REFERENCES AND NOTES

1. M. J. Rogers *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **82**, 1160 (1985).
2. M. C. Roberts, L. A. Koutsky, K. K. Holmes, D. J. LeBlanc, G. E. Kenny, *Antimicrob. Agents Chemother.* **28**, 141 (1985).
3. M. C. Roberts and G. E. Kenny, *ibid.* **29**, 350 (1986).
4. D. B. Clewell and C. Gawron-Burke, *Annu. Rev. Microbiol.* **40**, 635 (1986).
5. A. E. Franke and D. B. Clewell, *J. Bacteriol.* **145**, 494 (1981).
6. C. Gawron-Burke and D. B. Clewell, *Nature (London)* **300**, 281 (1982).
7. V. Burdett, J. Inamine, S. Rajagopalan, *J. Bacteriol.* **149**, 995 (1982).
8. C. Gawron-Burke and D. B. Clewell, *ibid.* **159**, 214 (1984).
9. T. L. Sladek and J. Maniloff, *ibid.* **155**, 734 (1983).
10. A. Liss and J. Maniloff, *Science* **173**, 725 (1971).
11. Mycoplasma medium consisted of 2.25% mycoplasma broth base without crystal violet (Gibco) supplemented with 0.5% CVA Enrichment (Gibco), 20% gamma globulin–free, heat-inactivated horse serum (Gibco), 0.5% dextrose, and 0.02% DNA (degraded free acid type IV) (Sigma). For plates, 1.5% agar (Gibco) was added to the above medium.
12. In subsequent experiments, we found that incubating for 2 hours before the addition of tetracycline was unnecessary as long as the colonies were added to prewarmed medium. Evidently the cells have a difficult time recovering from cold shock in the presence of tetracycline.
13. K. Dybvig, J. Alderette, G. H. Cassell, in preparation.
14. T. L. Sladek, J. A. Nowak, J. Maniloff, *J. Bacteriol.* **165**, 219 (1986).
15. A. Liss and J. Maniloff, *Virology* **55**, 118 (1973).
16. F. C. Minion *et al.*, *Infect. Immun.* **44**, 394 (1984).
17. J. A. Nowak and J. Maniloff, *J. Bacteriol.* **151**, 1420 (1982).
18. K. Haberer, G. Klotz, J. Maniloff, A. K. Kleinschmidt, *J. Virol.* **32**, 269 (1979).
19. K. Dybvig, A. Liss, J. Alderete, R. M. Cole, G. H. Cassell, *Isr. J. Med. Sci.*, in press.
20. J. M. Ranhand, W. O. Mitchell, T. J. Popkin, R. M. Cole, *J. Bacteriol.* **143**, 1194 (1980).
21. T. Maniatis, E. F. Fritsch, J. Sambrook, Eds., *Molecular Cloning* (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1982).
22. K. Dybvig and J. Maniloff, *J. Gen. Virol.* **64**, 1781 (1983).
23. K. Dybvig, T. L. Sladek, J. Maniloff, *J. Virol.* **59**, 584 (1986).
24. We thank D. Clewell for providing us with plasmid pAM120. We thank J. Alderete and L. Barrett for their laboratory assistance. Supported by NIH grants HL-19741 and RR-00959 and training grant award AI-07150.

30 October 1986; accepted 4 February 1987

Human Retinoblastoma Susceptibility Gene: Cloning, Identification, and Sequence

WEN-HWA LEE,* ROBERT BOOKSTEIN, FRANK HONG, LIH-JIUAN YOUNG, JIN-YUH SHEW, EVA Y.-H. P. LEE

Recent evidence indicates the existence of a genetic locus in chromosome region 13q14 that confers susceptibility to retinoblastoma, a cancer of the eye in children. A gene encoding a messenger RNA (mRNA) of 4.6 kilobases (kb), located in the proximity of esterase D, was identified as the retinoblastoma susceptibility (RB) gene on the basis of chromosomal location, homozygous deletion, and tumor-specific alterations in expression. Transcription of this gene was abnormal in six of six retinoblastomas examined: in two tumors, RB mRNA was not detectable, while four others expressed variable quantities of RB mRNA with decreased molecular size of about 4.0 kb. In contrast, full-length RB mRNA was present in human fetal retina and placenta, and in other tumors such as neuroblastoma and medulloblastoma. DNA from retinoblastoma cells had a homozygous gene deletion in one case and hemizygous deletion in another case, while the remainder were not grossly different from normal human control DNA. The gene contains at least 12 exons distributed in a region of over 100 kb. Sequence analysis of complementary DNA clones yielded a single long open reading frame that could encode a hypothetical protein of 816 amino acids. A computer-assisted search of a protein sequence database revealed no closely related proteins. Features of the predicted amino acid sequence include potential metal-binding domains similar to those found in nucleic acid–binding proteins. These results provide a framework for further study of recessive genetic mechanisms in human cancers.

RETINOBLASTOMA, THE MOST COMMON intraocular cancer of childhood, occurs in both hereditary and nonhereditary forms. The former is characterized by early age of onset and multiple tumor foci compared to the nonhereditary type, which occurs later with a single, unilateral tumor (*1*). Susceptibility to hereditary retinoblastoma is transmissible to offspring as an autosomal-dominant trait with 90% penetrance, and the tumor is a prototypic model for the study of genetic determination in cancer (*2*).

