(19)日本国特許庁(JP)　(12)特　許　公　報(B2)　(11)特許番号

特許第3646191号

(P3646191)

(45)発行日　平成17年5月11日(2005.5.11)　(24)登録日　平成17年2月18日(2005.2.18)

(51)Int.Cl.[7]	F I		
C12N 15/09	C12N	15/00	ZNAA
A61K 38/00	C07H	21/04	B
C07H 21/04	A61K	37/02	
// A61K 48/00	A61K	48/00	
C12P 21/02	C12P	21/02	C

請求項の数 7　(全 156 頁)　最終頁に続く

(21)出願番号　特願平9-69163
(22)出願日　平成9年3月5日(1997.3.5)
(65)公開番号　特開平9-308492
(43)公開日　平成9年12月2日(1997.12.2)
審査請求日　平成13年6月7日(2001.6.7)
(31)優先権主張番号　特願平8-63410
(32)優先日　平成8年3月19日(1996.3.19)
(33)優先権主張国　日本国(JP)

(73)特許権者　000206956
大塚製薬株式会社
東京都千代田区神田司町２丁目９番地
(74)代理人　100065215
弁理士　三枝　英二
(74)代理人　100076510
弁理士　掛樋　悠路
(74)代理人　100086427
弁理士　小原　健志
(74)代理人　100090066
弁理士　中川　博司
(74)代理人　100094101
弁理士　舘　泰光
(72)発明者　藤原　力
徳島県鳴門市鳴門町高島字中島１６１－８
最終頁に続く

(54)【発明の名称】ヒト遺伝子

(57)【特許請求の範囲】

【請求項１】

以下の(a)または(b)の蛋白質をコードする遺伝子：

(a) 配列番号：37で示されるアミノ酸配列からなる蛋白質、

(b) 配列番号：37で示されるアミノ酸配列において一部のアミノ酸乃至アミノ酸配列を置換、欠失もしくは付加したアミノ酸配列からなり且つ脳神経成長活性および／または神経再生活性を有する蛋白質。

【請求項２】

以下の(a)または(b)のDNAからなる遺伝子：

(a) 配列番号：38で示される塩基配列からなるDNAおよびこれと相補的な配列のDNA

（ｂ）　配列番号：38で示される塩基配列とストリンジェントな条件でハイブリダイズし且つ脳神経成長活性および／または神経再生活性を有する蛋白質をコードするDNA。

【請求項３】

配列番号：37で示されるアミノ酸配列からなる蛋白質をコードする請求項１に記載の遺伝子。

【請求項４】

配列番号：38で示される塩基配列からなるDNAである請求項２に記載の遺伝子。

【請求項５】

配列番号：39で示される塩基配列である請求項２に記載の遺伝子。

【請求項６】

請求項1または3に記載のアミノ酸配列からなる脳神経成長活性および／または神経再生活性を有する組換え体ポリペプチド。

【請求項７】

請求項2または4に記載のDNAを含むベクター。

【発明の詳細な説明】

【０００１】

【発明の属する技術分野】

本発明は、ヒトの疾患の予防、診断及び治療の指針として有用な遺伝子、より詳しくはラット、マウス、酵母、線虫やヒト遺伝子などと類似性を有する新規なヒト遺伝子に関し、該遺伝子のｃＤＮＡ解析、ｃＤＮＡ染色体へのマッピング及びｃＤＮＡの機能解析により、該遺伝子を用いた遺伝子診断並びに新しい治療法の開発に利用可能な新規なヒト遺伝子に関する。

【０００２】

【従来の技術】

生物の遺伝情報は、細胞の核内に存在するＡ、Ｃ、Ｇ及びＴの４種の塩基の列び(ＤＮＡ)として蓄積され、この遺伝情報は個々の生物の系統維持と個体発生のために保存されている。

【０００３】

ヒトの場合、その塩基数は約３０億(3×10^9)といわれ、その中に５－１０万の遺伝子があると推測されている。これらの遺伝情報は遺伝子(ＤＮＡ)からｍＲＮＡが転写され、次に蛋白質に翻訳されるという流れに沿い、調節蛋白質、構造蛋白質あるいは酵素が作られ、生命現象を維持している。上記遺伝子から蛋白質翻訳までの流れの異常は、細胞の増殖・分化などの生命維持システムの異常を惹起し、各種疾患の原因となるとされている。

【０００４】

これまでの遺伝子解析の結果から、インスリン受容体やＬＤＬ受容体などの各種受容体や細胞の増殖・分化に関わるもの、プロテアーゼ、ＡＴＰａｓｅ、スーパーオキシドディスムターゼのような代謝酵素などの医薬品開発にとって有用な素材となると思われる遺伝子が多く見つかっている。

【０００５】

しかしながら、ヒト遺伝子の解析とそれら解析された遺伝子の機能及び解析遺伝子と各種疾患との係わりについての研究は、まだ始まったばかりであり、不明な点が多く、更なる新しい遺伝子の解析、それらの遺伝子の機能の解析、解析された遺伝子と疾患の係わりの研究、惹いては解析された遺伝子の利用による遺伝子診断、該遺伝子の医薬用途への応用研究などが当業界で望まれている。

【０００６】

【発明が解決しようとする課題】

上記新たなヒト遺伝子が提供できれば、各細胞での発現レベルやその構造及び機能を解析でき、またその発現物の解析などにより、これらの関与する疾患、例えば遺伝子病、癌などの病態解明や診断、治療などが可能となると考えられ、本発明は、かかる新たなヒトの遺伝子を提供することを目的とする。

【０００７】

本発明者らは、上記目的より鋭意研究を重ねた結果、以下の知見を得、ここに本発明を完成するに至った。

【０００８】

即ち、本発明者らは、ヒト胎児脳、成人血管、胎盤などの各種組織より抽出したｍＲＮＡよりｃＤＮＡを合成し、これをベクターに組込んでライブラリーを構築し、該ライブラリーでトランスフォームした大腸菌コロニーを寒天培地上に形成させ、該コロニーをランダムにピックアップして９６ウェルマイクロプレートに移し、ヒト遺伝子を含む多数の大腸菌クローンを登録した。

【０００９】

登録した各クローンは、少量培養後、ＤＮＡを抽出精製し、抽出したｃＤＮＡを鋳型としてデオキシターミネーター法により４種の塩基特異的に停止する伸長反応を行ない、自動ＤＮＡシークエンサーにより、遺伝子の５’末端から約４００塩基配列を決定した。かくして得られた塩基配列情報について、公知のバクテリアから酵母、線虫、マウス及びヒトなどの動植物種に類似性を有する新規なファミリー遺伝子を検索した。

【００１０】

上記ｃＤＮＡ解析方法については、本発明者のひとりである藤原らによって細述されている(藤原　力，細胞工学，14，645-654(1995))。

【００１１】

このグループの中には新しいレセプター、ＤＮＡ結合ドメインを有する転写調節因子やシグナル伝達系因子、代謝酵素などがあり、遺伝子解析によって得られた本発明の新規な遺伝子と類似性ある遺伝子との相同性から、およそのその遺伝子産物、即ち蛋白質がどのような機能を有するか類推することが可能である。更に、その候補遺伝子を発現ベクターに組込み、リコンビナントを作製し、酵素活性や結合活性などの機能を調べることができる。

【００１２】

【課題を解決するための手段】

本発明によれば、配列番号：１、：４、：７、：１０、：１３、：１６、：１９、：２２、：２５、：２８、：３１、：３４、：３７及び：４０で示されるアミノ酸配列をコードする塩基配列を含むことを特徴とする新規なヒト遺伝子、前記各アミノ酸をコードする配列番号：２、：５、：８、：１１、：１４、：１７、：２０、：２３、：２６、：２９、：３２、：３５、：３８及び：４１で示される塩基配列を含むことを特徴とするヒト遺伝子、並びに配列番号：３、：６、：９、：１２、：１５、：１８、：２１、：２４、：２７、：３０、：３３、：３６、：３９及び：４２で示される塩基配列であることを特徴とする新規なヒト遺伝子が提供される。

【００１３】

以下、本明細書におけるアミノ酸、ペプチド、塩基配列、核酸などの略号による表示は、ＩＵＰＡＣ、ＩＵＢの規定、「塩基配列又はアミノ酸配列を含む明細書などの作成のためのガイドライン」（特許庁編）及び当該分野における慣用記号に従うものとする。

【００１４】

かかる本発明遺伝子の一具体例としては、後述する実施例１～１１に示される「ＧＥＮ－５０１Ｄ０８」、「ＧＥＮ－０８０Ｇ０１」、「ＧＥＮ－０２５Ｆ０７」、「ＧＥＮ－０７６Ｃ０９」、「ＧＥＮ－３３１Ｇ０７」、「ＧＥＮ－１６３Ｄ０９」、「ＧＥＮ－０７８Ｄ０５ＴＡ１３」、「ＧＥＮ－４２３Ａ１２」、「ＧＥＮ－０９２Ｅ１０」、「ＧＥＮ－４２８Ｂ１２」、「ＧＥＮ－０７３Ｅ０７」、「ＧＥＮ－０９３Ｅ０５」及び「ＧＥＮ－０７７Ａ０９」とそれぞれ名付けられた各クローンの有するＤＮＡ配列から演繹されるものを挙げることができ、それらの各塩基配列は、配列表に示される通りである。

【００１５】

また、これら各クローンは、同配列表に示される各アミノ酸でコードされるヌクレオチド（核酸）のオープンリーディングフレームを有しており、それぞれ後記実施例に示される分子量を有していると計算された。従って、本発明の各ヒト遺伝子を、以下後記実施例１～１１に示される通りにいうことがある。

【００１６】

【発明の実施の態様】

以下、本発明ヒト遺伝子につき詳述する。

【００１７】

上述の如く、本発明ヒト遺伝子のそれぞれは、ラット、マウス、酵母、線虫及び公知のヒト遺伝子などと類似性を有し、それら類似性ある遺伝子の情報に基づくヒト遺伝子の解析とそれら解析された遺伝子の機能及び解析遺伝子と各種疾患との係わりについての研究に利用でき、遺伝子と関係ある疾患への遺伝子診断並びに該遺伝子の医薬用途への応用研究

に用いることが可能である。

【００１８】

本発明遺伝子は、例えば配列番号：２に示されるように、一本鎖ＤＮＡ配列で表されるが、本発明はかかる一本鎖ＤＮＡ配列に相補的なＤＮＡ配列やこれらの両者を含むコンポーネントもまた包含する。尚、配列番号：３ｎ－１（ｎ＝１～１４の整数）に示す本発明遺伝子の配列は、これによりコードされる各アミノ酸残基を示すコドンの一つの組合わせ例であり、本発明遺伝子はこれに限らず、各アミノ酸残基に対して任意のコドンを組合わせ選択したＤＮＡ塩基配列を有することも勿論可能である。該コドンの選択は常法に従うことができ、例えば利用する宿主のコドン使用頻度を考慮することができる〔Ncl.Acids Res., 9, 43-74 (1981)〕。

【００１９】

更に本発明遺伝子には、上記で示されるアミノ酸配列の一部のアミノ酸乃至アミノ酸配列を置換、欠失、付加などにより改変してなり、同様の機能を有する同効物をコードするＤＮＡ配列もまた包含される。これらポリペプチドの製造、改変（変異）などは天然に生じることもあり、また翻訳後の修飾により、或は遺伝子工学的手法により天然の遺伝子（本発明遺伝子）を、例えばサイトスペシフィック・ミュータゲネシス〔Methods in Enzymology, 154, p350, 367-382 (1987)；同100, p468 (1983) ；Nucleic Acids Research, 12, p9441 (1984)；続生化学実験講座１「遺伝子研究法II」、日本生化学会編, p105 (1986)〕などの方法により改変したり、リン酸トリエステル法やリン酸アミダイト法などの化学合成手段〔J. Am. Chem. Soc., 89, p4801 (1967)；同91, p3350 (1969)；Science, 150, p178 (1968) ；Tetrahedron Lett., 22, p1859 (1981)；同24, p245 (1983)〕により変異させたＤＮＡを合成したり、それらの組合せにより収得することができる。

【００２０】

本発明遺伝子は、これを利用して、即ち例えばこれを微生物のベクターに組込み、形質転換された微生物を培養することによって、上記各遺伝子によりコードされる蛋白を容易にかつ安定して発現できる。

【００２１】

また本発明の遺伝子を利用して得られる各蛋白は、これを用いて、特異抗体を作成することもできる。ここで抗原として用いられるコンポーネントは、上記遺伝子工学的手法に従って大量に産生される蛋白を用いることができ、得られる抗体はポリクローナル抗体及びモノクローナル抗体のいずれでもよく、これら抗体はそれぞれの蛋白の精製、測定、識別などに有利に利用できる。

【００２２】

本発明遺伝子の製造は、本発明によって開示された本発明遺伝子についての配列情報によれば、一般的遺伝子工学的手法により容易に実施できる〔Molecular Cloning 2nd Ed, Cold Spring Harbor Laboratory Press (1989)；続生化学実験講座「遺伝子研究法I、II、III」、日本生化学会編 (1986) など参照〕。

【００２３】

これは、例えばヒトｃＤＮＡライブラリー（各遺伝子の発現される適当な起源細胞より常法に従い調製されたもの）から、本発明遺伝子に特有の適当なプローブや抗体を用いて所望クローンを選択することにより実施できる〔Proc. Natl. Acad. Sci. USA, 78, 6613 (1981) ; Science, 222, 778 (1983)など〕。

【００２４】

上記方法において、起源細胞としては、目的の遺伝子を発現する各種の細胞、組織やこれらに由来する培養細胞などが例示され、これからの全ＲＮＡの分離、ｍＲＮＡの分離や精製、ｃＤＮＡへの変換（合成）とそのクローニングなどはいずれも常法に従い実施できる。また、ｃＤＮＡライブラリーは市販されてもおり、本発明においてはそれらｃＤＮＡライブラリー、例えばクローンテック社（Clontech Lab. Inc.）より市販の各種ｃＤＮＡライブラリーなどを用いることもできる。

【００２５】

ｃＤＮＡライブラリーからの本発明遺伝子のスクリーニングは、前記通常の方法に従い実施することができる。該スクリーニング方法としては、例えばｃＤＮＡの産生する蛋白質に対して、該蛋白質特異抗体を使用した免疫的スクリーニングにより、対応するｃＤＮＡクローンを選択する方法、目的のＤＮＡ配列に選択的に結合するプローブを用いたプラークハイブリダイゼーション、コロニーハイブリダイゼーションなどやこれらの組合せを例示できる。ここで用いられるプローブとしては、本発明遺伝子のＤＮＡ配列に関する情報をもとにして化学合成されたＤＮＡ配列などを用いるのが一般的であり、勿論既に取得された本発明遺伝子やその断片もかかるプローブとして利用できる。

【００２６】

更に各細胞、組織より抽出、単離精製された天然抽出物の部分アミノ酸配列情報に基づき、センス・プライマー、アンチセンス・プライマーをスクリーニング用プローブとして用いることもできる。

【００２７】

また、本発明遺伝子の取得に際しては、ＰＣＲ法〔Science, 230, 1350-1354 (1985)〕によるＤＮＡ／ＲＮＡ増幅法が好適に利用できる。殊にライブラリーから全長のｃＤＮＡが得られ難いような場合に、レース法（ＲＡＣＥ：Rapid amplification of cDNA ends；実験医学、12(6), 35-38 (1994)）、殊に５′レース（５′ＲＡＣＥ）法（Frohman,M.A., et al., Proc. Natl. Acad. Sci., USA., 8, 8998-9002(1988)）の採用が好適である。かかるＰＣＲ法の採用に際して使用されるプライマーは、既に本発明によって明らかにされた本発明遺伝子の配列情報に基づいて適宜設定することができ、これは常法に従い合成することができる。

【００２８】

尚、増幅させたＤＮＡ／ＲＮＡ断片の単離精製は前記の通り常法に従うことができ、例えばゲル電気泳動法などによればよい。

上記で得られる本発明遺伝子或は各種ＤＮＡ断片などの塩基配列の決定も、常法に従うことができ、例えばジデオキシ法〔Proc. Natl. Acad. Sci. USA, 74, 5463-5467 (1977)〕やマキサム－ギルバート法〔Method in Enzymology, 65, 499 (1980)〕などにより行なうことができる。かかる塩基配列の決定は、市販のシークエンスキットなどを用いても容易に行ない得る。

【００２９】

本発明遺伝子の利用によれば、通常の遺伝子組換え技術〔例えば、Science, 224, p1431 (1984) ; Biochem. Biophys. Res. Comm., 130, p692 (1985) ；Proc. Natl. Acad. Sci. USA, 80, p5990 (1983)及び前記引用文献など参照〕に従うことにより、各組換え体蛋白を得ることができる。該蛋白の製造は、より詳細には、本発明遺伝子が宿主細胞中で発現できる組換えＤＮＡを作成し、これを宿主細胞に導入して形質転換し、該形質転換体を培養することにより行なわれる。

【００３０】

ここで宿主細胞としては、真核生物及び原核生物のいずれも用いることができる。該真核生物の細胞には、脊椎動物、酵母などの細胞が含まれ、脊椎動物細胞としては、例えばサルの細胞であるＣＯＳ細胞〔Cell, 23, 175-182 (1981)〕やチャイニーズ・ハムスター卵巣細胞及びそのジヒドロ葉酸レダクターゼ欠損株〔Proc. Natl. Acad. Sci. USA, 77, 4216-4220 (1980)〕などがよく用いられているが、これらに限定される訳ではない。

脊椎動物の発現ベクターとしては、通常発現しようとする遺伝子の上流に位置するプロモーター、ＲＮＡのスプライス部位、ポリアデニル化部位及び転写終了配列などを保有するものを使用でき、これは更に必要により複製起点を有していてもよい。該発現ベクターの例としては、例えば、ＳＶ４０の初期プロモーターを保有するｐＳＶ２dhfr〔Mol. Cell. Biol., 1, 854 (1981)〕などを例示できる。また、真核微生物としては、酵母が一般によく用いられ、中でもサッカロミセス属酵母を有利に利用できる。該酵母などの真核微生物の発現ベクターとしては、例えば酸性ホスファターゼ遺伝子に対するプロモーターを有するｐＡＭ８２〔Proc. Natl. Acad. Sci. USA, 80, 1-5 (1983)〕などを利用できる。

【００３１】
また、本発明遺伝子の発現ベクターとしては、原核生物遺伝子融合ベクターを好ましく利用することができ、該ベクターの具体例としては、例えば分子量２６０００のＧＳＴドメイン（S.japonicum 由来）を有するｐＧＥＸ－２ＴＫやｐＧＥＸ－４Ｔ－２などを例示することができる。
【００３２】
原核生物の宿主としては、大腸菌や枯草菌が一般によく用いられる。これらを宿主とする場合、本発明では、例えば該宿主菌中で複製可能なプラスミドベクターを用い、このベクター中に本発明遺伝子が発現できるように該遺伝子の上流にプロモーター及びＳＤ（シヤイン・アンド・ダルガーノ）塩基配列、更に蛋白合成開始に必要な開始コドン（例えばＡＴＧ）を付与した発現プラスミドを利用するのが好ましい。上記宿主としての大腸菌としては、エシエリヒア・コリ（Escherichia coli）Ｋ１２株などがよく用いられ、ベクターとしては一般にｐＢＲ３２２及びその改良ベクターがよく用いられるが、これらに限定されず公知の各種の菌株及びベクターをも利用できる。プロモーターとしては、例えばトリプトファン(trp)プロモーター、lppプロモーター、lacプロモーター、PL/PRプロモーターなどを使用できる。
【００３３】
かくして得られる所望の組換えＤＮＡの宿主細胞への導入方法及びこれによる形質転換方法としては、一般的な各種方法を採用できる。また得られる形質転換体は、常法に従い培養でき、該培養により本発明遺伝子によりコードされる目的の蛋白が生産、発現される。該培養に用いられる培地としては、採用した宿主細胞に応じて慣用される各種のものを適宜選択利用でき、その培養も宿主細胞の生育に適した条件下で実施できる。
【００３４】
上記により、形質転換体の細胞内、細胞外乃至は細胞膜上に目的とする組換え蛋白が発現、生産、蓄積乃至分泌される。
【００３５】
各組換え蛋白は、所望により、その物理的性質、化学的性質などを利用した各種の分離操作〔「生化学データーブックII」、1175-1259 頁、第１版第１刷、1980年 6月23日株式会社東京化学同人発行；Biochemistry, 25(25), 8274-8277 (1986); Eur. J. Biochem., 163, 313-321 (1987) など参照〕により分離、精製できる。該方法としては、具体的には例えば通常の再構成処理、蛋白沈澱剤による処理（塩析法）、遠心分離、浸透圧ショック法、超音波破砕、限外濾過、分子篩クロマトグラフィー（ゲル濾過）、吸着クロマトグラフィー、イオン交換クロマトグラフィー、アフィニティクロマトグラフィー、高速液体クロマトグラフィー（ＨＰＬＣ）などの各種液体クロマトグラフィー、透析法、これらの組合せなどを例示でき、特に好ましい上記方法としては所望の蛋白を結合させたカラムを利用したアフィニティクロマトグラフィーを例示できる。
【００３６】
また、本発明によって明らかにされた本発明遺伝子の配列情報を基にすれば、例えば該遺伝子の一部又は全部の塩基配列を利用することにより、各種ヒト組織における本発明遺伝子の発現の検出を行なうことができる。これは常法に従って行なうことができ、例えばＲＴ－ＰＣＲ（Reverse transcribed-Polymerase chain reaction）（Kawasaki, E.S., et al., Amplification of RNA. In PCR Protocol, A Guide to methods and applications, Academic Press, Inc., SanDiego, 21-27(1991)）によるＲＮＡ増幅により、またノーザンブロッティング解析（Molecular Cloning, Cold Spring Harbor Laboratory(1989)）などにより、いずれも良好に実施し得る。
【００３７】
尚、前記ＰＣＲ法を採用する場合において、用いられるプライマーは、本発明遺伝子のみを特異的に増幅できる本発明遺伝子に特有のものである限りなんら限定はなく、本発明遺伝情報に基いてその配列を適宜設定することができる。通常これは常法に従って２０～３０ヌクレオチド程度の部分配列を有するものとすることができる。その好適な例は、後記

実施例１－１１に示すとおりである。
【００３８】
しかして、本発明はかかる新規なヒト遺伝子に特有の検出に有用なプライマー及び／又はプローブをも提供するものである。
【００３９】
【発明の効果】
本発明によれば、新規なヒト遺伝子が提供され、該遺伝子を用いれば、該遺伝子の各種組織での発現の検出や、その構造及び機能を解析でき、また、該遺伝子でコードされるヒト蛋白の遺伝子工学的製造が可能となり、これらにより、その発現物の解析などや、これらの関与する疾患、例えば遺伝子病、癌などの病態解明や診断、治療などが可能となる。
【００４０】
【実施例】
以下、本発明を更に詳しく説明するため、実施例を挙げる。
【００４１】
【実施例１】
ＧＤＰ解離促進蛋白遺伝子
（１）ＧＤＰ解離促進蛋白遺伝子のクローニング及びＤＮＡシークエンシング
ヒト胎児脳、成人血管、胎盤の各組織より抽出したｍＲＮＡをクローンテック社より購入し、出発材料とした。
【００４２】
上記各ｍＲＮＡよりｃＤＮＡを合成し、ベクターλＺＡＰII（ストラタジーン社製）に挿入し、ｃＤＮＡライブラリーを構築した（大塚ＧＥＮリサーチ・インスティチュート、大塚製薬株式会社）。
【００４３】
インビボ・エキシジョン法（in vivo excision : Short, J. M., et al., Nucleic Acids Res., 16, 7583-7600 (1988)）によって、寒天培地上にヒト遺伝子を含む大腸菌コロニーを形成させ、ランダムにそのコロニーをピックアップし、９６ウエルマイクロプレートにヒト遺伝子を含む大腸菌クローンを登録した。登録されたクローンは、－８０℃にて保存した。
【００４４】
次に登録した各クローンを１．５ｍｌのＬＢ培地で一昼夜培養し、プラスミド自動抽出装置ＰＩ－１００（クラボウ社製）を用いてＤＮＡを抽出精製した。尚、コンタミした大腸菌のＲＮＡは、ＲＮase処理により分解除去した。最終的に３０μｌに溶解し、２μｌは、ミニゲルによりおおまかにＤＮＡのサイズ、量をチェックし、７μｌをシークエンス反応用に用い、残りの２１μｌは、プラスミドＤＮＡとして４℃に保存した。
【００４５】
この方法によれば、若干のプログラム変更によって、後記各実施例で示されるＦＩＳＨ（fluoresence in situ hybridization）のプローブ用としても使用可能なコスミドを抽出することができる。
【００４６】
続いて、Ｔ３、Ｔ７或は合成オリゴヌクレオチド・プライマーを用いるサンガーらのジデオキシターミネーター法（Sanger, F., et al., Proc. Natl. Acad. Sci. U.S.A., 74, 5463-5467 (1977)）或はジデオキシターミネーター法にＰＣＲ法を加味した方法であるサイクルシークエンス法（Carothers, A.M., et al., Bio. Techniques, 7, 494-499 (1989)）を実施した。これらは、少量のプラスミドＤＮＡ(およそ０．１－０．５μｇ)をテンプレート(鋳型)として４種の塩基特異的に停止する伸長反応させる方法である。
【００４７】
シークエンスプライマーとしては、ＦＩＴＣ（fluorescein isothiocyanate）の蛍光標識したものを使用し、Ｔａｑポリメラーゼにより通常約２５サイクル反応させた。そのＰＣＲ産物はポリアクリルアミド・尿素ゲルで分離し、蛍光標識したＤＮＡ断片を自動ＤＮＡ

シークエンサー、ＡＬＦ™ＤＮＡシークエンサー（ファルマシア社製）によりｃＤＮＡの５’末端側から約４００塩基の配列を決定した。

【００４８】

また、３’非翻訳領域は各遺伝子のヘテロジェナイティ(heterogeneity)が高く、個々の遺伝子を区別するには適しているので、場合によっては、３’側のシークエンスも行なった。

【００４９】

ＤＮＡシークエンサーで得られた膨大な塩基配列情報は、６４ビットのコンピューターＤＥＣ３４００に転送し、コンピューターによるホモロジー解析に用いた。ホモロジー解析には、ＵＷＧＣＧのＦＡＳＴＡプログラム（Pearson, W.R. and Lipman, D. J., Proc.Natl.Acad.Sci., USA., 85, 2444-2448(1988)）によるデーターベース（GenBank, EMBL）検索により行なった。

【００５０】

上記方法によって任意に選択し、ｃＤＮＡの配列解析により、ＧＥＮ－５０１Ｄ０８と名付けた０．８キロベースの挿入を持つ一つのクローンがヒトＲａｌＧＤＰ解離促進蛋白(ＲａｌＧＤＳ: Ral guanine nucleotide dissociation stimulator）の遺伝子のＣ末端領域に対して高い相同性を示した。ＲａｌＧＤＳは、シグナルの伝達経路になんらかの役割を演ずると考えられるので、更にそのヒト相同物を提供するｃＤＮＡ挿入部の完全塩基配列を決定した。

【００５１】

低分子ＧＴＰアーゼは、細胞の増殖、分化、転写を含む多くの細胞の機能に対してシグナルを送ることにおいて重大な役割を演じている（Bourne, H. R., et al., Nature, 348, 125-132 (1990); Bourne et al., Nature, 349, 117-127 (1991)）。

【００５２】

それらの間で、ｒａｓ遺伝子ファミリーによってコードされた蛋白は、分子スィッチとして機能すること、即ち、ｒａｓ遺伝子ファミリーの機能が生物学的に非活性なＧＤＰ結合蛋白或は活性ＧＤＰ結合蛋白などの結合蛋白の異なる条件下によって調整されること、これら２つの条件がＧＴＰアーゼ活性化蛋白(GAPs)或はＧＤＳによって誘導されることがよく知られている。前者の酵素は結合ＧＴＰの加水分解を刺激することによってＧＤＰ結合を誘導し、後者の酵素は、結合ＧＤＰを遊離することによって定められたＧＴＰ結合を誘導する（Bogusuki, M.S. and McCormick, F., Nature, 366, 643-654 (1993)）。

【００５３】

ＲａｌＧＤＳは、形質転換活性のないｒａｓ遺伝子ファミリーのメンバーの一つとして、ＲＡＳに対する特異的なＧＤＰ解離促進蛋白として、最初に見つけられた（Chardin, P. and Tavitian, A., EMBO J., 5, 2203-2208(1986); Albright, C. F., et al., EMBO J., 12, 339-347 (1993)）。

【００５４】

Ｒａｌに加えて、ＲａｌＧＤＳは、これらの蛋白との相互作用によってN-ras, H-ras, K-ras, Rapのエフェクター分子として機能することが見つけ出された（Spaargaren, M. and Bischoff, J. R., Proc. Natl. Acad. Sci. U.S.A., 91, １２６０９－１２６１３　（１９９４））。

配列番号：３にＧＥＮ－５０１Ｄ０８と命名されたｃＤＮＡクローンの核酸配列、配列番号：２にそのクローンのコードディング領域の核酸配列並びに配列番号：１に該核酸配列でコードされるアミノ酸配列を示す。

【００５５】

このｃＤＮＡは、８４２塩基からなり、１２２アミノ酸をコードする３６６塩基のオープン・リーディング・フレームを含んでいた。転写開始コドンは２８番目の核酸残基に位置していた。

【００５６】

一般のデータ・ベース中において知られているＲａｌＧＤＳ蛋白とこのｃＤＮＡによって

推定されたアミノ酸配列を比較した結果、このｃＤＮＡをコードする蛋白は、ヒトのＲａｌＧＤＳのＣ末端領域に対する相同物であることが明らかとなった。この新規な遺伝子でコードされるアミノ酸配列は、ｒａｓに結合するために必要であると考えられているＲａｌＧＤＳのＣ末端領域に対して３９．５％の同一性を持っていた。

【００５７】

従って前述のごとく、この遺伝子産物がｒａｓファミリー蛋白と相互作用をしているかもしれないこと、或はｒａｓ介在シグナル変換経路に影響を与えるかもしれないと考えられる。しかしながら、この新規な遺伝子は、ＧＤＳ活性領域をコードする領域を含んでおらず、ＧＤＳ蛋白と異なる機能を有しているようである。しかして、本発明者らはこの遺伝子をヒトＲａｌＧＤＳと名付けた。

【００５８】

（２）ノーザンブロット分析

正常ヒト組織におけるＲａｌＧＤＳ蛋白のｍＲＮＡの発現をランダム・オリゴヌクレオチド・プライミング法によって標識したヒトｃＤＮＡクローンをプローブとするノーザンブロットにより評価した。

【００５９】

ノーザンブロット分析は、製品使用法に従い、ヒトＭＴＮブロット（Human Multiple Tissue Nothern blot ; クローンテック社製、パロ・アルト、カリフォルニア、米国）を用いて実施した。

【００６０】

即ち、上記ＧＥＮ－５０１Ｄ０８クローンのＰＣＲ増幅産物を〔^{32}Ｐ〕－ｄＣＴＰ（ランダムプライムドＤＮＡラベリングキット、ベーリンガーマンハイム社）により標識してプローブとした。

【００６１】

ブロッティングは、４２℃で一晩、５０％ホルムアミド／５×ＳＳＣ／５０×デンハルツ溶液／０．１％ＳＤＳ溶液（１００μｇ／ｍｌ変性サケ精子ＤＮＡ含有）の溶液中でハイブリダイズした。２×ＳＳＣ／０．０１％ＳＤＳにて室温下にて２回洗浄後、次いで０．１×ＳＳＣ／０．０５％ＳＤＳにて５０℃下に４０分間で３回洗浄した。フィルターは－７０℃下に１８時間、Ｘ線フィルム（コダック社製）に対して露光した。

【００６２】

その結果、９００塩基の転写体が、試験した全てのヒト組織内に発現していることが明らかとなった。加えて、３．２キロ転写体が心臓と骨格筋内に特異的に観察された。これらの異なるサイズを有する転写体の発現は、二者択一的なスプライシングに負うか或は相同遺伝子に対するクロスーハイブリダイゼーションに負うのかもしれない。

【００６３】

（３）ＦＩＳＨによるコスミド・クローンと染色体の局在

ＦＩＳＨは、コスミド・ベクターｐＷＥ１５中にクローン化されたヒト染色体のライブラリーをｃＤＮＡクローンの０．８キロ塩基挿入部をプローブとしてスクリーニングすることにより実施した（Sambrook, J., et al., Molecular Cloning, 2nd Ed., pp.3.1-3.58, Cold Spring Harbor Laboratory Press. Cold Spring Harbor, New York(1989)）。

【００６４】

染色体の座位決定のためのＦＩＳＨは、Ｇバンドパターンを比較することによって確認するイナザワらの方法と同様に実施した（Inazawa, J.. et al., Genomics, 17, 153-162 (1993)）。

【００６５】

染色体のスクリーニングから得られたＧＥＮ－５０１Ｄ０８に対応するコスミドＤＮＡ０．５μｇをニック・トランスレーションによってビオチン１６ｄＵＴＰで標識し、プローブとした。

【００６６】

また、反復配列に起因するバック・グラウンド・ノイズを取り除くために、超音波処理し

たヒト胎盤ＤＮＡ（１０ｍｇ／ｍｌ）の０．５μｌをプローブ液の９．５μｌに加えた。混合液は８０℃で５分間変性させ、２０％デキストラン硫酸を含む４ｘＳＳＣの等量を混合した。ついでハイブリダイゼーション混合液を、変性したスライド上に播種し、パラフィンで覆った後、３７℃で１６－１８時間、湿箱内にてインキュベートした。５０％ホルムアミド／２ｘＳＳＣで３７℃で１５分間洗浄後、２ｘＳＳＣで１５分間、更に１ｘＳＳＣで１５分間それぞれ洗浄を行なった。

【００６７】

ついでスライドは、アビジン－ＦＩＴＣ（５μｇ／ｍｌ）を含む商品名「１％ＢｌｏｃｋＡｃｅ」（大日本製薬社製）を加えた４ｘＳＳＣ中で３７℃、４０分間インキュベートした。その後、スライドは４ｘＳＳＣ、０．０５％トリトンＸ－１００を含む４ｘＳＳＣで各々１０分間洗浄し、１％ＤＡＢＣＯ（ＰＢＳ（－）：グリセロールを１：９（ｖ：ｖ）に１％ＤＡＢＣＯ、シグマ社製）を含む抗除色液ＰＰＤ液（ＰＰＤ（和光カタログＮｏ．１６４－０１５３２１）１００ｍｇ及びＰＢＳ（－）（ｐＨ７．４）１０ｍｌを０．５Ｍ Na_2CO_3／０．５ＭＮａHCO_3（９：１，ｖ／ｖ）の緩衝液（ｐＨ９．０）でｐＨを８．０に調整した後、グリセリンを追加し、総容量を１００ｍｌとしたもの）中に浸して、ＤＡＰＩ(４，６－ジアミノ－２－フェニルインドール：シグマ社製)で対比染色した。

【００６８】

試験した１００以上の分裂中期の細胞から、他の染色体上のシグナルがない特異的なハイブリダイゼーションシグナルを染色体バンド６ｐ２１．３上に観察し、ＲａｌＧＤＳ遺伝子が染色体６ｐ２１．３に座位していることを確認した。

【００６９】

この実施例で得られた本発明の新規なヒトＲａｌＧＤＳ関連遺伝子の利用によれば、該遺伝子の各種組織での発現の検出や、ヒトＲａｌＧＤＳ蛋白の遺伝子工学的製造が可能となり、これらにより、発現蛋白とｒａｓ介在のシグナルの形質導入経路への役割の研究やこれらの関与する疾患、例えば癌などの病態解明や診断、治療などが可能となると考えられるほか、本発明ＲａｌＧＤＳ蛋白遺伝子の染色体の転座を同じくする強直性脊椎炎、心房中隔欠損症、色素性網膜症、失読症などの各種疾患の発生と進展との係わりを研究することが可能となる。

【００７０】

【実施例２】

細胞骨格関連蛋白２遺伝子(ＣＫＡＰ２遺伝子)

（１）細胞骨格関連蛋白２遺伝子のクローニング及びＤＮＡシークエンシング

実施例１－（１）と同様の方法でヒト胎児脳ｃＤＮＡライブラリーから任意に選択したｃＤＮＡクローンの配列解析により、ヒト細胞骨格関連蛋白遺伝子（cytoskeleton-associated protein：ＣＡＰ）のＣＡＰ－グリシン領域を含有するいくつかのＣＡＰファミリー遺伝子と相同性の高い塩基配列を有するクローンを見いだし、これをＧＥＮ－０８０Ｇ０１と名付けた。

【００７１】

ところで、細胞骨格は真核生物の細胞の細胞質や細胞膜のすぐ内側に存在し、線維状物質が複雑に絡み合った網目状の構造体である。該細胞骨格は、チューブリンからなる微小官、アクチンからなる微小線維、デスミンやビメンチンからなる中間径線維などから構成されている。細胞骨格は細胞の骨組み的な役割を担うだけでなく、線維系の重合・脱重合により細胞の形態的変化を起こすなど動的な機能も果たしている。細胞骨格は、細胞内小器官や細胞膜の受容器、イオン・チャンネルと結合し、それらの細胞内での移動や局在性の保持に重要な役割を果たしているほか、活性の調節や相互の情報伝達の機能を有するとされ、細胞全体の生理活性の調節において、非常に重要な位置を占めていると推測される。特に癌細胞は、正常細胞と違った形態、認識反応応答を示すことから、細胞の癌化と細胞骨格の質的変化の関係が注目される。

【００７２】

この細胞骨格の活性は、多くの細胞骨格関連蛋白（ＣＡＰｓ）によって調節されている。ＣＡＰｓのひとつのグループは、微小官との関連に貢献していると推定されている高く保存されたグリシン・モチーフによって特徴付けられている（CAP-GLY領域: Riehemann K. and Song C., Trends Biochem. Sci., 18, 82-83(1993)）。

【００７３】

ＣＡＰｓのこのグループのメンバーは、ＣＬＩＰ－１７０，１５０ｋＤａ　ＤＡＰ（dynein-associated protein, or dynactin）、D. melanogaster ＧＬＵＥＤ，S. cerevisiae ＢＩＫ１，リスティン（restin: Bilbe G., et al., EMBO J., 11, 2103-2113 (1992); Hilliker C., et al., Cytogenet. Cell Genet., 65, 172-176(1994)）とC. elegansの１１３．５ｋＤａ蛋白（Wilson R., et al., Nature, 368, 32-38(1994)）を含んでおり、後者の２つの蛋白を除いて、それらは、微小官と相互作用を及ぼすことのできることが直接的或は間接的証拠によって提言されている。

【００７４】

上記ＣＬＩＰ－１７０は、インビトロにおいて微小官に対する内分泌の小胞の結合にとって重要であり、内分泌細胞小器官とともに局在する（Rickard J.E. and Kreis T.E., J. Biol. Chem., 18, 82-83 (1990); Pierre P., et al., Cell, 70, 887-900 (1992)）。

【００７５】

上記ダイナクチンは、逆行性の小胞の輸送に働く（Schroer T.A. and Sheetz M.P., J. Cell Biol., 115, 1309-1318(1991)）か或は多分、有糸分裂期間中の染色体の動きに働くとされる（Pfarr C.M., et al., Nature, 345, 263-265(1990); Steuer E.R., et al., Nature, 345, 266-268 (1990); Wordeman L., et al., J. Cell Biol., 114, 285-294(1991)）細胞質のダイニン運動を構築している因子のひとつである。

【００７６】

哺乳動物のダイナクチンの相同遺伝子のショウジョウバエ(Drosophila)ＧＬＵＥＤは、ほとんど全ての細胞の生存にとって重要であり、いくつかの神経細胞の形成にとって重要である（Swaroop A., et al., Proc. Natl. Acad. Sci. U.S.A., 84, 6501-6505 (1987); Holzbaur E.L.F., et al., Nature, 351, 579-583 (1991)）。

【００７７】

ＢＩＫ１は微小官と相互作用を示し、酵母の有糸分裂期間中のスピンドル形成に重要な役割を演じている（Trueheart J., et al., Mol. Cell. Biol., 7, 2316-2326 (1987); Berlin V., et al., J. Cell. Biol., 111, 2573-2586 (1990)）。

【００７８】

現在、これらの遺伝子は、ＣＡＰファミリー(ＣＡＰｓ)と分類されている。

【００７９】

データ・ベースの検索の結果、４６３塩基(ポリＡシグナルを除いて)の上記ｃＤＮＡクローンは、レスティンとＣＬＩＰ－１７０をコードする遺伝子と有意な核酸配列相同性を示した。しかしながら、レスティン遺伝子と比較した時、このクローンは５’部分を欠いていたので、この欠けているセグメントを単離するために５’レースの技術を用いた（Frohman M.A., et al., Proc. Natl. Acad. Sci. U.S.A., 8, 8998-9002 (1988)）。

【００８０】

（２）５′レース法（５’ＲＡＣＥ：5'rapid amplification of cDNA ends)

本発明遺伝子の５′部分を含むｃＤＮＡクローンの単離・解析は、製品使用プロトコールの一部修飾させ、市販キット（5'-Rapid AmpliFinder RACE Kit, クローンテック社）を用いた５′レース法により、以下の通り単離した。

【００８１】

ここで用いた遺伝子特異的プライマーＰ１及びプライマーＰ２は常法に従い合成されたものであり、その塩基配列は下記表１に示す通りである。またアンカー・プライマーは、市販キットに付属のものを用いた。

【００８２】

【表１】

プライマー	塩基配列
プライマーＰ１	5'-ACACCAATCCAGTAGCCAGGCTTG-3'
プライマーＰ２	5'-CACTCGAGAATCTGTGAGACCTACATACATGACG-3'

ヒト胎児脳ポリ（Ａ）＋ＲＮＡの０．１μｇを逆転写酵素（Superscript TMII RNase H Reverse Transcriptase, Life Technologies社）を用いたランダムヘキサマーにより逆転写してｃＤＮＡを得、これをＰ１プライマー及びアンカープライマーを用いた第１のＰＣＲによるワタナベらの方法（Watanabe T., et al., Cell Genet., ）により増幅させた。

【００８３】

即ち、上記ポリ（Ａ）＋ＲＮＡの０．１μｇに２．５ｍＭｄＮＴＰ／１ｘＴａｑ緩衝液(宝酒造社製)／０．２μＭ　Ｐ１プライマー、０．２μＭアダプター・プライマー／ＥｘＴａｑ酵素(宝酒造社製)０．２５単位を併せて全量を５０μｌとし、これにアンカー・プライマーを加えて室温で反応させた後、ＰＣＲに供した。即ち、９４℃１分間、次いで９４℃４５秒、６０℃４５秒、７２℃２分のサイクルを３５サイクル行ない、７２℃で５分間反応させた。

【００８４】

その後、第１のＰＣＲ反応物５０μｌ中の１μｌを特異的ネスティッドＰ２プライマーとアンカー・プライマーを用いる第２のＰＣＲ反応により増幅させた。第２のＰＣＲ産物は、１．５％アガロース・ゲル電気泳動により分析した。

【００８５】

アガロース・ゲル電気泳動により、およそ６５０塩基の大きさを示す単一のバンドを検出し、このバンドの産物をベクター（ｐＴ７Ｂｌｕｅ（Ｒ）Ｔ－Ｖｅｃｔｏｒ，ノバゲン(Novagen) 社）に挿入し、適当なサイズの挿入がみられる複数のクローンを選別した。

【００８６】

ＰＣＲ反応物から得られた５’レース・クローンの６つが同じ配列を有しているが、異なる長さであった。ＧＥＮ－０８０Ｇ０１とＧＥＮ－０８０Ｇ０１４９の２つの重なりを持つｃＤＮＡクローンをシークェンシングすることによって、蛋白質をコードしている配列と、５’と３’フランキング配列と全長１０１５塩基の配列を決定し、該遺伝子を細胞骨格関連蛋白２遺伝子(ＣＫＡＰ２遺伝子)と命名した。

【００８７】

配列番号：６に上記ＧＥＮ－０８０Ｇ０１とＧＥＮ－０８０Ｇ０１４９の２つの重なっているｃＤＮＡクローンから得られた核酸配列、配列番号：５にそのクローンのコードディング領域の核酸配列、並びに配列番号：４に核酸配列によってコードされるアミノ酸配列をそれぞれ示す。

【００８８】

ＣＫＡＰ２遺伝子は、配列番号：６で示されるように、その５’非コード領域が相対的にＧＣ－リッチであり、（ＣＡＧ）４（ＣＧＧ）４（ＣＴＧ）（ＣＧＧ）の不完全なトリプレットが４０－６９塩基の位置に存在していた。

【００８９】

その２７４－２７６塩基の位置のＡＴＧは、推定されるスタート・コドンである。ストップ・コドン（ＴＧＡ）は、８５３－８５５塩基に位置していた。ポリアデニレーション・シグナル(ＡＴＴＡＡＡ)は、ポリ(Ａ)スタートの前の１６塩基にあった。推定されたオープン・リーディング・フレームは、２１，８００の計算された分子量を持つ１９３アミノ酸をコードする５７９塩基からなっている。

