questionable physiologic relevance, at least for B cells). Nevertheless, such exons were also included in the mutational analysis of the *1B4/Leu2* gene.

No homology between the *1B4/Leu2* and other genes was detected. This gene may not encode a protein product because only a short (55 aa) open reading frame is detectable that is not conserved in the mouse homolog, despite a high degree of conservation at the nucleotide sequence level (85% identity through exons 1-4; data not shown). This suggests evolutionary conservation of the primary nucleotide sequence independent of coding potential.

EST70/Leu1. The third gene (GenBank accession No. AF279660) located within the B-CLL MDR includes at least 3 exons, spanning approximately 25 kb of genomic DNA (bp 85 201-110 160 of the approximate 347-kb MDR consensus sequence, Figure 1). A transcript of approximately 1.2 kb (described previously[10,27]) and a less abundant 3.1-kb transcript were detected in all the lymphoid tissues tested (Figure 2 and data not shown). In some of the transcripts, exon 2 was excluded by alternative splicing (Figure 1C). Additional exons following exon 3 and further extending the gene were found by 3′ RACE experiments (see "Materials and methods" section; data not shown) but did not detect any transcript when tested by Northern analysis (nonetheless, they were included in the gene mutation analysis).

No homologies with previously identified genes or proteins were identified by BLAST searches with *EST70/Leu1* sequences. The longest open reading frame accounts for a putative product of 78 aa (or 68 aa, if exon 2 is included). Southern analysis of mouse genomic DNA with human-specific probes under low stringency conditions (allowing approximately 30% mismatches), as well as

From bloodjournal.hematologylibrary.org by guest on May 30, 2013. For personal use only.

mouse database searches using the *EST70/Leu1* sequences, suggested that this gene may not be conserved in the mouse genome (data not shown). This gene may also represent a noncoding RNA.

Candidate genes are not mutated in B-CLL

To investigate whether any of the 3 candidate genes was affected by mutations in B-CLL specimens in which one allele was retained, exon and exon/intron junction sequences were analyzed in 20 cases (the complete description of the sequences included in the mutation analysis is available at the B-CLL homepage). No sequence variant was detected, except for SNPs that were also found in normal samples (see SNPs summary at the B-CLL homepage). Absence of mutations affecting some of the *EST70/Leu1* and *1B4/Leu2* transcripts in hemizygously deleted patients had also been reported previously.[27]

MDR region is not mutated in B-CLL

To exclude the possibility that inactivating mutations may target unidentified gene regulatory elements, we searched for somatic mutations in the entire MDR area in B-CLL samples. If additional genes were present within the MDR and had escaped our attention, a complete mutation analysis would permit their recognition. Somatic sequence alterations were sought by FAMA and direct sequencing in 317 kb of the approximate 347-kb region in 5 hemizygously deleted B-CLL patients that displayed unmutated gene sequences (see above). A total of 221 kb were analyzed by FAMA, and 60 kb were analyzed by direct sequencing. A total of 36 kb consisting of low complexity repeated sequences were excluded from the analysis. Several sequence variants (200 in total) distributed along the entire sequence were detected. These variants were classified as SNPs, because they were also found in matching normal control DNA derived from the same patients (see SNPs summary and Bentley[48]).

Transcription analysis of the candidate genes in B-CLL

The absence of somatic mutations suggests that inactivation of these genes may occur through epigenetic mechanisms, such as DNA methylation, leading to loss of their expression in a tumor. To investigate this possibility, expression of the 3 B-CLL candidate gene transcripts was tested in a panel of normal tissues, tumor cell lines, and B-CLL-derived poly(A)$^+$ RNAs (Figure 2). Tumor cases, as well as cell lines representative of the germline or with deleted 13q14 configurations, were included in the analysis. When the levels of transcription of the *CAR* and *EST70/Leu1* genes were examined in nondeleted versus deleted patients, they were comparable to those observed in the positive normal tissues (data not shown and Figure 2), indicating that the retained allele of these genes is apparently normally expressed. Our analysis, however, would not detect 2-fold differences resulting from loss of expression of one allele, because of contaminating normal cells in the samples.

The *1B4/Leu2* gene showed no detectable expression in B-CLL samples, regardless of their 13q14 status, whereas expression was detected in other normal (CB33) or tumor-derived B cells (Figure 2). Low levels of *1B4/Leu2* transcription, comparable to the level detected in CB33 cells, were detected in Epstein-Barr virus-immortalized B-CLL cell lines, including one hemizygously deleted sample (MEC1; Figure 2). This finding is in contrast to what was observed in the primary tumor samples and may reflect peculiar features of the exceedingly small fraction of CLL cases that can grow in vitro, or gene expression changes associated with the immortalization process.

