

(19) 日本国特許庁(JP)

(12) 公表特許公報(A)

(11) 特許出願公表番号

特表2004-500822  
(P2004-500822A)

(43) 公表日 平成16年1月15日(2004.1.15)

(51) Int. Cl. <sup>7</sup>	F I	テーマコード (参考)
C 1 2 N 15/09	C 1 2 N 15/00 Z N A A	2 G O 4 5
A 6 1 K 38/00	A 6 1 P 35/00	4 B O 2 4
A 6 1 P 35/00	A 6 1 P 37/02	4 B O 6 3
A 6 1 P 37/02	A 6 1 P 43/00 1 O 5	4 B O 6 4
A 6 1 P 43/00	C O 7 K 14/47	4 B O 6 5
	審査請求 未請求 予備審査請求 有 (全 132 頁) 最終頁に続く	

(21) 出願番号	特願2001-560380 (P2001-560380)	(71) 出願人	301005050 インサイト・ゲノミックス・インコーポレイテッド アメリカ合衆国カリフォルニア州94304・パロアルト・ポータードライブ 3160
(86) (22) 出願日	平成13年2月15日 (2001.2.15)	(74) 代理人	100089266 弁理士 大島 陽一
(85) 翻訳文提出日	平成14年8月9日 (2002.8.9)	(72) 発明者	クリングラー、トッド・エム アメリカ合衆国カリフォルニア州94070・サンカルロス・ドーバーコート 28
(86) 国際出願番号	PCT/US2001/005075	(72) 発明者	スチュアート、エリザベス・エイ アメリカ合衆国カリフォルニア州94041・マウンテンビュー・ラザムストリート 1678
(87) 国際公開番号	W02001/061010		
(87) 国際公開日	平成13年8月23日 (2001.8.23)		
(31) 優先権主張番号	09/507,765		
(32) 優先日	平成12年2月18日 (2000.2.18)		
(33) 優先権主張国	米国 (US)		最終頁に続く

(54) 【発明の名称】 GTPアーゼ活性化タンパク質

(57) 【要約】

本発明は、哺乳動物核酸分子及びその断片を提供する。  
本発明はまた、この哺乳動物核酸分子を用いた、情報伝達、免疫系、及び細胞増殖、特に結腸癌に関連する症状、疾患、及び障害の特徴付け・診断・評価・治療・予防方法を提供する。更に、本発明は、この哺乳動物核酸分子によってコードされるタンパク質を産生させるための発現ベクター及び宿主細胞を提供する。

```

5' 10 19 28 37 46 55
   CAG ACC CCG CAC GCA GGT GGG GGC GCG CCG GGT CCG TGG CCA GAG CTG CAG AGA
   64 73 82 91 100 109
   GAC AAG GCG GCG GCG GGT GAT CTG CTG GGT GCA CTG AGG AAG AGG CCC TGG GTG
   118 127 136 145 154 163
   GTG CCC ATG GCT GGC CAG GAT CCT GCG CTG AGC CCG AGT CAC CCG TTC TAC GAC
   M A G Q D P A L S T S H E F Y E
   172 181 190 199 208 217
   GTG GGC AGA CAT GGC AAT CTG CAG GTG GCA GAG GAT GAC CCG TTT GGA AGA CGT
   V A R H G I L Q V A S D D R F G R R
   226 235 244 253 262 271
   GTT GTC ACG TTC AGC TGC TTC CCG ATG CCA CCG TTC CAC GAG CTG GAC CAC CAG
   V V T F S C C R K P P S H E L D H Q
   280 289 298 307 316 325
   CCG CTG CTG GAG TAT TTG AAG TAC ACA CTG GAC CAA TAC GTT GAG AAC GAT TAT
   R L L E Y L K Y T L D Q Y V E N D Y
  
```

## 【特許請求の範囲】

## 【請求項 1】

(a) SEQ ID NO: 30 のアミノ酸配列、  
 (b) SEQ ID NO: 30 のアミノ酸配列の生物学的に活性な部分、および  
 (c) SEQ ID NO: 30 のアミノ酸配列の免疫学的に活性な部分から選択される GTPアーゼ活性化タンパク質をコードする実質的に精製された哺乳動物核酸分子およびその断片、並びにその変異体およびその一部。

## 【請求項 2】

SEQ ID NO: 31 のアミノ酸配列をコードする請求項 1 に記載の哺乳動物核酸分子の変異体。

10

## 【請求項 3】

(a) SEQ ID NO: 1 - 29、  
 (b) 前記 (a) の断片、および  
 (c) 前記 (a) または前記 (b) の相補体から選択され、単離され精製された哺乳動物核酸分子。

## 【請求項 4】

請求項 3 のポリヌクレオチドを含むプローブを含むプローブ。

## 【請求項 5】

(a) SEQ ID NO: 28 の 662 番目のヌクレオチドの T から A への置換、  
 (b) SEQ ID NO: 28 の 1013 番目のヌクレオチドの T から C への置換、および  
 (c) SEQ ID NO: 28 の 1020 番目のヌクレオチドの C から G への置換から選択された置換を含む一塩基多型 (SNP)。

20

## 【請求項 6】

請求項 3 の核酸分子に機能的に結合したプロモーターを含む組換え核酸分子。

## 【請求項 7】

請求項 6 の組換え核酸分子で形質転換された細胞。

## 【請求項 8】

ポリペプチドを生産する方法であって、

(a) 前記ポリペプチドが発現する条件下で、請求項 7 の細胞を培養するステップと、  
 (b) 発現した前記ポリペプチドを回収するステップとを含むことを特徴とするポリペプチド生産方法。

30

## 【請求項 9】

サンプルにおいて哺乳動物核酸分子を検出するための方法であって、

(a) 請求項 4 のプローブを前記サンプルとハイブリダイズさせて、ハイブリダイゼーション複合体を形成させるステップと、  
 (b) 前記ハイブリダイゼーション複合体を検出するステップとを含み、  
 前記ハイブリダイゼーション複合体の検出が、前記サンプルにおける前記哺乳動物核酸分子の存在を示すことを特徴とする哺乳動物核酸分子の検出方法。

40

## 【請求項 10】

前記ハイブリダイゼーションの前に、前記核酸分子またはその断片を増幅するステップを更に含むことを特徴とする請求項 9 に記載の方法。

## 【請求項 11】

哺乳動物核酸分子またはその断片を用いて分子のライブラリをスクリーニングし、前記哺乳動物核酸分子に特異的に結合する少なくとも 1 つのリガンドを同定する方法であって、

(a) 特異的な結合が許容される条件下で、請求項 3 の哺乳動物核酸分子を前記分子のライブラリと結合させるステップと、  
 (b) 特異的な結合を検出して、前記哺乳動物核酸分子に特異的に結合するリガンドを同定するステップとを含むことを特徴とするリガンド同定方法。

## 【請求項 12】

50

前記ライブラリが、DNA分子、RNA分子、PNA、ペプチド、およびタンパク質から選択されることを特徴とする請求項11に記載の方法。

【請求項13】

哺乳動物核酸分子またはその断片を用いて、サンプルから前記哺乳動物核酸分子に特異的に結合するリガンドを精製する方法であって、

(a) 特異的な結合が許容される条件下で、請求項3の哺乳動物核酸分子またはその断片とサンプルとを結合させるステップと、

(b) 結合した前記哺乳動物核酸分子を回収するステップと、

(c) 前記リガンドから前記哺乳動物核酸分子を分離して、精製されたりガンドを得るステップとを含むことを特徴とするリガンド精製方法。

10

【請求項14】

(a) SEQ ID NO: 30のアミノ酸配列、

(b) SEQ ID NO: 30のアミノ酸配列の生物学的に活性な部分、および

(c) SEQ ID NO: 30のアミノ酸配列の免疫学的に活性な部分から選択されることを特徴とする実質的に精製されたタンパク質またはその一部。

【請求項15】

SEQ ID NO: 31のアミノ酸配列を含む請求項14のタンパク質の変異体。

【請求項16】

タンパク質またはその一部を用いて分子のライブラリをスクリーニングし、前記タンパク質に特異的に結合する少なくとも1つのリガンドを同定する方法であって、

(a) 特異的な結合が許容される条件下で、請求項14のタンパク質またはその一部を前記分子のライブラリと結合させるステップと、

(b) 特異的な結合を検出して、前記タンパク質に特異的に結合するリガンドを同定するステップとを含むことを特徴とするリガンド同定方法。

20

【請求項17】

前記ライブラリが、DNA分子、RNA分子、PNA、ペプチド、タンパク質、アゴニスト、アンタゴニスト、抗体、免疫グロブリン、インヒビター、薬剤化合物、および医薬品から選択されることを特徴とする請求項16に記載の方法。

【請求項18】

タンパク質またはその一部を用いて、サンプルから前記タンパク質に特異的に結合するリガンドを精製する方法であって、

(a) 特異的な結合が許容される条件下で、請求項14のタンパク質またはその一部をサンプルと結合させるステップと、

(b) 結合した前記タンパク質を回収するステップと、

(c) 前記リガンドから前記タンパク質を分離して、精製したりガンドを得るステップとを含むことを特徴とするリガンド精製方法。

30

【請求項19】

効果的な量の請求項14のタンパク質および医薬的に許容される賦形剤を含む医薬組成物。

【請求項20】

GTPAPの発現の変化に関連する疾患や症状を治療または防止する方法であって、そのような治療が必要な患者に請求項19の医薬組成物を投与することを特徴とする治療方法。

40

【請求項21】

前記疾患や症状が、シグナル伝達異常、免疫疾患、および癌を含む細胞増殖異常から選択されることを特徴とする請求項20に記載の方法。

【請求項22】

前記癌が結腸癌であることを特徴とする請求項21に記載の方法。

【発明の詳細な説明】

【0001】

50

(発明の技術分野)

本発明は、核酸分子およびその核酸分子にコードされたGTPアーゼ活性化タンパク質、およびその変異体に関連し、また、これらの分子を用いたシグナル伝達異常、免疫疾患、および特に結腸癌である細胞増殖異常の特徴付け・診断・予防・治療に関連する。

【0002】

(発明の背景)

グアニンヌクレオチド結合タンパク質(GTP結合タンパク質)は全ての真核細胞に存在し、代謝、細胞成長、分化、シグナル伝達、細胞骨格形成、および細胞内の小胞輸送や分泌などのプロセスにおいて機能する。高等動物では、GTP結合タンパク質は、免疫反応(Ausseilら(1988) J Immunol 140: 215-220)、アポトーシス、分化、および発癌を含む細胞増殖(Dhanasekaranら(1998) Oncogene 17: 1383-1394)などのプロセスを調節するシグナル伝達に参与している。

10

【0003】

GTP結合タンパク質のスーパーファミリーは幾つかのファミリーから構成されており、翻訳因子と、膜貫通シグナル伝達プロセスに関係するヘテロ三量体GTP結合タンパク質(Gタンパク質とも呼ぶ)と、発癌遺伝子Rasタンパク質と、rab、rap、rho、rac、smg21、smg25、YPT、SEC4、およびARFなどの遺伝子の産物を含むその他の低分子量GTP結合タンパク質と、チューブリンとに分類することができる(Kaziroら(1991) Ann Rev Biochem 60: 349-400)。

20

【0004】

低分子量(LMW)GTP結合タンパク質は、21~30 kDaの小さなタンパク質のクラスである。これらのタンパク質は、細胞成長、細胞周期、タンパク質分泌、および細胞内の小胞体相互作用を調節する。特にLMW GTP結合タンパク質は、受容体からの細胞外シグナルに应答して、様々な細胞機能に関係する分裂促進シグナルを伝達し、細胞タンパク質を活性化させる(Tavitian(1995) C R Seances Soc Biol Fil 189: 7-12)。このプロセスの際、GTPの加水分解により、GTP結合タンパク質が活性型(GTP結合)から不活性型(GDP結合)へ変換するだけでなく、エネルギーも供給される。

30

【0005】

LMW GTP結合タンパク質のRhoファミリー(Rho GTPアーゼ)には、アクチン細胞骨格再形成および細胞増殖に関連する様々な生体現象を調節するRhoA、Rac1、およびCdc42が含まれる。Rho GTPアーゼの主な調節分子群には、Rho GTPアーゼ活性化タンパク質(GAP)がある。Rho GAPは、Rho GTPアーゼのGTP結合型を選択的に認識し、内在性GTPアーゼ活性を刺激して結合GTPを加水分解してGDPにする。従って、Rho GAPは、Rho GTPアーゼの活性なGTP結合型から不活性なGDP結合型への変換を刺激して、Rho GTPアーゼの負の調節因子すなわちリプレッサーとして機能する(Zhangら(1997) J Biol Chem. 272: 21999-22007)。

40

【0006】

Rho GAPタンパク質は、Rho GAPドメインと呼ばれる約170~190のアミノ酸相同領域を共有している。この領域は、GAP活性に必要な最小の構造ドメインを含むと思われる(Zhengら(1993) J Biol Chem 24629-24634)。Rho GAPタンパク質は、このドメインにおけるアミノ酸同一性が20%~24%であるが、ある特定の残基が高度に保存されている。例えば、Rho GAPドメインのN末端に近接した一対のアルギニン残基は、Rho GAPタンパク質の間に高度に保存され、最大触媒作用に必要であると思われる(Leonardら(1998) J Biol Chem 273: 16210-16215)。更に、プロリンリッチのSH3ドメイン結合部位は、これらのタンパク質のN末端領域にも見られる(Leona

50

rd、前出)。

【0007】

癌に関連する遺伝子の同定および発癌に関係する遺伝メカニズムを理解することは、このような疾患の診断、予防、および治療に極めて重要である。特に結腸癌の場合、発癌遺伝子の活性化、癌抑制遺伝子の不活化、およびDNAミスマッチ修復遺伝子の変異の組み合わせが、正常な粘膜から結腸癌への悪化に関係する(Fearonら(1990) Cell 61 : 759-767 ; Chung(1995) Gastroenterology 109 : 1685-1699)。近年の研究により、正常な結腸粘膜に対する結腸腫瘍サンプルにおける染色体(遺伝子?)の対立遺伝子の喪失に基づいて、結腸癌に関連する第22染色体の特定の領域が同定された(Castellsら(1999) Gastroenterology 117 : 8331-837)。対立遺伝子欠失の最小領域が、0.57cMの間隔をカバーするマーカーD22S1171とD22S928との間に存在し、遺伝子座位22q13.33に対応することが、第22染色体のマイクロサテライトマーカーを用いて見出された。第22染色体上の様々な遺伝子が、結腸直腸の発癌に関連する癌抑制遺伝子の候補と提示されたが、それらの遺伝子には如何なる突然変異も存在しなかった(Castells他、前出)。

【0008】

GTPアーゼ活性化タンパク質をコードする核酸配列およびその変異体の発見により、癌、特に結腸癌を含む細胞増殖異常の特徴付け・診断・予防・治療に有用な組成物を提供する。

【0009】

(発明の開示)

総称して「GTPAP」、個別にはそれぞれ「GTPAP-1」および「GTPAP-2」と呼ぶGTPアーゼ活性化タンパク質をコードする核酸分子の発見に基づき、シグナル伝達異常、免疫疾患、および特に結腸癌である細胞増殖異常などの症状の特徴付け・診断・予防・治療に有用な組成物を提供することで当分野の要望に応える。

【0010】

本発明は、SEQ ID NO : 30のアミノ酸配列またはその一部と、SEQ ID NO : 30の生物学的に活性な部分と、SEQ ID NO : 30の免疫学的に活性な部分と、SEQ ID NO : 31を含むSEQ ID NO : 30の変異体とを含む本哺乳動物タンパク質をコードするSEQ ID NO : 1-29またはそれらの断片を含む単離され精製されたヒトおよびラット核酸分子を提供する。

【0011】

本発明はまた、本哺乳動物核酸分子またはその断片とハイブリダイズするプローブを提供する。本発明はまた、SEQ ID NO : 1-29の核酸分子に相補的な単離され精製された核酸分子を提供する。一実施態様では、このプローブは、一本鎖の相補的なRNA分子またはDNA分子である。

【0012】

本発明は更に、サンプルにおいて核酸分子を検出する方法であって、プローブをそのサンプルの少なくとも1つの核酸分子とハイブリダイズするステップと、ハイブリダイゼーション複合体を形成するステップと、そのハイブリダイゼーション複合体を検出するステップとを含み、そのハイブリダイゼーション複合体の存在がそのサンプルに核酸分子存在することと相関する方法を提供する。一実施態様では、この方法は、ハイブリダイゼーションの前にその核酸分子を増幅するステップを更に含む。この核酸分子若しくはその断片を、マイクロアレイ上のエレメント若しくは標的として用いることができる。

【0013】

本発明はまた、核酸配分子若しくはその断片を用いて、その核酸分子と特異的に結合する少なくとも1つのリガンドを同定するべく、分子のライブラリをスクリーニングする方法であって、特異的な結合が許容される条件下で、その核酸分子をその分子のライブラリと結合させるステップと、その特異的な結合を検出して、その核酸分子と特異的に結合するリ

ガンドを同定するステップとを含むスクリーニング方法を提供する。このようなライブラリには、DNA分子及びRNA分子、ペプチド、PNA、タンパク質等が含まれる。類似の方法で、核酸分子またはその断片を用いてリガンドを精製することができる。

【0014】

本発明はまた、本核酸分子の少なくともある断片を含む発現ベクターを提供する。別の実施態様では、この発現ベクターは宿主細胞内に含まれている。本発明は更に、タンパク質が発現する条件下で宿主細胞を培養するステップと、そのタンパク質をその宿主細胞から回収するステップとを含むタンパク質生産方法を提供する。

【0015】

本発明はまた、実質的に精製され哺乳動物GTPAPまたはその一部を提供する。更に、本発明は、SEQ ID NO: 30のアミノ酸配列と、SEQ ID NO: 30の生物学的に活性な部分と、SEQ ID NO: 30の免疫学的に活性な部分と、SEQ ID NO: 31を含むSEQ ID NO: 30の変異体とを有する単離され精製されたタンパク質を提供する。本発明はまた、医薬用担体と共に実質的に精製された哺乳動物GTPAPタンパク質またはその一部を含む医薬組成物を提供する。

10

【0016】

本発明はまた、GTPAPの発現の変化に関連した疾患または症状を治療するための方法を提供する。この方法は、医薬用担体と共に実質的に精製された哺乳動物GTPAPタンパク質またはその一部を含む医薬組成物を好適な量、そのような治療が必要な患者に投与することを含む。本発明の一実施例では、そのような疾患は、細胞のシグナル伝達異常、免疫疾患、および特に結腸癌である細胞増殖異常から選択される。本発明の別の実施例では、癌は結腸癌である。

20

【0017】

本発明は更に、その哺乳動物タンパク質の一部を用いて抗体を産生させる方法を提供する。本発明はまた、哺乳動物タンパク質若しくはその一部を用いて、そのタンパク質に特異的に結合する少なくとも1つのリガンドを同定するべく、分子のライブラリをスクリーニングする方法であって、特異的な結合が許容される条件下で、そのタンパク質を分子のライブラリと結合させるステップと、その特異的な結合を検出して、そのタンパク質と特異的に結合するリガンドを同定するスクリーニング方法を提供する。このようなライブラリには、DNA分子及びRNA分子、ペプチド、アゴニスト、アンタゴニスト、抗体、免疫グロブリン、薬剤、医薬品、その他のリガンドが含まれる。一実施態様では、この方法で同定されたりガンドがその哺乳動物タンパク質の活性を調節する。類似の方法で、そのタンパク質若しくはその一部を用いてリガンドを精製する。本発明はまた、特異的な結合が許容される条件下で、その哺乳動物タンパク質またはその一部をサンプルと結合させて、そのタンパク質とリガンドとの特異的な結合を検出して、その結合したタンパク質を回収し、そのリガンドからそのタンパク質を分離して、精製されたりガンドを得る方法を提供する。

30

【0018】

本発明はまた、哺乳動物のゲノムDNAの中にマーカー遺伝子を挿入して、その天然の哺乳動物核酸分子の発現を阻害する方法を提供する。本発明はまた、その哺乳動物核酸分子を用いて哺乳動物モデル系を作製する方法であって、その哺乳動物核酸分子を含むベクターを作製するステップと、そのベクターを全能性哺乳動物胚性幹細胞の中に導入するステップと、ゲノムDNAの中にそのベクターが組み込まれた胚性幹細胞を選択するステップと、この選択した細胞を哺乳動物胚盤胞の中に微量注入して、キメラ胚盤胞を形成するステップと、偽妊娠メスにキメラ胚盤胞を移植するステップと、このメスが、生殖細胞系の中に少なくとも1つの哺乳動物核酸分子の別の複製を含むキメラ哺乳動物を出産し、そのキメラ哺乳動物を交配して同型接合哺乳動物モデル系を作製するステップとを含む哺乳動物モデル系作製方法を提供する。

40

【0019】

(本発明の実施方法について)

50

本発明は、ここに開示した特定の装置及び材料、方法に限定されず、その実施形態を変更できることを理解されたい。また、ここで用いられる用語は、特定の実施例のみを説明する目的で用いるものであり、後述の請求の範囲によってのみ限定され、本発明の範囲を限定することを意図したものではないということも理解されたい。本明細書及び請求の範囲において単数形を表す「或る」、「その(この等)」は、文脈で明確に示していない場合は複数形を含むことに注意されたい。従って、例えば「或る宿主細胞」は当業者には周知の複数の宿主細胞を含む。

**【0020】**

本明細書で用いた全ての科学技術用語は、別の方法で定義されていない限り、本発明の属する技術分野の一般的な技術者が普通に解釈する意味と同じである。本明細書に記載の全ての文献は、本発明に関連して使用する可能性のある文献に記載された細胞系、プロトコル、試薬、ベクターを記述し開示するために引用した。従来の発明を引用したからと言って、本発明の新規性が損なわれると解釈されるものではない。

10

**【0021】**

(定義)

「GTPアーゼ活性化タンパク質」または「GTPAP」は、天然、合成、半合成或いは組み換え体などの全ての種(マウス、ウシ、ヒツジ、ブタ、齧歯類、イヌ、サルを含むが好ましくはヒトである哺乳動物)から得られる、実質的に精製されたタンパク質を指す。「GTPAPs」は、2つ以上のGTPAPを指す。

**【0022】**

「遺伝子発現の変化」および「発現の変化」は、サンプルにおける遺伝子発現の増大および減少、並びに転写されたメッセンジャーRNA転写産物の存在または不在を指す。

20

**【0023】**

「生物学的に活性」は、構造的または免疫学的、調節的、化学的な機能を有する、天然或いは組み換え、または合成分子であるタンパク質を指す。同様に、「免疫学的に活性」は、天然、組換え、または合成のGTPAP、或いはその任意のオリゴペプチドが、好適な動物または細胞において特定の免疫反応を誘導して特定の抗体に結合する能力を指す。

**【0024】**

「相補的な」は、プリン塩基とピリミジン塩基との自然な水素結合による塩基対形成を指す。例えば、配列A-C-G-Tは、その相補配列T-G-C-AまたはU-G-C-Aと水素結合する。2つの一本鎖分子の相補性は、ヌクレオチドの幾つかのみが結合する部分的な相補性と、ヌクレオチドの殆ど全てが結合する完全な相補性とがある。核酸鎖間の相補性の程度は、ハイブリダイゼーション及び増幅反応の効率及び強度に影響する。

30

**【0025】**

「誘導體」は、化学修飾された核酸分子やタンパク質配列を指す。化学修飾には、分子や配列の生物学的活性または寿命を維持或いは増大する、アルキル基またはアシル基、アミノ基による水素の置換や、グリコシル化、ポリエチレングリコール化、または類似の任意のプロセスが含まれる。

**【0026】**

「断片」は、有用な機能的特性を維持する核酸分子の任意の一部またはインサイト社クロームを指す。有用な断片には、ハイブリダイゼーションや増幅技術、または複製や転写、翻訳の調節に有用なオリゴヌクレオチドを含む。

40

**【0027】**

「ハイブリダイゼーション複合体」は、プリン塩基とピリミジン塩基との間の水素結合の形成による2つの核酸分子の複合体を指す。

**【0028】**

「リガンド」は、核酸分子やタンパク質の相補部位に特異的に結合する任意の分子や物質、または化合物を指す。このようなりガンドは、核酸及びタンパク質、炭水化物、脂肪、脂質を含む無機及び有機物質の少なくとも1つからなり、本発明の核酸分子やタンパク質の活性を安定化させたり、調節したりする。

50

## 【0029】

「調節する」は、分子と核酸分子がタンパク質の何れかとの特異的な結合によって、寿命や活性（生物学的または化学的、免疫学的）を変化させることを指す。

## 【0030】

「非保存的なアミノ酸の置換」は、置換されると元のタンパク質の特性、すなわちタンパク質の構造、特にその機能に影響を与える可能性が最も高い置換を指す。例えば、酸性のアスパラギン酸や塩基性のアルギニンなどのアミノ酸の中性アミノ酸への置換は、非保存的置換である。同様に、「保存的なアミノ酸の置換」は、置換されても元のタンパク質の特性への影響が最小、すなわちタンパク質の構造、特にその機能が保存され、あまり変化しない置換を指す。バリンとアラニンとの置換などの中性のアミノ酸と中性のアミノ酸との置換は保存的な置換である。また、アスパラギン酸とグルタミン酸との置換などの酸性のアミノ酸と酸性のアミノ酸との置換も保存的な置換である。

10

## 【0031】

「非保存的な核酸の置換」は、核酸分子におけるピリミジンヌクレオチドとプリンヌクレオチドとの置換を指す。同様に、「保存された核酸の置換」は、CとTのピリミジンヌクレオチドと別のピリミジンヌクレオチドとの置換や、AとGのプリンヌクレオチドと別のプリンヌクレオチドとの置換を指す。

## 【0032】

「核酸分子」は、核酸、オリゴヌクレオチド、ヌクレオチド、ポリヌクレオチド、またはそれらの任意の断片を指す。また、「核酸分子」は、ゲノム若しくは合成起源の二本鎖若しくは一本鎖のDNA或いはRNAであり、炭水化物または脂質、タンパク質、その他の物質と結合して、形質転換などの特殊な作用を起こしたり、ペプチド核酸(PNA)等の有用な成分を形成する。「オリゴヌクレオチド」は、アンプリマー(amplicimer)及びプライマー、オリゴマー、エレメント、標的、プローブと実質的に同一であり、好ましくは一本鎖である。核酸分子は別の核酸分子のスプライスバリエーションであり得る。

20

## 【0033】

「タンパク質」は、天然或いは合成のアミノ酸配列、オリゴペプチド、ペプチド、ポリペプチド、またはそれらの一部を指す。

## 【0034】

ここで用いる「一部」は、あらゆる目的に用いられるタンパク質の任意の一部を指すが、特に、その一部と特異的に結合する分子のライブラリのスクリーニングや抗体の生産に用いられる。「一部」の典型的な例は、SEQ ID NO: 46の初めの40の連続するアミノ酸である。

30

## 【0035】

ポリヌクレオチド配列に用いる「類似性」とは、標準的なアルゴリズムを用いてアラインメントされた少なくとも2つの配列間の一致する残基数から求められる。このようなプログラムは、2つの配列間のアラインメントを最適化するべく、標準的かつ再現性のある方法で比較する配列内にギャップを挿入可能なため、2つの配列間のより有意な比較を達成することが可能である。

## 【0036】

広く用いられている無料の配列比較アルゴリズム一式が、National Center for Biotechnology Information (NCBI)、(Bethesda, MD)によって提供される。Basic Local Alignment Search Tool (BLAST) (Altschul, ら (1990) J. Mol. Biol. 215: 403-410; Altschul (1993) J. Mol. Evol. 36: 290-300)は、NCBI (<http://WWW.ncbi.nlm.nih.gov/BLAST/>) および幾つかのほかの情報源から入手可能である。このBLASTソフトウェア一式には、既知のポリヌクレオチド配列と様々なデータベースの別のポリヌクレオチド配列とのアラインメントに用いられる「blastn」を含む、様々な配列分析プログラムが含まれる。「BLAST 2 Sequen

40

50

ces」と呼ばれるツールが入手可能であり、2つのヌクレオチド配列の対を直接比較するために用いられる。「BLAST 2 Sequences」は、<http://www.ncbi.nlm.nih.gov/gorf/b12.html>にアクセスして、対話形式で利用できる。「BLAST 2 Sequences」ツールは、blastn及びblastpの両方に用いることができる。BLASTプログラムは、一般的には、デフォルトを設定するギャップ及び他のパラメーターと共に用いられる。例えば、2つのヌクレオチド配列を比較する場合、ある者は、デフォルトパラメータに設定された「BLAST 2 Sequences」ツールVersion 2.0.09 (May-07-1999)でblastnを使用するであろう。そのようなデフォルトパラメータは、例えば、以下のようにする。

10

```

【0037】
Matrix: BLOSUM62
Reward for match: 1
Penalty for mismatch: -2
Open Gap: 5 及び Extension Gap: 2 penalties
Gap x drop-off: 50
Expect: 10
Word Size: 11
Filter: on

```

類似性は、例えば、特定の配列番号で決められた、所定の配列の全長に対して測定してもよいし、それより短い長さに対して、例えば、ある大きな所定の配列から得られた断片、例えば、連続する少なくとも、20または30、40、50、70、100、200のヌクレオチドの断片の長さに対して測定してもよい。このような長さは単なる例であり、配列表及び表、図面を含む明細書に記載の配列の任意の長さの断片を用いて、同一性のパーセントが測定される長さを示すことができる。

20

【0038】  
「サンプル」は、その最も広い意味で用いられる。ポリヌクレオチド、ポリペプチド、および抗体などを含むサンプルは、体液や、細胞が成長する細胞溶液、即ち培養液の可溶性画分や、細胞から単離或いは抽出した染色体、細胞小器官、または膜や、溶液中に存在する或いは基板に結合されたゲノムDNAまたはRNA、cDNAや、組織、細胞、組織プリント、またはフィンガープリント、皮膚、髪等を含み得る。

30

【0039】  
「特異的な結合」または「特異的に結合する」は、2つの分子間の相互作用を指す。ポリヌクレオチドの場合、特異的な結合には、センス鎖とアンチセンス鎖との間の水素結合、複製または転写に影響を与える一本鎖とタンパク質との間の水素結合、DNA分子の主溝または副溝への分子または化合物の挿入、転写因子、エンハンサー、およびリプレッサー等として機能する少なくとも1つの分子との相互作用が含まれる。ポリペプチドの場合、特異的な結合には、上記したようなポリヌクレオチドとの相互作用、或いはアゴニスト、抗体、またはアンタゴニストなどの分子や化合物との相互作用が含まれる。特異的な結合は、分子間の好適な化学的相互作用または分子相互作用を許容する構造的特性の存在に左右される。

40

【0040】  
「一塩基多型」またはSNPは、置換、挿入、または欠失により生じるポリヌクレオチドにおける1つのヌクレオチドの変化を指す。置換は上記したように保存的または非保存的であり、コードされるアミノ酸残基が変化する場合もあるし、変化しない場合もある。このような変化は、個人における特定の疾患や症状に対する素因となり得る。

【0041】  
「実質的に精製された」は、その自然環境から切り離されてから分離或いは単離された、自然環境では結合しているその他の成分が少なくとも約60%、好適には約75%、最も好適には、約90%取り除かれた核酸分子またはタンパク質を指す。

50

## 【0042】

「基板」は、核酸分子またはタンパク質が結合した任意の固体或いは半固体の支持物を指し、膜またはフィルター、チップ、スライド、ウエハ、ファイバー、磁気または非磁気ビーズ、ゲル、毛細管または他のチューブ、プレート、ポリマー、微小粒子が含まれる。この基板は、孔または溝、ピン、チャンネル、細孔を含む様々な表面形態を有する。

## 【0043】

「変異体」は、GTPAPをコードする核酸分子の変異体である核酸分子、またはそのような分子によってコードされるタンパク質を指す。このような変異体は、少なくとも100、最も好ましくは400以上のBLASTスコアから決定され得る。

## 【0044】

(発明)

本発明は、GTPAPsをコードする新規の哺乳動物核酸分子およびその核酸分子変異体の発見に基づき、シグナル伝達異常、免疫疾患、特に結腸癌である細胞増殖異常などの病態の特徴付け・診断・治療・予防においてその核酸分子やその断片、またはそのタンパク質やその一部を用いることに関連する。

## 【0045】

本発明のGTPAP1をコードする核酸は、アミノ酸配列アライメント用のコンピュータ検索を用いて子宮組織ライブラリ(UTRSNOR01)のインサイトクローン3068538H1において初めに同定された。コンセンサス配列SEQ ID NO: 28は、インサイトクローン908465R2; SEQ ID NO: 1 (COLNNOT09)、957130R6; SEQ ID NO: 2 (KIDNNOT05)、1580628H1; SEQ ID NO: 3 (DUODNOT01)、2631247F6; SEQ ID NO: 4 (COLNTUT15)、3068538H1; SEQ ID NO: 5 (UTRSNOR01)、3532286T6; SEQ ID NO: 6 (KIDNNOT25)、および1301520F6; SEQ ID NO: 7に由来する。

## 【0046】

本発明のヒトGTPアーゼ活性化タンパク質をコードする核酸分子、および核酸分子変異体は、本発明のGTPAP1およびGTPAP2をコードする核酸配列をLIFESEQデータベースおよびZOOSEQデータベース(Incyte Genomics, Palo Alto, CA)における核酸配列と比較して初めに同定された。

## 【0047】

一実施例では、本発明は、図1A - 図1Eに示されているSEQ ID NO: 30のアミノ酸配列を含むポリペプチドを含む。GTPAP1はアミノ酸433個の長さであって、N338残基にける潜在的なN結合グリコシル化部位を有する。潜在的なカゼインキナーゼIIリン酸化部位が、S169、T239、T292、S309、およびS382に見られ、潜在的なプロテインキナーゼCリン酸化部位が、S129、T239、およびS297に見られ、潜在的なチロシンキナーゼリン酸化部位が、Y60、Y101、およびY315に見られる。BLIMPS\_\_PFAM分析により、D260からP276のアミノ酸残基におけるRHG5 GTPアーゼ活性化タンパク質シグネチャが同定された。BLIMPS\_\_PRODOMにより、P210からA235及びL310からL350における2つのタンパク質GTPアーゼ活性化ドメインが同定された。BLAST\_\_PRODOM分析により、L8からS169における確率スコア $p = 4.3 \cdot 10^{-6}$ のRHG5 RhoGAPタンパク質に対する相同性と、T175からP387における確率スコア $p = 7.9 \cdot 10^{-29}$ のP50-RhoGAPに対する相同性とが同定された。

## 【0048】

別の実施例では、本発明は、図2A - 図2Eに示されているSEQ ID NO: 31のアミノ酸配列を含むポリペプチドを含む。GTPAP2はまた、アミノ酸433個の長さであって、GTPAP1と同じグリコシル化部位、リン酸化部位、モチーフ、およびシグネチャ配列を共有する。

## 【0049】

10

20

30

40

50

図3A - 図3Dに示されているように、GTPAP1は、GTPAP2 (インサイトID 404424.5; SEQ ID NO: 31)、ヒトRhoGAPタンパク質8 (GI 6572185; SEQ ID NO: 32) のアミノ酸配列の一部、およびヒトRhoGAPタンパク質GI 312212 (SEQ ID NO: 33) と配列相同性を有する。具体的には、GTPAP1とGTPAP2との同一性は99.8%であって、異なる残基は唯1つである。すなわち、GTPAP1の302位におけるグリシン残基がGTPAP2ではアルギニン残基に置換されている。この置換は、その位置のアミノ酸のコドンにおけるSNPにより生じている(図4Aおよび図4Bのポリヌクレオチドアラインメントを参照)。GTPAP1は、そのY101からL433においてGI 6572185と同一であるが、GTPAP1は、そのタンパク質のアミノ末端に更に100個のアミノ酸残基を含む。GTPAP1はまた、GI 312212と49.2%の同一性を有する。4つのタンパク質はそれぞれ、GTPAP1の約P176からP189アミノ酸残基までのプロリンリッチ領域であるSH3ドメイン結合部位を有する。4つのタンパク質はまた、RhoGAPドメインを含むC末端領域に実質的ににおいて相同である。具体的には、GTPAP1の配列F231RRSに含まれているGAPタンパク質の触媒活性に関連する2つのアルギニン残基は、4つ全てのタンパク質に保存されている。更に、GTPAP1の302位におけるグリシン残基がGTPAP2ではアルギニン残基に置換されているが、この置換が非保存的な置換であってこのタンパク質のRhoGAPドメインで生じていることに注目されたい。

10

#### 【0050】

20

図4Aおよび図4Bには、GTPAP1 (3068538; SEQ ID NO: 28) をコードするポリヌクレオチド、GTPAP2 (404424.5; SEQ ID NO: 29) をコードするポリヌクレオチド、およびこれらの配列が由来するインサイトcDNAクローンとの間のアラインメントが示されている。具体的には、このアラインメントには、2つの完全長のポリヌクレオチド間に見られる一塩基多型(SNP)の領域が含まれる。図4Aおよび図4Bでは、これらのSNPは で示した。これらの領域を含むcDNAクローンおよび対応ライブラリは、908465R2; SEQ ID NO: 1 (COLNNOT09)、957130R6; SEQ ID NO: 2 (KIDNNOT05)、1580628H1; SEQ ID NO: 3 (DUODNOT01)、2631247F6; SEQ ID NO: 4 (COLNTUT15)、1301520F6; SEQ ID NO: 7 (BRSTNOT07)、2465422F6; SEQ ID NO: 8 (THYRNOT08)、957130X313V1; SEQ ID NO: 9 (KIDNNOT05)、1580545H1; SEQ ID NO: 10 (DUODNOT01)、1891457H1; SEQ ID NO: 11 (BLADTUT07)、4649657H1; SEQ ID NO: 12 (PROSTUT04)、4002758H1; SEQ ID NO: 13 (HNT2AZS07)、957130R1; SEQ ID NO: 14 (KIDNNOT05)、2624365T6.comp; SEQ ID NO: 15 (KERANOT02)、2044444H1; SEQ ID NO: 16 (HIPONON02)、3416883T6; SEQ ID NO: 17 (PTHYNOT04)、1301520T6.comp; SEQ ID NO: 18 (BRSTNOT07)、1301520H1; SEQ ID NO: 19 (BRSTNOT07)、および1422635X305D1.comp; SEQ ID NO: 20 (KIDNNOT09) である。図4Aには、GTPAP1における662位のヌクレオチドTがGTPAP2ではAに非保存的に置換されているのが示されている。この変化は、6つのクローンの内、2つがその位置にTを含み、4つがその位置にAを含むことにより支持されている。図4Bには、1013位および1020位におけるSNPが示されている。1013位におけるSNPは、T(GTPAP1)からC(GTPAP2)への保存的な置換であり、12のクローンの内、2つがその位置にCを含み、10がその位置にTを含むことにより支持されている。1020位におけるSNPは、G(GTPAP1)からC(GTPAP2)への非保存的な置換であり、12のクローンの内、2つがその位置にCを含み、7つがそ

30

40

50

の位置にGを含むことにより支持されている。この領域を含む識別可能なクローンの2 / 12すなわち17%にのみ見られることから比較的稀であり、かつコードされるアミノ酸残基のコドンの5'位置、すなわち1020位におけるSNPにより、GTPAP2においてグリシンからアルギニンへの非保存的な置換が生じることから、1013位および1020位におけるSNPは重要である。1020位における変化は、タンパク質のこの領域に位置するコードされるタンパク質の触媒活性に影響を与え得る。

#### 【0051】

BLASTまたはBLAST2 (Altschulら (1997) *Nucleic Acids Res.* 25 : 3389 - 3402) を用いて、ZOOSEQデータベース (Incyte Pharmaceuticals, Palo Alto CA) において、ヒトRho GAPタンパク質をコードする核酸分子 (SEQ ID NO: 28およびSEQ ID NO: 29) に一致するラット、マウス、およびサル10のクローンを検索して、本発明のヒトGAPタンパク質をコードする核酸分子が更に同定された。図5A - 図5Eに、SEQ ID NO: 28とラット由来のこの配列の非ヒト変異配列との間のアライメントが示されている。ヒト配列 (SEQ ID NO: 28) に対してBLASTスコアが100以上である、701244926H1; SEQ ID NO: 21 (RALINOH02)、700950169H2; SEQ ID NO: 22 (RASPNON02)、701575974H1; SEQ ID NO: 23 (RALITXT25)、701274036H1; SEQ ID NO: 24 (RABFNON02)、700480528H1; SEQ ID NO: 25 (RALINON03)、700935753H201; SEQ ID NO: 26 (RALINON03)、および700936061H1; SEQ ID NO: 27 (RALINON03) が同定された。

#### 【0052】

核酸分子 (SEQ ID NO: 1 - 29) をハイブリダイゼーションおよび増幅技術に用いて、サンプルにおけるSEQ ID NO: 28、SEQ ID NO: 29、および類似分子間の同定および区別が可能である。これらの分子を用いて、遺伝子組換えによりヒトの症状、疾患、または障害を模した疾患動物モデルを作製したり、動物毒素研究、臨床検査、および被験動物 / 患者の治療プロファイルをモニタリングすることが可能である。

#### 【0053】

様々な組織においてGTPAPが発現するが、ノーザン分析によりその少なくとも68%が癌に関連し、32%が炎症および免疫反応に関連し、18%が細胞増殖の症状に関連することが示された。GTPAPが、結腸癌を含む胃腸組織 (23%) で発現することに注目されたい。30

#### 【0054】

GTPAPをコードするポリヌクレオチドが、結腸癌に関連する遺伝子を含む第22染色体の或る領域にマッピングされた。結腸癌に関連する第22染色体の或る領域を画定するパブリックドメイン配列のD22S1171およびDD22S92を用いて、BLAST2によりGTPAPをコードするポリヌクレオチド配列をこの領域内のゲノムDNAとアラインメントさせ、その一致が少なくとも100塩基対に亘って同一性が98%以上であることが分かった。40

#### 【0055】

Rho GAPタンパク質GI 6572185のアミノ酸配列の一部をコードするヒトDNA配列が、結腸癌に関連する遺伝子を含む第22染色体の或る領域に位置することが近年になって分かった。Rho GAPタンパク質は、正常な細胞におけるRho GTPアーゼ調節性シグナル伝達プロセスの負の調節因子すなわちサプレッサーとして作用することから癌抑制因子の可能性がある。

#### 【0056】

(本発明の特徴及び使用)

#### cDNAライブラリ

ここに開示する特定の実施例では、当分野で周知の方法を用いて哺乳動物の細胞及び組織50

から mRNA を単離し、これを用いて cDNA ライブラリを作製する。上記したインサート社クローンは、哺乳動物 cDNA ライブラリから単離された。本発明の代表的な少なくとも 1 つのライブラリの作製方法を以降に記載する実施例に示す。コンセンサス哺乳動物配列は、Phrap (P. Green, 前出) 及び GELVIEW 断片構築システム (Genetics Computer Group, Madison WI)、AUTO ASSEMBLER アプリケーション (PE Biosystems, Foster City CA) などのコンピュータプログラムを用いて、インサート社クローンを含む断片、伸長、及び / またはショットガン配列から化学的かつ / または電子的に構築した。クローン、伸長配列および / またはショットガン配列は、クラスターおよび / または主クラスターの中に電子的に組み入れられた。