【００９０】

更にＲＴ－ＰＣＲによってコード領域を増幅し、得られた合成配列がｃＤＮＡのキメラで

あるという可能性を排除した。
【００９１】
（２）他のＣＲＰｓとＣＫＡＰ２との類似性
ＣＫＡＰ２の配列は、レスティンとＣＬＩＰ－１７０の配列との相同性を明らかにしたが、相同性領域は、ＣＡＰ－ＧＬＹドメインの短い配列に限定されていた。アミノ酸レベルにおいて、推定されたＣＫＡＰ２は、他の５つのＣＡＰｓとも高い相同性を示した。
【００９２】
ＣＫＡＰ２は、アルファ－ヘリックス・ロッド、亜鉛フィンガー・モチーフのようないくつかのＣＡＰｓの特徴的な他のモチーフを保有していなかった。アルファ－ヘリックス・ロッドは、２量体形成や微小管結合能を増加させることに貢献していると思われた（Pierre P., et al., Cell, 70, 887-900, (1992)）。アルファ－ヘリックス領域の欠失は、ＣＫＡＰ２がホモ－或はヘテロ－・ダイマー形成することのできないことを意味するかもしれない。
【００９３】
これらの蛋白のＣＡＰ－ＧＬＹ領域を並べると、グリシンの他に保存された残基もまたＣＫＡＰ２中にあることが明らかとなった。また、ＣＡＰ－ＧＬＹ領域を持つＣＡＰｓは、細胞小器官の活動において、そして微小官とそれらの相互作用に関係すると考えられた。ＣＫＡＰ２は、上記の如くＣＡＰ－ＧＬＹ領域を含んでいるので、これらのファミリーの中に位置付けされる。
【００９４】
Ｇｌｕｅｄの変異体を用いた研究によれば、Ｇｌｕｅｄ産物は、ほとんど全ての細胞において重大な役割を演ずること（Swaroop A., et al., Proc. Natl. Acad. Sci. U.S.A., 84, 6501-6505 (1987)）及び神経細胞内において他の神経－特異的な機能を演じること（Meyerowizt E.M. and Kankel D.R., Dev. Biol., 62, 112-142 (1978)）が明らかにされている。これらの微小官関連蛋白は、小胞輸送と有糸分裂に働くと思われている。故に、神経細胞における小胞輸送システムの重要性は、これらのコンポーネントの欠損が異常な神経細胞系に導くかもしれない。
【００９５】
このようなことから、ＣＫＡＰ２は本質的な細胞機能と同様に特異的な神経細胞の機能に関係しているかもしれない。
【００９６】
（３）ノーザンブロット分析
正常ヒト組織におけるヒトＣＫＡＰ２ｍＲＮＡの発現を、実施例１－（２）と同様にしてＧＥＮ－０８０Ｇ０１クローン（５５３－１０１５塩基に相当する）をプローブとするノーザンブロットにより評価した。
【００９７】
その結果、試験したヒト心臓、脳、胎盤、肺、肝臓、骨格筋、腎臓、膵臓の８つの全ての組織において、ＣＫＡＰ２のｃＤＮＡのサイズに一致する１．０kbの転写体を検出した。該１．０ｋｂの転写体は、試験した他の組織より、心臓、脳において有意に高いレベルで発現されていた。３．４ｋｂと４．６ｋｂの２つの弱いバントがそれぞれ全ての組織において検出された。
【００９８】
ノーザン・ブロット分析によれば、上記３．４ｋｂと４．６ｋｂの転写体は、選択的スプライシングによって、ＣＫＡＰ２の１．０ｋｂ転写体をコードする同じ遺伝子から誘導されているか、もしくは、別の関連遺伝子から転写されたものである可能性も存在する。これらの転写体の特徴は、ＣＫＡＰ２が、ＣＡＰ－ＧＬＹ領域を有し、α－ヘリックスも有するような蛋白をもコードしているかもしれない。
【００９９】
（４）ダイレクト・Ｒ－バインディングＦＩＳＨによるコスミド・クローンと染色体上の局在

ＣＫＡＰ２ｃＤＮＡに対応する２つのコスミドを得た。それらの２つのコスミド・クローンを実施例１－（３）と同様にしてダイレクト・Ｒ－バインディングＦＩＳＨによってＣＫＡＰ２の染色体の座位のマッピングを行なった。

【０１００】

また、反復配列によるバックグラウンドの抑制のために、リヒター（Lichter P. et al., Proc. Natl. Sci. U.S.A., 87, 6634-6638(1990)）によって記載されたように２０倍過剰のヒトＣｏｔ－Ｉ　ＤＮＡ(ＢＲＬ社製)を加えた。プロビア１００フィルム(フジＩＳＯ１００；フジ・フィルム社製)を、顕微鏡写真撮影のために用いた。

【０１０１】

その結果、ＣＫＡＰ２は、染色体バンド１９ｑ１３．１１－ｑ１３．１２上にマップされた。

【０１０２】

ＤＮＡマーカーＤ１９Ｓ２２１とＤ１９Ｓ２２２の間のＣＡＤＡＳＩＬ（大脳皮質梗塞及び大脳白質萎縮を持つ大脳の常染色体優性動脈疾患）とＤ１９Ｓ２１５とＤ１９Ｓ２１６の間のＦＨＭ(家族性片麻痺性偏頭痛)の２つの常染色体性の優性神経疾患が、連鎖分析によってこの領域に局在されている。これら２つの疾患は、同じ遺伝子の対立遺伝子による疾患であるかもしれない（Tournier-Lasserve E. et al., Nature Genet., 3, 256-259(1993); Joutel A. et al., Nature Genet., 5, 40-45(1993)）。

【０１０３】

ＦＨＭと、或はＣＡＤＡＳＩＬ候補遺伝子としてＣＫＡＰ２を支持するための証拠はないが、その変異体は、なんらかの神経細胞の疾患に導くかもしれないと考えられる。

【０１０４】

この実施例により得られる本発明の新規なヒトＣＫＡＰ２遺伝子を用いれば、該遺伝子の各種組織での発現の検出や、ヒトＣＫＡＰ２の遺伝子工学的製造が可能となり、これらにより、前述したように多様で且つ細胞に本質的な活動に深く関与しているヒトＣＫＡＰ２システム乃至ヒトＣＫＡＰ２の機能の解析や、これが関与する各種神経疾患、例えば家族性偏頭痛、などの診断などを行なうことができ、またこれらの治療及び予防薬のスクリーニングや評価などをも行なうことができる。

【０１０５】

【実施例３】

ＯＴＫ２７遺伝子

（１）ＯＴＫ２７遺伝子のクローニング及びＤＮＡシークエンシング

実施例１－（１）と同様の方法でヒト胎児脳ｃＤＮＡライブラリーから任意に選択したｃＤＮＡクローンの配列解析とデータ・ベースの検索の結果、酵母核蛋白（Saccharomyces cerevisiae: Kolodrubetz, D. and Burgum, A., YEAST, 7, 79-90 (1991)）、ＮＨＰ２に対して高い相同性を有する蛋白をコードするｃＤＮＡクローン、ＧＥＮ－０２５Ｆ０７を見出し、ＯＴＫ２７と命名した。

【０１０６】

核蛋白質は、染色体、リボソームなどの細胞の基本構成成分であり、この核蛋白質は細胞増殖、生存能力に対して本質的な役割を演じていると考えられている。酵母の核蛋白質であるＮＨＰ２は、High-mobility group（ＨＭＧ）－like proteinで、ＨＭＧと同様に細胞の生存能力に対して必須の機能を有していると報告されている（Kolodrubetz, D. and Burgum, A., YEAST, 7, 79-90 (1991)）。

【０１０７】

上記酵母ＮＨＰ２に高い相同性を有する本発明の新規なヒト遺伝子、ＯＴＫ２７遺伝子もまた同様な機能を有していると推測される。

【０１０８】

上記該ＧＥＮ－０２５Ｆ０７クローンの塩基配列は、配列番号：９で示されるように１４９３塩基からなり、配列番号：８で示される３８４塩基のオープン・リーディング・フレームを含んでおり、これによりコードされるアミノ酸配列は、配列番号：７に示される通

り１２８アミノ酸からなっていた。開始コドンは、配列番号：９の塩基配列番号の９５－９７番目から始まり、４７９－４８１番目が終始コドンを示していた。

【０１０９】

ＯＴＫ２７ペプチドは、ＮＨＰ２とアミノ酸レベルにおいて３８％の高い相同性を示した。また、シロイヌナズナ（Arabidopsis thaliana: Newman, T., unpublished ; GENEMBL Accession No. T14197）からのｃＤＮＡの核酸配列から推定された蛋白に対して８３％同一であった。

【０１１０】

（２）ノーザンブロット分析

正常ヒト組織におけるヒトＯＴＫ２７ｍＲＮＡの発現を、ＯＴＫ２７ｃＤＮＡクローンのインサート部分をＰＣＲにより増幅し、該ＰＣＲ産物を精製し、〔^{32}Ｐ〕－ｄＣＴＰ（ランダムプライムドＤＮＡラベリングキット、ベーリンガーマンハイム社）により標識してプローブとし、実施例１－（２）に準じてノーザンブロッティングを行なった。

【０１１１】

ノーザンブロット分析の結果、この遺伝子からの２つの可能性ある転写体を示しているおよそ１．６キロ塩基と０．７キロ塩基の位置に２つのバンドを検出した。両サイズの転写体が試験した全ての正常成人組織に発現されていたが、０．７キロ塩基の転写体の発現は脳において有意に減少し、試験した他の組織に比べて心臓、骨格筋、睾丸においてより多く発現されていた。

【０１１２】

これらの２つの転写体の更なる試験のために睾丸ｃＤＮＡライブラリーから１１のｃＤＮＡクローンを単離し、実施例１－（１)と同様にそれらＤＮＡ配列を決定した。

【０１１３】

その結果、６つのクローンにおいて、それらは０．７キロ塩基の転写体に一致しており、ポリＡ配列は６００番目辺りの塩基に始まり、これらのクローンのうち２つのクローンにおいてポリＡ配列が５９８番目の塩基、３つのクローンが６０６番目の塩基、１つのクローンが６１３番目の塩基で始まっていた。

【０１１４】

これら６つのクローンにおいて、５８３－５８８番目の塩基で“ＴＡＴＡＡＡ”配列が、ポリＡシグナルとしておそらく認められた。また、この遺伝子の上流ポリＡシグナル“ＴＡＴＡＡＡ”は、脳においてほとんど影響せず、他の組織より上記した３つの組織においてより効果的であると認識され、各転写体の安定性は種々の組織間で異なっている可能性も考えられた。

【０１１５】

更に、ズー・ブロット分析の結果より、この遺伝子は他の脊椎動物においてもよく保存されていることが示された。この遺伝子は、正常成人組織のいたるところで発現されており、種間の広い範囲で保存されていたので、その遺伝子産物は重要な生理学的な役割を演じると考えられる。ＮＨＰ２を欠いている酵母は生存できないことから、ヒトの相同物もまた細胞の生存能力に対して本質的であるかもしれないと提言され得る。

【０１１６】

（３）ダイレクト・Ｒ－バィンディングＦＩＳＨによるＯＫＴ２７染色体の局在

プローブとしてＯＴＫ２７ｃＤＮＡのインサートを用いてcＤＮＡ ＯＴＫ２７に対応する１つのコスミド・クローンを全ヒト染色体コスミド・ライブラリー（５ゲノムに相当）から単離し、実施例１－（３）に準じてダイレクト・Ｒ－バィンディングＦＩＳＨを実施し、ＯＴＫ２７の染色体の局在を決定した。

【０１１７】

その結果、明瞭な２つのスポットが染色体１２ｑ２４．３のバンド上に観察された。

【０１１８】

本発明ＯＴＫ２７遺伝子は、ヒト核蛋白質、ＯＴＫ２７蛋白の発現及び検出に利用でき、かくして該蛋白の関与する各種の疾患の診断、病態解明、例えば細胞増殖や分化に関わる

ことからある種の癌の治療及び予防薬のスクリーニングや評価に利用できる。
【０１１９】
【実施例４】
ＯＴＫ１８遺伝子
（１）ＯＴＫ１８遺伝子のクローニング及びＤＮＡシークエンシング
亜鉛－フィンガー蛋白は、真核生物の転写調節を行なう蛋白の大きなファミリーと定義されており、進化的に保存された構造モチーフを含んでいる（Kadonaga, J. T., et al., Cell, 51, 1079-1090 (1987); Klung, A. and Rhodes, D., Trends Biol. Sci., 12, 464-469 (1987); Evans, R. M. and Hollenberg, S. M., Cell, 52, 1-3 (1988)）。
【０１２０】
亜鉛イオンと、システィンとヒスチジンの２残基の相互作用によって形作られたループ様モチーフである亜鉛フィンガーは、ＲＮＡ又はＤＮＡに対する蛋白の特異的結合に関係している。亜鉛フィンガー・モチーフは当初、カエル(Xenopus)の転写因子IIIAのアミノ酸配列内において確認されていた（Miller, J., et al., EMBO J., 4, 1609-1614 (1986)）。
【０１２１】
C_2H_2フィンガーモチーフは、通常いくつか繰り返され、その間には７或は８アミノ酸の介在配列が進化的に保存されている。この介在している広がりは、当初、ショウジョウバエのクルッペルセグメンテーション遺伝子において確認された（Rosenberg, U. B., et al., Nature, 319, 336-339 (1986)）。それ以来、脊椎動物ゲノムは、数百のC_2H_2亜鉛フィンガー蛋白をコードする遺伝子が発見されている。
【０１２２】
実施例１－（１）と同様の方法でヒト胎児脳ｃＤＮＡライブラリーから任意に選択したｃＤＮＡクローンの配列解析とデータ・ベースの検索の結果、亜鉛フィンガー構造を含むいくつかのクローンを確認し、ここにクルッペル・タイプの亜鉛フィンガー構造をもつクローンを見い出した。
【０１２３】
このクローンは転写体の５’部分が欠けていたので、プローブとしてｃＤＮＡクローンの約１．８キロ塩基のインサートを使用し、胎児脳ｃＤＮＡライブラリーをプラーク・ハイブリダイゼーションして、３つのクローンを単離し、実施例１－（１）と同様にそれらの塩基配列を決定した。
【０１２４】
この３つクローンのうちインサートの最も大きいクローンは、７１１アミノ酸をコードする２１３３塩基のオープン・リーデイング・フレームを持つ３７５４塩基長に及んでおり、該クローンは、ヒトＺＮＦ４１とショウジョウバエのクルッペル遺伝子に対して、亜鉛フィンガー領域で高い類似性を有するペプチドをコードする新規なヒト遺伝子を含んでいた。この遺伝子をＯＴＫ１８遺伝子（ＧＥＮ－０７６Ｃ０９クローン由来）と命名した。
【０１２５】
ＯＴＫ１８遺伝子のｃＤＮＡクローンの核酸配列は、配列番号：１２で示され、またコード領域を含む塩基配列は、配列番号：１１で示され、該ＯＴＫ１８遺伝子によりコードされる推定アミノ酸配列は、配列番号：１０で示される通りである。
【０１２６】
尚、配列番号：１２から、ＯＴＫ１８の推定されたペプチド配列は、その蛋白がそのカルボキシ側で１３のフィンガー・モチーフを含んでいることが判明した。
【０１２７】
（２）他の亜鉛フィンガー・モチーフを有する遺伝子との比較
ＯＴＫ１８、ヒトＺＮＦ４１、ショウジョウバエのクルッペル遺伝子の比較により、各フィンガー・モチーフは、共通配列ＣＸＥＣＧＫＡＦＸＱＫＳＸＬX_2ＨＱＲＸＨに対して大部分保存されていた。
【０１２８】

ヒトＺＮＦ４１とショウジョウバエのクルッペル遺伝子に対するＯＴＫ１８の亜鉛フィンガー・モチーフの共通配列の比較は、クルッペルタイプ・モチーフがＯＴＫ１８遺伝子によりコードされた蛋白内によく保存されていた。しかしながら、配列の類似性は亜鉛フィンガー領域に限定されており、他の領域では有意なホモロジーは示さなかった。

【０１２９】

また、亜鉛フィンガー領域は、標的ＤＮＡと特異的相互作用し、約５塩基の配列を認識して、ＤＮＡのヘリックスと特異的に結合する（Rhodes, D. and Klug, A., Cell, 46, 123-132 (1986)）。

【０１３０】

このことからマルチ・モデュール（亜鉛フィンガーがタンデムにつながっていること）は、ＤＮＡの広い範囲と相互作用することができるとする考えに基づいて、１３の繰り返し単位を含んでいるこの遺伝子産物の標的ＤＮＡは、およそ６５塩基長のひとつのＤＮＡ断片であることが推測される。

【０１３１】

（３）ノーザンブロット分析

ノーザンブロット分析は、実施例１－（２）に準じて、正常ヒト組織におけるヒトＯＴＫ１８ｍＲＮＡの発現を、ＯＴＫ１８ｃＤＮＡクローンのインサートをＰＣＲにより増幅し、該ＰＣＲ産物を精製し、〔^{32}Ｐ〕－ｄＣＴＰ（ランダムプライムドＤＮＡラベリングキット、ベーリンガーマンハイム社）により標識してプローブとしてＭＴＮブロットを用いて実施した。

【０１３２】

ノーザンブロット分析の結果、ＯＴＫ１８の転写体がおよそ４．３キロ塩基の長さであることが明らかとなり、これは種々の正常成人組織内いたるところに発現されていた。しかしながら、肝臓と末梢血リンパ球における発現のレベルは試験した他の臓器においてより低いと思われた。

（４）ダイレクト・Ｒ－バィンディングＦＩＳＨによるコスミド・クローンと染色体の局在

ＯＴＫ１８の染色体の局在を実施例１－（３）に準じて行なった。

【０１３３】

その結果、８サンプルで２つの完全なスポットが確認され、２３サンプルはどちらか一方或は両方に不完全なシグナル又は２つのスポットを呈した。全てのシグナルは染色体１９のｑ１３．４バンドに現われた。他の染色体上には２つのスポットは観察されなかった。

【０１３４】

このようにＦＩＳＨの結果、この遺伝子が染色体バンド１９ｑ１３．４に局在することが判った。この領域は亜鉛フィンガー領域のオリゴヌクレオチドとハイブリダイズする多くのＤＮＡセグメントを含んでいることが知られている（Hoovers, J. M. N., et al., Genomics, 12, 254-263 (1992)）。しかも、亜鉛フィンガー蛋白をコードする少なくとも一つの他の遺伝子は、この領域で見つけられている（Marine, J.-C., et al., Genomics, 21, 285-286 (1994)）。

【０１３５】

このことから染色体１９ｑ１３は、転写を調節する蛋白をコードする多くのマルチ遺伝子が集まっている部位であると推測される。

【０１３６】

この実施例により提供される新規なヒトＯＴＫ１８遺伝子を用いれば、該遺伝子の各種組織での発現の検出や、ヒトＯＴＫ１８蛋白の遺伝子工学的製造が可能となり、これらにより、前述したように多様で且つ細胞に基本的な活動に深く関与しているヒト転写調節蛋白質遺伝子システム乃至ヒト転写調節蛋白質の機能の解析や、これが関与する各種疾患、例えば発生分化異常の奇形や癌、或は神経系発生異常の精神、神経障害などの診断などを行なうことができ、またこれらの治療及び予防薬のスクリーニングや評価などをも行なうことができる。

【０１３７】
【実施例５】
ヒト２６Ｓプロテアソーム構成成分Ｐ４２蛋白質及びＰ２７蛋白質遺伝子
（１）ヒト２６Ｓプロテアソーム構成成分Ｐ４２蛋白質及びＰ２７蛋白質の各遺伝子のクローニング及びＤＮＡシークエンシング
多機能プロテアーゼであるプロテアソームは、酵母からヒトに至る真核生物に広く存在し、細胞内でエネルギー依存的にユビキチン結合蛋白質を分解する酵素である。該プロテアソームは、構造的には分子量２１～３１キロダルトンの種々の構成成分からなる２０Ｓプロテアソームと、３０～１１２キロダルトンの種々の構成成分からなる沈降係数２２ＳのＰＡ７００制御蛋白質群とで構成され、全体として沈降係数２６Ｓ、分子量約２００万ダルトンの巨大分子を形成している〔Rechsteiner, M., et al., J.Biol.Chem., 268, 6065-6068 (1993)、Yoshimura, T., et al., J.Struct.Biol., 111, 200-211 (1993)、Tanaka, K., et al., New Biologist, 4, 173-187 (1992)〕。
【０１３８】
その構造的及び機械的な解析からも尚、該プロテアソームの全容は明らかになっていないが、主として酵母とマウスを用いた研究から、以下の機能が報告され、その機能は次第に明らかになりつつある。
【０１３９】
即ち、細胞内におけるエネルギー依存性の蛋白質分解機構は、ユビキチン結合による蛋白質の選別にはじまるが、ユビキチン化蛋白質を分解するＡＴＰ依存的な活性は２０Ｓプロテアソームにはなく、２６Ｓプロテアソームにある〔Chu-Ping et al., J.Biol.Chem., 269, 3539-3547 (1994)〕。このことから、ヒト２６Ｓプロテアソームは、エネルギー依存性の蛋白質分解機構の解明に有用であると考えられる。
【０１４０】
細胞周期の調節に関与する因子は、一般に半減期が短く、厳格な量的制御を受けている場合が多い。実際、癌遺伝子産物であるＭｏｓ、Ｍｙｃ、Ｆｏｓなどはエネルギー、ユビキチン依存的に２６Ｓプロテアソームにより分解されることが明らかにされ〔Ishida, N., et al., FEBS Lett., 324, 345-348 (1993)、Hershko,A., and Ciechanover, A., Annu.Rev.Biochem., 61, 761-807 (1992)〕、細胞周期制御におけるプロテアソームの重要性が認識されつつある。
【０１４１】
また免疫系における重要性も指摘されており、クラスＩ主要組織適合複合体抗原提示において、プロテアソームが積極的に関与していることも示唆され〔Michalek MT., et al., Nature, 363, 552-554 (1993)〕、さらにアルツハイマー患者の脳内において、ユビキチン化蛋白質が異常に蓄積してるという現象〔Kitaguchi, N., et al., Nature, 361, 530-532 (1988)〕から、アルツハイマー病にもプロテアソームが関与していることが示唆されるなど、プロテアソームは多彩な機能を有するという点で、各種病態の診断や治療にその有用性が注目されてきている。
【０１４２】
また、２６Ｓプロテアソームの主たる機能は、ユビキチン化蛋白質を分解する活性である。中でもｃ－Ｍｙｃをはじめとする癌遺伝子産物やサイクリンのような細胞周期関連遺伝子産物の分解が、ユビキチン依存の経路で分解されることが判明している。また肝癌細胞、腎癌細胞、白血病細胞などにおいてプロテアソーム遺伝子は、正常細胞に比較して異常発現しており〔Kanayama, H., et al., Cancer Res., 51, 6677-6685 (1991)〕、腫瘍細胞核にプロテアソームが異常蓄積していることが観察されている。このことから、ヒトプロテアソームの構成成分はこれらの癌化のメカニズムの解明や癌の診断あるいは治療に有用となることが期待される。
【０１４３】
更に、プロテアソームの発現がインターフェロンγなどによって誘導され、細胞内での抗原提示に深く関与することが知られている〔Aki, M., et al., J.Biochem., 115, 257-26

9 (1994)〕。このことから、ヒトプロテアソームの構成成分は免疫系の抗原提示のメカニズムの解明や免疫制御剤の開発にも期待できる。

【０１４４】

加えて、プロテアソームはアルツハイマー患者の脳内に異常蓄積しているユビキチンとも深く関わっているとも考えられ、この点より、ヒトプロテアソームの構成成分は、アルツハイマー病の原因解明やその治療にもその有用性が示唆されている。

【０１４５】

このように、多機能が期待できるプロテアソームは、該酵素自体の利用以外にも、その各構成成分を抗原として抗体を作製すれば、かかる抗体が免疫測定法により各種病態の診断に用い得るとも考えられ、この診断分野での有用性も着目されている。

【０１４６】

ところで、ヒト２６Ｓプロテアソームの特性を有する蛋白質については、例えば特開平５－２９２９６４号に開示があり、また、ラットプロテアソーム構成蛋白質については、特開平５－２６８９５７号及び特開平５－３１７０５９号に開示があるが、ヒト２６Ｓプロテアソームの構成成分については知られていない。そこで本発明者らは、ヒト２６Ｓプロテアソームの構成成分について更なる検索を行い２つの新たなヒト２６Ｓプロテアソーム構成成分蛋白質であるヒト２６Ｓプロテアソーム構成成分Ｐ４２蛋白質及びヒト２６Ｓプロテアソーム構成成分Ｐ２７蛋白質を得、以下に示される方法にて該遺伝子のクローニング及びＤＮＡシークエンシングを行なった。

【０１４７】

（１）ヒト２６Ｓプロテアソーム構成成分Ｐ４２蛋白質及びＰ２７蛋白質の精製

新鮮なヒト腎臓約１００ｇを用い、特開平５－２９２９６４号公報に記載されているヒトプロテアソームの精製法に従い、バイオゲルＡ－１．５ｍ（５×９０ｃｍ、バイオラッド製）、ハイドロキシアパタイト（１．５×１５ｃｍ、バイオラッド製）、Ｑ－セファロース（１．５×１５ｃｍ、ファルマシア製）によるカラムクロマトグラフィー及びグリセロール密度勾配遠心法により、ヒトプロテアソームの精製を行なった。

【０１４８】

得られたヒトプロテアソームを、日立Ｌ６２００型高速液体クロマトグラフィー（ＨＰＬＣ）システムより、逆相ＨＰＬＣを行なった。カラムは、Shodex RS Pak D4-613（０．６×１５ｃｍ、昭和電工製）を用い、以下の２液によるグラジエント溶出を行なった。

【０１４９】

第１液：０．０６％トリフルオロ酢酸

第２液：０．０５％トリフルオロ酢酸、７０％アセトニトリル

各溶出画分につき、その一部をジチオスレイトール還元下、８．５％ＳＤＳ－ポリアクリルアミド電気泳動を行ない検出し、Ｐ４２蛋白質及びＰ２７蛋白質をそれぞれ単離精製した。

【０１５０】

精製したＰ４２及びＰ２７蛋白質について、それぞれ０．１Ｍトリス緩衝液（ｐＨ７．８）、２Ｍ尿素中、１μｇのトリプシンにより３７℃で８時間消化を行ない、得られた部分ペプチド断片を逆相ＨＰＬＣで分離し、エドマン分解により配列を決定した。その結果は、それぞれ下記表２に示す通りであった。

【０１５１】

【表２】

部分蛋白質	アミノ酸配列
Ｐ４２（１）	ＶＬＮＩＳＬＷ
（２）	ＴＬＭＥＬＬＮＱＭＤＧＦＤＴＬＨＲ
（３）	ＡＶＳＤＦＶＶＳＥＹＸＭＸＡ
（４）	ＥＶＤＰＬＶＹＮＸ
（５）	ＨＧＥＩＤＹＥＡＩＶＫ
（６）	ＬＳＸＧＦＮＧＡＤＬＲＮＶＸＴＥＡＧＭＦＡＩＸＡＤ
（７）	ＭＩＭＡＴＮＲＰＤＴＬＤＰＡＬＬＲＰＧＸＬ
（８）	ＩＨＩＤＬＰＮＥＱＡＲＬＤＩＬＫ
（９）	ＡＴＮＧＰＲＹＶＶＶＧ
（１０）	ＥＩＤＧＲＬＫ
（１１）	ＡＬＱＳＶＧＱＩＶＧＥＶＬＫ
（１２）	ＩＬＡＧＰＩＴＫ
（１３）	ＸＸＶＩＥＬＰＬＴＮＰＥＬＦＱＧ
（１４）	ＶＶＳＳＳＬＶＤＫ
（１５）	ＡＬＱＤＹＲＫ
（１６）	ＥＨＲＥＱＬＫ
（１７）	ＫＬＥＳＫＬＤＹＫＰＶＲ
Ｐ２７（１）	ＬＶＰＴＲ
（２）	ＡＫＥＥＥＩＥＡＱＩＫ
（３）	ＡＮＹＥＶＬＥＳＱＫ
（４）	ＶＥＤＡＬＨＱＬＨＡＲ
（５）	ＤＶＤＬＹＱＶＲ
（６）	ＱＳＱＧＬＳＰＡＱＡＦＡＫ
（７）	ＡＧＳＱＳＧＧＳＰＥＡＳＧＶＴＶＳＤＶＱＥ
（８）	ＧＬＬＧＸＮＩＩＰＬＱＲ

（２）ｃＤＮＡライブラリーの検索、クローンの単離及びｃＤＮＡ塩基配列の決定

本発明者らは、実施例１－（１）に示すように、ヒト胎児脳、血管動脈、胎盤ｃＤＮＡライブラリーの大量ＤＮＡ配列の決定から、約３万個のｃＤＮＡデーターを含むデーターベースを構築している。

【０１５２】

上記（１）で得られた各アミノ酸配列をもとにコンピューターによるＦＡＳＴＡ検索（アミノ酸配列に対して、ＤＮＡデーターベースから推測されたアミノ酸配列とのホモロジー検索）を行なった結果、Ｐ４２については、２つのアミノ酸配列（表２に示す（２）及び（７））が一致するクローン（ＧＥＮ－３３１Ｇ０７）を、またＰ２７についても、２つのアミノ酸配列（表２に示す（１）及び（８））が一致するクローン（ＧＥＮ－１６３Ｄ０９）を見出した。

【０１５３】

これら各クローンにつき、実施例２－（２）と同様にして、５′ＲＡＣＥ法により、５′側の配列を決定し、合わせて、それぞれの全配列を決定した。

【０１５４】

その結果、上記Ｐ４２クローンであるＧＥＮ－３３１Ｇ０７は、配列番号：１５に示され

る１５６６塩基配列から構成され、その中に配列番号：１４で示される１１６７塩基の翻訳領域を含み、これによってコードされるアミノ酸配列は、配列番号：１３に示す３８９アミノ酸配列であることが明らかとなった。

【０１５５】

Ｐ４２蛋白質は、コンピューターを用いたホモロジー検索の結果、ＡＡＡ（ATPases associated with a variety of cellular activities）蛋白質ファミリー（例えば、Ｐ４５、ＴＢＰ１、ＴＢＰ７、Ｓ４、ＭＳＳ１など）と、有意なホモロジーを有することが認められ、このことからＡＡＡ蛋白質ファミリーの新しい一員であると示唆された。

【０１５６】

また、Ｐ２７クローンであるＧＥＮ－１６３Ｄ０９は、配列番号：１８に示される１１２８塩基配列から構成され、その中に配列番号：１７で示される６６９塩基の翻訳領域を含み、これによってコードされるアミノ酸配列は、配列番号：１６に示す２２３アミノ酸配列であることが明らかとなった。

【０１５７】

Ｐ２７蛋白質は、コンピューターを用いたホモロジー検索の結果、パブリックなデーターバンク内にホモロジーを示す遺伝子は認められず、未知の機能を持つ新規な遺伝子と考えられた。

【０１５８】

しかして、上記Ｐ４２及びＰ２７の両遺伝子産物は、もともとプロテアソーム複合体の調節サブユニット成分として精製されたものであり、これらは、いずれも蛋白質分解を介して種々の生体機能にとって重要な役割、例えばＡＴＰ分解によるエネルギー供給のための役割を演じるていると考えられ、これらの点よりヒト２６Ｓプロテアソームの機能解明のみならず、上記生体機能の低下などに起因する各種病態の診断、治療などに有用であると考えられる。

【０１５９】

【実施例６】

ＢＮＡＰ遺伝子

（１）ＢＮＡＰ遺伝子のクローニング及びＤＮＡシークエンシング

ＤＮＡとヒストンで構成されるヌクレオソームは、真核生物の細胞において染色体を構成する基本構造で、種を越えてよく保存されている。この構造は、ＤＮＡの複製過程と転写に強く関連している。しかしながら、ヌクレオソームの形成については、完全には理解されておらず、ヌクレオソーム構築（nucleosome assembly: NAPs）に係わるいくつかの特異的な因子が確認されているのみである。

即ち、既に２つの酸性蛋白であるヌクレオプラスミン(nucleoplasmin)とＮ１が、ヌクレオソーム構築を容易にすることが確認されている（Kleinschmidt, J.A., et al., J. Biol. Chem., 260, 1166-1176 (1985): Dilworth, S. M., et al., Cell, 51, 1009-1018 (1987)）。

【０１６０】

モノクローナル抗体を用いて、酵母ＮＡＰ－Ｉの遺伝子が単離され、その組換え蛋白が、イン・ビトロにおいてヌクレオソーム構築活性を保有するかどうかが調べられた。

【０１６１】

更に近年、酵母のＮＡＰ－Ｉの哺乳動物相同物であるマウスのＮＡＰ－Ｉ遺伝子（Okuda, A.: registered in database as an accession number D12618）、マウスＤＮ３８遺伝子（Kato, K., Eur. J. Neurosci., 2, 704-711 (1990)）及びヒトヌクレオソーム構築蛋白（human nucleosome assembly protein: hNRP）がそれぞれクローニングされ（Simon, H. U., et al., Biochem. J., 297, 389-397(1994)）、ｈＮＲＰ遺伝子は、多くの組織において発現され、Ｔ－リンパ球の増殖に関連することが示された。

【０１６２】

本発明者らは、実施例１－（１）と同様の方法でヒト胎児脳ｃＤＮＡライブラリーから任意に選択したｃＤＮＡクローンの配列解析とデータ・ベースの検索の結果、１１２５塩基

ｃＤＮＡクローンのＧＥＮ－０７８Ｄ０５(ポリＡを含んでいない）が、ヌクレオソーム構築に関連しているヌクレオソーム構築蛋白(ＮＡＰｓ)の遺伝子であるマウスＮＡＰ－Ｉの遺伝子、マウスの部分ｃＤＮＡクローンＤＮ３８及びｈＮＲＰと有意に高い相同性を有することを解明した。

【０１６３】

該ＧＥＮ－０７８Ｄ０５は、５’領域を欠いていたので、全体のコード領域を得るために、実施例２－（２）と同様の方法で５’レース法を行なった。この５’レース法のために、下記表３に示す塩基配列のプライマーＰ１及びＰ２を用いた。

【０１６４】

【表３】

プライマー	塩基配列
プライマーＰ１	5’-TTGAAGAATGATGCATTAGGAACCAC-3’
プライマーＰ２	5’-CACTCGAGTGGCTGGATTTCAATTTCTCCAGTAG-3’

最初の５’レースの後、得られたものは配列長が１３００塩基のシングルバンドであった。この産物をｐＴ７Ｂｌｕｅ（Ｒ）Ｔ－Ｖｅｃｔｅｒに挿入し、適当なサイズのインサートを持ついくつかのクローンを選択した。

【０１６５】

２つの独立したＰＣＲ反応から得られた、５’レース・クローン、１０クローンを配列決定し、最も長いクローンＧＥＮ－０７８Ｄ０５ＴＡ１３(およそ１３００塩基長)を更に分析した。

【０１６６】

２つの重複しているｃＤＮＡクローンＧＥＮ－０７８Ｄ０５とＧＥＮ－０７８Ｄ０５ＴＡ１３の両ストランドを配列決定することによって、２つのクローンがまだ全体のコード領域をカバーしていないことを確認したので、更なる第２の５’レース法を実施した。２回目の５’レース法には、下記表４に示す塩基配列のプライマーＰ３及びＰ４を用いた。

【０１６７】

【表４】

プライマー	塩基配列
プライマーＰ３	5’-GTCGAGCTAGCCATCTCCTCTTCG-3’
プライマーＰ４	5’-CATGGGCGACAGGTTCCGAGACC-3’

２回目の５’レース法によって得られたクローンＧＥＮ－０７８Ｄ０５０８は、３００塩基長であった。このクローンは推定の開始コドンと３つの先行するイン－フレーム停止コドンを含んでいた。これらの重複している３つのクローンから２６３６塩基からなる全体のコード配列が明らかとなり、この遺伝子を脳特異的ヌクレオソーム構築蛋白(Brain-specific Nucleosome Assembly Protein:ＢＮＡＰ)遺伝子と命名した。

【０１６８】

ＢＮＡＰ遺伝子は、配列番号：２０で示される１５１８塩基のオープン・リーディング・フレームを含んでいて、これによってコードされるアミノ酸は、配列番号：１９で示されるように５０６アミノ酸残基を有し、ＢＮＡＰの全ｃＤＮＡクローンの核酸配列は、配列番号：２１で示されるとおりである。

【０１６９】

配列番号：２１で示されるように、該遺伝子の５’非コード領域は、総体的にＧＣリッチであった。開始コドンの候補配列は、塩基配列番号２６６－２６８番目、２８７－２８９番目及び３２９－３３１番目に見られた。これらの３つの配列は、すべて開始コドン周辺でよく保存されている配列を持っていた（Kozak, M., J. Biol. Chem., 266, 19867-19870 (1991)）。

【０１７０】

スキャンニング・モデルに従えば、ｃＤＮＡクローンの最初のＡＴＧ（塩基配列番号２６６－２６８番目）が開始コドンであるかもしれない。停止コドンは、塩基配列番号１７８４－１７８６番目に位置していた。

【０１７１】

３’非コード領域は全体的にＡＴリッチであり、２つのポリ・アデニレーション・シグナル(ＡＡＴＡＡＡ)は、塩基配列番号２６０６－２６１１番目と２６１０－２６１５番目にそれぞれ位置していた。

【０１７２】

最も長いオープン・リーディング・フレームは、５０６のアミノ酸をコードしている１５１８塩基からなり、ＢＮＡＰ遺伝子産物の計算された分子量は、５７，６００ダルトンであった。

【０１７３】

親水性プロットは、ＢＮＡＰが他のＮＡＰｓのように極めて親水性であることを示した。

【０１７４】

組換えＢＮＡＰの発現と精製、そしてＢＮＡＰの遺伝子配列が５’レースの段階によって３つのキメラのクローンを構成しているかもしれないと考えられる可能性を排除するために、センス・プライマーとして配列番号３２６－３５６番目の配列と、アンチ・センスプライマーとして配列番号１７５８－１７８６番目の配列とを用い、ＲＴ－ＰＣＲを実施した。

【０１７５】

その結果、単一の約１５００塩基対の産物が得られ、該配列がキメラではなく、単一転写産物であることが確認された。

【０１７６】

（２）ＢＮＡＰとＮＡＰｓとの比較

ＢＮＡＰの推定アミノ酸配列は、ｈＮＲＰと４６％の同一性と６５％の類似性を示した。

【０１７７】

推定されたＢＮＡＰ遺伝子産物は、既に報告されたＮＡＰｓにおいて見られたものとＢＮＡＰに特徴的なモチーフを持っていた。一般にＣ末端の半分はヒト及び酵母の間で強く、よく保存されていた。

【０１７８】

第１のモチーフ(ドメインＩ)は、ＫＧＩＰＤＹＷＬＩ（３０９－３１７番目のアミノ酸残基に相当）で、これはｈＮＲＰ（ＫＧＩＰＳＦＷＬＴ）及び酵母ＮＡＰ－Ｉ（ＫＧＩＰＥＦＷＬＴ）においても認められた。

【０１７９】

第２のモチーフ(ドメインII)は、ＡＳＦＦＮＦＦＳＰＰ（４３７－４４６番目のアミノ酸残基に相当）で、これはｈＮＲＰにおいては、ＤＳＦＦＮＦＦＡＰＰとして、酵母ＮＡＰ－Ｉにおいては、ＥＳＦＦＮＦＦＳＰとして表わされた。

【０１８０】

これら２つのモチーフもまたマウスＮＡＰ－ＩとＤＮ３８の推定ペプチドにおいて保存されていた。両者の保存されたモチーフは、親水性のクラスターで、４０２番目のＣｙｓもまた保存されていた。

【０１８１】

Ｎ末端の半分は酵母から哺乳動物の種にいたるまで厳格に保存されたモチーフを有しておらず、一方では、哺乳動物種の間で保存されたモチーフが見つけられた。

【０１８２】
例えば、哺乳動物特異的機能に関係しているかもしれない、ＨＤＬＥＲＫＹＡ（１３０－１３７番目のアミノ酸残基に相当）とＩＩＮＡＥＹＥＰＴＥＥＥＣＥＷ（１５０－１６４のアミノ酸残基に相当）は、厳格に保存されていた。
【０１８３】
ＮＡＰｓは、ヒストン又は他の塩基性蛋白に結合し易いと信じられている酸性のストレッチを有していた。全てのＮＡＰｓは３つの酸性のストレッチを持っていたが、それらのストレッチの位置は保存されていなかった。
【０１８４】
ＢＮＡＰはそのような３つの酸性ストレッチを有しておらず、代わりとして、９８アミノ酸残基のうち４１の酸性アミノ酸残基を含んでいるひとつの長い酸性のクラスターを有していた３つのくり返し配列（１９４－２０７番目、２０８－２２１番目及び２２２－２３５番目のアミノ酸残基に相当）があり、そのコンセンサスは、ＥｘｘＫＥｘＰＥＶＫｘＥＥＫ（ｘは保存されておらず、たいていは疎水性残基である)であった。
【０１８５】
更に、ＢＮＡＰ配列は、いくつかのＢＮＡＰ特異的モチーフを有していることが明らかとなった。即ち、４９アミノ酸残基のうち３３残基（６７％）がセリンである極端にセリンーリッチ領域(２４－７２番目のアミノ酸残基に相当)がＮ末端部分に見られた。それは核酸レベルでは不完全なＡＧＣくり返しとして見出された。
【０１８６】
このセリンーリッチ領域に続いて、１９のうち１０の塩基性アミノ酸残基を含んでいる塩基性領域(７１－８９番目のアミノ酸残基に相当）が現れた。
【０１８７】
ＢＮＡＰは核に局在すると考えられ、２つの可能性のある核局在シグナルが見られた(ＮＬＳｓ)。第一のシグナルはＢＮＡＰの塩基性ドメインに見つけられ、その配列ＹＲＫＫＲ（７５－７９番目のアミノ酸残基に相当）は、ＨＩＶ－１のＴａｔのＮＬＳ（ＧＲＫＫＲ）に類似していた。第二のシグナルは、Ｃ末端に位置しており、その配列ＫＫＹＲＫ（５０２－５０６番目のアミノ酸残基に相当）は、ＳＶ４０のラージＴ抗原のＮＬＳ（ＫＫＫＲＫ）に類似していた。これら２つの推定ＮＬＳｓがあることから、ＢＮＡＰは核に局在しているようであったが、他の塩基性クラスターがＮＬＳｓとして働くという可能性を排除できなかった。
【０１８８】
ＢＮＡＰはいくつかのリン酸化部位を有しており、リン酸化によりＢＮＡＰの活性を調節するかもしれない。
【０１８９】
（３）ノーザンブロット分析
ノーザンブロット分析は、実施例１－（２）に準じて、正常ヒト組織におけるＢＮＡＰｍＲＮＡの発現を、クローンＧＥＮ－０７８Ｄ０５ＴＡ１３(ＢＮＡＰ遺伝子の配列番号３２３から１５５８番目に相当する)をＰＣＲにより増幅し、該ＰＣＲ産物を精製し、［^{32}Ｐ］－ｄＣＴＰ（ランダムプライムドＤＮＡラベリングキット、ベーリンガーマンハイム社）により標識し、プローブとしてＭＴＮブロットに用いて試験した。
【０１９０】
ノーザンブロット分析の結果、試験した８つのヒト成人組織、心臓、脳、胎盤、肺、肝臓、骨格筋、腎臓及び膵臓のうち、脳においてＢＮＡＰの３．０キロの転写体が検出された(８時間露光)、そしてより長い露光(２４時間露光)で同じサイズのぼんやりしたバンドが心臓において検出された。
【０１９１】
ＢＮＡＰは試験した脳のいくつかの部位で等しく発現していたが、他の組織においては、７２時間露光によってもいかなるシグナルも検出されなかった。ｈＮＲＰｍＲＮＡは、試験したヒト組織においていたるところに発現されていたけれども、ＢＮＡＰｍＲＮＡは組

織特異的な手法で発現されていた。

【０１９２】

（４）ラディエーション・ハイブリッド・マッピング（Radiation Hybrid Mapping）

ラディエーション・ハイブリッド・マッピングによって、ＢＮＡＰのクローンの染色体マッピングを行なった（Cox, D.R., et al., Science, 250, 245-250 (1990)）。

【０１９３】

即ち、全ヒト・ゲノムのラディエーション・ハイブリッド・クローン(Ｇ３ＲＨ)のパネルをリサーチ・ジェネティックス社(Research Genetics, Inc., AL, U.S.A.)より購入し、製品の使用説明書に従い、表５に示す塩基配列のＡ１プライマーとＡ２プライマーを用いて染色体マッピング分析のためのＰＣＲ反応を実施した。

【０１９４】

【表５】

プライマー	塩基配列
Ａ１プライマー	5'-CCTAAAAAGTGTCTAAGTGCCAGTT-3'
Ａ２プライマー	5'-TCAGTGAAAGGGAAGGTAGAACAC-3'

得られた結果をインターネット上で使用可能なソフトウェアによって分析した（Boehnke, M., et al., Am. J. Hum. Genet., 46, 581-586 (1991)）。

【０１９５】

その結果、ＢＮＡＰ遺伝子はマーカーＤＸＳ９９０（ＬＯＤ＝１０００，ｃＲ８０００＝－０．００）に強く連鎖していた。マーカーＤＸＳ９９０は、染色体上のＸｑ２１．３－ｑ２２に局在したマーカーであることから、ＢＮＡＰは、いくつかのＸ染色体と連関する精神遅延が局在する染色体位置Ｘｑ２１．３－ｑ２２に局在されていることが確認された。

【０１９６】

ヌクレオソームの構造真核生物に特徴的で基本的な染色体の構造単位だけでなく、遺伝子発現の調節単位である。いくつかの結果から、転写活性の高い遺伝子は、ヌクレアーゼ処理に感受性があり、転写活性により染色体の構造が変化することを示唆している（Elgin, S.C.R., J. Biol. Chem., 263, 19259-19262(1988)）。

【０１９７】

ＮＡＰ－Ｉは、酵母、マウス、ヒトにおいてクローニングされており、イン・ビトロにおいてヌクレオソーム構築を促進する能力を持つ因子の一つである。実施された配列に関する研究において、特異抗体４Ａ８のエピトープを含んでいるＮＡＰｓがヒト、マウス、カエル、ショウジョウバエ、酵母(サッカロミセス・セレビシェ:S. cerevisiae)において検出された（Ishimi, Y., et al., Eur. J. Biochem., 162, 19-24 (1987)）。

【０１９８】

これらの実験においてＮＡＰｓは、ＳＤＳ－ＰＡＧＥ分析によって５０－６０ｋＤａの間に電気泳動されたが、組換えＢＮＡＰは、およそ８０ｋＤａの位置にゆっくりと電気泳動された。４Ａ８のエピトープは、２番目のよく保存された疎水性モチーフの中に局在されていることが示された。そして、ＦＮＦのトリプレットが、エピトープの一部として重要であることが同時に示された（Fujii-Nakata T., et al., J. Biol. Chem., 267, 20980-20986 (1992)）。

【０１９９】

ＢＮＡＰもまたこの一致モチーフをドメインIIに含んでいた。ドメインIIは、著しく疎水的である事実とドメインIIが免疫システムによって認識されることができる事実は、それがＢＮＡＰ表面でおそらく提示され、蛋白－蛋白相互作用に関与する可能性を持つことが

