

(19)日本国特許庁 (J P)

(12) **公表特許公報** (A) (11)特許出願公表番号

特表2003 - 508074

(P2003 - 508074A)

(43)公表日 平成15年3月4日(2003.3.4)

(51) Int.Cl ⁷	識別記号	F I	テ-マコード* (参考)
C 1 2 N 15/09	ZNA	A 0 1 K 67/027	2 G 0 4 5
A 0 1 K 67/027		C 0 7 K 14/47	4 B 0 2 4
C 0 7 K 14/47		16/18	4 B 0 6 3
16/18		C 1 2 N 1/15	4 B 0 6 4
C 1 2 N 1/15		1/19	4 B 0 6 5

審査請求 未請求 予備審査請求 (全129数) 最終頁に続く

(21)出願番号 特願2001 - 521734(P2001 - 521734)

(86)(22)出願日 平成12年9月4日(2000.9.4)

(85)翻訳文提出日 平成13年5月7日(2001.5.7)

(86)国際出願番号 PCT/FR00/02433

(87)国際公開番号 W001/018198

(87)国際公開日 平成13年3月15日(2001.3.15)

(31)優先権主張番号 99/11097

(32)優先日 平成11年9月3日(1999.9.3)

(33)優先権主張国 フランス(FR)

(71)出願人 サントル、ナショナル、ド、ラ、ルシエルシュ、シアンティフィク、(セ-エヌエルエス)

フランス国パリ、リュ、ミケ、ランジュ、3

(72)発明者 ジャン、バイセンバッハ

フランス国パリ、リュ、ド、ボジラール、163

(72)発明者 ジャミレ、アザン

フランス国パリ、リュ、デ、コルドリエール、22

(74)代理人 弁理士 吉武 賢次 (外4名)

最終頁に続く

(54)【発明の名称】 最も一般的な形態の常染色体性優性遺伝性痙性対麻痺に応答し得る S P G 遺伝子のクローニング、発現および同定

(57)【要約】

本発明はスパスチンをコードする S P G 遺伝子、および最も頻繁な形態の常染色体性優性家族性痙性対麻痺の原因であるそのいくつかの突然変異の同定および特性決定、その c D N A および対応するポリペプチドのクローニングおよび特性決定に関する。本発明はまた、ベクター、形質転換細胞、およびトランスジェニック動物、ならびに診断方法およびキット、および該ポリペプチドと直接的または間接的に相互作用し得る化学化合物または生化学化合物を選択する方法に関する。

【特許請求の範囲】**【請求項1】**

a) 配列番号1、配列番号2、配列番号72、配列番号106の配列、またはこれらの配列の1つの少なくとも15個の連続するヌクレオチドの配列、

b) 配列番号1、配列番号2、配列番号72、もしくは配列番号106の配列の相同体または変異体である核酸配列、および

c) a) および b) で定義される配列に対応する相補配列またはRNA配列、からなる群から選択される配列を含んでなることを特徴とする、精製または単離されたSPG4遺伝子の核酸。

【請求項2】

受託番号AB029006としてGenBankデータベースで確認される核酸を除く、請求項1に記載の精製または単離された核酸。

【請求項3】

配列番号2の配列、それに対して相補的な配列、またはその対応するRNAの配列についての、nt714~809(末端を含む)断片の少なくとも15個の連続するヌクレオチドの配列を少なくとも1つ含んでなる、請求項1または2に記載の精製または単離された核酸。

【請求項4】

ヒトの本来の多形性に対応する突然変異を含んでなる、請求項1~3のいずれか一項に記載の精製または単離された核酸。

【請求項5】

請求項1~4のいずれか一項に記載の核酸配列を含んでなる、プローブまたはプライマー。

【請求項6】

該配列が配列番号4~配列番号71の配列から選択される、請求項5に記載のプローブまたはプライマー。

【請求項7】

配列番号74~配列番号105の配列から選択される請求項1に記載の核酸配列を含んでなる、スプライス受容部位または供与部位。

【請求項 8】

cDNAまたはゲノムDNAライブラリーをスクリーニングするか、またはスパスチンをコードしている単離されたゲノムまたはcDNAをクローニングする方法であって、請求項1～7のいずれか一項に記載の核酸配列を用いる方法。

【請求項 9】

哺乳類、特にマウスのSPG4遺伝子のゲノムまたはcDNA配列を同定する、請求項8に記載の方法。

【請求項 10】

ヒトSPG4遺伝子が有する突然変異を同定する方法であって、請求項1～7のいずれか一項に記載の核酸配列を用いる方法。

【請求項 11】

常染色体性優性遺伝性痙性対麻痺の原因である突然変異を同定する、請求項10に記載の方法。

【請求項 12】

SPG4遺伝子の発現を促進および/または調節する核酸配列を同定する方法であって、請求項1～7のいずれか一項に記載の核酸配列を用いる方法。

【請求項 13】

請求項9～12のいずれか一項に記載の方法を用いて同定された核酸。

【請求項 14】

請求項1～4および13のいずれか一項に記載の核酸によりコードされているポリペプチド。

【請求項 15】

好ましくはその配列が受託番号AB029006としてGenBankデータベースで確認される584アミノ酸ペプチドを除く、請求項14に記載のポリペプチド。

【請求項 16】

a) 配列番号3、配列番号73、配列番号107の配列、またはこれらの配列の1つの少なくとも10個の連続するアミノ酸の配列、および

b) 配列番号3、配列番号73、または配列番号107の配列の相同体または

変異体である配列、
からなる群から選択されるアミノ酸配列を含んでなる、請求項14または15に
記載のポリペプチド。

【請求項17】

配列番号3の配列の a a 1 9 7 ~ 2 2 8 (末端を含む)断片の少なくとも8個
の連続するアミノ酸の配列を含んでなる、請求項14または15に記載のポリペ
プチド。

【請求項18】

ヒトの本来の多形性に対応する少なくとも1つの突然変異を有する、配列番号
3、配列番号73、配列番号107の配列、および少なくとも10個の連続する
アミノ酸のその断片の配列からなる群から選択されるアミノ酸配列を含んでなる
、請求項14または15に記載のポリペプチド。

【請求項19】

請求項1~4および13のいずれか一項に記載の核酸配列を含む、クローニン
グおよび/または発現ベクター。

【請求項20】

宿主細胞におけるその発現に必要なエレメントを含む、請求項19に記載のベ
クター。

【請求項21】

請求項19または20に記載のベクターで形質転換された、宿主細胞。

【請求項22】

請求項21に記載の細胞を含んでなる、ヒトを除く哺乳類。

【請求項23】

形質転換細胞を含んでなり、SPG4遺伝子の2つの対立遺伝子の少なくとも
一方の配列がヒトの本来の多形性に対応するか、または請求項10または11に
記載の方法を用いて同定される少なくとも1つの突然変異を含む、請求項22に
記載のヒトを除く哺乳類。

【請求項24】

核酸配列を検出および/または増幅するためのプローブまたはプライマーとし

での、請求項5、6および13のいずれか一項に記載の核酸配列の使用。

【請求項25】

ゲノムまたはcDNAライブラリーをスクリーニングするための、請求項1～7および13のいずれか一項に記載の核酸配列の使用。

【請求項26】

組換えまたは合成ポリペプチドを生産するための、請求項1～4および13のいずれか一項に記載の核酸配列の使用。

【請求項27】

組換えポリペプチドを発現させる条件下で請求項21に記載の形質転換細胞を培養してその組換えポリペプチドを回収する、組換えポリペプチドの生産方法。

【請求項28】

請求項27に記載の方法を用いて得られるポリペプチド。

【請求項29】

請求項14～18および28のいずれか一項に記載のポリペプチドを特異的に認識し得る、モノもしくはポリクローナル抗体、またはそれらの断片、キメラ抗体もしくは免疫複合体。

【請求項30】

請求項14～18および28のいずれか一項に記載のポリペプチドを検出および/または精製する方法であって、請求項29に記載の抗体を用いる方法。

【請求項31】

SPG4遺伝子に関連するAD-HSPの遺伝子型診断法であって、請求項1～7および13のいずれか一項に記載の核酸配列を用いる方法。

【請求項32】

患者由来の生物学的サンプルを用いる、SPG4遺伝子配列の少なくとも1つの突然変異の存在に関連するAD-HSPの遺伝子型診断法であって、

以下の工程：

- a) 必要に応じて、解析される生物学的サンプル由来のゲノムDNAを単離し、または生物学的サンプルのRNAからのcDNAを作製し、
- b) 請求項5および6のいずれか一項に記載のプライマーまたは請求項13に

記載の核酸を用いて、突然変異を含むと考えられるSPG4遺伝子の前記DNA配列を特異的増幅させ、

c) 得られた増幅産物の解析し、それらの配列とその対応する正常なSPG4遺伝子配列との比較する、
ことを含んでなる方法。

【請求項33】

SPG4遺伝子によりコードされているポリペプチドの異常な発現に関連するAD-HSPを診断する方法であって、

そのポリペプチドとその抗体との特異的な免疫学的複合体の形成を可能にする条件下で請求項29に記載の1以上の抗体と、試験される生物学的材料とを接触させて、形成される可能性のある免疫学的複合体を検出および/または定量する、方法。

【請求項34】

請求項14～18および28のいずれか一項に記載のポリペプチド、または請求項1～7および13のいずれか一項に記載の核酸と、直接的または間接的に相互作用でき、かつ/またはこれらのポリペプチドの発現または活性を調整することを可能にする、化学的もしくは生化学的化合物を選択する方法であって、

請求項1～7および13のいずれか一項に記載の核酸、請求項14～18および28のいずれか一項に記載のポリペプチド、請求項19および20のいずれか一項に記載のベクター、請求項21に記載の細胞、請求項22および23のいずれか一項に記載の哺乳類、または請求項29に記載の抗体と、候補化合物とを接触させて、そのポリペプチドの活性の改変を検出する、方法。

【請求項35】

SPG4遺伝子の発現または活性を調べるための、請求項1～7および13のいずれか一項に記載の核酸配列、請求項14～18および28のいずれか一項に記載のポリペプチド、請求項19および20のいずれか一項に記載のベクター、請求項21に記載の細胞、請求項22および23のいずれか一項に記載の哺乳類または請求項29に記載の抗体の使用。

【請求項36】

以下の化合物の群から選択される少なくとも一種の化合物を含んでなる、診断キットまたはパック。

- a) 請求項5および6のいずれか一項に記載の核酸、および
- b) 請求項29に記載の抗体。

【発明の詳細な説明】**【0001】****【発明の背景】****発明の分野**

本発明は最も一般的な形態の常染色体性優性遺伝性痙性対麻痺(HSP)の原因である、スパスチンをコードしているSPG4遺伝子の同定および特性決定、そのcDNAのクローニングおよび特性決定、およびまた対応するポリペプチドに関する。本発明はまたベクター、形質転換細胞、およびトランスジェニック動物、また診断方法ならびにキットおよび本発明のポリペプチドと直接的または間接的に相互作用し得る化学的または生化学的化合物を選択する方法に関する。

【0002】**背景技術**

優性遺伝性痙性対麻痺(HSP)は下肢の両側相称かつ進行性痙直を特徴とする中枢神経系の変性疾患である。それら自体は臨床的には歩行困難により現れ、両肢の全体麻痺に進行する可能性がある。この種の疾患の生理病理学はこれまでのところ割合に記載されていない。しかしながら、病理解剖学的データにより発病が脊髄の随意運動性を担う錐体索に限定されると推定できる(Reid, 1997)。HSPには様々な臨床型および遺伝子型のものがある。下肢の部分痙直に相当するいわゆる「単純」HSPは、その下肢痙直が神経学的または非神経学的タイプの他の臨床徴候に関連している「複合」HSPとは臨床上区別される(Bruyn et al., 1991)。遺伝学的見地から、HSPは常染色体性優性(AD-HSP)、常染色体性劣性(AR-HSP)またはX染色体性(X-HSP)様式により遺伝し得る。常染色体性優性様式により最も一般的に遺伝する「単純」型HSPが最もよく起こる(HSPの約80%)(Reid, 1997)。HSPの発病率は、疫学的研究が少なく、さらに臨床学的変化が多くあることから統計的に予測することは難しいが、デンマークでは0.9:100000、スペインのある地域では3~9.6:100000(Polo et al., 1991)またはノルウェイでは14:100000(Skre, 1974)(フランスでは約3:100000)と様々である。

【0003】

H S Pは、異なる家系間だけでなく、同じ家系の異なる患者間でも観察されるこの多くの臨床学的変化に加え、重要な遺伝子異質性によっても特徴づけられる。A D - H S Pの場合、これまでに4つの遺伝子座が第14染色体(遺伝子座S P G 3)(Hazan et al., 1993)、第2染色体(遺伝子座S P G 4)(Hazan et al., 1994; Hentali et al., 1994)、第15染色体(遺伝子座S P G 8)(Fink et al., 1995)および第8染色体(遺伝子座S P G 8)(Hedera et al., 1999)上で同定されている。A D - H S Pを示す多くの家系の調査では、第2染色体にある遺伝子がこの型の疾患の主要な遺伝子座であり、調査した家系の40~50%に見られることがわかった(The Hereditary Spastic Paraplegia Working Group, 1996; Durr et al., 1996)。いくつかの遺伝子座S P G 4関連H S P家系では先行的徴候が見られた: この徴候はデンマーク人の6家系においてR E D (Rapid Expansion Detection)手法を用いて示された(C A G) n反復配列の延長に関係していた(Nielsen et al., 1997)。しかしながら、調べたいずれの家系においてもこの方法またはY A C (酵母人工染色体)またはB A C (細菌人工染色体)クローン(Hazan et al., Genomics, 60 (3), 309-19, 1999)からなる物理的マップにおける(C A G) nタイプの配列の系統調査によりこの延長が確認される可能性はなかった。

【0004】

これまでのところ、2つの形態のX - H S Pおよび1つの形態のA R - H S Pの原因である3つの遺伝子が同定されている。ニューロン特異的細胞付着分子、L 1 - C A M (L 1細胞付着分子)をコードする、X q 2 8番(遺伝子座S P G 1)にある遺伝子の変異はその痙直が知的障害に関連している複合型H S Pを引き起こし、ミエリン層の構成性分子をコードする、X q 2 1番(遺伝子座S P G 2)にあるP L P (プロテオリピドタンパク質)遺伝子の変異は単純および複合型X - H S Pを引き起こす(Saugier-veber. P. et al., 1994)。さらに最近、対麻痺をコードする、1 6 q 2 4 . 3番(座S P G 7)にある遺伝子、A A A (様々な細胞活性に関連するA T Pアーゼ)タンパク質系(Confalonien et al., 1995)のミトコンドリアA T Pアーゼの変異は複合および単純型A R - H S Pに関連していた(Casari et al., 1998)。

【0005】

従って現在、最も一般的な形態のAD-HSPの原因である遺伝子を同定し、特性決定する必要性が大いに残っている。この遺伝子の同定により、特に、関係する家系の出生前スクリーニング用試験の将来性のほか、脊髄の神経管束に特異的なこれらの変性をもたらす分子機構のいくつかがよりよく理解され、また患者の治療処置に関して基本的応答を得ることが可能になりさえするであろう。

正確にはこれが本発明の目的である。

【0006】

【発明の概要】

遺伝子座SPG4関連HSP家系において組換え結果を調査することにより、D2S352およびD2S2347遺伝子マーカー間の位置範囲を決定した後、本発明者らは、約1.5Mbと見積もった物理的距離にわたるBACのコンテイングを確立し、SPG4範囲を配列決定することに基づく位置クローニング戦略に取り組み、候補領域にある総ての遺伝子を完全に同定した。2つのBAC、D(b336P14)およびG(B763N4)の配列の解析により、約100kbの距離にわたって延びる17のエキソンからなり、かつAAA系のタンパク質をコードする遺伝子と相同性を示す遺伝子の存在が明らかになった。この遺伝子の配列をAD-HSP家系の健康な人と影響を受けた人との間で比較することで患者の様々な突然変異を証明することができる。

【0007】

従って本発明の目的は最も一般的な型のAD-HSPの原因である、AAA系の新規細胞核要素をコードするSPG4(またはSPAST)遺伝子の同定および特性決定である。

【0008】

第一の態様において、本発明の目的は、

ヒトSPG4遺伝子のゲノム配列である配列番号1の配列；

配列番号1の配列の核酸の相同体または変異体である核酸配列；

それに対して相補的な配列；および

その対応するRNAの配列

からなる群から選択される配列の、少なくとも15個の連続するヌクレオチド、好ましくは20、25、30、35、40、45、50、75、100または200個の連続するヌクレオチドを含んでなることを特徴とする、精製または単離されたSPG4遺伝子の核酸である。

【0009】

【発明の具体的説明】

本発明はもちろんDNAおよびRNA配列の両方とそれらとハイブリダイズする配列、ならびにそれに対応する二本鎖DNAに関する。

【0010】

「核酸」、「核酸配列」または「核酸の配列」、「ポリヌクレオチド」、「オリゴヌクレオチド」、「ポリヌクレオチド配列」、および「ヌクレオチド配列」とは、本明細書において等しく使用されており、二本鎖DNA、一本鎖DNAならびにそのDNAの転写産物、および/またはRNA断片、核酸の断片または領域が定義できる、改変されていても改変されていなくてもよい、正確に連続したヌクレオチドをさす、異常なヌクレオチドを含んでも含まなくてもよいその単離された天然、または合成断片をいうものとする。「異常なヌクレオチドを含んでも含まなくてもよい単離された天然、または合成DNAおよび/またはRNA断片」とは、核酸の断片、セグメントまたは領域が定義できる、改変されていても改変されていなくてもよい、正確に連続したヌクレオチドを意味するものである。

【0011】

本発明はそれらの天然の染色体環境、すなわち自然状態におけるゲノムヌクレオチド配列に関するものでないものと理解されるべきである。それは単離および/または精製された、すなわちそれらの環境が少なくとも部分的に改変されており、それらが直接的または間接的に、例えばコピーにより分離された配列を含んでいる。

【0012】

「相同な核酸配列」とは、参照核酸配列に関して、特に欠失、末端切断、延長、キメラ融合および/または変異、特に点変異などのある改変を有し、かつ整列

後にその核酸配列が参照核酸配列と少なくとも80%、好ましくは90%または95%同一性を示す、配列をいうものとする。

【0013】

本発明の目的のため、2つの核酸またはアミノ酸配列間の「同一性%」とは、最良の整列後に得られた、比較する2つの配列間で同一であるヌクレオチドのまたはアミノ酸残基の割合%をいい、この割合%は単に統計的なものであり、2つの配列間の差はランダムにかつその長さ全体にわたって分布している。2つの核酸またはアミノ酸配列間の配列比較は通常それらを最良に整列した後にこれらの配列を比較することにより行われる。この比較はセグメントによりまたは「比較のウィンドウ」により行われ、配列の類似する局所領域を同定して比較する。手動にの他、Smith and Waterman (1981) [Ad. App. Math. 2:482]の局所相同アルゴリズムにより、Neddleman and Wunsch (1970) [J. Mol. Biol. 48:443]の局所相同アルゴリズムにより、Pearson and Lipman (1988) [Proc. Natl. Acad. Sci. USA 85:2444]の類似性調査方法により、およびこれらのアルゴリズムを用いるコンピュータプログラム (Wisconsin Genetics Software Package, Genetics Computer Group, 575 Science Dr. Madison, WIのGAP、BESTFIT、FASTAおよびTFASTA、またはBLAST NまたはBLAST P比較プログラムを用いて) により、比較する配列の最良の整列を行うことができる。

【0014】

2つの核酸またはアミノ酸配列間の同一性%は、これら2つの最良に整列された配列を、比較する核酸またはアミノ酸配列の領域がこれら2つの配列間の最良の整列のために参照核酸配列に対して付加または欠失を含んでなってもよい、比較のウィンドウにより比較することで求められる。同一性%は2つの配列間でヌクレオチドまたはアミノ酸残基が同一である同一位置数を求めて算出する。ここでは比較のウィンドウにおけるこの同一位置数を全位置数で割り、得られた結果に100を掛けてこれら2つの配列間の同一性%を得る。

【0015】

例えば、サイト<http://www.ncbi.nlm.nih.gov/gorf/bj2.html>で入手可能なBLASTプログラム「BLAST2配列」(Tatusova et al., "Blast2 sequences-

a new tool for comparing protein and nucleotide sequences", FEMS Microbiol. Lett. 174: 247-260)を用いてもよい。ここで用いるパラメーターは誤りにより得られるものであり(特に、パラメーター「オープンギャップペナルティー」; 5および「エクステンションギャップペナルティー」; 2: 選択されるマトリックスが、例えばプログラムにより推奨される「BLOSUM62」マトリックス)、比較する2つの配列間の同一性%はプログラムにより直接算出される。

【0016】

好ましくは、その相補配列が本発明の配列の1つと特異的にハイブリダイズし得る配列が含まれる。好ましくは特異的または高いストリンジェントハイブリダイゼーション条件とは、それらが2つの配列の一方とそのもう一方に相補的な配列間の整列後の同一性が少なくとも80%、好ましくは90%または95%であるよう保証するものである。

【0017】

高いストリンジェント条件下でのハイブリダイゼーションとは、温度およびイオン強度条件によって2つの相補DNA断片間のハイブリダイゼーションを継続し得るようにそれらを選択することを意味する。例えば、上述のポリヌクレオチド断片を定義するためにはハイブリダイゼーション工程の高いストリンジェント条件が以下のように有利である。

【0018】

DNA - DNAまたはDNA - RNAハイブリダイゼーションを2工程で行う: (1) 5×SSC (1×SSCは0.15M NaCl + 0.015Mクエン酸ナトリウム溶液に相当する)、50%ホルムアミド、7%ドデシル硫酸ナトリウム(SDS)、10×デンハート、5%硫酸デキストランおよび1%サケ精子DNAを含有するリン酸バッファー(20mM、pH7.5)中、42°Cで3時間のプレハイブリダイゼーション; (2) プローブの大きさによる温度で(すなわちプローブの大きさ > 100個ヌクレオチドの場合、42°C) 20時間の実際のハイブリダイゼーション、次いで2×SSC + 2%SDSにより20°Cで20分間、2回の洗浄および0.1×SSC + 0.1%SDSにより20°Cで20分間、1回の洗浄。最終洗浄は0.1×SSC + 0.1%SDSによりプローブの

大きさ > 100個ヌクレオチドの場合、60 で30分間行う。規定の大きさのポリヌクレオチドの高いストリンジェントハイブリダイゼーション条件は、Sambrook et al., 1989の教示に従い、当業者によって、より大きなまたはより小さなオリゴヌクレオチド用にも調整される。

【0019】

参照核酸配列の「変異体である核酸配列」または「変異体である核酸」とは、対立遺伝子変異体、すなわち参照核酸配列の各々の変異に対応する核酸配列群をいうものである。これらの天然の変異配列は哺乳類、特にヒトに存在する多形性、特に病状を引き起こしおよび/または進行させ得る多形性に対応している。

【0020】

本発明の配列は正常配列に関するものであるが、またそれらは正常配列に関して少なくとも1つの点突然変異、および好ましくは多くて10%の突然変異を含む範囲で変異させた配列に関する。

【0021】

特に変異核酸配列は配列番号1の配列のヒトSPG4遺伝子のゲノム配列の多形配列の、少なくとも15個の連続するヌクレオチド、好ましくは20、25、30、50、100または200個の連続するヌクレオチドのいずれもの配列および配列番号1の配列に関して、特にアミノ酸残基の末端切断、欠失、置換および/または付加に相当する少なくとも1つの突然変異を有する核酸配列を含み得る。この場合、ここでは少なくとも1つの突然変異を有する変異核酸配列はSPG4遺伝子と関連しているAD-HSP型の病状と関連している。

【0022】

好ましくは、本発明はその突然変異が正常配列によりコードされるポリペプチドのアミノ酸配列の改変を引き起こす変異核酸配列に関する。

【0023】

「変異核酸配列」とはまた、配列番号1のゲノム核酸配列のスプライス部位の突然変異の結果として生じるいずれのRNAまたはcDNAをも意味するものである。

【0024】

好ましくは、本発明は、

a) 配列番号1、配列番号2、配列番号72、配列番号106の配列、または配列番号1、配列番号2、配列番号72もしくは配列番号106の配列の少なくとも15個、好ましくは20、25、30、35、40、45、50、75、100もしくは200個の連続するヌクレオチドの配列；

b) 配列番号1、配列番号2、配列番号72または配列番号106の配列の相同体または変異体である核酸配列；および

c) a) および b) で定義される配列に対応する相補配列またはRNA配列（好ましくは、受託番号AB029006としてGenBankデータベースで確認される核酸を除く）

からなる群から選択される配列を含んでなる、精製または単離されたSPG4遺伝子の核酸に関する。

【0025】

その配列が受入番号AB029006としてGenBankデータベースで開示されている核酸はKazusa DNA Research Institute in Japan(Kikuno et al., DNA Research, 6, 197-205, 1999)により同定されたヒト脳mRNAライブラリーの100のcDNAのうちの1つの配列に対応している。

【0026】

好ましくは、本発明は、配列番号2の配列の、それに対して相補的な配列の、またはその対応するRNAの配列の、nt714~809（末端を含む）断片の少なくとも15個の連続するヌクレオチド、好ましくは20、25、30、50または75個の連続するヌクレオチドの配列を少なくとも1つ含んでなる、本発明の精製または単離された核酸に関する。

【0027】

好ましくは、本発明は次の群：

配列番号1の配列；

ヒト・スパスチンをコードしているcDNA配列である、配列番号2の配列；

図5に示されたネズミ・スパスチンをコードしている不完全cDNA配列を示す、「マウス」系、配列番号72、およびその完全配列を示す配列番号106の

配列；

配列番号1、配列番号2、配列番号72または配列番号106の配列の相同体または変異体である核酸配列；

それに対して相補的な配列；および

その対応するRNAの配列

から選択される配列を含んでなる、精製または単離された本発明の核酸に関する

。

【0028】

好ましくは、本発明は、ヒトの本来の多形性に対応する少なくとも1つの突然変異を含んでなる、本発明の精製または単離された核酸に関するものであり、特にその位置および性質は表5で確認できる。

【0029】

本発明の核酸配列を含んでなるプローブまたはプライマーもまた本発明の一部をなす。

【0030】

従って本発明は本発明のヌクレオチド配列から誘導できるものであり、さらに特にPCR法などの増幅方法、または関連する方法を用いて本発明のヌクレオチド配列、特に変異配列を証明することのできるプライマーセットに関する。

【0031】

本発明はまた、本発明のヌクレオチド配列から、特にそれらとハイブリダイズし得る配列から推定でき、さらにそのヌクレオチド配列を証明できる、特に正常配列と変異配列とを識別することのできるプローブセットに関する。

【0032】

本発明は特に、配列番号4～配列番号71の配列から選択される配列を有するプローブまたはプライマーに関する。

【0033】

本発明はまた、核酸配列を検出、同定、アッセイまたは増幅するためのプローブまたはプライマーとしての本発明の核酸配列の使用に関する。

【0034】

本発明によれば、核酸配列を検出、同定、アッセイまたは増幅するプロセスにおいてプローブまたはプライマーとして使用できるポリヌクレオチドの最小サイズは15塩基、好ましくは20塩基、またはより好ましくは25~30塩基である。

【0035】

本発明のプローブおよびプライマーセットは、当業者に十分公知な方法を用いて放射性または非放射性化合物で直接的または間接的に標識して検出および/または定量シグナルを得ればよい。

【0036】

本発明の非標識ポリヌクレオチド配列はプローブまたはプライマーとして直接使用できる。

【0037】

一般に、多くの用途に使用できる配列を得るためには配列を標識する。本発明のプライマーまたはプローブの標識は放射性元素または非放射性分子を用いて行われる。

【0038】

使用される放射性同位元素としては³²P、³³P、³⁵S、³Hまたは¹²⁵Iが挙げられる。非放射性物質としてはビオチン、アビジンまたはストレプトアビジンなどのリガンド、ジオキシゲニン、ハプテン、色素および放射発光性、化学発光性、生物発光性、蛍光性または燐光性薬剤などの発光剤から選択される。

【0039】

従って本発明のポリヌクレオチドは特に、PCR（ポリメラーゼ連鎖反応）法（Erlich. 1989; Innis et al., 1990、およびRofls et al., 1991）を用いるプロセスにおいてプライマーおよび/またはプローブとして使用できる。この手法では増幅しなければならない断片を組み立てるオリゴヌクレオチドプライマー対を選択する必要がある。例えば、米国特許第4,683,202号に記載される手法を挙げられる。例えばアガロースもしくはポリアクリルアミドゲル電気泳動後、またはゲル濾過もしくはイオン交換クロマトグラフィーなどのクロマトグラフ

イー法の後に増幅された断片を同定し、次いで配列決定することができる。増幅の特異性は、プライマーとして本発明のポリヌクレオチドのヌクレオチド配列、およびマトリックスとしてこれらの配列または誘導された増幅産物を含有するプラスミドを用いて制御できる。増幅されたヌクレオチド断片をハイブリダイゼーション反応において試薬として用い、生物学的サンプルにおいてその増幅されたヌクレオチド断片のものと相補的な配列を有する標的核酸の存在を証明することができる。

【0040】

本発明はまた本発明のプライマーを用いた増幅により得ることのできる核酸にも向けられる。

【0041】

PCR (PCR様) の代わりとして標的核酸を増幅する他の手法を、本発明のヌクレオチド配列を有するプライマー対を用いて使用することが有利である。「PCR様」とは、核酸配列の直接的または間接的再生産を用いる、またはその標識系が増幅されているあらゆる方法をいうものとする。これらの手法はもちろん公知であり、一般にそれらはポリメラーゼによるDNAの増幅を含んでいる：元のサンプルがRNAである場合、事前に逆転写を行うことが望ましい。現在、この増幅が可能な、例えばSDA (鎖置換増幅) 法(Walker et al., 1992)、Kwoh et al. in 1989により記載されたTAS (転写に基づく増幅系) 法、Guatelli et al. in 1990により記載された3SR (自立配列複製) 法、Kievitis et al. in 1991により記載されたNASBA (核酸配列に基づく増幅) 法、TMA (転写により媒介される増幅) 法、Landegren et al. in 1988により記載され、Barany et al. in 1991により改変された、易熱性リガーゼを用いるLCR (リガーゼ鎖反応) 手法、Segev in 1992により記載されたRCR (修復連鎖反応) 法、Duck et al. in 1990により記載されたCPR (サイクリングプローブ反応) 法、およびMiele et al. in 1983により記載され、特にChu et al. in 1986ならびにLizardi et al. in 1988、後にBurg et al.、さらにまたStone et al. in 1996により改変されたQ - - レプリカーゼ増幅法などの極めて多くの方法がある。

【0042】

検出される標的ポリヌクレオチドがmRNAである場合、本発明のプライマーを用いる増幅反応を行う前に、または本発明のプローブを用いる検出プロセスを行う前に逆転写型酵素を用いて生物学的サンプルに含まれるmRNA由来のcDNAを得ることが有利である。次いで得られたcDNAは本発明の増幅または検出プロセスにおいて用いられるプライマーまたはプローブの標的として役立つ。

【0043】

プローブハイブリダイゼーション法は種々の方法により行うことができる(Matthews et al., 1988)。最も一般的な方法は、種々の組織の細胞から、または培養中の細胞から抽出した核酸を支持体(ニトロセルロース、ナイロンまたはポリスチレン)に固定化し、さらに所定の条件下で固定化した標的核酸をプローブとともにインキュベートするものである。ハイブリダイゼーション後に過剰なプローブを排除し、形成されたハイブリッド分子を適当な方法(プローブと結合した放射能、蛍光性または酵素活性の測定)を用いて検出する。

【0044】

本発明の核酸プローブのもう1つの具体例によれば、後者を捕捉プローブとして用いることができる。本件では「捕捉プローブ」とよばれるプローブを支持体に固定化し、これを用いて調べる生物学的サンプルから得られた標的核酸を特異的ハイブリダイゼーションにより捕捉する。次いで容易に検出可能なエレメントで標識した「検出プローブ」とよばれる第2のプローブを用いて標的核酸を検出する。

【0045】

表3で確認される本発明のスプライス受容部位または供与部位配列(配列番号74~配列番号105)もまた本発明の一部をなす。

【0046】

もう1つの態様では、本発明は、cDNAまたはゲノムDNAライブラリーをスクリーニングする、またはスパスチンをコードしている単離されたゲノムまたはcDNAをクローニングする方法であって、本発明の核酸配列を用いる方法を含んでなる。

【0047】

これらの方法のうち、特に、

本発明の核酸配列を用いる cDNA ライブラリーのスクリーニングおよび単離された cDNA のクローニング (Sambrook et al., 1989; Suggs et al., 1981; Woo et al., 1979);

ゲノムライブラリー、例えば BAC のスクリーニング (Chumakov et al., 1992; Chumakov et al., 1995)、および所望によりスパスチンをコードしている SPG4 遺伝子の単離および染色体の限局化、さらに完全配列決定を可能にする本発明の配列を用いる FISH による遺伝子解析 (Cherif et al., 1990) が挙げられる。

【0048】

特に本発明によれば、これらの方法は他の哺乳類、特にマウスの SPG4 遺伝子のゲノム配列または cDNA を同定して得るために用いてよい。

【0049】

これらのスクリーニングおよび/またはクローニング方法は特に、本発明の核酸とゲノムまたは cDNA ライブラリーにある核酸とのハイブリダイゼーション工程を含んでなる。

【0050】

本発明はまた、配列番号 1 の配列の SPG4 遺伝子の発現を促進および/または調節する核酸配列を同定する方法であって、本発明の核酸を用いる方法を含んでなる。

【0051】

当業者に利用できるコンピュータツールにより、本発明のゲノム核酸配列を用い、遺伝子発現を制御するのに必要かつ十分なプロモーター調節ボックスを容易に同定することが可能となる。特に TATA、CCAAT および GC ボックス、またシス位で本発明の遺伝子発現を制御する刺激性調節配列 (「エンハンサー」) または阻害性調節配列 (「サイレンサー」) が挙げられ、これらの調節配列の中でも IRE、MRE および CRE が挙げられる。

【0052】

本発明はまた、ヒト SPG4 にある突然変異、特に常染色体性優性遺伝性痙性

対麻痺の原因である突然変異を同定する方法であって、本発明の核酸配列を用いる方法に関する。

【0053】

これらの突然変異を同定するこれらの方法は特に以下の工程を含んでなる：(i) 解析される生物学的サンプル由来のDNAの単離、または生物学的サンプルのmRNAからのcDNAの作製；(ii) 本発明のプライマーを用いる、突然変異を含むと考えられる標的DNAの特異的増幅；(iii) 増幅産物、特に増幅産物の大きさおよび/または配列の、参照配列に対する解析。

【0054】

「本発明の突然変異を同定する方法」とはまた、それに対して突然変異が同定された核酸を得ることが可能な方法をいうものとする。

【0055】

対応するタンパク質の発現を改変し得る突然変異を有する本発明のSPG4遺伝子のプロモーターおよび/または調節配列も本発明の一部をなす。

【0056】

本発明のこれまでの方法の1つを用いて得られることを特徴とする核酸、または高いストリンジェント条件(2つの配列の一方と、もう一方に相補的な配列との間の相同性が少なくとも80%)下で該核酸とハイブリダイズし得る核酸、特に変異体または相同体、特に配列番号1の配列のSPG4遺伝子の対立遺伝子変異体の、または配列番号2の配列のそのcDNAの核酸配列、またマウスなどの他の哺乳類の相同遺伝子のゲノム配列も本発明の一部をなす。

【0057】

本明細書において「Sp g 4」とはヒトSPG4遺伝子に相同なマウス遺伝子をいうものとする。

【0058】

もちろんゲノムライブラリーまたはcDNAをスクリーニングするための本発明の核酸配列の、プローブまたはプライマーとしての使用も本発明の主題の一部をなす。

【0059】

もう一つの態様では、本発明は、本発明の核酸によってコードされている精製または単離されたポリペプチドであって、好ましくはその配列が受託番号A B 0 2 9 0 0 6としてGenBankデータベースで確認される584アミノ酸ペプチドを除くものを含んでなる。

【0060】

本明細書において「ポリペプチド」とは、同等にタンパク質またはペプチドをいうのに使用される。

【0061】

好ましくは、本発明は、以下の群から選択されるアミノ酸を含んでなる本発明のポリペプチドに関する：

配列番号3の配列、ヒトSPG4遺伝子のcDNAの配列番号2の配列によってコードされているヒト・スパスチンに対応；

配列番号73の配列、マウスSp g 4遺伝子の不完全なcDNAの配列番号72の配列によってコードされているネズミ・スパスチン断片に対応（なおこの配列73は図4Aの「SPAST_マウス」の欄に示されている；

配列番号107の配列、マウスSp g 4遺伝子の完全cDNAの配列番号106の配列によってコードされているネズミ・スパスチンに対応；

配列番号3、配列番号73または配列番号107の配列のポリペプチドの相同体および変異体であるポリペプチド配列；および

少なくとも8、10、15、30または50個の連続するアミノ酸の、その断片配列。

【0062】

また好ましくは、本発明の主題は、

a) 配列番号3、配列番号73、配列番号107の配列、またはこれらの配列の1つの少なくとも10個の連続するアミノ酸の配列；および

b) 配列番号3、配列番号73または配列番号107の配列の相同体または変異体である配列

からなる群から選択されるアミノ酸配列を含んでなる、本発明のポリペプチドである。

【0063】

また好ましくは、本発明の主題は、配列番号3の配列のaa197~228(末端を含む)断片の少なくとも8個、好ましくは少なくとも10、15、20または30個の連続するアミノ酸の配列を含んでなる、本発明のポリペプチドである。

【0064】

また好ましくは、本発明の主題は、以下の群から選択されるアミノ酸配列を含んでなる本発明のポリペプチドである：

配列番号3、配列番号73および配列番号107の配列(なお、この配列はヒトにおける本来の多形性に対応する少なくとも1つの突然変異、特にその性質と位置が以下の表5で確認されるもの、または本発明に従いSPG4遺伝子の突然変異を同定する方法を用いて同定され得るものを有する)；および

少なくとも8、10、15、30または50個の連続するアミノ酸の、その断片配列。

【0065】

本発明は天然型のポリペプチドに関するものではない、すなわちそれらの環境中では得られないと理解されるべきである。特に本発明は天然源から精製によって得られるか、あるいは遺伝子組換えによるか、または化学合成によって得られ、従って非天然アミノ酸を含むペプチドに関する。本発明のヌクレオチド配列の1つを用いて行うことができる組換えペプチドの生産は、高い純度の所望のポリペプチドを得ることができるので特に有利である。

【0066】

「同種ポリペプチド」とは特に1以上の欠失もしくは末端切断、延長、キメラ融合および/または1以上の置換といった参照ポリペプチドに対してある改変を有するポリペプチドをいうものとし、そのアミノ酸配列は整列後に参照アミノ酸配列に対して少なくとも80%、好ましくは90%または95%の同一性を示す。

【0067】

「変異型ポリペプチド」(またはタンパク質変異体)とは、上記の変異型核酸

配列によってコードされているポリペプチドのセットをいうものとする。

【0068】

特に変異型ポリペプチドは配列番号1の配列のSPG4遺伝子の変異型ゲノム配列によってコードされているいずれかのポリペプチドを含み、そのアミノ酸配列は配列番号3の配列に対して特に末端切断、欠失、置換および/またはアミノ酸残基の付加に対応する少なくとも1つの突然変異を有する。この場合、少なくとも1つの突然変異を有する変異型ポリペプチドはAD-HSP型の病状に関連している。

【0069】

「変異型ポリペプチド」とはまた、配列番号1のゲノム核酸配列におけるスプライス部位の突然変異から生じたいずれかのポリペプチドをいうものとする。

【0070】

本発明はまた、本発明の核酸配列を含むクローニングおよび/または発現ベクターも含む。

【0071】

宿主細胞において該配列の発現および/もしくは分泌させるエレメント、または本細胞接近配列を含むことを特徴とする発明のベクターもまた本発明の一部をなす。

【0072】

本発明のプロモーターおよび/または調節配列を含むことを特徴とするベクターもまた本発明の一部をなす。

【0073】

該ベクターは好ましくはプロモーター、末端切断開始および終結シグナル、また転写調節に好適な領域を含む。それらは細胞内で安定して維持でき、かつ、所望により翻訳タンパク質の分泌を具現化する特定のシグナルを持つべきである。

【0074】

これら種々の制御シグナルは使用される宿主細胞に対して選択される。これを達成するため、本発明の核酸配列は選択された宿主内で自律的に複製するベクター、または選択された宿主に組み込まれるベクターに挿入することができる。

【0075】

自律的に複製する系のうち、宿主細胞に対してプラスミドまたはウイルス型の系が用いるのが好ましく、可能なウイルスベクターとしては特にアデノウイルス (Perricaudet et al., 1992)、レトロウイルス、レンチウイルス、ポックスウイルスまたはヘルペスウイルス (Epstein et al., 1992) がある。当業者ならばこれらの系の各々に対して使用可能な技術が分かる。

【0076】

宿主細胞の染色体への配列の組み込みが望まれる場合には、例えばプラスミドまたはウイルス型の系を使用すればよく、かかるウイルスとしては例えばレトロウイルス (Temin, 1986) または AAV (Carter, 1993) がある。

【0077】

非ウイルスベクターの中でも、付随VICAL、酵母での発現のための酵母人工染色体 (YAC)、ネズミ細胞での発現のためのマウス人工染色体 (MAC)、好ましくはヒト細胞での発現のためのヒト人工染色体 (HAC) によって開発された技術に従う裸のDNAまたは裸のRNA裸などの裸のポリヌクレオチドが好ましい。

【0078】

かかるベクターは当業者に一般に用いられている方法に従い製造され、それらから得られたクローンは、リポフェクション、エレクトロポレーションまたは熱ショックのような標準的な方法を用いて好適な宿主に導入することができる。

【0079】

本発明はまた宿主細胞、特に本発明のベクターで形質転換された原核および真核細胞、また本発明の形質転換細胞の1つを含んでなる、ヒトを除くトランスジェニック動物を含む。

【0080】

これらの目的に使用できる細胞の中では、もちろん細菌細胞 (Oline and Lee, 1993) が挙げられるが、酵母細胞 (Buckholz, 1993)、ならびに動物細胞、特に哺乳類細胞の培養物 (Edwards and Aruffo, 1993) および特にチャイニーズハムスター卵巣 (CHO) 細胞、また例えばバキュロウイルスを手段とする方法の使用

が可能となる昆虫細胞(Luckow, 1993)も挙げられる。本発明のタンパク質を発現するのに好ましい細胞宿主はCHO細胞である。

【0081】

本発明の哺乳類の中でも、本発明のポリペプチドを発現するマウス、ラットまたはウサギなどの動物が好ましい。

【0082】

本発明の哺乳類の中でも、SPG4遺伝子の2つの対立遺伝子のうち少なくとも一方の配列がヒトの本来の多形性に対応する少なくとも1つの突然変異を含むことを特徴とする形質転換細胞を含んでなるもの、特にその性質および位置が以下の表5で確認されるもの、または本発明に従うSPG4遺伝子の突然変異を同定するための方法を用いて同定され得るものもまた好ましい。

【0083】

本発明の哺乳類の中でも、本発明のスパスチンをコードする遺伝子が機能的でないか、またはノックアウトされていることを特徴とする、マウス、ラットまたはウサギなどの動物も好ましい。

【0084】

本明細書でさらに特に有利な動物モデルには、特に、

SPG4遺伝子の2つの対立遺伝子配列のうち少なくとも一方に、その位置および性質が表5で確認されるか、または本発明に従う方法を用いて同定される少なくとも1つの突然変異を有するトランスジェニック動物(これらのトランスジェニック動物は例えば胚幹細胞における相同組換え、これらの幹細胞の胚への導入、生殖系細胞に作用するキメラの選択および該キメラの増殖によって得られる)；

本発明の突然変異の1つが導入され得るSPG4遺伝子を過剰発現するトランスジェニック動物(好ましくはマウス)(このマウスは例えば本来、または組織種選択的に偏在する強いプロモーターの制御下での遺伝子コピーの徒論巢F_e句シオンにより、あるいはウイルス転写後に得られる)；

LOXP/CREレコンビナーゼ系(Rohlmann et al., 1996)またはこの遺伝子の発現を不活性化する他のいずれかの系を用いて不活性化することによる本発

明のSPG4遺伝子の欠失を作り出したトランスジェニック動物（好ましくはマウス）がある。

【0085】

本発明の細胞および哺乳類は以下に記載されるように本発明のポリペプチドを生産する方法に使用できるし、また解析およびDNA（ゲノムまたはcDNA）ライブラリースクリーニングのためモデルとしても使用できる。

【0086】

このように上記の形質転換細胞または哺乳類は本発明のポリペプチドと、本発明のポリペプチドの活性に直接的または間接的に関わる化学またはタンパク質化合物との間の相互作用を調べるためにモデルとして使用することができ、これは働きを持つようになる種々のメカニズムや相互作用を調べるためともなる。

【0087】

それらは特に本発明のポリペプチド、特に本発明の配列番号3の配列のヒト・スパスチンまたはその変異体と、補因子として、あるいは阻害剤、特に競合阻害剤として（すなわちこれらは本発明のポリペプチドの活性に対してアゴニストまたはアンタゴニスト活性を有する）相互作用する産物を選択するのに使用できる。好ましくは該形質転換細胞またはトランスジェニック動物は特に上記のSPG4遺伝子と関連している病状に対抗することができる産物の選択を可能にするモデルとして使用される。

【0088】

本発明はまた、本発明のポリペプチドと直接的または間接的に相互作用し得る、かつ/またはこれらのポリペプチドの発現もしくは活性を調整し得る化学または生化学化合物をスクリーニングするための、本発明の細胞、哺乳類またはポリペプチドの使用に関する。

【0089】

本発明はまた、組換えポリペプチドを合成するための、本発明の核酸配列の使用に関する。

【0090】

組換え型で本発明のポリペプチドを生産する方法はそれ自体本発明に含まれ、形質転換細胞、特に本発明の細胞または哺乳類が本発明の核酸配列によりコードされている組換えポリペプチドを発現させる条件下で培養して組換えポリペプチドを回収することを特徴とする。

【0091】

前記の生産方法を用いて得ることができる組換えポリペプチドもまた本発明の一部をなす。

【0092】

上記のようにして得られた組換えポリペプチドはグリコシル化形態および非グリコシル化形態のいずれであってもよく、天然の三次構造を持っていても持っていないともよい。

【0093】

これらのポリペプチドは当業者に公知の組換えポリペプチド生産法に従って、上記に定義された核酸配列に基づいて生産できる。この場合、使用される核酸配列は細胞宿主内でその発現が可能となるシグナルの制御下に置かれる。

【0094】

組換えポリペプチドの生産に有効な系は本発明のベクターおよび宿主細胞を必要とする。

【0095】

これらの細胞は上記で定義されたベクターに挿入されたヌクレオチド配列を宿主細胞に導入し、次いでトランスフェクトされたヌクレオチド配列の複製および/または発現させる条件下で該細胞を培養することによって得ることができる。

【0096】

使用される組換えポリペプチド精製法は当業者には公知である。組換えポリペプチドは細微溶解物および抽出物から、および/または培地上清から、分画、クロマトグラフィー法、モノクローナルまたはポリクローナル抗体を用いる免疫親和性技術などといった個々に使用される、または併用される方法により精製することができる。

【0097】

本発明のポリペプチドは、化学合成により（これは多くの公知のペプチド合成法の1つ、例えば固相を手段とする技術、もしくは部分的固相技術を用いる技術を用いる）、あるいは断片の縮合または通常の溶液合成によって得ることができる。

【0098】

固相合成法は当業者に十分公知である。特にStewart et al.(1984)およびBodansky (1984)を参照。

【0099】

化学合成によって得られ、かつ、対応する非天然アミノ酸を含み得るポリペプチドもまた本発明に含まれる。

【0100】

本発明のポリペプチドを特異的に認識し得ることを特徴とするモノクローナルもしくはポリクローナル抗体またはその断片、キメラ抗体、あるいは免疫複合体も本発明の一部をなす。

【0101】

特に通常の手順に従い遺伝子組換えまたはペプチド合成によって作製された特異的ポリクローナル抗体は本発明のポリペプチドに対して免疫化された動物の血清から得ることができる。

【0102】

本発明のあるポリペプチド、その変異体または免疫フラグメントを特異的に認識する抗体が有利であることが特に記載される。

【0103】

特異的モノクローナル抗体はKohler and Milstein, 1975によって記載された通常のハイブリドーマ培養法に従って得ることができる。

【0104】

本発明の抗体は例えばキメラ抗体、ヒト化抗体、またはFabもしくはF(ab')₂フラグメントである。それらはまた検出可能かつ/または定量可能なシグナルを得るために標識抗体または免疫複合体の形態であってもよい。

【0105】

本発明はまた、本発明のポリペプチドを検出および/または精製する方法であって、本発明の抗体を使用することを特徴とする方法に関する。

【0106】

本発明はまた、本発明の方法を用いて得られる精製ポリペプチドを含んでなる。

【0107】

さらにポリペプチドを精製するためのそれらの使用の他、本発明の抗体、特にモノクローナル抗体はまた生物学的サンプルにおいてこれらのポリペプチドを検出するのに使用できる。

【0108】

このようにそれらは例えば免疫蛍光もしくは金標識により、または酵素的免疫複合体を用い、特異的組織切片において本発明のポリペプチド、特に配列番号3の配列のポリペプチドまたはその変異体の配列を免疫細胞化学的または免疫組織化学的に解析する手段からなる。

【0109】

それらは特に、生物学的サンプルまたは組織におけるこれらのポリペプチドの異常な発現を証明することを可能にし、これによって疾病の進行の監視や分子診断に有用なものとなる。

【0110】

さらに一般的には、本発明の抗体は、本発明の正常なまたは変異型のポリペプチドの発現が認められるはずであるいずれの状況にも有利に使用できる。

【0111】

本発明に従い対立遺伝子変異性、突然変異、欠失、異形接合性の欠如、またはSPG4遺伝子のいずれかの遺伝的異常性を調べる方法であって、本発明の核酸配列または抗体を用いることを特徴とする方法もまた本発明の一部をなす。

【0112】

このように本発明はSPG4遺伝子に関連した病状の遺伝子型診断法であって、本発明の核酸配列を用いることを特徴とする方法を含む。

【0113】

好ましくは、本発明は、患者由来の生物学的サンプルを用いる、SPG4 遺伝子配列の少なくとも1つの突然変異の存在に関連する疾患の遺伝子型診断法であって、以下の工程：

a) 必要に応じて、解析される生物学的サンプル由来のゲノムDNAを単離し、または生物学的サンプルのRNAからのcDNAを作製し；

b) 本発明のプライマーを用いて、突然変異を含むと考えられるSPG4 遺伝子のDNA配列を特異的増幅させ；

c) 得られた増幅産物を解析し、それらの配列とその対応する正常なSPG4 遺伝子配列との比較する、
ことを含むことを特徴とする方法に関する。

【0114】

本発明はまた、SPG4 遺伝子によりコードされているポリペプチド、特に配列番号3の配列のポリペプチドの異常な発現に関連する疾病を診断する方法であって、そのポリペプチドとその抗体との特異的な免疫学的複合体の形成を可能にする条件下で本発明の1以上の抗体と試験される生物学的材料とを接触させて、形成される可能性のある免疫学的複合体を検出および/または定量することを特徴とする方法に関する。

