PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 :
A61K 38/57, C07K 16/38
G01N 33/53 // C07K 14/81

A1

(11) International Publication Number: **WO 95/33480**

(43) International Publication Date: 14 December 1995 (14.12.95)

(21) International Application Number: PCT/US95/07201

(22) International Filing Date: 6 June 1995 (06.06.95)

(30) Priority Data:

08/257,963	7 June 1994 (07.06.94)	US
08/367,841	30 December 1994 (30.12.94)	US

(81) Designated States: AM, AT, AU, BB, BG, BR, BY, CA, CH, CN, CZ, DE, DK, EE, ES, FI, GB, GE, HU, IS, JP, KE, KG, KP, KR, KZ, LK, LR, LT, LU, LV, MD, MG, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, TJ, TT, UA, UG, UZ, VN, European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG), ARIPO patent (KE, MW, SD, SZ, UG).

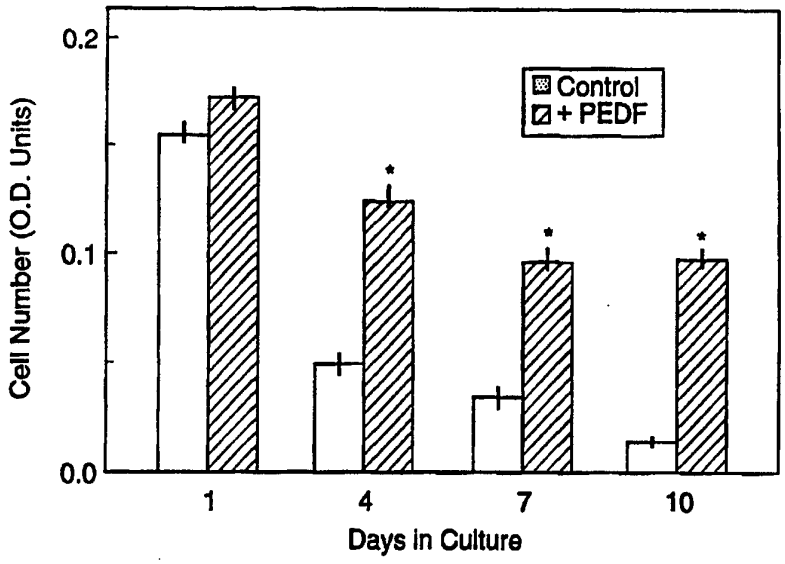
Published
With international search report.
Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(71) Applicant: THE GOVERNMENT OF THE UNITED STATES OF AMERICA, represented by THE SECRETARY, DEPARTMENT OF HEALTH AND HUMAN SERVICES [US/US]; Office of Technology Transfer, National Institutes of Health, Suite 325, Box 13, 6011 Executive Boulevard Rockville, MD 20852 (US).

(72) Inventors: CHADER, Gerald, J.; 9701 Singleton Drive, Bethesda, MD 20817 (US). BECERRA, Sofia, Patricia; 6218 Stoneham Court, Bethesda, MD 20817 (US). SCHWARTZ, Joan, P.; 6411 Wilson Lane, Bethesda, MD 20817 (US). TANIWAKI, Takayuki; 257 Congressional Lane, Rockville, MD 20852 (US).

(74) Agent: FEILER, William, S.; Morgan & Finnegan, L.L.P., 345 Park Avenue, New York, NY 10154 (US).

(54) Title: PIGMENT EPITHELIUM-DERIVED FACTOR: CHARACTERIZATION, GENOMIC ORGANIZATION AND SEQUENCE OF THE PEDF GENE



(57) Abstract

Nucleic acids encoding the neurotrophic protein known as pigment epithelium-derived factor (PEDF), a truncated version of PEDF referred to as rPEDF, and equivalent proteins, vectors comprising such nucleic acids, host cells into which such vectors have been introduced, recombinant methods for producing PEDF, rPEDF, and equivalent proteins, the rPEDF protein and equivalent proteins of rPEDF and PEDF -BP, -BX and BA, and the PEDF protein produced by recombinant methods. Effects and use of these variants on: 1) neuronal differentiation (neurotrophic effect), 2) neuron survival (neuronotrophic effect), and 3) glial inhibition (gliastatic effect) are described.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	**GB**	United Kingdom	**MR**	Mauritania
AU	Australia	**GE**	Georgia	**MW**	Malawi
BB	Barbados	**GN**	Guinea	**NE**	Niger
BE	Belgium	**GR**	Greece	**NL**	Netherlands
BF	Burkina Faso	**HU**	Hungary	**NO**	Norway
BG	Bulgaria	**IE**	Ireland	**NZ**	New Zealand
BJ	Benin	**IT**	Italy	**PL**	Poland
BR	Brazil	**JP**	Japan	**PT**	Portugal
BY	Belarus	**KE**	Kenya	**RO**	Romania
CA	Canada	**KG**	Kyrgystan	**RU**	Russian Federation
CF	Central African Republic	**KP**	Democratic People's Republic of Korea	**SD**	Sudan
CG	Congo			**SE**	Sweden
CH	Switzerland	**KR**	Republic of Korea	**SI**	Slovenia
CI	Côte d'Ivoire	**KZ**	Kazakhstan	**SK**	Slovakia
CM	Cameroon	**LI**	Liechtenstein	**SN**	Senegal
CN	China	**LK**	Sri Lanka	**TD**	Chad
CS	Czechoslovakia	**LU**	Luxembourg	**TG**	Togo
CZ	Czech Republic	**LV**	Latvia	**TJ**	Tajikistan
DE	Germany	**MC**	Monaco	**TT**	Trinidad and Tobago
DK	Denmark	**MD**	Republic of Moldova	**UA**	Ukraine
ES	Spain	**MG**	Madagascar	**US**	United States of America
FI	Finland	**ML**	Mali	**UZ**	Uzbekistan
FR	France	**MN**	Mongolia	**VN**	Viet Nam
GA	Gabon				

Pigment Epithelium-Derived Factor:

Characterization, Genomic Organization and Sequence of the PEDF gene

This application is a continuation-in-part of application Serial No. 08/257,963 filed on June 07, 1994, which is a continuation-in-part of application Serial No. 07/952,796 filed on September 24, 1992.

TECHNICAL FIELD OF THE INVENTION

This invention relates to a neurotrophic, neuronotrophic and gliastatic protein. More specifically, this invention relates to the biological properties of a protein known as pigment epithelium-derived factor (PEDF) and recombinant forms of the protein. This invention also relates to a truncated version of PEDF that is referred to as rPEDF. In addition to PEDF and rPEDF and functionally equivalent proteins, this invention relates to nucleic acids that encode rPEDF, and fragments thereof, to vectors comprising such nucleic acids, to host cells into which such vectors have been introduced, and to the use of these host cells to produce such proteins.

BACKGROUND OF THE INVENTION

Pigment epithelium-derived factor, otherwise known as pigment epithelium differentiation-factor, was identified in the conditioned medium of cultured fetal human retinal pigment epithelial cells as an extracellular neurotrophic agent capable of inducing neurite outgrowth in cultured human retinoblastoma cells (Tombran-Tink et al. (1989) *Invest. Ophthalmol. Vis. Sci.*, 30 (8), 1700-1707). The source of PEDF, namely the retinal pigment epithelium (RPE), may be crucial to the normal development and function of the neural retina. A variety of molecules, including growth factors, are synthesized and secreted by RPE cells. Given that the RPE develops prior to and lies adjacent to the neural retina, and that it functions as part of the blood-retina barrier (Fine et al. (1979) The Retina, Ocular Histology: A Text and Atlas, New

York, Harper & Row, 61-70), the RPE has been implicated in vascular, inflammatory, degenerative, and dystrophic diseases of the eye (Elner et al. (1990) *Am. J. Pathol.*, 136, 745-750). In addition to growth factors, nutrients and metabolites are also exchanged between the RPE and the retina. For example, the RPE supplies to the retina the well-known growth factors PDGF, FGF, TGF-α, and TGF-β (Campochiaro et al. (1988) *Invest. Ophthalmol. Vis. Sci.*, 29, 305-311; Plouet (1988) *Invest. Ophthalmol. Vis. Sci.*, 29, 106-114; Fassio et al. (1988) *Invest. Ophthalmol. Vis. Sci.*, 29, 242-250; Connor et al. (1988) *Invest. Ophthalmol. Vis. Sci.*, 29, 307-313). It is very likely that these and other unknown factors supplied by the RPE influence the organization, differentiation, and normal functioning of the retina.

In order to study and determine the effects of putative differentiation factors secreted by the RPE, cultured cells have been subjected to retinal extracts and conditioned medium obtained from cultures of human fetal RPE cells. For example, U.S. Patent No. 4,996,159 (Glaser) discloses a neovascularization inhibitor recovered from RPE cells that is of a molecular weight of about 57,000 +/- 3,000. Similarly, U.S. Patent Nos. 1,700,691 (Stuart), 4,477,435 (Courtois et al.), and 4,670,257 (Guedon born Saglier et al.) disclose retinal extracts and the use of these extracts for cellular regeneration and treatment of ocular disease. Furthermore, U.S. Patent Nos. 4,770,877 (Jacobson) and 4,534,967 (Jacobson et al.) describe cell proliferation inhibitors purified from the posterior portion of bovine vitreous humor.

PEDF only recently has been isolated from human RPE as a 50-kDa protein (Tombran-Tink et al. (1989) *Invest. Ophthalmol. Vis. Sci.*, 29, 414; Tombran-Tink et al. (1989) *Invest. Ophthalmol. Vis. Sci.*, 30, 1700-1707; Tombran-Tink et al. (1991) *Exp. Eye Res.*, 53, 411-414).

Specifically, PEDF has been demonstrated to induce the differentiation of human Y79 retinoblastoma cells, which are a neoplastic counterpart of normal retinoblasts (Chader (1987) *Cell Different.*, 20, 209-216). The differentiative changes induced by PEDF include the extension of a complex meshwork of neurites, and expression of neuronal markers such as neuron-specific enolase and neurofilament proteins. This is why the synthesis and secretion of PEDF protein by the RPE is believed to influence the development and differentiation of the neural retina. Furthermore, PEDF is only highly expressed in undifferentiated human retinal cells, like Y79 retinoblastoma cells, but is either absent or downregulated in their differentiated counterparts. Recently, it was reported that PEDF mRNA is expressed in abundance in quiescent human fetal W1 fibroblast cells and not expressed in their senescent counterparts (Pignolo et al., 1993).

Further study of PEDF and examination of its potential therapeutic use in the treatment of inflammatory, vascular, degenerative, and dystrophic diseases of the retina and central nervous system (CNS) necessitates the obtention of large quantities of PEDF. Unfortunately, the low abundance of PEDF in fetal human eye and furthermore, the rare availability of its source tissue, especially in light of restrictions on the use of fetal tissue in research and therapeutic applications, make further study of PEDF difficult at best. Therefore, there remains a need for large quantities of PEDF and equivalent proteins. Accordingly, the obtention of nucleic acids that encode PEDF and equivalent proteins, and the capacity to produce PEDF and equivalent proteins in large quantities would significantly impact upon the further study of PEDF, its structure, biochemical activity and cellular function, as well as the discovery and design of therapeutic uses for PEDF.

SUMMARY OF THE INVENTION

It is an object of the present invention to provide nucleic acids encoding for PEDF and functional fragments thereof, vectors comprising such nucleic acids, host cells into which such vectors have been introduced, and a recombinant method of producing PEDF and equivalent proteins. It is another object of the present invention to obtain the genomic DNA sequences encoding for PEDF, identify the intron-exon junctions, the chromosome location in the human genome, and to provide the regulatory regions of the gene which flank the genomic sequence. The present invention relates to such genomic PEDF DNA.

It is a further object of the present invention to provide structural characteristics of PEDF and its similarities to the serpin family of serine protease inhibitors, both structural and functional.

It is yet another object of the present invention to provide PEDF and equivalent proteins produced in accordance with such a recombinant method, wherein the PEDF and equivalent proteins so produced are free from the risks associated with the isolation of PEDF from naturally-occurring source organisms.

Another object of the present invention is to provide nucleic acids for a truncated version of PEDF, referred to as rPEDF, and equivalent proteins, vectors comprising such nucleic acids, host cells into which such vectors have been introduced, and a recombinant method of producing rPEDF and equivalent proteins. It is also an object of the present invention to provide rPEDF and equivalent proteins produced in accordance with such a recombinant method.

It is a further object of the invention to provide a PEDF protein having neuronotrophic and gliastatic activity. The neuronotrophic activity is seen in the prolonged survival of neuronal cells. The

gliastatic activity is observed in the inhibition of growth of glial cells in the presence of PEDF or active fragment thereof. It is another object of the invention to provide methods for treating neuronal cells so as to promote/enhance neuron survival and prevent growth of glial cells, comprising treating such cell populations with an effective amount of PEDF or an active fragment thereof.

It is yet another object of the present invention to provide antibodies which specifically recognize PEDF, either monoclonal or polyclonal antibodies, raised against native protein, the recombinant protein or an immunoreactive fragment thereof. It is an object of the invention to provide methods for detecting PEDF by immunoassay using such antibody preparation in determining aging and/or other degenerative diseases. Another object of the invention relates to a method of using PEDF antibodies to specifically inhibit PEDF activity.

These and other objects and advantages of the present invention, as well as additional inventive features, will be apparent from the description of the invention provided herein.

Descriptions of the Figures

Figure 1: Human PEDF Gene Structure: Restriction map and organization of the human PEDF gene. Exons 1-8 are indicated by black boxes and numbered 1-8. Introns and flanking DNA are represented by horizontal line and are labeled A-G. Positions of several genomic clones are shown above and below the diagrammed gene. Recognition sites for the restriction endonuclease, NotI ("N"), BamHI ("B") and EcoRI ("E") are indicated by vertical arrows.

Figure 2: Southern analysis of human genomic DNA (A) and P147 (B) restricted with Bam HI, EcoRI, HindIII and PstI endonuclease. Southern membranes from

Pulsed-field electrophoretic gel profiles were probed with radioactively labelled PEDF cDNA. The pattern of hybridization of P147 DNA is consistent with total human genomic DNA. Size markers are indicated.

Figure 3: 5' Flanking region of the PEDF gene. The first exon (capital letters) and the first 1050 bp of 5 prime flanking region are shown. Two Alu repetitive sequences are underlined. Possible binding sites for HNF-1, PEA3, Octomer (Oct), c/EBP are underlined and labeled. The putative AP-1 sites are shown in bold, and TREp/RAR are double underlined. The underlined (dashed) sequence in exon 1 was determined by the 5' RACE.

Figure 4: Northern Blot analysis of PEDF mRNA: Gene expression analysis of the human PEDF transcript in a number of human adult and fetal tissues. Tissues from which RNA was obtained are shown above corresponding lanes. Membranes contain 2 ug poly (A) RNA for each sample and were probed with radioactively labelled cDNA for human PEDF. A single 1.5 kb transcript is seen in both adult and fetal tissues with the greatest intensity of hybridization in liver, testis, skeletal muscle and ovary while the signal for brain, pancreas and thymus was significantly weaker than that for other tissues. No significant signal was detected for adult kidney and spleen. A significant difference in PEDF mRNA levels seen between adult and fetal kidney.

Figure 5: Evolutionary relatedness of the Human PEDF gene: Each lane represents a total of 8 ug of genomic DNA for each species digested with Eco RI. Southern blot analysis is shown with a PEDF probe. Hybridization signals for chicken (A), mammals (B) and primates (C) is shown. A large fragment of approximately 23 kb is seen in all primates and many mammalian species. In addition several polymorphisms are seen in the different mammalian species examined.

Figure 6A & 6B: Relationship between cell density plated and optical density measured by MTS assay. Different concentrations of postnatal-day 8 cerebellar granule cells were added to 96 well plate and cultured in serum-containing medium (6A), or chemically defined medium (6B). Optical density was measured on days in vitro (DIV) 1, 4, or 7. Square, DIV 1; Solid circle, DIV 4; Open circle, DIV7. The data are plotted as function of cell density (n=6).

Figure 7: Time course for PEDF stimulation of cell survival in chemically-defined medium. Postnatal-day 8 cerebellar granule cells were cultured in 96 well plate. PEDF was added at DIV 0 and the optical density was then measured on DIV 1, 4, 7, or 10. Solid bar, control; cross-hatched bar, PEDF treated (50ng/ml); striped bar, PEDF treated (500ng/ml). The data are expressed as optical density/well (means±SEM, n=6). Statistical analysis was done by two way ANOVA post-hoc Scheefe test. **P<0.0001 versus control.

Figure 8: Dose-response curve for PEDF in chemically defined medium. Different concentrations of PEDF were added on DIV 0 and MTS assay was carried out on DIV 7. The data are expressed as ratio to control (mean ± SEM, n=6). Statistical analysis was done by one way ANOVA post-hoc Scheffe F test. **P<0.0001 vesus control.

Figure 9: MTS assay of postnatal day 5 cerebellar granule cells at DIV 1 and DIV 2. Postnatal-day 5 cerebellar granule cells were cultured in 96 well plate using serum-containing medium without Ara-C (A), or chemically defined medium without F12(B). The MTS assay was carried out on DIV 1 and 2. Solid bar, control; Striped bar, PEDF treated (500ng/ml). The data are expressed as optical density/well (means ± SEM, n=6). Statistical analysis was done by two way ANOVA post-hoc Scheffe F test. **P<0.0005 vesus control.

Figure 10: BrdU incorporation into postnatal day 5 cerebellar granule cells. Postnatal-day 5 cerebellar granule cells were cultured in a 96 well plate using serum-containing medium (SCM) without Ara-C, or chemically defined medium (CDM) without F12. PEDF was added on DIV 0, BrdU was added on DIV 1 and the cells were fixed on DIV 2. Solid bar, control; Striped bar, PEDF treated (500ng/ml). The number of labeled nucleic acids are expressed as a percentage of total cell population (mean ± SEM). For each value, 3000 cells was counted at least.

Figure 11: Relationship between cell density and neurofilament content measured by ELISA. Different concentrations of postnatal-day 8 cerebellar granule cells are added to 96 wells and cultured. Optical density was measured on DIV 7. The data are plotted as a function of cell density.

Figure 12: Neurofilament ELISA assay in postnatal-day 8 cerebellar granule cells. Cells were cultured in a 96 well plate with or without PEDF using serum-containing medium (SCM) or chemically defined medium (CDM). After fixing cells on DIV 7, the neurofilament ELISA was carried out and the data are expressed as ratio to control (mean ± SEM, n=6 to 10). Solid bar, control; Striped bar, PEDF treated (500ng/ml). Statistical analysis was done by two way ANOVA post-hoc Scheffe F test. *P <0.05 vesus control.

Figure 13: Summary of PEDF neuronotrophic effects through 10 days in culture.

Figure 14: Effects of truncated peptides BP and BX on CGC viability.

Figure 15: Effect of PEDF on astroglia from cerebellum.

Figure 16: Effect of PEDF on cerebellar microglia.

Figure 17: Purification of PEDF-immunoreactive protein from bovine IPM. Washes of bovine IPM were subjected to A) TSK-3000 size-exclusion chromatography followed by B) Mono-S chromatography. Western blot inserts demonstrate the fractions containing PEDF.

Figure 18: Enzymatic deglycosylation of PEDF as demonstrated by Western blotting. PEDF treatment is given at the top of each lane. Numbers indicate positions of mol. wt. standards.

Figure 19: Antibody to rPEDF specifically recognizes native PEDF at a high titer. A) Western blot demonstrating effectiveness of the antibody to at least 1:50,000 dilution and that addition of excess rPEDF completely blocks band visualization. B) Slot-blot analysis shows the ability to detect $\leq$ 1 ng of native bovine PEDF protein.

Figure 20: Negative effect of PEDF antibody on neurite extension in Y-79 cells. Top row: bovine serum albumin (BSA) control cultures. Middle row: antibody effect on neurite-induction by native bovine PEDF protein. Bottom row: antibody effect on neurite induction by interphotoreceptor matrix (IPM).

Figure 21: Phase microscopy analysis of neurite outgrowth in the presence or absence of PEDF.

Figure 22: Phase microscopy analysis of neurite outgrowth in the presence of recombinant PEDF and native, isolated PEDF.

Figure 23: Schematic Diagram of C-terminal deletions of rPEDF.

DETAILED DESCRIPTION OF THE INVENTION

The present invention relates to a protein having novel, important and unobvious properties. Pigment epithelium-derived factor (PEDF) is a protein having neurotrophic, neuronotrophic and gliastatic characteristics. The present invention further relates to the DNA sequences coding for the PEDF gene, the genomic

DNA containing the PEDF gene and fragments of the PEDF gene encoding for protein fragments of PEDF having biological activity.

"Neurotrophic" activity is defined herein as the ability to induce differentiation of a neuronal cell population. For example, PEDF's ability to induce differentiation in cultured retinoblastoma cells is considered neurotrophic activity.

"Neuronotrophic" activity is defined herein as the ability to enhance survival of neuronal cell populations. For example, PEDF's ability to act as a neuron survival factor on neuronal cells is neuronotrophic activity.

"Gliastatic" activity is defined herein as the ability to inhibit glial cell growth and proliferation. For example, PEDF's ability to prevent growth and/or proliferation of glial cells is gliastatic activity.

Based upon the protein amino acid sequence elucidated in the present invention, PEDF has been found to have extensive sequence homology with the serpin gene family, members of which are serine protease inhibitors. Many members of this family have a strictly conserved domain at the carboxyl terminus which serves as the reactive site of the protein. These proteins are thus thought to be derived from a common ancestral gene. However the developmental regulation differs greatly among members of the serpin gene family and many have deviated from the classical protease inhibitory activity (Bock (1990) Plenum Press, New York Bock, S.C., *Protein Eng.* 4, 107-108; Stein et al. (1989) *Biochem. J.* 262, 103-107). Although PEDF shares sequence homology with serpins, analysis of the cDNA sequence indicates that it lacks the conserved domain and thus may not function as a classical protease inhibitor.

Genomic sequencing and analysis of PEDF has provided sequences of introns and exons as well as

approximately 4 kb of 5'-upstream sequence. The present invention demonstrates the localization of the gene for PEDF to 17p13.1 using both in situ hybridization and analyses of somatic cell hybrid panels (Tombran-Tink, et al., (1994) *Genomics*, 19:266-272). This is very close to the p53 tumor suppressor gene as well as to the chromosomal localization of a number of hereditary cancers unrelated to mutations in the p53 gene product. PEDF thus becomes a prime candidate gene for these cancers.

The full length genomic PEDF sequence is represented by SEQ ID NO:43. The PEDF gene encompasses approximately 16 Kb and contains 8 exons all of which have conventional consensus splice-sites. The 5' flanking region of the PEDF gene contains two Alu repetitive elements which cover approximately two thirds of the first 1050 bp of the putative promoter sequence. There are also several sequence motifs which may be recognized by members of several families of transcription factors. The presence of two possible binding sites for the ubiquitous octamer family of transcription factors, may explain the presence of PEDF in most tissues tested. The presence of other more specific elements, however, suggests that PEDF is under precise control and supports previous work including its effects on such diverse processes as neuronal differentiation and fibroblast senescence.

The genomic PEDF sequence or fragments thereof are useful as a probe for detecting the gene in a cell. In addition, such a probe is useful in a kit for identification of a cell type carrying the gene. Mutations, deletions or other alternations in the gene organization can be detected through the use of a DNA probe derived from the PEDF genomic sequence.

Tissue Distribution

Although PEDF is particularly highly expressed by RPE cells, it is detectable in most tissues, cell types, tumors, etc. by Northern and Western blot analyses.

It is readily detected, for example in vitreous and aqueous humors. The important question of subcellular localization of PEDF has also been addressed. Although the bulk of the PEDF appears to be secreted, we have used a PEDF antibody to probe cultured monkey RPE cells and found that PEDF is associated with the nucleus as well as with very specific cytoskeletal structures in the cytoplasm. Importantly, this varies as to the age of the cells and the specific cell-cycle state examined. For example, the protein appears to concentrate at the tips of the pseudopods of primate RPE cells that interact with the substratum during the initial stages of attachment. Later though, this staining disappears and there is appearance of the protein in association with specific cytoskeletal structures and the nucleus. Thus it appears that PEDF plays an important intracellular role in both nucleus and cytoplasm.

Involvement in Cell Cycle

The present invention indicates that there is expression in dividing, undifferentiated Y-79 cells and little or no expression in their quiescent, differentiated counterparts (Tombran-Tink, et al. (1994) *Genomics*, 19:266-272). Pignolo et al. (1993) *J. Biol. Chem.*, 268:2949-295) have demonstrated that the synthesis of PEDF in WI-38 fibroblast cells is restricted to the G_0 stage of the cell cycle in young cells. Moreover, in old senescent cells, PEDF messenger RNA is absent.

Production of Recombinant PEDF.

Segmentation of the PEDF polypeptide is basic to studies on structure-function. For this purpose, expression vectors containing fragments of PEDF coding sequences provide an excellent source for synthesizing and isolating different regions of the PEDF polypeptide. Expression of human fetal PEDF sequences was achieved with *E. coli* expression vectors and the human fetal PEDF cDNA. We have shown that the recombinant PEDF product (rPEDF) is

a biologically-active neurotrophic factor and is obtained in yields on the order of 1.3 mg/g of wet E. coli cells. Truncated peptides can also be made from appropriate molecular biological constructs and expressed in E. coli. Using these products, we have evidence that two distinct regions on the PEDF primary structure can be distinguished: 1) an "active site" conferring neurotrophic activity on the molecule that is located within amino acid residues 44-121 near the N-terminal of the protein and 2) a region near the C-terminal with homology to a serpin exposed loop i.e., the "classical" serpin active site. These results suggest 1) that the overall native conformation of PEDF is not required for neurite outgrowth and 2) that inhibition of serine proteases can not account for the biological activity of PEDF. We now have a series of truncated rPEDF constructs that span the protein sequence and can pinpoint the specific neurotrophic "active site" near the N-terminal.

Characterization with a highly specific polyclonal antibody.

Purified recombinant human PEDF was used to develop a polyclonal antibody ("Anti-rPEDF") that specifically blocks the PEDF-mediate neurotrophic activity. Furthermore, the anti-rPEDF completely blocks the IPM-induced neurotrophic activity.

Neuronotrophic properties of PEDF

In addition to demonstrating that native PEDF and rPEDF are neurotrophic in the Y-79 and Weri tumor cell systems, the present invention determined whether PEDF had an effect on normal neurons in primary culture. For this purpose, studies were conducted using cultures of normal cerebellar granule cells (CGCs) prepared from the 8-day postnatal rat. Cells treated with rPEDF did not respond to treatment by exhibiting a more neuronal morphological appearance. However, PEDF had a large effect on granule cell survival. Since these cells are not tumorous or

transformed cells, they have a finite life in culture, dying in about 21 days depending on the culture medium. PEDF-treated culture, however, contained up to 10-fold more cells after 10 days of culture in serum-free medium compared to non-treated culture (Figure 4). These results were determined; 1) by direct microscopic observation and cell counting and 2) use of an MTS (tetrazolium/formazan) assay which determines live cell numbers (See example 11). Thus, PEDF has a dramatic effect on CNS neuron survival and should be added to the short list of newly-emerging "neuronotrophic" proteins.

In General Tissue Culture Research:

Two problems that generally plague any tissue culture experiment using neurons and glia is that the neurons tend to die quickly and that glia tend to overrun the culture dish. PEDF or its peptides can help in both regards. Thus, one commercial use of PEDF might be as a general culture medium additive when CNS cells are to be cultured.

In CNS Transplantation Studies:

It is thought that transplantation of neurons may cure certain pathologies. For example, in Parkinson's disease, transplantation of specific fetal brain cells into patients could alleviate or cure the problems associated with the disease. One of the major problems to contend with, though, would be to prolong the life of the transplanted cells and to keep them differentiated, e.g. secreting the proper substances, etc. Pretreatment of the cells with PEDF could aid in both of these areas. Similarly, transfection of either neurons or astroglia with the PEDF gene before implantation can be a long-term source of PEDF at the transplantation site.

There is much activity in attempts at transplantation of neural retina and photoreceptor cells to help cure blindness. Attempts to date have not been fruitful both due to non-differentiation and death of the

grafts. Again, PEDF may help in both regards. Specifically, photoreceptor neurons to be transplanted can be pretreated with PEDF or the gene transfected into the cells before surgery. Alternatively, PEDF can be transfected at high levels into adjacent retinal pigment epithelial (RPE) cells where they can serve as a supranormal source of the protein. Several investigators have now shown that cultured RPE cells survive very well after transplantation into the interphotoreceptor space of test animals. Transfection of human RPE cells *in vitro* with the PEDF gene then use of them in retinal transplantation thus is feasible.

In Neurodegenerative Diseases:

Many neurodegenerative diseases and other insults to the CNS (brain and retina) are typified by death of neurons and overpopulation by glia (gliosis). PEDF can be used effectively in these conditions to prolong the life and functioning of the primary neurons and to stave off the glial advance. PEDF can be effective, for example, in blocking microglial activation in response to CNS injury as well as prolonging/sparing the lives of neurons.

In the retina, it is predictable that PEDF inhibits the Muller glial cells. Since Muller cells are similar to astroglia, PEDF would be similarly effective in blocking gliosis in conditions such as retinal detachment, diabetes, Retinitis Pigmentosa, etc. as well as sparing the lives of the retinal neurons.

In Glial Cancers:

Most of the major forms of cancer that strike the CNS involve glial elements, PEDF is a gliastatic factor that can be used in combination with other forms of therapy. For example, along with surgery, PEDF can effectively inhibit the spread or reoccurrence of the disease.

Genetic Analysis

The present invention relates to the determination of the organization of the human PEDF gene and its promoter and analysis of its evolutionary relatedness and expression in a variety of human fetal and adult tissues.

The present invention provides, among other things, a nucleic acid which encodes PEDF. In particular, a cDNA sequence is provided as set forth in SEQ ID NO:1. This cDNA sequence codes for PEDF, which has the amino acid sequence set forth in SEQ ID NO:2. Further genomic sequences are mapped in figure 1 and provided SEQ ID NO:43. Additional fragments of the genomic PEDF sequence are provided in SEQ ID NO: 9 through SEQ ID NO: 12. The location of intron-exon junctions are identified in table 1 and SEQ ID NO: 25 through SEQ ID NO: 40 and SEQ ID NO:43.

The term "nucleic acid" refers to a polymer of deoxyribonucleic acid (DNA) or ribonucleic acid (RNA), which can be derived from any source, can be single- or double-stranded, and can optionally contain synthetic, non-natural, or altered nucleotide which are capable of being incorporated into DNA or RNA polymers. The nucleic acid of the present invention is preferably a segment of DNA.

The present invention further provides truncated versions of PEDF. The largest of these is referred to as rPEDF, and comprises the amino acid sequence Met-Asn-Arg-Ile fused to $Asp^{44}...Pro^{418}$ of PEDF, the amino terminus of which has been deleted. The rPEDF protein comprises the amino acid sequence of SEQ ID NO:3. The present invention also provides a nucleic acid which encodes a protein comprising the amino acid sequence of rPEDF, i.e., the amino acid sequence of SEQ ID NO:3.

One who is skilled in the art will appreciate that more than one nucleic acid may encode any given

protein in view of the degeneracy of the genetic code and the allowance of exceptions to classical base pairing in the third position of the codon, as given by the so-called "Wobble rules". Accordingly, it is intended that the present invention encompass all nucleic acids that encode the amino acid sequences of SEQ ID NO:2 and SEQ ID NO:3, as well as equivalent proteins. The phrase "equivalent nucleic acids" is intended to encompass all of these nucleic acids.

It also will be appreciated by one skilled in the art that amino acid sequences may be altered without adversely affecting the function of a particular protein. In fact, some alterations in amino acid sequence may result in a protein with improved characteristics. The determination of which amino acids may be altered without adversely affecting the function of a protein is well within the ordinary skill in the art. Moreover, proteins that include more or less amino acids can result in proteins that are functionally equivalent. Accordingly, it is intended that the present invention encompass all amino acid sequences that result in PEDF protein or functional protein fragments thereof.

Some examples of possible equivalent nucleic acids and equivalent proteins include nucleic acids with substitutions, additions, or deletions which direct the synthesis of the rPEDF protein and equivalent protein fragments thereof; nucleic acids with different regulatory sequences that direct the production of rPEDF proteins; variants of rPEDF which possess different amino acids and/or a number of amino acids other than four fused to the amino terminal end of the protein; and PEDF and rPEDF and functional protein fragments thereof with amino acid substitutions, additions, deletions, modifications, and/or posttranslational modifications, such as glycosylations, that do not adversely affect activity. Since the neurotrophic activity has been correlated to a particular

portion of the PEDF protein fragments containing these residues are clearly within the scope of the present invention.

The present invention also provides a vector which comprises a nucleic acid of SEQ ID NO:1, a nucleic acid which encodes a protein comprising the amino acid sequence of SEQ ID NO:2 or an equivalent protein, a nucleic acid which encodes a protein comprising the amino acid sequence of SEQ ID NO:3 or conservatively modified variant proteins, and conservatively modified variant nucleic acids thereof.

In particular, the present invention provides the vector πFS17, which comprises the nucleic acid of SEQ ID NO:1, and the vector pEV-BH, which comprises a nucleic acid which encodes a protein comprising the amino acid sequence of SEQ ID NO:3. It will be appreciated by those skilled in the art that the cDNA inserts described can be present in alternative vectors. For example, inserts can be in vectors of different nature, such as phages, viral capsids, plasmids, cosmids, phagemids, YACs, or even attached to the outside of a phage or viral capsid. The vectors can differ in host range, stability, replication, and maintenance. Moreover, the vectors can differ in the types of control exerted over cloned inserts. For example, vectors can place cloned inserts under the control of a different promoter, enhancer, or ribosome binding site, or even organize it as part of a transposon or mobile genetic element.

The present invention also provides a host cell into which a vector, which comprises a nucleic acid of SEQ ID NO:1, a nucleic acid which encodes a protein comprising the amino acid sequence of SEQ ID NO:2 or an equivalent protein, a nucleic acid which encodes a protein comprising the amino acid of SEQ ID NO:3 or an equivalent protein, or an equivalent nucleic acid thereof, has been introduced. In particular, the host cell may have the vector πFS17,

which comprises the nucleic acid of SEQ ID NO:1, or the vector pEV-BH, which comprises a nucleic acid which encodes a protein comprising the amino acid sequence of SEQ ID NO:3.

The vectors of the present invention can be introduced into any suitable host cell, whether eukaryotic or prokaryotic. These host cells may differ in their preferred conditions for growth, their nutritive requirements, and their sensitivity to environmental agents. Any appropriate means of introducing the vectors into the host cells may be employed. In the case of prokaryotic cells, vector introduction may be accomplished, for example, by electroporation, transformation, transduction, conjugation, or mobilization. For eukaryotic cells, vectors may be introduced through the use of, for example, electroporation, transfection, infection, DNA coated microprojectiles, or protoplast fusion.

The form of the introduced nucleic acid may vary with the method used to introduce the vector into a host cell. For example, the nucleic acid may be closed circular, nicked, or linearized, depending upon whether the vector is to be maintained as an autonomously replicating element, integrated as provirus or prophage, transiently transfected, transiently infected as with a replication-disabled virus or phage, or stably introduced through single or double crossover recombination events.

The present invention also provides a method of producing PEDF, rPEDF, and equivalent proteins, which method comprises expressing the protein in a host cell. For example, a host cell into which has been introduced a vector which comprises a nucleic acid of SEQ ID NO:1, a nucleic acid which encodes a protein comprising the amino acid sequence of SEQ ID NO:2 or an equivalent protein, a nucleic acid which encodes a protein comprising the amino acid of SEQ ID NO:3 or an equivalent protein, or an

equivalent nucleic acid thereof, may be cultured under suitable conditions to produce the desired protein. In particular, a host cell into which has been introduced the vector πFS17, which comprises the nucleic acid of SEQ ID NO:1, or the vector pEV-BH, which comprises a nucleic acid which encodes a protein comprising the amino acid sequence of SEQ ID NO:3, may be cultured under suitable conditions to produce the proteins comprising the amino acid sequences of SEQ ID NO:2 and SEQ ID NO:3, respectively.

The present invention also provides recombinantly produced PEDF, and functional protein fragments thereof which have been produced in accordance with the aforementioned present inventive method of culturing an appropriate host cell to produce the desired protein. The production of a protein such as PEDF by recombinant means enables the obtention of large quantities of the protein in a highly purified state, free from any disease-causing agents which may accompany the protein isolated or purified from a naturally occurring source organism, and obviates the need to use, for example, fetal tissue as a source for such a protein.

Recombinant PEDF and functional protein fragments thereof may be supplied as active agents to cells by a variety of means, including, for example, the introduction of nucleic acids, such as DNA or RNA, which encode the protein and may be accordingly transcribed and/or translated within the host cell, the addition of exogenous protein, and other suitable means of administration as are known to those skilled in the art. In whatever form in which supplied, the active agent can be used either alone or in combination with other active agents, using pharmaceutical compositions and formulations of the active agent which are appropriate to the method of administration. Pharmaceutically acceptable excipients, i.e., vehicles, adjuvants, carriers or diluents, are well-known to those who are skilled in the art, and are readily

available. The choice of excipient will be determined in part by the particular compound, as well as by the particular method used to administer the compound. Accordingly, there is a wide variety of suitable formulations which can be prepared in the context of the present invention. However, pharmaceutically acceptable excipients not altering the neurotrophic, neuronotrophic and gliastatic activities of the recombinant protein are preferred.

The following examples serve to illustrate further the present invention and are not to be construed as limiting its scope in any way.

EXAMPLE 1

This example describes the trypsin digestion of PEDF and the amino acid sequencing of the resulting fragments.

