GENOMICS 18, 216–222 (1993)

Molecular Characterization of the Murine Neural Retina Leucine Zipper Gene, Nrl

QAIS FARJO,* ANNE U. JACKSON,† JUNZHE XU,* MICHELE GRYZENIA,* CRAIG SKOLNICK,* NEERAJ AGARWAL,‡ AND ANAND SWAROOP*,†,[1]

*Departments of *Ophthalmology and †Human Genetics, Kellogg Eye Center, University of Michigan, Ann Arbor, Michigan 48105; and ‡Department of Anatomy and Cell Biology and North Texas Eye Research Institute, Texas College of Osteopathic Medicine, Fort Worth, Texas 76107*

Received May 3, 1993; revised July 22, 1993

The NRL gene (D14S46E) is expressed in cells of human retina and encodes a putative DNA-binding protein of the leucine zipper family. Here we describe the analysis of the murine homolog of the NRL gene, Nrl. Various cDNAs resulting from alternate polyadenylation are characterized. The deduced polypeptide sequence is highly conserved between mouse and human, with an identical basic motif and leucine zipper domain. The nucleotide sequences in the 5′ and 3′-untranslated regions also show significant homology. The 3′-untranslated region contains a polymorphic AGG-trinucleotide repeat. The murine Nrl gene consists of three exons; of these, the first is untranslated. The 5′-upstream promoter region has no canonical TATA box, but contains consensus binding site sequences for several DNA-binding proteins. Analysis of RNA from adult mouse tissues confirms the retina-specific expression of Nrl. This study provides the basis for dissecting the *cis*-regulatory elements involved in the retina-specific expression and for the development of an experimental model to investigate the function or any diseases associated with this gene in humans.

INTRODUCTION

In mammalian systems, the study of tissue- or cell-type-specific genes has led to the identification of various classes of transcription factors. DNA-binding proteins are involved in transcription regulation and are believed to play a major role in cell proliferation and differentiation. The combinatorial and synergistic action of these proteins in distinct spatial and temporal patterns ensures the precise regulation of gene expression required for generating complex systems, particularly in producing neuronal diversity and complexity (Struhl, 1991; He and Rosenfeld, 1991). Based on DNA-binding motifs, at least 10 distinct classes of transcription factors have been defined (Mitchell and Tjian, 1989; Latchman, 1990). The members of these classes contain multiple functional domains that include a conserved region required for DNA binding. The proteins of the "leucine zipper" family influence the transcription of genes by interacting with DNA as homo- or heterodimers (Landschulz *et al.*, 1988; Kouzarides and Ziff, 1988). Transcription factors belonging to this class include AP1 and ATF/CREB proteins (Mitchell and Tjian, 1989; Hai *et al.*, 1989; Latchman, 1990).

Subtraction cDNA cloning is a useful technique for isolating tissue-specific genes. Using a biotin-based subtraction procedure, we enriched a human adult retina cDNA library by removing most of the constitutively expressed sequences (Swaroop *et al.*, 1991). Analysis of randomly isolated clones from the subtracted retinal library identified a cDNA for NRL (*n*eural *r*etina *l*eucine zipper), a retina-specific gene encoding a putative DNA-binding protein with a basic motif and a leucine zipper domain (Swaroop *et al.*, 1992). It was also demonstrated by *in situ* hybridization to adult baboon retinal sections that the gene is expressed in outer, inner, and ganglion cell layers of neural retina. The NRL gene (D14S46E) maps to human chromosome 14q11.1–q11.2 by somatic cell hybrid analysis and *in situ* hybridization to chromosome spreads (Yang-Feng and Swaroop, 1992) and may be a candidate for human retinopathies (Jackson *et al.*, 1993). The strong similarity of human NRL gene product with that of a v-*maf* transforming oncogene from an avian retrovirus (Nishizawa *et al.*, 1989), its evolutionary conservation, and tissue-specific expression suggest a significant function for NRL in retina. To understand the molecular mechanisms underlying the specificity of NRL expression and to define its function in the development of the visual pathway, we isolated the murine homolog (Nrl) of the human gene. In this report, we describe the sequence of murine Nrl cDNAs, the structural organization of the corresponding gene and promoter region, and its expression in adult tissues.

[1] To whom correspondence should be addressed at University of Michigan, 540 Kellogg Eye Center, 1000 Wall Street, Ann Arbor, MI 48105. Telephone: (313) 936-9547. Fax: (313) 936-2340.

