Hypomethylation of the interphotoreceptor retinoid-binding protein (IRBP) promoter and first exon is linked to expression of the gene

Adriana Albini[1,2], Judith Toffenetti[1], Zhen Zhu[2+], Gerald J.Chader[1] and Douglas M.Noonan[2]
[1]National Eye Institute, National Institutes of Health, Bethesda, MD, USA and [2]Instituto Nazionale per la Ricerca sul Cancro, Genova, Italy

Received May 14, 1990; Revised and Accepted August 2, 1990 EMBL accession no. X53044

ABSTRACT

The interphotoreceptor retinoid-binding protein (IRBP) is limited in expression to retinal photoreceptor cells and a subset of pinealocytes. We have obtained a genomic clone containing the entire coding region and 7 kb of 5′ flanking sequence. As a first step in studying IRBP gene regulation we have examined the CpG methylation patterns of the entire IRBP gene in expressing and non-expressing human cells. This has been done by isolation of high molecular weight DNA from Y-79 cells grown in suspension or attached to poly-D-lysine, which synthesize IRBP at different levels, and from human lymphocytes, which were shown by northern analysis to lack IRBP message. The DNA was digested by either *Hpa II*, *Msp I*, or *Hha I*. Southern blots were prepared with these digests and hybridized with probes made from fragments covering the complete genomic clone. Probes from the first exon, the introns and the 3′ end gave banding patterns which showed no differences between the expressing cells and the lymphocytes. A probe from the very 5′ end did not give a clear banding pattern, probably due to the presence of repetitive elements in the probe. However, a *Hind III* probe covering the 5′ flanking 3 kb and the beginning of the first exon hybridized with a 1.8 kb band in *Hpa II* digests of Y-79 cells which was not present in *Hpa II* digests of lymphocyte DNA. In addition, a 2.1 – 2.3 kb *Hha I* band was found only in the Y-79 DNA digests. Sequence analysis of the promoter region indicated that these bands were due to hypomethylation of sites within a CpG rich island from – 1578 to – 1108 in the promoter and hypomethylation of sites in the beginning of the first exon. A *Hha I* site between the CpG island and the first exon was not hypomethylated in the expressing Y-79 cells. We propose that hypomethylation of the CpG rich island of the IRBP promoter and the first exon is linked to the expression of this gene.

INTRODUCTION

Interphotoreceptor retinoid-binding protein (IRBP) is a large lipoglycoprotein widely distributed among vertebrate species (1). It has an extremely limited tissue distribution, however, as it is synthesized only in the photoreceptor cells of the retina, and in some cells in the pineal gland (2). Retinoblastoma tumor cells, which are postulated to be primitive stem cells of the retina (3), also synthesize this protein (4,5), as does the human retinoblastoma cell line Y-79 (1). The expression of IRBP in Y-79 cells can be modulated by external stimuli (6), and can be enhanced by attachment substrates such as poly-D-lysine (7). We have obtained a human genomic clone which contains the entire coding region of the IRBP protein, as well as approximately 7 kb of 5′ flanking sequence and some 3′ flanking DNA. The exon sequence of what appears to be an identical clone has been reported (8). Exon containing regions have been used to probe RNA northern and dot blots to demonstrate that Y-79 cells grown in suspension make IRBP, but when these cells are induced to attach to poly-D-lysine, they make several fold more IRBP. The extremely limited tissue distribution of this gene, as well its compact structure which we have completely cloned, provided us with a unique opportunity to study the relationship between the methylation state of the gene segments and the expression of the gene.

MATERIALS AND METHODS

IRBP Gene isolation

Two clones containing the human IRBP gene, λ4 and λ24, were isolated from a human lymphocyte λEMBL 3 genomic library (Clonetech) by screening with 3 oligonucleotide probes: E (CAGCACAACCAGCTGAGGGTGAAGCGGAGCCCC) and M (GGGGAGCGGACGGGCGGAGGCGCACTGTCTGTG) based on cDNA sequence (9) corresponding to the 3′ end and middle of the cDNA, respectively, and B (TTCCAGCCC^{A}T^{G}CCCTGGTGCTGGACATGGCCAAGGTGCTGGTGGACAACTA), based on N-terminal amino acid sequence (10) and the most common human codon usage (11). The clones were

+ World laboratory (project ACAB #MCD 1 I-IV) fellow from Peking Union Medical University, Beijing, China

