Cloning of the esterase D gene: A polymorphic gene probe closely linked to the retinoblastoma locus on chromosome 13

(gene cloning/genetic counseling/cancer/recessive mutation)

JEREMY SQUIRE*, THADDEUS P. DRYJA†, JAMES DUNN‡, AUDREY GODDARD*, THEO HOFMANN‡,
MARIA MUSARELLA‡, HUNTINGTON F. WILLARD‡, ANDREW J. BECKER‡, BRENDA L. GALLIE‡,
and ROBERT A. PHILLIPS*

*Department of Medical Biophysics, University of Toronto, and Hospital for Sick Children Research Institute, 555 University Avenue, Toronto, ON, Canada M5G 1X8; †Harvard Medical School, Massachusetts Eye & Ear Infirmary, Howe Laboratory of Ophthalmology, Boston, MA 02114; ‡Department of Ophthalmology, University of Toronto, and Research Institute, Hospital for Sick Children, 555 University Avenue, Toronto, ON, Canada M5G 1X8; §Department of Biochemistry, University of Toronto, Toronto, ON, Canada M5S 1A8; and ¶Department of Medical Genetics, University of Toronto, Toronto, ON, Canada M5S 1A8

Communicated by J. Tuvo Wilson, May 27, 1986

ABSTRACT The study of recessive oncopgenes such as those responsible for retinoblastoma and Wilms tumor is difficult because the gene products involved are unknown and because the diseases are not associated with unique cellular or molecular phenotypes suitable for genetic manipulation. Since the gene for esterase D (ESD) is known to be tightly linked to the retinoblastoma locus (RBI) in the q14.1 band of chromosome 13, we have cloned the ESD gene from a human cDNA library by using oligonucleotides specific for a partial amino acid sequence of the purified enzyme to provide a genetic marker for further studies on retinoblastoma. The putative ESD gene codes for a message of 1.2 kilobases, which is present in all cell types examined, and maps to 13q14.1, thus confirming that it is the ESD gene. Restriction enzyme analysis reveals a restriction fragment length polymorphism with Apa I; this polymorphism results from the heterozygosity of 32% of the individuals tested and is shown to be useful in identifying carriers of the mutation responsible for retinoblastoma. A preliminary screen of 24 retinoblastoma tumors by Southern blot did not reveal any homologous deletions or rearrangements of the ESD locus.

The recessive oncopgenes that initiate certain childhood tumors (1–3) may normally regulate the expression of other genes during development (4). Failure of the regulating gene to function normally may allow the inappropriate expression of other genes, perhaps traditional oncopgenes, leading to the development of a malignant state. Because of their importance in initiation of cancer, and their probable role in normal regulation, it would be of great interest to identify these recessive oncopgenes and to characterize their products.

One such recessive mutation that has been well characterized is retinoblastoma, at the genetic locus RBI. Since an experimental assay for the gene product of the RBI locus is not available, a cloned closely linked marker would provide a means to approach the retinoblastoma gene. Although other genes have been provisionally assigned to chromosome 13 (5, 6), the gene for esterase D ESD is the only one known to be tightly linked to retinoblastoma (7, 8). Among informative families (extensively studied families of patients with retinoblastoma), there are no reported genetic recombinations between the ESD and RBI loci (9). The only separation of the two loci is defined by a gross chromosome abnormality involving a deletion beginning in the q14 band and extending distally to q22 (10). Since this deletion affects the RBI locus and not the ESD locus, the ESD gene must lie closer to the centromere than RBI. In a careful prophase banding analysis of several patients with deletions, Ward et al. (8) regionally mapped the ESD and RBI loci in the upper half of the q14.11 subband of chromosome 13. Since each prophase band contains approximately 3000 kilobases (kb) of DNA, these two loci must be separated by less than this distance. It is unclear where other DNA markers mapped to the q14 region (11, 12) are located relative to the RBI and ESD loci. To obtain a close molecular probe to RBI for use in 'chromosomal walking' experiments, we undertook the molecular cloning of the ESD gene.