10

#### 【0057】

##### シーケンシング

核酸をシーケンシングする方法は当分野で周知であり、そのような方法を用いて本発明の任意の実施例を実施することができる。これらの方法は、DNA ポリメラーゼ I であるクレノウフラグメント、SEQUENASE、Taq DNA ポリメラーゼおよび熱耐性 T7 DNA ポリメラーゼ (Amersham Pharmacia Biotech (APB), Picataway NJ)、或いは ELONGASE 増幅システム (Life Technologies, Rockville MD) に用いられるような校正エクソヌクレアーゼとポリメラーゼとの組み合わせを用いることができる。配列の準備は、HYDRA マイクロディスペンサー (Robbins Scientific, Sunnyvale CA)、MICROLAB 2200 (Hamilton, Reno NV)、及び DNA ENGINE サーマルサイクラー (PTC 200; MJ Research, Watertown MA) などの装置を用いて自動的に行うのが望ましい。シーケンシングに用いる装置には、ABI 3700、377 または 373 DNA シーケンシングシステム (PE Biosystems)、及び MEGABACE 1000 DNA シーケンシングシステム (APB) 等がある。当分野で周知の様々なアルゴリズムを用いて、シーケンシングした配列を解析することができる。これらのアルゴリズムは、Ausubel (1997; Short Protocols in Molecular Biology, John Wiley & Sons, New York NY, unit 7.7) 及び Meyers (1995; Molecular Biology and Biotechnology, Wiley VCH, New York NY, pp. 856 - 853) に記載されている。

20

30

#### 【0058】

ショットガンシーケンシングを用いて、様々な由来するクローニングしたインサートから更に配列を作り出す。ショットガンシーケンシング方法は当分野で周知であり、熱耐性 DNA ポリメラーゼ類や非熱耐性 DNA ポリメラーゼ類、及び目的の核酸分子に隣接する代表的な領域から選択されたプライマーを用いる。当分野で周知の CONSED (Gordon (1998) Genome Res. 8: 195 - 202) などの様々なアルゴリズムやプログラムを用いて、組み立てが不完全な配列を調べる。ベクターやキメラ配列、または欠失配列を含む汚染配列を除去して、組み立てが不完全な配列を完全な配列に組み立てる。

40

#### 【0059】

##### 核酸配列の伸長

本発明の配列は、当分野で周知の様々な PCR 法を用いた方法で伸長することができる。例えば、XL-PCR キット (PE Biosystems) 及び入れ子プライマー (nested primer)、市販の cDNA またはゲノム DNA ライブラリを用いてヌクレオチド配列を伸長することが可能である。全ての PCR 系の方法に用いることができるように、プライマーは、OLIGO 4.06 プライマー分析ソフトウェア (National Biosciences, Plymouth MN) 等の市販のソフトウェアを用いて、ヌクレオチドの長さが約 22 ~ 30 個、GC 含量が約 50% 以上、約 55 ~ 68

50

の温度で標的配列とアニールするように設計することが可能である。調節エレメントを復活させるために配列を伸長する場合は、cDNAライブラリよりゲノムライブラリを用いる方が良い。

【0060】

(本哺乳動物核酸分子の使用)

ハイブリダイゼーション

本哺乳動物核酸分子及びその断片は、様々な目的のための様々なハイブリダイゼーション技術に用いることができる。プローブは、5'調節領域やGTPアーゼ活性化タンパク質グネチンなどの保存されたモチーフから作製可能であり、これらのプローブを、この哺乳動物タンパク質やアレル変位配列、または関連分子をコードする天然の分子を同定するためのプロトコルに用いることができる。通常は一本鎖であるDNA若しくはRNAからなるこのプローブは、任意のこの核酸配列と少なくとも50%の配列同一性を有することが望ましい。ハイブリダイゼーションプローブは、標識されたヌクレオチドの存在下でのPCR増幅、オリゴ標識化、ニックトランスレーション法、または末端標識化を利用して作製することができる。この核酸分子またはその断片を含むベクターを用いて、RNAポリメラーゼ及び標識したヌクレオチドを加えて*in vitro*でmRNAプローブを作製することができる。これらの方法はAmersham Pharmacia Biotech (APB)社が販売するキットを用いて行うことができる。

10

【0061】

ハイブリダイゼーションのストリンジェンシー(厳密性)は、プローブのGC含量、塩濃度、及び温度によって決まる。特に、塩濃度を下げる、またはハイブリダイゼーションの温度を上げて、ストリンジェンシーを高めることができる。ある膜系のハイブリダイゼーション用の溶液にホルムアミドなどの有機溶媒を加えて、反応が低い温度で起こるようにすることができる。ハイブリダイゼーションは、低いストリンジェンシーの緩衝液(5×SSC、1%のドデシル硫酸ナトリウム(SDS))で、60で行うことができるが、核酸配列間に不適正塩基対を含む複合体の形成を許容し得る。続く洗浄は、45(中程度のストリンジェンシー)或いは68(高いストリンジェンシー)の何れかの温度、0.2×SSC、0.1%SDSなどの高いストリンジェンシーで行う。高いストリンジェンシーでは、ハイブリダイゼーション複合体は、完全に相補的な核酸分子部分のみが安定して保持される。ある膜系のハイブリダイゼーションにおいて、好ましくは35%、最も好ましくは50%のホルムアミドをハイブリダイゼーション溶液に加えて、ハイブリダイゼーションを行う温度を下げたり、または、SarkosylやTriton X-100(Sigma Aldrich, St. Louis MO)などの界面活性剤及び変性したサケ精子DNAなどのブロック試薬を用いてバックグラウンドシグナルを低減することが可能である。ハイブリダイゼーションの条件や要素の選択については当分野で周知であり、Ausubel(前出)及びSambrook他(1989)Molecular Cloning, A Laboratory Manual, Cold Spring Harbor Press, Plainview NY.に記載されている。

20

30

【0062】

当分野で周知の方法でマイクロアレイを準備して分析することができる。オリゴヌクレオチドをマイクロアレイのプローブや標的として用いることができる。マイクロアレイを用いて、同時に極めて多数の遺伝子の発現レベルをモニタリングし、遺伝子変異体、突然変異及びSNP(一塩基多型)を同定することができる。このようなデータを用いて、遺伝子機能の解明や、症状及び疾患、または障害における遺伝子原理の解明や、症状及び疾患、障害の診断または治療、治療薬の開発、並びにこれらの治療薬の活性のモニタリングが可能である(例えば、Brennan他(1995)USPN 5,474,796; Schena他(1996)Proc. Natl. Acad. Sci. 93:10614-10619; Baldeschweiler他(1995)PCT出願WO95/251116; Shalon他(1995)PCT出願WO95/35505; Heller他(1997)Proc. Natl. Acad. Sci. 94:21

40

50

50-2155; and Heller 他 (1997) USPN 5,605,662 を参照)。

【0063】

ハイブリダイゼーションプローブはまた、天然のゲノム配列のマッピングに有用である。このプローブは、(1)特定の染色体、(2)染色体の特定の領域、(3)ヒト人工染色体(HAC)や酵母人工染色体(YAC)、細菌人工染色体(BAC)、細菌P1作製物、または単一の染色体、(5)これらの試料の任意のものから作製されたDNAライブラリなどの人工染色体作製物、にハイブリダイズすること可能である。

【0064】

発現

GTPAPをコードする多数の核酸分子をベクターにクローニングして、このタンパク質若しくはその一部を宿主細胞で発現させることができる。この核酸配列を、DNAシャフリング(Stemmer and Cramer i (1996) USPN 5,830,721)や部位特異的変異誘発などの方法によって、新規の制限部位を作り出したり、グリコシル化パターンを変えたり、優先コドンを変えて特定の宿主における発現を増大させたり、スプライスパリアントを作り出したり、半減期を延長する等の操作が可能である。この発現ベクターは、特定の宿主における各要素の効率に基づいて選択された様々なサンプルに由来する転写及び翻訳調節エレメント(プロモーター及びエンハンサー、特定の開始シグナル、ポリアデニル化3'配列)を含み得る。in vitro組み換えDNA技術、合成技術及び/またはin vivo遺伝子組み換え技術を組み合わせて、このベクターに核酸配列と調節エレメントをつなぐことができる。このような技術は、当分野で周知であり、Sambrook(前出、ch. 4, 8, 16 and 17)に記載されている。

10

20

【0065】

様々な宿主系を発現ベクターで形質転換することができる。以下に限定するものではないが、これらの中には組み換えバクテリオファージやプラスミド、またはコスミドDNA発現ベクターで形質転換された細菌と、酵母発現ベクターで形質転換された酵母と、バキュロウイルス発現ベクターで形質転換された昆虫細胞系と、ウイルスエレメント及び/または細菌エレメントを含む発現ベクターで形質転換された植物細胞系や動物細胞系が含まれる(Ausubel前出、unit 16)。例えば、アデノウイルス転写/翻訳複合体を哺乳動物細胞に用いることができる。配列をウイルスのゲノムのE1若しくはE3領域に結合させた後、この感染ウイルスを用いて形質転換させ、宿主細胞でタンパク質を発現させることができる。また、ラウス肉腫ウイルスエンハンサーやSV40、またはEBV系のベクターを用いてタンパク質を高発現させることができる。

30

【0066】

核酸配列のルーチンのクローニング及びサブクローニング、増殖は、多機能PBLUES CRIP Tベクター(Stratagene, La Jolla CA)またはPSPORT1プラスミド(Life Technologies)を用いて行うことができる。核酸配列をこれらのベクターの多数のクローニング部位に導入すると、lacZ遺伝子が破壊され、形質転換された細菌を確認するための比色法によるスクリーニングが可能となる。更に、これらのベクターは、クローニングされた配列におけるin vitroでの転写及びジデオキシ法によるシークエンシング、ヘルパーファージによる一本鎖の調整、入れ子状欠失の作製において有用である。

40

【0067】

長期に渡って組み換えタンパク質を産生させるために、同一或いは別のベクター上の選択マーカー遺伝子或いは可視マーカー遺伝子と共にこのベクターを持続的に細胞株に形質転換することができる。形質転換後、細胞を強化培地で約1~2日間増殖させてから選択培地に移す。選択マーカー、代謝拮抗物質、抗生物質、または除草剤耐性遺伝子は、関連する選択薬に対する抵抗性を与え、導入配列を確実に発現する細胞の増殖及び回収が可能となる。アントシアニン、緑色蛍光タンパク質(GFP)、グルクロニダーゼ、ルシフェ

50

ラーゼなどの可視マーカの発現によって同定された、或いは選択培地に生存することによって同定された耐性クローンを、培養技術を用いて増殖することができる。また、可視マーカを用いて、導入された遺伝子によって発現するタンパク質を定量することができる。宿主細胞が目的の哺乳動物核酸分子を含むか否かの決定は、DNA-DNAまたはDNA-RNAハイブリダイゼーション、或いはPCR増幅技術に基づいて行うことができる。

#### 【0068】

宿主細胞は、組み換えタンパク質を目的の形に修飾する能力に基づいて選択することができる。このような修飾には、アセチル化及びカルボキシル化、グリコシル化、リン酸化、脂質化、及びアシル化等が含まれる。「プレプロ」型を切断する翻訳後プロセッシングを利用して、タンパク質のターゲティング、折り畳み及び/または活性を特定することができる。翻訳後活性のための特定の細胞装置及び特徴的な機構を有する異なった宿主細胞 (Bethesda, MD) を選択して、組み換えタンパク質の正確な修飾及びプロセッシングを確実にすることが可能である。

10

#### 【0069】

##### 細胞培地からのタンパク質の回収

精製を容易にするために、ベクターに導入する異種部分は、グルタチオンSトランスフェラーゼ (GST)、カルモジュリン結合ペプチド (CBP)、6-His、FLAG、MYC等を含む。GST及びCBP、6-Hisはそれぞれ、グルタチオン及びカルモジュリン、金属キレート樹脂が結合した市販のアフィニティマトリックスを用いて精製される。FLAG及びMYCは、市販のモノクローナル抗体及びポリクローナル抗体を用いて精製される。目的のタンパク質配列と異種部分との間にタンパク質分解切断部位を設けて、生成の後の分離が容易にすることができる。組み換えタンパク質の発現及び精製の方法はAusubel (前出、unit 16) に記載され市販されている。

20

#### 【0070】

##### ペプチドの化学合成

タンパク質若しくはその一部は、組み換え方法以外の当分野で周知の化学的方法によって合成することもできる。固相技術を用いるペプチド合成は、パッチ式或いは連続的なフロープロセスによって行うことができる。連続的なフロープロセスでは、アミノ保護及び側鎖保護アミノ酸残基をリンカーを介して不溶性の高分子支持物に連続的に追加する。メチルアミン誘導体化ポリエチレングリコールなどのリンカーを、ポリ(スチレン-co-ジビニルベンゼン)に結合させて支持レジン形成する。このアミノ酸残基は、酸不安定Boc (t-butyl oxycarbonyl) 法若しくは塩基不安定Fmoc (9-fluorenylmethoxycarbonyl) 法によって保護されたN-である。保護されたアミノ酸のカルボキシル基をリンカーのアミンに結合して、この残基を固相支持レジンに結合させる。Boc若しくはFmocを用いた場合、トリフルオロ酢酸若しくはピペリジンを用いて保護基を除去する。カップリング試薬若しくは予め活性化されたアミノ酸誘導体を用いて、追加する各アミノ酸を結合された残基に付加してから、レジンを洗浄する。完全長のペプチドは、連続的な保護の停止、即ち誘導体化アミノ酸を結合させて合成し、ジクロロメタン及び/またはN,N-ジメチルホルムアミドで洗浄する。このペプチドは、ペプチドカルボキシル末端とリンカーとの間で切断され、ペプチド酸またはペプチドアミドが作られる (Novabiochem 1997/98 Catalog and Peptide Synthesis Handbook, San Diego CA pp. S1-S20)。ABI 431A ペプチドシンセサイザー (PE Biosystems) などの装置を用いて、ペプチドを自動合成することができる。タンパク質またはその一部は調整用の高性能液体クロマトグラフィーによって実質的に精製し、その組成をアミノ酸解析またはシーケンシングによって確認することができる (Creighton (1984) Proteins. Structures and Molecular Properties, WH Freeman, New York NY)。

30

40

50

## 【0071】

抗体の準備及びスクリーニング

ヤギ及びウサギ、ラット、マウス、ヒト等を含む様々な宿主は、GTPAP若しくはその任意の一部を注入して免疫することができる。フロイントなどのアジュバント及びミネラルゲルと、リゾレシチン及びpluronic polyol、ポリアニオン、ペプチド、油乳剤、キーホールリンペットヘモシニアン(KLH)、ジニトロフェノールなどの表面活性物質とを用いて免疫反応を高めることができる。オリゴペプチドやペプチド、またはタンパク質の一部を用いて、少なくとも約5個のアミノ酸、より好ましくは10個の天然のタンパク質の一部と同一のアミノ酸を含む抗体を誘発させる。キメラ分子に対する抗体を産生させるために、オリゴヌクレオチドをKLHなどのタンパク質と融合させることができる。

10

## 【0072】

モノクローナル抗体は、培地の連続細胞株によって抗体を産生させる任意の技術を用いて準備する。以下に限定するものではないが、このような技術には、ハイブリドーマ技術及びヒトB細胞ハイブリドーマ技術、EBV-ハイブリドーマ技術が含まれる(例えば、Kohler他(1975) *Nature* 256:495-497; Kozbor他(1985) *J. Immunol. Methods* 81:31-42; Cote他(1983) *Proc. Natl. Acad. Sci.* 80:2026-2030; and Cole他(1984) *Mol. Cell Biol.* 62:109-120. を参照)。

20

## 【0073】

別法では、当分野で周知の方法を用いる上記した一本鎖抗体を生産する技術で、エピトープ特異的一本鎖抗体を生産する。本哺乳動物タンパク質のエピトープに対して特異的に結合する部位を含む抗体断片を生産することが可能である。限定するものではないが、このような断片には、例えば、抗体分子のペプシン消化によって作製されたF(ab')<sub>2</sub>断片及びこのF(ab')<sub>2</sub>断片のジスルフィド架橋を減少させて作製したFab断片が含まれる。別法では、Fab発現ライブラリを作製して、目的の特異性を有するモノクローナルFab断片の高速かつ容易に同定できるようにする(例えば、Huse他(1989) *Science* 246:1275-1281を参照)。

30

## 【0074】

本哺乳動物タンパク質若しくはその一部を用いて、ファージミドまたはBリンパ球免疫グロブリン・ライブラリをスクリーニングして、目的の特異性を有する抗体を同定する。確立された特異性を有するモノクローナル抗体或いはポリクローナル抗体のいずれか一方を用いる、競合的結合またはイムノアッセイの様々なプロトコルが当分野で周知である。このようなイムノアッセイは通常、このタンパク質とその特異的な抗体との複合体形成の測定を行う。2つの非干渉エピトープに反応するモノクローナル抗体を用いる2部位モノクローナル系イムノアッセイが好ましいが、競合的結合によるアッセイを用いることもできる(Pound(1998) *Immunochemical Protocols*. Humana Press, Totowa NJ)。

40

## 【0075】

アッセイのための分子の標識化

多様な標識化及び接合技術は当分野で周知であり、様々な核酸やアミノ酸、及び抗体のアッセイに用いることができる。標識した分子の合成は、<sup>32</sup>P-dCTPまたはCy3-dCTP、Cy5-dCTPなどの標識したヌクレオチドや<sup>35</sup>Sメチオニンなどのアミノ酸を組み込むためのAmersham Pharmacia BiotechキットまたはPromega(Madison WI)を用いて行うことができる。ヌクレオチド及びアミノ酸は、BIODIPYまたはFITC(Molecular Probes, Eugene OR)などの試薬を用いて分子中に存在するアミン及びチオール基または他の基に化学的に結合させることで、様々な物質(蛍光剤または化学発光剤、色素産生剤など)で直接標識することができる。

50

## 【0076】

## (診断)

本核酸分子、断片、オリゴヌクレオチド、相補的なRNA及びDNA分子、PNAを用いて、遺伝子発現の変化やmRNAの過剰な発現の不在/存在を検出及び定量、または治療期間中のmRNAレベルのモニタリングを行うことができる。発現の変化に関連する症状や疾患には、細胞増殖異常、免疫疾患、およびシグナル伝達異常が含まれ、細胞増殖異常の中には日光性角化症及びアテローム性動脈硬化、滑液包炎、硬変、肝炎、混合型結合組織病(MCTD)、乾癬、原発性血小板血症、並びに腺癌及び白血病、リンパ腫、黒色腫、骨髄腫、肉腫、及び奇形癌、具体的には結腸癌が含まれ、免疫疾患の中には炎症及び日光性角化症、後天性免疫不全症候群(AIDS)及び副腎機能不全、成人呼吸窮迫症候群、アレルギー、貧血、喘息、自己免疫性甲状腺炎、気管支炎、滑液包炎、胆嚢炎、硬変、接触皮膚炎、クローン病、糖尿病、肺気腫、痛風、グレーブス病、肝炎、過好酸球増加症、過敏性大腸症候群、多発性硬化症、重症筋無力症、心筋または心膜炎、骨関節炎、骨粗しょう症、膵炎、リウマチ様関節炎、強皮症、潰瘍性大腸炎が含まれ、シグナル伝達異常の中には原発脳腫瘍、腺腫、下垂体切除、動脈瘤、血管奇形、血栓症、および頭部外傷による合併症などの病変から起こる視床下部及び下垂体の障害と、先端巨大症、巨人症、および不適当な抗利尿ホルモン(ADH)分泌症候群(SIADH)を含む下垂体亢進に関連した疾患と、甲状腺腫、粘液水腫、および細菌感染性急性甲状腺炎を含む甲状腺機能低下症に関連した疾患と、Conn病(慢性高カルシウム症)を含む副甲状腺機能亢進症とが含まれる。診断アッセイにハイブリダイゼーションまたは増幅技術を用いて、遺伝子発現の変化を検出するべく、患者からの生体サンプルの遺伝子発現レベルを標準的なサンプルの値と比較する。質的または量的なこのような比較法は当分野で周知である。

## 【0077】

例えば、本核酸分子またはプローブを標準的な方法で標識して、これをハイブリダイゼーション複合体の形成に好適な条件下で、患者からの生体サンプルに加える。インキュベーションの後、このサンプルを洗浄し、ハイブリダイゼーション複合体に関連する標識(シグナル)の量を定量して標準値と比較する。患者のサンプルにおける標識の量が標準値と著しく異なっている場合は、関連する症状や疾患、または異常症の存在が示唆される。

## 【0078】

遺伝子発現に関連する症状や疾患、または異常症の診断のための基準を設けるために、正常或いは標準的な発現プロフィールを確立する。この発現プロフィールは、動物かヒトの正常な被験体から採取した生体サンプルを、ハイブリダイゼーションまたは増幅に好適な条件下で、プローブと結合させることによって確立することができる。標準的なハイブリダイゼーションは、正常な被験体から得た値と、実質的に精製された標的配列を所定量用いた実験値とを比較することによって定量することができる。このように求めた標準値を、特定の症状や疾患、または異常症を示す患者のサンプルから得た値と比較することができる。標準値と特定の症状に関連する値との偏差からその症状を診断する。

## 【0079】

またこのようなアッセイを用いて、動物実験や臨床検査における特定の治療計画の効果を評価したり、患者個人の治療をモニタリングすることができる。病態が確認されると治療プロトコルを開始し、通常ペースで診断アッセイを繰り返して、被験者における発現のレベルが正常な患者に示される値に近づき始めたか否かを調べることが可能である。連続して行ったアッセイの結果から、数日から数ヶ月に渡る期間の治療効果を調べることができる。

## 【0080】

免疫学的方法

特異的なポリクローナル抗体若しくはモノクローナル抗体の何れかを用いるタンパク質の検出及び定量は当分野で周知である。このような技術には、ELISA(酵素結合免疫吸着検定法)及びラジオイムノアッセイ(RIA)、蛍光活性化セルソーター法(FACS)が含まれる。2つの非干渉エピトープに反応するモノクローナル抗体を用いる2部位モ

ノクローナル系イムノアッセイが好ましいが、競合的結合アッセイを用いることもできる（例えば、Coligan 他（1997）Current Protocols in Immunology, Wiley-Interscience, New York NY; および Pound 前出）。

【0081】

（治療）

GTPAPの所定の領域とGAPタンパク質の所定の領域との間に、配列およびモチーフの文脈における化学的および構造的な類似性が存在する。更に、GTPAPの発現が、細胞増殖、特に結腸癌、および免疫反応に密接に関連する。従って、GTPAPは、細胞シグナル伝達異常、免疫疾患、および細胞増殖異常（特に結腸癌）にある役割を果たすと考えられる。発現または活性の上昇に関連する症状の治療においては、タンパク質の発現または活性を低下させることが望ましい。また、発現または活性の低下に関連する症状の治療においては、タンパク質の発現または活性を増大させることが望ましい。

10

【0082】

一実施例において、本哺乳動物タンパク質の発現または活性の変化に関連する症状の治療または予防のために、本哺乳動物タンパク質、その一部、またはその誘導体を患者に投与することが可能である。限定するものではないがこのような疾患の例には、細胞増殖異常、免疫疾患、およびシグナル伝達異常が含まれ、細胞増殖異常の中には日光性角化症及びアテローム性動脈硬化、滑液包炎、硬変、肝炎、混合型結合組織病（MCTD）、乾癬、原発性血小板血症、並びに腺癌及び白血病、リンパ腫、黒色腫、骨髄腫、肉腫、及び奇形癌、具体的には結腸癌が含まれ、免疫疾患の中には炎症及び日光性角化症、後天性免疫不全症候群（AIDS）及び副腎機能不全、成人呼吸窮迫症候群、アレルギー、貧血、喘息、自己免疫性甲状腺炎、気管支炎、滑液包炎、胆嚢炎、硬変、接触皮膚炎、クローン病、糖尿病、肺気腫、痛風、グレーブス病、肝炎、過好酸球増加症、過敏性大腸症候群、多発性硬化症、重症筋無力症、心筋または心膜炎、骨関節炎、骨粗しょう症、膵炎、リウマチ様関節炎、強皮症、潰瘍性大腸炎が含まれ、シグナル伝達異常の中には原発脳腫瘍、腺腫、下垂体切除、動脈瘤、血管奇形、血栓症、および頭部外傷による合併症などの病変から起こる視床下部及び下垂体の障害と、先端巨大症、巨人症、および不適当な抗利尿ホルモン（ADH）分泌症候群（SIADH）を含む下垂体亢進に関連した疾患と、甲状腺腫、粘液水腫、および細菌感染性急性甲状腺炎を含む甲状腺機能低下症に関連した疾患と、Conn病（慢性高カルシウム症）を含む副甲状腺機能亢進症とが含まれる。

20

30

【0083】

別の実施例では、医薬用担体と共に実質的に精製された哺乳動物タンパク質を含む医薬組成物を患者に投与して、限定するものではないが上記した疾患を含むこの内在性タンパク質の活性、または発現の変化に関連する症状の治療または予防を行うことが可能である。

【0084】

更なる実施例では、この哺乳動物タンパク質の活性を調節するリガンドを患者に投与して、限定するものではないが上記した疾患を含むこのタンパク質の寿命や発現、または活性の変化に関連する症状の治療または予防を行うことが可能である。一実施態様では、この哺乳動物タンパク質に特異的に結合する抗体を、この哺乳動物タンパク質を発現する組織や細胞に薬剤を送達するためのターゲティング或いは送達機構として用いることもできる。

40

【0085】

更なる実施例では、この哺乳動物タンパク質またはその一部や誘導体を発現可能なベクターを患者に投与して、限定するものではないが上記した疾患を含むこのタンパク質の寿命や発現、または活性の変化に関連する症状の治療または予防を行うことが可能である。

【0086】

更なる実施例では、この核酸分子やその断片の相補配列を発現するベクターを患者に投与して、限定するものではないが上記した疾患を含むこのタンパク質の寿命や発現、または活性の変化に関連する症状の治療または予防を行うことが可能である。

50

## 【0087】

本核酸分子、またはそれに相補的な分子やその一部、そのタンパク質やその一部の内の任意のもの、これらの核酸分子やタンパク質を運ぶベクター、及びそれらのリガンドをその他の薬剤と共に投与することが可能である。併用療法に用いる薬剤の選択は、当業者が従来の薬学原理に従って行うことができる。薬剤を併用することによって、少量の各薬剤で特定の症状の予防または治療において相乗的な効果をあげることが可能である。

## 【0088】

核酸を用いる遺伝子発現の調節

遺伝子の発現は、哺乳動物遺伝子の5'または3'調節領域、或いは他の調節領域に対して相補的或いはアンチセンス分子を設計することで調節することが可能である。転写開始部位に対して設計されたオリゴヌクレオチドが好ましい。同様に、ポリメラーゼ、転写因子、または調節分子の結合を阻止する三重螺旋塩基対合で遺伝子発現を阻止することができる (Gee 他 In: Huber and Carr (1994) Molecular and Immunologic Approaches, Futura Publishing, Mt. Kisco NY, pp. 163-177)。また、リボソームとmRNAとの結合を阻止して翻訳が行われないように、相補的な分子を設計することも可能である。或るいは、核酸分子またはその断片のライブラリをスクリーニングして、翻訳されない調節配列に特異的に結合する核酸分子または断片を同定することも可能である。

10

## 【0089】

また、酵素活性をもつRNA分子であるリボザイムを用いて、RNAの特異的な切断を触媒してもよい。リボザイム作用のメカニズムは、まずリボザイム分子と相補的な標的RNAとの配列特異的なハイブリダイゼーションが起こり、次にGUAおよびGUU、GUCなどの部位においてヌクレオチド鎖が切断される。このような部位が一旦同定されたら、オリゴヌクレオチドを機能不全にしようとする二次構造特性について、同じ配列を有するオリゴヌクレオチドを評価することができる。また、候補標的としての適合性は、RNA分解酵素保護アッセイを用いて、相補的なオリゴヌクレオチドとのハイブリダイゼーションを検査して評価することができる。

20

## 【0090】

本発明の相補的な核酸およびリボザイムは、固相ホスホラミダイト化学合成法を用いて、in vitroまたはin vivoでの組換え発現によって調製することが可能である。更に、RNA分子は、その5'および/または3'末端に隣接配列を付加して、或いは分子のバックボーン内のホスホジエステル結合の代わりにホスホロチオネートまたは2'-O-メチルを用いて、細胞内の安定性および半減期が増大するように改変することができる。この改変はPNAの作製に固有であるが、他の核酸分子にも適用することができる。例えばイノシン、queosine、wybutosineなどの伝統的でない塩基を含めて、またはアセチル基、メチル基、チオ基でウリジン、アデニン、シチジン、グアニン、およびチミンを修飾して、内在性エンドヌクレアーゼに対する分子の有効性を低くする。

30

## 【0091】

スクリーニングアッセイ

本哺乳動物タンパク質をコードする核酸分子を用いて、分子または化合物のライブラリをスクリーニングして特異的な結合親和性を調べることが可能である。このライブラリは、生物系の核酸分子の活性、複製、転写、または翻訳を調節するDNA分子、RNA分子、PNA、ペプチド、転写因子などのタンパク質、エンハンサー、リプレッサー、およびその他のリガンドを含み得る。このアッセイは、特異的な結合が許容される条件下で、哺乳動物核酸分子またはその断片を分子のライブラリと結合させるステップと、特異的な結合を検出して、その核酸分子と特異的に結合する少なくとも1つの分子を同定するステップとを含む。

40

## 【0092】

50

同様に、本哺乳動物タンパク質またはその一部を用いて、任意の様々なスクリーニングアッセイで分子または化合物のライブラリをスクリーニングすることができる。このようなスクリーニングに用いるタンパク質の一部は、溶液中に遊離している、非生物的若しくは生物学的な基板に固定されている、または細胞内に位置していてもよい。そのタンパク質と分子との特異的な結合を測定することが可能である。スクリーニングするライブラリの種類によって、アッセイでそのタンパク質に特異的に結合するDNA、RNA、PNA分子、アゴニスト、アンタゴニスト、抗体、免疫グロブリン、インヒビター、ペプチド、タンパク質、薬剤、または他の任意のリガンドを同定することが可能である。極めて少ないアッセイ容量および極少量の試験化合物を用いるハイスループットのスクリーニング方法が、米国特許第5,876,946号に記載されており、言及することを以って本明細書の一部とする。この方法は、酵素阻害または受容体結合を調べるために多数の分子をスクリーニングする。

10

#### 【0093】

##### リガンドの精製

本核酸分子またはその断片を用いてサンプルからリガンドを精製することが可能である。哺乳動物核酸分子またはその断片を用いてリガンドを精製する方法は、特異的な結合が許容される条件下でその核酸分子またはその断片とサンプルとを結合させるステップと、特異的な結合を検出するステップと、結合した核酸分子を回収するステップと、好適な試薬を用いて精製されたリガンドから核酸分子を分離するステップとを含む。

#### 【0094】

同様に、本タンパク質またはその一部を用いてサンプルからリガンドを精製することが可能である。哺乳動物タンパク質またはその一部を用いてリガンドを精製する方法は、特異的な結合が許容される条件下でそのタンパク質またはその一部とサンプルとを結合させるステップと、そのタンパク質をリガンドとの間の特異的な結合を検出するステップと、結合したタンパク質を回収するステップと、好適なカオトロピック剤を用いて精製されたリガンドからタンパク質を分離するステップとを含む。

20

#### 【0095】

##### 薬理学

医薬組成物とは、所望の目的を達成するのに効果的な量の活性成分を含んでいる物質である。効果的な薬用量の決定は、当分野の技術者の能力による部分が大きい。どんな化合物であっても、初めは細胞培養アッセイ或いは動物モデルの何れかによって治療効果のある薬用量を推定する。また、動物モデルを使って、好適な濃度範囲および投与経路を決定する。次に、このような情報を用いて、ヒトへの効果的な投与経路および薬用量を決定する。

30

#### 【0096】

治療効果のある薬用量とは、症状または病態を改善するタンパク質またはインヒビターの量である。このような薬剤の薬用効果および毒性は、例えば、ED<sub>50</sub>（集団の50%に医薬的効果がある投与量）およびLD<sub>50</sub>（集団の50%に致命的である投与量）などの細胞培養または実験動物における標準的な製薬方法によって決定することができる。或る投与量における毒性効果と治療効果との比率が治療指数となり、LD<sub>50</sub>/ED<sub>50</sub>と示すことができる。高い治療指数を示す医薬組成物が好ましい。細胞培養アッセイおよび動物実験から得られたデータを用いて、ヒトへ適用する薬用量の範囲を決定する。

40

#### 【0097】

##### モデル系

動物モデルを用いて、ヒトに相当する暴露条件で動物モデルがヒトに類似の毒物反応を示すバイオアッセイを行うことができる。哺乳動物が最も一般的なモデルである。大抵の毒物研究は、低コストで入手が容易であり、かつ参照できる中毒学が豊富であることから、ラットやマウスなどの齧歯類を用いて行われる。齧歯類近交系は、目的の遺伝子の過剰或いは過少な発現の生理学的な原因を調査するのに有用なモデルであり、疾患の診断および治療方法の開発にも有用である。特定の遺伝子を大量に発現する（例えば、乳汁中に分泌

50

される)同系哺乳動物は、その遺伝子によって発現されるタンパク質の便利な供給源となり得る。

【0098】

#### 中毒学

中毒学とは、生物系における物質の影響を研究する学問である。殆どの毒物研究はラットまたはマウスを用いて、物質がヒトの健康に与える影響を推定する。生理機能及び行動、恒常性プロセス、致死率における質的及び量的変化を観察して、毒性プロファイルを作成し、その物質に曝露された後のヒトの健康状態を評価する。

【0099】

遺伝子毒物学は、物質が遺伝子の突然変異を引き起こす能力を同定し分析する。遺伝毒性物質は通常、核酸との相互作用を促進する共通の化学的或いは物理的特性を有し、突然変異した染色体が子孫に受け継がれるのが最大の害である。毒物研究によって、受胎前の両親のどちらか一方、または妊娠中の母、発生段階の生物に投与された場合の、子孫における構造的或いは機能的な異常の頻度を増加させる物質を同定することが可能である。マウス及びラットは生殖周期が短く、統計的必要性を満たす多数の子を出産する能力から、これらの試験にはマウス及びラットが用いられる場合が最も多い。

【0100】

急性毒物の検査は、被験体への物質の一回の投与に基づき、その物質による症状または致死率を決定する。この検査では、(1)初めの投与量の範囲を決定する実験と、(2)有効な投与量の範囲を狭める実験と、(3)用量応答曲線を確立する実験の3つの実験が行われる。

【0101】

長期に渡る毒性検査では繰り返し物質を投与する。このような検査には、ラットやイヌが一般的に用いられ、分類学上異なった種からデータを収集する。物質を高い投与濃度で3~4ヶ月間、毎日投与することで、発癌を除く、成体動物における殆どの毒性の種類が明らかになるという研究結果が多数報告されている。

【0102】

一年或いはそれ以上の長期に渡る慢性毒性検査は、物質に毒性がないこと、或いは物質の発癌の可能性の何れかを実証するために行われる。ラットで検査が行われる場合、少なくとも3つの検査グループと1つの対照グループが用いられ、最初から最後までである間隔で検査及びモニタリングが行われる。

【0103】

#### 遺伝子組換え動物モデル

目的の遺伝子を過剰或いは過小に発現する遺伝子組換え齧歯類を同系交配し、それを用いてヒト疾患モデルを作製したり、治療薬検査や毒物検査を行う(例えば、米国特許第4,736,866号、同第5,175,383号、および同第5,767,337号を参照)。場合によっては、導入遺伝子が、胚発生中若しくは出生後の特定の時期に特定の種類の組織で活性化され得る。導入遺伝子の発現は、実験的薬剤治療を施す前、その最中、またはその後の、遺伝子組換え動物における表現型や組織特異的なmRNAの発現の分析からモニタリングすることができる。

【0104】

#### 胚性幹細胞

齧歯類胚から単離された胚性幹細胞(ES細胞)は、胚を形成する可能性を維持している。ES細胞がキャリアとなる胚の中に導入されると、正常な発生が再開され、生まれる動物の全ての組織の一部を担うことになる。ES細胞は、実験用のノックアウトおよびノックイン齧歯類系を作製するのに好適な細胞である。マウス129/SvJ細胞株などのマウスES細胞は、マウスの初期胚から採取されてから当分野で周知の培養条件下で増殖されたものである。ノックアウト系に用いるベクターには、マーカー遺伝子配列を含むように破壊された疾患候補遺伝子が含まれる。このマーカー遺伝子配列は、in vivoでその転写および/または翻訳を阻害する。このベクターを、当分野で周知の電気穿孔法お

10

20

30

40

50

よびリポソーム法、マイクロインジェクション等の形質転換方法によってES細胞に導入する。齧歯類の内在性遺伝子が、細胞分裂の際の相同組換えや組み込みによって、破壊された疾患遺伝子に置換される。形質転換されたES細胞を同定し、C57BL/6マウス系などからのマウス細胞胚盤胞に微量注入するのが好ましい。この胚盤胞を複数の偽妊娠メスに外科的に導入すると、生まれるキメラ子孫がそれぞれ、その遺伝子型を有し、それらを交配してヘテロ接合系またはホモ接合系を作り出す。

#### 【0105】

また、in vitroでの神経細胞および造血系、心筋細胞などの様々な細胞型や組織の分化の研究に、ES細胞が用いられている(Bain他(1995) *Dev. Biol.* 168:342-357; Wiles and Keller(1991) *Development* 111:259-267; and Klug他(1996) *J. Clin. Invest.* 98:216-224)。近年の研究により、ヒト胚盤胞由来ES細胞をin vitroで操作して、内胚葉および中胚葉、外胚葉性の細胞型を含む8つの別の細胞系譜に分化可能であることが実証された(Thomsonら(1998) *Science* 282:1145-1147)。

#### 【0106】

##### ノックアウト分析

遺伝子ノックアウト分析では、ヒト疾患遺伝子候補のある領域が、ネオマイシン・ホスホトランスフェラーゼ遺伝子などの非哺乳動物遺伝子を含むように酵素によって改変される(neo; Capecchi(1989) *Science* 244:1288-1292)。挿入されたコード配列は、標的遺伝子の転写および翻訳を阻害して疾患候補タンパク質の生化学的合成を阻止する。この改変した遺伝子を培養胚性幹細胞(上記)に形質転換し、この形質転換細胞を齧歯類胚に注入し、この胚を偽妊娠メスに移植する。この遺伝子組換え子孫をクロス交配して、ホモ接合近交系を作り出す。

#### 【0107】

##### ノックイン分析

胚発達の初期段階に現われる分化全能性ES細胞を用いてノックインヒト化動物(ブタ)または遺伝子組換えヒト疾患動物モデル(マウスまたはラット)を作り出すことができる。ノックイン技術を用いて、ヒト遺伝子のある領域を動物ES細胞に注入し、そのヒト配列が組換えによって動物細胞ゲノムの中に組み込まれるようにする。ヒト遺伝子が組み込まれた全能性ES細胞を上記したように操作する。ヒトの症状に類似の情報を収集するべく、この近交系動物の研究および処置を行う。これらの方法を用いて、幾つかのヒト疾患モデルを作り出す(例えば、Lee他(1998) *Proc. Natl. Acad. Sci.* 95:11371-11376; Baudoin他(1998) *Genes Dev.* 12:1202-1216; and Zhuang他(1998) *Mol. Cell Biol.* 18:3340-3349)。

#### 【0108】

##### 非ヒト霊長類モデル

動物実験の分野は、生理学および遺伝学、化学、薬理学、統計学等の基本科学から得たデータおよび方法論を利用する。このようなデータは、非ヒト霊長類がヒトの健康に関連し得るため、非ヒト霊長類における治療薬の効果の評価に重要である。ヒトの代わりにサルをワクチンおよび薬剤の評価に用いる。サルの反応は、類似の条件下でのヒトへの暴露に相当する。カニクイザル(Macaca fascicularis、Macaca mulatta)、および一般的なマーモセット(Callithrix jacchus)は、このような実験に用いられる最も一般的な非ヒト霊長類(NHP)である。NHPの群体制りおよび維持には多大な費用が掛かるため、初期の研究および毒物学的検査は、通常は齧歯類モデルで行われる。薬物嗜癖などの行動を調べる研究では、NHPは最良の実験動物である。更に、個々のNHPおよびヒトは、多くの薬剤および毒物に対して異なった感受性を示すため、これらの物質を「高代謝型」および「低代謝型」に分類することができる。

## 【0109】

更なる実施例では、限定するものではないが、トリプレット遺伝子コードおよび特異的な塩基対相互作用などの特性を含む、現在知られている核酸配列の特性に新しい技術が依存する場合は、本哺乳動物タンパク質をコードする核酸分子を、開発中のあらゆる分子生物学技術に用いることが可能である。

## 【0110】

## 【実施例】

本発明は、記載した特定の装置及び物質、方法に限定されるものではないことを理解されたい。特定の実施例について説明するが、同等の実施例を用いて本発明を具現することも可能である。本発明の範囲は、前記請求の範囲によってのみ限定されるものであって、記載した実施例によって限定されるものではない。また、以下に記載の実施例は、本発明を例示するためのものであって本発明を限定するものではない。例示目的で、ヒト子宮cDNAライブラリUTRSNOR01の作製方法を記載する。

## 【0111】

## 1 代表的なcDNA配列の調整

UTRSNOR01 cDNAライブラリは、29歳の白人女性(標本#0909A)の腔式子宮摘出および膀胱ヘルニア修復術の際に採取した顕微鏡的に正常な子宮内膜から作製した。病理学的には、この子宮は1つの子宮内筋腫を示していた。子宮内膜は分泌期であって、子宮頸は、局所的な扁平化を伴う軽度の慢性子宮頸炎を示していた。患者の病歴には、甲状腺機能低下症、骨盤底弛緩、自動車事故による不完全T-12外傷性対麻痺、および自己導尿があった。家族歴には、良性高血圧症、II型糖尿病、および高脂血症があった。

## 【0112】

凍結組織を、POLYTRONホモジナイザー(PT-3000; Brinkmann Instruments, Westbury, NY)を用いて、TRIZOL試薬(TRIZOL 10ml当たり1gの組織; Life Technologies)、即ち単相溶液フェノールおよびグアニジンイソチオシアネートの溶液においてホモジナイズし溶解した。氷上で短時間インキュベートした後、クロロホルム(1:5 v/v)を加え、溶解物を遠心分離した。上側のクロロホルム層を取り除いて新しい試験管に移し、イソプロパノールでRNAを沈殿させ、DEPC処理した水に再懸濁し、37°Cで25分間DNAアーゼIで処理した。mRNAを、酸性フェノールクロロホルム(pH4.7)で一回再抽出して、0.3M酢酸ナトリウムおよび2.5倍量のエタノールで沈殿させた。次に、mRNAをOLIGOTEXキット(QIAGEN, Inc., Chatsworth, CA)を用いて単離し、これをcDNAライブラリの作製に用いた。