The genetic locus determining retinoblastoma susceptibility was assigned to band q14 of chromosome 13, along with the gene for the polymorphic marker enzyme esterase D, by examination of cytogenetic deletions (*3*). The close linkage of these loci was confirmed by studies of retinoblastoma pedigrees (*4*). The retinoblastoma susceptibility locus was further implicated in nonhereditary retinoblastoma by observations of frequent chromosome 13 abnormalities in tumor karyotypes and reduced esterase D activity in tumors (*5*). Benedict *et al.* proposed that inactivation of both alleles of a gene (RB) located in region 13q14 resulted in retinoblastoma, based in part on a case of hereditary retinoblastoma in which both RB alleles were inferred to be absent (*6*). Recently an assumption made in this case has been disproved (*7*), namely that the absence of esterase D activity implied loss of both esterase D and RB genes. Nonetheless, Cavenee *et al.* (*8*) found that chromosome 13 markers that were heterozygous in somatic cells often became homozygous or hemizygous in retinoblastoma tumors, and Dryja *et al.* (*9*) demonstrated homozygous 13q14 deletions at the molecular level in 2 out of 37 retinoblastoma tumors. These experiments provided evidence that the proposed RB gene indeed functions in a "recessive" manner at the cellular level (*1, 10*), in distinction to the "dominant" activities of classical oncogenes (*10, 11*) as measured, for example, by transfection assays.

Given the importance of RB as a novel type of cancer gene, we undertook to isolate this gene by "chromosome walking" from other chromosome 13 markers. Since nothing was known of RB gene expression or products, candidate genes would be identified on the basis of appropriate chromosomal location and presumed "recessive" behavior: that is, an intact RB gene should be expressed in retinal tissue but not in retinoblastomas. Our initial starting point was the gene encoding esterase D (*12*), which is linked to the retinoblastoma susceptibility locus in band 13q14.11 (*4, 13*) within an estimated 1500-kilobase range (*14*). The esterase D complementary DNA (cDNA) clone EL-22 was used as a probe to isolate its genomic DNA clones. Distal DNA segments of these genomic clones were used in turn to isolate additional genomic clones. At 20-kb intervals in walking regions, unique sequences were identified that were used as probes to isolate cDNA clones from fetal retina and placenta libraries. By alternately screening genomic and cDNA libraries, we have established a bidirectional chromosome walk covering 120 kb (Fig. 1). Two cDNA clones, SD-1 and SD-2, were isolated by means of probes 5′ to the esterase D gene. Chromosome walking 3′ to the esterase D gene was hampered by a 20-kb region

Department of Pathology, Experimental Pathology Program, School of Medicine, University of California at San Diego, La Jolla, CA 92093.

*To whom all correspondence should be addressed.

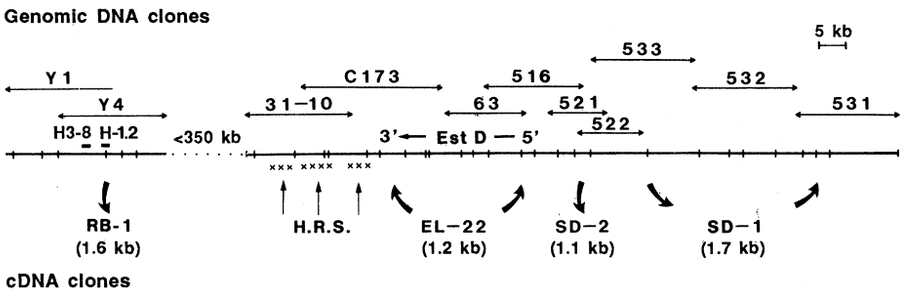

Fig. 1. Summary of chromosome walking and isolation of cDNA clones from this region. Starting points were EL-22 (the esterase D cDNA clone) (*14*), located at 13q14.11 (*13*), and H3-8, a 13q14-specific DNA probe (*15*). These were used to screen Charon 30, Charon 4A, EMBL-3, or cosmid genomic libraries (*17*) from DNA of human peripheral blood lymphocytes or Y79, a retinoblastoma cell line (*36*). Unique sequence subfragments of genomic clones were in turn used to screen λgt10 or gt11 cDNA libraries made from fetal retinal or placental mRNA, yielding SD-1, SD-2 and RB-1. Hybridization conditions were: 6× SSC/5× Denhardt's solution/0.1% SDS at 65°C, and washing conditions were: once in 2× SSC/0.1% SDS at room temperature for 15 minutes and twice in 0.2× SSC/0.1% SDS at 65°C for 30 minutes. DNA fragments containing repetitive sequences were determined by hybridizing with ^{32}P-labeled total human genomic DNA as probe. Those fragments failing to hybridize were considered unique sequence probes suitable for use in screening. Selected fragments were purified and labeled with ^{32}P-(α)-dATP by nick translation or random oligonucleotide primer methods (*37*). The consensus map resulting from digestion with several restriction enzymes has the Eco RI sites shown. A 20-kb region 3′ to the esterase D gene contained many highly repetitive sequences (H.R.S.), since most restriction fragments from this region showed positive hybridization with either human Alu-1 probe or total genomic DNA. The maximal distance between H3-8 and EL-22 was determined by demonstrating a large Sal I–digested DNA fragment that hybridized with both probes (*16*), and the intergene separation was adjusted for distances already covered by chromosome walking. The orientation of esterase D with respect to RB gene location (shown in the diagram with the 3′ end of EL-22 toward RB) should be regarded as tentative although it is supported by indirect evidence (*16*).

containing highly repetitive sequences. We obtained an additional probe, H3-8, which was isolated by random selection from a chromosome 13–specific genomic library (*15*). Preliminary results with orthogonal-field gel electrophoresis indicated that H3-8 was less than 500 kb from the esterase D gene (*16*). Starting from the location of H3-8, a second bidirectional chromosome walk was performed in a manner similar to that above, extending over 30 kb. A unique DNA fragment, H-1.2, identified two overlapping cDNA clones of 1.6 kb (RB-1) and 0.9 kb (RB-2) in human cDNA libraries (Fig. 1).

