

(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2004-527209

(P2004-527209A)

(43) 公表日 平成16年9月9日(2004.9.9)

(51) Int. Cl. ⁷	F I	テーマコード (参考)
C 1 2 N 15/09	C 1 2 N 15/00	Z N A A
A 6 1 K 38/00	A 6 1 K 45/00	2 G O 4 5
A 6 1 K 45/00	A 6 1 P 3/06	4 B O 2 4
A 6 1 P 3/06	A 6 1 P 5/06	4 B O 5 0
A 6 1 P 5/06	A 6 1 P 9/00	4 B O 6 3
		4 B O 6 4
	審査請求 未請求 予備審査請求 有	(全 318 頁) 最終頁に続く

(21) 出願番号	特願2002-513883 (P2002-513883)	(71) 出願人	304029608
(86) (22) 出願日	平成13年7月20日 (2001.7.20)		ソーントン、マイケル
(85) 翻訳文提出日	平成15年1月20日 (2003.1.20)		アメリカ合衆国カリフォルニア州9406
(86) 国際出願番号	PCT/US2001/023092		2-2612・ウッドサイド・メッドウェ
(87) 国際公開番号	W02002/008399		イロード 9
(87) 国際公開日	平成14年1月31日 (2002.1.31)	(74) 代理人	100089266
(31) 優先権主張番号	60/220,038		弁理士 大島 陽一
(32) 優先日	平成12年7月21日 (2000.7.21)	(72) 発明者	ソーントン、マイケル
(33) 優先権主張国	米国 (US)		アメリカ合衆国カリフォルニア州9406
(31) 優先権主張番号	60/222,112		2-2612・ウッドサイド・メッドウェ
(32) 優先日	平成12年7月28日 (2000.7.28)		イロード 9
(33) 優先権主張国	米国 (US)	(72) 発明者	ユエ、ヘンリー
(31) 優先権主張番号	60/222,831		アメリカ合衆国カリフォルニア州9408
(32) 優先日	平成12年8月4日 (2000.8.4)		7・サニーバイル・ルイスアベニュー 8
(33) 優先権主張国	米国 (US)		26
			最終頁に続く

(54) 【発明の名称】 ヒトキナーゼ

(57) 【要約】

本発明はヒトキナーゼ (PKIN) およびPKIN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、PKIN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】

下の (a) 乃至 (d) を有する群から選択した単離されたポリヌクレオチド。

(a) SEQ ID NO: 1 - 20 を有する群から選択したポリペプチド

(b) SEQ ID NO: 1 - 20 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含む天然のポリペプチド

(c) SEQ ID NO: 1 - 20 を有す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 - 20 を有す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

10

【請求項 2】

SEQ ID NO: 1 - 20 からなる群から選択された請求項 1 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】

SEQ ID NO: 21 - 40 からなる群から選択された請求項 4 の単離されたポリヌクレオチド。

20

【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、以下の過程を含む方法。

(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、

30

(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的に結合する単離された抗体。

【請求項 11】

以下の (a) 乃至 (d) を有する群から選択した単離されたポリヌクレオチド。

(a) SEQ ID NO: 21 - 40 を有す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: 21 - 40 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を含むポリヌクレオチド

40

(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、

50

(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ことを特徴とする方法。

10

【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。

20

【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有することを特徴とする成分。

【請求項 17】

前記ポリペプチドが、SEQ ID NO: 1 - 20 を有す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 PKIN 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。

30

【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曝露する過程と、

(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成分。

【請求項 21】

機能的な PKIN 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。

40

【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曝露する過程と、

(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

50

る賦形剤とを含むことを特徴とする成分。

【請求項 24】

機能的な P K I N 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ステップと、

(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。 10

【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。 20

【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。 30

【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、 40

(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 P K I N 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PKINの発現に関連する病状又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PKINの発現に関連する病状又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: 1-20を有する群から選択したアミノ酸配列またはその免疫抗原性断片を含むポリペプチド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: 1-20を有す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: 1-20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: 1-20を有す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】

10

20

30

40

50

F a b 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】

S E Q I D N O : 1 - 20 を有す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、S E Q I D N O : 1 - 20 を有す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。

10

【請求項 44】

S E Q I D N O : 1 - 20 を有す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、S E Q I D N O : 1 - 20 を有す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

20

【請求項 45】

S E Q I D N O : 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】

S E Q I D N O : 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】

S E Q I D N O : 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】

S E Q I D N O : 4 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 49】

S E Q I D N O : 5 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。

30

【請求項 50】

S E Q I D N O : 6 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 51】

S E Q I D N O : 7 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 52】

S E Q I D N O : 8 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 53】

S E Q I D N O : 9 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 54】

S E Q I D N O : 10 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。

40

【請求項 55】

S E Q I D N O : 11 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 56】

S E Q I D N O : 12 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 57】

S E Q I D N O : 13 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 58】

S E Q I D N O : 14 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 59】

S E Q I D N O : 15 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。

50

SEQ ID NO : 34 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。

【請求項 79】

SEQ ID NO : 35 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。

【請求項 80】

SEQ ID NO : 36 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。

【請求項 81】

SEQ ID NO : 37 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。 10

【請求項 82】

SEQ ID NO : 38 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。

【請求項 83】

SEQ ID NO : 39 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。

【請求項 84】

SEQ ID NO : 40 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。 20

【発明の詳細な説明】

【0001】

(技術分野)

本発明は、ヒトキナーゼ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癌、免疫疾患、成長および発達に影響を及ぼす疾患、心血管疾患、および脂質異常の診断・治療・予防に関する。本発明はさらに、ヒトキナーゼ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)

キナーゼは最も大きな既知の酵素スーパーファミリーを構成し、標的分子が多種多様である。キナーゼは、高エネルギーリン酸基のリン酸供与体からリン酸受容体への移動を触媒する。ヌクレオチドは通常、これらの反応においてリン酸供与体として働き、ほとんどのキナーゼはアデノシン三リン酸 (ATP) を利用する。リン酸受容体には様々な分子が可能であり、その中には、ヌクレオシド、ヌクレオチド、脂質、炭水化物、およびタンパク質が含まれる。タンパク質は、ヒドロキシアミノ酸においてリン酸化される。リン酸基が付加されると、受容体分子における局所電荷が変化し、それによって内部構造が変化し、分子間接触を引き起こし得る。可逆的なタンパク質リン酸化は、真核細胞におけるタンパク質活性の調節の主な方法である。一般に、タンパク質は、ホルモン、神経伝達物質、成長因子、および分化因子などの細胞外シグナルに応答したリン酸化により活性化される。活性化されたタンパク質は細胞の細胞応答を、細胞内シグナル伝達経路、並びにタンパク質リン酸化を調節するサイクリックヌクレオチド、カルシウム - カルモジュリン、イノシトール、および様々な分裂促進因子などのセカンドメッセンジャー分子によって開始させる。 30 40

【0003】

キナーゼは、解糖などの基本的な代謝プロセスから細胞周期の調節、分化、およびシグナル伝達カスケードによる細胞外環境との情報交換に至る細胞機能の全てに参与する。細胞内におけるタンパク質の不適切なリン酸化は、細胞周期の進行および細胞分化における変化に関連する。細胞周期における変化は、アポトーシスや癌の誘導に関連する。細胞分化における変化は、生殖系、免疫系、および骨格筋の疾患や障害に関連する。

プロテインキナーゼには 2 つのクラスがある。一方のクラスのプロテインチロシンキナー 50

ゼ (P T K) はチロシン残基をリン酸化し、他方のクラスであるプロテインセリン/トレオニンキナーゼ (S T K) はセリン残基およびトレオニン残基をリン酸化する。ある種の P T K および S T K は、両方のファミリーの構造的な特徴を有し、チロシン残基およびセリン/トレオニン残基の両方に対して特異性 (二重特異性) を有する。ほとんど全てのキナーゼは、キナーゼファミリー特有の特定の残基および配列モチーフを含む保存された 250 ~ 300 のアミノ酸からなる触媒ドメインを含む。プロテインキナーゼ触媒ドメインは、更に 11 のサブドメインに分類することができる。N 末端サブドメイン I - I V は、A T P 供与体分子に結合して方向性を与える 2 葉構造 (*two lobed structure*) に折り畳まれ、サブドメイン V は 2 葉にまたがる。C 末端サブドメイン V I - X I はタンパク質基質と結合し、リン酸を A T P からチロシン残基、セリン残基、またはトレオニン残基のヒドロキシル基に移動させる。11 のサブドメインのそれぞれは、サブドメインに特徴的なアミノ酸モチーフ或いは特定の触媒残基を含む。例えば、サブドメイン I は、8 個のアミノ酸からなる高グリシン A T P 結合共通モチーフを含み、サブドメイン I I は、触媒活性の最大化に必要な重要なリシン残基を含み、サブドメイン V I から I X は高度に保存された触媒中心を含む。P T K および S T K はまた、ヒドロキシアミノ酸特異性を付与し得るサブドメイン V I および V I I I に固有の配列モチーフを含む。

10

【 0 0 0 4 】

更に、キナーゼは、キナーゼドメイン内に含まれるか或いは隣接する、通常は 5 ~ 100 の範囲の残基からなる追加のアミノ酸配列によって分類され得る。これらの追加のアミノ酸配列はキナーゼ活性を調節し、基質特異性を決定する。(*Hardie, G. 及び S. Hanks (1995) The Protein Kinase Facts Book, Vol I, 17-20 ページ Academic Press, San Diego CA.* を参照)。具体的には、2 つのプロテインキナーゼシグネチャ配列が、キナーゼドメインで同定された。第 1 のプロテインキナーゼシグネチャ配列は、A T P 結合に関与するリシン残基活性部位を含み、第 2 のプロテインキナーゼシグネチャ配列は触媒活性に重要なアスパラギン酸残基を含む。分析するタンパク質がこの 2 つのプロテインキナーゼシグネチャを含む場合、そのタンパク質がプロテインキナーゼである確率は 100 % に近い (*PROSITE: PDOC00100, November 1995*)。

20

【 0 0 0 5 】

プロテインチロシンキナーゼ

プロテインチロシンキナーゼ (P T K) は、膜貫通受容体 P T K タンパク質または非膜貫通非受容体 P T K タンパク質のいずれかに分類され得る。膜貫通チロシンキナーゼは、ほとんどの成長因子に対する受容体として機能する。成長因子が受容体チロシンキナーゼ (R T K) に結合し、それによって受容体が、自己 (自己リン酸化) および特定の細胞内セカンドメッセンジャータンパク質をリン酸化する。受容体 P T K に結合する成長因子 (G F) は、上皮成長因子、血小板由来成長因子、繊維芽細胞成長因子、肝細胞成長因子、インスリンおよびインスリン類似成長因子、神経成長因子、血管内皮成長因子、およびマクローファージコロニー刺激因子が含まれる。

30

【 0 0 0 6 】

非膜貫通型非受容体 P T K は、膜貫通領域を含まない代わりに細胞膜受容体の細胞質ドメインとシグナル伝達複合体を形成する。非受容体 P T K を介して機能する受容体には、サイトカイン受容体およびホルモン受容体 (成長ホルモンおよびプロラクチン) や、T リンパ球および B リンパ球における抗原特異的受容体が含まれる。

40

【 0 0 0 7 】

多くの P T K は、P T K 活性が正常な細胞制御を受けない癌細胞における腫瘍遺伝子産物として初めに同定された。実際に、約 3 分の 1 の腫瘍遺伝子が P T K をコードする。細胞形質転換 (発癌) はチロシンリン酸化活性の上昇を伴う場合が多い (*Charbonneau, H. および Tonks, N. K. (1992) Annu. Rev. Cell Biol. 8: 463-93*)。したがって、P T K 活性の調節は、ある種の癌を制御するための重要な手段となり得る。

50

【0008】

チロシンキナーゼ基質は、チロシン-リン酸化cDNA発現ライブラリをスクリーニングするための抗ホスホチロシン抗体で同定できる。Srcで形質転換された線維芽細胞でチロシン-リン酸化されているFishは、そのような手法で同定されたチロシンキナーゼ基質である。Fishは5つのSH3ドメインおよびひとつのphox相同性(PX)ドメインをもっている。Fishがチロシンキナーゼのシグナルに關与し、アクチン細胞骨格に機能すると考えられる(Lock, P. 他(1998)EMBO J. 17: 4346-4357)。

【0009】

ホスホチロシン・ホスファターゼ含有のSH2ドメインであるSHP-2はいくつかの受容体チロシンキナーゼ(RTK)とサイトカイン受容体のポジティブなシグナル・トランスデューサである。ホスホチロシン・ホスファターゼは、細胞内シグナル経路重要なポジティブとネガティブの調節因子である。これは、細胞の分裂、遊走、分化、形質転換、存続または細胞死など特異性のある成長因子対応細胞に影響する。SIRP-サブファミリー・メンバーはSHP-2の細胞質内結合ドメインを有する。そのサブファミリーにはSIRP-1が含まれ、SIRP-1は膜貫通タンパク質で、活性化されたRTKの基質であり、RTKはSH2ドメインに結合する。SIRPは免疫抗原体認識分子に高いレベルの相同性をもっている。SIRP-サブファミリーは細胞質尾部をもっていない。SIRP-1遺伝子は398のアミノ酸のポリペプチドをコードする。一般的にSIRPファミリーは様々な生理学的・病理学的な過程を定めるシグナル調節に關与する(Kharitonov, A. 他(1997)Nature 386: 181-186)。調節可能な2つの領域には、脳の多様性および遺伝的個性(Sano, S. 他(1999)Biochem. J. 344 Pt 3: 667-675)、そして溶血性貧血のような疾患で欠陥がある自己認識(Oldenborg, P. - A. 他(2000)Science 288: 2051-2054)がある。

【0010】

プロテインセリン/トレオニンキナーゼ

プロテインセリン/トレオニンキナーゼ(STK)は非膜貫通型タンパク質である。STKのサブクラスは、ERK(細胞外シグナル調節キナーゼ)またはMAP(マイトジェン活性化プロテインキナーゼ)として知られ、様々なホルモンや成長因子下の細胞刺激によって活性化される。細胞刺激は、MEK(MAP/ERKキナーゼ)がリン酸化に導くシグナル伝達カスケードを誘導し、MEKのリン酸化の後に、セリンおよびトレオニンのリン酸化によりERKが活性化される。多数のタンパク質が活性化されたERKに対する下流のエフェクターであることから、細胞増殖および分化の調節や細胞骨格の調節に關与すると思われる。ERKの活性化は通常一過性である。細胞は、そのダウンレギュレーション(下方制御)に關係する二重特異性ホスファターゼを有する。また様々な研究から、上昇されたERK活性がある種の癌に關連することが分かった。その他のSTKには、サイクリックAMP依存性プロテインキナーゼ(PKA)、カルシウム-カルモジュリン(CaM)依存性プロテインキナーゼ、およびマイトジェン活性化プロテインキナーゼ(MAP)などのセカンドメッセンジャー依存性プロテインキナーゼや、サイクリン依存性プロテインキナーゼ、チェックポイントキナーゼおよび細胞周期キナーゼ、Numb關連キナーゼ(Nak)、ヒト融合型(hFu)、増殖關連キナーゼ、5'-AMP-活性化プロテインキナーゼ、およびアポトーシスに關与するキナーゼが含まれる。

【0011】

セカンドメッセンジャー依存性プロテインキナーゼは、サイクリックAMP(cAMP)、サイクリックGMP、イノシトール三リン酸、ホスファチジルイノシトール、3, 4, 5-三リン酸、サイクリックADPリボース、アラキドン酸、ジアシルグリセロールおよびカルシウム-カルモジュリンなどのセカンドメッセンジャーの効果を主に仲介する。PKAはホルモン誘導性細胞応答の仲介に關与し、ホルモン刺激に應答して細胞内で生成されるcAMPによって活性化される。cAMPは、研究した全ての動物細胞におけるホル

モン作用の細胞内メディエーターである。ホルモン誘導性細胞応答には、甲状腺ホルモン分泌、コルチゾル分泌、プロゲステロン分泌、グリコーゲン分解、骨再吸収、心拍数の調節、および心筋収縮力の調節が含まれる。PKAは全ての動物細胞に見られ、これらほとんどの細胞におけるcAMPの効果に関係すると思われる。PKA発現の変化は、癌、甲状腺疾患、糖尿病、アテローム性動脈硬化症、および心血管疾患に関連する (Isselbacher, K. J. 他、(1994) Harrison's Principles of Internal Medicine, McGraw-Hill, New York, NY, 416-431ページ, 1887)。

【0012】

カゼインキナーゼI (CKI) 遺伝子ファミリーは、セリン/トレオニンプロテインキナーゼの別のサブファミリーである。この引き続き拡大するキナーゼ群は、細胞代謝、DNAの複製、およびDNAの修復を含む様々な細胞プロセスおよび核プロセスの調節に関与する。CKI酵素は、真核細胞の膜、核、細胞質および細胞骨格や、哺乳動物細胞の紡錘体状有糸分裂に存在する (Fish, K. J. 他、(1995) J. Biol. Chem. 270:14875-14883)。CKIファミリーメンバーの全ては、9~76のアミノ酸からなる短いアミノ末端ドメイン、284のアミノ酸からなる高度に保存されたキナーゼドメイン、および24~200を超える範囲のアミノ酸からなる可変カルボキシル末端ドメインを含む (Cegiel ska, A. 他、(1998) J. Biol. Chem. 273:1357-1364)。このCKIファミリーは高度に関連するタンパク質である。これは、様々な試料からのカゼインキナーゼIのイソ型の同定によって確認されている。 、 、 、 の少なくとも5つのイソ型がある。ヒト胎盤cDNAライブラリからのCKI- はFishらによって同定された。このCKI- は416のアミノ酸からなる塩基性タンパク質であり、CKI- に最も類似している。組換え発現によって、CKI- が既知のCKI基質をリン酸化し、CKI特異的インヒビターであるCKI-7によって抑制されることが分かった。CKI- のヒト遺伝子は、酵母CKI遺伝子座であるHRR250の欠失によって起こる成長が遅い表現型 (slow-growth phenotype) を有する酵母を助けることができる (Fishら, 前出)。

【0013】

哺乳動物突然変異概日tauが、シリアンハムスターにおける概日リズムの周期を著しく短くするCKI- の半優性染色体対立遺伝子であることが分かった。tau遺伝子座はカゼインキナーゼI- によってコードされ、ショウジョウバエ概日遺伝子double-timeの相同体でもある。野生型およびtau突然変異CKI- 酵素の両方の研究から、突然変異酵素はその最大速度が著しく低下し、自己リン酸化状態であることが分かった。更に、in vitroにおけるCKI- は哺乳動物PERIODタンパク質と相互作用する能力を有するが、突然変異酵素はPERIODをリン酸化する能力に欠ける。Lowreyらが、CKI- が概日機構の中心をなす転写-翻訳系の自己調節ループ内の負のフィードバックシグナルを遅らせる重要な役割を果たしていると提案した。従って、CKI- 酵素は、概日リズム、時差ぼけおよび睡眠、更に概日調節下のその他の生理的プロセスおよび代謝プロセスに影響を与える医薬組成物の理想的な候補である (Lowrey, P. L. 他、(2000) Science 288:483-491)。ホメオドメイン相互作用プロテインキナーゼ (HIPK) はセリン/トレオニンキナーゼであり、DYRKキナーゼサブファミリーの新しいメンバーである (Hofmann, T. G. 他、(2000) Biochimie 82:1123-1127)。

【0014】

HIPKはホメオプロテインと相互作用するドメインから離れた、保存されたプロテインキナーゼドメインを含む。HIPKは核内キナーゼであり、またHIPK2は神経組織において高レベルで発現される (Kim, Y. H. 他、(1998) J. Biol. Chem. 273:25875-25879; Wang, Y. 他、(2001) Biochim. Biophys. Acta 1518:168-172)。HIPKは

10

20

30

40

50

ホメオドメイン転写因子の補抑制体として作用する。この補抑制体の活性は、細胞タンパク質機能の調節に重要なユビキチン結合やリン酸化のような翻訳後修飾において見られる (Kim, Y. H. 他、(1999) Proc. Nat. Acad. Sci. U.S.A. 96:12350-12355)。

【0015】

Caenorhabditis elegans UNC51の Maus 相同体であるセリン/トレオニンキナーゼは、小脳皮質の顆粒細胞から軸索形成を導く遺伝子発現プログラムを制御するのに不可欠であることが判明した (Tomoda, T. 他 (1999) Neuron 24:833-346)。UNC-51のヒトの相同体であるULK1 (「UNC-51 (C. elegans) - like kinase 1」の略) は、1050個のアミノ酸より構成され、計算された分子量は112.6 kDaで、8.80の等電点である。ULK1はUNC-51と41%の全体的な配列類似性があり、哺乳類、鳥類、爬虫類、両生類、および魚類を含む脊椎動物の中で高度に保存されている。ノーザンブロット分析により、Kuroyanagi 他は、骨格筋、心臓、脾臓、脳、胎盤、肝臓、腎臓、および肺を含めた成人組織でULK1が広範に発現されていることをした。一方、UNC-51は線虫の神経系に特異的に局在していた。FISHマッピングとRHマッピングでULK1がヒト染色体12q24.3に位置することが確認された (Kuroyanagi, H. 他 (1998) Genomics 51:76-85)。

【0016】

カルシウム - カルモジュリン依存性プロテインキナーゼ

カルシウム - カルモジュリン (CaM) 依存性キナーゼは、平滑筋の収縮、グリコーゲン分解 (ホスホリラーゼキナーゼ)、および神経伝達 (CaMキナーゼIおよびCaMキナーゼII) の調節に参与する。CaM依存性プロテインキナーゼは、細胞内の遊離カルシウムの濃度に応答して細胞内カルシウム受容体であるカルモジュリンによって活性化される。多くのCaMキナーゼはまた、リン酸化によって活性化される。ある種のCaMキナーゼはまた、自己リン酸化またはその他の調節キナーゼによって活性化される。CaMキナーゼIは、神経伝達物質関連タンパク質であるシナプシンIおよびII、遺伝子転写調節因子であるCREB、および囊胞性繊維コンダクタンズ調節タンパク質であるCFTRを含む様々な物質をリン酸化する (Haribabu, B. 他、(1995) EMBO Journal 14:3679-3686)。また、CaMキナーゼIIは、様々な部位においてシナプシンをリン酸化し、チロシンヒドロキシラーゼのリン酸化および活性化によって脳におけるカテコールアミンの合成を調節する。CaMキナーゼIIはまた、チロシンヒドロキシラーゼおよびトリプトファンヒドロキシラーゼのリン酸化/活性化によってカテコールアミンおよびセロトニンの合成を調節する (Fujisawa, H. (1990) BioEssays 12:27-29)。カルモジュリン結合プロテインキナーゼ類似タンパク質をコードするmRNAが哺乳動物前頭葉で多量に見つかった。このタンパク質は軸索および樹状突起の双方における小胞に結合し、主に出生後に蓄積される。このタンパク質のアミノ酸配列はCaM依存性STKに類似しており、このタンパク質はカルシウムの存在下でカルモジュリンと結合する (Godbout, M. 他、(1994) J. Neurosci. 14:1-13)。

【0017】

マイトジェン活性化プロテインキナーゼ

マイトジェン活性化プロテインキナーゼ (MAP) は、リン酸化カスケードによって細胞表面から核へのシグナル伝達を仲介する。MAPは、細胞内シグナル伝達経路を調節する別のSTKファミリーである。いくつかのサブグループが同定されており、それぞれが異なった基質特異性を有し、固有の細胞外刺激に応答する (Egan, S. E. および Weinberg, R. A. (1993) Nature 365:781-783)。MAPキナーゼ・カスケードには3つのキナーゼモジュールが含有する。すなわち、MAPK (MAP)、MAPKキナーゼ (MAP2K、MAPKKまたはMKK)、およびMKKキナーゼ (MAP3K、MAPKKKまたはMEKK) である (W 50

ang, X. S. 他、(1998) *Biochem. Biophys. Res. Commun.* 253:33-37)。細胞外調節キナーゼ(ERK)経路は、上皮成長因子(EGF)、紫外線、高浸透圧媒体、熱ショック、内毒素性リポ多糖(LPS)など成長因子および分裂促進因子で活性化される。緊密に関係し区別可能な経路、c-Jun N-末端キナーゼ(JNK)、またはストレス応答キナーゼ(SAPK)経路およびp38キナーゼ経路は、腫瘍壊因子(TNF)またはインターロイキン-1(IL-1)などストレス刺激性と前炎症性のサイトカインで活性される。MAPキナーゼ発現の変化は、癌、炎症、免疫疾患、および成長および発達に影響を及ぼす疾患を含む様々な疾患に關与する。MAPキナーゼシグナリング経路は哺乳類細胞および酵母に存在する。

【0018】

MAPKKK6(MAP3K6)は同定された多数のMAP3Kの一つである。骨格筋で分離されたMAP3K6は、1,280個のアミノ酸から成り、11のキナーゼ・サブドメインをもち、多種の組織に検出される。最も高い発現は心臓および骨格筋から確認された。MAP3K6はMAP3K5と45%のアミノ酸配列が同一で、またその触媒ドメインは82%のアミノ酸配列が同一である。

*in vivo*でのMAP3K6とMAP3K5との相互作用は共免疫沈降法で確認された。組換えMAP3K6はJNKを多少活性化するが、p38キナーゼまたはERK経路を活性化しないことが示された(Wang, X. S. ら、前出)。

【0019】

サイクリン依存性プロテインキナーゼ

サイクリン依存性プロテインキナーゼ(CDK)は、細胞周期を介して細胞の進行を調節するSTKである。細胞は、サイクリンと呼ばれる活性化タンパク質ファミリーの合成および分解による調節により、有糸分裂に入ったり有糸分裂から出たりする。小さな調節タンパク質であるサイクリンはCDKに結合してそのCDKを活性化させる。それによって、有糸分裂プロセスに關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は、活性化するために多数の入力が必要であるという点でユニークである。サイクリンの結合に加えて、CDKの活性化には、CDKにおける特定のトレオニン残基のリン酸化および特定のチロシン残基の脱リン酸化が必要である。

【0020】

NIMA(有糸分裂に決して入らない)関連キナーゼ(Nek)は細胞周期に關連するSTKの別のファミリーである。CDKおよびNekの両方は、動物細胞における複製、成熟、微小管形成中心である中心体の分離に關与する(Fry, A. M. 他、(1998) *EMBO J.* 17:470-481)。

【0021】

チェックポイントおよび細胞周期キナーゼ

細胞分裂のプロセスでは、細胞周期移行の順番およびタイミングは細胞周期チェックポイントの制御下にある。この細胞周期チェックポイントは、DNA複製および染色体分離などの重要なプロセスが正確に実行されるようにする。例えば放射線などによってDNAが損傷を受けた場合、チェックポイント経路が活性化され、修復の時間を与えるべく細胞周期が停止される。損傷が大きい場合には、アポトーシスが誘導される。このようなチェックポイントが存在しないと、損傷したDNAが異常な細胞に受け継がれ、癌などの増殖異常が引き起こされ得る。プロテインキナーゼは、このプロセスにおいて重要な役割を果たす。例えば、特定のキナーゼすなわちチェックポイントキナーゼ1(Chk1)が酵母および哺乳動物で同定された。このChk1は、酵母におけるDNA損傷によって活性化される。Chk1が活性化されると、G2/M移行において細胞が停止する。(Sanchez, Y. 他、(1997) *Science* 277:1497-1501.) 具体的には、Chk1が細胞分裂サイクルホスファターゼCDC25をリン酸化し、それによってサイクリン依存性キナーゼCdc2を脱リン酸化して活性化するCDC25の正常な機能が阻害される。Cdc2の活性化によって細胞が有糸分裂に入るのが調節される(Peng, C-Y 他、(1997) *Science* 277:1501-1505.)

10

20

30

40

50

。従って、Chk1の活性化によって、損傷した細胞が有糸分裂に入るのが阻止される。Chk1のようなチェックポイントキナーゼの同様の欠損によっても、G2/Mのような他のチェックポイントにおいてDNAが損傷した細胞が停止できないことにより癌が引き起こされ得る。

【0022】

細胞増殖関連キナーゼ

細胞増殖関連キナーゼは、ヒト巨核球細胞における細胞周期および細胞増殖の調節に関与する血清/サイトカイン誘導性STKである(Li, B.他、(1996) J. Biol. Chem. 271:19402-19408)。増殖関連キナーゼは、細胞分裂に関与するSTKのpolo(ショウジョウバエpolo遺伝子に由来する)ファミリーに関連する。増殖関連キナーゼは肺腫瘍組織においてダウンレギュレートされ、プロトオンコジーンであると思われる。プロトオンコジーンが正常な組織において制御を受けな

10

【0023】

5'-AMP-活性化プロテインキナーゼ

リガンド活性化STKプロテインキナーゼは、5'-AMP-活性化プロテインキナーゼ(AMPK)である(Gao, G.他、(1996) J. Biol. Chem. 271:8675-8681)。哺乳動物AMPKは、酵素であるアセチル-CoAカルボキシラーゼおよびヒドロキシメチルグルタリル-CoAレダクターゼのリン酸化による脂肪酸およびステロールの合成の調節因子であって、熱ショックやグルコースおよびATPの枯渇などの細胞内ストレスに対するこれらの経路の応答を仲介する。AMPKは、触媒サブユニットと、このサブユニットの活性を調節すると考えられている2つの非触媒サブユニットおよびサブユニットからなるヘテロ三量体複合体である。AMPKのサブユニットは、脳、心臓、脾臓、および肺などの非脂肪性組織において予想以上に広く分布している。この分布から脂質の代謝の調節以外にもある役割を果たしていると考えられる。

20

【0024】

アポトーシスにおけるキナーゼ

アポトーシスは、細胞死を導く高度に制御されたシグナル伝達経路であって、組織の発達および恒常性に極めて重要な役割を果たしている。このプロセスの調節不全は、自己免疫疾患、神経変性疾患、および癌を含む様々な疾患の病原に関連する。様々なSTKがこのプロセスにおいて重要な役割を果たす。ZIPキナーゼは、N末端プロテインキナーゼドメインに加えてC末端ロイジンジッパードメインを含むSTKである。このC末端ドメインは、ホモ二量体化およびキナーゼの活性化、ならびに転写因子のサイクリックAMP応答性エレメント結合タンパク質(ATF/CREB)ファミリーのメンバーである活性化転写因子ATF4などの転写因子との相互作用を仲介する(Sanjo, H.他、(1998) J. Biol. Chem., 273:29066-29071)。DRAK1およびDRAK2は、死関連プロテインキナーゼ(DAPキナーゼ)と相同性を共有するSTKであって、インターフェロン誘導性アポトーシスにおいて機能することが知られている(Sanjoら 前出)。ZIPキナーゼと同様に、DAPキナーゼも、N末端キナーゼドメインに加えてアンキリンリピート型のC末端タンパク質-タンパク質相互作用ドメインを含む。ZIP、DAP、およびDRAKキナーゼは、NIH3T3細胞に形質転換されるとアポトーシスに関連する形態変化を誘導する(Sanjoら, 前出)。しかしながら、これらのタンパク質のN末端キナーゼ触媒ドメインか或いはC末端ドメインのいずれかが欠失すると、アポトーシス活性がなくなる。このことから、キナーゼ活性に加えてC末端ドメインにおける活性も、アポトーシスに必要であり、調節因子または特定の基質と相互作用するドメインの可能性がある。

30

40

【0025】

RICKは、死受容体CD95を含む特定のアポトーシス経路を仲介するとして近年同定された別のSTKである(Inohara, N.他、(1998) J. Biol. C

50

hem. 273:12296-12300)。CD95は腫瘍壊死因子受容体スーパーファミリーのメンバーであって、免疫系の調節および恒常性において重要な役割を果たす(Nagata, S. (1997) Cell 88:355-365)。CD95受容体シグナル伝達経路は、様々な細胞内分子の受容体複合体への補充、およびそれに続くリガンド結合を伴う。このプロセスは、システインプロテアーゼカスパーゼ-8の補充を伴い、それによってカスパーゼカスケードが活性化され細胞死が起こる。RICKは、カスパーゼ様ドメインと相互作用するC末端の「カスパーゼ動員」ドメインとN末端のキナーゼ触媒ドメインから構成され、カスパーゼ8の動員におけるRICKの機能を示唆する。この解釈は、ヒト293T細胞におけるRICKの発現がカスパーゼ-8の活性化を促進し、CD95アポトーシス経路に関与する様々なタンパク質によるアポトーシスの誘導を増強するという事実によって支持される(Inoharaら, 前出)。

10

【0026】

ミトコンドリアプロテインキナーゼ

原核生物ヒスチジンプロテインキナーゼとアミノ酸配列で関連する、真核生物キナーゼの新規のクラスであるミトコンドリアプロテインキナーゼ(MPK)は、他の真核生物プロテインキナーゼとは配列類似性を有していないようである。これらのプロテインキナーゼは、ミトコンドリアマトリックス空間のみに存在し、原始的真核細胞によってエンドサイトーシスによって取り込まれた呼吸依存性細菌に存在した遺伝子が進化したものと思われる。MPKは、分枝鎖ケト酸デヒドロゲナーゼおよびピルビン酸デヒドロゲナーゼ複合体のリン酸化および不活化に関係する(Harris, R. A. 他、(1995) Adv. Enzyme Regul. 34:147-162)。5つのMPKがこれまで同定されている。4つのメンバーはピルビン酸デヒドロゲナーゼイソ酵素に対応し、解糖とクエン酸回路との中間において重要な調節酵素であるピルビン酸デヒドロゲナーゼ複合体の活性を調節する。5番目のメンバーは分枝鎖ケト酸デヒドロゲナーゼキナーゼに対応し、分枝鎖アミノ酸の除去のための経路の調節に重要である(Harris, R. A. 他、(1997) Adv. Enzyme Regul. 37:271-293)。飢餓状態および糖尿病状態において、ラットの肝臓、心臓および筋肉におけるピルビン酸デヒドロゲナーゼキナーゼの活性が著しく上昇することが知られている。この活性の上昇が、両方の病態におけるピルビン酸デヒドロゲナーゼ複合体のリン酸化および不活化、並びに糖新生のためのピルビン酸および乳酸の保存に貢献する(Harrisら, 前出)。

20

30

【0027】

非タンパク質基質に対するキナーゼ脂質キナーゼおよびイノシトールキナーゼ

脂質キナーゼは、脂質ヘッドグループのヒドロキシル残基をリン酸化する。ホスファチジルイノシトール(PI)のリン酸化に関与するキナーゼのファミリーが記載されており、それぞれのメンバーがイノシトール環の特定の炭素をリン酸化する(Leevers, S. 他、(1999) Curr. Opin. Cell. Biol. 11:219-225)。ホスファチジルイノシトールのリン酸化は、プロテインキナーゼCシグナル伝達経路の活性化に関与する。イノシトールリン脂質(ホスホイノシチド)細胞内シグナル伝達経路は、細胞膜においてシグナル伝達分子がGタンパク質結合受容体に結合することによって始まる。これによって、イノシトールキナーゼによる細胞膜の内側のホスファチジルイノシトール(PI)残基のリン酸化が起こり、それによってPI残基が二リン酸状態(PIP₂)に変換される。次にPIP₂がイノシトール三リン酸(IP₃)およびジアシルグリセロールに切断される。これらの2つの生成物は、シグナル伝達経路を分けるためのメディエーターとして作用する。これらの経路によって仲介される細胞応答には、バソプレッシンに応答する肝臓におけるグリコーゲン分解、アセチルコリンに応答する平滑筋の収縮、およびトロンピン誘導性血小板凝集がある。

40

【0028】

PIおよびその誘導体のD3位をリン酸化するPI3-キナーゼ(PI3K)は、細胞増殖、分化、および代謝に関与する成長因子シグナルカスケードにおいて重要な役割を果た

50

す。PI3Kは、アダプターサブユニットおよび触媒サブユニットからなるヘテロ二量体である。このアダプターサブユニットは足場タンパク質として作用し、特定のチロシン-リン酸化タンパク質、脂質成分、およびその他の細胞質因子と相互作用する。アダプターサブユニットがインスリン応答性基質(IRS)-1などのチロシンリン酸化標的と結合すると、触媒サブユニットが活性化され、PI(4,5)ニリン酸(PIP₂)をPI(3,4,5)P₃(PIP₃)に変換する。次にPIP₃が、PKA、プロテインキナーゼB(PKB)、プロテインキナーゼC(PKC)、グリコーゲンシンターゼキナーゼ(GSK)-3、およびP70リボソームs6キナーゼを含むその他の幾つかのタンパク質を活性化する。PI3Kもまた、細胞骨格形成タンパク質であるRac、rho、およびcdc42と直接相互作用する(Shepherd, P. R. 他、(1998) Biochem. J. 333:471-490)。obeseマウスおよびfatマウスなどの糖尿病動物モデルは、PI3kアダプターサブユニットのレベルが改変されている。アダプターサブユニットにおける特定の変異が、インスリン抵抗性のデンマーク人集団に見られることから、PI3Kがタイプ2糖尿病においてある役割を果たしていると思われる(Shepherd、前出)。

10

【0029】

脂質キナーゼリン酸化活性の例は、D-エリトロ-スフィンゴシンのスフィンゴ脂質代謝産物であるスフィンゴシン-1-リン酸(SPP)へのリン酸化である。SPPは、細胞外作用および細胞内作用の両方を持つ新規の脂質セカンドメッセンジャーであることが分かった(Kohama, T. 他、(1998) J. Biol. Chem. 273:23722-23728)。細胞外では、SPPはGタンパク質結合受容体EDG-1(内皮由来Gタンパク質結合受容体)のリガンドである。細胞内では、SPPは、細胞の成長、生存、運動性、および細胞骨格の変化を調節する。SPPのレベルは、D-エリトロ-スフィンゴシンを特異的にリン酸化してSPPにするスフィンゴシンキナーゼによって調節される。細胞シグナル伝達におけるスフィンゴシンキナーゼの重要性は、血小板由来成長因子(PDGF)、神経成長因子、およびプロテインキナーゼCの活性化を含む様々な刺激が、スフィンゴシンキナーゼの活性化によりSPPの細胞内レベルを上昇させるという事実、および酵素の競合阻害剤がPDGFによって誘導された細胞増殖を選択的に阻害するという事実から証明された(Kohamaら、前出)。

20

【0030】

プリンヌクレオチドキナーゼ

プリンヌクレオチドキナーゼであるアデニル酸キナーゼ(ATP:AMPホスホトランズフェラーゼ、もしくはAdK)およびグアニル酸キナーゼ(ATP:GMPホスホトランズフェラーゼ、もしくはGuK)が、ヌクレオチド代謝において重要な役割を果たし、ATPおよびGTPの合成および細胞内レベルの調節のそれぞれにおいて重要である。これらの2つの分子は、成長している細胞におけるDNA合成およびRNA合成の前駆物質であって、細胞における主な生化学エネルギーの主な供給源(ATP)であり、シグナル伝達経路(GTP)を提供する。これらの2つの分子の合成における様々なステップを妨げることが、癌および抗ウイルス治療のための多くの抗増殖剤の原理である(Pillwein, K. 他、(1990) Cancer Res. 50:1576-1579)。

30

40

【0031】

AdKはほとんどの細胞型に見られ、高い速度でATPを合成する骨格筋などの細胞に特に多く存在する。これらの細胞において、AdKは、細胞レベル下構造であるミトコンドリアおよび筋原繊維と物理的に関連する。ミトコンドリアはエネルギーを生産し、筋原繊維はそのエネルギーを利用する。近年の研究により、ATPを生成する代謝プロセスからATPを消費する細胞成分への高エネルギーのホスホリルの転移においてAdKが重要な役割を果たしていることが実証された(Zeleznikar, R. J. 他、(1995) J. Biol. Chem. 270:7311-7319)。従って、AdKは、細胞、特に癌細胞などの増殖や代謝速度が速い細胞におけるエネルギー生産の維持において中心的な役割を演じると考えられ、ある種の癌の治療においてその活性を阻害するた

50

めの候補を提供し得る。別法では、AdK活性の低下が、心不全や呼吸器不全を引き起こす様々な代謝筋エネルギー疾患の原因と考えられ、AdKの活性を上昇させて治療できるであろう。

【0032】

RNAおよびDNA合成のためのGTPの合成における重要なステップを提供するのに加えて、Gukはまた、GDPおよびGTPの調節によって細胞のシグナル伝達経路において重要な役割を果たす。具体的には、膜関連Gタンパク質に結合するGTPは細胞受容体の活性化を仲介し、アデニルシクラーゼが細胞内で活性化され、セカンドメッセンジャーであるサイクリックAMPが生産される。Gタンパク質に結合するGDPはこれらのプロセスを阻害する。また、GDPおよびGTPのレベルによって、細胞増殖の調節および発癌に關与するとして知られるp21^{ras}などのある種の腫瘍タンパク質の活性が調節される(Bos, J. L. (1989) *Cancer Res.* 49:4682-4689)。Gukの抑制によって起こるGDPに対するGTPの比率が高いと、p21^{ras}が活性化され発癌が促進される。GDPのレベルを上昇させGDPに対するGTPのレベルを低下させるためにGukの活性を高めることが、発癌を抑制する治療となり得る。

10

【0033】

Gukは、ヘルペスウイルス感染の治療に有用なある種の抗ウイルス剤のリン酸化および活性化における重要な酵素である。これらの薬剤には、グアニン相同体アシクロビルおよびbuciclovirが含まれる(Miller, W. H. およびMiller R. L. (1980) *J. Biol. Chem.* 255:7204-7207; Stenbergh, K. 他、(1986) *J. Biol. Chem.* 261:2134-2139)。感染細胞においてGukの活性を高めることが、これらの薬剤の効果を促進させる治療方法となり得るため、薬剤の服用を減らすことが可能であろう。

20

【0034】

ピリミジンキナーゼ

ピリミジンキナーゼは、デオキシチジンキナーゼ、およびチミジンキナーゼ1および2を含む。デオキシチジンキナーゼは核に存在し、チミジンキナーゼ1および2は細胞質に見られる(Johansson, M. 他、(1997) *Proc. Natl. Acad. Sci. U. S. A.* 94:11941-11945)。ピリミジンキナーゼによるデオキシリボヌクレオシドのリン酸化によって、DNA前駆物質のde novo合成のための代替の経路が提供される。リン酸化におけるピリミジンキナーゼの役割はプリンキナーゼと同様に、化学療法に重要な幾つかのヌクレオシド類似体の活性化に重要である(Arner E. S. およびEriksson, S. (1995) *Pharmacol. Ther.* 67:155-186)。

30

【0035】

新規のヒトキナーゼ、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、癌、免疫疾患、成長および発達に影響を及ぼす疾患、心血管疾患、および脂質異常の診断・治療・予防において有用であり、また、ヒトキナーゼ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。

40

【0036】

(発明の概要)

本発明は、総称して「PKIN」と呼び、また、個別にはそれぞれ「PKIN-1」、「PKIN-2」、「PKIN-3」、「PKIN-4」、「PKIN-5」、「PKIN-6」、「PKIN-7」、「PKIN-8」、「PKIN-9」、「PKIN-10」、「PKIN-11」、「PKIN-12」、「PKIN-13」、「PKIN-14」、「PKIN-15」、「PKIN-16」、「PKIN-17」、「PKIN-18」、「PKIN-19」および「PKIN-20」と呼ぶヒトキナーゼである精製ポリペプチドを特徴とする。或る実施態様において本発明は、(a) SEQ ID NO: 1-20を有す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

50

) SEQ ID NO: 1 - 20を有する群から選択したアミノ酸配列と少なくとも90%が同一である天然のアミノ酸配列を有するポリペプチド、(c) SEQ ID NO: 1 - 20を有す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 - 20を有す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 - 20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。

【0037】

また、本発明は(a) SEQ ID NO: 1 - 20を有す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 - 20を有する群から選択したアミノ酸配列と少なくとも90%が同一である天然のアミノ酸配列からなるポリペプチド、(c) SEQ ID NO: 1 - 20を有す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 - 20を有す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、を含む群から選択されたポリペプチドをコードする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 - 20を有す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: 11 - 20を有する群から選択される。

【0038】

本発明は更に、(a) SEQ ID NO: 1 - 20を有す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 - 20を有する群から選択したアミノ酸配列と少なくとも90%の相同性のある天然のアミノ酸配列からなるポリペプチド、(c) SEQ ID NO: 1 - 20を有す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 - 20を有す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有する群から選択したポリペプチドをコードするような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。

【0039】

更に本発明は、(a) SEQ ID NO: 1 - 20からなる一群から選択されたアミノ酸配列からなるポリペプチドと、(b) SEQ ID NO: 1 - 20からなる一群から選択されたアミノ酸配列と90%以上の配列同一性を有する天然のアミノ酸配列からなるポリペプチドと、(c) SEQ ID NO: 1 - 20からなる一群から選択されたアミノ酸配列を有するポリペプチドの生物学的に活性な断片と、(d) SEQ ID NO: 1 - 20とからなる一群から選択され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とで構成される一群から選択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。

【0040】

本発明は更に、(a) SEQ ID NO: 1 - 20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 - 20からなる群から選択したアミノ酸配列と少なくとも90%同一である天然のアミノ酸配列を含むポリペプチド、(c) SEQ ID NO: 1 - 20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、および(d) SEQ ID NO: 1 - 20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。

【0041】

10

20

30

40

50

本発明は更に、(a) SEQ ID NO: 21-40 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: 21-40 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%の同一性を有する天然のポリヌクレオチド配列を含むポリヌクレオチド、(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、および(e) (a) ~ (d) のRNA等価物からなる群から選択された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。

【0042】

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: 21-40 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: 21-40 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%の同一性を有する天然のポリヌクレオチド配列を含むポリヌクレオチド、(c) (a) に相補的なポリヌクレオチド、(d) (b) に相補的なポリヌクレオチド配列、および(e) (a) ~ (d) のRNA等価物からなる群から選択されたポリヌクレオチド配列を含む。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。

【0043】

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: 21-40 を有する群から選択したポリヌクレオチド配列からなるポリヌクレオチド、(b) SEQ ID NO: 21-40 を有す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%同一である天然のポリヌクレオチド配列からなるポリヌクレオチド、(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、または(e) (a) ~ (d) のRNA等価物を含む群から選択されたポリヌクレオチドの配列を有する。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。

【0044】

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: 1-20 を有す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: 1-20 を有する群から選択したアミノ酸配列と少なくとも90%同一である天然のアミノ酸配列を有するポリペプチド、(c) SEQ ID NO: 1-20 を有す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: 1-20 を有す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。一実施例では、SEQ ID NO: 1-20 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、患者にこの組成物を投与することを含む、機能的PKIN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。

【0045】

本発明はまた、(a) SEQ ID NO: 1-20 を有す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: 1-20 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を含むポリペプチド、(c) SEQ ID NO: 1-20 を有する群から選択したアミノ酸配列のポリペプチドの

生物学的活性断片、または (d) SEQ ID NO: 1 - 20 を有する群から選択したアミノ酸配列のポリペプチドの免疫抗原性断片を含む群から選択されたアゴニストとしてのポリペプチド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。更なる別法では、本発明は、この組成物の患者への投与を含む、機能的PKIN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。

【0046】

本発明は更に、(a) SEQ ID NO: 1 - 20 を有す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 - 20 を有する群から選択したアミノ酸配列と少なくとも90%が同一である天然のアミノ酸配列を有するポリペプチド、(c) SEQ ID NO: 1 - 20 を有す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 - 20 を有す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。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PKIN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。

10

20

【0047】

本発明は更に、(a) SEQ ID NO: 1 - 20 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 - 20 からなる群から選択したアミノ酸配列と少なくとも90%の同一性を有する天然のアミノ酸配列を含むポリペプチド、(c) SEQ ID NO: 1 - 20 を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 - 20 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。

30

【0048】

この方法は、(a) このポリペプチドを好適な条件下で少なくとも1つの化合物と結合させるステップと、(b) このポリペプチドとこの試験化合物との結合を検出して、このポリペプチドと特異的に結合する化合物を同定するステップとを含む。

【0049】

本発明は更に、(a) SEQ ID NO: 1 - 20 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 - 20 からなる群から選択したアミノ酸配列と少なくとも90%の同一性を有する天然のアミノ酸配列を含むポリペプチド、(c) SEQ ID NO: 1 - 20 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、および (d) SEQ ID NO: 1 - 20 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からなる群れから選択した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。

40

【0050】

50

本発明は更に、標的ポリヌクレオチドの変異発現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。標的ポリヌクレオチドは、SEQ ID NO: 21 - 40を有する群から選択した配列を含む。スクリーニング方法は、(a)標的ポリヌクレオチドを含むサンプルを化合物に曝す過程と、(b)標的ポリヌクレオチドの変異発現を検出する過程とを含む。

【0051】

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: 21 - 40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: 21 - 40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%の同一性を有する天然のポリヌクレオチド配列を含むポリヌクレオチド、(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、(v)(i)~(iv)のRNA等価物からなる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を含むプローブを用いて、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: 21 - 40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: 21 - 40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%の同一性を有する天然のポリヌクレオチド配列を含むポリヌクレオチド、(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、および(v)(i)~(iv)のRNA等価物からなる群から選択する。或いは、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、未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。

10

20

【0052】

(発明を実施するための形態)

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。しかし、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。

30

【0053】

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。

40

【0054】

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。

50

【0055】

(定義)

用語「PKIN」は、天然、合成、半合成或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PKINのアミノ酸配列を指す。

【0056】

用語「アゴニスト」は、PKINの生物学的活性を強めたり、模倣する分子を指す。このアゴニストは、PKINに直接相互作用するか、或いはPKINが関与する生物学的経路の成分と作用して、PKINの活性を調節するタンパク質、核酸、糖質、小分子、任意の他の化合物や組成物を含み得る。

10

【0057】

用語「対立遺伝子変異配列」は、PKIN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、単独或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。

【0058】

PKIN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、或いは置換が発生し、PKINと同じポリペプチド或いはPKIN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、PKIN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當或いは予期しないハイブリダイゼーション、並びにPKIN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或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PKINと機能的に等価となるアミノ酸残基の欠失、挿入、或いは置換を含み得る。意図的なアミノ酸置換は、生物学的或いは免疫学的にPKIN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。

20

30

【0059】

「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、また天然或いは合成分子を指す。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。

40

【0060】

用語「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。

【0061】

用語「アンタゴニスト」は、PKINの生物学的活性を阻害或いは減弱する分子である。アンタゴニストは、PKINに直接相互作用するか、或いはPKINが関与する生物学的経路の成分と作用して、PKINの活性を調節する抗体、核酸、糖質、小分子、任意の他の化合物や組成物などのタンパク質を含み得る。

【0062】

50

用語「抗体」は、エピトープの決定基と結合することができる、そのままの免疫グロブリンやその断片、例えば F a、F (a b ') 2 及び F v 断片を指す。P K I N ポリペプチドと結合する抗体は、免疫抗原として、そのままのポリペプチド、または関心のある小ペプチドを含む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、RNA の翻訳、または化学合成によって得られるポリペプチドまたはオリゴペプチドに由来し得るもので、好みに応じて担体タンパク質に抱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スカイガイのヘモシアニン（K L H）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。

【0063】

10

用語「抗原決定基」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。

【0064】

用語「アンチセンス」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、ホスホリチオ酸、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、或いは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。

20

【0065】

用語「生物学的に活性」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、天然或いは組換え体のP K I N、合成のP K I Nまたはそれらの任意のオリゴペプチドが、適当な動物或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。

30

【0066】

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

【0067】

用語「特定ポリヌクレオチド配列のある成分」及び「特定アミノ酸配列のある成分」は、それぞれ幅広く特定のポリヌクレオチド配列含有成分または特定のアミノ酸配列含有を指す。この成分には、乾燥製剤または水溶液が含まれ得る。P K I N 若しくはP K I N 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。

40

【0068】

「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（Applied Biosystems, Foster City

50

C A) を用いて 5 ' 及び / または 3 ' の方向に伸長され、再度シーケンシングされた核酸配列、または G E L V I E W 断片構築システム (G C G , M a d i s o n , W I) または P h r a p (U n i v e r s i t y o f W a s h i n g t o n , S e a t t l e W A) 等の断片構築用のコンピュータプログラムを用いて 1 つ或いはそれ以上の重複する c D N A や E S T、またはゲノム D N A 断片から構築された核酸配列を指す。伸長及びアセンブル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。

【 0 0 6 9 】

用語「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。

10

【 0 0 7 0 】

元残基	保的置換
Aa	Gly, Ser
Ag	His, Lys
Asn	Asp, Gln, His
Asp	Asn, Glu
Gs	Ala, Ser
元残基	保的置換
Aa	Gly, Ser
Ag	His, Lys
Asn	Asp, Gln, His
Asp	Asn, Glu
Gs	Ala, Ser
元残基	保的置換
Met	Ileu
Phe	His
Ser	Gly
Thr	Ser
Tip	Phe
Tyr	His
Val	Ile, Leu

20

30

【 0 0 7 1 】

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 シートや ヘリックス構造、(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。

【 0 0 7 2 】

「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。

【 0 0 7 3 】

用語「誘導体」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 (p e g y l a t i o n)、或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。

40

【 0 0 7 4 】

用語「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

50

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。

【0075】

用語「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。

【0076】

用語「エキソンシャッピング」は、異なったコード領域(エキソン)の組換えを指す。1つのエキソンがコードされたタンパク質の1つの構造的または機能的ドメインを代表し得るため、安定した基礎構造の新規な再分類を介して新しいタンパク質が組立てられることが可能であり、新しいタンパク質機能の進化を促進できる。

10

【0077】

用語「断片」は、PKIN または PKIN をコードするポリヌクレオチドの固有の部分であって、その親配列(parent sequence)と同一であるがその配列より長さが短いものを指す。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、或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或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。

20

【0078】

SEQ ID NO: 21 - 40の断片には、固有のポリヌクレオチド配列領域が含まれる。この領域は、SEQ ID NO: 21 - 40を特異的に同定するものであり、例えば同一ゲノム中のSEQ ID NO: 21 - 40以外の配列とは異なるものである。SEQ ID NO: 21 - 40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: 21 - 40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: 21 - 40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。

30

【0079】

SEQ ID NO: 1 - 20の断片はSEQ ID NO: 21 - 40の断片によってコードされている。SEQ ID NO: 1 - 20の断片はSEQ ID NO: 1 - 20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、配列番号1乃至20 SEQ ID NO: 1 - 20の断片は、SEQ ID NO: 1 - 20を特異認識する抗体を産出するための免疫抗原性ペプチドとして有用である。配列番号1乃至20 SEQ ID NO: 1 - 20の断片及び断片に対応するSEQ ID NO: 1 - 20の領域の正確な長さは、断片に対する意図した目的に基づき当業者が慣例的に決定することが可能である。

40

【0080】

用語「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。用語「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。

【0081】

用語「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。

【0082】

50

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。

【0083】

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 ソフトウェアパッケージ（一組の分子生物学的分析プログラム）(DNA STAR, 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。ポリヌクレオチド配列の対のアライメントの場合、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。「重み付けされた」残基重み付け表が、デフォルトとして選択された。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。

10

【0084】

或いは、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値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場合、デフォルトパラメータに設定された「BLAST 2 Sequences」ツール Version 2.0.12 (2000年4月21日)でblastnを使用できるであろう。デフォルトパラメータの設定例を以下に示す。

20

30

【0085】

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

40

【0086】

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。或いは、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

50

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。

【0087】

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。

10

【0088】

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。

【0089】

ポリペプチド配列間の一一致率は、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を用いる対方式のポリペプチド配列のアラインメントの場合、デフォルトパラメータは、Ktuple = 1、gap penalty = 3、window = 5、及び「diagonals saved」= 5と設定する。PAM250マトリクスが、デフォルトの残基重み付け表として選択される。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。

20

【0090】

或いは、NCBI BLASTソフトウェア一式を用いてもよい。例えばポリペプチド配列を2つ1組で比較をする場合、デフォルトパラメータとして設定されたblastpと共に「BLAST 2 Sequences」ツールVersion 2.0.12（2000年4月21日）を使用してもよい。デフォルトパラメータの設定例を以下に示す。

30

【0091】

Matrix: BLOSUM62
 Open Gap: 11 及び Extension Gap: 1 penalties
 Gap x drop-off: 50
 Expect: 10
 Word Size: 3
 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、配列或いは、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。

40

【0092】

「人工ヒト染色体」（HAC）は、約6 kb ~ 10 MbのサイズのDNA配列を含み得る、安定した染色体複製の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。