PEDF was purified from the medium of a primary culture of human fetal RPE cells by high performance liquid chromatography (HPLC). The HPLC-purified PEDF was then reduced and alkylated. Afterwards, it was dried and redissolved in 50 μl of CRA buffer (8 M urea, 0.4 M ammonium carbonate, pH 8.0), and 5 μl of 45 mM dithiothreitol (DTT) (Calbiochem, San Diego, CA) were added. After heating at 50°C for 15 minutes, the solution was cooled, and 5 μl of 100 mM iodoacetic acid (Sigma Chem. Co., St. Louis, MO) were added. After 15 minutes, the solution was diluted to a concentration of 2 M urea and subjected to trypsin digestion (Boehringer-Mannheim, Indianapolis, IN) for 22 hours at 37°C using an enzyme:substrate ratio of 1:25 (wt/wt). Tryptic peptides were separated by narrowbore, reverse-phase HPLC on a Hewlett-Packard 1090 HPLC, equipped with a 1040 diode array detector, using a Vydac 2.1 mm X 150 mm C18 column. A gradient of 5% B at 0 minutes, 33% B at 63 minutes, 60% B at 95 minutes, and 80% B at 105 minutes, with a flow rate of 150 μl/minute, was used. In this gradient, buffer

A was 0.06% trifluoroacetic acid/H_2O, and buffer B was 0.055% trifluoroacetic acid/acetonitrile. Chromatographic data at 210 and 277 nm, and UV spectra from 209 to 321 nm, of each peak were obtained. Samples for amino-terminal sequence analysis were applied to a polybrene precycled glass fiber filter and subjected to automated Edman degradation (Harvard Microchemical Facility, Boston, MA) on an ABI model 477A gas-phase protein sequencer (program NORMAL 1). The resulting phenylthiohydantoin amino acid fractions were manually identified using an on-line ABI Model 120A HPLC and Shimadzu CR4A integrator.

Trypsin digestion of purified PEDF and amino acid analysis of the resulting fragments yielded nonoverlapping peptide sequences, including the sequences JT-3 (SEQ ID NO:6):

```
Thr Ser Leu Glu Asp Phe Tyr Leu Asp Glu Glu Arg
  1               5                  10
Thr Val Arg Val Pro Met Met
         15
```

and JT-8 (SEQ ID NO:7):

```
Ala Leu Tyr Tyr Asp Leu Ile Ser Ser Pro Asp Ile
  1               5                  10
His Gly Thr Tyr Lys Glu Leu Leu Asp Thr Val Thr
         15                  20
Ala Pro Gln Xaa Asn
 25
```

EXAMPLE 2

This example describes the construction of oligonucleotides, based on the peptide sequences of Example 1, the use of the oligonucleotides in the isolation of PEDF cDNA, and the sequencing of PEDF cDNA.

Based on the JT-3 and JT-8 peptide sequences of Example 1 and codon usage data, the oligonucleotides oFS5665 (SEQ ID NO:4): 5'-AGYAAYTTYTAYGAYCTSTA-3' and oFS5667 (SEQ ID NO:5): 5'-CTYTCYTCRTCSAGRTARAA-3' were

constructed on an ABI 392 DNA/RNA Synthesizer and used as primers in a polymerase chain reaction (PCR).

A human fetal eye Charon BS cDNA library (obtained from Dr. A. Swaroop of the Kellog Eye Institute) was amplified once (Sambrook et al., Molecular Cloning: A Laboratory Manual, 2nd ed., Cold Spring Harbor Press, Cold Spring Harbor, NY (1989)) and screened by PCR (Friedman et al., Screening of λgt11 Libraries, In: PCR Protocols: A Guide to Methods and Applications, Innis et al., eds., Academic Press, NY (1990), pp. 253-260) using a Techne thermal cycler and standard reagents (GeneAMP, Perkin-Elmer Cetus), except that $MgSO_4$ was used at 3 mM. A PCR amplification fragment of about 350 bp was isolated on a 3% NuSieve 3:1 gel (FMC Biochemicals, Rockland, ME) using NA-45 DEAE-cellulose paper (Schleicher and Scheull) (Sambrook et al., supra). The fragment was labeled with α^{32}P-dCTP (Amersham Corp., Arlington Heights, IL) by random priming (Random Priming kit, Boehringer-Mannheim, Indianapolis, IN), and used to screen 200,000 plaque-forming units (PFUs) of the human fetal eye library.

Eight positive clones were isolated (Sambrook et al., supra), and DNA of the positive clones was purified according to Qiagen Maxi preparation protocols (Qiagen, Inc., Chatsworth, CA). The inserts of the positive clones were cut out with Not I (BRL, Gaithersburg, MD), circularized with T4 DNA ligase (New England Biolabs, Beverly, MA), transformed into Escherichia coli Epicurian Sure competent cells (Stratagene, Inc., La Jolla, CA), and plated onto Luria broth (LB) plates containing ampicillin and 5-bromo-4-chloro-3-indolyl-β-D-galactoside (X-gal).

White colonies were selected on the basis that such colonies should possess an insert, and plasmid DNA from single colony cultures were isolated by the Qiagen plasmid miniprep protocol. Purified plasmids were digested with EcoR I and Hind III (BRL). These restriction sites were added during library construction

through the ligation of linkers to the 5' and 3' ends of the insert, thus EcoR I- Hind III digestion excises the insert present in isolated plasmids. These fragments were electrophoresed on a 0.7% agarose gel to determine insert size. The plasmid possessing the largest insert, namely πFS17, was selected for mapping and subsequent sequencing using the Sequenase 2.0 sequencing kit (United States Biochemical Corp., Cleveland, OH) to confirm the identity of the clone. Sequence analysis was performed using the MacVector software package (International Biotechnologies, Inc.) and the GenBank® Sequence Data Bank (Intelligenetics, Mountain View, CA).

Sequence analysis of πFS17 revealed a base sequence comprising SEQ ID NO:1, with a long, open reading frame (ORF) encoding the 418 amino acids of SEQ ID NO:2, a typical ATG start codon, and a polyadenylation signal (not shown in SEQ ID NO:1). The coding sequence of the clone aligns exactly with all previously determined PEDF peptide sequences. The deduced amino acid sequence also contains a stretch of hydrophobic amino acids that could serve as a signal peptide. A comparison of the coding sequence and peptide sequence with the GenBank® Data Bank indicates that PEDF is a unique protein having significant homology to the serpin (serine protease inhibitor) gene family, which includes human [α]-1-antitrypsin. Although some of the members of this gene family exhibit neurotrophic activity (Monard et al. (1983) *Prog. Brain Res.*, 58, 359-364; Monard (1988) *TINS*, 11, 541-544), PEDF lacks homology to the proposed consensus sequence for the serpin reactive domain.

EXAMPLE 3

This example describes the construction of an expression vector for the production of recombinant PEDF.

An expression vector was constructed using the plasmid πFS17, which contains the full-length cDNA for human PEDF as described in Example 2. The PEDF coding

sequence was placed under the control of a bacteriophage lambda P_L promoter present in the plasmid pEV-vrf2 (Crowl et al., Gene, 38, 31-38 (1985)) to obtain the vector pEV-BH. This was accomplished by obtaining a BamH I-Hind III fragment of πFS17 comprising a portion of the PEDF coding region (namely, nucleotide 245 to 1490 of SEQ ID NO:1), digesting plasmid pEV-vrf2 with EcoR I-Hind III, rendering both fragments blunt by means of a fill-in reaction at the BamH I and EcoR I ends with DNA polymerase I (Klenow fragment), and ligating the resultant blunt-ended/compatible-ended fragments to each other. The resultant vector pEV-BH places a distance of 8 nucleotide between the Shine-Dalgarno (SD) sequence and the PEDF coding region. The construct specifies Met-Asn-Arg-Lle-Asp^{44}---Pro^{418} such that a protein of 379 amino acids, known as rPEDF, is encoded as indicated in SEQ ID NO:3. The amino acids at the amino terminus of the rPEDF protein do not occur in native PEDF and result from the fusion of nucleic acids during the construction of pEV-BH.

To verify production of the recombinant PEDF protein by pEV-BH, the plasmid was propagated in E. coli strain RRI (Maniatis et al. (1982) Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY), bearing the low copy-number compatible plasmid pRK248cIts that contains a gene for encoding a temperature-sensitive λcIAt2 repressor (Bernard et al. (1979) *Methods in Enzymology*, 68, 482-492). Protein induction was performed as described in Becerra et al. (1991) *Biochem.*, 30, 11707-11719, with the following modifications. Bacterial cells containing pEV-BH were grown in LB medium containing 50 μg/ml ampicillin at 32°C to early logarithmic phase, such that OD_{600nm}=0.2. The temperature of the culture was rapidly increased to 42°C by incubating the flask in a 65°C water bath, and the bacteria were subsequently grown at 42°C for 2-3 hours in

an air-flow incubator at 340 rpm. Aliquots were taken for absorbance readings at 600 nm.

Nascent proteins, synthesized following protein induction, were radiolabeled. After the temperature of the culture had reached 42°C, 150 μCi of L-[^{35}S]methionine (1040 Ci/mmol, Amersham Corp., Arlington Heights, IL) were added per ml of culture, and incubation was continued at 42°C for 10 minutes and 30 minutes. Cells were harvested by centrifugation and washed with TEN buffer (10 mM Tris-HCl, pH 7.5, 1 mM EDTA, and 100 mM NaCl). ^{35}S-labeled peptides from total bacterial extracts were resolved and analyzed on SDS-12% PAGE followed by fluorography. A band corresponding to a 42,820 M_r polypeptide was detected 10 and 30 minutes post-induction. The size obtained for the recombinant protein expressed by pEV-BH matched the expected size for the coding sequence subcloned in pEV-BH. In a similar manner, smaller fragments (BP = 28,000 M_r; BX = 24,000 M_r; BA = 9,000 M_r) can be synthesized and purified. BP peptide includes PEDF amino acids 44 through 269, BX peptide includes PEF amino acids 44 through 227, and BA peptide includes PEDF amino acids 44 through 121.

EXAMPLE 4

This example describes the construction of expression vectors containing the full-length PEDF cDNA.

In a manner similar to that described in Example 3 for the construction of pEV-BH, the PEDF ORF of plasmid πFS17 was placed under the control of the bacteriophage lambda P_L promoter present in the plasmids pRC23 and pEV-vrf1 (Crowl et al. Gene, 38, 31-38 (1985)). This was accomplished by obtaining the SfaN I-Hind III fragment of πFS17 comprising a portion of the PEDF cDNA (namely, nucleotide 107 to 1490 of SEQ ID NO:1), digesting the plasmids with EcoR I-Hind III, rendering the fragments blunt by means of a fill-in reaction at the SfaN I and EcoR I ends with DNA polymerase I (Klenow fragment), and ligating the resultant blunt-ended/compatible-ended

fragments to each other. The resulting vectors pRC-SH and pEV-SH place a distance of 14 and 8 nucleotide, respectively, between the SD sequence and the PEDF coding region. The construct pRC-SH encompasses the full-length PEDF ORF, and specifies a PEDF protein of 418 amino acids, with its naturally occurring amino terminus, as set forth in SEQ ID NO: 2. The construct pEV-SH encompasses the full-length PEDF ORF, and specifies a PEDF amino-terminal fusion protein of 425 amino acids, with Met-Asn-Glu-Leu-Gly-Pro-Arg (SEQ ID NO:8) preceding the PEDF sequence of SEQ ID NO:2. These additional amino acids at the amino terminus do not occur in native PEDF, and the codons in pEV-SH specifying these additional amino acids result from the fusion of nucleic acids during the construction of pEV-SH.

To verify production of the recombinant proteins specified by the two vectors, the vectors were introduced into E. coli strain RRI [pRK248cIts], and protein induction was performed and monitored by metabolic labeling with ^{35}S-methionine during induction in a manner similar to that set forth in Example 3. The induced expression of the proteins specified by pRC-SH and pEV-SH had a negative effect on bacterial cell growth. In comparison with bacterial cultures containing the parental plasmids, cultures containing pRC-SH and pEV-SH grew and divided more slowly. This negative effect on bacterial growth correlated with the distance between the initiation codon and the SD, which may suggest that a shorter such distance results in more efficient translation of the recombinant protein. A 46,000 M_r candidate polypeptide for PEDF was not detected in the media or cell lysates of bacterial cultures containing pRC-SH and pEV-SH. However, a 35,000 M_r protein was observed in extracts of cultures containing pRC-SH and pEV-SH, but not in extracts of cultures containing parental plasmids. This may indicate that the amino-terminal end of PEDF is protease-sensitive

and that recombinant full-length PEDF is metabolized in this particular host. Alternatively, failure to observe the anticipated-sized recombinant PEDF proteins may reflect an experimental artifact which could be overcome through the use of alternative expression vectors, hosts, inducible promoters, subcloning sites, methods of recombinant protein isolation or detection, or means of protein induction.

EXAMPLE 5

This example describes a method for producing large quantities of recombinantly produced PEDF.

A total of 1 g of E. coli cells containing rPEDF was resuspended in 50 ml 20mM Tris-HCl, pH 7.5, 20% sucrose, and 1 mM EDTA. The cells were maintained on ice for 10 minutes, sedimented by centrifugation at 4000 x g, and were resuspended in 50 ml of ice-cold water for 10 minutes. Lysed outer cell walls were separated from spheroplasts by centrifugation at 8000 x g.

The pelleted spheroplasts were resuspended in 10 ml of phosphate buffered saline (PBS) containing 5 mM EDTA, 1 μg/ml pepstatin and 20 μg/ml aprotinin. The suspension was probe-sonicated with a sonicator (Ultrasonics, Inc., model W-225) to lyse the cell membranes. Three bursts at 30 second pulses with a 30 second pause were performed while the sample was immersed in an ice-water bath. RNase TI (1300 units, BRL) and DNase I (500 μg, BRL) were added to the sonicated cell suspension, and the suspension was incubated at room temperature for 10 minutes. This suspension was diluted by the addition of 40 ml of phosphate buffered saline (PBS) containing 5 mM EDTA, 1 μg/ml pepstatin and 20 μg/ml aprotinin, and the crude inclusion bodies were sedimented by centrifugation at 13,000 x g for 30 minutes. The particulate material consisting of inclusion bodies was resuspended in 40 ml of PBS containing 25% sucrose, 5 mM EDTA, and 1% Triton X-100, incubated on ice for 10

minutes, and centrifuged at 24,000 x g for 10 minutes. The washing step was repeated three times. Finally, the inclusion bodies were resuspended in 10 ml of denaturation buffer containing 50 mM Tris-Cl, pH 8.0, 5 M guanidine-Cl, and 5 mM EDTA. The suspension was probe-sonicated briefly for 5 seconds in an ice-water bath. The resulting suspension was incubated on ice for an additional hour. After centrifugation at 12,000 x g for 30 minutes, the supernatant was added to 100 ml of renaturation buffer containing 50 mM Tris-Cl, pH 8.0, 20% glycerol, 1 mM DTT, 1 μg/ml pepstatin, and 20 μg/ml aprotinin, and stirred gently at 4°C overnight to renature the protein. The soluble and insoluble fractions were separated by centrifugation at 13,500 x g for 30 minutes.

The soluble fraction was further purified by concentrating it to 1 ml using a Centricon 30 microconcentrator (Amicon Div., W.R. Grace & Co., Beverly, MA), and dialyzing it against Buffer A (50 mM sodium phosphate, 1 mM DTT, 20% glycerol, 1 mM EDTA, 1 μg/ml pepstatin, and 1 mM benzamidine) at 4°C for 3 hours. The dialyzed extract was centrifuged at 14,000 rpm in an Eppendorf Centrifuge (Model 5415C) for ten minutes. The supernatant fraction was layered on a S-Sepharose fast-flow (Pharmacia, New Market, NJ) column (1 ml bed volume) pre-equilibrated with buffer A. The column was washed with two column-volumes of buffer A. Finally, recombinant rPEDF was eluted with a step gradient of 50, 100, 150, 200, 300, 400, 500, and 1000 mM NaCl in buffer A. Fractions of 1 ml were collected by gravity flow, and were dialyzed against buffer A. Fraction 300, containing recombinant rPEDF, was stored at -20°C. The recovery in fraction 300 was 50 μg per gram of packed cells, which represents 25% of the total protein.

Most of the rPEDF was recovered from the insoluble fraction by dissolving the fraction in 10 ml of 6M guanidinium-Cl in buffer B (50 mM Tris-Cl, pH 8.0, 1 mM

DTT, 2 mM EDTA). The solution was centrifuged at 10,000 x g for 5 minutes. The supernatant was layered onto a Superose-12 (Pharmacia, New Market, NJ) column attached in tandem to a second Superose-12 column (each column 2.6 cm x 95 cm) pre-equilibrated with buffer containing 4 M guanidinium-Cl in buffer B. The flow rate was 3 ml/minute. Recombinant rPEDF containing fractions from the Superose-12 column were pooled and dialyzed against buffer C (4 M urea, 50 mM sodium phosphate, pH 6.5, 1 mM benzamidine, 1 μg/ml pepstatin, 4 mM EDTA). The dialyzed fraction was passed through a 0.22 μm filter (Miller-GV, Millipore Corp., Bedford, MA). The filtered solution was layered onto a mono-S (Pharmacia, New Market, NJ) column (1 cm x 10 cm, d x h) pre-equilibrated with buffer C. The column was washed with buffer C, and recombinant rPEDF was eluted with a gradient of 0 mM - 500 mM NaCl in buffer C at 0.5 ml/min. Two-ml fractions were collected, and the peak fractions of recombinant rPEDF were pooled. The recovery in the pooled fractions was 0.5 mg of recombinant PEDF per gram of packed cells.

EXAMPLE 6

This example describes the use of purified recombinant PEDF as a differentiation agent.

Y79 cells (ATCC, HTB18) were grown in Eagle's Minimal Essential Medium with Earl's salts (MEM) supplemented with 15% fetal bovine serum and antibiotics (10,000 u/ml penicillin and 10 mg/ml streptomycin) at 37°C in a humidified incubator under 5% CO_2. Cells were propagated for two passages after receipt from the ATCC, and then frozen in the same medium containing 10% DMSO. A few of the frozen aliquots were used for each differentiation experiment. All experiments were performed in duplicate.

After thawing, the cells were kept, without further passaging, in the serum-containing medium until the appropriate number of cells were available. Cells

were collected by centrifugation and washed twofold in PBS, resuspended in PBS, and counted. At that point, 2.5 x 10^5 cells were plated into each well of a 6-well plate (Nunc, Inc., Roskilde, Denmark) with 2 ml of serum-free medium (MEM, supplemented with 1 mM sodium pyruvate, 10 mM HEPES, 1X non-essential amino acids, 1 mM L-glutamine, 0.1% ITS mix (5 μg/ml insulin, 5 μg/ml transferrin, 5 ng/ml selenium, Collaborative Research, Bedford, MA), and antibiotics as described above.

Differentiation effectors and control buffers were added 12-16 hours after plating, and the cultures were incubated and left undisturbed for 7 days. On the eighth day, cells were transferred to poly-D-lysine-coated six-well plates (Collaborative Research, Bedford, MA), and the old medium was replaced with 2 ml of fresh serum-free medium, upon attachment of the cells to the substrate. The cultures were maintained under these conditions for up to 11 days. Post-attachment cultures were examined daily for morphological evidence of differentiation as well as quantification of neurite outgrowth using an Olympus CK2 phase-contrast microscope.

In comparison with untreated cells, only Y79 cultures that were exposed to recombinant rPEDF showed any significant evidence of neuronal differentiation. Some neurite outgrowth (below 5%) was detectable in control cultures treated with the same buffer used to solubilize rPEDF, and no evidence of differentiation was found in cultures processed in the same manner without the addition of rPEDF or buffer (Figure 22A, "control"). Phase contrast microscopy of rPEDF treated cultures showed that between 50-65% of the cell aggregates had neurite extensions by day 3 post-attachment on poly-D-lysine (Figure 22B, "PEDF"). These 3-day neurite extensions appeared as short projections from pear-shaped cells at the edges of the cell aggregates. The number of differentiating aggregates, the number of differentiating

cells per aggregate, and the length of the neurite-like processes increased with post-attachment time. By day 5 post-attachment, about 75-85% of the aggregates showed signs of differentiation with neurites extending from most of their peripheral cells. rPEDF-treated cultures reached the maximum extent of differentiation on day 7 post-attachment, when 85-95% of the cells aggregate. At that time, two types of neuronal processes were observed, i.e., single neurites 2-3 fold longer than those observed on day 3 extending from peripheral cells of isolated aggregates, and much longer and thinner processes forming a branching network between neighbor cell aggregates. Upon extended incubation, i.e., beyond 10 days post-attachment, there was a marked decrease in the proportion of the network connections, and no further growth of the single neurites, although the viability of the cell aggregates was not severely affected, and remained at about 75-80% in different experiments. No differences were observed between purified native PEDF and recombinant PEDF (rPEDF) as seen in Figure 23.

The PEDF and rPEDF cDNA clones not only provide means to produce large quantities of the PEDF and rPEDF proteins but also serve as sources for probes that can be used to study the expression and regulation of the PEDF gene. In addition, these sequences can be used in the antisense technique of translation arrest to inhibit the translation of endogenous PEDF.

The recombinantly produced PEDF and rPEDF proteins and equivalent proteins can be used as potent neurotrophic agents in vitro and in vivo. Additional biochemical activities of these proteins as neurotrophic agents can be determined through standard in vitro tests, which will enable the development of other therapeutic uses for these proteins in the treatment of inflammatory, vascular, degenerative and dystrophic diseases of the retina. Given that these proteins are such potent

neurotrophic agents, it can be envisioned that these proteins could be modified for therapeutic utility in the treatment of tissues other than the retina, which also respond to neurotrophic factors. These proteins may even find more generic utility as "differentiation" factors for non-neural tissues and certain types of cancer.

EXAMPLE 7

Along with the 3,000 mol. wt. recombinant PEDF, smaller recombinant constructs have been synthesized to determine if they have neurotrophic activity. Smaller peptides could offer a variety of advantages over the full-length construct such as greater solubility, better membrane penetration, less antigenicity, greater ease in preparation, etc.

Figure 23 shows only three of the constructs that have been tested. BP, BX and BA are about 28,000, 24,000 and 9,000 mol. wts. respectively and represent C-terminal deletion mutants. All of these show neurotrophic activity similar to that depicted in Figures 21 and 22. The novel finding here is that even the 9,000 m.w. peptide (only about 20% of the full m.w. of the native protein) exhibits striking neurotrophic activity. Moreover, the active neurotrophic peptide represents sequences at the N-terminal rather than at the C-terminal which is known to contain the serpin active site. Thus, that the active site is at the N-terminal and activity can be elicited with such a small molecule are surprising findings that could not have been predicted based on any previous findings.

TABLE 1

Exon and Intron Organization of the human PEDF Gene

Exon Number	Exon Size (bp.)	5' Splice Donor	SEQ. ID. NO.	Intron size (Kb)
		Promotor ...aaggagta		
1	128	TATCCACAG/gtaaagtag...	25	4806bp
2	92	CCGGAGGAG/gtcagtagg...	26	2862bp
3	199	TCTCGCTGG/gtgagtgct...	27	980 bp
4	156	TTGAGAAGA/gtgagtcgc...	28	688 bp
5	204	ACTTCAAGG/gtgagcgcg...	29	2982bp
6	143	AGCTGCAAG/gtctgtggg...	30	1342bp
7	211	AGGAGATGA/gtatgtctg...	31	444 bp
8	377	TTTATCCCT/aacttctgt...	32	

3' Splice Acceptor	SEQ. ID. NO.	Intron No.
GCTGTAATC	33	1
...ttcttgcag/GCCCCAGGA	34	2
...tcctgccag/GGCTCCCCA	35	3
...ctctggcag/GAGCGGACG	36	4
...tcttctcag/AGCTGCGCA	37	5
...tctttccag/GGCAGTGGG	38	6
...ttgtctcag/ATTGCCCAG	39	7
...tctctacag/AGCTGCAAT	40	8

Table 1: Exons are in upper case and introns sequences in lower case. The 5' donor GT and 3' acceptor AG are underlined. Exon and intron sizes are given in bp and kb respectively.

EXAMPLE 8

Cloning and sequencing of the human PEDF gene.

Materials- Restriction enzymes, SuperScript® RT and Kanamycin were purchased from GIBCO-BRL (Gaithersburg, MD). Dynabeads® Oligo $dT_{(25)}$ were purchased from Dynal Inc. (Lake Success, NY). Retrotherm™ RT was obtained from Epicentre Technologies (Madison, WI). RNAsin® was purchased from Promega (Madison, WI). Taq polymerase was purchased from Perkin-Elmer (Norwalk, CT), or Stratagene (La Jolla, CA). The plasmid vector pBlueScript® used for subcloning was purchased from Stratagene (La Jolla, CA). Total RNA from neural retina and retinal pigment epithelium was purified from human tissue obtained from the National Disease Research Interchange (NDRI, Philadelphia, PA) as previously described (Chomczynki and Sacchi, 1987). [^{32}P]α -dATP and [^{32}P]γ-ATP (3000 Ci/mmol) used for labeling and sequencing (respectively) were purchased from Amersham) Arlington Hts, IL). Superbroth (Bacto-Tryptone 12g/L, yeast extract 24 g/L, K_2 HPO_4 12.5 g/L, $HK_2PO_4$3.8 g/L and glycerol 5 mL/L), denaturing solution (0.2 N NaOH, 1.5 M NaCl), neutralizing solution (1 M Tris-Cl pH 7.0, 1.5 M NaCl), 20X SSC (3.0 M NaCl, 0.3 mM sodium citrate), 10X TBE (1 M Tris-borate, 2 mM EDTA, pH 8.3), and 50X TAE (2 M Tris-acetate 50 mM EDTA, pH 8.0) were purchased from Quality Biologicals (Gaithersburg, MD). 20X SSPE (3M NaCl, 0.2 M NaH_2PO_4, 20 mM EDTA pH 7.4) was purchased from Digene Diagnostics, Inc. (Silver Spring, MD). Ampicillin was purchased from Sigma Chemical Co. (St. Louis, MO) dissolved in water and filter-sterilized.

Polymerase chain reaction (PCR). A 2X PCR mix was prepared containing 1.6 μmoles/mL of GeneAmp® dNTPs (400 μM each), 2X GeneAmp® PCR buffer and 50 U/mL Taq polymerase. These reagents were purchased from Perkin-Elmer (Norwalk, CT). In general, the template and

oligonucleotides (100 ng of each oligo) were mixed in 25 μL volume and 25 μL of the 2X mix were then added followed by 50 μL of mineral oil. The template was initially denatured for 2 min at 95°C, 30 sec annealing (temperature between 55 and 65°C depending on the primers) and an extension at 72°C for 1-5 min depending on the length of the product amplified.

cDNA synthesis on Dynabeads® oligo $(dT)_{25}$. The cDNA was synthesized on Dynabeads as previously described (Rodriguez and Chader 1992). The Dynabeads (0.5 mg) were washed with 100 μL of 10 mM Tris-Cl pH 7.0, 1 mM EDTA, 1 M KCl. The total RNA 30μL, (30μg,~1μL), in water was mixed with 30 μL of the above buffer and the equilibrated Dynabeads (0.5 mg) then heated to 55°C for 2 minutes. The poly+ A RNA was allowed to anneal to the beads for 15 min at room temperature and the excess RNA removed by binding the beads for 15 min at room temperature and the excess RNA removed by binding the beads to the MPC-E magnetic separator (Dynal Inc.). The beads with the annealed poly+ A mRNA were then suspended in 2.5 μL buffer A (200 mM Tris-Cl pH 8.3, 1.0 M KCl), 2.5 μL buffer B (30 mM $MgCl_2$, 15 mM MnCl), 20 μL 10 mM dNTP's (2.5 mM each), 1 μL RNAsin, 2 μL SuperScript RT, 5 μL of Retrotherm RT (1 Unit/μI) and 16 μL of H_2O to make a final volume of 50 μL. The reaction mixture was incubated at 40°C for 10 min, than at 65°C for 1 hr. The beads were again bound to the MPC-E magnetic separator and the excess RT reaction mix removed. The beads were then washed once with 100 μL 0.2N NaOH, once with 10X SSPE, and twice in 1X TE. The cDNA-containing beads were suspended in a final volume of 100 μL 1X TE.

5' Rapid Amplification of cDNA Ends (RACE). The 5'-RACE was performed using a modified method based on the 5'-AmpliFINDER RACE kit purchased from Clontech (Rodriguez et al. 1994). First, cDNA was synthesized on Dynabeads® Oligo $dT_{(25)}$ as described above (Rodriguez and Chader,

1992). The AmpliFINDER anchor primer (Clontech) was ligated to the 3' ends tips of the Dynabead-immobilized retinal pigment epithelium cDNA using the same conditions as for soluble cDNA described in the 5'-AmpliFINDER RACE kit. The Ampli-FINDER anchor primer was used in combination with an PEDF-specific primer #2744 to PCR amplify the 5' prime end. The amplification was done as described above with 2 μL of anchor-ligated human retinal pigment epithelium-Dynabeads cDNA used as template. The amplification was performed for 30 cycles.

Sequence of oligonucleotides. Oligonucleotide primers were synthesized in an Applied Biosystems Inc. (Foster City, CA) DNA synthesizer model 392. The oligonucleotides were deprotected and used without further purification.

Screening of genomic libraries. The human genomic cosmid library (Clontech) was plated on LB plates containing 150 mg/mL ampicillin, 20 mg/mL Kanamycin at a density of 10,000 colonies per plate. Nitrocellulose filters were used to lift the colonies and the filters were treated and hybridized as described in Sambrook et al., (1989). The library was probed with [^{32}P]-labeled PCR product obtained from amplifying a PEDF cDNA clone (Steele et al. 1993) using T7/T3 primers. This resulted in the isolation of the p10A cosmid. A λDASH™II library (Stratagene) was screened by Lark Sequencing Technologies Inc. (Houston, TX) using the insert from the PEDF cDNA clone mentioned above. This resulted in the isolation of the 7 Kb NotI-Not fragment (JT6A). A P-1 clone, p147, containing the entire PEDF gene and flanking regions was isolated using oligos 1590/1591 by Genome Systems (St. Louis, MO).

Cloning of PCR products: Four sets of primers, 603:604; 605:606; 2238:354 and 2213:2744 designed from the internal coding regions of the PEDF cDNA sequenced were synthesized

as decribed above for use as primers in a polymerase chain reaction (PCR) experiments. The primer sequences are as follows: 603: 5'-ACA AGC TGG CAG CGG CTG TC-3' (SEQ ID NO: 13), 604: 5'-CAG AGG TGC CAC AAA GCT GG-3' (SEQ ID NO: 14); 605: 5'-CCA GCT TTG TGG CAC CTC TG-3' (SEQ ID NO: 15), 606: 5'-CAT CAT GGG GAC CCT CAC GG-3' (SEQ ID NO: 16), 2213: 5'-AGG ATG CAG GCC CTG GTG CT-3' (SEQ ID NO: 17), 2744: 5'CCT CCT CCA CCA GCG CCC CT-3' (SEQ ID NO: 18); 2238: 5'-ATG ATG TCG GAC CCT AAG GCT GTT-3' (SEQ ID NO: 19), 354: 5'-TGG GGA CAG TGA GGA CCG CC-3' (SEQ ID NO: 20). The amplifications, subcloning and sequencing of the PCR products generated with primers 603:604 and 605:606 was performed by Lark Sequencing Technologies Inc. using human genomic DNA as template. The product generated from 603:604 is ~2 kb (jt8A) and expands from exon 3 to exon 5. The product generated using 605:606 is ~3.3 kb (jt 9) and expands from exon 5 to exon 6. The primers set 2213-2744 was used to amplify a ~ 2.5 Kb product (jt15; also referred to as JT115) from the P1 clone p147. This product was then sent to Lark Sequencing Technologies Inc. for subcloning and sequencing. The 2238:354 primers were used to amplify from exon 6 to exon 7 across intron E. This product was not subcloned but was sequenced directly and entirety by us.

DNA sequencing. The P-1 clone (p147), subclones of this clone and PCR products from this clone were sequenced. Most of the sequencing was performed by Lark Sequencing Technologies Inc. using standard sequencing techniques. All important areas (e.g. intron-exon boundaries), and junctions between clones were sequenced in our laboratory. DNA from the PCR products was prepared for sequencing using Wizard™ PCR Preps DNA purification kit purchased from Promega Corp. (Madison, WI). The P-1 clone, and plasmid subclones were purified using Qiagen Inc. (Chatsworth, CA) Midi plasmid purification kit. The purified PCR products and plasmids were sequenced using

the PRISM™ DyeDeoxy Terminator Cycle Sequencing Kit (Applied Biosystems a Division of Perkin-Elmer Corp., Foster City, CA), following the manufacturer's protocol. Typically, 0.5 pmoles of template and 3 pmoles of primer were used per sequencing reaction. The sequencing reaction products were purified using Select-D G-50 columns (5 Prime-3 Prime; Boulder, CO) and dried. Each sample was then dissolved in 5μL formamide, 1 μL 50 mM EDTA, heated and located in a Model 370A Automated Fluorescent Sequencer (ABI, Foster City, CA). All splice-sites junctions, intron F and junctions across clones were sequenced.

Southern blot. An EcoRI digested genomic (8 μg) blot of DNA from a variety of species was purchased from BIOS Laboratories, New Haven, CT. The blot was probed with the PEDF cDNA using standard techniques (Sambrook et al., 1989).

5' *RACE of PEDF.* The 5' RACE was performed as described above by ligating the anchor oligo to human retinal pigment epithelium cDNA previously synthesized on Dynabeads. The 5' end was amplified using the anchor primer (AmpliFinder's kit) and the PEDF-specific primer 2744. The amplification was performed for 30 cycles. One main band was observed at ~ 230 bp. The PCR products were cloned in pGEM-T (Promega Corp., Madison, WI) and sequenced. The longest of these clones was found to extend the 5' end of PEDF by 20 bp.

Isolation of the PEDF gene. The PEDF gene was isolated in a P-1 clone (p147) by Genome Systems (St. Louis, MO) using primers 1590 and 1591(1590: 5'-GGA CGC TGG ATT AGA AGG CAG CAA A-3' (SEQ ID NO: 23); and 1591: 5'-CCA CAC CCA GCC TAG TCC C-3' (SEQ ID NO: 24)). In order to determine if this clone contained the entire PEDF gene, both p147 and human genomic DNA were digested with BamHI, EcoHI, HindIII and PstI then separated by agarose

gel electrophoresis in a pulse field apparatus. The agarose gel was blotted and probed with the PEDF cDNA clone (Steele et al. (1993) *Proc. Natl. Acad. Sci. USA* **90**:1526-1530). Comparison of the band pattern between the P-1 clone and genomic DNA indicates that the entire PEDF gene is contained in this clone. Furthermore, this result is also indicative that there is only one gene for PEDF.

Sequence of the PEDF gene. A scale map of the gene is shown in Fig. 1. The PEDF gene was sequence in its entirety (SEQ ID NO:43). The clones jt1, jt14, jt6A and related PCR products (jt15, jt8A and jt9)(Fig. 1) were sequenced by Lark Sequencing Technologies Inc. The rest of the gene was sequenced by amplifying different portions of the gene using the p147 clone as template. All exons, intron-exon junctions and the entire intron F were sequenced in both directions in our laboratory as described above from PCR products generated from the P-1 clone, p147. The Not I site downstream from exon 1 was also confirmed by amplifying across it and sequencing the product. The gene expands approximately 16 Kb with 8 exons. All intron-exon junctions obey the AG/GT rule. The intron-exon junctions and flanking sequences are shown in Table I.

jt1: A 7.1 kb cosmid clone isolated from a human genomic cosmid library (Clontech) containing exon 7, exon 8 and the 3' flanking region of the PEDF gene. The 5' end of this clone, an area of approximately 2.1 Kb, is not part of PEDF. This was apparently caused by a rearrengement of the cosmid. This clone was sequenced entirely by Lark Sequencing Technologies Inc.

jt6A: This is a 7.2 kb Not I fragment isolated by Lark Sequencing Technologies Inc. from a λDASHII human genomic library (Statagene). This clone contained >6 Kb of the 5' flanking region, exon1 and 424 bp of intron A of the PEDF

gene. This clone was sequenced entirely by Lark Sequencing Technologies Inc.

jt8A: This cloned PCR product JT8A generated from genomic DNA using primers 603:604. This clones expands from exon 3 to exon 5 including exon 4 and introns C and D. It was amplified, cloned and sequenced entirely by Lark Sequencing Technologies Inc.

jt9: This cloned PCR product JT8A was generated from genomic DNA using primers 605:606. It contains the entire intron E and portions of exon 5 and exon 6. It was amplified, cloned and sequenced entirely by Lark Sequencing Technologies Inc.

jt15: This clone was obtained from a PCR product amplified using the primer pair 2213:2744 from p147. The clone expands from exon 2 to exon 3 across intron B. The PCR product was submitted to Lark Sequencing Technologies Inc. for subcloning and sequencing.

P1 clone p147: This clone was isolated by Genome Systems Inc. using oligonucleotides 1590:1591. This clone was used to obtain the sequence of intron F (2238:354), and the subclone jt14. It was also used to confirm the intron-exon boundaries initially obtained from the above mentioned clones. All the exons and intron boundaries were amplified (using p147 as template) using intron-specific oligos and the products sequenced. All splice junctions sequences were confirmed as well as the sizes of introns and exons.

jt14: This is a subclone of p147 containing most of intron A, exon 2 and a portion of intron B. This clone was isolated by us and sent to Lark Sequencing Technologies Inc. for sequencing.

Thus from the sequence analysis of all the above mentioned clones and PCR products the structure and size of exons and introns of the human PEDF gene were determined. The 5' splice donor and 3' splice acceptor sites in all junctions conform to the GT/AG consensus.

EXAMPLE 9

Analysis of the PEDF promoter.

