Biochemical and Biophysical Research Communications **260,** 174–180 (1999)
Article ID bbrc.1999.0753, available online at http://www.idealibrary.com on IDEAL®

Cloning and Characterization of a Novel Orphan G-Protein-Coupled Receptor Localized to Human Chromosome 2p16

Emma E. Tarttelin,* Lawrence S. Kirschner,‡ James Bellingham,* Judit Baffi,§ Susan E. Taymans,‡ Kevin Gregory-Evans,† Karl Csaky,§ Constantine A. Stratakis,‡ and Cheryl Y. Gregory-Evans*,[1]

**Section of Molecular Genetics, Division of Biomedical Sciences, Imperial College School of Medicine, London, United Kingdom SW7 2AZ; †Academic Unit of Ophthalmology, Imperial College School of Medicine, Western Eye Hospital, London, United Kingdom SW7 2AZ; ‡Unit on Genetics and Endocrinology, National Institute of Child Health and Human Development, National Institutes of Health, Bethesda, Maryland 20892; and §Section on Gene Therapy, Laboratory of Immunology, National Eye Institute, National Institutes of Health, Bethesda, Maryland 20892*

Received April 1, 1999

We report the identification and characterisation of a novel human orphan G-protein-coupled receptor (GPR) which maps to chromosome 2p16. We have determined the full-length coding sequence and genomic structure of a gene corresponding to the anonymous expressed sequenced tag, WI-31133. This gene encodes a novel protein that is 540 amino acids in length. Protein sequence analysis predicts the presence of seven transmembrane domains, a characteristic feature of GPRs. *In situ* hybridisation to human retina and Northern blot analysis of human retinal pigment epithelium (RPE) showed localisation of this transcript to the RPE and cells surrounding retinal arterioles. In contrast, the transcript was localised to the photoreceptor inner segments and the outer plexiform layer in mouse sections. Northern blot analysis demonstrated a 7 kb transcript highly expressed in the brain. No mutations were identified during a screen of patients suffering from Doyne's honeycomb retinal dystrophy (DHRD), an inherited retinal degeneration which maps to chromosome 2p16.

Transmembrane signal transduction systems pass information from the exterior of a cell to the interior by a complex series of protein-protein interactions and second messenger systems. Many such systems rely on a family of G-protein-coupled receptors (GPRs) which upon binding of a ligand, transduce a signal via a guanine-nucleotide binding protein (1). Characteristically, GPRs are integral membrane proteins with seven transmembrane α-helices which span the lipid bilayer. The primary function of cell surface receptors is to transmit and amplify extracellular signals to control cellular processes. GPRs function in many physiological processes including neurotransmission, olfaction, hormonal responses and vision. The first GPR to be cloned and characterised was rhodopsin, the photopigment molecule in the rod cells of the vertebrate retina (2). Mutations in the rhodopsin gene have been found to be the underlying cause of disease in a significant proportion of patients suffering from the degenerative retinopathy retinitis pigmentosa (3).

During the construction of a physical map of an inherited retinal dystrophy, the Doyne's honeycomb retinal dystrophy (DHRD) locus on human chromosome 2p16 (4, 5), we identified an anonymous expressed sequenced tag (EST) WI-31133, on the chromosome 2 integrated map at the Whitehead Institute Centre for Genome Research (http://www-genome.wi.mit.edu/). This EST was of interest to us since it potentially mapped within the critical region and is derived from a clone isolated from a human retina cDNA library. We have determined the full length coding sequence of WI-31133 and show that it encodes a novel G-protein-coupled receptor. In addition, we have investigated its tissue expression profile in conjunction with *in situ* hybridisation, and determined the genomic structure that encompasses the coding interval to enable us to screen patients suffering from DHRD/dominant drusen for potential disease-causative mutations.

[1] Correspondence to: Section of Molecular Genetics, Division of Biomedical Sciences, Imperial College School of Medicine, Sir Alexander Fleming Building, South Kensington, London, SW7 2AZ. UK. Fax: +44-(0)171-594-3100. E-mail: c.gregory-evans@ic.ac.uk.

The sequence data reported in this paper have been submitted to GenBank and have been assigned the accession numbers AF072693 (IMAGE clone 222124 sequence) and AF101472 (genomic sequence).