提言される。
【０２００】
ドメインＩもまた蛋白－蛋白相互作用に関与しているかもしれない。保存された領域の機能的な意味はいまだに不明であるが、これらのＮＡＰｓ間の比較的低いが総体的に保存されていることを考えると、それらはヌクレオソーム構築のためには本質的であらねばならないと考えられる。
【０２０１】
ｈＮＲＰ遺伝子は、甲状腺、胃、腎臓、腸、白血病、肺癌、乳癌などに発現される（Simon, H. U., et al., Biochem. J., 297, 389-397 (1994)）。一方、ＮＡＰｓは、いたるところに発現されていて、基本的なヌクレオソーム形成において重要な役割を演じていると考えられる。
【０２０２】
ＢＮＡＰは、脳特異的ヌクレオソーム形成に関与するかもしれず、その不全は、逸脱した神経細胞の機能の結果として神経学上の疾患や精神的な遅延を起こすかもしれない。
【０２０３】
ＢＮＡＰは、いくつかのＸ染色体と連鎖した精神的な遅延が、局在されるＸ染色体ｑ２１．３－ｑ２２上のマーカーに対して強く連鎖していた。このＸ染色体の中心周囲領域は、α－サラセミアや精神的遅延(ＡＴＲ－Ｘ)のような、精神的な遅延に対して責任ある遺伝子が多かった（Gibbons, R. J., et al., Cell, 80, 837-845 (1995)）。ヌクレオソーム構造を調節すると思われるＡＴＲ－Ｘ遺伝子の分析と同様に、本発明者らは、ＢＮＡＰがＸ染色体と連鎖した精神遅延のあるタイプに関係しているかもしれないと考える。
【０２０４】
本実施例によれば、新規なＢＮＡＰ遺伝子が提供され、該遺伝子を用いれば、該遺伝子の各種組織での発現の検出や、ＢＮＡＰ蛋白の遺伝子工学的製造が可能となり、これらにより、前述したように多様で且つ細胞に基本的な活動に深く関与している脳ヌクレオソーム形成の研究や脳神経細胞の機能の研究及びこれらが関与する各種神経学上の疾患や精神遅延の診断などを行なうことができ、これらの治療及び予防薬のスクリーニングや評価などをも行なうことができる。
【０２０５】
【実施例７】
ヒト骨格筋特異的ユビキチン結合酵素遺伝子（ＵＢＥ２Ｇ遺伝子）
ユビキチン・システムは、細胞プロセスに必須の酵素グループのひとつであり、酵母からヒトにいたるまで保存されている。該システムは、ユビキチン活性酵素(ubiquitin-activating enzymes: UBAs)、ユビキチン結合酵素(ubiquitin-conjugating enzymes: UBCs)、ユビキチン蛋白リガーゼ(ubiquitin protein ligases: UBRs)及び２６Ｓプロテアソーム粒子から構成されている。
【０２０６】
ユビキチンは、上記ＵＢＡから数種のＵＢＣに移されて活性化される。ＵＢＣは、ＵＢＲの関与を受けるものと、もしくは受けないで、ユビキチンを目標蛋白にトランスファーする。これらユビキチン化された目標蛋白は、多くの細胞応答、例えば蛋白分解、蛋白修飾、蛋白輸送、ＤＮＡ修復、細胞サイクル制御、複製制御、ストレス応答などや免疫応答を引き起こすとされている（Jentsch, S., et al., Biochim. Biophys. Acta, 1089, 127-139 (1991): Hershko, A. and Ciechanover, A., Annu. Rev. Biochem., 61, 761-807 (1992) : Jentsch, S., Annu. Rev. Genet., 26, 179-207 (1992): Ciechanover, A., Cell, 79, 13-21 (1994)）。
【０２０７】
ＵＢＣｓは、このシステムの重要な構成成分であり、異なる基質特異性を有し、異なった機能を調節すると思われる。例えば、サッカロミセス・セレビシェ(S. cerevisiae)のＵＢＣ７は、カドミウムによって誘導され、カドミウム中毒に対する抵抗性に関与していると思われる（Jungmann, J., et al., Nature, 361, 369-371 (1993)）。ＭＡＴ－α２の

分解もまたＵＢＣ７とＵＢＣ６によって実行される（Chen, P., et al., Cell, 74, 357-369 (1993)）。

【０２０８】

本実施例で得られた新規な遺伝子は、酵母細胞周期遺伝子ＵＢＣ９とＨｕｓ５に対して高い相同性を持っていて、ヒト骨格筋において強く発現するヒトＵＢＣ７様遺伝子であり、以下にそのクローニング及びＤＮＡのシークェンシングについて述べる。

【０２０９】

（１）ヒト骨格筋特異的ユビキチン結合酵素遺伝子（ＵＢＥ２Ｇ遺伝子）のクローニング及びＤＮＡシークエンシング

実施例１－（１）と同様の方法でヒト胎児脳ｃＤＮＡライブラリーから任意に選択したｃＤＮＡクローンの配列解析とデータ・ベースの検索の結果、ＧＥＮ－４２３Ａ１２ｃＤＮＡクローンが様々な種のユビキチン結合酵素(ubiquitin-conjugating enzymes: UBCs)をコードする遺伝子と有意に高い相同性を有する持つことが分かった。

【０２１０】

該ＧＥＮ－４２３Ａ１２は、５’側が欠けていたので、全体のコード領域を得るために実施例２－（２）と同様の方法で５’レース法を行なった。

【０２１１】

５’レース法のために、表６に示す塩基配列のＰ１プライマー及びＰ２プライマーを用いた。

【０２１２】

【表６】

プライマー	塩基配列
Ｐ１プライマー	5’-TAATGAATTTCATTTTAGGAGGTCGG-3’
Ｐ２プライマー	5’-ATCTTTTGGGAAAGTAAGATGAGCC-3’

５’レース産物をｐＴ７Ｂｌｕｅ（Ｒ）Ｔ－Ｖｅｃｔｅｒに挿入し、適当なサイズのインサートを持ついくつかのクローンを選択した。

【０２１３】

２つの独立したＰＣＲ反応から得られ、５’レース・クローンの４つのクローンが同じ配列を持っているが異なった長さであった。

【０２１４】

上記クローン配列決定によって、新しい遺伝子の翻訳配列と隣接する５’－と３’－フランキング配列を決定した。

【０２１５】

その結果、全配列が６１７塩基の新しい遺伝子が明らかとなり、この遺伝子をヒト骨格筋特異的ユビキチン結合酵素遺伝子（ＵＢＥ２Ｇ遺伝子）と命名した。

【０２１６】

また、この配列がキメラ・クローンである可能性を排除するために、ヒト胎児脳ｃＤＮＡライブラリーから増幅するために前記実施例６－（１）に準じて、センス・プライマーを用いＲＴ－ＰＣＲを実施した。その結果、単一のＰＣＲ産物が得られ、その配列がキメラでないことが確認された。

【０２１７】

ＵＢＥ２Ｇ遺伝子は、配列番号：２３で示される５１０塩基のオープン・リーディング・フレームを含んでいて、これによってコードされるアミノ酸は、配列番号：２２で示されるように１７０アミノ酸残基を有し、ＵＢＥ２Ｇの全ｃＤＮＡクローンの核酸配列は、配列番号：２４で示されるとおりである。

【０２１８】
配列番号：２４で示されるように推定された開始コドンは、ｃＤＮＡクローンの最初のＡＴＧトリプレットである塩基配列番号１９－２１番目に位置していた。このＡＴＧに先行したイン・フレーム停止コドンがなかったので、このクローンは、以下の理由から全体のオープン・リーディング・フレームを含んでいると演繹した。
【０２１９】
即ち、（ａ）サッカロミセス・セレビシェ（S. cerevisiae）ＵＢＣ７とアミノ酸配列の相同性が高く、酵母のＵＢＣ７の開始コドンと一致しておりこのＡＴＧを支持した。
【０２２０】
（ｂ）ＡＧＧＡＴＧＡの配列は、開始コドン周辺のコンセンサス配列（Ａ／Ｇ）ＣＣＡＴＧＧ（Kozak, M., J. Biol. Chem., 266, 19867-19870 (1991)）に類似していた。
【０２２１】
（２）ＵＢＥ２ＧとＵＢＣｓとのアミノ酸配列との比較
ＵＢＥ２ＧとＵＢＣｓとのアミノ酸配列の比較から、ユビキチンに結合することができる活性部位システィンは、９０番目のシスティン残基であるように思えた。これらの遺伝子でコードされるペプチドは、同じファミリーに属しているようである。
【０２２２】
（３）ノーザンブロット分析
ノーザンブロット分析は、実施例１－（２）に準じて、正常ヒト組織におけるＵＢＥ２ＧｍＲＮＡの発現を、ＵＢＥ２Ｇの全配列をＰＣＲにより増幅し、該ＰＣＲ産物を精製し、［^{32}Ｐ］－ｄＣＴＰ（ランダムプライムドＤＮＡラベリングキット、ベーリンガーマンハイム社）により標識し、プローブとして用い、メンブレンは、ＭＴＮブロットを用いて実施した。
【０２２３】
ノーザンブロット分析の結果、試験した１６のヒト成人組織、心臓、脳、胎盤、肺、肝臓、骨格筋、腎臓、膵臓、脾臓、甲状腺、膀胱、睾丸、卵巣、小腸、大腸及び末梢血白血球において、１８時間の露光の後、４．４キロ塩基対、２．４キロ塩基対、と１．６キロ塩基対の転写体が検出可能であった、そして骨格筋においてこれらの転写体の強い発現が観察された。
【０２２４】
（４）ラディエーション・ハイブリッド・マッピング（Radiation Hybrid Mapping）
ラディエーション・ハイブリッド・マッピングによって前記実施例６－（４）と同様の方法でＵＢＥ２Ｇのクローンの染色体マッピングを行なった。
【０２２５】
尚、染色体マッピング分析のためにＰＣＲ反応のために用いたＣ１プライマー及びＣ４プライマーは、配列番号：２４の配列の配列番号４１５－４３５番目及び配列番号５０９－５２８番目にそれぞれ相当しており、その塩基配列は下記表７に示す通りである。
【０２２６】
【表７】

プライマー	塩基配列
Ｃ１プライマー	5'-GGAGACTCACCTGCTAATGTT-3'
Ｃ４プライマー	5'-CTCAAAAGCAGTCTCTTGGC-3'

その結果、ＵＢＥ２Ｇ遺伝子はマーカーＤ１Ｓ４４６（ＬＯＤ＝１２．５２，ｃＲ８００＝８．６０）とＤ１Ｓ２３５（ＬＯＤ＝９．１４，ｃＲ８０００＝２２．４６）に連鎖していた。これらのマーカーは、染色体バンドｌｑ４２．１３－ｑ４２．３に局在した。

【０２２７】
ＵＢＥ２Ｇは、骨格筋において強く発現し、試験した全ての組織において非常に弱く発現された。全ての他のＵＢＣｓは細胞サイクルの調節のような必須の細胞機能に関係しており、それらＵＢＣｓは組織のいたるところに発現している。しかしながら、ＵＢＥ２Ｇの発現パターンは、筋肉特異的な役割が提言されるかもしれない。
【０２２８】
３つの異なるサイズの転写体が検出されたが、それらのどれがｃＤＮＡクローンに相当するものであるかを確認することはできなかった。ＵＢＥ２Ｇ産物の主要な構造は、酵母のＵＢＣ７に対して強い相同性を示した。また、線虫類のＵＢＣ７は、一方では酵母ＵＢＣ７に対して強い相同性を示した。それは、レプレッサーの分解に関連しており、更に酵母におけるカドミウムに対する抵抗を与える。これらの蛋白間の類似性から、それらが同じファミリーに属していることが提言される。
【０２２９】
ＵＢＥ２Ｇは、筋特異的蛋白の分解に関与しており、該遺伝子の不全は、結果として筋ジストロフィーのような疾患になると考えられる。近年、別の蛋白質分解酵素、カルパイン３（Calpain3）が肢体筋ジストロフィー・タイプ２Ａの原因であると証明された（Richard, I., et al.,Cell, 81, 27-40 (1995)）。該遺伝子がＵＢＥ２Ｇの染色体の位置からいかなる遺伝性筋疾患とも現在のところ有意な関連はないが、今後のリンケージ解析から、該遺伝子との関係が見つかる可能性があると思われる。
【０２３０】
本実施例によれば、新規なＵＢＥ２Ｇ遺伝子が提供され、該遺伝子を用いれば、該遺伝子の各種組織での発現の検出や、ＵＢＥ２Ｇ蛋白の遺伝子工学的製造が可能となり、これらにより、前述したように多様で且つ細胞に基本的な活動に深く関与している筋特異的蛋白分解の研究や骨格筋の機能の研究及びこれらが関与する各種筋疾患、例えば筋ジストロフィーの診断などを行なうことができ、これらの治療及び予防薬のスクリーニングや評価などをも行なうことができる。
【０２３１】
【実施例８】
ＴＭＰ－２遺伝子
（１）ＴＭＰ－２遺伝子のクローニング及びＤＮＡシークエンシング
実施例１－（１）と同様の方法でヒト胎児脳ｃＤＮＡライブラリーから任意に選択したｃＤＮＡクローンの配列解析とデータ・ベースの検索の結果、膜蛋白遺伝子(transmembrane protein:アクセッションＮｏ．；Ｕ１９８７８)に相同性の高いｃＤＮＡ配列を有するクローン（ＧＥＮ－０９２Ｅ１０）を見出した。
【０２３２】
しかして従来、膜蛋白遺伝子は、カエル(Xenopus laevis)とヒトにおいてクローニングされており、之等はフォリスタチン・モデュル(follistatin module)とＥＧＦ領域(epidermal growth factor domain)を有する細胞膜貫通型のタンパク質の遺伝子であると考えられる（アクセッションＮｏ．；Ｕ１９８７８)。
【０２３３】
上記タンパク質遺伝子の配列情報から、ＧＥＮ－０９２Ｅ１０クローンは、５’領域を欠いていたので、λｇｔ１０ｃＤＮＡライブラリー（Human Fetal Brain 5'-STRETCH PLUS cDNA; クローンテック社製）を、ＧＥＮ－０９２Ｅ１０クローンをプローブとしてスクリーニングすることにより、さらに５’上流を含むｃＤＮＡクローンを単離した。
【０２３４】
このｃＤＮＡクローンの両ストランドを配列決定することによって、全体のコード領域をカバーしている配列が明らかになり、この遺伝子をＴＭＰ－２遺伝子と命名した。
【０２３５】
ＴＭＰ－２遺伝子は、配列番号：２６で示される１１２２塩基のオープン・リーディング・フレームを含んでいて、これによってコードされるアミノ酸は、配列番号：２５で示さ

れるように３７４アミノ酸残基を有し、ＴＭＰ－２の全ｃＤＮＡクローンの核酸配列は、配列番号：２７で示されるとおり、１７２１塩基からなっていた。

【０２３６】

配列番号：２７で示されるように５’非コード領域は、総体的にＧＣリッチであった。開始コドンの候補配列は、いくつか存在したが、スキャンニング・モデルに従えば、ｃＤＮＡクローンの５番目のＡＴＧ（塩基配列番号３６８－３７０番目）が開始コドンと推定された。停止コドンは、塩基配列番号１４９０－１４９２番目に位置していた。ポリ・アデニレーション・シグナル(ＡＡＴＡＡＡ)は、塩基配列番号１７０３－１７０８番目位置していた。ＴＭＰ－２遺伝子産物の計算された分子量は、４１，４００ダルトンであった。

【０２３７】

上記したように膜蛋白遺伝子は、フォリスタチン・モデュルとＥＧＦ領域を有しているが、本発明の新規なヒト遺伝子においても、これらのモチーフが保存されていた。

【０２３８】

また、本発明のＴＭＰ－２遺伝子は、ＴＧＦ－α（transforming growth factor-α; Derynck, R., et al., Cell, 38, 287-297 (1984)）、beta-cellulin（Igarashi, K. and Folkman, J., Science, 259, 1604-1607 (1993)）、ヘパリン結合ＥＧＦ様成長因子（heparin-binding EGF-like growth factor; Higashiyama, S., et al., Science, 251, 936-939 (1991)）及びシュワノマ誘導成長因子（Schwannoma-derived growth factor; Kimura, H., et al., Nature, 348, 257-260 (1990)）と、ＥＧＦ領域を超えてアミノ酸レベルで相同性を有することなどから、細胞の増殖や細胞間のコミュニケーションに重要な役割を果たすと考えられる。

【０２３９】

（２）ノーザンブロット分析

ノーザンブロット分析は、実施例１－（２）に準じて、正常ヒト組織におけるＴＭＰ－２ｍＲＮＡの発現を、クローンＧＥＮ－０９２Ｅ１０をＰＣＲにより増幅し、該ＰＣＲ産物を精製し、［^{32}Ｐ］－ｄＣＴＰ（ランダムプライムドＤＮＡラベリングキット、ベーリンガーマンハイム社）により標識し、プローブとしてＭＴＮブロットを用いて実施した。

【０２４０】

ノーザン分析の結果、脳と前立腺において高い発現が検出された。該ＴＭＰ－２遺伝子のｍＲＮＡのサイズは、約２ｋｂであった。

【０２４１】

本発明によれば、新規なヒトＴＭＰ－２遺伝子が提供され、該遺伝子を用いれば、該遺伝子の各種組織での発現の検出や、ヒトＴＭＰ－２蛋白の遺伝子工学的製造が可能となり、これらにより、前述したように細胞増殖や細胞間コミュニケーションに深く関与している脳腫瘍や前立腺癌の研究及びこれらの疾患の診断等を行なうことができ、これらの治療及び予防薬のスクリーニングや評価等をも行なうことができる。

【０２４２】

【実施例９】

ヒトＮＰＩＫ遺伝子

（１）ヒトＮＰＩＫ遺伝子のクローニング及びＤＮＡシークエンシング

実施例１及び２と同様の方法で、ヒト胎児脳ｃＤＮＡライブラリーから任意に選択したｃＤＮＡクローンの配列解析とデータ・ベースの検索の結果、フォスファチジルイノシトール(phosphatidylinositol)３及び４キナーゼに保存されているアミノ酸配列をコードする遺伝子（Kunz, J., et al., Cell, 73, 585-596 (1993)）に高い相同性を有する２つのｃＤＮＡクローンを得、これらをＧＥＮ－４２８Ｂ１２ｃ１及びＧＥＮ－４２８Ｂ１２ｃ２と命名し、これらの全配列を、前記各実施例に準じて決定した。

【０２４３】

その結果、ＧＥＮ－４２８Ｂ１２ｃ１ｃＤＮＡクローンとＧＥＮ－４２８Ｂ１２ｃ２クローンは、５’末端の１２のアミノ酸残基のみが異なるコード配列を有することが判明し、ＧＥＮ－４２８Ｂ１２ｃ１ｃＤＮＡクローンが１２アミノ酸残基長かった。

【０２４４】
ヒトＮＰＩＫ遺伝子のＧＥＮ－４２８Ｂ１２ｃ１ｃＤＮＡ配列は、配列番号：３２で示される２４８７塩基のオープン・リーディング・フレームを含んでいて、これによってコードされるアミノ酸は、配列番号：３１で示されるように８２９アミノ酸残基を有し、全長ｃＤＮＡクローンの核酸配列は、配列番号：３３で示されるとおり、３３２４塩基からなっていた。
【０２４５】
配列番号：３３で示されるように推定された開始コドンは、ｃＤＮＡクローンの２番目のＡＴＧトリプレットである塩基配列番号１１５－１１７番目に位置していた。また、停止コドンは塩基配列番号の２６０２－２６０４番目に位置していた。ポリ・アデニレーション・シグナル(ＡＡＴＡＡＡ)は塩基配列番号３３０５－３３１０番目位置していた。
【０２４６】
一方、ヒトＮＰＩＫ遺伝子のＧＥＮ－４２８Ｂ１２ｃ２ｃＤＮＡ配列は、配列番号：２９で示される２４５１塩基のオープン・リーディング・フレームを含んでいて、これによってコードされるアミノ酸は、配列番号：２８で示されるように８１７アミノ酸残基を有し、全長ｃＤＮＡクローンの核酸配列は、配列番号：３０で示されるとおり、３６０２塩基からなっていた。
【０２４７】
配列番号：３０で示されるように推定された開始コドンは、ｃＤＮＡクローンの７番目のＡＴＧトリプレットである塩基配列番号４２９－４３１番目に位置していた。また、停止コドンは塩基配列番号の２８８０－２８８２番目に位置していた。ポリ・アデニレーション・シグナル(ＡＡＴＡＡＡ)は塩基配列番号３５８３－３５８８番目位置していた。
【０２４８】
（２）ノーザンブロット分析
ノーザンブロット分析は、実施例１－（２）に準じて、正常ヒト組織におけるヒトＮＰＩＫｍＲＮＡの発現を、ヒトＮＰＩＫの全配列をＰＣＲにより増幅し、該ＰＣＲ産物を精製し、［^{32}Ｐ］－ｄＣＴＰ（ランダムプライムドＤＮＡラベリングキット、ベーリンガーマンハイム社）により標識し、プローブとして用い、メンブレンは、ＭＴＮブロットを用いて実施した。
【０２４９】
ノーザン分析の結果、ヒトＮＰＩＫ遺伝子は、試験した１６のヒト成人組織の様々な臓器でその発現が認められ、約３．８ｋｂと約５ｋｂの転写体が検出可能であった。
【０２５０】
ＧＥＮ－４２８Ｂ１２ｃ２ｃＤＮＡ配列の開始コドンを含む、下記表８に示す塩基配列のプライマーＡと終止コドンを含むプライマーＢとを用いてHuman Fetal Brain Marathon-Ready cDNA（クロンテック社製）をテンプレートとしてＰＣＲを行ない、該ＰＣＲ産物の塩基配列を決定した。
【０２５１】
【表８】

プライマー	塩基配列
プライマーＡ	5'-ATGGGAGATACAGTAGTGGAGC-3'
プライマーＢ	5'-TCACATGATGCCGTTGGTGAG-3'

その結果、発現されたヒトＮＰＩＫｍＲＮＡには、ＧＥＮ－４２８Ｂ１２ｃ１ｃＤＮＡ配列の配列番号：３３の塩基配列番号１０６０－１１０４番目(アミノ酸残基の配列番号：３１のアミノ酸配列番号３１６－３３０番目)を欠くもの及び配列番号：３３の塩基配列

番号１８９７－１９１１番目(アミノ酸残基の配列番号：３１のアミノ酸配列番号５９５－５９９番目)を欠くものがあることが判明した。

【０２５２】

更にこの遺伝子には、ＧＥＮ－４２８Ｂ１２ｃ１ｃＤＮＡ配列の配列番号：３３の塩基配列番号１９４１－１９６６番目の領域に、下記表９に示されるポリモルフィズムが存在し(428B12c1.fasta)、これによりＧＥＮ－４２８Ｂ１２ｃ１アミノ酸残基の配列番号：３１のアミノ酸配列番号６１０－６１８番目のＩＱＤＳＣＥＩＴＴがＹＫＩＬＶＩＳＡに変異した変異体タンパク質がコードされることが判明した。

【０２５３】

【表９】

```
                                  1930       1940       1950      1959
                                    TGGATCAAGCCAATACAAGATTCTTGTGAA
                                    |||||||||||| ||||||||||||||||
    TCCATTTGGGAACAGGAGCGAGTGCCCCTTTGGATCAAGCC-ATACAAGATTCTTGTG--
 1900       1910       1920       1930       1940       1950

 1960       1970       1980
    ATTACGACTGATAGTGGCATG
    ||| || ||||||||||||||
    ATTTCGGCTGATAGTGGCATGATTGAACCAGTGGTCAATGCTGTGTCCATCCATCAGGTG
    1960       1970       1980       1990       2000       2010
```

（３）ＦＩＳＨによるヒトＮＰＩＫ遺伝子の染色体のマッピング

実施例１－（３）に準じてＦＩＳＨによるヒトＮＰＩＫ遺伝子の染色体のマッピングを行なった。

【０２５４】

その結果、ヒトＮＰＩＫ遺伝子は、染色体の座位が１ｑ２１．１－ｑ２１．３の位置にあることが判明した。

【０２５５】

本発明の新規なヒト遺伝子であるヒトＮＰＩＫ遺伝子は、上記したように、５’領域の異なる２つのｃＤＮＡが存在し、それらは、それぞれ８２９と８１７のアミノ酸をコードしうる領域を含んでいた。また、この遺伝子のｍＲＮＡに２つの欠失しうる部位が存在すること並びにＧＥＮ－４２８Ｂ１２ｃ１アミノ酸残基の配列番号：３１のアミノ酸配列番号６１０－６１８番目に相当する特定領域にポリモルフィズムが存在し、変異タンパク質がコードされることから、ヒトＮＰＩＫには、いくつかの組合わせから成るヒトＮＰＩＫ、ヒトＮＰＩＫ欠失体、ヒトＮＰＩＫ変異体及び／又はヒトＮＰＩＫ変異・欠失体の存在することが考えられる。

近年、上記ＰＩ３及び４キナーゼのファミリーに属する数種類のタンパク質に、タンパク質キナーゼ活性があることが示された（Dhand, R., et al., EMBO J., 13, 522-533 (1994): Stack, J. H. and Emr, S. D., J. Biol. Chem., 269, 31552-31562 (1994): Hartley, K. O., et al., Cell, 82, 849-856 (1995)）。

【０２５６】

また、このファミリーに属するタンパク質がＤＮＡの損傷の修復に関与していること（Hartley, K. o., et al., Cell, 82, 849-856 (1995)）、アタキシア（運動失調症）の原因遺伝子であることが明らかとなった（Savitsky, K., et al., Science, 268, 1749-1753 (1995)）。

【０２５７】

これらＰＩキナーゼのファミリーと高い相同性を有するヒトＮＰＩＫ遺伝子でコードされるタンパク質は、脂質あるいはタンパク質をリン酸化する新しい酵素であることが予想で

きる。

【０２５８】

本実施例によれば、新規ヒトＮＰＩＫ遺伝子が提供され、該遺伝子を用いれば、該遺伝子の各種組織での発現の検出や、ヒトＮＰＩＫ蛋白の遺伝子工学的製造が可能となり、これらにより、前述したように脂質あるいはタンパク質のリン酸化酵素の研究、ＤＮＡの修復の研究及びこれらが関与する疾患、例えば癌の研究や診断等を行なうことができ、これらのの治療及び予防薬のスクリーニングや評価等をも行なうことができる。

【０２５９】

（４）融合蛋白発現ベクターの構築

まず、ヒトＮＰＩＫ遺伝子（ＧＥＮ－４２８Ｂ１２ｃ２）のコード領域をサブクローニングするために、上記（１）で得られたＤＮＡ配列情報を基に、下記表１０に示す配列の２つのプライマー、Ｃ１プライマー及びＣ２プライマーを作製した。

【０２６０】

【表１０】

プライマー	塩基配列
Ｃ１プライマー	5′-CTCAGATCTATGGGAGATACAGTAGTGGAGC-3′
Ｃ２プライマー	5′-TCGAGATCTTCACATGATGCCGTTGGTGAG-3′

尚、上記Ｃ１及びＣ２プライマーは、共にＢｇｌIIサイトを含んでおり、Ｃ２プライマーはアンチセンスプライマーであった。

【０２６１】

上記各プライマーを利用して、ヒト胎児脳ｍＲＮＡから誘導されたｃＤＮＡを用いてＰＣＲを行ない、ｃＤＮＡを増幅させて、およそ２５００塩基長さの産物を得た。増幅させたｃＤＮＡをエタノール沈殿させた後、ｐＴ７BlueＴ－Vector（Novagen）社製に挿入し、サブクローニングした。この全配列を前記各実施例に準じて決定したところ、これは上記表９に示されるポリモルフィズムを有するものであった。

【０２６２】

該ｃＤＮＡをＢｇｌIIにより切断し、切断された産物をアガロースゲル電気泳動後、アガロースゲルより切り出し、GENECLEAN II KIT (BIO 101社)を用いて、ＤＮＡを回収した。回収されたｃＤＮＡをpBlueBacHis2Bベクター（Invitrogen社）のＢｇｌII部位に挿入し、サブクローニングした。

【０２６３】

得られた融合ベクターは、ＢｇｌII切断部位を有し、目的の遺伝子産物（およそ９１ｋｄ）とpBlueBacHis2Bベクター由来のポリヒスチジン領域及びAnti－Ｘpress™抗体（Invitrogen社）の認識エピトープを含む３８アミノ酸領域との融合蛋白を発現するベクターである。

【０２６４】

（５）昆虫細胞Ｓｆ－９へのトランスフェクション

本遺伝子の発現は、バキュロウイルス発現系を用いて行なった。バキュロウイルスは、環状２本鎖の昆虫病原性ウイルスであり、昆虫の細胞の中で多角体と呼ばれる封入体を大量に作ることができる。本発現系は、その特徴を生かして開発されたInvitrogen社のBac-N-Blue™ Transfection Kitを用いて以下の通り行なった。

【０２６５】

より詳しくは、ヒトＮＰＩＫをコードする領域を含むｐＢｌｕｅＢａｃＨｉｓ２Ｂ４μｇとＢａｃ－Ｎ－Ｂｌｕｅ™ＤＮＡ（Invitrogen社）１μｇをInsectin™Liposome(Invitrogen社)の存在下で、Ｓｆ－９細胞にコトランスフェクト（co-transfect）した。

【０２６６】

Ｂａｃ－Ｎ－ｂｌｕｅ™ＤＮＡには、ＬａｃＺ遺伝子が、ＢｌｕｅＢａｃＨｉｓ２Ｂと相同組換えを起こしたときにのみ発現するように組み込んであり、寒天下で培養すれば、青のプラークを選択することによって目的遺伝子を発現するウイルスを簡単に選択できるようにデザインされている。

【０２６７】

青のプラークを寒天ごと切り出し、培地４００μｌに懸濁させてウイルスを培地に拡散させた。遠心してウイルスを含む上清を得た後、Ｓｆ－９細胞に再度感染させ、力価を上昇させ、また多量のウイルス感染液を得た。

【０２６８】

（６）ヒトＮＰＩＫの調製

目的のヒトＮＰＩＫが発現しているか否かは、ウイルスを感染させてから３日後の細胞を用いて以下の通り確認した。即ち、細胞を回収してＰＢＳで洗浄し、ＳＤＳ－ＰＡＧＥローディング緩衝液と共に５分間煮沸してＳＤＳ－ＰＡＧＥを行なった後、Anti－Xpress抗体を用いたウエスタンブロッティング法によって、推定分子量位置に目的蛋白質を確認した。尚、ウイルス非感染の細胞を対照とした同試験では目的とするＮＰＩＫのバンドは認められなかった。

【０２６９】

より詳しくは、Ｓｆ－９細胞の７０乃至８０％集合した（semi-confluent）１７５ｃｍ^{2}フラスコ１５枚を用いて、大規模にウイルスを感染させて、目的蛋白質を発現させた。ウイルス感染３日後の細胞を回収し、ＰＢＳで洗浄後、細胞を緩衝液（２０ｍＭトリス／ＨＣｌ（ｐＨ７．５）、１ｍＭ　ＥＤＴＡ及び１ｍＭＤＴＴ）に懸濁させ、４℃下、３０秒間隔で３０秒ずつ４回超音波処理して破壊溶解させ、得られる細胞破砕物を遠心分離後、上清を回収し、上清中の蛋白質をAnti－Xpress抗体を用いて免疫沈降させ、プロティンＡセファロースビーズのスラリーとして得た。該スラリーをＳＤＳ－ＰＡＧＥローディング緩衝液と共に５分間煮沸し、ＳＤＳ－ＰＡＧＥによってＮＰＩＫの回収の確認と定量を行なった。尚、上記スラリーはそのままアッセイ系に供した。

【０２７０】

（７）ＰＩ４キナーゼ活性の確認

ＮＰＩＫの予想される活性は、ホスファチジルイノシトール（ＰＩ）のイノシトール環の４位にリン酸を付加する活性、即ちＰＩ４キナーゼ活性である。

【０２７１】

アッセイは、竹縄らの方法〔Yamakawa,A. and Takenawa,T., J.Biol.Chem., 263, 17555-17560 (1988)〕を用いて以下の通り検討した。即ち、ＮＰＩＫスラリー１０μｌ（２０ｍＭトリス／ＨＣｌ（ｐＨ７．５）、１ｍＭ　ＥＤＴＡ、１ｍＭ　ＤＴＴ及び５０％プロテインＡビーズ）、ＰＩ溶液（ＰＩ含有市販クロロホルム溶液５ｍｇをガラス容器に入れ、窒素気流下で乾固させ、２０ｍＭトリス／ＨＣｌ緩衝液（ｐＨ７．５）１ｍｌに溶解させた後、超音波でミセル化したもの）１０μｌ、反応用緩衝液（２１０ｍＭトリス／ＨＣｌ（ｐＨ７．５）、５ｍＭＥＧＴＡ及び１００ｍＭ　$MgCl_2$）１０μｌ及び蒸留水１０μｌを入れ、最後にＡＴＰ溶液（５００μＭ　ＡＴＰ５μｌ、蒸留水４．９μｌ及びγ－^{32}ＰＡＴＰ（６０００Ｃｉ／ｍｍｏｌ、ＮＥＮ社）０．１μｌ）１０μｌを入れて、３０℃にて反応を開始させた。反応は、２、５、１０及び２０分間行ない、後記アッセイにて、リン酸のＰＩへの取り込みが直線的に増加する領域にある１０分をアッセイのインキュベーションタイムと設定した。

【０２７２】

反応終了後、ＰＩを溶媒抽出法にて分画し、最終的にクロロホルムに再懸濁させ、薄層クロマトグラフィー（ＴＬＣ）にて展開し、反応生成物ＰＩ４Ｐの位置の放射活性をBio-Image analyzer（Fuji film社）を用いて定量化した。

【０２７３】

結果を図１に示す。図１は、上記ＴＬＣからの放射活性の定量結果を解析した図であり、右のレーン（２）はＮＰＩＫを含むウイルスを感染させたＳｆ－９細胞の細胞質画分を、

左のレーン（１）は感染していない同細胞質画分を示し、矢印のスポットがＰＩ４Ｐを示す。

【０２７４】

また、上記において、反応系にトリトン－Ｘ１００又はアデノシンの所定量を添加してＮＰＩＫのＰＩ４キナーゼ活性に及ぼす之等の影響を同様にして調べた。得られた結果は図２に示す通りであり、ＮＰＩＫは、トリトン－Ｘ１００によって活性化され、該トリトン－Ｘ１００によって活性化された活性がアデノシンによって阻害される、典型的なＰＩ４キナーゼ活性を有することが判った。

【０２７５】

尚、図２において、（１）は無添加を、（２）はトリトン－Ｘ１００の０．４％添加を、（３）はトリトン－Ｘ１００の０．４％＋アデノシンの２５μM添加を、（４）はトリトン－Ｘ１００の０．４％＋アデノシンの５０μM添加を、（５）はトリトン－Ｘ１００の０．４％＋アデノシンの１００μM添加を、（６）はトリトン－Ｘ１００の０．４％＋アデノシンの２００μM添加を、それぞれ示す。

【０２７６】

【実施例１０】

ｎｅｌ関連蛋白タイプ１(ＮＲＰ１）遺伝子及びｎｅｌ関連蛋白タイプ２(ＮＲＰ２）遺伝子

（１）ＮＲＰ１遺伝子及びＮＲＰ２遺伝子のクローニング及びＤＮＡシークエンシング

ＥＧＦ様反復配列は、多くの膜蛋白質において確認されており、成長調節と分化に関連する蛋白中において確認されている。このモチーフは蛋白間相互作用に関係しているように思われている。

【０２７７】

近年５つのＥＧＦ様反復配列を含んでいる新規なペプチド、ｎｅｌをコードする遺伝子がニワトリ胚のｃＤＮＡライブラリーからクローニングされた（Matsuhashi, S., et al., Dev. Dynamics., 203, 212-222(1995)）。この産物は、細胞外ドメインにそのＥＧＦ様反復配列を持っている膜分子であると考えられている。ｎｅｌｍＲＮＡの４．５ｋｂの転写体が胚期において様々な組織において発現され、生後は、脳と網膜にもっぱら発現されている。

【０２７８】

実施例１－（１）と同様の方法でヒト胎児脳ｃＤＮＡライブラリーから任意に選択したｃＤＮＡクローンの配列解析とデータ・ベースの検索の結果、上記ｎｅｌと有意に相同性の高い２つのｃＤＮＡクローンを見出し、それぞれＧＥＮ－０７３Ｅ０７とＧＥＮ－０９３Ｅ０５と名付けた。

【０２７９】

両クローンとも５’領域を欠いていたので、全体のコード領域を得るために実施例２－（２）と同様の方法で５’レース法を行なった。

【０２８０】

５’レース法のためのプライマーを、上記クローンのｃＤＮＡ配列から得た任意の配列を有するプライマーを合成し、アンカープライマーは市販のキットに付属のものをプライマーとして用いた。

【０２８１】

ＰＣＲ反応から得られた５’レース・クローンのいくつかを配列決定し、全体のコード領域をカバーしていると思われる配列を明らかにし、この遺伝子をそれぞれｎｅｌ関連蛋白１(nel-related protein type 1：ＮＲＰ１）遺伝子とｎｅｌ関連蛋白２(nel-related protein type 2：ＮＲＰ２）遺伝子とそれぞれ命名した。

【０２８２】

ＮＲＰ１遺伝子は、配列番号：３５で示される２４３０塩基のオープン・リーディング・フレームを含んでいて、これによってコードされる推定アミノ酸は、配列番号：３４で示されるように８１０アミノ酸残基を有しており、該ＮＲＰ１遺伝子の全ｃＤＮＡクローン

の核酸配列は、配列番号：３６で示されるとおり２９７７塩基からなっている。
【０２８３】
一方、ＮＲＰ２遺伝子は、配列番号：３８で示される２４４８塩基のオープン・リーディング・フレームを含んでいて、これによってコードされる推定アミノ酸は、配列番号：３７で示されるように８１６アミノ酸残基を有しており、該ＮＲＰ２遺伝子の全ｃＤＮＡクローンの核酸配列は、配列番号：３９で示されるとおり３１９８塩基からなっている。
【０２８４】
更にＲＴ－ＰＣＲによってコード領域を増幅し、得られた配列がｃＤＮＡのキメラであるという可能性を排除した。
【０２８５】
推定されたＮＲＰ１とＮＲＰ２遺伝子産物の両者は、膜挿入のためのシグナル・ペプチドとして機能することができる高い疎水性のＮ末端を有していた。ニワトリ胚のｎｅｌと比較すると、どちらも疎水性の膜貫通ドメインを保有していないように見えた。ＮＲＰ１、ＮＲＰ２及びｎｅｌ間の推定ペプチド配列を比較したところ、アミノ酸レベルでＮＲＰ２は８０％の相同性を有しており、ＮＲＰ１のそれの５０％よりｎｅｌにより近い関連があった。システィン－リッチ・ドメインとＥＧＦ様反復配列内のシスティン残基は、完全に保存されていた。
【０２８６】
ＮＲＰｓとニワトリ蛋白の最も目立つ差異は、ヒト相同物がｎｅｌの推定膜ドメインを欠いていることであったが、この欠けている領域においてもＮＲＰｓとｎｅｌの核酸配列は、大変類似していた。更に２つのＮＲＰは、６つのＥＧＦ様反復配列を保有していたが、一方、ｎｅｌは５つのみであった。
【０２８７】
マツハシらによって報告されたｎｅｌの他のユニークなモチーフ（Matsuhashi, S., et al., Dev. Dynamics., 203, 212-222(1995)）は、同じ位置でＮＲＰｓにおいても見つけられた。上記したように２つのＮＲＰの推定されたペプチドは、膜貫通タンパクではないことを示しており、ＮＲＰｓは分泌された蛋白或は翻訳後修飾によって膜に対して固定された蛋白であるかもしれない。
【０２８８】
本発明者らはＮＲＰｓがＥＧＦレセプターのような他の分子を刺激することによってリガンドとして機能するかもしれないと考えている。また本発明者らは、ｎｅｌの膜ドメイン領域をフレーム－シフトすることによってｎｅｌに余分のＥＧＦ様反復配列がコードされることを見つけ出した。
【０２８９】
ＮＲＰ２とｎｅｌを整列比較した時、このフレーム－シフトされた３アミノ酸配列は、ＮＲＰ２とｎｅｌ間の全領域で類似性を示した。このことは、ＮＲＰ２がｎｅｌのヒトの対応物であるかもしれないことを提言している。対照的に、ＮＲＰ１はヒトのｎｅｌ対応遺伝子ではなく類似性遺伝子であると考えられる。
【０２９０】
（２）ノーザンブロット分析
ノーザンブロット分析は、実施例１－（２）に準じて、正常ヒト組織におけるＮＲＰｓｍＲＮＡの発現を、両クローンしたｃＤＮＡｓの全体の配列をＰＣＲにより増幅し、該ＰＣＲ産物を精製し、［^{32}Ｐ］－ｄＣＴＰ（ランダムプライムドＤＮＡラベリングキット、ベーリンガーマンハイム社）により標識し、２つのプローブとしてＭＴＮブロットに用いて実施した。
【０２９１】
１６の成人組織と４つのヒト胎児組織において２つのＮＲＰｓの発現パターンを評価した。
【０２９２】
ノーザン分析の結果、ＮＲＰ１の３．５ｋｂの転写体が、胎児と成人の脳と腎臓において

弱く発現されていることが分かった。ＮＲＰ２の３．６ｋｂの転写体が、成人と胎児の脳においてのみ強く発現されたが、胎児の腎臓においてもわずかであるが発現されていた。

【０２９３】

このことからＮＲＰｓが成長調節のためのシグナル分子のような脳特異的役割を演ずることが提言される。加えて、これらの遺伝子は、腎臓において特別な機能を持っているかもしれない。

【０２９４】

（３）ＦＩＳＨによるＮＲＰ１遺伝子及びＮＲＰ２遺伝子の染色体のマッピング

実施例１－（３）に準じてＦＩＳＨによるＮＲＰ１遺伝子及びＮＲＰ２遺伝子の染色体のマッピングを行なった。

【０２９５】

その結果、ＮＲＰ１遺伝子は、染色体の座位が１１ｐ１５．１－ｐ１５．２に、ＮＲＰ２遺伝子は、染色体の座位が１２ｑ１３．１１－ｑ１３．１２の位置にそれぞれあることが判明した。

【０２９６】

本発明によれば、新規なヒトＮＲＰ１遺伝子及びＮＲＰ２遺伝子が提供され、該遺伝子を用いれば、該遺伝子の各種組織での発現の検出や、ヒトＮＲＰ１及びＮＲＰ２蛋白の遺伝子工学的製造が可能となり、脳の神経伝達経路関する研究に利用でき、更に脳の神経伝達経路に関連する各種疾患の診断等を行なうことができ、これらのの治療及び予防薬のスクリーニングや評価等をも行なうことができる。また、ＥＧＦドメインを持つこれらＮＲＰｓは脳内に成長因子的に働いている可能性が示唆され、各種脳内腫瘍の診断や治療、更に退行性神経疾患における神経の再生に奏効すると考えられる。

【０２９７】

【実施例１１】

ＧＳＰＴ１関連蛋白(ＧＳＰＴ１－ＴＫ）遺伝子

（１）ＧＳＰＴ１－ＴＫ遺伝子のクローニング及びＤＮＡシークエンシング

ヒトＧＳＰＴ１遺伝子は、細胞サイクルのＧ１期からＳ期転位に重要なＧＴＰ結合蛋白をコードする酵母のＧＳＴ１遺伝子のヒト相同遺伝子のひとつである。サッカロミセス・セレビシェ(Saccharomyces cerevisiae)の温度感受性ｇｓｔ１（Ｇ１－ｔｏ－Ｓ転位）変異体を補って完全にすることができる蛋白として、最初に特定された酵母のＧＳＴ１遺伝子は、酵母の染色体ライブラリーから単離された（Kikuchi,Y., Shimatake, H. and Kikuchi, A., EMBO J. 7, 1175-1182 (1988)）、該遺伝子は、ｃＡＭＰ依存性蛋白キナーゼの標的部位とＧＴＰａｓｅドメインを持つ蛋白をコードしていた。

【０２９８】

ヒトＧＳＰＴ１遺伝子は、酵母ＧＳＴ１遺伝子をプローブとして用いるハイブリダイゼーションによってＫＢ細胞のｃＤＮＡライブラリーから単離された（Hoshino, S., Miyazawa, H., Enomoto, T., Hanaoka, F., Kikucho, Y., Kikuchi, a. and Ui, m., EMBO J. 8, 3807-3814 (1989)）。該ＧＳＰＴ１遺伝子の推定蛋白もまた酵母ＧＳＴ１のそれと同様に、ＧＴＰ結合領域とＧＴＰａｓｅ活性センター部位を有しており、細胞の増殖に重要な役割を演じている。

【０２９９】

更に、染色体の再配列に対する中断点は急性非リンパ性白血病（ＡＮＬＬ）を持つ患者のＧＳＰＴ１遺伝子が染色体座位１６ｐ１３．３に観察されている(Ozawa, K., Murakami, Y., Eki, T., Yokoyama, K., Soeda, E., Hoshino, S., Ui, M. and Hanaoka, F., Somatic Cell and Molecular Genet., Vol.18, 189-194 (1992))。

【０３００】

実施例１－（１）と同様の方法でヒト胎児脳ｃＤＮＡライブラリーから任意に選択したｃＤＮＡクローンの配列解析とデータ・ベースの検索の結果、上記ＧＳＰＴ１に相同性の高い０．３ｋｂのｃＤＮＡ配列を有するクローンを見出し、ＧＥＮ－０７７Ａ０９と名付けた。ＧＥＮ－０７７Ａ０９クローンは、５’領域を欠いていると思われたので、全体のコ

ード領域を得るために実施例２－（２）と同様の方法で５’ＲＡＣＥ法を行なった。
【０３０１】
５’ＲＡＣＥR法のために、上記ｃＤＮＡクローンの分かっているｃＤＮＡ配列から得た表１１に示す塩基配列のＰ１プライマー及びＰ２プライマーを用い、また、アンカープライマーとしては市販のキットに付属のものをプライマーとして用いた。ＰＣＲ反応は、９４℃で４５秒、５８℃で４５秒、７２℃で２分を３５サイクル行ない、、最後に７２℃で７分間延長反応させた。
【０３０２】
【表１１】

プライマー	塩基配列
Ｐ１プライマー	5’-GATTTGTGCTCAATAATCACTATCTGAA-3’
Ｐ２プライマー	5’-GGTTACTAGGATCACAAAGTATGAATTCTGGAA-3’