In order to obtain some understanding as to the possible transcriptional elements that may regulating PEDF and guidance for future experiments on PEDF expression, we performed a theoretical analysis of the PEDF 5' flanking region (Fig. 3). The 5' flanking region of the PEDF gene lacks the classical TATAAA signal or TATA-box. However, it contains several interesting features and elements recognized by important transcription factors. There are two Alu repetitive elements from -164 to -591, and from -822 to -1050. Outside the Alu regions, there are two possible sites for the ubiquitous octamer family of transcription factors (Oct) at -29 (ATCCAAAT) and again at -113 (GTGCAAAT) which deviate by one base from the consensus ATGCAAAT (Parslow et al. (1984) *Proc. Natl. Acad. Sci. U.S.A.* **81**:2650-2654; Falkner et al. (1984) *Nature* **310**:71-74; Sturm et al. (1988) *Genes & Devel.* **2**:1582-1599; Faisst and Meyer (1992) *Nuc. Acids Res.* **20**:3-26). Another element of possible interest is located at -62. This element, GTAAAGTTAAC, which resembles the HNF-1 (hepatocyte nuclear factor) binding consensus GTAATNATTAAC (Frain, M., et al. (1989) *Cell* **59**:145-147). This is a homedomain-containing transcription factor which transactivates many predominately hepatic genes (Kuo et al. (1990) *Proc. Natl. Acad. Sci. USA* **87**:9838-9842) but has been implicated in endodermic differentiation (Baumhueter et al. (1990) *Genes Dev.* **4**:371-379). The sequence TCAGGTGATGCACCTGC at -202 is very similar to the artificial palindromic sequence (TREp) TCAGGTCATGACCTGA which is recognized by AP-1 and possibly transactivated by retinoic acid (Umescono et al. (1988) *Nature* **336**:262-265; Linney (1992) *Curr. Topics in Dev. Biol.* 27:309-350). The sequences TGAGTGCA at -22 and TGATGCA at -207 (within the TREp), are similar to the AP-1 consensus sequence TGACTCA

(Schüle, et al. (1990) *Cell* **61**:497-504). The sequence AGGTGATGCACCT at -204 contained within the TREp is also similar to the developmentally regulated RAR (retinoic acid receptor) motif whose consensus is AGGTCATGACCT (Faisst and Meyer (1992) *Nuc. Acids Res.* **20**:3-26). The PEA3 element (polyomavirus enhancer activator 3) AGGAAG/A (Martin et al. (1988) *Proc. Natl. Acad. Sci. USA* **85**:5839-5843; Faisst and Meyer (1992) *Nuc. Acids Res.* **20**:3-26) is present in tandem at -122 and -129, then again at -141. PEA3 is a member of the ETS family of transcription factors (Macleod et al. (1992) *TIBS* **17**:251-256) and its activity seems to be regulated by non-nuclear oncogenes (Wasylyk et al. (1989) *EMBO J.* **8**:3371-3378). One of the most interesting elements is located at -654 with the sequence GTGGTTATG. This element is within the consensus sequence GTGGT/AT/AT/AG recognized by the C/EBP (CAAT enhancer binding protein) family of transcription factors (Faisst and Meyer (1992) *Nuc. Acids Res.* **20**:3-26). This factor seems to be involved in terminal differentiation that leads to an adult phenotype (Vellanoweth et al. (1994) *Laboratory Investigation* **70**:784-799). Three possible CACCC boxes are present one at -845 and two in the reverse orientation at -826 and -905. These are all within the Alu repeat. A possible Sp1 site (CCCGGC) is present at -153 before the Alu repeat and a consensus Sp1 site GGCGGG is present -1030 inside the Alu repeat.

EXAMPLE 10

Expression of PEDF mRNA in Cultured Cells

Gene expression analysis

Multiple human tissue mRNA Northern blots (Clonetech) with 2 ug Poly-(A) RNA per lane were hybridize with a radioactively-labelled 667 bp PCR amplified PEDF product (Tombran-Tink et al., 1994 Genomics, 19:266-272). Blots were prehybridized for 15 min at 68°C in QuickHyb rapid hybridization solution (Stratagene, La Jolla, CA)

and hybridized for 1 hr at 68°C in the same solution containing 5 x 10^6 cpm DNA/ml. Hybridized blots were washed twice with 100 ml of 2XSSC, 0.1% SDS for 15 min at room temperature and once with 200 ml of 0.1XSSC, 0.1% SDS for 30 min at 68°C. The blots were autoradiographed at -70°C for 2 hr using Kodax XAR-5 film and DuPont intensifying screens.

Gene Expression:

In order to determine whether expression of the PEDF messenger RNA occurs in human tissues other than in cultured human fetal RPE cells, we analyzed multiple tissue human adult and fetal RNA blots containing equal amounts of poly-(A) RNA for each tissue examined. The results are shown in Figure 4. The PEDF probe identified a single primer 1.5 kb transcript of varying intensity of hybridization in 14 of the 16 adult tissue analyzed. No signal is detected in either adult kidney or peripheral blood leucocytes. Only a weak signal can be observed in adult brain, pancreas, spleen and thymus. The greatest amount of hybridization for PEDF messenger RNA is seen in human adult liver, skeletal muscle, testis and ovary. Surprisingly, only a very weak signal is observed in total brain RNA. In the fetal tissues examined, a very strong PEDF signal is seen in liver tissue, and interestingly a signal of significant intensity in fetal kidney as compared to no PEDF hybridization in adult kidney samples.

In contrast to the single 1.5 kb transcript observed in the adult tissues, an additional minor transcript of less than 500 bp is labelled variably and with lower intensity in fetal heart, lung and kidney. This may be due to partial degradation of the message or an alternative splicing phenomenon. PEDF is also only expressed in early passaged monkey RPE cells (1st - 5th passage) and not in late passaged cells (10th passage).

° These data demonstrate the relevance of PEDF to senescence.

EXAMPLE 11

Comparative Analysis Of PEDF In
A Variety Of Phylogenetically Related Species

Evolutionary conservation analysis

8 ug of genomic DNA from lymphocytes of a variety of species including a number of mammalian and primate species (BIOS laboratories, New Haven CT.) was digested with Eco-R1 and separated in 1% agarose gels. The gels were transblotted and membranes containing the digested DNA hybridized using the same procedure and conditions as that for Northern analysis.

Evolutionary conservation:

The evolutionary conservation of PEDF among a number of phylogenetically related species was examined. The results are presented in Figure 5. Using these high stringency hybridization conditions, a large EcoRI restriction fragment of approximately 23 kb is observed in aves, mammals and primates. No hybridization signals were seen in lower species (Figure 5A) possible due to weak homology of the human PEDF probe used. The EcoRI fragment for both chicken and mouse is somewhat smaller than that for humans. An interesting restriction pattern emerges in several of the mammalian species examined (Figure 5B). Several smaller restriction fragments ranging in size between 6 kb and 2 kb are seen. The larger fragments range in size between 9 kb and 23 kb and are seen in all primates species examined which has an additional strongly hybridizing polymorphic fragment at approximately 9 kb.

EXAMPLE 12

Neuronotrophic Effects of Pigment Epithelium Derived Factor On Cerebellar Granule Cells In Culture

Cell Culture

Cerebellar granule cells (CGC) were prepared from 5 or 8-day-old Sprague-Dawley rat pups as described by Novelli et al. (1988, *Brain Res.*, 451:205-212). In brief, tissue free of meninges was minced in a buffer containing 124 mM NaCl, 1mM NaH_2PO_4, 1.2 mM $MgSO_4$, 3 mg/ml bovine serum albumin (BSA), 27 μM phenol red, and 25 mM HEPES (pH 7.4), and centrifuged at 550 xg for 3 min. The tissue pellet from 10-20 animals was resuspended and trypsinized (15 min, 37°C) in 30ml of the same buffer containing 250 μg/ml trypsin; a further 15 ml of buffer was added containing 26 μg/ml DNase I, 166 ug/ml soybean trypsin inhibitor, and 0.5 mM additional $MgSO_4$ and the tissue was centrifuged again as described above. The pellet was resuspended in 1 ml of buffer supplemented with 80 μg/ml DNase, 0.52 mg/ml of trypsin inhibitor, and 1.6 mM additional $MgSO_4$, and triturated 60 times with a Pasteur pipette. The suspension was diluted with 2 ml of buffer containing 0.1 mM $CaCl_2$ and 1.3 mM additional $MgSO_4$, and undissociated material allowed to settle for 5 min. The supernatant was transferred to another tube, cells were recovered by brief centrifugation and resuspended in serum-containing medium (Eagle's basal medium with 25 mM KCl, 2 mM glutamine, 100 μg/ml gentamycin, and 10% heat inactivated fetal calf serum) or chemically defined medium (DMEM:F 12 (1:1) with 5 μg/ml insulin, 30 nM selenium, 100 μg/ml transferrin, 1000 nM putrescine, 20 nM progesterone, 50 U/ml penicillin, 50 μg/ml streptomycin, and 2 mM glutamine) (Bottenstein, 1985 Cell Culture in the Neurosciences, J.E. Bottenstein and G. Sato, eds. New York Plenum Publishing Corp. p. 3-43). Cells were plated in poly-L-lysine-coated 96 well plates (for MTS assay and

neurofilament ELISA assay) or 8-well chamber slides (for immunocytochemistry and BrdU labelling) at 2.5 x 10^5 cells/cm^2 and grown at 37°C in an atmosphere consisting of 5% CO_2 in air. After 1 day in culture, cytosine arabinoside (Ara-C) was added only to cells in serum-supplemented medium (final concentration 50μM).

MTS Assay

Cerebellar granule cells in 96 well plates were incubated in a CO_2 incubator for 4 hours with MTS (3-(4,5-dimethylthiazol-2-yl)-5-(3-carboxymethoxyphenyl)-2-(4-sulfophenyl)-2H-tetrazolium, inner salt) and PMS (phenazine methosulfate) final concentration; 333 μg/ml MTS and 25 μM PMS) (Promega Corp.). In the presence of PMS, MTS is converted to a water-soluble formazan by a dehydrogenase enzyme found in metabolically active cells (Cory et al. (1991) *Cancer Comm*, 3:207-212). The quantity of formazan product was determined by spectrophotometry at 490 nm.

Immunocytochemistry

After 7 days *in vitro* (DIV), the cells were washed three times in calcium-and magnesium-free phosphate-buffered saline (PBS) and fixed with 2% paraformaldehyde for 10 min, followed by 10 min at -20°C in 95% ethanol/5% acetic acid. Incubation with primary antibodies against NSE (neuron specific enolase), GABA, calbindin, or glial fibrillary acidic protein (GFAP) was carried out for 60 min at RT. Antibodies were applied at 1:1000-1:5000 in the presence of 2% normal goat serum and 0.2% BSA. The antibodies were visualized using the ABC system (Vector Laboratories) and diaminobenzidine. At least 20 fields were counted from 2-3 wells for each experiment. The average number of cells per field was then calculated to determine the ratio for the number of cells stained by the other antibodies relative to NSE-positive cells in control cultures.

Bromodeoxyridine (BrdU) Labeling

BrdU labeling was performed by the method of Gao et al. (1991 *Neuron*, 6: 705-715) with the following modification. The cells were plated in 8-well chamber slides and rPEDF added immediately. After 24 hours, BrdU (1:100; Amersham cell proliferation kit) was added to the culture medium for 24 hours, after which the cells were fixed in 2% paraformaldehyde (10 min), treated with 95% ethanol / 5 acetic acid (10 min), and incubated with an anti-BrdU monoclonal antibody (1:20 for 2 hrs). The cultures were then incubated with a horseradish peroxidase-conjugated goat anti-mouse secondary antibody for 60 min. After diaminobenzidine-peroxidase, the cells were mounted in Gel Mount. The mitotic index was determined by counting the percentage of labeled cells with a microscopy. For each value, a random sample of 3000 cells was counted.

Neurofilament ELISA Assay

The neurofilament ELISA was performed according to the method of Doherty et al. (1984 *J. Neurochem.*, 42:1116-1122) with slight modification. Cultures grown in 96-well microtiter plates were fixed with 4% paraformaldehyde in PBS at 4°C for 2 hr. The fixed cells were permeabilized by treatment for 15 min with 0.1% Triton X-100 in PBS, followed by incubation for 60 min with PBS containing 10% goat serum to block nonspecific binding. The cultures were then incubated with a monoclonal anti-neurofilament antibody overnight at 4°C (RMO-42 at 1:100; which stains only neurites in the cultures of cerebellar granule cells). After washing twice with PBS containing 10% goat serum, cells were incubated with secondary antibody (horseradish peroxidase-conjugated goat anti-mouse at 1:1000) for 1 hr. Following sequential washing with PBS and water, the cultures were incubated with 0.2% O-phenylenediamine and 0.02% H_2O_2 in 50

mM citrate buffer (pH 5.0) for 30 min. The reaction was stopped by adding an equal volume of 4.5 M H_2SO_4. Product formation was quantitated by reading the optical density (O.D.) of an aliquot of the reaction product at 490 nm using a microplate reader.

In order to validate the MTS assay as a measure of live cells, and to determine the range of cell number over which the results would be linear, the experiments shown in Figure 6 were carried out. In serum-containing medium (SCM) (Figure 6A), optical density (O.D.) was proportional to cell number plated over a range from 1-9 x 10^5 cells/cm_2. In contrast, for cells grown in chemically-defined medium (CDM) (Figure 6B), the linear range covered 1-5 x 10^5 cells/cm^2. For all subsequent experiments, cells were plated at 2.5 x 10^5 cells/cm^2, in the middle of the linear range for either type of culture medium.

Figure 7 shows that PEDF caused a significant increase in cell number by DIV4 with a larger difference at DIV7 and 10. However, the 2-3 fold increases were the result of large decreases in cell numbers in the control cultures. The dose-response curve in chemically-defined medium (Figure 8), showed that there is a statistically significant effect at 20ng/ml. Increasing the concentration of PEDF above 50 ng/ml did not produce further increases in CDM.

In order to determine whether the increase in O.D. (MTS assay) in response to PEDF reflected an increase in surviving cells or an increase in proliferation, a BrdU labeling study was performed using cultures from postnatal day 5 (P5) animals (a time when cerebellar granule cells are still dividing in the animal). Figure 9 shows the effect of PEDF on P5 CGC cultures at DIV1 and 2. Using the MTS assay, PEDF had no effect at DIV1 but caused a small increase in O.D. at DIV2 in either serum-containing medium or chemically defined medium. Therefore, BrdU was

added at day 1 and cells were fixed on day 2. The BrdU labeling index was 5% in SCM and 3% in CDM, under control conditions, and PEDF did not increase the BrdU labeling index in either culture medium (Figure 10). The lack of stimulation of the BrdU labeling index by PEDF implies that enhanced survival rather than increased cell division is responsible for the increased O.D. measured by the MTS assay after exposure to PEDF.

Immunocytochemistry was used to identify the cells present in cultures before and after treatment with PEDF. P8 cultures grown for 7 days with and without PEDF (500 ng/ml) were stained with four different antibodies: a polyclonal rabbit antibody to neuron-specific enolase (NSE), which recognizes all cerebellar neurons (Schmechel et al. (1978) *Science*, 199:313-315); a polyclonal antibody to GABA, which is synthesized in all cerebellar neurons except cerebellar granule cells (Gruol and Crimi (1988) *Dev. Brain Res.*, 41:135-146); an antibody to calbindin, which is a neuron-specific protein and GFAP, an intermediate filament protein present only in astrocytes. The results are summarized in Table 2. PEDF significantly increased the number of NSE-positive cells in both SCM (30% increase) and in CDM (60% increase). There was a small, not statistically significant, increase in the number of GABA-positive neurons and Purkinje cells (calbindin-positive). Thus, PEDF is neurotrophic only for granule neurons. In addition, PEDF significantly decreased the number of GFAP-positive astrocytes present in the cultures (30% decrease in SCM and 40% decrease in CDM). This "gliastatic" property of PEDF is further discussed in Example 14.

TABLE 2

Immunocytochemistry demonstrates that PEDF Increased The Number of NSE-Positive Cells (Neurons) But Decreased GFAP-Positive Cells (Glia)

Antigen	Treatment	SCM	CDM
NSE	Control	100.0 ± 6.2	100.0 ± 4.5
	PEDF	127.0 ± 5.9*	157.2 ± 7.4*
GABA	Control	2.8 ± 0.2	1.4 ± 0.2
	PEDF	3.2 ± 0.2	1.8 ± 0.2
Calbindin	Control	0.06 ± 0.01	0.07 ± 0.02
	PEDF	0.07 ± 0.02	0.12 ± 0.02
GFAP	Control	0.86 ± 0.07	0.99 ± 0.07
	PEDF	0.60 ± 0.03*	0.60 ± 0.06*

Postnatal-day 8 cerebellar granule cells were cultured in 8-well chamber slides. PEDF (500 ng/ml) was added at DIV 0, the cells were fixed on DIV 7, and the immunocytochemistry was carried out using antibodies against NSE, GABA, Calbindin and GFAP. At least 20 fields were counted from 2-3 wells for each experiment. Data are expressed as percent of control of NSE-positive cells. Each experiment value represents mean cell number ± SEM. *P<0.005 compared with each other control by using non-paired test.

In order to investigate the effects of PEDF on neurite outgrowth, a neurofilament ELISA assay was used. Immunocytochemistry had shown that the monoclonal antibody RMO-42, stained only the neurites of cerebellar granule cells in culture, so this antibody was used as a direct measure of neurofilament present only in processes and not the cell body (Figure 11). PEDF slightly increased neurofilament content, both in SCM and CDM, but the increase was directly proportional to the increase in cell number (Figure 12).

Figure 13 summarizes the data from this Example. By 10 days in culture, most untreated CGCs die (control) but 60% or more of the PEDF-treated cells remain viable. PEDF is thus a potent survival factor for brain neurons.

EXAMPLE 13

Neuronotrophic properties of rPEDF peptides, BP and BX.

Described in the previous sections on the "neuronotrophic" activity of PEDF is the fact that we can produce relatively large amounts of a recombinant PEDF (rPEDF) that exhibits potent neurotrophic activity. Using appropriate recombinant molecular biological technology, we can also produce smaller fragments of the PEDF molecule that can be tested for either neurotrophic or neuronotrophic activity. Figure 14 shows the effects of two of these truncated forms of PEDF on CGC viability. BX and BP are 24 and 28 kDa fragment from the amino-terminal portion of the PEDF molecule, respectively. Both fragments at 1x or 10x concentrations act as neuron-survival factors, significantly promoting the life of the CGC's. In this experiment, the peptide was given once at the beginning of the experiment and the cell number was determined 7 days later. We conclude that, along with the full PEDF molecule, smaller recombinant peptides near the N-terminal of the molecule are "neuronotrophic".

EXAMPLE 14

Gliastatic properties of PEDF

Along with neurons in the primary cultures of rat cerebellar granule cells are a small number of different types of glia. Glia are the "support" elements in the CNS for neurons, forming the architectural framework and the metabolic support system on which neurons depend. Glia are also of clinical importance since tumors of the brain are mostly formed by glia and gliosis is a problem in several neurodegenerative diseases. In our system, we first noticed an effect of PEDF on glia when we immunocytochemically stained the cultured mixed population of cells with antibodies specific for neurons and other antibodies specific for different types of glia. For this purpose, we used the standard markers Neuron-Specific Enolase (NSE) and others

to demonstrate the presence of neurons, Glial Fibrillary Acidic Protein (GFAP) to demonstrate the presence of astroglia and OX-42 to stain microglia. In this experiment (Table 2), we found the expected increase in NSE staining with PEDF treatment since we then knew that the neurons were living longer but we found an unexpected decrease in GFAP staining. This indicated the possibility of fewer astrocytes in the PEDF-treated cultures.

Because of the distinctive morphology of astroglia and microglia in the culture dishes and their selective staining for GFAP or OX-42, it is possible to individually count their numbers under the microscope under different experimental conditions. This has now been done as outlined in Figures 15 and 16. Figure 15 shows the effects of PEDF on numbers of astroglia in cultures obtained from rat brain at 2 weeks (2w) or 12 weeks (12w) in culture. Times given are 48 hrs, 96 hrs or 7 days after treatment with PEDF. Clearly, under all the conditions tested, PEDF treatment results in a dramatic decrease in the number of astroglia. Figure 16 shows a parallel analysis of microglia in the same cultures. Administration of PEDF for 48 hrs. or 7 days resulted in fewer numbers of the cells whether they has been cultured for 2 weeks (2W) or 12 weeks (12W). Thus, PEDF substantially decreases glial elements over a very long period of time while acting as a survival factor for neurons.

EXAMPLE 15

Characterization of Native Bovine PEDF

Since the specific antibody indicated the presence of PEDF in the adult IPM, we used bovine IPM washes as a source for purification of native PEDF. Although RPE and retinal cells express PEDF mRNA, anti-BH could not detect PEDF bands on Western transfers in these cell extracts, suggesting a rapid PEDF release into the IPM. We now estimate that PEDF is present in bovine IPM

at less than 1% of the total soluble protein (i.e. about 2-5 ng/bovine eye). At physiological temperatures, the PEDF protein in the IPM remains stable for extended periods of time and does not form non-reduced complexes resistant to SDS. Thus, its potential usefulness in culture experiments and transplantation *in vivo*. is greatly enhanced due to its stable nature.

Purification to apparent homogeneity is achieved by a simple two-step procedure (Figure 17). Components of IPm were fractionated by size-exclusion column chromatography (TSK-3000). The PEDF-immunoreactive fractions were pooled, applied to a cation-exchange column (Mono-S) and immunoreactivity was eluted with a NaCl linear gradient. Purification protocol is detailed in Materials and Methods. Elution profiles of each chromatography are shown in: panel A, TSK-3000 size-exclusion column chromatography, and panel B, mono-S column chromatography. Absorbance at 280 nm is represented by ____, and NaCl concentration by ---, PEDF-immunoreactivity was followed with antiserum Ab-rPEDF. The inserts correspond to Western blot analysis of the indicated fractions. Immunoreaction was performed with a 1:10,000 dilution of Ab-rPEDF and stained with 4-chloro-1-napthtol. Molecular size standards for the TSK-3000 chromatography were: BSA, bovine serum albumin (66,000); and CA, bovine carbonic anhydrase (29,000).

Starting with a wash of soluble IPM components, the first step involves removal of the most abundant protein, IRBP, by size exclusion chromatography. PEDF elutes as a monomeric polypeptide around 50 kDa in size. Since we have determined that PEDF's isoelectric point is 7.2-7.8, we have used S-sepharose column chromatography at pH 6.0 in the second step of our procedure to simultaneously purify and concentrate the protein. Purified protein is recovered at about 2 ug protein per adult bovine eye with a recovery of about 40%. Native

PEDF behaves like a monomeric glycoprotein with an apparent molecular weight of 49,500±1,000 on SDS-PAGE.

The purified protein is sensitive to glycosidase F, revealing N-linked oligosaccharides that account for up to 3,000-Mr of the native protein (Figure 18). To remove asparagine-linked oligosaccharides purified PEDF protein was treated with endoglycosidase H and *N*-Glycosidase F. Enzymatic reactions were performed as described in Materials and Methods with a total of 200 ng of PEDF protein in the presence or absence of β-mercaptoethanol. Reactions mixtures were applied to SDS-12.5% polyacrylamide gel. Photographs of western transfers of endoglycosidase H (left panel) and *N*-Glycosidase F (right panel) reactions are shown. Immunoblots were treated with antiserum Ab-rPEDF diluted 1:10,000. Addition in each reaction are indicated at the top. The numbers at the right side of each photograph indicate the migration of biotinylated SDS-PAGE standards: bovine serum albumin (66,200), ovalbumin (45,000) and bovine carbonic anhydrase (31,000). We have shown that purified bovine PEDF promotes neurite outgrowth on Y-79 cells and Weri retinoblastoma cells, and that this activity is blocked by Anti-rPEDF (see below).

The present invention provides the tools for determining the effect of authentic PEDF on the expression of neuronal and glial markers in the CGC cultures and Y-79 tumor cells including NSE, GFAP, neurofilament (NF-200) protein.

EXAMPLE 16

Pigment Epithelium-Derived Factor: Characterization Using A Highly Specific Polyclonal Antibody

We have used purified recombinant human PEDF produced in *E. coli* to develop polyclonal antibodies against PEDF. Anti-rPEDF specifically recognized one polypeptide on Western transfer of IPM wash from adult bovine eyes (Figure 19). Polyclonal antiserum to human

recombinant PEDF specifically recognizes rPEDF. Western transfer and slot blot of human rPEDF were treated with rabbit polyclonal antiserum to rPEDF, Ab-rPEDF. Photographs of immunostaining with 4-chloro-naphthol are shown. Panel A, Western transfers of 0.5 μg of rPEDF were used to assay increasing dilutions of antiserum. rPEDF protein was resolved by SDS-12.5% PAGE before transfer. Dilutions are indicated at the top of each lane. Diluted antiserum was preincubated with rPEDF at 5 μg/ml before using for immunodetection and is indicated as 1:10,000+rPEDF. The numbers to the left indicate the molecular weight of biotinylated SDS-PAGE standards. Panel B increasing amounts of rPEDF in 1% BSA/PBS were applied to a nitrocellulose membrane with a manifold. The membranes were treated with antiserum Anti-rPEDF and rabbit preimmune serum diluted 1:10,000. The numbers to the right indicate the amounts of rPEDF protein blotted on the membrane. The sera used in each paper are indicated at the top of the figure.

Anti-BH specifically recognizes human PEDF on Western transfers at dilutions as low as 1:50,000; importantly, it does not recognize serum α_1-antitrypsin. The antibody recognizes one major band on Western transfers of conditioned medium from juvenile monkey RPE cells in culture as well as of IPM from adult bovine eyes. Anti-rPEDF blocked the IPM-promoting neurotrophic activity (Figure 20). Human retinoblastoma Y-79 cells exponentially growing in serum containing medium were washed twice with PBS, and plated (2.5×10^5) cell per ml) in serum-free MEM supplemented with insulin, transferring and selenium (ITS mix, Collaborative Research Products). Effectors were then added to the cultures. After 7 days at 37°C in 5% CO_2, the cells were attached to poly-D-lysine coated plates with fresh serum-free medium. The differentiation state of the cultures was monitored at different intervals after attachment. Morphology characteristic of 9-day

post-attachment cultures is shown. Addition of effectors were as indicated in each panel at the following final concentrations: 125 μg/ml BSA, 1% IPM, and 100 ng/ml purified bovine PEDF. In order to block the neurite outgrowth inducing activity each effector was preincubated with an excess of antiserum Anti-rPEDF (1 μl) in 1% BSA/PBS at 4°C for at least 6 hours. All photographs are shown at x50 magnification.

The anti-rPEDF also blocked the neurite-outgrowth activity promoted by the purified PEDF. Our data indicate that PEDF is the only neurotrophic factor in the IPM. These results also suggest that the anti-rPEDF will be useful in probing the PEDF neurotrophic active site as well as the physiological role of PEDF in the IPM and other tissues (e.g. brain) as well. Further, these results indicate that PEDF is a bona fide component of the IPM and is probably the sole neurotrophic component in the extracellular matrix. Moreover, the protein is present in a wide range of tissues and extracellular spaces. The blocking antibody is useful in studies probing the physiological functions of PEDF.

EXAMPLE 17

Pigment Epithelium-Derived Factor: A Serpin With Neurotrophic Activity

The amino acid sequence derived from a fetal human PEDF cDNA shares identity of its primary structure (~30%) with the serine protease inhibitor (serpin) family, preserving 90% of the residues essential for the structural integrity of serpins. However, recombinant PEDF does not inhibit the serine proteases trypsin, chymotrypsin, elastase or cathepsin G. A natural target for PEDF has not yet been identified. We have analyzed proteins from the interphotoreceptor matrix (IPM), the space between the retinal pigment epithelium and the retina by immunodetection on Western blots with antibodies raised against PEDF and by zymography in gels containing

casein as a proteolytic substrate. Our results show that bovine IPM contains a stable, glycosylated PEDF polypeptide (50,000 Mr) at about 2-5μg per eye. Limited proteolysis of bovine PEDF produced a polypeptide of 46,000 Mr with trypsin, subtilisin, chymotrypsin and elastase, suggesting a globular structure with a hinge region susceptible to proteolytic cleavage. On the other hand, casein SDS-PAGE zymography revealed low protease activity in the IPM which migrated as a double of about 80,000 ± 5,000 Mr. The caseinolytic activities were inhibited 100% with 1 μg/ml aprotinin and 10mM PMSF added to the gel mixture, but were not affected by E64 or EDTA. Importantly, IPM protein did not react with antibody against plasminogen, a serine protease of about 80,000 Mr. When rPEDF protein was added at 1 μg/ml, the signal for these caseinolytic activities, as well as another serine protease activity of unknown origin, diminished by about 50%. Our results suggest the IPM as a natural extracellular site for a novel serine protease and the serpin PEDF, both present at ≤1% of the total protein.

All of the references cited herein are hereby incorporated in their entireties by reference.

The present invention discloses the general structural features of PEDF and beginnings of understanding of how these relate to function of the protein. PEDF possesses the structural features and general tertiary characteristics previously attributed to serpins but not its anti-protease activity. PEDF is a neurotrophic protein and appears to be the sole component of the IPM that promotes neurite-outgrowth on retinoblastoma cells. However, the reactive center for serine protease inhibition found near the carboxy terminal of classical serpins is not necessary for PEDF's neurotrophic biological activity. Specifically, a polypeptide chain containing a domain from the amino-terminal portion of the molecule (BA) is sufficient for

neurotrophic and neuron-survival activity. The present invention further allows for determination of whether the CGC neurons normally die by apoptosis and whether PEDF is an apoptosis inhibitor. In other words, the present invention allows one to determine by what mechanism PEDF "saves" neurons and "inhibits" glia growth or proliferation.

The present invention is useful in determining the specific neurotrophic "active site". Further, the use of rPEDF truncated peptides allows us to define the elements necessary for neuronotrophic and perhaps gliastatic activity of PEDF. The present invention further provides necessary tools to study the interactions of PEDF that trigger the signal for differentiation of retinoblastoma. Recent experiments demonstrate that ^{125}I-BH binds to retinoblastoma cells in competitive fashion only when added in medium that had been previously "conditioned" by retinoblastoma cells. This suggests that one or more co-factors produced by the cells could be required for binding. The present invention further provides the tools necessary to identify and characterize a putative cell-surface receptor for PEDF or for a PEDF complex from our CGC and retinoblastoma test systems.

Recombinant mutated proteins, proteolytic products and synthetic peptides have become instrumental in domain mapping of functional sites of proteins. Further, the recombinant proteins of the present invention allow the mapping of neurotrophic and neuronotrophic "active sites" on the PEDF molecule and the determination of the cellular transduction mechanism through which this interesting protein exerts its dramatic biological effects.

While this invention has been described with an emphasis upon preferred embodiments, it will be obvious to those of ordinary skill in the art that variations in the preferred nucleic acids coding for, and the amino acid

sequences of, PEDF, rPEDF, and equivalent proteins, (BP, BX, BA) the vectors utilizing any such nucleic acids, the recombinant methods of producing such proteins, and the methods of using such proteins, may be realized and that it is intended that the invention may be practiced otherwise than as specifically described herein. Accordingly, this invention includes all modifications encompassed within the spirit and scope of the invention as defined by the following claims.

SEQUENCE LIST

(1) GENERAL INFORMATION:

(i) APPLICANTS: Chader, Gerald J.; Becerra, Sofia Patricia; Schwartz, Joan P.; Taniwaki, Takayuki

(ii) TITLE OF INVENTION: PIGMENT EPITHELIUM DERIVED FACTOR: CHARACTERIZATION GENOMIC ORGANIZATION AND SEQUENCE OF THE PEDF GENE

(iii) NUMBER OF SEQUENCES: 43

(iv) CORRESPONDENCE ADDRESS:
(A) ADDRESSEE: Morgan & Finnegan, L.L.P.
(B) STREET: 345 Park Avenue
(C) CITY: New York
(D) STATE: New York
(E) COUNTRY: USA
(F) ZIP: 10154

(v) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Floppy Disk
(B) COMPUTER: IBM PC Compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: WORDPERFECT 5.1

(vi) CURRENT APPLICATION DATA:
(A) APPLICATION NO: TO BE ASSIGNED
(B) FILING DATE: 06-JUN-1995
(C) CLASSIFICATION:

(vii) PRIOR APPLICATION DATA:
(A) APPLICATION NO: 08/367,841
(B) FILING DATE: 30-DEC-1994

(vii) PRIOR APPLICATION DATA:
(A) APPLICATION NUMBER: 08/257,963
(B) FILING DATE: 07-JUN-1994

(vii) PRIOR APPLICATION DATA:
(A) APPLICATION NUMBER: 07/952,796
(B) FILING DATE: 24-SEP-1992

(viii) ATTORNEY/AGENT INFORMATION:
(A) NAME: DOROTHY R. AUTH
(B) REGISTRATION NUMBER: 36434
(C) REFERENCE/DOCKET NUMBER: 20264126PCT

(ix) TELECOMMUNICATION INFORMATION:
(A) TELEPHONE: (212) 758-4800
(B) TELEFAX: (212) 751-6849

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1512 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA to mRNA

(ix) FEATURE:
(A) NAME/KEY:
(B) LOCATION:
(D) OTHER INFORMATION: PEDF coding region

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

GCTGTAATCT GAAGCCTGCT GGACGCTGGA TTAGAAGGCA 40

GCAAAAAAAG CTCTGTGCTG GCTGGAGCCC CCTCAGTGTG 80

CAGGCTTAGA GGGACTAGGC TGGGTGTGGA GCTGCAGCGT 120

ATCCACAGGC CCCAGGATGC AGGCCCTGGT GCTACTCCTC 160

TGCATTGGAG CCCTCCTCGG GCACAGCAGC TGCCAGAACC 200

CTGCCAGCCC CCCGGAGGAG GGCTCCCCAG ACCCCGACAG 240

CACAGGGGCG CTGGTGGAGG AGGAGGATCC TTTCTTCAAA 280

GTCCCCGTGA ACAAGCTGGC AGCGGCTGTC TCCAACTTCG 320

GCTATGACCT GTACCGGGTG CGATCCAGCA TGAGCCCCAC 360

GACCAACGTG CTCCTGTCTC CTCTCAGTGT GGCCACGGCC 400

CTCTCGGCCC TCTCGCTGGG AGCGGAGCAG CGAACAGAAT 440

CCATCATTCA CCGGGCTCTC TACTATGACT TGATCAGCAG 480

CCCAGACATC CATGGTACCT ATAAGGAGCT CCTTGACACG 520

GTCACTGCCC CCCAGAAGAA CCTCAAGAGT GCCTCCCGGA 560

TCGTCTTTGA GAAGAAGCTG CGCATAAAAT CCAGCTTTGT 600

GGCACCTCTG GAAAAGTCAT ATGGGACCAG GCCCAGAGTC 640

CTGACGGGCA ACCCTCGCTT GGACCTGCAA GAGATCAACA 680

ACTGGGTGCA GGCGCAGATG AAAGGGAAGC TCGCCAGGTC 720

ACTGGGTGCA GGCGCAGATG AAAGGGAAGC TCGCCAGGTC 720

CACAAAGGAA ATTCCCGATG AGATCAGCAT TCTCCTTCTC 760

GGTGTGGCGC ACTTCAAGGG GCAGTGGGTA ACAAAGTTTG 800

ACTCCAGAAA GACTTCCCTC GAGGATTTCT ACTTGGATGA 840

AGAGAGGACC GTGAGGGTCC CCATGATGTC GGACCCTAAG 880

GCTGTTTTAC GCTATGGCTT GGATTCAGAT CTCAGCTGCA 920

AGATTGCCCA GCTGCCCTTG ACCGGAAGCA TGAGTATCAT 960

CTTCTTCCTG CCCCTGAAAG TGACCCAGAA TTTGACCTTG 1000

ATAGAGGAGA GCCTCACCTC CGAGTTCATT CATGACATAG 1040

ACCGAGAACT GAAGACCGTG CAGGCGGTCC TCACTGTCCC 1080

CAAGCTGAAG CTGAGTTACG AAGGCGAAGT CACCAAGTCC 1120

CTGCAGGAGA TGAAGCTGCA ATCCTTGTTT GATTCACCAG 1160

ACTTTAGCAA GATCACAGGC AAACCCATCA AGCTGACTCA 1200

GGTGGAACAC CGGGCTGGCT TTGAGTGGAA CGAGGATGGG 1240

GCGGGAACCA CCCCCAGCCC AGGGCTGCAG CCTGCCCACC 1280

TCACCTTCCC GCTGGACTAT CACCTTAACC AGCCTTTCAT 1320

CTTCGTACTG AGGGACACAG ACACAGGGGC CCTTCTCTTC 1360

ATTGGCAAGA TTCTGGACCC CAGGGGCCCC TAATATCCCA 1400

GTTTAATATT CCAATACCCT AGAAGAAAAC CCGAGGGACA 1440

GCAGATTCCA CAGGACACGA AGGCTGCCCC TGTAAGGTTT 1480

CAATGCATAC AATAAAAGAG CTTTATCCCT GC 1512

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 418 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(ix) FEATURE:

(A) NAME/KEY: CDS
(B) LOCATION: 117..1373
(D) OTHER INFORMATION: /note= "product = "pigment epithelial-derived factor" gene = "PEDF" codon_start = 1"

(ix) FEATURE:
(A) NAME/KEY:
(B) LOCATION:
(D) OTHER INFORMATION: PEDF amino acid sequence