0006-291X/99 $30.00

TABLE 1

Primer pair	Sequence 5′ to 3′	Encompassed bases	Size bp	Annealing temp °C	$MgCl_2$ mM
1A	CTTTCTGTTTCTGGGTGTGC				
1B	GTTGGTTCTGAATTTCCTG	30–406	376	55	1.5
2A	CACCTCGCTCCATGTGC				
2B	GAAGAGGTCACAGAAGG	211–448	238	55	2.0
3A	CTTCTTGTCCTTCTTCGATC				
3B	CGTGCGATTAGGCTGTTTC	361–637	276	55	1.5
4A	CACCAGTTCAGGCTTCATC				
4B	CAGGGTCTGAGCAATCATG	541–859	318	55	2.0
5A	GTCCAGTCTGATTGCTGG				
5B	GGACTCTTGGTATATCCAC	765–1047	282	55	1.5
6A	GTGGAGATCCCATCCAGTG				
6B	CTGGTAAAGAATGAAGCTCC	951–1240	289	55	2.0
7A	GGATTCCAAAGCCGTGGTC				
7B	GAGATAACATGTAGGCAGAG	1117–1475	358	55	2.0
8A	GCCTCCAATACATAGGCCTG				
8B	CTGCTCTCCTCCTGGGAAG	1341–1659	318	55	2.0
9A	CTCTGCTGGACATCAACAC				
9B	GTCCATTACTATCAGAAAC	1552–1852	300	50	2.0
10A	GTGCAGGAATATGACAGCAC				
10B	GTAGCTTCATTCCTGGCATC	1742–2073	332	55	2.0

MATERIALS AND METHODS

Database searches. EST WI-31133 was identified on the chromosome 2 integrated map at the Whitehead Institute Centre for Genome Research (http://www-genome.wi.mit.edu/). The sequence associated with WI-31133 (GDB:1064300; http://gdbwww.gdb.org/gdb/), is the 3′ sequence from IMAGE clone 222124 (GenBank H84878). IMAGE (6) clone 222124 was supplied by the UK HGMP Resource Centre (Hinxton, Cambridge CB10 1SB, UK).

Sequencing cloned DNA. Both strands of the insert of IMAGE clone 222124 were sequenced in their entirety using a T7 Sequencing Kit (Pharmacia) and Deaza G/A T7 Sequencing Mixes (Pharmacia) with [^{35}S]-α-dATP (Amersham). Products were separated by denaturing polyacrylamide gel electrophoresis (Sequagel, National Diagnostics) and visualised by autoradiography (Biomax MR, Kodak).

Peptide analysis. The ORF Finder (http://www.ncbi.nlm.nih.gov/gorf/gorf.html) was used to predict the open reading frame of the sequence generated from EST WI-31133. BLAST analysis (7; http://www.ncbi.nlm.nih.gov/gci-bin/BLAST/) was used to identify homology to the predicted protein sequence. The OMIGA1.1.3 (Oxford Molecular), and SOSUI (8; http://www.tuat.ac.jp/~mitaku/adv_sosui/) programs were used to predict secondary structure from the protein sequence.

PAC clone isolation and analysis. The insert of IMAGE clone 222124 was excised from its vector (pT7T3D) using *Eco*RI and *Hind*III, gel purified (QIAquick Gel Extraction Kit, Qiagen) and used to probe the RPCI1 PAC library (de Jong) on gridded filters (UK HGMP Resource Centre, Hinxton, Cambridge, UK). The probe was labelled with [^{32}P]-α-dCTP (Amersham) using an Oligolabelling Kit (Pharmacia) and hybridised against the filters in 0.5 M sodium phosphate buffer (pH 7.2), 7% SDS overnight at 65°C. The filters were washed with 0.2 × SSC, 0.01% SDS at 65°C prior to exposure against X-ray film at −80°C.

Positive PAC clones were restriction digested with *Eco*RI and the products were separated by 1% agarose gel electrophoresis prior to Southern blotting onto Hybond-N+ (Amersham) using 0.4 M NaOH. These filters were then probed as above.

PAC clone 135o20 from the RPCI1 library was sub-cloned into pUC18 (Pharmacia) and colony lifts were probed as above. Sequence data were generated as above.