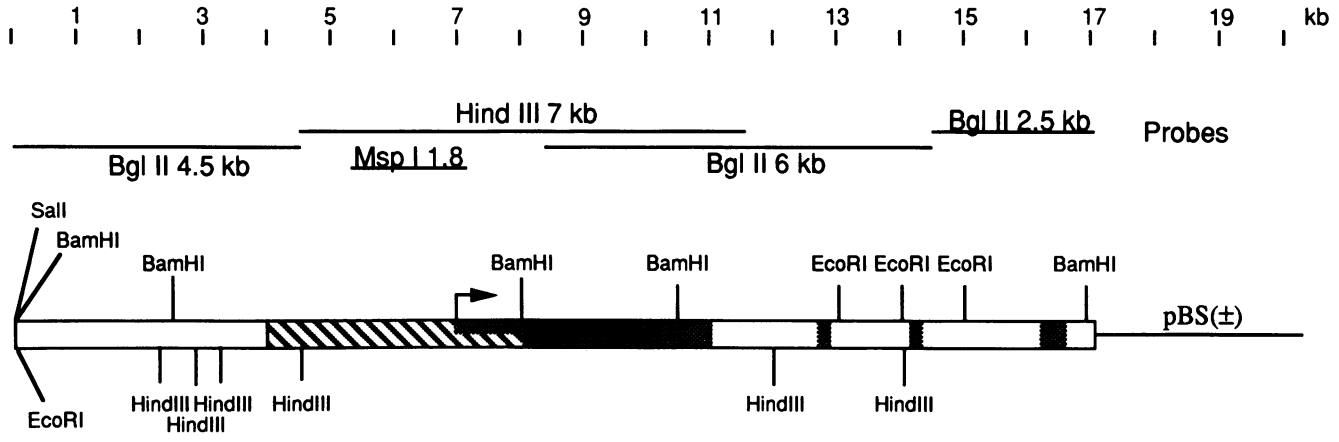

Figure 1. Schematic model of the entire IRBP gene subcloned into pBS(±). Shaded areas are the approximate locations of the exons, based on restriction mapping oligonucleotide hybridization, sequence data and exon sequences reported by Liou et al (8). Arrow indicates the beginning of transcription. Probes used in this study are shown as single lines above the model. The stripped portion of the model is the promoter region which has been sequenced, a model of this region is shown in figure 5.

confirmed by double-stranded Sanger dideoxy-sequencing of the λ phage using the oligonucleotides E and M as primers, the resulting sequence was identical to the published cDNA sequence. All three oligonucleotides hybridized to the 17 kb insert, indicating the clones contained the entire translated region. The insert from λ4 was subcloned into the pBS ± vector. Restriction maps were generated and approximate positions of the exons determined by oligonucleotide hybridization.

Sequence Analysis

A *Bam HI* fragment corresponding to portions of the first intron and the promoter was subcloned and sequenced on both strands using a directed deletion strategy and Sanger dideoxy sequencing (12).

Probes

Hind III, *Bgl II* and *Msp I* restriction fragments derived from the parent clones were used as probes to cover the entire span of the IRBP gene. Restriction fragments were purified and radiolabeled with ^{32}P-CTP by random oligonucleotide priming (Boehringer Mannheim).

Preparation of Southern and Northern Blots

Normal human lymphocytes were prepared by banding in ficoll. Residual red blood cells were lysed by addition of 84% NH_4Cl, then washed by resuspension in PBS. Y-79 cells were grown in suspension culture or as monolayers on poly-D-lysine in RMPI 1640 medium supplemented with 10% fetal calf serum, penicillin (100 U/ml), streptomycin (50 μg/ml) and glutamine (300 μg/ml) at 37°C in a 5% CO_2, humidified atmosphere. Genomic DNA was purified from the cells by standard methods (13). DNA was over-digested with 5 units/μg of either *Msp I* (recognition site CCGG, methylation insensitive), *Hpa II* (recognition site CCGG, methylation sensitive) or *Hha I* (recognition site GCGC, methylation sensitive). Digested DNA was then electrophoresed in 0.7% agarose gels and Southern blotted onto nylon (Hybond, Amersham) filters as described (13). Filters were probed with 2 million cpm of probe per ml of hybridization fluid (5× SSC, 5× Denhardt's solution, 0.1% SDS, 100μg/ml denatured Salmon