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

```
Met Gln Ala Leu Val Leu Leu Leu Cys Ile Gly Ala
 1               5                  10
Leu Leu Gly His Ser Ser Cys Gln Asn Pro Ala Ser
            15                  20
Pro Pro Glu Glu Gly Ser Pro Asp Pro Asp Ser Thr
 25                  30                  35
Gly Ala Leu Val Glu Glu Glu Asp Pro Phe Phe Lys
             40                  45
Val Pro Val Asn Lys Leu Ala Ala Ala Val Ser Asn
     50                  55                  60
Phe Gly Tyr Asp Leu Tyr Arg Val Arg Ser Ser Met
                 65                  70
Ser Pro Thr Thr Asn Val Leu Leu Ser Pro Leu Ser
            75                  80
Val Ala Thr Ala Leu Ser Ala Leu Ser Leu Gly Ala
 85                  90                  95
Glu Gln Arg Thr Glu Ser Ile Ile His Arg Ala Leu
            100                 105
Tyr Tyr Asp Leu Ile Ser Ser Pro Asp Ile His Gly
    110                 115                 120
Thr Tyr Lys Glu Leu Leu Asp Thr Val Thr Ala Pro
                125                 130
Gln Lys Asn Leu Lys Ser Ala Ser Arg Ile Val Phe
            135                 140
Glu Lys Lys Leu Arg Ile Lys Ser Ser Phe Val Ala
145                 150                 155
Pro Leu Glu Lys Ser Tyr Gly Thr Arg Pro Arg Val
                160                 165
Leu Thr Gly Asn Pro Arg Leu Asp Leu Gln Glu Ile
    170                 175                 180
Asn Asn Trp Val Gln Ala Gln Met Lys Gly Lys Leu
                185                 190
Ala Arg Ser Thr Lys Glu Ile Pro Asp Glu Ile Ser
            195                 200
```

```
Ile Leu Leu Leu Gly Val Ala His Phe Lys Gly Gln
205                 210                 215
Trp Val Thr Lys Phe Asp Ser Arg Lys Thr Ser Leu
            220                 225
Glu Asp Phe Tyr Leu Asp Glu Glu Arg Thr Val Arg
    230                 235                 240
Val Pro Met Met Ser Asp Pro Lys Ala Val Leu Arg
                245                 250
Tyr Gly Leu Asp Ser Asp Leu Ser Cys Lys Ile Ala
        255                 260
Gln Leu Pro Leu Thr Gly Ser Met Ser Ile Ile Phe
265                 270                 275
Phe Leu Pro Leu Lys Val Thr Gln Asn Leu Thr Leu
            280                 285
Ile Glu Glu Ser Leu Thr Ser Glu Phe Ile His Asp
    290                 295                 300
Ile Asp Arg Glu Leu Lys Thr Val Gln Ala Val Leu
                305                 310
Thr Val Pro Lys Leu Lys Leu Ser Tyr Glu Gly Glu
        315                 320
Val Thr Lys Ser Leu Gln Glu Met Lys Leu Gln Ser
325                 330                 335
Leu Phe Asp Ser Pro Asp Phe Ser Lys Ile Thr Gly
            340                 345
Lys Pro Ile Lys Leu Thr Gln Val Glu His Arg Ala
    350                 355                 360
Gly Phe Glu Trp Asn Glu Asp Gly Ala Gly Thr Thr
                365                 370
Pro Ser Pro Gly Leu Gln Pro Ala His Leu Thr Phe
        375                 380
Pro Leu Asp Tyr His Leu Asn Gln Pro Phe Ile Phe
385                 390                 395
Val Leu Arg Asp Thr Asp Thr Gly Ala Leu Leu Phe
            400                 405
Ile Gly Lys Ile Leu Asp Pro Arg Gly Pro
    410                 415
```

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 379 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(ix) FEATURE:
(A) NAME/KEY: Region
(B) LOCATION: 1..4

(D) OTHER INFORMATION: /note= "Met 1...Ile 4 is an N-terminal fusion to Asp 26...Pro 400 of SEQ ID NO:2; Met -18...Glu 25 of SEQ ID NO:2 is deleted"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

```
Met Asn Arg Ile Asp Pro Phe Phe Lys Val Pro Val
 1               5                  10
Asn Lys Leu Ala Ala Ala Val Ser Asn Phe Gly Tyr
            15                  20
Asp Leu Tyr Arg Val Arg Ser Ser Met Ser Pro Thr
 25                  30                  35
Thr Asn Val Leu Leu Ser Pro Leu Ser Val Ala Thr
                40                  45
Ala Leu Ser Ala Leu Ser Leu Gly Ala Glu Gln Arg
        50                  55                  60
Thr Glu Ser Ile Ile His Arg Ala Leu Tyr Tyr Asp
                    65                  70
Leu Ile Ser Ser Pro Asp Ile His Gly Thr Tyr Lys
            75                  80
Glu Leu Leu Asp Thr Val Thr Ala Pro Gln Lys Asn
 85                  90                  95
Leu Lys Ser Ala Ser Arg Ile Val Phe Glu Lys Lys
                100                 105
Leu Arg Ile Lys Ser Ser Phe Val Ala Pro Leu Glu
        110                 115                 120
Lys Ser Tyr Gly Thr Arg Pro Arg Val Leu Thr Gly
                    125                 130
Asn Pro Arg Leu Asp Leu Gln Glu Ile Asn Asn Trp
            135                 140
Val Gln Ala Gln Met Lys Gly Lys Leu Ala Arg Ser
145                 150                 155
Thr Lys Gln Ile Pro Asp Glu Ile Ser Ile Leu Leu
                160                 165
Leu Gly Val Ala His Phe Lys Gly Gln Trp Val Thr
        170                 175                 180
Lys Phe Asp Ser Arg Lys Thr Ser Leu Glu Asp Phe
                    185                 190
Tyr Leu Asp Glu Glu Arg Thr Val Arg Val Pro Met
            195                 200
Met Ser Asp Pro Lys Ala Val Leu Arg Tyr Gly Leu
205                 210                 215
Asp Ser Asp Leu Ser Cys Lys Ile Ala Gln Leu Pro
                220                 225
Leu Thr Gly Ser Met Ser Ile Ile Phe Phe Leu Pro
        230                 235                 240
Leu Lys Val Thr Gln Asn Leu Thr Leu Ile Glu Glu
                    245                 250
Ser Leu Thr Ser Glu Phe Ile His Asp Ile Asp Arg
            255                 260
Glu Leu Lys Thr Val Gln Ala Val Leu Thr Val Pro
265                 270                 275
Lys Leu Lys Leu Ser Tyr Glu Gly Glu Val Thr Lys
                280                 285
Ser Leu Gln Glu Met Lys Leu Gln Ser Leu Phe Asp
        290                 295                 300
Ser Pro Asp Phe Ser Lys Ile Thr Gly Lys Pro Ile
                    305                 310
```

```
Lys Leu Thr Gln Val Glu His Arg Ala Gly Phe Glu
        315                 320
Trp Asn Glu Asp Gly Ala Gly Thr Thr Pro Ser Pro
325                 330                 335
Gly Leu Gln Pro Ala His Leu Thr Phe Pro Leu Asp
            340                 345
Tyr His Leu Asn Gln Pro Phe Ile Phe Val Leu Arg
    350                 355                 360
Asp Thr Asp Thr Gly Ala Leu Leu Phe Ile Gly Lys
                365                 370
Ile Leu Asp Pro Arg Gly Pro
        375
```

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (synthetic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

AGYAAYTTYT AYGAYCTSTA 20

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (synthetic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

CTYTCYTCRT CSAGRTARAA 20

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 19 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

Thr Ser Leu Glu Asp Phe Tyr Leu Asp Glu Glu Arg
1 5 10
Thr Val Arg Val Pro Met Met
15

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 29 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

Ala Leu Tyr Tyr Asp Leu Ile Ser Ser Pro Asp Ile
1 5 10
His Gly Thr Tyr Lys Glu Leu Leu Asp Thr Val Thr
15 20
Ala Pro Gln Xaa Asn
25

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 7 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

Met Asn Glu Leu Gly Pro Arg
1 5

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 4421 Base Pairs
(B) TYPE: Nucleic Acid
(C) STRANDEDNESS: Double
(D) TOPOLOGY: Unknown

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Human

(ix) FEATURE:
(A) NAME/KEY: JT1
(B) LOCATION:
(C) IDENTIFICATION METHOD:
(D) OTHER INFORMATION: 7.1 kb Bam HI fragment Derived from human placental genomic DNA; Also referred to as JT101

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

GGATCCCTTG GTTGGGGTGT TGGGGAAGGC AGGGTTTTAA 40

CGGAAATCTC TCTCCATCTC TACAGAGCTG CAATCCTTGT 80

TTGATTCACC AGACTTTAGC AAGATCACAG GCAAACCCAT 120

CAAGCTGACT CAGGTGGAAC ACCGGGCTGG CTTTGAGTGG 160

AACGAGGATG GGGCGGGAAC CACCCCCAGC CCAGGGCTGC 200

AGCCTGCCCA CCTCACCTTC CCGCTGGACT ATCACCTTAA 240

CCAGCCTTTC ATCTTCGTAC TGAGGGACAC AGACACAGGG 280

GCCCTTCTCT TCATTGGCAA GATTCTGGAC CCCAGGGGCC 320

CCTAATATCC CAGTTTAATA TTCCAATACC CTAGAAGAAA 360

ACCCGAGGGA CAGCAGATTC CACAGGACAC GAAGGCTGCC 400

CCTGTAAGGT TTCAATGCAT ACAATAAAAG AGCTTTATCC 440

CTAACTTCTG TTACTTCGTT CCTCCTCCTA TTTTGAGCTA 480

TGCGAAATAT CATATGAAGA GAAACAGCTC TTGAGGAATT 520

TGGTGGTCCT CTACTTCTAG CCTGGTTTTA TCTAAACACT 560

GCAGGAAGTC ACCGTTCATA AGAACTCTTA GTTACCTGTG 600

TTGGATAAGG CACGGACAGC TTCTCTGCTC TGGGGGTATT 640

TCTGTACTAG GATCAGTGAT CCTCCCGGGA GGCCATTTCC 680

TGCCCCCATA ATCAGGGAAG CCTGCTCGTA AACAACACAT 720

GGACAGATAG GAGAGGCCAT TTGTAACTTA AGGAAACGGA 760

CCCGATACGT AAAGATTCTG AACATATTCT TTGTAAGGAG 800

GTATGCCTAT TTTACAAAGT ACAGCCGGGT GTGGTGGCTC 840

ATGGCTATAA TCCCAGCACT TTGGGAGGCC GAGGCGGGCG 880

GATCACCTGA GATCAGGAGT TTGAGACCAG CCTGACCAAC 920

ACGGAGAAAC CCCGTCTGTA CTAAAAATAC AAAATTAGCA 960

GGGTGTGGTG GTACATGCCT GTAATCCCAG CTACTGGGGA 1000

GGCTGAGGCA GGAGAATCAC TTGAACCCGG GAGGCGGAGG 1040

TTGCAGTGAG CCGAGATCAC GCCATTGCAC TCCAATCTAG 1080

GCAATAAGAG CAAAACTCCG TCTCAAACAA CAAAAAACCA 1120

AAGTATAACT GGGCTTTTTG AAGAACATGA AACATGCCCA 1160

GTGTCTGAAG TAGAATAACT ACCGAACTGT CCGTAGGACT 1200

AAACTTTTTC TTGAAAAAGC TCTACCAAAA AAAGTCACCG 1240

GCCACTCCCT TGTCACAGTT ATTAGACAGG AGGAGAAATG 1280

ATAATTCTAC TGCCCTTCAT TCTACAAATG TTTGAGTGCT 1320

AACTGTATTC CAGATTCTCA AAAAGCTATT GCCAGGTATC 1360

TCTGGGGCTA CTGATTTCCT GATCATAATG CAATGGCAAC 1400

CAACAGGCAC TTGGGCATGG TGAGGGTGGG CAAGCTTTCA 1440

AAAGCAGCGT GGATCTGGCA TTCTTTTCCA CGAATGCACC 1480

TCAACTACTT GGCACCAGTG GTAACACAGC AACCAGGGTT 1520

CCGACCTAGA GAATCCCGTA ACCTTCTGAC TGGAACGGGG 1560

TCTGGGCTGT CGCTACACAT CCTGGTGGAA GGCAGCTATC 1600

ATCCCTACCT TCTGCCTTCT GTCTCTTAAA TCTGAACCAC 1640

AAACAGCAAC GTCCATACCC TCAGCATTGT TAGAATCCCC 1680

TGCAGCCTCC AGTTCTCATA CTGTCTGTAT TCTACTCGCC 1720

AGTTTGGAGA GGTCTGGTGG AGAAAAGGAG TCTCTTTTCA 1760

GGCTTGACAA CAAATAGAAC TCAGGGCCGG GCGCGGTGGC 1800

TCACGCCTGT CATCCCAGCA CTGTGGGAGG CCGAAGCGGG 1840

CGGATCACCT GAGGTCGGGA GCTCAAGACC AGCCTGGCCA 1880

ACATGGAGAA ATCCCATCTT TACTAAAAAT ACAAAATTAG 1920

CCGGGCGTAC TGGCGAATGC CTGTAATGCC AGCTTCTCGG 1960

GAGGCTGAGG CAGGAGAATC GCTTGAACCT GGGAGGCAGA 2000

GGTTGCGGTG AGCCAAGACT GTGCCACTGT ACTCCAGCCT 2040

TGGTGACAGA GGGAGACTCT GTCTTAAGAA AAAAAGAAAA 2080

AAAAAAAAAA AGGGCCGGGC TCACGCCTGT AATCCCAGCA 2120

CTTTGGGAGG CCAAATCACC TGAGGCCGGG AGTTTGATAC 2160

CAACCTGACC AACATAGTGA AATCCCGTCT CTACTAAAAA 2200

TACAAAATTA GCCAGGCGTG GTGGCGGGCG CCTGTAATCC 2240

CAGCTACTCG GGAGGCTGAA GCAGGAGAAT CACTTGAACC 2280

CGGAAGGCGG AGGTTGCCGT AAGCCAAGAT CGCGCCATTG 2320

CGCTCCAGCC TGGGCAACAA GAGTGAAACT CCATCTCAAA 2360

AACAAAACAA AACAAAACAA AACCAACAAC TCAGAAGGAG 2400

GCATATGTGT TATAAAGTCT TTACTACAAC TTTGATTTTA 2440

TTAGTGGTTG GTTACTGACT CTGCCAAGAG TACAGAATGA 2480

AGGGCAGAGA GTAAGGACTG GAAAACTGGC AGGAAACACA 2520

CTGACAGCCG TCATCCCTGG AGGAAACTGC TCAATAAAAC 2560

GGCTCCATAT TTACTTCTCT GGTCACAGTT CATACTCCAC 2600

GATTTTAACA AAGGAGTCGA GGAAGCTAGA TACTGTAAGT 2640

GGAACGGTGT GTCTCTGGAG GTAAGCAGGC TTGCTGATTT 2680

CTTGTTTTAT AATTCTTTTT TAATTACAAT GTAACTACTA 2720

AGAGCTTCAG TTCCCACTGG AGTGGTGCAC ACATCTCATT 2760

ACTACTAAAA CCACAGGAAT GTTCCAGGGA AACAGACTAT 2800

CATCACTGAG CGAGGTGGAA TCCAGCCAAA ACCCCAGGCT 2840

AACATCCAGA TGCCTGCATA TCAGCTAAAA TCCTTTTAAA 2880

GGACTTGGAA TCTCCAGATA CTAGTTTTAA GTCTTTTCTG 2920

GGAACTGGGA GTTTGTACTG GAGGCCACTT AACTATTTCA 2960

AAAAATATTC ACCAAAATAG GTGTCTCTCT GACTGCAACG 3000

GTTTGAGTCC TCCTCAGCCC TCATATCCTA GGCTTCGGAC 3040

TGTTGGGAAA GTCTTATCTT CCTGACGAAA GCTCAGCAGC 3080

AACAGAACCT GTTATTTTTT TGTTGAGACA GGGTCTTACT 3120

CTGTCACCCA GGCTGGAGTG CAGTAGTGCG ATCTTGGCTC 3160

ACTGCAGCCT CAGCCTACCA GGCTCAGGTG ACCCTATCTC 3200

AGCTTCTCGA GTAGGTGGGA CTACAGGCAT GTGCCACCAT 3240

GCTCGGTGAA CTAAACAAAC TTTTTTGTAG TGATACGGTC 3280

TCACTATATT GCCCAGGCTG GTTTTGAACT CCTGGGCTCA 3320

AGTGATCCTC CCACCTCAGC GTCTCAAAGT ACTGGGATTA 3360

CAGGTGTGAG CCTCTACACT GGGCCTGCAG AACCTACACA 3400

GAATCCGCAC CTGGTCTGCA GAACCCACAC CCGACCCACA 3440

GAACCCACAC CCGACCCACA GAACCCACAT CTGGCAGCAG 3480

AACCTCTTAG TATTTTTTTT TTTTCTTTGA GATGGAGTCT 3520

GGCTCTGTCA CCCAGGCTGG AGTGCAGTGG CGCGATCTCG 3560

GCTCACTGCA AGCTCTTCCT CCCGGGTTCA CCCCATTCTC 3600

CTGCCTCAAC CTCCCGAGTA GCTGTGAATA CAGGCGTCCG 3640

CCACCACGCC CGACTAATTT TTTTGTATTT TTAGTAGAGA 3680

CGGGGTTTCA CCGTGTTAGC CAGGATGGTC TGGATCTCCT 3720

GACCTCGTGA TCTGCCTGCC TCGGCCTCCC AAAGTGCTGG 3760

GATTACAGGC TTGAGCCACC GCACCCGGCC TCTTATTTTT 3800

TTTTTTGAGA TGGAGTCTCA CACTGTCACC TGGGCTGGAG 3840

TGCAGTGGAG CGATCTCGGC TCACTGCAAC CTCCGCCTCC 3880

TGGGTTCAAG AGATTCTCCT GCCTCAGCCT CCCAAGTAGC 3920

TGGGATTACA GGTGCCCACC ACCACGCCTG GCTAGTTTTT 3960

TGTATTTTTA GTAAAGATGG GGTTTCACCA TGTTGGCCAG 4000

GCTGGTCTTG AACTCCTGAC ATCAGGTGAT CCGCCCACCT 4040

TAGCCTCCCA AAGTGCTGGG ATTACAGGCG TGAGCCACCA 4080

TACCTGGCCA GCAAAACCTC TTTAACTTGT GTTCCATGGG 4120

CTCCTTTTCT GTGGGTCAAA ATCCTCCTGG AACCCTACAA 4160

TGCAGGCCCT ACAGGGGTGG GTGGTAAGTC CAACAAACAG 4200

GATTTCATCT TCTGGAGCTC CTGGATTTCA TCGTCCCATG 4240

GGCCACAGTG CAGCGACAGA ACCTCCTCAG CTTTCTGTAT 4280

TGTGCTCAGG GCTTCGGGTA CTGCAAACCT GAGCCAAGGG 4320

AGGTAAGAGG AGTTAGTTCA CTGATTCGTG AGGCAAATGT 4360

TAATTGAGGG CCTACTCACA CACCGTGAAG AATGTAAGAT 4400

CATTTCTGTC ATCAAGGATC C 4421

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 7210 Base Pairs
(B) TYPE: Nucleic Acid
(C) STRANDEDNESS: Double
(D) TOPOLOGY: Unknown

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Human

(vii) IMMEDIATE SOURCE:
(A) LIBRARY: λDASH II

(ix) FEATURE:
(A) NAME/KEY: JT6A
(B) LOCATION:
(C) IDENTIFICATION METHOD:
(D) OTHER INFORMATION: 7.0 kb Not 1-Not fragment; Derived from human placental genomic DNA; also referred to as JT106

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

GATCTAGAGC GGCCGCAGGG TGGACTGTGC TGAGGAACCC 40

TGGGCCCAGC AGGGGTGGCA GCCCGCGCAG TGCCACGTTT 80

GGCCTCTGGC CGCTCGCCAG GCATCCTCCA CCCCGTGGTC 120

CCCTCTGACC TCGCCAGCCC TCCCCCGGGA CACCTCCACG 160

CCAGCCTGGC TCTGCTCCTG GCTTCTTCTT CTCTCTATGC 200

CTCAGGCAGC CGGCAACAGG GCGGCTCAGA ACAGCGCCAG 240

CCTCCTGGTT TGGGAGAAGA ACTGGCAATT AGGGAGTTTG 280

TGGAGCTTCT AATTACACAC CAGCCCCTCT GCCAGGAGCT 320

GGTGCCCGCC AGCCGGGGGC AGGCTGCCGG GAGTACCCAG 360

CTCCAGCTGG AGACAGTCAG TGCCTGAGGA TTTGGGGGAA 400

GCAGGTGGGG AAACCTTGGC ACAGGGCTGA CACCTTCCTC 440

TGTGCCAGAG CCCAGGAGCT GGGGCAGCGT GGGTGACCAT 480

GTGGGTGGGC ACGCTTCCCT GCTGGGGGTG CAGGGGGTCC 520

ACGTGGCAGC GGCCACCTGG AGCCCTAATG TGCAGCGGTT 560

AAGAGCAAGC CCCTGGAAGT CAGAGAGGCC TGGCATGGAG 600

TCTTGCTTCT TGCAAACGAG CCGTGTGGAG AGAGAGATAG 640

TAAATCAACA AAGGGAAATA CATGGTCTGT CCGAGGATGA 680

GCTGCCGGAG AGCAATGGTG AAAGTGAAGT GGGGGAGGGG 720

GCGGGGCTGG GAGGAAAAGC CTTGTGAGAA GGTGACACGA 760

GAGCACGGCC TTGAAGGGGA AGAAGGAGGG CACTATGGAG 800

GTCCCGGCGA AGCGTGGCCT GGCCGAGGAA CGGCATGTGC 840

AGAGGTCCTG CCGAGGAGCT CAAGACAAGT AGGGGACGGT 880

GGGGCTGGAG TGGAGAGAGT GAGTGGGAGG AGGAGTAGGA 920

GTCAGAGAGG AGCTCAGGAC AGATCCTTTA GGCTCTAGGG 960

ACACGATAAA CACAGTGTTT TTTGTCTTGT CAAGTGTGTC 1000

CTTTTTATTT TTTTGAAAGA GTCTCGCTCT GTAGCCCAGG 1040

CTGGAGTGCA GCGGTGCGAC CTCGGCTCAC TGCAACCTCT 1080

GCCTCCCGGG TCCAAGCAAT TCTCCTGCCT CAGCCTCCCG 1120

AGTAGCTGGG ATTACAGGCA CCCGCCACCA CGCACTGCTA 1160

ATTTTTGTAT TTTAGTAGAG ACCGGGTTTT GCCATGTTGG 1200

TCAGGCTGGT CTCGAACTCC TGACCTCAGG TGATCCGCCC 1240

GCCTCGGCCT CCCAGAGTGG TGTGAGCCAC TATGCCCTGC 1280

AGCACTTGTC AAGTCTTTCT CAGCGTTCCC CTCCTCTCCA 1320

CTGCAGCTCC CAGTGCCCCA GTCTGGGCCT CGTCTTCACT 1360

TCCTGGGATC CCTGACATTG CCTGCTAGGC TCTCCCTGTC 1400

TCTGGTCTGG CTGCCTTCAC TGTAACCTCC ACCCAGCAGG 1440

TACCTCTTCA GCACCTCCCA TGAACCCAGC AGAATACCAA 1480

GCCCTGGGGA TGCAGCAACG AACAGGTAGA CGCTGCACTC 1520

CAGCCTGGGC GACAGAGCAA GACTCCGCCT GAAGAAAAAA 1560

AAAAGGACCA GGCCGGGCGC GGTGGCTCAC GCCTGTAATC 1600

CCAGCACTTT GGGAGGCCGA GGTGGGTGGA TCATGAGGTC 1640

AGGAGTTCAA GACCAGCCTG GCCAAAATGG TGAAACCCCG 1680

TCTCTACTGA AAAATACAAA AATTAGCTGG GTGCAGTGGC 1720

GGGCGCCTGT AGTCTCAGCT ACTCAGGAGG CTGAGGCAGG 1760

ATAATTGCTT GACCCCAGGA GGCAGAGGTT GCAGTGAACC 1800

GAGATCACGC CACTGCACTC CAGCCTGGGC GACAGAGCAA 1840

GACTCTGCCT CAAAAAAAAG AATAAAAATA AAAAAAAGGA 1880

CCAGATACAG AAAACAGAAG GAGACGTACT ATGAAGGAAA 1920

TTGGAGAGCT TTTGGGATAC TGAGTAACTC AGGGTGGCCT 1960

TTCCCAGGGG ACATTTAGCT GAGAGATAGA CGGTATGAAG 2000

ACCTGACCGT TCAGAAACAG GGGAAGAGGC AGCAGCCCGG 2040

GCAAAGGCCT TTGGGGCAGG AAAGGGCTTG GATCACTGGA 2080

GAAGCAGAAA GATGGCCAGT GTGACCAGAG TGTGACAAAG 2120

TCAGAGAAAA CCAGGAAGAT GGAGCTGGAG ACACAGGCGG 2160

GGCCAGATCA CGAGGGTCCT CGCAGACCAG AGCAAGGGTT 2200

TGGATTTTAT TCCAAGTATG AAGGGAAGCT GCTGAAGTGT 2240

GTTTTCCTTT ACAATTTGTA GTTGAAATAT AATATGCAAA 2280

GTACACAAGT CTTAACTATA TGTAAGCTTA ATGAATGTTT 2320

CCATGAACCA AATACCGCTG TGCAACCATC ACCAGCTCAA 2360

GAGACGAACC CTTCTCCCTC CTCCTGACTG CCAGTAACAT 2400

AGTGGTTCAG CTCAAGAAAC AGAACTCTTC TGACTTCCCC 2440

TAACATAGCG GGTTTTCTTT TTTGTTTTGT TTTTTGTTGT 2480

TTTTTAAGAG ACAATGTCTT TATTATTTTT ATTTTTTTTT 2520

ATTTTTGAGA CGGAGTCTTG CTGTCGCCCA GGCTGGAGTG 2560

CAGTGGTGCG ATCTCGGCTC ACTGCAGGCT CTGCCCCCCG 2600

GGGTTCATGC CATTCTCCTG CCTCAGCCTC CCTAGCAGCT 2640

GGGACTACAG GTGCCCGCCA CCTCGCCCGG CTATTTTTTT 2680

GTATTTTTAG TGGAGACGGG GTTTCACCGT GTTAGCCAGG 2720

ATGGTCTCGA TCTCCTGACC TCGTGATCCG CCCACCTCGG 2760

CCTCCCAAAG TGCTGGGATT ACAGGCATGA GCCACCGCGC 2800

CCAGCCAAGA GACACGGTCT TGCTCTGTCG CCCAGGCTGG 2840

ATGGAGTGCC GTGGTGCGAT CACAGCTCGC GGCAGCCTTG 2880

ACATCCTGGG CTCAAGCAAC CTTCCTGCCT TGGCCTCCCA 2920

AATGTTGGGA TTATAGGCAT GAGCCACTGT GCTTGGCATC 2960

TATTCATCTT TAATGTCAAG CAGGCAATTG AATATTTGAT 3000

CAGGGATAGA ATTGTCTATT TGGGGGTATG CAGATGTGCT 3040

TCATGTCATG GAACTGGGCC GGGCGCGGTG GCTCATGCCT 3080

ATAATCCCAG CACTTTGGGA GGCCGAGGCA GGCGGATCAT 3120

AAGGTCAGGA GATCGAGACC ATCCGGGCCA ACACGGTGAA 3160

ACCCCGTCTC TACTAAAAAT ACAAAAATTA GGCAGGTGTG 3200

GTGGTGCGTG CCTGTAGTCC CAGCTACTCA GGGAGGCTGA 3240

GACAGGAGAA TTGATTGAAC CTGGGAGGCA GAGGTTGTAG 3280

TGAGCCAAGA TCGCGCCACT GCACTCCAGC CTGGGCGACA 3320

TGAGCGAGAC TCCGTCTCAA AAATAAACAA AAAAAAGTCA 3360

TGGAATTGAT GGAAATTGCC TAAGGGGAGA TGTAGAAGAA 3400

AAGGGGTCTC AGGATCAAGC CAGCAGAGAA GGCAGAAAAG 3440

GTAAGGTGTG TGAGGTGGCA GAAAAAGGGA AGAGTGTGGA 3480

CAGTGAGGGT TTCAAGGAGG AGGAACTGTC TACTGCCTCC 3520

TGCCAAGGAC GGAGGTGTCC ACTGCCAGTT GACATAAGGT 3560

CACCCATGAA CTTGGTGACA GGAATTTCAG TGGAGAAGTG 3600

GCCACAGACA CAAGTCTAGA ATTGAAATGG GAGCCGAGGC 3640

AGCGTAGACA AAAGAGGAAA CTGCTCCTTC CAGAGCGGCT 3680

CTGAGCGAGC ACCGAGAAAT GGGCAGTGGC TTTAGGGGAT 3720

GTAGCGTCAA GGAAGTGTCT TTTAAAGAAG TCGGGGGCCG 3760

GGCACGGTGG CTCACGCCTG TAGTCCCAGC ACTTTGGGAG 3800

GCCGAGGCAG GCAGATCACT TGAGGTCAGG AGTTCGAGAC 3840

CAGCCTGGCT AACACGATGA AACCCCGTCT CTACTAAAAA 3880

TACAAAAAAT TAGCTGGGCA CGGTGGCTCG TGCCTGTAAT 3920

CCCAGCACTT TGGGAGGCAG AGGTGGGCAG ATCACTTGAG 3960

GTCAGGAGTT TGAGACCAGC CTAGCCAACA TGGTGAAACC 4000

CCATCTCTAC TAAAACTACA AAAATTAGCC GGGAGTGGTG 4040

GCACGTGCCT GTAATCCCAG CCAGTCAGGA GGCTGAGGCA 4080

GGAGAATCAC TGGAATCCTG GAGGTGGAGG TGGCAGTGAG 4120

CCGAGATGGT ACCTCTGTAC TCCAGCCTGG GGGACAGAGT 4160

GAGACTCCGT CTCAAAAAAA AAAGAAGGTG GGGAAGGATC 4200

TTTGAGGGCC GGACACGCTG ACCCTGCAGG AGAGGACACA 4240

TTCTTCTAAC AGGGGTCGGA CAAAAGAGAA CTCTTCTGTA 4280

TAATTTATGA TTTTAAGATT TTTATTTATT ATTATTTTTT 4320

ATAGAGGCAA GCATTTTTCA CCACGTCACC CAGGCTGGTC 4360

TCCAACTCCT GGGCTCAAGT GTGCTGGGAT TATAGCCATG 4400

AGTCACCACA CCTGGCCCAG AAACTTTACT AAGGACTTAT 4440

TTAAATGATT TGCTTATTTG TGAATAGGTA TTTTGTTCAC 4480

GTGGTTCACA ACTCAAAAGC AACAAAAAGC ACCCAGTGAA 4520

AAGCCTTCCT CTCATTCTGA TTTCCAGTCA CTGGATTCTA 4560

CTCTTGGGAT GCAGTGTTTT TCATCTCTTT TTTGTATCCT 4600

TTTGGAAATA GTATTCTGCT TTAAAAAGCA AATACAGGCC 4640

AGGTATGGTG GCTCACTCCT GTAATCCCAG CACTTTGGGA 4680

GCCGAGGCAG GTGATCACCT AAGGTCAGGA GTTCAAGACC 4720

AGCCTGGCCA ATATGGTGAA ACCCTGTCTG TACCAAAACA 4760

CAAAAACAAA AACAAAAACA AAAATTAGCC GGGCGTGGTG 4800

GCGTGCTCCT GTAATCCCAG CTACTCAGGA GGCTGAGGCA 4840

GGAGAATCGC TTGAACCTGG GAGGCAGAGG TTGCAGTGAG 4880

CCGAGATTGT GCCACTGTAC TCCAGCCTGG GCCACAGAGC 4920

AAGGTTCCAT CTCAAACAAA ACAAAACAAA ACAAACAAAA 4960

AAACAAAACA AAAGCTAATA CAAACACATA TACAATAGAC 5000

AAAACTGTAA ATATTTTATT ATTTTTATTT TTTTTAGTAG 5040

AGACAGGGTT TCACCATGTT GGCCAGGATG GTCTCAAACT 5080

CCTGACCTCA GGTGATCCAC CCACCTCAGC CTCCCGATAG 5120

TTAGGATTAC AGGCATGAGC CACCACACCC GGCCTAAAAT 5160

TGTAAACGTT TTAGAAGAAA GTATAGATGA ATCCCTTCGT 5200

GATCTCGGGG AAGAAGAGAT TTTTTAAAAA AGATACCAAA 5240

AGAAGCACAA ATTATAAAAG AAAAGATTGA AAATGTTGGT 5280

GTTAAAATTA AAAACTTGTT TTAAAACAAG CTTGTGTAAC 5320

CCATGACCCA CAGGCTGCAT GTGGCCCAGA AAAGCTTTGA 5360

CTGCAGCCCA ACACAAATTC GTAAACTTTC CTAAAACATT 5400

ATGAGATTTT TTTTGAGATT TTGTTTTGTT TTGTTTTTTG 5440

TTTTTTTAGC TCATTCGGTA TCATTAATGT TAGCATATTT 5480

TACGTGGGGC CCAAGACAAT TCTTCTTCCA ATGTGTCTCA 5520

GGGGAGCCAA AAGATTGGAC ACCCCTGCCA TAAACATGAA 5560

AAGACAATGG CCGGGCACGG TGGCTCACGC CTGTAATCCC 5600

AGCACTTTGG GAGGCTGAGG GGGGCGGGAT CACCTGAGGT 5640

CAGGAGTTTG AGACAAGCGT GACCAATGTG GTGAAACCCT 5680

GTCTCTACTA AAAATACAAA AATTAGCCGG GCATGCTCGT 5720

GCACACCTAT AGTCCCAACT ACTCAGCAGG GTGAGGCAGG 5760

AGAACCTCTT GAACCCGGGA AGCGGAGGTT GCAGTGAGCC 5800

GACATTGCAC CCCTGCACTC CAGCCTGGGT GACAGAGTGA 5840

GTCTCCACTG GAAAAAAAAA AAAAAGAACA GTGTGATACA 5880

TTGACCTAAG GTTTAAGAAC ATGCAAACTG ATACTATATA 5920

TCACTTAGGG ACAAAAACTT ACATGGTAAA AGTAAAAAGA 5960

AATGTACGAA AATAATAAAA ATCAAATTCA AGATGGTGGT 6000

TATGGTGACG GGAAAGAACT GAGGCGGAAA TATAAGGTTG 6040

TCACTATATT GAGAAATTTT TCTATCTTTT TTTCTTTTTT 6080

CTTTTTTTGA GACGGGGTCT CGCTCTGTCG CCCAGGATGG 6120

AGTGCAGTGG TGTGATCTCA GCTCACTGCA ACCTCCGCCT 6160

CCCAGGTTTA AGTGATTCTC CTGCCTCAGA CTCCCAAGTA 6200

GCTGGGACTA CAGGTGCGCG CCAACACACC TGGGTAATTT 6240

TGTTTGTATT TTTAGTAGAG ATGGGGTTTC ACCGTGTTGA 6280

CTAGGCTGGT CTCGAACTCC TGACCTCAGG TGATCCCCCG 6320

GCCTCGGTCT CCCAAAGTGC TGGGATAACA AGCGTGAGCC 6360

ACTGCGCCCA GCTTTGTTTG CATTTTTAGG TGAGATGGGG 6400

TTTCACCACG TTGGCCAGGC TGGTCTTGAA CTCCTGACCT 6440

CAGGTGATGC ACCTGCCTCA GTCTCCCAAA GTGCTGGATT 6480

ACAGGCGTTA GCCCCTGCGC CCGGCCCCTG AAGGAAAATC 6520

TAAAGGAAGA GGAAGGTGTG CAAATGTGTG CGCCTTAGGC 6560

GTAATGGATG GTGGTGCAGC AGTGGGTTAA AGTTAACACG 6600

AGACAGTGAT GCAATCACAG AATCCAAATT GAGTGCAGGT 6640

CGCTTTAAGA AAGGAGTAGC TGTAATCTGA AGCCTGCTGG 6680

ACGCTGGATT AGAAGGCAGC AAAAAAAGCT CTGTGCTGGC 6720

TGGAGCCCCC TCAGTGTGCA GGCTTAGAGG GACTAGGCTG 6760

GGTGTGGAGC TGCAGCGTAT CCACAGGTAA AGCAGCTCCC 6800

CTGGCTGCTC TGATGCCAGG GACGGCGGGA GAGGCTCCCC 6840

TGGGCTGGGG GGACAGGGGA GAGGCAGGGG CACTCCAGGG 6880

AGCAGAAAAG AGGGGTGCAA GGGAGAGGAA ATGCGGAGAC 6920

AGCAGCCCCT GCAATTTGGG CAAAAGGGTG AGTGGATGAG 6960

AGAGGGCAGA GGGAGCTGGG GGGACAAGGC CGAAGGCCAG 7000

GACCCAGTGA TCCCCAAATC CCACTGCACC GACGGAAGAG 7040

GCTGGAAAGG CTTTTGAATG AAGTGAGTGG GAAACAGCGG 7080

AGGGGCGGTC ATGGGGAGGA AAGGGGAGCT AAGCTGCTGG 7120

GTCGGGTCTG AGCAGCACCC CAAGACTGGA GCCCGAGGCA 7160

AGGAGGCTCA CGGGAGCTGC TTCCACCAAG GGCAGTCAGG 7200

AAGGCGGCCG 7210

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1988 Base Pairs
(B) TYPE: Nucleic Acid
(C) STRANDEDNESS: Double
(D) TOPOLOGY: Unknown

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Human

(ix) FEATURE:
(A) NAME/KEY: JT8A
(B) LOCATION:
(C) IDENTIFICATION METHOD:
(D) OTHER INFORMATION: 2 kb PCR product using primers, SEQ ID: 13 and 14; Also referred to as JT108