In situ hybridisation. A 450 bp fragment was PCR amplified from IMAGE clone 222124 by oligonucleotides 5′-GCCTCCAATACATAGGCCTG-3′ (corresponding to bases 1379–1398 of AF072693) and 5′-CGGAGGGGACTGGAATCT-3′ (complementary to bases 1828–1811 of AF072693). The PCR product was subjected to agarose gel electrophoresis and the 450 bp fragment excised and gel purified with a QIAquick Gel Extraction Kit (QIAGEN), then cloned into pGEM-T (Promega). A full length opsin control probe was kindly provided by S. Jones (9).

Plasmids were linearised and RNA probes transcribed by either SP6 or T7 RNA polymerase to obtain sense and anti-sense probes. The probes were then labelled non-radioactively by digoxigenin-labelled nucleotides, according to the manufacturer's instructions (Boehringer Mannheim). Immediately before use, the probes were diluted in filtered hybridisation buffer (2 × SSC, 10% dextran sulphate, 100 μg/ml sheared salmon sperm DNA, 0.02% SDS, 50% formamide), to a concentration of 200 ng/ml, denatured for 2 minutes at 95°C and then quenched on ice.

8 μm paraffin sections of paraformaldehyde-fixed donor human and mouse eye tissue were collected on TESPA-treated slides. Sections were de-waxed in Histo-Clear (National Diagnostics), and the tissue rehydrated with a descending alcohol series to 30% ethanol/DEPC water. The sections were fixed with 4% paraformaldehyde in PBS at room temperature for 20 minutes and then treated with 50 μg/ml of proteinase K for 10 minutes at 37°C. Sections were dehydrated to 100% ethanol before pre-hybridisation for 30 minutes at 55°C. Hybridisation of the sections with the labelled probes was performed overnight at 55°C in a humid chamber containing 50% formamide and 2 × SSC.

Post-hybridisation washes were carried out in 50% formamide and 1 × SSC at 55°C and immunodetection of the mRNA:cRNA hybrids was performed using a 1:500 dilution of anti-digoxigenin alkaline phosphatase-conjugated Fab fragments (Boehringer Mannheim), followed by visualisation with the substrates fast red or BCIP/NBT (Boehringer Mannheim). Mouse sections were counter-stained with methyl green. Sections were viewed using bright field illumination on a Leitz Polyvar microscope.

Northern blot analysis. The 450 bp fragment generated for the *in situ* hybridisation was used to probe a Northern blot of poly(A)+ (Ambion mRNA purification kit, Austin) and total (SV Total RNA Isolation System, Promega) RNA extracted from human retinal pigment epithelium cell lines (RPE), (10), and three multiple tissue

