(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2004-507225
(P2004-507225A)
(43)公表日　平成16年3月11日(2004.3.11)

(51)Int.Cl.[7]	FI			テーマコード（参考）
C12N 15/09	C12N	15/00	ZNAA	2G045
A61K 38/00	A61K	45/00		4B024
A61K 45/00	A61K	49/00	A	4B063
A61K 49/00	A61P	1/04		4B064
A61P 1/04	A61P	1/16		4B065

審査請求 未請求　予備審査請求 有　（全 193 頁）　最終頁に続く

(21)出願番号　特願2002-500719(P2002-500719)
(86)(22)出願日　平成13年5月30日(2001.5.30)
(85)翻訳文提出日　平成14年11月29日(2002.11.29)
(86)国際出願番号　PCT/US2001/017581
(87)国際公開番号　WO2001/092527
(87)国際公開日　平成13年12月6日(2001.12.6)
(31)優先権主張番号　60/209,407
(32)優先日　平成12年6月1日(2000.6.1)
(33)優先権主張国　米国(US)
(31)優先権主張番号　60/250,326
(32)優先日　平成12年11月30日(2000.11.30)
(33)優先権主張国　米国(US)

(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
(74)代理人　100089266
弁理士　大島　陽一
(72)発明者　タング、ワイ・トム
アメリカ合衆国カリフォルニア州９５１１８・サンノゼ・ランウィックコート　４２３０

最終頁に続く

(54)【発明の名称】アポトーシス調節因子

(57)【要約】
本発明はヒトアポトーシス調節因子（ＡＰＲＧ）およびＡＰＲＧを同定し、コードするポリヌクレオチドを提供する。本発明はまた、発現ベクター、宿主細胞、抗体、アゴニストおよびアンタゴニストをも提供する。本発明はまた、ＡＰＲＧの異常発現に関連する疾患を診断、治療または予防する方法をも提供する。

【特許請求の範囲】
【請求項１】
以下の（ａ）乃至（ｄ）からなる群から選択した単離されたポリペプチド。
（ａ）配列番号１乃至３（ＳＥＱ ＩＤ ＮＯ：１－３）からなる群から選択したアミノ酸配列を含むポリペプチド
（ｂ）ＳＥＱ ＩＤ ＮＯ：１－３からなる群から選択したアミノ酸配列と少なくとも９０％が同一であるようなアミノ酸配列を有する天然のポリペプチド
（ｃ）ＳＥＱ ＩＤ ＮＯ：１－３からなる群から選択したアミノ酸配列を有するポリペプチドの生物学的活性断片
（ｄ）ＳＥＱ ＩＤ ＮＯ：１－３からなる群から選択したアミノ酸配列を有するポリペプチドの免疫抗原性断片
【請求項２】
ＳＥＱ ＩＤ ＮＯ：１－３からなる群から選択した請求項１に記載の単離されたポリペプチド。
【請求項３】
請求項１のポリペプチドをコードする単離されたポリヌクレオチド。
【請求項４】
請求項２のポリペプチドをコードする単離されたポリヌクレオチド。
【請求項５】
配列番号４乃至６（ＳＥＱ ＩＤ ＮＯ：４－６）からなる群から選択した請求項４に記載の単離されたポリヌクレオチド。
【請求項６】
請求項３に記載のポリヌクレオチドに機能的に結合したプロモーター配列を含む組換えポリヌクレオチド。
【請求項７】
請求項６に記載の組換えポリヌクレオチドを用いて形質転換した細胞。
【請求項８】
請求項６に記載の組換えポリヌクレオチドを含む遺伝形質転換体。
【請求項９】
請求項１のポリペプチドを生産する方法であって、
（ａ）前記ポリペプチドの発現に好適な条件下で、請求項１のポリペプチドをコードするポリヌクレオチドに機能的に結合されたプロモーター配列を含む組換えポリヌクレオチドで形質転換された細胞を培養する過程と、
（ｂ）そのように発現した前記ポリペプチドを回収する過程。
【請求項１０】
請求項１に記載のポリペプチドと特異結合するような単離された抗体。
【請求項１１】
以下の（ａ）乃至（ｄ）からなる群から選択した単離されたポリヌクレオチド。
（ａ）配列番号４乃至６（ＳＥＱ ＩＤ ＮＯ：４－６）からなる群から選択したポリヌクレオチド配列を有するポリヌクレオチド
（ｂ）ＳＥＱ ＩＤ ＮＯ：４－６からなる群から選択したポリヌクレオチド配列と少なくとも９０％が同一であるような天然のポリヌクレオチド
（ｃ）（ａ）のポリヌクレオチドに相補的なポリヌクレオチド
（ｄ）（ｂ）のポリヌクレオチドに相補的なポリヌクレオチド
（ｅ）（ａ）～（ｄ）のＲＮＡ等価物
【請求項１２】
請求項１１に記載のポリヌクレオチドの少なくとも６０の連続したヌクレオチドを含む単離されたポリヌクレオチド。
【請求項１３】
請求項１１に記載のポリヌクレオチドの配列を有する標的ポリヌクレオチドをサンプル中

から検出する方法であって、
（ａ）前記サンプル中の前記標的ポリヌクレオチドに相補的な配列を有する少なくとも２０の連続したヌクレオチドを含むプローブを用いて前記サンプルをハイブリダイズする過程（ただし、前記プローブと前記標的ポリヌクレオチドまたは断片の間でハイブリダイゼーション複合体が形成されるような条件下で、前記プローブが前記標的ポリヌクレオチドに特異的にハイブリダイズする）と、
（ｂ）前記ハイブリダイゼーション複合体の存在・不存在を検出し、該複合体が存在する場合にはオプションでその量を検出する過程からなり、 前記プローブと前記標的ポリヌクレオチドの間でハイブリダイゼーション複合体が形成されるような条件下で、前記プローブが前記標的ポリヌクレオチドに特異的にハイブリダイズすることを特徴とする方法。

【請求項１４】
前記プローブが少なくとも６０の連続したヌクレオチドを含むことを特徴とする請求項１３に記載の方法。

【請求項１５】
請求項１１に記載のポリヌクレオチドの配列を有する標的ポリヌクレオチドをサンプル中から検出する方法であって、
（ａ）ポリメラーゼ連鎖反応増幅を用いて前記標的ポリヌクレオチドまたはその断片を増幅する過程と、
（ｂ）前記標的ポリヌクレオチドまたはその断片の存在・不存在を検出し、該標的ポリヌクレオチドまたはその断片が存在する場合にはオプションでその量を検出する過程を含むことを特徴とする方法。

【請求項１６】
請求項１のポリペプチドと、薬剤として許容できる賦形剤とを含む組成物。

【請求項１７】
前記ポリペプチドが、ＳＥＱ ＩＤ ＮＯ：１－３からなる群から選択したアミノ酸配列を含むことを特徴とする請求項１６に記載の成分。

【請求項１８】
機能的なＡＰＲＧの発現の低下に関連する疾患や病態の治療方法であって、そのような治療が必要な患者に請求項１６の組成物を投与することを含むことを特徴とする治療方法。

【請求項１９】
請求項１に記載のポリペプチドのアゴニストとして有効性を確認するために化合物をスクリーニングする方法であって、
（ａ）請求項１のポリペプチドを含むサンプルを化合物に暴露する過程と、
（ｂ）前記サンプルにおいてアゴニスト活性を検出するステップとを含むことを特徴とするスクリーニング方法。

【請求項２０】
請求項１９に記載の方法によって同定したアゴニスト化合物と、薬剤として許容できる賦形剤とを含むことを特徴とする組成物。

【請求項２１】
機能的なＡＰＲＧの発現の低下に関連する疾患や病態の治療方法であって、そのような治療が必要な患者に請求項２０の組成物を投与することを含むことを特徴とする治療方法。

【請求項２２】
請求項１に記載のポリペプチドのアンタゴニストとして有効性を確認するために化合物をスクリーニングする方法であって、
（ａ）請求項１のポリペプチドを含むサンプルを化合物に暴露する過程と、
（ｂ）前記サンプルにおいてアンタゴニスト活性を検出するステップとを含むことを特徴とするスクリーニング方法。

【請求項２３】
請求項２２に記載の方法によって同定したアンタゴニスト化合物と、薬剤として許容できる賦形剤とを含む組成物。

【請求項２４】
機能的なＡＰＲＧの過剰な発現に関連する疾患や病態の治療方法であって、そのような治療が必要な患者に請求項２３の組成物を投与することを含むことを特徴とする治療方法。
【請求項２５】
請求項１に記載のポリペプチドに特異結合する化合物をスクリーニングする方法であって、
（ａ）適切な条件下で請求項１に記載のポリペプチドを少なくとも１つの試験化合物に結合させる過程と、
（ｂ）請求項１に記載のポリペプチドの試験化合物との結合を検出し、それによって請求項１に記載のポリペプチドに特異結合する化合物を同定する過程とを含むことを特徴とする方法。
【請求項２６】
請求項１に記載のポリペプチドの活性を調節する化合物をスクリーニングする方法であって、
（ａ）請求項１に記載のポリペプチドの活性が許容された条件下で、請求項１に記載のポリペプチドを少なくとも１つの試験化合物に結合させる過程と、
（ｂ）請求項１に記載のポリペプチドの活性を試験化合物の存在下で算定する過程と、
（ｃ）試験化合物の存在下での請求項１に記載のポリペプチドの活性を、試験化合物の不存在下での請求項１に記載のポリペプチドの活性と比較する過程とを含み、試験化合物の存在下での請求項１に記載のポリペプチドの活性の変化が、請求項１に記載のポリペプチドの活性を調節する化合物を標示することを特徴とする方法。
【請求項２７】
請求項５の配列を含む標的ポリヌクレオチドの発現を変化させるのに効果的な化合物をスクリーニングする方法であって、
（ａ）前記標的ポリヌクレオチドの発現に好適な条件下で、前記標的ポリヌクレオチドを含むサンプルを化合物に曝露する過程と、
（ｂ）前記標的ポリヌクレオチドの変異発現を検出する過程と、
（ｃ）可変量の前記化合物の存在下と前記化合物の不存在下で、前記標的ポリヌクレオチドの発現を比較する過程とを含むことを特徴とする方法。
【請求項２８】
試験化合物の毒性を算定する方法であって、
（ａ）核酸を含む生物学的サンプルを前記試験化合物で処理する過程と、
（ｂ）処理した前記生体サンプルの核酸と、請求項１１のポリヌクレオチドの少なくとも２０の連続するヌクレオチドを含むプローブをハイブリダイズさせるステップであって、このハイブリダイゼーションゼーションが、前記プローブと前記生体サンプルの標的ポリヌクレオチドとの間で特異的なハイブリダイゼーション複合体が形成される条件下で行われ、前記標的ポリヌクレオチドが、請求項１１のポリヌクレオチドのポリヌクレオチド配列またはその断片を含むポリヌクレオチドである、前記ステップと、
（ｃ）ハイブリダイゼーション複合体の収量を定量するステップと、
（ｄ）前記処理された生物学的サンプル中の前記ハイブリタイゼーション複合体の量を、処理されていない生物学的サンプル中の前記ハイブリタイゼーション複合体の量と比較する過程とを含み、前記処理された生物学的サンプル中の前記ハイブリタイゼーション複合体の量の差が、前記試験化合物の毒性を標示することを特徴とする方法。
【請求項２９】
生物学的サンプル中のＡＰＲＧの発現に関連する症状または疾患に対する診断試験法であって、
（ａ）前記抗体が前記ポリペプチドに結合し、抗体とポリペプチドとの複合体が形成されるのに適した条件下で、前記生物学的サンプルを請求項１０に記載の抗体と結合する過程と、
（ｂ）前記複合体を検出する過程とを含み、前記複合体の存在が、前記生物学的サンプル

中の前記ポリペプチドの存在と相関することを特徴とする方法。
【請求項３０】
前記抗体が、
（ａ）キメラ抗体
（ｂ）単鎖抗体
（ｃ）Ｆａｂ断片
（ｄ）Ｆ（ａｂ’）$_2$ 断片
（ｅ）ヒト化抗体 のいずれかであることを特徴とする請求項１０に記載の抗体。
【請求項３１】
請求項１０に記載の抗体と、許容できる賦形剤とを含む化合物。
【請求項３２】
被検者のＡＰＲＧの発現に関連する病状または疾患の診断方法であって、請求項３１に記載の化合物の有効量を前記被検者に投与する過程を含むことを特徴とする方法。
【請求項３３】
前記抗体が標識されることを特徴とする請求項３１に記載の化合物。
【請求項３４】
被検者のＡＰＲＧの発現に関連する病状または疾患の診断方法であって、請求項３３に記載の化合物の有効量を前記被検者に投与する過程を含むことを特徴とする方法。
【請求項３５】
請求項１０に記載の抗体の特異性を有するポリクローナル抗体を調製する方法であって、
（ａ）抗体反応を誘発する条件下で、ＳＥＱ ＩＤ ＮＯ：１－３からなる群から選択したアミノ酸配列含むポリペプチドまたはその免疫抗原性断片を用いて動物を免疫化する過程と、
（ｂ）前記動物から抗体を単離する過程と、
（ｃ）前記単離された抗体をポリペプチドでスクリーニングし、それによって、ＳＥＱ ＩＤ ＮＯ：１－３からなる群から選択したアミノ酸配列を有するポリペプチドに特異結合するようなポリクローナル抗体を同定する過程とを含むことを特徴とする方法。
【請求項３６】
請求項３５に記載の方法で産出した抗体。
【請求項３７】
請求項３６に記載の抗体及び適切なキャリアを含む化合物。
【請求項３８】
請求項１０に記載の抗体の特異性を有する抗体を用いてモノクローナル抗体を製造する方法であって、
（ａ）抗体反応を誘発する条件下で、ＳＥＱ ＩＤ ＮＯ：１－３からなる群から選択したアミノ酸配列含むポリペプチドまたはその免疫抗原性断片を用いて動物を免疫化する過程と、
（ｂ）前記動物から抗体産出細胞を単離する過程と、
（ｃ）前記抗体産出細胞を不死化の細胞と融合して、モノクローナル抗体を産出するハイブリドーマ細胞を形成する過程と、
（ｄ）前記ハイブリドーマ細胞を培養する過程と、
（ｅ）ＳＥＱ ＩＤ ＮＯ：１－３からなる群から選択したアミノ酸配列を有するポリペプチドに特異結合するような前記培養モノクローナル抗体から単離する過程とを含むことを特徴とする方法。
【請求項３９】
請求項３８に記載の方法で産出したモノクローナル抗体。
【請求項４０】
請求項３９に記載の抗体及び適切なキャリアを含む化合物。
【請求項４１】
Ｆａｂ発現ライブラリのスクリーニングにより前記抗体を産出することを特徴とする請求

項１０に記載の抗体。
【請求項４２】
組換え免疫グロブリンライブラリのスクリーニングにより前記抗体を産出することを特徴とする請求項１０に記載の抗体。
【請求項４３】
ＳＥＱ ＩＤ ＮＯ：１－３からなる群から選択したアミノ酸配列を有するポリペプチドを検出する方法であって、
（ａ）前記抗体と前記ポリペプチドの特異結合を許容する条件下で、サンプルを用いて請求項１０に記載の抗体をインキュベートする過程と、
（ｂ）特異結合を検出する過程とを含み、該特異結合が、ＳＥＱ ＩＤ ＮＯ：１－３からなる群から選択したアミノ酸配列を有するポリペプチドがサンプル中に存在することを示唆することを特徴とする方法。
【請求項４４】
ＳＥＱ ＩＤ ＮＯ：１－３からなる群から選択したアミノ酸配列を有するポリペプチドを精製する方法であって、
（ａ）前記抗体と前記ポリペプチドの特異結合を許容する条件下で、サンプルを用いて請求項１０に記載の抗体をインキュベートする過程と、
（ｂ）前記サンプルから前記抗体を分離し、ＳＥＱ ＩＤ ＮＯ：１－３からなる群から選択したアミノ酸配列を有する精製ポリペプチドを得る過程とを含むことを特徴とする方法。
【請求項４５】
ＳＥＱ ＩＤ ＮＯ：１のアミノ酸配列を有する請求項１に記載のポリペプチド。
【請求項４６】
ＳＥＱ ＩＤ ＮＯ：２のアミノ酸配列を有する請求項１に記載のポリペプチド。
【請求項４７】
ＳＥＱ ＩＤ ＮＯ：３のアミノ酸配列を有する請求項１に記載のポリペプチド。
【請求項４８】
ＳＥＱ ＩＤ ＮＯ：４のアミノ酸配列を有する請求項１１に記載のポリペプチド。
【請求項４９】
ＳＥＱ ＩＤ ＮＯ：５のアミノ酸配列を有する請求項１１に記載のポリペプチド。
【請求項５０】
ＳＥＱ ＩＤ ＮＯ：６のポリヌクレオチド配列を有する請求項１１に記載のポリヌクレオチド。
【発明の詳細な説明】
【０００１】
（技術分野）
本発明は、アポトーシス調節因子の核酸配列及びアミノ酸配列に関する。本発明はまた、これらの配列を利用した、細胞増殖異常と免疫学的異常および生殖器官疾患の診断・治療・予防および同じくこれらの配列を利用したアポトーシス調節因子の核酸配列及びアミノ酸配列の発現における外来性化合物の効果についての評価に関する。
【０００２】
（発明の背景）
組織の成長は組織増殖、細胞分化、および調節された細胞死（アポストーシス）の複雑かつ規則的なパターンを伴う。細胞増殖およびアポトーシスは細胞数およびその空間的組織を維持するために制御される。この制御機構は成長因子や分裂促進因子等の細胞外シグナルや、ＤＮＡの破壊または栄養的飢餓等の細胞内シグナルに反応し、細胞周期の進行を制御する適切なタンパク質の発現に依存する。細胞周期進行を直接的、または間接的に調節する分子は、成長因子およびその受容体、セカンドメッセンジャーシグナル伝達タンパク質、腫瘍遺伝子、腫瘍抑制タンパク質、分裂促進因子を含む様々なカテゴリーに分類される。癌は継続的な細胞増殖あるいはコントロール不可能な細胞増殖によって特徴づけられ

る。ある種の癌は通常のアポトーシス細胞死の抑制に関係している。

【０００３】

アポトーシスは、必要のないまたは破壊細胞が細胞死としてプログラムされる、遺伝的に制御されるプロセスである。細胞の選択的排除は細胞増殖や細胞分化の抑制に重要であるのと同様に形態形成や組織の再構築にも重要である。アポトーシスが欠如すると、細胞増殖の増加を伴う過形成その他の疾患を招く恐れがある。アポトーシスは免疫反応においても重要な要因とみなされている。細胞毒性のＴ細胞およびナチュラル・キラー細胞のような免疫細胞は、癌細胞およびウイルスに感染した細胞のアポトーシスを引き起こすことによって、疾病の広がりを防ぐ。更に、外来分子から自己分子を識別できない免疫細胞をアポトーシスによって除去して自己免疫反応を回避しなければならない。

【０００４】

アポトーシス細胞は特有のの形態学的変化を伴う。アポトーシスのホールマークには細胞の縮小、核および細胞質の凝縮、そして細胞質膜輪郭における大きな変化が含まれる。生化学的に、アポトーシス細胞は細胞内カルシウム濃度増加、染色体ＤＮＡの断片化、および新奇な細胞表面成分の発現によって特徴づけられる。

【０００５】

アポトーシスの分子メカニズムは高度に保存され、主要なタンパク質調節因子やアポトーシスのエフェクターの多くが同定されている。アポトーシスは一般に細胞内で変換されたシグナルに反応して進行し、遺伝子の発現およびタンパク質の活性パターンを変化させる。ホルモンやサイトカイン等のシグナル伝達分子は細胞表面の受容体との相互作用を介して、アポトーシスを活性したり、阻害することで知られる。

【０００６】

アポトーシスの特徴の一つとして染色体ＤＮＡの断片化があげられる。ＤＮＡ断片化因子（ＤＦＦ）は４０－ｋＤａのＤＦＦ４０／ＣＡＤと呼ばれるカスパーゼ活性化ヌクレアーゼと４５－ｋＤａの阻害因子ＤＦＦ４５／ＩＣＡＤの２つのサブユニットから構成されるタンパク質である。ＣＩＤＥ－ＡおよびＣＩＤＥ－Ｂと称されるＤＦＦ４５／ＩＣＡＤの２つのマウス相同体が最近記載されている（Ｉｎｏｈａｒａ，Ｎ．他．（１９９８）ＥＭＢＯ Ｊ．１７：２５２６－２５３３）。哺乳動物細胞でのＣＩＤＥ－ＡおよびＣＩＤＥ－Ｂの発現はアポトーシスを活性化したが、ＣＩＤＥ－Ａの発現だけでも、ＤＮＡの断片化を誘導した。更に、ＦＡＳ仲介アポトーシスはＣＩＤＥ－ＡおよびＣＩＤＥ－Ｂによって強化され、これらのタンパク質がアポトーシスを媒介するエフェクターであることをさらに裏付けた。

【０００７】

癌は不適切な細胞増殖を特徴とし、その理由は制御不能な細胞成長または不適当なアポトーシスに起因すると考えられる。その治療戦略として、細胞周期進行の再確立または癌細胞におけるアポトーシスの選択的な活性化のいずれかを採用し得る。分裂促進因子ストレスへの反応はしばしば転写レベルで制御され、様々な転写因子によって制御される。例えば脊椎動物転写因子のＲｅｌ／ＮＦ－ｋａｐｐａ Ｂファミリは放射線照射に対して炎症および免疫反応に中心的な役割を演じている。ＮＦ－ｋａｐｐａ Ｂ ファミリは、ｐ５０、ｐ５２、ＲｅｌＡ、ＲｅｌＢ、ｃＲｅｌおよび他の ＤＮＡ結合タンパク質を含む。ｐ５２ タンパク質はアポトーシスを誘発し、転写因子であるｃ－Ｊｕｎを増加調節し、ｃ－Ｊｕｎ Ｎ－末端キナーゼ１ （ＪＮＫ１）を活性化する （Ｓｕｎ，Ｌ．他．（１９９８） Ｇｅｎｅ ２０８：１５７－１６６）。ほとんどのＮＦ－ｋａｐｐａ Ｂ タンパク質はＤＮＡ結合ホモ二量体を形成する。多くの転写因子の二量体化は、ｂＺＩＰドメインと称される保存された配列が仲介される。ｂＺＩＰドメインは、塩基性領域とその後につづくロイシンジッパーによって特徴づけられる。

【０００８】

Ｆａｓ／Ａｐｏ－１受容体（ＦＡＳ）は腫瘍壊死因子（ＴＮＦ）受容体ファミリのメンバーである。そのリガンド（Ｆａｓ ｌｉｇａｎｄ）に結合すると、膜を貫通するＦＡＳは、死信号を送信するいくつかの細胞質内蛋白質を集めることによりアポトーシスを引き起

こす。そのようなタンパク質の一つであるＦＡＦ１関連のタンパク質因子１（ＦＡＦ１）がマウスから単離され、Ｌ細胞におけるＦＡＦ１の発現がＦＡＳに起因するアポトーシスを増強することが実証された（Ｃｈｕ，Ｋ．他．（１９９５）Ｐｒｏｃ．Ｎａｔｌ．Ａｃａｄ．Ｓｃｉ．ＵＳＡ９２：１１８９４－１１８９８）。その後、ＦＡＳ関連の因子はウズラやハエ等、数々の他の種族から単離されている（Ｆｒｏｈｌｉｃｈ，Ｔ．他．（１９９８）Ｊ．Ｃｅｌｌ Ｓｃｉ．１１１：２３５３－２３６３を参照）。Ｆａｓからの死シグナルの伝送に機能する他の細胞質タンパク質はＦａｓ関連のデスドメインタンパク質があり、ＦＡＤＤとして知られている。ＦＡＤＤはＦＡＳとＴＮＦ受容体の仲介者とし、カスパーゼ８を活性化する、両方のアポトーシス経路に死シグナルを送る（Ｂａｎｇ，Ｓ．他（２０００）Ｊ．Ｂｉｏｌ．Ｃｈｅｍ．２７５：３６２１７－３６２２２を参照）。

【０００９】

癌に対する免疫学的防御には腫瘍抑制因子による突然変異細胞におけるアポトーシスの誘導およびＴリンパ球による腫瘍抗原の認識が含まれる。細胞分裂促進ストレスへの反応はしばしば転写レベルで制御され、様々な転写因子によって制御される。例をあげると、脊椎動物転写因子のＲｅｌ／ＮＦ－ｋａｐｐａ Ｂファミリは放射線照射に対して炎症および免疫反応に中心的な役割を演じている。Ｒｅｌ／ＮＦ－ｋａｐｐａ Ｂファミリはｐ５０、ｐ５２、ＲｅｌＡ、ＲｅｌＢ、及びｃＲｅ、そして他のＤＮＡ結合タンパク質を含む。ｐ５２タンパク質はアポトーシスを誘導し、転写因子ｃ－Ｊｕｎを上方制御しｃ－ＪｕｎＮ末端キナーゼ１（ＪＮＫ１）を活性化させる（Ｓｕｎ，Ｌ他（１９９８）Ｇｅｎｅ ２０８：１５７－１６６を参照）。ほとんどのｋａｐｐａ Ｂタンパク質はＤＮＡに結合するホモ二量体またはヘテロ二量体を形成する。多くの転写因子の二量化は、ｂＺＩＰドメインとよばれている保存されたシーケンスによって媒介される。このドメインは、ロイシン・ジッパーとその後に続く塩基的な領域によって特徴づけられる。

【００１０】

成長因子とシグナル伝達機構

成長因子は典型的には、細胞増殖を促進するためにそれらのローカル環境中にいる細胞に作用する大きなポリペプチドである。成長因子は、細胞内シグナル伝達カスケードを開始させる、特定の細胞表面の受容体に結合してこれを活性化する。多くの成長因子受容体は、リガンドを結合すると、自ら燐酸化するチロシンキナーゼ受容体に分類される。受容体は、自らの燐酸化によって、ＳＨ２あるいはＳＨ３（Ｓｒｃ相同領域２あるいは３）の領域を含む蛋白質のようなシグナル伝達蛋白質と相互作用することができる。成長因子の下流に作用するその他の蛋白質には、ＳＰＲＹ（Ｓｐ１ａおよびリアノジン（ｒｙａｎｏｄｉｎｅ）受容体）ドメインの如きユニークな信号ドメインが含まれている（例えば、Ｓｃｈｕｌｔｚ，Ｊ．ｅｔ ａｌ．（１９９８）Ｐｒｏｃ．Ｎａｔｌ．Ａｃａｄ．Ｓｃｉ．ＵＳＡ９５：５８５７－５８６４を参照）。成長因子受容器に下流に作用する他の蛋白質は、ＳＰＲＹ（Ｓｐ１ａおよびｒｙａｎｏｄｉｎｅ受容器）領域のようなユニークなシグナリングドメインを含んでいる。その後、これらの蛋白質は、Ｒａｓ、ＲａｂおよびＲｈｏのような低分子量Ｇ－蛋白質、さらにＧＴＰａｓｅ活性促進蛋白質（ＧＡＰ）、グアニンヌクレオチド放出タンパク質（ＧＮＲＰ）、および他のグアニンヌクレオチド交換因子を調節する。低分子量Ｇ蛋白質は、細胞分裂促進活性化蛋白質キナーゼ（ＭＡＰキナーゼ）にカスケードを起させる分子のスイッチの役割をする。ＭＡＰキナーゼは、以下に考察するように早期に反応する遺伝子の転写を活性化する。

【００１１】

ほとんどの成長因子は、細胞の成長および細胞分裂に加え、他の多くの作用をもっている。これらは環境によって、細胞の増殖、存続、分化、細胞運動あるいは機能をコントロールすることができる。例えば、表皮成長因子（ＥＧＦ）は、細胞の運動と増殖を刺激することによって、傷から胃の粘膜を保護し潰瘍の治癒を加速する。ＥＧＦは、膜貫通のチロシンキナーゼＥＧＦ－Ｒ蛋白質と結合し、ＧＴＰ結合タンパク質であるＲａｓによってＲａｓ／Ｒａｆ／ＭＡＰキナーゼ経路の活性化を起こす一連の出来事を引き起こす。ＥＧＦ

によって活性化される可能性のある他の経路には、ホスファチジルイノシトール経路およびＪＡＫ／ＳＴＡＴシグナル経路が含まれる（Ｔａｒｎａｗｓｋｉ，Ａ．Ｓ．他（１９９８）Ｊ．Ｃｌｉｎ．Ｇａｓｔｒｏｅｎｔｅｒｏｌ．２７：Ｓ１２－Ｓ２０）。

【００１２】

成長因子に加えて、低分子シグナルペプチドやホルモンもまた細胞増殖に影響を与える。これらの分子は主として他の受容体クラスである三量体Ｇタンパク質結合受容体（ＧＰＣＲ）と結合する。これらの受容体は主に免疫、神経、神経内分泌の細胞表面上に見られる。リガンド結合が起こると、ＧＰＣＲは三量体Ｇタンパク質を活性化し、その結果、ホスホリパーゼＣ、Ca^{2+}、およびサイクリックＡＭＰ等の細胞内セカンドメッセンジャーのレベル増加を誘発する。ほとんどのＧＰＣＲ仲介シグナル経路は、直接的な分裂促進因子の効果を有する他のシグナル分子の分泌または分解を引き起こすことによって、間接的に細胞増殖を促進する（Ｓｍｉｔｈ，Ａ．等（１９９４）Ｃｅｌｌ ７６：９５９－９６２を参照）。

【００１３】

タンパク質キナーゼＣ（ＰＫＣ）は、ホルモンや成長因子に対するシグナル伝達を仲介調節することにより、様々な細胞タイプの増殖や分化の中心的な役割となっている。セリン／トレオニンキナーゼのＰＫＣファミリは、Ｃ末端で類似する触媒ドメインをもち、Ｎ末端調節ドメインでは相違する１２のアイソフォームを有する。ほとんどの細胞は複数のＰＫＣアイソフォームを発現することから、恒常的にあるいは細胞刺激が起こると細胞内の選択した場所に各イソ酵素を導くことによって、基質に対する各酵素の特異性が達成される。ＰＫＣイソ酵素をコンパートメント化する機能をもっている可能性のある、数々のＰＫＣ結合タンパク質および脂質が同定され、その内にはＲＡＣＫ１、血清減少反応（ｓｄｒ：ｓｅｒｕｍ ｄｅｐｒｉｖａｔｉｏｎ ｒｅｓｐｏｎｓｅ）タンパク質およびＳＲＢＣ（Ｃキナーゼと結合するｓｄｒ関連遺伝子生成物）が含まれる。興味深いのは、ｓｄｒおよびＳＲＢＣの両者が活性化されたＰＫＣのカベオラへの局在性を提供しているが、それぞれとも違ったイソ酵素に対する特異性を持つ。Ｓｄｒは特異的にＰＫＣαと相互作用し、ＳＲＢＣはＰＫＣδと相互作用する。ｓｄｒおよびＳＲＢＣの両者は成長停止の状態中に誘発され、無血清培養細胞から単離された。ｓｄｒおよびＳＲＢＣは、成長制御に重要な細胞内シグナル伝達サイトに活性化されたＰＫＣイソ酵素を導くことに関して重要であると考えられる（Ｍｉｎｅｏ，Ｃ．他（１９９８）Ｊ．Ｃｅｌｌ Ｂｉｏｌ．１４１：６０１－６１０，Ｉｚｕｍｉ，Ｙ．他（１９９７）Ｊ．Ｂｉｏｌ．Ｃｈｅｍ．２７２：７３８１－７３８９．）。

【００１４】

腫瘍遺伝子

腫瘍遺伝子（即ち「発癌遺伝子」）は成長因子の受容および伝達や、シグナルに反応した遺伝子発現の調節に関与する。例として、成長因子を刺激すると初期反応遺伝子と後期遺伝子の２つの遺伝子が活性化される。初期反応遺伝子はｍｙｃ、ｆｏｃ、およびｊｕｎを含み、その全ては遺伝子調節タンパク質をコードする。これらの調節タンパク質は遅滞反応遺伝子の転写を活性化し、この遺伝子はサイクリンおよびサイクリン依存キナーゼ等の細胞周期の進行に直接関与するタンパク質をコードする。遺伝子発現を直接調節する腫瘍遺伝子生成物は他に、Ｒｅｌ転写因子、Ｒｅｌ亜鉛フィンガータンパク質およびＴｒｅ腫瘍性タンパク質があげられる（例えば、Ｃａｏ，Ｔ．他．（１９９８）Ｊ．Ｃｅｌｌ Ｓｃｉ．１１１：１３１９－１３２９、Ｎａｋａｍｕｒａ，Ｔ．他（１９９２）Ｏｎｃｏｇｅｎｅ ７：７３３－７４１ を参照）。Ｃ３ＨＣ４ ＲＩＮＧフィンガーモチーフ等の、いくつかの腫瘍異伝子の保存領域が同定されている。例えば、Ｂｍｉ－１ 腫瘍性タンパク質のＣ３ＨＣ４ ＲＩＮＧフィンガードメインにおける突然変異はマウスのリンパ腫の誘発をブロックする（Ｈｅｍｅｎｗａｙ，Ｃ．Ｓ．（１９９８）Ｏｎｃｏｇｅｎｅ １６：２５４１－２５４７）。ＩＡＰ（アポトーシスの阻害因子）ファミリの作用に関与すると考えられるＢＩＲ反復等の、アポトーシス阻害モチーフが同定されている。腫瘍遺伝子の作用過多を引き起こす突然変異または染色体転座は、制御不能な細胞増

殖を起こし得る。

【００１５】

腫瘍抑制遺伝子

腫瘍抑制遺伝子は、細胞増殖の抑制に関与する。腫瘍抑制遺伝子の活性を減少させる突然変異は細胞増殖の増加を起こす結果となる。ヒト及び他の哺乳動物では、腫瘍抑制因子として網膜芽細胞腫（Ｒｂ）およびｐ５３タンパク質があげられる。腫瘍抑制因子はＤｒｏｓｏｐｈｉｌａ等の下等動物にも発見され、Ｄｉｓｃｓ－Ｌａｒｇｅ（Ｄｌｇ）やＨｙｐｅｒｐｌａｓｔｉｃ Ｄｉｓｃｓ（Ｈｙｄ）タンパク質は、発生中の成虫原器の未分化の上皮細胞の過形成を阻止する（例えば、Ｍａｎｓｆｉｅｌｄ，Ｅ．他（１９９４）Ｄｅｖ．Ｂｉｏｌ．１６５：５０７－５２６を参照）。大腸癌の約７５％にｐ５３遺伝子の不活性突然変異が見られ、５０％にｒａｓファミリ腫瘍遺伝子の過剰促進突然変異が見られるという事例から腫瘍抑制遺伝子および癌の発達における腫瘍遺伝子の重要性が認められる。

【００１６】

腫瘍抑制遺伝子はしばしば、「門番」として作用する（Ｋｉｎｚｌｅｒ，Ｋ．Ｗ．およびＶｏｇｅｌｓｔｅｉｎ，Ｂ．（１９９６）Ｃｅｌｌ ８７：１５９－１７０を参照）。通常、この門番は細胞分裂、細胞成長停止、および細胞死のバランスを維持する役割を持つ。外部シグナルによってこの門番を活性化、または不活性化したり、または細胞内での位置が変化する。時には、門番の不活性化が細胞増殖に必要なこともあり、また活性化が細胞の成長停止や分化に必要な場合もある。更に、この状況が反対になるケースもある。この門番と相互作用するタンパク質は、門番の活性または細胞内での位置を変更して、細胞発達のいずれの段階においても外来のシグナルに好適な応答ができるようにする。

【００１７】

この門番タンパク質の例として大腸癌抑制遺伝子（ＡＰＣ）タンパク質があげられる。ＡＰＣは広範に発現しているが、結腸直腸細胞の門番として機能していると考えられる。ＡＰＣタンパク質の突然変異は、家族性および散発的なフォームの大腸癌に関連している。これらの全ての突然変異は、ＡＰＣのＣ－末端における切断に関与している。ＡＰＣのＣ－末端は、ＥＢ１、ＲＰ１、および腫瘍抑制タンパク質Ｄｌｇを含んだ数々のタンパク質の結合サイトを提供している。ＡＰＣ及びこれらの結合タンパク質との間の相互作用はＡＰＣ活性を局在化、或いは調節に重要である可能性がある。例えばＥＢ１はＡＰＣを微小管に繋げ、染色体分離での欠陥が大腸での腫瘍形成における初期事象として関与していると考えられる（Ｂｅｒｒｅｕｔａ，Ｌ．（１９９８）Ｐｒｏｃ．Ｎａｔｌ．Ａｃａｄ．Ｓｃｉ．ＵＳＡ ９５：１０５９４－６６０１；Ｒｅｎｎｅｒ，Ｃ．他（１９９７）Ｊ．Ｉｍｍｕｎｏｌ．１５９：１２７６－１２８３を参照）。

【００１８】

門番の別例として、Ｅ２Ｆ転写因子があり、Ｅ２Ｆは、組織によって、細胞周期進行のポジティブな調節因子、あるいは細胞増殖の抑制因子として機能できる。成長停止および分化において細胞分割のバランスは、Ｅ２Ｆとの相互作用および調節をするタンパク質に関与すると見られる。これらのタンパク質にはＲｂ腫瘍抑制タンパク質およびＮＰＤＣ－１（神経増殖、分化、および制御）がある。ＲｂはＥ２Ｆの転写活性を抑制し、相応細胞の分化またはアポトーシスへと誘引する。ＮＰＤＣ－１は神経細胞の特異的な遺伝子で、成長を停止した細胞および分化細胞に発現する。ＮＰＤＣ－１ 遺伝子の生成物であるｎｐｄｃｆ－１はＥ２Ｆと相互作用して細胞増殖の阻止調節をする（Ｄｕｐｏｎｔ，Ｅ．他（１９９８）Ｊ．Ｎｅｕｒｏｓｃｉ．Ｒｅｓ．５１：２５７－２６７）。

【００１９】

細胞周期機構

分裂促進因子および成長因子に反応して、細胞周期を促進する分子機構は発芽酵母、分裂酵母、およびアフリカツメガエルＸｅｎｏｐｕｓ、等のモデルシステムを使用して非常に良く研究されている。本質的に、細胞周期はＧ１，Ｓ（ＤＮＡ合成）、Ｇ２、およびＭ（有糸分裂）．の４つの連続的な段階からなる。細胞周期を終了した細胞はＧ０と称される

静止期に入る。酵母での研究において、Ｓ期およびＭ期の終了は、細胞分裂後期促進複合体によって導かれる。その複合体は、サイクリンを分解するユビキチンに媒介されるタンパク質分解経路を介してサイクリンを分解するタンパク質の集合体である（例えば、Ｋｏｍｉｎａｍｉ，Ｋ．他（１９９８）ＥＭＢＯ Ｊ．１７：５３８８－５３９９を参照）。分裂酵母のＺｅｒ１ｐ ＲＮＡスプライシングタンパク質等の他の非キナーゼタンパク質は、細胞がＧ０期から出てＧ１またはＧ２期に入るために重要である（例えば、Ｕｒｕｓｈｉｙａｍａ，Ｓ．他（１９９７）Ｇｅｎｅｔｉｃｓ １４７：１０１－１１５を参照）。

【００２０】

有糸分裂からの細胞の出入を含め、各細胞周期の移行は、サイクリン依存キナーゼ（Ｃｄｋｓ）の活性化および阻止に依存する。ＣｄｋｓはキナーゼサブユニットのＣｄｋ、および活性サブユニットのサイクリンから構成されており、複合体は数々のレベルの調節に曝される。サイクリンはサイクリン依存タンパク質キナーゼに結合し活性化する。それによって、有糸分裂プロセスに関与する選択されたタンパク質がリン酸化され活性化される。Ｃｄｋサイクリン複合体はリン酸化によって活性および阻止される。更に、Ｃｄｋサイクリン複合体は、ＣＤＣ４およびＣＤＣ５３等の分子が関係する、目標をつけた分解によって調節される。他のタンパク質は有糸分裂に入ったり、その進展を調節する。例えば、ＢｅｒｒｙおよびＧｏｕｌｄは最近、Ｓ．ｐｏｍｂｅ酵母から新奇な１４２アミノ酸タンパク質、ｄｍｐ １ｐを同定した。これは適切な紡錘体の形成および有糸分裂のエントリに必要とされるが、サイクリンタイプタンパク質とは相互作用しない（Ｂｅｒｒｙ Ｌ．Ｄ．および Ｇｏｕｌｄ Ｋ．Ｌ．（１９９７）Ｊ．Ｃｅｌｌ Ｂｉｏｌ．１３７：１３３７－１３５４）。そのヒト相同体が最近報告されて以来、Ｄｉｍ１ｐは進化的に保存されていると考えらている（Ｌａｒｉｎ Ｄ．，他（１９９７）ＧＩ ２５６５２７５）。

【００２１】

アポトーシスのメカニズム

Ｂｃｌ－２タンパク質ファミリーは他の細胞質タンパク質と同様に、アポトーシスの主要な調節因子である。少なくとも１５のＢｃｌ－２ファミリーメンバーが３つのサブファミリの中にある。これらのタンパク質は哺乳動物細胞およびウイルスに同定され、それぞれ少なくとも４つのＢｃｌ－２相同体ドメイン（ＢＨ１からＢＨ４）のうちの一つをもつ。これらのドメインは高度に保存的ある。Ｂｃｌ－２ファミリタンパク質はＢＨ１およびＢＨ２ドメインを含有し、それらはプロサバイバル（生存促進）のサブファミリのメンバーにみられる一方、Ｂｃｌ－２に最もよく似たタンパク質には全て４つの保存されたドメインがあり、様々な細胞毒性環境に遭遇した場合アポトーシスの阻止を有効化する。プロサバイバルのサブファミリのメンバーは哺乳動物のＢｃｌ－２、Ｂｃｌ－x_L、Ｂｃｌ－ｗ、Ｍｃｌ－１、およびＡ１、ＮＦ－１３（ニワトリ）、ＣＥＤ－９（線虫Ｃａｅｎｏｒｈａｂｄｉｔｉｓ ｅｌｅｇａｎｓ）、およびウイルス タンパク質のＢＨＲＦ１、ＬＭＷ５－ＨＬ、ＯＲＦ１６、ＫＳ－Ｂｃｌ－２、およびＥ１Ｂ－１９Ｋ．がある。ＢＨ３ドメインはアポトーシス促進サブファミリタンパク質の機能に不可欠である。２つのアポトーシス促進サブファミリはＢａｘおよびＢＨ３であり、それぞれ Ｂａｘ、Ｂａｋ、および Ｂｏｋ（別称Ｍｔｄ）、およびＢｉｋ、Ｂｌｋ、Ｈｒｋ、ＢＮＩＰ３、Bim_L、Ｂａｄ、Ｂｉｄ、そしてＥｇｌ－１（Ｃ．ｅｌｅｇａｎｓ）を含む。Ｂａｘサブファミリのメンバーにはｂｈ１、ＢＨ２、およびＢＨ３が有り、Ｂｃｌ－２に近似している。それに比較してＢＨ３サブファミリは９－１６残基ドメインのみを有し、既知のタンパク質とは無関連であり、ＢｉｋおよびＢｌｋのみが配列の類似性を共有する。２つのアポトーシス促進サブファミリのタンパク質は生存促進サブファミリタンパク質のアンタゴニストとなり得る。これはアポトーシスに必要なＥｇｌ－１がＣＥＤ－９を介して結合し作用するＣ．ｅｌｅｇａｎｓがその例である（例えばＡｄａｍｓ，Ｊ．Ｍ．および Ｓ．Ｃｏｒｙ（１９９８）Ｓｃｉｅｎｃｅ ２８１：１３２２－１３２６を参照）。

【００２２】

アポトーシス促進および抗アポトーシスサブファミリタンパク質間のヘテロ二量体化はこれらのタンパク質サブファミリーの機能に関して、化学量論的効果を持つと見られ、各サブファミリのメンバーの相対的濃度がアポトーシスを調節することが示される。ヘテロ二量体化は生存促進タンパク質には必要がないが、ＢＨ３サブファミリには不可欠であり、Ｂａｘサブファミリではそれほど必須ではない。

【００２３】

Ｂｃｌ－２ タンパク質にはα型およびβ型の２つのアイソフォームがあり、このアイソフォームは選択的スプライシングによって生成される。Ｂｃｌ－２タンパク質はＢａｘ、Ｂａｋタンパク質、Ｂｃｌ－Ｘ アイソフォーム Ｂｃｌ－ｘ$_S$とホモ二量体およびヘテロ二量体を生成する。Ｂａｘ とのヘテロ二量体化は無損傷ＢＨ１ およびＢＨ２ ドメインを必要とし、生存促進活性に欠かせない。ＢＨ４ドメインも製造促進活性に関与していると考えられる。Ｂｃｌ－２は核エンベロープや小胞体の他、内外ミトコンドリア膜の中に見られ、様々な組織に発現する。濾胞性リンパ腫（慢性リンパ性白血病）との関与はＴ（１４；１８）（ｑ３２；ｑ２１）染色体転座に見いだされ、免疫グロブリン遺伝子領域に関係している。

【００２４】

Ｂｃｌ－ｘ タンパク質はアポトーシス細胞死の優勢調節因子である。選択的スプライシングによってＢｃｌ－ｘＢ、ロングアイソフォームおよびショートアイソフォームの３つのアイソフォームを形成する。ロングアイソフォームは細胞死抑制活性を示し、ショートアイソフォームはアポトーシスを促進する。Ｂｃｌ－ｘＬはＢａｘおよびＢａとヘテロ二量体を形成するが、Ｂａｘとのヘテロ二量体化は生存促進活性（抗アポトーシス）に必要とは考えられていない。Ｂｃｌ－ｘＳはＢｃｌ－２とヘテロ二量体を形成する。Ｂｃｌ－ｘはミトコンドリア膜および核周囲エンベロープに見られる。．Ｂｃｌ－ｘＳは、高率で再生を行っているリンパ球や他の細胞の発達において高いレベルで発現する。Ｂｃｌ－ｘＬはヒトの脳や他の組織の長寿命有糸分裂後の細胞に見られる。Ｂｃｌ－２とともにＢＨ１、ＢＨ２、およびＢＨ４ドメインは生存促進活性に関与している。

【００２５】

Ｂｃｌ－ｗタンパク質はほとんどの骨髄細胞株の細胞質内や多くの組織内に存在し、脳、結腸、唾液腺に高いレベルで発現する。このタンパク質は精巣、肝臓、心臓、胃、骨格筋、胎盤、リンパ球細胞株での発現は低い。Ｂｃｌ－ｗはＢＨ１、ＢＨ２、およびＢＨ４ドメインを含み、その全ては細胞生存促進活性に必要である。Ｂｃｌ－ｗ遺伝子機能が相同組換えを用いて破壊されたマウスは生存性を提示し、健康で通常な外観が観察され、雌では生殖機能が通常であったが、雄では不妊症が確認された。これらの男性は、精子形成の初期思春期前の段階では顕著な影響は見られず、通常の試験結果となった。しかし、精細管に不規則性および多くのアポトーシス細胞の存在が所見され、成熟した精子の生成に不能性が確認された。このマウスのモデルは、ヒト男性の不妊症にケースによって適用される可能性があり、精巣内のプログラムされた細胞死の変更が生殖能力の調節に適用され得ることを示している（Ｐｒｉｎｔ，Ｃ．Ｇ．他．（１９９８）Ｐｒｏｃ．Ｎａｔｌ．Ａｃａｄ．Ｓｃｉ．ＵＳＡ ９５：１２４２４－１２４３１を参照）。

【００２６】

ラット虚血性脳の研究によると、非虚血性脳では通常の低恒常性の発現レベルであるのに比較して、虚血性脳では過剰に発現していることが実証された。更に、Ｂｃｌ－ｗの活性機構を調べたｉｎ ｖｉｔｒｏ での研究では、Ｂｃｌ－ｗが存在する時に、Ｂａｘの組換えや高濃度のカルシウムの添加に、単離したラットの脳ミトコンドリアが反応できなかった。通常の反応ではミトコンドリアからチトクロームＣが放出される。また、組換えられたＢｃｌ－ｗタンパク質は、カルシウム誘発性のミトコンドリア膜での電位差の損失を阻止することが確認された。電位差の損失は透過性の変化を意味する。これらの発見から、Ｂｃｌ－ｗは虚血性神経細胞に対する保護因子として機能し、この保護機能をミトコンドリアの細胞死調節因子経路を介して達成し得ることが示される（Ｙａｎ，Ｃ．他（２０００）Ｊ．Ｃｅｒｅｂ．Ｂｌｏｏｄ Ｆｌｏｗ Ｍｅｔａｂ．２０：６２０－６３０

）。

【００２７】

ｂｆｌ－１遺伝子はＢｃｌ－２ファミリの付加メンバーであり、アポトーシスの抑制因子としても機能する。Ｂｆｌ－１タンパク質は１７５のアミノ酸を有し、Ｂｃｌ－２ファミリメンバーに存在するＢＨ１、ＢＨ２、およびＢＨ３保存ドメインを含む。このタンパク質はグルタミンの多いＮ末端領域も含むが、他のＢｃｌ－２ファミリメンバーとは異なってＮＨ１は含まない。マウスのＡ１タンパク質はＢｆｌ－１と高配列相同性持ち、Ｂｆｌ－１に存在する３の保存ドメインを持つ。ｐ５３腫瘍抑制タンパク質によって誘発されるアポトーシスはＢｆｌ－１によって抑制され、Ｂｃｌ－２、Ｂｃｌ－ｘＬ、およびＥＢＶ－ＢＨＲＦ１の活性作用と似ている（Ｄ’Ｓａ－Ｅｉｐｐｅｒ，Ｃ．他（１９９６）Ｃａｎｃｅｒ Ｒｅｓ．５６：３８７９－３８８２）。Ｂｆｌ－１は細胞内に存在し造血性分画、即ち血液、脾臓、および骨髄等に高度の発現、肺、小腸および精巣に中程度の発現、その他の組織に低度の発現が見られる。また、平滑筋細胞や、血液悪性腫瘍にも見られる。ｂｆｌ－１の発現レベルと胃癌の発達に相互関係のあることが注目され、ｂｆｌ－１の発現レベルと胃癌の発達に相互関係のあることが注目され、Ｂｆｌ－１タンパク質は癌組織細胞の生存促進あるいは癌において胃癌の発達に関与することを示している（Ｃｈｏｉ，Ｓ．Ｓ．他（１９９５）Ｏｎｃｏｇｅｎｅ １１：１６９３－１６９８）。

【００２８】

転写因子はアポトーシスの開始に重要な役割を握っている。多くの下流エフェクター分子、特にカスパーゼと称されるシステインプロテアーゼ等のプロテアーゼが、アポトーシスの初期および実行期に関与している。プロサバイバルおよびプロ好アポトーシスのＢｃｌ－２関連タンパク質の競合作用によってカスパーゼが活性化する（Ｐｒｉｎｔ，Ｃ．Ｇ．他．（１９９８）Ｐｒｏｃ．Ｎａｔｌ．Ａｃａｄ．Ｓｃｉ．ＵＳＡ ９５：１２４２４－１２４３１）。アポトーシスの促進シグナルで、タンパク分解カスパーゼカスケードを誘発するイニシエーターカスパーゼが活性化される。これにより標的タンパク質の加水分解および典型的なアポトーシス細胞死を誘因する。２つの活性化部位残基、システイン残基およびヒスチジン残基が触媒作用に関与している。カスパーゼは最も特異的なエンドペプチターゼの一種であり、アスパラギン酸残基の後を切断する。

【００２９】

カスパーゼは、大きいサブユニット（ｐ２０）および小さいサブユニット（ｐ１０）が小さなスペーサー領域で分かれており、可変のＮ末端プロドメイン構成からなる、不活性な酵素前駆体として合成される。このプロドメインはアポトーシスを正および負方向に影響を与えることができる補助因子と相互作用する。シグナルの活性化により、特異的アスパラギン酸残基（カスパーセ１の番号付けでＤ２９７の位置）の自己分解性開裂が起こり、スペーサー領域とプロドメインが除かれ、ｐ１０／ｐ２０のヘテロ二量体を残す。２つのヘテロ二量体は、小さいサブユニットを介して相互作用し、触媒活性の四量体を形成する。カスパーゼファミリメンバーの長いプロドメインは二量体およびプロカスパーゼの自己プロセッシングを促進することが示されている。カスパーゼにはプロドメインに「デスエフェクタードメイン」を含むものがあり、これにより他のカスパーゼおよびＦＡＤＤタンパク質と死受容体またはＴＮＦ受容体複合に関連した自己活性複合体として動員され得る。更に違うカスパーゼファミリメンバーから由来する二つの二量体は結合でき、その結果として得られる四量体の基質特異性を変える。カスパーゼの動員ドメイン（ＣＡＲＤ（ｃａｓｐａｓｅ ｒｅｃｒｕｉｔｍｅｎｔ ｄｏｍａｉｎ））は尖端カスパーゼにあるプロドメイン内に存在し、Ａｐａｆ－２、ＲＡＩＤＤおよび細胞内アポトーシス阻害タンパク質（ＩＡＰ）で保存されている（Ｈｏｆｍａｎｎ，Ｋ．他（１９９７）Ｔｒｅｎｄｓ Ｂｉｏｃｈｅｍ．Ｓｃｉ．２２：１５５－１５７）。アポトーシスのＣＡＲＤの調節因子としての機能は、Ａｐａｆ－１等のタンパク質をカスパーゼ－９と結合させると考えられる（Ｌｉ，Ｐ．他（１９９７）Ｃｅｌｌ ９１：４７９－４８９）。骨格筋および心筋で発現されており、ＣＡＲＤと共に働くアポトーシス抑制因子（ＡＲＣ）をコードするヒトｃＤＮＡが、同定されその特性が調べられている。ＡＲＣはアポトーシスの阻害

因子としての機能を持ち、カスパーゼと選択的に相互作用する（Ｋｏｓｅｋｉ，Ｔ．他（１９９８）Ｐｒｏｃ．Ｎａｔｌ．Ａｃａｄ．Ｓｃｉ．ＵＳＡ９５：５１５６－５１６０）。これらの相互作用の全ては、アポトーシスの制御に明瞭な効果を持つ（Ｃｈａｎ Ｓ．Ｌ．およびＭ．Ｐ．Ｍａｔｔｓｏｎ（１９９９）Ｊ．Ｎｅｕｒｏｓｃｉ．Ｒｅｓ．５８：１６７－１９０；Ｓａｌｖｅｓｏｎ，Ｇ．Ｓ．およびＶ．Ｍ．Ｄｉｘｉｔ（１９９９）Ｐｒｏｃ．Ｎａｔｌ．Ａｃａｄ．Ｓｃｉ．ＵＳＡ９６：１０９６４－１０９６７の概説を参照）。

【００３０】

新規のアポトーシス調節因子およびそれらをコードするポリヌクレオチドの発見により、新規の組成物を提供することで当分野の要望に応えることができる。この新規の組成物は、細胞増殖異常、免疫学的異常および生殖器器官疾患の診断・治療・予防において有用であり、また、アポトーシス調節因子の核酸配列及びアミノ酸配列の発現における外来性化合物の効果についての評価にも有用である。

【００３１】

（発明の要約）

本発明は、総称して「ＡＰＲＧ」、個別にはそれぞれ「ＡＰＲＧ－１」、「ＡＰＲＧ－２」、「ＡＰＲＧ－３」、「ＡＰＲＧ－４」、および「ＡＰＲＧ－５」と呼ぶアポトーシス調節因子である精製されたポリペプチドを提供する。或る実施態様において本発明は、（ａ）ＳＥＱ ＩＤ ＮＯ：１－３からなる群から選択したアミノ酸配列からなるポリペプチド、（ｂ）ＳＥＱ ＩＤ ＮＯ：１－３からなる群から選択したアミノ酸配列と少なくとも９０％が同一であるアミノ酸配列を有する天然のポリペプチド、（ｃ）ＳＥＱ ＩＤ ＮＯ：１－３からなる群から選択したアミノ酸配列を有するポリペプチドの生物学的活性断片、または（ｄ）ＳＥＱ ＩＤ ＮＯ：１－３からなる群から選択したアミノ酸配列を有するポリペプチドの免疫抗原性断片を含む群から選択した実質上単離されたポリペプチドを提供する。一実施態様では、ＳＥＱ ＩＤ ＮＯ：１－３のアミノ酸配列を含む実質上単離されたポリペプチドを提供する。

【００３２】

また、本発明は（ａ）ＳＥＱ ＩＤ ＮＯ：１－３（配列番号１乃至３）からなる群から選択したアミノ酸配列、（ｂ）ＳＥＱ ＩＤ ＮＯ：１－３からなる群から選択したアミノ酸配列と少なくとも９０％の相同性を有する天然のアミノ酸配列、（ｃ）ＳＥＱ ＩＤ ＮＯ：１－３からなる群から選択したアミノ酸配列の生物学的活性断片、または（ｄ）ＳＥＱ ＩＤ ＮＯ：１－３からなる群から選択したアミノ酸配列の免疫抗原性断片を含むポリペプチドをコードするような実質上単離されたポリヌクレオチドを提供する。一実施態様では、ポリヌクレオチドはＳＥＱ ＩＤ ＮＯ：１－３からなる群から選択したポリペプチドをコードする。別の実施態様では、ポリヌクレオチドはＳＥＱ ＩＤ ＮＯ：４－６（配列番号４乃至６）からなる群から選択される。

【００３３】

本発明は更に、（ａ）ＳＥＱ ＩＤ ＮＯ：１－３からなる群から選択したアミノ酸配列からなるポリペプチド、（ｂ）ＳＥＱ ＩＤ ＮＯ：１－３からなる群から選択したアミノ酸配列と少なくとも９０％の相同性を有するアミノ酸配列からなる天然のポリペプチド、（ｃ）ＳＥＱ ＩＤ ＮＯ：１－３からなる群から選択したアミノ酸配列を有するポリペプチドの生物学的活性断片、または（ｄ）ＳＥＱ ＩＤ ＮＯ：１－３からなる群から選択したアミノ酸配列を有するポリペプチドの免疫抗原性断片から構成される群から選択されたポリペプチドをコードするポリヌクレオチドと機能的に結合したプロモーター配列を有する組換えポリヌクレオチドを提供する。一実施態様では、本発明は組換えポリヌクレオチドを用いて形質転換した細胞を提供する。別の実施態様では、本発明は組換えポリヌクレオチドを含む遺伝形質転換体を提供する。

【００３４】

また、本発明は（ａ）ＳＥＱ ＩＤ ＮＯ：１－３からなる群から選択したアミノ酸配列、（ｂ）ＳＥＱ ＩＤ ＮＯ：１－３からなる群から選択したアミノ酸配列と少なくとも９０

％の相同性を有する天然のアミノ酸配列、（ｃ）ＳＥＱ ＩＤ ＮＯ：１－３からなる群から選択したアミノ酸配列の生物学的活性断片、または（ｄ）ＳＥＱ ＩＤ ＮＯ：１－３からなる群から選択したアミノ酸配列の免疫抗原性断片を含む実質上単離されたポリペプチドを製造する方法を提供する。製造方法は、（ａ）組換えポリヌクレオチドを用いて形質転換した細胞をポリペプチドの発現に適した条件下で培養する過程と、（ｂ）そのように発現したポリペプチドを受容する過程とを有し、組換えポリヌクレオチドはポリペプチドをコードするポリヌクレオチドに機能的に結合したプロモーター配列を有する。

【００３５】

本発明は更に、（ａ）ＳＥＱ ＩＤ ＮＯ：１－３からなる群から選択したアミノ酸配列を有するポリペプチド、（ｂ）ＳＥＱ ＩＤ ＮＯ：１－３からなる群から選択したアミノ酸配列と少なくとも９０％同一であるアミノ酸配列を有する天然のポリペプチド、（ｃ）ＳＥＱ ＩＤ ＮＯ：１－３からなる群から選択したアミノ酸配列を有するポリペプチドの生物学的活性断片、または（ｄ）ＳＥＱ ＩＤ ＮＯ：１－３からなる群から選択したアミノ酸配列を有するポリペプチドの免疫抗原性断片から構成される群から選択されたポリペプチドに特異結合するような実質上単離された抗体を提供する。

【００３６】

本発明は更に、（ａ）ＳＥＱ ＩＤ ＮＯ：４－６からなる群から選択したポリヌクレオチド配列を含むポリヌクレオチド、（ｂ）ＳＥＱ ＩＤ ＮＯ：４－６からなる群から選択したポリヌクレオチド配列と少なくとも９０％の相同性を有する天然のポリヌクレオチド配列、（ｃ）（ａ）に相補的なポリヌクレオチド配列、（ｄ）（ｂ）のポリヌクレオチドに相補的なポリヌクレオチド配列、および（ｅ）（ａ）～（ｄ）のＲＮＡ等価物を含む実質上単離されたポリヌクレオチドを提供する。一実施態様では、ポリヌクレオチドは少なくとも６０の連続したヌクレオチドを有する。

【００３７】

本発明は更に、サンプル中の標的ポリヌクレオチドを検出する方法を提供する。ここで、標的ポリヌクレオチドは（ａ）ＳＥＱ ＩＤ ＮＯ：４－６からなる群から選択したポリヌクレオチド配列、（ｂ）ＳＥＱ ＩＤ ＮＯ：４－６からなる群から選択したポリヌクレオチド配列と少なくとも９０％の相同性を有する天然のポリヌクレオチド配列、（ｃ）（ａ）に相補的なポリヌクレオチド配列、（ｄ）（ｂ）に相補的なポリヌクレオチド配列、または（ｅ）（ａ）～（ｄ）のＲＮＡ等価物を含む実質上単離されたポリヌクレオチドを提供する。検出方法は、（ａ）サンプル中の標的ポリヌクレオチドに相補的な配列を含む少なくとも２０の連続したヌクレオチドを含むプローブを用いて該サンプルをハイブリダイズする過程と、（ｂ）ハイブリダイゼーション複合体の存在・不存在を検出し、複合体が存在する場合にはオプションでその量を検出する過程からなり、プローブと標的ポリヌクレオチドあるいはその断片の間でハイブリダイゼーション複合体が形成されるような条件下で、プローブは標的ポリヌクレオチドに特異的にハイブリダイズする。一実施態様では、プローブは少なくとも６０の連続したヌクレオチドを含む。

【００３８】

本発明はまた、サンプル中の標的ポリヌクレオチドを検出する方法を提供する。ここで、標的ポリヌクレオチドは（ａ）ＳＥＱ ＩＤ ＮＯ：４－６からなる群から選択したポリヌクレオチド配列、（ｂ）ＳＥＱ ＩＤ ＮＯ：４－６からなる群から選択したポリヌクレオチド配列と少なくとも９０％の相同性を有する天然のポリヌクレオチド配列、（ｃ）（ａ）に相補的なポリヌクレオチド配列、（ｄ）（ｂ）のポリヌクレオチドに相補的なポリヌクレオチド、および（ｅ）（ａ）～（ｄ）のＲＮＡ等価物からなる群から選択された配列のポリヌクレオチドを有する。検出方法は、（ａ）ポリメラーゼ連鎖反応増幅を用いて標的ポリヌクレオチドまたはその断片を増幅する過程と、（ｂ）標的ポリヌクレオチドまたはその断片の存在・不存在を検出し、該標的ポリヌクレオチドまたはその断片が存在する場合にはオプションでその量を検出する過程を含む。

【００３９】

本発明は更に、有効量のポリペプチドと薬剤として許容できる賦形剤とを含む成分を提供

し、有効量のポリペプチドは、（ａ）ＳＥＱ ＩＤ ＮＯ：１－３を有する群から選択したアミノ酸配列、（ｂ）ＳＥＱ ＩＤ ＮＯ：１－３を有する群から選択したアミノ酸配列と少なくとも９０％の相同性を有する天然のアミノ酸配列、（ｃ）ＳＥＱ ＩＤ ＮＯ：１－３有する群から選択したアミノ酸配列の生物学的活性断片、または（ｄ）ＳＥＱ ＩＤ ＮＯ：１－３を有する群から選択したアミノ酸配列の免疫抗原性断片を含む。一実施例では、ＳＥＱ ＩＤ ＮＯ：１－３からなる一群から選択されたアミノ酸配列を含む組成物を提供する。更に、本発明は機能的ＡＰＲＧの発現の低下に関連した疾患やその症状の治療方法を提供するが、この方法にはそのような治療を必要とする患者にこの組成物を投与することが含まれる。

【００４０】

本発明はまた、（ａ）ＳＥＱ ＩＤ ＮＯ：１－３からなる群から選択したアミノ酸配列、（ｂ）ＳＥＱ ＩＤ ＮＯ：１－３からなる群から選択したアミノ酸配列と少なくとも９０％の相同性を有する天然のアミノ酸配列、（ｃ）ＳＥＱ ＩＤ ＮＯ：１－３からなる群から選択したアミノ酸配列の生物学的活性断片、または（ｄ）ＳＥＱ ＩＤ ＮＯ：１－３からなる群から選択したアミノ酸配列の免疫抗原性断片を含むポリペプチドのアゴニストとしての有効性を確認するために化合物をスクリーニングする方法を提供する。スクリーニング方法は、（ａ）ポリペプチドを有するサンプルを化合物に曝す過程と、（ｂ）サンプル中のアゴニスト活性を検出する過程とを含む。別法では、本発明は、この方法によって同定されたアゴニスト化合物と好適な医薬用賦形剤とを含む組成物を提供する。別法では、本発明は、この方法によって同定されたアゴニスト化合物と好適な医薬用賦形剤とを含む組成物を提供する。

【００４１】

本発明は更に、（ａ）ＳＥＱ ＩＤ ＮＯ：１－３からなる群から選択したアミノ酸配列、（ｂ）ＳＥＱ ＩＤ ＮＯ：１－３からなる群から選択したアミノ酸配列と少なくとも９０％の相同性を有する天然のアミノ酸配列、（ｃ）ＳＥＱ ＩＤ ＮＯ：１－３からなる群から選択したアミノ酸配列の生物学的活性断片、または（ｄ）ＳＥＱ ＩＤ ＮＯ：１－３からなる群から選択したアミノ酸配列の免疫抗原性断片を含むポリペプチドのアンタゴニストとしての有効性を確認するために化合物をスクリーニングする方法を提供する。スクリーニング方法は、（ａ）ポリペプチドを含むサンプルを化合物に曝露する過程と、（ｂ）サンプル中のアンタゴニスト活性を検出する過程とを含む。一実施態様で本発明は、この方法によって同定したアンタゴニスト化合物と薬剤として許容できる賦形剤とを含む成分を提供する。更なる別法では、本発明は、このような治療を必要とする患者へのこの組成物の投与を含む、機能的ＡＰＲＧの過剰な発現に関連した疾患やその症状の治療方法を提供する。

【００４２】

本発明は更に、（ａ）ＳＥＱ ＩＤ ＮＯ：１－３からなる群から選択したアミノ酸配列、（ｂ）ＳＥＱ ＩＤ ＮＯ：１－３からなる群から選択したアミノ酸配列と少なくとも９０％の相同性を有する天然のアミノ酸配列、（ｃ）ＳＥＱ ＩＤ ＮＯ：１－３からなる群から選択したアミノ酸配列の生物学的活性断片、または（ｄ）ＳＥＱ ＩＤ ＮＯ：１－３からなる群から選択したアミノ酸配列の免疫抗原性断片を含むポリペプチドに特異結合する化合物をスクリーニングする方法を提供する。スクリーニング方法は、（ａ）ポリペプチドを適切な条件下で少なくとも１つの試験化合物に結合させる過程と、（ｂ）試験化合物とのポリペプチドの結合を検出し、それによってポリペプチドに特異結合する化合物を同定する過程とを含む。

【００４３】

本発明は更に、（ａ）ＳＥＱ ＩＤ ＮＯ：１－３からなる群から選択したアミノ酸配列、（ｂ）ＳＥＱ ＩＤ ＮＯ：１－３からなる群から選択したアミノ酸配列と少なくとも９０％の相同性を有する天然のアミノ酸配列、（ｃ）ＳＥＱ ＩＤ ＮＯ：１－３からなる群から選択したアミノ酸配列の生物学的活性断片、または（ｄ）ＳＥＱ ＩＤ ＮＯ：１－３からなる群から選択したアミノ酸配列の免疫抗原性断片を含むポリペプチドの活性を調節す

る化合物をスクリーニングする方法を提供する。スクリーニング方法は、（ａ）ポリペプチドの活性が許容された条件下で、ポリペプチドを少なくとも１つの試験化合物と混合させる過程と、（ｂ）ポリペプチドの活性を試験化合物の存在下で算定する過程と、（ｃ）試験化合物の存在下でのポリペプチドの活性を試験化合物の不存在下でのポリペプチドの活性と比較する過程とを含み、試験化合物の存在下でのポリペプチドの活性の変化は、ポリペプチドの活性を調節する化合物であることを意味する。

【００４４】

更に本発明は、ＳＥＱ ＩＤ ＮＯ：４－６からなる一群から選択された配列を含む標的ポリヌクレオチドの発現を変化させるのに効果的な化合物をスクリーニングする方法であって、（ａ）この標的ポリヌクレオチドを含むサンプルを化合物に曝露するステップと、（ｂ）この標的ポリヌクレオチドの発現の変化を検出するステップとを含む、該スクリーニング方法を提供する。

【００４５】

本発明は更に、（ａ）核酸を含む生物学的サンプルを試験化合物で処理する過程と、（ｂ）（ｉ）ＳＥＱ ＩＤ ＮＯ：４－６からなる群から選択したポリヌクレオチド配列を含むポリヌクレオチド、（ｉｉ）ＳＥＱ ＩＤ ＮＯ：４－６からなる群から選択したポリヌクレオチド配列と少なくとも９０％の相同性を有するポリヌクレオチド配列を含む天然のポリヌクレオチド、（ｉｉｉ）（ｉ）に相補的な配列を有するポリヌクレオチド、（ｉｖ）（ｉｉ）のポリヌクレオチドに相補的なポリヌクレオチド、および（ｖ）（ｉ）～（ｉｖ）のＲＮＡ等価物を含む群から選択したポリヌクレオチドの少なくとも２０の連続したヌクレオチドから構成されるプローブを用いて、処理した生物学的サンプルの核酸をハイブリダイズする過程とを含む試験化合物の毒性の算定方法を提供する。ハイブリダイゼーションは、上記プローブと生物学的サンプル中の標的ポリヌクレオチドの間に特定のハイブリダイゼーション複合体が形成されるような条件下で発生し、上記標的ポリヌクレオチドは、（ｉ）ＳＥＱ ＩＤ ＮＯ：４－６からなる群から選択したポリヌクレオチド配列を含むポリヌクレオチド、（ｉｉ）ＳＥＱ ＩＤ ＮＯ：４－６からなる群から選択したポリヌクレオチド配列と少なくとも９０％の相同性を有するポリヌクレオチド配列を含む天然のポリヌクレオチド、（ｉｉｉ）（ｉ）のポリヌクレオチドに相補的な配列を有するポリヌクレオチド、（ｉｖ）（ｉｉ）のポリヌクレオチドに相補的なポリヌクレオチド、（ｖ）（ｉ）～（ｉｖ）のＲＮＡ等価物を含む群から選択する。あるいは、標的ポリヌクレオチドは、上記（ｉ）～（ｖ）からなる群から選択したポリヌクレオチド配列の断片と、（ｃ）ハイブリダイゼーション複合体の量を定量する過程と、（ｄ）処理した生物学的サンプルのハイブリダイゼーション複合体の量を、非処理の生物学的サンプルのハイブリダイゼーション複合体の量と比較する過程を含み、処理した生物学的サンプルのハイブリダイゼーション複合体の量の差は、試験化合物の毒性を示す

（発明の記載について）

本発明のタンパク質、ヌクレオチド配列及び方法について説明するが、その前に、説明した特定の装置、材料及び方法に本発明が限定されるものではなく、改変し得ることを理解されたい。また、ここで使用する専門用語は特定の実施例を説明する目的で用いたものに過ぎず、特許請求の範囲にのみ限定される本発明の範囲を限定することを意図したものではないことも併せて理解されたい。

【００４６】

請求の範囲及び明細書中で用いている単数形の「或る」及び「その（この）」の表記は、文脈から明らかにそうでないとされる場合を除いて複数のものを指す場合もあることに注意しなければならない。従って、例えば「或る宿主細胞」と記されている場合にはそのような宿主細胞が複数あることもあり、「或る抗体」と記されている場合には単数または複数の抗体、及び、当業者に公知の抗体の等価物等についても言及しているのである。

【００４７】

本明細書中で用いる全ての専門用語及び科学用語は、特に定義されている場合を除き、当業者に一般に理解されている意味と同じ意味を有する。本明細書で説明するものと類似あ

るいは同等の任意の装置、材料及び方法を用いて本発明の実施または試験を行うことができるが、ここでは好適な装置、材料、方法について説明する。本発明で言及する全ての刊行物は、刊行物中で報告されていて且つ本発明に関係があるであろう細胞、プロトコル、試薬及びベクターについて説明及び開示する目的で引用しているものである。本明細書のいかなる開示内容も、本発明が先行技術の効力によってこのような開示に対して先行する権利を与えられていないことを認めるものではない。

【００４８】

（定義）

用語「ＡＰＲＧ」は、天然、合成、半合成あるいは組換え体など全ての種（特にウシ、ヒツジ、ブタ、マウス、ウマ及びヒトを含む哺乳動物）から得られる実質的に精製されたＡＰＲＧのアミノ酸配列を指す。

【００４９】

用語「アゴニスト」は、ＡＰＲＧの生物学的活性を強めたり、模倣する分子を指す。アゴニストの例として、タンパク質、核酸、糖質、小分子その他 の任意の化合物や成分を挙げることができるが、これらはＡＰＲＧと直接相互作用することによって、あるいはＡＰＲＧが関与する生物学的経路の構成エレメントに作用することによって、ＡＰＲＧの活性を調節する。

【００５０】

用語「対立遺伝子変異配列」は、ＡＰＲＧをコードする遺伝子の別の形を指す。対立遺伝子変異体は、核酸配列における少なくとも１の突然変異から作製し得る。また、変異ＲＮＡまたはポリペプチドからも作製し得る。ポリペプチドの構造または機能は、変異することもしないこともある。遺伝子は、天然の対立遺伝子変異体を全く有しないか、１個若しくは数個の天然の対立遺伝子変異体を有し得る。一般に対立遺伝子変異体を生じさせる通常の突然変異性変化は、ヌクレオチドの自然欠失、付加または置換に帰するものである。これら各変化は、単独あるいは他 の変化と共に、所定の配列内で１回若しくは数回生じ得る。

【００５１】

ＡＰＲＧをコードする「変異」核酸配列は、様々なヌクレオチドの欠失、挿入、あるいは置換が起こっても、ＡＰＲＧと同じポリペプチドあるいはＡＰＲＧの機能特性の少なくとも１つを備えるポリペプチドを指す。この定義には、ＡＰＲＧをコードするポリヌクレオチド配列の正常な染色体の遺伝子座ではない位置での対立遺伝子変異配列との不適当あるいは予期しないハイブリダイゼーション、並びにＡＰＲＧをコードするポリヌクレオチドの特定のオリゴヌクレオチドプローブを用いて容易に検出可能なあるいは検出困難な多形性を含む。コードされたタンパク質も変異され得り、サイレント変化を生じＡＰＲＧと機能的に等価となるアミノ酸残基の欠失、挿入、あるいは置換を含み得る。意図的なアミノ酸置換は、生物学的あるいは免疫学的にＡＰＲＧの活性が保持される範囲で、残基の極性、電荷、溶解度、疎水性、親水性、及び／または両親媒性についての類似性に基づいて成され得る。例えば、負に帯電したアミノ酸にはアスパラギン酸及びグルタミン酸があり、正に帯電したアミノ酸にはリジン及びアルギニンがある。親水性値が近似している非荷電極性側鎖を有するアミノ酸には、 アスパラギンとグルタミン、セリンとスレオニンがある。親水性値が近似している非荷電側鎖を有するアミノ酸には、 ロイシンとイソロイシンとバリン、グリシンとアラニン、フェニルアラニンとチロシンがある。