【0115】

これらの方法は例えば、患者由来の生物学的サンプルを用いて上記の配列のうち少なくとも1つにおける突然変異の存在を調べることによる、本発明に従うSPG4 遺伝子における突然変異の存在に関連するAD-HSPの診断法、特に出生前診断法に向けられる。解析される核酸配列は等しくゲノムDNAであっても、cDNAであっても、あるいはmRNAであってもよい。

【0116】

本発明に基づく核酸または抗体はまた、患者における陽性診断、またはリスクのある個体、特に家系に疾病履歴を持つ個体における前徴候診断を可能にするために用いてもよい。

【0117】

もちろん野生型遺伝子に対してある遺伝子の突然変異を証明することが可能な

極めて多くの方法が存在する。それらは本質的に2つの主要なカテゴリーに分類できる。第1のタイプの方法は、変異型配列とその対応する野生型配列とを比較することによって突然変異の存在を検出するものであり、第2のタイプの方法は、例えば突然変異の存在による誤対合の証拠によって突然変異の存在を間接的に検出するものである。

【0118】

これらの方法では記載されている本発明のプロブおよびプライマーを使用できる。それらは通常、少なくとも15個のヌクレオチド、好ましくは20、25または30個のヌクレオチドを含んでなる精製された核酸ハイブリダイゼーション配列であり、本発明の核酸配列と特異的にハイブリダイズし得ることを特等とする。

【0119】

好ましくは特異的ハイブリダイゼーション条件は上記または実施例に記載されているようなものである。これらの核酸ハイブリダイゼーション配列の長さは15、20または30~200ヌクレオチド、好ましくは10~50ヌクレオチドの範囲であってよい。

【0120】

対立遺伝子変異性、突然変異、欠失、異形接合性の欠如、または遺伝的異常性を調べる方法の中でも、本発明の核酸配列を有するプライマー対を用いる、異常性があると思われる本発明の標的配列に関する少なくとも1つのいわゆるPCR（ポリメラーゼ連鎖反応）またはPCR様増幅工程を含む方法が好ましい。これらの増幅産物は標的とされる産物の検出およびアッセイを行う前に好適な制限酵素で処理してもよい。

【0121】

本発明のSPG4遺伝子の突然変異はその翻訳産物の種々の改変の原因であり得、これらの改変はおそらく診断的アプローチに使用できる。具体的にはこれらの突然変異に関連する抗原性の改変は特異的抗体の開発を可能とすると考えられる。変異型の遺伝子産物はこれらの方法を用いて識別できる。これらの改変は総て例えばRIAまたはELISAのような、正常なポリペプチドまたは変態を認

識するモノクローナルまたはポリクローナル抗体の使用に基づいたいくつかの十分公知な方法を用いて診断的アプローチに利用できる。

【0122】

このように本発明の主題はまた、以下の化合物群から選択される少なくとも1つの化合物：

a) 本発明の特にプライマーまたはプローブとしての核酸；および

b) 本発明の抗体

を含んでなることを特徴とする、本発明に従うSPG4遺伝子における突然変異の存在に関連するAD-HSPを特に診断するための診断キットまたはパックでもある。

【0123】

もう1つの態様では、本発明はSPG4遺伝子に関連するAD-HSPを予防および/または治療することができる化学化合物または生化学化合物を選択する方法であって、本発明の核酸配列、本発明のポリペプチド、本発明のベクター、本発明の細胞、本発明の哺乳類または本発明の抗体を使用することを特徴とする方法を含む。

【0124】

本発明のポリペプチドと、または本発明の核酸と直接的または間接的に相互作用し得る化学化合物または生化学化合物を選択し、かつ/またはこれらのポリペプチドの発現もしくは活性を調整することを可能とする方法であって、本発明のポリペプチド、本発明の形質転換細胞または本発明の哺乳類と候補化合物とを接触させて該ポリペプチドの活性の改変を検出することを含む方法も本発明に含まれる。

【0125】

例えば、限定するものではないが、マッチメーカー・ツー・ハイブリッド・システム2など、細菌または酵母のツーハイブリッド系を、マッチメーカー・ツー・ハイブリッド・システム2（カタログ番号K1604-1, Clontech）とともに提供されている取扱い説明書に従って用いて、本発明のポリペプチドと相互作用し得る分子を同定する方法が挙げられる。

【0126】

本発明に従いSPG4遺伝子のプロモーターおよび/または調節配列と相互作用するタンパク質をコードしている核酸はClontech製のマッチメーカー・ワン・ハイブリッド・システムキット(カタログ番号K1603-)とともに提供されている取扱い説明書に記載されるもののようなワンハイブリッド系を用いてスクリーニングおよび/または選択することができる。

【0127】

その他の態様では、本発明は、SPG4遺伝子の発現または活性を調べるための本発明の核酸もしくはポリペプチド、本発明のベクター、本発明の細胞、または本発明の哺乳類の使用を含む。

【0128】

本発明のその他の特徴および利点は実施例および図面を示す本明細書の以下の説明で明らかにする。なお、図面の説明は以下に示されている。

【0129】

【実施例】

例1：材料および方法

1) 候補領域のサブクローニングおよび配列決定

2種のヒトゲノムライブラリー、CITB_978_SKB (Research Genetic社が販売) およびRPCI-11 (Osoegawa et al., 1998)を起源とし、SPG4範囲を包含する12のBACを選択し、配列決定した(Hazan et al., Genomics, 60 (3), 309-19, 1999)。各BACの40 μgのDNAをCviJI制限酵素(CHIMERx)で部分的に消化し、0.4% LMPアガロースゲル(FMC)での電気泳動により分離した。そのサイズが3.5および10 kbの領域で変化するDNA画分を - アガラーゼ(Biolabs)で溶出し、予め、SmaIで消化して脱リン酸化したプラスミドベクターpBAM3と5×ベクターにつき1×挿入物の割合で連結した。エレクトロコンピテント大腸菌(E. coli) DH10B菌(GIBCO-BRL)をエレクトロポレーションにより種々の連結物で形質転換した。各BACより、20%が10 kbの挿入物を有するクローン、40%が5 kbの挿入物を有するクローンおよび40%が3 kbの挿入物を有するクローンからなる、約1000~

1500のサブクローン(8~10当量のゲノム)を単離した。これらのクローンの挿入物の末端をLICOR4200自動シーケンサーで配列決定した。各BACの配列をPhredおよびPhrapプログラムを用いていくつかのコンティグからなる主鎖に構成した。各コンティグ間のホールを標識したジデオキシヌクレオチドを用いてABI377シーケンサー(PE-Applied Biosystems)で配列決定した。これらの配列コンティグに含まれるエキソンをGRAIL II、GENSCAN、FGENEHおよびGenieコンピュータプログラムを用いて推定した。またその配列をBLASTNおよびBLASTXプログラムを用いてEMBLおよびGenBank核酸およびタンパク質データベースにおいて比較した。プロモーター配列の決定はTSSGおよびTSSWコンピュータプログラムを用いて行った。これら総ての配列解析の結果をGenotator配列注解プログラムを用いて視覚化した。

【0130】

2) cDNAクローニング

SPG4遺伝子のcDNAをMarathon cDNA増幅キット(Clontech)を用いて製造業者の教示に従い、胎児脳、成人脳および成人肝臓のポリA+RNAにおける5'および3'RACE-PCR試験により単離した。種々のプライマー対を用い、第1のPCRに次ぎに内部PCRを行った。そのプライマーの配列を以下の表1に示す。

【0131】

【表1】

表1 RACE-PCRおよびcDNA増幅に用いたプライマー

プライマー	配列(5'→3')	5'位のペア/PCR産物の大きさ		
SPA_5RACE5	CGGAGCTCCTCTTGGCTGCCATG (配列番号4)	nt 405		
SPA_5RACE6	AGAAGCGCTGGCAGAGCCACACGAAG (配列番号5)	nt 372		
SPA_5RACE7	AAGGCGACCAAACGCAGCAGCGCAAG (配列番号6)	nt 331		
SPA_3RACE1	AGGAGCAAGCTGTGGAATGGTATAAG (配列番号7)	nt 550		
SPA_3RACE2	TGGTTATGGCCAAGGACCGCTTACAAC (配列番号8)	nt 689		
SPA_3RACE3	CAAACGGACGTCTATAATGACAGTAC (配列番号9)	nt 747		
SPA_3RACE4	TTAGGAATGTGGACAGCAACCTTGC (配列番号10)	nt 1075		
SPA_3RACE5	CTTCTCTGAGGCCTGAGTTGTTAC (配列番号11)	nt 1207		
SPA_3RACE6	TGCTAGAATGACTGATGGATACTCAGG (配列番号12)	nt 1736		
SPA_3RACE7	AGATGCAGCACTGGGTCCATCCG (配列番号13)	nt 1787		
SPA_3RACE8	ATGAACGTCATCGGCTACAGAAACAG (配列番号14)	nt 2037		
SPA_Db	TAGCAGTGGCTGCCGCCGT (配列番号15)	nt 45	b+m	655 bp
SPA_Dm	AAGCGGTCCTTGGCCATAAC (配列番号16)	nt 700		
SPA_Dc	GGCGGCAGTGAGAGCTGTG (配列番号17)	nt 106	c+n	543 bp
SPA_Dn	CTAGCTCTTTCACACTGTTC (配列番号18)	nt 649		
SPA_Ad	AACAGGCCTTCGAGTACATC (配列番号19)	nt 487	d+n	746 bp
SPA_Am	CTGTGAACAACCTCAGGCCTC (配列番号20)	nt 1233		
SPA_Ac	ATGAGAAAGCAGGACAGAAG (配列番号21)	nt 532		
SPA_An	TGCCAAGTCTTGACCAGC (配列番号22)	nt 1175		
SPA_Ba	CTACAACTGCTACTCGTAAG (配列番号23)	nt 1036	a+m	763 bp
SPA_Bm	CAGTGCTGCATCTTTTGCC (配列番号24)	nt 1799		
SPA_Bb	TAGGAATGTGGACAGCAACC (配列番号25)	nt 1076		
SPA_Bn	AAAGCTGTTAGGTCATTCC (配列番号26)	nt 1780		
SPA_Ca	TGGAGATGACAGAGTACTTG (配列番号27)	nt 1550	a+m	766 bp
SPA_Cm	CTGGAATACTTTCATCTGC (配列番号28)	nt 2316		
SPA_Cb	ATGAGGCTGTTCTCAGGCG (配列番号29)	nt 1603		

【0132】

RACE-PCR産物をTA-クローニングキットを用いてクローニングし、その対応するクローンをABI377 (PE-Applied Biosystems)で配列決定した。SPG4転写物の配列は6人の健康な個体のリンパ芽球由来のcDNA集団から増幅したPCR産物の配列決定により確認した。

【0133】

3) 突然変異の検出

調査される家系の1患者と6人の対照個体のリンパ芽球株からRNA PLU SRキット(bioprobe Systems)を用いて全RNAを抽出した。cDNA合成を標準条件下で、ランダム6量体プライマー(Pharmacia)100 pmolおよびSuper script II逆転写酵素(Gibco BRL)200単位を用いて、RNA500 ng~1 µgにおいて行った。親および対照のcDNAにおいてSPG4オープンリーディングフレームの総てを包含する重複断片を生ずる4種のPCR増幅を行った。断片IをSPA__Db / SPA__Dmプライマーを用い、次いでSPA__Dc / SPA__Dnプライマーを用いて、内部PCRにより増幅した。断片II、IIIおよびIVを各々、SPA__Ad / SPA__Am、SPA__Ba / SPA__BmおよびSPA__Ca / SPA__Cmプライマー(表1にあるこれらのプライマーの配列を参照)を用いて増幅した。各々の増幅をcDNA4 µl(-1/7番目の調製物)、各プライマー20 pmol、dNTP200 µM、KCl50 mM、Tris10 mM、pH9、1.5 mM MgCl₂、tritonX-100 0.1%、ゼラチン0.01%およびTaqポリメラーゼ(Cetus-PE)2.5単位を含有する全量50 µlで行った。PCR反応は「ホットスタート」プロセスに従い行った: 94 °Cで5分間の最初の変性工程後にTaqポリメラーゼを92 °Cで加える。次ぎにサンプルを変性(94 °Cで40秒間)、ハイブリダイゼーション(55 °Cで50秒間、断片Iを除く: 58 °Cで50秒間)および伸長(72 °Cで1分間)、次ぎに最終伸長工程(72 °Cで5分間)35サイクルに付す。PCR産物を断片I、II、IIIおよびIV各々に対して、SPA__Dc / SPA__Dn、SPA__Ac / SPA__An、SPA__Bb / SPA__BnおよびSPA__Cb / SPA__Cmプライマーを用いてABI377自動シーケンサー(PE-Applied Biosystems)で配列決定した。

【0134】

また親および対照のSPG4遺伝子の17の推定エキソンを配列決定することにより突然変異を探し、または確認した。各エキソンをエキソン1(gSPAex1c / gSPAex1m)、およびgSPAex10a / gSPAex12m

およびgSPAex11a/gSPAex12mプライマー対を用いて同時増幅されるエキソン10、11および12を除く、対応する「a+m」プライマー対（以下の表2参照）を用いて増幅した。

【0135】

【表2】

表2 エキシソンの増幅および配列決定のためのPCRプライマー

エキソン	産物の大きさ	PCRプログラム	プライマー	配列(5'→3')(配列番号30~71)
1	1048 bp	0	gSPAex1c	GTGAGCCGAACTGCACATTG
			gSPAex1m	CAAAGTCGACAGCTACAGTGC
			gSPAex1d	GGAAGTGTAGTTGAGTGGGA
			gSPAex1n	AGATGAGGCTCCGACCTAC
2	624 bp	3	gSPAex2a	AATGCCACACTTGTAACTCTC
			gSPAex2m	TGTGAATATATCATAATTTGGG
			gSPAex2b	TACAGCAGTTCTCATGATG
3	812 bp	1	gSPAex3a	GACCAAATTGGTGCATGCATG
			gSPAex3m	ACATTTCCAATACATCCCAC
4	379 bp	3	gSPAex4a	ATTTGTCATTTACATGCAC
			gSPAex4m	TTAGAATGACTATACCTGAC
			gSPAex4n	TCAGGTTAAGTAAGACTC
5	830 bp	4	gSPAex5a	TTCTATCTACCTAGTGAC
			gSPAex5m	TTTTATAGCAAGTTGCCCTG
			gSPAex5b	CCTATGAAGATCCTGGTAC
6	484 bp	3	gSPAex6a	TGTCATGATTCTAACAAGGG
			gSPAex6m	TCTATTTCACTCCTGACATG
7	420 bp	2	gSPAex7a	GTCATAGGGCTTAGGCTTC
			gSPAex7m	ATCATACTACCCACTTTTCC
8	647 bp	3	gSPAex8a	TGTTTGGGAAGATGCTACTG
			gSPAex8m	CTACTGAAGATAACGTACATG
9	1268 bp	1	gSPAex9a	CATTGATTGCCATGTATTGG
			gSPAex9m	AGAAGGCCAGAAATACTCAG
			gSPAex9b	GTACTIONAATCGGTAATATGG
10]	1061 bp	4	gSPAex10a	CTCAAGTCTTAGGAATGCAG
11			gSPAex10b	GCACTTAACCAGGCTGTATG
12]	551 bp	3	gSPAex11a	CTCAGATGACTCACATAGC
			gSPAex12m	CTTACTAGACTAATTCTCCTG

13	1361 bp	4	gSPAex13a	CAGATTCAAGAAGACAGATC
			gSPAex13m	GCAATAATTCACCACACTTG
			gSPAex13n	GGTAGTTCTTGTCTTCTGCTC
14	985 bp	4	gSPAex14a	CAAGTGTGGTGAATTATTGC
			gSPAex14m	GAGCTGAAAAGTATTCAGC
			gSPAex14n	TGCAAAGGACATAGCCAGTG
15	1076 bp	1	gSPAex15a	AGCCTCTGGAGATAGTATGC
			gSPAex15m	CTAGAACAGGGGTCACAGTC
			gSPAex15n	TTGGACTTCTTAAACTTC
16	1404 bp	4	gSPAex16a	GCAGTATGCAAGAAATTGAAC
			gSPAex16m	GGCCTGTAATTTTCTTCTG
			gSPAex16b	GTAAGTGAATAGATACATGTAG
17	445 bp	3	gSPAex17a	GTGTAGCAGATCAACATAG
			gSPAex17m	CATCTTCAAGTTTGGTGAC

【0136】

アドバンテージGCゲノムPCRキット(Clontech)を用いて製造業者の教示に従い増幅されるエクソン1以外には、4種のわずかに異なるPCRプログラム(1、2、3および4)を用いてSPG4エクソン(表2参照)を増幅した。増幅は総て、ゲノムDNA100ng、各プライマー50pmol、dNTP250μM、1×TakaraバッファーおよびTakara La Taq Taqポリメラーゼ(Shuzo Co.)1単位を含有する容量50μlで行った。PCR反応は「ホットスタート」プロセスに従い行った: 96 で5分間の最初の変性工程後にTaqポリメラーゼを94 で加える。次にサンプルを変性(94 で40秒間)、ハイブリダイゼーション(プログラム1: 60 で50秒間、プログラム2: 58 で50秒間、プログラム3および4: 55 で50秒間)および伸長(プログラム1および4: 72 で1分間、プログラム2および3: 72 で40秒間)、次に最終伸長工程(72 で10分間)30サイクルに付す。これらのPCR産物の配列決定をPCRプライマーまたは「b」および「n」と呼ばれる内部プライマー(表2参照)のいずれかを用いてABI377シーケンサー(PE-Applied Biosystems)で行った。

【0137】

4) SPG4の同定

ネズミ・オルトログSPG4に対応するマウス胚盤胞およびE8胚cDNAライブラリー由来のcDNAクローン977312(EST AA560327)および568234(EST AA107866)をIMAGEコンソーシアムを用いて単離し、ABI377シーケンサー(PE-Applied Biosystems)で実験室で配列決定した。SPG4およびそのネズミ・オルトログSPG4の発現プロフィールを調べるために、種々の胎児ならびに成人ヒト組織、およびまたマウス組織(MTCパネル、Clontech)由来のcDNAコレクションを、ヒトcDNAについてはSPA__Ca/SPA__Cmプライマー対を用いた、およびマウスcDNAについてはSPA__Ca/spam(spam:5'-ACCGAAGTCAAGAGCCTATC-3')対を用いたPCRにより製造業者のプロトコールに従って調べた。PCR条件は、サンプルを成人ヒト組織ならびにマウス組織由来のcDNAについては32サイクル、および胎児組織由来のcDNAについては28サイクルに付したことを除き、リンパ芽球株cDNA由来のSPG4の増幅に用いるもの(§突然変異の検出参照)である。増幅産物を2%LMPアガロースゲルでの電気泳動により泳動させた。

【0138】

5) 患者由来の筋肉生検の組織学的解析

Casari et al., 1998に記載された標準的な方法によりSPG4遺伝子座に関する家系の患者由来の筋肉生検において組織学的かつ組織酵素解析を行った。

【0139】

6) 公開データベースにおける受託番号

SPG4(またはSPAST)cDNAおよび推定タンパク質配列、GenBank/EMBL AJ246001;不完全SpG4 cDNAクローン、GenBank/EMBL AJ246002;SPG4(またはSPAST)遺伝子、GenBank/EMBL AJ246003。

【0140】

例2: SPG4遺伝子範囲の配列解析

組換え結果の解析によりSPG4候補領域をD2S352およびD2S234

7 マーカー (1 9、 2 0) 間 0 c M の遺伝子範囲に限定することが可能になる。3 7 B A C からなる S P G 4 範囲の予備配列決定マップを作製した (Hazan et al., in press in Genomics) ; 候補領域は約 1 . 5 M b の物理的距離に及んでいる。クローン A および E 間の 4 k b の単一ホールを除く、 S P G 4 領域にわたって伸びている 1 2 の重複 B A C を選択し、配列決定した (図 1 A) 。注目される領域の約 7 0 % に及んでいるこれらの 7 の B A C (A、 B、 C、 D、 E、 F および G) はすでに配列決定されている。これらの 7 の B A C 配列を核酸およびタンパク質データベースのものと比較し、 4 種のエキソン推定プログラムを用いて解析した。これらの予備的配列解析により、キサンチンデヒドロゲナーゼ、ステロイド S - レダクターゼ 2 および T G F 結合タンパク質をコードする遺伝子に対応する 3 つを含む、 1 4 の可能性ある転写単位を示すことが可能になる。配列解析により検出された 1 4 の遺伝子のうち 9 つのものはすでに E S T (「 Expressed Sequence Tag 」) データベースにおいて確認されており、 S P G 4 範囲に位置していた (Hazan et al., in press in Genomics) : 残る 5 つの遺伝子は候補領域の配列決定により同定のみできる。これら 5 つの新規遺伝子の 1 つはそのコード領域の 3 ' において A A A タンパク質系をコードする遺伝子と相同性を示した (Confalonieri et al., 1995) 。より完全な配列解析により、 S P G 4 (または S P A S T) と呼ばれるこの遺伝子が 1 7 のエキソンからなり、かつ 2 つの隣接する B A C クローン、 D および G に包含される約 9 0 k b の領域にわたっていることが示された (図 1 B 参照) 。この遺伝子の最初の 3 つの推定エキソンは用いられる 4 種の推定プログラムのうちの 2 種、 G R A I L II および G E N S C A N (それらはマウス胚盤胞 E S T、 A A 5 6 0 3 2 7 と強い相同性を示す) により B A C D において同定された。最後の 1 4 のエキソンは B A C G において見出せる。エキソン 7 ~ 1 7 から推定されるタンパク質配列は Y t a 6 p (Schnail et al., 1994)、 T B P 6 (Schnail et al., 1994) および E n d 1 3 酵母タンパク質、また S K D 1 マウスタンパク質 (Perier et al., 1994) も含む、 A A A 系サブクラスと有意に相同である。

【 0 1 4 1 】

4 種のエキソン推定プログラムのうち F G E N E H が最も信頼性がありかつ最

も能力が優れており、2 p 2 1 ~ 2 2 のこの染色体領域の大部分の遺伝子の検出を可能にすると思われる。この所見はまた S P G 4 遺伝子にも当てはまる。この 1 5 のエキソンはこのプログラムを用いて示されるが、4、9 または 1 1 のエキソンだけは各々、Genie、GRAIL II および GENSCAN プログラムを用いて位置決定する。次ぎにこの遺伝子のゲノム構成 (図 1 B) を S P G 4 c DNA の配列を決定することにより確認する。イントロン / エキソン連結については以下の表 3 に示す : エキソンサイズは 4 1 b p (エキソン 1 6) ~ 1 . 4 1 0 k b (エキソン 1 7) の範囲にわたっており、イントロンサイズは 1 4 0 b p (イントロン 1 1) ~ 2 3 . 2 4 7 k b (イントロン 1) の範囲にわたっている。

【 0 1 4 2 】

【表 3】

表 3 SPG4 遺伝子のイントロン / エキソン 構成

エキソン / イントロン	エキソンサイズ (bp)	cDNAに おける位置	スプライス受容部位 (配列番号 74~89)	スプライス供与部位 (配列番号 90~105)	イントロンサイズ (bp)
1	540	1		TGAGAAAG/gtaactaggggctcgg	23 247
2	87	541	atTTTTatTTTaaag/CAGGACAG	AGGACAAAG/gtaagattgtattgt	1 943
3	84	628	aatTTTTctTfcag/GTGAACAG	ACTTCTAG/gtaicaaTTaattgtat	9 190
4	96	712	cttctctgtgcatag/AGAAAGATG	CCAGTCAG/gtgggttaggTaaac	15 745
5	188	808	actTTTccTctgcag/AAAGTGGA	CTCATAAG/gtattciggacagta	876
6	134	996	TTTgtatccTTtaag/GGTA CTCC	GTGGACAA/gtaagTTTgccaTct	283
7	94	1 130	aggctgtttcttag/TGGAACAG	GGCCTGAG/gtaagaactTTatatt	10 735
8	75	1 224	agTatataTTTtttag/TTGTTCAC	CAATGCTG/gtaagggTtctctca	1 385
9	72	1 299	ctTgtattTTTaaag/GCTAAAGC	CAAAATAC/gtgagTctctgtttc	8 083
10	76	1 371	taatgctttgttttag/GTGGGAGA	TTTTATAG/gtaagaacataTTttc	238
11	92	1 447	ctTgtattTcctctag/ATGAAGTT	TTGATGGT/gtaagTgtgattatg	140
12	80	1 539	gattTTTgctTgtag/GTACAGTC	GTTCTCAG/gtagggagatttatat	4 715
13	43	1 619	ggattTTTTtttttag/GCGTTTCA	ATGAGGAG/gtatgatactgtgTtt	1 389
14	80	1 662	TTTaaataTTTttcag/ACAAGACT	C TTGCTAG/gtgagTaatTTggatt	1 521
15	71	1 742	tcctTccctcTcag/AATGACTG	TATCCGAG/gtagTatacaagagc	2 210
16	41	1 813	ctTTtatTTTttacag/AACTAAAA	CCAGTGAG/gtatagTattttcaaa	7 115
17	1 410	1 854	ctTTTaaaaaatctag/ATGAGAAA		

【 0 1 4 3 】

それぞれ上段はエキソンの配列を、下段はイントロンの配列を示す。

例3：SPG4 cDNAの同定

成人肝臓ならびに脳および胎児脳 cDNA のコレクションにおいて 5' および 3' RACE - PCR により数度の連続増幅を行い、SPG4 転写物の同定を行った。転写物の 5' 領域が GC 高含量であることから (90% の GC は nt 263 位の前に置かれた 60 bp にある)、総ての 5' RACE - PCR により SPG4 cDNA の nt 263 位で終結する (図 2) 増幅産物を得た。コード領域の総てを包含する 4 つの重複 PCR 産物を 6 人の対照個体のリンパ芽球由来の cDNA から増幅し、SPG4 転写物の配列を確かめる補助手段により完全に配列決定した。PCR および RACE - PCR 産物総ての配列を調節することにより 125 bp の非翻訳 5' 領域 (「5' 非翻訳領域」の代わりに 5' UTR) を前に置いた 1848 bp のオープンリーディングフレーム、さらに後に続くポリ A テール上流 ~ 35 bp の nt 3227 - 3232 位間のポリアデニル化部位を含む (図 2) 1290 bp の 3' UTR 領域を含んでなる、3263 bp の配列を再構築することができる。SPG4 cDNA と EST データバンクとの比較により、第 2 のポリアデニル化部位を含んでなるさらに延長された 3' 非コード領域 (+ 180 bp) を含む EST N47973 をはじめとする 6 人のヒト EST との有意な相同性を検出することができる。翻訳開始部位を、プリンガイニシエーター ATG の 3 nt 上流に配置され、それ自身が停止コドンの前に置かれることとなる翻訳開始に「好適な関係」として定義されるコサック共通配列 (CTGTGAatgA) の存在により同定した。3263 bp の cDNA 配列は SPG4 遺伝子の 17 のエキソンから推定される転写配列と同一である。TSSG および TSSW コンピュータプログラムを用いる 5' 領域の配列解析によりエキソン 1 の nt 1 位の 43 bp 上流に配置された TATA ボックス型のプロモーター配列の存在が示される。

【0144】

例4：SPG4 遺伝子における突然変異

SPG4 遺伝子座が関連した家系の 14 人の患者 (各家系から 1 人の患者) のリンパ芽球由来の SPG4 cDNA においてヘテロ接合突然変異を調べた。14 人の患者と、また 6 人の健康な対照個体において SPG4 cDNA のオープ

ンリーディングフレームを包含する4つの重複PCR断片I、II、IIIおよびIVを増幅し、配列決定した。PCR断片IVのアガロースゲル電気泳動により、スイスの同じ地域に源を発する家系2992、5226および5330の3人の患者ではスプライス部位の微量欠失または突然変異を示す同じ強度の3つのバンドが示された；家系2992および5330の2人の健康な個体には2つの付加的なバンドは存在しなかった(図3A)。これらの3家系の患者ではエキソン16のゲノム配列によりイントロン15のスプライス受容部位(AG)のヘテロ接合AG突然変異が示された(図3B)；この突然変異はエキソン16の欠失の後、異常転写物におけるリーディングフレームシフトを引き起こすものである。夫および妻をはじめとする健康な構成員のうち、スプライス部位のこの突然変異を有するものはいない。これらのスイスの3家系の総ての患者における同一突然変異の同定により、おそらくはハプロタイプの研究により示された共通の祖先の存在が立証される。

【0145】

アミノ酸置換をタンパク質配列(S362C、C448YおよびR499C)に導入する3つの点突然変異、1210C G、1468G Aおよび1620C Tが、各々家系624、4014および618の患者におけるPCR断片II IおよびIVを配列決定することにより示された。これらの3つの置換総てにはタンパク質配列におけるシステインの欠失または挿入を誘導するシステイン残基が含まれる。465個のアミノ酸(aa)からなる末端切断型タンパク質を誘導する停止コドンを出現させる1bpの欠失、1520delTをA家系の患者において検出した。以下の表4に概略を示す5つの突然変異はいずれも、彼らが本明細書において調べた7家系の健康な兄弟姉妹または配偶者のいずれであろうと、調べた対照個体では見られなかった。これらの5つの突然変異はAAA系の総ての構成員においてATPアーゼ活性を担うと推定されるいくつかのタンパク質モチーフからなる非常に保存されたドメイン、またはAAAカセット(Beyer, 1997)のタンパク質配列に重大な影響を与える。

【0146】

【表4】

表4
AD-HSP患者のSPG4における突然変異

家系	位置	突然変異 ^a	アミノ酸変化 ^b	結果
624	エキソン7	1210C→G	S362C	ミスセンス
4014	エキソン11	1468G→A	C448Y	ミスセンス
A	エキソン11	1520 delT	466STOPコドン	ナンセンス
618	エキソン13	1620C→T	R499C	ミスセンス
2992	イントロン15	1813-2a→g	Δ aa564→aa576 (PTC+7aa)	エキソン16の欠失+シフト
5226	イントロン15	1813-2a→g	Δ aa564→aa576 (PTC+7aa)	エキソン16の欠失+シフト
5330	イントロン15	1813-2a→g	Δ aa564→aa576 (PTC+7aa)	エキソン16の欠失+シフト

^a nt位置はSPG4 cDNAの配列をさす。

^b aa位置はスパスチン配列をさす。

上段ではエキソンの塩基を、下段ではイントロンの塩基を示す。

PTC+7aa=エキソン16の7aa下流にある「成熟前終結コドン」

【0147】

上記のこれらの5つの突然変異に加え、その他の36の家系のAD-HSP患

者においてヘテロ接合突然変異の研究を行い、SPG4遺伝子の発現産物を改変するまたは改変すると思われるその他の34の突然変異を示すことができた。

【0148】

これらその他の34の突然変異の特徴を、上記の最初の5つの変異も加えて以下の表5に概略を示している。

【0149】

【表5】

表5

AD-HSP患者のSPG4における突然変異

家系	位置	突然変異 ^a	アミノ酸変化 ^b	結果
624	エキソン7	1210C→G	S362C	ミスセンス
6958	エキソン8	1233G→A	G370R	ミスセンス
214	エキソン8	1267T→G	F381C	ミスセンス
1002	エキソン8	1283T→G	N386R	ミスセンス
027	エキソン8	1288A→G	K388R	ミスセンス
019	エキソン10	1401C→G	L426V	ミスセンス
4014	エキソン11	1468G→A	C448Y	ミスセンス
148	エキソン11	1504G→T	R460L	ミスセンス
618	エキソン13	1620C→T	R499C	ミスセンス
636	エキソン15	1788G→A	D555N	ミスセンス
627	エキソン15	1792C→T	A556V	ミスセンス
2971	エキソン3	702C→T	Q193STOP	ナンセンス
3655	エキソン5	873A→T	K229STOP	ナンセンス
1010	エキソン5	907C→A	S261STOP	ナンセンス
3938	エキソン5	932C→G	Y269STOP	ナンセンス
6922	エキソン10	1416C→T	R431STOP	ナンセンス
616	エキソン10	1416C→T	R431STOP	ナンセンス
605	エキソン15	1809C→T	R562STOP	ナンセンス
030	エキソン2	578-579insA	PTC+2aa	シフト+ナンセンス
615	エキソン5	852del11	PTC+18aa	シフト+ナンセンス
042	エキソン5	882-883insA	PTC+12aa	シフト+ナンセンス
032	エキソン5	906delT	PTC+17aa	シフト+ナンセンス
189	エキソン9	1299delG	PTC+3aa	シフト+ナンセンス
3686	エキソン9	1340del5	PTC+35aa	シフト+ナンセンス
625	エキソン9	1340del5	PTC+35aa	シフト+ナンセンス
A	エキソン11	1520delT	PTC+7aa	シフト+ナンセンス
115	エキソン12	1574delGG	PTC+2aa	シフト+ナンセンス
3266	エキソン13	1634del22	PTC+18aa	シフト+ナンセンス
149	エキソン14	1684-1685insTT	PTC+9aa	シフト+ナンセンス
645	エキソン14	1685del4	PTC+7aa	シフト+ナンセンス

家系	位置	突然変異 ^a	アミノ酸変化 ^b	結果
029	イントロン4	808-2a→g	?	スプライス部位変異
162	イントロン6	1129+2t→g	?	スプライス部位変異
125	イントロン7	1223+1g→t	?	スプライス部位変異
143	イントロン8	1299+1g→a	?	スプライス部位変異
1620	イントロン11	1538+5g→a	(PTC+6aa)	エキソン11の欠失+シフト
1006	イントロン11	1538+3del4	?	スプライス部位変異
1605	イントロン13	1661+1g→t	?	スプライス部位変異
1012	イントロン13	1662-2a→t	?	スプライス部位変異
1626	イントロン15	1812+1g→a	?	スプライス部位変異
2992	イントロン15	1813-2a→g	Δ aa564→aa576 (PTC+7aa)	エキソン16の欠失+シフト
5226	イントロン15	1813-2a→g	Δ aa564→aa567 (PTC+7aa)	エキソン16の欠失+シフト
5330	イントロン15	1813-2a→g	Δ aa564→aa567 (PTC+7aa)	エキソン16の欠失+シフト
1611	イントロン16	1853+1g→a	?	スプライス部位変異

^a nt位置はSPG4cDNA配列についている。

^b aa位置はスパスチン配列についている。

上段では エキシンの塩基を、下段では イントロンの塩基を示す。

PTC+n aaは、突然変異のアミノ酸n個下流の「成熟前終結コドン」。

【0150】

例5：スパスチンのタンパク質配列の解析

SPG4のオープンリーディングフレームはスパスチンと呼ばれ、その分子量が約67.2キロダルトン(kD)である616aaタンパク質をコードする。タンパク質データベースのこのアミノ酸配列をBLASTプログラムを用いて比較することにより、スパスチンのC末端においてAAA系のいくつかの構成員と著しく相同である領域を示すことができた。AAAカセットに含まれるAAA系の「典型的」モチーフはProDomおよびPrositeタンパク質データベースにおける配列比較により342位と599位との間に確認される(図2参照)。Walker AならびにBモチーフおよびまたAAAタンパク質の最小共通モチーフを含む3つの保存された典型的ドメインはそれぞれ、aa382~389、437~442および480~498位のAAAカセットにおいて確認される(図2)。ATP結合ドメインに対応し、p-ループとも呼ばれるWalker Aモチーフ、「GPPGNGKT」およびBモチーフ、「IIFIDE」はスパスチンをはじめとしてAAA系の総ての構成員の間で非常に保存され

る。

【0151】

進化上極めて遠い生物由来のこのA T Pアーゼ系の150のタンパク質に存在するA A Aカセットを比較することにより、この群のタンパク質を確認されたA A Aカセット数(1または2)およびこれらの種々のカセット間の配列相同性(Beyer, 1997)に関して、いくつかの部分群に分類することができる。A A A系の総てのタンパク質のうち、スパスチンはA A Aの特定のサブクラスと、より詳しくは問題の生物体のゲノムの完全な配列決定によりその大部分が同定された次のタンパク質: シーノラブディチス・エレガンス(*Caenorhabditis elegans*)の2つのタンパク質、O16299およびQ18128; サッカロミセス・セレビスエ(*Saccharomyces cerevisiae*)の26Sプロテアソーム2つのサブユニット、Yta6p(Q02845)およびTBP6(P40326)(Schnail et al., 1994); シゾサッカロミセス・ポンベ(*Schizosaccharomyces pombe*)のプロテアソームのサブユニット(O43078); S. セレビスエのSAP1(P39955)およびEND13(P52917)タンパク質およびネズミSKD1タンパク質(P46467)(Perier et al., 1994)とより著しい相同性を示している。スパスチンとこれら8種のタンパク質との多数のアライメントを図4Aに示している。A A Aカセットを含む257個のアミノ酸(aa342~599位)のうち、スパスチンはYta6p(Q02845)酵母タンパク質、O16299線虫類タンパク質およびTBP6(P40326)酵母タンパク質とそれぞれ、52%、51%および50%の配列同一性を示している。ProDomデータベースにおいてスパスチンのタンパク質配列を解析することにより同様の結果が得られた。ここでは酵母の26Sプロテアソームの推定サブユニットで見られる相同な3つのドメインの存在(92、179および6226と呼ばれ、aa342~409位、411~509位および512~599位に対応している)が示された。さらにこのA A A部分群の構成員は共通してロイシンジッパー型モチーフを含んでおり、このうち2つがPrositeデータベースにおいて配列を解析することにより、aa50~78位および508~529位のスパスチンのタンパク質配列において検出できた(図2参照)。またこの解析によりaa478~4

86位間に確認されるヘリックスループの二量体化モチーフ - ヘリックス型の存在も推定できた。

【0152】

スパスチンのタンパク質配列とAFG3、RCA1およびYME1酵母タンパク質などのミトコンドリアメタロプロテアーゼ、およびさらに稀少型AR-HSPに関連するパラプレギンのものとを比較することにより、AAA系のこれら5つの構成員間の相同性がAAAカセットを含む257aa領域に限定されることがわかる(図4B)。この領域においてスパスチンとパラプレギンとの配列同一性はわずか29%であるが、タンパク質配列のこの同一部分でパラプレギンとAFG3酵母タンパク質とは57%同一である。この配列比較によりスパスチンがパラプレギンおよび他のミトコンドリアメタロプロテアーゼのように、同じAAA部分群に属していないことが示される。さらに、タンパク質の細胞下分布が推定できるPSORTIIプログラムを用いるスパスチン配列のコンピュータ解析によりスパスチンが核タンパク質であることが示されることは明らかである。可能性ある核分布シグナル(NLS)、RGKKKはaa7~11位間に示されたが、パラプレギンでの観察とは異なり、ミトコンドリアへの引き入れの特徴であるシグナルペプチドは検出されなかった。

【0153】

例6：SPG4およびそのネズミ類似遺伝子Spg4の発現プロフィール

ESTデータバンクにおいてSPG4の核酸配列を比較することにより、SPG4と著しい相同性を示す数種類のヒト、ネズミおよびラットESTが検出できた。2つのネズミEST、AA560327およびAA107866に対応するマウス胚盤胞およびE8胚cDNAクローンをIMAGEコンソーシアムから得て、完全に配列決定した。これらのcDNAクローンの配列を構成することにより、1514bpの不完全オープンリーディングフレームを含む1689bpの一致配列が再構築できた。ヒトSPG4cDNAとこのマウスcDNAとを比較することにより、ネズミ転写物が翻訳開始コドンを含む5'末端において約460bpを欠失していることがわかった。マウスオープンリーディングフレームの後にポリAテール~20bp上流にあるポリアダニル化部位を含む175bpの

3'非コード領域(3'UTR)が続いている(図5)。SPG4の核酸配列およびヒトスパスチンのタンパク質配列はマウスcDNAおよび推定タンパク質配列とそれぞれ、89%(nt460~1982位間)および96%(aa113~616位間)の同一性を示す。このようにながりの相同性であることから、マウス転写物がSPG4のネズミ類似遺伝子に対応していることが確認でき、従ってSpq4と呼ばれている。

【0154】

種々のヒトおよびネズミ組織(Clontech)のmRNAを含んでなるノーザンプロットのSPG4およびSpq4cDNAクローンとのハイブリダイゼーションでは10日間の暴露後、マウス精巢の2.5kb転写物に対応する極めて弱いバンド以外、いずれの納得のゆく結果も得られなかった。この遺伝子の発現レベルが低いため、SPG4およびSpq4の発現プロフィールを種々の成人および胎児組織由来の標準cDNAコレクションにおけるPCR試験により求めた(図6A~6C参照)。ネズミSpq4遺伝子はマウス成人組織、さらにマウス胚E7~E17期においても至る所で発現される(図6A)。Spq4のより高い発現は肝臓、骨格筋および精巢、さらに胚E15期においても検出された。胚発生のSpq4の初期発現は公開ESTデータベースの胚盤胞、E8胚および胚癌腫cDNAライブラリー由来のESTの存在により確認された。ヒトSPG4遺伝子はそれ自体、成人(図6B)および胎児(図6C)組織の至る所で発現されるが、胎児脳で顕著に発現されるであろう。

【0155】

例7: SPG4遺伝子座が関連したAD-HSPにおける酸化的リン酸化障害の否定

スパスチン突然変異がパラプレギンの場合と同様にミトコンドリアにおける酸化的リン酸化(OXPHOS)障害を誘発するか否かを確認するため、SPG4遺伝子座が関連した1AD-HSP家系の患者において筋生検を行った。筋生検の形態学的かつ組織酵素解析ではミトコンドリアにおけるOXPHOS障害に特徴的なRRF(「破れた赤色線維(ragged red fiber)」)型筋繊維は全く示されなかった。総ての筋繊維が正常であろうという事実、さらにはスパスチンの核分

布の推定によっても、SPG4遺伝子座が関連したAD-HSPが、SPG7遺伝子座が関連したAR-HSPとは異なり、OXPHOS型のミトコンドリア疾患ではないことが示されているようである。

【0156】

1. 5 Mb領域の配列決定に基づく位置的クローニングアプローチを用いて、発明者らは染色体バンド2p21~p22にすでに確認されている最も一般的な型のAD-HSPの原因となるSPG4(またはSPAST)遺伝子を同定した。SPG4遺伝子座との関連を示すAD-HSPを有する41家系の患者においてスパスチンと呼ばれる発現産物を改変するまたは改変すると思われる39の突然変異が検出できた。スパスチンは核分布すると考えられ、かつ酵母の26Sプロテアソームのサブユニットと著しい相同性を示すAAAタンパク質系の新規構成員である。AAAカセットとよばれる230~250aaのドメインに限定された高い相同性にもかかわらず、このタンパク質系の多くの構成員は液胞内のタンパク質の輸送、細胞周期制御、細胞小器官の発生、すなわち転写制御などといった極めて多様な細胞機構に関係し得る。しかしながら、これら総ての細胞機構はタンパク質複合体の構成、機能または分解に関係しており、このことによりAAA系の構成員がいわゆる「シャペロン(chaperon)」タンパク質であることが示唆される。

【0157】

参考文献

【表6】

- Barany, F., (1991), Proc. Natl. Acad. Sci. USA, 88, 189-193.
- Beyer, A. Sequence analysis of the AAA protein family. Protein Sci. 6, 2043-2058 (1997).
- Bodansky M., Principles of peptide synthesis, (1984).
- Bruyn, R.P.M. & Scheltens, P.H. Hereditary spastic paraparesis (Strumpell-Lorrain) in Handbook of clinical neurology Vol. 15 (ed. de Jong, J.M.B.V.) 301-318 (Elsevier Science Publishers B.V., 1991).
- Buckholz, R.G. Curr. Op. Biotechnology 4 : 538-542, 1993.
- Burg, J.L. et al. (1996), Mol. and Cell. Probes, 10, 257-271.
- Carter, B.J. Curr. Op. Biotechnology 3 : 533-539, 1993.
- Casari, G. et al. Spastic paraplegia and OXPHOS impairment caused by mutations in Paraplegin, a nuclear-encoded mitochondrial metalloprotease. Cell 93, 973-983 (1998).
- Cherif D., Julier, C., Delattre, O., Derré, J., Lathrop, G.M., and Berger, R. Proc. Natl. Acad. Sci. USA. 87 : 6639-6643, 1990.
- Chu, B.C.F. et al. (1986), Nucleic Acids Res., 14, 5591-5603.
- Chumakov, I., Rigault, P., Guillou, S., Ougen, P., Billault, A., Guasconi, G., Gervy, P., Le Chumakov, I.M., Rigault, P., Le Gall, I., et al. Nature 377 : 175-183, 1995.
- Confalonieri, F. & Duguet, M. A 200-amino acid ATPase module in search of a basic function. BioEssays 17, 639-650 (1995).
- Duck, P. et al. (1990), Biotechniques, 9, 142-147.
- Durr, A. et al. Phenotype of autosomal spastic paraplegia linked to chromosome 2. Brain 119, 1487-1496 (1996).
- Edwards, C.P., and Aruffo, A. Curr. Op. Biotechnology 4 : 558-563, 1993.
- Epstein, A. Médecine/Sciences 8 : 902-911, 1992.
- Erlich, H.A., (1989), New York : Stockton Press.
- Guatelli J.C. et al. Proc. Natl. Acad. Sci. USA 87: 1874-1878, 1990 et al. Cell 85 : 281-290, 1996.
- Fink, J.K. et al. Autosomal dominant familial spastic paraplegia : tight linkage to chromosome 15q. Am. J. Hum. Genet. 56, 188-192 (1995).
- Hazan, J., Lamy, C., Melki, J., Munnich, A., de Recondo, J., & Weissenbach, J. Autosomal dominant familial spastic paraplegia is genetically heterogeneous and one locus maps to chromosome 14q. Nature Genet. 5, 163-167 (1993).
- Hazan, J. et al. Linkage of a new locus for autosomal dominant familial spastic paraplegia to chromosome 2p. Hum. Mol. Genet. 3, 1569-1573 (1994).

【表 7】

- Hedera, P. et al. Novel locus for autosomal dominant hereditary spastic paraplegia, on chromosome 8q. *Am. J. Hum. Genet.* 64, 563-569 (1999).
- Heinzlef, O. et al. Mapping of a complicated familial spastic paraplegia to locus SPG4 on chromosome 2p. *J. Med. Genet.* 35, 89-93 (1998).
- Hentati, A. et al. Linkage of a locus for autosomal dominant familial spastic paraplegia to chromosome 2p markers. *Hum. Mol. Genet.* 3, 1867-1871 (1994).
- The Hereditary Spastic Paraplegia Working Group. Hereditary spastic paraplegia : advances in genetic research. *Neurology* 46, 1507-1514 (1996).
- Innis, M.A. et al. (1990), Academic Press.
- Jouet, M. et al. X-linked spastic paraplegia (SPG1), MASA syndrome and X-linked hydrocephalus result from mutations in the L1 gene. *Nature Genet.* 7, 402-407 (1994).
- Kievitits, T. et al. (1991), *J. Virol. Methods*, 35, 273-286.
- Köhler et Milstein. *Nature* 256, 495-497, 1975.
- Kwoh, D.Y. et al. (1989), *Proc. Natl. Acad. Sci. USA*, 86, 1173-1177.
- Landegren U., Kaiser R., Sanders J. & Hood L. *Science* 241 : 1077-1080, 1988.
- Lizardi, P.M. et al. (1988), *Bio/technology*, 6, 1197-1202.
- Luckow, V.A. (1993), *Curr. Op. Biotechnology* 4, 564-572.
- Matthews, J.A. et al. (1988), *Anal. Biochem.*, 169 : 1-25.
- Miele, E.A. et al. (1983), *J. Mol. Biol.*, 171 : 281-295.
- Nielsen, J.E. et al. CAG repeat expansion in autosomal dominant pure spastic paraplegia linked to chromosome 2p21-24. *Hum. Mol. Genet.* 6, 1811-1816 (1997).
- Olins, P.O., and Lee, S.C. *Curr. Op. Biotechnology* 4 : 520-525, 1993.
- Osoegawa, K. et al. An improved approach for construction of bacterial artificial chromosome libraries. *Genomics* 52, 1-8 (1998).
- Perier, F. et al. Identification of a novel mammalian member of the NSF/CDC48p/Pas1p/TBP-1 family through heterologous expression in yeast. *FEBS lett.* 351, 286-290 (1994).
- Perricaudet, M., Stratford-Perricaudet, L. and Briand, P. *La Recherche* 23 : 471-473, 1992.
- Polo, J.M., Calleja, J., Combarros, O. & Berciano, J. Hereditary ataxias and paraplegias in Cantabria, Spain. An epidemiological and clinical study. *Brain* 114, 855-866 (1991).
- Reid, E. Pure hereditary spastic paraplegia. *J. Med. Genet.* 34, 499-503 (1997).
- Rohlmann, A., Gotthardt, M., Willnow, T.E., Hammer, R.E., and Herz, J. *Nature Biotech.* 14 : 1562-1565, 1996.
- Rolfs, A. et al. (1991), Berlin : Springer-Verlag.

【表 8】

- Sambrook, J., Fritsch, E.F., and Maniatis, T. *Molecular cloning : a laboratory manual*.
- Saugier-veber, P. et al. X-linked spastic paraplegia and Pelizaeus-Merzbacher disease are allelic disorders at the proteolipid protein locus. *Nature Genet.* 6, 257-262 (1994).
- Schnall, R. et al. Identification of a set of yeast genes coding for a novel family of putative ATPases with high similarity to constituents of the 26S protease complex. *Yeast* 10, 1141-1155 (1994).
- Scott, W.K. et al. Locus heterogeneity, anticipation, and reduction of the chromosome 2p minimal candidate region in autosomal dominant familial spastic paraplegia. *Neurogenetics* 1, 95-102 (1997).
- Sec. Ed. Cold Spring Harbor Lab., Cold Spring Harbor, New York.
- Segev, D., (1992), Kessler C. Springer Verlag, Berlin, New-York, 197-205.
- Skre, H. Hereditary spastic paraplegia in Western Norway. *Clin. Genet.* 6, 165-183 (1974).
- Stone, B.B. et al. (1996). *Mol. and Cell. Probes*, 10 : 359-370.
- Stewart J.M. et Yound J.D., *solid phase peptides synthesis*, Pierce Chem. Company, Rockford, 111, 2ème éd., (1984).
- Suggs S.V., Wallace R.B., Hirose T., Kawashima E.H. and Itakura K. *PNAS* 78 : 6613-6617, 1981.
- Temin, H.M. Retrovirus vectors for gene transfer. In Kucherlapati R., ed. *Gene Transfer*, New York, Plenum Press, 149-187, 1986.
- Walker G.T., Fraiser M.S., Schram J.L., Little M.C., Nadeau J.G., & Malinowski D.P. *Nucleic Acids Res.* 20 : 1691-1696, 1992.
- Werderlin, L. Hereditary ataxias. Occurrence and clinical features. *Acta Neurol. Scand.* 73 (Suppl. 106) (1986).
- Woo S.L.C. *Methods Enzymol.* 68 : 389, 1979.