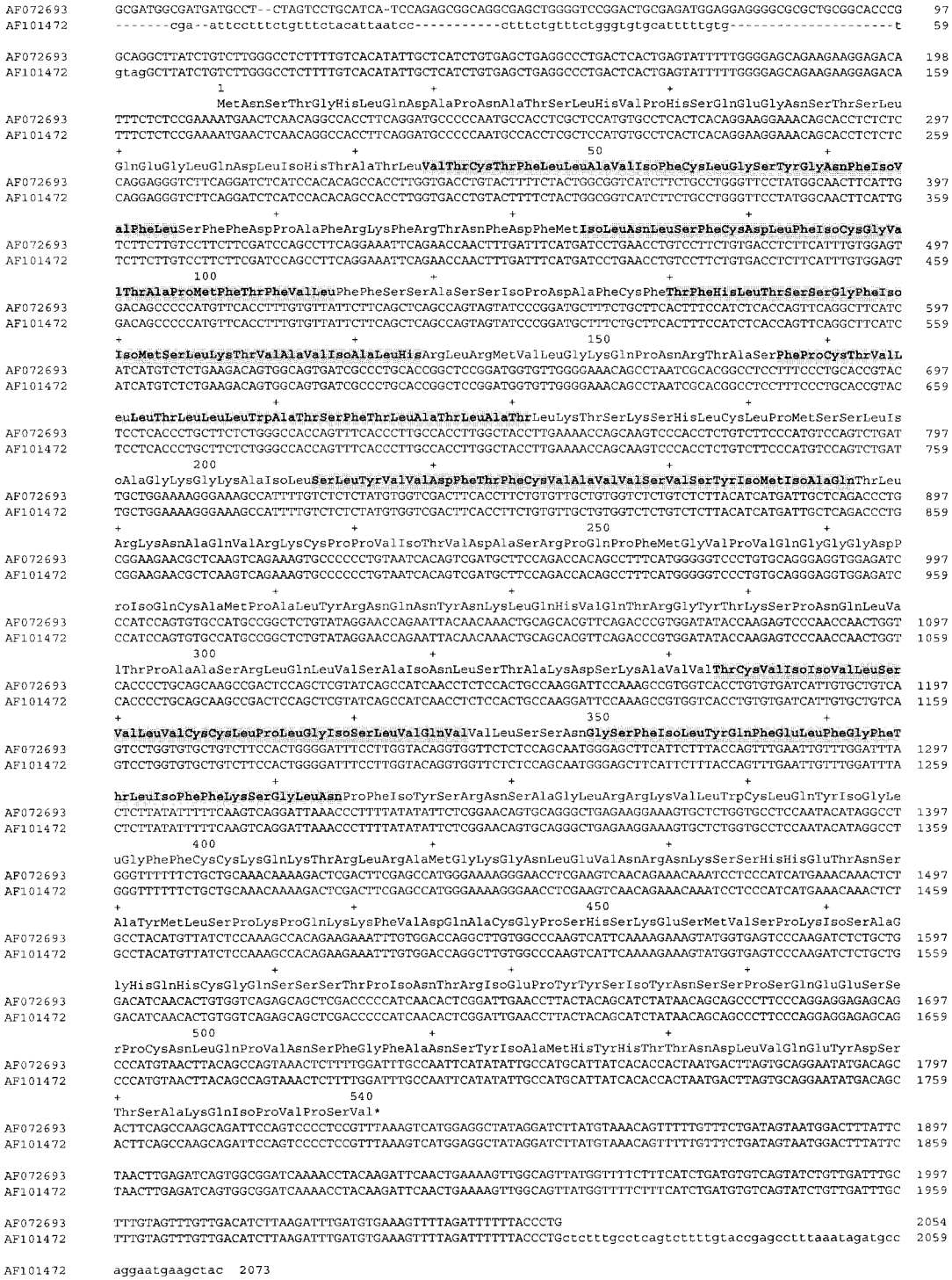

FIG. 1. Alignment of the nucleotide sequence of IMAGE clone 222124 (AF072693) and its corresponding genomic sequence (AF101472) derived from PAC clone 135o20. Lower case letters indicate intron acceptor splice site and the 3′ flanking region. The stop codon is indicated by an asterisk (*). The predicted peptide translation is shown above the nucleotide sequence with every tenth amino acid being indicated (+). Probable amino acids in seven transmembrane domains are shaded.

Northern blots (MTN blots, Clontech). Subsequently, two further MTN blots of multiple brain regions (Clontech) were probed with the same fragment. The probe was labeled using a random prime kit (T7 Quick Prime Kit, Pharmacia), with [^{32}P]-α-dCTP. After pre-hybridisation (QuickHyb, Stratagene), blots were hybridised at 60°C for 1 hour, then washed in 2 × SSC and 0.1% SDS at room temperature, then at 50°C in 0.1 × SSC and 0.1% SDS. The blots were then exposed against X-ray films at −80°C.