0888-7543/93 $6.00

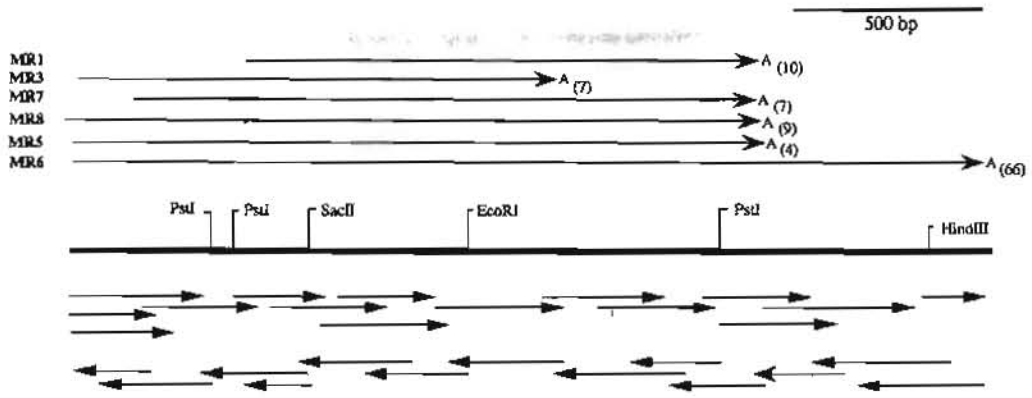

FIG. 1. Murine Nrl cDNAs and sequencing strategy. Thick bar with restriction enzyme sites represents the longest cDNA, MR6. Approximate size and location of additional cDNAs are indicated above the bar, with the number of (A) residues at the 3'-end in parentheses. The length and direction of arrows below the central bar show the extent and orientation of the nucleotide sequence obtained with a particular primer.

MATERIALS AND METHODS

Retinal RNA was prepared from 1-month-old Balb/c mice, as described (Agarwal *et al.*, 1990). Total RNAs from other adult tissues (obtained from C57/BL6 mice) were kindly provided by Dr. Miriam Meisler (University of Michigan). Poly(A)$^+$ RNA from retina was used to construct an oligo(dT)primed directional cDNA library in Charon BS(−) phage vector (Swaroop and Weissman, 1988) by the procedure described (Swaroop, 1993). A mouse genomic library in pWE15 cosmid vector was purchased from Stratagene (La Jolla, CA). Methods for screening of cDNA and genomic libraries, preparation of phage, cosmid, or plasmid DNA, restriction enzyme analysis, and Southern and Northern analyses were performed essentially as described (Swaroop *et al.*, 1988; Sambrook *et al.*, 1989). DNA sequencing was performed by the dideoxy chain termination method using double-stranded plasmid DNA (Zagursky *et al.*, 1985) with Sequenase enzyme (U.S. Biochemicals, Cleveland, OH). DNA sequences were analyzed with MacVector (IBI-Kodak, New Haven, CT), DNASTAR (Madison, WI), and GCG (Madison, WI) programs. RNA–PCR analysis of total RNA from mouse tissues was performed with kits from Perkin-Elmer–Cetus (Norwalk, CT) or Invitrogen (San Diego, CA) using the manufacturer's protocols. All commonly used enzymes and reagents were obtained from New England Biolabs (Beverly, MA), GIBCO-BRL (Gaithersburg, MD), Stratagene, or Boehringer-Mannheim (Indianapolis, IN).

RESULTS

Isolation and Analysis of cDNA Clones

The murine cDNA library was screened with the insert from human NRL cDNA, AS321 (Swaroop *et al.*, 1992), which was labeled with ^{32}P using random primers (Sambrook *et al.*, 1989). Six positive phage clones were plaque-purified and subsequently transferred to Bluescript KS(−) plasmid by *Not*I digestion of phage DNA followed by recircularization (Swaroop and Weissman, 1988). Figure 1 shows the restriction map and sequencing strategy for murine cDNAs. The lengths and positions of various clones are indicated. The major transcript observed in mouse retina is about 2 kb (Swaroop *et al.*, 1992). The sizes and polyadenylation sites of cDNAs MR1, MR5, MR7, and MR8 are consistent with the size of this transcript. MR6 cDNA represents only a minor species (<10%). The complete sequence of the longest 2.5-kb MR6 cDNA, obtained from both strands, is shown in Figure 2. The sequences of other cDNAs are indicated relative to this sequence (see Fig. 2 legend). All cDNAs have a poly(A) tail at the 3'-end and are generated by polyadenylation at alternative sites. An imperfect yet functionally active signal ATTAAA (Wickens, 1990), also observed in the human NRL cDNA, is present upstream of the polyadenylation site in MR1, MR5, MR7, and MR8. The cDNA clone, MR3, results from polyadenylation at a site preceded by an AATTAA upstream sequence, whereas polyadenylation in MR6 cDNA follows the sequence TATAAA.