【配列表】

LISTE DE SÉQUENCES

<110> CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE - CNRS

<120> CLONAGE, EXPRESSION ET CARACTÉRISATION DU GÈNE SPG4
RESPONSABLE DE LA FORME LA PLUS FRÉQUENTE DE PARAPLÉGIE
SPASTIQUE AUTOSOMIQUE DOMINANTE

<130> D18374

<150> FR 99 11097
<151> 1999-09-03

<160> 107

<170> PatentIn Vers. 2.0

<210> i
<211> 110000
<212> DNA
<213> Homo sapiens
<220>
<221> exon
<222> (9932)...(10471)
<220>
<221> intron
<222> (10472)...(33718)
<220>
<221> exon
<222> (33719)...(33805)
<220>
<221> intron
<222> (33806)...(35748)
<220>
<221> exon
<222> (35749)...(35832)
<220>
<221> intron
<222> (35833)...(45022)
<220>
<221> exon
<222> (45023)...(45118)
<220>
<221> intron
<222> (45119)...(60863)
<220>
<221> exon
<222> (60864)...(61051)
<220>
<221> intron
<222> (61052)...(61927)
<220>
<221> exon
<222> (61928)...(62061)
<220>
<221> intron
<222> (62062)...(62344)
<220>
<221> exon

<222> (62345)...(62438)
<220>
<221> intron
<222> (62439)...(73173)
<220>
<221> exon
<222> (73174)...(73248)
<220>
<221> intron
<222> (73249)...(74633)
<220>
<221> exon
<222> (74634)...(74705)
<220>
<221> intron
<222> (74706)...(82788)
<220>
<221> exon
<222> (82789)...(82864)
<220>
<221> intron
<222> (82865)...(83102)
<220>
<221> exon
<222> (83103)...(83194)
<220>
<221> intron
<222> (83195)...(83334)
<220>
<221> exon
<222> (83335)...(83414)
<220>
<221> intron
<222> (83415)...(88129)
<220>
<221> exon
<222> (88130)...(88172)
<220>
<221> intron
<222> (88173)...(89561)
<220>
<221> exon
<222> (89562)...(89641)
<220>
<221> intron
<222> (89642)...(91162)
<220>
<221> exon
<222> (91163)...(91233)
<220>
<221> intron
<222> (91234)...(93443)
<220>
<221> exon
<222> (93444)...(93484)
<220>
<221> intron
<222> (93485)...(100599)
<220>

<221> exon
<222> (100600)...(102009)

<400> 1							
taactgactc	tgctgtgttt	ccttggcatt	atagctaato	aaattgagca	ggtcaggtaa	60	
cagtttatac	ttacacctac	tatttcaaaa	ccatgagctc	attcacattt	tcactgaagt	120	
aacaaatcct	ccataaacta	gaaaaatctca	aactgggtgac	tgggagtttt	ggttttgttt	180	
ttttgttggt	ttattttatt	ttattttatt	ttctagatgg	agtcttgctc	tgttgccag	240	
gctggaatgc	aatggcatga	tctcagctca	ctgcaacctc	cacctttcgg	gttcaagcaa	300	
tcctcctgcc	tcaatcttcc	aagtagctgg	agctacagga	atgagctgcc	gcacctggcc	360	
tggtttgttt	gtttttaaat	tttgaggcca	ggtgcagtgg	cccatactctg	tgatcccagc	420	
actttgggag	accaaggcag	gccgattact	tgaggtcagg	agttcaagac	cagccaggcc	480	
aacatggtaa	aacctgtcgg	ctactaaaaa	cacaaaaatt	aggccgggca	cggtggtca	540	
cgtctataat	cccagcactt	tgggaggcca	agtgtagtgg	atcacctgag	gtcaggagggt	600	
caagaccagc	ctggcaaaaca	tggtgaaacc	ccgtctctat	taaaaataca	aaaattagcc	660	
gggtgtggtg	gcgggcacct	gtaatcccag	ctattgggga	ggtcaggcca	ggagaatcgc	720	
ttgaacctgg	gagtgaggag	ttgcagtggag	tcgagatggt	gccactgacc	tccagcctgg	780	
gcaataagag	cgagacgcca	tctcaaaaat	ataataataa	ataataaaa	atacaaaaag	840	
tagctaggct	ttgtggcagg	cacctgtagt	cccagctact	taggaggctg	aggcaggaga	900	
attgcttgaa	cccaggaggc	ggaggttgca	gtgagccaag	attgtgccac	tgctgtccag	960	
cctgggtgac	agagcaagac	tccttctcaa	aaataataa	ataataaat	aaataataa	1020	
ataataaaaa	ttttgagctg	ggcatgaaag	ctgaggcagg	aggatccctt	gagcccagca	1080	
gtttgagacc	ccagtgagct	ataattctga	cactgcactt	cagcctggct	gcagaggga	1140	
gaccctgtat	ctaaaaagaa	taaaataaca	atgatttttg	agccaataac	tcttagccaa	1200	
tagatttcac	ataaaaattt	agaattctgg	tttctcttga	aaaattaaaa	aatctgaca	1260	
atgttaagct	tcacattcct	gaaaggcaaa	aatcagtgga	agctagctgg	gtgctgtggc	1320	
tcaccctat	aaaccagtg	ctttgggagg	ctagggtaag	tggattggtt	gagtcaggga	1380	
gtttgagatc	agcctggcca	acacagtcaa	acccatctc	tacaaaacat	acaaaatta	1440	
gctgggatat	ggtggtgtgc	acttgtggtc	ccaactactc	aggaggctga	ggcaggagga	1500	
ccacttgagc	cctggagggt	gagggggcag	tgagccatga	ttgtgccact	gcactccagc	1560	
ctggacgaca	gagagacatt	gtctcaaaaa	gaaaaaaaaa	aatcagctca	gtgggagctg	1620	
agtaacagct	gtcctgttca	attacaggat	gcaactcttt	agcttctcat	agtttccatc	1680	
attcacatca	tacatgcatt	tttgcattgc	acacaaccca	cacatggaac	ccatatatgt	1740	
tatatgcctg	accactgttg	ctattggaag	ttttggccac	tgcattaaac	tataaactcc	1800	
atcttattaa	tcctgacacc	ccacttattg	cctgatata	tgtccgtttc	ttaatatcta	1860	
ttcatagaac	aatgaaatga	ataaatatgtg	ccacattgtg	gactcaattc	aggagatga	1920	
ctaatacttc	acaattatgt	ttttccttct	taatacagag	attggataat	aattccccaa	1980	
ttgttaactc	tctcatccct	cccctccaac	catatctttt	catttgtttt	tcttatttta	2040	
ggttttctgc	ctttttcaaa	atcagccatt	tcctcaactg	actctacatg	tgccattttt	2100	
tttttttttt	ttttactaat	tttttagtt	gaaaagaggt	ccttaatatc	tgtcattggt	2160	
ccacacttga	aatctagaat	ctctcgaatt	gaaggtctga	agagttcctt	agaaggcaag	2220	
tagggccgaa	atatccaaaa	aaaactattt	gatggtaggc	actgtggtaa	ataaatatgg	2280	
gtfaaagtga	taaagtgacc	gaatgtcctg	gattagttag	cacagtacct	agctccttct	2340	
ttttgtcctt	gtgaaactgt	ctcagattcc	attcaagatt	aagtgtcctg	aaagttctta	2400	
caagctgaaa	actgggaggc	tctcacactg	taggtagaat	ggctagcagg	ggtgggatct	2460	
catctaggca	cattgcagaa	ccagggtaac	accaaggcat	tatttttttt	ctttctttac	2520	
ttttcttctt	tttttttaga	gacagggtct	cattatgtta	cccaggctgg	attccaattc	2580	
ctgagctcaa	gtgactcctc	tgagtagctg	ggacttcagg	tgtgcaccgc	tgtgtctgaa	2640	
ggcactgttt	ttttgttttt	ttttgttttt	tttgacacaa	atttaggaag	atgttaattc	2700	
acaacagtct	tgagactgag	atataattcc	aaggagcaga	agatgtgagt	ttagagcact	2760	
aaaaaatgag	attccagtag	agtcagaaat	ctgaaatggc	attacagata	taagagacaa	2820	
aaacaaaatg	tattgagctc	tgtcatgttg	caggcatcat	gatggagggt	ttagatgtac	2880	
tctttcattt	tgtaattttt	atagaggaa	taactagaat	agcaacccca	gtcccactaa	2940	
ctctaaatcc	catattttta	ccatacaaaa	aaagagcaaa	agtgacagaaa	agcacagtca	3000	
atattaagta	caaccagata	gcagagactc	agtaaatggg	aggccggagg	cccgaaccc	3060	
aacataatgc	ccatgaaatga	aagcccatca	cttgcgcttc	aggggctaac	aatatactta	3120	
tttcataatt	aaaaatagaaa	cagactatgt	aaaaatatta	ttcttgagat	cccagatttt	3180	
taccttaaat	tactgaggca	gtaagtgtaa	ttaactaata	tgtgatgttg	ggcaataaac	3240	
agactcttag	agccccaat	tctttatttt	aaaaactga	ggccagatga	ggtggcttac	3300	

acatgtaatc	ccagcactgg	ggatgccaaa	gcaggcagat	cacttgagct	tgagcttagg	3360
agtttgagat	cagcctgggc	aacacgggta	agacctcatc	tctaaacaaa	atacaaaaat	3420
tagcacaaca	tggtgggtgtg	cacotgtgtt	cctggctact	caggaggctg	aggtggaagg	3480
atcgcttgag	cccagagagt	tgaggggtgca	gtgagccatg	atcttgccat	tgcaactctac	3540
actctagtct	gggtgacaga	gtgagactct	gtcttggggg	aaaaacaaaag	agatgataat	3600
gcttaactga	agtagcaata	ttttaaaaag	gcaactaaaag	ttcatctgct	tagttcagaa	3660
tatgggttcg	ataaatatta	gcaagtagta	gtagtcatca	tcatcactgt	cactgctgtt	3720
ctctccttaa	acctaaagcat	gtttttgtt	tttttgagac	agtgtctcac	tctgtcacc	3780
aagctggagt	tcagtgggtg	gatcttggct	cactccaacc	tctgcctccc	aggttcaagt	3840
gattctcctg	tctcaccctc	ccaagtagct	gggaccacag	acacgtacca	caaccacacc	3900
cagctaattt	ttcgtatttt	tggtagaggt	ggggtttcac	catggtggcc	aggtgggtct	3960
tgaactcctg	acctcaagtg	atccacctgt	ctcggcctcc	caaagtgtctg	ggattacagg	4020
cgttagccac	agcatccagc	cttaagcatg	ttaattaagt	ttttataatt	cagcaaaaatg	4080
gttggaanaa	gctgtcttaa	atgagatgct	taagctgccc	tctgaacatg	aggtagaagg	4140
aaattctaca	cataatcatt	gtgctaaatt	acttgcaaaag	atggccacaa	caattcctcc	4200
tatcctcata	tatatgcccc	tttgcaatgt	gactttgcta	cttctctatc	aagatgtgga	4260
gcttattttc	ccatataattg	cactagagtt	ggccttctga	cttgctttga	caatggaatg	4320
tagtacaat	gacactgtgc	aactttggat	tttaggtttc	gagagaactt	acaccttcca	4380
ctcacactct	cttggaacc	agatgcaatg	taaagaagtc	agggtatcc	tgctagagac	4440
atatgtccca	gctaatagcc	acaatcaacc	ctgaaacata	tgaatgaggc	tagctaggcc	4500
atacagccat	tcogtcaagc	catcagatga	ctacatccac	aggaaatgatc	cacaggcaag	4560
gccatcagaa	gaaccatcca	gotgaaactta	ccccaaattg	ctgagtcaca	aagttgtgtg	4620
taataaaatg	totgctatct	taagccagtg	agttttggag	tggtatatta	catagcatca	4680
gaaatctaac	acaatcatta	tgtttgaaatc	atttttcaaa	tttctcatat	ttatataatg	4740
agtaccataa	gcaaggtgtc	aggctggatg	caaaaagtga	ggcaaaaatgt	ataaagtgtg	4800
accactgcct	tcagtaagtt	tacaatctat	atcaagaggt	gatgaagtgt	ttaaataatc	4860
atcctgcagg	gcaatatagt	ataagagcca	cagagtaaca	caaccatatt	gtcataacaa	4920
ctgaaaaaca	agatcatttc	tgctggaggt	gataatggaa	taatttatca	agaatataac	4980
agagctggac	gcggtggctc	acacctgtaa	tcccagcact	ttgggaggcc	aaggaaaggtg	5040
gatcacaagg	tcaggagttc	gagaccatcc	tggtcaaac	gatgaaaccg	tctctactaa	5100
aaatacaaga	aattagccgg	gcgtgggtggc	acgcgtctgt	agtcgcagct	actcaggagg	5160
ctgagggcagg	agaaccactt	gaacatggga	agcagaggtt	gcattgagct	gagatcgtgc	5220
catggcactc	cagcctgggt	gacagagtga	gactcagctc	caaaaaaaaaa	aaaaaaaaaaa	5280
aaaataaac	attagaggta	agtcttgaag	gactttgaca	gtggaaagtag	gaggcgaggc	5340
cattctaaagt	gaatgaaaaa	tgacaggaga	gtaattgtag	tcctggaaaa	gagcaaaagta	5400
ggtacagacc	aacagtctat	attagctaga	gtatagtgaa	agtgcagagg	aaatgtcggga	5460
gaaaccattct	ttattcaaaa	actatcttcc	tcatggccaa	gcatagtggc	tcatgcctgt	5520
aatcccagca	ttttgggagg	tcaaggtgag	aggatcactt	gagctcagga	attcaagacc	5580
atctggggca	acctagtgag	acctcatctc	aaactaaaaa	caaaaaattc	agacagatgc	5640
agtggctcac	acctgtaatc	ccagaacttt	ggtaggctga	ggcgggcgga	tcacgaggtc	5700
aggagatcaa	gacctcctg	gacaaactgg	agaaacccca	tctctattaa	aaatacaaaa	5760
ttagctgggc	atggtggcac	atccctgtaa	tcccagctac	tcgggaggct	gaggcaggag	5820
aatcgcttga	accagggagt	cgagggttgc	agtgagccga	gatcgacca	ctgcactcca	5880
gtctggcgac	agagcggagc	tccatcttaa	aaaataaata	aattttaaaa	aaaactacc	5940
cagcatgggtg	gtgcatgcct	gtagtcccag	ttactcagga	ggctgaggca	agagggtggt	6000
ttgagccagg	gaggtcaagg	ctgcagtgag	ctctgatggc	gccactgtac	tccagcttgg	6060
gtgacagagt	gagaccttgt	ctcaaaaaaca	aaaaaaaaaa	caaaaaacca	acaaatctcc	6120
ttgttagtat	catggtgagt	aaaaaataaa	ataaaaaatag	aaataaactg	aacatggtgg	6180
ctcatgcctg	taatcctagc	actttggaag	gctgaaagtgg	gaggattgct	tgagggtctgg	6240
agttcaaaaac	tggtctgggc	aacacgggtga	gagagacctt	gtctctacaa	aagaactttt	6300
aaaacaaaaa	atagataatt	taaaaaaatt	aaaaaaaaaca	aaataataaaa	aaataatcaa	6360
gtatcaactt	gattccaggc	actgcttact	actctagtgt	tatactgtag	atgtggaagc	6420
tgagttaactc	atccaagatc	accgaaagtg	atggaacaca	gatctaaatg	caaccagtct	6480
gactccagga	ccatttaacc	attctactat	tgggccctat	cttggttaag	ttagaaaagta	6540
agttactttc	tttagtggta	aagactggag	ggataacagg	gaagatagtt	atttaagaaa	6600
aaaaactggc	atcaaaactaa	atatccatca	atagttgaac	agtaaaaatag	gttgtggtaa	6660
attcatataa	tggaatacta	tatagcagtg	aaaatgtacc	acagttatag	aaatcaacag	6720
ggaggaatth	caacacttaa	ttattaagta	ggtagccagg	catagcggtt	tatgcctgta	6780
atcccagcac	tttgggagac	caagacagga	ggattacttg	agcccagggg	ttcgagatca	6840

acctgggcaa	cagtgagact	ccatctctat	ttcttataaa	taaaataaat	gaaatthtaa	6900
aaatthttgag	gagggaaaagc	aaacaaggga	tacttgaat	atgattacat	ttccataaag	6960
tcaaagtgag	gcaaaatcat	acaagacatt	gtttagaat	acataaatac	actgcaaact	7020
aaaaatgaga	cactagaatg	attaatataa	aattcaggat	agtggcttcc	tctagaggaa	7080
gagacaagac	attgagatta	gggaggagct	cacagagtgc	ttcagaggag	tggttacatt	7140
cattthttctt	aaatggaatg	ctgcttatta	tttttcttta	aattgtgcat	ttaagtaaca	7200
cacttcttgt	ttatatgata	tatgtataaa	tgtaatthtt	ttttttgaga	tggagtthcg	7260
ctcttgttgc	ccaggctgga	gtgcaatggc	actatcttgg	ctractgcaa	cctccacttc	7320
ctgggttcaa	gtgattctcc	tgctcagcc	tcccagtag	ctgggattac	aggcatgccc	7380
cacccatgccc	ggctagtthtt	gtatthttta	tagaгааааg	gtttctccat	gttggtcagg	7440
ctggtctcga	actcccagcc	tcagggtgatc	cgctgcctt	ggcctcccaa	agtgttggga	7500
ttacaggtgt	gagccaccgt	gccaggccct	gaatcagatt	taaaagaggg	catttcatta	7560
aaaaaaattht	tttgttthtt	gctthttgaga	cagagtctcg	ctctgtcggc	caggctgcag	7620
tgcatthggca	tgatcttggc	tcaccgcggc	ctcagcctcc	caggthcaag	tgattctcct	7680
gcctcagcct	cgcatctagtt	gagattacag	gaatgcacca	ccaccacagg	aatgcacctg	7740
tctaaacttht	gtatthtttag	tatagaggga	gtthttgccat	gttagccagg	ctgctcttga	7800
actcctgacc	tcoggtgatc	ggctcccaaa	ggctcccaaa	gttctgggat	tacagggctg	7860
agccacccaca	cccggccgaa	agagggcatt	tcagaatgag	ggtctagcat	aagcacagag	7920
aagggggagc	aataagaggg	aaacagggag	taggtcattt	ttgcaatagc	ctgtgacatt	7980
tgtagggcag	tactggcggg	gaataattaa	gtaaaattgg	ctggtgctgt	ggctcatgcc	8040
tgtaatccca	gcactthttgg	aggccgaggg	gggcaggtht	cttgagccca	ggaattcaag	8100
accaaccctgg	gaaacatagc	aagaccctgt	ctcaacaaaa	aagtaaaaaa	attagctggg	8160
ggcgcgatgg	ggtggctcat	gcctgtaatc	ccaacacttht	ggaaggctga	ggcagggcga	8220
ttgcttgagc	ccaggagtht	gagaccagcc	tgggcaacat	ggtgaaacc	tggtctata	8280
aagaatacaa	aaattagctg	ggcccagtht	cgthgtcctg	tgatcccagc	tactcgggag	8340
gctgaggtgg	aaggtatcac	tgagccaggg	aggtggaggt	tcagtgagth	catgthttt	8400
gcccactgc	actccagcct	gggcaatgga	gtgaaaccct	gtccaaaaaa	taaaaaata	8460
aagctgtggc	agaatgtgga	gattcttggga	agctggaagc	tctcatgggg	catttgaaa	8520
cctcacattg	taaataacgg	agthctthtt	tcagthttgg	thccttagth	ttaggaaa	8580
agaaataatt	atggctaact	caagtaaaaa	gagaaagaga	agagaaaaaa	gacgtggaga	8640
tagagagaga	gggagagaga	ggaaaagacg	aaaggaagga	agggggaaa	gagagaggaa	8700
gagagaaaca	gagaaacaga	ctgattagth	tattggatag	attacataac	caagtgacca	8760
gtcaggaaacc	cagcagctct	gggggagctc	aatgtgatgc	attgataaac	ccgctcttaa	8820
gagcactcgt	ttccagthta	thtctattcg	gtgggtctcc	agccaagatt	ccaggtccca	8880
ggagaactct	actgacctag	tgthttgctt	cgctthttcg	gtctgggtht	tgthcttga	8940
gctcattaga	atacagggag	cagagacaag	caggtagtht	cccaaaggaa	gggatgctga	9000
gtagattaaa	aaaaaagtht	agattcttca	gtaaacatag	ggatggtaac	tatgcaaaac	9060
ctaagatttc	ccttattcaa	ataaattatc	thtcatatta	gacatctaaa	tatgcactaa	9120
thtagthtaa	ccctggtht	agthtgatct	atcacactga	gctaacattt	thgttctgt	9180
tgthttgcag	gacctgaagt	thcttatctt	cacaattgct	thctctcaa	ataattccca	9240
gattthtaaat	thttatthta	thttthctgg	agacggagth	thctctctc	gcccaggctg	9300
gagthcagth	gcccgatctc	agctcactth	cagcctctgc	ctcccagtht	caagcgattc	9360
thccggcctta	gccttccaac	cagctgggag	tacagggcgg	cgccccacg	cccggctaath	9420
ttaatthcca	gattgatatc	cattgcttct	gagatgggth	aattatcctt	cggagaagac	9480
thaggthcgc	tggcagaaaa	agatgaaaga	aatctaaaga	aacgacgaca	ctgagagagg	9540
agcctagcga	accagcagag	cgaccccagg	ccgcaattcc	ccctccgtg	gatcgattac	9600
gaaggcttcc	tggcagagag	thctccaggg	tgccgacgth	agccgaactg	caatttggga	9660
actgtagtht	agthggaaag	ccgagagggc	ggggccgcac	acgcgtacag	gggccccgth	9720
caacaaagac	gcccgtgccc	cgccgctgccc	ggagaaaaac	acgggaagac	gtgcccgtgc	9780
gcccgcggcc	ctgggagcca	ccagggcggc	gagagagacg	cgacaggaag	ggaggggccc	9840
gagccaccga	ctgcagaggg	agaagggtht	gtgctcctgg	ccgaggaagg	agaaaagggc	9900
ggggccggcc	ggcagcgtgc	ggcagthcgg	agctctgag	accggcgggc	acacggggth	9960
ctgtggcccc	cgccgtagca	gtggctgccc	ccgtcctth	gttcccgtc	gtctcgggga	10020
ggcggthtat	ggcggcggcc	gcagthgag	ctgtgaaatg	attctccggg	tggaagggg	10080
aagaagaaa	gctccggcgg	cgccagcaac	ccggtgctc	ccaggcctcc	gccccctg	10140
ctggcccccc	ccccccgc	cgccggggcc	gcccctccg	ccgagthcgc	gcataagcgg	10200
aacctgtact	atthctccta	cccgtgtht	gtagctthc	cgctgctg	thtggthcgc	10260
thccacctgg	gctcctctt	cgthgtgctc	tgccagcgt	thctcccgc	cctcatggca	10320
gccaagagga	gctccgggg	cgccgcaaga	cctgcctcgg	cctcggcccc	ggcgcggth	10380

ccgggccccg	aggccgagcg	cgtccgagtc	ttccacaaac	aggccttcga	gtacatctcc	10440
attgccctgc	gcatcgatga	ggatgagaaa	ggtaactagg	gggctggggg	agggggcggc	10500
ggcgccggga	agaaggcggg	ggggtcgccg	ggggagggca	acacctgcgt	ccctttctcg	10560
cgggagggga	cggtgcaccc	ccggaattga	fatgccccgg	gagactgctt	tcccgtagggt	10620
cggagcctca	tcttctagta	ttcttaaaac	ctctcccctt	tcagggcact	gtagctgtcg	10680
actttgtttc	agacaccagc	cttccccccac	acttctgcat	gaccaggtc	actatgagac	10740
acccagacgt	gttgatgaca	gtgacatttg	tcctagagtg	accacactga	tcctttctag	10800
cactgtgaag	agtgtgcagc	ttcctctgaa	ccaaggtttc	caaaaggttt	tgatattgaa	10860
gaagcagtg	cgcttactg	gcttttaatg	aaagcagagt	attgtagtgt	cagaaaaaaa	10920
gaacaaatgg	tgacaatttt	gaaagaaata	gctgcatatg	actgcagttg	aatttgcatc	10980
atthtaataca	agataatcat	tttatcagta	caacgattcc	tgaatacttt	ttcaatgaag	11040
ttatatttag	cataaaactt	ttccccctgt	tgctttgatt	ttaattaaaa	cagtattcca	11100
aagtagcctt	taatttccaa	gttgaaatgt	ttgatgaatg	gattgctgtaa	acttaacat	11160
accactttac	agtaaaacct	aaaacaacta	tgtatgtttc	tgaatgaaag	caaggatact	11220
acatctttcg	ggtttctttt	aagctaactt	tttttttttt	ttgagacgga	atctcgccct	11280
gtcgcccaag	ctggagtgca	atggtgcagt	ctcggtcac	tgcaacctcc	gactccctgg	11340
ttccagcgat	tcctctgcct	cagcctcccg	agtagctggg	attacaggca	cgaccacca	11400
cgctgggcta	atthttgtat	ttttagtaga	gtagggattt	caccatgttg	gccaggatgg	11460
tctccatctc	ctgacctcgt	gatcgggtcg	cctcgccctc	ccaaagtgtc	ggtaatcacg	11520
gcatgagcca	ccgcgcctgg	ccttaagcta	acatttttta	ttatagtgic	caggcattgt	11580
gctattagct	ttgcatgtgt	tatctttttt	cttttataaa	aaaatagcaa	ccatcccaga	11640
ccatgaaaag	tgttatttaa	tcctcacaat	aactttgtga	gatgaaggta	ttattgggat	11700
cagcatttta	gagatgaaga	aatgaggcc	caaaagataa	aggaggttat	tccaaacctg	11760
tgctaataat	gaaatttctt	ttatgaaaaa	taagtgaat	taggagaagt	ctaactttta	11820
caattctccc	cttttatcct	tacttccagt	atgctgagat	cctgtctctc	cctctgccaa	11880
aacaccccac	ttttctacca	cacctcaatt	agatactcac	ttgcattgtc	cattagttaa	11940
aacagaaaca	atctgcactt	cattcatagt	gtctgtctct	actgccaact	ccaaaaactg	12000
ttctagaatt	tcaatcttgt	tgaacctgt	ttcctttgtg	gggctgggga	gtgggaagt	12060
ggatataagg	aaagaggcag	ttattcattg	ttttggacag	taaggaaaga	gtgacggtta	12120
agagaggtca	aggagggtgt	taacatttaa	gaatactatg	tgthttaga	aggaattttt	12180
ttgttaactg	ctccattctt	tttttttttt	ttttttgaga	tggagtctcc	ctgtgtcacc	12240
cagactggag	tgcagtggcg	tgatctcagc	ttactgcaac	ctctgcctcc	cagacgcagg	12300
cgattctctt	gcctcagctc	cccagtagc	tggaaaccaca	ggtgcgccct	atcacggctg	12360
gctaattttt	tgtatthtta	gtagagatgg	tgthtaccca	tgthtggccag	actggtcacg	12420
aactcctgac	ctcaggtgat	tcgcctgcct	gacacctcca	agtgcttggga	ttacagggtgt	12480
cagccaactgt	gcccaccacg	caagctccat	tctttatcac	ctcttaagaa	catccaggat	12540
cccttgggga	gaattaaaac	ggttgcaaa	ttttagaata	gaggaaactg	tttaagcgta	12600
gatcattttt	tgggccaagtc	atggtggctc	acgcgtgtaa	tcccaacct	ttgggaggcc	12660
ggggcggtg	gatcatgagg	tcaagagttt	gagaccatcc	tggccaacat	ggtgaaatcc	12720
catctctact	aaaaatacaa	aaattagctg	ggcatggtga	cacatgcctg	tagtcccagc	12780
tactcaggag	gctgaggcaa	gagaagtgtc	tgaacctgga	agggtggagg	tgcaagtgagc	12840
cgagatcgtg	ccactacact	ccagcctggg	cggcagagca	agactccgtc	ttgaaaaaaa	12900
aaagtaggtc	atthttggct	gggcacgggtg	gctcatgcct	gtaattccag	cactttggga	12960
ggctgaggtg	ggtggattgc	ttgagcccag	gagthttaga	ccagcctggg	caacatagtg	13020
aaacctgtc	tttgtagaaa	atacaaagat	tagctaggcg	cagtgccaaa	tgctgtagtg	13080
cccagctact	tggggggctg	aggtaggagg	atcacttgag	ctcaggttgt	tcaggctgca	13140
atgagctgag	atcgtgcaac	tgcaactccga	actgggtgac	aggagtaaaa	ctgtctcaaa	13200
aaaaaaaaaa	aaaaaaaaacca	aaaaaaaaaaa	aactgtttta	attgttttat	ttaggaaagag	13260
aaagtcagaa	catgcaagga	aatthttttt	atthtthtat	ttttgagacg	gagctctcgt	13320
cagttgccca	ggctggagtg	caatggtatg	atcttggctc	actgcaacct	ctgcctccc	13380
gattcaagcg	atthtctcctg	ctcagcctcc	tgagcagctg	ggattacagg	tgatgtccac	13440
cacgcccagc	taatthttgt	gtthtttagta	gagatggggt	tccaccatgt	tggccaggct	13500
ggtthtcgagc	tcctgacctc	aaagtgaaccg	ccctccttgg	cctcccaaag	tgctgggatt	13560
acaggtgtga	gcccgggtga	ccgaccacaa	ggaaatthta	gttaacactg	ttggttgatg	13620
ggagttggga	ggtaggataa	aaggagaaat	taaggaaaa	ctaggcatga	aaaataaaa	13680
accctgagct	cttagatttg	aagaaatagc	agttccatgt	gaggaataag	tggagaagaat	13740
agaattcaga	cctcaggtct	caggctgggtg	acttaaatct	ttcagtatca	catatatgaa	13800
tatatctatt	atacttaaat	cctcctaaac	atthttatth	ttcagttgga	tatatthaat	13860
atataaaaa	aattatthta	tttatthtag	agacagtgth	tcactctctt	acccaagctg	13920

gagtggatg	tgatcatagc	tcactgcagc	ctcaaggctc	attcctaggc	ttaagtgatc	13980
ctctttttt	tttttttgc	agagatagga	tcttgttatg	ttagccaggc	tggagaaatt	14040
tcaacaatat	tttgaacaat	aaaaaaaaa	taaattaggt	tttattgtaa	agtggatgt	14100
tttaagtttac	gccattctcc	tgcgtaatgc	atatttcata	ctcttcctac	tgataatggt	14160
tctgttcaca	actttttctt	tatattttga	tttctttct	cttttctttt	ttttttttt	14220
tgagacagag	cctcgccttg	tcgcccaggc	tagagtgcag	tggcgcgac	tccgctcact	14280
gctagctccg	cctcccgggt	tcacgccatt	ctcctgcctc	agcttcccga	gtagctggga	14340
ctacaggcgc	ccgccaccat	gcccagctaa	ttttttgtat	tttttagtaga	gacggggttt	14400
caactgtgta	gccaggatgg	tctcgatctc	ctgacctcgt	gatccaccog	cctcggcctc	14460
ccaaagtgtc	gggattacag	gcgtgagcca	ccgcgcctgg	cacatatttt	gatttcta	14520
gtggacatca	gaatgggctg	ttgacagctc	ttttttttt	ttttgaggca	gagtctcgt	14580
tcttcaccca	ggctggagtg	caatggcgcg	atctcagctc	actgcaacct	ccacctcca	14640
ggttcaagcg	attctcctgc	ctcagcctcc	ccagtaactg	ggattacagg	catgtgccac	14700
cccacaccca	gccaattttt	gtatttttag	tagagacagc	atttcaccat	gttggccagg	14760
ctgtttctga	actcctggcc	tcaagtgatc	gcccagcctt	ggcctcacia	actgtagga	14820
ttacaggtgt	gagccaccgt	gcccagcctt	gacaggtctt	tagtttgatt	ttagttcaac	14880
aactgatgcc	gtaatatgcc	aaattaaatt	agttcagact	gaaacggatt	acttaaagat	14940
tcattttct	ttaaaaatga	agtaaaactt	tagccggatg	tgggtggcggg	tgtgtgta	15000
cccagctact	cgggaggggtg	aggcaggaga	atcgcttgaa	cccaggagg	ggagggtgca	15060
gtgagctgag	atctcaccga	ttgcactcct	gctgggtga	gaagagtga	actccatcaa	15120
aaaaaaaaa	aaagcctggg	catggtggct	caagccgggt	gcagtgccc	atgcctgtta	15180
tcccagcaact	ttgggaggca	gaggcaggcg	gatcacttgt	agtcagaagt	tgcgcaccag	15240
cctggccaac	atattagtgga	aaccctcct	ctactaaaaa	tacaaaaatt	agctgggcat	15300
agtgggtgca	gcctgtagtc	ccgggtactt	gggacgctga	cgcaggacaa	ttgcttgaa	15360
tccagaggtg	gagattgcgg	tgggccaaga	tcacgccact	gcacgccagc	ctgggcaaca	15420
gagcaagact	ccatctctca	aaaaaaaaa	aaaagaaaa	ctgagtttat	attgttatgg	15480
ttcttatcta	tcttaaat	tttcttttagg	agattgaata	tttttgaact	ttagtcttaa	15540
tccagtggtt	gaaagagagt	gcacttttga	agtctgcctc	ttggctgtcc	ttgacaacac	15600
aaaccttagt	tccaagagaa	tgtaattctt	cctctttctc	agtgtctcaa	aatatata	15660
atcaaataga	aacgagttga	ataggcagtc	tcttcaaagg	tttccctact	ctgtgggtta	15720
ctatcaagga	gctggttaata	tcatgcactg	ccattccctt	ggcaacatga	cttatctttt	15780
ttttttttt	ttctttttt	agacagagtc	ttgtctgtc	accaggctg	gagtgcagt	15840
gcgagatctc	agctcgtctc	aagccccgcc	tcccacgttc	acgccatttt	cctgcctcag	15900
cctcccgaagt	agctgggact	acagggcctt	gccatcacac	ctggctaatt	ttttttgtat	15960
tttttagtag	agacggggtt	tcaccgtgtt	agccagtag	gtctcgatct	cctgtcctcg	16020
tgatccgccc	accttggcct	cccaaagtgt	tgggattaca	ggtgtgagcc	actgocctg	16080
gcctacttat	cttctaattt	aactgaaaac	caatttattt	gattcagtga	aatggcatca	16140
aactgtagta	gtgttaattg	aaatatttgg	taccttgaaa	tgtaaatgc	caaatata	16200
ctcattttaa	tgaaatctcg	tgtaaatgtg	ttttatag	tgactatgtt	tattctgaat	16260
tttattctta	tggcatacta	aaaaaaaaa	aaattttttt	tttttttgt	aatggagctc	16320
tgctctgtca	cccacgctgg	agtgcagtg	cgcaatctcg	gttcaactga	acctccacct	16380
cccgggttca	agcgattctc	ctgcctcagt	ctcccagta	gctggtaacta	caggcgtgca	16440
ccactatgcc	tggctaata	ttttgtattt	ttaggagaga	cagggttca	ccatgttggt	16500
caggctgatc	togaactcct	gacctgaagt	gatccgcctg	cctcagcctc	ccaaagtgtc	16560
gggattacag	gtatgaccca	ctgcaaccaa	cccatactca	aatttgacac	tgaattttca	16620
taaaggcctt	aatttatctg	aaaccaaac	atttcaaaag	aggaatagca	cagcaaatc	16680
tgttgactta	atgtagaggt	atgtgaagtc	tatttattaa	agcaaatatt	aatggaggc	16740
cagttaat	gtacagctct	gcattttaga	tatttgagaa	atatttattt	cctctccagt	16800
gagatgtgtt	aaaacattag	ttatgtgatt	aacaaatag	tgtacatacg	tatatatgta	16860
catacacatt	ttgagacagg	gccttggctc	gttgttcagg	ctggagtgca	atgacacat	16920
cttagctcat	tgaagcttca	gcctcgcagg	ctcaatogat	ccacccacct	cagcctccct	16980
agtagctggg	tctacagaca	tataccacca	tgcttggctg	atttttta	ttttttaga	17040
gatgggtgatc	ttgccctgtt	gactaagttg	gtaaatattt	taattgttga	actttcttgg	17100
aagactgaaa	acctgtgata	gccattttat	ataaggagaa	gctgaggttc	aaagagtga	17160
ctcatagcag	aaacaaaaat	agaatttaag	tgaatggact	caaaatattg	tactttttac	17220
tttatactgc	aggtttttat	gttgtaatgc	tggttaag	ctccttggaa	tatttggagg	17280
aaaagagaag	gttgtataaa	tggttctttg	gatttactag	aacatcat	gttctgcagt	17340
gctcctgtgg	gtagacaagc	cggaaatctc	ctgggtaaca	caatgggtga	ggttctctag	17400
gtgacgtttg	atttctcaag	tacataggac	taaacagaaa	aggcctagta	tgttatatga	17460

atgagagatc	aagtttctca	ggatattcta	gggctaaggg	atcaggcatc	gaagacagaa	17520
attgtctaaa	taaaatTTTT	tctattcata	gTTTTaaagg	gctaaagggT	caggcattga	17580
agacagaaat	tgtctaagta	aaatatTTTT	ctgttcatag	TTTTaagctg	tgtatatgtg	17640
catgtgtgta	TTTTaaataa	cttctgcaat	gaaaagagtc	aaactctgta	aaatatTTga	17700
agagatttat	tctgagccaa	atatgagtga	ccagtggccc	atgacacagc	cccagtagat	17760
actaagaaca	tctgtccaag	gtggtcaggg	tatagcttga	TTTTatacac	TTtagggaga	17820
cataagacgt	cagTTaaaca	tgtaaagtgt	acattggTtc	catctggaaa	ggcaggaaaa	17880
ctagaagttg	gggaggcttc	caggtcgtag	gcagattcaa	agattttctg	attggcaatt	17940
ggtcaaaaaga	gcttatctaa	agtcctggaa	tccatagaag	ggagtgtctg	gTTTTaaata	18000
ataggTTgtg	gctaccaagg	TTTTtatTgt	acagattgaag	cctccaggta	gcaggcttca	18060
gagagaatac	attgtaaatg	TTtcttatga	gactTTaaaa	ggTggcagac	tcttaagTTa	18120
atTTTTctcct	ggTtcaggta	aaagactTgg	aaagggaag	gattctctac	agaacgTaaa	18180
TTTTccccac	aagagaaagc	TTTgcagggc	catttccagaa	tatgtcaaaG	aaataataatt	18240
tagggTaaaa	tacttcaatt	TgtTTTTattt	atTTattTtat	TTTTtgagac	agagtctcgc	18300
tctgtTgccc	aggctggagt	gcagtggcat	gatctcggct	cactgcaagc	tccgcctcct	18360
gtgttccagc	cgTtctcctg	cctcagcctc	ctgagtagct	gggactacag	gcgctggcca	18420
ccatgcccag	ctatTTTTttt	TgtatTTTTa	gtagagacgg	gatttcca	TgtTggccag	18480
gatggTctcg	atctcttgac	ctcgtgatTc	acccgcctcg	gcctcccaaa	gtgctgggat	18540
tacaggcgtg	agccactgtg	cccggcctca	aaataactca	atTTctTTtca	Tggcctgcta	18600
tctgacgtga	Tgtctgacta	gagTcagggc	gggaattTgg	cgtcttattg	ctacaaaaca	18660
tcttaataTc	tctgtTTTTaa	TgtTaatgct	gatcagTtgt	ccctgaattc	caaagggaag	18720
agggtatatg	agtcatgtcc	aaccccact	tctcattatg	gcctgaacta	gTTTTttagg	18780
TTaactTTTT	aatgcctTtg	gcaaggggag	ggTccatgag	tcagtTgggg	gtcttagagt	18840
TTtattTTTT	tctTaccggt	TTataaaaaag	TTaatgaaaa	TTatcatatt	tcataattct	18900
atataattca	atattgtact	TTatTTaaaa	ctcacgtata	aaatagctgt	ccatattctgt	18960
TTtcagaaga	Tgaggatgga	gggtagaaat	cagaagtgtc	agattTggta	atTTtctTac	19020
actgctgaaa	acctatacct	acctctTtga	aaggattaat	TtcagactTg	ctTtctTtgg	19080
gcctaattgat	tctactTTga	agTTtctctg	TTtaaaactaa	ggaataaatc	Tgataaatgg	19140
acattccagat	gataccatac	TTTTccaaaa	gataacattg	ctTTtgatta	catatgcaat	19200
aaacatttca	cattTTTTct	caagattatt	tactggcatc	Tgaccaaaag	acacaaaaaa	19260
gcagccactg	TTaaggactt	TatccctTgt	tctgtTTTTa	gctggTTTgt	TgtTgtattt	19320
Ttctctcatg	Tgaatacagt	Taaaccctat	Taaactggat	tccccatatt	actgttagTt	19380
gtcctgatga	caattaaagga	TagttagTaa	atggatattg	aatcattTTa	TTTTtTtagt	19440
agTgatggcg	TTtcgctgtg	TtgGCCaggt	TggTctcgaa	ctcctggcct	caagtgatct	19500
gcctgtgtTg	gcccctcaaa	gtgttaggat	Tacaggcgtg	agccactgca	cctggctgga	19560
Tattgaatta	Ttgaaagaag	atgttccta	acacttctct	cctctTTtcc	gctgtctTac	19620
tctctctgtt	ccaccocgta	gtagtTgggt	TTTTgatTc	Tggcaaaagta	tagattTct	19680
actctcttat	attagTTTTc	TtcaactTtt	TTTTtaaaaa	TgtTTTTaaa	gcaagataga	19740
gacagggTct	Tgccgtattg	cccaggctgg	tctTgaactc	ctgagctcaa	gcaatcctcc	19800
cacctcacc	cataaagtgc	ggagattaca	ggcatgagcc	accatgctctg	gccagTTTTc	19860
Ttctactctg	acctataata	gtcctgcaaa	gccagtgaag	ctgttaatat	gctgacgtag	19920
ctctTTTTtc	tctcgTTTTgt	gaattattaa	caatcgctga	tcttacacat	catatacaat	19980
aaaacatctt	Tgaacactgt	cacatctcag	cagctcattc	TggtTaatga	ggaaagaaaa	20040
atgtcaaaat	ctgtgatTTt	cttagaggTt	attaaatgtt	Ttacagctat	gtagatattc	20100
TgttagactTg	ctgtactTtac	Ttacataaact	TTtctgctct	tctgcagagg	gagagaatta	20160
acttccataag	TgggctTTgt	caatgcccctc	cagTcagcga	tctccaggaa	caaaactTatc	20220
TatggtTgag	caagtggggT	Ttattaccca	ctgtagccag	ggagaacaca	catatgTaat	20280
aaccatggta	Tgtcatagta	aagggtgtTta	ggaaagatagg	attTgggctc	gTttattTgg	20340
TgatTTTaaG	gagggtTTTaa	agaagcaggg	TTTgtctctg	attTggatgc	Taccaggagc	20400
TgCGaataat	tctatgaata	aaatbcataa	aacctatcta	gaagaaaaga	ctagagTgag	20460
gtctaagct	gtagtagTTa	aaaagcaaaG	gtcactcctt	atctggaaaa	ggggaatatt	20520
TggtattTca	TggTttagac	agtgtTcagT	gttcatgTtt	Tgcctgtgtt	tagacataat	20580
TgttagagTgG	tctTgtTcca	ccgtggTcac	agagtggcat	TattTcatgc	Tgatattctg	20640
TaagTgccag	gtcaaatacc	aaggTttagc	TgatagtagT	aggccagctc	ctggatgTaa	20700
TaagctgtTt	TtctctTtct	cagctTTTTt	TTTTgtTgtt	TTTTTTTTaa	aatagagTct	20760
ggctctgtca	cccaggctgg	agTtCagTgg	gggcaatctc	agctcactgc	agcctctgcc	20820
Tcctggattc	aagcgaactc	cctgcctcag	cctcccaact	agctggaaTt	acaggcgcaa	20880
gtcaccacac	ccggctaatt	TTTgtattTt	tagtagagac	gggTttccac	catgtTgagc	20940
aggctgtct	cgaactccgg	acctcaagca	atccatctgc	ctcagcctcc	caaagtgcTg	21000

agattacagg	catgagtcac	tgcacctggc	ctgtttctta	gatttgaggg	tcaactttta	21060
cccttttcga	actgtggggc	ttcttatgga	aattgacatt	taagtcctga	ccatatagga	21120
tottgggcaa	gttacttaac	tatgtatgct	gagatgtttt	tttaaatgct	tagtgcttgc	21180
acataatagg	tactcagtaa	atgagaacta	ttattataaa	atcaatagta	cttttaagat	21240
tacagctaga	ggttatgtta	gaacattggt	aactctctgt	actagtttct	gttacacttt	21300
ttaggatgca	gtagcagcat	aaacgagata	tggggaagaa	cgagaggtat	ttaaacagtg	21360
atataggccg	ggcgtggtgg	ctcacgcttg	taatcccagc	actttgggag	gctgaggtgg	21420
gcagattacg	aggtcaggag	attgagacca	tcctggctaa	cacggtgaaa	ccccatctct	21480
actaaaaata	caaaaacaaa	attagccggg	cgtggtagtg	ggcgcctgta	atcccagcta	21540
tttgggaggg	tgaggcagga	gagtgccgag	aaocccaggag	acagagcttg	cagtgagccg	21600
agatcgcgcc	actgcacccc	agcctggggc	acagagcgag	actctgtctc	aaaaaaaaaa	21660
aaaaaaaaaa	attaaaacaa	aaatatttgt	gttaattgtg	atgacaaaaa	aaaaaaaaaga	21720
gatggaagtc	tctccctaac	ctcactcctc	atthagtgctc	atggcttttt	tctttttttt	21780
tttttttttt	ttgagacaaa	gcctcactct	gtcaccaggy	ctggactgca	gtgggtgcaat	21840
ctcagctcac	tgcaggctct	gcctcccaga	ttccaagcaat	tctcctgcct	cagcctcaca	21900
agtagatggg	actacaggca	catggcacca	tgcccagcta	atTTTTtTgTg	tagtttttag	21960
tagagacagg	gtttcactat	gttggccagg	ctggctctga	actcctgagc	tcaagtgatc	22020
ccccctcctt	ggcctcccaa	agtgetggga	ttacaggcgt	gagccctgct	cccagactcc	22080
tggctttttt	tttttttttt	aatgaaaaat	tcaaaatgct	cttttttttt	tttttttttt	22140
tttgagacag	agctctgctc	tgtcgcccag	gctgaagtgc	agtgggtgga	tctcggtcca	22200
ctgcaacctc	tgccctccag	gttgaagtga	ttctcctgcc	tcagcctctc	gaacagctga	22260
gattacaggt	acgcgccacc	acgcctggct	aatTTTTtTgT	tttttaatag	agatgggggt	22320
ttgtcatggt	gtccaggcta	gccttgaact	cctcgectcg	tgatgaccac	ctgccttggc	22380
ctccccaaat	gctaggatta	taggcgtgag	tcgctgcacc	tggccacaga	aatTTTTtTgT	22440
agaagataaaa	taaggtgaca	tttttaaggg	tcaaagaaaa	tgtcaaaaac	tagaatgatg	22500
tctttacata	gggtttaaaa	ctttccaaat	taacagggaa	aataattctt	tacctgaaa	22560
ataaatgttt	gctagtgaaa	gcaaatacaa	tctttttact	aatgttttta	ttaaattttt	22620
ttttcttgta	gacacagggg	cctactgtat	taccaggcct	ggtcttgaat	gcttggcctc	22680
aaagcaatcct	cctacctcag	tcttccaaaag	tgctgggact	acagacatga	gccatcacac	22740
tgggtctttt	ttaccaaatt	atagtagaaa	gcactttttc	tctaattggtg	aactatgaga	22800
gaattaatca	ggggctatta	gtaattcatc	cctgaattaa	tcagtattaa	taatgctttg	22860
tggctccatgt	agtttgctgg	ggattaacac	acatgaaaag	tctaccagga	gatttttttt	22920
tttctttgag	aacagggggc	ataatcagta	gtccttaaat	gaaatggact	attcccattt	22980
cattatattgt	tgccataggct	ggactcgagc	tccttggctt	aagtgatctt	cccacttcag	23040
actctcaagc	agctgggact	ataggtgtgt	gccaccatac	ctggcttaat	tgagtgtttg	23100
ttttttgttt	ttttttctga	gacagagtct	cgttgtgtcg	ccaggctag	agtgcagtgg	23160
cacaatctcg	gctcactgca	acctctgttt	ccaggttca	aacgattctc	ctgcctcagc	23220
ctcccagaatg	gctgggatta	caggcgcctg	ccaccatgcc	tggctaattt	ttgtattttt	23280
agtagagatg	gggtttcacc	atcttggcca	ggctgatctc	gaactcttga	cctcatgatc	23340
cacacacctt	ggcctcccaa	agtgtctggga	ttataggcgt	gagccaccgc	gcccggcttt	23400
aattgagatt	tttagatatc	tattactctg	ctaattttgt	cacttgcaag	ttgccatcag	23460
aaaattgtag	gaaaatggat	atatttggtt	cttggaatgg	tt:gtgtgag	aatacttaag	23520
gattaaatag	ataagtaaaa	ctgggtggct	ttatataaca	tagatgagca	aatgtcagga	23580
acatacaact	gtgcacacag	ttcaggagaa	ggaggattta	agttaatcaa	caaatttact	23640
aagtataata	aagatactaa	aagttagtgt	tccataccac	tttattactt	aaagtatcat	23700
catataccct	attttatatg	atttttgcca	caagtacag	ttaggtaaaa	gaaactttg	23760
cttttcagggt	aaggagtttg	acgcccagac	agathtaactg	acttttccaa	aatcatattg	23820
ctattaaatg	gtggaacaag	gacttaaatc	tttgcttctc	aactcacata	cttgcaaaaa	23880
catatcctct	cactctaccc	caagctaccc	atgttttgac	ccttcttctg	gcaatctggg	23940
tctcactaat	atttgaaaga	aaacgtacag	tagataattt	gcaagttaat	ctgttacgca	24000
tatctcttac	ctctatttaa	agatgaatat	cagcatttct	gttgtttcta	cagtaacata	24060
ctaaaaaata	gtcagtcoca	ggtgcagtgg	ctcacacgtg	taatcctagc	acttttggaa	24120
gttgagcgag	gaggatcact	tgaagccagg	agttogagac	tagcctgggt	atgcaagacc	24180
ccattttctgt	tttttttttt	tgtttgtttt	gttttgtttt	gttttgtttt	tttgagacgg	24240
agtctcgtct	tgtcgcccag	gctggagtgc	agtggcggga	tctoggctca	ctgcaagctc	24300
cgectccggg	gttccagcca	ttctcctgcc	tcagcctccc	aagttagctgg	gactacaggc	24360
gcccgcact	acgcccggct	aatTTTTtTgT	atTTTTtagTg	gagacggggg	ttcaccgttt	24420
tagccgggat	ggfctcgatc	tcctgacctc	gtgatcccca	tttctttttc	actgcaacct	24480
ttgcctccca	ggttcaagtg	attgtcgtgc	ctcagcctcc	caagttagctt	gggattacag	24540