MATERIALS AND METHODS

Esterase D Protein Purification. Packed erythrocytes (14 liters) were used to obtain 4 mg of pure esterase D (13). Because the amino terminus was blocked, an amino acid sequence of 25 residues was obtained from the largest cyanogen bromide cleavage peptide fragment CN3 using an Edman sequenator followed by HPLC. The sequence was obtained using a Beckman Model 890C sequencer.

Oligonucleotide Screening of DNA Library. Two oligonucleotides, a 20-mer and a 17-mer, were chosen as probes based on the amino acid sequence of CN3. Because cyanogen bromide cleavage occurs at the carboxyl side of methionine residues, we also included an AUG in one oligonucleotide. The degeneracy of the genetic code required a redundancy of 256 for the 20-mer and 128 for the 17-mer. Both were synthesized using an Applied Biosystems (Foster City, CA) model 380A DNA synthesizer and were radiolabeled at their 5' ends using [γ-32P]ATP and T4 polynucleotide kinase to a specific activity of 1–5 × 106 cpm/µg. A human cDNA library, generously provided by Okayama and Berg (14) and made from human simian virus 40-transformed fibroblast mRNA in the pcD vector (14), was probed. Initially 106 colonies at a density of 109 colonies per 25-cm2 dish were screened using the 20-mer. Secondary colony purification was carried out at dilutions below 103 colonies per plate.

Nitrocellulose filters with blotted 'colony lifts' were prehybridized overnight in 6x SSC (1× SSC = 0.15 M NaCl/15 mM sodium citrate, pH 7.0) at 68°C, then washed at 0°C. Hybridization was done for 24 hr in prehybridization buffer containing 0.5–1.0 × 106 cpm of probe per ml. The filters were washed for 2 hr in 6x SSC with two or three changes of buffer. Prehybridization, hybridization, and washing of filters was done at 44°C for the 20-mer probe and at 50°C for the 17-mer probe. Independent verification using the 17-mer was carried out on DNA obtained

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviation: kb, kilobase(s).
from rapid plasmid preparations using alkaline lysis (15) followed by Southern blotting and oligonucleotide hybridization.

**DNA Sequencing.** All sequencing was done by the Sanger dideoxy chain-termination method (16). The 3′ BamHI restriction fragment of pESD-14.1.1, which was positive for both oligonucleotides, was inserted into pEMBL (17) and sequenced in both directions, using the M13 single-stranded primer and unidirectional digestion with exonuclease III to generate targeted deletions in the cloned DNA segment (18). The 117-nucleotide fragment on the 3′ side of the internal BamHI recognition site of pESD-14.1.1 was sequenced and found to contain a complete open reading frame of 39 amino acids.

**Cell Lines and Retinoblastoma Tumors.** The fibroblasts AG 1142 and AG 2718 were obtained from the Human Genetic Mutant Cell Repository (19). The lymphoblastoid line of the patient EL was generously provided by Hans Ochs. Retinoblastoma tumors were surgically removed from patients and grown as xenografts in nude mice and in tissue culture. The panel of somatic cell hybrids used for mapping the ESD gene has been described in detail elsewhere (20, 21).

**Blot Hybridization.** DNA was extracted from 0.5 to 1.0 × 10⁸ fibroblasts, lymphoblastoid lines, or retinoblastoma tumors by the standard method of phenol and isomyl alcohol extraction followed by ethanol precipitation (22). Samples of DNA were digested with restriction enzymes (Boehringer Mannheim) using conditions recommended by the supplier. Digested genomic DNA was separated by gel electrophoresis in 1.0% agarose, and blotted onto nitrocellulose membranes (23). DNA probes were radiolabeled with [³²P]dCTP by nick-translation (24) to a specific activity of 3 × 10⁸ cpm/µg. Filters were washed under high stringency conditions (0.1× SSC/0.2% Na₂SO₄, 65°C). Autoradiography was carried out overnight at −70°C using one or two intensifying screens. Probes were stripped with water at 60°C for 30 min. DNA hybridization ratios were estimated by band grain density on autoradiograms using a Bio-Rad 620 Video Densitometer and Bio-Rad Model 3392A Integrator.