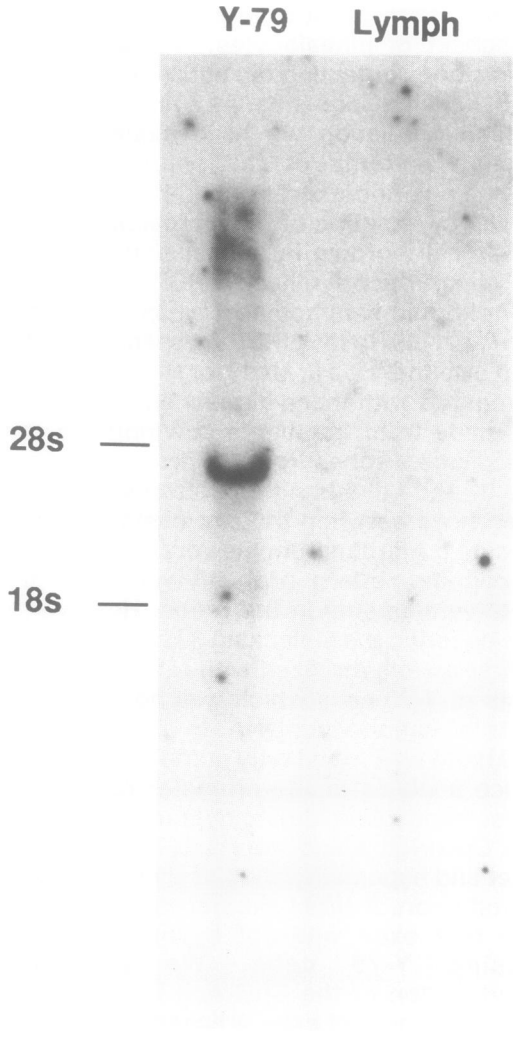

Figure 2. Northern analysis using the *Hind III* 7 kb probe, which includes th first exon of IRBP. Y-79: 20 μg of total RNA prepared from Y-79 cells grow in suspension. Lymph: 20 μg of total RNA prepared from normal human bloo lymphocytes. Positions of the ribosomal RNA subunits are shown as marker

