MOLECULAR AND CELLULAR BIOLOGY, Sept. 1990, p. 4658–4666
0270-7306/90/094658-09$02.00/0

Vol. 10, No. 9

Evi-2, a Common Integration Site Involved in Murine Myeloid Leukemogenesis

ARTHUR M. BUCHBERG,[1] HENDRICK G. BEDIGIAN,[2] NANCY A. JENKINS,[1] AND NEAL G. COPELAND[1]*

Mammalian Genetics Laboratory, NCI-Frederick Cancer Research and Development Center, ABL-Basic Research Program, Frederick, Maryland 21702,[1] and The Jackson Laboratory, Bar Harbor, Maine 04609[2]

Received 6 April 1990/Accepted 13 June 1990

BXH-2 mice have the highest incidence of spontaneous retrovirally induced myeloid leukemia of any known inbred strain and, as such, represent a valuable model system for identifying cellular proto-oncogenes involved in myeloid disease. Chronic murine leukemia viruses often induce disease by insertional activation or mutation of cellular proto-oncogenes. These loci are identified as common viral integration sites in tumor DNAs. Here we report on the characterization of a novel common viral integration site in BXH-2 myeloid leukemias, designated *Evi-2*. Within the cluster of viral integration sites that define *Evi-2*, we identified a gene that has the potential for encoding a novel protein of 223 amino acids. This putative proto-oncogene possesses all of the structural features of a transmembrane protein. Within the transmembrane domain is a "leucine zipper," suggesting that *Evi-2* is involved in either homopolymer or heteropolymer formation, which may play an important role in the normal functioning of *Evi-2*. Interestingly, the human homolog of *Evi-2* has recently been shown to be tightly linked to the von Recklinghausen neurofibromatosis locus, suggesting a role for *Evi-2* in human disease as well.

Chronic murine leukemia viruses induce disease in susceptible hosts via insertional activation or mutation of cellular proto-oncogenes (13). These proto-oncogenes can represent either known or previously unidentified cellular genes. Thus, the somatically acquired proviruses in tumors serve as useful retrotransposon tags for identifying and cloning cellular proto-oncogenes involved in neoplastic disease.

In practice, the genes activated or mutated by proviral integration in tumors are identified by cloning somatic proviral DNA-host cellular DNA junction fragments from retrovirally induced tumors. Unique cellular DNA probes flanking each provirus are then used to screen DNA from other tumors to determine whether the locus is a common viral integration site. Since retroviruses integrate into many distinct sites in the host genome, detection of a common viral integration site suggests that the locus encodes a gene whose activation or mutation by viral integration predisposes the target cell to neoplastic disease. Many of the proto-oncogenes identified as common viral integration sites in animal model systems are also likely to be involved in human disease, making this a very powerful approach for identifying novel cellular proto-oncogenes involved in human neoplasia.

Several mechanisms have been identified whereby viral integration can activate or mutate expression of cellular proto-oncogenes (41). Retroviral integration upstream of (or in 5′ noncoding sequences) and in the same transcriptional orientation as the proto-oncogene results in initiation of transcription from promoter sequences located in the 3′ viral long terminal repeat (LTR). Viral integration can also occur upstream from the proto-oncogene in the opposite transcriptional orientation or downstream from the proto-oncogene in the same transcriptional orientation. In both cases, gene expression is elevated because of the presence of enhancer sequences located in the LTR. Viruses can integrate into 3′ noncoding sequences of the gene, in which case RNA termination usually occurs in the poly(A) termination sequences present in the LTR. Although it is not clear how premature RNA termination affects proto-oncogene expression, one possibility is that it enhances the stability of the mRNA. Viruses can also integrate into an intron and result in production of a truncated gene product because of transcript initiation from the promoter sequences in the viral LTR. Recently, it has also been shown that viruses can integrate into the coding region of a proto-oncogene in such a fashion that it totally ablates gene expression (5, 23).

Experimental results from many laboratories suggest that the number of genes activated or mutated by proviral integration in murine tumors is large and that the repertoire of activated or mutated genes varies as a function of tumor cell type, host strain background, and virus isolate (13, 39; unpublished data). Furthermore, many of the cellular proto-oncogenes that are activated or mutated by viral integration in murine tumors appear to be causally associated primarily with T-cell lymphomas (13, 39), suggesting that many of the proto-oncogenes involved in virally induced B-cell or myeloid disease are novel cellular proto-oncogenes.

One inbred mouse strain that should be particularly valuable for identifying proto-oncogenes involved in myeloid disease is BXH-2. BXH-2 is a recombinant inbred strain derived from a cross between C57BL/6J and C3H/HeJ mice (4). Although the leukemia incidence in the two parental strains is low, greater than 95% of BXH-2 mice die of granulocytic leukemia by 1 year of age (3; unpublished data). The high incidence of granulocytic leukemia is causally associated with high levels of expression of an ecotropic murine leukemia virus that is horizontally transmitted in this strain (26).

Southern blot analysis of BXH-2 tumor DNAs has shown that the tumors are monoclonal and contain primarily somatic ecotropic proviruses, consistent with their postulated mode of induction by viral integration (3; unpublished data). Similar studies also suggest that the known proto-oncogenes are not the targets of viral activation or mutation in BXH-2

* Corresponding author.

tumors (unpublished data), consistent with tumor induction by activation or mutation of a novel set of proto-oncogenes.

In the studies described here, we characterized a novel common viral integration site in BXH-2 tumor DNAs that we have designated *Evi-2* (ecotropic viral integration site 2). Within this locus, we identified a gene that can encode a protein of 223 amino acids and has all of the structural motifs of a transmembrane protein. Within the proposed membrane-spanning hydrophobic domain is a "leucine zipper" (32), providing a potential surface for protein-protein contacts. The interaction of Evi-2 with itself or other membrane proteins may thus be an important requirement for the normal functioning of Evi-2. *Evi-2* has previously been localized to mouse chromosome 11 in a region that shows extensive synteny with human chromosome 17 (7, 8). Interestingly, the human homolog of *Evi-2* has recently been shown to be closely linked to the von Recklinghausen neurofibromatosis (NF1) locus (P. O'Connell, D. Viscochil, A. M. Buchberg, J. Fountain, R. M. Cawthon, M. Culver, J. Stevens, D. C. Rich, D. H. Ledbetter, M. Wallace, J. C. Carey, N. A. Jenkins, N. G. Copeland, F. S. Collins, and R. White, Genomics, in press; R. C. Cawthon, P. O'Connell, A. M. Buchberg, D. Viskochil, B. Weiss, M. Culver, J. Stevens, N. A. Jenkins, N. G. Copeland, and R. White, Genomics, in press), suggesting a role for *Evi-2* in this human disease.

MATERIALS AND METHODS

Mice. The BXH-2 recombinant inbred strain was derived and maintained by Benjamin A. Taylor (The Jackson Laboratory, Bar Harbor, Maine).

DNA isolation and Southern blots analysis. High-molecular-weight genomic DNAs were prepared from frozen mouse tissues (26). Bacteriophage and plasmid DNAs were purified by using standard procedures (36). Restriction endonuclease digestions, agarose gel electrophoresis, Southern blot transfers, hybridization, and washes were performed as previously described (7, 8, 26).

Screening of genomic and cDNA libraries. The partial genomic library from tumor 80-597 was generated by cleavage of tumor DNA with *Eco*RI and subsequent enrichment of DNA fragments of 17 to 18 kilobases (kb) by agarose gel electrophoresis. The size-selected DNA fragments were inserted into the *Eco*RI site of λEMBL3. The phage DNA was packaged by using Gigapack (Stratagene), and 4.4×10^5 recombinant phage were screened. A total genomic library was prepared from *Sau*3AI partial digestion of C57BL/6J spleen DNA and cloned into λFix (Stratagene). Approximately 2×10^6 recombinant phage were screened. The cDNA library (35), a gift from Elizabeth Lacy, was constructed from C57BL/6J brain cDNA cloned into λgt10. Approximately 2×10^6 recombinant phage were screened. The recombinant phage libraries were screened by using standard procedures (36).

Probes. The pEco probe is a 400-base-pair (bp) *Sma*I fragment derived from the *env* gene of pAKV623 (10). Several unique sequence probes from the *Evi-2* locus were used to restriction map the locus, as well as to determine the orientations of the proviruses. These probes included RI-G (probe A), a 1.1-kb *Eco*RI fragment isolated from a lambda genomic clone; p597.1 (probe B), a 500-bp *Pst*I genomic fragment; pBK3.3 (probe C), a 163-bp cDNA fragment subcloned into pBluescript which represents the unique 5′ sequence of pBK3; pXS1.9 (probe D), a 1.9-kb *Xba*I-*Sst*I genomic fragment subcloned into pBluescript; and probe E, a 400-bp *Bam*HI-*Eco*RI genomic fragment isolated from a lambda genomic clone.