【００５２】

「アミノ酸」または「アミノ酸配列」の語は、オリゴペプチド、ペプチド、ポリペプチド若しくはタンパク質の配列またはその断片を指し、天然または合成分子を指す。ここで、「アミノ酸配列」は天然のタンパク質分子のアミノ酸配列を指すものであり、「アミノ酸配列」及び類似の語は、アミノ酸配列を、列挙したタンパク質分子に関連する完全な本来のアミノ酸配列に限定しようとするものではない。

【００５３】

「増幅」は、核酸配列の追加複製に関連する。増幅は通常、当業者によく知られたポリメ

ラーゼ連鎖反応（ＰＣＲ）技術を用いて行う。

【００５４】

用語「アンタゴニスト」は、ＡＰＲＧの生物学的活性を阻害あるいは減弱する分子である。アンタゴニストとしては、抗体などのタンパク質、核酸、糖質、小分子またはその他の任意の化合物や成分を挙げることができるが、これらはＡＰＲＧと直接相互作用することによって、あるいはＡＰＲＧが関与する生物学的経路の構成エレメントに作用することによって、ＡＰＲＧの活性を調節する。

【００５５】

「抗体」の語は、無損傷免疫グロブリンやその断片、例えばＦａ、Ｆ（ａｂ’）$_2$ 及びＦｖ断片を指すが、これらはエピトープの決定基と結合することができる。ＡＰＲＧポリペプチドと結合する抗体は、抗体を免疫する小ペプチドを含む無傷の分子またはその断片を用いて作製可能である。動物（マウス、ラット、ウサギ等）を免疫化するために用いるポリペプチドまたはオリゴペプチドは、翻訳または化学合成されたＲＮＡに由来し得るもので、好みに応じて担体タンパク質に接合することも可能である。通常用いられる担体であってペプチドと化学結合するものは、ウシ血清アルブミン、サイログロブリン及びキーホールリンペットヘモシアニン（ＫＬＨ）等がある。結合ペプチドは、動物を免疫化するために用いる。

【００５６】

「抗原決定基」の語は、特定の抗体と接触している分子の領域（即ちエピトープ）を指す。タンパク質またはタンパク質断片を用いて宿主動物を免疫化する場合、タンパク質の多数の領域が、抗原決定基（タンパク質の特定の領域または３次元構造）に特異結合する抗体の産生を誘導し得る。抗原決定基は、抗体への結合において無損傷抗原（即ち免疫応答を誘導するために用いられる免疫原）と競合し得る。

【００５７】

「アンチセンス」の語は、特定の核酸配列の「センス」（コーディング）鎖と塩基対を形成することが可能な任意の成分を指す。アンチセンス成分には、ＤＮＡや、ＲＮＡや、ペプチド核酸（ＰＮＡ）や、ホスホロチオ酸、メチルホスホン酸またはベンジルホスホン酸等の修飾されたバックボーン連鎖を有するオリゴヌクレオチドや、２’－メトキシエチル糖または２’－メトキシエトキシ糖等の修飾された糖類を有するオリゴヌクレオチドや、あるいは５－メチルシトシン、２－デオキシウラシルまたは７－デアザ－２’－デオキシグアノシン等の修飾された塩基を有するオリゴヌクレオチドが含まれうる。アンチセンス分子は、化学合成または転写を含む任意の方法で製造することができる。相補的アンチセンス分子は、ひとたび細胞に導入されたら、細胞が形成した天然の核酸配列と塩基対を形成し、転写または翻訳を妨害する二重鎖を形成する。「負」若しくは「マイナス（－）」の語が参照ＤＮＡ分子のアンチセンス鎖を、「正」若しくは「プラス（＋）」がセンス鎖を指すことがある。

【００５８】

「生物学的に活性」の語は、天然分子の構造的機能、調節機能または生化学的機能を有するタンパク質を指す。同様に、用語「免疫学的に活性」または「免疫原性」は、天然あるいは組換え体のＡＰＲＧ、合成のＡＰＲＧまたはそれらの任意のオリゴペプチドが、適当な動物あるいは細胞の特定の免疫応答を誘発して特定の抗体と結合する能力を指す。

【００５９】

「相補（的）」または「相補性」の語は、塩基対形成によってアニーリングする２つの一本鎖核酸の間の関係を指す。例えば、配列「５’Ａ－Ｇ－Ｔ３’」は、相補配列「３’Ｔ－Ｃ－Ａ５’」と対を形成する。

【００６０】

「所定のポリヌクレオチド配列を含む成分」及び「所定のアミノ酸配列を含む成分」は、所定のポリヌクレオチド配列またはアミノ酸配列を含む広範囲の任意の成分を指す。この成分には、乾燥製剤または水溶液が含まれ得る。ＡＰＲＧ 若しくはＡＰＲＧ の断片をコードするポリヌクレオチド配列を含む組成物は、ハイブリダイゼーションプローブとして

使用され得る。このプローブは、凍結乾燥状態で保存可能であり、糖質などの安定化剤と結合させることが可能である。ハイブリダイゼーションにおいては、塩（例えばＮａＣｌ）、界面活性剤（例えばドデシル硫酸ナトリウム；ＳＤＳ）及びその他 の構成エレメント（例えばデンハート液、脱脂粉乳、サケの精子のＤＮＡ等）を含む水溶液中にプローブを分散させることができる。

【００６１】

「コンセンサス配列」は、不要な塩基を分離するためにＤＮＡ配列の解析を繰り返し行い、ＸＬ－ＰＣＲキット（ＰＥ Ｂｉｏｓｙｓｔｅｍｓ，Ｆｏｓｔｅｒ Ｃｉｔｙ ＣＡ）を用いて５’及び／または３’の方向に伸長され、再度シークエンシングされた核酸配列、またはＧＥＬＶＩＥＷ 断片構築システム（ＧＣＧ，Ｍａｄｉｓｏｎ，ＷＩ）またはＰｈｒａｐ（Ｕｎｉｖｅｒｓｉｔｙ ｏｆ Ｗａｓｈｉｎｇｔｏｎ，Ｓｅａｔｔｌｅ ＷＡ）等の断片構築用のコンピュータプログラムを用いて１つあるいはそれ以上の重複するｃＤＮＡやＥＳＴ、またはゲノムＤＮＡ断片から構築された核酸配列を指す。伸長及び構築の両方を行ってコンセンサス配列を決定する配列もある。

【００６２】

「保存的なアミノ酸置換」は、置換がなされた時に元のタンパク質の特性を殆ど損なわないと予測されるような置換、即ちタンパク質の構造と特に機能が保存され、そのような置換による大きな変化がない置換を指す。下表は、タンパク質中で元のアミノ酸と置換可能で、保存アミノ酸置換と認められるアミノ酸を示している。

元の残基	保存的な置換
Ａｌａ	Ｇｌｙ，Ｓｅｔ
Ａｒｇ	Ｈｉｓ，Ｌｙｓ
Ａｓｎ	Ａｓｐ，Ｇｌｎ，Ｈｉｓ
Ａｓｐ	Ａｓｎ，Ｇｌｕ
Ｃｙｓ	Ａｌａ，Ｓｅｒ
Ｇｌｎ	Ａｓｎ，Ｇｌｕ，Ｈｉｓ
Ｇｌｕ	Ａｓｐ，Ｇｌｎ，Ｈｉｓ
Ｇｌｙ	Ａｌａ
Ｈｉｓ	Ａｓｎ，Ａｒｇ，Ｇｌｎ，Ｇｌｕ
Ｉｌｅ	Ｌｅｕ，Ｖａｌ
Ｌｅｕ	Ｉｌｅ，Ｖａｌ
Ｌｙｓ	Ａｒｇ，Ｇｌｎ，Ｇｌｕ
Ｍｅｔ	Ｌｅｕ，Ｉｌｅ
Ｐｈｅ	Ｈｉｓ，Ｍｅｔ，Ｌｅｕ，Ｔｒｐ，Ｔｙｒ
Ｓｅｒ	Ｃｙｓ，Ｔｈｒ
Ｔｈｒ	Ｓｅｒ，Ｖａｌ
Ｔｒｐ	Ｐｈｅ，Ｔｙｒ
Ｔｙｒ	Ｈｉｓ，Ｐｈｅ，Ｔｒｐ
Ｖａｌ	Ｉｌｅ．Ｌｅｕ，Ｔｈｒ

保存アミノ酸置換では通常、（ａ）置換領域におけるポリペプチドのバックボーン構造、例えばβシートやα螺旋構造、（ｂ）置換部位における分子の電荷または疎水性、及び／または（ｃ）側鎖の大部分を保持する。

【００６３】

「欠失」は、結果的に１個若しくは数個のアミノ酸またはヌクレオチドが失われてなくなるようなアミノ酸またはヌクレオチド配列における変化を指す。

【００６４】

「誘導体」の語は、ポリペプチド配列またはポリヌクレオチド配列の化学修飾を指す。例えば、アルキル基、アシル基、ヒドロキシル基またはアミノ基による水素の置換は、ポリヌクレオチド配列の化学修飾に含まれ得る。ポリヌクレオチド誘導体は、天然分子の生物学的または免疫学的機能を少なくとも１つは保持しているポリペプチドをコードする。ポ

リペプチド誘導体は、グリコシル化、ポリエチレングリコール化（ｐｅｇｙｌａｔｉｏｎ）、あるいは任意の同様なプロセスであって誘導起源のポリペプチドから少なくとも１つの生物学的若しくは免疫学的機能を保持しているプロセスによって、修飾されたポリペプチドである。

【００６５】

「検出可能な標識」は、測定可能な信号を生成することができ、ポリヌクレオチドまたはポリペプチドに共有結合または非共有結合するようなレポーター分子または酵素を指す。

【００６６】

「示差発現」は少なくとも２つの異なったサンプルを比較することによって決められる、増加、または非調節、あるいは減少、下方調節、または欠損遺伝子またはタンパク発現を指す。このような比較は例えば、治療後サンプルと未治療のサンプルまたは病態のサンプルと正常サンプルの間で行われ得る。

【００６７】

「断片」は、ＡＰＲＧまたはＡＰＲＧをコードするポリヌクレオチドの固有部分であって、親配列と同一配列であるが親配列よりも長さが短い配列を指す。

【００６８】

断片は、画定された配列の全長から１ヌクレオチド／アミノ酸残基を差し引いた長さよりも短い長さを有し得る。例えば或る断片は、５～１０００の連続したヌクレオチドまたはアミノ酸残基を有し得る。プローブ、プライマー、抗原、治療用分子として、あるいはその他 の目的のために用いられる断片は、少なくとも５、１０、１５、２０、２５、３０、４０、５０、６０、７５、１００、１５０、２５０若しくは５００の連続したヌクレオチドあるいはアミノ酸残基長さであり得る。断片は、分子の特定領域から優先的に選択し得る。例えば、ポリペプチド断片は、所定の配列に示すような最初の２５０または５００アミノ酸（またはポリペプチドの最初の２５％または５０％）から選択された或る長さの連続したアミノ酸を有し得る。これらの長さは明らかに例として挙げているものであり、本発明の実施例では、配列表、表及び図面を含む明細書に裏付けされた任意の長さであってよい。

【００６９】

ＳＥＱ ＩＤ ＮＯ：４－６の断片は、例えば、この断片を得たゲノム内の他 の配列とは異なる、ＳＥＱ ＩＤ ＮＯ：４－６を明確に同定する固有のポリヌクレオチド配列の領域を含む。ＳＥＱ ＩＤ ＮＯ：４－６のある断片は、例えば、ハイブリダイゼーションや増幅技術、またはＳＥＱ ＩＤ ＮＯ：４－６を関連ポリヌクレオチド配列から区別する類似の方法に有用である。ある断片と一致するＳＥＱ ＩＤ ＮＯ：４－６の正確な断片の長さや領域は、その断片の目的に基づいて当分野で一般的な技術によって日常的に測定できる。

【００７０】

ＳＥＱ ＩＤ ＮＯ：１－３のある断片は、ＳＥＱ ＩＤ ＮＯ：４－６のある断片によってコードされる。ＳＥＱ ＩＤ ＮＯ：１－３のある断片は、ＳＥＱ ＩＤ ＮＯ：１－３を特異的に同定する固有のアミノ酸配列の領域を含む。例えば、ＳＥＱ ＩＤ ＮＯ：１－３のある断片は、ＳＥＱ ＩＤ ＮＯ：１－３を特異的に認識する抗体の開発における免疫原性ペプチドとして有用である。ある断片と一致するＳＥＱ ＩＤ ＮＯ： １－１２の正確な断片の長さや領域は、その断片の目的に基づいて当分野で一般的な技術によって日常的に決定できる。

【００７１】

「完全長」ポリヌクレオチド配列とは、少なくとも１つの翻訳開始コドン（例えばメチオニン）、オープンリーディングフレーム及び翻訳終止コドンを有する配列である。「完全長」ポリヌクレオチド配列は、「完全長」ポリペプチド配列をコードする。

【００７２】

「相同性」の語は、配列類似性即ち２つ以上のポリヌクレオチド配列または２つ以上のポリペプチド配列の配列間で互換可能な配列同一性である。

【００７３】
ポリヌクレオチド配列に適用される「一致率」または「一致％」の語は、標準化された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つ以上のポリヌクレオチド配列間で一致する残基の割合を意味する。このようなアルゴリズムは、２配列間のアラインメントを最適化するために比較する配列において、標準化された再現性のある方法でギャップを挿入するので、２つの配列をより有意に比較できる。
【００７４】
ポリヌクレオチド配列間の一致率は、ＭＥＧＡＬＩＧＮ ｖｅｒｓｉｏｎ ３．１２ｅ配列アラインメントプログラムに組込まれているようなＣＬＵＳＴＡＬ Ｖアルゴリズムのデフォルトのパラメータを用いて決定できる。このプログラムは、ＬＡＳＥＲＧＥＮＥ ソフトウエアパッケージ（一組の分子生物学的分析プログラム）（ＤＮＡＳＴＡＲ，Ｍａｄｉｓｏｎ ＷＩ）の一部である。このＣＬＵＳＴＡＬ Ｖは、Ｈｉｇｇｉｎｓ，Ｄ．Ｇ．及びＰ．Ｍ．Ｓｈａｒｐ（１９８９）ＣＡＢＩＯＳ ５：１５１－１５３、Ｈｉｇｇｉｎｓ，Ｄ．Ｇ．他（１９９２）ＣＡＢＩＯＳ ８：１８９－１９１に記載されている。ポリペプチド配列を２つ１組でアラインメントする際のデフォルトパラメータは、Ｋｔｕｐｌｅ＝２、ｇａｐ ｐｅｎａｌｔｙ＝５、ｗｉｎｄｏｗ＝４、「ｄｉａｇｏｎａｌｓ ｓａｖｅｄ」＝４と設定する。残基「重み付け」表をデフォルトとして選択する。ＣＬＵＳＴＡＬ Ｖは、アラインメントされたポリヌクレオチド配列対間の「類似率」として一致率を報告する。
【００７５】
あるいは、米国国立バイオテクノロジー情報センター（ＮＣＢＩ）のＢａｓｉｃ Ｌｏｃａｌ Ａｌｉｇｎｍｅｎｔ Ｓｅａｒｃｈ Ｔｏｏｌ（ＢＬＡＳＴ）が一般的に用いられ、且つ、無料で入手可能な配列比較アルゴリズム一式を提供している（Ａｌｔｓｃｈｕｌ，Ｓ．Ｆ．他．（１９９０）Ｊ．Ｍｏｌ．Ｂｉｏｌ．２１５：４０３－４１０）。このアルゴリズムは、幾つかの情報源から入手可能であり、メリーランド州ベセスダにあるＮＣＢＩ及びインターネット（ｈｔｔｐ：／／ｗｗｗ．ｎｃｂｉ．ｎｌｍ．ｎｉｈ．ｇｏｖ／ＢＬＡＳＴ／）からも入手可能である。ＢＬＡＳＴソフトウェア一式には様々な配列分析プログラムが含まれており、既知のポリヌクレオチド配列を種々のデータベースから得た別のポリヌクレオチド配列とアラインメントする「ｂｌａｓｔｎ」もその１つである。その他 にも、２つのヌクレオチド配列を対で直接比較するために用いる「ＢＬＡＳＴ ２ Ｓｅｑｕｅｎｃｅｓ」と称されるツールも利用可能である。「ＢＬＡＳＴ ２ Ｓｅｑｕｅｎｃｅｓ」は、ｈｔｔｐ：／／ｗｗｗ．ｎｃｂｉ．ｎｌｍ．ｎｉｈ．ｇｏｖ／ｇｏｒｆ／ｂｌ２．ｈｔｍｌにアクセスして対話形式で利用することが可能である。「ＢＬＡＳＴ ２ Ｓｅｑｕｅｎｃｅｓ」ツールは、ｂｌａｓｔｎ と ｂｌａｓｔｐ（後述）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、ギャップ及びデフォルト設定に設定された他 のパラメータと共に用いる。例えば、２つのヌクレオチド配列を比較するために、デフォルトパラメータとして設定された「ＢＬＡＳＴ ２ Ｓｅｑｕｅｎｃｅｓ」ツールＶｅｒｓｉｏｎ ２．０．１２（２０００年４月２１日）を用いてｂｌａｓｔｎを実行してもよい。デフォルトパラメータの設定例を以下に示す。
【００７６】
一致率は、完全に画定された（例えば特定の配列番号で画定された）配列長さと比較して測定し得る。あるいは、より短い長さ、例えばより大きな画定された配列から得られた断片（例えば少なくとも２０、３０、４０、５０、７０、１００または２００の連続したヌクレオチドの断片）の長さと比較して一致率を測定してもよい。ここに挙げた長さは単なる例示的なものに過ぎず、表、図及び配列リストを含めた本明細書に記載された配列に裏付けられた任意の配列長さの断片を用いて、一致率を測定し得る長さを説明し得ることを理解されたい。
【００７７】
高度の相同性を示さない核酸配列が、それにもかかわらず遺伝子コードの縮重が原因で類似のアミノ酸配列をコードする場合がある。この縮重を利用して核酸配列内で変化を生じ

させて、全ての核酸配列が実質上同一のタンパク質をコードするような多数の核酸配列を生成し得ることを理解されたい。

【００７８】

ポリペプチド配列に適用される「一致率」または「一致％」の語は、標準化された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以上のポリペプチド配列間で一致する残基の割合を意味する。ポリペプチド配列アラインメントの方法は公知である。保存的アミノ酸置換を考慮するアラインメント方法もある。既に詳述したこのような保存的置換は通常、置換部位の電荷及び疎水性を保存するので、ポリペプチドの構造を（従って機能も）保存する。

【００７９】

ポリペプチド配列間の一致率は、ＭＥＧＡＬＩＧＮ ｖｅｒｓｉｏｎ ３．１２ｅ配列アラインメントプログラムに組込まれているようなＣＬＵＳＴＡＬ Ｖアルゴリズムのデフォルトのパラメータを用いて決定できる（既に説明したのでそれを参照されたい）。ＣＬＵＳＴＡＬ Ｖを用いて、ポリペプチド配列を２つ１組でアラインメントする際のデフォルトパラメータは、Ｋｔｕｐｌｅ＝１、ｇａｐ ｐｅｎａｌｔｙ＝３、ｗｉｎｄｏｗ＝５、「ｄｉａｇｏｎａｌｓ ｓａｖｅｄ」＝５と設定する。デフォルトの残基重み付け表としてＰＡＭ２５０マトリクスを選択する。ポリヌクレオチドアラインメントと同様に、ＣＬＵＳＴＡＬ Ｖは、アラインメントされたポリペプチド配列対間の「類似率」として一致率を報告する。

【００８０】

あるいは、ＮＣＢＩ ＢＬＡＳＴソフトウェア一式を用いてもよい。例えば、２つのポリペプチド配列を対で比較をする場合、ある者は、デフォルトパラメータで設定された「ＢＬＡＳＴ ２ Ｓｅｑｕｅｎｃｅｓ」ツールＶｅｒｓｉｏｎ ２．０．１２（Ａｐｒ－２１－２０００）でｂｌａｓｔｐを使用するであろう。ｃメータの設定例を以下に示す。

【００８１】

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

一致率は、完全に画定された（例えば特定の配列番号で画定された）ポリペプチド配列の長さと比較して測定し得る。一致率は、配列あるいは、より短い長さ、例えばより大きな画定されたポリペプチド配列から得られた断片（例えば少なくとも１５、２０、３０、４０、５０、７０または１５０の連続した残基の断片）の長さと比較して一致率を測定してもよい。ここに挙げた長さは単なる例示的なものに過ぎず、表、図及び配列リストを含めた本明細書に記載された配列に裏付けられた任意の配列長さの断片を用いて、或る長さであってその長さに対して一致率を測定し得る長さを説明し得ることを理解されたい。

【００８２】

「ヒト人工染色体（ＨＡＣ）」は、約６ｋｂ（キロベース）～１０Ｍｂのサイズの ＤＮＡ配列を含み得る、染色体の複製、分離及び維持に必要な全てのエレメントを含む直鎖状の小染色体である。

【００８３】

「ヒト化抗体」の語は、非抗体結合領域におけるアミノ酸配列はヒト抗体により近づくように変異させた抗体分子であって、本来の結合能力はそのまま保持しているような抗体分子を指す。

【００８４】

「ハイブリダイゼーション」は、所定のハイブリダイゼーション条件下で塩基対を形成す

ることによって、一本鎖ポリヌクレオチドが相補的鎖とアニーリングするプロセスを指す。特異的ハイブリダイゼーションは、２つの核酸配列が高い相同性を共有することを示すものである。特異的ハイブリダイゼーション複合体は許容されるアニーリング条件下で形成され、「洗浄」ステップ後もハイブリダイズされたままである。洗浄ステップは、ハイブリダイゼーションプロセスのストリンジェンシーを決定する際に特に重要であり、更にストリンジェントな条件では、非特異結合（即ち完全には一致しない核酸鎖間の対の結合）が減少する。核酸配列のアニーリングに対する許容条件は、本技術分野における当業者が慣例的に決定する。許容条件はハイブリダイゼーション実験の間は一定でよいが、洗浄条件は所望のストリンジェンシーを得るように、従ってハイブリダイゼーション特異性も得るように実験中に変更することができる。アニーリングが許容される条件は、例えば、温度が６８℃で、約６×ＳＳＣ、約１％（ｗ／ｖ）のＳＤＳ、並びに約１００μｇ／ｍｌのせん断して変性したサケ精子ＤＮＡが含まれる。

【００８５】

一般に、ハイブリダイゼーションのストリンジェンシーは或る程度、洗浄ステップを実行する温度を基準にして表すことができる。このような洗浄温度は通常、所定のイオン強度及びｐＨにおける特定の配列の融点（Ｔｍ）より約５～２０℃低くなるように選択する。このＴｍは、所定のイオン強度及びｐＨの下で、完全に一致するプローブに標的配列の５０％がハイブリダイズする温度である。T_m を計算する式及び核酸のハイブリダイゼーション条件はよく知られており、Ｓａｍｂｒｏｏｋ 他 （１９８９） Ｍｏｌｅｃｕｌａｒ Ｃｌｏｎｉｎｇ：Ｌａｂｏｒａｔｏｒｙ Ｍａｎｕａｌ，第２版，１－３巻，Ｃｏｌｄ Ｓｐｒｉｎｇ Ｈａｒｂｏｒ Ｐｒｅｓｓ，Ｐｌａｉｎｖｉｅｗ ＮＹに記載されており、特に２巻の９章を参照されたい。

【００８６】

本発明のポリヌクレオチド間の高いストリンジェンシーのハイブリダイゼーションでは、約０．２ｘ ＳＳＣ及び約０．１％のＳＤＳの存在の下、約６８℃で１時間の洗浄過程を含む。別法では、６５℃、６０℃、５５℃、４２℃の温度で行う。ＳＳＣ濃度は、約０．１％のＳＤＳ存在下で、約０．１～２×ＳＳＣの範囲で変化し得る。通常は、ブロッキング剤を用いて非特異ハイブリダイゼーションを阻止する。このような遮断剤には、例えば、約１００～２００μｇ／ｍｌの変性サケ精子ＤＮＡがある。特定条件下で、例えばＲＮＡとＤＮＡのハイブリダイゼーションに有機溶剤、例えば約３５～５０％ｖ／ｖの濃度のホルムアミドを用いることもできる。洗浄条件の有用なバリエーションは、当業者には自明であろう。ハイブリダイゼーションは、特に高ストリンジェント条件下では、ヌクレオチド間の進化的な類似性を示唆し得る。このような類似性は、ヌクレオチド及びヌクレオチドにコードされるポリペプチドに対する類似の役割を強く示唆している。

【００８７】

「ハイブリダイゼーション複合体」の語は、相補的塩基対間の水素結合の形成力によって２つの核酸配列間に形成された複合体を指す。ハイブリダイゼーション複合体は、溶解状態で形成し得る（C_0t または R_0t 解析等）。あるいは、一方の核酸配列が溶解状態で存在し、もう一方の核酸配列が固体支持体（例えば紙、膜、フィルタ、チップ、ピンまたはガラススライド、あるいは他 の適切な基質であって細胞若しくはその核酸が固定される基質）に固定されているような２つの核酸配列間に形成され得る。

【００８８】

「挿入」及び「付加」の語は、１個若しくは数個のアミノ酸残基またはヌクレオチド配列を各々付加するようなアミノ酸またはヌクレオチド配列における変化を指す。

【００８９】

「免疫応答」は、炎症、外傷、免疫異常症、伝染性疾患または遺伝性疾患に関連する症状を指し得る。これらの症状は、細胞及び全身の防御系に作用し得る種々の因子、例えばサイトカイン、ケモカイン、その他 のシグナル伝達分子の発現によって特徴づけることができる。

用語「免疫原性断片」は、例えば哺乳動物などの生きている動物に導入すると、免疫反応

を引き起こすＡＰＲＧのポリペプチド断片またはオリゴペプチド断片を指す。用語「免疫原性断片」はまた、本明細書で開示するまたは当分野で周知のあらゆる抗体生産方法に有用なＡＰＲＧのポリペプチド断片またはオリゴペプチド断片を含む。

【００９０】

「マイクロアレイ」の語は、基質上の複数のポリヌクレオチド、ポリペプチドまたはその他 の化合物の構成を指す。

【００９１】

「エレメント」または「アレイエレメント」の語は、マイクロアレイ上に定義された固有の位置にあるハイブリダイズ可能なポリヌクレオチド、ポリペプチドその他 の化合物を指す。

用語「調節」は、ＡＰＲＧの活性の変化を指す。例えば、調節によって、ＡＰＲＧのタンパク質活性、あるいは結合特性、またはその他 の生物学的特性、機能的特性あるいは免疫学的特性の変化が起こる。

【００９２】

「核酸」及び「核酸配列」の語は、ヌクレオチド、オリゴヌクレオチド、ポリヌクレオチドまたはこれらの断片を指す。「核酸」及び「核酸配列」の語は、ゲノム起源または合成起源のＤＮＡまたはＲＮＡであって一本鎖または二本鎖であるかあるいはセンス鎖またはアンチセンス鎖を表し得るようなＤＮＡまたはＲＮＡや、ペプチド核酸（ＰＮＡ）や、任意のＤＮＡ様またはＲＮＡ様物質を指すこともある。

【００９３】

「機能的に結合した」は、第１核酸配列が第２核酸配列と機能的な関係があるように配置された状態を指す。例えば、プロモーターがコード配列の転写または発現に影響を及ぼす場合には、そのプロモーターはそのコード配列に機能的に結合している。機能的に結合したＤＮＡ配列は非常に近接するか、あるいは連続し得る。そして、２つのタンパク質コード領域を結合する必要がある場合は、同一のリーディングフレーム内にある。

【００９４】

「ペプチド核酸」（ＰＮＡ）は、アンチセンス分子または抗遺伝子物質であって、リジンを末端とするアミノ酸残基のペプチドバックボーンに結合した、少なくとも約５ヌクレオチドの長さのオリゴヌクレオチドからなるものを指す。末端のリジンは、成分に溶解性を与える。ＰＮＡは、相補的一本鎖ＤＮＡまたはＲＮＡに優先的に結合して転写の拡張を停止するものであり、ポリエチレングリコール化して細胞におけるＰＮＡの寿命を延長し得る。

【００９５】

ＡＰＲＧの「翻訳後修飾」には、脂質化、グリコシル化、リン酸化、アセチル化、ラセミ化、蛋白分解性切断及びその他 の当分野で既知の修飾を含まれ得る。これらのプロセスは、合成あるいは生化学的に生じ得る。生化学的修飾は、ＡＰＲＧの酵素環境に依存し、細胞の種類によって異なり得る。

【００９６】

「プローブ」とは、同一配列あるいは対立遺伝子核酸配列、関連する核酸配列の検出に用いる、ＡＰＲＧやそれらの相補配列、またはそれらの断片をコードする核酸配列のことである。プローブは、単離されたオリゴヌクレオチドまたはポリヌクレオチドであって、検出可能な標識またはレポーター分子に結合したものである。典型的な標識には、放射性アイソトープ、リガンド、化学発光試薬及び酵素がある。「プライマー」は、短い核酸、通常はＤＮＡオリゴヌクレオチドであり、相補的塩基対を形成することで標的ポリヌクレオチドにアニーリングされ得る。プライマーは次に、ＤＮＡポリメラーゼ酵素によって標的ＤＮＡ鎖に延在し得る。プライマー対は、例えばポリメラーゼ連鎖反応（ＰＣＲ）による核酸配列の増幅（及び同定）に用い得る。

【００９７】

本発明に用いるようなプローブ及びプライマーは通常、既知の配列の少なくとも１５の連続したヌクレオチドを含んでいる。特異性を高めるために長めのプローブ及びプライマー

、例えば開示した核酸配列の少なくとも２０、２５、３０、４０、５０、６０、７０、８０、９０、１００または１５０の連続したヌクレオチドからなるようなプローブ及びプライマーを用いてもよい。これよりもかなり長いプローブ及びプライマーもある。表、図面及び配列リストを含む本明細書に裏付けされた任意の長さのヌクレオチドを用いることができるものと理解されたい。

【００９８】

プローブ及びプライマーの調製及び使用方法については、Ｓａｍｂｒｏｏｋ，Ｊ．他．（１９８９）Ｍｏｌｅｃｕｌａｒ Ｃｌｏｎｉｎｇ：Ａ Ｌａｂｏｒａｔｏｒｙ Ｍａｎｕａｌ，第２版，１－３巻，Ｃｏｌｄ Ｓｐｒｉｎｇ Ｈａｒｂｏｒ Ｐｒｅｓｓ，Ｐｌａｉｎｖｉｅｗ ＮＹ、Ａｕｓｕｂｅｌ，Ｆ．Ｍ．ら、（１９８７）Ｃｕｒｒｅｎｔ Ｐｒｏｔｏｃｏｌｓ ｉｎ Ｍｏｌｅｃｕｌａｒ Ｂｉｏｌｏｇｙ，Ｇｒｅｅｎｅ Ｐｕｂｌｉ．Ａｓｓｏｃ．＆Ｗｉｌｅｙ－Ｉｎｔｅｒｓｃｉｅｎｃｅｓ，Ｎｅｗ Ｙｏｒｋ ＮＹ、Ｉｎｎｉｓ他．（１９９０）ＰＣＲ Ｐｒｏｔｏｃｏｌｓ，Ａ Ｇｕｉｄｅ ｔｏ Ｍｅｔｈｏｄｓ ａｎｄ Ａｐｐｌｉｃａｔｉｏｎｓ，Ａｃａｄｅｍｉｃ Ｐｒｅｓｓ，Ｓａｎ Ｄｉｅｇｏ ＣＡ等を参照されたい。ＰＣＲプライマー対は、その目的のためのコンピュータプログラム、例えばＰｒｉｍｅｒ（Ｖｅｒｓｉｏｎ ０．５，１９９１，Ｗｈｉｔｅｈｅａｄ Ｉｎｓｔｉｔｕｔｅ ｆｏｒ Ｂｉｏｍｅｄｉｃａｌ Ｒｅｓｅａｒｃｈ，Ｃａｍｂｒｉｄｇｅ ＭＡ）を用いるなどして既知の配列から得ることができる。

【００９９】

プライマーとして用いるオリゴヌクレオチドの選択は、そのような目的のために本技術分野でよく知られているソフトウェアを用いて行う。例えばＯＬＩＧＯ ４．０６ソフトウェアは、各１００ヌクレオチドまでのＰＣＲプライマー対の選択に有用であり、オリゴヌクレオチド及び最大５，０００までの大きめのポリヌクレオチドであって３２キロベースまでのインプットポリヌクレオチド配列から得たものを分析するのにも有用である。類似のプライマー選択プログラムには、拡張能力のための追加機能が組込まれている。例えば、ＰｒｉｍＯＵプライマー選択プログラム（テキサス州ダラスにあるテキサス大学南西部医療センターのゲノムセンターから一般向けに入手可能）は、メガベース配列から特定のプライマーを選択することが可能であり、従ってゲノム全体の範囲でプライマーを設計するのに有用である。Ｐｒｉｍｅｒ３プライマー選択プログラム（マサチューセッツ州ケンブリッジのＷｈｉｔｅｈｅａｄ Ｉｎｓｔｉｔｕｔｅ／ＭＩＴゲノム研究センターから一般向けに入手可能）ではユーザーが「ミスプライミング・ライブラリ」をインプットすることができ、ここでプライマー結合部位として避けたい配列はユーザーが指定する。Ｐｒｉｍｅｒ３は特に、マイクロアレイのためのオリゴヌクレオチドの選択に有用である。（後二者のプライマー選択プログラムのソースコードは、各自のソースから得てユーザー固有のニーズを満たすように変更してもよい。）ＰｒｉｍｅｒＧｅｎプログラム（英国ケンブリッジ市の英国ヒトゲノムマッピングプロジェクトーリソースセンターから一般向けに入手可能）は、多数の配列アラインメントに基づいてプライマーを設計し、それによって、アラインメントされた核酸配列の最大保存領域または最小保存領域の何れかとハイブリダイズするようなプライマーの選択を可能にする。従って、このプログラムは、固有な、および保存されたオリゴヌクレオチド及びポリヌクレオチドの断片の同定に有用である。上記選択方法のいずれかによって同定したオリゴヌクレオチド及びポリヌクレオチドの断片は、ハイブリダイゼーション技術において、例えばＰＣＲまたはシークエンシングプライマーとして、マイクロアレイエレメントとして、あるいは核酸のサンプルにおいて完全または部分的相補的ポリヌクレオチドを同定する特異プローブとして有用である。オリゴヌクレオチドの選択方法は、上記の方法に限定されるものではない。

【０１００】

「組換え核酸」は天然配列ではない配列であるかあるいは人為的に組み合わせなければ離隔しているような配列の２以上のセグメントを人為的に組み合わせて産出した配列を有する配列である。この人為的組合せはしばしば化学合成によって達成するが、より一般的に

は核酸の単離セグメントの人為的操作によって、例えばのＳａｍｂｒｏｏｋらの文献（前出）に記載されているような遺伝子工学的手法によって達成する。組換え核酸の語は、単に核酸の一部を付加、置換または欠失した変異核酸も含む。しばしば組換え核酸には、プロモーター配列に機能的に結合した核酸配列が含まれる。このような組換え核酸は、ベクターの不可欠なエレメントであって例えばある細胞を形質転換するために用いられるようなものであり得る。

【０１０１】

あるいはこのような組換え核酸は、ウイルスベクターの不可欠なエレメントであって例えばワクシニアウイルスに基づくものであり得る。ワクシニアウイルスは哺乳動物のワクチン接種に用いることが可能で、その際に組換え核酸が発現して哺乳動物の防御免疫応答を誘導する。

【０１０２】

「調節エレメント」は、通常は遺伝子の未翻訳領域に由来する核酸配列であり、エンハンサー、プロモーター、イントロン及び５’及び３’の未翻訳領域（ＵＴＲ）を含む。調節エレメントは、転写、翻訳またはＲＮＡ安定性を調節する宿主またはウイルスタンパク質と相互作用する。

【０１０３】

「レポーター分子」とは、核酸、アミノ酸または抗体を標識するのに用いられる化学的または生化学的成分である。レポーター分子には、放射性核種、酵素、蛍光剤、化学発光剤、発色剤、基質、補助因子、阻害因子、磁気粒子及びその他 の当分野で既知の成分がある。

【０１０４】

ＤＮＡ配列に関する「ＲＮＡ等価物」は、窒素塩基チミンが全てウラシルに置換されていることと、糖のバックボーンがデオキシリボースではなくリボースから構成されていることを除いて、参照ＤＮＡ配列と同一のヌクレオチド線形配列から構成されている。

【０１０５】

「サンプル」の語は、その最も広い意味で用いられる。ＡＰＲＧ、ＡＰＲＧをコードする核酸、またはその断片を含むと推定されるサンプルは、体液と、細胞からの抽出物や細胞から単離された染色体や細胞内小器官、膜と、細胞と、溶液中に存在するまたは基板に固定されたゲノムＤＮＡ、ＲＮＡ、ｃＤＮＡと、組織と、組織プリント等を含み得る。

【０１０６】

「特異結合」または「特異的に結合する」の語は、タンパク質またはペプチドと、アゴニスト、抗体、アンタゴニスト、小分子、任意の天然成分または合成結合成分との間の相互作用を指す。この相互作用は、タンパク質の特定の構造（例えば抗原決定基即ちエピトープ）であって結合分子が認識するものが存在するか否かに依存している。例えば、抗体がエピトープ「Ａ」に対して特異的である場合、標識された遊離したＡ及びその抗体を含む反応において、エピトープＡ（つまり遊離し、標識されていないＡ）を含むポリペプチドの存在が、抗体に結合している標識されたＡの量を低減させる。

【０１０７】

「実質上精製された」の語は、自然環境から取り除かれ、あるいは単離または分離された核酸またはアミノ酸配列であって、自然に会合するその他 の構成エレメントの少なくとも約６０％、好ましくは少なくとも約７５％、最も好ましいのは少なくとも約９０％が遊離しているものを指す。

【０１０８】

「置換」は、１個若しくは数個のアミノ酸またはヌクレオチドを各々別のアミノ酸またはヌクレオチドに置換することを意味する。

【０１０９】

「基質」は、任意の好適な固体または半固体の支持体を指すものであって、膜、フィルタ、チップ、スライド、ウエハ、ファイバー、磁性または非磁性ビーズ、ゲル、管、プレート、ポリマー、微細粒子、毛管が含まれる。基質は、凹み、溝、ピン、チャネル、孔等、

様々な表面形態を有することができ、基質表面にはポリヌクレオチドやポリペプチドが結合する。

【０１１０】

「転写イメージ」は、所与の時間、条件での固有の細胞タイプまたは組織による遺伝子発現の集合的パターンを指す。

【０１１１】

「形質転換」は、外来性のＤＮＡが受入細胞に入り込むプロセスを表す。形質転換は、本技術分野で知られている種々の方法に従って自然条件または人工条件下で生じ得るものであり、外来性の核酸配列を原核または真核宿主細胞に挿入する任意の既知の方法を基にし得る。形質転換の方法は、形質転換する宿主細胞の種類によって選択する。限定するものではないが形質転換方法には、バクテリオファージまたはウイルス感染、電気穿孔法（エレクトロポレーション）、熱ショック、リポフェクション及び微粒子銃を用いる方法がある。「形質転換された」細胞には、導入されたＤＮＡが自律的に複製するプラスミドとしてあるいは宿主染色体の一部として複製可能である安定的に形質転換された細胞が含まれる。さらに、限られた時間に一時的に導入ＤＮＡ若しくは導入ＲＮＡを発現する細胞も含まれる。

【０１１２】

ここで用いる「遺伝形質転換体」とは任意の有機体であり、限定するものではないが動植物を含み、有機体の１個若しくは数個の細胞が、ヒトの関与によって、例えば本技術分野でよく知られている形質転換技術によって導入された異種核酸を有する。核酸の細胞への導入は、直接または間接的に、細胞の前駆物質に導入することによって、計画的な遺伝子操作によって、例えば微量注射法によってあるいは組換えウイルスの導入によって行う。遺伝子操作の語は、古典的な交雑育種あるいはｉｎ ｖｉｔｒｏ受精を指すものではなく、組換えＤＮＡ分子の導入を指すものである。本発明に基づいて予期される遺伝形質転換体には、バクテリア、シアノバクテリア、真菌及び動植物がある。本発明の単離されたＤＮＡは、本技術分野で知られている方法、例えば感染、形質移入、形質転換またはトランス接合によって宿主に導入することができる。本発明のＤＮＡをこのような有機体に移入する技術はよく知られており、前出のＳａｍｂｒｏｏｋ 他 （１９８９）等の参考文献に与えられている。

【０１１３】

特定の核酸配列の「変異体」は、核酸配列１本全部の長さに対して特定の核酸配列と少なくとも４０％の相同性を有する核酸配列であると定義する。その際、デフォルトパラメータに設定した「ＢＬＡＳＴ ２ Ｓｅｑｕｅｎｃｅｓ」ツールＶｅｒｓｉｏｎ ２．０．９（１９９９年５月７日）を用いてｂｌａｓｔｎを実行する。このような核酸対は、所定の長さに対して、例えば少なくとも５０％、６０％、７０％、８０％、８５％、９０％、９１％、９２％、９３％、９４％、９５％、９６％、９７％、９８％、９９％またはそれ以上の相同性を示し得る。或る変異体は、例えば「対立遺伝子」変異体（前述）、「スプライス」変異体、「種」変異体または「多形性」変異体として説明し得る。スプライス変異体は参照分子とかなりの相同性を有し得るが、ｍＲＮＡプロセッシング中のエキソンの交互スプライシングによって通常多数のあるいは僅かな数のポリヌクレオチドを有することになる。対応するポリペプチドは、追加機能ドメインを有するかあるいは参照分子に存在するドメインが欠落していることがある。種変異体は、種相互に異なるポリヌクレオチド配列である。結果的に生じるポリペプチドは通常、相互にかなりのアミノ酸相同性を有する。多形性変異体は、与えられた種の個体間で特定の遺伝子のポリヌクレオチド配列が異なる。また、多形性変異体は、１つのヌクレオチド塩基によってポリヌクレオチド配列が変化する「１塩基多形性」（ＳＮＰ）を含み得る。ＳＮＰの存在は、例えば特定の個体群、病状または病状性向を示し得る。

【０１１４】

特定のポリペプチド配列の「変異体」は、ポリペプチド配列の１本の長さ全体で特定のポリペプチド配列に対して少なくとも４０％の相同性を有するポリペプチド配列として画定

される。ここで、デフォルトパラメータに設定した「ＢＬＡＳＴ ２ Ｓｅｑｕｅｎｃｅｓ」ツールＶｅｒｓｉｏｎ ２．０．９（１９９９年５月７日）を用いてｂｌａｓｔｐを実行する。このようなポリペプチド対は、所定の長さに対して、例えば少なくとも５０％、６０％、７０％、８０％、９０％、９１％、９２％、９３％、９４％、９５％、９６％、９７％、９８％、９９％またはそれ以上の配列同一性を示し得る。

【０１１５】

（発明）

本発明は、新規のヒトアポトシス調節因子（ＡＰＲＧ）及びＡＰＲＧをコードするポリヌクレオチドの発見に基づき、これらの組成物を利用した細胞増多症と免疫学的異常および生殖器官疾患の診断、治療、及び予防に関する。

【０１１６】

表１は、本発明の完全長ポリヌクレオチド配列及びポリペプチド配列の命名の概略である。各ポリヌクレオチド及びその対応するポリペプチドは、１つのＩｎｃｙｔｅプロジェクト識別番号（ＩｎｃｙｔｅプロジェクトＩＤ）と相関する。各ポリペプチド配列は、ポリペプチド配列識別番号（ポリペプチドＳＥＱ ＩＤ ＮＯ）とＩｎｃｙｔｅポリペプチド配列番号（ＩｎｃｙｔｅポリペプチドＩＤ）によって表示した。各ポリヌクレオチド配列は、ポリヌクレオチド配列識別番号（ポリヌクレオチドＳＥＱ ＩＤ ＮＯ）とＩｎｃｙｔｅポリヌクレオチド配列番号（ＩｎｃｙｔｅポリヌクレオチドＩＤ）によって表示した。

【０１１７】

表２は、ＧｅｎＢａｎｋタンパク質（ｇｅｎｐｅｐｔ）データベースに対するＢＬＡＳＴ分析によって同定されたような、本発明のポリペプチドとの相同性を有する配列を示している。列１と２は発明したポリペプチドをポリペプチド配列識別番号（ポリペプチドＳＥＱ ＩＤ ＮＯ）と対応するＩｎｃｙｔｅポリペプチド配列番号（ＩｎｃｙｔｅポリペプチドＩＤ）によって表示した。列３は、ＧｅｎＢａｎｋの最も近い相同体のＧｅｎＢａｎｋの識別番号（Ｇｅｎｂａｎｋ ＩＤ ＮＯ：）を示す。列４は、各ポリペプチドとそのＧｅｎＢａｎｋ相同体との間の一致を表す確率スコアを示す。列５は、ＧｅｎＢａｎｋ相同体のアノテーションを示し、更に該当箇所には適当な引用文も示す。これらを引用することを以って本明細書の一部とする。

【０１１８】

表３は、本発明のポリペプチドの様々な構造的特徴を示す。列１と２は発明した個々のポリペプチドをポリペプチド配列識別番号（ＳＥＱ ＩＤ ＮＯ）と対応するＩｎｃｙｔｅポリペプチド配列番号（ＩｎｃｙｔｅポリペプチドＩＤ）を示す。列３は、各ポリペプチドのアミノ酸残基数を示す。列４および列５はそれぞれ、ＧＣＧ配列分析ソフトウェアパッケージのＭＯＴＩＦＳプログラム（Ｇｅｎｅｔｉｃｓ Ｃｏｍｐｕｔｅｒ Ｇｒｏｕｐ，Ｍａｄｉｓｏｎ ＷＩ）によって決定された、リン酸化およびグリコシル化の可能性のある部位を示す。列６は、シグネチャ（ｓｉｇｎａｔｕｒｅ）配列、ドメイン、およびモチーフを含むアミノ酸残基を示す。列７は、タンパク質の構造／機能の分析のための分析方法を示し、該当箇所にはさらに分析方法に利用した検索可能なデータベースを示す。

【０１１９】

表２及び３は共に、本発明の各々のポリペプチドの特性を要約しており、それら特性が請求の範囲に記載されたポリペプチドがタンパク質ホスファターゼであることを確立している。例えば、ＳＥＱ ＩＤ ＮＯ：１はゼブラフィシュのデスエフェクタードメインタンパク質（Ｄｅｄｄ１）（ＧｅｎＢａｎｋ ＩＤ ｇ７６７３６３８）と５１％の同一性を有することがＢａｓｉｃ Ｌｏｃａｌ Ａｌｉｇｎｍｅｎｔ Ｓｅａｒｃｈ Ｔｏｏｌ（ＢＬＡＳＴ）によって示された（表２参照）。ＢＬＡＳＴ確率スコアは５．１ｅ－７０であり、これは観察されたポリペプチド配列アラインメントが偶然に得られる確率を示している。ＳＥＱ ＩＤ ＮＯ：１はまた、デスエフェクタードメインを有する。これは、隠れマルコフモデル（ＨＭＭ）を基にした保存されたタンパク質ファミリードメインのＰＦＡＭデータベースにおいて、統計的に有意な一致を検索して決定された（表３参照）。ＢＬＡＳＴ

解析から得られたデータによって、ＳＥＱ ＩＤ ＮＯ：１がデスエフェクタードメインを有するアポトーシス調節因子であることが裏付けられた。

【０１２０】

代わりの例として、ＳＥＱ ＩＤ ＮＯ：２はヒトのカスパーゼ動員ドメイン タンパク質７（ｇ１０１９８２０９）と３６％の同一性を有することがＢａｓｉｃ Ｌｏｃａｌ Ａｌｉｇｎｍｅｎｔ Ｓｅａｒｃｈ Ｔｏｏｌ（ＢＬＡＳＴ）によって示された（表２参照）。ＢＬＡＳＴ確率スコアは１．１ｅ－６６であり、これは観察されたポリペプチド配列アラインメントが偶然に得られる確率を示している。

【０１２１】

別の例において、ＳＥＱ ＩＤ ＮＯ：３はヒトＴ細胞白血病ウイルスＩ型（ＨＬＶ－Ｉ）Ｔａｘ－結合タンパク質（ＴＸＢＰ１５１）（ＧｅｎＢａｎｋ ＩＤ ｇ５７７６５４５）に３３％の同一性を有するが、これはＢａｓｉｃ Ｌｏｃａｌ Ａｌｉｇｎｍｅｎｔ Ｓｅａｒｃｈ Ｔｏｏｌ（ＢＬＡＳＴ）によって決定される（表２参照）。ＢＬＡＳＴ確率スコアは２．１ｅ－７３であり、これは観察されたポリペプチド配列アラインメントが偶然に得られる確率を示している。Ｔａｘ－結合タンパク質はＮＩＨ ３Ｔ３の腫瘍壊死因子によって誘発されるアポトーシスを阻止することが実証されている（Ｄｅ Ｖａｌｃｋ，Ｄ．他（１９９９）Ｏｎｃｏｇｅｎｅ １８：４１８２－４１９０）。ＭＯＴＩＦＳ、ＨＭＭＥＲ、及び ＢＬＩＭＰＳ解析よりのデータは、ＳＥＱ ＩＤ ＮＯ：３ がアポトーシス関連のタンパク質であることのより実証的な証拠を提供する。ＳＥＱ ＩＤ ＮＯ：３ はＰｒｏｓｉｔｅ の定義に一致するパターンを検索するＭＯＴＩＦＳ プログラムによって決定されるＲＧＤ モチーフ 及び ロイシン・ジッパー・モチーフを含む（表３参照）。

【０１２２】

ＳＥＱ ＩＤ ＮＯ．３ は配列をＢＬＯＣＫＳ、ＰＲＩＮＴＳ、ＤＯＭＯ、ＰＲＯＤＯＭおよびＰＦＡＭ データベースの配列と対応させて遺伝子ファミリー、配列相同性および構造的フィンガープリント領域を検索するＢＬｏｃｋｓ ＩＭＰｒｏｖｅｄ サーチャー（ＢＬＩＭＰＳ）によって裏付けられたＮＤＰ５２核ドメインタンパク質との配列相同性を含む（表３参照）。この相同性により、ＳＥＱ ＩＤ ＮＯ：３は核に位置付けられることが示される。これにより、ＳＥＱ ＩＤ ＮＯ：３のタンパク質は核標識タンパク質として使用され得る。ＳＥＱ ＩＤ ＮＯ：１－３の解析のためのアルゴリズムおよびパラメータを表７に記載する。

【０１２３】

表４に示すように、本発明の完全長ポリヌクレオチド配列は、ｃＤＮＡ配列またはゲノムＤＮＡ由来のコード（エキソン）配列を用いて、あるいはこれら２種類の配列を任意に組み合わせて構築した。列１と２は本発明のポリヌクレオチドをポリヌクレオチド配列識別番号（Ｐｏｌｙｎｕｃｌｅｏｔｉｄｅ ＳＥＱ ＩＤ ＮＯ）と対応するＩｎｃｙｔｅポリヌクレオチド コンセンサス配列番号（ＩｎｃｙｔｅポリヌクレオチドＩＤ）を示す。列３は、各ポリヌクレオチド配列の長さ（塩基対単位）を示す。列４は、例えば、ＳＥＱ ＩＤ ＮＯ：４－６を同定するため、あるいはＳＥＱ ＩＤ ＮＯ：４－６と関連するポリヌクレオチド配列とを区別するためのハイブリダイゼーションまたは増幅技術に有用なポリヌクレオチド配列の断片を示す。列５はｃＤＮＡ配列、ゲノムＤＮＡから予想されたコード配列（エキソン）及び／またはｃＤＮＡ及びゲノムＤＮＡを共に有する配列集合に対応する識別番号を示している。これらの配列は、本発明の完全長ポリヌクレオチド配列を構築するのに用いた。表４の列６および列７はそれぞれ、列５の配列に対応するｃＤＮＡ配列およびゲノム配列の開始ヌクレオチド（５’）位置および終了ヌクレオチド（３’）位置を示す。

【０１２４】

表４の列５の識別番号は、特に例えばＩｎｃｙｔｅ ｃＤＮＡとそれに対応するｃＤＮＡライブラリに照会し得る。例えば、７０１８４８２Ｈ１ はインサイトｃＤＮＡ 配列の識別番号であり、ＫＩＤＮＮＯＣ０１ はそれが由来するｃＤＮＡ ライブラリの識別番号で

ある。ｃＤＮＡライブラリが示されていないＩｎｃｙｔｅ ｃＤＮＡは、プールされているｃＤＮＡライブラリ（例えば、７２０７０２３８Ｖ１）に由来する。あるいは列５の識別番号は、完全長ポリヌクレオチド配列の構築に寄与するＧｅｎＢａｎｋ ｃＤＮＡまたはＥＳＴに照会し得る。あるいは列５の識別番号は、ゲノムＤＮＡのＧｅｎｓｃａｎ分析により予想されるコード領域に照会し得る。例えば、ＧＮＮ．ｇ８０９９７９９＿００００２０＿００６ は、Ｇｅｎｓｃａｎ推定コード配列の識別番号であって、ｇ８０９９７９９ がＧｅｎｓｃａｎ分析によって得られたＧｅｎＢａｎｋの配列の識別番号である。このＧｅｎｓｃａｎ推定コード配列は、配列を組み立てる前に編集する場合がある（例４を参照）。または列５の識別番号は、「エキソンスティッチング（ｅｘｏｎ－ｓｔｉｔｃｈｉｎｇ）」アルゴリズムにより結び合わせたｃＤＮＡ及びＧｅｎｓｃａｎ予想エキソンの両方からなる群を意味する場合もある。または、列５の識別番号は、「エキソンストレッチング」アルゴリズムによって集められたｃＤＮＡおよびＧｅｎｓｃａｎ推定エキソンの両方からなる群の場合もある（実施例５を参照）。場合によっては、列５に示されている配列の範囲と重複するＩｎｃｙｔｅ ｃＤＮＡの範囲が得られ、最終的なコンセンサス配列が決定されるが、それに相当するＩｎｃｙｔｅ ｃＤＮＡの識別番号は示されていない。

【０１２５】

表５は、Ｉｎｃｙｔｅ ｃＤＮＡ配列を用いて構築された完全長ポリヌクレオチド配列のための代表的なｃＤＮＡライブラリを示している。代表的なｃＤＮＡライブラリは、上記のポリヌクレオチド配列を構築及び確認するために用いられるＩｎｃｙｔｅ ｃＤＮＡ配列によって最も頻繁に代表されるＩｎｃｙｔｅ ｃＤＮＡライブラリである。ｃＤＮＡライブラリを作製するために用いた組織及びベクターを表５に示し、表６で説明している。

【０１２６】

本発明はまた、ＡＰＲＧの変異体も含む。好適なＡＰＲＧの変異体は、ＡＰＲＧの機能的あるいは構造的特徴の少なくともどちらか一方を有し、かつＡＰＲＧアミノ酸配列に対して少なくとも約８０％のアミノ酸配列同一性、あるいは少なくとも約９０％のアミノ酸配列同一性、更には少なくとも約９５％のアミノ酸配列同一性を有する。

本発明はまた、ＡＰＲＧをコードするポリヌクレオチドを提供する。特定の実施例において、本発明は、ＡＰＲＧをコードするＳＥＱ ＩＤ ＮＯ：４－６ からなる一群から選択された配列を含むポリヌクレオチド配列を提供する。配列表に示したＳＥＱ ＩＤ ＮＯ：４－６のポリヌクレオチド配列は、窒素系塩基のチミンがウラシルに置換され、糖鎖のバックボーンがデオキシリボースではなくリボースからなる等価ＲＮＡ配列を含む。

【０１２７】

本発明はまた、ＡＰＲＧ をコードするポリヌクレオチド配列の変異配列を含む。詳細には、このようなポリヌクレオチド配列の変異配列は、ＡＰＲＧをコードするポリヌクレオチド配列と少なくとも７０％のポリヌクレオチド配列同一性、あるいは少なくとも８５％のポリヌクレオチド配列同一性、更には少なくとも９５％ものポリヌクレオチド配列同一性を有する。本発明の特定の実施形態は、ＳＥＱ ＩＤ ＮＯ：４－４２からなる一群から選択された核酸配列と少なくとも７０％のポリヌクレオチド配列同一性、あるいは少なくとも８５％のポリヌクレオチド配列同一性、更には少なくとも９５％ものポリヌクレオチド配列同一性を有するＳＥＱ ＩＤ ＮＯ：４－６ からなる一群から選択された配列を含むポリヌクレオチド配列の変異配列を提供する。上記したポリヌクレオチド変異配列は何れも、ＡＰＲＧ の機能的あるいは構造的特徴の少なくとも１つを有するアミノ酸配列をコードする。

【０１２８】

遺伝暗号の縮重により作り出され得るＡＰＲＧ をコードする種々のポリヌクレオチド配列には、既知の自然発生する任意の遺伝子のポリヌクレオチド配列と最小の類似性しか有しないものも含まれることを、当業者は理解するであろう。したがって本発明には、可能コドン選択に基づく組合せの選択によって産出し得るようなありとあらゆる可能性のあるポリヌクレオチド配列変異体を網羅し得る。これらの組み合わせは、天然のＡＰＲＧのポ

リヌクレオチド配列に適用される標準的なトリプレット遺伝暗号を基に作られ、全ての変異が明確に開示されていると考慮する。

【０１２９】

ＡＰＲＧ をコードするヌクレオチド配列及びその変異配列は一般に、好適に選択されたストリンジェントな条件下で、天然のＡＰＲＧ のヌクレオチド配列とハイブリダイズ可能であるが、非天然のコドンを含めるなどの実質的に異なった使い方のコドンを有するＡＰＲＧ あるいはその誘導体をコードするヌクレオチド配列を作ることは有利となり得る。宿主が特定のコドンを利用する頻度に基づいて、特定の真核または原核宿主に発生するペプチドの発現率を高めるようにコドンを選択することが可能である。コードされたアミノ酸配列を変えないで、ＡＰＲＧ 及びその誘導体をコードするヌクレオチド配列を実質的に変更する別の理由は、天然の配列から作られる転写物より例えば長い半減期など好ましい特性を備えるＲＮＡ転写物を作ることにある。

【０１３０】

本発明はまた、ＡＰＲＧ 及びその誘導体をコードするＤＮＡ配列またはそれらの断片を完全に合成化学によって作り出すことも含む。作製後にこの合成配列を、当分野で良く知られた試薬を用いて、種々の入手可能な発現ベクター及び細胞系の何れの中にも挿入可能である。更に、合成化学を用いてＡＰＲＧ またはその任意の断片をコードする配列に突然変異を誘導し得る。

【０１３１】

更に本発明には、種々のストリンジェントな条件下で、請求項に記載されたポリヌクレオチド配列、特に、ＳＥＱ ＩＤ ＮＯ：４－６及びそれらの断片とハイブリダイズ可能なポリヌクレオチド配列が含まれる（例えば、Ｗａｈｌ，Ｇ．Ｍ．及びＳ．Ｌ．Ｂｅｒｇｅｒ（１９８７）Ｍｅｔｈｏｄｓ Ｅｎｚｙｍｏｌ．１５２：３９９－４０７；Ｋｉｍｍｅｌ，Ａ．Ｒ．（１９８７）Ｍｅｔｈｏｄｓ Ｅｎｚｙｍｏｌ．１５２：５０７－５１１を参照）。アニーリング及び洗浄条件を含むハイブリダイゼーションの条件は、「定義」に記載されている。

【０１３２】

ＤＮＡシークエンシングの方法は当分野でよく知られており、本発明の何れの実施例もＤＮＡシークエンシング方法を用いて実施可能である。ＤＮＡシークエンシング方法には酵素を用いることができ、例えばＤＮＡポリメラーゼＩのクレノウ断片、ＳＥＱＵＥＮＡＳＥ（ＵＳ Ｂｉｏｃｈｅｍｉｃａｌ，Ｃｌｅｖｅｌａｎｄ ＯＨ）、Ｔａｑポリメラーゼ（Ａｐｐｌｉｅｄ Ｂｉｏｓｙｓｔｅｍｓ）、熱安定性Ｔ７ポリメラーゼ（Ａｍｅｒｓｈａｍ，Ｐｈａｒｍａｃｉａ Ｂｉｏｔｅｃｈ，Ｐｉｓｃａｔａｗａｙ ＮＪ）を用いることができる。あるいは、例えばＥＬＯＮＧＡＳＥ増幅システム（Ｌｉｆｅ Ｔｅｃｈｎｏｌｏｇｉｅｓ，Ｇａｉｔｈｅｒｓｂｕｒｇ ＭＤ）において見られるように、ポリメラーゼと校正エキソヌクレアーゼを併用することができる。好適には、ＭＩＣＲＯＬＡＢ２２００液体転移システム（Ｈａｍｉｌｔｏｎ，Ｒｅｎｏ，ＮＶ）、ＰＴＣ２００サーマルサイクラー（ＭＪ Ｒｅｓｅａｒｃｈ，Ｗａｔｅｒｔｏｗｎ ＭＡ）及びＡＢＩ ＣＡＴＡＬＹＳＴ ８００サーマルサイクラー（Ａｐｐｌｉｅｄ Ｂｉｏｓｙｓｔｅｍｓ）等の装置を用いて配列の調製を自動化する。次に、ＡＢＩ ３７３ あるいは ３７７ ＤＮＡシークエンシングシステム（Ａｐｐｌｉｅｄ Ｂｉｏｓｙｓｔｅｍｓ）、ＭＥＧＡＢＡＣＥ １０００ ＤＮＡシークエンシングシステム（Ｍｏｌｅｃｕｌａｒ Ｄｙｎａｍｉｃｓ，Ｓｕｎｎｙｖａｌｅ ＣＡ）または当分野でよく知られている他 の方法を用いてシークエンシングを行う。結果として得られた配列を当分野でよく知られている種々のアルゴリズムを用いて分析する（Ａｕｓｕｂｅｌ，Ｆ．Ｍ．（１９９７）Ｓｈｏｒｔ Ｐｒｏｔｏｃｏｌｓ ｉｎ Ｍｏｌｅｃｕｌａｒ Ｂｉｏｌｏｇｙ，Ｊｏｈｎ Ｗｉｌｅｙ ＆ Ｓｏｎｓ，Ｎｅｗ Ｙｏｒｋ ＮＹ，ｕｎｉｔ ７．７、Ｍｅｙｅｒｓ，Ｒ．Ａ．（１９９５）Ｍｏｌｅｃｕｌａｒ Ｂｉｏｌｏｇｙ ａｎｄ Ｂｉｏｔｅｃｈｎｏｌｏｇｙ，Ｗｉｌｅｙ ＶＣＨ，Ｎｅｗ Ｙｏｒｋ ＮＹ，８５６－８５３ページ）。

【０１３３】

当分野で周知のＰＣＲ法をベースにした種々の方法で、部分的なヌクレオチド配列を利用して、ＡＰＲＧ をコードする核酸配列を伸長し、プロモーターや調節エレメントなどの上流にある配列を検出する。例えば、使用し得る方法の１つである制限部位ＰＣＲ法は、ユニバーサルプライマー及びネステッドプライマーを用いてクローニングベクター内のゲノムＤＮＡから未知の配列を増幅する方法である（Ｓａｒｋａｒ，Ｇ．（１９９３）ＰＣＲ Ｍｅｔｈｏｄｓ Ａｐｐｌｉｃ．２：３１８３２２を参照．）。別の方法に逆ＰＣＲ法があり、これは広範な方向に伸長させたプライマーを用いて環状化した鋳型から未知の配列を増幅する方法である。鋳型は、既知のゲノム遺伝子座及びその周辺の配列を含む制限断片から得る（Ｔｒｉｇｌｉａ，Ｔ．他（１９８８）Ｎｕｃｌｅｉｃ Ａｃｉｄｓ Ｒｅｓ．１６：８１８６）。第３の方法としてキャプチャＰＣＲ法があり、これはヒト及び酵母菌人工染色体ＤＮＡの既知の配列に隣接するＤＮＡ断片をＰＣＲ増幅する方法に関与している（Ｌａｇｅｒｓｔｒｏｍ，Ｍ．他（１９９１）ＰＣＲ Ｍｅｔｈｏｄｓ Ａｐｐｌｉｃ．１：１１１１１９を参照）。この方法では、ＰＣＲを行う前に複数の制限酵素の消化及び連結反応を用いて未知の配列領域内に組換え二本鎖配列を挿入することが可能である。また、未知の配列を検索するために用い得る別の方法については当分野で知られている（Ｐａｒｋｅｒ，Ｊ．Ｄ．他．（１９９１）Ｎｕｃｌｅｉｃ Ａｃｉｄｓ Ｒｅｓ．１９：３０５５３０６０を参照）。．更に、ＰＣＲ、ネステッドプライマー及びＰｒｏｍｏｔｅｒＦｉｎｄｅｒ（商標）ライブラリ（Ｃｌｏｎｔｅｃｈ，Ｐａｌｏ Ａｌｔｏ ＣＡ）を用いてゲノムＤＮＡをウォーキングすることができる。この手順は、ライブラリをスクリーニングする必要がなく、イントロン／エキソン接合部を見付けるのに有用である。全てのＰＣＲベースの方法に対して、市販されているソフトウェア、例えばＯＬＩＧＯ ４．０６プライマー分析ソフトウェア（Ｎａｔｉｏｎａｌ Ｂｉｏｓｃｉｅｎｃｅｓ，Ｐｌｙｍｏｕｔｈ ＭＮ）あるいは別の好適なプログラムを用いて、長さが約２２～３０ヌクレオチド、ＧＣ含有率が約５０％以上、温度約６８℃～７２℃で鋳型に対してアニーリングするようにプライマーを設計し得る。

【０１３４】

完全長ｃＤＮＡをスクリーニングする際は、より大きなｃＤＮＡを含むようにサイズ選択されたライブラリを用いるのが好ましい。更に、ランダムプライマーのライブラリは、しばしば遺伝子の５’領域を有する配列を含み、オリゴｄ（Ｔ）ライブラリが完全長ｃＤＮＡを作製できない状況に対して好適である。ゲノムライブラリは、５’非転写調節領域への配列の伸長に有用であろう。

【０１３５】

市販されているキャピラリー電気泳動システムを用いて、シークエンシングまたはＰＣＲ産物のサイズを分析し、またはそのヌクレオチド配列を確認することができる。具体的には、キャピラリーシークエンシングは、電気泳動による分離のための流動性ポリマーと、４つの異なるヌクレオチドに特異的であるような、レーザで活性化される蛍光色素と、放出された波長の検出に利用するＣＣＤカメラとを有し得る。出力／光の強度は、適切なソフトウェア（Ａｐｐｌｉｅｄ Ｂｉｏｓｙｓｔｅｍｓ社のＧＥＮＯＴＹＰＥＲ、ＳＥＱＵＥＮＣＥ ＮＡＶＩＧＡＴＯＲ等）を用いて電気信号に変換し得る。サンプルのロードからコンピュータ分析及び電子データ表示までの全プロセスがコンピュータ制御可能である。キャピラリー電気泳動法は、特定のサンプルに少量しか存在しないようなＤＮＡ小断片のシークエンシングに特に適している。

【０１３６】

本発明の別の実施例では、ＡＰＲＧ をコードするポリヌクレオチド配列またはその断片を組換えＤＮＡ分子にクローニングして、適切な宿主細胞内に ＡＰＲＧ、その断片または機能的等価物を発現させることが可能である。遺伝暗号固有の縮重により、実質的に同じあるいは機能的に等価のアミノ酸配列をコードする別のＤＮＡ配列が作られ得り、これらの配列をＡＰＲＧ のクローン化及び発現に利用可能である。

【０１３７】

種々の目的でＡＰＲＧ をコードする配列を変えるために、当分野で一般的に知られてい

る方法を用いて、本発明のヌクレオチド配列を組換えることができる。この目的には、遺伝子産物のクローン化、プロセッシング及び／または発現の調節が含まれるが、これらに限定されるものではない。遺伝子断片及び合成オリゴヌクレオチドのランダムなフラグメンテーション及びＰＣＲ再アセンブリによるＤＮＡシャッフリングを用い、ヌクレオチド配列を組み換えることが可能である。例えば、オリゴヌクレオチド仲介特定部位突然変異誘導を利用して、新規な制限部位の生成、グリコシル化パターンの変更、コドン優先の変更、スプライス変異体の生成等を行う突然変異を導入し得る。

【０１３８】

本発明のヌクレオチドは、ＭＯＬＥＣＵＬＡＲＢＲＥＥＤＩＮＧ（Ｍａｘｙｇｅｎ Ｉｎｃ．，Ｓａｎｔａ Ｃｌａｒａ ＣＡ；米国特許第５，８３７，４５８号；Ｃｈａｎｇ，Ｃ．－Ｃ．他（１９９９）Ｎａｔ．Ｂｉｏｔｅｃｈｎｏｌ．１７：７９３－７９７；Ｃｈｒｉｓｔｉａｎｓ，Ｆ．Ｃ．他（１９９９）Ｎａｔ．Ｂｉｏｔｅｃｈｎｏｌ．１７：２５９－２６４；Ｃｒａｍｅｒｉ，Ａ．他（１９９６）Ｎａｔ．Ｂｉｏｔｅｃｈｎｏｌ．１４：３１５－３１９を参照）などのＤＮＡシャフリング技術の対象となり得るもので、生物学的または酵素的な活性、あるいは他 の分子や化合物と結合する能力などのＡＴＲＳの生物学的特性を変更あるいは改良することができる。ＤＮＡシャッフリングは、遺伝子断片のＰＣＲ仲介再組換えを用いて遺伝子変異体のライブラリを生成するプロセスである。ライブラリはその後、その遺伝子変異体を所望の特性に同定するような選択またはスクリーニングにかける。次にこれらの好適な変異体をプールし、更に反復してＤＮＡシャッフリング及び選択／スクリーニングを行ってもよい。このように、遺伝の多様性は「人為的」品種改良及び急速な分子の進化を経て創生される。例えば、ランダムポイント突然変異を有する単一の遺伝子の断片を組み換えて、スクリーニングし、その後所望の特性が最適化されるまでシャッフリングすることができる。あるいは、所定の遺伝子を同種または異種のいずれかから得た同一遺伝子ファミリーの相同遺伝子と組み換えて、それによって天然に存在する複数の遺伝子の遺伝多様性を、指図された制御可能な方法で最大化させることができる。

【０１３９】

別の実施例によれば、ＡＰＲＧ をコードする配列は、当分野で周知の化学的方法を用いて、全体あるいは一部が合成可能である（例えば、Ｃａｒｕｔｈｅｒｓ．Ｍ．Ｈ．ら（１９８０）（１９８０）Ｎｕｃｌ．Ａｃｉｄｓ Ｒｅｓ．Ｓｙｍｐ．Ｓｅｒ ７：２１５－２２３；及びＨｏｒｎ，Ｔ．他（１９８０）Ｎｕｃｌ．Ａｃｉｄｓ Ｒｅｓ．Ｓｙｍｐ．Ｓｅｒ．７：２２５－２３２を参照）。別法として、化学的方法を用いてＡＰＲＧ 自体またはその断片を合成することが可能である。例えば、種々の液相または固相技術を用いてペプチド合成を行うことができる（Ｃｒｅｉｇｈｔｏｎ，Ｔ．（１９８４）Ｐｒｏｔｅｉｎｓ，Ｓｔｒｕｃｔｕｒｅｓ ａｎｄ Ｍｏｌｅｃｕｌａｒ Ｐｒｏｐｅｒｔｉｅｓ，ＷＨ Ｆｒｅｅｍａｎ，Ｎｅｗ Ｙｏｒｋ ＮＹ，５５－６０ページ、Ｒｏｂｅｒｇｅ，Ｊ．Ｙ．他（１９９５）Ｓｃｉｅｎｃｅ ２６９：２０２－２０４）。自動合成はＡＢＩ ４３１Ａペプチドシンセサイザ（Ｐｅｒｋｉｎ Ｅｌｍｅｒ）を用いて達成し得る。更にＡＰＲＧ のアミノ酸配列または任意のその一部は、直接的な合成の際の変更、及び／または化学的方法を用いた他のタンパク質または任意のその一部からの配列との組み合わせにより、天然のポリペプチド配列を有するポリペプチドまたは変異体ポリペプチドを作製することが可能である。

【０１４０】

ペプチドは、分離用高速液体クロマトグラフィーを用いて実質上精製可能である（Ｃｈｉｅｚ，Ｒ．Ｍ．ａｎｄ Ｆ．Ｚ．Ｒｅｇｎｉｅｒ（１９９０）Ｍｅｔｈｏｄｓ Ｅｎｚｙｍｏｌ．１８２：３９２－４２１等を参照）。合成ペプチドの組成は、アミノ酸分析またはシークエンシングによって確認することができる（前出のＣｒｅｉｇｈｔｏｎ，２８－５３ページ）。

【０１４１】

生物学的に活性なＡＰＲＧ を発現させるために、ＡＰＲＧ をコードするヌクレオチド配

列またはその誘導体を好適な発現ベクターに挿入する。この発現ベクターは、好適な宿主に挿入されたコーディング配列の転写及び翻訳の調節に必要なエレメントを含む。必要な要素には、ベクター及びＡＰＲＧ をコードするポリヌクレオチド配列におけるエンハンサー、構成型及び発現誘導型プロモーター、５’及び３’の非翻訳領域などの調節配列がある。このような要素は、長さ及び特異性が様々である。特定の開始シグナルによって、ＡＰＲＧ をコードする配列のより効果的な翻訳を達成することが可能である。このようなシグナルには、ＡＴＧ開始コドンと、コザック配列などの近傍の配列が含まれる。ＡＰＲＧ をコードする配列及びその開始コドン、上流の調節配列が好適な発現ベクターに挿入された場合は、更なる転写調節シグナルや翻訳調節シグナルは必要なくなるであろう。しかしながら、コーディング配列あるいはその断片のみが挿入された場合は、インフレームのＡＴＧ開始コドンを含む外来性の翻訳調節シグナルが発現ベクターに含まれるようにすべきである。外来性の翻訳要素及び開始コドンは、様々な天然物及び合成物を起源とし得る。用いられる特定の宿主細胞系に好適なエンハンサーを含めることで発現の効率を高めることが可能である（Ｓｃｈａｒｆ，Ｄ．他．（１９９４）Ｒｅｓｕｌｔｓ Ｐｒｏｂｌ．Ｃｅｌｌ Ｄｉｆｆｅｒ．２０：１２５－１６２．）。

【０１４２】

当業者に周知の方法を用いて、ＡＰＲＧ をコードする配列、好適な転写及び翻訳調節エレメントを含む発現ベクターを作製することが可能である。これらの方法には、ｉｎ ｖｉｔｒｏ組換えＤＮＡ技術、合成技術、及びｉｎ ｖｉｖｏ遺伝子組換え技術が含まれる（例えば、Ｓａｍｂｒｏｏｋ，Ｊ．他．（１９８９）Ｍｏｌｅｃｕｌａｒ Ｃｌｏｎｉｎｇ．Ａ Ｌａｂｏｒａｔｏｒｙ Ｍａｎｕａｌ，Ｃｏｌｄ Ｓｐｒｉｎｇ Ｈａｒｂｏｒ Ｐｒｅｓｓ，Ｐｌａｉｎｖｉｅｗ ＮＹ，４章及び８章，及び１６－１７章；及びＡｕｓｕｂｅｌ，Ｆ．Ｍ．他．（１９９５）Ｃｕｒｒｅｎｔ Ｐｒｏｔｏｃｏｌｓ ｉｎ Ｍｏｌｅｃｕｌａｒ Ｂｉｏｌｏｇｙ，Ｊｏｈｎ Ｗｉｌｅｙ ＆ Ｓｏｎｓ，Ｎｅｗ Ｙｏｒｋ ＮＹ，ｃｈ．９章、１３章及び１６章を参照）。