50

【0093】

用語「ヒト化抗体」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。

【0094】

用語「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。アニーリングが許容される条件下で、特異的なハイブリダイゼーション複合体が形成され、「洗浄ステップ」の後もハイブリダイズしたままである。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。

10

【0095】

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。この洗浄温度は通常、所定のイオン強度とpHにおける特定の配列の熱融点（ T_m ）より約5～20℃低く選択される。この T_m 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: A 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。

20

【0096】

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0.2×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℃、または42℃の温度で行うことができる。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。このようなブロッキング剤には、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。これらの洗浄条件の有用な改変は、当業者には周知である。特に高いストリンジェントな条件でのハイブリダイゼーションは、ヌクレオチド間の進化における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。

30

40

【0097】

用語「ハイブリダイゼーション複合体」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（CotまたはRot解析等）。或いは、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、ピンまたはガラスライド、或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。

【0098】

用語「挿入」及び「付加」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

50

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。

【0099】

用語「免疫応答」は、炎症性疾患及び外傷、免疫異常、感染症、遺伝病などに関連する症状を指す。これらの症状は、細胞系及び全身防衛系に影響を及ぼすサイトカイン及びケモカイン、他のシグナル伝達分子などの様々な因子の発現という特徴をもつ。

【0100】

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PKIN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PKINのポリペプチド断片またはオリゴペプチド断片を含む。

10

【0101】

用語「マイクロアレイ」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。

【0102】

用語「エレメント」または「アレイエレメント」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。

【0103】

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

20

【0104】

用語「核酸」及び「核酸配列」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。用語「核酸」及び「核酸配列」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或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類似またはRNA類似物質を指すこともある。

【0105】

用語「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、或いは連続し得る。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。

30

【0106】

用語「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。

40

【0107】

PKIN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み得る。これらのプロセスは、合成或いは生化学的に生じ得る。生化学的修飾は、PKINの酵素環境に依存し、細胞の種類によって異なり得る。

【0108】

用語「プローブ」とは、同一配列或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PKIN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

50

イソトープ、リガンド、化学発光試薬及び酵素がある。用語「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。

【0109】

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。

10

【0110】

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. 他, (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。

20

【0111】

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み込まれている。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・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またはシーケンシングブラ

30

40

50

イマーとして、マイクロアレイエレメントとして、或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。

【0112】

用語「組換え核酸」は、天然配列ではない配列であるか或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で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。

10

【0113】

或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。

【0114】

「調節エレメント」は、通常は遺伝子の非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'の非翻訳領域(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タンパク質またはウイルスタンパク質と相互作用する。

20

【0115】

「レポーター分子」は、核酸、アミノ酸または抗体の標識に用いられる化学的または生化学的な部分である。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。

【0116】

本明細書において、DNA配列に対する「RNA等価物」とは、基準となるDNA配列と同じ直鎖の核酸配列から構成されるが、窒素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。

30

【0117】

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。

【0118】

用語「サンプル」は、その最も広い意味で用いられる。PKIN、PKIN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。

40

【0119】

用語「特異結合」または「特異的に結合する」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（すなわち、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。

【0120】

用語「実質上精製された」は、自然環境から取り除かれ、或いは単離または分離された核

50

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。

【0121】

用語「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。

【0122】

用語「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、チャンネル、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。

10

【0123】

用語「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。

【0124】

用語「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。用語「形質転換された」細胞には、限られた時間内に挿入されたDNAやRNAを発現するような一時的に形質転換された細胞のみならず、安定的に形質転換された細胞であってその中に挿入されたDNAが自律的に複製するプラスミドとして或いは宿主の染色体の一部として複製可能であるものも含まれる。

20

【0125】

本明細書における「遺伝子組換え生物」とは、当分野で周知の遺伝子組換え技術などを用いて、人間が生物の1つ以上の細胞に異種の核酸を導入した任意の生物であり、動物及び植物を含むが、それらに限定されるものではない。微量注入や組換えウイルスに感染させるなどの慎重な遺伝子操作によって、細胞の前駆体に直接或いは間接的に異種核酸を細胞に導入する。遺伝子操作の語は、古典的な交雑育種或いは*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。

30

【0126】

特定の核酸配列の「変異配列」とは、デフォルトパラメータ設定の「BLAST 2 Sequences」ツールVersion 2.0.9（1999年5月7日）を用いるblastnによって、ある核酸配列のある長さに対する該特定の核酸配列の同一性が、少なくとも40%と決定された核酸配列のことである。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或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或いは参照分子に存在するドメイ

40

50

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「単一ヌクレオチド多形性」(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。

【0127】

特定の核酸配列の「変異配列」とは、デフォルトパラメータ設定の「BLAST 2 Sequences」ツールVersion 2.0.9(1999年5月7日)を用いるblastnによって、ある核酸配列のある長さに対する該特定の核酸配列の同一性が、少なくとも40%と決定された核酸配列のことである。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。

10

【0128】

(発明)

本発明は、新規のヒトキナーゼ(PKIN)及びPKIN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癌、免疫疾患、成長および発達に影響を及ぼす疾患、心血管疾患、および脂質異常の診断、治療、及び予防に関する。

【0129】

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。

20

【0130】

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もって本明細書の一部とする。

30

【0131】

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。表2及び表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ヒトキナーゼであることを確立している(表2参照)。例えば、SEQ ID NO:2は、Basic Local Alignment Se

40

50

arch Tool (BLAST)で決定されるようにマウスの t o u s l e d 様類似キナーゼ (GenBank ID g 2 8 5 3 0 3 1) と 9 7 % 同一である (表 2 参照)。BLAST 確率スコアは 0 . 0 であり、これは観測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 : 2 はまた、真核細胞キナーゼ活性部位ドメインを有するが、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、MOTIFS、及び PROFILE SCAN 解析よりのデータは、SEQ ID NO : 2 が攪乱された類似キナーゼであることで、さらに実証的な証拠を提供する。別の例において、SEQ ID NO : 1 0 は BLAST によって同定されるようにヒトのセリン/トレオニンタンパク質キナーゼ (GenBank ID g 3 6 6 1 5) に 6 3 % の同一性がある (表 2 参照)。BLAST 確率スコアは 7.7×10^{-122} であり、これは観測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 0 はまた、キナーゼ活性部位ドメインを有するが、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。BLAST、MOTIFS、及び PROFILE SCAN 解析よりのデータは、SEQ ID NO : 1 0 がセリン/トレオニン・キナーゼである、さらに実証的な証拠を提供する。「セリン/トレオニン・キナーゼ」がある特定のキナーゼであることに注意する (表 2 参照)。別の例において、SEQ ID NO : 1 6 は BLAST によって同定される予想されるヒトの受容体プロテインチロシン・キナーゼ (GenBank ID g 5 5 1 6 0 8) に 5 3 % の同一性を有する (表 2 参照)。BLAST 確率スコアは 4.1×10^{-290} であり、これは観測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 6 はまた、真核細胞キナーゼ活性部位ドメインを有するが、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。BLIMPS および PROFILE SCAN 分析から得たデータは、SEQ ID NO : 1 6 が受容体チロシン・キナーゼであることを裏づける証拠を更に提供する。別の例において、SEQ ID NO : 1 9 は BLAST によって同定されるようにラットのカルシウム/カルモジュリン依存性タンパク質キナーゼ (GenBank ID g 1 8 3 6 1 6 1) に 9 3 % 同一である (表 2 参照)。BLAST 確率スコアは 6.0×10^{-257} であり、それは、偶然に観測されたポリペプチド配列を得る確率を示す。SEQ ID NO : 1 9 はまた、真核細胞タンパク質キナーゼ活性部位ドメインを有するが、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、MOTIFS 及び PROFILE SCAN 解析よりのデータは、SEQ ID NO : 1 9 がタンパク質キナーゼである、さらに実証的な証拠を提供する。SEQ ID NO : 1、SEQ ID NO : 3 - 9、SEQ ID NO : 1 1 - 1 5、SEQ ID NO : 1 7 - 1 8 及び SEQ ID NO : 2 0 については、同様の方法で分析し、注釈を付けた。SEQ ID NO : 1 - 2 0 の解析のためのアルゴリズム及びパラメータが表 7 で記述されている。

【 0 1 3 2 】

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、或いはこれら 2 種類の配列を任意に組み合わせ構築した。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 : 2 1 - 4 0 を同定するため、或いは SEQ ID NO : 2 1 - 4 0 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アセンブル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。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ライブラリを指す場合もある。例えば、2564295H1はIncyte cDNA配列の識別番号であり、ADRETUT01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、プールされたcDNAライブラリから得られた（例えば71191190V1）。或いは、列5の識別番号が、全長ポリヌクレオチド配列の構築に寄与したGenBankのcDNAまたはEST（例えばg1164223）を意味する場合がある。更に、列5の識別番号はENSEMBL（The Sanger Centre, ケンブリッジ（英国））データベースから得られた配列（例えば、「ENST」名称を含む配列群）を指す場合もある。或いは、列5の識別番号は、NCBI RefSeq Nucleotide Sequence Records データベースから由来する場合もあり（即ち「NM」または「NT」の命名を含む配列）、またNCBI RefSeq Protein Sequence Recordsから由来する場合もある（即ち「NP」の命名を含む配列）。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がある。例えば、FL__XXXXXX__N₁__N₂__YYYYY__N₃__N₄ は「ステッチされた」配列であり、その内、XXXXXXはアルゴリズムが適用される配列のクラスターの識別番号であり、YYYYYはアルゴリズムにより生み出される予測の数であり、N₁, N₂, N₃, ... がある場合には、解析中に手で編集された特別のエキソンを表す（実施例5参照）。または列5の識別番号は、「エキソンストレッチング」アルゴリズムにより結び合わせたcDNA及びGenscan予想エキソンの両方の集合を意味する場合がある。例えば、FLXXXXXX__gAAAAA__1__Nは「ストレッチ」配列の識別番号である。ここでXXXXXXはIncyteプロジェクト識別番号、gAAAAAは「エキソンストレッチング」アルゴリズムを適用したヒトゲノム配列のGenBank識別番号、gBBBBBBは一番近いGenBankタンパク質相同体のGenBank識別番号、またはNCBI RefSeq識別番号、Nは特定のエキソンである（実施例5を参照）。あるRefSeq配列が「エキソンストレッチング」アルゴリズムのためのタンパク質相同体として使用された場合は、RefSeq識別番号（「NM」、「NP」、または「NT」によって表される）が、GenBank識別（即ち、gBBBBBB）の代わりに使用される場合もある。

【0133】

また、接頭コードは、手で編集された構成配列、ゲノムDNA配列から予測された構成配列、または組み合わされた配列解析方法から由来する構成配列を同定する。以下の表は、構成配列の接頭コードと、接頭コードに対応する配列分析方法の例を列記する（実施例4と5を参照）。

10

20

30

40

接頭コード	分析法またはプログラムの例
GNJCGENST	例えば GENSCAN (Stanford University, CA, USA) または FCENES (Computer Graphics Group, The Sanger Centre, Cambridge, UK) を用いたゲノム配列からのエキソン予測
GH	ゲノム配列の再編成分析
EL	ストレッチまたはストレッチされたゲノム配列 (実施例を参照)。
INCY	EST 配列がゲノムのマッピングからの完全長とエキソンの予想エキソンと塩基を予想するために、ゲノム位置と EST 構成データが組合せられる。

10

20

【0134】

場合によっては、最終コンセンサスポリヌクレオチド配列を確認するための列 5 に示すような配列の適用範囲と重複する Incyte cDNA の適用範囲が得られたが、関連する Incyte cDNA 識別番号は示さなかった。

【0135】

表 5 は、Incyte cDNA 配列を用いてアセンブル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アセンブル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。

30

【0136】

本発明はまた、PKIN の変異体も含む。好適な PKIN の変異体は、PKIN の機能的或いは構造的特徴の少なくともどちらか一方を有し、かつ PKIN 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、或いは少なくとも約 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。

【0137】

本発明はまた、PKIN 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、PKIN をコードする SEQ ID NO : 21 - 40 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 : 21 - 40 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。

40

【0138】

本発明はまた、PKIN 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、PKIN 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、或いは少なくとも 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 : 21 - 40 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、或いは少なくとも

50

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: 21-40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、PKINの機能的或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。

【0139】

遺伝暗号の縮重により作り出され得るPKIN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PKIN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。

10

【0140】

PKIN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PKIN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PKIN或いはその誘導体をコードするヌクレオチド配列を作る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、特定の真核又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、PKIN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。

20

【0141】

本発明はまた、PKIN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、PKINまたはその任意の断片をコードする配列の中に突然変異を導入することも可能である。

【0142】

更に、合成化学を用いて、PKINまたはその任意の断片をコードする配列の中に突然変異を導入することも可能である。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: 21-40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; and 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。

30

【0143】

DNAシーケ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シーケンシング方法を用いて実施可能である。DNAシーケ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。或いは、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、MICROLAB 2200液体転移システム (Hamilton, Reno, NV)、PTC 200サーマルサイクラー (MJ Research, Watertown MA)及びABI CATAL

40

50

YST 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 或いは 377 DNA シークエンシングシステム (PE Biosystems)、MEGABACE 1000 DNA シークエンシングシステム (Molecular Dynamics, Sunnyvale CA) または当分野で周知の他の方法を用いてシークエンシングを行う。得られた配列を当分野で周知の様々なアルゴリズムを用いて分析する (例えば、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 ページを参照)。

【0144】

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、ランダムに初回抗原刺激を受けたライブラリは、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。

【0145】

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。

【0146】

本発明の別の実施例では、PKIN 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 PKIN、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、実質的に同じ或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 PKIN のクローン化及び発現に利用可能である。

【0147】

種々の目的で PKIN 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。

【0148】

本発明のヌクレオチドを、MOLECULAR BREEDING (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; Cramer, A. 他 (1996) Nat. Biotechnol. 14:315-319) などの DNA シャッフリング技術を用いてシ

10

20

30

40

50

ャフリングして、PKINの生物学的または酵素的な活性、或いは他の分子や化合物と結合する能力などのPKINの生物学的特性を変更或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。

【0149】

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生み出される。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。或いは、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。

10

【0150】

別の実施例によれば、PKIN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、全体或いは一部が合成可能である(例えば、Caruthers, M. H. 他、(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PKIN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等を参照)。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PKIN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。

20

30

【0151】

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ケンシングによって確認することができる(前出のCreighton, 28-53ページ等を参照)。

【0152】

生物学的に活性なPKINを発現させるために、PKIN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。これらのエレメントには、ベクター及びPKIN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、構造型及び発現誘導型のプロモーター、5'及び3'の非翻訳領域などの調節配列が含まれる。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、PKIN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。PKIN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、コーディング配列或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

40

50

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。(例えば、Scharf, D. 他 (1994) Results Probl. Cell Differ. 201-18-162. を参照)。

【0153】

当業者に周知の方法を用いて、PKIN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章及び13章1-4章を参照)。種々の発現ベクター/宿主系を利用して、PKIN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) EMBO J. 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等を参照)。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、Buller, 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等を参照)。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。

【0154】

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。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、PKIN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、PKIN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。PKINをコードする配列の、ベクターの多数のクローニング部位へのライゲーション反応によって、la 50

c Z 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を同定する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、ジデオキシのシーケンシング、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: 5503 - 5509 を参照)。例えば、抗体の産生のためなどに多量の PKIN が必要な場合は、PKIN の発現を高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。

【0155】

PKIN の発現に酵母の発現系の使用が可能である。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或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 *Saccharomyces cerevisiae* 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* 121 - 181 - 184. を参照）。

【0156】

植物系を使用して PKIN を発現することも可能である。PKIN 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: 3073 - 311) 由来のオメガリーダー配列と組み合わせる用いられるような CaMV 由来の 35S および 19S プロモーターによって促進される。或いは、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 (Coruzzi, G. 他 (1984) *EMBO J.* 3: 1671 - 1680、Broglie, R. 他 (1984) *Science* 224: 838 - 843、Winter, J. 他 (1991) *Results Probl. Cell Differ.* 17: 85 - 105 等を参照)。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（『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 ページ等を参照）。

【0157】

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 PKIN 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 PKIN 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. Shenk (1984) *Proc. Natl. Acad. Sci. USA* 81: 3655 - 3659 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。

【0158】

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 6 kb ~ 10 Mb 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）。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 PKIN 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、PKIN をコードする配列を株化細胞に形質転換するこ

10

20

30

40

50

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、同じ或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。

【0159】

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; 及びLowy, I. 他(1980) Cell 22: 817-823を参照)。また、選択の基礎として代謝拮抗物質、抗生物質或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(Wigler, M. 他、(1980) Proc. Natl. Acad. Sci. USA 77: 3567-3570; 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(Hartman, S.C. 及びR.C. Mulligan(1988) Proc. Natl. Acad. Sci. USA 85: 8047-8051等を参照)。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)、グルクロニダーゼ及びその基質グルクロニド、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、形質転換体を同定するだけでなく、特定のベクター系に起因する一過性或いは安定したタンパク質発現を定量することも可能である(Rhodes, C.A. (1995) Methods Mol. Biol. 55: 121-131等を参照)。

【0160】

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、PKIN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、PKIN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PKIN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。

【0161】

一般に、PKINをコードする核酸配列を含み、PKIN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或いはDNA-RNAハイブリダイゼーションや、PCR増幅法、核酸或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、或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。

【0162】

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PKINの発現の検出及び計測のための免疫学的方法は、当分野で周知である。このような技法には、酵素に結合したイムノソルベントアッセイ(ELISA)、ラジオイムノアッセイ(RIA)、蛍光標示式細胞分取器(FACS)などがある。PKIN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

10

20

30

40

50

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その他のアッセイは、当分野では十分に知られている。(例えば、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; 及び Pound, J. D. (1990) Immunochemical Protocols, Human Press, Totowa NJ)。

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。PKIN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或いは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。別法として、PKINをコードする配列、またはその任意の断片をmRNAプローブの生成の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或いは標識には、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。

【0163】

PKIN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、或いはその両者に依存する。PKIN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PKIN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。

【0164】

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, Manassas, 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。

【0165】

本発明の別の実施例では、PKINをコードする自然或いは変更された、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PKINタンパク質が、PKINの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促進し得る。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)、

10

20

30

40

50

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、PKIN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、PKIN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。

【0166】

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PKIN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³⁵S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。

【0167】

本発明のPKINまたはその断片を用いて、PKIN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、PKIN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。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PKIN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)の5章等を参照)。同様に、化合物は、PKINが結合する天然受容体、或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、分泌タンパク質或いは細胞膜上のタンパク質の何れか一方としてPKIN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。PKINを発現する細胞またはPKINを含有する細胞膜断片を試験化合物と接触させて、PKIN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。

【0168】

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、溶液中の或いは固体支持物に固定されたPKINと結合させるステップと、PKIN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、無細胞再構成系、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。

【0169】

本発明のPKINまたはその断片を用いて、PKIN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、PKINが少なくとも1つの試験化合物と結合する、PKIN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PKINの活性が試験化合物不在下でのPKINの活性と比較する。試験化合物の存在下でのPKINの活性の変化は、PKIN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PKINの活性に適した条件下でPKINを含む *in vitro* または無細胞再構成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、PKIN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。

【0170】

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、PKIN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が「ロックアウト」される。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）。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ロックアウトする（Marth, J.D. (1996) *Clin. Invest.* 97:1999-2002; Wagner, K.U. 他 (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。

10

20

【0171】

PKIN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* 282:1145-1147参照）。

【0172】

PKIN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「ロックイン」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。ロックイン技術を用いて、PKIN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PKINを乳汁内に分泌するなどPKIN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) *Biootechnol. Annu. Rev.* 4:55-74）。

30

【0173】

（治療）

PKINのある領域とヒトキナーゼのある領域との間に、例えば配列及びモチーフの内容における化学的及び構造的類似性が存在する。更に、PKINの発現は明らかに膀胱癌、前立腺、卵巣、脳、結腸、回腸、陰茎、皮膚、副腎腫瘍および消化器官、および腫瘍の各組織に密接に関連する。

40

【0174】

従って、PKINは、癌、免疫疾患、成長および発達に影響を及ぼす疾患、心血管疾患、および脂質異常においてある役割を果たすと考えられる。PKINの発現若しくは活性の増大に関連する疾患の治療においては、PKINの発現または活性を低下させることが望ましい。また、PKIN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、PK

50

INの発現または活性を増大させることが望ましい。

【0175】

従って、一実施例において、PKIN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PKIN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癌、免疫疾患、成長及び発達に影響を及ぼす障害、アテローム性動脈硬化症及びその他の心血管疾患、及び脂質異常が含まれ、癌の中には腺癌及び白血病、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、頸部、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、多発骨髄腫などの白血病、悪性リンパ腫などのリンパ腫が含まれ、免疫疾患の中には、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、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ー(APECED)、気管支炎、胆嚢炎、接触皮膚炎、クローン病、アトピー性皮膚炎、皮膚筋炎、糖尿病、肺気腫、リンパ球毒素性一時性リンパ球減少症、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、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、リウマチ様関節炎、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、血小板減少性紫斑病、潰瘍性大腸炎、ブドウ膜炎、ウェルナー症候群、癌、血液透析および体外循環の合併症、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、成長および発達に影響を及ぼす疾患の中には日光性角化症及びアテローム性動脈硬化、滑液包炎、硬変、肝炎、混合型結合組織病(MCTD)、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、並びに腺癌及び白血病、リンパ腫、黒色腫、骨髄腫、肉腫、及び奇形癌、具体的には、副腎、膀胱、骨、骨髄、脳、乳房、頸部、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、子宮の癌や、尿細管性アシドーシス、貧血、クッシング症候群、軟骨形成不全性小人症、デュシェンヌ ベッカー型筋ジストロフィー、癩癧、性腺形成異常、WAGR症候群(ウィルムス腫瘍、無虹彩症、尿生殖器異常、精神薄弱)、スミス マジェニス症候群(Smith-Magenis syndrome)、脊髄形成異常症候群、遺伝性粘膜上皮異形成、遺伝性角皮症、シャルコー マリー ツース病及び神経線維腫症などの遺伝性神経病、甲状腺機能低下症、水頭症、舞踏病(Sydenham's chorea)及び脳性小児麻痺などの発作障害、脊髄二分裂、無脳症、頭蓋脊椎披裂、先天性緑内障、白内障、感覚神経性聴力損失が含まれ、心血管疾患の中には、動静脈瘻、アテローム硬化、高血圧、脈管炎、レイノー病、静脈奇形、動脈解離、静脈瘤、血栓静脈炎及び静脈血栓、血管の腫瘍、血栓崩壊の合併症、バルーン血管形成術、血管置換術、バイパス手術、うっ血性心不全、虚血性心疾患、狭心症、心筋梗塞、高血圧性心疾患、変性弁膜性心疾患、石灰化大動脈弁狭窄症、先天性2尖大動脈弁、僧帽弁輪状石灰化、僧帽弁逸脱、リウマチ熱、リウマチ性心疾患、感染性心内膜炎、非細菌性血栓性心内膜炎、全身性エリテマトーデスの心内膜炎、カルチノイド心疾患、心筋症、心筋炎、心膜炎、腫瘍性心疾患、先天性心臓疾患、及び心臓移植の合併症と、先天性肺異常、肺拡張不全、肺うっ血及び肺水腫、肺動脈塞栓症、肺出血、肺梗塞、肺高血圧症、血管硬化症、閉塞性肺疾患、拘束型肺疾患、慢性閉塞性肺疾患、肺気腫、慢性気管支炎、気管支喘息、細気管支拡張症、細菌性肺炎、ウイルス性肺炎及びマイコプラズマ肺炎、肺膿瘍、肺結核、びまん性間質性疾患、塵肺症、サルコイド症、特発性肺線維症、剥離性間質性肺炎、過敏症肺炎、肺好酸球増加閉塞性細気管支炎 器質性肺炎(pulmonary eosinophilia bronchiolitis obliterans-organizing pneumonia)、びまん性肺出血症候群、グッドパスチャー症候群、特発性肺血鉄症、肺併発膠原血管病、肺胞たんぱく症、肺腫瘍、炎症性及び非炎症性胸水、気胸症、

胸膜腫瘍、薬物性肺疾患、放射線性肺疾患及び肺移植の合併症などが含まれ、脂質異常の中には、脂肪肝、胆汁うっ滞、原発性胆汁性肝硬変、カルニチン欠乏症、カルニチンパルミトイルトランスフェラーゼ欠乏症、ミオアデニレートデミナーゼ欠乏症、高トリグリセリド血症、ファブリー病などの脂質貯蔵病、ゴーシェ病、ニーマンピック病、変染色性白質ジストロフィー、副腎性白質ジストロフィー、 G_{M_2} にガングリオシドーシス、セロイドリポフスチン症、無リポ蛋白血症、タンジアー病、リポ蛋白過剰血症、糖尿病、脂肪異栄養症、脂肪腫症、急性皮下脂肪組織炎、播種性脂肪組織壊死症、有痛脂肪症、リポイド副腎過形成、リポイドネフローゼ、脂肪腫、アテローム性動脈硬化症、高コレステロール血症、高トリグリセリド血症を伴った高コレステロール血症、原発性低リポ蛋白血症、甲状腺機能低下症、腎臓病、肝疾患、レシチン-コレステロールアシルトランスフェラーゼ欠乏症、脳髄黄色腫症、シトステロール血症、低コレステロール血症、テイサク病、サンドホフ病、高脂血症、脂肪過剰血症、脂質筋障害、肥満症が含まれる。

10

【0176】

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PKIN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、PKIN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することも可能である。

【0177】

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PKIN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PKIN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。

20

【0178】

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PKIN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、PKIN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。

【0179】

更なる実施例では、PKIN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PKIN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、上記した癌、免疫疾患、成長及び発達に影響を及ぼす障害、アテローム性動脈硬化症及びその他の心血管疾患、及び脂質異常が含まれる。一実施態様では、PKINと特異的に結合する抗体が直接アンタゴニストとして、或いはPKIN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或いは送達機構として間接的に用いられ得る。

30

【0180】

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PKIN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、PKINをコードするポリヌクレオチドの相補配列を発現するベクターを患者に投与することも可能である。

【0181】

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 of 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。

40

【0182】

PKIN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PKIN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PKINと特異的に結合するものを同定が可能である。PKIN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

50

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。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、PKIN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、ブルニックポリオール、ポリアニオン、ペプチド、油性乳剤、スカシガイのヘモシニアン、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 ゲラン桿菌）及びコリネバクテリウム パルヴムが特に好ましい。

10

【0183】

PKIN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。PKIN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。

【0184】

PKIN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（Kohler, G. 他、(1975) Nature 256: 495-497、Kozbor, D. 他、(1985) J. Immunol. Methods 81: 31-42、Cote, R. J. 他、(1983) Proc. Natl. Acad. Sci. USA 80: 2026-2030、Cole, S. P. 他、(1984) Mol. Cell Biol. 62: 109-120等を参照）。

20

【0185】

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、Morrisson, S. L. 他、(1984) Proc. Natl. Acad. Sci. 81: 6851-6855; Neuberger, M. S. 他、(1984) Nature 312: 604-608; Takeda, S. 他、(1985) Nature 314: 452-454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、PKIN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等を参照）。

30

【0186】

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、或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（Orlandi, R. 他、(1989) Proc. Natl. Acad. Sci. USA 86: 3833-3837、Winter, G. 他、(1991) Nature 349: 293-299等を参照）。PKINに対する特異的な結合部位を含む抗体断片も得ることができ。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減らすことによって作製されるFab断片とがある。或いは、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

40

50

(Huse, W. D. 他 (1989) Science 256:1275-1281 等を参照)。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、PKIN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PKIN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、競合的な結合アッセイも利用することができる(Pound、前出)。

【0187】

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、PKINに対する抗体の親和性を評価する。親和性を結合定数 K_a で表すが、この K_a は、平衡状態の下でPKIN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PKIN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医薬の K_a は、PKINに対する抗体の平均親和性または結合活性を表す。特定のPKINエピトープに単一特異的なモノクローナル抗体試薬の K_a は、親和性の真の測定値を表す。 K_a 値が $10^9 \sim 10^{12}$ liter/molの高親和性抗体医薬は、PKIN抗体複合体が過酷な処理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。 K_a 値が $10^6 \sim 10^7$ liter/molの低親和性抗体医薬は、PKIN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. and Cryer, A. (1991) A Practical Guide to Monoclonal Antibodies, John Wiley & Sons, New York NY)。

【0188】

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1~2 mg/mlの特異的な抗体、好ましくは5~10 mg/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、PKIN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Colliganらの文献等を参照)。

【0189】

本発明の別の実施例では、PKIN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、PKIN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、PKIN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)。

【0190】

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 Scanl

10

20

30

40

50

on, 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 等を参照)。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. 等を参照)。

【0191】

本発明の別の実施例では、PKIN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) Science 270: 475-480、Bordignon, C. 他、(1995) Science 270: 470-475)、嚢胞性繊維症 (Zabner, J. 他、(1993) Cell 75: 207-216; Crystal, R. G. 他、(1995) Hum. Gene Therapy 6: 643-666、Crystal, R. G. 他、(1995) Hum. Gene Therapy 6: 667-703)、サラセミア (thalassaemia)、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) Science 270: 404-410、Verma, I. M. and Somia, N. (1997) Nature 389: 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) Nature 335: 395-396、Poeschl, 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。PKIN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PKIN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。

【0192】

本発明の更なる実施例では、PKINの欠損による疾患や異常症は、PKIN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PKIN欠損細胞に導入することによって治療する。in vivo 或いは ex vivo 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) がある。

【0193】

PKIN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。PKIN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 -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) Science 268:1766-1769; Rossi, F.M.V. and H.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. and H.M. Blau, 前出)、または(iii)正常な個体に由来するPKIN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。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. および A.J. Eb (1973) Virology 52:456-467) 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。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。

【0194】

本発明の別の実施例では、PKIN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PKIN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親和性を有するエンベロープ遺伝子またはVSVG等の汎親和性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. and A.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)。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 supe

r n a t a n t」)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. 他、(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)。

10

【0195】

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、PKINの発現に関連する1つ或いは複数の遺伝子異常を有する細胞にPKIN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は当分野では周知である。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臍臓の無損傷の臍島の中に導入するために用途が広いことが証明された(Csete, M.E. 他、(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. および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。別法では、ヘルペス系遺伝子治療の送達系を用いて、PKINの発現に関連する1つ或いは複数の遺伝子異常を有する標的細胞にPKINをコードするポリヌクレオチドを送達する。単純疱疹ウイルス(HSV)系のベクターは、HSV親和性の中枢神経細胞にPKINを導入する際に特に重要である。ヘルペス系ベクターの作製及びパッケージングは当分野では周知である。複製適格性の単純疱疹ウイルス(HSV)I型系のベクターは、霊長類の眼にレポーター遺伝子を送達するために用いられてきた(Liu, X. 他(1999) Exp. Eye Res. 169:385-395)。HSV-1ウイルスベクターの作製は、DeLucaに付与された米国特許第5,804,413号(Herpes simplex virus swains for gene transfer)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。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コントロールの下で、細胞に導入される少なくとも1つの内在性遺伝子を含むゲノムからなる組換え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

20

30

40

【0196】

別法では、ウイルス(正の一本鎖RNAウイルス)ベクターを用いてPKIN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.

50

B i o t e c h . 9 : 4 6 4 - 4 6 9) 。 ウィルスRNAの複製中に、通常はウィ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、PKINをコードする配列を ウィルスゲノムのキャプシッ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PKINをコードするRNAが産生され、高いレベルでPKINが合成される。通常はウィ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、 ウィ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. 他、(1997) *Virology* 228 : 74 - 83)。様々な宿主に ウィルスを導入できることから、様々なタイプの細胞にPKINを導入することができる。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。 ウィルスの感染性cDNAクローンの処置方法、 ウィルスのcDNA及びRNAの形質移入方法及び ウィルスの感染方法は、当業者に公知である。

10

【0197】

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. E. 及び B. I. Carr, Molecular and Immunologic Approaches, Futura Publishing Co., Mt. Kisco, NY, 163 - 177* ページ等を参照)。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。

20

【0198】

リボザイムは酵素性RNA分子である。RNAの特異性切断を触媒する目的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参与している。例えば、人工的に作製されたハンマーヘッド型リボザイム分子が、PKINをコードする配列の内ヌクレオチド鎖分解性の切断を特異的且つ効果的に触媒できる可能性がある。

30

【0199】

任意のRNA標的内の特異的リボザイム切断部位を、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する。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。

40

【0200】

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。或いは、PKIN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。或いは、相補的RNAを構成的或

50

いは誘導的に合成するようなこれら c D N A 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。

【0201】

細胞内の安定性を高め、半減期を長くするために R N A 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、P N A 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 (q u e o s i n e) 、ワイプトシン (w y b u t o s i n e) 等を加えることができる。

10

【0202】

本発明の更なる実施例は、P K I N 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、P K I N 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、P K I N 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、P K I N 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、P K I N 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。

20

【0203】

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。P K I N 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。P K I N 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、P K I N 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 *S c h i z o s a c c h a r o m y c e s p o m b e* 遺伝子発現系 (A t k i n s , D . 他、(1 9 9 9) 米国特許第 5 , 9 3 2 , 4 3 5 号、A r n d t , G . M . 他、(2 0 0 0) *N u c l e i c A c i d s R e s .* 2 8 : E 1 5) または *H e L a* 細胞等のヒト細胞系 (C l a r k e , M . L . 他、(2 0 0 0) *B i o c h e m . B i o p h y s . R e s . C o m m u n .* 2 6 8 : 8 - 1 3) 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

30

40

50

(Bruce, T.W. 他、(1997) 米国特許第5,686,242号、Bruce, T.W. 他(2000) 米国特許第6,022,691号)。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。移植、リポソム注射、またはポリカチオンアミノ高分子による移動はよく知られる方法を適用できる(例えば、Goldman, C.K. 他(1997) *Nat. Biotechnol.* 15:462-466 参照)。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。

10

【0204】

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、PKIN、PKINの抗体、擬態、アゴニスト、アンタゴニスト、またはPKINのインヒビターなどからなる。

【0205】

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。

20

【0206】

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、従来の低分子量有機薬剤)の場合には、速効製剤のエアロゾル輸送が当分野で周知である。高分子(例えばより大きなペプチドやタンパク質)の場合には、肺の肺泡領域を介する肺輸送の技術が近年向上したため、インスリン等の薬剤を実際に血中に輸送することが可能となった(Patton, J.S. 他、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。

30

【0207】

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。

【0208】

PKIN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、PKIN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) *Science* 285:1569-1572)。各成分において治療的効果の投与量は、neoplastic cells など細胞内培養エッセイまたマウス、ラット、ウサギ、イヌ、サル、ブタなど動物モデルで初期に予測できる。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。

40

【0209】

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PKINまたはその断片、PKINの抗体、PKIN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ED₅₀(集団の50%の医薬的有効量)またはLD₅₀

50

0 (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、毒性を殆ど或いは全く含まず、 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。

【0210】

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、或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、薬剤の組み合わせ、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、或いは2週間に1度の間隔で投与し得る。

10

【0211】

通常薬用量は投与の経路によって異なるが、約0.1~100,000 μ gまでの最大約1グラムまでである。特定の薬用量及び送達の方法に関するガイダンスは文献に記載されており、一般に当分野の実務者は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とは異なったヌクレオチドの製剤を利用するであろう。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。

20

【0212】

(診断)

別の実施例では、PKINに特異的に結合する抗体が、PKINの発現によって特徴付けられる疾患の診断、またはPKINやPKINのアゴニストまたはアンタゴニスト、インヒビター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用いられる。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。PKINの診断アッセイには、抗体及び標識を用いてヒトの体液或いは細胞や組織から採取されたものからPKIN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。

30

【0213】

PKIN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、変わった或いは異常なレベルのPKINの発現を診断する元となるものを提供する。正常或いは標準的なPKIN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PKIN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、及び疾患生検組織からの各サンプルのPKINの発現の量が基準値と比較される。基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。

40

【0214】

本発明の別の実施例によれば、PKIN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PKIN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、PKINの存在の有無、更に過剰な発現を調べ、治療中のPKIN値の調節を監視する。

【0215】

50

一実施形態では、PKIN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、PKIN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或いは増幅のストリンジエントは、プローブがPKINをコードする自然界の配列のみを同定するかどうか、或いは対立遺伝子や関連配列を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。プローブはまた、関連する配列の検出に利用され、PKIN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: 21-40の配列、或いはPKIN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。

10

【0216】

PKINをコードするDNAに対して特異的なハイブリダイゼーションプローブの作製方法には、PKIN及びPKIN誘導体をコードするポリヌクレオチド配列をmRNAプローブの作製のためのベクターにクローニングする方法がある。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、³²Pまたは³⁵S等の放射性核種、或いはアビジン/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。

20

【0217】

PKINをコードするポリヌクレオチド配列を用いて、PKINの発現に関連する疾患を診断することが可能である。特に限定するものではないが、このような疾患には癌、免疫疾患、成長及び発達に影響を及ぼす障害、アテローム性動脈硬化症及びその他の心血管疾患、及び脂質異常が含まれ、癌の中には腺癌及び白血病、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、頸部、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、多発骨髄腫などの白血病、悪性リンパ腫などのリンパ腫が含まれ、免疫疾患の中には、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、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ー(APECED)、気管支炎、胆嚢炎、接触皮膚炎、クローン病、アトピー性皮膚炎、皮膚筋炎、糖尿病、肺気腫、リンパ球毒素性一時的リンパ球減少症、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、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、リウマチ様関節炎、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、血小板減少性紫斑病、潰瘍性大腸炎、ブドウ膜炎、ウェルナー症候群、癌、血液透析および体外循環の合併症、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、成長および発達に影響を及ぼす疾患の中には日光性角化症及びアテローム性動脈硬化、滑液包炎、硬変、肝炎、混合型結合組織病(MCTD)、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、並びに腺癌及び白血病、リンパ腫、黒色腫、骨髄腫、肉腫、及び奇形癌、具体的には、副腎、膀胱、骨、骨髄、脳、乳房、頸部、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、子宮の癌や、尿細管性アシドーシス、貧血、クッシング症候群、軟骨形成不全性小人症、デュシェンヌベッカー型筋ジストロフィー、癲癇、性腺形成異常、WAGR症候群(ウィルムス腫瘍、無虹彩症、尿

30

40

50

生殖器異常、精神薄弱)、スミス マジェニス症候群 (Smith - Magenis syndrome)、脊髄形成異常症候群、遺伝性粘膜上皮異形成、遺伝性角皮症、シャルコー マリー ツース病及び神経線維腫症などの遺伝性神経病、甲状腺機能低下症、水頭症、舞蹈病 (Sydenham's chorea) 及び脳性小児麻痺などの発作障害、脊髄二分裂、無脳症、頭蓋脊椎披裂、先天性緑内障、白内障、感覚神経性聴力損失が含まれ、心血管疾患の中には、動静脈瘤、アテローム硬化、高血圧、脈管炎、レイノー病、静脈奇形、動脈解離、静脈瘤、血栓静脈炎及び静脈血栓、血管の腫瘍、血栓崩壊の合併症、バルーン血管形成術、血管置換術、バイパス手術、うっ血性心不全、虚血性心疾患、狭心症、心筋梗塞、高血圧性心疾患、変性弁膜性心疾患、石灰化大動脈弁狭窄症、先天性2尖大動脈弁、僧帽弁輪状石灰化、僧帽弁逸脱、リウマチ熱、リウマチ性心疾患、感染性心内膜炎、非細菌性血栓性心内膜炎、全身性エリテマトーデスの心内膜炎、カルチノイド心疾患、心筋症、心筋炎、心膜炎、腫瘍性心疾患、先天性心臓疾患、及び心臓移植の合併症と、先天性肺異常、肺拡張不全、肺うっ血及び肺水腫、肺動脈塞栓症、肺出血、肺梗塞、肺高血圧症、血管硬化症、閉塞性肺疾患、拘束型肺疾患、慢性閉塞性肺疾患、肺気腫、慢性気管支炎、気管支喘息、細気管支拡張症、細菌性肺炎、ウイルス性肺炎及びマイコプラズマ肺炎、肺膿瘍、肺結核、びまん性間質性疾患、塵肺症、サルコイド症、特発性肺線維症、剥離性間質性肺炎、過敏症肺炎、肺好酸球増加閉塞性細気管支炎 器質性肺炎 (pulmonary eosinophilia bronchiolitis obliterans - organizing pneumonia)、びまん性肺出血症候群、グッドパスチャー症候群、特発性肺血鉄症、肺併発膠原血管病、肺胞たんぱく症、肺腫瘍、炎症性及び非炎症性胸水、気胸症、胸膜腫瘍、薬物性肺疾患、放射線性肺疾患及び肺移植の合併症などが含まれ、脂質異常の中には、脂肪肝、胆汁うっ滞、原発性胆汁性肝硬変、カルニチン欠乏症、カルニチンパルミトイルトランスフェラーゼ欠乏症、ミオアデニレートデミネラーゼ欠乏症、高トリグリセリド血症、ファブリー病などの脂質貯蔵病、ゴーシェ病、ニーマン ピック病、変染色性白質ジストロフィー、副腎性白質ジストロフィー、GM₂ にガングリオシドーシス、セロイドリポフスチン症、無リポ蛋白血症、タンジアー病、リポ蛋白過剰血症、糖尿病、脂肪異常栄養症、脂肪腫症、急性皮下脂肪組織炎、播種性脂肪組織壊死症、有痛脂肪症、リポイド副腎過形成、リポイドネフローゼ、脂肪腫、アテローム性動脈硬化症、高コレステロール血症、高トリグリセリド血症を伴った高コレステロール血症、原発性低リポ蛋白血症、甲状腺機能低下症、腎臓病、肝疾患、レシチン - コレステロールアシルトランスフェラーゼ欠乏症、脳腱黄色腫症、シトステロール血症、低コレステロール血症、テイ サックス病、サンドホフ病、高脂血症、脂肪過剰血症、脂質筋障害、肥満症が含まれる。PKINをコードするポリヌクレオチド配列は、サーザン法やノーザン法、ドットプロット法、或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PKINの発現を検出するため、患者から採取した体液或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。このような定性分析法または定量分析法は、当業者では周知される。

【0218】

ある実施態様では、PKIN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。PKIN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或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PKIN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、或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。

【0219】

P K I N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或いは増幅に好適な条件下で、動物或いはヒトの何れかの正常な被験者から抽出された体液或いは細胞と、P K I Nをコードする配列或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。

【0220】

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。

10

【0221】

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。

【0222】

P K I N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、P C R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、或いは*in vitro*で産出し得る。オリゴマーは、好ましくはP K I Nをコードするポリヌクレオチドの断片、或いはP K I N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D N A或いはR N A配列の検出、定量、或いはその両方のため用いることが可能である。

20

【0223】

或る実施態様において、P K I N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、一塩基多型性(S N P)を検出し得る。S N P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S N Pの検出方法には、S S C P (single-stranded conformation polymorphism)及び蛍光S S C P (f S S C P)法がある。S S C Pでは、P K I N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(P C R)でD N Aを増幅する。D N A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。D N A内のS N Pは、一本鎖形状のP C R生成物の2次及び3次構造に差異を生じさせる。

30

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。f S C C P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D N Aシーケ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S N P (*in silico* S N P, *is* S N P)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D N A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、D N Aの実験室での調整及び統計モデル及びD N A配列クロマトグラムの自動分析を用いたシーケ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M A S S A R R A Yシステム(Sequenom, Inc., San Diego CA)を用いた質量分析によりS N Pを検出し、特徴付ける。P K I N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

40

50

た結果の補間もある(例えば、Melby, P. C. 他(1993) J. Immunol. Methods 159:235244; Duplaa, C. 他(1993) Anal. Biochem. 212:229236. 参照)。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、本明細書に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ァイ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ァイルに基づき、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。

10

【0225】

別の実施例では、PKIN、PKINの断片、PKIN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ァイルをモニターまたは測定することが可能である。

20

【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 他、米国特許第5,840,484号の「Comparative Gene Transcript Analysis」を参照。この特許に言及することを以って本明細書の一部とする)。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ァイルを提供し得る。

30

【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 *in vivo*、または株化細胞の場合には *in vitro* での遺伝子発現を反映する。

40

【0228】

本発明のポリヌクレオチドの発現のプロファイ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、*in vitro* 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. 他、(1999) Mol. Carcinog. 24:153-159、Steiner, S. 及び N. L. Anderson (2000) Toxicol. Lett. 112-113:467-471、当該文献は特に引用することをも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

50

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の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未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ケンシングする。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、PKINに特異的な抗体を用いてPKIN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. Biochem. 270:103-111、Mendoza, 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。

10

【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ケンシングし、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。

20

【0235】

当分野で周知の方法でマイクロアレイを準備して使用し、分析する。(例えば、Brennan, T. M. 他 (1995) 米国特許第5,474,796号; Schena, M. 他 (1996) *Proc. Natl. Acad. Sci.* 93:10614 - 10619; Baldeschweiler 他 (1995) PCT出願番号WO95/251116; Shalon, D. 他 (1995) PCT出願番号WO95/35505; Heller, R. A. 他 (1997) *Proc. Natl. Acad. Sci.* 94:2150 - 2155; 及び Heller, M. J. 他 (1997) 米国特許第5,605,662号を参照)。様々なタイプのマイクロアレイが公知であり、詳細については、*DNA Microarrays: A Practical Approach*, M. Schena, 編集。(1999) Oxford University Press, Londonに記載されている。該文献は、特に引用することをもって本明細書の一部となす。

30

本発明の別の実施例ではまた、PKINをコードする核酸配列を用いて、天然のゲノム配列をマッピングするのに有用なハイブリダイゼーションプローブを作製す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。コード化された配列は、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にマッピングされる。(例えば、Harrington, J. J. et al. 他 (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

40

50

染色体領域の遺伝または制限酵素断片長多型 (RFLP) と関連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. USA 83: 7353-7357を参照)。

【0236】

蛍光原位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関連し得る(前出のHeinz-Ulrich, ら (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PKINをコードする遺伝子の位置と特定の疾患との相関性、或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。

10

【0237】

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或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等を参照)。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。

20

【0238】

本発明の別の実施例では、PKIN、その触媒作用断片或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る。PKIN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。

30

【0239】

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(例えば、Geysen, 他 (1984) PCT application WO84/03564参照)。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、PKIN、或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PKINが、当分野で周知の方法で検出される。精製されたPKIN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。

40

【0240】

別の実施例では、PKINと結合可能な中和抗体がPKIN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、PKIN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。

【0241】

別の実施例では、発展途上の分子生物学技術にPKIN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。

【0242】

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もって本発明を最大限に利用で

50

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。

【0243】

本明細書において開示した全ての特許、特許出願及び刊行物、特に米国特許第60/220,038号、第60/222,112号、第60/222,831号及び第60/224,729号は、言及することをもって本明細書の一部となす。

【0244】

(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) に含まれているcDNAライブラリに由来し、表4の列5に示されている。組織の一部を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する一方、この組織の別の一部をホモジナイズしてフェノールに溶解するか、或いはTRIZOL (Life Technologies)、グアニジニウムイソチオシアネート及びフェノールの単相溶液などの好適な変性剤の混合液に溶解した。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、或いは別の方法を用いて、溶解物からRNAを沈殿させた。

【0245】

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。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)を用いて当分野で周知の推奨方法または類似の方法でcDNAを合成してcDNAライブラリを作製した。(例えば、Ausubel, 1997, 前出, ユニット5.1-6.6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。殆どのライブラリでは、SEPHACRYL S1000またはSEPHAROSE CL2B、SEPHAROSE CL4Bカラムクロマトグラフィー(Amersham Pharmacia Biotech)、アガロースゲル電気泳動法によってcDNAの大きさ(300~1000bp)を選択した。PBLUESCRIPTプラスミド(Stratagene)またはpSPORT1プラスミド(Life Technologies)、pcDNA2.1プラスミド(Invitrogen Carlsbad CA)、PBK-CMVプラスミド(Stratagene)、pINCYプラスミド(Incyte Pharmaceuticals, Palo Alto CA)、またはそれらの誘導体などの好適なプラスミドのポリリンカーの適合性制限酵素部位にcDNAを結合させた。組換えプラスミドは、Stratagene社のXL1-Blue、XL1-BlueMRFまたはSOLR、或いはLife Technologies社のDH5、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。

【0246】

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、或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、凍結乾燥して或いは凍結乾燥せずに、4℃で保管した。

【0247】

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 Oy, Helsinki, Finland) を用いて蛍光分析的に定量した。

【0248】

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のシークエンシング反応は標準的な方法で行うか、またはHYDRAマイクロディスペンサー (Robbins Scientific) 或いはMICROLAB 2200 (Hamilton) 液体移送装置と共にABICATALYST 800 (PE Biosystems) サーマルサイ클ラー或いはPTC-200 thermal cycler (MJ Research) などの高処理装置を用いて行った。cDNAのシークエンシング反応は、Amersham Pharmacia Biotech社の試薬、またはABI PRISM BIGDYE Terminator cycle sequencing ready reactionキット (PE Biosystems) などのABIシークエンシングキットに含まれる試薬を用いて行った。cDNAシークエンシングの反応物の電気泳動的による分離及び標識したポリヌクレオチドの検出は、MEGABACE 1000 DNAシークエンシングシステム (Molecular Dynamics)、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 (PE Biosystems)、または当分野で周知のその他の配列解析システムを用いて行った。cDNA配列内の読み枠は、標準的な方法 (Ausubel, 1997, 前出, unit 7.7) 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。

【0249】

Incyteの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)配列の除去し、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次構造を分析する確率的アプローチである。例えば、Eddy, S.R. (1996) Curr. Opin. Struct. Biol. 6:361-365を参照)。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。或いは、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。或いは、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Prositate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。

10

【0250】

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も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）。

20

【0251】

完全長ポリヌクレオチド及びポリペプチド配列の構築及び分析に用いる上記のプログラムは、SEQ ID NO: 21 - 40からのポリヌクレオチド配列の断片を同定するためにも使用できる。ハイブリダイゼーション及び増幅技術に有用な約20～約4000のヌクレオチドの断片は、表4の列4に記載した。

30

【0252】

4 ゲノムDNAからのコード配列の同定及び編集

推定ヒトキナーゼは、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て同定された。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び構築された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、どの配列がヒトキナーゼ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モデルにおいてヒトキナーゼについて問合せて分析した。潜在的なヒトキナーゼが、ヒトキナーゼとしてアノテーションが付けられたIncyte 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

40

50

する。BLAST分析はまた、任意のIncycyte 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。Incycyte 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、Incycyte 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。或いは、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。

【0253】

5 cDNA配列データを使ったゲノム配列データの構築 ステッチ配列 (Stitched 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ク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エク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クソンは、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。

【0254】

「ストレッチ」配列

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。まず、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Incycyte cDNA配列または実施例4に記載のGenscanエク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。

【0255】

6 PKIN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: 21 - 40を組み立て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、Incycyteの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: 21 - 40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

10

20

30

40

50

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。

【0256】

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。ヒトゲノム地図及びその他の公衆に利用可能な資源、例えばNCBIの「GeneMap'99」ウェブサイト(<http://www.ncbi.nlm.nih.gov/genemap/>)を用いて、以前に同定した病変遺伝地図が上記の間隔内またはその近傍にあるかどうかを決定することができる。

【0257】

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或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。検索の基準はプロダクト積スコアであり、次式で定義される。

【数1】

(BLASTスコア×配列一致率)

$5 \times (\text{長さ(配列1)}, \text{長さ(配列2)})$ の最小値

【0258】

プロダクト積スコアは、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。プロダクト積スコアは、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プロダクト積スコアを計算する。プロダクト積スコアは、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。例えばプロダクト積スコア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。プロダクト積スコア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。プロダクト積スコア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。

【0259】

或いは、PKIN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。各ヒト組織は、心血管系、結合組織、消化系、胚構造、内分泌系、外分泌腺、女性生殖器、男性生殖器、生殖細胞、血液および免疫系、肝臓、筋骨格系、神経系、膵臓、呼吸器系、感覚器官、皮膚、顎口腔系、分類不能/混合、または尿管などの生物/組織のカテゴリの一つに分類される。各カテゴリのライブラリ数を数えて、全カテゴリ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/病状カテゴリ即ち癌、細胞系、発達、炎症、神経性、外傷、心血管、鬱血、その他の一つに分類される。各カテゴリのライブラリ数を数えて、全カテゴリの総ライブラリ数で除する。得られるパーセンテージは、PKIN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。

10

【0260】

8 PKIN コードされたポリヌクレオチドの伸張

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。開始プライマーは、OLIGO 4.06ソフトウェア (National Biosciences) 或いは他の適切なプログラムを用いて、約222個から約30個のヌクレオチドの長さで約50%以上のGC含量を有し、かつ約68~72の温度で標的配列にアニールするように設計した。ヘアピン構造及びプライマー・プライマー二量体が生じないようにヌクレオチドを伸長した。

20

【0261】

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、付加的プライマー或いはプライマーのネステッドセットを設計した。

【0262】

当分野で既知の方法を利用したPCR法で高い忠実度で増幅した。PCRはPTC-200 thermal cycler (MJ Research, Inc.)を用いて96ウェルブロックプレートで行った。反応混合液は、鋳型DNA及び200 nmolの各プライマー、Mg²⁺と(NH₄)₂SO₄と2-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ーで増幅を行った。

30

ステップ1	94	で3分間
ステップ2	94	で15秒間
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

40

【0263】

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ーで増幅を行った。

ステップ1	94	で3分間
ステップ2	94	で15秒間
ステップ3	57	で1分間

50

- 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】

各穴のDNA濃度は、透明蛍光計プレート(Corning Costar, Acton MA)に100 µl PICO GREEN 滴定試薬(0.25% (v/v) PICO GREEN、分子プローブ、1X TEに溶解するEugene OR) および希薄されないPCR製品0.5 µlを入れ、DNAが試薬に結合するのを待つ。サンプルの蛍光を計測してDNAの濃度を定量するようにプレートをFluoroskan II (Labsystems Oy, Helsinki, Finland)でスキャンする。反応混合物の5~10 µlのアリコットを1%のアガロースミニゲル上での電気泳動によって解析し、何れの反応物が配列を伸長することに成功したかを決定する。

【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、pUC18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。シヨットガン・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 Biolabs, Beverly MA)を用いてpUC18ベクター(Amersham Pharmacia Biotech)に再連結し、PfuDNA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、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PCRによってDNAを増幅した。その際用いたパラメータは次の通りである。

- ステップ1 94 で3分間
 ステップ2 94 で15秒間
 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

【0267】

DNAは、上記のPICO GREEN試薬(Molecular Probes)によって定量した。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。

【0268】

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、或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。

【0269】

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: 21 - 40 から得たハイブリダイゼーションプローブを利用して、cDNA、ゲノムDNAまたはmRNA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を、OLIGO4.06ソフトウェア(National Bioscience)のような最新式のソフトウェアを用いてデザインし、50 pmolの各オリゴマーと、250 μ Ciの[³²P]アデノシン三リン酸(Amersham, Chicago, IL)及びT4ポリヌクレオチドキナーゼ(DuPont NEN, Boston MA)とを組み合わせて用い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超細繊維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、Xba I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107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。

10

【0270】

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°Cで16時間かけて行う。非特異的シグナルを取り除くため、例えば、最大0.0x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の条件下で、プロットを順次室温にて洗浄する。オートラジオグラフィ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。

20

【0271】

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。或いは、ドットプロット法またはスロットプ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) Science 270: 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)。

30

40

【0272】

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(DNA STAR)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。或いは、レーザ脱離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

50

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。

【0273】

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、0.05 pg/ μ lのオリゴ(dT)プライマー(21mer)、1x第一鎖合成バッファー、0.03 unit/ μ lのRNAアーゼ阻害因子、500 μ MのdATP、500 μ MのdGTP、500 μ MのdTTP、40 μ MのdCTP、40 μ M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。特異的制御ポリ(A)+RNAは、37で2時間インキュベートした後、in vitro転写により非コード酵母ゲノムDNAから合成する。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、85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CHROMASPIN 30ゲル濾過スピナラム(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、14 μ lの5xSSC/0.2%SDS中で再懸濁する。

【0274】

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5 μ g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。

【0275】

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ガラス上に固定する。顕微鏡スライドガラス(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。スライドガラスは、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のオープンで硬化させる。

【0276】

米国特許第5,807,522号で説明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。該特許は、引用を以って本明細書の一部となす。平均濃度が100ng/ μ lのアレイエレメントDNA1 μ l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。

