GENOMICS 9, 446-460 (1991)

cDNA Sequence and Genomic Structure of *EVI2B*, a Gene Lying within an Intron of the Neurofibromatosis Type 1 Gene

RICHARD M. CAWTHON,* LONE B. ANDERSEN,† ARTHUR M. BUCHBERG,‡ GANGFENG XU,* PETER O'CONNELL,* DAVID VISKOCHIL,* ROBERT B. WEISS,* MARGARET R. WALLACE,† DOUGLAS A. MARCHUK,† MELANIE CULVER,* JEFFREY STEVENS,* NANCY A. JENKINS,‡ NEAL G. COPELAND,‡ FRANCIS S. COLLINS,† AND RAY WHITE*

*Howard Hughes Medical Institute, University of Utah, Salt Lake City, Utah 84132; †Howard Hughes Medical Institute and Departments of Internal Medicine and Human Genetics, University of Michigan, Ann Arbor, Michigan 48109; and ‡ABL-Basic Research Program, NCI-Frederick Cancer Research and Development Center, Frederick, Maryland 21702

Received August 6, 1990

The gene responsible for neurofibromatosis type 1 (NF1), one of the more common inherited human disorders, was identified recently, and segments of it were cloned. Two translocation breakpoints that interrupt the *NF1* gene in NF1 patients flank a 60-kb segment of DNA that contains the *EVI2A* locus (previously reported as the *EVI2* locus), the human homolog of a mouse gene, *Evi-2A*, implicated in retrovirus-induced murine myeloid tumors. *EVI2A* lies within an intron of the *NF1* gene and is transcribed from telomere toward centromere, opposite to the direction of transcription of the *NF1* gene. Here we describe a second locus, *EVI2B*, also located between the two *NF1* translocation breakpoints. Full-length cDNAs from the *EVI2B* locus detect a 2.1-kb transcript in bone marrow, peripheral blood mononuclear cells, and fibroblasts. Sequencing studies predict an *EVI2B* protein of 448 amino acids that is proline-rich and contains an N-terminal signal peptide, an extracellular domain with four potential glycosylation sites, a single hydrophobic transmembrane domain, and a cytoplasmic hydrophilic domain. At the level of genomic DNA the *EVI2B* locus lies within the same intron of the *NF1* gene as *EVI2A* and contains a 57-bp 5′ exon that is noncoding, an 8-kb intron, and a 2078-bp 3′ exon that includes the entire open reading frame. *EVI2B* is transcribed in the same direction as *EVI2A;* its 5′ exon lies only 4 kb downstream from the 3′ exon of the *EVI2A* locus. In the mouse the 5′ exon of the homologous gene, *Evi-2B*, lies approximately 2.8 kb from the 3′ end of *Evi-2A*, in the midst of a cluster of viral integration sites identified in retrovirus-induced myeloid tumors; thus, *Evi-2B* may function as an oncogene in these tumors. © 1991 Academic Press, Inc.

INTRODUCTION

Neurofibromatosis type 1 (NF1), originally described by von Recklinghausen, is an autosomal dominant disorder occurring at a frequency of 1 in 3000 (Crowe *et al.*, 1956). NF1 is a clinical diagnosis defined by criteria agreed upon in a consensus conference at the National Institutes of Health (Stumpf *et al.*, 1987). The most common physical signs of this condition are multiple subcutaneous nodules (neurofibromas) and hyperpigmented patches of skin (café-au-lait spots). Patients with NF1 are at significantly increased risk for development of malignancies (e.g., see Riccardi and Eichner, 1986), including malignant schwannoma, pheochromocytoma, and rhabdomyosarcoma.

Recently, the *NF1* gene was identified and segments of it were cloned (Cawthon *et al.*, 1990a; Viskochil *et al.*, 1990a; Wallace *et al.*, 1990a). Two balanced translocations involving chromosome band 17q11.2 have been found in two unrelated NF1 patients (Schmidt *et al.*, 1987; Ledbetter *et al.*, 1989; Menon *et al.*, 1989); the breakpoints on chromosome 17 lie approximately 60 kb apart (O'Connell *et al.*, 1990). An RNA PCR-based assay indicated that both translocations prevent transcription of the *NF1* gene (Wallace *et al.*, 1990a) on the affected chromosome 17, and hybridization studies showed that *NF1* cDNAs are interrupted by both the t(17;22) break (Viskochil *et al.*, 1990a) and the more centromeric t(1;17) break (Xu *et al.*, 1990). We have previously reported the mapping of the human homolog of a mouse gene implicated in leukemogenesis, *Evi-2A* (*e*cotropic *v*iral *i*ntegration site *2A*, originally designated *Evi-2:* Buchberg *et al.*, 1988, 1989, 1990), into the region between the two NF1 translocation breakpoints (O'Connell *et al.*, 1990) and the characterization of the human *EVI2A* transcript (previously referred to as *EVI2:* Cawthon *et al.*, 1990b). The orientation of *EVI2A* (Cawthon *et al.*, 1990b) is opposite that of *NF1* (Cawthon *et al.*,

0888-7543/91 $3.00

1990a; Viskochil *et al.*, 1990a; Wallace *et al.*, 1990a), and by hybridization studies and sequence comparison, *NF1* exons that map between the translocation breakpoints do not overlap with *EVI2A* exons. Therefore, *EVI2A* is transcribed from the anti-sense strand of the *NF1* gene and is entirely contained within an intron of the *NF1* gene.

Here we show that a second transcript, *EVI2B*, also maps into the interval between the two NF1 translocation breakpoints. We examine the pattern of *EVI2B* gene expression in several tissues, report the sequences of human *EVI2B* cDNA and genomic clones, predict features of the encoded protein product, and show the location and orientation of the gene on the genomic map of the breakpoint region. We show that in mouse genomic DNA, many ecotropic retroviral DNA integrations that are in the vicinity of the murine *Evi-2A* locus and are associated with myeloid tumor formation are located and oriented in a fashion consistent with disrupting or altering the regulation of the murine homolog of human *EVI2B* (hence the choice of *EVI2B* as the name for the new locus); these results suggest that *Evi-2B* may function as an oncogene in murine myeloid tumors. (*EVI2B* is fully characterized for the first time here; however, this locus was referred to briefly in previous publications by other names: *RC1* in Cawthon *et al.*, 1990a,b, and Viskochil *et al.*, 1990b; *NF1-c2* in Wallace *et al.*, 1990a,b).

MATERIALS AND METHODS

cDNA Libraries and Plaque Screening

Two cDNA libraries were screened. A normal adult human bone marrow cDNA library in λgt10, catalogue number HL1058a, was purchased from CLONTECH Laboratories, Inc. (Palo Alto, CA). Phage from the library were incubated with *Escherichia coli* strain LE392 and plated according to the manufacturer's instructions. A human cDNA library in λgt11 (Bonthron *et al.*, 1985) constructed from the lymphoid cell line GM-1715 (Human Genetic Mutant Cell Repository, Camden, NJ), derived from an immunodeficient patient (partial adenosine deaminase deficiency and impaired T cell immunity), was also screened. In addition, 5×10^5 pfu (plaque forming units) from this library were incubated with *E. coli* strain Y1088 and plated at a density of 25,000 pfu per 150-mm plate (176.7 cm^2). Duplicate plaque lifts were made with 0.2-μm pore-size Biotrans filters (Pall Biodyne) or 0.45-μm pore-size Immobilon-NC nitrocellulose filters (Millipore Corp.) by the method of Benton and Davis (1977). Radiolabeling of DNA probes was done according to Feinberg and Vogelstein (1984). First, inserts were separated from vector by restriction enzyme digestion and gel electrophoresis in low-melting agarose (NuSieve GTG agarose, FMC Bio-Products, Rockland, ME). DNA in gel slices was used directly for radiolabeling.

For the screening of the bone marrow cDNA library, the cDNA library filters were prehybridized in 50% formamide, 5× SSC, 50 m*M* $NaPO_4$ (pH 6.5), 2× Denhardt's solution, 5% dextran sulfate, and 500 μg/ml sheared, denatured human DNA at 42°C for 2 h. Separate prehybridization of the radiolabeled probes was simultaneously carried out in a solution of the same composition except that 50 μg/ml of DNA from the vector used to clone the probe was included. After prehybridization the probe and filters were combined and hybridization was carried out for 4–16 h at 42°C. When genomic cosmids were used as probes, washes were in 2× SSC, 0.1% SDS at room temperature. When cDNAs were used as probes, washes were in 0.1× SSC, 0.1% SDS at 55–65°C. After washing, filters were placed on X-Omat AR film with an intensifying screen at −70°C.

For the screening of the lymphoid cell line cDNA library, filters were prehybridized and hybridized in 6× SSC, 0.5% SDS, 2× Denhardt's solution, 1 m*M* EDTA, and 100 μg/ml sheared, denatured salmon sperm DNA at 65°C. Before hybridization was carried out, the radiolabeled probes were preannealed with 250 μg/ml sheared human placental DNA for 2 h. Hybridization was carried out for 12–24 h at 65°C, after which the filters were washed at 65°C to a final stringency of 0.1× SSC, 0.1% SDS. The signal was visualized on X-Omat AR film (Kodak) with an intensifying screen at −70°C overnight.

Cell Lines and Tissues

Lymphoblastoid cell lines from normal and NF1 individuals, established by transformation of peripheral blood mononuclear cells with Epstein–Barr virus, and somatic cell hybrid lines containing various portions of chromosome 17 were maintained in culture as described by O'Connell *et al.* (1990). Skin fibroblasts were grown in Dulbecco's modified Eagle's medium supplemented with 10% fetal bovine serum, and passaged every 5–7 days with a trypsin/EDTA solution. Mononuclear cells were prepared from whole peripheral blood by the LeucoPREP procedure according to the manufacturer's instructions (Becton Dickinson and Company, Lincoln Park, NJ). Bone marrow aspirates (performed at University Hospital, University of Utah Medical Center, with informed consent of the patients and Institutional Review Board approval) drawn into tubes with heparin were spun at 2000 rpm for 10 min in a tabletop centrifuge; the pellets were then used for preparation of total RNA (see below). Other tissues removed from patients for therapeutic

reasons or at autopsy were immediately placed in liquid nitrogen and stored at −135°C until used for RNA preparation. Some tissue specimens were obtained from the National Neurological Research Bank, VAMC Wadsworth (Los Angeles, CA) which is sponsored by NINCDS/NIMH, NMSS, HD Foundation, TS Association, and the Veterans Administration.

RNA Preparation and Analysis

Total RNA from peripheral blood mononuclear cells, bone marrow aspirates, and cultured cells was prepared either by the guanidinium–acid–phenol method as described by Chomczynski and Sacchi (1987) or by the vanadyl ribonucleoside complex (VRC) method described by Berger and Birkenmeier (1979). RNA from tissues that had been frozen in liquid nitrogen was prepared either by the guanidinium–acid–phenol method or by a procedure involving centrifugation in cesium chloride (Sambrook *et al.*, 1989).