【０１４３】

種々の発現ベクター／宿主系を利用して、ＡＰＲＧ をコードする配列の保持及び発現が可能である。限定するものではないがこのような発現ベクター／宿主系には、組換えバクテリオファージ、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せた細菌や、酵母菌発現ベクターで形質転換させた酵母菌や、ウイルス発現ベクター（例えばバキュロウイルス）に感染した昆虫細胞系や、ウイルス発現ベクター（例えばカリフラワーモザイクウイルスＣａＭＶまたはタバコモザイクウイルスＴＭＶ）または細菌発現ベクター（例えばＴｉまたはｐＢＲ３２２プラスミド）で形質転換させた植物細胞系、動物細胞系などの微生物等がある（前出のＳａｍｂｒｏｏｋ、前出のＡｕｓｕｂｅｌ、Ｖａｎ Ｈｅｅｋｅ，Ｇ．ａｎｄ Ｓ．Ｍ．Ｓｃｈｕｓｔｅｒ（１９８９）Ｊ．Ｂｉｏｌ．Ｃｈｅｍ．２６４：５５０３－５５０９、；Ｅｎｇｅｌｈａｒｄ、Ｅ．Ｋ．他．（１９９４）Ｐｒｏｃ．Ｎａｔｌ．Ａｃａｄ．Ｓｃｉ．ＵＳＡ ９１：３２２４－３２２７、Ｓａｎｄｉｇ，Ｖ．他．（１９９６）Ｈｕｍ．Ｇｅｎｅ Ｔｈｅｒ．７：１９３７－１９４５、Ｔａｋａｍａｔｓｕ，Ｎ．（１９８７）ＥＭＢＯＪ．６：３０７－３１１、；『マグローヒル科学技術年鑑』（Ｔｈｅ ＭｃＧｒａｗ Ｈｉｌｌ Ｙｅａｒｂｏｏｋ ｏｆ Ｓｃｉｅｎｃｅ ａｎｄ Ｔｅｃｈｎｏｌｏｇｙ）（１９９２）ＭｃＧｒａｗ Ｈｉｌｌ Ｎｅｗ Ｙｏｒｋ ＮＹ，１９１－１９６ページ、Ｌｏｇａｎ，Ｊ．ａｎｄ Ｔ．Ｓｈｅｎｋ（１９８４）Ｐｒｏｃ．Ｎａｔｌ．Ａｃａｄ．Ｓｃｉ．ＵＳＡ ８１：３６５５－３６５９、Ｈａｒｒｉｎｇｔｏｎ，Ｊ．Ｊ．他．（１９９７）Ｎａｔ．Ｇｅｎｅｔ．１５：３４５－３５５）。レトロウイルス、アデノウイルス、ヘルペスウイルスまたはワクシニアウイルス由来の発現ベクター、または種々の細菌性プラスミド由来の発現ベクターを用いて、ヌクレオチド配列を標的器官、組織または細胞集団へ輸送することができる（Ｄｉ Ｎｉｃｏｌａ，Ｍ．他（１９９８）Ｃａｎｃｅｒ Ｇｅｎ．Ｔｈｅｒ．５（６）：３５０－３５６、Ｙｕ，Ｍ．他．（１９９３）Ｐｒｏｃ．Ｎａｔｌ．Ａｃａｄ．Ｓｃｉ．ＵＳＡ ９０（１３）：６３４０－６３４４、Ｂｕｌｌｅ

r, R.M. 他.（1985）Nature 317（6040）：813－815；McGregor, D.P. 他.（1994）Mol. Immunol. 31（3）：219－226、Verma, I.M. and N. Somia（1997）Nature 389：239－242）。本発明は使用される宿主細胞によって限定されるものではない。

【0144】

細菌系では、多数のクローニングベクター及び発現ベクターが、APRGをコードするポリヌクレオチド配列の使用目的に応じて選択可能である。例えば、APRG をコードするポリヌクレオチド配列の日常的なクローニング、サブクローニング、増殖には、PBLUESCRIPT（Stratagene, La Jolla CA）またはpSPORT1プラスミド（GIBCO BRL）などの多機能の大腸菌ベクターを用いることができる。ベクターの多数のクローニング部位に APRG をコードする配列をライゲーションするとlacZ遺伝子が破壊され、組換え分子を含む形質転換された細菌の同定のための比色スクリーニング法が可能となる。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転写、ジデオキシのシークエンシング、ヘルパーファージによる一本鎖の救出、入れ子状態の欠失の生成にも有用であろう （例えば、Van Heeke, G. および S.M. Schuster（1989）J. Biol. Chem. 264：55035509を参照）。例えば、抗体の産生のためなどに多量の APRG が必要な場合は、APRG の発現をハイレベルで誘導するベクターが使用できる。例えば、強力な誘導T5バクテリオファージプロモーターまたは誘導T7バクテリオファージプロモーターを含むベクターが使用できる。

【0145】

APRG の発現に酵母の発現系の使用が可能である。α因子、アルコールオキシダーゼ、PGHプロモーター等の構成型あるいは誘導型のプロモーターを含む多数のベクターが、酵母菌サッカロミセス－セレビジエまたはPichia pastorisに使用可能である。更に、このようなベクターは、発現したタンパク質の分泌か細胞内への保持のどちらかを誘導し、安定した増殖のために宿主ゲノムの中に外来配列を組み込む（例えば、Ausubel, 1995, 前出、Bitter, G.A. 他.（1987）Methods Enzymol. 153：516－544、及びScorer. C. A. 他.（1994）Bio／Technology 12：181－184.）。

【0146】

植物系も APRG の発現に使用可能である。APRG をコードする配列の転写は、ウイルスプロモーター、例えば単独あるいはTMV（Takamatsu, N.（1987）EMBO J 6：307－311を参照）由来のオメガリーダー配列と組み合わせて用いられるようなCaMV由来の35S及び19Sプロモーターによって促進される。あるいは、RUBISCOの小サブユニット等の植物プロモーターまたは熱ショックプロモーターを用いてもよい（例えば、Coruzzi, G. 他.（1984）EMBO J. 3：1671－1680；Broglie, R. 他.（1984）Science 224：838－843；および Winter, J. 他.（1991）Results Probl. Cell Differ. 17：85－105を参照）。これらの構成物は、直接DNA形質転換または病原体を媒介とする形質移入によって、植物細胞内に導入可能である（例えば、The McGraw Hill Yearbook of Science and Technology 他、（1992）McGraw Hill, New York NY, 191196ページを参照）。

【0147】

哺乳動物細胞においては、多数のウイルスベースの発現系を利用し得る。アデノウイルスが発現ベクターとして用いられる場合、後発プロモーター及び3連リーダー配列からなるアデノウイルス転写物／翻訳複合体に APRG をコードする配列を結合し得る。非必須のE1またはE3領域へウイルスのゲノムを挿入し、宿主細胞で APRG を発現する感染ウイルスを得ることが可能である（例えば、Logan, J. および T. Shen

ｋ（１９８４）Ｐｒｏｃ．Ｎａｔｌ．Ａｃａｄ．Ｓｃｉ．ＵＳＡ８１：３６５５３６５９を参照）。更に、ラウス肉腫ウイルス（ＲＳＶ）エンハンサー等の転写エンハンサーを用いて、哺乳動物宿主細胞における発現を増大させ得る。ＳＶ４０またはＥＢＶをベースにしたベクターを用いてタンパク質を高レベルで発現させることもできる。

【０１４８】

ヒト人工染色体（ＨＡＣ）を用いて、プラスミドに含まれ且つプラスミドから発現するものより大きなＤＮＡの断片を輸送することもできる。治療のために約６ｋｂ～１０ＭｂのＨＡＣｓを作製し、従来の輸送方法（リポソーム、ポリカチオンアミノポリマー、またはベシクル）で供給する（例えば、Ｈａｒｒｉｎｇｔｏｎ．Ｊ．Ｊ．他（１９９７）Ｎａｔ Ｇｅｎｅｔ．１５：３４５－３５５．を参照）。

【０１４９】

哺乳動物系の組換えタンパク質の長期にわたる産生のためには、株化細胞における ＡＰＲＧ の安定した発現が望ましい。例えば、発現ベクターを用いて、ＡＰＲＧ をコードする配列を株化細胞に形質転換することが可能である。このような発現ベクターは、ウイルス起源の複製及び／または内在性の発現要素や、同じあるいは別のベクターの上の選択マーカー遺伝子を含む。ベクターの導入後、選択培地に移す前に強化培地で約１～２日間細胞を増殖させることができる。選択可能マーカーの目的は選択培地への抵抗性を与えることであり、選択可能マーカーが存在することにより、導入された配列をうまく発現するような細胞の成長及び回収が可能となる。安定的に形質転換された細胞の耐性クローンは、その細胞型に適した組織培養技術を用いて増殖可能である。

【０１５０】

任意の数の選択系を用いて、形質転換細胞株を回収できる。限定するものではないがこのような選択系には、ｔｋ－単純細胞のために用いられるヘルペスウイルスチミジンキナーゼ遺伝子と、ａｐｒ－細胞のために用いられるアデニンホスホリボシルトランスフェラーゼ遺伝子がある（例えば、Ｗｉｇｌｅｒ，Ｍ．他（１９７７）Ｃｅｌｌ １１：２２３－２３２；Ｌｏｗｙ，Ｉ．ら（１９８０）Ｃｅｌｌ ２２：８１７－８２３等を参照）。また、選択の基礎として代謝拮抗物質、抗生物質あるいは除草剤への耐性を用いることができる。例えばｄｈｆｒはメトトレキセートに対する耐性を与え、ｎｅｏはアミノグリコシッドネオマイシン及びＧ－４１８に対する耐性を与え、ａｌｓはクロルスルフロンに対する耐性を、ｐａｔはホスフィノトリシンアセチルトランスフェラーゼに対する耐性を各々与える（例えば、Ｗｉｇｌｅｒ，Ｍ．他（１９８０）Ｐｒｏｃ．Ｎａｔｌ．Ａｃａｄ．Ｓｃｉ．ＵＳＡ７７：３５６７３５７０；Ｃｏｌｂｅｒｅ－Ｇａｒａｐｉｎ，Ｆ．他（１９８１）Ｊ．Ｍｏｌ．Ｂｉｏｌ．１５０：１－１４ 等を参照）。この他の選択可能な遺伝子、例えば、代謝のための細胞の必要条件を変えるｔｒｐＢ及びｈｉｓＤは、文献に記載されているアニトシアニン、緑色蛍光タンパク質（ＧＦＰ；Ｃｌｏｎｔｅｃｈ）、βグルクロニダーゼ及びその基質ＧＵＳ，ルシフェラーゼ及びその基質ルシフェリンなどの可視マーカーが用いられる。可視マーカー、例えばアントシアニン、緑色蛍光タンパク質（ＧＦＰ；Ｃｌｏｎｔｅｃｈ）、βグルクロニダーゼ及びその基質βグルクロニド、またはルシフェラーゼ及びその基質ルシフェリン等を用いてもよい。これらのマーカーを用いて、トランスフォーマントを特定するだけでなく、特定のベクター系に起因する一過性あるいは安定したタンパク質発現を定量することが可能である（Ｒｈｏｄｅｓ，Ｃ．Ａ．（１９９５）Ｍｅｔｈｏｄｓ Ｍｏｌ．Ｂｉｏｌ．５５：１２１１３１）。

【０１５１】

マーカー遺伝子発現の存在／不存在によって目的の遺伝子の存在が示唆されても、その遺伝子の存在及び発現の確認が必要な場合もある。例えば、ＡＰＲＧをコードする配列がマーカー遺伝子配列の中に挿入された場合、ＡＰＲＧをコードする配列を含む形質転換された細胞は、マーカー遺伝子機能の欠落により特定可能である。または、１つのプロモーターの制御下でマーカー遺伝子がＡＰＲＧをコードする配列と一列に配置することも可能である。誘導または選択に応答したマーカー遺伝子の発現は通常、タンデム遺伝子の発現も

示す。

【０１５２】

一般に、ＡＰＲＧ をコードする核酸配列を含み、ＡＰＲＧ を発現する宿主細胞は、当業者に周知の種々の方法を用いて特定することが可能である。これらの方法には、ＤＮＡ－ＤＮＡあるいはＤＮＡ－ＲＮＡハイブリダイゼーションや、ＰＣＲ法、核酸あるいはタンパク質の検出及び／または数量化のための膜系、溶液ベース、あるいはチップベースの技術を含むタンパク質生物学的試験法または免疫学的アッセイが含まれるが、これらに限定されるものではない。

【０１５３】

特異的なポリクローナル抗体またはモノクローナル抗体のどちらかを用いる ＡＰＲＧ の発現の検出及び計測のための免疫学的な方法は、当分野で周知である。このような技法には、酵素に結合した免疫吸着剤検定法（ＥＬＩＳＡ）、ラジオイムノアッセイ（ＲＩＡ）、フローサイトメーター（ＦＡＣＳ）などがある。ＡＰＲＧ 上の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た、２部位のモノクローナルベースイムノアッセイ（ｔｗｏ－ｓｉｔｅ，ｍｏｎｏｃｌｏｎａｌ－ｂａｓｅｄ ｉｍｍｕｎｏａｓｓａｙ）が好ましいが、競合の結合アッセイも用いることもできる。これらのアッセイ及びこれ以外のアッセイは、当分野で公知である（Ｈａｍｐｔｏｎ．Ｒ．他．（１９９０）Ｓｅｒｏｌｏｇｉｃａｌ Ｍｅｔｈｏｄｓ，ａ Ｌａｂｏｒａｔｏｒｙ Ｍａｎｕａｌ．ＡＰＳ Ｐｒｅｓｓ．Ｓｔ Ｐａｕｌ．ＭＮ，Ｓｅｃｔ．ＩＶ、Ｃｏｌｉｇａｎ，Ｊ．Ｅ．他．（１９９７）Ｃｕｒｒｅｎｔ Ｐｒｏｔｏｃｏｌｓ ｉｎ Ｉｍｍｕｎｏｌｏｇｙ，Ｇｒｅｅｎｅ Ｐｕｂ．Ａｓｓｏｃｉａｔｅｓ ａｎｄ Ｗｉｌｅｙ－Ｉｎｔｅｒｓｃｉｅｎｃｅ，Ｎｅｗ Ｙｏｒｋ ＮＹ、Ｐｏｕｎｄ，Ｊ．Ｄ．（１９９８）Ｉｍｍｕｎｏｃｈｅｍｉｃａｌ Ｐｒｏｔｏｃｏｌｓ，Ｈｕｍａｎｓ Ｐｒｅｓｓ，Ｔｏｔｏｗａ ＮＪ）。

【０１５４】

多岐にわたる標識方法及び結合方法が、当業者に知られており、様々な核酸アッセイおよびアミノ酸アッセイにこれらの方法を用い得る。ＡＰＲＧ をコードするポリヌクレオチドに関連する配列を検出するための、標識されたハイブリダイゼーションプローブあるいは ＰＣＲ プローブを生成する方法には、オリゴ標識化、ニックトランスレーション、末端標識化、または標識されたヌクレオチドを用いるＰＣＲ増幅が含まれる。あるいは、ＡＰＲＧ をコードする配列またはその任意の断片を、ｍＲＮＡプローブを産出するためのベクターにクローニングすることも可能である。このようなベクターは、当分野において知られており、市販もされており、Ｔ７、Ｔ３またはＳＰ６等の好適なＲＮＡポリメラーゼ及び標識されたヌクレオチドを加えて、ｉｎ ｖｉｔｒｏでＲＮＡプローブの合成に用いることができる。このような方法は、例えばＡｍｅｒｓｈａｍ Ｐｈａｒｍａｃｉａ Ｂｉｏｔｅｃｈ、Ｐｒｏｍｅｇａ（Ｍａｄｉｓｏｎ ＷＩ）、Ｕ．Ｓ．Ｂｉｏｃｈｅｍｉｃａｌ等から市販されている種々のキットを用いて実行することができる。検出を容易にするために用い得る好適なレポーター分子あるいは標識には、基質、補助因子、インヒビター、磁気粒子のほか、放射性核種、酵素、蛍光剤、化学発光剤、発色剤等がある。

【０１５５】

ＡＰＲＧ をコードするヌクレオチド配列で形質転換された宿主細胞は、細胞培地でのこのタンパク質の発現及び回収に好適な条件下で培養される。形質転換細胞から製造されたタンパク質が分泌されるか細胞内に留まるかは、使用される配列、ベクター、あるいはその両者に依存する。ＡＰＲＧ をコードするポリヌクレオチドを含む発現ベクターは、原核細胞膜及び真核細胞膜を透過する ＡＰＲＧ の分泌を誘導するシグナル配列を含むように設計できることは、当業者には理解されよう。

【０１５６】

更に、宿主細胞株の選択は、挿入した配列の発現を調節する能力または発現したタンパク質を所望の形に処理する能力によって行い得る。限定するものではないがこのようなポリペプチドの修飾には、アセチル化、カルボキシル化、グリコシル化、リン酸化、脂質化及

びアシル化がある。タンパク質の「プレプロ」または「プロ」形を切断するような翻訳後処理を利用して、タンパク質のターゲティング、折りたたみ及び／または活性を特定することも可能である。翻訳後の活性のための固有の細胞装置及び特徴のある機構を有する種々の宿主細胞（例えばＣＨＯ、ＨｅＬａ、ＭＤＣＫ、ＭＥＫ２９３、ＷＩ３８等）は、Ａｍｅｒｉｃａｎ Ｔｙｐｅ Ｃｕｌｔｕｒｅ Ｃｏｌｌｅｃｔｉｏｎ（ＡＴＣＣ，Ｂｅｔｈｅｓｄａ，ＶＡ）から入手可能であり、外来タンパク質の正しい修飾及び処理を確実にするように選択し得る。

【０１５７】

本発明の別の実施例では、ＡＰＲＧ をコードする自然あるいは変更された、または組換えの核酸配列を上記した任意の宿主系の融合タンパク質の翻訳となる異種配列に結合させる。例えば、市販の抗体によって認識できる異種部分を含むキメラＡＴＲＳ タンパク質が、ＡＴＲＳ 活性のインヒビターに対するペプチドライブラリのスクリーニングを容易にする。また、異種タンパク質部分及び異種ペプチド部分も、市販されている親和性基質を用いて融合タンパク質の精製を促進し得る。限定されるものではないがこのような部分には、グルタチオンＳトランスフェラーゼ（ＧＳＴ）、マルトース結合タンパク質（ＭＢＰ）、チオレドキシン（Ｔｒｘ）、カルモジュリン結合ペプチド（ＣＢＰ）、６－Ｈｉｓ、ＦＬＡＧ、ｃ－ｍｙｃ、赤血球凝集素（ＨＡ）がある。ＧＳＴは固定化グルタチオン上で、ＭＢＰはマルトース上で、Ｔｒｘはフェニルアルシンオキシド上で、ＣＢＰはカルモジュリン上で、そして６－Ｈｉｓは金属キレート樹脂上で、同族の融合タンパク質の精製を可能にする。ＦＬＡＧ、ｃ－ｍｙｃ及び赤血球凝集素（ＨＡ）は、これらのエピトープ標識を特異的に認識する市販されているモノクローナル抗体及びポリクローナル抗体を用いて、融合タンパク質の免疫親和性精製を可能にする。また、ＡＰＲＧ をコードする配列と異種タンパク質配列との間にあるタンパク質分解切断部位を融合タンパク質が含むように遺伝子操作すると、ＡＰＲＧ が精製の後に異種部分から切断され得る。融合タンパク質の発現及び精製方法は、前出のＡｕｓｕｂｅｌ（１９９５）１０章に記載されている。市販されている種々のキットを用いて融合タンパク質の発現及び精製を促進することもできる。

【０１５８】

本発明の別の実施例では、ＴＮＴ ウサギ網状赤血球可溶化液またはコムギ胚芽抽出系（Ｐｒｏｍｅｇａ）を用いて ｉｎ ｖｉｔｒｏ で放射能標識した ＡＰＲＧ の合成が可能である。これらの系は、Ｔ７、Ｔ３またはＳＰ６プロモーターと機能的に結合したタンパク質コード配列の転写及び翻訳を結合する。翻訳は、例えば^{35}Ｓメチオニンのような放射能標識したアミノ酸前駆体の存在下で起こる。

【０１５９】

本発明の ＡＰＲＧ またはその断片を用いて、ＡＰＲＧ に特異結合する化合物をスクリーニングすることができる。少なくとも１つまたは複数の試験化合物を用いて、ＡＰＲＧ への特異的な結合をスクリーニングすることが可能である。試験化合物の例には、抗体、オリゴヌクレオチド、タンパク質（例えば受容体）または小分子が挙げられる。

【０１６０】

一実施例では、このように同定された化合物は、例えばリガンドやその断片などの ＡＰＲＧ の天然のリガンド、または天然の基質、構造的または機能的な擬態性または自然結合パートナーに密接に関連している（Ｃｏｌｉｇａｎ，Ｊ．Ｅ．他（１９９１）Ｃｕｒｒｅｎｔ Ｐｒｏｔｏｃｏｌｓ ｉｎ Ｉｍｍｕｎｏｌｏｇｙ １（２）：Ｃｈａｐｔｅｒ ５．を参照） 同様に、化合物は、ＡＰＲＧ が結合する天然受容体、あるいは例えばリガンド結合部位などの少なくとも受容体のある断片に密接に関連し得る。何れの場合も、既知の技術を用いてこの化合物を合理的に設計することができる。一実施例では、このような化合物に対するスクリーニングには、分泌タンパク質あるいは細胞膜上のタンパク質の何れか一方としてＡＰＲＧ を発現する好適な細胞の作製が含まれる。好適な細胞には、哺乳動物、酵母、ショウジョウバエ、大腸菌からの細胞が含まれる。ＡＰＲＧ を発現する細胞または ＡＰＲＧ を含有する細胞膜断片を試験化合物と接触させて、ＡＰＲ

Ｇ または化合物の何れかの結合、刺激または阻害を分析する。

【０１６１】

あるアッセイは、単に試験化合物をポリペプチドに実験的に結合させ、結合を、蛍光色素、放射性同位体、酵素抱合体またはその他 の検出可能な標識により検出することができる。例えば、このアッセイは、少なくとも１つの試験化合物を、溶液中のあるいは固体支持物に固定された ＡＰＲＧ と結合させるステップと、ＡＰＲＧ とこの化合物との結合を検出するステップを含み得る。別法では、標識された競合物の存在下での試験化合物の結合の検出及び測定を行うことができる。更にこのアッセイでは、無細胞再構成標本、化学ライブラリまたは天然の生成混合物を用いて実施することができ、試験化合物は、溶液中で遊離させるか固体支持体に固定させる。

【０１６２】

本発明のＡＰＲＧまたはその断片を用いて、ＡＰＲＧ の活性を調整する化合物をスクリーニングすることが可能である。このような化合物には、アゴニスト、アンタゴニスト、或るいは部分的または逆アゴニスト等が含まれる。一実施例では、ＡＰＲＧ が少なくとも１つの試験化合物と結合する、ＡＰＲＧ の活性が許容される条件下でアッセイを実施し、試験化合物の存在下でのＡＰＲＧの活性が試験化合物不在下でのＡＰＲＧの活性と比較する。試験化合物の存在下でのＡＰＲＧの活性の変化は、ＡＰＲＧ の活性を調整する化合物の存在を示唆する。別法では、試験化合物をＡＰＲＧの活性に適した条件下でＡＰＲＧを含むｉｎ ｖｉｔｒｏまたは細胞遊離系と結合させてアッセイを実施する。これらアッセイの何れかにおいて、ＡＰＲＧの活性を調整する試験化合物は間接的に結合することが可能であり、試験化合物と直接接触する必要がない。少なくとも１つから複数の試験化合物をスクリーニングすることができる。

【０１６３】

別の実施例では、胚性幹細胞（ＥＳ細胞）における相同組換えを用いて動物モデル系内で、ＡＰＲＧまたはその哺乳動物相同体をコードするポリヌクレオチドを「ノックアウト」する。このような技術は当技術分野において周知であり、ヒト疾患動物モデルの作製に有用である（米国特許第５，１７５，３８３号及び第５，７６７，３３７号等を参照）。（米国特許第５，１７５，３８３号及び第５，７６７，３３７号等を参照）。例えば１２９／ＳｖＪ細胞株等のマウスＥＳ細胞は初期のマウス胚に由来し、培地で増殖させることができる。このＥＳ細胞は、ネオマイシンホスホトランスフェラーゼ遺伝子（ｎｅｏ： Ｃａｐｅｃｃｈｉ，Ｍ．Ｒ．（１９８９）Ｓｃｉｅｎｃｅ ２４４：１２８８－１２９２）等のマーカー遺伝子で破壊した目的の遺伝子を含むベクターで形質転換する。このベクターは、相同組換えにより宿主ゲノムの対応する領域に組み込まれる。別法では、Ｃｒｅ－ｌｏｘＰ系を用いて相同組換えを行い、組織特異的または発生段階特異的に目的遺伝子をノックアウトする（Ｍａｒｔｈ，Ｊ．Ｄ．（１９９６）Ｃｌｉｎ．Ｉｎｖｅｓｔ．９７：１９９９－２００２；Ｗａｇｎｅｒ，Ｋ．Ｕ．他（１９９７）（１９９７）Ｎｕｃｌｅｉｃ Ａｃｉｄｓ Ｒｅｓ．２５：４３２３－４３３０）。形質転換したＥＳ細胞を同定し、例えばＣ５７ＢＬ／６マウス系等から採取したマウス細胞胚盤胞に微量注入する。胚盤胞を偽妊娠メスに外科的に導入し、得られるキメラ子孫の遺伝形質を決め、これを交配させてヘテロ接合性系またはホモ接合性系を作製する。このようにして作製した遺伝子組換え動物は、可能性のある治療薬や毒性薬剤で検査することができる。

【０１６４】

ＡＰＲＧをコードするポリヌクレオチドをｉｎ ｖｉｔｒｏでヒト胚盤胞由来のＥＳ細胞において操作することが可能である。ヒトＥＳ細胞は、内胚葉、中胚葉及び外胚葉の細胞の種類を含む少なくとも８つの別々の細胞系統に分化する可能性を有する。これらの細胞系統は、例えば神経細胞、造血系統及び心筋細胞に分化する（Ｔｈｏｍｓｏｎ，Ｊ．Ａ．他（１９９８）Ｓｃｉｅｎｃｅ ２８２：１１４５－１１４７）。

【０１６５】

ＡＰＲＧをコードするポリヌクレオチドを用いて、ヒト疾患をモデルとした「ノックイン」ヒト化動物（ブタ）または遺伝子組換え動物（マウスまたはラット）を作製することが

可能である。ノックイン技術を用いて、ＡＰＲＧをコードするポリヌクレオチドの或る領域を動物ＥＳ細胞に注入し、注入した配列を動物細胞ゲノムに組み込ませる。形質転換細胞を胞胚に注入し、胞胚を上記のように移植する。遺伝子組換え子孫または近交系について研究し、可能性のある医薬品を用いて処理し、ヒトの疾患の治療に関する情報を得る。別法では、例えば ＡＰＲＧ を乳汁内に分泌するなどＡＰＲＧを過剰に発現する哺乳動物近交系は、便利なタンパク質源となり得る（Ｊａｎｎｅ，Ｊ．他（１９９８）Ｂｉｏｔｅｃｈｎｏｌ．Ａｎｎｕ．Ｒｅｖ．４：５５－７４）。

【０１６６】

（治療）

ＡＰＲＧ のある領域とアポトーシス調節因子のある領域との間に、例えば配列及びモチーフの文脈における化学的及び構造的類似性が存在する。更に、ＡＰＲＳ の発現は、癌組織に密接に関連する。従って、ＡＰＲＧ は、細胞異常増殖、免疫学的異常または生殖器官疾患に関与すると考えられる。ＡＰＲＧ の発現または活性の増大に関連する疾患の治療においては、ＡＰＲＧ の発現または活性を低下させることが望ましい。また、ＡＰＲＧ の発現または活性の低下に関連する疾患の治療においては、ＡＰＲＧ の発現または活性を増大させることが望ましい。

【０１６７】

従って、或る実施例において、ＡＰＲＧ の発現または活性の低下に関連した疾患の治療または予防のために、患者にＡＰＲＧ またはその断片や誘導体を投与することが可能である。限定するものではないが、このような疾患には、細胞増殖異常として、日光性角化症、動脈硬化、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）、骨髄線維症、発作性夜間ヘモグロビン尿症、真性多血症、乾癬、原発性血小板血症、また特に腺癌及び白血病、リンパ腫、黒色腫、骨髄腫、肉腫、及び奇形癌、具体的には、副腎、膀胱、骨、骨髄、脳、乳房、頚部、胆嚢、神経節、消化管、心臓、腎臓、肝臓、肺、筋肉、卵巣、膵臓、副甲状腺、陰茎、前立腺、唾液腺、皮膚、脾臓、精巣、胸腺、甲状腺、子宮の癌があり、自己免疫／炎症性疾患には、後天性免疫不全症候群（ＡＩＤＳ）、アジソン病、成人呼吸窮迫症候群、アレルギー、強直性脊椎炎、アミロイド症、貧血症、喘息、アテローム性動脈硬化症、自己免疫性溶血性貧血、自己免疫性甲状腺炎、自己免疫性多腺性内分泌カンジダ性外胚葉ジストロフィ（ＡＰＥＣＥＤ）、気管支炎、胆嚢炎、接触皮膚炎、クローン病、アトピー皮膚炎、皮膚筋炎、真性糖尿病、肺気腫、リンパ球毒素性一時性リンパ球減少症、胎児赤芽球症、結節性紅斑、萎縮性胃炎、糸球体腎炎、グッドパスチャー症候群、痛風、グレーブス病、橋本甲状腺炎、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、心筋または心膜炎症、骨関節炎、骨粗しょう症、膵炎、多発性筋炎、乾癬、ライター症候群、慢性関節リウマチ炎、強皮症、シェーグレン症候群、全身性アナフィラキシー、全身性エリテマトーデス、全身性硬化症、血小板減少症、潰瘍性大腸炎、ウェルナー症候群、癌合併症、血液透析、体外循環、ウイルス感染症、細菌感染症、真菌感染症、寄生虫感染症、原虫感染症、蠕虫感染症、外傷が含まれる。生殖系の疾患には、プロラクチン産生障害、不妊症（卵管疾患、排卵欠損症、子宮内膜症）、性周期障害、月経周期障害、多嚢胞性卵巣症候群、卵巣過刺激症候群、子宮内膜癌、卵巣癌、子宮筋腫、自己免疫疾患、子宮外妊娠、催奇形、乳がん、繊維嚢胞性乳房疾患、乳漏症、精子形成破壊、精子異常生理、精巣癌、前立腺癌、良性前立腺肥大、前立腺炎、パイロニー病、性交不能、男性乳癌、女性化乳房がふくまれる。

【０１６８】

別の実施例では、ＡＰＲＧ またはその断片や誘導体を発現し得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 ＡＰＲＧ の発現または活性の低下に関連した疾患を治療または予防することも可能である。

【０１６９】

更に別の実施例では、限定するものではないが上に列記した疾患を含むＴＲＩＣＨの発現または活性の低下に関連した疾患の治療または予防のために、実質的に精製されたＴＲＩＣＨを含む組成物を好適な医薬用担体と共に患者に投与することも可能である。

【０１７０】
更に別の実施例では、限定するものではないが上に列記した疾患を含むＡＰＲＧの発現または活性の低下に関連した疾患の治療または予防のために、ＡＰＲＧの活性を調節するアゴニストを患者に投与することも可能である。
【０１７１】
更なる実施例では、ＡＰＲＧの発現または活性の増大に関連した疾患の治療または予防のために、患者にＡＰＲＧのアンタゴニストを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、上記した細胞増殖異常、遺伝学的異常および生殖器官の疾患が含まれる。一実施態様では、ＡＰＲＧと特異的に結合する抗体が直接アンタゴニストとして、あるいはＡＰＲＧを発現する細胞または組織に薬剤を運ぶターゲッティングあるいは運搬機構として間接的に用いられ得る。
【０１７２】
別の実施例では、ＡＰＲＧをコードするポリヌクレオチドの相補体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ＡＰＲＧの発現または活性の増大に関連した疾患を治療または予防することも可能である。
【０１７３】
別の実施例では、本発明の任意のタンパク質、アンタゴニスト、抗体、アゴニスト、相補配列、またはベクターを、別の好適な治療薬と組み合わせて投与することもできる。併用療法で用いる好適な治療薬は、当業者が従来の医薬原理に従ってを選択し得る。治療薬と組み合わせることにより、上記した種々の疾患の治療または予防に相乗効果をもたらし得る。この方法を用いることにより少量の各薬剤で医薬効果をあげることが可能となり、それによって副作用の可能性を低減し得る。
【０１７４】
ＡＰＲＧのアンタゴニストは、当分野で一般的な方法を用いて製造することが可能である。詳しくは、精製されたＡＰＲＧを用いて抗体を作ったり、治療薬のライブラリをスクリーニングしてＡＰＲＧと特異的に結合するものの同定が可能である。ＴＲＩＣＨの抗体も、当分野で一般的な方法を用いて製造することが可能である。このような抗体には、ポリクローナル抗体、モノクローナル抗体、キメラ抗体、一本鎖、Ｆａｂフラグメント、及びＦａｂ発現ライブラリによって作られたフラグメントが含まれる。但し、これらに限定されるものではない。中和抗体（即ち二量体の形成を阻害する抗体）は通常、治療用に好適である。
【０１７５】
抗体の産生のためには、ヤギ、ウサギ、ラット、マウス、ヒト及びその他 のものを含む種々の宿主が、ＡＰＲＧまたは任意の断片、または免疫原性の特性を備えるそのオリゴペプチドの注入によって免疫化され得る。宿主の種に応じて、種々のアジュバントを用いて免疫応答を高めることもできる。限定するものではないがこのようなアジュバントには、フロイントアジュバントと、水酸化アルミニウム等のミネラルゲルアジュバントと、リゾレシチン、プルロニックポリオル、ポリアニオン、ペプチド、油性乳剤、キーホールリンペットヘモシニアン、ジニトロフェノール等の界面活性剤とがある。ヒトに用いられるアジュバントの中では、ＢＣＧ（カルメット‐ゲラン杆菌）及びコリネバクテリウム‐パルヴムが特に好ましい。
【０１７６】
ＡＰＲＧに対する抗体を誘発するために用いられるオリゴペプチド、ペプチド、または断片は、少なくとも約５個のアミノ酸からなり、一般的には約１０個以上のアミノ酸からなるものが好ましい。これらのオリゴペプチド、ペプチドまたは断片は、天然のタンパク質のアミノ酸配列の一部と同一であり且つ小さな天然の分子の全アミノ酸配列を含むことが望ましい。ＡＰＲＧアミノ酸の短いストレッチは、ＫＬＨなどの別のタンパク質の配列と融合し、キメラ分子に対する抗体が産生され得る。
【０１７７】
ＡＰＲＧに対するモノクローナル抗体は、培地内の連続した細胞株によって、抗体分子

を産生する任意の技術を用いて作製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技術には、ハイブリドーマ技術、ヒトＢ細胞ハイブリドーマ技術及びＥＢＶ－ハイブリドーマ技術がある（例えば、Ｋｏｈｌｅｒ，Ｇ．他（１９７５）Ｎａｔｕｒｅ ２５６：４９５４９７；Ｋｏｚｂｏｒ，Ｄ．他．（１９８５）Ｊ．Ｉｍｍｕｎｏｌ．Ｍｅｔｈｏｄｓ ８１：３１－４２；Ｃｏｔｅ，Ｒ．Ｊ．他（１９８３）Ｐｒｏｃ．Ｎａｔｌ．Ａｃａｄ．Ｓｃｉ．ＵＳＡ ８０：２０２６２０３０；およびＣｏｌｅ，Ｓ．Ｐ．ｅｔ ａｌ．（１９８４）Ｍｏｌ．Ｃｅｌｌ Ｂｉｏｌ．６２：１０９－１２０を参照）。

【０１７８】

更に、「キメラ抗体」作製のために発達したヒト抗体遺伝子にマウス抗体遺伝子をスプライシングするなどの技術が、好適な抗原特異性及び生物学的活性を備える分子を得るために用いられる（例えば、Ｍｏｒｒｉｓｏｎ，Ｓ．Ｌ．他（１９８４）Ｐｒｏｃ．Ｎａｔｌ．Ａｃａｄ．Ｓｃｉ．ＵＳＡ ８１：６８５１６８５５；Ｎｅｕｂｅｒｇｅｒ，Ｍ．Ｓ．他．（１９８４）Ｎａｔｕｒｅ ３１２：６０４６０８；Ｔａｋｅｄａ，Ｓ．他．（１９８５）Ｎａｔｕｒｅ ３１４：４５２４５４を参照）。別法では、当分野で周知の方法を用いて、一本鎖抗体の産生のための記載された技術を適用して、ＡＰＲＧ 特異性一本鎖抗体を生成する。関連特異性を有するがイディオタイプ組成が異なるような抗体を、ランダムな組合せの免疫グロブリンライブラリからチェーンシャッフリングによって産生することもできる（Ｂｕｒｔｏｎ Ｄ．Ｒ．（１９９１）Ｐｒｏｃ．Ｎａｔｌ．Ａｃａｄ．Ｓｃｉ．ＵＳＡ ８８：１０１３４－１０１３７）。

【０１７９】

抗体の産生は、リンパ球集団におけるｉｎ ｖｉｖｏ産生の誘導によって、あるいは免疫グロブリンライブラリのスクリーニングまたは文献に開示されているような高特異結合試薬のパネルのスクリーニングによっても行い得る（例えば、Ｏｒｌａｎｄｉ，Ｒ．他（１９８９）Ｐｒｏｃ．Ｎａｔｌ．Ａｃａｄ．Ｓｃｉ．ＵＳＡ ８６：３８３３８３７；Ｗｉｎｔｅｒ，Ｇ他（１９９１）Ｎａｔｕｒｅ ３４９：２９３２９９．を参照）。

【０１８０】

ＡＰＲＧのための特異結合部位を有する抗体を産生することもできる。例えば、限定するものではないが、このような断片には、抗体分子のペプシン消化によって作製されるＦ（ａｂ’）$_2$ 断片と、Ｆ（ａｂ’）$_2$ 断片のジスルフィド架橋を還元することによって作製されるＦａｂ断片とがある。あるいは、Ｆａｂ発現ライブラリを作製することによって、モノクローナルＦａｂ断片を所望の特異性と迅速且つ容易に同定することが可能となる（例えば、Ｈｕｓｅ，Ｗ．Ｄ．他（１９８９）Ｓｃｉｅｎｃｅ ２４６：１２７５－１２８１等を参照）。

【０１８１】

種々のイムノアッセイを用いてスクリーニングし、所望の特異性を有する抗体を同定することができる。隔離された特異性を有するポリクローナル抗体またはモノクローナル抗体の何れかを用いる競合的な結合、または免疫放射線活性のための数々のプロトコルが、当分野では周知である。通常このようなイムノアッセイには、ＡＰＲＧ とその特異性抗体との間の複合体調整の計測が含まれる。二つの非干渉性 ＡＰＲＧ エピトープに対して反応性のモノクローナル抗体を用いる、２部位モノクローナルベースのイムノアッセイが一般に利用されるが、競合的結合アッセイも利用することができる（Ｐｏｕｎｄ、前出）。

【０１８２】

ラジオイムノアッセイ技術と共にＳｃａｔｃｈａｒｄ分析などの様々な方法を用いて、ＡＰＲＧ に対する抗体の親和性を評価する。親和性を結合定数Ｋａで表すが、このＫａは、平衡状態の下で ＡＰＲＧ 抗体複合体のモル濃度を遊離抗体と遊離抗原のモル濃度で除して得られる値である。多数の ＡＰＲＧ エピトープに対して親和性が不均一なポリクローナル抗体試薬の Ｋａ は、ＡＰＲＧ に対する抗体の平均親和性または結合活性を表す。モノクローナル抗体は特定の ＡＰＲＧ エピトープに対して単一特異的であり、モノク

ローナル抗体試薬のために決定したＫａは、親和性の真の測定値を表す。Ｋａ値が10^{9}～10^{12}ｌｉｔｅｒ／ｍｏｌの高親和性抗体試薬は、ＡＰＲＧ 抗体複合体が激しい操作に耐えなければならないイムノアッセイに用いるのが好ましい。Ｋａ値が10^{6}～10^{7}ｌｉｔｅｒ／ｍｏｌの低親和性抗体試薬は、ＡＰＲＧ が抗体から最終的に活性化状態で解離する必要がある免疫精製（ｉｍｍｕｎｏｐｕｒｉｆｉｃａｔｉｏｎ）及び類似の処理に用いるのが好ましい（Ｃａｔｔｙ，Ｄ．（１９８８）Ａｎｔｉｂｏｄｉｅｓ，Ｖｏｌｕｍｅ Ｉ：Ａ Ｐｒａｃｔｉｃａｌ Ａｐｐｒｏａｃｈ．ＩＲＬ Ｐｒｅｓｓ，Ｗａｓｈｉｎｇｔｏｎ，ＤＣ；Ｌｉｄｄｅｌｌ，Ｊ．Ｅ．及びＣｒｙｅｒ，Ａ．（１９９１）Ａ Ｐｒａｃｔｉｃａｌ Ｇｕｉｄｅ ｔｏ Ｍｏｎｏｃｌｏｎａｌ Ａｎｔｉｂｏｄｉｅｓ，Ｊｏｈｎ Ｗｉｌｅｙ ＆ Ｓｏｎｓ，Ｎｅｗ Ｙｏｒｋ ＮＹ）。

【０１８３】

ポリクローナル抗体試薬の抗体価及び結合活性を更に評価して、後に使う或る適用例に対するこのような試薬の品質及び適性を決定することができる。例えば、少なくとも１～２ｍｇ／ｍｌの特異的な抗体、好ましくは５～１０ｍｇ／ｍｌの特異的な抗体を含むポリクローナル抗体試薬は一般に、ＡＰＲＧ 抗体複合体を沈殿させなければならない処理に用いられる。抗体の特異性、抗体価、結合活性、様々な適用例における抗体の品質や使用に対する指針については、一般に入手可能である。（前出のＣａｔｔｙの文献、同Ｃｏｌｉｇａｎ らの文献等を参照）。

【０１８４】

本発明の別の実施例では、ＡＰＲＧ をコードするポリヌクレオチド、またはその任意の断片や相補配列が、治療目的で使用することができる。ある実施態様では、ＡＰＲＧ をコードする遺伝子のコーディング領域や調節領域に相補的な配列やアンチセンス分子（ＤＮＡ及びＲＮＡ、修飾ヌクレオチド）を設計して遺伝子発現を変更することができる。このような技術は当分野では周知であり、センスまたはアンチセンスオリゴヌクレオチドまたは大きな断片が、ＡＰＲＧ をコードする配列の制御領域から、またはコード領域に沿ったさまざまな位置から設計可能である。（Ａｇｒａｗａｌ，Ｓ．，ｅｄ．（１９９６）Ａｎｔｉｓｅｎｓｅ Ｔｈｅｒａｐｅｕｔｉｃｓ，Ｈｕｍａｎａ Ｐｒｅｓｓ Ｉｎｃ．，Ｔｏｔａｗａ ＮＪを参照。）

治療に用いる場合、アンチセンス配列を好適な標的細胞に導入するのに好適な任意の遺伝子送達系を用いることができる。アンチセンス配列は、転写時に標的タンパク質をコードする細胞配列の少なくとも一部に相補的な配列を発現する発現プラスミドの形で細胞内に輸送することが可能である（Ｓｌａｔｅｒ，Ｊ．Ｅ．他．（１９９８）Ｊ．Ａｌｌｅｒｇｙ Ｃｌｉｎ．Ｉｍｍｕｎｏｌ．１０２（３）：４６９－４７５ 及びＳｃａｎｌｏｎ，Ｋ．Ｊ．他．（１９９５）９（１３）：１２８８－１２９６．等を参照）。アンチセンス配列はまた、例えばレトロウイルスやアデノ関連ウイルスベクター等のウイルスベクターを用いて細胞内に導入することもできる（Ｍｉｌｌｅｒ，Ａ．Ｄ．（１９９０）Ｂｌｏｏｄ ７６：２７１、前出のＡｕｓｕｂｅｌ、Ｕｃｋｅｒｔ，Ｗ．ａｎｄ Ｗ．Ｗａｌｔｈｅｒ（１９９４）Ｐｈａｒｍａｃｏｌ．Ｔｈｅｒ．６３（３）：３２３－３４７）。その他の遺伝送達機構には、リポソーム系、人工的なウイルスエンベロープ及び当分野で公知のその他の系が含まれる（Ｒｏｓｓｉ，Ｊ．Ｊ．（１９９５）Ｂｒ．Ｍｅｄ．Ｂｕｌｌ．５１（１）：２１７－２２５；Ｂｏａｄｏ、Ｒ．Ｊ．他．（１９９８）Ｊ．Ｐｈａｒｍ．Ｓｃｉ．８７（１１）：１３０８－１３１５、Ｍｏｒｒｉｓ，Ｍ．Ｃ．他．（１９９７）Ｎｕｃｌｅｉｃ Ａｃｉｄｓ Ｒｅｓ．２５（１４）：２７３０－２７３６．）。

【０１８５】

本発明の別の実施例では、ＡＰＲＧ をコードするポリヌクレオチドを、体細胞若しくは生殖細胞の遺伝子治療に用いることが可能である。遺伝子治療を行うことにより、（ｉ）遺伝子欠損症（例えばＸ染色体鎖遺伝（Ｃａｖａｚｚａｎａ－Ｃａｌｖｏ，Ｍ．他．（２０００）Ｓｃｉｅｎｃｅ ２８８：６６９－６７２）により特徴付けられる重度の複合型免疫欠損（ＳＣＩＤ）－Ｘ１の場合）、先天性アデノシンデアミナーゼ（ＡＤＡ）欠

損症に関連する重度の複合型免疫欠損（Ｂｌａｅｓｅ，Ｒ．Ｍ．他．（１９９５）Ｓｃｉｅｎｃｅ２７０：４７５－４８０、Ｂｏｒｄｉｇｎｏｎ，Ｃ．他．（１９９５）Ｓｃｉｅｎｃｅ２７０：４７０－４７５）、嚢胞性繊維症（Ｚａｂｎｅｒ，Ｊ．他．（１９９３）Ｃｅｌｌ７５：２０７－２１６：Ｃｒｙｓｔａｌ、Ｒ．Ｇ．他．（１９９５）Ｈｕｍ．Ｇｅｎｅ Ｔｈｅｒａｐｙ６：６４３－６６６、Ｃｒｙｓｔａｌ，Ｒ．Ｇ．ら．（１９９５）Ｈｕｍ．Ｇｅｎｅ Ｔｈｅｒａｐｙ６：６６７－７０３）、サラセミア（ｔｈａｌａｓｓａｍｉａ）、家族性高コレステロール血症、第ＶＩＩＩ因子若しくは第ＩＸ因子欠損に起因する血友病（Ｃｒｙｓｔａｌ，Ｒ．Ｇ．（１９９５）Ｓｃｉｅｎｃｅ２７０：４０４－４１０、Ｖｅｒｍａ，Ｉ．Ｍ．ａｎｄ Ｓｏｍｉａ．Ｎ．（１９９７）Ｎａｔｕｒｅ３８９：２３９－２４２）を治療し、（ｉｉ）条件的致死性遺伝子産物を発現させ（例えば制御不能な細胞増殖に起因する癌の場合）、（ｉｉｉ）細胞内の寄生虫（例えばヒト免疫不全ウイルス（ＨＩＶ）（Ｂａｌｔｉｍｏｒｅ，Ｄ．（１９８８）Ｎａｔｕｒｅ３３５：３９５－３９６、Ｐｏｅｓｃｂｌａ，Ｅ．他．（１９９６）Ｐｒｏｃ．Ｎａｔｌ．Ａｃａｄ．Ｓｃｉ．ＵＳＡ．９３：１１３９５－１１３９９）、Ｂ型若しくはＣ型肝炎ウイルス（ＨＢＶ、ＨＣＶ）、Ｃａｎｄｉｄａ ａｌｂｉｃａｎｓ及びＰａｒａｃｏｃｃｉｄｉｏｉｄｅｓ ｂｒａｓｉｌｉｅｎｓｉｓ等の真菌寄生虫、並びにＰｌａｓｍｏｄｉｕｍ ｆａｌｃｉｐａｒｕｍ及びＴｒｙｐａｎｏｓｏｍａ ｃｒｕｚｉ等の原虫寄生体に対する防御機能を有するタンパク質を発現させることができる。）ＡＰＲＧ の発現若しくは調節に必要な遺伝子の欠損が疾患を引き起こす場合、導入した細胞の好適な集団から ＡＰＲＧ を発現させて、遺伝子欠損によって起こる症状の発現を緩和することが可能である。

【０１８６】

本発明の更なる実施例では、ＡＰＲＧ の欠損による疾患や異常症は、ＡＰＲＧ をコードする哺乳動物発現ベクターを作製して、これらのベクターを機械的手段によって ＡＰＲＧ 欠損細胞に導入することによって治療する。ｉｎ ｖｉｖｏあるいはｅｘ ｖｉｔｒｏの細胞に用いる機械的導入技術には、（ｉ）個々の細胞内への直接的なＤＮＡ微量注射法、（ｉｉ）遺伝子銃、（ｉｉｉ）リポソームを介した形質移入、（ｉｖ）受容体を介した遺伝子導入、及び（ｖ）ＤＮＡトランスポソンの使用（Ｍｏｒｇａｎ，Ｒ．Ａ．および Ｗ．Ｆ．Ａｎｄｅｒｓｏｎ（１９９３）Ａｎｎｕ．Ｒｅｖ．Ｂｉｏｃｈｅｍ．６２：１９１－２１７、Ｉｖｉｃｓ，Ｚ．（１９９７）Ｃｅｌｌ ９１：５０１－５１０；Ｂｏｕｌａｙ，Ｊ－Ｌ．および Ｈ．Ｒｅｃｉｐｏｎ（１９９８）Ｃｕｒｒ．Ｏｐｉｎ．Ｂｉｏｔｅｃｈｎｏｌ．９：４４５－４５０を参照）がある。

【０１８７】

ＡＰＲＧ の発現に影響を及ぼし得る発現ベクターには、限定するものではないが、ＰＣＤＮＡ ３．１、ＥＰＩＴＡＧ、ＰＲＣＣＭＶ２、ＰＲＥＰ、ＰＶＡＸベクター（Ｉｎｖｉｔｒｏｇｅｎ，Ｃａｒｌｓｂａｄ ＣＡ）、ＰＣＭＶ－ＳＣＲＩＰＴ、ＰＣＭＶ－ＴＡＧ、ＰＥＧＳＨ／ＰＥＲＶ（Ｓｔｒａｔａｇｅｎｅ，Ｌａ Ｊｏｌｌａ ＣＡ）、ＰＴＥＴ－ＯＦＦ、ＰＴＥＴ－ＯＮ、ＰＴＲＥ２、ＰＴＲＥ２－ＬＵＣ、ＰＴＫ－ＨＹＧ（Ｃｌｏｎｔｅｃｈ，Ｐａｌｏ Ａｌｔｏ ＣＡ）が含まれる。ＡＰＲＧ を発現させるために、（ｉ）恒常的に活性なプロモーター（例えば、サイトメガロウイルス（ＣＭＶ）、ラウス肉腫ウイルス（ＲＳＶ）、ＳＶ４０ウイルス、チミジンキナーゼ（ＴＫ）、若しくはβ－アクチン遺伝子等）、（ｉｉ）誘導性プロモーター（例えば、市販されているＴ－ＲＥＸプラスミド（Ｉｎｖｉｔｒｏｇｅｎ）に含まれている、テトラサイクリン調節性プロモーター（Ｇｏｓｓｅｎ，Ｍ．ａｎｄ Ｈ．Ｂｕｊａｒｄ（１９９２）Ｐｒｏｃ．Ｎａｔｌ．Ａｃａｄ．Ｓｃｉ．Ｕ．Ｓ．Ａ．８９：５５４７－５５５１；Ｇｏｓｓｅｎ，Ｍ．他（１９９５）Ｓｃｉｅｎｃｅ２６８：１７６６－１７６９；Ｒｏｓｓｉ，Ｆ．Ｍ．Ｖ．ａｎｄ Ｈ．Ｍ．Ｂｌａｕ（１９９８）Ｃｕｒｒ．Ｏｐｉｎ．Ｂｉｏｔｅｃｈｎｏｌ．９：４５１－４５６を参照）、エクジソン誘導性プロモーター（市販されているプラスミドＰＶＧＲＸＲ及びＰＩＮＤに含まれている：Ｉｎｖｉｔｒｏｇｅｎ）、ＦＫ５０６／ラパマイシン誘導性プロモーター、またはＲＵ４８６／ミフェプリ

ストーン誘導性プロモーター（Ｒｏｓｓｉ，Ｆ．Ｍ．Ｖ．ａｎｄＨ．Ｍ．Ｂｌａｕ，前出）、または（ｉｉｉ）正常な個体に由来するＡＰＲＧをコードする内在性遺伝子の天然のプロモーター若しくは組織特異的プロモーターを用いることが可能である。ＡＰＲＧを発現させるために、（ｉ）恒常的に活性なプロモーター（例えば、サイトメガロウイルス（ＣＭＶ）、ラウス肉腫ウイルス（ＲＳＶ）、ＳＶ４０ウイルス、チミジンキナーゼ（ＴＫ）、若しくはβ－アクチン遺伝子等）、（ｉｉ）誘導性プロモーター（例えば、市販されているＴ－ＲＥＸプラスミド（Ｉｎｖｉｔｒｏｇｅｎ）に含まれている、テトラサイクリン調節性プロモーター（Ｇｏｓｓｅｎ，Ｍ．ａｎｄＨ．Ｂｕｊａｒｄ（１９９２）Ｐｒｏｃ．Ｎａｔｌ．Ａｃａｄ．Ｓｃｉ．Ｕ．Ｓ．Ａ．８９：５５４７－５５５１；Ｇｏｓｓｅｎ，Ｍ．他（１９９５）Ｓｃｉｅｎｃｅ２６８：１７６６－１７６９；Ｒｏｓｓｉ，Ｆ．Ｍ．Ｖ．ａｎｄＨ．Ｍ．Ｂｌａｕ（１９９８）Ｃｕｒｒ．Ｏｐｉｎ．Ｂｉｏｔｅｃｈｎｏｌ．９：４５１－４５６））、エクジソン誘導性プロモーター（市販されているプラスミドＰＶＧＲＸＲ及びＰＩＮＤに含まれている：Ｉｎｖｉｔｒｏｇｅｎ）、ＦＫ５０６／ラパマイシン誘導性プロモーター、またはＲＵ４８６／ミフェプリストーン誘導性プロモーター（Ｒｏｓｓｉ，Ｆ．Ｍ．Ｖ．ａｎｄＨ．Ｍ．Ｂｌａｕ，前出）、または（ｉｉｉ）正常な個体に由来するＡＰＲＧをコードする内在性遺伝子の天然のプロモーター若しくは組織特異的プロモーターを用いることが可能である。

【０１８８】

市販のリポソーム形質転換キット（例えばＩｎｖｉｔｒｏｇｅｎ社のＰｅｒＦｅｃｔ Ｌｉｐｉｄ Ｔｒａｎｓｆｅｃｔｉｏｎ Ｋｉｔ）を用いれば、当業者は経験にそれほど頼らないでもポリヌクレオチドを培養中の標的細胞に導入することが可能になる。別法では、リン酸カルシウム法（Ｇｒａｈａｍ．Ｆ．Ｌ．ａｎｄＡ．Ｊ．Ｅｂ（１９７３）Ｖｉｒｏｌｏｇｙ５２：４５６－４６７）若しくは電気穿孔法（Ｎｅｕｍａｎｎ，Ｂ．他．（１９８２）ＥＭＢＯ Ｊ．１：８４１－８４５）を用いて形質転換を行う。（１９８２）ＥＭＢＯ Ｊ．１：８４１８４５）．　初代培養細胞にＤＮＡを導入するためには、標準化された哺乳動物の形質移入プロトコルの修飾が必要である。

【０１８９】

本発明の別の実施例では、ＡＰＲＧ の発現に関連する遺伝子欠損によって起こる疾患や異常症は、（ｉ）レトロウイルス末端反復配列（ＬＴＲ）プロモーター若しくは独立したプロモーターのコントロール下で ＡＰＲＧ をコードするポリヌクレオチドと、（ｉｉ）好適なＲＮＡパッケージングシグナルと、（ｉｉｉ）追加のレトロウイルス・シス作用性ＲＮＡ配列及び効率的なベクターの増殖に必要なコーディング配列を伴うＲｅｖ応答性エレメント（ＲＲＥ）とからなるレトロウイルスベクターを作製して治療することができる。レトロウイルスベクター（例えばＰＦＢ及びＰＦＢＮＥＯ）はＳｔｒａｔａｇｅｎｅ社から市販されており、刊行データ（Ｒｉｖｉｅｒｅ，Ｉ．ら．（１９９５）Ｐｒｏｃ．Ｎａｔｌ．Ａｃａｄ．Ｓｃｉ．Ｕ．Ｓ．Ａ．９２：６７３３－６７３７）に基づいている。上記データを引用することをもって本明細書の一部とする。ベクターは、好適なベクター産生細胞系（ＶＰＣＬ）において増殖され、ＶＰＣＬは、標的細胞上の受容体に対する向性を有するエンベロープ遺伝子またはＶＳＶｇ等の乱交雑エンベロープタンパク質を発現する（Ａｒｍｅｎｔａｎｏ，Ｄ．他．（１９８７）Ｊ．Ｖｉｒｏｌ．６１：１６４７－１６５０、Ｂｅｎｄｅｒ，Ｍ．Ａ．他．（１９８７）Ｊ．Ｖｉｒｏｌ．６１：１６３９－１６４６、Ａｄａｍ，Ｍ．Ａ．ａｎｄＡ．Ｄ．Ｍｉｌｌｅｒ（１９８８）Ｊ．Ｖｉｒｏｌ．６２：３８０２－３８０６、Ｄｕｌｌ，Ｔ．他．（１９９８）Ｊ．Ｖｉｒｏｌ．７２：８４６３－８４７１、Ｚｕｆｆｅｒｅｙ，Ｒ．他．（１９９８）Ｊ．Ｖｉｒｏｌ．７２：９８７３－９８８０）。（１９８７）Ｊ．Ｖｉｒｏｌ．６１：１６４７－１６５０；Ｂｅｎｄｅｒ、Ｍ．Ａ．等．（１９８７）Ｊ．Ｖｉｒｏｌ．６１：１６３９－１６４６；Ａｄａｍ，Ｍ．Ａ．ａｎｄＡ．Ｄ．Ｍｉｌｌｅｒ（１９８８）Ｊ．Ｖｉｒｏｌ．６２：３８０２－３８０６；Ｄｕｌｌ，Ｔ．等．（１９９８）Ｊ．Ｖｉｒｏｌ．７２：８４６３－８４７１；Ｚ

ufferey, R. 等（１９９８）J. Virol. ７２：９８７３－９８８０）．ＲＩＧＧに付与された米国特許第５，９１０，４３４号（「Method for obtaining retrovirus packaging cell lines producing high transducing efficiency retroviral supernatant」）において、レトロウイルスパッケージング細胞系を得るための方法が開示されており、引用することをもって本明細書の一部とする。レトロウイルスベクターの増殖、細胞集団（例えばＣＤ４$^{+}$ Ｔ細胞）の形質導入、及び形質導入した細胞の患者への戻しは、遺伝子治療の分野では当業者に公知の方法であり、多数の文献に記載されている（Ranga, U. ら.（１９９７）J. Virol. ７１：７０２０－７０２９、Bauer, G. 他．（１９９７）Blood ８９：２２５９－２２６７、Bonyhadi, M. L.（１９９７）J. Virol. ７１：４７０７－４７１６、Ranga, U. 他．（１９９８）Proc. Natl. Acad. Sci. U. S. A. ９５：１２０１－１２０６、Su, L.（１９９７）Blood ８９：２２８３－２２９０）。

【０１９０】

別法では、アデノウイルス系遺伝子治療の送達系を用いて、ＡＰＲＧ の発現に関連する１あるいは複数の遺伝子異常を有する細胞に ＡＰＲＧ をコードするポリヌクレオチドを送達する。アデノウイルス系ベクターの作製及びパッケージングについては、当業者に公知である。複製欠損型アデノウイルスベクターは、免疫調節タンパク質をコードする遺伝子を膵臓の無損傷の膵島内に導入するために可変性であることが証明された（Csete, M. E. ら.（１９９５）Transplantation ２７：２６３－２６８）。使用できる可能性のあるアデノウイルスベクターは、Armentanoに付与された米国特許第５，７０７，６１８号（”Adenovirus vectors for gene therapy”）に記載されており、引用することをもって本明細書の一部とする。アデノウイルスベクターについては、Antinozzi, P. A. 他．（１９９９）Annu. Rev. Nutr. １９：５１１－５４４ 及び Verma, I. M. and N. Somia（１９９７）Nature １８：３８９：２３９－２４２も参照されたい。両文献は、引用することをもって本明細書の一部とする。（１９９９）Annu. Rev. Nutr. １９：５１１－５４４、Verma, I. M. 及び N. Somia（１９９７）Nature １８：３８９：２３９－２４２も参照されたい。両文献は、引用することをもって本明細書の一部とする。

【０１９１】

別法では、ヘルペス系遺伝子治療の送達系を用いて、ＡＰＲＧ の発現に関連する１つあるいは複数の遺伝子異常を有する標的細胞に ＡＰＲＧ をコードするポリヌクレオチドを送達する。ＨＳＶが向性を有するような中枢神経系の細胞に ＡＰＲＧ を導入する際には、単純ヘルペスウイルス（ＨＳＶ）系のベクターの使用は特に役立つ。ヘルペス系ベクターの作製及びパッケージングは、当業者に公知である。複製適格性単純ヘルペスウイルス（ＨＳＶ）Ｉ型系のベクターは、レポーター遺伝子を霊長類の眼に送達するために用いられてきた（Liu, X. 他．（１９９９）Exp. Eye Res. １６９：３８５－３９５を参照）．．ＨＳＶ－１ウイルスベクターの作製についても、DeLucaに付与された米国特許第５，８０４，４１３号（”Herpes simplex virus swains for gene transfer”）に開示されており、該特許の引用をもって本明細書の一部とする。米国特許第５，８０４，４１３号には、ヒト遺伝子治療を含む目的のために好適なプロモーターの制御下において細胞に導入される少なくとも１つの外在性遺伝子を有するゲノムを含む組換えＨＳＶ ｄ９２についての記載がある。上記特許はまた、ＩＣＰ４、ＩＣＰ２７及びＩＣＰ２２のために除去される組換えＨＳＶ系統の作製及び使用について開示している。ＨＳＶベクターについては、Goins, W. F. 他．（１９９９）J. Virol. ７３：５１９－５３２ 及び Xu, H. 他．（１９９４）Dev. Biol. １６３：１５２－１６１も参照されたい。両文献は、引用をもって本明細書の一部とする。 クローン化ヘルペスウイルス配列の操

作、巨大ヘルペスウイルスのゲノムの異なった部分を含む多数のプラスミドを形質移入した後の組換えウイルスの産生、ヘルペスウイルスの成長及び増殖、並びにヘルペスウイルスの細胞への感染は、当業者に公知の技術である。

【０１９２】

別法では、αウイルス（正の一本鎖ＲＮＡウイルス）ベクターを用いて ＡＰＲＧ をコードするポリヌクレオチドを標的細胞に送達する。プロトタイプのαウイルスであるセムリキ森林熱ウイルス（Ｓｅｍｌｉｋｉ Ｆｏｒｅｓｔ Ｖｉｒｕｓ，ＳＦＶ）の生物学的研究が広範に行われており、遺伝子導入ベクターがＳＦＶゲノムに基づいていることが分かった（Ｇａｒｏｆｆ，Ｈ．およびＫ．－Ｊ．Ｌｉ（１９９８）Ｃｕｎ．Ｏｐｉｎ．Ｂｉｏｔｅｃｈ．９：４６４－４６９）。αウイルスＲＮＡの複製中に、通常はウイルスのキャプシッドタンパク質をコードするサブゲノムＲＮＡが作り出される。このサブゲノムＲＮＡは、完全長のゲノムＲＮＡより高いレベルに複製されるため、酵素活性（例えばプロテアーゼ及びポリメラーゼ）を有するウイルスタンパク質に比べてキャプシッドタンパク質が過剰産生される。同様に、ＡＰＲＧ をコードする配列をαウイルスゲノムのカプシドをコードする領域に導入することによって、ベクター導入細胞において多数の ＡＰＲＧ をコードするＲＮＡが産生され、高いレベルで ＡＰＲＧ が合成される。通常はαウイルスの感染が数日以内での細胞溶解に関係する一方で、シンドビスウイルス（ＳＩＮ）の変異体を有するハムスター正常腎臓細胞（ＢＨＫ－２１）の持続的な感染を確立する能力は、αウイルスの溶解複製を遺伝子治療に適用できるように好適に変更可能であることを示唆している（Ｄｒｙｇａ，Ｓ．Ａ．ら．（１９９７）Ｖｉｒｏｌｏｇｙ ２２８：７４－８３を参照）。様々な宿主にαウイルスを導入できることから、様々なタイプの細胞に ＡＰＲＧ を導入することできる。或る集団におけるサブセットの細胞の特定形質導入は、形質導入前に細胞の選別を必要とし得る。αウイルスの感染性ｃＤＮＡクローンの処置方法、αウイルスのｃＤＮＡ及びＲＮＡの形質移入方法及びαウイルスの感染方法は、当業者に公知である。

【０１９３】

転写開始部位由来のオリゴヌクレオチドを用いて遺伝子発現を阻害することも可能である。転写開始部位とは例えば開始部位から数えて約－１０と約＋１０の間である。同様に、三重らせん塩基対の形成方法を用いて阻害が可能となる。三重らせん塩基対形成は、ポリメラーゼ、転写因子または調節分子の結合のために十分に開くような二重らせんの能力を阻害するので、三重らせん塩基対形成は有用である。三重らせんＤＮＡを用いる最近の治療の進歩については文献に記載がある （Ｇｅｅ，Ｊ．Ｅ．他．（１９９４）ｉｎ：Ｈｕｂｅｒ，Ｂ．Ｉ．Ｃａｒｒ，Ｍｏｌｅｃｕｌａｒ ａｎｄ Ｉｍｍｕｎｏｌｏｇｉｃ Ａｐｐｒｏａｃｈｅｓ，Ｆｕｔｕｒａ Ｐｕｂｌｉｓｈｉｎｇ，Ｍｔ．Ｋｉｓｃｏ，ＮＹ １６３－１７７ページを参照）。 相補配列またはアンチセンス分子もまた、転写物がリボソームに結合するのを阻止することによってｍＲＮＡの翻訳を阻止するように設計することができる。

【０１９４】

リボザイムは酵素性ＲＮＡ分子であり、ＲＮＡの特異的切断を触媒するためにリボザイムを用いることもできる。リボザイム作用のメカニズムは、ヌクレオチド鎖切断に先立つ相補的標的ＲＮＡへのリボザイム分子の配列特異性ハイブリダイゼーションに関与している。例えば、ＡＰＲＧ をコードする配列のヌクレオチド鎖切断を、特異的且つ効果的に触媒する組換え型のハンマーヘッド型リボザイム分子が含まれる。

【０１９５】

任意のＲＮＡ標的内の特異的リボザイム切断部位は、ＧＵＡ、ＧＵＵ、ＧＵＣ配列を含めたリボザイム切断部位に対する標的分子をスキャンすることによって先ず同定される。一度同定すると、オリゴヌクレオチドを機能不全にするような２次構造の特徴に対して切断部位を含む標的遺伝子の領域に対応する１５～２０リボヌクレオチドの短いＲＮＡ配列を、評価することが可能になる。候補標的の適合性の評価も、リボヌクレアーゼ保護アッセイを用いて相補的オリゴヌクレオチドとのハイブリダイゼーションの実施容易性をテスト

することによって行うことができる。

【０１９６】

本発明の相補リボ核酸分子及びリボザイムは、核酸分子合成のために当分野でよく知られている任意の方法を用いて作製し得る。任意の方法には、固相フォスフォアミダイト化合物等のオリゴヌクレオチドを化学的に合成する方法がある。あるいは、ＡＰＲＧ をコードするＤＮＡ配列のｉｎ ｖｉｔｒｏ及びｉｎ ｖｉｖｏ転写によってＲＮＡ分子を産出し得る。このようなＤＮＡ配列は、Ｔ７やＳＰ６等の好適なＲＮＡポリメラーゼプロモーターを用いて多様なベクター内に取り込むことが可能である。あるいは、相補的ＲＮＡを構成的あるいは誘導的に合成するようなこれらｃＤＮＡ産物を、細胞系、細胞または組織内に導入することができる。

【０１９７】

細胞内の安定性を高め、半減期を長くするためにＲＮＡ分子を修飾することができる。限定するものではないが可能な修飾には、分子の５’末端、３’末端、あるいはその両方においてフランキング配列を追加したり、分子の主鎖内においてホスホジエステラーゼ結合ではなくホスホロチオネートまたは２’ Ｏ－メチルを使用したりすることが含まれる。この概念は、ＰＮＡの産出に固有のものであり、これら全ての分子に拡大することができる。それには、内在性エンドヌクレアーゼによって容易には認識されないアデニン、シチジン、グアニン、チミン、及びウリジンにアセチル－、メチル－、チオ－及び同様の修飾をしたものの他 、非従来型塩基、例えばイノシン、クエオシン（ｑｕｅｏｓｉｎｅ）、ワイブトシン（ｗｙｂｕｔｏｓｉｎｅ）等を加えることができる。

【０１９８】

本発明の更なる実施例は、ＡＰＲＧ をコードするポリヌクレオチドの発現の変化に有効な化合物をスクリーニングする方法を含む。限定するものではないが特異ポリヌクレオチドの変異発現に有効な化合物には、オリゴヌクレオチド、アンチセンスオリゴヌクレオチド、三重らせん形成オリゴヌクレオチド、転写因子その他 のポリペプチド転写制御因子、及び特異ポリヌクレオチド配列と相互作用し得る非高分子化学的実体がある。有効な化合物は、ポリヌクレオチド発現のインヒビターまたはエンハンサーのいずれかとして作用することによりポリヌクレオチド発現を変異し得る。従って、ＡＰＲＧ の発現または活性の増加に関連する疾患の治療においては、ＡＰＲＧ をコードするポリヌクレオチドの発現を特異的に阻害する化合物が治療上有用であり、ＡＰＲＧ の発現または活性の低下に関連する疾患の治療においては、ＡＰＲＧ をコードするポリヌクレオチドの発現を特異的に促進する化合物が治療上有用であり得る。

【０１９９】

特異ポリヌクレオチドの変異発現における有効性に対して、少なくとも１個から複数個の試験化合物をスクリーニングし得る。試験化合物は、当分野で通常知られている任意の方法により得られる。このような方法には、ポリヌクレオチドの発現を変異させる場合と、既存の、市販のまたは専売の、天然または非天然の化合物ライブラリから選択する場合と、標的ポリヌクレオチドの化学的及び／または構造的特性に基づく化合物を合理的にデザインする場合と、組合せ的にまたは無作為に生成した化合物のライブラリから選択する場合に有効であることが知られているような化合物の化学修飾がある。ＡＰＲＧ をコードするポリヌクレオチドを含むサンプルは、少なくとも１つの試験化合物に曝露して得る。サンプルには例えば、無傷細胞、透過化処理した細胞、無細胞再構成系または再構成生化学系があり得る。ＡＰＲＧ をコードするポリヌクレオチドの発現における変化は、当分野で周知の任意の方法でアッセイする。通常、ＡＰＲＧ をコードするポリヌクレオチドの配列に相補的なヌクレオチド配列を有するプローブを用いたハイブリダイゼーションにより、特定のヌクレオチドの発現を検出する。ハイブリダイゼーション量を定量し、それによって１つ以上の試験化合物に曝露される及び曝露されないポリヌクレオチドの発現の比較に対する基礎を形成し得る。試験化合物に曝露されるポリヌクレオチドの発現における変化の検出は、ポリヌクレオチドの発現を変異する際に試験化合物が有効であることを示している。特異ポリヌクレオチドの変異発現に有効な化合物に対して、例えばＳｃｈｉ

ｚｏｓａｃｃｈａｒｏｍｙｃｅｓ ｐｏｍｂｅ遺伝子発現系（Ａｔｋｉｎｓ，Ｄ．他．（１９９９）米国特許第５，９３２，４３５号、Ａｒｎｄｔ，Ｇ．Ｍ．他．（２０００）Ｎｕｃｌｅｉｃ Ａｃｉｄｓ Ｒｅｓ．２８：Ｅ１５）またはＨｅＬａ細胞等のヒト細胞系（Ｃｌａｒｋｅ，Ｍ．Ｌ．他．（２０００）Ｂｉｏｃｈｅｍ．Ｂｉｏｐｈｙｓ．Ｒｅｓ．Ｃｏｍｍｕｎ．２６８：８－１３）を用いてスクリーニングを実行する。本発明の特定の実施例は、特異的ポリヌクレオチド配列に対するアンチセンス活性のためのオリゴヌクレオチド（デオキシリボヌクレオチド、リボヌクレオチド、ペプチド核酸、修飾オリゴヌクレオチド）の組み合わせライブラリをスクリーニングすることに関与している（Ｂｒｕｉｃｅ，Ｔ．Ｗ．他．（１９９７）米国特許第５，６８６，２４２号、Ｂｒｕｉｃｅ，Ｔ．Ｗ．他．（２０００）米国特許第６，０２２，６９１号）。

【０２００】

ベクターを細胞または組織に導入する多数の方法が利用可能であり、ｉｎ ｖｉｖｏ、ｉｎ ｖｉｔｒｏ及びｅｘ ｖｉｖｏの使用に対して同程度に適している。ｅｘ ｖｉｖｏ治療の場合、ベクターを患者から採取した幹細胞内に導入し、クローニング増殖して同一患者に自家移植で戻すことができる。トランスフェクション、リボソーム注入またはポリカチオンアミノポリマーによる輸送は、当分野でよく知られている方法を用いて実行することができる（例えば、（１９９７）Ｎａｔ Ｂｉｏｔｅｃｈｎｏｌ．１５：４６２－４６６．を参照）。

【０２０１】

上記の治療方法はいずれも、例えば、ヒト、イヌ、ネコ、ウシ、ウマ、ウサギ、サル等の哺乳動物を含めて治療が必要な全ての対象に適用できる。

【０２０２】

本発明の追加実施例は、通常薬剤として許容できる賦形剤で処方される活性成分を有する成分の投与に関連する。賦形剤には例えば、糖、でんぷん、セルロース、ゴム及びタンパク質がある。様々な処方が通常知られており、詳細はＲｅｍｉｎｇｔｏｎ’ｓ Ｐｈａｒｍａｃｅｕｔｉｃａｌ Ｓｃｉｅｎｃｅｓ（Ｍａａｃｋ Ｐｕｂｌｉｓｈｉｎｇ，Ｅａｓｔｏｎ ＰＡ）の最新版に記載されている。このような組成物は、ＡＰＲＧ、ＡＰＲＧの抗体、擬態、アゴニスト、アンタゴニスト、またはＡＰＲＧのインヒビターなどからなる。