【0277】

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回洗浄し、

蒸留水で3回洗浄する。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。

【0278】

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1.8cm²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。このアレイを含むチャンバーを、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、第2洗浄緩衝液中(0.1×SSC)において45 で10分間それぞれ3回洗浄し、その後乾燥させる。

10

【0279】

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でスペクトル線を生じ得るInnova 70混合ガス10W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。

20

【0280】

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。

30

【0281】

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配列を含め、その位置におけるシグナルの強度がハイブリダイズする種の重量比1:100,000に相関するようにする。異なる試料(例えば検査細胞及びコントロール細胞を代表する)からの2つのサンプルを、各々異なる蛍光体で標識し、他と異なって発現する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させる場合には、較正は2つの蛍光体を有する較正するcDNAのサンプルを標識して、ハイブリダイゼーション混合液に各々等量を加えて行う。

40

【0282】

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-デジタル(A/D)変換ボード(Analog Devices, Inc., Norwood MA)を用いてデジタル化される。デ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

50

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（発光スペクトルの重なり起因する）光学的クロストークを補正する。

【0283】

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 (Incycyte) である。

【0284】

1.1 相補的ポリヌクレオチド

P K I N をコードする配列或いはその任意の一部に対して相補的な配列は、天然の P K I N 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、これより小さな或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。Oligo 4.06 ソフトウェア (National Biosciences) 及び P K I N 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 P K I N をコードする転写物に結合するのを阻害する。

【0285】

1.2 P K I N の発現

P K I N 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 P K I N が発現するために、抗生物質耐性及び c D N A 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 c D N A 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 - D チオガラクトピラノシド (IPTG) で誘発されると P K I N を発現する。真核細胞での P K I N 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の多角体遺伝子を、相同組換え或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、P K I N をコードする c D N A 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 c D N A 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. USA 91:3224-3227、Sandig, V. 他、(1996) Hum. Gene Ther. 7:1937-1945. 等を参照)。

【0286】

殆どの発現系では、P K I N が、例えばグルタチオン S トランスフェラーゼ (G S T)、または F L A G や 6 - H i s 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。G S T は日本住血吸虫からの 26 k D a 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、G S T 部分を特定の操作部位で P K I N からタンパク質分解的に切断できる。F L A G 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 F L A G 抗体 (Eastman Kodak) を用いて免疫親和性精製を可能にする。6 ヒス

10

20

30

40

50

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PKINを直接用いて以下の実施例16、17、18及び19のアッセイを行うことができる。

【0287】

1.3 機能的アッセイ

PKIN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PKIN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pCR3.1プラスミド(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。5~10µgの組換えベクターを、例えば内皮由来か造血由来のヒト細胞株にリポソーム製剤或いは電気穿孔法によって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或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、プ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nuclear DNA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。

【0288】

遺伝子発現におけるPKINの影響は、PKIN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が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。PKIN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。

【0289】

1.4 PKIN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; 例えば、Harrington, M. G., (1990) Methods Enzymol. 182: 488-495を参照)または他の精製技術で実質的に精製されたPKINを用いて、標準的なプロトコルでウサギを免疫化して抗体を作り出す。

別法では、PKINアミノ酸配列をLASERGENEソフトウェア(DNA STAR)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、或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)。

【0290】

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等を参照)。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PKIN活性を検査するには、ペプチドまたはPKIN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。

10

【0291】

15 特異的抗体を用いる天然PKINの精製

天然PKIN或いは組換えPKINを、PKIN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クラムは、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PKIN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。

【0292】

PKINを含む培養液をイムノアフィニティークラムに通し、PKIN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PKINとの結合を切るような条件で(例えば、pH2~3のバッファー、或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塩で)溶出させ、PKINを回収する。

20

【0293】

16 PKINと相互作用する分子の同定

PKINまたは生物学的に活性であるPKIN断片を、¹²⁵I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PKINと共にインキュベートし、洗浄して、標識したPKIN複合体を有する全てのウェルをアッセイする。様々なPKIN濃度で得られたデータを用いて、候補分子と結合したPKINの数量及び親和性、会合についての値を計算する。

30

【0294】

別法では、PKIN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。PKINはまた、ハイスループット型の酵母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)。

40

【0295】

17 PKIN 活性の実証

一般的に、プロテインキナーゼ活性は、³²P ATP環境下のPKINによるタンパク質基質リン酸化の定量分析で測定される。PKINはタンパク質基質、³²P-ATP、及び適切なキナーゼバッファーと共にインキュベートされる。基質に組み込まれた³²Pは、電気泳動法で遊離³²P-ATPより分離され、組み込まれた³²Pは同位性カウンタ

50

で計数する。組み込まれた³²Pの量は、PKINの活性に比例する。リン酸化された特異的アミノ酸残基の定量は、加水分解タンパク質のホスホアミド酸分析によってなされる。

【0296】

ある実施態様では、プロテインキナーゼ活性はガンマリリン酸塩がアデノシン三リン酸(ATP)からタンパク質基質中のセリン、トレオニン、またはチロシン残基に転移する量を定量化することによって測定される。反応はビオチン標識されたペプチド基質を有するプロテインキナーゼサンプルと³²P-ATPとの間で起こる。反応に続き、溶液中のアビジンがビオチン標識された³²Pペプチド生成物に結合させる目的で添加される。結合試料は、次に生成物アビジン複合体を保持し、遊離の³²P-ATPの透過させる膜を用いて、遠心限外濾過プロセスを行う。残余として³²Pペプチド生成物を含む遠心分離ユニットのリザーバーは、次にシンチレーションカウンタで計数される。この方法は、選択されたペプチド基質及びキナーゼ反応バッファによってあらゆるタイプのプロテインキナーゼのアッセイを可能とする。このアッセイは、キット形態(ASUA, Affinity Ultrafiltration Separation Assay, Transbio Corporation, Baltimore MD, 米国特許番号 5, 869, 275)で提供される。示唆された基質及びその各々の酵素は、次のヒストンH1(シグマ)及びp34^{cdc2}キナーゼ、アネキシンI、アンジオテンシン(シグマ)及びEGF受容体キナーゼ、アネキシンII及びsrcキナーゼ、ERK1及びERK2基質及びMEK、ミエリン塩基性蛋白質及びERKを含む(Pearson, J. D. 他、(1991) Methods Enzymol. 200: 62-81)が、これには限るものではない。

【0297】

別の実施態様では、PKINのタンパク質活性は、PKIN、50µlのキナーゼバッファ、ミエリン塩基性蛋白質(MBP)若しくは合成ペプチド基質のような1µgの基質、1mMのDTT、10µgのATP、及び0.5-5µCiの[³²P]ATPを含むアッセイに於いて*in vitro*で示される。反応液を30で30分間インキュベートし、ピペットでP81ペーパーに移して反応を停止させる。組み込まれていない[³²P]は、洗浄によって取り除かれ、組み込まれた放射活性は放射能シンチレーションカウンタを用いて測定される。別法では、SDSローディングバッファの存在下で100に加熱して反応を停止し、オートラジオグラフにより12%SDSポリアクリルアミドゲル上で視覚化する。組み込まれた³²Pの量は、PKINの活性に比例する。

【0298】

さらに別の実施例では、アデニレートキナーゼ若しくはグアニル酸キナーゼが、同位性カウンタを用いて、標識された³²P-ATPのADP若しくはGDPへの取り込みを測定でき得る。キナーゼバッファ中の酵素は、適切なヌクレオチドモノリン酸塩基質(AMP若しくはGMP)及びリン酸塩ドナーとしての³²P標識ATPと共にインキュベートされる。反応は37でインキュベートされ、トリクロロ酢酸の添加によって終了される。酸抽出物は、中和されゲル電気泳動にかけられて、一リン酸、二リン酸、三リン酸のヌクレオチド画分に分離する。二リン酸ヌクレオチド画分が取り除かれ、計数される。回復した放射性活性は酵素の活性に比例する。

【0299】

さらに別の実施例では、PKINのための別のアッセイは、シンチレーション近接アッセイ(SPA)、シンチレーションプレート技術、及びフィルタ結合アッセイを含む。適用できる基質はグルタチオン・トランスフェラーゼでタグされる組換えタンパク質またはビオチンでタグされる合成ペプチド基質を含む。微小有機分子、タンパク質またはペプチドなどPKIN活性のインヒビターはそのようなエッセイで同化される。

【0300】

18 プロテインキナーゼ活性の強化/抑制
PKIN活性化若しくは抑制のアゴニスト若しくはアンタゴニストは、実施例17で記述

したアッセイを用いてテストしても良い。アゴニストは、PKIN活性の増加を引き起こし、アンタゴニストはPKIN活性の減少を引き起こす。

【0301】

19 キナーゼ結合エッセイ

FLAG-CD44 cyt 融合タンパク質へのPKINの結合は、PKINを抗PKIN抗体を抱合したイムノアフィニティビーズとインキュベーションした後に、¹²⁵I 標識されたFLAG-CD44 cyt 融合タンパク質(5,000 cpm/ng タンパク質)の存在下で、0.5 mlの結合バッファー(20 mM Tris-HCl (pH 7.4)、150 mM NaCl、0.1%のウシ血清アルブミン、および0.05% Triton X-100)を一部のビーズ(10~20 ngのタンパク質が付いている)と4°Cで5時間インキュベートする。結合させた後、ビーズを結合バッファーでよく洗浄し、ビーズに結合した放射線をシンチレーションカウンターで測定する(Bourgignon, L.Y.W. 他(2001) J. Biol. Chem. 276:7327-7336)。組み込まれた³²Pの量は、PKINの活性に比例する。

【0302】

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。

【0303】

(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。

【0304】

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 相同志のアンノテーションを示す。各ポリペプチドとそのGenBank 相同志が一致する確率スコアも併せて示す。

【0305】

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。

【0306】

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。

【0307】

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。

【0308】

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。

【0309】

表7は、PKIN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】

10

20

30

40

表1

Incyte プロジェクト ID	ボリベブチド ID NO:	ボリベブチド SEQ	Incyte ボリベブチド ID	ボリスクレオチド 配列識別番号	Incyte リスクレオチド ID
2564295	1		2564295CD1	21	2564295CB1
2837050	2		2837050CD1	22	2837050CB1
7474590	3		7474590CD1	23	7474590CB1
7474594	4		7474594CD1	24	7474594CB1
7477585	5		7477585CD1	25	7477585CB1
7477587	6		7477587CD1	26	7477587CB1
7594537	7		7594537CD1	27	7594537CB1
7046749	8		7046749CD1	28	7046749CB1
7478559	9		7478559CD1	29	7478559CB1
698381	10		698381CD1	30	698381CB1
7474637	11		7474637CD1	31	7474637CB1
770260	12		770260CD1	32	770260CB1
1797506	13		1797506CD1	33	1797506CB1
1851973	14		1851973CD1	34	1851973CB1
7474604	15		7474604CD1	35	7474604CB1
7474721	16		7474721CD1	36	7474721CB1
7478815	17		7478815CD1	37	7478815CB1
7477141	18		7477141CD1	38	7477141CB1
2190612	19		2190612CD1	39	2190612CB1
7477549	20		7477549CD1	40	7477549CB1

10

20

30

40

【表 2】

表2-1

ポリペプチド	Incyte ポリペプチド ID	GenBank	種差スコア	GenBank 相同体
1	2564295CD1	g186555	0.0	インシュリン受容体・関係受容体[ヒト]
2	2837050CD1	g2853031	0.0	Tousled 様キナーゼ[ハツカネズミ]
3	7474590CD1	g6453611	5.1e-86	プロテインキナーゼ (突然変異型)[ハツカネズミ]
4	7474594CD1	g3879221	5.6e-99	Genefinder を用いた予測。カゼイン・キナーゼ I に類似、[線虫 (Caenorhabditis elegans)]。
5	7477585CD1	g348245	3.5e-62	プロテインキナーゼ[ヒト]
6	7477587CD1	g312998	7.4e-73	プロテインキナーゼ[ヒト]
7	7594537CD1	g485398	0.0	90kDa-ジアシルグリセロール・キナーゼ[トブネズミ]
8	7046749ICD1	g3089349	0.0	Cdc25C 関連プロテインキナーゼ c-TAK1 [ヒト]
9	7478559CD1	g7960111 g9998952	4.2e-114 1.00E-123	エタノールアミン・キナーゼ [ヒト] [ヒト] エタノールアミン・キナーゼ Lykidis, A. 他 Overexpression of a mammalian chanolamine-specific kinase accelerates the CDP-ethanolamine pathway J. Biol. Chem. 276, 2174-2179 (2001)
10	1698381CD1	g36615	7.7e-122	[ヒト] セリン/トレオニンプロテイン・キナーゼ Meyerson, M. 他 (1992) EMBO J. 11:2909-2917
11	7474637CD1	g1181079 g1401232	0.0 0	[ヒト] ジアシルグリセロール・キナーゼ δ Sakane, F. 他 (1996) J. Biol. Chem. 271:8394-8401 [キヌゲネズミ属] ジアシルグリセロール・キナーゼ ϵ Klauck, T.M. 他 糖鎖コルチコイド誘導ジアシルグリセロール・キナーゼのクローニングと特性決定 J. Biol. Chem. 271, 19781-19788 (1996)

10

20

30

40

表2-2

ポリペプチド	Incyte ポリペプチド ID	GenBank	確率スコア	GenBank 相同体
12	7170260CD1	g8101585	3.5e-126	[ハツカネズミ] 精製特異的セリンキナーゼ-3 Zuercher, G. 他 (2000) Mech. Dev. 93:175-177
13	1797506CD1	g3300094	4.5e-227	[ヒト] プロテインキナーゼ/エンドリボヌクレオ ーゼ Tirasophon, W. 他 (1998) Genes Dev. 12:1812-1824 [ヒト] プロテインキナーゼ/リボヌクレオ ーゼ IRE1 β Iwawaki, T. 他 ER ストレスの ER 膜貫通キナーゼ/リボヌク レアーゼ IRE1 による翻訳制御 Nat. Cell Biol. 3, 158-164 (2001) [分枝酵母] プロテインキナーゼ Samejima, I., 及び Yanagida, M. (1994) Mol. Cell Biol. 14:6361-71 [シロイヌナズナ] IRE 相同体; プロテインキナ ーゼ様タンパク質 Sato, S., Nakamura, Y., Kaneko, T., Kato, T. 他 Structural analysis of Arabidopsis thaliana chromosome 3. I. Sequence features of the regions of 4,504,864 bp covered by sixty P1 and TAC clones DNA Res. 7, 131-135 (2000)
14	1851973CD1	g1853976	1.3e-37	[ハツカネズミ] Raf プロテインキナーゼ関係の プロテインキナーゼ Therrien, M. 他 (1995) Cell 83:879-888
15	7474604CD1	g1171250	2.0e-218	[ヒト] 受容体プロテインチロシン・キナーゼ Fox, G.M. 他 (1995) Oncogene 10:897-905
16	7474721CD1	g551608	4.1e-290	[ヒト] ヘリキナーゼ I Ruzzo, A. 他 (1998) Biochem. J. 331(Pt 2):607-613
17	7478815CD1	g2873349	0.0	

10

20

30

40

表2-3

ポリペプチド	Incyte ポリペプチド ID	GenBank	融解スコア	GenBank 相合体
18	747714ICD1	97239696 911385416	1.6e-87 0	[ヒト]ミオシン軽鎖キナーゼ Garcia, J.G. 他 (1997) Am. J. Respir. Cell Mol. Biol. 16:489-494 Garcia, J.G.N. 他 (1996) Biochem. Biophys. Res. Commun. 1:1-1 [ハツカネズミ] 横紋筋-特異的セリン/トレオニン・プロテインキナーゼ Hsieh, C.M. 他 Striated Muscle Preferentially Expressed Genes alpha and beta Are Two Serine/Threonine Protein Kinases Derived from the Same Gene as the Aortic Preferentially Expressed Gene-1 J. Biol. Chem. 275 (47), 36966-36973 (2000)
19	2190612CD1	91836161	6.0e-257	[ラット]Ca2+/カルモジュリン依存性プロテインキナーゼ IV キナーゼ Okuno, S., Kitani, T. および Fujisawa, H. (1996) J. Biochem. 119:1176-1181
20	7477549CD1	95006445 92736151	3.6e-179 0	[ヒト]CDC42 結合プロテインキナーゼβ Moncrieff, C.L. 他 (1999) Genomics 57:297-300 [ドブネズミ] 筋線維性ジストロフィー・キナーゼ関係 Cdc42 結合キナーゼ Leung, T 他 Myotonic dystrophy kinase-related Cdc42-binding kinase acts as a Cdc42 effector in promoting cytoskeletal reorganization Mol. Cell Biol. 18, 130-140 (1998)

10

20

30

40

表2-4

ポリペプチド	Incyte ポリペプチド ID	GenBank ID NO	確率スコア	GenBank 相同体
20		92217968	1.40E-161	[ヒト] 筋緊張性ジストロフィ・プロテインキナーゼ Kedra, D. 他 The germinal center kinase gene and a novel CDC25-like gene are located in the vicinity of the PYGM gene on 11q13 Hum.Genet. 100, 611-619 (1997)

10

20

30

40

表3-1

SEQ ID NO:	Incyte プラドID	ポリペ	アミノ酸残基数	潜在的リジン酸化部位	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	分析手法及びデータベース
1	2564295CD1		1297	S151 S238 S271 S49 S564 S666 S741 S758 S827 S887 S900 S93 S962 T223 T348 T475 T486 T494 T581 T582 T629 T64 Y454 Y652 T1014 T1020 T1063 S1163 T1171 T1187 S1245 T1275 T1284 S1073 T1128 S1253 T1145	N311 N411 N47 N492 N528 N616 N634 N756 N885 N898 N949	シグネチャ配列、ドメインおよびモチーフ プロテインキナーゼドメイン DM00004 P14617 980-1238; S980-F1239 受容体 シグナル トランスフェラーゼ・チロシンプロテイン キナーゼ 膜貫通 糖タンパク質 ATP 結合リ酸化 PD006834: A603-R745, F760-I818 受容体 前駆体 シグナル トランスフェラーゼ・チロシンプロテイン キナーゼ 膜貫通 糖タンパク質 ATP 結合リ酸化 PD005347: Q466-P602 推定上 インスリン様ペプチド受容体 前駆体 EC 2.7.1.112 トランスフェラーゼ・チロシンプロテインキナーゼ 膜貫通 糖タンパク質 ATP 結合 リン酸化 シグナル PD146134: L344-E495, V773-G899, D513-C799, E825-R855 前駆体 シグナル インスリン様 受容体 トランスフェラーゼ・チロシンプロテイン キナーゼ 膜貫通 糖タンパク質 ATP 結合 PD004354: V330-G410 受容体チロシンキナーゼ BL00239: G464-P473, E1030-E1077, M1092-R1114, A1117-E1142, D1144-Y1193, N1198-I1242 受容体チロシンキナーゼ BL00240F: T1143-E1190 受容体チロシンキナーゼ BL00790H: S831-L856 チロシンキナーゼ触媒ドメイン PR00109: M1059-R1072, Y1105-V1123, L1154-L1164, S1173-G1195, C1217-F1239 プロテインキナーゼシグネチャおよびプロファイル プロテインキナーゼ tyf.prf: E1091-T1143 受容体 チロシンキナーゼ クラスII シグネチャ receptor_tyr_kin_li.prf: R1119-G1167 シグナルペプチド: M1-D25 膜貫通ドメイン: Y922-Y944, Furin様システインリッチ領域 G173-K329 受容体 Lドメイン N47-N170, G346-N472 真核生物プロテインキナーゼドメイン プロテインキナーゼ: I979-E1248 Protein_Kinase_Atp L985-K1013	BLAST-DOMO BLAST-PRODOM BLAST-PRODOM BLAST-PRODOM BLAST-PRODOM BLIMPS-BLOCKS BLIMPS-BLOCKS BLIMPS-PRINTS PROFILESKAN PROFILESKAN HMMEER HMMEER HMMEER-PFAM HMMEER-PFAM HMMEER-PFAM MOTIFS

【表7】

表3-2

SEQ ID NO.	Incyte 承バ ブチド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	シグナペプチド配列, ドメインおよびモチーフ	分析方法及びデ ータベース
2	2837050CD1	718	S165 S186 S194 S238 S246 S257 S275 S298 S44 S46 S509 S605 S632 T176 T269 T344 T403 T488 T558 T78 Y571 Y97	N340 N36 N548 N630 N713 N714	シグナペプチド配列, ドメインおよびモチーフ プロテインキナーゼドメイン DM00004 P34314 736- 1002; L409-D677 TOUSLED様キナーゼ PD102959; M2-E183 キナーゼタンパク質 TOUSLED様 PD013350; M237- D400, Q287-L409 TOUSLED様キナーゼ KIAA0137 タンパク質 PD035377; K184-T236 TOUSLED様キナーゼ 複製 精巢 転写 PD026280; A682-N718 チロシンキナーゼ触媒ドメイン PR00109; L490-K503, V608-N630 プロテインキナーゼシグナペプチドおよびプロフィール プロ テインキナーゼ tyr.prf: E512-S570 真核生物プロテインキナーゼドメイン プロテインキナー ゼ: Y408-L687 Protein_Kinase_Atp: L414-K437 Protein_Kinase_St: I534-L546	BLAST-DOMO BLAST-PRODOM BLAST-PRODOM BLAST-PRODOM BLAST-PRODOM BLIMPS-PRINTS
3	7474590CD1	497	S17 S286 S291 S3 S314 S356 S372 S375 S381 S382 S409 S440 S447 S5 S70 T225 T265 T427 T445 T461	N243	プロテインキナーゼ触媒ドメイン DM00004 P27448 58- 297; V30-T265 チロシンキナーゼ触媒ドメイン PR00109; Y136-V154, V202-S224, L244-A266 プロテインキナーゼシグナペプチドおよびプロフィール protein_kinase_tyr.prf: Q94-G174 真核生物プロテインキナーゼドメイン プロテインキナー ゼ: Y28-L275 Protein_Kinase_St: V142-V154	MOTIFS MOTIFS BLAST-DOMO BLIMPS-PRINTS PROFILESCAN HMMER-PFAM
4	7474594CD1	741	S397 S402 S471 S592 S641 S652 S656 S737 T237 T274 T292 T308 T388 T587	N119 N291	プロテインキナーゼドメイン DM00004 P48730 11- 265; K144-Y392 カゼインキナーゼに類似 PD115501; F332-D422, L130-T233 真核生物プロテインキナーゼドメイン プロテインキナー ゼ: W140-F374 Protein_Kinase_Atp: I146-K169 シグナル切断: M1-L19	MOTIFS BLAST-DOMO BLAST-PRODOM HMMER-PFAM MOTIFS SPSCAN

10

20

30

40

【表 8】

表3-3

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
5	7477585CD1	645	S251 S273 S277 S372 S414 S454 S47 S490 S522 S600 S64 S84 S97 T211 T302 T329 T340 T538 T547 Y368	N598 N71	プロテインキナーゼドメイン DM000004 P51957 8-251: L35-S277 チロシンキナーゼ触媒ドメイン PRO0109 : T108-Q121, Y148-L166, Y256-A278 プロテインキナーゼシグネチャおよびプロファイル protein_kinase_tyr.prf : Q134-S185 真核生物プロテインキナーゼドメイン プロテインキナーゼ : Y29-L287 Protein_Kinase_St: I154-L166 プロテインキナーゼドメイン DM000004 P53350 55-295: R99-A267, S253-L310 チロシンキナーゼ触媒ドメイン PRO0109 : Y212-L230 プロテインキナーゼシグネチャおよびプロファイル protein_kinase_tyr.prf : F198-G250 膜貫通ドメイン transmem_domain: L555-S575 真核生物プロテインキナーゼドメイン プロテインキナーゼ : Y97-A267, S268-F319 Protein_Kinase_Alp: I103-K126 Protein_Kinase_St: L218-L230	BLAST-DOMO BLIMPS-PRINTS PROFILES SCAN HMMER-PFAM MOTIFS BLAST-DOMO BLIMPS-PRINTS PROFILES SCAN HMMER HMMER-PFAM MOTIFS MOTIFS
6	7477587CD1	623	S32 S393 S439 S54 S61 S67 S80 S93 T195 T367 T454 T463 T584			
7	7594537CD1	797	S11 S136 S165 S208 S25 S294 S380 S670 S675 S684 S81 T2 T26 T274 T298 T312 T320 T388 T518 T62 T625 T689 T743 T762 T766 Y449	N546 N646 N793	ホルボール-エステルおよび DAG 結合ドメイン DM01331 P49621 326-792: V321-K789 キナーゼ ジアシルグリセロール ホルボール-エステル結合 トラスフェラーゼ ジグリセロイド DAG 多重遺伝子ファミリー DKG PD002939: I575-F755 有望なジアシルグリセロールキナーゼ EC 2.7.1.107 ジグリセリド DKG DAG 仮説上のプロテイントランスフェラーゼ カルシウム結合 ホルボール-エステル結合 PD078865: A118-G236, L10-D85, T50-S81 キナーゼ ジアシルグリセロール ホルボール-エステル結合 プロテイントラスフェラーゼ ジグリセロイド DAG 多重遺伝子ファミリー PD002780: P431-W555	BLAST-DOMO BLAST-PRODOM BLAST-PRODOM BLAST-PRODOM BLAST-PRODOM

【表9】

10

20

30

40

表3-4

SEQ ID NO:	Incyte ポリヌチド ID	アミノ酸残基数	潜在的リジン酸化部位	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
8	7046749ICD1	749	S141 S2 S24 S346 S374 S417 S424 S444 S456 S457 S461 S49 S494 S495 S516 S634 S653 S659 S664 S730 T118 T283 T302 T33 T36 T508 T512 T519 T535 T614 T618 T623 T82 T9 Y113	N386 N399 N400 N479 N533 N637	ジアシルグリセロールキナーゼ、 β EC 2.7.1.107 ジグリセリド キナーゼ DGK DAG 90KD トランスフェラーゼ カルシウム結合 ホルボールエステル結合 多重遺伝子ファミリー PD119174; D352-H430 ジアシルグリセロールキナーゼ 触媒ドメイン PF00781; H331-Q336 P431-Y462 R483-L497 P509-Y532 K539-V539 N577-Y613 L655-G668 L747-Q758 ジアシルグリセロールキナーゼ 触媒ドメイン DAGKC; P431-W555 ジアシルグリセロールキナーゼ アクセサリードメイン DAGKa; I575-P755 ホルボールエステル/ジアシルグリセロール結合ドメイン d DAG PF-gIind; H238-C287, H303-C351 EF hand eIhand; K146-M174, I191-T219 ホルボールエステル/ジアシルグリセロール結合ドメイン BL00479; Q264-C279, L514-L526, H238-G260 ホルボールエステル/ジアシルグリセロール結合ドメイン dag_pe_binding_domain.prf; Y250-G378 Dag_Pe_Binding_Domain; H238-C287 Ef_Hand; D155-L167, D200-W212 プロテインキナーゼドメイン DM000004 P27448 58-297; L62-L303 キナーゼ セリン/トレオニン/プロテイン プロテイントランスフェラーゼ ATP 結合 セリン/トレオニン 推定上 KINI EMK PARI PD004300; G633-L749 キナーゼ セリン/トレオニン/プロテイン セリン/トレオニン 推定上トランスフェラーゼ ATP 結合タンパク質 EMK P78 CDC25C PD008571; S413-E632 キナーゼ セリン/トレオニン/プロテイン 推定上 セリン/トレオニン トランスフェラーゼ ATP 結合タンパク質 PARI KP78 EMK PD005838; I312-R412 セリン/トレオニン キナーゼ PD119193; S551-P622	BLAST-PRODOM BLIMPS-PFAM HMMER-PFAM HMMER-PFAM HMMER-PFAM BLIMPS-BLOCKS PROFILES SCAN MOTIFS MOTIFS BLAST-DOMO BLAST-PRODOM BLAST-PRODOM BLAST-PRODOM BLAST-PRODOM BLAST-PRODOM

10

20

30

40

【表 10】

表3-5

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
9	7478559CD1	386	S237 S259 S355 S38 S380 T20 T322 T85 T93 Y271	N168	チロシンキナーゼ触媒ドメイン PRO0109: Y173-L191, V239-Q261 プロテインキナーゼシグネチャおよびプロファイル protein_kinase_tyr-pyf: K122-G212 真核生物プロテインキナーゼドメイン プロテインキナーゼ: Y60-E85 真核生物プロテインキナーゼドメイン プロテインキナーゼ: F137-I312 Protein Kinase St: I179-L191 d0 コリン; キナーゼ; YDR147W; B0285.10; DM01931 P35790 128-455; D256-K376, F131-P300 d0 コリン; キナーゼ; YDR147W; B0285.10; DM01931 P46560 1-305; E125-A289 キナーゼ コリン トランスフェラーゼ タンパク質 多重遺伝子ファミリー 推定上類似 染色体 III PD003547: V222-L382, V109-E240 キナーゼ トランスフェラーゼ コリン PD02952: V243-I256, H263-N292 コリン/エタノールアミン キナーゼ Choline_kinase: T85-T356	BLIMPS-PRINTS PROFILES SCAN HMMER-PFAM HMMER-PFAM MOTIFS BLAST-DOMO BLAST-DOMO BLAST-PRODOM BLIMPS-PRODOM HMMER-PFAM HMMER-PFAM
10	1698381CD1	342	S180 S205 S238 S284 S288 S38 T247 Y15 Y211	N23	真核生物 プロテインキナーゼドメイン pkinase:Y4-F286, プロテインキナーゼ シグネチャおよびプロファイル protein_kinase_tyrosine: E90-G154 プロテインキナーゼドメイン DM00004 Q00532 7-278; K6-C277 プロテインキナーゼドメイン DM00004 Q00526 6-286; K6-F286 プロテインキナーゼドメイン DM00004 P23437 6-286; K6-G218 プロテインキナーゼドメイン DM00004 P51958 6-277; K6-G218 キナーゼ トランスフェラーゼ タンパク質 セリン/トレオニン タンパク質 ATP 結合 II リン酸化 カゼイン アルファ鎖 PD002608: V161-F286	BLAST-DOMO BLAST-DOMO BLAST-DOMO BLAST-DOMO BLAST-DOMO BLAST-PRODOM BLAST-PRODOM

10

20

30

40

表3-6

SEQ NO.	Incyte 配列バッチ ID	アミノ酸残基数	潜在的リジン酸化部位	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
11	7474637C01	1164	S114 S119 S152 S258 S39 S399 S41 S432 S450 S511 S56 S586 S587 S591 S608 S623 S654 S66 S664 S695 S766 S820 S873 S958 S967 S1075 T316 T419 T486 T514 T518 T659 T678 T863 T908 T955 T1046 T1118	N124 N314 N651 N1059 N1122	子ロシキンキナーゼ 触媒ドメイン シグネチャ PROJ0109: F116-T134 セリン/トレオニン/チロシンキナーゼ活性部位 シグネチャ C122-H134 ホルボールエステル/ジアシルグリセロール結合ドメイン: F188-A259 SEIGNAL_Cleavage/MI-A32 ホルボールエステル/ジアシルグリセロール結合ドメイン (C1 ドメイン): H176-C225, H248-C298 PH ドメイン: S66-T158 DAG キナーゼ触媒ドメイン: P332-W457 DAG キナーゼ触媒ドメイン: V770-A927 ホルボールエステルおよび DAG 結合ドメイン DM01331 P49621 326-792: P332-H505, V770-E865, R869-I946, C279-L313, G198-C225 ホルボールエステルおよび DAG 結合ドメイン DM01331 Q09103 683-1148: V330-H459, T752-R988, C279-P310, A2-I61 ホルボールエステルおよび DAG 結合ドメイン DM01331 P23743 308-734: P332-L500, V770-F869, P872-L946 ホルボールエステルおよび DAG 結合ドメイン DM01331 S9282 52-782: C279-H505, V770-L946 キナーゼ ジアシルグリセロール結合 リヒート 多重糖伝子 PD040467: S458-C769 ジアシルグリセロールホルボールエステル結合 キナーゼ E1A ジグリセロイド DAG トランスフェラーゼ E1A ジグリセロイド DAG トランスフェラーゼ E1A ジグリセロイド DAG トランスフェラーゼ E1A ジグリセロイド DAG 極微糖伝子ファミリ DKK PD002939: V770-E926 キナーゼ ジアシルグリセロールホルボールエステル結合プロテイントランスフェラーゼ ジグリセロイド DAG 多重糖伝子ファミリ PD002780: V330-W457	BLIIPS_PRINTS MOTIFS PROFILESCAN SPSCAN HMMER_PFAM HMMER_PFAM HMMER_PFAM HMMER_PFAM BLAST_DOMO BLAST_DOMO BLAST_DOMO BLAST_PRODOM BLAST_PRODOM BLAST_PRODOM BLAST_PRODOM

10

20

30

40

【表 1 2】

表3-7

SEQ ID No.	Incyte ポリプ チドID	アミノ酸残 基数	潜在的リン酸化部位 残基	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
11					ホルボールエステル/ジアシルグリセロール結合ドメイン シグネチャ配列、ドメインおよびモチーフ ホルボールエステル/ジアシルグリセロール結合ドメイン シグネチャ配列、ドメインおよびモチーフ PF00781 : H176-G198, H202-C217, L415-L427 シグネチャ配列、ドメインおよびモチーフ PF00781 : K278-K283, P332-F363, R384-L398, C410-Y433, O441-T461, N772-Y808, L848-G861, V919-O930 シグネチャ配列、ドメインおよびモチーフ PR00008: H202-A213, I1214-K226 キナーゼタンパク質ドメイン PD00584: K74-R84, L386-G395, L466-L473 ホルボールエステル/ジアシルグリセロール結合ドメイン: H176-C225 真核生物プロテインキナーゼドメイン プロテインキナーゼ: Y10-L265 プロテインキナーゼ シグネチャおよびプロファイル protein_kinase_tyrosine : G82-H162 プロテインキナーゼドメイン DM00004 P27446 58-297: K14-I256 プロテインキナーゼドメイン DM00004 I48609 55-294: K14- S255 プロテインキナーゼドメイン DM00004 O05512 55-294: K14- S255 プロテインキナーゼドメイン DM00004 J14446 20-261: Q11- I256 チロシンキナーゼ 触媒ドメイン シグネチャ PR00109: Y124-L142 プロテインキナーゼ ATP 結合領域シグネチャ: I16-K39 E870 プロテインキナーゼシグネチャおよびプロファイル protein_kinase_tyr_drf : E652-G709 プロテインキナーゼドメイン DM00004 Q09499 536-784: P561-A811 プロテインキナーゼドメイン DM00004 P3236 676-970: V564-O732, I740-A811 キナーゼ; トレオニン; ATP; セリン; DM0630 Q09499 786-924: V814-Y949	BLIMPS_BLOCKS BLIMPS_PFAM BLIMPS_PRINTS BLIMPS_PRODOM MOTIFS HMMER_PFAM PROFILES SCAN BLAST_DOMO BLAST_DOMO BLAST_DOMO BLIMPS_PRINTS MOTIFS HMMER_PFAM PROFILES SCAN BLAST_DOMO BLAST_DOMO BLAST_DOMO
12	7170260CD1	268	S161 S188 S255 S29 T15 Y124 Y21			
13	1797506CD1	965	S234 S326 S527 S530 S607 S636 S741 S841 S879 S884 S92 T111 T143 T155 T174 T202 T215 T229 T29 T372 T619 T685 T82 T922 T932 T963 Y173	N227		

10

20

30

40

【表 1 3】

表3-8

SEQ ID No.	Incyte ポリバ ブチド 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
14	1851973CD1	329	S264 S270 S293 S31 S311 S320 S7	N73	キナーゼ；トレオニン；ATP；セリン； DM06305 P32361 972-1114; Q813-L946 プロテインキナーゼ/エンドリンボスクレアーゼ 推定上セ リン/トレオニン プロテインキナーゼ C41C4.4 染色体 II 前駆体 トランスフェラーゼ PD152704; T197-L422, L88-E190 セリン/トレオニン プロテインキナーゼ 前駆体 膜貫通 シグナル トランスフェラーゼ ATP 結合 タンパク質 IRE1 糖タンパク質 PD032590; W821-Y949 チロシンキナーゼ 触媒ドメイン シグネチャ PR00109; H666-I684, G721-L731, V743-D765 セリン/トレオニンプロテインキナーゼ活性部位シグネチ ャ： I672-I684 ホスホリラーゼキナーゼ ファミリ シグネチャ PR01049; P812-R823 真核生物プロテインキナーゼドメイン プロテインキナー ゼ： F35-V180 プロテインキナーゼ シグネチャおよびプロファイル protein_kinase_tyrosine : M132-R184 プロテインキナーゼドメイン DM00004 P43565 796-1240; I37-R184 プロテインキナーゼドメイン DM00004 A56155 714-1002; V38-L177 プロテインキナーゼドメイン DM00004 P38679 238-527; V38-S178 プロテインキナーゼドメイン DM00004 P53894 353-658; V38-S178 チロシンキナーゼ 触媒ドメイン シグネチャ PR00109; M110-H123, V146-I164 セリン/トレオニンプロテインキナーゼ活性部位シグネチ ャ： I152-I164	BLAST_DOMO BLAST_PRODUM BLAST_PRODUM BLIMPS_PRINTS MOTIFS BLIMPS_PRINTS HMMER_PFIAM PROFLESCAN BLAST_DOMO BLAST_DOMO BLAST_DOMO BLAST_DOMO BLIMPS_PRINTS MOTIFS

10

20

30

40

【表 1 4】

表3-9

SRQ ID	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
15	7474604CD1	945	S110 S157 S193 S246 S289 S290 S31 S329 S356 S359 S405 S411 S611 S623 S636 S645 S67 S934 T170 T2 T217 T322 T42 T47 T496 T712 T839	N140 N155 N382 N631 N756 N888	真核生物プロテインキナーゼドメイン プロテインキナーゼ : L661-N920 プロテインキナーゼ シグネチャおよびプロフィール protein_kinase_tyrosine : K757-L801 プロテインキナーゼドメイン DM000004 P27966 85-332: I663-F916 プロテインキナーゼドメイン DM000004 P15056 458-705: I663-F916, プロテインキナーゼドメイン DM000004 P10398 312-559: I667-F916 プロテインキナーゼドメイン DM000004 B26126 305-552: I667-F916 キナーゼ RAS1 の抑制因子 KSR1 IIB タンパク質 PD103125 : V390-P557, K501-L661 L222-P323 キナーゼ RAS の抑制因子 KSR ホルボールエステル結合 RASI KSR1 IIB PD017776 : L21-E344 S485-T519 チロシンキナーゼ 触媒ドメイン シグネチャ PR00109 : Y771-Y789, W867-I877, M894-F916 セリン/トレオニンプロテインキナーゼ活性部位シグネチャ : I777-Y789	HMMER_PPFAM PROFILES SCAN BLAST_DOMO BLAST_DOMO BLAST_DOMO BLAST_PRODOM BLAST_PRODOM BLIMPS_PRINTS MOTIFS
16	7474721CD1	1009	S184 S203 S244 S293 S325 S44 S473 S62 S625 S682 S686 S805 S825 S851 S980 T108 T121 T133 T162 T214 T224 T232 T32 T423 T488 T551 T616 T619 Y504 Y766 Y801	N311 N486	真核生物プロテインキナーゼドメイン プロテインキナーゼ : V645-H897 Ephrin 受容体 リガンド結合ドメイン EPH_1hd: E35-C211 プロテインキナーゼ シグネチャおよびプロフィール protein_kinase_tyrosine : Q746-A799 受容体チロシンキナーゼ クラスV DM000501 S3174 33-382: V36-G394 受容体チロシンキナーゼ クラスV DM000501 P54759 33-382: V36-G394 受容体チロシンキナーゼ クラスV DM000501 4861 34-382: I37-G394 受容体チロシンキナーゼ クラスV DM000501 48612 34-382: I37-G394 キナーゼ 受容体 前駆体 チロシン タンパク質 EPHIN トランスフェラーゼ ATP 結合 リン酸化 膜貫通 糖タンパク質 PD001495 : E35-C211	HMMER_PPFAM HMMER_PPFAM PROFILES SCAN BLAST_DOMO BLAST_DOMO BLAST_DOMO BLAST_PRODOM

10

20

30

40

【表 15】

表3-10

SEQ ID NO:	Incyte ポリヌクレオチド 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
16	7477742ICD1	1009			キナーゼ 受容体 前駆体 チロシン タンパク質 Ephrin トランスフェラーゼ ATP 結合 リン酸化 膜貫通 糖タンパク質 PD149648 : A213-A284 EPH フォスホリタータンパク質 PD002683 : P339-T451 キナーゼ 受容体 前駆体 チロシン タンパク質 Ephrin トランスフェラーゼ ATP 結合 リン酸化 膜貫通 シグナル PD001551 : C285-R336 受容体チロシンキナーゼ BL00239 : E694-Q741, I747-R769, A772-S797, E798-Y847, G852-I896 受容体チロシンキナーゼ BL00790 : L751-A772, S805-W837, E838-G862, F863-K911, A955-R998, E835-N56, D65-P116, K172-A225, P252-Q276, C282-P329, R351-L377, C390-S433 シグナルドメイン : MI-A33 膜貫通ドメイン : V568-W589	BLAST_PRODOM BLAST_PRODOM BLAST_PRODOM BLIMPS_BLOCKS BLIMPS_BLOCKS
17	7478815CD1	917	S243 S364 S379 S449 S503 S547 S551 S772 S787 S791 S810 S826 S896 T114 T161 T172 T275 T35 T508 T523 T569 T625 T722 T726 T811 T877 Y27 Y497	N122 N208 N655	シグナルドメイン : MI-A33 シグナル Cleavage: MI-A33 ヘキソキナーゼ ヘキソキナーゼ ; E16-V463 Q464-L910 ヘキソキナーゼシグネチャ ヘキソキナーゼ ; I577-R642, V130-R195 ヘキソキナーゼ DM00597P27881465-915: Q466-A913, D17-Q464 ヘキソキナーゼ DM00597P52789465-915: Q466-A913, D17-Q464 ヘキソキナーゼ DM00597S48809465-915: Q466-A913, D17-Q464 ヘキソキナーゼ DM00597P27595465-915: Q466-Q911, D17-Q466 ヘキソキナーゼ トランスフェラーゼ キナーゼ 解糖 ATP 結合タイプ アロステリック酵素 HK 複製 PD001109 : Q466-D886, E699-A907, E16-D439, D251-R462 ヘキソキナーゼタンパク質 BL00378: V22-K49, V509-I545, V207-G250, M255-D266, Y724-G769, S892-V906 ヘキソキナーゼファミリー シグネチャ PRO0475: L529-I545, L597-F622, I650-Y666, V226-E240, Q291-M313, V818-I840, M890-V906 ヘキソキナーゼ L597-F622	HMMER HMMER SPSCAN HMMER PFAM PROFLIESCAN BLAST_DOMO BLAST_DOMO BLAST_DOMO BLAST_DOMO BLAST_PRODOM BLIMPS_BLOCKS BLIMPS_PRINTS MOTIFS

10

20

30

40

【表 16】

表3-11

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
18	7477141CD1	2380	S143 S166 S241 S277 S278 S285 S299 S343 S480 S537 S553 S568 S602 S711 S736 S996 S1033 S1035 S1037 S1062 S1127 S1523 S1571 S1245 S1435 S1468 S1506 S1586 S1609 S1679 S1691 S1747 S1117 S1527 S1557 S1578 S1594 S1613 S1736 S1747 S1876 S1947 S2137 S2171 S2253 S2321 S2058 S2062 S2165 S2269 S680 S754 S986 T108 T153 T158 T170 T350 T408 T476 T498 T578 T614 T692 T803 T862 T957 T1068 T1082 T1311 T1493 T1602 T1981 T2080 T1301 T1856 T1901 T2069 T2101 T2144 T2348 T1608 T2343 Y632 Y772 Y822	N37 N1675 N1847 N1874 N2099 N2299	真核生物プロテインキナーゼドメイン プロテインキナーゼ : Y714-F967, Y2079-I2331 プロテインキナーゼドメイン DM000004 S0757 5152-5396; D715-D952, E2083-I2322 プロテインキナーゼドメイン DM000004 P53355 15-257; O718-D952, E2083-I2322 プロテインキナーゼドメイン DM000004 JN0588 727-969; I716-D952, L2082-I2312 プロテインキナーゼドメイン DM000004 P07313 298-541; O718-R953, G2088-S2321 チロシンキナーゼ 触媒ドメイン シグネチャ PR00109 : Y822-V840 シグナルペプチド : M52-A70 真核生物プロテインキナーゼドメイン プロテインキナーゼ : Y2079-I2331 プロテインキナーゼ ATP 結合領域シグネチャ : I720-K743 セリン/トレオニンプロテインキナーゼ活性部位シグネチャ : V828-V840, V2194-I2206	HMMER_PPFAM BLAST_DOMO BLAST_DOMO BLAST_DOMO BLAST_DOMO BLIMPS_PRINTS HMMER HMMER_PPFAM MOTIFS MOTIFS
19	2190612CD1	505	S100 S117 S160 S330 S419 S425 S437 S458 S69 S74 S82 T108 T26 T430 T58	N147	真核生物プロテインキナーゼドメイン : Y128-V409 プロテインキナーゼ シグネチャおよびプロファイル protein_kinase_tyrosine : Q251-N303 プロテインキナーゼドメイン DM000004 A57156 130-399; L130-V400 プロテインキナーゼドメイン DM000004 P50526 136-399; E133-I398 プロテインキナーゼドメイン DM000004 P38990 135-438; E133-I320, N303-V400 プロテインキナーゼドメイン DM000004 P43637 52-334; I134-I378	HMMER_PPFAM PROFLESCAN BLAST_DOMO BLAST_DOMO BLAST_DOMO BLAST_DOMO

10

20

30

40

【表 17】

表3-12

SEQ ID NO:	Incyte プラト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
20	7477549CD1	1572	S161 S280 S307 S363 S407 S430 S471 S545 S625 S629 S646 S675 S711 S730 S737 S807 S811 S815 S841 S1058 S1294 S1162 S1500 S1405 S1414 S1556 T455 T590 T673 T888 T956 T1088 T1378		シグネチャ配列、ドメインおよびモチーフ キナーゼタンパク質 ベータ Ca2+/カルモジュリン依存性 Ca2+/カルモジュリン依存性 CAM キナーゼ IV アイソフォーム Δ ホスホリラーゼ B PD031900: M1-O127 キナーゼタンパク質 ベータ Ca2+/カルモジュリン依存性 Ca2+/カルモジュリン依存性 CAM キナーゼ IV アイソフォーム Δ ホスホリラーゼ B PD019141: V409-F463 キナーゼタンパク質 Ca2+/カルモジュリン依存性 IV アイソフォーム Δ ホスホリラーゼ B グリコーゲン合成 A PD027014: E464-S505 チロシンキナーゼ触媒ドメインシグネチャ PR00109: Y265-L283, G312-I322 ATP/GTP 結合部位モチーフ A(P-loop) G485-S492 プロテインキナーゼATP 結合領域シグネチャ: I134-K157 セリン/トレオニンプロテインキナーゼ活性部位シグネチャ: I271-L283 ホルボールエステル ジアシルグリセロール結合ドメイン: C900-S963 真核生物プロテインキナーゼドメイン プロテインキナーゼ: F71-F337 プロテインキナーゼドメイン DM00004 Q09013 83-336: I73-R325 プロテインキナーゼドメイン DM00004 S4286 75-498: I73-H252, V232-Y398 プロテインキナーゼドメイン DM00004 L38133 90-369: E72-L220, V232-G324 プロテインキナーゼドメイン DM00004 P53894 353-658: L74-G215, V232-R325 ホルボールエステル 結合 キナーゼ ジストロフィー キナーゼ関連 GM42 結合 類似 セリン/トレオニン タンパク質 ジンギスカン PDI50840: W1355-G1462 ホルボールエステル 結合 キナーゼ ジストロフィー キナーゼ関連 GM42 結合 類似 セリン/トレオニン タンパク質 ジンギスカン PDI51400: T1039-R1140	BLAST_PRODOM BLAST_PRODOM BLAST_PRODOM BLIMPS_PRINTS MOTIFS MOTIFS MOTIFS PROFILES SCAN HMMER_Pfam BLAST_DOMO BLAST_DOMO BLAST_DOMO BLAST_DOMO BLAST_PRODOM BLAST_PRODOM

10

20

30

40

表3-13

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
					キナーゼ RH9 関連 コイルドコイルタンパク質形成 ホルモンレスステル結合 ジストロフィー キナーゼ 関連 CDC42 結合 P0006715: T944-V1038, H433-L456 ホルボールエステル結合 ジストロフィー キナーゼ 関連 CDC42 結合 キナーゼ ジンギスカン 筋緊張性 筋緊張性 PD011252:S694-S815 チロシンキナーゼ 触媒ドメイン シグネチャ PR00109: C257-E279, M148-S161, S185-L203 ホルボールエステル/ジアシルグリセロール結合 dom DAG PE-bind: H887-C935 ホルボールエステル/ジアシルグリセロール結合ドメイン: H887-C935. プロテインキナーゼ ATP 結合領域 シグネチャ I77-K100 セリン/トレオニン/プロテインキナーゼ活性部位シグネチャ: Y191-L203 CNH ドメイン: L1100-K1380 プロテインキナーゼ C 末端 ドメイン: P351-D366 PH ドメイン PH: T936-R1074 signal_cleavage: M1-S37	BLAST_PRODOM BLAST_PRODOM BLIMPS_PRINTS HMMER_PPFAM MOTIFS MOTIFS MOTIFS HMMER_PPFAM HMMER_PPFAM HMMER_PPFAM SPSCAN

10

20

30

40

表4-1

ポリスクレオチド SEQ ID NO.:	Incyte ポリスクレオチド ID	配列長	選択された断片	配列断片	5'位置	3'位置
21	2564295GB1	4298	701-1736, 3536-3629, 1- 356, 2349- 2589, 3956- 4298, 2841- 3428	FL2564295_g7160581_000014_g 387060_1_15-16 FL2564295_g7160581_000014_g 387060_1_16-17 FL2564295_g7160581_000014_g 387060_1_7-8 55078393J1 FL2564295_g7160581_000014_g 387060_1_8-9 55078386J1 FL2564295_g7160581_000014_g 387060_1_18-19 2564295H1 (ADRETTUT01) g186554_CD FL2564295_g7160581_000014_g 387060_1_9-10 3599581H1 (DRCTNOT01) FL2564295_g7160581_000014_g 387060_1_1-2 FL2564295_g7160581_000014_g 387060_1_10-11 FL2564295_g7160581_000014_g 387060_1_20-21 FL2564295_g7160581_000014_g 387060_1_11-12 FL2564295_g7160581_000014_g 387060_1_2-3 FL2564295_g7160581_000014_g 387060_1_12-13 FL2564295_g7160581_000014_g 387060_1_3-4 FL2564295_g7160581_000014_g 387060_1_4-5	3200 3253 1938 37 2167 1 3594 4048 442 2335 3453 441 2531 3884 2573 994 2794 1298 1441	3482 3593 2334 717 2530 709 3883 4298 4250 2572 3756 1297 2793 4250 2930 1440 3093 1585 1800

10

20

30

40

【表 20】

表4-2

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
21				FL2564295_g7160581_000014_g 387060_1_14-15 FL2564295_g7160581_000014_g 387060_1_19-20	3094 3754	3252 4018
22	2837050CBI	2863	1-430, 2346- 2863	6854541H1 (BRAIFEN08) 91164223 71191190V1 7728560H1 (UTRCDIE01) 79 71972220V1 2227 71972389V1 2180 6881340H1 (BRAHTR03) 1555 7401101H1 (SINIDME01) 1293 CBI_g8103343_000001.eddit 1 FL7474590_g7630344_000002_g 6779549_1_1 1116	782 1 1439 79 2227 2180 1555 598 1 1 1512 858 340 1 1281 1180 1988 612 1305 1900 1241 1 1384 2019 2552 1849 679 1 1662	1467 496 2085 681 2863 2857 2209 1293 1494 1116 2341 1544 959 426 1779 1590 2534 1305 1900 686 1930 2552 2517 1390 1980 2164
23	7474590CBI	1494	1-1494			
24	7474594CBI	2341	682-792, 1- 262, 1522- 2341, 1254- 1373, 339-361	5503665J1 6949237H1 (BRAITDR02) 858 8016740J1 (BMARTXE01) 340 GNN_g8247875_000031_002 1 7278940H1 (BMARTXE01) 1281 GNN_g6689704_000006_002 1180 71975408V1 1988 55030002H1 612 55030074J1 1241 1406660F6 (LATRTU02) 1 6329987H1 (BRANDIN01) 1384 71987367V1 2019 6704049H1 (DRGCNT02) 1849 55030089H1 679 98671962.eddit 1 5823464F7 (PROSTUS23) 1662	2341 1512 858 340 1 1281 1180 1988 612 1305 1900 1241 1 1384 2019 2552 1849 679 1 1662	2341 1544 959 426 1779 1590 2534 1305 1900 686 1930 2552 2517 1390 1980 2164
25	7477585CBI	2552	1-465, 1075- 1150			
26	7477587CBI	2176	1276-1873, 1- 286			

10

20

30

40

表4-3

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
27	7594537CB1	4277	2383-2614, 1- 511-1170, 1- 518, 2763- 2834, 1714- 1859, 3119- 4277	7594537H1 (LIVRNC07) 7328693H1 (UFRCDIE01)	130 1	766 351
28	70467491CB1	2616	1717-2616, 1- 425	2395018F6 (THPIA2T01) FL70467491_g7708222_g759580 0	2015 1	2520 2250
29	7478559CB1	1253	1215-1253, 1- 53	93770955 7661715J1 (OVARNOE02) 95769093	1 655 314	321 1253 804
30	1698381CB1	1790	1-146, 892- 1313, 1659- 1790, 186-237	1698381F6 (BLADTUT05) 71870273V1 55068293J1 1186	523 1	1019 786 1790
31	7474637CB1	4132	3420-3535, 1- 377, 4035- 4132, 1301- 2486	1698381T6 (BLADTUT05) 4129796F6 (CARGDIT01) 55076747H1 55075847H1 55075848H1 55077477H1 GBI_g8247425_000008_000011. edit 55076756J1 GNN.g6648263_002.edit5p 6286993H2 (EPIPUNA01) 721743H2 (THYRDIE01) 6766106H1 (BRAUNOR01) 55061367H1 6766106J1 (BRAUNOR01) 1732420H1 (LIVRUT01) 55046242J2 3152909F6 (TLYMYT02)	774 3639 2871 1379 1623 1045 504 3148 2805 1041 35 1893 2149 368 1 694 1	1363 4132 3468 1783 1987 1472 1126 3745 3026 1168 503 2356 2841 909 157 1137 145
32	7170260CB1	1137	877-1137			

10

20

30

40

表4-4

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
32				7659273J1 (OVARNOE02) 55046250H1 3343082F7 (SPLNNO709)	416 692 144	971 1108 555
33	1797506CB1	3365	1-1032, 3340- 3365, 1532- 1735	1513994T6 (PANCTT01) FL1797506_g7458755_000012_g 3766209	2793 1	3365 2898
34	1851973CB1	2049	1-125, 1836- 2049, 806-915	7667239H1 (URETTUC01) 55075655J1 55077237J1 55067487H1 1454205F1 (PENITUT01) 1454205T6 (PENITUT01)	1289 547 378 1 1179 1436	1800 1222 1221 532 1617 2049
35	7474604CB1	2962	1-1526, 1757- 2114, 2481- 2962	55075789J1 8104459J1 (MIXDDIE02) 55056946J2 6884701F6 (BRAHTR03) 55067076J1 55075383J1	1760 1 1734 2255 1214 651	2440 746 2433 2962 1763 1335
36	7474721CB1	3112	2395-3112, 1353-1459, 2014-2280	6802884F6 (COLENO03) 71976507V1 55057353J1 GB1:g6996165_000001.raw GB1:g6996165.raw 55062828H1 71980671V1 1418	2055 1564 314 1910 140 1 1418	2826 2315 980 3112 1735 712 2051
37	7478815CB1	3650	862-1366, 1826-1999, 1- 787, 3623- 3650	55076655H1 6934749H1 (SINTMR02) 238539R6 (SINTNOT02) 614864T6 (COLNTUT02) 70845065V1 70842842V1	1 1710 3159 3004 1862 2420	658 2388 3647 3614 2441 3073