gcatgtgcca	ccatgcctgg	ctaatttttg	tattttttgg	tagagatgga	gtttctggcc	24600
aaagtggcca	ggctgggtctc	gaacgcctag	octcaagtga	tctaccctcc	ttagcctccc	24660
aaagtgtctgg	gattacaggt	gtgagccact	gtgcccggcc	ccatttctac	aacaattaa	24720
aaatattagc	ccagtgtagt	ggtgcatgtc	tctagtccca	gctaetcaga	aggctgaagt	24780
gaaaggattg	cttgagccca	gaatttcaag	gctacagtga	gctatgataa	tggcattgca	24840
ctccagcttg	ggtgacagag	tgagaccctg	tctctaataaa	atgaagtata	atagtgcaca	24900
agtatagaac	ttgaaaatct	tccttaacct	taccataagg	gaaatgatta	ctaataagtt	24960
tcttaacttt	ttgtacttac	ataaacataa	atattcatca	gagaaaaaaa	tatgcaaac	25020
aatgtgcaat	ctttttcact	taccatattt	tggaaattttt	ttcatttcaa	tatatttgat	25080
cttccttgtg	ttttttcagt	ttgtttgttt	ttgtcaccca	ggttggagtg	aagtggtaag	25140
aacatggcct	atgacagcct	caatctcctg	ctcccattca	gcccctcaag	tagctgggac	25200
tacaggtaca	tgtcaccacg	cccggctaag	ttttattttt	atthttggtag	agatgggggt	25260
tcaccatggt	gcccaggctg	atcttgaatt	cctgggctca	agtgaaccgc	ccacctcagc	25320
ctcccgaagt	gttgagagaa	caggtgtgaa	ccacatgct	ccacctctta	gtctttacaa	25380
tctgcaaaac	ctcataagtg	gctaatagag	gaataagta	aagcaaaagg	ggatctcact	25440
gattagaact	gtgtttttag	gctgggtcgg	gtggctcacg	cctgtaatcg	caacattttg	25500
ggaggctaag	tgggagatc	acttgagccc	gggagttcaa	gaccagcctg	ggcaatatag	25560
tgagagacc	tgtctttaga	aaaaaattaa	ccaggtgtgg	tgggtcacac	ctgtggtccc	25620
agctattcaa	gaggctgagg	tgggagaatc	gcttaagcct	aggaggcggg	ggttgcaagt	25680
agatcatacc	actgctctct	agcatgggtg	acagagcgag	accaggtcta	aaaaaaaag	25740
tattttttcg	ttttttcca	actcatgtac	acccgccacc	ccaccctcgc	ttttttttt	25800
tctgacattg	ggtcttctc	tgtcaccacg	gctagagtgc	agtagcacia	tcaactcact	25860
gcagcctcct	cttcctggac	tcagcagctc	ctaccaccctc	agcctcccaa	gtagctggga	25920
ccacaggtgt	gcaccatcat	gcctggctaa	ttttgtact	ttttgtagag	atagggtttc	25980
accatgtgg	tcaccatctc	taactcctgg	cctcagtcag	tccacttttg	cctcggcatg	26040
agccactgtg	tgagcccac	gttttttatt	aatggatatt	tggattgttt	ccatctattg	26100
tgaataatgt	ggctatgaac	attgggtctaa	atatctgttt	aagtcccggc	tttcaatact	26160
tttgatata	taccatggag	tagaattact	gaattatatg	gtaactttct	gtttaaattt	26220
ttgaaactg	aacctgtttt	ccataggggc	tgaccatttt	tgcatcccca	ccagcaggtg	26280
acaagggttc	cagtttctcc	acatttggtta	tttttcattt	tttaaaataa	tagtcatcct	26340
aaagggtatg	aagtggatc	tcgttgtgat	tttgatttgc	atgtattttt	ctaataactt	26400
atgatgctca	gcattttgtc	atgtacttat	gtaccatttg	tgtatcttct	ttggaaaaat	26460
gtctattaat	gttcttttcc	cattttttaa	ttgggttgtt	tttatgttta	tcaattttgt	26520
aaacatttta	agctctgcag	cataactact	caacctgtc	acatggtaag	attgaccag	26580
taaaacttta	tgtacaaaaa	taggcagctt	actagattta	atcttagtcc	atagtttgct	26640
aaggcatgca	ttagataatg	tagttacact	attggctaatt	aattttaaact	acaagtggtt	26700
gtaagtttct	gcccccacia	ttctttcttg	atttgatgta	gtctgggtgg	ttgaatttga	26760
gtgttatatg	gggtcataga	gtaaaagag	aaatgtctat	gagaaaactag	ggactgttgg	26820
gagctaagt	taaaaggattt	tggaggcctt	ttgtgactg	gagaccattg	gaagattgga	26880
ttttctgcta	taaatgtatc	taaaggataa	tcagtgtgag	ttatgggctg	tagtttgcca	26940
accctacat	taagggattg	aaataaattg	aatatgggtt	tcagttcttg	tatggctcgg	27000
ctcagttctt	ttttttttt	ttttttttt	ttttttttt	agatggagtc	tcactctggt	27060
gcccaggctg	cagtgccatg	gcacgatctc	agctcgtctc	aacctccgcc	tcctggcttc	27120
aagtgattct	cctgcctcag	cctccagagt	agctaggact	acaggcacgt	gccaccacat	27180
ctggctaatt	ttttttttt	tcaagacaga	gtttcgctct	tgtcaccacg	cctggagtg	27240
aatggcacga	tcttggctca	ccgcagcctc	cgctcctgg	gttcaagtga	ttctcctgcc	27300
tcagcctccc	aaatagctga	ggcaacaggc	gtgcgccacc	acgcctagct	aattttctg	27360
tttttagtag	agatggggtt	tcaccatctt	ggccaggctg	gtctcgaact	cctgatctca	27420
tgatccaccc	gcctcagcct	cccaaagtgc	tgggattaca	ggcgtgagcc	acctcacctg	27480
gctgtttttt	gtgtgtgttt	tcagttagaga	tgggcttca	ctgtattagc	cagaatggtc	27540
ttgatctct	gacctcgtga	ttcacctgcc	tcagcctccc	aaagtgtctg	aattatagge	27600
atgagccace	acgcctggct	gtggctcagt	tcttaactgt	tcattcattc	aggttcccag	27660
ccaaaatctt	ttgccacagc	ttctcctggg	catgctctga	aatcattttt	ttgtctgttt	27720
agttgcatga	gtccactgaa	agctttgctt	atthttctcca	ccactttcag	agttcacagt	27780
cacttcaagt	agacgaggtc	aggagatcga	gaccatcctg	gccaacatgg	tgaaaccccg	27840
tctctactaa	aaatacaaaa	cttagctggg	cggtgtagcg	tgtgectgta	gtcccagcta	27900
ctcaggaggc	tgaggcagga	aaattgcttg	aaccggagag	gcagaggttg	cagtaagccg	27960
agatcatacc	atthcactcc	agcctgggcg	acagagtggg	acactgtcaa	aaaaaaaaa	28020
aaaaaaaaaa	aggaagaaa	agtgctccaa	atgctgggct	catctctgct	ggcttctctc	28080

ctttttatat	ttcgataaatt	tttaaaataa	taaccatcct	aatgggcttg	atattggctcc	28140
tcctttctaa	attttggtaa	ttttcattgt	tttggaaact	cactaatccc	ttcaaacaga	28200
tgtttttaaa	acttttttcc	tagttctctc	tctttttttt	tttttttttt	tttttgagac	28260
agggctctcc	tcagtggtgc	gacctccctt	actgcaacct	ctccctcctg	ggttcagca	28320
gogattctcc	tgctcagcc	tcccaaatag	ctgggagtac	aggcaccac	ccccacgccc	28380
agctaattgt	tgtattttta	gtagagattg	ggtttcatca	tgttggccag	gctggtctcg	28440
aaactcctgac	ccaggtgat	cgcctgcct	cagcctcca	aagtgcctggg	attacagcca	28500
tgagccaccg	cacctggctc	tagttctctt	ctaactttac	tatagtaatc	aaaattttag	28560
tgctgatggt	ataatccaag	cccaaatgga	tattcttata	cattaaatgt	tggaatatca	28620
tgtttgttaa	aaatcagatc	tgctgggcac	ggtggctcaa	tgctgttat	cccagcactt	28680
tgggaggccg	aggcaggtgg	atcccctgag	atcaggagtt	cgagaccagc	otgaccaaca	28740
tggagaaaact	ccgtctctac	taaaaataca	aaattagccg	ggcgtggtgg	cacgtgcctt	28800
tcattcccagc	tactgggag	gctgaggcag	aagaattgct	tgaacctgtg	aggcggaggt	28860
tccagggagc	tgagatcgtg	ccattgtact	ccagcctggg	caacaagagc	gaaatccat	28920
ctcaaaaaaa	aaaaatcaga	tctgtcccta	tggttttatc	ttcactgcat	gtcatatata	28980
aatggaagca	tgtagtatgt	agccttttgt	gtctcgcttc	tctccctttg	cataatattt	29040
ttgagattta	tcacttttcc	tacatatgtc	agtagtttgt	ttctttttaa	tgctgaatag	29100
tattccacgt	gtggagatac	cacaatttgt	tactccattc	actaaatgat	ttgggttgtt	29160
tcctgttttt	gttattgtttg	ttgttgttgt	ttaatattt	aattaatttt	ttttttgaga	29220
tggagtcttg	ctctgttgcc	caggctagag	tgcagtggcg	caatcacagc	tcactgcaac	29280
ctctgcctcc	ccggttcaag	gcctcagcct	gcctcagcct	cccaagttagc	tgtgattaca	29340
ggtgcccgcc	accacaacca	gctaattttt	gtatttttag	tagagatggg	atctcaccat	29400
gttggccagg	ctggtcttga	actcttccct	ccttggcttc	ccaaaatgct	gggattacag	29460
gtgtgagcca	ccatgccccat	cctcctgttt	ttaaatttta	tgaataaagg	tactgtaaat	29520
tcagtacag	gtcttagaag	tttgcatttt	tcttgggtaa	atatgtagaa	gaagagattg	29580
ttgactcatg	tgtaaatat	atgtttaatg	tcataagaaa	ctaccaaact	gtttttccaa	29640
ctgggtgcca	ttttgtttcc	taccagcaaa	cataatgaaa	ggtccattac	cctttgtctc	29700
gtaaatatttt	ggtattttta	tgttcttttg	tttttttggc	attcaaaatt	gggtgcattg	29760
ttgtatcttt	ggatttttaat	ttgtacttcc	ctaacaacta	atgatgtgga	gcatcttttt	29820
atacgtttac	ttgccaatta	tgctcttaatt	cctttggaat	acttatattt	gactctgcct	29880
actatatata	tatatatata	tatttttatg	ctatatatag	agratatata	tttttatact	29940
atataataatt	atataatact	tataataatt	tataatact	atataataat	atataataat	30000
atattttttt	ttttttttga	gatgaagtca	tgctctgttg	accaggctgg	atgcagtggc	30060
atgatcttgg	cttacccgaa	cctcaacctc	gcctcccag	gttcaagtga	ttctcctgcc	30120
tcagcttcct	gagttagctg	gactacagac	atgcgccacc	atgcccggct	aatttttgta	30180
tttttagtag	aggtggagtt	tcgccatggt	gtccaggctg	gtcttcaact	cctgacctca	30240
gggtgatccac	ccgcttcagc	ctcccaaaaag	tgctgggatt	acaggcatga	cccaccgtgc	30300
ccagctggaa	taagcaatct	taaaaagcag	tttgttgatt	tctggtgaat	agagaaaatg	30360
tgacaaaact	tgactcaata	gtttcatttc	ttttttgttt	gtttgtttgt	ttttttgaga	30420
cagagtcttg	ctctgtcacc	caggctggcg	tgcagtggct	caactgcaagc	tctgcttcct	30480
gggttcacgc	cattctcctg	cctcagcctc	cggagttagct	gggactacag	gcgcccgcca	30540
ccatgcctgg	ctaattttgtt	ttttgtattt	ttagtagaga	cagggtttca	ctgtgttagc	30600
caggatggtc	tcgatctcct	gacctcgtga	tccaccctgc	tcggcctccc	aaagtgcctg	30660
gattacagcc	atgagccacc	gccccctggc	tacttttatt	tcttaagcat	atatcttaaa	30720
gaaactgcag	gtatgatta	ggggacaggt	agaaaaatgt	tacagcattg	tttgaatgt	30780
tgaaaaactg	gaaacaacct	caatatctta	ataatagaaa	aaatggttga	ggctggctgt	30840
gggtggtcat	gcctgtaatc	ctagcacttt	tggaggctga	ggcgggcgga	tcacctgag	30900
tcgggagttt	gagtcagcc	tggccaacat	ggtgaaaccc	cgtctctact	aaaaatacaa	30960
aaaattagcc	aggtgtggg	gtgcacacct	gtaatcccag	ctacttggga	ggctgaggca	31020
caagaatcac	ttgagcctgg	gaaatggagg	ttgtagttag	ccagatcat	gcctactgca	31080
ctccagcctg	ggcaatagag	tgagactccg	tctcaaaaaa	aaaaaaaaaa	aaaaaaaaagg	31140
aaggagcatg	ggaatgggaa	gaggcactaa	aaaaggaaact	tcttttgtat	ctgttacatt	31200
gfactcattt	tgtttttaaa	attttaagcc	aaatatgaca	ttgttatcac	ttttgttgcc	31260
actcagaata	caagtgtctat	tacactattt	tctataattt	tccctttttc	aaactaaaaa	31320
acaaaaatac	atgtaaacag	gtgttttctt	tatgaaaaac	aggaaggtga	gcatccaaat	31380
aatttttagtt	aattttataac	aaactcttaa	taatttcttt	ttcttctctt	tctttttttt	31440
tttttttttt	tgagtcaggg	ttcttgcctt	ttgctcaggt	tagagtgcaa	tggtacaatc	31500
tggtacaatc	acatcttagt	gcaacttcta	cctcttcagc	tcaagtgacc	tctctcctca	31560
gtccccaag	tagatgggac	cacagatgtg	tatccccaat	catggttaat	ttttaaat	31620

tttgtagag	ataaaatctc	accatgttgc	ccaagctggg	cttgagctcc	tgggctcaag	31680
cagtcctcct	gccttggcct	cccaaagtgc	tgggattata	ggcgtgagca	aactgtgcc	31740
agtcttcata	aacttttctt	tacatgtcct	tatcaagtac	tttttgagca	cctactgtca	31800
accagccgta	gtatataata	ctgataattc	tataacataa	gaaattgacc	tgtttaaggg	31860
gatgaatatg	gtagttatt	ttcaggggtga	agaaacagaa	gaatcgggga	gtagtagcat	31920
agtcataagg	agacggcatt	tcttagtcac	tttgtgtggt	gtttataata	agacttcatg	31980
tggccaggtg	cggtggctcg	tgagcgctcg	taatcccagc	actttgggag	gccgaggttg	32040
gaggatcaca	acaaggtcag	gagatcgaga	ctatcctgaa	caacatggtg	aaaccccgtc	32100
tctactaaaa	atacaaaaat	tagctgggtg	tgggtggcacg	tgctgtaac	cccagctact	32160
cgggaggctg	aggcaggaga	atcgcttga	ccaggagtc	ggagggttga	gtgaaaccaag	32220
atcgtgccac	tgcactgcag	cctgggctac	agagagagac	tccatctaaa	taataaata	32280
aataaataaa	taaatagact	tcgtctgttt	catcaagagt	cattgtatta	tattgatttt	32340
gaaatttcaa	ataataataa	ggaatagtat	ggaaatTTTT	gtttgattaa	atggggacat	32400
gaggccaggt	gtggctcact	aggcttga	tcccagcact	tcgggagacc	agggcaggag	32460
gattacttga	gcctaggagt	tcaagaccag	cctgggcaac	atggcaagac	cctatctcta	32520
caaaagaaat	acacaaaaat	tagctaagtg	tggtggcaca	tgctgtagt	cttagctact	32580
cgggaggcca	agggaaagg	atggatcact	tgagcccagg	agttggaggg	tgcaagtgc	32640
taggattgtg	cccctgcact	ctagtgtggg	tgacagcaag	actctgttta	aaaaaaaaaa	32700
aaaatgaggc	cgggcgcggt	ggctcacctc	tgtaactga	gcactttgag	agggcaaggc	32760
gggtggatca	cctgagttca	agatcagcct	ggccaacatg	gtgaaaacc	cgctcttata	32820
cttaaattac	aaagattagc	tggtgtgat	ggtgcacacc	tgtaaccca	gctactccag	32880
aggctgagac	agcagaatcg	cttgaacccg	ggagggggag	gttgcaatga	gccgagattg	32940
caccactgca	ctccagccta	ggtgacagag	caagactctg	tctcaaaaa	aaaagaaaa	33000
aagtctaaga	agttaaTTTT	cattcagaca	aatgttcaaa	aataatagaa	ataaaaacaaa	33060
aaattaatTT	tcaaatatgt	gatttgagta	taacgatcac	tttacagaag	ttatctaca	33120
tactTTTTga	ttatgctgag	TTTTtagtgc	cttttaaaagc	ttaacttgta	tcaggaatta	33180
atgttttaaa	ctagttgttg	ggttcaaga	aaatagaaat	gtgttattaa	atgccacact	33240
tgtaactctca	tacagcagtt	ctcatgatgc	aatcaattgt	taataaaaa	cttctctag	33300
agaggaggtg	cccagcaatg	aacacattat	atcagatttg	aatatgagat	taacaatgc	33360
TTTTgtttc	atgtttattg	aaatcatatt	tgtatttttc	aaaagtatat	atacttaaaa	33420
atctctaaat	aaaatgtcct	tctgtaatac	agcatttttag	gttataagga	tcaataccat	33480
gggctgataa	gcaaaatagg	actcacggcg	ccaaaatgtt	gataaacatc	atgaccatat	33540
tcccataaat	ggagttacat	atacacatta	caatgataaa	aataactaaga	agttatatag	33600
taatgtttct	cagacttggt	cacaaccctg	TTTTtatgt	attactctc	aaaagcatga	33660
tttgcaatat	ttagtgtact	cttcatacga	tctatacaaa	taatTTTTta	TTTTaaagca	33720
ggacagaagg	agcaagctgt	ggaatggtat	aagaaaggta	tgaaagaact	ggaaaagga	33780
atagctgtta	tagttacagg	acaaggtaa	attgtatttg	tttatagcca	tcccaatta	33840
tgatataattc	acatgattgt	ccagatttca	gatctattta	tttatttatt	tatttttctt	33900
tctttctttt	ctttctttt	ctttttttt	ttttttttg	agacagggtc	tttctctgt	33960
tacccaggct	ggagtgcagt	gatgtgatca	tagctcactg	caacctcaaa	cttctgggct	34020
caagtgatcc	tctgcctct	gcttcccag	tagctgggac	cacaggcgca	cactaccata	34080
cttagctaat	tattattatt	attattatta	ttattattac	tagtttttga	gacagagtct	34140
cgctcttttg	cccaggctgt	agtgcaatgg	tgcaatctcg	gctcactgca	atctccgct	34200
cccggattca	ggctatTTat	tctcctgcct	gagactcctg	agtagctggg	attacaggtg	34260
catgccaccg	tgcccagcta	atTTTTgtat	TTTTtagtga	gatggggttt	tgccaogttg	34320
gccaggctgg	tttogaactc	ctgatctcaa	gtgatctgcc	cacctcagcc	tcccaagtg	34380
gtgggagtac	aggcatgagc	cactgtgggc	cacacttagc	taattaaaa	aaatTTTTt	34440
gtagagacag	ggtctcactg	tgtttctgtg	actgatctta	aacttctggc	ctcaagtgat	34500
cctcctgcct	cagctttcca	aaattgtggg	attacaggca	tgagccactg	taccagcca	34560
gattgtatTT	atcttaataa	atgtatTTgg	gccaggcaca	gtggctcatg	cctgtaatcc	34620
cagcactttg	ggaggccgag	gtgggcgaat	caactgggat	caggagtctg	agaccagcct	34680
gatcaacatg	gagaaacct	gtctctacta	aaaattcaaa	attagccagg	cgtggtgtcg	34740
cgcaacctgta	attgcagcta	cttgggaggc	tgaggcagga	gaattgcttg	gaccagggag	34800
gcgagggttg	tgtaagcca	agatcgcgcc	attgcactcc	agcctgggca	acaagagtg	34860
aactctgtct	cagaaagaaa	gaaaagcaaa	acagtatttc	atttaattaa	aaatgccttg	34920
tattatacca	gtaagagtaa	aaaacaattt	aattaaagta	tgatataatg	tcttttcatc	34980
acttagaatt	tgaatTTtaa	tcttattgaa	gaggctgttt	tagttttatt	tagataggt	35040
tataacaaca	atattaatat	ctgtgaccaa	atgggtgcat	gcatgggcaa	tgaaaagcta	35100
tgttgtgact	gcagcatggt	gtagatgatt	agtaagttag	cattggttgt	aaaatgaaac	35160

ccaatttcaa	agattttgaa	acgtggggaa	aaaatatttt	agaccccatg	aaataagacc	35220
tgaaatctat	gaaataccac	agttaggttg	aaatcatcat	gagaaactgt	aactattttt	35280
ttttataggt	attatatgta	ttctttgaga	cagggtctca	ccctggtgcc	caggctggaa	35340
tgctgtgatg	tggtcacctc	tcactgcaac	atccacttcc	tggggctgaa	gtgatccttg	35400
caactcacc	tcccatgctc	ggctaatttt	tttttttttt	tttttaagta	gagaccagat	35460
ctctttatgt	tgcccaggct	ggtcctgaa	tcctgggctc	aagtgattct	gctttgtctg	35520
cccagagtgc	tgggattaca	gaagtga	acaacacctg	gcctaaagat	agtattttta	35580
aaatttcttc	tgttttcttt	cttttgggta	tacattttct	tctttttttt	aaaaaaaaat	35640
ttctgtataa	agactgtgac	tccccatgaa	agtagtttgg	gtgataattt	atcgtgaaac	35700
aatattagtt	gggaaatgta	gatattttaa	tttaattttt	tctttcaggt	gaacagtgtg	35760
aaagagctag	acgccttcaa	gctaaaatga	tgactaattt	ggttatggcc	aaggaccgct	35820
tacaactctc	aggtatcaat	taatgtataa	ttttagtggg	gatgtattgg	aaatgtgtgt	35880
tcaatgaaac	tttaatttgt	agaaagaaat	agatcagtga	ttgaaaatgt	ggtccaggct	35940
ttttaacata	aagaaaacgt	ataacatata	caaaaacaga	agagcataat	ggacttcttt	36000
ttttttttga	gatggacgct	tgctctgtca	cccaggctgg	agtgcagtgg	cgtgatcttg	36060
gctcactgca	agctctacct	cccaggttca	cgccattctc	ctgcctcagc	ctcctgagta	36120
gctgggacta	caggcaccgc	ccaccaagcc	ccgctaattt	ttttgtagtt	ttagttagaga	36180
cagggtttca	ccatggttagc	cagaatggtg	tccatctcct	gacctcgtga	tccaccgccc	36240
toggcctccc	gaagtgtctgg	agttacaggc	gtgagccaac	gcgcccagcc	gagcataatg	36300
aacttcta	tattcttctc	tcagcttga	taattatcac	ctcactatca	gtcttgattt	36360
atttatacca	agggctagca	aactgcagct	gatgggcca	cccagctcac	tgccatattac	36420
tataaataag	gttttattgg	cacacagcca	ttcctttcct	ttaggtattg	tctatggctt	36480
tcatgctacg	gtgacagaat	tgaatagtgg	atacagagac	aacctgccc	tcaaaaacta	36540
acatattttac	tgctctggccc	tgagaaagcc	tgccatcttc	acttgacatc	ccacttcaact	36600
gtcccccaac	ttcttattat	tttgaaaacaa	attattgact	tattttccat	ttataaata	36660
ttcatattat	acctctaaaa	gataaggatg	tcaaaaataa	atcacagaaa	taccattatc	36720
acacctaaaa	aataactcca	aatttctttt	aacattatcc	aaaaaaactg	accccccaa	36780
aatgcaattt	ccaaatttct	aattttctca	aaatttaata	atttttacta	tttttaata	36840
atgtgtttgt	atcaggatct	taaaaagatc	cgatgtctc	ttttaatgta	caggttcac	36900
tgtttttcat	tttcttccaa	tttatttgg	gatgaaacca	ggtccttctc	atgtaattat	36960
tcctacagtc	tgacttttgc	tgcatgcate	ttttgtggta	tgttttaaca	tgttcttctt	37020
tcctcccacg	ttaaatgata	gttgatata	gagccttgat	cacattgaaa	gttgattttt	37080
atttttta	gattagacta	cttcttaggt	ggttgtgttt	gtgtgttcoat	cattagttgt	37140
tcocatctgt	gaggttagca	gcagttaact	atgactgatg	tccagtgata	ctttaattct	37200
gtcatctctt	cagcattcat	caactgtaat	agaaaacttt	actgtctctc	tgtttgatgg	37260
ctgaggtttt	aaatatata	taataagtta	aatctaccac	tcaaaaaaag	tagaacaaaa	37320
cctactacat	atgagtggct	ctaccatacg	ttgggtggatg	gaaatagtag	atatttggtta	37380
agcaggtaaa	ccttattttat	ttgtagccct	gtgacttggg	gtaagttatt	ttgcatgtct	37440
gggtcttttt	tatctaattt	gtaatatgaa	gctattatta	ttattattat	tattattatt	37500
attattatta	ttattattat	tattcgagat	ggagctctgc	tttgttgcoc	aggctggagt	37560
gcagtgggtc	aatctcggct	caatgcaccc	tccacctcct	gggttcaaac	tattctcgtg	37620
attcagctctc	ccgagtagct	aggattacag	gggctacca	ccacacctaa	ttttgtatt	37680
tttagtaggg	acggggtttc	accatgttgg	ccaggctggt	ctcaaaactcc	tgacatcaag	37740
tgatccatcc	acctcagcct	ccaaaagtgc	ttggtagggc	cgggcacggt	ggctaaccgc	37800
tgtaatccca	gcactttggg	aggccgaggc	gggtggatca	cgaggtcagg	agattgagac	37860
catcctggct	aacacgggtga	aaccccgctc	ctagtaaaaa	aatacaaaaa	actagccagg	37920
tgtgtgtggc	ggcacctgta	gtcccagcta	cttgggaggc	tgaggcagga	gaatggcatg	37980
aaaccaggag	gcggagcttg	cagtgagccg	agatcatgcc	actgcaactcc	aggctgggca	38040
aagagcgag	actctgtctc	aaaaaaaaata	aataaaaata	atagataaat	aataataata	38100
aagtgtctgg	attacaggca	tgagocacca	cgcccagctg	aagctaatat	tattagctaa	38160
taataatata	tcataatttt	gaccattgaa	ttattaaagg	aacctaaacc	atagtagtaa	38220
gtgctttaaa	aatgttgggt	tttattttatc	tgtttaacctg	actgtcatca	cctcctttgc	38280
ttaggtagaa	agabatattc	ataaacacat	acatatacat	tttattcatt	cttttaattca	38340
tttactccta	tatggtcatt	gtgtttgtgg	gatttttggc	cgcacatagt	taaaaatgca	38400
gccgttactg	ctttcagatg	agtaacaagg	tagtgttccc	tggcttctgt	ggctgacagg	38460
atgtgccttc	ttcctttcta	aatggaggtt	attactgtgt	cagatataat	taaatagtgt	38520
attnaaagtt	gttgcaataa	aataattgat	ggttctaatt	ggtactttct	acgtgtttta	38580
tctttaaagc	ttttcagtgt	atataaaag	tatatatcat	acagaataaa	tttgattgtg	38640
gagcattttg	taacatcttt	caaaaatta	aagagttggt	ttttatgttt	tgtttgatgt	38700

tatgatatag	gacattagaa	gtattaatat	caattagaga	ctcatotttg	aatgtgactt	38760
tgtactttct	tatttgggtt	agtagaggag	agaacaaaaa	gaagatatgt	aatgtaatat	38820
gaaactaggg	atthaaagat	ttaacgtttt	ggatatttta	aatgttgggt	ctgttttcac	38880
cctcaaaaat	gctattacct	atthtgaaa	tatctttaa	agtgtggagt	ggtagatggg	38940
aaaggtgaca	tcggctgggc	gcggtggctc	acacctataa	tcccagcact	ttgggagggc	39000
aaggtggcgg	gatcacctga	ggtcaggagt	tcacgaccag	tctggaaca	tggtgaaacc	39060
taatctctac	taaaaaatac	aaaaattagc	cagcgtgggt	gtacgcacct	gtaatcccag	39120
ctactcagga	ggctgagtca	ggaggatggc	ttgaaccag	aaaaaggagg	ttgcattgag	39180
ccaagatggc	accactgcac	tccagactgg	gtgacagaga	gagactccgt	ctcaaaaaaa	39240
aaaaaacgcc	gggactgtg	gctcacacct	gtaactctag	cactttggga	gactgagggc	39300
agcagatcac	ttgaggtcag	gagttcgaga	ccagctgggt	gcctgggcaa	catagtga	39360
ccccgtctct	actaaaaata	caaaaaaatt	agccggggct	agtggcgggc	gcctgtagtc	39420
ccagctactt	gggaggctga	ggaaggagaa	tggcgtgaac	ccgggagggc	gagcttgca	39480
tgagccgaga	tcccgccact	gcactccagc	ctgggtgaca	gagcagagact	ccgtctcaa	39540
aaaaaaaaaa	aaaaaaaaaa	tagccgggca	tgggtggcgg	cacctcta	ctcacctact	39600
tgggaggttg	aggcatgaga	attgctcgag	cctgggaggc	agaggttgca	gtaagcagag	39660
accacaccac	tgactcttag	cctggatgac	agatgagac	ttgtctcaa	aacaaaaaca	39720
aacaaacca	accacctctt	gtccactcct	aatctccctt	actgtgtttc	acttattttt	39780
ttcccataa	tttgtcacct	tctaataaac	tgtgtaactt	atthtttatg	tttattgtgt	39840
actatctgtc	ttctccctgt	tggaatata	actctatgga	ggagggatcc	ttgtctatat	39900
cccaagtgcc	tacaacagt	tctagcacat	gctagtacc	gaataaatat	ttgccaagt	39960
aatgaagtgg	tttataacc	ttggcactat	acctgatata	gtgtagggac	ttaataaatg	40020
aatgttaaat	gaagctatct	taaaaatata	tatafatata	taacagccag	atagccaat	40080
ttgccgacat	caaaattaat	ctthttcatct	attgcagggt	ttacagttaa	gctttgctct	40140
aaggttcaact	atcactaaac	caaaaataaa	taatatggat	ggtcttttcc	catthttaac	40200
agagaggcta	catcctctca	aaccagtcag	tattaggaat	tcccaggatc	attctctcca	40260
ttttccatcc	taattagata	cctatctggc	accacatctc	tgaaaagctc	caatthttaa	40320
ccagcacata	ataaacaaat	atgttactat	taatacacgt	ttgtacacaa	gtcattthta	40380
ttgtcttagg	tatatccact	cccagacctt	ctgagttatt	gaaagacagg	aatthgtgat	40440
gtcacctctt	ttttcttttt	tttgagacgg	agtcttgcc	tatcacccag	gcagagatgc	40500
aatggcgtga	tctcggtcca	ctgcaacctc	tgccctccgg	gttcaagcga	ttatcctgcc	40560
tcacctctct	gagtagctgg	gattcacagg	ggcctagtt	tgggtthttg	gtthtaaca	40620
ggatggtctc	ctctttgaga	agttgccatg	agttagccat	gatgatacct	gggacaata	40680
tggttccacag	agaagtaaga	gccaatata	cgcttaaga	gaaagtacct	ggtatgtttg	40740
tgaagcagca	aagaggccaa	tgtgttccag	caaggagatt	atagtagaga	tgaagtca	40800
gaactcatgc	ggagagggat	tgtcaaca	cgcagggctg	taaaggccac	catgaggact	40860
tgagctthta	ccgcctgtga	aatggggagc	agttacaaag	atthtttaaga	aagcagccca	40920
gtgcagtgcc	tcatgcctat	aatcccagct	ctthgggagg	ctgaggtggg	tggatcacct	40980
gaggtcagga	gttcgagacc	agcctggcca	acatggcgaa	accttgtctt	tactaaaaat	41040
agaaaaatta	gctgcgtgtg	gtggccatgc	ttgtaatccc	aaatactcgg	gaggtcgagg	41100
cagaagaatc	acttgaacct	gggagcgcca	ggttgcagtg	agccgaggtt	gcgccactgc	41160
cctccagcct	gggcaacaaa	gcgagactcc	gtctcaaagc	cagaaaaaaa	gattthtagg	41220
aaaggaatga	aatactctga	cttaaattht	agaaggatca	ctctggttgc	tatgttgcat	41280
tatagactct	aagggtagaa	ccagggagag	caattaggac	actgttacag	taatctagga	41340
gaaaagtgat	agaggcttgc	accaggtag	tatcaggaga	aatggtaaca	agtggttgga	41400
ttatgattat	agttgaaggt	cgaaacaacta	gatttgctga	tagattggat	aaaaatggctc	41460
tctctttatg	gatctcataa	tacagttagga	gaaacagaca	cataaatagt	catcactgca	41520
ggatgggtat	tagaaatatt	ttgcaaggg	gtaaaagggtg	ctaaagaaagc	ttctctgtgt	41580
gacaaagctg	gttaaagaaa	gaaaagaaag	cttcctagaa	gtgaaacct	aaataacac	41640
ttagataaat	gaataaaagt	tatgaaagaa	tgtgagggct	gggcacagtg	gctcatacct	41700
gtaatcctag	cactttggga	ggttgagtcg	ggggtattgc	ttgagtcag	gagttgaaa	41760
ccagcctgga	cagcgtactg	agaccocatg	tctaaattta	aaaaaaaaaa	aaatagccag	41820
gagtggtgat	gtgcaacctg	agtcctagct	acttacttgg	gaggtcgagg	caggaggatt	41880
gcttgagctc	aggaggtcga	agccatgagt	agccatgagt	acaccactgc	actctagtct	41940
gggtgacaga	gtgagacct	gtctcaaga	aaatgaaaag	aaaaagaatt	tgagaaaaga	42000
tatgtcagga	aaatgtgatg	acatgaacaa	atactcaaag	gcaagaaaag	catggtgatg	42060
aggggaggg	acaggttgag	acatgaagtt	gggattgccc	aggcgtaaag	ggtataggca	42120
gatcatggat	ggcctggtht	gtcatagga	ggagcttggg	ctcttatctg	tgggcaatgg	42180
gaagccacta	aagggthtta	agtagaagag	tgttatatgg	taaggtthtt	ctctccagta	42240

gatcactcag	atgactgtgg	gggttgatg	tgaggaggtg	aagcagcaaa	gttactgctt	42300
taatttaaat	cattgattaa	tctgtacaat	cctagtcat	ccaaaaagaa	acattagtcc	42360
tttaacagaa	gtgaattggg	tctgggctca	gtggctcaag	ccgtaatcc	cagtactttg	42420
ggaggccgag	gcgggcagat	caactgaggt	caggagtttg	agacaagtct	ggccaacatg	42480
gcaaaacctt	gtatctacta	aaattacaaa	aattagctgg	gcatgggggc	acccacctgt	42540
aatcccagct	attcgggagg	ctgaggcagg	agactctctt	gaacccataa	ggcagaggtt	42600
gcaatgagct	gaggtcgcac	cactggcctc	cagcctgggc	gacagagtga	gactccgtct	42660
ccaaaaaaa	gaaaaaatat	tgagcagtt	tcacagatgc	tgtttactgt	tatactgtat	42720
gtgtctatga	ctcctcctcc	aagaaaaaaa	aatgaattg	gagcaggttc	acagaagcaa	42780
tgtactgtta	cactgtatgt	gtctgtttct	acacataata	atctgaattc	tgtgtacacg	42840
aaaagaaatc	ctatagcttt	ttattcctag	ctataaaaac	taagaatata	atttctttct	42900
ttttattaca	taatgaatgg	ttctgttaac	ttttgttaa	aatttcattg	aggggagtat	42960
taattccat	ttatgtacgg	atgtctacat	ttacaaatca	gtgtattttt	tgatttatgc	43020
ttttactgag	acaaaggggt	tctgtctcag	catggctcatt	taaagagttt	atcattgaga	43080
aaaatcagat	gaccaacctg	ttagctcaaa	aaaaaaaaa	cctccaaggt	atattgtatc	43140
agccagttct	aggatacaaa	agccatgcag	tactttgtgt	tttgtgcaa	aaagggtagc	43200
tgctatttga	cctgtcccaa	agcagtggt	agttgtaccg	taaaccaagc	atggtagctg	43260
tttgtcaaac	tttagaaatg	aaagtttaag	agagtttaata	tataggtgct	gcatttttta	43320
tgtattcatt	gacttgctgg	tcagaagaa	aagaatcaat	tatgattcag	cacaatactc	43380
cacttgggga	agagagtgc	gcagtagttt	agagtgtcag	ggatcaaac	gctacctct	43440
gggttccag	ttgttagact	taagagacc	agatcttggg	aggggtttt	gttgttattt	43500
tagatggggg	taaggtgaa	atactccctt	ttaattgatt	aatataaaaa	ataggtaatg	43560
tagacgtatg	tcacaaaatt	ttaacaagtt	tgagtgtgaa	aagcatccat	cccattctat	43620
gccctggcta	cccatttccc	ctccctagag	gcggctgtta	ttatcagttt	ctctgggtcc	43680
tttttagagg	taatcttttt	gtatgcaact	gtgcacatgt	atgtttacac	aaatggtagg	43740
gtattgtata	tataacagt	tactgcagta	tgcttattg	actttgcttt	attgctttta	43800
ttttctgcct	gacatatgtt	aatgtggcca	aatagggcat	ttcccagaca	atccttatag	43860
tttcattctg	tgctacttta	aatgctgttc	tccttccctg	aaatgtcctt	tttttttttc	43920
ttttcttttt	ttttttttga	gacagagtct	cgctctgtca	cccaggctgg	agtgcagtgg	43980
cacaatctca	gctcactgca	acctccacct	ccocggttta	agccattccc	ctgcctcagg	44040
ttcccgagta	gttgggacta	caggcgcgtg	ccaccacgcc	cggatacttt	tttgtttgtt	44100
tgttttagta	gagacagggg	ttcactgtgt	tagccaggat	cgctcaatc	tcctgacctc	44160
ctgatctgcc	tgctcagccc	tccc aaagtg	ctgggattac	agacgtgagc	cactgcgctt	44220
ggctgtcctt	tcttttaatc	ttttcaaatt	ctagccagtt	catagcccac	ttgctctgta	44280
aagctttttc	agtgatcttt	cttttgtacc	atgcagcttt	ttaaaaaatc	tacagtttta	44340
ttataagttg	ctaattatat	ttaataaata	cctaaatggt	agaaattgat	tcatttttta	44400
atctccgtag	tgccaggaac	tgtggtcat	gtctgtaatc	cctgcacttt	gggaggctaa	44460
ggcgggagga	tcgcttaagg	ccaggagctt	gagcccagcc	tgtggaacat	atggagacc	44520
cgtctctaca	aaaaataaat	tagctgggca	tagtgggtct	tgtatgtagt	gtcaactact	44580
caggagactg	aggcagaagg	ataacttgag	ctcagggagt	tgaggccaca	gtgagccatg	44640
atgggtccac	tgtacttcag	cctgggtgac	agagcaagac	caagaccgat	aattatctcc	44700
atggtatttg	ttttggacgt	attagatatt	cagtaaata	tttctggtga	tgatagtgat	44760
atctaactac	ctcaggttac	ctaacctgaa	taaggcctat	ttttaacata	gccattcact	44820
tgtttaattc	tgttattttc	gtgacttcat	ttttactaat	tataatttaa	aagataattt	44880
tatataaatg	caaaaaactt	ttatcatgta	acaactctgt	aacaccttga	gtaatttgtc	44940
atttcacatg	cacattttat	ttgttcatta	tctttttctt	ttttgtttat	tttttctggt	45000
ttttaccttc	tctgttgcat	agagaagatg	caaccagttt	tgccattttc	caagtcacaa	45060
acggacgtct	ataatgacag	tactaaactg	gcatgccgca	atggacatct	ccagtcaggt	45120
gggttttaggt	taactaacat	aaaataataa	agcttgcag	caaagtaaga	gtcttactta	45180
acctgataat	gttgatttga	ttttataatg	gtaggtttta	ttgttcatgt	tttcacaggg	45240
ctgtgttgaa	aatatagtac	ctttatcctc	atagtatata	aacatgcaac	aagtccaggt	45300
tagtcatctt	aaacttgctt	tatttgcata	ttttgcaacc	tctgggtttt	cagattataa	45360
aatatgaagg	tatacattcc	cgtaagtgat	taaaaataac	ttttttctg	ttgtggcttt	45420
tgtgttttta	tatggatagc	tatactgatt	tttttcatga	aatagtgttt	tctaagacac	45480
attttaattt	tttatgtatc	aatttttatt	gcatatcaat	tttcattttat	agtatatgtg	45540
aatgtcttat	agtacttgat	agaagaattt	tgaaacttca	catagtgagg	agaacatta	45600
cagtattatt	tgctatggat	taactctatc	cttcataaac	ttttggatat	caactctac	45660
agaagattaa	atgtctttct	gttttgtttt	ctttgggtag	gggactctcc	atttcttagg	45720
ttctctgaca	tctgagctag	ctttttgttt	aattaaaact	tttttttag	ttgacagata	45780

atgtacctgt	tcatggggta	cacagtgatg	ttttgataca	tatagtatat	ggtgattaga	45840
tcaaggtaat	tatcatttcc	atcatcccaa	acatttatca	ttccttgtgt	tggaaacatt	45900
cagtatcttc	cttctagcat	ttgaaatgat	ataatataatt	atattgttac	ctataatcat	45960
cctgcagtga	tgtagaacaa	tggaacttac	tctttctatc	tagctatacc	tttgtatcct	46020
ttccctaccc	ttttcagcct	ctagtatcct	ctgttctact	ttttatctct	atgagattaa	46080
ctttttaaaa	ataatattct	gtgtgtatat	ataccacatt	ttcctttatc	attcgtctgt	46140
tcttgacac	ctaggttgat	tctgtatctt	ggctgtttgt	aatagtgtgt	cagtaaaccat	46200
gggggttaga	tgtctgttccg	atatgatgat	tttcttccct	ttggataaat	tcccaatagt	46260
gggattgttg	ggtcataatgg	tagttctgtt	tgtagttttt	gaggaacctc	catactcttc	46320
tccatagtag	ctgtactagt	ttagattccc	aaaagttagcg	tataagagtt	cactttctac	46380
acatcctcac	cagcatttgg	tacttttttt	gtctttttga	taatagccat	cttaactggg	46440
gtaagatacc	ttgttgtggg	tttgatttgc	atcttccctca	taattaacga	tgttgagcat	46500
tttttcatat	ctttgttttt	gacggtccag	ctagttttat	tggttacttt	tttttttttt	46560
tttttagcaat	agtttcatgt	gcagggtcca	aattatattt	aatattactt	ttttcaaat	46620
ataccagaac	aaccttaggt	taggcataaa	aacaactgcc	ttccttttct	tttggcataa	46680
taattatata	tatatTTTTA	attacacttt	acgttctagg	gtgcatgtgc	acaacatgca	46740
ggtttgttac	atatgtatac	aagtgcctatg	ttggtgtgcc	gcaccatta	acttgtcatt	46800
tacattaggt	atatctccta	atgctacccc	tcccctctcc	cctcacccca	cgacaggccc	46860
catgtgtgga	tgttccccac	cctgtgtcca	agtgttctca	ttgttcagtt	cccatctgtg	46920
agtgagaaca	tacggtgttt	ggttttctgt	ccttgcgata	gtttgcctag	aatgatggtt	46980
tccagcttca	tccatgtccc	tacaaaggac	atgaaactcat	cctttttgat	ggctgcttag	47040
tattccatgg	tgtatatgtg	ccataTTTTc	ttagtcagat	ctatcattga	tggacatttg	47100
cattgattcc	aagtctttgc	tatttgaat	agtgcacag	taaacataca	tgtgcatgtg	47160
tctttatagc	agcatgattt	ataatccttt	gggtataatac	ccagtaatgg	gatggctggg	47220
tcaaagtgta	tttctagttc	tagatccttg	aggaatcgcc	acactgtctt	ccacaatggt	47280
tgaactagtt	tacgttccca	tcaacagtggt	aaaagtgttc	ctgtttctcc	acatcctctc	47340
cagcacctgt	tgtttcctga	gtttttaatg	atcgccattc	taactgggtg	gagacgggat	47400
ctcattgtgg	ttttgatttg	catttctctg	atggccagtg	atgatgagca	tttttcatg	47460
tgtcttttgg	ctgcataaat	gtcttctttt	gagaagcatc	tgttccatata	ctttgcccac	47520
tttttataggg	ggttgtttga	ttttttctt	gtaaattttt	ttaaattctt	tgtagattct	47580
ggatatcagc	cctttgtcag	atgggtagat	tgcataaaatt	ttctccatt	ctatagggtg	47640
cctgttccact	ctgatggtag	tttcttttgc	tgtgcagaaa	ctcttttagtt	taattagatc	47700
ccatttgtca	attttggctt	ttgttgccat	tgtctttggg	gttttaggca	tgaagtccct	47760
gcccattgcca	gtgcccattg	cagtgtcctg	aatggatttg	ccttagatttt	cttctagggt	47820
ttttatagtt	ttagaactaa	catttaagtc	tttaatccat	cttgaattaa	ttttgtata	47880
aggtgtaagg	aagggatcca	gtttcagctt	tgtacgtttg	gctagccagt	tttcccagca	47940
ccatttatta	aataggaat	cctttcccca	tttattgttt	ttgtcaggtt	tgtcaaagat	48000
cacatgggtg	tagatgtgtg	gtattatttc	tgagggctct	gttctgttcc	attggtctat	48060
atctctgttt	tggtatcagt	accatgctgt	tttgattact	gtaccttctg	agtatagttt	48120
gaagtcaagg	agcaccatgc	ctccagcttt	gttcttttgg	cttaggattg	tcttgccaat	48180
gcaggctctt	ttttggttcc	atatgaactt	taaagttagtt	ttttccaatt	ctgtgaagaa	48240
agtcatattgt	agcctgatgg	ggatggcatt	gaatctataa	attacottgg	gcagtgtggc	48300
cattttcatg	atattgatcc	ttcctatcca	taagcatgga	atgttcttcc	atgtgtctgt	48360
gtcctctttt	atttcattga	gcagtgtttt	gtagttctcc	ttgaagaggt	ccttccacatc	48420
ccttgtaagt	tggattccta	ggatatttat	tctctttgaa	gcaattgtga	atgggagttc	48480
actcatgatt	tggctgtttg	tctgttattg	gtgtatagga	atgcttgtga	tttttgcaca	48540
ttgattttgt	atcctgagac	tttctgtgaag	ttgcttatca	gcttaaggag	atttgggct	48600
gagatgatgg	ggttttctaa	atatacaatc	atgcatctc	caaacagggg	caatttgact	48660
ttcttttccct	aatgggatac	cctttatttc	tttctcctgc	ctgattgccc	tggccagaac	48720
ttccaacaact	atgttgaaca	ggagtgtgga	gagaggcat	ccctgtcttg	tgctagcttt	48780
caaagggaaat	gcttccagtt	tttgcccatt	cagtatgata	tttgctgtgg	gtttgtcata	48840
aatagctctt	attattttga	gatacactcc	atcaatacct	agtttattga	gagtttttag	48900
catgaagggc	tgttgaattt	tgtcaaaagg	cctttctgca	tctattgaga	taatcatgtg	48960
gtttttgtct	ttggttctgt	ttatatgatg	gattacgttt	attgatttgc	atatgttgaa	49020
ccagccttgc	atcccaggga	tgaagccaac	ttgattgcgt	tggataagct	ttttgatgtg	49080
ctgctggatt	cggtttgcga	gtattttatt	gaggattttt	gcgttgatgt	tcatacagga	49140
tattgttcta	aaattctctt	tgttgtgtct	ctgccaggct	ttggatcag	gatgatgtg	49200
gcctcataaa	atgagttagg	gaggattccc	tctttttcta	ttgattggaa	tcatttcaga	49260
aggaatggta	ccagctcctc	tttgtacctc	tgttagaatt	cagctgtgaa	tccatctggt	49320

cctggacttt	ttttggttag	taggctatta	attattgcct	caatttcaga	gcctgttatt	49380
ggctattca	gggattcagc	ttcttcctgg	tttagccttg	ggaggctgta	tgtgtccagg	49440
gatttatcca	ttcttcctag	atttcttagt	ttatttgagt	agagggtgtt	atagttattc	49500
ctgatggtag	tttgcatttc	tgtggaatcg	gtggtgatat	ccccctttatc	atctcttatt	49560
gcactctattt	gattctcttc	tcttttcttc	tttattagtc	ttggtagcag	tctatcagtt	49620
ttgttgatct	tttcaaaaaa	ccagctcctg	gattcattga	ttttttgaaa	ggttttttgt	49680
gtctctatct	ccttcatttc	tgctctgac	ttagtattt	ctgcctctct	gctagctttt	49740
gaagggtgtt	gctcttgctt	ctctagttct	tttaatggtg	atggtagggt	gtcaatttta	49800
gatctttcct	gttttctctt	gtgggcattt	agtgtataa	atctccctct	acacactact	49860
ttaaagtgtg	cccaagatt	ctgatatggt	gtgtctttgt	tctcgttggg	ttcaaaagac	49920
atcttttattt	ctgccttcat	tttcttata	accagtagt	cattcaggag	cagggtgttc	49980
agtttccatg	tagttgtgtg	gttttgctg	agtctcttaa	tctgagttc	tagtttgatt	50040
gcactgtggc	ctgagagaca	gtttgttga	atctctgttc	ttttacattt	gctgaggagt	50100
gctttagttc	caactatgtg	gtcaattttg	gaataggtgt	ggtgtgtgtc	tgagaagaat	50160
gtatattctg	ttgatttggg	gttttagagtt	ctgtagatgt	ctattaggtc	cacttggtgc	50220
agagctgagt	tcagttcctg	gatctgtctt	gttgactgt	ctaataattga	cagtgggtg	50280
ttgaagtctc	ccagtattat	tgtgtgggag	tttaagtctc	ttgttaggtc	tctagggact	50340
tgctttatga	atctgggtgc	tctgtattg	gggtcatata	tatttaggat	agttagctct	50400
tcttgttgaa	ttgatccctt	tagcattata	tgatggcctt	ctttgtctct	tttgatcttt	50460
gttggtttaa	agtctgtttt	atcagagagt	tggattgcaa	accctgcttt	ttttgttttc	50520
catttgcttg	gtgatcttc	ctccatccct	ttattttgag	cctatgtgtg	tctgtgacg	50580
tgagatgggt	ttcctgaata	cagcacactg	atgggtcttg	actcgttatc	caatttgcca	50640
gtctgtgtct	tttaattgga	gcatttagcc	catttccatt	taaggttaat	attggtatgt	50700
ttgaatttga	tcctgtcatt	atgatgttag	ctggttattt	tgctcgttag	ttgatgcagt	50760
ttcttcctag	cctcagatgg	ctttacaatt	tggcatggtt	ttgcagtggc	ttgtaccggt	50820
tgttcatttc	catgttcagt	gcttcctca	ggagctcctg	taagcaggcc	tggtagttag	50880
aaaatctgtc	agcatttgct	tgtctttaa	ggattttatt	tctccttcac	ttatgaagct	50940
tagtttggtc	ggatagaaa	ctctgggttg	aaaattcttt	cctttaagaa	tgttgaatat	51000
tggcccccac	tctctctctg	tttttagagt	ttctgccaag	agatcagctg	ttagtctgat	51060
gggcttccct	ttgtgggtaa	cccgcctttt	ctctctggct	gcccttaaca	ttttttccta	51120
catttcaact	ttggtgaate	tgacaattat	gtgtcttgg	gttgccttc	ttgagttaga	51180
tctttgtggc	attctctgta	tttctgaat	ttgaatgttg	gcctgccttg	ctagggttgg	51240
gaacttctcc	tggaataat	cctgcagagt	gttttccaac	ttggttccat	tctccccctg	51300
actttcaggt	acaccctca	gacatagatt	tggtcttttc	acatagctct	atatttcttg	51360
gaggttctgt	tcgtttcctt	ttactctttt	ttctctaaac	ttctctctg	gcttcatttc	51420
attcatttga	tcttcaatca	ctgataccct	ttcttccact	tgatcgaatc	ggctactgaa	51480
gctcatgcat	gcactcagta	gttttcgtgc	catggttttc	agctccatca	ggccatttaa	51540
ggctctctcc	atgctgttta	ttctagttag	ccatttgtct	aatctttttt	caaggttttt	51600
agcttctttg	caatggtttc	gaacatcctc	ctttagctcg	gagaactttg	ttattcccca	51660
tcgtctgagg	cctaactctg	tcagcttgtc	aaagtctttc	tctgtctagc	tttgttccgt	51720
tgctgggtgag	gagctgtggt	cctttggagg	agaagaggcg	ctctgaattt	tagaattttc	51780
agcttttctg	ctctgttttc	tccccatctt	tgtggcttta	tctacctttg	gtctttgatg	51840
atggtgacgt	acagatgggg	ttttgggtg	gatgtccttt	ctgtttggtt	atcttctctc	51900
taacactcag	gaccctcagc	tgcaggtctg	ttggagtttg	ctggaggtcc	actccagacc	51960
ctgtttgcct	gggtatcacc	agcggaggct	gcagaacagc	aaatggtgca	gaacgacaaa	52020
tgttgctgtc	tgatccttcc	tctgaaacct	tcgtctgaca	ggggtaccca	ggtatatgag	52080
gcgtcagtca	gcccgtatgg	ggaggtgtct	cccagttagg	ctacttgggg	atcagggacc	52140
cacttgagga	ggcagctctg	cogttcgcgg	atctcaaac	ccatgctggg	agaactacta	52200
ctctctttag	agctgtcaga	cagggacctt	taagtctgta	gaagttactg	ctgccttttg	52260
ttcagctatc	ccatgtcccc	agaggtggag	tctacagagg	cagtcaggcc	tccctgagct	52320
gtgggtgact	ccacccagtt	caagcttctt	agctgctttg	tttaccact	caagcctcag	52380
caatggcaga	tgcccccgcc	tccagcctct	ttgcccctt	gcagttcgat	ctcagactgc	52440
tgtctctagca	gtgagcgagg	ttccgttggc	atgggacct	ctgagccagg	catgggat	52500
aatctcctgg	tgtgcccgtt	gctaagacca	ttggaaaagc	ccagtttag	ggtgggagtc	52560
tccctatttt	ccaggtacca	tatgtcacgg	cttcccttg	ctaggagagg	gaattcccca	52620
accccttgcg	cttcccaggt	gaggcaatac	cccgccttc	agctcaact	atgtgggctg	52680
cacccactgt	ctgacaatcc	ccagtgagat	gaaccagtt	cctcagttg	aaatgcagaa	52740
atcagctgtc	ttctgtgtcg	ctcacactgg	gagctgcaga	ctggagctgt	tccatattgg	52800
ccatcttgg	accctgctt	cttcattcat	atgtaataca	aaacttctaa	ggttttagtg	52860