本発明に用いられる成分は、任意の数の経路によって投与することができ、限定するものではないが経路には、経口、静脈内、筋肉内、動脈内、骨髄内、クモ膜下腔内、心室内、肺、経皮、皮下、腹腔内、鼻腔内、腸内、局所、舌下または直腸がある。

【０２０３】

肺から投与する成分は、液状または乾燥粉末状で調製し得る。このような成分は通常、患者が吸入する直前にエアロゾル化する。小分子（例えば伝統的な低分子量有機薬）の場合には、速効製剤のエアロゾル送達は当分野で公知である。高分子（例えばより大きなペプチド及びタンパク質）の場合には、当該分野において肺の肺胞領域を介しての肺送達が最近向上したことにより、インスリン等の薬剤を実質的に血液循環へ輸送することを可能にした（Ｐａｔｔｏｎ，Ｊ．Ｓ．ら，米国特許第５，９９７，８４８号等を参照）。肺送達は、針注射なしに投与する点で優れており、有毒な可能性のある浸透エンハンサーの必要性をなくす。

【０２０４】

本発明での使用に適した成分には、所定の目的を達成するために必要なだけの量の活性成分を含有する成分が含まれる。有効投与量の決定は、当業者の能力の範囲内で行う。

【０２０５】

ＡＰＲＧまたはその断片を含む高分子を直接細胞内に送達するべく、特殊な形態に組成物が調製されるのが好ましい。例えば、細胞不透過性高分子を含むリポソーム製剤は、細胞融合及び高分子の細胞内送達を促進し得る。別法では、ＡＰＲＧまたはその断片をＨＩＶ Ｔａｔ－１タンパク質の陽イオンＮ末端部に結合することもできる。このようにして生成された融合タンパク質は、マウスモデル系の脳を含む全ての組織の細胞に形質導入す

ることがわかっている（Ｓｃｈｗａｒｚｅ，Ｓ．Ｒ．他．（１９９９）Ｓｃｉｅｎｃｅ２８５：１５６９－１５７２）。

【０２０６】

任意の化合物に対して、細胞培養アッセイ、例えば新生物性細胞の細胞培養アッセイにおいて、あるいは、動物モデル、例えばマウス、ウサギ、イヌまたはブタ等において、先ず治療の有効投与量を推定することができる。動物モデルはまた、好適な濃度範囲及び投与経路を決定するためにも用い得る。このような情報を用いて、次にヒトに対する有益な投与量及び投与経路を決定することができる。

【０２０７】

医学的に効果的な薬用量は、症状や容態を回復させる、たとえばＡＰＲＧまたはその断片、ＡＰＲＧの抗体、ＡＰＲＧのアゴニストまたはアンタゴニスト、インヒビターなどの活性処方成分の量に関連する。治療有効度及び毒性は、細胞培養または動物実験における標準的な薬剤手法によって、例えばED_{50}（集団の５０％の医薬的有効量）またはLD_{50}（集団の５０％の致死量）を測定するなどして決定することができる。毒性効果の薬用効果に対する投与量の比は、治療指数であり、LD_{50}/ED_{50}比として表すことができる。高い治療指数を示すような成分が望ましい。細胞培養アッセイ及び動物実験から得られたデータは、ヒトに用いるための投与量の範囲を調剤するのに用いられる。このような組成物が含まれる投与量は、毒性を殆どあるいは全く含まず、ED_{50}を含むような血中濃度の範囲にあることが好ましい。用いられる投与形態、患者の感受性及び投与の経路によって、投与量はこの範囲内で様々に変わる。

【０２０８】

正確な投与量は、治療が必要な被験者に関する要素を考慮して、現場の医者が決定することになる。効果的なレベルの活性成分を与え、あるいは所望の効果を維持するべく、投与量及び投与を調節する。被験者に関する要素としては、疾患の重症度、患者の通常の健康状態、患者の年齢、体重及び性別、投与の時間及び頻度、薬剤の配合、反応感受性及び治療に対する応答等を考慮する。作用期間が長い成分は、特定の製剤の半減期及びクリアランス率によって３～４日毎に１度、１週間に１度、あるいは２週間に１度の間隔で投与し得る。

【０２０９】

通常の投与量は、投与の経路にもよるが約０．１～１００，０００μｇであり、合計で約１ｇまでとする。特定の投与量及び送達方法に関するガイダンスは文献に記載されており、現場の医者は通常それを利用することができる。当業者は、タンパク質またはインヒビターに対する処方とは異なる、ヌクレオチドに対する処方を利用することになる。同様に、ポリヌクレオチドまたはポリペプチドの送達は、特定の細胞、状態、位置等に特異的なものとなる。

【０２１０】

（診断）

別の実施例では、ＡＰＲＧの発現によって特徴付けられる疾患の診断のために、あるいはＡＰＲＧやＡＰＲＧのアゴニスト、アンタゴニストまたは阻害剤で治療を受けている患者をモニターするためのアッセイにおいて、ＡＰＲＧを特異的に結合する抗体が用いられることがある。診断目的に有用な抗体は、上記の治療の箇所で記載した方法と同じ方法で調合される。ＡＰＲＧの診断アッセイには、抗体及び標識を用いてヒトの体液あるいは細胞や組織から採取されたものか他ＡＰＲＧを検出する方法が含まれる。この抗体は修飾されたものもされていないものも可能であり、レポーター分子との共有結合または非共有結合で標識化できる。レポーター分子としては広くさまざまな種類が本分野で知られており、また使用可能であるが、そのうちのいくつかは上記で説明されている。

【０２１１】

ＡＰＲＧを測定するためのＥＬＩＳＡ，ＲＩＡ，及びＦＡＣＳを含む種々のプロトコルは、当分野では周知であり、変わったあるいは異常なレベルのＡＰＲＧの発現を診断する元となるものを提供する。正常あるいは標準的なＡＰＲＧの発現の値は、複合体の形

成に適した条件下で、正常な哺乳動物、例えばヒトなどの被験者から採取した体液または細胞とＡＰＲＧ に対する抗体とを結合させることによって決定する。標準複合体形成量は、種々の方法、例えば測光法で定量できる。被験者、対照及び疾患の生検組織からのサンプルの ＡＰＲＧ の発現の量が基準値と比較される。標準値と被験者との偏差が疾患を診断するパラメータとなる。

【０２１２】

本発明の別の実施例によれば、ＡＰＲＧ をコードするポリヌクレオチドを診断のために用いることもできる。用いることができるポリヌクレオチドには、オリゴヌクレオチド配列、相補的ＲＮＡ及びＤＮＡ分子、そしてＰＮＡが含まれる。このポリヌクレオチドを用いて、疾患と相関し得る ＡＰＲＧ を発現する生検組織における遺伝子の発現を検出し定量する。この診断アッセイを用いて、ＡＰＲＧ の存在の有無、更に過剰な発現を調べ、治療中の ＡＰＲＧ 値の調節を監視する。

【０２１３】

一実施形態では、ＡＰＲＧ または近縁の分子をコードする遺伝子配列を含むポリヌクレオチド配列を検出可能なＰＣＲプローブを用いたハイブリダイゼーションによって、ＡＰＲＧ をコードする核酸配列を同定することが可能である。例えば５’調節領域である高度に特異的な領域か、例えば保存されたモチーフであるやや特異性の低い領域から作られているかのプローブの特異性と、ハイブリダイゼーションあるいは増幅のストリンジェントは、プローブが ＡＰＲＧ をコードする自然界の配列のみを同定するかどうか、あるいは対立遺伝子や関連配列コードする自然界の配列のみを同定するかどうかによって決まるであろう。

【０２１４】

プローブはまた、関連する配列の検出に利用され、ＡＰＲＧ をコードする任意の配列と少なくとも５０％の配列同一性を有し得る。目的の本発明のハイブリダイゼーションプローブには、ＤＮＡあるいはＲＮＡが可能であり、ＳＥＱ ＩＤ ＮＯ：４－６の配列、あるいは ＡＰＲＧ 遺伝子のプロモーター、エンハンサー、イントロンを含むゲノム配列に由来し得る。

【０２１５】

ＡＰＲＧ をコードするＤＮＡに対する特異的ハイブリダイゼーションプローブを作製する方法には、ＡＰＲＧまたはＡＰＲＧ誘導体をコードするポリヌクレオチド配列を、ｍＲＮＡプローブを作製するためのベクターにクローニングする方法が含まれる。ｍＲＮＡプローブ作製のためのベクターは、当業者に知られており、市販されており、好適なＲＮＡポリメラーゼ及び好適な標識されたヌクレオチドを加えることによって、ｉｎ ｖｉｔｒｏでＲＮＡプローブを合成するために用いられ得る。ハイブリダイゼーションプローブは、種々のレポーターの集団によって標識され得る。レポーター集団の例としては、^{32}Ｐまたは^{35}Ｓ等の放射性核種、あるいはアビジン／ビオチン結合系を介してプローブに結合されたアルカリホスファターゼ等の酵素標識などが挙げられる。

【０２１６】

ＡＰＲＧをコードするポリヌクレオチド配列を用いて、ＡＰＲＧの発現に関連する疾患を診断することが可能である。限定するものではないが、このような疾患には、細胞増殖異常として、日光性角化症、動脈硬化、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）、骨髄線維症、発作性夜間ヘモグロビン尿症、真性多血症、乾癬、原発性血小板血症、また特に腺癌及び白血病、リンパ腫、黒色腫、骨髄腫、肉腫、及び奇形癌、具体的には、副腎、膀胱、骨、骨髄、脳、乳房、頚部、胆嚢、神経節、消化管、心臓、腎臓、肝臓、肺、筋肉、卵巣、膵臓、副甲状腺、陰茎、前立腺、唾液腺、皮膚、脾臓、精巣、胸腺、甲状腺、子宮の癌があり、自己免疫／炎症性疾患には、後天性免疫不全症候群（ＡＩＤＳ）、アジソン病、成人呼吸窮迫症候群、アレルギー、強直性脊椎炎、アミロイド症、貧血症、喘息、アテローム性動脈硬化症、自己免疫性溶血性貧血、自己免疫性甲状腺炎、自己免疫性多腺性内分泌カンジダ性外胚葉ジストロフィ（ＡＰＥＣＥＤ）、気管支炎、胆嚢炎、接触皮膚炎、クローン病、アトピー皮膚炎、皮膚筋炎、真性糖尿

病、肺気腫、リンパ球毒素性一時性リンパ球減少症、胎児赤芽球症、結節性紅斑、萎縮性胃炎、糸球体腎炎、グッドパスチャー症候群、痛風、グレーブス病、橋本甲状腺炎、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、心筋または心膜炎症、骨関節炎、骨粗しょう症、膵炎、多発性筋炎、乾癬、ライター症候群、慢性関節リウマチ炎、強皮症、シェーグレン症候群、全身性アナフィラキシー、全身性エリテマトーデス、全身性硬化症、血小板減少症、潰瘍性大腸炎、ウェルナー症候群、癌合併症、血液透析、体外循環、ウイルス感染症、細菌感染症、真菌感染症、寄生虫感染症、原虫感染症、蠕虫感染症、外傷が含まれる。生殖系の疾患には、プロラクチン産生障害、不妊症（卵管疾患、排卵欠損症、子宮内膜症）、性周期障害、月経周期障害、多嚢胞性卵巣症候群、卵巣過刺激症候群、子宮内膜癌、卵巣癌、子宮筋腫、自己免疫疾患、子宮外妊娠、催奇形、乳がん、繊維嚢胞性乳房疾患、乳漏症、精子形成破壊、精子異常生理、精巣癌、前立腺癌、良性前立腺肥大、前立腺炎、パイロニー病、性交不能、男性乳癌、女性化乳房がふくまれる。ＡＰＲＧをコードするポリヌクレオチド配列は、サザーン法やノーザン法、ドットブロット法、あるいはその他の膜系の技術、ＰＣＲ法、ディップスティック（ｄｉｐｓｔｉｃｋ）、ピン（ｐｉｎ）、ＥＬＩＳＡ式アッセイ、及び変異ＡＰＲＧの発現を検出するために患者から採取した体液あるいは組織を利用するマイクロアレイに使用することが可能である。このような定性方法または定量方法は、当分野で公知である。

【０２１７】

ある実施態様では、ＡＰＲＧをコードするヌクレオチド配列は、関連する疾患、特に上記した疾患を検出するアッセイにおいて有用であろう。ＡＰＲＧをコードするヌクレオチド配列は、標準的な方法で標識化され、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、患者から採取した体液あるいは組織のサンプルに加えることができるであろう。好適なインキュベーション期間が経過したらサンプルを洗浄し、シグナルを定量して標準値と比較する。患者のサンプルのシグナルの量が、対照サンプルと較べて著しく変わっている場合は、サンプル内のＡＡＰＲＧをコードするヌクレオチド配列の変異レベルにより、関連する疾患の存在が明らかになる。このようなアッセイは、動物実験、臨床試験における特定の治療効果を推定するため、あるいは個々の患者の治療をモニターするために用いることもできる。

【０２１８】

ＡＰＲＧの発現に関連する疾患の診断の基準となるものを提供するために、発現の正常すなわち標準的なプロファイルが確立される。これは、ハイブリダイゼーションあるいは増幅に好適な条件下で、動物あるいはヒトの何れかの正常な被験者から抽出された体液あるいは細胞と、ＡＰＲＧをコードする配列あるいはその断片とを結合させることにより達成され得る。実質的に精製されたポリヌクレオチドを既知量用いて行った実験から得た値を正常な対象から得た値と比較することにより、標準ハイブリダイゼーションを定量することができる。このようにして得た標準値は、疾患の徴候を示す患者から得たサンプルから得た値と比較することができる。標準値からの偏差を用いて疾患の存在を確定する。

【０２１９】

疾患の存在が確定されて治療プロトコルが開始されると、患者の発現レベルが正常な被検者に観察されるレベルに近づき始めたかどうかを測定するため、ハイブリダイゼーションアッセイを通常ベースで繰り返し得る。連続アッセイから得られた結果を用いて、数日から数ヶ月の期間にわたる治療の効果を示し得る。

【０２２０】

癌に関しては、個体からの生体組織における異常な量の転写物（過少発現または過剰発現）の存在は、疾患の発生素質を示したり、実際に臨床的症状が現れる前に疾患を検出する方法を提供したりし得る。この種のより明確な診断により、医療の専門家が予防方法または積極的な治療法を早くから利用し、それによって癌の発生または更なる進行を防止することが可能となる。

【０２２１】

ＡＰＲＧをコードする配列から設計されたオリゴヌクレオチドのさらなる診断への利用に

は、ＰＣＲの利用が含まれ得る。これらのオリゴマーは、化学的に合成するか、酵素により生産するか、あるいはｉｎ ｖｉｔｒｏで産出し得る。オリゴマーは、好ましくはＡＰＲＧをコードするポリヌクレオチドの断片、あるいはＡＰＲＧをコードするポリヌクレオチドと相補的なポリヌクレオチドの断片を含み、最適な条件下で、特定の遺伝子や条件を識別するために利用される。また、オリゴマーは、やや緩いストリンジェント条件下で、近縁のＤＮＡあるいはＲＮＡ配列の検出、定量、あるいはその両方のため用いることが可能である。

【０２２２】

或る実施態様において、ＡＰＲＧをコードするポリヌクレオチド配列由来のオリゴヌクレオチドプライマーを用いて、一塩基多型（ＳＮＰ）を検出し得る。ＳＮＰは、多くの場合にヒトの先天性または後天性遺伝病の原因となるような置換、挿入及び欠失である。限定するものではないがＳＮＰの検出方法には、ＳＳＣＰ（ｓｉｎｇｌｅ－ｓｔｒａｎｄｅｄ ｃｏｎｆｏｒｍａｔｉｏｎ ｐｏｌｙｍｏｒｐｈｉｓｍ）及び蛍光ＳＳＣＰ（ｆＳＳＣＰ）法がある。ＳＳＣＰでは、ＡＰＲＧをコードするポリヌクレオチド配列由来のオリゴヌクレオチドプライマーを用いたポリメラーゼ連鎖反応（ＰＣＲ）でＤＮＡを増幅する。ＤＮＡは例えば、病変組織または正常組織、生検サンプル、体液その他 に由来し得る。ＤＮＡ内のＳＮＰは、一本鎖形状のＰＣＲ生成物の２次及び３次構造に差異を生じさせる。差異は非変性ゲル中でのゲル電気泳動法を用いて検出可能である。ｆＳＣＣＰでは、オリゴヌクレオチドプライマーを蛍光性に標識する。それによってＤＮＡシークエンシング機などの高処理機器でアンプリマー（ａｍｐｌｉｍｅｒ）の検出が可能になる。更に、インシリコＳＮＰ（ｉｎ ｓｉｌｉｃｏ ＳＮＰ, ｉｓＳＮＰ）と呼ばれる配列データベース分析法は、一般的なコンセンサス配列に配列されるような個々の重畳するＤＮＡ断片の配列を比較することにより、多形性を同定し得る。これらのコンピュータベースの方法は、ＤＮＡの実験室での調整及び統計モデル及びＤＮＡ配列クロマトグラムの自動分析を用いたシークエンシングのエラーに起因する配列の変異をフィルタリングして除去する。別の態様では、例えば高処理ＭＡＳＳＡＲＲＡＹシステム（Ｓｅｑｕｅｎｏｍ, Ｉｎｃ．, Ｓａｎ Ｄｉｅｇｏ ＣＡ）を用いた質量分析によりＳＮＰを検出し、特徴付ける。

【０２２３】

ＡＰＲＧの発現を定量するために用い得る方法には、ヌクレオチドの放射標識またはビオチン標識、調節核酸の相互増幅（ｃｏａｍｐｌｉｆｉｃａｔｉｏｎ）及び標準曲線から得た結果の補間もある （例えば、Ｍｅｌｂｙ, Ｐ．Ｃ．ら（１９９３） Ｊ．Ｉｍｍｕｎｏｌ．Ｍｅｔｈｏｄｓ, １５９：２３５－２４４；Ｄｕｐｌａａ, Ｃ．ら（１９９３） Ａｎａｌ．Ｂｉｏｃｈｅｍ．２１２： ２２９－２３６を参照）。（１９９３） Ｊ．Ｉｍｍｕｎｏｌ．Ｍｅｔｈｏｄｓ １５９：２３５－２４４；Ｄｕｐｌａａ, Ｃ．他．（１９９３） Ａｎａｌ．Ｂｉｏｃｈｅｍ．２１２：２２９２３６．） 目的のオリゴマーが種々の希釈液中に存在し、分光光度法または比色反応によって定量が迅速になるような高処理フォーマットのアッセイを行うことによって、複数のサンプルの定量速度を加速することができる。

【０２２４】

更に別の実施例では、本明細書で記載した任意のポリヌクレオチド配列由来のオリゴヌクレオチドまたはより長い断片を、マイクロアレイにおけるエレメントとして用いることができる。多数の遺伝子の関連発現レベルを同時にモニターする転写イメージング技術にマイクロアレイを用いることが可能である。これについては、以下に記載する。マイクロアレイはまた、遺伝変異体、突然変異及び多形性の同定に用いることができる。この情報を用いることで、遺伝子機能を決定し、疾患の遺伝的根拠を理解し、疾患を診断し、遺伝子発現の機能としての疾病の進行／後退をモニターし、疾病治療における薬剤の活性を開発及びモニターすることができる。特に、患者にとって最もふさわしく、有効的な治療法を選択するために、この情報を用いて患者の薬理ゲノムプロフィールを開発することができる。例えば、患者の薬理ゲノムプロフィールに基づき、患者に対して高度に有効的で副作用を殆ど示さない治療薬を選択することができる。

【０２２５】
別の実施例では、ＡＰＲＧ、ＡＰＲＧの断片、ＡＰＲＧに特異的な抗体をマイクロアレイ上のエレメントとして用いることができる。マイクロアレイを用いて、上記のようなタンパク質－タンパク質相互作用、薬剤－標的相互作用及び遺伝子発現プロフィールをモニターまたは測定することが可能である。
【０２２６】
或る実施例は、或る組織または細胞タイプの転写イメージを生成するような本発明のポリヌクレオチドの使用に関連する。転写イメージは、特定の組織または細胞タイプにより遺伝子発現の包括的パターンを表す。包括的遺伝子発現パターンは、所与の条件下で所与の時間に発現した遺伝子の数及び相対存在量を定量することにより分析し得る（Ｓｅｉｌｌｉａｍｅｒらの米国特許第５，８４０，４８４号”Ｃｏｍｐａｒａｔｉｖｅ Ｇｅｎｅ Ｔｒａｎｓｃｒｉｐｔ Ａｎａｌｙｓｉｓ”を参照。該特許は特に引用することを以って本明細書の一部となす）。従って、特定の組織または細胞タイプの転写または逆転写全体に本発明のポリヌクレオチドまたはその補体をハイブリダイズすることにより、転写イメージを生成し得る。或る実施例では、本発明のポリヌクレオチドまたはその補体がマイクロアレイ上のエレメントのサブセットを複数含むような高処理フォーマットでハイブリダイゼーションを発生させる。結果として得られる転写イメージは、遺伝子活性のプロフィールを提供し得る。
【０２２７】
転写イメージは、組織、株化細胞、生検またはその生物学的サンプルから単離した転写物を用いて生成し得る。転写イメージは従って、組織または生検サンプルの場合には<u>ｉｎ ｖｉｖｏ</u>、または株化細胞の場合には<u>ｉｎ ｖｉｔｒｏ</u>での遺伝子発現を反映する。
【０２２８】
本発明のポリヌクレオチドの発現のプロフィールを作製する転写イメージはまた、工業的または天然の環境化合物の毒性試験のみならず、<u>ｉｎ ｖｉｔｒｏ</u>モデル系及び薬剤の前臨床評価に関連して使用し得る。全ての化合物は、作用及び毒性のメカニズムを暗示する、しばしば分子フィンガープリントまたは毒性サインと称されるような特徴的な遺伝子発現パターンを惹起する（Ｎｕｗａｙｓｉｒ，Ｅ．Ｆ．ら．（１９９９）Ｍｏｌ．Ｃａｒｃｉｎｏｇ．２４：１５３－１５９、Ｓｔｅｉｎｅｒ，Ｓ．ａｎｄ Ｎ．Ｌ．Ａｎｄｅｒｓｏｎ（２０００）Ｔｏｘｉｃｏｌ．Ｌｅｔｔ．１１２－１１３：４６７－４７１、該文献は特に引用することを以って本明細書の一部となす）。試験化合物が、既知の毒性を有する化合物のシグネチャと同一のシグネチャを有する場合には、毒性特性を共有している可能性がある。フィンガープリントまたはシグネチャは、多数の遺伝子及び遺伝子ファミリーからの発現情報を含んでいる場合には、最も有用且つ正確である。理想的には、発現のゲノム全域にわたる測定が最高品質のシグネチャを提供する。自己の発現が任意の試験された化合物により変化しない遺伝子が同様に重要であっても、このような遺伝子の発現レベルを用いて残りの発現データを規準化する。規準化手順は、異なる化合物で処理した後の発現データの比較に有用である。毒性シグネチャの要素に遺伝子機能を割り当てることが毒性メカニズムの解釈に役立つが、毒性の予測につながるシグネチャの統計的に一致させるのに遺伝子機能の知識は必要とされない（例えば２０００年２月２９日にＮａｔｉｏｎａｌ Ｉｎｓｔｉｔｕｔｅ ｏｆ Ｅｎｖｉｒｏｎｍｅｎｔａｌ Ｈｅａｌｔｈ Ｓｃｉｅｎｃｅｓより発行されたＰｒｅｓｓ Ｒｅｌｅａｓｅ ００－０２を参照されたい。これについてはｈｔｔｐ：／／ｗｗｗ．ｎｉｅｈｓ．ｎｉｈ．ｇｏｖ／ｏｃ／ｎｅｗｓ／ｔｏｘｃｈｉｐ．ｈｔｍで入手可能である）。従って、中毒学的スクリーニングの際に毒性シグネチャを用いて、全ての発現した遺伝子配列を含めることは重要且つ望ましいことである。
【０２２９】
或る実施例では、核酸を含有する生物学的サンプルを試験化合物で処理することにより試験化合物の毒性を算定する。処理した生物学的サンプル中で発現した核酸は、本発明のポリヌクレオチドに特異的な１つ若しくは複数のプローブでハイブリダイズし、それによっ

て本発明のポリヌクレオチドに対応する転写レベルを定量し得る。処理した生物学的サンプル中の転写レベルを、非処理生物学的サンプル中のレベルと比較する。両サンプルの転写レベルの差は、処理されたサンプル中で試験化合物が引き起こす毒性反応を示す。

【０２３０】

別の実施例は、本発明のポリペプチド配列を用いて組織または細胞タイプのプロテオームを分析することに関連する。プロテオームの語は、特定の組織または細胞タイプでのタンパク質発現の包括的パターンを指す。プロテオームの各タンパク質成分は、個々に更に分析の対象とすることができる。プロテオーム発現パターン即ちプロフィールは、所与の条件下で所与の時間に発現したタンパク質の数及び相対存在量を定量することにより分析し得る。従って細胞のプロテオームのプロフィールは、特定の組織または細胞タイプのポリペプチドを分離及び分析することにより作成し得る。或る実施例では、１次元等電点電気泳動によりサンプルからタンパク質を分離し、２次元ドデシル硫酸ナトリウムスラブゲル電気泳動により分子量に応じて分離するような２次元ゲル電気泳動により分離が達成される（前出のＳｔｅｉｎｅｒ および Ａｎｄｅｒｓｏｎ）。タンパク質は、通常クーマシーブルーまたはシルバーまたは蛍光染色などの物質を用いてゲルで染色することにより、分散した、独特な位置にある点としてゲル中で可視化される。各タンパク質スポットの光学密度は、通常サンプル中のタンパク質レベルに比例する。異なるサンプル、例えば試験化合物または治療薬で処理または非処理のいずれかの生物学的サンプルから得られるタンパク質スポットの光学密度を比較し、処理に関連するタンパク質スポット密度の変化を同定する。スポット内のタンパク質は、例えば化学的または酵素的切断とそれに続く質量分析を用いる標準的な方法を用いて部分的にシークエンシングする。スポット内のタンパク質の同一性は、その部分配列を、好適には少なくとも５個の連続するアミノ酸残基を、本発明のポリペプチド配列と比較することにより決定し得る。場合によっては、決定的なタンパク質同定のための更なる配列が得られる。

【０２３１】

プロテオームのプロファイルは、ＡＰＲＧに特異的な抗体を用いてＡＰＲＧ発現レベルを定量することによっても作成可能である。或る実施例では、マイクロアレイ上でエレメントとして抗体を用い、マイクロアレイをサンプルに曝して各アレイ要素へのタンパク質結合レベルを検出することによりタンパク質発現レベルを定量する（Ｌｕｅｋｉｎｇ，Ａ．他（１９９９）Ａｎａｌ．Ｂｉｏｃｈｅｒｎ．２７０：１０３－１１１、Ｍｅｎｄｏｚｅ，Ｌ．Ｇ．他．（１９９９）Ｂｉｏｔｅｃｈｎｉｑｕｅｓ ２７：７７８－７８８）。検出は当分野で既知の様々な方法で行うことができ、例えば、チオールまたはアミノ反応性蛍光化合物を用いてサンプル中のタンパク質を反応させ、各アレイのエレメントにおける蛍光結合の量を検出し得る。

【０２３２】

プロテオームレベルでの毒性サインも中毒学的スクリーニングに有用であり、転写レベルでの毒性サインと並行に分析するべきである。或る組織の或るタンパク質に対しては、転写とタンパク質の存在量の相関が乏しいこともあるので（Ａｎｄｅｒｓｏｎ，Ｎ．Ｌ．および Ｊ．Ｓｅｉｌｈａｍｅｒ（１９９７）Ｅｌｅｃｔｒｏｐｈｏｒｅｓｉｓ １８：５３３－５３７）、転写イメージにはそれ程影響しないがタンパク質のプロフィールを変化させるような化合物の分析においてプロテオーム毒性サインは有用たり得る。更に、体液中での転写の分析はｍＲＮＡ急速な分解により困難であるので、タンパク質のプロフィール作成はこのような場合により信頼でき、情報価値がある。

【０２３３】

別の実施例では、タンパク質を含有する生物学的サンプルを試験化合物で処理することにより試験化合物の毒性を算定する。処理された生物学的サンプル中で発現したタンパク質は、各タンパク質の量を定量し得るように分離する。各タンパク質の量を、非処理生物学的サンプル中の対応するタンパク質の量と比較する。両サンプルのタンパク質量の差は、処理サンプル中の試験化合物に対する反応を示す。個々のタンパク質は、個々のタンパク質のアミノ酸残基をシークエンシングし、これら部分配列を本発明のポリペプチドと比較

することにより同定する。

【０２３４】

別の実施例では、タンパク質を含有する生物学的サンプルを試験化合物で処理することにより試験化合物の毒性を算定する。生物学的サンプルから得たタンパク質は、本発明のポリペプチドに特異的な抗体を用いてインキュベートする。抗体により認識されたタンパク質の量を定量する。処理された生物学的サンプル中のタンパク質の量を、非処理生物学的サンプル中のタンパク質の量と比較する。両サンプルのタンパク質量の差は、処理サンプル中の試験化合物に対する反応を示す。

【０２３５】

マイクロアレイは、本技術分野でよく知られている方法を用いて調製し、使用し、そして分析する（Ｂｒｅｎｎａｎ，Ｔ．Ｍ．他．（１９９５）の米国特許第５，４７４，７９６号、Ｓｃｈｅｎａ，Ｍ．他．（１９９６）Ｐｒｏｃ．Ｎａｔｌ．Ａｃａｄ．Ｓｃｉ．ＵＳＡ９３：１０６１４－１０６１９、Ｂａｌｄｅｓｃｈｗｅｉｌｅｒらの（１９９５）ＰＣＴ出願第ＷＯ９５／２５１１１６号、Ｓｈａｌｏｎ，Ｄ．らの（１９９５）ＰＣＴ出願第ＷＯ９５／３５５０５号、Ｈｅｌｌｅｒ，Ｒ．Ａ．他．（１９９７）Ｐｒｏｃ．Ｎａｔｌ．Ａｃａｄ．Ｓｃｉ．ＵＳＡ９４：２１５０－２１５５、Ｈｅｌｌｅｒ，Ｍ．Ｊ．らの（１９９７）米国特許第５，６０５，６６２号等を参照）。様々なタイプのマイクロアレイが周知であり、詳細については、ＤＮＡ Ｍｉｃｒｏａｒｒａｙｓ：Ａ Ｐｒａｃｔｉｃａｌ Ａｐｐｒｏａｃｈ，Ｍ．Ｓｃｈｅｎａ，ｅｄ．（１９９９）Ｏｘｆｏｒｄ Ｕｎｉｖｅｒｓｉｔｙ Ｐｒｅｓｓ，Ｌｏｎｄｏｎに記載されている。該文献は、特別に引用することを以って本明細書の一部となす。

【０２３６】

本発明の別の実施例では、天然のゲノム配列をマッピングする際に有効なハイブリダイゼーションプローブを産出するため、ＡＰＲＧをコードする核酸配列を用いることが可能である。コード配列または非コード配列のいずれかを用いることができ、或る例では、コード配列全体で非コード配列が好ましい。例えば、多重遺伝子ファミリーのメンバー内でのコード配列の保存により、染色体マッピング中に望ましくないクロスハイブリダイゼーションが生じる可能性がある。核酸配列は、特定の染色体、染色体の特定領域または人工形成の染色体、例えば、ヒト人工染色体（ＨＡＣ）、酵母人工染色体（ＹＡＣ）、細菌人工染色体（ＢＡＣ）、細菌Ｐ１産物、あるいは単一染色体ｃＤＮＡライブラリに対してマッピングされる（例えば、Ｈａｒｒｉｎｇｔｏｎ，Ｊ．Ｊ．他．（１９９７）Ｎａｔ Ｇｅｎｅｔ．１５：３４５－３５５、Ｐｒｉｃｅ，Ｃ．Ｍ．（１９９３）Ｂｌｏｏｄ Ｒｅｖ．７：１２７－１３４、Ｔｒａｓｋ，Ｂ．Ｊ．（１９９１）Ｔｒｅｎｄｓ Ｇｅｎｅｔ．７：１４９－１５４等を参照）。一度マッピングすると、本発明の核酸配列を用いて例えば病状の遺伝を特定の染色体領域の遺伝または制限酵素断片長多型（ＲＦＬＰ）と相関させるような遺伝子連鎖地図を発生させ得る（例えば、Ｌａｎｄｅｒ，Ｅ．Ｓ．ａｎｄ Ｄ．Ｂｏｔｓｔｅｉｎ（１９８６）Ｐｒｏｃ．Ｎａｔｌ．Ａｃａｄ．Ｓｃｉ．ＵＳＡ８３：７３５３－７３５７を参照）。

【０２３７】

蛍光原位置ハイブリッド形成法（ＦＩＳＨ）は、他の物理的及び遺伝地図データと相関し得る（前出のＨｅｉｎｚ－Ｕｌｒｉｃｈ，他．（１９９５）ｉｎ Ｍｅｙｅｒｓ，９６５－９６８ページ．等を参照）。遺伝地図データの例は、種々の科学雑誌あるいはＯｎｌｉｎｅ Ｍｅｎｄｅｌｉａｎ Ｉｎｈｅｒｉｔａｎｃｅ ｉｎ Ｍａｎ（ＯＭＩＭ）のウェブサイトに見ることができる。物理的な染色体地図上のＡＰＲＧをコードする遺伝子の位置と特定の疾患との相関性、あるいは特定の疾患に対する素因が、このような疾患と関連するＤＮＡ領域の決定に役立つため、更なる位置を決定するクローニングが行われる。

【０２３８】

確定した染色体マーカーを用いた結合分析等の物理的マッピング技術及び染色体標本原位置ハイブリッド形成法を用いて、遺伝地図を拡張することができる。例えばマウスなど別

の哺乳動物の染色体上に遺伝子を配置することにより、正確な染色体の遺伝子座がわかっていない場合でも関連するマーカーを明らかにし得る。この情報は、位置クローニングその他 の遺伝子発見技術を用いて疾患遺伝子を探す研究者にとって価値がある。疾患または症候群が、血管拡張性失調症の１１ｑ２２－２３領域等、特定の遺伝子領域への遺伝的結合によって大まかに位置決めがなされると、該領域に対するいかなるマッピングも、更なる調査のための関連遺伝子あるいは調節遺伝子を表すことができる（Ｇａｔｔｉ，Ｒ．Ａ．他．（１９８８）Ｎａｔｕｒｅ ３３６：５７７－５８０）。転座、反転等に起因する、健常者、保有者、感染者の三者間における染色体位置の相違を発見するために本発明のヌクレオチド配列を用いてもよい。

【０２３９】

本発明の別の実施例では、ＡＰＲＧ、その触媒作用断片あるいは免疫原断片またはそのオリゴペプチドを、種々の任意の薬剤スクリーニング技術における化合物のライブラリのスクリーニングに用いることができる。薬剤スクリーニングに用いる断片は、溶液中に遊離しているか、固体支持物に固定されるか、細胞表面上に保持されるか、細胞内に位置することになろう。ＡＰＲＧと検査する薬剤との結合による複合体の形成を測定してもよい。

【０２４０】

別の薬剤スクリーニング方法は、目的のタンパク質に対して好適な結合親和性を有する化合物を高い処理能力でスクリーニングするために用いられる（Ｇｅｙｓｅｎ，他 （１９８４） ＰＣＴ出願番号 ＷＯ８４／０３５６４）。この方法においては、多数の異なる小さな試験用化合物を固体基質上で合成する。試験用化合物は、ＡＰＲＧ、あるいはその断片と反応してから洗浄される。次ぎに、結合されたＡＰＲＧが、当分野で周知の方法で検出される。精製されたＡＰＲＧはまた、前記した薬剤をスクリーニングする技術に用いられるプレート上で直接被覆することもできる。別法では、非中和抗体を用いてペプチドを捕捉し、ペプチドを固体支持物に固定することもできる。

【０２４１】

別の実施例では、ＡＰＲＧと結合可能な中和抗体がＡＰＲＧと結合するため試験用化合物と特に競合する、競合的薬剤スクリーニングアッセイを用いることができる。この方法では、抗体が、ＡＰＲＧと１つ以上の抗原決定因子を共有するどのペプチドの存在も検出する。

【０２４２】

別の実施例では、発展途上の分子生物学技術にＡＰＲＧをコードするヌクレオチド配列を用いて、限定はされないが、現在知られているトリプレット暗号及び特異的な塩基対相互作用などのヌクレオチド配列の特性に依存する新しい技術を提供することができる。

【０２４３】

更に詳細説明をしなくとも、当業者であれば以上の説明を以って本発明を最大限に利用できるであろう。従って、これ以下に記載する実施例は単なる例示目的にすぎず、いかようにも本発明を限定するものではない。

【０２４４】

前述した及び以下に記載する全ての特許出願、特許、刊行物、特に米国特許出願第６０／２５０，３２６号、および同第６０／２０９，４０７号に言及することをもって本明細書の一部とする。

【０２４５】

（実施例）

１　ｃＤＮＡライブラリの作製

Ｉｎｃｙｔｅ ｃＤＮＡは、ＬＩＦＥＳＥＱ ＧＯＬＤデータベース（Ｉｎｃｙｔｅ Ｇｅｎｏｍｉｃｓ，Ｐａｌｏ Ａｌｔｏ ＣＡ）に記載されたｃＤＮＡライブラリに由来するものであり、表４の列５に列記した。ホモジナイズしてグアニジニウムイソチオシアネート溶液に溶解した組織もあり、また、ホモジナイズしてフェノールまたは好適な変性剤の混合液に溶解した組織もある。変性剤の混合液は、例えばフェノールとグアニジニウムイソチオシアネートの単相溶液であるＴＲＩＺＯＬ（Ｌｉｆｅ Ｔｅｃｈｎｏｌｏｇｉｅｓ

）等である。結果として得られた溶解物は、塩化セシウムクッション上で遠心分離するかクロロホルムで抽出した。イソプロパノールか、酢酸ナトリウムとエタノールか、いずれか一方、あるいは別の方法を用いて、溶解物からＲＮＡを沈殿させた。

【０２４６】

ＲＮＡの純度を高めるため、ＲＮＡのフェノールによる抽出及び沈殿を必要な回数繰り返した。場合によっては、ＤＮアーゼでＲＮＡを処理した。殆どのライブラリでは、オリゴｄ（Ｔ）連結常磁性粒子（Ｐｒｏｍｅｇａ）、ＯＬＩＧＯＴＥＸラテックス粒子（ＱＩＡＧＥＮ，Ｖａｌｅｎｃｉａ ＣＡ）またはＯＬＩＧＯＴＥＸ ｍＲＮＡ精製キット（ＱＩＡＧＥＮ）を用いて、ポリ（Ａ＋） ＲＮＡを単離した。別法では、別のＲＮＡ単離キット、例えばＰＯＬＹ（Ａ）ＰＵＲＥ ｍＲＮＡ精製キット（Ａｍｂｉｏｎ，Ａｕｓｔｉｎ ＴＸ）を用いて組織溶解物からＲＮＡを直接単離した。

【０２４７】

場合によってはＳｔｒａｔａｇｅｎｅ社へのＲＮＡ提供を行い、対応するｃＤＮＡライブラリをＳｔｒａｔａｇｅｎｅ社が作製することもあった。そうでない場合は、ＵＮＩＺＡＰベクターシステム（Ｓｔｒａｔａｇｅｎｅ）またはＳＵＰＥＲＳＣＲＩＰＴプラスミドシステム（Ｌｉｆｅ Ｔｅｃｈｎｏｌｏｇｉｅｓ）を用いて本技術分野で公知の推奨方法または類似の方法でｃＤＮＡを合成し、ｃＤＮＡライブラリを作製した（前出のＡｕｓｕｂｅｌ，１９９７，ｕｎｉｔ ５．１－６．６等を参照）。逆転写は、オリゴｄ（Ｔ）またはランダムプライマーを用いて開始した。合成オリゴヌクレオチドアダプターを二本鎖ｃＤＮＡに連結反応させ、好適な制限酵素でｃＤＮＡを消化した。殆どのライブラリに対して、ｃＤＮＡのサイズ（３００～１０００ｂｐ）選択は、ＳＥＰＨＡＣＲＹＬ Ｓ１０００、ＳＥＰＨＡＲＯＳＥ ＣＬ２ＢまたはＳＥＰＨＡＲＯＳＥ ＣＬ４Ｂカラムクロマトグラフィー（Ａｍｅｒｓｈａｍ Ｐｈａｒｍａｃｉａ Ｂｉｏｔｅｃｈ）、あるいは調製用アガロースゲル電気泳動法を用いて行った。合成オリゴヌクレオチドアダプターを二本鎖ｃＤＮＡに連結反応させ、好適な制限酵素または酵素でｃＤＮＡを消化した。好適なプラスミドは、例えばＰＢＬＵＥＳＣＲＩＰＴプラスミド（Ｓｔｒａｔａｇｅｎｅ）、ｐＳＰＯＲＴ１プラスミド（Ｌｉｆｅ Ｔｅｃｈｎｏｌｏｇｉｅｓ）またはｐｌＮＣＹ（Ｉｎｃｙｔｅ Ｐｈａｒｍａｃｅｕｔｉｃａｌｓ，Ｐａｌｏ Ａｌｔｏ ＣＡ）等である。組換えプラスミドは、Ｓｔｒａｔａｇｅｎｅ社のＸＬ１－Ｂｌｕｅ、ＸＬ１－ＢＩｕｅＭＲＦまたはＳＯＬＲ、あるいはＬｉｆｅ Ｔｅｃｈｎｏｌｏｇｉｅｓ社のＤＨ５α、ＤＨ１０ＢまたはＥｌｅｃｔｒｏＭＡＸ ＤＨ１０Ｂを含むコンピテント大腸菌細胞に形質転換した。

【０２４８】

２　ｃＤＮＡクローンの単離

ＵＮＩＺＡＰベクターシステム（Ｓｔｒａｔａｇｅｎｅ）を用いたｉｎ ｖｉｖｏ切除によって、あるいは細胞溶解によって、実施例１のようにして得たプラスミドを宿主細胞から回収した。ＭａｇｉｃまたはＷＩＺＡＲＤ Ｍｉｎｉｐｒｅｐｓ ＤＮＡ精製システム（Ｐｒｏｍｅｇａ）、ＡＧＴＣ Ｍｉｎｉｐｒｅｐ精製キット（Ｅｄｇｅ Ｂｉｏｓｙｓｔｅｍｓ，Ｇａｉｔｈｅｒｓｂｕｒｇ ＭＤ）、ＱＩＡＧＥＮ社のＱＩＡＷＥＬＬ ８ Ｐｌａｓｍｉｄ、ＱＩＡＷＥＬＬ ８ Ｐｌｕｓ Ｐｌａｓｍｉｄ及びＱＩＡＷＥＬＬ ８ Ｕｌｔｒａ Ｐｌａｓｍｉｄ 精製システム、Ｒ．Ｅ．Ａ．Ｌ． Ｐｒｅｐ ９６プラスミドキットの中から少なくとも１つを用いて、プラスミドを精製した。沈殿させた後、０．１ｍｌの蒸留水に再懸濁して、凍結乾燥してあるいは凍結乾燥せずに、４℃で保管した。

【０２４９】

別法では、高処理フォーマットにおいて直接結合ＰＣＲ法を用いて宿主細胞溶解物からプラスミドＤＮＡを増幅した（Ｒａｏ，Ｖ．Ｂ．（１９９４）Ａｎａｌ．Ｂｉｏｃｈｅｍ．２１６：１－１４）。宿主細胞の溶解及び熱サイクリング過程は、単一反応混合液中で行った。サンプルを処理し、それを３８４穴プレート内で保管し、増幅したプラスミドＤＮＡの濃度をＰＩＣＯＧＲＥＥＮ色素（Ｍｏｌｅｃｕｌａｒ Ｐｒｏｂｅｓ，Ｅｕｇｅｎｅ ＯＲ）及びＦｌｕｏｒｏｓｋａｎ ＩＩ蛍光スキャナ（Ｌａｂｓｙｓｔｅｍｓ Ｏ

ｙ，Ｈｅｌｓｉｎｋｉ，Ｆｉｎｌａｎｄ）を用いて蛍光分析的に定量した。

【０２５０】

３　シークエンシング及び分析

実施例２に記載したようにプラスミドから回収したＩｎｃｙｔｅ ｃＤＮＡを、以下に示すようにシークエンシングした。ｃＤＮＡのシークエンス反応は、標準的方法あるいは高処理装置、例えばＡＢＩ ＣＡＴＡＬＹＳＴ ８００ サーマルサイクラー（Ａｐｐｌｉｅｄ Ｂｉｏｓｙｓｔｅｍｓ）またはＰＴＣ－２００ サーマルサイクラー（ＭＪ Ｒｅｓｅａｒｃｈ）をＨＹＤＲＡマイクロディスペンサー（Ｒｏｂｂｉｎｓ Ｓｃｉｅｎｔｉｆｉｃ）またはＭＩＣＲＯＬＡＢ ２２００（Ｈａｍｉｌｔｏｎ）液体転移システムと併用して処理した。ｃＤＮＡのシークエンス反応は、Ａｍｅｒｓｈａｍ Ｐｈａｒｍａｃｉａ Ｂｉｏｔｅｃｈ社が提供する試薬またはＡＢＩシークエンシングキット、例えばＡＢＩ ＰＲＩＳＭ ＢＩＧＤＹＥ Ｔｅｒｍｉｎａｔｏｒ ｃｙｃｌｅ ｓｅｑｕｅｎｃｉｎｇ ｒｅａｄｙ ｒｅａｃｔｉｏｎ ｋｉｔ（Ａｐｐｌｉｅｄ Ｂｉｏｓｙｓｔｅｍｓ）に与えられた試薬を用いて準備した。ｃＤＮＡのシークエンス反応の電気泳動的分離及び標識したポリヌクレオチドの検出には、ＭＥＧＡＢＡＣＥ １０００ ＤＮＡシークエンシングシステム（Ｍｏｌｅｃｕｌａｒ Ｄｙｎａｍｉｃｓ）か、標準ＡＢＩプロトコル及び塩基対呼び出しソフトウェアを用いるＡＢＩ ＰＲＩＳＭ ３７３または３７７シークエンシングシステム（Ａｐｐｌｉｅｄ Ｂｉｏｓｙｓｔｅｍｓ）か、あるいはその他 の本技術分野でよく知られている配列解析システムを用いた。ｃＤＮＡ配列内のリーディングフレームは、標準的方法（前出のＡｕｓｕｂｅｌ，１９９７，ｕｎｉｔ ７．７に概説）を用いて決定した。ｃＤＮＡ配列の幾つかを選択して、実施例８に記載した方法で配列を伸長させた。

【０２５１】

Ｉｎｃｙｔｅ ｃＤＮＡ配列に由来するポリヌクレオチド配列は、ベクター、リンカー及びポリ（Ａ）配列を除去し、あいまいな塩基対をマスクすることによって有効性を確認した。その際、ＢＬＡＳＴ、動的プログラミング及び隣接ジヌクレオチド頻度分析に基づくアルゴリズム及びプログラムを用いた。次に、ＢＬＡＳＴ、ＦＡＳＴＡ及びＢＬＩＭＰＳに基づくプログラムを用いて、プログラム中の注釈を得るべく、公共のデータベース、例えばＧｅｎＢａｎｋの霊長類及びげっ歯類、哺乳動物、脊椎動物、真核生物のデータベースと、ＢＬＯＣＫＳ、ＰＲＩＮＴＳ、ＤＯＭＯ、ＰＲＯＤＯＭ及びＰＦＡＭ等のＨｉｄｄｅｎ Ｍａｒｋｏｖ Ｍｏｄｅｌ（ＨＭＭ）ベースのタンパク質ファミリーデータベースの選択に対するＩｎｃｙｔｅ ｃＤＮＡ配列またはその翻訳を問い合わせた ＨＭＭは、遺伝子ファミリーのコンセンサス１次構造を解析する確率的アプローチである（Ｅｄｄｙ，Ｓ．Ｒ．（１９９６）Ｃｕｒｒ．Ｏｐｉｎ．Ｓｔｒｕｃｔ．Ｂｉｏｌ．６：３６１－３６５等を参照）。Ｅｄｄｙ，Ｓ．Ｒ．（１９９６）Ｃｕｆｆ．Ｏｐｉｎ．Ｓｔｒｕｃｔ．Ｂｉｏｌ．６：３６１－３６５等を参照）の選択に対するＩｎｃｙｔｅ ｃＤＮＡ配列またはその翻訳を問い合わせた。問合せは、ＢＬＡＳＴ、ＦＡＳＴＡ、ＢＬＩＭＰＳ及びＨＭＭＥＲに基づくプログラムを用いて行った。Ｉｎｃｙｔｅ ｃＤＮＡ配列は、完全長のポリヌクレオチド配列を産出するように構築した。あるいは、ＧｅｎＢａｎｋ ｃＤＮＡ、ＧｅｎＢａｎｋ ＥＳＴ、ステッチされた配列、ストレッチされた配列またはＧｅｎｓｃａｎ予測コード配列（実施例４及び５を参照）を用いてＩｎｃｙｔｅ ｃＤＮＡの集団を完全長まで伸長させた。Ｐｈｒｅｄ、Ｐｈｒａｐ及びＣｏｎｓｅｄに基づくプログラムを用いて構築し、ＧｅｎＭａｒｋ、ＢＬＡＳＴ及びＦＡＳＴＡに基づくプログラムを用いてｃＤＮＡの集団をオープンリーディングフレームに対してスクリーニングした。対応する完全長ポリペプチド配列を誘導するべく完全長ポリヌクレオチド配列を翻訳した。あるいは、本発明のポリペプチドは完全長翻訳ポリペプチドの任意のメチオニン残基で開始し得る。引き続いて、ＧｅｎＢａｎｋタンパク質データベース（ｇｅｎｐｅｐｔ）、ＳｗｉｓｓＰｒｏｔ、ＢＬＯＣＫＳ、ＰＲＩＮＴＳ、ＤＯＭＯ、ＰＲＯＤＯＭ及びＰｒｏｓｉｔｅ等のデータベース、ＰＦＡＭ等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に対する問合せによって完全長ポリペプチド配列を分析した。完全長ポリヌクレオチド配列はまた、ＭＡＣＤＮＡＳＩＳ ＰＲＯソフトウェア

（日立ソフトウェアエンジニアリング，Ｓｏｕｔｈ Ｓａｎ Ｆｒａｎｃｉｓｃｏ ＣＡ）及びＬＡＳＥＲＧＥＮＥソフトウェア（ＤＮＡＳＴＡＲ）を用いて分析した。ポリヌクレオチド及びポリペプチド配列アラインメントは、アラインメントした配列と配列の一致率も計算するＭＥＧＡＬＩＧＮマルチシークエンスアラインメントプログラム（ＤＮＡＳＴＡＲ）に組み込まれているようなＣＬＵＳＴＡＬアルゴリズムによって特定されるデフォルトパラメータを用いて生成する。

【０２５２】

Ｉｎｃｙｔｅ ｃＤＮＡ及び完全長配列の分析及びアセンブリに利用したツール、プログラム及びアルゴリズムの概略と、適用可能な説明、引用文献、閾値パラメータを表７に示す。用いたツール、プログラム及びアルゴリズムを表７の列１に、それらの簡単な説明を列２に示す。列３は好適な引用文献であり、全ての文献はそっくりそのまま引用を以って本明細書の一部となす。適用可能な場合には、列４は２つの配列が一致する強さを評価するために用いたスコア、確率値その他 のパラメータを示す（スコアが高ければ高いほど２配列間の相同性が高くなる）。

【０２５３】

完全長のポリヌクレオチド配列およびポリペプチド配列の組み立て及び分析に用いる上記のプログラムは、ＳＥＱ ＩＤ ＮＯ：４－６のポリヌクレオチド配列断片の同定にも利用できる。ハイブリダイゼーション及び増幅技術に有用である約２０～約４０００ヌクレオチドの断片を表４の列４に示した。

【０２５４】

４ ゲノムＤＮＡからのコード配列の同定及び編集

推定上のアポトーシス調節因子は、公共のゲノム配列データベース（例えば、ｇｂｐｒｉやｇｂｈｔｇ）においてＧｅｎｓｃａｎ遺伝子同定プログラムを実行して初めに同定された。Ｇｅｎｓｃａｎは、様々な生物からゲノムＤＮＡ配列を分析する汎用遺伝子同定プログラムである（Ｂｕｒｇｅ，Ｃ．およびＳ．Ｋａｒｌｉｎ（１９９７）Ｊ．Ｍｏｌ．Ｂｉｏｌ．２６８：７８－９４及びＢｕｒｇｅ，Ｃ．及びＳ．Ｋａｒｌｉｎ（１９９８）Ｃｕｆｆ．Ｏｐｉｎ．Ｓｔｒｕｃｔ．Ｂｉｏｌ．８：３４６－３５４参照）。プログラムは予測エキソンを連結し、メチオニンから停止コドンに及ぶ構築されたｃＤＮＡ配列を形成する。Ｇｅｎｓｃａｎの出力は、ポリヌクレオチド及びポリペプチド配列のＦＡＳＴＡデータベースである。Ｇｅｎｓｃａｎが一度に分析する配列の最大範囲は、３０ｋｂに設定した。これらのＧｅｎｓｃａｎ推定ｃＤＮＡ配列の内、どの配列がアポトーシス調節因子をコードするかを決定するために、コードされたポリペプチドをＰＦＡＭモデルにおいてアポトーシス調節因子について問合せて分析した。潜在的なアポトーシス調節因子が、アポトーシス調節因子としてアノテーションが付けられたインサイトｃＤＮＡ配列に対する相同性を基に同定された。こうして選択されたＧｅｎｓｃａｎ予測配列は、次にＢＬＡＳＴ分析により公共データベースｇｂｐｒｉ及びｇｂｈｔｇと比較した。必要であれば、ｇｅｎｐｅｐｔからヒットしたトップのＢＬＡＳＴと比較することによりＧｅｎｓｃａｎ予測配列を編集し、余分なまたは取り除かれたエキソンなどのＧｅｎｓｃａｎにより予測された配列のエラーを修正する。ＢＬＡＳＴ分析はまた、任意のＩｎｃｙｔｅ ｃＤＮＡまたはＧｅｎｓｃａｎ予測配列の公共ｃＤＮＡ適用範囲の発見に用いられるので、転写の証拠を提供する。Ｉｎｃｙｔｅ ｃＤＮＡ適用範囲が利用できる場合には、この情報を用いてＧｅｎｓｃａｎ予測配列を修正または確認した。完全長ポリヌクレオチド配列は、実施例３に記載された構築プロセスを用いて、Ｉｎｃｙｔｅ ｃＤＮＡ配列及び／または公共のｃＤＮＡ配列でＧｅｎｓｃａｎ予測コード配列を構築することにより得た。あるいは、完全長ポリヌクレオチド配列は編集または非編集のＧｅｎｓｃａｎ予測コード配列に完全に由来する。

【０２５５】

５ ゲノム配列データのｃＤＮＡ配列データへの構築

ステッチ配列（Ｓｔｉｃｈｅｄ Ｓｅｑｕｅｎｃｅ）

部分ｃＤＮＡ配列は、実施例４に記載のＧｅｎｓｃａｎ遺伝子同定プログラムにより予測

されたエキソンを用いて伸長させた。実施例３に記載されたように構築された部分ｃＤＮＡは、ゲノムＤＮＡにマッピングし、関連するｃＤＮＡ及び１つ若しくは複数のゲノム配列から予測されたＧｅｎｓｃａｎエキソンを含むクラスタに分解した。ｃＤＮＡ及びゲノム情報を統合するべくグラフ理論及び動的プログラミングに基づくアルゴリズムを用いて各クラスタを分析し、引き続いて確認、編集または伸長して完全長配列を産出するような潜在的スプライス変異体を生成した。間隔全体の長さがクラスタ中の２以上の配列に存在するような配列を同定し、そのように同定された間隔は推移により等しいと考えられた。例えば、１つのｃＤＮＡ及び２つのゲノム配列に間隔が存在する場合、３つの間隔は全て等しいと考えられる。このプロセスは、無関係であるが連続したゲノム配列をｃＤＮＡ配列により結び合わせて架橋し得る。このようにして同定された区間を、親配列（ｐａｒｅｎｔ　ｓｅｑｕｅｎｃｅ）に沿って現われるようにステッチアルゴリズムで縫い合わせ、可能な最も長い配列および変異配列を作製する。１種類の親配列に沿って発生した間隔と間隔との連鎖（ｃＤＮＡ－ｃＤＮＡまたはゲノム配列－ゲノム配列）は、親の種類を変える連鎖（ｃＤＮＡ－ゲノム配列）に優先した。結果として得られるステッチ配列は、ＢＬＡＳＴ分析により公共データベースｇｅｎｐｅｐｔ及びｇｂｐｒｉに翻訳されて比較された。Ｇｅｎｓｃａｎにより予測された不正確なエキソンは、ｇｅｎｐｅｐｔからヒットしたトップのＢＬＡＳＴと比較することにより修正した。必要な場合には、追加ｃＤＮＡ配列を用いるかゲノムＤＮＡの検査により配列を更に伸長させた。

【０２５６】

ストレッチ配列（Ｓｔｒｅｔｃｈｅｄ　Ｓｅｑｕｅｎｃｅ）

部分ＤＮＡ配列は、ＢＬＡＳＴ分析に基づくアルゴリズムにより完全長まで伸長された。先ず、ＢＬＡＳＴプログラムを用いて、ＧｅｎＢａｎｋの霊長類、げっ歯類、哺乳動物、脊椎動物及び真核生物のデータベースなどの公共データベースに対し、実施例３に記載されたように構築された部分ｃＤＮＡを問い合わせた。次に、最も近いＧｅｎＢａｎｋタンパク質相同体をＢＬＡＳＴ分析によりＩｎｃｙｔｅ　ｃＤＮＡ配列または実施例４に記載のＧｅｎＳｃａｎエキソン予測配列のいずれかと比較した。結果として得られる高スコアリングセグメント対（ＨＳＰ）を用いてキメラタンパク質を産出し、翻訳した配列をＧｅｎＢａｎｋタンパク質相同体上にマッピングした。元のＧｅｎＢａｎｋタンパク質相同体に関連して、キメラタンパク質内で挿入または削除が起こり得る。ＧｅｎＢａｎｋタンパク質相同体、キメラタンパク質またはその両方をプローブとして用い、公共のヒトゲノムデータベースから相同ゲノム配列を検索した。このようにして、部分的なＤＮＡ配列を相同ゲノム配列の付加によりストレッチすなわち伸長した。結果として得られるストレッチ配列を検査し、完全遺伝子を含んでいるか否かを決定した。

【０２５７】

６　ＡＰＲＧをコードするポリヌクレオチドの染色体マッピング

ＳＥＱ　ＩＤ　ＮＯ：４－６を構築するために用いた配列を、ＢＬＡＳＴ及びＳｍｉｔｈ－Ｗａｔｅｒｍａｎアルゴリズムを用いて、Ｉｎｃｙｔｅ　ＬＩＦＥＳＥＱデータベース及び公共のドメインデータベースの配列と比較した。ＳＥＱ　ＩＤ　ＮＯ：４－６と一致するこれらのデータベースの配列を、Ｐｈｒａｐ（表７）などの構築アルゴリズムを使用して、連続及び重複した配列のクラスターに組み入れた。スタンフォード・ヒトゲノムセンター（ＳＨＧＣ）、ホワイトヘッド・ゲノム研究所（ＷＩＧＲ）、Ｇｅｎｅｔｈｏｎ等の公的な情報源から入手可能な放射線ハイブリッド及び遺伝地図データを用いて、クラスタ化された配列が前もってマッピングされたかを測定した。マッピングされた配列がクラスタに含まれている結果、個々の配列番号を含めてそのクラスタの全配列が地図上の位置に割り当てられた。

【０２５８】

地図上の位置は、ヒト染色体の範囲または間隔として表される。センチモルガン間隔の地図上の位置は、染色体のｐアームの末端に関連して測定する。（センチモルガン（ｃＭ）は、染色体マーカー間の組換え頻度に基づく計測単位である。平均して、１ｃＭは、ヒト中のＤＮＡの１メガベース（Ｍｂ）にほぼ等しい。尤も、この値は、組換えのホットスポ

ット及びコールドスポットに起因して広範囲に変化する。ｃＭ距離は、配列が各クラスタ内に含まれるような放射線ハイブリッドマーカーに対して境界を提供するようなＧｅｎｅｔｈｏｎによってマッピングされた遺伝マーカーに基づく。ＮＣＢＩ「ＧｅｎｅＭａｐ９９」（ｈｔｔｐ：／／ｗｗｗ．ｎｃｂｉ．ｎｌｍ．ｎｉｈ．ｇｐｖ／ｇｅｎｅｍａｐ）などの一般個人が入手可能なヒト遺伝子マップおよびその他 の情報源を用いて、上記した区間が既に同定されている疾患遺伝子マップ内若しくは近傍に位置するかを決定できる。

【０２５９】

７　ポリヌクレオチド発現の分析

ノーザン分析は、転写された遺伝情報の存在を検出するために用いられる実験技術であり、特定の細胞種または組織からのＲＮＡが結合される膜への標識されたヌクレオチド配列のハイブリダイゼーションに関与している。（前出のＳａｍｂｒｏｏｋ，７章、同Ａｕｓｕｂｅｌ．Ｆ．Ｍ．ら，４章及び１６章等を参照。）

ＢＬＡＳＴに適用する類似のコンピュータ技術を用いて、ＧｅｎＢａｎｋやＬｉｆｅＳｅｑ（Ｉｎｃｙｔｅ Ｐｈａｒｍａｃｅｕｔｉｃａｌｓ）等のヌクレオチドデータベースにおいて同一または関連分子を検索する。ノーザン分析は、多数膜系ハイブリダイゼーションよりも非常に速い。更に、特定の同一を厳密なあるいは相同的なものとして分類するか否かを決定するため、コンピュータ検索の感度を変更することができる。検索の基準は積スコアであり、次式で定義される。

【数１】

$$\frac{(\text{BLAST スコア} \times \text{配列一致率})}{5 \times (\text{長さ(配列1)，長さ(配列2)})\text{の最小値}}$$

積スコアは、２つの配列間の類似度及び配列が一致する長さの両方を考慮している。積スコアは、０～１００の規準化された値であり、次のようにして求める。ＢＬＡＳＴスコアにヌクレオチドの配列一致率を乗じ、その積を２つの配列の短い方の長さの５倍で除する。高スコアリングセグメント対（ＨＳＰ）に一致する各塩基に＋５のスコアを割り当て、各不適性塩基対に－４を割り当てることにより、ＢＬＡＳＴスコアを計算する。２つの配列は、２以上のＨＳＰを共有し得る（ギャップにより隔離され得る）。２以上のＨＳＰがある場合には、最高ＢＬＡＳＴスコアの塩基対を用いて積スコアを計算する。積スコアは、断片的重畳とＢＬＡＳＴアラインメントの質とのバランスを表す。例えば積スコア１００は、比較した２つの配列の短い方の長さ全体にわたって１００％一致する場合のみ得られる。積スコア７０は、一端が１００％一致し、７０％重畳しているか、他 端が８８％一致し、１００％重畳しているかのいずれかの場合に得られる。積スコア５０は、一端が１００％一致し、５０％重畳しているか、他 端が７９％一致し、１００％重畳しているかのいずれかの場合に得られる。

【０２６０】

あるいは、ＡＰＲＧをコードするポリヌクレオチド配列は、由来する組織に対して分析する。例えば或る完全長配列は、Ｉｎｃｙｔｅ ｃＤＮＡ配列（実施例３を参照）と少なくとも一部は重畳するように構築される。各ｃＤＮＡ配列は、ヒト組織から作製されたｃＤＮＡライブラリに由来する。各ｃＤＮＡ配列は、ヒト組織から作製されたｃＤＮＡライブラリに由来する。各ヒト組織は、以下の生物／組織カテゴリー即ち心血管系、結合組織、消化器系、胎芽構造、内分泌系、外分泌腺、女性生殖器、男性生殖器、生殖細胞、血液及び免疫系、肝、筋骨格系、神経系、膵臓、呼吸器系、感覚器、皮膚、顎口腔系、非分類性／混合性または尿路の１つに分類される。各カテゴリーのライブラリ数を数えて、全カテゴリーの総ライブラリ数で除する。同様に、各ヒト組織は、以下の疾患／病状カテゴリー即ち 癌、細胞系、発達、炎症、神経性、外傷、心血管、鬱血、その他 の１つに分類される。各カテゴリーのライブラリ数を数えて、全カテゴリーの総ライブラリ数で除する。得られるパーセンテージは、ＡＰＲＧをコードするｃＤＮＡの疾患特異的な発現を反映する。ｃＤＮＡ配列およびｃＤＮＡライブラリ／組織の情報は、ＬＩＦＥＳＥＱ ＧＯＬＤ デ

ータベース（Ｉｎｃｙｔｅ Ｇｅｎｏｍｉｃｓ, Ｐａｌｏ Ａｌｔｏ ＣＡ）から得ることができる。

【０２６１】

８　ＡＰＲＧをコードするポリヌクレオチドの伸長

完全長のポリヌクレオチド配列もまた、完全長分子の適切な断片から設計したオリゴヌクレオチドプライマーを用いて該断片を伸長させて生成した。一方のプライマーは既知の断片の５’伸長を開始するべく合成し、他 方のプライマーは既知の断片の３’伸長を開始するべく合成した。開始プライマーの設計は、長さが約２２～３０ヌクレオチド、ＧＣ含有率が５０％以上となり、約６８～７２℃の温度で標的配列にアニーリングするように、ＯＬＩＧＯ ４．０６ソフトウェア（Ｎａｔｉｏｎａｌ Ｂｉｏｓｃｉｅｎｃｅｓ）あるいは別の適切なプログラムを用いて、ｃＤＮＡから設計した。ヘアピン構造及びプライマー－プライマー二量体を生ずるようなヌクレオチドの区間は全て回避した。

【０２６２】

配列を伸長するために、選択されたヒトｃＤＮＡライブラリを用いた。２段階以上の伸長が必要または望ましい場合には、付加的プライマーあるいはプライマーのネステッドセットを設計した。

【０２６３】

高忠実度の増幅が、当業者によく知られている方法を利用したＰＣＲ法によって得られた。ＰＣＲは、ＰＴＣ－２００ サーマルサイクラー（ＭＪ Ｒｅｓｅａｒｃｈ, Ｉｎｃ.）を用いて９６穴プレート内で行った。反応混合液は、鋳型ＤＮＡ及び２００ ｎｍｏｌの各プライマー、Mg^{2+}と$(NH_4)_2SO_4$とβ－メルカプトエタノールを含むバッファー、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）を含む。プライマーの組、ＰＣＩ ＡとＰＣＩ Ｂに対して以下のパラメータで増幅を行った。ステップ １： ９４℃で３分間、ステップ ２： ９４℃で１５分間、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。ステップ ６： ６８℃で５分間、ステップ ７： ４℃で保存。別法では、プライマー対、Ｔ７とＳＫ＋に対して以下のパラメータで増幅を行った。ステップ １： ９４℃で３分間、ステップ ２： ９４℃で１５分間、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、ステップ ７： ４℃で保存。

【０２６４】

各ウェルのＤＮＡ濃度は、１Ｘ ＴＥ及び０．５μｌの希釈していないＰＣＲ産物に溶解した１００μｌのＰＩＣＯＧＲＥＥＮ定量試薬（０．２５（ｖ／ｖ） ＰＩＣＯＧＲＥＥＮ；Ｍｏｌｅｃｕｌａｒ Ｐｒｏｂｅｓ, Ｅｕｇｅｎｅ ＯＲ）を不透明な蛍光光度計プレート（Ｃｏｍｉｎｇ Ｃｏｓｔａｒ, Ａｃｔｏｎ ＭＡ）の各ウェルに分配してＤＮＡが試薬と結合できるようにして測定する。サンプルの蛍光を計測してＤＮＡの濃度を定量するべくプレートをＦｌｕｏｒｏｓｋａｎ ＩＩ（Ｌａｂｓｙｓｔｅｍｓ Ｏｙ, Ｈｅｌｓｉｎｋｉ, Ｆｉｎｌａｎｄ）でスキャンした。反応混合物のアリコート５～１０μｌを１％アガロースミニゲル上で電気泳動法によって解析し、どの反応が配列の伸長に成功したかを決定した。

【０２６５】

伸長させたヌクレオチドは、脱塩及び濃縮して３８４穴プレートに移し、ＣｖｉＪＩコレラウイルスエンドヌクレアーゼ（Ｍｏｌｅｃｕｌａｒ Ｂｉｏｌｏｇｙ Ｒｅｓｅａｒｃｈ, Ｍａｄｉｓｏｎ ＷＩ）を用いて消化し、ｐＵＣ １８ベクター（Ａｍｅｒｓｈａｍ Ｐｈａｒｍａｃｉａ Ｂｉｏｔｅｃｈ）への再連結反応前に音波処理またはせん断した。ショットガン・シークエンシングのために、消化したヌクレオチドを低濃度（０．６～０．８％）のアガロースゲル上で分離し、断片を切除し、寒天をＡｇａｒ ＡＣＥ（Ｐｒｏｍｅｇａ）で消化した。伸長させたクローンをＴ４リガーゼ（Ｎｅｗ Ｅｎｇｌａｎｄ Ｂｉ

ｏｌａｂｓ，Ｂｅｖｅｒｌｙ ＭＡ）を用いてｐＵＣ １８ベクター（Ａｍｅｒｓｈａｍ Ｐｈａｒｍａｃｉａ Ｂｉｏｔｅｃｈ）に再連結し、Ｐｆｕ ＤＮＡポリメラーゼ（Ｓｔｒａｔａｇｅｎｅ）で処理して制限部位の張出部（ｏｖｅｒｈａｎｇ）を満たし、大腸菌細胞に形質移入した。形質移入した細胞を選択して抗生物質を含む培地に移し、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３８４ウェルプレートに３７℃で一晩培養した。

【０２６６】

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ＰＣＲ増幅した。ステップ １： ９４℃で３分間、ステップ ２： ９４℃で１５分間、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２０回繰り返す。ステップ ６： ７２℃で５分間、ステップ ７： ４℃で保存。上記したようにＰＩＣＯＧＲＥＥＮ試薬（Ｍｏｌｅｃｕｌａｒ Ｐｒｏｂｅｓ）でＤＮＡを定量化した。ＤＮＡの回収率が低いサンプルは、上記と同一の条件を用いて再増幅した。サンプルは２０％ジメチルスルホキシド（１：２，ｖ／ｖ）で希釈し、ＤＹＥＮＡＭＩＣ エネルギートランスファー シークエンシングプライマー、及びＤＹＥＮＡＭＩＣ ＤＩＲＥＣＴ ｋｉｔ（Ａｍｅｒｓｈａｍ Ｐｈａｒｍａｃｉａ Ｂｉｏｔｅｃｈ）またはＡＢＩ ＰＲＩＳＭ ＢＩＧＤＹＥ ターミネーターサイクルシークエンシング反応キット（Ｔｅｒｍｉｎａｔｏｒ ｃｙｃｌｅ ｓｅｑｕｅｎｃｉｎｇ ｒｅａｄｙ ｒｅａｃｔｉｏｎ ｋｉｔ）（Ａｐｐｌｉｅｄ Ｂｉｏｓｙｓｔｅｍｓ）を用いてシークエンシングした。

【０２６７】

同様に、上記手順を用いて完全長ヌクレオチド配列を検証し、あるいはそのような伸長のために設計されたオリゴヌクレオチド及び適切なゲノムライブラリを用いて５’調節配列を得る。

【０２６８】

９　個々のハイブリダイゼーションプローブの標識及び使用

ＳＥＱ ＩＤ ＮＯ：４－６から導き出されたハイブリダイゼーションプローブを用いて、ｃＤＮＡ、ｍＲＮＡ、またはゲノムＤＮＡをスクリーニングする。約２０塩基対からなるオリゴヌクレオチドの標識について特に記載するが、より大きなヌクレオチド断片に対しても事実上同一の手順が用いられる。オリゴヌクレオチドは、ＯＬＩＧＯ ４．０６ソフトウェア（Ｎａｔｉｏｎａｌ Ｂｉｏｓｃｉｅｎｃｅｓ）等の最新ソフトウェアを用いて設計し、各オリゴマー５０ｐｍｏｌと、［γ－^{32}Ｐ］アデノシン３リン酸（Ａｍｅｒｓｈａｍ Ｐｈａｒｍａｃｉａ Ｂｉｏｔｅｃｈ）２５０μＣｉと、Ｔ４ポリヌクレオチドキナーゼ（ＤｕＰｏｎｔ ＮＥＮ，Ｂｏｓｔｏｎ ＭＡ）を混合することにより標識する。標識したオリゴヌクレオチドは、ＳＥＰＨＡＤＥＸ Ｇ－２５超細繊分子サイズ排除デキストラン ビードカラム（Ａｍｅｒｓｈａｍ Ｐｈａｒｍａｃｉａ Ｂｉｏｔｅｃｈ）を用いて十分に精製する。Ａｓｅ Ｉ、Ｂｇｌ ＩＩ、Ｅｃｏ ＲＩ、Ｐｓｔ Ｉ、Ｘｂａ１またはＰｖｕ ＩＩ（ＤｕＰｏｎｔ ＮＥＮ）のいずれか１つのエンドヌクレアーゼで消化されたヒトゲノムＤＮＡの典型的な膜ベースのハイブリダイゼーション解析において、毎分10^{7}カウントの標識されたプローブを含むアリコットを用いる。Ａｓｅ Ｉ，Ｂｇｌ ＩＩ，Ｅｃｏ ＲＩ，Ｐｓｔ Ｉ，Ｘｂａ Ｉ，ｏｒ Ｐｖｕ ＩＩ（ＤｕＰｏｎｔ ＮＥＮ）．

各消化物から得たＤＮＡは、０．７％アガロースゲル上で分画してナイロン膜（Ｎｙｔｒａｎ Ｐｌｕｓ，Ｓｃｈｌｅｉｃｈｅｒ ＆ Ｓｃｈｕｅｌｌ，Ｄｕｒｈａｍ ＮＨ）に移す。ハイブリダイゼーションは、４０℃で１６時間行う。非特異的シグナルを除去するため、例えば０．１×クエン酸ナトリウム食塩水及び０．５％ドデシル硫酸ナトリウムに一致する条件下で、ブロットを室温で順次洗浄する。オートラジオグラフィーまたはそれに代わるイメージング手段を用いてハイブリダイゼーションパターンを視覚化し、比較する。