The sequence of murine Nrl cDNAs reveals an open reading frame of 237 amino acids, as in the human NRL. The first methionine codon exists in the consensus translation initiation sequence (Kozak, 1987). The derived polypeptide contains consensus sites for a number of protein kinases, including cAMP-dependent protein kinase and protein kinase C (Kemp and Pearson, 1990). The murine sequence shows high homology to human NRL (about 90% in the coding region). Comparison of nucleotide sequence of MR6 cDNA to human DD6 clone (a 2-kb NRL cDNA with longer 3'-untranslated region, generated by alternative polyadenylation) demonstrates significant homology even in the 5'- and 3'-untranslated regions (data not shown; GenBank Accession No. for DD6 cDNA is M95925). As expected, the murine Nrl sequence is also homologous to the v-*maf* oncogene and its product (Nishizawa *et al.*, 1989). Figure 3 shows the comparison of polypeptide sequences derived from human and mouse NRL and v-*maf* cDNAs. Complete identity between the Nrl polypeptides is observed in the predicted basic motif and leucine zipper domain. A remarkable similarity is also seen in the NH_2-terminal region, with strong conservation of putative proline- and serine-rich domains. This region is also highly homologous to the corresponding sequence in the v-*maf* oncogene product.

The 3'-untranslated region of the cDNA sequence contains a trinucleotide sequence, $(AGG)_{13}$, that can be further extended by two G → A substitutions and one nucleotide (A) deletion. This sequence shows polymorphic variation among strains of mice and has been used

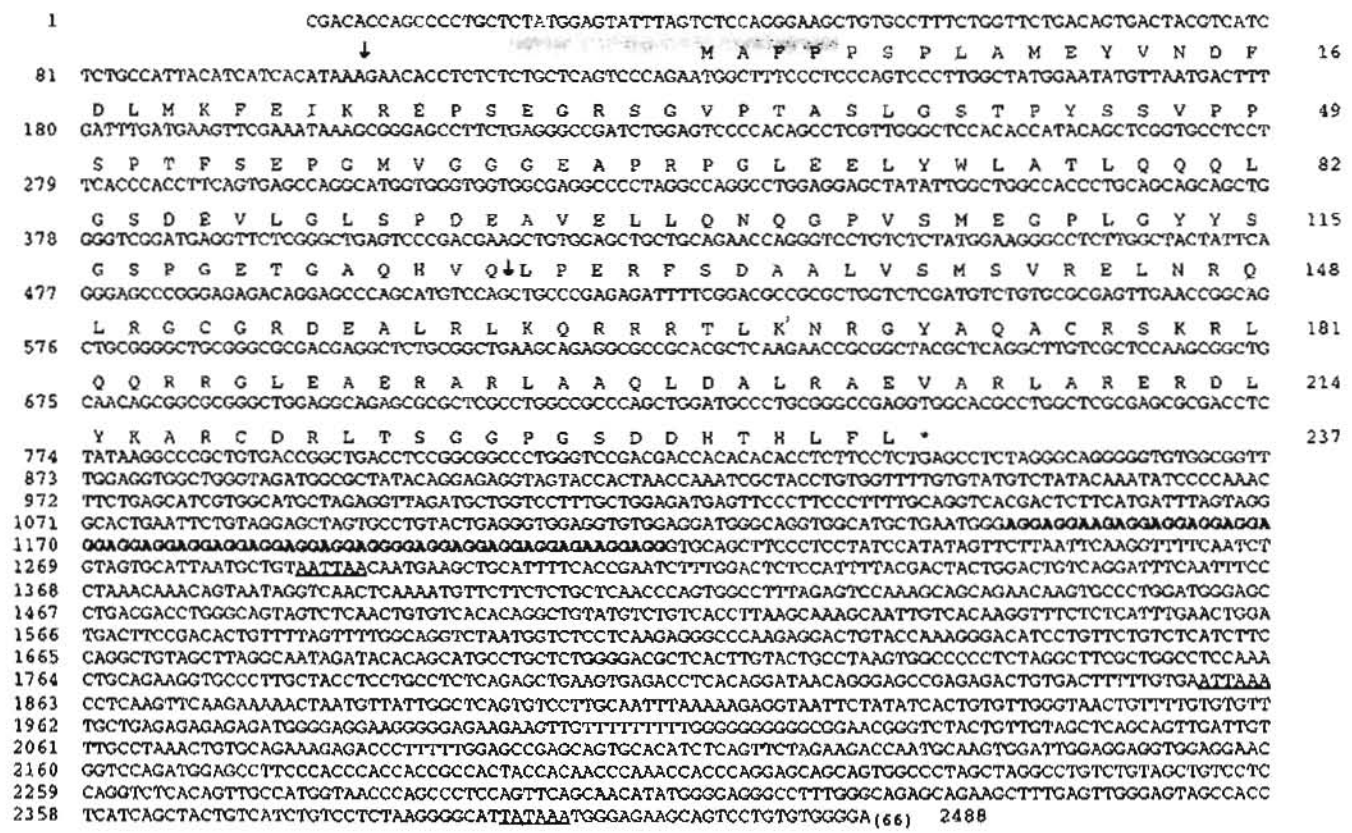

FIG. 2. Complete nucleotide sequence of the MR6 cDNA and derived polypeptide. The numbers on the right correspond to the amino acids. Vertical arrows indicate the position of introns. The trinucleotide repeat in the 3'-untranslated region is represented in bold. MR3 has two additional nucleotides at the 5'-end, whereas MR7 begins at position 152. MR5 and MR8 contain 10 or 15 additional 5'-nucleotides, respectively, that are not present in the genomic sequence and are perhaps artifacts of cDNA cloning. The positions of polyadenylation site in the cDNAs are MR3, bp 1314; MR7, bp 1872; MR1 and MR8, bp 1875; MR5, bp 1882. Polyadenylation signal sequences are underlined. The sequence has been deposited in the GenBank Database under Accession No. L14935.

