Positional cloning in Xp22 : towards the isolation of the gene involved in X-linked retinoschisis

Vosse, E. van de

CHAPTER 4.2

A novel human serine-threonine phosphatase related to the *Drosophila retinal degeneration C* (*rdg*C) gene is selectively expressed in sensory neurons of neural crest origin

Montini, E., Rugarli, E.I., Van de Vosse, E., Andolfi, G., Mariani, M., Puca, A.A., Consalez, G.G., Den Dunnen, J.T., Ballabio, A., and Franco, B.

Human Molecular Genetics 6:1137-1145, 1997

A novel human serine-threonine phosphatase related to the *Drosophila retinal degeneration* C (*rdgC*) gene is selectively expressed in sensory neurons of neural crest origin

E. Montini[1,+], E. I. Rugarli[1,+], E. Van de Vosse[3], G. Andolfi[1], M. Mariani[2], A. A. Puca[1], G. G. Consalez[2], J. T. den Dunnen[3], A. Ballabio[1] and B. Franco[1,*]

[1]Telethon Institute of Genetics and Medicine (TIGEM) and [2]Department of Biological Technological Research (DIBIT), San Raffaele Biomedical Science Park, Milan, Italy and [3]MGC, Department of Human Genetics, Leiden University, Leiden, the Netherlands

Received February 14, 1997; Revised and Accepted April 8, 1997

Through our transcriptional mapping effort in the Xp22 region, we have isolated by exon trapping a new transcript highly homologous to the *Drosophila retinal degeneration* C (*rdgC*) gene. *rdgC* encodes a serine/threonine phosphatase protein and is required in *Drosophila* to prevent light-induced retinal degeneration. This human gene is the first mammalian member of the serine-threonine phosphatase with EF hand motif gene family, and was thus named *PPEF* (Protein Phosphatase with EF calcium-binding domain). The expression pattern of the mouse *Ppef* gene was studied by RNA *in situ* hybridization on embryonic tissue sections. While *rdgC* is expressed in the visual system of the fly, as well as in the mushroom bodies of the central brain, we found that *Ppef* is highly expressed in sensory neurons of the dorsal root ganglia (DRG) and neural crest-derived cranial ganglia. The selective pattern of expression makes *PPEF* an important marker for sensory neuron differentiation and suggests a role for serine-threonine phosphatases in mammalian development.

INTRODUCTION

Our group is involved in the construction of a transcription map of the human Xp22 region. To achieve this aim, we have constructed a detailed physical map of a 35 Mb region spanning human chromosome Xp22.3–Xp21.3. The backbone of the map is represented by a single contiguous stretch of 585 overlapping yeast artificial chromosome (YAC) clones covering the entire region (1). Several disease loci have been mapped in this region including Retinoschisis (RS), Nance–Horan syndrome (NHS), Coffin–Lowry syndrome (CLS), and Keratosis Follicularis Spinulosa Decalvans (KFSD) (1). As a first step toward building a transcription map of this region, we decided to concentrate our efforts on YAC clone 939H7 which spans the entire critical region for RS and partially spans the critical regions for NHS, CLS and KFSD.

This effort led us to the isolation of a gene highly homologous to the *Drosophila retinal degeneration* C (*rdgC*) gene. The *rdgC* gene encodes a serine/threonine phosphatase protein (2) and is required in *Drosophila* to prevent light-induced retinal degeneration (3). *rdgC* is expressed in the visual system of the fly, as well as in the mushroom bodies of the central brain. We named this new gene *PPEF*, for Protein Phosphatase with EF hand motif.

To test the involvement of *PPEF* in the pathogenesis of RS, we isolated the full-length cDNA, established the genomic structure, and searched for mutations in RS patients by PTT (protein truncation test) and SSCP (single strand conformation polymorphism) analyses. Identification of the mouse homolog of this gene allowed us to perform RNA *in situ* hybridization studies on mouse embryo tissue sections, revealing a very specific pattern of expression localized in sensory neurons of cranial and dorsal root ganglia.

RESULTS

Identification and characterization of the PPEF cDNA

One hundred cosmid clones were selected by screening an X-specific cosmid library with *Alu* PCR products derived from YAC 939H7. Cosmid clones were grouped and used for exon trapping experiments. Several exon trapping products were sequenced and used to search non-redundant DNA and protein databases through the BLAST-X, BLAST-P and TBLAST-N algorithms. One of them (clone 3pn1D2) was found to be identical to EST H18854, which shows significant homology to the *Drosophila rdgC* protein (accession no. M89628). Clone 3pn1D2 mapped back to YAC939H7 in the RS critical region and to cosmid 44C1 containing marker *DXS999* (Fig. 1). The same transcript was subsequently identified in our laboratory, using a bioinformatic approach aimed at the identification of human homologs of *Drosophila* genes involved in the generation of