gagaagagat	agagtaaaaag	gatttctata	gacagaagaa	acagttggtc	atcaactctt	52920
tcoccttggtg	cttcacatct	ccctaagga	cttcttatgt	tggtttggtc	ttacagtata	52980
gtcagtgcca	gtttccctct	tcoccttct	tectctgtct	gatttaaaat	gctgtttcaa	53040
gtatcgatat	aagtattttg	cctgtttctt	ttagcgtggc	tgtgaagggc	tgacattttc	53100
agaaggcact	tactgaaaaa	aaaaaaacaa	agaaatgtaa	gagtccatca	catataaata	53160
gttaagtttc	taaaatatgt	atttgagatc	ccagtaattc	tactaggata	aatagcaaaa	53220
attctccagc	cctgaagagt	tggctctgtc	ttcctttcct	tgttatcttg	attctctttt	53280
atctcatttc	acatgctaga	ccccatcatg	ttttcctgct	gcattcccca	ctccaccctc	53340
caaagcaatg	ttctcttctc	gcaacctgtt	tgaaaaaatt	gaactttgtc	ttatctcata	53400
tcocctagtac	aagttgaata	tcocctaata	gaaatctgaa	atccaaaatg	ctctaaaatt	53460
cgactcctttt	ttttttgaga	tgaagtcttc	ctctgtcact	caggctggag	tgcggtggcg	53520
caatcttggc	tcactgcaac	ctccgcagcc	tcocctaacc	ccatccctgt	gggttcaagc	53580
gattctccca	cttcactcctc	ccaagtagct	gggactacag	gcacctgcca	ccacacccaa	53640
ctaatttttt	gtattttttag	tagagatgag	gtttcaccat	gttggtcaag	ctagtctgag	53700
actcctgacc	ftaagtgatc	caoctgcocct	ggcctcccaa	agtgttgagg	ttacagtcat	53760
gagccactgt	gocctagccaa	atatccaaaa	ctttttgagc	gcrgacatga	tgctcaaagg	53820
aaatctcga	gtattttgca	ttttggattt	ttgggttaag	gatgctgaac	cagtaagtat	53880
aatgtaaacca	ttgcaaaaatg	caaaaaaatg	taaaaaacct	aagcagttct	ggtcccaagc	53940
atctgcata	agggatactc	aaocccataat	cttttctttt	ctgtttctcg	gttggaaggg	54000
catattggct	ttatagctaa	tttacacagc	attgatgttt	aatacacagt	gagccaagt	54060
agatcactca	gacctattag	gtgtcactca	gtgtcactca	cttctggaac	attctgtgat	54120
gttttacttg	gatgagttct	ttcacctctc	ttgggaaata	gtcataccaa	agtctgctta	54180
ttactacaat	gtgttatcta	ttatacattg	tcocattttc	tactcaaaat	tactagacag	54240
gcaagaaaaa	agtaaagtat	aaoccttact	caagaaaaaa	aagcaatcag	tagaaaactgt	54300
gagtgggccc	agatattgga	tttagcagac	aaagacttca	aagcagctat	tataatagt	54360
ttaagaat	gaaaagaaat	atggtatcag	ttaaacagga	aatcctaagta	gatgatataa	54420
actagacaat	aaaaataata	ttctaaagt	gaaaagtgt	gttactgaaa	ttaaaaattt	54480
acagacaaca	gcctcaacag	cagattagag	atagcaaaag	aaagattcag	tgactttgaa	54540
tgcaaggtctg	tagaatattac	taaattgatg	actctcacgt	agcaactttc	acccgtagtt	54600
tggtttcata	tacaatgctt	taactctttt	tgctattttt	tctcttactt	tctgtgatgg	54660
aaacattttg	tgctttgatt	ttaatgggtg	tatcacagtt	atatacatct	ttacaaactc	54720
atcaaatgt	atacttttatt	tatatattta	tttttttttg	agatagggtc	tgactctttc	54780
tcocaggctg	gagtaacagct	gaaattagtt	gggactacag	gtgggacca	ctgtgccagc	54840
taattttgt	attttttgta	gagatgggtg	ttcccatgt	tgccagctct	ggctcaaac	54900
tcctgggctc	aatcattttct	gccgcctcag	actcccagag	tggtgggatt	ataggtgtga	54960
gctactgtgc	ctggctggta	tagctttttt	tttttttttt	tttttttttg	tgacagagtc	55020
tcactctgtc	gcccaagggtg	gagtgcaagt	gcgcgatctg	ggctcactgc	aacctctgcc	55080
tcocgggttc	aaccatttct	cctgcctcag	cctcctgagt	agctgggact	acagatgcgt	55140
gccaccatac	ccagctgatt	tttgattct	tagtagaaac	agagtttcac	catgttgccc	55200
aggatgatct	cgatctcctg	acctcgatcc	acctgccttg	gcctcccaca	gtgctggcat	55260
tacaggcctg	agccaccatg	cctggctgtt	ggtatacttt	aaatggatgt	aattcattgc	55320
agattatacc	tcagttaaatt	tttatttagt	ttttgagact	gagttgtctc	gttgcccagg	55380
ctggagtgca	gtggcaagat	ttcagctcac	tgcaacctct	gcctcctggg	tttaagcgat	55440
tttctgtctc	cagcctcccc	ggtagctgga	attagaggtg	tggtgccaca	tgatcagatt	55500
atttttgtgt	ttttagtaaa	gatggggttt	caccatgttg	gccaggctgg	tctcgaactc	55560
ctgacttcag	gtgatccacc	cgctcggcc	tctcaaagtg	ctgggattac	aggtgtgagc	55620
caccacgcct	ggccttaatt	tttaaaact	gtaaggctta	taaagaaaag	aatattcccc	55680
ttctgtttct	ttcctctcac	gtagcaacct	tcacccttag	tttggtttca	tatacagtgt	55740
tttaactctt	tttgctattt	ttttctctta	ctttctacta	tattttccaaa	tacaatgctt	55800
ctataatgat	tccttttttt	tttctatcag	tttttgataa	tcattgactc	cttatggctca	55860
aagaagactt	aattcccttc	tgtaactctt	catatattaa	tataactaat	atataatgtg	55920
tttagttttg	atataactaa	aaataactat	atgtatata	aactatata	gtataagcta	55980
tctgtatata	gttatatatg	tatatgcaca	tacatatata	gttacatatg	tatacacaca	56040
tacatatata	gttacatatg	tatatgcata	tacatatata	gttacatctc	tatatgcata	56100
tacatatata	gttacatatg	tatatgcata	cacatatata	gttacatatg	tatatgcata	56160
tacatatata	gttacatatg	tatatgtata	tgtatataca	tatatagtta	tatatgtatt	56220
agttatatca	aaataactat	gtatatatag	ctgtaaatgt	atatataaac	tatatgtata	56280
cagttatata	gtattttgta	tatgtgtgta	tacatatata	gttttttttg	tttttttttt	56340
ttgttttttt	tgtttttttt	tggagatgga	gtcttgcocct	gtccccagc	ctggaatgca	56400

gtgggtgctat	cttggccttac	cgcaacctct	gcttcccagg	ttcaagcaat	tctcctgctt	56460
cagtctcccc	agtacctggg	attacaggca	cgtggcacca	cgccaggcta	atthtttgta	56520
tcttttagtag	aaatgggggt	tcacctggtt	ggccaggctg	ttctcaaac	cctgacctcg	56580
tgatccaccc	gcctcggcct	cccaaagtgc	tgggattaca	ggcgtgagca	ccgcgcctgg	56640
tccctataca	tatatagtta	tatataccta	tagttattht	tagttatata	aaaataacta	56700
tatgtatatg	taactatata	tatagtata	atagtatata	tatactatat	agtgtgtata	56760
tatagtatat	atactatata	tagtgtgtat	atatagtata	tatatagttg	atatacgtg	56820
tatacactat	atactatata	gtgtatatat	agtatatgta	gtatatatag	tataatagg	56880
atagtatata	tagtatata	agtgtatata	tactgtatat	atagtgttac	atagtatact	56940
atatagtata	catatagtac	actgtatagt	atataatagta	tagtatata	agtatacata	57000
gtatactata	tatagtata	tatacatagt	atactatata	gtatatagag	tatatataca	57060
gtatactata	tagtatatag	agtatatata	cagtatacta	tatcgtgtgt	atagagtata	57120
tatacagtat	actatatacgt	gtgtatagag	tatatataca	gtatactata	tcgtgtgtat	57180
agagtatata	tacagtatac	tatatcgtgt	gtatagagta	tatatacagt	atactatata	57240
gtgtgtatag	agtatatata	cagtatacta	tatcgtgtgt	atagagtata	tatacagtat	57300
actatactcg	gtgtatagag	tatatataca	gtatactata	tagtgtgtat	agagtatata	57360
tacagtatac	tatatagttg	gtatagagta	tatatacagt	atactatata	gtgtgtatag	57420
agtatatata	cagtatacta	tatagtgtgt	atagcgtata	tatacagtat	actataatag	57480
gtgtatagcg	tatatataca	gtatactata	tagtgtgtat	agcgtatata	tacagtatac	57540
tgtatagttg	gtatagcgtg	tatatacagt	atactgtata	gtgtgtatag	cgatatata	57600
gtatactgta	tagcgtatag	agtatatata	tagtatactg	tatagtgtat	agagtatata	57660
tatagtatac	tgtatagttg	atagagtata	tatagttaca	tatacatata	gthttttta	57720
gtaatatcaa	aaaaccatag	ctaagatttt	tatgatttag	caaaatattg	tttaccocaa	57780
agccacgtat	tatatattaa	tgatattttt	tcccttttga	gthttttctc	cccacaggtt	57840
ttctgatctt	ttttttgtt	ctctaacttt	atattatagg	tatatactct	tttctctggt	57900
taatttgctt	tttaactaga	agacctcctt	tagtagttct	tttagtgaat	atctgtgggt	57960
aaactcttag	tctttccagt	gtctgaaata	ccctcatttt	aatgtgacag	gtcacttttc	58020
tcacctcaac	atthttggaca	tacctcattt	ccaacaaaca	tttgttactt	ctgatagaga	58080
gtctgctgtc	attctcatac	ttgttccttg	tggaaatgct	tatatataaa	gtctcagaac	58140
tatatattht	taatatat	acatathttt	atthtttaa	aaatgtataa	ttaagttagt	58200
agtagttagta	gtagttagtag	tagtagtagt	agtagttagt	ttttttttt	tagatggagc	58260
ctccatctcc	cggttcaag	tgattctcct	gcctcagcct	cctgagtagc	tgggattaca	58320
ggcacacacc	accatgcctg	gctaattttt	gtattttgag	tggagatggg	gthtcaccat	58380
ggtggccagg	ctgatcttaa	actcctgacc	tcagatggac	caactgcttt	ggcctccaa	58440
agcctagga	ttacaggcat	gaaccactgc	gtctggcctt	aagttattat	tgactattga	58500
tgccctggtg	ctctatcaag	tagtaggtct	tatttattct	agaacctcgt	tttaagttat	58560
cctgaatgtg	gctttttttt	ttttttttcc	taagacggag	tctcactctg	ttgccaggc	58620
tggagtgcaa	tggagtgatc	ttggctcacc	gcaactctctg	cctccttggg	tcaagcaatt	58680
ctcctgcctc	agcctcctga	gtagctgaga	ttacaggcgc	acgccaccac	acctggctaa	58740
tttttggtt	tttagtagag	acgggtttc	accacattgg	tcaggcctgt	ctcaaactcc	58800
tgacctgtg	atccgcccac	cttggcctcc	gaaagtctctg	ggattacagg	cgtaagccc	58860
cacgcccagc	ctaattgtgga	ttttgtttt	ttttttaact	tgaattttat	ttatactcc	58920
tgattttata	atagtttaact	tgtagtataa	gtaaacctgt	tgtctaaata	aataagctga	58980
tttctatac	tagtttagag	tctttagttt	tcttaaaccg	tattaagaga	tttgattaca	59040
aaagtaaaat	tgaggctttt	gtctttttt	ttggccactc	taatttgctg	actgacttaa	59100
catttgctct	cacagttaac	attgtaaata	aactctgatg	ttgttgctg	agtatgtaa	59160
aatattatth	aggagtgggtg	agatttaaac	atthtcaaac	ctgtaatata	tatagtacag	59220
caacactgga	caaataacct	tttcaactta	atcagaggtg	ttctataaat	gttcatttgc	59280
atataatacag	cccatataat	acttagcata	tgtaaaagca	catctttata	ctctttcaat	59340
taaaattgat	tattattatt	atthttttt	tgagacagag	tcctctcag	tcacccaggc	59400
tggattgcag	tggtgtgatc	agggctcact	gcaatcgggt	agctggaact	acagggatgt	59460
gccaccacgc	ccagctatth	ttgtatthtt	agtgagatg	gggtttcacc	atgttggcca	59520
ggctggctct	gaactccata	cctcaggtga	tctgcccgcc	tcagcctccc	aaatagctgg	59580
gataacaggc	gtgagctact	gcaccgggcc	taaaattgat	tagtttttaa	gacccctctg	59640
gtgggaacag	actttcatag	tgthttatagt	tataatagtt	tggaaatcaa	gtagtctgag	59700
gtactthtt	tttctctgt	acatcctatc	cctaagcct	ttactatatt	ttgtttctta	59760
tttgagagct	ttatthgctt	cataacttht	ttttttttt	ttttttttg	agacagagtt	59820
tcaactctgtc	accaggtctg	gagtgacagta	gtgcaatctc	ggcttactgc	aacctccgcc	59880
tctgtagttc	aagtgattct	cctgcctcag	cctccgaagt	agctgggatt	ataggccccc	59940

accaccacgt	ctgactgatt	ttttgtat	tttagcagaga	cgaggtttca	ccatgtttggc	60000
cagcctgggc	ttgaaactact	gacatcaggt	gatccgtcca	cctcagatc	ccaaagtgc	60060
gggattacag	gtgtgagcca	tcgcgcccgg	cgcttcataa	cctttttata	ttgtctttct	60120
gttccccaaa	atgaaagtgt	ttcttcttaa	aaactactgt	ttttgttttt	ttgttttttt	60180
ttttat	gatggaatct	ctctctgtcg	ccaggctgga	gtgcagtgcc	acgatctcgg	60240
ctcacttcaa	cctctgactc	cgtggttcaa	gcaattctcc	tgccctagcc	ttcccaggtag	60300
ctgggattac	aggcacaat	actgttttta	ataagaaaat	tagaataggg	acaggcacia	60360
tggtcactc	ttgtaatccc	agcactttgg	gaggccaagg	tgaggaggatt	gcctgagccc	60420
aggagtccga	gaccagccta	ggcaacatag	tgagtggagac	cctgtttcta	taaaaaaat	60480
ttttaattag	tccggtctgg	tggtgcacia	ctgatgtccc	agctacttgg	gagtttgagg	60540
tgaggagtag	gatcacttga	gcccaggaag	attaaggcca	caatgagctg	tgatcatgcc	60600
actgcaactc	agcccaggcc	acaggtttca	aaaaagaga	aaattagaat	tatacgttag	60660
ggggaaaaaa	cctaataacc	atattcctat	ctacctagtg	accacccta	tgaagatcct	60720
ggtacatggt	ctcattgaaa	tcttat	aaat	gaattaaaa	aatatattat	60780
taccttgggt	ttacaaatgt	ttgcttgtct	ttatgttcag	ctacaat	ctaatcacia	60840
tggttttact	ttttccttgt	cagaaagtgg	agctgttcca	aaaagaaaag	acccttaac	60900
acacactagt	aattcactgc	ctcgttcaaa	aacagttatg	aaaactggat	ctgcaggcct	60960
ttcaggccac	catagagcac	ctagttacag	tggtttatcc	atggtttctg	gagtgaaca	61020
gggatctggg	cctgctccta	ccactcataa	ggatctctgg	gacagtaact	ttaatgtctg	61080
tctttttgca	aatagaaaa	tttttaagat	actattcctg	cttaagttga	tcataagtac	61140
tttataat	tttagagaat	ggataagttt	ccataaagtt	aaat	ttttttttt	61200
tgagactcct	gttttcagg	ctggagtcca	atggcgcgat	ctcggctcat	tgcaacctcc	61260
gcctcccagg	ttcaagcgat	tctcctcct	cagctccca	agtagctggg	attacaggca	61320
tgaccaccca	cgctcagcta	at	tttagtagag	acggggttc	tccatgttgg	61380
tcaggctggt	ctcaaactcc	cgacctcagg	tgatccgcct	gcctcagcct	cccaaagtac	61440
tggtgattaca	ggcgtgagcc	accgtgccc	aagttaatt	tttttaaat	ccacagggca	61500
acttgctata	aaacttttt	tgacctcat	tagacattta	ataccaagt	ttgcttacc	61560
agtctg	atataatgct	gtatatttca	ttttaaatt	at	ttatttcaat	61620
ctttgt	gttttatacc	agtaacatac	ctagagttt	gtcatgattc	taacaagggt	61680
taaaat	ttttat	taaaaaactaa	atctctggat	aaaaatcctc	ataactgact	61740
taatg	fatagttctc	tagtgaatac	agttttacct	tcaggtaaat	aaatatacaa	61800
tttat	tttataccct	ttttcctatt	tttaaagcct	gaattctgtg	aactttaagg	61860
ttact	tatgaaaagt	gtaaagtta	ggttgtat	tcataataaa	at	61920
ctttaagggt	actccgaaa	caaataggac	aaataaacct	tctacccta	caactgctac	61980
tcgtaagaaa	aaagacttga	agaat	gaatgtggac	agcaacctg	ctaacttat	62040
aatgaatgaa	attgtggaca	agtaagttt	gccatctaaa	tgttttat	tgatttttt	62100
atatt	tttacttata	aaacatgtca	ggagtgaat	agataataaa	tacctgtct	62160
ggtttacata	cttaat	atgatagttt	tcaattataa	atgtagaaa	ccattgcttt	62220
actgatt	ctatagttta	acagttaca	ttaaaaataa	ctatagtca	taggcttag	62280
gcttcactct	gtaataactg	ggccctgtt	gtatcgtaga	actaactgag	gtctgtttc	62340
ttagt	agctgttaa	tttgatgata	tagctgttca	agacttggca	aaacaagcat	62400
tgcaagaaat	tgttattctt	ccttctctga	ggcctgaggt	aagaactta	tattatcatt	62460
ttctataat	accatctgtt	actgaatcca	tagtagtagt	agtagtaaag	aaatatttga	62520
gctatactaa	aataat	tcataataag	taacaataga	tttaatgtt	taaaaaagat	62580
aaaacattaa	caattgatgt	tgagaagtta	ttatagaata	ggaaaatgga	tatgagttcc	62640
taatttcaa	tcttttata	tttttaggaa	aagtggtag	tatgatataa	ttttgtttg	62700
ttttggagac	ggagtctcac	tcattgtgtc	gccaggctg	gagtgcagtg	gcatgatctt	62760
agctcactgc	aacctctgcg	tcccagaatt	caagtgatc	tctctcctca	gcctccaag	62820
tagctgggat	tacaggcatg	tgccactatg	cccagctaat	ttttgtaatt	ttagtagaga	62880
tggggttttg	ctatattggc	caggctggtc	tttcatacct	gatgtcaagt	tatccacca	62940
tttcggcctc	ccagagtgtt	aggattacaa	gtgtgagtca	ctacatctgg	ccaaat	63000
atatcaagggt	gagagagatt	taaaattaag	ataaggta	aaaattagcc	tagtgtgggg	63060
gcgcaagcct	gtagtcccaa	ctactgggga	ggctgtggca	ggagaattgc	ttgaaccagg	63120
aggcagaggt	tgcaagtgagc	caagatggca	ccactgcact	ccatcctggg	tgacagagcg	63180
agatgtcatc	tcaaaaacia	aacaggccgg	gtacggtagt	tcacgcctgt	aatccagca	63240
ctttgggagg	ctgaggcggg	cagatcacga	ggtcaagaga	ttgagaccat	cctggcta	63300
atggtgaaac	cctgtcttta	ctaaaattac	aaaaattagc	tggtgtgtac	ggtgtgtaac	63360
cccagctact	cggtatgctg	aggcaggaga	atcgctttaa	tggtgtgtac	ggaggtgca	63420
gtgagccgag	atcacgccac	tgcaactccag	cctggcgaca	gagcagagact	ccgtctcaaa	63480

aaaataaata	aaaatttaa	aagataaata	cataaaaata	aataaataat	attaagagaa	63540
ggaaatcagg	caggtagtgg	cccctgacac	aatgagtttt	cccagaattg	gattgcttgg	63600
aaatgccgct	caaagagtgt	ggtaaactcc	atcgaaggct	aaataccaac	gtgacagtga	63660
taataaacaa	gtactttatg	ggaaagtfff	tttttaatt	atftttaaaa	agagagaaat	63720
tgtactggag	aaaagaggaa	ttcaggtaga	ataattcttt	tttttttttt	ttttgagatg	63780
gaattttgct	cctgttacc	aggctggagt	gcagtggtct	gatgttggct	cactgcaacc	63840
tctgctttg	gggtcaagt	gattctcatg	cctcatcctc	ctgagtagct	gggtctatag	63900
gcattgcacc	caacacctgg	ctaatttttg	tgtttttagt	agagataggg	ttttaccatt	63960
ttggccaggc	tggtctcaaa	ctcctggctt	cagatgatcc	gcccaccttg	gcttcccaaa	64020
gtgctgggat	tacaggcaga	ggccactgtg	cctggcatga	agaacaattt	ttaaaagaa	64080
gacttttaag	gatattaagt	catcaaagta	gatagagcca	ttaaatgatg	ggtagaaacc	64140
taactttcca	tcccataatt	tatgacttat	taaggaagat	aggcaatcct	gcagttgata	64200
ataaatatftt	ggctttccat	acttagcacc	gttttgaatt	tttccagttc	acagatggta	64260
tatttagtgc	tgttccaata	atattgcaaa	aaataaatct	gaagactcac	ttctgggtca	64320
cagttatttc	actattaaca	ttaaaatcct	acggacctac	ctgcaacatg	tagtggaat	64380
aagttgtgtg	gcacgtttgt	gtgggtgcat	tattaaataa	atgtgcaaa	gttttatggc	64440
ttgtccaatt	tttccatagt	cttgagtcac	tcacatactg	cctttatggt	gtttatcact	64500
tgtaagtgg	taccaactgt	actagaactt	actgaatatt	tttcaagtct	atfttactta	64560
ttactcattc	taagtacata	aaatcactgg	tctgatatgc	aagttagggt	ttcgctaata	64620
cccagtaaaa	taattagata	attattaaaa	tagaaagcat	ttgtcattat	gtgaactgga	64680
tttatgtcat	gtaccaactc	aggtcacact	ttgggaaacg	tagcttacgt	attgagctaa	64740
cattttgctg	gagttatata	agattagtaa	atgctataga	ctaaatattg	aaaaaaaaagc	64800
ttgcttaaag	aggatttaca	taattgtaga	cctgaaaagg	atcttagcag	tcttctgatc	64860
gattcattga	taaagaaact	agtcctaaga	gctaaaatga	cctgctcaag	atftttagtta	64920
agtggttagc	ctaggaata	cctttctaag	attagagctg	cctggttagt	gctatagctg	64980
ccacctaat	cttacattta	tttattaagg	gacatgcttt	ttacggtaga	aaaatataatc	65040
taaaattgtg	aaggtaatac	caaggaaata	tttaggctat	gcaagaaaag	tcatcagct	65100
tttgctttac	ccaggcctat	agagaacata	tctaaaataa	taagtaacat	gtgaatttga	65160
atfttagcat	atfttcatgg	tagtgacata	atgtgcaatt	aagaagctat	aataatgtaa	65220
aaatgccagt	tgcaagtatt	ttcaaataaa	tttgaataac	ctgacctcag	gaggggcaga	65280
aaccttagta	aacttagcaa	actgagacct	aaggaaagc	aagatctgtg	gtagagccag	65340
gtctaaaggt	ttatatcctg	gtacaaactt	gtctatacat	caccacacct	gccaaatgga	65400
acttttttcc	cttgtagttc	agatttccaa	gaaagggat	attattgtcc	tagctcagcg	65460
tttcaacaata	gactcacaaa	atgaatcttg	ttaagcctat	gaatgggtaa	gacttatccc	65520
atacacagcc	tctattgtgt	aaaaaaaaatca	gctttattcc	taattccctg	tagtaggaag	65580
cttgcataga	ccacaataga	gacattgtta	gattatatca	agaaggtgag	tgccaggcat	65640
tcttgagtat	actggaattt	ttatfttcatt	gtagaaacac	agaatcataa	ttagtcatct	65700
aaaactgggt	ttaaggagat	aataactaaag	atactcactt	gaggtagatt	gttctcagaa	65760
ttacttaaaa	taagtgtctt	aaatcctgct	cattcatgaa	gccatagccc	tggaaggaga	65820
tattgaaaaa	cttctccatc	ttccttcaat	cccataggac	catgttttta	attgtagttt	65880
aacattaaac	ataatttttg	aaagttagcca	atattaatct	ttgacaaaaa	actcaaaacta	65940
aatagcagaa	aatagaacat	gccctctcct	ctcgcccaaa	agtaagacta	cattttaaag	66000
tatatcacag	tgatctgtct	aaatattfttt	ctcaaatatt	ttactaagaa	agcgtatatt	66060
cttctgtcct	tatttcatag	atctcttctg	tatcctgtgc	agcttgtcta	tcatttctgag	66120
cttttatatt	tcttattttg	ttctcttgct	gcttaaatgt	gattctttga	gatttataga	66180
agaatgtgat	tgftctcoag	ttttaaaaag	ctcttctcta	tactactaaa	tttccatgaa	66240
tttctttctt	tctttttttt	tttttgattg	agacagctct	gctgtgttct	ccagcgtggc	66300
cttaggctta	agtgatcctc	ctgtctcagc	ctcccaagta	gctgagatta	cagaaatttc	66360
catgaatttc	caaatatgat	gttatctttc	atataagctt	aaacataagc	ctttcttctt	66420
tcaaaaatgt	tagaacttac	ctctttaaat	aatcttataa	agccatttct	taatttttgt	66480
tacctttctg	aatctattaa	cttgcatatc	atagtaaatg	aaccttagc	tocatcatac	66540
cacacacctt	ctaactcctg	cctcccactt	tttctcacc	aggctggagt	acagtgaata	66600
tttagtgctt	ttaaatatta	ttcattacca	agccaagtag	cgtaactatt	tttctttgta	66660
cagccttttt	ttcccctaaa	taattgtgtt	atfttttcat	ttgcttcat	ttcatcttat	66720
cgattgataa	ttcttctgtt	ccatttccagt	ctaatttcca	catggccaaa	ccaatctaaa	66780
aacctgttaa	ctcaggtctt	tttttttttt	ttttttccca	gaataccttt	tcttggaaacc	66840
accttcttct	tttcatctga	gaccaaagc	tttccatgcc	tggttcaaa	tggttcttct	66900
ctaactgtct	ttcactacoa	tcataatttc	cttttctctt	tttcttctct	ttattatgta	66960
tctctgtct	tttaaattcc	atattttcca	ctttcatgat	ttattttctc	atfttgatgg	67020

ctacatcttt	cagtaactga	gaaaaggttg	tgtgagggaa	tattttgaga	tgctgtagtt	67080
ttgaaaatat	ctttattctg	ctgtcttctc	attaataaatt	tgactggatg	tcaagattta	67140
agttggaaac	tattttccat	gagtattttg	aaggcattat	ctattgtctt	ctagcttcca	67200
gtgctgcttt	gtattagaga	cagaagatca	ggaatcagga	tagcattgga	cttcttattt	67260
ctatttccca	aagctttttg	gatcttctct	ttattcttat	cttctgatat	tttataatga	67320
tatgcttgat	gcatttactt	tttcttttct	tcttccctct	tgtacatttc	tttctttttt	67380
taggtggggg	caggggggaa	ggagtctcac	tcttctcacc	caggctagag	tgcaagtgtg	67440
caacctcagc	tcactgcaac	ctccacctct	ggggttcaag	cgattctcct	gcctcagcct	67500
cccagtagc	tggtattaca	ggcatctgcc	accacgccca	gctaattttt	gtatttttag	67560
tagagacggg	gtttcaggct	ggtcttgaac	toctgacctc	aggtgatccg	ccctcctcag	67620
cctcccaaaag	tgctggggatt	agaggcatga	gccactgcgc	ccagcccctc	ttggaccttt	67680
ctagatattc	atccttaaat	tctgagaaaag	gttctataat	tctttaattt	aaaaaaagt	67740
tttttgtttc	atctttctca	cttttttttt	tttttttttt	ttttaagaca	cagtcttgct	67800
ctgtcaccoca	ggctggagtg	cagtggtgcg	atctcagctc	actgcgccct	tcgctcctg	67860
ggttcgagtg	attctcctgc	ctcagcctcc	caagcagctg	ggattacag	tgcccaccac	67920
caccatgtcc	ggctaattgt	tgtattttta	gtagagacgg	agtttcccca	tgttggccag	67980
gctgggttca	aactcctgac	ctcaggtgat	ccacctgctt	tgccctccca	gagtgtcggg	68040
attacaggca	tgagccaccg	tgcccagccc	tattctgat	tcatagatgc	agtgtctttt	68100
atgtctctga	gaggttttgt	caagttttct	tttgtttact	ttattatctg	tttcttcaag	68160
attctcctct	tttctttttt	gccttaagct	tttcatatt	ggaagcttct	ctcaaatgtc	68220
tggttattct	ggttgtccac	ttatattacg	tgaacaata	aaaaagtga	ttaggactct	68280
gtgctgtaag	taagtaaat	gttgatagtg	agagccatt	aatcacagga	tgatcaagca	68340
gcaagcctac	tgtttatttg	ggtttctcca	aatggctata	gctgtcagtc	tttttttct	68400
ggggttattt	ggctctctca	gaaaagaatt	ttctagtctc	ttgtttggag	aatacaagct	68460
tgactatttg	tgttccagaa	aatgggtggg	aggagacgac	ttgttttct	gttaggttga	68520
gccatgtgaa	attactaggt	tttgttttgt	tttctgtttt	gttttgagac	agagtctcac	68580
tctgttgccc	aggctggagt	gccttagcat	gatcttggct	cactgcaacc	tctgcctccc	68640
aggttcaagc	aattctcatg	cctcagcctc	caaagtagct	gggattatag	gcattgacca	68700
ccacgcctgg	ctaatttttg	tatttttagt	agagacgggg	ttttgccatg	ttgaccagtc	68760
tcagtctgaa	ctcctgacct	caggtgatcc	acctgccttg	gcctcccaaa	gtgctgggat	68820
tacaggtatg	agccaccaca	cctgggctga	aattactgtt	tttataggtc	aaaaacagtt	68880
gagggatagg	catggtggct	cccacctata	atcccagcac	tttcagaggc	caaggcagga	68940
agattgcttg	agtccaggag	ttcgacacca	gcctgggcaa	tgtagtgaag	cccactctat	69000
atttaaaaaa	aaaaaataga	tgaatatctt	tgtttgacga	taacatgatt	gcatatgtag	69060
aaaatcccaa	agaaccaaca	aaaagagctc	ctagaactaa	taagtgatta	tgacaaggtg	69120
tagaatacaa	agttaatata	caaagtccaca	ttgctttttt	atctaccagc	agtgaaccaac	69180
tggaatttga	aattaaaaa	caatactgct	gggtgcagtg	gctcacacct	gtaatcccag	69240
cactttggga	ggccgaggtg	ggcagatcac	tctagtctcag	gggttcaaga	ccagcctgac	69300
caatattggtg	aaaccccctc	tctactaaaa	atacaaaaat	tagctgggca	tggtggcggg	69360
tgctgtaat	ctcagctact	caggaggtcg	aggcaggaga	atcgcttgaa	ccaggaggc	69420
agaggttgca	gtgagccaag	atcgcgccat	tgactccag	cctgggggac	agagcaagac	69480
tccgtctcaa	aaaaaaaaac	acaatacctt	tcataatgag	actaataaaa	tgaatatgt	69540
gtagttctaa	caaagtgtgt	tgtagaagat	ctatatgaga	agaattatag	cactcatgaa	69600
agaaaatcaa	gatctaagta	aactgagaga	tattccatgt	aaatggacag	ggagactaaa	69660
tattattgag	atgtcagttc	ttcccagtt	catatatcga	ttcagtgag	tcccagtcaa	69720
accccagcca	gttattttgt	ggatactggc	aaactaaagt	ttatatgaaa	aggcaaaaaga	69780
cctagaacag	ccaacacagt	attgaagaag	aaaaaagtca	gaggactgaa	actacccaat	69840
ttcaagactt	actgtaaagc	tacattaatc	aagacagcat	gtcattggca	aaagaataga	69900
caaataaatc	agtagaattg	gacagagagc	ctagaatca	accacacacg	ataaagtcaa	69960
ctgatctttg	gcaaagggac	aaagacaatt	cagtggagaa	aagatagcct	tttcaacaaa	70020
tggtatagga	caactggaca	tccacatgca	aaaaagttaa	tctagacaca	gacctgacaa	70080
ctttcacaaa	aataaatgga	tcatagacct	aaatgtaaca	tgcgaaactga	aacttctaga	70140
ggataacata	gggaaaaatc	taggataaca	tgagaaaaat	ttttggtttg	gcagtgactt	70200
cttaggtaca	ataccaaaac	atgatccttg	aaaaaaaaaa	tcagtatgtt	gaactttgtt	70260
taaattataa	acttctgctc	tgtgtaagat	gctgttagga	gaatgaaaag	acatgcagca	70320
gagtggggaga	ctttacaaaa	ttcattatct	gatgaaaggac	cagtatccaa	aataacaaa	70380
gaacttttaa	aactcaacaa	taagaaaata	tacaaccag	taaataaatg	ggcaaaatat	70440
ctgaacagac	acctcaccaa	ggaagataga	tagatgacaa	caagcatatg	aatatatgct	70500
caacatcatg	tcgttaggga	aattgcgcat	taaaaacaaa	acaagatacc	ctgcatcccc	70560

tattagaatg	gctgaaatct	aaaacactga	caacacccaa	ttctggcagg	gatgtggagc	70620
agcagaaact	ttaattcatt	gctgatggaa	atctaaaatg	gtagaaccat	tttggagggt	70680
agttggacag	ttttttacag	aactaaagac	agtttgacag	tttcttacaa	aactcttacc	70740
atatgggtcca	gcagtccttac	tccttagtat	ttaccccaat	aagtttaaaa	tgtacatcca	70800
ataaaaaaaaa	ctgcacatga	atattttctag	cagcattatt	catagttgcc	aaaacttggg	70860
agcagtcacg	gcacccataa	gtaggggaa	ggataaacag	acttttggtat	atcatgtaat	70920
gtagtattat	tcagcaataa	aaagaaatga	tctatcaagc	cacaaaaata	tatggaggaa	70980
ccataaatgc	atattgctaa	atgaaagaag	ccagtcctgaa	gaggctacac	tataggattc	71040
tgaactatag	atgtttttgga	aaaggcaaaa	ctatggaaac	agtaaataga	tcagtggttg	71100
ccaagggaga	cagggagaga	tgaataggtg	gagcacagtg	gatttttaag	gcagtgaaac	71160
tgttctttat	gataatccaa	tggtggatac	atgtcattat	acctttgtca	aaaccacag	71220
aatataaaac	ataagagtga	accctaagt	aaaatatgga	cttcagttaa	taataatata	71280
tgaatatttt	ttcattagtt	ctaacaagt	tactacacta	atacaagata	ttcagagtga	71340
gggaaatttg	aaaggaatga	gaggttatat	gggaactctg	tactttctgc	tcaattttct	71400
gtaaacctta	aatcaactaa	atttttat	atttttat	ttattttttt	ttaattttta	71460
attttttgag	atggagtttc	actcttgtg	cccaggctgg	aatacaatgg	cacgatctcg	71520
gctcactgaa	acctccgct	ccagggttca	agcgattctc	atgcctcaac	ctcccagta	71580
gctgagatta	caggcattgt	ccaccacgcc	cagctaattt	tgcagtttta	gtagagacag	71640
ggtttctcca	tgttggtcag	gctggtcttg	aactcctgat	ctcaggggat	ctgcccgcct	71700
cggcctccca	aatagctgtg	attataggcg	tgagcagctg	cgcccagcag	gttttttttt	71760
ttaagttgga	tattagccat	ttcatatgat	tcaacttaaa	agtacataca	ccttcactgt	71820
tattaagtg	taggtgagat	gttttcagtc	tggagctcta	cccttgattc	ctgccatgcc	71880
tagtgcctct	gaatctggag	actctgactt	atttctttag	agaacgaaac	tcctgccttc	71940
tgccttggtg	gtgattggtg	cctgcttgac	tgcttttgg	ggggagttoc	tcataccaac	72000
tttcaatcag	ttcctgtgtt	tgtttgtttg	ttgtttttg	agacggagtt	ttgctctgtc	72060
gcccgggttg	gagtgacgtg	gcgcaatctt	ggctcaatac	acgctccgcc	tcccgggttc	72120
acgccatttt	cctgcctcag	cctcctgagt	agctgggacg	acagggcccc	gccaccacgc	72180
ccggctaatt	tttttgtatt	tttagtagag	acggggttct	atcgtgttag	ccaggatggt	72240
ctcgatctcc	tgacctcgtg	atccgcccgc	cttggcctcc	caaagtctctg	ggattacagg	72300
cgtgacgact	gtaagccacc	gtgcccggcc	agtctctctg	tttttgactg	cctgccttac	72360
tactgctttc	tgtggtgcct	gataacatcc	aattcctgaa	ccttcctggg	atttttgttc	72420
acgtcagcgt	gctctctgca	tttaggtatc	ccttcacaag	taggcattta	ggttttaagc	72480
tctgctaagt	gatttaccac	actttatcag	ttctccattt	tgtggaattc	attgtgttaa	72540
tctccttttc	tatttttgtt	tggagatttc	atataatttt	tattcattta	gtattttttg	72600
gtgggatttc	taagtagaga	aacatgccta	tgttcaatat	gtcttgttta	agcagctctg	72660
tttcaatttt	accaccgagg	agttggtttt	atttcttttt	tttttttttt	tttttttttt	72720
ttttgagaca	gggtctcact	ttgtcagcca	ggctggagta	cagtggcaca	atcaaagctc	72780
actgcaacct	tggcctccca	agctcaagt	atctcccac	ctcagcctcc	tgaatagctg	72840
ggagctcagg	tgatccagca	caccacagct	attttttttt	tttttaattt	ttttagaag	72900
cagagttccc	ctatgttgcc	caagctggtc	ttgaactctt	gggctcatgt	aatcctcctg	72960
tctttgcccc	ctaaagtctg	ggattacagg	tgtgagccac	cacaccagc	tgtttttaat	73020
tattacatta	atltatgatt	atgtggttcc	tttaaagcta	tgggcagctc	tgtttgggaa	73080
gatgctactg	aaaaaaggat	gctttttaga	tggcaagag	tacttaaaat	gtctctagaa	73140
tcatagttgt	aaactaaagt	atataatttt	tagttgttca	cagggcttag	agctcctgcc	73200
agagggtctg	tactctttgg	tccacctggg	aatgggaaga	caatgctggt	aagggtctct	73260
ttcaaatttg	agttttctgt	tgagatattt	gggataatat	gaaaaaaga	aactttatct	73320
tgtccttgag	tctattattt	acgacttget	ttttgctatt	gtacactttt	gttttttttt	73380
tttgtttgtt	ttgttttgtt	ttgagtgatc	tgggctcacc	gcaacctccg	cctcctgggt	73440
tcaagcgatt	ttcatgcctc	agcctoccaa	gtagctggga	ttacaggcac	acgccaccac	73500
acctgtctag	gttttttatt	ttggcagag	acaggatttc	gccatggttg	ccaggctggt	73560
gtcgaactcc	tgacctcaag	cagtccctct	acctgggctt	cctaaagtga	tgggattaca	73620
ggcatgagcc	actgcacctg	gctgatacac	ttttaagttt	ttcagctact	tttcaatgta	73680
gaagtagatg	gaaaaccatg	tacgttatct	tcagtagtgt	gtttttggtt	ggttaaattt	73740
gacagtatga	ttgtcattat	tttttgtaaa	ttaaattttt	acctggaaga	gcttacctta	73800
ctatattgag	tatctttcta	accctgatt	tttgcctcta	ctatcataat	aactttattt	73860
aagtaatcag	atgtttatag	cttttttttt	tttttaagta	ttcttttggc	agaagttttt	73920
atcaggctct	ggataacctct	ttcctctgca	tagtctctct	ggatggaaga	aacaaagagg	73980
gaaagagtaa	cttttcttta	gatgtttgtc	tttctcaaag	cagttatctt	tgtatatcta	74040
agaagagagg	agaataaac	tgtctctctt	tttttttaaa	tctctctcta	ctcattctct	74100

ctcaggagga	gagtagaaag	aagcacagct	cttcctataa	cctgtcotta	ttactgagaa	74160
aggaacacat	tgattgccat	gtattgggga	ttgtattata	ccttacattt	ttatTTTTat	74220
TTTTattaa	TTTTTTTT	tagacagtct	tactctatca	cccaagctgg	agtgcagtga	74280
tgtgatctcg	gctcactacc	atctcttctc	tctgggttca	agcaattctg	ccacgtcagc	74340
ttcctgagta	gctgggatta	caggeatgcy	ccaccacgcc	tgactaattt	ttgtatTTTT	74400
attagagatg	ggatttcacc	atgttgcca	ggctggtctc	aaactcttgg	cctcaagtga	74460
tgccctcg	tcggcctcog	aaagtgcgg	gattacaggc	atgagccacc	acacctggcc	74520
tcatagctta	cattttttaga	gaatctttc	tagtacttaa	atcggtaaat	atggttatct	74580
tttaaagtga	atatattgaa	ctaatttaa	atTTgtctt	gtgattttta	aaggctaaag	74640
cagttagctgc	agaatcgaat	gcaacctctc	ttaatataag	tgctgcaagt	ttaaacttcaa	74700
aatacgtgag	tgctctgttt	ccaatattgt	cgatttttaa	gttactgtct	aaatgttact	74760
gtgttaactg	taaatggtaa	tatttcatga	aaatattttt	ctaggagctt	atctattgta	74820
tctattatTT	acatatgatg	aatatctatc	ttcagagtag	aaagtattgt	acatttTgtg	74880
tgtcaaatac	tgtattagtt	tactggggcc	atgtaataaa	ataccgtaaa	ctggctggct	74940
taaacatcag	caatttattg	tctcacagtt	tcggaaggta	gaagtctaag	atcagtcaaa	75000
atgttggcag	gyttgcttcc	tctgagggc	tgtgagggaa	aatgtatgtt	gtgtgcctct	75060
ctcctggctt	ctggtggctt	gctggcaatc	TTTTgtattc	cttggcttgt	agatgcatcc	75120
ctcctatctc	tgtctttatc	tttatgtggc	attctcctcg	tgtctgtcac	catgtccaaa	75180
tttccccgtg	ggattaggac	ccacccta	gatctcaatt	taagttgtc	taagcaaca	75240
attctatgtc	caaataatgt	cacattcata	gctactaggg	ataggcttc	aacacgtttt	75300
tgagaacac	agttcaacc	attaacaa	actatcactt	tccacttaag	cttcaagtaa	75360
agtggatttt	atctcaagga	gccaccagat	aggaacacag	atctgatggc	ataaactgag	75420
tatttctggc	cttctgattc	tgatctaaaa	tatgacagaa	gattttccct	gTTTTaattt	75480
TTTTTTTT	ttataatag	agacacagtc	tcgctatggt	gccaggctac	tttcaaatc	75540
ctgggcttaa	gcaatcctcc	tgccctggac	tctcaaagtg	ctggaattac	aagtgtgagc	75600
caccacacct	ggccagtaga	ttttccctgc	tttcttttga	ttgtttataa	ttttgttttc	75660
TTTTccatt	cacctctgc	tgaccctata	gtattattca	aagaagtgtt	cagtctagtt	75720
ttgggtagg	gcaagcataa	ctacagtgtc	taaagagagt	aatttgcctg	gtgtgcagaa	75780
ctagtgtgta	aatataactg	gtgcattgca	aaactgtgaa	gtagttctcg	tcaaacctta	75840
cactgctttg	tctttctccc	tctctcctt	tctcttggtt	gccccctccc	cctcccaatg	75900
ataccttagt	ctctgcttgc	ctttatcaaa	acctttatga	ttggccgggc	acagtggccc	75960
atgcctataa	tcttagcatt	ttgggaggct	gaggcaggaa	gatcacctga	gcccagaagt	76020
tgagagaccg	cctaggcaac	atggtgaaac	cccctctctt	ccaaaaataa	aaaatagagc	76080
caggtgtggt	ggcatgcacc	tgtagtcca	gctactcagg	aagctgaggc	gagaggatct	76140
cttagcctg	agaggtttag	gctgcagtga	actgtgatca	tgccactgca	ctccagcctg	76200
agtgacagtg	tgagaccctg	tctcaaaagaa	acaaaacaag	ggggggcatg	gtggctaact	76260
cctgtaatcc	cagtactttg	ggagactgag	gcaggaggat	tgcttgaggc	caggagttca	76320
ataccaccct	gggcaacata	gtgagacccc	catctctaca	aaaaataaaa	aatttagctg	76380
gacatgccag	gcaatacgtg	gtcccagcaa	atcaggaggc	tgagggtgga	ggatcacttg	76440
agtcaaggag	gttgaggttg	cagtgcagca	cgatcatgcc	aatgcattcc	aacctgggcy	76500
tcagagcaag	accagctctc	aaaaacaaa	caaaacaaac	ttttatggtt	gaaagtgttt	76560
tggaacaact	acttaaactg	aaatgtgaat	ctctgatgaa	agaacatggt	acctgtaaaa	76620
gtttgaagtg	tcagcatttg	ttgcaccgaa	atccagaggt	gaggccaggt	gtggtggctc	76680
acgcctgtaa	tcccagcact	ttgggaagcc	aaggctggca	gatcacctga	ggtcaggagt	76740
tcaaaaccag	cctggccaac	attgcaaaa	ccogtctcga	ctaaaaatac	aaaaatttagc	76800
cagacttgg	ggtgtgcgoc	tgtaatccca	gctattcggg	aggctgagac	acgagaattg	76860
ctcgaacc	ggaggcagag	gttgcagtga	gccgagatgg	caccactgca	ctccagcctg	76920
ggcaactgag	tgagactctg	tctcaaaaa	aaaaaaaaaa	aaaaatccag	aggtgaatcc	76980
agaggtgatc	accacatgat	atccagatag	cctctttcca	tgagaggctc	aaaggataat	77040
tttactgtct	acagttttgc	agcagagaaa	acttgatttt	atcagtacac	caagagcaga	77100
tctatgtctt	cggaaacagac	atgagatcag	aattgtctag	ctgctatgaa	cagcatgttc	77160
tctccctgta	cctatagaca	tgtatgggaa	acttatttgt	aaggttgtat	aatgagcagt	77220
gagttaaagc	aaacttgaca	tgttgaccat	agttgttatg	gcattggact	aaagtagcct	77280
ccatcactat	gatagagata	gccttgggtg	tggaaccata	gtttttgaa	tctgtttgct	77340
aagactccct	cttcccttagg	cagtcttca	aatattatta	tgctttttcc	tgtcatggga	77400
ttcctcttca	tggagtttcc	tttctcagga	acactatctt	ctcctggtta	atTTgtactt	77460
attcttcaga	taccagttga	aatgttacaa	ccttccctga	ccctccaaac	tattccccgt	77520
tattattctc	ctagaccac	ttgcacctca	tttctactct	cactggagtt	gcaattcata	77580
ttcattgata	ggattatTTT	aattgtatct	gatatactg	tcacctccac	tagaagatgg	77640