Candidate RB genes were used as probes in RNA blotting analysis to detect relevant messenger RNA (mRNA) transcripts (Table 1). Polyadenylated RNA was prepared (*17*) from human fetal retinas (obtained by dissection of about 25 first- or second-trimester fetal eyes) and from portions of normal human placentas. Primary retinoblastoma tumor samples large enough to yield sufficient mRNA for analysis are not usually available. Polyadenylated RNA was isolated from cultured cells of six retinoblastomas, three neuroblastomas and one medulloblastoma; mRNA was also obtained from a primary medulloblastoma specimen. Esterase D transcripts (1.4 kb) were detected in all tumor and tissue samples (Fig. 2), consistent with the known "constitutive" expression of esterase D (*12*). Esterase D hybridization was subsequently used as a positive control.

Neither SD-1 nor SD-2 seemed promising as candidate RB genes, because transcripts hybridizing to these clones were not detected in retina and placenta mRNA samples nor from any retinoblastomas (Table 1). Since SD-1 and SD-2 were obtained from placental (*18*) and fetal retinal cDNA libraries, respectively, their expression in cognate mRNA was expected. While the lack of detection may be due to low copy number of transcripts, both clones require further characterization to explain these findings.

RB-1 hybridized with a 4.6-kb mRNA transcript in fetal retina and placenta (Fig. 2). In three of six retinoblastoma samples (Fig. 2A, lanes 1, 2, and 5), abnormal mRNA transcripts measuring approximately 4.0 kb were detected by RB-1. In two retinoblastomas (lanes 3 and 6), mRNA transcripts were not observed. RNA from one sample (lane 4) showed a faint band of about 4.5 kb, slightly smaller in size than

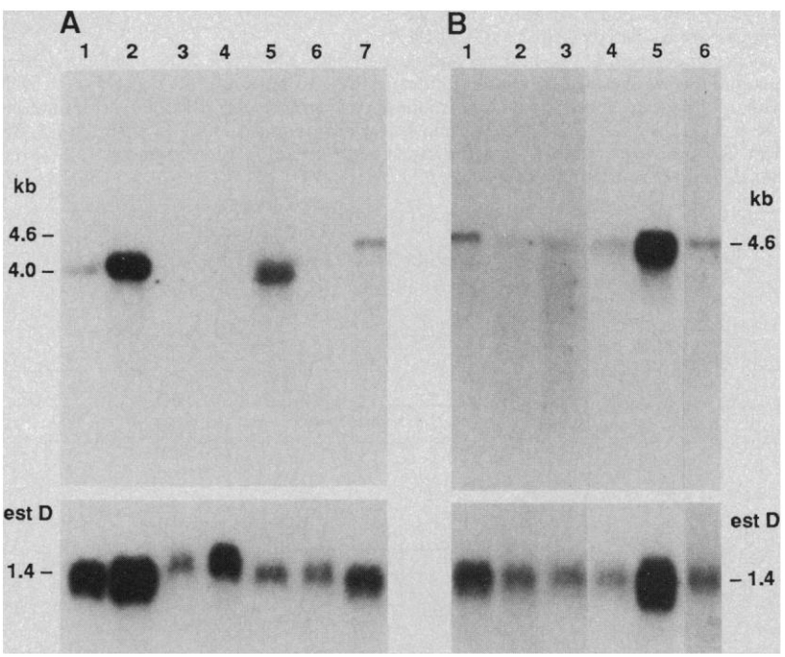

Fig. 2. RNA blot analysis of RB gene transcripts in tumors and normal tissues. Polyadenylated RNA (2 to 5 μg) prepared from cultured retinoblastoma cells (lanes A1 to A6), fetal retina (lane A7), neuroblastoma cell lines (lanes B1 to B3), a medulloblastoma cell line and a fresh tumor (lanes B4 and 5), and human placenta (lane B6) were separated by electrophoresis in 1% formaldehyde–agarose gels and transferred to nitrocellulose filters with 20× SSC. Filters were hybridized with ^{32}P-labeled RB-1 DNA (**top panel**) in 50% formamide/6× SSC/5× Denhardt's solution/0.1% SDS/denatured salmon sperm DNA (100 μg/ml)/10 m*M* phosphate buffer (*p*H 7.0) at 42°C for 24 hours. Washing followed in 2× SSC/0.1% SDS at room temperature for 20 minutes and twice in 0.1× SSC/0.1% SDS at 65°C for 30 minutes. Autoradiography was done with Kodak XAR-5 film at −70°C for 3 days with an intensifying screen. Filters were then rehybridized with ^{32}P-labeled EL-22 DNA as described above, and exposed for 3 days (**bottom panel**). The apparent slight variation in mobility of esterase D mRNA transcripts reflects the required overloading.

that of control mRNA (retina; Fig. 2A, lane 7). Results from retinoblastoma samples 3, 4, and 6 were duplicated by overloading lanes and prolonged autoradiographic exposure, confirming that mRNA transcripts were indeed absent in tumors 3 and 6; an additional faint band of 3.8 kb was seen in lane 4. Esterase D mRNA was detected in all cases. Three neuroblastomas and two medulloblastomas displayed identical transcripts of 4.6 kb, equivalent to that in normal tissues (Fig. 2B). Altered RB-1–related gene expression (manifested as a shortened transcript or absence of transcript) was thus found in six of six retinoblastomas, but not in two normal tissues or two other related human tumors of neurectodermal origin, strongly suggesting that RB-1 represented part of the putative RB gene. We sought a complete cDNA clone by rescreening the two cDNA libraries with RB-1 as probe. Clone RB-5, with a 3.5-kb insert, was then isolated; it gave identical results as RB-1 in mRNA hybridization. Restriction enzyme analysis suggested that RB-5 and RB-1 overlapped in a 0.4-kb region, and together defined a DNA segment of about 4.6 kb, a size close to that of the normal mRNA transcript. For ease of analysis, the following probes were generated: RB-1 was cleaved by Eco RI into 0.9-kb (RB0.9) and 0.7-kb (RB0.7) fragments. RB-5 was cleaved by Bgl II into three fragments called RB710 (0.7 kb, Eco RI–Bgl II), RB1.8 (1.8 kb, Bgl II–Bgl II), and RB1.0 (1.0 kb, Bgl II–Eco RI).