## 【0113】

このmRNAを、cDNA合成およびプラスミドクローニングのためにSUPERSCRIPTプラスミドシステム(Life Technologies)の推奨プロトコルに従って処理した。このcDNAを、SEPHAROSE CL4Bカラム(Amersham Pharmacia Biotech)上で分画化し、400bpを越える大きさのcDNAを、pINCY1プラスミド(Incyte Pharmaceuticals, Palo Alto, CA)に結合させた。次に組換えプラスミドを、DH5αコンピテント細胞(Life Technologies)に形質転換した。

## 【0114】

## 2 cDNAの単離およびシーケンシング

プラスミドDNAを細胞から遊離させ、REAL Prep 96 プラスミドキット(Qiagen)を用いて精製した。このキットにより、マルチチャンネル試薬ディスプレイを用いて96穴プレートの中の96のサンプルを同時に精製することが可能となる。以下の変更点を除いて推奨プロトコルに従った。その変更点は、(1)細菌を、25 mg/Lのカルベニシリンと0.4%のグリセロールを含む1 mlの滅菌TERRIFIC BROTH (Life Technologies)で培養し、(2)接種した後、19時

間インキュベートし、インキュベーションの最後に0.3 mlの溶解緩衝液で溶解し、(3)イソプロパノール沈殿の後、プラスミドDNAペレットを0.1 mlの蒸留水に懸濁した。プロトコルのステップを終了した後、サンプルを96穴プレートに移し4で保管した。

#### 【0115】

このcDNAは、MICROLAB 2200 (Hamilton, Reno NV)と共にPeltier DNA ENGINEサーマルサイクラー (MJ Research, Watertown, MA)を用いて調製した。cDNAは、ABI PRISM 377 DNAシーケンシングシステム (PE Biosystems)を用いて、SangerおよびCoulson (1975; J Mol Biol 94: 441f)の方法によってシーケンシングし、リーディングフレームを決定した。 10

#### 【0116】

##### 3 同定及び伸長、構築、解析

コンセンサス配列 (SEQ ID NO: 28)は、ヒトcDNA配列のZOOSEQデータベース (Incyte Genomics)のインサイトクローン番号: 908465R2 (COLNNOT09)、957130R6 (KIDNNOT05)、1301520F6 (BRSTNOT07)、1580628H1 (DUODNOT01)、2631247F6 (COLNTUT15)、3068538H1 (UTRSNOR01)、および3532286T6 (KIDNNOT25)から組み立て、これを用いて更なるGTPAP関連配列を同定するためにLIFESEQデータベースおよびZOOSEQデータベース (Incyte Genomics)を検索した。MACDNASIS PROソフトウェア (Hitachi Software Engineering)を用いたSEQ ID NO: 28及びSEQ ID NO: 29の翻訳により、翻訳領域であるSEQ ID NO: 30及びSEQ ID NO: 31を解明した。 20

#### 【0117】

当分野で周知のアルゴリズムを用いるソフトウェアプログラムを組み合わせて用いて、cDNAシーケンシングから得たポリヌクレオチド配列を構築および分析した。表1の列1には、用いたツール、プログラム、およびアルゴリズムが示されている。列2には、それらの簡単な説明が示されている。列3には、適当な問合せ先が示されている。これらに言及することを以って本明細書の一部とする。列4には、該当する場合にのみ、2つの配列間の一致の程度を評価するために用いたスコア、確率値、およびその他のパラメーターが示されている (スコアが高ければ高いほど2つの配列間の相溶性が高い)。配列は、MACDNASIS PROソフトウェア (Hitachi Software Engineering, South San Francisco CA)およびLASERGENEソフトウェア (DNASTAR)を用いて分析した。 30

#### 【0118】

ポリヌクレオチド配列は、BALST、動的アルゴリズム、およびジヌクレオチド分布による解析法に基づいたアルゴリズムおよびプログラムを用いて、ベクター、リンカー、ポリA配列を除去し、曖昧な塩基対をマスクして確認した。次に、これらの配列を、GenBankの霊長類、齧歯類、哺乳類、脊椎動物、および真核生物のデータベースや、BLOCKS、PRINTS、DOMO、PRODOM、およびPFAMなどの適当なデータベースに問合せ、BLAST、FASTA、およびBLIMPSに基づいたプログラムを用いてアノテーションを取得した。これらの配列を、Phred、PHRAP、およびConsedに基づいたプログラムを用いて完全長のポリヌクレオチドに組み立て、GeneMark、BLAST、およびFASTAに基づいたプログラムを用いてオープンリーディングフレームを調べるためにスクリーニングした。完全長のポリヌクレオチドを翻訳して対応する完全長のアミノ酸配列を得、これらの完全長のアミノ酸配列を、GenBankデータベース (上記)、SwissProt、BLOCKS、PRINTS、DOMO、PRODOM、Prosite、およびPFAMなどの隠れマルコフモデル (HMM)ベースのタンパク質ファミリーデータベースなどのデータベースに問合せ分析した。 40 50

HMMは、遺伝子ファミリーの共通の一次構造を分析する確率的アプローチである (Eddy, S. R. (1996) Curr. Opin. Str. Biol. 6: 361-365を参照)。

【0119】

完全長のポリヌクレオチド配列およびアミノ酸配列を組立ておよび分析するための上記したプログラムを用いて、SEQ ID NO: 28およびSEQ ID NO: 29からポリヌクレオチド配列断片を同定した。ハイブリダイゼーションおよび増幅技術に有用な約20~4000のヌクレオチドの断片を、本明細書の(発明)の部分に記載した。

【0120】

#### 4 核酸変異体の同定

哺乳動物GTPアーゼ活性化タンパク質をコードする核酸分子(SEQ ID NO: 28及び29)の変異体である核酸分子をBLAST若しくはBLAST2 (Altschul, 前出; デフォルトパラメータを備えたNCI-BLASTN version 2.0.4)を用いて同定し、SEQ ID NO: 28及び29とアラインメントするクローンをLIFSEQデータベースまたはZOOSEQデータベース(Incyte Genomics)において見出した。哺乳動物核酸分子変異体は、BLASTスコアによって選択した。BLASTスコアは、核酸高スコア対(HSP)で一致する各塩基には+5のスコアを与え、不一致の塩基には-4のスコアを与えて計算した。BLASTアラインメントは視覚的に検査でき、BLASTスコアが100を超えるクローンは、Phrap (Green, 前出)を用いて互いにアラインメントした。

【0121】

#### 5 ノーザン分析

ノーザン分析は、遺伝子の転写物の存在を検出するために用いる実験用技術であり、標識したプローブを特定の細胞型或いは組織から抽出したRNAを固定した膜にハイブリダイズさせることが含まれる。

【0122】

BLASTに用いられる類似性コンピュータ技術を用いて、LIFSEQデータベースまたはZOOSEQデータベース(Incyte Genomics)などのヌクレオチドデータベースにおいて、同一或いは関連する分子を検索する。この配列に基づいた解析は、多くの膜系ハイブリダイゼーションより非常に速度が速い。さらにコンピュータ検索の感度を調節して、任意の特定の一致を、同一または類似と分類することができる。BLAST検索の基準は、( $\%配列同一性 \times \%最大BLASTスコア$ ) / 100と定義される積スコアである。積スコアは、2つの配列間の類似度及び配列一致の長さの両方を考慮する。例えば、積スコアが40の場合、その一致は1~2%誤差の範囲内で正確であり、70以上であればその一致は正確であろう。類似する分子或いは関連する分子は通常、8~40の範囲の積スコアを示す分子を選択することにより同定される。

【0123】

ノーザン分析の結果は、この哺乳動物タンパク質をコードする転写物が発生したライブラリの分布割合(パーセント)として報告される。分析には、器官/組織及び疾患によるcDNAライブラリの分類も含まれる。器官/組織のカテゴリーには、心血管、皮膚、発生、内分泌、胃腸、造血/免疫、筋骨格、神経、生殖、泌尿器が含まれる。疾患のカテゴリーには、癌、炎症/外傷、細胞増殖、神経が含まれる。カテゴリー別に、目的の配列を発現するライブラリ数をカウントし、この数をこの哺乳動物タンパク質をコードする核酸分子の核酸分子変異体を含むと同定された全てのライブラリ数で除した。

【0124】

#### 6 核酸分子の伸長

SEQ ID NO: 28およびSEQ ID NO: 29の組み立てに用いた核酸分子の少なくとも1つは、オリゴヌクレオチドプライマーを用いてインサイト社cDNAクローンを伸長して作製した。一方のプライマーは既知の断片の5'の伸長を開始するために合成し、他方のプライマーは既知の断片の3'の伸長を開始するために合成した。開始プライ

10

20

30

40

50

マーは、OLIGO 4.06ソフトウェア(National Biosciences)を用いて、約22~約30個のヌクレオチドの長さ、約50%以上のGC含量で、約55~68の温度で標的配列にアニールするように設計した。ヘアピン構造及びプライマー-プライマー二量体が生じる断片は排除した。選択されたヒトcDNAライブラリを用いてこの分子を伸長した。2段階以上の伸長が必要な場合は、追加の或いは組みになった入れ子のプライマー(nested primer)を設計する。

#### 【0125】

DNA ENGINE サーマルサイクラー(MJ Research)を用いて96ウェルプレートでPCR法を実施し、高い忠実度の増幅を達成した。反応混合液には、鋳型DNA及び200 nmolの各プライマー、反応緩衝液( $Mg^{2+}$ 、 $(NH_4)_2SO_4$ 、 $\beta$ -メルカプトエタノールを含む)、Taq DNAポリメラーゼ(Amersham Pharmacia Biotech)、ELONGASE酵素(Life Technologies)、Pfu DNAポリメラーゼ(Stratagene)を含まれている。プラスミドから選択したプライマーの組みに対して以下のパラメーターで増幅を行った。

ステップ1	94	で3分間	
ステップ2	94	で15秒	
ステップ3	60	で1分間	
ステップ4	68	で2分間	
ステップ5	ステップ2、3、及び4を20回繰り返す		20
ステップ6	68	で5分間	
ステップ7	4	で保管。	

別法では、プライマーの組、T7とSK+(Stratagene)に対して以下のパラメーターで増幅を行った。

ステップ1	94	で3分間	
ステップ2	94	で15秒	
ステップ3	57	で1分間	
ステップ4	68	で2分間	
ステップ5	ステップ2、3、及び4を20回繰り返す		
ステップ6	68	で5分間	30
ステップ7	4	で保管。	

#### 【0126】

各ウェルのDNA濃度は、 $1 \times TE$ 及び0.5  $\mu l$ の希釈していないPCR産物に溶解した100  $\mu l$ のPICO GREEN定量試薬(0.25% (v/v); Molecular Probes)を不透明な蛍光光度計プレート(Corning Science Products, Corning NY)の各ウェルに分注して、DNAがその試薬と結合できるようにして測定する。このプレートをFluoroskan II(Labsystems Oy, Helsinki, Finland)でスキャンして、サンプルの蛍光を測定してDNAの濃度を定量化する。反応混合物の5~10  $\mu l$ のアリコットを1%のアガロースミニゲル上で電気泳動させて解析し、何れの反応がより長い配列の伸長に成功したかを決定する。

#### 【0127】

伸長した配列を脱塩及び濃縮してから384穴プレートに移し、CviJIコレラウィルスエンドヌクラーゼ(Molecular Biology Research, Madison WI)で消化し、pUC18ベクター(Amersham Pharmacia Biotech)に再連結する前に音波処理または切断した。ショットガンシーケンシングのために、消化した断片を0.6~0.8%の濃度のアガロースゲル上に分離させ、UV光で視認できるように断片を切断し、寒天をAGARACE(Promega)で消化/除去した。T4 DNAリガーゼ(New England Biolabs, Beverly MA)を用いて伸長した断片をpUC18ベクター(Amersh

10

20

30

40

50

am Pharmacia Biotech) に再連結し、Pfu DNAポリメラーゼ (Stratagene) で制限部位の延び出しを処理してから、コンピテント大腸菌細胞に形質転換した。形質転換細胞が抗生物質を含む培地で選択され、それぞれのコロニーを切りとって、LB/2Xカルベニシリン培養液の入った384ウェルプレートの中で、37 で一晩培養した。

#### 【0128】

この細胞を溶解して、Taq DNAポリメラーゼ (Amersham Pharmacia Biotech) 及びPfu DNAポリメラーゼ (Stratagene) を用いて以下の手順でDNAを増幅した。

- ステップ1 94 で3分間
- ステップ2 94 で15秒
- ステップ3 60 で1分間
- ステップ4 72 で2分間
- ステップ5 ステップ2、3、及び4を29回繰り返す
- ステップ6 72 で5分間
- ステップ7 4 で保管。

上記したようにPICO GREEN試薬 (Molecular Probes) でDNAを定量した。DNA回収率の悪いサンプルは、上記した条件で再び増幅した。サンプルを20%のジメチルサルホサイド (dimethyl sulphoxide) (1:2, v/v) で希釈し、DYENAMICエネルギー移動シークエンシングプライマー及びDYENAMIC DIRECTキット (Amersham Pharmacia Biotech) またはABI PRISM BIGDYE Terminator cycle sequencing ready reactionキット (PE Biosystems) を用いてシークエンシングした。

#### 【0129】

同様に上述の手順で、SEQ ID NO: 1 - 29のヌクレオチド配列を利用し、外側に伸長するように設計したオリゴヌクレオチドとゲノムDNAライブラリを用いて調節配列を得た。

#### 【0130】

### 7 ハイブリダイゼーション技術および分析

#### ポリヌクレオチドの基板への固定

ポリヌクレオチドを以下に示す方法で基板に固定する。ポリヌクレオチドの混合液をジェル電気泳動により分画後、毛細管輸送によってナイロン膜に移す。別法では、ポリヌクレオチドを個別にベクターに結合して、それを細菌宿主細胞に挿入してライブラリを形成する。次にポリヌクレオチドを以下の方法で基板に配列する。第1の方法では、個別のクローンを含む細菌細胞を機械的に摘み上げてナイロン膜に整列させる。このナイロン膜を、選択薬 (用いるベクターによって異なるが、カルベニシリン、カナマイシン、アンピシリン、またはクロラムフェニコールなど) を含むLB寒天培地に載置し、37 で16時間インキュベートする。この膜を寒天培地から取り除き、次にこの膜をコロニー側を上にして10% SDS、変性溶液 (1.5 M NaCl, 0.5 M NaOH)、中和液 (1.5 M NaCl, 1 M Tris, pH 8.0) に入れ、2xSSCにおいて10分間づつ2回インキュベートする。次にこの膜を、STRATALINKER UVクロスリナー (Stratagene) でUV照射する。

#### 【0131】

第2の方法では、インサートに隣接したベクター配列に相補的なプライマーを用いて、PCRを30サイクル行い細菌ベクターからポリヌクレオチドを増幅する。PCR増幅により、拡散の濃度を開始時の1~2 ngから最終的に5 µgまで増大させる。約400 bp~約5000 bpの増幅した拡散をSEPHACRYL-400ビーズ (Amersham Pharmacia Biotech) を用いて精製する。精製した拡散を、手動で或いはドット/スロットプロットングマニフォールド及び吸入装置でナイロン膜に配

10

20

30

40

50

列し、上記した変性、中和、及びUV照射によって固定する。精製した拡散を、USPN 5, 807, 522に開示された方法でポリマーコートスライドガラスに機械的に配列して固定する。ポリマーコートスライドガラスは、顕微鏡スライドガラス(Corning, Acton MA)を0.1% SDS及びアセトンで超音波洗浄した後、4%フッ化水素酸(VWR Scientific Products, West Chester PA)においてエッチングし、次に95%エタノールにおいて0.05%アミノプロピルシランでコーティングし、オーブンで110 で硬化させて準備する。処理中及び処理後にこのスライドガラスを蒸留水で十分に洗浄する。次に、核酸分子をスライド上に配列し、STRATALINKER UVクロスリンカー(Stratagene)を用いてこの核酸分子のアレイをUV照射してスライド上に固定する。次にこのアレイを、室温で0.2% SDSで洗浄し、蒸留水で3回すすぐ。非特異的な接合部位を、アレイをPBS(Tropix, Bedford MA)に於ける0.2%カゼインにおいて60 で30分間インキュベートしてブロックし、次にこのアレイを0.2% SDSで洗浄し、前記したように蒸留水で洗い流す。

#### 【0132】

##### メンブレンを用いるハイブリダイゼーションのためのプローブの準備

本配列表のポリヌクレオチドに由来するハイブリダイゼーションプローブは、メンブレンハイブリダイゼーションにおいてcDNA、mRNA、またはゲノムDNAをスクリーニングするために用いられる。プローブを準備するために、ポリヌクレオチドを45 µl TEバッファにおいて40~50 ngの濃度に希釈し、100 で5分間加熱して変性し、短時間遠心分離する。次に、変性したポリヌクレオチドをREDIPRIMEチューブ(Amersham Pharmacia Biotech)に加え、青色が十分に拡散するまで軽く混合し、短時間遠心分離する。5 µlの[<sup>32</sup>P] dCTPをチューブに加え、その内容物を37 で10分間インキュベートする。この標識化反応を5 µlの0.2 M EDTAを加えてストップし、PROBEQUANT G-50マイクロカラム(Amersham Pharmacia Biotech)を用いてプローブを組み込まれなかったヌクレオチドから精製する。精製したプローブを100 で5分間加熱してから氷上で2分間冷却し、以下に記載するようにメンブレンハイブリダイゼーションに用いる。

#### 【0133】

##### ポリマーコートスライドを用いるハイブリダイゼーションのためのプローブの準備

サンプルから単離したmRNAに由来するハイブリダイゼーションプローブを用いて、アレイハイブリダイゼーションにおいて本配列表のポリヌクレオチドをスクリーニングする。プローブは、GEMbrightキット(Incyte Pharmaceuticals)を用いて、mRNAを9 µl TEバッファにおいて200 ngの濃度に希釈し、5 µlの5xバッファ、1 µlの0.1 M DTT、3 µlのCy3またはCy5標識混合物、1 µlのRNアーゼインヒビター、1 µlの逆転写酵素、及び5 µlの1x酵母コントロールmRNAを加えて準備する。酵母コントロールmRNAは、非コード酵母ゲノムDNA(W. Lei, unpublished)から*in vitro*転写によって合成する。量的コントロールとして、0.002 ng、0.02 ng、0.2 ng、及び2 ngのコントロールmRNAからなるセットを、サンプルmRNAに対してそれぞれ1:100, 000、1:10, 000、1:1000、及び1:100 (w/w)の比率で逆転写反応液に希釈する。mRNAの異なった発現パターンを検査するために、第2のセットのコントロールmRNAを、1:3、3:1、1:10、10:1、1:25、及び25:1 (w/w)の比率で逆転写反応液に希釈する。反応液を混合し、37 で2時間インキュベートする。次に反応液を85 で20分間インキュベートし、2つの連続したCHROMA SPIN+TE 30カラム(Clontech)を用いてプローブを精製する。精製したプローブを、DEPC処理水で90 µlに希釈して、2 µlの1 mg/mlグリコーゲン、60 µlの5 M酢酸ナトリウム、300 µlの100%エタノールを加えてエタノール沈殿させる。このプローブを20, 800 x gにおいて20

分間遠心分離し、ペレットを12  $\mu$ lの再懸濁バッファに再懸濁させ、65  $^{\circ}$ Cで5分間加熱し、完全に混合する。上記したようにプローブを加熱及び混合してから氷上で保管する。プローブを以下に記載するように高密度のアレイハイブリダイゼーションに用いる。

#### 【0134】

##### メンブレンを用いるハイブリダイゼーション

膜を、1% Sarkosyl及び1x高リン酸バッファ(0.5M NaCl、0.1M Na<sub>2</sub>HPO<sub>4</sub>、5mM EDTA、pH 7)を含むハイブリダイゼーション溶液において55  $^{\circ}$ Cで2時間プレハイブリダイゼーションする。15mlの新しいハイブリダイゼーション溶液に希釈したプローブを膜に加える。55  $^{\circ}$ Cで16時間、膜にプローブをハイブリダイズさせる。ハイブリダイゼーションの後、膜を1mM Tris(pH 8.0)、1% Sarkosylにおいて25  $^{\circ}$ Cで15分間洗浄し、1mM Tris(pH 8.0)において25  $^{\circ}$ Cで15分間づつ4回洗浄する。ハイブリダイゼーション複合体を検出するために、XOMAT-ARフィルム(Eastman Kodak, Rochester NY)を膜に一晩-70  $^{\circ}$ Cで露出し、現像して目で確認する。

10

#### 【0135】

##### ポリマーコートスライドを用いるハイブリダイゼーション

プローブを65  $^{\circ}$ Cで5分間加熱し、5415Cマイクロ遠心分離器(Eppendorf Scientific, Westbury NY)を用いて毎分9400回転で5分間遠心分離し、アレイの表面に18  $\mu$ lのアリコットを載せカバースリップで覆う。このアレイを、顕微鏡スライドより僅かに大きいキャピティを有する防水容器に移す。この容器の角から140  $\mu$ lの5xSSCを加えてその内部を100%の湿度に保つ。アレイを含む容器を60  $^{\circ}$ Cで約6時間半インキュベートする。このアレイを、1xSSC、0.1% SDSにおいて45  $^{\circ}$ Cで10分間洗浄し、0.1xSSCにおいて45  $^{\circ}$ Cで10分間づつ3回洗浄した後、乾燥させる。

20

#### 【0136】

ハイブリダイゼーション反応を、絶対ハイブリダイゼーション(absolute hybridization)方式またはディファレンシャルハイブリダイゼーション方式で行う。絶対ハイブリダイゼーション方式の場合、1つのサンプルからのプローブをアレイの要素にハイブリダイズさせ、ハイブリダイゼーション複合体が形成された後にシグナルを検出する。シグナルの強度がサンプルにおけるプローブmRNAのレベルに相関する。ディファレンシャルハイブリダイゼーション方式の場合、2つの生体サンプルにおける遺伝子のセットの異なった発現を分析する。2つのサンプルからのプローブを準備し、異なった標識成分で標識する。2つの標識したプローブの混合物をアレイ要素にハイブリダイズさせ、2つの異なった標識からの蛍光シグナルをそれぞれ別に検出できる条件下でシグナルを検査する。アレイ上の要素に、2つの生体サンプルに由来するプローブのそれぞれが実質的に同数ハイブリダイズする場合、明瞭な複合蛍光が得られる(Shalon W 095/35505)。

30

#### 【0137】

ハイブリダイゼーション複合体をInnova 70混合ガス10Wレーザーを備えた顕微鏡(Coherent, Santa Clara CA)で検出する。この顕微鏡は、Cy3を励起するために488nmのスペクトル線が発生し、Cy5を励起するために632nmのスペクトル線が発生することができる。励起レーザー光を20倍の顕微鏡対物レンズ(Nikon, Melville NY)を用いてアレイに集束させる。アレイを含むスライドを顕微鏡上のコンピュータ制御されたXYステージ上に載せ、20  $\mu$ mの解像度で対物レンズを通してラスタースキャンする。ディファレンシャルハイブリダイゼーション方式の場合、2つの蛍光物質を順番にレーザーで励起する。波長に応じて放出された光を2つの蛍光物質に対応する2つの光電子増倍管検出器(PMT R1477, Hamamatsu Photonics Systems, Bridgewater NJ)に分割する。アレイと光電子増倍管との間に配置された好適なフィルターを用いてシグナルをフィルタリングする。用いられる蛍光物質の最大発光波長は、Cy3が565nmで

40

50

あり、Cy5が650nmである。スキャンの感度は、プローブ混合物に加えられた酵母コントロールmRNAによって生成されるシグナル強度を用いて校正する。アレイのある位置に相補的なDNA配列が含まれていると、その位置におけるシグナルの強度が、ハイブリダイズ種(hybridizing species)の重量比で1:100,000の関係となり得る。

#### 【0138】

光電子増倍管の出力をIBM互換性パーソナルコンピュータにインストールした12ビットRTI-835Hアナログ/デジタル変換ボードを用いてデジタル化する(Analog Devices, Norwood MA)。このデジタル化したデータを、イメージとして表示する。この場合、シグナル強度を、20色線形変換を用いて、青(弱いシグナル)から赤(強いシグナル)の疑似色スケールにマッピングする。データはまた量的にも分析する。2つの異なった蛍光物質が同時に励起されて測定された場合、データを、各蛍光物質の発光スペクトルを用いて蛍光物質間の光学クロストーク(発光スペクトルの重複による)をまず補正する。各スポットからのシグナルがグリッドの各要素の中心に来るように、蛍光物質シグナル強度の上にグリッドを重ね合わせる。各要素内の蛍光シグナルを統合して、シグナルの平均強度に対応する数値を求める。シグナル分析に用いるソフトウェアはGEMTOOLSプログラム(Incyte Pharmaceuticals)である。

10

#### 【0139】

### 8 相補的な核酸配列

本核酸分子に相補的な配列、或いはその断片を用いて、遺伝子の発現を検出したり、低下させたり、阻害することができる。約15~約30個の塩基対からなるオリゴヌクレオチドの使用について記載するが、それより小さい或いは大きい配列の断片、またはその誘導体(PNA)の場合でも同じ方法を用いることができる。オリゴヌクレオチドは、SEQ ID NO: 1-29を用いて、Oligo4.06ソフトウェア(National Biosciences)で設計する。プロモーターの結合を阻害して転写を阻止するために、相補的なヌクレオチドを、最も好ましくはオープンリーディングフレームの開始コードンの前の約10個のヌクレオチドである最もユニークな5'配列に結合するように設計する。翻訳を阻害するために、相補的なオリゴヌクレオチドを、本哺乳動物タンパク質をコードするmRNAにリボソームが結合するのを阻害するように設計する。

20

30

#### 【0140】

### 9 哺乳動物タンパク質の発現

本哺乳動物タンパク質の発現及び精製は、細菌若しくはウイルス系の発現系を用いて行うことができる。細菌で発現させるために、抗生物質耐性遺伝子とcDNAを高レベルで転写させる誘導性プロモーターとを含むベクターにcDNAをサブクローニングする。このようなプロモーターには、以下に限定するものではないが、lacオペレーター調節エレメントに関連するT5またはT7バクテリオファージプロモーター及びtrp-lac(ctac)ハイブリッドプロモーターが含まれる。組み換えベクターを、例えばBL21(DE3)などの好適な細菌宿主に形質転換する。イソプロピル-Dチオガラクトピラノシド(IPTG)で誘導されると、抗生物質耐性細菌がこの哺乳動物タンパク質を発現する。真核細胞での発現は、Autographica californica核多面性ウイルス(ACMNPV)という組み換えバキュロウイルスをSpodoptera frugiperda(Sf9)昆虫細胞に感染させて行う。バキュロウイルスのポリヘドリン遺伝子を、相同組み換え或いは転移プラスミド中間体が関係する細菌媒介遺伝子転移の何れかによって、この哺乳動物cDNAと置換する。ウイルスの感染力は維持され、強力なポリヘドリンプロモーターによって高いレベルでcDNAが転写される。

40

#### 【0141】

大抵の発現系では、この哺乳動物タンパク質がグルタチオン-S-トランスフェラーゼ(GST)またはペプチドエピトプタグ(FLAG)などとの融合タンパク質として合成されるため、未精製の細胞溶解物から親和性を用いた迅速な一段階の組み換え融合タンパ

50

ク質の精製を行うことができる。GSTによって、タンパク質の活性及び抗原性を維持する条件下で、固定されたグルタチオンでの融合タンパク質の精製が可能となる(Amersham Pharmacia Biotech)。精製の後、GST部分をこの哺乳動物タンパク質から蛋白分解的に遺伝子操作された特定部位で切断できる。アミノ酸8個のペプチドからなるFLAGによって、市販のモノクローナル及びポリクローナル抗FLAG抗体(Eastman Kodak)を用いたイムノアフィニティー精製が可能となる。6個の連続するヒスチジン残基のストレッチである6-Hisによって、金属キレート樹脂(QIAGEN)での精製が可能となる。タンパク質の発現及び精製の方法は、Ausubel(1995, 前出, ch 10, 16)に記載されている。これらの方法で精製したこの哺乳動物タンパク質を直接用いて以下のアッセイを行うことができる。

10

## 【0142】

10 機能のアッセイ

GTPAP活性のアッセイは、Rho GTPアーゼにおける内在性GTPアーゼ活性の刺激を測定する。GTPアーゼ活性は、RhoAおよびCdc42-GTPアーゼを用いて、文献(Lira(1997) J Biol Chem 272 : 32830-32835)に記載してあるように測定する。アッセイは、GTPAPを含めず、および様々な濃度のGTPAPの存在下で実施する。GTPAPの存在下でのGTPアーゼの活性量からGTPAPが不在下での活性量を差引いた値が、サンプルにおけるGTPAPの活性に比例する。

20

## 【0143】

別法では、タンパク質の機能は、哺乳動物細胞培養でGTPAPをコードする配列を生理学的に高レベルで発現させてアッセイする。この核酸配列を、強力なサイトメガロウイルスプロモーターを含むPCMV SPORTベクター(Life Technologies)にサブクローニングし、この組み換えベクター5~10µgを、当分野で周知の形質転換法方によってヒトの内皮細胞株或いは造血細胞株に形質転換する。更に、蛍光マーカーとなるCD64-GFP(Clontech, Palo Alto CA)をコードする配列を含むプラスミド1~2µgを同時に形質転換して、フローサイトメトリー(FCM)で形質転換細胞を同定できるようにする。

## 【0144】

発現における導入された遺伝子の影響は、精製されたこれらの形質転換細胞集団を用いて評価することができる。形質転換細胞の表面で発現されるCD64-GFPが、ヒト免疫グロブリンG(IgG)の保存された領域と結合するため、この形質転換細胞を、ヒトIgGか抗CD64抗体の何れかで被覆された磁気ビーズを用いて分離することができる(DYNAL, Lake Success, NY)。mRNAを細胞から単離して、ハイブリダイゼーション技術で分析することができる。

30

## 【0145】

11 GTPAPに特異的な抗体の生産

ポリアクリルアミドゲル電気泳動法でGTPAPを精製し、これを用いてマウスやウサギを免疫する。以下の標準的なプロトコルを用いて抗体を生産する。別法では、GTPAPのアミノ酸配列をLASERGENEソフトウェア(DNA STAR)を用いて解析し、免疫原性の高い領域を決定する。通常はC末端付近或いは親水性領域内に存在する免疫原性エピトープを選択して合成し、これを用いて抗体を増大させる。通常は15残基程度の長さのエピトープを、Fmoc法でABI 431Aペプチドシンセサイザー(PE Biosystems)で合成し、これをN-マレイミドベンゾイル-N-ヒドロキシスクシンイミドエステルとの反応でKLH(Sigma-Aldrich)に結合させて免疫原性を高める。

40

## 【0146】

エピトープ-KLH複合体をフロイントの完全アジュバントと共に投与してウサギを免疫する。その後、間隔を置いてフロイントの不完全アジュバントで免疫化を繰り返す。マウスの場合は少なくとも7週間、ウサギの場合は少なくとも12週間が経過した後、抗血清

50

を採取して抗ペプチド活性を検査する。この検査には、このペプチドをプラスチックに結合し、1%ウシ血清アルブミン(BSA)でブロッキング処理し、ウサギ抗血清と反応させ、洗浄し、さらに放射性ヨウ素標識されたヤギ抗ウサギIgGと反応させる。当分野で周知の方法を用いて抗体力価及び形成された複合体の収量を測定する。

【0147】

#### 1.2 特異的な抗体を用いる天然タンパク質の精製

天然或いは組み換えの哺乳動物タンパク質を、このタンパク質に特異的な抗体を用いてイムノアフィニティークロマトグラフィで実質的に精製する。イムノアフィニティークラムは、この抗体をCNBr-活性化SEPHAROSEレジン(Amersham Pharmacia Biotech)に共有結合させて作製する。このタンパク質を含む培養液をイムノアフィニティークラムに通し、このタンパク質を優先的に吸着できるように、界面活性剤の存在下で高イオン強度緩衝液でそのカラムを洗浄する。結合後、緩衝液(pH 2~3)或いは高濃度の尿素やチオシアネートイオンを用いてそのカラムからこのタンパク質を溶離させて抗体とこのタンパク質との結合を切断し、このタンパク質を回収する。

【0148】

#### 1.3 ポリヌクレオチドの第22染色体の結腸癌領域へのマッピング

第22染色体のパブリックドメインゲノムDNA配列データは、GenBankから入手でき、文献(Dunham, I.ら(1999) Nature 402: 489-495)に記載されている。インサイトID No. 440424.5(SEQ ID NO: 29)を含むインサイトテンプレートを、BLAST2を用いてパブリックドメインゲノム配列と整列させ、2つの配列間に100塩基対以上に亘って98%以上の同一性を有する配列を特定した。前出のCastellsらによって同定された結腸癌に関連する第22染色体の或る領域が、パブリックドメインのマーカーD22S1171およびD22S92によって画定される。これらのマーカーを、BALST2を用いて第22染色体の配列をアラインメントし、この領域に関連するゲノムDNAの座位を決定する。上記したSEQ ID NO: 29に一致したゲノムDNA配列がこの領域内にマッピングされた。

【0149】

#### 1.4 本核酸分子若しくはタンパク質と特異的に結合する分子のスクリーニング

本核酸分子やその断片、或いは本タンパク質やその断片を、<sup>32</sup>P-dCTP、Cy3-dCTP、Cy5-dCTP(Amersham Pharmacia Biotech)、またはBIODIPYやFITC(Molecular Probes)でそれぞれ標識する。予め基板上に配列した候補分子のライブラリを、標識した核酸分子またはタンパク質の存在下でインキュベートする。核酸分子或いはアミノ酸配列の何れかが存在する条件下でインキュベートした後、その基板を洗浄し、特異的な結合或いは複合体形成を示唆する標識が保持されている基板の全ての部分をアッセイし、結合している分子を同定する。様々な濃度の核酸若しくはタンパク質で得られたデータを用いて、標識した核酸またはタンパク質と結合した分子との親和性を計算する。

【0150】

#### 1.5 GTPAPと相互作用する分子の同定

GTPAP若しくは生物学的活性を有するその断片を、<sup>125</sup>I Bolton-Hunter試薬(Bolton他(1973) Biochem. J. 133: 529-539)で標識する。マルチウェルプレートのウェルにおいて、候補リガンド分子を標識したGTPAPと共にインキュベートし、次に洗浄してから、標識した複合体を有する全てのウェルをアッセイする。様々な濃度のGTPAPで得られたデータを用いて、GTPAPと結合した候補リガンド分子の関連性、数量、親和性を計算する。

【0151】

(表の簡単な説明)

表1は、GTPAPの分析に用いたツール、プログラム、およびアルゴリズム、並びに該

10

20

30

40

50

当する説明、参照、および閾値パラメーターを示す。

【図面の簡単な説明】

【図1A】

本哺乳動物タンパク質のヒトアミノ酸配列 (SEQ ID NO: 30) をコードするヒト核酸分子 (SEQ ID NO: 28) を示す。このアラインメントは、MACDNASIS PROソフトウェア (Hitachi Software Engineering, South San Francisco CA) を用いて作成した。

【図1B】

本哺乳動物タンパク質のヒトアミノ酸配列 (SEQ ID NO: 30) をコードするヒト核酸分子 (SEQ ID NO: 28) を示す。このアラインメントは、MACDNASIS PROソフトウェア (Hitachi Software Engineering, South San Francisco CA) を用いて作成した。 10

【図1C】

本哺乳動物タンパク質のヒトアミノ酸配列 (SEQ ID NO: 30) をコードするヒト核酸分子 (SEQ ID NO: 28) を示す。このアラインメントは、MACDNASIS PROソフトウェア (Hitachi Software Engineering, South San Francisco CA) を用いて作成した。

【図1D】

本哺乳動物タンパク質のヒトアミノ酸配列 (SEQ ID NO: 30) をコードするヒト核酸分子 (SEQ ID NO: 28) を示す。このアラインメントは、MACDNASIS PROソフトウェア (Hitachi Software Engineering, South San Francisco CA) を用いて作成した。 20

【図1E】

本哺乳動物タンパク質のヒトアミノ酸配列 (SEQ ID NO: 30) をコードするヒト核酸分子 (SEQ ID NO: 28) を示す。このアラインメントは、MACDNASIS PROソフトウェア (Hitachi Software Engineering, South San Francisco CA) を用いて作成した。

【図2A】

本哺乳動物タンパク質のヒトアミノ酸配列 (SEQ ID NO: 31) をコードするヒト核酸分子 (SEQ ID NO: 29) を示す。このアラインメントは、MACDNASIS PROソフトウェア (Hitachi Software Engineering) を用いて作成した。 30

【図2B】

本哺乳動物タンパク質のヒトアミノ酸配列 (SEQ ID NO: 31) をコードするヒト核酸分子 (SEQ ID NO: 29) を示す。このアラインメントは、MACDNASIS PROソフトウェア (Hitachi Software Engineering) を用いて作成した。

【図2C】

本哺乳動物タンパク質のヒトアミノ酸配列 (SEQ ID NO: 31) をコードするヒト核酸分子 (SEQ ID NO: 29) を示す。このアラインメントは、MACDNASIS PROソフトウェア (Hitachi Software Engineering) を用いて作成した。 40

【図2D】

本哺乳動物タンパク質のヒトアミノ酸配列 (SEQ ID NO: 31) をコードするヒト核酸分子 (SEQ ID NO: 29) を示す。このアラインメントは、MACDNASIS PROソフトウェア (Hitachi Software Engineering) を用いて作成した。

【図2E】

本哺乳動物タンパク質のヒトアミノ酸配列 (SEQ ID NO: 31) をコードするヒト核酸分子 (SEQ ID NO: 29) を示す。このアラインメントは、MACDNASIS PROソフトウェア (Hitachi Software Engineering) を用いて作成した。 50

S PROソフトウェア (Hitachi Software Engineering) を用いて作成した。

【図3A】

ヒトRhoGAPタンパク質、GTPAP2 (クローンID 404424; SEQ ID NO: 31)、GTPAP1 (クローンID 3068538; SEQ ID NO: 30)、Rho GTPアーゼ活性化タンパク質8 (GI 6572185; SEQ ID NO: 32)、およびrhoGAPタンパク質 (GI 312212; SEQ ID NO: 33) との間の化学的および構造的類似性を実証する。このアラインメントは、LASER GENEソフトウェアのMEGALIGNプログラム (DNASTAR, Madison WI) を用いて作成した。

10

【図3B】

ヒトRhoGAPタンパク質、GTPAP2 (クローンID 404424; SEQ ID NO: 31)、GTPAP1 (クローンID 3068538; SEQ ID NO: 30)、Rho GTPアーゼ活性化タンパク質8 (GI 6572185; SEQ ID NO: 32)、およびrhoGAPタンパク質 (GI 312212; SEQ ID NO: 33) との間の化学的および構造的類似性を実証する。このアラインメントは、LASER GENEソフトウェアのMEGALIGNプログラム (DNASTAR, Madison WI) を用いて作成した。

【図3C】

ヒトRhoGAPタンパク質、GTPAP2 (クローンID 404424; SEQ ID NO: 31)、GTPAP1 (クローンID 3068538; SEQ ID NO: 30)、Rho GTPアーゼ活性化タンパク質8 (GI 6572185; SEQ ID NO: 32)、およびrhoGAPタンパク質 (GI 312212; SEQ ID NO: 33) との間の化学的および構造的類似性を実証する。このアラインメントは、LASER GENEソフトウェアのMEGALIGNプログラム (DNASTAR, Madison WI) を用いて作成した。

20

【図3D】

ヒトRhoGAPタンパク質、GTPAP2 (クローンID 404424; SEQ ID NO: 31)、GTPAP1 (クローンID 3068538; SEQ ID NO: 30)、Rho GTPアーゼ活性化タンパク質8 (GI 6572185; SEQ ID NO: 32)、およびrhoGAPタンパク質 (GI 312212; SEQ ID NO: 33) との間の化学的および構造的類似性を実証する。このアラインメントは、LASER GENEソフトウェアのMEGALIGNプログラム (DNASTAR, Madison WI) を用いて作成した。

30

【図4A】

ヒトGTPAPをコードする核酸分子間のアラインメント、並びに一塩基多型 (SNP) の領域を示す。SNPは で示されている。このアラインメントは、PHRAPソフトウェア (Phil Green, University of Washington, Seattle WA) を用いて作成した。

【図4B】

ヒトGTPAPをコードする核酸分子間のアラインメント、並びに一塩基多型 (SNP) の領域を示す。SNPは で示されている。このアラインメントは、PHRAPソフトウェア (Phil Green, University of Washington, Seattle WA) を用いて作成した。

40

【図5A】

PHRAPソフトウェア (Phil Green, University of Washington, Seattle WA) を用いて作成したアラインメントであって、ヒト3068538CB1 (SEQ ID NO: 28) 核酸配列、およびラット701244926H1 (SEQ ID NO: 21)、700950169H2 (SEQ ID NO: 22)、70157597H1 (SEQ ID NO: 23)、701274036

50

H1 (SEQ ID NO: 24)、700480528 (SEQ ID NO: 25)、700935753H1 (SEQ ID NO: 26)、および700936061H1 (SEQ ID NO: 27) 核酸配列との間の化学的および構造的類似性を示す。

【図5B】

PHRAPソフトウェア(Phil Green, University of Washington, Seattle WA)を用いて作成したアラインメントであって、ヒト3068538CB1 (SEQ ID NO: 28) 核酸配列、およびラット701244926H1 (SEQ ID NO: 21)、700950169H2 (SEQ ID NO: 22)、70157597H1 (SEQ ID NO: 23)、701274036H1 (SEQ ID NO: 24)、700480528 (SEQ ID NO: 25)、700935753H1 (SEQ ID NO: 26)、および700936061H1 (SEQ ID NO: 27) 核酸配列との間の化学的および構造的類似性を示す。

10

【図5C】

PHRAPソフトウェア(Phil Green, University of Washington, Seattle WA)を用いて作成したアラインメントであって、ヒト3068538CB1 (SEQ ID NO: 28) 核酸配列、およびラット701244926H1 (SEQ ID NO: 21)、700950169H2 (SEQ ID NO: 22)、70157597H1 (SEQ ID NO: 23)、701274036H1 (SEQ ID NO: 24)、700480528 (SEQ ID NO: 25)、700935753H1 (SEQ ID NO: 26)、および700936061H1 (SEQ ID NO: 27) 核酸配列との間の化学的および構造的類似性を示す。

20

【図5D】

PHRAPソフトウェア(Phil Green, University of Washington, Seattle WA)を用いて作成したアラインメントであって、ヒト3068538CB1 (SEQ ID NO: 28) 核酸配列、およびラット701244926H1 (SEQ ID NO: 21)、700950169H2 (SEQ ID NO: 22)、70157597H1 (SEQ ID NO: 23)、701274036H1 (SEQ ID NO: 24)、700480528 (SEQ ID NO: 25)、700935753H1 (SEQ ID NO: 26)、および700936061H1 (SEQ ID NO: 27) 核酸配列との間の化学的および構造的類似性を示す。

30

【図5E】

PHRAPソフトウェア(Phil Green, University of Washington, Seattle WA)を用いて作成したアラインメントであって、ヒト3068538CB1 (SEQ ID NO: 28) 核酸配列、およびラット701244926H1 (SEQ ID NO: 21)、700950169H2 (SEQ ID NO: 22)、70157597H1 (SEQ ID NO: 23)、701274036H1 (SEQ ID NO: 24)、700480528 (SEQ ID NO: 25)、700935753H1 (SEQ ID NO: 26)、および700936061H1 (SEQ ID NO: 27) 核酸配列との間の化学的および構造的類似性を示す。

【表1】

40

表 1-1

プログラム名	説明	引用文献	パラメーター閾値
ABI FACTURA	核酸配列においてベクター配列を除去して不定の塩基をマスクするプログラム。	Perkin-Elmer Applied Biosystems, Foster City, CA.	
ABI/PARACEL FDF	Fast Data Finder は、アミノ酸または核酸配列の比較及び注釈付け (annotation) に有用である。	Perkin-Elmer Applied Biosystems, Foster City, CA; Paracel Inc., Pasadena, CA.	不一致<50%
ABI AutoAssembler	核酸配列を構築するプログラム。	Perkin-Elmer Applied Biosystems, Foster City, CA.	
BLAST	Basic Local Alignment Search Tool は、アミノ酸及び核酸配列の配列類似性検索に有用であり、blastp 及び blastn、blastx、tblastn、tblastx の 5 つのアプリケーションがある。	Altschul, S.F. 他 (1990) J. Mol. Biol. 215:403-410; Altschul, S.F. 他 (1997) Nucleic Acids Res. 25: 3389-3402.	ESTs: 確率値=1.0E-8 以下 完全長配列: 確率値=1.0E-10 以下
FASTA	Pearson 及び Lipman アルゴリズムは、問合わせの配列と同種の配列群との類似性を検索する。FASTA は、fasta 及び tfasta、fastx、tfastx、ssearch の少なくとも 5 つのアプリケーションを含む。	Pearson, W.R. and D.J. Lipman (1988) Proc. Natl. Acad. Sci. 85:2444-2448; Pearson, W.R. (1990) Methods Enzymol. 183: 63-98; Smith, T.F. and M. S. Waterman (1981) Adv. Appl. Math. 2:482-489.	ESTs: fasta E 値=1.06E-6 構築された ESTs: fasta 同一性=95%以上、一致長さ=200 塩基以上、fastx E 値=1.0E-8 以下 完全長配列: fastx スコア=100 以上
BLIMPS	BLOCKS IMPROVED Searcher は、BLOCKS 及び PRINTS、DOMO、PRODOM、PFAM データベースにおける配列に対して、ある配列の一致性を調べ、遺伝子ファミリー及び配列同一性、構造的フィンガープリント領域を探索する。	Henikoff, S. and J.G. Henikoff, Nucl. Acid Res., 19:6565-72, 1991. J.C., Henikoff and S. Henikoff (1996) Methods Enzymol. 266:88-105; Artwood, T.K. 他 (1997) J. Chem. Inf. Comput. Sci. 37: 417-424.	スコア=1000 以上 スコア/強度の比率=0.75 以上 該当する場合は、確率値=1.0E-3 以下
HMMER	PFAM などのタンパク質ファミリーコンセンサス配列の隠れマルコフモデル(HMM)に基づいたデータベースに対して問合せ配列を検索するアルゴリズム。	Krogh, A. 他 (1994) J. Mol. Biol. 235:1501-1531; Sonnhammer, E.L.L. 他 (1998) Nucleic Acids Res. 26:320-322. Durbin, R. 他 (1998) Our World View, in a Nutshell, Cambridge Univ. Press, pp. 1-350.	PFAM ヒットの場、スコア=10-50 ビット、タンパク質ファミリーによってそれぞれ異なる

10

20

30

表1-2

プログラム名	説明	引用文献	パラメーター閾値
ProfileScan	Prosite で定義された配列パターンと一致するタンパク質配列における構造及び配列のモチーフを検索するアルゴリズム。	Gribskov, M. 他 (1988) CABIOS 4: 61-66; Gribskov, 他 (1989) Methods Enzymol. 183: 146-159; Bairoch, A. 他 (1997) Nucleic Acids Res. 25: 217-221	標準化された質のスコアを特定 の Prosite モチーフに対する GCG 指定 "HIGH" 値 通常、スコア=1.4-2.1
Phred	高い感度及び確率で自動配列決定機のトレースを調べる塩基読み出しアルゴリズムである。	Ewing, B. 他 (1998) Genome Res. 8: 175-185; Ewing, B. and P. Green (1998) Genome Res. 8: 186-194.	
Phrap	Smith-Waterman アルゴリズムの効率的なインプリメンテーションに基づく SWAT や CrossMatch を含む Phrap Revised Assembly プログラムであって、配列相同性の検索及び DNA 配列の構築に有用である。	Smith, T.F. and M. S. Waterman (1981) Adv. Appl. Math. 2: 482-489; Smith, T.F. and M.S. Waterman (1981) J. Mol. Biol. 147: 195-197; Green, P., University of Washington, Seattle, WA.	スコア=120 以上 一致長さ=56 以上
Consed	Phrap で構築したものの表示及び編集をするためのグラフィックツールである。	Gordon, D. 他 (1998) Genome Res. 8: 195-202.	
SPScan	タンパク質配列をスキャンして分岐シグナルペプチドの存在を調べる重み付けマトリクス解析プログラムである。	Nielson, H. 他 (1997) Protein Engineering 10: 1-6; Claverie, J.M. and S. Audic (1997) CABIOS 12: 431-439.	スコア=3.5 以上
Motifs	Prosite で定義された配列と一致したパターンについてアミノ酸配列を検索するプログラムである。	Bairoch 他、前出; Wisconsin Package Program Manual, version 9, page M51-59, Genetics Computer Group, Madison, WI.	