Total RNA was electrophoresed through 1–1.2% agarose gels containing formaldehyde and blotted overnight onto Duralon uncharged nylon membranes (Stratagene) or unto Hybond N nylon membranes (Amersham) essentially as described by Maniatis *et al.* (1982), except that 0.37 *M* formaldehyde was used in some of the gels as described by Selden (1989) and in the electrophoresis buffer as well instead of 2.2 *M* formaldehyde. Filters were crosslinked with uv irradiation in a Stratalinker (Stratagene). Radiolabeling of probes for RNA analysis, hybridization, washing, and X-ray film exposures of RNA blots were performed as per the methods described above for screening cDNA libraries, except that preannealing of the probes was not performed in all cases.

Sequencing

All DNA sequencing was based on the dideoxy-termination method of Sanger (Sanger *et al.*, 1977). The insert from *EVI2B* cDNA clone *D* (EVI2B*D*) (λgt10) was released by digestion with *Eco*RI, digested with *Sau*3AI and *Alu*I, and subcloned into M13mp18; single-strand preparations were sequenced with fluorescently tagged M13 primers on an Applied Biosystems Inc. Model 370A DNA sequencer (ABI, Foster City, CA), using T7 DNA polymerase or *Taq*I DNA polymerase. To obtain sequence from the ends of clones EVI2B*C* and EVI2B*F*, the inserts were released with *Eco*RI and subcloned into pUC19; double-stranded plasmid preparations were made by the cesium chloride/ethidium bromide centrifugation method (Maniatis *et al.*, 1982), denatured by treatment with 0.2 *N* NaOH, neutralized, and sequenced from M13 primers on the ABI machine. To confirm sequences obtained by the above methods and to bridge gaps remaining in the cDNA sequences, oligonucleotide primers were synthesized on the basis of sequences obtained above and used in combination with primers based on pUC19 or λgt10 vector sequences to make single-stranded DNA preparations from each of the above cDNA clones and EVI2B*E* by the asymmetric polymerase chain reaction (Gyllensten, 1989). These single-strand preps were then sequenced on the ABI machine.

The inserts from EVI2B cDNA clones *A* and *B* in λgt11 were released by digestion with *Eco*RI and subcloned into the *Eco*RI site of Bluescript. Small-scale plasmid preparations (Holmes and Quigley, 1981) were subsequently used as template for dideoxy manual sequencing with Sequenase Version 2.0 (United States Biochemical Corp., Cleveland, OH). Essentially the same strategy as that described above was employed to sequence fully both strands of each clone.

DNA sequence from cosmid clone cEVI20 was obtained by multiplexed dideoxy sequencing of random shotgun clones as described previously (Cawthon *et al.*, 1990b).

DNA and Peptide Sequence Analysis

DNA sequences were aligned, examined for open reading frames, and compared to DNA sequences in the EMBL and GenBank databases using the IntelliGenetics Suite of programs (IntelliGenetics Inc., Mountain View, CA) and the Wisconsin GCG Analysis Package (Devereux *et al.*, 1984). The predicted *EVI2B* peptide was analyzed for hydropathy by the method of Kyte and Doolittle (1982) and compared to peptide sequences in the Swiss-Prot and PIR protein databases, again using the IntelliGenetics Suite. The most likely site for cleavage of the putative signal peptide was determined using the SIGSEQ1 and SIGSEQ2 programs of Folz and Gordon (1987), based on a weight-matrix method developed by von Heijne (1986).

Mapping EVI2B cDNAs onto Human and Mouse Genomic DNA

Human genomic cosmid clones of cEVI20 and cEVI36 were digested with *Eco*RI or *Bgl*II, and mouse genomic λ phage clones spanning the *Evi-2* region were digested with *Eco*RI or *Xba*I. The digested clones were then electrophoresed through Tris–acetate–agarose gels and Southern-blotted onto nylon filters. To map the 3′ exon of *EVI2B*, a PCR product containing the entire open reading frame of *EVI2B* was radiolabeled (Feinberg and Vogelstein, 1984) and hybridized to and washed from these filters essentially as described in the next section. To map the 5′ exon of *EVI2B*, the exon was synthesized as a 60-bp oligonucleotide, 5′ end-labeled with polynucleotide kinase and [γ^{32}P]ATP, and hybridized overnight to the

filters in 6× SSCP, 1% SDS, 5× Denhardt's and 0.5 mg/ml denatured salmon sperm DNA. The filters were washed in 2× SSCP, 1% SDS and visualized by autoradiography with X-Omat film and two intensifying screens at −70°C.

Screening for NF1 Mutations

Human DNA samples (5–10 μg) were restriction enzyme-digested essentially according to instructions supplied by the manufacturer (Molecular Biology Resources, Bethesda Research Laboratories, and Boehringer-Mannheim). Restriction enzyme-digested samples were fractionated on 0.8–1% agarose gels in TEAC or TBE buffer and transferred either to Gelman Biotrace RP filters in 0.4 *M* NaOH (Reed and Mann, 1985) or to Hybond N in 10× SSC essentially as described by Southern (1975). Radiolabeling of DNA probes was done according to Feinberg and Vogelstein (1984). DNA plugs for PFGE were prepared in low-melting-point agarose (Schwartz and Cantor, 1984) and run on a variety of boxes including FIGE (Carle and Olsen, 1984), TAFE (Gardiner *et al.*, 1986), and CHEF (Chu *et al.*, 1986). Electrophoresis took place in 1.0–1.5% agarose gels in 0.5× TBE buffer; DNA gels were depurinated by two 5-min washes in 0.25 *M* HCl, immediately before transfer.

Hybridizations took place in 50% formamide, 5× SSC, 50 m*M* $NaPO_4$ (pH 6.5), 2× Denhardt's solution, and 200 μg/ml sheared, denatured human DNA at 42°C or in 1 *M* NaCl, 10% dextran sulfate, 1× Denhardt's solution, 1% SDS, and 100 μg/ml sheared, denatured salmon sperm DNA at 65°C. Washes were in 0.1× SSC, 0.1% SDS at 55–65°C.

PCR–SSCP (polymerase chain reaction–single-strand conformation polymorphism) screening was performed as described by Orita *et al.* (1989). Radiolabeling of the PCR products was accomplished by incorporation of [^{32}P]dCTP during the PCR reaction as described.

RESULTS

Isolation of EVI2B cDNA Clones

The inserts of two cosmid clones, cEVI20 and cEVI36, are overlapping genomic DNA segments that span the 60-kb interval between the two NF1 translocation breakpoints (O'Connell *et al.*, 1990). The insert from cEVI36 was used to screen an oligo(dT)-primed cDNA library from normal human bone marrow in λgt10. Thirteen positive clones (from 1×10^6 plated plaques) were picked and plaque-purified (Cawthon *et al.*, 1990b); rescreening these 13 with the murine *Evi-2A* cDNA BK3 as probe identified 4 as being related to *Evi-2A*. Subsequently, cDNA clones representing sequences further 5′ in the human *EVI2A* mRNA were synthesized using the anchored PCR technique (Loh *et al.*, 1989); these new *EVI2A* cDNAs were then shown to hybridize to an additional fifth clone from the original 13 bone marrow cDNA clones (Cawthon *et al.*, 1990b). Therefore, 8 bone marrow cDNA clones appeared to be unrelated to *EVI2A*. The 3.7-kb insert from the largest of these 8 hybridized to all 8 clones and to none of the 5 *EVI2A*-related clones. These results indicate that all 8 cDNA clones belong to a class independent from *EVI2A* cDNAs.

Cosmid cEVI20 served as the starting point for an approach that resulted in the subsequent cloning of two additional independent cDNA clones related to the eight described above. A jump clone, pEH1 (Wallace *et al.*, 1990b), was obtained by screening a jumping library with a 1.2-kb fragment of a 4.4-kb *Eco*RI subclone from the cosmid cEVI20 (O'Connell *et al.*, 1990). The insert of pEH1, consisting of 4.6 kb partially overlapping with cEVI20 (beginning of jump, BOJ) (Fig. 5) as well as 5.7 kb derived from the end of jump approximately 50–60 kb further telomeric along the chromosome, was used as a probe against an RNA panel. A transcript was detected in B lymphoblasts, and EH1 was subsequently used as a probe to screen the human B-lymphoblast cDNA library in λgt11. Of 5×10^5 pfu screened, five positive clones were picked, plaque-purified, and subcloned. Mapping experiments showed that these cDNA clones were identified by the portion of pEH1 that is the beginning of jump. The insert size of four of these clones was 0.9 kb and the size of the fifth clone was 1.2 kb. Sequencing of the clones showed the four 0.9-kb clones to be identical (EVI2B*A*), whereas the 1.2-kb clone (EVI2B*B*) and the 0.9-kb clones showed an overlap of 378 bp.

Mapping the EVI2B cDNAs between the NF1 Translocation Breakpoints

Somatic cell hybrid lines containing NF1 translocation chromosomes provide a means of precisely mapping genes relative to the translocation breakpoints. The cell line NF13 contains the derivative chromosome 22 [der(22)] from an NF1 patient with a balanced translocation, t(17;22)(q11.2;q11.2) (Ledbetter *et al.*, 1989). Cell line DCR1 contains the derivative chromosome 1 [der(1)] from an NF1 patient with a different balanced translocation, t(1;17)(p34.3;q11.2) (Menon *et al.*, 1989). In both NF13 and DCR1 cell lines, the only chromosome 17 content is the portion of the long arm of 17 from the breakpoint to the telomere. Therefore, a gene mapping between the breakpoints would be expected to hybridize to DNA from the hybrid cell containing the proximal translocation breakpoint, but would show no hybridization to DNA from the hybrid containing the distal translocation breakpoint. Figure 1 shows the hybrid-

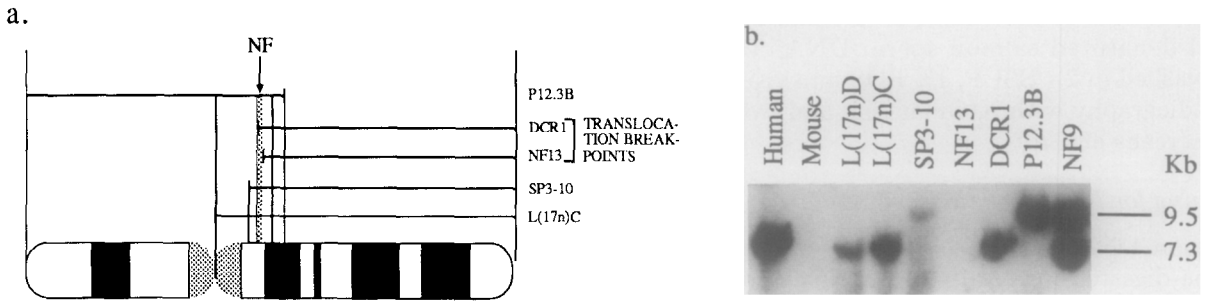

FIG. 1. (**a**) Chromosome 17 content of somatic cell hybrid cell lines constituting the mapping panel for the *NF1* region. The name of each cell line is listed on the right; the portion of chromosome 17 present in each cell line is indicated by a line segment overlying the diagram of the chromosome; the *NF1* region between the translocation breakpoints is shaded (not shown to scale). (**b**) Hybridization of radiolabeled *EVI2B* cDNA *D* to Southern blots of *Eco*RI-digested total genomic DNAs from human and mouse and somatic cell hybrid DNA from the mapping panel. The 9.5- and 7.3-kb bands are indicated. Cell line L(17n)D contains an entire human non-NF1 chromosome 17; cell line NF9 includes the translocated chromosome 17 proximal to the t(17;22) break and contains the (nontranslocated) chromosome 17 homolog as well.