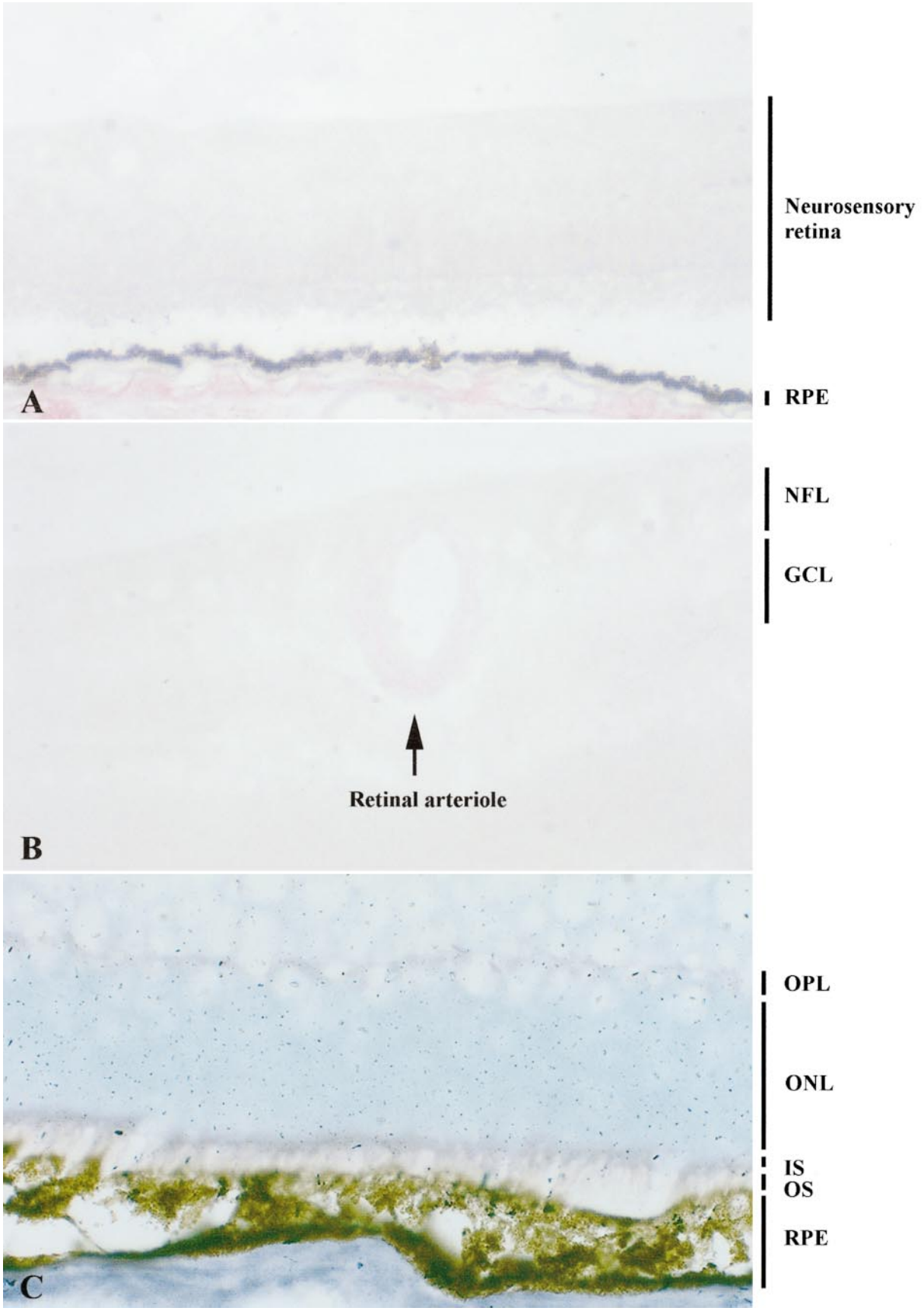

FIG. 2. Expression of WI-31133 in human and mouse retina. (A) Human eye section showing fast red signal in the choroid but not in the neural retina. (B) High magnification (100×) of human retinal arteriole showing perivascular staining. (C) Mouse eye section hybridised with the human probe showing the presence of transcript in the photoreceptor inner segment and the outer plexiform layer. RPE = retinal pigment epithelium, NFL = nerve fibre layer, GCL = ganglion cell layer, OPL = outer plexiform layer, ONL = outer nuclear layer, IS = inner segments, OS = outer segments.