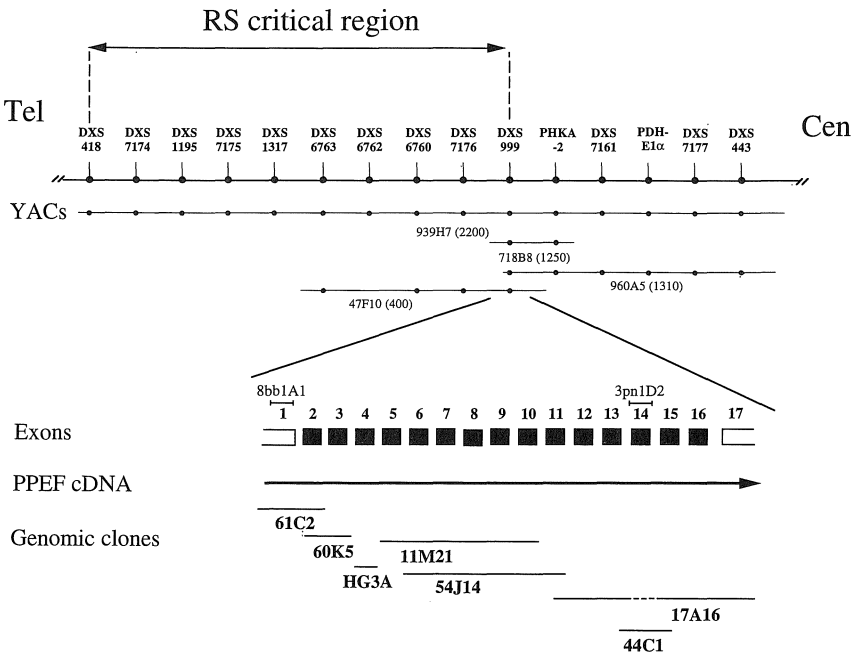

Figure 1. Map of the RS critical region. The map displays the order of markers and the YACs positive for two or more markers. The RS critical region spans from *DXS418* distally to *DXS999* proximally. YAC clones are indicated by thin bars. Cosmid and phage clones are indicated by thick bars. The number and position of the exons are shown.

mutant phenotypes, and found to correspond to DRES10 (*Drosophila* related expressed sequence #10) (4).

The I.M.A.G.E. cDNA clone 51064 corresponding to EST H18854 was used to screen both an infant (5) and a fetal (Clontech) brain cDNA library. Six different cDNA clones were isolated and characterized by end-sequencing, restriction mapping and PCR, using specific primers in combination with vector primers. Subsequently, an additional cDNA (nt19) was isolated by screening a teratocarcinoma/neuron (mature hNT neuron, Stratagene 937233) cDNA library with a different exon trapping product (clone 8bb1a1). Characterization of each of the cDNA clones allowed us to establish a cDNA contig of 2872 bp (base pairs) (accession no. X97867). The authenticity of the 5′ end of the cDNA was validated by sequencing the corresponding genomic region. The putative initiation codon was identified at position 484 and is located within a nucleotide sequence that fulfills Kozak's criteria for an initiation codon (6). The first in-frame stop codon (TAA) was identified at nucleotide 2443, predicting a protein product of 653 amino acids. Sequence analysis of the predicted protein product revealed the presence of two putative functional domains. The first domain comprises amino acids 150–438 and shares sequence similarity with the catalytic domain of phosphoprotein phosphatases of the serine-threonine kind (Fig. 2A). The second domain is present at the carboxyl end of the predicted protein (amino acids 566–643) and contains potential Ca^{++}-binding sites as defined by the EF hand motif (7) (Fig. 2A and B). On the basis of these homologies, we therefore named the gene *PPEF*, for Phosphoprotein Phosphatase with EF hand motif.

Sequence analysis and comparison of *PPEF* with previously identified phosphoprotein phosphatases revealed conservation of several invariant residues including Asp59, Asp88, His61 and His139. Site-directed-mutagenesis experiments revealed that substitution of these residues results in abolishment of phosphatase activity (8,9). These data suggest that PPEF might be a functional phosphatase.

A BLASTX search with PPEF disclosed the highest homology with the *Drosophila rdgC* gene, which also encodes for a serine-threonine phosphatase with EF motif and, when mutated, causes light-induced retinal degeneration in the fly. *PPEF* and *rdgc* are 60% identical at the nucleotide level, and 62% similar and 42% identical at the protein level. To the best of our knowledge, the *rdgC* and PPEF proteins are the only two molecules in which a phosphoprotein phosphatase domain coexists with EF hand motifs. Furthermore, within the phosphoprotein phosphatase domain, PPEF displays much higher homology with *rdgC* than with other members of the phosphoprotein phosphatase gene family (data not shown), indicating that PPEF and *rdgC* belong to a distinct subfamily of serine-threonine phosphatases.

PPEF genomic structure and RS patient analysis

The complete genomic structure of the *PPEF* gene was determined using the available cosmid clones. Genomic clones corresponding to exon 4 were obtained by screening a total phage genomic library using a specific PCR probe. *Eco*RI and *Hin*dIII