ization results for cDNA EVI2B*D*. EVI2B*D* detects a single-copy locus. There is an *Eco*RI restriction fragment length polymorphism with two common alleles in the population (Andersen *et al.*, 1990), identical to that detected by the genomic fragment clone pHU39 (O'Connell *et al.*, 1990). Both alleles, 9.5 and 7.3 kb, are represented in Fig. 1. The absence of the human band in the NF13 lane shows that *EVI2B* sequences are missing from the portion of the long arm of chromosome 17 that is distal to the t(17;22) breakpoint. The presence of the human band in the DCR1 lane indicates that *EVI2B* sequences are present in the segment of chromosome 17 distal to the t(1;17) breakpoint. The NF9 hybrid cell line is also derived from the patient with the t(17;22) balanced translocation. Its karyotype includes the portion of the translocated chromosome 17 that is proximal to the break, and the homologous chromosome 17 as well (D. Ledbetter, personal communication). Since both the 9.5- and 7.3-kb alleles are seen in the NF9 cell line, we can conclude that the t(17;22) translocation does not involve a deletion that removes the *EVI2B* locus. *EVI2B*, therefore, maps to a region between the NF1 translocation breakpoints.

Analysis of EVI2B mRNA Expression

EVI2B cDNAs identify a 2.1-kb transcript on RNA blots of bone marrow and other tissues (Fig. 2). With autoradiographic exposures of 1–3 days the *EVI2B* transcript was most abundant in 5-μg samples of total RNA from peripheral blood mononuclear cells, bone marrow, Epstein–Barr virus (EBV)-transformed lymphoblastoid cell lines, and T cells and detectable in fibroblasts. No significant difference in the level of expression of *EVI2B* between T cells and EBV-transformed lymphoblasts (of B-cell origin) was detected. With prolonged autoradiographic exposures (10 days) *EVI2B* transcripts were detected in peripheral nerve, neurofibroma, malignant schwannoma, adrenal gland, brain frontal cortex, and liver (data not shown); however, the low abundance of *EVI2B* transcripts in these pathologic specimens and autopsied tissues could be due to the presence of peripheral blood cells in the vasculature of these tissues. *EVI2B* transcript was not detected in a choriocarcinoma cell line.

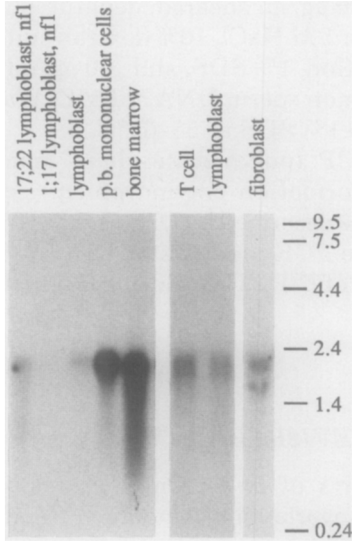

FIG. 2. RNA blots of human tissues, probed with *EVI2B* cDNAs. Approximately 5 μg of total RNA was loaded in each lane. The mobilities of RNA standards of known size are shown to the right of the blots. EBV-transformed B cells grown in tissue culture are labeled "lymphoblast." In the fourth lane, "p.b." indicates "peripheral blood." T cells were fractionated from whole blood. Fibroblasts were grown in tissue culture. Lanes **1** and **2** contain lymphoblast RNA from the NF1 patients with the t(17;22) and t(1;17) balanced translocations; all other samples are from non-NF1 subjects. Lanes **1–5** were exposed to film for 2 days; lanes **6** and **7** for 3 days; and lane **8** for 5 days.

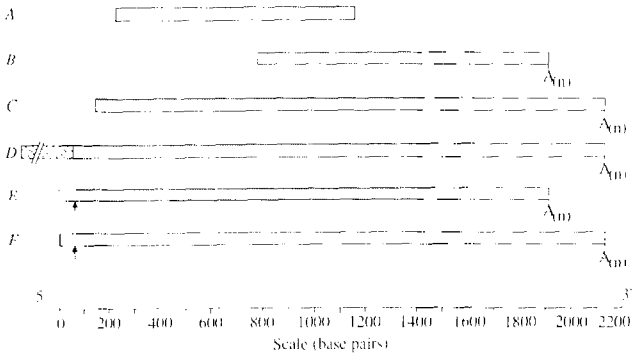

FIG. 3. Human *EVI2B* cDNA clones. The scale is in nucleotide base pairs. *A* and *B* are B-lymphoblast cDNAs. *C*, *D*, *E*, and *F* are bone marrow cDNAs. The stippled segments contain the *EVI2B* open reading frame. The solid white segments are 5′ and 3′ noncoding *EVI2B* sequences. The cross-hatched segment in clone *D*, representing *EVI2B* intron sequences, is approximately 1.6 kb long and is not shown in its entirety. $A_{(n)}$ designates the presence of a poly(A) tail. Clones *E* and *F* are independent full-length cDNAs, differing at the 3′ end but ending at the same base at the 5′ end. In the genome an intron, approximately 8 kb in length, interrupts the sequence of clones *E* and *F* at the position indicated by the arrow.

A second transcript, approximately 1.9 kb, was detected in fibroblasts (Fig. 2) and, after longer (5-day) autoradiographic exposures, also detected in EBV-transformed lymphoblasts (data not shown). Apparently, the relative proportion of the 2.1- and 1.9-kb transcripts varies in different tissues. Characterization of cDNA clones from the bone marrow and B-lymphoblast cDNA libraries (see next section) suggests that the 1.9-kb transcript probably is identical to the 5′ 1.9 kb of the 2.1-kb transcript.

The blot shown in Fig. 2 includes RNA samples from the two NF1 lymphoblastoid cell lines containing, respectively, the t(17;22) and t(1;17) balanced translocations. *EVI2B* transcripts of abnormal mobility were not detected in these samples. Furthermore, EBV-transformed lymphoblast RNAs from 20 unrelated NF1 patients and 7 non-NF1 controls were screened with cDNA clone EVI2B*D* or a mixture of EVI2B*A* and *B*. No *EVI2B* transcripts of aberrant mobility were found, and no significant difference in the levels of expression of *EVI2B* between NF1 samples and controls was detected.

cDNA and Genomic Nucleotide Sequences and Predicted Amino Acid Sequence of EVI2B

Figure 3 is a schematic representation of several of the *EVI2B* cDNA clones that we have isolated. EVI2B*E* and EVI2B*F* appear to be independent, full-length *EVI2B* cDNA clones. The EVI2B*E* and EVI2B*F* sequences are identical for 1915 bp, beginning with the first base pair at their 5′ ends. EVI2B*E* terminates 3′ of bp 1915 in a poly(A) tract containing 23 A's, but EVI2B*F* continues without a poly(A) tract through bp 2135, then ends with a tract of 12 A's. The size of EVI2B*F* fits well with the size of the major transcript seen on RNA blots; EVI2B*E* may represent the smaller size transcript. Clone EVI2B*B* also appears to be derived from the smaller transcript. Its sequence is identical to the sequence of a segment of the EVI2B*F* cDNA, but like clone EVI2B*E*, its 3′ end terminates approximately 220 bp 5′ of the 3′ end of the EVI2B*F* sequence; however, the poly(A) tract for EVI2B*B* begins one base further 3′ than does the poly(A) tract of EVI2B*E*. Clones EVI2B*C* and EVI2B*D* terminate at their 3′ ends at the same base as clone EVI2B*F*. The sequences of clones EVI2B*B* and EVI2B*C* do not continue as far 5′ as those for EVI2B*E* or EVI2B*F*, presumably due to incomplete first-strand cDNA synthesis during the construction of the cDNA libraries. EVI2B*A* is identical in sequence to a segment of EVI2B*F*; it lacks a poly(A) tract, and it overlaps EVI2B*B* for 378 bp. EVI2B*D* is 3.7 kb long and appears to be derived from an unprocessed *EVI2B* RNA (see below).

Comparing the cDNA sequences to genomic sequences obtained from a cloned cosmid, cEVI20 (Weiss, in preparation), we were able to determine that in the genome an intron interrupts the sequence of cDNA clones EVI2B*E* and EVI2B*F* at the point indicated in Fig. 3. The first 57 bp at the 5′ end of the EVI2B*E* and EVI2B*F* cDNA clones are identical to a 57-bp genomic sequence that is contained within one cosmid subclone; further 3′ the genomic sequence diverges from the cDNA sequence. The sequence around the divergence point (with capital letters for sequence shared by cDNAs and genome, small letters for sequences present only in the genome, and a slash to mark the divergence point) is GAG/gtaaga, which approximates the primate 5′ splice site consensus sequence NAG/gtragt (Shapiro and Senapathy, 1987). The next 94 bp of EVI2B*E* and EVI2B*F* cDNA sequence are identical to the last 94 bp of a sequenced portion of another cEVI20 cosmid subclone; however, the genomic sequence 5′ of these 94 bp diverges from the cDNA sequence. The sequence around this divergence is cccgctatag/ATAACG, consistent with a 3′ splice junction. Mapping studies (presented below) indicate that this intron is approximately 8 kb long and that there are no additional introns. cDNA EVI2B*D* apparently is derived from an unprocessed *EVI2B* transcript; it is 3.7 kb long, lacks the 57-bp 5′ exon found in clones EVI2B*E* and EVI2B*F*, and matches cDNA EVI2B*F* starting just 3′ of the 3′ splice junction described above and continuing up to the beginning of the poly(A) tails (2078 bp). In addition, 5′ of the sequence shared with EVI2B*F*, the sequence of clone EVI2B*D* matches the genomic sequence lying

immediately 5′ of this 3′ splice junction in a cEVI20 subclone.

Although we have not yet sequenced all of the genomic DNA that gives rise to the full-length *EVI2B* cDNAs, experiments indicate that the cDNA sequences 3′ of the 57 bp exon described above are not interrupted in the genome. First, a PCR experiment was performed using as primers an oligonucleotide located in the already identified intron, within the 40 bp adjacent to the 3′ splice junction, in combination with a primer located in the EVI2B*F* cDNA sequence within the 25 bp adjacent to the poly(A) tail. Single amplified products, indistinguishable in size by agarose gel electrophoresis and ethidium bromide staining, were obtained from genomic DNA, from cosmid cEVI36, and from cDNA clone EVI2B*D* (data not shown). Furthermore, six pairs of oligonucleotide primers based on sequences located within cDNA clone EVI2B*D* have been used to PCR-amplify a contiguous set of overlapping segments (approximate length 400 bp per segment) that includes all of the EVI2B*F* cDNA sequence between the known 3′ splice junction and the poly(A) tail. For each of the primer pairs, the amplified products from genomic DNA, cEVI36, and cDNA EVI2B*D* were indistinguishable in size by gel electrophoresis in 4% agarose gels and ethidium bromide staining. These results indicate that the terminal 2078 bp of cDNA EVI2B*F* are most likely uninterrupted in the genome, and therefore constitute a single exon.