【０２６９】
１０　マイクロアレイ
マイクロアレイの表面上でアレイエレメントの連鎖または合成は、フォトリソグラフィ、圧電印刷（インクジェット印刷、前出のＢａｌｄｅｓｃｈｗｅｉｌｅｒ等を参照）、機械的マイクロスポッティング技術及びこれらから派生したものを用いて達成することが可能である。上記各技術において基質は、均一且つ非多孔性の固体とするべきである（Ｓｃｈｅｎａ（１９９９）前出）。推奨する基質には、シリコン、シリカ、スライドガラス、ガラスチップ及びシリコンウエハがある。あるいは、ドットブロット法またはスロットブロット法に類似のアレイを利用して、熱的、紫外線的、化学的または機械的結合手順を用いて基質の表面にエレメントを配置及び結合させてもよい。通常のアレイは、手作業で、または利用可能な方法や機械を用いて作製でき、任意の適正数のエレメントを有し得る（Ｓｃｈｅｎａ，Ｍ．他．（１９９５）Ｓｃｉｅｎｃｅ２７０：４６７－４７０、Ｓｈａｌｏｎ．Ｄ．他．（１９９６）Ｇｅｎｏｍｅ Ｒｅｓ．６：６３９－６４５、Ｍａｒｓｈａｌｌ，Ａ．ａｎｄ Ｊ．Ｈｏｄｇｓｏｎ（１９９８）Ｎａｔ．Ｂｉｏｔｅｃｈｎｏｌ．１６：２７－３１．を参照）。
【０２７０】
完全長ｃＤＮＡ、発現配列タグ（ＥＳＴ）、またはその断片またはオリゴマーは、マイクロアレイのエレメントと成り得る。ハイブリダイゼーションに好適な断片またはオリゴマーを、レーザＧＥＮＥソフトウェア（ＤＮＡＳＴＡＲ）等の本技術分野で公知のソフトウェアを用いて選択することが可能である。アレイエレメントは、生物学的サンプル中でポリヌクレオチドを用いてハイブリダイズされる。生物学的サンプル中のポリヌクレオチドは、検出を容易にするために蛍光標識またはその他 の分子タグに抱合される。ハイブリダイゼーション後、生物学的サンプルからハイブリダイズされていないヌクレオチドを除去し、蛍光スキャナを用いて各アレイエレメントにおいてハイブリダイゼーションを検出する。あるいは、レーザ脱着及び質量スペクトロメトリを用いてもハイブリダイゼーションを検出し得る。マイクロアレイ上のエレメントにハイブリダイズする各ポリヌクレオチドの相補性の度合及び相対存在度は、算定し得る。一実施例におけるマイクロアレイの調整及び使用について、以下に詳述する。
【０２７１】
組織または細胞サンプルの調製
グアニジウムチオシアネート法を用いて組織サンプルから全ＲＮＡを単離し、オリゴ（ｄＴ）セルロース法を用いてポリ（Ａ）$^{+}$ＲＮＡを精製する。各ポリ（Ａ）$^{+}$ＲＮＡサンプルを、ＭＭＬＶ逆転写酵素、０．０５ｐｇ／μｌのオリゴ（ｄＴ）プライマー（２１ｍｅｒ）、１×第一鎖合成バッファー、０．０３ｕｎｉｔ／μｌのＲＮアーゼ阻害因子、５００μＭのｄＡＴＰ、５００μＭのｄＧＴＰ、５００μＭのｄＴＴＰ、４０μＭのｄＣＴＰ、４０μＭのｄＣＴＰ－Ｃｙ３（ＢＤＳ）またはｄＣＴＰ－Ｃｙ５（Ａｍｅｒｓｈａｍ Ｐｈａｒｍａｃｉａ Ｂｉｏｔｅｃｈ）を用いて逆転写する。逆転写反応は、ＧＥＭＢＲＩＧＨＴキット（Ｉｎｃｙｔｅ）を用いてポリ（Ａ）$^{+}$ＲＮＡ含有の２５体積ｍｌ内で行う。特異的対照ポリ（Ａ）$^{+}$ＲＮＡは、非コード酵母ゲノムＤＮＡからｉｎ ｖｉｔｒｏ転写により合成する。各反応サンプル（１つはＣｙ３、もう１つはＣｙ５標識）は、２．５ｍｌの０．５Ｍ水酸化ナトリウムで処理し、８５０℃で２０分間インキュベートし、反応を停止させてＲＮＡを分解させる。サンプルは、２つの連続するＣＨＲＯＭＡ ＳＰＩＮ ３０ゲル濾過スピンカラム（ＣＬＯＮＴＥＣＨ Ｌａｂｏｒａｔｏｒｉｅｓ，Ｉｎｃ．（ＣＬＯＮＴＥＣＨ），Ｐａｌｏ Ａｌｔｏ ＣＡ）を用いて精製する。結合後、２つの反応サンプルは、１ｍｌのグリコーゲン（１ｍｇ／ｍｌ）を用いて析出させたエタノール、６０ｍｌの酢酸ナトリウム及び３００ｍｌの１００％エタノールである。サンプルは次に、ＳｐｅｅｄＶＡＣ（Ｓａｖａｎｔ Ｉｎｓｔｒｕｍｅｎｔｓ Ｉｎｃ．，Ｈｏｌｂｒｏｏｋ ＮＹ）を用いて乾燥して仕上げ、１４μｌの５×ＳＳＣ／０．２％ＳＤＳ中で再懸濁する。
【０２７２】

マイクロアレイの調製

本発明の配列を用いて、アレイエレメントを生成する。各アレイエレメントは、クローン化ｃＤＮＡインサートによりベクター含有細菌性細胞から増幅する。ＰＣＲ増幅は、ｃＤＮＡインサートの側面に位置するベクター配列に相補的なプライマーを用いる。３０サイクルのＰＣＲで１～２ｎｇの初期量から５μｇより大きい最終量までアレイエレメントを増幅する。増幅されたアレイエレメントは、ＳＥＰＨＡＣＲＹＬ－４００（Ａｍｅｒｓｈａｍ Ｐｈａｒｍａｃｉａ Ｂｉｏｔｅｃｈ）を用いて精製される。

【０２７３】

精製したアレイエレメントは、ポリマーコートされたスライドグラス上に固定する。顕微鏡スライドグラス（Ｃｏｒｎｉｎｇ）は、処理中及び処理後に０．１％のＳＤＳ及びアセトン中で超音波をかけ、蒸留水で非常に良く洗って洗浄する。スライドグラスは、４％フッ化水素酸（ＶＷＲ Ｓｃｉｅｎｔｉｆｉｃ Ｐｒｏｄｕｃｔｓ Ｃｏｒｐｏｒａｔｉｏｎ（ＶＷＲ），Ｗｅｓｔ Ｃｈｅｓｔｅｒ ＰＡ）中でエッチングし、蒸留水中で広範囲にわたって洗浄し、９５％エタノール中で０．０５％アミノプロピルシラン（Ｓｉｇｍａ）を用いてコーティングする。コーティングしたスライドガラスは、１１０℃のオブンで硬化させる。

【０２７４】

米国特許第５，８０７，５２２号に記載されている方法を用いて、コーティングしたガラス基板にアレイエレメントを付加する。この特許に引用することを以って本明細書の一部とする。平均濃度が１００ｎｇ／μｌのアレイエレメントＤＮＡ１μｌを高速ロボット装置により開口キャピラリープリントエレメントに充填する。装置はここで、スライド毎に約５ｎｌのアレイエレメントサンプルを加える。

【０２７５】

マイクロアレイには、ＳＴＲＡＴＡＬＩＮＫＥＲ ＵＶ架橋剤（Ｓｔｒａｔａｇｅｎｅ）を用いてＵＶ架橋する。マイクロアレイは、室温において０．２％ＳＤＳで１度洗浄し、蒸留水で３度洗浄する。リン酸緩衝生理食塩水（ＰＢＳ）（Ｔｒｏｐｉｘ，Ｉｎｃ．，Ｂｅｄｆｏｒｄ ＭＡ）中の０．２％カゼイン中において６０℃で３０分間マイクロアレイをインキュベートした後、前に行ったように０．２％ＳＤＳ及び蒸留水で洗浄することにより、非特異結合部位をブロックする。

【０２７６】

ハイブリダイゼーション

ハイブリダイゼーション反応は、５×ＳＳＣ，０．２％ＳＤＳハイブリダイゼーション緩衝液中のＣｙ３及びＣｙ５標識したｃＤＮＡ合成生成物を各０．２μｇ含む９μｌのサンプル混合体を有する。サンプル混合体は、６５℃まで５分間加熱し、マイクロアレイ表面上で等分して１．８cm^2 のカバーガラスで覆う。アレイは、顕微鏡スライドより僅かに大きい空洞を有する防水チェンバーに移す。チェンバーのコーナーに１４０μｌの５×ＳＳＣを加えることにより、チェンバー内部を湿度１００％に保持する。アレイを含むチェンバーは、６０℃で約６．５時間インキュベートする。アレイは、第１洗浄緩衝液中（１×ＳＳＣ，０．１％ＳＤＳ）において４５℃で１０分間洗浄し、第２洗浄緩衝液中（０．１×ＳＳＣ）において４５℃で１０分間各々３度洗浄して乾燥させる。

【０２７７】

検出

レポーター標識ハイブリダイゼーション複合体は、Ｃｙ３の励起のためには４８８ｎｍ、Ｃｙ５の励起のためには６３２ｎｍでスペクトル線を発生し得るＩｎｎｏｖａ ７０混合ガス１０ Ｗレーザ（Ｃｏｈｅｒｅｎｔ，Ｉｎｃ．，Ｓａｎｔａ Ｃｌａｒａ ＣＡ）を備えた顕微鏡で検出する。２０×顕微鏡対物レンズ（Ｎｉｋｏｎ，Ｉｎｃ．，Ｍｅｌｖｉｌｌｅ ＮＹ）を用いて、アレイ上に励起レーザ光の焦点を当てる。アレイを含むスライドを顕微鏡のコンピュータ制御のＸ－Ｙステージに置き、対物レンズを通過してラスタースキャンする。本実施例で用いた１．８ｃｍ×１．８ｃｍのアレイは、２０μｍの解像度でスキャンした。

【０２７８】
２つの異なるスキャンのうち、混合ガスマルチラインレーザは２つの蛍光色素を連続的に励起する。発光された光は、波長に基づき分離され、２つの蛍光色素に対応する２つの光電子増倍管検出器（ＰＭＴ Ｒ１４７７，Ｈａｍａｍａｔｓｕ Ｐｈｏｔｏｎｉｃｓ Ｓｙｓｔｅｍｓ，Ｂｒｉｄｇｅｗａｔｅｒ ＮＪ）に送られる。アレイと光電子増倍管間に設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ーする。用いる蛍光色素の最大発光の波長は、Ｃｙ３では５６５ｎｍ、Ｃｙ５では６５０ｎｍである。装置は両方の蛍光色素からのスペクトルを同時に記録し得るが、レーザ源において好適なフィルターを用いて、蛍光色素１つにつき１度スキャンし、各アレイを通常２度スキャンする。
【０２７９】
スキャンの感度は通常、既知濃度のサンプル混合体に添加されるｃＤＮＡ対照種により生成されるシグナル強度を用いて較正する。アレイ上の特定の位置には相補的ＤＮＡ配列が含まれ、その位置におけるシグナルの強度をハイブリダイジング種の重量比１：１００，０００に相関させる。異なる源泉（例えば試験される細胞及び対照細胞など）からの２つのサンプルを、各々異なる蛍光色素で標識し、他と異なって発現した遺伝子を同定するために単一のアレイにハイブリダイズする場合には、その較正を、２つの蛍光色素で較正するｃＤＮＡのサンプルを標識し、ハイブリダイゼーション混合体に各々等量を加えることによって行う。
【０２８０】
光電子増倍管の出力は、ＩＢＭコンパチブルＰＣコンピュータにインストールされた１２ビットＲＴＩ－８３５Ｈアナログ－ディジタル（Ａ／Ｄ）変換ボード（Ａｎａｌｏｇ Ｄｅｖｉｃｅｓ，Ｉｎｃ．，Ｎｏｒｗｏｏｄ ＭＡ）を用いてディジタル化される。ディジタル化されたデータは、青色（低シグナル）から赤色（高シグナル）までの擬似カラー範囲へのリニア２０色変換を用いてシグナル強度がマッピングされたようなイメージとして表示される。データは、定量的にも分析される。２つの異なる蛍光色素を同時に励起及び測定する場合には、各蛍光色素の発光スペクトルを用いて、データは先ず蛍光色素間の光学的クロストーク（発光スペクトルの重なりに起因する）を補正する。
【０２８１】
グリッドが蛍光シグナルイメージ上に重ねられ、それによって各スポットからのシグナルはグリッドの各エレメントに集められる。各エレメント内の蛍光シグナルは統合され、シグナルの平均強度に応じた数値が得られる。シグナル分析に用いるソフトウェアは、ＧＥＭＴＯＯＬＳ遺伝子発現分析プログラム（Ｉｎｃｙｔｅ）である。
【０２８２】
１１　相補的ポリヌクレオチド
ＡＰＲＧをコードする配列あるいはその任意の一部に対して相補的な配列は、天然のＡＰＲＧの発現を低下させるため即ち阻害するために用いられる。約１５～３０塩基対を含むオリゴヌクレオチドの使用について記すが、これより小さなあるいは大きな配列の断片の場合でも本質的に同じ方法を用いることができる。Ｏｌｉｇｏ４．０６ソフトウェア（Ｎａｔｉｏｎａｌ Ｂｉｏｓｃｉｅｎｃｅｓ）及びＡＰＲＧのコーディング配列を用いて、適切なオリゴヌクレオチドを設計する。転写を阻害するためには、最も独特な５’配列から相補的オリゴヌクレオチドを設計し、これを用いてプロモーターがコーディング配列に結合するのを阻害する。翻訳を阻害するためには、相補的なオリゴヌクレオチドを設計して、リボソームがＡＰＲＧをコードする転写物に結合するのを阻害する。
【０２８３】
１２　ＡＰＲＧの発現
ＡＰＲＧの発現及び精製は、細菌若しくはウイルスを基にした発現系を用いて行うことができる。細菌でＡＰＲＧが発現するために、抗生物質耐性及びｃＤＮＡの転写レベルを高める誘導性のプロモーターを含む好適なベクターに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、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（ｔａｃ）ハイブリッドプロモー

ターが含まれるが、これらに限定するものではない。組換えベクターを、ＢＬ２１（ＤＥ３）等の好適な細菌宿主に形質転換する。抗生物質耐性をもつ細菌が、イソプロピルβ－Ｄチオガラクトピラノシド（ＩＰＴＧ）で誘発されるとＡＰＲＧを発現する。真核細胞でのＡＰＲＧの発現は、昆虫細胞株または哺乳動物細胞株に一般にバキュロウイスルスとして知られているＡｕｔｏｇｒａｐｈｉｃａ ｃａｌｉｆｏｒｎｉｃａ核多角体病ウイルス（ＡｃＭＮＰＶ）を感染させて行う。バキュロウイルスの非必須ポリヘドリン遺伝子を、相同組換えあるいは転移プラスミドの媒介を伴う細菌の媒介による遺伝子転移のどちらかによって、ＡＰＲＧをコードするｃＤＮＡと置換する。ウイルスの感染力は維持され、強い多角体プロモーターによって高いレベルのｃＤＮＡの転写が行われる。組換えバキュロウイルスは、多くの場合はＳｐｏｄｏｐｔｅｒａ ｆｒｕｇｉｐｅｒｄａ（Ｓｆ９）昆虫細胞に感染に用いられるが、ヒト肝細胞の感染にも用いられることもある。後者の感染の場合は、バキュロウイルスの更なる遺伝的変更が必要になる。（Ｅｎｇｅｌｈａｒｄ．Ｅ．Ｋ．他．（１９９４）Ｐｒｏｃ．Ｎａｔｌ．Ａｃａｄ．Ｓｃｉ．ＵＳＡ９１：３２２４－３２２７、Ｓａｎｄｉｇ，Ｖ．他．（１９９６）Ｈｕｍ．Ｇｅｎｅ Ｔｈｅｒ．７：１９３７－１９４５．）。

【０２８４】

殆どの発現系では、ＡＰＲＧが、例えばグルタチオンＳトランスフェラーゼ（ＧＳＴ）、またはＦＬＡＧや６－Ｈｉｓなどのペプチドエピトープ標識で合成された融合タンパク質となるため、未精製の細胞溶解物からの組換え融合タンパク質の親和性ベースの精製が素早く１回で行うことができる。ＧＳＴは日本住血吸虫からの２６ｋＤａの酵素であり、タンパク質の活性及び抗原性を維持した状態で、固定化グルタチオン上で融合タンパク質の精製を可能とする（Ａｍｅｒｓｈａｍ Ｐｈａｒｍａｃｉａ Ｂｉｏｔｅｃｈ）。精製の後、ＧＳＴ部分を特定の操作部位でＡＰＲＧからタンパク分解的に切断できる。ＦＬＡＧは８アミノ酸のペプチドであり、市販されているモノクローナル及びポリクローナル抗ＦＬＡＧ抗体（Ｅａｓｔｍａｎ Ｋｏｄａｋ）を用いて免疫親和性精製を可能にする。６ヒスチジン残基が連続して伸長した６－Ｈｉｓは、金属キレート樹脂（ＱＩＡＧＥＮ）上での精製を可能にする。タンパク質の発現及び精製の方法は、前出のＡｕｓｕｂｅｌ（１９９５）１０章、１６章に記載されている。これらの方法で精製したＡＰＲＧを直接用いて以下の実施例１６、及び１７のアッセイを行うことができる。

【０２８５】

１３　機能的アッセイ

ＡＰＲＧの機能は、哺乳動物細胞培養系において生理学的に高められたレベルでのＡＰＲＧをコードする配列の発現によって評価する。ｃＤＮＡを、ｃＤＮＡを高いレベルで発現する強いプロモーターを含む哺乳動物発現ベクターにサブクローニングする。選り抜きのベクターには、ｐＣＭＶ ＳＰＯＲＴプラスミド（Ｌｉｆｅ Ｔｅｃｈｎｏｌｏｇｉｅｓ）及びｐＣＲ ３．１プラスミド（Ｉｎｖｉｔｒｏｇｅｎ）が含まれ、どちらもサイトメガロウイルスプロモーターを有する。リポソーム製剤あるいは電気穿孔法を用いて、５～１０μｇの組換えベクターをヒト細胞株、例えば内皮由来または造血由来の細胞株に一時的に形質移入する。更に、標識タンパク質をコードする配列を含む１～２μｇのプラスミドを同時に形質移入する。標識タンパク質の発現により、形質移入細胞と非形質移入細胞を区別する手段が与えられる。また、標識タンパク質の発現によって、ｃＤＮＡの組換えベクターからの発現を正確に予想できる。標識タンパク質は、例えば緑色蛍光タンパク質（ＧＦＰ；Ｃｌｏｎｔｅｃｈ）、ＣＤ６４またはＣＤ６４－ＧＦＰ融合タンパク質から選択できる。自動化された、レーザー光学に基づく技術であるフローサイトメトリー（ＦＣＭ）を用いて、ＧＦＰまたはＣＤ６４－ＧＦＰを発現する形質移入された細胞を同定し、その細胞のアポトーシス状態や他 の細胞特性を評価する。ＦＣＭは、細胞死に先行するかあるいは同時に発生する現象を診断する蛍光分子の取込を検出して計量する。このような現象として挙げられるのは、プロピジウムヨウ化物によるＤＮＡ染色によって計測される核ＤＮＡ内容物の変化、前方光散乱と９０°側方光散乱によって計測される細胞サイズと顆粒状性の変化、ブロモデオキシウリジンの取込量の低下によって計測されるＤＮＡ合成

の下方調節、特異抗体との反応性によって計測される細胞表面及び細胞内におけるタンパク質の発現の変化、及び蛍光複合アネキシンＶタンパク質の細胞表面への結合によって計測される原形質膜組成の変化とがある。フローサイトメトリー法については、Ｏｒｍｅｒｏｄ，Ｍ．Ｇ．（１９９４）Ｆｌｏｗ Ｃｙｔｏｍｅｔｒｙ Ｏｘｆｏｒｄ，Ｎｅｗ Ｙｏｒｋ，ＮＹ．に記述がある。

【０２８６】

遺伝子発現におけるＡＰＲＧの影響は、ＡＰＲＧをコードする配列とＣＤ６４またはＣＤ６４－ＧＦＰのどちらかが形質移入された高度に精製された細胞集団を用いて評価することができる。ＣＤ６４またはＣＤ６４－ＧＦＰは、形質転換された細胞表面で発現し、ヒト免疫グロブリンＧ（ＩｇＧ）の保存領域と結合する。形質転換された細胞と形質転換されない細胞とは、ヒトＩｇＧかＣＤ６４に対する抗体のどちらかで被覆された磁気ビードを用いて分離することができる（ＤＹＮＡＬ．Ｌａｋｅ Ｓｕｃｃｅｓｓ．ＮＹ）。ｍＲＮＡは、当分野で周知の方法で細胞から精製することができる。ＡＰＲＧ及び目的の他の遺伝子をコードするｍＲＮＡの発現は、ノーザン分析やマイクロアレイ技術で分析することができる。

【０２８７】

１４　ＡＴＲＳに特異的な抗体の作製

ポリアクリルアミドゲル電気泳動法（ＰＡＧＥ；Ｈａｒｒｉｎｇｔｏｎ，Ｍ．Ｇ．（１９９０）Ｍｅｔｈｏｄｓ Ｅｎｚｙｍｏｌ．１８２：４８８－４９５等を参照）または他の精製技術を用いて実質上精製されたＡＰＲＧを用いて、標準プロトコルでウサギを免疫化して抗体を産出する。

【０２８８】

別法では、ＡＰＲＧアミノ酸配列をＬＡＳＥＲＧＥＮＥソフトウェア（ＤＮＡＳＴＡＲ）を用いて解析して免疫原性の高い領域を決定し、対応するオリゴペプチドを合成してこれを用いて当業者に周知の方法で抗体を生産する。Ｃ末端付近の、あるいは隣接する親水性領域内のエピトープなどの適切なエピトープの選択については、当分野で周知である（例えば、前出のＡｕｓｕｂｅｌ，１９９５，１１章を参照）。

【０２８９】

通常は、長さ約１５残基のオリゴペプチドを、Ｆｍｏｃケミストリを用いるＡＢＩ ４３１Ａ ペプチドシンセサイザ（Ａｐｐｌｉｅｄ Ｂｉｏｓｙｓｔｅｍｓ）を用いて合成し、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（ＭＢＳ）を用いた反応によってＫＬＨ（Ｓｉｇｍａ－Ａｌｄｒｉｃｈ，Ｓｔ．Ｌｏｕｉｓ ＭＯ）に結合させて、免疫抗原性を高める（前出のＡｕｓｕｂｅｌ，１９９５）。完全フロイントアジュバントにおいてオリゴペプチド－ＫＬＭ複合体を用いてウサギを免疫化する。得られた抗血清の抗ペプチド活性及び抗ＡＰＲＧ活性を検査するには、ペプチドまたはＡＰＲＧを基板に結合し、１％ＢＳＡを用いてブロッキング処理し、ウサギ抗血清と反応させて洗浄し、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。

【０２９０】

１５　特異的抗体を用いる天然ＡＰＲＧの精製

天然ＡＰＲＧあるいは組換えＡＰＲＧを、ＡＰＲＧに特異的な抗体を用いるイムノアフィニティークロマトグラフィにより実質的に精製する。イムノアフィニティーカラムは、ＣＮＢｒ－活性化ＳＥＰＨＡＲＯＳＥ（Ａｍｅｒｓｈａｍ Ｐｈａｒｍａｃｉａ Ｂｉｏｔｅｃｈ）のような活性化クロマトグラフィー用レジンと抗ＡＰＲＧ抗体とを共有結合させることにより形成する。結合後に、製造者の使用説明書に従って樹脂をブロックし、洗浄する。

【０２９１】

ＡＰＲＧを含む培養液をイムノアフィニティーカラムに通し、ＡＰＲＧを優先的に吸着できる条件で（例えば、界面活性剤の存在下において高イオン強度のバッファーで）そのカラムを洗浄する。そのカラムを、抗体とＡＰＲＧとの結合を切るような条件で（例えば、ｐＨ２～３のバッファー、あるいは高濃度の尿素またはチオシアン酸塩イオンのようなカ

オトロピックイオンで）溶出させ、ＡＰＲＧを回収する。

【０２９２】

１６　ＡＰＲＧと相互作用する分子の同定

ＡＰＲＧまたは生物学的に活性であるＡＴＲＳ断片を、^{125}Ｉボルトンハンター試薬で標識する。（例えば Ｂｏｌｔｏｎ Ａ．Ｅ．ａｎｄ Ｗ．Ｍ．Ｈｕｎｔｅｒ（１９７３）Ｂｉｏｃｈｅｍ．Ｊ．１３３：５２９－５３９を参照。）マルチウェルプレートに予め配列しておいた候補の分子を、標識したＡＰＲＧと共にインキュベートし、洗浄して、標識したＡＰＲＧ複合体を有する全てのウェルをアッセイする。様々なＡＰＲＧ濃度で得られたデータを用いて、候補分子と結合したＡＰＲＧの数量及び親和性、会合についての値を計算する。

【０２９３】

別法では、ＡＰＲＧと相互作用する分子を、Ｆｉｅｌｄｓ，Ｓ．及びＯ．Ｓｏｎｇ（１９８９，Ｎａｔｕｒｅ ３４０：２４５－２４６）に記載の酵母２－ハイブリッドシステム（ｙｅａｓｔ ｔｗｏ－ｈｙｂｒｉｄ ｓｙｓｔｅｍ）やＭＡＴＣＨＭＡＫＥＲシステム（Ｃｌｏｎｔｅｃｈ）などの２－ハイブリッドシステムに基づいた市販のキットを用いて分析する。

ＡＰＲＧはまた、高処理型の酵母２ハイブリッドシステムを使用するＰＡＴＨＣＡＬＬＩＮＧプロセス（ＣｕｒａＧｅｎ Ｃｏｒｐ．，Ｎｅｗ Ｈａｖｅｎ ＣＴ）に用いて、遺伝子の２つの大きなライブラリによってコードされるタンパク質間の全ての相互作用を決定することができる（Ｎａｎｄａｂａｌａｎ，Ｋ．他 （２０００）米国特許第６，０５７，１０１号）。

【０２９４】

１７　ＡＰＲＧ活性の実証

このアッセイは、ＡＰＲＧを発現するスイスマウス３Ｔ３細胞において新規に合成されたＤＮＡの量として細胞増殖の度合いを測定する。当分野で公知の形質移入方法を用いて、ＡＰＲＧをコードするポリヌクレオチドを含むプラスミドを静止状態の３Ｔ３培養細胞に移入する。一過性に形質移入された細胞は次に、放射性ＤＮＡ前駆分子である［^{3}Ｈ］チミジンの存在下でインキュベートする。必要に応じて、培養した細胞に`ＡＰＲＧリガンドの変化する量を加える。酸－沈殿可能ＤＮＡへの［^{3}Ｈ］チミジンの混合は、適切な時間間隔で測定され、混合量は新規に合成されたＤＮＡの量に直接比例する。

【０２９５】

ＡＰＲＧ活性の別方は、ＡＰＲＧが哺乳動物細胞培養系において生理学的に高められたレベルで発現するときアポトーシスの誘導を評価するものである。ｃＤＮＡを、ｃＤＮＡを高いレベルで発現する強いプロモーターを含む哺乳動物発現ベクターにサブクローニングする。選り抜きのベクターには、ｐＣＭＶ ＳＰＯＲＴ（Ｌｉｆｅ Ｔｅｃｈｎｏｌｏｇｉｅｓ，Ｇａｉｔｈｅｒｓｂｕｒｇ，ＭＤ）及びｐＣＲ ３．１（Ｉｎｖｉｔｒｏｇｅｎ，Ｃａｒｌｓｂａｄ，ＣＡ）が含まれ、どちらもサイトメガロウイルスプロモーターを有する。リポソーム製剤あるいは電気穿孔法を用いて、５～１０μｇの組換えベクターをヒト細胞株、特に最適な内皮由来または造血由来の細胞株に一時的に形質移入する。更に、標識タンパク質をコードする配列を含む１～２μｇのプラスミドを同時に形質移入する。標識タンパク質の発現により、形質移入細胞と非形質移入細胞を区別する手段が与えられる。また、標識タンパク質の発現によって、ｃＤＮＡの組換えベクターからの発現を正確に予想できる。標識タンパク質は、例えば緑色蛍光タンパク質（ＧＦＰ）（Ｃｌｏｎｔｅｃｈ，Ｐａｌｏ Ａｌｔｏ，ＣＡ）、ＣＤ６４またはＣＤ６４－ＧＦＰ融合タンパク質から選択できる。自動化された、レーザー光学に基づく技術であるフローサイトメトリー（ＦＣＭ）を用いて、ＧＦＰまたはＣＤ６４－ＧＦＰを発現する形質移入された細胞を同定し、それらの細胞のアポトーシス状態を評価する。ＦＣＭは、細胞死に先行するかあるいは同時に発生する現象を診断する蛍光分子の取込を検出して計量する。このような現象として挙げられるのは、プロピジウムヨウ化物によるＤＮＡ染色によって計測される核ＤＮＡ内容物の変化、前方光散乱と９０°側方光散乱によって計測される細胞サイズ

と顆粒状性の変化、ブロモデオキシウリジンの取込量の低下によって計測されるＤＮＡ合成の下方調節、特異抗体との反応性によって計測される細胞表面及び細胞内におけるタンパンク質の発現の変化、及び蛍光複合アネキシンＶタンパク質の細胞表面への結合によって計測される原形質膜組成の変化とがある。

【０２９６】

別方として、ＡＰＲＧの活性は、哺乳動物細胞培養系において生理学的に高められたレベルでのＡＰＲＧが発現したときに増殖停止の誘導によって評価する方法があげられる。ｃＤＮＡの発現レベルを高める強いプロモーターを含む哺乳動物発現ベクターにＡＰＲＧｃＤＮＡをサブクｒ－ニングする。これらの構成体を当分野公知の方法を用いてＮＩＨ３Ｔ６ または Ｃ６のような形質転換株化細胞に安定た形質移入を行なう。ハイグロマイシン耐性等の選択可能マーカーをコードする配列を含むプラズミドを同時に形質移入する。標識タンパク質の発現により、形質移入細胞と非形質移入細胞を区別する手段が与えられる。また、標識タンパク質の発現によって、ｃＤＮＡの組換えベクターからの発現を正確に予想できる。ＡＰＲＧを発現している細胞をベクター単独の非形質移入または形質移入された細胞でコントロール細胞と増殖停止についてその特性を比較する。かかる特性は、［^{3}Ｈ］チミジンの新規に合成されたＤＮＡへの混合の減少、ダブリング及び生成時間の低下、培養液飽和密度の減少を含むが、それに限定しない。

【０２９７】

別法では、ＡＰＲＧ活性のアッセイでＤＮＡ合成時における放射標識のＤＮＡへの取り込み量を測定するため、またはアポトーシスをともなうＤＮＡのフラグメンテーションの測定するために［α^{32}Ｐ］ＡＴＰ等の放射標識されたヌクレオチドを使用する。当分野で一般的な方法を用いて、哺乳動物細胞は ＡＰＲＧをコードするｃＤＮＡを含むプラスミドとともに形質移入される。 細胞を次に、放射能標識したヌクレオチドとともに時間の長さをそれぞれ変えてインキュベードする。 染色体 ＤＮＡを回収し放射活性を放射能シンチレーションカウンタを用いて測定する。放射能標識の染色体ＤＮＡへの組み入れは細胞周期の活性度に比例する。ＡＰＲＧ がアポトーシスを促進するかどうかを判断するには、染色体ＤＮＡを上記の方法で回収し、当分野で既知のポリアクリルアミドゲル電気泳動法で分析する。ＤＮＡのフラグメンテーションを非感染コントロール細胞と比較することによって定量され、ＡＰＲＧ活性に比例する。

【０２９８】

当業者は、本発明の範囲及び精神から逸脱することなく本発明の記載した方法及びシステムの種々の改変を行い得る。本発明について説明するにあたり特定の好適実施例に関連して説明を行ったが、本発明の範囲が、そのような特定の実施例に不当に制限されるべきではないことを理解されたい。実際に、分子生物学または関連分野の専門家には明らかな、本明細書に記載されている本発明の実施方法の様々な改変は、特許請求の範囲内にあるものとする。

【０２９９】

（表の簡単な説明）

表１は、本発明の完全長ポリヌクレオチド配列及びポリペプチド配列の命名法の概略を示す。

【０３００】

表２は、ＧｅｎＢａｎｋ識別番号及び本発明のポリペプチドに最も近いＧｅｎＢａｎｋ相同体の注釈を示す。各ポリペプチドとそのＧｅｎＢａｎｋ相同体が一致する確率スコアも併せて示す。

【０３０１】

表３は、予測されるモチーフ及びドメインを含む本発明のポリヌクレオチド配列の構造的特徴を、ポリペプチドの分析に用いるための方法、アルゴリズム及び検索可能なデータベースと共に示す。

【０３０２】

表４は、本発明のポリヌクレオチド配列を構築するために用いたｃＤＮＡやゲノムＤＮＡ

断片を、ポリヌクレオチド配列の選択した断片と共に示す。

【０３０３】

表５は、本発明のポリヌクレオチドの代表的なｃＤＮＡライブラリを示す。

【０３０４】

表６は、表５に示したｃＤＮＡライブラリの作製に用いた組織及びベクターを説明する付表である。

【０３０５】

表７は、本発明のポリヌクレオチドとポリペプチドの分析に用いたツール、プログラム、アルゴリズムを、適用可能な説明、引用文献及び閾値パラメータと共に示す。

【表１】

表１

Incyte プロジェクト ID	ポリペプチド SEQ ID NO:	Incyte ポリペプチド ID	ポリヌクレオチド SEQ ID NO:	Incyte ポリヌクレオチド ID
3102521	1	3102521CD1	4	3102521CB1
7474984	2	7474984CD1	5	7474984CB1
2458545	3	2458545CD1	6	2458545CB1

【表２】

表2

ポリペプチド SEQ ID NO:	Incyte ポリペプチド ID	GenBank ID NO:	確率スコア	GenBank 相同体
1	3102521CD1	g7673638	5.10E-70	[Danio rerio (ゼブラフィッシュ)] Dedd1 (デスエフェクタードメインタンパク質) Inohara, N. および G. Nunez (2000) Genes with homology to mammalian apoptosis regulators identified in zebrafish Cell Death Differ. 7:509-510
2	7474984CD1	g10198209	1.10E-66	[ヒト] (AF298548) カスパーゼ動員ドメイン 7
3	2458545CD1	g5776545	2.10E-73	tax1-結合タンパク質 TXBP151 [ヒト]
		g11245470	1.00E-116	[fl] [Xenopus laevis (アフリカツメガエル)] 核ドメイン 10 タンパク質 NDP52 Korioth, F. 他 (1995) Molecular characterization of NDP52, a novel protein of the nuclear domain 10, which is redistributed upon virus infection and interferon treatment. J. Cell Biol. 130:1-13.

【表３】

表 3

SEQ ID NO:	Incyte ポリペプチド ID	アミノ酸 残基数	潜在的リン酸化部位	潜在的グリコシル化部位	シグネチャ配列、ドメインおよびモチーフ	分析方法及びデータベース
1	3102521CD1	326	S64 S122 S123 S129 S136 S219 S257 S298 T38 T154 T323 Y258		シグナルペプチド: M1-A56	SPScan
					デスエフェクタードメイン: L24-V110	HMMER-PFAM
					FLDBD1 (デスエフェクタードメイン) PD181655: D243-I312	BLAST-PRODOM
					KE05 タンパク質, FLDED1 (デスエフェクタードメインタンパク質) PD124668: W12-Q182	BLAST-PRODOM
2	7474984CD1	655	S18 S207 S217 S268 S316 S448 S463 S468 S491 S521 S525 S527 S554 S566 S585 S603 T177 T282 T291 T36 T377 T634 T640 T649	N380 N558 N583	膜貫通ドメイン: Y202-W225	HMMER
					PYRIN ドメイン PD136606: A6-Y98 (P = 1.1e-5)	BLAST-PRODOM
					ATP/GTP-結合部位モチーフ A (P-loop): G173-T180	MOTIFS
3	2458545CD1	691	T28 S43 S68 S72 T129 S156 S208 T239 S282 S305 T376 S383 S468 T520 T521 T537 S539 S543 S593 S595 S597 S612 T639 T652 S667 S683 T190 S241 T257	N618	RGD モチーフ, Rgd:R557-D559	MOTIFS
					ロイシンジッパーモチーフ、Leucine_Zipper:L163-L184 L475-L496 L482-L503	MOTIFS
					GHMP キナーゼ 推定上 ATP-結合タンパク質ドメイン, GHMP キナーゼ:A146-G172	HMMER-PFAM
					亜鉛フィンガー、 C2H2 タイプ zf-C2H2:K654-H679	HMMER-PFAM
					NDP52 核ドメインタンパク質 PD034365:V15-D134	BLIMPS-PRODOM

【表４】

表 4

ポリヌクレオチド SEQ ID NO:	Incyte ポリヌクレオチド ID	配列長	選択された断片	配列断片	5' 位置	3' 位置
4	3102521CB1	1230	1175-1230	7207023BV1	593	1230
				7018482H1 (KIDNNOC01)	1	668
5	7474984CB1	2460	1-945, 2400-2460, 1355-2193	GNN.g8099799_000020_006	747	1974
				55084116J2	204	1051
				58005233H1	1659	2184
				72482573D1	1	581
				58005209H1	1661	2460
6	2458545CB1	2986	1-40, 2494-2528, 2946-2986, 2183-2225, 959-1210	414248IH1 (BRSTTMT01)	801	1090
				1915072H1 (PROSTUT04)	1460	1727
				258220SF6 (KIDNTUT13)	900	1422
				1679569T7 (STOMFET01)	2304	2986
				1679569F6 (STOMFET01)	1749	2371
				3341764F6 (SPLNNOT09)	534	1086
				483375LH1 (BRAVTXT03)	1370	1674
				3472904F6 (LUNGNOT27)	418	726
				433358?H1 (KIDCTMT01)	504	788
				1419213F1 (KIDNNOT09)	1552	2076
				2458545F6 (ENDANOT01)	1	473
				4424620H1 (BRAPDIT01)	1187	1454

【表 5】

表 5

ポリヌクレオチド SEQ ID NO:	Incyte プロジェクト ID	代表的ライブラリ
4	3102521CB1	ADRETUT05
6	2458545CB1	BRSTNOT04

【表 6 】

表 6

ライブラリ	ベクター	ライブラリの説明
ADRETUT05	pINCY	ライブラリは、一側副腎摘出中に 52 歳の白人女性から採取した副腎腫瘍組織の分離 RNA を使用して構築された。病理は褐色細胞腫を示した。
BRSTNOT04	PSPORT1	ライブラリは 62 才の東インド人女性の片側性拡大単純乳房切除術時に採取した乳房組織から単離した RNA を用いて作製した。関連する腫瘍組織の病理は、浸潤性グレード 3 の乳管癌を示した。患者の病歴には良性高血圧症、高脂血症および血尿があった。家族歴には心臓血管系疾患、脳血管系疾患、高脂血症および肝癌がある。

【表 7】

表7-1

プログラム	説明	引用文献	パラメータ閾値
ABI FACTURA	ベクター配列を除き、核酸配列のまぎらわしい塩基を遮蔽するプログラム	Applied Biosystems, Foster City, CA.	
ABI/PARACEL FDF	アミノ酸または核酸配列の比較、注釈に有用なファーストデータファインダ。	Applied Biosystems, Foster City, CA. Paracel Inc., Pasadena, CA.	不一致<50%
ABIオートアセンブラ	核酸配列の構築を行うプログラム。	Applied Biosystems, Foster City, CA.	
BLAST	アミノ酸および核酸配列の配列類似性検索に有用なベーシックローカルアラインメント検索ツール。ブラストには blastp、blastn、blastx、tblastnおよびtblastxの５つの機能がある。	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	一群の同じタイプの配列と問合せ配列との間の類似性について検索する Pearson および Lipman アルゴリズム。FASTA には最小5つの機能(: fasta、tfasta、fastx、tfastx および ssearch)がある。	Pearson, W.R. and D.J. Lipman (1988) Proc. Natl. Acad Sci. USA 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*EST*: fasta同一性=95%以上 一致した長さ=200塩基以上 *ESTs:* fasta E値=1.06E-8 *完全長配列:* fastx スコア=100以上
BLIMPS	配列をBLOCKS、PRINTS、DOMO、PRODOMおよびPFAMデータベースの配列と対応させて遺伝子ファミリ、配列相同性および構造的フィンガープリント領域を検索する BLocks IMProved サーチャー。	Henikoff, S. and J.G. Henikoff (1991) Nucleic Acids Res. 19:6565-6572; Henikoff, J.G. and S. Henikoff (1996) Methods Enzymol. 266:88-105; and Attwood, T.K. 他 (1997) J. Chem. Inf. Comput. Sci. 37:417-424.	確率値= 1.0E-3以下
HMMER	問合せ配列を PFAM のようなタンパク質ファミリ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 (1988) 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: 確率値= 1.0E-3 以下 シグナルペプチドヒット：スコア=0以上

表7-2

プログラム	説明	引用文献	パラメータ閾値
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の構造的モチーフおよび配列モチーフを検索するアルゴリズム。	Gribskov, M. 他 (1988) CABIOS 4:61-66; Gribskov, M. 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≧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	高度感受性および確立性のある自動シークエンサートレースを調べる塩基コーリングアルゴリズム。	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実行に基づいたプログラムである、SWAT と CrossMatch を含むPhils 改訂構築プログラムで、配列相同性検索や DNA配列の構築に有用である。	Smith, T.F. および M.S. Waterman (1981) Adv. Appl. Math. 2:482-489; Smith, T.F. および M.S. Waterman (1981) J. Mol. Biol. 147:195-197; Green, P., University of Washington, Seattle, WA.	スコア=120以上 一致した長さ=56以上
Consed	Phrap 構築の表示、編集用グラフィカルツール。	Gordon, D. 他 (1998) Genome Res. 8:195-202.	
SPScan	タンパク質配列を走査し、分泌シグナルペプチドの存在を調べる重量マトリックス分析プログラム。	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
TMAP	重量マトリックスを使ってタンパク質配列上の膜貫通部分を描写し、定位(方向)を決定するプログラム。	Persson, B. および P. Argos (1994) J. Mol. Biol. 237:182-192; Persson, B. および P. Argos (1996) Protein Sci. 5:363-371.	
TMHMMER	隠れMarkov モデル (HMM) を使ってタンパク質配列上の膜貫通部分を描写し、定位(方向)を決定するプログラム。	Sonnhammer, E.L. 他 (1998) Proc. Sixth Intl. Conf. on Intelligent Systems for Mol. Biol., Glasgow 他, 編集, The Am. Assoc. for ArtificialIntelligence Press, Menlo Park, CA, 175-182ページ	
Motifs	Prositeで定義されたものと一致するパターンのアミノ酸配列を検索するプログラム。	Bairoch, A. 他 (1997) Nucleic Acids Res. 25:217-221; Wisconsin Package Program Manual, version 9, pagePage M51-59, Genetics Computer Group, Madison, WI	

【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
6 December 2001 (06.12.2001)

PCT

(10) International Publication Number
WO 01/92527 A2

(51) International Patent Classification⁷: C12N 15/12, C07K 14/47, 16/18, G01N 33/68, 33/53, A61K 38/17, C12Q 1/68

(21) International Application Number: PCT/US01/17581

(22) International Filing Date: 30 May 2001 (30.05.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/209,407 1 June 2000 (01.06.2000) US
60/250,326 30 November 2000 (30.11.2000) 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

(75) Inventors/Applicants (*for US only*): **TANG, Y., Tom** [US/US]; 4230 Renwick Court, San Jose, CA 95118 (US). **AZIMZAI, Yalda** [US/US]; 5518 Boulder Canyon Drive, Castro Valley, CA 94552 (US). **YUE, Henry** [US/US]; 826 Lois Avenue, Sunnyvale, CA 94087 (US). **BURFORD, Neil** [GB/US]; 105 Wildwood Circle, Durham, CT 06422 (US). **DING, Li** [CN/US]; 3353 Alma Street #146, Palo Alto, CA 94306 (US). **ELLIOTT, Vicki, S.** [US/US]; 3770 Bolton Place Way, San Jose, CA 95121 (US). **PATTERSON, Chandra** [US/US]; 490 Sherwood Way #1, Menlo Park, CA 94025 (US). **BAUGHN, Mariah, R.** [US/US]; 14244 Santiago Road, San Leandro, CA 94577 (US).

(74) Agents: **HAMLET-COX, Diana** 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, BG, BR, BY, BZ, CA, CH, CN, CO, 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.

(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, LU, MC, NL, PT, SE, TR). OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

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 01/92527 A2

(54) Title: REGULATORS OF APOPTOSIS

(57) Abstract: The invention provides human apoptosis regulators (APRG) and polynucleotides which identify and encode APRG. The invention also provides expression vectors, host cells, antibodies, agonists, and antagonists. The invention also provides methods for diagnosing, treating, or preventing disorders associated with aberrant expression of APRG.

REGULATORS OF APOPTOSIS

TECHNICAL FIELD

This invention relates to nucleic acid and amino acid sequences of apoptosis regulators and to the use of these sequences in the diagnosis, treatment, and prevention of cell proliferative, immunological, and reproductive disorders, and in the assessment of the effects of exogenous compounds on the expression of nucleic acid and amino acid sequences of apoptosis regulators.

BACKGROUND OF THE INVENTION

Tissue growth involves complex and ordered patterns of cell proliferation, cell differentiation, and regulated cell death (apoptosis). Cell proliferation and apoptosis are regulated to maintain both the number and the spatial organization of cells. This regulation depends on appropriate expression of proteins which control cell cycle progression in response to extracellular signals, such as growth factors and other mitogens, and intracellular cues, such as DNA damage or nutrient starvation. Molecules which directly or indirectly modulate cell cycle progression fall into several categories, including growth factors and their receptors, second messenger and signal transduction proteins, oncogene products, tumor-suppressor proteins, and mitosis-promoting factors. Cancers are characterized by continuous or uncontrolled cell proliferation. Some cancers are associated with suppression of normal apoptotic cell death.

Apoptosis is the genetically controlled process by which unneeded or defective cells undergo programmed cell death. Selective elimination of cells is as important for morphogenesis and tissue remodeling as is cell proliferation and differentiation. Lack of apoptosis may result in hyperplasia and other disorders associated with increased cell proliferation. Apoptosis is also a critical component of the immune response. Immune cells such as cytotoxic T-cells and natural killer cells prevent the spread of disease by inducing apoptosis in tumor cells and virus-infected cells. In addition, immune cells that fail to distinguish self molecules from foreign molecules must be eliminated by apoptosis to avoid an autoimmune response.

Apoptotic cells undergo distinct morphological changes. Hallmarks of apoptosis include cell shrinkage, nuclear and cytoplasmic condensation, and alterations in plasma membrane topology. Biochemically, apoptotic cells are characterized by increased intracellular calcium concentration, fragmentation of chromosomal DNA, and expression of novel cell surface components.

The molecular mechanisms of apoptosis are highly conserved, and many of the key protein regulators and effectors of apoptosis have been identified. Apoptosis generally proceeds in response

to a signal which is transduced intracellularly and results in altered patterns of gene expression and protein activity. Signaling molecules such as hormones and cytokines are known both to stimulate and to inhibit apoptosis through interactions with cell surface receptors.

Fragmentation of chromosomal DNA is one of the hallmarks of apoptosis. DNA fragmentation factor (DFF) is a protein composed of two subunits, a 40-kDa, caspase-activated nuclease termed DFF40/CAD, and its 45-kDa inhibitor DFF45/ICAD. Two mouse homologs of DFF45/ICAD, termed CIDE-A and CIDE-B, have recently been described (Inohara, N. et al.(1998) EMBO J. 17:2526-2533). CIDE-A and CIDE-B expression in mammalian cells activated apoptosis, while expression of CIDE-A alone induced DNA fragmentation. In addition, FAS-mediated apoptosis was enhanced by CIDE-A and CIDE-B, further implicating these proteins as effectors that mediate apoptosis.

Cancers are characterized by inappropriate cell proliferation, which may be due to uncontrolled cell growth or inadequate apoptosis. Strategies for treatment may involve either reestablishing control over cell cycle progression, or selectively stimulating apoptosis in cancerous cells (Nigg, E.A. (1995) BioEssays 17:471-480). Response to mitogenic stresses is frequently controlled at the level of transcription and is coordinated by various transcription factors. For example, the Rel/NF-kappa B family of vertebrate transcription factors plays a pivotal role in inflammatory and immune responses to radiation. The NF-kappa B family includes p50, p52, RelA, RelB, cRel, and other DNA-binding proteins. The p52 protein induces apoptosis, upregulates the transcription factor c-Jun, and activates c-Jun N-terminal kinase 1 (JNK1) (Sun, L. et al. (1998) Gene 208:157-166). Most NF-kappa B proteins form DNA-binding homodimers or heterodimers. Dimerization of many transcription factors is mediated by a conserved sequence known as the bZIP domain, characterized by a basic region followed by a leucine zipper.

The Fas/Apo-1 receptor (FAS) is a member of the tumor necrosis factor (TNF) receptor family. Upon binding its ligand (Fas ligand), the membrane-spanning FAS induces apoptosis by recruiting several cytoplasmic proteins that transmit the death signal. One such protein, termed FAS-associated protein factor 1 (FAF1), was isolated from mice, and it was demonstrated that expression of FAF1 in L cells potentiated FAS-induced apoptosis (Chu, K. et al. (1995) Proc. Natl. Acad. Sci. USA 92:11894-11898). Subsequently, FAS-associated factors have been isolated from numerous other species, including quail and fly (Frohlich, T. et al. (1998) J. Cell Sci. 111:2353-2363). Another cytoplasmic protein that functions in the transmittal of the death signal from Fas is the Fas-associated death domain protein, also known as FADD. FADD transmits the death signal in both FAS-mediated and TNF receptor-mediated apoptotic pathways by activating caspase-8 (Bang, S. et al. (2000) J. Biol. Chem. 275:36217-36222).

Immunological defenses against cancer include induction of apoptosis in mutant cells by tumor suppressors, and the recognition of tumor antigens by T lymphocytes. Response to mitogenic stresses is frequently controlled at the level of transcription and is coordinated by various transcription factors. The Rel/NF-kappa B family of vertebrate transcription factors, for example, plays a pivotal role in inflammatory and immune responses to radiation. The NF-kappa B family includes p50, p52, RelA, RelB, and cRel and other DNA-binding proteins. The p52 protein induces apoptosis, upregulates transcription factor c-Jun, and activates c-Jun N-terminal kinase 1 (JNK1) (Sun, L. et al. (1998) Gene 208:157-166). Most NF-kappa B proteins form DNA-binding homodimers or heterodimers. Dimerization of many transcription factors is mediated by a conserved sequence known as the bZIP domain, characterised by a basic region followed by a leucine zipper.

Growth Factors and Signal Transduction Machinery

Growth factors are typically large, secreted polypeptides that act on cells in their local environment to promote cell proliferation. Growth factors bind to and activate specific cell surface receptors that initiate intracellular signal transduction cascades. Many growth factor receptors are classified as receptor tyrosine kinases that undergo autophosphorylation upon ligand binding. Autophosphorylation enables the receptor to interact with signal transduction proteins such as SH2 or SH3 (Src homology regions 2 or 3) domain-containing proteins. Other proteins that act downstream of growth factor receptors contain unique signaling domains such as the SPRY (Sp1a and ryanodine receptor) domain. (See, for example, Schultz, J. et al. (1998) Proc. Natl. Acad. Sci. USA 95:5857-5864.) These proteins then modulate the activity state of small G-proteins, such as Ras, Rab, and Rho, along with GTPase activating proteins (GAPs), guanine nucleotide releasing proteins (GNRPs), and other guanine nucleotide exchange factors. Small G proteins act as molecular switches that turn on mitogen-activated protein kinase (MAP kinase) cascades. MAP kinase activates transcription of the early-response genes discussed below.

Most growth factors also have a multitude of other actions besides the regulation of cell growth and division: they can control the proliferation, survival, differentiation, migration, or function of cells depending on the circumstance. For example, epidermal growth factor (EGF) protects gastric mucosa against injury and accelerates ulcer healing by stimulating cell migration and proliferation. EGF binds the transmembrane protein tyrosine kinase EGF-R to trigger a series of events that results in activation of the Ras/Raf/MAP kinase pathway by the GTP-binding protein Ras. Other pathways potentially activated by EGF include the phosphatidylinositol pathway and the JAK/STAT signaling pathway (Tarnawski, A.S. et al. (1998) J. Clin. Gastroenterol. 27:S12-S20).

In addition to growth factors, small signaling peptides and hormones also influence cell proliferation. These molecules bind primarily to another class of receptor, the trimeric G-protein

coupled receptor (GPCR), found predominantly on the surface of immune, neuronal, and neuroendocrine cells. Upon ligand binding, the GPCR activates a trimeric G protein which in turn triggers increased levels of intracellular second messengers such as phospholipase C, Ca^{2+}, and cyclic AMP. Most GPCR-mediated signaling pathways indirectly promote cell proliferation by causing the secretion or breakdown of other signaling molecules that have direct mitogenic effects (Smith, A. et al. (1994) Cell 76:959-962).

Protein kinase C (PKC) plays a central role in the control of proliferation and differentiation of various cell types by mediating the signal transduction response to hormones and growth factors. The PKC family of serine/threonine kinases includes twelve different isoforms which have similar catalytic domains at their C-termini, but differ in their N-terminal regulatory domains. Since most cells express multiple PKC isoforms, the specificity of each enzyme for its substrate is achieved by targeting individual isoenzymes to a select location in the cell, either constitutively or upon cell stimulation. A variety of PKC-binding proteins and lipids have been identified that may function to compartmentalize PKC isoenzymes, including RACK1, serum deprivation response (sdr) protein, and SRBC (sdr-related gene product that binds C-kinase). Interestingly, both sdr and SRBC appear to provide localization of activated PKC to caveolae, but each has specificity for a different isoenzyme; sdr interacts specifically with PKCα and SRBC interacts with PKCδ. Both sdr and SRBC are induced during stages of growth arrest, and were originally isolated from serum-deprived cultured cells. Thus, sdr and SRBC appear to be important for targeting activated PKC isoenzymes to subcellular signaling sites important in growth control. (Mineo, C. et al. (1998) J. Cell Biol. 141:601-610; and Izumi, Y. et al. (1997) J. Biol. Chem. 272:7381-7389.)

Oncogenes

Oncogenes (i.e. "cancer-causing genes") are involved in the reception and transduction of growth factor signals and in the modulation of gene expression in response to these signals. For example, stimulation of a cell by growth factor activates two sets of genes, the early-response genes and the delayed-response genes. Early-response gene products include myc, fos, and jun, all of which encode gene regulatory proteins. These regulatory proteins activate the transcription of the delayed-response genes which encode proteins directly involved in cell cycle progression, such as the cyclins and cyclin dependent kinase discussed below. Additional oncogene products which directly regulate gene expression include the Rel transcription factor, the Ret zinc finger protein, and the Tre oncoprotein. (See, for example, Cao, T. et al. (1998) J. Cell Sci. 111:1319-1329; and Nakamura, T. et al. (1992) Oncogene 7:733-741.) Some conserved regions of oncogenes have been identified, such as the C3HC4 RING finger motif. Mutations in the C3HC4 RING finger domain of the Bmi-1 oncoprotein, for example, block lymphoma induction in mice (Hemenway, C.S. (1998) Oncogene 16:2541-2547).

Apoptosis inhibition motifs have also been identified, such as the BIR repeat implicated in the activity of the IAP (Inhibitor of Apoptosis) family. Mutations or chromosomal translocations which result in hyperactivation of oncogenes result in uncontrolled cell proliferation.

Tumor Suppressors

Tumor suppressor genes are involved in inhibition of cell proliferation. Mutations which decrease the activity of tumor suppressor genes result in increased cell proliferation. In humans and other mammals, tumor suppressors include the retinoblastoma (Rb) and p53 proteins. Tumor suppressors have also been discovered in lower animals such as Drosophila, in which the Discs-Large (Dlg) and Hyperplastic Discs (Hyd) proteins inhibit hyperplasia of undifferentiated epithelial cells in developing imaginal discs. (See, for example, Mansfield, E. et al. (1994) Dev. Biol. 165:507-526.) The importance of tumor suppressor genes and oncogenes in the development of cancer is demonstrated by the fact that about 75% of colorectal cancers have inactivating mutations in the p53 gene and about 50% have a hyper-activating mutation in a ras family oncogene.

Tumor supressor genes often act as "gatekeepers" (Kinzler, K.W. and Vogelstein, B. (1996) Cell 87:159-170). Normally, the gatekeeper is responsible for maintaining a balance of cell division, growth arrest, and death. External signals may activate or inactivate the gatekeeper, or alter its location within the cell. In some cases, inactivation of the gatekeeper is necessary for cell proliferation, and activation is necessary for cell growth arrest and differentiation. In other cases, the situation is reversed. Proteins which interact with the gatekeeper modify its activity or intracellular location to provide the appropriate response to external signals at any stage in the cell's development.

An example of a gatekeeper protein is the adenomatous polyposis coli (APC) protein. Though APC is expressed ubiquitously, it appears to function as a gatekeeper in colorectal cells. Mutations in the APC protein are linked to familial and sporadic forms of colon cancer. All of these mutations involve truncations in the APC C-terminus, which serves as a binding site for several proteins, including EB1, RP1, and the tumor suppressor protein Dlg. The interactions between APC and these binding proteins may be important for localizing or regulating APC activity. For example, EB1 appears to link APC to microtubules, and a defect in chromosome segregation has been implicated as an early event in colorectal tumorigenesis (Berrueta, L. (1998) Proc. Natl. Acad. Sci. USA 95:10596-10601; and Renner, C. et al. (1997) J. Immunol. 159:1276-1283).

Another example of a gatekeeper is the E2F transcription factor, which can function either as a positive regulator of cell cycle progression or as a suppressor of cell proliferation, depending on the tissue. The balance of cell division over growth arrest and differentiation appears to involve proteins which interact with and modulate E2F. These proteins include the Rb tumor suppressor protein and NPDC-1 (neural proliferation, differentiation, and control). Rb acts to repress transcriptional activity

of E2F, leading to differentiation or apoptosis in the responding cell. NPDC-1 is a neural specific gene expressed in growth arrested and differentiated cells. The NPDC-1 gene product, npdc1-1, interacts with E2F to down-regulate cell proliferation (Dupont, E. et al. (1998) J. Neurosci. Res. 51:257-267).

Cell Cycle Machinery

The molecular machinery which drives the cell cycle in response to mitogens and growth factors has been extensively studied in model systems such as budding yeast, fission yeast, and the African clawed frog, Xenopus. Essentially, the cell cycle is comprised of four successive phases: G1, S (DNA synthesis), G2, and M (mitosis). Cells which exit the cell cycle enter a quiescent phase called G0. Studies in yeast have shown that exit from S and M phases is driven by the anaphase-promoting complex, an assembly of proteins that degrades cyclins via the ubiquitin-mediated protein degradation pathway. (See, for example, Kominami, K. et al. (1998) EMBO J. 17:5388-5399.) Other non-kinase proteins, such as the Zer1p RNA splicing protein in fission yeast, are important for exit of the cell from G0 and entry into G1 or G2. (See, for example, Urushiyama, S. et al. (1997) Genetics 147:101-115.)

Several cell cycle transitions, including the entry and exit of a cell from mitosis, are dependent upon the activation and inhibition of cyclin-dependent kinases (Cdks). The Cdks are composed of a kinase subunit, Cdk, and an activating subunit, cyclin, in a complex that is subject to many levels of regulation. Cyclins bind and activate cyclin-dependent protein kinases which then phosphorylate and activate selected proteins involved in the mitotic process. The Cdk-cyclin complex is both activated and inhibited by phosphorylation. In addition, the Cdk-cyclin complex is regulated by targeted degradation involving molecules such as CDC4 and CDC53. Other proteins mediate entry into or progression through mitosis. For example, Berry and Gould recently identified a novel, 142 amino acid protein from the yeast S. pombe, termed dmp1p, that is required for proper spindle formation and entry into mitosis, but does not interact with cyclin-type proteins (Berry L.D. and Gould K.L. (1997) J. Cell Biol. 137:1337-1354). Dmp1p appears to be evolutionarily conserved, since a human homolog has recently been described (Larin D., et al. (1997) GI 2565275).

Apoptosis Machinery

The Bcl-2 family of proteins, as well as other cytoplasmic proteins, are key regulators of apoptosis. There are at least 15 Bcl-2 family members within 3 subfamilies. These proteins have been identified in mammalian cells and in viruses, and each possesses at least one of four Bcl-2 homology domains (BH1 to BH4), which are highly conserved. Bcl-2 family proteins contain the BH1 and BH2 domains, which are found in members of the pro-survival subfamily, while those proteins which are most similar to Bcl-2 have all four conserved domains, enabling inhibition of apoptosis following

encounters with a variety of cytotoxic challenges. Members of the pro-survival subfamily include Bcl-2, Bcl-x_L, Bcl-w, Mcl-1, and A1 in mammals; NR-13 (chicken); CED-9 (Caenorhabditis elegans); and viral proteins BHRF1, LMW5-HL, ORF16, KS-Bcl-2, and E1B-19K. The BH3 domain is essential for the function of pro-apoptosis subfamily proteins. The two pro-apoptosis subfamilies, Bax and BH3, include Bax, Bak, and Bok (also called Mtd); and Bik, Blk, Hrk, BNIP3, Bim_L, Bad, Bid, and Egl-1 (C. elegans): respectively. Members of the Bax subfamily contain the BH1, BH2, and BH3 domains, and resemble Bcl-2 rather closely. In contrast, members of the BH3 subfamily have only the 9-16 residue BH3 domain, being otherwise unrelated to any known protein, and only Bik and Blk share sequence similarity. The proteins of the two pro-apoptosis subfamilies may be the antagonists of pro-survival subfamily proteins. This is illustrated in C. elegans where Egl-1, which is required for apoptosis, binds to and acts via CED-9 (for review, see Adams, J.M. and S. Cory (1998) Science 281:1322-1326).

Heterodimerization between pro-apoptosis and anti-apoptosis subfamily proteins seems to have a titrating effect on the functions of these protein subfamilies, which suggests that relative concentrations of the members of each subfamily may act to regulate apoptosis. Heterodimerization is not required for a pro-survival protein; however, it is essential in the BH3 subfamily, and less so in the Bax subfamily.

The Bcl-2 protein has 2 isoforms, alpha and beta, which are formed by alternative splicing. It forms homodimers and heterodimers with Bax and Bak proteins and the Bcl-X isoform Bcl-x_S. Heterodimerization with Bax requires intact BH1 and BH2 domains, and is necessary for pro-survival activity. The BH4 domain seems to be involved in pro-survival activity as well. Bcl-2 is located within the inner and outer mitochondrial membranes, as well as within the nuclear envelope and endoplasmic reticulum, and is expressed in a variety of tissues. Its involvement in follicular lymphoma (type II chronic lymphatic leukemia) is seen in a chromosomal translocation T(14;18) (q32;q21) and involves immunoglobulin gene regions.

The Bcl-x protein is a dominant regulator of apoptotic cell death. Alternative splicing results in three isoforms, Bcl-xB, a long isoform, and a short isoform. The long isoform exhibits cell death repressor activity, while the short isoform promotes apoptosis. Bcl-xL forms heterodimers with Bax and Bak, although heterodimerization with Bax does not seem to be necessary for pro-survival (anti-apoptosis) activity. Bcl-xS forms heterodimers with Bcl-2. Bcl-x is found in mitochondrial membranes and the perinuclear envelope. Bcl-xS is expressed at high levels in developing lymphocytes and other cells undergoing a high rate of turnover. Bcl-xL is found in adult brain and in other tissues' long-lived post-mitotic cells. As with Bcl-2, the BH1, BH2, and BH4 domains are involved in pro-survival activity.

The Bcl-w protein is found within the cytoplasm of almost all myeloid cell lines and in numerous tissues, with the highest levels of expression in brain, colon, and salivary gland. This protein is expressed in low levels in testis, liver, heart, stomach, skeletal muscle, and placenta, and a few lymphoid cell lines. Bcl-w contains the BH1, BH2, and BH4 domains, all of which are needed for its cell survival promotion activity. Although mice in which Bcl-w gene function was disrupted by homologous recombination were viable, healthy, and normal in appearance, and adult females had normal reproductive function, the adult males were infertile. In these males, the initial, prepuberty stage of spermatogenesis was largely unaffected and the testes developed normally. However, the seminiferous tubules were disorganized, contained numerous apoptotic cells, and were incapable of producing mature sperm. This mouse model may be applicable to some cases of human male sterility and suggests that alteration of programmed cell death in the testes may be useful in modulating fertility (Print, C.G. et al. (1998) Proc. Natl. Acad. Sci. USA 95:12424-12431).

Studies in rat ischemic brain found Bcl-w to be overexpressed relative to its normal low constitutive level of expression in nonischemic brain. Furthermore, in vitro studies to examine the mechanism of action of Bcl-w revealed that isolated rat brain mitochondria were unable to respond to an addition of recombinant Bax or high concentrations of calcium when Bcl-w was also present. The normal response would be the release of cytochrome c from the mitochondria. Additionally, recombinant Bcl-w protein was found to inhibit calcium-induced loss of mitochondrial transmembrane potential, which is indicative of permeability transition. Together these findings suggest that Bcl-w may be a neuro-protectant against ischemic neuronal death and may achieve this protection via the mitochondrial death-regulatory pathway (Yan, C. et al. (2000) J. Cereb. Blood Flow Metab. 20:620-630).

The bfl-1 gene is an additional member of the Bcl-2 family, and is also a suppressor of apoptosis. The Bfl-1 protein has 175 amino acids, and contains the BH1, BH2, and BH3 conserved domains found in Bcl-2 family members. It also contains a Gln-rich NH2-terminal region and lacks an NH domain 1, unlike other Bcl-2 family members. The mouse A1 protein shares high sequence homology with Bfl-1 and has the 3 conserved domains found in Bfl-1. Apoptosis induced by the p53 tumor suppressor protein is suppressed by Bfl-1, similar to the action of Bcl-2, Bcl-xL, and EBV-*BHRF1* (D'Sa-Eipper, C. et al. (1996) Cancer Res. 56:3879-3882). *Bfl-1 is found intracellularly*, with the highest expression in the hematopoietic compartment, i.e. blood, spleen, and bone marrow; moderate expression in lung, small intestine, and testis; and minimal expression in other tissues. It is also found in vascular smooth muscle cells and hematopoietic malignancies. A correlation has been noted between the expression level of bfl-1 and the development of stomach cancer, suggesting that the Bfl-1 protein is involved in the development of stomach cancer, either in the promotion of cancerous cell survival or in

cancer (Choi, S.S. et al. (1995) Oncogene 11:1693-1698).

Transcription factors play an important role in the onset of apoptosis. A number of downstream effector molecules, particularly proteases such as the cysteine proteases called caspases, are involved in the initiation and execution phases of apoptosis. The activation of the caspases results from the competitive action of the pro-survival and pro-apoptosis Bcl-2-related proteins (Print, C.G. et al. (1998) Proc. Natl. Acad. Sci. USA 95:12424-12431). A pro-apoptotic signal can activate initiator caspases that trigger a proteolytic caspase cascade, leading to the hydrolysis of target proteins and the classic apoptotic death of the cell. Two active site residues, a cysteine and a histidine, have been implicated in the catalytic mechanism. Caspases are among the most specific endopeptidases, cleaving after aspartate residues.

Caspases are synthesized as inactive zymogens consisting of one large (p20) and one small (p10) subunit separated by a small spacer region, and a variable N-terminal prodomain. This prodomain interacts with cofactors that can positively or negatively affect apoptosis. An activating signal causes autoproteolytic cleavage of a specific aspartate residue (D297 in the caspase-1 numbering convention) and removal of the spacer and prodomain, leaving a p10/p20 heterodimer. Two of these heterodimers interact via their small subunits to form the catalytically active tetramer. The long prodomains of some caspase family members have been shown to promote dimerization and auto-processing of procaspases. Some caspases contain a "death effector domain" in their prodomain by which they can be recruited into self-activating complexes with other caspases and FADD protein-associated death receptors or the TNF receptor complex. In addition, two dimers from different caspase family members can associate, changing the substrate specificity of the resultant tetramer. A caspase recruitment domain (CARD) is found within the prodomain of several apical caspases and is conserved in several apoptosis regulatory molecules such as Apaf-2, RAIDD, and cellular inhibitors of apoptosis proteins (IAPs) (Hofmann, K. et al. (1997) Trends Biochem. Sci. 22:155-157). The regulatory role of CARD in apoptosis may be to allow proteins such as Apaf-1 to associate with caspase-9 (Li, P. et al. (1997) Cell 91:479-489). A human cDNA encoding an apoptosis repressor with a CARD (ARC) which is expressed in both skeletal and cardiac muscle has been identified and characterized. ARC functions as an inhibitor of apoptosis and interacts selectively with caspases (Koseki, T. et al. (1998) Proc. Natl. Acad. Sci. USA 95:5156-5160). All of these interactions have clear effects on the control of apoptosis (reviewed in Chan S.L. and M.P. Mattson (1999) J. Neurosci. Res. 58:167-190; Salveson, G.S. and V.M. Dixit (1999) Proc. Natl. Acad. Sci. USA 96:10964-10967).

The discovery of new apoptosis regulators and the polynucleotides encoding them satisfies a need in the art by providing new compositions which are useful in the diagnosis, prevention, and treatment of cell proliferative, immunological, and reproductive disorders, and in the assessment of the

effects of exogenous compounds on the expression of nucleic acid and amino acid sequences of apoptosis regulators.

SUMMARY OF THE INVENTION

The invention features purified polypeptides, apoptosis regulators, referred to collectively as "APRG" and individually as "APRG-1," "APRG-2," and "APRG-3." In one aspect, the invention provides an isolated polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-3, b) a naturally occurring polypeptide comprising an amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-3, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-3, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-3. In one alternative, the invention provides an isolated polypeptide comprising the amino acid sequence of SEQ ID NO:1-3.

The invention further provides an isolated polynucleotide encoding a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-3, b) a naturally occurring polypeptide comprising an amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-3, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-3, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-3. In one alternative, the polynucleotide encodes a polypeptide selected from the group consisting of SEQ ID NO:1-3. In another alternative, the polynucleotide is selected from the group consisting of SEQ ID NO:4-6.

Additionally, the invention provides a recombinant polynucleotide comprising a promoter sequence operably linked to a polynucleotide encoding a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-3, b) a naturally occurring polypeptide comprising an amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-3, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-3, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-3. In one alternative, the invention provides a cell transformed with the recombinant polynucleotide. In another alternative, the invention provides a transgenic organism comprising the recombinant polynucleotide.

The invention also provides a method for producing a polypeptide selected from the group

consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-3, b) a naturally occurring polypeptide comprising an amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-3, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-3, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-3. The method comprises a) culturing a cell under conditions suitable for expression of the polypeptide, wherein said cell is transformed with a recombinant polynucleotide comprising a promoter sequence operably linked to a polynucleotide encoding the polypeptide, and b) recovering the polypeptide so expressed.

Additionally, the invention provides an isolated antibody which specifically binds to a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-3, b) a naturally occurring polypeptide comprising an amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-3, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-3, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-3.

The invention further provides an isolated polynucleotide selected from the group consisting of a) a polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:4-6, b) a naturally occurring polynucleotide comprising a polynucleotide sequence at least 90% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:4-6, c) a polynucleotide complementary to the polynucleotide of a), d) a polynucleotide complementary to the polynucleotide of b), and e) an RNA equivalent of a)-d). In one alternative, the polynucleotide comprises at least 60 contiguous nucleotides.

Additionally, the invention provides a method for detecting a target polynucleotide in a sample, said target polynucleotide having a sequence of a polynucleotide selected from the group consisting of a) a polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:4-6, b) a naturally occurring polynucleotide comprising a polynucleotide sequence at least 90% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:4-6, c) a polynucleotide complementary to the polynucleotide of a), d) a polynucleotide complementary to the polynucleotide of b), and e) an RNA equivalent of a)-d). The method comprises a) hybridizing the sample with a probe comprising at least 20 contiguous nucleotides comprising a sequence complementary to said target polynucleotide in the sample, and which probe specifically hybridizes to said target polynucleotide, under conditions whereby a hybridization complex is formed between said probe and said target polynucleotide or fragments thereof, and b) detecting the presence or absence of

said hybridization complex, and optionally, if present, the amount thereof. In one alternative, the probe comprises at least 60 contiguous nucleotides.

The invention further provides a method for detecting a target polynucleotide in a sample, said target polynucleotide having a sequence of a polynucleotide selected from the group consisting of a) a polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:4-6, b) a naturally occurring polynucleotide comprising a polynucleotide sequence at least 90% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:4-6, c) a polynucleotide complementary to the polynucleotide of a), d) a polynucleotide complementary to the polynucleotide of b), and e) an RNA equivalent of a)-d). The method comprises a) amplifying said target polynucleotide or fragment thereof using polymerase chain reaction amplification, and b) detecting the presence or absence of said amplified target polynucleotide or fragment thereof, and, optionally, if present, the amount thereof.

The invention further provides a composition comprising an effective amount of a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-3, b) a naturally occurring polypeptide comprising an amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-3, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-3, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-3, and a pharmaceutically acceptable excipient. In one embodiment, the composition comprises an amino acid sequence selected from the group consisting of SEQ ID NO:1-3. The invention additionally provides a method of treating a disease or condition associated with decreased expression of functional APRG, comprising administering to a patient in need of such treatment the composition.

The invention also provides a method for screening a compound for effectiveness as an agonist of a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-3, b) a naturally occurring polypeptide comprising an amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-3, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-3, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-3. The method comprises a) exposing a sample comprising the polypeptide to a compound, and b) detecting agonist activity in the sample. In one alternative, the invention provides a composition comprising an agonist compound identified by the method and a pharmaceutically acceptable excipient. In another alternative, the invention provides a method of

treating a disease or condition associated with decreased expression of functional APRG, comprising administering to a patient in need of such treatment the composition.

Additionally, the invention provides a method for screening a compound for effectiveness as an antagonist of a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-3, b) a naturally occurring polypeptide comprising an amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-3, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-3, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-3. The method comprises a) exposing a sample comprising the polypeptide to a compound, and b) detecting antagonist activity in the sample. In one alternative, the invention provides a composition comprising an antagonist compound identified by the method and a pharmaceutically acceptable excipient. In another alternative, the invention provides a method of treating a disease or condition associated with overexpression of functional APRG, comprising administering to a patient in need of such treatment the composition.

The invention further provides a method of screening for a compound that specifically binds to a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-3, b) a naturally occurring polypeptide comprising an amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-3, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-3, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-3. The method comprises a) combining the polypeptide with at least one test compound under suitable conditions, and b) detecting binding of the polypeptide to the test compound, thereby identifying a compound that specifically binds to the polypeptide.

The invention further provides a method of screening for a compound that modulates the activity of a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-3, b) a naturally occurring polypeptide comprising an amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-3, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-3, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-3. The method comprises a) combining the polypeptide with at least one test compound under conditions permissive for the activity of the polypeptide, b) assessing the

activity of the polypeptide in the presence of the test compound, and c) comparing the activity of the polypeptide in the presence of the test compound with the activity of the polypeptide in the absence of the test compound, wherein a change in the activity of the polypeptide in the presence of the test compound is indicative of a compound that modulates the activity of the polypeptide.

The invention further provides a method for screening a compound for effectiveness in altering expression of a target polynucleotide, wherein said target polynucleotide comprises a sequence selected from the group consisting of SEQ ID NO:4-6, the method comprising a) exposing a sample comprising the target polynucleotide to a compound, and b) detecting altered expression of the target polynucleotide.