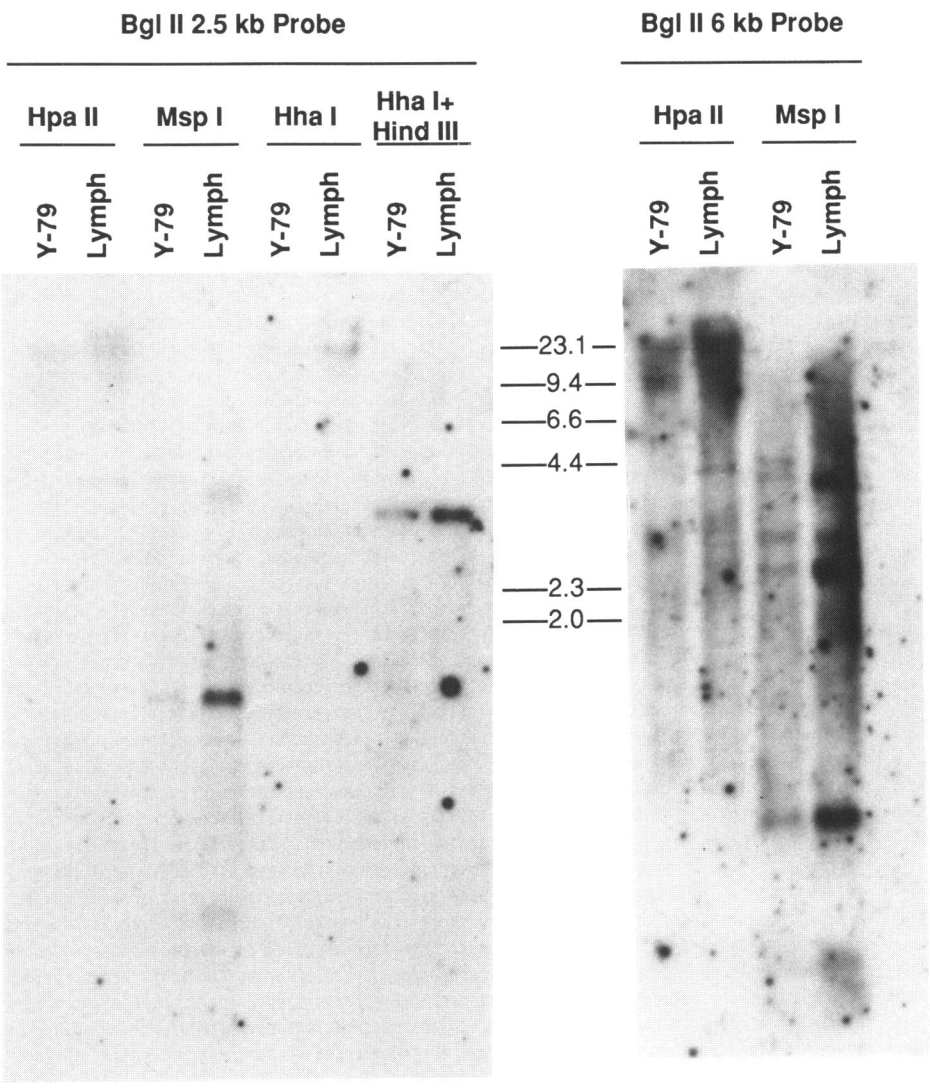

Figure 3. Southern blot analysis of genomic DNA from Y-79 cells grown in suspension (Y-79) and lymphocytes (Lymph). *Hpa II*- DNA was digested with *Hpa II* prior to electrophoresis. *Msp I*- DNA was digested with *Msp I* prior to electrophoresis. *Hha I*- DNA was digested with *Hha I* prior to electrophoresis. *Hha I+Hind III*- DNA was double digested with *Hha I* and *Hind III* prior to electrophoresis. Probes used were the *Bgl II* 2.5 kb, covering the most 3′ portions of the gene, or a *Bgl II* 6 kb fragment containing much of the first exon, the second and third exons and the first two introns (see figure 1) as indicated. Positions of *Hind III* digested lambda phage DNA are shown as size markers.

Sperm DNA) at 42°C (50% Formamide) or 68°C (without formamide). Filters were then washed twice in 2× SSC 0.1% SDS for 20 minutes at 22°C, and twice in 0.2× SSC, 0.1% SDS at 42°C for 20 minutes. Filters were then air dried and exposed to film. Total cellular RNA was isolated from these cells by the method of Chomczynsi and Sacchi (14). RNA was electrophoresed in 0.7% denaturing formaldehyde gels and transferred to nitrocellulose (13). Filters were probed under the same conditions as Southern blots.

RESULTS

A restriction map of the IRBP genomic clone is shown in figure 1. This figure shows the approximate locations of the exons, as well as the fragments used to probe Southern and northern blots. The isolated clones contain the entire translated sequence, as well as 7 kb of 5′ flanking sequence. Y-79 cells have been previously shown to synthesize IRBP in suspension culture, these cells have also been found to synthesize 3−5 fold more IRBP and IRBP mRNA when grown as monolayers attached to poly-D-lysine (7). Northern blots of RNA extracted from Y-79 cells and lymphocytes were probed with the 7 kb *Hind III* fragment containing the first exon of the gene (figure 2). Y-79 cells showed a single message at 4.4 kb. No IRBP mRNA was detected in the lymphocytic cells.

The probes to the 3′ domain (*Bgl II* 2.5 kb), the intronic region and the first exon (*Bgl II* 6 kb) gave distinct banding patterns in southern blots (figure 3) but did not show substantial differences between lymphocytes or Y-79 cells. No significant changes in banding patterns of *Hha I* digested DNA were observed using these probes (not shown). These data indicate that there are no alterations in methylation of these regions of the gene in non-expressing and expressing cells. The *Bgl II* 4.5 kb fragment encoding the most 5′ portions of the isolated clone,

Figure 4. Nucleotide sequence of the IRBP promoter region. The proposed 5′ mRNA cap site (8) is numbered +1 and marked with an arrow. CpG dinucleotide pairs are underlined, the protein coding region is in italics and marked with a dashed arrow. *Msp I/Hpa II* sites (CCGG) are indicated with *, *Hha I* sites (GCGC) with #.

as well as a *Sma I* fragment containing the 5′ 7 kb did not give a distinct banding patterns (data not shown), indicating the most 5′ portion of the clone probably contains a repetitive element (such as an alu repeat) present elsewhere in the genome.