ＰＣＲ反応から得られた５’ＲＡＣＥ・クローンのいくつかを配列決定し、この５’ＲＡＣＥ・クローンとＧＥＮ－０７７Ａ０９クローンの２つの重複しているｃＤＮＡクローンの塩基配列を決定することによって、全体のコード領域をカバーしていると考えられる配列を明らかにし、これをＧＳＰＴ１関連蛋白「ＧＳＰＴ１－ＴＫ遺伝子」と命名した。
【０３０３】
ＧＳＰＴ１－ＴＫ遺伝子は、配列番号：４１で示される１４９７塩基のオープン・リーディング・フレームを含んでいて、これによってコードされる推定アミノ酸は、配列番号：４０で示されるとおり４９９アミノ酸残基を有していた。
【０３０４】
ＧＳＰＴ１－ＴＫ遺伝子の全ｃＤＮＡクローンの核酸配列は、配列番号：４２で示されるとおり２０５７塩基からなり、計算された分子量は５５，７４０ダルトンであった。
【０３０５】
オープン・リーディング・フレームの最初のメチオニン（ＡＴＧ）は、イン－フレームに停止コドンが存在していなかったが、このＡＴＧは転写の開始に対するコザックの一致配列に類似する配列によって囲まれていたので、この配列番号１４４－１４６番目に位置するＡＴＧトリプレットを開始コドンと結論づけた。
【０３０６】
また、１つのポリアデニールシグナルＡＡＴＡＡＡがポリアデニレーション部位から１３塩基上流に認められた。
【０３０７】
ヒトＧＳＰＴ１－ＴＫは、Ｎ末端近くにグルタミン酸リッチ領域を含んでおり、ＧＳＰＴ１－ＴＫのこの領域に存在する２０のグルタミン酸のうち１８が、保存されていて、ヒトＧＳＰＴ１蛋白のそれと完全に整列している。ＧＳＰＴ１－ＴＫ蛋白は、ヒトＧＳＰＴ１蛋白のそれらの領域と同様にグアニン核酸結合と水解を担うＧＴＰ結合蛋白のいくつかの領域(Ｇ１、Ｇ２、Ｇ３、Ｇ４及びＧ５)が保存されていた。
【０３０８】
よって、ヒトＧＳＰＴ１－ＴＫは、細胞サイクルのＧ１期からＳ期転位に重要な役割を演じると考えられるヒトＧＳＰＴ１とＤＮＡ配列において８９．４％、推定されたアミノ酸配列において９２．４％の同一性を明らかにした。また、酵母のＧＳＴ１とはアミノ酸配列において５０．８％の同一性を示した。
【０３０９】
（２）ノーザンブロット分析
ノーザンブロット分析は、実施例１－（２）に準じて、正常ヒト組織におけるＧＳＰＴ１

－ＴＫｍＲＮＡの発現を、ＧＥＮ－０７７Ａ０９ｃＤＮＡクローンをＰＣＲにより増幅し、該ＰＣＲ産物を精製し、［^{32}Ｐ］－ｄＣＴＰ（ランダムプライムドＤＮＡラベリングキット、ベーリンガーマンハイム社）により標識し、プローブとしてＭＴＮブロットに用いて実施した。

【０３１０】

ノーザン分析の結果、様々の組織において２．７ｋｂの主要な転写体が検出された。ヒトＧＳＰＴ１－ＴＫ発現レベルは、脳と睾丸において最も発現が高いように思われた。

【０３１１】

（３）ＦＩＳＨによるＧＳＰＴ１－ＴＫ遺伝子の染色体のマッピング

実施例１－（３）に準じてＦＩＳＨによるＧＳＰＴ１－ＴＫ遺伝子の染色体のマッピングを行なった。

【０３１２】

その結果、ＧＳＰＴ１－ＴＫ遺伝子は、染色体の座位が１９ｐ１３．３の位置にあることが判明した。この染色体局在部位には急性リンパ性白血病（acute lymphocytic leukemia: ＡＬＬ）と急性ミエロイド白血病（acute myeloid leukemia: ＡＭＬ）からの相互の座位が高くしばしば観察されている。加えて、急性非リンパ性白血病（acute nonlymphocytic leukemia: ＡＮＬＬ）がヒトＧＳＰＴ１の領域を含んでいる再配列と関連されることが報告されている（Ozawa, K., Murakami, Y., Eki, T., Yokoyama, K., Soeda, E., Hoshino, S., Ui, M. and Hanaoka, F., Somatic Cell and Molecular Genet., Vol.18, 189-194 (1992)）。

【０３１３】

これらのことからＡＬＬとＡＭＬがこの遺伝子と関連している最も候補遺伝子であることが考えられる。

【０３１４】

本発明によれば、新規なヒトＧＳＰＴ１－ＴＫ遺伝子が提供され、該遺伝子を用いれば、該遺伝子の各種組織での発現の検出や、ヒトＧＳＰＴ１－ＴＫ蛋白の遺伝子工学的製造が可能となり、これらは、前述したように細胞増殖に関する研究に用いることが出来、この遺伝子の染色体の座位に関連する各種疾患、例えば急性骨髄性白血病等の診断等を行なうことも可能となる。即ち、これは、この遺伝子の転座により、ＧＳＰＴ１－ＴＫ遺伝子が破壊され、更にこの転座により発現される融合タンパクが之等の疾患の原因となっていると考えられるためである。

【０３１５】

更に、上記融合タンパクに対する抗体を作成し、細胞内局在を明らかにして、之等疾患に特異的な発現を調べることにより、その診断や治療も可能となると予想される。従って本発明遺伝子の利用によれば、これらの治療及び予防薬のスクリーニングや評価等をも行なうことができると考えられる。

【０３１６】

【配列表】

【０３１７】

配列番号：１
配列の長さ：１２２
配列の型：アミノ酸
トポロジー：直線状
配列の種類：蛋白
配列：

```
Met Glu Leu Gly Glu Asp Gly Ser Val Tyr Lys Ser Ile Leu Val Thr
 1               5                  10                  15
Ser Gln Asp Lys Ala Pro Ser Val Ile Ser Arg Val Leu Lys Lys Asn
            20                  25                  30
Asn Arg Asp Ser Ala Val Ala Ser Glu Tyr Glu Leu Val Gln Leu Leu
        35                  40                  45
Pro Gly Glu Arg Glu Leu Thr Ile Pro Ala Ser Ala Asn Val Phe Tyr
    50                  55                  60
Pro Met Asp Gly Ala Ser His Asp Phe Leu Leu Arg Gln Arg Arg Arg
65                  70                  75                  80
Ser Ser Thr Ala Thr Pro Gly Val Thr Ser Gly Pro Ser Ala Ser Gly
                85                  90                  95
Thr Pro Pro Ser Glu Gly Gly Gly Gly Ser Phe Pro Arg Ile Lys Ala
            100                 105                 110
Thr Gly Arg Lys Ile Ala Arg Ala Leu Phe
        115                 120
```

【０３１８】

配列番号：２
配列の長さ：３６６
配列の型：核酸
鎖の数：一本鎖
トポロジー：直線状
配列の種類：DNA (cDNA)
配列：

ATGGAGTTGG GGGAAGATGG CAGTGTCTAT AAGAGCATTT TGGTGACAAG CCAGGACAAG 60
GCTCCAAGTG TCATCAGTCG TGTCCTTAAG AAAAACAATC GTGACTCTGC AGTGGCTTCA 120
GAGTATGAGC TGGTACAGCT GCTACCAGGG GAGCGAGAGC TGACTATCCC AGCCTCGGCT 180
AATGTATTCT ACCCCATGGA TGGAGCTTCA CACGATTTCC TCCTGCGGCA GCGGCGAAGG 240
TCCTCTACTG CTACACCTGG CGTCACCAGT GGCCCGTCTG CCTCAGGAAC TCCTCCGAGT 300
GAGGGAGGAG GGGGCTCCTT TCCCAGGATC AAGGCCACAG GGAGGAAGAT TGCACGGGCA 360
CTGTTC 366

【０３１９】

配列番号：３
配列の長さ：８４２
配列の型：核酸
鎖の数：一本鎖
トポロジー：直線状
配列の種類：DNA (genomic)
起源：
　生物名：ヒト胎児脳ｃＤＮＡライブラリー
　株名：ＧＥＮ－５０１Ｄ０８
配列の特徴：
　特徴を表わす記号：CDS
　存在位置：２８．．３９３
　特徴を決定した方法：Ｅ
配列：

```
CCCACGAGCC GTATCATCCG AGTCCAG ATG GAG TTG GGG GAA GAT GGC AGT        51
                               Met Glu Leu Gly Glu Asp Gly Ser
                                 1               5
GTC TAT AAG AGC ATT TTG GTG ACA AGC CAG GAC AAG GCT CCA AGT GTC       99
Val Tyr Lys Ser Ile Leu Val Thr Ser Gln Asp Lys Ala Pro Ser Val
     10                  15                  20
ATC AGT CGT GTC CTT AAG AAA AAC AAT CGT GAC TCT GCA GTG GCT TCA      147
Ile Ser Arg Val Leu Lys Lys Asn Asn Arg Asp Ser Ala Val Ala Ser
 25                  30                  35                  40
GAG TAT GAG CTG GTA CAG CTG CTA CCA GGG GAG CGA GAG CTG ACT ATC      195
Glu Tyr Glu Leu Val Gln Leu Leu Pro Gly Glu Arg Glu Leu Thr Ile
                 45                  50                  55
CCA GCC TCG GCT AAT GTA TTC TAC CCC ATG GAT GGA GCT TCA CAC GAT      243
Pro Ala Ser Ala Asn Val Phe Tyr Pro Met Asp Gly Ala Ser His Asp
             60                  65                  70
```

```
TTC CTC CTG CGG CAG CGG CGA AGG TCC TCT ACT GCT ACA CCT GGC GTC      291
Phe Leu Leu Arg Gln Arg Arg Arg Ser Ser Thr Ala Thr Pro Gly Val
         75                  80                  85
ACC AGT GGC CCG TCT GCC TCA GGA ACT CCT CCG AGT GAG GGA GGA GGG      339
Thr Ser Gly Pro Ser Ala Ser Gly Thr Pro Pro Ser Glu Gly Gly Gly
     90                  95                 100
GGC TCC TTT CCC AGG ATC AAG GCC ACA GGG AGG AAG ATT GCA CGG GCA      387
Gly Ser Phe Pro Arg Ile Lys Ala Thr Gly Arg Lys Ile Ala Arg Ala
105                 110                 115                 120
CTG TTC TGAGGAGGAA GCCCCTTTTT TTACAGAAGT CATGGTGTTC ATACCAGATG      443
Leu Phe

TGGGTAGCCA TCCTGAATGG TGGCAATTAT ATCACATTGA GACAGAAATT CAGAAAGGGA    503
GCCAGCCACC CTGGGGCAGT GAAGTGCCAC TGGTTTACCA GACAGCTGAG AAATCCAGCC    563
CTGTCGGAAC TGGTGTCTTA TAACCAAGTT GGATACCTGT GTATAGCTTG CCACCTTCCA    623
TGAGTGCAGC ACACAGGTAG TGCTGGAAAA ACGCATCAGT TTCTGATTCT TGGCCATATC    683
CTAACATGCA AGGGCCAAGC AAAGGCTTCA AGGCTCTGAG CCCCAGGGCA GAGGGGAATG    743
GCAAAATGTA GGTCCTGGCA GGAGCTCTTC TTCCCACTCT GGGGGTTTCT ATCACTGTGA    803
CAACACTAAG ATAATAAACC AAAACACTAC CTGAATTCT                           842
```

【０３２０】

配列番号：４
配列の長さ：１９３
配列の型：アミノ酸
トポロジー：直線状
配列の種類：蛋白
配列：

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

【０３２１】

配列番号：５

配列の長さ：５７９

配列の型：核酸

鎖の数：一本鎖

トポロジー：直線状

配列の種類：DNA（cDNA）

配列：

ATGGAACTGG AGCTGTATGG AGTTGACGAC AAGTTCTACA GCAAGCTGGA TCAAGAGGAT 60

GCGCTCCTGG GCTCCTACCC TGTAGATGAC GGCTGCCGCA TCCACGTCAT TGACCACAGT 120

GGCGCCCGCC TTGGTGAGTA TGAGGACGTG TCCCGGGTGG AGAAGTACAC GATCTCACAA 180

GAAGCCTACG ACCAGAGGCA AGACACGGTC CGCTCTTTCC TGAAGCGCAG CAAGCTCGGC 240

CGGTACAACG AGGAGGAGCG GGCTCAGCAG GAGGCCGAGG CCGCCCAGCG CCTGGCCGAG 300

GAGAAGGCCC AGGCCAGCTC CATCCCCGTG GGCAGCCGCT GTGAGGTGCG GGCGGCGGGA 360

CAATCCCCTC GCCGGGGCAC CGTCATGTAT GTAGGTCTCA CAGATTTCAA GCCTGGCTAC 420

TGGATTGGTG TCCGCTATGA TGAGCCACTG GGGAAAAATG ATGGCAGTGT GAATGGGAAA 480

CGCTACTTCG AATGCCAGGC CAAGTATGGC GCCTTTGTCA AGCCAGCAGT CGTGACGGTG 540

GGGGACTTCC CGGAGGAGGA CTACGGGTTG GACGAGATA 579

【０３２２】

配列番号：６
配列の長さ：１０１５
配列の型：核酸
鎖の数：一本鎖
トポロジー：直線状
配列の種類：DNA (genomic)
起源：
　生物名：ヒト胎児脳ｃＤＮＡライブラリー
　株名：ＧＥＮ－０８０Ｇ０１
配列の特徴：
　特徴を表わす記号：CDS
　存在位置：２７４．．８５２
　特徴を決定した方法：Ｅ
配列：

```
TGATTGGTCA GGCACGGAGC AGGAGGCGGG CTGATAGCCC AGCAGCAGCA GCGGCGGCGG     60
CGGCTGCGGA GCGGGTGTGA GGCGGCTGGA CCGCGCTGCA GGCATCCGCG GGCGCGGCAA    120
GATGGAGGTG ACGGGGGTGT CGGCACCACG GTGACCGTTT TCATCAGCAG CTCCCTCAGC    180
ACCTTCCGCT CCGAGAAGCG ATACAGCCGC AGCCTCACCA TCGCTGAGTT CAAGTGTAAA    240
CTGGAGTTGC TGGTGGGCAG CCCTGCTTCC TGC ATG GAA CTG GAG CTG TAT GGA      294
                                     Met Glu Leu Glu Leu Tyr Gly
                                       1               5
GTT GAC GAC AAG TTC TAC AGC AAG CTG GAT CAA GAG GAT GCG CTC CTG       342
Val Asp Asp Lys Phe Tyr Ser Lys Leu Asp Gln Glu Asp Ala Leu Leu
         10                  15                  20
GGC TCC TAC CCT GTA GAT GAC GGC TGC CGC ATC CAC GTC ATT GAC CAC       390
Gly Ser Tyr Pro Val Asp Asp Gly Cys Arg Ile His Val Ile Asp His
     25                  30                  35
AGT GGC GCC CGC CTT GGT GAG TAT GAG GAC GTG TCC CGG GTG GAG AAG       438
Ser Gly Ala Arg Leu Gly Glu Tyr Glu Asp Val Ser Arg Val Glu Lys
```

```
 40                  45                  50                  55
TAC ACG ATC TCA CAA GAA GCC TAC GAC CAG AGG CAA GAC ACG GTC CGC      486
Tyr Thr Ile Ser Gln Glu Ala Tyr Asp Gln Arg Gln Asp Thr Val Arg
                 60                  65                  70
TCT TTC CTG AAG CGC AGC AAG CTC GGC CGG TAC AAC GAG GAG GAG CGG      534
Ser Phe Leu Lys Arg Ser Lys Leu Gly Arg Tyr Asn Glu Glu Glu Arg
             75                  80                  85
GCT CAG CAG GAG GCC GAG GCC GCC CAG CGC CTG GCC GAG GAG AAG GCC      582
Ala Gln Gln Glu Ala Glu Ala Ala Gln Arg Leu Ala Glu Glu Lys Ala
         90                  95                 100
CAG GCC AGC TCC ATC CCC GTG GGC AGC CGC TGT GAG GTG CGG GCG GCG      630
Gln Ala Ser Ser Ile Pro Val Gly Ser Arg Cys Glu Val Arg Ala Ala
    105                 110                 115
GGA CAA TCC CCT CGC CGG GGC ACC GTC ATG TAT GTA GGT CTC ACA GAT      678
Gly Gln Ser Pro Arg Arg Gly Thr Val Met Tyr Val Gly Leu Thr Asp
120                 125                 130                 135
TTC AAG CCT GGC TAC TGG ATT GGT GTC CGC TAT GAT GAG CCA CTG GGG      726
Phe Lys Pro Gly Tyr Trp Ile Gly Val Arg Tyr Asp Glu Pro Leu Gly
                140                 145                 150
AAA AAT GAT GGC AGT GTG AAT GGG AAA CGC TAC TTC GAA TGC CAG GCC      774
Lys Asn Asp Gly Ser Val Asn Gly Lys Arg Tyr Phe Glu Cys Gln Ala
            155                 160                 165
AAG TAT GGC GCC TTT GTC AAG CCA GCA GTC GTG ACG GTG GGG GAC TTC      822
Lys Tyr Gly Ala Phe Val Lys Pro Ala Val Val Thr Val Gly Asp Phe
        170                 175                 180
CCG GAG GAG GAC TAC GGG TTG GAC GAG ATA TGACACCTAA GGAATTCCCC        872
Pro Glu Glu Asp Tyr Gly Leu Asp Glu Ile
    185                 190
TGCTTCAGCT CCTAGCTCAG CCACTGACTG CCCCTCCTGT GTGTGCCCAT GGCCCTTTTC    932

TCCTGACCCC ATTTTAATTT TATTCATTTT TTCCTTTGCC ATTGATTTTT GAGACTCATG    992
CATTAAATTC ACTAGAAACC CAG                                          1015
```

【０３２３】

配列番号：７

配列の長さ：１２８

配列の型：アミノ酸

トポロジー：直線状

配列の種類：蛋白

配列：

```
Met Thr Glu Ala Asp Val Asn Pro Lys Ala Tyr Pro Leu Ala Asp Ala
 1               5                  10                  15
His Leu Thr Lys Lys Leu Leu Asp Leu Val Gln Gln Ser Cys Asn Tyr
            20                  25                  30
Lys Gln Leu Arg Lys Gly Ala Asn Glu Ala Thr Lys Thr Leu Asn Arg
        35                  40                  45
Gly Ile Ser Glu Phe Ile Val Met Ala Ala Asp Ala Glu Pro Leu Glu
    50                  55                  60
Ile Ile Leu His Leu Pro Leu Leu Cys Glu Asp Lys Asn Val Pro Tyr
65                  70                  75                  80
Val Phe Val Arg Ser Lys Gln Ala Leu Gly Arg Ala Cys Gly Val Ser
                85                  90                  95
Arg Pro Val Ile Ala Cys Ser Val Thr Ile Lys Glu Gly Ser Gln Leu
            100                 105                 110
Lys Gln Gln Ile Gln Ser Ile Gln Gln Ser Ile Glu Arg Leu Leu Val
        115                 120                 125
```

【０３２４】

配列番号：８
配列の長さ：３８４
配列の型：核酸
鎖の数：一本鎖
トポロジー：直線状
配列の種類：DNA (genomic)
配列：

ATGACTGAGG CTGATGTGAA TCCAAAGGCC TATCCCCTTG CCGATGCCCA CCTCACCAAG 60
AAGCTACTGG ACCTCGTTCA GCAGTCATGT AACTATAAGC AGCTTCGGAA AGGAGCCAAT 120
GAGGCCACCA AAACCCTCAA CAGGGGCATC TCTGAGTTCA TCGTGATGGC TGCAGACGCC 180
GAGCCACTGG AGATCATTCT GCACCTGCCG CTGCTGTGTG AAGACAAGAA TGTGCCCTAC 240
GTGTTTGTGC GCTCCAAGCA GGCCCTGGGG AGAGCCTGTG GGGTCTCCAG GCCTGTCATC 300
GCCTGTTCTG TCACCATCAA AGAAGGCTCG CAGCTGAAAC AGCAGATCCA ATCCATTCAG 360
CAGTCCATTG AAAGGCTCTT AGTC 384

【０３２５】

配列番号：９
配列の長さ：１４９３
配列の型：核酸
鎖の数：一本鎖
トポロジー：直線状
配列の種類：DNA (genomic)
起源：
　生物名：ヒト胎児脳ｃＤＮＡライブラリー
　株名：ＧＥＮ－０２５Ｆ０７
配列の特徴：
　特徴を表わす記号：CDS
　存在位置：９５．．４７８
　特徴を決定した方法：Ｅ
配列：

```
ATCCGTGTCC TTGCGGTGCT GGGCAGCAGA CCGTCCAAAC CGACACGCGT GGTATCCTCG     60
CGGTGTCCGG CAAGAGACTA CCAAGACAGA CGCT ATG ACT GAG GCT GAT GTG         112
                                      Met Thr Glu Ala Asp Val
                                        1               5
AAT CCA AAG GCC TAT CCC CTT GCC GAT GCC CAC CTC ACC AAG AAG CTA       160
Asn Pro Lys Ala Tyr Pro Leu Ala Asp Ala His Leu Thr Lys Lys Leu
            10                  15                  20
CTG GAC CTC GTT CAG CAG TCA TGT AAC TAT AAG CAG CTT CGG AAA GGA       208
Leu Asp Leu Val Gln Gln Ser Cys Asn Tyr Lys Gln Leu Arg Lys Gly
        25                  30                  35
GCC AAT GAG GCC ACC AAA ACC CTC AAC AGG GGC ATC TCT GAG TTC ATC       256
Ala Asn Glu Ala Thr Lys Thr Leu Asn Arg Gly Ile Ser Glu Phe Ile
    40                  45                  50
GTG ATG GCT GCA GAC GCC GAG CCA CTG GAG ATC ATT CTG CAC CTG CCG       304
Val Met Ala Ala Asp Ala Glu Pro Leu Glu Ile Ile Leu His Leu Pro
```

```
 55                  60                  65                  70
CTG CTG TGT GAA GAC AAG AAT GTG CCC TAC GTG TTT GTG CGC TCC AAG      352
Leu Leu Cys Glu Asp Lys Asn Val Pro Tyr Val Phe Val Arg Ser Lys
                 75                  80                  85
CAG GCC CTG GGG AGA GCC TGT GGG GTC TCC AGG CCT GTC ATC GCC TGT      400
Gln Ala Leu Gly Arg Ala Cys Gly Val Ser Arg Pro Val Ile Ala Cys
            90                  95                 100
TCT GTC ACC ATC AAA GAA GGC TCG CAG CTG AAA CAG CAG ATC CAA TCC      448
Ser Val Thr Ile Lys Glu Gly Ser Gln Leu Lys Gln Gln Ile Gln Ser
        105                 110                 115
ATT CAG CAG TCC ATT GAA AGG CTC TTA GTC TAAACCTGTG GCCTCTGCCA        498
Ile Gln Gln Ser Ile Glu Arg Leu Leu Val
    120                 125
CGTGCTCCCT GCCAGCTTCC CCCCTGAGGT TGTGTATCAT ATTATCTGTG TTAGCATGTA    558
GTATTTTCAG CTACTCTCTA TTGTTATAAA ATGTAGTACT AAATCTGGTT TCTGGATTTT    618
TGTGTTGTTT TTGTTCTGTT TTACAGGGTT GCTATCCCCC TTCCTTTCCT CCCTCCCTCT    678
GCCATCCTTC ATCCTTTTAT CCTCCCTTTT TGGAACAAGT GTTCAGAGCA GACAGAAGCA    738
GGGTGGTGGC ACCGTTGAAA GGCAGAAAGA GCCAGGAGAA AGCTGATGGA GCCAGGACAG    798
AGATCTGGTT CCAGCTTTCA GCCACTAGCT TCCTGTTGTG TGCGGGGTGT GGTGGAATTA    858
AACAGCATTC ATTGTGTGTC CCTGTGCCTG GCACACAGAA TCATTCATAC GTGTTCAAGT    918
GATCAAGGGG TTTCATTTGC TCTTGGGGGA TTAGGTATCA TTTGGGGAGG AAGCATGTGT    978
TCTGTGAGGT TGTTCGGCTA TGTCCAAGTG TCGTTTACTA ATGTACCCCT GCTGTTTGCT   1038
TTTGGTAATG TGATGTTGAT GTTCTCCCCC TACCCACAAC CATGCCCTTG AGGGTAGCAG   1098
GGCAGCAGCA TACCAAAGAG ATGTGCTGCA GGACTCCGGA GGCAGCCTGG GTGGGTGAGC   1158
CATGGGGCAG TTGACCTGGG TCTTGAAAGA GTCGGGAGTG ACAAGCTCAG AGAGCATGAA   1218
CTGATGCTGG CATGAAGGAT TCCAGGAAGA TCATGGAGAC CTGGCTGGTA GCTGTAACAG   1278
AGATGGTGGA GTCCAAGGAA ACAGCCTGTC TCTGGTGAAT GGGACTTTCT TTGGTGGACA   1338
CTTGGCACCA GCTCTGAGAG CCCTTCCCCT GTGTCCTGCC ACCATGTGGG TCAGATGTAC   1398
TCTCTGTCAC ATGAGGAGAG TGCTAGTTCA TGTGTTCTCC ATTCTTGTGA GCATCCTAAT   1458

AAATCTGTTC CATTTTGAAA AAAAAAAAAA AAAAA                              1493
```

【０３２６】

配列番号：１０
配列の長さ：７１１
配列の型：アミノ酸
トポロジー：直線状
配列の種類：蛋白
配列：

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

180 185 190

Lys Gln Pro Gln Lys Cys Cys Leu Phe Thr Glu Ser Leu Lys Leu Asn

195 200 205

Leu Glu Val Asn Gly Gln Asn Glu Ser Asn Asp Thr Glu Gln Leu Asp

210 215 220

Asp Val Val Gly Ser Gly Gln Leu Phe Ser His Ser Ser Ser Asp Ala

225 230 235 240

Cys Ser Lys Asn Ile His Thr Gly Glu Thr Phe Cys Lys Gly Asn Gln

245 250 255

Cys Arg Lys Val Cys Gly His Lys Gln Ser Leu Lys Gln His Gln Ile

260 265 270

His Thr Gln Lys Lys Pro Asp Gly Cys Ser Glu Cys Gly Gly Ser Phe

275 280 285

Thr Gln Lys Ser His Leu Phe Ala Gln Gln Arg Ile His Ser Val Gly

290 295 300

Asn Leu His Glu Cys Gly Lys Cys Gly Lys Ala Phe Met Pro Gln Leu

305 310 315 320

Lys Leu Ser Val Tyr Leu Thr Asp His Thr Gly Asp Ile Pro Cys Ile

325 330 335

Cys Lys Glu Cys Gly Lys Val Phe Ile Gln Arg Ser Glu Leu Leu Thr

340 345 350

His Gln Lys Thr His Thr Arg Lys Lys Pro Tyr Lys Cys His Asp Cys

355 360 365

Gly Lys Ala Phe Phe Gln Met Leu Ser Leu Phe Arg His Gln Arg Thr

370 375 380

His Ser Arg Glu Lys Leu Tyr Glu Cys Ser Glu Cys Gly Lys Gly Phe

385 390 395 400

Ser Gln Asn Ser Thr Leu Ile Ile His Gln Lys Ile His Thr Gly Glu

405 410 415

Arg Gln Tyr Ala Cys Ser Glu Cys Gly Lys Ala Phe Thr Gln Lys Ser
420 425 430
Thr Leu Ser Leu His Gln Arg Ile His Ser Gly Gln Lys Ser Tyr Val
435 440 445
Cys Ile Glu Cys Gly Gln Ala Phe Ile Gln Lys Ala His Leu Ile Val
450 455 460
His Gln Arg Ser His Thr Gly Glu Lys Pro Tyr Gln Cys His Asn Cys
465 470 475 480
Gly Lys Ser Phe Ile Ser Lys Ser Gln Leu Asp Ile His His Arg Ile
485 490 495
His Thr Gly Glu Lys Pro Tyr Glu Cys Ser Asp Cys Gly Lys Thr Phe
500 505 510
Thr Gln Lys Ser His Leu Asn Ile His Gln Lys Ile His Thr Gly Glu
515 520 525
Arg His His Val Cys Ser Glu Cys Gly Lys Ala Phe Asn Gln Lys Ser
530 535 540
Ile Leu Ser Met His Gln Arg Ile His Thr Gly Glu Lys Pro Tyr Lys
545 550 555 560
Cys Ser Glu Cys Gly Lys Ala Phe Thr Ser Lys Ser Gln Phe Lys Glu
565 570 575
His Gln Arg Ile His Thr Gly Glu Lys Pro Tyr Val Cys Thr Glu Cys
580 585 590
Gly Lys Ala Phe Asn Gly Arg Ser Asn Phe His Lys His Gln Ile Thr
595 600 605
His Thr Arg Glu Arg Pro Phe Val Cys Tyr Lys Cys Gly Lys Ala Phe
610 615 620
Val Gln Lys Ser Glu Leu Ile Thr His Gln Arg Thr His Met Gly Glu
625 630 635 640
Lys Pro Tyr Glu Cys Leu Asp Cys Gly Lys Ser Phe Ser Lys Lys Pro

645 650 655

Gln Leu Lys Val His Gln Arg Ile His Thr Gly Glu Arg Pro Tyr Val

660 665 670

Cys Ser Glu Cys Gly Lys Ala Phe Asn Asn Arg Ser Asn Phe Asn Lys

675 680 685

His Gln Thr Thr His Thr Arg Asp Lys Ser Tyr Lys Cys Ser Tyr Ser

690 695 700

Val Lys Gly Phe Thr Lys Gln

705 710

【０３２７】

配列番号：１１

配列の長さ：２１３３

配列の型：核酸

鎖の数：一本鎖

トポロジー：直線状

配列の種類：DNA(genomic)

配列：

ATGCCTGCTG ATGTGAATTT ATCCCAGAAG CCTCAGGTCC TGGGTCCAGA GAAGCAGGAT 60

GGATCTTGCG AGGCATCAGT GTCATTTGAG GACGTGACCG TGGACTTCAG CAGGGAGGAG 120

TGGCAGCAAC TGGACCCTGC CCAGAGATGC CTGTACCGGG ATGTGATGCT GGAGCTCTAT 180

AGCCATCTCT TCGCAGTGGG GTATCACATT CCCAACCCAG AGGTCATCTT CAGAATGCTA 240

AAAGAAAAGG AGCCGCGTGT GGAGGAGGCT GAAGTCTCAC ATCAGAGGTG TCAAGAAAGG 300

GAGTTTGGGC TTGAAATCCC ACAAAAGGAG ATTTCTAAGA AAGCTTCATT TCAAAAGGAT 360

ATGGTAGGTG AGTTCACAAG AGATGGTTCA TGGTGTTCCA TTTTAGAAGA ACTGAGGCTG 420

GATGCTGACC GCACAAAGAA AGATGAGCAA AATCAAATTC AACCCATGAG TCACAGTGCT 480

TTCTTCAACA AGAAAACATT GAACACAGAA AGCAATTGTG AATATAAGGA CCCTGGGAAA 540

ATGATTCGCA CGAGGCCCCA CCTTGCTTCT TCACAGAAAC AACCTCAGAA ATGTTGCTTA 600

TTTACAGAAA GTTTGAAGCT GAACCTAGAA GTGAACGGTC AGAATGAAAG CAATGACACA 660

GAACAGCTTG ATGACGTTGT TGGGTCTGGT CAGCTATTCA GCCATAGCTC TTCTGATGCC 720

TGCAGCAAGA ATATTCATAC AGGAGAGACA TTTTGCAAAG GTAACCAGTG TAGAAAAGTC 780

TGTGGCCATA AACAGTCACT CAAGCAACAT CAAATTCATA CTCAGAAGAA ACCAGATGGA 840

TGTTCTGAAT GTGGGGGGAG CTTCACCCAG AAGTCACACC TCTTTGCCCA ACAGAGAATT 900

CATAGTGTAG GAAACCTCCA TGAATGTGGC AAATGTGGAA AAGCCTTCAT GCCACAACTA 960

AAACTCAGTG TATATCTGAC AGATCATACA GGTGATATAC CCTGTATATG CAAGGAATGT 1020

GGGAAGGTCT TTATTCAGAG ATCAGAATTG CTTACGCACC AGAAAACACA CACTAGAAAG 1080

AAGCCCTATA AATGCCATGA CTGTGGAAAA GCCTTTTTCC AGATGTTATC TCTCTTCAGA 1140

CATCAGAGAA CTCACAGTAG AGAAAAACTC TATGAATGCA GTGAATGTGG CAAAGGCTTC 1200

TCCCAAAACT CAACCCTCAT TATACATCAG AAAATTCATA CTGGTGAGAG ACAGTATGCA 1260

TGCAGTGAAT GTGGGAAAGC CTTTACCCAG AAGTCAACAC TCAGCTTGCA CCAGAGAATC 1320

CACTCAGGGC AGAAGTCCTA TGTGTGTATC GAATGCGGGC AGGCCTTCAT CCAGAAGGCA 1380
CACCTGATTG TCCATCAAAG AAGCCACACA GGAGAAAAAC CTTATCAGTG CCACAACTGT 1440
GGGAAATCCT TCATTTCCAA GTCACAGCTT GATATACATC ATCGAATTCA TACAGGGGAG 1500
AAACCTTATG AATGCAGTGA CTGTGGAAAA ACCTTCACCC AAAAGTCACA CCTGAATATA 1560
CACCAGAAAA TTCATACTGG AGAAAGACAC CATGTATGCA GTGAATGCGG GAAAGCCTTC 1620
AACCAGAAGT CAATACTCAG CATGCATCAG AGAATTCACA CCGGAGAGAA GCCTTACAAA 1680
TGCAGTGAAT GTGGGAAAGC CTTCACTTCT AAGTCTCAAT TCAAAGAGCA TCAGCGAATT 1740
CACACGGGTG AGAAACCCTA TGTGTGCACT GAATGTGGGA AGGCCTTCAA CGGCAGGTCA 1800
AATTTCCATA AACATCAAAT AACTCACACT AGAGAGAGGC CTTTTGTCTG TTACAAATGT 1860
GGGAAGGCTT TTGTCCAGAA ATCAGAGTTG ATTACCCATC AAAGAACTCA CATGGGAGAG 1920
AAACCCTATG AATGCCTTGA CTGTGGGAAA TCGTTCAGTA AGAAACCACA ACTCAAGGTG 1980
CATCAGCGAA TTCACACGGG AGAAAGACCT TATGTGTGTT CTGAATGTGG AAAGGCCTTC 2040
AACAACAGGT CAAACTTCAA TAAACACCAA ACAACTCATA CCAGAGACAA ATCTTACAAA 2100
TGCAGTTATT CTGTGAAAGG CTTTACCAAG CAA 2133

【０３２８】

配列番号：１２
配列の長さ：３７５４
配列の型：核酸
鎖の数：一本鎖
トポロジー：直線状
配列の種類：DNA(genomic)
起源：
　生物名：ヒト胎児脳ｃＤＮＡライブラリー
　株名：ＧＥＮ－０７６Ｃ０９
配列の特徴：
　特徴を表わす記号：CDS
　存在位置：３４６．．２４７８
　特徴を決定した方法：Ｅ
配列：

```
GCTAAGCCTA TGTCGCTTAC TGGACGCTGA AGTGATTGGG AATATTAGCA GTGGGGGTTC     60
TGTAGGGTCA GGAAGGGGCG GCTGGCTTTG GGGGAGTGAT GAGGGGCTTG TTGGGGGTGG    120
GGGTGCGTGA TAAAGGGATT TCTCGGCTGA AGACGAGGCT GTGAGGCTTC TGCAGAACCC    180
CCAGGTCAGG CCACATCATT GAGGCTGCAG GATCTCTCTT CATAGCCCAG TACGACTCTC    240
CGCCGTGTCC CTGGTTGGAA AATCCAAACA CCTATCCAGC TTCTGGCTCC TGGGAAAAGT    300
GGAGTTGTCA GCAAGAGAGA CCGAGAGTAG AAGCCCAGAG TGGAG ATG CCT GCT        354
                                                  Met Pro Ala
                                                    1
GAT GTG AAT TTA TCC CAG AAG CCT CAG GTC CTG GGT CCA GAG AAG CAG       402
Asp Val Asn Leu Ser Gln Lys Pro Gln Val Leu Gly Pro Glu Lys Gln
      5                  10                  15
GAT GGA TCT TGC GAG GCA TCA GTG TCA TTT GAG GAC GTG ACC GTG GAC       450
Asp Gly Ser Cys Glu Ala Ser Val Ser Phe Glu Asp Val Thr Val Asp
 20                  25                  30                  35
TTC AGC AGG GAG GAG TGG CAG CAA CTG GAC CCT GCC CAG AGA TGC CTG       498
```

```
Phe Ser Arg Glu Glu Trp Gln Gln Leu Asp Pro Ala Gln Arg Cys Leu
                  40                  45                  50
TAC CGG GAT GTG ATG CTG GAG CTC TAT AGC CAT CTC TTC GCA GTG GGG      546
Tyr Arg Asp Val Met Leu Glu Leu Tyr Ser His Leu Phe Ala Val Gly
              55                  60                  65
TAT CAC ATT CCC AAC CCA GAG GTC ATC TTC AGA ATG CTA AAA GAA AAG      594
Tyr His Ile Pro Asn Pro Glu Val Ile Phe Arg Met Leu Lys Glu Lys
          70                  75                  80
GAG CCG CGT GTG GAG GAG GCT GAA GTC TCA CAT CAG AGG TGT CAA GAA      642
Glu Pro Arg Val Glu Glu Ala Glu Val Ser His Gln Arg Cys Gln Glu
      85                  90                  95
AGG GAG TTT GGG CTT GAA ATC CCA CAA AAG GAG ATT TCT AAG AAA GCT      690
Arg Glu Phe Gly Leu Glu Ile Pro Gln Lys Glu Ile Ser Lys Lys Ala
100                 105                 110                 115
TCA TTT CAA AAG GAT ATG GTA GGT GAG TTC ACA AGA GAT GGT TCA TGG      738
Ser Phe Gln Lys Asp Met Val Gly Glu Phe Thr Arg Asp Gly Ser Trp
                  120                 125                 130
TGT TCC ATT TTA GAA GAA CTG AGG CTG GAT GCT GAC CGC ACA AAG AAA      786
Cys Ser Ile Leu Glu Glu Leu Arg Leu Asp Ala Asp Arg Thr Lys Lys
              135                 140                 145
GAT GAG CAA AAT CAA ATT CAA CCC ATG AGT CAC AGT GCT TTC TTC AAC      834
Asp Glu Gln Asn Gln Ile Gln Pro Met Ser His Ser Ala Phe Phe Asn
          150                 155                 160
AAG AAA ACA TTG AAC ACA GAA AGC AAT TGT GAA TAT AAG GAC CCT GGG      882
Lys Lys Thr Leu Asn Thr Glu Ser Asn Cys Glu Tyr Lys Asp Pro Gly
      165                 170                 175
AAA ATG ATT CGC ACG AGG CCC CAC CTT GCT TCT TCA CAG AAA CAA CCT      930
Lys Met Ile Arg Thr Arg Pro His Leu Ala Ser Ser Gln Lys Gln Pro
180                 185                 190                 195
```

```
CAG AAA TGT TGC TTA TTT ACA GAA AGT TTG AAG CTG AAC CTA GAA GTG      978
Gln Lys Cys Cys Leu Phe Thr Glu Ser Leu Lys Leu Asn Leu Glu Val
                200                 205                 210
AAC GGT CAG AAT GAA AGC AAT GAC ACA GAA CAG CTT GAT GAC GTT GTT     1026
Asn Gly Gln Asn Glu Ser Asn Asp Thr Glu Gln Leu Asp Asp Val Val
            215                 220                 225
GGG TCT GGT CAG CTA TTC AGC CAT AGC TCT TCT GAT GCC TGC AGC AAG     1074
Gly Ser Gly Gln Leu Phe Ser His Ser Ser Ser Asp Ala Cys Ser Lys
        230                 235                 240
AAT ATT CAT ACA GGA GAG ACA TTT TGC AAA GGT AAC CAG TGT AGA AAA     1122
Asn Ile His Thr Gly Glu Thr Phe Cys Lys Gly Asn Gln Cys Arg Lys
    245                 250                 255
GTC TGT GGC CAT AAA CAG TCA CTC AAG CAA CAT CAA ATT CAT ACT CAG     1170
Val Cys Gly His Lys Gln Ser Leu Lys Gln His Gln Ile His Thr Gln
260                 265                 270                 275
AAG AAA CCA GAT GGA TGT TCT GAA TGT GGG GGG AGC TTC ACC CAG AAG     1218
Lys Lys Pro Asp Gly Cys Ser Glu Cys Gly Gly Ser Phe Thr Gln Lys
                280                 285                 290
TCA CAC CTC TTT GCC CAA CAG AGA ATT CAT AGT GTA GGA AAC CTC CAT     1266
Ser His Leu Phe Ala Gln Gln Arg Ile His Ser Val Gly Asn Leu His
            295                 300                 305
GAA TGT GGC AAA TGT GGA AAA GCC TTC ATG CCA CAA CTA AAA CTC AGT     1314
Glu Cys Gly Lys Cys Gly Lys Ala Phe Met Pro Gln Leu Lys Leu Ser
        310                 315                 320
GTA TAT CTG ACA GAT CAT ACA GGT GAT ATA CCC TGT ATA TGC AAG GAA     1362
Val Tyr Leu Thr Asp His Thr Gly Asp Ile Pro Cys Ile Cys Lys Glu
    325                 330                 335
TGT GGG AAG GTC TTT ATT CAG AGA TCA GAA TTG CTT ACG CAC CAG AAA     1410
Cys Gly Lys Val Phe Ile Gln Arg Ser Glu Leu Leu Thr His Gln Lys
```

```
340                 345                 350                 355
ACA CAC ACT AGA AAG AAG CCC TAT AAA TGC CAT GAC TGT GGA AAA GCC     1458
Thr His Thr Arg Lys Lys Pro Tyr Lys Cys His Asp Cys Gly Lys Ala
                360                 365                 370
TTT TTC CAG ATG TTA TCT CTC TTC AGA CAT CAG AGA ACT CAC AGT AGA     1506
Phe Phe Gln Met Leu Ser Leu Phe Arg His Gln Arg Thr His Ser Arg
            375                 380                 385
GAA AAA CTC TAT GAA TGC AGT GAA TGT GGC AAA GGC TTC TCC CAA AAC     1554
Glu Lys Leu Tyr Glu Cys Ser Glu Cys Gly Lys Gly Phe Ser Gln Asn
        390                 395                 400
TCA ACC CTC ATT ATA CAT CAG AAA ATT CAT ACT GGT GAG AGA CAG TAT     1602
Ser Thr Leu Ile Ile His Gln Lys Ile His Thr Gly Glu Arg Gln Tyr
    405                 410                 415
GCA TGC AGT GAA TGT GGG AAA GCC TTT ACC CAG AAG TCA ACA CTC AGC     1650
Ala Cys Ser Glu Cys Gly Lys Ala Phe Thr Gln Lys Ser Thr Leu Ser
420                 425                 430                 435
TTG CAC CAG AGA ATC CAC TCA GGG CAG AAG TCC TAT GTG TGT ATC GAA     1698
Leu His Gln Arg Ile His Ser Gly Gln Lys Ser Tyr Val Cys Ile Glu
                440                 445                 450
TGC GGG CAG GCC TTC ATC CAG AAG GCA CAC CTG ATT GTC CAT CAA AGA     1746
Cys Gly Gln Ala Phe Ile Gln Lys Ala His Leu Ile Val His Gln Arg
            455                 460                 465
AGC CAC ACA GGA GAA AAA CCT TAT CAG TGC CAC AAC TGT GGG AAA TCC     1794
Ser His Thr Gly Glu Lys Pro Tyr Gln Cys His Asn Cys Gly Lys Ser
        470                 475                 480
TTC ATT TCC AAG TCA CAG CTT GAT ATA CAT CAT CGA ATT CAT ACA GGG     1842
Phe Ile Ser Lys Ser Gln Leu Asp Ile His His Arg Ile His Thr Gly
    485                 490                 495
GAG AAA CCT TAT GAA TGC AGT GAC TGT GGA AAA ACC TTC ACC CAA AAG     1890
```

```
Glu Lys Pro Tyr Glu Cys Ser Asp Cys Gly Lys Thr Phe Thr Gln Lys
500                 505                 510                 515
TCA CAC CTG AAT ATA CAC CAG AAA ATT CAT ACT GGA GAA AGA CAC CAT     1938
Ser His Leu Asn Ile His Gln Lys Ile His Thr Gly Glu Arg His His
                520                 525                 530
GTA TGC AGT GAA TGC GGG AAA GCC TTC AAC CAG AAG TCA ATA CTC AGC     1986
Val Cys Ser Glu Cys Gly Lys Ala Phe Asn Gln Lys Ser Ile Leu Ser
            535                 540                 545
ATG CAT CAG AGA ATT CAC ACC GGA GAG AAG CCT TAC AAA TGC AGT GAA     2034
Met His Gln Arg Ile His Thr Gly Glu Lys Pro Tyr Lys Cys Ser Glu
        550                 555                 560
TGT GGG AAA GCC TTC ACT TCT AAG TCT CAA TTC AAA GAG CAT CAG CGA     2082
Cys Gly Lys Ala Phe Thr Ser Lys Ser Gln Phe Lys Glu His Gln Arg
    565                 570                 575
ATT CAC ACG GGT GAG AAA CCC TAT GTG TGC ACT GAA TGT GGG AAG GCC     2130
Ile His Thr Gly Glu Lys Pro Tyr Val Cys Thr Glu Cys Gly Lys Ala
580                 585                 590                 595
TTC AAC GGC AGG TCA AAT TTC CAT AAA CAT CAA ATA ACT CAC ACT AGA     2178
Phe Asn Gly Arg Ser Asn Phe His Lys His Gln Ile Thr His Thr Arg
                600                 605                 610
GAG AGG CCT TTT GTC TGT TAC AAA TGT GGG AAG GCT TTT GTC CAG AAA     2226
Glu Arg Pro Phe Val Cys Tyr Lys Cys Gly Lys Ala Phe Val Gln Lys
            615                 620                 625
TCA GAG TTG ATT ACC CAT CAA AGA ACT CAC ATG GGA GAG AAA CCC TAT     2274
Ser Glu Leu Ile Thr His Gln Arg Thr His Met Gly Glu Lys Pro Tyr
        630                 635                 640
GAA TGC CTT GAC TGT GGG AAA TCG TTC AGT AAG AAA CCA CAA CTC AAG     2322
Glu Cys Leu Asp Cys Gly Lys Ser Phe Ser Lys Lys Pro Gln Leu Lys
    645                 650                 655
```

```
GTG CAT CAG CGA ATT CAC ACG GGA GAA AGA CCT TAT GTG TGT TCT GAA     2370
Val His Gln Arg Ile His Thr Gly Glu Arg Pro Tyr Val Cys Ser Glu
660                 665                 670                 675
TGT GGA AAG GCC TTC AAC AAC AGG TCA AAC TTC AAT AAA CAC CAA ACA     2418
Cys Gly Lys Ala Phe Asn Asn Arg Ser Asn Phe Asn Lys His Gln Thr
                680                 685                 690
ACT CAT ACC AGA GAC AAA TCT TAC AAA TGC AGT TAT TCT GTG AAA GGC     2466
Thr His Thr Arg Asp Lys Ser Tyr Lys Cys Ser Tyr Ser Val Lys Gly
            695                 700                 705
TTT ACC AAG CAA TGAATTCCTA GTGCATCAGC ATATTCATAA ATGAAATATA         2518
Phe Thr Lys Gln
        710
CTCCGAGTTT CTTGAAGAAG AGAACATCTT CTCAGAATCA GGTCTAATTA TATGTTATTG   2578
AATTCATGCT TCAGAAAAAC TCTAGGGATG CACTGCATGT GTGAACACAT GATAAAAAAG   2638
TCATGCTTTA TTTTAGTGAG GGCAATTACA GAGAAAAGAG TAAGCAGAAA TGTCCTTCTG   2698
AGTACTGGCC TCATTAAGGA TTATAAATTT TCTCCCCGGG AAGAAACCCT GACTAACGCA   2758
TTGAGAAAAG CCTTTCTGTA AAGAATGGTA CAAGACAGGT TGTTACTCGA TTATTTATAG   2818
TAAAATATGT GGGAAATTAT ATCAATGATA ACCCTGTTTA TTGTGGGATA TCAATATTTT   2878
TAAAGTGCCA ACACAGTCAT GATAGGACAA TATTTTATGT GTGTGTGTGC GCCTTATGTA   2938
TATAAGCATA TATATAATAT ATAAGCATAT TATTATATAC AGGTTGAGTA TCCCTTCTCC   2998
AAAATGCCTG GGATCAGAAG CATTTTGGAT TTCAGATACT TACAGATTTT GGAATATTTG   3058
CATTATATTT ATTGGTTGAG CATCCCTAAT CTGAAAATCC AAGATTAAAT GCTCCAATTA   3118
GCATTTCCTT TGAGCGTCAT GTTAGAGTTC AAAAAGTTTC AGATTTTGGG TTTTCAGATT   3178
AGGAATACCC AACCTGTATG TACGTATATT TCTGTATCTA TGTATGTATA TATATGCATA   3238
TGCAGACATA TGTATATGGT CTGGTCAGCA TATGTGTATG TATGCGTATG TATGTATGTA   3298
TGTATGCCCT CAGTGCAGTG GGGTTTGCTG CAGAATTCAC TGCATAGCAG GAGATGTAAG   3358
CAGATGAGTT ATTTTTTAAG AGAATCTAAT CTAATTGTTT TTATAAAAAT TATTCCCTAT   3418
TGAATATTTA TATAATGAGG TTGTATCAAC AATGATTAAC TCCTTTATTA TACATACACA   3478
TGAATGTGCA TTTTTGGTAA ATGCATAAAT GAGATTCTAT AATGTTTACT GATCTTTATA   3538

TTACAGATTT TCTCTTCTTT TAGGATTAGC TCAGCTTGCC CCCCCTTTCC ATCTCCACCA   3598
TCTATAGTGA GCCTCTCCAT AATTAGTGCC AACCATTAGT CTCGTTCATA TTTTTACACC   3658
AGGAGTCAAC AAACTGTGCC ATTGGCCAAA TATGGCCTCC CAACTGTTTT TTTAAAATAA   3718
AGTTTTATTG GAACACAAAA AAAAAAAAAA AAAAAA                             3754
```