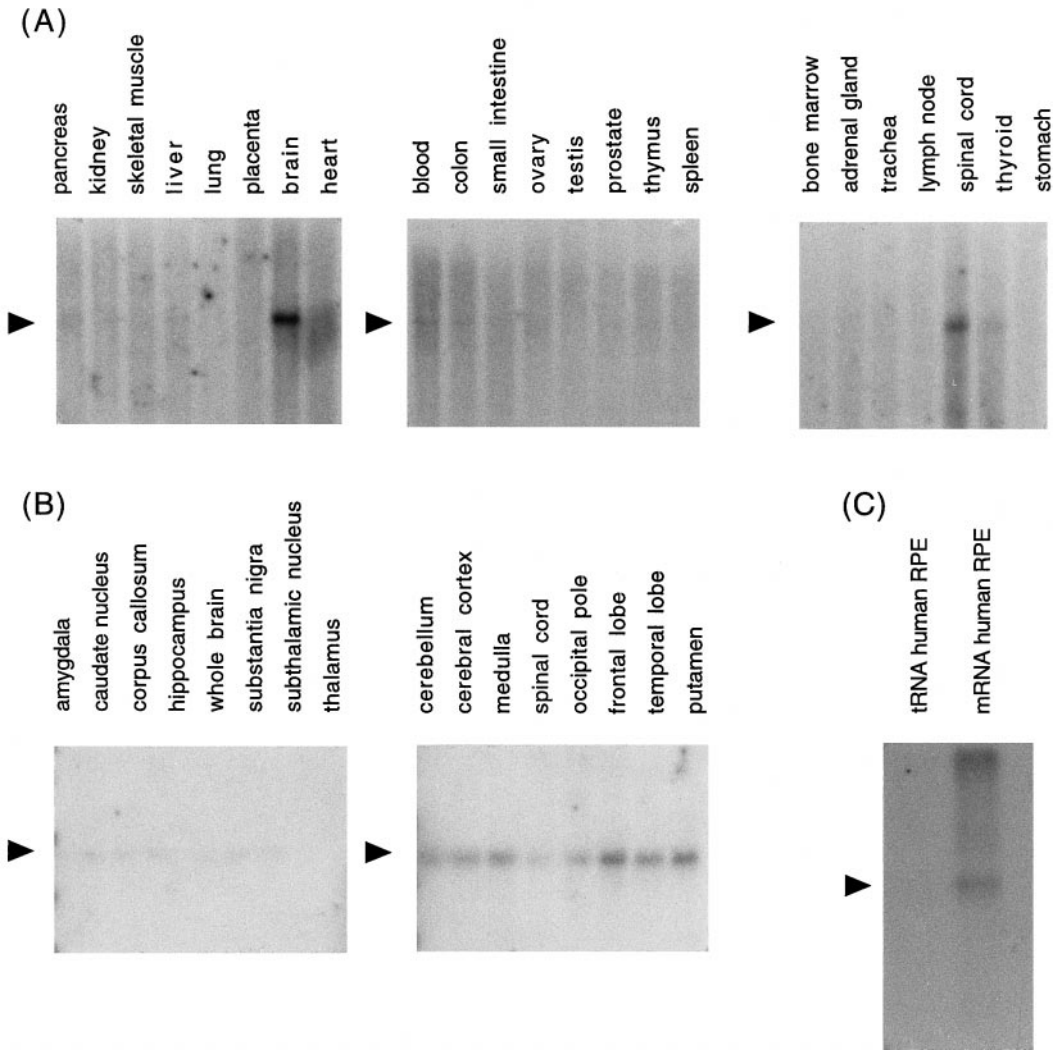

FIG. 3. Distribution of WI-31133 in human tissues as determined by Northern blot analysis. Each lane contains 2 μg poly (A)+ RNA, except the total RNA from human RPE which contains 20 μg and the poly (A)+ RNA from human RPE which contains 7 μg. (A) shows the distribution of the 7 kb transcript in 23 different human tissues, whilst (B) shows expression of the transcript in sub-regions of the brain. (C) Expression of WI-31133 is observed in poly (A)+ RNA from human RPE cells, but not in total RNA from human RPE.

Clinical appraisal. Doyne's honeycomb retinal dystrophy (DHRD; 11) is a rare condition leading to blindness usually in the fifth to sixth decade of life. Clinically, pre-senile (under 50 years of age) lipofuscin deposits (drusen) are seen accumulating under the neurosensory retina. This eye condition is similar to Malattia Leventinese which also maps to chromosome 2p16 (12) and the much commoner condition age-related macular degeneration. We have studied patients from six generations of one, 8 generation pedigree (13), and a number of affected individuals from a further 5 unrelated pedigrees. As well as classical features, it was noted that blindness was not always linked to advancing age. Occasional patients with DHRD have significant visual loss in early childhood.

Mutation detection. Direct sequencing of PCR products amplified from patient genomic DNA was used as the mutation detection method. Ten pairs of oligonucleotide primers were designed to amplify the entire coding region, extending from within the intron in the 5′UTR to the 3′UTR (see Table 1). PCR products were purified through Sephacryl S-400 Microspin columns (Pharmacia), following the manufacturer's instructions. Both strands of each fragment were sequenced with the relevant PCR primers, using a T7 Sequencing Kit (Pharmacia) and Deaza G/A T7 Sequencing Mixes (Pharmacia), with [^{33}P]-α-dATP (Amersham). Products were separated and visualised as described previously.