10

20

30



【 1 E 】

1306 1315 1324 1333 1342 1351  
 TCG GAA CAG GCG ACC AGG GCG GCC CCT TTG CAG TAA GCT GTC CCA CCG ACA GAA  
 W E Q G S R A A P L Q E A V P R T Q  
 1360 1369 1378 1387 1386 1405  
 GCC ACG GGC CTC ACC AAG CTT ACC CTA CCT CCG AGT CCC CCG ATG CCA CCC AGA  
 A T G L T K P T L P S F L M A A R  
 1414 1423 1432 1441 1450 1459  
 AGA COT CTC TAG TGT TGC GAA CAC TCT GTA TAT TTC GAG CTA CCA CCC ACA CCT  
 R R L  
 1468 1477 1486 1495 1504 1513  
 GTG TGT GCA CTT GTA TGT TTT ATA AAC TTG CCA TCT GTA AAA RIR ACC AOC GAT  
 1522 1531 1540 1549  
 TAG ATG AAT TGA GAA CTT TCT AAT GAA AAA AAA AAA 3'

【 2 A 】

11 20 29 38 47 56  
 5' AGA CCC GGC ACG CAG GTG GGG GCC GGC GGG GTC GGT GGA CCA CAG CAG CAG AGA  
 65 74 83 92 101 110  
 GTC AAG CCG GCG GAG GCT GCT GTC GGT GCA GTG AGG AAG AGG CCC TGG GTG  
 119 128 137 146 155 164  
 GTC CCC AAG GCT GGC CAG GAT CCA GCG CAG AAT CAC CCG TTC TAC TAC  
 M A G Q D P A L S T S H P F Y D  
 173 182 191 200 209 218  
 GTC GCC AAG CAT GGC AAT CAG GTG GCA CCG GGT GAC GGC CTT GSA AGA CTT  
 V A K H G I L Q V A G D D R F C R  
 227 236 245 254 263 272  
 GTT GTC ACG TTC ACG TGC TGC CCA CCC TCC CAC CAG CAG CAG GAC GAC  
 V V T F S C C R H F S H E L D H  
 281 290 299 308 317 326  
 CCG CAG CAG TAT TAT AAG TAC ACA CTG GAC CAA TAC GTT GAG AAC GAT TAT  
 R L L E Y L K Y T L D Q Y V E N D Y

【 2 B 】

344 353 362 371 380  
 ACC ATC CTC TAT TTC CAC TAC GAG CAG AAC AGC CAG CCA AAG CCA TTC CTC GGC  
 T I V Y F H Y G L N S R N K P S L C  
 388 398 407 416 425 434  
 GCG CTC CAG ACC GCA TAC AAG CAG TTC GAT AAG AAG TAC AAG AAC TCG AAG  
 W L Q S A Y K E F D R K Y R K N L R  
 443 452 461 470 479 488  
 GCC CTC TAC GTC GTC CCG ACC AGC TTC ATC AAG GTC CAG CAG TCG AAC ATC TTG  
 A L Y V V H P T S F I K V L W N I L  
 497 506 515 524 533 542  
 AAG CCC CTC ATC ATC CAC AAG TTT GGG AAG AAA GTC ATC TAT TTC AAC TAC CTS  
 K P L I S H K F G K K V I Y F N Y L  
 551 560 569 578 587 596  
 AGT GAG CTC CAC GAA CAC CTA AAA TAC GAC CAG CTS GTC ATC CCA CCA GAA GTT  
 S E L H E H L K Y D Q L V I P P E V  
 605 614 623 632 641 650  
 TTG CCG TAC GAT AAG CAG CTC CAG ACC CAG CAC GAG GGC CCG ACG CCG CCA CCG  
 L R Y D E K L Q S L H E G R T P V P

【 2 C 】

658 668 677 686 695 704  
 ALC AAG ACA CCA CCG CCG CCG ACC ACG CAG CCG CCG CCG CCG TTT GGC GTC AAT  
 T K T P P R P L P T Q Q F G V S  
 713 722 731 740 749 758  
 CTS CAA TAC CTC AAA GAC AAA AAT CAA GGC GAA CTC ATC CCC CCA CCG CTS ACG  
 L Q Y L E K D K N Q G E L I P F V L R  
 767 776 785 794 803 812  
 TTC ACA GTC ACG TAC CTS AGA GAG AAA GGC CTC CCG ACC GAG GCG CTS TTC GAG  
 F T V T Y L R E K C L R T E G L F R  
 821 830 839 848 857 866  
 AKA TCC GCC AAG GTC ACC GTC CCG GAG ATC CAG AAG CTC TAC AAC CAA GAG  
 K S A S V Q T V R E I Q H I Y N Q  
 875 884 893 902 911 920  
 AAG CCC GTC AAG TTT GAC TAC CAG GAG AAT CAC ATC CCA CCG GTC ATC CTS  
 K P V N F D D Y G D I H I F A V I L  
 929 938 947 956 965 974  
 AAG ACC TTC CCA GAG CTS CCG CCG CCG CCG CCG CCG CCG CCG CCG CCG CCG  
 K T F L R E L P Q P L L T F Q A Y E

【 2 D 】

983 992 1001 1010 1019 1028  
 CRG AIT CTC GGG AYC ACC TGT GTG GAG AGC AGC CTC GGC ACT CCC TTC CUC  
 Q I H G I T C V E S S L R V T R C R  
 1037 1046 1055 1064 1073 1082  
 CAG AYC TTA GGG AGC CTC CCA GAG CAC AAC TAC GTC CTC GGC TAC CTC ATC  
 Q I L R S L P E H N Y V V L R Y L W  
 1091 1100 1109 1118 1127 1136  
 GCC TTC CGT GGT GCG TCC GGG GAG AGC ATC TTC AAC AAA ATG AAC AGC TCT  
 G F L H A V S R E S I F N K K N S S  
 1145 1154 1163 1172 1181 1190  
 AAC CTC GGC TGT CTC TTC GGG CTC AAT TTC ATC TGG TGG CCA TTC CAG GGG CTC ATC  
 N L A C V F G L N L I W P S Q G V S  
 1199 1208 1217 1226 1235 1244  
 TCC CTC AGT GGC CTC GTC CCC CTC AAC ATG TTC ACT GAA CTC GTC ATC GAG TAC  
 S L S A L V P L N H F T E L L L E Y  
 1253 1262 1271 1280 1289 1298  
 TAT GAA AAG AYC AYC TTC ACC ACC CGG GAG GCA CTC GGG GAG CAC GGC CTG GCA CCA  
 Y E K I F S T F S A P G E H G L A P

【 2 E 】

1307 1316 1325 1334 1343 1352  
 TCC GAA CAG GGG AGC AGC GCA GGC CCT TTG CAG GAG GCT GTG CCA CGG ACA CAA  
 W B Q G S R A A P L Q E A V P R T Q  
 1361 1370 1379 1388 1397 1406  
 GCC ACG GGC CTC ACC AAG CCT ACC CTA CCT CCG AGT CUC CAG ATC GCA GGC AAA  
 A T G L V K P T L P P S P L M A A R  
 1415 1424 1433 1442 1451 1460  
 AGA CCT CTC TGG TGT TGC GAA CAC TCT GTA TGT TTC GAG CTA CCT CCC ACA CCT  
 R R L  
 1469 1478 1487 1496 1505 1514  
 GTC TGT SCA CTT GTA TGT TTT GTA AAC TTG GCA TCT CTA AAA ATA ACC AGC CAT  
 1523 1532  
 TGG AYC AAT TCA GAA CCT TCT AAT G 3'

【 3 A 】

1 M A 404424  
 1 M A 3068538  
 1 M D P I S E L Q D D L T I D D T S R A L N Q L K L A S I D E 9312212  
 3 404424  
 3 3068538  
 1 96572185  
 31 K N W P S D E M P D P P K S D D S K S S P E L V H L K W 9312212  
 H S H P F Y D V A R H G I L Q V A G D D R F G R R V V T F S C 404424  
 11 S H P F Y D V A R H G I L Q V A G D D R F G R R V V T F S C 3068538  
 1 96572185  
 61 D D P V Y D I A R H Q I V E V A G D D K Y G R K I I V F S I A 9312212  
 41 C R M P P S H E L D H Q R L L E Y L K Y T L D Q Y V E N D Y 404424  
 41 C R M P P S H E L D H Q R L L E Y L K Y T L D Q Y V E N D Y 3068538  
 1 96572185  
 91 C R M P P S H Q L D H S K L G Y L K R H T L D Q Y V E N D Y 9312212  
 71 T I V Y F H Y G L N S R N K P S L G W L Q S A Y K E F D R K 404424  
 71 T I V Y F H Y G L N S R N K P S L G W L Q S A Y K E F D R K 3068538  
 1 96572185  
 121 T L L Y L H I G L T S I D N K P S L S W L R D A Y R E F D R K 9312212

【 3 B 】

101 Y K K N L K A L Y V V H P T S F I K V L W N I L K P L I S I I 404424  
 101 Y K K N I K A L Y V V H P T S F I K V L W N I L K P L I S I H 3068538  
 1 Y K K N L K A L Y V V H P T S F I K V L W N I L K P L I S I H 96572185  
 151 Y K K N I K A L Y I V H P T M P I K T L L I L F K P E L I S I P 9312212  
 131 K F G K K V I Y F N Y L S E L H E H L K Y D Q L V I P P E V 404424  
 131 K F G K K V I Y F N Y L S E L H E H L K Y D Q L V I P P E V 3068538  
 31 K F G K K V I Y F N Y L S E L H E H L K Y D Q L V I P P E V 96572185  
 181 K F S Q K I F V I V N Y L S E L S E H V K L E Q L G L P R Q V 9312212  
 161 L R Y D E K L Q S I H E G R T P P P T K T P P P R P L P T 404424  
 161 L R Y D E K L Q S L H E G R T P P P T K T P P P R P L P T 3068538  
 61 L R Y D E K L Q S L H E G R T P P T K T P P P R P L P T 96572185  
 211 L K Y P D F L K S T Q K S P A T A E - K P M P P R P P L L E N 9312212  
 191 Q Q F G V S L Q Y L K D K N - Q G E L I P P V L R F T V T Y 404424  
 191 Q Q F G V S L Q Y L K D K N - Q G E L I P P V L R F T V T Y 3068538  
 91 Q Q F G V S L Q Y L K D K N - Q G E L I P P V L R F T V T Y 96572185  
 240 Q Q F G V S L Q H I Q E K N P E Q H P I P I V L R N T V A Y 9312212  
 220 L R E K G L R T E G L F R R S A S V Q T V R F I Q R I Y N Q 404424  
 220 L R E K G L R T E G L F R R S A S V Q T V R F I Q R I Y N Q 3068538  
 120 L R E K G L R T E G L F R R S A S V Q T V R F I Q R I Y N Q 96572185  
 270 L Q A H A L T E G I F R R S A N T Q I V R F I Q K Y N M 9312212





【 5 E 】

3068538CE1                   atttgatctggmatalccougggggtctctccctgglygcctttklygcccccgaalgttccctcaact  
70157574HL                   .....  
700480528HL                   .....  
700935753HL                   acttcaatctggccatccccaaggggtgcttccctcagggcccttgggtctctctgaacttcttcaagaact  
700935061HL                   aattgatctggccatccccaaggggtgcttccctcagggcccttgggtctctctgaacttcttcaagaagct  
1270                   1280                   1290                   1300                   1310                   1320                   1330  
3068538CE1                   gctgatacgaatcactatga\*aaagatcttccagcaccgccagcactggggagcagcagcctgcaaccatg  
70157574HL                   .....  
700480528HL                   .....  
700935753HL                   gctgatacgaatcactatgaagaatcttctccagtcaccaagggccctggaggagcactccgggtatctctg  
700935061HL                   gctgatacgaatcactatgaagaatcttctccagtcaccaagggccctggaggagcactccgggtatctctg  
1340                   1350                   1360                   1370                   1380                   1390                   1400  
3068538CE1                   ggagcagggagcagggagcagcccttggcagggtgtrnccagcagcagcagcagcagcagcagcagcag  
70157574HL                   .....  
700480528HL                   .....  
700935753HL                   cgaacagcaacagactgggtctctgttaccosaaagattccacacacagcagcagcagcagcagcagcag  
700935061HL                   cgaacagcaacagactgggtctctgttaccosaaagattccacacacagcagcagcagcagcagcagcag  
1410                   1420                   1430                   1440                   1450                   1460                   1470

【国際公開パンフレット】

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization  
International Bureau



(43) International Publication Date  
23 August 2001 (23.08.2001)

PCT

(10) International Publication Number  
WO 01/61010 A2

(51) International Patent Classification: C12N 15/12,  
C07K 14/47, C12Q 1/68, G01N 33/68, A61K 38/00,  
C07K 16/18, C12N 15/11, A01K 67/027

STEWART, Elizabeth, A. [US/US]; 1678 Latham Street,  
Mountain View, CA 94041 (US); YUE, Henry [US/US];  
826 Lois Avenue, Sunnyvale, CA 94087 (US); BAUGHN,  
Mariah, R. [US/US]; 14244 Santiago Road, San Leandro,  
CA 94577 (US).

(21) International Application Number: PCT/US01/05075

(22) International Filing Date: 15 February 2001 (15.02.2001)

(74) Agents: TURNER, Christopher et al.; Incyte Genomics,  
Inc., 3160 Porter Drive, Palo Alto, CA 94304 (US).

(25) Filing Language: English

(81) Designated States (national): AF, AG, AI, AM, AT, AU,  
AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ,  
DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR,  
HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR,  
LS, LU, LV, MA, MD, MG, MK, MN, MW, MX, MY,  
NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM,  
TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(26) Publication Language: English

(30) Priority Data:  
09/507,765 18 February 2000 (18.02.2000) US

(84) Designated States (regional): ARIPO patent (GH, GM,  
KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW); Eurasian  
patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM); European  
patent (AT, BE, CH, CY, CZ, DE, DK, ES, FI, FR, GB, GR, IE,  
IT, LI, LU, MC, NL, PT, SE, TR); OAPI patent (BF, BJ, CF,  
CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

(71) Applicant (for all designated States except US): INCYTE  
GENOMICS, INC. [US/US]; 3160 Porter Drive, Palo  
Alto, CA 94304 (US).

(72) Inventors and  
(75) Inventors/Applicants (for US only): KLINGLER, Tod,  
M. [US/US]; 28 Dover Court, San Carlos, CA 94070 (US).

[Continued on next page]

(54) Title: GTPASE ACTIVATING PROTEINS

	10	19	28	37	46	55
5'	CAG ACC CCG CAC GCA GGT GGG GGC CGG CGG GGT CCG TGG CCA GAG CTG CAG AGA					
	64	73	82	91	100	109
	GAC AAG GCG GCG GCG GCT GCT GTG CTG CCG GCA GTG AGG AAG ACG CCC TCG GTG					
	118	127	136	145	154	163
	GTG CCC ATG GCT GGC CAG GAT CCT GCG CTG AGC ACG AGT CAC CCG TTC TAC GAC					
	M A G Q D P A L S T S H P F Y D					
	172	181	190	199	208	217
	GTG GCC AGA CAT GGC APT CTG CAG GTG GCA GGG GAT GAC CGC TTT GGA AGA CGT					
	V A R H G I L Q V A G D D R F G R R					
	226	235	244	253	262	271
	GTT CTC AGC TTC AGC TGC CCG ATG CCA CCC TCC CAC GAG CTG GAC CAG CAG					
	V V T F S C C R M F P S H E I D H Q					
	280	289	298	307	316	325
	CGG CTG CTG GAG TAT TTG AAG TAC ACA CTG CAC CAA TAC GGT GAG AAC GAT TAT					
	R L L E Y L K Y T L D Q Y V E N D Y					

(57) Abstract: The invention provides mammalian nucleic acid molecules and fragments thereof. It also provides for the use of the mammalian nucleic acid molecules for the characterization, diagnosis, evaluation, treatment, or prevention of conditions, diseases and disorders associated with cell signaling, the immune system, and cell proliferation, particularly colon cancer. The invention additionally provides expression vectors and host cells for the production of the protein encoded by the mammalian nucleic acid molecules.



WO 01/61010 A2

**WO 01/61010 A2** 

**Published:**

— without international search report and to be republished upon receipt of that report

*For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.*

WO 01/61010

PCT/US01/05075

**GTPASE ACTIVATING PROTEINS****FIELD OF THE INVENTION**

This invention relates to nucleic acid molecules and the encoded GTPase activating proteins, and variants thereof, and to the use of these molecules in the characterization, diagnosis, prevention, and treatment of cell signaling, immune, and cell proliferative disorders, particularly colon cancer.

**BACKGROUND OF THE INVENTION**

Guanine nucleotide binding proteins (GTP-binding proteins) are present in all eukaryotic cells and function in processes including metabolism, cellular growth, differentiation, signal transduction, cytoskeletal organization, and intracellular vesicle transport and secretion. In higher organisms they are involved in signaling that regulates such processes as the immune response (Ausset et al (1988) *J Immunol* 140:215-220), apoptosis, differentiation, and cell proliferation including oncogenesis (Dhanasekaran et al. (1998) *Oncogene* 17:1383-1394).

The superfamily of GTP-binding proteins consists of several families and may be grouped as translational factors, heterotrimeric GTP-binding proteins involved in transmembrane signaling processes (also called G-proteins), proto-oncogene Ras proteins, and other low molecular weight GTP-binding proteins including the products of *rab*, *rap*, *rho*, *rac*, *smg21*, *smg25*, *YPI*, *SEC4*, and *ARF* genes, and tubulin (Kaziro et al. (1991) *Ann Rev Biochem* 60:349-400).

The low molecular weight (LMW) GTP-binding proteins are a class of small proteins of 21-30 kDa. These proteins regulate cell growth, cell cycle control, protein secretion, and intracellular vesicle interaction. In particular, LMW GTP-binding proteins activate cellular proteins by transducing mitogenic signals involved in various cell functions in response to extracellular signals from receptors (Tavittian (1995) *C R Seances Soc Biol Fil* 189:7-12). During this process, the hydrolysis of GTP acts as an energy source as well as the means of converting the GTP-binding protein from an active (GTP-bound) to an inactive (GDP-bound) form.

The Rho family of LMW GTP-binding proteins ( Rho GTPases) includes RhoA, Rac1, and Cdc42 which regulate a variety of biological events related to actin-cytoskeletal reorganization and cell proliferation. A key group of regulatory molecules for the Rho GTPases is the Rho GTPase-activating proteins (GAPs). Rho GAPs preferentially recognize the GTP-bound form of a Rho GTPase and stimulate the intrinsic GTPase activity to hydrolyze the bound GTP to GDP. Rho GAPs therefore function as negative regulators or suppressors of Rho GTPase by stimulating the conversion of the Rho GTPase from the active GTP-bound form to the inactive GDP-bound form (Zhang, et al. (1997) *J Biol Chem* 272:21999-22007).

Rho GAP proteins share an approximately 170-190 amino acid homology region, designated as the Rho GAP domain, that appears to contain the minimum structural domain necessary for GAP activity (Zheng, et al. (1993) *J Biol Chem* 268:24629-24634). Rho GAP proteins share 20-24% amino

WO 01/61010

PCT/US01/05075

acid identity in this domain, however, certain specific residues are highly conserved. For example, a pair of arginine residues located near the N-terminus of the Rho GAP domain appear to be highly conserved among Rho GAP proteins and are necessary for maximum catalysis (Leonard et al. (1998) *J Biol Chem* 273:16210-16215). In addition, a proline-rich, SH3 domain-binding site is also found in the N-terminal region of these proteins (Leonard, *supra*).

The identification of genes associated with cancer and understanding the genetic mechanisms underlying carcinogenesis are critical to the diagnosis, prevention, and treatment of these diseases. In colon cancer particularly, it is known that a combination of activation of oncogenes, inactivation of tumor suppressor genes, and alteration of DNA mismatch repair genes is involved in the progression from normal mucosa to colon cancer (Fearon et al. (1990) *Cell* 61:759-767; Chung (1995) *Gastroenterology* 109:1685-1699). Recently a study was conducted which identified a specific region of chromosome 22 associated with colon cancer based on chromosomal (gene?) allelic losses in colon tumor samples relative to normal colon mucosa (Castells et al. (1999) *Gastroenterology* 117:8331-837). Using microsatellite markers from chromosome 22, a minimal region of allelic deletion was identified between markers D22S1171 and D22S928 covering an interval of 0.57 cM and corresponding to the cytogenetic location 22q13.33. Various genes on chromosome 22 have been proposed as candidate tumor-suppressor genes associated with colorectal carcinogenesis, however, no mutations have been found in any of these genes (Castells et al. *supra*).

The discovery of nucleic acid sequences encoding GTPase-activating proteins and variants thereof provides new compositions that are useful in the characterization, diagnosis, prevention, and treatment of cell proliferative disorders, including cancer and, in particular, colon cancer.

#### SUMMARY OF THE INVENTION

The invention is based on the discovery of nucleic acid molecules encoding GTPase-activating protein referred to collectively as "GTPAP" and individually as "GTPAP1" and "GTPAP2" and variants thereof, which satisfies a need in the art by providing compositions useful in the characterization, diagnosis, prevention, and treatment of conditions such as cell signaling, immune, and cell proliferative disorders, particularly colon cancer.

The invention provides isolated and purified human and rat nucleic acid molecules comprising SEQ ID NOs:1-29, and fragments thereof, encoding the mammalian protein comprising the amino acid sequence of SEQ ID NO:30, or portions thereof, a biologically active portion of SEQ ID NO:30, an immunologically active portion of SEQ ID NO:30, and a variant of SEQ ID NO:30 comprising SEQ ID NO:31.

The invention further provides a probe that hybridizes to the mammalian nucleic acid molecules or fragments thereof. The invention also provides isolated and purified nucleic acid molecules that are complementary to the nucleic acid molecules of SEQ ID NOs:1-29. In one aspect,

WO 01/61010

PCT/US01/05075

the probe is a single stranded complementary RNA or DNA molecule.

The invention further provides a method for detecting a nucleic acid molecule in a sample, the method comprising the steps of hybridizing a probe to at least one nucleic acid molecule of a sample, forming a hybridization complex; and detecting the hybridization complex, wherein the presence of the  
5 hybridization complex indicates the presence of the nucleic acid molecule in the sample. In one aspect, the method further comprises amplifying the nucleic acid molecule prior to hybridization. The nucleic acid molecule or a fragment thereof may comprise either an element or a target on a microarray.

The invention also provides a method for using a nucleic acid molecule or a fragment thereof to screen a library of molecules to identify at least one ligand that specifically binds the nucleic acid  
10 molecule, the method comprising combining the nucleic acid molecule with a library of molecules under conditions allowing specific binding, and detecting specific binding, thereby identifying a ligand that specifically binds the nucleic acid molecule. Such libraries include DNA and RNA molecules, peptides, PNAs, proteins, and the like. In an analogous method, the nucleic acid molecule or a fragment thereof is used to purify a ligand.

The invention also provides an expression vector containing at least a fragment of the nucleic acid molecule. In another aspect, the expression vector is contained within a host cell. The invention further provides a method for producing a protein, the method comprising the steps of culturing the host  
15 cell under conditions for the expression of the protein and recovering the protein from the host cell culture.

The invention also provides substantially purified mammalian GTPAP or a portion thereof.  
20 The invention further provides isolated and purified proteins having the amino acid sequence of SEQ ID NO:30, a biologically active portion of SEQ ID NO:30, and an immunologically active portion of SEQ ID NO:30, and a variant of SEQ ID NO:30 comprising SEQ ID NO:31. Additionally, the invention provides a pharmaceutical composition comprising a substantially purified mammalian GTPAP protein  
25 or a portion thereof in conjunction with a pharmaceutical carrier.

The invention also provides a method for treating a disease or condition associated with altered expression of GTPAP, comprising administering to a patient in need of such treatment an effective amount of a pharmaceutical composition comprising a substantially purified mammalian GTPAP protein or a portion thereof in conjunction with a pharmaceutical carrier. In one embodiment of the  
30 invention, the disease or condition is selected from cell signaling, immune, and cell proliferative disorders, particularly colon cancer. In another embodiment of the invention, the cancer is a colon cancer.

The invention further provides a method for using at least a portion of the mammalian protein to produce antibodies. The invention also provides a method for using a mammalian protein or a

WO 01/61010

PCT/US01/05075

portion thereof to screen a library of molecules to identify at least one ligand that specifically binds the protein, the method comprising combining the protein with the library of molecules under conditions allowing specific binding, and detecting specific binding, thereby identifying a ligand that specifically binds the protein. Such libraries include DNA and RNA molecules, peptides, agonists, antagonists, antibodies, immunoglobulins, drug compounds, pharmaceutical agents, and other ligands. In one aspect, the ligand identified using the method modulates the activity of the mammalian protein. In an analogous method, the protein or a portion thereof is used to purify a ligand. The method involves combining the mammalian protein or a portion thereof with a sample under conditions to allow specific binding, detecting specific binding between the protein and ligand, recovering the bound protein, and separating the protein from the ligand to obtain purified ligand.

The invention further provides a method for inserting a marker gene into the genomic DNA of a mammal to disrupt the expression of the natural mammalian nucleic acid molecule. The invention also provides a method for using the mammalian nucleic acid molecule to produce a mammalian model system, the method comprising constructing a vector containing the mammalian nucleic acid molecule; introducing the vector into a totipotent mammalian embryonic stem cell; selecting an embryonic stem cell with the vector integrated into genomic DNA; microinjecting the selected cell into a mammalian blastocyst, thereby forming a chimeric blastocyst; transferring the chimeric blastocyst into a pseudopregnant dam, wherein the dam gives birth to a chimeric mammal containing at least one additional copy of mammalian nucleic acid molecule in its germ line; and breeding the chimeric mammal to generate a homozygous mammalian model system.

#### BRIEF DESCRIPTION OF THE FIGURES AND TABLE

Table 1 shows the tools, programs, and algorithms used to analyze GTPAP, along with applicable descriptions, references, and threshold parameters.

Figures 1A, 1B, 1C, 1D, and 1E show the human nucleic acid molecule (SEQ ID NO:28) encoding the human amino acid sequence (SEQ ID NO:30) of the mammalian protein. The alignment was produced using MACDNASIS PRO software (Hitachi Software Engineering, South San Francisco CA).

Figures 2A, 2B, 2C, 2D, and 2E show the human nucleic acid molecule (SEQ ID NO:29) encoding the human amino acid sequence (SEQ ID NO:31) of the mammalian protein. The alignment was produced using MACDNASIS PRO software (Hitachi Software Engineering)

Figures 3A, 3B, 3C, and 3D demonstrate the chemical and structural similarity among human RhoGAP proteins, GTPAP2 (clone ID 404424;SEQ ID NO:31), GTPAP1 (clone ID 3068538;SEQ ID NO:30), Rho GTPase activating protein 8 (GI 6572185;SEQ ID NO:32), and rhoGAP protein (GI 312212;SEQ ID NO:33). The alignments were produced using the MEGALIGN program

WO 01/61010

PCT/US01/05075

(DNASTAR, Madison WI)

Figures 4A and 4B demonstrate the alignments among nucleic acids encoding human GTPAP and indicating regions of single nucleotide polymorphisms (SNPs). SNPs are indicated by (A). The alignments were produced using PHRAP software (Phil Green, University of Washington, Seattle WA).

5 Figures 5A, 5B, 5C, 5D, and 5E demonstrate the chemical and structural similarity among human, 3068538CB1 (SEQ ID NO:28); and rat, 701244926H1 (SEQ ID NO:21), 700950169H2 (SEQ ID NO:22), 70157597H1 (SEQ ID NO:23), 701274036H1 (SEQ ID NO:24), 700480528 (SEQ ID NO:25), 700935753H1 (SEQ ID NO:26), and 700936061H1 (SEQ ID NO:27) nucleic acid sequences, produced using FLIRAP software (Phil Green, University of Washington, Seattle WA).

10

#### DESCRIPTION OF THE INVENTION

It is understood that this invention is not limited to the particular machines, materials and methods described. It is also to be understood that the terminology used herein is for the purpose of describing particular embodiments only and is not intended to limit the scope of the present invention which will be limited only by the appended claims. As used herein, the singular forms "a," "an," and "the" include plural reference unless the context clearly dictates otherwise. For example, a reference to "a host cell" includes a plurality of such host cells known to those skilled in the art.

Unless defined otherwise, all technical and scientific terms used herein have the same meanings as commonly understood by one of ordinary skill in the art to which this invention belongs. All publications mentioned herein are cited for the purpose of describing and disclosing the cell lines, protocols, reagents and vectors which are reported in the publications and which might be used in connection with the invention. Nothing herein is to be construed as an admission that the invention is not entitled to antedate such disclosure by virtue of prior invention.

#### Definitions

25 "GTPase-activating protein" or "GTPAP", refers to a substantially purified protein obtained from any mammalian species, including murine, bovine, ovine, porcine, rodent, canine, simian, and preferably the human species, and from any source, whether natural, synthetic, semi-synthetic, or recombinant. "GTPAPs" refers to more than one GTPAP.

"Altered gene expression" and "altered expression" refers to the increase or decrease of gene expression and the presence or absence of transcribed messenger RNA transcribed in a sample.

30 "Biologically active" refers to a protein having structural, immunological, regulatory, or chemical functions of a naturally occurring, recombinant or synthetic molecule. Likewise, "immunologically active" refers to the capability of the natural, recombinant, or synthetic GTPAP, or of any oligopeptide thereof, to induce a specific immune response in appropriate animals or cells and to

WO 01/61010

PCT/US01/05075

bind with specific antibodies.

"Complementary" refer to the natural hydrogen bonding by base pairing between purines and pyrimidines. For example, the sequence A-C-G-T forms hydrogen bonds with its complements T-G-C-A or U-G-C-A. Two single-stranded molecules may be considered partially complementary, if only some of the nucleotides bond, or completely complementary, if nearly all of the nucleotides bond. The degree of complementarity between nucleic acid strands affects the efficiency and strength of the hybridization and amplification reactions.

"Derivative" refers to the chemical modification of a nucleic acid molecule or amino acid sequence. Chemical modifications can include replacement of hydrogen by an alkyl, acyl, or amino group or glycosylation, pegylation, or any similar process that retains or enhances biological activity or lifespan of the molecule or sequence.

"Fragment" refers to an invertebrate clone or any part of a nucleic acid molecule that retains a usable, functional characteristic. Useful fragments include oligonucleotides which may be used in hybridization or amplification technologies or in regulation of replication, transcription or translation.

"Hybridization complex" refers to a complex between two nucleic acid molecules by virtue of the formation of hydrogen bonds between purines and pyrimidines.

"Ligand" refers to any molecule, agent, or compound that will bind specifically to a complementary site on a nucleic acid molecule or protein. Such ligands stabilize or modulate the activity of nucleic acid molecules or proteins of the invention and may be composed of at least one of the following: inorganic and organic substances including nucleic acids, proteins, carbohydrates, fats, and lipids.

"Modulates" refers to a change in activity (biological, chemical, or immunological) or lifespan resulting from specific binding between a molecule and either a nucleic acid molecule or a protein.

"Non-conservative amino acid substitutions" are those substitutions that, when made, are most likely to effect the properties of the original protein, i.e., the structure and especially the function of the protein. For example, the substitution of an acidic or basic amino acid, such as aspartate or arginine, for a neutral amino acid, such as glycine or alanine represents a non-conservative substitution. Like wise, "conservative amino acid substitutions" are those substitutions that, when made, least interfere with the properties of the original protein, i.e., the structure and especially the function of the protein is conserved and not significantly changed by such substitutions. The substitution of one neutral amino acid for another, such as valine for alanine is a conservative substitution, as well as the substitution of one acid amino acid for another, such as aspartate for glutamate.

"Non-conservative nucleic acid substitution" refers to the substitution of a pyrimidine

WO 01/61010

PCT/US01/05075

nucleotide for a purine nucleotide in a nucleic acid molecule. Likewise a "conservative nucleic acid substitution" refers to the substitution of one pyrimidine or one purine nucleotide for another, such as C for T or A for G.

"Nucleic acid molecule" refers to a nucleic acid, oligonucleotide, nucleotide, polynucleotide or any fragment thereof. It may be DNA or RNA of genomic or synthetic origin, double-stranded or single-stranded, and combined with carbohydrate, lipids, protein or other materials to perform a particular activity such as transformation or form a useful composition such as a peptide nucleic acid (PNA). "Oligonucleotide" is substantially equivalent to the terms amplicon, primer, oligomer, element, target, and probe and is preferably single stranded. The nucleic acid molecules may be splice variants of another nucleic acid molecule.

"Protein" refers to an amino acid sequence, oligopeptide, peptide, polypeptide or portions thereof whether naturally occurring or synthetic.

"Portion", as used herein, refers to any part of a protein used for any purpose, but especially for the screening of a library of molecules that specifically bind to that portion, or for the production of antibodies. An exemplary portion is the first 40 contiguous amino acids of SEQ ID NO:46.

"Similarity" as applied to polynucleotide sequences, refers to the quantified residue matches between at least two polynucleotide sequences aligned using a standardized algorithm. Such an algorithm may insert, in a standardized and reproducible way, gaps in the sequences being compared in order to optimize alignment between two sequences, and therefore achieve a more meaningful comparison of the two sequences.

A suite of commonly used and freely available sequence comparison algorithms is provided by the National Center for Biotechnology Information (NCBI), Bethesda MD. BLAST, Basic Local Alignment Search Tool (Altschul et al., (1990) J. Mol. Biol. 215:403-410; Altschul (1993) J. Mol. Evol. 36:290-300), is available from NCBI (<http://www.ncbi.nlm.nih.gov/BLAST/>) and several other sources. The BLAST software suite includes various sequence analysis programs including "blastn", that is used to align a known polynucleotide sequence with other polynucleotide sequences from a variety of databases. Also available is a tool called "BLAST 2 sequences" that is used for direct pairwise comparison of two nucleotide sequences. "BLAST 2 sequences" can be accessed and used interactively at <http://www.ncbi.nlm.nih.gov/ncbi/blast2.html>. The "BLAST 2 sequences" tool can be used for both blastn and blastp. BLAST programs are commonly used with gap and other parameters set to default settings. For example, to compare two nucleotide sequences, one may use blastn with the "BLAST 2 Sequences" tool Version 2.0.9 (May-07-1999) set at default parameters. Such default parameters may be, for example: *Matrix: BLOSUM62; Reward for match: 1; Penalty for mismatch: -2; Open Gap: 5 and Extension Gap: 2 penalties; Gap x drop-off: 50; Expect: 10; Word Size: 11; and Filter: on.*

WO 01/61010

PCT/US01/05075

Similarity may be measured over the length of an entire defined sequence, for example, as defined by a particular SEQ ID number, or may be measured over a shorter length, for example, over the length of a fragment taken from a larger, defined sequence, for instance, a fragment of at least 20, at least 30, at least 40, at least 50, at least 70, at least 100, or at least 200 contiguous nucleotides.

5 Such lengths are exemplary only, and it is understood that any fragment length supported by the sequences shown herein, in the tables, figures, or Sequence Listing, may be used to describe a length over which percentage identity may be measured.

"Sample" is used herein in its broadest sense. A sample containing polynucleotides, polypeptides, antibodies and the like may comprise a bodily fluid; a soluble fraction of a cell  
10 preparation, or media in which cells were grown; a chromosome, an organelle, or membrane isolated or extracted from a cell; genomic DNA, RNA, or cDNA in solution or bound to a substrate; a cell; a tissue; a tissue print; a fingerprint, skin or hair; and the like.

"Specific binding" or "specifically binding" refers to the interaction between two molecules. In the case of a polynucleotide, specific binding may involve hydrogen bonding between sense and  
15 antisense strands or between one strand and a protein which affects its replication or transcription, intercalation of a molecule or compound into the major or minor groove of the DNA molecule, or interaction with at least one molecule which functions as a transcription factor, enhancer, repressor, and the like. In the case of a polypeptide, specific binding may involve interactions with polynucleotides, as described above or with molecules or compounds such as agonists, antibodies,  
20 antagonists, and the like. Specific binding is dependent upon the presence of structural features that allow appropriate chemical or molecular interactions between molecules.

"Single Nucleotide Polymorphism", or SNP, refers to a single nucleotide alteration in a polynucleotide that occurs as a result of a substitution, insertion or deletion. The substitution may be conservative or non-conservative, as defined above, and may or may not result in a change in an  
25 encoded amino acid residue. Such changes may predispose an individual to a specific disease or condition.

"Substantially purified" refers to nucleic acid molecules or proteins that are removed from their natural environment and are isolated or separated, and are at least about 60% free, preferably about 75% free, and most preferably about 90% free, from other components with which they are naturally  
30 associated.

"Substrate" refers to any rigid or semi-rigid support to which nucleic acid molecules or proteins are bound and includes membranes, filters, chips, slides, wafers, fibers, magnetic or nonmagnetic beads, gels, capillaries or other tubing, plates, polymers, and microparticles with a variety of surface forms including wells, trenches, pits, channels and pores.

WO 01/61010

PCT/US01/05075

"Variant" refers to nucleic acid molecules that are variants of a nucleic acid molecule that encodes GTPAP, or to the proteins encoded by these molecules. Such variants may be determined by BLAST score, wherein the score is at least 100, and most preferably at least 400.