The cDNA sequence, genomic splice-junction sequences, predicted amino acid sequence, and a hydropathy plot of human *EVI2B* are shown in Figs. 4a and 4b. The intervening sequence lies between nucleotide positions −21 and −22 of the cDNA sequence. A single, large open reading frame (ORF) extends from the A at the +1 position to the T at +1344. The ATG starting at position +1 is the first ATG in this reading frame, and it matches the Kozak consensus initiation sequence 5′-RNNATGG-3′ for eukaryotic mRNAs (Kozak, 1984). There are frequent stop codons in both of the remaining two reading frames.

The *EVI2B* ORF predicts a 448-amino-acid protein. The N-terminal segment has features suggesting a signal peptide (von Heijne, 1985; Gierasch, 1989), with cleavage most likely occurring after the serine residue at position 21 (von Heijne, 1986; Folz and Gordon, 1987). Next is a 181-residue hydrophilic, presumably extracellular domain, with four potential N-glycosylation sites (nx(s/t)). Then follows a segment of 24 hydrophobic amino acids with an average hydropathy score of 2.6 by a Kyte and Doolittle (1982) analysis that indicates it is a transmembrane domain. The remaining hydrophilic domain of 222 amino acid residues is presumably cytoplasmic. Like the hydrophilic domain amino-terminal to the transmembrane domain, this hydrophilic domain carboxy-terminal of the transmembrane domain also has four potential N-glycosylation sites. *EVI2B* is proline-rich; 47 of the 448 amino acids are proline (10.5%), and there are five proline doublets, one stretch of four consecutive prolines, and one stretch of five consecutive prolines. There are six cysteine residues, one in the putative signal peptide and five carboxy-terminal to the transmembrane domain. Structural motifs characteristic of known oncogenes have been sought, but not detected. Comparison of the cDNA sequence to sequences in the EMBL and GenBank databases, comparison of the predicted peptide sequence to peptides obtained by translating the DNA database sequences in all six reading frames, and comparison of the predicted peptide to protein sequences in the PIR and Swiss-Prot databases have shown that *EVI2B* is a novel gene with no significant specific similarities to previously reported DNA or protein sequences.

The 3′ untranslated sequence of the *EVI2B* mRNA contains four instances of the AUUUA pentamer. This motif is common to the 3′ untranslated regions of many oncogene and cytokine mRNAs, and it appears to promote rapid cytoplasmic degradation of mRNAs (Shaw and Kamen, 1986; Greenberg *et al.*, 1986; Wilson and Treisman, 1988; Brawerman, 1989; Malter, 1989).

Location of EVI2B with Respect to the NF1 Translocation Breakpoints

As shown in Fig. 5, the entire *EVI2B* full-length cDNA sequence (i.e., the entire sequence of the EVI2B*F* cDNA) lies within the interval between the translocation breakpoints and consists of two exons separated in genomic DNA by an 8-kb intron. Transcription is in the direction indicated by the arrow; i.e., the 5′ end of the sense strand is toward the telomere of 17q. This genomic organization was deduced from the following experiments.

The *Bam*HI and *Eco*RI restriction map of the breakpoint region and the *Bgl*II restriction map of cosmid cEVI36 (O'Connell *et al.*, 1990) are shown in Fig. 5. To locate the *EVI2B* 3′ exon on this restriction map, cDNA EVI2BC, which consists of the terminal 2 kb of the 3′ exon, was radiolabeled and hybridized to Southern blots of *Bam*HI- and *Eco*RI-digested cEVI20 and cEVI36 cosmid clones. In *Bam*HI/*Eco*RI doubly digested samples, only the 1.2-kb *Bam*HI centromeric end fragment of cosmid cEVI20 and the 6-kb *Eco*RI/*Bam*HI fragment of cEVI36 that this cEVI20 end fragment overlaps hybridized to the EVI2BC probe. The centromeric end of cEVI20 has been sequenced for 267 bp; comparison to the sequence of the *EVI2B* cDNAs showed that these 267 bp are identical to bp 893–1159 of the EVI2B*F* sequence (Fig. 4a) and

```
                                                                5' CAAAATGTCCAGTTAGAACAAGAATTTAGCATTCTGCAAAAGAA -215
GTTAACAGCTGAGgtaagaaaattttcacacaatttggtttgatcttgctttatataaaatttatttttaaaactgtcattagtgttttaaacttaaactctgatagt -108
cta...aatataatgaaaagtcaaagttttaactagacaccaatgacgcctaactgtctttctctttcattataaacccgctatagATAACGAGGAAATATTCTGAA   -1

ATG GAT CCC AAA TAT TTC ATC TTA ATT TTG TTT TGT GGA CAC CTG AAC AAT ACA TTT TTT TCA AAG ACA GAG ACA ATT ACA   81
M  D  P  K  Y  F  I  L  I  L  F  C  G  H  L  N  N  T  F  F  S ▲ K  T  E  T  I  T    27

ACA GAG AAG CAG TCA CAG CCT ACC TTA TAC ACA TCA TCA ATG TCA CAG GTA TTG GCT AAT TCT CAA AAC ACA ACA GGG AAT  162
T  E  K  Q  S  Q  P  T  L  Y  T  S  S  M  S  Q  V  L  A  N  S  Q  N  T  T  G  N     54

CCT TTG GGT CAA CCA ACA CAA TTC AGC GAC ACT TTT TCT GGA CAA TCA ATA TCA CCT GCC AAA GTC ACT GCT GGA CAA CCA  243
P  L  G  Q  P  T  Q  F  S  D  T  F  S  G  Q  S  I  S  P  A  K  V  T  A  G  Q  P     81

ACA CCA GCT GTC TAT ACC TCT TCT GAA AAA CCA GAA GCA CAT ACT TCT GCT GGA CAA CCA CTT GCC TAC AAC ACC AAA CAA  324
T  P  A  V  Y  T  S  S  E  K  P  E  A  H  T  S  A  G  Q  P  L  A  Y  N  T  K  Q    108

CCA ACA CCA ATA GCC AAC ACC TCC TCC CAG CAA GCC GTG TTC ACC TCT GCC AGA CAA CTA CCA TCT GCC CGT ACT TCT ACC  405
P  T  P  I  A  N  T  S  S  Q  Q  A  V  F  T  S  A  R  Q  L  P  S  A  R  T  S  T    135

ACA CAA CCA CCA AAG TCA TTT GTC TAT ACT TTT ACT CAA CAA TCA TCA TCT GTC CAG ATC CCT TCT AGA AAA CAA ATA ACT  486
T  Q  P  P  K  S  F  V  Y  T  F  T  Q  Q  S  S  S  V  Q  I  P  S  R  K  Q  I  T    162

GTT CAT AAT CCA TCC ACA CAA CCA ACA TCA ACT GTC AAA AAT TCA CCT AGG AGT ACA CCA GGA TTT ATC TTA GAT ACT ACC  567
V  H  N  P  S  T  Q  P  T  S  T  V  K  N  S  P  R  S  T  P  G  F  I  L  D  T  T    189

AGT AAC AAA CAA ACC CCA CAA AAA AAC AAT TAT AAT TCA ATA GCT GCC ATA CTA ATT GGT GTA CTT CTG ACT TCT ATG TTG  648
S  N  K  Q  T  P  Q  K  N  N  Y  N  S  I  A  A  I  L  I  G  V  L  L  T  S  M  L    216

GTA GCT ATA ATC ATC ATT GTA CTT TGG AAA TGC TTA AGG AAA CCA GTT TTA AAT GAT CAA AAT TGG GCA GGT AGA TCT CCA  729
V  A  I  I  I  I  V  L  W  K  C  L  R  K  P  V  L  N  D  Q  N  W  A  G  R  S  P    243

TTT GCT GAT GGA GAA ACC CCT GAC ATT TGT ATG GAT AAC ATC AGA GAA AAT GAA ATA TCC ACA AAA CGT ACA TCA ATC ATT  810
F  A  D  G  E  T  P  D  I  C  M  D  N  I  R  E  N  E  I  S  T  K  R  T  S  I  I    270

TCA CTT ACA CCC TGG AAA CCA AGC AAA AGC ACA CTT TTA GCA GAT GAC TTA GAA ATT AAG TTG TTT GAA TCA AGT GAA AAC  891
S  L  T  P  W  K  P  S  K  S  T  L  L  A  D  D  L  E  I  K  L  F  E  S  S  E  N    297

ATT GAA GAC TCC AAC AAC CCC AAA ACA GAG AAA ATA AAA GAT CAA GTA AAT GGT ACA TCA GAA GAT AGT GCT GAT GGT TCA  972
I  E  D  S  N  N  P  K  T  E  K  I  K  D  Q  V  N  G  T  S  E  D  S  A  D  G  S    324

ACA GTT GGA ACT GCT GTT TCT TCT TCA GAT GAT GCA GAT CTG CCT CCA CCA CCT CCC CTT CTG GAT TTG GAA GGA CAG GAA 1053
T  V  G  T  A  V  S  S  S  D  D  A  D  L  P  P  P  P  P  L  L  D  L  E  G  Q  E    351

AGT AAC CAA TCT GAC AAA CCC ACA ATG ACA ATT GTA TCT CCT CTT CCA AAT GAT TCT ACT AGT CTC CCT CCA TCT CTG GAC 1134
S  N  Q  S  D  K  P  T  M  T  I  V  S  P  L  P  N  D  S  T  S  L  P  P  S  L  D    378

TGT CTC AAT CAA GAC TGT GGA GAT CAT AAA TCT GAG ATA ATA CAA TCA TTT CCA CCG CTT GAC TCA CTT AAC TTG CCC CTG 1215
C  L  N  Q  D  C  G  D  H  K  S  E  I  I  Q  S  F  P  P  L  D  S  L  N  L  P  L    405

CCA CCA GTA GAT TTT ATG AAA AAC CAA GAA GAT TCC AAC CTT GAG ATC CAG TGT CAG GAG TTC TCT ATT CCT CCC AAC TCT 1296
P  P  V  D  F  M  K  N  Q  E  D  S  N  L  E  I  Q  C  Q  E  F  S  I  P  P  N  S    432

GAT CAA GAT CTT AAT GAA TCC CTG CCA CCT CCA CCT GCA GAA CTG TTA TAA ATATTACAACTTGCTTTTTAGCTGATCTTCCATCCTCAA 1386
D  Q  D  L  N  E  S  L  P  P  P  P  A  E  L  L  .                                                        448

ATGACTCTTTTTTCTTTATATGTTAACATATATAAAATGGCAACTGATAGTCAATTTTGATTTTTATTCAGGAACTATCTGAAATCTGCTCAGAGCCTATGTGCATA 1493
GATGAAACTTTTTTTTAAAAAAAGTTATTTAACAGTAATCTATTTACTAATTATAGTACCTATCTTTAAAGTATAGTACATTTTACATATGTAAATGGTATGTTTCA 1600
ATAATTTAAGAACTCTGAAACAATCTACATATACTTATTACCCAGTACAGTTTTTTTTCCCCTGAAAAGCTGTGTATAAAATTATGGTGAATAAACTTTTATGTTTC 1707
CATTTCAAAGACCAGGGTGGAGAGGAATAAGAGACTAAGTATATGCTTCAAGTTTTAAATTAATACCTCAAGTATTAAATAAATATTCCAAGTTTGTGGGAATGGGA 1814
GATTAAAATGCATGTTTGAGAATAGAGAAATTTTCTTCTTGGTTTCATTGCAAAGAGTAAAACAAACATGTTAAAACATCAACTGAAGGGTTGGGTTAGGAACATTT 1921
ACCCTGAAAAAAATATGAGGATGCATCATAAAATGTAAATATTTTCCTACCATGTTGGGGGGGCACAAATTTTAAAACTGGCATCTTTACAAGTTTCTTCTTTATAA 2028
ACACCCAAACAAAATCAAGTTTTATAAAGAAAAAAAAAAAA 3'                                                              2057
```