We were most interested in the 5′ flanking 2−3 kb, which should contain the promoter region. We have subcloned a *Bam HI* fragment containing this region and sequenced 4 kb of the 3′ end of this fragment (figure 4). There are no TATA or CAATT boxes upstream from the mRNA cap site, and searches of the Genbank data base did not find any sequences similar to the IRBP promoter region. A repeat motif of $GT_{(4-7)}$ is present. A CpG rich island is found from −1578 to −1108. Numerous CpG dinucleotides (as well as *Msp I-Hpa II* and *Hha I* sites) are also found in the first exon. This clustering of CpG dinucleotides indicates this gene has 'HTF islands' which are associated with many mammalian genes (15). Although the CpG contents of the CpG rich-island (32% of expected) and the first exon (36% of expected) is lower than classical HTF islands, they are substantially higher than the sequences flanking the CpG island (both 12% of expected). When a *Hind III* probe to the promoter region and first exon was hybridized to southern blots, a 1.8 kb band was observed in all *Msp I* digests but only in the *Hpa II* digests from Y-79 cells, not from the lymphocytes (figure 5). This appeared to correspond to a 1.8 kb segment of the promoter region from −1578 to +248 as shown in figure 6. Digestion of the entire subcloned IRBP gene with *Msp I* yields only one 1.8 kb fragment. When this band was isolated and used as a probe, it confirmed that the 1.8 band in the *Hpa II* digests was from the promoter segment, rather than a combination band due to internally methylated sites (figure 7). Identical results were

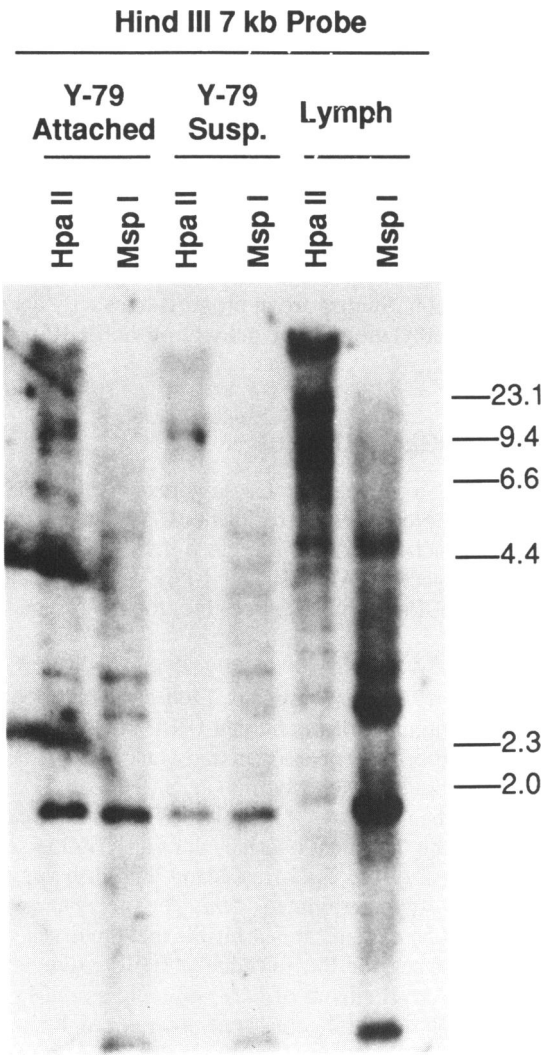

Figure 5. Southern blot analysis of Y-79 cell (Y-79) and lymphocyte (Lymph) genomic DNA using a 7 kb *Hind III* probe containing the first exon and 2.5 kb of 5′ flanking DNA. Y-79 Attached: Y-79 cells grown in attachment culture on poly-D-lysine, these cells synthesize approximately 3–5 fold more IRBP mRNA than Y-79 cells grown in suspension (7). Y-79 Susp: Y-79 cells grown in suspension culture. *Hpa II*, *Msp I*: DNA was digested with these enzymes prior to electrophoresis. Positions of *Hind III* digested lambda phage DNA are shown as size markers.

found with DNA isolated from Y-79 cells grown in suspension and those grown as monolayers attached to poly-D-lysine (figure 5) where IRBP mRNA expression is enhanced several fold (7).