#### THE INVENTION

5 The invention is based on the discovery of new mammalian nucleic acid molecules that encode GTPAPs and nucleic acid molecule variants thereof, and on the use of the nucleic acid molecules, or fragments thereof, and protein, or portions thereof, as compositions in the characterization, diagnosis, treatment, or prevention of conditions such as cell signaling, immune, and cell proliferative disorders, particularly colon cancer.

10 Nucleic acids encoding GTPAP1 of the present invention were first identified in Incyte Clone 3068538H1 from the uterine tissue library UTRSNOR01 using a computer search for amino acid sequence alignments. A consensus sequence, SEQ ID NO:28, was derived from Incyte Clones 908465R2;SEQ ID NO:1 (COLNNOT09), 957130R6;SEQ ID NO:2 (KIDNNOT05), 1580628H1;SEQ ID NO:3 (DUODNOT01), 2631247F6;SEQ ID NO:4 (COLNTUT15),  
15 3068538H1;SEQ ID NO:5 (UTRSNOR01), 3532286T6;SEQ ID NO:6 (KIDNNOT25), and 1301520F6;SEQ ID NO:7.

Nucleic acid molecules encoding the human GTPase-activating protein and nucleic acid molecule variants of the present invention were identified by comparing nucleic acid sequences encoding GTPAP1 and GTPAP2 of the present invention with nucleic acid sequences in the LIFESFQ  
20 and ZOOSEQ databases (Incyte Pharmaceuticals, Palo Alto CA).

In one embodiment, the invention encompasses a polypeptide comprising the amino acid sequence of SEQ ID NO:30, as shown in Figures 1A, 1B, 1C, 1D, and 1E. GTPAP1 is 433 amino acids in length and has a potential N-linked glycosylation site at N338. Potential protein kinase phosphorylation sites are found for casein kinase II at S169, T239, T292, S309, and S382, for protein  
25 kinase C at S129, T239, and S297, and for tyrosine kinase at Y60, Y101, and Y315. BLIMPS\_PPFAM analysis identified an RHG5 GTPase-activator protein signature from amino acid residues D260 through P276. BLIMPS\_PRODOME identified two protein GTPase-activator domains from P210 through A235 and from L310 through L350. BLAST\_PRODOME analysis also identified homology to an RHG5 RhoGAP protein from L8 through S169 with a probability score of  $p=4.3^{60}$ , and to a P50-  
30 RhoGAP from T175 through P387 with a probability score of  $p=7.9^{29}$ .

In another embodiment, the invention encompasses a polypeptide comprising the amino acid sequence of SEQ ID NO:31, as shown in 2A, 2B, 2C, 2D, and 2E. GTPAP2 is also 433 amino acids in length and shares the same glycosylation sites, phosphorylation sites, motifs and signature sequences as GTPAP1.

WO 01/61010

PCT/US01/05075

As shown in Figures 3A, 3B, 3C, and 3D, GTPAP1 shares sequence homology with GTPAP2 (Incye ID 404424.5:SEQ ID NO:31), a partial amino acid sequence of human RhoGAP protein 8 (GI 6572185:SEQ ID NO:32), and human RhoGAP protein GI 312212 (SEQ ID NO:33). In particular, GTPAP1 and GTPAP2 share 99.8% identity, differing in only one amino acid residue at position 302 in which an arginine residue in GTPAP2 is substituted for a glycine residue at that position in GTPAP1. This substitution results from a SNP in the codon for the amino acid at that position (see polynucleotide alignments in Figures 4A and 4B described below). GTPAP1 is identical to GI 6572185 from residues Y101 through L433 of GTPAP1, however GTPAP1 contains an additional 100 amino acid residues on the N-terminus of the protein. GTPAP1 also shares 49.2% identity with GI 312212. The four proteins share a proline-rich region from about amino acid residue P176 through P189 of GTPAP1, the SH3 domain-binding site. The four proteins also share substantial homology in the C-terminal region containing the Rho GAP domain. In particular, the two arginine residues associated with the catalytic activity in GAP proteins contained in the sequence F231RRS of GTPAP1, are conserved in all four proteins. It is also noteworthy that the single amino acid change at position 302 between GTPAP1 (glycine) and GTPAP2 (arginine) is a non-conservative change that occurs in the Rho GAP domain of the protein.

Figures 4A and 4B show the alignments between the polynucleotides encoding GTPAP1 (3068538:SEQ ID NO:28) and GTPAP2 (404424.5:SEQ ID NO:29), and Incyte cDNA clones from which these sequences were derived. In particular, the alignments encompass regions in which single nucleotide polymorphism (SNP) changes are found between the two full length polynucleotides. These SNPs are identified by ( $\Delta$ ) in Figures 4A and 4B. The cDNA clones encompassing these regions (and their associated libraries) are: 908465R2; SEQ ID NO:1 (COLNNOT09), 957130R6; SEQ ID NO:2 (KIDNNOT05), 1580628H1; SEQ ID NO:3 (DUODNOT01), 2631247F6; SEQ ID NO:4 (COLNTUT15), 1301520F6; SEQ ID NO:7 (BRSTNOT07), 2465422F6; SEQ ID NO:8 (THYRNOT08), 957130X313V1; SEQ ID NO:9 (KIDNNOT05), 1580545H1; SEQ ID NO:10 (DUODNOT01), 1891457H1; SEQ ID NO:11 (BLADTUT07), 4649657H1; SEQ ID NO:12 (PROSTUT04), 4002758H1; SEQ ID NO:13 (HNT2AZS07), 957130R1; SEQ ID NO:14 (KIDNNOT05), 2624365T6.comp; SEQ ID NO:15 (KERANOT02), 2044444H1; SEQ ID NO:16 (HIPONOT02), 3416883T6; SEQ ID NO:17 (PTHYNOT04), 1301520T6.comp; SEQ ID NO:18 (BRSTNOT07), 1301520H1; SEQ ID NO:19 (BRSTNOT07), and 1422635X305D1.comp; SEQ ID NO:20 (KIDNNOT09). Figure 4A demonstrates a SNP at nucleotide position 662 showing a non-conservative change from T to A in GTPAP2 relative to GTPAP1. This change is supported by 2 of 6 clones containing T in this position, and 4 of 6 clones containing A in this position. Figure 4B demonstrates SNPs at positions 1013 and 1020. The SNP at position 1013 shows a conservative change of T to C in GTPAP2 and is supported by 2 of 12 clones containing a C in this position

WO 01/61010

PCT/US01/05075

compared with 10 of 12 clones containing a T in this position. The SNP at position 1020 shows a non-conservative change of G to C in GTPAP2 and is also supported by 2 of 12 clones containing a C in this position compared with 7 of 12 clones containing G in this position. The SNPs at positions 1013 and 1020 are particularly significant because they are relatively rare, being found in only 2/12 or 17% of the identifiable clones encompassing this region, and because the SNP at position 1020 occurs in the 5' position of the codon for an encoded amino acid residue, resulting in a non-conservative change from glycine to arginine in GTPAP2. The change at position 1020 has potential consequences on the catalytic activity of the encoded protein which is located in this region of the protein.

10 Additional nucleic acid molecules encoding the mammalian GAP protein of the present invention were identified by using BLAST or BLAST2 (Altschul *et al.* (1997) *Nucleic Acids Res.* 25:3389-3402) against the ZOOSEQ database (Incyte Pharmaceuticals, Palo Alto CA) to align rat, mouse, and monkey clones with the human Rho GAP protein encoding nucleic acid molecules (SEQ ID NO:28 and SEQ ID NO:29). Figures 5A, 5B, 5C, 5D, and 5E show the alignments between SEQ ID  
15 NO:28 and non-human variants of this sequence from rat. The following clones, SEQ ID NOs:, (and their libraries) were identified by a BLAST score of at least 100 relative to the human sequence, SEQ ID NO:28: 701244926H1; SEQ ID NO:21 (RALINOH02), 700950169H2; SEQ ID NO:22 (RASPNON02), 701575974H1; SEQ ID NO:23 (RALITXT25), 701274036H1; SEQ ID NO:24 (RABFNON02), 700480328H1; SEQ ID NO:25 (RALINON03), 700935753H1; SEQ ID NO:26  
20 (RALINON03), and 700936061H1; SEQ ID NO:27 (RALINON03).

The nucleic acid sequences, SEQ ID NOs:1-29 may be used in hybridization and amplification technologies to identify and distinguish among SEQ ID NOs:28 and 29 and similar molecules in a sample. The molecules may be used to mimic human conditions, diseases, or disorders, produce transgenic animal models for these conditions, or to monitor animal toxicology studies, clinical trials,  
25 and subject/patient treatment profiles.

Northern analysis shows the expression of GTPAP in various tissues, at least 68% of which are associated with cancer, 32% of which are associated with inflammation and the immune response, and 18% of which are associated with cell proliferative conditions. Of particular note is the expression of GTPAP in gastrointestinal tissues (23%) including colon cancer.

30 Polynucleotides encoding GTPAP were mapped to a region of chromosome 22 containing genes associated with colon cancer. Using the public domain sequence tagged sites D22S1171 and DD22S92 previously found to define a region of chromosome 22 associated with colon cancer, polynucleotide sequences encoding GTPAP were aligned by BLAST2 analysis with genomic DNA within this region and matches found containing a minimum of 98% identity over at least 100 base  
35 pairs.

WO 01/61010

PCT/US01/05075

The DNA human sequence encoding the partial amino acid sequence for Rho GAP protein GI 6572185 was recently discovered to be localized to the region of chromosome 22 containing genes associated with colon cancer. Rho GAP proteins are potential tumor suppressors by virtue of their role as negative regulators or suppressors of Rho GTPase modulated signaling processes in normal cells.

##### 5 Characterization and Use of the Invention

###### cDNA libraries

In a particular embodiment disclosed herein, mRNA was isolated from mammalian cells and tissues using methods that are well known to those skilled in the art and used to prepare the cDNA libraries. The Incyte clones listed above were isolated from mammalian cDNA libraries. At least one library preparation representative of the invention is described in the EXAMPLES below. The consensus mammalian sequences were chemically and/or electronically assembled from fragments including Incyte clones and extension and/or shotgun sequences using computer programs such as PHRAP (Green, supra), GELVIEW Fragment Assembly system (Genetics Computer Group, Madison WI), and AUTOASSEMBLER application (PE Biosystems, Foster City CA). Clones, extension and/or shotgun sequences are electronically assembled into clusters and/or master clusters.

###### Sequencing

Methods for sequencing nucleic acids are well known in the art and may be used to practice any of the embodiments of the invention. These methods employ enzymes such as the Klenow fragment of DNA polymerase I, SEQUENASE, Taq DNA polymerase and thermostable T7 DNA polymerase (Amersham Pharmacia Biotech, Piscataway NJ), or combinations of polymerases and proofreading exonucleases such as those found in the ELONGASE amplification system (Life Technologies, Gaithersburg MD). Preferably, sequence preparation is automated with machines such as the HYDRA microdispenser (Robbins Scientific, Sunnyvale CA), MICROLAB 2200 system (Hamilton, Reno NV), and the DNA ENGINE thermal cycler (PTC200; MJ Research, Watertown MA). Machines used for sequencing include the ABI PRISM 3700, 377 or 373 DNA sequencing systems (PE Biosystems), the MEGABACE 1000 DNA sequencing system (Amersham Pharmacia Biotech), and the like. The sequences may be analyzed using a variety of algorithms which are well known in the art and described in Ausubel et al. (1997; Short Protocols in Molecular Biology, John Wiley & Sons, New York NY, unit 7.7) and Meyers (1995; Molecular Biology and Biotechnology, Wiley VCH, New York NY, pp. 856-853).

Shotgun sequencing is used to generate more sequence from cloned inserts derived from multiple sources. Shotgun sequencing methods are well known in the art and use thermostable DNA polymerases, heat-labile DNA polymerases, and primers chosen from representative regions flanking the nucleic acid molecules of interest. Incomplete assembled sequences are inspected for identity using various algorithms or programs such as CONSED (Gordon (1998) Genome Res. 8:195-202) which are well known in the art.

WO 01/61010

PCT/US01/05075

Contaminating sequences including vector or chimeric sequences or deleted sequences can be removed or restored, respectively, organizing the incomplete assembled sequences into finished sequences.

#### Extension of a Nucleic Acid Sequence

The sequences of the invention may be extended using various PCR-based methods known in the art.

5 For example, the XL-PCR kit (PE Biosystems), nested primers, and commercially available cDNA or genomic DNA libraries may be used to extend the nucleic acid sequence. For all PCR-based methods, primers may be designed using commercially available software, such as OLIGO 4.06 primer analysis software (National Biosciences, Plymouth MN) to be about 22 to 30 nucleotides in length, to have a GC content of about 50% or more, and to anneal to a target molecule at temperatures from about 55°C to about  
10 68°C. When extending a sequence to recover regulatory elements, it is preferable to use genomic, rather than cDNA libraries.

#### **USE OF THE MAMMALIAN NUCLEIC ACID MOLECULE**

##### Hybridization

The mammalian nucleic acid molecule and fragments thereof can be used in hybridization

15 technologies for various purposes. A probe may be designed or derived from unique regions such as the 5' regulatory region or from a conserved motif such as the GTPase-activating protein signature and used in protocols to identify naturally occurring molecules encoding the mammalian protein, allelic variants, or related molecules. The probe may be DNA or RNA, is usually single stranded and should have at least 50% sequence identity to any of the nucleic acid sequences. The probe may comprise at least 16 contiguous  
20 nucleotides of a nucleic acid molecule. Hybridization probes may be produced using oligolabeling, nick translation, end-labeling, or PCR amplification in the presence of labeled nucleotide. A vector containing the nucleic acid molecule or a fragment thereof may be used to produce an mRNA probe *in vitro* by addition of an RNA polymerase and labeled nucleotides. These procedures may be conducted using commercially available kits such as those provided by Amersham Pharmacia Biotech.

25 The stringency of hybridization is determined by G+C content of the probe, salt concentration, and temperature. In particular, stringency can be increased by reducing the concentration of salt or raising the hybridization temperature. In solutions used for some membrane based hybridizations, addition of an organic solvent such as formamide allows the reaction to occur at a lower temperature. Hybridization can be performed at low stringency with buffers, such as 5xSSC with 1% sodium dodecyl sulfate (SDS) at 60°  
30 C, which permits the formation of a hybridization complex between nucleic acid sequences that contain some mismatches. Subsequent washes are performed at higher stringency with buffers such as 0.2xSSC with 0.1% SDS at either 45° C (medium stringency) or 68° C (high stringency). At high stringency, hybridization complexes will remain stable only where the nucleic acid molecules are completely complementary. In some membrane-based hybridizations, preferably 35% or most preferably 50%, formamide can be added to the

WO 01/61010

PCT/US01/05075

hybridization solution to reduce the temperature at which hybridization is performed, and background signals can be reduced by the use of other detergents such as Sarkosyl or Triton X-100 (Sigma Aldrich, St. Louis MO) and a blocking agent such as denatured salmon sperm DNA. Selection of components and conditions for hybridization are well known to those skilled in the art and are reviewed in Ausubel (supra) 5 and Sambrook et al. (1989) Molecular Cloning, A Laboratory Manual, Cold Spring Harbor Press, Plainview NY.

Microarrays may be prepared and analyzed using methods known in the art. Oligonucleotides may be used as either probes or targets in a microarray. The microarray can be used to monitor the expression level of large numbers of genes simultaneously and to identify genetic variants, mutations, and single 10 nucleotide polymorphisms. Such information may be used to determine gene function; to understand the genetic basis of a condition, disease, or disorder; to diagnose a condition, disease, or disorder; and to develop and monitor the activities of therapeutic agents. (See, e.g., Brennan et al. (1995) USPN 5,474,796; Schena et al. (1996) Proc. Natl. Acad. Sci. 93:10614-10619; Baldeschweiler et al. (1995) PCT application WO95/251116; Shafiq et al. (1995) PCT application WO95/35505; Heller et al. (1997) Proc. Natl. Acad. 15 Sci. 94:2150-2155; and Heller et al. (1997) USPN 5,605,662.)

Hybridization probes are also useful in mapping the naturally occurring genomic sequence. The probes may be hybridized to: 1) a particular chromosome, 2) a specific region of a chromosome, 3) artificial chromosome constructions such as human artificial chromosomes (HAC), yeast artificial chromosomes (YAC), bacterial artificial chromosomes (BAC), bacterial P1 constructions, or single 20 chromosomes, or cDNA libraries made from any of these sources.

#### Expression

A multitude of nucleic acid molecules encoding GTPAP may be cloned into a vector and used to express the protein, or portions thereof, in host cells. The nucleic acid sequence can be engineered by such methods as DNA shuffling (Stemmer and Cramer (1996) USPN 5,830,721) and site-directed mutagenesis to 25 create new restriction sites, alter glycosylation patterns, change codon preference to increase expression in a particular host, produce splice variants, extend half-life, and the like. The expression vector may contain transcriptional and translational control elements (promoters, enhancers, specific initiation signals, and polyadenylated 3' sequence) from various sources that have been selected for their efficiency in a particular host. The vector, nucleic acid molecule, and regulatory elements are combined using in vitro recombinant 30 DNA techniques, synthetic techniques, and/or in vivo genetic recombination techniques well known in the art and described in Sambrook (supra, ch. 4, 8, 16 and 17).

A variety of host systems may be transformed with an expression vector. These include, but are not limited to, bacteria transformed with recombinant bacteriophage, plasmid, or cosmid DNA expression vectors; yeast transformed with yeast expression vectors; insect cell systems transformed with baculovirus

WO 01/61010

PCT/US01/05075

expression vectors; plant cell systems transformed with expression vectors containing viral and/or bacterial elements, or animal cell systems (Ausubel *supra*, unit 16). For example, an adenovirus transcription/translation complex may be utilized in mammalian cells. After sequences are ligated into the E1 or E3 region of the viral genome, the infective virus is used to transform and express the protein in host cells. The Rous sarcoma virus enhancer or SV40 or EBV-based vectors may also be used for high-level protein expression.

Routine cloning, subcloning, and propagation of nucleic acid sequences can be achieved using the multifunctional PBLUESCRIPT vector (Stratagene, La Jolla CA) or PSPORT1 plasmid (Life Technologies). Introduction of a nucleic acid sequence into the multiple cloning site of these vectors disrupts the *lacZ* gene and allows colorimetric screening for transformed bacteria. In addition, these vectors may be useful for *in vitro* transcription, dideoxy sequencing, single strand rescue with helper phage, and creation of nested deletions in the cloned sequence.

For long term production of recombinant proteins, the vector can be stably transformed into cell lines along with a selectable or visible marker gene on the same or on a separate vector. After transformation, cells are allowed to grow for about 1 to 2 days in enriched media and then are transferred to selective media. Selectable markers, antimetabolite, antibiotic, or herbicide resistance genes, confer resistance to the relevant selective agent and allow growth and recovery of cells that successfully express the introduced sequences. Resistant clones identified either by survival on selective media or by the expression of visible markers, such as anthocyanins, green fluorescent protein (GFP),  $\beta$  glucuronidase, luciferase, and the like, may be propagated using culture techniques. Visible markers are also used to quantify the amount of protein expressed by the introduced genes. Verification that the host cell contains the desired mammalian nucleic acid molecule is based on DNA-DNA or DNA-RNA hybridizations or PCR amplification techniques.

The host cell may be chosen for its ability to modify a recombinant protein in a desired fashion. Such modifications include acetylation, carboxylation, glycosylation, phosphorylation, lipidation, acylation and the like. Post-translational processing that cleaves a "prepro" form may also be used to specify protein targeting, folding, and/or activity. Different host cells available from the ATCC (Manassas MD) which have specific cellular machinery and characteristic mechanisms for post-translational activities may be chosen to ensure the correct modification and processing of the recombinant protein.

#### 30 Recovery of Proteins from Cell Culture

Heterologous moieties engineered into a vector for ease of purification include glutathione S-transferase (GST), calmodulin binding peptide (CBP), 6-His, FLAG, MYC, and the like. GST, CBP, and 6-His are purified using commercially available affinity matrices such as immobilized glutathione, calmodulin, and metal-chelate resins, respectively. FLAG and MYC are purified using commercially available

WO 01/61010

PCT/US01/05075

monoclonal and polyclonal antibodies. A proteolytic cleavage site may be located between the desired protein sequence and the heterologous moiety for ease of separation following purification. Methods for recombinant protein expression and purification are discussed in Ausubel *supra*, unit 16) and are commercially available.

#### 5 Chemical Synthesis of Peptides

Proteins or portions thereof may be produced not only by recombinant methods, but also by using chemical methods well known in the art. Solid phase peptide synthesis may be carried out in a batchwise or continuous flow process which sequentially adds  $\alpha$ -amino- and side chain-protected amino acid residues to an insoluble polymeric support via a linker group. A linker group such as methylamine-derivatized polyethylene glycol is attached to poly(styrene-co-divinylbenzene) to form the support resin. The amino acid residues are *N*- $\alpha$ -protected by acid labile Boc (*t*-butoxycarbonyl) or base-labile Fmoc (9-fluorenylmethoxycarbonyl). The carboxyl group of the protected amino acid is coupled to the amine of the linker group to anchor the residue to the solid phase support resin. Trifluoroacetic acid or piperidine are used to remove the protecting group in the case of Boc or Fmoc, respectively. Each additional amino acid is added to the anchored residue using a coupling agent or pre-activated amino acid derivative, and the resin is washed. The full length peptide is synthesized by sequential deprotection, coupling of derivitized amino acids, and washing with dichloromethane and/or *N,N*-dimethylformamide. The peptide is cleaved between the peptide carboxy terminus and the linker group to yield a peptide acid or amide. (Novabiochem 1997/98 Catalog and Peptide Synthesis Handbook, San Diego CA pp. S1-S20). Automated synthesis may also be carried out on machines such as the ABI 431A peptide synthesizer (PE Biosystems). A protein or portion thereof may be substantially purified by preparative high performance liquid chromatography and its composition confirmed by amino acid analysis or by sequencing (Creighton (1984) Proteins, Structures and Molecular Properties, WH Freeman, New York NY).

#### Preparation and Screening of Antibodies

Various hosts including goats, rabbits, rats, mice, humans, and others may be immunized by injection with GTPAP or any portion thereof. Adjuvants such as Freund's, mineral gels, and surface active substances such as lysolecithin, pluronic polysols, polyanions, peptides, oil emulsions, keyhole limpet hemocyanin (KLH), and dinitrophenol may be used to increase immunological response. The oligopeptide, peptide, or portion of protein used to induce antibodies should consist of at least about five amino acids, more preferably ten amino acids, which are identical to a portion of the natural protein. Oligopeptides may be fused with proteins such as KLH in order to produce the chimeric molecule.

Monoclonal antibodies may be prepared using any technique that provides for the production of antibodies by continuous cell lines in culture. These include, but are not limited to, the hybridoma technique, the human B-cell hybridoma technique, and the EBV-hybridoma technique. (See, e.g., Kohler *et al.* (1975)

WO 01/61010

PCT/US01/05075

Nature 256:493-497; Kozbor *et al.* (1985) *J. Immunol. Methods* 81:31-42; Cote *et al.* (1983) *Proc. Natl. Acad. Sci.* 80:2026-2030; and Cole *et al.* (1984) *Mol. Cell. Biol.* 62:109-120.)

Alternatively, techniques described for the production of single chain antibodies may be adapted, using methods known in the art, to produce epitope specific single chain antibodies. Antibody fragments which contain specific binding sites for epitopes of the mammalian protein may also be generated. For example, such fragments include, but are not limited to, F(ab)<sup>2</sup> fragments produced by pepsin digestion of the antibody molecule and Fab fragments generated by reducing the disulfide bridges of the F(ab)<sup>2</sup> fragments. Alternatively, Fab expression libraries may be constructed to allow rapid and easy identification of monoclonal Fab fragments with the desired specificity (Huse *et al.* (1989) *Science* 246:1275-1281).

The mammalian protein or a portion thereof may be used in screening assays of phagemid or B-lymphocyte immunoglobulin libraries to identify antibodies having the desired specificity. Numerous protocols for competitive binding or immunoassays using either polyclonal or monoclonal antibodies with established specificities are well known in the art. Such immunoassays typically involve the measurement of complex formation between the protein and its specific antibody. A two-site, monoclonal-based immunoassay utilizing monoclonal antibodies reactive to two non-interfering epitopes is preferred, but a competitive binding assay may also be employed (Pound (1998) *Immunochemical Protocols*, Humana Press, Totowa NJ).

#### Labeling of Molecules for Assay

A wide variety of labels and conjugation techniques are known by those skilled in the art and may be used in various nucleic acid, amino acid, and antibody assays. Synthesis of labeled molecules may be achieved using Promega (Madison WI) or Amersham Pharmacia Biotech kits for incorporation of a labeled nucleotide such as <sup>32</sup>P-dCTP, Cy3-dCTP or Cy5-dCTP or amino acid such as <sup>35</sup>S-methionine. Nucleotides and amino acids may be directly labeled with a variety of substances including fluorescent, chemiluminescent, or chromogenic agents, and the like, by chemical conjugation to amines, thiols and other groups present in the molecules using reagents such as BIODIPY or FITC (Molecular Probes, Eugene OR).

#### **DIAGNOSTICS**

The nucleic acid molecules, fragments, oligonucleotides, complementary RNA and DNA molecules, and PNAs may be used to detect and quantify altered gene expression, absence/presence vs. excess, expression of mRNAs or to monitor mRNA levels during therapeutic intervention. Conditions, diseases or disorders associated with altered expression include a cell proliferative disorder such as actinic keratosis, arteriosclerosis, atherosclerosis, bursitis, cirrhosis, hepatitis, mixed connective tissue disease (MCTD), psoriasis, primary thrombocythemia, and cancers including adenocarcinoma, leukemia, lymphoma, melanoma, myeloma, sarcoma, teratocarcinoma, and, in particular, cancer of the colon; an immune disorder such as inflammation, actinic keratosis, acquired immunodeficiency syndrome (AIDS), Addison's

WO 01/61010

PCT/US01/05075

disease, adult respiratory distress syndrome, allergies, anemia, asthma, autoimmune thyroiditis, bronchitis, bursitis, cholecystitis, cirrhosis, contact dermatitis, Crohn's disease, diabetes mellitus, emphysema, gout, Graves' disease, hepatitis, hypereosinophilia, irritable bowel syndrome, multiple sclerosis, myasthenia gravis, myocardial or pericardial inflammation, osteoarthritis, osteoporosis, pancreatitis, rheumatoid arthritis, scleroderma, and ulcerative colitis; and a cell signaling disorder including endocrine disorders such as disorders of the hypothalamus and pituitary resulting from lesions such as primary brain tumors, adenomas, hypophysectomy, aneurysms, vascular malformations, thrombosis, and complications due to head trauma; disorders associated with hyperpituitarism including acromegaly, gigantism, and syndrome of inappropriate antidiuretic hormone (ADH) secretion (SIADH); disorders associated with hypothyroidism including goiter, myxedema, acute thyroiditis associated with bacterial infection; and disorders associated with hyperparathyroidism including Conn disease (chronic hypercalcemia). The diagnostic assay may use hybridization or amplification technology to compare gene expression in a biological sample from a patient to standard samples in order to detect altered gene expression. Qualitative or quantitative methods for this comparison are well known in the art.

15 For example, the nucleic acid molecule or probe may be labeled by standard methods and added to a biological sample from a patient under conditions for the formation of hybridization complexes. After an incubation period, the sample is washed and the amount of label (or signal) associated with hybridization complexes, is quantified and compared with a standard value. If the amount of label in the patient sample is significantly altered in comparison to the standard value, then the presence of the associated condition, disease or disorder is indicated.

20 In order to provide a basis for the diagnosis of a condition, disease or disorder associated with gene expression, a normal or standard expression profile is established. This may be accomplished by combining a biological sample taken from normal subjects, either animal or human, with a probe under conditions for hybridization or amplification. Standard hybridization may be quantified by comparing the values obtained using normal subjects with values from an experiment in which a known amount of a substantially purified target sequence is used. Standard values obtained in this manner may be compared with values obtained from samples from patients who are symptomatic for a particular condition, disease, or disorder. Deviation from standard values toward those associated with a particular condition is used to diagnose that condition.

25 Such assays may also be used to evaluate the efficacy of a particular therapeutic treatment regimen in animal studies and in clinical trial or to monitor the treatment of an individual patient. Once the presence of a condition is established and a treatment protocol is initiated, diagnostic assays may be repeated on a regular basis to determine if the level of expression in the patient begins to approximate that which is observed in a normal subject. The results obtained from successive assays may be used to show the efficacy of treatment over a period ranging from several days to months.

35 Immunological Methods

WO 01/61010

PCT/US01/05075

Detection and quantification of a protein using either specific polyclonal or monoclonal antibodies are known in the art. Examples of such techniques include enzyme-linked immunosorbent assays (ELISAs), radioimmunoassays (RIAs), and fluorescence activated cell sorting (FACS). A two-site, monoclonal-based immunoassay utilizing monoclonal antibodies reactive to two non-interfering epitopes is preferred, but a competitive binding assay may be employed (Coligan *et al.* (1997) *Current Protocols in Immunology*, Wiley-Interscience, New York NY; Pound, *supra*).

#### 15 THERAPEUTICS

Chemical and structural similarity, e.g., in the context of sequences and motifs, exists between regions of GTPAP and GAP proteins. In addition, the expression of GTPAP is closely associated with cell proliferation, particularly colon cancer, and with inflammation and the immune response. Therefore, GTPAP appears to play a role in cell signaling, immune, and cell proliferative disorders, particularly colon cancer. In the treatment of conditions associated with increased expression or activity, it is desirable to decrease expression or protein activity. In the treatment of conditions associated with decreased expression or activity, it is desirable to increase expression or protein activity.

15 In one embodiment, the mammalian protein or a portion or derivative thereof may be administered to a subject to treat or prevent a condition associated with altered expression or activity of the mammalian protein. Examples of such conditions include, but are not limited to, a cell proliferative disorder such as actinic keratosis, arteriosclerosis, atherosclerosis, bursitis, cirrhosis, hepatitis, mixed connective tissue disease (MCTD), psoriasis, primary thrombocythemia, and cancers including adenocarcinoma, leukemia, 20 lymphoma, melanoma, myeloma, sarcoma, teratocarcinoma, and, in particular, cancer of the colon; an immune disorder such as inflammation, actinic keratosis, acquired immunodeficiency syndrome (AIDS), Addison's disease, adult respiratory distress syndrome, allergies, anemia, asthma, autoimmune thyroiditis, bronchitis, bursitis, cholecystitis, cirrhosis, contact dermatitis, Crohn's disease, diabetes mellitus, emphysema, gout, Graves' disease, hepatitis, hypereosinophilia, irritable bowel syndrome, multiple 25 sclerosis, myasthenia gravis, myocardial or pericardial inflammation, osteoarthritis, osteoporosis, pancreatitis, rheumatoid arthritis, scleroderma, and ulcerative colitis; and a cell signaling disorder including endocrine disorders such as disorders of the hypothalamus and pituitary resulting from lesions such as primary brain tumors, adenomas, hypophysectomy, aneurysms, vascular malformations, thrombosis, and complications due to head trauma; disorders associated with hyperpituitarism including 30 acromegaly, gigantism, and syndrome of inappropriate antidiuretic hormone (ADH) secretion (SIADH); disorders associated with hypothyroidism including goiter, myxedema, acute thyroiditis associated with bacterial infection; and disorders associated with hyperparathyroidism including Conn disease (chronic hypercalcemia).

In another embodiment, a pharmaceutical composition comprising a substantially purified 35 mammalian protein in conjunction with a pharmaceutical carrier may be administered to a subject to treat or

WO 01/61010

PCT/US01/05075

prevent a condition associated with altered expression or activity of the endogenous protein including, but not limited to, those provided above.

In a further embodiment, a ligand which modulates the activity of the mammalian protein may be administered to a subject to treat or prevent a condition associated with altered lifespan, expression, or activity of the protein including, but not limited to, those listed above. In one aspect, an antibody which specifically binds the mammalian protein may be used as a targeting or delivery mechanism for bringing a pharmaceutical agent to cells or tissue that express the mammalian protein.

In an additional embodiment, a vector capable of expressing the mammalian protein or a portion or derivative thereof may be administered to a subject to treat or prevent a condition associated with altered lifespan, expression, or activity of protein including, but not limited to, those described above.

In a still further embodiment, a vector expressing the complement of the nucleic acid molecule or fragments thereof may be administered to a subject to treat or prevent a condition associated with altered lifespan, expression, or activity of the protein including, but not limited to, those described above.

Any of the nucleic acid molecules, complementary molecules and fragments thereof, proteins or portions thereof, vectors delivering these nucleic acid molecules or proteins, and their ligands may be administered in combination with other therapeutic agents. Selection of the agents for use in combination therapy may be made by one of ordinary skill in the art according to conventional pharmaceutical principles. A combination of therapeutic agents may act synergistically to affect prevention or treatment of a particular condition at a lower dosage of each agent.

#### 20 Modification of Gene Expression Using Nucleic Acids

Gene expression may be modified by designing complementary or antisense molecules (DNA, RNA, or PNA) to the control, 5', 3', or other regulatory regions of the mammalian gene. Oligonucleotides designed with reference to the transcription initiation site are preferred. Similarly, inhibition can be achieved using triple helix base-pairing that inhibits the binding of polymerases, transcription factors, or regulatory molecules (Gee *et al.* In: Huber and Carr (1994) *Molecular and Immunologic Approaches*, Futura Publishing, Mt. Kisco NY, pp. 163-177). A complementary molecule may also be designed to block translation by preventing binding between ribosomes and mRNA. In one alternative, a library of nucleic acid molecules or fragments thereof may be screened to identify those which specifically bind a regulatory, nontranslated sequence.

30 Ribozymes, enzymatic RNA molecules, may also be used to catalyze the specific cleavage of RNA. The mechanism of ribozyme action involves sequence-specific hybridization of the ribozyme molecule to complementary target RNA followed by endonucleolytic cleavage at sites such as GUA, GUU, and GUC. Once such sites are identified, an oligonucleotide with the same sequence may be evaluated for secondary structural features that would render the oligonucleotide inoperable. The suitability of candidate targets may

WO 01/61010

PCT/US01/05075

also be evaluated by testing their hybridization with complementary oligonucleotides using ribonuclease protection assays.

Complementary nucleic acids and ribozymes of the invention may be prepared via recombinant expression, *in vitro* or *in vivo*, or using solid phase phosphoramidite chemical synthesis. In addition, RNA molecules may be modified to increase intracellular stability and half-life by addition of flanking sequences at the 5' and/or 3' ends of the molecule or by the use of phosphorothioate or 2' O-methyl rather than phosphodiesterase linkages within the backbone of the molecule. Modification is inherent in the production of PNAs and can be extended to other nucleic acid molecules. Either the inclusion of nontraditional bases such as inosine, queosine, and wybutosine, and or the modification of adenine, cytidine, guanine, thymine, and uridine with acetyl-, methyl-, thio- groups renders the molecule less available to endogenous endonucleases.

#### Screening Assays

The nucleic acid molecule encoding the mammalian protein may be used to screen a library of molecules or compounds for specific binding affinity. The libraries may be DNA molecules, RNA molecules, PNAs, peptides, proteins such as transcription factors, enhancers, repressors, and other ligands that regulate the activity, replication, transcription, or translation of the nucleic acid molecule in the biological system. The assay involves combining the mammalian nucleic acid molecule or a fragment thereof with the library of molecules under conditions allowing specific binding, and detecting specific binding to identify at least one molecule that specifically binds the nucleic acid molecule.

Similarly the mammalian protein or a portion thereof may be used to screen libraries of molecules or compounds in any of a variety of screening assays. The portion of the protein employed in such screening may be free in solution, affixed to an abiotic or biotic substrate (e.g. borne on a cell surface), or located intracellularly. Specific binding between the protein and molecule may be measured. Depending on the kind of library being screened, the assay may be used to identify DNA, RNA, or PNA molecules, agonists, antagonists, antibodies, immunoglobulins, inhibitors, peptides, proteins, drugs, or any other ligand, which specifically binds the protein. One method for high throughput screening using very small assay volumes and very small amounts of test compound is described in USPN 5,876,946, incorporated herein by reference, which screens large numbers of molecules for enzyme inhibition or receptor binding.

#### Purification of Ligand

The nucleic acid molecule or a fragment thereof may be used to purify a ligand from a sample. A method for using a mammalian nucleic acid molecule or a fragment thereof to purify a ligand would involve combining the nucleic acid molecule or a fragment thereof with a sample under conditions to allow specific binding, detecting specific binding, recovering the bound nucleic acid molecule, and using an appropriate agent to separate the nucleic acid molecule from the purified ligand.

WO 01/61010

PCT/US01/05075

Similarly, the protein or a portion thereof may be used to purify a ligand from a sample. A method for using a mammalian protein or a portion thereof to purify a ligand would involve combining the protein or a portion thereof with a sample under conditions to allow specific binding, detecting specific binding between the protein and ligand, recovering the bound protein, and using an appropriate chaotropic agent to separate  
5 the protein from the purified ligand.

#### Pharmacology

Pharmaceutical compositions are those substances wherein the active ingredients are contained in an effective amount to achieve a desired and intended purpose. The determination of an effective dose is well within the capability of those skilled in the art. For any compound, the therapeutically effective dose may be  
10 estimated initially either in cell culture assays or in animal models. The animal model is also used to achieve a desirable concentration range and route of administration. Such information may then be used to determine useful doses and routes for administration in humans.

A therapeutically effective dose refers to that amount of protein or inhibitor that ameliorates the symptoms or condition. Therapeutic efficacy and toxicity of such agents may be determined by standard  
15 pharmaceutical procedures in cell cultures or experimental animals, e.g., ED<sub>50</sub> (the dose therapeutically effective in 50% of the population) and LD<sub>50</sub> (the dose lethal to 50% of the population). The dose ratio between toxic and therapeutic effects is the therapeutic index, and it may be expressed as the ratio, LD<sub>50</sub>/ED<sub>50</sub>. Pharmaceutical compositions that exhibit large therapeutic indexes are preferred. The data obtained from cell culture assays and animal studies are used in formulating a range of dosage for human  
20 use.

#### **MODEL SYSTEMS**

Animal models may be used as bioassays where they exhibit a toxic response similar to that of humans and where exposure conditions are relevant to human exposures. Mammals are the most common models, and most toxicity studies are performed on rodents such as rats or mice because of low cost,  
25 availability, and abundant reference toxicology. Inbred rodent strains provide a convenient model for investigation of the physiological consequences of under- or over-expression of genes of interest and for the development of methods for diagnosis and treatment of diseases. A mammal inbred to over-express a particular gene (for example, secreted in milk) may also serve as a convenient source of the protein expressed by that gene.

#### Toxicology

Toxicology is the study of the effects of agents on living systems. The majority of toxicity studies are performed on rats or mice to help predict the effects of these agents on human health. Observation of qualitative and quantitative changes in physiology, behavior, homeostatic processes, and lethality are used to generate a toxicity profile and to assess the consequences on human health following exposure to the agent.

WO 01/61010

PCT/US01/05075

Genetic toxicology identifies and analyzes the ability of an agent to produce genetic mutations.

Genotoxic agents usually have common chemical or physical properties that facilitate interaction with nucleic acids and are most harmful when chromosomal aberrations are passed along to progeny.

Toxicological studies may identify agents that increase the frequency of structural or functional

5 abnormalities in progeny if administered to either parent before conception, to the mother during pregnancy, or to the developing organism. Mice and rats are most frequently used in these tests because of their short reproductive cycle which produces the number of organisms needed to satisfy statistical requirements.

Acute toxicity tests are based on a single administration of the agent to the subject to determine the symptomatology or lethality of the agent. Three experiments are conducted: 1) an initial dose-range-finding  
10 experiment, 2) an experiment to narrow the range of effective doses, and 3) a final experiment for establishing the dose-response curve.

Prolonged toxicity tests are based on the repeated administration of the agent. Rat and dog are commonly used in these studies to provide data from species in different families. With the exception of carcinogenesis, there is considerable evidence that daily administration of an agent at high-dose

15 concentrations for periods of three to four months will reveal most forms of toxicity in adult animals.

Chronic toxicity tests, with a duration of a year or more, are used to demonstrate either the absence of toxicity or the carcinogenic potential of an agent. When studies are conducted on rats, a minimum of three test groups plus one control group are used, and animals are examined and monitored at the outset and at intervals throughout the experiment.

#### 20 Transgenic Animal Models

Transgenic rodents that over-express or under-express a gene of interest may be inbred and used to model human diseases or to test therapeutic or toxic agents. (See USPN 4,736,866; USPN 5,175,383; and USPN 5,767,337). In some cases, the introduced gene may be activated at a specific time in a specific tissue type during fetal development or postnatally. Expression of the transgene is monitored by analysis of

25 phenotype or tissue-specific mRNA expression in transgenic animals before, during, and after challenge with experimental drug therapies.

#### Embryonic Stem Cells

Embryonic stem cells (ES) isolated from rodent embryos retain the potential to form an embryo.

When ES cells are placed inside a carrier embryo, they resume normal development and contribute to all  
30 tissues of the live-born animal. ES cells are the preferred cells used in the creation of experimental knockout and knockin rodent strains. Mouse ES cells, such as the mouse 129/SvJ cell line, are derived from the early mouse embryo and are grown under culture conditions well known in the art. Vectors for knockout strains contain a disease gene candidate modified to include a marker gene that disrupts transcription and/or translation *in vivo*. The vector is introduced into ES cells by transformation methods such as

WO 01/61010

PCT/US01/05075

electroporation, liposome delivery, microinjection, and the like which are well known in the art. The endogenous rodent gene is replaced by the disrupted disease gene through homologous recombination and integration during cell division. Transformed ES cells are identified, and preferably microinjected into mouse cell blastocysts such as those from the C57BL/6 mouse strain. The blastocysts are surgically transferred to pseudopregnant dams and the resulting chimeric progeny are genotyped and bred to produce heterozygous or homozygous strains.

ES cells are also used to study the differentiation of various cell types and tissues *in vitro*, such as neural cells, hematopoietic lineages, and cardiomyocytes (Bain *et al.* (1995) *Dev. Biol.* 168:342-357; Wiles and Keller (1991) *Development* 111:259-267; and Klug *et al.* (1996) *J. Clin. Invest.* 98:216-224). Recent developments demonstrate that ES cells derived from human blastocysts may also be manipulated *in vitro* to differentiate into eight separate cell lineages, including endoderm, mesoderm, and ectodermal cell types (Thomson (1998) *Science* 282:1145-1147).

#### Knockout Analysis

In gene knockout analysis, a region of a human disease gene candidate is enzymatically modified to include a non-mammalian gene such as the neomycin phosphotransferase gene (*neo*; Capeochli (1989) *Science* 244:1288-1292). The inserted coding sequence disrupts transcription and translation of the targeted gene and prevents biochemical synthesis of the disease candidate protein. The modified gene is transformed into cultured embryonic stem cells (described above), the transformed cells are injected into rodent blastulae, and the blastulae are implanted into pseudopregnant dams. Transgenic progeny are crossbred to obtain homozygous inbred lines.