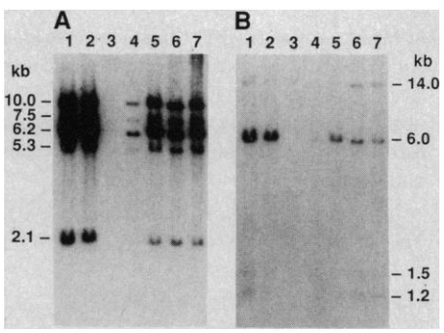

Fig. 3. DNA blot analysis of genomic DNA fragments in retinoblastoma tumors. About 10 μg of DNA extracted from cultured cells derived from six retinoblastomas (lanes 1 to 6) and normal human DNA (lane 7) were digested with restriction enzyme Hind III, separated by electrophoresis in 0.75% agarose gels, and transferred to nitrocellulose filters with 10× SSC. Filters were hybridized with ^{32}P-labeled RB-5 DNA (**A**) or ^{32}P-labeled RB0.9 DNA (**B**) in 40% formamide/6× SSC/5× Denhardt's solution/0.1% SDS/denatured salmon sperm DNA (100 μg/ml)/ 100 m*M* phosphate buffer (*p*H 7.0) at 42°C for 24 hours. Filters were washed twice in 2× SSC/0.1% SDS at room temperature for 15 minutes and once in 0.2× SSC/0.1% SDS at 60°C for 30 minutes. DNA quantitation was verified by washing filters at 95°C to remove probes and then hybridizing as above with a ^{32}P-labeled human growth hormone gene fragment (*38*). Band intensities indicated that lanes 5 to 7 had twice as much DNA as lane 4 and half as much DNA as lanes 1 to 3. With this adjustment, RB gene copy number was estimated at two for retinoblastoma samples 1, 2, 5, and 6, one for sample 4, and zero for sample 3.

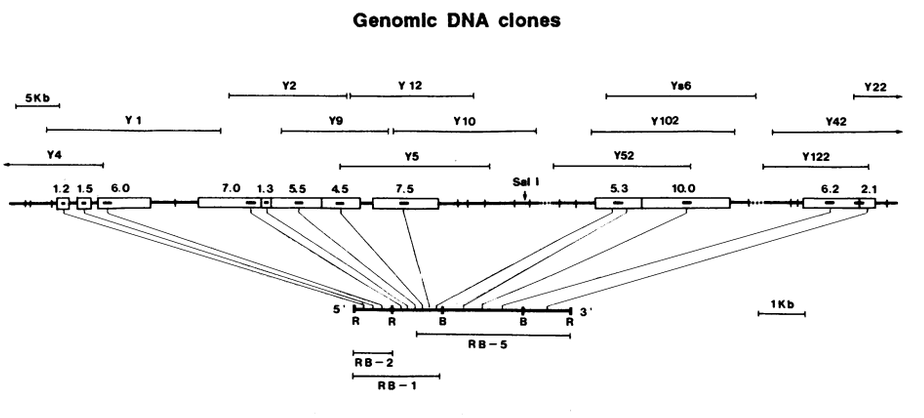

Fig. 4. Genomic organization of the RB gene and physical relationship to its mRNA transcript. A simplified restriction map of RB cDNA is shown (**bottom**), where R = Eco RI site and B = Bgl II site. Transcriptional orientation was determined as described (*19*). RB-1 and RB-5 cDNA clones were then used to screen an EMBL-3 genomic phage library constructed from retinoblastoma cell line Y79 as described in Fig. 1. About 20 relevant phage clones were isolated, of which the nonredundant ones are shown (**top**). Characterization of these clones included Hind III restriction mapping by the *cos* method (*39*), and DNA blotting analysis with cDNA subfragments as probes to establish their physical relationship with genomic fragments. The genomic map (middle) was constructed from this data and from data in Fig. 3. Hind III cleavage sites (vertical tics) and Hind III fragments (boxes) containing RB gene exons (solid bars) are shown. A single Sal I site is present among all clones. Dotted lines represent intron regions not yet mapped. Additional exons may be present beyond the 5′ end of the diagram.

The mRNA transcript orientation was determined by generating single-stranded M13 clones from both strands of RB1.8, which were then used as probes in RNA blotting analysis (*19*). This indicated that the DNA sense strand runs 5′ to 3′ from the 0.9-kb fragment of RB-1 to the 1.0-kb fragment of RB-5.