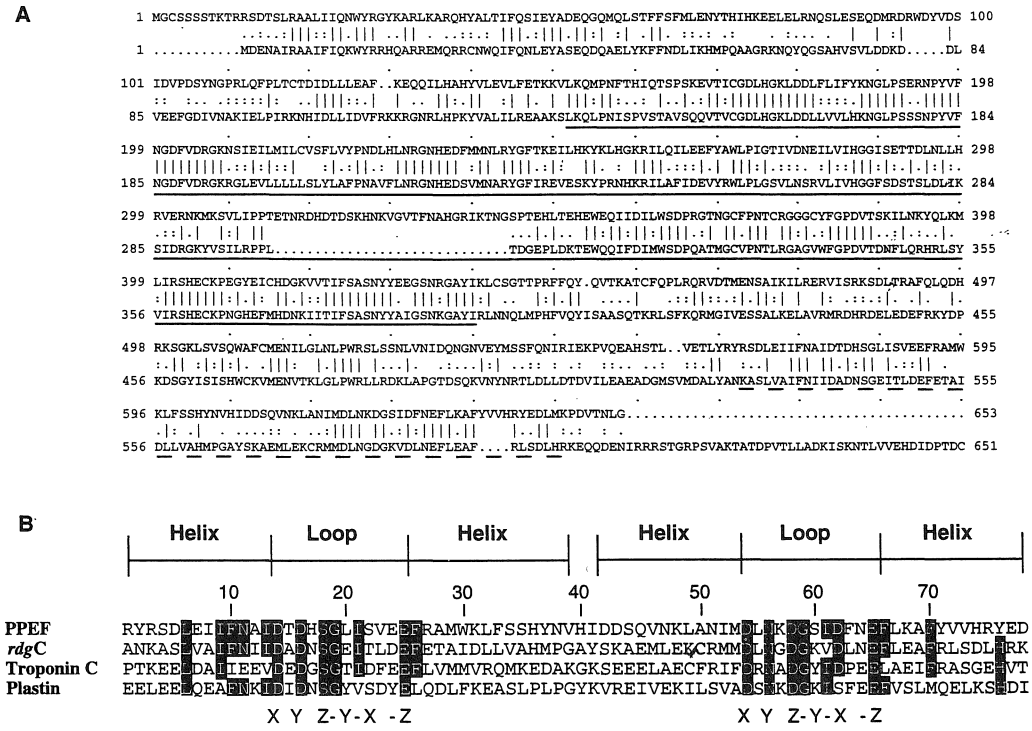

Figure 2. Sequence analysis of the PPEF predicted protein. (A) Amino acid sequence comparison between PPEF (top) and *Drosophila rdg*C (bottom) proteins. The region of homology shared with other members of the protein phosphatase family is underlined. The EF calcium-binding domain is indicated by a dashed line. (**B**) Amino acid identity within the EF hand motifs between PPEF *rdg*C, Troponin C and Plastin. The positions of the Ca^{++} chelating side chains are labeled X–Z.

cosmid and phage fragments hybridizing to the cDNA were subcloned in pBluescript and sequenced using primers designed from the cDNA sequence. Seventeen exons were identified and the sequences of all intron–exon boundaries were determined. Exon sizes and splice junction sequences are shown in Table 1. Junction sequences are in agreement with 5′ and 3′ splice site consensus motifs, with the exception of the exon 15 5′ splice site which shows the sequence GCAAGTG, instead of the consensus 5′ splice site sequence $G_{100}T_{100}A_{62}A_{68}G_{84}T_{63}$. Differences in the GT dinucleotide have been reported in 0.13% of splice site sequences (10). The exon trapping products 8bb1A1 and 3pn1D2 were found to correspond to the first untranslated exon and to exon 14, respectively (Fig. 1). Figure 1 shows the map of the genomic clones and the position of the exons.

Sequence homology and mapping data identified the *PPEF* gene as a good candidate for retinoschisis and thus, a mutation study in RS patients was undertaken. Thirty-seven unrelated male patients with clinical features of RS, but with no reported chromosomal abnormality involving the Xp22 region, were tested for small rearrangements or point mutations in the *PPEF* gene by Southern and SSCP analyses (data not shown). Furthermore, PTT analysis was carried out on nine different RS patients (data not shown). No anomalies were found with either technique, thus suggesting that *PPEF* is not involved in the pathogenesis of RS.

PPEF is alternatively spliced

RT-PCR experiments, using nested primers, on RNA isolated from lymphocytes of RS patients and normal controls revealed several different size products (Fig. 3B). Similar results were obtained using RT-PCR on RNA isolated from seven other tissues (data not shown). Sequence analysis of these different products and comparison with the cDNA sequence confirmed that they were the result of alternative splicing. Figure 3A shows a schematic representation of all the splice variants detected. RT-PCR1a corresponds to a product in which exon 5 was spliced out, resulting in a truncated protein of 600 amino acids. RT-PCR1b is the result of using a cryptic splice site within exon 11. This product lacks 84 bp and, therefore, encodes a protein deleted of 28 amino acids. RT-PCR2 corresponds to a product in which exons 12 and 13 were spliced out. This splicing event causes the premature termination of the protein. Finally, RT-PCR3 corresponds to a product in which only exon 12 was spliced out. This alternative spliced variant results in an in-frame deletion of 186 bp and, thus, in a protein which is missing 86 amino acids. These variant forms have never been identified during cDNA library screening, and may not have any functional significance.

Table 1. Splice junction sequences of the PPEF gene

Exon number	Splice junctions 3′ splice site	5′ splice site	Exon size (bp)
1	ttttctttcagCCCTCTG	AAACAGgtaatatgtt	103
2	tcctcattagAATCTAT	ACACATgtgagtactg	168
3	gggtctgcagCACTGA	ATGCAGgtctgttttg	128
4	ttttgtccagTTATCC	AGCTAGgtaagtaaaa	61
5	ccttttccagAATTAA	CAACAGgtaagtggaa	161
6	ccgctcacagATACTT	TCTGTGgtaagtttca	115
7	tgcactgcagGTGATT	TACAAGgtaaatgatg	47
8	cttccacagAATGGT	TCTGAGgtaacccag	167
9	ttttttttagGTATGG	TATAAGgtaagacatg	37
10	ttacttacagCTACAT	AACAAGgtaagaagta	150
11	tgttttatagATGAAA	GAACAGgtaggtaatc	153
12	ttgttccaagATTATT	GGGAAGgtaagctaaa	86
13	ctttaaccagGTGGTG	CCAAAGgtgtgtatac	143
4	tttctttgcagAGTGG	AATCAGgtaacaaatt	107
15	ttgttgttagGAAAAC	CAAGAGgcaagtgaaa	164
16	tttattttagGCTCAT	ACTCAGgtaaataaat	85
last	attcttttagGCCTGA	–	

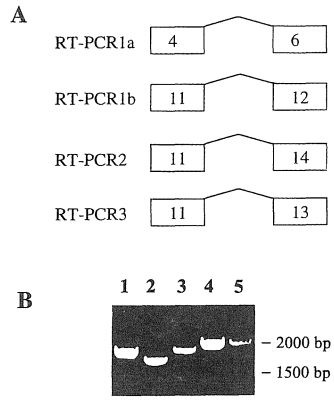

Figure 3. (**A**) Schematic representation of different splice variants identified by RT-PCR experiments, using nested primers. (**B**) RT-PCR analysis on RNA isolated from lymphocytes of RS patients and normal controls. Lane 1 corresponds to the products RT-PCR1a and RT-PCR1b. Lane 2 to RT-PCR2, lane 3 to RT-PCR3. Lane 4 corresponds to a product identical to the cDNA described in the text. Lane 5 contains a 100 bp marker.

Expression studies

The expression pattern of the *PPEF* transcript was determined by both Northern analysis on human tissues and RNA *in situ* hybridization on embryonic mouse tissue sections. Northern blot analysis on both human adult and fetal tissues detected two transcripts (2.7 and 4.3 kb, respectively) selectively expressed in the brain. These transcripts were found to be strongly up-regulated during fetal life (Fig. 4).