RESULTS AND DISCUSSION

The DHRD interval on human chromosome 2p16 is defined by the markers D2S2352 (telomeric) and D2S2251 (centromeric) (5). During the construction of a physical map of this region we identified an anonymous EST WI-31133, that co-segregates with the polymorphic marker D2S2251 (AFM205te7) on the chromosome 2 integrated radiation hybrid and genetic map at the Whitehead Institute Centre for Genome Research (http://www-genome.wi.mit.edu/). WI-31133 is derived from the 3′ sequence of IMAGE clone 222124 (GenBank: H84878) that originates from a retina cDNA library.

IMAGE clone 222124 (GenBank AF072693) contains a 2054 bp insert (excluding 20 bp poly-A tail), with a 1623 bp open reading frame (Fig. 1). The proposed ATG start codon occurs at 211 bp in a weak context; AAAatgA as opposed to RNNatgG (14), though is reinforced by the presence of in-frame stop codons 5′ to the start codon at positions 145 bp and 166 bp of the cDNA sequence. The open reading frame from 211 bp to 1833 bp predicts a 540 amino acid polypeptide with a molecular weight of 59.4 kDa. Sequence analysis (7; http://www.ncbi.nlm.nih.gov/gci-bin/BLAST/) of the predicted protein indicates weak homology to heptahelical G-protein-coupled receptors (GPRs). The conceptual translation of IMAGE clone 222124 is most closely related to a putative *C. elegans* neuropeptide Y receptor (GenBank: U41028), showing 24% identity; 25% identity to the rat galanin receptor type 3 (GenBank: AF031522), and 23% identity to the porcine growth hormone secretagogue receptor type 1b (GenBank: U60180). This homology to GPRs was confirmed by a Kyte-Dolittle hydropathy plot (OMIGA1.1.3, Oxford Molecular) which predicted the presence of seven transmembrane spanning regions, with the amino acids involved in the alpha-helices predicted by the SOSUI program (8).

We identified 6 positive clones from the RPCI1 PAC library that contain the genomic sequence of IMAGE clone 222124. Southern blot analysis of these clones indicated that they contained two EcoRI fragments positive for the 222124 cDNA-approximately 1.4 kb and 4.4 kb. One of the PACs, 135o20 was sub-cloned into pUC18 and a clone containing the 4.4 kb fragment identified. This was sequenced using primers designed when sequencing the 222124 cDNA, and it was found that the entire coding region was contained within a single exon together with 109 bp of 5′UTR and the whole 3′UTR (GenBank AF101472). The intron acceptor site in the 5′UTR conforms to the consensus, while the absence of a poly-A tail in the genomic sequence suggests that this gene is not a retropseudogene (15).

In order to identify the site of expression of this gene within the retina, the distribution of transcripts was visualized by *in situ* hybridisation to human adult eye sections. Transcripts were localized only to the perivascular cells, surrounding retinal arterioles, in the ganglion cell/nerve fibre layer (Fig. 2A and 2B). No transcript was detectable by *in situ* techniques in the rest of the human neurosensory retina or RPE. In contrast, in mouse sections staining was found in the inner segments of the photoreceptors and in the outer plexiform layer (Fig. 2C). Staining around retinal arterioles, as seen in human sections, was not seen in mouse sections. The significance of the different pattern of staining observed for WI-31133 between mouse and man has yet to be determined. The *in situ* probe was designed to the 3′ end of the human gene. This sequence is possibly different in the mouse gene, which has as yet to be cloned. Thus, the staining observed in the mouse could be imprecise and may not reflect the true distribution of WI-31133 in the mouse retina.

Northern blot analysis of poly (A)+ RNA from 23 different human tissues and total and poly (A)+ RNA from human RPE was performed at high stringency using a 450 bp probe derived from the 3′ end of the cDNA sequence. A transcript of approximately 7 kb was found to be expressed at high levels in the brain and spinal cord, with subsequent analysis of multiple brain regions showing expression in several areas of the brain (Fig. 3A and 3B). Expression of the transcript was also observed in human RPE, although only at detectable levels in the poly (A)+ RNA (Fig. 3C). Heavy RPE pigmentation could explain why this was not detected in our human *in situ* section studies.