#### Knockin Analysis

Totipotent ES cells, present in the early stages of embryonic development, can be used to create knockin humanized animals (pigs) or transgenic animal models (mice or rats) of human diseases. With knockin technology, a region of a human gene is injected into animal ES cells, and the human sequence integrates into the animal cell genome by recombination. Totipotent ES cells that contain the integrated human gene are handled as described above. Inbred animals are studied and treated to obtain information on the analogous human condition. These methods have been used to model several human diseases. (See, e.g., Lee *et al.* (1998) *Proc. Natl. Acad. Sci.* 95:11371-11376; Baudoin *et al.* (1998) *Genes Dev.* 12:1202-1216; and Zhang *et al.* (1998) *Mol. Cell Biol.* 18:3340-3349.)

#### 30 Non-Human Primate Model

The field of animal testing deals with data and methodology from basic sciences such as physiology, genetics, chemistry, pharmacology and statistics. These data are paramount in evaluating the effects of therapeutic agents on non-human primates as they can be related to human health. Monkeys are used as human surrogates in vaccine and drug evaluations, and their responses are relevant to human exposures

WO 01/61010

PCT/US01/05075

under similar conditions. Cynomolgus monkeys (*Macaca fascicularis*, *Macaca mulatta*) and common marmosets (*Callithrix jacchus*) are the most common non-human primates (NHPs) used in these investigations. Since great cost is associated with developing and maintaining a colony of NHPs, early research and toxicological studies are usually carried out in rodent models. In studies using behavioral measures such as drug addiction, NHPs are the first choice test animal. In addition, NHPs and individual humans exhibit differential sensitivities to many drugs and toxins and can be classified as "extensive metabolizers" and "poor metabolizers" of these agents.

In additional embodiments, the nucleic acid molecules that encode the mammalian protein may be used in any molecular biology techniques that have yet to be developed, provided the new techniques rely on properties of nucleic acid molecules that are currently known, including, but not limited to, such properties as the triplet genetic code and specific base pair interactions.

#### EXAMPLES

It is to be understood that this invention is not limited to the particular machines, materials and methods described. Although particular embodiments are described, equivalent embodiments may be used to practice the invention. The described embodiments are not intended to limit the scope of the invention which is limited only by the appended claims. The examples below are provided to illustrate the subject invention and are not included for the purpose of limiting the invention. For purposes of example, the preparation of the human uterine tissue cDNA library, UTRSNOR01, is described.

##### I Representative cDNA sequence preparation

The UTRSNOR01 cDNA library was constructed from microscopically normal endometrium obtained from a 29-year-old Caucasian female (specimen #0909A) during a vaginal hysterectomy and cystocele repair. Pathology of the uterus indicated a single intramural uterine leiomyoma. The endometrium was in secretory phase, and the cervix showed mild chronic cervicitis with focal squamous metaplasia. Patient history included hypothyroidism, pelvic floor relaxation, an incomplete T-12 injury from a motor vehicle accident causing paraplegia, and self-catheterization. Family history included benign hypertension, diabetes type II, and hyperlipidemia.

The frozen tissue was homogenized and lysed in TRIZOL reagent (1 gm tissue/10 ml TRIZOL; Cat. #10296-028; Life Technologies, Gaithersburg MD), a monophasic solution of phenol and guanidine isothiocyanate, using a POLYTRON homogenizer (PT-3000; Brinkmann Instruments, Westbury NY). After a brief incubation on ice, chloroform was added (1:5 v/v), and the lysate was centrifuged. The upper chloroform layer was removed to a fresh tube, and the RNA precipitated with isopropanol, resuspended in DEPC-treated water, and treated with DNase for 25 min at 37°C. The mRNA was re-extracted once with acid phenol-chloroform, pH 4.7, and precipitated using 0.3M sodium acetate and 2.5 volumes of ethanol.

WO 01/61010

PCT/US01/05075

The mRNA was isolated with the OLIGOTEX kit (Qiagen, Chatsworth CA) and used to construct the cDNA library.

The mRNA was handled according to the recommended protocols in the SUPERSCRIP<sup>T</sup> plasmid system for cDNA synthesis and plasmid Cloning (Life Technologies). The cDNAs were fractionated on a 5 SEPHAROSECL4B column (Amersham Pharmacia Biotech), and those cDNAs exceeding 400 bp were ligated into pINC<sup>Y</sup> 1 plasmid (Incyte Pharmaceuticals Palo Alto, CA). The plasmid was subsequently transformed into DH5 $\alpha$  competent cells (Life Technologies).

## II Isolation and Sequencing of cDNA clones

Plasmid DNA was released from the cells and purified using the REAL Prep 96 plasmid kit 10 (Qiagen). This kit enabled the simultaneous purification of 96 samples in a 96-well block using multi-channel reagent dispensers. The recommended protocol was employed except for the following changes: 1) the bacteria were cultured in 1 ml of sterile Terrific Broth (Life Technologies) with carbenicillin at 25 mg/L and glycerol at 0.4%; 2) after inoculation, the cultures were incubated for 19 hours and at the end of incubation, the cells were lysed with 0.3 ml of lysis buffer; and 3) following isopropanol precipitation, the 15 plasmid DNA pellet was resuspended in 0.1 ml of distilled water. After the last step in the protocol, samples were transferred to a 96-well block for storage at 4° C.

The cDNAs were prepared using a MICROLAB 2200 (Hamilton, Reno, NV) in combination with a Peltier DNA ENGINE thermal cyclers (MJ Research, Watertown, MA). The cDNAs were sequenced by the method of Sanger and Coulson (1975, J Mol Biol 94:441-448,) using an ABI PRISM 377 DNA 20 Sequencing Systems (PE Biosystems), and the reading frame was determined.

## III Identification, Extension, Assembly, and Analyses

The consensus sequence (SEQ ID NO:28) was assembled from Incyte Clones (library) 908465R2 (COLNNOT09), 957130R6 (KIDNNOT05), 1301520F6 (BRSTNOT07), 1580628H1 (DIODNOT01), 2531247F6 (COLNUT15), 3068538H1 (UTRSNOR01), and 3532286T6 (KIDNNOT25) from the 25 LIFESEQ database (Incyte Pharmaceuticals) of human cDNA sequences and used to identify additional sequences in the LIFESEQ and ZOOSEQ databases (Incyte Pharmaceuticals) related to GTPAP. Translation of SEQ ID NOs:28 and 29 using MACDN-ASIS PRO software (Hitachi Software Engineering) elucidated the coding regions, SEQ ID NOs:30 and 31, respectively.

The polynucleotide sequences derived from cDNA sequencing were assembled and analyzed using a 30 combination of software programs which utilize algorithms well known to those skilled in the art. Table 1 summarizes the tools, programs, and algorithms used and provides applicable descriptions, references, and threshold parameters. The first column of Table 1 shows the tools, programs, and algorithms used, the second column provides brief descriptions thereof, the third column presents appropriate references, all of which are incorporated by reference herein in their entirety, and the fourth column presents, where

WO 01/61010

PCT/US01/05075

applicable, the scores, probability values, and other parameters used to evaluate the strength of a match between two sequences (the higher the score, the greater the homology between two sequences). Sequences were analyzed using MACDNASIS PRO software (Hitachi Software Engineering, South San Francisco CA) and LASERGENE software (DNASTAR).

- 5 The polynucleotide sequences were validated by removing vector, linker, and polyA sequences and by masking ambiguous bases, using algorithms and programs based on BLAST, dynamic programming, and dinucleotide nearest neighbor analysis. The sequences were then queried against a selection of public databases such as the GenBank primate, rodent, mammalian, vertebrate, and eukaryote databases, and BLOCKS, PRINTS, DOMO, PRODOM, and PFAM to acquire annotation using programs based on
- 10 BLAST, FASTA, and BLIMPS. The sequences were assembled into full length polynucleotide sequences using programs based on Phred, PHRAP, and Consed, and were screened for open reading frames using programs based on GeneMark, BLAST, and FASTA. The full length polynucleotide sequences were translated to derive the corresponding full length amino acid sequences, and these full length sequences were subsequently analyzed by querying against databases such as the GenBank databases (described above),
- 15 SwissProt, BLOCKS, PRINTS, DOMO, PRODOM, Prosite, and Hidden Markov Model (HMM)-based protein family databases such as PFAM. HMM is a probabilistic approach which analyzes consensus primary structures of gene families. (See, e.g., Eddy, S.R. (1996) *Curr. Opin. Str. Biol.* 6:361-365.)

The programs described above for the assembly and analysis of full length polynucleotide and amino acid sequences were also used to identify polynucleotide sequence fragments from SEQ ID NOs:28 and 29. Fragments from about 20 to about 4000 nucleotides which are useful in hybridization and amplification technologies were described in The Invention section above.

#### IV Identification of Nucleic Acid Variants

- Nucleic acid molecules which are variants of the nucleic acid molecules encoding the mammalian GTPase-activating protein (SEQ ID NOs:28 and 29) were identified by using BLAST or BLAST2
- 25 (Altschul, *supra*; NCI-BLASTN version 2.0.4 with default parameters), to identify clones in the LIFESEQ or ZOOSEQ database (Incyte Pharmaceuticals) which aligned with SEQ ID NOs:28 and 29. Mammalian nucleic acid molecule variants were selected by BLAST score. The BLAST score is calculated by scoring +5 for every base that matches in a nucleic acid High Scoring Pair (HSP) and -4 for every mismatch. The BLAST alignments were visually inspected and those clones with BLAST scores greater than 100 were
- 30 aligned together using PHRAP (Green, *supra*).

#### V Northern Analysis

Northern analysis is a laboratory technique used to detect the presence of a transcript of a gene and involves the hybridization of a labeled probe to a membrane on which RNAs from a particular cell type or tissue have been bound.

WO 01/61010

PCT/US01/05075

Analogous computer techniques applying BLAST2 (NCI-BLASTN version 2.0.4 with default parameters) were used to search for identical or related molecules in nucleotide databases such as the LIFESEQ database or ZOOSEQ database (both Incyte Pharmaceuticals). Sequence-based analysis is much faster than membrane-based hybridization, and the sensitivity of the computer search can be modified to determine whether any particular match is categorized as exact or similar. The basis of the BLAST search is the product score which is defined as: (percent sequence identity x percent maximum BLAST score) divided by 100. The product score takes into account both the degree of similarity between two sequences and the length of the sequence match. For example, with a product score of 40, the match will be exact within a 1% to 2% error, and with a product score of at least 70, the match will be exact. Similar or related molecules are usually identified by selecting those which show product scores between 8 and 40.

The results of northern analyses are reported as a fraction or a percentage distribution of libraries in which the transcript encoding the mammalian protein occurred. Analysis involved the categorization of cDNA libraries by organ/tissue and disease. The organ/tissue categories included cardiovascular, dermatologic, developmental, endocrine, gastrointestinal, hematopoietic/immune, musculoskeletal, nervous, reproductive, and urologic. The disease categories included cancer, inflammation/trauma, cell proliferation, and neurological. For each category, the number of libraries expressing the sequence was counted and divided by the total number of libraries identified containing nucleic acid molecule variants of the nucleic acid molecule encoding the mammalian protein.

#### VI Extension of Nucleic Acid Molecules

At least one of the nucleic acid molecules used to assemble SEQ ID NOs:28 and 29 was produced by extension of an Incyte cDNA clone using oligonucleotide primers. One primer was synthesized to initiate 5' extension of the known fragment, and the other, to initiate 3' extension. The initial primers were designed using OLIGO 4.06 software (National Biosciences) to be about 22 to 30 nucleotides in length, to have a GC content of about 50%, and to anneal to the target sequence at temperatures of about 55°C to about 68°C. Any fragment which would result in hairpin structures and primer-primer dimerizations was avoided. Selected human cDNA libraries were used to extend the molecule. If more than one extension is needed, additional or nested sets of primers are designed.

High fidelity amplification was obtained by performing PCR in 96-well plates using the DNA ENGINE thermal cycler (MJ Research). The reaction mix contained DNA template, 200 nmol of each primer, reaction buffer containing Mg<sup>2+</sup>, (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>, and β-mercaptoethanol, Taq DNA polymerase (Amersham Pharmacia Biotech), ELONGASE enzyme (Ligo Technologies), and Pfu DNA polymerase (Stratagene), with the following parameters for primer pair selected from the plasmid: Step 1: 94°C, 3 min; Step 2: 94°C, 15 sec; Step 3: 60°C, 1 min; Step 4: 68°C, 2 min; Step 5: Steps 2, 3, and 4 repeated 20 times; Step 6: 68°C, 5 min; Step 7: storage at 4°C. In the alternative, parameters for the primer pair, T7

WO 01/61010

PCT/US01/05075

and SK+ (Stratagene), were as follows: Step 1: 94°C, 3 min; Step 2: 94°C, 15 sec; Step 3: 57°C, 1 min; Step 4: 68°C, 2 min; Step 5: Steps 2, 3, and 4 repeated 20 times; Step 6: 68°C, 5 min; Step 7: storage at 4°C.

The concentration of DNA in each well was determined by dispensing 100 µl PICOGREEN  
5 quantitation reagent (0.25% (w/v); Molecular Probes) dissolved in 1X TE and 0.5 µl of undiluted PCR  
product into each well of an opaque fluorimeter plate (Corning Costar, Acton MA) and allowing the DNA to  
bind to the reagent. The plate was scanned in a Fluoroskan II (Labsystems Oy, Helsinki, Finland) to  
measure the fluorescence of the sample and to quantify the concentration of DNA. A 5 µl to 10 µl aliquot of  
the reaction mixture was analyzed by electrophoresis on a 1% agarose mini-gel to determine which reactions  
10 were successful in producing longer sequence.

The extended sequences were desalted, concentrated, transferred to 384-well plates, digested with  
CviII cholera virus endonuclease (Molecular Biology Research, Madison WI), and sonicated or sheared  
prior to religation into pUC 18 vector (Amersham Pharmacia Biotech). For shotgun sequencing, the  
digested fragments were separated on about 0.6-0.8% agarose gels, fragments were excised as visualized  
15 under UV light, and agar removed/digested with AGARACE (Promega). Extended fragments were religated  
using T4 DNA ligase (New England Biolabs, Beverly MA) into pUC 18 vector (Amersham Pharmacia  
Biotech), treated with Pfu DNA polymerase (Stratagene) to fill-in restriction site overhangs, and transformed  
into competent *E. coli* cells. Transformed cells were selected on antibiotic-containing media, and individual  
colonies were picked and cultured overnight at 37°C in 384-well plates in LB/2x carbenicillin liquid media.

The cells were lysed, and DNA was amplified using Taq DNA polymerase (Amersham Pharmacia  
Biotech) and Pfu DNA polymerase (Stratagene) with the following parameters: Step 1: 94°C, 3 min; Step 2:  
20 94°C, 15 sec; Step 3: 60°C, 1 min; Step 4: 72°C, 2 min; Step 5: steps 2, 3, and 4 repeated 29 times; Step 6:  
72°C, 5 min; Step 7: storage at 4°C. DNA was quantified by PICOGREEN reagent (Molecular Probes) as  
described above. Samples with low DNA recoveries were reamplified using the conditions described above.  
25 Samples were diluted with 20% dimethylsulphoxide (1:2, v/v), and sequenced using DYENAMIC energy  
transfer sequencing primers and the DYENAMIC DIRECT kit (Amersham Pharmacia Biotech) or the ABI  
PRISM BIGDYE Terminator cycle sequencing ready reaction kit (PE Biosystems).

In like manner, a nucleic acid molecule of SEQ ID NOs:1-29 is used to obtain regulatory sequences  
using the procedure above, oligonucleotides designed for outward extension, and a genomic DNA library.

### 30 VII Hybridization Technologies and Analyses

#### Immobilization of Polynucleotides on a Substrate

Polynucleotides are applied to a substrate by one of the following methods. A mixture of  
polynucleotides is fractionated by gel electrophoresis and transferred to a nylon membrane by capillary  
transfer. Alternatively, the polynucleotides are individually ligated to a vector and inserted into bacterial

WO 01/61010

PCT/US01/05075

host cells to form a library. The polynucleotides are then arranged on a substrate by one of the following methods. In the first method, bacterial cells containing individual clones are robotically picked and arranged on a nylon membrane. The membrane is placed on LB agar containing selective agent (carbenicillin, kanamycin, ampicillin, or chloramphenicol depending on the vector used) and incubated at 37°C for 16 hr. 5 The membrane is removed from the agar and consecutively placed colony side up in 10% SDS, denaturing solution (1.5 M NaCl, 0.5 M NaOH), neutralizing solution (1.5 M NaCl, 1 M Tris, pH 8.0), and twice in 2xSSC for 10 min each. The membrane is then UV irradiated in a STRATALINKER UV-crosslinker (Stratagene).

In the second method, polynucleotides are amplified from bacterial vectors by thirty cycles of PCR 10 using primers complementary to vector sequences flanking the insert. PCR amplification increases a starting concentration of 1-2 ng nucleic acid to a final quantity greater than 5 µg. Amplified nucleic acids from about 400 bp to about 5000 bp in length are purified using SEPHACRYL-400 beads (Amersham Pharmacia Biotech). Purified nucleic acids are arranged on a nylon membrane manually or using a dot/slot blotting manifold and suction device and are immobilized by denaturation, neutralization, and UV irradiation as 15 described above. Purified nucleic acids are robotically arranged and immobilized on polymer-coated glass slides using the procedure described in USPN 5,807,522. Polymer-coated slides are prepared by cleaning glass microscope slides (Corning, Acton MA) by ultrasound in 0.1% SDS and acetone, etching in 4% hydrofluoric acid (VWR Scientific Products, West Chester PA), coating with 0.05% aminopropyl silane (Sigma Aldrich) in 95% ethanol, and curing in a 110°C oven. The slides are washed extensively with 20 distilled water between and after treatments. The nucleic acids are arrayed onto the slide and then immobilized by exposing the array to UV irradiation using a STRATALINKER UV-crosslinker (Stratagene). Arrays are then washed at room temperature in 0.2% SDS and rinsed three times in distilled water. Non-specific binding sites are blocked by incubation of arrays in 0.2% casein in PBS (Tropix, Bedford MA) for 30 min at 60°C; then the arrays are washed in 0.2% SDS and rinsed in distilled water as 25 before.

#### Probe Preparation for Membrane Hybridization

Hybridization probes derived from the polynucleotides of the Sequence Listing are employed for screening cDNAs, mRNAs, or genomic DNA in membrane-based hybridizations. Probes are prepared by diluting the polynucleotides to a concentration of 40-50 ng in 45 µl TE buffer, denaturing by heating to 30 100°C for five min, and briefly centrifuging. The denatured polynucleotide is then added to a REDIPRIME tube (Amersham Pharmacia Biotech), gently mixed until blue color is evenly distributed, and briefly centrifuged. Five microliters of [<sup>32</sup>P]dCTP is added to the tube, and the contents are incubated at 37°C for 10 min. The labeling reaction is stopped by adding 5 µl of 0.2M EDTA, and probe is purified from unincorporated nucleotides using a PROBEQUANT G-50 microcolumn (Amersham Pharmacia Biotech).

WO 01/61010

PCT/US01/05075

The purified probe is heated to 100°C for five min, snap cooled for two min on ice, and used in membrane-based hybridizations as described below.

Probe Preparation for Polymer Coated Slide Hybridization

Hybridization probes derived from mRNA isolated from samples are employed for screening  
5 polynucleotides of the Sequence Listing in array-based hybridizations. Probe is prepared using the  
GEMbright kit (Incyte Pharmaceuticals) by diluting mRNA to a concentration of 200 ng in 9 µl TE buffer  
and adding 5 µl 5x buffer, 1 µl 0.1 M DTT, 3 µl Cy3 or Cy5 labeling mix, 1 µl RNase inhibitor, 1 µl  
reverse transcriptase, and 5 µl 1x yeast control mRNAs. Yeast control mRNAs are synthesized by *in vitro*  
transcription from noncoding yeast genomic DNA (W. Lei, unpublished). As quantitative controls, one set  
10 of control mRNAs at 0.002 ng, 0.02 ng, 0.2 ng, and 2 ng are diluted into reverse transcription reaction  
mixture at ratios of 1:100,000, 1:10,000, 1:1000, and 1:100 (w/w) to sample mRNA respectively. To  
examine mRNA differential expression patterns, a second set of control mRNAs are diluted into reverse  
transcription reaction mixture at ratios of 1:3, 3:1, 1:10, 10:1, 1:25, and 25:1 (w/w). The reaction mixture is  
mixed and incubated at 37°C for two hr. The reaction mixture is then incubated for 20 min at 85°C, and  
15 probes are purified using two successive CHROMA SPIN+TE 30 columns (Cintech). Purified probe is  
ethanol precipitated by diluting probe to 90 µl in DEPC-treated water, adding 2 µl 1 mg/ml glycogen, 60 µl 5  
M sodium acetate, and 300 µl 100% ethanol. The probe is centrifuged for 20 min at 20,800xg, and the  
pellet is resuspended in 12 µl resuspension buffer, heated to 65°C for five min, and mixed thoroughly. The  
probe is heated and mixed as before and then stored on ice. Probe is used in high density array-based  
20 hybridizations as described below.

Membrane-based Hybridization

Membranes are pre-hybridized in hybridization solution containing 1% Sarkosyl and 1x high  
phosphate buffer (0.5 M NaCl, 0.1 M Na<sub>2</sub>HPO<sub>4</sub>, 5 mM EDTA, pH 7) at 55°C for two hr. The probe,  
diluted in 15 ml fresh hybridization solution, is then added to the membrane. The membrane is hybridized  
25 with the probe at 55°C for 16 hr. Following hybridization, the membrane is washed for 15 min at 25°C in  
1mM Tris (pH 8.0), 1% Sarkosyl, and four times for 15 min each at 25°C in 1mM Tris (pH 8.0). To detect  
hybridization complexes, XOMAT-AR film (Eastman Kodak, Rochester NY) is exposed to the membrane  
overnight at -70°C, developed, and examined visually.

Polymer Coated Slide-based Hybridization

30 Probe is heated to 65°C for five min, centrifuged five min at 9400 rpm in a 5415C microcentrifuge  
(Eppendorf Scientific, Westbury NY), and then 18 µl is aliquoted onto the array surface and covered with a  
coverslip. The arrays are transferred to a waterproof chamber having a cavity just slightly larger than a  
microscope slide. The chamber is kept at 100% humidity internally by the addition of 140 µl of 5xSSC in a  
corner of the chamber. The chamber containing the arrays is incubated for about 6.5 hr at 60°C. The

WO 01/61010

PCT/US01/05075

arrays are washed for 10 min at 45°C in 1xSSC, 0.1% SDS, and three times for 10 min each at 45°C in 0.1xSSC, and dried.

Hybridization reactions are performed in absolute or differential hybridization formats. In the absolute hybridization format, probe from one sample is hybridized to array elements, and signals are detected after hybridization complexes form. Signal strength correlates with probe mRNA levels in the sample. In the differential hybridization format, differential expression of a set of genes in two biological samples is analyzed. Probes from the two samples are prepared and labeled with different labeling moieties. A mixture of the two labeled probes is hybridized to the array elements, and signals are examined under conditions in which the emissions from the two different labels are individually detectable. Elements on the array that are hybridized to substantially equal numbers of probes derived from both biological samples give a distinct combined fluorescence (Shalon WO95/35505).

Hybridization complexes are detected with a microscope equipped with an Innova 70 mixed gas 10 W laser (Coherent, Santa Clara CA) capable of generating spectral lines at 488 nm for excitation of Cy3 and at 632 nm for excitation of Cy5. The excitation laser light is focused on the array using a 20X microscope objective (Nikon, Melville NY). The slide containing the array is placed on a computer-controlled X-Y stage on the microscope and raster-scanned past the objective with a resolution of 20 micrometers. In the differential hybridization format, the two fluorophores are sequentially excited by the laser. Emitted light is split, based on wavelength, into two photomultiplier tube detectors (PMT R1477, Hamamatsu Photonics Systems, Bridgewater NJ) corresponding to the two fluorophores. Appropriate filters positioned between the array and the photomultiplier tubes are used to filter the signals. The emission maxima of the fluorophores used are 565 nm for Cy3 and 650 nm for Cy5. The sensitivity of the scans is calibrated using the signal intensity generated by the yeast control mRNAs added to the probe mix. A specific location on the array contains a complementary DNA sequence, allowing the intensity of the signal at that location to be correlated with a weight ratio of hybridizing species of 1:100,000.

The output of the photomultiplier tube is digitized using a 12-bit RTL-835H analog-to-digital (A/D) conversion board (Analog Devices, Norwood MA) installed in an IBM-compatible PC computer. The digitized data are displayed as an image where the signal intensity is mapped using a linear 20-color transformation to a pseudocolor scale ranging from blue (low signal) to red (high signal). The data is also analyzed quantitatively. Where two different fluorophores are excited and measured simultaneously, the data are first corrected for optical crosstalk (due to overlapping emission spectra) between the fluorophores using each fluorophore's emission spectrum. A grid is superimposed over the fluorescence signal image such that the signal from each spot is centered in each element of the grid. The fluorescence signal within each element is then integrated to obtain a numerical value corresponding to the average intensity of the signal. The software used for signal analysis is the GUMTOOLS program (Lucyte Pharmaceuticals).

WO 01/61010

PCT/US01/05075

**VIII Complementary Nucleic Acid Molecules**

Molecules complementary to the nucleic acid molecule, or a fragment thereof, are used to detect, decrease, or inhibit gene expression. Although use of oligonucleotides comprising from about 15 to about 30 base pairs is described, the same procedure is used with larger or smaller fragments or their derivatives (PNAs). Oligonucleotides are designed using OLIGO 4.06 software (National Biosciences) and SEQ ID NOs:1-29. To inhibit transcription by preventing promoter binding, a complementary oligonucleotide is designed to bind to the most unique 5' sequence, most preferably about 10 nucleotides before the initiation codon of the open reading frame. To inhibit translation, a complementary oligonucleotide is designed to prevent ribosomal binding to the mRNA encoding the mammalian protein.

**IX Expression of the Mammalian Protein**

Expression and purification of the mammalian protein are achieved using bacterial or virus-based expression systems. For expression in bacteria, cDNA is subcloned into a vector containing an antibiotic resistance gene and an inducible promoter that directs high levels of cDNA transcription. Examples of such promoters include, but are not limited to, the *trp-lac (tac)* hybrid promoter and the T5 or T7 bacteriophage promoter in conjunction with the *lac* operator regulatory element. Recombinant vectors are transformed into bacterial hosts, e.g., BL21(DE3). Antibiotic resistant bacteria express the mammalian protein upon induction with isopropyl beta-D-thiogalactopyranoside (IPTG). Expression in eukaryotic cells is achieved by infecting *Spodoptera frugiperda (Sf9)* insect cells with recombinant baculovirus, *Autographica californica* nuclear polyhedrosis virus. The polyhedrin gene of baculovirus is replaced with the mammalian cDNA by either homologous recombination or bacterial-mediated transposition involving transfer plasmid intermediates. Viral infectivity is maintained and the strong polyhedrin promoter drives high levels of cDNA transcription.

In most expression systems, the mammalian protein is synthesized as a fusion protein with, e.g., GST or a peptide epitope tag, such as FLAG, permitting rapid, single-step, affinity-based purification of recombinant fusion protein from crude cell lysates. GST enables the purification of fusion proteins on immobilized glutathione under conditions that maintain protein activity and antigenicity (Amersham Pharmacia Biotech). Following purification, the GST moiety can be proteolytically cleaved from the mammalian protein at specifically engineered sites. FLAG, an 8-amino acid peptide, enables immunoprecipitation purification using commercially available monoclonal and polyclonal anti-FLAG antibodies (Eastman Kodak). 6-His, a stretch of six consecutive histidine residues, enables purification on metal-chelate resins (Qiagen). Methods for protein expression and purification are discussed in Ausubel (*supra*, unit 16). Purified mammalian protein obtained by these methods can be used directly in the following activity assay.

**X Functional Assays**

WO 01/61010

PCT/US01/05075

An assay for GTPAP activity measures the stimulation of intrinsic GTPase activity in a Rho GTPase. GTPase activity is using RhoA or Cdc42-GTPase is measured as described in Li et al. (1997) J Biol Chem 272:32830-32835. The assay is conducted in and absence of GTPAP and again in the presence of varying amounts of GTPAP. The amount of GTPase activity measured in the presence of GTPAP minus 5 the amount measured in the absence of GTPAP is proportional to the activity of GTPAP in the sample.

In the alternative, protein function is assessed by expressing the sequences encoding GTPAP at physiologically elevated levels in mammalian cell culture. The nucleic acid molecule is subcloned into PCMV SPORT vector (Life Technologies), which contains the strong cytomegalovirus promoter, and 5-10 µg of the vector is transformed into an endothelial or hematopoietic human cell line using transformation 10 methods well known in the art. An additional 1-2 µg of a plasmid containing sequence encoding CD64-GFP (Clontech, Palo Alto CA) is co-transformed to provide a fluorescent marker to identify transformed cells using flow cytometry (FCM).

The influence of the introduced genes on expression can be assessed using purified populations of these transformed cells. Since CD64-GFP, which is expressed on the surface of transformed cells, binds to 15 conserved regions of human immunoglobulin G (IgG), the transformed cells are separated using magnetic beads coated with either human IgG or antibody against CD64 (DYNAL, Lake Success NY). mRNA is purified from the cells and analyzed by hybridization techniques.

#### XI Production of GTPAP Specific Antibodies

GTPAP is purified using polyacrylamide gel electrophoresis and used to immunize mice or rabbits. 20 Antibodies are produced using the protocols below. Alternatively, the amino acid sequence of GTPAP is analyzed using LASERGENE software (DNASTAR) to determine regions of high immunogenicity. An immunogenic epitope, usually found near the C-terminus or in a hydrophilic region is selected, synthesized, and used to raise antibodies. Typically, epitopes of about 15 residues in length are produced using an ABI Model 431A peptide synthesizer (PE Biosystems) using Fmoc-chemistry and coupled to KLH (Sigma- 25 Aldrich) by reaction with N-maleimidobenzoyl-N-hydroxysuccinimide ester to increase immunogenicity.

Rabbits are immunized with the epitope-KLH complex in complete Freund's adjuvant.

Immunizations are repeated at intervals thereafter in incomplete Freund's adjuvant. After a minimum of seven weeks for mouse or twelve weeks for rabbit, antisera are drawn and tested for antipeptide activity.

Testing involves binding the peptide to plastic, blocking with 1% bovine serum albumin, reacting with rabbit 30 antisera, washing, and reacting with radio-iodinated goat anti-rabbit IgG. Methods well known in the art are used to determine antibody titer and the amount of complex formation.

#### XII Purification of Naturally Occurring Protein Using Specific Antibodies

Naturally occurring or recombinant mammalian protein is substantially purified by immunoadfinity chromatography using antibodies specific for the protein. An immunoadfinity column is constructed by

WO 01/61010

PCT/US01/05075

covalently coupling the antibody to CNBr-activated SEPHAROSE resin (Amersham Pharmacia Biotech). Media containing the protein is passed over the immunoaffinity column, and the column is washed using high ionic strength buffers in the presence of detergent to allow preferential absorbance of the protein. After coupling, the protein is eluted from the column using a buffer of pH 2-3 or a high concentration of urea or thiocyanate ion to disrupt antibody/protein binding, and the protein is collected.

#### XIII Mapping of Polynucleotides to the Colon Cancer Region of Chromosome 22

Public domain genomic DNA sequence data for chromosome 22 was obtained from Genbank and was described in Dunham, I. et al. (1999) Nature 402:489-495. Incyte templates, including Incyte ID No. 440424.5 (SEQ ID NO:29), were aligned with the public domain genomic sequence using BLAST2 and a match between any two sequences was determined by a 98% identity over at least 100bp. A region of chromosome 22 associated with colon cancer which was identified by Castells et al., *supra* is defined by the public domain sequence tagged sites D22S1171 and D22S92. These markers were aligned by BLAST2 with the sequence of chromosome 22 to determine the coordinates of genomic DNA associated with this region. Genomic DNA sequences which had been matched to SEQ ID NO:29 as described above was mapped within this region.

#### XIV Screening Molecules for Specific Binding with the Nucleic Acid Molecule or Protein

The nucleic acid molecule, or fragments thereof, or the protein, or portions thereof, are labeled with <sup>32</sup>P-dCTP, Cy3-dCTP, or Cy5-dCTP (Amersham Pharmacia Biotech), or with BIODIPY or FITC (Molecular Probes, Eugene OR), respectively. Libraries of candidate molecules previously arrayed on a substrate are incubated in the presence of labeled nucleic acid molecule or protein. After incubation under conditions for either a nucleic acid or amino acid sequence, the substrate is washed, and any position on the substrate retaining label, which indicates specific binding or complex formation, is assayed, and the binding molecule is identified. Data obtained using different concentrations of the nucleic acid or protein are used to calculate affinity between the labeled nucleic acid or protein and the bound molecule.

#### 25 XV Identification of Molecules Which Interact with GTPAP

GTPAP, or biologically active portions thereof, are labeled with <sup>125</sup>I Bolton-Hunter reagent (Bolton et al. (1973) Biochem. J. 133:529-539). Candidate ligand molecules previously arrayed in the wells of a multi-well plate are incubated with the labeled GTPAP, washed, and any wells with labeled GTPAP complex are assayed. Data obtained using different concentrations of GTPAP are used to calculate values for the number, affinity, and association of GTPAP with the candidate ligand molecules.

WO 01/61010

PCT/US01/05075

Table 1

Program	Description	Reference	Parameter Threshold
ABI FACTURA	A program that removes vector sequences and masks ambiguous bases in nucleic acid sequences.	Perkin-Elmer Applied Biosystems, Foster City, CA.	
ABI/PARACEL EDF	A Fast Data Finder useful in comparing and annotating amino acid or nucleic acid sequences.	Perkin-Elmer Applied Biosystems, Foster City, CA.; Paracel Inc., Pasadena, CA.	Mismatch <50%
ABI AutoAssembler	A program that assembles nucleic acid sequences.	Perkin-Elmer Applied Biosystems, Foster City, CA.	
BLAST	A Basic Local Alignment Search Tool useful in sequence similarity search for amino acid and nucleic acid sequences. BLAST includes five functions: blastp, blastn, blastx, tblastn, and tblastx.	Altschul, S.F. et al. (1990) <i>J. Mol. Biol.</i> 215:403-410; Altschul, S.F. et al. (1997) <i>Nucleic Acids Res.</i> 25: 3389-3402.	EST: Probability values=1.0E-8 or less Full Length sequencer: Probability values=1.0E-10 or less
FASTA	A Pearson and Lipman algorithm that searches for similarity between a query sequence and a group of sequences of the same type. FASTA compares at least five functions: fasta, tblastn, blastx, tblastx, and search.	Pearson, W.R. and D.J. Lipman (1988) <i>Proc. Natl. Acad. Sci.</i> 85:2444-2448; Pearson, W.R. (1990) <i>Methods Enzymol.</i> 183: 63-98; and Smith, T.F. and M. S. Waterman (1981) <i>Adv. Appl. Math.</i> 2:482-489.	EST: fasta E value=1.0E-6 Assembled EST: fasta Identity=95% or greater and Match length=200 bases or greater; fasta, E value=1.0E-8 or less Full Length sequencer: fasta score=100 or greater
BLIMPS	A Blocks Improved Searcher that matches a sequence against those in BLOCKS, PRINTS, DORO, PRODOM, and PFAM databases to search for gene families, sequence homology, and structural fingerprint regions.	Henikoff, S. and J.G. Henikoff, <i>Nucl. Acid Res.</i> 19:6565-72, 1991; J.G. Henikoff and S. Henikoff (1996) <i>Methods Enzymol.</i> 266:88-108; and Atwood, T.K. et al. (1997) <i>J. Chem. Inf. Comput. Sci.</i> 37: 417-424.	Scores 1000 or greater; Ratio of Score/Strength = 0.75 or larger; and, if applicable, Probability values=1.0E-3 or less
HMMER	An algorithm for searching a query sequence against hidden Markov model (HMM)-based databases or protein family consensus sequences, such as PFAM.	Krogh, A. et al. (1994) <i>J. Mol. Biol.</i> 233:1301-1332; Stormann, E.L.J., et al. (1988) <i>Nucleic Acids Res.</i> 20:320-322.	Scores 10-50 bits for PFAM hits, depending on individual protein families

WO 01/61010

PCT/US01/05075

Table 1 (cont.)

Program	Description	Reference	Parameter Threshold
ProfileScan	An algorithm that searches for structural and sequence motifs in protein sequences that match sequence patterns defined in Prosite.	Grubkov, M. et al. (1988) CABIOS 4:51-66; Grubkov, et al. (1989) Methods Enzymol. 183:148-159; Banoch, A. et al. (1997) Nucleic Acids Res. 25: 217-221.	Normalized quality score > OCC-specified "HC(H)" value for that particular Prosite motif. Generally, score=1.4-2.1.
Phred	A base-calling algorithm that examines automated sequencer traces with high sensitivity and probability.	Ewing, B. et al. (1998) Genome Res. 8:175-185; Ewing, B. and P. Green (1998) Genome Res. 8:186-194.	
Phrap	A Phile Revised Assembly Program including SWAT and CrossMatch, programs based on efficient implementation of the Smith-Waterman algorithm, useful in searching sequence homology and assembling DNA sequences.	Smith, T.F. and M. S. Waterman (1981) Adv. Appl. Math. 2:482-489; Smith, T.F. and M. S. Waterman (1981) J. Mol. Biol. 147:195-197; and Green, P., University of Washington, Seattle, WA.	Score=120 or greater; Match length= 56 or greater
Consed	A graphical tool for viewing and editing Phrap assemblies	Gordon, D. et al. (1998) Genome Res. 8:195-202.	
SPScan	A weight matrix analysis program that scans protein sequences for the presence of secretory signal peptides.	Nielson, H. et al. (1987) Protein Engineering 1(1):6; Charic, J.M. and S. Audic (1997) CABIOS 12:431-439.	Score=3.5 or greater
Motifs	A program that searches amino acid sequences for patterns that matched those defined in Prosite.	Bairoch et al. SUBSP: Wisconsin Package Program Manual, version 9, page M31-59, Genetics Computer Group, Madison, WI.	

WO 01/61010

PCT/US01/05075

What is claimed is:

1. A substantially purified nucleic acid molecule and fragments thereof encoding a GTPase activating protein, variants, and portions thereof, selected from:
  - a) an amino acid sequence of SEQ ID NO:30,
  - 5 b) a biologically active portion of the amino acid sequence of SEQ ID NO:30, and
  - c) an immunologically portion of the amino acid sequence of SEQ ID NO:30.
2. A variant of the mammalian nucleic molecule of claim 1 that encodes an amino acid sequence of SEQ ID NO:31.
3. An isolated and purified mammalian nucleic acid molecule selected from:
  - 10 a) SEQ ID NOs:1-29,
  - b) a fragment of a), and
  - c) a complement of a) or b).
4. A probe comprising probe comprising a polynucleotide of claim 3.
5. A single nucleotide polymorphism comprising a substitution selected from:
  - 15 a) A for T at nucleotide 662 of SEQ ID NO:28,
  - b) C for T at nucleotide 1013 of SEQ ID NO:28, and
  - c) G for C at nucleotide 1020 of SEQ ID NO:28.
6. A recombinant nucleic acid molecule comprising a promoter operably linked to the nucleic acid molecule of claim 3.
- 20 7. A cell transformed with the recombinant nucleic acid molecule of claim 6.
8. A method of producing a polypeptide, the method comprising:
  - a) culturing the cell of claim 7 under conditions for expression of the polypeptide, and
  - b) recovering the polypeptide so expressed.
9. A method for detecting the mammalian nucleic acid molecule in a sample, the method comprising:
  - 25 a) hybridizing the sample with the probe of claim 4 to form a hybridization complex; and
  - b) detecting the hybridization complex, wherein the hybridization complex indicates the presence of the mammalian nucleic acid molecule in the sample.
10. The method of claim 9 further comprising amplifying the nucleic acid molecule or a fragment thereof prior to hybridization.
- 30 11. A method of using the mammalian nucleic acid molecule or a fragment thereof to screen a library of molecules to identify at least one ligand that specifically binds the mammalian nucleic acid molecule, the method comprising:
  - a) combining the mammalian nucleic acid molecule of claim 3 with the library of molecules under conditions to allow specific binding, and

WO 01/61010

PCT/US01/05075

- b) detecting specific binding, thereby identifying a ligand that specifically binds the mammalian nucleic acid molecule.
12. The method of claim 11 wherein the library is selected from DNA molecules, RNA molecules, PNAs, peptides, and proteins.
- 5 13. A method of using the mammalian nucleic acid molecule or a fragment thereof to purify a ligand that specifically binds the mammalian nucleic acid molecule from a sample, the method comprising:
- a) combining the mammalian nucleic acid molecule or a fragment thereof of claim 3 with a sample under conditions to allow specific binding;
- b) recovering the bound mammalian nucleic acid molecule; and
- 10 c) separating the mammalian nucleic acid molecule from the ligand, thereby obtaining purified ligand.
14. A substantially purified protein or a portion thereof selected from:
- a) an amino acid sequence of SEQ ID NO:30;
- b) a biologically active portion of the amino acid sequence of SEQ ID NO:30; and
- 15 c) an immunologically active portion of the amino acid sequence of SEQ ID NO:30.
- 15 A variant of the protein of claim 14 comprising an amino acid sequence of SEQ ID NO:31.
16. A method for using a protein or a portion thereof to screen a library of molecules to identify at least one ligand that specifically binds the protein, the method comprising:
- a) combining the protein or a portion thereof of claim 14 with the library of molecules under
- 20 conditions to allow specific binding, and
- b) detecting specific binding, thereby identifying a ligand that specifically binds the protein.
17. The method of claim 16 wherein the library is selected from DNA molecules, RNA molecules, PNAs, peptides, proteins, agonists, antagonists, antibodies, immunoglobulins, inhibitors, drug compounds and pharmaceutical agents.
- 25 18. A method of using a protein or a portion thereof to purify a ligand that specifically binds the protein from a sample, the method comprising:
- a) combining the protein or a portion thereof of claim 14 with a sample under conditions to allow specific binding,
- b) recovering the bound protein, and
- 30 c) separating the protein from the ligand, thereby obtaining purified ligand.
19. A pharmaceutical composition comprising an effective amount of the protein of claim 14 and a pharmaceutically acceptable excipient.

WO 01/61010

PCT/US01/05075

20. A method for treating or preventing a disease or condition associated with altered expression of GTPAF, comprising administering to a patient in need of such treatment the pharmaceutical composition of claim 19.
21. The method of claim 20, wherein the disease or condition is selected from cell signaling, immune, and cell proliferative disorders, including cancer.
22. The method of claim 21, wherein the cancer is a colon cancer.