Previous zoo blot experiments indicated evolutionary conservation of the *PPEF* gene in several species including hamster, rat, mouse, pig, chicken and cow (data not shown). To obtain a probe for RNA *in situ* hybridization studies on mouse tissue sections, we screened a mouse embryonic E11.5 day cDNA library (Clontech) using probe c14-c22 obtained by PCR with primers c14 and c22. This PCR product corresponds to the region with the highest sequence homology (63.2% at the nucleotide level) between the human *PPEF* and the *Drosophila rdgC* genes. This screening led to the isolation of a partial cDNA clone (EM800; accession no. Y08234). Sequence analysis revealed 82.4% identity at the nucleotide level, and 83.1% similarity and 76.4% identity at the protein level between mouse and human (data not shown). The homology between the mouse and *Drosophila* proteins was revealed to be 65.3% similarity and 43.1% identity.

RNA *in situ* hybridization was performed on mouse embryonic tissue sections from embryonic day 10.5 (E10.5) to embryonic day 16.5 (E16.5). These experiments revealed an expression pattern overtly different from that displayed by the *rdgC* gene in the fly. *Ppef* is almost exclusively expressed in the peripheral nervous system, within sensory neurons of neural crest origin. *Ppef* expression is up-regulated at E12.5 in dorsal root ganglia (DRG), and in some sensory cranial ganglia (Fig. 5A). Sensory neurons of the vertebrate peripheral nervous system have two distinct embryological origins. Several studies, mainly in the chick, have shown that neurons of the DRG, of the dorso-medial part of the trigeminal ganglion, and of the superior ganglia of cranial nerves IX and X (jugular ganglion) are derived from the neural crest, while neurons of the ventrolateral part of the trigeminal ganglion, the geniculate, vestibuloacustic, petrosal and nodose ganglia, are derived from ectodermal placodes, i.e., thickening of the surface ectoderm (11). *Ppef* expression was found to be restricted to neuronal populations of neural crest-derived sensory ganglia. In fact, *Ppef* is highly expressed in neurons of the DRG (Fig. 5C and D), in distinct neuronal populations of the trigeminal ganglion (Fig. 5B), and in the superior ganglia of the IX and X cranial nerves. No expression was observed in the geniculate, vestibuloacustic (Fig. 5B), petrosal, and nodose ganglia.

Although sensory ganglia are already formed in the mouse well before E12.5, this stage is believed to correspond with the start of neurogenesis in these structures (12). An enlarged view of *Ppef* expression within DRG and the trigeminal ganglion (Fig. 5B) clearly demonstrated that most, but not all, neurons express *Ppef*. A slightly decreased expression could still be detected in sensory ganglia at E16.5. We do not know if *Ppef* expression in sensory neurons persists in adult life. The discrepancy in the strength of the hybridization signals on Northern blot in adult and fetal brain is probably due to the inclusion of the cranial sensory ganglia in the brain sample of fetal origin. The only sites of *Ppef* expression, outside sensory neurons, were found in the inner ear (Fig. 5E) and in a small group of neurons located at the midbrain/pons junction (Fig. 5F). We did not detect any expression signals above background level in the embryonic and adult mouse retina (data not shown).

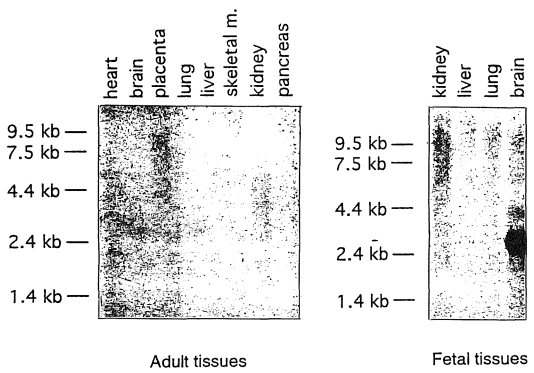

Figure 4. Northern blot analysis of the PPEF gene. PolyA$^+$ RNA from multiple adult and fetal tissues hybridized to cDNA corresponding to the PPEF gene. A 2.7 kb transcript is detected in both adult and fetal brain, while a less abundant 4.3 kb band is detected only in fetal brain.

Linkage mapping of *Ppef* in mouse

The finding in the N2 progeny of the BSS backcross (13) of male individuals carrying only B alleles (hemizygotes) led to the unequivocal assignment of *Ppef* to the murine X chromosome. The analysis of the strain distribution pattern (SDP) observed in the same progeny permitted the localization of *Ppef* to the distal third of the chromosome (Fig. 6). *Ppef* is in linkage with DXXrf132 (Spencer *et al.*, in preparation) (θ = 1.41; LOD = 19.1), and was mapped using a probe containing dbEST 122118 (accession no. F06456) with homology to the *Saccharomyces cerevisiae* GTR1 gene. In the BSS map, *Ppef* is located telomeric to DXXrf132. *Ppef* maps ~1.5 cM centromeric to *Grpr*, the gastrin releasing peptide receptor gene, identified using DXMit20 primers (Korobova and Arnheim, unpublished) and *Piga* (θ = 1.37; LOD = 19.7). The locus order defined in mouse between the *Amel* (Korobova and Arnheim, unpublished) and *Ppef* loci correlates with the physical map established in human (1). However, the position of the *Oa1* gene with respect to the *Clc4-1-Amel-(Piga-Grpr)-Ppef* linkage group confirms the prior notion of a rearrangement (inversion) within this region of human–mouse synteny (14).