The invention further provides a method for assessing toxicity of a test compound, said method comprising a) treating a biological sample containing nucleic acids with the test compound; b) hybridizing the nucleic acids of the treated biological sample with a probe comprising at least 20 contiguous nucleotides of a polynucleotide selected from the group consisting of i) a polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:4-6, ii) a naturally occurring polynucleotide comprising a polynucleotide sequence at least 90% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:4-6, iii) a polynucleotide having a sequence complementary to i), iv) a polynucleotide complementary to the polynucleotide of ii), and v) an RNA equivalent of i)-iv). Hybridization occurs under conditions whereby a specific hybridization complex is formed between said probe and a target polynucleotide in the biological sample, said target polynucleotide selected from the group consisting of i) a polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:4-6, ii) a naturally occurring polynucleotide comprising a polynucleotide sequence at least 90% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:4-6, iii) a polynucleotide complementary to the polynucleotide of i), iv) a polynucleotide complementary to the polynucleotide of ii), and v) an RNA equivalent of i)-iv). Alternatively, the target polynucleotide comprises a fragment of a polynucleotide sequence selected from the group consisting of i)-v) above; c) quantifying the amount of hybridization complex; and d) comparing the amount of hybridization complex in the treated biological sample with the amount of hybridization complex in an untreated biological sample, wherein a difference in the amount of hybridization complex in the treated biological sample is indicative of toxicity of the test compound.

BRIEF DESCRIPTION OF THE TABLES

Table 1 summarizes the nomenclature for the full length polynucleotide and polypeptide sequences of the present invention.

Table 2 shows the GenBank identification number and annotation of the nearest GenBank

homolog for polypeptides of the invention. The probability score for the match between each polypeptide and its GenBank homolog is also shown.

Table 3 shows structural features of polypeptide sequences of the invention, including predicted motifs and domains, along with the methods, algorithms, and searchable databases used for analysis of the polypeptides.

Table 4 lists the cDNA and/or genomic DNA fragments which were used to assemble polynucleotide sequences of the invention, along with selected fragments of the polynucleotide sequences.

Table 5 shows the representative cDNA library for polynucleotides of the invention.

Table 6 provides an appendix which describes the tissues and vectors used for construction of the cDNA libraries shown in Table 5.

Table 7 shows the tools, programs, and algorithms used to analyze the polynucleotides and polypeptides of the invention, along with applicable descriptions, references, and threshold parameters.

DESCRIPTION OF THE INVENTION

Before the present proteins, nucleotide sequences, and methods are described, it is understood that this invention is not limited to the particular machines, materials and methods described, as these may vary. 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.

It must be noted that as used herein and in the appended claims, the singular forms "a," "an," and "the" include plural reference unless the context clearly dictates otherwise. Thus, for example, a reference to "a host cell" includes a plurality of such host cells, and a reference to "an antibody" is a reference to one or more antibodies and equivalents thereof known to those skilled in the art, and so forth.

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. Although any machines, materials, and methods similar or equivalent to those described herein can be used to practice or test the present invention, the preferred machines, materials and methods are now described. 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

"APRG" refers to the amino acid sequences of substantially purified APRG obtained from any species, particularly a mammalian species, including bovine, ovine, porcine, murine, equine, and human, and from any source, whether natural, synthetic, semi-synthetic, or recombinant.

The term "agonist" refers to a molecule which intensifies or mimics the biological activity of APRG. Agonists may include proteins, nucleic acids, carbohydrates, small molecules, or any other compound or composition which modulates the activity of APRG either by directly interacting with APRG or by acting on components of the biological pathway in which APRG participates.

An "allelic variant" is an alternative form of the gene encoding APRG. Allelic variants may result from at least one mutation in the nucleic acid sequence and may result in altered mRNAs or in polypeptides whose structure or function may or may not be altered. A gene may have none, one, or many allelic variants of its naturally occurring form. Common mutational changes which give rise to allelic variants are generally ascribed to natural deletions, additions, or substitutions of nucleotides. Each of these types of changes may occur alone, or in combination with the others, one or more times in a given sequence.

"Altered" nucleic acid sequences encoding APRG include those sequences with deletions, insertions, or substitutions of different nucleotides, resulting in a polypeptide the same as APRG or a polypeptide with at least one functional characteristic of APRG. Included within this definition are polymorphisms which may or may not be readily detectable using a particular oligonucleotide probe of the polynucleotide encoding APRG, and improper or unexpected hybridization to allelic variants, with a locus other than the normal chromosomal locus for the polynucleotide sequence encoding APRG. The encoded protein may also be "altered," and may contain deletions, insertions, or substitutions of amino acid residues which produce a silent change and result in a functionally equivalent APRG. Deliberate amino acid substitutions may be made on the basis of similarity in polarity, charge, solubility, hydrophobicity, hydrophilicity, and/or the amphipathic nature of the residues, as long as the biological or immunological activity of APRG is retained. For example, negatively charged amino acids may include aspartic acid and glutamic acid, and positively charged amino acids may include lysine and arginine. Amino acids with uncharged polar side chains having similar hydrophilicity values may include: asparagine and glutamine; and serine and threonine. Amino acids with uncharged side chains having similar hydrophilicity values may include: leucine, isoleucine, and valine; glycine and alanine; and phenylalanine and tyrosine.

The terms "amino acid" and "amino acid sequence" refer to an oligopeptide, peptide, polypeptide, or protein sequence, or a fragment of any of these, and to naturally occurring or synthetic molecules. Where "amino acid sequence" is recited to refer to a sequence of a naturally occurring

protein molecule, "amino acid sequence" and like terms are not meant to limit the amino acid sequence to the complete native amino acid sequence associated with the recited protein molecule.

"Amplification" relates to the production of additional copies of a nucleic acid sequence. Amplification is generally carried out using polymerase chain reaction (PCR) technologies well known in the art.

The term "antagonist" refers to a molecule which inhibits or attenuates the biological activity of APRG. Antagonists may include proteins such as antibodies, nucleic acids, carbohydrates, small molecules, or any other compound or composition which modulates the activity of APRG either by directly interacting with APRG or by acting on components of the biological pathway in which APRG participates.

The term "antibody" refers to intact immunoglobulin molecules as well as to fragments thereof, such as Fab, $F(ab')_2$, and Fv fragments, which are capable of binding an epitopic determinant. Antibodies that bind APRG polypeptides can be prepared using intact polypeptides or using fragments containing small peptides of interest as the immunizing antigen. The polypeptide or oligopeptide used to immunize an animal (e.g., a mouse, a rat, or a rabbit) can be derived from the translation of RNA, or synthesized chemically, and can be conjugated to a carrier protein if desired. Commonly used carriers that are chemically coupled to peptides include bovine serum albumin, thyroglobulin, and keyhole limpet hemocyanin (KLH). The coupled peptide is then used to immunize the animal.

The term "antigenic determinant" refers to that region of a molecule (i.e., an epitope) that makes contact with a particular antibody. When a protein or a fragment of a protein is used to immunize a host animal, numerous regions of the protein may induce the production of antibodies which bind specifically to antigenic determinants (particular regions or three-dimensional structures on the protein). An antigenic determinant may compete with the intact antigen (i.e., the immunogen used to elicit the immune response) for binding to an antibody.

The term "antisense" refers to any composition capable of base-pairing with the "sense" (coding) strand of a specific nucleic acid sequence. Antisense compositions may include DNA; RNA; peptide nucleic acid (PNA); oligonucleotides having modified backbone linkages such as phosphorothioates, methylphosphonates, or benzylphosphonates; oligonucleotides having modified sugar groups such as 2'-methoxyethyl sugars or 2'-methoxyethoxy sugars; or oligonucleotides having modified bases such as 5-methyl cytosine, 2'-deoxyuracil, or 7-deaza-2'-deoxyguanosine. Antisense molecules may be produced by any method including chemical synthesis or transcription. Once introduced into a cell, the complementary antisense molecule base-pairs with a naturally occurring nucleic acid sequence produced by the cell to form duplexes which block either transcription or translation. The designation "negative" or "minus" can refer to the antisense strand, and the

designation "positive" or "plus" can refer to the sense strand of a reference DNA molecule.

The term "biologically active" refers to a protein having structural, regulatory, or biochemical functions of a naturally occurring molecule. Likewise, "immunologically active" or "immunogenic" refers to the capability of the natural, recombinant, or synthetic APRG, or of any oligopeptide thereof, to induce a specific immune response in appropriate animals or cells and to bind with specific antibodies.

"Complementary" describes the relationship between two single-stranded nucleic acid sequences that anneal by base-pairing. For example, 5'-AGT-3' pairs with its complement, 3'-TCA-5'.

A "composition comprising a given polynucleotide sequence" and a "composition comprising a given amino acid sequence" refer broadly to any composition containing the given polynucleotide or amino acid sequence. The composition may comprise a dry formulation or an aqueous solution. Compositions comprising polynucleotide sequences encoding APRG or fragments of APRG may be employed as hybridization probes. The probes may be stored in freeze-dried form and may be associated with a stabilizing agent such as a carbohydrate. In hybridizations, the probe may be deployed in an aqueous solution containing salts (e.g., NaCl), detergents (e.g., sodium dodecyl sulfate; SDS), and other components (e.g., Denhardt's solution, dry milk, salmon sperm DNA, etc.).

"Consensus sequence" refers to a nucleic acid sequence which has been subjected to repeated DNA sequence analysis to resolve uncalled bases, extended using the XL-PCR kit (Applied Biosystems, Foster City CA) in the 5' and/or the 3' direction, and resequenced, or which has been assembled from one or more overlapping cDNA, EST, or genomic DNA fragments using a computer program for fragment assembly, such as the GELVIEW fragment assembly system (GCG, Madison WI) or Phrap (University of Washington, Seattle WA). Some sequences have been both extended and assembled to produce the consensus sequence.

"Conservative amino acid substitutions" are those substitutions that are predicted to 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 table below shows amino acids which may be substituted for an original amino acid in a protein and which are regarded as conservative amino acid substitutions.

Original Residue	Conservative Substitution
Ala	Gly, Ser
Arg	His, Lys
Asn	Asp, Gln, His
Asp	Asn, Glu
Cys	Ala, Ser
Gln	Asn, Glu, His

Glu	Asp, Gln, His
Gly	Ala
His	Asn, Arg, Gln, Glu
Ile	Leu, Val
Leu	Ile, Val
Lys	Arg, Gln, Glu
Met	Leu, Ile
Phe	His, Met, Leu, Trp, Tyr
Ser	Cys, Thr
Thr	Ser, Val
Trp	Phe, Tyr
Tyr	His, Phe, Trp
Val	Ile, Leu, Thr

Conservative amino acid substitutions generally maintain (a) the structure of the polypeptide backbone in the area of the substitution, for example, as a beta sheet or alpha helical conformation, (b) the charge or hydrophobicity of the molecule at the site of the substitution, and/or (c) the bulk of the side chain.

A "deletion" refers to a change in the amino acid or nucleotide sequence that results in the absence of one or more amino acid residues or nucleotides.

The term "derivative" refers to a chemically modified polynucleotide or polypeptide. Chemical modifications of a polynucleotide can include, for example, replacement of hydrogen by an alkyl, acyl, hydroxyl, or amino group. A derivative polynucleotide encodes a polypeptide which retains at least one biological or immunological function of the natural molecule. A derivative polypeptide is one modified by glycosylation, pegylation, or any similar process that retains at least one biological or immunological function of the polypeptide from which it was derived.

A "detectable label" refers to a reporter molecule or enzyme that is capable of generating a measurable signal and is covalently or noncovalently joined to a polynucleotide or polypeptide.

"Differential expression" refers to increased or upregulated; or decreased, downregulated, or absent gene or protein expression, determined by comparing at least two different samples. Such comparisons may be carried out between, for example, a treated and an untreated sample, or a diseased and a normal sample.

A "fragment" is a unique portion of APRG or the polynucleotide encoding APRG which is identical in sequence to but shorter in length than the parent sequence. A fragment may comprise up to the entire length of the defined sequence, minus one nucleotide/amino acid residue. For example, a fragment may comprise from 5 to 1000 contiguous nucleotides or amino acid residues. A fragment used as a probe, primer, antigen, therapeutic molecule, or for other purposes, may be at least 5, 10, 15, 16, 20, 25, 30, 40, 50, 60, 75, 100, 150, 250 or at least 500 contiguous nucleotides or amino acid residues in length. Fragments may be preferentially selected from certain regions of a molecule. For

example, a polypeptide fragment may comprise a certain length of contiguous amino acids selected from the first 250 or 500 amino acids (or first 25% or 50%) of a polypeptide as shown in a certain defined sequence. Clearly these lengths are exemplary, and any length that is supported by the specification, including the Sequence Listing, tables, and figures, may be encompassed by the present embodiments.

A fragment of SEQ ID NO:4-6 comprises a region of unique polynucleotide sequence that specifically identifies SEQ ID NO:4-6, for example, as distinct from any other sequence in the genome from which the fragment was obtained. A fragment of SEQ ID NO:4-6 is useful, for example, in hybridization and amplification technologies and in analogous methods that distinguish SEQ ID NO:4-6 from related polynucleotide sequences. The precise length of a fragment of SEQ ID NO:4-6 and the region of SEQ ID NO:4-6 to which the fragment corresponds are routinely determinable by one of ordinary skill in the art based on the intended purpose for the fragment.

A fragment of SEQ ID NO:1-3 is encoded by a fragment of SEQ ID NO:4-6. A fragment of SEQ ID NO:1-3 comprises a region of unique amino acid sequence that specifically identifies SEQ ID NO:1-3. For example, a fragment of SEQ ID NO:1-3 is useful as an immunogenic peptide for the development of antibodies that specifically recognize SEQ ID NO:1-3. The precise length of a fragment of SEQ ID NO:1-3 and the region of SEQ ID NO:1-3 to which the fragment corresponds are routinely determinable by one of ordinary skill in the art based on the intended purpose for the fragment.

A "full length" polynucleotide sequence is one containing at least a translation initiation codon (e.g., methionine) followed by an open reading frame and a translation termination codon. A "full length" polynucleotide sequence encodes a "full length" polypeptide sequence.

"Homology" refers to sequence similarity or, interchangeably, sequence identity, between two or more polynucleotide sequences or two or more polypeptide sequences.

The terms "percent identity" and "% identity," as applied to polynucleotide sequences, refer to the percentage of 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.

Percent identity between polynucleotide sequences may be determined using the default parameters of the CLUSTAL V algorithm as incorporated into the MEGALIGN version 3.12e sequence alignment program. This program is part of the LASERGENE software package, a suite of molecular biological analysis programs (DNASTAR, Madison WI). CLUSTAL V is described in Higgins, D.G. and P.M. Sharp (1989) CABIOS 5:151-153 and in Higgins, D.G. et al. (1992) CABIOS 8:189-191.

For pairwise alignments of polynucleotide sequences, the default parameters are set as follows: Ktuple=2, gap penalty=5, window=4, and "diagonals saved"=4. The "weighted" residue weight table is selected as the default. Percent identity is reported by CLUSTAL V as the "percent similarity" between aligned polynucleotide sequences.

Alternatively, a suite of commonly used and freely available sequence comparison algorithms is provided by the National Center for Biotechnology Information (NCBI) Basic Local Alignment Search Tool (BLAST) (Altschul, S.F. et al. (1990) J. Mol. Biol. 215:403-410), which is available from several sources, including the NCBI, Bethesda, MD, and on the Internet at http://www.ncbi.nlm.nih.gov/BLAST/. 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/gorf/bl2.html. The "BLAST 2 Sequences" tool can be used for both blastn and blastp (discussed below). 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 2.0.12 (April-21-2000) 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

Filter: on

Percent identity 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. 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.

Nucleic acid sequences that do not show a high degree of identity may nevertheless encode similar amino acid sequences due to the degeneracy of the genetic code. It is understood that changes in

a nucleic acid sequence can be made using this degeneracy to produce multiple nucleic acid sequences that all encode substantially the same protein.

The phrases "percent identity" and "% identity," as applied to polypeptide sequences, refer to the percentage of residue matches between at least two polypeptide sequences aligned using a standardized algorithm. Methods of polypeptide sequence alignment are well-known. Some alignment methods take into account conservative amino acid substitutions. Such conservative substitutions, explained in more detail above, generally preserve the charge and hydrophobicity at the site of substitution, thus preserving the structure (and therefore function) of the polypeptide.

Percent identity between polypeptide sequences may be determined using the default parameters of the CLUSTAL V algorithm as incorporated into the MEGALIGN version 3.12e sequence alignment program (described and referenced above). For pairwise alignments of polypeptide sequences using CLUSTAL V, the default parameters are set as follows: Ktuple=1, gap penalty=3, window=5, and "diagonals saved"=5. The PAM250 matrix is selected as the default residue weight table. As with polynucleotide alignments, the percent identity is reported by CLUSTAL V as the "percent similarity" between aligned polypeptide sequence pairs.

Alternatively the NCBI BLAST software suite may be used. For example, for a pairwise comparison of two polypeptide sequences, one may use the "BLAST 2 Sequences" tool Version 2.0.12 (April-21-2000) with blastp set at default parameters. Such default parameters may be, for example:

Matrix: BLOSUM62

Open Gap: 11 and Extension Gap: 1 penalties

Gap x drop-off: 50

Expect: 10

Word Size: 3

Filter: on

Percent identity may be measured over the length of an entire defined polypeptide 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 polypeptide sequence, for instance, a fragment of at least 15, at least 20, at least 30, at least 40, at least 50, at least 70 or at least 150 contiguous residues. 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.

"Human artificial chromosomes" (HACs) are linear microchromosomes which may contain DNA sequences of about 6 kb to 10 Mb in size and which contain all of the elements required for chromosome replication, segregation and maintenance.

The term "humanized antibody" refers to an antibody molecule in which the amino acid sequence in the non-antigen binding regions has been altered so that the antibody more closely resembles a human antibody, and still retains its original binding ability.

"Hybridization" refers to the process by which a polynucleotide strand anneals with a complementary strand through base pairing under defined hybridization conditions. Specific hybridization is an indication that two nucleic acid sequences share a high degree of complementarity. Specific hybridization complexes form under permissive annealing conditions and remain hybridized after the "washing" step(s). The washing step(s) is particularly important in determining the stringency of the hybridization process, with more stringent conditions allowing less non-specific binding, i.e., binding between pairs of nucleic acid strands that are not perfectly matched. Permissive conditions for annealing of nucleic acid sequences are routinely determinable by one of ordinary skill in the art and may be consistent among hybridization experiments, whereas wash conditions may be varied among experiments to achieve the desired stringency, and therefore hybridization specificity. Permissive annealing conditions occur, for example, at 68°C in the presence of about 6 x SSC, about 1% (w/v) SDS, and about 100 µg/ml sheared, denatured salmon sperm DNA.

Generally, stringency of hybridization is expressed, in part, with reference to the temperature under which the wash step is carried out. Such wash temperatures are typically selected to be about 5°C to 20°C lower than the thermal melting point (T_m) for the specific sequence at a defined ionic strength and pH. The T_m is the temperature (under defined ionic strength and pH) at which 50% of the target sequence hybridizes to a perfectly matched probe. An equation for calculating T_m and conditions for nucleic acid hybridization are well known and can be found in Sambrook, J. et al. (1989) Molecular Cloning: A Laboratory Manual, 2nd ed., vol. 1-3, Cold Spring Harbor Press, Plainview NY; specifically see volume 2, chapter 9.

High stringency conditions for hybridization between polynucleotides of the present invention include wash conditions of 68°C in the presence of about 0.2 x SSC and about 0.1% SDS, for 1 hour. Alternatively, temperatures of about 65°C, 60°C, 55°C, or 42°C may be used. SSC concentration may be varied from about 0.1 to 2 x SSC, with SDS being present at about 0.1%. Typically, blocking reagents are used to block non-specific hybridization. Such blocking reagents include, for instance, sheared and denatured salmon sperm DNA at about 100-200 µg/ml. Organic solvent, such as formamide at a concentration of about 35-50% v/v, may also be used under particular circumstances, such as for RNA:DNA hybridizations. Useful variations on these wash conditions will be readily apparent to those of ordinary skill in the art. Hybridization, particularly under high stringency conditions, may be suggestive of evolutionary similarity between the nucleotides. Such similarity is strongly indicative of a similar role for the nucleotides and their encoded polypeptides.

The term "hybridization complex" refers to a complex formed between two nucleic acid sequences by virtue of the formation of hydrogen bonds between complementary bases. A hybridization complex may be formed in solution (e.g., C_0t or R_0t analysis) or formed between one nucleic acid sequence present in solution and another nucleic acid sequence immobilized on a solid support (e.g., paper, membranes, filters, chips, pins or glass slides, or any other appropriate substrate to which cells or their nucleic acids have been fixed).

The words "insertion" and "addition" refer to changes in an amino acid or nucleotide sequence resulting in the addition of one or more amino acid residues or nucleotides, respectively.

"Immune response" can refer to conditions associated with inflammation, trauma, immune disorders, or infectious or genetic disease, etc. These conditions can be characterized by expression of various factors, e.g., cytokines, chemokines, and other signaling molecules, which may affect cellular and systemic defense systems.

An "immunogenic fragment" is a polypeptide or oligopeptide fragment of APRG which is capable of eliciting an immune response when introduced into a living organism, for example, a mammal. The term "immunogenic fragment" also includes any polypeptide or oligopeptide fragment of APRG which is useful in any of the antibody production methods disclosed herein or known in the art.

The term "microarray" refers to an arrangement of a plurality of polynucleotides, polypeptides, or other chemical compounds on a substrate.

The terms "element" and "array element" refer to a polynucleotide, polypeptide, or other chemical compound having a unique and defined position on a microarray.

The term "modulate" refers to a change in the activity of APRG. For example, modulation may cause an increase or a decrease in protein activity, binding characteristics, or any other biological, functional, or immunological properties of APRG.

The phrases "nucleic acid" and "nucleic acid sequence" refer to a nucleotide, oligonucleotide, polynucleotide, or any fragment thereof. These phrases also refer to DNA or RNA of genomic or synthetic origin which may be single-stranded or double-stranded and may represent the sense or the antisense strand, to peptide nucleic acid (PNA), or to any DNA-like or RNA-like material.

"Operably linked" refers to the situation in which a first nucleic acid sequence is placed in a functional relationship with a second nucleic acid sequence. For instance, a promoter is operably linked to a coding sequence if the promoter affects the transcription or expression of the coding sequence. Operably linked DNA sequences may be in close proximity or contiguous and, where necessary to join two protein coding regions, in the same reading frame.

"Peptide nucleic acid" (PNA) refers to an antisense molecule or anti-gene agent which comprises an oligonucleotide of at least about 5 nucleotides in length linked to a peptide backbone of

amino acid residues ending in lysine. The terminal lysine confers solubility to the composition. PNAs preferentially bind complementary single stranded DNA or RNA and stop transcript elongation, and may be pegylated to extend their lifespan in the cell.

"Post-translational modification" of an APRG may involve lipidation, glycosylation, phosphorylation, acetylation, racemization, proteolytic cleavage, and other modifications known in the art. These processes may occur synthetically or biochemically. Biochemical modifications will vary by cell type depending on the enzymatic milieu of APRG.

"Probe" refers to nucleic acid sequences encoding APRG, their complements, or fragments thereof, which are used to detect identical, allelic or related nucleic acid sequences. Probes are isolated oligonucleotides or polynucleotides attached to a detectable label or reporter molecule. Typical labels include radioactive isotopes, ligands, chemiluminescent agents, and enzymes. "Primers" are short nucleic acids, usually DNA oligonucleotides, which may be annealed to a target polynucleotide by complementary base-pairing. The primer may then be extended along the target DNA strand by a DNA polymerase enzyme. Primer pairs can be used for amplification (and identification) of a nucleic acid sequence, e.g., by the polymerase chain reaction (PCR).

Probes and primers as used in the present invention typically comprise at least 15 contiguous nucleotides of a known sequence. In order to enhance specificity, longer probes and primers may also be employed, such as probes and primers that comprise at least 20, 25, 30, 40, 50, 60, 70, 80, 90, 100, or at least 150 consecutive nucleotides of the disclosed nucleic acid sequences. Probes and primers may be considerably longer than these examples, and it is understood that any length supported by the specification, including the tables, figures, and Sequence Listing, may be used.

Methods for preparing and using probes and primers are described in the references, for example Sambrook, J. et al. (1989) Molecular Cloning: A Laboratory Manual, 2nd ed., vol. 1-3, Cold Spring Harbor Press, Plainview NY; Ausubel, F.M. et al. (1987) Current Protocols in Molecular Biology, Greene Publ. Assoc. & Wiley-Intersciences, New York NY; Innis, M. et al. (1990) PCR Protocols, A Guide to Methods and Applications, Academic Press, San Diego CA. PCR primer pairs can be derived from a known sequence, for example, by using computer programs intended for that purpose such as Primer (Version 0.5, 1991, Whitehead Institute for Biomedical Research, Cambridge MA).

Oligonucleotides for use as primers are selected using software known in the art for such purpose. For example, OLIGO 4.06 software is useful for the selection of PCR primer pairs of up to 100 nucleotides each, and for the analysis of oligonucleotides and larger polynucleotides of up to 5,000 nucleotides from an input polynucleotide sequence of up to 32 kilobases. Similar primer selection programs have incorporated additional features for expanded capabilities. For example, the PrimOU

primer selection program (available to the public from the Genome Center at University of Texas South West Medical Center, Dallas TX) is capable of choosing specific primers from megabase sequences and is thus useful for designing primers on a genome-wide scope. The Primer3 primer selection program (available to the public from the Whitehead Institute/MIT Center for Genome Research, Cambridge MA) allows the user to input a "mispriming library," in which sequences to avoid as primer binding sites are user-specified. Primer3 is useful, in particular, for the selection of oligonucleotides for microarrays. (The source code for the latter two primer selection programs may also be obtained from their respective sources and modified to meet the user's specific needs.) The PrimeGen program (available to the public from the UK Human Genome Mapping Project Resource Centre, Cambridge UK) designs primers based on multiple sequence alignments, thereby allowing selection of primers that hybridize to either the most conserved or least conserved regions of aligned nucleic acid sequences. Hence, this program is useful for identification of both unique and conserved oligonucleotides and polynucleotide fragments. The oligonucleotides and polynucleotide fragments identified by any of the above selection methods are useful in hybridization technologies, for example, as PCR or sequencing primers, microarray elements, or specific probes to identify fully or partially complementary polynucleotides in a sample of nucleic acids. Methods of oligonucleotide selection are not limited to those described above.

A "recombinant nucleic acid" is a sequence that is not naturally occurring or has a sequence that is made by an artificial combination of two or more otherwise separated segments of sequence. This artificial combination is often accomplished by chemical synthesis or, more commonly, by the artificial manipulation of isolated segments of nucleic acids, e.g., by genetic engineering techniques such as those described in Sambrook, supra. The term recombinant includes nucleic acids that have been altered solely by addition, substitution, or deletion of a portion of the nucleic acid. Frequently, a recombinant nucleic acid may include a nucleic acid sequence operably linked to a promoter sequence. Such a recombinant nucleic acid may be part of a vector that is used, for example, to transform a cell.

Alternatively, such recombinant nucleic acids may be part of a viral vector, e.g., based on a vaccinia virus, that could be use to vaccinate a mammal wherein the recombinant nucleic acid is expressed, inducing a protective immunological response in the mammal.

A "regulatory element" refers to a nucleic acid sequence usually derived from untranslated regions of a gene and includes enhancers, promoters, introns, and 5' and 3' untranslated regions (UTRs). Regulatory elements interact with host or viral proteins which control transcription, translation, or RNA stability.

"Reporter molecules" are chemical or biochemical moieties used for labeling a nucleic acid, amino acid, or antibody. Reporter molecules include radionuclides; enzymes; fluorescent,

chemiluminescent, or chromogenic agents; substrates; cofactors; inhibitors; magnetic particles; and other moieties known in the art.

An "RNA equivalent," in reference to a DNA sequence, is composed of the same linear sequence of nucleotides as the reference DNA sequence with the exception that all occurrences of the nitrogenous base thymine are replaced with uracil, and the sugar backbone is composed of ribose instead of deoxyribose.

The term "sample" is used in its broadest sense. A sample suspected of containing APRG, nucleic acids encoding APRG, or fragments thereof may comprise a bodily fluid; an extract from a cell, chromosome, organelle, or membrane isolated from a cell; a cell; genomic DNA, RNA, or cDNA, in solution or bound to a substrate; a tissue; a tissue print; etc.

The terms "specific binding" and "specifically binding" refer to that interaction between a protein or peptide and an agonist, an antibody, an antagonist, a small molecule, or any natural or synthetic binding composition. The interaction is dependent upon the presence of a particular structure of the protein, e.g., the antigenic determinant or epitope, recognized by the binding molecule. For example, if an antibody is specific for epitope "A," the presence of a polypeptide comprising the epitope A, or the presence of free unlabeled A, in a reaction containing free labeled A and the antibody will reduce the amount of labeled A that binds to the antibody.

The term "substantially purified" refers to nucleic acid or amino acid sequences that are removed from their natural environment and are isolated or separated, and are at least 60% free, preferably at least 75% free, and most preferably at least 90% free from other components with which they are naturally associated.

A "substitution" refers to the replacement of one or more amino acid residues or nucleotides by different amino acid residues or nucleotides, respectively.

"Substrate" refers to any suitable rigid or semi-rigid support including membranes, filters, chips, slides, wafers, fibers, magnetic or nonmagnetic beads, gels, tubing, plates, polymers, microparticles and capillaries. The substrate can have a variety of surface forms, such as wells, trenches, pins, channels and pores, to which polynucleotides or polypeptides are bound.

A "transcript image" refers to the collective pattern of gene expression by a particular cell type or tissue under given conditions at a given time.

"Transformation" describes a process by which exogenous DNA is introduced into a recipient cell. Transformation may occur under natural or artificial conditions according to various methods well known in the art, and may rely on any known method for the insertion of foreign nucleic acid sequences into a prokaryotic or eukaryotic host cell. The method for transformation is selected based on the type of host cell being transformed and may include, but is not limited to, bacteriophage or viral infection,

electroporation, heat shock, lipofection, and particle bombardment. The term "transformed cells" includes stably transformed cells in which the inserted DNA is capable of replication either as an autonomously replicating plasmid or as part of the host chromosome, as well as transiently transformed cells which express the inserted DNA or RNA for limited periods of time.

A "transgenic organism," as used herein, is any organism, including but not limited to animals and plants, in which one or more of the cells of the organism contains heterologous nucleic acid introduced by way of human intervention, such as by transgenic techniques well known in the art. The nucleic acid is introduced into the cell, directly or indirectly by introduction into a precursor of the cell, by way of deliberate genetic manipulation, such as by microinjection or by infection with a recombinant virus. The term genetic manipulation does not include classical cross-breeding, or in vitro fertilization, but rather is directed to the introduction of a recombinant DNA molecule. The transgenic organisms contemplated in accordance with the present invention include bacteria, cyanobacteria, fungi, plants and animals. The isolated DNA of the present invention can be introduced into the host by methods known in the art, for example infection, transfection, transformation or transconjugation. Techniques for transferring the DNA of the present invention into such organisms are widely known and provided in references such as Sambrook et al. (1989), supra.

A "variant" of a particular nucleic acid sequence is defined as a nucleic acid sequence having at least 40% sequence identity to the particular nucleic acid sequence over a certain length of one of the nucleic acid sequences using blastn with the "BLAST 2 Sequences" tool Version 2.0.9 (May-07-1999) set at default parameters. Such a pair of nucleic acids may show, for example, at least 50%, at least 60%, at least 70%, at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% or greater sequence identity over a certain defined length. A variant may be described as, for example, an "allelic" (as defined above), "splice," "species," or "polymorphic" variant. A splice variant may have significant identity to a reference molecule, but will generally have a greater or lesser number of polynucleotides due to alternative splicing of exons during mRNA processing. The corresponding polypeptide may possess additional functional domains or lack domains that are present in the reference molecule. Species variants are polynucleotide sequences that vary from one species to another. The resulting polypeptides will generally have significant amino acid identity relative to each other. A polymorphic variant is a variation in the polynucleotide sequence of a particular gene between individuals of a given species. Polymorphic variants also may encompass "single nucleotide polymorphisms" (SNPs) in which the polynucleotide sequence varies by one nucleotide base. The presence of SNPs may be indicative of, for example, a certain population, a disease state, or a propensity for a disease state.

A "variant" of a particular polypeptide sequence is defined as a polypeptide sequence having at least 40% sequence identity to the particular polypeptide sequence over a certain length of one of the polypeptide sequences using blastp with the "BLAST 2 Sequences" tool Version 2.0.9 (May-07-1999) set at default parameters. Such a pair of polypeptides may show, for example, at least 50%, at least 60%, at least 70%, at least 80%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% or greater sequence identity over a certain defined length of one of the polypeptides.

THE INVENTION

The invention is based on the discovery of new human apoptosis regulators (APRG), the polynucleotides encoding APRG, and the use of these compositions for the diagnosis, treatment, or prevention of cell proliferative, immunological, and reproductive disorders.

Table 1 summarizes the nomenclature for the full length polynucleotide and polypeptide sequences of the invention. Each polynucleotide and its corresponding polypeptide are correlated to a single Incyte project identification number (Incyte Project ID). Each polypeptide sequence is denoted by both a polypeptide sequence identification number (Polypeptide SEQ ID NO:) and an Incyte polypeptide sequence number (Incyte Polypeptide ID) as shown. Each polynucleotide sequence is denoted by both a polynucleotide sequence identification number (Polynucleotide SEQ ID NO:) and an Incyte polynucleotide consensus sequence number (Incyte Polynucleotide ID) as shown.

Table 2 shows sequences with homology to the polypeptides of the invention as identified by BLAST analysis against the GenBank protein (genpept) database. Columns 1 and 2 show the polypeptide sequence identification number (Polypeptide SEQ ID NO:) and the corresponding Incyte polypeptide sequence number (Incyte Polypeptide ID) for polypeptides of the invention. Column 3 shows the GenBank identification number (Genbank ID NO:) of the nearest GenBank homolog. Column 4 shows the probability score for the match between each polypeptide and its GenBank homolog. Column 5 shows the annotation of the GenBank homolog along with relevant citations where applicable, all of which are expressly incorporated by reference herein.

Table 3 shows various structural features of the polypeptides of the invention. Columns 1 and 2 show the polypeptide sequence identification number (SEQ ID NO:) and the corresponding Incyte polypeptide sequence number (Incyte Polypeptide ID) for each polypeptide of the invention. Column 3 shows the number of amino acid residues in each polypeptide. Column 4 shows potential phosphorylation sites, and column 5 shows potential glycosylation sites, as determined by the MOTIFS program of the GCG sequence analysis software package (Genetics Computer Group, Madison WI). Column 6 shows amino acid residues comprising signature sequences, domains, and motifs. Column 7

shows analytical methods for protein structure/function analysis and in some cases, searchable databases to which the analytical methods were applied.

Together, Tables 2 and 3 summarize the properties of polypeptides of the invention, and these properties establish that the claimed polypeptides are apoptosis regulators. For example, SEQ ID NO:1 is 51% identical to zebrafish Dedd1, a death effector domain protein (GenBank ID g7673638) as determined by the Basic Local Alignment Search Tool (BLAST). (See Table 2.) The BLAST probability score is 5.1e-70, which indicates the probability of obtaining the observed polypeptide sequence alignment by chance. SEQ ID NO:1 also contains a death effector domain as determined by searching for statistically significant matches in the hidden Markov model (HMM)-based PFAM database of conserved protein family domains. (See Table 3.) Data from BLAST analyses provide further corroborative evidence that SEQ ID NO:1 is a regulator of apoptosis having a death effector domain.

In an alternative example, SEQ ID NO:2 is 36% identical to human caspase recruitment domain protein 7 (g10198209) as determined by the Basic Local Alignment Search Tool (BLAST). (See Table 2.) The BLAST probability score is 1.1e-66, which indicates the probability of obtaining the observed polypeptide sequence alignment by chance.

In an alternative example, SEQ ID NO:3 is 33% identical to human T-cell leukemia virus type-I (HLV-I) Tax-binding protein (TXBP151) (GenBank ID g5776545) as determined by the Basic Local Alignment Search Tool (BLAST). (See Table 2.) The BLAST probability score is 2.1e-73, which indicates the probability of obtaining the observed polypeptide sequence alignment by chance. Tax-binding protein has been shown to inhibit apoptosis induced by tumor necrosis factor in NIH 3T3 cells (De Valck, D. et al. (1999) Oncogene 18:4182-4190). Data from MOTIFS, HMMER-PFAM and BLIMPS analyses provide further corroborative evidence that SEQ ID NO:3 is an apoptosis-related protein. SEQ ID NO:3 contains a RGD motif and a Leucine Zipper motif as determined by the MOTIFS program that searches for patterns that match those defined in Prosite. (See Table 3.)

SEQ ID NO:3 also contains sequence homolgy to the NDP52 nuclear domain protein as determined by the Blocks IMProved Searcher (BLIMPS) that matches a sequence against those in BLOCKS, PRINTS, DOMO, PRODOM, and PFAM databases to search for gene families, sequence homology, and structural fingerprint regions. (See Table 3.) This homology indicates that SEQ ID NO:3 can be localized to the nucleus. Therefore, the protein of SEQ ID NO:3 may be used as a nuclear marker protein. The algorithms and parameters for the analysis of SEQ ID NO:1-3 are described in Table 7.

As shown in Table 4, the full length polynucleotide sequences of the present invention were assembled using cDNA sequences or coding (exon) sequences derived from genomic DNA, or any

combination of these two types of sequences. Columns 1 and 2 list the polynucleotide sequence identification number (Polynucleotide SEQ ID NO:) and the corresponding Incyte polynucleotide consensus sequence number (Incyte Polynucleotide ID) for each polynucleotide of the invention. Column 3 shows the length of each polynucleotide sequence in basepairs. Column 4 lists fragments of the polynucleotide sequences which are useful, for example, in hybridization or amplification technologies that identify SEQ ID NO:4-6 or that distinguish between SEQ ID NO:4-6 and related polynucleotide sequences. Column 5 shows identification numbers corresponding to cDNA sequences, coding sequences (exons) predicted from genomic DNA, and/or sequence assemblages comprised of both cDNA and genomic DNA. These sequences were used to assemble the full length polynucleotide sequences of the invention. Columns 6 and 7 of Table 4 show the nucleotide start (5') and stop (3') positions of the cDNA and/or genomic sequences in column 5 relative to their respective full length sequences.

The identification numbers in Column 5 of Table 4 may refer specifically, for example, to Incyte cDNAs along with their corresponding cDNA libraries. For example, 7018482H1 is the identification number of an Incyte cDNA sequence, and KIDNNOC01 is the cDNA library from which it is derived. Incyte cDNAs for which cDNA libraries are not indicated were derived from pooled cDNA libraries (e.g., 72070238V1). Alternatively, the identification numbers in column 5 may refer to GenBank cDNAs or ESTs which contributed to the assembly of the full length polynucleotide sequences. Alternatively, the identification numbers in column 5 may refer to coding regions predicted by Genscan analysis of genomic DNA. For example, GNN.g8099799_000020_006 is the identification number of a Genscan-predicted coding sequence, with g8099799 being the GenBank identification number of the sequence to which Genscan was applied. The Genscan-predicted coding sequences may have been edited prior to assembly. (See Example IV.) Alternatively, the identification numbers in column 5 may refer to assemblages of both cDNA and Genscan-predicted exons brought together by an "exon stitching" algorithm. (See Example V.) Alternatively, the identification numbers in column 5 may refer to assemblages of both cDNA and Genscan-predicted exons brought together by an "exon-stretching" algorithm. (See Example V.) In some cases, Incyte cDNA coverage redundant with the sequence coverage shown in column 5 was obtained to confirm the final consensus polynucleotide sequence, but the relevant Incyte cDNA identification numbers are not shown.

Table 5 shows the representative cDNA libraries for those full length polynucleotide sequences which were assembled using Incyte cDNA sequences. The representative cDNA library is the Incyte cDNA library which is most frequently represented by the Incyte cDNA sequences which were used to assemble and confirm the above polynucleotide sequences. The tissues and vectors which were used to construct the cDNA libraries shown in Table 5 are described in Table 6.

The invention also encompasses APRG variants. A preferred APRG variant is one which has at least about 80%, or alternatively at least about 90%, or even at least about 95% amino acid sequence identity to the APRG amino acid sequence, and which contains at least one functional or structural characteristic of APRG.

The invention also encompasses polynucleotides which encode APRG. In a particular embodiment, the invention encompasses a polynucleotide sequence comprising a sequence selected from the group consisting of SEQ ID NO:4-6, which encodes APRG. The polynucleotide sequences of SEQ ID NO:4-6, as presented in the Sequence Listing, embrace the equivalent RNA sequences, wherein occurrences of the nitrogenous base thymine are replaced with uracil, and the sugar backbone is composed of ribose instead of deoxyribose.

The invention also encompasses a variant of a polynucleotide sequence encoding APRG. In particular, such a variant polynucleotide sequence will have at least about 70%, or alternatively at least about 85%, or even at least about 95% polynucleotide sequence identity to the polynucleotide sequence encoding APRG. A particular aspect of the invention encompasses a variant of a polynucleotide sequence comprising a sequence selected from the group consisting of SEQ ID NO:4-6 which has at least about 70%, or alternatively at least about 85%, or even at least about 95% polynucleotide sequence identity to a nucleic acid sequence selected from the group consisting of SEQ ID NO:4-6. Any one of the polynucleotide variants described above can encode an amino acid sequence which contains at least one functional or structural characteristic of APRG.

It will be appreciated by those skilled in the art that as a result of the degeneracy of the genetic code, a multitude of polynucleotide sequences encoding APRG, some bearing minimal similarity to the polynucleotide sequences of any known and naturally occurring gene, may be produced. Thus, the invention contemplates each and every possible variation of polynucleotide sequence that could be made by selecting combinations based on possible codon choices. These combinations are made in accordance with the standard triplet genetic code as applied to the polynucleotide sequence of naturally occurring APRG, and all such variations are to be considered as being specifically disclosed.

Although nucleotide sequences which encode APRG and its variants are generally capable of hybridizing to the nucleotide sequence of the naturally occurring APRG under appropriately selected conditions of stringency, it may be advantageous to produce nucleotide sequences encoding APRG or its derivatives possessing a substantially different codon usage, e.g., inclusion of non-naturally occurring codons. Codons may be selected to increase the rate at which expression of the peptide occurs in a particular prokaryotic or eukaryotic host in accordance with the frequency with which particular codons are utilized by the host. Other reasons for substantially altering the nucleotide sequence encoding APRG and its derivatives without altering the encoded amino acid sequences include

the production of RNA transcripts having more desirable properties, such as a greater half-life, than transcripts produced from the naturally occurring sequence.

The invention also encompasses production of DNA sequences which encode APRG and APRG derivatives, or fragments thereof, entirely by synthetic chemistry. After production, the synthetic sequence may be inserted into any of the many available expression vectors and cell systems using reagents well known in the art. Moreover, synthetic chemistry may be used to introduce mutations into a sequence encoding APRG or any fragment thereof.

Also encompassed by the invention are polynucleotide sequences that are capable of hybridizing to the claimed polynucleotide sequences, and, in particular, to those shown in SEQ ID NO:4-6 and fragments thereof under various conditions of stringency. (See, e.g., Wahl, G.M. and S.L. Berger (1987) Methods Enzymol. 152:399-407; Kimmel, A.R. (1987) Methods Enzymol. 152:507-511.) Hybridization conditions, including annealing and wash conditions, are described in "Definitions."

Methods for DNA sequencing are well known in the art and may be used to practice any of the embodiments of the invention. The methods may employ such enzymes as the Klenow fragment of DNA polymerase I, SEQUENASE (US Biochemical, Cleveland OH), Taq polymerase (Applied Biosystems), thermostable T7 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 MICROLAB 2200 liquid transfer system (Hamilton, Reno NV), PTC200 thermal cycler (MJ Research, Watertown MA) and ABI CATALYST 800 thermal cycler (Applied Biosystems). Sequencing is then carried out using either the ABI 373 or 377 DNA sequencing system (Applied Biosystems), the MEGABACE 1000 DNA sequencing system (Molecular Dynamics, Sunnyvale CA), or other systems known in the art. The resulting sequences are analyzed using a variety of algorithms which are well known in the art. (See, e.g., Ausubel, F.M. (1997) Short Protocols in Molecular Biology, John Wiley & Sons, New York NY, unit 7.7; Meyers, R.A. (1995) Molecular Biology and Biotechnology, Wiley VCH, New York NY, pp. 856-853.)

The nucleic acid sequences encoding APRG may be extended utilizing a partial nucleotide sequence and employing various PCR-based methods known in the art to detect upstream sequences, such as promoters and regulatory elements. For example, one method which may be employed, restriction-site PCR, uses universal and nested primers to amplify unknown sequence from genomic DNA within a cloning vector. (See, e.g., Sarkar, G. (1993) PCR Methods Applic. 2:318-322.) Another method, inverse PCR, uses primers that extend in divergent directions to amplify unknown sequence from a circularized template. The template is derived from restriction fragments comprising a

known genomic locus and surrounding sequences. (See, e.g., Triglia, T. et al. (1988) Nucleic Acids Res. 16:8186.) A third method, capture PCR, involves PCR amplification of DNA fragments adjacent to known sequences in human and yeast artificial chromosome DNA. (See, e.g., Lagerstrom, M. et al. (1991) PCR Methods Applic. 1:111-119.) In this method, multiple restriction enzyme digestions and ligations may be used to insert an engineered double-stranded sequence into a region of unknown sequence before performing PCR. Other methods which may be used to retrieve unknown sequences are known in the art. (See, e.g., Parker, J.D. et al. (1991) Nucleic Acids Res. 19:3055-3060). Additionally, one may use PCR, nested primers, and PROMOTERFINDER libraries (Clontech, Palo Alto CA) to walk genomic DNA. This procedure avoids the need to screen libraries and is useful in finding intron/exon junctions. 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) or another appropriate program, to be about 22 to 30 nucleotides in length, to have a GC content of about 50% or more, and to anneal to the template at temperatures of about 68°C to 72°C.

When screening for full length cDNAs, it is preferable to use libraries that have been size-selected to include larger cDNAs. In addition, random-primed libraries, which often include sequences containing the 5' regions of genes, are preferable for situations in which an oligo d(T) library does not yield a full-length cDNA. Genomic libraries may be useful for extension of sequence into 5' non-transcribed regulatory regions.

Capillary electrophoresis systems which are commercially available may be used to analyze the size or confirm the nucleotide sequence of sequencing or PCR products. In particular, capillary sequencing may employ flowable polymers for electrophoretic separation, four different nucleotide-specific, laser-stimulated fluorescent dyes, and a charge coupled device camera for detection of the emitted wavelengths. Output/light intensity may be converted to electrical signal using appropriate software (e.g., GENOTYPER and SEQUENCE NAVIGATOR, Applied Biosystems), and the entire process from loading of samples to computer analysis and electronic data display may be computer controlled. Capillary electrophoresis is especially preferable for sequencing small DNA fragments which may be present in limited amounts in a particular sample.

In another embodiment of the invention, polynucleotide sequences or fragments thereof which encode APRG may be cloned in recombinant DNA molecules that direct expression of APRG, or fragments or functional equivalents thereof, in appropriate host cells. Due to the inherent degeneracy of the genetic code, other DNA sequences which encode substantially the same or a functionally equivalent amino acid sequence may be produced and used to express APRG.

The nucleotide sequences of the present invention can be engineered using methods generally

known in the art in order to alter APRG-encoding sequences for a variety of purposes including, but not limited to, modification of the cloning, processing, and/or expression of the gene product. DNA shuffling by random fragmentation and PCR reassembly of gene fragments and synthetic oligonucleotides may be used to engineer the nucleotide sequences. For example, oligonucleotide-mediated site-directed mutagenesis may be used to introduce mutations that create new restriction sites, alter glycosylation patterns, change codon preference, produce splice variants, and so forth.

The nucleotides of the present invention may be subjected to DNA shuffling techniques such as MOLECULARBREEDING (Maxygen Inc., Santa Clara CA; described in U.S. Patent Number 5,837,458; Chang, C.-C. et al. (1999) Nat. Biotechnol. 17:793-797; Christians, F.C. et al. (1999) Nat. Biotechnol. 17:259-264; and Crameri, A. et al. (1996) Nat. Biotechnol. 14:315-319) to alter or improve the biological properties of APRG, such as its biological or enzymatic activity or its ability to bind to other molecules or compounds. DNA shuffling is a process by which a library of gene variants is produced using PCR-mediated recombination of gene fragments. The library is then subjected to selection or screening procedures that identify those gene variants with the desired properties. These preferred variants may then be pooled and further subjected to recursive rounds of DNA shuffling and selection/screening. Thus, genetic diversity is created through "artificial" breeding and rapid molecular evolution. For example, fragments of a single gene containing random point mutations may be recombined, screened, and then reshuffled until the desired properties are optimized. Alternatively, fragments of a given gene may be recombined with fragments of homologous genes in the same gene family, either from the same or different species, thereby maximizing the genetic diversity of multiple naturally occurring genes in a directed and controllable manner.

In another embodiment, sequences encoding APRG may be synthesized, in whole or in part, using chemical methods well known in the art. (See, e.g., Caruthers, M.H. et al. (1980) Nucleic Acids Symp. Ser. 7:215-223; and Horn, T. et al. (1980) Nucleic Acids Symp. Ser. 7:225-232.) Alternatively, APRG itself or a fragment thereof may be synthesized using chemical methods. For example, peptide synthesis can be performed using various solution-phase or solid-phase techniques. (See, e.g., Creighton, T. (1984) Proteins, Structures and Molecular Properties, WH Freeman, New York NY, pp. 55-60; and Roberge, J.Y. et al. (1995) Science 269:202-204.) Automated synthesis may be achieved using the ABI 431A peptide synthesizer (Applied Biosystems). Additionally, the amino acid sequence of APRG, or any part thereof, may be altered during direct synthesis and/or combined with sequences from other proteins, or any part thereof, to produce a variant polypeptide or a polypeptide having a sequence of a naturally occurring polypeptide.

The peptide may be substantially purified by preparative high performance liquid

chromatography. (See, e.g., Chlez, R.M. and F.Z. Regnier (1990) Methods Enzymol. 182:392-421.) The composition of the synthetic peptides may be confirmed by amino acid analysis or by sequencing. (See, e.g., Creighton, supra, pp. 28-53.)

In order to express a biologically active APRG, the nucleotide sequences encoding APRG or derivatives thereof may be inserted into an appropriate expression vector, i.e., a vector which contains the necessary elements for transcriptional and translational control of the inserted coding sequence in a suitable host. These elements include regulatory sequences, such as enhancers, constitutive and inducible promoters, and 5' and 3' untranslated regions in the vector and in polynucleotide sequences encoding APRG. Such elements may vary in their strength and specificity. Specific initiation signals may also be used to achieve more efficient translation of sequences encoding APRG. Such signals include the ATG initiation codon and adjacent sequences, e.g. the Kozak sequence. In cases where sequences encoding APRG and its initiation codon and upstream regulatory sequences are inserted into the appropriate expression vector, no additional transcriptional or translational control signals may be needed. However, in cases where only coding sequence, or a fragment thereof, is inserted, exogenous translational control signals including an in-frame ATG initiation codon should be provided by the vector. Exogenous translational elements and initiation codons may be of various origins, both natural and synthetic. The efficiency of expression may be enhanced by the inclusion of enhancers appropriate for the particular host cell system used. (See, e.g., Scharf, D. et al. (1994) Results Probl. Cell Differ. 20:125-162.)

Methods which are well known to those skilled in the art may be used to construct expression vectors containing sequences encoding APRG and appropriate transcriptional and translational control elements. These methods include in vitro recombinant DNA techniques, synthetic techniques, and in vivo genetic recombination. (See, e.g., Sambrook, J. et al. (1989) Molecular Cloning, A Laboratory Manual, Cold Spring Harbor Press, Plainview NY, ch. 4, 8, and 16-17; Ausubel, F.M. et al. (1995) Current Protocols in Molecular Biology, John Wiley & Sons, New York NY, ch. 9, 13, and 16.)

A variety of expression vector/host systems may be utilized to contain and express sequences encoding APRG. These include, but are not limited to, microorganisms such as bacteria transformed with recombinant bacteriophage, plasmid, or cosmid DNA expression vectors; yeast transformed with yeast expression vectors; insect cell systems infected with viral expression vectors (e.g., baculovirus); plant cell systems transformed with viral expression vectors (e.g., cauliflower mosaic virus, CaMV, or tobacco mosaic virus, TMV) or with bacterial expression vectors (e.g., Ti or pBR322 plasmids); or animal cell systems. (See, e.g., Sambrook, supra; Ausubel, supra; Van Heeke, G. and S.M. Schuster (1989) J. Biol. Chem. 264:5503-5509; Engelhard, E.K. et al. (1994) Proc. Natl. Acad. Sci. USA 91:3224-3227; Sandig, V. et al. (1996) Hum. Gene Ther. 7:1937-1945; Takamatsu, N. (1987) EMBO

J. 6:307-311; The McGraw Hill Yearbook of Science and Technology (1992) McGraw Hill, New York NY, pp. 191-196; Logan, J. and T. Shenk (1984) Proc. Natl. Acad. Sci. USA 81:3655-3659; and Harrington, J.J. et al. (1997) Nat. Genet. 15:345-355.) Expression vectors derived from retroviruses, adenoviruses, or herpes or vaccinia viruses, or from various bacterial plasmids, may be used for delivery of nucleotide sequences to the targeted organ, tissue, or cell population. (See, e.g., Di Nicola, M. et al. (1998) Cancer Gen. Ther. 5(6):350-356; Yu, M. et al. (1993) Proc. Natl. Acad. Sci. USA 90(13):6340-6344; Buller, R.M. et al. (1985) Nature 317(6040):813-815; McGregor, D.P. et al. (1994) Mol. Immunol. 31(3):219-226; and Verma, I.M. and N. Somia (1997) Nature 389:239-242.) The invention is not limited by the host cell employed.

In bacterial systems, a number of cloning and expression vectors may be selected depending upon the use intended for polynucleotide sequences encoding APRG. For example, routine cloning, subcloning, and propagation of polynucleotide sequences encoding APRG can be achieved using a multifunctional E. coli vector such as PBLUESCRIPT (Stratagene, La Jolla CA) or PSPORT1 plasmid (Life Technologies). Ligation of sequences encoding APRG into the vector's multiple cloning site disrupts the *lacZ* gene, allowing a colorimetric screening procedure for identification of transformed bacteria containing recombinant molecules. 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. (See, e.g., Van Heeke, G. and S.M. Schuster (1989) J. Biol. Chem. 264:5503-5509.) When large quantities of APRG are needed, e.g. for the production of antibodies, vectors which direct high level expression of APRG may be used. For example, vectors containing the strong, inducible SP6 or T7 bacteriophage promoter may be used.

Yeast expression systems may be used for production of APRG. A number of vectors containing constitutive or inducible promoters, such as alpha factor, alcohol oxidase, and PGH promoters, may be used in the yeast Saccharomyces cerevisiae or Pichia pastoris. In addition, such vectors direct either the secretion or intracellular retention of expressed proteins and enable integration of foreign sequences into the host genome for stable propagation. (See, e.g., Ausubel, 1995, supra; Bitter, G.A. et al. (1987) Methods Enzymol. 153:516-544; and Scorer, C.A. et al. (1994) Bio/Technology 12:181-184.)

Plant systems may also be used for expression of APRG. Transcription of sequences encoding APRG may be driven by viral promoters, e.g., the 35S and 19S promoters of CaMV used alone or in combination with the omega leader sequence from TMV (Takamatsu, N. (1987) EMBO J. 6:307-311). Alternatively, plant promoters such as the small subunit of RUBISCO or heat shock promoters may be used. (See, e.g., Coruzzi, G. et al. (1984) EMBO J. 3:1671-1680; Broglie, R. et al. (1984) Science 224:838-843; and Winter, J. et al. (1991) Results Probl. Cell Differ. 17:85-105.) These constructs can

be introduced into plant cells by direct DNA transformation or pathogen-mediated transfection. (See, e.g., The McGraw Hill Yearbook of Science and Technology (1992) McGraw Hill, New York NY, pp. 191-196.)

In mammalian cells, a number of viral-based expression systems may be utilized. In cases where an adenovirus is used as an expression vector, sequences encoding APRG may be ligated into an adenovirus transcription/translation complex consisting of the late promoter and tripartite leader sequence. Insertion in a non-essential E1 or E3 region of the viral genome may be used to obtain infective virus which expresses APRG in host cells. (See, e.g., Logan, J. and T. Shenk (1984) Proc. Natl. Acad. Sci. USA 81:3655-3659.) In addition, transcription enhancers, such as the Rous sarcoma virus (RSV) enhancer, may be used to increase expression in mammalian host cells. SV40 or EBV-based vectors may also be used for high-level protein expression.

Human artificial chromosomes (HACs) may also be employed to deliver larger fragments of DNA than can be contained in and expressed from a plasmid. HACs of about 6 kb to 10 Mb are constructed and delivered via conventional delivery methods (liposomes, polycationic amino polymers, or vesicles) for therapeutic purposes. (See, e.g., Harrington, J.J. et al. (1997) Nat. Genet. 15:345-355.)

For long term production of recombinant proteins in mammalian systems, stable expression of APRG in cell lines is preferred. For example, sequences encoding APRG can be transformed into cell lines using expression vectors which may contain viral origins of replication and/or endogenous expression elements and a selectable marker gene on the same or on a separate vector. Following the introduction of the vector, cells may be allowed to grow for about 1 to 2 days in enriched media before being switched to selective media. The purpose of the selectable marker is to confer resistance to a selective agent, and its presence allows growth and recovery of cells which successfully express the introduced sequences. Resistant clones of stably transformed cells may be propagated using tissue culture techniques appropriate to the cell type.

Any number of selection systems may be used to recover transformed cell lines. These include, but are not limited to, the herpes simplex virus thymidine kinase and adenine phosphoribosyltransferase genes, for use in *tk* and *aprt* cells, respectively. (See, e.g., Wigler, M. et al. (1977) Cell 11:223-232; Lowy, I. et al. (1980) Cell 22:817-823.) Also, antimetabolite, antibiotic, or herbicide resistance can be used as the basis for selection. For example, *dhfr* confers resistance to methotrexate; *neo* confers resistance to the aminoglycosides neomycin and G-418; and *als* and *pat* confer resistance to chlorsulfuron and phosphinotricin acetyltransferase, respectively. (See, e.g., Wigler, M. et al. (1980) Proc. Natl. Acad. Sci. USA 77:3567-3570; Colbere-Garapin, F. et al. (1981) J. Mol. Biol. 150:1-14.) Additional selectable genes have been described, e.g., *trpB* and *hisD*, which alter cellular requirements for metabolites. (See, e.g., Hartman, S.C. and R.C. Mulligan (1988) Proc. Natl. Acad. Sci. USA

85:8047-8051.) Visible markers, e.g., anthocyanins, green fluorescent proteins (GFP; Clontech), β glucuronidase and its substrate β-glucuronide, or luciferase and its substrate luciferin may be used. These markers can be used not only to identify transformants, but also to quantify the amount of transient or stable protein expression attributable to a specific vector system. (See, e.g., Rhodes, C.A. (1995) Methods Mol. Biol. 55:121-131.)

Although the presence/absence of marker gene expression suggests that the gene of interest is also present, the presence and expression of the gene may need to be confirmed. For example, if the sequence encoding APRG is inserted within a marker gene sequence, transformed cells containing sequences encoding APRG can be identified by the absence of marker gene function. Alternatively, a marker gene can be placed in tandem with a sequence encoding APRG under the control of a single promoter. Expression of the marker gene in response to induction or selection usually indicates expression of the tandem gene as well.

In general, host cells that contain the nucleic acid sequence encoding APRG and that express APRG may be identified by a variety of procedures known to those of skill in the art. These procedures include, but are not limited to, DNA-DNA or DNA-RNA hybridizations, PCR amplification, and protein bioassay or immunoassay techniques which include membrane, solution, or chip based technologies for the detection and/or quantification of nucleic acid or protein sequences.

Immunological methods for detecting and measuring the expression of APRG 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 on APRG is preferred, but a competitive binding assay may be employed. These and other assays are well known in the art. (See, e.g., Hampton, R. et al. (1990) Serological Methods, a Laboratory Manual, APS Press, St. Paul MN, Sect. IV; Coligan, J.E. et al. (1997) Current Protocols in Immunology, Greene Pub. Associates and Wiley-Interscience, New York NY; and Pound, J.D. (1998) Immunochemical Protocols, Humana Press, Totowa NJ.)

A wide variety of labels and conjugation techniques are known by those skilled in the art and may be used in various nucleic acid and amino acid assays. Means for producing labeled hybridization or PCR probes for detecting sequences related to polynucleotides encoding APRG include oligolabeling, nick translation, end-labeling, or PCR amplification using a labeled nucleotide. Alternatively, the sequences encoding APRG, or any fragments thereof, may be cloned into a vector for the production of an mRNA probe. Such vectors are known in the art, are commercially available, and may be used to synthesize RNA probes in vitro by addition of an appropriate RNA polymerase such as T7, T3, or SP6 and labeled nucleotides. These procedures may be conducted using a variety of commercially available

kits, such as those provided by Amersham Pharmacia Biotech, Promega (Madison WI), and US Biochemical. Suitable reporter molecules or labels which may be used for ease of detection include radionuclides, enzymes, fluorescent, chemiluminescent, or chromogenic agents, as well as substrates, cofactors, inhibitors, magnetic particles, and the like.

Host cells transformed with nucleotide sequences encoding APRG may be cultured under conditions suitable for the expression and recovery of the protein from cell culture. The protein produced by a transformed cell may be secreted or retained intracellularly depending on the sequence and/or the vector used. As will be understood by those of skill in the art, expression vectors containing polynucleotides which encode APRG may be designed to contain signal sequences which direct secretion of APRG through a prokaryotic or eukaryotic cell membrane.

In addition, a host cell strain may be chosen for its ability to modulate expression of the inserted sequences or to process the expressed protein in the desired fashion. Such modifications of the polypeptide include, but are not limited to, acetylation, carboxylation, glycosylation, phosphorylation, lipidation, and acylation. Post-translational processing which cleaves a "prepro" or "pro" form of the protein may also be used to specify protein targeting, folding, and/or activity. Different host cells which have specific cellular machinery and characteristic mechanisms for post-translational activities (e.g., CHO, HeLa, MDCK, HEK293, and WI38) are available from the American Type Culture Collection (ATCC, Manassas VA) and may be chosen to ensure the correct modification and processing of the foreign protein.

In another embodiment of the invention, natural, modified, or recombinant nucleic acid sequences encoding APRG may be ligated to a heterologous sequence resulting in translation of a fusion protein in any of the aforementioned host systems. For example, a chimeric APRG protein containing a heterologous moiety that can be recognized by a commercially available antibody may facilitate the screening of peptide libraries for inhibitors of APRG activity. Heterologous protein and peptide moieties may also facilitate purification of fusion proteins using commercially available affinity matrices. Such moieties include, but are not limited to, glutathione S-transferase (GST), maltose binding protein (MBP), thioredoxin (Trx), calmodulin binding peptide (CBP), 6-His, FLAG, *c-myc*, and hemagglutinin (HA). GST, MBP, Trx, CBP, and 6-His enable purification of their cognate fusion proteins on immobilized glutathione, maltose, phenylarsine oxide, calmodulin, and metal-chelate resins, respectively. FLAG, *c-myc*, and hemagglutinin (HA) enable immunoaffinity purification of fusion proteins using commercially available monoclonal and polyclonal antibodies that specifically recognize these epitope tags. A fusion protein may also be engineered to contain a proteolytic cleavage site located between the APRG encoding sequence and the heterologous protein sequence, so that APRG may be cleaved away from the heterologous moiety following purification. Methods for fusion protein

expression and purification are discussed in Ausubel (1995, supra, ch. 10). A variety of commercially available kits may also be used to facilitate expression and purification of fusion proteins.

In a further embodiment of the invention, synthesis of radiolabeled APRG may be achieved in vitro using the TNT rabbit reticulocyte lysate or wheat germ extract system (Promega). These systems couple transcription and translation of protein-coding sequences operably associated with the T7, T3, or SP6 promoters. Translation takes place in the presence of a radiolabeled amino acid precursor, for example, ^{35}S-methionine.

APRG of the present invention or fragments thereof may be used to screen for compounds that specifically bind to APRG. At least one and up to a plurality of test compounds may be screened for specific binding to APRG. Examples of test compounds include antibodies, oligonucleotides, proteins (e.g., receptors), or small molecules.

In one embodiment, the compound thus identified is closely related to the natural ligand of APRG, e.g., a ligand or fragment thereof, a natural substrate, a structural or functional mimetic, or a natural binding partner. (See, e.g., Coligan, J.E. et al. (1991) Current Protocols in Immunology 1(2): Chapter 5.) Similarly, the compound can be closely related to the natural receptor to which APRG binds, or to at least a fragment of the receptor, e.g., the ligand binding site. In either case, the compound can be rationally designed using known techniques. In one embodiment, screening for these compounds involves producing appropriate cells which express APRG, either as a secreted protein or on the cell membrane. Preferred cells include cells from mammals, yeast, Drosophila, or E. coli. Cells expressing APRG or cell membrane fractions which contain APRG are then contacted with a test compound and binding, stimulation, or inhibition of activity of either APRG or the compound is analyzed.

An assay may simply test binding of a test compound to the polypeptide, wherein binding is detected by a fluorophore, radioisotope, enzyme conjugate, or other detectable label. For example, the assay may comprise the steps of combining at least one test compound with APRG, either in solution or affixed to a solid support, and detecting the binding of APRG to the compound. Alternatively, the assay may detect or measure binding of a test compound in the presence of a labeled competitor. Additionally, the assay may be carried out using cell-free preparations, chemical libraries, or natural product mixtures, and the test compound(s) may be free in solution or affixed to a solid support.

APRG of the present invention or fragments thereof may be used to screen for compounds that modulate the activity of APRG. Such compounds may include agonists, antagonists, or partial or inverse agonists. In one embodiment, an assay is performed under conditions permissive for APRG activity, wherein APRG is combined with at least one test compound, and the activity of APRG in the presence of a test compound is compared with the activity of APRG in the absence of the test

compound. A change in the activity of APRG in the presence of the test compound is indicative of a compound that modulates the activity of APRG. Alternatively, a test compound is combined with an in vitro or cell-free system comprising APRG under conditions suitable for APRG activity, and the assay is performed. In either of these assays, a test compound which modulates the activity of APRG may do so indirectly and need not come in direct contact with the test compound. At least one and up to a plurality of test compounds may be screened.

In another embodiment, polynucleotides encoding APRG or their mammalian homologs may be "knocked out" in an animal model system using homologous recombination in embryonic stem (ES) cells. Such techniques are well known in the art and are useful for the generation of animal models of human disease. (See, e.g., U.S. Patent Number 5,175,383 and U.S. Patent Number 5,767,337.) For example, mouse ES cells, such as the mouse 129/SvJ cell line, are derived from the early mouse embryo and grown in culture. The ES cells are transformed with a vector containing the gene of interest disrupted by a marker gene, e.g., the neomycin phosphotransferase gene (neo; Capecchi, M.R. (1989) Science 244:1288-1292). The vector integrates into the corresponding region of the host genome by homologous recombination. Alternatively, homologous recombination takes place using the Cre-loxP system to knockout a gene of interest in a tissue- or developmental stage-specific manner (Marth, J.D. (1996) Clin. Invest. 97:1999-2002; Wagner, K.U. et al. (1997) Nucleic Acids Res. 25:4323-4330). Transformed ES cells are identified and 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. Transgenic animals thus generated may be tested with potential therapeutic or toxic agents.

Polynucleotides encoding APRG may also be manipulated in vitro in ES cells derived from human blastocysts. Human ES cells have the potential to differentiate into at least eight separate cell lineages including endoderm, mesoderm, and ectodermal cell types. These cell lineages differentiate into, for example, neural cells, hematopoietic lineages, and cardiomyocytes (Thomson, J.A. et al. (1998) Science 282:1145-1147).

Polynucleotides encoding APRG can also be used to create "knockin" humanized animals (pigs) or transgenic animals (mice or rats) to model human disease. With knockin technology, a region of a polynucleotide encoding APRG is injected into animal ES cells, and the injected sequence integrates into the animal cell genome. Transformed cells are injected into blastulae, and the blastulae are implanted as described above. Transgenic progeny or inbred lines are studied and treated with potential pharmaceutical agents to obtain information on treatment of a human disease. Alternatively, a mammal inbred to overexpress APRG, e.g., by secreting APRG in its milk, may also serve as a convenient source of that protein (Janne, J. et al. (1998) Biotechnol. Annu. Rev. 4:55-74).

THERAPEUTICS

Chemical and structural similarity, e.g., in the context of sequences and motifs, exists between regions of APRG and apoptosis regulators. In addition, the expression of APRG is closely associated with cancerous tissue. Therefore, APRG appears to play a role in cell proliferative, immunological, and reproductive disorders. In the treatment of disorders associated with increased APRG expression or activity, it is desirable to decrease the expression or activity of APRG. In the treatment of disorders associated with decreased APRG expression or activity, it is desirable to increase the expression or activity of APRG.

Therefore, in one embodiment, APRG or a fragment or derivative thereof may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of APRG. Examples of such disorders include, but are not limited to, a cell proliferative disorder such as actinic keratosis, arteriosclerosis, atherosclerosis, bursitis, cirrhosis, hepatitis, mixed connective tissue disease (MCTD), myelofibrosis, paroxysmal nocturnal hemoglobinuria, polycythemia vera, psoriasis, primary thrombocythemia, and cancers including adenocarcinoma, leukemia, lymphoma, melanoma, myeloma, sarcoma, teratocarcinoma, and, in particular, a cancer of the adrenal gland, bladder, bone, bone marrow, brain, breast, cervix, gall bladder, ganglia, gastrointestinal tract, heart, kidney, liver, lung, muscle, ovary, pancreas, parathyroid, penis, prostate, salivary glands, skin, spleen, testis, thymus, thyroid, and uterus; an immunological disorder such as acquired immunodeficiency syndrome (AIDS), Addison's disease, adult respiratory distress syndrome, allergies, ankylosing spondylitis, amyloidosis, anemia, asthma, atherosclerosis, autoimmune hemolytic anemia, autoimmune thyroiditis, autoimmune polyendocrinopathy-candidiasis-ectodermal dystrophy (APECED), bronchitis, cholecystitis, contact dermatitis, Crohn's disease, atopic dermatitis, dermatomyositis, diabetes mellitus, emphysema, episodic lymphopenia with lymphocytotoxins, erythroblastosis fetalis, erythema nodosum, atrophic gastritis, glomerulonephritis, Goodpasture's syndrome, gout, Graves' disease, Hashimoto's thyroiditis, hypereosinophilia, irritable bowel syndrome, multiple sclerosis, myasthenia gravis, myocardial or pericardial inflammation, osteoarthritis, osteoporosis, pancreatitis, polymyositis, psoriasis, Reiter's syndrome, rheumatoid arthritis, scleroderma, Sjögren's syndrome, systemic anaphylaxis, systemic lupus erythematosus, systemic sclerosis, thrombocytopenic purpura, ulcerative colitis, uveitis, Werner syndrome, a complication of cancer, hemodialysis, and extracorporeal circulation, viral, bacterial, fungal, parasitic, protozoal, and helminthic infections, and trauma; and a reproductive disorder such as a disorder of prolactin production, infertility, including tubal disease, ovulatory defects, and endometriosis, a disruption of the estrous cycle, a disruption of the menstrual cycle, polycystic ovary syndrome, ovarian hyperstimulation syndrome, endometrial and ovarian tumors, uterine fibroids, autoimmune disorders, ectopic pregnancies, and teratogenesis; cancer of the breast, fibrocystic breast

disease, and galactorrhea; disruptions of spermatogenesis, abnormal sperm physiology, cancer of the testis, cancer of the prostate, benign prostatic hyperplasia, prostatitis, Peyronie's disease, impotence, carcinoma of the male breast, and gynecomastia.

In another embodiment, a vector capable of expressing APRG or a fragment or derivative thereof may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of APRG including, but not limited to, those described above.

In a further embodiment, a composition comprising a substantially purified APRG in conjunction with a suitable pharmaceutical carrier may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of APRG including, but not limited to, those provided above.

In still another embodiment, an agonist which modulates the activity of APRG may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of APRG including, but not limited to, those listed above.

In a further embodiment, an antagonist of APRG may be administered to a subject to treat or prevent a disorder associated with increased expression or activity of APRG. Examples of such disorders include, but are not limited to, those cell proliferative, immunological, and reproductive disorders described above. In one aspect, an antibody which specifically binds APRG may be used directly as an antagonist or indirectly as a targeting or delivery mechanism for bringing a pharmaceutical agent to cells or tissues which express APRG.

In an additional embodiment, a vector expressing the complement of the polynucleotide encoding APRG may be administered to a subject to treat or prevent a disorder associated with increased expression or activity of APRG including, but not limited to, those described above.

In other embodiments, any of the proteins, antagonists, antibodies, agonists, complementary sequences, or vectors of the invention may be administered in combination with other appropriate therapeutic agents. Selection of the appropriate agents for use in combination therapy may be made by one of ordinary skill in the art, according to conventional pharmaceutical principles. The combination of therapeutic agents may act synergistically to effect the treatment or prevention of the various disorders described above. Using this approach, one may be able to achieve therapeutic efficacy with lower dosages of each agent, thus reducing the potential for adverse side effects.

An antagonist of APRG may be produced using methods which are generally known in the art. In particular, purified APRG may be used to produce antibodies or to screen libraries of pharmaceutical agents to identify those which specifically bind APRG. Antibodies to APRG may also be generated using methods that are well known in the art. Such antibodies may include, but are not limited to, polyclonal, monoclonal, chimeric, and single chain antibodies, Fab fragments, and fragments produced

by a Fab expression library. Neutralizing antibodies (i.e., those which inhibit dimer formation) are generally preferred for therapeutic use.

For the production of antibodies, various hosts including goats, rabbits, rats, mice, humans, and others may be immunized by injection with APRG or with any fragment or oligopeptide thereof which has immunogenic properties. Depending on the host species, various adjuvants may be used to increase immunological response. Such adjuvants include, but are not limited to, Freund's, mineral gels such as aluminum hydroxide, and surface active substances such as lysolecithin, pluronic polyols, polyanions, peptides, oil emulsions, KLH, and dinitrophenol. Among adjuvants used in humans, BCG (bacilli Calmette-Guerin) and Corynebacterium parvum are especially preferable.

It is preferred that the oligopeptides, peptides, or fragments used to induce antibodies to APRG have an amino acid sequence consisting of at least about 5 amino acids, and generally will consist of at least about 10 amino acids. It is also preferable that these oligopeptides, peptides, or fragments are identical to a portion of the amino acid sequence of the natural protein. Short stretches of APRG amino acids may be fused with those of another protein, such as KLH, and antibodies to the chimeric molecule may be produced.

Monoclonal antibodies to APRG may be prepared using any technique which provides for the production of antibody molecules 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, G. et al. (1975) Nature 256:495-497; Kozbor, D. et al. (1985) J. Immunol. Methods 81:31-42; Cote, R.J. et al. (1983) Proc. Natl. Acad. Sci. USA 80:2026-2030; and Cole, S.P. et al. (1984) Mol. Cell Biol. 62:109-120.)

In addition, techniques developed for the production of "chimeric antibodies," such as the splicing of mouse antibody genes to human antibody genes to obtain a molecule with appropriate antigen specificity and biological activity, can be used. (See, e.g., Morrison, S.L. et al. (1984) Proc. Natl. Acad. Sci. USA 81:6851-6855; Neuberger, M.S. et al. (1984) Nature 312:604-608; and Takeda, S. et al. (1985) Nature 314:452-454.) Alternatively, techniques described for the production of single chain antibodies may be adapted, using methods known in the art, to produce APRG-specific single chain antibodies. Antibodies with related specificity, but of distinct idiotypic composition, may be generated by chain shuffling from random combinatorial immunoglobulin libraries. (See, e.g., Burton, D.R. (1991) Proc. Natl. Acad. Sci. USA 88:10134-10137.)