to localize the Nrl gene to mouse chromosome 14 (Bespalova *et al.*, 1993).

Murine Nrl Gene

A 1.1-kb *Eco*RI insert from MR5 cDNA (including the coding region) was used to screen a mouse cosmid library. One of the positive clones (Mcos 2C) with a 35-kb genomic insert was characterized further. The restriction map of the cosmid clone was generated by partial and complete restriction enzyme digestions and by probing the Southern blots of digests with T3 and T7 primer sequences flanking the genomic insert (Sambrook *et al.*, 1989). Hybridization of Southern blots to various parts of Nrl cDNA also provided information on the structure of the gene. Figure 4 shows the restriction map of Mcos2C and the organization of Nrl gene. Three exons, including a 5'-untranslated exon, constitute the mature Nrl mRNA. The exon–intron boundaries were obtained by sequence analysis using appropriate subclones and oligonucleotide primers (Fig. 5). The sequence at the exon–intron junctions matched the published consensus sequences (Mount, 1982). In addition to the GT and AG nucleotides at the donor and accepter splice sites, the adjacent sequences followed the consensus pattern. The complete sequence of the exons and flanking region was determined and confirmed by comparison to the cDNA sequence (see Fig. 2).

To delineate the regulatory and/or promoter elements, we also obtained the sequence in the 5'-upstream region of the Nrl gene (Fig. 6). Based on primer-extension analysis using human retinal RNA and comparison of human and murine sequences (data not shown), we assigned the transcription initiation site 26 bp upstream of the first basepair of the cDNA sequence. This site is identical to the consensus CAP sequence CANYY (Bucher and Trifonov, 1986). Although a canonical TATA -element is not present, the binding site sequences for a number of DNA-binding proteins are observed in the immediate upstream region. These include a CCAAT box, an octamer motif, an AP2 binding site, two E box elements (helix–loop–helix binding proteins), and sites for AP1 or a related transcription complex (Mitchell and Tjian, 1989; Blackwell and Weintraub, 1990). This suggests a complex interaction of proteins for regulating the development and tissue-specific expression of the Nrl gene.

Expression Analysis

Northern analysis of RNA from various human tissues and cell lines demonstrated that the NRL gene is

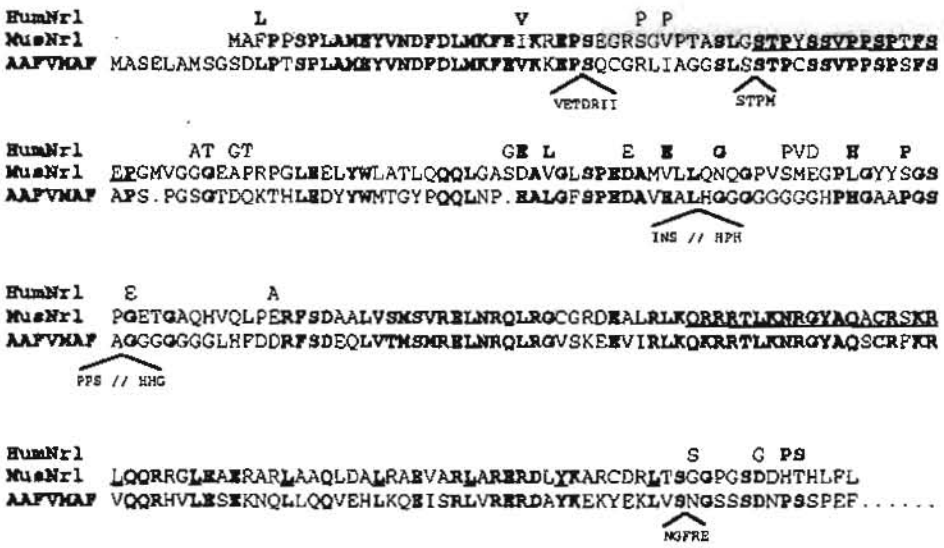

FIG. 3. Comparison of deduced polypeptide sequences from human and mouse Nrl (MusNrl and HumNrl, respectively) and from v-*maf* oncogene (AAFVMAF). Amino acid residues in bold indicate identity. The proline-serine-rich domain in NH_2-terminal region and amino acids in the basic motif and the leucine zipper domain are underlined. Additional residues in v-*maf* gene product are indicated by insertions below the sequence.