**DNA Probes.** The ESD probe refers to the pESD-14.1.1 Xho I fragment; MYC probe refers to the pHISR 9-kb EcoRI-HindIII fragment (25); the T-cell receptor (TCR) probe refers to the pYT35 Ava I-EcoRV fragment (26).

**RESULTS**

**Isolation of Esterase D cDNA Clone.** Because of its widespread use as a polymorphic genetic marker, esterase D has been studied in many laboratories, and a purification scheme for this enzyme has been described by Scott and Wright (13). Applying their protocol to human erythrocytes, we obtained 4 mg of enzyme, which was homogeneous by gel electrophoresis. The monomer molecular size of the purified enzyme was 33 kDa; the active form of the enzyme is a dimer. Although the amino terminus was blocked, we obtained a partial amino acid sequence from a large, 11-kDa, cyanogen bromide fragment (CN3). A sequence of 25 amino acids was obtained and is shown [Fig. 1A (Top)]. The sequences enclosed by the dashed rectangles in Fig. 1A (Middle) are regions chosen for the synthesis of oligonucleotides. Both oligonucleotides were synthesized as mixtures; the 17-mer had a redundancy of 128, and the 20-mer had a redundancy of 256. A human cDNA library was screened using the 20-mer as a probe, and positive colonies were identified and purified by two rounds of colony hybridization. Of the 20 colonies that repeatedly scored positively for hybridization with the 20-

---

![Fig. 1](image-url)  
**Fig. 1.** (A) The amino acid sequence of the amino-terminal portion of the purified esterase D peptide fragment CN3 and a partial nucleotide sequence of the cDNA insert of pESD-14.1.1 and derived amino acids. The codons for the partial amino acid sequence from CN3 (Top) and their expected mRNA sequences (Middle) were derived. Oligonucleotides were synthesized for use as mixed redundant probes from two regions of CN3 (enclosed in dashed rectangles). Those probes were used to isolate the plasmid pESD-14.1.1 from a human fibroblast cDNA library. DNA from this plasmid was sequenced (Bottom) to confirm its derivation from CN3. The cross-hatched rectangle indicates its position on pESD-14.1.1. The amino acid at position 8 (+) was not identifiable (cysteine was not detectable with the amino acid sequencing protocol used). (B) Restriction map and location of the partial amino acid sequence of CN3 on pESD-14.1.1 cDNA. The fragment containing the oligonucleotide hybridizing sequence was identified by restriction mapping and Southern blotting. The location of the sequenced 25 amino acids of CN3 is indicated by the horizontal solid bar. Forty nucleotides of pcD vector DNA are present in the most 3′ position. The Pst I site is immediately on the 5′ side of the pcD oligo(dG) tract. Restriction enzymes: P, Pst I; B, BamHI; S, Sac I.
mer, only 4 also hybridized with the 17-mer. Analysis of these 4 clones revealed that they had the same restriction map and were homologous to each other on Southern blots. The clone with the largest insert (1.2 kb), pESD-14.1.1, was selected for further analysis. The DNA sequence was determined for that region of the clone containing the sequences hybridizing to the oligonucleotides. Fig. 1A shows the nucleotide sequence and the amino acid sequence associated with the single open reading frame. The nucleotide sequence confirms that the cDNA indeed codes for the fragment derived from the purified protein. Position 8 in the peptide was not identified because cysteine was not detected in the method used for amino acid sequence analysis.