【０３２９】

配列番号：１３
配列の長さ：３８９
配列の型：アミノ酸
トポロジー：直線状
配列の種類：蛋白
配列：

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

180 185 190
Cys Asn Phe Leu Lys Val Val Ser Ser Ser Ile Val Asp Lys Tyr Ile
195 200 205
Gly Glu Ser Ala Arg Leu Ile Arg Glu Met Phe Asn Tyr Ala Arg Asp
210 215 220
His Gln Pro Cys Ile Ile Phe Met Asp Glu Ile Asp Ala Ile Gly Gly
225 230 235 240
Arg Arg Phe Ser Glu Gly Thr Ser Ala Asp Arg Glu Ile Gln Arg Thr
245 250 255
Leu Met Glu Leu Leu Asn Gln Met Asp Gly Phe Asp Thr Leu His Arg
260 265 270
Val Lys Met Thr Met Ala Thr Asn Arg Pro Asp Thr Leu Asp Pro Ala
275 280 285
Leu Leu Arg Pro Gly Arg Leu Asp Arg Lys Ile His Ile Asp Leu Pro
290 295 300
Asn Glu Gln Ala Arg Leu Asp Ile Leu Lys Ile His Ala Gly Pro Ile
305 310 315 320
Thr Lys His Gly Glu Ile Asp Tyr Glu Ala Ile Val Lys Leu Ser Asp
325 330 335
Gly Phe Asn Gly Ala Asp Leu Arg Asn Val Cys Thr Glu Ala Gly Met
340 345 350
Phe Ala Ile Arg Ala Asp His Asp Phe Val Val Gln Glu Asp Phe Met
355 360 365
Lys Ala Val Arg Lys Val Ala Asp Ser Lys Lys Leu Glu Ser Lys Leu
370 375 380
Asp Tyr Lys Pro Val
385

【０３３０】

配列番号：１４

配列の長さ：１１６７

配列の型：核酸

鎖の数：一本鎖

トポロジー：直線状

配列の種類：ＤＮＡ(genomic)

配列：

ATGGCGGACC CTAGAGATAA GGCGCTTCAG GACTACCGCA AGAAGTTGCT TGAACACAAG 60

GAGATCGACG GCCGTCTTAA GGAGTTAAGG GAACAATTAA AAGAACTTAC CAAGCAGTAT 120

GAAAAGTCTG AAAATGATCT GAAGGCCCTA CAGAGTGTTG GGCAGATCGT GGGTGAAGTG 180

CTTAAACAGT TAACTGAAGA AAAATTCATT GTTAAAGCTA CCAATGGACC AAGATATGTT 240

GTGGGTTGTC GTCGACAGCT TGACAAAAGT AAGCTGAAGC CAGGAACAAG AGTTGCTTTG 300

GATATGACTA CACTAACTAT CATGAGATAT TTGCCGAGAG AGGTGGATCC ACTGGTTTAT 360

AACATGTCTC ATGAGGACCC TGGGAATGTT TCTTATTCTG AGATTGGAGG GCTATCAGAA 420

CAGATCCGGG AATTAAGAGA GGTGATAGAA TTACCTCTTA CAAACCCAGA GTTATTTCAG 480

CGTGTAGGAA TAATACCTCC AAAAGGCTGT TTGTTATATG GACCACCAGG TACGGGAAAA 540

ACACTCTTGG CACGAGCCGT TGCTAGCCAG CTGGACTGCA ATTTCTTAAA GGTTGTATCT 600

AGTTCTATTG TAGACAAGTA CATTGGTGAA AGTGCTCGTT TGATCAGAGA AATGTTTAAT 660

TATGCTAGAG ATCATCAACC ATGCATCATT TTTATGGATG AAATAGATGC TATTGGTGGT 720

CGTCGGTTTT CTGAGGGTAC TTCAGCTGAC AGAGAGATTC AGAGAACGTT AATGGAGTTA 780

CTGAATCAAA TGGATGGATT TGATACTCTG CATAGAGTTA AAATGACCAT GGCTACAAAC 840

AGACCAGATA CACTGGATCC TGCTTTGCTG CGTCCAGGAA GATTAGATAG AAAAATACAT 900

ATTGATTTGC CAAATGAACA AGCAAGATTA GACATACTGA AAATCCATGC AGGTCCCATT 960

ACAAAGCATG GTGAAATAGA TTATGAAGCA ATTGTGAAGC TTTCGGATGG CTTTAATGGA 1020

GCAGATCTGA GAAATGTTTG TACTGAAGCA GGTATGTTCG CAATTCGTGC TGATCATGAT 1080

TTTGTAGTAC AGGAAGACTT CATGAAAGCA GTCAGAAAAG TGGCTGATTC TAAGAAGCTG 1140

GAGTCTAAAT TGGACTACAA ACCTGTG 1167

【０３３１】

配列番号：１５
配列の長さ：１５６６
配列の型：核酸
鎖の数：一本鎖
トポロジー：直線状
配列の種類：ｃＤＮＡ
起源：
　生物名：ヒト胎児脳ｃＤＮＡライブラリー
　株名：ＧＥＮ－３３１Ｇ０７
配列の特徴：
　特徴を表わす記号：ＣＤＳ
　存在位置：１７．．１１８３
　特徴を決定した方法：Ｅ
配列：

```
GAGACGGCTT CTCATC ATG GCG GAC CCT AGA GAT AAG GCG CTT CAG GAC        49
                  Met Ala Asp Pro Arg Asp Lys Ala Leu Gln Asp
                    1               5                  10
TAC CGC AAG AAG TTG CTT GAA CAC AAG GAG ATC GAC GGC CGT CTT AAG      97
Tyr Arg Lys Lys Leu Leu Glu His Lys Glu Ile Asp Gly Arg Leu Lys
            15                  20                  25
GAG TTA AGG GAA CAA TTA AAA GAA CTT ACC AAG CAG TAT GAA AAG TCT      145
Glu Leu Arg Glu Gln Leu Lys Glu Leu Thr Lys Gln Tyr Glu Lys Ser
        30                  35                  40
GAA AAT GAT CTG AAG GCC CTA CAG AGT GTT GGG CAG ATC GTG GGT GAA      193
Glu Asn Asp Leu Lys Ala Leu Gln Ser Val Gly Gln Ile Val Gly Glu
    45                  50                  55
GTG CTT AAA CAG TTA ACT GAA GAA AAA TTC ATT GTT AAA GCT ACC AAT      241
Val Leu Lys Gln Leu Thr Glu Glu Lys Phe Ile Val Lys Ala Thr Asn
60                  65                  70                  75
```

GGA CCA AGA TAT GTT GTG GGT TGT CGT CGA CAG CTT GAC AAA AGT AAG 289
Gly Pro Arg Tyr Val Val Gly Cys Arg Arg Gln Leu Asp Lys Ser Lys
80 85 90
CTG AAG CCA GGA ACA AGA GTT GCT TTG GAT ATG ACT ACA CTA ACT ATC 337
Leu Lys Pro Gly Thr Arg Val Ala Leu Asp Met Thr Thr Leu Thr Ile
95 100 105
ATG AGA TAT TTG CCG AGA GAG GTG GAT CCA CTG GTT TAT AAC ATG TCT 385
Met Arg Tyr Leu Pro Arg Glu Val Asp Pro Leu Val Tyr Asn Met Ser
110 115 120
CAT GAG GAC CCT GGG AAT GTT TCT TAT TCT GAG ATT GGA GGG CTA TCA 433
His Glu Asp Pro Gly Asn Val Ser Tyr Ser Glu Ile Gly Gly Leu Ser
125 130 135
GAA CAG ATC CGG GAA TTA AGA GAG GTG ATA GAA TTA CCT CTT ACA AAC 481
Glu Gln Ile Arg Glu Leu Arg Glu Val Ile Glu Leu Pro Leu Thr Asn
140 145 150 155
CCA GAG TTA TTT CAG CGT GTA GGA ATA ATA CCT CCA AAA GGC TGT TTG 529
Pro Glu Leu Phe Gln Arg Val Gly Ile Ile Pro Pro Lys Gly Cys Leu
160 165 170
TTA TAT GGA CCA CCA GGT ACG GGA AAA ACA CTC TTG GCA CGA GCC GTT 577
Leu Tyr Gly Pro Pro Gly Thr Gly Lys Thr Leu Leu Ala Arg Ala Val
175 180 185
GCT AGC CAG CTG GAC TGC AAT TTC TTA AAG GTT GTA TCT AGT TCT ATT 625
Ala Ser Gln Leu Asp Cys Asn Phe Leu Lys Val Val Ser Ser Ser Ile
190 195 200
GTA GAC AAG TAC ATT GGT GAA AGT GCT CGT TTG ATC AGA GAA ATG TTT 673
Val Asp Lys Tyr Ile Gly Glu Ser Ala Arg Leu Ile Arg Glu Met Phe
205 210 215
AAT TAT GCT AGA GAT CAT CAA CCA TGC ATC ATT TTT ATG GAT GAA ATA 721
Asn Tyr Ala Arg Asp His Gln Pro Cys Ile Ile Phe Met Asp Glu Ile

```
220                 225                 230                 235
GAT GCT ATT GGT GGT CGT CGG TTT TCT GAG GGT ACT TCA GCT GAC AGA      769
Asp Ala Ile Gly Gly Arg Arg Phe Ser Glu Gly Thr Ser Ala Asp Arg
                240                 245                 250
GAG ATT CAG AGA ACG TTA ATG GAG TTA CTG AAT CAA ATG GAT GGA TTT      817
Glu Ile Gln Arg Thr Leu Met Glu Leu Leu Asn Gln Met Asp Gly Phe
            255                 260                 265
GAT ACT CTG CAT AGA GTT AAA ATG ACC ATG GCT ACA AAC AGA CCA GAT      865
Asp Thr Leu His Arg Val Lys Met Thr Met Ala Thr Asn Arg Pro Asp
        270                 275                 280
ACA CTG GAT CCT GCT TTG CTG CGT CCA GGA AGA TTA GAT AGA AAA ATA      913
Thr Leu Asp Pro Ala Leu Leu Arg Pro Gly Arg Leu Asp Arg Lys Ile
    285                 290                 295
CAT ATT GAT TTG CCA AAT GAA CAA GCA AGA TTA GAC ATA CTG AAA ATC      961
His Ile Asp Leu Pro Asn Glu Gln Ala Arg Leu Asp Ile Leu Lys Ile
300                 305                 310                 315
CAT GCA GGT CCC ATT ACA AAG CAT GGT GAA ATA GAT TAT GAA GCA ATT     1009
His Ala Gly Pro Ile Thr Lys His Gly Glu Ile Asp Tyr Glu Ala Ile
                320                 325                 330
GTG AAG CTT TCG GAT GGC TTT AAT GGA GCA GAT CTG AGA AAT GTT TGT     1057
Val Lys Leu Ser Asp Gly Phe Asn Gly Ala Asp Leu Arg Asn Val Cys
            335                 340                 345
ACT GAA GCA GGT ATG TTC GCA ATT CGT GCT GAT CAT GAT TTT GTA GTA     1105
Thr Glu Ala Gly Met Phe Ala Ile Arg Ala Asp His Asp Phe Val Val
        350                 355                 360
CAG GAA GAC TTC ATG AAA GCA GTC AGA AAA GTG GCT GAT TCT AAG AAG     1153
Gln Glu Asp Phe Met Lys Ala Val Arg Lys Val Ala Asp Ser Lys Lys
    365                 370                 375
CTG GAG TCT AAA TTG GAC TAC AAA CCT GTG TAATTTACTG TAAGATTTTT       1203
```

Leu Glu Ser Lys Leu Asp Tyr Lys Pro Val
380 385

GATGGCTGCA TGACAGATGT TGGCTTATTG TAAAAATAAA GTTAAAGAAA ATAATGTATG 1263
TATTGGCAAT GATGTCATTA AAAGTATATG AATAAAAATA TGAGTAACAT CATAAAAATT 1323
AGTAATTCAA CTTTTAAGAT ACAGAAGAAA TTTGTATGTT TGTTAAAGTT GCATTTATTG 1383
CAGCAAGTTA CAAAGGGAAA GTGTTGAAGC TTTTCATATT TGCTGCGTGA GCATTTTGTA 1443
AAATATTGAA AGTGGTTTGA GATAGTGGTA TAAGAAAGCA TTTCTTATGA CTTATTTTGT 1503
ATCATTTGTT TTCCTCATCT AAAAAGTTGA ATAAAATCTG TTTGATTCAG TTCTCCTAAA 1563
AAA 1566

【０３３２】

配列番号：１６
配列の長さ：２２３
配列の型：アミノ酸
トポロジー：直線状
配列の種類：蛋白
配列：

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

180 185 190

Ile Arg Arg Gly Glu Lys His Gln Leu Arg Leu Val Pro Thr Arg Trp

195 200 205

Ala Gly Lys Gly Leu Leu Gly Cys Asn Ile Ile Pro Leu Gln Arg

210 215 220

【０３３３】

配列番号：１７

配列の長さ：６６９

配列の型：核酸

鎖の数：一本鎖

トポロジー：直線状

配列の種類：ＤＮＡ(genomic)

配列：

ATGTCCGACG AGGAAGCGAG GCAGAGCGGA GGCTCCTCGC AGGCCGGCGT CGTGACTGTC 60

AGCGACGTCC AGGAGCTGAT GCGGCGCAAG GAGGAGATAG AAGCGCAGAT CAAGGCCAAC 120

TATGACGTGC TGGAAAGCCA AAAAGGCATT GGGATGAACG AGCCGCTGGT GGACTGTGAG 180

GGCTACCCCC GGTCAGACGT GGACCTGTAC CAAGTCCGCA CCGCCAGGCA CAACATCATA 240

TGCCTGCAGA ATGATCACAA GGCAGTGATG AAGCAGGTGG AGGAGGCCCT GCACCAGCTG 300

CACGCTCGCG ACAAGGAGAA GCAGGCCCGG GACATGGCTG AGGCCCACAA AGAGGCCATG 360

AGCCGCAAAC TGGGTCAGAG TGAGAGCCAG GGCCCTCCAC GGGCCTTCGC CAAAGTGAAC 420

AGCATCAGCC CCGGCTCCCC AGCCAGCATC GCGGGTCTGC AAGTGGATGA TGAGATTGTG 480

GAGTTCGGCT CTGTGAACAC CCAGAACTTC CAGTCACTGC ATAACATTGG CAGTGTGGTG 540

CAGCACAGTG AGGGGAAGCC CCTGAATGTG ACAGTGATCC GCAGGGGGGA AAAACACCAG 600

CTTAGACTTG TTCCAACACG CTGGGCAGGA AAAGGACTGC TGGGCTGCAA CATTATTCCT 660

CTGCAAAGA 669

【０３３４】

配列番号：１８
配列の長さ：１１２８
配列の型：核酸
鎖の数：一本鎖
トポロジー：直線状
配列の種類：ｃＤＮＡ
起源：
　生物名：ヒト胎児脳ｃＤＮＡライブラリー
　株名：ＧＥＮ－１６３Ｄ０９
配列の特徴：
　特徴を表わす記号：ＣＤＳ
　存在位置：１２５．．７９３
　特徴を決定した方法：Ｅ
配列：

```
ACTGTTCTCG CGTTCGCGGA CGGCTGTGGT GTTTTGGCGC ATGGGCGGAG CGTAGTTACG     60
GTCGACTGGG GCGTCGTCCC TAGCCCGGGA GCCGGGTCTC TGGAGTCGCG GCCCGGGGTT    120
CACG ATG TCC GAC GAG GAA GCG AGG CAG AGC GGA GGC TCC TCG CAG GCC      169
     Met Ser Asp Glu Glu Ala Arg Gln Ser Gly Gly Ser Ser Gln Ala
       1               5                  10                  15
GGC GTC GTG ACT GTC AGC GAC GTC CAG GAG CTG ATG CGG CGC AAG GAG      217
Gly Val Val Thr Val Ser Asp Val Gln Glu Leu Met Arg Arg Lys Glu
                  20                  25                  30
GAG ATA GAA GCG CAG ATC AAG GCC AAC TAT GAC GTG CTG GAA AGC CAA      265
Glu Ile Glu Ala Gln Ile Lys Ala Asn Tyr Asp Val Leu Glu Ser Gln
             35                  40                  45
AAA GGC ATT GGG ATG AAC GAG CCG CTG GTG GAC TGT GAG GGC TAC CCC      313
Lys Gly Ile Gly Met Asn Glu Pro Leu Val Asp Cys Glu Gly Tyr Pro
         50                  55                  60
CGG TCA GAC GTG GAC CTG TAC CAA GTC CGC ACC GCC AGG CAC AAC ATC      361
```

```
Arg Ser Asp Val Asp Leu Tyr Gln Val Arg Thr Ala Arg His Asn Ile
 65                  70                  75
ATA TGC CTG CAG AAT GAT CAC AAG GCA GTG ATG AAG CAG GTG GAG GAG      409
Ile Cys Leu Gln Asn Asp His Lys Ala Val Met Lys Gln Val Glu Glu
 80                  85                  90                  95
GCC CTG CAC CAG CTG CAC GCT CGC GAC AAG GAG AAG CAG GCC CGG GAC      457
Ala Leu His Gln Leu His Ala Arg Asp Lys Glu Lys Gln Ala Arg Asp
                100                 105                 110
ATG GCT GAG GCC CAC AAA GAG GCC ATG AGC CGC AAA CTG GGT CAG AGT      505
Met Ala Glu Ala His Lys Glu Ala Met Ser Arg Lys Leu Gly Gln Ser
            115                 120                 125
GAG AGC CAG GGC CCT CCA CGG GCC TTC GCC AAA GTG AAC AGC ATC AGC      553
Glu Ser Gln Gly Pro Pro Arg Ala Phe Ala Lys Val Asn Ser Ile Ser
        130                 135                 140
CCC GGC TCC CCA GCC AGC ATC GCG GGT CTG CAA GTG GAT GAT GAG ATT      601
Pro Gly Ser Pro Ala Ser Ile Ala Gly Leu Gln Val Asp Asp Glu Ile
    145                 150                 155
GTG GAG TTC GGC TCT GTG AAC ACC CAG AAC TTC CAG TCA CTG CAT AAC      649
Val Glu Phe Gly Ser Val Asn Thr Gln Asn Phe Gln Ser Leu His Asn
160                 165                 170                 175
ATT GGC AGT GTG GTG CAG CAC AGT GAG GGG AAG CCC CTG AAT GTG ACA      697
Ile Gly Ser Val Val Gln His Ser Glu Gly Lys Pro Leu Asn Val Thr
                180                 185                 190
GTG ATC CGC AGG GGG GAA AAA CAC CAG CTT AGA CTT GTT CCA ACA CGC      745
Val Ile Arg Arg Gly Glu Lys His Gln Leu Arg Leu Val Pro Thr Arg
            195                 200                 205
TGG GCA GGA AAA GGA CTG CTG GGC TGC AAC ATT ATT CCT CTG CAA AGA      793
Trp Ala Gly Lys Gly Leu Leu Gly Cys Asn Ile Ile Pro Leu Gln Arg
        210                 215                 220
```

TGATTGTCCC TGGGGAACAG TAACAGGAAA GCATCTTCCC TTGCCCTGGA CTTGGGTCTA 853
GGGATTTCCA ACTTGTCTTC TCTCCCTGAA GCATAAGGAT CTGGAAGAGG CTTGTAACCT 913
GAACTTCTGT GTGGTGGCAG TACTGTGGCC CACCAGTGTA ATCTCCCTGG ATTAAGGCAT 973
TCTTAAAAAC TTAGGCTTGG CCTCTTTCAC AAATTAGGCC ACGGCCCTAA ATAGGAATTC 1033
CCTGGATTGT GGGCAAGTGG GCGGAAGTTA TTCTGGCAGG TACTGGTGTG ATTATTATTA 1093
TTATTTTTAA TAAAGAGTTT TACAGTGCTG ATATG 1128

【０３３５】

配列番号：１９
配列の長さ：５０６
配列の型：アミノ酸
トポロジー：直線状
配列の種類：蛋白
配列：

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

180 185 190

Glu Glu Asn Pro Lys Glu Asn Pro Glu Val Lys Ala Glu Glu Lys Glu

195 200 205

Val Pro Lys Glu Ile Pro Glu Val Lys Asp Glu Glu Lys Glu Val Ala

210 215 220

Lys Glu Ile Pro Glu Val Lys Ala Glu Glu Lys Ala Asp Ser Lys Asp

225 230 235 240

Cys Met Glu Ala Thr Pro Glu Val Lys Glu Asp Pro Lys Glu Val Pro

245 250 255

Gln Val Lys Ala Asp Asp Lys Glu Gln Pro Lys Ala Thr Glu Ala Lys

260 265 270

Ala Arg Ala Ala Val Arg Glu Thr His Lys Arg Val Pro Glu Glu Arg

275 280 285

Leu Arg Asp Ser Val Asp Leu Lys Arg Ala Arg Lys Gly Lys Pro Lys

290 295 300

Arg Glu Asp Pro Lys Gly Ile Pro Asp Tyr Trp Leu Ile Val Leu Lys

305 310 315 320

Asn Val Asp Lys Leu Gly Pro Met Ile Gln Lys Tyr Asp Glu Pro Ile

325 330 335

Leu Lys Phe Leu Ser Asp Val Ser Leu Lys Phe Ser Lys Pro Gly Gln

340 345 350

Pro Val Ser Tyr Thr Phe Glu Phe His Phe Leu Pro Asn Pro Tyr Phe

355 360 365

Arg Asn Glu Val Leu Val Lys Thr Tyr Ile Ile Lys Ala Lys Pro Asp

370 375 380

His Asn Asp Pro Phe Phe Ser Trp Gly Trp Glu Ile Glu Asp Cys Lys

385 390 395 400

Gly Cys Lys Ile Asp Arg Arg Arg Gly Lys Asp Val Thr Val Thr Thr

405 410 415

Thr Gln Ser Arg Thr Thr Ala Thr Gly Glu Ile Glu Ile Gln Pro Arg
420 425 430
Val Val Pro Asn Ala Ser Phe Phe Asn Phe Phe Ser Pro Pro Glu Ile
435 440 445
Pro Met Ile Gly Lys Leu Glu Pro Arg Glu Asp Ala Ile Leu Asp Glu
450 455 460
Asp Phe Glu Ile Gly Gln Ile Leu His Asp Asn Val Ile Leu Lys Ser
465 470 475 480
Ile Tyr Tyr Tyr Thr Gly Glu Val Asn Gly Thr Tyr Tyr Gln Phe Gly
485 490 495
Lys His Tyr Gly Asn Lys Lys Tyr Arg Lys
500 505

【０３３６】

配列番号：２０
配列の長さ：１５１８
配列の型：核酸
鎖の数：一本鎖
トポロジー：直線状
配列の種類：ＤＮＡ(genomic)
配列：

ATGGCAGAAG CAGATTTTAA AATGGTCTCG GAACCTGTCG CCCATGGGGT TGCCGAAGAG 60
GAGATGGCTA GCTCGACTAG TGATTCTGGG GAAGAATCTG ACAGCAGTAG CTCTAGCAGC 120
AGCACTAGTG ACAGCAGCAG CAGCAGCAGC ACTAGTGGCA GCAGCAGCGG CAGCGGCAGC 180
AGCAGCAGCA GCAGCGGCAG CACTAGCAGC CGCAGCCGCT TGTATAGAAA GAAGAGGGTA 240
CCTGAGCCTT CCAGAAGGGC GCGGCGGGCC CCGTTGGGAA CAAATTTCGT GGATAGGCTG 300
CCTCAGGCAG TTAGAAATCG TGTGCAAGCG CTTAGAAACA TTCAAGATGA ATGTGACAAG 360
GTAGATACCC TGTTCTTAAA AGCAATTCAT GATCTTGAAA GAAAATATGC TGAACTCAAC 420
AAGCCTCTGT ATGATAGGCG GTTTCAAATC ATCAATGCAG AATACGAGCC TACAGAAGAA 480
GAATGTGAAT GGAATTCAGA GGATGAGGAG TTCAGCAGTG ATGAGGAGGT GCAGGATAAC 540
ACCCCTAGTG AAATGCCTCC CTTAGAGGGT GAGGAAGAAG AAAACCCTAA AGAAAACCCA 600
GAGGTGAAAG CTGAAGAGAA GGAAGTTCCT AAAGAAATTC CTGAGGTGAA GGATGAAGAA 660
AAGGAAGTTG CTAAAGAAAT TCCTGAGGTA AAGGCTGAAG AAAAAGCAGA TTCTAAAGAC 720
TGTATGGAGG CAACCCCTGA AGTAAAAGAA GATCCTAAAG AAGTCCCCCA GGTAAAGGCA 780
GATGATAAAG AACAGCCTAA AGCAACAGAG GCTAAGGCAA GGGCTGCAGT AAGAGAGACT 840
CATAAAAGAG TTCCTGAGGA AAGGCTTCGG GACAGTGTAG ATCTTAAAAG AGCTAGGAAG 900
GGAAAGCCTA AAAGAGAAGA CCCTAAAGGC ATTCCTGACT ATTGGCTGAT TGTTTTAAAG 960
AATGTTGACA AGCTCGGGCC TATGATTCAG AAGTATGATG AGCCCATTCT GAAGTTCTTG 1020
TCGGATGTTA GCCTGAAGTT CTCAAAACCT GGCCAGCCTG TAAGTTACAC CTTTGAATTT 1080
CATTTTCTAC CCAACCCATA CTTCAGAAAT GAGGTGCTGG TGAAGACATA TATAATAAAG 1140
GCAAAACCAG ATCACAATGA TCCCTTCTTT TCTTGGGGAT GGGAAATTGA AGATTGCAAA 1200
GGCTGCAAGA TAGACCGGAG AAGAGGAAAA GATGTTACTG TGACAACTAC CCAGAGTCGC 1260
ACAACTGCTA CTGGAGAAAT TGAAATCCAG CCAAGAGTGG TTCCTAATGC ATCATTCTTC 1320

AACTTCTTTA GTCCTCCTGA GATTCCTATG ATTGGGAAGC TGGAACCACG AGAAGATGCT 1380
ATCCTGGATG AGGACTTTGA AATTGGGCAG ATTTTACATG ATAATGTCAT CCTGAAATCA 1440
ATCTATTACT ATACTGGAGA AGTCAATGGT ACCTACTATC AATTTGGCAA ACATTATGGA 1500
AACAAGAAAT ACAGAAAA 1518

【０３３７】

配列番号：２１
配列の長さ：２６３６
配列の型：核酸
鎖の数：一本鎖
トポロジー：直線状
配列の種類：ＤＮＡ(genomic)
起源：
　生物名：ヒト胎児脳ｃＤＮＡライブラリー
　株名：ＧＥＮ－０７８Ｄ０５
配列の特徴：
　特徴を表わす記号：ＣＤＳ
　存在位置：２６６．．１７８３
　特徴を決定した方法：Ｅ
配列：

```
GATTCGGCTG CGGTACATCT CGGCACTCTA GCTGCAGCCG GGAGAGGCCT TGCCGCCACC     60
GCTGTCGCCC AAGCCTCCAC TGCCGCTGCC ACCTCAGCGC CGGCCTCTGC ATCCCCAGCT    120
CCAGCTCCGC TCTGCGCCGC TGCTGCCATC GCCGCTGCCA CCTCCGCAGC CCGGGCCTCC    180
GCCGCCGCCA CCCAAGCATC CGTGAGTCAT TTTCTGCCCA TCTCTGGTCG CGCGGTCTCC    240
CTGGTAGAGT TTGTAGGCTT GCAAG ATG GCA GAA GCA GAT TTT AAA ATG GTC      292
                           Met Ala Glu Ala Asp Phe Lys Met Val
                             1               5
TCG GAA CCT GTC GCC CAT GGG GTT GCC GAA GAG GAG ATG GCT AGC TCG      340
Ser Glu Pro Val Ala His Gly Val Ala Glu Glu Glu Met Ala Ser Ser
 10                  15                  20                  25
ACT AGT GAT TCT GGG GAA GAA TCT GAC AGC AGT AGC TCT AGC AGC AGC      388
Thr Ser Asp Ser Gly Glu Glu Ser Asp Ser Ser Ser Ser Ser Ser Ser
                 30                  35                  40
ACT AGT GAC AGC AGC AGC AGC AGC AGC ACT AGT GGC AGC AGC AGC GGC      436
Thr Ser Asp Ser Ser Ser Ser Ser Ser Thr Ser Gly Ser Ser Ser Gly
```

```
                45                  50                  55
AGC GGC AGC AGC AGC AGC AGC AGC GGC AGC ACT AGC AGC CGC AGC CGC      484
Ser Gly Ser Ser Ser Ser Ser Ser Gly Ser Thr Ser Ser Arg Ser Arg
        60                  65                  70
TTG TAT AGA AAG AAG AGG GTA CCT GAG CCT TCC AGA AGG GCG CGG CGG      532
Leu Tyr Arg Lys Lys Arg Val Pro Glu Pro Ser Arg Arg Ala Arg Arg
    75                  80                  85
GCC CCG TTG GGA ACA AAT TTC GTG GAT AGG CTG CCT CAG GCA GTT AGA      580
Ala Pro Leu Gly Thr Asn Phe Val Asp Arg Leu Pro Gln Ala Val Arg
90                  95                  100                 105
AAT CGT GTG CAA GCG CTT AGA AAC ATT CAA GAT GAA TGT GAC AAG GTA      628
Asn Arg Val Gln Ala Leu Arg Asn Ile Gln Asp Glu Cys Asp Lys Val
                110                 115                 120
GAT ACC CTG TTC TTA AAA GCA ATT CAT GAT CTT GAA AGA AAA TAT GCT      676
Asp Thr Leu Phe Leu Lys Ala Ile His Asp Leu Glu Arg Lys Tyr Ala
            125                 130                 135
GAA CTC AAC AAG CCT CTG TAT GAT AGG CGG TTT CAA ATC ATC AAT GCA      724
Glu Leu Asn Lys Pro Leu Tyr Asp Arg Arg Phe Gln Ile Ile Asn Ala
        140                 145                 150
GAA TAC GAG CCT ACA GAA GAA GAA TGT GAA TGG AAT TCA GAG GAT GAG      772
Glu Tyr Glu Pro Thr Glu Glu Glu Cys Glu Trp Asn Ser Glu Asp Glu
    155                 160                 165
GAG TTC AGC AGT GAT GAG GAG GTG CAG GAT AAC ACC CCT AGT GAA ATG      820
Glu Phe Ser Ser Asp Glu Glu Val Gln Asp Asn Thr Pro Ser Glu Met
170                 175                 180                 185
CCT CCC TTA GAG GGT GAG GAA GAA GAA AAC CCT AAA GAA AAC CCA GAG      868
Pro Pro Leu Glu Gly Glu Glu Glu Glu Asn Pro Lys Glu Asn Pro Glu
                190                 195                 200
GTG AAA GCT GAA GAG AAG GAA GTT CCT AAA GAA ATT CCT GAG GTG AAG      916
```

```
Val Lys Ala Glu Glu Lys Glu Val Pro Lys Glu Ile Pro Glu Val Lys
            205                 210                 215
GAT GAA GAA AAG GAA GTT GCT AAA GAA ATT CCT GAG GTA AAG GCT GAA      964
Asp Glu Glu Lys Glu Val Ala Lys Glu Ile Pro Glu Val Lys Ala Glu
        220                 225                 230
GAA AAA GCA GAT TCT AAA GAC TGT ATG GAG GCA ACC CCT GAA GTA AAA     1012
Glu Lys Ala Asp Ser Lys Asp Cys Met Glu Ala Thr Pro Glu Val Lys
    235                 240                 245
GAA GAT CCT AAA GAA GTC CCC CAG GTA AAG GCA GAT GAT AAA GAA CAG     1060
Glu Asp Pro Lys Glu Val Pro Gln Val Lys Ala Asp Asp Lys Glu Gln
250                 255                 260                 265
CCT AAA GCA ACA GAG GCT AAG GCA AGG GCT GCA GTA AGA GAG ACT CAT     1108
Pro Lys Ala Thr Glu Ala Lys Ala Arg Ala Ala Val Arg Glu Thr His
                270                 275                 280
AAA AGA GTT CCT GAG GAA AGG CTT CGG GAC AGT GTA GAT CTT AAA AGA     1156
Lys Arg Val Pro Glu Glu Arg Leu Arg Asp Ser Val Asp Leu Lys Arg
            285                 290                 295
GCT AGG AAG GGA AAG CCT AAA AGA GAA GAC CCT AAA GGC ATT CCT GAC     1204
Ala Arg Lys Gly Lys Pro Lys Arg Glu Asp Pro Lys Gly Ile Pro Asp
        300                 305                 310
TAT TGG CTG ATT GTT TTA AAG AAT GTT GAC AAG CTC GGG CCT ATG ATT     1252
Tyr Trp Leu Ile Val Leu Lys Asn Val Asp Lys Leu Gly Pro Met Ile
    315                 320                 325
CAG AAG TAT GAT GAG CCC ATT CTG AAG TTC TTG TCG GAT GTT AGC CTG     1300
Gln Lys Tyr Asp Glu Pro Ile Leu Lys Phe Leu Ser Asp Val Ser Leu
330                 335                 340                 345
AAG TTC TCA AAA CCT GGC CAG CCT GTA AGT TAC ACC TTT GAA TTT CAT     1348
Lys Phe Ser Lys Pro Gly Gln Pro Val Ser Tyr Thr Phe Glu Phe His
                350                 355                 360
```

```
TTT CTA CCC AAC CCA TAC TTC AGA AAT GAG GTG CTG GTG AAG ACA TAT     1396
Phe Leu Pro Asn Pro Tyr Phe Arg Asn Glu Val Leu Val Lys Thr Tyr
            365                 370                 375
ATA ATA AAG GCA AAA CCA GAT CAC AAT GAT CCC TTC TTT TCT TGG GGA     1444
Ile Ile Lys Ala Lys Pro Asp His Asn Asp Pro Phe Phe Ser Trp Gly
        380                 385                 390
TGG GAA ATT GAA GAT TGC AAA GGC TGC AAG ATA GAC CGG AGA AGA GGA     1492
Trp Glu Ile Glu Asp Cys Lys Gly Cys Lys Ile Asp Arg Arg Arg Gly
    395                 400                 405
AAA GAT GTT ACT GTG ACA ACT ACC CAG AGT CGC ACA ACT GCT ACT GGA     1540
Lys Asp Val Thr Val Thr Thr Thr Gln Ser Arg Thr Thr Ala Thr Gly
410                 415                 420                 425
GAA ATT GAA ATC CAG CCA AGA GTG GTT CCT AAT GCA TCA TTC TTC AAC     1588
Glu Ile Glu Ile Gln Pro Arg Val Val Pro Asn Ala Ser Phe Phe Asn
                430                 435                 440
TTC TTT AGT CCT CCT GAG ATT CCT ATG ATT GGG AAG CTG GAA CCA CGA     1636
Phe Phe Ser Pro Pro Glu Ile Pro Met Ile Gly Lys Leu Glu Pro Arg
            445                 450                 455
GAA GAT GCT ATC CTG GAT GAG GAC TTT GAA ATT GGG CAG ATT TTA CAT     1684
Glu Asp Ala Ile Leu Asp Glu Asp Phe Glu Ile Gly Gln Ile Leu His
        460                 465                 470
GAT AAT GTC ATC CTG AAA TCA ATC TAT TAC TAT ACT GGA GAA GTC AAT     1732
Asp Asn Val Ile Leu Lys Ser Ile Tyr Tyr Tyr Thr Gly Glu Val Asn
    475                 480                 485
GGT ACC TAC TAT CAA TTT GGC AAA CAT TAT GGA AAC AAG AAA TAC AGA     1780
Gly Thr Tyr Tyr Gln Phe Gly Lys His Tyr Gly Asn Lys Lys Tyr Arg
490                 495                 500                 505
AAA TAAGTCAATC TGAAAGATTT TTCAAGAATC TTAAAATCTC AAGAAGTGAA          1833
Lys
```

GCAGATTCAT ACAGCCTTGA AAAAAGTAAA ACCCTGACCT GTAACCTGAA CACTATTATT 1893
CCTTATAGTC AAGTTTTTGT GGTTTCTTGG TAGTCTATAT TTTAAAAATA GTCCTAAAAA 1953
GTGTCTAAGT GCCAGTTTAT TCTATCTAGG CTGTTGTAGT ATAATATTCT TCAAAATATG 2013
TAAGCTGTTG TCAATTATCT AAAGCATGTT AGTTTGGTGC TACACAGTGT TGATTTTTGT 2073
GATGTCCTTT GGTCATGTTT CTGTTAGACT GTAGCTGTGA AACTGTCAGA ATTGTTAACT 2133
GAAACAAATA TTTGCTTGAA AAAAAAAGTT CATGAAGTAC CAATGCAAGT GTTTTATTTT 2193
TTTTCTTTTT TCCAGCCCAT AAGACTAAGG GTTTAAATCT GCTTGCACTA GCTGTGCCTT 2253
CATTAGTTTG CTATAGAAAT CCAGTACTTA TAGTAAATAA AACAGTGTAT TTTGAAGTTT 2313
GACTGCTTGA AAAAGATTAG CATACATCTA ATGTGAAAAG ACCACATTTG ATTCAACTGA 2373
GACCTTGTGT ATGTGACATA TAGTGGCCTA TAAATTTAAT CATAATGATG TTATTGTTTA 2433
CCACTGAGGT GTTAATATAA CATAGTATTT TTGAAAAAGT TTCTTCATCT TATATTGTGT 2493
AATTGTAAAC TAAAGATACC GTGTTTTCTT TGTATTGTGT TCTACCTTCC CTTTCACTGA 2553
AAATGATCAC TTCATTTGAT ACTGTTTTTC ATGTTCTTGT ATTGCAACCT AAAATAAATA 2613
AATATTAAAG TGTGTTATAC TAT 2636

【０３３８】

配列番号：２２
配列の長さ：１７０
配列の型：アミノ酸
トポロジー：直線状
配列の種類：蛋白
配列：

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

【０３３９】

配列番号：２３
配列の長さ：５１０
配列の型：核酸
鎖の数：一本鎖
トポロジー：直線状
配列の種類：ＤＮＡ(genomic)
配列：

ATGACGGAGC TGCAGTCGGC ACTGCTACTG CGAAGACAGC TGGCAGAACT CAACAAAAAT 60
CCAGTGGAAG GCTTTTCTGC AGGTTTAATA GATGACAATG ATCTCTACCG ATGGGAAGTC 120
CTTATTATTG GCCCTCCAGA TACACTTTAT GAAGGTGGTG TTTTTAAGGC TCATCTTACT 180
TTCCCAAAAG ATTATCCCCT CCGACCTCCT AAAATGAAAT TCATTACAGA AATCTGGCAC 240
CCAAATGTTG ATAAAAATGG TGATGTGTGC ATTTCTATTC TTCATGAGCC TGGGGAAGAT 300
AAGTATGGTT ATGAAAAGCC AGAGGAACGC TGGCTCCCTA TCCACACTGT GGAAACCATC 360
ATGATTAGTG TCATTTCTAT GCTGGCAGAC CCTAATGGAG ACTCACCTGC TAATGTTGAT 420
GCTGCGAAAG AATGGAGGGA AGATAGAAAT GGAGAATTTA AAAGAAAAGT TGCCCGCTGT 480
GTAAGAAAAA GCCAAGAGAC TGCTTTTGAG 510

【０３４０】

配列番号：２４
配列の長さ：６１７
配列の型：核酸
鎖の数：一本鎖
トポロジー：直線状
配列の種類：ＤＮＡ(genomic)
起源：
　生物名：ヒト胎児脳ｃＤＮＡライブラリー
　株名：ＧＥＮ－４２３Ａ１２
配列の特徴：
　特徴を表わす記号：ＣＤＳ
　存在位置：１９．．５２８
　特徴を決定した方法：Ｅ
配列：

```
GGGCCCTCGG CAGGGAGG ATG ACG GAG CTG CAG TCG GCA CTG CTA CTG CGA      51
                     Met Thr Glu Leu Gln Ser Ala Leu Leu Leu Arg
                       1               5                  10
AGA CAG CTG GCA GAA CTC AAC AAA AAT CCA GTG GAA GGC TTT TCT GCA      99
Arg Gln Leu Ala Glu Leu Asn Lys Asn Pro Val Glu Gly Phe Ser Ala
            15                  20                  25
GGT TTA ATA GAT GAC AAT GAT CTC TAC CGA TGG GAA GTC CTT ATT ATT     147
Gly Leu Ile Asp Asp Asn Asp Leu Tyr Arg Trp Glu Val Leu Ile Ile
        30                  35                  40
GGC CCT CCA GAT ACA CTT TAT GAA GGT GGT GTT TTT AAG GCT CAT CTT     195
Gly Pro Pro Asp Thr Leu Tyr Glu Gly Gly Val Phe Lys Ala His Leu
    45                  50                  55
ACT TTC CCA AAA GAT TAT CCC CTC CGA CCT CCT AAA ATG AAA TTC ATT     243
Thr Phe Pro Lys Asp Tyr Pro Leu Arg Pro Pro Lys Met Lys Phe Ile
60                  65                  70                  75
```

```
ACA GAA ATC TGG CAC CCA AAT GTT GAT AAA AAT GGT GAT GTG TGC ATT      291
Thr Glu Ile Trp His Pro Asn Val Asp Lys Asn Gly Asp Val Cys Ile
                80                  85                  90
TCT ATT CTT CAT GAG CCT GGG GAA GAT AAG TAT GGT TAT GAA AAG CCA      339
Ser Ile Leu His Glu Pro Gly Glu Asp Lys Tyr Gly Tyr Glu Lys Pro
            95                  100                 105
GAG GAA CGC TGG CTC CCT ATC CAC ACT GTG GAA ACC ATC ATG ATT AGT      387
Glu Glu Arg Trp Leu Pro Ile His Thr Val Glu Thr Ile Met Ile Ser
        110                 115                 120
GTC ATT TCT ATG CTG GCA GAC CCT AAT GGA GAC TCA CCT GCT AAT GTT      435
Val Ile Ser Met Leu Ala Asp Pro Asn Gly Asp Ser Pro Ala Asn Val
    125                 130                 135
GAT GCT GCG AAA GAA TGG AGG GAA GAT AGA AAT GGA GAA TTT AAA AGA      483
Asp Ala Ala Lys Glu Trp Arg Glu Asp Arg Asn Gly Glu Phe Lys Arg
140                 145                 150                 155
AAA GTT GCC CGC TGT GTA AGA AAA AGC CAA GAG ACT GCT TTT GAG          528
Lys Val Ala Arg Cys Val Arg Lys Ser Gln Glu Thr Ala Phe Glu
                160                 165                 170
TGACATTTAT TTAGCAGCTA GTAACTTCAC TTATTTCAGG GTCTCCAATT GAGAAACATG    588
GCACTGTTTT TCCTGCACTC TACCCACCG                                      617
```