RNA isolation and northern (RNA) blots. RNA was isolated from mouse tissue as previously described (48) by using the guanidinium isothiocyanate method. Ten micrograms of total RNA was size fractionated on 1.2% formaldehyde gels (33). Northern blots were hybridized and washed by the method of Church and Gilbert (11).

DNA sequencing. DNA clones were sequenced by the dideoxy-chain termination method (53) by using a Sequenase kit (United States Biochemical Corp., Cleveland, Ohio) on double-stranded plasmid DNA. Sequencing primers were either the M13 sequencing primer or synthetic oligomers derived from previously determined sequences.

Restriction mapping. Restriction mapping of the genomic clones was performed by standard techniques using single and double digestions of genomic or phage DNA as previously described (36). The order of the restriction fragments was determined by partial digestion of end-labeled clones (36).

Computer analysis. Sequence analysis was performed by using the software package of the Genetics Computer Group (16). Transmembrane domains were identified by using the program ALOM (allocation of membrane proteins), based on the algorithm derived by Klein et al. (29). ALOM is part of the IDEAS (integrated data base and extended analysis system for nucleic acids and proteins) sequence analysis package (28). The hydrophobicity plot was generated by the method described by Kyte and Doolittle (31) by using a program developed by Gary Smythers (Program Resources Inc., Frederick Cancer Research and Development Center). The predicted cleavage site of the signal peptide was determined by using the computer programs SIGSEQ1 and SIGSEQ2 (18).

Nucleotide sequence accession numbers. The nucleotide sequences of BK3 and BK4 have been submitted to GENBANK and assigned the accession numbers M34896 and M34897, respectively.

RESULTS

Identification of five potential common viral integration sites in BXH-2 myeloid tumors. As a first step in identifying novel proto-oncogenes associated with BXH-2 tumorigenesis, we screened 67 independent tumor DNAs containing detectable somatic ecotropic proviruses by Southern blot analysis to identify subclasses containing common integration sites. These different subclasses were constructed by grouping tumors that exhibited similar-size proviral DNA-cellular DNA junction fragments. We reasoned that if these tumors were induced by ecotropic proviral integration, then at least some of these somatic ecotropic proviruses should be located within or near a proto-oncogene(s) and that we should also be able to identify and estimate the number of different common viral integration sites involved in BXH-2 myeloid disease by using this approach. Each tumor DNA was cleaved with *Pvu*II, *Sac*I, *Eco*RI, or *Bcl*I and subjected to Southern blot analysis with an ecotropic virus-specific envelope probe (pEco). *Eco*RI and *Bcl*I do not cleave within the ecotropic provirus, while *Pvu*II and *Sac*I produce single 3′ junction fragments that are detectable with pEco (49; unpublished data). This analysis identified five distinct potential common viral integration sites in BXH-2 tumor DNAs (data not shown).

Cloning of a potential common viral integration site from BXH-2 myeloid tumors. A provirus representing one of the

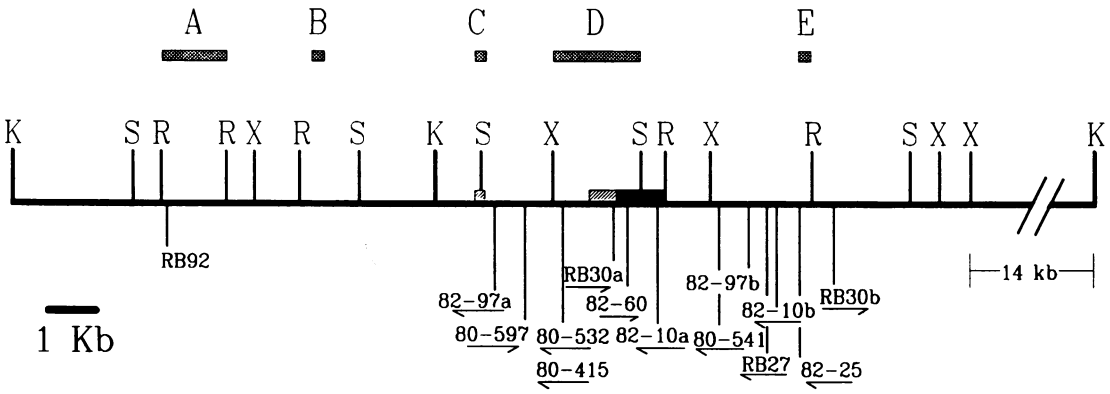

FIG. 1. Restriction map of the *Evi-2* locus. The restriction map of the *Evi-2* locus (solid line) was generated by using *Eco*RI (R), *Kpn*I (K), *Sst*I (S), and *Xba*I (X). The *Kpn*I site at the right end of the map was determined by Southern blot analysis of genomic DNA. The cross-hatched boxes (A to E) above the restriction map represent the locations of five unique sequence probes from the *Evi-2* locus (described in Materials and Methods). The solid box represents the transcribed region that is common to both *Evi-2* transcripts (see Results). The two hatched boxes represent the regions unique to each of two *Evi-2* cDNA clones (see Results). The locations and orientations of proviral integrations in each tumor are shown beneath the restriction map. The pairs of proviral integrations in tumors 82-10, 82-97, and RB30 are arbitrarily labeled a and b.

five potential common integration sites was cloned from tumor 80-597. This tumor contained a single somatic ecotropic provirus which, upon *Eco*RI digestion and hybridization with pEco, produced a 17.5-kb fragment that should contain the entire somatic ecotropic provirus, as well as 5′- and 3′-flanking cellular DNAs. Genomic DNA from tumor 80-597 was cleaved with *Eco*RI, size fractionated on a 0.8% agarose gel, and cloned into the *Eco*RI site of λEMBL3. Three independent clones were identified with pEco, and one was subsequently cloned into the *Eco*RI site of plasmid pBR325 for further study. Restriction mapping of the clone indicated that it contains a full-length ecotropic provirus, as well as 4.8 and 3.9 kb of 5′- and 3′-flanking cellular DNAs, respectively.

Two unique cellular sequences flanking the provirus were identified, a 500-bp *Pst*I fragment (probe B) located 5′ of the provirus and a 1.9-kb *Xba*I-*Sst*I fragment (probe D) located 3′ of the provirus (Fig. 1). These probes were subsequently used to screen a partial *Sau*3AI-digested C57BL/6J genomic library cloned in λFix to identify additional sequences from the region that would be useful for screening other BXH-2 tumors for virally induced rearrangements. Six clones were identified that collectively span approximately 30 kb of genomic DNA. A restriction endonuclease map of this region is shown in Fig. 1.

The cloned locus represents a common viral integration site in BXH-2 tumors. Sixty-nine BXH-2 tumor DNAs were analyzed for virally induced rearrangements in this potential common integration site by digestion with *Kpn*I, followed by Southern blot hybridization with two unique sequence probes, B and C (Fig. 1). The two probes recognize 9.0- and 26.0-kb fragments, respectively, in normal genomic DNA, allowing us to scan 35 kb of DNA for virally induced rearrangements. Among the 69 tumors analyzed, 11 (15.9%, including tumor 80-597) contained rearrangements. Ten tumors with rearrangements were identified with probe C (Fig. 2A), while one tumor with a rearrangement was identified with probe B (data not shown).These results confirmed that this locus represents a common viral integration site, which we have designated *Evi-2* (ecotropic viral integration site 2). Among the 11 tumors containing rearrangements, 14 viral integrations were detected (see below). The integrations spanned 14 kb, and most (13 of 14) occurred within an 8-kb region (Fig. 1).

Since ecotropic proviruses contain *Kpn*I sites within both LTRs, the sizes of the rearranged *Kpn*I fragments can be used to predict the locations of viral integration sites within the *Evi-2* locus (Fig. 1). Surprisingly, three tumors (82-10, 82-97, and RB30) each contained two rearranged fragments (Fig. 2A, lanes 5, 8, and 10). Further Southern blot analysis demonstrated that these tumors contained two independent integrations within the *Evi-2* locus (Fig. 1). On the basis of the results shown in Fig. 2, it was not possible to determine whether each pair of integration events occurred in two different alleles of the same tumor cell or in two different subpopulations of tumor cells. However, the results demonstrated that the integrations were not present in the same allele.