10

20

30

40

表4-5

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
37				72026676V1 55075416J1 9657793 70863076V1 2605255F6 (LJNGTUF07) 7355120H1 (HEARN003) GBI:98014664 97242948 CD 3012344H1 (MUSCNO107) 71179707V1 7642405J1 (SEMTDE01) 70775995V1 55024095J1 (PKINDNV04) 6854667H1 (BRAIFEN08) 7188730H2 (BRATD1C01) 70780513V1 70780809V1	966 358 919 2471 3378 7201 1 63 7488 6783 6728 1 914 1441 1353 500 384	1742 926 1011 3154 3650 7767 260 6763 7772 7436 7294 498 1558 1937 1820 981 919
38	7477141CBI	7789	1-699, 6785- 6880, 7767- 7789, 7184- 7214, 1237- 6218	55121415H1 55121423J1 7992167H1 (UTRSDIC01) 71999521V1 6822270H1 (SINTNOR01) GNN:94755212.010.edit 6594083H1 (LUNGFER02) 716449388 (PLACNOR01) 71583419V1 7402224H1 (SINIDME01) 7694930H1 (LNODTDE01) 7978995H1 (LSUBDMG01)	4574 4413 3402 1448 857 1 2835 3204 713 289 1082 1478	5373 5274 4043 1590 1407 4567 3147 3711 1385 795 1448 2186
39	2190612CBI	1937	727-1188, 1- 643, 1731- 1761			
40	7477549CBI	5373	4983-5373, 1- 1612, 2046- 2470, 4414- 4442, 2596- 2647, 2814- 3056			

10

20

30

40

【表 2 4】

表5

ポリスクレオパド SEQ ID NO:	Incyte プロジェクトID	代表的ライブラリ
21	2564295CB1	ADRETUT01
22	2837050CB1	THYRNGT03
24	7474594CB1	EMARTXE01
25	7477585CB1	BRALN0N02
26	7477587CB1	PROSTUS23
27	7594537CB1	LIVRNDCC7
28	70467491CB1	PROSNGT18
29	7478559CB1	OVARNGE02
30	1698381CB1	BLADTUT05
31	7474637CB1	EPIFUNA01
32	7170260CB1	OVARNGE02
33	1797506CB1	COLENGR03
34	1851973CB1	PENITUT01
35	7474604CB1	BRAHTR03
36	7474721CB1	COLENGR03
37	7478815CB1	SINITUT03
38	7477141CB1	SKIRNGR01
39	2190612CB1	ADRETUT07
40	7477549CB1	SINTNGR01

10

20

30

40

【表 2 5】

表6-1

ライブラリ	ベクター	ライブラリの説明
ADRETT01	PSPORT	ライブラリは、非一側性 (amitara) 副腎摘出術中の60歳のトルコ人男性の右副腎腫瘍組織から単離したRNAを使用して作製した。病理学的検査では、骨髄質に位置する限局性の出血性の転移性腎細胞癌を示した。患者は副腎皮質不全、切開創ヘルニア、および非アルコール性脂肪肝炎を示した。患者の病歴には腎細胞癌がある。家族の病歴には肝臓癌がある。
ADRETT07	pINCY	ライブラリは、一側性副腎摘出術中の43歳の白人女性から除去された副腎腫瘍組織から単離したRNAを用いて作製した。病理学的検査では褐色細胞腫が見られた。
BLADTUT05	pINCY	ライブラリは、根治的膀胱切除、根治的肺切除、及び尿造口術中の66歳白人男性から取り除かれた前立腺腫瘍組織から単離されたRNAを用いて作製された。病理学的検査では膀胱前壁にグレード3の移行性細胞癌があった。患者病歴には肺腫瘍と寛解期にあるタバコ中毒がある。家族病歴には悪性乳房腫瘍、肺結核、脳血管症、アテローム性冠動脈疾患および肺がんがある。
BMARTX01	pINCY	この5'に偏向してランダムプライムされたライブラリは、4才白人女子から除去した悪性骨髄神経芽細胞腫に由来する、処理されたSH-SY5Y (Scheiting AG) 細胞から単離されたRNAから作製された。培養液は10% 仔ウシ血清のDMEM/F12であった。80% 培養飽和密度の後、細胞は8時間、100 μM の6-ヒドロキシドナーパミン (6-OHDA) で処理された。
BRAHTR03	PGDMA2.1	このランダムプライムされたライブラリは肝内胆管癌で死した55才白人女性から抽出した原始皮質 (archaeocortex) と後部海馬組織から単離したRNAから作製した。病理学的では、主に円盤出面上に軽度の髄膜線維症、帯状皮質白質と視床に散在性軸索スフェアロイド (scattered axonal spheroids)、および嗅内皮質と周辺水道反白質領域 (periaqueductal gray region) に少数の散在した神経原線維濃縮体が見られた。関連する腫瘍組織の病理学的検査では、よく分化した肝臓の肝内胆管癌と残余の腫瘍または再発の腫瘍であった。患者病歴は肝内胆管癌、手術後 Budd-Chiari 症候群、胆汁うっ滞腹水症、水胸症、脱水症、栄養不良、乏尿また 慢性肝臓障害があった。過去の手術は胆嚢切除および肝臓の85%切除がある。
BRAHND02	pINCY	この視床線維ライブラリはある視床線維ライブラリの424万の独立型クローンから作製された。最初のRNAは、心不全疾患から死亡した35歳の白人男子から切除された視床組織から作製された。病理学検査では、中程度の軟膜線維症と大脳新皮質の複発微小腫瘍が見られた。顕微鏡によって大脳半球は、焦点石灰化がある中程度の軟膜線維症と判別した。大脳半球全体に縮んだわずかに好酸性の単体ニューロンの形跡および萎縮がある。さらに、大脳皮質全体に散在して、周辺神経節性を伴う複数の小さな空洞発生がある。

10

20

30

40

【表 2 6】

表6-2

ライブラム	ベクター	ライブラリの説明
COLEND03	PCDNA2.1	ライブラリの説明 当患者の病歴には、拡張型心筋症、鬱血性心不全、心肥大、及び肥大した脾臓と肝臓が含まれる。このライブラリは、極めて長時間(48時間/一回)の再アニーリングによるハイブリダイゼーションを用いたことを除いては、Soares 他, PNAS (1994) 91:9228 および Bonaldo 他, Genome Research (1996) 6:791 を適応した条件を用いて2回にわたりにノーマライズした。
EPIPLNA01	PSPFORT	ライブラリは、13才の白人女性から除去した結腸上皮組織から単離した RNA を用いて作製した。
LIVRN0007	pINCY	ライブラリは、17才のラテンアメリカ系男子から採取した未処理の前立腺上皮細胞組織から単離した RNA を用いて作製した。血清検査は陰性だった。
OVARND02	PCDNA2.1	ライブラリは、二人の異なる提供者からプールした cDNA を使って作製された。cDNA は、Palau 症候群で死亡した 20 週間の白人男子胎児(提供者 A)、と無脳症で死亡した 18 週間の白人女子胎児(提供者 B)の肝臓組織から生成された。家族歴には胎児の母(提供者 B)に僧帽弁逸脱症がある。
PENITE01	pINCY	この5'に偏向してランダムプライムされたライブラリは、47才の白人女性の複式子宮全摘手術、両側性卵管卵巣摘出術、膈懸垂と固定化、および付随的虫垂切除術時に除去された右卵巣組織から単離された RNA を用いて作製された。この患者は閉経前月経過多症を示した。患者病歴は骨関節痛、尿管妊娠、また左股の小児麻痺整骨療法を受けたことがある。過去の手術は胃腸吻合、口蓋の整形手術、扁桃腺アデノイド切除、拡張術と橋脚術、胆嚢摘出術、また膀胱再建術があった。患者の使用薬剤には、ビタミン剤、鉄、亜鉛があった。家族病歴は、父親に良性高血圧症およびII型糖尿病が、また兄弟姉妹にII型糖尿病がある。
		ライブラリは、64才の白人男性の陰茎切除中に、陰茎組織から除いた腫瘍組織から単離した RNA を用いて作製した。病理学検査は、包皮の内側壁を占め、陰茎龜頭まで広がったカビ状の浸潤性グレード4の扁平上皮癌がある。患者病歴は人腸の良性腫瘍、アテローム性冠動脈疾患、狭心症、風痛、また肥満症がある。家族歴は悪性個頭腫、慢性リンパ性白血病、および慢性肝臓病がある。

10

20

30

40

表6-3

ライブラム	ベクター	ライブラリの説明
PROSMOT18	pLNCY	ライブラリは、根治的膀胱切除、及び尿注回中の58才白人男性から取り除かれた炎症性前立腺腫瘍組織から単離されたRNAを用いて作製された。病理学検査では、腺癌性過形成があり、この組織はグレード3の移行性細胞癌に適合していた。患者の病歴には狭心症と気腫がある。家族歴には慢性心筋梗塞、アテローム性冠動脈疾患およびII型糖尿病がある。
PROSTUS23	pLNCY	このサブトラクシオンハイブリダイゼーションされた前立腺腫瘍ライブラリは、プールされた前立腺組織ライブラリの1000万のクローンと2回のサブトラクシオンハイブリダイゼーションにかけられた、プールされた前立腺腫瘍ライブラリの1000万のクローンをを用いて作製した。サブトラクシオンハイブリダイゼーション用開始ライブラリは、58才(N)、61才(B)、66才(O)、および68才(D)の白人男子からリンパ節切除を伴う前立腺切除術時に採取した前立腺腫瘍から単離したmRNAを用いた4個の前立腺腫瘍ライブラリから等しい数のクローンをプールすることによって作製した。病理はすべてのドナーにおいて腺癌を示した。病歴は、ドナーAでは、PSAの上昇、硬結、およびタバコの濫用、ドナーBはPSAの上昇、硬結、前立腺肥大、腎不全、骨関節炎、腎動脈狭窄、良性のHTN、血小板減少症、高脂血症、タバコ/アルコール濫用、およびC型肝炎(保有者)、ドナーCは、PSAの上昇、硬結、およびタバコの濫用、また、ドナーDは、PSAの上昇、硬結、高コレステロール血症、および腎結石である。サブトラクシオンハイブリダイゼーションプロトコルは、前立腺組織、前立腺上皮細胞および二人のドナーによる前立腺ストローマからの総単離細胞から由来する3個の前立腺組織ライブラリから得られた同じ数のcDNAクローンをプールすることによって作製した。サブトラクシオンハイブリダイゼーションの条件はSwarcopら、NAR 19 (1991):1954 および Donaldo らの Genome Research 6 (1996):791 の方法に基づいたものである。(1998) Genome Res. 8:186-194.
SKINOR01	PCDNA2.1	このランダムプライムライブラリは17才の白人女性の両側小乳房形成術時に採取した皮膚組織から単離したRNAを用いて作製した。患者の病歴には乳房肥大が含まれる。家族歴には良性高血圧が含まれている。
SINTUT03	pLNCY	ライブラリは49才白人女性の腹膜組織、腹膜癒着剥離、回腸切除および永久的結腸造設術時に採取した回腸腫瘍組織から単離したRNAを用いて作製した。病歴はグレード4の腺癌を示した。患者の病歴には良性高血圧がある。前に行われた手術には、複式子宮全摘出、両側性卵巣摘出、限局性リンパ節切除、付随的虫垂切除および極肥がある。家族歴には良性高血圧、脳血管疾患、高脂血症、アテローム硬化型冠動脈血管疾患、高脂血症、II型糖尿病および胃癌がある。
SINTOR01	PCDNA2.1	この無作為プライムライブラリは31才白人女性のRoux-en-Y胃吻合術時に小腸組織から単離されたRNAを用いて作製された。患者の病歴には臨床的肥満がある。
THYNOT03	pLNCY	ライブラリは28才白人女性の甲状腺全摘出術時に左甲状腺から取り除かれた甲状腺組織から単離したRNAを用いて作製した。病歴は左甲状腺に小結節の腺腫性過形成が見られた。関連腫瘍組織の病理は、顕著な濾胞状腺癌を示しており、左甲状腺に被包性塊が形成されていた。

10

20

30

40

表7-1

【表 2 9】

プログラム	説明	引用文献	パラメータ閾値
ABI FACTURA	核酸配列においてベクター配列を除去して、特定の塩基をマスクするプログラム	Applied Biosystems, Foster City, CA.	
ABI/PARACELI FDF	Fast Data Finderは、アミノ酸または核酸配列の比較、注釈付けに有用である。	Applied Biosystems, Foster City, CA. Paracel Inc., Pasadena, CA.	不一致<50%
ABI グートアセンブラ	核酸配列を構築するプログラム。	Applied Biosystems, Foster City, CA.	
BLAST	Basic Local Alignment Search Toolはアミノ酸および核酸配列の配列類似性検索に有用であり、BLASTには5つの機能がある：blastp、blastn、blastx、tblastnおよびtblastx。	Altschul, S.F. 他(1990) J.Mol.Biol. 215:403-410; Altschul, S.F. 他(1997) (1997) Nucleic Acids Res. 25:3389-3402.	ESYs: 確率値=1.0E-8 以下 完全長配列: 確率値=1.0E-10以下
FASTA	問合せ配列と同種の配列群配との類似性を検索するPearson およびLipman アルゴリズム。FASTAは最低5つの機能からなる：fasta、tfasta、fastx、tblastxおよびsssearch。	Pearson, W.R. and D.J.Lipman (1988) Proc.Nat.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.	ESYs:fasta E 値=1.0E-6 構築された ESYs:fasta 同一性=95%以上 一致した長さ=200塩基以上 fastx E 値=1.0E-8 以下 完全長配列: fastx スコア=100以上;
BLIMPS	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、配列相同性および情造的フィングア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; およびAttwood, T.K. 他。(1997) J.Chem.Inf.Comput.Sci. 37:417-424.	確率値=1.0E-3以下
HMMER	PFAMのようなタンパク質ファミリーコンセンサス配列の隠れマルコフモデル(HMM)に基づいたデータベースに対して問合せ配列を検索するアルゴリズム。	Krogh, A. 他。(1994) J.Mol.Biol.235:1501-1531; Somnhammer, 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 シグナルペプティドヒット: スコア<0以上。

10

20

30

40

表7-2

プログラム	説明	引用文献	パラメータ関数
PROFILESSCAN	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とその特定のPrositeモチーフに対するGCG指定のJIGHPH順 通常、スコア=1.4-2.1.
Phred	高い感度と確率で自動配列決定機のトレースを調べ る塩基誤出しアルゴリズム。	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なインプリンメンテーションに基づくSWATやCrossMatchを含むPhils Revised Assembly プログラムで、配列相同性の検索やDNA配列の構築に有用である。	Smith, T.F. 及び M.S. Waterman (1981) Adv. Appl. Math. 2:482-489; Smith, T.F. 及び M.S. Waterman (1981) J. Mol. Biol. 147:195-197; and Green, P., University of Washington, Seattle, WA.	スコア=120以上 一致した長さ=56以上
Consed	Phrap で構築したものの表示、編集用グラフィカルツール。	Gordon, D. 他(1998) Genome Res. 8:195-202.	
SPスキヤン	タンパク質配列をスキヤンして、分泌シグナルペプチドの存在を調べる重み付けマトリックス解析プログラム。	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	加重マトリックスを用いて蛋白質配列での既知通セグメントの描写および配向を決定するプログラム。	Persson, B. and P. Argos (1994) J. Mol. Biol. 237:182-192; Persson, B. および P. Argos (1996) Protein Sci. 5:363-371.	
TMFMMER	隠れMarkov モデル(HMM)を使って蛋白質配列の膜貫通セグメントを描写し、配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., Biol., Glasgow 他, eds., The Am. Assoc. for Artificial Intelligence Press, Menlo Park, CA, pp. 175-182.	
MOTIFS	Prositeで定義された配列と一致したペプタ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, 第9版, M51-59頁, Genetics Computer Group, Madison, WI	

10

20

30

40

【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
31 January 2002 (31.01.2002)

PCT

(10) International Publication Number
WO 02/08399 A2

(51) International Patent Classification: C12N 9/00

(21) International Application Number: PCT/US01/23092

(22) International Filing Date: 20 July 2001 (20.07.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/220,038 21 July 2000 (21.07.2000) US
60/222,112 28 July 2000 (28.07.2000) US
60/222,831 4 August 2000 (04.08.2000) US
60/224,729 11 August 2000 (11.08.2000) US

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

(71) Applicant and
(72) Inventor: THORNTON, Michael [US/US]; 9 Madway Road, Woodside, CA 94062 (US).

(72) Inventors: and

(75) Inventors/Applicants (for US only): YUE, Henry [US/US]; 826 Lois Avenue, Sunnyvale, CA 94087 (US). KHAN, Farrak, A. [IN/US]; 3617 Central Road #102, Glenview, IL 60025 (US). GURURAJAN, Rajagopal [IN/US]; 5591 Dent Avenue, San Jose, CA 95118 (US). HAFALIA, April, J. A. [US/US]; 2227 Calle de Primavera, Santa Clara, CA 95054 (US). WALIA, Narinder, K. [US/US]; 890 Davis Street #205, San Leandro, CA 94577 (US). PATTERSON, Chandra [US/US]; 490 Sherwood Way #1, Menlo Park, CA 94025 (US). RAMKUMAR, Jayalaxmi [IN/US]; 34359 Maybird Circle, Fremont, CA 94555 (US). GANDHI, Ameena, R. [US/US]; 837 Roble Avenue, #1, Menlo Park, CA 94025 (US). POLICKY, Jennifer, L. [US/US]; 1511 Jarvis Court, San Jose, CA 95118 (US). BAUGHN, Mariah, R. [US/US]; 14244 Santiago Road, San Leandro, CA 94577 (US). TRIBOLEY, Catherine, M. [US/US]; 1121 Tennessee Street, #5, San Francisco, CA 94107 (US). BANDMAN, Olga [US/US]; 366 Anna Avenue, MT. View, CA 94043 (US). NGUYEN, Dannie, B. [US/US]; 1403 Ridgewood Drive, San Jose, CA 95118 (US). LI, Yan [CN/US]; 3885 Corina Way, Palo Alto, CA 94303 (US). BURFORD, Neil [US/US]; 105 Wildwood Circle, Durham, CT 06422

(US). LAL, Preeti [US/US]; P.O. Box 5142, Santa Clara, CA 95056 (US). DING, Li [CN/US]; 3353 Alma Street #146, Palo Alto, CA 94306 (US). YAO, Monique, G. [US/US]; 111 Frederick Court, Mountain View, CA 94043 (US). ELLIOTT, Vicki, S. [US/US]; 3770 Pebble Place Way, San Jose, CA 95121 (US). RECIPON, Shirley, A. [US/US]; 85 Fortuna Avenue, San Francisco, CA 94115 (US). KEARNEY, Liam [IE/US]; 50 Woodside Avenue, San Francisco, CA 94127 (US). LI, Dyung, Aina, M. [US/US]; 233 Coy Drive, San Jose, CA 95123 (US). GREENWALD, Sara, R. [US/US]; 21 Bucareli Drive, San Francisco, CA 94132 (US). TANG, Y., Tom [US/US]; 4230 Ranwick Court, San Jose, CA 95118 (US). XU, Yuming [US/US]; 1739 Walnut Drive, Mountain View, CA 94040 (US). WALSH, Roderick, T. [IE/GB]; 8 Boundary Court, St. Lawrence Road, Canterbury, Kent CT1 3EZ (GB). GIETZEN, Kimberly, J. [US/US]; 691 Los Huecos Drive, San Jose, CA 95123 (US). YANG, Junming [CN/US]; 7125 Bark Lane, San Jose, CA 95129 (US). HILLMAN, Jennifer, L. [US/US]; 230 Monroe Drive, #17, Mountain View, CA 94040 (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, TI, 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, GQ, 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.

(54) Title: HUMAN KINASES

(57) Abstract: The invention provides human human kinases (PKIN) and polynucleotides which identify and encode PKIN. 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 PKIN.



WO 02/08399 A2

WO 02/08399

PCT/US01/23092

HUMAN KINASES**TECHNICAL FIELD**

This invention relates to nucleic acid and amino acid sequences of human kinases and to the
5 use of these sequences in the diagnosis, treatment, and prevention of cancer, immune disorders,
disorders affecting growth and development, cardiovascular diseases, and lipid disorders, and in the
assessment of the effects of exogenous compounds on the expression of nucleic acid and amino acid
sequences of human kinases.

BACKGROUND OF THE INVENTION

Kinases comprise the largest known enzyme superfamily and vary widely in their target
molecules. Kinases catalyze the transfer of high energy phosphate groups from a phosphate donor to
a phosphate acceptor. Nucleotides usually serve as the phosphate donor in these reactions, with most
kinases utilizing adenosine triphosphate (ATP). The phosphate acceptor can be any of a variety of
15 molecules, including nucleosides, nucleotides, lipids, carbohydrates, and proteins. Proteins are
phosphorylated on hydroxyamino acids. Addition of a phosphate group alters the local charge on the
acceptor molecule, causing internal conformational changes and potentially influencing intermolecular
contacts. Reversible protein phosphorylation is the primary method for regulating protein activity in
eukaryotic cells. In general, proteins are activated by phosphorylation in response to extracellular
20 signals such as hormones, neurotransmitters, and growth and differentiation factors. The activated
proteins initiate the cell's intracellular response by way of intracellular signaling pathways and second
messenger molecules such as cyclic nucleotides, calcium-calmodulin, inositol, and various mitogens,
that regulate protein phosphorylation.

Kinases are involved in all aspects of a cell's function, from basic metabolic processes, such
25 as glycolysis, to cell-cycle regulation, differentiation, and communication with the extracellular
environment through signal transduction cascades. Inappropriate phosphorylation of proteins in cells
has been linked to changes in cell cycle progression and cell differentiation. Changes in the cell cycle
have been linked to induction of apoptosis or cancer. Changes in cell differentiation have been linked
to diseases and disorders of the reproductive system, immune system, and skeletal muscle.

30 There are two classes of protein kinases. One class, protein tyrosine kinases (PTKs),
phosphorylates tyrosine residues, and the other class, protein serine/threonine kinases (STKs),
phosphorylates serine and threonine residues. Some PTKs and STKs possess structural
characteristics of both families and have dual specificity for both tyrosine and serine/threonine

WO 02/08399

PCT/US01/23092

residues. Almost all kinases contain a conserved 250-300 amino acid catalytic domain containing specific residues and sequence motifs characteristic of the kinase family. The protein kinase catalytic domain can be further divided into 11 subdomains. N-terminal subdomains I-IV fold into a two-lobed structure which binds and orients the ATP donor molecule, and subdomain V spans the two lobes. C-terminal subdomains VI-XI bind the protein substrate and transfer the gamma phosphate from ATP to the hydroxyl group of a tyrosine, serine, or threonine residue. Each of the 11 subdomains contains specific catalytic residues or amino acid motifs characteristic of that subdomain. For example, subdomain I contains an 8-amino acid glycine-rich ATP binding consensus motif, subdomain II contains a critical lysine residue required for maximal catalytic activity, and subdomains VI through IX comprise the highly conserved catalytic core. PTKs and STKs also contain distinct sequence motifs in subdomains VI and VIII which may confer hydroxyamino acid specificity.

In addition, kinases may also be classified by additional amino acid sequences, generally between 5 and 100 residues, which either flank or occur within the kinase domain. These additional amino acid sequences regulate kinase activity and determine substrate specificity. (Reviewed in Hardie, G. and S. Hanks (1995) The Protein Kinase Facts Book, Vol I, pp. 17-20 Academic Press, San Diego CA.). In particular, two protein kinase signature sequences have been identified in the kinase domain, the first containing an active site lysine residue involved in ATP binding, and the second containing an aspartate residue important for catalytic activity. If a protein analyzed includes the two protein kinase signatures, the probability of that protein being a protein kinase is close to 100% (PROSITE: PDOC00100, November 1995).

Protein Tyrosine Kinases

Protein tyrosine kinases (PTKs) may be classified as either transmembrane, receptor PTKs or nontransmembrane, nonreceptor PTK proteins. Transmembrane tyrosine kinases function as receptors for most growth factors. Growth factors bind to the receptor tyrosine kinase (RTK), which causes the receptor to phosphorylate itself (autophosphorylation) and specific intracellular second messenger proteins. Growth factors (GF) that associate with receptor PTKs include epidermal GF, platelet-derived GF, fibroblast GF, hepatocyte GF, insulin and insulin-like GFs, nerve GF, vascular endothelial GF, and macrophage colony stimulating factor.

Nontransmembrane, nonreceptor PTKs lack transmembrane regions and, instead, form signaling complexes with the cytosolic domains of plasma membrane receptors. Receptors that function through non-receptor PTKs include those for cytokines and hormones (growth hormone and prolactin), and antigen-specific receptors on T and B lymphocytes.

Many PTKs were first identified as oncogene products in cancer cells in which PTK

WO 02/08399

PCT/US01/23092

activation was no longer subject to normal cellular controls. In fact, about one third of the known oncogenes encode PTKs. Furthermore, cellular transformation (oncogenesis) is often accompanied by increased tyrosine phosphorylation activity (Charbonneau, H. and N.K. Tonks (1992) *Annu. Rev. Cell Biol.* 8:463-493). Regulation of PTK activity may therefore be an important strategy in controlling
5 some types of cancer.

Substrates for tyrosine kinases can be identified using anti-phosphotyrosine antibodies to screen tyrosine-phosphorylated cDNA expression libraries. Fish, so named for tyrosine-phosphorylated in Src-transformed fibroblast, is a tyrosine kinase substrate which has been identified by such a technique. Fish has five SH3 domains and a phospho homology (PX) domain. Fish is
10 suggested to be involved in signalling by tyrosine kinases and have a role in the actin cytoskeleton (Lock, P. et al (1998) *EMBO J.* 17:4346-4357).

SHP-2, an SH2-domain-containing phosphotyrosine phosphatase, is a positive signal transducer for several receptor tyrosine kinases (RTKs) and cytokine receptors. Phosphotyrosine phosphatases are critical positive and negative regulators in the intracellular signalling pathways that
15 result in growth-factor-specific cell responses such as mitosis, migration, differentiation, transformation, survival or death. Signal-regulatory proteins (SIRPs) comprise a new gene family of at least 15 members, consisting of two subtypes distinguished by the presence or absence of a cytoplasmic SHP-2-binding domain. The SIRP-alpha subfamily members have a cytoplasmic SHP2-binding domain and includes SIRP-alpha-1, a transmembrane protein, a substrate of activated RTKs
20 and which binds to SH2 domains. SIRPs have a high degree of homology with immune antigen recognition molecules. The SIRP-beta subfamily lacks the cytoplasmic tail. The SIRP-beta-1 gene encodes a polypeptide of 398 amino acids. SIRP family members are generally involved in regulation of signals which define different physiological and pathological processes (Kharitonov, A. et al (1997) *Nature* 386:181-186). Two possible areas of regulation include determination of brain diversity and
25 genetic individuality (Sano, S et al (1999) *Biochem. J.* 344 Pt 3:667-675) and recognition of self which fails in diseases such as hemolytic anemia (Oldenborg, P.-A et al (2000) *Science* 288:2051-2054).

Protein Serine/Threonine Kinases

Protein serine/threonine kinases (STKs) are nontransmembrane proteins. A subclass of STKs are known as ERKs (extracellular signal regulated kinases) or MAPs (mitogen-activated
30 protein kinases) and are activated after cell stimulation by a variety of hormones and growth factors. Cell stimulation induces a signaling cascade leading to phosphorylation of MEK (MAP/ERK kinase) which, in turn, activates ERK via serine and threonine phosphorylation. A varied number of proteins represent the downstream effectors for the active ERK and implicate it in the control of cell

WO 02/08399

PCT/US01/23092

proliferation and differentiation, as well as regulation of the cytoskeleton. Activation of ERK is normally transient, and cells possess dual specificity phosphatases that are responsible for its down-regulation. Also, numerous studies have shown that elevated ERK activity is associated with some cancers. Other STKs include the second messenger dependent protein kinases such as the
5 cyclic-AMP dependent protein kinases (PKA), calcium-calmodulin (CaM) dependent protein kinases, and the mitogen-activated protein kinases (MAP); the cyclin-dependent protein kinases; checkpoint and cell cycle kinases; Numb-associated kinase (Nak); human Fused (hFu); proliferation-related kinases; 5'-AMP-activated protein kinases; and kinases involved in apoptosis.

The second messenger dependent protein kinases primarily mediate the effects of second
10 messengers such as cyclic AMP (cAMP), cyclic GMP, inositol triphosphate, phosphatidylinositol, 3,4,5-triphosphate, cyclic ADP ribose, arachidonic acid, diacylglycerol and calcium-calmodulin. The PKAs are involved in mediating hormone-induced cellular responses and are activated by cAMP produced within the cell in response to hormone stimulation. cAMP is an intracellular mediator of hormone action in all animal cells that have been studied. Hormone-induced cellular responses include
15 thyroid hormone secretion, cortisol secretion, progesterone secretion, glycogen breakdown, bone resorption, and regulation of heart rate and force of heart muscle contraction. PKA is found in all animal cells and is thought to account for the effects of cAMP in most of these cells. Altered PKA expression is implicated in a variety of disorders and diseases including cancer, thyroid disorders, diabetes, atherosclerosis, and cardiovascular disease (Isselbacher, K.J. et al. (1994) Harrison's
20 Principles of Internal Medicine, McGraw-Hill, New York NY, pp. 416-431, 1887).

The casein kinase I (CKI) gene family is another subfamily of serine/threonine protein kinases. This continuously expanding group of kinases have been implicated in the regulation of numerous cytoplasmic and nuclear processes, including cell metabolism, and DNA replication and repair. CKI enzymes are present in the membranes, nucleus, cytoplasm and cytoskeleton of
25 eukaryotic cells, and on the mitotic spindles of mammalian cells (Fish, K.J. et al. (1995) *J. Biol. Chem.* 270:14875-14883).

The CKI family members all have a short amino-terminal domain of 9-76 amino acids, a highly conserved kinase domain of 284 amino acids, and a variable carboxyl-terminal domain that ranges from 24 to over 200 amino acids in length (Cegielska, A. et al. (1998) *J. Biol. Chem.* 273:1357-1364).
30 The CKI family is comprised of highly related proteins, as seen by the identification of isoforms of casein kinase I from a variety of sources. There are at least five mammalian isoforms, α , β , γ , δ , and ϵ . Fish et al., identified CKI-epsilon from a human placenta cDNA library. It is a basic protein of 416 amino acids and is closest to CKI-delta. Through recombinant expression, it was determined to

WO 02/08399

PCT/US01/23092

phosphorylate known CKI substrates and was inhibited by the CKI-specific inhibitor CKI-7. The human gene for CKI-epsilon was able to rescue yeast with a slow-growth phenotype caused by deletion of the yeast CKI locus, HRR250 (Fish et al., *supra*).

The mammalian circadian mutation tau was found to be a semidominant autosomal allele of
5 CKI-epsilon that markedly shortens period length of circadian rhythms in Syrian hamsters. The tau
locus is encoded by casein kinase I-epsilon, which is also a homolog of the *Drosophila* circadian gene
double-time. Studies of both the wildtype and tau mutant CKI-epsilon enzyme indicated that the
mutant enzyme has a noticeable reduction in the maximum velocity and autophosphorylation state.
Further, *in vitro*, CKI-epsilon is able to interact with mammalian PERIOD proteins, while the mutant
10 enzyme is deficient in its ability to phosphorylate PERIOD. Lowrey et al., have proposed that CKI-
epsilon plays a major role in delaying the negative feedback signal within the transcription-translation-
based autoregulatory loop that composes the core of the circadian mechanism. Therefore the CKI-
epsilon enzyme is an ideal target for pharmaceutical compounds influencing circadian rhythms, jet-lag
and sleep, in addition to other physiologic and metabolic processes under circadian regulation (Lowrey,
15 P.L. et al. (2000) *Science* 288:483-491).

Homeodomain-interacting protein kinases (HIPKs) are serine/threonine kinases and novel
members of the DYRK kinase subfamily (Hofmann, T.G. et al. (2000) *Biochimie* 82:1123-1127).
HIPKs contain a conserved protein kinase domain separated from a domain that interacts with
homeoproteins. HIPKs are nuclear kinases, and HIPK2 is highly expressed in neuronal tissue (Kim,
20 Y.H. et al. (1998) *J. Biol. Chem.* 273:25875-25879; Wang, Y. et al. (2001) *Biochim. Biophys. Acta*
1518:168-172). HIPKs act as corepressors for homeodomain transcription factors. This corepressor
activity is seen in posttranslational modifications such as ubiquitination and phosphorylation, each of
which are important in the regulation of cellular protein function (Kim, Y.H. et al. (1999) *Proc. Natl.*
Acad. Sci. USA 96:12350-12355).

The murine homology to *Caenorhabditis elegans* UNC51, a serine/threonine kinase, has
been determined to be required to signal the program of gene expression leading to axon formation
from granule cells of the cerebellar cortex (Tomoda, T. et al (1999) *Neuron* 24:833-346. The human
homolog of UNC-51, ULK1, for UNC-51 (*C. elegans*)-like kinase I, is composed of 1050 amino acids,
has a calculated MV of 112.6 kDa and a pI of 8.80. ULK1 has 41% overall sequence similarity to
30 UNC-51 and is highly conserved among vertebrates including mammals, birds, reptiles, amphibians,
and fish. By Northern blot analysis, Kuroyanagi et al have shown ULK1 to be ubiquitously expressed
in adult tissues, including skeletal muscle, heart, pancreas, brain, placenta, liver, kidney, and lung while
UNC-51 has been specifically located in the nervous system of *C. elegans*. Fish and RH mapping

WO 02/08399

PCT/US01/23092

confirmed the localization of ULK1 to human chromosome 12q24.3. (Kuroyanagi, H. et al (1998) Genomics 51:76-85.

Calcium-Calmodulin Dependent Protein Kinases

Calcium-calmodulin dependent (CaM) kinases are involved in regulation of smooth muscle contraction, glycogen breakdown (phosphorylase kinase), and neurotransmission (CaM kinase I and CaM kinase II). CaM dependent protein kinases are activated by calmodulin, an intracellular calcium receptor, in response to the concentration of free calcium in the cell. Many CaM kinases are also activated by phosphorylation. Some CaM kinases are also activated by autophosphorylation or by other regulatory kinases. CaM kinase I phosphorylates a variety of substrates including the neurotransmitter-related proteins synapsin I and II, the gene transcription regulator, CREB, and the cystic fibrosis conductance regulator protein, CFTR (Haribabu, B. et al. (1995) EMBO J. 14:3679-3686). CaM kinase II also phosphorylates synapsin at different sites and controls the synthesis of catecholamines in the brain through phosphorylation and activation of tyrosine hydroxylase. CaM kinase II controls the synthesis of catecholamines and serotonin, through phosphorylation/activation of tyrosine hydroxylase and tryptophan hydroxylase, respectively (Fujisawa, H. (1990) BioEssays 12:27-29). The mRNA encoding a calmodulin-binding protein kinase-like protein was found to be enriched in mammalian forebrain. This protein is associated with vesicles in both axons and dendrites and accumulates largely postnatally. The amino acid sequence of this protein is similar to CaM-dependent STKs, and the protein binds calmodulin in the presence of calcium (Godbout, M. et al. (1994) J. Neurosci. 14:1-13).

Mitogen-Activated Protein Kinases

The mitogen-activated protein kinases (MAP) which mediate signal transduction from the cell surface to the nucleus via phosphorylation cascades are another STK family that regulates intracellular signaling pathways. Several subgroups have been identified, and each manifests different substrate specificities and responds to distinct extracellular stimuli (Egan, S.E. and R.A. Weinberg (1993) Nature 365:781-783). There are 3-kinase modules comprising the MAP kinase cascade: MAPK (MAP), MAPK kinase (MAP2K, MAPKK, or MKK), and MKK kinase (MAP3K, MAPKKK, OR MEKK) (Wang,X.S. et al (1998) Biochem. Biophys. Res. Commun. 253:33-37). The extracellular-regulated kinase (ERK) pathway is activated by growth factors and mitogens, for example, epidermal growth factor (EGF), ultraviolet light, hyperosmolar medium, heat shock, endotoxic lipopolysaccharide (LPS). The closely related though distinct parallel pathways, the c-Jun N-terminal kinase (JNK), or stress-activated kinase (SAPK) pathway, and the p38 kinase pathway are activated by stress stimuli and proinflammatory cytokines such as tumor necrosis factor (TNF) and interleukin-1

WO 02/08399

PCT/US01/23092

(IL-1). Altered MAP kinase expression is implicated in a variety of disease conditions including cancer, inflammation, immune disorders, and disorders affecting growth and development. MAP kinase signaling pathways are present in mammalian cells as well as in yeast.

MAPKKK6 (MAP3K6) is one of numerous MAP3Ks identified. Isolated from skeletal muscle, MAP3K6 is 1,280 amino acids in length with 11 kinase subdomains and is detected in several tissues. The highest expression has been found in heart and skeletal muscle. MAP3K6 has 45% amino acid sequence identity with MAP3K5, while their catalytic domains share 82% identity. MAP3K6 interaction with MAP3K5 *in vivo* was confirmed by coimmunoprecipitation. Recombinant MAP3K6 has been shown to weakly activate the JNK but not the p38 kinase or ERK pathways

(Wang, X.S. et al. *supra*)

Cyclin-Dependent Protein Kinases

The cyclin-dependent protein kinases (CDKs) are STKs that control the progression of cells through the cell cycle. The entry and exit of a cell from mitosis are regulated by the synthesis and destruction of a family of activating proteins called cyclins. Cyclins are small regulatory proteins that bind to and activate CDKs, which then phosphorylate and activate selected proteins involved in the mitotic process. CDKs are unique in that they require multiple inputs to become activated. In addition to cyclin binding, CDK activation requires the phosphorylation of a specific threonine residue and the dephosphorylation of a specific tyrosine residue on the CDK.

Another family of STKs associated with the cell cycle are the NIMA (never in mitosis)-related kinases (Neks). Both CDKs and Neks are involved in duplication, maturation, and separation of the microtubule organizing center, the centrosome, in animal cells (Fry, A.M. et al. (1998) EMBO J. 17:470-481).

Checkpoint and Cell Cycle Kinases

In the process of cell division, the order and timing of cell cycle transitions are under control of cell cycle checkpoints, which ensure that critical events such as DNA replication and chromosome segregation are carried out with precision. If DNA is damaged, e.g. by radiation, a checkpoint pathway is activated that arrests the cell cycle to provide time for repair. If the damage is extensive, apoptosis is induced. In the absence of such checkpoints, the damaged DNA is inherited by aberrant cells which may cause proliferative disorders such as cancer. Protein kinases play an important role in this process. For example, a specific kinase, checkpoint kinase 1 (Chk1), has been identified in yeast and mammals, and is activated by DNA damage in yeast. Activation of Chk1 leads to the arrest of the cell at the G2/M transition (Sanchez, Y. et al. (1997) Science 277:1497-1501). Specifically, Chk1 phosphorylates the cell division cycle phosphatase CDC25, inhibiting its normal function which is

WO 02/08399

PCT/US01/23092

to dephosphorylate and activate the cyclin-dependent kinase Cdc2. Cdc2 activation controls the entry of cells into mitosis (Peng, C.-Y. et al. (1997) *Science* 277:1501-1505). Thus, activation of Chk1 prevents the damaged cell from entering mitosis. A similar deficiency in a checkpoint kinase, such as Chk1, may also contribute to cancer by failure to arrest cells with damaged DNA at other checkpoints such as G2/M.

Proliferation-Related Kinases

Proliferation-related kinase is a serum/cytokine inducible STK that is involved in regulation of the cell cycle and cell proliferation in human megakaryocytic cells (Li, B. et al. (1996) *J. Biol. Chem.* 271:19402-19408). Proliferation-related kinase is related to the polo (derived from *Drosophila* polo gene) family of STKs implicated in cell division. Proliferation-related kinase is downregulated in lung tumor tissue and may be a proto-oncogene whose deregulated expression in normal tissue leads to oncogenic transformation.

5'-AMP-activated protein kinase

A ligand-activated STK protein kinase is 5'-AMP-activated protein kinase (AMPK) (Gao, G. et al. (1996) *J. Biol. Chem.* 271:8675-8681). Mammalian AMPK is a regulator of fatty acid and sterol synthesis through phosphorylation of the enzymes acetyl-CoA carboxylase and hydroxymethylglutaryl-CoA reductase and mediates responses of these pathways to cellular stresses such as heat shock and depletion of glucose and ATP. AMPK is a heterotrimeric complex comprised of a catalytic alpha subunit and two non-catalytic beta and gamma subunits that are believed to regulate the activity of the alpha subunit. Subunits of AMPK have a much wider distribution in non-lipogenic tissues such as brain, heart, spleen, and lung than expected. This distribution suggests that its role may extend beyond regulation of lipid metabolism alone.

Kinases in Apoptosis

Apoptosis is a highly regulated signaling pathway leading to cell death that plays a crucial role in tissue development and homeostasis. Deregulation of this process is associated with the pathogenesis of a number of diseases including autoimmune disease, neurodegenerative disorders, and cancer. Various STKs play key roles in this process. ZIP kinase is an STK containing a C-terminal leucine zipper domain in addition to its N-terminal protein kinase domain. This C-terminal domain appears to mediate homodimerization and activation of the kinase as well as interactions with transcription factors such as activating transcription factor, ATF4, a member of the cyclic-AMP responsive element binding protein (ATF/CREB) family of transcriptional factors (Sanjo, H. et al. (1998) *J. Biol. Chem.* 273:29066-29071). DRAK1 and DRAK2 are STKs that share homology with the death-associated protein kinases (DAP kinases), known to function in interferon- γ induced

WO 02/08399

PCT/US01/23092

apoptosis (Sanjo et al., *supra*). Like ZIP kinase, DAP kinases contain a C-terminal protein-protein interaction domain, in the form of ankyrin repeats, in addition to the N-terminal kinase domain. ZIP, DAP, and DRAX kinases induce morphological changes associated with apoptosis when transfected into NIH3T3 cells (Sanjo et al., *supra*). However, deletion of either the N-terminal kinase catalytic domain or the C-terminal domain of these proteins abolishes apoptosis activity, indicating that in addition to the kinase activity, activity in the C-terminal domain is also necessary for apoptosis, possibly as an interacting domain with a regulator or a specific substrate.

RICK is another STK recently identified as mediating a specific apoptotic pathway involving the death receptor, CD95 (Inohara, N. et al. (1998) *J. Biol. Chem.* 273:12296-12300). CD95 is a member of the tumor necrosis factor receptor superfamily and plays a critical role in the regulation and homeostasis of the immune system (Nagata, S. (1997) *Cell* 88:355-365). The CD95 receptor signaling pathway involves recruitment of various intracellular molecules to a receptor complex following ligand binding. This process includes recruitment of the cysteine protease caspase-8 which, in turn, activates a caspase cascade leading to cell death. RICK is composed of an N-terminal kinase catalytic domain and a C-terminal "caspase-recruitment" domain that interacts with caspase-like domains, indicating that RICK plays a role in the recruitment of caspase-8. This interpretation is supported by the fact that the expression of RICK in human 293T cells promotes activation of caspase-8 and potentiates the induction of apoptosis by various proteins involved in the CD95 apoptosis pathway (Inohara et al., *supra*).

20 Mitochondrial Protein Kinases

A novel class of eukaryotic kinases, related by sequence to prokaryotic histidine protein kinases, are the mitochondrial protein kinases (MPKs) which seem to have no sequence similarity with other eukaryotic protein kinases. These protein kinases are located exclusively in the mitochondrial matrix space and may have evolved from genes originally present in respiration-dependent bacteria which were endocytosed by primitive eukaryotic cells. MPKs are responsible for phosphorylation and inactivation of the branched-chain alpha-ketoacid dehydrogenase and pyruvate dehydrogenase complexes (Harris, R.A. et al. (1995) *Adv. Enzyme Regul.* 34:147-162). Five MPKs have been identified. Four members correspond to pyruvate dehydrogenase kinase isozymes, regulating the activity of the pyruvate dehydrogenase complex, which is an important regulatory enzyme at the interface between glycolysis and the citric acid cycle. The fifth member corresponds to a branched-chain alpha-ketoacid dehydrogenase kinase, important in the regulation of the pathway for the disposal of branched-chain amino acids. (Harris, R.A. et al. (1997) *Adv. Enzyme Regul.* 37:271-293). Both starvation and the diabetic state are known to result in a great increase in the activity of the pyruvate

WO 02/08399

PCT/US01/23092

dehydrogenase kinase in the liver, heart and muscle of the rat. This increase contributes in both disease states to the phosphorylation and inactivation of the pyruvate dehydrogenase complex and conservation of pyruvate and lactate for gluconeogenesis (Harris (1995) supra).

5 KINASES WITH NON-PROTEIN SUBSTRATES

Lipid and Inositol kinases

Lipid kinases phosphorylate hydroxyl residues on lipid head groups. A family of kinases involved in phosphorylation of phosphatidylinositol (PI) has been described, each member phosphorylating a specific carbon on the inositol ring (Leevers, S.J. et al. (1999) *Curr. Opin. Cell Biol.* 11:219-225). The phosphorylation of phosphatidylinositol is involved in activation of the protein kinase C signaling pathway. The inositol phospholipids (phosphoinositides) intracellular signaling pathway begins with binding of a signaling molecule to a G-protein linked receptor in the plasma membrane. This leads to the phosphorylation of phosphatidylinositol (PI) residues on the inner side of the plasma membrane by inositol kinases, thus converting PI residues to the biphosphate state (PIP₂). PIP₂ is then cleaved into inositol triphosphate (IP₃) and diacylglycerol. These two products act as mediators for separate signaling pathways. Cellular responses that are mediated by these pathways are glycogen breakdown in the liver in response to vasopressin, smooth muscle contraction in response to acetylcholine, and thrombin-induced platelet aggregation.

PI 3-kinase (PI3K), which phosphorylates the D3 position of PI and its derivatives, has a central role in growth factor signal cascades involved in cell growth, differentiation, and metabolism. PI3K is a heterodimer consisting of an adapter subunit and a catalytic subunit. The adapter subunit acts as a scaffolding protein, interacting with specific tyrosine-phosphorylated proteins, lipid moieties, and other cytosolic factors. When the adapter subunit binds tyrosine phosphorylated targets, such as the insulin responsive substrate (IRS)-1, the catalytic subunit is activated and converts PI (4,5) bisphosphate (PIP₂) to PI (3,4,5) P₃ (PIP₃). PIP₃ then activates a number of other proteins, including PKA, protein kinase B (PKB), protein kinase C (PKC), glycogen synthase kinase (GSK)-3, and p70 ribosomal s6 kinase. PI3K also interacts directly with the cytoskeletal organizing proteins, Rac, rho, and cdc42 (Shepherd, P.R. et al. (1998) *Biochem. J.* 333:471-490). Animal models for diabetes, such as *obese* and *fat* mice, have altered PI3K adapter subunit levels. Specific mutations in the adapter subunit have also been found in an insulin-resistant Danish population, suggesting a role for PI3K in type-2 diabetes (Shepard, supra).

An example of lipid kinase phosphorylation activity is the phosphorylation of

WO 02/08399

PCT/US01/23092

D-erythro-sphingosine to the sphingolipid metabolite, sphingosine-1-phosphate (SPP). SPP has emerged as a novel lipid second-messenger with both extracellular and intracellular actions (Kohama, T. et al. (1998) *J. Biol. Chem.* 273:23722-23728). Extracellularly, SPP is a ligand for the G-protein coupled receptor EDG-1 (endothelial-derived, G-protein coupled receptor). Intracellularly, SPP regulates cell growth, survival, motility, and cytoskeletal changes. SPP levels are regulated by sphingosine kinases that specifically phosphorylate D-erythro-sphingosine to SPP. The importance of sphingosine kinase in cell signaling is indicated by the fact that various stimuli, including platelet-derived growth factor (PDGF), nerve growth factor, and activation of protein kinase C, increase cellular levels of SPP by activation of sphingosine kinase, and the fact that competitive inhibitors of the enzyme selectively inhibit cell proliferation induced by PDGF (Kohama et al., supra).

Purine Nucleotide Kinases

The purine nucleotide kinases, adenylate kinase (ATP:AMP phosphotransferase, or AdK) and guanylate kinase (ATP:GMP phosphotransferase, or GuK) play a key role in nucleotide metabolism and are crucial to the synthesis and regulation of cellular levels of ATP and GTP, respectively. These two molecules are precursors in DNA and RNA synthesis in growing cells and provide the primary source of biochemical energy in cells (ATP), and signal transduction pathways (GTP). Inhibition of various steps in the synthesis of these two molecules has been the basis of many antiproliferative drugs for cancer and antiviral therapy (Pillwein, K. et al. (1990) *Cancer Res.* 50:1576-1579).

AdK is found in almost all cell types and is especially abundant in cells having high rates of ATP synthesis and utilization such as skeletal muscle. In these cells AdK is physically associated with mitochondria and myofibrils, the subcellular structures that are involved in energy production and utilization, respectively. Recent studies have demonstrated a major function for AdK in transferring high energy phosphoryls from metabolic processes generating ATP to cellular components consuming ATP (Zeleznikar, R.J. et al. (1995) *J. Biol. Chem.* 270:7311-7319). Thus AdK may have a pivotal role in maintaining energy production in cells, particularly those having a high rate of growth or metabolism such as cancer cells, and may provide a target for suppression of its activity to treat certain cancers. Alternatively, reduced AdK activity may be a source of various metabolic, muscle-energy disorders that can result in cardiac or respiratory failure and may be treatable by increasing AdK activity.

GuK, in addition to providing a key step in the synthesis of GTP for RNA and DNA synthesis, also fulfills an essential function in signal transduction pathways of cells through the regulation of GDP and GTP. Specifically, GTP binding to membrane associated G proteins mediates the activation of cell receptors, subsequent intracellular activation of adenylyl cyclase, and production of the second

WO 02/08399

PCT/US01/23092

messenger, cyclic AMP. GDP binding to G proteins inhibits these processes. GDP and GTP levels also control the activity of certain oncogenic proteins such as p21^{ras} known to be involved in control of cell proliferation and oncogenesis (Bos, J.L. (1989) Cancer Res. 49:4682-4689). High ratios of GTP:GDP caused by suppression of GuK cause activation of p21^{ras} and promote oncogenesis.

- 5 Increasing GuK activity to increase levels of GDP and reduce the GTP:GDP ratio may provide a therapeutic strategy to reverse oncogenesis.

GuK is an important enzyme in the phosphorylation and activation of certain antiviral drugs useful in the treatment of herpes virus infections. These drugs include the guanine homologs acyclovir and bucidovir (Miller, W.H. and R.L. Miller (1980) J. Biol. Chem. 255:7204-7207; Stenberg, K. et al. 10 (1986) J. Biol. Chem. 261:2134-2139). Increasing GuK activity in infected cells may provide a therapeutic strategy for augmenting the effectiveness of these drugs and possibly for reducing the necessary dosages of the drugs.

Pyrimidine Kinases

The pyrimidine kinases are deoxycytidine kinase and thymidine kinase 1 and 2. Deoxycytidine 15 kinase is located in the nucleus, and thymidine kinase 1 and 2 are found in the cytosol (Johansson, M. et al. (1997) Proc. Natl. Acad. Sci. USA 94:11941-11945). Phosphorylation of deoxyribonucleosides by pyrimidine kinases provides an alternative pathway for de novo synthesis of DNA precursors. The role of pyrimidine kinases, like purine kinases, in phosphorylation is critical to the activation of several chemotherapeutically important nucleoside analogues (Arner E.S. and S. Eriksson (1995) Pharmacol. 20 Ther. 67:155-186).

The discovery of new human kinases, 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 cancer, immune disorders, disorders affecting growth and development, cardiovascular diseases, and lipid disorders, and in the assessment of the effects of exogenous compounds on the expression of 25 nucleic acid and amino acid sequences of human kinases.

SUMMARY OF THE INVENTION

The invention features purified polypeptides, human kinases, referred to collectively as "PKIN" and individually as "PKIN-1," "PKIN-2," "PKIN-3," "PKIN-4," "PKIN-5," "PKIN-6," 30 "PKIN-7," "PKIN-8," "PKIN-9," "PKIN-10," "PKIN-11," "PKIN-12," "PKIN-13," "PKIN-14," "PKIN-15," "PKIN-16," "PKIN-17," "PKIN-18," "PKIN-19," and "PKIN-20." 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:1-20, b) a

WO 02/08399

PCT/US01/23092

polypeptide comprising a naturally occurring amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-20, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-20, and d) an immunogenic fragment of a polypeptide having an amino acid sequence
5 selected from the group consisting of SEQ ID NO:1-20. In one alternative, the invention provides an isolated polypeptide comprising the amino acid sequence of SEQ ID NO:1-20.

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:1-20, b) a polypeptide comprising a naturally occurring amino acid sequence
10 at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-20, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-20, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-20. In one alternative, the polynucleotide encodes a polypeptide selected from the group consisting of SEQ ID NO:1-20. In
15 another alternative, the polynucleotide is selected from the group consisting of SEQ ID NO:21-40.

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:1-20, b) a polypeptide comprising a naturally occurring amino acid sequence at least
20 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-20, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-20, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-20. In one alternative, the invention provides a cell transformed with the recombinant polynucleotide. In another alternative, the
25 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:1-20, b) a polypeptide comprising a naturally occurring amino acid sequence at least
30 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-20, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-20, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-20. The method comprises a) culturing a cell under conditions suitable for expression of the polypeptide, wherein said cell is

WO 02/08399

PCT/US01/23092

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:1-20, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-20, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-20, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-20.

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:21-40, b) a polynucleotide comprising a naturally occurring polynucleotide sequence at least 90% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:21-40, 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:21-40, b) a polynucleotide comprising a naturally occurring polynucleotide sequence at least 90% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:21-40, 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

WO 02/08399

PCT/US01/23092

NO:21-40, b) a polynucleotide comprising a naturally occurring polynucleotide sequence at least 90% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:21-40, 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:1-20, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-20, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-20, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-20, and a pharmaceutically acceptable excipient. In one embodiment, the composition comprises an amino acid sequence selected from the group consisting of SEQ ID NO:1-20. The invention additionally provides a method of treating a disease or condition associated with decreased expression of functional PKIN, 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:1-20, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-20, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-20, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-20. 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 PKIN, 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

WO 02/08399

PCT/US01/23092

amino acid sequence selected from the group consisting of SEQ ID NO:1-20, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-20, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-20, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-20. 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 PKIN, 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:1-20, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-20, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-20, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-20. 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:1-20, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-20, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-20, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-20. 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

WO 02/08399

PCT/US01/23092

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 polynucleotide sequence selected from the group consisting of SEQ ID NO:21-40, 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:21-40, ii) a polynucleotide comprising a naturally occurring polynucleotide sequence at least 90% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:21-40, 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:21-40, ii) a polynucleotide comprising a naturally occurring polynucleotide sequence at least 90% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:21-40, 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

WO 02/08399

PCT/US01/23092

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.

5 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.

10 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

15 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.

20 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.

25 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
30 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.

WO 02/08399

PCT/US01/23092

DEFINITIONS

"PKIN" refers to the amino acid sequences of substantially purified PKIN 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.

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

An "allelic variant" is an alternative form of the gene encoding PKIN. 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
15 in a given sequence.

"Altered" nucleic acid sequences encoding PKIN include those sequences with deletions, insertions, or substitutions of different nucleotides, resulting in a polypeptide the same as PKIN or a polypeptide with at least one functional characteristic of PKIN. Included within this definition are polymorphisms which may or may not be readily detectable using a particular oligonucleotide probe of the polynucleotide encoding PKIN, and improper or unexpected hybridization to allelic variants, with a locus other than the normal chromosomal locus for the polynucleotide sequence encoding PKIN. 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 PKIN. 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 PKIN 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
20 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

WO 02/08399

PCT/US01/23092

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.

- 5 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 PKIN. Antagonists may include proteins such as antibodies, nucleic acids, carbohydrates, small molecules, or any other compound or composition which modulates the activity of PKIN either by
10 directly interacting with PKIN or by acting on components of the biological pathway in which PKIN 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 PKIN polypeptides can be prepared using intact polypeptides or using fragments
15 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
20 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
30 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

WO 02/08399

PCT/US01/23092

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 PKIN, 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 PKIN or fragments of PKIN 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

WO 02/08399

PCT/US01/23092

	Asn	Asp, Gln, His
	Asp	Asn, Glu
	Cys	Ala, Ser
	Gln	Asn, Glu, His
5	Glu	Asp, Gln, His
	Gly	Ala
	His	Asn, Arg, Gln, Glu
	Ile	Leu, Val
	Leu	Ile, Val
10	Lys	Arg, Gln, Glu
	Met	Leu, Ile
	Phe	His, Met, Leu, Trp, Tyr
	Ser	Cys, Thr
	Thr	Ser, Val
15	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.