These data identified two hypomethylation sites found in the promoter of the expressing Y-79 cells, one in the CpG rich island and one in the first exon, as indicated in figure 6. When genomic DNA was digested with *Hha I* and probed with the 7 kb *Hind III* fragment or the *Msp I*-1.8 kb fragment, a band of 2.1–2.3 kb was found only in the Y-79 cell DNA (figure 7). A very faint 1.6 kb band was also seen in some blots. Since digestion of the entire subcloned IRBP gene does not give a band in this region as predicted by the sequence of the promoter region, this 2.1–2.3 kb band must contain internally methylated *Hha I* sites. There was no change in the banding pattern when the DNA was digested with both *Hind III* and *Hha I* (figure 7), indicating the band was 3′ to the *Msp I* site at −1578 in the promoter (figure 6). It was difficult to determine the exact 3′ border of this band in the first exon due to the closeness of several *Hha I* sites in this region. The faint 1.6 kb band may be due to a minor demethylation of the *Hha I* site at −776, although this site appears to remain methylated even in the expressing Y-79 cells. Sequence analysis indicates that the *Hpa II*/*Msp I* site at −1578 and the *Hha I* site at −1284 are within an 'island' of CpG rich sequence. Our data indicate that there is a specific demethylation of the CpG sites in this 'island' in cells which express IRBP, as well as hypomethylation within the beginning of the first exon. Analysis of other portions of the gene found no other differences in the methylation of *Hpa II* or *Hha I* sites between expressing and non-expressing cells.

DISCUSSION

Previous studies on the bovine (16) and human (8) IRBP gene indicate an atypical gene structure. Only four exons encode the 145 kd protein with a large first exon containing most of the transcribed sequence. As seen in this study, the 5′ flanking region is also unusual in that it does not contain canonical TATA or CAATT boxes, and does not have similarity to tissue-specific promoters studied to date.

Our results show that Y-79 cells expressing IRBP are hypomethylated in a CpG rich island in the promoter region and in the beginning of the first exon of the IRBP gene, when compared to non-expressing normal human lymphocytes. It is interesting to note that the hypomethylation was confined to these

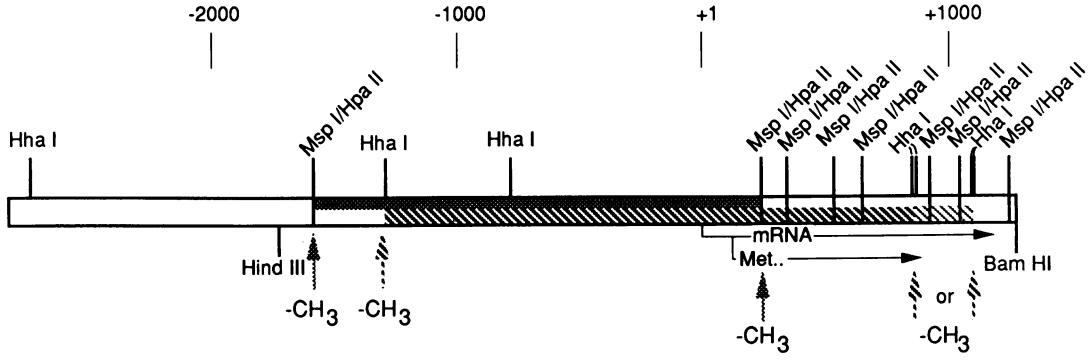

Figure 6. A model of the IRBP promoter region based on the sequence, restriction mapping and hybridization studies. +1: mRNA 5′ cap site as indicated by Liou et al (8). Shaded region: location of the 1.8 kb *Msp I*/*Hpa II* fragment which is demethylated in Y-79 cells as indicated by the shaded arrows. Stripped region: location of the 2.1–2.3 *Hha I* fragment which is present only in Y-79 cells, due to demethylation at the points indicated by the stripped arrows. This fragment contains an internal, methylated *Hha I* site even in the Y-79 cell genomic DNA.