The tumors containing *Evi-2* rearrangements were also analyzed by Southern blot analysis following *Eco*RI digestion and hybridization with each of the unique sequence probes shown in Fig. 1. Since ecotropic proviruses do not contain *Eco*RI sites, the size of the altered allele can be used to predict whether the rearrangements result from the integration of full-length or truncated proviruses. Among the 13 rearrangements characterized, six (80-597, 80-532, 80-415, 82-10a, 80-541, and RB27) appeared to result from integration of full-length ecotropic proviruses (Fig. 2B, lanes 2 to 5; data not shown). Another six rearrangements (82-97a, RB30a, 82-60, 82-10b, 82-25, and RB30b) appeared to result from integration of an altered provirus. Subsequent restriction enzyme analysis indicated that the altered provirus was a result of a 3.8-kb deletion that included pEco sequences (Fig. 2B, lanes 7, 8, and 10; data not shown). This deleted provirus may be inherited horizontally in the BXH-2 strain, like the ecotropic virus (26). However, since this truncated provirus does not hybridize with pEco, it is impossible to determine whether it is inherited horizontally or carried in the germ line. One rearrangement (RB92) resulted from a 3.0-kb insertion. The nature of this insertion was not characterized.

The orientations of proviruses within the *Evi-2* locus were determined by further analysis with a restriction endonuclease, such as *Xba*I or *Hin*dIII, which cuts asymmetrically within the provirus (49). No strongly preferred provirus orientation was identified; four proviruses were in the 5′-to-3′ orientation with respect to the *Evi-2* restriction map, while

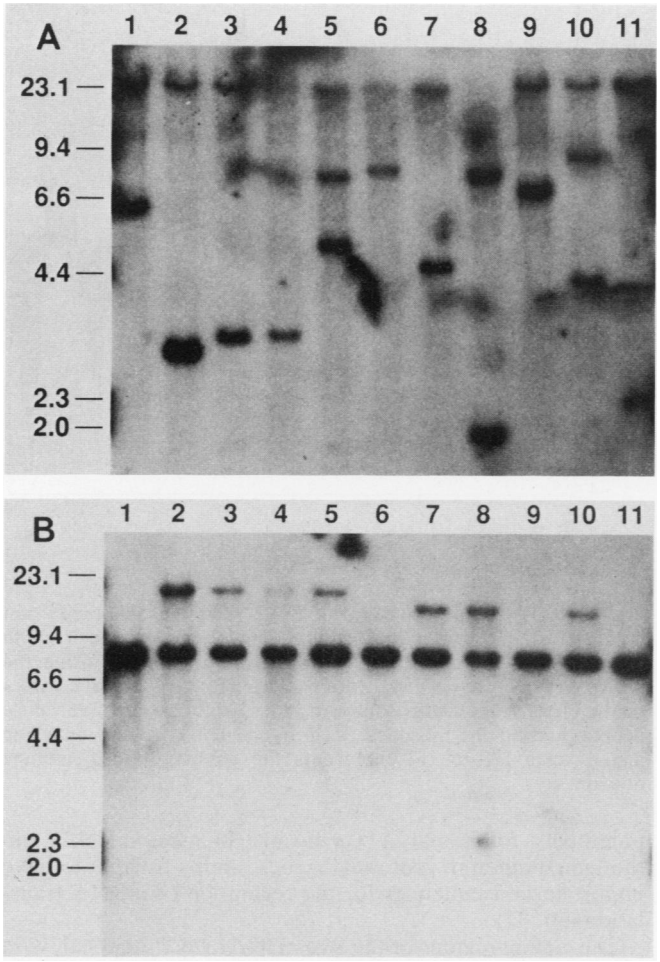

FIG. 2. *Evi-2* rearrangements in BXH-2 tumor DNAs. BXH-2 tumor DNAs were cleaved with *Kpn*I (A) or *Eco*RI (B) and analyzed for rearrangements by Southern blot analysis with probe C (Fig. 1). DNA (5 μg per lane) from tumors 80-541 (lane 1), 80-597 (lane 2), 80-415 (lane 3), 80-532 (lane 4), 82-10 (lane 5), 82-25 (lane 6), 82-60 (lane 7), 82-97 (lane 8), RB27 (lane 9), RB30 (lane 10), and RB92 (lane 11) are shown. The restriction fragment common to all lanes (>23 kb in A and 8.6 kb in B) corresponds to the unrearranged germ line *Evi-2* locus. The additional restriction fragments seen in each lane represent virally induced rearrangements of the *Evi-2* locus. The sizes (in kilobases) on the left correspond to the mobilities of *Hin*dIII-cleaved lambda DNA fragments run in parallel lanes of the same gels.

eight proviruses were oriented in the 3′-to-5′ direction (Fig. 1).

Identification of *Evi-2* transcripts in normal adult tissues. To identify transcripts from the *Evi-2* locus, we first screened each of the *Evi-2* unique-sequence probes (Fig. 1) for the presence of evolutionary well-conserved sequences that might suggest the presence of a transcribed region. A "zoo blot" containing DNAs from many different species was hybridized with each of the unique sequence probes. The results of this analysis indicated that probe D was evolutionarily well conserved; it hybridized to DNAs from chickens, mice, mink, cats, sheep, hamsters, dogs, rats, and humans (unpublished data). Probe D is located within the cluster of viral insertion sites defining *Evi-2* (Fig. 1), further suggesting that it is a likely candidate to encode *Evi-2* transcripts.

Since RNA was not available from tumors containing

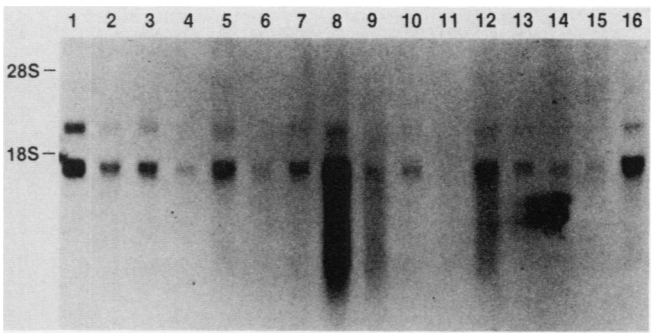

FIG. 3. *Evi-2* expression in normal mouse tissues. Total RNAs (10 μg) obtained from several mouse tissues were examined by Northern blot analysis with probe D. RNAs were isolated from adult brain (lane 1), spleen (lane 2), thymus (lane 3), heart (lane 4), lung (lane 5), small intestine (lane 6), skin (lane 7), lipopolysaccharide-activated macrophages (lane 8), bone and marrow (lane 9), muscle (lane 10), submaxillary gland (lane 11), mammary gland (lane 12), placenta (lane 13), 20-day-old embryos (lane 14), testis (lane 15), and ovary (lane 16). The RNAs in lanes 8, 9, and 12 were slightly degraded, and lanes 2 and 11 contained approximately 50% less RNA than the others as determined by hybridization of a chicken β-actin probe (data not shown). The positions of 18S and 28S RNAs are shown.

Evi-2 rearrangements, we sought to identify *Evi-2* transcripts in normal mouse tissues. Northern blots containing RNAs from several normal adult mouse tissues and 20-day-old embryos were hybridized with probe D. The Northern analysis identified two *Evi-2* transcripts of 1.8 and 2.2 kb in almost every tissue examined (Fig. 3). The relative hybridization intensities of the two transcripts differed; the larger transcript was less abundant than the smaller transcript. The highest levels of *Evi-2* expression were observed in brain tissue, macrophages, and ovaries. These results suggest that *Evi-2* encodes a gene that is normally expressed in adult mouse tissues.

Characterization of the *Evi-2* coding region. *Evi-2* cDNA clones were obtained by screening a C57BL/6J brain cDNA library with probe D. Eight clones were isolated, and two were subsequently subcloned into pBluescript for further analysis. The two subclones, pBK3 and pBK4, were sequenced (Fig. 4). BK3 was 1,217 bp long, and BK4 was 1,613 bp long. Neither clone contained a poly(A) tail nor a poly(A) addition site. Both clones ended at the same *Eco*RI site that we subsequently determined to represent a naturally occurring *Eco*RI site in the two cDNAs that was not completely methylated by *Eco*RI methylase during generation of the cDNA library. The absence of a poly(A) region or a poly(A) addition consensus sequence indicates that the two cDNA clones do not represent full-length mRNA. The 3′ 1,046-bp segments of both cDNAs were identical except for a single nucleotide substitution at position 404. However, the 5′ ends of the two cDNAs differed.