"Exon shuffling" refers to the recombination of different coding regions (exons). Since an exon may represent a structural or functional domain of the encoded protein, new proteins may be assembled through the novel reassortment of stable substructures, thus allowing acceleration of the

WO 02/08399

PCT/US01/23092

evolution of new protein functions.

A "fragment" is a unique portion of PKIN or the polynucleotide encoding PKIN 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
5 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
10 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:21-40 comprises a region of unique polynucleotide sequence that
15 specifically identifies SEQ ID NO:21-40, for example, as distinct from any other sequence in the genome from which the fragment was obtained. A fragment of SEQ ID NO:21-40 is useful, for example, in hybridization and amplification technologies and in analogous methods that distinguish SEQ ID NO:21-40 from related polynucleotide sequences. The precise length of a fragment of SEQ ID NO:21-40 and the region of SEQ ID NO:21-40 to which the fragment corresponds are routinely
20 determinable by one of ordinary skill in the art based on the intended purpose for the fragment.

A fragment of SEQ ID NO:1-20 is encoded by a fragment of SEQ ID NO:21-40. A fragment of SEQ ID NO:1-20 comprises a region of unique amino acid sequence that specifically identifies SEQ ID NO:1-20. For example, a fragment of SEQ ID NO:1-20 is useful as an immunogenic peptide for the development of antibodies that specifically recognize SEQ ID NO:1-20.
25 The precise length of a fragment of SEQ ID NO:1-20 and the region of SEQ ID NO:1-20 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
30 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

WO 02/08399

PCT/US01/23092

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.

5 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 10 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 15 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 20 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/bi2.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 25 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

30 *Open Gap: 5 and Extension Gap: 2 penalties*

Gap x drop-off: 50

Expect: 10

Word Size: 11

WO 02/08399

PCT/US01/23092

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

WO 02/08399

PCT/US01/23092

*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)

WO 02/08399

PCT/US01/23092

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. 5 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, 10 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 15 sequences by virtue of the formation of hydrogen bonds between complementary bases. A hybridization complex may be formed in solution (e.g., C₀t or R₀t 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).

20 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 25 cellular and systemic defense systems.

An "immunogenic fragment" is a polypeptide or oligopeptide fragment of PKIN 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 PKIN which is useful in any of the antibody production methods disclosed herein or known in the art. 30

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.

WO 02/08399

PCT/US01/23092

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

5 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.

10 "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.

15 "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.

20 "Post-translational modification" of an PKIN 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 PKIN.

"Probe" refers to nucleic acid sequences encoding PKIN, 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. 25 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).

30 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

WO 02/08399

PCT/US01/23092

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

WO 02/08399

PCT/US01/23092

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
5 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 used to vaccinate a mammal wherein the recombinant nucleic acid is
10 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,
15 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
20 other moieties known in the art.

An "RNA equivalent," in reference to a DNA sequence, is composed of the same linear
25 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 PKIN, nucleic acids encoding PKIN, or fragments thereof may comprise a bodily fluid; an extract from a cell,
25 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
30 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.

WO 02/08399

PCT/US01/23092

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.

5 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,
10 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
15 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
20 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
25 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,
30 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.

WO 02/08399

PCT/US01/23092

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
5 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
10 polynucleotides due to alternate 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
15 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
20 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
25 identity over a certain defined length of one of the polypeptides.

THE INVENTION

The invention is based on the discovery of new human human kinases (PKIN), the polynucleotides encoding PKIN, and the use of these compositions for the diagnosis, treatment, or
30 prevention of cancer, immune disorders, disorders affecting growth and development, cardiovascular diseases, and lipid 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

WO 02/08399

PCT/US01/23092

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.

5 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.

15 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.

20 Together, Tables 2 and 3 summarize the properties of polypeptides of the invention, and these properties establish that the claimed polypeptides are human kinases. For example, SEQ ID NO:2 is 97% identical to mouse tousled-like kinase (GenBank ID g2853031) as determined by the Basic Local Alignment Search Tool (BLAST). (See Table 2.) The BLAST probability score is 0.0, which indicates the probability of obtaining the observed polypeptide sequence alignment by chance. SEQ ID NO:2 also contains an eukaryotic protein kinase 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 BLIMPS, MOTIFS, and PROFILESCAN analyses provide further corroborative evidence that SEQ ID NO:2 is a tousled-like kinase. In an alternative example, SEQ ID NO:10 is 63% identical to human serine/threonine protein kinase (GenBank ID g36615) as determined by the Basic Local Alignment Search Tool (BLAST). (See Table 2.) The

WO 02/08399

PCT/US01/23092

BLAST probability score is $7.7e-122$, which indicates the probability of obtaining the observed polypeptide sequence alignment by chance. SEQ ID NO:10 also contains an eukaryotic protein kinase domain as determined by searching for statistically significant matches in the hidden Markov model (HMM)-based PFAM database of conserved protein domains. (See Table 3.) Data from BLIMPS, MOTIFS, and PROFILESCAN analyses provide further corroborative evidence that SEQ ID NO:10 is a serine/threonine kinase. Note that "serine/threonine kinase" is a specific class of kinases. In an alternative example, SEQ ID NO:16 is 53% identical to human receptor protein-tyrosine kinase (GenBank ID g551608) as determined by the Basic Local Alignment Search Tool (BLAST). (See Table 2.) The BLAST probability score is $4.1e-290$, which indicates the probability of obtaining the observed polypeptide sequence alignment by chance. SEQ ID NO:16 also contains an eukaryotic protein kinase 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 BLIMPS and PROFILESCAN analyses provide further corroborative evidence that SEQ ID NO:16 is a receptor tyrosine kinase. In an alternative example, SEQ ID NO:19 is 93% identical to rat Calcium/calmodulin-dependent protein kinase isoform IV (GenBank ID g1836161) as determined by the Basic Local Alignment Search Tool (BLAST). (See Table 2.) The BLAST probability score is $6.0e-257$, which indicates the probability of obtaining the observed polypeptide sequence alignment by chance. SEQ ID NO:19 also contains an eukaryotic protein kinase domain as determined by searching for statistically significant matches in the hidden Markov model (HMM)-based PFAM database of conserved protein domains. (See Table 3.) Data from BLIMPS, MOTIFS, and PROFILESCAN analyses provide further corroborative evidence that SEQ ID NO:19 is a protein kinase. SEQ ID NO:1, SEQ ID NO:3-9, SEQ ID NO:11-15, SEQ ID NO:17-18, and SEQ ID NO:20 were analyzed and annotated in a similar manner. The algorithms and parameters for the analysis of SEQ ID NO:1-20 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:21-40 or that distinguish between SEQ ID NO:21-40 and related polynucleotide sequences. Column 5 shows identification numbers corresponding to cDNA

WO 02/08399

PCT/US01/23092

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, 2564295H1 is the identification number of an Incyte cDNA sequence, and ADRETUT01 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., 71191190V1). Alternatively, the identification numbers in column 5 may refer to GenBank cDNAs or ESTs (e.g., g1164223) which contributed to the assembly of the full length polynucleotide sequences. In addition, the identification numbers in column 5 may identify sequences derived from the ENSEMBL (The Sanger Centre, Cambridge, UK) database (*i.e.*, those sequences including the designation "ENST"). Alternatively, the identification numbers in column 5 may be derived from the NCBI RefSeq Nucleotide Sequence Records Database (*i.e.*, those sequences including the designation "NM" or "NT") or the NCBI RefSeq Protein Sequence Records (*i.e.*, those sequences including the designation "NP"). 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. For example, FL_XXXXX_N₁_N₂_YYYY_N₃_N₄ represents a "stitched" sequence in which XXXXX is the identification number of the cluster of sequences to which the algorithm was applied, and YYYY is the number of the prediction generated by the algorithm, and N_{1,2,3,4}, if present, represent specific exons that may have been manually edited during analysis (See Example V). Alternatively, the identification numbers in column 5 may refer to assemblages of exons brought together by an "exon-stretching" algorithm. For example, FLXXXXX_gAAAAA_gBBBBB_1_N is the identification number of a "stretched" sequence, with XXXXX being the Incyte project identification number, gAAAAA being the GenBank identification number of the human genomic sequence to which the "exon-stretching" algorithm was applied, gBBBBB being the GenBank identification number or NCBI RefSeq identification number of the nearest GenBank protein homolog, and N referring to specific exons (See Example V). In instances where a RefSeq sequence was used as a protein homolog for the "exon-stretching" algorithm, a RefSeq identifier (denoted by "NM," "NP," or "NT") may be used in place of the GenBank identifier (*i.e.*, gBBBBB).

Alternatively, a prefix identifies component sequences that were hand-edited, predicted from

WO 02/08399

PCT/US01/23092

genomic DNA sequences, or derived from a combination of sequence analysis methods. The following Table lists examples of component sequence prefixes and corresponding sequence analysis methods associated with the prefixes (see Example IV and Example V).

Prefix	Type of analysis and/or examples of programs
GNN, GFG, ENST	Exon prediction from genomic sequences using, for example, GENSCAN (Stanford University, CA, USA) or FGENES (Computer Genomics Group, The Sanger Centre, Cambridge, UK).
GBI	Hand-edited analysis of genomic sequences.
FL	Stitched or stretched genomic sequences (see Example V).
INCY	Full length transcript and exon prediction from mapping of EST sequences to the genome. Genomic location and EST composition data are combined to predict the exons and resulting transcript.

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 PKIN variants. A preferred PKIN 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 PKIN amino acid sequence, and which contains at least one functional or structural characteristic of PKIN.

The invention also encompasses polynucleotides which encode PKIN. In a particular embodiment, the invention encompasses a polynucleotide sequence comprising a sequence selected from the group consisting of SEQ ID NO:21-40, which encodes PKIN. The polynucleotide sequences of SEQ ID NO:21-40, 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 PKIN. In particular, such a variant polynucleotide sequence will have at least about 70%, or alternatively at least

WO 02/08399

PCT/US01/23092

about 85%, or even at least about 95% polynucleotide sequence identity to the polynucleotide sequence encoding PKIN. 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:21-40 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:21-40. Any one of the polynucleotide variants described above can encode an amino acid sequence which contains at least one functional or structural characteristic of PKIN.

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 PKIN, 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 PKIN, and all such variations are to be considered as being specifically disclosed.

Although nucleotide sequences which encode PKIN and its variants are generally capable of hybridizing to the nucleotide sequence of the naturally occurring PKIN under appropriately selected conditions of stringency, it may be advantageous to produce nucleotide sequences encoding PKIN 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 PKIN 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 PKIN and PKIN 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 PKIN 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:21-40 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-

WO 02/08399

PCT/US01/23092

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 PKIN 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

WO 02/08399

PCT/US01/23092

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
5 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
10 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
15 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 PKIN may be cloned in recombinant DNA molecules that direct expression of PKIN, or
20 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 PKIN.

The nucleotide sequences of the present invention can be engineered using methods generally known in the art in order to alter PKIN-encoding sequences for a variety of purposes including, but
25 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
30 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 Cramer, A. et al. (1996) Nat. Biotechnol. 14:315-319) to alter or improve

WO 02/08399

PCT/US01/23092

the biological properties of PKIN, 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 PKIN 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, PKIN 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 PKIN, 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., Chiez, 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 PKIN, the nucleotide sequences encoding PKIN 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 PKIN. Such elements may vary in their strength and specificity. Specific initiation signals may also be used to achieve more efficient translation of sequences encoding PKIN. Such signals

WO 02/08399

PCT/US01/23092

include the ATG initiation codon and adjacent sequences, e.g. the Kozak sequence. In cases where sequences encoding PKIN 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,

5 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.)

10 Methods which are well known to those skilled in the art may be used to construct expression vectors containing sequences encoding PKIN 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.)

15 A variety of expression vector/host systems may be utilized to contain and express sequences encoding PKIN. 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); 20 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 25 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, 30 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.

WO 02/08399

PCT/US01/23092

In bacterial systems, a number of cloning and expression vectors may be selected depending upon the use intended for polynucleotide sequences encoding PKIN. For example, routine cloning, subcloning, and propagation of polynucleotide sequences encoding PKIN 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 PKIN 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 PKIN are needed, e.g. for the production of antibodies, vectors which direct high level expression of PKIN 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 PKIN. 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 PKIN. Transcription of sequences encoding PKIN 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; Brögile, 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 PKIN 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 PKIN in host cells. (See, e.g., Logan, J. and T. Shenk (1984) *Proc.*

WO 02/08399

PCT/US01/23092

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 PKIN in cell lines is preferred. For example, sequences encoding PKIN 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 *ap^r* 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 chloresulfuron 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

WO 02/08399

PCT/US01/23092

is also present, the presence and expression of the gene may need to be confirmed. For example, if the sequence encoding PKIN is inserted within a marker gene sequence, transformed cells containing sequences encoding PKIN can be identified by the absence of marker gene function. Alternatively, a marker gene can be placed in tandem with a sequence encoding PKIN 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 PKIN and that express PKIN 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 PKIN 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 PKIN 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 PKIN include oligolabeling, nick translation, end-labeling, or PCR amplification using a labeled nucleotide. Alternatively, the sequences encoding PKIN, 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.

WO 02/08399

PCT/US01/23092

Host cells transformed with nucleotide sequences encoding PKIN 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 PKIN may be designed to contain signal sequences which direct secretion of PKIN 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 PKIN 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 PKIN protein containing a heterologous moiety that can be recognized by a commercially available antibody may facilitate the screening of peptide libraries for inhibitors of PKIN 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 PKIN encoding sequence and the heterologous protein sequence, so that PKIN 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 PKIN may be achieved in

WO 02/08399

PCT/US01/23092

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, ³⁵S-methionine.

5 PKIN of the present invention or fragments thereof may be used to screen for compounds that specifically bind to PKIN. At least one and up to a plurality of test compounds may be screened for specific binding to PKIN. 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
10 PKIN, 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 PKIN 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
15 these compounds involves producing appropriate cells which express PKIN, either as a secreted protein or on the cell membrane. Preferred cells include cells from mammals, yeast, Drosophila, or E. coli. Cells expressing PKIN or cell membrane fractions which contain PKIN are then contacted with a test compound and binding, stimulation, or inhibition of activity of either PKIN or the compound is analyzed.

20 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 PKIN, either in solution or affixed to a solid support, and detecting the binding of PKIN to the compound. Alternatively, the assay may detect or measure binding of a test compound in the presence of a labeled competitor.

25 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.

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

WO 02/08399

PCT/US01/23092

in vitro or cell-free system comprising PKIN under conditions suitable for PKIN activity, and the assay is performed. In either of these assays, a test compound which modulates the activity of PKIN 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.

5 In another embodiment, polynucleotides encoding PKIN 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
10 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
15 (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
20 therapeutic or toxic agents.

Polynucleotides encoding PKIN 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.
25 (1998) Science 282:1145-1147).

Polynucleotides encoding PKIN 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 PKIN 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
30 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 PKIN, e.g., by secreting PKIN in its milk, may also serve as a convenient source of that protein (Janne, J. et al. (1998) Biotechnol. Annu. Rev. 4:55-74).

WO 02/08399

PCT/US01/23092

THERAPEUTICS

Chemical and structural similarity, e.g., in the context of sequences and motifs, exists between regions of PKIN and human kinases. In addition, the expression of PKIN is closely associated with bladder cancer, prostatic, ovarian, brain, colon, ileum, penis, skin, adrenal tumor, digestive, and cancerous tissues. Therefore, PKIN appears to play a role in cancer, immune disorders, disorders affecting growth and development, cardiovascular diseases, and lipid disorders. In the treatment of disorders associated with increased PKIN expression or activity, it is desirable to decrease the expression or activity of PKIN. In the treatment of disorders associated with decreased PKIN expression or activity, it is desirable to increase the expression or activity of PKIN.

Therefore, in one embodiment, PKIN 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 PKIN. Examples of such disorders include, but are not limited to, a cancer, such as adenocarcinoma, leukemia, lymphoma, melanoma, myeloma, sarcoma, teratocarcinoma, and, in particular, cancers 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, leukemias such as multiple myeloma and lymphomas such as Hodgkin's disease; an immune 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, hyper eosinophilia, 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, complications of cancer, hemodialysis, and extracorporeal circulation, viral, bacterial, fungal, parasitic, protozoal, and helminthic infections, and trauma; a growth and developmental 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

WO 02/08399

PCT/US01/23092

particular, cancers 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, renal tubular acidosis, anemia, Cushing's syndrome, achondroplastic dwarfism, Duchenne and Becker muscular dystrophy, epilepsy,

5 gonadal dysgenesis, WAGR syndrome (Wilms' tumor, aniridia, genitourinary abnormalities, and mental retardation), Smith-Magenis syndrome, myelodysplastic syndrome, hereditary mucoepithelial dysplasia, hereditary keratodermas, hereditary neuropathies such as Charcot-Marie-Tooth disease and neurofibromatosis, hypothyroidism, hydrocephalus, seizure disorders such as Sydenham's chorea and cerebral palsy, spina bifida, anencephaly, craniorachischisis, congenital glaucoma, cataract, and

10 sensorineural hearing loss; a cardiovascular disease, such as arteriovenous fistula, atherosclerosis, hypertension, vasculitis, Raynaud's disease, aneurysms, arterial dissections, varicose veins, thrombophlebitis and phlebothrombosis, vascular tumors, and complications of thrombolysis, balloon angioplasty, vascular replacement, and coronary artery bypass graft surgery, congestive heart failure, ischemic heart disease, angina pectoris, myocardial infarction, hypertensive heart disease,

15 degenerative valvular heart disease, calcific aortic valve stenosis, congenitally bicuspid aortic valve, mitral annular calcification, mitral valve prolapse, rheumatic fever and rheumatic heart disease, infective endocarditis, nonbacterial thrombotic endocarditis, endocarditis of systemic lupus erythematosus, carcinoid heart disease, cardiomyopathy, myocarditis, pericarditis, neoplastic heart disease, congenital heart disease, and complications of cardiac transplantation, congenital lung

20 anomalies, atelectasis, pulmonary congestion and edema, pulmonary embolism, pulmonary hemorrhage, pulmonary infarction, pulmonary hypertension, vascular sclerosis, obstructive pulmonary disease, restrictive pulmonary disease, chronic obstructive pulmonary disease, emphysema, chronic bronchitis, bronchial asthma, bronchiectasis, bacterial pneumonia, viral and mycoplasmal pneumonia, lung abscess, pulmonary tuberculosis, diffuse interstitial diseases, pneumoconioses, sarcoidosis, idiopathic

25 pulmonary fibrosis, desquamative interstitial pneumonitis, hypersensitivity pneumonitis, pulmonary eosinophilia bronchiolitis obliterans-organizing pneumonia, diffuse pulmonary hemorrhage syndromes, Goodpasture's syndromes, idiopathic pulmonary hemosiderosis, pulmonary involvement in collagen-vascular disorders, pulmonary alveolar proteinosis, lung tumors, inflammatory and noninflammatory pleural effusions, pneumothorax, pleural tumors, drug-induced lung disease, radiation-

30 induced lung disease, and complications of lung transplantation; and a lipid disorder such as fatty liver, cholestasis, primary biliary cirrhosis, carnitine deficiency, carnitine palmitoyltransferase deficiency, myoadenylate deaminase deficiency, hypertriglyceridemia, lipid storage disorders such as Fabry's disease, Gaucher's disease, Niemann-Pick's disease, metachromatic leukodystrophy,

WO 02/08399

PCT/US01/23092

adrenoleukodystrophy, GM₂ gangliosidosis, and ceroid lipofuscinosis, abetalipoproteinemia, Tangier disease, hyperlipoproteinemia, diabetes mellitus, lipodystrophy, lipomatosis, acute panniculitis, disseminated fat necrosis, adiposis dolorosa, lipid adrenal hyperplasia, minimal change disease, lipomas, atherosclerosis, hypercholesterolemia, hypercholesterolemia with hypertriglyceridemia, 5 primary hypolipoproteinemia, hypothyroidism, renal disease, liver disease, lecithin:cholesterol acyltransferase deficiency, cerebrotendinous xanthomatosis, sitosterolemia, hypocholesterolemia, Tay-Sachs disease, Sandhoff's disease, hyperlipidemia, hyperlipemia, lipid myopathies, and obesity.

In another embodiment, a vector capable of expressing PKIN or a fragment or derivative thereof may be administered to a subject to treat or prevent a disorder associated with decreased 10 expression or activity of PKIN including, but not limited to, those described above.

In a further embodiment, a composition comprising a substantially purified PKIN 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 PKIN including, but not limited to, those 15 provided above.

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

In a further embodiment, an antagonist of PKIN may be administered to a subject to treat or prevent a disorder associated with increased expression or activity of PKIN. Examples of such 20 disorders include, but are not limited to, those cancer, immune disorders, disorders affecting growth and development, cardiovascular diseases, and lipid disorders described above. In one aspect, an antibody which specifically binds PKIN 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 PKIN.

In an additional embodiment, a vector expressing the complement of the polynucleotide encoding PKIN may be administered to a subject to treat or prevent a disorder associated with 25 increased expression or activity of PKIN 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 30 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

WO 02/08399

PCT/US01/23092

efficacy with lower dosages of each agent, thus reducing the potential for adverse side effects.

An antagonist of PKIN may be produced using methods which are generally known in the art. In particular, purified PKIN may be used to produce antibodies or to screen libraries of pharmaceutical agents to identify those which specifically bind PKIN. Antibodies to PKIN 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 PKIN 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 PKIN 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 PKIN 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 PKIN 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 PKIN-specific single

WO 02/08399

PCT/US01/23092

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 PKIN 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 PKIN and its specific antibody. A two-site, monoclonal-based immunoassay utilizing monoclonal antibodies reactive to two non-interfering PKIN 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 PKIN. Affinity is expressed as an association constant, K_a , which is defined as the molar concentration of PKIN-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 PKIN epitopes, represents the average affinity, or avidity, of the antibodies for PKIN. The K_a determined for a preparation of monoclonal antibodies, which are monospecific for a particular PKIN 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 PKIN-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 PKIN, preferably in active form, from the antibody (Catty, D. (1988) *Antibodies, Volume I: A Practical Approach*, IRL Press, Washington DC; Liddell,

WO 02/08399

PCT/US01/23092

J.E. and A. Cryer (1991) A Practical Guide to Monoclonal Antibodies, John Wiley & Sons, New York NY).

5 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 PKIN-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.)

10 In another embodiment of the invention, the polynucleotides encoding PKIN, 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 PKIN. Such technology is well known in the art, and antisense oligonucleotides or larger fragments
15 can be designed from various locations along the coding or control regions of sequences encoding PKIN. (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
20 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 Clin. 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
25 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 PKIN may be used for
30 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

WO 02/08399

PCT/US01/23092

- (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), thalassemias, 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 PKIN expression or regulation causes disease, the expression of PKIN from an appropriate population of transduced cells may alleviate the clinical manifestations caused by the genetic deficiency.
- 15 In a further embodiment of the invention, diseases or disorders caused by deficiencies in PKIN are treated by constructing mammalian expression vectors encoding PKIN and introducing these vectors by mechanical means into PKIN-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
- 20 (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 PKIN include, but are not limited to, the pCDNA 3.1, EPITAG, PRCCMV2, PREP, PVAX vectors (Invitrogen, Carlsbad CA),
- 25 PCMV-SCRIPT, PCMV-TAG, PEGSH/PERV (Stratagene, La Jolla CA), and PTET-OFF, PTET-ON, PTRE2, PTRE2-LUC, PTK-HYG (Clontech, Palo Alto CA). PKIN 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 promoter (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
- 30 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.

WO 02/08399

PCT/US01/23092

and Blau, H.M. *supra*)), or (iii) a tissue-specific promoter or the native promoter of the endogenous gene encoding PKIN 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
5 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.

10 In another embodiment of the invention, diseases or disorders caused by genetic defects with respect to PKIN expression are treated by constructing a retrovirus vector consisting of (i) the polynucleotide encoding PKIN 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
15 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.

25 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).

30 In the alternative, an adenovirus-based gene therapy delivery system is used to deliver polynucleotides encoding PKIN to cells which have one or more genetic abnormalities with respect to the expression of PKIN. 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

WO 02/08399

PCT/US01/23092

be versatile for importing 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 PKIN to target cells which have one or more genetic abnormalities with respect to the expression of PKIN. The use of herpes simplex virus (HSV)-based vectors may be especially valuable for introducing PKIN 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 1-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 PKIN 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 PKIN into the alphavirus

WO 02/08399

PCT/US01/23092

genome in place of the capsid-coding region results in the production of a large number of PKIN-coding RNAs and the synthesis of high levels of PKIN 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 PKIN 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 PKIN.

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

WO 02/08399

PCT/US01/23092

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 PKIN. 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 PKIN. 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 PKIN expression or activity, a compound which specifically inhibits expression of the polynucleotide encoding PKIN may be therapeutically useful, and in the treatment of disorders associated with decreased PKIN expression or activity, a compound which specifically promotes expression of the polynucleotide encoding PKIN 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 PKIN is exposed to at least one test compound thus obtained. The sample

WO 02/08399

PCT/US01/23092

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 PKIN 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 PKIN. 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 (Bruce, T.W. et al. (1997) U.S. Patent No. 5,686,242; Bruce, 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 PKIN, antibodies to PKIN, and mimetics, agonists, antagonists, or inhibitors of PKIN.

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,

WO 02/08399

PCT/US01/23092

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 PKIN or fragments thereof. For example, liposome preparations containing a cell-impermeable macromolecule may promote cell fusion and intracellular delivery of the macromolecule. Alternatively, PKIN 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 PKIN or fragments thereof, antibodies of PKIN, and agonists, antagonists or inhibitors of PKIN, 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 ED₅₀ (the dose therapeutically effective in 50% of the population) or LD₅₀ (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 LD₅₀/ED₅₀ ratio. Compositions which exhibit large therapeutic indices are preferred. The data obtained from cell culture assays and animal studies are

WO 02/08399

PCT/US01/23092

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 ED₅₀ 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.

5 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
10 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.

15 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 PKIN may be used for the
20 diagnosis of disorders characterized by expression of PKIN, or in assays to monitor patients being treated with PKIN or agonists, antagonists, or inhibitors of PKIN. Antibodies useful for diagnostic purposes may be prepared in the same manner as described above for therapeutics. Diagnostic assays for PKIN include methods which utilize the antibody and a label to detect PKIN in human body fluids or in extracts of cells or tissues. The antibodies may be used with or without modification, and
25 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 PKIN, including ELISAs, RIAs, and FACS, are known in the art and provide a basis for diagnosing altered or abnormal levels of PKIN expression. Normal or standard values for PKIN expression are established by combining body fluids or cell extracts taken
30 from normal mammalian subjects, for example, human subjects, with antibodies to PKIN under conditions suitable for complex formation. The amount of standard complex formation may be quantitated by various methods, such as photometric means. Quantities of PKIN expressed in subject, control, and disease samples from biopsied tissues are compared with the standard values. Deviation

WO 02/08399

PCT/US01/23092

between standard and subject values establishes the parameters for diagnosing disease.

In another embodiment of the invention, the polynucleotides encoding PKIN 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 PKIN may be correlated with disease. The diagnostic assay may be used to determine absence, presence, and excess expression of PKIN, and to monitor regulation of PKIN levels during therapeutic intervention.

In one aspect, hybridization with PCR probes which are capable of detecting polynucleotide sequences, including genomic sequences, encoding PKIN or closely related molecules may be used to identify nucleic acid sequences which encode PKIN. 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 PKIN, 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 PKIN 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:21-40 or from genomic sequences including promoters, enhancers, and introns of the PKIN gene.

Means for producing specific hybridization probes for DNAs encoding PKIN include the cloning of polynucleotide sequences encoding PKIN or PKIN 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 ³²P or ³⁵S, or by enzymatic labels, such as alkaline phosphatase coupled to the probe via avidin/biotin coupling systems, and the like.

Polynucleotide sequences encoding PKIN may be used for the diagnosis of disorders associated with expression of PKIN. Examples of such disorders include, but are not limited to, a cancer, such as adenocarcinoma, leukemia, lymphoma, melanoma, myeloma, sarcoma, teratocarcinoma, and, in particular, cancers 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, leukemias such as multiple myeloma and lymphomas such as Hodgkin's disease; an immune disorder, such as acquired immunodeficiency syndrome (AIDS), Addison's disease, adult respiratory distress

WO 02/08399

PCT/US01/23092

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 lymphocytotoxicity, 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, complications of cancer, hemodialysis, and extracorporeal circulation, viral, bacterial, fungal, parasitic, protozoal, and helminthic infections, and trauma; a growth and developmental 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, cancers 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, renal tubular acidosis, anemia, Cushing's syndrome, achondroplastic dwarfism, Duchenne and Becker muscular dystrophy, epilepsy, gonadal dysgenesis, WAGR syndrome (Wilms' tumor, aniridia, genitourinary abnormalities, and mental retardation), Smith-Magenis syndrome, myelodysplastic syndrome, hereditary mucoepithelial dysplasia, hereditary keratodermas, hereditary neuropathies such as Charcot-Marie-Tooth disease and neurofibromatosis, hypothyroidism, hydrocephalus, seizure disorders such as Sydenham's chorea and cerebral palsy, spina bifida, anencephaly, craniorachischisis, congenital glaucoma, cataract, and sensorineural hearing loss; a cardiovascular disease, such as arteriovenous fistula, atherosclerosis, hypertension, vasculitis, Raynaud's disease, aneurysms, arterial dissections, varicose veins, thrombophlebitis and phlebothrombosis, vascular tumors, and complications of thrombolysis, balloon angioplasty, vascular replacement, and coronary artery bypass graft surgery, congestive heart failure, ischemic heart disease, angina pectoris, myocardial infarction, hypertensive heart disease, degenerative valvular heart disease, calcific aortic valve stenosis, congenitally bicuspid aortic valve, mitral annular calcification, mitral valve prolapse, rheumatic fever and rheumatic heart disease, infective endocarditis, nonbacterial thrombotic endocarditis, endocarditis of systemic lupus erythematosus, carcinoid heart disease, cardiomyopathy,

WO 02/08399

PCT/US01/23092

myocarditis, pericarditis, neoplastic heart disease, congenital heart disease, and complications of cardiac transplantation, congenital lung anomalies, atelectasis, pulmonary congestion and edema, pulmonary embolism, pulmonary hemorrhage, pulmonary infarction, pulmonary hypertension, vascular sclerosis, obstructive pulmonary disease, restrictive pulmonary disease, chronic obstructive pulmonary disease, emphysema, chronic bronchitis, bronchial asthma, bronchiectasis, bacterial pneumonia, viral and mycoplasmal pneumonia, lung abscess, pulmonary tuberculosis, diffuse interstitial diseases, pneumoconioses, sarcoidosis, idiopathic pulmonary fibrosis, desquamative interstitial pneumonitis, hypersensitivity pneumonitis, pulmonary eosinophilia bronchiolitis obliterans-organizing pneumonia, diffuse pulmonary hemorrhage syndromes, Goodpasture's syndromes, idiopathic pulmonary

10 hemosiderosis, pulmonary involvement in collagen-vascular disorders, pulmonary alveolar proteinosis, lung tumors, inflammatory and noninflammatory pleural effusions, pneumothorax, pleural tumors, drug-induced lung disease, radiation-induced lung disease, and complications of lung transplantation; and a lipid disorder such as fatty liver, cholestasis, primary biliary cirrhosis, carnitine deficiency, carnitine palmitoyltransferase deficiency, myoadenylate deaminase deficiency, hypertriglyceridemia, lipid

15 storage disorders such Fabry's disease, Gaucher's disease, Niemann-Pick's disease, metachromatic leukodystrophy, adrenoleukodystrophy, GM₂ gangliosidosis, and ceroid lipofuscinosis, abetalipoproteinemia, Tangier disease, hyperlipoproteinemia, diabetes mellitus, lipodystrophy, lipomatoses, acute panniculitis, disseminated fat necrosis, adiposis dolorosa, lipoid adrenal hyperplasia, minimal change disease, lipomas, atherosclerosis, hypercholesterolemia, hypercholesterolemia with

20 hypertriglyceridemia, primary hypolipoproteinemia, hypothyroidism, renal disease, liver disease, lecithin:cholesterol acyltransferase deficiency, cerebrotendinous xanthomatosis, sitosterolemia, hypocholesterolemia, Tay-Sachs disease, Sandhoff's disease, hyperlipidemia, hyperlipemia, lipid myopathies, and obesity. The polynucleotide sequences encoding PKIN may be used in Southern or northern analysis, dot blot, or other membrane-based technologies; in PCR technologies; in dipstick,

25 pin, and multiformat ELISA-like assays; and in microarrays utilizing fluids or tissues from patients to detect altered PKIN expression. Such qualitative or quantitative methods are well known in the art.

In a particular aspect, the nucleotide sequences encoding PKIN may be useful in assays that detect the presence of associated disorders, particularly those mentioned above. The nucleotide sequences encoding PKIN may be labeled by standard methods and added to a fluid or tissue sample

30 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 PKIN in the sample

WO 02/08399

PCT/US01/23092

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 PKIN, 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 PKIN, 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 PKIN 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 PKIN, or a fragment of a polynucleotide complementary to the polynucleotide encoding PKIN, 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 PKIN 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

WO 02/08399

PCT/US01/23092

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 PKIN 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 fSSCP, 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 PKIN 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; Duplax, 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

WO 02/08399

PCT/US01/23092

profile.

In another embodiment, PKIN, fragments of PKIN, or antibodies specific for PKIN 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.

5 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
10 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
15 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
20 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)
25 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
30 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

WO 02/08399

PCT/US01/23092

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
5 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
10 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
15 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
20 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
25 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
30 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.

WO 02/08399

PCT/US01/23092

A proteomic profile may also be generated using antibodies specific for PKIN to quantify the levels of PKIN 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; Mendozze, 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*

WO 02/08399

PCT/US01/23092

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 PKIN 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 PKIN 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,

WO 02/08399

PCT/US01/23092

inversion, etc., among normal, carrier, or affected individuals.

In another embodiment of the invention, PKIN, 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 PKIN 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 PKIN, or fragments thereof, and washed. Bound PKIN is then detected by methods well known in the art. Purified PKIN 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 PKIN specifically compete with a test compound for binding PKIN. In this manner, antibodies can be used to detect the presence of any peptide which shares one or more antigenic determinants with PKIN.

In additional embodiments, the nucleotide sequences which encode PKIN 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 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, including U.S. Ser. No. 60/220,038, U.S. Ser. No. 60/222,112, U.S. Ser. No. 60/222,831, and U.S. Ser. No. 60/224,729 are hereby expressly incorporated by reference.

30

EXAMPLES

I. Construction of cDNA Libraries

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

WO 02/08399

PCT/US01/23092

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
5 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
10 (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
15 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-
20 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
25 CA), or derivatives thereof. Recombinant plasmids were transformed into competent *E. coli* cells including XL1-Blue, XL1-BlueMRF, or SOLR from Stratagene or DH5 α , 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*
30 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

WO 02/08399

PCT/US01/23092

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.

WO 02/08399

PCT/US01/23092

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 Conseq, 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:21-40. 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 human kinases were initially identified by running the Genscan gene identification program against public genomic sequence databases (e.g., gbpr1 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

WO 02/08399

PCT/US01/23092

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 human kinases, the encoded polypeptides were analyzed by querying against

5 PFAM models for human kinases. Potential human kinases were also identified by homology to Incyte cDNA sequences that had been annotated as human kinases. 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

10 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

15 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

20 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

25 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

30 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

WO 02/08399

PCT/US01/23092

type (cDNA to genomic sequence). The resultant stitched sequences were translated and compared by BLAST analysis to the genpept and gbpr 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.

5 **“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.

15 **VI. Chromosomal Mapping of PKIN Encoding Polynucleotides**

The sequences which were used to assemble SEQ ID NO:21-40 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:21-40 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

WO 02/08399

PCT/US01/23092

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.

5 **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.)

10 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:

15

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

The product score takes into account both the degree of similarity between two sequences and the
 20 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
 25 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
 30 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 PKIN are analyzed with respect to the tissue sources from which they were derived. For example, some full length sequences are

WO 02/08399

PCT/US01/23092

assembled, at least in part, with overlapping Incyte cDNA sequences (see Example III). 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 PKIN. cDNA sequences and cDNA library/tissue information are found in the LIFSEQ GOLD database (Incyte Genomics, Palo Alto CA).

VIII. Extension of PKIN 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

WO 02/08399

PCT/US01/23092

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 CviII cholera virus endonuclease (Molecular Biology Research, Madison WI), and sonicated or sheared prior to religation into pUC 18 vector (Amersham Pharmacia Biotech). For shotgun sequencing, the digested 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% dimethylsulfoxide (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:21-40 are employed to screen cDNAs,

WO 02/08399

PCT/US01/23092

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 μ Ci of $[\gamma\text{-}^{32}\text{P}]$ 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 (Schna (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., Schna, 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.

WO 02/08399

PCT/US01/23092

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/ μ l oligo-(dT) primer (21mer), 1X first strand buffer, 0.03 units/ μ l RNase inhibitor, 500 μ M dATP, 500 μ M dGTP, 500 μ M dTTP, 40 μ M dCTP, 40 μ M 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 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 μ l 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 μ g. 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

WO 02/08399

PCT/US01/23092

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 μ l of the array element DNA, at an average concentration of 100 ng/ μ l, 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 μ l of sample mixture consisting of 0.2 μ g 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 cm² 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 μ l 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,

WO 02/08399

PCT/US01/23092

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 PKIN-encoding sequences, or any parts thereof, are used to detect, decrease, or inhibit expression of naturally occurring PKIN. 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 PKIN. 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 PKIN-encoding transcript.

XII. Expression of PKIN

Expression and purification of PKIN is achieved using bacterial or virus-based expression systems. For expression of PKIN in bacteria, cDNA is subcloned into an appropriate vector containing an antibiotic resistance gene and an inducible promoter that directs high levels of cDNA

WO 02/08399

PCT/US01/23092

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 PKIN upon induction with isopropyl beta-D-thiogalactopyranoside (IPTG). Expression of PKIN 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 PKIN 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, PKIN 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 PKIN 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 PKIN obtained by these methods can be used directly in the assays shown in Examples XVI, XVII, XVIII, and XIX where applicable.

XIII. Functional Assays

PKIN function is assessed by expressing the sequences encoding PKIN 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 μ g 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 μ g of an additional plasmid containing sequences encoding a

WO 02/08399

PCT/US01/23092

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 PKIN on gene expression can be assessed using highly purified populations of cells transfected with sequences encoding PKIN 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 PKIN and other genes of interest can be analyzed by northern analysis or microarray techniques.

XIV. Production of PKIN Specific Antibodies

PKIN 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 PKIN 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-

WO 02/08399

PCT/US01/23092

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-PKIN activity by, for example, binding the peptide or PKIN 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 PKIN Using Specific Antibodies

Naturally occurring or recombinant PKIN is substantially purified by immunoaffinity chromatography using antibodies specific for PKIN. An immunoaffinity column is constructed by covalently coupling anti-PKIN 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 PKIN are passed over the immunoaffinity column, and the column is washed under conditions that allow the preferential absorbance of PKIN (e.g., high ionic strength buffers in the presence of detergent). The column is eluted under conditions that disrupt antibody/PKIN 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 PKIN is collected.

XVI. Identification of Molecules Which Interact with PKIN

PKIN, or biologically active fragments thereof, are labeled with ¹²⁵I 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 PKIN, washed, and any wells with labeled PKIN complex are assayed. Data obtained using different concentrations of PKIN are used to calculate values for the number, affinity, and association of PKIN with the candidate molecules.

Alternatively, molecules interacting with PKIN 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).

PKIN 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 PKIN Activity

Generally, protein kinase activity is measured by quantifying the phosphorylation of a protein

WO 02/08399

PCT/US01/23092

substrate by PKIN in the presence of [γ - 32 P]ATP. PKIN is incubated with the protein substrate, 32 P-ATP, and an appropriate kinase buffer. The 32 P incorporated into the substrate is separated from free 32 P-ATP by electrophoresis and the incorporated 32 P is counted using a radioisotope counter. The amount of incorporated 32 P is proportional to the activity of PKIN. A determination of the specific amino acid residue phosphorylated is made by phosphoamino acid analysis of the hydrolyzed protein.

In one alternative, protein kinase activity is measured by quantifying the transfer of gamma phosphate from adenosine triphosphate (ATP) to a serine, threonine or tyrosine residue in a protein substrate. The reaction occurs between a protein kinase sample with a biotinylated peptide substrate and gamma 32 P-ATP. Following the reaction, free avidin in solution is added for binding to the biotinylated 32 P-peptide product. The binding sample then undergoes a centrifugal ultrafiltration process with a membrane which will retain the product-avidin complex and allow passage of free gamma 32 P-ATP. The reservoir of the centrifuged unit containing the 32 P-peptide product as retentate is then counted in a scintillation counter. This procedure allows assay of any type of protein kinase sample, depending on the peptide substrate and kinase reaction buffer selected. This assay is provided in kit form (ASUA, Affinity Ultrafiltration Separation Assay, Transbio Corporation, Baltimore MD, U.S. Patent No. 5,869,275). Suggested substrates and their respective enzymes include but are not limited to: Histone H1 (Sigma) and p34^{cdc2}kinase, Annexin I, Angiotensin (Sigma) and EGF receptor kinase, Annexin II and *src* kinase, ERK1 & ERK2 substrates and MEK, and myelin basic protein and ERK (Pearson, J.D. et al. (1991) *Methods Enzymol.* 200:62-81).

In another alternative, protein kinase activity of PKIN is demonstrated in an assay containing PKIN, 50 μ l of kinase buffer, 1 μ g substrate, such as myelin basic protein (MBP) or synthetic peptide substrates, 1 mM DTT, 10 μ g ATP, and 0.5 μ Ci [γ - 32 P]ATP. The reaction is incubated at 30°C for 30 minutes and stopped by pipetting onto P81 paper. The unincorporated [γ - 32 P]ATP is removed by washing and the incorporated radioactivity is measured using a scintillation counter. Alternatively, the reaction is stopped by heating to 100°C in the presence of SDS loading buffer and resolved on a 12% SDS polyacrylamide gel followed by autoradiography. The amount of incorporated 32 P is proportional to the activity of PKIN.

In yet another alternative, adenylate kinase or guanylate kinase activity may be measured by the incorporation of 32 P from [γ - 32 P]ATP into ADP or GDP using a gamma radioisotope counter. The enzyme, in a kinase buffer, is incubated together with the appropriate nucleotide mono-phosphate substrate (AMP or GMP) and 32 P-labeled ATP as the phosphate donor. The reaction is incubated at 37°C and terminated by addition of trichloroacetic acid. The acid extract is neutralized and subjected

WO 02/08399

PCT/US01/23092

to gel electrophoresis to separate the mono-, di-, and triphosphonucleotide fractions. The diphosphonucleotide fraction is excised and counted. The radioactivity recovered is proportional to the enzyme activity.

In yet another alternative, other assays for PKIN include scintillation proximity assays (SPA), scintillation plate technology and filter binding assays. Useful substrates include recombinant proteins tagged with glutathione transferase, or synthetic peptide substrates tagged with biotin. Inhibitors of PKIN activity, such as small organic molecules, proteins or peptides, may be identified by such assays.

XVIII. Enhancement/Inhibition of Protein Kinase Activity

Agonists or antagonists of PKIN activation or inhibition may be tested using assays described in section XVII. Agonists cause an increase in PKIN activity and antagonists cause a decrease in PKIN activity.

XIX. Kinase Binding Assay

Binding of PKIN to a FLAG-CD44 cyt fusion protein can be determined by incubating PKIN to anti-PKIN-conjugated immunaffinity beads followed by incubating portions of the beads (having 10-20 ng of protein) with 0.5 ml of a binding buffer (20 mM Tris-HCL (pH 7.4), 150 mM NaCl, 0.1% bovine serum albumin, and 0.05% Triton X-100) in the presence of ¹²⁵I-labeled FLAG-CD44cyt fusion protein (5,000 cpm/ng protein) at 4 °C for 5 hours. Following binding, beads were washed thoroughly in the binding buffer and the bead-bound radioactivity measured in a scintillation counter (Bourguignon, L.Y.W. et al. (2001) J. Biol. Chem. 276:7327-7336). The amount of incorporated ³²P is proportional to the amount of bound PKIN.

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.

WO 02/08399

PCT/US01/23092

Table 1

Incyte Project ID	Polypeptide SEQ. ID. NO:	Incyte Polypeptide ID	Polynucleotide SEQ. ID. NO:	Incyte Polynucleotide ID
2564295	1	2564295CD1	21	2564295CB1
2837050	2	2837050CD1	22	2837050CB1
7474590	3	7474590CD1	23	7474590CB1
7474594	4	7474594CD1	24	7474594CB1
7477585	5	7477585CD1	25	7477585CB1
7477587	6	7477587CD1	26	7477587CB1
7594537	7	7594537CD1	27	7594537CB1
70467491	8	70467491CD1	28	70467491CB1
7078569	9	7078569CD1	29	7078569CB1
1598381	10	1598381CD1	30	1598381CB1
7074637	11	7074637CD1	31	7074637CB1
7170260	12	7170260CD1	32	7170260CB1
1707456	13	1707456CD1	33	1707456CB1
1851072	14	1851072CD1	34	1851072CB1
7474664	15	7474664CD1	35	7474664CB1
7474724	16	7474724CD1	36	7474724CB1
7478815	17	7478815CD1	37	7478815CB1
7477141	18	7477141CD1	38	7477141CB1
2190612	19	2190612CD1	39	2190612CB1
7477549	20	7477549CD1	40	7477549CB1

WO 02/08399

PCT/US01/23092

Table 2

Polypeptide SEQ ID NO:	Incyte Polypeptide ID	GenBank ID NC_001080	Probability score	GenBank Homolog
1	236423CD1	g168355	0.0	Kanflin receptor-related receptor [Homo sapiens]
2	283705CD1	g2853031	0.0	Tousled-like kinase [Mus musculus]
3	747459CD1	g6453611	5.1e-86	Protein Kinase (mutant form) [Mus musculus]
4	747459CD1	g3879221	5.6e-99	Predicted using GeneFinder similar to casein kinase I [Caenorhabditis elegans]
5	7477585CD1	g348245	3.5e-62	Protein serine/threonine kinase [Homo sapiens]
6	7477587CD1	g312938	7.4e-73	Protein Kinase [Homo sapiens]
7	7594337CD1	g485398	0.0	90kDa diacylglycerol kinase [Rattus norvegicus]
8	7046749CD1	g3089349	0.0	Cdc25C associated protein kinase C-TAK1 [Homo sapiens]
9	7478559CD1	g7960111 g9998852	4.2e-114 1.00E-123	Ethanolamine kinase [Homo sapiens] [Homo sapiens] ethanolamine kinase Lykidiis, A. et al. Overexpression of a mammalian ethanolamine-specific kinase activates the GTP-ethanolamine pathway J. Biol. Chem. 276, 2174-2179 (2001)
10	1698381CD1	g366715	7.7e-132	Protein kinase Meyerson, D. et al. (1992) EMBO J. 11, 2909-2917
11	7474637CD1	g1181079 g1401232	0.0 0	[Homo sapiens] diacylglycerol kinase Saito, F. et al. (1996) J. Biol. Chem. 271, 8394-8401 (Cricketidae gen. sp.) diacylglycerol kinase eta Klauck, T.M. et al. Cloning and characterization of a glucocorticoid-induced diacylglycerol kinase T. J. Biol. Chem. 271, 19781-19788 (1996)

Table 2 (cont.)

Polypeptide SEQ ID NO:	Incyte Polypeptide ID	GenBank ID NO:	Probability score	GenBank Homology
12	7170260CD1	g8101585	3.5e-126	[Mus musculus] testis specific serine kinase Zscharer, G. et al. (2000) Mech. Dev. 93:175-177
13	1797506CD1	g3300094	4.5e-227	[Homo sapiens] protein kinase/endoribonuclease Tirasophon, W. et al. (1998) Genes Dev. 12:1812-1824
		g12407081	0	[Homo sapiens] protein kinase/ribonuclease IRE1 beta Kawaguchi, T. et al. Translational control by the ER transmembrane kinase/Ribonuclease IRE1 under ER stress Nat. Cell Biol. 3: 159-164 (2001) [Schizosaccharomyces pombe] protein kinase Samejima, I., and Yanagida, M. (1994) Mol. Cell Biol. 14:1931-1937 [Homo sapiens] protein kinase-like protein Sato, S., Nakamura, Y., Kaneko, T., Katoh, T. et al. Structural analysis of Arabidopsis thaliana chromosome 3. I. Sequence features of the regions of 4,504,864 bp covered by sixty F1 and PAC clones Genes Dev. 12:1211-1220 [Mus musculus] protein kinase related to Raf protein kinases Theerrien, M. et al. (1995) Cell 83:879-888
14	1851973CD1	g1853976	1.3e-37	[Homo sapiens] receptor protein-tyrosine kinase Fox, G. M. et al. (1995) Proc Natl Acad Sci USA 92:12905-12909
		g9294469	5.00E-47	[Homo sapiens] protein kinase I Fuzeo, R. et al. (1998) Biochem. J. 331 (Pt. 2):607-613
15	7474604CD1	g1171250	2.0e-218	[Homo sapiens] protein kinase related to Raf protein kinases Theerrien, M. et al. (1995) Cell 83:879-888
16	7474721CD1	g551608	4.1e-220	[Homo sapiens] receptor protein-tyrosine kinase Fox, G. M. et al. (1995) Proc Natl Acad Sci USA 92:12905-12909
17	7478815CD1	g2873349	0.0	[Homo sapiens] protein kinase I Fuzeo, R. et al. (1998) Biochem. J. 331 (Pt. 2):607-613

WO 02/08399

PCT/US01/23092

Table 2 (cont.)

Polypeptide SEQ ID NO:	Incyte Polypeptide ID	GenBank ID NO:	Probability score	GenBank Homolog
18	7477549CDL	97239696	1.6e-87	[Homo sapiens] myosin light chain Kawaguchi, Y., et al. (1997) Garcia, J.G. et al. (1997) Am. J. Respir. Cell Mol. Biol. 16:489-494 Garcia, J.G.N. et al. (1996) Biochem. Biophys. Res. Commun. 1:1-1 [Mus musculus] striated muscle- specific serine/threonine protein kinase Huang, C.M. et al. Striated Muscle Preferentially Expressed Genes alpha and beta Are Two Serine/Threonine Protein Kinases Derived from the Same Gene as the Aortic Preferentially Expressed Gene- 1 T. Biol. Chem. 275 (47), 36966-36973 (2000)
19	2190612CDL	91836161	6.0e-257	[Rattus sp.] Ca2+/calmodulin- dependent protein kinase IV kinase Okuno, S., Kitani, T. and Fujisawa, H. (1996) J. Biochem. 119:1176-1181 (2000)
20	7477549CDL	95006445	3.6e-179	[Homo sapiens] CDC42-binding protein kinase beta Kawaguchi, Y., et al. (1999) Genomics 57:297-300 [Rattus norvegicus] myotonic dystrophy kinase-related Cdc42-binding kinase Leung, T. et al. Myotonic dystrophy kinase-related Cdc42-binding kinase acts as a Cdc42 effector in promoting cytoskeletal reorganization Maj. Cell. Biol. 18, 130-140 (1998)

WO 02/08399

PCT/US01/23092

Table 2 (cont.)

Polypeptide ID NO: SEQ ID NO:	Incyte Polypeptide ID	GenBank ID NO: 92217968	Probability score 1.408E-161	GenBank Homolog
20				[Homo sapiens] myotonic dystrophy protein kinase Kedra, D. et al. The germinal center kinase gene and a novel CDC25-like gene are located in the vicinity of the PKM gene on 11q13 Hum. Genet. 100, 611-619 (1997)

Table 3 (cont.)

SEQ ID NO.	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
2	2837050CD1	718	S185 S186 S194 S228 S238 S247 S252 S259 S265 S346 S349 S365 S632 T176 T269 T344 T403 T488 T558 T78 Y571 Y97	N240 N246 N548 N630 N713 N714	L002L L499-D97 PROTEIN KINASE DOMAIN DM00004 P34314 736-737 PROTEIN KINASE ST: I534-L546 KINASE PROTEIN TOUNSLEDLIKE PD013350: M237-D400 C287-L409 TOUNSLEDLIKE KINASE KIAA0137 PROTEIN PD035377: K184-T236 TOUNSLEDLIKE KINASE MUI/TIPLE TESTIS TRANSSCRIPT PD026280: A692-N718 Tyrosine kinase catalytic domain PRO0109: L490-K502, V608-W630 Protein Kinases Signatures and Profile Protein Kinase: Pf: E517-S570 Protein Kinase: Pf: E517-S570 Y408-L687 Protein Kinase Domain Ekinase: Protein Kinase: Pf: L414-K437 Protein Kinase: Pf: I534-L546 PROTEIN KINASE DOMAIN DM00004 P27448 58-297: V20-T265 Tyrosine kinase catalytic domain PRO0109: Y136-V154, V202-E224, L244-A286 Protein Kinases Signatures and Profile Protein Kinase: Pf: Q94-G174 Protein Kinase: Pf: Q94-G174 Y28-L275 Protein Kinase: Pf: V142-V154 PROTEIN KINASE DOMAIN DM00004 P48730 11-265: K144-Y392 SIMILAR TO CASEIN KINASES PDI15501: F312-D422, L130-T233 Eukaryotic protein Kinase Domain Ekinase: BL00-F374 Protein Kinase: Pf: T145-K169 Signature Sequences: M-419	BLAST-PRODOM BLAST-PRODOM BLAST-PRODOM BLAST-PRODOM BLAST-PRODOM BLIMFS-PRINTS PROFILESCAN HMMER-PFAM MOTIFS MOTIFS BLAST-DMO BLIMFS-PRINTS PROFILESCAN HMMER-PFAM MOTIFS BLAST-DMO BLAST-PRODOM HMMER-PFAM MOTIFS SPECAN
3	7474590CD1	497	S17 S286 S291 S3 S314 S356 S372 S375 S381 S382 S409 S440 S445 S446 S447 T25 T26 T427 T445 T461	R243	PROTEIN KINASE ST: I534-L546 PROTEIN KINASE DOMAIN DM00004 P27448 58-297: V20-T265 Tyrosine kinase catalytic domain PRO0109: Y136-V154, V202-E224, L244-A286 Protein Kinases Signatures and Profile Protein Kinase: Pf: Q94-G174 Protein Kinase: Pf: Q94-G174 Y28-L275 Protein Kinase: Pf: V142-V154 PROTEIN KINASE DOMAIN DM00004 P48730 11-265: K144-Y392 SIMILAR TO CASEIN KINASES PDI15501: F312-D422, L130-T233 Eukaryotic protein Kinase Domain Ekinase: BL00-F374 Protein Kinase: Pf: T145-K169 Signature Sequences: M-419	BLAST-DMO BLIMFS-PRINTS PROFILESCAN HMMER-PFAM MOTIFS BLAST-DMO BLAST-PRODOM HMMER-PFAM MOTIFS SPECAN
4	7474594CD1	741	S397 S402 S471 S592 S641 S652 S656 S737 T237 T274 T292 T308 T388 T567	K119 N291	PROTEIN KINASE ST: I534-L546 PROTEIN KINASE DOMAIN DM00004 P27448 58-297: V20-T265 Tyrosine kinase catalytic domain PRO0109: Y136-V154, V202-E224, L244-A286 Protein Kinases Signatures and Profile Protein Kinase: Pf: Q94-G174 Protein Kinase: Pf: Q94-G174 Y28-L275 Protein Kinase: Pf: V142-V154 PROTEIN KINASE DOMAIN DM00004 P48730 11-265: K144-Y392 SIMILAR TO CASEIN KINASES PDI15501: F312-D422, L130-T233 Eukaryotic protein Kinase Domain Ekinase: BL00-F374 Protein Kinase: Pf: T145-K169 Signature Sequences: M-419	BLAST-DMO BLAST-PRODOM HMMER-PFAM MOTIFS SPECAN

Table 3 (cont.)