We surmised that some alterations in mRNA expression or size might be due to small deletions in genomic DNA encoding the candidate RB gene. Consequently, genomic DNA was extracted from corresponding tumor cells described for Fig. 2, digested with Hind III, and hybridized in DNA blotting analysis with probes RB-5 (Fig. 3A) and RB0.9 (Fig. 3B). Normal human genomic DNA was used as a control. Four of six retinoblastomas had genomic patterns identical to that of normal DNA. The RB gene was completely absent in sample 3. Decreased band intensities in DNA from sample 4 suggested the presence of a single copy of the RB gene, which was consistent with a cytogenetic feature of the tumor [del(13q12-q14)] (*20*). Despite the absence of RB mRNA transcripts, genomic DNA from sample 6 was not detectably altered. Thus only gross deletions but not small partial deletions of the RB gene were seen in these tumors, and detection of an intact genome did not imply normal gene expression. Such a disparity may be due to mutations in the promotor region. Explanations for decreased mRNA transcript size include small deletions within exons (which might not sufficiently alter gel mobility of large exon-containing Hind III fragments), or point mutations in exon splicing sequences or other mRNA processing signals.

The numerous large Hind III bands seen on DNA blotting analysis suggested that the RB locus was spread over a rather large genomic region. To further clarify genomic structure, more than 20 phage clones were isolated from a human genomic DNA library with RB-1 and RB-5 as probes (Fig. 4). These clones were characterized by restriction mapping and hybridization to subfragment cDNA probes. In conjunction with data from genomic DNA blotting (Fig. 3), the Hind III restriction map of the RB gene was constructed (Fig. 4). The RB gene consists of at least 12 exons scattered over more than 100 kb of DNA. One large intron of at least 20 kb is located between exons from the RB-1/RB-5 overlap region.

Friend *et al.* (*21*) recently described a cDNA clone with properties that were attributed to the RB gene. This clone detected a 4.7-kb mRNA transcript present in adenovirus 12–transformed retinal cells but absent in four of four retinoblastoma cell lines. Deletions involving part or all of this gene

were observed in five retinoblastomas, but none of these were internal homozygous deletions. These findings were not sufficient to precisely identify the RB gene. While sequence data were not included in their report, the restriction map of their cDNA clone was similar to ours.

Genes with important developmental functions are often found to have similar counterparts in phylogenetically related organisms. We examined genomic DNA from five other vertebrate species by hybridizing with probe RB-5. Under standard hybridization stringencies, homologous sequences were detected in all samples, with weaker hybridization intensity as evolutionary distance increased (Fig. 5). The homologous chicken gene was just barely detectable. Thus the RB gene is measurably divergent in vertebrates, and may not be conserved in more distant species.

We noted above that RB gene expression, while specifically altered in retinoblastoma, was not confined to normal fetal retina but was also seen in at least one other normal unrelated tissue, placenta. For a more extensive survey, mRNA from fetal and adult rat tissues was prepared and analyzed by RNA blotting (Fig. 6). A 4.6-kb mRNA transcript (presumably the normal size for rat) was detected in all tissues, though quantity varied markedly. A second species of transcript, approximately 2.3 kb in size, was also apparent in fetal rat brain. This short transcript may represent either differential processing of the RB gene or transcription of a separate but closely related gene.

Sequence analysis of clones RB-1 and RB-5 was performed by the dideoxy-terminator method of Sanger (*22*) to yield the reconstructed complete cDNA sequence (Fig. 7). Different deletion templates were generated by the "cyclone" method (*22*) in single-stranded M13 phage clones, which yielded greater than 95% of the sequence. The remaining gaps were sequenced by primer extension on both strands. The complete sequence contained 4523 nucleotides, with a short poly(A) tail and a polyadenylation signal sequence (AATAA) near the 3′ end. An open reading frame was present from the 5′ end to base 2688, with numerous additional in-frame stop codons further downstream. Translation from the first methionine codon (base 241) yielded a hypothetical protein of 816 amino acids (94,000 daltons in size). The second in-frame methionine was at base 346. Since the nucleotide sequence surrounding the first ATG is not typical of other known mRNAs (*23*), the start codon assignment should be regarded as tentative. A computer search of the National Biological Research Foundation protein sequence database detected no strong homology with any of more than 4000 published amino acid sequences. However, a number of nucleic acid–binding proteins and viral proteins showed weak sequence homology, with a yeast DNA-directed RNA polymerase (*24*) having the highest homology score.

The predicted protein sequence included ten potential glycosylation sites (*25*) but a candidate transmembrane domain (at least 20 consecutive hydrophobic residues) was not found. The amino acid hydropathy plot showed a mildly hydrophobic region near the putative amino terminus and a hydrophilic region at the carboxyl terminus (Fig. 8). We identified two pairs of short amino acid sequences that were bracketed by cysteine and histidine residues in the manner of metal-binding domains found in nucleic acid–binding proteins (*26*) (Fig. 7). A region of 54 amino acids from position 663 to 716 contains 14 proline residues (26%); proline-rich regions have also been observed in nuclear oncogene proteins *myc* and *myb* (*27*). While the significance of these observations is not well established, they suggest that the RB gene product may be a nucleic acid–binding protein.

Early hypotheses suggested that retinoblastoma may arise by inactivation of both alleles of a gene that normally functions to suppress tumor formation (*1, 28*). We have isolated a gene in chromosome region 13q14 that is either deleted or aberrantly expressed in each of six retinoblastomas examined. Tumor-specific alterations in gene expression provide the best evidence for identifying this gene as RB, and the examples of homozygous deletion and absence of mRNA expression support its postulated recessive nature. Further confirmation of RB gene identity and function awaits an appropriate biological assay. For example, retinoblastoma cells can be transfected with the normal RB gene and observed for loss of malignant phenotype. However, cells may fail to "revert" because, by the concept of multistep carcinogenesis, altering the ini-

Table 1. The mRNA expression of genes in region 13q14. Abbreviations are as follows: Est, esterase D; +, detectable mRNA expression; −, no detectable expression; (+), expression expected; ND, not done; alt, altered mRNA size.