Discussion

In an effort to identify the B-CLL-associated tumor suppressor gene located on chromosome 13q14, we analyzed approximately 347 kb of completely annotated sequence encompassing the B-CLL MDR and characterized the genes located in this region. We have also examined their mutational and expression status in B-CLL. Moreover, we investigated the mutational status of the entire MDR interval in B-CLL patients.

The 3 genes described here may represent all the active transcriptional units present in the MDR in B cells. Although the transcriptional characterization of the MDR through a combination of database homology searches, exon trapping, and computer predictions revealed a total of 74 unique EST clusters, the majority (67) of these sequences were not bona fide transcripts either because of lack of expression (at least in B cells) or because they represented "noise" dbEST hits (ie, clones derived from unprocessed nuclear RNA fragments or from contaminating genomic DNA primed at poly(A) stretches within low complexity regions, such as Alu elements, etc). The remaining 7 EST clusters recognized RNA species on Northern blots. These corresponded to a RING finger-encoding gene (*CAR*) and 2 apparently noncoding genes (*1B4/Leu2* and *EST70/Leu1*). Characterization of these genes revealed several novel transcripts described here for the first time (detailed in Figure 1C, rows c-f).

On the basis of its structural features and the well-proven link between RING-finger containing proteins and tumorigenesis,[47] the *CAR* gene was a credible candidate for the B-CLL-associated tumor suppressor. However, no alterations were discovered in the gene exon or exon-intron junction sequences in 20 tumor cases, indicating that this gene is not affected by a typical mechanism involving loss of one allele and mutation of the retained allele. In addition, when transcription of *CAR* was tested in B-CLL samples with loss of 13q14, we found that the wild-type nondeleted allele was expressed at levels comparable to normal cell line controls, although a 50% reduction would not be detectable by our analysis. Finally, we observed that some B-CLL patients with 13q14 deletion retained the *CAR* gene (Figure 1B). On the basis of these results, the *CAR* gene is unlikely to represent the B-CLL-associated tumor suppressor gene.

The *EST70/Leu1* locus probably is a new member of a growing family of noncoding RNA genes. Such genes exert their function through their RNA product and play important roles in various cellular processes, including transcriptional (and posttranscriptional) regulation or regulation of enzymatic activities.[49-51] A link between sterile RNA transcripts and tumorigenesis has also been proposed in the case of the *H19* and tropomyosin genes.[52,53] Despite the fact that B-CLL samples with monoallelic deletion at 13q14 invariably lose one allele of the *EST70/Leu1* gene, the retained allele lacks somatic mutations, indicating that the *EST70/Leu1* gene cannot be implicated in B-CLL pathogenesis through a classic 2-hit inactivation mechanism. Because transcription of the retained unmutated alleles is maintained in the tumors, it remains possible that haplo-insufficiency may contribute to B-CLL development, in a way analogous to other tumor suppressor genes in the mouse.[54,55]

It is likely that the *1B4/Leu2* gene also produces a sterile transcript. *1B4/Leu2* is highly conserved in the mouse genome,

From bloodjournal.hematologylibrary.org by guest on May 30, 2013. For personal use only.

strongly supporting its functional significance. *1B4/Leu2* sequences are not mutated in B-CLL, but, in contrast to *CAR* and *EST70/Leu1*, no expression was detected in the B-CLL cases tested, regardless of the presence or absence of chromosome 13q14 deletion. It is possible that the lack of expression of *1B4/Leu2* in B-CLL may reflect the lack of transcription of this gene in the normal cellular counterpart of B-CLL. This possibility is difficult to address because various human lymphocyte subpopulations are heterogeneous for *1B4/Leu2* expression and the normal counterpart of B-CLL is currently unknown.[1] However, epigenetic silencing may be tumor associated and, in alternative to or together with gene loss, contribute to the lack of *1B4/Leu2* expression in all B-CLL cases. Thus, this gene remains a candidate for the B-CLL associated tumor suppressor.