expressed only in retinal tissue and retinoblastoma cell lines, and its expression was detected in all neuronal layers by *in situ* hybridization to baboon retinal sections (Swaroop *et al.*, 1992). The gene was shown to be conserved during vertebrate evolution, and expressed in bovine, mouse, and rat retina. To confirm the specificity of expression, we performed Northern analysis of total RNA from adult C57/BL6 mice tissues (liver, brain, kidney, lung, pancreas, spleen, ovary/oviduct, parotid gland, submaxillary gland, and mammary gland). No transcript was detected using the MR5 cDNA as a probe (data not shown). Furthermore, RNA–PCR analysis using total RNA from these tissues and retina identified an amplified Nrl product only in retina (data not shown). Northern analysis of RNA by Dr. Kirk Beisel (Boys Town National Research Hospital, Omaha, NE) also showed no expression of Nrl in other mouse tissues, including heart, testes, and cochlea (personal comm., April 30, 1992). Nrl expression in retina (as detected by Northern analysis) was not altered in response to different lengths of exposure of mice to a light–dark cycle (data not shown).

DISCUSSION

The molecular hierarchy of transcription factors involved in the development of retinal neuronal cells is currently not understood. Recently, while analyzing

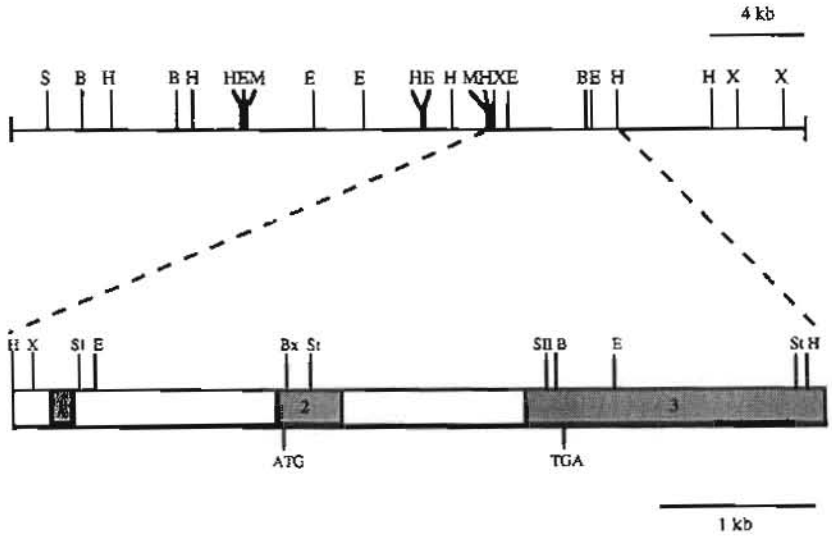

FIG. 4. Restriction map of Mcos2C cosmid showing the exon–intron structure of the murine Nrl gene. Appropriate scales are indicated. Exons are represented by shaded areas. Positions of initiation and termination codons are identified. Abbreviations for restriction enzymes are B, *Bss*HII; Bx, *Bst*XI; E, *Eco*RI; H, *Hin*dIII; M, *Mlu*I; S, *Sal*I; SI, *Sac*I; SII, *Sac*II; St, *Stu*I; X, *Xho*I.

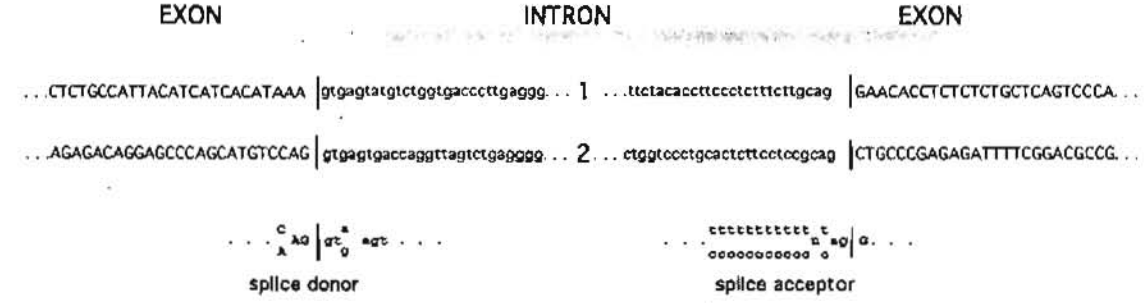

FIG. 5. Sequences at the exon–intron boundaries in the murine Nrl gene. Appropriate splice site signals are present. Junction sequences match with the consensus splice donor and acceptor sequences.