【０３４１】

配列番号：２５
配列の長さ：３７４
配列の型：アミノ酸
トポロジー：直線状
配列の種類：蛋白
配列：

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

180 185 190
Pro Leu Cys Ala Ser Asp Gly Lys Ser Tyr Asp Asn Ala Cys Gln Ile
195 200 205
Lys Glu Ala Ser Cys Gln Lys Gln Glu Lys Ile Glu Val Met Ser Leu
210 215 220
Gly Arg Cys Gln Asp Asn Thr Thr Thr Thr Thr Lys Ser Glu Asp Gly
225 230 235 240
His Tyr Ala Arg Thr Asp Tyr Ala Glu Asn Ala Asn Lys Leu Glu Glu
245 250 255
Ser Ala Arg Glu His His Ile Pro Cys Pro Glu His Tyr Asn Gly Phe
260 265 270
Cys Met His Gly Lys Cys Glu His Ser Ile Asn Met Gln Glu Pro Ser
275 280 285
Cys Arg Cys Asp Ala Gly Tyr Thr Gly Gln His Cys Glu Lys Lys Asp
290 295 300
Tyr Ser Val Leu Tyr Val Val Pro Gly Pro Val Arg Phe Gln Tyr Val
305 310 315 320
Leu Ile Ala Ala Val Ile Gly Thr Ile Gln Ile Ala Val Ile Cys Val
325 330 335
Val Val Leu Cys Ile Thr Arg Lys Cys Pro Arg Ser Asn Arg Ile His
340 345 350
Arg Gln Lys Gln Asn Thr Gly His Tyr Ser Ser Asp Asn Thr Thr Arg
355 360 365
Ala Ser Thr Arg Leu Ile
370

【０３４２】

配列番号：２６

配列の長さ：１１２２

配列の型：核酸

鎖の数：一本鎖

トポロジー：直線状

配列の種類：ＤＮＡ(genomic)

配列：

ATGGTGCTGT GGGAGTCCCC GCGGCAGTGC AGCAGCTGGA CACTTTGCGA GGGCTTTTGC 60

TGGCTGCTGC TGCTGCCCGT CATGCTACTC ATCGTAGCCC GCCCGGTGAA GCTCGCTGCT 120

TTCCCTACCT CCTTAAGTGA CTGCCAAACG CCCACCGGCT GGAATTGCTC TGGTTATGAT 180

GACAGAGAAA ATGATCTCTT CCTCTGTGAC ACCAACACCT GTAAATTTGA TGGGGAATGT 240

TTAAGAATTG GAGACACTGT GACTTGCGTC TGTCAGTTCA AGTGCAACAA TGACTATGTG 300

CCTGTGTGTG GCTCCAATGG GGAGAGCTAC CAGAATGAGT GTTACCTGCG ACAGGCTGCA 360

TGCAAACAGC AGAGTGAGAT ACTTGTGGTG TCAGAAGGAT CATGTGCCAC AGATGCAGGA 420

TCAGGATCTG GAGATGGAGT CCATGAAGGC TCTGGAGAAA CTAGTCAAAA GGAGACATCC 480

ACCTGTGATA TTTGCCAGTT TGGTGCAGAA TGTGACGAAG ATGCCGAGGA TGTCTGGTGT 540

GTGTGTAATA TTGACTGTTC TCAAACCAAC TTCAATCCCC TCTGCGCTTC TGATGGGAAA 600

TCTTATGATA ATGCATGCCA AATCAAAGAA GCATCGTGTC AGAAACAGGA GAAAATTGAA 660

GTCATGTCTT TGGGTCGATG TCAAGATAAC ACAACTACAA CTACTAAGTC TGAAGATGGG 720

CATTATGCAA GAACAGATTA TGCAGAGAAT GCTAACAAAT TAGAAGAAAG TGCCAGAGAA 780

CACCACATAC CTTGTCCGGA ACATTACAAT GGCTTCTGCA TGCATGGGAA GTGTGAGCAT 840

TCTATCAATA TGCAGGAGCC ATCTTGCAGG TGTGATGCTG GTTATACTGG ACAACACTGT 900

GAAAAAAAGG ACTACAGTGT TCTATACGTT GTTCCCGGTC CTGTACGATT TCAGTATGTC 960

TTAATCGCAG CTGTGATTGG AACAATTCAG ATTGCTGTCA TCTGTGTGGT GGTCCTCTGC 1020

ATCACAAGGA AATGCCCCAG AAGCAACAGA ATTCACAGAC AGAAGCAAAA TACAGGGCAC 1080

TACAGTTCAG ACAATACAAC AAGAGCGTCC ACGAGGTTAA TC 1122

【０３４３】

配列番号：２７
配列の長さ：１７２１
配列の型：核酸
鎖の数：一本鎖
トポロジー：直線状
配列の種類：ＤＮＡ(genomic)
起源：
　生物名：ヒト胎児脳ｃＤＮＡライブラリー
　株名：ＧＥＮ－０９２Ｅ１０
配列の特徴：
　特徴を表わす記号：ＣＤＳ
　存在位置：３６８．．１４８９
　特徴を決定した方法：Ｅ
配列：

```
CTGCGGGGCG CCTTGACTCT CCCTCCACCC TGCCTCCTCG GGCTCCACTC GTCTGCCCCT     60
GGACTCCCGT CTCCTCCTGT CCTCCGGCTT CCCAGAGCTC CCTCCTTATG GCAGCAGCTT    120
CCCGCGTCTC CGGCGCAGCT TCTCAGCGGA CGACCCTCTC GCTCCGGGGC TGAGCCAGTC    180
CCTGGATGTT GCTGAAACTC TCGAGATCAT GCGCGGGTTT GGCTGCTGCT TCCCCGCCGG    240
GTGCCACTGC CACCGCCGCC GCCTCTGCTG CCGCCGTCCG CGGGATGCTC AGTAGCCCGC    300
TGCCCGGCCC CCGCGATCCT GTGTTCCTCG GAAGCCGTTT GCTGCTGCAG AGTTGCACGA    360
ACTAGTC ATG GTG CTG TGG GAG TCC CCG CGG CAG TGC AGC AGC TGG ACA       409
       Met Val Leu Trp Glu Ser Pro Arg Gln Cys Ser Ser Trp Thr
         1               5                  10
CTT TGC GAG GGC TTT TGC TGG CTG CTG CTG CTG CCC GTC ATG CTA CTC       457
Leu Cys Glu Gly Phe Cys Trp Leu Leu Leu Leu Pro Val Met Leu Leu
 15                  20                  25                  30
ATC GTA GCC CGC CCG GTG AAG CTC GCT GCT TTC CCT ACC TCC TTA AGT       505
Ile Val Ala Arg Pro Val Lys Leu Ala Ala Phe Pro Thr Ser Leu Ser
                 35                  40                  45
```

```
GAC TGC CAA ACG CCC ACC GGC TGG AAT TGC TCT GGT TAT GAT GAC AGA      553
Asp Cys Gln Thr Pro Thr Gly Trp Asn Cys Ser Gly Tyr Asp Asp Arg
            50                  55                  60
GAA AAT GAT CTC TTC CTC TGT GAC ACC AAC ACC TGT AAA TTT GAT GGG      601
Glu Asn Asp Leu Phe Leu Cys Asp Thr Asn Thr Cys Lys Phe Asp Gly
        65                  70                  75
GAA TGT TTA AGA ATT GGA GAC ACT GTG ACT TGC GTC TGT CAG TTC AAG      649
Glu Cys Leu Arg Ile Gly Asp Thr Val Thr Cys Val Cys Gln Phe Lys
    80                  85                  90
TGC AAC AAT GAC TAT GTG CCT GTG TGT GGC TCC AAT GGG GAG AGC TAC      697
Cys Asn Asn Asp Tyr Val Pro Val Cys Gly Ser Asn Gly Glu Ser Tyr
95                  100                 105                 110
CAG AAT GAG TGT TAC CTG CGA CAG GCT GCA TGC AAA CAG CAG AGT GAG      745
Gln Asn Glu Cys Tyr Leu Arg Gln Ala Ala Cys Lys Gln Gln Ser Glu
                115                 120                 125
ATA CTT GTG GTG TCA GAA GGA TCA TGT GCC ACA GAT GCA GGA TCA GGA      793
Ile Leu Val Val Ser Glu Gly Ser Cys Ala Thr Asp Ala Gly Ser Gly
            130                 135                 140
TCT GGA GAT GGA GTC CAT GAA GGC TCT GGA GAA ACT AGT CAA AAG GAG      841
Ser Gly Asp Gly Val His Glu Gly Ser Gly Glu Thr Ser Gln Lys Glu
        145                 150                 155
ACA TCC ACC TGT GAT ATT TGC CAG TTT GGT GCA GAA TGT GAC GAA GAT      889
Thr Ser Thr Cys Asp Ile Cys Gln Phe Gly Ala Glu Cys Asp Glu Asp
    160                 165                 170
GCC GAG GAT GTC TGG TGT GTG TGT AAT ATT GAC TGT TCT CAA ACC AAC      937
Ala Glu Asp Val Trp Cys Val Cys Asn Ile Asp Cys Ser Gln Thr Asn
175                 180                 185                 190
TTC AAT CCC CTC TGC GCT TCT GAT GGG AAA TCT TAT GAT AAT GCA TGC      985
Phe Asn Pro Leu Cys Ala Ser Asp Gly Lys Ser Tyr Asp Asn Ala Cys
```

```
                195                 200                 205
CAA ATC AAA GAA GCA TCG TGT CAG AAA CAG GAG AAA ATT GAA GTC ATG     1033
Gln Ile Lys Glu Ala Ser Cys Gln Lys Gln Glu Lys Ile Glu Val Met
            210                 215                 220
TCT TTG GGT CGA TGT CAA GAT AAC ACA ACT ACA ACT ACT AAG TCT GAA     1081
Ser Leu Gly Arg Cys Gln Asp Asn Thr Thr Thr Thr Thr Lys Ser Glu
        225                 230                 235
GAT GGG CAT TAT GCA AGA ACA GAT TAT GCA GAG AAT GCT AAC AAA TTA     1129
Asp Gly His Tyr Ala Arg Thr Asp Tyr Ala Glu Asn Ala Asn Lys Leu
    240                 245                 250
GAA GAA AGT GCC AGA GAA CAC CAC ATA CCT TGT CCG GAA CAT TAC AAT     1177
Glu Glu Ser Ala Arg Glu His His Ile Pro Cys Pro Glu His Tyr Asn
255                 260                 265                 270
GGC TTC TGC ATG CAT GGG AAG TGT GAG CAT TCT ATC AAT ATG CAG GAG     1225
Gly Phe Cys Met His Gly Lys Cys Glu His Ser Ile Asn Met Gln Glu
                275                 280                 285
CCA TCT TGC AGG TGT GAT GCT GGT TAT ACT GGA CAA CAC TGT GAA AAA     1273
Pro Ser Cys Arg Cys Asp Ala Gly Tyr Thr Gly Gln His Cys Glu Lys
            290                 295                 300
AAG GAC TAC AGT GTT CTA TAC GTT GTT CCC GGT CCT GTA CGA TTT CAG     1321
Lys Asp Tyr Ser Val Leu Tyr Val Val Pro Gly Pro Val Arg Phe Gln
        305                 310                 315
TAT GTC TTA ATC GCA GCT GTG ATT GGA ACA ATT CAG ATT GCT GTC ATC     1369
Tyr Val Leu Ile Ala Ala Val Ile Gly Thr Ile Gln Ile Ala Val Ile
    320                 325                 330
TGT GTG GTG GTC CTC TGC ATC ACA AGG AAA TGC CCC AGA AGC AAC AGA     1417
Cys Val Val Val Leu Cys Ile Thr Arg Lys Cys Pro Arg Ser Asn Arg
335                 340                 345                 350
ATT CAC AGA CAG AAG CAA AAT ACA GGG CAC TAC AGT TCA GAC AAT ACA     1465
```

```
Ile His Arg Gln Lys Gln Asn Thr Gly His Tyr Ser Ser Asp Asn Thr
                355                 360                 365
ACA AGA GCG TCC ACG AGG TTA ATC TAA AGGGAGCATG TTTCACAGTG          1512
Thr Arg Ala Ser Thr Arg Leu Ile
            370
GCTGGACTAC CGAGAGCTTG GACTACACAA TACAGTATTA TAGACAAAAG AATAAGACAA  1572
GAGATCTACA CATGTTGCCT TGCATTTGTG GTAATCTACA CCAATGAAAA CATGTACTAC  1632
AGCTATATTT GATTATGTAT GGATATATTT GAAATAGTAT ACATTGTCTT GATGTTTTTT  1692
CTGTAATGTA AATAAACTAT TTATATCAC                                    1721
```

【０３４４】

配列番号：２８
配列の長さ：８１７
配列の型：アミノ酸
トポロジー：直線状
配列の種類：蛋白
配列：

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

180 185 190
Ala Ile Lys Pro Tyr Ile Val His Arg Cys Arg Gln Ser Ile Asn Phe
195 200 205
Ser Leu Gln Cys Ala Leu Leu Leu Gly Ala Tyr Ser Ser Asp Met His
210 215 220
Ile Ser Thr Gln Arg His Ser Arg Gly Thr Lys Leu Arg Lys Leu Ile
225 230 235 240
Leu Ser Asp Glu Leu Lys Pro Ala His Arg Lys Arg Glu Leu Pro Ser
245 250 255
Leu Ser Pro Ala Pro Asp Thr Gly Leu Ser Pro Ser Lys Arg Thr His
260 265 270
Gln Arg Ser Lys Ser Asp Ala Thr Ala Ser Ile Ser Leu Ser Ser Asn
275 280 285
Leu Lys Arg Thr Ala Ser Asn Pro Lys Val Glu Asn Glu Asp Glu Glu
290 295 300
Leu Ser Ser Ser Thr Glu Ser Ile Asp Asn Ser Phe Ser Ser Pro Val
305 310 315 320
Arg Leu Ala Pro Glu Arg Glu Phe Ile Lys Ser Leu Met Ala Ile Gly
325 330 335
Lys Arg Leu Ala Thr Leu Pro Thr Lys Glu Gln Lys Thr Gln Arg Leu
340 345 350
Ile Ser Glu Leu Ser Leu Leu Asn His Lys Leu Pro Ala Arg Val Trp
355 360 365
Leu Pro Thr Ala Gly Phe Asp His His Val Val Arg Val Pro His Thr
370 375 380
Gln Ala Val Val Leu Asn Ser Lys Asp Lys Ala Pro Tyr Leu Ile Tyr
385 390 395 400
Val Glu Val Leu Glu Cys Glu Asn Phe Asp Thr Thr Ser Val Pro Ala
405 410 415

Arg Ile Pro Glu Asn Arg Ile Arg Ser Thr Arg Ser Val Glu Asn Leu
420 425 430
Pro Glu Cys Gly Ile Thr His Glu Gln Arg Ala Gly Ser Phe Ser Thr
435 440 445
Val Pro Asn Tyr Asp Asn Asp Asp Glu Ala Trp Ser Val Asp Asp Ile
450 455 460
Gly Glu Leu Gln Val Glu Leu Pro Glu Val His Thr Asn Ser Cys Asp
465 470 475 480
Asn Ile Ser Gln Phe Ser Val Asp Ser Ile Thr Ser Gln Glu Ser Lys
485 490 495
Glu Pro Val Phe Ile Ala Ala Gly Asp Ile Arg Arg Arg Leu Ser Glu
500 505 510
Gln Leu Ala His Thr Pro Thr Ala Phe Lys Arg Asp Pro Glu Asp Pro
515 520 525
Ser Ala Val Ala Leu Lys Glu Pro Trp Gln Glu Lys Val Arg Arg Ile
530 535 540
Arg Glu Gly Ser Pro Tyr Gly His Leu Pro Asn Trp Arg Leu Leu Ser
545 550 555 560
Val Ile Val Lys Cys Gly Asp Asp Leu Arg Gln Glu Leu Leu Ala Phe
565 570 575
Gln Val Leu Lys Gln Leu Gln Ser Ile Trp Glu Gln Glu Arg Val Pro
580 585 590
Leu Trp Ile Lys Pro Ile Gln Asp Ser Cys Glu Ile Thr Thr Asp Ser
595 600 605
Gly Met Ile Glu Pro Val Val Asn Ala Val Ser Ile His Gln Val Lys
610 615 620
Lys Gln Ser Gln Leu Ser Leu Leu Asp Tyr Phe Leu Gln Glu His Gly
625 630 635 640
Ser Tyr Thr Thr Glu Ala Phe Leu Ser Ala Gln Arg Asn Phe Val Gln

645 650 655

Ser Cys Ala Gly Tyr Cys Leu Val Cys Tyr Leu Leu Gln Val Lys Asp

660 665 670

Arg His Asn Gly Asn Ile Leu Leu Asp Ala Glu Gly His Ile Ile His

675 680 685

Ile Asp Phe Gly Phe Ile Leu Ser Ser Ser Pro Arg Asn Leu Gly Phe

690 695 700

Glu Thr Ser Ala Phe Lys Leu Thr Thr Glu Phe Val Asp Val Met Gly

705 710 715 720

Gly Leu Asp Gly Asp Met Phe Asn Tyr Tyr Lys Met Leu Met Leu Gln

725 730 735

Gly Leu Ile Ala Ala Arg Lys His Met Asp Lys Val Val Gln Ile Val

740 745 750

Glu Ile Met Gln Gln Gly Ser Gln Leu Pro Cys Phe His Gly Ser Ser

755 760 765

Thr Ile Arg Asn Leu Lys Glu Arg Phe His Met Ser Met Thr Glu Glu

770 775 780

Gln Leu Gln Leu Leu Val Glu Gln Met Val Asp Gly Ser Met Arg Ser

785 790 795 800

Ile Thr Thr Lys Leu Tyr Asp Gly Phe Gln Tyr Leu Thr Asn Gly Ile

805 810 815

Met

【０３４５】

配列番号：２９

配列の長さ：２４５１

配列の型：核酸

鎖の数：一本鎖

トポロジー：直線状

配列の種類：ＤＮＡ(genomic)

配列：

ATGGGAGATA CAGTAGTGGA GCCTGCCCCC TTGAAGCCAA CTTCTGAGCC CACTTCTGGC 60

CCACCAGGGA ATAATGGGGG GTCCCTGCTA AGTGTCATCA CGGAGGGGGT CGGGGAACTA 120

TCAGTGATTG ACCCTGAGGT GGCCCAGAAG GCCTGCCAGG AGGTGTTGGA GAAAGTCAAG 180

CTTTTGCATG GAGGCGTGGC AGTCTCTAGC AGAGGCACCC CACTGGAGTT GGTCAATGGG 240

GATGGTGTGG ACAGTGAGAT CCGTTGCCTA GATGATCCAC CTGCCCAGAT CAGGGAGGAG 300

GAAGATGAGA TGGGGGCCGC TGTGGCCTCA GGCACAGCCA AAGGAGCAAG AAGACGGCGG 360

CAGAACAACT CAGCTAAACA GTCTTGGCTG CTGAGGCTGT TTGAGTCAAA ACTGTTTGAC 420

ATCTCCATGG CCATTTCATA CCTGTATAAC TCCAAGGAGC CTGGAGTACA AGCCTACATT 480

GGCAACCGGC TCTTCTGCTT TCGCAACGAG GACGTGGACT TCTATCTGCC CCAGTTGCTT 540

AACATGTACA TCCACATGGA TGAGGACGTG GGTGATGCCA TTAAGCCCTA CATAGTCCAC 600

CGTTGCCGCC AGAGCATTAA CTTTTCCCTC CAGTGTGCCC TGTTGCTTGG GGCCTATTCT 660

TCAGACATGC ACATTTCCAC TCAACGACAC TCCCGTGGGA CCAAGCTACG GAAGCTGATC 720

CTCTCAGATG AGCTAAAGCC AGCTCACAGG AAGAGGGAGC TGCCCTCCTT GAGCCCGGCC 780

CCTGATACAG GGCTGTCTCC CTCCAAAAGG ACTCACCAGC GCTCTAAGTC AGATGCCACT 840

GCCAGCATAA GTCTCAGCAG CAACCTGAAA CGAACAGCCA GCAACCCTAA AGTGGAGAAT 900

GAGGATGAGG AGCTCTCCTC CAGCACCGAG AGTATTGATA ATTCATTCAG TTCCCCTGTT 960

CGACTGGCTC CTGAGAGAGA ATTCATCAAG TCCCTGATGG CGATCGGCAA GCGGCTGGCC 1020

ACGCTCCCCA CCAAAGAGCA GAAAACACAG AGGCTGATCT CAGAGCTCTC CCTGCTCAAC 1080

CATAAGCTCC CTGCCCGAGT CTGGCTGCCC ACTGCTGGCT TTGACCACCA CGTGGTCCGT 1140

GTACCCCACA CACAGGCTGT TGTCCTCAAC TCCAAGGACA AGGCTCCCTA CCTGATTTAT 1200

GTGGAAGTCC TTGAATGTGA AAACTTTGAC ACCACCAGTG TCCCTGCCCG GATCCCCGAG 1260

AACCGAATTC GGAGTACGAG GTCCGTAGAA AACTTGCCCG AATGTGGTAT TACCCATGAG 1320

CAGCGAGCTG GCAGCTTCAG CACTGTGCCC AACTATGACA ACGATGATGA GGCCTGGTCG 1380
GTGGATGACA TAGGCGAGCT GCAAGTGGAG CTCCCCGAAG TGCATACCAA CAGCTGTGAC 1440
AACATCTCCC AGTTCTCTGT GGACAGCATC ACCAGCCAGG AGAGCAAGGA GCCTGTGTTC 1500
ATTGCAGCAG GGGACATCCG CCGGCGCCTT TCGGAACAGC TGGCTCATAC CCCGACAGCC 1560
TTCAAACGAG ACCCAGAAGA TCCTTCTGCA GTTGCTCTCA AAGAGCCCTG GCAGGAGAAA 1620
GTACGGCGGA TCAGAGAGGG CTCCCCCTAC GGCCATCTCC CCAATTGGCG GCTCCTGTCA 1680
GTCATTGTCA AGTGTGGGGA TGACCTTCGG CAAGAGCTTC TGGCCTTTCA GGTGTTGAAG 1740
CAACTGCAGT CCATTTGGGA ACAGGAGCGA GTGCCCCTTT GGATCAAGCC AATACAAGAT 1800
TCTTGTGAAA TTACGACTGA TAGTGGCATG ATTGAACCAG TGGTCAATGC TGTGTCCATC 1860
CATCAGGTGA AGAAACAGTC ACAGCTCTCC TTGCTCGATT ACTTCCTACA GGAGCACGGC 1920
AGTTACACCA CTGAGGCATT CCTCAGTGCA CAGCGCAATT TTGTGCAAAG TTGTGCTGGG 1980
TACTGCTTGG TCTGCTACCT GCTGCAAGTC AAGGACAGAC ACAATGGGAA TATCCTTTTG 2040
GACGCAGAAG GCCACATCAT CCACATCGAC TTTGGCTTCA TCCTCTCCAG CTCACCCCGA 2100
AATCTGGGCT TTGAGACGTC AGCCTTTAAG CTGACCACAG AGTTTGTGGA TGTGATGGGC 2160
GGCCTGGATG GCGACATGTT CAACTACTAT AAGATGCTGA TGCTGCAAGG GCTGATTGCC 2220
GCTCGGAAAC ACATGGACAA GGTGGTGCAG ATCGTGGAGA TCATGCAGCA AGGTTCTCAG 2280
CTTCCTTGCT TCCATGGCTC CAGCACCATT CGAAACCTCA AAGAGAGGTT CCACATGAGC 2340
ATGACTGAGG AGCAGCTGCA GCTGCTGGTG GAGCAGATGG TGGATGGCAG TATGCGGTCT 2400
ATCACCACCA AACTCTATGA CGGCTTCCAG TACCTCACCA ACGGCATCAT G 2451

【０３４６】

配列番号：３０
配列の長さ：３６０２
配列の型：核酸
鎖の数：一本鎖
トポロジー：直線状
配列の種類：ＤＮＡ(genomic)
起源：
　生物名：ヒト胎児脳ｃＤＮＡライブラリー
　株名：ＧＥＮ－４２８Ｂ１２ｃ２
配列の特徴：
　特徴を表わす記号：ＣＤＳ
　存在位置：４２９．．２８７９
　特徴を決定した方法：Ｅ
配列：

```
GGTGGCTCAC GCCTGTAATC CCAGCACTTT GGGAGGACAA GGCAGATCCC TTGAGCCCAG     60
GAGGTAGAGG CTGCAGTGAG CTGTGATGGT GCCACTGCAC TCCAGCCTGG GCAATGAAGC    120
AAGACCCTAT CTGAAAAAAA AAATTTTTAA AAAAGGCAAA GATGGGCCTG GGGCACCAAA    180
TATTCCAGAG GAAAGGGAAC GTGTGTACTC CTTGAGGTGG GGAACATGAC CCACTTGAGG    240
TGCAGAAAGA AGACTTGTAT GGGGCTGGTG CAGCCTCCGC GGCCGCTGTC AGGGAAGCGC    300
AGGCGGCCAA TGGAACCCGG GAGCGGTCGC TGCTGCTGAG GCGGCAGTGT CGGCAGTCCA    360
ACCGCGACTG CCCGCACCCC CTCCGCGGGG TCCCCCAGAG CTTGGAAGCT CGAAGTCTGG    420
CTGTGGCC ATG GGA GAT ACA GTA GTG GAG CCT GCC CCC TTG AAG CCA ACT      470
         Met Gly Asp Thr Val Val Glu Pro Ala Pro Leu Lys Pro Thr
           1               5                  10
TCT GAG CCC ACT TCT GGC CCA CCA GGG AAT AAT GGG GGG TCC CTG CTA      518
Ser Glu Pro Thr Ser Gly Pro Pro Gly Asn Asn Gly Gly Ser Leu Leu
 15                  20                  25                  30
AGT GTC ATC ACG GAG GGG GTC GGG GAA CTA TCA GTG ATT GAC CCT GAG      566
Ser Val Ile Thr Glu Gly Val Gly Glu Leu Ser Val Ile Asp Pro Glu
```

```
                    35                  40                  45
GTG GCC CAG AAG GCC TGC CAG GAG GTG TTG GAG AAA GTC AAG CTT TTG      614
Val Ala Gln Lys Ala Cys Gln Glu Val Leu Glu Lys Val Lys Leu Leu
                50                  55                  60
CAT GGA GGC GTG GCA GTC TCT AGC AGA GGC ACC CCA CTG GAG TTG GTC      662
His Gly Gly Val Ala Val Ser Ser Arg Gly Thr Pro Leu Glu Leu Val
            65                  70                  75
AAT GGG GAT GGT GTG GAC AGT GAG ATC CGT TGC CTA GAT GAT CCA CCT      710
Asn Gly Asp Gly Val Asp Ser Glu Ile Arg Cys Leu Asp Asp Pro Pro
        80                  85                  90
GCC CAG ATC AGG GAG GAG GAA GAT GAG ATG GGG GCC GCT GTG GCC TCA      758
Ala Gln Ile Arg Glu Glu Glu Asp Glu Met Gly Ala Ala Val Ala Ser
    95                  100                 105                 110
GGC ACA GCC AAA GGA GCA AGA AGA CGG CGG CAG AAC AAC TCA GCT AAA      806
Gly Thr Ala Lys Gly Ala Arg Arg Arg Arg Gln Asn Asn Ser Ala Lys
                115                 120                 125
CAG TCT TGG CTG CTG AGG CTG TTT GAG TCA AAA CTG TTT GAC ATC TCC      854
Gln Ser Trp Leu Leu Arg Leu Phe Glu Ser Lys Leu Phe Asp Ile Ser
            130                 135                 140
ATG GCC ATT TCA TAC CTG TAT AAC TCC AAG GAG CCT GGA GTA CAA GCC      902
Met Ala Ile Ser Tyr Leu Tyr Asn Ser Lys Glu Pro Gly Val Gln Ala
        145                 150                 155
TAC ATT GGC AAC CGG CTC TTC TGC TTT CGC AAC GAG GAC GTG GAC TTC      950
Tyr Ile Gly Asn Arg Leu Phe Cys Phe Arg Asn Glu Asp Val Asp Phe
    160                 165                 170
TAT CTG CCC CAG TTG CTT AAC ATG TAC ATC CAC ATG GAT GAG GAC GTG      998
Tyr Leu Pro Gln Leu Leu Asn Met Tyr Ile His Met Asp Glu Asp Val
175                 180                 185                 190
GGT GAT GCC ATT AAG CCC TAC ATA GTC CAC CGT TGC CGC CAG AGC ATT     1046
```

Gly Asp Ala Ile Lys Pro Tyr Ile Val His Arg Cys Arg Gln Ser Ile
195 200 205

AAC TTT TCC CTC CAG TGT GCC CTG TTG CTT GGG GCC TAT TCT TCA GAC 1094
Asn Phe Ser Leu Gln Cys Ala Leu Leu Leu Gly Ala Tyr Ser Ser Asp
210 215 220

ATG CAC ATT TCC ACT CAA CGA CAC TCC CGT GGG ACC AAG CTA CGG AAG 1142
Met His Ile Ser Thr Gln Arg His Ser Arg Gly Thr Lys Leu Arg Lys
225 230 235

CTG ATC CTC TCA GAT GAG CTA AAG CCA GCT CAC AGG AAG AGG GAG CTG 1190
Leu Ile Leu Ser Asp Glu Leu Lys Pro Ala His Arg Lys Arg Glu Leu
240 245 250

CCC TCC TTG AGC CCG GCC CCT GAT ACA GGG CTG TCT CCC TCC AAA AGG 1238
Pro Ser Leu Ser Pro Ala Pro Asp Thr Gly Leu Ser Pro Ser Lys Arg
255 260 265 270

ACT CAC CAG CGC TCT AAG TCA GAT GCC ACT GCC AGC ATA AGT CTC AGC 1286
Thr His Gln Arg Ser Lys Ser Asp Ala Thr Ala Ser Ile Ser Leu Ser
275 280 285

AGC AAC CTG AAA CGA ACA GCC AGC AAC CCT AAA GTG GAG AAT GAG GAT 1334
Ser Asn Leu Lys Arg Thr Ala Ser Asn Pro Lys Val Glu Asn Glu Asp
290 295 300

GAG GAG CTC TCC TCC AGC ACC GAG AGT ATT GAT AAT TCA TTC AGT TCC 1382
Glu Glu Leu Ser Ser Ser Thr Glu Ser Ile Asp Asn Ser Phe Ser Ser
305 310 315

CCT GTT CGA CTG GCT CCT GAG AGA GAA TTC ATC AAG TCC CTG ATG GCG 1430
Pro Val Arg Leu Ala Pro Glu Arg Glu Phe Ile Lys Ser Leu Met Ala
320 325 330

ATC GGC AAG CGG CTG GCC ACG CTC CCC ACC AAA GAG CAG AAA ACA CAG 1478
Ile Gly Lys Arg Leu Ala Thr Leu Pro Thr Lys Glu Gln Lys Thr Gln
335 340 345 350

```
AGG CTG ATC TCA GAG CTC TCC CTG CTC AAC CAT AAG CTC CCT GCC CGA     1526
Arg Leu Ile Ser Glu Leu Ser Leu Leu Asn His Lys Leu Pro Ala Arg
                355                 360                 365
GTC TGG CTG CCC ACT GCT GGC TTT GAC CAC CAC GTG GTC CGT GTA CCC     1574
Val Trp Leu Pro Thr Ala Gly Phe Asp His His Val Val Arg Val Pro
            370                 375                 380
CAC ACA CAG GCT GTT GTC CTC AAC TCC AAG GAC AAG GCT CCC TAC CTG     1622
His Thr Gln Ala Val Val Leu Asn Ser Lys Asp Lys Ala Pro Tyr Leu
        385                 390                 395
ATT TAT GTG GAA GTC CTT GAA TGT GAA AAC TTT GAC ACC ACC AGT GTC     1670
Ile Tyr Val Glu Val Leu Glu Cys Glu Asn Phe Asp Thr Thr Ser Val
    400                 405                 410
CCT GCC CGG ATC CCC GAG AAC CGA ATT CGG AGT ACG AGG TCC GTA GAA     1718
Pro Ala Arg Ile Pro Glu Asn Arg Ile Arg Ser Thr Arg Ser Val Glu
415                 420                 425                 430
AAC TTG CCC GAA TGT GGT ATT ACC CAT GAG CAG CGA GCT GGC AGC TTC     1766
Asn Leu Pro Glu Cys Gly Ile Thr His Glu Gln Arg Ala Gly Ser Phe
                435                 440                 445
AGC ACT GTG CCC AAC TAT GAC AAC GAT GAT GAG GCC TGG TCG GTG GAT     1814
Ser Thr Val Pro Asn Tyr Asp Asn Asp Asp Glu Ala Trp Ser Val Asp
            450                 455                 460
GAC ATA GGC GAG CTG CAA GTG GAG CTC CCC GAA GTG CAT ACC AAC AGC     1862
Asp Ile Gly Glu Leu Gln Val Glu Leu Pro Glu Val His Thr Asn Ser
        465                 470                 475
TGT GAC AAC ATC TCC CAG TTC TCT GTG GAC AGC ATC ACC AGC CAG GAG     1910
Cys Asp Asn Ile Ser Gln Phe Ser Val Asp Ser Ile Thr Ser Gln Glu
    480                 485                 490
AGC AAG GAG CCT GTG TTC ATT GCA GCA GGG GAC ATC CGC CGG CGC CTT     1958
Ser Lys Glu Pro Val Phe Ile Ala Ala Gly Asp Ile Arg Arg Arg Leu
```

```
495                 500                 505                 510
TCG GAA CAG CTG GCT CAT ACC CCG ACA GCC TTC AAA CGA GAC CCA GAA     2006
Ser Glu Gln Leu Ala His Thr Pro Thr Ala Phe Lys Arg Asp Pro Glu
                515                 520                 525
GAT CCT TCT GCA GTT GCT CTC AAA GAG CCC TGG CAG GAG AAA GTA CGG     2054
Asp Pro Ser Ala Val Ala Leu Lys Glu Pro Trp Gln Glu Lys Val Arg
            530                 535                 540
CGG ATC AGA GAG GGC TCC CCC TAC GGC CAT CTC CCC AAT TGG CGG CTC     2102
Arg Ile Arg Glu Gly Ser Pro Tyr Gly His Leu Pro Asn Trp Arg Leu
        545                 550                 555
CTG TCA GTC ATT GTC AAG TGT GGG GAT GAC CTT CGG CAA GAG CTT CTG     2150
Leu Ser Val Ile Val Lys Cys Gly Asp Asp Leu Arg Gln Glu Leu Leu
    560                 565                 570
GCC TTT CAG GTG TTG AAG CAA CTG CAG TCC ATT TGG GAA CAG GAG CGA     2198
Ala Phe Gln Val Leu Lys Gln Leu Gln Ser Ile Trp Glu Gln Glu Arg
575                 580                 585                 590
GTG CCC CTT TGG ATC AAG CCA ATA CAA GAT TCT TGT GAA ATT ACG ACT     2246
Val Pro Leu Trp Ile Lys Pro Ile Gln Asp Ser Cys Glu Ile Thr Thr
                595                 600                 605
GAT AGT GGC ATG ATT GAA CCA GTG GTC AAT GCT GTG TCC ATC CAT CAG     2294
Asp Ser Gly Met Ile Glu Pro Val Val Asn Ala Val Ser Ile His Gln
            610                 615                 620
GTG AAG AAA CAG TCA CAG CTC TCC TTG CTC GAT TAC TTC CTA CAG GAG     2342
Val Lys Lys Gln Ser Gln Leu Ser Leu Leu Asp Tyr Phe Leu Gln Glu
        625                 630                 635
CAC GGC AGT TAC ACC ACT GAG GCA TTC CTC AGT GCA CAG CGC AAT TTT     2390
His Gly Ser Tyr Thr Thr Glu Ala Phe Leu Ser Ala Gln Arg Asn Phe
    640                 645                 650
GTG CAA AGT TGT GCT GGG TAC TGC TTG GTC TGC TAC CTG CTG CAA GTC     2438
```

```
Val Gln Ser Cys Ala Gly Tyr Cys Leu Val Cys Tyr Leu Leu Gln Val
655                 660                 665                 670
AAG GAC AGA CAC AAT GGG AAT ATC CTT TTG GAC GCA GAA GGC CAC ATC     2486
Lys Asp Arg His Asn Gly Asn Ile Leu Leu Asp Ala Glu Gly His Ile
                675                 680                 685
ATC CAC ATC GAC TTT GGC TTC ATC CTC TCC AGC TCA CCC CGA AAT CTG     2534
Ile His Ile Asp Phe Gly Phe Ile Leu Ser Ser Ser Pro Arg Asn Leu
            690                 695                 700
GGC TTT GAG ACG TCA GCC TTT AAG CTG ACC ACA GAG TTT GTG GAT GTG     2582
Gly Phe Glu Thr Ser Ala Phe Lys Leu Thr Thr Glu Phe Val Asp Val
        705                 710                 715
ATG GGC GGC CTG GAT GGC GAC ATG TTC AAC TAC TAT AAG ATG CTG ATG     2630
Met Gly Gly Leu Asp Gly Asp Met Phe Asn Tyr Tyr Lys Met Leu Met
    720                 725                 730
CTG CAA GGG CTG ATT GCC GCT CGG AAA CAC ATG GAC AAG GTG GTG CAG     2678
Leu Gln Gly Leu Ile Ala Ala Arg Lys His Met Asp Lys Val Val Gln
735                 740                 745                 750
ATC GTG GAG ATC ATG CAG CAA GGT TCT CAG CTT CCT TGC TTC CAT GGC     2726
Ile Val Glu Ile Met Gln Gln Gly Ser Gln Leu Pro Cys Phe His Gly
                755                 760                 765
TCC AGC ACC ATT CGA AAC CTC AAA GAG AGG TTC CAC ATG AGC ATG ACT     2774
Ser Ser Thr Ile Arg Asn Leu Lys Glu Arg Phe His Met Ser Met Thr
            770                 775                 780
GAG GAG CAG CTG CAG CTG CTG GTG GAG CAG ATG GTG GAT GGC AGT ATG     2822
Glu Glu Gln Leu Gln Leu Leu Val Glu Gln Met Val Asp Gly Ser Met
        785                 790                 795
CGG TCT ATC ACC ACC AAA CTC TAT GAC GGC TTC CAG TAC CTC ACC AAC     2870
Arg Ser Ile Thr Thr Lys Leu Tyr Asp Gly Phe Gln Tyr Leu Thr Asn
    800                 805                 810
```

GGC ATC ATG TGA CACGCTCCTC AGCCCAGGAG TGGTGGGGGG TCCAGGGCAC 2922

Gly Ile Met *

815

CCTCCCTAGA GGGCCCTTGT CTGAGAAACC CCAAACCAGG AAACCCCACC TACCCAACCA 2982

TCCACCCAAG GGAAATGGAA GGCAAGAAAC ACGAAGGATC ATGTGGTAAC TGCGAGAGCT 3042

TGCTGAGGGG TGGGAGAGCC AGCTGTGGGG TCCAGACTTG TTGGGGCTTC CCTGCCCCTC 3102

CTGGTCTGTG TCAGTATTAC CACCAGACTG ACTCCAGGAC TCACTGCCCT CCAGAAAACA 3162

GAGGTGACAA ATGTGAGGGA CACTGGGGCC TTTCTTCTCC TTGTAGGGGT CTCTCAGAGG 3222

TTCTTTCCAC AGGCCATCCT CTTATTCCGT TCTGGGGCCC AGGAAGTGGG GAAGAGTAGG 3282

TTCTCGGTAC TTAGGACTTG ATCCTGTGGT TGCCACTGGC CATGCTGCTG CCCAGCTCTA 3342

CCCCTCCCAG GGACCTACCC CTCCCAGGGA CCGACCCCTG GCCCAAGCTC CCCTTGCTGG 3402

CGGGCGCTGC GTGGGCCCTG CACTTGCTGA GGTTCCCCAT CATGGGCAAG GCAAGGGAAT 3462

TCCCACAGCC CTCCAGTGTA CTGAGGGTAC TGGCCTAGCC ATGTGGAATT CCCTACCCTG 3522

ACTCCTTCCC CAAACCCAGG GAAAAGAGCT CTCAATTTTT TATTTTTAAT TTTTGTTTGA 3582

AATAAAGTCC TTAGTTAGCC 3602

【０３４７】

配列番号：３１
配列の長さ：８２９
配列の型：アミノ酸
トポロジー：直線状
配列の種類：蛋白
配列：

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

180 185 190

Asn Met Tyr Ile His Met Asp Glu Asp Val Gly Asp Ala Ile Lys Pro

195 200 205

Tyr Ile Val His Arg Cys Arg Gln Ser Ile Asn Phe Ser Leu Gln Cys

210 215 220

Ala Leu Leu Leu Gly Ala Tyr Ser Ser Asp Met His Ile Ser Thr Gln

225 230 235 240

Arg His Ser Arg Gly Thr Lys Leu Arg Lys Leu Ile Leu Ser Asp Glu

245 250 255

Leu Lys Pro Ala His Arg Lys Arg Glu Leu Pro Ser Leu Ser Pro Ala

260 265 270

Pro Asp Thr Gly Leu Ser Pro Ser Lys Arg Thr His Gln Arg Ser Lys

275 280 285

Ser Asp Ala Thr Ala Ser Ile Ser Leu Ser Ser Asn Leu Lys Arg Thr

290 295 300

Ala Ser Asn Pro Lys Val Glu Asn Glu Asp Glu Glu Leu Ser Ser Ser

305 310 315 320

Thr Glu Ser Ile Asp Asn Ser Phe Ser Ser Pro Val Arg Leu Ala Pro

325 330 335

Glu Arg Glu Phe Ile Lys Ser Leu Met Ala Ile Gly Lys Arg Leu Ala

340 345 350

Thr Leu Pro Thr Lys Glu Gln Lys Thr Gln Arg Leu Ile Ser Glu Leu

355 360 365

Ser Leu Leu Asn His Lys Leu Pro Ala Arg Val Trp Leu Pro Thr Ala

370 375 380

Gly Phe Asp His His Val Val Arg Val Pro His Thr Gln Ala Val Val

385 390 395 400

Leu Asn Ser Lys Asp Lys Ala Pro Tyr Leu Ile Tyr Val Glu Val Leu

405 410 415

Glu Cys Glu Asn Phe Asp Thr Thr Ser Val Pro Ala Arg Ile Pro Glu
420 425 430
Asn Arg Ile Arg Ser Thr Arg Ser Val Glu Asn Leu Pro Glu Cys Gly
435 440 445
Ile Thr His Glu Gln Arg Ala Gly Ser Phe Ser Thr Val Pro Asn Tyr
450 455 460
Asp Asn Asp Asp Glu Ala Trp Ser Val Asp Asp Ile Gly Glu Leu Gln
465 470 475 480
Val Glu Leu Pro Glu Val His Thr Asn Ser Cys Asp Asn Ile Ser Gln
485 490 495
Phe Ser Val Asp Ser Ile Thr Ser Gln Glu Ser Lys Glu Pro Val Phe
500 505 510
Ile Ala Ala Gly Asp Ile Arg Arg Arg Leu Ser Glu Gln Leu Ala His
515 520 525
Thr Pro Thr Ala Phe Lys Arg Asp Pro Glu Asp Pro Ser Ala Val Ala
530 535 540
Leu Lys Glu Pro Trp Gln Glu Lys Val Arg Arg Ile Arg Glu Gly Ser
545 550 555 560
Pro Tyr Gly His Leu Pro Asn Trp Arg Leu Leu Ser Val Ile Val Lys
565 570 575
Cys Gly Asp Asp Leu Arg Gln Glu Leu Leu Ala Phe Gln Val Leu Lys
580 585 590
Gln Leu Gln Ser Ile Trp Glu Gln Glu Arg Val Pro Leu Trp Ile Lys
595 600 605
Pro Ile Gln Asp Ser Cys Glu Ile Thr Thr Asp Ser Gly Met Ile Glu
610 615 620
Pro Val Val Asn Ala Val Ser Ile His Gln Val Lys Lys Gln Ser Gln
625 630 635 640
Leu Ser Leu Leu Asp Tyr Phe Leu Gln Glu His Gly Ser Tyr Thr Thr

645 650 655

Glu Ala Phe Leu Ser Ala Gln Arg Asn Phe Val Gln Ser Cys Ala Gly

660 665 670

Tyr Cys Leu Val Cys Tyr Leu Leu Gln Val Lys Asp Arg His Asn Gly

675 680 685

Asn Ile Leu Leu Asp Ala Glu Gly His Ile Ile His Ile Asp Phe Gly

690 695 700

Phe Ile Leu Ser Ser Ser Pro Arg Asn Leu Gly Phe Glu Thr Ser Ala

705 710 715 720

Phe Lys Leu Thr Thr Glu Phe Val Asp Val Met Gly Gly Leu Asp Gly

725 730 735

Asp Met Phe Asn Tyr Tyr Lys Met Leu Met Leu Gln Gly Leu Ile Ala

740 745 750

Ala Arg Lys His Met Asp Lys Val Val Gln Ile Val Glu Ile Met Gln

755 760 765

Gln Gly Ser Gln Leu Pro Cys Phe His Gly Ser Ser Thr Ile Arg Asn

770 775 780

Leu Lys Glu Arg Phe His Met Ser Met Thr Glu Glu Gln Leu Gln Leu

785 790 795 800

Leu Val Glu Gln Met Val Asp Gly Ser Met Arg Ser Ile Thr Thr Lys

805 810 815

Leu Tyr Asp Gly Phe Gln Tyr Leu Thr Asn Gly Ile Met

820 825

【０３４８】

配列番号：３２

配列の長さ：２４８７

配列の型：核酸

鎖の数：一本鎖

トポロジー：直線状

配列の種類：ＤＮＡ(genomic)