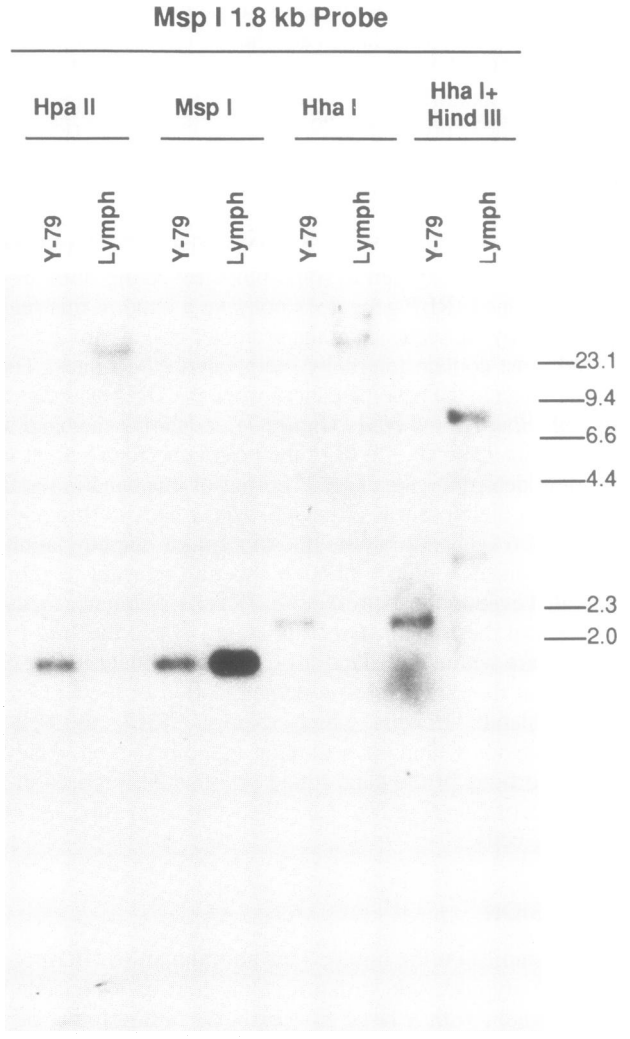

Figure 7. Southern blot analysis of genomic DNA from Y-79 cells grown in suspension (Y-79) and lymphocytes (Lymph) using the *Msp I* 1.8 kb probe containing immediate 5′ flanking DNA (see figures 1 and 5). Enzyme digests are as indicated in figure 3. Positions of *Hind III* digested lambda phage DNA are shown as size markers.

two areas; the *Hha I* site lying between the hypomethylated CpG rich island in the promoter and the hypomethylated portions of the first exon remained methylated even in the expressing cells. No other differences in methylation were found in the other regions of the gene. Early studies indicated methylation of promoter sequences was important in gene inactivation (reviewed in 15). Recent studies have found that alterations in methylation state are important in controlling the activity of various promoters (17, 18, 19, 20, 21, 22). Inhibition of promoter activity by methylation has been proposed as a reversible regulatory signal (23). *In vitro* methylation can substantially inhibit *in vivo* activation of promoter-reporter gene constructs (19, 20). These methylation patterns can be affected by cellular differentiation (21, 22). Control of gene expression by hypomethylation of promoter regions may explain the differentiative effects of methylation inhibiting agents. These studies indicate that methylation state should be considered an important factor in studying promoter activity in transient transfection systems; abnormally active promoters in non-expressing cells may be due to the hypomethylated state of bacterial-derived DNA. Our data are in agreement with methylation state playing an important role in IRBP gene expression. Clearly, however, other factors are involved in the regulation of the level of IRBP expression, as no differences in methylation patterns were noted between Y-79 cells grown in suspension and those grown on poly-D-lysine, where substantial differences in IRBP mRNA expression are found. We propose that hypomethylation of these critical regions is permissive for IRBP expression, perhaps by allowing specific nuclear factors to bind, but that external influences on the differentiation state of the cells can further enhance the levels of IRBP expression. Studies are in progress to exactly determine the effects of methylation on the activity of the IRBP promoter region *in vivo*.

ACKNOWLEDGEMENTS

We would like to thank Dr C. Jaworski for the database searching. D.M. Noonan is the recipient of a CNR research grant 'Ingegneria Genetica'. Zhen Zhu is a fellow of the World Laboratory, project ACAB (MCD 1 I-IV). We are grateful to Dr M.A. Gonda, Program Resources Inc, for DNA sequencing.