Two groups[23,26] reported exclusion of the *EST70/Leu1* and *1B4/Leu2* genes from the deleted interval in 3 B-CLL cases when analyzing a total of 39 tumor samples with deletion at 13q14. However, this result was never observed by us and by 2 other groups[10,13,25] when mapping the deleted interval in more than 300 B-CLL samples with 13q14 deletion. This discrepancy could either reflect different technical approaches used in mapping the deleted area (PCR[23,26] versus Southern blotting hybridization and/or fluorescence in situ hybridization[10,13,25]) or point to the involvement of a second tumor suppressor locus in a small minority of B-CLL cases with deletion at 13q14.

In conclusion, the results presented here leave open several possibilities on the underlying pathogenetic mechanisms in B-CLL-associated deletions. First, it remains formally possible that 13q14 deletions are not of pathogenetic significance in CLL, but rather they represent epiphenomena in the disease. However, the presence of these lesions as sole cytogenetic abnormality, their high frequency, and their clonal nature strongly argue in favor of their selection during tumorigenesis. Second, the B-CLL-associated gene is within the MDR but has escaped detection because of its atypical features, including unusually large introns potentially spanning the entire MDR, lack of homologous sequences in the public domain, or lack of polyadenylation, as in the case of small nucleolar RNA genes (not represented in the dbEST collection of cDNAs). Alternatively, reduced dosage of the product of one or more of the candidate genes may be sufficient to contribute to B-CLL development. Reduced dosage can occur as a result of monoallelic loss in the case of *EST70/Leu1*, or as a consequence of allelic loss or transcriptional silencing in the case of *1B4/Leu2*. The complete MDR sequence presented here will be useful in the continuing search for transcriptional units within this region and for the design of gene targeting experiments that will explore the role of haplo-insufficiency in mice.

Acknowledgments

We thank Andrey Rzhetsky for discussions during various phases of the project; Jing Chen, Hye Park, Juyan Zhou, Gifty Asamani, Kendra Hill, and Vladan Miljkovic for expert technical assistance.

Sequence data from this article have been deposited with the EMBL/GenBank Data Libraries under Accession Nos. AF272953 (MDR sequence), AF279658 (*1B4/Leu2*), AF279659 (*CAR*), and AF279660 (*EST70/Leu1*). The B-CLL homepage can be found at http://genome2.cpmc.columbia.edu/~bcll.