配列：

ATGAGATTCT TGGAAGCTCG AAGTCTGGCT GTGGCCATGG GAGATACAGT AGTGGAGCCT 60

GCCCCCTTGA AGCCAACTTC TGAGCCCACT TCTGGCCCAC CAGGGAATAA TGGGGGGTCC 120

CTGCTAAGTG TCATCACGGA GGGGGTCGGG GAACTATCAG TGATTGACCC TGAGGTGGCC 180

CAGAAGGCCT GCCAGGAGGT GTTGGAGAAA GTCAAGCTTT TGCATGGAGG CGTGGCAGTC 240

TCTAGCAGAG GCACCCCACT GGAGTTGGTC AATGGGGATG GTGTGGACAG TGAGATCCGT 300

TGCCTAGATG ATCCACCTGC CCAGATCAGG GAGGAGGAAG ATGAGATGGG GGCCGCTGTG 360

GCCTCAGGCA CAGCCAAAGG AGCAAGAAGA CGGCGGCAGA ACAACTCAGC TAAACAGTCT 420

TGGCTGCTGA GGCTGTTTGA GTCAAAACTG TTTGACATCT CCATGGCCAT TTCATACCTG 480

TATAACTCCA AGGAGCCTGG AGTACAAGCC TACATTGGCA ACCGGCTCTT CTGCTTTCGC 540

AACGAGGACG TGGACTTCTA TCTGCCCCAG TTGCTTAACA TGTACATCCA CATGGATGAG 600

GACGTGGGTG ATGCCATTAA GCCCTACATA GTCCACCGTT GCCGCCAGAG CATTAACTTT 660

TCCCTCCAGT GTGCCCTGTT GCTTGGGGCC TATTCTTCAG ACATGCACAT TTCCACTCAA 720

CGACACTCCC GTGGGACCAA GCTACGGAAG CTGATCCTCT CAGATGAGCT AAAGCCAGCT 780

CACAGGAAGA GGGAGCTGCC CTCCTTGAGC CCGGCCCCTG ATACAGGGCT GTCTCCCTCC 840

AAAAGGACTC ACCAGCGCTC TAAGTCAGAT GCCACTGCCA GCATAAGTCT CAGCAGCAAC 900

CTGAAACGAA CAGCCAGCAA CCCTAAAGTG GAGAATGAGG ATGAGGAGCT CTCCTCCAGC 960

ACCGAGAGTA TTGATAATTC ATTCAGTTCC CCTGTTCGAC TGGCTCCTGA GAGAGAATTC 1020

ATCAAGTCCC TGATGGCGAT CGGCAAGCGG CTGGCCACGC TCCCCACCAA AGAGCAGAAA 1080

ACACAGAGGC TGATCTCAGA GCTCTCCCTG CTCAACCATA AGCTCCCTGC CCGAGTCTGG 1140

CTGCCCACTG CTGGCTTTGA CCACCACGTG GTCCGTGTAC CCCACACACA GGCTGTTGTC 1200

CTCAACTCCA AGGACAAGGC TCCCTACCTG ATTTATGTGG AAGTCCTTGA ATGTGAAAAC 1260

TTTGACACCA CCAGTGTCCC TGCCCGGATC CCCGAGAACC GAATTCGGAG TACGAGGTCC 1320

GTAGAAAACT TGCCCGAATG TGGTATTACC CATGAGCAGC GAGCTGGCAG CTTCAGCACT 1380
GTGCCCAACT ATGACAACGA TGATGAGGCC TGGTCGGTGG ATGACATAGG CGAGCTGCAA 1440
GTGGAGCTCC CCGAAGTGCA TACCAACAGC TGTGACAACA TCTCCCAGTT CTCTGTGGAC 1500
AGCATCACCA GCCAGGAGAG CAAGGAGCCT GTGTTCATTG CAGCAGGGGA CATCCGCCGG 1560
CGCCTTTCGG AACAGCTGGC TCATACCCCG ACAGCCTTCA AACGAGACCC AGAAGATCCT 1620
TCTGCAGTTG CTCTCAAAGA GCCCTGGCAG GAGAAAGTAC GGCGGATCAG AGAGGGCTCC 1680
CCCTACGGCC ATCTCCCCAA TTGGCGGCTC CTGTCAGTCA TTGTCAAGTG TGGGGATGAC 1740
CTTCGGCAAG AGCTTCTGGC CTTTCAGGTG TTGAAGCAAC TGCAGTCCAT TTGGGAACAG 1800
GAGCGAGTGC CCCTTTGGAT CAAGCCAATA CAAGATTCTT GTGAAATTAC GACTGATAGT 1860
GGCATGATTG AACCAGTGGT CAATGCTGTG TCCATCCATC AGGTGAAGAA ACAGTCACAG 1920
CTCTCCTTGC TCGATTACTT CCTACAGGAG CACGGCAGTT ACACCACTGA GGCATTCCTC 1980
AGTGCACAGC GCAATTTTGT GCAAAGTTGT GCTGGGTACT GCTTGGTCTG CTACCTGCTG 2040
CAAGTCAAGG ACAGACACAA TGGGAATATC CTTTTGGACG CAGAAGGCCA CATCATCCAC 2100
ATCGACTTTG GCTTCATCCT CTCCAGCTCA CCCCGAAATC TGGGCTTTGA GACGTCAGCC 2160
TTTAAGCTGA CCACAGAGTT TGTGGATGTG ATGGGCGGCC TGGATGGCGA CATGTTCAAC 2220
TACTATAAGA TGCTGATGCT GCAAGGGCTG ATTGCCGCTC GGAAACACAT GGACAAGGTG 2280
GTGCAGATCG TGGAGATCAT GCAGCAAGGT TCTCAGCTTC CTTGCTTCCA TGGCTCCAGC 2340
ACCATTCGAA ACCTCAAAGA GAGGTTCCAC ATGAGCATGA CTGAGGAGCA GCTGCAGCTG 2400
CTGGTGGAGC AGATGGTGGA TGGCAGTATG CGGTCTATCA CCACCAAACT CTATGACGGC 2460
TTCCAGTACC TCACCAACGG CATCATG 2487

【０３４９】

配列番号：３３
配列の長さ：３３２４
配列の型：核酸
鎖の数：一本鎖
トポロジー：直線状
配列の種類：ＤＮＡ(genomic)
起源：
　生物名：ヒト胎児脳ｃＤＮＡライブラリー
　株名：ＧＥＮ－４２８Ｂ１２ｃ１
配列の特徴：
　特徴を表わす記号：ＣＤＳ
　存在位置：１１５．．２６０１
　特徴を決定した方法：Ｅ
配列：

```
CCGGAATTCC GGGAAGGCCG GAGCAAGTTT TGAAGAAGTC CCTATCAGAT TACACTTGGT     60
TGACTACTCC GGAGCAGCCA CTAAGAGGGA TGAACAGGCC TGCGTGGAAA TTGA ATG      117
                                                          Met
                                                           1
AGA TTC TTG GAA GCT CGA AGT CTG GCT GTG GCC ATG GGA GAT ACA GTA      165
Arg Phe Leu Glu Ala Arg Ser Leu Ala Val Ala Met Gly Asp Thr Val
              5                  10                  15
GTG GAG CCT GCC CCC TTG AAG CCA ACT TCT GAG CCC ACT TCT GGC CCA      213
Val Glu Pro Ala Pro Leu Lys Pro Thr Ser Glu Pro Thr Ser Gly Pro
         20                  25                  30
CCA GGG AAT AAT GGG GGG TCC CTG CTA AGT GTC ATC ACG GAG GGG GTC      261
Pro Gly Asn Asn Gly Gly Ser Leu Leu Ser Val Ile Thr Glu Gly Val
     35                  40                  45
GGG GAA CTA TCA GTG ATT GAC CCT GAG GTG GCC CAG AAG GCC TGC CAG      309
Gly Glu Leu Ser Val Ile Asp Pro Glu Val Ala Gln Lys Ala Cys Gln
```

```
 50                  55                  60                  65
GAG GTG TTG GAG AAA GTC AAG CTT TTG CAT GGA GGC GTG GCA GTC TCT      357
Glu Val Leu Glu Lys Val Lys Leu Leu His Gly Gly Val Ala Val Ser
                 70                  75                  80
AGC AGA GGC ACC CCA CTG GAG TTG GTC AAT GGG GAT GGT GTG GAC AGT      405
Ser Arg Gly Thr Pro Leu Glu Leu Val Asn Gly Asp Gly Val Asp Ser
             85                  90                  95
GAG ATC CGT TGC CTA GAT GAT CCA CCT GCC CAG ATC AGG GAG GAG GAA      453
Glu Ile Arg Cys Leu Asp Asp Pro Pro Ala Gln Ile Arg Glu Glu Glu
        100                 105                 110
GAT GAG ATG GGG GCC GCT GTG GCC TCA GGC ACA GCC AAA GGA GCA AGA      501
Asp Glu Met Gly Ala Ala Val Ala Ser Gly Thr Ala Lys Gly Ala Arg
    115                 120                 125
AGA CGG CGG CAG AAC AAC TCA GCT AAA CAG TCT TGG CTG CTG AGG CTG      549
Arg Arg Arg Gln Asn Asn Ser Ala Lys Gln Ser Trp Leu Leu Arg Leu
130                 135                 140                 145
TTT GAG TCA AAA CTG TTT GAC ATC TCC ATG GCC ATT TCA TAC CTG TAT      597
Phe Glu Ser Lys Leu Phe Asp Ile Ser Met Ala Ile Ser Tyr Leu Tyr
                150                 155                 160
AAC TCC AAG GAG CCT GGA GTA CAA GCC TAC ATT GGC AAC CGG CTC TTC      645
Asn Ser Lys Glu Pro Gly Val Gln Ala Tyr Ile Gly Asn Arg Leu Phe
            165                 170                 175
TGC TTT CGC AAC GAG GAC GTG GAC TTC TAT CTG CCC CAG TTG CTT AAC      693
Cys Phe Arg Asn Glu Asp Val Asp Phe Tyr Leu Pro Gln Leu Leu Asn
        180                 185                 190
ATG TAC ATC CAC ATG GAT GAG GAC GTG GGT GAT GCC ATT AAG CCC TAC      741
Met Tyr Ile His Met Asp Glu Asp Val Gly Asp Ala Ile Lys Pro Tyr
    195                 200                 205
ATA GTC CAC CGT TGC CGC CAG AGC ATT AAC TTT TCC CTC CAG TGT GCC      789
```

```
Ile Val His Arg Cys Arg Gln Ser Ile Asn Phe Ser Leu Gln Cys Ala
210                 215                 220                 225
CTG TTG CTT GGG GCC TAT TCT TCA GAC ATG CAC ATT TCC ACT CAA CGA      837
Leu Leu Leu Gly Ala Tyr Ser Ser Asp Met His Ile Ser Thr Gln Arg
                230                 235                 240
CAC TCC CGT GGG ACC AAG CTA CGG AAG CTG ATC CTC TCA GAT GAG CTA      885
His Ser Arg Gly Thr Lys Leu Arg Lys Leu Ile Leu Ser Asp Glu Leu
            245                 250                 255
AAG CCA GCT CAC AGG AAG AGG GAG CTG CCC TCC TTG AGC CCG GCC CCT      933
Lys Pro Ala His Arg Lys Arg Glu Leu Pro Ser Leu Ser Pro Ala Pro
        260                 265                 270
GAT ACA GGG CTG TCT CCC TCC AAA AGG ACT CAC CAG CGC TCT AAG TCA      981
Asp Thr Gly Leu Ser Pro Ser Lys Arg Thr His Gln Arg Ser Lys Ser
    275                 280                 285
GAT GCC ACT GCC AGC ATA AGT CTC AGC AGC AAC CTG AAA CGA ACA GCC     1029
Asp Ala Thr Ala Ser Ile Ser Leu Ser Ser Asn Leu Lys Arg Thr Ala
290                 295                 300                 305
AGC AAC CCT AAA GTG GAG AAT GAG GAT GAG GAG CTC TCC TCC AGC ACC     1077
Ser Asn Pro Lys Val Glu Asn Glu Asp Glu Glu Leu Ser Ser Ser Thr
                310                 315                 320
GAG AGT ATT GAT AAT TCA TTC AGT TCC CCT GTT CGA CTG GCT CCT GAG     1125
Glu Ser Ile Asp Asn Ser Phe Ser Ser Pro Val Arg Leu Ala Pro Glu
            325                 330                 335
AGA GAA TTC ATC AAG TCC CTG ATG GCG ATC GGC AAG CGG CTG GCC ACG     1173
Arg Glu Phe Ile Lys Ser Leu Met Ala Ile Gly Lys Arg Leu Ala Thr
        340                 345                 350
CTC CCC ACC AAA GAG CAG AAA ACA CAG AGG CTG ATC TCA GAG CTC TCC     1221
Leu Pro Thr Lys Glu Gln Lys Thr Gln Arg Leu Ile Ser Glu Leu Ser
    355                 360                 365
```

```
CTG CTC AAC CAT AAG CTC CCT GCC CGA GTC TGG CTG CCC ACT GCT GGC     1269
Leu Leu Asn His Lys Leu Pro Ala Arg Val Trp Leu Pro Thr Ala Gly
370                 375                 380                 385
TTT GAC CAC CAC GTG GTC CGT GTA CCC CAC ACA CAG GCT GTT GTC CTC     1317
Phe Asp His His Val Val Arg Val Pro His Thr Gln Ala Val Val Leu
                390                 395                 400
AAC TCC AAG GAC AAG GCT CCC TAC CTG ATT TAT GTG GAA GTC CTT GAA     1365
Asn Ser Lys Asp Lys Ala Pro Tyr Leu Ile Tyr Val Glu Val Leu Glu
            405                 410                 415
TGT GAA AAC TTT GAC ACC ACC AGT GTC CCT GCC CGG ATC CCC GAG AAC     1413
Cys Glu Asn Phe Asp Thr Thr Ser Val Pro Ala Arg Ile Pro Glu Asn
        420                 425                 430
CGA ATT CGG AGT ACG AGG TCC GTA GAA AAC TTG CCC GAA TGT GGT ATT     1461
Arg Ile Arg Ser Thr Arg Ser Val Glu Asn Leu Pro Glu Cys Gly Ile
    435                 440                 445
ACC CAT GAG CAG CGA GCT GGC AGC TTC AGC ACT GTG CCC AAC TAT GAC     1509
Thr His Glu Gln Arg Ala Gly Ser Phe Ser Thr Val Pro Asn Tyr Asp
450                 455                 460                 465
AAC GAT GAT GAG GCC TGG TCG GTG GAT GAC ATA GGC GAG CTG CAA GTG     1557
Asn Asp Asp Glu Ala Trp Ser Val Asp Asp Ile Gly Glu Leu Gln Val
                470                 475                 480
GAG CTC CCC GAA GTG CAT ACC AAC AGC TGT GAC AAC ATC TCC CAG TTC     1605
Glu Leu Pro Glu Val His Thr Asn Ser Cys Asp Asn Ile Ser Gln Phe
            485                 490                 495
TCT GTG GAC AGC ATC ACC AGC CAG GAG AGC AAG GAG CCT GTG TTC ATT     1653
Ser Val Asp Ser Ile Thr Ser Gln Glu Ser Lys Glu Pro Val Phe Ile
        500                 505                 510
GCA GCA GGG GAC ATC CGC CGG CGC CTT TCG GAA CAG CTG GCT CAT ACC     1701
Ala Ala Gly Asp Ile Arg Arg Arg Leu Ser Glu Gln Leu Ala His Thr
```

```
    515                 520                 525
CCG ACA GCC TTC AAA CGA GAC CCA GAA GAT CCT TCT GCA GTT GCT CTC     1749
Pro Thr Ala Phe Lys Arg Asp Pro Glu Asp Pro Ser Ala Val Ala Leu
530                 535                 540                 545
AAA GAG CCC TGG CAG GAG AAA GTA CGG CGG ATC AGA GAG GGC TCC CCC     1797
Lys Glu Pro Trp Gln Glu Lys Val Arg Arg Ile Arg Glu Gly Ser Pro
                550                 555                 560
TAC GGC CAT CTC CCC AAT TGG CGG CTC CTG TCA GTC ATT GTC AAG TGT     1845
Tyr Gly His Leu Pro Asn Trp Arg Leu Leu Ser Val Ile Val Lys Cys
            565                 570                 575
GGG GAT GAC CTT CGG CAA GAG CTT CTG GCC TTT CAG GTG TTG AAG CAA     1893
Gly Asp Asp Leu Arg Gln Glu Leu Leu Ala Phe Gln Val Leu Lys Gln
        580                 585                 590
CTG CAG TCC ATT TGG GAA CAG GAG CGA GTG CCC CTT TGG ATC AAG CCA     1941
Leu Gln Ser Ile Trp Glu Gln Glu Arg Val Pro Leu Trp Ile Lys Pro
    595                 600                 605
ATA CAA GAT TCT TGT GAA ATT ACG ACT GAT AGT GGC ATG ATT GAA CCA     1989
Ile Gln Asp Ser Cys Glu Ile Thr Thr Asp Ser Gly Met Ile Glu Pro
610                 615                 620                 625
GTG GTC AAT GCT GTG TCC ATC CAT CAG GTG AAG AAA CAG TCA CAG CTC     2037
Val Val Asn Ala Val Ser Ile His Gln Val Lys Lys Gln Ser Gln Leu
                630                 635                 640
TCC TTG CTC GAT TAC TTC CTA CAG GAG CAC GGC AGT TAC ACC ACT GAG     2085
Ser Leu Leu Asp Tyr Phe Leu Gln Glu His Gly Ser Tyr Thr Thr Glu
            645                 650                 655
GCA TTC CTC AGT GCA CAG CGC AAT TTT GTG CAA AGT TGT GCT GGG TAC     2133
Ala Phe Leu Ser Ala Gln Arg Asn Phe Val Gln Ser Cys Ala Gly Tyr
        660                 665                 670
TGC TTG GTC TGC TAC CTG CTG CAA GTC AAG GAC AGA CAC AAT GGG AAT     2181
```

```
Cys Leu Val Cys Tyr Leu Leu Gln Val Lys Asp Arg His Asn Gly Asn
    675                 680                 685
ATC CTT TTG GAC GCA GAA GGC CAC ATC ATC CAC ATC GAC TTT GGC TTC     2229
Ile Leu Leu Asp Ala Glu Gly His Ile Ile His Ile Asp Phe Gly Phe
690                 695                 700                 705
ATC CTC TCC AGC TCA CCC CGA AAT CTG GGC TTT GAG ACG TCA GCC TTT     2277
Ile Leu Ser Ser Ser Pro Arg Asn Leu Gly Phe Glu Thr Ser Ala Phe
                710                 715                 720
AAG CTG ACC ACA GAG TTT GTG GAT GTG ATG GGC GGC CTG GAT GGC GAC     2325
Lys Leu Thr Thr Glu Phe Val Asp Val Met Gly Gly Leu Asp Gly Asp
            725                 730                 735
ATG TTC AAC TAC TAT AAG ATG CTG ATG CTG CAA GGG CTG ATT GCC GCT     2373
Met Phe Asn Tyr Tyr Lys Met Leu Met Leu Gln Gly Leu Ile Ala Ala
        740                 745                 750
CGG AAA CAC ATG GAC AAG GTG GTG CAG ATC GTG GAG ATC ATG CAG CAA     2421
Arg Lys His Met Asp Lys Val Val Gln Ile Val Glu Ile Met Gln Gln
    755                 760                 765
GGT TCT CAG CTT CCT TGC TTC CAT GGC TCC AGC ACC ATT CGA AAC CTC     2469
Gly Ser Gln Leu Pro Cys Phe His Gly Ser Ser Thr Ile Arg Asn Leu
770                 775                 780                 785
AAA GAG AGG TTC CAC ATG AGC ATG ACT GAG GAG CAG CTG CAG CTG CTG     2517
Lys Glu Arg Phe His Met Ser Met Thr Glu Glu Gln Leu Gln Leu Leu
                790                 795                 800
GTG GAG CAG ATG GTG GAT GGC AGT ATG CGG TCT ATC ACC ACC AAA CTC     2565
Val Glu Gln Met Val Asp Gly Ser Met Arg Ser Ile Thr Thr Lys Leu
            805                 810                 815
TAT GAC GGC TTC CAG TAC CTC ACC AAC GGC ATC ATG TGA CACGCTCCTC      2614
Tyr Asp Gly Phe Gln Tyr Leu Thr Asn Gly Ile Met  *
        820                 825                 830
```

AGCCCAGGAG TGGTGGGGGG TCCAGGGCAC CCTCCCTAGA GGGCCCTTGT CTGAGAAACC 2674
CCAAACCAGG AAACCCCACC TACCCAACCA TCCACCCAAG GGAAATGGAA GGCAAGAAAC 2734
ACGAAGGATC ATGTGGTAAC TGCGAGAGCT TGCTGAGGGG TGGGAGAGCC AGCTGTGGGG 2794
TCCAGACTTG TTGGGGCTTC CCTGCCCCTC CTGGTCTGTG TCAGTATTAC CACCAGACTG 2854
ACTCCAGGAC TCACTGCCCT CCAGAAAACA GAGGTGACAA ATGTGAGGGA CACTGGGGCC 2914
TTTCTTCTCC TTGTAGGGGT CTCTCAGAGG TTCTTTCCAC AGGCCATCCT CTTATTCCGT 2974
TCTGGGGCCC AGGAAGTGGG GAAGAGTAGG TTCTCGGTAC TTAGGACTTG ATCCTGTGGT 3034
TGCCACTGGC CATGCTGCTG CCCAGCTCTA CCCCTCCCAG GGACCTACCC CTCCCAGGGA 3094
CCGACCCCTG GCCCAAGCTC CCCTTGCTGG CGGGCGCTGC GTGGGCCCTG CACTTGCTGA 3154
GGTTCCCCAT CATGGGCAAG GCAAGGGAAT TCCCACAGCC CTCCAGTGTA CTGAGGGTAC 3214
TGGCCTAGCC ATGTGGAATT CCCTACCCTG ACTCCTTCCC CAAACCCAGG GAAAAGAGCT 3274
CTCAATTTTT TATTTTTAAT TTTTGTTTGA AATAAAGTCC TTAGTTAGCC 3324

【０３５０】

配列番号：３４
配列の長さ：８１０
配列の型：アミノ酸
トポロジー：直線状
配列の種類：蛋白
配列：

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

```
            180                 185                 190
Arg Asn Gln Lys His Gly Leu Phe Lys Gly Ile Ile Gln Asp Gly Lys
        195                 200                 205
Ile Ile Phe Met Pro Asn Gly Tyr Ile Thr Gln Cys Pro Asn Leu Asn
    210                 215                 220
His Thr Cys Pro Thr Cys Ser Asp Phe Leu Ser Leu Val Gln Gly Ile
225                 230                 235                 240
Met Asp Leu Gln Glu Leu Leu Ala Lys Met Thr Ala Lys Leu Asn Tyr
                245                 250                 255
Ala Glu Thr Arg Leu Ser Gln Leu Glu Asn Cys His Cys Glu Lys Thr
            260                 265                 270
Cys Gln Val Ser Gly Leu Leu Tyr Arg Asp Gln Asp Ser Trp Val Asp
        275                 280                 285
Gly Asp His Cys Arg Asn Cys Thr Cys Lys Ser Gly Ala Val Glu Cys
    290                 295                 300
Arg Arg Met Ser Cys Pro Pro Leu Asn Cys Ser Pro Asp Ser Leu Pro
305                 310                 315                 320
Val His Ile Ala Gly Gln Cys Cys Lys Val Cys Arg Pro Lys Cys Ile
                325                 330                 335
Tyr Gly Gly Lys Val Leu Ala Glu Gly Gln Arg Ile Leu Thr Lys Ser
            340                 345                 350
Cys Arg Glu Cys Arg Gly Gly Val Leu Val Lys Ile Thr Glu Met Cys
        355                 360                 365
Pro Pro Leu Asn Cys Ser Glu Lys Asp His Ile Leu Pro Glu Asn Gln
    370                 375                 380
Cys Cys Arg Val Cys Arg Gly His Asn Phe Cys Ala Glu Gly Pro Lys
385                 390                 395                 400
Cys Gly Glu Asn Ser Glu Cys Lys Asn Trp Asn Thr Lys Ala Thr Cys
                405                 410                 415
```

Glu Cys Lys Ser Gly Tyr Ile Ser Val Gln Gly Asp Ser Ala Tyr Cys
420 425 430
Glu Asp Ile Asp Glu Cys Ala Ala Lys Met His Tyr Cys His Ala Asn
435 440 445
Thr Val Cys Val Asn Leu Pro Gly Leu Tyr Arg Cys Asp Cys Val Pro
450 455 460
Gly Tyr Ile Arg Val Asp Asp Phe Ser Cys Thr Glu His Asp Glu Cys
465 470 475 480
Gly Ser Gly Gln His Asn Cys Asp Glu Asn Ala Ile Cys Thr Asn Thr
485 490 495
Val Gln Gly His Ser Cys Thr Cys Lys Pro Gly Tyr Val Gly Asn Gly
500 505 510
Thr Ile Cys Arg Ala Phe Cys Glu Glu Gly Cys Arg Tyr Gly Gly Thr
515 520 525
Cys Val Ala Pro Asn Lys Cys Val Cys Pro Ser Gly Phe Thr Gly Ser
530 535 540
His Cys Glu Lys Asp Ile Asp Glu Cys Ser Glu Gly Ile Ile Glu Cys
545 550 555 560
His Asn His Ser Arg Cys Val Asn Leu Pro Gly Trp Tyr His Cys Glu
565 570 575
Cys Arg Ser Gly Phe His Asp Asp Gly Thr Tyr Ser Leu Ser Gly Glu
580 585 590
Ser Cys Ile Asp Ile Asp Glu Cys Ala Leu Arg Thr His Thr Cys Trp
595 600 605
Asn Asp Ser Ala Cys Ile Asn Leu Ala Gly Gly Phe Asp Cys Leu Cys
610 615 620
Pro Ser Gly Pro Ser Cys Ser Gly Asp Cys Pro His Glu Gly Gly Leu
625 630 635 640
Lys His Asn Gly Gln Val Trp Thr Leu Lys Glu Asp Arg Cys Ser Val

645 650 655

Cys Ser Cys Lys Asp Gly Lys Ile Phe Cys Arg Arg Thr Ala Cys Asp

660 665 670

Cys Gln Asn Pro Ser Ala Asp Leu Phe Cys Cys Pro Glu Cys Asp Thr

675 680 685

Arg Val Thr Ser Gln Cys Leu Asp Gln Asn Gly His Lys Leu Tyr Arg

690 695 700

Ser Gly Asp Asn Trp Thr His Ser Cys Gln Gln Cys Arg Cys Leu Glu

705 710 715 720

Gly Glu Val Asp Cys Trp Pro Leu Thr Cys Pro Asn Leu Ser Cys Glu

725 730 735

Tyr Thr Ala Ile Leu Glu Gly Glu Cys Cys Pro Arg Cys Val Ser Asp

740 745 750

Pro Cys Leu Ala Asp Asn Ile Thr Tyr Asp Ile Arg Lys Thr Cys Leu

755 760 765

Asp Ser Tyr Gly Val Ser Arg Leu Ser Gly Ser Val Trp Thr Met Ala

770 775 780

Gly Ser Pro Cys Thr Thr Cys Lys Cys Lys Asn Gly Arg Val Cys Cys

785 790 795 800

Ser Val Asp Phe Glu Cys Leu Gln Asn Asn

805 810

【０３５１】

配列番号：３５

配列の長さ：２４３０

配列の型：核酸

鎖の数：一本鎖

トポロジー：直線状

配列の種類：ＤＮＡ(genomic)

配列：

ATGCCGATGG ATTTGATTTT AGTTGTGTGG TTCTGTGTGT GCACTGCCAG GACAGTGGTG 60

GGCTTTGGGA TGGACCCTGA CCTTCAGATG GATATCGTCA CCGAGCTTGA CCTTGTGAAC 120

ACCACCCTTG GAGTTGCTCA GGTGTCTGGA ATGCACAATG CCAGCAAAGC ATTTTTATTT 180

CAAGACATAG AAAGAGAGAT CCATGCAGCT CCTCATGTGA GTGAGAAATT AATTCAGCTG 240

TTCCAGAACA AGAGTGAATT CACCATTTTG GCCACTGTAC AGCAGAAGCC ATCCACTTCA 300

GGAGTGATAC TGTCCATTCG AGAACTGGAG CACAGCTATT TTGAACTGGA GAGCAGTGGC 360

CTGAGGGATG AGATTCGGTA TCACTACATA CACAATGGGA AGCCAAGGAC AGAGGCACTT 420

CCTTACCGCA TGGCAGATGG ACAATGGCAC AAGGTTGCAC TGTCAGTTAG CGCCTCTCAT 480

CTCCTGCTCC ATGTCGACTG TAACAGGATT TATGAGCGTG TGATAGACCC TCCAGATACC 540

AACCTTCCCC CAGGAATCAA TTTATGGCTT GGCCAGCGCA ACCAAAAGCA TGGCTTATTC 600

AAAGGGATCA TCCAAGATGG GAAGATCATC TTTATGCCGA ATGGATATAT AACACAGTGT 660

CCAAATCTAA ATCACACTTG CCCAACCTGC AGTGATTTCT TAAGCCTGGT GCAAGGAATA 720

ATGGATTTAC AAGAGCTTTT GGCCAAGATG ACTGCAAAAC TAAATTATGC AGAGACAAGA 780

CTTAGTCAAT TGGAAAACTG TCATTGTGAG AAGACTTGTC AAGTGAGTGG ACTGCTCTAT 840

CGAGATCAAG ACTCTTGGGT AGATGGTGAC CATTGCAGGA ACTGCACTTG CAAAAGTGGT 900

GCCGTGGAAT GCCGAAGGAT GTCCTGTCCC CCTCTCAATT GCTCCCCAGA CTCCCTCCCA 960

GTACACATTG CTGGCCAGTG CTGTAAGGTC TGCCGACCAA AATGTATCTA TGGAGGAAAA 1020

GTTCTTGCAG AAGGCCAGCG GATTTTAACC AAGAGCTGTC GGGAATGCCG AGGTGGAGTT 1080

TTAGTAAAAA TTACAGAAAT GTGTCCTCCT TTGAACTGCT CAGAAAAGGA TCACATTCTT 1140

CCTGAGAATC AGTGCTGCCG TGTCTGTAGA GGTCATAACT TTTGTGCAGA AGGACCTAAA 1200

TGTGGTGAAA ACTCAGAGTG CAAAAACTGG AATACAAAAG CTACTTGTGA GTGCAAGAGT 1260

GGTTACATCT CTGTCCAGGG AGACTCTGCC TACTGTGAAG ATATTGATGA GTGTGCAGCT 1320

AAGATGCATT ACTGTCATGC CAATACTGTG TGTGTCAACC TTCCTGGGTT ATATCGCTGT 1380
GACTGTGTCC CAGGATACAT TCGTGTGGAT GACTTCTCTT GTACAGAACA CGATGAATGT 1440
GGCAGCGGCC AGCACAACTG TGATGAGAAT GCCATCTGCA CCAACACTGT CCAGGGACAC 1500
AGCTGCACCT GCAAACCGGG CTACGTGGGG AACGGGACCA TCTGCAGAGC TTTCTGTGAA 1560
GAGGGCTGCA GATACGGTGG AACGTGTGTG GCTCCCAACA AATGTGTCTG TCCATCTGGA 1620
TTCACAGGAA GCCACTGCGA GAAAGATATT GATGAATGTT CAGAGGGAAT CATTGAGTGC 1680
CACAACCATT CCCGCTGCGT TAACCTGCCA GGGTGGTACC ACTGTGAGTG CAGAAGCGGT 1740
TTCCATGACG ATGGGACCTA TTCACTGTCC GGGGAGTCCT GTATTGACAT TGATGAATGT 1800
GCCTTAAGAA CTCACACCTG TTGGAACGAT TCTGCCTGCA TCAACCTGGC AGGGGGTTTT 1860
GACTGTCTCT GCCCCTCTGG GCCCTCCTGC TCTGGTGACT GTCCTCATGA AGGGGGGCTG 1920
AAGCACAATG GCCAGGTGTG GACCTTGAAA GAAGACAGGT GTTCTGTCTG CTCCTGCAAG 1980
GATGGCAAGA TATTCTGCCG ACGGACAGCT TGTGATTGCC AGAATCCAAG TGCTGACCTA 2040
TTCTGTTGCC CAGAATGTGA CACCAGAGTC ACAAGTCAAT GTTTAGACCA AAATGGTCAC 2100
AAGCTGTATC GAAGTGGAGA CAATTGGACC CATAGCTGTC AGCAGTGTCG GTGTCTGGAA 2160
GGAGAGGTAG ATTGCTGGCC ACTCACTTGC CCCAACTTGA GCTGTGAGTA TACAGCTATC 2220
TTAGAAGGGG AATGTTGTCC CCGCTGTGTC AGTGACCCCT GCCTAGCTGA TAACATCACC 2280
TATGACATCA GAAAAACTTG CCTGGACAGC TATGGTGTTT CACGGCTTAG TGGCTCAGTG 2340
TGGACGATGG CTGGATCTCC CTGCACAACC TGTAAATGCA AGAATGGAAG AGTCTGTTGT 2400
TCTGTGGATT TTGAGTGTCT TCAAAATAAT 2430

【０３５２】

配列番号：３６
配列の長さ：２９７７
配列の型：核酸
鎖の数：一本鎖
トポロジー：直線状
配列の種類：ＤＮＡ(genomic)
起源：
　生物名：ヒト胎児脳ｃＤＮＡライブラリー
　株名：ＧＥＮ－０７３Ｅ０７
配列の特徴：
　特徴を表わす記号：ＣＤＳ
　存在位置：１０３．．２５３２
　特徴を決定した方法：Ｅ
配列：

```
TAGCAAGTTT GGCGGCTCCA AGCCAGGCGC GCCTCAGGAT CCAGGCTCAT TTGCTTCCAC     60
CTAGCTTCGG TGCCCCCTGC TAGGCGGGGA CCCTCGAGAG CG ATG CCG ATG GAT       114
                                                 Met Pro Met Asp
                                                   1
TTG ATT TTA GTT GTG TGG TTC TGT GTG TGC ACT GCC AGG ACA GTG GTG       162
Leu Ile Leu Val Val Trp Phe Cys Val Cys Thr Ala Arg Thr Val Val
  5                  10                  15                  20
GGC TTT GGG ATG GAC CCT GAC CTT CAG ATG GAT ATC GTC ACC GAG CTT       210
Gly Phe Gly Met Asp Pro Asp Leu Gln Met Asp Ile Val Thr Glu Leu
                 25                  30                  35
GAC CTT GTG AAC ACC ACC CTT GGA GTT GCT CAG GTG TCT GGA ATG CAC       258
Asp Leu Val Asn Thr Thr Leu Gly Val Ala Gln Val Ser Gly Met His
             40                  45                  50
AAT GCC AGC AAA GCA TTT TTA TTT CAA GAC ATA GAA AGA GAG ATC CAT       306
Asn Ala Ser Lys Ala Phe Leu Phe Gln Asp Ile Glu Arg Glu Ile His
```

```
          55                  60                  65
GCA GCT CCT CAT GTG AGT GAG AAA TTA ATT CAG CTG TTC CAG AAC AAG      354
Ala Ala Pro His Val Ser Glu Lys Leu Ile Gln Leu Phe Gln Asn Lys
     70                  75                  80
AGT GAA TTC ACC ATT TTG GCC ACT GTA CAG CAG AAG CCA TCC ACT TCA      402
Ser Glu Phe Thr Ile Leu Ala Thr Val Gln Gln Lys Pro Ser Thr Ser
 85                  90                  95                  100
GGA GTG ATA CTG TCC ATT CGA GAA CTG GAG CAC AGC TAT TTT GAA CTG      450
Gly Val Ile Leu Ser Ile Arg Glu Leu Glu His Ser Tyr Phe Glu Leu
                105                 110                 115
GAG AGC AGT GGC CTG AGG GAT GAG ATT CGG TAT CAC TAC ATA CAC AAT      498
Glu Ser Ser Gly Leu Arg Asp Glu Ile Arg Tyr His Tyr Ile His Asn
            120                 125                 130
GGG AAG CCA AGG ACA GAG GCA CTT CCT TAC CGC ATG GCA GAT GGA CAA      546
Gly Lys Pro Arg Thr Glu Ala Leu Pro Tyr Arg Met Ala Asp Gly Gln
        135                 140                 145
TGG CAC AAG GTT GCA CTG TCA GTT AGC GCC TCT CAT CTC CTG CTC CAT      594
Trp His Lys Val Ala Leu Ser Val Ser Ala Ser His Leu Leu Leu His
    150                 155                 160
GTC GAC TGT AAC AGG ATT TAT GAG CGT GTG ATA GAC CCT CCA GAT ACC      642
Val Asp Cys Asn Arg Ile Tyr Glu Arg Val Ile Asp Pro Pro Asp Thr
165                 170                 175                 180
AAC CTT CCC CCA GGA ATC AAT TTA TGG CTT GGC CAG CGC AAC CAA AAG      690
Asn Leu Pro Pro Gly Ile Asn Leu Trp Leu Gly Gln Arg Asn Gln Lys
                185                 190                 195
CAT GGC TTA TTC AAA GGG ATC ATC CAA GAT GGG AAG ATC ATC TTT ATG      738
His Gly Leu Phe Lys Gly Ile Ile Gln Asp Gly Lys Ile Ile Phe Met
            200                 205                 210
CCG AAT GGA TAT ATA ACA CAG TGT CCA AAT CTA AAT CAC ACT TGC CCA      786
```

```
Pro Asn Gly Tyr Ile Thr Gln Cys Pro Asn Leu Asn His Thr Cys Pro
        215                 220                 225
ACC TGC AGT GAT TTC TTA AGC CTG GTG CAA GGA ATA ATG GAT TTA CAA      834
Thr Cys Ser Asp Phe Leu Ser Leu Val Gln Gly Ile Met Asp Leu Gln
    230                 235                 240
GAG CTT TTG GCC AAG ATG ACT GCA AAA CTA AAT TAT GCA GAG ACA AGA      882
Glu Leu Leu Ala Lys Met Thr Ala Lys Leu Asn Tyr Ala Glu Thr Arg
245                 250                 255                 260
CTT AGT CAA TTG GAA AAC TGT CAT TGT GAG AAG ACT TGT CAA GTG AGT      930
Leu Ser Gln Leu Glu Asn Cys His Cys Glu Lys Thr Cys Gln Val Ser
                265                 270                 275
GGA CTG CTC TAT CGA GAT CAA GAC TCT TGG GTA GAT GGT GAC CAT TGC      978
Gly Leu Leu Tyr Arg Asp Gln Asp Ser Trp Val Asp Gly Asp His Cys
            280                 285                 290
AGG AAC TGC ACT TGC AAA AGT GGT GCC GTG GAA TGC CGA AGG ATG TCC     1026
Arg Asn Cys Thr Cys Lys Ser Gly Ala Val Glu Cys Arg Arg Met Ser
        295                 300                 305
TGT CCC CCT CTC AAT TGC TCC CCA GAC TCC CTC CCA GTA CAC ATT GCT     1074
Cys Pro Pro Leu Asn Cys Ser Pro Asp Ser Leu Pro Val His Ile Ala
    310                 315                 320
GGC CAG TGC TGT AAG GTC TGC CGA CCA AAA TGT ATC TAT GGA GGA AAA     1122
Gly Gln Cys Cys Lys Val Cys Arg Pro Lys Cys Ile Tyr Gly Gly Lys
325                 330                 335                 340
GTT CTT GCA GAA GGC CAG CGG ATT TTA ACC AAG AGC TGT CGG GAA TGC     1170
Val Leu Ala Glu Gly Gln Arg Ile Leu Thr Lys Ser Cys Arg Glu Cys
                345                 350                 355
CGA GGT GGA GTT TTA GTA AAA ATT ACA GAA ATG TGT CCT CCT TTG AAC     1218
Arg Gly Gly Val Leu Val Lys Ile Thr Glu Met Cys Pro Pro Leu Asn
            360                 365                 370
```

```
TGC TCA GAA AAG GAT CAC ATT CTT CCT GAG AAT CAG TGC TGC CGT GTC     1266
Cys Ser Glu Lys Asp His Ile Leu Pro Glu Asn Gln Cys Cys Arg Val
        375                 380                 385
TGT AGA GGT CAT AAC TTT TGT GCA GAA GGA CCT AAA TGT GGT GAA AAC     1314
Cys Arg Gly His Asn Phe Cys Ala Glu Gly Pro Lys Cys Gly Glu Asn
    390                 395                 400
TCA GAG TGC AAA AAC TGG AAT ACA AAA GCT ACT TGT GAG TGC AAG AGT     1362
Ser Glu Cys Lys Asn Trp Asn Thr Lys Ala Thr Cys Glu Cys Lys Ser
405                 410                 415                 420
GGT TAC ATC TCT GTC CAG GGA GAC TCT GCC TAC TGT GAA GAT ATT GAT     1410
Gly Tyr Ile Ser Val Gln Gly Asp Ser Ala Tyr Cys Glu Asp Ile Asp
                425                 430                 435
GAG TGT GCA GCT AAG ATG CAT TAC TGT CAT GCC AAT ACT GTG TGT GTC     1458
Glu Cys Ala Ala Lys Met His Tyr Cys His Ala Asn Thr Val Cys Val
            440                 445                 450
AAC CTT CCT GGG TTA TAT CGC TGT GAC TGT GTC CCA GGA TAC ATT CGT     1506
Asn Leu Pro Gly Leu Tyr Arg Cys Asp Cys Val Pro Gly Tyr Ile Arg
        455                 460                 465
GTG GAT GAC TTC TCT TGT ACA GAA CAC GAT GAA TGT GGC AGC GGC CAG     1554
Val Asp Asp Phe Ser Cys Thr Glu His Asp Glu Cys Gly Ser Gly Gln
    470                 475                 480
CAC AAC TGT GAT GAG AAT GCC ATC TGC ACC AAC ACT GTC CAG GGA CAC     1602
His Asn Cys Asp Glu Asn Ala Ile Cys Thr Asn Thr Val Gln Gly His
485                 490                 495                 500
AGC TGC ACC TGC AAA CCG GGC TAC GTG GGG AAC GGG ACC ATC TGC AGA     1650
Ser Cys Thr Cys Lys Pro Gly Tyr Val Gly Asn Gly Thr Ile Cys Arg
                505                 510                 515
GCT TTC TGT GAA GAG GGC TGC AGA TAC GGT GGA ACG TGT GTG GCT CCC     1698
Ala Phe Cys Glu Glu Gly Cys Arg Tyr Gly Gly Thr Cys Val Ala Pro
```

```
                520                 525                 530
AAC AAA TGT GTC TGT CCA TCT GGA TTC ACA GGA AGC CAC TGC GAG AAA     1746
Asn Lys Cys Val Cys Pro Ser Gly Phe Thr Gly Ser His Cys Glu Lys
        535                 540                 545
GAT ATT GAT GAA TGT TCA GAG GGA ATC ATT GAG TGC CAC AAC CAT TCC     1794
Asp Ile Asp Glu Cys Ser Glu Gly Ile Ile Glu Cys His Asn His Ser
    550                 555                 560
CGC TGC GTT AAC CTG CCA GGG TGG TAC CAC TGT GAG TGC AGA AGC GGT     1842
Arg Cys Val Asn Leu Pro Gly Trp Tyr His Cys Glu Cys Arg Ser Gly
565                 570                 575                 580
TTC CAT GAC GAT GGG ACC TAT TCA CTG TCC GGG GAG TCC TGT ATT GAC     1890
Phe His Asp Asp Gly Thr Tyr Ser Leu Ser Gly Glu Ser Cys Ile Asp
                585                 590                 595
ATT GAT GAA TGT GCC TTA AGA ACT CAC ACC TGT TGG AAC GAT TCT GCC     1938
Ile Asp Glu Cys Ala Leu Arg Thr His Thr Cys Trp Asn Asp Ser Ala
            600                 605                 610
TGC ATC AAC CTG GCA GGG GGT TTT GAC TGT CTC TGC CCC TCT GGG CCC     1986
Cys Ile Asn Leu Ala Gly Gly Phe Asp Cys Leu Cys Pro Ser Gly Pro
        615                 620                 625
TCC TGC TCT GGT GAC TGT CCT CAT GAA GGG GGG CTG AAG CAC AAT GGC     2034
Ser Cys Ser Gly Asp Cys Pro His Glu Gly Gly Leu Lys His Asn Gly
    630                 635                 640
CAG GTG TGG ACC TTG AAA GAA GAC AGG TGT TCT GTC TGC TCC TGC AAG     2082
Gln Val Trp Thr Leu Lys Glu Asp Arg Cys Ser Val Cys Ser Cys Lys
645                 650                 655                 660
GAT GGC AAG ATA TTC TGC CGA CGG ACA GCT TGT GAT TGC CAG AAT CCA     2130
Asp Gly Lys Ile Phe Cys Arg Arg Thr Ala Cys Asp Cys Gln Asn Pro
                665                 670                 675
AGT GCT GAC CTA TTC TGT TGC CCA GAA TGT GAC ACC AGA GTC ACA AGT     2178
```

```
Ser Ala Asp Leu Phe Cys Cys Pro Glu Cys Asp Thr Arg Val Thr Ser
            680                 685                 690
CAA TGT TTA GAC CAA AAT GGT CAC AAG CTG TAT CGA AGT GGA GAC AAT     2226
Gln Cys Leu Asp Gln Asn Gly His Lys Leu Tyr Arg Ser Gly Asp Asn
        695                 700                 705
TGG ACC CAT AGC TGT CAG CAG TGT CGG TGT CTG GAA GGA GAG GTA GAT     2274
Trp Thr His Ser Cys Gln Gln Cys Arg Cys Leu Glu Gly Glu Val Asp
    710                 715                 720
TGC TGG CCA CTC ACT TGC CCC AAC TTG AGC TGT GAG TAT ACA GCT ATC     2322
Cys Trp Pro Leu Thr Cys Pro Asn Leu Ser Cys Glu Tyr Thr Ala Ile
725                 730                 735                 740
TTA GAA GGG GAA TGT TGT CCC CGC TGT GTC AGT GAC CCC TGC CTA GCT     2370
Leu Glu Gly Glu Cys Cys Pro Arg Cys Val Ser Asp Pro Cys Leu Ala
                745                 750                 755
GAT AAC ATC ACC TAT GAC ATC AGA AAA ACT TGC CTG GAC AGC TAT GGT     2418
Asp Asn Ile Thr Tyr Asp Ile Arg Lys Thr Cys Leu Asp Ser Tyr Gly
            760                 765                 770
GTT TCA CGG CTT AGT GGC TCA GTG TGG ACG ATG GCT GGA TCT CCC TGC     2466
Val Ser Arg Leu Ser Gly Ser Val Trp Thr Met Ala Gly Ser Pro Cys
        775                 780                 785
ACA ACC TGT AAA TGC AAG AAT GGA AGA GTC TGT TGT TCT GTG GAT TTT     2514
Thr Thr Cys Lys Cys Lys Asn Gly Arg Val Cys Cys Ser Val Asp Phe
    790                 795                 800
GAG TGT CTT CAA AAT AAT TGAAGTATTT ACAGTGGACT CAACGCAGAA            2562
Glu Cys Leu Gln Asn Asn
805                 810
GAATGGACGA AATGACCATC CAACGTGATT AAGGATAGGA ATCGGTAGTT TGGTTTTTTT   2622
GTTTGTTTTG TTTTTTTAAC CACAGATAAT TGCCAAAGTT TCCACCTGAG GACGGTGTTT   2682
CGGAGGTTGC CTTTTGGACC TACCACTTTG CTCATTCTTG CTAACCTAGT CTAGGTGACC   2742

TACAGTGCCG TGCATTTAAG TCAATGGTTG TTAAAAGAAG TTTCCCGTGT TGTAAATCAT   2802
GTTTCCCTTA TCAGATCATT TGCAAATACA TTTAAATGAT CTCATGGTAA ATGGTTGATG   2862
TATTTTTTGG GTTTATTTTG TGTACTAACC ATAATAGAGA GAGACTCAGC TCCTTTTATT   2922
TATTTTGTTG ATTTATGGAT CAAATTCTAA AATAAAGTTG CCTGTTGTGA CTTTT        2977
```