DISCUSSION

PPEF is the first mammalian member of the protein phosphatase with EF hand motif gene family

A wide variety of cellular functions, including cell signalling, gene expression, membrane transport and secretion and cell division, are regulated by the reversible phosphorylation of proteins on serine and threonine residues (15). The phosphatases that catalyze the dephosphorylation of these amino acids are a crucial component of this regulation. The serine-threonine phosphatases belong to a rapidly expanding gene family, and six different mammalian members (PP1, PP2A, PP2B, PP2C, PP4 and PP5) have been described so far. They can be distinguished depending on their ability to dephosphorylase either the α- or the β-subunit of phosphorylase kinase, and on their sensitivity to specific inhibitors (16). These phosphatases have been highly conserved during evolution from yeasts to vertebrates (17).

In *Drosophila*, the *rdgC* gene encodes a serine/threonine phosphatase. This phosphatase shares 30% homology with the catalytic domain of PP1, PP2A and PP2B, but is unique due to the presence of five Ca^{++}-binding sites, as defined by the EF hand motif in the C-terminus. Owing to its particular features, *rdgC* is likely to be a member of a novel subfamily of protein phosphatases, characterized by the coexistence of the catalytic phosphatase domain and Ca^{++}-binding sites. Very little is known about the function of this phosphatase and no evidence for vertebrate homologs has been produced so far. The *rdgC* gene is required in *Drosophila* to prevent light-induced degeneration of the retina. *Drosophila rdgC* mutants show normal retinal morphology and photoreceptor physiology at a young age. The retina of one-day-old *rdgC* mutants has wild type structure, but by three days, the photoreceptors R1–R6 begin degenerating. By five days, degeneration of photoreceptors R1–R6 is complete and photoreceptors R7 and R8 begin showing signs of degeneration (3). The *rdgC* gene is thought to be involved in the regeneration of rhodopsin and is expressed in the retina, ocelli, optic lobes and in the mushroom bodies of the central brain (2).

We report here the isolation of the first mammalian member of the serine-threonine phosphatase with EF hand motif gene family. The new gene was named *PPEF* (Protein Phosphatase with EF calcium-binding domain) and is highly homologous to the *Drosophila rdgC* gene. In order to study the expression pattern of PPEF during embryonic development, we isolated a partial murine homologous cDNA. This cDNA shows 82.4% identity at the nucleotide level and 83.1% similarity and 76.4% identity at the protein level with the human *PPEF* gene. Linkage mapping experiments, performed in the mouse, established that this gene is located on the murine X chromosome in a region synthenic with human Xp22 (18). These data strongly indicate that the murine *Ppef* gene may be considered a *bona fide* ortholog for the human *PPEF*.

PPEF expression is restricted to sensory neurons of neural crest origin

In the fly, *rdgC* is highly expressed in the compound eye as well as in the ocelli, the other photoreceptor-containing organ, and is thought to participate in a rhodopsin-initiated pathway that regulates photoreceptor membrane renewal. The tissue