random clones from a subtracted human retina cDNA library, we identified a gene, named NRL, that shows a retina-specific pattern of expression (Swaroop *et al.*, 1992). NRL expression was detected in all cellular layers of neural retina. Analysis of NRL gene product identified a leucine zipper domain, preceded by a region of basic amino acids, and revealed a strong homology to the product of an avian retroviral transforming oncogene, v-*maf* (Nishizawa *et al.*, 1989). These results, combined with the evolutionary conservation of its gene, suggested that NRL gene product may function as a DNA-binding protein with a significant role in retinal development. Since NRL transcript is specifically detected in retina, the molecular events regulating its expression may provide significant insights into the differentiation of retinal neurons. Genetic manipulation of the mouse embryo has recently permitted molecular studies on the regulation of neuronal development (Rossant, 1990). Therefore, to elucidate the regulation of NRL gene expression and the function of its product, we have isolated and characterized murine Nrl cDNAs and the corresponding gene.

Earlier, we observed that the major Nrl transcript in mouse retina is 2 kb, compared to 1.3 kb in human (Swaroop *et al.*, 1992). The structural analysis of murine cDNAs reveals that the 2-kb transcript is the result of polyadenylation at a later site, generating a larger 3′-untranslated region. As in human NRL, this major polyadenylation site follows an ATTAAA signal. An AATTAA sequence upstream of the polyadenylation site in the 1.3-kb MR3 cDNA is probably responsible for inefficient polyadenylation (Wickens, 1990). The next available polyadenylation signal, ATTAAA, is therefore utilized in the murine Nrl gene. The longer 2.5-kb cDNA, MR6, may be the result of leakiness of the transcription termination machinery and utilizes an upstream TATAAA sequence (Wickens, 1990). Variations in the lengths of the 3′-untranslated regions may cause differences in the stability or translation efficiency of Nrl mRNAs (see Jackson and Standart, 1990). It will be of interest to determine whether these species are differentially expressed in retinal cells.

The human and murine NRL gene products are of the same size (237 amino acids) and show a high degree of conservation. Particularly remarkable is the complete identity of the putative DNA-binding domain (basic motif and leucine zipper). The proline-, serine-, and threonine-rich amino-terminal region of the polypeptide also shows strong homology (51/56 amino acids). This region is similar to the transcription activation domain present in jun proteins (Lewin, 1991). It should be noted that the homology of Nrl gene product to that of v-*maf* extends beyond the putative DNA-binding region and is also observed in the NH_2-terminal activation domain. This suggests functional similarity of Nrl protein to that of the transforming v-*maf* oncoprotein.

An intriguing observation from sequence analysis is the presence of a long AGG-trinucleotide repeat in the 3′-untranslated region of the murine Nrl mRNA. A rudimentary part of this repeat is also seen in the human NRL (GGGAGGTGGAGGAGG) (Swaroop *et al.*, 1992). Alterations in the number of trinucleotide repeats in sev-

FIG. 6. The immediate 5′-upstream sequence (putative promoter region) of the murine Nrl gene. The asterisk shows the first basepair of the cDNA sequence (see Fig. 2). +1 indicates the site of transcription initiation, based on primer extension results with the human NRL gene and comparisons between the human and murine sequences (data not shown). Positions of consensus DNA sequence elements recognized by various transcription factors are appropriately identified.

eral human genes are associated with inherited diseases, particularly those involving neuronal dysfunction (Richards and Sutherland, 1992; Suthers *et al.*, 1992; Huntington's Disease Collaborative Research Group, 1993). The sequence polymorphism of AGG repeat, observed in mouse strains, has been used to localize the murine Nrl gene to chromosome 14, in a region homologous to human chromosome 14 (Bespalova *et al.*, 1993). The possibility of an unstable Nrl genomic region because of alterations (particularly expansion) in the trinucleotide repeat may be explored further. It will be of interest to determine whether mutations in this gene are associated with functional defect in retinal neurons.

The mature Nrl transcript is derived from three exons that span a genomic region of about 6 kb. The sequence at the exon–intron junctions in the Nrl gene conforms to the consensus splice sites. The organization of gene in mice and humans is almost identical (Jackson and Swaroop, unpublished data). The first AUG codon that initiates an open reading frame of 237 amino acids is detected in the second exon. The first exon, therefore, is untranslated. A remarkable sequence conservation is observed between the murine and human 5'- and 3'-untranslated regions. These regions are implicated in fine-tuning a variety of functions, which include translation efficiency of mRNA (Grens and Scheffler, 1990), mRNA degradation (Fen and Daniel, 1991), and stimulation of transforming potential (Gishizky *et al.*, 1991).