Expression and Mapping of Esterase D. Esterase D appears to be constitutively produced in all tissues. The pESD-14.1.1 probe detects a 1.2-kb mRNA in retinoblastoma, neuroblastoma, HeLa cell, ovarian tumor, and bladder tumors (data not shown). To establish unequivocally that the isolated gene was ESD, it was mapped to chromosome 13 by four different techniques. (i) The pESD-14.1.1 probe hybridized to identical HindIII restriction fragments in blots of total human DNA and in DNA isolated from a sorted chromosome 13 library (11) (data not shown). (ii) The presence of the gene was examined in 25 mouse–human hybrids (20, 21). There was no discordance in the association of the presence of the isolated clone and the presence of chromosome 13; all other chromosomes showed discordancies in five or more hybrids (data not shown). In addition, there was complete concordance in these hybrids between the presence of ESD DNA and the presence of the gene for human propionyl-CoA carboxylase (α chain), which has been independently assigned to chromosome 13 (21). If there are multiple copies of the ESD gene or pseudogenes, they must also be on chromosome 13. (iii) The probe was tested for linkage to the retinoblastoma locus using families with hereditary retinoblastoma. A total of 35 individuals were studied, and the probe detected a polymorphic Apa I recognition site with two codominant allelic restriction fragments of 8.6 and 7.2 kb (Fig. 2). This restriction fragment length polymorphism must be distinct from the allelic isozymes of ESD detected by starch gel electrophoresis (27) because its allele frequency (0.8 and 0.2) is different from the isozyme allele frequency in Caucasians (0.9 and 0.1) and because all four possible haplotypes determined by the two polymorphisms have been found in humans. Three hereditary retinoblastoma families who carry this polymorphism have been identified. For eight Apa I heterozygotes from these families a cross-over between RBI and ESD would have been detected, but none was observed. These data are consistent with the known close linkage between the loci for ESD and hereditary retinoblastoma (28) and confirm the assignment of these DNA sequences to chromosome 13. (iv) We studied by quantitative hybridization the pESD-14.1.1 sequences in cell lines containing different deletions of chromosome 13. As shown in Fig. 3, two copies of pESD-14.1.1 are present in normal control fibroblasts and in the lymphoblastoid cells from patient EL who has a known deletion in chromosome 13 beginning between the ESD and RBI loci (10). ESD is present in single copy in two lines, AG1142 (29) and AG2718 (37); both lines have known deletions involving most of the q14 region affecting both the ESD and RBI loci. The lymphoblastoid cells from patient number RB449, who has a deletion distal to band 13q14 and who does not have retinoblastoma, also have two copies of the ESD gene (data not shown). The line AG1142 is particularly interesting because its deletion is shorter than the EL deletion and appears cytogenetically to begin at the same place as the EL deletion. As shown in Fig. 3, EL cells have two copies of pESD-14.1.1 and AG1142 has only a single copy. On the basis of these quantitative hybridization results and published cytogenetic analyses of deletions involving ESD, we can map the ESD gene to a point near the proximal breakpoint in the chromosome band 13q14.1, the same location proposed by Ward et al. (8) on the basis of cytogenetic analysis. The prediction of the amino acid sequence and the mapping to 13q14.1 allows unequivocal identification of this probe as the ESD gene.

Since the ESD locus is very close to the RBI locus, one might expect the ESD gene to be affected by some of the somatic mutations known to be associated with the RBI locus in all retinoblastoma tumors. Indeed, Benedict et al. (30) have suggested that one of their patients has a submicroscopic deletion affecting both loci; the retinoblastoma tumor that developed was monosomic, lacked ESD activity, and apparently retained only the deleted chromosome 13. In addition, Cowell et al. (31) reported several patients with half-normal levels of esterase D activity in their erythrocytes, again suggesting the possibility of microdeletion in q14.1 affecting both ESD and RBI. If mutations at the RBI locus are similar to mutations at other loci, approximately 10% may involve deletions (32), some of which may be quite large and include the ESD locus. Fig. 4 shows a Southern blot of DNA from 11 retinoblastoma tumors. As shown in this figure, none of these tumors has a homozygous deletion or alteration for any of the five major hybridizing fragments detected in a HindIII digest. A total of 24 tumors have been examined using HindIII, and all showed apparently normal germ-line configuration. However, it is important to note that a small percentage of tumors could have ESD present in a hemizygous state. Using standard techniques, only bands associated with deletion or rearrangement would be detectable; attempts to look for tumors with simple deletions resulting in single copy ESD would require a detailed analysis by quantitative hybridization.