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

ACAAGCTGGC AGCGGCTGTC TCCAACTTCG GCTATGACCT 40

GTACCGGGTG CGATCCAGCA NGAGCCCCAC GACCAACGTG 80

CTCCTGTCTC CTCTCAGTGT GGCCACGGCC CTCTCGGCCC 120

TCTCGCTGGG TGAGTGCTCA GATGCAGGAA GCCCCAGGCA 160

GACCTGGAGA GGCCCCCTGT GGCCTCTGCG TAAACGTGGC 200

TGAGTTTATT GACATTTCAG TTCAGCGAGG GGTGAAGTAG 240

CACCAGGGGC CTGGCCTGGG GGTCCCAGCT GTGTAAGCAG 280

GAGCTCAGGG GCTGCACACA CACGATTCCC CAGCTCCCCG 320

AAAGGGGCTG GGCACCACTG ACATGGCGCT TGGCCTCAGG 360

GTTCGCTTAT TGACACAGTG ACTTCAAGGC ACATTCTTGC 400

ATTCCTTAAC CAAGCTGGTG CTAGCCTAGG TTCCTGGGAT 440

GTAACTGCAA ACAAGCAGGT GTGGGCTTGC CCTCACCGAG 480

GACACAGCTG GGTTCACAGG GGAACTAATA CCAGCTCACT 520

ACAGAATAGT CTTTTTTTTT TNTTTTTTTN NNCTTTCTGA 560

GACGGAGTCT CGCTTTGTCN CCAAGGCTGG AGTGCAGTGG 600

TGTGATCTCA GCTCACTGCA ACCTCTGCCT CCCTGGTTCA 640

AGGAATTCTC CTGCCTCAGC CTCCAGAGTA GCTGGGATTA 680

CAGGCACCTG CCATCATGCC CAGCTAATTT TTGTATTTTT 720

AGTAGAGACG GGGTTTCACC ATGTTGCCTA GGCTGGTCTC 760

AAACTCCCGG GCTCAAGCGA TCCACCCGCC TTGGCCTCCC 800

AAAGTGCTGG GATTACAGGC GTGAGCCACC GCGCCTGGCC 840

AGAATAATCT TAAGGGCTAT GATGGGAGAA GTACAGGGAC 880

TGGTACCTCT CACTCCCTCA CTCCCACCTT CCAGGCCTGA 920

TGCCTTTAAC CTACTTCAGG AAAATCTCTA AGGATGAANA 960

TTCCTTGGCC ACCTAGATTG TCTTGAAGAT CAGCCTACTT 1000

GGGCTCTCAG CAGACAAAAA AGATGAGTAT AGTGTCTGTG 1040

TTCTGGGAGG GGGCTTGATT TGGGGCCCTG GTGTGCAGTT 1080

ATCAACGTCC ACATCCTTGT CTCTGGCAGG AGCGGAGCAG 1120

CGAACAGAAT CCATCATTCA CCGGGCTCTC TACTATGACT 1160

TGATCAGCAG CCCAGACATC CATGGTACCT ATAAGGAGCT 1200

CCTTGACACG GTCACTGCCC CCCAGAAGAA CCTCAAGAGT 1240

GCCTCCCGGA TCGTCTTTGA GAAGAGTGAG TCGCCTTTGC 1280

AGCCCAAGTT GCCTGAGGCA TGNGGGNTCC ATGCTGCAGG 1320

CTGGGGGGGT CTTTTTTTTT TTTTTNNNNA GACGGAGTCT 1360

CGCTCTGTTG CCCAGGCTGG AGTGCAGTGG CGNGATCTCG 1400

GCTCACTGCA ACCTCCACCT CCCGGGTTCA CACCATCCTC 1440

CTGCCTCAGC CTCCCGAGTA GCTGGGACTG CAGGNGCCCA 1480

GCTAATCTTT NTTGTATTTT TAGCAGAGAC GGGGTTTCAC 1520

CGTGTTTGCC AGGATAGTCT CGATCTCCTG ACCTGGTGTT 1560

CTGCCCGCCT CGACCTCCCA AAGTGCTGGG ATTACAGGTG 1600

TGAGCCACCG CGCTCGGCCC GTTTCTAAAC AATAGATCAT 1640

GTGTGCCCAG GCCTGGCCTG GCACTGGTGT GGAGGAAGGG 1680

CCCGTGAGCC CAAAGAGGCT CAGAAAGAGG AAGTGGGCTG 1720

CAGGAGACGG TGGGAGGGGC NGGGAGGGCA GTGGCGCGAT 1760

GTGGGGAAAT CTGCTGCCCC CCTGGCCAGT GCCTGGGGAT 1800

GCCAGCAGAA GTCCTGGCAA GTCACAGGAA GATGCTGGCT 1840

GGGAAGTCAG GGCCTGCTGA GCGCTAAACC AGAACCCGAG 1880

CCTGGCAGGC TCTCAAAGAC GGGATGCTTG TCGTNGAGTC 1920

TCATANGCTA ACCTCTGCTC CGCCTCTTCT CAGAGCTGCG 1960

CATAAAATCC AGCTTTGTGG CACCTCTG 1988

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 3267 Base Pairs
(B) TYPE: Nucleic Acid
(C) STRANDEDNESS: Double
(D) TOPOLOGY: Unknown

(ii) MOLECULE TYPE: Genomic DNA

(ix) FEATURE:
(A) NAME/KEY: JT109
(B) LOCATION:
(C) IDENTIFICATION METHOD:
(D) OTHER INFORMATION: 3.3 kb PCR product using primers, SEQ ID No: 15 and 16

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

GATTCCAGCT TTGTGGCACC TCTGGAAAAG TCATATGGGA 40

CCAGGCCCAG AGTCCTGACG GGCAACCCTC GCTTGGACCT 80

GCAAGAGATC AACAACTGGG TGCAGGCGCA GATGAAAGGG 120

AAGCTCGCCA GGTCCACAAA GGAAATTCCC GATGAGATCA 160

GCATTCTCCT TCTCGGTGTG GCGCACTTCA AGGGTGAGCG 200

CGTCTCCAAT TCTTTTTCAT TTATTTTACT GTATTTTAAC 240

TAATTAATTA ATTCGATGGA GTCTTACTCT GTAGCCCTAA 280

CTGGAGTGCA GTGGTGCGAT CTCAGCTCAA TGCAACCTCC 320

GCCTCCCAGG TTCAAGCAAT TCTTGTGCCT CAGCCTCCCG 360

AGTAGCTGGG ATTACAGGGA TGTACCACCA CTCCCGGCTA 400

ATTTTTTGTA TTTAATAGAC ATGGGGTTTC ACCATGTTGG 440

CCAGGCTGGT CTCGAACTCC TGAGCTCAGG TGGTCTGCCC 480

GCCTCAGCCT CCCAAAGTGC TAGGATTACA AGCTTGAGCC 520

ACCACGCCCA GCCCTTTTTA TTTTTAAATT AAGAGACAAG 560

GTGTTGCCAT GATGCCCAGG CTGGTCTCGA ACTCCTGGGC 600

TCAAGTAATC CTCCCACCTT GGCCTCCCAA AGTGCTGGGA 640

TTACAGGCAT GAGCCACCGC GCCCGGCCCT TTTACATTTA 680

TTTATTTATT TTTTGAGACA GAGTCTTGCT CTGTCACCCA 720

GGCTGGAGTG CAGTGGCGCG ATCTCGGCTC ACTGCAAGCT 760

CTGCCTTCCA GGTTCACACC ATTCTCCTGC CTCGACCTCC 800

CGAGTAGCTG GGACTACAGG CGCCCGCCAC TGCGCCCTAC 840

TAATTTTTTG TATTTTTAGT AGAGACGGGG TTTCACCGTG 880

GTCTCGATCT CCTGACCTCG TGATCCACCC GCCTCAGCCT 920

CCCAAAGTGC TGGGATTACA GGCGTGAGCC ACTGCGCCCG 960

GCCCTTTTAC ATTTATTTTT AAATTAAGAG ACAGGGTGTC 1000

ACTATGATGC CGAGGCTGGT CTCGAACTCC TGAGCTGAAG 1040

TGATCCTCCC ACCTCGGCCT CCCAAAATGC TGGGATTACC 1080

ATGTCCAACT TTCCACTTCT TGTTTGACCA AGGATGGATG 1120

GCAGACATCA GAAGGGGCTT GGAAAGGGAG GTGTCAAAGA 1160

CCTTGCCCAG CATGGAGTCT GGGTCACAGC TGGGGGAGGA 1200

TCTGGGAACT GTGCTTGCCT GAAGCTTACC TGCTTGTCAT 1240

CAAATCCAAG GCAAGGCGTG AATGTCTATA GAGTGAGAGA 1280

CTTGTGGAGA CAGAAGAGCA GAGAGGGAGG AAGAATGAAC 1320

CTGGGTCTGT TTGGGGCTTT CCCAGCTTTT GAGTCAGACA 1360

AGATTTATTT ATTTATTTAA GATGGAGTCT CATTCTGTTG 1400

CCCAGGCTGG AGTGCAGTGG TGCCATCTTG GCTCACTACA 1440

GCCTCCCCAC CTCCCAGGTT CAAGTGCTTC TCCTGCCTCA 1480

GCCTCCCGAG TAGTTGGGAT TACAGGCGCC CGCCACCACA 1520

CCCAGCTAAT TTTTGTATTT TCAGTAGAGA TGGGGTTTCG 1560

CCATGCTGGC CAGGCTGTTC TCGAAAACTC CTGACCTCAG 1600

ATGATCCACC CGCCTCGGCC TCCCACAGTG CTGGGATTAC 1640

AGGCGTGAGC CACTGCGCTG GCCAAATCAG ACAAGGTTTA 1680

AATCCCAGCT CTGCCTGTAC TAGCTGAGGA ACTCTGCACA 1720

CATTTCATAA CCTTTCTGGG CCTACGTTCT CACCTTTAAC 1760

GTGAGGATAA TATATCTACT TCATAGACAC CTTTTTATGT 1800

TGTCTCCAAG TTTTCTAACA GCTCTAGTTC TGTACCCAAG 1840

ACATGGCAGG TGGCCAACGA CATCCTTCTA GGCTGTGGTG 1880

ATGTGTTTGG AGCTTGTTCC ACGGGTCTTG TGTGGGGCCA 1920

GCCCTGTTCA GATAAGGCCT TGTGGGGTGG CCTGGGGTAG 1960

GGGGAGGGGT TGGGCAAACT CTCCCTTAAA ACGCTTTGTA 2000

ACCATCTGAG GCACCAGCAA GAGCGGCCCC CGAGCCTGGA 2040

CAAAATCCAA ACGGCTTCCT ACTTCAAGCA CTGATGTCTA 2080

GTGAGTGAAG GAACAGCTCT GGGTCCAGGA TATTATAGGT 2120

CACATTAAAC TAAAGGGGCT TGGCCATCAG CTGGCTTCCA 2160

GAGCGTCAGC CAGTTACTTC ACCTCTTTGG CTTTGGCCTG 2200

TTTTCAGCTA CAAGAGGACT TAATCCAGAG GACCTCAGAG 2240

GTCCTTCCCA GCTCAGACCT TCTTTGACTG TCTCCCAGAG 2280

ACACTGCTGT AGGAGTGCAC ACCAGTTTAC TTTTCTTTCT 2320

TTTGTTTTTG AGATGGAGTT TCGCTCTTTT TGCCTAGGCT 2360

GGAGTGCTGT GGTGTGATCT CAGCTCACTG CAACCTCTGG 2400

CTCCCAGGTT CAAGTGATTC TCCTGTCTCT GCCTCCCGAG 2440

TAGCTGGGAT TACAGACACC CACCACTGCA CCCGGCTAGT 2480

TTTTGTATTT TCAGTAGAGA TGGGGTTTCG CCATGCTGGC 2520

CAGGCTGTTC TCGAAAACTC CTGACCTCAG ATGATCCATC 2560

CGCCTTGGCC TCCCAAAGTG CTGAGATTAC AGATGTGAGG 2600

CACCACACCC GGCCATTTTT GTATTTTTAG TAGAGACGGG 2640

GTTTTGCCAT GTTGGCCACG CTGGTCTCAA ACTCCTGACC 2680

TCAAGTGATC TGCCCACCTT GGCCTCCTGA AGGGCTGGGA 2720

CTACAGGCGT GAGTCACCGT GCCCGGCCAT TTTTGTATTT 2760

TTAGGACAGC GTTTTTTCAT GTTGGCCAGG CTGGTCTCAA 2800

ACTCCTGACC TCAAGTGATC CACCCACCCC GGCCTCCCAA 2840

TATGCTGGGA TTCCAGGTGT GAGTTACCAT GCCCGGCTAC 2880

CACTTTACTT TTCCTGCAGG CTATCACAGA ACGTGTACAA 2920

TCTAGACTCT AATCAACCAA ATCAACGTCT TGCCATCGGA 2960

GTTTGCTGGT GAAGGGCACT TGGGGTCCTG GAAATAACTG 3000

TAGGCTCCAA GCCACACACA CTGAGATAGG CCTATTCCCT 3040

GAGGCCTCAG AGCCCCTGAC AGCTAAGCTC CCTTGAGTCG 3080

GGCAATTTTC AACAACGTGC TCTGGGGACA CAGCATGGCG 3120

CCACTGTCTT TCTGGTCTCC TGGGGCTCAG ACTATGTCAT 3160

ACACTTCTTT CCAGGGCAGT GGGTAACAAA GTTTGACTCC 3200

AGAAAGACTT CCCTCGAGGA TTTCTACTTG GATGAAGAGA 3240

GGACCGTGAG GGTCCCCATG ATGAATC 3267

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 Base Pairs
(B) TYPE: Nucleic Acid
(C) STRANDEDNESS: Unkown
(D) TOPOLOGY: Unknown

(ii) MOLECULE TYPE: Oligonucleotide

(ix) FEATURE:
(A) NAME/KEY: 603
(B) LOCATION:
(C) IDENTIFICATION METHOD:
(D) OTHER INFORMATION: primer in a polymerase chain reaction

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

ACAAGCTGGC AGCGGCTGTC 20

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 Base Pairs
(B) TYPE: Nucleic Acid
(C) STRANDEDNESS: Unkown
(D) TOPOLOGY: Unknown

(ii) MOLECULE TYPE: Oligonucleotides

(ix) FEATURE:
(A) NAME/KEY: 604
(B) LOCATION:
(C) IDENTIFICATION METHOD:
(D) OTHER INFORMATION: primer in a polymerase chain reaction

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

CAGAGGTGCC ACAAAGCTGG 20

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 Base Pairs
(B) TYPE: Nucleic Acid
(C) STRANDEDNESS: Unkown
(D) TOPOLOGY: Unknown

(ii) MOLECULE TYPE: Oligonucleotides

(ix) FEATURE:
(A) NAME/KEY: 605
(B) LOCATION:
(C) IDENTIFICATION METHOD:
(D) OTHER INFORMATION: primer in a polymerase chain reaction

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

CCAGCTTTGT GGCACCTCTG 20

(2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 Base Pairs
(B) TYPE: Nucleic Acid
(C) STRANDEDNESS: Unknown
(D) TOPOLOGY: Unknown

(ii) MOLECULE TYPE: Oligonucleotide

(ix) FEATURE:
(A) NAME/KEY: 606
(B) LOCATION:
(C) IDENTIFICATION METHOD:
(D) OTHER INFORMATION: primer in a polymerase chain reaction

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

CATCATGGGG ACCCTCACGG 20

(2) INFORMATION FOR SEQ ID NO:17:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 Base Pairs
(B) TYPE: Nucleic Acid
(C) STRANDEDNESS: Unknown
(D) TOPOLOGY: Unknown

(ii) MOLECULE TYPE: Oligonucleotide

(ix) FEATURE:
(A) NAME/KEY: 2213
(B) LOCATION:
(C) IDENTIFICATION METHOD:
(D) OTHER INFORMATION: primer in a polymerase chain reaction

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

AGGATGCAGG CCCTGGTGCT 20

(2) INFORMATION FOR SEQ ID NO:18:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 Base Pairs
(B) TYPE: Nucleic Acid
(C) STRANDEDNESS: Unknown
(D) TOPOLOGY: Unknown

(ii) MOLECULE TYPE: Oligonucleotide

(ix) FEATURE:
(A) NAME/KEY: 2744
(B) LOCATION:
(C) IDENTIFICATION METHOD:
(D) OTHER INFORMATION: primer in a polymerase chain reaction

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

CCTCCTCCAC CAGCGCCCCT 20

(2) INFORMATION FOR SEQ ID NO:19:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 24 Base Pairs
(B) TYPE: Nucleic Acid
(C) STRANDEDNESS: Uknown
(D) TOPOLOGY: Unknown

(ii) MOLECULE TYPE: Oligonucleotide

(ix) FEATURE:
(A) NAME/KEY: 2238
(B) LOCATION:
(C) IDENTIFICATION METHOD:
(D) OTHER INFORMATION: primer in a polymerase chain reaction

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

ATGATGTCGG ACCCTAAGGC TGTT 24

(2) INFORMATION FOR SEQ ID NO:20:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 Base Pairs
(B) TYPE: Nucleic Acid
(C) STRANDEDNESS: Unknown
(D) TOPOLOGY: Unknown

(ii) MOLECULE TYPE: Oligonucleotide

(ix) FEATURE:
(A) NAME/KEY: 354
(B) LOCATION:
(C) IDENTIFICATION METHOD:
(D) OTHER INFORMATION: primer in a polymerase chain reaction

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

TGGGGACAGT GAGGACCGCC 20

(2) INFORMATION FOR SEQ ID NO:21:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 24 Base Pairs
(B) TYPE: Nucleic Acid
(C) STRANDEDNESS: Unknown
(D) TOPOLOGY: Unknown

(ii) MOLECULE TYPE: Oligonucleotide

(ix) FEATURE:
(A) NAME/KEY: JT10 - UP01
(B) LOCATION:
(C) IDENTIFICATION METHOD:
(D) OTHER INFORMATION: primer in a polymerase chain reaction

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

GGTGTGCAAA TGTGTGCGCC TTAG 24

(2) INFORMATION FOR SEQ ID NO:22:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 24 Base Pairs
(B) TYPE: Nucleic Acid
(C) STRANDEDNESS: Unkown
(D) TOPOLOGY: Unknown

(ii) MOLECULE TYPE: Oligonucleotide

(ix) FEATURE:
(A) NAME/KEY: JT10 - DP01
(B) LOCATION:
(C) IDENTIFICATION METHOD:

(D) OTHER INFORMATION: primer in a polymerase chain reaction

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

GGGAGCTGCT TTACCTGTGG ATAC 24

(2) INFORMATION FOR SEQ ID NO:23:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 25 Base Pairs
(B) TYPE: Nucleic Acid
(C) STRANDEDNESS: Unknown
(D) TOPOLOGY: Unknown

(ii) MOLECULE TYPE: Oligonucleotide

(ix) FEATURE:
(A) NAME/KEY: 1590
(B) LOCATION:
(C) IDENTIFICATION METHOD:
(D) OTHER INFORMATION: primer in a polymerase chain reaction

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

GGACGCTGGA TTAGAAGGCA GCAAA 25

(2) INFORMATION FOR SEQ ID NO:24:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 19 Base Pairs
(B) TYPE: Nucleic Acid
(C) STRANDEDNESS: Unknown
(D) TOPOLOGY: Unknown

(ii) MOLECULE TYPE: Oligonucleotide

(ix) FEATURE:
(A) NAME/KEY: 1591
(B) LOCATION:
(C) IDENTIFICATION METHOD:
(D) OTHER INFORMATION: primer in a polymerase chain reaction

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

CCACACCCAG CCTAGTCCC 19

(2) INFORMATION FOR SEQ ID NO:25:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 18 Base Pairs
(B) TYPE: Nucleic Acid
(C) STRANDEDNESS: Double
(D) TOPOLOGY: Unknown

(ii) MOLECULE TYPE: Genomic DNA

(ix) FEATURE:
(A) NAME/KEY: 5' splice site of EXON 1
(B) LOCATION:
(C) IDENTIFICATION METHOD:
(D) OTHER INFORMATION: 5' Splice Donor site is located between nucleotides 9 and 10

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

TATCCACAGG TAAAGTAG 18

(2) INFORMATION FOR SEQ ID NO:26:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 18 Base Pairs
(B) TYPE: Nucleic Acid
(C) STRANDEDNESS: Double
(D) TOPOLOGY: Unknown

(ii) MOLECULE TYPE: Genomic DNA
(ix) FEATURE:
(A) NAME/KEY: 5' splice site of EXON 2
(B) LOCATION:
(C) IDENTIFICATION METHOD:
(D) OTHER INFORMATION: 5' Splice Donor site is located between nucleotides 9 and 10

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

CCGGAGGAGG TCAGTAGG 18

(2) INFORMATION FOR SEQ ID NO:27:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 18 Base Pairs
(B) TYPE: Nucleic Acid
(C) STRANDEDNESS: Double
(D) TOPOLOGY: Unknown

(ii) MOLECULE TYPE: Genomic DNA

(ix) FEATURE:
(A) NAME/KEY: 5' splice site of EXON 3
(B) LOCATION:
(C) IDENTIFICATION METHOD:
(D) OTHER INFORMATION: 5' Splice Donor site is located between nucleotides 9 and 10

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

TCTCGCTGGG TGAGTGCT 18

(2) INFORMATION FOR SEQ ID NO:28:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 18 Base Pairs
(B) TYPE: Nucleic Acid
(C) STRANDEDNESS: Double
(D) TOPOLOGY: Unknown

(ii) MOLECULE TYPE: Genomic DNA

(ix) FEATURE:
(A) NAME/KEY: 5' splice site of EXON 4
(B) LOCATION:
(C) IDENTIFICATION METHOD:
(D) OTHER INFORMATION: 5' Splice Donor site is located between nucleotides 9 and 10

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:

TTGAGAAGAG TGAGTCGC 18

(2) INFORMATION FOR SEQ ID NO:29:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 18 Base Pairs
(B) TYPE: Nucleic Acid
(C) STRANDEDNESS: Double
(D) TOPOLOGY: Unknown

(ii) MOLECULE TYPE: Genomic DNA

(ix) FEATURE:
(A) NAME/KEY: 5' splice site of EXON 5
(B) LOCATION:
(C) IDENTIFICATION METHOD:
(D) OTHER INFORMATION: 5' Splice Donor site is located between nucleotides 9 and 10

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:

ACTTCAAGGG TGAGCGCG 18

(2) INFORMATION FOR SEQ ID NO:30:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 18 Base Pairs
(B) TYPE: Nucleic Acid
(C) STRANDEDNESS: Double
(D) TOPOLOGY: Unknown

(ii) MOLECULE TYPE: Genomic DNA

(ix) FEATURE:
(A) NAME/KEY: 5' splice site of EXON 6
(B) LOCATION:
(C) IDENTIFICATION METHOD:
(D) OTHER INFORMATION: 5' Splice Donor site is located between nucleotides 9 and 10

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:

AGCTGCAAGG TCTGTGGG 18

(2) INFORMATION FOR SEQ ID NO:31:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 18 Base Pairs
(B) TYPE: Nucleic Acid
(C) STRANDEDNESS: Double
(D) TOPOLOGY: Unknown

(ii) MOLECULE TYPE: Genomic DNA

(ix) FEATURE:
(A) NAME/KEY: 5' splice site of EXON 7
(B) LOCATION:
(C) IDENTIFICATION METHOD:
(D) OTHER INFORMATION: 5' Splice Donor site is located between nucleotides 9 and 10

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:31:

AGGAGATGAG TATGTCTG 18

(2) INFORMATION FOR SEQ ID NO:32:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 18 Base Pairs
(B) TYPE: Nucleic Acid
(C) STRANDEDNESS: Double
(D) TOPOLOGY: Unknown

(ii) MOLECULE TYPE: Genomic DNA

(ix) FEATURE:
(A) NAME/KEY: 5' splice site of EXON 8
(B) LOCATION:
(C) IDENTIFICATION METHOD:
(D) OTHER INFORMATION: 5' Splice Donor site is located between nucleotides 9 and 10

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:32:

TTTATCCCTA ACTTCTGT 18

(2) INFORMATION FOR SEQ ID NO:33:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 9 Base Pairs
(B) TYPE: Nucleic Acid
(C) STRANDEDNESS: Double
(D) TOPOLOGY: Unknown

(ii) MOLECULE TYPE: Genomic DNA

(ix) FEATURE:
(A) NAME/KEY: 3' splice site of INTRON 1
(B) LOCATION:
(C) IDENTIFICATION METHOD:
(D) OTHER INFORMATION: 3' Splice Acceptor site is located between nucleotides 9 and 10

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:33:

GGACGCTGG 9

(2) INFORMATION FOR SEQ ID NO:34:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 18 Base Pairs
(B) TYPE: Nucleic Acid
(C) STRANDEDNESS: Double
(D) TOPOLOGY: Unknown

(ii) MOLECULE TYPE: Genomic DNA

(ix) FEATURE:
(A) NAME/KEY: 3' splice site of INTRON 2
(B) LOCATION:
(C) IDENTIFICATION METHOD:
(D) OTHER INFORMATION: 3' Splice Acceptor site is located between nucleotides 9 and 10

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:34:

TTCTTGCAGG CCCCAGGA 18

(2) INFORMATION FOR SEQ ID NO:35:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 18 Base Pairs
(B) TYPE: Nucleic Acid
(C) STRANDEDNESS: Double
(D) TOPOLOGY: Unknown

(ii) MOLECULE TYPE: Genomic DNA

(ix) FEATURE:
(A) NAME/KEY: 3' splice site of INTRON 3
(B) LOCATION:
(C) IDENTIFICATION METHOD:
(D) OTHER INFORMATION: 3' Splice Acceptor site is located between nucleotides 9 and 10

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:35:

TCCTGCCAGG GCTCCCCA 18

(2) INFORMATION FOR SEQ ID NO:36:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 18 Base Pairs
(B) TYPE: Nucleic Acid
(C) STRANDEDNESS: Double
(D) TOPOLOGY: Unknown

(ii) MOLECULE TYPE: Genomic DNA

(ix) FEATURE:
(A) NAME/KEY: 3' splice site of INTRON 4
(B) LOCATION:
(C) IDENTIFICATION METHOD:
(D) OTHER INFORMATION: 3' Splice Acceptor site is located between nucleotides 9 and 10

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:36:

CTCTGGCAGG AGCGGACG 18

(2) INFORMATION FOR SEQ ID NO:37:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 18 Base Pairs
(B) TYPE: Nucleic Acid
(C) STRANDEDNESS: Double
(D) TOPOLOGY: Unknown

(ii) MOLECULE TYPE: Genomic DNA

(ix) FEATURE:
(A) NAME/KEY: 3' splice site of INTRON 5
(B) LOCATION:
(C) IDENTIFICATION METHOD:
(D) OTHER INFORMATION: 3' Splice Acceptor site is located between nucleotides 9 and 10

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:37:

TCTTCTCAGA GCTGCGCA 18

(2) INFORMATION FOR SEQ ID NO:38:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 18 Base Pairs
(B) TYPE: Nucleic Acid
(C) STRANDEDNESS: Double
(D) TOPOLOGY: Unknown

(ii) MOLECULE TYPE: Genomic DNA

(ix) FEATURE:
(A) NAME/KEY: 3' splice site of INTRON 6
(B) LOCATION:
(C) IDENTIFICATION METHOD:
(D) OTHER INFORMATION: 3' Splice Acceptor site is located between nucleotides 9 and 10

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:38:

TCTTTCCAGG GCAGTGGG 18

(2) INFORMATION FOR SEQ ID NO:39:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 18 Base Pairs
(B) TYPE: Nucleic Acid
(C) STRANDEDNESS: Double
(D) TOPOLOGY: Unknown

(ii) MOLECULE TYPE: Genomic DNA
(ix) FEATURE:
(A) NAME/KEY: 3' splice site of INTRON 7
(B) LOCATION:
(C) IDENTIFICATION METHOD:
(D) OTHER INFORMATION: 3' Splice Acceptor site is located between nucleotides 9 and 10

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:39:

TTGTCTCAGA TTGCCCAG 18

(2) INFORMATION FOR SEQ ID NO:40:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 18 Base Pairs
(B) TYPE: Nucleic Acid
(C) STRANDEDNESS: Double
(D) TOPOLOGY: Unknown

(ii) MOLECULE TYPE: Genomic DNA

(ix) FEATURE:
(A) NAME/KEY: 3' splice site of INTRON 8
(B) LOCATION:
(C) IDENTIFICATION METHOD:
(D) OTHER INFORMATION: 3' Splice Acceptor site is located between nucleotides 9 and 10

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:40:

TCTCTACAGA GCTGCAAT 18

(2) INFORMATION FOR SEQ ID NO:41:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 737 Base Pairs
(B) TYPE: Nucleic Acid
(C) STRANDEDNESS: Double
(D) TOPOLOGY: Unknown

(ii) MOLECULE TYPE: Genomic DNA

(ix) FEATURE:
(A) NAME/KEY: PEDF Promoter
(B) LOCATION:
(C) IDENTIFICATION METHOD:
(D) OTHER INFORMATION: EXON begins at 614 and ends at 728 of PEDF GENE

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:41:

TTCTTTTTTT GAGACGGGGT CTCGCTCTGC TCGCCCAGGA 40

TGGAGTGCAG TGGTGTGATC TCAGCTCACT GCAACCTCCG 80

CCTCCCAGGT TTAAGTGATT CTCCTGCCTC AGACTCCCAA 120

GTAGCTGGGA CTACAGGTGC GCGCCAACAC ACCTGGGTAA 160

TTTTGTTTGT ATTTTTAGTA GAGATGGGGT TTCACCGTGT 200

TGACTAGGCT GGTCTCGAAC CTCCTGACCT CAGGTGATCC 240

CCCGGCCTCG GTCTCCCAAA GTGCTGGGGA TAACAAGCGT 280

GAGCCACTGC GCCCAGCTTT GTTTGCATTT TTAGGTGAGA 320

TGGGGTTTCA CCACGTTGGC CAGGCTGGTC TTGAACTCCT 360

GACCTCAGGT GATGCACCTG CCTCAGTCTC CCAAAGTGCT 400

GGATTACAGG CGTTAGCCCC TGCGCCCGGC CCCTGAAGGA 440

AAATCTAAAG GAAGAGGAAG GTGTGCAAAT GTGTGCGCCT 480

TAGGCGTAAT GGATGGTGGT GCAGCAGTGG GTTAAAGTTA 520

ACACGAGACA GTGATGCAAT CACAGGAATC CAAATTGAGT 560

GCAGGTCGCT TTAAGAAAGG AGTAGCTGTA ATCTGAAGCC 600

ATCTGAAGCC TGCTGGACGC TGGATTAGAA GGCAGCAAAA 640

AAAGCTCTGT GCTGGCTGGA GCCCCCTCAG TGCAGGCTTA 680

GAGGGACTAG GCTGGGTGTG GAGCTGCAGC GTATCCACAG 720

GCCCCAGGGT AAAGTAG 737

(2) INFORMATION FOR SEQ ID NO:42:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 88 Base Pairs
(B) TYPE: Nucleic Acid
(C) STRANDEDNESS: Double
(D) TOPOLOGY: Unknown

(ii) MOLECULE TYPE: Genomic DNA

(ix) FEATURE:
(A) NAME/KEY: PEDF Promoter
(B) LOCATION:
(C) IDENTIFICATION METHOD:
(D) OTHER INFORMATION: EXON PEDF GENE begins at 9

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:42:

TTCTTGCAGA TGCAGGCCCT GGTGCTACTC CTCTGCATTG 40

GAGCCCTCCT CGGGCACAGC AGCTGCCAGA ACCCTGCCAG 80

CCCCCCGG 88

(2) INFORMATION FOR SEQ ID NO:43:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22481 Base Pairs
(B) TYPE: Nucleic Acid
(C) STRANDEDNESS: Double
(D) TOPOLOGY: Unknown

(ii) MOLECULE TYPE: Genomic DNA

(ix) FEATURE:
(A) NAME/KEY: P1-147
(B) LOCATION:
(C) IDENTIFICATION METHOD:
(D) OTHER INFORMATION: full length genomic sequence for PEDF plus flanking sequences.

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:43:

GCGGCCGCAG GGTGGACTGT GCTGAGGAAC CCTGGGCCCA 40

GCAGGGGTGG CAGCCCGCGC AGTGCCACGT TTGGCCTCTG 80

GCCGCTCGCC AGGCATCCTC CACCCCGTGG TCCCCTCTGA 120

CCTCGCCAGC CCTCCCCCGG GACACCTCCA CGCCAGCCTG 160

GCTCTGCTCC TGGCTTCTTC TTCTCTCTAT GCCTCAGGCA 200

GCCGGCAACA GGGCGGCTCA GAACAGCGCC AGCCTCCTGG 240

TTTGGGAGAA GAACTGGCAA TTAGGGAGTT TGTGGAGCTT 280

CTAATTACAC ACCAGCCCCT CTGCCAGGAG CTGGTGCCCG 320

CCAGCCGGGG GCAGGCTGCC GGGAGTACCC AGCTCCAGCT 360

GGAGACAGTC AGTGCCTGAG GATTTGGGGG AAGCAGGTGG 400

GGAAACCTTG GCACAGGGCT GACACCTTCC TCTGTGCCAG 440

AGCCCAGGAG CTGGGGCAGC GTGGGTGACC ATGTGGGTGG 480

GCACGCTTCC CTGCTGGGGG TGCAGGGGGT CCACGTGGCA 520

GCGGCCACCT GGAGCCCTAA TGTGCAGCGG TTAAGAGCAA 560

GCCCCTGGAA GTCAGAGAGG CCTGGCATGG AGTCTTGCTT 600

CTTGCAAACG AGCCGTGTGG AGAGAGAGAT AGTAAATCAA 640

CAAAGGGAAA TACATGGTCT GTCCGAGGAT GAGCTGCCGG 680

AGAGCAATGG TGAAAGTGAA GTGGGGGAGG GGGCGGGGCT 720

GGGAGGAAAA GCCTTGTGAG AAGGTGACAC GAGAGCACGG 760

CCTTGAAGGG GAAGAAGGAG GGCACTATGG AGGTCCCGGC 800

GAAGCGTGGC CTGGCCGAGG AACGGCATGT GCAGAGGTCC 840

TGCCGAGGAG CTCAAGACAA GTAGGGGACG GTGGGGCTGG 880

AGTGGAGAGA GTGAGTGGGA GGAGGAGTAG GAGTCAGAGA 920

GGAGCTCAGG ACAGATCCTT TAGGCTCTAG GGACACGATA 960

AACACAGTGT TTTTTGTCTT GTCAAGTGTG TCCTTTTTAT 1000

TTTTTTGAAA GAGTCTCGCT CTGTAGCCCA GGCTGGAGTG 1040

CAGCGGTGCG ACCTCGGCTC ACTGCAACCT CTGCCTCCCG 1080

GGTCCAAGCA ATTCTCCTGC CTCAGCCTCC CGAGTAGCTG 1120

GGATTACAGG CACCCGCCAC CACGCACTGC TAATTTTTGT 1160

ATTTTAGTAG AGACCGGGTT TTGCCATGTT GGTCAGGCTG 1200

GTCTCGAACT CCTGACCTCA GGTGATCCGC CCGCCTCGGC 1240

CTCCCAGAGT GGTGTGAGCC ACTATGCCCT GCAGCACTTG 1280

TCAAGTCTTT CTCAGCGTTC CCCTCCTCTC CACTGCAGCT 1320

CCCAGTGCCC CAGTCTGGGC CTCGTCTTCA CTTCCTGGGA 1360

TCCCTGACAT TGCCTGCTAG GCTCTCCCTG TCTCTGGTCT 1400

GGCTGCCTTC ACTGTAACCT CCACCCAGCA GGTACCTCTT 1440

CAGCACCTCC CATGAACCCA GCAGAATACC AAGCCCTGGG 1480

GATGCAGCAA CGAACAGGTA GACGCTGCAC TCCAGCCTGG 1520

GCGACAGAGC AAGACTCCGC CTGAAGAAAA AAAAAAGGAC 1560

CAGGCCGGGC GCGGTGGCTC ACGCCTGTAA TCCCAGCACT 1600

TTGGGAGGCC GAGGTGGGTG GATCATGAGG TCAGGAGTTC 1640

AAGACCAGCC TGGCCAAAAT GGTGAAACCC CGTCTCTACT 1680

GAAAAATACA AAAATTAGCT GGGTGCAGTG GCGGGCGCCT 1720

GTAGTCTCAG CTACTCAGGA GGCTGAGGCA GGATAATTGC 1760

TTGACCCCAG GAGGCAGAGG TTGCAGTGAA CCGAGATCAC 1800

GCCACTGCAC TCCAGCCTGG GCGACAGAGC AAGACTCTGC 1840

CTCAAAAAAA AGAATAAAAA TAAAAAAAAG GACCAGATAC 1880

AGAAAACAGA AGGAGACGTA CTATGAAGGA AATTGGAGAG 1920

CTTTTGGGAT ACTGAGTAAC TCAGGGTGGC CTTTCCCAGG 1960

GGACATTTAG CTGAGAGATA GACGGTATGA AGACCTGACC 2000

GTTCAGAAAC AGGGGAAGAG GCAGCAGCCC GGGCAAAGGC 2040

CTTTGGGGCA GGAAAGGGCT TGGATCACTG GAGAAGCAGA 2080

AAGATGGCCA GTGTGACCAG AGTGTGACAA AGTCAGAGAA 2120

AACCAGGAAG ATGGAGCTGG AGACACAGGC GGGGCCAGAT 2160

CACGAGGGTC CTCGCAGACC AGAGCAAGGG TTTGGATTTT 2200

ATTCCAAGTA TGAAGGGAAG CTGCTGAAGT GTGTTTTCCT 2240

TTACAATTTG TAGTTGAAAT ATAATATGCA AAGTACACAA 2280

GTCTTAACTA TATGTAAGCT TAATGAATGT TTCCATGAAC 2320

CAAATACCGC TGTGCAACCA TCACCAGCTC AAGAGACGAA 2360

CCCTTCTCCC TCCTCCTGAC TGCCAGTAAC ATAGTGGTTC 2400

AGCTCAAGAA ACAGAACTCT TCTGACTTCC CCTAACATAG 2440

CGGGTTTTCT TTTTTGTTTT GTTTTTTGTT GTTTTTTAAG 2480

AGACAATGTC TTTATTATTT TTATTTTTTT TTATTTTTGA 2520

GACGGAGTCT TGCTGTCGCC CAGGCTGGAG TGCAGTGGTG 2560

CGATCTCGGC TCACTGCAGG CTCTGCCCCC CGGGGTTCAT 2600

GCCATTCTCC TGCCTCAGCC TCCCTAGCAG CTGGGACTAC 2640

AGGTGCCCGC CACCTCGCCC GGCTATTTTT TTGTATTTTT 2680

AGTGGAGACG GGGTTTCACC GTGTTAGCCA GGATGGTCTC 2720

GATCTCCTGA CCTCGTGATC CGCCCACCTC GGCCTCCCAA 2760

AGTGCTGGGA TTACAGGCAT GAGCCACCGC GCCCAGCCAA 2800

GAGACACGGT CTTGCTCTGT CGCCCAGGCT GGATGGAGTG 2840

CCGTGGTGCG ATCACAGCTC GCGGCAGCCT TGACATCCTG 2880

GGCTCAAGCA ACCTTCCTGC CTTGGCCTCC CAAATGTTGG 2920

GATTATAGGC ATGAGCCACT GTGCTTGGCA TCTATTCATC 2960

TTTAATGTCA AGCAGGCAAT TGAATATTTG ATCAGGGATA 3000

GAATTGTCTA TTTGGGGGTA TGCAGATGTG CTTCATGTCA 3040

TGGAACTGGG CCGGGCGCGG TGGCTCATGC CTATAATCCC 3080

AGCACTTTGG GAGGCCGAGG CAGGCGGATC ATAAGGTCAG 3120

GAGATCGAGA CCATCCGGGC CAACACGGTG AAACCCCGTC 3160

TCTACTAAAA ATACAAAAAT TAGGCAGGTG TGGTGGTGCG 3200

TGCCTGTAGT CCCAGCTACT CAGGGAGGCT GAGACAGGAG 3240

AATTGATTGA ACCTGGGAGG CAGAGGTTGT AGTGAGCCAA 3280

GATCGCGCCA CTGCACTCCA GCCTGGGCGA CATGAGCGAG 3320

ACTCCGTCTC AAAAATAAAC AAAAAAAAGT CATGGAATTG 3360

ATGGAAATTG CCTAAGGGGA GATGTAGAAG AAAAGGGGTC 3400

TCAGGATCAA GCCAGCAGAG AAGGCAGAAA AGGTAAGGTG 3440

TGTGAGGTGG CAGAAAAAGG GAAGAGTGTG GACAGTGAGG 3480

GTTTCAAGGA GGAGGAACTG TCTACTGCCT CCTGCCAAGG 3520

ACGGAGGTGT CCACTGCCAG TTGACATAAG GTCACCCATG 3560

AACTTGGTGA CAGGAATTTC AGTGGAGAAG TGGCCACAGA 3600

CACAAGTCTA GAATTGAAAT GGGAGCCGAG GCAGCGTAGA 3640

CAAAAGAGGA AACTGCTCCT TCCAGAGCGG CTCTGAGCGA 3680

GCACCGAGAA ATGGGCAGTG GCTTTAGGGG ATGTAGCGTC 3720

AAGGAAGTGT CTTTTAAAGA AGTCGGGGGC CGGGCACGGT 3760

GGCTCACGCC TGTAGTCCCA GCACTTTGGG AGGCCGAGGC 3800

AGGCAGATCA CTTGAGGTCA GGAGTTCGAG ACCAGCCTGG 3840

CTAACACGAT GAAACCCCGT CTCTACTAAA AATACAAAAA 3880

ATTAGCTGGG CACGGTGGCT CGTGCCTGTA ATCCCAGCAC 3920

TTTGGGAGGC AGAGGTGGGC AGATCACTTG AGGTCAGGAG 3960

TTTGAGACCA GCCTAGCCAA CATGGTGAAA CCCCATCTCT 4000

ACTAAAACTA CAAAAATTAG CCGGGAGTGG TGGCACGTGC 4040

CTGTAATCCC AGCCAGTCAG GAGGCTGAGG CAGGAGAATC 4080

ACTGGAATCC TGGAGGTGGA GGTGGCAGTG AGCCGAGATG 4120

GTACCTCTGT ACTCCAGCCT GGGGGACAGA GTGAGACTCC 4160

GTCTCAAAAA AAAAAGAAGG TGGGGAAGGA TCTTTGAGGG 4200

CCGGACACGC TGACCCTGCA GGAGAGGACA CATTCTTCTA 4240

ACAGGGGTCG GACAAAAGAG AACTCTTCTG TATAATTTAT 4280

GATTTTAAGA TTTTTATTTA TTATTATTTT TTATAGAGGC 4320

AAGCATTTTT CACCACGTCA CCCAGGCTGG TCTCCAACTC 4360

CTGGGCTCAA GTGTGCTGGG ATTATAGCCA TGAGTCACCA 4400

CACCTGGCCC AGAAACTTTA CTAAGGACTT ATTTAAATGA 4440

TTTGCTTATT TGTGAATAGG TATTTTGTTC ACGTGGTTCA 4480

CAACTCAAAA GCAACAAAAA GCACCCAGTG AAAAGCCTTC 4520

CTCTCATTCT GATTTCCAGT CACTGGATTC TACTCTTGGG 4560

ATGCAGTGTT TTTCATCTCT TTTTTGTATC CTTTTGGAAA 4600

TAGTATTCTG CTTTAAAAAG CAAATACAGG CCAGGTATGG 4640

TGGCTCACTC CTGTAATCCC AGCACTTTGG GAGGCCGAGG 4680

CAGGTGATCA CCTAAGGTCA GGAGTTCAAG ACCAGCCTGG 4720

CCAATATGGT GAAACCCTGT CTGTACCAAA ACACAAAAAC 4760

AAAAACAAAA ACAAAAATTA GCCGGGCGTG GTGGCGTGCT 4800

CCTGTAATCC CAGCTACTCA GGAGGCTGAG GCAGGAGAAT 4840

CGCTTGAACC TGGGAGGCAG AGGTTGCAGT GAGCCGAGAT 4880

TGTGCCACTG TACTCCAGCC TGGGCCACAG AGCAAGGTTC 4920

CATCTCAAAC AAAACAAAAC AAAACAAACA AAAAAACAAA 4960

ACAAAAGCTA ATACAAACAC ATATACAATA GACAAAACTG 5000

TAAATATTTT ATTATTTTTA TTTTTTTTAG TAGAGACAGG 5040

GTTTCACCAT GTTGGCCAGG ATGGTCTCAA ACTCCTGACC 5080

TCAGGTGATC CACCCACCTC AGCCTCCCGA TAGTTAGGAT 5120

TACAGGCATG AGCCACCACA CCCGGCCTAA AATTGTAAAC 5160

GTTTTAGAAG AAAGTATAGA TGAATCCCTT CGTGATCTCG 5200

GGGAAGAAGA GATTTTTTAA AAAAGATACC AAAAGAAGCA 5240

CAAATTATAA AAGAAAAGAT TGAAAATGTT GGTGTTAAAA 5280

TTAAAAACTT GTTTTAAAAC AAGCTTGTGT AACCCATGAC 5320

CCACAGGCTG CATGTGGCCC AGAAAAGCTT TGACTGCAGC 5360

CCAACACAAA TTCGTAAACT TTCCTAAAAC ATTATGAGAT 5400

TTTTTTTGAG ATTTTGTTTT GTTTTGTTTT TTGTTTTTTT 5440

AGCTCATTCG GTATCATTAA TGTTAGCATA TTTTACGTGG 5480

GGCCCAAGAC AATTCTTCTT CCAATGTGTC TCAGGGGAGC 5520

CAAAAGATTG GACACCCCTG CCATAAACAT GAAAAGACAA 5560

TGGCCGGGCA CGGTGGCTCA CGCCTGTAAT CCCAGCACTT 5600

TGGGAGGCTG AGGGGGGCGG GATCACCTGA GGTCAGGAGT 5640

TTGAGACAAG CGTGACCAAT GTGGTGAAAC CCTGTCTCTA 5680

CTAAAAATAC AAAAATTAGC CGGGCATGCT CGTGCACACC 5720

TATAGTCCCA ACTACTCAGC AGGGTGAGGC AGGAGAACCT 5760

CTTGAACCCG GGAAGCGGAG GTTGCAGTGA GCCGACATTG 5800

CACCCCTGCA CTCCAGCCTG GGTGACAGAG TGAGTCTCCA 5840

CTGGAAAAAA AAAAAAAAGA ACAGTGTGAT ACATTGACCT 5880

AAGGTTTAAG AACATGCAAA CTGATACTAT ATATCACTTA 5920

GGGACAAAAA CTTACATGGT AAAAGTAAAA AGAAATGTAC 5960

GAAAATAATA AAAATCAAAT TCAAGATGGT GGTTATGGTG 6000

ACGGGAAAGA ACTGAGGCGG AAATATAAGG TTGTCACTAT 6040

ATTGAGAAAT TTTTCTATCT TTTTTTCTTT TTTCTTTTTT 6080

TGAGACGGGG TCTCGCTCTG TCGCCCAGGA TGGAGTGCAG 6120

TGGTGTGATC TCAGCTCACT GCAACCTCCG CCTCCCAGGT 6160

TTAAGTGATT CTCCTGCCTC AGACTCCCAA GTAGCTGGGA 6200

CTACAGGTGC GCGCCAACAC ACCTGGGTAA TTTTGTTTGT 6240

ATTTTTAGTA GAGATGGGGT TTCACCGTGT TGACTAGGCT 6280

GGTCTCGAAC TCCTGACCTC AGGTGATCCC CCGGCCTCGG 6320

TCTCCCAAAG TGCTGGGATA ACAAGCGTGA GCCACTGCGC 6360

CCAGCTTTGT TTGCATTTTT AGGTGAGATG GGGTTTCACC 6400

ACGTTGGCCA GGCTGGTCTT GAACTCCTGA CCTCAGGTGA 6440

TGCACCTGCC TCAGTCTCCC AAAGTGCTGG ATTACAGGCG 6480

TTAGCCCCTG CGCCCGGCCC CTGAAGGAAA ATCTAAAGGA 6520

AGAGGAAGGT GTGCAAATGT GTGCGCCTTA GGCGTAATGG 6560

ATGGTGGTGC AGCAGTGGGT TAAAGTTAAC ACGAGACAGT 6600

GATGCAATCA CAGAATCCAA ATTGAGTGCA GGTCGCTTTA 6640

AGAAAGGAGT AGCTGTAATC TGAAGCCTGC TGGACGCTGG 6680

ATTAGAAGGC AGCAAAAAAA GCTCTGTGCT GGCTGGAGCC 6720

CCCTCAGTGT GCAGGCTTAG AGGGACTAGG CTGGGTGTGG 6760

AGCTGCAGCG TATCCACAGG TAAAGCAGCT CCCTGGCTGC 6800

TCTGATGCCA GGGACGGCGG GAGAGGCTCC CCTGGGCTGG 6840

GGGGACAGGG GAGAGGCAGG GGCACTCCAG GGAGCAGAAA 6880

AGAGGGGTGC AAGGGAGAGG AAATGCGGAG ACAGCAGCCC 6920

CTGCAATTTG GGCAAAAGGG TGAGTGGATG AGAGAGGGCA 6960

GAGGGAGCTG GGGGGACAAG GCCGAAGGCC AGGACCCAGT 7000

GATCCCCAAA TCCCACTGCA CCGACGGAAG AGGCTGGAAA 7040

GGCTTTTGAA TGAAGTGAGT GGGAAACAGC GGAGGGGCGG 7080

TCATGGGGAG GAAAGGGGAG CTAAGCTGCT GGGTCGGGTC 7120

TGAGCAGCAC CCCAAGACTG GAGCCCGAGG CAAGGAGGCT 7160

CACGGGAGCT GCTTCCACCA AGGGCAGTCA GGAAGGCGGC 7200

CGCCCTGCAG CCCAGCCCTG GCCCCTGCTC CCTCGGCTCC 7240

CTGCTACTTT TTCAAAATCA GCTGGTGCTG ACTGTTAAGG 7280

CAATTTCCCA GCACCACCAA ACCGCTGGCC TCGGCGCCCT 7320

GGCTGAGGGC TGGGATGGAG GACAGCTGGG TCCTTCTAGC 7360

CAGCCCCCAC CCACTCTCTT TGGCTACATG AGTCAAGGCT 7400

GGGCGACCAA TGAGGTTGTG GCCTCCGGCA AACAATGACC 7440

ACTATTTAGG CCGGCAGGTG TATAGGGCGT GGGGGCCCAG 7480

CTGCCAGTGC TGGAGACAAG GGCTGTCCGA GATGAACCCT 7520

TTCTGCTGCC TGCCAAGCCA CTGGGAGGGG TAGGTCTCAG 7560

CAGGATTCCC AGAAACCCCG CCCCTGTCCA GCCTAGGCCC 7600

CCCACCCGGT GTTAGCTAAC CCAACGTTAG CCCCCAGGTT 7640

CCGTGGGGTT GGGGGGCAGG GAGTCCTATT CTTGGGGCTG 7680

CTGCTTCTGG GGTGTGGGGA AGTGCAACTC CACGGCACCC 7720

TGGGCTGACT CATTCAGCTT CTAAAGCTTC AGGAAACATT 7760

GTTTGGGGCT GGGTCACCAT GGGTGGGCCA GAGAGGACCC 7800

CTCAATCCCC TCCGGAGAGC CAGGGGAGGG GGAGGTGCCC 7840

TTCCCCATGC TATCTCCGAG GCCCACTGCC ATGTGGCTGA 7880

AGGCTGTGCG GTTCTGGGAA GAGGGGGAGG TGGCGGTGGA 7920

GGCTGTTTGT CTCCTAACTG GGCTTAATCT GAAACACATG 7960

TATTGGCTTG AGTTGATCCG CCTCACGTGG AGGCAAGATC 8000

ACAAAAGCTT CTGTGTTTCT TGATGTGGGC AATTGTCAGA 8040

AAATAAGGCC TGACCTTGGC CCAGCAGGGA GGGTATCTAC 8080

CTCTCCCTGA GCCCTCCCCC GCCTGCTAGG ACGAGAGCGG 8120

GGCTTGGATA CTGCCCTTTG GACAGGATGG CATCATTGTC 8160

TGTGGCTGCA GCCAGCCAGC GGTCGCCTGC TCAGCCCATG 8200

AGCAACCACT GTGGACAGGG TATTGCGTGT GTGCTGAGGG 8240

GCGTCCATGC AGACCCCCAC GCTTGCCCTC TCACTGCCCT 8280

TGTAGGGTTT TCAATCATCT CTCCTCTTCC CTTATCCAGA 8320

TGGCTTGAAG TGGAGGATTC AGACTTGCCG TTAATACTCT 8360

GGGTCCCTGT GTCTAGCTCG GGGCCACCTT TGGACCCATG 8400

TCCCTTCCCT GCCAGGCTCC CTCACCTCAC CTCAGCCTAC 8440

CCACATTGTG ACAATCATCT ACCACCTGAT CTGGGGTTTG 8480

GGCTTAGATT CTGTAGGCAC CAAGACTAAA GTCGCTCCTT 8520

CAAGTCCATT TGAATTGTGA CTTTAGTTTC CTTAAATACT 8560

ATGCCAGGAT AATGGCCAGG GATGGTGGCT CACGCCTGTA 8600

CTCCTGGCAC TTTGGGATGC TGGTGGATCA CCTGAGATCA 8640

GGATTCCAGG CCAGCCTGGC CAACACGGTG AAACCCCATC 8680

TCTACTAAAA CATAAAAATT AACCAGGTGT GGTGGCGGGC 8720

ACCTGTAATC CCAGCTACTC AGGAGACTGA GGCAGGAGAA 8760

TTGCTTGAAC CCGGGAGGTG GAAGTTGCAC TGAGCTGAGA 8800

TCGCGCCACT GCACTTTAGC CTGGGCGACA AGAGTGAAAC 8840

TCTGTCTCAA AAACAAAAAA AACTATGCCG GGATGAGCCT 8880

GTCTCCTCCC TTAATTTCTT ACTTGGGCCA GAGGAACTAG 8920

AACTAACAAC TTCTCTTCTA GCCTTGCCTC CTGTGTACCT 8960

CACTGAATTT TTGGTCTCTA ATAAACCAGT CTGCAGAGGC 9000

TCAGGGGAGG CAGGCTCCTG GCAGCTGGGT GGGGCTGGCC 9040

CCAGCCGGGT GGAGACCAGC TGTAGGCCTG GATGGTGGTG 9080

AGGCCTCTGT CTTGCACTGC AGAAAGCTTT TCCTGTTGTC 9120

TACACGAAAG TTTTCTCCCT GCATGTCAGG GCAGCCACGT 9160

GCAAGAGCAG CTGGCTGGGA ACGCAGAGGT CTGCGGCTCG 9200

AGGCGGGGTT TAGAAAGAAA ACCAGGCTGC TTCCTGCTGC 9240

CCGTCCTGCC TTAAGCTGAG TAAACTCAAA GGCAATCTTC 9280

TTTCATGCCT CACGATATTG TCCAGTGGAT TATCTGATTT 9320

AATTTGAAGG ACGAGAGCCA ACAATCACAC AACGTCCTCC 9360

CAAATTTTCT GATCCACTTT GTTCTGGGAA GTCAAAAAGT 9400

GCGTGTGCTG TGTGGGTGGA TGTTTGTGTA TATAAATGGA 9440

TAATGAAGGA TGATGTGTTG GGGGCCAGGG CAGGGGAGAC 9480

AACGCTGTTC AGATTCTACA TTTTTTTTTC CTTTTTTTTT 9520

TTTTTTTGAG ATGGAGTCTT GCTCTGTTGC CCAGCCTGGA 9560

GTGCAGTGGC GCGATCTCAG CTCACTGCAA CCTCCACTTC 9600

CTGGATTCAA GTGATTCTCC TGCCTTAGCC TCCCAAGTAG 9640

CTGGGATTAC AGGCATGCGC CACCACACCC GGCTAATTTT 9680

TGTATTTTTA GTAGAGATGG GGTTTCTCCA TGTTGGCCAG 9720

GATGGTCTCA AACTCCTGAC CTCAGGTGAT CTACCCGCCT 9760

CGGCCTCTCA AAGTGCTGGG ATTACAGGTT TGAGCCACTG 9800

CGCCTGGCCT TTTTTTTTTT TTTTGAGATG GAGTTTTCAC 9840

TCTTGTTGCC CAGGCTGGAG TGCAGTGGTG CGATCTTGGC 9880

TCACTGCAAC CTCCACCTCC CAAGTTCAAG TGATTCTCCA 9920

GCCTTAGCCC TCCAAGTAGC TGGGACTACA GGTGTGTGCC 9960

ACCATGCCTG GCTATTTTAT TTTATTTTAT TTTATTTATT 10000

TATTTTTGAG ACTAAGTCTT GCTCTGTTGC CCAGGCTGGA 10040

GTGCAGTGGC ATAATCGGCT CACTGCAACC TCTGCCTCCC 10080

AGGTTCAAGT GATTCTCCTG CCTCAGCCTC CTGAGTAACT 10120

GGGATTACAG GGGCCTGCCA CCACGCCTGG CTACTTTTTG 10160

TATTTTTAGT ATAGATGGGG TTTCACCATG TTGGCCAGGC 10200

TGGTCTCGAA CTCCTGACCT CAGGCTATCC GCCTGCCTCA 10240

GCCTCCCAAA GTGCTGGGAT TACAGGCATG AGCCACTGTG 10280

CTCGGTAGTT GTTTTATTTT AATAGTAGGT TATTTTATTT 10320

CCATTTTACA AGAGAAAAAA TGGTGATTTA AAGAGCTACT 10360

AAGACACAGC ACTGAGACCA TGTGTGATGG CATGCGCCTG 10400

CAGTCCCAGC TACTCACGAG GCTGAGGCAG GAGGATCACA 10440

TGAGGTCAGG AGTTCCAGGC TGTGGAGTGC TATGGTTGTG 10480

TAGTGAATAG CCACTACACT CCAGCCTGGG CAGCACAGCA 10520

AGATCTTGTC TCCCAAAAAA AAAAAAAAAA AAAAATTTCA 10560

AATGTGAACC CAGGATCTCT GACCCTAGGC CCTGCACTCC 10600

TAACCATGGG AGGAAGAGCT CTTGAAAGGG AACTGTGGGA 10640

GAAGGGAATG AGCTGCCTTG TGAGGCCACA GAAGTCCAAA 10680

GACAGCTTGA GAATTTGGAG GGACAGCACG TGCCGGACTG 10720

GGTGCCTCTA TGCTTGGTAT CCGGTGATTC CATGGAGGAG 10760

ACCTGGGTTC TGCCCCATTC TCCTGGGAGG GGTTGCCCAA 10800

AGTCTTATCA CCGGAGTGGG TCAGCTGCCT CCAGGACAAA 10840

GCTTTAGCAT ACACTTGTGC TGGGCCATAC TCCACGTGGA 10880

GAAGCCCTGC TGGGGCTGGG GCCCCACTGC TCTGGATCTT 10920

TAAAAGCTAT TGGTTCAGGG GCCAGGTGTA ATGGCTCACA 10960

CCTATAACCC TAGCACTTTG GGAGGCTGAA GCAGGTGGAT 11000

AGCCTGAGGT CAGGAGTTTG AGACAAGCCT GATGAACGTG 11040

GTGAAACCCC ATCGCTATTA AAATACAAAA AATTAGCCGG 11080

GCATGGTGGC AGGTGCCTGT AATTCCAGCT ACTTGGGAGG 11120

CTGAGGCGGG AGAATCGCTT GAACCCAGGA GGCGGAGGTT 11160

GCAGTGAGCC AAGATCGCTC CACTGTACTC CAGCCTGGGC 11200

GACAGAGCCA GACTCTGTTT CAAAAAATAA AATATAAATA 11240

AATAAATAAA TAAATAAATA AATAAATAAA AGCTTTAGGC 11280

TTAAAGGAGG GTCCCCTGAC GCAGACAGTG GAACAAAAGC 11320

ACAAGCTTAT GGTATGACTG TGGGCCCTGA GGCAGGGGGA 11360

GGGGCGGGAG AACCTTGCTG GGAGGGATGG GCCATCAAGC 11400

TGAGGGTCCA CTTCTGGGGG CCTGGAGGGG TGAGGGGTGG 11440

TCGCTGCAGG GGGTGGGGGA AAGTGACTAG CCCTGCCCAA 11480

CCCCTGGGTC CTGGCTGGGG TGGCCAGGAA GGGGTAGCGG 11520

GGCAGTGCAG TGTCGGGGGA GAGCGGCTTG CTGCCTCGTT 11560

CTTTTCTTGC AGGCCCCAGG ATGCAGGCCC TGGTGCTACT 11600

CCTCTGCATT GGAGCCCTCC TCGGGCACAG CAGCTGCCAG 11640

AACCCTGCCA GCCCCCCGGA GGAGGTCAGT AGGCAGGCGG 11680

GGAGGGCGTG GTCAGCATTC CCCGCCCCTC CTTGGCAGGC 11720

AGCACGGGAA ACAGGACAGG GAACCCGGAC CCAGGTTCCA 11760

GGCCAGGCTT GGGCCTTTAT TTCTCTAGGG CTGGAGTTTC 11800

TCCAGCAGCA AAACAGAGAG AAAATGTCTT GCCTTGCCTT 11840

TCAGGGGATG GAGTAGGGAC ATGAATAAGA TCCCAAAAGA 11880

GTAAAAATCT GAAGCACTTT TAACAAGTCC AGGGCAATTC 11920

TCCTGCCTCA GCTTCCCAAG CAGCTGGGAT TACAGGCATG 11960

CACCACCAAG CCCGGCTCAT TTTGTATTTT TAGTAGAGAC 12000

GGGGTTTCTC CATGTTGGTC AGGCTGGTCT CGAACTCCCG 12040

ACCTCAAGTG ATTCTCCTGC CTCGGCCTCC CAAAGTGCCG 12080

GGATGACAGG TGTGAGCCAC CGCACCTGGC CAGGATCTTT 12120

TCTCATTACC TTGTCTTCCT AGTGGGGGCT CCACTGAGCA 12160

GGTCATGTTC CCGGACATTT GTTCGGATAC TGACCAGGCT 12200

GTGGCAGGGA GTGAGGGTAT GGAGTGACCT CTCTCCTGCC 12240

CAGAAAGGGC GCAGCTGGGT TCCCAAGGCA GATACAGGCA 12280

CATGGAGGGA AGCCTGGGCC ATATGAGTGT TATGGGGTGA 12320

GTGTTGGCGG AGGCCCACCC TTGAGGGACA AGAGCAGCTG 12360

GGCATCTTGG CGAGAGCCCT GGACTTTCGT GAGGTCAGAG 12400

TATGAATTCT GCGTCTCCCT CTTCCTAGCT TTGTGACCCT 12440

AGACAACCCT TACCTCAGTC TTTGCTTCCT TGCCTATGAA 12480

ATGGGATAAA AACACCCATT CTACAGGGCC ATGTGGCCAC 12520

TCATTTATTT CTCATCTACC AAACACCTAC TCGACAGGGG 12560

CTGGCAATGG GCGGAAATAA AAACTCAGTT CTGCCGGGTG 12600

CGGTGGCTCA CACCTGTAAT CCCAGCAGTG TGGGAGGCGG 12640

AGCAGGACGA TCCCTTGAAT CCAGGAGTTT GAGACCAGCA 12680

TAGGCAACAT AGTGAGACCC CTGTCTCTAC ACAAAAGCAA 12720

AAATTACCAG GCGTGGTGGC AAGTGCTTGT GGTACTACCT 12760

ACTTGGGAAG CTGAGGTGGG AGGATCACTT GAGCCCAGGA 12800

GATTAAGACT GCAGTGAGGG GCCGGGCGCG GTGGCTCACG 12840

CCTGTAATCC CAGCACTTTG GGAGGTGGAG GTGGGTGGAT 12880

CACGAGGTCA GGAGATCGAG ACCATCCTGG CTAACACGGT 12920

GAAACCCCGT CTCTACTAAA AATACAAAAA ATTAGCTGGG 12960

TGTGGTGGGG GGCGCCTGTA GTCCCAGCTA CTCGGGAGGC 13000

TGAGGCAGGA GAATGGCGTG AACCCGGGAG GTGGAGGTTG 13040

CAGTGAGCTG AGCTCGCACC ACTGCACTCC AGCCTGGGCG 13080

ACAGAGTGAG ACTCCGTCTC AAAAAAAAAA AAAAAAAAAA 13120

GAAAGAAAGA AAAACTGAGT TCTTTTTTTT AACTTTCTTT 13160

TTTTAGAGAC AGAGTCTCAC TCCATCACCC ATGCTGGAGT 13200

ACAGTGGTGC GATCTTGGCT CACTGCAATC TTGGCCTCCT 13240

GAGTTCAACC AATTCTCATG CCTCAGCCTC CCAAATAGCT 13280

GGGACCACAG GCACGTGCCA CCACGCCCAG CTAATTTTTT 13320

GGGTATTTTT AGTAGAGATG GGGCCTCACC ATGTTGCTCA 13360

GGTTGGTCTG AAACTCCTGA GCTCAAGTGA TCCATCTTCC 13400

TCGGCCTGCC AAAGTGCTGG GATTATAGGC ATAAGCCACT 13440