Cells	Probes			
	Est	SD-1	SD-2	RB-1
Placenta	+	− (+)	−	+
Fetal retina	+	−	− (+)	+
Retinoblastoma				
1	+	−	−	+/alt
2	+	−	−	+/alt
3	+	−	−	−
4	+	−	−	±/alt
5	+	−	−	+/alt
6	+	−	−	−
Neuroblastoma				
1	+	−	−	+
2	+	−	−	+
3	+	−	−	+
Medulloblastoma				
1	+	ND	ND	+
2	+	ND	ND	+

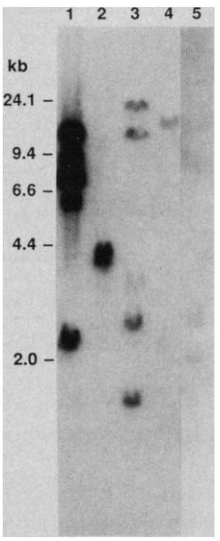

Fig. 5. Evolutionary conservation of the RB gene in vertebrates. About 10 μg of DNA extracted from human placenta (lane 1), calf retina (lane 2), feline fibroblasts (lane 3), mouse NIH 3T3 fibroblasts (lane 4), and chicken embryo fibroblasts (lane 5) were digested with Hind III and separated by electrophoresis in 0.75% agarose gels. DNA fragments were then transferred to nitrocellulose and hybridized with ^{32}P-labeled RB-5 cDNA as described for Fig. 3. At increased hybridization stringency (50% formamide), the signal for mouse and chicken became undetectable despite 7 days autoradiography.

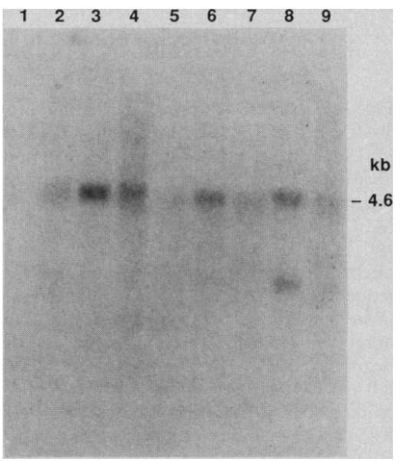

Fig. 6. Tissue expression of the RB gene in rat. Polyadenylated RNA was isolated from adult liver, brain, kidney, ovary, and spleen (lanes 1 to 5, respectively); from the placenta (lane 6); and from fetal (16-day gestation) rat tissues (liver, brain, and kidney in lanes 7 to 9, respectively). About 5 μg of mRNA was loaded into each lane of a 1% formaldehyde–agarose gel and analyzed as described for Fig. 2. The probe was a ^{32}P-labeled RB-1 (similar results were obtained with RB-5).

Fig. 7. Nucleotide and predicted amino acid sequences of the RB gene. RB-1 cDNA (1.6 kb) and Bgl II–digested RB-5 cDNA (yielding RB710, RB1.8, and RB1.0), were subcloned into M13 mp11, mp19, and mp10. Contiguous deletion mutants were constructed and sequenced (*22*) to obtain about 95% of the complete sequence. Confirmation of unclear regions and closing of sequence gaps were accomplished by synthesizing new oligonucleotide primers or by subcloning small fragments into new M13 phage. Sequences were entered directly into and analyzed on a microcomputer with a gel reader and DNA sequence software (International Biotechnologies Inc., New Haven, CT). Each of 4523 nucleotides was sequenced an average of five times. Eco RI linkers are noted at both ends of the sequence. A short poly(A) tail (broken underline) and the polyadenylation signal sequence (AATAA) (solid underline) are present near the 3′ end. The translated amino acid sequence (*40*) is shown, numbered from the first methionine residue. Translation from a hypothetical methionine 5′ to the beginning of RB-1 is also possible (*41*), since by primer extension methods about 100 nucleotides of RB mRNA are not included at the 5′ end of RB-1. However, there is a near-consensus ribosomal binding sequence starting at base 25 (underline). Potential glycosylation sites (sequences of the form N-X-[S or T]) are indicated (*). Two paired potential metal binding domains (sequences of the form [H or C]-$X_{3\text{-}5}$-[H or C] separated by 27 to 31 residues) are shown (boxes). The boundaries of a proline-rich region are marked (arrows). --, Termination codon.

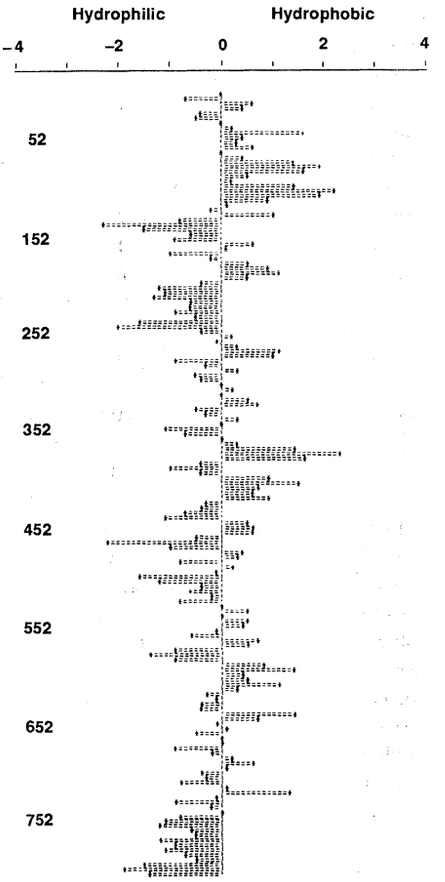

Fig. 8. Hydropathy plot of the predicted amino acid sequence of the RB gene. The predicted sequence shown in Fig. 7 was analyzed according to the methods of Kyte and Doolittle (*42*) to illustrate the regional predominance of hydrophilic or hydrophobic residues. Amino acid position numbers appear at the left.

tiating factor does not necessarily reverse the process (*10*). Ultimately, a definitive animal model may be necessary, consisting perhaps of a transgenic mouse carrying an anti-sense RB gene.