NOTE ADDED IN PROOF

After submission of this manuscript Liou et al reported that a 1.3 kb *Pvu II* fragment of the human IRBP promoter was able to confer tissue-specific expression in transgenic mice (Liou, G.I., Geng, L., Al-Ubaidi, M.R., Matragoon, S., Hanten, G., Baehr, W. and Overbeek, P.A. (1990) J. Biol. Chem., 265, 8373–8376). This fragment overlaps the transcription start site and contains most of the CpG rich island within the promoter, including the hypomethylated *Hha I* site but not the hypomethylated *Hpa II* site. It is probable that the tissue specific methylation patterns for the introduced IRBP promoter were restored during development of the transgenic animals, studies of the methylation state of the transgene promoter in these animals would be very informative.

REFERENCES

1. Inouye, L., Albini, A., Chader, G., Redmond, T. and Nickerson, J. (1989) Exp. Eye. Res., 49, 171–180.
2. van Veen, T., Katial, A., Shinohara, T., D.J., B., Wiggert, B., Chader, G.J. and Nickerson, J. (1986) FEBS letters., 208, 133–137.
3. Kyritsis, A.P., Tsokos, M., Triche, T.J. and Chader, G.J. (1984) Nature., 307, 471–473.
4. Fong, S., Balakier, H., Canton, M., Bridges, C.D.B. and Gallie, B. (1988) Cancer Res., 48, 1124–1128.
5. Rodrigues, M.M., Wiggert, B., Sheilds, J., Donoso, L., Bardenstein, D., Katz, N., Friendly, D. and Chader, G. (1987) Ophthalmology., 94, 378–387.
6. Kyritsis, A.P., Wiggert, B., Lee, L. and Chader, G.J. (1985) J. Cell Phys., 124, 233–239.
7. Albini, A., Toffenetti, J., Noonan, D. and Chader, G. (1990) Inves. Ophthal. Vis. Sci., 31(suppl.), 597.
8. Liou, G., Ma, D., Yang, Y., Geng, L., Zhu, C. and Baehr, W. (1989) J. Biol. Chem., 264, 8200–8206.
9. Liou, G., Fong, S., Gosden, J., Tuinen, P.v., Ledbetter, D., Christie, S., Rout, D., Bhattacharya, S., Cook, R., Li, Y., Wang, C. and Bridges, C. (1987) Somatic Cell Mol. Gen., 13, 315–323.
10. Fong, S., Cook, R., Alvarez, R., Liou, G., Landers, R. and Bridges, C. (1986) FEBS Lett., 205, 309–312.
11. Lathe, R. (1985) J. Mol. Biol., 183, 142–149.
12. Sanger, F., Coulson, A.R., Barell, B.G., Smith, A.J.A. and Roe, B.A. (1980) J. Mol. Biol., 143, 161–178.
13. Maniatis, T., Fritsch, E.F. and Sambrook, J. (1982) Molecular Cloning: A laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
14. Chomczynsi, P. and Sacchi, N. (1987) Anal. Biochem., 162, 156–159.

15. Bird, A.P. (1986) Nature., 321, 209–213.
16. Borst, D.E., Redmond, T.M., Elser, J.E., Gonda, M.A., Wiggert, B., Chader, G.J. and Nickerson, J.M. (1989) J. Biol. Chem., 264, 1115–1123.
17. Doerfler, W., Hoeveler, A., Weisshaar, B., Dobrzanski, P., Knebel, D., Langner, K., Achten, S. and Muller, U. (1989) Cell Biophys., 15, 21–7.
18. Toth, M., Lichtenberg, U. and Doerfler, W. (1989) Proc. Natl. Acad. Sci. U.S.A., 86, 3728–3732.
19. Rachal, M., Yoo, H., Becker, F. and Lapeyre, J. (1989) Nucleic Acids Res., 17, 5135–5147.
20. Gaido, M. and Strobl, J. (1989) Biochim. Biophys. Acta., 1008, 234–242.
21. Wotton, D., Flanagan, B. and Owen, M. (1989) Proc. Natl. Acad. Sci. U.S.A., 86, 4195–4199.
22. Levy-Wilson, B. and Fortier, C. (1989) J. Biol. Chem., 264, 9891–9896.
23. Doerfler, W., Weisshaar, B., Hoeveler, A., Knebel, D., Muller, U., Dobrzanski, P., Lichtenberg, U., Achten, S. and Hermann, R. (1988) Gene., 74, 129–133.