FIG. 4a. cDNA and genomic nucleotide sequences and predicted amino acid sequence of *EVI2B*. The *EVI2B F* cDNA nucleotide sequence is capitalized; genomic sequences from the intron are in lowercase. The cDNA sequence is numbered beginning with the putative initiating ATG codon and ending with the G nucleotide at position 2057; this G is followed by a series of 12 A's that end at the *Eco*RI cloning site of the vector. Sequences in the intron are not numbered; ellipses . . . represent the omitted portion of the 8-kb intron. In the open reading frame the one-letter amino acid translation is shown below each codon; the arrowhead between predicted amino acids 21 and 22 indicates the most likely cleavage site for the putative signal peptide; single underlining denotes potential sites for N-linked glycosylation; and double underlining indicates the strongly hydrophobic, presumably transmembrane segment of the predicted protein. In the 3′ noncoding sequence single underlining identifies potential polyadenylation signals, and double underlining marks a reiterated pentamer shared by rapidly degraded mRNAs.

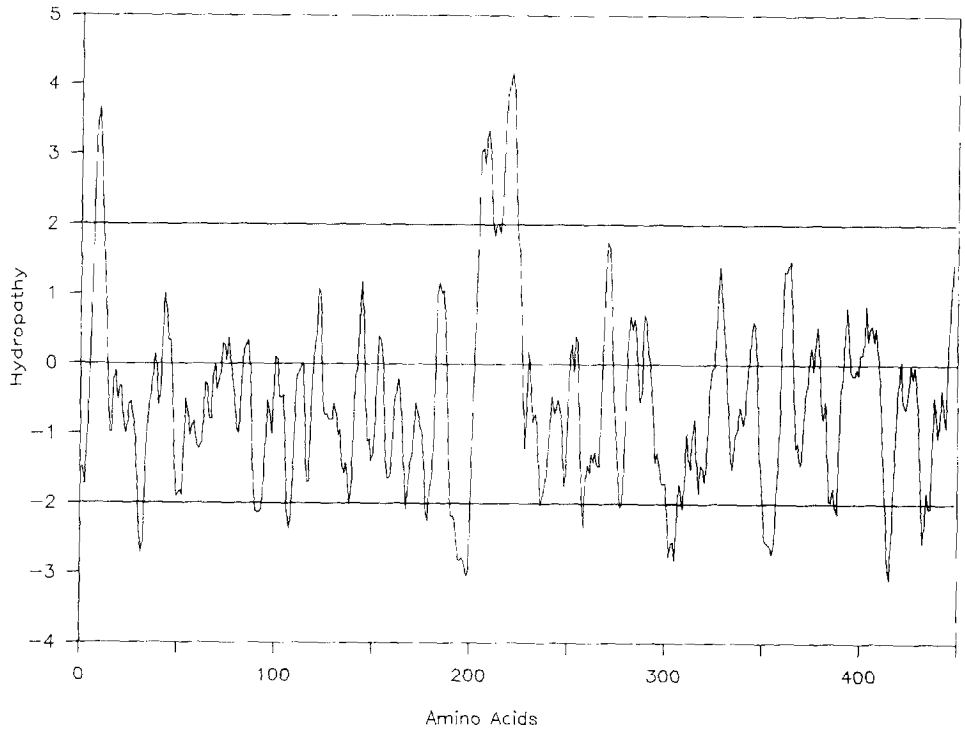

FIG. 4b. Hydropathy profile of predicted EVI2B protein. Deflections above zero indicate hydrophobic regions.

that the cDNA sequence is oriented with its 3′ end toward the centromere. In line with this, 199 bp of the beginning of the jumping clone EH1, from which the cDNA clones EVI2B*A* and EVI2B*B* were derived, was sequenced. These base pairs are identical to bp 1138–1336 of EVI2B*F*, also contained within EVI2B*B*. Due to the directionality of jumping (Collins *et al.*, 1987) the orientation of this segment was known and thus the 3′ end of the cDNA clones could be determined to be toward the centromere.

To locate the 5′ exon of the EVI2B*F* cDNA clone on the restriction map of the breakpoint interval, an oli-

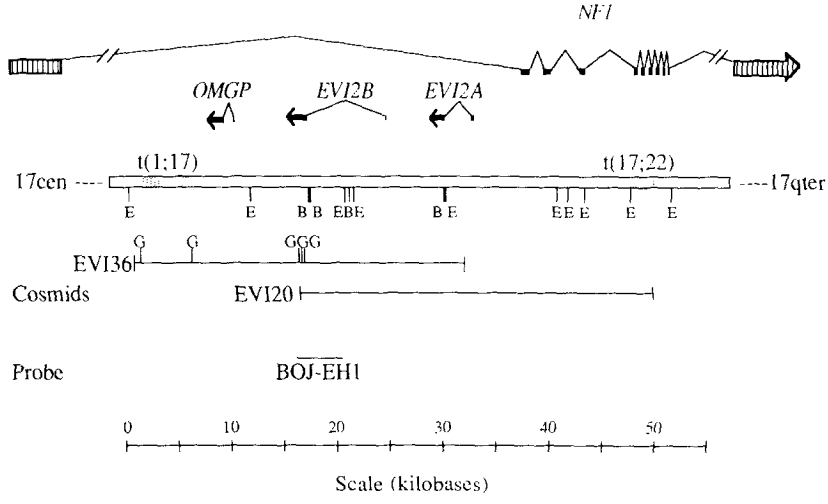

FIG. 5. Location of transcripts in the translocation-breakpoint region. The long open bar is a *Bam*HI (B) and *Eco*RI (E) restriction map of the region of chromosome 17q containing the two NF1 translocation breakpoints. The centromere is toward the left, the telomere toward the right. The t(1;17) and t(17;22) breakpoints lie somewhere within the respective stippled sections. Below the restriction map are shown two genomic cosmids that span the region and a 4.6-kb segment of a jump clone, pEH1. The *Bgl*II (G) restriction map of cosmid cEVI36 is included. Above the *Eco*RI/*Bam*HI restriction map cDNA clones are represented as heavy arrows overlying the restriction fragments to which they hybridize. The length of an arrow or heavy bar is scaled to the length of the isolated cDNA; the direction of the arrow indicates the direction of transcription.

gonucleotide containing it was synthesized, end-labeled with ^{32}P, and hybridized to a Southern blot of *Eco*RI-digested cosmids cEVI20 and cEVI36. The only *Eco*RI fragment hybridized by this oligonucleotide is approximately 8.5 kb long and shared by both cosmids; this result places the *EVI2B* 5′ exon somewhere within this restriction fragment (data not shown); in Fig. 5 the exon is shown overlying this *Eco*RI fragment. To determine the orientation and approximate location of the exon within the *Eco*RI fragment, PCR was performed using primers based on the sequenced genomic segment containing the exon, in combination with primers whose location on the map was known; cosmids cEVI20 and cEVI36 were used as template DNA. Primers from the *EVI2B* 5′ exon in combination with primers from the 3′ end of the 3′ exon of the *EVI2A* gene yielded single DNA bands by agarose gel electrophoresis and ethidium bromide staining. The sizes of these bands and the known positions of the *EVI2A* primers relative to the telomeric end of the *Eco*RI fragment locate and orient the *EVI2B* 5′ exon within the *Eco*RI fragment as shown in Fig. 5. The *EVI2B* 5′ exon lies only 4 kb centromeric of the 3′ end of the *EVI2A* polyadenylation site.

Taken together, the placements of the 5′ and 3′ exons on the restriction map of the breakpoint region indicate that the *EVI2B* intron is approximately 8 kb in size.

As shown in Fig. 5, the polyadenylation site of *EVI2B* cDNA clone *F* lies approximately 5 kb upstream from the 5′ most sequences known to be a part of the *OMGP* 5′ exon (Viskochil *et al.*, 1990). *OMGP* is the gene encoding the oligodendrocyte-myelin-glycoprotein (Mikol *et al.*, 1990).

Identification of the Murine Homolog of EVI2B: Genomic Organization and Proximity to Evi-2A

Evi-2A (in previous publications designated *Evi-2*) was originally identified as a common site of viral integration in virally induced myeloid tumors in BXH-2 mice (Buchberg *et al.*, 1988, 1989, 1990); this finding suggested that *Evi-2A* could be a proto-oncogene. Since the *EVI2B* locus in the human was found to be located quite near human *EVI2A*, the possibility was raised that the murine homolog of *EVI2B* is similarly located close to murine *Evi-2A*, is affected by the viral integrations, and contributes to tumor formation independently or in conjunction with murine *Evi-2A*. Figure 6 is a map of murine genomic DNA from chromosome 11, showing the location and orientation of the integrated proviruses within the *Evi-2A* locus with respect to the exons of *Evi-2A*. In general, for any particular tumor only a single viral integration occurred in the vicinity of *Evi-2A* (exceptions are described below); the figure compiles the various sites found in surveying the DNA from several tumors. As can be seen, four viral integrations occur within the intron of *Evi-2A*, three occur within the 3′ exon, one occurs 5′ of the gene, while six occur 3′ of *Evi-2A* (Buchberg *et al.*, 1990). To determine the location of the murine homolog of *EVI2B*, we first used a cDNA probe representing the 3′ exon of human *EVI2B*. In the mouse the genomic DNA segment identified by this probe lies 5.6 kb from the 3′ end of *Evi-2A*, a location homologous to that in the human. Moreover, using a 60-bp probe representing the 5′ exon of human *EVI2B*, the corresponding murine genomic DNA segment was localized to a 500-bp *Bam*HI–*Eco*RI fragment located 2.8 kb from the 3′ end of *Evi-2A*. The location of this 5′ exon puts it in the midst of the viral integrations occurring 3′ of *Evi-2A* (Fig. 6).