DNA sequence analysis of a genomic clone derived from the transcribed region (probe D) demonstrated that BK4 is encoded by a single exon. The unique 5′ end of BK3 is located 2.5 kb upstream of the BK4 exon (Fig. 1), indicating that BK3 is generated via splicing of an upstream exon into a splice acceptor site located at position −37 to −21 in BK4 (Fig. 4). Probes from the unique 5′ end of BK4 hybridized only to the 2.2-kb *Evi-2* transcript, providing evidence that BK4 sequences are present in mature mRNA and are not an artifact resulting from the presence of unspliced RNA (data not shown).

```
A -587 CGGACGCACTTCACATTCTCATTTGATGCACACTTGCTCTCAGGCTCCACCAACCTCCAC -528
  -527 CCTCCCTCTCCCACAAAAAAAGCCACTTGTTATATCTTTTAAAACTTGAAACGAAACTTT -468
  -467 CAAGGCCATGTTCTCCTTTCCTTCAAGATGAGGCTGTGCTTGTCAGAGGGTGGTACTTCT -408
  -407 GTGAGAAGGAGCCTTCAAAGAGGAAGCACAGCACAAAAGCAGAAGCATCGGACTCCAAGC -348
  -347 CCAGGTCCAGGGCACTCCCTTAAAAGGAGAATTTGTAGAGCTATGGTGCAAGGTTAAACA -288
  -287 TGGTCAATAGATTTGTAATGCTAACTTACAGCATTACAGAATGTACAACAAAAGTTTTCA -228
  -227 ATTTTTATGTACAATCTAAGAACAGAAATGAAAATTCAGTTAAAAAACAAAACAACAAAT -168
  -167 TCAGTTGAGAAGAATCCAATAACAAAAGGTAGGCATGGATCTCTGGCGTCCTTGTGCTAT -108
  -107 CTCAAGTGCAGCCTGTTAACAACAATTACTCCACGTGCCAGACACTGTTCTTCCACAAAC  -48
   -47 TAACAACTGATATTGCTGTCCCAACAGATAGCAGTGGACCTGTAGGTATGGAGCACAAAG   13
                                                 M  E  H  K  G

    14 GACAGTACCTGCATCTTGTTTTCCTGATGACAACAGTGTGGGCCTCGTCGTCTTCTGGAA   73
     6  Q  Y  L  H  L  V  F  L  M  T  T  V  W  A  S  S  S  S  G  T   25

    74 CAAGACCAAACTACACACACCTGTGGGCTAGCAGTGTCACTGCCTCAGGTTCCAGTAATC  133
    26  R  P  N  Y  T  H  L  W  A  S  S  V  T  A  S  G  S  S  N  Q   45

   134 AGAATGGCTCCAGCAGACATCCGAGTGACAATAACACAAACCTTGTCACTCCTGCCGTGG  193
    46  N  G  S  S  R  H  P  S  D  N  N  T  N  L  V  T  P  A  V  G   65

   194 GTCACAAGGTGAGCGCCACAGACAAGCCTGCATCATCTCCCCCTGTGCCTTTAGCTTCCA  253
    66  H  K  V  S  A  T  D  K  P  A  S  S  P  P  V  P  L  A  S  T   85

   254 CATCTACACTCAAGTCCTCCACGCCCCATGCCTTCAGAAACAGTTCTCCAACAGCAGAGA  313
    86  S  T  L  K  S  S  T  P  H  A  F  R  N  S  S  P  T  A  E  I  105

   314 TCAAAAGTCAGGGGGAAACGTTTAAAAAGGAAGTCTGTGAGGAGAACACCAGCAACACGG  373
   106  K  S  Q  G  E  T  F  K  K  E  V  C  E  E  N  T  S  N  T  A  125

   374 CCATGCTAATTTGCTTAATTGTAATTGCAGTGCTTTTCCTTATCTGTACCTTTCTATTTC  433
   126  M  L  I  C  L  I  V  I  A  V  L  F  L  I  C  T  F  L  F  L  145

   434 TATCAACTGTGGTTCTGGCAAACAAAGTCTCATCTCTCAAAAGGTCAAAGCAAGTAGGCA  493
   146  S  T  V  V  L  A  N  K  V  S  S  L  K  R  S  K  Q  V  G  K  165

   494 AGCGCCAGCCGCGGAGCAACGGTGACTTTCTGGCAAGTAGTGGGCTCTGGACTGCTGAGT  553
   166  R  Q  P  R  S  N  G  D  F  L  A  S  S  G  L  W  T  A  E  S  185

   554 CGGACACTTGGAAAAGAGCAAAGGAGCTCACAGGCTCCAACCTCCTATTGCAATCTCCTG  613
   186  D  T  W  K  R  A  K  E  L  T  G  S  N  L  L  L  Q  S  P  G  205

   614 GTGTGCTTACAGCAGCCAGGGAGAGAAAGCACGAAGAAGGAACTGAAAAACTCAACTAGG  673
   206  V  L  T  A  A  R  E  R  K  H  E  E  G  T  E  K  L  N  *     225

   674 CTGTGGGTTGGAAAGGGAAATGCAAAGCAGCAATGAGGAAGGCAGAGTTAGAAATAAAGC  733
   734 TGGCTGGTGGTATTTCCCATCCACGTTTGCTCAGATGATCCAAAGTCTGACATGCTCTTT  793
   794 GTTTAGCGTCAGACCAGAAAGGGAGATGTATGCTTGCTCAAGGTTTATATCTGTTTCGTA  853
   854 CTTCATATTTTTAAGAAAGTAAAACAACAGCAGCTAAAGAAACTAATAGAAATAACATGA  913
   914 GGTAATTAAACATTCCTGTTCAGCAGTGAATGACCTCTGTCTGAATGTACTTGAACTTGG  973
   974 AAGAGTTCTTTTAAGCCCTTAGAGCCAAGCAGATGTCAAATTCCCAATGTGTA        1026

B -191 AAAGGAAATGTGTCATCTGTGGTTTGGTTTTAAGAGTGGAAAGCTAGCTGTACATATCCT  -132
  -131 CTTTAAAAAACCCTGCAGATTTAAGTTTCACAGTCTCCAAAGTATTCTACTTAAAATAAG   -72
   -71 AAGAGCAGAACTGAGGTGGCTTACGAAGTGACGGCTGGGCGAGCTCTGAAG           -21
```

FIG. 4. *Evi-2* cDNA sequence. (A) The entire nucleotide and deduced amino acid sequences (single-letter code) of *Evi-2* cDNA BK4 are shown. Nucleotide numbering starts with the first nucleotide of the presumed initiating ATG of the longest ORF. The deduced amino acid sequence is numbered from the first amino acid of the ORF. The presumed signal sequence 1 to 19 is doubly underlined. The most likely cleavage site is after amino acid 19; however, the predicted cleavage site could also be at amino acid position 21 to 25, as predicted by SIGSEQ1 and SIGSEQ2 (18). The five potential N-linked glycosylation sites are shown as singly underlined amino acid residues. The putative transmembrane domain is shaded. The amino acid sequence potentially involved in formation of a leucine zipper is in boldface. The splice acceptor site used by BK3 is shown as the underlined nucleotides at position −37 to −21. *, Location of termination codon. (B) Nucleotide sequence of the 5′ exon of BK3. The numbering is the distance from the presumed start site of translation.

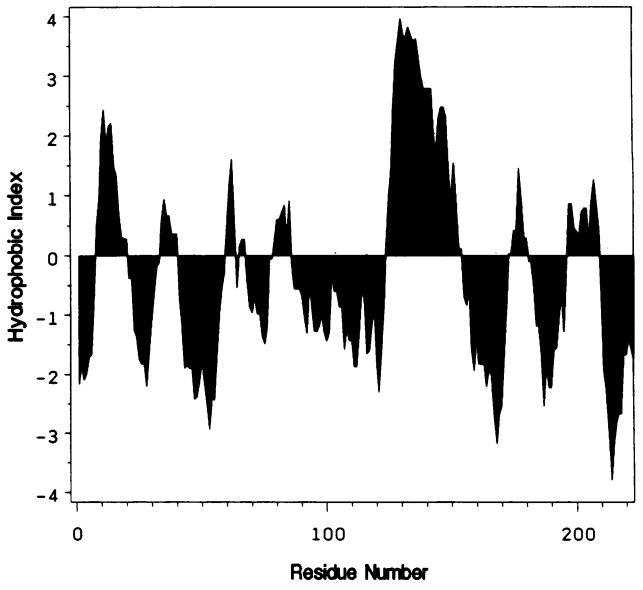

FIG. 5. Hydrophobicity plot of Evi-2. The hydrophobicity plot was calculated by using the algorithm of Kyte and Doolittle (31) with a window setting of 7. The calculation was done by using the University of Wisconsin software package (16). The positive values on the *y* axis denote hydrophobic residues, while the negative values denote hydrophilic residues. The hydrophobic region spanning amino acids 125 to 151 represents the putative transmembrane domain.