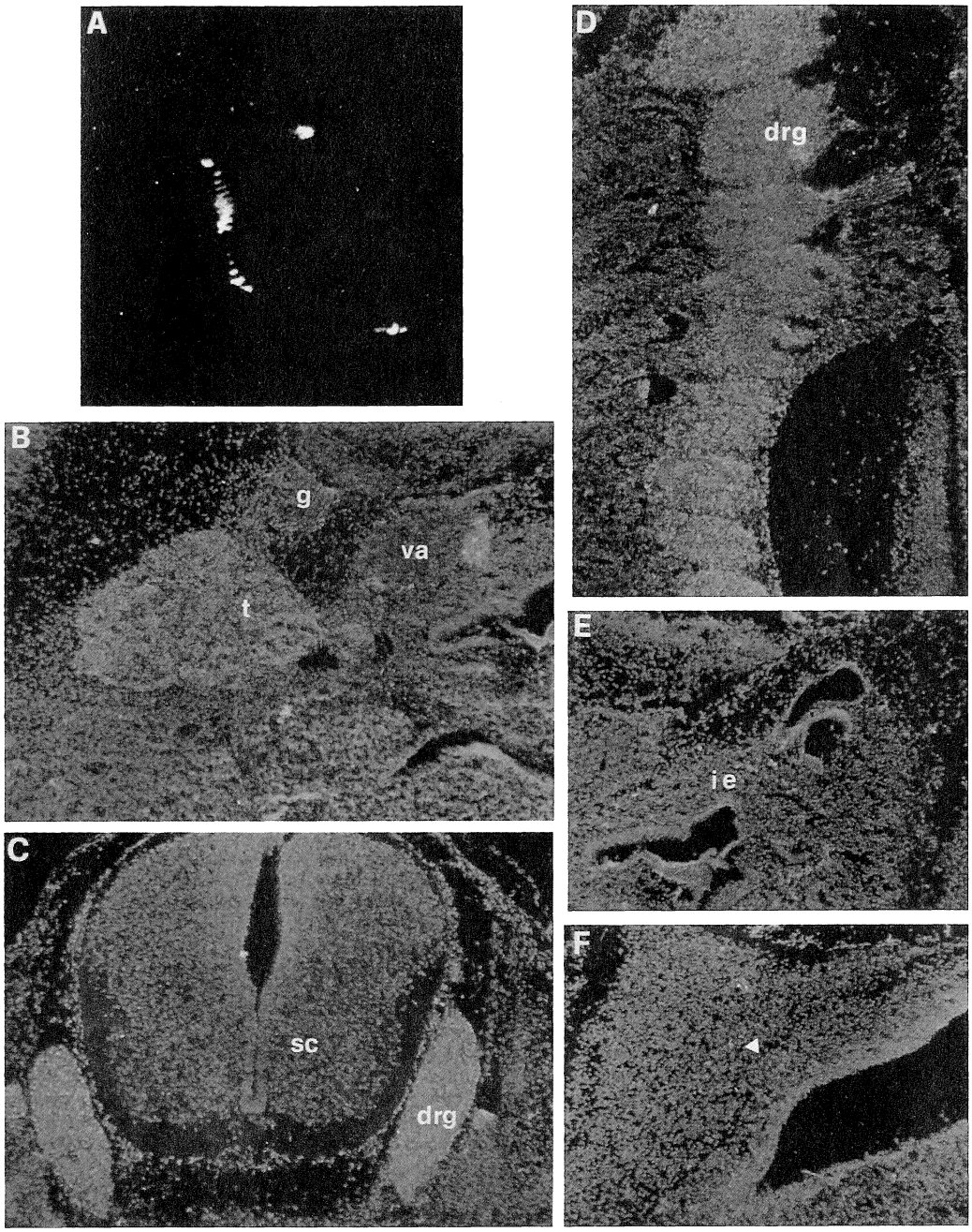

Figure 5. Expression pattern of PPEF, as revealed by *in situ* hybridization in a 13.5 day old mouse embryo. (**A**) Autoradiography of a sagittal section shows expression restricted to the trigeminal ganglion and the dorsal root ganglia. Sagittal sections through the head show PPEF expression in the trigeminal ganglion (**B**). A strong hybridization signal is present in the dorsal root ganglia, as displayed in a sagittal section (**C**) and in a transverse section (**D**). Ppef positive signal is also present in the inner ear (**E**) and in a group of neurons located at the pons/midbrain junction (arrow in **F**). Abbreviations: tg, trigeminal ganglion; g, geniculate ganglion; va, vestibuloacoustic ganglion; drg, dorsal root ganglia; sc, spinal cord; ie, inner ear.

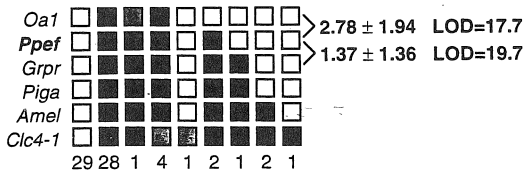

Figure 6. Haplotype and linkage analysis of *Ppef* and flanking loci in the murine X chromosome through the analysis of the BSS backcross (Jackson Laboratory, Bar Harbor, ME). Empty squares indicate the *Mus spretus* allele; solid squares indicate the C57BL/6J allele; stippled squares indicate genotype not determined. Numbers to the right, between rows, indicate recombination fractions ± standard error, and LOD scores. Columns represent different haplotypes observed on the X chromosome. Numbers below columns define the number of individuals in the progeny sharing each haplotype.

localization of the mouse *Ppef* gene was studied by RNA *in situ* hybridization on embryonic tissue sections. In contrast with the expression pattern of the *Drosophila rdgC* gene, we found that *Ppef* is selectively expressed in neurons of the DRG, in distinct neuronal populations of the trigeminal ganglion, and in the superior ganglia of the IX and X cranial nerves. No expression was observed in the geniculate, vestibuloacustic, petrosal and nodose ganglia. This selective pattern of expression correlates with embryological origin, with Ppef being a marker specific for sensory neurons of neural crest origin. Several genes have been found to be expressed in sensory neurons and RNA *in situ* hybridization studies have allowed the study of the distribution of specific transcripts within sensory ganglia. In addition, immunocytochemical and histochemical techniques have been used extensively to biochemically define distinct neuronal populations in sensory ganglia by the presence of peptides, enzymes and specific carbohydrate groups (19,20). So far, however, the only other genes that show an expression pattern restricted to sensory neurons are the neurotrophin receptor genes (21).

For their survival, sensory neurons of the developing peripheral nervous system rely upon specific members of the Nerve Growth Factor (NGF) family of neurotrophins, which are secreted by their targets. It is believed that different neuronal populations in DRG are responsive to different neurotrophins and express different neurotrophin receptors (22,23). *In vitro* studies of neurons from cranial sensory ganglia have shown that there is a difference in the response of placode-derived and neural crest-derived neurons to neurotrophins; neural crest-derived neurons are responsive to NGF, while placode-derived neurons respond to Brain-Derived Neurotrophic Factor (BDNF) (24,25).

The selective expression of *Ppef* in cranial sensory ganglia of neural crest-origin and the lack of expression in placode-derived ganglia suggest that this gene is expressed by NGF-responsive neurons. NGF exerts its effect by eliciting a phosphorylation cascade through activation of the TrkA tyrosine-kinase receptor (26). Although the phosphorylation events mediated by the binding of NGF to the TrkA receptor have been extensively studied, very little is known about dephosphorylation pathways and the phosphatases involved. It is an appealing hypothesis that PPEF might be involved in the specific signalling pathway initiated by NGF in sensory neurons. Consistently, up-regulation of *Ppef* expression coincides with the time at which neurogenesis and TrkA receptor expression begins in sensory ganglia.

In conclusion, our study provides evidence for the presence of a mammalian protein phosphatase with EF hand motifs. Although this phosphatase is highly homologous to the *Drosophila rdgC* gene, expression studies seem to suggest that it represents a related gene with a different function. We do not exclude the possibility that additional members of this protein phosphatase gene family will be identified in the near future. Alternatively, *PPEF* might be the evolutionary equivalent of the *Drosophila rdg*C gene, but we must assume that the two genes have diverged in terms of function. Since mapping data placed *PPEF* within the critical region for RS, we have excluded its involvement in the pathogenesis of X-linked juvenile RS by searching mutations in RS patients. Future experiments, including the identification of sub-populations of sensory neurons expressing PPEF phosphatase and of the natural substrate of the protein, will greatly contribute to the understanding of the biological role of the *PPEF* gene in mammalian development and will allow testing for its possible involvement in NGF signal transduction.