tctccatgcg	ggcagagact	atcaccatgt	gttcttcaact	tcaattttca	gtagttggct	77700
gtgagtaggt	attgaataaa	tatttgtgga	gtaatcataa	tgaggtatag	atattattct	77760
catagctcta	ttttattaat	tagggaatta	cagaattcag	tgatctgctc	aggatctcat	77820
aaccaggaag	tgggaaacta	ggatttgagc	tccagtgagt	gtggcctttc	atataaaata	77880
ttacagcaac	cattctcttt	tttttttttt	ttttttttga	gatggagtct	tgctctggag	77940
tgcaagtggc	tgatctcggc	tcactgcagc	ctctgcctcc	cgggttcaag	tgatccccct	78000
gccttagcct	gctgagtagc	tgggactaca	ggcaccgcc	acgacacttg	gctaattttt	78060
gtattttagt	agagacgggg	tttcaccatg	ttggccagga	tggtctcgat	ctcctgacct	78120
tgtgatccgc	ccgcatctac	ctcccaaaagt	gctggattat	agggcgaagc	caccacacc	78180
ggcctacagc	aaccattctc	ttttatccat	acttttttca	agagtactgt	ttcatcttca	78240
atcttgaagt	aacaacatag	cattcatgat	cttaaccctcc	aattctgata	ctgcctgaat	78300
atcccagcac	aagtttactt	ttaagaaagt	tgaggctagg	tggtgtggct	catgcctgta	78360
ggctgggcaa	tttgggaggc	caaggcaggt	ggatcacttg	agctcaggag	ttcaagacca	78420
gtggcgtgtg	catggcgaaa	cctgtctct	accagaaata	caaaaaatta	gtcgggcgtg	78480
ttggaggttg	cctgtggctc	cagccacttg	ggagactgaa	gtgggaggat	ttcttgagct	78540
agagtggat	agtttgetgt	gagccgagag	atcatgccac	tgactccag	cctgggtggc	78600
attttgtcag	cccattctca	aaaagaaaa	gaaaaagaaa	attgaaatgt	ctagctctatc	78660
taatttttac	ttctatctaa	tacaattttt	tccttatgtc	taactgaaat	ctgctttttc	78720
ttcaaatata	atacttgatt	taacaaaaact	caactttttt	ttttttttat	gagacagcct	78780
attttagttg	taggaactta	atggtatatac	tgcttccctc	cagtccccag	aatagttact	78840
atattttacc	tctttctatg	gtctcatgcc	agtttgtcag	tatgcctaga	taagaactga	78900
gcagtggtct	tcagatgtga	cctgactttg	aagactttaa	aaggaaagcat	tggtccaggc	78960
ttgagaccac	acacctgtga	tcccagcact	ttgggaggcc	gaggtgggca	ggtcaggagt	79020
actaaatact	ccatgattac	atgggaaaac	cccatgttcc	tctactaaat	acaaactctc	79080
acatgcctgt	aatcccagct	actcgggagg	ctgaggcagg	gatgagctgg	gcattgtggc	79140
ggcggagggt	gcygtgagcc	gagattgcgc	cattccactc	agaatcactt	gagcccagga	79200
aaactccatc	tcaaaaaaaaa	aaagcagcag	catgtgttaa	cagcctgggc	aacaaaagtg	79260
tcctgatcta	tccctgagaaa	gcttttggga	gcaactgcat	tattatgtag	atgttgtgtc	79320
gtgttattaa	aataatctaga	ttattttcca	caaaaaatca	catagtcatg	gacaactatt	79380
tggtgtatta	ttgtttaacc	ttgtttatg	aataactaac	gttacatag	tatcttaaca	79440
cataagtata	cagctgtata	cagctcagtg	gattaccaca	atgtagaaaa	gtatttatag	79500
tcaccaccca	ggtcaagaaa	taaattgtta	cctgtggccc	aagcgaatat	actttcataa	79560
caccatcttt	atocactcac	tcctctccct	caaaaccact	taaaatccct	ccaggcactc	79620
aaagctagca	tgcccttgaa	ctttatataa	atctaagat	agactactaa	catcatagac	79680
gtatttggct	tctttcatca	gcattgtatt	tgtagagatt	gtaggatttt	gtgtgtatgt	79740
ttgtagttgt	gcttttttac	acagatttaa	tttttataat	atccagattg	ttgcaagtag	79800
ttaaacctgct	tatctattat	taaaaaaaaa	aaaacgaaac	tttcttatat	gttcgaacag	79860
ttacagtggt	ttttccacac	cgtcttcaaa	atgtaaagtt	ttcacatagt	tcttaccagt	79920
tgctgctaga	tttaaaact	agggaaaaaa	aaatcagaga	tggtcttcaa	tacatcagta	79980
acctccacta	gaagatggtc	tccatgtgga	agtaaatatt	agttaataat	attaactgtc	80040
aagtaaaatc	ttctgtaagt	ttctttaaat	attttaataa	tctcttgtat	tatctgtgtc	80100
ctcaatattc	taaagttagtt	aaaagtaact	ataaaatagt	atcatagtac	ttaaaatggt	80160
tttacttctc	atgtgaaatt	ttacaagtcg	ttactctatt	acctgttttt	ctgatcacat	80220
aagacaggg	ctgttctgtc	gcccaggctg	tgactcatt	tatttatgga	tttatttttt	80280
agcctcagcc	tcctgggctc	aagtgatcct	gagtaacagt	cgtgatcat	ggctcactgc	80340
ataggcacat	gcccaccatgc	ccagctaatt	cccaccttag	catcccaagt	agctgggact	80400
gctcacacct	gtaatcccag	cactttggga	ttaaaaaatt	ctggggggcc	gaatgcggtg	80460
gttcgagacc	agcctggcta	acatggtgaa	ggccgaggca	ggcgaatcac	aaggtcagga	80520
agcgggggtg	ggtggcaggc	gcctgtaatc	accctgtctc	tactaaaaat	acaaaaaatt	80580
ttgcttgaac	ctgggaggca	gagattgcag	ccacttactc	aggaggctga	ggcaggagag	80640
attttttttt	gtagtgacaa	ggtgtcactg	tgagccgaga	ctccatctca	aaaaaaaaaa	80700
ctcaagtgat	cctcccattt	cggcctccca	tggtgccagg	gctgggtctca	aacttctggg	80760
tgccctggtct	tcaagttggt	attaagcat	aagtgttagg	atcacaggca	tgagtcactg	80820
gaaagtattg	ttgtggaagt	taggagatag	gtttaccacc	attatgcaca	tggtataatg	80880
gacaaggtct	cactttttcg	ccccaggccg	ggattctagc	ctagcttttt	atfttttttg	80940
aacctccaac	tctcaggttc	aagcaattct	aagtgcagtt	gtgcgatctc	ggctcactgc	81000
acaggcatgc	gccaccacgc	ccggctaatt	cccacgtcag	cctcccaggt	agctgggatt	81060
catgttggcc	aggctggtct	tgaactcccg	tttgtagttt	tagtagacac	agggtttcac	81120
			acctcaggtg	atccacccac	cttggcctcc	81180

caaagtgctg	ggattacagg	catgagccac	cgccccggc	ctctagcgtg	acttttacat	81240
cctgaactga	ccttaagaaa	gtataacttt	aggcctgttt	catctgtaaa	atgttaatgt	81300
cataggagat	gatcttttga	gatttctttc	agctctgata	attttgtgtg	tgtgtgtgtg	81360
tgtgtgtgtg	tgtgagatgg	agtcttgctc	tgtcgccggg	gctggagtgc	agtggtagca	81420
tctcggctca	ctgcaagctc	cacctcctgg	gttcacgcca	ttctcctgcc	tcagcctccc	81480
gagtagctgg	gactacaggg	gcctgccacc	acgcctggct	aatTTTTTgt	atTTTTagta	81540
gagacgaggt	ttccaccgtg	tagccaggat	ggtctcgatc	tcctgacctc	gtgatccgcc	81600
cgccctagcc	tcccaaagtg	ctgggattac	aggctcgagc	cgccgcgccc	ggcctagaat	81660
atTTgtttcg	attattctaa	atctggtgac	atTTctTTTg	ttTTtaagtt	aaatcttcag	81720
taaaaagaat	aatgcccacc	tagaggacag	aaaaatTTTT	acagtagatt	atcacagacc	81780
tcatgactca	ttactttggg	gtataaaatg	gcctttgtat	ggtgtcagca	cctgggaaatg	81840
tctcaagggg	tgttcattga	ccttctggac	tatctggaaa	tatTTtgata	tttattggct	81900
gggcgcggtg	gctcatgcoct	gtaattccag	cacctcgggg	ggccaagtta	ggtggatcac	81960
ttgaggtcag	gagtttgaga	tcagcctggc	caactaagtt	agttgactat	ttgtgaaatt	82020
tccctctcca	aatagaaaag	ataggattgg	aaacattatt	cagaaggaaag	aagTTTTaaa	82080
gaagggcagg	cttaaagact	atctaatagaa	aatgacaata	tggtaaaacc	ccatctttac	82140
taaaaataca	aaaattagcc	aggcgtgggtg	gtgggcgcct	gtaatccag	ctacctggga	82200
tgttgagggg	tgagaatcac	ttgaaccag	gaggcaaaaag	tcacagtgag	ccgagattgc	82260
accactgcac	tccagactgg	gctatagagc	aaaactctgt	ctcaaaaaaa	aaaaggaaaa	82320
aaaaaaaaaga	aatTTtgata	tttatgtgag	aactgactttt	cacggtgttc	ttaatagcgc	82380
aagTTTTgtt	taggagagca	cattccaact	tacttgcttc	tataaatata	ccgtgtaatc	82440
tagggcttaa	agaatatgta	caatgtcttt	ttctctcccc	tagtcttccc	ctttctctcac	82500
tagttatata	cttgactgaa	gagctatttc	attctcaagt	cttaggaatg	cagggtgaag	82560
caaaacagat	gaatTTTTaa	agcacttaac	cagcctgtat	gaaatcacag	tctgttgtct	82620
aaaattgtaa	gggacgggta	gtagtactct	cccctttctc	aaaccaaatc	tttggttggt	82680
ttaaggaaagg	gaaattaaat	tcctgtgtgc	tagatTTTca	acataaaatt	taaaaaactg	82740
gaataatggt	gcattTTtatg	tgtataacag	tataatgctt	tgtTTtaggt	gggagaagga	82800
gagaaattgg	tgagggctct	ttttgctgtg	gctcgagaac	ttcaaccttc	tataatTTTT	82860
ataggtaaaga	acatattttc	caactaagtt	attgactatt	tgtgaaattt	ccctctccaa	82920
ataggaaga	tacgattgga	aacattattc	aggaggaaga	agTTTTaaag	aagggcaagc	82980
ttaaagacta	tctaataaat	ttagtaggac	ccactatatt	aataagtagt	aaactagatt	83040
aatctcagat	gactcacata	gcttggtctt	taattaaagt	cttatacttg	tatttctctc	83100
agatgaagtt	gatagccttt	tgtgtgaaag	aagagaaggg	gagcacgatg	ctagtagacg	83160
cctaaaaact	gaaattctaa	tagaatttga	tgggttaagt	gttgattatg	atattTTTaa	83220
tgtggcagca	ttttagata	ttttctctatt	aatggccaa	ggttaaaaaat	acaaatatct	83280
ttatatttgt	tattactttt	ctaaatgaat	tgaaaaaaga	ttttttgctt	gtaggtacag	83340
tctgctggag	atgacagagt	acttgtaatg	ggtgcaacta	atggccaca	agagcttgat	83400
gaggtgttct	tcaggtaggg	agatttatat	ggaatatcat	gcatttatta	cagacaata	83460
ttactcatgt	gtccatctta	catattattt	ccttactctc	agTTTTaaga	ctaaattcac	83520
tattttcttc	cagtactatc	tctagcctct	tgttaccaac	tacataagga	ttttgaggtc	83580
aacagcttga	tatcaggaga	attagctctag	taaaggttta	gttacggttt	taggcaaatt	83640
ggtcagttct	tctgggttgc	attaaattat	ctttaaactt	gaaaattgat	tcttacogtc	83700
ttttattggg	cttcatgaga	aaataatcaa	atcagggttc	atggaaaatg	tgttttctgt	83760
tcttggtgtg	ttgagcagtt	tattcccttc	atactggaag	gacggtttgg	gtcagcataa	83820
aattcttggg	catcctttcc	caaggaattt	acaatatgac	tcattgtttt	ctagattggc	83880
tgtagtagtg	taaaagtctg	aatccaaatt	gattTTTTca	aatTTgtaa	ataacttgat	83940
tttttctct	ctttactgga	tgggttacgt	gggtatattg	caccaggtta	gtgagcatag	84000
ggtcgagttag	gtagTTTTtc	aaccacacc	tcactgcttt	cttccccat	gtagtagtcc	84060
gcagtgtgta	ttgttcccat	gtttatgtcc	atgtgtactc	aatTTTtagc	tccccactta	84120
tgagaacatg	agctatTTTg	tttactgttc	ccacattaat	tggectagta	ttaggcctc	84180
cacctccatc	catgttgctg	caaaggacat	tatttcattc	ttttttatgg	ctgcgtagta	84240
catagtatct	catggtgtgt	acgtaccaca	ttttctttat	ccagtccacc	attgatggac	84300
acttagattg	attctatgtc	tgctgtcatg	aatagtagac	caataaacat	gaaatgcatg	84360
tctttttggg	ataatgatct	atTTtctttt	gggtatatac	ccagtaatgg	gatttcaggg	84420
tcaaaggttag	atTTgtTTTa	tgtTTTTTga	gaaatcttca	aactgctttc	tacaatgact	84480
gaaactttaca	ttcctaccag	cagtacatac	gctccacaac	ctcaccaca	tctgttattt	84540
ttttactttt	tccgcacaat	ctgtagacaa	tatgggattt	tttttgccat	tttattaata	84600
gccattctga	ctgggtgtgag	atatctcatt	gtgattTTTga	tttgcatTTT	tctgatgggt	84660
agtgatgatg	agcattTTTT	catgTTTgtt	gtccccctcg	atgtctTTTg	agaagtgtct	84720

atcatttcc	tttttacc	ttttttaata	gggttgtgtg	tttttagctt	gttcaattgt	84780
ttgaattcct	tatagattct	ggatatcaga	ccattgttag	atgcagttct	gtagattgtc	84840
tgtttactct	gttgatggtt	cttttgcctg	tcgagaagct	cttttagtta	attaggtccc	84900
acttgtcaat	ttttattttt	gttgcaattg	tgtttgggga	cttagccaaa	atctcttct	84960
gaggctgatg	tcaagaagtg	tatttcctag	ttttcttcta	gcatttttat	actttgagat	85020
cttacattta	aatctttaat	ccaccttgag	ttaatttttg	tatatgggga	aaagtaaggg	85080
tccagtttca	tttttctaca	tatggctagc	cagtaatccc	catgccattt	attgaatagg	85140
gagtcctttt	cctcattgct	tgtttttgc	agtcttaca	aagatcagat	ggttataggt	85200
gtgtggcttt	atcttcaat	tttctgtcct	gttctttgg	ttgtgtgtc	tgtttttata	85260
ccagaattat	gctgtttggg	ctgctgtgta	gttcggttta	aagtcaggta	acgtgatact	85320
tacagctttg	ttctttttgc	ttaggattgc	tttgtctatt	tggactctt	ttttgcttcc	85380
atatgaattt	tagaataatt	tttctaatt	ttgtgaaaa	caacattgat	agtttgatag	85440
aaataccatt	gaatctgtaa	attgctttgg	gcagtatagt	cattttacga	ccagataaat	85500
caggcaagag	aaggaaacaa	aaggcatcca	aataggaaaa	gaagtcacga	tactctcact	85560
cttcaactgg	aatatgattc	tatactttga	aaaccctaaa	aattccgtca	aaaggctact	85620
agaactgata	aatgatttta	ccccagatt	caggatacaa	aatcagtgta	caaaaaatta	85680
gtaccatttc	tatacgctaa	taatgtccag	gctaagagtc	aaatcaagaa	cacagtccca	85740
tttacaatag	ccacaaagaa	aatgaaataa	tggtgggca	cagtggctca	tgctgtaat	85800
cccagcactt	tgaggagctg	aggcaggtgg	atcacctgag	gtcaggagtt	cgagaccagc	85860
ctgaccaaca	tggtgaaacc	ccgtctctac	taaaaaatac	aaaaattagc	caggcgtggt	85920
ggtggggcgc	tgtaatccca	tctgtctggg	aggctgaggc	aggagaattg	cttgaaccga	85980
ggaggcatag	gttgcaatga	gcccagatcg	tgccaactgca	ctgcagcctg	ggtgtgaaag	86040
agcaagaatc	cgtttccaaa	attaaaaaaa	aaaaaagaaa	aaaaaaaaga	caatggagta	86100
cctaggaata	cggctgatga	aggaggtgaa	agaatctgta	tgaggagaag	tgtaaagcac	86160
tgctaaaata	aatcagaggt	gacacaaaata	aatggaaaaa	cactccatgc	tcatggattg	86220
gaagaaatcaa	tatcgttaaa	atgggaaatt	tatcttaatt	ttttcccta	gttcattttt	86280
ctgttttctt	ttgcacataa	attatgtgtt	atgttgcctc	ctttgtttt	ccctatctgt	86340
aattttcttc	ttcatccttt	taaaccttta	ggtctttttg	ttttgttctg	ttttcccaag	86400
cctgttgcct	ttattcttta	ctgggttagt	atctaactct	ctttctgcta	tatctttgt	86460
aactcttctt	ctcttgggtg	tttttatttc	attactttcc	tgacatctgt	tattctactt	86520
ttcatcttct	gttttatttc	tttgagcatc	tgtatgttat	ttccataagc	tcttacattt	86580
cactgggctc	ttgtttcata	gaaaaaactt	ttttggcaag	gacagaccac	ttgaggccag	86640
gagtttgaga	ccagcctggc	caacatagta	caaccctgtc	tctgctaaaa	atacaaaaaa	86700
ttagctgggt	gtggtcacac	acacctgtaa	tccagctac	tcaggaggct	gataatcgct	86760
tgaaccaggg	aggcagaggt	tgcaagtgagc	caagacggcg	gcactgtact	ctagcctggg	86820
ggacagagca	aggctctgtc	tcaaaaaaaaa	aaagaaaaaa	ttttttggct	aggcgtcatg	86880
gtgggtcaca	cctgtaatcc	cagcactttg	ggaggccaga	tcacttgagg	ccgggaattc	86940
cagaccagcc	tgcgcaactg	ggcaaaacc	catacttata	aaaaaaactt	aaggataaaa	87000
aaaattaaat	ttttaaaaaa	ttaaattatt	ttatttctgt	gaaagatact	tggtcataat	87060
tttagtttgc	tctatggcaa	aatcttctca	tgcatgttct	tcattttatc	tagtttatgt	87120
ttattgtgtt	ttgtttgctc	ttttttgggg	gcaggttttg	ttgttgttgt	tttttttttt	87180
ttttttttaa	atgagccctg	tcctgggttt	tccttttttg	cttattattt	ttctttgaa	87240
aaaaggagtt	ttcttggaaa	acctagtttt	ggaaggatat	tgtaggggaa	gggataggt	87300
tcttgagtag	tccatgtata	cactaaacta	gcatgcagcc	tcctcactt	acagtgaaag	87360
cctacctaata	aatgaagtat	atgtgtatgt	aaataagaga	tagggttgat	ttattcttct	87420
actaaaaaat	attggctcact	gcaagacatt	gtcttcccgc	cagactgtta	gatgccttct	87480
tataaagata	attcgtctta	gagtttcttt	gttcattctc	accttctctg	ccattcctat	87540
gctatcaatt	agggtttata	tggtctgcag	agccaggctt	catatgcttg	tttaaatgtg	87600
gagggaactgg	ccttattacc	ctttaggggtg	tgctatctct	ttttgggagg	aaaactgtac	87660
tctgacagtc	atgtccatga	tcttttagogg	cagatccttt	gtgtgtcttc	ctttgtcagt	87720
gtaattttca	ctgccctagg	cagttcacct	ttatgtattt	tagttccata	aatgttacct	87780
agttttgtga	tgaagtgtat	gtttctgttt	ctattatcct	tgctgatttt	agtttttttc	87840
agattcaaga	agacagatct	acttatatca	gatattttgt	ttgttttaaa	gcattactgc	87900
tattaaaaaac	atatacatac	ataaatgatt	aattctaaca	ataccgaatt	aaggatttta	87960
aacattaact	agagttttaa	aagagtaaca	aaaatatttt	tacattgata	actacaaaaa	88020
tgcatccag	tgccctgaa	attatatttt	aaaagctttt	cctgtcattt	gctgtttcag	88080
ctttaaattc	aaaattatat	ttctaaaagt	gctgtatttt	tttttttagg	cgtttcatca	88140
aacgggtata	tggtgtctta	ccaaatgagg	aggtatgtat	ctgtgtttga	attttttttg	88200
tttttagagca	gaaacaagaa	ctaccatctt	gacaatatta	agtcttccaa	tccatggtac	88260

agctactttg	gaaaacagtt	tagtggtttc	ttaaaaagat	gtacataagc	ttactgtcag	88320
agccagcaat	ttcactccta	agaatttacc	ccagagaaac	aaaaatgtat	gtccacacaa	88380
agacttgTac	aagaattttt	atagcagcaa	tattaataat	agccagaact	acaaatgatc	88440
taaaatgttt	atftttgtga	ataaacaaat	gtggTatatt	catttaatgg	atactgttat	88500
ccattataca	tgctgcatac	tcacttcata	tattaagttt	tatgaatgaa	actccaaaac	88560
attataagtg	aaagaagcaa	gatgatatat	gtcgcatac	ataggattct	gtttatatgc	88620
agaaaaacgca	aatfttattt	ttatttattt	atftttttga	gatggagtct	tgctctgtcg	88680
ctcaggctgg	agtgcagtgg	ogtgatctcg	gctcactgca	acctcacct	cctgggttca	88740
agtgattctc	ctgcctcagc	ctcctgagta	gctgggactg	caggcatgtg	ccaccaggcc	88800
tggctaattt	tttgtgtttt	tagtagagac	gggtttcact	gtgttagcca	ggatgggtgtg	88860
atccgccttc	ctcagcctcc	caaagtgtcg	ggattacagg	catgagtcac	cgtgcccgcc	88920
cagaaatcga	aaatttctag	aaacgTaaag	cagatcagtg	ggttgtcagg	gtgagagttg	88980
ggaaTgagga	ttaaactgcaa	acaggTatga	gtgaaactaaa	aagtgttcta	aaactggatt	89040
gtggTgatga	ttgcacatct	ataaatgtat	taaaactcat	tgaattatat	acttacaatg	89100
ggtgaaatTc	atagtagatg	aattcatacc	tctataaaaac	tggttttttg	cagcaaaaata	89160
tttattgtcta	tgtttttatt	caagtTggtg	gaattattgc	tatgttttta	tccaggTgtg	89220
ctgtctcagc	cctataatcc	caacactttg	ggaggctgag	gtgggaggat	cgcttgagcc	89280
aggaattcaa	gaccagcctg	ggcaacatag	ggagaccctg	tctctacaaa	aactttttta	89340
gaaatTaaac	aggcatgtgtg	gcacgtgcct	ttggTcccag	ttactcggga	ggctgagatg	89400
ggaggattgc	ttgaaccag	gatatcgagg	ctatagtTgag	ctatgatTgt	gccctgcact	89460
ccagcctggg	taacagcaca	agaccctgtc	tcaatataaa	aaaagaaaag	aatcattaat	89520
tctgaaatta	gactgaatga	tcatftttta	atatttttca	gacaagacta	cttttgctta	89580
aaaatctgtt	atgtaaacaa	ggaagtccat	tgacccaaaa	agaactagca	caacttgcta	89640
ggtgagTaat	ttggattTgg	tttatcttac	agctftttatt	tattftttgt	aaataattct	89700
ttfttttagtt	attTaaagTat	atctTaaagTat	Tgaaatgagt	aattcatTga	tcaagaact	89760
ttctctcatc	ctctacctcc	tttgctfttt	tcttacctct	tgttcttata	tatatatgtg	89820
gaatctaaat	ttcactggct	atgtcctttg	caagacgtga	tctaattgatg	atagtatatt	89880
ctftttataa	atfttattaaa	acctgtaaagt	ggtattTaaag	taattTaaac	atfttacacct	89940
taggcaacat	ggtTaaaccc	catctctaca	aaaatataaaa	Tgagtcaagga	atggtggcac	90000
atacctgtag	ttcccagctac	ttgggaggtt	gatgtgggag	gatcgttca	gcccgggagg	90060
TggaggtTga	gcaacagagt	gagaccctgt	ctcaaaaaaa	atftTaaaca	tttcatattg	90120
cacactaaca	cagatatttt	tataTaaact	Tgtttctggc	atatgtgTaa	tactftttcag	90180
ctctftTaat	aatftTaat	atfttgataat	ttTaaataatt	ttaaattcag	ctattTgaat	90240
aatctTaaac	tttatgaaaca	gtgtgaaagTat	gaaactgagt	ataggagatg	cgctcatgaa	90300
ttftattccc	Tgaaaaatgt	cataaaagTaa	tataaaagTat	tcaatgagTat	TgtgagTatt	90360
gaataatftt	áacaagaact	ctgagTtccc	agTatccctat	gacaaaagaa	caaattactt	90420
ggaagtctct	tctactccag	acctftctgt	ctatccccct	atftTaaaaa	atgtacattg	90480
atTgtccctt	atccccatcaa	gagagTgttc	tttgccaaaga	Taaacgctaa	Ttagagaagt	90540
aatTgtgttt	atgtgtcagt	Tggccaccag	tagTttgtag	atgttacctt	tccaggTgac	90600
atgtttacaa	tttaatgaag	cctctggaga	tagtatgcct	taatctaaag	gaatggttct	90660
aatgattaaa	atftTgaggca	ttactacatt	Tgtftttcagc	aaatggccct	ttctattccc	90720
ttaaaatcaa	acctcataa	ctatgggttt	gctgacaaag	ggaaactagt	ttfttactct	90780
atttatatgga	cattftTaaat	ctgcagagat	gaaaataatt	Tgctgtttca	gctgggtgcg	90840
gtggctcatg	ogtTaatct	cactactttg	ggaggctgtg	gcgggtggat	caoctgaggt	90900
caggagTtctg	agaccagcct	gaccaacatg	gcgaaacgcc	atctatacta	aatacaaaaa	90960
aaaaatTtagc	caggtgcggt	ggcgcatgcc	tataatccca	gctacctgag	aggctgaggt	91020
aggagaatctg	ctccaggagg	Tggagatcac	agtgagctga	gatcatgcca	Ttgcactcca	91080
gctTgggcaa	caagagcaaa	actccatctc	aaaaaaaagc	gggaggggaa	ataattTgct	91140
gtttcttccc	tcccttccc	agaatgactg	atggatactc	aggaagtgac	ctaacagctt	91200
Tggcaaaaga	Tgcagcactg	ggtcctatcc	gaggtagTat	tacaagagct	Taaaacattt	91260
agaactattt	attataaccac	cttTaaagtt	TaaagaagTcc	aaaaaaatct	accaagagat	91320
ttfttttttc	ttftTggagac	agggtctcac	tctgttgccc	aggctgtagt	gcagtggcac	91380
aatcatagtt	cactgcagcc	Tcgaaacccc	caggctcagg	Tgattctccc	accagcctt	91440
cccagtagct	gggactacag	gcatgcacca	tcatgcccag	ataattattt	tattfttattt	91500
ctTgtagaga	cagggtttca	ccatgttgcc	caagctggTc	Ttgaactccc	gggtccaagt	91560
gagcagcctg	cctTggcctc	ccaaagtgtc	gggattacag	gcatgagcca	ccatgcccag	91620
cctTaaaacc	atcttagct	cacagatcat	acaaaaacag	tctatgagcc	agtagactgt	91680
gaccctgtt	ctaggcaaga	ggTttcattt	ctTgatagTat	agTaaagcagc	agaaaaactca	91740
aaaagaaaac	ttTaaacaaa	aaagaataat	gTaaaggaaaa	Tgttcttatc	Tgtftttctgg	91800

ctcttttcag	tattctgaca	tcatttgaaa	ttaatgtgtc	taaaggaaat	tctataagag	91860
aagttttcac	ttttttaatt	atatcttttg	agggattgca	gcaaaaatatt	tattgctatg	91920
tttttattca	ggttacttgg	ttgattctga	aataccacga	ccaattctttt	ttaggttagg	91980
ttaggataaa	ctcacacaga	tattttcagc	tacatttcca	cagccagcat	cggtagtggg	92040
tggttgatgc	tctcaggtat	gcactcagat	tattagctat	gatgataaag	taatcataat	92100
agctccatcc	ctttctctca	ttcaagaaag	tacaattgaa	tgtagtgtga	cgagaaagat	92160
gttatcatag	tgtctactgt	aatttgcatt	aaaagcctac	attatacaaa	cctttttggt	92220
ttactagtta	aaaattacca	gtaacacctc	ccgtagtcaa	atagggtgct	gattaagaac	92280
tgctactttg	caaaaataaaa	aatgtaatct	tgagagtata	ttttgaaact	ctggatgaac	92340
aaattttctt	gctgctggag	cttaaactct	gttcagttag	tggatttaga	aacagttagct	92400
aatgcataaaa	tgagttgtat	ttcttagaaa	cgttttggct	attctaagggt	agacaaaatt	92460
tggttttaaa	aaaagattcg	catccattaa	agtacagaaa	atggccgggt	gcagtgctc	92520
acacctataa	tcccagcact	ttggggagcc	aaggtgggag	gatcgtgtgg	gcctaatagt	92580
ttgagactag	actgggcaac	atagggagat	cctgtctcta	caaaaaata	taataaatta	92640
accaggcccg	ttggcatgag	cctgtaatcc	cagctactcg	ggagcttgaa	gtaggaggat	92700
cacttgagcc	tggaaggtgc	aggcagcagt	gagccatggt	catgccactc	actgcactcc	92760
agcctgggtg	ggtgacatag	tgaggtgctg	tcttaaaaag	taaataaata	ggccggcac	92820
ggtggtctcac	gcctgtaatc	ccagcacttt	gggaggccaa	gaugggtgga	tcatgaggtc	92880
aggagttoga	gaccagcttg	gccaatatgg	tgaaacccca	tctctactga	aaatacaaaa	92940
attagccagg	cgtaggttgg	cgcacctgta	atcccagcta	ctcaggaggc	tgaggcagga	93000
gtatcgcttg	aacctggggg	gtgaggttgc	agtaagccga	gcactgcagc	ctgggcaaca	93060
gagcaagact	ctgtctcaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaagtaa	atacataaag	93120
tgcaagagtaa	aaacaaaaaa	agactaatgc	attttgtaaa	gaacaagttg	cattccttta	93180
agttcccttt	caaatttgtg	aactattggt	tttgggcagt	atgcaagaaa	ttgaacactt	93240
tccagttatt	ccaaagaagg	aatattctct	tctcagcatt	tataaatgtg	atttgcctct	93300
aaagttaaca	tgtgtctctt	tttttaata	taatgatttg	tactgaatag	atacatgtag	93360
atcattgtac	ttggttttgc	ccttcaacaa	tttcaactgc	aaaaatgtatg	tatttttaag	93420
tgectgactt	ttatgtttta	cagaactaaa	accagaaacag	gtgaagaata	tgtctgccag	93480
tgaggtatag	tattttacaa	tgatattttc	tttgtctctc	atattgtaag	acataataaa	93540
gacatacata	tatgaatgtg	tgtgtgtgtg	gttttttttt	tttttttttt	ttttttgaga	93600
cagctctggca	tgatatcggc	tcactgcaac	ctccacctcc	tggattcaag	ctattctcct	93660
gcctcagcct	cccaagtagc	tacaggcatg	taccaccaag	cccagctaata	ttttgtattt	93720
ttagtagaga	gggcaattca	ccatgttggc	cagtctggcc	tcaagactcct	gacctcaagt	93780
gatccaccaca	actcggcctc	ccaaagtgct	gggattacag	gcttgagccg	ctgcatctgg	93840
ctgttttggt	tgtttgtttg	tttgtttggt	tgtttgagac	ggagtctcac	tctgtgcccc	93900
agactggagat	gcagtggtgc	gatctcagct	cactgcaagc	tctgcctccc	gggttcacgc	93960
cattctcctg	cctcagcctc	cctagtagct	gggactacag	gcgccacca	ccacgcctgg	94020
ctaatttttt	gtacttttag	tagagacggg	gtttcacctg	gttagccagg	atggtctctg	94080
tctcctgatc	tcgtgatctg	cctgcctcag	cctcccaaaag	tgctgggatt	acagggctga	94140
gccactcggc	ccggcctttt	tttttttttt	tttttgagac	acagggctct	gctctgtct	94200
caggtcggag	tgcaagtggg	caatcgtygc	tgactgtagc	cttgacctct	tggctctcaag	94260
cagtcctctt	gtctcaaaact	ttttagtagc	tgggactaca	ggcaagtgcc	acaacgcccc	94320
actaattttt	taattttctg	tagagatggt	gtctctccct	gttgcccagg	gtggtctcag	94380
actcttggtc	tcaagtaatc	ctcctgcctc	agtcctctcag	agtgcctggg	actaattttt	94440
tttaattggat	ttttctagtt	gattagaagc	tagaaaaaat	taactttgct	tcatttcaaa	94500
atatggaaac	ctatgaaata	gtcaaaattt	tttttggtaa	tagacaaaac	atatatacat	94560
aaagtcattt	tctatcaaat	catattacat	aaataccaaa	cagaagaaaa	ttacaggccg	94620
ggtgcgggtg	ctcacgcctg	taatcccagc	actttggggag	gccgaggtgg	gcagatcact	94680
tgaggtcagg	agttcaagac	cagcctggcc	aacatggtga	aacctgtct	ctatcaaaaa	94740
tacaaaaaaa	ttaggccggg	ggcgggtggc	cacacctgta	atcccagcac	tttgggaggc	94800
cgagggcggc	ggatcacgag	gtcaggagat	cgagaccatg	ctggctaaca	tggtgaaacc	94860
ctgtctccac	taaaaaatac	aaaaaaatta	ccggggcgctg	gtggcgggca	cctgtagtcc	94920
cagctacttg	ggaggctgag	gcaggagaat	ggcgtgaacc	tgggaagtgg	agcttgcaag	94980
gagccgagat	cgtgccactg	cactccagcc	tggcgcagag	agcaagactc	catctcaaaa	95040
aaaaaaaaaa	atagctgggc	atggtggcac	gtgcctgtgg	tcctggctac	ttgggaagct	95100
gagggcaggag	aatcaactga	acttggggag	cagaggttgc	agtgaagcaa	gattgcccac	95160
ttgcactcta	gcctgggcga	cagagtgaga	cccagctctca	aaaaaaaaaa	aattacatga	95220
aaacttgggg	ggtttataca	ataaaaaata	cctataacta	tacgttaatc	ctttaataat	95280
gttctatftt	gtttttcaat	aacctttatt	tttttaatta	atfttctttt	tttgagaag	95340

actctggctc	tgtcaccag	gotgcagtgc	agtggcacca	tcttggctca	ctgcagcctc	95400
caacctccag	gttcaagcga	ttctcctgcc	tcagcctcct	gcgtagctga	gattacaggc	95460
atgggccaacc	acgcccagct	aatTTTTTTT	TTTTTgtat	TTTTagtaga	gatggggTTT	95520
cactatgTtg	gccaggtggg	tctcgaactc	ctgacctcag	gcgatcgcca	gcctcggcct	95580
ccctaagtgc	aaggattaca	ggcgtgagcc	accatgcctg	gccaatttat	taccgtttct	95640
taatatggaa	gacctcctta	tgagatgctg	aagcatttca	TTTgaaaaaa	gTtgcatgta	95700
aaatataaaa	ttgggtatgc	agaatgTtca	taactctgtg	agcaaaattt	tgaaatcagg	95760
cttgaattac	agtcagcctt	ttctatccgt	ggattctgca	ttcatggatt	caaccaactg	95820
cagatcaaaa	atacttgggg	gaaaaaagca	tctgtgttga	acatgtacag	acttttattc	95880
ttgttattat	tocctaataca	atatagtata	aaaactattt	acattgcatt	tacattgtat	95940
taggtgttat	aagtaatcta	gagatgattt	aaagtataca	ggaggattgt	gtaggttGta	96000
tgcaaatact	gcaccatttt	acataaggga	ctttaagcgt	ttgocggattt	tgctatctat	96060
gggggaacca	atttccgaga	gatactgagg	gacagctgta	tatttGtaac	ttatTTTTta	96120
ttccctaat	tgCagcagct	gttgagggga	cagtgaactg	ttaacacaga	taacaagtgt	96180
attcaagtac	atTTTggagt	ttgtttgttt	TTTTTTTca	ggcttttatt	cgctgtcgg	96240
atgaggcacc	atcttgaat	TTTTTTTTT	aagaaagctt	ttagtTTTct	ttctttcttt	96300
TTTTTTTTT	gtgagatgga	gtcttgctct	gtttctaggc	tgagtgagg	tgCagtggtg	96360
tgatctcagc	tcactgcaac	ctccgcctac	tggttcaag	tgattctctc	gcctcagcct	96420
cccgagttagc	tgggactaca	ggtgcacgcc	accacgcca	gctaattttt	TTTTTTTTT	96480
TTTTTTTTT	tattttagt	agagacagag	ttccacctg	ttggccagga	tggtctcgat	96540
ctcctgacct	tgggatccgc	ccaacttggc	ctcccaaagt	gctagggtatt	acaggcatga	96600
gccaccacgc	ccggctgaaa	gcttttagtt	ttctaactta	tttaatttaa	tttaatttaa	96660
tttattTTta	ttttattttt	tgggacagtg	tctcacttgg	ttgcccaggc	tggagtgcag	96720
tggtaaatc	atggctcact	gcaccctcta	gctcctggc	tcaggcaatc	ctcctgcctc	96780
agcctTTTga	gttGctggga	ctacgggat	gtaccaccac	actcagctaa	atTTTtaatt	96840
TTTTgtagag	atggggtctc	actatgttgc	ctaggctgg	ctcagattcc	tcaagcattc	96900
ctccacttg	cacctcccaa	aatgctggga	ttacaggtgt	gacaccgtgc	cagacttgaa	96960
atTTTtaatc	ccacaccta	aaatataatt	ttatccacca	TTTTTaaaa	gtcataacat	97020
tatttattaa	aaatttagat	ggtaaaaacta	aaaattaaag	cttaataaaa	ctactgagtg	97080
atTTcaagg	aagaatatta	ctggctcctt	ttgtgagcat	cccattaaat	atataattc	97140
aggTtatatt	gttacaatat	ttggttctac	tgtataactt	TTTTTTTgag	actgagtttc	97200
gctgttGttg	cccaggctgg	agtGcaatgg	tgcaatctcg	gctcactgca	acctccgct	97260
cctgggttca	agtGattctc	ctgctcagc	ctcctgagta	gctgggatta	caggcatgcg	97320
ccactatgcc	tggctaattt	tgtgtttttg	tagagacggg	gtttcactat	gttggtcagg	97380
ctgatctgga	actcccaacc	tcaggagatc	cgccaccctc	ggcctcccaa	agtgtcggga	97440
ttgcagggtg	gagccaccat	gcctggctct	actatatact	ttcattcagt	tgTTTctttt	97500
taactagtg	gTTTTgggat	taataatttg	ataatgacct	ttagctgtta	ttgcttaact	97560
atgagTtaat	TTTTataaag	cacttaaagt	agctggcacg	tagtaaacac	tatgtaaaga	97620
tccattaaat	aactTTaaaa	aatataaaac	tgatagtggc	atTTTattat	agagattaag	97680
gtaatccatt	ctctcatttc	catttatggg	atgagacgta	aacacaagta	gtttgtctc	97740
taaaaactgta	tactagaatt	ttatataccc	attatttgat	gcaactTTaa	taccaaagtg	97800
tattgtcagT	taccggtgaa	tatatataaa	tttaggtaag	gaaaacccaa	cttggtcgca	97860
cacagtggct	cacgcctgta	atcctaccac	tttgggaggc	cgaggcagtt	ggatcacctg	97920
aggTcttgag	ttcgagacca	gcctggccaa	cgtggcgaaa	ccccgtctcc	actaaaaata	97980
caaaaattag	ccgggcgtgg	tggcacatgc	ctgtaatccc	agctgctcgg	gaggctgagg	98040
aaggagaatt	gcttgaacct	ggggaggTgg	aggttgCagt	gagctgagat	cgtaccactg	98100
cactccagtc	tgtgcgacag	gagcgagact	ccatctccaa	aaaaaaaaaa	aagaaaaacc	98160
caacttatct	tttacagttt	ataatagtag	aagttcaaat	aattggTttg	gaatttctgt	98220
atTTTtaagg	ttagtactaa	aattgttggT	tataaattgg	ggtacaatat	actttgtttt	98280
ttaagtacct	tgtgtatcta	atTTaacttt	aagTccttta	ttatTTTggt	ttgataagac	98340
aactTTTcta	cttatttccc	cccttaactg	aaccagctac	catctgcctt	ttcctgttg	98400
tatacattag	tctcttaccg	taaaatatca	tataagTtTc	atataatac	acatatcaaa	98460
ctatagactt	aaagtacaat	tacatcaaca	tctTTTaaaa	ccttaatttc	tggccaggcg	98520
tggTtGttca	tgcctataat	cccagcactt	tgggaggctc	aggccattgg	atccccggag	98580
ctcaggaatt	cgagaccagc	ctggataaca	tggTgaaatc	ctcatctcta	caaaaattac	98640
aaaaagTtag	ctgggcgtgg	tggcgcacac	ctatagTcct	agctacttgg	gaggatctct	98700
taagcctggg	aggcagaggT	tgCagtgagc	cgagatcatg	ccactgcact	ccagcctggg	98760
taacagTgag	accctgtctc	aaaacaaaa	tattaatttc	tataacaaat	aaattttatt	98820
tagtaatatc	atgtTTaata	cccatgttac	attcattatt	cttctacact	ttggcctTTa	98880

ggaagaggca	gtgtgtgtgt	tacagttaaa	taggaaacaa	gtagggttca	tatagtgtct	98940
atggggtttt	ctttgggtggg	gggttgctgg	gtttttgggg	tttttttttt	tgagacaggg	99000
tatcactctg	tctcccaggc	tgaaatacat	tgccatgctc	atggctcact	gcagcctcaa	99060
ctcccagac	tcaagtgatc	ctcccacctc	agcttcccta	gtagctgaga	ctacaggcgt	99120
gcgccaccat	gcctagctaa	attttgtaca	ttttgtagag	atgaggcttt	gccatgttgc	99180
ccaggctggt	ctctaactcc	tgggctcaag	tgatccactt	gcttcagcct	cccaaagttc	99240
tgggattaac	aggtgcgagc	cactgcagcc	agcctatatt	ttactatat	gtttttttct	99300
ttttggctaa	aattttccag	attagtttac	aaagtacaag	tgtaggtgat	atctcatgga	99360
gatatcaaag	atgatatgaa	gttagattgg	gtttttaaga	gtagttttta	aaatacggat	99420
aaataccagt	tgttggagtt	ttgtttaaag	ttctttaact	tcttattttt	gggcaaaaat	99480
acaggtatac	gctagaaatg	atttttaaca	caggtcattt	atgccaact	gcattttgcc	99540
ttaatttttt	tttttttttt	tttttttgag	aggagctctc	actcattgcc	caagctggag	99600
tgcaatggca	cgatctcggc	tcacttacia	cctccacctc	ccaggttcaa	gcgattctcc	99660
tgctcagcc	tctgagtag	ctgggattac	aggcacctgc	ctggctaatt	tttgtatttt	99720
tagtggagac	gggtttcac	catgttggcc	aggctggtct	cgactgctg	acctcaggtg	99780
atccacccgc	cttggcctcc	caaagtgcgg	ggattacagg	tgtgagccac	cggtgcttgc	99840
caaaattttt	atataatttt	ctattgcctg	gactctgtga	acctatccat	tttgcctttt	99900
aaaaatactt	aggtgtaaat	atagatattc	attaactcag	cattgtttta	atctataatt	99960
ccaaaggcaa	tttaaaagat	cagaaaataa	gaccaaaata	atataaaaa	gcatacttta	100020
ggctgggcaa	agtggctcac	gtctgtaatc	ccagcacttt	gggaggccaa	ggtgggaggga	100080
tcacctgagg	ttaggagttc	aggaccaacc	tggccatcat	ggcgaacccc	tgtctctact	100140
aaaaatacaa	aaattagctg	ggcatggagg	catgtgcctg	taatcccagc	tactcgggag	100200
gctgagacag	aagaattgct	tgaaccttgg	agggggaggt	tgcatatctg	agtggtgaaa	100260
ttgtgattct	ttttttctct	ttgtctgtat	ttttgaactt	ttctataaat	gattgtgttt	100320
tgtttttata	ttggaaaaat	attatgcttt	caaagttaa	tacctatgaa	actaaacaca	100380
agtaataaat	atattagtat	agcatttatt	aaggtttctt	gtgtagcaga	tcaacataga	100440
aaataatatt	aaatggctga	cataattttc	taagaataca	tacacgtata	ttttttataa	100500
cattaagaaa	cagcagcatc	attactttta	tccatcattt	cgtaaccac	catataacctg	100560
ttgatcattt	gtattgtcat	gtgcttttta	aaaactaga	tgagaaatat	tcgattatct	100620
gacttactgt	aatccttgaa	aaaaataaaa	ggcagcgtca	gcctcaaac	tttagaagcg	100680
tacatacgtt	ggaacaagga	ctttggagat	accactgttt	aaggaaatac	ctttgtaaac	100740
ctgcagaaca	ttttacttaa	aagaggaaac	acaagatctt	caatgaacgt	catcggctac	100800
agaaacagcc	taagttfaca	ggacttttta	gagtcctfaca	tatttgtgca	ccaaacttga	100860
agatgaacca	gaaaacagac	ttaaacaaaa	tatacaatgc	aaatgtaatt	ttttgtgttt	100920
taaggccttg	ccttgatggt	cacagttatc	ccaatggaca	ctaagttaga	gcacaacaaa	100980
acctgattct	ggtcttcttt	accaatataa	tcataatgta	ataataaatt	tgtatattgt	101040
gttgacagatg	aaagtattcc	aggaacagtg	aatggtagaa	gacacaagaa	catttgtttg	101100
tttgtctctt	gatgtttttt	cttaaaaatag	taatttctcc	tacttttctt	ttctactgtt	101160
gtcttaacta	caggtgattg	gaatgcctaaa	cactcttaag	tttattttct	tttttctgtt	101220
tataaattca	gtgtgcctaaa	tgaaactttt	ttcctaagta	actgtaatag	gaaaagttt	101280
attttgagag	tttcttcttc	ataaatctac	agacattaaa	caattgttgt	gttcttttta	101340
ccttttattt	ttctattacc	ttgctacca	acagtttaga	tagcaatata	atagcaaaaa	101400
agcaaatatg	gtaaaaataga	gaaggtttga	aggtttgagt	tactctgtca	tataacatgt	101460
agatcagctc	tcatgtgacc	tcagatattt	ttttttctaa	tgtatttgtc	agaaatctgt	101520
tgtagactgt	taacttcttc	ctgatggat	ttatttctg	caagaattat	tctgatattt	101580
aagagagcca	attttaactg	ctgtgaaaat	gtttccagtg	caagagaagg	gaaatactag	101640
gaaactaagac	atttctaat	tattgcttat	tactttctta	attttacagg	ataattataa	101700
gcaagtggaa	ctaccatctt	ttattcttaa	taattattaa	tcccttcaat	gaaactttaa	101760
aaaaactgaa	ttttatatac	tggcatacat	ttttctagtt	ccttctgctt	gctttattaa	101820
ctcaaaagtt	ctagttctag	tctgttgatc	tgcttttgg	tctcccaaaa	tgtacagtaa	101880
ttccatttgt	ttgtataaat	atgcctggat	tttcattata	aaaaatgcat	tgtagggagt	101940
agagactcat	atcatggcct	tttaaatatt	gtaataaagg	caaatagata	tttgcctta	102000
gtttactggt	taaaagtttg	tttacagaac	ttttctctgg	tgcttaaatg	atgctatgta	102060
aaatgtcatg	agtggaaaga	atattttag	tagtaacaag	aatttttcat	ttaggaaaga	102120
tttcttaggt	tttgaaagaa	tacattaaaa	taaaaaactt	gccctacta	ggtaagaact	102180
ttataatgaa	gacatacatt	cttcttaatt	ttactcttgc	tcttgttaaa	gatttgtttg	102240
aatatagaag	atgcatgatt	tctgggtttt	tttttttttt	tgagacagag	tttgcctctt	102300
gttgccaggg	ctggagtgca	atggcgcaat	ctogactcac	cacaacctcc	gcctcccagg	102360
ttcaagcaat	tctcctgcct	cagcctcccg	agtagctggg	attacaggca	tgccccaacta	102420