The presence of binding sites for several transcription factors in the promoter region suggests that the Nrl gene is under stringent regulation. The topology, affinity, concentration, and interaction of various DNA-binding proteins probably determine the exclusive expression of the gene in adult neural retina (see Mitchell and Tjian, 1989; Kemler *et al.*, 1991). CCAAT-enhancer binding protein (C/EBP), octamer binding transcription factors (OTFs), and AP2 have been implicated in morphogenesis and differentiation (Okamoto *et al.*, 1990; Umek *et al.*, 1991; Mitchell *et al.*, 1991). The presence of two E-box sequences (CANNTG) for binding to helix–loop–helix proteins and two potential target sites (TGATCTCA and TGAGTCC) for AP1 or related transcription factors suggest further selective regulation. Interestingly, we can also recognize a sequence element, CC(AT)$_5$GG, that is reminiscent of the binding site consensus sequence for mammalian serum response factor (Treisman, 1986). It should be mentioned that the expression of NRL gene in human Y79 retinoblastoma cell line responds to serum (A. Swaroop, unpublished data). Additional expression specificity may also be imparted by DNA sequence elements that bind to specific retinal proteins.

Analysis of RNA from a number of adult mouse tissues confirms the specificity of Nrl expression in retina. Studies are in progress to determine its expression during embryonic development and during differentiation of retinal neurons. The isolation and characterization of cDNAs and the gene for mouse Nrl, described here, now allow further investigations of the regulation of its expression by transgenic studies. They also form the basis for understanding the physiological function of NRL and for developing an experimental model for any inherited retinopathy caused by altered function of the human gene.

ACKNOWLEDGMENTS

We thank Drs. Miriam Meisler and Linda Samuelson for providing total RNA from adult C57/BL6 mouse tissues, Dr. Peter Hitchcock for many helpful discussions, and Ms. Dorothy Giebel for typing the manuscript. Oligonucleotides used in the study were synthesized at the U-M Medical School core facility. This research was supported in part by grants from the George Gund and Retinitis Pigmentosa Foundations (to A.S.), a predoctoral fellowship from National Science Foundation (to A.J.), and summer research fellowships from U-M Biomedical Research Program (to Q.F. and M.G.). We also acknowledge NIH Grants EY07003 (CORE), M01-RR00042 (General Clinical Research Center), and P30 HG00209 (Human Genome Center).

REFERENCES

Agarwal, N., Nir, I., and Papermaster, D. S. (1990). Opsin synthesis and mRNA levels in dystrophic retinas devoid of outer segments in retinal degeneration slow (rds) mice. *J. Neurosci.* **10**: 3275–3285.

Bespalova, I. N., Farjo, Q., Mortlock, D., Jackson, A., Meisler, M. H., Swaroop, A., and Burmeister, M. (1993). Mapping of the neural retina leucine zipper (Nrl) gene to mouse chromosome 14. *Mamm. Genome,* in press.

Blackwell, T. K., and Weintraub, H. (1990). Differences and similarities in DNA-binding preferences of MyoD and E2A protein complexes revealed by binding site selection. *Science* **250**: 1104–1110.

Bucher, P., and Trifonov, E. N. (1986). Compilation and analysis of eukaryotic POL II promoter sequences. *Nucleic Acids Res.* **14**: 10009–10026.

Fen, Z., and Daniel, T. O. (1991). 5' untranslated sequences determine degradative pathway for alternate PDGF B/c-*sis* mRNA's. *Oncogene* **6**: 953–959.

Gishizky, M. L., McLaughlin, J., Pendergast, A. M., and Witte, O. N. (1991). The 5' non-coding region of the BCR/ABL oncogene augments its ability to stimulate the growth of immature lymphoid cells. *Oncogene* **6**: 1299–1306.

Grens, A., and Scheffler, I. E. (1990). The 5'- and 3'-untranslated regions of ornithine decarboxylase mRNA affect the translational efficiency. *J. Biol. Chem.* **265**: 11810–11816.

Hai, T., Liu, F., Coukos, W. J., and Green, M. R. (1989). Transcription factor ATF cDNA clones: An extensive family of leucine zipper proteins able to selectively form DNA-binding heterodimers. *Genes Dev.* **3**: 2083–2090.

He, X., and Rosenfeld, M. G. (1991). Mechanisms of complex transcriptional regulation: Implications for brain development. *Neuron* **7**: 183–196.

Huntington's Disease Collaborative Research Group. (1993). A novel gene containing a trinucleotide repeat that is expanded and unstable on Huntington's Disease chromosomes. *Cell* **72**: 971–983.

Jackson, A., Zheng, K., Yang-Feng, T. L., and Swaroop, A. (1993). Identification of candidate genes for eye diseases: Studies on a neural retina-specific gene encoding a putative DNA binding protein of leucine zipper family. *In* "Retinal Degeneration: Clinical and Laboratory Applications" (J. G. Hollyfield, Ed.), Plenum, New York, in press.

Jackson, R. J., and Standart, N. (1990). Do the poly(A) tail and 3' untranslated region control mRNA translation? *Cell* **62**: 15–24.