References

1. Caligaris-Cappio F, Hamblin TJ. B-cell chronic lymphocytic leukemia: a bird of a different feather. J Clin Oncol. 1999;17:399-408.
2. Dohner H, Stilgenbauer S, Dohner K, Bentz M, Lichter P. Chromosome aberrations in B-cell chronic lymphocytic leukemia: reassessment based on molecular cytogenetic analysis. J Mol Med. 1999;77:266-281.
3. Gaidano G, Newcomb EW, Gong JZ, et al. Analysis of alterations of oncogenes and tumor suppressor genes in chronic lymphocytic leukemia. Am J Pathol. 1994;144:1312-1319.
4. Lo Coco F, Ye BH, Lista F, et al. Rearrangements of the BCL6 gene in diffuse large cell non-Hodgkin's lymphoma. Blood. 1994;83:1757-1759.
5. Gaidano G, Ballerini P, Gong JZ, et al. p53 mutations in human lymphoid malignancies: association with Burkitt lymphoma and chronic lymphocytic leukemia. Proc Natl Acad Sci U S A. 1991; 88:5413-5417.
6. Stankovic T, Weber P, Stewart G, et al. Inactivation of ataxia telangiectasia mutated gene in B-cell chronic lymphocytic leukaemia. Lancet. 1999; 353:26-29.
7. Starostik P, Manshouri T, O'Brien S, et al. Deficiency of the ATM protein expression defines an aggressive subgroup of B-cell chronic lymphocytic leukemia. Cancer Res. 1998;58:4552-4557.
8. Bullrich F, Rasio D, Kitada S, et al. ATM mutations in B-cell chronic lymphocytic leukemia. Cancer Res. 1999;59:24-27.
9. Schaffner C, Stilgenbauer S, Rappold GA, Dohner H, Lichter P. Somatic ATM mutations indicate a pathogenic role of ATM in B-cell chronic lymphocytic leukemia. Blood. 1999;94:748-753.
10. Kalachikov S, Migliazza A, Cayanis E, et al. Cloning and gene mapping of the chromosome 13q14 region deleted in chronic lymphocytic leukemia. Genomics. 1997;42:369-377.
11. Migliazza A, Cayanis E, Bosch-Albareda F, et al. Molecular pathogenesis of B-cell chronic lymphocytic leukemia: analysis of 13q14 chromosomal deletions. Curr Top Microbiol Immunol. In press.
12. Liu Y, Hermanson M, Grander D, et al. 13q deletions in lymphoid malignancies. Blood. 1995;86: 1911-1915.
13. Stilgenbauer S, Nickolenko J, Wilhelm J, et al. Expressed sequences as candidates for a novel tumor suppressor gene at band 13q14 in B-cell chronic lymphocytic leukemia and mantle cell lymphoma. Oncogene. 1998;16:1891-1897.
14. Cigudosa JC, Rao PH, Calasanz MJ, et al. Characterization of nonrandom chromosomal gains and losses in multiple myeloma by comparative genomic hybridization. Blood. 1998;91:3007-3010.
15. Avet-Loiseau H, Li JY, Morineau N, et al. Monosomy 13 is associated with the transition of monoclonal gammopathy of undetermined significance to multiple myeloma. Intergroupe Francophone du Myelome. Blood. 1999;94:2583-2589.
16. Rosenwald A, Ott G, Krumdiek AK, et al. A biological role for deletions in chromosomal band 13q14 in mantle cell and peripheral t-cell lymphomas? Genes Chromosomes Cancer. 1999;26: 210-214.
17. Siu LL, Wong KF, Chan JK, Kwong YL. Comparative genomic hybridization analysis of natural killer cell lymphoma/leukemia. Recognition of consistent patterns of genetic alterations. Am J Pathol. 1999;155:1419-1425.
18. La Starza R, Wlodarska I, Aventin A, et al. Molecular delineation of 13q deletion boundaries in 20 patients with myeloid malignancies. Blood. 1998;91:231-237.
19. Li C, Larsson C, Futreal A, et al. Identification of two distinct deleted regions on chromosome 13 in prostate cancer. Oncogene. 1998;16:481-487.
20. Yin Z, Spitz MR, Babaian RJ, Strom SS, Troncoso P, Kagan J. Limiting the location of a putative human prostate cancer tumor suppressor gene at chromosome 13q14.3. Oncogene. 1999; 18:7576-7583.
21. Gupta VK, Schmidt AP, Pashia ME, Sunwoo JB, Scholnick SB. Multiple regions of deletion on chromosome arm 13q in head-and-neck squamous-cell carcinoma. Int J Cancer. 1999;84:453-457.
22. Tamura K, Zhang X, Murakami Y, et al. Deletion of three distinct regions on chromosome 13q in human non-small-cell lung cancer. Int J Cancer. 1997;74:45-49.
23. Bullrich F, Veronese ML, Kitada S, et al. Minimal region of loss at 13q14 in B-cell chronic lymphocytic leukemia. Blood. 1996;88:3109-3115.
24. Bouyge-Moreau I, Rondeau G, Avet-Loiseau H, et al. 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. 1997;46:183-190.
25. Corcoran MM, Rasool O, Liu Y, et al. Detailed molecular delineation of 13q14.3 loss in B-cell chronic lymphocytic leukemia. Blood. 1998;91: 1382-1390.
26. Rondeau G, Moreau I, Bezieau S, Cadoret E, Moisan JP, Devilder MC. Exclusion of Leu1 and Leu2 genes as tumor suppressor genes in 13q14.3-deleted B-CLL. Leukemia. 1999;13: 1630-1632.
27. Liu Y, Corcoran M, Rasool O, et al. Cloning of two candidate tumor suppressor genes within a 10 kb region on chromosome 13q14, frequently deleted in chronic lymphocytic leukemia. Oncogene. 1997;15:2463-2473.
28. Deininger PL. Random subcloning of sonicated

From bloodjournal.hematologylibrary.org by guest on May 30, 2013. For personal use only.

DNA: application to shotgun DNA sequence analysis. Anal Biochem. 1983;129:216-223.