【０３５３】

配列番号：３７
配列の長さ：８１６
配列の型：アミノ酸
トポロジー：直線状
配列の種類：蛋白
配列：

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

```
            180                 185                 190
Gln Arg Asn Asn Ala His Gly Tyr Phe Lys Gly Ile Met Gln Asp Val
        195                 200                 205
Gln Leu Leu Val Met Pro Gln Gly Phe Ile Ala Gln Cys Pro Asp Leu
    210                 215                 220
Asn Arg Thr Cys Pro Thr Cys Asn Asp Phe His Gly Leu Val Gln Lys
225                 230                 235                 240
Ile Met Glu Leu Gln Asp Ile Leu Ala Lys Thr Ser Ala Lys Leu Ser
                245                 250                 255
Arg Ala Glu Gln Arg Met Asn Arg Leu Asp Gln Cys Tyr Cys Glu Arg
            260                 265                 270
Thr Cys Thr Met Lys Gly Thr Thr Tyr Arg Glu Phe Glu Ser Trp Ile
        275                 280                 285
Asp Gly Cys Lys Asn Cys Thr Cys Leu Asn Gly Thr Ile Gln Cys Glu
    290                 295                 300
Thr Leu Ile Cys Pro Asn Pro Asp Cys Pro Leu Lys Ser Ala Leu Ala
305                 310                 315                 320
Tyr Val Asp Gly Lys Cys Cys Lys Glu Cys Lys Ser Ile Cys Gln Phe
                325                 330                 335
Gln Gly Arg Thr Tyr Phe Glu Gly Glu Arg Asn Thr Val Tyr Ser Ser
            340                 345                 350
Ser Gly Val Cys Val Leu Tyr Glu Cys Lys Asp Gln Thr Met Lys Leu
        355                 360                 365
Val Glu Ser Ser Gly Cys Pro Ala Leu Asp Cys Pro Glu Ser His Gln
    370                 375                 380
Ile Thr Leu Ser His Ser Cys Cys Lys Val Cys Lys Gly Tyr Asp Phe
385                 390                 395                 400
Cys Ser Glu Arg His Asn Cys Met Glu Asn Ser Ile Cys Arg Asn Leu
                405                 410                 415
```

Asn Asp Arg Ala Val Cys Ser Cys Arg Asp Gly Phe Arg Ala Leu Arg
420 425 430
Glu Asp Asn Ala Tyr Cys Glu Asp Ile Asp Glu Cys Ala Glu Gly Arg
435 440 445
His Tyr Cys Arg Glu Asn Thr Met Cys Val Asn Thr Pro Gly Ser Phe
450 455 460
Met Cys Ile Cys Lys Thr Gly Tyr Ile Arg Ile Asp Asp Tyr Ser Cys
465 470 475 480
Thr Glu His Asp Glu Cys Ile Thr Asn Gln His Asn Cys Asp Glu Asn
485 490 495
Ala Leu Cys Phe Asn Thr Val Gly Gly His Asn Cys Val Cys Lys Pro
500 505 510
Gly Tyr Thr Gly Asn Gly Thr Thr Cys Lys Ala Phe Cys Lys Asp Gly
515 520 525
Cys Arg Asn Gly Gly Ala Cys Ile Ala Ala Asn Val Cys Ala Cys Pro
530 535 540
Gln Gly Phe Thr Gly Pro Ser Cys Glu Thr Asp Ile Asp Glu Cys Ser
545 550 555 560
Asp Gly Phe Val Gln Cys Asp Ser Arg Ala Asn Cys Ile Asn Leu Pro
565 570 575
Gly Trp Tyr His Cys Glu Cys Arg Asp Gly Tyr His Asp Asn Gly Met
580 585 590
Phe Ser Pro Ser Gly Glu Ser Cys Glu Asp Ile Asp Glu Cys Gly Thr
595 600 605
Gly Arg His Ser Cys Ala Asn Asp Thr Ile Cys Phe Asn Leu Asp Gly
610 615 620
Gly Tyr Asp Cys Arg Cys Pro His Gly Lys Asn Cys Thr Gly Asp Cys
625 630 635 640
Ile His Asp Gly Lys Val Lys His Asn Gly Gln Ile Trp Val Leu Glu

645 650 655

Asn Asp Arg Cys Ser Val Cys Ser Cys Gln Asn Gly Phe Val Met Cys

660 665 670

Arg Arg Met Val Cys Asp Cys Glu Asn Pro Thr Val Asp Leu Phe Cys

675 680 685

Cys Pro Glu Cys Asp Pro Arg Leu Ser Ser Gln Cys Leu His Gln Asn

690 695 700

Gly Glu Thr Leu Tyr Asn Ser Gly Asp Thr Trp Val Gln Asn Cys Gln

705 710 715 720

Gln Cys Arg Cys Leu Gln Gly Glu Val Asp Cys Trp Pro Leu Pro Cys

725 730 735

Pro Asp Val Glu Cys Glu Phe Ser Ile Leu Pro Glu Asn Glu Cys Cys

740 745 750

Pro Arg Cys Val Thr Asp Pro Cys Gln Ala Asp Thr Ile Arg Asn Asp

755 760 765

Ile Thr Lys Thr Cys Leu Asp Glu Met Asn Val Val Arg Phe Thr Gly

770 775 780

Ser Ser Trp Ile Lys His Gly Thr Glu Cys Thr Leu Cys Gln Cys Lys

785 790 795 800

Asn Gly His Ile Cys Cys Ser Val Asp Pro Gln Cys Leu Gln Glu Leu

805 810 815

【０３５４】

配列番号：３８

配列の長さ：２４４８

配列の型：核酸

鎖の数：一本鎖

トポロジー：直線状

配列の種類：ＤＮＡ(genomic)

配列：

ATGGAGTCTC GGGTCTTACT GAGAACATTC TGTTTGATCT TCGGTCTCGG AGCAGTTTGG 60

GGGCTTGGTG TGGACCCTTC CCTACAGATT GACGTCTTAA CAGAGTTAGA ACTTGGGGAG 120

TCCACGACCG GAGTGCGTCA GGTCCCGGGG CTGCATAATG GGACGAAAGC CTTTCTCTTT 180

CAAGATACTC CCAGAAGCAT AAAAGCATCC ACTGCTACAG CTGAACAGTT TTTTCAGAAG 240

CTGAGAAATA AACATGAATT TACTATTTTG GTGACCCTAA AACAGACCCA CTTAAATTCA 300

GGAGTTATTC TCTCAATTCA CCACTTGGAT CACAGGTACC TGGAACTGGA AAGTAGTGGC 360

CATCGGAATG AAGTCAGACT GCATTACCGC TCAGGCAGTC ACCGCCCTCA CACAGAAGTG 420

TTTCCTTACA TTTTGGCTGA TGACAAGTGG CACAAGCTCT CCTTAGCCAT CAGTGCTTCC 480

CATTTGATTT TACACATTGA CTGCAATAAA ATTTATGAAA GGGTAGTAGA AAAGCCCTCC 540

ACAGACTTGC CTCTAGGCAC AACATTTTGG CTAGGACAGA GAAATAATGC GCATGGATAT 600

TTTAAGGGTA TAATGCAAGA TGTCCAATTA CTTGTCATGC CCCAGGGATT TATTGCTCAG 660

TGCCCAGATC TTAATCGCAC CTGTCCAACT TGCAATGACT TCCATGGACT TGTGCAGAAA 720

ATCATGGAGC TACAGGATAT TTTAGCCAAA ACATCAGCCA AGCTGTCTCG AGCTGAACAG 780

CGAATGAATA GATTGGATCA GTGCTATTGT GAAAGGACTT GCACCATGAA GGGAACCACC 840

TACCGAGAAT TTGAGTCCTG GATAGACGGC TGTAAGAACT GCACATGCCT GAATGGAACC 900

ATCCAGTGTG AAACTCTAAT CTGCCCAAAT CCTGACTGCC CACTTAAGTC GGCTCTTGCG 960

TATGTGGATG GCAAATGCTG TAAGGAATGC AAATCGATAT GCCAATTTCA AGGACGAACC 1020

TACTTTGAAG GAGAAAGAAA TACAGTCTAT TCCTCTTCTG GAGTATGTGT TCTCTATGAG 1080

TGCAAGGACC AGACCATGAA ACTTGTTGAG AGTTCAGGCT GTCCAGCTTT GGATTGTCCA 1140

GAGTCTCATC AGATAACCTT GTCTCACAGC TGTTGCAAAG TTTGTAAAGG TTATGACTTT 1200

TGTTCTGAAA GGCATAACTG CATGGAGAAT TCCATCTGCA GAAATCTGAA TGACAGGGCT 1260

GTTTGTAGCT GTCGAGATGG TTTTAGGGCT CTTCGAGAGG ATAATGCCTA CTGTGAAGAC 1320

ATCGATGAGT GTGCTGAAGG GCGCCATTAC TGTCGTGAAA ATACAATGTG TGTCAACACC 1380
CCGGGTTCTT TTATGTGCAT CTGCAAAACT GGATACATCA GAATTGATGA TTATTCATGT 1440
ACAGAACATG ATGAGTGTAT CACAAATCAG CACAACTGTG ATGAAAATGC TTTATGCTTC 1500
AACACTGTTG GAGGACACAA CTGTGTTTGC AAGCCGGGCT ATACAGGGAA TGGAACGACA 1560
TGCAAAGCAT TTTGCAAAGA TGGCTGTAGG AATGGAGGAG CCTGTATTGC CGCTAATGTG 1620
TGTGCCTGCC CACAAGGCTT CACTGGACCC AGCTGTGAAA CGGACATTGA TGAATGCTCT 1680
GATGGTTTTG TTCAATGTGA CAGTCGTGCT AATTGCATTA ACCTGCCTGG ATGGTACCAC 1740
TGTGAGTGCA GAGATGGCTA CCATGACAAT GGGATGTTTT CACCAAGTGG AGAATCGTGT 1800
GAAGATATTG ATGAGTGTGG GACCGGGAGG CACAGCTGTG CCAATGATAC CATTTGCTTC 1860
AATTTGGATG GCGGATATGA TTGTCGATGT CCTCATGGAA AGAATTGCAC AGGGGACTGC 1920
ATCCATGATG GAAAAGTTAA GCACAATGGT CAGATTTGGG TGTTGGAAAA TGACAGGTGC 1980
TCTGTGTGCT CATGTCAGAA TGGATTCGTT ATGTGTCGAC GGATGGTCTG TGACTGTGAG 2040
AATCCCACAG TTGATCTTTT TTGCTGCCCT GAATGTGACC CAAGGCTTAG TAGTCAGTGC 2100
CTCCATCAAA ATGGGGAAAC TTTGTATAAC AGTGGTGACA CCTGGGTCCA GAATTGTCAA 2160
CAGTGCCGCT GCTTGCAAGG GGAAGTTGAT TGTTGGCCCC TGCCTTGCCC AGATGTGGAG 2220
TGTGAATTCA GCATTCTCCC AGAGAATGAG TGCTGCCCGC GCTGTGTCAC AGACCCTTGC 2280
CAGGCTGACA CCATCCGCAA TGACATCACC AAGACTTGCC TGGACGAAAT GAATGTGGTT 2340
CGCTTCACCG GGTCCTCTTG GATCAAACAT GGCACTGAGT GTACTCTCTG CCAGTGCAAG 2400
AATGGCCACA TCTGTTGCTC AGTGGATCCA CAGTGCCTTC AGGAACTG 2448

【０３５５】

配列番号：３９
配列の長さ：３１９８
配列の型：核酸
鎖の数：一本鎖
トポロジー：直線状
配列の種類：ＤＮＡ(genomic)
起源：
　生物名：ヒト胎児脳ｃＤＮＡライブラリー
　株名：ＧＥＮ－０９３Ｅ０５
配列の特徴：
　特徴を表わす記号：ＣＤＳ
　存在位置：９７．．２５４４
　特徴を決定した方法：Ｅ
配列：

```
TTGGGAGGAG CAGTCTCTCC GCTCGTCTCC CGGAGCTTTC TCCATTGTCT CTGCCTTTAC     60
AACAGAGGGA GACGATGGAC TGAGCTGATC CGCACC ATG GAG TCT CGG GTC TTA      114
                                        Met Glu Ser Arg Val Leu
                                          1               5
CTG AGA ACA TTC TGT TTG ATC TTC GGT CTC GGA GCA GTT TGG GGG CTT      162
Leu Arg Thr Phe Cys Leu Ile Phe Gly Leu Gly Ala Val Trp Gly Leu
             10                  15                  20
GGT GTG GAC CCT TCC CTA CAG ATT GAC GTC TTA ACA GAG TTA GAA CTT      210
Gly Val Asp Pro Ser Leu Gln Ile Asp Val Leu Thr Glu Leu Glu Leu
         25                  30                  35
GGG GAG TCC ACG ACC GGA GTG CGT CAG GTC CCG GGG CTG CAT AAT GGG      258
Gly Glu Ser Thr Thr Gly Val Arg Gln Val Pro Gly Leu His Asn Gly
     40                  45                  50
ACG AAA GCC TTT CTC TTT CAA GAT ACT CCC AGA AGC ATA AAA GCA TCC      306
Thr Lys Ala Phe Leu Phe Gln Asp Thr Pro Arg Ser Ile Lys Ala Ser
```

```
 55                  60                  65                  70
ACT GCT ACA GCT GAA CAG TTT TTT CAG AAG CTG AGA AAT AAA CAT GAA      354
Thr Ala Thr Ala Glu Gln Phe Phe Gln Lys Leu Arg Asn Lys His Glu
                 75                  80                  85
TTT ACT ATT TTG GTG ACC CTA AAA CAG ACC CAC TTA AAT TCA GGA GTT      402
Phe Thr Ile Leu Val Thr Leu Lys Gln Thr His Leu Asn Ser Gly Val
            90                  95                 100
ATT CTC TCA ATT CAC CAC TTG GAT CAC AGG TAC CTG GAA CTG GAA AGT      450
Ile Leu Ser Ile His His Leu Asp His Arg Tyr Leu Glu Leu Glu Ser
        105                 110                 115
AGT GGC CAT CGG AAT GAA GTC AGA CTG CAT TAC CGC TCA GGC AGT CAC      498
Ser Gly His Arg Asn Glu Val Arg Leu His Tyr Arg Ser Gly Ser His
    120                 125                 130
CGC CCT CAC ACA GAA GTG TTT CCT TAC ATT TTG GCT GAT GAC AAG TGG      546
Arg Pro His Thr Glu Val Phe Pro Tyr Ile Leu Ala Asp Asp Lys Trp
135                 140                 145                 150
CAC AAG CTC TCC TTA GCC ATC AGT GCT TCC CAT TTG ATT TTA CAC ATT      594
His Lys Leu Ser Leu Ala Ile Ser Ala Ser His Leu Ile Leu His Ile
                155                 160                 165
GAC TGC AAT AAA ATT TAT GAA AGG GTA GTA GAA AAG CCC TCC ACA GAC      642
Asp Cys Asn Lys Ile Tyr Glu Arg Val Val Glu Lys Pro Ser Thr Asp
            170                 175                 180
TTG CCT CTA GGC ACA ACA TTT TGG CTA GGA CAG AGA AAT AAT GCG CAT      690
Leu Pro Leu Gly Thr Thr Phe Trp Leu Gly Gln Arg Asn Asn Ala His
        185                 190                 195
GGA TAT TTT AAG GGT ATA ATG CAA GAT GTC CAA TTA CTT GTC ATG CCC      738
Gly Tyr Phe Lys Gly Ile Met Gln Asp Val Gln Leu Leu Val Met Pro
    200                 205                 210
CAG GGA TTT ATT GCT CAG TGC CCA GAT CTT AAT CGC ACC TGT CCA ACT      786
```

```
Gln Gly Phe Ile Ala Gln Cys Pro Asp Leu Asn Arg Thr Cys Pro Thr
215                 220                 225                 230
TGC AAT GAC TTC CAT GGA CTT GTG CAG AAA ATC ATG GAG CTA CAG GAT      834
Cys Asn Asp Phe His Gly Leu Val Gln Lys Ile Met Glu Leu Gln Asp
                235                 240                 245
ATT TTA GCC AAA ACA TCA GCC AAG CTG TCT CGA GCT GAA CAG CGA ATG      882
Ile Leu Ala Lys Thr Ser Ala Lys Leu Ser Arg Ala Glu Gln Arg Met
            250                 255                 260
AAT AGA TTG GAT CAG TGC TAT TGT GAA AGG ACT TGC ACC ATG AAG GGA      930
Asn Arg Leu Asp Gln Cys Tyr Cys Glu Arg Thr Cys Thr Met Lys Gly
        265                 270                 275
ACC ACC TAC CGA GAA TTT GAG TCC TGG ATA GAC GGC TGT AAG AAC TGC      978
Thr Thr Tyr Arg Glu Phe Glu Ser Trp Ile Asp Gly Cys Lys Asn Cys
    280                 285                 290
ACA TGC CTG AAT GGA ACC ATC CAG TGT GAA ACT CTA ATC TGC CCA AAT     1026
Thr Cys Leu Asn Gly Thr Ile Gln Cys Glu Thr Leu Ile Cys Pro Asn
295                 300                 305                 310
CCT GAC TGC CCA CTT AAG TCG GCT CTT GCG TAT GTG GAT GGC AAA TGC     1074
Pro Asp Cys Pro Leu Lys Ser Ala Leu Ala Tyr Val Asp Gly Lys Cys
                315                 320                 325
TGT AAG GAA TGC AAA TCG ATA TGC CAA TTT CAA GGA CGA ACC TAC TTT     1122
Cys Lys Glu Cys Lys Ser Ile Cys Gln Phe Gln Gly Arg Thr Tyr Phe
            330                 335                 340
GAA GGA GAA AGA AAT ACA GTC TAT TCC TCT TCT GGA GTA TGT GTT CTC     1170
Glu Gly Glu Arg Asn Thr Val Tyr Ser Ser Ser Gly Val Cys Val Leu
        345                 350                 355
TAT GAG TGC AAG GAC CAG ACC ATG AAA CTT GTT GAG AGT TCA GGC TGT     1218
Tyr Glu Cys Lys Asp Gln Thr Met Lys Leu Val Glu Ser Ser Gly Cys
    360                 365                 370
```

```
CCA GCT TTG GAT TGT CCA GAG TCT CAT CAG ATA ACC TTG TCT CAC AGC     1266
Pro Ala Leu Asp Cys Pro Glu Ser His Gln Ile Thr Leu Ser His Ser
375                 380                 385                 390
TGT TGC AAA GTT TGT AAA GGT TAT GAC TTT TGT TCT GAA AGG CAT AAC     1314
Cys Cys Lys Val Cys Lys Gly Tyr Asp Phe Cys Ser Glu Arg His Asn
                395                 400                 405
TGC ATG GAG AAT TCC ATC TGC AGA AAT CTG AAT GAC AGG GCT GTT TGT     1362
Cys Met Glu Asn Ser Ile Cys Arg Asn Leu Asn Asp Arg Ala Val Cys
            410                 415                 420
AGC TGT CGA GAT GGT TTT AGG GCT CTT CGA GAG GAT AAT GCC TAC TGT     1410
Ser Cys Arg Asp Gly Phe Arg Ala Leu Arg Glu Asp Asn Ala Tyr Cys
        425                 430                 435
GAA GAC ATC GAT GAG TGT GCT GAA GGG CGC CAT TAC TGT CGT GAA AAT     1458
Glu Asp Ile Asp Glu Cys Ala Glu Gly Arg His Tyr Cys Arg Glu Asn
    440                 445                 450
ACA ATG TGT GTC AAC ACC CCG GGT TCT TTT ATG TGC ATC TGC AAA ACT     1506
Thr Met Cys Val Asn Thr Pro Gly Ser Phe Met Cys Ile Cys Lys Thr
455                 460                 465                 470
GGA TAC ATC AGA ATT GAT GAT TAT TCA TGT ACA GAA CAT GAT GAG TGT     1554
Gly Tyr Ile Arg Ile Asp Asp Tyr Ser Cys Thr Glu His Asp Glu Cys
                475                 480                 485
ATC ACA AAT CAG CAC AAC TGT GAT GAA AAT GCT TTA TGC TTC AAC ACT     1602
Ile Thr Asn Gln His Asn Cys Asp Glu Asn Ala Leu Cys Phe Asn Thr
            490                 495                 500
GTT GGA GGA CAC AAC TGT GTT TGC AAG CCG GGC TAT ACA GGG AAT GGA     1650
Val Gly Gly His Asn Cys Val Cys Lys Pro Gly Tyr Thr Gly Asn Gly
        505                 510                 515
ACG ACA TGC AAA GCA TTT TGC AAA GAT GGC TGT AGG AAT GGA GGA GCC     1698
Thr Thr Cys Lys Ala Phe Cys Lys Asp Gly Cys Arg Asn Gly Gly Ala
```

```
    520                 525                 530
TGT ATT GCC GCT AAT GTG TGT GCC TGC CCA CAA GGC TTC ACT GGA CCC     1746
Cys Ile Ala Ala Asn Val Cys Ala Cys Pro Gln Gly Phe Thr Gly Pro
535                 540                 545                 550
AGC TGT GAA ACG GAC ATT GAT GAA TGC TCT GAT GGT TTT GTT CAA TGT     1794
Ser Cys Glu Thr Asp Ile Asp Glu Cys Ser Asp Gly Phe Val Gln Cys
                555                 560                 565
GAC AGT CGT GCT AAT TGC ATT AAC CTG CCT GGA TGG TAC CAC TGT GAG     1842
Asp Ser Arg Ala Asn Cys Ile Asn Leu Pro Gly Trp Tyr His Cys Glu
            570                 575                 580
TGC AGA GAT GGC TAC CAT GAC AAT GGG ATG TTT TCA CCA AGT GGA GAA     1890
Cys Arg Asp Gly Tyr His Asp Asn Gly Met Phe Ser Pro Ser Gly Glu
        585                 590                 595
TCG TGT GAA GAT ATT GAT GAG TGT GGG ACC GGG AGG CAC AGC TGT GCC     1938
Ser Cys Glu Asp Ile Asp Glu Cys Gly Thr Gly Arg His Ser Cys Ala
    600                 605                 610
AAT GAT ACC ATT TGC TTC AAT TTG GAT GGC GGA TAT GAT TGT CGA TGT     1986
Asn Asp Thr Ile Cys Phe Asn Leu Asp Gly Gly Tyr Asp Cys Arg Cys
615                 620                 625                 630
CCT CAT GGA AAG AAT TGC ACA GGG GAC TGC ATC CAT GAT GGA AAA GTT     2034
Pro His Gly Lys Asn Cys Thr Gly Asp Cys Ile His Asp Gly Lys Val
                635                 640                 645
AAG CAC AAT GGT CAG ATT TGG GTG TTG GAA AAT GAC AGG TGC TCT GTG     2082
Lys His Asn Gly Gln Ile Trp Val Leu Glu Asn Asp Arg Cys Ser Val
            650                 655                 660
TGC TCA TGT CAG AAT GGA TTC GTT ATG TGT CGA CGG ATG GTC TGT GAC     2130
Cys Ser Cys Gln Asn Gly Phe Val Met Cys Arg Arg Met Val Cys Asp
        665                 670                 675
TGT GAG AAT CCC ACA GTT GAT CTT TTT TGC TGC CCT GAA TGT GAC CCA     2178
```

```
Cys Glu Asn Pro Thr Val Asp Leu Phe Cys Cys Pro Glu Cys Asp Pro
    680                 685                 690
AGG CTT AGT AGT CAG TGC CTC CAT CAA AAT GGG GAA ACT TTG TAT AAC     2226
Arg Leu Ser Ser Gln Cys Leu His Gln Asn Gly Glu Thr Leu Tyr Asn
695                 700                 705                 710
AGT GGT GAC ACC TGG GTC CAG AAT TGT CAA CAG TGC CGC TGC TTG CAA     2274
Ser Gly Asp Thr Trp Val Gln Asn Cys Gln Gln Cys Arg Cys Leu Gln
                715                 720                 725
GGG GAA GTT GAT TGT TGG CCC CTG CCT TGC CCA GAT GTG GAG TGT GAA     2322
Gly Glu Val Asp Cys Trp Pro Leu Pro Cys Pro Asp Val Glu Cys Glu
            730                 735                 740
TTC AGC ATT CTC CCA GAG AAT GAG TGC TGC CCG CGC TGT GTC ACA GAC     2370
Phe Ser Ile Leu Pro Glu Asn Glu Cys Cys Pro Arg Cys Val Thr Asp
        745                 750                 755
CCT TGC CAG GCT GAC ACC ATC CGC AAT GAC ATC ACC AAG ACT TGC CTG     2418
Pro Cys Gln Ala Asp Thr Ile Arg Asn Asp Ile Thr Lys Thr Cys Leu
    760                 765                 770
GAC GAA ATG AAT GTG GTT CGC TTC ACC GGG TCC TCT TGG ATC AAA CAT     2466
Asp Glu Met Asn Val Val Arg Phe Thr Gly Ser Ser Trp Ile Lys His
775                 780                 785                 790
GGC ACT GAG TGT ACT CTC TGC CAG TGC AAG AAT GGC CAC ATC TGT TGC     2514
Gly Thr Glu Cys Thr Leu Cys Gln Cys Lys Asn Gly His Ile Cys Cys
                795                 800                 805
TCA GTG GAT CCA CAG TGC CTT CAG GAA CTG TGAAGTTAAC TGTCTCATGG       2564
Ser Val Asp Pro Gln Cys Leu Gln Glu Leu
            810                 815
GAGATTTCTG TTAAAAGAAT GTTCTTTCAT TAAAAGACCA AAAAGAAGTT AAAACTTAAA   2624
TTGGGTGATT TGTGGGCAGC TAAATGCAGC TTTGTTAATA GCTGAGTGAA CTTTCAATTA   2684
TGAAATTTGT GGAGCTTGAC AAAATCACAA AAGGAAAATT ACTGGGGCAA AATTAGACCT   2744
```

CAAGTCTGCC TCTACTGTGT CTCACATCAC CATGTAGAAG AATGGGCGTA CAGTATATAC 2804
CGTGACATCC TGAACCCTGG ATAGAAAGCC TGAGCCCATT GGATCTGTGA AAGCCTCTAG 2864
CTTCACTGGT GCAGAAAATT TTCCTCTAGA TCAGAATCTT CAGAATCAGT TAGGTTCCTC 2924
ACTGCAAGAA ATAAAATGTC AGGCAGTGAA TGAATTATAT TTTCAGAAGT AAAGCAAAGA 2984
AGCTATAACA TGTTATGTAC AGTACACTCT GAAAAGAAAT CTGAAACAAG TTATTGTAAT 3044
GATAAAAATA ATGCACAGGC ATGGTTACTT AATATTTTCT AACAGGAAAA GTCATCCCTA 3104
TTTCCTTGTT TTACTGCACT TAATATTATT TGGTTGAATT TGTTCAGTAT AAGCTCGTTC 3164
TTGTGCAAAA TTAAATAAAT ATTTCTCTTA CCTT 3198

【０３５６】

配列番号：４０
配列の長さ：４９９
配列の型：アミノ酸
トポロジー：直線状
配列の種類：蛋白
配列：

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

180 185 190

Gly Phe Glu Lys Gly Gly Gln Thr Arg Glu His Ala Met Phe Gly Lys

195 200 205

Thr Ala Gly Val Lys His Leu Ile Val Leu Ile Asn Lys Met Asp Asp

210 215 220

Pro Thr Val Asn Trp Gly Ile Glu Arg Tyr Glu Glu Cys Lys Glu Lys

225 230 235 240

Leu Val Pro Phe Leu Lys Lys Val Gly Phe Ser Pro Lys Lys Asp Ile

245 250 255

His Phe Met Pro Cys Ser Gly Leu Thr Gly Ala Asn Ile Lys Glu Gln

260 265 270

Ser Asp Phe Cys Pro Trp Tyr Thr Gly Leu Pro Phe Ile Pro Tyr Leu

275 280 285

Asn Asn Leu Pro Asn Phe Asn Arg Ser Ile Asp Gly Pro Ile Arg Leu

290 295 300

Pro Ile Val Asp Lys Tyr Lys Asp Met Gly Thr Val Val Leu Gly Lys

305 310 315 320

Leu Glu Ser Gly Ser Ile Phe Lys Gly Gln Gln Leu Val Met Met Pro

325 330 335

Asn Lys His Asn Val Glu Val Leu Gly Ile Leu Ser Asp Asp Thr Glu

340 345 350

Thr Asp Phe Val Ala Pro Gly Glu Asn Leu Lys Ile Arg Leu Lys Gly

355 360 365

Ile Glu Glu Glu Glu Ile Leu Pro Glu Phe Ile Leu Cys Asp Pro Ser

370 375 380

Asn Leu Cys His Ser Gly Arg Thr Phe Asp Val Gln Ile Val Ile Ile

385 390 395 400

Glu His Lys Ser Ile Ile Cys Pro Gly Tyr Asn Ala Val Leu His Ile

405 410 415

His Thr Cys Ile Glu Glu Val Glu Ile Thr Ala Leu Ile Ser Leu Val
420 425 430
Asp Lys Lys Ser Gly Glu Lys Ser Lys Thr Arg Pro Arg Phe Val Lys
435 440 445
Gln Asp Gln Val Cys Ile Ala Arg Leu Arg Thr Ala Gly Thr Ile Cys
450 455 460
Leu Glu Thr Phe Lys Asp Phe Pro Gln Met Gly Arg Phe Thr Leu Arg
465 470 475 480
Asp Glu Gly Lys Thr Ile Ala Ile Gly Lys Val Leu Lys Leu Val Pro
485 490 495
Glu Lys Asp

【０３５７】

配列番号：４１
配列の長さ：１４９７
配列の型：核酸
鎖の数：一本鎖
トポロジー：直線状
配列の種類：ＤＮＡ(genomic)
配列：

ATGGAACTTT CAGAACCTGT TGTAGAAAAT GGAGAGGTGG AAATGGCCCT AGAAGAATCA 60
TGGGAGCACA GTAAAGAAGT AAGTGAAGCC GAGCCTGGGG GTGGTTCCTC GGGAGATTCA 120
GGGCCCCCAG AAGAAAGTGG CCAGGAAATG ATGGAGGAAA AAGAGGAAAT AAGAAAATCC 180
AAATCTGTGA TCGTACCCTC AGGTGCACCT AAGAAAGAAC ACGTAAATGT AGTATTCATT 240
GGCCATGTAG ACGCTGGCAA GTCAACCATC GGAGGACAGA TAATGTTTTT GACTGGAATG 300
GCTGACAAAA GAACACTGGA GAAATATGAA AGAGAAGCTG AGGAAAAAAA CAGAGAAACC 360
TGGTATTTGT CCTGGGCCTT AGATACAAAT CAGGAGGAAC GAGACAAGGG TAAAACAGTC 420
GAAGTGGGTC GTGCCTATTT TGAAACAGAA AGGAAACATT TCACAATTTT AGATGCCCCT 480
GGCCACAAGA GTTTTGTCCC AAATATGATT GGTGGTGCTT CTCAAGCTGA TTTGGCTGTG 540
CTGGTCATCT CTGCCAGGAA AGGAGAGTTT GAAACTGGAT TTGAAAAAGG TGGACAGACA 600
AGAGAACATG CGATGTTTGG CAAAACGGCA GGAGTAAAAC ATTTAATAGT GCTTATTAAT 660
AAGATGGATG ATCCCACAGT AAATTGGGGC ATCGAGAGAT ATGAAGAATG TAAAGAAAAA 720
CTGGTGCCCT TTTTGAAAAA AGTAGGCTTT AGTCCAAAAA AGGACATTCA CTTTATGCCC 780
TGCTCAGGAC TGACCGGAGC AAATATTAAA GAGCAGTCAG ATTTCTGCCC TTGGTACACT 840
GGATTACCAT TTATTCCGTA TTTGAATAAC TTGCCAAACT TCAACAGATC AATTGATGGA 900
CCAATAAGAC TGCCAATTGT GGATAAGTAC AAAGATATGG GCACTGTGGT CCTGGGAAAG 960
CTGGAATCCG GGTCCATTTT TAAAGGCCAG CAGCTCGTGA TGATGCCAAA CAAGCACAAT 1020
GTAGAAGTTC TTGGAATACT TTCTGATGAT ACTGAAACTG ATTTTGTAGC CCCAGGTGAA 1080
AACCTCAAAA TCAGACTGAA GGGAATTGAA GAAGAAGAGA TTCTTCCAGA ATTCATACTT 1140
TGTGATCCTA GTAACCTCTG CCATTCTGGA CGCACGTTTG ATGTTCAGAT AGTGATTATT 1200
GAGCACAAAT CCATCATCTG CCCAGGTTAT AATGCGGTGC TGCACATTCA TACTTGTATT 1260
GAGGAAGTTG AGATAACAGC GTTAATCTCC TTGGTAGACA AAAAATCAGG GGAAAAAAGT 1320

AAGACACGAC CCCGCTTCGT GAAACAAGAT CAAGTATGCA TTGCTCGTTT AAGGACAGCA 1380
GGAACCATCT GCCTCGAGAC GTTCAAAGAT TTTCCTCAGA TGGGTCGTTT TACTTTAAGA 1440
GATGAGGGTA AGACCATTGC AATTGGAAAA GTTCTGAAAT TGGTCCCAGA GAAGGAC 1497

【０３５８】

配列番号：４２
配列の長さ：２０５７
配列の型：核酸
鎖の数：一本鎖
トポロジー：直線状
配列の種類：ＤＮＡ(genomic)
起源：
　生物名：ヒト胎児脳ｃＤＮＡライブラリー
　株名：ＧＥＮ－０７７Ａ０９
配列の特徴：
　特徴を表わす記号：ＣＤＳ
　存在位置：１４４．．１６４０
　特徴を決定した方法：Ｅ
配列：

```
TCCCGGCCGG CTCCGGCAGC AACGATGAAG CCTGCACCGG CGCGGGATAC CCTCAAGGTA     60
AAAGGATGGG ACGGGGGGCA CCTGTGGAAC CTTCCCGAGA GGAACCGTTA GTGTCGCTTG    120
AAGGTTCCAA TTCAGCCGTT ACC ATG GAA CTT TCA GAA CCT GTT GTA GAA         170
                          Met Glu Leu Ser Glu Pro Val Val Glu
                            1               5
AAT GGA GAG GTG GAA ATG GCC CTA GAA GAA TCA TGG GAG CAC AGT AAA       218
Asn Gly Glu Val Glu Met Ala Leu Glu Glu Ser Trp Glu His Ser Lys
 10                  15                  20                  25
GAA GTA AGT GAA GCC GAG CCT GGG GGT GGT TCC TCG GGA GAT TCA GGG       266
Glu Val Ser Glu Ala Glu Pro Gly Gly Gly Ser Ser Gly Asp Ser Gly
                 30                  35                  40
CCC CCA GAA GAA AGT GGC CAG GAA ATG ATG GAG GAA AAA GAG GAA ATA       314
Pro Pro Glu Glu Ser Gly Gln Glu Met Met Glu Glu Lys Glu Glu Ile
             45                  50                  55
AGA AAA TCC AAA TCT GTG ATC GTA CCC TCA GGT GCA CCT AAG AAA GAA       362
```

```
Arg Lys Ser Lys Ser Val Ile Val Pro Ser Gly Ala Pro Lys Lys Glu
         60                  65                  70
CAC GTA AAT GTA GTA TTC ATT GGC CAT GTA GAC GCT GGC AAG TCA ACC      410
His Val Asn Val Val Phe Ile Gly His Val Asp Ala Gly Lys Ser Thr
     75                  80                  85
ATC GGA GGA CAG ATA ATG TTT TTG ACT GGA ATG GCT GAC AAA AGA ACA      458
Ile Gly Gly Gln Ile Met Phe Leu Thr Gly Met Ala Asp Lys Arg Thr
 90                  95                 100                 105
CTG GAG AAA TAT GAA AGA GAA GCT GAG GAA AAA AAC AGA GAA ACC TGG      506
Leu Glu Lys Tyr Glu Arg Glu Ala Glu Glu Lys Asn Arg Glu Thr Trp
                110                 115                 120
TAT TTG TCC TGG GCC TTA GAT ACA AAT CAG GAG GAA CGA GAC AAG GGT      554
Tyr Leu Ser Trp Ala Leu Asp Thr Asn Gln Glu Glu Arg Asp Lys Gly
            125                 130                 135
AAA ACA GTC GAA GTG GGT CGT GCC TAT TTT GAA ACA GAA AGG AAA CAT      602
Lys Thr Val Glu Val Gly Arg Ala Tyr Phe Glu Thr Glu Arg Lys His
        140                 145                 150
TTC ACA ATT TTA GAT GCC CCT GGC CAC AAG AGT TTT GTC CCA AAT ATG      650
Phe Thr Ile Leu Asp Ala Pro Gly His Lys Ser Phe Val Pro Asn Met
    155                 160                 165
ATT GGT GGT GCT TCT CAA GCT GAT TTG GCT GTG CTG GTC ATC TCT GCC      698
Ile Gly Gly Ala Ser Gln Ala Asp Leu Ala Val Leu Val Ile Ser Ala
170                 175                 180                 185
AGG AAA GGA GAG TTT GAA ACT GGA TTT GAA AAA GGT GGA CAG ACA AGA      746
Arg Lys Gly Glu Phe Glu Thr Gly Phe Glu Lys Gly Gly Gln Thr Arg
                190                 195                 200
GAA CAT GCG ATG TTT GGC AAA ACG GCA GGA GTA AAA CAT TTA ATA GTG      794
Glu His Ala Met Phe Gly Lys Thr Ala Gly Val Lys His Leu Ile Val
            205                 210                 215
```

```
CTT ATT AAT AAG ATG GAT GAT CCC ACA GTA AAT TGG GGC ATC GAG AGA      842
Leu Ile Asn Lys Met Asp Asp Pro Thr Val Asn Trp Gly Ile Glu Arg
        220                 225                 230
TAT GAA GAA TGT AAA GAA AAA CTG GTG CCC TTT TTG AAA AAA GTA GGC      890
Tyr Glu Glu Cys Lys Glu Lys Leu Val Pro Phe Leu Lys Lys Val Gly
    235                 240                 245
TTT AGT CCA AAA AAG GAC ATT CAC TTT ATG CCC TGC TCA GGA CTG ACC      938
Phe Ser Pro Lys Lys Asp Ile His Phe Met Pro Cys Ser Gly Leu Thr
250                 255                 260                 265
GGA GCA AAT ATT AAA GAG CAG TCA GAT TTC TGC CCT TGG TAC ACT GGA      986
Gly Ala Asn Ile Lys Glu Gln Ser Asp Phe Cys Pro Trp Tyr Thr Gly
                270                 275                 280
TTA CCA TTT ATT CCG TAT TTG AAT AAC TTG CCA AAC TTC AAC AGA TCA     1034
Leu Pro Phe Ile Pro Tyr Leu Asn Asn Leu Pro Asn Phe Asn Arg Ser
            285                 290                 295
ATT GAT GGA CCA ATA AGA CTG CCA ATT GTG GAT AAG TAC AAA GAT ATG     1082
Ile Asp Gly Pro Ile Arg Leu Pro Ile Val Asp Lys Tyr Lys Asp Met
        300                 305                 310
GGC ACT GTG GTC CTG GGA AAG CTG GAA TCC GGG TCC ATT TTT AAA GGC     1130
Gly Thr Val Val Leu Gly Lys Leu Glu Ser Gly Ser Ile Phe Lys Gly
    315                 320                 325
CAG CAG CTC GTG ATG ATG CCA AAC AAG CAC AAT GTA GAA GTT CTT GGA     1178
Gln Gln Leu Val Met Met Pro Asn Lys His Asn Val Glu Val Leu Gly
330                 335                 340                 345
ATA CTT TCT GAT GAT ACT GAA ACT GAT TTT GTA GCC CCA GGT GAA AAC     1226
Ile Leu Ser Asp Asp Thr Glu Thr Asp Phe Val Ala Pro Gly Glu Asn
                350                 355                 360
CTC AAA ATC AGA CTG AAG GGA ATT GAA GAA GAA GAG ATT CTT CCA GAA     1274
Leu Lys Ile Arg Leu Lys Gly Ile Glu Glu Glu Glu Ile Leu Pro Glu
```

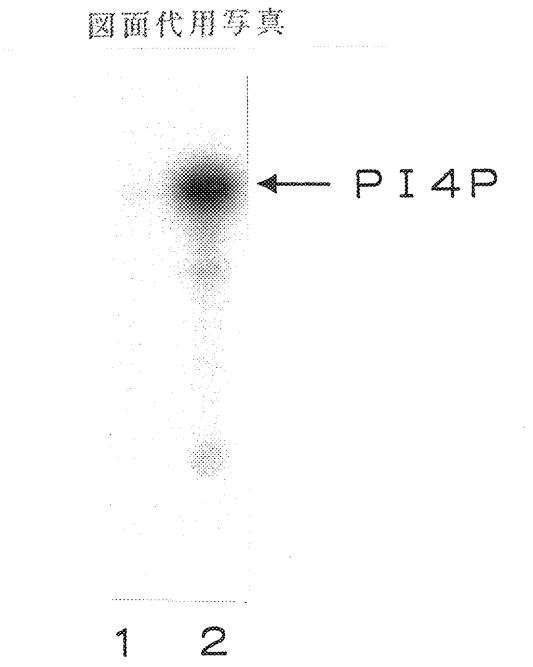
```
            365                 370                 375
TTC ATA CTT TGT GAT CCT AGT AAC CTC TGC CAT TCT GGA CGC ACG TTT     1322
Phe Ile Leu Cys Asp Pro Ser Asn Leu Cys His Ser Gly Arg Thr Phe
        380                 385                 390
GAT GTT CAG ATA GTG ATT ATT GAG CAC AAA TCC ATC ATC TGC CCA GGT     1370
Asp Val Gln Ile Val Ile Ile Glu His Lys Ser Ile Ile Cys Pro Gly
    395                 400                 405
TAT AAT GCG GTG CTG CAC ATT CAT ACT TGT ATT GAG GAA GTT GAG ATA     1418
Tyr Asn Ala Val Leu His Ile His Thr Cys Ile Glu Glu Val Glu Ile
410                 415                 420                 425
ACA GCG TTA ATC TCC TTG GTA GAC AAA AAA TCA GGG GAA AAA AGT AAG     1466
Thr Ala Leu Ile Ser Leu Val Asp Lys Lys Ser Gly Glu Lys Ser Lys
                430                 435                 440
ACA CGA CCC CGC TTC GTG AAA CAA GAT CAA GTA TGC ATT GCT CGT TTA     1514
Thr Arg Pro Arg Phe Val Lys Gln Asp Gln Val Cys Ile Ala Arg Leu
            445                 450                 455
AGG ACA GCA GGA ACC ATC TGC CTC GAG ACG TTC AAA GAT TTT CCT CAG     1562
Arg Thr Ala Gly Thr Ile Cys Leu Glu Thr Phe Lys Asp Phe Pro Gln
        460                 465                 470
ATG GGT CGT TTT ACT TTA AGA GAT GAG GGT AAG ACC ATT GCA ATT GGA     1610
Met Gly Arg Phe Thr Leu Arg Asp Glu Gly Lys Thr Ile Ala Ile Gly
    475                 480                 485
AAA GTT CTG AAA TTG GTC CCA GAG AAG GAC TAAGCAATTT TCTTGATGCC       1660
Lys Val Leu Lys Leu Val Pro Glu Lys Asp
490                 495
TCTGCAAGAT ACTGTGAGGA GAATTGACAG CAAAAGTTCA CCACCTACTC TTATTTACTG   1720
CCCATTGATT GACTTTTCTT CATATTTTGC AAAGAGAAAT TTCACAGCAA AAATTCATGT   1780
TTTGTCAGCT TTCTCATGTT GAGATCTGTT ATGTCACTGA TGAATTTACC CTCAAGTTTC   1840
CTTCCTCTGT ACCACTCTGC TTCCTTGGAC AATATCAGTA ATAGCTTTGT AAGTGATGTG   1900

GACGTAATTG CCTACAGTAA TAAAAAAATA ATGTACTTTA ATTTTTCATT TTCTTTTAGG   1960
ATATTTAGAC CACCCTTGTT CCACGCAAAC CAGAGTGTGT CAGTGTTTGT GTGTGTGTTA   2020
AAATGATAAC TAACATGTGA ATAAAATACT CCATTTG                            2057
```

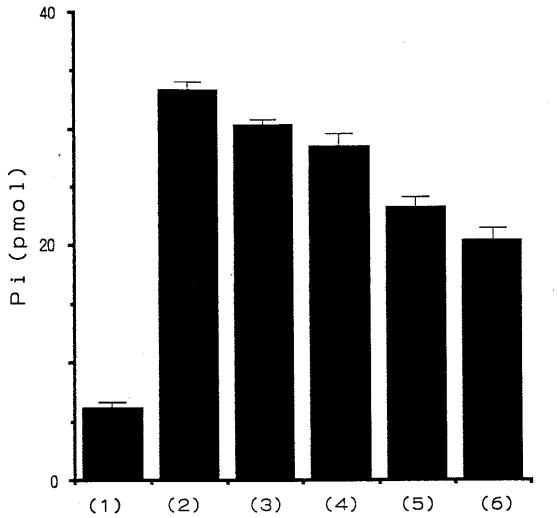
【図面の簡単な説明】

【図１】実施例９の（６）に従う蛋白発現試験における放射活性の定量結果を解析した図面に代わる写真である。
【図２】実施例９の（６）に従う蛋白発現試験において、トリトン－Ｘ１００、アデノシンの添加が、発現蛋白の活性に及ぼす影響を調べたグラフである。

【図１】



【図２】



フロントページの続き

(51)Int.Cl.[7]	FI	
Ｃ１２Ｑ　1/68	Ｃ１２Ｑ　1/68	Ａ
(Ｃ１２Ｐ　21/02	Ｃ１２Ｐ　21/02	
Ｃ１２Ｒ　1:19　)	Ｃ１２Ｒ　1:19	

(72)発明者　渡辺　武
徳島県板野郡藍住町富吉字富吉　リッチ・デ・藍住８０１号
(72)発明者　堀江　正人
徳島県徳島市南田宮４丁目４－４４　ラベンダーハイツ３０８
(72)発明者　片桐　豊雅
東京都豊島区西巣鴨１－２－１９－４０１

審査官　田村　明照

(56)参考文献　Developmental Dynamics, Vol.203, pp.212-222 (1995)

(58)調査した分野(Int.Cl.[7], ＤＢ名)
C12N 15/00
SwissProt/PIR/GeneSeq