GCACCTAGCT CCCAATTTTT ATATTTATAT TTATTTTTAT 13480

TTACTTATTT ATTTTTTGAG ACAGGGTCTC ACTCTGTCAC 13520

CCAGGCTGGA GTACAGTGGC ACTATCTCAG CTCACTGCAA 13560

CCTCTGCCTC CTGGGTTCAA GCGAATCTCG TGCCTCAGCC 13600

TCCTGAGTAG CTGGGATTAC AGGCATGCAC CACCATGCCC 13640

CGTTAATTTT TTTGTATTTT TAGTAGAGAC GGGTTTCACC 13680

GTGTTGCCCA GGATGGTCTC GAACTCCTGA CCTCAAGTGA 13720

TTCACCCACC TCAGCCTCCC AAAGTGCTGG GATTATAGGT 13760

GTGAGCCACT CGGCTGATGG TTTTTAAAAA GTGGGTCATG 13800

GGGCTGGGCG CGGTGGCTCA TGCCTGTAAT CCCAGCACTT 13840

TGGTAGACCG AGGCGGGTGG ATCACAAGGT CAGGAGATCG 13880

AGACCATCCT GCCTAACACG GTGAAACCCC GTCTCTACTA 13920

AAAATACAAA AAATTACCCA GGCATGGTGG TGGGCGCCTG 13960

TAGTCCCAGC TACTCGGGAG GCTGAGGCAG GAGAATGGCG 14000

TGAACCTGGG AGGCGGAGCT TGCAGTGAGC CGAGATCACG 14040

CCACCGTACT CCAGCCTGAG CGACAGAGCG AGACTCCGTC 14080

TCAAAAAAAA AAAAAAAAAG TGGGTCATAG GTTTCGGCTT 14120

ATAGGTCACA AGTGTTTAAA CCTGGCCATG AGGCCAGGCG 14160

CAGTGGCGCA TGCCTGTAAT CCCAGCCATT TGGGAGGCTA 14200

AGGCAGGAAA ATCGCTTGAA CCGGGGAGGT GGAGGTTGCA 14240

GTGAGCTGAG ATCGCGCCAC TGAACTCTAG CCTGGGTGAC 14280

ACAGTAAGAC TCTGTCTCAA ATAAAAAAAA AAACAGCTGA 14320

TCTCTCTTCT GCGCTGTCTC TCCACAGAGA GCTCATGCGT 14360

GATCAGGGAG TAAAACTCAT TCCCGTTTTA GGCCAAACAC 14400

AGAAAAATTA GGAAGGACAG CCCCAAGGGG CCAGAACCAC 14440

CACCCTACAC AAAGCCGTGA GGAGACAGTC CCTGTGCATC 14480

TCTGCGAGTC CCTGAACTCA AACCCAAGAC TTCCTGTCTC 14520

CTGCCAGGGC TCCCCAGACC CCGACAGCAC AGGGGCGCTG 14560

GTGGAGGAGG AGGATCCTTT CTTCAAAGTC CCCGTGAACA 14600

AGCTGGCAGC GGCTGTCTCC AACTTCGGCT ATGACCTGTA 14640

CCGGGTGCGA TCCAGCATGA GCCCCACGAC CAACGTGCTC 14680

CTGTCTCCTC TCAGTGTGGC CACGGCCCTC TCGGCCCTCT 14720

CGCTGGGTGA GTGCTCAGAT GCAGGAAGCC CCAGGCAGAC 14760

CTGGAGAGGC CCCCTGTGGC CTCTGCGTAA ACGTGGCTGA 14800

GTTTATTGAC ATTTCAGTTC AGCGAGGGGT GAAGTAGCAC 14840

CAGGGGCCTG GCCTGGGGGT CCCAGCTGTG TAAGCAGGAG 14880

CTCAGGGGCT GCACACACAC GATTCCCCAG CTCCCCGAAA 14920

GGGGCTGGGC ACCACTGACA TGGCGCTTGG CCTCAGGGTT 14960

CGCTTATTGA CACAGTGACT TCAAGGCACA TTCTTGCATT 15000

CCTTAACCAA GCTGGTGCTA GCCTAGGTTC CTGGGATGTA 15040

ACTGCAAACA AGCAGGTGTG GGCTTGCCCT CACCGAGGAC 15080

ACAGCTGGGT TCACAGGGGA ACTAATACCA GCTCACTACA 15120

GAATAGTCTT TTTTTTTTNT TTTTTTNNNC TTTCTGAGAC 15160

GGAGTCTCGC TTTGTCNCCA AGGCTGGAGT GCAGTGGTGT 15200

GATCTCAGCT CACTGCAACC TCTGCCTCCC TGGTTCAAGG 15240

AATTCTCCTG CCTCAGCCTC CAGAGTAGCT GGGATTACAG 15280

GCACCTGCCA TCATGCCCAG CTAATTTTTG TATTTTTAGT 15320

AGAGACGGGG TTTCACCATG TTGCCTAGGC TGGTCTCAAA 15360

CTCCCGGGCT CAAGCGATCC ACCCGCCTTG GCCTCCCAAA 15400

GTGCTGGGAT TACAGGCGTG AGCCACCGCG CCTGGCCAGA 15440

ATAATCTTAA GGGCTATGAT GGGAGAAGTA CAGGGACTGG 15480

TACCTCTCAC TCCCTCACTC CCACCTTCCA GGCCTGATGC 15520

CTTTAACCTA CTTCAGGAAA ATCTCTAAGG ATGAAAATTC 15560

CTTGGCCACC TAGATTGTCT TGAAGATCAG CCTACTTGGG 15600

CTCTCAGCAG ACAAAAAAGA TGAGTATAGT GTCTGTGTTC 15640

TGGGAGGGGG CTTGATTTGG GGCCCTGGTG TGCAGTTATC 15680

AACGTCCACA TCCTTGTCTC TGGCAGGAGC GGAGCAGCGA 15720

ACAGAATCCA TCATTCACCG GGCTCTCTAC TATGACTTGA 15760

TCAGCAGCCC AGACATCCAT GGTACCTATA AGGAGCTCCT 15800

TGACACGGTC ACTGCCCCCC AGAAGAACCT CAAGAGTGCC 15840

TCCCGGATCG TCTTTGAGAA GAGTGAGTCG CCTTTGCAGC 15880

CCAAGTTGCC TGAGGCATGT GGGCTCCATG CTGCAGGCTG 15920

GGGGGGTCTT TTTTTTTTTT GGGGAAAGAC GGAGTCTCGC 15960

TCTGTTGCCC AGGTTGGAGT GAAGTGGCGT GATCTCGGTT 16000

CACTGAAACC CCCACCTCCC GGGTTCACAC CATCCTCCTG 16040

CCTCAGCCTC CCGAGTAGCT GGGACTGCAG GNGCCCAGCT 16080

AATCTTTNTT GTATTTTTAG CAGAGACGGG GTTTCACCGT 16120

GTTTGCCAGG ATAGTCTCGA TCTCCTGACC TGGTGTTCTG 16160

CCCGCCTCGA CCTCCCAAAG TGCTGGGATT ACAGGTGTGA 16200

GCCACCGCGC TCGGCCCGTT TCTAAACAAT AGATCATGTG 16240

TGCCCAGGCC TGGCCTGGCA CTGGTGTGGA GGAAGGGCCC 16280

GTGAGCCCAA AGAGGCTCAG AAAGAGGAAG TGGGCTGCAG 16320

GAGACGGTGG GAGGGGCAGG GAGGGCAGTG GCGCGATGTG 16360

GGGAAATCTG CTGCCCCCCT GGCCAGTGCC TGGGGATGCC 16400

AGCAGAAGTC CTGGCAAGTC ACAGGAAGAT GCTGGCTGGG 16440

AAGTCAGGGC CTGCTGAGCG CTAAACCAGA ACCCGAGCCT 16480

GGCAGGCTCT CAAAGACGGG ATGCTTGTCG TCGAGTCTCA 16520

TACGCTAACC TCTGCTCCGC CTCTTCTCAG AGCTGCGCAT 16560

AAAATCCAGC TTTGTGGCAC CTCTGGAAAA GTCATATGGG 16600

ACCAGGCCCA GAGTCCTGAC GGGCAACCCT CGCTTGGACC 16640

TGCAAGAGAT CAACAACTGG GTGCAGGCGC AGATGAAAGG 16680

GAAGCTCGCC AGGTCCACAA AGGAAATTCC CGATGAGATC 16720

AGCATTCTCC TTCTCGGTGT GGCGCACTTC AAGGGTGAGC 16760

GCGTCTCCAA TTCTTTTTCA TTTATTTTAC TGTATTTTAA 16800

CTAATTAATT AATTCGATGG AGTCTTACTC TGTAGCCCTA 16840

ACTGGAGTGC AGTGGTGCGA TCTCAGCTCA ATGCAACCTC 16880

CGCCTCCCAG GTTCAAGCAA TTCTTGTGCC TCAGCCTCCC 16920

GAGTAGCTGG GATTACAGGG ATGTACCACC ACTCCCGGCT 16960

AATTTTTTGT ATTTAATAGA CATGGGGTTT CACCATGTTG 17000

GCCAGGCTGG TCTCGAACTC CTGAGCTCAG GTGGTCTGCC 17040

CGCCTCAGCC TCCCAAAGTG CTAGGATTAC AAGCTTGAGC 17080

CACCACGCCC AGCCCTTTTT ATTTTTAAAT TAAGAGACAA 17120

GGTGTTGCCA TGATGCCCAG GCTGGTCTCG AACTCCTGGG 17160

CTCAAGTAAT CCTCCCACCT TGGCCTCCCA AAGTGCTGGG 17200

ATTACAGGCA TGAGCCACCG CGCCCGGCCC TTTTACATTT 17240

ATTTATTTAT TTTTTGAGAC AGAGTCTTGC TCTGTCACCC 17280

AGGCTGGAGT GCAGTGGCGC GATCTCGGCT CACTGCAAGC 17320

TCTGCCTTCC AGGTTCACAC CATTCTCCTG CCTCGACCTC 17360

CCGAGTAGCT GGGACTACAG GCGCCCGCCA CTGCGCCCTA 17400

CTAATTTTTT GTATTTTTAG TAGAGACGGG GTTTCACCGT 17440

GGTCTCGATC TCCTGACCTC GTGATCCACC CGCCTCAGCC 17480

TCCCAAAGTG CTGGGATTAC AGGCGTGAGC CACTGCGCCC 17520

GGCCCTTTTA CATTTATTTT TAAATTAAGA GACAGGGTGT 17560

CACTATGATG CCGAGGCTGG TCTCGAACTC CTGAGCTGAA 17600

GTGATCCTCC CACCTCGGCC TCCCAAAATG CTGGGATTAC 17640

CATGTCCAAC TTTCCACTTC TTGTTTGACC AAGGATGGAT 17680

GGCAGACATC AGAAGGGGCT TGGAAAGGGA GGTGTCAAAG 17720

ACCTTGCCCA GCATGGAGTC TGGGTCACAG CTGGGGGAGG 17760

ATCTGGGAAC TGTGCTTGCC TGAAGCTTAC CTGCTTGTCA 17800

TCAAATCCAA GGCAAGGCGT GAATGTCTAT AGAGTGAGAG 17840

ACTTGTGGAG ACAGAAGAGC AGAGAGGGAG GAAGAATGAA 17880

CACTGGGTCT GTTTGGGGCT TTCCCAGCTT TTGAGTCAGA 17920

CAAGATTTAT TTATTTATTT AAGATGGAGT CTCATTCTGT 17960

TGCCCAGGCT GGAGTGCAGT GGTGCCATCT TGGCTCACTA 18000

CAGCCTCCCC ACCTCCCAGG TTCAAGTGCT TCTCCTGCCT 18040

CAGCCTCCCG AGTAGTTGGG ATTACAGGCG CCCGCCACCA 18080

CACCCAGCTA ATTTTTGTAT TTTCAGTAGA GATGGGGTTT 18120

CGCCATGCTG GCCAGGCTGT TCTCGAAAAC TCCTGACCTC 18160

AGATGATCCA CCCGCCTCGG CCTCCCACAG TGCTGGGATT 18200

ACAGGCGTGA GCCACTGCGC TGGCCAAATC AGACAAGGTT 18240

TAAATCCCAG CTCTGCCTGT ACTAGCTGAG GAACTCTGCA 18280

CACATTTCAT AACCTTTCTG GGCCTACGTT CTCACCTTTA 18320

ACGTGAGGAT AATATATCTA CTTCATAGAC ACCTTTTTAT 18360

GTTGTCTCCA AGTTTTCTAA CAGCTCTAGT TCTGTACCCA 18400

AGACATGGCA GGTGGCCAAC GACATCCTTC TAGGCTGTGG 18440

TGATGTGTTT GGAGCTTGTT CCACGGGTCT TGTGTGGGGC 18480

CAGCCCTGTT CAGATAAGGC CTTGTGGGGT GGCCTGGGGT 18520

AGGGGGAGGG GTTGGGCAAA CTCTCCCTTA AAACGCTTTG 18560

TAACCATCTG AGGCACCAGC AAGAGCGGCC CCCGAGCCTG 18600

GACAAAATCC AAACGGCTTC CTACTTCAAG CACTGATGTC 18640

TAGTGAGTGA AGGAACAGCT CTGGGTCCAG GATATTATAG 18680

GTCACATTAA ACTAAAGGGG CTTGGCCATC AGCTGGCTTC 18720

CAGAGCGTCA GCCAGTTACT TCACCTCTTT GGCTTTGGCC 18760

TGTTTTCAGC TACAAGAGGA CTTAATCCAG AGGACCTCAG 18800

AGGTCCTTCC CAGCTCAGAC CTTCTTTGAC TGTCTCCCAG 18840

AGACACTGCT GTAGGAGTGC ACACCAGTTT ACTTTTCTTT 18880

CTTTTGTTTT TGAGATGGAG TTTCGCTCTT TTTGCCTAGG 18920

CTGGAGTGCT GTGGTGTGAT CTCAGCTCAC TGCAACCTCT 18960

GGCTCCCAGG TTCAAGTGAT TCTCCTGTCT CTGCCTCCCG 19000

AGTAGCTGGG ATTACAGACA CCCACCACTG CACCCGGCTA 19040

GTTTTTGTAT TTTCAGTAGA GATGGGGTTT CGCCATGCTG 19080

GCCAGGCTGT TCTCGAAAAC TCCTGACCTC AGATGATCCA 19120

TCCGCCTTGG CCTCCCAAAG TGCTGAGATT ACAGATGTGA 19160

GGCACCACAC CCGGCCATTT TTGTATTTTT AGTAGAGACG 19200

GGGTTTTGCC ATGTTGGCCA CGCTGGTCTC AAACTCCTGA 19240

CCTCAAGTGA TCTGCCCACC TTGGCCTCCT GAAGGGCTGG 19280

GACTACAGGC GTGAGTCACC GTGCCCGGCC ATTTTTGTAT 19320

TTTTAGGACA GCGTTTTTTC ATGTTGGCCA GGCTGGTCTC 19360

AAACTCCTGA CCTCAAGTGA TCCACCCACC CCGGCCTCCC 19400

AATATGCTGG GATTCCAGGT GTGAGTTACC ATGCCCGGCT 19440

ACCACTTTAC TTTTCCTGCA GGCTATCACA GAACGTGTAC 19480

AATCTAGACT CTAATCAACC AAATCAACGT CTTGCCATCG 19520

GAGTTTGCTG GTGAAGGGCA CTTGGGGTCC TGGAAATAAC 19560

TGTAGGCTCC AAGCCACACA CACTGAGATA GGCCTATTCC 19600

CTGAGGCCTC AGAGCCCCTG ACAGCTAAGC TCCCTTGAGT 19640

CGGGCAATTT TCAACAACGT GCTCTGGGGA CACAGCATGG 19680

CGCCACTGTC TTTCTGGTCT CCTGGGGCTC AGACTATGTC 19720

ATACACTTCT TTCCAGGGCA GTGGGTAACA AAGTTTGACT 19760

CCAGAAAGAC TTCCCTCGAG GATTTCTACT TGGATGAAGA 19800

GAGGACCGTG AGGGTCCCCA TGATGTCGGA CCCTAAGGCT 19840

GTTTTACGCT ATGGCTTGGA TTCAGATCTC AGCTGCAAGG 19880

TCTGTGGGGA TAGGGGCAGG GTGGGGGGTG GATGGAGGGA 19920

GAGGATAGAG AAGCAAAACA GGGTAGTGGG AATAAAATGA 19960

CCTTTGAGAT CCGACAGCTG TCTACATGTC GCCTGCTGTG 20000

TGACTTTGAG CAGGTTAATA ACATGTCTGA GCTTTCCTCC 20040

TCTTAAGATG GGGCAGGGGA TCGTTACCAA CACTTACCCT 20080

CCCAGGGTTT GTTGTAAGGA CGAATAAGGT AATAGGAAAT 20120

GGGCCCTCAG CACTGGGCAC CCACATGTTT GTTCTCTTGA 20160

GACTCCTATT TCTAGAATTT AAAGCCAAAC TTTGAAAAAT 20200

AATGACAAAC TCCAAATCGT TGGCATCTTT TTTTTTTTTT 20240

GAGACAGTCT CGCTCTGTCG GCCAGGCTGG AGTCCAGTGG 20280

CACGATCTCG GCTCACCACA ACCTCCGCCC CCGCTGGGTT 20320

AAAGCGATTC TCTTGCCTCA GCCTCCTGAG TAGCTGGGAT 20360

TACAGGCGTG TGCCTCCATG CCTGGCTAAT TTTATACAGA 20400

CGGGGTTTCT CCATGTTGGT CAGGCTGGTC TCAAACTCCC 20440

AAACTCAGGT GATCCGCCTG CCTCGGTCTC CCAAAACACA 20480

GGGGATTCCA GGCATGAGCC ACCACGCTTG GCCAATCGTT 20520

GGCATTCTAA GGCTTTCAGT GTACCTGACT TCTTTTAGTT 20560

CTAAGTCTGT AACTGTTAAC CTTTCTTGGG CCACGGCTAT 20600

CACACGGATC TCTCTGGGAA TCTGACGACA GTGCCTCAAA 20640

CCCGAGGGAG CACCGCCAGG TGTGCACACA CGTTTCTGTC 20680

AACGATTTCG GAGGACTCTT GGGATCCCTG AACACCATCT 20720

GTTCCATGGG ACCTTAGGTT AAGAGCCTCT GTTCAAAGGA 20760

GGCTTTTGCT CTTGGTGGGT GGATGGGGTG AAGTCTCCAA 20800

GCCCTCTTRC GGSCCCTTCG GTATTCCTAT NCCCCGGTTC 20840

TGCCCTGTCT TAGTCCAGTG CTCTCTATTT AACAAATGAG 20880

CAGTAAATGT ACACCGATGG ACTTTGGGAG ACAATAAAGA 20920

CCTGATATTC AATTCTAGCT CCTTAAACCA CAGGAGAACA 20960

TTCTTTCAGC AGACAACTTC AGTTGGTATT AGGCCAAGGT 21000

AAGAAAGGCC AACAGCATCC TTTTCTGAAG AAACCTCAGG 21040

AGATGGCTCT CTGCCAGAAA GCTATAACCT GGAAGGGGAA 21080

TTGTAAAATA GATGAGGGGC TGGATGAAGG ACGAGACCAG 21120

GGCCCCGTCA CGGGAGAGGG AAGGCAGCTC CTGGCTGTGT 21160

CTGTCCCCCG GCTTTTGGGC TCTGAAGGAC TAACCACATG 21200

CTTTCTCACT TGTCTCAGAT TGCCCAGCTG CCCTTGACCG 21240

GAAGCATGAG TATCATCTTC TTCCTGCCCC TGAAAGTGAC 21280

CCAGAATTTG ACCTTGATAG AGGAGAGCCT CACCTCCGAG 21320

TTCATTCATG ACATAGACCG AGAACTGAAG ACCGTGCAGG 21360

CGGTCCTCAC TGTCCCCAAG CTGAAGCTGA GTTACGAAGG 21400

CGAAGTCACC AAGTCCCTGC AGGAGATGAG TATGTCTGAA 21440

GACCCTTTCG CTCTTGGTGG GTGGATGGGG TGGGGCAGGG 21480

TCTTTGGGCC TTCCACTGTG CTAAGCAGAA CGCAAGGGCT 21520

CCACAGGCTT GTAGGGGGGC CGTGGATGAG TCCTTAATCC 21560

TCATCGTGCC AGAAGGGAAG GCTGAACTGC CTTCTCTCAT 21600

CAGACTCATT CCTCAGCCTC ACGAGCAGAC CTCCCTGACA 21640

GGCGCTCACA ACACTGCCTC TCAAGACGAG TCTGTCTGAC 21680

CTGTTTTCTC ATCTTGACCT AACTTGCTAA ATGCTCCTGG 21720

GCAAGTCACT CCACCCTCGG TCAGCTCAGA CCTCTTCAGG 21760

CCTCAGAGAA AGTCAACAGT GCTGCGCCAT CCCAGCTTGC 21800

TTGCAAAGGG ATCCCTTGGT TGGGGTGTTG GGGAAGGCAG 21840

GGTTTTAACG GAAATCTCTC TCCATCTCTA CAGAGCTGCA 21880

ATCCTTGTTT GATTCACCAG ACTTTAGCAA GATCACAGGC 21920

AAACCCATCA AGCTGACTCA AGGTGGAACA CCGGGCTGGC 21960

TTTGAGTGGA ACGAGGATGG GGCGGGAACC ACCCCCAGCC 22000

CAGGGCTGCA GCCTGCCCAC CTCACCTTCC CGCTGGACTA 22040

TCACCTTAAC CAGCCTTTCA TCTTCGTACT GAGGGACACA 22080

GACACAGGGG CCCTTCTCTT CATTGGCAAG ATTCTGGACC 22120

CCAGGGGCCC CTAATATCCC AGTTTAATAT TCCAATACCC 22160

TAGAAGAAAA CCCGAGGGAC AGCAGATTCC ACAGGACACG 22200

AAGGCTGCCC CTGTAAGGTT TCAATGCATA CAATAAAAGA 22240

GCTTTATCCC TAACTTCTGT TACTTCGTTC CTCCTCCTAT 22280

TTTGAGCTAT GCGAAATATC ATATGAAGAG AAACAGCTCT 22320

TGAGGAATTT GGTGGTCCTC TACTTCTAGC CTGGTTTTAT 22360

CTAAACACTG CAGGAAGTCA CCGTTCATAA GAACTCTTAG 22400

TTACCTGTGT TGGATAAGGC ACGGACAGCT TCTCTGCTCT 22440

GGGGGTATTT CTGTACTAGG ATCAGTGATC CTCCCGGGAG 22480

G 22481

CLAIMS

1. A method of enhancing neuron cell survival comprising:

treating a cell population comprising neurons with an effective amount of pigment epithelium - derived factor; and

enhancing neuronal cell survival in said population.

2. A method of inhibiting glial cell proliferation comprising:

treating a cell population in comprising glial cells with an effective amount of pigment epithelium derived factor; and

inhibiting glial cell proliferation in said population.

3. The method according to claim 1 wherein the neuronal cells are in a tissue cell culture.

4. The method according to claim 1 further comprising:

setting up a cell culture; and

treating said cell culture with an effective amount of PEDF.

5. The method according to claim 1, wherein the cells treated comprise a component of tissue being transplanted into a subject.

6. The method according to claim 6, wherein the cells are fetal brain cells.

7. The method according to claim 2, wherein the glial cells are part of a tumor growth.

8. The method according to claim 2, wherein glial cell growth inhibited is a gliosis.

9. Purified antibodies or antigen-binding fragments of said antibodies raised against a purified pigment epithelium-derived factor or an antigenic fragment thereof.

10. The isolated antibodies or antibody fragments of claim 9, wherein said antibodies are polyclonal.

11. The antibodies or antibody fragments of claim 9, wherein said antibodies are monoclonal.

12. The antibodies or antibody fragments of claim 9, wherein said antibodies are labeled with a detectable label.

13. A method of inhibiting pigment epithelium derived factor comprising:

treating cells or a population of cells with an effective amount of antibody or antigen binding fragments of said antibodies of claim 9; and

inhibiting pigment epithelium derived factor biological activity.

14. A method of determining levels of pigment epithelium - derived factor in a fluid, cellular or tissue sample, said method comprising:

A. contacting said sample with purified antibodies or antigen-binding fragments according to claim 9 under conditions in which an immune complex forms between said antibodies or antigen binding fragments and any pigment epithelium-derived factor present in said sample;

B. separating excess antibodies or antigen binding fragments and thereby from immune complexes; and

C. determining the level of immune complexes determining levels of pigment epithelium - derived factor.

FIG. 1

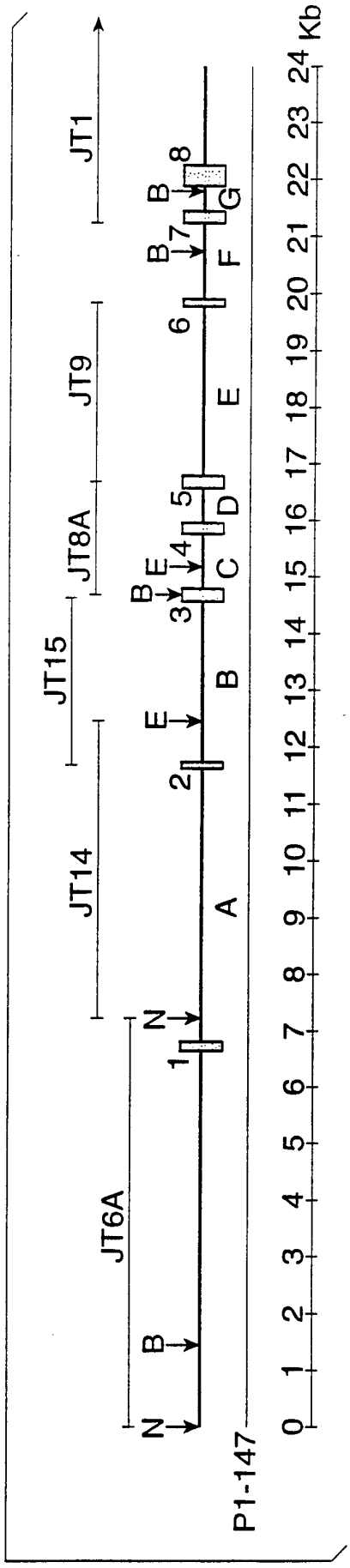

JT1
JT9
JT8A
JT15
JT14
JT6A
P1-147
A
B
C
D
E
F
G
1
2
3
4
5
6
7
8
N
B
E
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 Kb

FIG. 2B

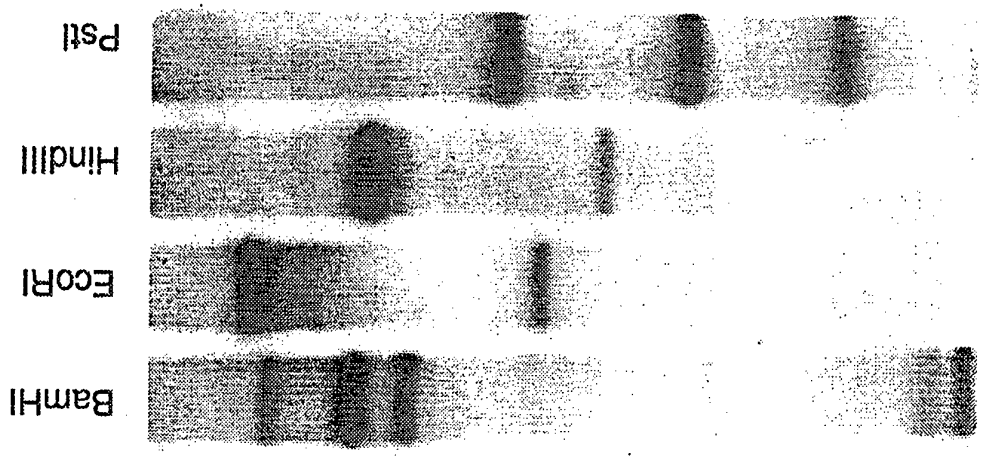


FIG. 2A

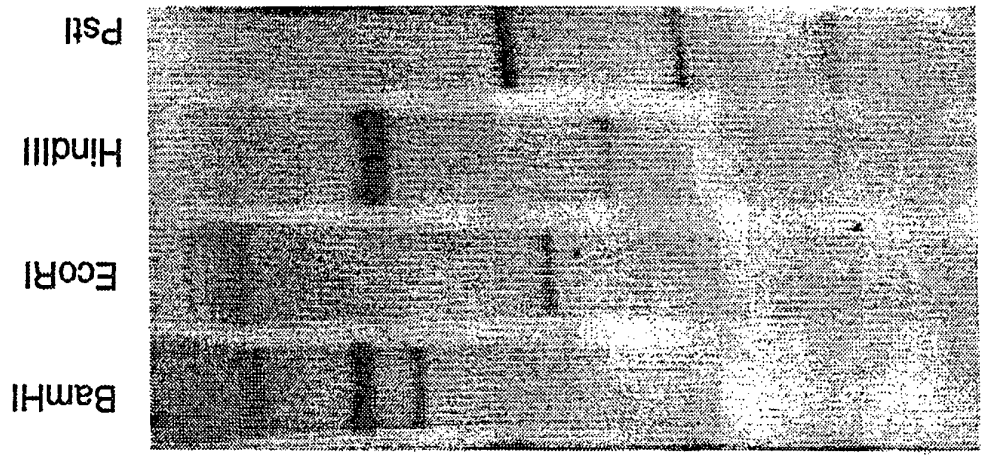


FIG. 3

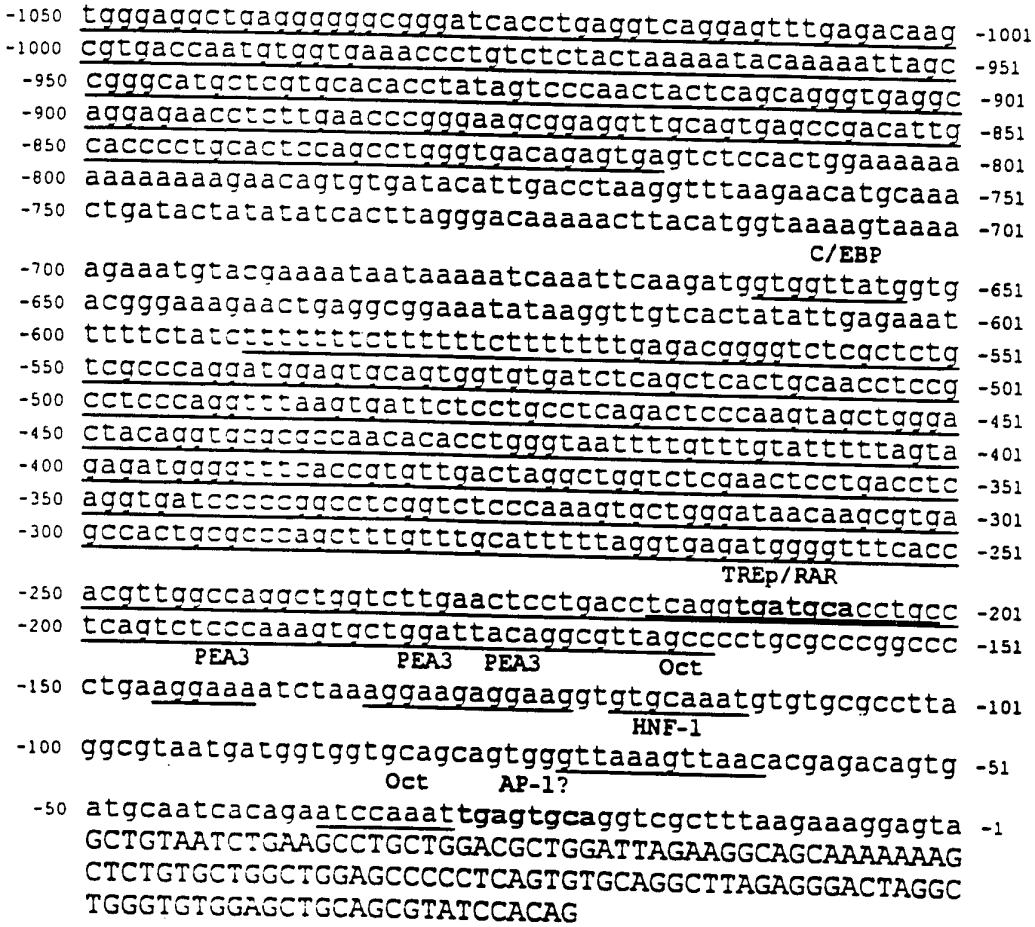

-1050 tgggaggctgaggggggcgggatcacctgaggtcaggagtttgagacaag -1001
-1000 cgtgaccaatgtggtgaaaccctgtctctactaaaaatacaaaaattagc -951
-950 cgggcatgctcgtgcacacctatagtcccaactactcagcagggtgaggc -901
-900 aggagaacctcttgaacccgggaagcggaggttgcagtgagccgacattg -851
-850 caccctgcactccagcctgggtgacagagtgagtctccactggaaaaaa -801
-800 aaaaaaaagaacagtgtgatacattgacctaaggtttaagaacatgcaaa -751
-750 ctgatactatatatcacttagggacaaaaacttacatggtaaaagtaaaa -701
C/EBP
-700 agaaatgtacgaaaataataaaaatcaaattcaagatggtggttatggtg -651
-650 acgggaaagaactgaggcggaaatataaggttgtcactatattgagaaat -601
-600 ttttctatctttttttctttttttctttttttgagacggggtctcgctctg -551
-550 tcgcccaggatggagtgcagtggtgtgatctcagctcactgcaacctccg -501
-500 cctcccaggtttaagtgattctcctgcctcagactcccaagtagctggga -451
-450 ctacaggtgcgcgccaacacacctgggtaattttgtttgtatttttagta -401
-400 gagatggggtttcaccgtgttgactaggctggtctcgaactcctgacctc -351
-350 aggtgatccccccggcctcggtctcccaaagtgctgggataacaagcgtga -301
-300 gccactgcgcccagctttgtttgcatttttaggtgagatggggtttcacc -251
TREp/RAR
-250 acgttggccaggctggtcttgaactcctgacctcaggtgatgcacctgcc -201
-200 tcagtctcccaaagtgctggattacaggcgttagccccctgcgcccggccc -151
PEA3 PEA3 PEA3 Oct
-150 ctgaaggaaaatctaaaggaagaggaaggtgtgcaaatgtgtgcgcctta -101
HNF-1
-100 ggcgtaatgatggtggtgcagcagtgggttaaagttaacacgagacagtg -51
Oct AP-1?
-50 atgcaatcacagaatccaaattgagtgcaggtcgctttaagaaaggagta -1
GCTGTAATCTGAAGCCTGCTGGACGCTGGATTAGAAGGCAGCAAAAAAAG
CTCTGTGCTGGCTGGAGCCCCCTCAGTGTGCAGGCTTAGAGGGACTAGGC
TGGGTGTGGAGCTGCAGCGTATCCACAG

FIG. 4A

FIG. 4B

FIG. 4C

FIG. 5A

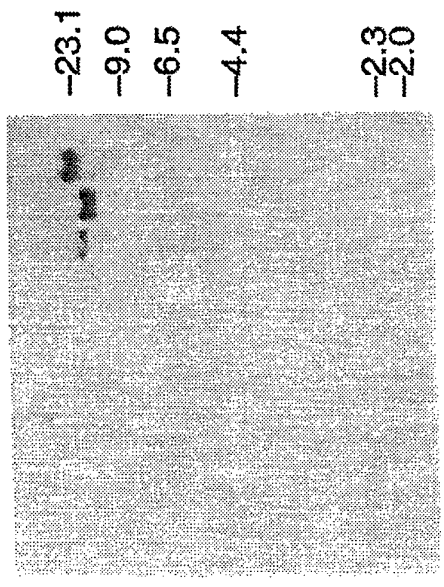


FIG. 5B

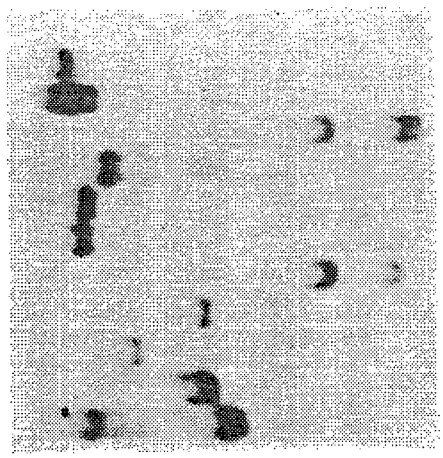


FIG. 5C

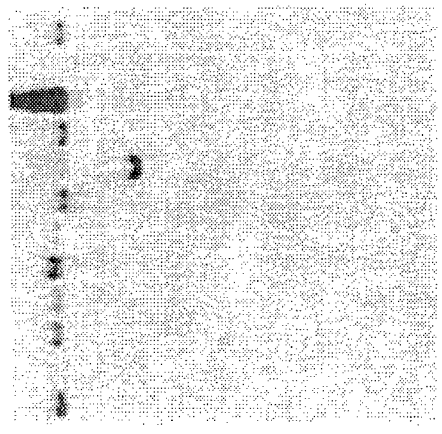


FIG. 6A

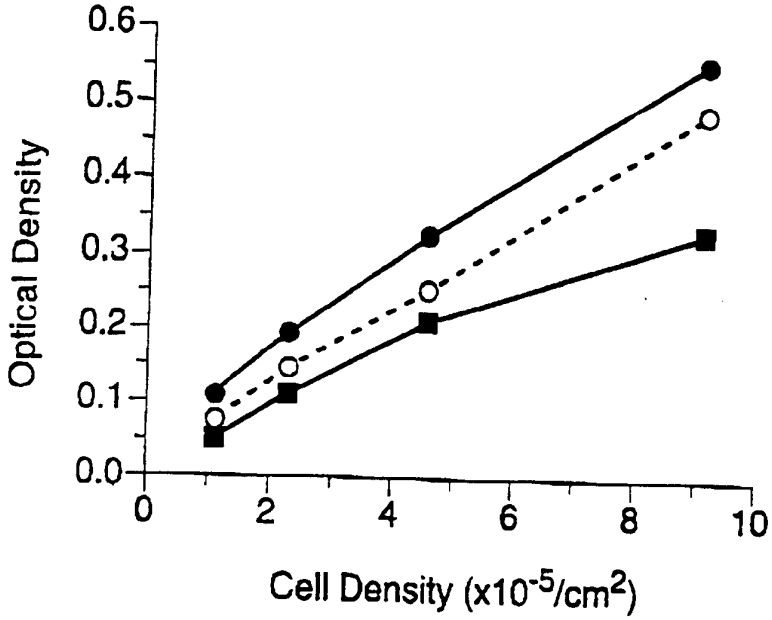

Optical Density
0.6
0.5
0.4
0.3
0.2
0.1
0.0
0
2
4
6
8
10
Cell Density (x10-5/cm2)

FIG. 6B

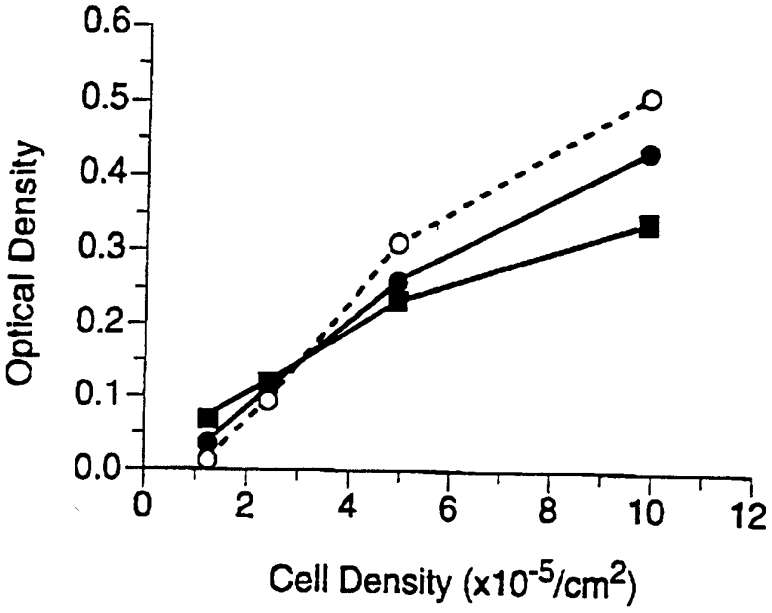

Optical Density
0.6
0.5
0.4
0.3
0.2
0.1
0.0
0
2
4
6
8
10
12
Cell Density (x10-5/cm2)

FIG. 7

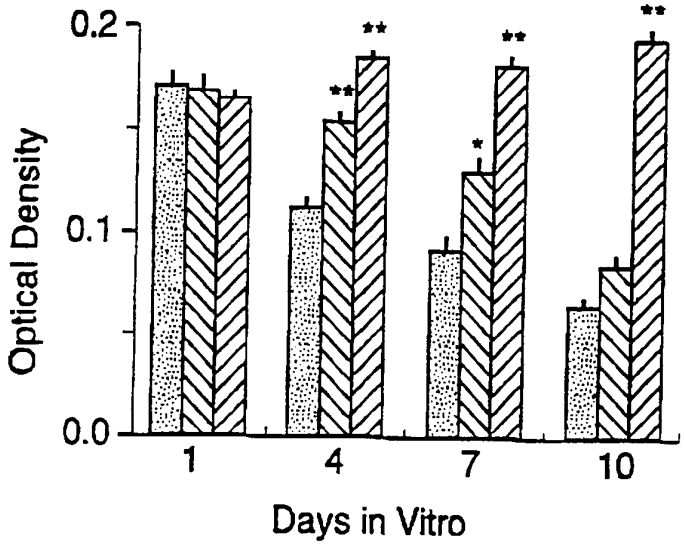

0.2
0.1
0.0
Optical Density
1
4
7
10
Days in Vitro
**
**
*
**
**

FIG. 8

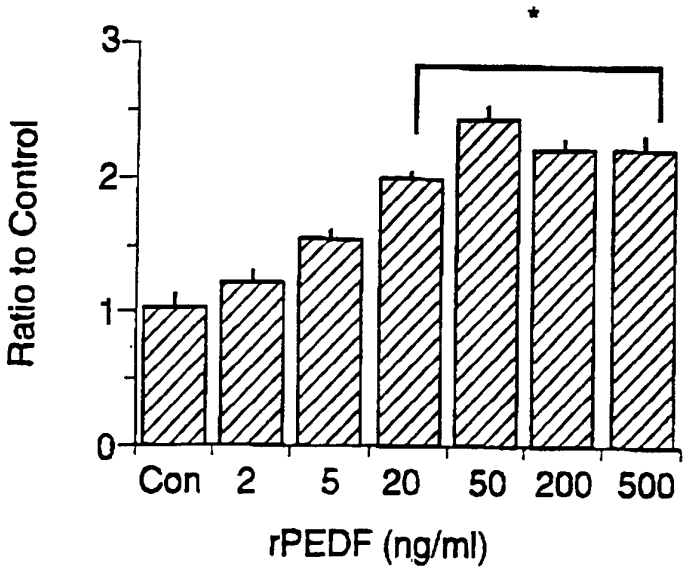

3
2
1
0
Ratio to Control
Con
2
5
20
50
200
500
rPEDF (ng/ml)
*

FIG. 9

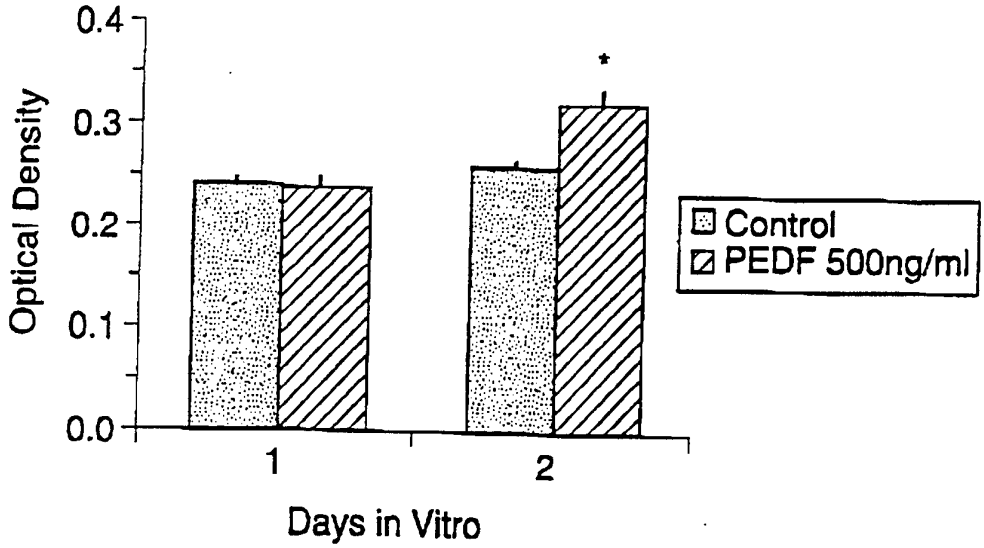

0.4
0.3
0.2
0.1
0.0
Optical Density
*
1
2
Days in Vitro
Control
PEDF 500ng/ml

FIG. 10

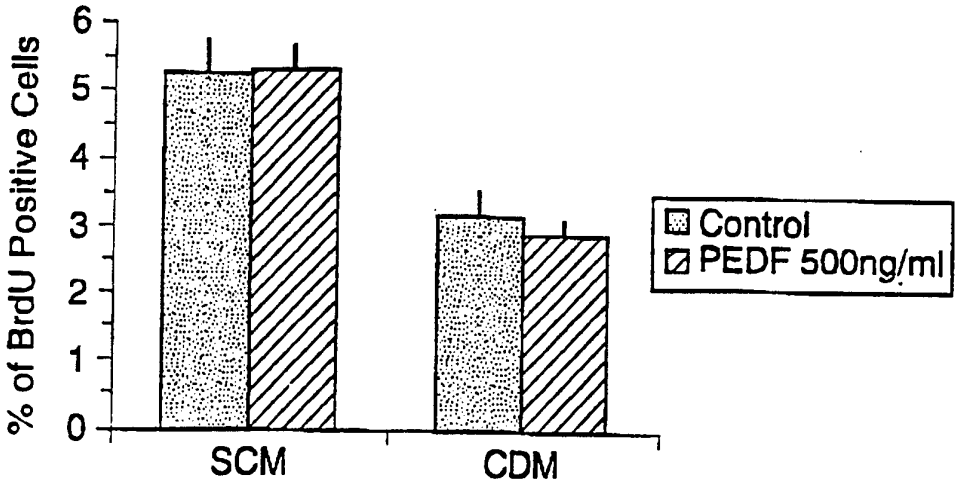

6
5
4
3
2
1
0
% of BrdU Positive Cells
SCM
CDM
Control
PEDF 500ng/ml

FIG. 11

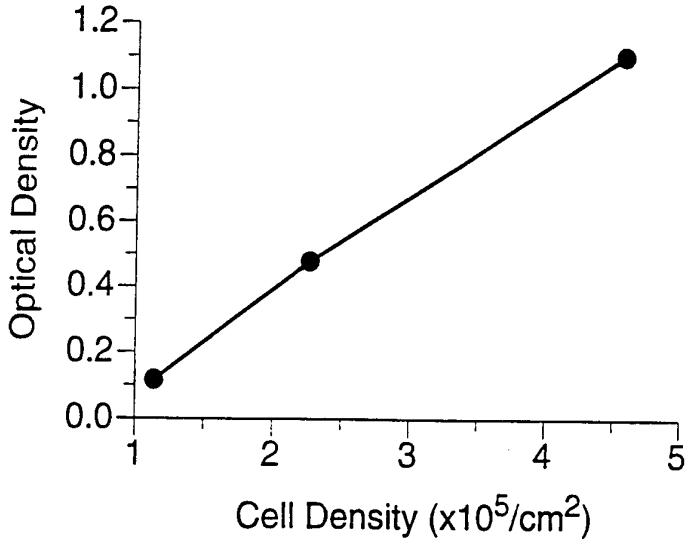

1.2
1.0
0.8
0.6
0.4
0.2
0.0
Optical Density
1
2
3
4
5
Cell Density (x10^5/cm^2)