29. Russo JJ, Bohenzky RA, Chien MC, et al. Nucleotide sequence of the Kaposi sarcoma-associated herpesvirus (HHV8). Proc Natl Acad Sci U S A. 1996;93:14862-14867.
30. Ewing B, Green P. Base-calling of automated sequencer traces using phred. II. Error probabilities. Genome Res. 1998;8:186-194.
31. Ewing B, Hillier L, Wendl MC, Green P. Base-calling of automated sequencer traces using phred. I. Accuracy assessment. Genome Res. 1998;8: 175-185.
32. Gordon D, Abajian C, Green P. Consed: a graphical tool for sequence finishing. Genome Res. 1998;8:195-202.
33. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. J Mol Biol. 1990;215:403-410.
34. Burge C, Karlin S. Prediction of complete gene structures in human genomic DNA. J Mol Biol. 1997;268:78-94.
35. Kawata A, Han T, Dadey B, et al. Establishment and characterization of the tumors of chronic lymphocytic leukemia cell line in nude and SCID mice. Leuk Res. 1993;17:883-894.
36. Montserrat E, Binet JL, Catovsky D, et al. 5th International Workshop on Chronic Lymphocytic Leukemia. Leuk Res. 1992;16:717-720.
37. Cheson BD, Bennett JM, Grever M, et al. National Cancer Institute-sponsored Working Group guidelines for chronic lymphocytic leukemia: revised guidelines for diagnosis and treatment. Blood. 1996;87:4990-4997.
38. Kempkes B, Spitkovsky D, Jansen-Durr P, et al. B-cell proliferation and induction of early G1-regulating proteins by Epstein-Barr virus mutants conditional for EBNA2. EMBO J. 1995;14:88-96.
39. Stacchini A, Aragno M, Vallario A, et al. MEC1 and MEC2: two new cell lines derived from B-chronic lymphocytic leukaemia in prolymphocytoid transformation. Leuk Res. 1999;23:127-136.
40. Wendel-Hansen V, Sallstrom J, De Campos-Lima PO, et al. Epstein-Barr virus (EBV) can immortalize B-CLL cells activated by cytokines. Leukemia. 1994;8:476-484.
41. Miller SA, Dykes DD, Polesky HF. A simple salting out procedure for extracting DNA from human nucleated cells. Nucleic Acids Res. 1988;16:1215.
42. Hanke M, Wink M. Direct DNA sequencing of PCR-amplified vector inserts following enzymatic degradation of primer and dNTPs [published erratum appears in Biotechniques 1995;18(4):636]. Biotechniques. 1994;17:858-860.
43. Verpy E, Biasotto M, Meo T, Tosi M. Efficient detection of point mutations on color-coded strands of target DNA. Proc Natl Acad Sci U S A. 1994; 91:1873-1877.
44. Verpy E, Biasotto M, Brai M, Misiano G, Meo T, Tosi M. Exhaustive mutation scanning by fluorescence-assisted mismatch analysis discloses new genotype-phenotype correlations in angiodema. Am J Hum Genet. 1996;59:308-319.
45. Chirgwin JM, Przybyla AE, MacDonald RJ, Rutter WJ. Isolation of biologically active ribonucleic acid from sources enriched in ribonuclease. Biochemistry. 1979;18:5294-5299.
46. Kapanadze B, Kashuba V, Baranova A, et al. 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. 1998;426:266-270.
47. Saurin AJ, Borden KL, Boddy MN, Freemont PS. Does this have a familiar RING? Trends Biochem Sci. 1996;21:208-214.
48. Bentley DR. The human genome project—an overview. Med Res Rev. 2000;20:189-196.
49. Erdmann VA, Szymanski M, Hochberg A, de Groot N, Barciszewski J. Collection of mRNA-like non-coding RNAs. Nucleic Acids Res. 1999;27: 192-195.
50. Lanz RB, McKenna NJ, Onate SA, et al. A steroid receptor coactivator, SRA, functions as an RNA and is present in an SRC-1 complex. Cell. 1999; 97:17-27.
51. Eddy SR. Noncoding RNA genes. Curr Opin Genet Dev. 1999;9:695-699.
52. Rastinejad F, Conboy MJ, Rando TA, Blau HM. Tumor suppression by RNA from the 3′ untranslated region of alpha-tropomyosin. Cell. 1993;75: 1107-1117.
53. Hao Y, Crenshaw T, Moulton T, Newcomb E, Tycko B. Tumour-suppressor activity of H19 RNA. Nature. 1993;365:764-767.
54. Fero ML, Randel E, Gurley KE, Roberts JM, Kemp CJ. The murine gene p27Kip1 is haploinsufficient for tumour suppression. Nature. 1998; 396:177-180.
55. Venkatachalam S, Shi YP, Jones SN, et al. Retention of wild-type p53 in tumors from p53 heterozygous mice: reduction of p53 dosage can promote cancer formation. EMBO J. 1998;17: 4657-4667.