WO 01/61010

PC1/LS01/05075

1/21

```

5'  CAG ACC CCG CAC GCA GGT GGG GGC CGG CGG GGT CCG TGG CCA GAG CTG CAG AGA 55
-----
      10      19      28      37      46
GAC AAG GCG GCG GCT GCT GTC CTG GGT GCA GTG AGG AAG AGG CCC TCG GTG 109
-----
      64      73      82      91
      118      127      136      145      154      163
GTG CCC ATG GCT GGC CAG GAT CCT GCG CTG AGC ACG AGT CAC CCG TTC TAC GAC
-----
      M A G Q D P A L S T S H P F Y D
      172      181      190      199      208      217
GTG GCC AGA CAT GGC ATT CTG CAG GTG GCA GGG GAT GAC CGC TTT OGA AGA CGT
-----
      V A R H G I L Q V A G D D R F G R R
      226      235      244      253      262      271
GTT GTC ACG TTC AGC TGC TCC CCG ATG CCA CCC TCC CAC GAG CTG GAC CAC CAG
-----
      V V T F S C C R M P P S H E L D H Q
      280      289      298      307      316      325
CGG CTG CTG GAG TAT TTG ANG TAC ACA CTG GAC CAA TAC GTT GAG AAC GAT TAT
-----
      R L L E E Y L K Y T L D D Q Y V E N D Y

```

FIGURE 1A

WO 01/61010

PCT/US01/05075

2/21

```

334 ACC ATC GTC TAT TTC CAC TAC GGG CTG AAC AGC CCG AAC AAG CCT TTC CTG GGC 379
---
T I V Y F H Y G L N S R N K P S L G 370
---
388 TGG CTC CAG AGC GCA TAC AAG GAG TTC GAT AGG AAG TAC AAG AAG AAC TTG AAG 433
---
W L Q S A Y K E F D R K Y K K N L K 424
---
442 451 460 469 478 487
GCC CTC TAC GTG GTG CAC CCC ACC AGC TTC ATC AAG GTC CTG TGG AAC ATC TTG
---
A L Y V V H P T S F I K V L W N I L 478
---
496 505 514 523 532 541
AAG CCC CTC ATC AGT CAC AAG TTT GGG AAG AAA GTC ATC TAT TTC AAC TAC CTG
---
K P L I S H K F G K K V I Y F N Y L 532
---
550 559 568 577 586 595
AGT GAG CTC CAC GAA CAC CTT AAA TAC GAC CAG CTG GTC ATC CCT CCC GAA GTT
---
S E L H E H L K Y D Q L V I P P E V 586
---
604 613 622 631 640 649
TTG CCG TAC GAT GAG AAG CTC CAG AGC CTG CAC GAG GGC CCG ACG CCG CCT CCC
---
L R Y D E K L Q S L H E G R T P P P 640

```

FIGURE 1B



WO 01/61010

PCT/US01/05075

4/21

982	988	991	1000	1009	1018	1027
CAG	ATT	CTC	GGG	AAT	ACC	TGT
---	---	---	---	---	---	---
Q	I	L	G	I	T	C
---	---	---	---	---	---	---
1036	1045	1054	1063	1072	1081	1081
CAG	ATC	TTA	CGG	AGC	CTC	CCA
---	---	---	---	---	---	---
Q	I	L	R	S	L	P
---	---	---	---	---	---	---
1090	1099	1108	1117	1126	1135	1135
GGC	TTC	CTG	CAT	CGG	GTG	TCC
---	---	---	---	---	---	---
G	F	L	H	A	V	S
---	---	---	---	---	---	---
1144	1153	1162	1171	1180	1189	1189
AAC	CTG	GCC	TGT	GTC	TTC	GGG
---	---	---	---	---	---	---
N	L	A	C	V	F	G
---	---	---	---	---	---	---
1198	1207	1216	1225	1234	1243	1243
TCC	CTG	AGT	GCC	CTT	GTG	CCC
---	---	---	---	---	---	---
S	L	S	A	L	V	P
---	---	---	---	---	---	---
1252	1261	1270	1279	1288	1297	1297
TAT	GAA	AAG	ATC	TTC	AGC	ACC
---	---	---	---	---	---	---
Y	E	K	I	F	S	T
---	---	---	---	---	---	---

FIGURE 1D

```

1306      1315      1324      1333      1342      1351
TGG GAA CAG GGG AGC AGG GCA GCC CCT TTG CAG CAG GCT GTG CCA CGG ACA CAA
-----
W E Q G S R A A P L Q E A V P R T Q
-----
1360      1369      1378      1387      1396      1405
GCC ACG GGC CTC ACC AAG CCT ACC CTA CCT CCG AGT CCC CTG ATG GCA GCC AGA
-----
A T G L T K P T L P P S P L M A A R
-----
1414      1423      1432      1441      1450      1459
AGA CGT CTC TAG TGT TGC GAA CAC TCT GTA TAT TTC GAG CTA CCT CCC ACA CCT
-----
R R L
1468      1477      1486      1495      1504      1513
GTC TGT GCA CTT GTA TGT TTT ATA AAC TTG GCA TCT GTA AAA ATA ACC AGC CAT
-----
1522      1531      1540      1549
TAG ATG AAT TCA GRA CCT TCT AAT GAA AAA AAA AAA 3'
-----

```

FIGURE 1E

WO 01/61010

PCT/US01/05075

6/21

```

11      20      29      38      47      56
5'  AGA CCC GGC ACG CAG CAG GTG GGG GCC GGC GGC GGC GGC CCA GAG CTG CAG AGA
-----
65      74      83      92     101     110
GAC AAG GCG GCG GGT GCT CTG GGT GCA GTG AGG AAG AGG CCC TCG GTG
-----
119     128     137     146     155     164
GTG CCC ATG GCT GGC CAG GAT CCT GCG CTG AGC ACG AGT CAC CCG TTC TAC GAC
-----
M  A  G  Q  D  P  A  L  S  T  S  H  P  F  Y  D
-----
173     182     191     200     209     218
GTG GCC AGA CAT GGC AAT CTG CAG GTG GCA GGG GAT GAC CGC TTT GGA AGA CGT
-----
V  A  R  H  G  I  L  Q  V  A  G  D  D  R  F  G  R  R
-----
227     236     245     254     263     272
GTT GTC ACG TTC AGC TGC TGC CGG ATG CCA CCC TCC CAC GAG CTG GAC CAC CAG
-----
V  V  T  F  S  C  C  R  M  P  P  S  H  E  L  D  H  Q
-----
281     290     299     308     317     326
CGG CTG CTG GAG TAT TTG AAG TAC ACA CTG GAC CAA TAC TAC GTT GAG AAC GAT TAT
-----
R  L  L  E  Y  L  K  Y  T  L  D  Q  Y  V  E  N  D  Y

```

FIGURE 2A

WO 01/61010

PCT/US01/05075

7/21

```

335 ACC ATC TAT TTT CAC TAC GGG CTG AAC AGC CGG AAC AAG CCT TTC CTG GGC 380
---
344 T I V Y F H Y G L N S R N K P S L G 371
---
389 TGG CTC CAG AGC GCA TAC AAG GAG TTC GAT AGS AAG TAC AAG AAG AAC TTG AAG 434
---
398 W L Q S A Y K E F D R K Y K K N L K 425
---
443 GCC CTC TAC GTG GTG CAC CCC ACC ACC TTC ATC AAG GTC CTG TGG AAC ATC TTG 488
---
452 A L Y V V H P T S F I K V L W N I L 479
---
497 AAG CCC CTC ATC AGT CAC AAG TTT GGG AAG AAA GTC ATC TAT TTC AAC TAC CTG 542
---
506 K P L I S H K P G K K V I Y F N Y L 533
---
551 AGT GAG CTC CAC GAA CAC CTT AAA TAC GAC CAG CTG GTC ATC CCT CCC GAA GTT 596
---
560 S E L H E H L K Y D Q L V I P P E V 587
---
605 TTG CGG TAC GAT GAG AAG CTC CAG AGC CTG CAC GAG GGC CGG ACG CCT CCC 641
---
614 L R Y D E K L Q S L H E G R T P P P 650

```

FIGURE 2B

WO 01/61010

PCT/US01/05075

8/21

```

659 ACC AAG ACA CCA CCG CCG CGG CCC CCC CTG CCG ACA CAG CAG TTT GGC GTC AGT 704
---
T K T P P P R P P L P T Q Q F G V S
---
713 CTG CAA TAC CTC AAA GAC AAA AAT CAA GGC GAA CTC ATC CCC CCT GTG CTG AGG 758
---
L Q Y L K D K N Q G E L I P P V L R
---
767 TTC ACA GTG ACG TAC CTG AGA GAG AAA GGC CTG CGC ACC GAG GGC CTG TTC CGG 812
---
F T V T Y L R E K G L R T E G L F R
---
821 AGA TCC GCC AGC GTG CAG ACC GTC CGC GAG ATC CAG AGG CTC TAC AAC CAA GGG 866
---
R S A S V Q T V R E I Q R L Y N Q G
---
875 AAG CCC CTG AAC TTT GAC GAC TAC GGG GAC APT CAC ATC CCT GGC GTG ATC CTG 920
---
K P V N F D D Y G D I H I P A V I L
---
929 AAG ACC TTC CTG CGA GAG CTG CCC CAG CCG CPT CTG ACC TTC CAG GCC TAC GAG 974
---
K T F L R E L P Q P L L T F Q A Y E

```

FIGURE 2C

WO 01/61010

PCT/US01/05075

9/21

983	CAG ATT CTC GGG ATC ACC TGT GTG GAG AGC AGC CTG CGC GTC ACT CGC TGC CGC	1001	1010	1019	1028
---	---	---	---	---	---
Q	I L G I T C V E S S L R V T R C R				
1037	CAG ATC TTA CGG AGC CTC CCA GAG CAC AAC TAG GTC GTC CTC CGC TAC CTC ATG	1055	1064	1073	1082
---	---	---	---	---	---
Q	I L R S L P E H N Y V L R Y L M				
1091	GCC TTC CTG CAT GCG GTG TCC CGG GAG AGC ATC TTC AAC AAA ATG AAC AGC TCT	1109	1118	1127	1136
---	---	---	---	---	---
G	F L H A V S R E S I F N K M N S S				
1145	AAC CTG GCC TGT GTC TTC GGG CTG AAT TTG ATC TGG CCA TCC CAG GGG GTC TCC	1163	1172	1181	1190
---	---	---	---	---	---
N	L A C V F G L N L I W P S Q G V S				
1199	TCC CTG AGT GCC CTT GTG CCC CTG AAC ATG TTC ACT GAA CTG CTG ATC GAG TAC	1217	1226	1235	1244
---	---	---	---	---	---
S	L S A L V P L N M F T E L L I E Y				
1253	TAT GAA AAG ATC TTC AGC ACC CCG GAG GCA CCT GGG GAG CAC GGC CTG GCA CCA	1271	1280	1289	1298
---	---	---	---	---	---
Y	E K I F S T P E A P G E H G L A P				

FIGURE 2D

WO 01/61010

PCT/US01/05075

10/21

```

1307      1316      1325      1334      1343      1352
TGG GAA CAG GGG ACC AGG GCA GCC CCT TTG CAG GAG GCT GTG CCA CGG ACA CAA
-----
W E Q G S R A A P L Q E A V P R T Q
1361      1370      1379      1388      1397      1406
GCC ACG GGC CTC ACC AAG CCT ACC CTA CCT CCG AGT CCC CTG ATG GCA GCC AGA
-----
A T G L T K P T L P P S P L M A A R
1415      1424      1433      1442      1451      1460
AGA CGT CTC TAG TGT TGC GAA CAC TCT GTA TGT TTC GAG CTA CCT CCC ACA CCT
-----
R R L
1469      1478      1487      1496      1505      1514
GTC TGT GCA CTT GTA TGT TTT GTA AAC TTG GCA TCT GTA AAA ATA ACC AGC CAT
-----
TAG AUG AAT TCA GAA CCT TCT AAT G 3'
-----

```

FIGURE 2E



WO 01/61010

PCT/US01/05075

12/21

101	Y	K	K	N	L	K	A	L	Y	V	V	H	P	T	S	F	I	K	V	L	W	N	I	L	K	P	L	I	S	H	404424
101	Y	K	K	N	L	K	A	L	Y	V	V	H	P	T	S	F	I	K	V	L	W	N	I	L	K	P	L	I	S	H	3068538
1	Y	K	K	N	L	K	A	L	Y	V	V	H	P	T	S	F	I	K	V	L	W	N	I	L	K	P	L	I	S	H	96572185
151	Y	K	K	N	I	K	A	L	Y	V	H	P	T	S	F	I	K	T	L	I	L	F	K	P	L	I	S	F	9312212		
131	K	F	G	K	K	V	I	Y	F	N	Y	L	S	E	L	H	E	H	L	K	Y	D	Q	L	V	I	P	P	E	V	404424
131	K	F	G	K	K	V	I	Y	F	N	Y	L	S	E	L	H	E	H	L	K	Y	D	Q	L	V	I	P	P	E	V	3068538
31	K	F	G	K	K	V	I	Y	F	N	Y	L	S	E	L	H	E	H	L	K	Y	D	Q	L	V	I	P	P	E	V	96572185
181	K	F	G	Q	K	I	F	Y	V	N	Y	L	S	E	L	S	E	H	V	K	L	E	Q	L	G	I	P	R	Q	V	9312212
161	L	R	Y	D	E	K	L	Q	S	L	H	E	G	R	T	P	P	P	T	K	T	P	P	P	R	P	P	L	P	T	404424
161	L	R	Y	D	E	K	L	Q	S	L	H	E	G	R	T	P	P	P	T	K	T	P	P	P	R	P	P	L	P	T	3068538
61	L	R	Y	D	E	K	L	Q	S	L	H	E	G	R	T	P	P	P	T	K	T	P	P	P	R	P	P	L	P	T	96572185
211	L	R	Y	D	D	F	L	K	S	T	Q	K	S	P	A	T	A	P	-	K	P	M	P	P	R	P	P	L	P	N	9312212
191	Q	F	G	V	S	L	Q	Y	L	K	D	K	N	-	Q	G	E	L	I	P	P	V	L	R	F	F	T	V	T	Y	404424
191	Q	F	G	V	S	L	Q	Y	L	K	D	K	N	-	Q	G	E	L	I	P	P	V	L	R	F	F	T	V	T	Y	3068538
91	Q	F	G	V	S	L	Q	Y	L	K	D	K	N	-	Q	G	E	L	I	P	P	V	L	R	F	F	T	V	T	Y	96572185
240	Q	F	G	V	S	L	Q	H	L	Q	E	K	N	P	E	Q	E	P	I	P	I	V	L	R	E	T	V	A	Y	9312212	
220	L	R	E	K	G	L	R	T	E	G	L	F	R	R	S	A	S	V	Q	T	V	R	E	I	Q	R	L	Y	N	Q	404424
220	L	R	E	K	G	L	R	T	E	G	L	F	R	R	S	A	S	V	Q	T	V	R	E	I	Q	R	L	Y	N	Q	3068538
120	L	R	E	K	G	L	R	T	E	G	L	F	R	R	S	A	S	V	Q	T	V	R	E	I	Q	R	L	Y	N	Q	96572185
270	L	Q	A	H	A	L	T	E	G	I	F	R	R	S	A	N	T	Q	V	I	R	E	V	Q	K	Y	N	M	9312212		

FIGURE 3B

WO 01/61010

PCT/US01/05075

13/21

250 GKPVNFDYDDYGGDIHIPAVILLKTTFLRELPPQL 404424  
 250 GKPVNFDYDDYGGDIHIPAVILLKTTFLRELPPQL 3068538  
 150 GKPVNFDYDDYGGDIHIPAVILLKTTFLRELPPQL 96572185  
 300 GLPVDFDQYNELHLPAVILLKTTFLRELPEPL 9312212  
  
 280 LTFQAYEQIILGITCVESLRLVTRCRQILRS 404424  
 280 LTFQAYEQIILGITCVESLRLVTRGCRQILRS 3068538  
 180 LTFQAYEQIILGITCVESLRLVTRGCRQILRS 96572185  
 330 LTFDLYPHVVGFLNIDESQRVVPATLQVLQT 9312212  
  
 310 LPEHNYVVLRYLMLGFLHAVSRRESIFNKKMNS 404424  
 310 LPEHNYVVLRYLMLGFLHAVSRRESIFNKKMNS 3068538  
 210 LPEHNYVVLRYLMLGFLHAVSRRESIFNKKMNS 96572185  
 360 LPEENYQVLRFLTAFLLVQISAHSDQNKMTN 9312212  
  
 340 SNLACVFGNLNIWPSQGVSSLSALVPLNMF 404424  
 340 SNLACVFGNLNIWPSQGVSSLSALVPLNMF 3068538  
 240 SNLACVFGNLNIWPSQGVSSLSALVPLNMF 96572185  
 390 TINLAVVFGPNLWAKDAITLKAINPTF 9312212

FIGURE 3C

WO 01/61010

PCT/US01/05075

14/21

370	TE	L	L	L	I	E	Y	Y	E	K	I	F	S	T	P	E	A	P	G	E	H	G	L	A	P	W	E	Q	G	S	404424
370	TE	L	L	L	I	E	Y	Y	E	K	I	F	S	T	P	E	A	P	G	E	H	G	L	A	P	W	E	Q	G	S	3068538
270	TE	L	L	L	I	E	Y	Y	E	K	I	F	S	T	P	E	A	P	G	E	H	G	L	A	P	W	E	Q	G	S	96572185
420	T	K	F	L	L	D	H	Q	G	E	L	F	P	S	P	D	P	S	G	-	-	-	-	-	-	-	-	-	-	9312212	
400	R	A	A	P	L	Q	E	A	V	P	R	T	Q	A	T	G	L	T	K	P	T	L	P	P	S	P	L	M	A	A	404424
400	R	A	A	P	L	Q	E	A	V	P	R	T	Q	A	T	G	L	T	K	P	T	L	P	P	S	P	L	M	A	A	3068538
300	R	A	A	P	L	Q	E	A	V	P	R	T	Q	A	T	G	L	T	K	P	T	L	P	P	S	P	L	M	A	A	96572185
439	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	9312212		
430	R	R	R	L	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	404424	
430	R	R	R	L	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	3068538	
330	R	R	R	L	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	96572185	
439	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	9312212	

FIGURE 3D

WO 01/61010

PCT/US01/05075

424.5  
8538CB1  
5422F6  
1247F6  
130X313V1  
0528H1  
0545H1  
1457H1  
9657H1

15/21

```

caagacaccacgcccggccccccgctgcccacacagcagtttggcgtcagtcgtgcaatacctcaagac
caagacacctcggccggccccccgctgcccacacagcagtttggcgtcagtcgtgcaatacctcaagac
caagaaacttggggg*ggccccgggtgcc.....
caagacaccacgcccggccccccgctgcccacacagcagtttggcgtcagtcgtgcaatacctcaagac
caagacacctcggccggccccccgctgcccacacagcagtttggcgtcagtcgtgcaatacctcaagac
caagacancacgcccggccccccgctgcccacacagcagtttggcgtcagtcgtgcaatacctcaaa...
caagacaccacgcccggccccccgctgcccacacagcagtttggcgtcagtcgtgcaatacctcaa...
caagacaccacgcccggccccccgctgcccacacagcagtttggcgtcagtcgtgcaatacctcaagac
caagacaccacgcccggccccccgctgcccacacagcagtttggcgtcagtcgtgcaatacctcaagac
..... 660 670 680 690 700 710 720

```

FIGURE 4A

WO 01/61010

PCT/US01/05075

16/21

424.5 GAGCAGCCCTGCCGCTCA\*CTGCTGCCGCCAGAT\*CTTACGGAGCCTCCAGAGCACAACACTACGTCGTCC  
 853.8GR1 GAGCAGCCCTGCCGCTCA\*ctggctgCCGCCAGAT\*CTTACGGAGCCTCCAGAGCACAACACTACGTCGTCC  
 130X313V1 na\*cagc\*tnegttaa\*tt\*gnlccgcn\*at\*nttaag\*ago\*tcccagagcanaa\*lanqt\*gtcc  
 465R2 GAGCAGCCCTGCCGCTCA\*ctggctgCCGCCAGAT\*CTTACGGAGCCTCCAGAGCACAACACTACGTCGTCC  
 2758H1 GAGCAGCCCTGCCGCTCA\*ctggctgCCGCCAGAT\*CTTACGGAGCCTCCAGAGCACAACACTACGTCGTCC  
 130R6 GAGCAGCCCTGCCGCTCA\*ctggctgCCGCCAGAT\*CTTACGGAGCCTCCAGAGCACAACACTACGTCGTCC  
 130R1 GAGCAGCCCTGCCGCTCA\*ctggctgCCGCCAGAT\*CTTACGGAGCCTCCAGAGCACAACACTACGTCGTCC  
 130H1 GAGCAGCCCTGCCGCTCA\*CTCCCTGCCGCCAGAT\*CTTACGGAGCCTCCAGAGCACAACACTACGTCGTCC  
 4365T6.comp GAGCAGCCCTGCCGCTCA\*CTCCCTGCCGCCAGAT\*CTTACGGAGCCTCCAGAGCACAACACTACGTCGTCC  
 4444H1 GAGCAGCCCTGCCGCTCA\*CTCCCTGCCGCCAGAT\*CTTACGGAGCCTCCAGAGCACAACACTACGTCGTCC  
 6883T6.comp GAGCAGCCCTGCCGCTCA\*ctgctgCCGCCAGAT\*CTTACGGAGCCTCCAGAGCACAACACTACGTCGTCC  
 1520T6.comp tTcccagcCTGCCGCTCA\*ctggctgCCGCCAGAT\*CTTACGGAGCCTCCAGAGCACAACACTACGTCGTCC  
 1520H1 ...cagcCTGCCGCTCA\*ctggctgCCGCCAGAT\*CTTACGGAGCCTCCAGAGCACAACACTACGTCGTCC  
 1520T6 .....cagcCTGCCGCTCA\*ctggctgCCGCCAGAT\*CTTACGGAGCCTCCAGAGCACAACACTACGTCGTCC  
 2635X305D1.comp .....c\*antgntgcccgaagat\*\*taccgagcmtcccagAGCACAACACTACGTCGTCC  
 1010 A 1020 1030 1040 1050 1060 1070  
 A

FIGURE 4B









WO 01/61010

PCT/US01/05075

21/21

3068538CB1	atttgatctggccatccccaggggtctctccctcctgagtgcccttgcccccctgaacatgtccactgaact	
701575974H1	.....	
700480528H1	acttgatctggccatccccaggggtgcttccctgagcgcctggttccctgaaacttgttcacagagct	
700935753H1	acttgatctggccatccccaggggtgcttccctcctgagcgcctggttccctgaaacttgttcacagagct	1270 1280 1290 1300 1310 1320 1330
700936061H1		
3068538CB1	gctgactcaggtactatga*aaagatcttcagcaccctggaggcaacctggggagcagggcctggcaccatg	
701575974H1	.....	
700480528H1	gctgactcaggtactatgacaaag*tcctcagtgccccagagggcctggggagcaccatccgggatactgt	
700935753H1	gctgactcaggtactatgacaaag*tcctcagtgccccagagggcctggggagcaccatccgggatactgt	1340 1350 1360 1370 1380 1390 1400
700936061H1		
3068538CB1	ggaacagggagcagggcagcccttctgacagggctgtgccacggacacaaagccagggcctccaccaag	
701575974H1	.....	
700480528H1	cgaacgaaacagggctgctctgtt	
700935753H1	cgaacgaaacagggctgctctgttaccaaagaattcacacagcagggcactccccgggctca...	1410 1420 1430 1440 1450 1460 1470
700936061H1		

FIGURE 5E

WO 01/61010

PCT/US01/05075

<110> INCYTE GENOMICS, INC.  
KLINGER, Tod M.  
STEWART, Elizabeth A.  
YUE, Henry  
BAUGHN, Mariah R.

<120> GTPASE ACTIVATING PROTEINS

<130> PC-0010 PC1

<140> Unassigned

<141> Herewith

<150> 09/507,765

<151> 2000-02-18

<160> 33

<170> PERL Program

<210> 1

<211> 476

<212> DNA

<213> Homo sapiens

<220>

<221> misc\_feature

<223> Incyte ID No: 908465R2

<220>

<221> unsure

<222> 314, 343, 446

<223> a, t, c, g, or other

<400> 1  
ctctctctgca ggcctgcgca ccgagggcct gttccggaga tccgccagcg tgcagaccgt 60  
cccgagatc cagaggtctt acaaccaagg gaagcccttg aactttgacg actacggga 120  
cattcacatc cctgcctgga tccatgaagc attcctcgga ggcctgccc agccgcttct 180  
gaacttccag gcttaecgagc agattctctg gatcacctgt gttgagagca gcttgcctgt 240  
cactgctgc cctgattctt tacggagcct cccagagcac aactacgtag tccctcctta 300  
cctcagggc ttcttgatc cctgtcccgc gtagagatc ttnaaaaaa tgaacagctc 360  
taaacggcc tgtgtctctg gctgcaattt gatctgcat ccagggggtc tctctctga 420  
gtgccccttg gccctgaac atgttnactg aactgctgat ccagttacta gaaaag 476

<210> 2

<211> 321

<212> DNA

<213> Homo sapiens

<220>

<221> misc\_feature

<223> Incyte ID No: 957130R6

<400> 2  
cagagctgtc attctcaaac cctgtttctc ctcccgggac gggatgcagc gctgcccctg 60  
gcccctctgg agctcagcag ggagcccact gcccttccag gtttggagag cagccttcgt 120  
gtcactggct gctgacagat cttaacggag ctcccagagc acaactacgt cgtctctcgc 180  
taactatgg gcttctgca tgcagggttc cggggagca tcttaacaat aatgaaagc 240  
tctaacctgg cctgtgtott cgggctgaat ttgtcttggc catcccaggg ggtctctctc 300

WO 01/61010 PCT/US01/05075

ctgagtgcc ttgtgccct g 321

<210> 3  
 <211> 195  
 <212> DNA  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 1586628H1

<230>  
 <231> unsure  
 <232> 85, 96, 98, 102-103, 106, 108, 136, 161, 163  
 <233> a, t, c, g, or other

<400> 3  
 atctattca actacctgag tggctctac gaacacctt aatacgacca gctggtctac 60  
 cctccgaag ttgtccgta cgtatagaag ctccanacc tnnacnagg ccgaccccg 120  
 ctcccacca agacancacc gccgcggccc ccgctgccc nannagcgtt tggcgtcagt 180  
 ctgcaatacc tcaaa 195

<210> 4  
 <211> 487  
 <212> DNA  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 2631247F6

<230>  
 <231> unsure  
 <232> 483  
 <233> a, t, c, g, or other

<400> 4  
 ccggaacaag ccttcccgg gctggctcca gaggccatc aaggagtcc ataggaagta 60  
 caagaagaac ttgaagccc tctactggt gcaccccacc agcttcata agtctctgtg 120  
 gaacatcttg aagcccctca tcaatcaca gtttgggaag aaagtcatc atttcaacta 180  
 cctgagttag ctccaagAAC acbtaaaata ccaccagctg gtcctcctc ccgaagtttt 240  
 gcgtctgat gagaagctcc agagcctgca ccagggccgg ccgcccctc ccaccnagac 300  
 accaccgccg cggcccctgc tgcaccaca gcagtttggc gtcagtctgc aatacctcaa 360  
 agacaaaact caagggcaac tcatccccc ctgtctgag ttccacagta cgtaccctgag 420  
 agagaagggc ctgcgcaccg agggcctggt ccggagattc ggcagcgtg cagaccctcc 480  
 gcbagga 487

<210> 5  
 <211> 259  
 <212> DNA  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 3068538H1

<230>  
 <231> unsure  
 <232> 9, 17, 33, 39, 41, 48, 51, 86, 158, 198, 200, 219, 228  
 <233> a, t, c, g, or other

<400> 5

WO 01/61010

PCT/US01/05075

```

gcacggtggc gccgcgnggg ccggtggcca cactgcaba nacaccagcc nccggcggt 60
gctgtctgg gtcagctgg gaacangcc cgggtgtgc ccctgctgg ccaggtacct 120
gccgtgagca cgagtcacc gttctaccac gtggccahac atggcattct gcaggtggca 180
ggggatgacc gcttggman acgtgctgtc acgttcagnt gctgcccnat gccaccctcc 240
caccgctgg accaccagc
<210> 6
<211> 388
<212> DNA
<213> Homo sapiens
<220>
<221> misc_feature
<222> Incyte ID No: 3532286T6
<220>
<221> unsure
<222> 6, 9, 17, 23, 40, 65, 72, 77, 103, 171, 188, 193, 206, 316, 331, 247,
272, 274, 301, 348, 364, 382
<223> a, t, c, g, or other
<400> 6
acagggatng gaggtanctc canacacagc gttgcgacn ctayagacgt cttctggctg 60
ccatnagggg antcggnggt aggttaggtc fggtgagccc cgggctctgt gtcggtgcca 120
cagctccctg caaaggggct gccctgttcc cctgttccca tgggtccagg ccgtgtccc 180
caggtgcctc caggtgctc aagatcintt catagnactc gatcagcagc ccagtgaaca 240
tgctcaggg caacaaggca ctcagggagg anaaccctc ggatgcccag atcaaatcca 300
ncccgaaagc acagggcagc cttagagctgt ccattttggt gaagangctc tccggggaca 360
ccgntgcag gaagcccagc angtagcg
<210> 7
<211> 512
<212> DNA
<213> Homo sapiens
<220>
<221> misc_feature
<222> Incyte ID No: 1301520F6
<220>
<221> unsure
<222> 452, 471
<223> a, t, c, g, or other
<400> 7
cngcttgct gtbaactggt gccgccagat ctacggagc ctcccagagc adaaactcgt 60
cgtctccgc lacctcagc gcttccctga tccggtgtcc cgggagagca tcttcaaaa 120
aatgaacagc tctaacctgg cctgtgtctt cgggctgaat ttgatctggc catcccaggg 180
ggctctctcc ctgagtgccc ttgtgccctt gaacctgttc acggaactg tgatcgagta 240
ctatgaaaag atcttcagca ccccgagggc acctggggag ccgggctctg caccatggca 300
acagggagag agggcagccc ctctgcagga gctctgtcca cggacacag cccacggccc 360
tcaccagccc taacctactc cagatccctt gatggcagcc agaagacgta tctaggtttg 420
cgaacctctc gtatattctg agctactccc anaactgtct gtgacttgta ngllttatax 480
acttggcctc tgtaaaaata accagcatta ga
<210> 8
<211> 480
<212> DNA
<213> Homo sapiens
<220>
<221> misc_feature

```

WO 01/61010

PCT/US01/05075

<233> Incyte ID No: 246542276

<220>

<221> unsure

<222> 2, 231, 368, 470

<223> a, t, c, g, or other

<400> 8

```

anggcctgca ggcctgcctg gcgggggcag nagggtgagg gtctgtatht tccccgagtt 60
tatttcaatc ttgtttgat gtctttaat tgatcctgtt gagaggagta acattctgag 120
acloacagtg gaggagcctg tttcagggtt atgggcctg gglgtttct cggagcgcgg 180
cagcctgagc tcatcccccg tttccctcc: caggtaaaag nagaacttga aggccctcta 240
cgtgttgca: cccccacagc ttcatcaagc tccgttggaa cabcttgag cccctcaata 300
gtcccagtt tgggagaaa gtcatctatt tcaactcct gagtgagct caccacacc 360
ttaaataaga ccagctggtc atccctccg aacttttgg gtaagatgag aactccagag 420
cotgcaccag ggcgggagc gcggattoca acaagaaact tggcggggg ccccggtgcc 480

```

<210> 9

<211> 655

<212> DNA

<213> Homo sapiens

<220>

<221> misc\_feature

<223> Incyte ID No: 957130XJ13V1

<220>

<221> unsure

<222> 352, 425, 437-438, 440, 479, 481, 486, 493, 504, 512, 515, 534, 539, 550, 593, 603, 617, 623, 625, 630, 633, 640, 649

<223> a, t, c, g, or other

<400> 9

```

gctgggag aagtcctct atttcaacta cctgagctg ctoacgaa: acottaataa 60
gaccagctg gtcctccctc ccgaagtttt ccgtaccga gagaaactcc agagcctgca 120
cagggccgg ccgcgcctc ccaccaagac acctccgng cggccctccg tccccaaca 180
gcagtttgg gtcagctgc aatacctcaa agacaaaa: caaggcgaac tcatccccc 240
tgtctgggg tcaagctga cttactgaga gagaaggcc tgcgaccga gggcctgttc 300
cggagattcg ccaacgtgta gaccctcgc aatttcagag gtctcaacc agggagcct 360
gtgaacttg ccgaactcag gggcattcac atccctgag tgatccgaa gacctcctg 420
gaganttgc caagcymtn ttaccttcca ggctacgag aattctcgg ggaactctt 480
nttganacag ctncgtgtaa ttgntccgc cnatnttca agctcccca gcaaatcag 540
tgtcctcgn taactcattg gtctcctag ggtccggga agattcaaa naaatgaaa 600
gtntaaactg cctctntctg ggnthaattn atnccgcaan ccaaggggnt cccc 655

```

<210> 10

<211> 194

<212> DNA

<213> Homo sapiens

<220>

<221> misc\_feature

<223> Incyte ID No: 1580545H1

<220>

<221> unsure

<222> 161

<223> a, t, c, g, or other

<400> 10

```

actatttca actacctgag ttagctccc gaacacctta aatcagaca gctggctac 60
ctccagag ttgtgggta agatggagag ctccagacc tgcacgggg ccggacccc 120

```

WO 01/61010

PCT/US01/05075

ctctcccaaca agacaccacc gccggggccc ccgctgccca nacagcagtt tggcgtcagt 180  
cgcgaatacc tcaa 194

<210> 11  
<211> 281  
<212> DNA  
<213> Homo sapiens

<220>  
<221> misc\_feature  
<223> Incyte ID No: 1821457H1

<400> 11  
gctgttcata cctcccgaag ttttgggta cgtatgaaag ctccagacc tgcacgagg 60  
ccggagcccg cctcccaca agacaccacc gccggggcct ccgctgccca nacagcagtt 120  
tggcgtcagt ctgcaatacc tcaagacaaa aaatccagc gaactctcc cccctgtgt 180  
gaggttcaca gtagcgtacc tgaagagaaa aggcctggc accgagggcc tgttccggag 240  
atccgcagc gtgcagacc tcgcgagat ccagaggctc t 281

<210> 12  
<211> 268  
<212> DNA  
<213> Homo sapiens

<220>  
<221> misc\_feature  
<223> Incyte ID No: 4649657H1

<220>  
<221> unsure  
<222> 25  
<223> a, t, c, g, or other

<400> 12  
agcagtttgg cgtcagcttg caatacctca sagacaaaa tcaaggcga tcaiccccc 60  
tgtgctgagg ttcacagtga cgtacctgag agagaaggg ctggccacc aggcctgttt 120  
ccggagatcc gccngcgtgc agnccytcc cggatccag aggtctaca accaagggaa 180  
gccctgaaac ttgacgact acggggacat tcaatccct gccctgatcc tgaagacctt 240  
cctgcagag atgccccagc cgtttcig 268

<210> 13  
<211> 271  
<212> DNA  
<213> Homo sapiens

<220>  
<221> misc\_feature  
<223> Incyte ID No: 4002758H1

<220>  
<221> unsure  
<222> 8, 75, 120, 163, 178, 191, 201, 256, 265  
<223> a, t, c, g, or other

<400> 13  
gggaagcncg tgaactttga cgaactacgg gacattcaaa tccctgccc gatcctgaag 60  
accttctgc gaganctgcc ccagccgctt ctgaccttcc aggcctaca gcagattctn 120  
gggatacct gtagggagag cagcctgcgt gtaactggct gngccagat ctttaagggc 180  
ctcccaggac ncaactaegt ngtcctccgc tactctatgg gcttcttcca tgcggtytcc 240  
agggagagca tcttcnaca aactgnacagc t 271

<210> 14

WO 01/61010

PCT/US01/05075

<211> 498  
<212> DNA  
<213> Homo sapiens

<220>  
<221> misc\_feature  
<223> Incyte ID No: 957130R1

<220>  
<221> unsure  
<222> 387, 441, 479  
<223> a, t, c, g, or other

<400> 14  
cagagctgctc cttcctacac cctgttctc ccccgggac gggatgcagc gctgcccctg 60  
gccctctgg agctcagcag gggcccocat gccctccag gttggagag cagcctggct 120  
gtcactggct gccccagat cttacgggag ccccagagc acaclactg cgtcctccc 180  
aacctcattg gcttctctga tgggtgtcc gggcagac cttcaccac aatgaaacgc 240  
ctaaacctgg cctgtgctt cgggcctgat ttgatctgg cctcccaggg ggtcctctcc 300  
ctgggtgcc ctgtgcccct gaacatgttc actgactgc tgatcagta ctatgaaag 360  
actctcagca ccccgaggc acctgggag caggcctgg caccatgga acaggggagc 420  
agggcagccc ctttgacgga ngctgtgca cggacaaagc cagggtctca acaagctaac 480  
ctacctcg 488

<210> 15  
<211> 567  
<212> DNA  
<213> Homo sapiens

<220>  
<221> misc\_feature  
<223> Incyte ID No: 262436ST6

<400> 15  
acagatgcc agtttacaaa acatacaagt gccacagac gtttgggagc tagctcgaac 60  
atacagagt gttcccaaca ctagagcgt cttctggctg ccatacgggg actcgggggt 120  
agggtaggct tggtagggcc cgtggcttgt gtcctgggca cagcctctct caaaggggct 180  
gccctgctcc cctgttccca tggtagcagg ccgtgctccc cagctcctc cggggtgctg 240  
aagatcttt calagtactc gatacagcagl tcagtgaaac lgttcagggg cacaagggca 300  
ctcagggagg agaccctctg ggaagggcag atcaaatca gccccagagc acagggcagg 360  
ttagagctgt tcattttgtt gaagatgctc tcccgggaca ccgcagcag gaagcccatg 420  
aggtagcgga ggcagacgta gtltgctct gggaggtccc gtaagatctg gccgcagca 480  
gtgacgcgca gctgctctc cacaaggtg atcccagaa cctgctgta ggcttggag 540  
gtcagaagcg gctggggcag cctcagc 567

<210> 16  
<211> 265  
<212> DNA  
<213> Homo sapiens

<220>  
<221> misc\_feature  
<223> Incyte ID No: 204444H1

<400> 16  
ggagagctg ccccggcgc ttctgacct ccagcctac gaggagattc tcgggatcac 60  
ctgttggag agcagcctgc gcgtcactcg clyccggcag atcttacgga gcctcccaga 120  
gcacaactac gtgctctcc gctacctcat gggcttctgt catcgggtgt cccgggagag 180  
cattctaac aaatgaaac gctctaacct ggcctgtgtc ttggggctga atttgactg 240  
gccatcccag ggggtctct cctg 265

<210> 17

WO 01/61010 PCT/US01/05075  
 <211> 518  
 <212> DNA  
 <213> Homo sapiens  
 <220>  
 <221> misc\_feature  
 <223> Incyte ID No: 3416882T6  
 <230>  
 <221> unsure  
 <222> 496  
 <223> a, t, c, g, or other  
 <400> 17  
 tacaaaacat acaagtgcac agacagggtg gggaggtagc bogaaacata cagagtgttc 60  
 gcaacactag agacgltctc tggctgccat cagggaaact ggggttaggg tagcctgggt 120  
 ggggcccgtg gcttgygtcc gtggcacacc ctctgcana gggcttccc tctccccty 180  
 ttcccaggtt gccagggcgt gctcccaggg tgcctccggg gtgctgaaga tctttctca 240  
 glactcgatc agcagttccg tgaacatggt caggggcaaa agggcaetca gggaggagac 300  
 cccctgggat ggcacagaca aattcagccc gaagacacag gccaggttag agctgttcat 360  
 ttgtttgaag atgctctccc gggcacaccg atgcaggaag cccatgggtt agcggaggac 420  
 gacytattgt gctctggggg gctcgttaaga tctgggggca gcagttgaca cgcaggctgc 480  
 tctccacacc tgggaangca tggggctccc tgcctgac 518  
 <210> 18  
 <211> 548  
 <212> DNA  
 <213> Homo sapiens  
 <220>  
 <221> misc\_feature  
 <223> Incyte ID No: 1301520T6  
 <220>  
 <221> unsure  
 <222> 509  
 <223> a, t, c, g, or other  
 <400> 18  
 ggttctgaat tcatctaatg gctggttatt ttacagatg ccaagtttat aaaaatata 60  
 agtgcacaga caggtgtggg aggtagctcg aaataacag agtgttcgca acactagaga 120  
 cgtctctctg ctgcccctag gggactggga ggttaggttag gcttggtagg gccctggctt 180  
 tgtgtccctg gacacagctc ctgcacaagg gctgcoctgc tcccctgttc caatgggtcc 240  
 aggcctgtct cccaggtgc ctccggggly ctgaagatct ttcaatagta ctgatcagc 300  
 agttcagtga acatgttcag gggcaaacgg gactctaggg agsagacccc ctggatagc 360  
 cagatcaaat tcaqcccga gcaacaggcc aggttagagc tgttcatttt gttgaagatg 420  
 ctctcccggg aaccgcgatg caggaagccc atgaggttag gtaggagcac gtattgtgct 480  
 ctggagggtt ccgttagatc tggcggcanc agtgaacccc agcttgggat tcaattctcg 540  
 agcttacg 548  
 <210> 19  
 <211> 244  
 <212> DNA  
 <213> Homo sapiens  
 <220>  
 <221> misc\_feature  
 <223> Incyte ID No: 1301520H1  
 <400> 19  
 cagcctgggt gtcactggct gccgccaagt cttaccggag ctcccagagc acaactaagt 60  
 cgtctctccc tacctcatgy gcttctctga tgcggtgtcc cgggagagca tcttcaaca 120

WO 01/61010

PCT/US01/05075

aatgaacagc tctaacctgg cctgtgtctt cgggctgaaat ttgatctggc caccaccagg 180  
 ggtctctccu ctgagtgccc ttgtgcccci gaacatgllc aotgraactgc tgcctgagta 240  
 clat 244

<210> 20  
 <211> 521  
 <212> DNA  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 1422635X305D1

<220>  
 <221> unsure  
 <222> 495, 516, 519  
 <223> a, t, c, g, or other

<400> 20  
 callcgaagg ttctgaattc atctaatggc tggtratttt lccagatgcc aagtttaoaa 60  
 aacatacaag tgcacagaca ggtgtggagg gtgctogaa acatacagag tgttcycaac 120  
 actagagacg tctctgtgct gccatcaggg gactcggagg tagggttagc ttggtgaggc 180  
 cctgtgcttg tctcctgtgc acagctctct gccaaagggc tgcctgtctc cctgttccc 240  
 atggtgccag gccgtgtctc ccaggtgctc ccgggggtct gaagatcttt tcalagtact 300  
 cgateagcag ttcagtgaa: atgttcaggc gccaaagggc actcagggag gagaccctct 360  
 yggatggcca galcaaatc agcccyaaag cacagggccag gttagagctg ttcattttgt 420  
 tgaagatgct ctcccggac accgatgca ggagccca: y agttagcggg gaagcagcga 480  
 gttgtgctct gggangctcc gtaaatcttg cggcancant g 521