DISCUSSION

Development of retinoblastoma almost certainly requires mutations in both alleles at the RBI locus (1, 2). In approximately 70% of retinoblastoma tumors, somatic chromosome changes lead to homozygosity of all markers on chromosome 13, resulting in homozygous mutations at RBI. Homozygous deletions beyond a certain size will be cellular lethals because they will delete other essential genes closely linked to the RBI locus. In support of this suggestion, patients with large, karyotypically detectable, germ-line deletions have fewer tumors than patients with a normal karyotype (33). It has been suggested that patients with congenital deletions are not susceptible to the most common somatic events that generate homozygosity (2, 34). Presumably, the remaining normal
alleles in deletion patients is inactivated by a point mutation or small deletion within the RBI locus. Thus, homozygous deletions affecting ESD and RBI may be lethal. On the basis of our failure to detect homozygous deletions in 24 tumors, the possibility that the patient reported by Benedict et al. (30) represents two independent mutations, one in RBI and one in ESD, rather than a single deletion affecting both loci must be considered.

The natural substrate for esterase D is O-acetylated sialic acid, and a metabolic role for the enzyme has been postulated in the reutilization of sialic acids (35). Conservation and ubiquitous expression of the ESD gene suggests an important normal function. We have found esterase D to be expressed in all cell types tested and to be highly conserved. At high stringency hybridization, Pst I-digested DNA had a similar restriction pattern in both rodents and human DNA, and only one major band differed between HindIII and EcoRI digests. This stability is consistent with the wide tissue distribution found for the O-acetyl esterase, supporting the proposed identification of Varki et al. (35).

ESD is the fourth DNA fragment mapped to the q14 band on chromosome 13. The others are random DNA fragments, pTD2 (12), H2-42 (11), and H3-8 (11). The precise order of these markers in relation to each other and to ESD and RBI is unknown. We found two copies of the H3-8 sequence present in EL fibroblasts (data not shown). Thus, this probe, like ESD, is proximal to the RBI locus, but its relation to ESD is unknown. As with the ESD probe, no abnormalities were observed in the H3-8 sequence in the 24 tumors studied. We have not examined the hybridization of H2-42 and pTD2 with our panel of tumor.

On the basis of published results and the data presented here, it is only possible to give a crude estimate of the separation of the ESD and RBI loci. Our preliminary results suggest that ESD is more than 5 kb from the RBI locus; otherwise we should have detected abnormalities in the ESD gene in some of the tumors examined. As described above, Ward et al. (8) have placed the ESD and RBI loci in the top half of the small chromosome subband 13q14.11, a region encompassing at most 3000 kb. Thus, we estimate that ESD and RBI are separated by more than 5 kb and less than 1500 kb. As ESD is the only well-characterized genetic marker in the immediate vicinity of RBI, the cDNA probe spanning more than 30 kb of this genomic region will provide an ideal starting place for chromosomal walking experiments. Fur-
thermore, the availability of the EL deletion will allow orientation with respect to the centromere in initial steps of the walk. Application of the techniques for manipulating large fragments of DNA (36) should allow ordering of ESD and the other DNA fragments on chromosome 13 and an estimation of the distances between the markers and the breakpoints of various deletions associated with retinoblastoma. The ESD probe will be immediately useful for improved genetic counseling for patients with retinoblastoma because of the much larger numbers of heterozygotes now detectable using the Apa I polymorphism.

We thank A. M. Lambonwah and R. Gravel for invaluable advice concerning oligonucleotides; M. Lalande and S. Latt for providing their sorted chromosome 13 library; M. Cantor, J. Cocking, and J. Chin for excellent technical assistance; R. Tinoco for help with labeling; A. Cunningham for purification of ESD; and S. Bree for determination of the amino acid sequence. This work was supported by grants from the Hospital for Sick Children Foundation, the Medical Research Council (MRC) of Canada, and the National Cancer Institute of Canada. J.S. is supported by the Canadian Institute of the Blind. A.G. is supported by an MRC Studentship. B.L.G. is a research associate of the Ontario Cancer Treatment and Research Foundation. H.F.W. is a research scholar of the MRC.