ccccagctaa	ttttgtat	ttagtagaga	tggggtttct	ccatgttgg	caggctggtc	102480
ttgaactcct	gacctcaggt	gatccgcctg	cctcggcctc	ccaaagtgc	gggattacag	102540
gcataagcca	ctgcgcccag	ccagaagatg	catgatttct	taggatcata	tgctgtttgt	102600
agccataagg	taaatacatg	ctctccaat	catgactttg	gaactccctg	aataataaaa	102660
atgagagt	agataaatag	gggaaaaaaa	atttttttca	agccagagct	atgcatatgt	102720
tagtgatgg	gtagtatccc	tttaaggtct	caaacattac	aacatcaatt	atgaaatact	102780
gataacgaaa	ggtagtaatg	aaataataat	gatgaaaaaga	attgagaagt	tctaaattaa	102840
gacatttcag	ttaaagctcat	aaaatttcat	tgttttcatt	taaaagatta	acgttattga	102900
tacttggata	actggcta	catattaaag	gactatgtgg	ttccagctca	acttttaata	102960
tattgtctcc	tttaaaacta	tcatggttat	aattctattg	ggaaagactt	ttagataaca	103020
aagatttcaa	atgttaaaag	agataaaagt	caggttaata	ctatcttaaa	cactgagtca	103080
gaaaatcatt	actgtataga	agttgctttc	ctgatcaagt	ctgaacttca	gctagtgtca	103140
gagaactatt	ttctatgact	taactcta	caagttttat	tttaagctgt	ttctttgata	103200
gaaggccat	gaaaatagag	taatgatata	gtaggagata	agggattgg	ttggctcttt	103260
tcaataaaga	tagaagttgc	tgaagtttct	tgaattaata	atgacttaga	ttgtgacctt	103320
ttagattcgg	tgttgagctc	tgtgtgtgat	aacttcctaa	aagataatgc	ttaaacatta	103380
agcattagtg	tgctcttcat	gttaatatgg	cagagttttg	taaaactaat	taaaacttac	103440
tgatataattg	gactttgagc	caagggaaag	aatgagta	atctttccag	atatcttaag	103500
ggtaaaagct	tattctaaga	cagctctgtcc	atbgagaata	ttagatttct	gacttgcaaa	103560
tatgtttgta	ctccagaaga	attagaggaa	aagcagatac	tagaattcta	atttaattac	103620
atatacagcc	gtctttgttt	atagtgtaga	atctttata	ttttgtacaa	aaactaattc	103680
ttttggtaaa	atgaaccatt	tacagttcgg	ttttggactc	tgagtcaaa	gattttcctt	103740
taaatacattg	tctcaatttt	agtctggctc	tttctacttt	tcttcagaag	aatgaatta	103800
aagggtacag	ttgcataaag	tggtttttta	tcctaagtga	ttggaataaa	atgataaa	103860
ttattttgc	tttacttttt	tactttaaaa	ctttttgata	ttttaggggt	tgtagtctga	103920
taatgaagga	gttgtgtgta	ttggactctt	agtaacaatt	ataaacgctt	acaaaatat	103980
agaaagcaat	gattgggtgg	ctctgcagag	caatcaaaac	aaggtagaaa	ctgcaaaagc	104040
ctataatgga	agagattcag	gctgggtgtg	gtggctcatg	cctgtttggg	aggccagtat	104100
gggagaatcg	cttactactca	ggaattcaag	accagcctgg	ggaatatagg	gagaccctgt	104160
ttctgcca	aaaaaaaa	aaaaaaaaac	taaaactaaa	attagccagg	tttgggtggc	104220
tacacctgta	gtcccgggta	ctcaagaggc	tgaggttaga	ggatcgctta	agcccgggag	104280
gaggaggtg	cagtgcagcca	agatcacacc	accacactcc	agccaggtt	acagagaacc	104340
tgtctcaaaa	ataaaaaagt	aaacgagatt	caacttaacc	agcttttacc	catagggcaa	104400
tttccagttt	gtgcagttca	cttggatata	gaattcaggt	agaatgtgac	agttttgaca	104460
ggctgaggaa	tcatttggga	cttctagaac	acctgaaaat	tagaagagaa	attttgggaa	104520
ggagagtgcc	tcagaaagta	agctctaaaa	tctgcctaca	tattcttttc	aaatccttgg	104580
tgatttccaa	aattgtgctg	gtgcagtg	attattttaa	ggaaccagg	agaaagcaag	104640
ttaaagtcta	aaaaactcag	cagcgattgc	agctgcccc	ggacagagag	tttgaatttc	104700
aagttccact	aagaaggact	gagtaaatc	ttgggttttc	ccattgaacc	cccaaatgc	104760
cacgccttaa	ggaagaatga	cagcatccta	gaactaaagg	ctgtgcttca	gactaagga	104820
caaaatggaa	ataaacttac	tgtaatgaag	cttaaaatca	aatctcacag	catcaaggtt	104880
atccattagt	aatttaaaac	cagaacaaa	cacagcattc	ttgagaagaa	aacaattcag	104940
tcatcacagc	atataatccaa	agctcagtac	ataataaaca	gttactaagc	atgcaaagaa	105000
aaattatgtg	atccttttga	ctcaacttcc	ctgatggaaa	aaaaaaagt	gtccatacca	105060
tactcaacag	aaaatataca	gaagcagatg	cacaagtgc	ctctatgaga	gtgattttaa	105120
aatagcgtta	aaatatgtta	aaagaattta	caagaagaga	taatgggtga	atagataggg	105180
agtttcagca	gaaaaataag	aattgaaaac	caaacagagg	ccaggcaccg	tgctcacgc	105240
ccgtaatcac	agcactttgg	gaggecaagg	caggtggatc	acctaaggtc	aggagttaa	105300
gaccagcctg	accaacatgg	tgaaccocgg	tctctactaa	aaatacaaaa	ctagccgggc	105360
atgggtgatg	catgcccata	atcctagcta	cttgggaggc	tgaagcagga	gaattgcttg	105420
aacctgggag	gtagaggttg	cagtgaacca	agatcatgcc	attgactcc	agcctgggca	105480
agaagagcaa	aactccatct	caaaaaataa	taataataat	tggtctggcg	cgggtgctca	105540
cgctgtaat	cccagcactt	tgggagccg	agacaggcag	atcacaaggt	caggagattg	105600
agaccatcct	ggctaacatg	gtgaaacccc	gtctctacta	aaaatacaaaa	aaattagcct	105660
ggcatggtgg	cgggcacctg	tagtcccagc	tactcgggag	gctgaagcag	gagaatggca	105720
tgaccccag	aggcggaaact	tgcaatgagc	ccagatcgcg	ccactgca	ccagcctggg	105780
caaccgtctc	aaaaataata	ataataataa	ttgaaaaata	ataaaaaata	accactgga	105840
aagtcttgta	atgaaaacta	cagtatctga	aattaaaaat	ttaaatggat	ggacttaaca	105900
gtaaaactag	cacaacagga	aagactgaaa	aagacaagta	agtatcaaac	caaagcagag	105960

agagaaagaa	aaaagaaata	gcagaggttt	gagacctttg	gaacattatc	agttcctgta	106020
at ttgaa gta	aagaaggaaa	ggaagagaat	ggggcaagag	aaatctttga	agcaatgatg	106080
cgtgaaat t t	tccccaaagt	gttgaaagac	atcaacgtac	agatctaaga	agttcactga	106140
atccccaaagc	agaataaata	caacaaaagac	tacatctggg	tacatcacat	tgccaaaaaa	106200
aaaaaaaaaaaa	aaaatcttaa	aagcaatcaa	ggtggggagt	gggaaggagt	agcattacat	106260
tcagataaac	aagagtgatg	gctaacctca	tcagaaatga	tggaattaag	aaaaatgatgg	106320
aatgacatct	ttaagtgca	aagatgaaac	aaaggttaac	ttaaaactct	gtatccagtt	106380
aaaatatcct	tcaagatgtt	aaggcaatgg	ccaggctcag	tggcttatgc	ctgtaatccc	106440
aaacttttgg	ggggctgaga	caggagaatc	ccttgaggcc	atgagtgacc	agcctgggca	106500
acctagtgag	aaccccatca	aaaaataaaa	ttagctgggc	atgggctgtg	gtcgtgctac	106560
tgcaactccag	cctgggggac	agagtcccta	tctctgaaa	aacttaagac	aaaaatcttt	106620
ccagatgaaa	acagaatttg	ttgctggcag	actgacttta	tacggaatgc	tttaagaaat	106680
tcttcaggct	gaaggaaaat	ataaatctgg	agtgaaaaac	gtagtatctg	aaattaaaaat	106740
tttctttgga	tgagataata	agtaaatgga	tttcccaatg	ggaacctgta	tctccaggat	106800
ggaattaaga	at tttgaaat	gataaacatg	ttggtaaaata	agaaagacat	ttttcttttt	106860
tat ttttctta	aatatcattg	ctgtttaaag	caaacaaaaa	tagtattgca	gtgtttataa	106920
tatatgtata	agtaaaatat	aggacagaaa	tagcatgaag	ggtcagagag	ggcaataaaat	106980
gtaattatata	tgagataata	ttactatatt	gtcacaaaag	tgagatagca	ttagtgtaaag	107040
gtagattgtg	ataagaatgc	at tttgtaat	ctctcacccc	atcactgggg	tgca gtagat	107100
ctaaaaagcc	at tagagaag	ataaaatggg	tcactaagaa	ttatgtgatt	agtcaaatga	107160
aatcaaggaa	agaaaaacag	agcaaaagaa	tgtaagataa	aacagtggga	tggaagagct	107220
gaaattaata	atgtcactaa	attgcttacc	ttctccttcc	ctgaaactgtt	gtggat tttt	107280
tttttttccag	ggggttgagg	ttttttttta	gagacagagt	ctctattgcc	ccaggctgga	107340
cagcagtggc	gtaatcagtc	tcaaaatccc	tgca gctcca	aactcctggg	ctcaagtgat	107400
cctcctgcct	cggcctccca	aaatgctggg	attacaggca	ggcagagcca	ccacaccctg	107460
ccagaactat	tgttttaaaag	tcattaggtg	gttaagccct	cgagactatg	tgcaagtttg	107520
atgattcggt	aggagcactc	acaggactca	gcatatagtc	ttattcccaa	caatgat tta	107580
ttagagcaaa	aggatatgaa	gcaaaattcag	caaagggaaa	aaaggaatta	agtgaagtcc	107640
agaggaaacc	aagcaccagc	ttctaagggt	tctctcctaa	tgaa gtcaca	caggatgtca	107700
ctccagcaac	aagt gacagc	atgtgaaagt	gttaagcctc	actgttagac	atagtaagct	107760
tgttagggaa	tagagggaat	ctcccctcaa	at tccaagtc	ccagatgcct	ggcaaggccc	107820
aaccttgcca	gcaggacttt	ctgggtagag	cag tcccagg	caogtgtttg	cacagtggtt	107880
cagatcaatg	tgcttacatt	gggtagaggg	gaacctatgga	agtc caaatt	tggggtgcag	107940
agacctaaata	gggtgaagac	agtgctctaca	atgatggaca	gctaggtatg	aggtgtcagg	108000
gacagattga	ggcagttatt	cacatgggtg	atagggggcaa	cctggaaatga	ggaataagct	108060
caagcacaga	aaggggtgtc	catatgggg	aaaggggtca	gcagagatgg	tagattgggt	108120
gcacaccaaaa	gacttgatgg	aataaaagtg	atatactaac	cacagagaag	gtaattataa	108180
atacagaaaag	gggggaaact	agaatgaaac	ctggagttca	gcttgaattg	ggtttaagaa	108240
agtgaattca	tggtttaaaa	tctataaaaa	tagatgaaat	atcggctggg	cacagtggtg	108300
cacgcatata	atcccagcac	tttggggcga	cgagggcggc	agatcgcttg	agatcaggag	108360
tttgagacca	gcctggccaa	cactgcgaaa	ccccatctct	agtaaaaaata	caaaaaattag	108420
ccaggtgtgg	tggtacacac	ctgtagtccc	agctacttgg	gaggctgagg	cacgagaatc	108480
acttgaaacct	gggagggcaga	ggttgcagtg	ggctgagatc	gcaccactgc	actccagcct	108540
gggcaacaga	gcgagatcct	gcctcaaata	aaaatagagg	acatataagc	atataaatat	108600
acacgtgtgt	atgtgtccaa	atatgtatat	tcctagtct	gtccaccaag	gtggccttgg	108660
agcagttatg	ctccaataat	aatgagcaca	taaagtacc	atatcttgcc	ttccaaattc	108720
ttcactgtct	tagtctgctt	gggctgcatt	acaaaatacc	atagactggg	cagcttaaat	108780
aacagaaatt	tatcctcaca	gttctagaag	cttggaagtc	caagattaaa	gtaccagcca	108840
gtttggtttc	tagtgagggc	tttcttctg	ggttgcagat	ggccaccttc	tcaccgtgtc	108900
cttgtatggc	agacagcaca	agctctctgg	tttccctttt	taaaagggca	ttaatcccat	108960
catgacagtc	ccatcctcat	tatctcatct	aaccttaggt	acttcccaaa	ggctgaaatca	109020
ccaaagacca	tcacattgct	ggtgaaggct	tcaacatag	aatttgaggg	acacgaatat	109080
tcagtcacata	acatcaacta	aaggaaacca	gactctttga	taaaatggct	aaattcaggg	109140
ctggggcaga	gaaaatacat	gagtggtgaa	cttcttgtgc	cagagagaaa	aagtgc ccaa	109200
agattgatga	ggatgaatca	ttgaaatgac	acacagatta	aaagggttcc	cactggacaa	109260
at tttgagcat	caaaaatagt	aatagtagta	attaattata	acccatcaga	agaaataaac	109320
catgagctca	tgtgaaatata	tgaatacaaa	cataa caaaa	ttacaagcat	aatgaggaat	109380
gtgatattta	tatggtttaa	aggtaacctc	ccaggccggg	tcagtaact	ctcacctgta	109440
atcccagcac	tttgggaggg	caaggcaggt	agatcacctg	aggtcaggca	tttgagacca	109500

gcctgcacaa	catggtgaaa	cctgactct	actaaaaata	cataacgga	gccgggctg	109560
gtggcacgtg	tctataatct	gccactgatt	agggtgtgga	ttttcccaag	caggggataa	109620
tagtagtacc	tatgtcaaaag	gctgttatga	ggattaatg	agctaacaca	taatcgtgct	109680
ttttttttt	ttttttttt	ttgagacaga	gtcttgcact	gtcgcctggg	ctggagtgca	109740
atggcacgat	ctcggccac	tgcaacctct	gcctcccagg	ttcaagtgat	tctcctgcct	109800
cagcctcctg	agtagctggg	attacaggct	cctgccacca	cacctggcta	ttttcaatag	109860
agacgggggt	tcactatgtt	ggccaggcta	gtctcaaaaa	cctgacctcg	tgatccacc	109920
gctttggcct	cccaaagtgc	tgggattaca	ggcatgagcc	actgcaccog	gcttttttt	109980
ttttttttg	agatggaatc					110000

<210> 2
 <211> 3263
 <212> DNA
 <213> Homo sapiens

<400> 2						
gctcctgaga	cggcgggca	cacgggggtc	tgtggccccc	gccgtagcag	tggctgccgc	60
cgtcgttgg	ttcccgtcg	tctgcgggag	gcgggttatg	gcggcggcgg	cagtggagagc	120
tgtgaatgaa	ttctccgggt	ggaacgaggga	agaagaaagg	ctccggcggc	gccagcaacc	180
cggtgcctcc	caggcctccg	cccccttgcc	tggcccccg	ccctcccggc	gccgggcccgg	240
cccctccgcc	cgagtcccg	cataagcggg	acctgtacta	tttctcctac	ccgctgtttg	300
taggtctcgc	gctgctgcgt	ttggtgcct	tccacctggg	gctcctcttc	gtgtggctct	360
gccagcgtt	ctcccgcgcc	ctcatggcag	ccaagaggag	ctccggggcc	gcgccagcac	420
ctgcctcggc	ctcggccccg	gcgccgggtc	cgggcggcga	ggccgagcgc	gtccgagtct	480
tccacaaaca	ggccttcogag	tacatctcca	ttgcctcgcg	catcgatgag	gatgagaaaag	540
caggacagaa	ggagcaagct	gtggaatggt	ataagaaagg	tattgaagaa	ctggaaaaag	600
gaatagctgt	tatagttaca	ggacaagggt	aacagtgtga	aagagctaga	cgccttcaag	660
ctaaaatgat	gactaatttg	gttatggcca	aggaccgctt	acaacttcta	gagaagatgc	720
aaccagtttt	gccattttcc	aagtcacaaa	cgagcgtcta	taatgacagt	actaacttgg	780
cagcccgcaa	tctctttaat	cagtcagaaa	gtggagctgt	tccaaaaaga	aaagaccctt	840
taacacacac	tagtaattca	ctgcctcgtt	caaaaacagt	tatgaaaact	ggatctgcag	900
gcctttcagg	ccaccataga	gcacctagtt	acagtggttt	atccatggtt	tctggagtga	960
aacagggatc	tggtcctgct	cctaccactc	ataagggtac	tccgaaaaaca	aataggacaa	1020
ataaaccttc	taccctatac	actgctactc	gtaagaaaaa	agacttgaag	aattttagga	1080
atgtggacag	caaccttgct	aaccttataa	tgaatgaaat	tgtggacaat	ggaacagctg	1140
ttaaatttga	tgataatagct	ggtcaagact	tggcaaaaca	agcattgcaa	gaaattgtta	1200
ttcttcttc	tctgaggcct	gagttgttca	cagggcttag	agctcctgcc	agagggctgt	1260
tactctttgg	tccacctggg	ataagggaaga	caatgctggc	taaagcagta	gctgcagaat	1320
cgaatgcaac	cttctttaat	ataagtctg	caagtttaac	ttcaaaatac	gtgggagaag	1380
gagagaaatt	ggtgagggtc	ctttttgctg	tggctcgaga	acttcaacct	tctataattt	1440
ttatagatga	agttgatagc	cttttgtgtg	aaagaagaga	aggggagcac	gatgctagta	1500
gacgcctaaa	aactgaattt	ctaataagaat	ttgatggtgt	acagtctgct	ggagatgaca	1560
gagtaactgt	aatgggtgca	actaataggc	cacaagagct	tgatgaggct	gttctcaggc	1620
gtttcatcaa	acgggtatat	gtgtctttac	caaatgagga	gacaagacta	cttttgctta	1680
aaaatctggt	atgtaaacaa	ggaagtccat	tgaccaaaaa	agaactagca	caacttgcta	1740
gaatgactga	tggatactca	ggaagtgacc	taacagcttt	ggcaaaagat	gcagcactgg	1800
gtcctatccg	agaactaaaa	ccagaacagg	tgaagaatat	gtctgccagt	gagatgagaa	1860
atattcogatt	atctgacttc	actgaatcct	tgaaaaaaat	aaaacgcagc	gtcagccctc	1920
aaactttaga	agcgtacata	cgttggaaca	aggactttgg	agataccact	gtttaaggaa	1980
atacctttgt	aaacctgcag	aacattttac	ttaaagagg	aaacacaaga	tcttcaatga	2040
acgtcatcgg	ctacagaaac	agcctaagtt	tacaggactt	tttagagtct	tacatatttg	2100
tgcaccaaac	ttgaagatga	accagaaaac	agacttaaac	aaaatataca	atgcaaatgt	2160
aattttttgt	tgtttaaggc	cttgccctga	tggtcacagt	tatcccaatg	gacactaagt	2220
tagagcacaa	caaaaacctga	ttctggtcct	ctttaccaat	ataatcataa	tgtaataaat	2280
aatttgtata	ttgtgttgca	gatgaaagta	ttccaggaac	agtgaatggt	agaagacaca	2340
agaacatttg	tttgtttgtc	ttctgatgtt	ttttcttaaa	atagtaattt	ctcctacttt	2400
tctttttctac	tgttgtctta	actacagggtg	attggaatgc	caaacactct	taagtttatt	2460
ttcttttttc	gttttataaa	ttcagtggtc	caaatgaaac	ttttttccta	agtaactgta	2520
ataggaaaaa	gtttattttg	agagtttctt	cttcataaat	ctacagacat	taaacatttg	2580

```

ttgtgttctt tttacctttt atttttctat taccttgcta ccaaacagtt tagatagcaa 2640
tataaatagca aaaaagcaaa tatggtaaaa tagagaaggt ttgaagggtt gagttactct 2700
gtcatataaac atgtagatca gtcttcatgt gacctgcagt attttttttt ctaatgtatt 2760
tgtcagaat ctgttgtaga ctgttaactt cttcctgatg gaattttatt tctgcaagaa 2820
ttattctgat atttaagaga gccaatTTta actgctgtga aaatgtttcc agtgaagag 2880
aagggaata ctaggaacta agacatttct aatttattgc ttattacttt cttaatTTta 2940
caggataatt ataagcaagt ggaactacca tcttttattc ttaataatta ttaatccctt 3000
caatgaaact ttaaaaaaac tgaattTTta tacatggcat acatttttct agttccttct 3060
gcttgcttta ttaactcaaa agttctagtt ctagtctggt gatctgcctt ttgttctccc 3120
aaaatgtaca gtaattccat ttgtttgtat aaatagcct ggattttcat tataaaaatg 3180
tcattgtagg gagtagagac tcatatcatg gccttttaaa tattgtaata aaggcaata 3240
gatatttgcc cttagtttac tgg 3263

```

<210> 3

<211> 616

<212> PRT

<213> Homo sapiens

<400> 3

```

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

```

Thr Thr Ala Thr Arg Lys Lys Lys Asp Leu Lys Asn Phe Arg Asn Val
 305 310 315 320
 Asp Ser Asn Leu Ala Asn Leu Ile Met Asn Glu Ile Val Asp Asn Gly
 325 330 335
 Thr Ala Val Lys Phe Asp Asp Ile Ala Gly Gln Asp Leu Ala Lys Gln
 340 345 350
 Ala Leu Gln Glu Ile Val Ile Leu Pro Ser Leu Arg Pro Glu Leu Phe
 355 360 365
 Thr Gly Leu Arg Ala Pro Ala Arg Gly Leu Leu Leu Phe Gly Pro Pro
 370 375 380
 Gly Asn Gly Lys Thr Met Leu Ala Lys Ala Val Ala Ala Glu Ser Asn
 385 390 395 400
 Ala Thr Phe Phe Asn Ile Ser Ala Ala Ser Leu Thr Ser Lys Tyr Val
 405 410 415
 Gly Glu Gly Glu Lys Leu Val Arg Ala Leu Phe Ala Val Ala Arg Glu
 420 425 430
 Leu Gln Pro Ser Ile Ile Phe Ile Asp Glu Val Asp Ser Leu Leu Cys
 435 440 445
 Glu Arg Arg Glu Gly Glu His Asp Ala Ser Arg Arg Leu Lys Thr Glu
 450 455 460
 Phe Leu Ile Glu Phe Asp Gly Val Gln Ser Ala Gly Asp Asp Arg Val
 465 470 475 480
 Leu Val Met Gly Ala Thr Asn Arg Pro Gln Glu Leu Asp Glu Ala Val
 485 490 495
 Leu Arg Arg Phe Ile Lys Arg Val Tyr Val Ser Leu Pro Asn Gly Glu
 500 505 510
 Thr Arg Leu Leu Leu Leu Lys Asn Leu Leu Cys Lys Gln Gly Ser Pro
 515 520 525
 Leu Thr Gln Lys Glu Leu Ala Gln Leu Ala Arg Met Thr Asp Gly Tyr
 530 535 540
 Ser Gly Ser Asp Leu Thr Ala Leu Ala Lys Asp Ala Ala Leu Gly Pro
 545 550 555 560
 Ile Arg Glu Leu Lys Pro Glu Gln Val Lys Asn Met Ser Ala Ser Glu
 565 570 575
 Met Arg Asn Ile Arg Leu Ser Asp Phe Thr Glu Ser Leu Lys Lys Ile
 580 585 590
 Lys Arg Ser Val Ser Pro Gln Thr Leu Glu Ala Tyr Ile Arg Trp Asn
 595 600 605
 Lys Asp Phe Gly Asp Thr Thr Val
 610 615

<210> 4
 <211> 23
 <212> ADN
 <213> Artificial Sequence
 <220>
 <223> Amorce

<400> 4
 cggagctcct cttggtgcc atg

23

<210> 5
 <211> 26
 <212> ADN
 <213> Artificial Sequence

<220>
 <223> Amorce

<400> 5 agaagcgctg gcagagccac acgaag	26
<210> 6 <211> 27 <212> ADN <213> Artificial Sequence <220> <223> Amorce	
<400> 6 aaggcgacca aacgcagcag cgcgaag	27
<210> 7 <211> 26 <212> ADN <213> Artificial Sequence <220> <223> Amorce	
<400> 7 aggagcaagc tgtggaatgg tataag	26
<210> 8 <211> 27 <212> ADN <213> Artificial Sequence <220> <223> Amorce	
<400> 8 tggttatggc caaggaccgc ttacaac	27
<210> 9 <211> 26 <212> ADN <213> Artificial Sequence <220> <223> Amorce	
<400> 9 caaacggacg tctataatga cagtac	26
<210> 10 <211> 25 <212> ADN <213> Artificial Sequence <220> <223> Amorce	
<400> 10 ttaggaatgt ggacagcaac cttgc	25
<210> 11 <211> 25 <212> ADN <213> Artificial Sequence <220> <223> Amorce	

<400> 11 cttctctgag gcctgagttg ttcac	25
<210> 12 <211> 27 <212> ADN <213> Artificial Sequence <220> <223> Amorce	
<400> 12 tgctagaatg actgatggat actcagg	27
<210> 13 <211> 24 <212> ADN <213> Artificial Sequence <220> <223> Amorce	
<400> 13 agatgcagca ctgggtccta tccg	24
<210> 14 <211> 26 <212> ADN <213> Artificial Sequence <220> <223> Amorce	
<400> 14 atgaacgtca tcggctacag aaacag	26
<210> 15 <211> 19 <212> ADN <213> Artificial Sequence <220> <223> Amorce	
<400> 15 tagcagtggc tgccgccgt	19
<210> 16 <211> 20 <212> ADN <213> Artificial Sequence <220> <223> Amorce	
<400> 16 aagcggctct tggccataac	20
<210> 17 <211> 19 <212> ADN <213> Artificial Sequence <220> <223> Amorce	

<400> 17 ggcggcagtg agagctgtg	19
<210> 18 <211> 20 <212> ADN <213> Artificial Sequence <220> <223> Amorce	
<400> 18 ctagctcttt cacactgttc	20
<210> 19 <211> 20 <212> ADN <213> Artificial Sequence <220> <223> Amorce	
<400> 19 aacaggcctt cgagtacatc	20
<210> 20 <211> 20 <212> ADN <213> Artificial Sequence <220> <223> Amorce	
<400> 20 ctgtgaacaa ctcaggcctc	20
<210> 21 <211> 20 <212> ADN <213> Artificial Sequence <220> <223> Amorce	
<400> 21 atgagaaagc aggacagaag	20
<210> 22 <211> 18 <212> ADN <213> Artificial Sequence <220> <223> Amorce	
<400> 22 tgccaagtct tgaccagc	18
<210> 23 <211> 20 <212> ADN <213> Artificial Sequence <220> <223> Amorce	

<400> 23 ctacaactgc tactcgtaag	20
<210> 24 <211> 19 <212> ADN <213> Artificial Sequence <220> <223> Amorce	
<400> 24 cagtgctgca tcttttgcc	19
<210> 25 <211> 20 <212> ADN <213> Artificial Sequence <220> <223> Amorce	
<400> 25 taggaatgtg gacagcaacc	20
<210> 26 <211> 20 <212> ADN <213> Artificial Sequence <220> <223> Amorce	
<400> 26 aaagctgtta ggtcacttcc	20
<210> 27 <211> 20 <212> ADN <213> Artificial Sequence <220> <223> Amorce	
<400> 27 tggagatgac agagtacttg	20
<210> 28 <211> 19 <212> ADN <213> Artificial Sequence <220> <223> Amorce	
<400> 28 ctggaataact ttcactgac	19
<210> 29 <211> 19 <212> ADN <213> Artificial Sequence <220> <223> Amorce	

<400> 29 atgaggctgt tctcaggcg	19
<210> 30 <211> 20 <212> ADN <213> Artificial Sequence <220> <223> Amorce	
<400> 30 gtgagccgaa ctgcacattg	20
<210> 31 <211> 21 <212> ADN <213> Artificial Sequence <220> <223> Amorce	
<400> 31 caaagtcgac agctacagtg c	21
<210> 32 <211> 20 <212> ADN <213> Artificial Sequence <220> <223> Amorce	
<400> 32 ggaactgtag ttgagtggga	20
<210> 33 <211> 19 <212> ADN <213> Artificial Sequence <220> <223> Amorce	
<400> 33 agatgaggct cgcacctac	19
<210> 34 <211> 20 <212> ADN <213> Artificial Sequence <220> <223> Amorce	
<400> 34 aatgccacac ttgtaatctc	20
<210> 35 <211> 22 <212> ADN <213> Artificial Sequence <220> <223> Amorce	

<400> 35 tgtgaatata tcataatttg gg	22
<210> 36 <211> 19 <212> ADN <213> Artificial Sequence <220> <223> Amorce	
<400> 36 tacagcagtt ctcatgatg	19
<210> 37 <211> 21 <212> ADN <213> Artificial Sequence <220> <223> Amorce	
<400> 37 gaccaaattg gtgcatgcat g	21
<210> 38 <211> 20 <212> ADN <213> Artificial Sequence <220> <223> Amorce	
<400> 38 acatttccaa tacatcccac	20
<210> 39 <211> 20 <212> ADN <213> Artificial Sequence <220> <223> Amorce	
<400> 39 atttgcatt tcacatgcac	20
<210> 40 <211> 20 <212> ADN <213> Artificial Sequence <220> <223> Amorce	
<400> 40 ttagaatgac tatacctgac	20
<210> 41 <211> 18 <212> ADN <213> Artificial Sequence <220> <223> Amorce	

<400> 41 tcaggttaag taagactc	18
<210> 42 <211> 19 <212> ADN <213> Artificial Sequence <220> <223> Amorce	
<400> 42 ttcctatcta cctagtgac	19
<210> 43 <211> 20 <212> ADN <213> Artificial Sequence <220> <223> Amorce	
<400> 43 ttttatagca agttgcctg	20
<210> 44 <211> 19 <212> ADN <213> Artificial Sequence <220> <223> Amorce	
<400> 44 cctatgaaga tcttggtac	19
<210> 45 <211> 20 <212> ADN <213> Artificial Sequence <220> <223> Amorce	
<400> 45 tgtcatgatt ctaacaagg	20
<210> 46 <211> 20 <212> ADN <213> Artificial Sequence <220> <223> Amorce	
<400> 46 tctatttcac tctgacatg	20
<210> 47 <211> 19 <212> ADN <213> Artificial Sequence <220> <223> Amorce	

<400> 47 gtcatagggc ttaggettcc	19
<210> 48 <211> 20 <212> ADN <213> Artificial Sequence <220> <223> Amorce	
<400> 48 atcatactac ccacttttcc	20
<210> 49 <211> 20 <212> ADN <213> Artificial Sequence <220> <223> Amorce	
<400> 49 tgtttgggaa gatgctactg	20
<210> 50 <211> 21 <212> ADN <213> Artificial Sequence <220> <223> Amorce	
<400> 50 ctactgaaga taacgtacat g	21
<210> 51 <211> 20 <212> ADN <213> Artificial Sequence <220> <223> Amorce	
<400> 51 cattgattgc catgtattgg	20
<210> 52 <211> 20 <212> ADN <213> Artificial Sequence <220> <223> Amorce	
<400> 52 agaaggccag aaatactcag	20
<210> 53 <211> 22 <212> ADN <213> Artificial Sequence <220> <223> Amorce	

<400> 53 gtactttaat cggtaaatat gg	22
<210> 54 <211> 20 <212> ADN <213> Artificial Sequence <220> <223> Amorce	
<400> 54 ctcaagtctt aggaatgcag	20
<210> 55 <211> 20 <212> ADN <213> Artificial Sequence <220> <223> Amorce	
<400> 55 gcacttaacc aggctgtatg	20
<210> 56 <211> 19 <212> ADN <213> Artificial Sequence <220> <223> Amorce	
<400> 56 ctcagatgac tcacatagc	19
<210> 57 <211> 22 <212> ADN <213> Artificial Sequence <220> <223> Amorce	
<400> 57 ctttactaga ctaattctcc tg	22
<210> 58 <211> 20 <212> ADN <213> Artificial Sequence <220> <223> Amorce	
<400> 58 cagattcaag aagacagatc	20
<210> 59 <211> 20 <212> ADN <213> Artificial Sequence <220> <223> Amorce	

<400> 59 gcaataattc accacacttg	20
<210> 60 <211> 20 <212> ADN <213> Artificial Sequence <220> <223> Amorce	
<400> 60 ggtagttcctt gtttctgctc	20
<210> 61 <211> 20 <212> ADN <213> Artificial Sequence <220> <223> Amorce	
<400> 61 caagtgtggt gaattattgc	20
<210> 62 <211> 19 <212> ADN <213> Artificial Sequence <220> <223> Amorce	
<400> 62 gagctgaaaa gtattcagc	19
<210> 63 <211> 20 <212> ADN <213> Artificial Sequence <220> <223> Amorce	
<400> 63 tgcaaaggac atagccagtg	20
<210> 64 <211> 20 <212> ADN <213> Artificial Sequence <220> <223> Amorce	
<400> 64 agcctctgga gatagtatgc	20
<210> 65 <211> 20 <212> ADN <213> Artificial Sequence <220> <223> Amorce	

<400> 65 ctagaacagg ggtcacagtc	20
<210> 66 <211> 18 <212> ADN <213> Artificial Sequence <220> <223> Amorce	
<400> 66 ttggacttct taaacttc	18
<210> 67 <211> 21 <212> ADN <213> Artificial Sequence <220> <223> Amorce	
<400> 67 gcagtatgca agaaattgaa c	21
<210> 68 <211> 19 <212> ADN <213> Artificial Sequence <220> <223> Amorce	
<400> 68 ggcctgtaat tttcttctg	19
<210> 69 <211> 21 <212> ADN <213> Artificial Sequence <220> <223> Amorce	
<400> 69 gtactgaata gatacatgta g	21
<210> 70 <211> 19 <212> ADN <213> Artificial Sequence <220> <223> Amorce	
<400> 70 gtgtagcaga tcaacatag	19
<210> 71 <211> 20 <212> ADN <213> Artificial Sequence <220> <223> Amorce	

<400> 71
catcttcaag tttggtgcac

20

<210> 72
<211> 1689
<212> ADN
<213> Mus musculus
<220>
<223> ADN Spg4 incomplet.

<400> 72
aggccgagag cgtcccgctc ttccacaagc aggccttcga gtacatctcc attgccctgc 60
gcatcgacga ggaagagaaa gcaggacaga aggaacaagc tgtggaatgg tataagaaag 120
gtatcgaaaga actggaaaaa ggaatcgctg ttatagttac gggccaaggt gaacagtatg 180
aaagagctag acgtcttcaa gcaaaaatga tgactaattt agttatggcc aaggaccgtt 240
tacaacttct agagaagctg caaccagttt tgcaattttc caagtccacag acggacgtct 300
ataacgagag tactaacctg acatgccgca atggacatct ccagtcagaa agtggagcag 360
ttccgaagag gaaagacccc ttaacacatg ctagtaattc attgcctcga tcaaaaactg 420
tctgaaaag tggctccgca gggtctcccg gtcaccacag ggcgcctagt tgcagtggtt 480
tgtccatggt ttctggagca agaccgggac ctggtcctgc agctaccaca cataagggta 540
ctccaaaacc aaatagaacc aacaaacctt ctactccac aactgcagtt cggaaaaaga 600
aagacttgaa aaattttagg aatgtggaca gcaatcttgc taaccttata atgaatgaaa 660
ttgttgacia tgggacagct gtaagtttg atgacatagc cgggcaggag ctggcaaacg 720
aagcgtgca ggagattgtc atccttcctt ctctgcccgc tgagttgttc acagggctca 780
gagctcctgc tagaggcttg ttactcttgc gtcgccagc aaacggaaaa acaatgctgg 840
ctaaagcagt agctgcagag tctaattgca cctttttcaa cataagtgtc gccagtttaa 900
cttcaaaata tgtgggagaa ggagagaaat tgggtgagagc tctctttgct gtggctcgag 960
aacttcaacc atctataatt ttatagatg aagtgacagc tcttttgtgt gagagacggg 1020
aaggggagca cgacgctagc agacggctaa agacggaatt tttaatagaa tttgacgggg 1080
tgcaatctgc tggagatgac agagtacttg taatgggtgc aactaacagg ccccaagagc 1140
ttgatgaagc tgttctcagg cgtttcatta aacgggtata tgtgtcctta ccaaatgagg 1200
agacaagact ccttctgctt aaaaacctgt tgtgtaaaaca aggaagtcca ctgacccaaa 1260
aagaactcgc acagcttgct agaattgaccg atggatactc tgggaagtgt ctgaccgctt 1320
tggccaagga tgcagccctg ggtcctatcc gagaactgaa gccagagcag gtgaagaata 1380
tgtctgccag tgagatgaga aatattcgat tatctgactt cacagaatcc ttaaaaaaga 1440
taaaacgcag tgtgagctct cagaccctag aagcatacat acgctggaac aaggattttg 1500
gagacaccac tgtttaaagg aatggatgcc tctgtgagcc catagaacat cgcacttcac 1560
aggaacaag agctttggct acaggaacct agacttcggt tacaggacgt tttagagttt 1620
tcattttgt gcaccaaact tgaagagga caagaagaca gacctaaata aaatatgcaa 1680
tatgaatgg 1689

<210> 73
<211> 504
<212> PRT
<213> Mus musculus
<220>
<223> Spastin murine incomplète.

<400> 73
Ala Glu Ser Val Arg Val Phe His Lys Gln Ala Phe Glu Tyr Ile Ser
1 5 10 15
Ile Ala Leu Arg Ile Asp Glu Glu Glu Lys Ala Gly Gln Lys Glu Gln
20 25 30
Ala Val Glu Trp Tyr Lys Lys Gly Ile Glu Glu Leu Glu Lys Gly Ile
35 40 45
Ala Val Ile Val Thr Gly Gln Gly Glu Gln Tyr Glu Arg Ala Arg Arg
50 55 60
Leu Gln Ala Lys Met Met Thr Asn Leu Val Met Ala Lys Asp Arg Leu
65 70 75 80

<220>
 <223> Site accepteur d'épissage du gène SPG4.

<400> 74
 atttttttatt ttaaagcagg acag 24

<210> 75
 <211> 24
 <212> ADN
 <213> Homo sapiens
 <220>
 <223> Site accepteur d'épissage du gène SPG4.

<400> 75
 aatttttttc tttcaggtga acag 24

<210> 76
 <211> 24
 <212> ADN
 <213> Homo sapiens
 <220>
 <223> Site accepteur d'épissage du gène SPG4.

<400> 76
 cttctctggt gcatagagaa gatg 24

<210> 77
 <211> 24
 <212> ADN
 <213> Homo sapiens
 <220>
 <223> Site accepteur d'épissage du gène SPG4.

<400> 77
 actttttcct tgtcagaaag tgga 24

<210> 78
 <211> 24
 <212> ADN
 <213> Homo sapiens
 <220>
 <223> Site accepteur d'épissage du gène SPG4.

<400> 78
 ttttgtatcc ttaagggtta ctcc 24

<210> 79
 <211> 24
 <212> ADN
 <213> Homo sapiens
 <220>
 <223> Site accepteur d'épissage du gène SPG4.

<400> 79
 aggtcttggt tcttagtgga acag 24

<210> 80
 <211> 24
 <212> ADN

<213> Homo sapiens
 <220>
 <223> Site accepteur d'épissage du gène SPG4.

<400> 80
 agtataatatt ttttagttgt tcac 24

<210> 81
 <211> 24
 <212> ADN
 <213> Homo sapiens
 <220>
 <223> Site accepteur d'épissage du gène SPG4.

<400> 81
 cttgtgattt ttaaaggcta aagc 24

<210> 82
 <211> 24
 <212> ADN
 <213> Homo sapiens
 <220>
 <223> Site accepteur d'épissage du gène SPG4.

<400> 82
 taatgctttg ttttaggtgg gaga 24

<210> 83
 <211> 24
 <212> ADN
 <213> Homo sapiens
 <220>
 <223> Site accepteur d'épissage du gène SPG4.

<400> 83
 cttgtatttc ctctagatga agtt 24

<210> 84
 <211> 24
 <212> ADN
 <213> Homo sapiens
 <220>
 <223> Site accepteur d'épissage du gène SPG4.

<400> 84
 gattttttgc ttgtaggtac agtc 24

<210> 85
 <211> 24
 <212> ADN
 <213> Homo sapiens
 <220>
 <223> Site accepteur d'épissage du gène SPG4.

<400> 85
 ggattttttt ttttaggcgt ttca 24

<210> 86

<211> 24
 <212> ADN
 <213> Homo sapiens
 <220>
 <223> Site accepteur d'épissage du gène SPG4.

<400> 86
 ttttaatat tttcagacaa gaçt 24

<210> 87
 <211> 24
 <212> ADN
 <213> Homo sapiens
 <220>
 <223> Site accepteur d'épissage du gène SPG4.

<400> 87
 tccttcctt cctcagaatg actg 24

<210> 88
 <211> 24
 <212> ADN
 <213> Homo sapiens
 <220>
 <223> Site accepteur d'épissage du gène SPG4.

<400> 88
 cttttatggt ttacagaact aaaa 24

<210> 89
 <211> 24
 <212> ADN
 <213> Homo sapiens
 <220>
 <223> Site accepteur d'épissage du gène SPG4.

<400> 89
 ctttttaaaa atctagatga gaaa 24

<210> 90
 <211> 24
 <212> ADN
 <213> Homo sapiens
 <220>
 <223> Site donneur d'épissage du gène SPG4.

<400> 90
 tgagaaaggt aactaggggg ctgg 24

<210> 91
 <211> 24
 <212> ADN
 <213> Homo sapiens
 <220>
 <223> Site donneur d'épissage du gène SPG4.

<400> 91
 aggacaaggt aagattgtat ttgt 24

<210> 92
 <211> 24
 <212> ADN
 <213> Homo sapiens
 <220>
 <223> Site donneur d'épissage du gène SPG4.

<400> 92
 acttctaggt atcaattaat gtat 24

<210> 93
 <211> 24
 <212> ADN
 <213> Homo sapiens
 <220>
 <223> Site donneur d'épissage du gène SPG4.

<400> 93
 ccagtcaggt gggtttaggt taac 24

<210> 94
 <211> 24
 <212> ADN
 <213> Homo sapiens
 <220>
 <223> Site donneur d'épissage du gène SPG4.

<400> 94
 ctcataaggt attctgggac agta 24

<210> 95
 <211> 24
 <212> ADN
 <213> Homo sapiens
 <220>
 <223> Site donneur d'épissage du gène SPG4.

<400> 95
 gtggacaagt aagttttgcc atct 24

<210> 96
 <211> 24
 <212> ADN
 <213> Homo sapiens
 <220>
 <223> Site donneur d'épissage du gène SPG4.

<400> 96
 ggcctgaggt aagaacttta tatt 24

<210> 97
 <211> 24
 <212> ADN
 <213> Homo sapiens
 <220>
 <223> Site donneur d'épissage du gène SPG4.

<400> 97
 caatgctggt aagggttctc ttca 24

<210> 98
<211> 24
<212> ADN
<213> Homo sapiens
<220>
<223> Site donneur d'épissage du gène SPG4.

<400> 98
caaaatacgt gagtgctctg tttc 24

<210> 99
<211> 24
<212> ADN
<213> Homo sapiens
<220>
<223> Site donneur d'épissage du gène SPG4.

<400> 99
ttttataggt aagaacatat tttc 24

<210> 100
<211> 24
<212> ADN
<213> Homo sapiens
<220>
<223> Site donneur d'épissage du gène SPG4.

<400> 100
ttgatggtgt aagtgtgat tatg 24

<210> 101
<211> 24
<212> ADN
<213> Homo sapiens
<220>
<223> Site donneur d'épissage du gène SPG4.

<400> 101
gttctcaggt agggagattt atat 24

<210> 102
<211> 24
<212> ADN
<213> Homo sapiens
<220>
<223> Site donneur d'épissage du gène SPG4.

<400> 102
atgaggaggt atgtatctgt gttt 24

<210> 103
<211> 24
<212> ADN
<213> Homo sapiens
<220>
<223> Site donneur d'épissage du gène SPG4.

<400> 103
cttgctaggt gagtaatttg gatt 24

<210> 104
 <211> 24
 <212> ADN
 <213> Homo sapiens
 <220>
 <223> Site donneur d'épissage du gène SPG4.

<400> 104
 tatccgaggt aggtatacaa gagc 24

<210> 105
 <211> 24
 <212> ADN
 <213> Homo sapiens
 <220>
 <223> Site donneur d'épissage du gène SPG4.

<400> 105
 ccagtgaggt atagtatttt acaa 24

<210> 106
 <211> 2152
 <212> ADN
 <213> Mus musculus
 <220>
 <223> ADN Spg4 complet

<400> 106
 gctcccgggc tccggcgggc gcgcgagcgg ctccgtgggc ccccgcggcc gcagtgagcag 60
 tggccgcccgc cgcgcgttgg tcgcccgtcgg tctgcgggaa gcgggttatg gcggcggcgg 120
 cagtgaggagc tgtgaatgag ttctccggcc ggacgacgga agaagaaagg ctccggcggc 180
 gcgagcccgg cgcccggcgc gccctccgcc cccgcgcggg tccccgcccc tgccgcggc 240
 ccggcccctg cggccggctc gccgcctaag cggaaacctgt cttctttctc gtcccggctg 300
 gtcgtcggct tcgccctgct gcgcctgctg gectgccacc tggggctcct cttcgcgtgg 360
 ctctgccagc gcttctcccg cgcctcatg gccccaaga ggagctccgg gaccgcggc 420
 gcgcccgcct cgcctccgc cccagcggcc ggaccgggtg gcgaggccga gaggctccgc 480
 gtcttcacaca agcaggcctt cgagtacatc tccattgccc tgcgcatcga cgaggaaagag 540
 aaagcaggac agaaggaaca agctgtggaa tggataaga aaggtatcga agaactggaa 600
 aaaggaatcg ctgttatagt tacgggcaa ggtgaacagt atgaaagagc tagacgtctt 660
 caagccaaaa tgatgactaa tttagttatg gccaaaggacc gtttacaact tctagagaag 720
 ctgcaaccag ttttgcaatt ttccaagtca cagacggacg tctateacga gagtactaac 780
 ctgacatgcc gcaatggaca tctccagtca gaaagtggag cagttccgaa gaggaaagac 840
 cccttaacac atgctagtaa ttcattgcct cgatcaaaaa ctgtcctgaa aagtggctcc 900
 gcagggctct ccggtcacca caggggcctt agttgcagtg gtttgtccat ggtttctgga 960
 gcaagaccgg gacctgtcc tcagctacc acacataagg gtactccaaa accaaataga 1020
 accaacaac cttctactcc cacaactgca gttcggaaaa agaaagactt gaaaaatatt 1080
 aggaatgtgg acagcaatct tgctaaccctt ataataatgaa aaattgttga caatgggaca 1140
 gctgttaagt ttgatgacat agccgggagc gagctggcaa agcaagcgtc gcaggagatt 1200
 gtcatccttc cttctctgog gcctgagttg ttcacagggc tcagagctcc tgctagaggc 1260
 ttgttactct tcggtcggcc aggaaacgga aaaaacatgc tggctaaagc agtagctgca 1320
 gagtctaagt cgaccttttt caacataagt gctgccagtt taacttcaaa atatgtggga 1380
 gaaggagaga aattggtgag agctctcttt gctgtggctc gagaacttca accatctata 1440
 atttttatag atgaagttga cagtcttttg tgtgagagac gggaaagggga gcacgacgct 1500
 agcagacggc taaagacgga atttttaata gaatttgacg ggggtcaatc tgctggagat 1560
 gacagagtac ttgtaatggg tgcaactaac agggcccaag agcttgatga agctgttctc 1620
 aggcgtttca ttaaaccggg atatgtgtcc ttaccaaag aggagacaag actccttctg 1680
 cttaaaaacc tgttgtgtaa acaaggaagt ccaactgacc aaaaagaact cgcacagctt 1740
 gctagaatga ccgatggata ctctggaagt gatctgaccg ctttgcccaa ggatgcagcc 1800
 ctgggtccta tccgagaact gaagccagag caggtgaaga atatgtctgc cagtgagatg 1860

agaaatattc gattatctga cttcacagaa tccttaaaaa agataaaacg cagtgtgagt 1920
 cctcagacct tagaagcata catacgtgga aacaaggatt ttggagacac cactgtttaa 1980
 aggaatggat gcctctgtga gcccatagaa catcgcactt cacaggaaac aagagctttg 2040
 gctacaggaa cccagacttc gtttacagga cgttttagag ttttcatttt tgtgcaccaa 2100
 acttgaagag gaacaagaag acagacctaa ataaaatag caatatgaat gg 2152

<210> 107

<211> 614

<212> PRT

<213> Mus musculus

<220>

<223> Spastin murine complete

<400> 107

Met Ser Ser Pro Ala Gly Arg Arg Lys Lys Lys Gly Ser Gly Gly Ala
 1 5 10 15
 Ser Pro Ala Pro Ala Arg Pro Pro Pro Pro Ala Ala Val Pro Ala Pro
 20 25 30
 Ala Ala Gly Pro Ala Pro Ala Ala Gly Ser Pro Pro Lys Arg Asn Leu
 35 40 45
 Ser Ser Phe Ser Ser Pro Leu Val Val Gly Phe Ala Leu Leu Arg Leu
 50 55 60
 Leu Ala Cys His Leu Gly Leu Leu Phe Ala Trp Leu Cys Gln Arg Phe
 65 70 75
 Ser Arg Ala Leu Met Ala Ala Lys Arg Ser Ser Gly Thr Ala Pro Ala
 85 90 95
 Pro Ala Ser Pro Ser Pro Pro Ala Pro Gly Pro Gly Gly Glu Ala Glu
 100 105 110
 Ser Val Arg Val Phe His Lys Gln Ala Phe Glu Tyr Ile Ser Ile Ala
 115 120 125
 Leu Arg Ile Asp Glu Glu Glu Lys Ala Gly Gln Lys Glu Gln Ala Val
 130 135 140
 Glu Trp Tyr Lys Lys Gly Ile Glu Glu Leu Glu Lys Gly Ile Ala Val
 145 150 155
 Ile Val Thr Gly Gln Gly Glu Gln Tyr Glu Arg Ala Arg Arg Leu Gln
 165 170 175
 Ala Lys Met Met Thr Asn Leu Val Met Ala Lys Asp Arg Leu Gln Leu
 180 185 190
 Leu Glu Lys Leu Gln Pro Val Leu Gln Phe Ser Lys Ser Gln Thr Asp
 195 200 205
 Val Tyr Asn Glu Ser Thr Asn Leu Thr Cys Arg Asn Gly His Leu Gln
 210 215 220
 Ser Glu Ser Gly Ala Val Pro Lys Arg Lys Asp Pro Leu Thr His Ala
 225 230 235 240
 Ser Asn Ser Leu Pro Arg Ser Lys Thr Val Leu Lys Ser Gly Ser Ala
 245 250 255
 Gly Leu Ser Gly His His Arg Ala Pro Ser Cys Ser Gly Leu Ser Met
 260 265 270
 Val Ser Gly Ala Arg Pro Gly Pro Gly Pro Ala Ala Thr Thr His Lys
 275 280 285
 Gly Thr Pro Lys Pro Asn Arg Thr Asn Lys Pro Ser Thr Pro Thr Thr
 290 300
 Ala Val Arg Lys Lys Lys Asp Leu Lys Asn Phe Arg Asn Val Asp Ser
 305 310 315 320
 Asn Leu Ala Asn Leu Ile Met Asn Glu Ile Val Asp Asn Gly Thr Ala
 325 330 335
 Val Lys Phe Asp Asp Ile Ala Gly Gln Glu Leu Ala Lys Gln Ala Leu
 340 345 350

Gln Glu Ile Val Ile Leu Pro Ser Leu Arg Pro Glu Leu Phe Thr Gly
 355 360 365
 Leu Arg Ala Pro Ala Arg Gly Leu Leu Leu Phe Gly Pro Pro Gly Asn
 370 375 380
 Gly Lys Thr Met Leu Ala Lys Ala Val Ala Ala Glu Ser Asn Ala Thr
 385 390 395 400
 Phe Phe Asn Ile Ser Ala Ala Ser Leu Thr Ser Lys Tyr Val Gly Glu
 405 410 415
 Gly Glu Lys Leu Val Arg Ala Leu Phe Ala Val Ala Arg Glu Leu Gln
 420 425 430
 Pro Ser Ile Ile Phe Ile Asp Glu Val Asp Ser Leu Leu Cys Glu Arg
 435 440 445
 Arg Glu Gly Glu His Asp Ala Ser Arg Arg Leu Lys Thr Glu Phe Leu
 450 455 460
 Ile Glu Phe Asp Gly Val Gln Ser Ala Gly Asp Asp Arg Val Leu Val
 465 470 475 480
 Met Gly Ala Thr Asn Arg Pro Gln Glu Leu Asp Glu Ala Val Leu Arg
 485 490 495
 Arg Phe Ile Lys Arg Val Tyr Val Ser Leu Pro Asn Glu Glu Thr Arg
 500 505 510
 Leu Leu Leu Leu Lys Asn Leu Leu Cys Lys Gln Gly Ser Pro Leu Thr
 515 520 525
 Gln Lys Glu Leu Ala Gln Leu Ala Arg Met Thr Asp Gly Tyr Ser Gly
 530 535 540
 Ser Asp Leu Thr Ala Leu Ala Lys Asp Ala Ala Leu Gly Pro Ile Arg
 545 550 555 560
 Glu Leu Lys Pro Glu Gln Val Lys Asn Met Ser Ala Ser Glu Met Arg
 565 570 575
 Asn Ile Arg Leu Ser Asp Phe Thr Glu Ser Leu Lys Lys Ile Lys Arg
 580 585 590
 Ser Val Ser Pro Gln Thr Leu Glu Ala Tyr Ile Arg Trp Asn Lys Asp
 595 600 605
 Phe Gly Asp Thr Thr Val
 610