SEQ ID NO.	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
5	7477585CD1	643	S251, S273, S274, S47, S49, S52, S60, S64, S84, S97, T211, T302, T323, T340, T538, T547, Y368	NS98, N71	PROTEIN KINASE DOMAIN DM00004 P51957 8-251; Tyrosine kinase catalytic domain PRO0109; T108-Q124, Y148-L166, Y256-S278 Protein Kinases signatures and profile Protein Kinase Tyr.p.f.: O134-S185 Eukaryotic protein kinase domain pkinase: Y29-L287 PROTEIN_KINASE_St: T154-L166	BLAST-DMO BLAST-PRINTS PROFILES-SCAN HMMER-PFAM MOTIFS
6	7477587CD1	623	S22, S33, S39, S81, S83, S84, S85, S86, S87, S88, S89, T454, T463, T564	N546, N646, N793	PROTEIN KINASE DOMAIN DM00004 P53350 95-133; R39-R497, S293-T319 Tyrosine kinase catalytic domain PRO0109; Y212-L236 Protein Kinases signatures and profile Protein Kinase Tyr.p.f.: E198-Q250 Transmembrane domain transmem_domain: L555-S575 Eukaryotic protein kinase domain pkinase: Y97-A287, S288-F319 PROTEIN_KINASE_ACP: I103-R126 PROTEIN_KINASE_St: I218-L230	BLAST-DMO BLAST-PRINTS PROFILES-SCAN HMMER HMMER-PFAM MOTIFS MOTIFS
7	7594537CD1	797	S11, S13, S16, S208, S25, S294, S380, S670, S675, S684, S81, T2, T26, T274, T298, T312, T320, T388, T518, T74, T77, T86, T743, T762, T766, Y449	N546, N646, N793	PHOREB-ESTER AND DAG BINDING DOMAIN DM01311 P49521 326-792; V321-K789 KINASE DIACYLGLYCEROL PHORBOL-ESTER BINDING TRANSFERASE DIGLYCERIDE DAG MULTIGENE FAMILY DKG_P0002939; I575-P755 PROBABLE DIACYLGLYCEROL KINASE EC 2.7.1.107 DIGLYCERIDE DKG DAG HYDROPHILIC PROTEIN BINDING MULTIGENE FAMILY PHORBOL-ESTER BINDING P0078689; A118-Q296, I140-D85, I50-S81 KINASE DIACYLGLYCEROL PHORBOL-ESTER BINDING PROTEIN TRANSFERASE DIGLYCERIDE DAG MULTIGENE FAMILY PD002780; P431-W555	BLAST-DMO BLAST-PRODOM BLAST-PRODOM BLAST-PRODOM BLAST-PRODOM BLAST-PRODOM

Table 3 (cont.)

SEQ ID NO.	Incyte PD	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
					DIACYLGLYCEROL KINASE. RFR. EC 2.7.1.107 DIGLYCERIDE KINASE DSK DAG 90 KD TRANSFERASE CALCIUMBINDING PHORBOL-ESTER BINDING MULTIGENE FAMILY PD119174: D352- H430 Diacylglycerol kinase catalytic domain PF00781: H31-Q336 P431-Y462 R483-L497 P498-532 X539-V559 N577-Y613 L655-G666 L707-Q752 Diacylglycerol kinase catalytic domain DAGK: P431-N555 Diacylglycerol kinase accessory domain DAGKa: I575-P755 Phorbol esters/diacylglycerol binding domain d DAG-PE-bind: H238-C287. H303-C351 EF Hand sfhand: K146-W174. I191-T219 Phorbol esters/diacylglycerol binding domain Phorbol esters/diacylglycerol binding domain Phorbol esters/diacylglycerol binding domain Phorbol esters/diacylglycerol binding domain Phorbol esters/diacylglycerol binding domain Dag-PE binding domain prf: Y250-G378 Dag-PE binding domain: H238-C287 Ef-Hand: D155-L167. D200-W212	BLAST-PRODOM BLIMPS-PFAM HMMER-PFAM HMMER-PFAM HMMER-PFAM HMMER-PFAM HMMER-PFAM HMMER-PFAM BLIMPS-SHOCKS PROFILES-CAN MOTIFS MOTIFS
8	7046749ICD1	749	S141 S2 S24 S346 S372 S417 S424 S444 S456 S457 S461 S49 S454 S495 S516 S634 S653 S659 S664 S730 T118 T32 T638 T639 T519 T535 T614 T618 T623 T62 T9 Y113	N326 N329 N400 N479 N533 N637	PROTEIN KINASE DOMAIN DM00004 P27448 56- 297: L62-L300 KINASE SERINE/THREONINE/PROTEIN TRANSFERASE ATTBINDING SERINE/THREONINE PUTATIVE KIN. EMK PARL PD004300: G633-L749 KINASE SERINE/THREONINE/PROTEIN SERINE/THREONINE PUTATIVE TRANSFERASE ATTBNDING PROTEIN EMK P78 CDC25C KINASE P78: S413-S652 KINASE SERINE/THREONINE/PROTEIN SERINE/THREONINE TRANSFERASE ATTBINDING PROTEIN PARL K678 EMK PD005638: E312-R412 SERINE/THREONINE KINASE PD19193: S551- P622	BLAST-DOMO BLAST-PRODOM BLAST-PRODOM BLAST-PRODOM BLAST-PRODOM BLAST-PRODOM

WO 02/08399

PCT/US01/23092

Table 3 (cont.)

SEQ ID NO.	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and
9	7478559CD1	366	S237 S289 S355 S38 S380 T20 T322 T85 T93 Y271	N138	Protein Kinase St: K178-L191 Protein Kinase St: Y031-L041 E0285_I0 DM01931 P3790 128-455; D256-K376; F131-E300 G3 CHOLINE_KINASE; YDR147M; R0285_I0; DM01931 P4656 1-305; E125-A289 KINASE CHOLINE TRANSFERASE PROTEIN MUTIGENE FAMILY PUTATIVE LIKE CHROMOSOME III_P003547; V229-I382; V109-E240 KINASE TRANSFERASE CHOLINE_P002952; V243-E236; E263-N292 G11; Methanotamine Kinase Choline_Kinase; T85-T356	BLIAPS-PAIPTS PROFILIBSCAN HMER-PFAM HMER-PFAM
						Protein Kinase St: K178-L191 Protein Kinase St: Y031-L041 E0285_I0 DM01931 P3790 128-455; D256-K376; F131-E300 G3 CHOLINE_KINASE; YDR147M; R0285_I0; DM01931 P4656 1-305; E125-A289 KINASE CHOLINE TRANSFERASE PROTEIN MUTIGENE FAMILY PUTATIVE LIKE CHROMOSOME III_P003547; V229-I382; V109-E240 KINASE TRANSFERASE CHOLINE_P002952; V243-E236; E263-N292 G11; Methanotamine Kinase Choline_Kinase; T85-T356
10	1698381CD1	342	S180 S205 S238 S284 S288 S38 T247 Y15 Y211	N23	Eukaryotic protein kinase domain_pkinase;Y4-F286; F286; Protein kinases signatures and profile Protein_kinase_lycosine; E90-G154 PROTEIN_KINASE_DOMAIN_DM00004 Q00532 7-278; K6-G277 PROTEIN_KINASE_DOMAIN_DM00004 Q00526 6-286; K6-E286 PROTEIN_KINASE_DOMAIN_DM00004 P23437 6-286; K6-G218 PROTEIN_KINASE_DOMAIN_DM00004 P51958 6-277; K6-G218 KINASE TRANSFERASE PROTEIN SERINE/THREONINE PROTEIN APP-BINDING II PHOSPHORYLATION CASEIN ALPHA CHAIN_P002608; V161-F286	HMER-PFAM PROFILIBSCAN BLAST_DOMO BLAST_DOMO BLAST_DOMO BLAST_DOMO BLAST_DOMO BLAST_PRODOM

WO 02/08399

PCT/US01/23092

Table 3 (cont.)

SEQ ID No.	Incyte Polypeptide	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Domains
12	7170260CD1	268	S161 S188 S255 S29 T15 Y124 Y21		<p>Signature Sequences, Domains and Motifs</p> <p>Phorbol esters/Diacetyl glycerol binding domain H176-G198, H202-G217, I415-I427 Proteins BT00479: Diacetyl glycerol kinase catalytic domain (presumed) PF00781: K278-K283, P332-P363, R384-I398, C410-Y433, Q441-T451, N772-Y909, I948-G961, V919-Q930 Diacetyl glycerol/cholesterol-ester binding signature motif: H207-P208, P210-P213, P214-P218, K74-I84, I386-G395, I466-I473 Phorbol esters/Diacetyl glycerol binding domain: H176-C225 Eukaryotic protein kinase domain pkinase: Y10-L265 Protein kinases signatures and profile Protein kinase Tyrosine: S82-H162 KINASE DOMAIN DM00004 P27448 58-297: BLAST_DOMO PROTEIN KINASE DOMAIN DM00004 I466609 55-294: BLAST_DOMO K14-S255 PROTEIN KINASE DOMAIN DM00004 Q05512 55-294: BLAST_DOMO K14-S255 PROTEIN KINASE DOMAIN DM00004 JC1446 20-261: BLAST_DOMO Q11-I256 Tyrosine kinase catalytic domain signature PRO109: Y124-L142 Protein kinases ATP-binding region signature: I16-K39</p>	<p>BLMFS_BGCNS BLMFS_PFBM BLMFS_PRINTS BLMFS_PRODOM MOTIFS HMNER_PFBM PROFILESSCAN BLAST_DOMO BLAST_DOMO BLAST_DOMO BLMFS_PRINTS MOTIFS</p>
13	1797506CD1	965	S234 S236 S537 S530 S606 S636 S741 S841 S879 S884 S92 T111 T143 T155 T174 T202 T215 T229 T25 T372 T619 T685 T82 T922 T932 T963 T173	N227	<p>Signature Sequences, Domains and Motifs</p> <p>Phorbol esters/Diacetyl glycerol binding domain H176-G198, H202-G217, I415-I427 Proteins BT00479: Diacetyl glycerol kinase catalytic domain (presumed) PF00781: K278-K283, P332-P363, R384-I398, C410-Y433, Q441-T451, N772-Y909, I948-G961, V919-Q930 Diacetyl glycerol/cholesterol-ester binding signature motif: H207-P208, P210-P213, P214-P218, K74-I84, I386-G395, I466-I473 Phorbol esters/Diacetyl glycerol binding domain: H176-C225 Eukaryotic protein kinase domain pkinase: Y10-L265 Protein kinases signatures and profile Protein kinase Tyrosine: S82-H162 KINASE DOMAIN DM00004 P27448 58-297: BLAST_DOMO PROTEIN KINASE DOMAIN DM00004 I466609 55-294: BLAST_DOMO K14-S255 PROTEIN KINASE DOMAIN DM00004 Q05512 55-294: BLAST_DOMO K14-S255 PROTEIN KINASE DOMAIN DM00004 JC1446 20-261: BLAST_DOMO Q11-I256 Tyrosine kinase catalytic domain signature PRO109: Y124-L142 Protein kinases ATP-binding region signature: I16-K39</p>	<p>HMNER_PFBM PROFILESSCAN BLAST_DOMO BLAST_DOMO BLAST_DOMO BLMFS_PRINTS MOTIFS</p>

WO 02/08399

PCT/US01/23092

Table 3 (cont.)

SEQ ID No. / Invertebrate ID	Amrro Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Programs
14	1851973CD1 329	S264 S270 S293 S31 S31.1 S320 S7	N73	<p>KINASE: THREONINE; AMP; SERINE: PROTEIN KINASE/ENDORIBONUCLEASE PUTATIVE SERINE/THREONINE PROTEIN KINASE C41C4.4 CHROMOSOME II PRECURSOR TRANSFERASE PD152704; T197-L422, L88-E190 SERINE/THREONINE PROTEIN KINASE PRECURSOR TRANSFERASE SIGNAL TRANSFERASE ATP-DEPENDENT PHOSPHORYLASE G1COPROTEIN PD132590; G821-Y849 Tyrosine kinase catalytic domain signature PRO0109; H656-I684, G721-I731, V743-D765 Serine/Threonine protein kinases active-site signature: I672-I684 Phosphorylase Kinase family signature PRO1042; H42-THREONINE protein kinase domain signature: R35-V180 Protein kinases signatures and profile PROTEIN_KINASE_Tyrosine: M132-RL84 PROTEIN_KINASE_DOMAIN DM00004 F43565 796-1240; I37-R184 PROTEIN_KINASE_DOMAIN DM00004 A56155 714-1002; V38-L177 PROTEIN_KINASE_DOMAIN DM00004 F38679 238-52706; V38-L177 PROTEIN_KINASE_DOMAIN DM00004 P53894 353-658; V38-S178 Tyrosine kinase catalytic domain signature PRO0109; M110-H123, Y146-I164 Serine/Threonine protein kinases active-site signature: I152-I164</p>	<p>BLAST_DOMO BLAST_PRODOM BLAST_PRODOM BLIMPS_PRINTS MOTIFS BLIMPS_PRINTS HMGR_PFBM PROFILESCAN BLAST_DOMO BLAST_DOMO BLAST_DOMO BLIMPS_PRINTS MOTIFS</p>

WO 02/08399

PCT/US01/23092

Table 3 (cont.)

SEQ ID No.	Incye Polypeptide	Amino Acid Residues	Potential Glycosylation Sites	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and
15	7474604CD1	945	S110 S157 S193 S246 S289 S290 S31 S329 S356 S359 S405 S411 S611 S623 S636 S645 S67 S934 T170 T2 T217 T322 T339 S47 T496 T712 T839	S140 M155 M362 M631 N756 N888	Eukaryotic protein kinase domain pkinase: E661-2920 Protein kinases signatures and profile Protein kinase tyrosine: K757-L801 PROTEIN KINASE DOMAIN DM000004 E27966 85-332: I663-F916 PROTEIN KINASE DOMAIN DM000004 L15058 458-705: I663-F916 PROTEIN KINASE DOMAIN DM000004 F10398 312-326: I663-F916 PROTEIN KINASE DOMAIN DM000004 E26126 305-532: I667-F916 KINASE SUPPRESSOR OF RAS1 KSR1 HB PROTEIN PD10125: V390-P557, K501-L661 L222-P223 KINASE SUPPRESSOR OF RAS KSR PHORBOL-ESTER BINDING RAS1 KSR1 HB PD017776: L21-E344 S485-T519 Tyrosine kinase catalytic domain signature K501-L661 Y771-Y789, W867-I877, M894-F916 Serine/threonine protein kinases active-site signature: L777-Y789 Eukaryotic protein kinase domain pkinase: V645-H897 Ephrin receptor ligand binding domain EPH_1bd: E35-C211 Protein kinases signatures and profile PROTEIN KINASE TYROSINE: Q746-A799 RECEPTOR TYROSINE KINASE CLASS V DM0501 S51741 133-382: V746-G394 RECEPTOR TYROSINE KINASE CLASS V DM0501 F54759 133-382: V746-G394 RECEPTOR TYROSINE KINASE CLASS V DM0501 L486 2134-382: I37-G394 RECEPTOR TYROSINE KINASE CLASS V DM0501 I486 2134-382: I37-G394 KINASE RECEPTOR PRECURSOR TYROSINE PROTEIN EPHRIN TRANSFERASE ATP-BINDING PHOSPHORYLATION TRANSMEMBRANE GLYCOPROTEIN PD001495: E35-C211	HMMER_PFAM PROFILESSCAN BLAST_DOMO BLAST_DOMO BLAST_DOMO BLAST_DOMO BLAST_PRODUM BLAST_PRODUM BLIMPS_PRTMVS MOTIFS HMMER_PFAM HMMER_PFAM PROFILESSCAN BLAST_DOMO BLAST_DOMO BLAST_DOMO BLAST_PRODUM	
16	7474721CD1	1009	S184 S203 S244 S293 S325 S44 S473 S62 S625 S682 S686 S805 S825 S831 S890 T162 T244 T224 T232 T32 T423 T488 T551 T616 T619 Y504 Y766 Y801	M311 N486			

WO 02/08399

PCT/US01/23092

Table 3 (cont.)

SEQ ID NO.	Incyte Polypeptide	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Domains and Motifs	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
18	747714ICD1	2380	S143 S166 S241 S277 S278 S285 S299 S343 S480 S537 S553 S568 S602 S711 S716 S996 S1033 S1035 S157 S162 S1127 S157 S162 S1127 S1435 S1468 S1492 S1586 S1609 S1679 S1691 S1747 S1117 S1537 S1557 S1578 S1594 S1613 S1736 S1747 S1876 S1947 S2057 S2078 S2083 S2321 S2359 S2662 S7154 S986 T108 T153 T158 T170 T350 T408 T476 T498 T578 T614 T692 T803 T822 T831 T1008 T1082 T1311 T1363 T1366 T1981 T2083 T1301 T1856 T1901 T2069 T2101 T2144 T2348 T1608 T2343 Y632 Y772 Y822	N378 N1575 N1847 N1374 N2099 N2299	Eukaryotic Protein Kinase domain PKinase: Y714-F967, Y2079-L2331 PROTEIN KINASE DOMAIN DM000004 S07571 S152-5396; D715-D952, E2083-L2322 PROTEIN KINASE DOMAIN DM000004 P53355 15-257; Q718-D952, E2083-L2322 PROTEIN KINASE DOMAIN DM000004 JN0583 727-969; I716-D952, L2082-L2312 PROTEIN KINASE DOMAIN DM000004 F07313 298-541 Tyrosine kinase catalytic domain signature P00109; Y822-V840 signal peptide: M53-A70 Eukaryotic protein kinase domain PKinase: Y2079-L2331 Protein kinases ATP-binding region Signatures: I720-K743 Serine/threonine protein kinases active- V828-V840, V2194-L2206	HMMER_PFBM BLAST_DOMO BLAST_DOMO BLAST_DOMO BLAST_DOMO HMMER_PFBM MOTIFS MOTIFS
19	2190612CD1	505	S100 S117 S160 S330 S419 S425 S437 S458 S49 S483 S488 T26 P430 T58	N147	Eukaryotic protein kinase domain: Y128-V409 Protein kinases signatures and profile Protein Kinase Tyrosine: Q251-N303 PROTEIN KINASE DOMAIN DM000004 AS7156 130-399; PROTEIN KINASE DOMAIN DM000004 P50526 136-399; E133-I398 PROTEIN KINASE DOMAIN DM000004 P38990 135-438; E133-E320, N303-V400 PROTEIN KINASE DOMAIN DM000004 P43637 52-334; I134-I378	HMMER_PFBM PROFILIBSCAN BLAST_DOMO BLAST_DOMO BLAST_DOMO BLAST_DOMO

WO 02/08399

PCT/US01/23092

Table 3 (cont.)

SEQ ID NO.	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
					KINASE HQ ASSOCIATED COILED COIL PROTEIN FORMING PHORBOL ESTER BINDING DYSTROPHY KINASE RELATED CDC42 BINDING PD006745: 7944-11038, H433-L456 PHORBOL ESTER BINDING DYSTROPHY KINASE RELATED CDC42 BINDING KINASE GENGHIS KHAN MYOTOMIC MYOTOMIC PD01192: S694-S615 Tyrosine Kinase catalytic domain signature R001091: C257-E273, M448-S161, S185-L203 Phorbol esters/diacylglycerol binding dom DAG PE-bind: H807-C938 Phorbol esters/diacylglycerol binding domain: H887-C935, Protein kinases ATP-binding region signature I77-K100 Serrine/Threonine protein kinases active-site signature: Y191-L203 CMH domain: L1100-K1380 Protein Kinase C-ε serine domain: P351-D366 Et domain FR: P358-F374 SH3-like cleavage site: M-537	BLAST_PRODOM BLAST_PRODOM BLIMPS_PRINTS HMMER_PPFAM HMMER_PPFAM HMMER_PPFAM HMMER_PPFAM HMMER_PPFAM

WO 02/08399

PCT/US01/23092

Table 4

Polynucleotide SEQ ID NO:	Incyte Polynucleotide ID	Sequence Length	Selected Fragment(s)	Sequence Fragments	5' Position	3' Position
21	2564295CBL	4298	701-1736, 3536-3829, 1-386, 2349-4589, 3936-4848, 2841-3428	FL2564295_g7160581_000014_g 387060_1_15-16 FL2564295_g7160581_000014_g 387060_1_16-17 FL2564295_g7160581_000014_g 387060_1_7-8 5507839351 FL2564295_g7160581_000014_g 387060_1_8-9 5507838671 FL2564295_g7160581_000014_g 387060_1_18-19 2564295HL (ADRFRTFVLI) GI86554_CD FL2564295_g7160581_000014_g 387060_1_9-10 3595581HL (DRGFRNFTVLI) FL2564295_g7160581_000014_g 387060_1_1-2 FL2564295_g7160581_000014_g 387060_1_10-11 FL2564295_g7160581_000014_g 387060_1_20-21 FL2564295_g7160581_000014_g 387060_1_11-12 FL2564295_g7160581_000014_g 387060_1_2-3 FL2564295_g7160581_000014_g 387060_1_12-13 FL2564295_g7160581_000014_g 387060_1_3-4 FL2564295_g7160581_000014_g 387060_1_4-5	3200 3253 1938 37 2167 1 3594 4048 442 442 2335 3453 441 2531 3884 2573 994 2794 1298 1441	3482 3593 2334 717 2530 709 3883 4298 4250 2572 3756 1297 2793 4250 2930 1440 3093 1585 1800

WO 02/08399

PCT/US01/23092

Table 4 (cont.)

Polynucleotide SEQ ID NO:	Incyte Polynucleotide ID	Sequence Length	Selected Fragment(s)	Sequence Fragments	5' Position	3' Position		
21				FL2564295_g7160581_000014_s	3094	3232		
				387060_1_14-15				
				FL2564295_g7160581_000014_s	3754	4018		
22	2837050CBL	2863	1-430, 2346-2863	6854541HI (BRALPEN08)	782	1467		
				61164223	1	496		
				71191190V1	1439	2085		
				7728560HI (UFRCDIE01)	79	681		
				71972220V1	2227	2863		
				71972389V1	2180	2857		
23	7474590CBL	1494	1-1494	6881340HI (BRAHTR03)	1555	2209		
				7401101HI (SINDHE01)	598	1393		
				GBL-68103343_000001_eb1e	1	1494		
24	7474594CBL	2341	682-792, 1-262, 1522-2341, 1254-1373, 339-361	FL7474590_g7630344_000002_g	1	1116		
				6779549_1_1				
				55053685V1	1512	2341		
				6849237HI (BRATUN02)	858	1544		
				8016740V1 (BRARFY01)	240	929		
				ONS-68247875_000031_002	1	426		
				7278930HI (BRARCE01)	1281	1779		
25	7477585CBL	2552	1-465, 1075-1150	688-96689704_000006_002	1180	1590		
				71975409V1	1988	2534		
				55030074V1	612	1305		
				55030074V1	1241	1900		
				34066660F6 (LATPTU02)	1	686		
				632987HI (BRANDT01)	1384	1830		
				71987367V1	2019	2552		
				6704049HI (DRSCNOF02)	1849	2517		
				55030089HI	679	1390		
				98671962_e31e	1	1980		
26	7477587CBL	2176	1276-1873, 1-286	5823464F7 (PROSTUS23)	1662	2164		

WO 02/08399

PCT/US01/23092

Table 4 (cont.)

Polynucleotide SEQ ID NO:	Incyte Polynucleotide ID	Sequence Length	Selected Fragment(s)	Sequence Fragments	5' Position	3' Position
27	7594537CB1	4277	2383-2814,	7594537HL (LIVRNC007)	130	766
			611-1170, 1-	7228693HL (UFRCDIE01)	1	351
			518, 2763-2834, 1714-1859, 3119-4277			
28	70467491CB1	2616	1717-2616, 1-	2395018F6 (THPLAZT01)	2015	2520
			485	FL70467491_gf7708222_gf7595800	1	2250
29	7478559CB1	1253	1215-1253, 1-	g3770955	1	321
			53	76617157L (OVARNOE02)	655	1253
30	1698381CB1	1790	1-146, 892-	95769093	314	804
			1313, 1659-	1598381F6 (BRADYU05)	523	1019
			1790, 186-237	5506829371	1	786
31	7474637CB1	4132	1599381F6	71870273V1	1186	1790
			3430-3535, 1-	1599381F6 (BRADYU05)	776	1363
			377, 4035-	4122796F6 (CARGDIF01)	3639	4132
			412, 1501-	55076747H1	2871	3468
			2486	55075847H1	1379	1783
				55075848H1	1623	1987
				55077477H1	1045	1472
32	7170260CB1	1137	877-1137	GB1_g8247425_000008_000011.edit	504	1126
				5507675671	3148	3745
				GMN_g5648263_002.edit5p	2805	3026
				6286939H2 (EFTPONA01)	1041	1168
				7721743H2 (THZRDIE01)	35	503
				6766106H1 (BRAUNOR01)	1893	2256
				55061367H1	2149	2841
				67661067L (BRAUNOR01)	368	909
				1752420H1 (LIVRVT01)	1	157
				55046242J2	694	1137
	3152309F6 (MLVMTX02)	1	145			

Table 4 (cont.)

Polynucleotide SEQ ID NO:	Incyte Polynucleotide ID	Sequence Length	Selected Fragment(s)	Sequence Fragments	5' Position	3' Position
32				7659273UL (OVARNOE02)	415	971
				55046250H1	692	1108
				3343082F7 (SPLANCT09)	144	555
33	1797506CB1	3365	1-1032, 3340-3365, 1532-1735	P111797506_g7458755_000013_g	2793	3365
				3766209	1	2898
34	1851973CB1	2049	1-125, 1836-2049, 806-915	7667239H1 (URENUTC01)	1289	1800
				5507565571	547	1222
				5507725771	378	1221
				55067487H1	1	531
				1454265F1 (PENICUT01)	1179	1617
35	747460CB1	2962	1-1526, 1757-2114, 2481-2962	1454265F1 (PENICUT01)	1179	1617
				1454265G6 (PENICUT02)	1436	2049
				550757691	1760	2440
				8104269H1 (MIXDDIE02)	1	746
				596594692	1734	2833
				6867701F6 (BRANFTR03)	2255	2962
				5507076C1	1214	1763
				55075283J1	651	1335
36	7474721CB1	3112	2395-3112, 1353-1459, 2014-2280	6802884F6 (COLMOR03)	2055	2826
				71976507V1	1564	2315
				55075253J1	314	980
				GB1.GF6956165_000001.raw	1910	3112
				GB1.GF6956165.raw	140	1735
				55062828H1	1	712
				71990671V1	1418	2051
				55076655H1	1	658
				6934749H1 (SINUTMR02)	1710	2388
				238539R6 (SINUTMR02)	3159	3547
37	7478815CB1	3650	862-1366, 1826-1989, 1-787, 3623-3650	614864T6 (COLMUT02)	3004	3614
				70845065V1	1862	2441
				70842842V1	2420	3075

WO 02/08399

PCT/US01/23092

Table 4 (cont.)

Polynucleotide SEQ ID NO:	Incyte Polynucleotide ID	Sequence Length	Selected Fragment(s)	Sequence Fragments	5' Position	3' Position
37				72026676VL	966	1742
				55075416VL	358	926
				6657793	919	1011
38	7477141CB1	7789	1-699, 6785-6880, 7767-7789, 7184-7214, 1237-6218	70863076VL	2471	3154
				260525F6 (LUNGCTP07)	3378	3650
				7351520HL (HEARNON03)	7201	7767
				GBI-98014664	1	260
				97242948 CD	63	6763
				3012344HL (MUSCHN07)	7488	7772
39	2190612CB1	1937	727-1188, 1-643, 1731-1761	71179707VL	6783	7436
				76424052L (SEAVYDE01)	6728	7294
				7075595VL	1	498
				550240957L (PKTINW04)	914	1558
				6864657HL (BRATFER08)	1441	1937
				71887308Z (BRATDEC01)	1353	1820
				70780513VL	500	861
				70780609VL	384	529
				55121415HL	4374	5373
				55121422VL	4413	5274
40	7477549CB1	5373	4983-5373, 1-1612, 2046-2470, 4414-4442, 2356-2647, 2814-3056	7992167HL (OVRSDIC01)	3402	4043
				71999521VL	1448	1590
				6822270HL (SIMNOR01)	857	1407
				GNN-94755212.O10.ed1L	1	4567
				6594063HL (LUNGFER02)	2835	3147
				71644493HL (PLACNOR01)	3204	3711
				71583419VL	713	1385
				7402224HL (SINDMBS1)	289	795
				7694930HL (LNOPTVB01)	1082	1448
				7978995HL (LSUBDDK01)	1478	2186

WO 02/08399

PCT/US01/23092

Table 5

Polynucleotide SEQ. ID NO:	Incyte Project ID	Representative Library
21	2564295CE1	ADREN001
22	2837050CE1	THYM003
24	7474594CE1	BRAD001
25	7477585CE1	BRAD002
26	7477587CE1	PROST023
27	7594537CE1	LIVRN007
28	70467494CE1	PROSN018
29	7478559CE1	OVARN002
30	1698381CE1	BLADN005
31	7474637CE1	EPIT001
32	7170260CE1	OVARN002
33	1797506CE1	COLON003
34	1851973CE1	PRN001
35	7474604CE1	BRAD003
36	7474721CE1	COLON003
37	7478615CE1	SINT003
38	7477141CE1	SKIRN001
39	2190612CE1	ADREN007
40	7477549CE1	SINT001

WO 02/08399

PCT/US01/23092

Table 6

Library	Vector	Library Description
ADRETT001	ESFORT	Library was constructed using RNA isolated from right adrenal tumor tissue removed from a 30-year-old Turkish male during unilateral adrenalectomy. Pathology indicated a metastatic renal cell carcinoma that formed a circumscribed, spongy, hemorrhagic nodule adjacent to renal cell carcinoma. This patient presented with corticoidrenal insufficiency, including bilateral hemihypertrophy, hepatomegaly, and hypokalemia. Pathology included renal cell carcinoma. Family history included liver disease.
ADRETT007	PINCY	Library was constructed using RNA isolated from adrenal tumor tissue removed from a 43-year-old Caucasian female during a unilateral adrenalectomy. Pathology indicated pheochromocytoma.
BLADTT005	PINCY	Library was constructed using RNA isolated from bladder tumor tissue removed from a 66-year-old Caucasian male during a radical prostatectomy, radical cystectomy, and urinary diversion. Pathology indicated grade 3 transitional cell carcinoma on the anterior wall of the bladder. Patient history included lung neoplasm and tobacco abuse in remission. Family history included malignant breast neoplasm, tuberculosis, cerebrovascular disease, atherosclerotic coronary artery disease, and lung cancer.
BMAR00201	PINCY	This 5' biased random primed library was constructed using RNA isolated from treated SH-SY5Y cells derived from a metastatic bone marrow neuroblastoma, removed from a 4-year-old Caucasian female (Schering AG). The medium was MEM/HAM'S F12 with 10% fetal calf serum. After reaching about 80% confluency cells were treated with 6-Hydroxydopamine (6-OHDA) at 100 microm for 8 hours.
BRADTT003	PCDR2.1	This random primed library was constructed using RNA isolated from archaocortex, anterior hippocampus tissue removed from a 55-year-old Caucasian female who died from cholangiocarcinoma. Pathology indicated mild meningeal fibrosis predominately over the convexities, scattered axonal spheroids in the white matter of the cingulate cortex and the thalamus, and a few scattered neurofibrillary tangles in the entorhinal cortex and the periaqueductal gray region. Pathology for the associated tumor tissue indicated well-differentiated cholangiocarcinoma of the liver with residual or relapsed tumor. Patient history included cholangiocarcinoma, post-operative Budd-Chiari syndrome, biliary ascites, hydrothorax, dehydration, malnutrition, oliguria and acute renal failure. Previous surgeries included cholecystectomy and resection of 85% of the liver.
BRALM002	PINCY	This thalamus tissue library was constructed from 4.24 million independent clones from a thalamus tissue library. Starting RNA was made from thalamus tissue removed from a 35-year-old Caucasian male who died from cardiac failure. Pathology indicated moderate leptomeningeal fibrosis and multiple microinfarctions of the cerebral neocortex. Microscopically, the cerebral hemisphere revealed moderate fibrosis of the leptomeninges with focal calcifications. There was evidence of shrunken and slightly eosinophilic pyramidal neurons throughout the cerebral hemispheres. Scattered throughout the cerebral cortex, there were multiple small microscopic areas of cavitation with surrounding

WO 02/08399

PCT/US01/23092

Table 6 (cont.)

Library	Vector	Library Description
COLENR03	PCDNA2.1	gliosis. Patient history included dilated cardiomyopathy, congestive heart failure, cardiomegaly and an enlarged spleen and liver. The library was normalized in two rounds using conditions adapted from Soares et al., FMS (1994) 91:9228-9232 and Bonaldo et al., Genome Research (1996) 6:791, except that a significantly longer (48 hours/round) reannealing hybridization was used.
BPIPUNA01	PSPORT	Library was constructed using RNA isolated from colon epithelium tissue removed from a 13-year-old Caucasian female who died from a motor vehicle accident.
LIVRNGC07	pUNCY	Library was constructed using RNA isolated from untreated prostatic epithelial cell tissue removed from a 17-year-old Hispanic male. Serologies were negative.
OVARNOE02	PCDNA2.1	Library was constructed using pooled cDNA from two different donors. cDNA was generated using RNA isolated from liver tissue removed from a 20-week-old Caucasian male fetus who died from Patau's Syndrome (donor A) and a 16-week-old Caucasian female fetus who died from anencephaly (donor B). Family history included mitral valve prolapse in the mother of donor B.
PEMTUT01	pUNCY	This 5' biased random primed library was constructed using RNA isolated from right ovary tissue removed from a 47-year-old Caucasian female during total abdominal hysterectomy, bilateral salpingo-oophorectomy, incisional hernia repair, and panniculectomy. The patient presented with premenopausal menorrhagia. Patient history included osteoarthritis, tubal pregnancy, and polio osteopathy of the left leg. Previous surgeries included gastroenterostomy, plastic repair of the palate, adenotonsillectomy, dilation and curettage, cholecystectomy, and bladder reconstruction. Patient medications included vitamins, iron, and zinc. Family history included benign hypertension and type II diabetes in the father, and type II diabetes in the sibling(s).
		Library was constructed using RNA isolated from tumor tissue removed from the penis of a 64-year-old Caucasian male during penile amputation. Pathology indicated a fungating invasive grade 4 squamous cell carcinoma involving the inner wall of the foreskin and extending onto the glans penis. Patient history included benign neoplasm of the large bowel, stenotic coronary artery disease, angina pectoris, gout, and obesity. Family history included malignant pharyngeal neoplasm, chronic lymphocytic leukemia, and chronic liver disease.

WO 02/08399

PCT/US01/23092

Table 6 (cont.)

Library	Vector	Library Description
PROST018	pINCY	Library was constructed using RNA isolated from diseased prostate tissue removed from a 58-year-old Caucasian male during a radical cystectomy, radical prostatectomy, and gastrectomy. Pathology indicated adenofibromatous hyperplasia; this tissue was associated with a grade 3 transitional cell carcinoma. Patient history included angina and emphysema. Family history included acute myocardial infarction, atherosclerotic coronary artery disease, and type II diabetes.
PROST023	pINCY	This subtracted prostate tumor library was constructed using 10 million clones from a pooled prostate tumor library that was subjected to 2 rounds of subtractive hybridization with 10 million clones from a pooled prostate tissue library. The starting library for subtraction was constructed by pooling equal numbers of clones from 4 prostate tumor libraries using MASH isolated from prostate tumor removed from Caucasian patients aged 57 (A), 66 (B), 68 (C), and 68 (D) during prostatectomy with lymph node excision. Pathology indicated adenocarcinoma in all cases. History included elevated PSA, induration and tobacco abuse in donor A, elevated PSA, induration, prostate hyperplasia, renal failure, osteoarthritis, renal artery stenosis, benign prostatic hyperthrombocytopenia, hyperlipidemia, tobacco/alcohol abuse and hepatitis C (carrier) in donor B; elevated PSA, induration, and kidney calculus in donor C; and elevated PSA, induration, hypercholesterolemia, and kidney calculus in donor D. The hybridization probe for subtraction was constructed by pooling equal numbers of cDNA clones from 3 prostate tissue libraries derived from prostate tissue, prostate epithelial cells, and fibroblasts from prostate stroma from 3 different donors. Subtractive hybridization conditions were based on the methodologies of Swarcop et al., NAR 19 (1991):1954 and Ronaldo, et al., Genome Research 6 (1996):791.
SKIN00001	PCDNA2.1	Random-primed library was constructed using RNA isolated from skin tissue removed from the breast of a 17-year-old Caucasian female during bilateral reduction mammoplasty. Patient history included breast hypertrophy. Family history included benign hypertension.
SINIT003	pINCY	Library was constructed using RNA isolated from ileal tumor tissue obtained from a 49-year-old Caucasian female during destruction of peritoneal tissue, peritoneal adhesiolysis, ileum resection, and permanent colostomy. Pathology indicated grade 4 adenocarcinoma. Patient history included benign hypertension. Previous surgeries included total abdominal hysterectomy, bilateral salpingo-oophorectomy, regional lymph node excision, an incidental appendectomy, and dilation and curettage. Family history included benign hypertension, cerebrovascular disease, hyperlipidemia, atherosclerotic coronary artery disease, hyperlipidemia, type II diabetes, and stomach cancer.
SITN00001	PCDNA2.1	This random primed library was constructed using RNA isolated from small intestine tissue removed from a 31-year-old Caucasian female during Roux-en-Y gastric bypass. Patient history included clinical obesity.

Table 6 (cont.)

Library	Vector	Library Description
THYROID03	P1NGY	Library was constructed using RNA isolated from thyroid tissue removed from the left thyroid of a 28-year-old Caucasian female during a complete thyroidectomy. Pathology indicated a small nodule of adenomatous hyperplasia present in the left thyroid. Pathology for the associated tumor tissue indicated dominant follicular adenoma, forming a well-encapsulated mass in the left thyroid.

Table 7

Program	Description	Reference	Parameter/Threshold
ABI FACTURA	A program that removes vector sequences and marks 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 $\leq 50\%$
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: 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.	ESY: 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 at least five functions: fasta, tfasta, fastx, tblastx, and search.	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.	ESY: fasta E value=1.0E-6 Assembled ESY: 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
BLMPS	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. 19:5565-5572; Henikoff, J.G. and S. Henikoff (1996) Methods Enzymol. 265:88-105; and Altschul, S.F. et al. (1997) J. Chem. Inf. Comput. Sci. 37:471-474.	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. (1988) Nucleic Acids Res. 26:320-322; Durbin, R. et al. (1998) Our Word View, in a Nucleic Acid Database, Cambridge Univ. Press, pp. 1-330.	PFAM hits: Probability value= 1.0E-5 or less Signal peptide hits: Score= 0 or greater

WO 02/08399

PCT/US01/23092

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.	Gribkov, M. et al. (1988) C.A.BIOS 4:61-66; Gribkov, M. et al. (1989) Methods Enzymol. 183:146-159; Baroch, A. et al. (1997) Nucleic Acids Res. 25:217-221.	Normalized quality score=CCG-specified "HIGH" value for that particular Prosite motif. Generally, score=1.4-2.1.
Phred	A base-calling algorithm that examines automated sequencer traces with high sensitivity and probability.	Ewing, B. et al. (1998) Genome Res. 8:175-185; Ewing, B. and P. Green (1998) Genome Res. 8:186-194.	
Phrap	A Phis 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 (1983) 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.	
SFSscan	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. 5:563-571.	
TMEMMER	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 Int. 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).	Baroch, A. et al. (1997) Nucleic Acids Res. 25:217-221; Wisconsin Package Program Manual, version 9, page M51-59, Genetics Computer Group, Madison, WI.	

WO 02/08399

PCT/US01/23092

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
- 5 SEQ ID NO:1-20,
 - b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-20,
 - c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-20, and
- 10 d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-20.
2. An isolated polypeptide of claim 1 selected from the group consisting of SEQ ID NO:1-20.
- 15 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
- 20 NO:21-40.
6. A recombinant polynucleotide comprising a promoter sequence operably linked to a polynucleotide of claim 3.
- 25 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
- 30 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.

WO 02/08399

PCT/US01/23092

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
 - 5 of SEQ ID NO:21-40,
 - b) a polynucleotide comprising a naturally occurring polynucleotide sequence at least 90% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:21-40,
 - c) a polynucleotide complementary to a polynucleotide of a),
 - d) a polynucleotide complementary to a polynucleotide of b), and
 - 10 e) an RNA equivalent of a)-d).
12. An isolated polynucleotide comprising at least 60 contiguous nucleotides of a polynucleotide of claim 11.
- 15 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
 - 20 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.
- 25 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
 - 30 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

WO 02/08399

PCT/US01/23092

excipient.

17. A composition of claim 16, wherein the polypeptide has an amino acid sequence selected from the group consisting of SEQ ID NO:1-20.

5

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

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.

25

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 PKIN, comprising administering to a patient in need of such treatment a composition of claim 23.

30

25. A method of screening for a compound that specifically binds to the polypeptide of claim

WO 02/08399

PCT/US01/23092

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
5 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
10 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
15 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
20 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
25 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
30 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;

WO 02/08399

PCT/US01/23092

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.

5

29. A diagnostic test for a condition or disease associated with the expression of PKIN 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

10 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:

- 15 a) a chimeric antibody,
b) a single chain antibody,
c) a Fab fragment,
d) a F(ab')₂ fragment, or
e) a humanized antibody.

20 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 PKIN in a subject, comprising administering to said subject an effective amount of the composition of claim 31.

25 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 PKIN in a subject, comprising administering to said subject an effective amount of the composition of claim 33.

30 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:1-20, or an immunogenic fragment thereof, under conditions to elicit

WO 02/08399

PCT/US01/23092

- 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:1-20.
- 5
36. An antibody produced by a method of claim 35.
37. A composition comprising the antibody of claim 36 and a suitable carrier.
- 10
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:1-20, 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:1-20.
- 15
39. A monoclonal antibody produced by a method of claim 38.
- 20
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.
- 30
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

WO 02/08399

PCT/US01/23092

group consisting of SEQ ID NO:1-20 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
- 5 having an amino acid sequence selected from the group consisting of SEQ ID NO:1-20 in the sample.

44. A method of purifying a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-20 from a sample, the method comprising:

- a) incubating the antibody of claim 10 with a sample under conditions to allow specific binding
- 10 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:1-20.

15 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.

20 48. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:4.

49. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:5.

25 50. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:6.

51. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:7.

52. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:8.

30 53. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:9.

54. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:10.

WO 02/08399

PCT/US01/23092

55. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:11.
56. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:12.
- 5 57. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:13.
58. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:14.
59. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:15.
- 10 60. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:16.
61. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:17.
- 15 62. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:18.
63. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:19.
64. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:20.
- 20 65. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:21.
66. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:22.
- 25 67. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:23.
68. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:24.
69. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:25.
- 30 70. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:26.
71. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:27.

WO 02/08399

PCT/US01/23092

72. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:28.
73. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:29.
- 5 74. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:30.
75. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:31.
76. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:32.
- 10 77. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:33.
78. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:34.
- 15 79. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:35.
80. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:36.
81. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:37.
- 20 82. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:38.
83. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:39.
- 25 84. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:40.

WO 02/08399

PCT/US01/23092

<110> INCYTE GENOMICS, INC.
YUE, Henry
KHAN, Farrah A.
GURURAJAN, Rajagopal
HAFALIA, April J. A.
WALLA, Nazinder K.
PATTERSON, Chandra
RAMKUMAR, Jayalaxmi
GANDHI, Ameena R.
POLICKY, Jennifer L.
BAUGHN, Mariah R.
TRIBOULEY, Catherine M.
THORNTON, Michael
BANDMAN, Olga
NGUYEN, Dannel B.
LU, Yan
BURFORD, Neil
LAL, Preeti
DING, Li
YAO, Monique G.
ELLIOTT, Vicki S.
RECIFFON, Shirley A.
KEARNEY, Liam
LU, Dying Aina M.
GREENWALD, Sara R.
TANG, Y. Tom
XU, Yuming
WALSH, Roderick T.
GIETZEN, Kimberly J.
YANG, Junming
HILLMAN, Jennifer L.

<120> HUMAN KINASES

<130> PI-0162 PCT

<140> To Be Assigned
<141> Herewith

<150> 60/220,038; 60/222,112; 60/222,831; 60/224,729
<151> 2000-07-21; 2000-07-28; 2000-08-04; 2000-08-11

<160> 40
<170> PERL Program

<210> 1
<211> 1297
<212> ERT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2564295CD1

<400> 1
Met Ala Val Pro Ser Leu Trp Pro Trp Gly Ala Cys Leu Pro Val

WO 02/08399

PCT/US01/23092

	20		25		30
Pro Ser Leu Asp	Ile Arg Ser Glu Val Ala Glu Leu Arg Gln Leu				
	35		40		45
Glu Asn Cys Ser	Val Val Glu Gly His Leu Gln Ile Leu Leu Met				
	50		55		60
Phe Thr Ala Thr	Gly Glu Asp Phe Arg Gly Leu Ser Phe Pro Arg				
	65		70		75
Leu Thr Gln Val	Thr Asp Tyr Leu Leu Leu Phe Arg Val Tyr Gly				
	80		85		90
Leu Glu Ser Leu	Arg Asp Leu Phe Pro Asn Leu Ala Val Ile Arg				
	95		100		105
Gly Thr Arg Leu	Phe Leu Gly Tyr Ala Leu Val Ile Phe Glu Met				
	110		115		120
Pro His Leu Arg	Asp Val Ala Leu Pro Ala Leu Gly Ala Val Leu				
	125		130		135
Arg Gly Ala Val	Arg Val Glu Lys Asn Gln Glu Leu Cys His Leu				
	140		145		150
Ser Thr Ile Asp	Trp Gly Leu Leu Gln Pro Ala Pro Gly Ala Asn				
	155		160		165
His Ile Val Gly	Asn Lys Leu Gly Glu Glu Cys Ala Asp Val Cys				
	170		175		180
Pro Gly Val Leu	Gly Ala Ala Gly Glu Pro Cys Ala Lys Thr Thr				
	185		190		195
Phe Ser Gly His	Thr Asp Tyr Arg Cys Trp Thr Ser Ser His Cys				
	200		205		210
Gln Arg Val Cys	Pro Cys Pro His Gly Met Ala Cys Thr Ala Arg				
	215		220		225
Gly Glu Cys Cys	His Thr Glu Cys Leu Gly Gly Cys Ser Gln Pro				
	230		235		240
Glu Asp Pro Arg	Ala Cys Val Ala Cys Arg His Leu Tyr Phe Gln				
	245		250		255
Gly Ala Cys Leu	Trp Ala Cys Pro Pro Gly Thr Tyr Gln Tyr Glu				
	260		265		270
Ser Trp Arg Cys	Val Thr Ala Glu Arg Cys Ala Ser Leu His Ser				
	275		280		285
Val Pro Gly Arg	Ala Ser Thr Phe Gly Ile His Gln Gly Ser Cys				
	290		295		300
Leu Ala Gln Cys	Pro Ser Gly Phe Thr Arg Asn Ser Ser Ser Ile				
	305		310		315
Phe Cys His Lys	Cys Glu Gly Leu Cys Pro Lys Glu Cys Lys Val				
	320		325		330
Gly Thr Lys Thr	Ile Asp Ser Ile Gln Ala Ala Gln Asp Leu Val				
	335		340		345
Gly Cys Thr His	Val Glu Gly Ser Leu Ile Leu Asn Leu Arg Gln				
	350		355		360
Gly Tyr Asn Leu	Glu Pro Gln Leu Gln His Ser Leu Gly Leu Val				
	365		370		375
Glu Thr Ile Thr	Gly Phe Leu Lys Ile Lys His Ser Phe Ala Leu				
	380		385		390
Val Ser Leu Gly	Phe Phe Lys Asn Leu Lys Leu Ile Arg Gly Asp				
	395		400		405
Ala Met Val Asp	Gly Asn Tyr Thr Leu Tyr Val Leu Asp Asn Gln				
	410		415		420
Asn Leu Gln Gln	Leu Gly Ser Trp Val Ala Ala Gly Leu Thr Ile				
	425		430		435
Pro Val Gly Lys	Ile Tyr Phe Ala Phe Asn Pro Arg Leu Cys Leu				

WO 02/08399

PCT/US01/23092

	440		445		450
Glu His Ile Tyr Arg Leu Glu Glu Val Thr Gly Thr Arg Gly Arg					
	455		460		465
Gln Asn Lys Ala Glu Ile Asn Pro Arg Thr Asn Gly Asp Arg Ala					
	470		475		480
Ala Cys Gln Thr Arg Thr Leu Arg Phe Val Ser Asn Val Thr Glu					
	485		490		495
Ala Asp Arg Ile Leu Leu Arg Trp Glu Arg Tyr Glu Pro Leu Glu					
	500		505		510
Ala Arg Asp Leu Leu Ser Phe Ile Val Tyr Tyr Lys Glu Ser Pro					
	515		520		525
Phe Gln Asn Ala Thr Glu His Val Gly Pro Asp Ala Cys Gly Thr					
	530		535		540
Gln Ser Trp Asn Leu Leu Asp Val Glu Leu Pro Leu Ser Arg Thr					
	545		550		555
Gln Glu Pro Gly Val Thr Leu Ala Ser Leu Lys Pro Trp Thr Gln					
	560		565		570
Tyr Ala Val Phe Val Arg Ala Ile Thr Leu Thr Thr Glu Glu Asp					
	575		580		585
Ser Pro His Gln Gly Ala Gln Ser Pro Ile Val Tyr Leu Arg Thr					
	590		595		600
Leu Pro Ala Ala Pro Thr Val Pro Gln Asp Val Ile Ser Thr Ser					
	605		610		615
Asn Ser Ser Ser His Leu Leu Val Arg Trp Lys Pro Pro Thr Gln					
	620		625		630
Arg Asn Gly Asn Leu Thr Tyr Tyr Leu Val Leu Trp Gln Arg Leu					
	635		640		645
Ala Glu Asp Gly Asp Leu Tyr Leu Asn Asp Tyr Cys His Arg Gly					
	650		655		660
Leu Arg Leu Pro Thr Ser Asn Asn Asp Pro Arg Phe Asp Gly Glu					
	665		670		675
Asp Gly Asp Pro Glu Ala Glu Met Glu Ser Asp Cys Cys Pro Cys					
	680		685		690
Gln His Pro Pro Pro Gly Gln Val Leu Pro Pro Leu Glu Ala Gln					
	695		700		705
Glu Ala Ser Phe Gln Lys Lys Phe Glu Asn Phe Leu His Asn Ala					
	710		715		720
Ile Thr Ile Pro Ile Ser Pro Trp Lys Val Thr Ser Ile Asn Lys					
	725		730		735
Ser Pro Gln Arg Asp Ser Gly Arg His Arg Arg Ala Ala Gly Pro					
	740		745		750
Leu Arg Leu Gly Gly Asn Ser Ser Asp Phe Glu Ile Gln Glu Asp					
	755		760		765
Lys Val Pro Arg Glu Arg Ala Val Leu Ser Gly Leu Arg His Phe					
	770		775		780
Thr Glu Tyr Arg Ile Asp Ile His Ala Cys Asn His Ala Ala His					
	785		790		795
Thr Val Gly Cys Ser Ala Ala Thr Phe Val Phe Ala Arg Thr Met					
	800		805		810
Pro His Arg Glu Ala Asp Gly Ile Pro Gly Lys Val Ala Trp Glu					
	815		820		825
Ala Ser Ser Lys Asn Ser Val Leu Leu Arg Trp Leu Glu Pro Pro					
	830		835		840
Asp Pro Asn Gly Leu Ile Leu Lys Tyr Glu Ile Lys Tyr Arg Arg					
	845		850		855
Leu Gly Glu Glu Ala Thr Val Leu Cys Val Ser Arg Leu Arg Tyr					

WO 02/08399

PCT/US01/23092

	860		865		870									
Ala	Lys	Phe	Gly	Gly	Val	His	Leu	Ala	Leu	Leu	Pro	Pro	Gly	Asn
	875													885
Tyr	Ser	Ala	Arg	Val	Arg	Ala	Thr	Ser	Leu	Ala	Gly	Asn	Gly	Ser
	890								895					900
Trp	Thr	Asp	Ser	Val	Ala	Phe	Tyr	Ile	Leu	Gly	Pro	Glu	Glu	Glu
	905								910					915
Asp	Ala	Gly	Gly	Leu	His	Val	Leu	Leu	Thr	Ala	Thr	Pro	Val	Gly
	920								925					930
Leu	Thr	Leu	Leu	Ile	Val	Leu	Ala	Ala	Leu	Gly	Phe	Phe	Tyr	Gly
	935								940					945
Lys	Lys	Arg	Asn	Arg	Thr	Leu	Tyr	Ala	Ser	Val	Asn	Pro	Glu	Tyr
	950								955					960
Phe	Ser	Ala	Ser	Asp	Met	Tyr	Val	Pro	Asp	Glu	Trp	Glu	Val	Pro
	965								970					975
Arg	Glu	Gln	Ile	Ser	Ile	Ile	Arg	Glu	Leu	Gly	Gln	Gly	Ser	Phe
	980								985					990
Gly	Met	Val	Tyr	Glu	Gly	Leu	Ala	Arg	Gly	Leu	Glu	Ala	Gly	Glu
	995								1000					1005
Glu	Ser	Thr	Pro	Val	Ala	Leu	Lys	Thr	Val	Asn	Glu	Leu	Ala	Ser
	1010								1015					1020
Pro	Arg	Glu	Cys	Ile	Glu	Phe	Leu	Lys	Glu	Ala	Ser	Val	Met	Lys
	1025								1030					1035
Ala	Phe	Lys	Cys	His	His	Val	Val	Arg	Leu	Leu	Gly	Val	Val	Ser
	1040								1045					1050
Gln	Gly	Gln	Pro	Thr	Leu	Val	Ile	Met	Glu	Leu	Met	Thr	Arg	Gly
	1055								1060					1065
Asp	Leu	Lys	Ser	His	Leu	Arg	Ser	Leu	Arg	Pro	Glu	Ala	Glu	Asn
	1070								1075					1080
Asn	Pro	Gly	Leu	Pro	Gln	Pro	Ala	Leu	Gly	Glu	Met	Ile	Gln	Met
	1085								1090					1095
Ala	Gly	Glu	Ile	Ala	Asp	Gly	Met	Ala	Tyr	Leu	Ala	Ala	Asn	Lys
	1100								1105					1110
Phe	Val	His	Arg	Asp	Leu	Ala	Ala	Arg	Asn	Cys	Met	Val	Ser	Gln
	1115								1120					1125
Asp	Phe	Thr	Val	Lys	Ile	Gly	Asp	Phe	Gly	Met	Thr	Arg	Asp	Val
	1130								1135					1140
Tyr	Glu	Thr	Asp	Tyr	Tyr	Arg	Lys	Gly	Gly	Lys	Gly	Leu	Leu	Pro
	1145								1150					1155
Val	Arg	Trp	Met	Ala	Pro	Glu	Ser	Leu	Lys	Asp	Gly	Ile	Phe	Thr
	1160								1165					1170
Thr	His	Ser	Asp	Val	Trp	Ser	Phe	Gly	Val	Val	Leu	Trp	Glu	Ile
	1175								1180					1185
Val	Thr	Leu	Ala	Glu	Gln	Pro	Tyr	Gln	Gly	Leu	Ser	Asn	Glu	Gln
	1190								1195					1200
Val	Leu	Lys	Phe	Val	Met	Asp	Gly	Gly	Val	Leu	Glu	Glu	Leu	Glu
	1205								1210					1215
Gly	Cys	Pro	Leu	Gln	Leu	Gln	Glu	Leu	Met	Ser	Arg	Cys	Trp	Gln
	1220								1225					1230
Pro	Asn	Pro	Arg	Leu	Arg	Pro	Ser	Phe	Thr	His	Ile	Leu	Asp	Ser
	1235								1240					1245
Ile	Gln	Glu	Glu	Leu	Arg	Pro	Ser	Phe	Arg	Leu	Leu	Ser	Phe	Tyr
	1250								1255					1260
Tyr	Ser	Pro	Glu	Cys	Arg	Gly	Ala	Arg	Gly	Ser	Leu	Pro	Thr	Thr
	1265								1270					1275
Asp	Ala	Glu	Pro	Asp	Ser	Ser	Pro	Thr	Pro	Arg	Asp	Cys	Ser	Pro