FIG. 12

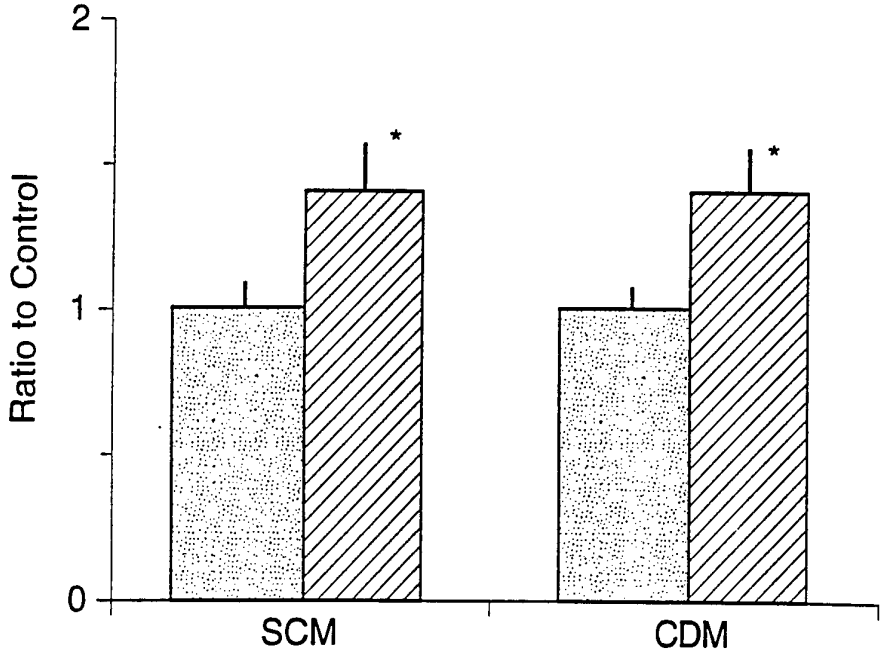

2
1
0
Ratio to Control
*
*
SCM
CDM

FIG. 13

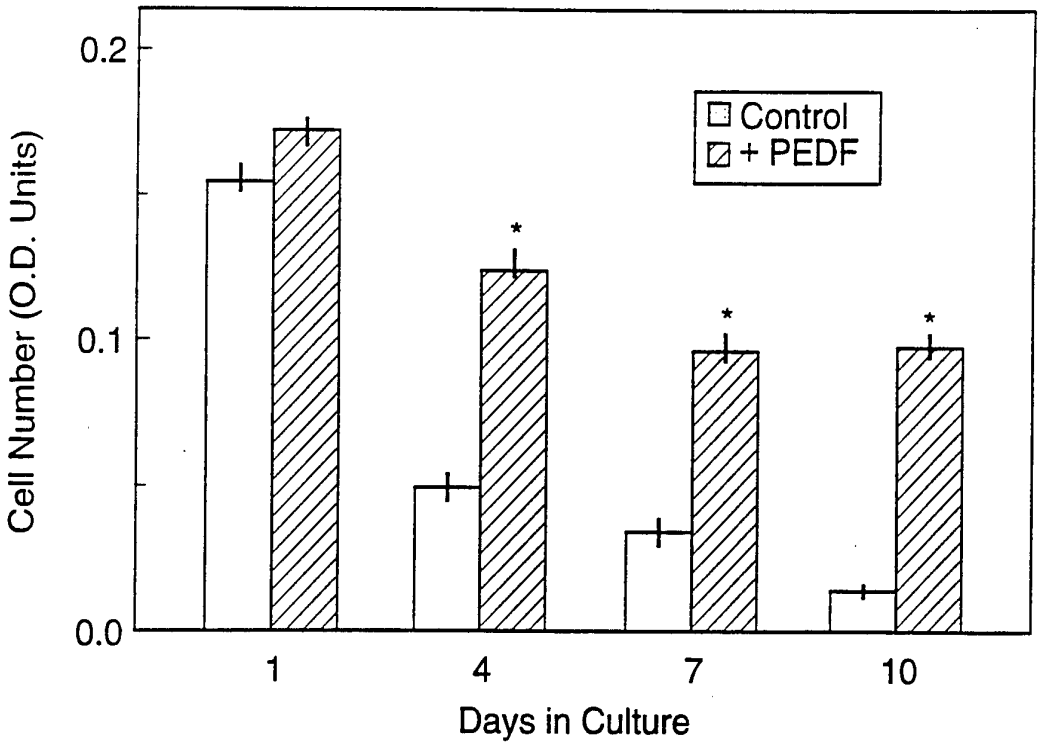

0.2
0.1
0.0
Cell Number (O.D. Units)
Control
+ PEDF
*
*
*
1
4
7
10
Days in Culture

FIG. 14

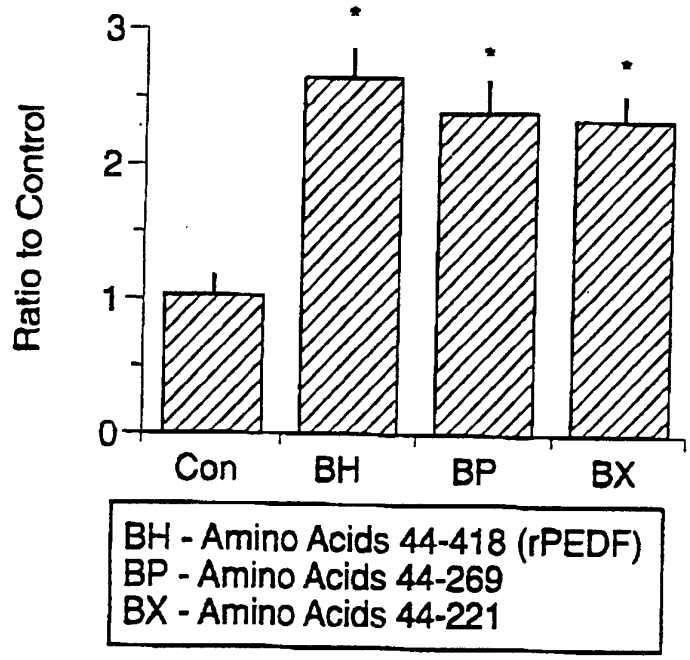

Ratio to Control
3
2
1
0
Con
BH
BP
BX
BH - Amino Acids 44-418 (rPEDF)
BP - Amino Acids 44-269
BX - Amino Acids 44-221

FIG. 15

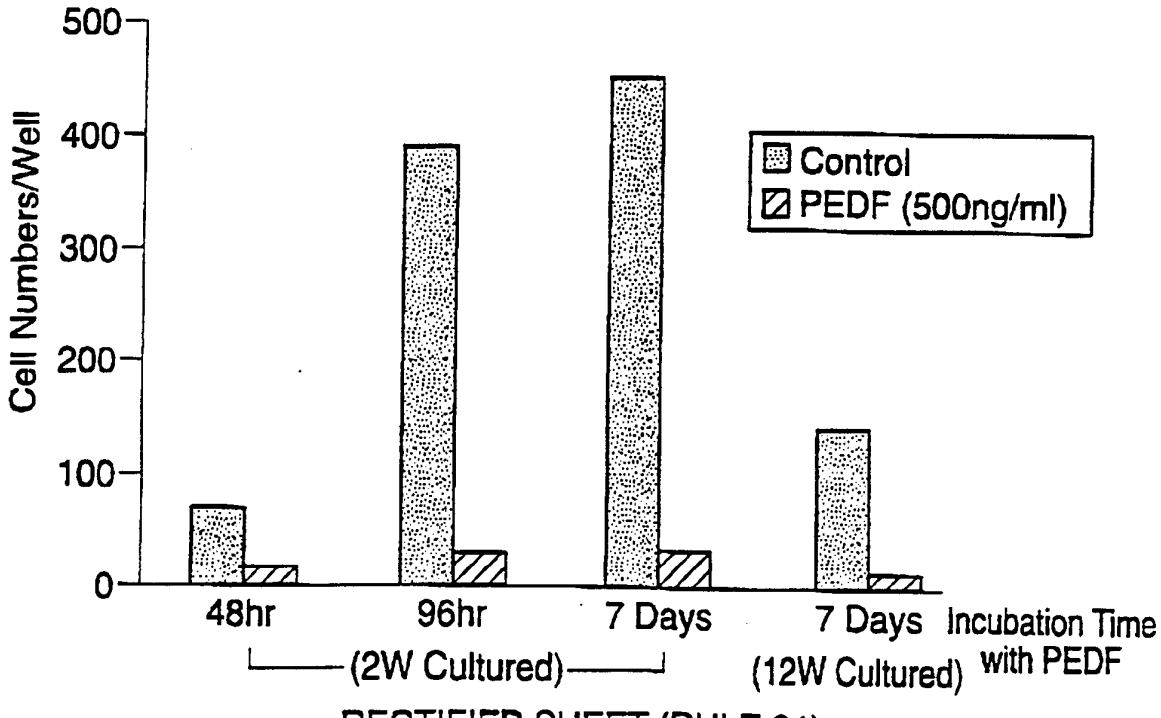

Cell Numbers/Well
500
400
300
200
100
0
Control
PEDF (500ng/ml)
48hr
96hr
7 Days
7 Days
Incubation Time with PEDF
(2W Cultured)
(12W Cultured)

FIG. 16

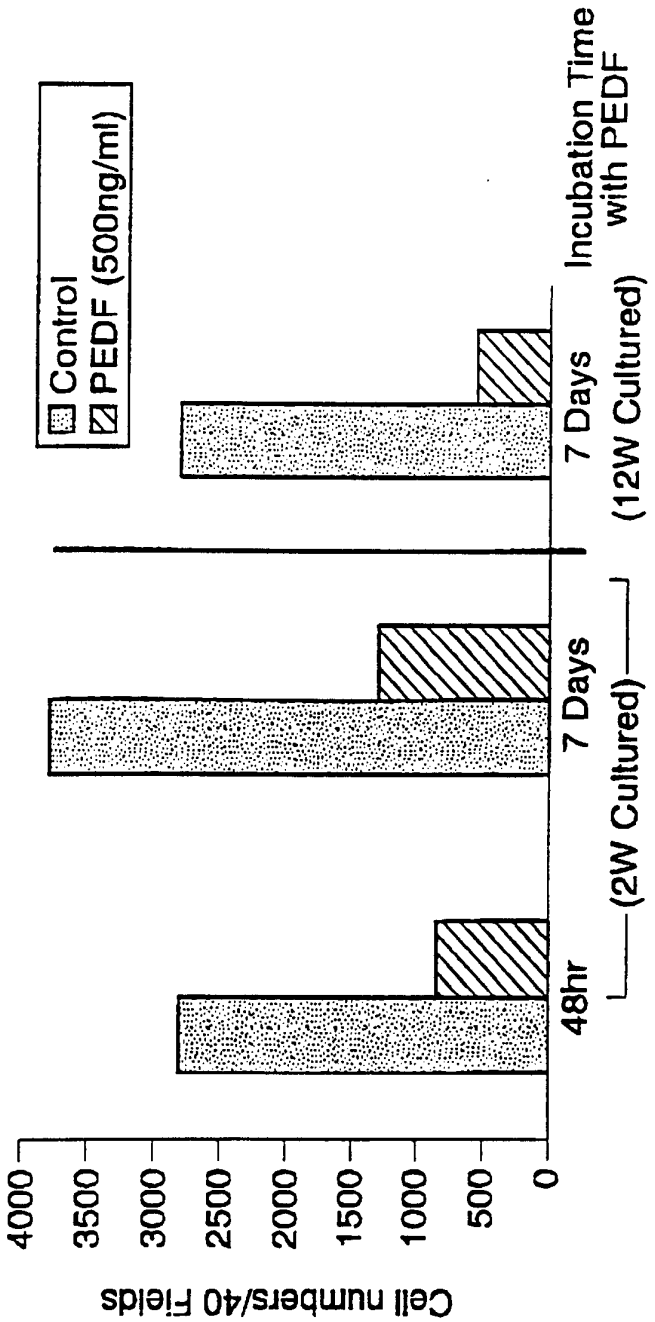
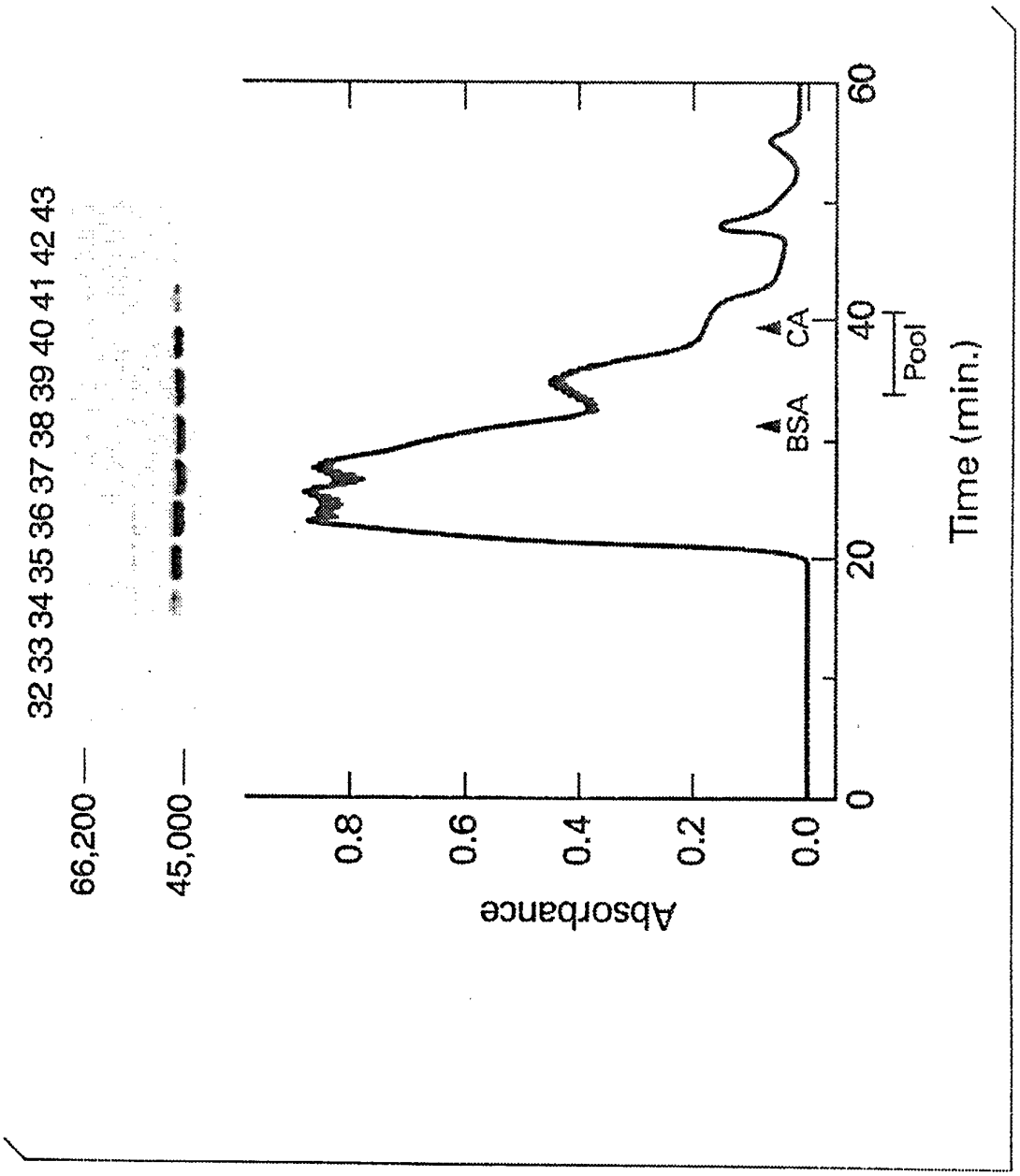

Control
PEDF (500ng/ml)
Cell numbers/40 Fields
4000
3500
3000
2500
2000
1500
1000
500
0
48hr
7 Days
(2W Cultured)
7 Days
(12W Cultured)
Incubation Time with PEDF

FIG. 17A


66,200
45,000
32 33 34 35 36 37 38 39 40 41 42 43
Absorbance
0.0
0.2
0.4
0.6
0.8
0
20
40
60
Time (min.)
BSA
CA
Pool

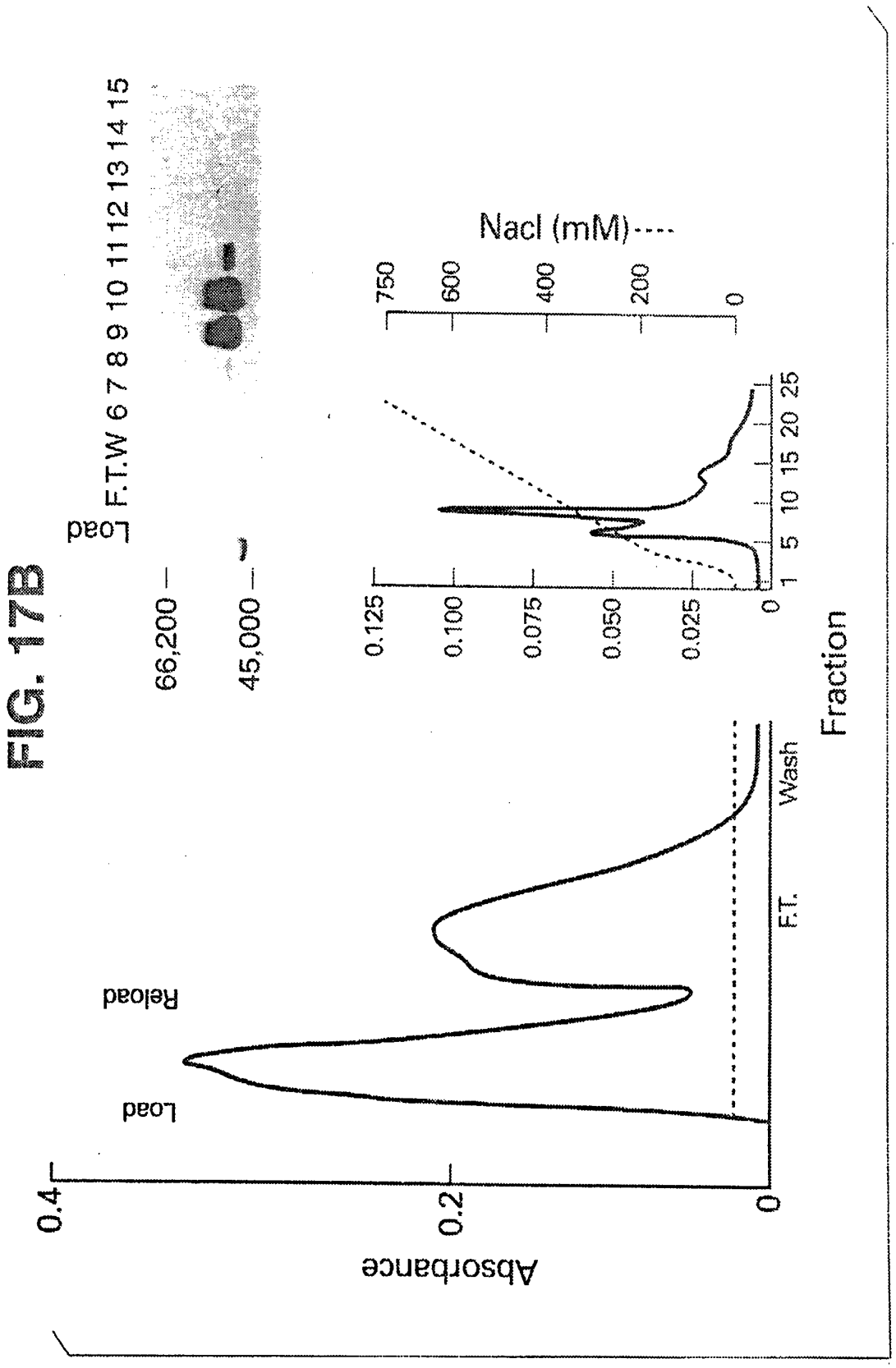

Load F.T.W 6 7 8 9 10 11 12 13 14 15
66,200
45,000
Nacl (mM)
750
600
400
200
0
0.125
0.100
0.075
0.050
0.025
0
1 5 10 15 20 25
Fraction
Absorbance
0.4
0.2
0
Load
Reload
F.T.
Wash

FIG. 17B

FIG. 18A

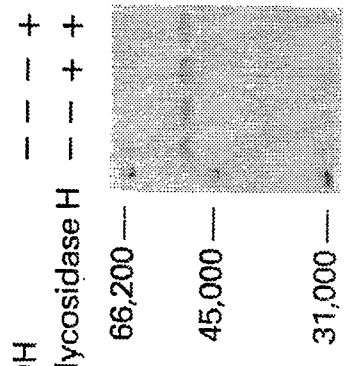


FIG. 18B

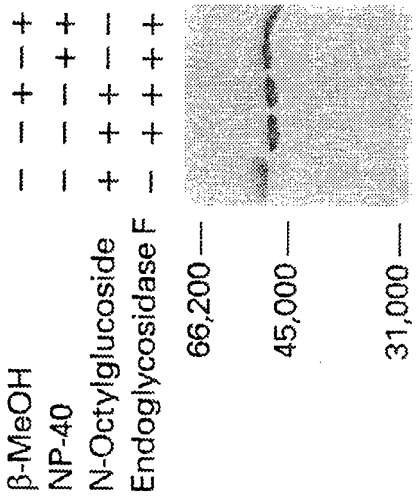


FIG. 19B

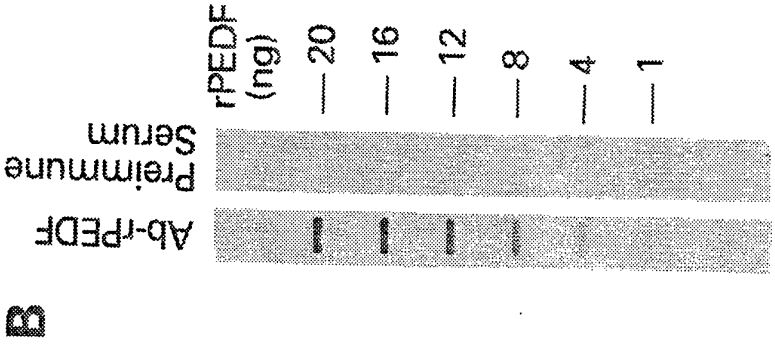


FIG. 19A

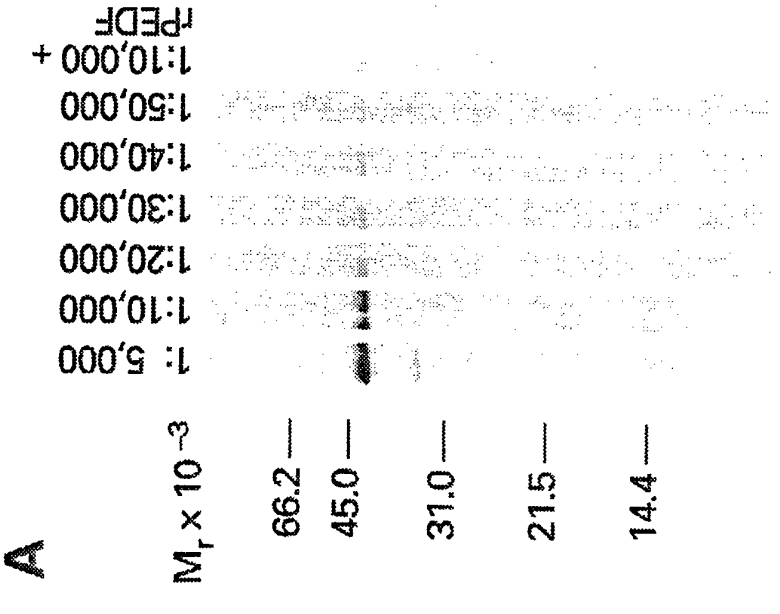


FIG. 20

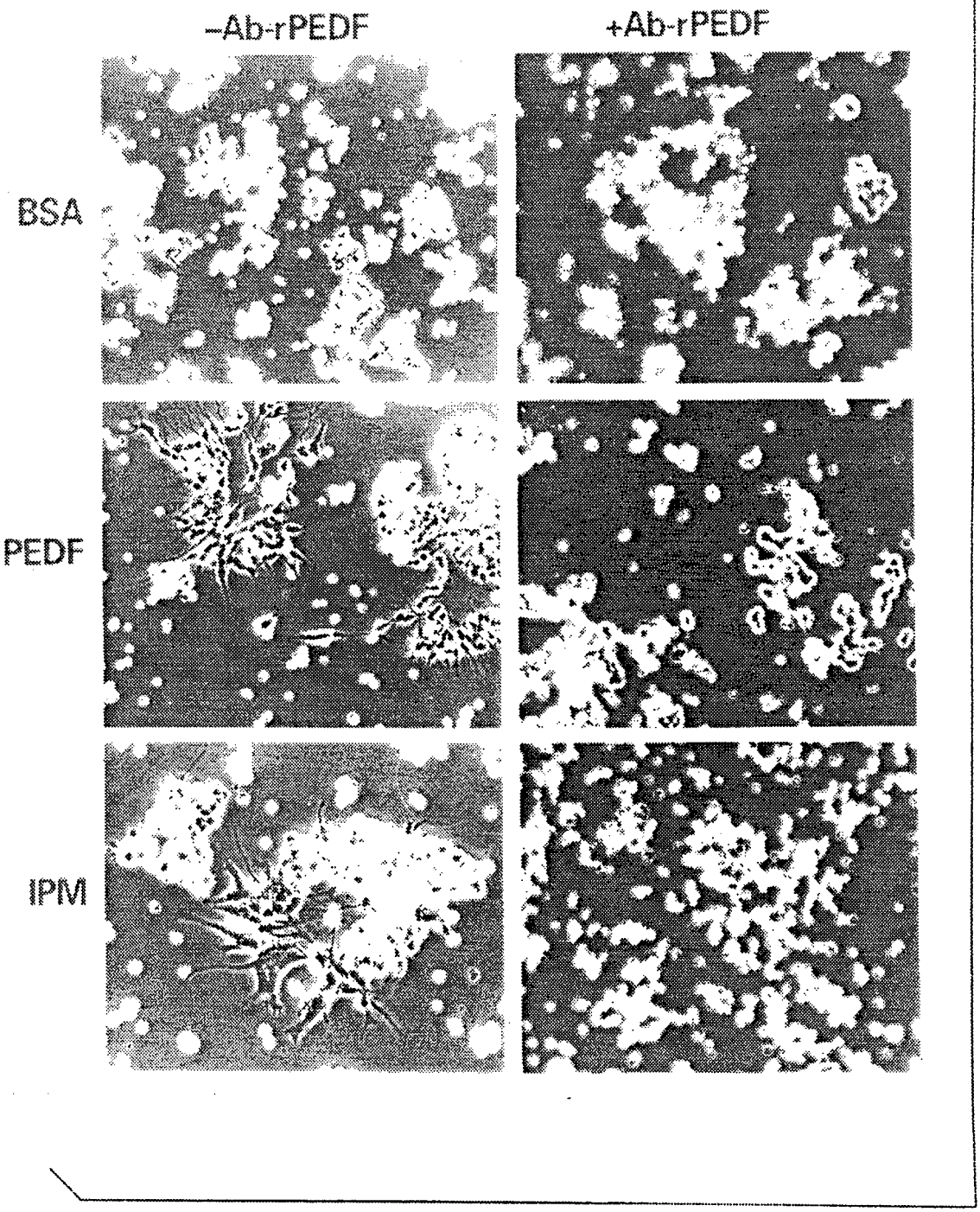


FIG. 21A

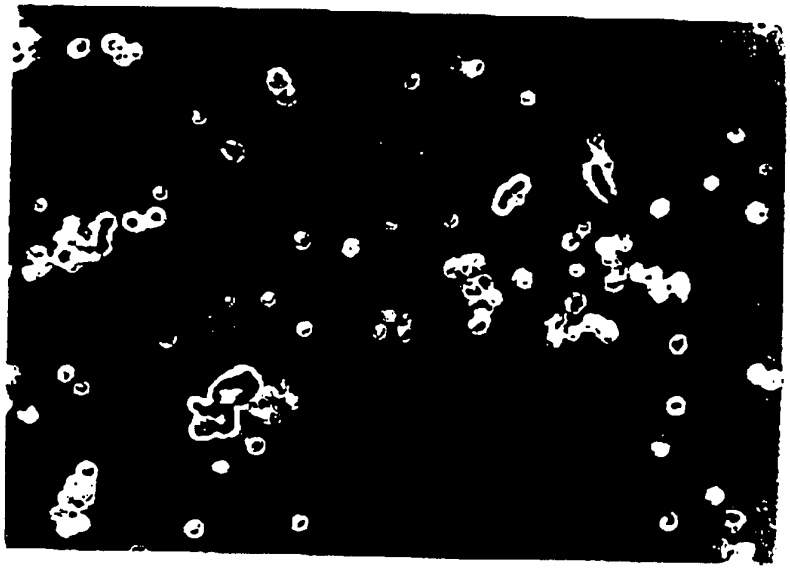


FIG. 21B

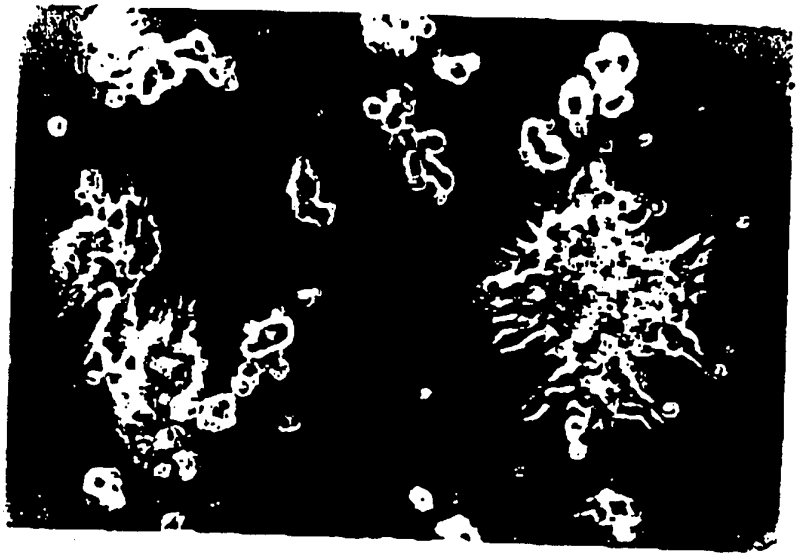


FIG. 22A

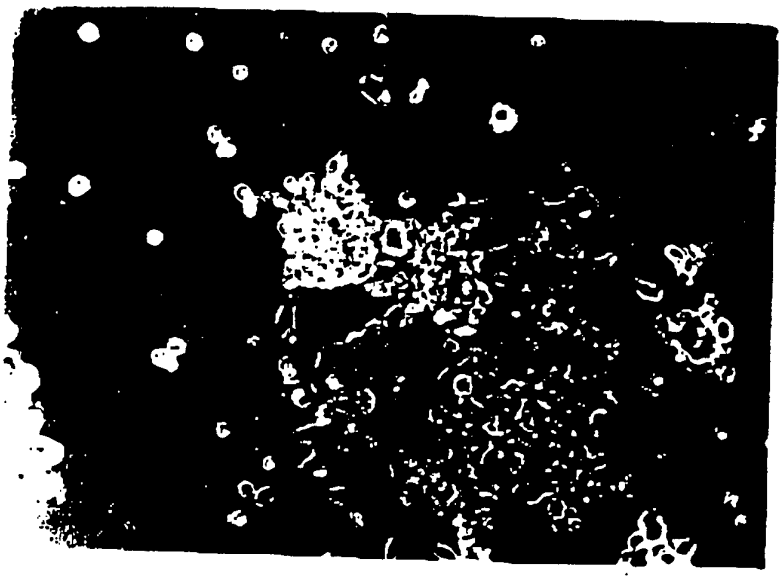


FIG. 22B

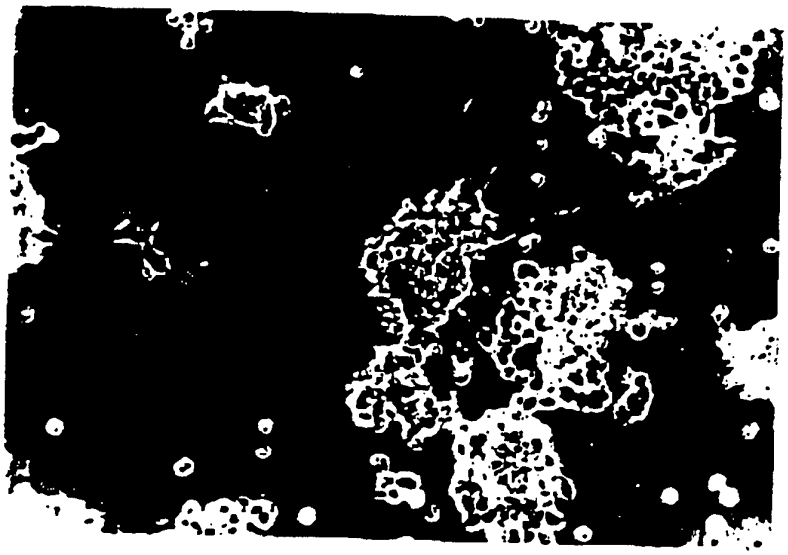
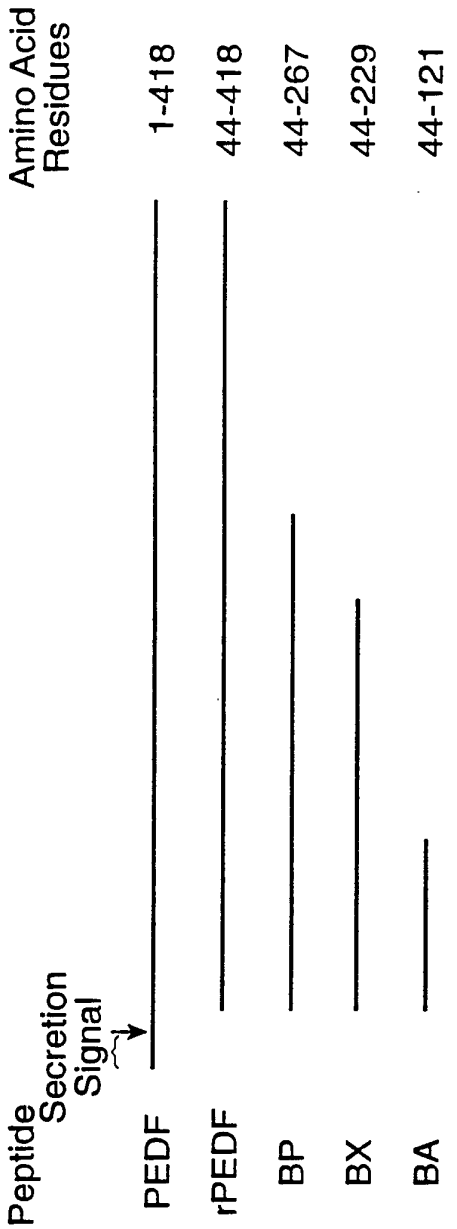


FIG. 23


Amino Acid Residues
Peptide
Secretion Signal
PEDF
rPEDF
BP
BX
BA
1-418
44-418
44-267
44-229
44-121

INTERNATIONAL SEARCH REPORT

Internati Application No
PCT/US 95/07201

A. CLASSIFICATION OF SUBJECT MATTER
IPC 6 A61K38/57 C07K16/38 G01N33/53 //C07K14/81

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 6 C07K A61K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	THE FASEB JOURNAL, vol. 8, no. 7, 19 April 1994 page A1302 BECERRA ET AL 'PIGMENT EPITHELIUM-DERIVED FACTOR:CHARACTERIZATION USING A HIGHLY SPECIFIC POLYCLONAL ANTIBODY' see abstract 252 ---	9-14
X	WO,A,93 24529 (UNIV SOUTHERN CALIFORNIA) 9 December 1993 see page 5, line 2 - line 30 see page 20, line 19 - page 22, line 26 --- -/--	1,3-6

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
3 October 1995	22.11.95

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax (+31-70) 340-3016	Sitch, W

Form PCT/ISA/210 (second sheet) (July 1992)

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	THE JOURNAL OF BIOLOGICAL CHEMISTRY, vol. 268, no. 31, 5 November 1993 pages 23148-23156, BECERRA ET AL 'OVEREXPRESSION OF FETAL HUMAN PIGMENT EPITHELIUM-DERIVED FACTOR IN ESCHERICHIA COLI.A FUNCTIONALLY ACTIVE NEUROTROPHIC FACTOR' see page 23148,abstract ---	
A	PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES,USA, vol. 90, February 1992 pages 1526-1530, STEELE ET AL 'PIGMENT EPITHELIUM-DERIVED FACTOR:NEUROTROPHIC ACTIVITY AND IDENTIFICATION AS A MEMBER OF THE SERINE PROTEASE INHIBITOR GENE FAMILY' see page 1526,abstract ---	
A	DATABASE CHEMICAL ABSTRACTS FILE SERVER STN KARLSRUHE ABSTRACT NO.117:45182, GAUR ET AL 'RPE CONDITIONED MEDIUM STIMULATES PHOTORECEPTOR CELL SURVIVAL,NEURITE OUTGROWTH AND DIFFERENTIATION IN VITRO' & EXP.EYE.RES. (1992) 54 (5),645-59 see abstract ---	
A	DATABASE CHEMICAL ABSTRACTS FILE SERVER STN KARLSRUHE ABSTRACT NO.118:188996, KLAIDMAN ET AL 'EFFECTS OF MEDIUM CONDITIONED BY RETINAL PIGMENTED EPITHELIAL CELLS ON NEUROTRANSMITTER PHENOTYPE IN RETINOBLASTOMA CELLS' & CANCER LETT. (SHANNON,IREL.) (1993) 68 (2-3), 207-13 see abstract ---	
P,X	SOCIETY FOR NEUROSCIENCE ABSTRACTS, vol. 20, no. 1-2, November 1994 page 873 SUGITA ET AL 'EFFECTS OF PIGMENT EPITHELIUM-DERIVED FACTOR (PEDF) ON ASTROCYTES AND MICROGLIA IN CULTURE' see abstract 365.7 -----	1-8

2

INTERNATIONAL SEARCH REPORT

Inte ional application No.
PCT/US95/07201

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. [X] Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

Remark: Although claims 1,2,5-8 and 13 partially, in so far as they relate to an in vivo method, are directed to a method of treatment of the human/animal body the search has been carried out and based on the alleged effects of the compound/composition.

2. [ ] Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

3. [ ] Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. [ ] As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.

2. [ ] As all searchable claims could be searches without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. [ ] As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

4. [ ] No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

[ ] The additional search fees were accompanied by the applicant's protest.

[ ] No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT

Information on patent family members

Internati Application No
PCT/US 95/07201

Patent document cited in search report	Publication date	Patent family member(s)		Publication date
WO-A-9324529	09-12-93	AU-B-	4406993	30-12-93
		CA-A-	2137377	09-12-93
		EP-A-	0662087	12-07-95