MATERIALS AND METHODS

cDNA identification

YAC clone 939H7 was converted into cosmid clones by hybridization of long range *Alu*-PCR product obtained with a variety of human-specific *Alu* primers. PCR amplification was performed according to Gu *et al.* (27). Cosmid clones were grouped (10 clones per group), digested with *Bam*HI/*Bgl*II, cloned in the pSPL3 vector, and used for the exon amplification experiments as described previously (28) and by Montini *et al.* (manuscript in preparation). In order to identify the *PPEF* full-length transcript, three human cDNA libraries were screened: a teratocarcinoma/neuron cDNA library (mature hNT neuron, Stratagene 937233), a fetal brain cDNA library (Clontech HL3003a), and the Bento Soares infant brain 1NIB arrayed cDNA library. For the isolation of a partial murine cDNA clone, an 11.5 day embryo (Clontech ML3003a) mouse cDNA library was used. Plating, hybridization and washing conditions were performed as previously described (29). Primers used to obtain the probe used to screen the mouse cDNA library were as follows:

c14, AAGTCCTGAAGCAAATGCCG;
c22, GCCATACCTCAGATTCATC.

cDNA sequence analysis

cDNA sequence analysis and nucleotide and protein database searches were performed as previously described (4). Data on similarity/identity were obtained using the Bestfit program of the GCG software package, version 8.1. The multiple alignment analyses were generated using the PileUp program of the Wisconsin GCG software package, ver. 8.1.

Expression studies

Commercial Northern blots (Clontech) containing human RNA from fetal and adult tissues were hybridized and washed using the conditions recommended by the manufacturer. Mouse embryo tissue sections were prepared and RNA *in situ* hybridization experiments were performed as previously described (30). cDNA clone EM800 was linearized with appropriate restriction enzymes to transcribe either sense or antisense ^{35}S-labelled riboprobes. Slides were exposed for 10 days. Micrographs are double

exposures: red represents the *in situ* hybridization signal, and blue shows the nuclei stained with Hoechst 33258 dye.

RNA isolation and RT-PCR

RNA isolation from blood and RT-PCR experiments were carried out as previously described (31). RT-PCR was carried out with nested primers. Primers 1F+1R were used for the first round, the resulting PCR products were then reamplified using primers 2F+2R. Primers used for RT-PCR:

1F, 5′-TGAAGGCCAGACAACACTATG-3′;
1R, 5′-ACTAGGTCCAACCCAGTTTCT-3′;
2F, 5′-TAATACGACTCACTATAGGAACAGACCACCAT GGAATATGCTGATGAACAAGGC-3′
2R, 5′-CTTTTCTCTGCTACTGACTATGAA-3′.

Primers for sequencing:

909, 5′-CTTGGAAGAATTCTATGCCTGG-3′;
947, 5′-ATTGTACCGATTGGGAGCCA-3′;
1528,5′-CCATAGTATCCACTCTTTGGCGA-3′;
1501,5′-CCTCTTCGCCAAAGAGTGGATA-3′.

Linkage mapping in mouse

Genetic mapping was achieved utilizing a (C57BL/6j × SPRET/Ei)F1 × SPRET/Ei (BSS) backcross generated and distributed by the Jackson Laboratory (Bar Harbor, ME) (13). An *Msp*I RFLP was identified by hybridization of C57BL/6j and SPRET/Ei parental DNAs cut with each of the six restriction enzymes (*Eco*RI, *Eco*RV, *Kpn*I, *Msp*I, *Taq*I and *Xba*I). Four Southern panels containing *Msp*I-cut parental DNAs and N2 progeny ($n = 94$) DNAs were hybridized with a *Ppef* cDNA probe. The resulting strain distribution pattern (SDP) was analyzed with the Map Manager 2.6 program (32).

ACKNOWLEDGEMENTS

We thank Drs A. A. B. Bergen, B. Dallapiccola, A. Gal and A. Hanauer for providing us with samples from RS patients. We thank the Netherlands Organisation for Scientific Research (NWO) (J.T.d.D. and E.v.d.V.). We thank the TIGEM sequencing facilities for technical support and M. Smith for preparation of the manuscript. This work was supported by the Italian Telethon Foundation and the EC under Grant Nos BMH4-CT96-1134 and BMH4-CT96-0889.