<210> 21  
 <211> 213  
 <212> DNA  
 <213> Rattus norvegicus

<220>  
 <221> misc\_feature  
 <223> incyte ID No: 701244926H1

<400> 21  
 ccggcctgcc ctggggccag ccaggtgtgc ggttagagta gctgagatca ggagagggtc 60  
 tcygtggacc gagctgcaga gacaaggaag cagcagccac actgagggcc acaggaggac 120  
 cctcagtggt gttctatggt ggcctggacc ccacgtcag cacaaatcac ccattctatg 180  
 atgtggccag acacggcacc ctgcaggtgg cag 213

<210> 22  
 <211> 273  
 <212> DNA  
 <213> Rattus norvegicus

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 700950169H2

<400> 22  
 ctggggccag ccaggtgtgc ggttagagta gctgagatca ggagagggtc tgggtggtcc 60  
 gagctgcaga gacaaggaag cagcagccac actgagggcc acaggaggac cctcagtggt 120  
 gttcattggt ggcctggacc ccacgtctag caaaatcac ccattctatg atgtggccag 180  
 tccagtcacc ctgcaggtgg cagggggtga ccccccgggg agacgcacct tcaatttcag 240  
 ctgctgcagg ttgccacct tgcaccagct caa 273

<210> 23  
 <211> 325

WO 01/61010

PCT/US01/05075

```

<212> DNA
<213> Rattus norvegicus

<220>
<221> misc_feature
<223> Incyte ID No: 701575974H1

<400> 23
cagccctccc caccagaagc ccgccaactc ggcgcctct gctaccacag cagttcggcg 60
lcagtttgca atacctcaga gacaaaaatc aaggtgaact cctccccct gtgcctgcgtt 120
ggacggtgac atactctaga gaaaaaggac tgcacactga aggcctgttc cggagatcac 180
ccagcgccca gactgtccgc cagggtcagc ggcctctaga tcaagggagc cctgtgaact 240
ttagatgatta tggtagacag caccctccag ctgtgattct aaagacatll ctttagagac 300
tgccccagcc actgctgacc tcca                                     325

<210> 24
<211> 233
<212> DNA
<213> Rattus norvegicus

<220>
<221> misc_feature
<223> Incyte ID No: 701274036H1

<400> 24
atcaaggllg actcctcccc cctgtgctgc gttggacggt gacatatctg agagaaaaag 60
ggaagcctgt gaactttgat gattatggcg acatgcacct cccagctgtg attctaaaga 120
catttctctg agagctgccc cagccactgc tgaacttcca agcctaagag cagattctct 180
ggtaccacag tglggagagc agcctggagc tgaccactgc ccgctgata ctg                                     233

<210> 25
<211> 128
<212> DNA
<213> Rattus norvegicus

<220>
<221> misc_feature
<223> Incyte ID No: 700480528H1

<400> 25
ctggaggagc ctcccagaac acaactatgc cgtctctcgc taactcaggg gcttctctga 60
tgagggtgtc ctggagagca ttccaaacaa gatgaacagc tctaacctgg catgtgtgtt 120
tgggctga

<210> 26
<211> 212
<212> DNA
<213> Rattus norvegicus

<220>
<221> misc_feature
<223> Incyte ID No: 700935753H1

<400> 26
tttcaaaana gatgaacagc tctaacctgg catgtgtgtt tgggctgaac ttgatctggc 60
catcccaggg ggtggcttcc ctgagcgccc tggttctctt gaacttgto acagagctgc 120
tgatagagta ctatgacaaa gtcttcagtg cccaggaggg ccttggggag cacatccggg 180
atctgtctga aacgaascag gctggctctg tt                                     212

<210> 27
<211> 251
<212> DNA

```

WO 01/61010

PCT/US01/05075

<213> Rattus norvegicus

<220>

<221> misc\_feature

<223> Incyte ID No: 700936061H1

<400> 27

```

tttcaaacaa gatgaacagc tetaaacctgy catgtgtgtt tgggctgaac ttgatctgyc 60
catccccggg ggtggcttcc ctgagcgcgc tgggtccctc gaacttgctc acagagctgc 120
tgatagagta ctatgacaaa gtcttcaagt cccaggaggg cctctggggg cacatccggc 180
atactgtcga aacgaacacg gctggtccctg ttaccaaaaga attcacacag acgggcactc 240
ccccggctcc a

```

<210> 28

<211> 1549

<212> DNA

<213> Homo sapiens

<220>

<221> misc\_feature

<223> Incyte ID No: 3068538CR1

<400> 28

```

gcagaccggg cacqacagtg ggggcggggg gggctcgtgg ccagagctgc agagagacaa 60
ggggcggggg gctgclytgc tgggtccagt gaggagagag cctctgtag tgcctatgac 120
tggccaggat cclgctgta gcacgagtca cccgttctac gacgtggcca gacatggcat 180
tclgacagtg gaaggggagt accgctttgg aagacgtgtt gtcacgttca gctgctgccc 240
gatgcaacc tcccacgagc tggaccccga ggggtctgtc gactattga agtacaacct 300
ggcaaacac gttgagacgy atkaiacacat cglctatttc caclacggg tgaacagccg 360
gaataagcct tccctgggct ggtccacagc cgtatacaag gacttcgata ggaagtcaaa 420
gaagaacttg aagcctctct acgtggtgca ccccacagc ttctcaag tctgtggaaa 480
catcttgaag cccctatca gtacaagtt tgggaagaaa gtcctctatt tcaactacct 540
gagtgagctc cacgaacacc ttaaatacga ccagctggtc atccctccg aaglttttgg 600
gtacgatgag sagctccaga gcctgcacga ggcctggag cccctccca ccagagacc 660
tcccgccggg ccccggctgc ccaacacaga gtttgggctc agtctcaat acctcaaga 720
caaaagctca ggcgaactca tccccctgtt gctgaggttc acagtgcgt acctgagaga 780
gaaagagctg cgcacagagg gcctlytccg gayatccgc ayctgcaga ccgtccgca 840
gatccagagc ctctacaacc aaggyaagcc cgtgaacttt gacgactag gggcaattca 900
catccctgcc gtgactctga agacccttct gcgagagctg ccccagctgc ttctgacttt 960
ccagggctac gacgagattc tgggattcac ctgttggag acagagctgc gtgtcactgg 1020
ctgcccggag atcttaagga gctctccaga gcacaactac gtcgtccctc gctcaactat 1080
gggcttccctg catgcccgtt cccgggagag catcttcaac aaaatgaca gctcaacct 1140
ggcctgtgtc ttccggctga atttgatctg gccatcccag ggggtctct cctbgaatgc 1200
ccttctgccc ctgaactgt tcactgaact gctgatggag tactatgaaa agatcttcag 1260
cacccggggg gcaactgggg agcaagcctt ggcacatgg gaaacggggg gcagggcagc 1320
cccttgcag gaggctgtgc kacggaact agcccgggc ctcaacagc ctacactacc 1380
ctcagctccc ctgatggcag ccagaagacg tctctagttt lgcgaact ctgtatatt 1440
cgagctactt cccacactg tctgtgact tctatgtttt atnaacttgy catctgtaca 1500
aataaccago cattaagatga attcagpaco ttotatgaa aaaaaaaaa 1549

```

<210> 29

<211> 1539

<212> DNA

<213> Homo sapiens

<220>

<221> misc\_feature

<223> Incyte ID No: 404424.5

<400> 29

```

gcagaccggg cacgacagtg ggggcggggg gggctcgtgg accagagctg cagagagaca 60
agggcgsggg gctgctgtg ctgggtgag tggagagag gccctcgtg gtgccatgg 120

```

WO 01/61010

PCT/US01/05075

```

ctggccaggga tccgctgctg agcaccgagtc acccgttcta cgaagctggcc agaatggcca 186
ttctgcaagtl ggcaggggat gaccgclltg gaagacgtgt tctcacttc agctgtgccc 240
ggatgcccacc ctcccaagag ctggaccacc agcggclgtt ggaglatctg aagtacacac 306
tggacaataa cgtttgagac gattakaca tctgtctatt ccactacggg ctgaacagcc 360
ggaacaagcc tccctgggc tggctccaga ggcatacaa ggagttgat aggaagtac 420
agaagaactt gaagcccttc taccggtgct accccaccag ctctactcaag gtcctgtgga 480
acatcttgaa cccctcactc agtccacaagt ttgggaagaa agtcatctat ttcaactacc 540
lgagtggctt ccaagtaaac cttaaatcag accagctggt cactccctcc gaagttttgc 600
ggtagatga gaagctccag agcctgcacg agggccggac gcccctccc accaagaac 660
caccgctggc gccccgctg ccccaacagc agtttgggtt cagctgcaa tacctcaaa 720
acaaaaatca agcgaactt atccccctg tgcctgagtt cacagtgcg tacctgagag 780
agaagggcct ggcaccggag ggcctgttcc ggagatccac caccgtccag accgaccgag 840
agatccagag gctctaacac caagggaaag ccgtgacct tgcagactac gggacactt 900
cactccctgc cplgactctg aaagccttcc tccggagagt gcccaagccg ctctgacct 960
tccaggccta cgaagcagatt ctgggatca cctgtgtgga gaccgactc cgcctactc 1020
gctgcccaca gatcttaccg agcctcccag agcaacaata cgttgtctc cgtctactca 1080
tggcttctct ccactcgggtg tccccggaga gcalcttaa caaatgaac agctataac 1140
lygctctgtt cttcgggctg aatttgatct ggcactccc ggggtctct tccctgagtg 1200
ccctgtgccc cctgaacatg ttaactgaac tgcctatgaa gtactatgaa agatcttca 1260
gcacccccga ggcacttggg uagcaccgac tggcaccatg ggaacaggg agcagggccg 1320
cccctttgca ggagcctgtg ccaagggcac aagccacagg cctcaccag cctaccctac 1380
ctccagctcc cctgatggca gccagaagac gctctatgt ttggaacac tctgtatgt 1440
tccagctacc tccacacct gctctgtcac ttgtatgtl tgtaaactg gcactgtaa 1500
aaataaccag ccattagatg nattcagaac cttctaagt 1539

```

```

<210> 30
<211> 433
<212> PRT
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<222> Incyta ID No: 3068538CD1

```

```

<400> 30
Met Ala Gly Gln Asp Pro Ala Leu Ser Thr Ser His Pro Phe Tyr
1 5 10 15
Asp Val Ala Arg His Gly Ile Leu Gln Val Ala Gly Asp Asp Arg
20 25 30
Phe Gly Arg Arg Val Thr Phe Ser Cys Cys Arg Met Pro Pro
35 40 45
Ser His Glu Leu Asp His Gln Arg Leu Leu Glu Tyr Leu Lys Tyr
50 55 60
Thr Leu Asp Gln Tyr Val Gln Asn Asp Tyr Thr Ile Val Tyr Phe
65 70 75
His Tyr Gly Leu Asn Ser Arg Asn Lys Pro Ser Leu Gly Trp Leu
80 85 90
Gln Ser Ala Tyr Lys Glu Phe Asp Arg Lys Tyr Lys Lys Asn Leu
95 100 105
Lys Ala Leu Tyr Val Val His Pro Thr Ser Phe Ile Lys Val Leu
110 115 120
Trp Asn Ile Leu Lys Pro Leu Ile Ser His Lys Phe Gly Lys Lys
125 130 135
Val Ile Tyr Phe Asn Tyr Leu Ser Glu Leu His Glu His Leu Lys
140 145 150
Tyr Asp Gln Leu Val Ile Pro Pro Glu Val Leu Arg Tyr Asp Glu
155 160 165
Lys Leu Gln Ser Leu His Glu Gly Arg Thr Pro Pro Pro Thr Lys
170 175 180
Thr Pro Pro Pro Arg Pro Pro Leu Pro Thr Gln Gln Phe Gly Val
185 190 195
Ser Leu Gln Tyr Leu Lys Asp Lys Asn Gln Gly Glu Leu Ile Pro

```

WO 01/61010

PCT/US01/05075

```

200          205          210
Pro Val Leu Arg Phe Thr Val Thr Tyr Leu Arg Glu Lys Gly Leu
225          220          225
Arg Thr Glu Gly Leu Phe Arg Arg Ser Ala Ser Val Gln Thr Val
230          235          240
Arg Glu Ile Gln Arg Leu Tyr Asn Gln Gly Lys Pro Val Asn Phe
245          250          255
Asp Asp Tyr Gly Asp Ile His Ile Pro Ala Val Ile Leu Lys Thr
260          265          270
Phe Met Arg Glu Leu Pro Gln Pro Leu Leu Thr Phe Gln Ala Tyr
275          280          285
Glu Gln Ile Leu Gly Ile Thr Cys Val Glu Ser Ser Leu Arg Val
290          295          300
Thr Gly Cys Arg Gln Ile Leu Arg Ser Leu Pro Glu His Asn Tyr
305          310          315
Val Val Leu Arg Tyr Leu Met Gly Phe Leu His Ala Val Ser Arg
320          325          330
Glu Ser Ile Phe Asn Lys Met Asn Ser Ser Asn Leu Ala Cys Val
335          340          345
Phe Gly Leu Asn Leu Ile Trp Pro Ser Gln Gly Val Ser Ser Leu
350          355          360
Ser Ala Leu Val Pro Leu Asn Met Phe Thr Glu Leu Leu Ile Glu
365          370          375
Tyr Tyr Glu Lys Ile Phe Ser Thr Pro Glu Ala Pro Gly Glu His
380          385          390
Gly Leu Ala Pro Trp Glu Gln Gly Ser Arg Ala Ala Pro Leu Gln
395          400          405
Glu Ala Val Pro Arg Thr Gln Ala Thr Gly Leu Thr Lys Pro Thr
410          415          420
Leu Pro Pro Ser Pro Leu Met Ala Ala Arg Arg Arg Leu
425          430          435
<210> 11
<211> 433
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No. 404424.5.pseq

<400> 31
Met Ala Gly Gln Asp Pro Ala Leu Ser Thr Ser His Pro Phe Tyr
1      5      10      15
Asp Val Ala Arg His Gly Ile Leu Gln Val Ala Gly Asp Asp Arg
20     25     30     35
Phe Gly Arg Arg Val Val Thr Phe Ser Cys Cys Arg Met Pro Pro
35     40     45     50
Ser His Glu Leu Asp His Gln Arg Leu Leu Glu Tyr Leu Lys Tyr
50     55     60     65
Thr Leu Asp Gln Tyr Val Glu Asn Asp Tyr Thr Ile Val Tyr Phe
65     70     75     80
His Tyr Gly Leu Asn Ser Arg Asn Lys Pro Ser Leu Gly Trp Leu
80     85     90     95
Gln Ser Ala Tyr Lys Glu Phe Asp Arg Lys Tyr Lys Lys Asn Leu
95    100   105   110
Lys Ala Leu Tyr Val Val His Pro Thr Ser Phe Ile Lys Val Leu
110   115   120   125
Trp Asn Ile Leu Lys Pro Leu Ile Ser His Lys Phe Gly Lys Lys
125   130   135   140
Val Ile Tyr Phe Asn Tyr Leu Ser Glu Leu His Glu His Leu Lys
140   145   150   155
Tyr Asp Gln Leu Val Ile Pro Pro Glu Val Leu Arg Tyr Asp Glu
155   160   165   170

```

WO 01/61010

PCT/US01/05075

```

155          160          165
Lys Leu Gln Ser Leu His Glu Gly Arg Thr Pro Pro Pro Thr Lys
170
Thr Pro Pro Pro Arg Pro Pro Leu Pro Thr Gln Gln Phe Gly Val
185
Ser Leu Gln Tyr Leu Lys Asp Lys Asn Gln Gly Glu Leu Ile Pro
200
Pro Val Leu Arg Phe Thr Val Thr Tyr Leu Arg Glu Lys Gly Leu
215
Arg Thr Glu Gly Leu Phe Arg Arg Ser Ala Ser Val Gln Thr Val
230
Arg Glu Ile Gln Arg Leu Tyr Asn Gln Gly Lys Pro Val Asn Phe
245
Asp Asp Tyr Gly Asp Ile His Ile Pro Ala Val Ile Leu Lys Thr
260
Phe Leu Arg Glu Leu Pro Gln Pro Leu Leu Thr Phe Gln Ala Tyr
275
Glu Gln Ile Leu Gly Ile Thr Cys Val Glu Ser Ser Leu Arg Val
290
Thr Arg Cys Arg Glu Ile Leu Arg Ser Leu Pro Glu His Asn Tyr
305
Val Val Leu Arg Tyr Leu Met Gly Phe Leu His Ala Val Ser Arg
320
Glu Ser Ile Phe Asn Lys Met Asn Ser Ser Asn Leu Ala Cys Val
335
Phe Gly Leu Asn Leu Ile Trp Pro Ser Gln Gly Val Ser Ser Leu
350
Ser Ala Leu Val Pro Leu Asn Met Phe Thr Glu Leu Leu Ile Glu
365
Tyr Tyr Glu Lys Ile Phe Ser Thr Pro Glu Ala Pro Gly Glu His
380
Gly Leu Ala Pro Trp Gln Gln Gly Ser Arg Ala Ala Pro Leu Gln
395
Glu Ala Val Pro Arg Thr Gln Ala Thr Gly Leu Thr Lys Pro Thr
410
Leu Pro Pro Ser Pro Leu Met Ala Ala Arg Arg Arg Leu
425          430
<210> 32
<211> 333
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: g6572185

<400> 32
Tyr Lys Lys Asn Leu Lys Ala Leu Tyr Val Val His Pro Thr Ser
1
1
Phe Ile Lys Val Leu Trp Asn Ile Leu Lys Pro Leu Ile Ser His
20
25
Lys Phe Gly Lys Lys Val Ile Tyr Phe Asn Tyr Leu Ser Glu Leu
35
40
His Glu His Leu Lys Tyr Asp Gln Leu Val Ile Pro Pro Glu Val
50
55
Leu Arg Tyr Asp Glu Lys Leu Gln Ser Leu His Glu Gly Arg Thr
65
70
Pro Pro Pro Thr Lys Thr Pro Pro Pro Arg Pro Pro Leu Pro Thr
80
85
Gln Gln Phe Gly Val Ser Leu Gln Tyr Leu Lys Asp Lys Asn Gln
95
100
Gly Glu Leu Ile Pro Pro Val Leu Arg Phe Thr Val Thr Tyr Leu

```

WO 01/61010

PCT/US01/05075

```

110          115          120
Arg Glu Lys Gly Leu Arg Thr Glu Gly Leu Phe Arg Arg Ser Ala
125          130          135
Ser Val Gln Thr Val Arg Glu Ile Gln Arg Leu Tyr Asn Gln Gly
140          145          150
Lys Pro Val Asn Phe Asp Asp Tyr Gly Asp Ile His Ile Pro Ala
155          160          165
Val Ile Leu Lys Thr Phe Leu Arg Glu Leu Pro Gln Pro Leu Leu
170          175          180
Thr Phe Gln Ala Tyr Glu Gln Ile Leu Gly Ile Thr Cys Val Glu
185          190          195
Ser Ser Leu Arg Val Thr Gly Cys Arg Gln Ile Leu Arg Ser Leu
200          205          210
Pro Glu His Asn Tyr Val Val Leu Arg Tyr Leu Met Gly Phe Hea
215          220          225
His Ala Val Ser Arg Glu Ser Ile Phe Asn Lys Met Asn Ser Ser
230          235          240
Asn Leu Ala Cys Val Phe Gly Leu Asn Leu Ile Trp Pro Ser Gln
245          250          255
Gly Val Ser Ser Leu Ser Ala Leu Val Pro Leu Asn Met Phe Thr
260          265          270
Glu Leu Leu Ile Glu Tyr Tyr Glu Lys Ile Phe Ser Thr Pro Glu
275          280          285
Ala Pro Gly Glu His Gly Leu Ala Pro Trp Glu Gln Gly Ser Arg
290          295          300
Ala Ala Pro Leu Gln Glu Ala Val Pro Arg Thr Gln Ala Thr Gly
305          310          315
Leu Thr Lys Pro Thr Leu Pro Pro Ser Pro Leu Met Ala Ala Arg
320          325          330
Arg Arg Leu

```

```

<210> 33
<211> 439
<212> PRT
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Incyte ID No: g312212

```

```

<400> 33
Met Asp Pro Leu Ser Glu Leu Gln Asp Asp Leu Thr Leu Asp Asp
1          5          10          15
Thr Ser Glu Ala Leu Asn Gln Leu Lys Leu Ala Ser Ile Asp Glu
20          25          30
Lys Asn Trp Pro Ser Asp Glu Met Pro Asp Phe Pro Lys Ser Asp
35          40          45
Asp Ser Lys Ser Ser Ser Pro Glu Leu Val Thr His Leu Lys Trp
50          55          60
Asp Asp Pro Tyr Tyr Asp Ile Ala Arg His Gln Ile Val Glu Val
65          70          75
Ala Gly Asp Asp Lys Tyr Gly Arg Lys Ile Ile Val Phe Ser Ala
80          85          90
Cys Arg Met Pro Pro Ser His Gln Leu Asp His Ser Lys Leu Leu
95          100          105
Gly Tyr Leu Lys His Thr Leu Asp Gln Tyr Val Glu Ser Asp Tyr
110          115          120
Thr Leu Leu Tyr Leu His His Gly Leu Thr Ser Asp Asn Lys Pro
125          130          135
Ser Leu Ser Trp Leu Arg Asp Ala Tyr Arg Glu Phe Asp Arg Lys
140          145          150
Tyr Lys Lys Asn Ile Lys Ala Leu Tyr Ile Val His Pro Thr Met

```

WO 01/61010

PCT/US01/05075

Phe	Ile	Lys	Thr	Leu	Leu	Ile	Leu	Phe	Lys	Pro	Leu	Ile	Ser	Phe	155	160	165
Lys	Phe	Gly	Gln	Lys	Ile	Phe	Tyr	Val	Asn	Tyr	Leu	Ser	Gln	Leu	170	175	180
Ser	Glu	His	Val	Lys	Leu	Glu	Gln	Leu	Gly	Ile	Pro	Arg	Gln	Val	185	190	195
Leu	Lys	Tyr	Asp	Asp	Phe	Leu	Lys	Ser	Thr	Gln	Lys	Ser	Pro	Ala	200	205	210
Thr	Ala	Pro	Lys	Pro	Met	Pro	Pro	Arg	Pro	Pro	Leu	Pro	Asn	Gln	215	220	225
Gln	Phe	Gly	Val	Ser	Leu	Gln	His	Leu	Gln	Glu	Lys	Asn	Pro	Glu	230	235	240
Gln	Glu	Pro	Ile	Pro	Ile	Val	Leu	Arg	Glu	Thr	Val	Ala	Tyr	Leu	245	250	255
Gln	Ala	His	Ala	Leu	Thr	Thr	Glu	Gly	Ile	Phe	Arg	Arg	Ser	Ala	260	265	270
Asn	Thr	Gln	Val	Val	Arg	Glu	Val	Gln	Gln	Lys	Tyr	Asn	Met	Gly	275	280	285
Leu	Pro	Val	Asp	Phe	Asp	Gln	Tyr	Asn	Glu	Leu	His	Leu	Pro	Ala	290	295	300
Val	Ile	Leu	Lys	Thr	Phe	Leu	Arg	Glu	Leu	Pro	Glu	Pro	Leu	Leu	305	310	315
Thr	Phe	Asp	Leu	Tyr	Pro	His	Val	Val	Gly	Phe	Leu	Asn	Ile	Asp	320	325	330
Glu	Ser	Gln	Arg	Val	Pro	Ala	Thr	Leu	Gln	Val	Leu	Gln	Thr	Leu	335	340	345
Pro	Glu	Glu	Asn	Tyr	Gln	Val	Leu	Arg	Phe	Leu	Thr	Ala	Phe	Leu	350	355	360
Val	Gln	Ile	Ser	Ala	His	Ser	Asp	Gln	Asn	Lys	Met	Thr	Asn	Thr	365	370	375
Asn	Leu	Ala	Val	Val	Phe	Gly	Pro	Asn	Leu	Leu	Trp	Ala	Lys	Asp	380	385	390
Ala	Ala	Ile	Thr	Leu	Lys	Ala	Ile	Asn	Pro	Ile	Asn	Thr	Phe	Thr	395	400	405
Lys	Phe	Leu	Leu	Asp	His	Gln	Gly	Glu	Leu	Phe	Pro	Ser	Pro	Asp	410	415	420
Pro	Ser	Gly	Leu												425	430	435

## 【国際公開パンフレット(コレクトバージョン)】

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization  
International Bureau(43) International Publication Date  
23 August 2001 (23.08.2001)

PCT

(10) International Publication Number  
WO 01/61010 A3

- (51) International Patent Classification: C12N 15/12, C07K 1447, C12Q 168, G01N 33/68, A61K 38/00, C07K 16/18, C12N 15/11, A01K 67/027
- (74) Agents: TURNER, Christopher et al.; Incyte Genomics, Inc., 3160 Porter Drive, Palo Alto, CA 94304 (US).
- (81) Designated States (national): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BC, BR, BY, BZ, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MY, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.
- (84) Designated States (regional): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LI, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CI, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).
- (21) International Application Number: PCT/US01/03075
- (22) International Filing Date: 15 February 2001 (15.02.2001)
- (75) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data: 09/507,765 18 February 2000 (18.02.2000) US
- (71) Applicant (for all designated States except US): INCYTE GENOMICS, INC. [US/US]; 3160 Porter Drive, Palo Alto, CA 94304 (US).
- (72) Inventors; and
- (73) Inventors/Applicants (for US only): KLINGLER, Fred, M. [US/US]; 38 Dover Court, San Carlos, CA 94070 (US); STEWART, Elizabeth, A. [US/US]; 1678 Latham Street, Mountain View, CA 94041 (US); YLE, Henry [US/US]; 826 Lois Avenue, Sunnyvale, CA 94087 (US); BAUGHN, Mariah, R. [US/US]; 14244 Santiago Road, San Leandro, CA 94527 (US).
- Published:  
— with international search report
- (88) Date of publication of the international search report: 25 April 2002
- For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.



WO 01/61010 A3

(54) Title: GTPASE ACTIVATING PROTEINS

(57) Abstract: The invention provides mammalian nucleic acid molecules and fragments thereof. It also provides for the use of the mammalian nucleic acid molecules for the characterization, diagnosis, evaluation, treatment, or prevention of conditions, diseases and disorders associated with cell signaling, the immune system, and cell proliferation, particularly with cancer. The invention additionally provides expression vectors and host cells for the production of the protein encoded by the mammalian nucleic acid molecules.

## 【 国際調査報告 】

INTERNATIONAL SEARCH REPORT		International Application No. PCT/US 01/05075
A. CLASSIFICATION OF SUBJECT MATTER		
IPC 7	C12N15/12 C07K16/18	C07K14/47 C12N15/11 C12Q1/68 A61K38/00 G01N33/68 A61K38/00 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 C12Q G01N A61K A01K
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)		
EPO-internal, WPI Data, PAJ, BIOSIS		
C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	BARFOOT ET AL: "Cloning and expression of a human CDC42 GTPase-activating protein reveals a functional SH3-binding domain." JOURNAL OF BIOLOGICAL CHEMISTRY, vol. 268, no. 35, 1993, pages 26059-26062, XP002177181 ISSN: 0021-9258 the whole document	1,3,4, 6-14, 16-18
X	DATABASE EM EST [Online] ENBL Heidelberg, Germany; AC/ID A1188551, 14 November 1998 (1998-11-14) NCI-CGAP: "GTPase-activating protein" XP002177183 abstract	1,3,4, 6-12
	---	
	--- --	
<input checked="" type="checkbox"/> Further documents are listed in the continuation of box C. <input checked="" type="checkbox"/> Parent family members are listed in annex.		
* Special categories of cited documents:		
<p>"A" document defines the general state of the art which is not considered to be of particular relevance</p> <p>"B" special document but published on or after the international filing date</p> <p>"L" document which may throw doubts on priority claims of which is cited to establish the publication date of another claim or other special reason has specified</p> <p>"O" document relating to an oral disclosure, use, exhibition or other means</p> <p>"P" document published prior to the international filing date but after the priority date claimed</p> <p>"T" later document published after the international filing date or priority date and not identified with the application but cited to understand the principle of the invention</p> <p>"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</p> <p>"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.</p> <p>"Z" documents members of the same parent family</p>		
Date of the actual completion of the international search		Date of mailing of the international search report
11 September 2001		15. 01. 02
Name and mailing address of the ISA European Patent Office, P.O. Box 5518, Palisade 2 CH - 2200 HV Rivier Tel: +31-70 340-2042, Telex: 316514 EPON Fax: +31-70 340-3016		Authorised officer Oderwald, H

Form PCT/ISA/C18 (pre-01.06.00) Sheet 1 (of 1)

## INTERNATIONAL SEARCH REPORT

International Application No.  
PCT/US 01/05075

C. (Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT	
Category*	Relevant to claim No.
A	LAMARCHE N ET AL: "GAPS FOR RHO-RELATED GTPASES" TRENDS IN GENETICS, ELSEVIER SCIENCE PUBLISHERS B.V. AMSTERDAM, NL, vol. 10, no. 12, 1 December 1994 (1994-12-01), pages 436-440, XP002070373 ISSN: 0168-9525 the whole document
A	CICCHETTI P ET AL: "IDENTIFICATION OF A PROTEIN THAT BINDS TO THE SH3 REGION OF ABI AND IS SIMILAR TO BCR AND GAP-RHO" SCIENCE, AMERICAN ASSOCIATION FOR THE ADVANCEMENT OF SCIENCE, US, vol. 257, 7 August 1992 (1992-08-07), pages 803-806, XP000578247 ISSN: 0036-8075 the whole document
A	MAHESHWAR M M ET AL: "The GAP-related domain of tuberlin, the product of the TSC2 gene, is a target for missense mutations in tuberous sclerosis." HUMAN MOLECULAR GENETICS, vol. 6, no. 11, 1997, pages 1991-1996, XP002177182 ISSN: 0964-6906 the whole document
P,X, L	WO 01 05970 A (INCYTE GENOMICS INC ; PATTERSON CHANDRA (US); AZIMZAI YALDA (US); R) 25 January 2001 (2001-01-25) L: priority see SEQ ID NO: 22 (Fig. 21, 22) and SEQ ID NO: 88 (Fig. 83) page 4, line 35 -page 58, line 29; claims 1-7,9,11-17,25,26
	1,3,4, 6-14, 16-19

2

Form PCT/ISA210 (continuation of second sheet) (July 1997)

INTERNATIONAL SEARCH REPORT		International application No. PCT/US 01/05075
<b>Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)</b>		
This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:		
1.	<input type="checkbox"/>	Claims Nos.: because they relate to subject matter not required to be searched by this Authority, namely:
2.	<input type="checkbox"/>	Claims Nos.: because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:
3.	<input type="checkbox"/>	Claims Nos.: because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).
<b>Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)</b>		
This International Searching Authority found multiple inventions in this international application, as follows:		
see additional sheet		
1.	<input type="checkbox"/>	As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.
2.	<input type="checkbox"/>	As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3.	<input type="checkbox"/>	As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:
4.	<input checked="" type="checkbox"/>	No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.: Claims 1, 14, 16-19 completely, 3, 4, 6-13 partially.
Remark on Protest		
	<input type="checkbox"/>	The additional search fees were accompanied by the applicant's protest.
	<input type="checkbox"/>	No protest accompanied the payment of additional search fees.

International Application No. PCT/US 01/05075

## FURTHER INFORMATION CONTINUED FROM PCTISA/ 210

This International Searching Authority found multiple (groups of) inventions in this international application, as follows:

1. Claims: {1, 14, 16-19 complete}, {3, 4, 6-13 partially}

A purified nucleic acid molecule as set forth in SEQ ID NO: 28 encoding a GTPase activating protein as set forth in SEQ ID NO: 30. A protein with an amino acid sequence as set forth in SEQ ID NO: 30. A probe, a recombinant nucleic acid molecule comprising a promoter operably linked to said nucleic acid molecule, a transformed cell, a method of producing a polypeptide, methods of detection, methods to screen a library, a method for purifying a ligand, a pharmaceutical composition comprising said nucleic acid and protein.

2. Claims: {2, 15 complete}, {3, 4, 6-13 partially}

A variant of a nucleic acid molecule as set forth in SEQ ID NO: 29 encoding a GTPase activating protein as set forth in SEQ ID NO: 31. A probe, a recombinant nucleic acid molecule comprising a promoter operably linked to said nucleic acid molecule, a transformed cell, a method of producing a polypeptide, methods of detection, methods to screen a library, a method for purifying a ligand comprising said nucleic acid.

3. Claims: {3, 4, 6-13 partially}

A nucleic acid molecule as set forth in SEQ ID NO: 1. A probe, a recombinant nucleic acid molecule comprising a promoter operably linked to said nucleic acid molecule, a transformed cell, a method of producing a polypeptide, methods of detection, methods to screen a library, a method for purifying a ligand comprising said nucleic acid.

4. Claims: {3, 4, 6-13 partially}

same as invention 3 but comprising SEQ ID NO: 2.

5. Claims: {3, 4, 6-13 partially}

same as invention 3 but comprising SEQ ID NO: 3.

6. Claims: {3, 4, 6-13 partially}

same as invention 3 but comprising SEQ ID NO: 4.

7. Claims: {3, 4, 6-13 partially}

International Application No. PCT/US 01/05075

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

- same as invention 3 but comprising SEQ ID NO: 5.
8. Claims: {3, 4, 6-13 partially}  
same as invention 3 but comprising SEQ ID NO: 6.
9. Claims: {3, 4, 6-13 partially}  
same as invention 3 but comprising SEQ ID NO: 7.
10. Claims: {3, 4, 6-13 partially}  
same as invention 3 but comprising SEQ ID NO: 8.
11. Claims: {3, 4, 6-13 partially}  
same as invention 3 but comprising SEQ ID NO: 9.
12. Claims: {3, 4, 6-13 partially}  
same as invention 3 but comprising SEQ ID NO: 10.
13. Claims: {3, 4, 6-13 partially}  
same as invention 3 but comprising SEQ ID NO: 11.
14. Claims: {3, 4, 6-13 partially}  
same as invention 3 but comprising SEQ ID NO: 12.
15. Claims: {3, 4, 6-13 partially}  
same as invention 3 but comprising SEQ ID NO: 13.
16. Claims: {3, 4, 6-13 partially}  
same as invention 3 but comprising SEQ ID NO: 14.
17. Claims: {3, 4, 6-13 partially}  
same as invention 3 but comprising SEQ ID NO: 15.
18. Claims: {3, 4, 6-13 partially}

International Application No. PCT/US 01/05075

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

- same as invention 3 but comprising SEQ ID NO: 16.
19. Claims: {3, 4, 6-13 partially}  
same as invention 3 but comprising SEQ ID NO: 17.
20. Claims: {3, 4, 6-13 partially}  
same as invention 3 but comprising SEQ ID NO: 18.
21. Claims: {3, 4, 6-13 partially}  
same as invention 3 but comprising SEQ ID NO: 19.
22. Claims: {3, 4, 6-13 partially}  
same as invention 3 but comprising SEQ ID NO: 20.
23. Claims: {3, 4, 6-13 partially}  
same as invention 3 but comprising SEQ ID NO: 21.
24. Claims: {3, 4, 6-13 partially}  
same as invention 3 but comprising SEQ ID NO: 22.
25. Claims: {3, 4, 6-13 partially}  
same as invention 3 but comprising SEQ ID NO: 23.
26. Claims: {3, 4, 6-13 partially}  
same as invention 3 but comprising SEQ ID NO: 24.
27. Claims: {3, 4, 6-13 partially}  
same as invention 3 but comprising SEQ ID NO: 25.
28. Claims: {3, 4, 6-13 partially}  
same as invention 3 but comprising SEQ ID NO: 26.
29. Claims: {3, 4, 6-13 partially}

International Application No. PCT/US 01/05075

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

same as invention 3 but comprising SEQ ID NO: 27.

30. Claim : {5 complete}

A single nucleotide polymorphism comprising a substitution at position 562, 1013 or 1020 of SEQ ID NO: 28.

**INTERNATIONAL SEARCH REPORT**  
Information on patent family members

International Application No.  
PCT/US 01/05075

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 0105970 A	25-01-2001	AU 6223700 A WO 0105970 A2	05-02-2001 25-01-2001

## フロントページの続き

(51) Int.Cl. <sup>7</sup>	F I	テーマコード(参考)
C 0 7 K 14/47	C 1 2 N 1/15	4 C 0 8 4
C 1 2 N 1/15	C 1 2 N 1/19	4 H 0 4 5
C 1 2 N 1/19	C 1 2 N 1/21	
C 1 2 N 1/21	C 1 2 P 21/02	C
C 1 2 N 5/10	C 1 2 Q 1/68	A
C 1 2 P 21/02	G 0 1 N 33/15	Z
C 1 2 Q 1/68	G 0 1 N 33/50	Z
G 0 1 N 33/15	G 0 1 N 33/53	D
G 0 1 N 33/50	G 0 1 N 33/53	M
G 0 1 N 33/53	G 0 1 N 33/566	
G 0 1 N 33/566	C 1 2 N 5/00	A
	A 6 1 K 37/02	

(81) 指定国 AP(GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), EA(AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), EP(AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OA(BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG), AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW

(72) 発明者 ユエ、ヘンリー

アメリカ合衆国カリフォルニア州 9 4 0 8 7 ・ サニーベイル・ルイスアベニュー 8 2 6

(72) 発明者 ボーグン、マライア・アール

アメリカ合衆国カリフォルニア州 9 4 5 7 7 ・ サンレアンドロ・サンティアゴロード 1 4 2 4 4

F ターム(参考) 2G045 AA40 DA12 DA13 DA14 DA36 FB02 FB03  
 4B024 AA01 AA11 BA80 CA04 CA09 CA11 DA02 DA05 EA02 EA04  
 GA11 GA18 GA19 HA03 HA14  
 4B063 QA01 QA18 QQ20 QQ42 QR32 QR35 QR38 QR48 QR55 QR72  
 QR82 QS11 QS25 QS34 QS39 QX02 QX07  
 4B064 AG01 CA02 CA10 CA19 CC24 DA04 DA14  
 4B065 AA01X AA90X AA93Y AB01 AC14 BA01 CA24 CA44 CA46  
 4C084 AA01 AA02 AA06 AA07 BA01 BA22 CA53 DC50 ZB072 ZB212  
 ZB262  
 4H045 AA10 AA20 AA30 BA10 CA40 DA86 EA22 EA28 EA51 FA74

专利名称(译)	GTP酶活化蛋白		
公开(公告)号	<a href="#">JP2004500822A</a>	公开(公告)日	2004-01-15
申请号	JP2001560380	申请日	2001-02-15
[标]申请(专利权)人(译)	洞察Genomics公司		
申请(专利权)人(译)	洞察基因组公司		
[标]发明人	クリングラートッドエム スチュアートエリザベスエイ ユエヘンリー ボーグンマライアール		
发明人	クリングラー、トッド・エム スチュアート、エリザベス・エイ ユエ、ヘンリー ボーグン、マライア・アール		
IPC分类号	G01N33/50 A61K38/00 A61P35/00 A61P37/02 A61P43/00 C07K14/47 C12N1/15 C12N1/19 C12N1/21 C12N5/10 C12N15/09 C12N15/12 C12P21/02 C12Q1/68 G01N33/15 G01N33/53 G01N33/566		
CPC分类号	C07K14/4722 A01K2217/05 A01K2217/075 A61K38/00		
FI分类号	C12N15/00.ZNA.A A61P35/00 A61P37/02 A61P43/00.105 C07K14/47 C12N1/15 C12N1/19 C12N1/21 C12P21/02.C C12Q1/68.A G01N33/15.Z G01N33/50.Z G01N33/53.D G01N33/53.M G01N33/566 C12N5/00.A A61K37/02		
F-TERM分类号	2G045/AA40 2G045/DA12 2G045/DA13 2G045/DA14 2G045/DA36 2G045/FB02 2G045/FB03 4B024 /AA01 4B024/AA11 4B024/BA80 4B024/CA04 4B024/CA09 4B024/CA11 4B024/DA02 4B024/DA05 4B024/EA02 4B024/EA04 4B024/GA11 4B024/GA18 4B024/GA19 4B024/HA03 4B024/HA14 4B063 /QA01 4B063/QA18 4B063/QQ20 4B063/QQ42 4B063/QR32 4B063/QR35 4B063/QR38 4B063/QR48 4B063/QR55 4B063/QR72 4B063/QR82 4B063/QS11 4B063/QS25 4B063/QS34 4B063/QS39 4B063 /QX02 4B063/QX07 4B064/AG01 4B064/CA02 4B064/CA10 4B064/CA19 4B064/CC24 4B064/DA04 4B064/DA14 4B065/AA01X 4B065/AA90X 4B065/AA93Y 4B065/AB01 4B065/AC14 4B065/BA01 4B065/CA24 4B065/CA44 4B065/CA46 4C084/AA01 4C084/AA02 4C084/AA06 4C084/AA07 4C084 /BA01 4C084/BA22 4C084/CA53 4C084/DC50 4C084/ZB072 4C084/ZB212 4C084/ZB262 4H045 /AA10 4H045/AA20 4H045/AA30 4H045/BA10 4H045/CA40 4H045/DA86 4H045/EA22 4H045/EA28 4H045/EA51 4H045/FA74		
优先权	09/507765 2000-02-18 US		
外部链接	<a href="#">Espacenet</a>		

#### 摘要(译)

本发明提供哺乳动物核酸分子及其片段。本发明还提供了使用该哺乳动物核酸分子表征，诊断，评估，治疗或预防与信息传递，免疫系统和细胞增殖，特别是结肠癌有关的症状，疾病和病症的方法。到。此外，本发明提供了用于产生由哺乳动物核酸分子编码的蛋白质的表达载体和宿主细胞。

57	ACC	CGG	CAC	GCA	GGT	GCG	GGC	CTG	CGG	GGT	CGG	TGG	CCA	GAG	CTG	CAG	AGA				
64	AAG	GCG	GCG	GCG	GCT	GCT	GTV	CTG	GGT	GCA	GTG	AGG	AAG	AGG	CCC	TGG	GTC				
118	CGC	ATG	GCT	AGC	CAG	GAT	CGT	CGG	CTG	AGC	ACG	AGT	CAC	CGG	TTC	TAC	GAC				
			M	A	G	Q	D	P	R	L	S	T	S	K	P	P	Y	E			
172	GCC	AGA	CAT	GGC	ATT	CTG	CAG	GTC	GCA	GAG	GAT	GAC	CGC	TTT	GAA	AGA	CGT				
			J	A	R	H	G	I	L	Q	V	A	G	D	D	R	F	G	R	R	
226	CTC	ACG	TTT	AGC	AGC	TGC	CGG	ATG	CCA	CGC	TCC	CAC	GAG	CTG	GAC	CAC	CAG				
			J	V	T	F	S	C	C	R	M	P	P	S	H	E	L	D	H	Q	
280	CTG	CTG	GAG	TAT	TTT	AGG	TAC	ACA	CTG	GAC	CAA	TAC	GTT	GAG	AAC	GAT	TAT				
			K	L	L	S	Y	L	K	Y	T	T	L	D	Q	Y	V	E	N	D	Y