Search for EVI2B Mutations Specific to NF1 Patients

As yet, no mutations in *EVI2B* specific to NF1 patients have been found, although several techniques capable of detecting different types of mutation have been applied in the course of this search. All NF1 patients were diagnosed by the criteria of Stumpf *et al.* (1987). To screen for large-scale deletions or insertions in the region of *EVI2B*, *EVI2B* cDNA clones were used to probe Southern transfers of genomic DNA from 38 independent NF1 lymphoblastoid cell lines and several non-NF1 controls, after digestion of the DNA with rarely cutting restriction endonucleases (*Nru*I, *Sac*II, and *Not*I) followed by pulsed-field gel electrophoresis. To screen for medium-scale deletions or insertions, these same probes were used to screen Southern transfers of genomic DNA from 126 unrelated NF1 patients after digestion with restriction enzymes with six-base recognition sites (*Eco*RI and *Bam*HI) and standard agarose gel electrophoresis. To screen for small deletions, small insertions, and point mutations in the ORF of *EVI2B*, we analyzed genomic DNA from four sporadic NF1 patients and their eight (unaffected) parents, by the PCR–SSCP method of Orita *et al.* (1989); for each individual, four pairs of oligonucleotide primers were used to amplify four overlapping segments (approximately 380 bp each) encompassing the entire *EVI2B* ORF, and the amplified products were heat-denatured and subjected to electrophoresis through 4–5% nondenaturing polyacrylamide gels. The *EVI2B* ORFs from an additional 40 unrelated NF1 patients were also examined by PCR–SSCP; however, for these samples the entire ORF was amplified as part of a single 1.5-kb PCR product, then digested with either *Alu* I or *Sau* 3A I prior to heat denaturation and gel electrophoresis. None of the above procedures identified a restric-

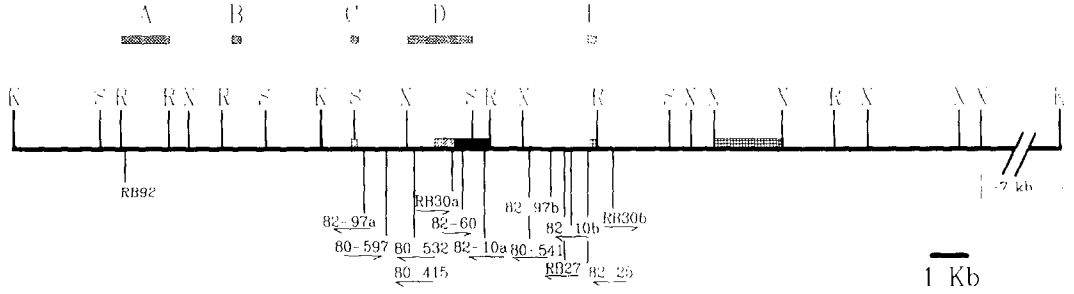

FIG. 6. Restriction map of the murine *Evi-2A–Evi-2B* region. The restriction map of the region, represented by the solid line, was generated using *Eco*RI (R), *Kpn*I (K), *Sst*I (S), and *Xba*I (X). The *Kpn*I site at the right end of the map was determined by Southern blot analysis of genomic DNA. The cross-hatched boxes (A–E) above the restriction map represent the location of five unique sequence probes from the region (12). The checkered boxes represent the two exons of *Evi-2B*. The solid box represents the transcribed region that is shared in common by the two *Evi-2A* transcripts found in mouse tissues; the two hatched boxes represent the regions unique to each of two *Evi-2A* cDNA clones. The location and orientation 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.

tion fragment of PCR product specific to NF1 patients.

DISCUSSION

The proximity of murine *EVI2B* to a common site of viral integration in virally induced myeloid tumors suggests that this locus may encode a proto-oncogene. BXH-2 mice express chronic leukemia viruses beginning at Days 14–16 of gestation (Bedigian *et al.*, 1981). These viruses induce disease in susceptible hosts via the insertional activation or mutation of cellular proto-oncogenes (reviewed by Copeland *et al.*, 1989). Since retroviruses integrate into many distinct sites in the host genome, the detection of a common viral integration site in tumors suggests that the locus encodes a gene whose activation or mutation by viral integration predisposes the target cell to neoplastic disease. *Evi-2A* was found to be a site of viral integration in 11 of 69 BXH-2 tumors analyzed; interestingly, three of these tumors had two independent integrations within the *Evi-2A* locus (Buchberg *et al.*, 1990). A transcription unit was identified, within the region defined by the viral integration sites, that is capable of encoding a 223-amino-acid protein that has all the structural hallmarks of a transmembrane protein containing a leucine zipper within the transmembrane domain (Buchberg *et al.*, 1990). However, since we did not have a source of RNA from the tumors containing viral integrations within the *Evi-2A* locus, we were unable to determine whether viral integrations altered *Evi-2A* expression.

The location and orientation of the eight integrated proviruses located 5′ or within *Evi-2A* gene (see Fig. 6) are consistent with their being involved in the alteration of transcription of the *Evi-2A* locus. However, viral integrations that occur at the 3′ end of genes usually are oriented in the same transcriptional orientation as the gene they affect (Copeland *et al.*, 1989), though there have been reports of viral integrations in the opposite transcriptional orientation (Morris *et al.*, 1990). Of the six proviral integrations 3′ of *Evi-2A*, one is in the same orientation, four are in the opposite orientation, and one has an unknown orientation. The large number of integrations in the 3′ end of *Evi-2A* in the orientation opposite from that of *Evi-2A* locus suggests either a novel mechanism of gene activation or the presence of another transcript in the vicinity (e.g., *Evi-2B*) that the viruses are affecting. Additionally, of the three tumors that had two independent integrations, one integration occurred within the *Evi-2A* coding region while the second integration occurred in the 3′ end of *Evi-2A*, suggesting that either two hits are better than one or, alternatively, two genes are being affected.

Thus it is possible that these viral integrations are affecting the expression of *Evi-2B*, as well as or instead of *Evi-2A*. It has been previously observed that viral integrations can affect the expression of several closely linked genes (Peters *et al.*, 1989). For example, integration of mouse mammary tumor virus (MMTV) between two fibroblast growth factor-related protooncogenes, *Int-1* and *Hst*, which are located within 17 kb of one another on mouse chromosome 7, can activate the expression of both genes (Peters *et al.*, 1989). The activation of *Int-1* and *Hst* expression by MMTV is thought to be important in the induction of mammary tumors by MMTV. Integration of Moloney-MuLV into the *Pvt-1* locus has been shown to activate the expression of *c-myc* (Tsichlis *et al.*, 1989), which is located 260 kb proximal of *Pvt-1* (Henglein *et al.*, 1989), 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* (Shtivelman *et al.*, 1989). Viral integrations in *Pvt-1* not only activate the ex-

pression of *c-myc* but also affect the expression of *Pvt-1* (Tsichlis *et al.*, 1989). While 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 mouse and rat (Corcoran *et al.*, 1984; Selten *et al.*, 1984; Steffen, 1984; Steffen and Nacar, 1988). Thus, the viral integrations found in the vicinity of the two Evi-2 loci in murine myeloid tumors may contribute to tumorigenesis by affecting the expression of the *Evi-2B* locus, in addition to or instead of affecting the expression of *Evi-2A*. It is also possible that the viral integrations are affecting the expression of the murine homolog of the *NF1* gene and that murine *NF1* contributes to the production of the myeloid tumors.

Although the predicted protein product of *EVI2B* has no significant sequence homology to the products of known oncogenes, it appears to be a glycosylated protein with a single membrane-spanning domain. Examples of transmembrane proteins that when mutated contribute to carcinogenesis include certain growth factor receptors (Ullrich *et al.*, 1984; Bargmann *et al.*, 1986) and cell adhesion molecules (Fearon *et al.*, 1990). Also, overexpression of growth factor receptors of normal structure can contribute to transformation in cell culture (Di Fiore *et al.*, 1987). Studies of the relative abundance of *Evi-2A*, *Evi-2B*, and *NF1* transcripts in murine myeloid tumors containing viral integrations in the Evi-2 region may show specific activation of just one locus, supporting the designation of that locus as a proto-oncogene. Rigorous tests of whether *EVI2B* has oncogenic potential may be possible by using expression vectors to deliver a normal or mutated *EVI2B* gene product to normal or transformed cells in culture, then looking for any effect on the growth properties of the cells. *EVI2B* endogenous to a cultured cell line could be knocked out using an expression vector that produces RNA that is anti-sense to the *EVI2B* mRNA; the consequences to cell growth could then be measured. Analogous studies could be performed in the intact mouse using transgene or homologous recombination technology.

In light of the possible roles of *EVI2A* and *EVI2B* in murine leukemogenesis and the abundance of *EVI2A* (Cawthon *et al.*, 1990b) and *EVI2B* transcripts in peripheral white blood cells, it is worth noting that juvenile chronic myelogenous leukemias (JCMLs) have been reported to occur more often in NF1 patients than in the general population (Bader and Miller, 1978; Mays *et al.*, 1980; Clark and Hutter, 1982). Perhaps certain mutations affect *NF1* and *EVI2A* or *NF1* and *EVI2B* genes to produce the NF1/JCML phenotype; mutations affecting *EVI2A* and/or *EVI2B* without affecting *NF1* may predispose the individual to acquiring particular leukemias. Our screens for mutations in the *EVI2A* (Cawthon *et al.*, 1990b) and *EVI2B* loci in NF1 patients were completely negative. Furthermore, examination of lymphoblast RNA showed no significant differences in the abundance of *EVI2A* or *EVI2B* transcripts between several unrelated NF1 individuals and non-NF1 controls. However, we have not yet thoroughly examined the sequence and expression of the EVI2 region in NF1/JCML patients or in leukemia patients.

We are not aware of another example of three active genes (*EVI2A*, *EVI2B*, and *OMGP;* Viskochil *et al.*, 1990b) located on the anti-sense strand of an intron of another gene (*NF1*). Some of the possible consequences of this arrangement have been discussed (Cawthon *et al.*, 1990a; Wallace *et al.*, 1990a). One interesting situation that may produce *NF1* in a small fraction of patients could involve a mutation that alters the splicing and transport to the cytoplasm of the *EVI2B* transcript. If the levels of this transcript in the nucleus become elevated, it might hybridize to the complementary segment of the *NF1* transcript, thereby promoting rapid degradation or interfering with the splicing of that intron. Such a mutation could be dominant with respect to *NF1*, yet recessive with respect to the normal function of *EVI2B*, since one copy of the normal *EVI2B* allele (provided by the chromosome 17 homolog) may be adequate to provide that function. Thus, in addition to yielding a wealth of information about the regulation of the *NF1* locus and mechanisms of mutation in *NF1*, the study of these four genes may generate principles of gene organization, regulation, and interaction that can be applied to other yet to be discovered systems of genes embedded within other genes.