A second line of investigation will be to define the nature of alterations in the RB gene, especially the origin and consequences of slightly shortened mRNA transcripts. In this way, it may be possible to determine the function of the putative RB protein, its substrates (if any), and a mechanism by which alterations of the RB gene might induce retinoblastoma. Comings (*28*) speculated that the RB gene product "down-regulated" other growth-promoting genes or proto-oncogenes, the overexpression of which might lead to transformation. N-*myc* is amplified and expressed at high levels in several primary retinoblastomas and neuroblastomas (*29*); furthermore, N-*myc* expression appears to be developmentally regulated (*30*). For these reasons, N-*myc* has been proposed as a possible target for regulation by RB (*29*). Since we did not demonstrate any abnormality of the RB gene in three neuroblastomas, N-*myc* might not be directly related to the RB gene.

A developmental role for RB may be indirectly inferred from the "embryonal" characteristics of retinoblastoma as well as several observations in this report, such as conservation of the RB gene among vertebrates and its pattern of expression in rat tissues. Further investigation in this area may resolve an apparent paradox: how does a gene with ubiquitous tissue expression control the development of specific tumors? If we temporarily restrict our consideration

to retinoblastoma, several alternatives may be suggested: (i) only undifferentiated retinoblasts provide the proper "substrate" for RB gene function (perhaps in differentiation), and its expression in other tissues is vestigial or has other purposes; (ii) the RB gene functions in many tissues but inactivation of this gene is lethal during embryogenesis except when it occurs in a nonvital organ such as the eye; or (iii) inactivation of the RB gene occurs only in retinoblasts by unknown mechanisms.

The RB gene is of more general interest than simply that attendant to a rare childhood cancer. First, recessive genes similar to RB may control other unusual embryonal childhood cancers such as nephroblastoma (Wilm's tumor), hepatoblastoma, embryonal rhabdomyosarcoma, and neuroblastoma (*31*). Like retinoblastoma, all of these tumors resemble massive overgrowth of tissues found in normal embryogenesis. Nephroblastoma has been associated with deletions of chromosome region 11p13, and adjacent polymorphic markers become homozygous in these tumors (*32*). At a great phylogenetic distance, a recessive gene called "lethal giant larva" has been isolated in *Drosophila* that causes overgrowth of primitive neurectoderm (*33*), a behavior highly reminiscent of the RB gene.

Second, the role of the RB gene in other tumors may now be explored. Retinoblastoma patients have a high rate of second malignancies occuring at a variety of sites (*34*); osteosarcoma is the most common. Reduction to homozygosity in region 13q14 has been reported in osteosarcomas even without prior retinoblastoma (*35*), which suggests a common oncogenic mechanism for the two tumors despite their histologic and ontologic dissimilarity. In contrast, neuroblastoma and medulloblastoma, which are considered closely related to retinoblastoma, apparently do not involve alterations of RB. Further study of the RB gene will likely provide insight into unifying mechanisms of oncogenesis.