Antibodies may also be produced by inducing in vivo production in the lymphocyte population or by screening immunoglobulin libraries or panels of highly specific binding reagents as disclosed in the literature. (See, e.g., Orlandi, R. et al. (1989) Proc. Natl. Acad. Sci. USA 86:3833-3837; Winter, G. et al. (1991) Nature 349:293-299.)

Antibody fragments which contain specific binding sites for APRG may also be generated. For example, such fragments include, but are not limited to, $F(ab')_2$ fragments produced by pepsin digestion of the antibody molecule and Fab fragments generated by reducing the disulfide bridges of the F(ab')2 fragments. Alternatively, Fab expression libraries may be constructed to allow rapid and easy identification of monoclonal Fab fragments with the desired specificity. (See, e.g., Huse, W.D. et al. (1989) Science 246:1275-1281.)

Various immunoassays may be used for screening to identify antibodies having the desired specificity. Numerous protocols for competitive binding or immunoradiometric assays 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 APRG and its specific antibody. A two-site, monoclonal-based immunoassay utilizing monoclonal antibodies reactive to two non-interfering APRG epitopes is generally used, but a competitive binding assay may also be employed (Pound, supra).

Various methods such as Scatchard analysis in conjunction with radioimmunoassay techniques may be used to assess the affinity of antibodies for APRG. Affinity is expressed as an association constant, K_a, which is defined as the molar concentration of APRG-antibody complex divided by the molar concentrations of free antigen and free antibody under equilibrium conditions. The K_a determined for a preparation of polyclonal antibodies, which are heterogeneous in their affinities for multiple APRG epitopes, represents the average affinity, or avidity, of the antibodies for APRG. The K_a determined for a preparation of monoclonal antibodies, which are monospecific for a particular APRG epitope, represents a true measure of affinity. High-affinity antibody preparations with K_a ranging from about 10^9 to 10^{12} L/mole are preferred for use in immunoassays in which the APRG-antibody complex must withstand rigorous manipulations. Low-affinity antibody preparations with K_a ranging from about 10^6 to 10^7 L/mole are preferred for use in immunopurification and similar procedures which ultimately require dissociation of APRG, preferably in active form, from the antibody (Catty, D. (1988) Antibodies, Volume I: A Practical Approach, IRL Press, Washington DC; Liddell, J.E. and A. Cryer (1991) A Practical Guide to Monoclonal Antibodies, John Wiley & Sons, New York NY).

The titer and avidity of polyclonal antibody preparations may be further evaluated to determine the quality and suitability of such preparations for certain downstream applications. For example, a polyclonal antibody preparation containing at least 1-2 mg specific antibody/ml, preferably 5-10 mg specific antibody/ml, is generally employed in procedures requiring precipitation of APRG-antibody complexes. Procedures for evaluating antibody specificity, titer, and avidity, and guidelines for antibody quality and usage in various applications, are generally available. (See, e.g., Catty, supra, and Coligan et al. supra.)

In another embodiment of the invention, the polynucleotides encoding APRG, or any fragment or complement thereof, may be used for therapeutic purposes. In one aspect, modifications of gene expression can be achieved by designing complementary sequences or antisense molecules (DNA, RNA, PNA, or modified oligonucleotides) to the coding or regulatory regions of the gene encoding APRG. Such technology is well known in the art, and antisense oligonucleotides or larger fragments can be designed from various locations along the coding or control regions of sequences encoding APRG. (See, e.g., Agrawal, S., ed. (1996) Antisense Therapeutics, Humana Press Inc., Totawa NJ.)

In therapeutic use, any gene delivery system suitable for introduction of the antisense sequences into appropriate target cells can be used. Antisense sequences can be delivered intracellularly in the form of an expression plasmid which, upon transcription, produces a sequence complementary to at least a portion of the cellular sequence encoding the target protein. (See, e.g., Slater, J.E. et al. (1998) J. Allergy Cli. Immunol. 102(3):469-475; and Scanlon, K.J. et al. (1995) 9(13):1288-1296.) Antisense sequences can also be introduced intracellularly through the use of viral vectors, such as retrovirus and adeno-associated virus vectors. (See, e.g., Miller, A.D. (1990) Blood 76:271; Ausubel, supra; Uckert, W. and W. Walther (1994) Pharmacol. Ther. 63(3):323-347.) Other gene delivery mechanisms include liposome-derived systems, artificial viral envelopes, and other systems known in the art. (See, e.g., Rossi, J.J. (1995) Br. Med. Bull. 51(1):217-225; Boado, R.J. et al. (1998) J. Pharm. Sci. 87(11):1308-1315; and Morris, M.C. et al. (1997) Nucleic Acids Res. 25(14):2730-2736.)

In another embodiment of the invention, polynucleotides encoding APRG may be used for somatic or germline gene therapy. Gene therapy may be performed to (i) correct a genetic deficiency (e.g., in the cases of severe combined immunodeficiency (SCID)-X1 disease characterized by X-linked inheritance (Cavazzana-Calvo, M. et al. (2000) Science 288:669-672), severe combined immunodeficiency syndrome associated with an inherited adenosine deaminase (ADA) deficiency (Blaese, R.M. et al. (1995) Science 270:475-480; Bordignon, C. et al. (1995) Science 270:470-475), cystic fibrosis (Zabner, J. et al. (1993) Cell 75:207-216; Crystal, R.G. et al. (1995) Hum. Gene Therapy 6:643-666; Crystal, R.G. et al. (1995) Hum. Gene Therapy 6:667-703), thalassaemias, familial hypercholesterolemia, and hemophilia resulting from Factor VIII or Factor IX deficiencies (Crystal, R.G. (1995) Science 270:404-410; Verma, I.M. and N. Somia (1997) Nature 389:239-242)), (ii) express a conditionally lethal gene product (e.g., in the case of cancers which result from unregulated cell proliferation), or (iii) express a protein which affords protection against intracellular parasites (e.g., against human retroviruses, such as human immunodeficiency virus (HIV) (Baltimore, D. (1988) Nature 335:395-396; Poeschla, E. et al. (1996) Proc. Natl. Acad. Sci. USA. 93:11395-11399), hepatitis B or C virus (HBV, HCV); fungal parasites, such as Candida albicans and Paracoccidioides

brasiliensis; and protozoan parasites such as Plasmodium falciparum and Trypanosoma cruzi). In the case where a genetic deficiency in APRG expression or regulation causes disease, the expression of APRG from an appropriate population of transduced cells may alleviate the clinical manifestations caused by the genetic deficiency.

In a further embodiment of the invention, diseases or disorders caused by deficiencies in APRG are treated by constructing mammalian expression vectors encoding APRG and introducing these vectors by mechanical means into APRG-deficient cells. Mechanical transfer technologies for use with cells in vivo or ex vivo include (i) direct DNA microinjection into individual cells, (ii) ballistic gold particle delivery, (iii) liposome-mediated transfection, (iv) receptor-mediated gene transfer, and (v) the use of DNA transposons (Morgan, R.A. and W.F. Anderson (1993) Annu. Rev. Biochem. 62:191-217; Ivics, Z. (1997) Cell 91:501-510; Boulay, J-L. and H. Récipon (1998) Curr. Opin. Biotechnol. 9:445-450).

Expression vectors that may be effective for the expression of APRG include, but are not limited to, the PCDNA 3.1, EPITAG, PRCCMV2, PREP, PVAX vectors (Invitrogen, Carlsbad CA), PCMV-SCRIPT, PCMV-TAG, PEGSH/PERV (Stratagene, La Jolla CA), and PTET-OFF, PTET-ON, PTRE2, PTRE2-LUC, PTK-HYG (Clontech, Palo Alto CA). APRG may be expressed using (i) a constitutively active promoter, (e.g., from cytomegalovirus (CMV), Rous sarcoma virus (RSV), SV40 virus, thymidine kinase (TK), or β-actin genes), (ii) an inducible promoter (e.g., the tetracycline-regulated promotor (Gossen, M. and H. Bujard (1992) Proc. Natl. Acad. Sci. USA 89:5547-5551; Gossen, M. et al. (1995) Science 268:1766-1769; Rossi, F.M.V. and H.M. Blau (1998) Curr. Opin. Biotechnol. 9:451-456), commercially available in the T-REX plasmid (Invitrogen)); the ecdysone-inducible promoter (available in the plasmids PVGRXR and PIND; Invitrogen); the FK506/rapamycin inducible promoter; or the RU486/mifepristone inducible promoter (Rossi, F.M.V. and Blau, H.M. supra)), or (iii) a tissue-specific promoter or the native promoter of the endogenous gene encoding APRG from a normal individual.

Commercially available liposome transformation kits (e.g., the PERFECT LIPID TRANSFECTION KIT, available from Invitrogen) allow one with ordinary skill in the art to deliver polynucleotides to target cells in culture and require minimal effort to optimize experimental parameters. In the alternative, transformation is performed using the calcium phosphate method (Graham, F.L. and A.J. Eb (1973) Virology 52:456-467), or by electroporation (Neumann, E. et al. (1982) EMBO J. 1:841-845). The introduction of DNA to primary cells requires modification of these standardized mammalian transfection protocols.

In another embodiment of the invention, diseases or disorders caused by genetic defects with respect to APRG expression are treated by constructing a retrovirus vector consisting of (i) the

polynucleotide encoding APRG under the control of an independent promoter or the retrovirus long terminal repeat (LTR) promoter, (ii) appropriate RNA packaging signals, and (iii) a Rev-responsive element (RRE) along with additional retrovirus *cis*-acting RNA sequences and coding sequences required for efficient vector propagation. Retrovirus vectors (e.g., PFB and PFBNEO) are commercially available (Stratagene) and are based on published data (Riviere, I. et al. (1995) Proc. Natl. Acad. Sci. USA 92:6733-6737), incorporated by reference herein. The vector is propagated in an appropriate vector producing cell line (VPCL) that expresses an envelope gene with a tropism for receptors on the target cells or a promiscuous envelope protein such as VSVg (Armentano, D. et al. (1987) J. Virol. 61:1647-1650; Bender, M.A. et al. (1987) J. Virol. 61:1639-1646; Adam, M.A. and A.D. Miller (1988) J. Virol. 62:3802-3806; Dull, T. et al. (1998) J. Virol. 72:8463-8471; Zufferey, R. et al. (1998) J. Virol. 72:9873-9880). U.S. Patent Number 5,910,434 to Rigg ("Method for obtaining retrovirus packaging cell lines producing high transducing efficiency retroviral supernatant") discloses a method for obtaining retrovirus packaging cell lines and is hereby incorporated by reference. *Propagation of retrovirus vectors, transduction of a population of cells* (e.g., $CD4^+$ T-cells), and the return of transduced cells to a patient are procedures well known to persons skilled in the art of gene therapy and have been well documented (Ranga, U. et al. (1997) J. Virol. 71:7020-7029; Bauer, G. et al. (1997) Blood 89:2259-2267; Bonyhadi, M.L. (1997) J. Virol. 71:4707-4716; Ranga, U. et al. (1998) Proc. Natl. Acad. Sci. USA 95:1201-1206; Su, L. (1997) Blood 89:2283-2290).

In the alternative, an adenovirus-based gene therapy delivery system is used to deliver polynucleotides encoding APRG to cells which have one or more genetic abnormalities with respect to the expression of APRG. The construction and packaging of adenovirus-based vectors are well known to those with ordinary skill in the art. Replication defective adenovirus vectors have proven to be versatile for inporting genes encoding immunoregulatory proteins into intact islets in the pancreas (Csete, M.E. et al. (1995) Transplantation 27:263-268). Potentially useful adenoviral vectors are described in U.S. Patent Number 5,707,618 to Armentano ("Adenovirus vectors for gene therapy"), hereby incorporated by reference. For adenoviral vectors, see also Antinozzi, P.A. et al. (1999) Annu. Rev. Nutr. 19:511-544 and Verma, I.M. and N. Somia (1997) Nature 18:389:239-242, both incorporated by reference herein.

In another alternative, a herpes-based, gene therapy delivery system is used to deliver polynucleotides encoding APRG to target cells which have one or more genetic abnormalities with respect to the expression of APRG. The use of herpes simplex virus (HSV)-based vectors may be especially valuable for introducing APRG to cells of the central nervous system, for which HSV has a tropism. The construction and packaging of herpes-based vectors are well known to those with ordinary skill in the art. A replication-competent herpes simplex virus (HSV) type I-based vector has

been used to deliver a reporter gene to the eyes of primates (Liu, X. et al. (1999) Exp. Eye Res. 169:385-395). The construction of a HSV-1 virus vector has also been disclosed in detail in U.S. Patent Number 5,804,413 to DeLuca ("Herpes simplex virus strains for gene transfer"), which is hereby incorporated by reference. U.S. Patent Number 5,804,413 teaches the use of recombinant HSV d92 which consists of a genome containing at least one exogenous gene to be transferred to a cell under the control of the appropriate promoter for purposes including human gene therapy. Also taught by this patent are the construction and use of recombinant HSV strains deleted for ICP4, ICP27 and ICP22. For HSV vectors, see also Goins, W.F. et al. (1999) J. Virol. 73:519-532 and Xu, H. et al. (1994) Dev. Biol. 163:152-161, hereby incorporated by reference. The manipulation of cloned herpesvirus sequences, the generation of recombinant virus following the transfection of multiple plasmids containing different segments of the large herpesvirus genomes, the growth and propagation of herpesvirus, and the infection of cells with herpesvirus are techniques well known to those of ordinary skill in the art.

In another alternative, an alphavirus (positive, single-stranded RNA virus) vector is used to deliver polynucleotides encoding APRG to target cells. The biology of the prototypic alphavirus, Semliki Forest Virus (SFV), has been studied extensively and gene transfer vectors have been based on the SFV genome (Garoff, H. and K.-J. Li (1998) Curr. Opin. Biotechnol. 9:464-469). During alphavirus RNA replication, a subgenomic RNA is generated that normally encodes the viral capsid proteins. This subgenomic RNA replicates to higher levels than the full length genomic RNA, resulting in the overproduction of capsid proteins relative to the viral proteins with enzymatic activity (e.g., protease and polymerase). Similarly, inserting the coding sequence for APRG into the alphavirus genome in place of the capsid-coding region results in the production of a large number of APRG-coding RNAs and the synthesis of high levels of APRG in vector transduced cells. While alphavirus infection is typically associated with cell lysis within a few days, the ability to establish a persistent infection in hamster normal kidney cells (BHK-21) with a variant of Sindbis virus (SIN) indicates that the lytic replication of alphaviruses can be altered to suit the needs of the gene therapy application (Dryga, S.A. et al. (1997) Virology 228:74-83). The wide host range of alphaviruses will allow the introduction of APRG into a variety of cell types. The specific transduction of a subset of cells in a population may require the sorting of cells prior to transduction. The methods of manipulating infectious cDNA clones of alphaviruses, performing alphavirus cDNA and RNA transfections, and performing alphavirus infections, are well known to those with ordinary skill in the art.

Oligonucleotides derived from the transcription initiation site, e.g., between about positions -10 and +10 from the start site, may also be employed to inhibit gene expression. Similarly, inhibition can be achieved using triple helix base-pairing methodology. Triple helix pairing is useful because it causes

inhibition of the ability of the double helix to open sufficiently for the binding of polymerases, transcription factors, or regulatory molecules. Recent therapeutic advances using triplex DNA have been described in the literature. (See, e.g., Gee, J.E. et al. (1994) in Huber, B.E. and B.I. Carr, Molecular and Immunologic Approaches, Futura Publishing, Mt. Kisco NY, pp. 163-177.) A complementary sequence or antisense molecule may also be designed to block translation of mRNA by preventing the transcript from binding to ribosomes.

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. For example, engineered hammerhead motif ribozyme molecules may specifically and efficiently catalyze endonucleolytic cleavage of sequences encoding ΔPRG.

Specific ribozyme cleavage sites within any potential RNA target are initially identified by scanning the target molecule for ribozyme cleavage sites, including the following sequences: GUA, GUU, and GUC. Once identified, short RNA sequences of between 15 and 20 ribonucleotides, corresponding to the region of the target gene containing the cleavage site, may be evaluated for secondary structural features which may render the oligonucleotide inoperable. The suitability of candidate targets may also be evaluated by testing accessibility to hybridization with complementary oligonucleotides using ribonuclease protection assays.

Complementary ribonucleic acid molecules and ribozymes of the invention may be prepared by any method known in the art for the synthesis of nucleic acid molecules. These include techniques for chemically synthesizing oligonucleotides such as solid phase phosphoramidite chemical synthesis. Alternatively, RNA molecules may be generated by in vitro and in vivo transcription of DNA sequences encoding ΔPRG. Such DNA sequences may be incorporated into a wide variety of vectors with suitable RNA polymerase promoters such as T7 or SP6. Alternatively, these cDNA constructs that synthesize complementary RNA, constitutively or inducibly, can be introduced into cell lines, cells, or tissues.

RNA molecules may be modified to increase intracellular stability and half-life. Possible modifications include, but are not limited to, the addition of flanking sequences at the 5' and/or 3' ends of the molecule, or the use of phosphorothioate or 2' O-methyl rather than phosphodiesterase linkages within the backbone of the molecule. This concept is inherent in the production of PNAs and can be extended in all of these molecules by the inclusion of nontraditional bases such as inosine, queosine, and wybutosine, as well as acetyl-, methyl-, thio-, and similarly modified forms of adenine, cytidine, guanine, thymine, and uridine which are not as easily recognized by endogenous endonucleases.

An additional embodiment of the invention encompasses a method for screening for a

compound which is effective in altering expression of a polynucleotide encoding APRG. Compounds which may be effective in altering expression of a specific polynucleotide may include, but are not limited to, oligonucleotides, antisense oligonucleotides, triple helix-forming oligonucleotides, transcription factors and other polypeptide transcriptional regulators, and non-macromolecular chemical entities which are capable of interacting with specific polynucleotide sequences. Effective compounds may alter polynucleotide expression by acting as either inhibitors or promoters of polynucleotide expression. Thus, in the treatment of disorders associated with increased APRG expression or activity, a compound which specifically inhibits expression of the polynucleotide encoding APRG may be therapeutically useful, and in the treatment of disorders associated with decreased APRG expression or activity, a compound which specifically promotes expression of the polynucleotide encoding APRG may be therapeutically useful.

At least one, and up to a plurality, of test compounds may be screened for effectiveness in altering expression of a specific polynucleotide. A test compound may be obtained by any method commonly known in the art, including chemical modification of a compound known to be effective in altering polynucleotide expression; selection from an existing, commercially-available or proprietary library of naturally-occurring or non-natural chemical compounds; rational design of a compound based on chemical and/or structural properties of the target polynucleotide; and selection from a library of chemical compounds created combinatorially or randomly. A sample comprising a polynucleotide encoding APRG is exposed to at least one test compound thus obtained. The sample may comprise, for example, an intact or permeabilized cell, or an in vitro cell-free or reconstituted biochemical system. Alterations in the expression of a polynucleotide encoding APRG are assayed by any method commonly known in the art. Typically, the expression of a specific nucleotide is detected by hybridization with a probe having a nucleotide sequence complementary to the sequence of the polynucleotide encoding APRG. The amount of hybridization may be quantified, thus forming the basis for a comparison of the expression of the polynucleotide both with and without exposure to one or more test compounds. Detection of a change in the expression of a polynucleotide exposed to a test compound indicates that the test compound is effective in altering the expression of the polynucleotide. A screen for a compound effective in altering expression of a specific polynucleotide can be carried out, for example, using a Schizosaccharomyces pombe gene expression system (Atkins, D. et al. (1999) U.S. Patent No. 5,932,435; Arndt, G.M. et al. (2000) Nucleic Acids Res. 28:E15) or a human cell line such as HeLa cell (Clarke, M.L. et al. (2000) Biochem. Biophys. Res. Commun. 268:8-13). A particular embodiment of the present invention involves screening a combinatorial library of oligonucleotides (such as deoxyribonucleotides, ribonucleotides, peptide nucleic acids, and modified oligonucleotides) for antisense activity against a specific polynucleotide sequence (Bruice, T.W. et al. (1997) U.S. Patent No. 5,686,242; Bruice, T.W. et al. (2000) U.S.

Patent No. 6,022,691).

Many methods for introducing vectors into cells or tissues are available and equally suitable for use in vivo, in vitro, and ex vivo. For ex vivo therapy, vectors may be introduced into stem cells taken from the patient and clonally propagated for autologous transplant back into that same patient. Delivery by transfection, by liposome injections, or by polycationic amino polymers may be achieved using methods which are well known in the art. (See, e.g., Goldman, C.K. et al. (1997) Nat. Biotechnol. 15:462-466.)

Any of the therapeutic methods described above may be applied to any subject in need of such therapy, including, for example, mammals such as humans, dogs, cats, cows, horses, rabbits, and monkeys.

An additional embodiment of the invention relates to the administration of a composition which generally comprises an active ingredient formulated with a pharmaceutically acceptable excipient. Excipients may include, for example, sugars, starches, celluloses, gums, and proteins. Various formulations are commonly known and are thoroughly discussed in the latest edition of Remington's Pharmaceutical Sciences (Maack Publishing, Easton PA). Such compositions may consist of APRG, antibodies to APRG, and mimetics, agonists, antagonists, or inhibitors of APRG.

The compositions utilized in this invention may be administered by any number of routes including, but not limited to, oral, intravenous, intramuscular, intra-arterial, intramedullary, intrathecal, intraventricular, pulmonary, transdermal, subcutaneous, intraperitoneal, intranasal, enteral, topical, sublingual, or rectal means.

Compositions for pulmonary administration may be prepared in liquid or dry powder form. These compositions are generally aerosolized immediately prior to inhalation by the patient. In the case of small molecules (e.g. traditional low molecular weight organic drugs), aerosol delivery of fast-acting formulations is well-known in the art. In the case of macromolecules (e.g. larger peptides and proteins), recent developments in the field of pulmonary delivery via the alveolar region of the lung have enabled the practical delivery of drugs such as insulin to blood circulation (see, e.g., Patton, J.S. et al., U.S. Patent No. 5,997,848). Pulmonary delivery has the advantage of administration without needle injection, and obviates the need for potentially toxic penetration enhancers.

Compositions suitable for use in the invention include compositions wherein the active ingredients are contained in an effective amount to achieve the intended purpose. The determination of an effective dose is well within the capability of those skilled in the art.

Specialized forms of compositions may be prepared for direct intracellular delivery of macromolecules comprising APRG or fragments thereof. For example, liposome preparations containing a cell-impermeable macromolecule may promote cell fusion and intracellular delivery of the

macromolecule. Alternatively, APRG or a fragment thereof may be joined to a short cationic N-terminal portion from the HIV Tat-1 protein. Fusion proteins thus generated have been found to transduce into the cells of all tissues, including the brain, in a mouse model system (Schwarze, S.R. et al. (1999) Science 285:1569-1572).

For any compound, the therapeutically effective dose can be estimated initially either in cell culture assays, e.g., of neoplastic cells, or in animal models such as mice, rats, rabbits, dogs, monkeys, or pigs. An animal model may also be used to determine the appropriate concentration range and route of administration. Such information can then be used to determine useful doses and routes for administration in humans.

A therapeutically effective dose refers to that amount of active ingredient, for example APRG or fragments thereof, antibodies of APRG, and agonists, antagonists or inhibitors of APRG, which ameliorates the symptoms or condition. Therapeutic efficacy and toxicity may be determined by standard pharmaceutical procedures in cell cultures or with experimental animals, such as by calculating the ED_{50} (the dose therapeutically effective in 50% of the population) or LD_{50} (the dose lethal to 50% of the population) statistics. The dose ratio of toxic to therapeutic effects is the therapeutic index, which can be expressed as the LD_{50}/ED_{50} ratio. Compositions which exhibit large therapeutic indices are preferred. The data obtained from cell culture assays and animal studies are used to formulate a range of dosage for human use. The dosage contained in such compositions is preferably within a range of circulating concentrations that includes the ED_{50} with little or no toxicity. The dosage varies within this range depending upon the dosage form employed, the sensitivity of the patient, and the route of administration.

The exact dosage will be determined by the practitioner, in light of factors related to the subject requiring treatment. Dosage and administration are adjusted to provide sufficient levels of the active moiety or to maintain the desired effect. Factors which may be taken into account include the severity of the disease state, the general health of the subject, the age, weight, and gender of the subject, time and frequency of administration, drug combination(s), reaction sensitivities, and response to therapy. Long-acting compositions may be administered every 3 to 4 days, every week, or biweekly depending on the half-life and clearance rate of the particular formulation.

Normal dosage amounts may vary from about 0.1 μg to 100,000 μg, up to a total dose of about 1 gram, depending upon the route of administration. Guidance as to particular dosages and methods of delivery is provided in the literature and generally available to practitioners in the art. Those skilled in the art will employ different formulations for nucleotides than for proteins or their inhibitors. Similarly, delivery of polynucleotides or polypeptides will be specific to particular cells, conditions, locations, etc.

DIAGNOSTICS

In another embodiment, antibodies which specifically bind APRG may be used for the diagnosis of disorders characterized by expression of APRG, or in assays to monitor patients being treated with APRG or agonists, antagonists, or inhibitors of APRG. Antibodies useful for diagnostic purposes may be prepared in the same manner as described above for therapeutics. Diagnostic assays for APRG include methods which utilize the antibody and a label to detect APRG in human body fluids or in extracts of cells or tissues. The antibodies may be used with or without modification, and may be labeled by covalent or non-covalent attachment of a reporter molecule. A wide variety of reporter molecules, several of which are described above, are known in the art and may be used.

A variety of protocols for measuring APRG, including ELISAs, RIAs, and FACS, are known in the art and provide a basis for diagnosing altered or abnormal levels of APRG expression. Normal or standard values for APRG expression are established by combining body fluids or cell extracts taken from normal mammalian subjects, for example, human subjects, with antibodies to APRG under conditions suitable for complex formation. The amount of standard complex formation may be quantitated by various methods, such as photometric means. Quantities of APRG expressed in subject, control, and disease samples from biopsied tissues are compared with the standard values. Deviation between standard and subject values establishes the parameters for diagnosing disease.

In another embodiment of the invention, the polynucleotides encoding APRG may be used for diagnostic purposes. The polynucleotides which may be used include oligonucleotide sequences, complementary RNA and DNA molecules, and PNAs. The polynucleotides may be used to detect and quantify gene expression in biopsied tissues in which expression of APRG may be correlated with disease. The diagnostic assay may be used to determine absence, presence, and excess expression of APRG, and to monitor regulation of APRG levels during therapeutic intervention.

In one aspect, hybridization with PCR probes which are capable of detecting polynucleotide sequences, including genomic sequences, encoding APRG or closely related molecules may be used to identify nucleic acid sequences which encode APRG. The specificity of the probe, whether it is made from a highly specific region, e.g., the 5' regulatory region, or from a less specific region, e.g., a conserved motif, and the stringency of the hybridization or amplification will determine whether the probe identifies only naturally occurring sequences encoding APRG, allelic variants, or related sequences.

Probes may also be used for the detection of related sequences, and may have at least 50% sequence identity to any of the APRG encoding sequences. The hybridization probes of the subject invention may be DNA or RNA and may be derived from the sequence of SEQ ID NO:4-6 or from genomic sequences including promoters, enhancers, and introns of the APRG gene.

Means for producing specific hybridization probes for DNAs encoding APRG include the cloning of polynucleotide sequences encoding APRG or APRG derivatives into vectors for the production of mRNA probes. Such vectors are known in the art, are commercially available, and may be used to synthesize RNA probes in vitro by means of the addition of the appropriate RNA polymerases and the appropriate labeled nucleotides. Hybridization probes may be labeled by a variety of reporter groups, for example, by radionuclides such as ^{32}P or ^{35}S, or by enzymatic labels, such as alkaline phosphatase coupled to the probe via avidin/biotin coupling systems, and the like.

Polynucleotide sequences encoding APRG may be used for the diagnosis of disorders associated with expression of APRG. Examples of such disorders include, but are not limited to, a cell proliferative disorder such as actinic keratosis, arteriosclerosis, atherosclerosis, bursitis, cirrhosis, hepatitis, mixed connective tissue disease (MCTD), myelofibrosis, paroxysmal nocturnal hemoglobinuria, polycythemia vera, psoriasis, primary thrombocythemia, and cancers including adenocarcinoma, leukemia, lymphoma, melanoma, myeloma, sarcoma, teratocarcinoma, and, in particular, a cancer of the adrenal gland, bladder, bone, bone marrow, brain, breast, cervix, gall bladder, ganglia, gastrointestinal tract, heart, kidney, liver, lung, muscle, ovary, pancreas, parathyroid, penis, prostate, salivary glands, skin, spleen, testis, thymus, thyroid, and uterus; an immunological disorder such as acquired immunodeficiency syndrome (AIDS), Addison's disease, adult respiratory distress syndrome, allergies, ankylosing spondylitis, amyloidosis, anemia, asthma, atherosclerosis, autoimmune hemolytic anemia, autoimmune thyroiditis, autoimmune polyendocrinopathy-candidiasis-ectodermal dystrophy (APECED), bronchitis, cholecystitis, contact dermatitis, Crohn's disease, atopic dermatitis, dermatomyositis, diabetes mellitus, emphysema, episodic lymphopenia with lymphocytotoxins, erythroblastosis fetalis, erythema nodosum, atrophic gastritis, glomerulonephritis, Goodpasture's syndrome, gout, Graves' disease, Hashimoto's thyroiditis, hypereosinophilia, irritable bowel syndrome, multiple sclerosis, myasthenia gravis, myocardial or pericardial inflammation, osteoarthritis, osteoporosis, pancreatitis, polymyositis, psoriasis, Reiter's syndrome, rheumatoid arthritis, scleroderma, Sjögren's syndrome, systemic anaphylaxis, systemic lupus erythematosus, systemic sclerosis, thrombocytopenic purpura, ulcerative colitis, uveitis, Werner syndrome, a complication of cancer, hemodialysis, and extracorporeal circulation, viral, bacterial, fungal, parasitic, protozoal, and helminthic infections, and trauma; and a reproductive disorder such as a disorder of prolactin production, infertility, including tubal disease, ovulatory defects, and endometriosis, a disruption of the estrous cycle, a disruption of the menstrual cycle, polycystic ovary syndrome, ovarian hyperstimulation syndrome, endometrial and ovarian tumors, uterine fibroids, autoimmune disorders, ectopic pregnancies, and teratogenesis; cancer of the breast, fibrocystic breast disease, and galactorrhea; disruptions of spermatogenesis, abnormal sperm physiology, cancer of the testis, cancer of the prostate, benign prostatic hyperplasia, prostatitis, Peyronie's disease, impotence,

carcinoma of the male breast, and gynecomastia. The polynucleotide sequences encoding APRG may be used in Southern or northern analysis, dot blot, or other membrane-based technologies; in PCR technologies; in dipstick, pin, and multiformat ELISA-like assays; and in microarrays utilizing fluids or tissues from patients to detect altered APRG expression. Such qualitative or quantitative methods are well known in the art.

In a particular aspect, the nucleotide sequences encoding APRG may be useful in assays that detect the presence of associated disorders, particularly those mentioned above. The nucleotide sequences encoding APRG may be labeled by standard methods and added to a fluid or tissue sample from a patient under conditions suitable for the formation of hybridization complexes. After a suitable incubation period, the sample is washed and the signal is quantified and compared with a standard value. If the amount of signal in the patient sample is significantly altered in comparison to a control sample then the presence of altered levels of nucleotide sequences encoding APRG in the sample indicates the presence of the associated disorder. Such assays may also be used to evaluate the efficacy of a particular therapeutic treatment regimen in animal studies, in clinical trials, or to monitor the treatment of an individual patient.

In order to provide a basis for the diagnosis of a disorder associated with expression of APRG, a normal or standard profile for expression is established. This may be accomplished by combining body fluids or cell extracts taken from normal subjects, either animal or human, with a sequence, or a fragment thereof, encoding APRG, under conditions suitable for hybridization or amplification. Standard hybridization may be quantified by comparing the values obtained from normal subjects with values from an experiment in which a known amount of a substantially purified polynucleotide is used. Standard values obtained in this manner may be compared with values obtained from samples from patients who are symptomatic for a disorder. Deviation from standard values is used to establish the presence of a disorder.

Once the presence of a disorder is established and a treatment protocol is initiated, hybridization 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 the 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.

With respect to cancer, the presence of an abnormal amount of transcript (either under- or overexpressed) in biopsied tissue from an individual may indicate a predisposition for the development of the disease, or may provide a means for detecting the disease prior to the appearance of actual clinical symptoms. A more definitive diagnosis of this type may allow health professionals to employ preventative measures or aggressive treatment earlier thereby preventing the development or further

progression of the cancer.

Additional diagnostic uses for oligonucleotides designed from the sequences encoding APRG may involve the use of PCR. These oligomers may be chemically synthesized, generated enzymatically, or produced in vitro. Oligomers will preferably contain a fragment of a polynucleotide encoding APRG, or a fragment of a polynucleotide complementary to the polynucleotide encoding APRG, and will be employed under optimized conditions for identification of a specific gene or condition. Oligomers may also be employed under less stringent conditions for detection or quantification of closely related DNA or RNA sequences.

In a particular aspect, oligonucleotide primers derived from the polynucleotide sequences encoding APRG may be used to detect single nucleotide polymorphisms (SNPs). SNPs are substitutions, insertions and deletions that are a frequent cause of inherited or acquired genetic disease in humans. Methods of SNP detection include, but are not limited to, single-stranded conformation polymorphism (SSCP) and fluorescent SSCP (fSSCP) methods. In SSCP, oligonucleotide primers derived from the polynucleotide sequences encoding APRG are used to amplify DNA using the polymerase chain reaction (PCR). The DNA may be derived, for example, from diseased or normal tissue, biopsy samples, bodily fluids, and the like. SNPs in the DNA cause differences in the secondary and tertiary structures of PCR products in single-stranded form, and these differences are detectable using gel electrophoresis in non-denaturing gels. In fSCCP, the oligonucleotide primers are fluorescently labeled, which allows detection of the amplimers in high-throughput equipment such as DNA sequencing machines. Additionally, sequence database analysis methods, termed in silico SNP (isSNP), are capable of identifying polymorphisms by comparing the sequence of individual overlapping DNA fragments which assemble into a common consensus sequence. These computer-based methods filter out sequence variations due to laboratory preparation of DNA and sequencing errors using statistical models and automated analyses of DNA sequence chromatograms. In the alternative, SNPs may be detected and characterized by mass spectrometry using, for example, the high throughput MASSARRAY system (Sequenom, Inc., San Diego CA).

Methods which may also be used to quantify the expression of APRG include radiolabeling or biotinylating nucleotides, coamplification of a control nucleic acid, and interpolating results from standard curves. (See, e.g., Melby, P.C. et al. (1993) J. Immunol. Methods 159:235-244; Duplaa, C. et al. (1993) Anal. Biochem. 212:229-236.) The speed of quantitation of multiple samples may be accelerated by running the assay in a high-throughput format where the oligomer or polynucleotide of interest is presented in various dilutions and a spectrophotometric or colorimetric response gives rapid quantitation.

In further embodiments, oligonucleotides or longer fragments derived from any of the

polynucleotide sequences described herein may be used as elements on a microarray. The microarray can be used in transcript imaging techniques which monitor the relative expression levels of large numbers of genes simultaneously as described below. The microarray may also be used to identify genetic variants, mutations, and polymorphisms. This information may be used to determine gene function, to understand the genetic basis of a disorder, to diagnose a disorder, to monitor progression/regression of disease as a function of gene expression, and to develop and monitor the activities of therapeutic agents in the treatment of disease. In particular, this information may be used to develop a pharmacogenomic profile of a patient in order to select the most appropriate and effective treatment regimen for that patient. For example, therapeutic agents which are highly effective and display the fewest side effects may be selected for a patient based on his/her pharmacogenomic profile.

In another embodiment, APRG, fragments of APRG, or antibodies specific for APRG may be used as elements on a microarray. The microarray may be used to monitor or measure protein-protein interactions, drug-target interactions, and gene expression profiles, as described above.

A particular embodiment relates to the use of the polynucleotides of the present invention to generate a transcript image of a tissue or cell type. A transcript image represents the global pattern of gene expression by a particular tissue or cell type. Global gene expression patterns are analyzed by quantifying the number of expressed genes and their relative abundance under given conditions and at a given time. (See Seilhamer et al., "Comparative Gene Transcript Analysis," U.S. Patent Number 5,840,484, expressly incorporated by reference herein.) Thus a transcript image may be generated by hybridizing the polynucleotides of the present invention or their complements to the totality of transcripts or reverse transcripts of a particular tissue or cell type. In one embodiment, the hybridization takes place in high-throughput format, wherein the polynucleotides of the present invention or their complements comprise a subset of a plurality of elements on a microarray. The resultant transcript image would provide a profile of gene activity.

Transcript images may be generated using transcripts isolated from tissues, cell lines, biopsies, or other biological samples. The transcript image may thus reflect gene expression in vivo, as in the case of a tissue or biopsy sample, or in vitro, as in the case of a cell line.

Transcript images which profile the expression of the polynucleotides of the present invention may also be used in conjunction with in vitro model systems and preclinical evaluation of pharmaceuticals, as well as toxicological testing of industrial and naturally-occurring environmental compounds. All compounds induce characteristic gene expression patterns, frequently termed molecular fingerprints or toxicant signatures, which are indicative of mechanisms of action and toxicity (Nuwaysir, E.F. et al. (1999) Mol. Carcinog. 24:153-159; Steiner, S. and N.L. Anderson (2000) Toxicol. Lett. 112-113:467-471, expressly incorporated by reference herein). If a test compound has a

signature similar to that of a compound with known toxicity, it is likely to share those toxic properties. These fingerprints or signatures are most useful and refined when they contain expression information from a large number of genes and gene families. Ideally, a genome-wide measurement of expression provides the highest quality signature. Even genes whose expression is not altered by any tested compounds are important as well, as the levels of expression of these genes are used to normalize the rest of the expression data. The normalization procedure is useful for comparison of expression data after treatment with different compounds. While the assignment of gene function to elements of a toxicant signature aids in interpretation of toxicity mechanisms, knowledge of gene function is not necessary for the statistical matching of signatures which leads to prediction of toxicity. (See, for example, Press Release 00-02 from the National Institute of Environmental Health Sciences, released February 29, 2000, available at http://www.niehs.nih.gov/oc/news/toxchip.htm.) Therefore, it is important and desirable in toxicological screening using toxicant signatures to include all expressed gene sequences.

In one embodiment, the toxicity of a test compound is assessed by treating a biological sample containing nucleic acids with the test compound. Nucleic acids that are expressed in the treated biological sample are hybridized with one or more probes specific to the polynucleotides of the present invention, so that transcript levels corresponding to the polynucleotides of the present invention may be quantified. The transcript levels in the treated biological sample are compared with levels in an untreated biological sample. Differences in the transcript levels between the two samples are indicative of a toxic response caused by the test compound in the treated sample.

Another particular embodiment relates to the use of the polypeptide sequences of the present invention to analyze the proteome of a tissue or cell type. The term proteome refers to the global pattern of protein expression in a particular tissue or cell type. Each protein component of a proteome can be subjected individually to further analysis. Proteome expression patterns, or profiles, are analyzed by quantifying the number of expressed proteins and their relative abundance under given conditions and at a given time. A profile of a cell's proteome may thus be generated by separating and analyzing the polypeptides of a particular tissue or cell type. In one embodiment, the separation is achieved using two-dimensional gel electrophoresis, in which proteins from a sample are separated by isoelectric focusing in the first dimension, and then according to molecular weight by sodium dodecyl sulfate slab gel electrophoresis in the second dimension (Steiner and Anderson, supra). The proteins are visualized in the gel as discrete and uniquely positioned spots, typically by staining the gel with an agent such as Coomassie Blue or silver or fluorescent stains. The optical density of each protein spot is generally proportional to the level of the protein in the sample. The optical densities of equivalently positioned protein spots from different samples, for example, from biological samples either treated or

untreated with a test compound or therapeutic agent, are compared to identify any changes in protein spot density related to the treatment. The proteins in the spots are partially sequenced using, for example, standard methods employing chemical or enzymatic cleavage followed by mass spectrometry. The identity of the protein in a spot may be determined by comparing its partial sequence, preferably of at least 5 contiguous amino acid residues, to the polypeptide sequences of the present invention. In some cases, further sequence data may be obtained for definitive protein identification.

A proteomic profile may also be generated using antibodies specific for APRG to quantify the levels of APRG expression. In one embodiment, the antibodies are used as elements on a microarray, and protein expression levels are quantified by exposing the microarray to the sample and detecting the levels of protein bound to each array element (Lueking, A. et al. (1999) Anal. Biochem. 270:103-111; Mendoze, L.G. et al. (1999) Biotechniques 27:778-788). Detection may be performed by a variety of methods known in the art, for example, by reacting the proteins in the sample with a thiol- or amino-reactive fluorescent compound and detecting the amount of fluorescence bound at each array element.

Toxicant signatures at the proteome level are also useful for toxicological screening, and should be analyzed in parallel with toxicant signatures at the transcript level. There is a poor correlation between transcript and protein abundances for some proteins in some tissues (Anderson, N.L. and J. Seilhamer (1997) Electrophoresis 18:533-537), so proteome toxicant signatures may be useful in the analysis of compounds which do not significantly affect the transcript image, but which alter the proteomic profile. In addition, the analysis of transcripts in body fluids is difficult, due to rapid degradation of mRNA, so proteomic profiling may be more reliable and informative in such cases.

In another embodiment, the toxicity of a test compound is assessed by treating a biological sample containing proteins with the test compound. Proteins that are expressed in the treated biological sample are separated so that the amount of each protein can be quantified. The amount of each protein is compared to the amount of the corresponding protein in an untreated biological sample. A difference in the amount of protein between the two samples is indicative of a toxic response to the test compound in the treated sample. Individual proteins are identified by sequencing the amino acid residues of the individual proteins and comparing these partial sequences to the polypeptides of the present invention.

In another embodiment, the toxicity of a test compound is assessed by treating a biological sample containing proteins with the test compound. Proteins from the biological sample are incubated with antibodies specific to the polypeptides of the present invention. The amount of protein recognized by the antibodies is quantified. The amount of protein in the treated biological sample is compared with the amount in an untreated biological sample. A difference in the amount of protein between the two samples is indicative of a toxic response to the test compound in the treated sample.

Microarrays may be prepared, used, and analyzed using methods known in the art. (See, e.g.,

Brennan, T.M. et al. (1995) U.S. Patent No. 5,474,796; Schena, M. et al. (1996) Proc. Natl. Acad. Sci. USA 93:10614-10619; Baldeschweiler et al. (1995) PCT application WO95/251116; Shalon, D. et al. (1995) PCT application WO95/35505; Heller, R.A. et al. (1997) Proc. Natl. Acad. Sci. USA 94:2150-2155; and Heller, M.J. et al. (1997) U.S. Patent No. 5,605,662.) Various types of microarrays are well known and thoroughly described in DNA Microarrays: A Practical Approach, M. Schena, ed. (1999) Oxford University Press, London, hereby expressly incorporated by reference.

In another embodiment of the invention, nucleic acid sequences encoding APRG may be used to generate hybridization probes useful in mapping the naturally occurring genomic sequence. Either coding or noncoding sequences may be used, and in some instances, noncoding sequences may be preferable over coding sequences. For example, conservation of a coding sequence among members of a multi-gene family may potentially cause undesired cross hybridization during chromosomal mapping. The sequences may be mapped to a particular chromosome, to a specific region of a chromosome, or to artificial chromosome constructions, e.g., human artificial chromosomes (HACs), yeast artificial chromosomes (YACs), bacterial artificial chromosomes (BACs), bacterial P1 constructions, or single chromosome cDNA libraries. (See, e.g., Harrington, J.J. et al. (1997) Nat. Genet. 15:345-355; Price, C.M. (1993) Blood Rev. 7:127-134; and Trask, B.J. (1991) Trends Genet. 7:149-154.) Once mapped, the nucleic acid sequences of the invention may be used to develop genetic linkage maps, for example, which correlate the inheritance of a disease state with the inheritance of a particular chromosome region or restriction fragment length polymorphism (RFLP). (See, for example, Lander, E.S. and D. Botstein (1986) Proc. Natl. Acad. Sci. USA 83:7353-7357.)

Fluorescent in situ hybridization (FISH) may be correlated with other physical and genetic map data. (See, e.g., Heinz-Ulrich, et al. (1995) in Meyers, supra, pp. 965-968.) Examples of genetic map data can be found in various scientific journals or at the Online Mendelian Inheritance in Man (OMIM) World Wide Web site. Correlation between the location of the gene encoding APRG on a physical map and a specific disorder, or a predisposition to a specific disorder, may help define the region of DNA associated with that disorder and thus may further positional cloning efforts.

In situ hybridization of chromosomal preparations and physical mapping techniques, such as linkage analysis using established chromosomal markers, may be used for extending genetic maps. Often the placement of a gene on the chromosome of another mammalian species, such as mouse, may reveal associated markers even if the exact chromosomal locus is not known. This information is valuable to investigators searching for disease genes using positional cloning or other gene discovery techniques. Once the gene or genes responsible for a disease or syndrome have been crudely localized by genetic linkage to a particular genomic region, e.g., ataxia-telangiectasia to 11q22-23, any sequences mapping to that area may represent associated or regulatory genes for further investigation. (See, e.g.,

Gatti, R.A. et al. (1988) Nature 336:577-580.) The nucleotide sequence of the instant invention may also be used to detect differences in the chromosomal location due to translocation, inversion, etc., among normal, carrier, or affected individuals.

In another embodiment of the invention, APRG, its catalytic or immunogenic fragments, or oligopeptides thereof can be used for screening libraries of compounds in any of a variety of drug screening techniques. The fragment employed in such screening may be free in solution, affixed to a solid support, borne on a cell surface, or located intracellularly. The formation of binding complexes between APRG and the agent being tested may be measured.

Another technique for drug screening provides for high throughput screening of compounds having suitable binding affinity to the protein of interest. (See, e.g., Geysen, et al. (1984) PCT application WO84/03564.) In this method, large numbers of different small test compounds are synthesized on a solid substrate. The test compounds are reacted with APRG, or fragments thereof, and washed. Bound APRG is then detected by methods well known in the art. Purified APRG can also be coated directly onto plates for use in the aforementioned drug screening techniques. Alternatively, non-neutralizing antibodies can be used to capture the peptide and immobilize it on a solid support.

In another embodiment, one may use competitive drug screening assays in which neutralizing antibodies capable of binding APRG specifically compete with a test compound for binding APRG. In this manner, antibodies can be used to detect the presence of any peptide which shares one or more antigenic determinants with APRG.

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

Without further elaboration, it is believed that one skilled in the art can, using the preceding description, utilize the present invention to its fullest extent. The following preferred specific embodiments are, therefore, to be construed as merely illustrative, and not limitative of the remainder of the disclosure in any way whatsoever.

The disclosures of all patents, applications, and publications mentioned above and below, in particular U.S. Ser. No. 60/250,326, and U.S. Ser. No. 60/209,407, are hereby expressly incorporated by reference.

EXAMPLES

1. Construction of cDNA Libraries

Incyte cDNAs were derived from cDNA libraries described in the LIFESEQ GOLD database

(Incyte Genomics, Palo Alto CA) and shown in Table 4, column 5. Some tissues were homogenized and lysed in guanidinium isothiocyanate, while others were homogenized and lysed in phenol or in a suitable mixture of denaturants, such as TRIZOL (Life Technologies), a monophasic solution of phenol and guanidine isothiocyanate. The resulting lysates were centrifuged over CsCl cushions or extracted with chloroform. RNA was precipitated from the lysates with either isopropanol or sodium acetate and ethanol, or by other routine methods.

Phenol extraction and precipitation of RNA were repeated as necessary to increase RNA purity. In some cases, RNA was treated with DNase. For most libraries, poly(A)+ RNA was isolated using oligo d(T)-coupled paramagnetic particles (Promega), OLIGOTEX latex particles (QIAGEN, Chatsworth CA), or an OLIGOTEX mRNA purification kit (QIAGEN). Alternatively, RNA was isolated directly from tissue lysates using other RNA isolation kits, e.g., the POLY(A)PURE mRNA purification kit (Ambion, Austin TX).

In some cases, Stratagene was provided with RNA and constructed the corresponding cDNA libraries. Otherwise, cDNA was synthesized and cDNA libraries were constructed with the UNIZAP vector system (Stratagene) or SUPERSCRIPT plasmid system (Life Technologies), using the recommended procedures or similar methods known in the art. (See, e.g., Ausubel, 1997, supra, units 5.1-6.6.) Reverse transcription was initiated using oligo d(T) or random primers. Synthetic oligonucleotide adapters were ligated to double stranded cDNA, and the cDNA was digested with the appropriate restriction enzyme or enzymes. For most libraries, the cDNA was size-selected (300-1000 bp) using SEPHACRYL S1000, SEPHAROSE CL2B, or SEPHAROSE CL4B column chromatography (Amersham Pharmacia Biotech) or preparative agarose gel electrophoresis. cDNAs were ligated into compatible restriction enzyme sites of the polylinker of a suitable plasmid, e.g., PBLUESCRIPT plasmid (Stratagene), PSPORT1 plasmid (Life Technologies), PCDNA2.1 plasmid (Invitrogen, Carlsbad CA), PBK-CMV plasmid (Stratagene), or pINCY (Incyte Genomics, Palo Alto CA), or derivatives thereof. Recombinant plasmids were transformed into competent E. coli cells including XL1-Blue, XL1-BlueMRF, or SOLR from Stratagene or DH5α, DH10B, or ElectroMAX DH10B from Life Technologies.

II. Isolation of cDNA Clones

Plasmids obtained as described in Example I were recovered from host cells by in vivo excision using the UNIZAP vector system (Stratagene) or by cell lysis. Plasmids were purified using at least one of the following: a Magic or WIZARD Minipreps DNA purification system (Promega); an AGTC Miniprep purification kit (Edge Biosystems, Gaithersburg MD); and QIAWELL 8 Plasmid, QIAWELL 8 Plus Plasmid, QIAWELL 8 Ultra Plasmid purification systems or the R.E.A.L. PREP 96 plasmid purification kit from QIAGEN. Following precipitation, plasmids were resuspended in 0.1 ml of

distilled water and stored, with or without lyophilization, at 4 °C.

Alternatively, plasmid DNA was amplified from host cell lysates using direct link PCR in a high-throughput format (Rao, V.B. (1994) Anal. Biochem. 216:1-14). Host cell lysis and thermal cycling steps were carried out in a single reaction mixture. Samples were processed and stored in 384-well plates, and the concentration of amplified plasmid DNA was quantified fluorometrically using PICOGREEN dye (Molecular Probes, Eugene OR) and a FLUOROSKAN II fluorescence scanner (Labsystems Oy, Helsinki, Finland).

III. Sequencing and Analysis

Incyte cDNA recovered in plasmids as described in Example II were sequenced as follows. Sequencing reactions were processed using standard methods or high-throughput instrumentation such as the ABI CATALYST 800 (Applied Biosystems) thermal cycler or the PTC-200 thermal cycler (MJ Research) in conjunction with the HYDRA microdispenser (Robbins Scientific) or the MICROLAB 2200 (Hamilton) liquid transfer system. cDNA sequencing reactions were prepared using reagents provided by Amersham Pharmacia Biotech or supplied in ABI sequencing kits such as the ABI PRISM BIGDYE Terminator cycle sequencing ready reaction kit (Applied Biosystems). Electrophoretic separation of cDNA sequencing reactions and detection of labeled polynucleotides were carried out using the MEGABACE 1000 DNA sequencing system (Molecular Dynamics); the ABI PRISM 373 or 377 sequencing system (Applied Biosystems) in conjunction with standard ABI protocols and base calling software; or other sequence analysis systems known in the art. Reading frames within the cDNA sequences were identified using standard methods (reviewed in Ausubel, 1997, supra, unit 7.7). Some of the cDNA sequences were selected for extension using the techniques disclosed in Example VIII.

The polynucleotide sequences derived from Incyte cDNAs were validated by removing vector, linker, and poly(A) sequences and by masking ambiguous bases, using algorithms and programs based on BLAST, dynamic programming, and dinucleotide nearest neighbor analysis. The Incyte cDNA sequences or translations thereof 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 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, for example, Eddy, S.R. (1996) Curr. Opin. Struct. Biol. 6:361-365.) The queries were performed using programs based on BLAST, FASTA, BLIMPS, and HMMER. The Incyte cDNA sequences were assembled to produce full length polynucleotide sequences. Alternatively, GenBank cDNAs, GenBank ESTs, stitched sequences, stretched sequences, or Genscan-predicted coding sequences (see Examples IV and V) were used to extend Incyte cDNA assemblages to full length.

Assembly was performed using programs based on Phred, Phrap, and Consed, and cDNA assemblages 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 polypeptide sequences. Alternatively, a polypeptide of the invention may begin at any of the methionine residues of the full length translated polypeptide. Full length polypeptide sequences were subsequently analyzed by querying against databases such as the GenBank protein databases (genpept), SwissProt, BLOCKS, PRINTS, DOMO, PRODOM, Prosite, and hidden Markov model (HMM)-based protein family databases such as PFAM. Full length polynucleotide sequences are also analyzed using MACDNASIS PRO software (Hitachi Software Engineering, South San Francisco CA) and LASERGENE software (DNASTAR). Polynucleotide and polypeptide sequence alignments are generated using default parameters specified by the CLUSTAL algorithm as incorporated into the MEGALIGN multisequence alignment program (DNASTAR), which also calculates the percent identity between aligned sequences.

Table 7 summarizes the tools, programs, and algorithms used for the analysis and assembly of Incyte cDNA and full length sequences and provides applicable descriptions, references, and threshold parameters. The first column of Table 7 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 applicable, the scores, probability values, and other parameters used to evaluate the strength of a match between two sequences (the higher the score or the lower the probability value, the greater the identity between two sequences).

The programs described above for the assembly and analysis of full length polynucleotide and polypeptide sequences were also used to identify polynucleotide sequence fragments from SEQ ID NO:4-6. Fragments from about 20 to about 4000 nucleotides which are useful in hybridization and amplification technologies are described in Table 4, column 4.

IV. Identification and Editing of Coding Sequences from Genomic DNA

Putative apoptosis regulators were initially identified by running the Genscan gene identification program against public genomic sequence databases (e.g., gbpri and gbhtg). Genscan is a general-purpose gene identification program which analyzes genomic DNA sequences from a variety of organisms (See Burge, C. and S. Karlin (1997) J. Mol. Biol. 268:78-94, and Burge, C. and S. Karlin (1998) Curr. Opin. Struct. Biol. 8:346-354). The program concatenates predicted exons to form an assembled cDNA sequence extending from a methionine to a stop codon. The output of Genscan is a FASTA database of polynucleotide and polypeptide sequences. The maximum range of sequence for Genscan to analyze at once was set to 30 kb. To determine which of these Genscan predicted cDNA

sequences encode apoptosis regulators, the encoded polypeptides were analyzed by querying against PFAM models for apoptosis regulators. Potential apoptosis regulators were also identified by homology to Incyte cDNA sequences that had been annotated as apoptosis regulators. These selected Genscan-predicted sequences were then compared by BLAST analysis to the genpept and gbpri public databases. Where necessary, the Genscan-predicted sequences were then edited by comparison to the top BLAST hit from genpept to correct errors in the sequence predicted by Genscan, such as extra or omitted exons. BLAST analysis was also used to find any Incyte cDNA or public cDNA coverage of the Genscan-predicted sequences, thus providing evidence for transcription. When Incyte cDNA coverage was available, this information was used to correct or confirm the Genscan predicted sequence. Full length polynucleotide sequences were obtained by assembling Genscan-predicted coding sequences with Incyte cDNA sequences and/or public cDNA sequences using the assembly process described in Example III. Alternatively, full length polynucleotide sequences were derived entirely from edited or unedited Genscan-predicted coding sequences.

V. Assembly of Genomic Sequence Data with cDNA Sequence Data

"Stitched" Sequences

Partial cDNA sequences were extended with exons predicted by the Genscan gene identification program described in Example IV. Partial cDNAs assembled as described in Example III were mapped to genomic DNA and parsed into clusters containing related cDNAs and Genscan exon predictions from one or more genomic sequences. Each cluster was analyzed using an algorithm based on graph theory and dynamic programming to integrate cDNA and genomic information, generating possible splice variants that were subsequently confirmed, edited, or extended to create a full length sequence. Sequence intervals in which the entire length of the interval was present on more than one sequence in the cluster were identified, and intervals thus identified were considered to be equivalent by transitivity. For example, if an interval was present on a cDNA and two genomic sequences, then all three intervals were considered to be equivalent. This process allows unrelated but consecutive genomic sequences to be brought together, bridged by cDNA sequence. Intervals thus identified were then "stitched" together by the stitching algorithm in the order that they appear along their parent sequences to generate the longest possible sequence, as well as sequence variants. Linkages between intervals which proceed along one type of parent sequence (cDNA to cDNA or genomic sequence to genomic sequence) were given preference over linkages which change parent type (cDNA to genomic sequence). The resultant stitched sequences were translated and compared by BLAST analysis to the genpept and gbpri public databases. Incorrect exons predicted by Genscan were corrected by comparison to the top BLAST hit from genpept. Sequences were further extended with additional cDNA sequences, or by inspection of genomic DNA, when necessary.

"Stretched" Sequences

Partial DNA sequences were extended to full length with an algorithm based on BLAST analysis. First, partial cDNAs assembled as described in Example III were queried against public databases such as the GenBank primate, rodent, mammalian, vertebrate, and eukaryote databases using the BLAST program. The nearest GenBank protein homolog was then compared by BLAST analysis to either Incyte cDNA sequences or GenScan exon predicted sequences described in Example IV. A chimeric protein was generated by using the resultant high-scoring segment pairs (HSPs) to map the translated sequences onto the GenBank protein homolog. Insertions or deletions may occur in the chimeric protein with respect to the original GenBank protein homolog. The GenBank protein homolog, the chimeric protein, or both were used as probes to search for homologous genomic sequences from the public human genome databases. Partial DNA sequences were therefore "stretched" or extended by the addition of homologous genomic sequences. The resultant stretched sequences were examined to determine whether it contained a complete gene.

VI. Chromosomal Mapping of APRG Encoding Polynucleotides

The sequences which were used to assemble SEQ ID NO:4-6 were compared with sequences from the Incyte LIFESEQ database and public domain databases using BLAST and other implementations of the Smith-Waterman algorithm. Sequences from these databases that matched SEQ ID NO:4-6 were assembled into clusters of contiguous and overlapping sequences using assembly algorithms such as Phrap (Table 7). Radiation hybrid and genetic mapping data available from public resources such as the Stanford Human Genome Center (SHGC), Whitehead Institute for Genome Research (WIGR), and Généthon were used to determine if any of the clustered sequences had been previously mapped. Inclusion of a mapped sequence in a cluster resulted in the assignment of all sequences of that cluster, including its particular SEQ ID NO:, to that map location.

Map locations are represented by ranges, or intervals, of human chromosomes. The map position of an interval, in centiMorgans, is measured relative to the terminus of the chromosome's p-arm. (The centiMorgan (cM) is a unit of measurement based on recombination frequencies between chromosomal markers. On average, 1 cM is roughly equivalent to 1 megabase (Mb) of DNA in humans, although this can vary widely due to hot and cold spots of recombination.) The cM distances are based on genetic markers mapped by Généthon which provide boundaries for radiation hybrid markers whose sequences were included in each of the clusters. Human genome maps and other resources available to the public, such as the NCBI "GeneMap'99" World Wide Web site (http://www.ncbi.nlm.nih.gov/genemap/), can be employed to determine if previously identified disease genes map within or in proximity to the intervals indicated above.

VII. Analysis of Polynucleotide Expression

Northern analysis is a laboratory technique used to detect the presence of a transcript of a gene and involves the hybridization of a labeled nucleotide sequence to a membrane on which RNAs from a particular cell type or tissue have been bound. (See, e.g., Sambrook, supra, ch. 7; Ausubel (1995) supra, ch. 4 and 16.)

Analogous computer techniques applying BLAST were used to search for identical or related molecules in cDNA databases such as GenBank or LIFESEQ (Incyte Genomics). This analysis is much faster than multiple membrane-based hybridizations. In addition, 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 search is the product score, which is defined as:

$$\frac{\text{BLAST Score} \times \text{Percent Identity}}{5 \times \text{minimum } \{\text{length}(\text{Seq. } 1), \text{length}(\text{Seq. } 2)\}}$$

The product score takes into account both the degree of similarity between two sequences and the length of the sequence match. The product score is a normalized value between 0 and 100, and is calculated as follows: the BLAST score is multiplied by the percent nucleotide identity and the product is divided by (5 times the length of the shorter of the two sequences). The BLAST score is calculated by assigning a score of +5 for every base that matches in a high-scoring segment pair (HSP), and -4 for every mismatch. Two sequences may share more than one HSP (separated by gaps). If there is more than one HSP, then the pair with the highest BLAST score is used to calculate the product score. The product score represents a balance between fractional overlap and quality in a BLAST alignment. For example, a product score of 100 is produced only for 100% identity over the entire length of the shorter of the two sequences being compared. A product score of 70 is produced either by 100% identity and 70% overlap at one end, or by 88% identity and 100% overlap at the other. A product score of 50 is produced either by 100% identity and 50% overlap at one end, or 79% identity and 100% overlap.

Alternatively, polynucleotide sequences encoding APRG are analyzed with respect to the tissue sources from which they were derived. For example, some full length sequences are assembled, at least in part, with overlapping Incyte cDNA sequences (see Example II). Each cDNA sequence is derived from a cDNA library constructed from a human tissue. Each human tissue is classified into one of the following organ/tissue categories: cardiovascular system; connective tissue; digestive system; embryonic structures; endocrine system; exocrine glands; genitalia, female; genitalia, male; germ cells; hemic and immune system; liver; musculoskeletal system; nervous system; pancreas; respiratory system; sense organs; skin; stomatognathic system; unclassified/mixed; or urinary tract. The number of

libraries in each category is counted and divided by the total number of libraries across all categories. Similarly, each human tissue is classified into one of the following disease/condition categories: cancer, cell line, developmental, inflammation, neurological, trauma, cardiovascular, pooled, and other, and the number of libraries in each category is counted and divided by the total number of libraries across all categories. The resulting percentages reflect the tissue- and disease-specific expression of cDNA encoding APRG. cDNA sequences and cDNA library/tissue information are found in the LIFESEQ GOLD database (Incyte Genomics, Palo Alto CA).

VIII. Extension of APRG Encoding Polynucleotides

Full length polynucleotide sequences were also produced by extension of an appropriate fragment of the full length molecule using oligonucleotide primers designed from this fragment. One primer was synthesized to initiate 5' extension of the known fragment, and the other primer was synthesized to initiate 3' extension of the known fragment. The initial primers were designed using OLIGO 4.06 software (National Biosciences), or another appropriate program, to be about 22 to 30 nucleotides in length, to have a GC content of about 50% or more, and to anneal to the target sequence at temperatures of about 68°C to about 72°C. Any stretch of nucleotides which would result in hairpin structures and primer-primer dimerizations was avoided.

Selected human cDNA libraries were used to extend the sequence. If more than one extension was necessary or desired, additional or nested sets of primers were designed.

High fidelity amplification was obtained by PCR using methods well known in the art. PCR was performed in 96-well plates using the PTC-200 thermal cycler (MJ Research, Inc.). The reaction mix contained DNA template, 200 nmol of each primer, reaction buffer containing Mg^{2+}, $(NH_4)_2SO_4$, and 2-mercaptoethanol, Taq DNA polymerase (Amersham Pharmacia Biotech), ELONGASE enzyme (Life Technologies), and Pfu DNA polymerase (Stratagene), with the following parameters for primer pair PCI A and PCI B: 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, the parameters for primer pair T7 and SK+ 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 quantitation reagent (0.25% (v/v) PICOGREEN; Molecular Probes, Eugene OR) dissolved in 1X TE and 0.5 µl of undiluted PCR product into each well of an opaque fluorimeter plate (Corning Costar, Acton MA), 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 gel to determine which reactions were successful in extending the sequence.

The extended nucleotides were desalted and concentrated, transferred to 384-well plates, digested with CviJI 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 nucleotides were separated on low concentration (0.6 to 0.8 %) agarose gels, fragments were excised, and agar digested with Agar ACE (Promega). Extended clones were religated using T4 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 transfected 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 carb liquid media.

The cells were lysed, and DNA was amplified by PCR using Taq DNA polymerase (Amersham Pharmacia Biotech) and Pfu DNA polymerase (Stratagene) with the following parameters: Step 1: 94°C, 3 min; Step 2: 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 same conditions as described above. Samples were diluted with 20% dimethysulfoxide (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 (Applied Biosystems).

In like manner, full length polynucleotide sequences are verified using the above procedure or are used to obtain 5' regulatory sequences using the above procedure along with oligonucleotides designed for such extension, and an appropriate genomic library.

IX. Labeling and Use of Individual Hybridization Probes

Hybridization probes derived from SEQ ID NO:4-6 are employed to screen cDNAs, genomic DNAs, or mRNAs. Although the labeling of oligonucleotides, consisting of about 20 base pairs, is specifically described, essentially the same procedure is used with larger nucleotide fragments. Oligonucleotides are designed using state-of-the-art software such as OLIGO 4.06 software (National Biosciences) and labeled by combining 50 pmol of each oligomer, 250 μCi of [γ-^{32}P] adenosine triphosphate (Amersham Pharmacia Biotech), and T4 polynucleotide kinase (DuPont NEN, Boston MA). The labeled oligonucleotides are substantially purified using a SEPHADEX G-25 superfine size exclusion dextran bead column (Amersham Pharmacia Biotech). An aliquot containing 10^7 counts per minute of the labeled probe is used in a typical membrane-based hybridization analysis of human genomic DNA digested with one of the following endonucleases: Ase I, Bgl II, Eco RI, Pst I, Xba I, or

Pvu II (DuPont NEN).

The DNA from each digest is fractionated on a 0.7% agarose gel and transferred to nylon membranes (Nytran Plus, Schleicher & Schuell, Durham NH). Hybridization is carried out for 16 hours at 40°C. To remove nonspecific signals, blots are sequentially washed at room temperature under conditions of up to, for example, 0.1 x saline sodium citrate and 0.5% sodium dodecyl sulfate. Hybridization patterns are visualized using autoradiography or an alternative imaging means and compared.

X. Microarrays

The linkage or synthesis of array elements upon a microarray can be achieved utilizing photolithography, piezoelectric printing (ink-jet printing, See, e.g., Baldeschweiler, supra.), mechanical microspotting technologies, and derivatives thereof. The substrate in each of the aforementioned technologies should be uniform and solid with a non-porous surface (Schena (1999), supra). Suggested substrates include silicon, silica, glass slides, glass chips, and silicon wafers. Alternatively, a procedure analogous to a dot or slot blot may also be used to arrange and link elements to the surface of a substrate using thermal, UV, chemical, or mechanical bonding procedures. A typical array may be produced using available methods and machines well known to those of ordinary skill in the art and may contain any appropriate number of elements. (See, e.g., Schena, M. et al. (1995) Science 270:467-470; Shalon, D. et al. (1996) Genome Res. 6:639-645; Marshall, A. and J. Hodgson (1998) Nat. Biotechnol. 16:27-31.)

Full length cDNAs, Expressed Sequence Tags (ESTs), or fragments or oligomers thereof may comprise the elements of the microarray. Fragments or oligomers suitable for hybridization can be selected using software well known in the art such as LASERGENE software (DNASTAR). The array elements are hybridized with polynucleotides in a biological sample. The polynucleotides in the biological sample are conjugated to a fluorescent label or other molecular tag for ease of detection. After hybridization, nonhybridized nucleotides from the biological sample are removed, and a fluorescence scanner is used to detect hybridization at each array element. Alternatively, laser desorption and mass spectrometry may be used for detection of hybridization. The degree of complementarity and the relative abundance of each polynucleotide which hybridizes to an element on the microarray may be assessed. In one embodiment, microarray preparation and usage is described in detail below.

Tissue or Cell Sample Preparation

Total RNA is isolated from tissue samples using the guanidinium thiocyanate method and poly(A)$^+$ RNA is purified using the oligo-(dT) cellulose method. Each poly(A)$^+$ RNA sample is reverse transcribed using MMLV reverse-transcriptase, 0.05 pg/μl oligo-(dT) primer (21mer), 1X first

strand buffer, 0.03 units/μl RNase inhibitor, 500 μM dATP, 500 μM dGTP, 500 μM dTTP, 40 μM dCTP, 40 μM dCTP-Cy3 (BDS) or dCTP-Cy5 (Amersham Pharmacia Biotech). The reverse transcription reaction is performed in a 25 ml volume containing 200 ng poly(A)$^+$ RNA with GEMBRIGHT kits (Incyte). Specific control poly(A)$^+$ RNAs are synthesized by in vitro transcription from non-coding yeast genomic DNA. After incubation at 37°C for 2 hr, each reaction sample (one with Cy3 and another with Cy5 labeling) is treated with 2.5 ml of 0.5M sodium hydroxide and incubated for 20 minutes at 85°C to the stop the reaction and degrade the RNA. Samples are purified using two successive CHROMA SPIN 30 gel filtration spin columns (CLONTECH Laboratories, Inc. (CLONTECH), Palo Alto CA) and after combining, both reaction samples are ethanol precipitated using 1 ml of glycogen (1 mg/ml), 60 ml sodium acetate, and 300 ml of 100% ethanol. The sample is then dried to completion using a SpeedVAC (Savant Instruments Inc., Holbrook NY) and resuspended in 14 μl 5X SSC/0.2% SDS.

Microarray Preparation

Sequences of the present invention are used to generate array elements. Each array element is amplified from bacterial cells containing vectors with cloned cDNA inserts. PCR amplification uses primers complementary to the vector sequences flanking the cDNA insert. Array elements are amplified in thirty cycles of PCR from an initial quantity of 1-2 ng to a final quantity greater than 5 μg. Amplified array elements are then purified using SEPHACRYL-400 (Amersham Pharmacia Biotech).

Purified array elements are immobilized on polymer-coated glass slides. Glass microscope slides (Corning) are cleaned by ultrasound in 0.1% SDS and acetone, with extensive distilled water washes between and after treatments. Glass slides are etched in 4% hydrofluoric acid (VWR Scientific Products Corporation (VWR), West Chester PA), washed extensively in distilled water, and coated with 0.05% aminopropyl silane (Sigma) in 95% ethanol. Coated slides are cured in a 110°C oven.

Array elements are applied to the coated glass substrate using a procedure described in US Patent No. 5,807,522, incorporated herein by reference. 1 μl of the array element DNA, at an average concentration of 100 ng/μl, is loaded into the open capillary printing element by a high-speed robotic apparatus. The apparatus then deposits about 5 nl of array element sample per slide.

Microarrays are UV-crosslinked using a STRATALINKER UV-crosslinker (Stratagene). Microarrays are washed at room temperature once in 0.2% SDS and three times in distilled water. Non-specific binding sites are blocked by incubation of microarrays in 0.2% casein in phosphate buffered saline (PBS) (Tropix, Inc., Bedford MA) for 30 minutes at 60° C followed by washes in 0.2% SDS and distilled water as before.

Hybridization

Hybridization reactions contain 9 µl of sample mixture consisting of 0.2 µg each of Cy3 and Cy5 labeled cDNA synthesis products in 5X SSC, 0.2% SDS hybridization buffer. The sample mixture is heated to 65°C for 5 minutes and is aliquoted onto the microarray surface and covered with an 1.8 cm^2 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 5X SSC in a corner of the chamber. The chamber containing the arrays is incubated for about 6.5 hours at 60°C. The arrays are washed for 10 min at 45°C in a first wash buffer (1X SSC, 0.1% SDS), three times for 10 minutes each at 45°C in a second wash buffer (0.1X SSC), and dried.

Detection

Reporter-labeled hybridization complexes are detected with a microscope equipped with an Innova 70 *mixed gas* 10 W laser (Coherent, Inc., 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, Inc., 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. The 1.8 cm x 1.8 cm array used in the present example is scanned with a resolution of 20 micrometers.

In two separate scans, a mixed gas multiline laser excites the two fluorophores sequentially. 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. Each array is typically scanned twice, one scan per fluorophore using the appropriate filters at the laser source, although the apparatus is capable of recording the spectra from both fluorophores simultaneously.

The sensitivity of the scans is typically calibrated using the signal intensity generated by a cDNA control species added to the sample mixture at a known concentration. 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. When two samples from different sources (e.g., representing test and control cells), each labeled with a different fluorophore, are hybridized to a single array for the purpose of identifying genes that are differentially expressed, the calibration is done by labeling samples of the calibrating cDNA with the two fluorophores and adding identical amounts of each to the hybridization mixture.

The output of the photomultiplier tube is digitized using a 12-bit RTI-835H analog-to-digital (A/D) conversion board (Analog Devices, Inc., 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 GEMTOOLS gene expression analysis program (Incyte).

XI. Complementary Polynucleotides

Sequences complementary to the APRG-encoding sequences, or any parts thereof, are used to detect, decrease, or inhibit expression of naturally occurring APRG. Although use of oligonucleotides comprising from about 15 to 30 base pairs is described, essentially the same procedure is used with smaller or with larger sequence fragments. Appropriate oligonucleotides are designed using OLIGO 4.06 software (National Biosciences) and the coding sequence of APRG. To inhibit transcription, a complementary oligonucleotide is designed from the most unique 5' sequence and used to prevent promoter binding to the coding sequence. To inhibit translation, a complementary oligonucleotide is designed to prevent ribosomal binding to the APRG-encoding transcript.

XII. Expression of APRG

Expression and purification of APRG is achieved using bacterial or virus-based expression systems. For expression of APRG in bacteria, cDNA is subcloned into an appropriate 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 suitable bacterial hosts, e.g., BL21(DE3). Antibiotic resistant bacteria express APRG upon induction with isopropyl beta-D-thiogalactopyranoside (IPTG). Expression of APRG in eukaryotic cells is achieved by infecting insect or mammalian cell lines with recombinant Autographica californica nuclear polyhedrosis virus (AcMNPV), commonly known as baculovirus. The nonessential polyhedrin gene of baculovirus is replaced with cDNA encoding APRG 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. Recombinant baculovirus is used to infect Spodoptera frugiperda (Sf9) insect cells in most cases, or human hepatocytes, in some cases. Infection of the latter requires additional genetic modifications to baculovirus. (See Engelhard, E.K. et al. (1994) Proc. Natl. Acad. Sci. USA 91:3224-3227; Sandig, V. et al. (1996) Hum. Gene Ther. 7:1937-1945.)

In most expression systems, APRG is synthesized as a fusion protein with, e.g., glutathione S-transferase (GST) or a peptide epitope tag, such as FLAG or 6-His, permitting rapid, single-step, affinity-based purification of recombinant fusion protein from crude cell lysates. GST, a 26-kilodalton enzyme from Schistosoma japonicum, 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 APRG at specifically engineered sites. FLAG, an 8-amino acid peptide, enables immunoaffinity 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 (1995, supra, ch. 10 and 16). Purified APRG obtained by these methods can be used directly in the assays shown in Examples XVI and XVII, where applicable.