WO 02/08399

PCT/US01/23092

1280 1285 1290
 Gln Asn Gly Gly Pro Gly His
 1295

<210> 2
 <211> 718
 <212> PRT
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 2837050CD1

<400> 2
 Met Met Glu Glu Leu His Ser Leu Asp Pro Arg Arg Gln Glu Leu
 1 5 10 15
 Leu Glu Ala Arg Phe Thr Arg Val Gly Val Ser Lys Gly Pro Leu
 20 25 30
 Asn Ser Glu Ser Ser Asn Gln Ser Leu Cys Ser Val Gly Ser Leu
 35 40 45
 Ser Asp Lys Glu Val Glu Thr Pro Glu Lys Lys Gln Asn Asp Gln
 50 55 60
 Arg Asn Arg Lys Arg Lys Ala Glu Pro Tyr Glu Thr Ser Gln Gly
 65 70 75
 Lys Gly Thr Pro Arg Gly His Lys Ile Ser Asp Tyr Phe Glu Arg
 80 85 90
 Arg Val Glu Gln Pro Leu Tyr Gly Leu Asp Gly Ser Ala Ala Lys
 95 100 105
 Glu Ala Thr Glu Glu Gln Ser Ala Leu Pro Thr Leu Met Ser Val
 110 115 120
 Met Leu Ala Lys Pro Arg Leu Asp Thr Glu His Val Ala Gln Arg
 125 130 135
 Gly Ala Gly Leu Cys Phe Thr Phe Val Ser Ala Gln Gln Asn Ser
 140 145 150
 Pro Ser Ser Thr Gly Ser Gly Asn Thr Glu His Ser Cys Ser Ser
 155 160 165
 Gln Lys Gln Ile Ser Ile Gln His Arg Gln Thr Gln Ser Asp Leu
 170 175 180
 Thr Ile Glu Lys Ile Ser Ala Leu Glu Asn Ser Lys Asn Ser Asp
 185 190 195
 Leu Glu Lys Lys Glu Gly Arg Ile Asp Asp Leu Leu Arg Ala Asn
 200 205 210
 Cys Asp Leu Arg Arg Gln Ile Asp Glu Gln Gln Lys Met Leu Glu
 215 220 225
 Lys Tyr Lys Glu Arg Leu Asp Arg Cys Val Thr Met Ser Lys Lys
 230 235 240
 Leu Leu Ile Glu Lys Ser Lys Gln Glu Lys Met Ala Cys Arg Asp
 245 250 255
 Lys Ser Met Gln Asp Arg Leu Arg Leu Gly His Phe Thr Thr Val
 260 265 270
 Arg His Gly Ala Ser Phe Thr Glu Gln Trp Thr Asp Gly Tyr Ala
 275 280 285
 Phe Gln Asn Leu Ile Lys Gln Gln Glu Arg Ile Asn Ser Gln Arg
 290 295 300
 Glu Glu Ile Glu Arg Gln Arg Lys Met Leu Ala Lys Arg Lys Pro
 305 310 315

WO 02/08399

PCT/US01/23092

Pro Ala Met Gly Gln Ala Pro Pro Ala Thr Asn Glu Gln Lys Gln
 320 325 330
 Arg Lys Ser Lys Thr Asn Gly Ala Glu Asn Glu Thr Leu Thr Leu
 335 340 345
 Ala Glu Tyr His Glu Gln Glu Glu Ile Phe Lys Leu Arg Leu Gly
 350 355 360
 His Leu Lys Lys Glu Glu Ala Glu Ile Gln Ala Glu Leu Glu Arg
 365 370 375
 Leu Glu Arg Val Arg Asn Leu His Ile Arg Glu Leu Lys Arg Ile
 380 385 390
 His Asn Glu Asp Asn Ser Gln Phe Lys Asp His Pro Thr Leu Asn
 395 400 405
 Asp Arg Tyr Leu Leu His Leu Leu Gly Arg Gly Gly Phe Ser
 410 415 420
 Glu Val Tyr Lys Ala Phe Asp Leu Thr Glu Gln Arg Tyr Val Ala
 425 430 435
 Val Lys Ile His Gln Leu Asn Lys Asn Trp Arg Asp Glu Lys Lys
 440 445 450
 Glu Asn Tyr His Lys His Ala Cys Arg Glu Tyr Arg Ile His Lys
 455 460 465
 Glu Leu Asp His Pro Arg Ile Val Lys Leu Tyr Asp Tyr Phe Ser
 470 475 480
 Leu Asp Thr Asp Ser Phe Cys Thr Val Leu Glu Tyr Cys Glu Gly
 485 490 495
 Asn Asp Leu Asp Phe Tyr Leu Lys Gln His Lys Leu Met Ser Glu
 500 505 510
 Lys Glu Ala Trp Ser Ile Ile Met Gln Ile Val Asn Ala Leu Lys
 515 520 525
 Tyr Leu Asn Glu Ile Lys Pro Pro Ile Ile His Tyr Asp Leu Lys
 530 535 540
 Pro Gly Asn Ile Leu Leu Val Asn Gly Thr Val Cys Gly Glu Arg
 545 550 555
 Lys Ile Thr Asp Phe Gly Leu Ser Lys Ile Met Asp Asp Asp Ser
 560 565 570
 Tyr Asn Ser Val Gly Gly Met Glu Leu Thr Ser Gln Gly Ala Gly
 575 580 585
 Thr Tyr Trp Tyr Leu Pro Pro Glu Cys Phe Val Val Glu Lys Glu
 590 595 600
 Pro Pro Lys Ile Ser Asn Lys Val Asp Val Trp Ser Val Gly Val
 605 610 615
 Ile Phe Tyr Gln Cys Leu Ser Gly Gly Lys Pro Phe Gly His Asn
 620 625 630
 Gln Ser Gln Gln Asp Ile Leu Gln Glu Asn Thr Ile Leu Lys Ala
 635 640 645
 Ala Glu Val Gln Phe Pro Pro Lys Pro Val Val Thr Pro Glu Ala
 650 655 660
 Lys Ala Phe Ile Arg Arg Cys Leu Ala Tyr Arg Lys Glu Asp Cys
 665 670 675
 Ile Asp Ala Gln Gln Leu Ala Cys Asp Pro Tyr Leu Leu Pro His
 680 685 690
 Ile Arg Lys Ser Val Ser Thr Ser Ser Pro Ala Gly Ala Ala Ile
 695 700 705
 Ala Ser Thr Ser Gly Ala Ser Asn Asn Ser Ser Ser Asn
 710 715

<210> 3

WO 02/08399

PCT/US01/23092

<211> 497
 <212> PRT
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 7474590CD1

<400> 3

```

Met Tyr Ser Asp Ser Glu Asp Glu Ser Ser Ser Glu Leu Ser Thr Val
1      5      10      15
Leu Ser Met Phe Glu Glu Lys Glu Phe Thr Arg Gln Tyr Thr Val
20     25     30
Leu Lys Thr Leu Ser Gln His Gly Thr Thr Glu Val Arg Leu Cys
35     40     45
Ser His His Leu Thr Gly Val Thr Val Ala Val Lys Ala Leu Lys
50     55     60
Tyr Gln Arg Trp Trp Glu Pro Lys Val Ser Glu Val Glu Ile Met
65     70     75
Lys Met Leu Ser His Pro Asn Ile Val Ser Leu Leu Gln Val Ile
80     85     90
Glu Thr Glu Gln Asn Ile Tyr Leu Ile Met Glu Val Ala Gln Gly
95     100    105
Thr Gln Leu His Asn Arg Val Gln Glu Ala Arg Cys Leu Lys Glu
110    115    120
Asp Glu Ala Arg Ser Ile Phe Val Gln Leu Leu Ser Ala Ile Gly
125    130    135
Tyr Cys His Gly Glu Gly Val Val His Arg Asp Leu Lys Pro Asp
140    145    150
Asn Val Ile Val Asp Glu His Gly Asn Val Lys Ile Val Asp Phe
155    160    165
Gly Leu Gly Ala Arg Phe Met Pro Gly Gln Lys Leu Glu Arg Leu
170    175    180
Cys Gly Ala Phe Gln Phe Ile Pro Pro Glu Ile Phe Leu Gly Leu
185    190    195
Pro Tyr Asp Gly Pro Lys Val Asp Ile Trp Ala Leu Gly Val Leu
200    205    210
Leu Tyr Tyr Met Val Thr Gly Ile Phe Pro Phe Val Gly Ser Thr
215    220    225
Leu Ser Glu Ile Ser Lys Glu Val Leu Gln Gly Arg Tyr Glu Ile
230    235    240
Pro Tyr Asn Leu Ser Lys Asp Leu Arg Ser Met Ile Gly Leu Leu
245    250    255
Leu Ala Thr Asn Ala Arg Gln Arg Pro Thr Ala Gln Asp Leu Leu
260    265    270
Ser His Pro Trp Leu Gln Glu Gly Glu Lys Thr Ile Thr Phe His
275    280    285
Ser Asn Gly Asp Thr Ser Phe Pro Asp Pro Asp Ile Met Ala Ala
290    295    300
Met Lys Asn Ile Gly Phe His Val Gln Asp Ile Arg Glu Ser Leu
305    310    315
Lys His Arg Lys Phe Asp Glu Thr Met Ala Thr Tyr Asn Leu Leu
320    325    330
Arg Ala Glu Ala Cys Gln Asp Asp Gly Asn Tyr Val Gln Thr Lys
335    340    345
Leu Met Asn Pro Gly Met Pro Pro Phe Pro Ser Val Thr Asp Ser

```

WO 02/08399

PCT/US01/23092

350 355 360
 Gly Ala Phe Ser Leu Pro Pro Arg Arg Arg Ala Ser Glu Pro Ser
 365 370 375
 Phe Lys Val Leu Val Ser Ser Thr Glu Glu His Gln Leu Arg Gln
 380 385 390
 Thr Gly Gly Thr Asn Ala Pro Phe Pro Pro Lys Lys Thr Pro Thr
 395 400 405
 Met Gly Arg Ser Gln Lys Gln Lys Arg Ala Met Thr Ala Pro Cys
 410 415 420
 Ile Cys Leu Leu Arg Asn Thr Tyr Ile Asp Thr Glu Asp Ser Ser
 425 430 435
 Phe Cys Thr Ser Ser Gln Ala Glu Lys Thr Ser Ser Asp Pro Glu
 440 445 450
 Lys Ser Glu Thr Ser Thr Ser Cys Pro Leu Thr Pro Arg Gly Trp
 455 460 465
 Arg Lys Trp Lys Lys Arg Ile Val Ala Cys Ile Gln Thr Leu Cys
 470 475 480
 Cys Cys Thr Leu Pro Gln Lys Lys Cys Pro Arg Ser Val His Pro
 485 490 495
 Gln Lys

<210> 4
 <211> 741
 <212> PRT
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 7474594CD1

<400> 4
 Met Ser Gly Leu Val Leu Met Leu Ala Ala Arg Cys Ile Val Gly
 1 5 10 15
 Ser Ser Pro Leu Cys Arg Cys Arg Arg Arg Pro Arg Arg Ile
 20 25 30
 Gly Ala Gly Pro Gly Arg Asp Asp Pro Gly Arg Lys Ala Ala Ala
 35 40 45
 Ala Gly Gly Ser Gly Ser Pro Asn Ala Ala Leu Ser Arg Pro Arg
 50 55 60
 Pro Ala Pro Ala Pro Gly Asp Ala Pro Pro Arg Ala Ala Ala Ser
 65 70 75
 Ala Ala Ala Ala Ala Ala Ala Ala Ala Gly Thr Glu Gln Val Asp
 80 85 90
 Gly Pro Leu Arg Ala Gly Pro Ala Asp Thr Pro Pro Ser Gly Trp
 95 100 105
 Arg Met Gln Cys Leu Ala Ala Ala Leu Lys Asp Glu Thr Asn Met
 110 115 120
 Ser Gly Gly Gly Glu Gln Ala Asp Ile Leu Pro Ala Asn Tyr Val
 125 130 135
 Val Lys Asp Arg Trp Lys Val Leu Lys Lys Ile Gly Gly Gly Gly
 140 145 150
 Phe Gly Glu Ile Tyr Glu Ala Met Asp Leu Leu Thr Arg Glu Asn
 155 160 165
 Val Ala Leu Lys Val Glu Ser Ala Gln Gln Pro Lys Gln Val Leu
 170 175 180

WO 02/08399

PCT/US01/23092

Lys Met Glu Val Ala Val Leu Lys Lys Leu Gln Gly Lys Asp His
 185 190 195
 Val Cys Arg Phe Ile Gly Cys Gly Arg Asn Glu Lys Phe Asn Tyr
 200 205 210
 Val Val Met Gln Leu Gln Gly Arg Asn Leu Ala Asp Leu Arg Arg
 215 220 225
 Ser Gln Pro Arg Gly Thr Phe Thr Leu Ser Thr Thr Leu Arg Leu
 230 235 240
 Gly Lys Gln Ile Leu Glu Ser Ile Glu Ala Ile His Ser Val Gly
 245 250 255
 Phe Leu His Arg Asp Ile Lys Pro Ser Asn Phe Ala Met Gly Arg
 260 265 270
 Leu Pro Ser Thr Tyr Arg Lys Cys Tyr Met Leu Asp Phe Gly Leu
 275 280 285
 Ala Arg Gln Tyr Thr Asn Thr Thr Gly Asp Val Arg Pro Pro Arg
 290 295 300
 Asn Val Ala Gly Phe Arg Gly Thr Val Arg Tyr Ala Ser Val Asn
 305 310 315
 Ala His Lys Asn Arg Glu Met Gly Arg His Asp Asp Leu Trp Ser
 320 325 330
 Leu Phe Tyr Met Leu Val Glu Phe Ala Val Gly Gln Leu Pro Trp
 335 340 345
 Arg Lys Ile Lys Asp Lys Glu Gln Val Gly Met Ile Lys Glu Lys
 350 355 360
 Tyr Glu His Arg Met Leu Leu Lys His Met Pro Ser Glu Phe His
 365 370 375
 Leu Phe Leu Asp His Ile Ala Ser Leu Asp Tyr Phe Thr Lys Pro
 380 385 390
 Asp Tyr Gln Leu Ile Met Ser Val Phe Glu Asn Ser Met Lys Glu
 395 400 405
 Arg Gly Ile Ala Glu Asn Glu Ala Phe Asp Trp Glu Lys Ala Gly
 410 415 420
 Thr Asp Ala Leu Leu Ser Thr Ser Thr Ser Thr Pro Pro Gln Gln
 425 430 435
 Asn Thr Arg Gln Thr Ala Ala Met Phe Gly Val Val Asn Val Thr
 440 445 450
 Pro Val Pro Gly Asp Leu Leu Arg Glu Asn Thr Glu Asp Val Leu
 455 460 465
 Gln Gly Glu His Leu Ser Asp Gln Glu Asn Ala Pro Pro Ile Leu
 470 475 480
 Pro Gly Arg Pro Ser Glu Gly Leu Gly Pro Ser Pro His Leu Val
 485 490 495
 Pro His Pro Gly Gly Pro Glu Ala Glu Val Trp Glu Glu Thr Asp
 500 505 510
 Val Asn Arg Asn Lys Leu Arg Ile Asn Ile Gly Lys Val Thr Ala
 515 520 525
 Ala Arg Ala Lys Gly Val Gly Gly Leu Phe Ser His Pro Arg Phe
 530 535 540
 Pro Ala Leu Cys Pro Cys Pro Val Pro Pro Lys His Pro Val Pro
 545 550 555
 Gly His Leu Pro Ala Cys Pro Ala Ser Val Ser Arg Ser Leu Pro
 560 565 570
 Ala Leu Ala Ser Leu Cys Leu Pro Ser Ser Ser Ser Val Ser
 575 580 585
 Phe Thr Leu Arg Arg Pro Ser Ala His Ser Arg Leu Ile Ser Pro
 590 595 600

WO 02/08399

PCT/US01/23092

Ser Ser Trp His Ser Pro Leu Leu Gln Ser Pro Cys Val Glu Glu
605 610 615
Glu Gln Ser Arg Gly Met Gly Val Pro Ser Ser Pro Val Arg Ala
620 625 630
Pro Pro Asp Ser Pro Thr Thr Pro Val Arg Ser Leu Arg Tyr Arg
635 640 645
Arg Val Asn Ser Pro Glu Ser Glu Arg Leu Ser Thr Ala Asp Gly
650 655 660
Arg Val Glu Leu Pro Glu Arg Arg Trp Val Trp Gly Gln Gly His
665 670 675
Gly Trp Gly Pro Arg Pro Ser Pro Pro Ser Arg Gly Trp Ser Gly
680 685 690
Gly Lys Val Arg Cys Val Ala Glu Val Gly Arg Pro Trp Glu Val
695 700 705
Leu Arg Gly Leu Tyr Leu Gly Leu Gly Ser Asp Ser Val Gly Ala
710 715 720
Arg Asp Arg Ala Trp Glu Asn Gln Trp Gly Ile Gln Arg Gly Pro
725 730 735
Gly Ser Cys Gln Glu Thr
740

<210> 5

<211> 645

<212> PFT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 7477585CD1

<400> 5

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

WO 02/08399

PCT/US01/23092

	185		190		195									
Pro	His	Tyr	Met	Ser	Pro	Glu	Ala	Leu	Lys	His	Gln	Gly	Tyr	Asp
	200								205					210
Thr	Lys	Ser	Asp	Ile	Trp	Ser	Leu	Ala	Cys	Ile	Leu	Tyr	Glu	Met
	215								220					225
Cys	Cys	Met	Asn	His	Ala	Phe	Ala	Gly	Ser	Asn	Phe	Leu	Ser	Ile
	230								235					240
Val	Leu	Lys	Ile	Val	Glu	Gly	Asp	Thr	Pro	Ser	Leu	Pro	Glu	Arg
	245								250					255
Tyr	Pro	Lys	Glu	Leu	Asn	Ala	Ile	Met	Glu	Ser	Met	Leu	Asn	Lys
	260								265					270
Asn	Pro	Ser	Leu	Arg	Pro	Ser	Ala	Ile	Glu	Ile	Leu	Lys	Ile	Pro
	275								280					285
Tyr	Leu	Asp	Glu	Gln	Leu	Gln	Asn	Leu	Met	Cys	Arg	Tyr	Ser	Glu
	290								295					300
Met	Thr	Leu	Glu	Asp	Lys	Asn	Leu	Asp	Cys	Gln	Lys	Glu	Ala	Ala
	305								310					315
His	Ile	Ile	Asn	Ala	Met	Gln	Lys	Arg	Ile	His	Leu	Gln	Thr	Leu
	320								325					330
Arg	Ala	Leu	Ser	Glu	Val	Gln	Lys	Met	Thr	Pro	Arg	Glu	Arg	Met
	335								340					345
Arg	Leu	Arg	Lys	Leu	Gln	Ala	Ala	Asp	Glu	Lys	Ala	Arg	Lys	Leu
	350								355					360
Lys	Lys	Ile	Val	Glu	Glu	Lys	Tyr	Glu	Glu	Asn	Ser	Lys	Arg	Met
	365								370					375
Gln	Glu	Leu	Arg	Ser	Arg	Asn	Phe	Gln	Gln	Leu	Ser	Val	Asp	Val
	380								385					390
Leu	His	Glu	Lys	Thr	His	Leu	Lys	Gly	Met	Glu	Glu	Lys	Glu	Glu
	395								400					405
Gln	Pro	Glu	Gly	Arg	Leu	Ser	Cys	Ser	Pro	Gln	Asp	Glu	Asp	Glu
	410								415					420
Glu	Arg	Trp	Gln	Gly	Arg	Glu	Glu	Glu	Ser	Asp	Glu	Pro	Thr	Leu
	425								430					435
Glu	Asn	Leu	Pro	Glu	Ser	Gln	Pro	Ile	Pro	Ser	Met	Asp	Leu	His
	440								445					450
Glu	Leu	Glu	Ser	Ile	Val	Glu	Asp	Ala	Thr	Ser	Asp	Leu	Gly	Tyr
	455								460					465
His	Glu	Ile	Pro	Glu	Asp	Pro	Leu	Val	Ala	Glu	Glu	Tyr	Tyr	Ala
	470								475					480
Asp	Ala	Phe	Asp	Ser	Tyr	Cys	Val	Glu	Ser	Asp	Glu	Glu	Glu	Glu
	485								490					495
Glu	Ile	Ala	Leu	Glu	Arg	Pro	Glu	Lys	Glu	Ile	Arg	Asn	Glu	Gly
	500								505					510
Ser	Gln	Pro	Ala	Tyr	Arg	Thr	Asn	Gln	Gln	Asp	Ser	Asp	Ile	Glu
	515								520					525
Ala	Leu	Ala	Arg	Cys	Leu	Glu	Asn	Val	Leu	Gly	Cys	Thr	Ser	Leu
	530								535					540
Asp	Thr	Lys	Thr	Ile	Thr	Thr	Met	Ala	Glu	Asp	Met	Ser	Pro	Gly
	545								550					555
Pro	Pro	Ile	Phe	Asn	Ser	Val	Met	Ala	Arg	Thr	Lys	Met	Lys	Arg
	560								565					570
Met	Arg	Glu	Ser	Ala	Met	Gln	Lys	Leu	Gly	Thr	Glu	Val	Phe	Glu
	575								580					585
Glu	Val	Tyr	Asn	Tyr	Leu	Lys	Arg	Ala	Arg	His	Gln	Asn	Ala	Ser
	590								595					600
Glu	Ala	Glu	Ile	Arg	Glu	Cys	Leu	Glu	Lys	Val	Val	Pro	Gln	Ala

WO 02/08399

PCT/US01/23092

	605	610	615
Ser Asp Cys Phe Glu Val Asp Gln Leu Leu Tyr Phe Glu Glu Gln			
620	625	630	
Leu Leu Ile Thr Met Gly Lys Glu Pro Thr Leu Gln Asn His Leu			
635	640	645	

<210> 6
 <211> 623
 <212> PRT
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 7477587CD1

<400> 6
 Met Trp Ala Pro Gly Thr Arg Gln Gln Gly Gly Pro Glu Met Ala
 1 5 10 15
 His Ile Gln Asn Val Glu Ala His Thr Ser Ser Ala Leu Trp Gly
 20 25 30
 Arg Ser Pro Arg Lys Pro Pro Thr Pro His Ala Arg Glu Ser Leu
 35 40 45
 Ser Phe Pro Leu Glu Arg Pro Arg Ser Gly Arg Ser Ala Val Val
 50 55 60
 Ser Ala Arg Leu Arg Gln Ser Pro Arg Met Glu Pro Arg Pro Arg
 65 70 75
 Arg Arg Arg Arg Ser Arg Pro Leu Val Ala Ala Phe Leu Arg Asp
 80 85 90
 Pro Gly Ser Gly Arg Val Tyr Arg Arg Gly Lys Leu Ile Gly Lys
 95 100 105
 Gly Ala Phe Ser Arg Cys Tyr Lys Leu Thr Asp Met Ser Thr Ser
 110 115 120
 Ala Val Phe Ala Leu Lys Val Val Pro Cys Gly Gly Ala Gly Ala
 125 130 135
 Gly Trp Leu Arg Pro Gln Gly Lys Val Glu Arg Glu Ile Ala Leu
 140 145 150
 His Ser Arg Leu Arg Pro Arg Asn Ile Val Ala Phe His Gly His
 155 160 165
 Phe Ala Asp Arg Asp His Val Tyr Met Val Leu Glu Tyr Cys Ser
 170 175 180
 Arg Gln Ser Leu Ala His Val Leu Arg Ala Arg Gln Ile Leu Thr
 185 190 195
 Glu Pro Glu Val Arg Asp Tyr Leu Arg Gly Leu Val Ser Gly Leu
 200 205 210
 Arg Tyr Leu His Gln Arg Cys Ile Leu His Arg Asp Leu Lys Leu
 215 220 225
 Ser Asn Phe Phe Leu Asn Lys Asn Met Glu Val Lys Ile Gly Asp
 230 235 240
 Leu Gly Leu Ala Ala Lys Val Gly Pro Gly Gly Arg Cys His Arg
 245 250 255
 Tyr Thr Val Leu Thr Gly Thr Pro Pro Phe Met Ala Ser Pro Leu
 260 265 270
 Ser Glu Met Tyr Gln Asn Ile Arg Glu Gly His Tyr Pro Glu Pro
 275 280 285
 Ala His Leu Ser Ala Asn Ala Arg Arg Leu Ile Val His Leu Leu

WO 02/08399

PCT/US01/23092

```

290                295                300
Ala Pro Asn Pro Ala Glu Arg Pro Ser Leu Asp His Leu Leu Gln
305                310                315
Asp Asp Phe Phe Thr Gln Gly Phe Thr Pro Asp Arg Leu Pro Ala
320                325                330
His Ser Cys His Ser Pro Pro Ile Phe Ala Ile Pro Pro Pro Leu
335                340                345
Gly Arg Ile Phe Arg Lys Val Gly Gln Arg Leu Leu Thr Gln Cys
350                355                360
Arg Pro Pro Cys Pro Phe Thr Pro Lys Glu Ala Ser Gly Pro Gly
365                370                375
Glu Gly Gly Pro Asp Pro Asp Ser Met Glu Trp Asp Gly Glu Ser
380                385                390
Ser Leu Ser Ala Lys Glu Val Pro Cys Leu Glu Gly Pro Ile His
395                400                405
Leu Val Ala Gln Gly Thr Leu Gln Ser Asp Leu Ala Ala Thr Gln
410                415                420
Asp Pro Leu Gly Glu Gln Gln Pro Ile Leu Trp Ala Pro Lys Trp
425                430                435
Val Asp Tyr Ser Ser Lys Tyr Gly Phe Gly Tyr Gln Leu Leu Asp
440                445                450
Gly Gly Arg Thr Gly Arg His Pro His Gly Pro Ala Thr Pro Arg
455                460                465
Arg Tyr Leu Leu Ser Thr Tyr Cys Ala His Leu Gln Val Leu Pro
470                475                480
Ala Cys Gln Val Cys Tyr Met Pro Asn Cys Gly Arg Leu Glu Ala
485                490                495
Phe Ala Leu Arg Asp Val Pro Gly Leu Leu Gly Ala Lys Leu Ala
500                505                510
Val Leu Gln Leu Phe Ala Gly Cys Leu Arg Arg Arg Leu Arg Glu
515                520                525
Glu Gly Thr Leu Pro Thr Pro Val Pro Pro Ala Gly Pro Gly Leu
530                535                540
Cys Leu Leu Arg Phe Leu Ala Ser Glu His Ala Leu Leu Leu Leu
545                550                555
Phe Ser Asn Gly Met Val Gln Val Ser Phe Ser Gly Val Pro Ala
560                565                570
Gln Leu Val Leu Ser Gly Glu Gly Glu Gly Leu Gln Leu Thr Leu
575                580                585
Trp Glu Gln Gly Ser Pro Gly Thr Ser Tyr Ser Leu Asp Val Pro
590                595                600
Arg Ser His Gly Cys Ala Pro Thr Thr Gly Gln His Leu His His
605                610                615
Ala Leu Arg Met Leu Gln Ser Ile
620

```

```

<210> 7
<211> 797
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7594537CD1

<400> 7

```

WO 02/08399

PCT/US01/23092

Met Thr Asn Gln Glu Lys Trp Ala His Leu Ser Pro Ser Glu Phe
 1 5 10 15
 Ser Gln Leu Gln Lys Tyr Ala Glu Tyr Ser Thr Lys Lys Leu Lys
 20 25 30
 Asp Val Leu Glu Glu Phe His Gly Asn Gly Val Leu Ala Lys Tyr
 35 40 45
 Asn Pro Glu Gly Thr Ile Asp Phe Glu Gly Phe Lys Leu Phe Met
 50 55 60
 Lys Thr Phe Leu Glu Ala Glu Leu Pro Asp Asp Phe Thr Ala His
 65 70 75
 Leu Phe Met Ser Phe Ser Asn Lys Phe Pro His Ser Ser Pro Met
 80 85 90
 Val Lys Ser Lys Pro Ala Leu Leu Ser Gly Gly Leu Arg Met Asn
 95 100 105
 Lys Gly Ala Ile Thr Pro Pro Arg Thr Thr Ser Pro Ala Asn Thr
 110 115 120
 Cys Ser Pro Glu Val Ile His Leu Lys Asp Ile Val Cys Tyr Leu
 125 130 135
 Ser Leu Leu Glu Arg Gly Arg Pro Glu Asp Lys Leu Glu Phe Met
 140 145 150
 Phe Arg Leu Tyr Asp Thr Asp Gly Asn Gly Phe Leu Asp Ser Ser
 155 160 165
 Glu Leu Glu Asn Ile Ile Ser Gln Met Met His Val Ala Glu Tyr
 170 175 180
 Leu Glu Trp Asp Val Thr Glu Leu Asn Pro Ile Leu His Glu Met
 185 190 195
 Met Glu Glu Ile Asp Tyr Asp His Asp Gly Thr Val Ser Leu Glu
 200 205 210
 Glu Trp Ile Gln Gly Gly Met Thr Thr Ile Pro Leu Leu Val Leu
 215 220 225
 Leu Gly Leu Glu Asn Asn Val Lys Asp Asp Gly Gln His Val Trp
 230 235 240
 Arg Leu Lys His Phe Asn Lys Pro Ala Tyr Cys Asn Leu Cys Leu
 245 250 255
 Asn Met Leu Ile Gly Val Gly Lys Gln Gly Leu Cys Cys Ser Phe
 260 265 270
 Cys Lys Tyr Thr Val His Glu Arg Cys Val Ala Arg Ala Pro Pro
 275 280 285
 Ser Cys Ile Lys Thr Tyr Val Lys Ser Lys Arg Asn Thr Asp Val
 290 295 300
 Met His His Tyr Trp Val Glu Gly Asn Cys Pro Thr Lys Cys Asp
 305 310 315
 Lys Cys His Lys Thr Val Lys Cys Tyr Gln Gly Leu Thr Gly Leu
 320 325 330
 His Cys Val Trp Cys Gln Ile Thr Leu His Asn Lys Cys Ala Ser
 335 340 345
 His Leu Lys Pro Glu Cys Asp Cys Gly Pro Leu Lys Asp His Ile
 350 355 360
 Leu Pro Pro Thr Thr Ile Cys Pro Val Val Leu Gln Thr Leu Pro
 365 370 375
 Thr Ser Gly Val Ser Val Pro Glu Glu Arg Gln Ser Thr Val Lys
 380 385 390
 Lys Glu Lys Ser Gly Ser Gln Gln Pro Asn Lys Val Ile Asp Lys
 395 400 405
 Asn Lys Met Gln Arg Ala Asn Ser Val Thr Val Asp Gly Gln Gly
 410 415 420

WO 02/08399

PCT/US01/23092

Leu Gln Val Thr Pro Val Pro Gly Thr His Pro Leu Leu Val Phe
 425 430 435
 Val Asn Pro Lys Ser Gly Gly Lys Gln Gly Glu Arg Ile Tyr Arg
 440 445 450
 Lys Phe Gln Tyr Leu Leu Asn Pro Arg Gln Val Tyr Ser Leu Ser
 455 460 465
 Gly Asn Gly Pro Met Pro Gly Leu Asn Phe Phe Arg Asp Val Pro
 470 475 480
 Asp Phe Arg Val Leu Ala Cys Gly Gly Asp Gly Thr Val Gly Trp
 485 490 495
 Val Leu Asp Cys Ile Glu Lys Ala Asn Val Gly Lys His Pro Pro
 500 505 510
 Val Ala Ile Leu Pro Leu Gly Thr Gly Asn Asp Leu Ala Arg Cys
 515 520 525
 Leu Arg Trp Gly Gly Tyr Glu Gly Glu Asn Leu Met Lys Ile
 530 535 540
 Leu Lys Asp Ile Glu Asn Ser Thr Glu Ile Met Leu Asp Arg Trp
 545 550 555
 Lys Phe Glu Val Ile Pro Asn Asp Lys Asp Glu Lys Gly Asp Pro
 560 565 570
 Val Pro Tyr Ser Ile Ile Asn Asn Tyr Phe Ser Ile Gly Val Asp
 575 580 585
 Ala Ser Ile Ala His Arg Phe His Ile Met Arg Glu Lys His Pro
 590 595 600
 Glu Lys Phe Asn Ser Arg Met Lys Asn Lys Phe Trp Tyr Phe Glu
 605 610 615
 Phe Gly Thr Ser Glu Thr Phe Ser Ala Thr Cys Lys Lys Leu His
 620 625 630
 Glu Ser Val Glu Ile Glu Cys Asp Gly Val Gln Ile Asp Leu Ile
 635 640 645
 Asn Ile Ser Leu Glu Gly Ile Ala Ile Leu Asn Ile Pro Ser Met
 650 655 660
 His Gly Gly Ser Asn Leu Trp Gly Glu Ser Lys Lys Arg Arg Ser
 665 670 675
 His Arg Arg Ile Glu Lys Lys Gly Ser Asp Lys Arg Thr Thr Val
 680 685 690
 Thr Asp Ala Lys Glu Leu Lys Phe Ala Ser Glu Asp Leu Ser Asp
 695 700 705
 Gln Leu Leu Glu Val Val Gly Leu Glu Gly Ala Met Glu Met Gly
 710 715 720
 Gln Ile Tyr Thr Gly Leu Lys Ser Ala Gly Arg Arg Leu Ala Gln
 725 730 735
 Cys Ser Cys Val Val Ile Arg Thr Ser Lys Ser Leu Pro Met Gln
 740 745 750
 Ile Asp Gly Glu Pro Trp Met Gln Thr Pro Cys Thr Ile Lys Ile
 755 760 765
 Thr His Lys Asn Gln Ala Pro Met Leu Met Gly Pro Pro Pro Lys
 770 775 780
 Thr Gly Leu Phe Cys Ser Leu Val Lys Arg Thr Arg Asn Arg Ser
 785 790 795
 Lys Glu

<210> 8
 <211> 749
 <212> PRT

WO 02/08399

PCT/US01/23092

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 70467491CD1

<400> 8

```

Met Ser Thr Arg Thr Pro Leu Pro Thr Val Asn Glu Arg Asp Thr
1      5      10      15
Glu Asn Ala Val Leu Pro His Thr Ser His Gly Asp Gly Arg Gln
20     25     30
Glu Val Thr Ser Arg Thr Ser Arg Ser Gly Ala Arg Cys Arg Asn
35     40     45
Ser Ile Ala Ser Cys Ala Asp Glu Gln Pro His Ile Gly Asn Tyr
50     55     60
Arg Leu Leu Lys Thr Ile Gly Lys Gly Asn Phe Ala Lys Val Lys
65     70     75
Leu Ala Arg His Ile Leu Thr Gly Arg Glu Lys Asn Val Arg Ile
80     85     90
Ser Lys Glu Ile Asp Asn Phe Leu Gly Lys His Asp Leu Pro Lys
95     100    105
Leu Thr Leu Glu Lys Asn Arg Tyr Thr Ser Val Thr Thr Glu Val
110    115    120
Glu Lys Val Val Asn Ile Leu Pro Asn Leu Glu Phe Met Ile Glu
125    130    135
Phe Phe Glu Ile Tyr Ser Ile Gly Glu Val Phe Asp Tyr Leu Val
140    145    150
Ala His Gly Arg Met Lys Glu Lys Glu Ala Arg Ser Lys Phe Arg
155    160    165
Gln Ile Val Ser Ala Val Gln Tyr Cys His Gln Lys Arg Ile Val
170    175    180
His Arg Asp Leu Lys Ala Glu Asn Leu Leu Leu Asp Ala Asp Met
185    190    195
Asn Ile Lys Ile Ala Asp Phe Gly Phe Ser Asn Glu Phe Thr Val
200    205    210
Gly Gly Lys Leu Asp Thr Phe Cys Gly Ser Pro Pro Tyr Ala Ala
215    220    225
Pro Glu Leu Phe Gln Gly Lys Lys Tyr Asp Gly Pro Glu Val Asp
230    235    240
Val Trp Ser Leu Gly Val Ile Leu Tyr Thr Leu Val Ser Gly Ser
245    250    255
Leu Pro Phe Asp Gly Gln Asn Leu Lys Glu Leu Arg Glu Arg Val
260    265    270
Leu Arg Gly Lys Tyr Arg Ile Pro Phe Tyr Met Ser Thr Asp Cys
275    280    285
Glu Asn Leu Leu Lys Arg Phe Leu Val Leu Asn Pro Ile Lys Arg
290    295    300
Gly Thr Leu Glu Gln Ile Met Lys Asp Arg Trp Ile Asn Ala Gly
305    310    315
His Glu Glu Asp Glu Leu Lys Pro Phe Val Glu Pro Glu Leu Asp
320    325    330
Ile Ser Asp Gln Lys Arg Ile Asp Ile Met Val Gly Met Gly Tyr
335    340    345
Ser Gln Glu Glu Ile Gln Glu Ser Leu Ser Lys Met Lys Tyr Asp
350    355    360
Glu Ile Thr Ala Thr Tyr Leu Leu Leu Gly Arg Lys Ser Ser Glu

```

16/68

WO 02/08399

PCT/US01/23092

365 370 375
 Leu Asp Ala Ser Asp Ser Ser Ser Ser Sar Asn Leu Ser Leu Ala
 380 385 390
 Lys Val Arg Pro Ser Ser Asp Leu Asn Asn Ser Thr Gly Gln Ser
 395 400 405
 Pro His His Lys Val Gln Arg Ser Val Ser Ser Ser Gln Lys Gln
 410 415 420
 Arg Arg Tyr Ser Asp His Ala Gly Pro Ala Ile Pro Ser Val Val
 425 430 435
 Ala Tyr Pro Lys Arg Ser Gln Thr Ser Thr Ala Asp Ser Asp Leu
 440 445 450
 Lys Glu Asp Gly Ile Ser Ser Arg Lys Ser Ser Gly Ser Ala Val
 455 460 465
 Gly Gly Lys Gly Ile Ala Pro Ala Ser Pro Met Leu Gly Asn Ala
 470 475 480
 Ser Asn Pro Asn Lys Ala Asp Ile Pro Glu Arg Lys Lys Ser Ser
 485 490 495
 Thr Val Pro Ser Ser Asn Thr Ala Ser Gly Gly Met Thr Arg Arg
 500 505 510
 Asn Thr Tyr Val Cys Ser Glu Arg Thr Thr Ala Asp Arg His Ser
 515 520 525
 Val Ile Gln Asn Gly Lys Glu Asn Ser Thr Ile Pro Asp Gln Arg
 530 535 540
 Thr Pro Val Ala Ser Thr His Ser Ile Ser Ser Ala Ala Thr Pro
 545 550 555
 Asp Arg Ile Arg Phe Pro Arg Gly Thr Ala Ser Arg Ser Thr Phe
 560 565 570
 His Gly Gln Pro Arg Glu Arg Arg Thr Ala Thr Tyr Asn Gly Pro
 575 580 585
 Pro Ala Ser Pro Ser Leu Ser His Glu Ala Thr Pro Leu Ser Gln
 590 595 600
 Thr Arg Ser Arg Gly Ser Thr Asn Leu Phe Ser Lys Leu Thr Ser
 605 610 615
 Lys Leu Thr Arg Arg Leu Pro Thr Glu Tyr Glu Arg Asn Gly Arg
 620 625 630
 Tyr Glu Gly Ser Ser Arg Asn Val Ser Ala Glu Gln Lys Asp Glu
 635 640 645
 Asn Lys Glu Ala Lys Pro Arg Ser Leu Arg Phe Thr Trp Ser Met
 650 655 660
 Lys Thr Thr Ser Ser Met Asp Pro Gly Asp Met Met Arg Glu Ile
 665 670 675
 Arg Lys Val Leu Asp Ala Asn Asn Cys Asp Tyr Glu Gln Arg Glu
 680 685 690
 Arg Phe Leu Leu Phe Cys Val His Gly Asp Gly His Ala Glu Asn
 695 700 705
 Leu Val Gln Trp Glu Met Glu Val Cys Lys Leu Pro Arg Leu Ser
 710 715 720
 Leu Asn Gly Val Arg Phe Lys Arg Ile Ser Gly Thr Ser Ile Ala
 725 730 735
 Phe Lys Asn Ile Ala Ser Lys Ile Ala Asn Glu Leu Lys Leu
 740 745

<210> 9
 <211> 386
 <212> PRT
 <213> Homo sapiens

WO 02/08399

PCT/US01/23092

<220>
 <221> misc_feature
 <223> Incyte ID No: 7478559CD1

 <400> 9
 Met Ala Val Pro Pro Ser Ala Pro Gln Pro Arg Ala Ser Phe His
 1 5 10 15
 Leu Arg Arg His Thr Pro Cys Pro Gln Cys Ser Trp Gly Met Glu
 20 25 30
 Glu Lys Ala Ala Ala Ser Ala Ser Cys Arg Glu Pro Pro Gly Pro
 35 40 45
 Pro Arg Ala Ala Ala Val Ala Tyr Phe Gly Ile Ser Val Asp Pro
 50 55 60
 Asp Asp Ile Leu Pro Gly Ala Leu Arg Leu Ile Gln Glu Leu Arg
 65 70 75
 Pro His Trp Lys Pro Glu Gln Val Arg Thr Lys Arg Phe Met Asp
 80 85 90
 Gly Ile Thr Asn Lys Leu Val Ala Cys Tyr Val Glu Glu Asp Met
 95 100 105
 Gln Asp Cys Val Leu Val Arg Val Tyr Gly Glu Arg Thr Glu Leu
 110 115 120
 Leu Val Asp Arg Glu Asn Glu Val Arg Asn Phe Gln Leu Leu Arg
 125 130 135
 Ala His Ser Cys Ala Pro Lys Leu Tyr Cys Thr Phe Gln Asn Gly
 140 145 150
 Leu Cys Tyr Glu Tyr Met Gln Gly Val Ala Leu Glu Pro Glu His
 155 160 165
 Ile Arg Glu Pro Arg Leu Phe Arg Leu Ile Ala Leu Glu Met Ala
 170 175 180
 Lys Ile His Thr Ile His Ala Asn Gly Ser Leu Pro Lys Pro Ile
 185 190 195
 Leu Trp His Lys Met His Asn Tyr Phe Thr Leu Val Lys Asn Glu
 200 205 210
 Ile Asn Pro Ser Leu Ser Ala Asp Val Pro Lys Val Glu Val Leu
 215 220 225
 Glu Arg Glu Leu Ala Trp Leu Lys Glu His Leu Ser Gln Leu Glu
 230 235 240
 Ser Pro Val Val Phe Cys His Asn Asp Leu Leu Cys Lys Asn Ile
 245 250 255
 Ile Tyr Asp Ser Ile Lys Gly His Val Arg Phe Ile Asp Tyr Glu
 260 265 270
 Tyr Ala Gly Tyr Asn Tyr Gln Ala Phe Asp Ile Gly Asn His Phe
 275 280 285
 Asn Glu Phe Ala Gly Val Asn Glu Val Asp Tyr Cys Leu Tyr Pro
 290 295 300
 Ala Arg Glu Thr Gln Leu Gln Trp Leu His Tyr Tyr Leu Gln Ala
 305 310 315
 Gln Lys Gly Met Ala Val Thr Pro Arg Glu Val Gln Arg Leu Tyr
 320 325 330
 Val Gln Val Asn Lys Phe Ala Leu Ala Ser His Phe Phe Trp Ala
 335 340 345
 Leu Trp Ala Leu Ile Gln Asn Gln Tyr Ser Thr Ile Asp Phe Asp
 350 355 360
 Phe Leu Arg Tyr Ala Val Ile Arg Phe Asn Gln Tyr Phe Lys Val
 365 370 375
 Lys Pro Gln Ala Ser Ala Leu Glu Met Pro Lys

WO 02/08399

PCT/US01/23092

380

385

<210> 10
<211> 342
<212> FRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1698381CD1

```

<400> 10
Met Glu Lys Tyr Glu Lys Leu Ala Lys Thr Gly Glu Gly Ser Tyr
 1          5          10          15
Gly Val Val Phe Lys Cys Arg Asn Lys Thr Ser Gly Gln Val Val
 20          25          30
Ala Val Lys Lys Phe Val Glu Ser Glu Asp Asp Pro Val Val Lys
 35          40          45
Lys Ile Ala Leu Arg Glu Ile Arg Met Leu Lys Gln Leu Lys His
 50          55          60
Pro Asn Leu Val Asn Leu Ile Glu Val Phe Arg Arg Lys Arg Lys
 65          70          75
Met His Leu Val Phe Glu Tyr Cys Asp His Thr Leu Leu Asn Glu
 80          85          90
Leu Glu Arg Asn Pro Asn Gly Val Ala Asp Gly Val Ile Lys Ser
 95          100         105
Val Leu Trp Gln Thr Leu Gln Ala Leu Asn Phe Cys His Ile His
 110         115         120
Asn Cys Ile His Arg Asp Ile Lys Pro Glu Asn Ile Leu Ile Thr
 125         130         135
Lys Gln Gly Ile Ile Lys Ile Cys Asp Phe Gly Phe Ala Gln Ile
 140         145         150
Leu Ile Pro Gly Asp Ala Tyr Thr Asp Tyr Val Ala Thr Arg Trp
 155         160         165
Tyr Arg Ala Pro Glu Leu Leu Val Gly Asp Thr Gln Tyr Gly Ser
 170         175         180
Ser Val Asp Ile Trp Ala Ile Gly Cys Val Phe Ala Glu Leu Leu
 185         190         195
Thr Gly Gln Pro Leu Trp Pro Gly Lys Ser Asp Val Asp Gln Leu
 200         205         210
Tyr Leu Ile Ile Arg Thr Leu Gly Lys Leu Ile Pro Arg His Gln
 215         220         225
Ser Ile Phe Lys Ser Asn Gly Phe Phe His Gly Ile Ser Ile Pro
 230         235         240
Glu Pro Glu Asp Met Glu Thr Leu Glu Glu Lys Phe Ser Asp Val
 245         250         255
His Pro Val Ala Leu Asn Phe Met Lys Gly Cys Leu Lys Met Asn
 260         265         270
Pro Asp Asp Arg Leu Thr Cys Ser Gln Leu Leu Glu Ser Ser Tyr
 275         280         285
Phe Asp Ser Phe Gln Glu Ala Gln Ile Lys Arg Lys Ala Arg Asn
 290         295         300
Glu Gly Arg Asn Arg Arg Gln Gln Asn Gln Leu Leu Pro Leu
 305         310         315
Ile Pro Gly Ser His Ile Ser Pro Thr Pro Asp Gly Arg Lys Gln
 320         325         330

```

WO 02/08399

PCT/US01/23092

Val Leu Gln Leu Lys Phe Asp His Leu Pro Asn Ile
335 340

<210> 11
<211> 1164
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7474637CD1

<400> 11
Met Ala Gly Ala Gly Gly Gln His His Pro Pro Gly Ala Ala Gly
1 5 10 15
Gly Ala Ala Ala Gly Ala Gly Ala Ala Val Thr Ser Ala Ala Ala
20 25 30
Ser Ala Gly Pro Gly Glu Asp Ser Ser Asp Ser Glu Ala Glu Gln
35 40 45
Glu Gly Pro Gln Lys Leu Ile Arg Lys Val Ser Thr Ser Gly Gln
50 55 60
Ile Arg Thr Lys Thr Ser Ile Lys Glu Gly Gln Leu Leu Lys Gln
65 70 75
Thr Ser Ser Phe Gln Arg Trp Lys Lys Arg Tyr Phe Lys Leu Arg
80 85 90
Gly Arg Thr Leu Tyr Tyr Ala Lys Asp Ser Lys Ser Leu Ile Phe
95 100 105
Asp Glu Val Asp Leu Ser Asp Ala Ser Val Ala Glu Ala Ser Thr
110 115 120
Lys Asn Ala Asn Asn Ser Phe Thr Ile Ile Thr Pro Phe Arg Arg
125 130 135
Leu Met Leu Cys Ala Glu Asn Arg Lys Glu Met Glu Asp Trp Ile
140 145 150
Ser Ser Leu Lys Ser Val Gln Thr Arg Glu Pro Tyr Glu Val Ala
155 160 165
Gln Phe Asn Val Glu His Phe Ser Gly Met His Asn Trp Tyr Ala
170 175 180
Cys Ser His Ala Arg Pro Thr Phe Cys Asn Val Cys Arg Glu Ser
185 190 195
Leu Ser Gly Val Thr Ser His Gly Leu Ser Cys Glu Val Cys Lys
200 205 210
Phe Lys Ala His Lys Arg Cys Ala Val Arg Ala Thr Asn Asn Cys
215 220 225
Lys Trp Thr Thr Leu Ala Ser Ile Gly Lys Asp Ile Ile Glu Asp
230 235 240
Glu Asp Gly Val Ala Met Pro His Gln Trp Leu Glu Gly Asn Leu
245 250 255
Pro Val Ser Ala Lys Cys Ala Val Cys Asp Lys Thr Cys Gly Ser
260 265 270
Val Leu Arg Leu Gln Asp Trp Lys Cys Leu Trp Cys Lys Thr Met
275 280 285
Val His Thr Ala Cys Lys Asp Leu Tyr His Pro Ile Cys Pro Leu
290 295 300
Gly Gln Cys Lys Val Ser Ile Ile Pro Pro Ile Ala Leu Asn Ser
305 310 315
Thr Asp Ser Asp Gly Phe Cys Arg Ala Thr Phe Ser Phe Cys Val

WO 02/08399

PCT/US01/23092

320 325 330
 Ser Pro Leu Leu Val Phe Val Asn Ser Lys Ser Gly Asp Asn Gln
 335 340 345
 Gly Val Lys Phe Leu Arg Arg Phe Lys Gln Leu Leu Asn Pro Ala
 350 355 360
 Gln Val Phe Asp Leu Met Asn Gly Gly Pro His Leu Gly Leu Arg
 365 370 375
 Leu Phe Gln Lys Phe Asp Asn Phe Arg Ile Leu Val Cys Gly Gly
 380 385 390
 Asp Gly Ser Val Gly Trp Val Leu Ser Glu Ile Asp Lys Leu Asn
 395 400 405
 Leu Asn Lys Gln Cys Gln Leu Gly Val Leu Pro Leu Gly Thr Gly
 410 415 420
 Asn Asp Leu Ala Arg Val Leu Gly Trp Gly Gly Ser Tyr Asp Asp
 425 430 435
 Asp Thr Gln Leu Pro Gln Ile Leu Glu Lys Leu Glu Arg Ala Ser
 440 445 450
 Thr Lys Met Leu Asp Arg Trp Ser Ile Met Thr Tyr Glu Leu Lys
 455 460 465
 Leu Pro Pro Lys Ala Ser Leu Leu Pro Gly Pro Pro Glu Ala Ser
 470 475 480
 Glu Glu Phe Tyr Met Thr Ile Tyr Glu Asp Ser Val Ala Thr His
 485 490 495
 Leu Thr Lys Ile Leu Asn Ser Asp Glu His Ala Val Val Ile Ser
 500 505 510
 Ser Ala Lys Thr Leu Cys Glu Thr Val Lys Asp Phe Val Ala Lys
 515 520 525
 Val Glu Lys Thr Tyr Asp Lys Thr Leu Glu Asn Ala Val Val Ala
 530 535 540
 Asp Ala Val Ala Ser Lys Cys Ser Val Leu Asn Glu Lys Leu Glu
 545 550 555
 Gln Leu Leu Gln Ala Leu His Thr Asp Ser Gln Ala Ala Pro Val
 560 565 570
 Leu Pro Gly Leu Ser Pro Leu Ile Val Glu Glu Asp Ala Val Glu
 575 580 585
 Ser Ser Ser Glu Glu Ser Leu Gly Glu Ser Lys Glu Gln Leu Gly
 590 595 600
 Asp Asp Val Thr Lys Pro Ser Ser Gln Lys Ala Val Lys Pro Arg
 605 610 615
 Glu Ile Met Leu Arg Ala Asn Ser Leu Lys Lys Ala Val Arg Gln
 620 625 630
 Val Ile Glu Glu Ala Gly Lys Val Met Asp Asp Pro Thr Val His
 635 640 645
 Pro Cys Glu Pro Ala Asn Gln Ser Ser Asp Tyr Asp Ser Thr Glu
 650 655 660
 Thr Asp Glu Ser Lys Glu Glu Ala Lys Asp Asp Gly Ala Lys Glu
 665 670 675
 Ser Ile Thr Val Lys Thr Ala Pro Arg Ser Pro Asp Ala Arg Ala
 680 685 690
 Ser Tyr Gly His Ser Gln Thr Asp Ser Val Pro Gly Pro Ala Val
 695 700 705
 Ala Ala Ser Lys Glu Asn Leu Pro Val Leu Asn Thr Arg Ile Ile
 710 715 720
 Cys Pro Gly Leu Arg Ala Gly Leu Ala Ala Ser Ile Ala Gly Ser
 725 730 735
 Ser Ile Ile Asn Lys Met Leu Leu Ala Asn Ile Asp Pro Phe Gly

WO 02/08399

PCT/US01/23092

	740		745		750
Ala Thr Pro Phe Ile Asp Pro Asp Leu Asp Ser Val Asp Gly Tyr					
	755		760		765
Ser Glu Lys Cys Val Met Asn Asn Tyr Phe Gly Ile Gly Leu Asp					
	770		775		780
Ala Lys Ile Ser Leu Glu Phe Asn Asn Lys Arg Glu Glu His Pro					
	785		790		795
Glu Lys Cys Arg Ser Arg Thr Lys Asn Leu Met Trp Tyr Gly Val					
	800		805		810
Leu Gly Thr Arg Glu Leu Leu Gln Arg Ser Tyr Lys Asn Leu Glu					
	815		820		825
Gln Arg Val Gln Leu Glu Cys Asp Gly Gln Tyr Ile Pro Leu Pro					
	830		835		840
Ser Leu Gln Gly Ile Ala Val Leu Asn Ile Pro Ser Tyr Ala Gly					
	845		850		855
Gly Thr Asn Phe Trp Gly Gly Thr Lys Glu Asp Asp Ile Phe Ala					
	860		865		870
Ala Pro Ser Phe Asp Asp Lys Ile Leu Glu Val Val Ala Ile Phe					
	875		880		885
Asp Ser Met Gln Met Ala Val Ser Arg Val Ile Lys Leu Gln His					
	890		895		900
His Arg Ile Ala Gln Cys Arg Thr Val Lys Ile Thr Ile Phe Gly					
	905		910		915
Asp Glu Gly Val Pro Val Gln Val Asp Gly Glu Ala Trp Val Gln					
	920		925		930
Pro Pro Gly Ile Ile Lys Ile Val His Lys Asn Arg Ala Gln Met					
	935		940		945
Leu Thr Arg Asp Arg Ala Phe Glu Ser Thr Leu Lys Ser Trp Glu					
	950		955		960
Asp Lys Gln Lys Cys Asp Ser Gly Lys Pro Val Leu Arg Thr His					
	965		970		975
Leu Tyr Ile His His Ala Ile Asp Leu Ala Thr Glu Glu Val Ser					
	980		985		990
Gln Met Gln Leu Cys Ser Gln Ala Ala Glu Glu Leu Ile Thr Arg					
	995		1000		1005
Ile Cys Asp Ala Ala Thr Ile His Cys Leu Leu Glu Gln Glu Leu					
	1010		1015		1020
Ala His Ala Val Asn Ala Cys Ser His Ala Leu Asn Lys Ala Asn					
	1025		1030		1035
Pro Arg Cys Pro Glu Ser Leu Thr Arg Asp Thr Ala Thr Glu Ile					
	1040		1045		1050
Ala Ile Asn Val Lys Ala Leu Tyr Asn Glu Thr Glu Ser Leu Leu					
	1055		1060		1065
Val Gly Arg Val Pro Leu Gln Leu Glu Ser Pro His Glu Glu Arg					
	1070		1075		1080
Val Ser Asn Ala Leu His Ser Val Glu Val Glu Leu Gln Lys Leu					
	1085		1090		1095
Thr Glu Ile Pro Trp Leu Tyr Tyr Ile Leu His Pro Asn Glu Asp					
	1100		1105		1110
Glu Glu Pro Pro Met Asp Cys Thr Lys Arg Asn Asn Arg Ser Thr					
	1115		1120		1125
Val Phe Arg Ile Val Pro Lys Phe Lys Lys Glu Lys Val Gln Lys					
	1130		1135		1140
Gln Lys Thr Ser Ser Gln Pro Gly Ser Gly Asp Thr Glu Ser Gly					
	1145		1150		1155
Ser Cys Glu Ala Asn Ser Pro Gly Asn					