A long open reading frame (ORF) was identified in the region common to both cDNAs (Fig. 4). This ORF was entirely contained within the sequences of both *Evi-2* cDNAs. The sequence surrounding the first in-frame ATG of the ORF conforms to the consensus eucaryotic translation start site, suggesting that this ATG represents the start of translation (30). This ORF has the potential to encode a 223-amino-acid protein with a predicted molecular mass of 24 kilodaltons. The 5′ leader sequence of BK4 contains 10 ATG codons, at least 3 of which are located within good consensus start sites. However, all of these ATGs are closely followed by stop codons. The presence of multiple potentially functional ATGs is rare in most mRNAs, although common in proto-oncogenes, and is thought to have important consequences for the regulation of mRNA translation (30, 37).

The coding regions of the two cDNAs were identical, with the exception of a single valine-to-alanine substitution resulting from a T-to-C change at position 404. This conservative substitution might reflect a polymorphism that is still segregating in C57BL/6J mice or a cloning artifact. Sequencing of the corresponding genomic region (probe D) revealed that it encoded a valine at this position, indicating that if this substitution is a cloning artifact, valine is the amino acid normally encoded at this position.

***Evi-2* has all of the structural motifs of a transmembrane protein.** Neither the nucleotide sequence nor the predicted amino acid sequence of *Evi-2* shares any homology with known genes whose sequences are present in the GENBANK, NBRF, or EMBL data base, consistent with the hypothesis that *Evi-2* encodes a novel proto-oncogene. The predicted coding region of *Evi-2* has all of the structural motifs of a transmembrane protein (Fig. 4 and 5). The amino-terminal 19 amino acids conform to the consensus sequence for a signal peptide (62). The most likely cleavage site of the signal peptide is after the alanine at position 19; however, it is also possible that cleavage could occur after amino acid 21, 22, 23, 24, or 25, as predicted by von Heijne (63), who used the computer programs (SIGSEQ1 and SIGSEQ2) developed by Folz and Gordon (18). The putative 105-amino-acid extracellular domain of *Evi-2* has five potential N-linked glycosylation sites (Fig. 4). Additionally, this region is extremely rich in serine and threonine residues (23 and 14, respectively, of 105 amino acids), suggesting the presence of numerous O-linked glycosylation sites as well (Fig. 4) (27). The presumed extracellular domain ends with a region rich in hydrophobic amino acids (average hydropho-

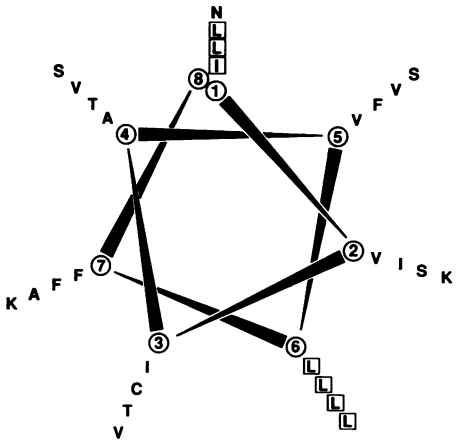

FIG. 6. Helical-wheel analysis of the hydrophobic region of Evi-2. The amino acid sequence of the hydrophobic region of Evi-2 is displayed in a helical wheel. The view is end to end along the axis of the α helix. The helix begins at amino acid 131 (isoleucine) at position 1 and continues to amino acid 158 (lysine). The amino acids forming the potential leucine zipper are boxed.

bicity, 2.7 over a 27-amino-acid region [31]), followed by several positively charged amino acids. This hydrophobic sequence is typical of transmembrane domains, and multiple basic residues mark the cytoplasmic end of the transmembrane domain (Fig. 5). The putative cytoplasmic domain of *Evi-2* does not contain any potential N-linked glycosylation sites, supporting the proposed orientation of the protein.

Interestingly, a heptad repeat of leucine residues potentially capable of forming a leucine zipper (32) is present in the region of the transmembrane domain. The hydrophobicity plot (Fig. 5) indicates that the leucine zipper is likely to cross the membrane boundary with the last turn of the helix occurring within the cytoplasm. This intracellular region might be important in initiating the formation of the helix. It has been noted recently that leucine residues are very common in proteins (6), especially in hydrophobic regions; thus, the presence of the leucine zipper in *Evi-2* should be interpreted cautiously. However, this region is 100% conserved with respect to the human homolog of the *Evi-2* gene (Cawthon et al., Genomics, in press), and furthermore, all of the leucines are clustered on two of the seven spokes of the helical wheel (Fig. 6), suggesting that this leucine zipper plays a vital role in *Evi-2* function. Several transmembrane proteins have been reported to have potential leucine zippers within or just distal to their transmembrane domains. This group includes the family of K^+ channel proteins, the glucose-transporter glycoproteins, and the F glycoproteins of paramyxoviruses (9, 38, 67). The first two groups of proteins have multiple membrane-spanning domains, while *Evi-2* and the F proteins appear to have only a single membrane-spanning region. It has been proposed that the leucine zipper is important in protein dimerization (32, 43); thus, the presence of leucine zippers in these proteins suggests that they are involved in protein-protein interactions. It has been demonstrated or hypothesized for several of these transmembrane proteins that dimers or multimers are the functional units (9, 67), further implicating these leucine zippers as protein dimerization sites.

A helical-wheel analysis (32) performed for the transmembrane region of *Evi-2* (Fig. 6) showed that two spokes of the helical wheel are composed of leucine (or isoleucine) residues. One of these spokes starts with isoleucine at amino acid 131, and the other spoke starts with the leucine at amino acid 136. The presence of two spokes capable of forming leucine zippers raises the possibility that *Evi-2* is involved in multimer formation rather than simply dimerization, since there are two nucleation sites for interactions with either itself or other proteins. It is also noteworthy that the only cysteine residues found in *Evi-2* are located proximal to or within the leucine zipper (at amino acid positions 117, 129, and 140) and are conserved in the human homolog of *Evi-2* (Cawthon et al., Genomics, in press). These cysteines might be involved in formation of disulfide bridges once contact is made via the leucine zipper and possibly represent a mechanism to regulate *Evi-2* function.

DISCUSSION

We describe here the isolation and characterization of a novel common site of viral integration, *Evi-2*, from myeloid tumors of BXH-2 mice. Within the cluster of viral integration sites that define *Evi-2*, we identified a gene that can encode a 24-kilodalton protein possessing all of the structural features of a transmembrane protein. These results suggest that expression of this gene is altered by viral integration and that this altered expression predisposes cells to myeloid disease.Additional experiments are required to prove definitively that this gene encodes a proto-oncogene involved in myeloid disease.

Northern blot analysis demonstrated that *Evi-2* is normally expressed in mice (Fig. 3). Since we had no RNAs from the original 11 tumors that had viral integrations within the *Evi-2* locus, we screened an additional 70 newly isolated BXH-2 tumors for viral integrations within the *Evi-2* locus. None of the newly isolated BXH-2 tumors that provided a source of RNA contained rearrangements in *Evi-2*. As a result, we were unable to observe the effects of viral integration on *Evi-2* expression. Nonetheless, many of the viral integrations can be predicted to have profound effects on *Evi-2* expression. For example, four viral integrations appear to have occurred in *Evi-2* intron sequences and three others appear to be located in exon sequences. Additional sequencing studies are required to determine whether the latter three viral integrations occurred in coding or in 5′- or 3′-untranslated sequences. However, polymerase chain reaction amplification of DNA from one of these tumors (82-10) suggested that the provirus is located in untranslated sequences (unpublished data). One integration was detected 5′ of the gene, while six integrations were detected 3′ of the gene. However, because the 5′ and 3′ ends of *Evi-2* have yet to be cloned, it is still possible that these viral integrations are located within transcribed sequences.