XIII. Functional Assays

APRG function is assessed by expressing the sequences encoding APRG at physiologically elevated levels in mammalian cell culture systems. cDNA is subcloned into a mammalian expression vector containing a strong promoter that drives high levels of cDNA expression. Vectors of choice include PCMV SPORT (Life Technologies) and PCR3.1 (Invitrogen, Carlsbad CA), both of which contain the cytomegalovirus promoter. 5-10 μg of recombinant vector are transiently transfected into a human cell line, for example, an endothelial or hematopoietic cell line, using either liposome formulations or electroporation. 1-2 μg of an additional plasmid containing sequences encoding a marker protein are co-transfected. Expression of a marker protein provides a means to distinguish transfected cells from nontransfected cells and is a reliable predictor of cDNA expression from the recombinant vector. Marker proteins of choice include, e.g., Green Fluorescent Protein (GFP; Clontech), CD64, or a CD64-GFP fusion protein. Flow cytometry (FCM), an automated, laser optics-based technique, is used to identify transfected cells expressing GFP or CD64-GFP and to evaluate the apoptotic state of the cells and other cellular properties. FCM detects and quantifies the uptake of fluorescent molecules that diagnose events preceding or coincident with cell death. These events include changes in nuclear DNA content as measured by staining of DNA with propidium iodide; changes in cell size and granularity as measured by forward light scatter and 90 degree side light scatter; down-regulation of DNA synthesis as measured by decrease in bromodeoxyuridine uptake; alterations in expression of cell surface and intracellular proteins as measured by reactivity with specific antibodies; and alterations in plasma membrane composition as measured by the binding of fluorescein-conjugated Annexin V protein to the cell surface. Methods in flow cytometry are discussed in Ormerod, M.G. (1994) Flow Cytometry, Oxford, New York NY.

The influence of APRG on gene expression can be assessed using highly purified populations of cells transfected with sequences encoding APRG and either CD64 or CD64-GFP. CD64 and CD64-GFP are expressed on the surface of transfected cells and bind to conserved regions of human immunoglobulin G (IgG). Transfected cells are efficiently separated from nontransfected cells using magnetic beads coated with either human IgG or antibody against CD64 (DYNAL, Lake Success NY). mRNA can be purified from the cells using methods well known by those of skill in the art. Expression of mRNA encoding APRG and other genes of interest can be analyzed by northern analysis or microarray techniques.

XIV. Production of APRG Specific Antibodies

APRG substantially purified using polyacrylamide gel electrophoresis (PAGE; see, e.g., Harrington, M.G. (1990) Methods Enzymol. 182:488-495), or other purification techniques, is used to immunize rabbits and to produce antibodies using standard protocols.

Alternatively, the APRG amino acid sequence is analyzed using LASERGENE software (DNASTAR) to determine regions of high immunogenicity, and a corresponding oligopeptide is synthesized and used to raise antibodies by means known to those of skill in the art. Methods for selection of appropriate epitopes, such as those near the C-terminus or in hydrophilic regions are well described in the art. (See, e.g., Ausubel, 1995, supra, ch. 11.)

Typically, oligopeptides of about 15 residues in length are synthesized using an ABI 431A peptide synthesizer (Applied Biosystems) using FMOC chemistry and coupled to KLH (Sigma-Aldrich, St. Louis MO) by reaction with N-maleimidobenzoyl-N-hydroxysuccinimide ester (MBS) to increase immunogenicity. (See, e.g., Ausubel, 1995, supra.) Rabbits are immunized with the oligopeptide-KLH complex in complete Freund's adjuvant. Resulting antisera are tested for antipeptide and anti-APRG activity by, for example, binding the peptide or APRG to a substrate, blocking with 1% BSA, reacting with rabbit antisera, washing, and reacting with radio-iodinated goat anti-rabbit IgG.

XV. Purification of Naturally Occurring APRG Using Specific Antibodies

Naturally occurring or recombinant APRG is substantially purified by immunoaffinity chromatography using antibodies specific for APRG. An immunoaffinity column is constructed by covalently coupling anti-APRG antibody to an activated chromatographic resin, such as CNBr-activated SEPHAROSE (Amersham Pharmacia Biotech). After the coupling, the resin is blocked and washed according to the manufacturer's instructions.

Media containing APRG are passed over the immunoaffinity column, and the column is washed under conditions that allow the preferential absorbance of APRG (e.g., high ionic strength buffers in the presence of detergent). The column is eluted under conditions that disrupt antibody/APRG binding (e.g., a buffer of pH 2 to pH 3, or a high concentration of a chaotrope, such as urea or thiocyanate ion),

and APRG is collected.

XVI. Identification of Molecules Which Interact with APRG

APRG, or biologically active fragments thereof, are labeled with ^{125}I Bolton-Hunter reagent. (See, e.g., Bolton A.E. and W.M. Hunter (1973) Biochem. J. 133:529-539.) Candidate molecules previously arrayed in the wells of a multi-well plate are incubated with the labeled APRG, washed, and any wells with labeled APRG complex are assayed. Data obtained using different concentrations of APRG are used to calculate values for the number, affinity, and association of APRG with the candidate molecules.

Alternatively, molecules interacting with APRG are analyzed using the yeast two-hybrid system as described in Fields, S. and O. Song (1989) Nature 340:245-246, or using commercially available kits based on the two-hybrid system, such as the MATCHMAKER system (Clontech).

APRG may also be used in the PATHCALLING process (CuraGen Corp., New Haven CT) which employs the yeast two-hybrid system in a high-throughput manner to determine all interactions between the proteins encoded by two large libraries of genes (Nandabalan, K. et al. (2000) U.S. Patent No. 6,057,101).

XVII. Demonstration of APRG Activity

An assay for APRG activity measures cell proliferation as the amount of newly initiated DNA synthesis in Swiss mouse 3T3 cells. A plasmid containing polynucleotides encoding APRG is transfected into quiescent 3T3 cultured cells using methods well known in the art. The transiently transfected cells are then incubated in the presence of [^{3}H]thymidine, a radioactive DNA precursor. Where applicable, varying amounts of APRG ligand are added to the transfected cells. Incorporation of [^{3}H]thymidine into acid-precipitable DNA is measured over an appropriate time interval, and the amount incorporated is directly proportional to the amount of newly synthesized DNA.

An alternative assay for APRG activity measures the induction of apoptosis when APRG is expressed at physiologically elevated levels in mammalian cell culture systems. cDNA is subcloned into a mammalian expression vector containing a strong promoter that drives high levels of cDNA

Green Fluorescent Protein (GFP) (Clontech, Palo Alto, CA), CD64, or a CD64-GFP fusion protein. Flow cytometry (FCM), an automated, laser optics-based technique, is used to identify transfected cells expressing GFP or CD64-GFP and to evaluate their apoptotic state. FCM detects and quantifies the uptake of fluorescent molecules that diagnose events preceding or coincident with cell death. These events include changes in nuclear DNA content as measured by staining of DNA with propidium iodide; changes in cell size and granularity as measured by forward light scatter and 90 degree side light scatter; down-regulation of DNA synthesis as measured by decrease in bromodeoxyuridine uptake; alterations in expression of cell surface and intracellular proteins as measured by reactivity with specific antibodies; and alterations in plasma membrane composition as measured by the binding of fluorescein-conjugated Annexin V protein to the cell surface.

Alternatively, APRG activity may be measured by the induction of growth arrest when APRG is expressed at physiologically elevated levels in transformed mammalian cell lines. APRG cDNA is subcloned into a mammalian expression vector containing a strong promoter that drives high levels of cDNA expression, and these constructs are stably transfected into a transformed cell line, such as NIH 3T6 or C6, using methods known in the art. An additional plasmid, containing sequences which encode a selectable marker, such as hygromycin resistance, are co-transfected. Expression of a marker protein provides a means to distinguish transfected cells from nontransfected cells and is a reliable predictor of cDNA expression from the recombinant vector. Cells expressing APRG are compared with control cells, either non-transfected or transfected with vector alone, for characteristics associated with growth arrest. Such characteristics can include, but are not limited to, a reduction in [^{3}H]-thymidine incorporation into newly synthesized DNA, lower doubling and generation times, and decreased culture saturation density.

Alternatively, an assay for APRG activity uses radiolabeled nucleotides, such as [α^{32}P]ATP, to measure either the incorporation of radiolabel into DNA during DNA synthesis, or fragmentation of DNA that accompanies apoptosis. Mammalian cells are transfected with plasmid containing cDNA encoding APRG by methods well known in the art. Cells are then incubated with radiolabeled nucleotide for various lengths of time. Chromosomal DNA is collected, and radioactivity detected using a scintillation counter. Incorporation of radiolabel into chromosomal DNA is proportional to the degree of stimulation of the cell cycle. To determine if APRG promotes apoptosis, chromosomal DNA is collected as above, and analyzed using polyacrylamide gel electrophoresis, by methods well known in the art. Fragmentation of DNA is quantified by comparison to untransfected control cells, and is proportional to the apoptotic activity of APRG.

Various modifications and variations of the described methods and systems of the invention will

be apparent to those skilled in the art without departing from the scope and spirit of the invention. Although the invention has been described in connection with certain embodiments, it should be understood that the invention as claimed should not be unduly limited to such specific embodiments. Indeed, various modifications of the described modes for carrying out the invention which are obvious to those skilled in molecular biology or related fields are intended to be within the scope of the following claims.

Table 1

Incyte Project ID	Polypeptide SEQ ID NO:	Incyte Polypeptide ID	Polynucleotide SEQ ID NO:	Incyte Polynucleotide ID
3102521	1	3102521CD1	4	3102521CB1
7474984	2	7474984CD1	5	7474984CB1
2458545	3	2458545CD1	6	2458545CB1

Table 2

Polypeptide SEQ ID NO:	Incyte Polypeptide ID	GenBank ID NO:	Probability score	GenBank Homolog
1	3102523CD1	g7671638	5.10E-70	(Danio rerio) Deddl (death effector domain protein) Inohara, N. and G. Nunez (2000) Genes with homology to mammalian apoptosis regulators identified in zebrafish Cell Death Differ. 7:509-510
2	7474984CD1	gi01198309	1.10E-66	(Homo sapiens) (AF298548) caspase recruitment domain protein 7
3	2458545CD1	g5776545	2.10E-73	taxi-binding protein TXBP151 (Homo sapiens)
		gi12445470	1.00E-116	(gi) (Xenopus laevis) nuclear domain-10 protein NDP52 Korioth, F. et al. (1995) Molecular characterization of NDP52, a novel protein of the nuclear domain 10, which is redistributed upon virus infection and interferon treatment. J. Cell Biol. 130:1-13.

Table 3

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
1	3102521CD1	326	S64 S122 S123 S126 S136 S219 S257 S298 T38 T154 T323 Y258		Signal peptide: M1-A56	SPSCAN
					death effector domain: I24-Y110	HMMER-PFAM
					FLAMDD1 (death effector domain protein) PD184655: D243-L312	BLAST-PRODOM
					KB05 protein, FLDED1 (death effector domain protein) PD124668: W13-Q182	BLAST-PRODOM
2	7474984CD1	655	S18 S207 S217 S268 S316 S448 S463 S468 S491 S521 S525 S527 S554 S566 S585 S603 T177 T282 T291 T36 T377 T634 T640 T649	N380 N558 N583	Transmembrane domain: Y203-W225	HMMER
					PYRIN domain PD136606: A6-Y98 (E = 1.1e-5)	BLAST-PRODOM
					ATP/GTP-binding site motif A (P-loop): G171-T180	MOTIFS
3	2458545CD1	691	T28 S43 S68 S72 T129 S156 S208 T239 S282 S305 T376 S383 S468 T520 T521 T537 S539 S543 S593 S595 S597 S612 T639 T652 S667 S689 T190 S241 T257	N618	RGD motif. Rgd:R557-D559	MOTIFS
					Leucine zipper motif, Leucine_zipper:L163-L184 L475-L496 L482-L503	MOTIFS
					cGMP kinases putative ATP-binding protein domain, cGMP_kinases:A146-G172	HMMER-PFAM
					Zinc finger, C2H2 type zf-C2H2:K634-H679	HMMER-PFAM
					NDP52 NUCLEAR DOMAIN PROTEIN PD034365:V15-D134	BLIMPS-PRODOM

Table 4

Polynucleotide SEQ ID NO:	Incyte Polynucleotide ID	Sequence Length	Selected Fragment(s)	Sequence Fragments	5' Position	3' Position
4	3102551CB1	1230	1175-1230	7207023BY1	503	1230
				7012482H1 (KIDNNOC01)	1	668
5	7474984CB1	2460	1-945, 2400-2460, 1355-2193	GNH_g8099799_000020_006	747	1974
				5509411602	204	1051
				5909233H1	1659	2184
				7248257ED1	1	581
				5800520HH1	1661	2460
6	3450545CB1	2986	1-10, 2494-2538, 2936-2986, 2183-2225, 959-1210	4143481H1 (BRSTNMT01)	801	1090
				1915072H1 (PROSTUT04)	1460	1727
				2582205F6 (KIDNTUT13)	900	1422
				1679569T7 (STOMFET01)	2304	2986
				1679569F6 (STOMFET01)	1749	2371
				3341764F6 (SPLNNOT09)	534	1086
				4033751H1 (BRAYTUT03)	1370	1674
				3472904F6 (LUNGNOT37)	418	726
				4333581H1 (KIDCTUT04)	504	788
				1413211F1 (KIDNNOT09)	1552	2076
				2455545F6 (ENDANOT01)	1	473
				4424620H1 (BRAFDIT01)	1187	1954

Table 5

Polynucleotide SEQ ID NO:	Incyte Project ID	Representative Library
4	3102321CB1	ADRSTUT05
6	2458545CB1	BRSTNOT06

Table 6

Library	Vector	Library Description
ADRGTUT05	pINCY	Library was constructed using RNA isolated from adrenal tumor tissue removed from a 52-year-old Caucasian female during a unilateral adrenalectomy. Pathology indicated a pheochromocytoma.
BRSTNOT04	pSPORT1	Library was constructed using RNA isolated from breast tissue removed from a 62-year-old East Indian female during a unilateral extended simple mastectomy. Pathology for the associated tumor tissue indicated an invasive grade 3 ductal carcinoma. Patient history included benign hypertension, hypercholesterolemia, and hematuria. Family history included cerebrovascular and cardiovascular disease, hypercholesterolemia, and liver cancer.

Table 7

Program	Description	Reference	Parameter Threshold
ABI FACTURA	A program that removes vector sequences and masks ambiguous bases in nucleic acid sequences.	Applied Biosystems, Foster City, CA.	
ABI/PARACEL FDF	A Fast Data Finder useful in comparing and annotating amino acid or nucleic acid sequences.	Applied Biosystems, Foster City, CA; Paracel Inc., Pasadena, CA.	Mismatch <50%
ABI AutoAssembler	A program that assembles nucleic acid sequences.	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) J. Mol. Biol. 215:403-410; Altschul, S.F. et al. (1997) Nucleic Acids Res. 25:3389-3402.	*ESTs:* Probability value= 1.0E-8 or less *Full Length sequences:* Probability value= 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 comprises as least five functions: fasta, tfasta, fastx, tfastx, and ssearch.	Pearson, W.R. and D.J. Lipman (1988) Proc. Natl. Acad Sci. USA 85:2444-2448; Pearson, W.R. (1990) Methods Enzymol. 183:63-98; and Smith, T.F. and M.S. Waterman (1981) Adv. Appl. Math. 2:482-489.	*ESTs:* fasta E value=1.06E-6 *Assembled ESTs:* fasta Identity= 95% or greater and Match length=200 bases or greater; fasta E value=1.0E-8 or less *Full Length sequences:* fasta score=100 or greater
BLIMPS	A BLocks IMProved Searcher that matches a sequence against those in BLOCKS, PRINTS, DOMO, PRODOM, and PFAM databases to search for gene families, sequence homology, and structural fingerprint regions.	Henikoff, S. and J.G. Henikoff (1991) Nucleic Acids Res. 19:6565-6572; Henikoff, J.G. and S. Henikoff (1996) Methods Enzymol. 266:88-105; and Attwood, T.K. et al. (1997) J. Chem. Inf. Comput. Sci. 37:417-424.	Probability value= 1.0E-3 or less
HMMER	An algorithm for searching a query sequence against hidden Markov model (HMM)-based databases of protein family consensus sequences, such as PFAM.	Krogh, A. et al. (1994) J. Mol. Biol. 235:1501-1531; Sonnhammer, E.L.L. et al. (1998) Nucleic Acids Res. 26:320-322; Durbin, R. et al. (1998) Our World View, in a Nutshell, Cambridge Univ. Press, pp. 1-350.	*PFAM hits:* Probability value= 1.0E-3 or less *Signal peptide hits:* Score= 0 or greater

Table 7 (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.	Gribskov, M. et al. (1988) CABIOS 4:61-66; Gribskov, M. et al. (1989) Methods Enzymol. 183:146-159; Bairoch A. et al. (1997) Nucleic Acids Res. 25:217-221.	Normalized quality score≥GCG-specified "HIGH" value for that particular Prosite motif. Generally, score=1.4-2.1.
Pfrcd	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 Phils 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. (1997) Protein Engineering 10:1-6; Claverie, J.M. and S. Audic (1997) CABIOS 12:431-439.	Score=3.5 or greater
TMAP	A program that uses weight matrices to delineate transmembrane segments on protein sequences and determine orientation.	Persson, B. and P. Argos (1994) J. Mol. Biol. 237:182-192; Persson, B. and P. Argos (1996) Protein Sci. 5:363-371.	
TMHMMER	A program that uses a hidden Markov model (HMM) to delineate transmembrane segments on protein sequences and determine orientation.	Sonnhammer, E.L. et al. (1998) Proc. Sixth Intl. Conf. on Intelligent Systems for Mol. Biol., Glasgow et al., eds., The Am. Assoc. for Artificial Intelligence Press, Menlo Park, CA, pp. 175-182.	
Motifs	A program that searches amino acid sequences for patterns that matched those defined in Prosite.	Bairoch, A. et al. (1997) Nucleic Acids Res. 25:217-221; Wisconsin Package Program Manual, version 9, page M51-59, Genetics Computer Group, Madison, WI.	

What is claimed is:

1. An isolated polypeptide selected from the group consisting of:

a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-3,

b) a naturally occurring polypeptide comprising an amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-3,

c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-3, and

d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-3.

2. An isolated polypeptide of claim 1 selected from the group consisting of SEQ ID NO:1-3.

3. An isolated polynucleotide encoding a polypeptide of claim 1.

4. An isolated polynucleotide encoding a polypeptide of claim 2.

5. An isolated polynucleotide of claim 4 selected from the group consisting of SEQ ID NO:4-6.

6. A recombinant polynucleotide comprising a promoter sequence operably linked to a polynucleotide of claim 3.

7. A cell transformed with a recombinant polynucleotide of claim 6.

8. A transgenic organism comprising a recombinant polynucleotide of claim 6.

9. A method for producing a polypeptide of claim 1, the method comprising:

a) culturing a cell under conditions suitable for expression of the polypeptide, wherein said cell is transformed with a recombinant polynucleotide, and said recombinant polynucleotide comprises a promoter sequence operably linked to a polynucleotide encoding the polypeptide of claim 1, and

b) recovering the polypeptide so expressed.

10. An isolated antibody which specifically binds to a polypeptide of claim 1.

11. An isolated polynucleotide selected from the group consisting of:

a) a polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:4-6,

b) a naturally occurring polynucleotide comprising a polynucleotide sequence at least 90% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:4-6,

c) a polynucleotide complementary to a polynucleotide of a),

d) a polynucleotide complementary to a polynucleotide of b), and

e) an RNA equivalent of a)-d).

12. An isolated polynucleotide comprising at least 60 contiguous nucleotides of a polynucleotide of claim 11.

13. A method for detecting a target polynucleotide in a sample, said target polynucleotide having a sequence of a polynucleotide of claim 11, the method comprising:

a) hybridizing the sample with a probe comprising at least 20 contiguous nucleotides comprising a sequence complementary to said target polynucleotide in the sample, and which probe specifically hybridizes to said target polynucleotide, under conditions whereby a hybridization complex is formed between said probe and said target polynucleotide or fragments thereof, and

b) detecting the presence or absence of said hybridization complex, and, optionally, if present, the amount thereof.

14. A method of claim 13, wherein the probe comprises at least 60 contiguous nucleotides.

15. A method for detecting a target polynucleotide in a sample, said target polynucleotide having a sequence of a polynucleotide of claim 11, the method comprising:

a) amplifying said target polynucleotide or fragment thereof using polymerase chain reaction amplification, and

b) detecting the presence or absence of said amplified target polynucleotide or fragment thereof, and, optionally, if present, the amount thereof.

16. A composition comprising a polypeptide of claim 1 and a pharmaceutically acceptable excipient.

17. A composition of claim 16, wherein the polypeptide has an amino acid sequence selected from the group consisting of SEQ ID NO:1-3.

18. A method for treating a disease or condition associated with decreased expression of functional APRG, comprising administering to a patient in need of such treatment the composition of claim 16.

19. A method for screening a compound for effectiveness as an agonist of a polypeptide of claim 1, the method comprising:

a) exposing a sample comprising a polypeptide of claim 1 to a compound, and

b) detecting agonist activity in the sample.

20. A composition comprising an agonist compound identified by a method of claim 19 and a pharmaceutically acceptable excipient.

21. A method for treating a disease or condition associated with decreased expression of functional APRG, comprising administering to a patient in need of such treatment a composition of claim 20.

22. A method for screening a compound for effectiveness as an antagonist of a polypeptide of claim 1, the method comprising:

a) exposing a sample comprising a polypeptide of claim 1 to a compound, and

b) detecting antagonist activity in the sample.

23. A composition comprising an antagonist compound identified by a method of claim 22 and a pharmaceutically acceptable excipient.

24. A method for treating a disease or condition associated with overexpression of functional APRG, comprising administering to a patient in need of such treatment a composition of claim 23.

25. A method of screening for a compound that specifically binds to the polypeptide of claim 1, said method comprising the steps of:

a) combining the polypeptide of claim 1 with at least one test compound under suitable conditions, and

b) detecting binding of the polypeptide of claim 1 to the test compound, thereby identifying a

compound that specifically binds to the polypeptide of claim 1.

26. A method of screening for a compound that modulates the activity of the polypeptide of claim 1, said method comprising:

a) combining the polypeptide of claim 1 with at least one test compound under conditions permissive for the activity of the polypeptide of claim 1,

b) assessing the activity of the polypeptide of claim 1 in the presence of the test compound, and

c) comparing the activity of the polypeptide of claim 1 in the presence of the test compound with the activity of the polypeptide of claim 1 in the absence of the test compound, wherein a change in the activity of the polypeptide of claim 1 in the presence of the test compound is indicative of a compound that modulates the activity of the polypeptide of claim 1.

27. A method for screening a compound for effectiveness in altering expression of a target polynucleotide, wherein said target polynucleotide comprises a sequence of claim 5, the method comprising:

a) exposing a sample comprising the target polynucleotide to a compound, under conditions suitable for the expression of the target polynucleotide,

b) detecting altered expression of the target polynucleotide, and

c) comparing the expression of the target polynucleotide in the presence of varying amounts of the compound and in the absence of the compound.

28. A method for assessing toxicity of a test compound, said method comprising:

a) treating a biological sample containing nucleic acids with the test compound;

b) hybridizing the nucleic acids of the treated biological sample with a probe comprising at least 20 contiguous nucleotides of a polynucleotide of claim 11 under conditions whereby a specific hybridization complex is formed between said probe and a target polynucleotide in the biological sample, said target polynucleotide comprising a polynucleotide sequence of a polynucleotide of claim 11 or fragment thereof;

c) quantifying the amount of hybridization complex; and

d) comparing the amount of hybridization complex in the treated biological sample with the amount of hybridization complex in an untreated biological sample, wherein a difference in the amount of hybridization complex in the treated biological sample is indicative of toxicity of the test compound.

29. A diagnostic test for a condition or disease associated with the expression of APRG in a biological sample comprising the steps of:

a) combining the biological sample with an antibody of claim 10, under conditions suitable for the antibody to bind the polypeptide and form an antibody:polypeptide complex; and

b) detecting the complex, wherein the presence of the complex correlates with the presence of the polypeptide in the biological sample.

30. The antibody of claim 10, wherein the antibody is:

a) a chimeric antibody,

b) a single chain antibody,

c) a Fab fragment,

d) a $F(ab')_2$ fragment, or

e) a humanized antibody.

31. A composition comprising an antibody of claim 10 and an acceptable excipient.

32. A method of diagnosing a condition or disease associated with the expression of APRG in a subject, comprising administering to said subject an effective amount of the composition of claim 31.

33. A composition of claim 31, wherein the antibody is labeled.

34. A method of diagnosing a condition or disease associated with the expression of APRG in a subject, comprising administering to said subject an effective amount of the composition of claim 33.

35. A method of preparing a polyclonal antibody with the specificity of the antibody of claim 10 comprising:

a) immunizing an animal with a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-3, or an immunogenic fragment thereof, under conditions to elicit an antibody response;

b) isolating antibodies from said animal; and

c) screening the isolated antibodies with the polypeptide, thereby identifying a polyclonal antibody which binds specifically to a polypeptide having an amino acid sequence selected from the

group consisting of SEQ ID NO:1-3.

36. An antibody produced by a method of claim 35.

37. A composition comprising the antibody of claim 36 and a suitable carrier.

38. A method of making a monoclonal antibody with the specificity of the antibody of claim 10 comprising:

a) immunizing an animal with a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-3, or an immunogenic fragment thereof, under conditions to elicit an antibody response;

b) isolating antibody producing cells from the animal;

c) fusing the antibody producing cells with immortalized cells to form monoclonal antibody-producing hybridoma cells;

d) culturing the hybridoma cells; and

e) isolating from the culture monoclonal antibody which binds specifically to a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-3.

39. A monoclonal antibody produced by a method of claim 38.

40. A composition comprising the antibody of claim 39 and a suitable carrier.

41. The antibody of claim 10, wherein the antibody is produced by screening a Fab expression library.

42. The antibody of claim 10, wherein the antibody is produced by screening a recombinant immunoglobulin library.

43. A method for detecting a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-3 in a sample, comprising the steps of:

a) incubating the antibody of claim 10 with a sample under conditions to allow specific binding of the antibody and the polypeptide; and

b) detecting specific binding, wherein specific binding indicates the presence of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-3 in

the sample.

44. A method of purifying a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-3 from a sample, the method comprising:

a) incubating the antibody of claim 10 with a sample under conditions to allow specific binding of the antibody and the polypeptide; and

b) separating the antibody from the sample and obtaining the purified polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-3.

45. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:1.

46. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:2.

47. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:3.

48. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:4.

49. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:5.

50. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:6.

<110> INCYTE GENOMICS, INC.
TANG, Y. Tom
AZIMZAI, Yalda
YUE, Henry
BURFORD, Neil
DING, Li
ELLIOTT, Vicki S.
PATTERSON, Chandra
BAUGHN, Mariah R.

<120> REGULATORS OF APOPTOSIS

<130> PI-0307 PCT

<140> To Be Assigned
<141> Herewith

<150> 60/209,407; 60/250,326
<151> 2000-06-01; 2000-11-30

<160> 6
<170> PERL Program

<210> 1
<211> 326
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 3102521CD1

<400> 1

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

```
Leu Ser Tyr Leu Asp Ala Phe Trp Gly Asp Tyr Leu Ser Gly Ala
                260                 265                 270
Leu Leu Gln Ala Leu Arg Gly Val Phe Leu Thr Glu Ala Leu Arg
                275                 280                 285
Glu Ala Val Gly Arg Glu Ala Val Arg Leu Leu Val Ser Val Asp
                290                 295                 300
Glu Ala Asp Tyr Glu Ala Gly Arg Arg Arg Leu Leu Leu Met Glu
                305                 310                 315
Glu Glu Gly Gly Arg Arg Pro Thr Glu Ala Ser
                320                 325

<210> 2
<211> 655
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7474984CD1

<400> 2
Met Ala Met Ala Lys Ala Arg Lys Pro Arg Glu Ala Leu Leu Trp
  1               5                  10                  15
Ala Leu Ser Asp Leu Glu Glu Asn Asp Phe Lys Lys Leu Lys Leu
                20                  25                  30
Tyr Leu Arg Asp Met Thr Leu Ser Glu Gly Gln Pro Pro Leu Ala
                35                  40                  45
Arg Gly Glu Leu Glu Gly Leu Ile Pro Val Asp Leu Ala Glu Leu
                50                  55                  60
Leu Ile Ser Lys Tyr Gly Glu Lys Glu Ala Val Lys Val Val Leu
                65                  70                  75
Lys Gly Leu Lys Val Met Asn Leu Leu Glu Leu Val Asp Gln Leu
                80                  85                  90
Ser His Ile Cys Leu His Asp Tyr Arg Glu Val Tyr Arg Glu His
                95                 100                 105
Val Arg Cys Leu Glu Glu Trp Gln Glu Ala Gly Val Asn Gly Arg
               110                 115                 120
Tyr Asn Gln Val Leu Leu Val Ala Lys Pro Ser Ser Glu Ser Pro
               125                 130                 135
Glu Ser Leu Ala Cys Pro Phe Pro Glu Gln Glu Leu Glu Ser Val
               140                 145                 150
Thr Val Glu Ala Leu Phe Asp Ser Gly Glu Lys Pro Ser Leu Ala
               155                 160                 165
Pro Ser Leu Val Val Leu Gln Gly Ser Ala Gly Thr Gly Lys Thr
               170                 175                 180
Thr Leu Ala Arg Lys Met Val Leu Asp Trp Ala Thr Gly Thr Leu
               185                 190                 195
Tyr Pro Gly Arg Phe Asp Tyr Val Phe Tyr Val Ser Cys Lys Glu
               200                 205                 210
Val Val Leu Leu Leu Glu Ser Lys Leu Glu Gln Leu Leu Phe Trp
               215                 220                 225
Cys Cys Gly Asp Asn Gln Ala Pro Val Thr Glu Ile Leu Arg Gln
               230                 235                 240
Pro Glu Arg Leu Leu Phe Ile Leu Asp Gly Phe Asp Glu Leu Gln
               245                 250                 255
Arg Pro Phe Glu Glu Lys Leu Lys Lys Arg Gly Leu Ser Pro Lys
               260                 265                 270
Glu Ser Leu Leu His Leu Leu Ile Arg Arg His Thr Leu Pro Thr
               275                 280                 285
Cys Ser Leu Leu Ile Thr Thr Arg Pro Leu Ala Leu Arg Asn Leu
               290                 295                 300
Glu Pro Leu Leu Lys Gln Ala Arg His Val His Ile Leu Gly Phe
               305                 310                 315
Ser Glu Glu Glu Arg Ala Arg Tyr Phe Ser Ser Tyr Phe Thr Asp
               320                 325                 330
Glu Lys Gln Ala Asp Arg Ala Phe Asp Ile Val Gln Lys Asn Asp
               335                 340                 345
```

```
Ile Leu Tyr Lys Ala Cys Gln Val Pro Gly Ile Cys Trp Val Val
            350                 355                 360
Cys Ser Trp Leu Gln Gly Gln Met Glu Arg Gly Lys Val Val Leu
            365                 370                 375
Glu Thr Pro Arg Asn Ser Thr Asp Ile Phe Met Ala Tyr Val Ser
            380                 385                 390
Thr Phe Leu Pro Pro Asp Asp Asp Gly Gly Cys Ser Glu Leu Ser
            395                 400                 405
Arg His Arg Val Leu Arg Ser Leu Cys Ser Leu Ala Ala Glu Gly
            410                 415                 420
Ile Gln His Gln Arg Phe Leu Phe Glu Glu Ala Glu Leu Arg Lys
            425                 430                 435
His Asn Leu Asp Gly Pro Arg Leu Ala Ala Phe Leu Ser Ser Asn
            440                 445                 450
Asp Tyr Gln Leu Gly Leu Ala Ile Lys Lys Phe Tyr Ser Phe Arg
            455                 460                 465
His Ile Ser Phe Gln Asp Phe Phe His Ala Met Ser Tyr Leu Val
            470                 475                 480
Lys Glu Asp Gln Ser Arg Leu Gly Lys Glu Ser Arg Arg Glu Val
            485                 490                 495
Gln Arg Leu Leu Glu Val Lys Glu Gln Glu Gly Asn Asp Glu Met
            500                 505                 510
Thr Leu Thr Met Gln Phe Leu Leu Asp Ile Ser Lys Lys Asp Ser
            515                 520                 525
Phe Ser Asn Leu Glu Leu Lys Phe Cys Phe Arg Ile Ser Pro Cys
            530                 535                 540
Leu Ala Gln Asp Leu Lys His Phe Lys Glu Gln Met Glu Ser Met
            545                 550                 555
Lys His Asn Arg Thr Trp Asp Leu Glu Phe Ser Leu Tyr Glu Ala
            560                 565                 570
Lys Ile Lys Asn Leu Val Lys Gly Ile Gln Met Asn Asn Val Ser
            575                 580                 585
Phe Lys Ile Lys His Ser Asn Glu Lys Lys Ser Gln Ser Gln Asn
            590                 595                 600
Leu Phe Ser Val Lys Ser Ser Leu Ser His Gly Pro Lys Glu Glu
            605                 610                 615
Gln Lys Cys Pro Ser Val His Gly Gln Lys Glu Gly Lys Asp Asn
            620                 625                 630
Ile Ala Gly Thr Gln Lys Glu Ala Ser Thr Gly Lys Gly Arg Gly
            635                 640                 645
Thr Glu Glu Thr Pro Lys Asn Thr Tyr Ile
            650                 655
```

<210> 3
<211> 691
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2458545CD1

```
<400> 3
Met Glu Glu Ser Pro Leu Ser Arg Ala Pro Ser Arg Gly Gly Val
  1               5                  10                  15
Asn Phe Leu Asn Val Ala Arg Thr Tyr Ile Pro Asn Thr Lys Val
            20                  25                  30
Glu Cys His Tyr Thr Leu Pro Pro Gly Thr Met Pro Ser Ala Ser
            35                  40                  45
Asp Trp Ile Gly Ile Phe Lys Val Glu Ala Ala Cys Val Arg Asp
            50                  55                  60
Tyr His Thr Phe Val Trp Ser Ser Val Pro Glu Ser Thr Thr Asp
            65                  70                  75
Gly Ser Pro Ile His Thr Ser Val Gln Phe Gln Ala Ser Tyr Leu
            80                  85                  90
Pro Lys Pro Gly Ala Gln Leu Tyr Gln Phe Arg Tyr Val Asn Arg
            95                 100                 105
```

```
Gln Gly Gln Val Cys Gly Gln Ser Pro Pro Phe Gln Phe Arg Glu
                110                 115                 120
Pro Arg Pro Met Asp Glu Leu Val Thr Leu Glu Glu Ala Asp Gly
            125                 130                 135
Gly Ser Asp Ile Leu Leu Val Val Pro Lys Ala Thr Val Leu Gln
        140                 145                 150
Asn Gln Leu Asp Glu Ser Gln Gln Glu Arg Asn Asp Leu Met Gln
    155                 160                 165
Leu Lys Leu Gln Leu Glu Gly Gln Val Thr Glu Leu Arg Ser Arg
170                 175                 180
Val Gln Glu Leu Glu Arg Ala Leu Ala Thr Ala Arg Gln Glu His
            185                 190                 195
Thr Glu Leu Met Glu Gln Tyr Lys Gly Ile Ser Arg Ser His Gly
        200                 205                 210
Glu Ile Thr Glu Glu Arg Asp Ile Leu Ser Arg Gln Gln Gly Asp
    215                 220                 225
His Val Ala Arg Ile Leu Glu Leu Glu Asp Asp Ile Gln Thr Ile
230                 235                 240
Ser Glu Lys Val Leu Thr Lys Glu Val Glu Leu Asp Arg Leu Arg
            245                 250                 255
Asp Thr Val Lys Ala Leu Thr Arg Glu Gln Glu Lys Leu Leu Gly
        260                 265                 270
Gln Leu Lys Glu Val Gln Ala Asp Lys Glu Gln Ser Glu Ala Glu
    275                 280                 285
Leu Gln Val Ala Gln Gln Glu Asn His His Leu Asn Leu Asp Leu
290                 295                 300
Lys Glu Ala Lys Ser Trp Gln Glu Glu Gln Ser Ala Gln Ala Gln
            305                 310                 315
Arg Leu Lys Asp Lys Val Ala Gln Met Lys Asp Thr Leu Gly Gln
        320                 325                 330
Ala Gln Gln Arg Val Ala Glu Leu Glu Pro Leu Lys Glu Gln Leu
    335                 340                 345
Arg Gly Ala Gln Glu Leu Ala Ala Ser Ser Gln Gln Lys Ala Thr
350                 355                 360
Leu Leu Gly Glu Glu Leu Ala Ser Ala Ala Ala Ala Arg Asp Arg
            365                 370                 375
Thr Ile Ala Glu Leu His Arg Ser Arg Leu Glu Val Ala Glu Val
        380                 385                 390
Asn Gly Arg Leu Ala Glu Leu Gly Leu His Leu Lys Glu Glu Lys
    395                 400                 405
Cys Gln Trp Ser Lys Glu Arg Ala Gly Leu Leu Gln Ser Val Glu
410                 415                 420
Ala Glu Lys Asp Lys Ile Leu Lys Leu Ser Ala Glu Ile Leu Arg
            425                 430                 435
Leu Glu Lys Ala Val Gln Glu Glu Arg Thr Gln Asn Gln Val Phe
        440                 445                 450
Lys Thr Glu Leu Ala Arg Glu Lys Asp Ser Ser Leu Val Gln Leu
    455                 460                 465
Ser Glu Ser Lys Arg Glu Leu Thr Glu Leu Arg Ser Ala Leu Arg
470                 475                 480
Val Leu Gln Lys Glu Lys Glu Gln Leu Gln Glu Glu Lys Gln Glu
            485                 490                 495
Leu Leu Glu Tyr Met Arg Lys Leu Glu Ala Arg Leu Glu Lys Val
        500                 505                 510
Ala Asp Glu Lys Trp Asn Glu Asp Ala Thr Thr Glu Asp Glu Glu
    515                 520                 525
Ala Ala Val Gly Leu Ser Cys Pro Ala Ala Leu Thr Asp Ser Glu
530                 535                 540
Asp Glu Ser Pro Glu Asp Met Arg Leu Pro Pro Tyr Gly Leu Cys
            545                 550                 555
Glu Arg Gly Asp Pro Gly Ser Ser Pro Ala Gly Pro Arg Glu Ala
        560                 565                 570
Ser Pro Leu Val Val Ile Ser Gln Pro Ala Pro Ile Ser Pro His
    575                 580                 585
Leu Ser Gly Pro Ala Glu Asp Ser Ser Ser Asp Ser Glu Ala Glu
590                 595                 600
Asp Glu Lys Ser Val Leu Met Ala Ala Val Gln Ser Gly Gly Glu
```

```
                605                 610                 615
Glu Ala Asn Leu Leu Leu Pro Glu Leu Gly Ser Ala Phe Tyr Asp
                620                 625                 630
Met Ala Ser Gly Phe Thr Val Gly Thr Leu Ser Glu Thr Ser Thr
                635                 640                 645
Gly Gly Pro Ala Thr Pro Thr Trp Lys Glu Cys Pro Ile Cys Lys
                650                 655                 660
Glu Arg Phe Pro Ala Glu Ser Asp Lys Asp Ala Leu Glu Asp His
                665                 670                 675
Met Asp Gly His Phe Phe Phe Ser Thr Gln Asp Pro Phe Thr Phe
                680                 685                 690
Glu
```

<210> 4
<211> 1230
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 3102521CB1

<400> 4
```
cgcaggcgta ataatagaga aggtgccaga aagatccaaa acaagtggct gcggccgtcg 60
cccaggagtc atcggacgcc agaaatctggc cgggttctga gattgttccg actcctccc 120
ccgggaatgg cgctatccgg gtcgaccccg gccccgtgct gggacgagga tgagtgcctg 180
gactactacg ggatgctgtc gattcaccgt atgttcgagg tggtgggcgg gcaactgacc 240
gagtgcgagc tggagctcct ggccttttgg ctggatgagg ctcctggcgc cgccggaggc 300
ttagcccggg cccgcagcgg cctagagctc ctgctggagc tggagccgcg cgggcagtgc 360
gacgagacta acctgcggct gctgggccaa ctcctgcgcg tgctggccgg ccacgacctg 420
ctgccgcacc tggcggcgaa gcggcgcgcgg ccagtgtctc cagaacgcta tagctatggc 480
acctccagct cttcaaagag gacagagggt agctgccgtc gccgtcggca gtccagcagt 540
tctgcaaatt ctcagcaggg tcagtgggag acaggctccc cccaaccaa gcggcagcgg 600
cggagtcggg gccggcccag tggtggtgcc agacggcggc ggagaggggc cccagccgca 660
ccccagcagc agtcagagcc cgccagacct tcctctgaag gcaaagtgac ctgtgacatc 720
cggctccggg ttcgagcaga gtactgcgag catgggccag ccttggagca gggcgtggca 780
tccccgccgc ccccaggcgct ggcgcgtcag ctggacgtgt ttgggcaggc cacccgcagtg 840
ctgcgctcaa gggacctggg ctctgtggtt tgtgacatca agttctcaga gctctcctat 900
atggacgcct tatgggcgga ctacctgagt ggcgccctgc tgcaggccct gcgggggctg 960
ttcctgactg aggccctgcg agaggctgtg ggccgggagg ctgttcgcct gatggtcagt 1020
gtggatgagg ctgactatga ggctggccgg cgccgcctgt tgctgatgga ggaggaaggg 1080
gggcggcgcc cgacagaggc ctcctgatcc aggactggca ggattgatcc cacctccaag 1140
tctccgggcc accttctcct gggaggacga ccatctctac ccctagagga ctgtcaactct 1200
agcatctttg aggactgcga caggaccggg                                   1230
```

<210> 5
<211> 2460
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7474984CB1

<400> 5
```
gcaccaccaa acagaagtga actagtgagt atgggctaag agagcccaaa cttggacctg 60
tagagctgtc ggaccaggaa aggggatctg tttcgtctca gtcccaggc tttgcttact 120
gggctcctgg atcagggagc tgagttctcg ctgcctcacc tccagctctc caagtctgaa 180
ctgtggtcac tggtcttctg ghctggactt gatccttccc ccagatcacc atggccatgg 240
ccaaggccag aaagccccgg gaggcattgc tctgggcctt gagtgacctt gaggagaacg 300
atttcaagaa gttaaagctc tacttacggg atatgacect gtctgagggc cagccccccac 360
tggccagagg ggagttggag ggcctgattc cggtggacct ggcagaatta ctgatttcaa 420
agtatggaga aaaggaggct gtgaaagttg tcttcaaggg cttgaaggtc akgaacctgt 480
tggaacttgt ggaccagctc agccatattt gtctgcatga ttacagagaa gtatacccgg 540
agcatgtgcg ctgcctagag gaatggcagg aagcaggagt caatggcaga tacaaccagg 600
tgctcctggt ggccaagccc agctcagaga gcccagaatc acttgcctgc ccttcccgg 660
```

agcaggagct ggagtctgtc acggtggagg ctctatttga ttcaggggaa aagccctcac 720
tggccccatc cttagttgtg ctacagggct cggctggcac tggaaagaca actttcgcca 780
gaaaaatggt gttggactgg gccacggtta ctctgtaccc aggcaggttt gattatgtcl 840
tttatgtaag ctgcaaagaa gtggtcctgc tgctggagag caaactggag cagctcctlt 900
tctggtgctg cgggtacaat caagccccig tcacagagat tctgaggcag ccaggagcgc 960
tcctgttcat cctggatggc tttgatgagc tgcaggaggc ctttgaagaa aagttgaaga 1020
agaggggttt gagtcccaag gagagcctgc tgcacctict aattaggaga catacactcc 1080
ccacgtgctc ccttctcatc accacccggc cctggctttt gaggaatctg gagccctgc 1140
tgaaacaagc acgtcatgtc catatcctag gcttctctga ggaggagagg gcggagtact 1200
tcagtcccta tttcacggat gagaagcaag ctgaccgtgc cttcgacatt gtacagaaaa 1260
atgacattct ctacaaagcg tgtccaggtt caggcatttg ctgggtgctg tgctcctggc 1320
tgcaggggca gatggagaga ggcaaagttg tcttagagac acctagaaac agcactgaca 1380
tattcatggc ttacgtctcc ascttttgc cgcccgatga tgatggggac tgctccgagc 1440
tttcccggca cagggtcctg aggagtctgt gatccctagc agctgaaggg attcagcacc 1500
agaggttcct atttgaagaa gctgagctca ggaaacataa tttagatggc cccaggcttg 1560
cgctttcct gagtagtaac gactaccaat tgggacttgc catcaagaag ttctcacgct 1620
tccgccacat cagcttccag gacttttttc atgccatgtc ttacctggtg aaagaggacc 1680
aaaycctgct ggggaaggag tcccgcagag aagtgcaaag gctgctggag gtaaagtgag 1740
aggaaggggaa tgatgagatg accctcacta tgcagttttt actggacatc tcgaaaaaag 1800
acagcttctc gaacttggag ctcaagttct gcttcagaac ttstcccigt ttagcgcagg 1860
atctgaagca ttttaaagaa cagatggaat ctatgaagca caacaggacc tgggatttgg 1920
aattccgcct gtatgaaget aaaataaaga atctggtaaa aggtatttag atgaacaatg 1980
tatcattcaa gataaaacat tcaaatgaaa agaaatcaca gagccagaat ttattttctg 2040
tcaaaagcag cttgagtcat ggacctaagg aggagcaaaa atgtcctict gtccatggac 2100
agaaggaggg caaagataat atagcaggaa cacaaaagga agcttctact ggaaaaggca 2160
gagggacaga ggaaacacca aaaaatactt acatataaac ctaarraaat atgtacaaga 2220
tctctatgag gaaaatgaca acactctgag gaaagaaatc aaagaagttg taaatgaatg 2280
ggaagatttt ccatgttcat gaatagaaag actcaatatt gttaagatgt tagttttccc 2340
cagtctgatc tgtagattca atgcaatccc aatcaagatc ccagcaaata acttttgtgga 2400
cattgataag atgatttctaa tgtttatgtg caaaggcaaa agactcagaa tagacaaaaa 2460

<210> 6
<211> 2986
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2458545CB1

<400> 6
ctccagaagt cggagtgctg tttttgttgt tcgtgaaagg tgaggggaac agctgatccg 60
tctgttggga ggacagatat ctcaaggcca ggatggaaga atcaccacta agccgggcac 120
catccgtgg tggagtcaac tttctcaatg tagccggac ctacatcccc aacaccaagg 180
tggaatgtca ctacaccctt ccccaggca ccatgcccag tgccagtgac tggattggca 240
tcttcaaggt ggaggctgcc tgtgttcggg attaccacac atttgtgtgg tcttccgtgc 300
ctgaaagtac aactgatggt tccccattc acaccagtgt ccagttccaa gccagctacc 360
tgcccaaacc aggagctcag ctctaccagt tccgatatgt gaaccgccag ggccagghgt 420
gtgggcagag ccccccttts cagttccgag agccaaggcc catggatgaa ctggtgaccc 480
tggaggaggc tgatggggc tctgacatcc tgctqqttgt cccaaqqca actgtgttac 540
agaaccagct cgatgagagc cagcaagaac ggaatgacct gatgcagctg aagctacagc 600
tggagggaca ggtgacagag ctgaggagcc gagtgcagga gctcgagagg gctctggcaa 660
ctgccaggca ggagcacacg gagctgatgg aacagtacaa ggggatttcc cggtcccatg 720
ggggagtcac agaagagagg gacatcctga gccggcaaca gggagaccat gtggcacgca 780
tcctggagct agaggatgac atccagacca tcagtgagaa agtgctgacg aaggaagtgg 840
agctggacag gcttagagac acagtgaagg ccctgacttg ggaacaagag aagctccttg 900
ggcaacttgaa agaagtacaa gcagacaagg agcaaagtga ggctgagctc caagtggcac 960
aacaggagaa ccatcactta aatttggacc tgaaggaggc gaagagctgg caagaggagc 1020
agagtgctca ggctcagcga ctgaaagaca agytggccca gatgaaggac accctaggcc 1080
aggccagca gcgggtggcc gagctggagc ccttgaagga gcagcttcga ggggcccagg 1140
agttgcagc ctcaagccag cagaaagcca ccctttctgg ggaggagttg gccagtgcag 1200
cagcagccag ggaccgcacc atagccgaac tacaccgcag ccgcctggaa gtggctgaag 1260
ttaacggcag gctggctgag ctcggtttgc acttgaagga agaaaaatgc caatggagca 1320
aggagcgggc aaggctgctg cagagtgtgg aggcagagaa ggacaagatc ctgaagctga 1380
gtgcagagat acttcgattg gagaaggcag ttcaggagga gaggaccaaa aaccaagtgt 1440
tcaagactga gctggccggg gagaaggatt ctagcctgat acagttgtca gaaaglaagc 1500
gggagctgac agagctgcgg tcagcctgc gtgtgctcca gaaggaaaag gagcagttac 1560

```
aggaggagaa acaggaattg ctagagtaca tgagaaagct agaggcccgc ctggagaagg 1620
tggcagatga gaagtggaat gaggatgcca ccacagagga tgaggaggcc gctgtggggc 1680
tgagctgccc ggcagctctg acagactcag aggacgagtc cccagaagac atgaggvctcc 1740
cacccatggg actttgtgag cgtggagacc caggctcctc tcctgctggg cctcgagagg 1800
cttctcccct tgttgtcatc agccagccag ctcccatttc tcctcacctc tctgggccag 1860
ctgaggacag tagctctgac tcggaggctg aagatgagaa gtccagtcctg atggcagctg 1920
tgcagagtgg gggtgaggag gccaacttac tgctttctga actgggcagt gcatttctatg 1980
acatggccag tggctttaca gtgggtaccc tgtcagaaac cagcactggg ggccctgcca 2040
ccccccatg gaaggagtgt cctatctgta aggagcgctt tcctgctgag agtgacaagg 2100
atgccctgga ggaccacatg gatggacact tctttttcag cacccaggac ccattcacct 2160
ttgagtgatc ttacttccte gtacatgcac aaatacacac tcatgcacac acacactcac 2220
acacatgcat acacttaggt ttcatgccca ttttctatca cactgggctc catgatattc 2280
tgttccctaa gaactgcttc tgtgtgccct gttttcatcc cnagatttct cagttcatcc 2340
tctcctaccl ggctcttttg tcccagggag gggtcctgtt cggaacccagt ggctgaattt 2400
atcccctgaa agtggttttg gaggaaccgg gatggaggag gccttcccct gtgggaatag 2460
aatcgtccac tcctagccct ggttgctttct gatacncagc cactgcacac acacactcac 2520
actcacactc ccttgtctga tgcccaaag ccaattcctg gggcacccta ccctctctta 2580
tttggagttt ccgttggttt accigagttt tctctgyggt ctgccacagag gcagcagcat 2640
ggacatcatg gcctctcagg tccctttgg ttctcagttt cattggttcc tcttttctgtt 2700
ccccccattga ctttctgtgcc ccaccctagc cttttccata accttaggta ttcagtttgg 2760
aggggttttt tgtatttttg aggattcctg tattctgtat cctctcctcg catctcctca 2820
catggaaaga aataatgtat ttgtgccttc tgtgaggaat gggggggaaca agtggttccca 2880
ggtatcccca tttccaggtc cccctccgt ctccaggtcc ccccacagca ataaaagctt 2940
ccccctgata tccatccctt tgtagtttga acaaatatat ttatac 2986
```

【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
6 December 2001 (06.12.2001)

PCT

(10) International Publication Number
WO 01/092527 A3

(51) International Patent Classification⁷: C12N 15/12, C07K 14/47, 16/18, G01N 33/68, 33/53, A61K 38/17, C12Q 1/68

(21) International Application Number: PCT/US01/17581

(22) International Filing Date: 30 May 2001 (30.05.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/209,407 1 June 2000 (01.06.2000) US
60/250,326 30 November 2000 (30.11.2000) 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

(75) Inventors/Applicants (*for US only*): TANG, Y., Tom [US/US]; 1230 Renwick Court, San Jose, CA 95118 (US); AZIMZAI, Yalda [US/US]; 5518 Boulder Canyon Drive, Castro Valley, CA 94552 (US); YUE, Henry, Yi [US/US]; 826 Lois Avenue, Sunnyvale, CA 94087 (US); BURFORD, Neil [GB/US]; 105 Windwood Circle, Durham, CT 06422 (US); DING, Li [CN/US]; 3353 Alma Street #146, Palo Alto, CA 94306 (US); ELLIOTT, Vicki, S. [US/US]; 3770 Polton Place Way, San Jose, CA 95121 (US); ARVIZU, Chandra, S. [US/US]; 1706 Morocco Drive, San Jose, CA 95125 (US); BAUGHN, Mariah, R. [US/US]; 14243 Santiago Road, San Leandro, CA 94577 (US).

(74) Agents: HAMLET-COX, Diana 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, BG, BR, BY, BZ, CA, CH, CN, CO, 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.

(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, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published:
— *with international search report*

(88) Date of publication of the international search report:
14 August 2003

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 01/092527 A3

(54) Title: REGULATORS OF APOPTOSIS

(57) Abstract: The invention provides human apoptosis regulators (APRG) and polynucleotides which identify and encode APRG. The invention also provides expression vectors, host cells, antibodies, agonists, and antagonists. The invention also provides methods for diagnosing, treating, or preventing disorders associated with aberrant expression of APRG.

【国際調査報告】

INTERNATIONAL SEARCH REPORT

International Application No
PCT/US 01/17581

A. CLASSIFICATION OF SUBJECT MATTER
IPC 7 C12N15/12 C07K14/47 C07K16/18 G01N33/68 G01N33/53
A61K38/17 C12Q1/68

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 G01N A61K C12Q

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

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

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No
A	PARK J R: "CYTOKINE REGULATION OF APOPTOSIS IN HEMATOPOIETIC PRECURSOR CELLS" CURRENT OPINION IN HEMATOLOGY, RAPID SCIENCE PUBLISHERS, PHILADELPHIA, PA, US, vol. 3, no. 3, May 1996 (1996-05), pages 191-196, XP001095101 ISSN: 1065-6251 the whole document -/--	

[X] Further documents are listed in the continuation of box C. [X] Patent family members are listed in annex.

* Special categories of cited documents :
"A" document defining the general state of the art which is not considered to be of particular relevance
"E" earlier document but published on or after the international filing date
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
"O" document referring to an oral disclosure, use, exhibition or other means
"P" document published prior to the international filing date but later than the priority date claimed
"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.
"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
28 October 2002	21/01/2003

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

Form PCT/ISA/210 (second sheet) (July 1998)

INTERNATIONAL SEARCH REPORT

International Application No
PCT/US 01/17581

C (Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	ERL W ET AL: "NUCLEAR FACTOR-KAPPAB REGULATES INDUCTION OF APOPTOSIS AND INHIBITOR OF APOPTOSIS PROTEIN-1 EXPRESSION IN VASCULAR SMOOTH MUSCLE CELLS" CIRCULATION RESEARCH, GRUNE AND STRATTON, BALTIMORE, US, vol. 6, no. 84, 2 April 1999 (1999-04-02), pages 668-677. XP001079701 ISSN: 0009-7330 the whole document	
P,X	WO 00 56752 A (HUMAN GENOME SCIENCES INC ;NI JIAN (US); RUBEN STEVEN M (US); YOUN) 28 September 2000 (2000-09-28) see gene 1 and seq.ID.'s 2 and 11	1,3,4, 6-27, 29-45
P,X	WO 01 04300 A (HELIX RES INST ;ISOGAI TAKAO (JP); KAWAI YURI (JP); MIYOSHI SOUSUK) 18 January 2001 (2001-01-18) the whole document -& EP 1 201 754 A (HELIX RES INST ;FUJISAWA PHARMACEUTICAL CO (JP)) 2 May 2002 (2002-05-02)	1,3,4, 6-27, 29-45
P,X	WO 01 22920 A (HUMAN GENOME SCIENCES INC ;ROSEN CRAIG A (US); BARASH STEVEN C (US) 5 April 2001 (2001-04-05) see seq.ID's 339 and 4616	1,3,4, 6-27, 29-45

Form PCT/ISA/210 (continuation of second sheet) (July 1998)

INTERNATIONAL SEARCH REPORT

International application No.
PCT/JS 01/17581

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

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

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

see FURTHER INFORMATION sheet PCT/ISA/210

2. [X] Claims Nos.: 20, 23
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:

see FURTHER INFORMATION sheet PCT/ISA/210

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

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

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

see additional sheet

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

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

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

4. [X] 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.:

45 and 48 completely; (1-44) partially

Remark on Protest

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

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

Form PCT/ISA/210 (continuation of first sheet (1)) (July 1998)

International Application No. PCT/US 01 17581

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

Continuation of Box I.1

Although claims 18, 21, and 24 are directed to a method of treatment of the human/animal body, the search has been carried out and based on the alleged effects of the compound/composition.

Although claims 32-34 are directed to a diagnostic method practised on the human/animal body, the search has been carried out and based on the alleged effects of the compound/composition.

Continuation of Box I.2

Claims Nos.: 20, 23

Present claims 20 and 23 relate to products defined by reference to a desirable characteristic or property, namely having (ant)agonistic activity towards the protein(s) of claim 1, and being identified by the methods of claims 19 and 22, respectively.
The application provides no support within the meaning of Article 6 PCT and/or disclosure within the meaning of Article 5 PCT for any such products. In the present case, the claims so lack support, and the application so lacks disclosure, that a meaningful search is impossible. Independent of the above reasoning, the claims also lack clarity (Article 6 PCT). An attempt is made to define the product by reference to a result to be achieved. Again, this lack of clarity in the present case is such as to render a meaningful search impossible. Consequently, the present search report does not extend to said claims.

The applicant's attention is drawn to the fact that claims, or parts of claims, relating to inventions in respect of which no international search report has been established need not be the subject of an international preliminary examination (Rule 66.1(e) PCT). The applicant is advised that the EPO policy when acting as an International Preliminary Examining Authority is normally not to carry out a preliminary examination on matter which has not been searched. This is the case irrespective of whether or not the claims are amended following receipt of the search report or during any Chapter II procedure.

International Application No. PCT/US 01 47581

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

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

1. Claims: 45 and 48 completely, and 1-44 partially

Polypeptide with 90% sequence identity to seq.ID.1 and fragments thereof, nucleic acid encoding it (e.g. seq.ID.4), expression vector, host, transgenic, method for producing, antibody, methods of detecting the nucleic acid or the polypeptide, composition of the polypeptide or the antibody, method for screening for (ant)agonists, modulators, or binding compounds of the polypeptide, or modulators of the expression of the nucleic acid, method for assessing the toxicity, and method of in vitro diagnosis.

2. Claims: 46 and 49 completely, and 1-44 partially

Polypeptide with 90% sequence identity to seq.ID.2 and fragments thereof, nucleic acid encoding it (e.g. seq.ID.5), expression vector, host, transgenic, method for producing, antibody, methods of detecting the nucleic acid or the polypeptide, composition of the polypeptide or the antibody, method for screening for (ant)agonists, modulators, or binding compounds of the polypeptide, or modulators of the expression of the nucleic acid, method for assessing the toxicity, and method of in vitro diagnosis.

3. Claims: 47 and 50 completely, and 1-44 partially

Polypeptide with 90% sequence identity to seq.ID.3 and fragments thereof, nucleic acid encoding it (e.g. seq.ID.6), expression vector, host, transgenic, method for producing, antibody, methods of detecting the nucleic acid or the polypeptide, composition of the polypeptide or the antibody, method for screening for (ant)agonists, modulators, or binding compounds of the polypeptide, or modulators of the expression of the nucleic acid, method for assessing the toxicity, and method of in vitro diagnosis.

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No
PCT/US 01/17581

Patent document cited in search report		Publication date		Patent family member(s)		Publication date
WO 0056752	A	28-09-2000	AU	3528400	A	09-10-2000
			EP	1171449	A1	16-01-2002
			WO	0056752	A2	28-09-2000
WO 0104300	A	18-01-2001	AU	5849300	A	30-01-2001
			AU	5849400	A	30-01-2001
			AU	5849500	A	30-01-2001
			AU	5850500	A	30-01-2001
			EP	1067182	A2	10-01-2001
			EP	1197554	A1	17-04-2002
			EP	1203816	A1	08-05-2002
			EP	1201754	A1	02-05-2002
			WO	0104312	A1	18-01-2001
			WO	0104299	A1	18-01-2001
			WO	0104300	A1	18-01-2001
			WO	0104301	A1	18-01-2001
			JP	2002017376	A	22-01-2002
WO 0122920	A	05-04-2001	AU	7721500	A	30-04-2001
			EP	1265582	A2	18-12-2002
			WO	0122920	A2	05-04-2001

Form PCT/ISA/210 (patent family annex) (July 1992)

フロントページの続き

(51)Int.Cl.[7]	FI		テーマコード（参考）
Ａ６１Ｐ 1/16	Ａ６１Ｐ 1/18		４Ｃ０８４
Ａ６１Ｐ 1/18	Ａ６１Ｐ 3/10		４Ｃ０８５
Ａ６１Ｐ 3/10	Ａ６１Ｐ 5/14		４Ｈ０４５
Ａ６１Ｐ 5/14	Ａ６１Ｐ 7/00		
Ａ６１Ｐ 7/00	Ａ６１Ｐ 7/06		
Ａ６１Ｐ 7/06	Ａ６１Ｐ 7/08		
Ａ６１Ｐ 7/08	Ａ６１Ｐ 9/10		
Ａ６１Ｐ 9/10	Ａ６１Ｐ 9/10	１０１	
Ａ６１Ｐ 11/00	Ａ６１Ｐ 11/00		
Ａ６１Ｐ 11/06	Ａ６１Ｐ 11/06		
Ａ６１Ｐ 13/02	Ａ６１Ｐ 13/02		
Ａ６１Ｐ 13/08	Ａ６１Ｐ 13/08		
Ａ６１Ｐ 13/12	Ａ６１Ｐ 13/12		
Ａ６１Ｐ 15/00	Ａ６１Ｐ 15/00		
Ａ６１Ｐ 15/08	Ａ６１Ｐ 15/08		
Ａ６１Ｐ 15/10	Ａ６１Ｐ 15/10		
Ａ６１Ｐ 17/00	Ａ６１Ｐ 17/00		
Ａ６１Ｐ 17/06	Ａ６１Ｐ 17/06		
Ａ６１Ｐ 19/00	Ａ６１Ｐ 19/00		
Ａ６１Ｐ 19/04	Ａ６１Ｐ 19/04		
Ａ６１Ｐ 19/06	Ａ６１Ｐ 19/06		
Ａ６１Ｐ 21/00	Ａ６１Ｐ 21/00		
Ａ６１Ｐ 25/28	Ａ６１Ｐ 25/28		
Ａ６１Ｐ 29/00	Ａ６１Ｐ 29/00		
Ａ６１Ｐ 31/04	Ａ６１Ｐ 29/00	１０１	
Ａ６１Ｐ 31/10	Ａ６１Ｐ 31/04		
Ａ６１Ｐ 31/12	Ａ６１Ｐ 31/10		
Ａ６１Ｐ 31/18	Ａ６１Ｐ 31/12		
Ａ６１Ｐ 33/00	Ａ６１Ｐ 31/18		
Ａ６１Ｐ 33/02	Ａ６１Ｐ 33/00		
Ａ６１Ｐ 35/00	Ａ６１Ｐ 33/02		
Ａ６１Ｐ 35/02	Ａ６１Ｐ 35/00		
Ａ６１Ｐ 37/02	Ａ６１Ｐ 35/02		
Ａ６１Ｐ 37/08	Ａ６１Ｐ 37/02		
Ａ６１Ｐ 43/00	Ａ６１Ｐ 37/08		
Ｃ０７Ｋ 14/47	Ａ６１Ｐ 43/00	１０１	
Ｃ０７Ｋ 16/18	Ａ６１Ｐ 43/00	１０５	
Ｃ１２Ｎ 1/15	Ａ６１Ｐ 43/00	１１１	
Ｃ１２Ｎ 1/19	Ｃ０７Ｋ 14/47		
Ｃ１２Ｎ 1/21	Ｃ０７Ｋ 16/18		
Ｃ１２Ｎ 5/10	Ｃ１２Ｎ 1/15		
Ｃ１２Ｐ 21/02	Ｃ１２Ｎ 1/19		
Ｃ１２Ｐ 21/08	Ｃ１２Ｎ 1/21		
Ｃ１２Ｑ 1/68	Ｃ１２Ｐ 21/02	Ｃ	
Ｃ１２Ｓ 3/14	Ｃ１２Ｐ 21/08		
Ｇ０１Ｎ 33/15	Ｃ１２Ｑ 1/68	Ａ	
Ｇ０１Ｎ 33/50	Ｃ１２Ｓ 3/14		

Ｇ０１Ｎ	33/53	Ｇ０１Ｎ	33/15 Ｚ
Ｇ０１Ｎ	33/566	Ｇ０１Ｎ	33/50 Ｚ
		Ｇ０１Ｎ	33/53 Ｄ
		Ｇ０１Ｎ	33/53 Ｍ
		Ｇ０１Ｎ	33/566
		Ｃ１２Ｎ	5/00 Ａ
		Ａ６１Ｋ	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,CO,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 アジムザイ、ヤルダ
アメリカ合衆国カリフォルニア州９４５５２・カストロバレー・ボールダーキャニオンドライブ ５５１８
(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９４０８７・サニーベイル・ルイスアベニュー ８２６
(72)発明者 バーフォード、ニール
アメリカ合衆国コネチカット州０６４２２・ダラム・ワイルドウッドサークル １０５
(72)発明者 ディング、リー
アメリカ合衆国カリフォルニア州９４３０６・パロアルト・＃１４６・アルマストリート ３３５３
(72)発明者 エリオット、ビッキ・エス
アメリカ合衆国カリフォルニア州９５１２１・サンノゼ・ポルトンプレイスウェイ ３７７０
(72)発明者 アービズ、チャンドラ・エス
アメリカ合衆国カリフォルニア州９５１２５・サンノゼ・モロッコドライブ １７０６
(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９４５７７・サンレアンドロ・サンティアゴロード １４２４４

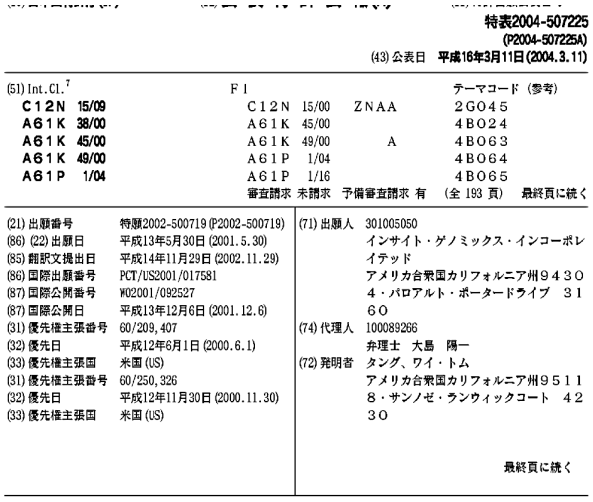
Ｆターム(参考) 2G045 AA40 BA11 BB50 DA13 DA36 FB02 FB03
4B024 AA01 AA11 BA41 BA61 BA80 CA04 CA11 DA01 DA02 DA05 DA11 GA01 GA11 HA08 HA11
4B063 QA01 QA05 QA18 QQ02 QQ08 QQ42 QQ52 QQ79 QQ91 QR08 QR32 QR35 QR42 QR48 QR55 QR62 QR72 QR77 QS24 QS32 QS34 QS36
4B064 AG01 AG26 AG27 CA02 CA05 CA10 CA11 CA19 CA20 CC24 DA01 DA13
4B065 AA01X AA57X AA87X AA90Y AB01 AB02 BA01 BA08 CA24 CA25 CA44
4C084 AA02 AA07 AA17 BA01 BA08 BA22 BA23 CA18 DA53 DC50 NA14 ZA151 ZA361 ZA451 ZA511 ZA551 ZA591 ZA681 ZA751 ZA811 ZA891 ZA941 ZA961 ZA971 ZB071 ZB111 ZB131 ZB151 ZB211 ZB212 ZB261 ZB271 ZB331 ZB351 ZB371 ZB381 ZC061 ZC311 ZC351 ZC412 ZC551
4C085 KA03 KA04 KA05 KA26 KA27 KA29 KB52 KB59 LL18
4H045 AA10 AA11 AA20 AA30 BA09 CA40 DA75 DA76 EA20 EA50 FA72 FA73 FA74

专利名称(译)	细胞凋亡调节因子		
公开(公告)号	JP2004507225A	公开(公告)日	2004-03-11
申请号	JP2002500719	申请日	2001-05-30
[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	タングワイトム アジムザイヤルダ ユエヘンリー バーフォードニール ディングリー エリオットビッキエス アービズチャンドラエス ボーグンマライアアール		
发明人	タング、ワイ·トム アジムザイ、ヤルダ ユエ、ヘンリー バーフォード、ニール ディング、リー エリオット、ビッキ·エス アービズ、チャンドラ·エス ボーグン、マライア·アール		
IPC分类号	G01N33/50 A61K38/00 A61K45/00 A61K49/00 A61P1/04 A61P1/16 A61P1/18 A61P3/10 A61P5/14 A61P7/00 A61P7/06 A61P7/08 A61P9/10 A61P11/00 A61P11/06 A61P13/02 A61P13/08 A61P13/12 A61P15/00 A61P15/08 A61P15/10 A61P17/00 A61P17/06 A61P19/00 A61P19/04 A61P19/06 A61P21/00 A61P25/28 A61P29/00 A61P31/04 A61P31/10 A61P31/12 A61P31/18 A61P33/00 A61P33/02 A61P35/00 A61P35/02 A61P37/02 A61P37/08 A61P43/00 C07K14/47 C07K16/18 C12N1/15 C12N1/19 C12N1/21 C12N5/10 C12N15/09 C12N15/12 C12P21/02 C12P21/08 C12Q1/68 C12S3/14 G01N33/15 G01N33/53 G01N33/566		
CPC分类号	A61K38/00 A61P1/04 A61P1/16 A61P1/18 A61P11/00 A61P11/06 A61P13/02 A61P13/08 A61P13/12 A61P15/00 A61P15/08 A61P15/10 A61P17/00 A61P17/06 A61P19/00 A61P19/04 A61P19/06 A61P21/00 A61P25/28 A61P29/00 C07K14/4747 C07K16/18 C12Q1/6876		
FI分类号	C12N15/00.ZNA.A A61K45/00 A61K49/00.A A61P1/04 A61P1/16 A61P1/18 A61P3/10 A61P5/14 A61P7/00 A61P7/06 A61P7/08 A61P9/10 A61P9/10.101 A61P11/00 A61P11/06 A61P13/02 A61P13/08 A61P13/12 A61P15/00 A61P15/08 A61P15/10 A61P17/00 A61P17/06 A61P19/00 A61P19/04 A61P19/06 A61P21/00 A61P25/28 A61P29/00 A61P29/00.101 A61P31/04 A61P31/10 A61P31/12 A61P31/18 A61P33/00 A61P33/02 A61P35/00 A61P35/02 A61P37/02 A61P37/08 A61P43/00.101 A61P43/00.105 A61P43/00.111 C07K14/47 C07K16/18 C12N1/15 C12N1/19 C12N1/21 C12P21/02.C C12P21/08 C12Q1/68.A C12S3/14 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/BA11 2G045/BB50 2G045/DA13 2G045/DA36 2G045/FB02 2G045/FB03 4B024/AA01 4B024/AA11 4B024/BA41 4B024/BA61 4B024/BA80 4B024/CA04 4B024/CA11 4B024/DA01 4B024/DA02 4B024/DA05 4B024/DA11 4B024/GA01 4B024/GA11 4B024/HA08 4B024/HA11 4B063/QA01 4B063/QA05 4B063/QA18 4B063/QQ02 4B063/QQ08 4B063/QQ42 4B063/QQ52 4B063/QQ79 4B063/QQ91 4B063/QR08 4B063/QR32 4B063/QR35 4B063/QR42 4B063/QR48 4B063/QR55 4B063/QR62 4B063/QR72 4B063/QR77 4B063/QS24 4B063/QS32 4B063/QS34 4B063/QS36 4B064/AG01 4B064/AG26 4B064/AG27 4B064/CA02 4B064/CA05 4B064/CA10 4B064/CA11 4B064/CA19 4B064		

	/CA20 4B064/CC24 4B064/DA01 4B064/DA13 4B065/AA01X 4B065/AA57X 4B065/AA87X 4B065/AA90Y 4B065/AB01 4B065/AB02 4B065/BA01 4B065/BA08 4B065/CA24 4B065/CA25 4B065/CA44 4C084/AA02 4C084/AA07 4C084/AA17 4C084/BA01 4C084/BA08 4C084/BA22 4C084/BA23 4C084/CA18 4C084/DA53 4C084/DC50 4C084/NA14 4C084/ZA151 4C084/ZA361 4C084/ZA451 4C084/ZA511 4C084/ZA551 4C084/ZA591 4C084/ZA681 4C084/ZA751 4C084/ZA811 4C084/ZA891 4C084/ZA941 4C084/ZA961 4C084/ZA971 4C084/ZB071 4C084/ZB111 4C084/ZB131 4C084/ZB151 4C084/ZB211 4C084/ZB212 4C084/ZB261 4C084/ZB271 4C084/ZB331 4C084/ZB351 4C084/ZB371 4C084/ZB381 4C084/ZC061 4C084/ZC311 4C084/ZC351 4C084/ZC412 4C084/ZC551 4C085/KA03 4C085/KA04 4C085/KA05 4C085/KA26 4C085/KA27 4C085/KA29 4C085/KB52 4C085/KB59 4C085/LL18 4H045/AA10 4H045/AA11 4H045/AA20 4H045/AA30 4H045/BA09 4H045/CA40 4H045/DA75 4H045/DA76 4H045/EA20 4H045/EA50 4H045/FA72 4H045/FA73 4H045/FA74
优先权	60/209407 2000-06-01 US 60/250326 2000-11-30 US
外部链接	Espacenet

摘要(译)

本发明提供了鉴定和编码APRG的人凋亡调节剂（APRG）和多核苷酸。 本发明还提供表达载体，宿主细胞，抗体，激动剂和拮抗剂。 本发明还提供了诊断，治疗或预防与APRG异常表达有关的疾病的方法。



特表2004-507225
(P2004-507225A)
(43) 公表日 平成16年3月11日 (2004.3.11)

(51) Int.Cl.[7]	F I		テーマコード（参考）
C12N 15/09	C12N 15/00	ZNAA	2G045
A61K 38/00	A61K 45/00		4B024
A61K 45/00	A61K 49/00	A	4B063
A61K 49/00	A61P 1/04		4B064
A61P 1/04	A61P 1/16		4B065

審査請求 未請求 予備審査請求 有 （全 193 頁） 最終頁に続く

(21) 出願番号	特願2002-500719 (P2002-500719)
(86) (22) 出願日	平成13年5月30日 (2001.5.30)
(85) 翻訳文提出日	平成14年11月29日 (2002.11.29)
(86) 国際出願番号	PCT/US2001/017581
(87) 国際公開番号	WO2001/092527
(87) 国際公開日	平成13年12月6日 (2001.12.6)
(31) 優先権主張番号	60/209,407
(32) 優先日	平成12年6月1日 (2000.6.1)
(33) 優先権主張国	米国 (US)
(31) 優先権主張番号	60/250,326
(32) 優先日	平成12年11月30日 (2000.11.30)
(33) 優先権主張国	米国 (US)

(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
(74) 代理人 100089266
弁理士 大島 陽一
(72) 発明者 タング、ワイ・トム
アメリカ合衆国カリフォルニア州９５１１８・サンノゼ・ランウィックコート ４２３０

最終頁に続く