【図面の簡単な説明】

【図1】

S P G 4 の範囲およびS P G 4 のゲノム構成の物理的マップを示す。

図1A： 太字で示されるD 2 S 3 5 2およびD 2 S 2 3 4 7遺伝子マーカーによって1.5 Mbの候補領域が定められている。多形マーカーおよびその他のSTSの位置は標準的な文字で示され、ESTの位置はイタリックで示されている。予備配列決定マップを構成するBACクローンが四角で示され、名称が上に示され、決定されていればクローンの正確な大きさが下に示されている。BAC A、B、Cなどの名称の後には、クローンがBACライブラリーCITB_978__SKBに由来する場合には「b」で始まり、またはライブラリーRPSI-11に由来する場合には「B」で始まるクローン名を含む括弧で示されている。

図1B： SAC D (b 3 3 6 P 1 4) およびG (B 5 6 3 N 4) が重複するS P G 4 遺伝子の模式図。エキソン黒い四角で示され、上にそれらの名称が

示されている。

図1C： 7つのSPG4遺伝子座関連AD-HSP家系で同定されている5つの突然変異がエキソン7、11および13、ならびにイントロン15のスプライス受容部位の位置にある。

【図2】

スパスチンのSPG4 cDNAの核酸およびタンパク質配列を示す。

下にある数字とともに17本の垂直な線が種々のエキソン間の連結を示している。ATG開始コドンはnt126~128番に位置し、終結に関する停止コドンはnt1974~1976番に位置している。エキソン16の欠失を含むこれまでに確認された5つの突然変異はイタリックで示されている(nt1210、nt1468、nt1520、nt1620およびエキソン16の欠失に関するnt1813~1853)。ポリアデニル化部位は下線を施したイタリックで示されている。推定核局在化シグナル(NLS)、RGKKK、またProDomデータベースにおける解析によって推定された3つの保存されたドメインはそれぞれaa7~11番(NLS)、342~409番(ドメイン92)、411~509番(ドメイン179)、および512~599番(ドメイン6226)に位置している。Prositeデータベースにおける配列比較によって推定された4つのモチーフはaa50~78番と508~529番の2つの「ロイシンジッパー」、aa382~386番のATP結合部位(またはWalker Aモチーフ)、およびaa478~486番の「ヘリックス-ループ-ヘリックス」二量化ドメインである。Walker AおよびBモチーフ「GPPGNGKT」および「IIFIDE」、またAAA最小共通[欠文]に下線で示されている。

【図3A】

3つのSPG4遺伝子座関連AD-HPS家系の影響を受けた個々のものにおけるスプライス部位突然変異の同定。

図3Aは、リンパ芽球cDNAを用いるSPG4 cDNAの断片IVのPCR増幅を示す：ウェルM、サイズマーカーVII(Boehringer)；ウェル1、影響を受けていない家系2992のメンバー；ウェル2、家系2992の患者；ウェ

ル3、影響を受けていない家系5330のメンバー；ウェル4、家系5330の患者；ウェル5、家系5226の患者；ウェル6、陰性対照（ヒトゲノムDNA）。

【図3B】

3つのSPG4遺伝子座関連AD-HP5家系の影響を受けた個々のものにおけるスプライス部位突然変異の同定。

図3Bは、イントロン15のスプライス受容部位の突然変異の配列グラフを示す。上は対照個体および下は家系2992の患者のゲノム配列。nt1813~4番のアスタリスクは患者のイントロン15のスプライス受容部位の非保存的ヌクレオチドに影響を及ぼすA/Cの多形性を示している。

【図4A】

スパスチンの相同性。同一の残基は網掛け領域で強調されている。

種々の生物に由来する8つのタンパク質のCLUSTALWによって作り出され、ヒト・スパスチンとネズミ・スパスチン（配列番号73）と強い配列相同性を有する複数のアライメント。

【図4B】

スパスチンの相同性。同一の残基は網掛け領域で強調されている。

酵母メタロプロテアーゼAFG3、RCA1およびYME1、ならびにヒト・プララプレジンおよびスパスチンのCLUSTALWによるアライメント。

【図5】

SPG4 cDNAおよびそのマウス・オルトログSp g 4の核酸配列（配列番号73）のBLASTNによるアライメント。ネズミcDNAのポリアデニル化部位は下線を施したイタリックで示されている。停止コドンはネズミcDNAではnt1515~1517番に位置し、ヒトcDNAではnt1974~1976番に位置している。

【図6】

SPG4およびそのネズミ・オルトログSp g 4の発現のPCR解析。

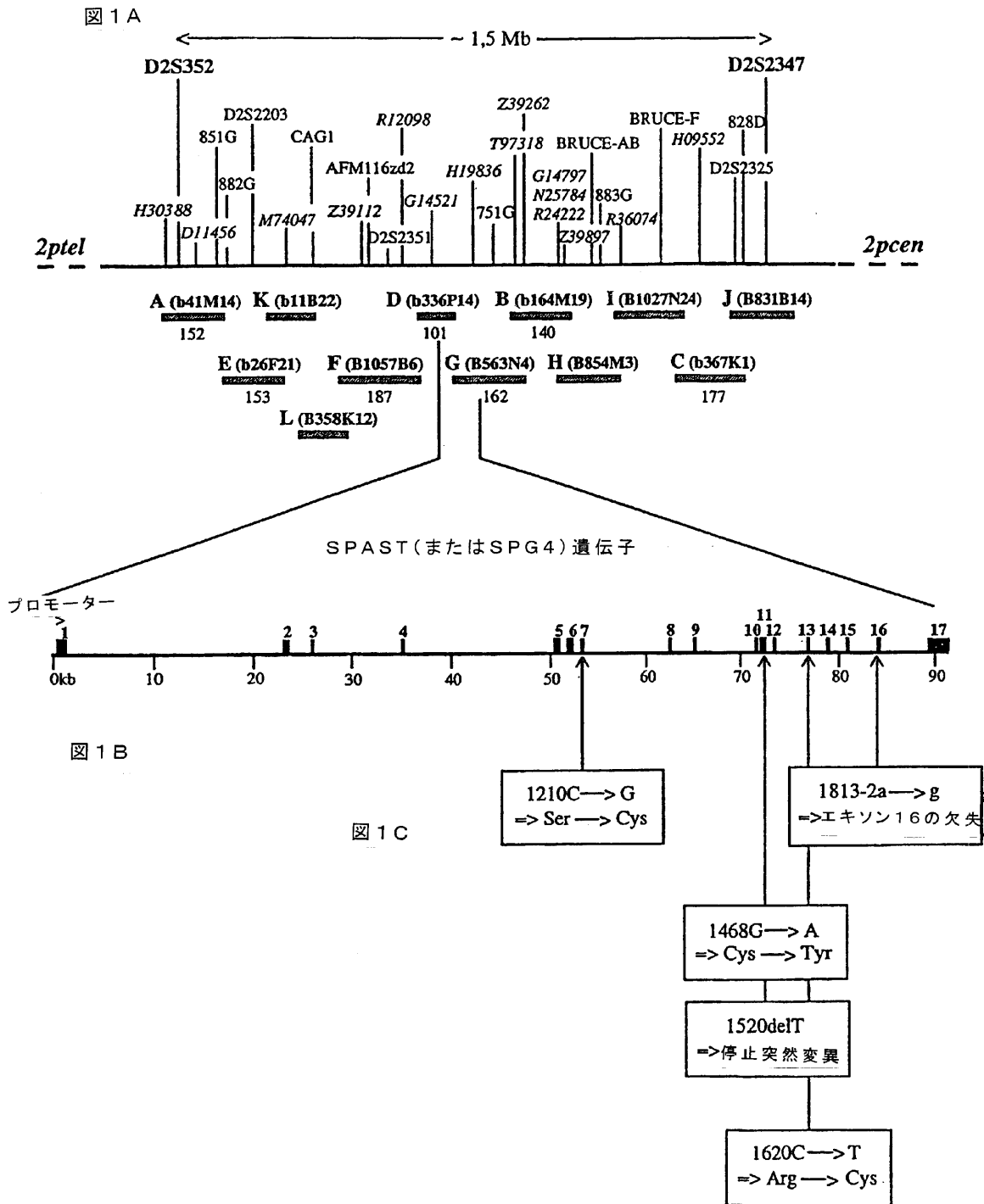
図6A： 複数のマウス組織に由来するcDNAコレクション。ウェルM、サイズマーカーV(Boehringer)；ウェル1、心臓；ウェル2、脳；ウェル3、脾臓

; ウェル4、肺; ウェル5、肝臓; ウェル6、骨格筋; ウェル7、腎臓; ウェル8、精巣; ウェル9、E7 7日胚; ウェル10、E11 11日胚; ウェル11、E15 15日胚; ウェル12、E17 17日胚; ウェル13、陰性対照(ヒトゲノムDNA)。

図6B: 複数のヒト組織に由来するcDNAコレクション。ウェルM、サイズマーカーVII(Boehringer); ウェル1、脳; ウェル2、心臓; ウェル3、腎臓; ウェル4、肝臓; ウェル5、肺; ウェル6、膵臓; ウェル7、胎盤; ウェル8、骨格筋; ウェル9、陰性対照(ヒトゲノムDNA); ウェル10、陰性対照(DNAを含まず)。

図6C: 複数のヒト胎児組織に由来するcDNAコレクション。ウェルM、サイズマーカーV(Boehringer); ウェル1、脳; ウェル2、心臓; ウェル3、腎臓; ウェル4、肝臓; ウェル5、肺; ウェル6、骨格筋; ウェル7、脾臓; ウェル8、胸腺; ウェル9、陰性対照(ヒトゲノムDNA); ウェル10、陰性対照(DNAを含まず)。

【図1】

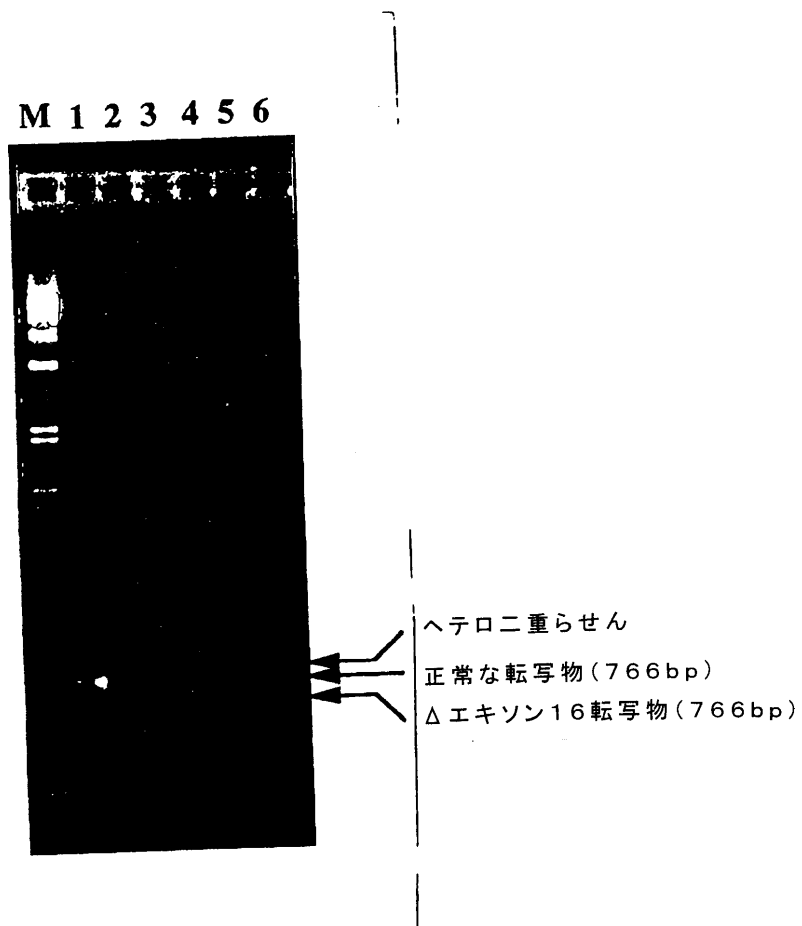


【図2】

GCTCCTGAGACCGGGGGCACACGGGGTCTGTGGCCCCCGCTAGCAGTGGCTGCCCGCTCGCTTGGTTCCCGTCCGCTGCGGGAGGCGGG 95
 1 TTATGGCGGGCGCGCAGTGTAGAGCTGTGAATGAATCTCCGGGTGGACGAGGGAAGAAAGGCTCCGGGGCGCCAGCAACCCGGTGCCTCC 190
 M N S P G G R G K K K G S G G A S N P V P P
 23 CAGGCOCTCCGCCCTTGCCTGGCCCCCGCCCTCCCGCCCGGGCCCGCCCTCCGCCGAGTCCGCCGATAAGCGGAACCTGTACTATTCT 285
 R P P P P C L A P A P P A A G P A P P P E S P H K R N L Y Y F S
 CCTACCCGCTGTTTGTAGGCTTCGCGGTGCTGGTTCGCTTCCACCTGGGGCTCCCTTCCTGCTGGCTCTGCCAGCGCTTCTCCCGCGCC 380
 Y P L F V G P A L L R L V A F H L G L L F V W L C Q R F S R A
 55 CTCAATGGCAGCAAGAGGAGCTCCGGGGCCGCCAGCACCTGCCTCGGCTCCGGCCCGGGCCGGTCCGGGGCGGAGGGCCGAGCCGCTCCG 475
 L M A A K R S S G A A P A P A S A P A P V P G G E A E R V R
 86 AGTCTTCCACAAAACAGGCTTCGAGTACATCTCCATTGCCCTGCCCATCGATGAGGATGAGAAAGTAGGACAGAAGGACCAAGCTGTGGAATGGT 570
 V F H K Q A F E Y I S I A L R I D E D E K A 2 G Q K E Q A V E W Y
 118 ATAAGAAAGGTATTGAAGAACTGGAAGAAAGGAATAGCTGTATAGTTACAGGACAAGTAGAACAGTGTGAAAGAGCTAGACGCCCTTCAAGCTAAA 665
 K K G I E E L E K G I A V I V T G Q G 3 E Q C E R A R R L Q A K
 150 ATGATGACTAATTTGGTTATGGCCAAAGGACCGCTTACAACCTCTAGTGAAGATCAACCACTTTTGGCCATTTTCCAACTCACAACCGGACGCTTA 760
 M M T N L V M A K D R L Q L L E 4 K M Q P V L P F S K S T P T A
 181 TAATGACAGTACTAAGTGGCATGCCCAATGGACATCTCCAGTCAAGTAGGAGCTGTCCAAAAGAAAAGGCCCTTAAACACACTAGTA 855
 L R A A K R S S G A A P A P A S A P A P V P G G E A E R V R
 213 N D S T N L A C R N G H L Q S E 5 S G A V P K R K D P L T H T S N
 ATTCACTGCCCTGTTCAAAAACAGTTATGAAAACAGGATCTGCAGGCTTTCAGGCCACCATAGAGCACCTAGTTACAGTGGTTTATCCATGGTT 950
 245 S L P R S K T V H K T G S A G L S G H H R A P S Y S G L S M V
 TCTGGAGTGAACAGGATCTGGTCCTTACCACCTATAAGTGTACTCCGAAAACAATAGGACAAAATAACCTTCTACCCCTACAACCTGC 1045
 276 S G V K Q G S G P A P T T H K G 6 T P K T N R T R K P S T P T A
 TACTCGTAAGAAAAAGACTTGAAGAAATTTAGGAATGGACAGCAACCTTGTCAACCTTATAATGAATGAATTTGGGACATTTGGAACAGCTG 1140
 308 T R K K K D L K N F R N V D S N L A N L I M N E I V D N G 7 T A V
 TTAATAATTTGATGATAGCTGGTCAAGACTTGGCAAAAACAGCAATGCAAGAAATTTTATCTTCTCTGAGGCCCTGATTTGTTCCAGGG 1235
 340 K F D D I A G Q D L A K Q A L Q E I V I L P S L R P E L E F T G
 CTTAGAGCTCCTGCCAGAGGCTGTTACTCTTTGGTCCACCTGGGAATGGGAAGACAATGCTGCTTAAAGCAGTAGCTGCAGAAATCGAATGCAAC 1330
 371 L R A P A R G L L L F G P P G N G K T M L A 9 K A V A A E S N A T
 CTTCTTTAATAAAGTGTGCAAGTTTAACTTCAAATACTTGGGAGAAAGGAGAAATTTGGTGGGGCTCTTTTGGCTGGCTCGAGAATTC 1425
 403 P F N I S A A S L T S K Y V 10 G E G E K L V R A L F A V A R E L Q
 AACCTTCTATAATTTTATAGTTGAAGTTGATAGCCTTTTGTGAAAGAAGGAGGAGCAGCATGCTAGTAGACCCCTAAAACCTGAAATTT 1520
 435 P S I I F I D I E V D S L L C E R R E G E H D A S R R L K T E F
 CTAATAGAAATTTGATGGTGTACAGTCTGCTGGAGATGACAGAGTACTTGTAAATGGGTGCAACTAATAGGCCACAAGAGCTTGAAGAGGCTGTCT 1615
 466 L I E F D G V 120 S A G D D R V L V M G A T N R P Q E L D E A V L
 CAGCGTTTCATCAAAACGGGTATATGTCTTTACCAATGAGGAGCAAGACTACTTTTGGTTAAAAATCTGTATGTAACAAGGAAGTCCAT 1710
 498 R 13 R F I K R V Y V S L P N E E T 14 R L L L L K N L L L C K Q G S P L
 TGACCCAAAAGAACTAGCACAACTTGTAGTATGACTGATGGATATCAGGAAGTACCTAACAGCTTTGGCAAAAGATGGAGCACTCGGTCCT 1805
 530 T O K E L A Q L A R M 15 T D G Y S G S D L T A L A K D A A L G P
 ATCCGAGTAACTAAAACCAACAGGTTGAAGAAATATGCTGCCAGTCAATGAGAAATATTCGATTATCTGACTTCACTGAATCCTTGAAAAAAAT 1900
 561 I R E 16 L K P E Q V K N M S A S E M 17 R N I R L S D F T E S L K K I
 AAAACGAGCGTACGCCCTCAAACCTTAGAAGCGTACATACGTTGGAACAAGGACTTTGGAGATACCCTGTTAAGGAAATACCTTTGTA AAC 1995
 593 K R S V S P Q T L E A Y I R W N K D F G D T T V *
 TGCAGAACATTTTACTTAAAGAGGAAACACAAGATCTTCAATGAACCTCATCCGCTACAGAAAACGCCCTAAGTTTACAGGACTTTTATAGACTCT 2090
 TACATATTTGTGCACCAACTTGAAGATGAACCGAAAACAGACTTAAACAAAATATACAAATGCAATGTAATTTTGTGTTTAAAGCCCTTGC 2185
 CTTGATGGTCACTTTATCCCAA TGGACACTAAGTTAGAGCACAACAAACCTGATCTGGTCTTCTTTACCAATATAATCATATGTAATAAT 2280
 AATTTGATATTTGTTGCGATGAAAGTATCCAGGAACAGTGAATGGTAGAAGACACAAGAACATTTGTTGTTGCTCTCTGATGTTTTTTC 2375
 TTAATAATGATATTTCTCTACTTTCTTTTCTACTGTGTCTTAACTACAGGTTGATGGAATGCCAAACACTCTTAAGTTTATTTCTTTTTTC 2470
 GTTTATAAATTCAGTGTGCCAAATGAAACCTTTTCTTAAGTAACTGTAATAGGAAAAAGTTTATTTTGGAGTTTCTTCTCATRAATCTACA 2565
 GACATTAACCAATTTGTTGCTTTTACCCTTTATTTCTATTACCTTGTACCAACAGTTTAGATAGCAATATAATAGCAAAAAGCAAAA 2660
 TATGGTAAATAGAGAAGGTTTGAAGGTTTGAAGTACTCTGTATATAACATGATGATCAGTCTTCAATGACCTGCAGTATTTTCTTTCTAAT 2755
 GTATTTGTCAGAAATCTGTTGTAGACTGTTAACTTCTTCTGATGGAATTTATTTTCTGCAAGAAATTTATTTCTGATATTTAAGAGACCAATTTA 2850
 ACTGCTGTGAAAATGTTTCCAGTGAAGAGAAGGAAATACTAGGAATAAGACATTTCTAATTTATGCTTATTTACTTCTTAAATTTTACAGGA 2945
 TAATTAAGCAAGTGAACCTACATCTTTTATCTTAAATTTAATTCCTTCAATGAACTTTAAAAAACCTGAAATTTTATACATGGCAT 3040
 ACATTTTCTAGTTCTTCTGCTTGTCTTATTAACCTAAAAGTTCTAGTCTTAGTCTGTTGATCTGCCTTTTGTCTCCCAAAATGTACAGTAAAT 3135
 TCCATTTGTTGTAATAATAGCCTGGATTTTCAATTAATAAATGTCATTGTAGGGAGTAGAGACTCATATCATGGCCTTTAAATAATTTGTAATA 3230
 AAGGCAATAGATATTTGCCCTTACTTACTCG 3263

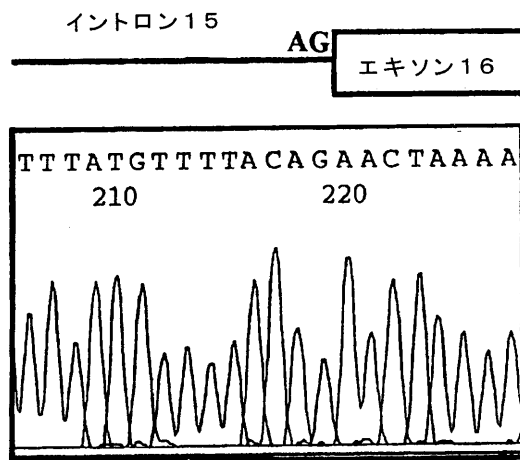
FIGURE 2

【図3A】

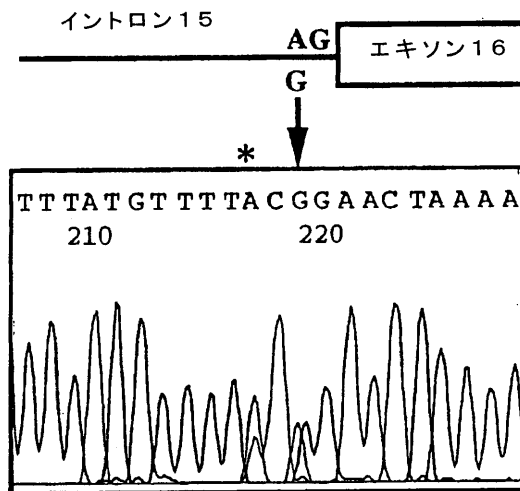


【図3B】

正常なスプライス受容部位



変異型スプライス受容部位



【図4B】

```

AFG3_YEAST ---MMMWRYANGAFR-----SLTSL-----FGKASRISTVKF-----VLKSRMPVHQRLQTLGSLGATR-----NTIHR--S
RCAL_YEAST -MLLLSWSRIATKVVRRPVRFRSYGLTHIKSLHTQYRLNLRLQENKSGKNEDONNEDAKLNKEIETDEEVEAIKKQVREYIEQTKMNTIPANWK
PARAPLEGIN MAVLLLLLRALRRCPGCPFR-----FLWGPFAWS-PGFPARPGRRP-----YMASRFFGD--LABAGGRALQ-----S
YME1_YEAST ---MNVSKILVSPVTVTN-----VLRIFAPRLPQIGASLLVQKKWA-----LRSKKFFRYSEKNSGMPF-----K
SPAST_HUMAN ---MNSPGGHRGKKGSGGASN-----PVFPRFPFCIAPAPFAAGFAFPF-----ESPFRNLYYFSYFLVGFALLLEVAFHLGLLFLVHLG

AFG3_YEAST TQIRSFHISNTRLN-----ERRPKK--GCKNGKNDNNSKEDGKDRRNEFGSL-----EYFRSKRFANTMFLTIGFTLIF
RCAL_YEAST EQKKLDESRRLEDAVLEKQESNRQRRKKEEENGPSKAKSNRTKQCYFEGNNSRNIPFPFPPFPKPELNDPNSFVSKNVNLPQIGLTFLL
PARAPLEGIN LQLRLLPFFEGINGLLIKQH--LVQNFVRLWQLGGTFYFHTSRLKQNKERKDRSKGKAPEDD-----ERRRRERDDQMYRERLRLLV
YME1_YEAST QRSSEKSKAKKST-----LSSIDNSQF--PPSNTNDRTKQANVAVSHAMLATREOE-----RKDLTSPDAQAAFYKLLLSQSNY
SPAST_HUMAN QRFRRALMAAKRSSGAAPF---ASASAPAPVF-GGEAEVRVFRKQAFYISIALRLDEDEKAG-----QKEQAVWEYKKGIEELEKCIIV

AFG3_YEAST TLLTFSSNNSGDDSNRVLTFQDFKTKYLEKGLVSKIYVHKF-LVEAELVN-----TKQVVSFTIGSV---DIFPEQ--MDQIQDLLNI
RCAL_YEAST LSLFLDLLNSLEEQS-EITWQDFRKKLLAKGYVAKLIVHKS-MVKVMLNDNGRNQADNYGRNFYFTIGSI---DSFEK--LQKAQDELDI
PARAPLEGIN IAVVMSLLNALSTSGGISWDFVEMLAKGEVQRVQVPSDDVVEVYLPQAVVGRPRALMWRMQVANI---DKFEK--LRAKDELDNI
YME1_YEAST PQYVVSFRFETFGIASSPECMELYMALQRIGRHEADAQRH-LLYASSAG-----AVNFSLASSSSNQ---SGYGH--FESMYSFLYG
SPAST_HUMAN IVTGQGGQCEARRRLLQAKMMTNLVMKDRLLLEKMQVLPFSKSGTQVYNDSTN--LACRNGHLQSESGAVFRKDELTRTNSLFRSKTVMKT

AFG3_YEAST PFRDRIDIRYIERSSPFFELFF---LPTLILGLCYFITRKRINSSPPNANGGGCGGLGCMFVYKSKARLENKETDIK--ISKNVA-CDE
RCAL_YEAST DKDFRDPVLVYQGNNAKAMEQI---LPTVLMHAGIIMLTRS-----AQAGGSRGGIFGLSRKAKKFTEDTVK--IKKVA-CDE
PARAPLEGIN EAKDRIDVSKYKRTGFFGNALYSV---GMTAVGLAILWVFRLAG-----MTGREGGSAFNQLKMARFTIVDGMKGGK--VSKVA-MRE
YME1_YEAST SRKE--PLWVVVSESTFTVSRW---VKWLVVGLTYFSEG-----FKYTEHTTLKSSVANKSVDAKTN--VKKDDVC-CDE
SPAST_HUMAN GSAGLSGHRAPSYSGLSMVSGVKGSGPAPTRTKGTPKTRNTEFSEFTATRKKKDLKFRNVDSNLANLIMNEVLDNGTAVKDDIAQDL

AFG3_YEAST KQELMRFHFLKN-GKYTKLGAIPRAISGKATLKKKAGANVFLSVSGSEFVEMFVVAARVDTQSMASIFLI
RCAL_YEAST KEELMRFSELKE-SRYERMGAIPRAISGKATLKKKAGVYFVSGSEFVEMFVVAARVDTQSMASIFLI
PARAPLEGIN KLEVRFDYKLS-ERELQLGARVPRALVLCGKATLKKKAGVYFVSGSEFVEMFVVAARVDTQSMASIFLI
YME1_YEAST RABLEYDFLKD-TKYESLGGKAPVLTGKATLKKKAGVDFVSGSEFVEMFVVAARVDTQSMASIFLI
SPAST_HUMAN KQALQTLFLPSLRELFTGLRAPA-RLLEKATLKKKAGVDFVSGSEFVEMFVVAARVDTQSMASIFLI

AFG3_YEAST AIGKEGKGGALGGANDEREATLNQVVMFTTS--DQVVVLAGRFDVNLMPGFDRIQIDSDVNGQQIYLVRLKRLNLDPELLT
RCAL_YEAST AICKAKGK-NEGGANDEREATLNQVVMFTTS--DQVVVLAGRFDVNLMPGFDRIQIDSDVNGQQIYLVRLKRLNLDPELLT
PARAPLEGIN AVGKKSTT-MSGFSNTEEEQLNLQVVMGTT--DQVVVLAGRADIQLMPGFDRIQIDSDVNGQQIYLVRLKRLNLDPELLT
YME1_YEAST AIGKEGK--DQYAKQTLNLQVVMFTTS--DQVVVLAGRADIQLMPGFDRIQIDSDVNGQQIYLVRLKRLNLDPELLT
SPAST_HUMAN SLLCEREG---EHDASREKTRFVAVVQVQAGDVRVVMGARRPQEFVLF--FKRVVYSLMEETLILLKRLNLDPELLTQK

AFG3_YEAST DDMNLSGKLTLEPFTADIANACNELIAARNDOPYITIHFFEGIERVLAGLEKTRVSKEEKRSVAYEACHVCGWFLKYADPFLKV
RCAL_YEAST DLKN---REELPFTADIANACNELIAARNDOPYITIHFFEGIERVLAGLEKTRVSKEEKRSVAYEACHVCGWFLKYADPFLKV
PARAPLEGIN FYSQ---REELPFTADIANACNELIAAREGTSVHTLHFEYIERVLAGTAKKSKLSEEGKVVAFHESCHLVGWMLDHTVANKV
YME1_YEAST FTIT---RCPGLSIAELANLVQVAVYACQKAVSVDMSEFKKILMGASRRTHVTDAAKATSEHACHAKAYTMGATFLKA
SPAST_HUMAN ELAQ---LRMDVSDLTALAKDELGPITRELKFEQVKN--MSSEMNRNIELSDFTESKKIKR--SVSPQTLTYIRNWKDFGDTTV--

AFG3_YEAST SIIPRGQCALGYAOYLPEDQYILISEQVFRHMMIALGGVSEELHFF--SVTSGAHDFFKVVQMANANVTISLGHSPKIGILSFDQNDG--NEKV
RCAL_YEAST SIIPRGQCALGYAOYLPEDQYILISEQVFRHMMIALGGVSEELHFF--SVTSGAHDFFKVVQMANANVTISLGHSPKIGILSFDQNDG--NEKV
PARAPLEGIN SITPRTNAAALGFACOMLRDQHLFTKQLFERMCMALGGVSEELHFF--SVTSGAHDFFKVVQMANANVTISLGHSPKIGILSFDQNDG--NEKV
YME1_YEAST TILPRGRALGITQLPEMFKVDITKRECOARLDVCHGKIAEELIYKDNVTSGCGSDLSQATGTARAMVTQYCSMDVGVVNLSENW
SPAST_HUMAN -----

AFG3_YEAST NKFPNSKARTIDLEKVSIVDDAHRACTELTKLNLQKVDLVAKELLRKEAITREDMIRLLGFRF-FKEERNAFEKYLD-----PKSNTEP
RCAL_YEAST TKPFSDETGDILDSVYRIVQECRDPCTKLLKKAEDVERIAQVLLKKEVLTREDMIDLLGKRF-FERNDAFDKYLNDYETERIKKEEKNKR
PARAPLEGIN RRPFSQGLQDMDBEABLLVAKAYRETEKVLQDNLQKALANALKEVINYEDIALGFPFBGKFKHIAFQWIDAQF-SQBSLSEETST
YME1_YEAST -ESWSNKIRDIADNEVIELLKDSEERARRLLTKKRVELHRLAQQGLIYETLDAHEIQVCKEK-EDLKTSTNVTVVGFD--S-DEKKGIDDK
SPAST_HUMAN -----

AFG3_YEAST PDAPAATH-----
RCAL_YEAST NPKFSTN-----
PARAPLEGIN QQPPLGGEETWFK
YME1_YEAST FKIRTMLNA-----
SPAST_HUMAN -----

```

FIGURE 4B

【図5】

```

Human: 1 [...] 459
Mouse: 1 AGGCCGAGAGCGCTCCGCGCTCTCCACAAGCAGGCCCTCGAGTACATCTCCATTGCCCTGC 60
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Human: 460 AGCCGAGCGCGCTCCGAGTCTTCCACAACAGGCCCTCGAGTACATCTCCATTGCCCTGC 519
Mouse: 61 GCATCGAGGAGGAGAGAAAAGCAGGACAGAGGAACAAGCTGTGGAATGGTATAAGAAAG 120
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Human: 520 GCATCGATGAGSATGAGAAAAGCAGGACAGAAAGGAGCAAGCTGTGGAATGGTATAAGAAAG 579
Mouse: 121 GTATCGAAGAACTGGAAAAAGGAATCGCTGTTATAGTACCGGCCAAGGTGAACAGTATG 180
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Human: 580 GTATGGAAGAACTGGAAAAAGGAATAGCTGTTATAGTTACAGGACAAGGTGAACAGTGTG 639
Mouse: 181 AAAGAGCTAGACGCTTCAACCCAAAATGATGACTAATTTAGTTATGGCCAAGGACCGCT 240
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Human: 640 AAAGAGCTAGACGCTTCAAGCTAAAATGATGACTAATTTGGTTATGGCCAAGGACCGCT 699
Mouse: 241 TACAACCTCTAGAGAACTGCAACCAAGTTTGCAATTTCCAAGTACAGACGAGCGCTCT 300
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Human: 700 TACAACCTCTAGAGAACTGCAACCAAGTTTGCAATTTCCAAGTACAGACGAGCGCTCT 759
Mouse: 301 ATACCGAGAGTACTAACCTGACATGCGCAATGGACATCTCCAGTCAGAAAGTGGAGGAG 360
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Human: 760 ATAAATGACAGTACTAACTTGGCATGCCCAATGGACATCTCCAGTCAGAAAGTGGAGCTG 819
Mouse: 361 TTCCGAAAGAGAAAAGACCCCTTAAACACATGCTAGTAATTCATGTCCCTCGATCAAAAAC 420
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Human: 820 TTCCAAAAGAAAAGACCCCTTAAACACACTAGTAATTCAGTCCCTCGTCAAAAACAG 879
Mouse: 421 TCCTGAAAAGTGGCTCCGAGGGCTCTCCGGTCAACACAGGGCGCTAGTTGCAGTGTG 480
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Human: 880 TTATGAAAACATGGATCTCGAGGCCCTTCAGGCCACCATAGAGCACTAGTTACAGTGGTT 939
Mouse: 481 TGTCCATGGTTTCTGGAGCAAGACCGGGACCTGGTCTCGAGCTACCACACATAAGGGTA 540
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Human: 940 TATCCATGGTTTCTGGAGTGAACAGGGATCTGGTCCGCTCCTACCACATAAGGGTA 999
Mouse: 541 CTCCAAAACCAATAGAACCAACAACTTCTACTCCCAACTGCAGTTCGGAAGAAAGA 600
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Human: 1000 CTCCAAAACCAATAGAACCAAACTTCTACTCCCAACTGCAGTTCGGAAGAAAGA 1059
Mouse: 601 AAGACTTGAAAATTTTAGGAATGGGACAGCAATCTTGCTAACCTTATAATGAATGAAA 660
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Human: 1060 AAGACTTGAAAATTTTAGGAATGGGACAGCAACTTGCTAACCTTATAATGAATGAAA 1119
Mouse: 661 TTGTTGACAAATGGGACAGCTGTAAAGTTGATGACATAGCCGGGACAGGAGCTGGCAAG 720
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Human: 1120 TTGTTGACAAATGGGACAGCTGTAAAGTTGATGATAGCTGCTCAAGACTTGGCAAAAC 1179
Mouse: 721 AAGCGCTGCAGGAGATTGTCATCTTCTCTCGCGCCCTGAGTTGTTCCACAGGGCTCA 780
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Human: 1180 AAGCATTGCAAGAAATTTGTTATCTTCTCTCTGAGGCCCTGAGTTGTTCCACAGGGCTTA 1239
Mouse: 781 GAGCTCCTGCTAGAGGCTTGTACTCTTCCGTCGCGCAGGAAACGGAAAAACATGCTGG 840
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Human: 1240 GAGCTCCTGCGCAGAGGCTGTACTCTTGGTCCACCTGGGSAATGGGAAGACAATGCTGG 1299
Mouse: 841 CTAAAGCAGTAGCTGCAGAGTCTAATGGGACCTTTTCAACATAAGTGCTCCAGTTTAA 900
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Human: 1300 CTAAAGCAGTAGCTGCAGAAATCGAAATCGCAACTTCTTAATATAAGTGCTGCAAGTTTAA 1359
Mouse: 901 CTTCAAAATATGTGGGAGAAGGAGAGAAATGGTGAGAGCTCTCTTGTCTGGTCCGAG 960
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Human: 1360 CTTCAAAATACGTGGGAGAAGGAGAGAAATGGTGAGAGCTCTTGTGTGGTGGCTCGAG 1419
Mouse: 961 AACTTCAACCATCTATAATTTTATAGATGAAGTTGACAGCTCTTTTGTGTGAGAGACGGG 1020
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Human: 1420 AACTTCAACCTCTATAATTTTATAGATGAAGTTGATAGCCTTTGTGTGAAAGAAAG 1479
Mouse: 1021 AAGGGGAGCAGCAGCTAGCAGACGGCTAAAGACCGAATTTTAAATAGAATTTGACGGGG 1080
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Human: 1480 AAGGGGAGCAGCAGCTAGTACGCGCTAAAACCTGAATTTCTAATAGAATTTGATGGTG 1539
Mouse: 1081 TGCAATCTGCTGGAGATGACAGAGTACTTSTAATGGGTGCAACTAACAGGCCCAAGAGC 1140
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Human: 1540 TACAGTCTGCTGGAGATGACAGAGTACTTGAATGGTGCACATTAATAGGCCAAGAGC 1599
Mouse: 1141 TTGATGAAGCTGTTCTCAGGGCTTTCATTAACGGGATATAGTGTCTTACCAATGAGG 1200
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Human: 1600 TTGATGAGGCTGTTCTCAGGGCTTTCATCAAACGGGATATAGTGTCTTACCAATGAGG 1659
Mouse: 1201 AGACAAGACTCCTTCTGCTTAAAACCTGTTGTGTAACAAGGAAGTCCACTGAOCCAAA 1260
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Human: 1660 AGACAAGACTCCTTCTGCTTAAAATCTGTTATGTAACAAGGAAGTCCATTGACCCAAA 1719
Mouse: 1261 AAGAACTCGCACAGCTTGCTAGAATGACCGATGGTACTCTGGAAGTGTCTGACCGCTT 1320
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Human: 1720 AAGAACTCGCACAGCTTGCTAGAATGACTGATGGATACTCAGGAAGTGAACCTAACAGCTT 1779
Mouse: 1321 TGGCCAAGGATGCAAGCCCTGGGTCTTATCCGAGAACTGAAGCCAGAGCAGGTGAAGAATA 1380
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Human: 1780 TGGCCAAGGATGCAAGCACTGGGTCTTATCCGAGAACTAAAACAGAAAGGTGAAGAATA 1839
Mouse: 1381 TGTCTGCCAGTGAATGAGAAATATTCGATTATCTGACTTCACAGAACTTAAAARGA 1440
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Human: 1840 TGTCTGCCAGTGAATGAGAAATATTCGATTATCTGACTTCACAGAACTTAAAARGA 1899
Mouse: 1441 TAAAGCCAGTGTGAGTCTCAGACCTTAGAAGCATACATACGCTGGAAACAAGGATTTTG 1500
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Human: 1900 TAAAGCCAGTGTGAGTCTCAGACCTTAAACTTTAGAAGCGTACATACGTTGGAAACAAGCTTTG 1959
Mouse: 1501 GAGACACCAGTGTAAAGGAAT 1523
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Human: 1960 GAGATACCAGTGTAAAGGAAT 1982
Human: 1983 [...] 3263
Mouse: 1524 GGATGCCCTCTGTAGCCCATAGAACAATCGCACTTCACAGGAAACAAGAGCTTTGGCTACA 1583
1584 GGAACCCAGACTCGTTTACAGGACGTTTATAGATTTTCAATTTTGTGCACCAACTTGA 1643
1644 AGAGGAACAAGAGACAGACCTAAATAAAATATGCAATATGAATGG 1689

```

FIGURE 5


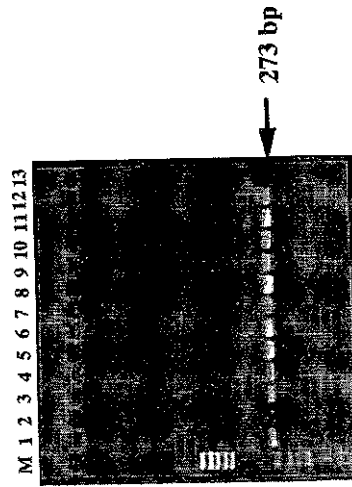
【 6 A】

FIGURE 6A




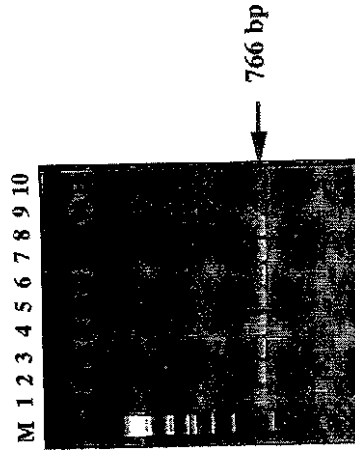
【 6 B】

FIGURE 6B




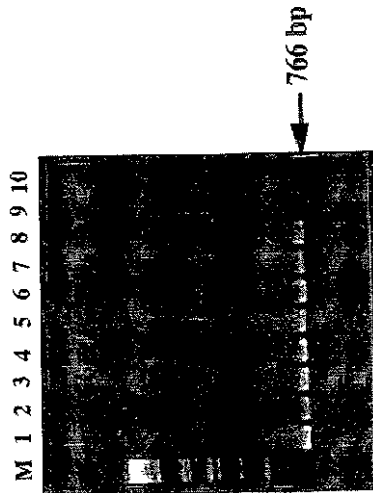
【 6 C】

FIGURE 6C



【国際調査報告】

INTERNATIONAL SEARCH REPORT		International Application No. PCT/FR 00/02433
A. CLASSIFICATION OF SUBJECT MATTER IPC 7 C12N15/12 C12N5/10 C07K14/47 C07K16/18 C12Q1/68 G01N33/50 A01K67/027		
According to International Patent Classification (IPC) or to both national classification and IPC		
B. FIELDS SEARCHED Minimum documentation searched (classification system followed by classification symbols) IPC 7 C12N C07K A01K G01N C12Q Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched Electronic data base consulted during the International search (name of data base and, where practical, search terms used) EMBL, STRAND, BIOSIS		
C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	KIKUNO R. ET AL.: "Prediction of the coding sequences of unidentified human genes. XIV. The complete sequences of 100 cDNA clones from brain which code for large proteins in vitro." DNA RES., vol. 6, 30 June 1999 (1999-06-30), pages 197-205, XP000852618 the whole document -& DATABASE EMBL 'Online! Accession No. AB029006, 4 August 1999 (1999-08-04) OHARA O.: "Homo sapiens mRNA for KIAA1083 protein" XP002140994 the whole document	1,4,5, 14,16, 18-21, 24-30
<input checked="" type="checkbox"/> Further documents are listed in the continuation of box C. <input type="checkbox"/> Patent family members are listed in annex.		
* Special categories of cited documents: *A* document defining the general state of the art which is not considered to be of particular relevance *E* earlier document but published on or after the international filing date *L* document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified) *O* document referring to an oral disclosure, use, exhibition or other means *P* document published prior to the international filing date but later than the priority date claimed *T* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention *X* document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone *Y* document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art. *Z* document member of the same patent family		
Date of the actual completion of the international search		Date of mailing of the international search report
4 January 2001		25/01/2001
Name and mailing address of the ISA European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016		Authorized officer Galli, I

INTERNATIONAL SEARCH REPORT

 International Application No
 PCT/FR 00/02433

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT		
Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	DATABASE EMBL SEQUENCES 'Online! Accession No. HS1179930, 29 March 1997 (1997-03-29) MYERS R.M.: "human STS SHGC-44567" XP002156510 brin complémentaire aux nucléotides 2204-2575 de la séquence no. 2 ---	1, 4, 5, 14, 16, 18-21, 24-30
P, X	HAZAN J. ET AL.: "Spastin, a new AAA protein, is altered in the most frequent form of autosomal dominant spastic paraplegia" NAT. GENET., vol. 23, November 1999 (1999-11), pages 296-303, XP000914979 the whole document ---	1-36
A	HEINZLEF O. ET AL.: "Mapping of a complicated familial spastic paraplegia to locus SPG4 on chromosome 2p" J. MED. GENET., vol. 35, no. 2, February 1998 (1998-02), pages 89-93, XP000914971 cited in the application the whole document -----	1-36

INTERNATIONAL SEARCH REPORT

International application No. PCT/FR 00/02433
--

Continuation of Box I.2

Claims 13-15, 19-36 partly concern nucleic acid sequences promoting and/or regulating the expression of the SPG4 gene, as well as the corresponding polypeptides, without providing therein a real technical characterisation. Said claims concern also nucleic acids and polypeptides derived from any other gene controlling SPG4, which is neither defined nor characterised in the description. Consequently, said claims are ambiguous and vague, and their subject matter is not sufficiently disclosed and supported in accordance with PCT Article 5 and 6. Consequently, it not possible to carry out any meaningful search for speculative claims whereof the subject matter consists merely of a list of results to be obtained in future.

The applicant's attention is drawn to the fact that claims, or parts of claims, in respect of which no search report has been established need not be the subject of a preliminary examination report (PCT Rule 66.1(e)). The applicant is advised that the line of conduct adopted by the EPO acting in its capacity as International Preliminary Examination Authority is not to proceed with an examination of a subject matter in respect of which no search has been carried out. Such an attitude will remain unchanged, notwithstanding whether the claims have been modified or not, either after the report has been received or during any procedure under Chapter II.

フロントページの続き

(51)Int.Cl. ⁷	識別記号	F I	テ-マ-コ-ト' (参考)	
C 1 2 N	1/19	C 1 2 N	1/21	4 H 0 4 5
	1/21	C 1 2 P	21/02	C
	5/10		21/08	
C 1 2 P	21/02	C 1 2 Q	1/02	A
	21/08		1/68	Z
C 1 2 Q	1/02	G 0 1 N	33/15	Z
	1/68		33/50	D
G 0 1 N	33/15		33/53	Z N A A
	33/50	C 1 2 N	15/00	A
	33/53		5/00	
F タ-ム(参考)	2G045 AA40 DA36 FB03			
	4B024 AA01 AA11 BA31 BA41 CA03			
	CA04 DA02 EA04 HA12			
	4B063 QA07 QA17 QA19 QQ43 QR32			
	QR55 QR62 QS25 QS34			
	4B064 AG01 AG27 CA10 CA19 CC24			
	DA13			
	4B065 AA26X AA90X AA99Y AB01			
	BA02 CA24 CA25 CA46			
	4H045 AA10 AA11 BA10 CA40 DA75			
	DA86 EA27 EA50 FA74			

专利名称(译)	克隆，表达和鉴定SPG基因，可以对最常见形式的常染色体显性痉挛性截瘫做出反应		
公开(公告)号	JP2003508074A	公开(公告)日	2003-03-04
申请号	JP2001521734	申请日	2000-09-04
[标]申请(专利权)人(译)	山藤露娜那边ドラルシエルシュシアンティフィクセーエヌエルエス肖醒		
申请(专利权)人(译)	中心Nashonaru，德，拉，RECHERCHE，Shiantifiku，(NTT节省LSI)		
[标]发明人	ジャンバイセンバツハ ジャミレアザン		
发明人	ジャン、バイセンバツハ ジャミレ、アザン		
IPC分类号	A01K67/027 C07K14/47 C07K16/18 C12N1/15 C12N1/19 C12N1/21 C12N5/10 C12N15/09 C12N15/12 C12P21/02 C12P21/08 C12Q1/02 C12Q1/68 G01N33/15 G01N33/50 G01N33/53		
CPC分类号	A01K2217/05 A61P21/00 C07K14/47		
FI分类号	A01K67/027 C07K14/47 C07K16/18 C12N1/15 C12N1/19 C12N1/21 C12P21/02.C C12P21/08 C12Q1/02 C12Q1/68.A G01N33/15.Z G01N33/50.Z G01N33/53.D C12N15/00.ZNA.A C12N5/00.A		
F-TERM分类号	2G045/AA40 2G045/DA36 2G045/FB03 4B024/AA01 4B024/AA11 4B024/BA31 4B024/BA41 4B024/CA03 4B024/CA04 4B024/DA02 4B024/EA04 4B024/HA12 4B063/QA07 4B063/QA17 4B063/QA19 4B063/QQ43 4B063/QR32 4B063/QR55 4B063/QR62 4B063/QS25 4B063/QS34 4B064/AG01 4B064/AG27 4B064/CA10 4B064/CA19 4B064/CC24 4B064/DA13 4B065/AA26X 4B065/AA90X 4B065/AA99Y 4B065/AB01 4B065/BA02 4B065/CA24 4B065/CA25 4B065/CA46 4H045/AA10 4H045/AA11 4H045/BA10 4H045/CA40 4H045/DA75 4H045/DA86 4H045/EA27 4H045/EA50 4H045/FA74		
优先权	1999011097 1999-09-03 FR		
其他公开文献	JP2003508074A5		
外部链接	Espacenet		

摘要(译)

本发明提供了鉴定和表征编码spastin的SPG基因及其一些突变，这些突变引起常染色体显性家族性痉挛性截瘫的最常见形式，其cDNA和相应多肽的克隆。并且关于表征。本发明还涉及载体，转化的细胞和转基因动物，以及诊断方法和试剂盒，以及用于选择可以直接或间接与多肽相互作用的化学或生化化合物的方法。

プライマー	配列(5'→3')	5'位のペア/PCR産物の大きさ
SPA_5RACE5	CGGAGCTCCTCTGGCTGCCATG (配列番号4)	nt 405
SPA_5RACE6	AGAAGCGCTGGCAGAGCCACGAAG (配列番号5)	nt 372
SPA_5RACE7	AAGGCGACCAAAACGAGCAGCGCGAAG (配列番号6)	nt 331
SPA_3RACE1	AGGAGCAAGCTGTGGAATGGTATAAG (配列番号7)	nt 550
SPA_3RACE2	TGGTTATGGCCAAGGACCGCTTACAAC (配列番号8)	nt 689
SPA_3RACE3	CAAACGGACGCTATAATGACAGTAC (配列番号9)	nt 747
SPA_3RACE4	TTAGGAATGTGGACAGCAACCTTGC (配列番号10)	nt 1075
SPA_3RACE5	CTTCTCTGAGGCCTGAGTTGTTAC (配列番号11)	nt 1207
SPA_3RACE6	TGCTAGAATGACTGATGGATACTCAGG (配列番号12)	nt 1736
SPA_3RACE7	AGATGCAGCACTGGGTCCTATCCG (配列番号13)	nt 1787
SPA_3RACE8	ATGAACGTCATCGGCTACAGAAACAG (配列番号14)	nt 2037