REFERENCES AND NOTES

1. A. G. Knudson, *Proc. Natl. Acad. Sci. U.S.A.* **68**, 820 (1971); F. Vogel, *Hum. Genet.* **52**, 1 (1979); A. L. Murphree and W. F. Benedict, *Science* **223**, 1028 (1984).
2. E. Matsunaga, *Am. J. Hum. Genet.* **30**, 406 (1978); J. Francois, M. T. Matton, S. DeBie, Y. Tanaka, D. Vandenbulcke, *Ophthalmologica* **70**, 405 (1976); A. G. Knudson, *Cancer* **35**, 1022 (1975).
3. K. P. Lele, L. S. Penrose, H. B. Stallard, *Ann. Hum. Genet.* **27**, 171 (1963); U. Francke, *Birth Defects* **12**, 131 (1976); J. J. Yunis and N. Ramsay, *Am. J. Dis. Child.* **132**, 161 (1978); R. S. Sparkes *et al.*, *Science* **208**, 1042 (1980); L. S. Strong, V. M. Riccardi, R. E. Ferrell, R. S. Sparkes, *ibid.* **213**, 1501 (1981).
4. R. S. Sparkes *et al.*, *Science* **219**, 971 (1983).
5. W. F. Benedict, A. Banerjee, C. Mark, A. L. Murphree, *Cancer Genet. Cytogenet.* **10**, 311 (1983); G. Balaban, F. Gilbert, W. Nichols, A. T. Meadows, J. Shields, *ibid.* **6**, 213 (1982).
6. W. F. Benedict *et al.*, *Science* **219**, 973 (1983).
7. W.-H. Lee, R. E. Bookstein, W. Wheatley, W. Benedict, E. Y.-H. P. Lee, *Hum. Genet.*, in press.
8. W. K. Cavenee *et al.*, *Nature (London)* **305**, 779 (1983).
9. T. P. Dryja, J. M. Rapaport, J. M. Joyce, R. A. Petersen, *Proc. Natl. Acad. Sci. U.S.A.* **83**, 7391 (1986).
10. J. M. Bishop, *Science* **235**, 305 (1987); R. Sager, *Cancer Res.* **46**, 1573 (1986).
11. P. Duesberg, *Science* **228**, 669 (1985); G. Klein and E. Klein, *Nature (London)* **315**, 190 (1985).
12. W.-H. Lee, W. Wheatley, W. F. Benedict, C. M. Huang, E. Y.-H. P. Lee, *Proc. Natl. Acad. Sci. U.S.A.* **83**, 6790 (1986).
13. P. Ward *et al.*, *J. Med. Genet.* **21**, 92 (1984).
14. E. Y.-H. P. Lee and W.-H. Lee, *Proc. Natl. Acad. Sci. U.S.A.* **83**, 6337 (1986).
15. M. Lalande *et al.*, *Cancer Genet. Cytogenet.* **13**, 283 (1984).
16. R. Bookstein *et al.*, in preparation.
17. T. Maniatus, E. F. Fritsch, J. Sambrook, *Molecular Cloning: A Laboratory Manual*, (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY (1982).
18. J. Millan, *J. Biol Chem.* **262**, 3112 (1986).
19. Transcriptional orientation was determined by RNA blotting analysis with M13mp19 phage probes containing single-stranded RB1.8 DNA inserts. To label the M13 probe, a 17-mer hybridization primer was first annealed with single-stranded M13 DNA. The second strand was synthesized by DNA polymerase (large fragment) with four triphosphates including ^{32}P-(α)-dATP. By limiting the amount of one triphosphate, the distal end of the second strand will not be synthesized and the insert will remain single-stranded. Two recombinant M13 phage containing inserts with opposite orientations were labeled. Under nondenaturing conditions, only one hybridized with corresponding mRNA, indicating that RB-1 was 5′ on the DNA sense strand and RB-5 was 3′.
20. R. C. McFall, T. W. Sery, M. Makedon, *Cancer Res.* **37**, 1003 (1977); T. W. Sery *et al.*, in preparation.
21. S. H. Friend *et al.*, *Nature (London)* **323**, 643 (1986).
22. F. Sanger, S. Nicklen, A. R. Coulsen, *Proc. Natl. Acad. Sci. U.S.A.* **74**, 5463 (1977); R. M. K. Dale, B. A. McClure, J. P. Houchins, *Plasmid* **13**, 31 (1985).
23. M. Kozak, *Nucleic Acids Res.* **12**, 857 (1984).
24. L. A. Allison, M. Moyle, M. Shales, C. J. Ingles, *Cell* **42**, 599 (1985).
25. P. V. Wagh and O. P. Bahl, *CRC Crit. Rev. Biochem.* **10**, 307 (1981).
26. J. M. Berg, *Science* **232**, 485 (1986).
27. R. Weiss, N. Teich, H. Varmus, J. Coffin, *RNA Tumor Viruses* (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY 1985).
28. D. E. Comings, *Proc. Natl. Acad. Sci. U.S.A.* **70**, 3324 (1973).
29. W.-H. Lee, A. L. Murphree, W. F. Benedict, *Nature (London)* **309**, 458 (1984); M. Schwab *et al.*, *ibid.* **305**, 245 (1983).
30. A. Jakobovitz *et al.*, *ibid.* **318**, 188 (1985).
31. A. Koufos *et al.*, *ibid.* **309**, 170 (1984).
32. S. H. Orkin, D. S. Goldman, S. E. Sallan, *ibid.*, p. 172; E. R. Fearon, B. Vogelstein, A. P. Feinberg, *ibid.*, p. 176.
33. B. M. Mechler, W. McGinnis, W. J. Gehring, *EMBO J.* **4**, 1551 (1985).
34. D. H. Abramson, R. M. Ellsworth, F. D. Kitchin, G. Tung, *Ophthalmology* **91**, 1351 (1984); R. D. Jensen and R. W. Miller, *N. Engl. J. Med.* **285**, 307 (1971); N. R. Schimke, J. T. Lowman, G. A. B. Cowan, *Cancer* **34**, 2077 (1974).
35. M. F. Hansen *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **82**, 6216 (1985).
36. T. W. Reid *et al.*, *J. Natl. Cancer Inst.* **53**, 347 (1974).
37. P. W. J. Rigby, M. Dieckmann, C. Rhodes, P. Berg, *J. Mol. Biol.* **113**, 237 (1977); A. P. Feinberg and B. Vogelstein, *Anal. Biochem.* **132**, 6 (1983).
38. F. M. DeNoto, D. D. Moore, H. M. Goodman, *Nucleic Acids Res.* **9**, 3719 (1981).
39. H. R. Rackwitz, G. Zehetner, A. M. Frischauf, H. Lehrach, *Gene* **30**, 195 (1984).
40. Single-letter abbreviations for the amino acid residues are: A, Ala; C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; M, Met; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; W, Trp; and Y, Tyr.
41. The additional 80 amino acids translated from nucleotides 1 through 240 are as follows: NSGDSGPEDL PLVRLEFEET EEPDFTALCQ KLKIPDHVRE RAWLTWEKVS SVDGVLGGYI QKKKELWGIC IFIARVDLDE. The predicted peptide is neither strongly hydrophobic nor hydrophilic in this region.
42. J. Kyte and R. F. Doolittle, *J. Mol. Biol.* **157**, 105 (1982).

43. We are very grateful to the late Dr. P. Lampert, former department chairman of Pathology, for his full support at the initiation of this project. We thank S. Latt for probe H3-8, T. Sery and R. Phillips for retinoblastoma cell lines, C. Lin for construction of cDNA libraries, T. Dryja for helpful suggestions, A. Hsueh for critical comments, and G. Walter for interest and encouragement. Supported by grants to W.H.L. from the National Eye Institute (NEI), the Knights Templar Foundation, Inc., and the Cancer Research Coordinating Committee, University of California. R.E.B. is a recipient of a Physician Scientist Award from NEI. E.L. was supported by a fellowship from the Giannini Foundation, Bank of America.

27 January 1987; accepted 12 February 1987