ACKNOWLEDGMENTS

We thank V. M. Riccardi and D. Fults for providing NF1 tissues; M. Robertson for the fluorescent sequencing of *EVI2B* cDNAs on the Applied Biosystems Inc. Model 370A DNA sequencer and K. Murphy for assistance with the manual sequencing; and S. Nothwehr for analyzing *EVI2B* with the SIGSEQ programs. R. Foltz prepared the figures. This study was supported in part by the National Cancer Institute, Department of Health and Human Services, under Contract N01-CO-74101 with ABL and in part by NIH Grant NS23410 to F.S.C. L.B.A. is a recipient of a fellowship from the University of Copenhagen, Denmark; D.V. and M.R.W. are supported by the National Neurofibromatosis Foundation; R.M.C. is an Associate, P.O'C. a Senior Associate, F.S.C. an Associate Investigator, and R.W. an Investigator in the Howard Hughes Medical Institute.

REFERENCES

1. ANDERSEN, L. B., WALLACE, M. R., MARCHUK, D. A., CAWTHON, R. M., ODEH, H. M., LETCHER, R., WHITE, R., AND COLLINS, F. S. (1990). A polymorphic cDNA probe on

chromosome 17q11.2, located within the *NF1* gene (D17S376). *Nucleic Acids Res.* **19,** in press.

2. BADER, J. L., AND MILLER, R. W. (1978). Neurofibromatosis and childhood leukemia. *J. Pediatr.* **92:** 925–929.
3. BARGMANN, C. I., HUNG, M.-C., AND WEINBERG, R. A. (1986). Multiple independent activations of the *neu* oncogene by a point mutation altering the transmembrane domain of p185. *Cell* **45:** 649–657.
4. BARKER, D., WRIGHT, E., NGUYEN, K., CANNON, L., FAIN, P., GOLDGAR, D., BISHOP, D. T., CAREY, J., BATY, B., KIVLIN, J., WILLARD, H., WAYE, J. S., GREIG, G., LEINWAND, L., NAKAMURA, Y., O'CONNELL, P., LEPPERT, M., LALOUEL, J.-M., WHITE, R., AND SKOLNICK, M. (1987). Gene for von Recklinghausen neurofibromatosis is in the pericentromeric region of chromosome 17. *Science* **236:** 1100–1102.
5. BEDIGIAN, H. G., TAYLOR, B. A., AND MEIER, H. (1981). Expression of murine leukemia viruses in the highly lymphomatous BXH-2 recombinant inbred mouse strain. *J. Virol.* **39:** 632–640.
6. BENTON, W. D., AND DAVIS, R. W. (1977). Screening lambda-gt recombinant clones by hybridization to single plaques in situ. *Science* **196:** 180–182.
7. BERGER, S. L., AND BIRKENMEIER, C. S. (1979). Inhibition of intractable nucleases with ribonucleoside-vanadyl complexes: Isolation of messenger RNA from resting lymphocytes. *Biochemistry* **18:** 5143–5149.
8. BONTHRON, D. T., MARKHAM, A. F., GINSBURG, D., AND ORKIN, S. H. (1985). Identification of a point mutation in the adenosine deaminase gene responsible for immunodeficiency. *J. Clin. Invest.* **76:** 894–897.
9. BRAWERMAN, G. (1989). mRNA decay: Finding the right targets. *Cell* **57:** 9–10.
10. BUCHBERG, A. M., BEDIGAN, H. G., TAYLOR, B. A., BROWNELL, E., IHLE, J. N., NAGATA, S., JENKINS, N. A., AND COPELAND, N. (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.
11. BUCHBERG, A. M., BROWNELL, E., NAGATA, S., JENKINS, N. A., AND COPELAND, N. G. (1989). A comprehensive genetic map of murine chromosome 11 reveals extensive linkage conservation between mouse and human. *Genetics* **122:** 153–161.
12. BUCHBERG, A. M., BEDIGIAN, H. G., JENKINS, N. A., AND COPELAND, N. G. (1990). *Evi-2,* A common integration site involved in murine myeloid leukemogenesis. *Mol. Cell. Biol.* **10:** 4658–4666.
13. CARLE, G. F., AND OLSON, M. V. (1984). Separation of chromosomal DNA molecules from yeast by orthogonal field alteration gel electrophoresis. *Nucleic Acids Res.* **12:** 5647–5664.
14. CAWTHON, R. M., WEISS, R., XU, G., VISKOCHIL, D., CULVER, M., STEVENS, J., ROBERTSON, M., DUNN, D., GESTELAND, R., O'CONNELL, P., AND WHITE, R. (1990a). A major segment of the neurofibromatosis type 1 gene: cDNA sequence, genomic structure, and point mutations. *Cell* **62:** 193–201.
15. CAWTHON, R., O'CONNELL, P., BUCHBERG, A., VISKOCHIL, D., WEISS, R., CULVER, M., STEVENS, J., JENKINS, N., COPELAND, N., AND WHITE, R. (1990b). Identification and characterization of transcripts from the neurofibromatosis 1 region: The sequence and genomic structure of *EVI2* and mapping of other transcripts. *Genomics* **7:** 555–565.
16. CHOMCZYNSKI, P., AND SACCHI, N. (1987). Single-step method of RNA isolation by acid guanidinium thiocyanate–phenol–chloroform extraction. *Anal. Biochem.* **162:** 156–159.
17. CHU, G., VOLLRATH, D., AND DAVIS, R. (1986). Separation of large DNA molecules by contour-clamped homogeneous electric fields. *Science* **234:** 1582–1585.
18. CLARK, R. D., AND HUTTER, J. J., JR. (1982). Familial neurofibromatosis and juvenile chronic myelogenous leukemia. *Hum. Genet.* **60:** 230–232.
19. COLLINS, F. S., DRUMM, M. L., COLE, J. L., LOCKWOOD, W. K., VANDE WOUDE, G. F., AND IANNUZZI, M. C. (1987). Construction of a general human chromosome jumping library, with application to cystic fibrosis. *Science* **235:** 1046–1049.
20. COLLINS, F. S., O'CONNELL, P., PONDER, B. A., AND SEIZINGER, B. R. (1989). Progress towards identifying the neurofibromatosis (NF1) gene. *Trends Genet.* **5:** 217–225.
21. COPELAND, N. G., BUCHBERG, A. M., GILBERT, D. J., AND JENKINS, N. A. (1989). Recombinant inbred mouse strains: Models for studying the molecular genetic basis of myeloid tumorigenesis. *In* "Mechanism in Myeloid Tumorigenesis" (G. L. C. Shen-Ong, M. Potter, and N. G. Copeland, Eds.), Current Topics in Microbiology and Immunology, Vol. 149, pp. 45–57, Springer-Verlag, Berlin/Heidelberg.
22. CORCORAN, L. M., ADAMS, J. M., DUNN, A. R., AND CORY, S. (1984). Murine T lymphomas in which the cellular *myc* oncogene has been activated by retroviral insertion. *Cell* **37:** 113–122.
23. CROWE, F. W., SCHULL, W. T., AND NEEL, J. F. (1956). "A Clinical, Pathological, and Genetic Study of Multiple Neurofibromatosis," Thomas, Springfield, IL.
24. DEVEREUX, J., HAEBERLI, P., AND SMITHIES, O. (1984). A comprehensive set of sequence analysis programs for the VAX. *Nucleic Acids Res.* **12:** 387–395.
25. DI FIORE, P. P., PIERCE, J. H., FLEMING, T. P., HAZAN, R., ULLRICH, A., KING, C. R., SCHLESSINGER, J., AND AARONSON, S. A. (1987). Overexpression of the human EGF receptor confers an EGF-dependent transformed phenotype to NIH 3T3 cells. *Cell* **51:** 1063–1070.
26. FEARON, E. R., CHO, K. R., NIGRO, J. M., KERN, S. E., SIMONS, J. W., RUPPERT, J. M., HAMILTON, S. R., PREISINGER, A. C., THOMAS, G., KINZLER, K. W., AND VOGELSTEIN, B. (1990). Identification of a chromosome 18q gene which is altered in colorectal cancers. *Science* **247:** 49–56.
27. FEINBERG, A., AND VOGELSTEIN, B. (1984). A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. Addendum. *Anal. Biochem.* **137:** 266–267.
28. FOLZ, R. J., AND GORDON, J. I. (1987). Computer-assisted predictions of signal peptidase processing sites. *Biochem. Biophys. Res. Commun.* **146:** 870–877.
29. FOUNTAIN, J. W., WALLACE, M. R., BRUCE, M. S., SEIZINGER, B. R., MENON, A. G., GUSELLA, J. F., MICHELS, V. V., SCHMIDT, M. A., DEWALD, G. W., AND COLLINS, F. S. (1989). Physical mapping of a translocation breakpoint in neurofibromatosis. *Science* **244:** 1085–1087.
30. FROHMAN, M. A., DUSH, M. K., AND MARTIN, G. R. (1988). Rapid production of full-length cDNAs from rare transcripts: Amplification using a single gene-specific oligonucleotide primer. *Proc. Natl. Acad. Sci. USA* **85:** 8998–9002.
31. GARDINER, K., LAAS, W., AND PATERSON, D. (1986). Fractionation of large mammalian DNA restriction fragments using vertical pulsed-field gradient electrophoresis. *Somatic Cell Mol. Genet.* **12:** 185–195.
32. GIERASCH, L. M. (1989). Signal sequences. *Biochemistry* **28:** 923–930.
33. GOLDGAR, D. E., GREEN, P., PARRY, D. M., AND MULVIHILL,

J. J. (1989). Multipoint linkage analysis in neurofibromatosis type 1: An international collaboration. *Amer. J. Hum. Genet.* **44:** 6–12.