On the basis of the locations of viral integration sites with respect to the coding region of *Evi-2*, it is apparent that most, if not all, viral integrations would leave the *Evi-2* coding region intact, suggesting that viral integration alters the regulation of *Evi-2* expression, rather than producing a mutated form of the protein or totally ablating *Evi-2* expression. The data also suggest that there are multiple ways in which viral integration can affect *Evi-2* expression. For example, proviruses are located in positions where they could act (i) through an enhancer mechanism (proviruses RB92, 82-97a, 80-532, 80-415, 82-97b, 80-541, 82-10b, RB27, 82-25, and RB30b), (ii) by promoting *Evi-2* expression from viral LTRs (proviruses 80-597, RB30a, and 82-60), or (iii) by producing premature *Evi-2* RNA termination (provirus 82-10a).

Multiple mechanisms for virus-induced proto-oncogene

activation have been reported for several other cellular proto-oncogenes. For example, *Myb* expression is altered by viral integration in introns located both 5′ and 3′ of the exons homologous to v-*myb* (56, 65, 66); *Int-1*, *Int-2*, and *Hst* expression is altered by virus integration in 5′ or 3′ noncoding sequences (17, 42, 45, 46); and c-*myc* expression is altered by virus integration in 5′- or 3′-noncoding sequences, 5′- or 3′-untranslated sequences, or intron sequences (14, 24, 25, 34, 40, 50, 52, 54, 59). It is interesting that most of the viral integrations located in 3′-noncoding regions are oriented in the same transcriptional orientation as the gene being activated. However, in *Evi-2*, most proviruses located downstream of the known end of the gene are located in the opposite transcriptional orientation. The significance of this observation remains to be determined.

Surprisingly, 3 of 11 BXH-2 tumors with *Evi-2* rearrangements sustained two viral integrations in the *Evi-2* locus. Double integrations have been reported for other common integration sites as well, including p53, *Ahi-1*, *Mlvi-1*, *Mlvi-2*, and *Pvt-1* (5, 20, 23, 47, 61). In the tumor suppressor gene p53, in which viral integration apparently leads to gene inactivation, viral integration into both alleles of a tumor cell could totally ablate p53 expression, resulting in subsequent tumor formation (5, 23). However, with *Evi-2*, it is unlikely that viral integrations lead to inactivation of *Evi-2* expression, since in all cases, one of the integrations occurs 3′ of the known gene. In *Ahi-1*, *Mlvi-1*, and *Mlvi-2*, the double integrations appear to occur in different subpopulations of tumor cells. In cell line YAC-1, with two *Pvt-1* integrations, a normal unrearranged allele, in addition to two *Pvt-1* rearrangements, was observed, indicating that both integrations occurred in the same allele or that the cell line contains multiple copies of chromosome 15 that carry *Pvt-1*. For *Evi-2*, we did not determine whether the double integrations occur in the same tumor cell or in different subpopulations of tumor cells. If they occur in the same tumor cell, our Southern results indicate that the double integrations occur in two separate alleles (Fig. 2), suggesting that they provide some further selective advantage to tumor cells.

Other genes located in the vicinity of *Evi-2* may also be affected by viral integration, and their altered expression may play an important role in the induction of myeloid disease. For example, integration of mouse mammary tumor virus between two fibroblast growth factor-related proto-oncogenes, *Int-1* and *Hst*, which are located within 17 kb of one another on mouse chromosome 7, can activate expression of both genes (46). Activation of *Int-1* and *Hst* expression by mouse mammary tumor virus is thought to be important in the induction of mammary tumors by mouse mammary tumor virus. Integration of Moloney murine leukemia virus into the *Pvt-1* locus has been shown to activate the expression of c-*myc* (60), which is located 260 kb proximal of *Pvt-1* (22), leading to the development of T-cell lymphomas in rats. The *Pvt-1* locus is part of a large gene that occupies a minimum of 200 kb of DNA and begins 57 kb distal of c-*myc* (57). Viral integrations in *Pvt-1* not only activate c-*myc* expression but also affect *Pvt-1* expression (60). Although the role of the *Pvt-1* gene in tumor formation is unknown, it is clear that activation of c-*myc* is an important event in the induction of T-cell lymphomas in mice and rats (14, 50, 54, 58, 59).

Northern analysis has identified two *Evi-2* transcripts of 1.8 and 2.2 kb in virtually every tissue examined. Sequence analysis of the *Evi-2* cDNAs suggests that the two transcripts contain a common coding region but differ at their 5′ noncoding ends. We do not know whether these two transcripts are initiated from two distinct promoters or whether they are initiated from one promoter, followed by differential splicing of 5′-untranslated leader sequences.

The 5′ leader of the 2.2-kb transcript contains multiple ATG codons that are out of frame with the ORF. This motif has been observed in other proto-oncogenes (30, 37) and is thought to be involved in fine-tuning the expression of genes which have important functions in regulating cell growth and differentiation. The 1.8-kb transcript does not contain these multiple ATG codons. However, since we did not identify the 5′ end of this transcript, it is possible that these codons are present in the region not cloned. In this regard, it is important to note that the human homolog of *Evi-2* encodes a single 1.6-kb mRNA that does not contain multiple 5′ ATG codons but instead has multiple AUUUA sequences in the 3′-untranslated region (Cawthon et al., Genomics, in press). This motif is known to be involved in including rapid mRNA turnover (55). Similar AUUUA motifs are not present in the known 3′-untranslated region of murine *Evi-2*.

The presence of two *Evi-2* mRNAs suggests a mechanism whereby viral integration can affect *Evi-2* expression. Namely, viral integration could affect the relative abundance or nature of *Evi-2* transcripts. This mechanism would be analogous to results obtained with c-*myc*, in which viral integrations or chromosomal translocations located within or near c-*myc* can alter the relative abundance of c-*myc* transcripts initiated from two different promoters (50), induce expression from a cryptic promoter (25, 54), or alter the translation start site (21).

Among 14 viral integration events characterized at the *Evi-2* locus, six appeared to result from integration of a provirus with a 3.8-kb deletion. The deletion includes the ecotropic virus-specific *env* sequences (pEco), suggesting that if this is a simple deletion, the deletion includes most of the *pol* and *env* genes. Two other retrovirus-induced diseases have also been shown to be causally associated with the presence of a defective retrovirus. Feline acquired immunodeficiency syndrome is induced by a defective feline leukemia virus that contains a deletion in *pol* (44), while a similar disease in mice is induced by an ecotropic virus that is defective in *pol* and *env* (1). When the helper-free defective ecotropic virus is inoculated into young adult mice, 1,000-fold expansion of the infected cell population is observed (24). The expanded cell population is oligoclonal, suggesting that additional genetic changes are required to produce the final tumor cell population, which is monoclonal. These studies raise the possibility that the defective ecotropic provirus identified in BXH-2 myeloid leukemias has a dual role in the induction of myeloid disease. The defective virus could initially infect and greatly expand an oligoclonal population of myeloid cells that are subsequently converted into a monoclonal proliferation via retroviral insertional mutagenesis of cellular proto-oncogenes, such as *Evi-2*. Proto-oncogene activation via insertional mutagenesis could result from integration of either the wild-type ecotropic helper virus or the defective virus. Transmission of this defective virus in BXH-2 mice could explain the unique susceptibility of these mice to myeloid disease.

The *Evi-2* locus has previously been mapped to the distal half of mouse chromosome 11 in a region that is conserved with human chromosome 17 (7, 8). This synteny has allowed us to predict that the human homolog of *Evi-2* would map in the pericentric region of chromosome 17. Recent studies have confirmed this prediction, showing that the human homolog of *Evi-2* lies between two translocation breakpoints identified in patients with NF1 (O'Connell et al., unpub-

lished data; Cawthon et al., unpublished data). These breakpoints were mapped to human chromosome 17q11.2 and lie within or near the NF1 gene, identifying *Evi-2* as a candidate for the NF1 gene.

This is not the first example of implication of a common viral integration site in both murine and human diseases. The *Pvt-1* locus was originally identified as a variant (6;15) translocation in murine plasmacytomas (64) and was subsequently shown to serve as a site for chromosome translocation in human Burkitt's lymphomas (19) and as a common viral integration site in murine T-cell lymphomas (20).