Kemler, I., Bucher, E., Seipel, K., Muller-Immergluck, M. M., and Schaffner, W. (1991). Promoters with the octamer DNA motif (ATGCAAAT) can be ubiquitous or cell type-specific depending on

binding affinity of the octamer site and Oct-factor concentration. *Nucleic Acids Res.* **19**: 237–242.

Kemp, B. E., and Pearson, R. B. (1990). Protein kinase recognition sequence motifs. *Trends Biochem. Sci.* **15**: 342–346.

Kouzarides, T., and Ziff, E. (1988). The role of the leucine zipper in the fos-jun interaction. *Nature* **336**: 646–651.

Kozak, M. (1987). An analysis of 5'-noncoding sequences from 699 vertebrate messenger RNAs. *Nucleic Acids Res.* **15**: 8125–8148.

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.

Latchman, D. S. (1990). Eukaryotic transcription factors. *Biochem. J.* **270**: 281–289.

Lewin, B. (1991). Oncogenic conversion by regulatory changes in transcription factors. *Cell* **64**: 303–312.

Mitchell, P. J., and Tjian, R. (1989). Transcriptional regulation in mammalian cells by sequence-specific DNA binding proteins. *Science* **245**: 371–378.

Mitchell, P. J., Timmons, P. M., Hebert, J. M., Rigby, P. W. J., and Tjian, R. (1991). Transcription factor AP-2 is expressed in neural crest cell lineages during mouse embryogenesis. *Genes Dev.* **5**: 105–119.

Mount, S. M. (1982). A catalogue of splice junction sequences. *Nucleic Acids Res.* **10**: 459–472.

Nishizawa, M., Kataoka, K., Goto, N., Fujiwara, K. T., and Kawai, S. (1989). v-*maf*, a viral oncogene that encodes a "leucine zipper" motif. *Proc. Natl. Acad. Sci. USA* **86**: 7711–7715.

Okamoto, K., Okazawa, H., Okuda, A., Sakai, M., Muramatsu, M., and Hamada, H. (1990). A novel octamer binding transcription factor is differentially expressed in mouse embryonic cells. *Cell* **60**: 461–472.

Richards, R. I., and Sutherland G. R. (1992). Heritable unstable DNA sequences. *Nature Genet.* **1**: 7–9.

Rossant, J. (1990). Manipulating the mouse genome: Implications for neurobiology. *Neuron* **2**: 323–334.

Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989). "Molecular Cloning: A Laboratory Manual," 2nd ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Struhl, K. (1991). Mechanisms for diversity in gene expression patterns. *Neuron* **7**: 177–181.

Suthers, G. K., Husan, S. M., and Davies, K. E. (1992). Instability versus predictability: The molecular diagnosis of myotonic dystrophy. *J. Med. Genet.* **29**: 761–765.

Swaroop, A. (1993). Construction of directional cDNA libraries from human retinal tissue/cells and their enrichment for specific genes using an efficient subtraction procedure. *In* "Methods in Neurosciences" (P. Hargrave, Ed.), Vol. 15, pp. 285–300, Academic Press, San Diego.

Swaroop, A., Hogan, B. L. M., and Francke, U. (1988). Molecular analysis of the cDNA for human SPARC/Osteonectin/BM-40: Sequence, expression, and localization of the gene to chromosome 6q31–q33. *Genomics* **2**: 37–47.

Swaroop, A., and Weissman, S. M. (1986). Charon BS (+) and (−), versatile lambda phage vectors for constructing directional cDNA libraries and their efficient transfer to plasmids. *Nucleic Acids Res.* **16**: 8739.

Swaroop, A., Xu, J., Agarwal, N., and Weissman, S. M. (1991). A simple and efficient cDNA library subtraction procedure: Isolation of human retina-specific cDNA clones. *Nucleic Acids Res.* **19**: 1954.

Swaroop, A., Xu, J., Pawar, H., Jackson, A., Skolnick, C., and Agarwal, N. (1992). A conserved retina specific gene encodes a basic motif/leucine zipper domain. *Proc. Natl. Acad. Sci. USA* **89**: 266–270.

Treisman, R. (1986). Identification of a protein-binding site that mediates transcriptional response of the c-*fos* gene to serum factors. *Cell* **46**: 567–574.

Umek, R. M., Friedman, A. D., and McKnight, S. L. (1991). CCAAT-enhancer binding protein: A component of a differentiation switch. *Science* **251**: 288–292.

Wickens, M. (1990). How the messenger got its tail: Addition of poly(A) in the nucleus. *Trends Biochem. Sci.* **15**: 277–281.

Yang-Feng, T. L., and Swaroop, A. (1992). Neural retina-specific leucine zipper gene NRL (D14S46E) maps to human chromosome 14q11.1–q11.2. *Genomics* **14**: 491–492.

Zagursky, R. J., Baumeister, K., Lomax, N., and Berman, M. L. (1985). Rapid and easy sequencing of large linear double-stranded DNA and supercoiled plasmid DNA. *Gene Anal. Techn.* **2**: 89–94.