A number of different cell types are present surrounding retinal arterioles. In particular, all retinal vessels are separated from surrounding neural elements by the cytoplasm of glial cells. These cells provide a support framework for vessels, act as insulation from neighboring neurons and provide a pathway for nutrition from the circulation to neurites. Large arterioles in the ganglion cell/nerve fiber layer are surrounded mainly by stellate, astrocytic glial cells compared to capillaries that are surrounded by Müller glial cells (16). Since the WI-31133 transcript was highly expressed in all areas of the brain that we tested, it is possible that this also corresponds to astrocyte glial cell staining. This will however need to be confirmed with glial cell-specific markers such as glial acid fibrillary protein (GFAP).

To determine whether WI-31133 is the disease-causative gene in DHRD, direct sequencing was carried out on genomic DNA from individuals from six different families. Both affected and normal individuals were screened in each case. No mutations were found to be segregating with the disease phenotype in any of the families examined. Two polymorphisms were found in both normal and affected individuals during the screen; a T → C substitution at nucleotide 489 (ATT–ATC) which preserves the isoleucine residue and a T → C substitution at nucleotide 1557 (AGT–AGC) which preserves the serine residue. Therefore we have identified a putative GPR that is expressed in the brain and RPE, but does not appear to be associated with the phenotype at the DHRD locus on human chromosome 2p16, where this gene localises.

ACKNOWLEDGMENTS

We gratefully acknowledge the financial support of the Wellcome Trust (Grants 043825/Z/95 and 054517/Z/98) while conducting this work. The authors also thank the United Kingdom Human Genome Mapping Project Resource Centre, Hinxton, Cambridge, CB10 1SB for the supply of IMAGE clone 222124, the gridded RPCI1 PAC library, and the individual clones isolated from it.

REFERENCES

1. Dohlman, H. G., Thorner, J., Caron, M. G., and Lefkowitz, R. J. (1991) *Ann. Rev. Biochem*. **60,** 653–688.
2. Nathans, J., and Hogness, D. S. (1983) *Cell* **34,** 807–814.
3. Gal, A., Apfelstedt-Sylla, E., Janecke, A. R., and Zrenner, E. (1997) *Progress in Retinal and Eye Research* **16**(1), 51–79.
4. Gregory, C. Y., Evans, K., Wijesuriya, S. D., Kermani, S., Jay, M. R., Plant, C., Cox, N., Bird, A. C., and Bhattacharya, S. S. (1996). *Hum. Mol. Genet.* **5**(7), 1055–1059.
5. Kermani, S., Gregory-Evans, K., Tarttelin, E. E., Bellingham, J., Plant, C., Bird, A. C., Fox, M., Bhattacharya, S. S., and Gregory-Evans, C. Y. (1999) *Hum. Genet.* **104,** 77–82.
6. Lennon, G. G., Auffray, C., Polymeropoulos, M., and Soares, M. B. (1996) *Genomics* **33,** 151–152.
7. Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D. J. (1997) *Nuc. Acids Res.* **25,** 3389–3402.
8. Hirokawa, T., Boon-Chieng, S., and Mitaku, S. (1998) *Bioinformatics* **14**(4), 378–379.
9. Foreman, D. M., Jones, S., McKechnie, N. M., Williams, G., and Boulton, M. E. (1996) *Ophthalmic Res.* **28,** 296–302.
10. Dunn, K. C., Aotaki-Keen, A. E., Putkey, F. R., and Hjelmeland, L. M. (1996) *Exp. Eye Res.* **62,** 155–169.
11. Doyne, R. W. (1899) *Trans. Ophthalmol. Soc. UK.* **19,** 71.
12. Heon, E., Piguet, B., Munier, F., Sneed, S. R., Morgan, C. M., Forni, S., Pescia, G., Schorderet, D., Taylor, C. M., Streb, L. M., Wiles, C. D., Nishimura, D. Y., Sheffield, V. C., and Stone, E. M. (1996) *Arch. Ophthalmol.* **114,** 193–198.
13. Evans, K., Gregory, C. Y., Wijesuriya, S. D., Kermani, S., Jay, M. R., Plant, C., and Bird, A. C. (1997) *Arch. Ophthalmol.* **115,** 904–910.
14. Kozak, M. (1996) *Mamm. Genome* **7,** 563–574.
15. Weiner, A. M., Deininger, P. L., and Efstratiadis, A. (1986) *Ann. Rev. Biochem*. **55,** 631–661.
16. Robinson, S. R., and Dreher, Z. (1989) *Neurosci. Lett.* **106,** 261–268.