NF1 is an autosomal dominant disease occurring at a frequency of 1 in 3,000 individuals (15). The disease is characterized by hyperpigmented patches of skin (café-au-lait spots) and multiple subcutaneous nodules (neurofibromas) that are thought to represent preneoplastic lesions. NF1 patients have a significantly increased risk of developing malignancies such as neurofibrosarcomas (e.g., see reference 51). They also have been reported to have a higher incidence of juvenile myeloid leukemia (2, 12). The postulated role of *Evi-2* in murine neoplastic disease, combined with its chromosomal location in humans and the increased frequency of malignancies observed in patients with NF1, makes *Evi-2* a good candidate for the NF1 gene. Whether or not this turns out to be the case, *Evi-2* has provided a molecular entry into a very interesting and important region of human chromosome 17 and has raised several possibilities for *Evi-2* involvement in human disease.

ACKNOWLEDGMENTS

We thank Elizabeth Lacy for the gift of the C57BL/6J brain cDNA library, Fritz Propst for providing the RNA samples, Steven Nothwehr for analyzing *Evi-2* with the SIGSEQ programs, and Linda Byrd for helpful discussion and sharing unpublished results. We thank Linda Siracusa, Peter Johnson, Peter O'Connell, Richard Cawthon, and Brian Stanton for many helpful and valuable discussions and Gary W. Smythers for help with the computer analysis and for writing the program that produced the hydrophobicity plot. We also thank Marge Strobel and David Kingsley for providing the C57BL/6J genomic library, Linda Cleveland and Mary Dickinson for technical assistance, and the members of the Mammalian Genetics Laboratory for critical review of the manuscript.

This work was supported in part by the National Cancer Institute under Public Health Service contract NO1-CO-74101 with ABL and by the National Institutes of Health under Public Health Service grant CA31101 to H.G.B. The NCI-Frederick Cancer Research and Development Center and The Jackson Laboratory are fully accredited by the American Association for Accreditation of Laboratory Animal Care.

ADDENDUM IN PROOF

Recent studies (D. Viskochil et al., Cell **62**:187–192, 1990; R. M. Cawthon et al., Cell **62**:193–201, 1990) indicate that the *Evi-2* gene described here is embedded in a large intron of the *NF-1* gene. Viral integrations in the *Evi-2* locus may thus affect expression of the *NF-1* gene as well.

LITERATURE CITED

1. **Aziz, D. C., Z. Hanna, and P. Jolicoeur.** 1989. Severe immunodeficiency disease induced by a defective murine leukaemia virus. Nature (London) **338**:505–508.
2. **Bader, J. L., and R. W. Miller.** 1978. Neurofibromatosis and childhood leukemia. J. Pediatr. **92**:925–929.
3. **Bedigian, H. G., D. A. Johnson, N. A. Jenkins, N. G. Copeland, and R. Evans.** 1984. Spontaneous and induced leukemias of myeloid origin in recombinant inbred BXH mice. J. Virol. **51**:586–594.
4. **Bedigian, H. G., B. A. Taylor, and H. Meier.** 1981. Expression of murine leukemia viruses in the highly lymphomatous BXH-2 recombinant inbred mouse strain. J. Virol. **39**:632–640.
5. **Ben David, Y., V. R. Prideaux, V. Chow, S. Benchimol, and A. Bernstein.** 1988. Inactivation of the p53 oncogene by internal deletion or retroviral integration in erythroleukemic cell lines induced by Friend leukemia virus. Oncogene **3**:179–185.
6. **Brendel, V., and S. Karlin.** 1989. Too many leucine zippers? Nature (London) **341**:574–575.
7. **Buchberg, A. M., H. G. Bedigian, B. A. Taylor, E. Brownell, J. N. Ihle, S. Nagata, N. A. Jenkins, and N. G. Copeland.** 1988. Localization of *Evi-2* to chromosome 11: linkage to other proto-oncogene and growth factor loci using interspecific backcross mice. Oncogene Res. **2**:149–165.
8. **Buchberg, A. M., E. Brownell, S. Nagata, N. A. Jenkins, and N. G. Copeland.** 1989. A comprehensive genetic map of murine chromosome *11* reveals extensive linkage conservation between mouse and human. Genetics **122**:153–161.
9. **Buckland, R., and F. Wild.** 1989. Leucine zipper motif extends. Nature (London) **338**:547.
10. **Chattopadhyay, S. K., M. R. Lander, E. Rands, and D. R. Lowy.** 1980. Structure of endogenous murine leukemia virus DNA in mouse genomes. Proc. Natl. Acad. Sci. USA **77**:5774–5778.
11. **Church, G. M., and W. Gilbert.** 1984. Genomic sequencing. Proc. Natl. Acad. Sci. USA **81**:1991–1995.
12. **Clark, R. D., and J. J. Hutter, Jr.** 1982. Familial neurofibromatosis and juvenile chronic myelogenous leukemia. Hum. Genet. **60**:230–232.
13. **Copeland, N. G., A. M. Buchberg, D. J. Gilbert, and N. A. Jenkins.** 1989. Recombinant inbred mouse strains: models for studying the molecular genetic basis of myeloid tumorigenesis. Curr. Top. Microbiol. Immunol. **149**:45–57.
14. **Corcoran, L. M., J. M. Adams, A. R. Dunn, and S. Cory.** 1984. Murine T lymphomas in which the cellular *myc* oncogene has been activated by retroviral insertion. Cell **37**:113–122.
15. **Crowe, F. W., W. T. Schull, and J. F. Neel.** 1956. A clinical pathological and genetic study of multiple neurofibromatosis. C. C. Thomas, Springfield, Ill.
16. **Devereux, J., P. Haeberli, and O. Smithies.** 1984. A comprehensive set of sequence analysis programs for the VAX. Nucleic Acids Res. **12**:387–395.
17. **Dickson, C., R. Smith, S. Brookes, and G. Peters.** 1984. Tumorigenesis by mouse mammary tumor virus: proviral activation of a cellular gene in the common integration region *int-2*. Cell **37**:529–536.
18. **Folz, R. J., and J. I. Gordon.** 1987. Computer-assisted predictions of signal peptidase processing sites. Biochem. Biophys. Res. Commun. **146**:870–877.
19. **Graham, M., and J. M. Adams.** 1986. Chromosome 8 breakpoint far 3′ of the c-*myc* oncogene in a Burkitt's lymphoma 2;8 variant translocation is equivalent to the murine *pvt-1* locus. EMBO J. **5**:2845–2851.
20. **Graham, M., J. M. Adams, and S. Cory.** 1985. Murine T lymphomas with retroviral inserts in the chromosomal 15 locus for plasmacytoma variant translocations. Nature (London) **314**: 740–743.
21. **Hann, S. R., M. W. King, D. L. Bentley, C. W. Anderson, and R. N. Eisenman.** 1988. A non-AUG translational initiation in c-*myc* exon 1 generates an N-terminally distinct protein whose synthesis is disrupted in Burkitt's lymphomas. Cell **52**:185–195.
22. **Henglein, B., H. Synovzik, P. Groitl, G. W. Bornkamn, P. Hartl, and M. Lipp.** 1989. Three breakpoints of variant t(2;8) translocations in Burkitt's lymphoma cells fall within a region 140 kilobases distal from c-*myc*. Mol. Cell. Biol. **9**:2105–2113.
23. **Hicks, G. G., and M. Mowat.** 1988. Integration of Friend murine leukemia virus into both alleles of the p53 oncogene in an erythroleukemic cell line. J. Virol. **62**:4752–4755.
24. **Huang, M., C. Simard, and P. Jolicoeur.** 1989. Immunodeficiency and clonal growth of target cells induced by helper-free defective retrovirus. Science **246**:1614–1617.
25. **Isfort, R., R. L. Witter, and H.-J. Kung.** 1987. C-*myc* activation in an unusual retrovirus-induced avian T-lymphoma resembling Marek's disease: proviral insertion 5′ of exon one enhances the