34. GREENBERG, M. E., HERMANOWSKI, A. L., AND ZIFF, E. B. (1986). Effect of protein synthesis inhibitors on growth factor activation of c-fos, c-myc, and actin gene transcription. *Mol. Cell. Biol.* **6:** 1050–1057.
35. GYLLENSTEN, U. (1989). Direct sequencing of *in vitro* amplified DNA. *In* "PCR Technology" (H. Erlich, Ed.), pp. 45–60, Stockton Press, New York.
36. HALUSKA, F. G., TSUJIMOTO, Y., AND CROCE, C. M. (1987a). The t(8;14) chromosome translocation of the Burkitt lymphoma cell line Daudi occurred during immunoglobulin gene rearrangement and involved the heavy chain diversity region. *Proc. Natl. Acad. Sci. USA* **84:** 6835–6839.
37. HALUSKA, F. G., TSUJIMOTO, Y., AND CROCE, C. M. (1987b). Oncogene activation by chromosome translocation in human malignancy. *Annu. Rev. Genet.* **21:** 321–345.
38. HENGLEIN, B., SYNOVZIK, H., GROITL, P., BORNKAMM, G., HARTL, P., AND LIPP, M. (1989). Three breakpoints of variant t(2;8) translocations in Burkitt's lymphoma cell fall within a region 140 kilobases distant from c-*myc*. *Mol. Cell Biol.* **9:** 2105–2113.
39. HOLMES, D. S., AND QUIGLEY, M. (1981). A rapid boiling method for the preparation of bacterial plasmids. *Anal. Biochem.* **114:** 193.
40. KOZAK, M. (1984). Compilation and analysis of sequences upstream from the translational start site in eukaryotic mRNAs. *Nucleic Acids Res.* **12:** 857–872.
41. KYTE, J., AND DOOLITTLE, R. F. (1982). A simple method for displaying the hydropathic character of a protein. *J. Mol. Biol.* **157:** 105–132.
42. LANDSCHULZ, W. H., JOHNSON, P. F., AND MCKNIGHT, S. L. (1988). The leucine zipper: A hypothetical structure common to a new class of DNA binding proteins. *Science* **240:** 1759–1764.
43. LEDBETTER, D. H., RICH, D. C., O'CONNELL, P., LEPPERT, M., AND CAREY, J. C. (1989). Precise localization of NF1 to 17q11.2 by balanced translocation. *Amer. J. Hum. Genet.* **44:** 20–24.
44. LOCKE, J., KOTARSKI, M. A., AND TARTOF, K. D. (1988). Dosage-dependent modifiers of position effect variegation in Drosophila and a mass action model that explains their effect. *Genetics* **120:** 181–198.
45. LOH, E. Y., ELLIOTT, J. F., CWIRLA, S., LANIER, L. L., AND DAVIS, M. M. (1989). Polymerase chain reaction with single-sided specificity: Analysis of T cell receptor delta chain. *Science* **243:** 217–220.
46. MALTER, J. S. (1989). Identification of an AUUUA-specific messenger RNA binding protein. *Science* **246:** 664–666.
47. MANIATIS, T., FRITSCH, E. F., AND SAMBROOK, J. (1982). "Molecular Cloning: A Laboratory Manual," Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
48. MAYS, J. A., NEERHOUT, R. C., BAGBY, G. C., AND KOLER, R. D. (1980). Juvenile chronic granulocytic leukemia: Emphasis on cutaneous manifestations and underlying neurofibromatosis. *Amer. J. Dis. Child.* **134:** 654–658.
49. MENON, A. G., LEDBETTER, D. H., RICH, D. C., SEIZINGER, B. R., ROULEAU, G. A., MICHAELS, V. F., SCHMIDT, M. S., DEWALD, G., DELLATORRE, C. M., HAINES, J. L., AND GUSELLA, J. F. (1989). Characterization of a translocation within the von Recklinghausen neurofibromatosis region of chromosome 17. *Genomics* **5:** 245–249.
50. MIKOL, D. D., GULCHER, J. R., AND STEFANSSON, K. (1990). The oligodendrocyte-myelin glycoprotein belongs to a distinct family of proteins and contains the HNK-1 carbohydrate. *J. Cell Biol.* **110:** 471–480.
51. MORRIS, D. W., BARRY, P. A., BRADSHAW, H. D., JR., AND CARDIFF, R. D. (1990). Insertion mutation of the int-1 and int-2 loci by mouse mammary tumor virus in premalignant and malignant neoplasms from the GR strain. *J. Virol.* **64:** 1794–1802.
52. O'CONNELL, P., LEACH, R., CAWTHON, R., CULVER, M., STEVENS, J., VISKOCHIL, D., FOURNIER, R. E. K., RICH, D., LEDBETTER, D., AND WHITE, R. (1989). Two von Recklinghausen neurofibromatosis translocations map within a 600-kb segment of 17q11.2. *Science* **244:** 1087–1088.
53. O'CONNELL, P., VISKOCHIL, D., BUCHBERG, A. M., FOUNTAIN, J., CAWTHON, R., CULVER, M., STEVENS, J., RICH, D. C., LEDBETTER, D. H., WALLACE, M., CAREY, J. C., JENKINS, N. A., COPELAND, N. G., COLLINS, F. S., AND WHITE, R. (1990). The human homologue of murine *Evi-2* lies between two translocation breakpoints associated with von Recklinghausen neurofibromatosis. *Genomics* **7:** 547–554.
54. ORITA, M., SUZUKI, Y., SEKIYA, T., AND HAYASHI, K. (1989). Rapid and sensitive detection of point mutations and DNA polymorphisms using the polymerase chain reaction. *Genomics* **5:** 874–879.
55. PETERS, G., S. BROOKES, R. SMITH, M. PLACZEK, AND DICKSON, C. (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.
56. REED, K. C., AND MANN, D. A. (1985). Rapid transfer of DNA from agarose gels to nylon membranes. *Nucleic Acids Res.* **13:** 7207–7221.
57. RICCARDI, V. M., AND EICHNER, J. E. (1986). "Neurofibromatosis: Phenotype, Natural History, and Pathogenesis," Johns Hopkins Univ. Press, Baltimore, MD.
58. SAMBROOK, J., FRITSCH, E. F., AND MANIATIS, T. (1989). "Molecular Cloning: A Laboratory Manual," 2nd ed., Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
59. SANGER, F., NICKLEN, S., AND COULSON, A. R. (1977). DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA* **74:** 5463–5467.
60. SCHMIDT, M. A., MICHELS, V. V., AND DEWALD, G. W. (1987). Cases of neuro-fibromatosis with rearrangements of chromosome 17 involving band 17q11.2. *Amer. J. Med. Genet.* **28:** 771–777.
61. SCHWARTZ, D. C., AND CANTOR, C. R. (1984). Separation of yeast chromosome-sized DNAs by pulsed-field gel electrophoresis. *Cell* **37:** 67–75.
62. SEIZINGER, B., ROULEAU, G., OZELIUS, L., LANE, A., FARYNIARZ, A., CHAO, M., HUSON, S., KORF, B., PARRY, D., PERICAK-VANCE, M., COLLINS, F., HOBBS, W., FALCONE, B., IANNAZZI, J., ROY, J., ST. GEORGE-HYSLOP, P., TANZI, R., BOTHWELL, M., UPADHYAYA, M., HARPER, P., GOLDSTEIN, A., HOOVER, D., BADER, J., SPENCE, M., MULVIHILL, J., AYLSWORTH, A., VANCE, J., ROSSENWASSER, G., GASKELL, P., ROSES, A., MARTUZA, R., BREAKEFIELD, X., AND GUSELLA, J. (1987). Genetic linkage of von Recklinghausen neurofibromatosis to the nerve growth factor receptor gene. *Cell* **49:** 589–594.
63. SELDEN, R. F. (1989). Analysis of RNA by northern hybridization. *In* "Current Protocols in Molecular Biology" (F. M. Ausubel, R. Brent, R. E. Kingston, D. D. Moore, J. G. Seid-

man, J. A. Smith, and K. Struhl, Eds.), pp. 4.9.1–4.9.8, Wiley, New York.

64. SELTEN, G., CUYPERS, H. Y., ZIJLSTRA, M., MELIEF, C., AND BERNS, A. (1984). Involvement of c-myc in MuLV-induced T cell lymphomas in mice: Frequency and mechanisms of activation. *EMBO J.* **3:** 3215–3222.

65. SHAPIRO, M., AND SENAPATHY, P. (1987). RNA splice junctions of different classes of eukaryotes: Sequence statistics and functional implications in gene expression. *Nucleic Acids Res.* **15:** 7155.

66. SHAW, G., AND KAMEN, R. (1986). A conserved AU sequence from the 3′ untranslated region of GM-CSF mRNA mediates selective mRNA degradation. *Cell* **46:** 659–667.

67. SHTIVELMAN, E., HENGLEIN, B., GROITL, P., LIPP, M., AND BISHOP, J. M. (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.

68. SOUTHERN, E. M. (1975). Detection of specific sequences among DNA fragments separated by gel electrophoresis. *J. Mol. Biol.* **98:** 503–517.

69. STEFFEN, D. (1984). Proviruses are adjacent to c-myc in some murine leukemia virus-induced lymphomas. *Proc. Natl. Acad. Sci. USA* **81:** 2097–2101.

70. STEFFEN, D. L., AND NACAR, E. Q. (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.

71. STUMPF, D. A., ALKSNE, J. F., ANNEGERS, J. F., BROWN, S. S., CONNEALLY, P. M., HOUSMAN, D., LEPPERT, M., MILLER, J. P., MOSS, M. L., PILEGGI, A. J., RAPIN, I., STROHMAN, R. C., SWANSON, L. W., AND ZIMMERMAN, A. (1987). Neurofibromatosis, "NIH Consensus Development Conference Statement," Vol. 6, No. 12.

72. TSICHLIS, P. N., SHEPHERD, B. M., AND BEAR, S. E. (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.

73. ULLRICH, A., *et al.* (1984). Human epidermal growth factor receptor cDNA sequence and aberrant expression of the amplified gene in A431 epidermoid carcinoma cells. *Nature (London)* **309:** 418–425.

74. VISKOCHIL, D., BUCHBERG, A. M., XU, G., CAWTHON, R. M., STEVENS, J., WOLFF, R. K., CULVER, M., CAREY, J. C., COPELAND, N. G., JENKINS, N. A., WHITE, R., AND O'CONNELL, P. (1990a). Deletions and a translocation interrupt a cloned gene at the neurofibromatosis type 1 locus. *Cell* **62:** 187–192.

75. VISKOCHIL, D., CAWTHON, R., O'CONNELL, P., XU, G., STEVENS, J., CULVER, M., CAREY, J., AND WHITE, R. (1990b). The gene encoding the oligodendrocyte-myelin glycoprotein is embedded within the neurofibromatosis type 1 gene. *Mol. Cell Biol.* **11,** in press.

76. VON HEIJNE, G. (1985). Signal sequences: The limits of variation. *J. Mol. Biol.* **184:** 99–105.

77. VON HEIJNE, G. (1986). A new method for predicting signal sequence cleavage sites. *Nucleic Acids Res.* **14:** 4683–4690.

78. WALLACE, M. R., MARCHUK, D. A., ANDERSEN, L. B., LETCHER, R., ODEH, H. M., SAULINO, A. M., FOUNTAIN, J. W., BRERETON, A., NICHOLSON, J., MITCHELL, A. L., BROWNSTEIN, B. H., AND COLLINS, F. S. (1990a). Type 1 neurofibromatosis gene: Identification of a large transcript disrupted in three NF1 patients. *Science* **249:** 181–186.

79. WALLACE, M. R., ANDERSEN, L. B., FOUNTAIN, J. W., ODEH, H. M., VISKOCHIL, D., MARCHUK, D. A., O'CONNELL, P., WHITE, R., AND COLLINS, F. S. (1990b). A chromosome jump crosses a translocation breakpoint in the von Recklinghausen neurofibromatosis region. *Genes, Chromosomes, and Cancer,* in press.

80. WILSON, T., AND TREISMAN, R. (1988). Removal of poly(A) and consequent degradation of c-fos mRNA facilitated by 3′ AU-rich sequences. *Nature (London)* **336:** 396–399.

81. XU, G., O'CONNELL, P., VISKOCHIL, D., CAWTHON, R., ROBERTSON, M., CULVER, M., DUNN, D., STEVENS, J., GESTELAND, R., WHITE, R., AND WEISS, R. (1990). The neurofibromatosis type 1 gene encodes a protein related to GAP. *Cell* **62:** 599–608.