REFERENCES

1. Ferrero, G.B., Franco, B., Roth, E.J., Firulli, B.A., Borsani, G., Delmas-Mata, J., Weissenbach, J., Halley, G., Schlessinger, D., Chinault, A.C., Zoghbi, H.Y., L., N.D. and Ballabio, A. (1995) An integrated physical and genetic map of a 35 Mb region on chromosome Xp22.3–Xp21.3. *Hum. Mol. Genet.* **4**, 1821–1827.
2. Steele, F.R., Washburn, T., Rieger, R. and O'Tousa, J.E. (1992) Drosophila *retinal degeneration C (rdgC)* encodes a novel serine/threonine protein phosphatase. *Cell* **69**, 669–676.
3. Steele, F. and O'Tousa, J.E. (1990) Rhodopsin activation causes retinal degeneration in drosophila *rdgC* mutant. *Neuron* **4**, 883–890.
4. Banfi, S., Borsani, G., Rossi, E., Bernard, L., Guffanti, A., Rubboli, F., Marchitiello, A., Giglio, S., Coluccia, E., Zollo, M., Zuffardi, O. and Ballabio, A. (1996) Identification and mapping of human cDNAs homologous to *Drosophila* mutant genes through EST database searching. *Nature Genet.* **13**, 167–174.
5. Soares, M.B., Bonaldo, M.F., Jelene, P., Su, L., Lawton, L. and Efstratiadid, A. (1994) Construction and characterization of a normalized cDNA library. *Proc. Natl. Acad. Sci. USA* **91**, 9228–9232.
6. Kozak, M. (1984) Compilation and analysis of sequences upstream from the translational start site in eukaryotic mRNAs. *Nucleic Acids Res.* **12**, 857–873.
7. Kretsinger, R.H. (1980) Structure and evolution of calcium-modulated proteins. *CRC Crit. Rev. Biochem.* **8**, 119–174.
8. Barton, G.J., Cohen, P.T.W. and Barford, D. (1994) Conservation analysis and structure prediction of the protein serine/threonine phosphatases: Sequence similarity with diadenosine tetraphosphatase from *Escherichia coli* suggests homology to the protein phosphatases. *Eur. J. Biochem.* **220**, 225–237.
9. Barford, D. (1996) Molecular mechanisms of the protein serine/threonine phosphatases. *Trends Biochem. Sci.* **251**, 407–412.
10. Shapiro, M.B. and Senapathy, P. (1987) RNA splice junctions of different classes of eukaryotes: sequence statistics and functional implications in gene expression. *Nucleic Acids Res.* **15**, 7155–7174.
11. Le Douarin, N.M., Kalcheim, C. and Teillet, M.-A. (1992). The cellular and molecular basis of early sensory ganglion development. In Scott, S.A. (ed.), *Sensory Neurons: Diversity, Development, and Plasticity.* Oxford University Press, New York, pp. 143–170.
12. Altman, J. and Bayer, S.A. (1984) The development of the rat spinal cord. *Adv. Anat. Embryol. Cell. Biol.* **85**, 1–164.
13. Rowe, L.B., Nadeau, J.H., Turner, R., Frankel, W.N., Letts, V.A., Eppig, J.T., Ko, M.S.H., Thurston, S.J. and Birkenmeier, E.H. (1994) Maps from two interspecific backcross DNA panels available as a community genetic mapping resource. *Mamm. Genome* **5**, 253–274.
14. Dinulos, M.B., Bassi, M.T., Rugarli, E.I., Chapman, V., Ballabio, A. and Disteche, C.M. (1996) A new region of conservation is defined between human and mouse X chromosomes. *Genomics* **35**, 244–247.
15. Fischer, E.H. and Krebs, E.G. (1990) *Biochim. Biophys. Acta* **1000**, 297–301.
16. Cohen, P.T.W., Brewis, N.D., Hughes, V. and Mann, D.J. (1990) Protein serine/threonine phosphatases; an expanding family. *FEBS Lett.* **268**, 355–359.
17. Chen, M.X., McPartlin, A.E., Brown, L., Chen, Y.H., Barker, H.M. and Cohen, P.T.W. (1994) A novel human protein serine/threonine phosphatase, which possesses four tetratricopeptide repeat motifs and localizes to the nucleus. *EMBO J.* **13**, 4278–4290.
18. Herman, G.E., Boyd, Y., Brown, S.D.M., de Gouyon, B., Haynes, A. and Quaderi, N. (1996) Encyclopedia of the mouse genome V. Mouse X chromosome. *Mamm. Genome* **6**, S317–S330.
19. Lawson, S.N. (1992). Morphological and biochemical cell types of sensory neurons. In Scott, S.A. (ed.), *Sensory Neurons: Diversity, Development, and Plasticity.* Oxford University Press, New York, pp. 27–59.
20. Hunt, S.P., Mantyh, P.W. and Priestley, J.V. (1992) The organization of biochemically characterized sensory neurons. In Scott, S.A. (ed.), *Sensory Neurons: Diversity, Development, and Plasticity.* Oxford University Press, New York, pp. 60–76.
21. Martin-Zanca, D., Barbacid, M. and Parada, L.F. (1990) Expression of the *trk* proto-oncogene is restricted to the sensory cranial and spinal ganglia of neural crest origin in mouse development. *Genes Dev.* **4**, 683–694.
22. McMahon, S.B., Armanini, M.P., Ling, L.H. and Phillips, H.S. (1994) Expression and coexpression of Trk receptors in subpopulations of adult primary sensory neurons projecting to identified peripheral targets. *Neuron* **12**, 1161–1171.
23. Snider, W.D. and Wright, D.E. (1996) Neurotrophins cause a new sensation. *Neuron* **16**, 229–232.
24. Lindsay, R.M., Thoenen, H. and Barde, Y.-A. (1985) Placode and neural crest-derived sensory neurons are responsive at early developmental stages to brain-derived neurotrophic factor. *Dev. Biol.* **112**, 319–328.
25. Davies, A.M. and Lindsay, R.M. (1985) The cranial sensory ganglia in culture: differences in the response of placode-derived and neural crest-derived neurons to nerve growth factor. *Dev. Biol.* **111**, 62–72.
26. Bothwell, M. (1995) Functional interactions of neurotrophins and neurotrophin receptors. *Annu. Rev. Neurosci.* **18**, 223–253.
27. Gu, Y., Ferrero, G.B., Schofield, T., Franco, B., Lee, C.C., Graves, M., Arenson, A., Ballabio, A., Chinault, A.C. and Nelson, D.L. (1995) Towards a cosmid-based map of the distal short arm of the human X chromosome. Abstract presented at the 6th X Chromosome Workshop, Banff, Alberta, Canada, pp. B3.
28. Church, D.M., Stotler, C.J., Rutter, J.L., Murrell, J.R., Trofatter, J.A. and Buckler, A.J. (1994) Isolation of genes from complex sources of mammalian genomic DNA using exon amplification. *Nature Genet.* **6**, 98–105.

29. Franco, B., Guioli, S., Pragliola, A., Incerti, B., Bardoni, B., Tonlorenzi, R., Carrozzo, R., Maestrini, E., Pieretti, M., Taillon-Miller, P., Brown, C.J., Willard, H.F., Lawrence, C., Persico, M.G., Camerino, G. and Ballabio, A. (1991) A gene deleted in Kallmann's syndrome shares homology with neural cell adhesion and axonal path-finding molecules. *Nature* **353**, 529–536.
30. Rugarli, E.I., Lutz, B., Kuratani, S.C., Wawersik, S., Borsani, G., Ballabio, A. and Eichele, G. (1993) Expression pattern of the Kallmann syndrome gene in the olfactory system suggests a role in neuronal targeting. *Nature Genet.* **4**, 19–25.
31. den Dunnen, J.T., Roest, P.A.M., Van der Luijt, R.B. and Hogervorst, F.B.L. (1996). The Protein Truncation Test (PTT) for rapid detection of translation-terminating mutations. In Pfeifer, G.P. (ed.), *Technologies for Detection of DNA Damage and Mutations*. Plenum Press, New York, pp. 323–341.
32. Manly, K.F. and Elliott, R.W. (1991) RI manager, a microcomputer program for the analysis of data from recombinant inbred strains. *Mamm. Genome* **1**, 123–126.