expression of an intron promoter. Oncogene Res. 2:81–94.
26. **Jenkins, N. A., N. G. Copeland, B. A. Taylor, H. G. Bedigian, and B. K. Lee.** 1982. Ecotropic murine leukemia virus DNA content of normal and lymphomatous tissues of BXH-2 recombinant inbred mice. J. Virol. 42:379–388.
27. **Johnson, D., A. Lanahan, C. R. Buck, A. Sehgal, C. Morgan, E. Mercer, M. Bothwell, and M. Chao.** 1986. Expression and structure of the human NGF receptor. Cell 47:545–554.
28. **Kanehisa, M., P. Klein, P. Greif, and C. DeLisi.** 1984. Computer analysis and structure prediction of nucleic acids and proteins. Nucleic Acids Res. 12:414–428.
29. **Klein, P., M. Kanehisa, and C. DeLisi.** 1985. The detection and classification of membrane-spanning proteins. Biochim. Biophys. Acta 815:468–476.
30. **Kozak, M.** 1987. An analysis of 5'-noncoding sequences from 699 vertebrate messenger RNAs. Nucleic Acids Res. 15:8125–8148.
31. **Kyte, J., and R. F. Doolittle.** 1982. A simple method for displaying the hydropathic character of a protein. J. Mol. Biol. 157:105–132.
32. **Landschulz, W. H., P. F. Johnson, and S. L. McKnight.** 1988. The leucine zipper: a hypothetical structure common to a new class of DNA binding proteins. Science 240:1759–1764.
33. **Lehrach, H., D. Diamond, J. M. Wozney, and H. Boedtker.** 1977. RNA molecular weight determinations by gel electrophoresis under denaturing conditions, a critical reexamination. Biochemistry 16:4743–4751.
34. **Li, Y., C. A. Holland, J. W. Hartley, and N. Hopkins.** 1984. Viral integration near c-*myc* in 10–20% of MCF 247-induced AKR lymphomas. Proc. Natl. Acad. Sci. USA 81:6808–6811.
35. **Lonberg, N., S. N. Gettner, E. Lacy, and D. R. Littman.** 1988. Mouse brain CD4 transcripts encode only the COOH-terminal half of the protein. Mol. Cell. Biol. 8:2224–2228.
36. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
37. **Marth, J. D., R. W. Overell, K. E. Meier, E. G. Krebs, and R. M. Perlmutter.** 1988. Translational activation of the *lck* proto-oncogene. Nature (London) 332:171–173.
38. **McCormack, K., J. T. Campanelli, M. Ramaswami, M. K. Mathew, M. A. Tanouye, L. E. Iverson, and B. Rudy.** 1989. Leucine-zipper motif update. Nature (London) 340:103.
39. **Mucenski, M. L., D. J. Gilbert, B. A. Taylor, N. A. Jenkins, and N. G. Copeland.** 1987. Common sites of viral integration in lymphomas arising in AKXD recombinant inbred mouse strains. Oncogene Res. 2:33–48.
40. **Nottenburg, C., E. Stubblefield, and H. E. Varmus.** 1987. An aberrant avian leukosis virus provirus inserted downstream from the chicken c-*myc* coding sequence in a bursal lymphoma results from intrachromosomal recombination between two proviruses and deletion of cellular DNA. J. Virol. 61:1828–1833.
41. **Nusse, R.** 1986. The activation of cellular oncogenes by retroviral insertion. Trends Genet. 2:244–247.
42. **Nusse, R., A. van Ooyen, D. Cox, Y. K. Fung, and H. Varmus.** 1984. Mode of proviral activation of a putative mammary oncogene (*int-1*) on mouse chromosome 15. Nature (London) 307:131–136.
43. **O'Shea, E. K., R. Rutkowski, and P. S. Kim.** 1989. Evidence that the leucine zipper is a coiled coil. Science 243:538–542.
44. **Overbaugh, J., P. R. Donahue, S. L. Quackenbush, E. A. Hoover, and J. I. Mullins.** 1988. Molecular cloning of a feline leukemia virus that induces fatal immunodeficiency disease in cats. Science 239:906–910.
45. **Peters, G., S. Brookes, M. Placzek, M. Schuermann, R. Michalides, and C. Dickson.** 1989. A putative *int* domain for mouse mammary tumor virus on mouse chromosome 7 is a 5' extension of *int-2*. J. Virol. 63:1448–1450.
46. **Peters, G., S. Brookes, R. Smith, M. Placzek, and C. Dickson.** 1989. The mouse homolog of the *hst/k-FGF* gene is adjacent to *int-2* and is activated by proviral insertion in some virally induced mammary tumors. Proc. Natl. Acad. Sci. USA 86:5678–5682.
47. **Poirier, Y., C. Kozak, and P. Jolicoeur.** 1988. Identification of a common helper provirus integration site in Abelson murine leukemia virus-induced lymphoma DNA. J. Virol. 62:3985–3992.
48. **Propst, F., M. P. Rosenberg, A. Iyer, K. Kaul, and G. F. Vande Woude.** 1987. c-*mos* proto-oncogene RNA transcripts in mouse tissues: structural features, developmental regulation, and localization in specific cell types. Mol. Cell. Biol. 7:1629–1637.
49. **Rands, E., D. R. Lowy, M. R. Lander, and S. K. Chattopadhyay.** 1981. Restriction endonuclease mapping of ecotropic murine leukemia viral DNAs: size and sequence heterogeneity of the long terminal repeat. Virology 108:445–452.
50. **Reicin, A., J.-Q. Yang, K. B. Marcu, E. Fleissner, C. F. Koehne, and P. V. O'Donnell.** 1986. Deregulation of the c-*myc* oncogene in virus-induced thymic lymphomas of AKR/J mice. Mol. Cell. Biol. 6:4088–4092.
51. **Riccardi, V. M., and J. E. Eicher.** 1986. Neurofibromatosis: phenotype, natural history and pathogenesis. Johns Hopkins University Press, Baltimore.
52. **Robinson, H. L., and G. C. Gagnon.** 1986. Patterns of proviral insertion and deletion in avian leukosis virus-induced lymphomas. J. Virol. 57:28–36.
53. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA 74:5463–5467.
54. **Selten, G., H. Y. Cuypers, M. Zijlstra, C. Melief, and A. Berns.** 1984. Involvement of c-*myc* in MuLV-induced T cell lymphomas in mice: frequency and mechanisms of activation. EMBO J. 3:3215–3222.
55. **Shaw, G., and R. Kamen.** 1986. A conserved AU sequence from the 3' untranslated region of GM-CSF mRNA mediates selective mRNA degradation. Cell 46:659–667.
56. **Shen-Ong, G. L. C., H. C. Morse III, M. Potter, and J. F. Mushinski.** 1986. Two modes of c-*myb* activation in virus-induced mouse myeloid tumors. Mol. Cell. Biol. 6:380–392.
57. **Shtivelman, E., B. Henglein, P. Groitl, M. Lipp, and J. M. Bishop.** 1989. Identification of a human transcription unit affected by the variant chromosomal translocations 2;8, and 8;22 of Burkitt lymphoma. Proc. Natl. Acad. Sci. USA 86:3257–3260.
58. **Steffen, D.** 1984. Proviruses are adjacent to c-*myc* in some murine leukemia virus-induced lymphomas. Proc. Natl. Acad. Sci. USA 81:2097–2101.
59. **Steffen, D. L., and E. Q. Nacar.** 1988. Nucleotide sequence of the first exon of the rat c-*myc* gene: proviral insertions in murine leukemia virus-induced lymphomas do not affect exon 1. Virology 164:53–63.
60. **Tsichlis, P. N., B. M. Shepherd, and S. E. Bear.** 1989. Activation of the *Mlvi-1/mis1/pvt-1* locus in Moloney murine leukemia virus-induced T-cell lymphomas. Proc. Natl. Acad. Sci. USA 86:5487–5491.
61. **Tsichlis, P. N., P. G. Strauss, and M. A. Lohse.** 1985. Concerted DNA rearrangements in Moloney murine leukemia virus-induced thymomas: a potential synergistic relationship in oncogenesis. J. Virol. 56:258–267.
62. **von Heijne, G.** 1985. Signal sequences: the limits of variation. J. Mol. Biol. 184:99–105.
63. **von Heijne, G.** 1986. A new method for predicting signal sequence cleavage sites. Nucleic Acids Res. 14:4683–4690.
64. **Webb, E., J. M. Adams, and S. Cory.** 1984. Variant (6;15) translocation in a murine plasmacytoma occurs near an immunoglobulin kappa gene but far from the *myc* oncogene. Nature (London) 312:777–779.
65. **Weinstein, Y., J. L. Cleveland, D. S. Askew, U. R. Rapp, and J. N. Ihle.** 1987. Insertion and truncation of c-*myb* by murine leukemia virus in a myeloid cell line derived from cultures of normal hematopoietic cells. J. Virol. 61:2339–2343.
66. **Weinstein, Y., J. N. Ihle, S. Lavu, and E. P. Reddy.** 1986. Truncation of the c-*myb* gene by a retroviral integration in an interleukin 3-dependent myeloid leukemia cell line. Proc. Natl. Acad. Sci. USA 83:5010–5014.
67. **White, M. K., and M. J. Weber.** 1989. Leucine-zipper motif update. Nature (London) 340:103–104.