WO 02/08399

PCT/US01/23092

1160

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

<220>
 <221> misc_feature
 <223> Incyte ID No: 7170260CD1

<400> 12
 Met Glu Asp Phe Leu Leu Ser Asn Gly Tyr Gln Leu Gly Lys Thr
 1 5 10 15
 Ile Gly Glu Gly Thr Tyr Ser Lys Val Lys Glu Ala Phe Ser Lys
 20 25 30
 Lys His Gln Arg Lys Val Ala Ile Lys Val Ile Asp Lys Met Gly
 35 40 45
 Gly Pro Glu Glu Phe Ile Gln Arg Phe Leu Pro Arg Glu Leu Gln
 50 55 60
 Ile Val Arg Thr Leu Asp His Lys Asn Ile Ile Gln Val Tyr Glu
 65 70 75
 Met Leu Glu Ser Ala Asp Gly Lys Ile Cys Leu Val Met Glu Leu
 80 85 90
 Ala Glu Gly Gly Asp Val Phe Asp Cys Val Leu Asn Gly Gly Pro
 95 100 105
 Leu Pro Glu Ser Arg Ala Lys Ala Leu Phe Arg Gln Met Val Glu
 110 115 120
 Ala Ile Arg Tyr Cys His Gly Cys Gly Val Ala His Arg Asp Leu
 125 130 135
 Lys Cys Glu Asn Ala Leu Leu Gln Gly Phe Asn Leu Lys Leu Thr
 140 145 150
 Asp Phe Gly Phe Ala Lys Val Leu Pro Lys Ser His Arg Glu Leu
 155 160 165
 Ser Glu Thr Phe Cys Gly Ser Thr Ala Tyr Ala Ala Pro Glu Val
 170 175 180
 Leu Gln Gly Ile Pro His Asp Ser Lys Lys Gly Asp Val Trp Ser
 185 190 195
 Met Gly Val Val Leu Tyr Val Met Leu Cys Ala Ser Leu Pro Phe
 200 205 210
 Asp Asp Thr Asp Ile Pro Lys Met Leu Trp Gln Gln Gln Lys Gly
 215 220 225
 Val Ser Phe Pro Thr His Leu Ser Ile Ser Ala Asp Cys Gln Asp
 230 235 240
 Leu Leu Lys Arg Leu Leu Glu Pro Asp Met Ile Leu Arg Pro Ser
 245 250 255
 Ile Glu Glu Val Ser Trp His Pro Trp Leu Ala Ser Thr
 260 265

<210> 13
 <211> 965
 <212> PRT
 <213> Homo sapiens

<220>
 <221> misc_feature

23/68

WO 02/08399

PCT/US01/23092

<223> Incyte ID No: 1797506CD1

<400> 13

Met Arg Arg Ala Gly Ile Gly Glu Asp Ser Arg Leu Gly Leu Gln
 1 5 10 15
 Ala Gln Pro Gly Ala Glu Pro Ser Pro Gly Arg Ala Gly Thr Glu
 20 25 30
 Arg Ser Leu Gly Gly Thr Gln Gly Pro Gly Gln Pro Cys Ser Cys
 35 40 45
 Pro Gly Ala Met Ala Ser Ala Val Arg Gly Ser Arg Pro Trp Pro
 50 55 60
 Arg Leu Gly Leu Gln Leu Gln Phe Ala Ala Leu Leu Leu Gly Thr
 65 70 75
 Leu Ser Pro Gln Val His Thr Leu Arg Pro Glu Asn Leu Leu Leu
 80 85 90
 Val Ser Thr Leu Asp Gly Ser Leu His Ala Leu Ser Lys Gln Thr
 95 100 105
 Gly Asp Leu Lys Trp Thr Leu Arg Asp Asp Pro Val Ile Glu Gly
 110 115 120
 Pro Met Tyr Val Thr Glu Met Ala Phe Leu Ser Asp Pro Ala Asp
 125 130 135
 Gly Ser Leu Tyr Ile Leu Gly Thr Gln Lys Gln Gln Gly Leu Met
 140 145 150
 Lys Leu Pro Phe Thr Ile Pro Glu Leu Val His Ala Ser Pro Cys
 155 160 165
 Arg Ser Ser Asp Gly Val Phe Tyr Thr Gly Arg Lys Gln Asp Ala
 170 175 180
 Trp Phe Val Val Asp Pro Glu Ser Gly Glu Thr Gln Met Thr Leu
 185 190 195
 Thr Thr Glu Gly Pro Ser Thr Pro Arg Leu Tyr Ile Gly Arg Thr
 200 205 210
 Gln Tyr Thr Val Thr Met His Asp Pro Arg Ala Pro Ala Leu Arg
 215 220 225
 Trp Asn Thr Thr Tyr Arg Arg Tyr Ser Ala Pro Pro Met Asp Gly
 230 235 240
 Ser Pro Gly Lys Tyr Met Ser His Leu Ala Ser Cys Gly Met Gly
 245 250 255
 Leu Leu Leu Thr Val Asp Pro Gly Ser Gly Thr Val Leu Trp Thr
 260 265 270
 Gln Asp Leu Gly Val Pro Val Met Gly Val Tyr Thr Trp His Gln
 275 280 285
 Asp Gly Leu Arg Gln Leu Pro His Leu Thr Leu Ala Arg Asp Thr
 290 295 300
 Leu His Phe Leu Ala Leu Arg Trp Gly His Ile Arg Leu Pro Ala
 305 310 315
 Ser Gly Pro Arg Asp Thr Ala Thr Leu Phe Ser Thr Leu Asp Thr
 320 325 330
 Gln Leu Leu Met Thr Leu Tyr Val Gly Lys Asp Glu Thr Gly Phe
 335 340 345
 Tyr Val Ser Lys Ala Leu Val His Thr Gly Val Ala Leu Val Pro
 350 355 360
 Arg Gly Leu Thr Leu Ala Pro Ala Asp Gly Pro Thr Thr Asp Glu
 365 370 375
 Val Thr Leu Gln Val Ser Gly Glu Arg Glu Gly Ser Pro Ser Thr
 380 385 390
 Ala Val Arg Tyr Pro Ser Gly Ser Val Ala Leu Pro Ser Gln Trp

WO 02/08399

PCT/US01/23092

	395		400		405									
Leu	Leu	Ile	Gly	His	Glu	Leu	Pro	Pro	Val	Leu	His	Thr	Thr	
	410												420	
Met	Leu	Arg	Val	His	Pro	Thr	Leu	Gly	Ser	Gly	Thr	Ala	Glu	Thr
	425													435
Arg	Pro	Pro	Glu	Asn	Thr	Gln	Ala	Pro	Ala	Phe	Phe	Leu	Glu	Leu
	440													450
Leu	Ser	Leu	Ser	Arg	Glu	Lys	Leu	Trp	Asp	Ser	Glu	Leu	His	Pro
	455													465
Glu	Glu	Lys	Thr	Pro	Asp	Ser	Tyr	Leu	Gly	Leu	Gly	Pro	Gln	Asp
	470													480
Leu	Leu	Ala	Ala	Ser	Leu	Thr	Ala	Val	Leu	Leu	Gly	Gly	Trp	Ile
	485													495
Leu	Phe	Val	Met	Arg	Gln	Gln	Gln	Glu	Thr	Pro	Leu	Ala	Pro	Ala
	500													510
Asp	Phe	Ala	His	Ile	Ser	Gln	Asp	Ala	Gln	Ser	Leu	His	Ser	Gly
	515													525
Ala	Ser	Arg	Arg	Ser	Gln	Lys	Arg	Leu	Gln	Ser	Pro	Ser	Pro	Glu
	530													540
Ser	Pro	Pro	Ser	Ser	Pro	Pro	Ala	Glu	Gln	Leu	Thr	Val	Val	Gly
	545													555
Lys	Ile	Ser	Phe	Asn	Pro	Lys	Asp	Val	Leu	Gly	Arg	Gly	Ala	Gly
	560													570
Gly	Thr	Phe	Val	Phe	Arg	Gly	Gln	Phe	Glu	Gly	Arg	Ala	Val	Ala
	575													585
Val	Lys	Arg	Leu	Leu	Arg	Glu	Cys	Phe	Gly	Leu	Val	Arg	Arg	Glu
	590													600
Val	Gln	Leu	Leu	Gln	Glu	Ser	Asp	Arg	His	Pro	Asn	Val	Leu	Arg
	605													615
Tyr	Phe	Cys	Thr	Glu	Arg	Gly	Pro	Gln	Phe	His	Tyr	Ile	Ala	Leu
	620													630
Glu	Leu	Cys	Arg	Ala	Ser	Leu	Gln	Glu	Tyr	Val	Glu	Asn	Pro	Asp
	635													645
Leu	Asp	Arg	Gly	Gly	Leu	Glu	Pro	Glu	Val	Val	Leu	Gln	Gln	Leu
	650													660
Met	Ser	Gly	Leu	Ala	His	Leu	His	Ser	Leu	His	Ile	Val	His	Arg
	665													675
Asp	Leu	Lys	Pro	Gly	Asn	Ile	Leu	Ile	Thr	Gly	Pro	Asp	Ser	Gln
	680													690
Gly	Leu	Gly	Arg	Val	Val	Leu	Ser	Asp	Phe	Gly	Leu	Cys	Lys	Lys
	695													705
Leu	Pro	Ala	Gly	Arg	Cys	Ser	Phe	Ser	Leu	His	Ser	Gly	Ile	Pro
	710													720
Gly	Thr	Glu	Gly	Trp	Met	Ala	Pro	Glu	Leu	Leu	Gln	Leu	Leu	Pro
	725													735
Pro	Asp	Ser	Pro	Thr	Ser	Ala	Val	Asp	Ile	Phe	Ser	Ala	Gly	Cys
	740													750
Val	Phe	Tyr	Tyr	Val	Leu	Ser	Gly	Gly	Ser	His	Pro	Phe	Gly	Asp
	755													765
Ser	Leu	Tyr	Arg	Gln	Ala	Asn	Ile	Leu	Thr	Gly	Ala	Pro	Cys	Leu
	770													780
Ala	His	Leu	Glu	Glu	Glu	Val	His	Asp	Lys	Val	Val	Ala	Arg	Asp
	785													795
Leu	Val	Gly	Ala	Met	Leu	Ser	Pro	Leu	Pro	Gln	Pro	Arg	Pro	Ser
	800													810
Ala	Pro	Gln	Val	Leu	Ala	His	Pro	Phe	Phe	Trp	Ser	Arg	Ala	Lys

WO 02/08399

PCT/US01/23092

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

<210> 15

<211> 945

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 7474604CD1

<400> 15

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

WO 02/08399

PCT/US01/23092

185 190 195
 Val Pro Ser Lys Cys Val Gln His Tyr Cys His Thr Ser Pro Thr
 200 205 210
 Pro Gly Ala Pro Val Tyr Thr His Val Asp Arg Leu Thr Val Asp
 215 220 225
 Ala Tyr Pro Gly Leu Cys Pro Pro Pro Pro Leu Glu Ser Gly His
 230 235 240
 Arg Ser Leu Pro Pro Ser Pro Arg Gln Arg His Ala Val Arg Thr
 245 250 255
 Pro Pro Arg Thr Pro Asn Ile Val Thr Thr Val Thr Pro Pro Gly
 260 265 270
 Thr Pro Pro Met Arg Lys Lys Asn Lys Leu Lys Pro Pro Gly Thr
 275 280 285
 Pro Pro Pro Ser Ser Arg Lys Leu Ile His Leu Ile Pro Gly Phe
 290 295 300
 Thr Ala Leu His Arg Ser Lys Ser His Glu Phe Gln Leu Gly His
 305 310 315
 Arg Val Asp Glu Ala His Thr Pro Lys Ala Lys Lys Lys Ser Lys
 320 325 330
 Pro Leu Asn Leu Lys Ile His Ser Ser Val Gly Ser Cys Glu Asn
 335 340 345
 Ile Pro Ser Gln Gln Arg Ser Pro Leu Leu Ser Glu Arg Ser Leu
 350 355 360
 Arg Ser Phe Phe Val Gly His Ala Pro Phe Leu Pro Ser Thr Pro
 365 370 375
 Pro Val His Thr Glu Ala Asn Phe Ser Ala Asn Thr Leu Ser Val
 380 385 390
 Pro Arg Trp Ser Pro Gln Ile Pro Arg Arg Asp Leu Gly Asn Ser
 395 400 405
 Ile Lys His Arg Phe Ser Thr Lys Tyr Trp Met Ser Gln Thr Cys
 410 415 420
 Thr Val Cys Gly Lys Gly Met Leu Phe Gly Leu Lys Cys Lys Asn
 425 430 435
 Cys Lys Leu Lys Cys His Asn Lys Cys Thr Lys Glu Ala Pro Pro
 440 445 450
 Cys His Leu Leu Ile Ile His Arg Gly Asp Pro Ala Arg Leu Val
 455 460 465
 Arg Thr Glu Ser Val Pro Cys Asp Ile Asn Asn Pro Leu Arg Lys
 470 475 480
 Pro Pro Arg Tyr Ser Asp Leu His Ile Ser Gln Thr Leu Pro Lys
 485 490 495
 Thr Asn Lys Ile Asn Lys Asp His Ile Pro Val Pro Tyr Gln Pro
 500 505 510
 Asp Ser Ser Ser Asn Pro Ser Ser Thr Thr Ser Ser Thr Pro Ser
 515 520 525
 Ser Pro Ala Pro Pro Leu Pro Pro Ser Ala Thr Pro Pro Ser Pro
 530 535 540
 Leu His Pro Ser Pro Gln Cys Thr Arg Gln Gln Lys Asn Phe Asn
 545 550 555
 Leu Pro Ala Ser His Tyr Tyr Lys Tyr Lys Gln Gln Phe Ile Phe
 560 565 570
 Pro Asp Val Val Pro Val Pro Glu Thr Pro Thr Arg Ala Pro Gln
 575 580 585
 Val Ile Leu His Pro Val Thr Ser Asn Pro Ile Leu Glu Gly Asn
 590 595 600
 Pro Leu Leu Gln Ile Glu Val Glu Pro Thr Ser Glu Asn Glu Glu

WO 02/08399

PCT/US01/23092

```

605          610          615
Val His Asp Glu Ala Glu Glu Ser Glu Asp Asp Phe Glu Glu Met
620          625          630
Asn Leu Ser Leu Leu Ser Ala Arg Ser Phe Pro Arg Lys Ala Ser
635          640          645
Gln Thr Ser Ile Phe Leu Gln Glu Trp Asp Ile Pro Phe Glu Gln
650          655          660
Leu Glu Ile Gly Glu Leu Ile Gly Lys Gly Arg Phe Gly Gln Val
665          670          675
Tyr His Gly Arg Trp His Gly Glu Val Ala Ile Arg Leu Ile Asp
680          685          690
Ile Glu Arg Asp Asn Glu Asp Gln Leu Lys Ala Phe Lys Arg Glu
695          700          705
Val Met Ala Tyr Arg Gln Thr Arg His Glu Asn Val Val Leu Phe
710          715          720
Met Gly Ala Cys Met Ser Pro Pro His Leu Ala Ile Ile Thr Ser
725          730          735
Leu Cys Lys Gly Arg Thr Leu Tyr Ser Val Val Arg Asp Ala Lys
740          745          750
Ile Val Leu Asp Val Asn Lys Thr Arg Gln Ile Ala Gln Glu Ile
755          760          765
Val Lys Gly Met Gly Tyr Leu His Ala Lys Gly Ile Leu His Lys
770          775          780
Asp Leu Lys Ser Lys Asn Val Phe Tyr Asp Asn Gly Lys Val Val
785          790          795
Ile Thr Asp Phe Gly Leu Phe Ser Ile Ser Gly Val Leu Gln Ala
800          805          810
Gly Arg Arg Glu Asp Lys Leu Arg Ile Gln Asn Gly Trp Leu Cys
815          820          825
His Leu Ala Pro Glu Ile Ile Arg Gln Leu Ser Pro Asp Thr Glu
830          835          840
Glu Asp Lys Leu Pro Phe Ser Lys His Ser Asp Val Phe Ala Leu
845          850          855
Gly Thr Ile Trp Tyr Glu Leu His Ala Arg Glu Trp Pro Phe Lys
860          865          870
Thr Gln Pro Ala Glu Ala Ile Ile Trp Gln Met Gly Thr Gly Met
875          880          885
Lys Pro Asn Leu Ser Gln Ile Gly Met Gly Lys Glu Ile Ser Asp
890          895          900
Ile Leu Leu Phe Cys Trp Ala Phe Glu Gln Glu Glu Arg Pro Thr
905          910          915
Phe Thr Lys Leu Met Asp Met Leu Glu Lys Leu Pro Lys Arg Asn
920          925          930
Arg Arg Leu Ser His Pro Gly His Phe Trp Lys Ser Ala Glu Leu
935          940          945

```

<210> 16
<211> 1009
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7474721CD1

WO 02/08399

PCT/US01/23092

<400> 16
Met Glu Thr Cys Ala Gly Pro His Pro Leu Arg Leu Phe Leu Cys
1 5 10 15
Arg Met Gln Leu Cys Leu Ala Leu Leu Leu Gly Pro Trp Arg Pro
20 25 30
Gly Thr Ala Glu Glu Val Ile Leu Leu Asp Ser Lys Ala Ser Gln
35 40 45
Ala Glu Leu Gly Trp Thr Ala Leu Pro Ser Asn Gly Trp Glu Glu
50 55 60
Ile Ser Gly Val Asp Glu His Asp Arg Pro Ile Arg Thr Tyr Gln
65 70 75
Val Cys Asn Val Leu Glu Pro Asn Gln Asp Asn Trp Leu Gln Thr
80 85 90
Gly Trp Ile Ser Arg Gly Arg Gly Gln Arg Ile Phe Val Glu Leu
95 100 105
Gln Phe Thr Leu Arg Asp Cys Ser Ser Ile Pro Gly Ala Ala Gly
110 115 120
Thr Cys Lys Glu Thr Phe Asn Val Tyr Tyr Leu Glu Thr Glu Ala
125 130 135
Asp Leu Gly Arg Gly Arg Pro Arg Leu Gly Gly Ser Arg Pro Arg
140 145 150
Lys Ile Asp Thr Ile Ala Ala Asp Glu Ser Phe Thr Gln Gly Asp
155 160 165
Leu Gly Glu Arg Lys Met Lys Leu Asn Thr Glu Val Arg Glu Ile
170 175 180
Gly Pro Leu Ser Arg Arg Gly Phe His Leu Ala Phe Gln Asp Val
185 190 195
Gly Ala Cys Val Ala Leu Val Ser Val Arg Val Tyr Tyr Lys Gln
200 205 210
Cys Arg Ala Thr Val Arg Gly Leu Ala Thr Phe Pro Ala Thr Ala
215 220 225
Ala Glu Ser Ala Phe Ser Thr Leu Val Glu Val Ala Gly Thr Cys
230 235 240
Val Ala His Ser Glu Gly Glu Pro Gly Ser Pro Pro Arg Met His
245 250 255
Cys Gly Ala Asp Gly Glu Trp Leu Val Pro Val Gly Arg Cys Ser
260 265 270
Cys Ser Ala Gly Phe Gln Glu Arg Gly Asp Ile Cys Glu Ala Cys
275 280 285
Pro Pro Gly Phe Tyr Lys Val Ser Pro Arg Arg Arg Val Cys Ser
290 295 300
Pro Cys Pro Glu His Ser Arg Ala Leu Glu Asn Ala Ser Thr Phe
305 310 315
Cys Val Cys Gln Asp Ser Tyr Ala Arg Ser Pro Thr Asp Pro Pro
320 325 330
Ser Ala Ser Cys Thr Arg Gly Pro Pro Ser Ala Pro Arg Asp Leu
335 340 345
Gln Tyr Ser Leu Ser Arg Ser Pro Leu Val Leu Arg Leu Arg Trp
350 355 360
Leu Pro Pro Ala Asp Ser Gly Gly Arg Ser Asp Val Thr Tyr Ser
365 370 375
Leu Leu Cys Leu Arg Cys Gly Arg Glu Gly Pro Ala Gly Ala Cys
380 385 390
Glu Pro Cys Gly Pro Arg Val Ala Phe Leu Pro Arg Gln Ala Gly
395 400 405
Leu Arg Glu Arg Ala Ala Thr Leu Leu His Leu Arg Pro Gly Ala

WO 02/08399

PCT/US01/23092

	410		415		420
Arg Tyr Thr Val	Arg Val Ala Val Leu	Asn Gly Val Ser Gly Pro			
425		430		435	
Ala Ala Ala Leu Val	Pro Val Gly Ala Val Ser Ile Asn Pro Gly				
440		445		450	
Thr Val Gly Pro Val	Pro Val Ala Gly Val Ile Arg Asp Arg Val				
455		460		465	
Glu Pro Gln Ser Val	Ser Leu Ser Trp Arg Glu Pro Ile Pro Ala				
470		475		480	
Gly Ala Pro Gly Ala	Asn Asp Thr Glu Tyr Glu Ile Arg Tyr Tyr				
485		490		495	
Glu Lys Val Gln Ser	Glu Gln Thr Tyr Ser Met Val Lys Thr Gly				
500		505		510	
Ala Pro Thr Val Thr	Val Thr Asn Leu Lys Pro Ala Thr Arg Tyr				
515		520		525	
Val Phe Gln Ile Arg	Ala Ala Ser Pro Gly Pro Ser Trp Glu Ala				
530		535		540	
Gln Ser Phe Asn Pro	Ser Ile Glu Val Gln Thr Leu Gly Glu Ala				
545		550		555	
Ala Ser Gly Ser Arg	Asp Gln Ser Pro Ala Ile Val Val Thr Val				
560		565		570	
Val Thr Ile Ser Ala	Leu Leu Val Leu Gly Ser Val Met Ser Val				
575		580		585	
Leu Ala Ile Trp Arg	Arg Pro Cys Ser Tyr Gly Lys Gly Gly Gly				
590		595		600	
Asp Ala His Asp Glu	Glu Glu Leu Tyr Phe His Phe Lys Val Pro				
605		610		615	
Thr Arg Arg Thr Phe	Leu Asp Pro Gln Ser Cys Gly Asp Leu Leu				
620		625		630	
Gln Ala Val His Leu	Phe Ala Lys Glu Leu Asp Ala Lys Ser Val				
635		640		645	
Thr Leu Glu Arg Ser	Leu Gly Gly Gly Arg Phe Gly Glu Leu Cys				
650		655		660	
Cys Gly Cys Leu Gln	Leu Pro Gly Arg Gln Glu Leu Leu Val Ala				
665		670		675	
Val His Met Leu Arg	Asp Ser Ala Ser Asp Ser Gln Arg Leu Gly				
680		685		690	
Phe Leu Ala Glu Ala	Leu Thr Leu Gly Gln Phe Asp His Ser His				
695		700		705	
Ile Val Arg Leu Glu	Gly Val Val Thr Arg Gly Ser Thr Leu Met				
710		715		720	
Ile Val Thr Glu Tyr	Met Ser His Gly Ala Leu Asp Gly Phe Leu				
725		730		735	
Arg Arg His Glu Gly	Gln Leu Val Ala Gly Gln Leu Met Gly Leu				
740		745		750	
Leu Pro Gly Leu Ala	Ser Ala Met Lys Tyr Leu Ser Glu Met Gly				
755		760		765	
Tyr Val His Arg Gly	Leu Ala Ala Arg His Val Leu Val Ser Ser				
770		775		780	
Asp Leu Val Cys Lys	Ile Ser Gly Phe Gly Arg Gly Pro Arg Asp				
785		790		795	
Arg Ser Glu Ala Val	Tyr Thr Thr Met Ser Gly Arg Ser Pro Ala				
800		805		810	
Leu Trp Ala Ala Pro	Glu Thr Leu Gln Phe Gly His Phe Ser Ser				
815		820		825	
Ala Ser Asp Val Trp	Ser Phe Gly Ile Ile Met Trp Glu Val Met				

WO 02/08399

PCT/US01/23092

830 835 840
 Ala Phe Gly Glu Arg Pro Tyr Trp Asp Met Ser Gly Gln Asp Val
 845 850 855
 Ile Lys Ala Val Glu Asp Gly Phe Arg Leu Pro Pro Pro Arg Asn
 860 865 870
 Cys Pro Asn Leu Leu His Arg Leu Met Leu Asp Cys Trp Gln Lys
 875 880 885
 Asp Pro Gly Glu Arg Pro Arg Phe Ser Gln Ile His Ser Ile Leu
 890 895 900
 Ser Lys Met Val Gln Asp Pro Glu Pro Pro Lys Cys Ala Leu Thr
 905 910 915
 Thr Cys Pro Arg Pro Pro Thr Pro Leu Ala Asp Arg Ala Phe Ser
 920 925 930
 Thr Phe Pro Ser Phe Gly Ser Val Gly Ala Trp Leu Glu Ala Leu
 935 940 945
 Asp Leu Cys Arg Tyr Lys Asp Ser Phe Ala Ala Ala Gly Tyr Gly
 950 955 960
 Ser Leu Glu Ala Val Ala Glu Met Thr Ala Gln Arg Asp Leu Val
 965 970 975
 Ser Leu Gly Ile Ser Leu Ala Glu His Arg Glu Ala Leu Leu Ser
 980 985 990
 Gly Ile Ser Ala Leu Gln Ala Arg Val Leu Gln Leu Gln Gly Gln
 995 1000 1005
 Gly Val Gln Val

<210> 17

<211> 917

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 7478815CD1

<400> 17

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

WO 02/08399

PCT/US01/23092

Gly Leu Thr Phe Ser Phe Pro Cys Arg Gln Thr Lys Leu Glu Glu
 155 160 165
 Gly Val Leu Leu Ser Trp Thr Lys Lys Phe Lys Ala Arg Gly Val
 170 175 180
 Gln Asp Thr Asp Val Val Ser Arg Leu Thr Lys Ala Met Arg Arg
 185 190 195
 His Lys Asp Met Asp Val Asp Ile Leu Ala Leu Val Asn Asp Thr
 200 205 210
 Val Gly Thr Met Met Thr Cys Ala Tyr Asp Asp Pro Tyr Cys Glu
 215 220 225
 Val Gly Val Ile Ile Gly Thr Gly Thr Asn Ala Cys Tyr Met Glu
 230 235 240
 Asp Met Ser Asn Ile Asp Leu Val Glu Gly Asp Glu Gly Arg Met
 245 250 255
 Cys Ile Asn Thr Glu Trp Gly Ala Phe Gly Asp Asp Gly Ala Leu
 260 265 270
 Glu Asp Ile Arg Thr Glu Phe Asp Arg Glu Leu Asp Leu Gly Ser
 275 280 285
 Leu Asn Pro Gly Lys Gln Leu Phe Glu Lys Met Ile Ser Gly Leu
 290 295 300
 Tyr Leu Gly Glu Leu Val Arg Leu Ile Leu Leu Lys Met Ala Lys
 305 310 315
 Ala Gly Leu Leu Phe Gly Gly Glu Lys Ser Ser Ala Leu His Thr
 320 325 330
 Lys Gly Lys Ile Glu Thr Arg His Val Ala Ala Met Glu Lys Tyr
 335 340 345
 Lys Glu Gly Leu Ala Asn Thr Arg Glu Ile Leu Val Asp Leu Gly
 350 355 360
 Leu Glu Pro Ser Glu Ala Asp Cys Ile Ala Val Gln His Val Cys
 365 370 375
 Thr Ile Val Ser Phe Arg Ser Ala Asn Leu Cys Ala Ala Ala Leu
 380 385 390
 Ala Ala Ile Leu Thr Arg Leu Arg Glu Asn Lys Lys Val Glu Arg
 395 400 405
 Leu Arg Thr Thr Val Gly Met Asp Gly Thr Leu Tyr Lys Ile His
 410 415 420
 Pro Gln Tyr Pro Lys Arg Leu His Lys Val Val Arg Lys Leu Val
 425 430 435
 Pro Ser Cys Asp Val Arg Phe Leu Leu Ser Glu Ser Gly Ser Thr
 440 445 450
 Lys Gly Ala Ala Met Val Thr Ala Val Ala Ser Arg Val Gln Ala
 455 460 465
 Gln Arg Lys Gln Ile Asp Arg Val Leu Ala Leu Phe Gln Leu Thr
 470 475 480
 Arg Glu Gln Leu Val Asp Val Gln Ala Lys Met Arg Ala Glu Leu
 485 490 495
 Glu Tyr Gly Leu Lys Lys Lys Ser His Gly Leu Ala Thr Val Arg
 500 505 510
 Met Leu Pro Thr Tyr Val Cys Gly Leu Pro Asp Gly Thr Glu Lys
 515 520 525
 Gly Lys Phe Leu Ala Leu Asp Leu Gly Gly Thr Asn Phe Arg Val
 530 535 540
 Leu Leu Val Lys Ile Arg Ser Gly Arg Arg Ser Val Arg Met Tyr
 545 550 555
 Asn Lys Ile Phe Ala Ile Pro Leu Glu Ile Met Gln Gly Thr Gly
 560 565 570

WO 02/08399

PCT/US01/23092

Glu Glu Leu Phe Asp His Ile Val Gln Cys Ile Ala Asp Phe Leu
 575 580 585
 Asp Tyr Met Gly Leu Lys Gly Ala Ser Leu Pro Leu Gly Phe Thr
 590 595 600
 Phe Ser Phe Pro Cys Arg Gln Met Ser Ile Asp Lys Gly Thr Leu
 605 610 615
 Ile Gly Trp Thr Lys Gly Phe Lys Ala Thr Asp Cys Glu Gly Glu
 620 625 630
 Asp Val Val Asp Met Leu Arg Glu Ala Ile Lys Arg Arg Asn Glu
 635 640 645
 Phe Asp Leu Asp Ile Val Ala Val Val Asn Asp Thr Val Gly Thr
 650 655 660
 Met Met Thr Cys Gly Tyr Glu Asp Pro Asn Cys Glu Ile Gly Leu
 665 670 675
 Ile Ala Gly Thr Gly Ser Asn Met Cys Tyr Met Glu Asp Met Arg
 680 685 690
 Asn Ile Glu Met Val Glu Gly Gly Glu Gly Lys Met Cys Ile Asn
 695 700 705
 Thr Glu Trp Gly Gly Phe Gly Asp Asn Gly Cys Ile Asp Asp Ile
 710 715 720
 Arg Thr Arg Tyr Asp Thr Glu Val Asp Glu Gly Ser Leu Asn Pro
 725 730 735
 Gly Lys Gln Arg Tyr Glu Lys Met Thr Ser Gly Met Tyr Leu Gly
 740 745 750
 Glu Ile Val Arg Gln Ile Leu Ile Asp Leu Thr Lys Gln Gly Leu
 755 760 765
 Leu Phe Arg Gly Gln Ile Ser Glu Arg Leu Arg Thr Arg Gly Ile
 770 775 780
 Phe Glu Thr Lys Phe Leu Ser Gln Ile Glu Ser Asp Arg Leu Ala
 785 790 795
 Leu Leu Gln Val Arg Arg Ile Leu Gln Gln Leu Gly Leu Asp Ser
 800 805 810
 Thr Cys Glu Asp Ser Ile Val Val Lys Glu Val Cys Gly Ala Val
 815 820 825
 Ser Arg Arg Ala Ala Gln Leu Cys Gly Ala Gly Leu Ala Ala Ile
 830 835 840
 Val Glu Lys Arg Arg Glu Asp Gln Gly Leu Glu His Leu Arg Ile
 845 850 855
 Thr Val Gly Val Asp Gly Thr Leu Tyr Lys Leu His Pro His Phe
 860 865 870
 Ser Arg Ile Leu Gln Glu Thr Val Lys Glu Leu Ala Pro Arg Cys
 875 880 885
 Asp Val Thr Phe Met Leu Ser Glu Asp Gly Ser Gly Lys Gly Ala
 890 895 900
 Ala Leu Ile Thr Ala Val Ala Lys Arg Leu Gln Gln Ala Gln Lys
 905 910 915
 Glu Asn

<210> 18
 <211> 2380
 <212> PRT
 <213> Homo sapiens
 <220>
 <221> misc_feature

WO 02/08399

PCT/US01/23092

<223> Incyte ID No: 7477141CD1

<400> 18

```

Met Asn His Pro Trp Pro Ser Leu Asp Cys His Leu Lys Ala
 1      5      10      15
Arg Ser Gly His Ala Leu Leu Ser Trp Pro Gly Gly Trp Ala Phe
 20      25      30
Pro Ile Ser Arg Glu Gln Asn Ala Ser Leu Ser Leu Cys Leu Ser
 35      40      45
Val Ser Leu Cys Val Arg Met Cys Val Ser Leu Thr Leu Cys Val
 50      55      60
Ser Ala Leu Cys Val Ala Pro Val Ala Ala Phe Pro Ser Ala His
 65      70      75
Pro Glu Ser Arg Ser Leu Ala Val Leu Ala Pro Leu Gln Asp Val
 80      85      90
Asp Val Gly Ala Gly Glu Met Ala Leu Phe Glu Cys Leu Val Ala
 95      100     105
Gly Pro Thr Asp Val Glu Val Asp Trp Leu Cys Arg Gly Arg Leu
 110     115     120
Leu Gln Pro Ala Leu Leu Lys Cys Lys Met His Phe Asp Gly Arg
 125     130     135
Lys Cys Lys Leu Leu Leu Thr Ser Val His Glu Asp Asp Ser Gly
 140     145     150
Val Tyr Thr Cys Lys Leu Ser Thr Ala Lys Asp Glu Leu Thr Cys
 155     160     165
Ser Ala Arg Leu Thr Val Arg Pro Ser Leu Ala Pro Leu Phe Thr
 170     175     180
Arg Leu Leu Glu Asp Val Glu Val Leu Glu Gly Arg Ala Ala Arg
 185     190     195
Phe Asp Cys Lys Ile Ser Gly Thr Pro Pro Pro Val Val Thr Trp
 200     205     210
Thr His Phe Gly Cys Pro Met Glu Glu Ser Glu Asn Leu Arg Leu
 215     220     225
Arg Gln Asp Gly Gly Leu His Ser Leu His Ile Ala His Val Gly
 230     235     240
Ser Glu Asp Glu Gly Leu Tyr Ala Val Ser Ala Val Asn Thr His
 245     250     255
Gly Gln Ala His Cys Ser Ala Gln Leu Tyr Val Glu Glu Pro Arg
 260     265     270
Thr Ala Ala Ser Gly Pro Ser Ser Lys Leu Glu Lys Met Pro Ser
 275     280     285
Ile Pro Glu Glu Pro Glu Gln Gly Glu Leu Glu Arg Leu Ser Ile
 290     295     300
Pro Asp Phe Leu Arg Pro Leu Gln Asp Leu Glu Val Gly Leu Ala
 305     310     315
Lys Glu Ala Met Leu Glu Cys Gln Val Thr Gly Leu Pro Tyr Pro
 320     325     330
Thr Ile Ser Trp Phe His Asn Gly His Arg Ile Gln Ser Ser Asp
 335     340     345
Asp Arg Arg Met Thr Gln Tyr Arg Asp Val His Arg Leu Val Phe
 350     355     360
Pro Ala Val Gly Pro Gln His Ala Gly Val Tyr Lys Ser Val Ile
 365     370     375
Ala Asn Lys Leu Gly Lys Ala Ala Cys Tyr Ala His Leu Tyr Val
 380     385     390
Thr Asp Val Val Pro Gly Pro Pro Asp Gly Ala Pro Gln Val Val

```

35/68

WO 02/08399

PCT/US01/23092

395 400 405
 Ala Val Thr Gly Arg Met Val Thr Leu Thr Trp Asn Pro Pro Arg
 410 415 420
 Ser Leu Asp Met Ala Ile Asp Pro Asp Ser Leu Thr Tyr Thr Val
 425 430 435
 Gln His Gln Val Leu Gly Ser Asp Gln Trp Thr Ala Leu Val Thr
 440 445 450
 Gly Leu Arg Glu Pro Gly Trp Ala Ala Thr Gly Leu Arg Lys Gly
 455 460 465
 Val Gln His Ile Phe Arg Val Leu Ser Thr Thr Val Lys Ser Ser
 470 475 480
 Ser Lys Pro Ser Pro Pro Ser Glu Pro Val Gln Leu Leu Glu His
 485 490 495
 Gly Pro Thr Leu Glu Glu Ala Pro Ala Met Leu Asp Lys Pro Asp
 500 505 510
 Ile Val Tyr Val Val Glu Gly Gln Pro Ala Ser Val Thr Val Thr
 515 520 525
 Phe Asn His Val Glu Ala Gln Val Val Trp Arg Ser Cys Arg Gly
 530 535 540
 Ala Leu Leu Glu Ala Arg Ala Gly Val Tyr Glu Leu Ser Gln Pro
 545 550 555
 Asp Asp Asp Gln Tyr Cys Leu Arg Ile Cys Arg Val Ser Arg Arg
 560 565 570
 Asp Met Gly Ala Leu Thr Cys Thr Ala Arg Asn Arg His Gly Thr
 575 580 585
 Gln Thr Cys Ser Val Thr Leu Glu Leu Ala Glu Ala Pro Arg Phe
 590 595 600
 Glu Ser Ile Met Glu Asp Val Glu Val Gly Ala Gly Glu Thr Ala
 605 610 615
 Arg Phe Ala Val Val Val Glu Gly Lys Pro Leu Pro Asp Ile Met
 620 625 630
 Trp Tyr Lys Asp Glu Val Leu Leu Thr Glu Ser Ser His Val Ser
 635 640 645
 Phe Val Tyr Glu Glu Asn Glu Cys Ser Leu Val Val Leu Ser Thr
 650 655 660
 Gly Ala Gln Asp Gly Gly Val Tyr Thr Cys Thr Ala Gln Asn Leu
 665 670 675
 Ala Gly Glu Val Ser Cys Lys Ala Glu Leu Ala Val His Ser Ala
 680 685 690
 Gln Thr Ala Met Glu Val Glu Gly Val Gly Glu Asp Glu Asp His
 695 700 705
 Arg Gly Arg Arg Leu Ser Asp Phe Tyr Asp Ile His Gln Glu Ile
 710 715 720
 Gly Arg Gly Ala Phe Ser Tyr Leu Arg Arg Ile Val Glu Arg Ser
 725 730 735
 Ser Gly Leu Glu Phe Ala Ala Lys Phe Ile Pro Ser Gln Ala Lys
 740 745 750
 Pro Lys Ala Ser Ala Arg Arg Glu Ala Arg Leu Leu Ala Arg Leu
 755 760 765
 Gln His Asp Cys Val Leu Tyr Phe His Glu Ala Phe Glu Arg Arg
 770 775 780
 Arg Gly Leu Val Ile Val Thr Glu Leu Cys Thr Glu Glu Leu Leu
 785 790 795
 Glu Arg Ile Ala Arg Lys Pro Thr Val Cys Glu Ser Glu Ile Arg
 800 805 810
 Ala Tyr Met Arg Gln Val Leu Glu Gly Ile His Tyr Leu His Gln

WO 02/08399

PCT/US01/23092

1235 1240 1245
 Gln Gly Glu Ala Glu Pro Arg Gly Arg His Arg Arg Ala Gly Ala
 1250 1255 1260
 Pro Leu Glu Ile Pro Val Ala Arg Leu Gly Ala Arg Arg Leu Gln
 1265 1270 1275
 Glu Ser Pro Ser Leu Ser Ala Leu Ser Glu Ala Gln Pro Ser Ser
 1280 1285 1290
 Pro Ala Arg Pro Ser Ala Pro Lys Pro Ser Thr Pro Lys Ser Ala
 1295 1300 1305
 Glu Pro Ser Ala Thr Thr Pro Ser Asp Ala Pro Gln Pro Pro Ala
 1310 1315 1320
 Pro Gln Pro Ala Gln Asp Lys Ala Pro Glu Pro Arg Pro Glu Pro
 1325 1330 1335
 Val Arg Ala Ser Lys Pro Ala Pro Pro Gln Ala Leu Gln Thr
 1340 1345 1350
 Leu Ala Leu Pro Leu Thr Pro Tyr Ala Gln Ile Ile Gln Ser Leu
 1355 1360 1365
 Gln Leu Ser Gly His Ala Gln Gly Pro Ser Gln Gly Pro Ala Ala
 1370 1375 1380
 Pro Pro Ser Glu Pro Lys Pro His Ala Ala Val Phe Ala Arg Val
 1385 1390 1395
 Ala Ser Pro Pro Pro Gly Ala Pro Glu Lys Arg Val Pro Ser Ala
 1400 1405 1410
 Gly Gly Pro Pro Val Leu Ala Glu Lys Ala Arg Val Pro Thr Val
 1415 1420 1425
 Pro Pro Arg Pro Gly Ser Ser Leu Ser Ser Ile Glu Asn Leu
 1430 1435 1440
 Glu Ser Glu Ala Val Phe Glu Ala Lys Phe Lys Arg Ser Arg Glu
 1445 1450 1455
 Ser Pro Leu Ser Leu Gly Leu Arg Leu Leu Ser Arg Ser Arg Ser
 1460 1465 1470
 Glu Glu Arg Gly Pro Phe Arg Gly Ala Glu Glu Glu Asp Gly Ile
 1475 1480 1485
 Tyr Arg Pro Ser Pro Ala Gly Thr Pro Leu Glu Leu Val Arg Arg
 1490 1495 1500
 Pro Glu Arg Ser Arg Ser Val Gln Asp Leu Arg Ala Val Gly Glu
 1505 1510 1515
 Pro Gly Leu Val Arg Arg Leu Ser Leu Ser Leu Ser Gln Arg Leu
 1520 1525 1530
 Arg Arg Thr Pro Pro Ala Gln Arg His Pro Ala Trp Glu Ala Arg
 1535 1540 1545
 Gly Gly Asp Gly Glu Ser Ser Glu Gly Gly Ser Ser Ala Arg Gly
 1550 1555 1560
 Ser Pro Val Leu Ala Met Arg Arg Arg Leu Ser Phe Thr Leu Glu
 1565 1570 1575
 Arg Leu Ser Ser Arg Leu Gln Arg Ser Gly Ser Ser Glu Asp Ser
 1580 1585 1590
 Gly Gly Ala Ser Gly Arg Ser Thr Pro Leu Phe Gly Arg Leu Arg
 1595 1600 1605
 Arg Ala Thr Ser Glu Gly Glu Ser Leu Arg Arg Leu Gly Leu Pro
 1610 1615 1620
 His Asn Gln Leu Ala Ala Gln Ala Gly Ala Thr Thr Pro Ser Ala
 1625 1630 1635
 Glu Ser Leu Gly Ser Glu Ala Ser Ala Thr Ser Gly Ser Ser Ala
 1640 1645 1650
 Pro Gly Glu Ser Arg Ser Arg Leu Arg Trp Gly Phe Ser Arg Pro

WO 02/08399

PCT/US01/23092

1655 1660 1665
 Arg Lys Asp Lys Gly Leu Ser Pro Pro Asn Leu Ser Ala Ser Val
 1670 1675 1680
 Gln Glu Glu Leu Gly His Gln Tyr Val Arg Ser Glu Ser Asp Phe
 1685 1690 1695
 Pro Pro Val Phe His Ile Lys Leu Lys Asp Gln Val Leu Leu Glu
 1700 1705 1710
 Gly Glu Ala Ala Thr Leu Leu Cys Leu Pro Ala Ala Cys Pro Ala
 1715 1720 1725
 Pro His Ile Ser Trp Met Lys Asp Lys Lys Ser Leu Arg Ser Glu
 1730 1735 1740
 Pro Ser Val Ile Ile Val Ser Cys Lys Asp Gly Arg Gln Leu Leu
 1745 1750 1755
 Ser Ile Pro Arg Ala Gly Lys Arg His Ala Gly Leu Tyr Glu Cys
 1760 1765 1770
 Ser Ala Thr Asn Val Leu Gly Ser Ile Thr Ser Ser Cys Thr Val
 1775 1780 1785
 Ala Val Ala Arg Val Pro Gly Lys Leu Ala Pro Pro Glu Val Pro
 1790 1795 1800
 Gln Thr Tyr Gln Asp Thr Ala Leu Val Leu Trp Lys Pro Gly Asp
 1805 1810 1815
 Ser Arg Ala Pro Cys Thr Tyr Thr Leu Glu Arg Arg Val Asp Gly
 1820 1825 1830
 Glu Ser Val Trp His Pro Val Ser Ser Gly Ile Pro Asp Cys Tyr
 1835 1840 1845
 Tyr Asn Val Thr His Leu Pro Val Gly Val Thr Val Arg Phe Arg
 1850 1855 1860
 Val Ala Cys Ala Asn Arg Ala Gly Gln Gly Pro Phe Ser Asn Ser
 1865 1870 1875
 Ser Glu Lys Val Phe Val Arg Gly Thr Gln Asp Ser Ser Ala Val
 1880 1885 1890
 Pro Ser Ala Ala His Gln Glu Ala Pro Val Thr Ser Arg Pro Ala
 1895 1900 1905
 Arg Ala Arg Pro Pro Asp Ser Pro Thr Ser Leu Ala Pro Pro Leu
 1910 1915 1920
 Ala Pro Ala Ala Pro Thr Pro Pro Ser Val Thr Val Ser Pro Ser
 1925 1930 1935
 Ser Pro Pro Thr Pro Pro Ser Gln Ala Leu Ser Ser Leu Lys Ala
 1940 1945 1950
 Val Gly Pro Pro Pro Gln Thr Pro Pro Arg Arg His Arg Gly Leu
 1955 1960 1965
 Gln Ala Ala Arg Pro Ala Glu Pro Thr Leu Pro Ser Thr His Val
 1970 1975 1980
 Thr Pro Ser Glu Pro Lys Pro Phe Val Leu Asp Thr Gly Thr Pro
 1985 1990 1995
 Ile Pro Ala Ser Thr Pro Gln Gly Val Lys Pro Val Ser Ser Ser
 2000 2005 2010
 Thr Pro Val Tyr Val Val Thr Ser Phe Val Ser Ala Pro Pro Ala
 2015 2020 2025
 Pro Glu Pro Pro Ala Pro Glu Pro Pro Pro Glu Pro Thr Lys Val
 2030 2035 2040
 Thr Val Gln Ser Leu Ser Pro Ala Lys Glu Val Val Ser Ser Pro
 2045 2050 2055
 Gly Ser Ser Pro Arg Ser Ser Pro Arg Pro Glu Gly Thr Thr Leu
 2060 2065 2070
 Arg Gln Gly Pro Pro Gln Lys Pro Tyr Thr Phe Leu Glu Glu Lys

WO 02/08399

PCT/US01/23092

Ala Asp Gly Gly Pro Glu Pro Thr Arg Asn Gly Val Asp Pro Pro
35 40 45
Pro Arg Ala Arg Ala Ala Ser Val Ile Pro Gly Ser Thr Ser Arg
50 55 60
Leu Leu Pro Ala Arg Pro Ser Leu Ser Ala Arg Lys Leu Ser Leu
65 70 75
Gln Glu Arg Pro Ala Gly Ser Tyr Leu Glu Ala Gln Ala Gly Pro
80 85 90
Tyr Ala Thr Gly Pro Ala Ser His Ile Ser Pro Arg Ala Trp Arg
95 100 105
Arg Pro Thr Ile Glu Ser His His Val Ala Ile Ser Asp Ala Glu
110 115 120
Asp Cys Val Gln Leu Asn Gln Tyr Lys Leu Gln Ser Glu Ile Gly
125 130 135
Lys Gly Ala Tyr Gly Val Val Arg Leu Ala Tyr Asn Glu Ser Glu
140 145 150
Asp Arg His Tyr Ala Met Lys Val Leu Ser Lys Lys Lys Leu Leu
155 160 165
Lys Gln Tyr Gly Phe Pro Arg Arg Pro Pro Pro Arg Gly Ser Gln
170 175 180
Ala Ala Gln Gly Gly Pro Ala Lys Gln Leu Leu Pro Leu Glu Arg
185 190 195
Val Tyr Gln Glu Ile Ala Ile Leu Lys Lys Leu Asp His Val Asn
200 205 210
Val Val Lys Leu Ile Glu Val Leu Asp Asp Pro Ala Glu Asp Asn
215 220 225
Leu Tyr Leu Val Phe Asp Leu Leu Arg Lys Gly Pro Val Met Glu
230 235 240
Val Pro Cys Asp Lys Pro Phe Ser Glu Glu Gln Ala Arg Leu Tyr
245 250 255
Leu Arg Asp Val Ile Leu Gly Leu Glu Tyr Leu His Cys Gln Lys
260 265 270
Ile Val His Arg Asp Ile Lys Pro Ser Asn Leu Leu Leu Gly Asp
275 280 285
Asp Gly His Val Lys Ile Ala Asp Phe Gly Val Ser Asn Gln Phe
290 295 300
Glu Gly Asn Asp Ala Gln Leu Ser Ser Thr Ala Gly Thr Pro Ala
305 310 315
Phe Met Ala Pro Glu Ala Ile Ser Asp Ser Gly Gln Ser Phe Ser
320 325 330
Gly Lys Ala Leu Asp Val Trp Ala Thr Gly Val Thr Leu Tyr Cys
335 340 345
Phe Val Tyr Gly Lys Cys Pro Phe Ile Asp Asp Phe Ile Leu Ala
350 355 360
Leu His Arg Lys Ile Lys Asn Glu Pro Val Val Phe Pro Glu Glu
365 370 375
Pro Glu Ile Ser Glu Glu Leu Lys Asp Leu Ile Leu Lys Met Leu
380 385 390
Asp Lys Asn Pro Glu Thr Arg Ile Gly Val Pro Asp Ile Lys Leu
395 400 405
His Pro Trp Val Thr Lys Asn Gly Glu Glu Pro Leu Pro Ser Glu
410 415 420
Glu Glu His Cys Ser Val Val Glu Val Thr Glu Glu Glu Val Lys
425 430 435
Asn Ser Val Arg Leu Ile Pro Ser Trp Thr Thr Val Ile Leu Val
440 445 450

WO 02/08399

PCT/US01/23092

Lys Ser Met Leu Arg Lys Arg Ser Phe Gly Asn Pro Phe Glu Pro
 455 460 465
 Gln Ala Arg Arg Glu Glu Arg Ser Met Ser Ala Pro Gly Asn Leu
 470 475 480
 Leu Val Lys Glu Gly Phe Gly Glu Gly Gly Lys Ser Pro Glu Leu
 485 490 495
 Pro Gly Val Gln Glu Asp Glu Ala Ala Ser
 500 505

<210> 20
 <211> 1572
 <212> PRT
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 7477549CD1

<400> 20
 Met Glu Arg Arg Leu Arg Ala Leu Glu Gln Leu Ala Arg Gly Glu
 1 5 10 15
 Ala Gly Gly Cys Pro Gly Leu Asp Gly Leu Leu Asp Leu Leu Leu
 20 25 30
 Ala Leu His His Glu Leu Ser Ser Gly Pro Leu Arg Arg Glu Arg
 35 40 45
 Ser Val Ala Gln Phe Leu Ser Trp Ala Ser Pro Phe Val Ser Lys
 50 55 60
 Val Lys Glu Leu Arg Leu Gln Arg Asp Asp Phe Glu Ile Leu Lys
 65 70 75
 Val Ile Gly Arg Gly Ala Phe Gly Glu Val Thr Val Val Arg Gln
 80 85 90
 Arg Asp Thr Gly Gln Ile Phe Ala Met Lys Met Leu His Lys Trp
 95 100 105
 Glu Met Leu Lys Arg Ala Glu Thr Ala Cys Phe Arg Glu Glu Arg
 110 115 120
 Asp Val Leu Val Lys Gly Asp Ser Arg Trp Val Thr Thr Leu His
 125 130 135
 Tyr Ala Phe Gln Asp Glu Glu Tyr Leu Tyr Leu Val Met Asp Tyr
 140 145 150
 Tyr Ala Gly Gly Asp Leu Leu Thr Leu Leu Ser Arg Phe Glu Asp
 155 160 165
 Arg Leu Pro Pro Glu Leu Ala Gln Phe Tyr Leu Ala Glu Met Val
 170 175 180
 Leu Ala Ile His Ser Leu His Gln Leu Gly Tyr Val His Arg Asp
 185 190 195
 Val Lys Pro Asp Asn Val Leu Leu Asp Val Asn Gly His Ile Arg
 200 205 210
 Leu Ala Asp Phe Gly Ser Cys Leu Arg Leu Asn Thr Asn Gly Met
 215 220 225
 Val Asp Ser Ser Val Ala Val Gly Thr Pro Asp Tyr Ile Ser Pro
 230 235 240
 Glu Ile Leu Gln Ala Met Glu Glu Gly Lys Gly His Tyr Gly Pro
 245 250 255
 Gln Cys Asp Trp Trp Ser Leu Gly Val Cys Ala Tyr Glu Leu Leu
 260 265 270
 Phe Gly Glu Thr Pro Phe Tyr Ala Glu Ser Leu Val Glu Thr Tyr

WO 02/08399

PCT/US01/23092

275	280	285
Gly Lys Ile Met Asn His Glu Asp His Leu Gln Phe Pro Pro Asp		
290	295	300
Val Pro Asp Val Pro Ala Ser Ala Gln Asp Leu Ile Arg Gln Leu		
305	310	315
Leu Cys Arg Gln Glu Glu Arg Leu Gly Arg Gly Gly Leu Asp Asp		
320	325	330
Phe Arg Asn His Pro Phe Phe Glu Gly Val Asp Trp Glu Arg Leu		
335	340	345
Ala Ser Ser Thr Ala Pro Tyr Ile Pro Glu Leu Arg Gly Pro Met		
350	355	360
Asp Thr Ser Asn Phe Asp Val Asp Asp Asp Thr Leu Asn His Pro		
365	370	375
Gly Thr Leu Pro Pro Pro Ser His Gly Ala Phe Ser Gly His His		
380	385	390
Leu Pro Phe Val Gly Phe Thr Tyr Thr Ser Gly Ser His Ser Pro		
395	400	405
Glu Ser Ser Ser Glu Ala Trp Ala Ala Leu Glu Arg Lys Leu Gln		
410	415	420
Cys Leu Glu Gln Glu Lys Val Glu Leu Ser Arg Lys His Gln Glu		
425	430	435
Ala Leu His Ala Pro Thr Asp His Arg Glu Leu Glu Gln Leu Arg		
440	445	450
Lys Glu Val Gln Thr Leu Arg Asp Arg Leu Pro Glu Met Leu Arg		
455	460	465
Asp Lys Ala Ser Leu Ser Gln Thr Asp Gly Pro Pro Ala Gly Ser		
470	475	480
Pro Gly Gln Asp Ser Asp Leu Arg Gln Glu Leu Asp Arg Leu His		
485	490	495
Arg Glu Leu Ala Glu Gly Arg Ala Gly Leu Gln Ala Gln Glu Gln		
500	505	510
Glu Leu Cys Arg Ala Gln Gly Gln Gln Glu Glu Leu Leu Gln Arg		
515	520	525
Leu Gln Glu Ala Gln Glu Arg Glu Ala Ala Thr Ala Ser Gln Thr		
530	535	540
Arg Ala Leu Ser Ser Gln Leu Glu Glu Ala Arg Ala Ala Gln Arg		
545	550	555
Glu Leu Glu Ala Gln Val Ser Ser Leu Ser Arg Gln Val Thr Gln		
560	565	570
Leu Gln Gly Gln Trp Glu Gln Arg Leu Glu Glu Ser Ser Gln Ala		
575	580	585
Lys Thr Ile His Thr Ala Ser Glu Thr Asn Gly Met Gly Pro Pro		
590	595	600
Glu Gly Gly Pro Gln Glu Ala Gln Leu Arg Lys Glu Val Ala Ala		
605	610	615
Leu Arg Glu Gln Leu Glu Gln Ala His Ser His Arg Pro Ser Gly		
620	625	630
Lys Glu Glu Ala Leu Cys Gln Leu Gln Glu Glu Asn Arg Arg Leu		
635	640	645
Ser Arg Glu Gln Glu Arg Leu Glu Ala Glu Leu Ala Gln Glu Gln		
650	655	660
Glu Ser Lys Gln Arg Leu Glu Gly Glu Arg Arg Glu Thr Glu Ser		
665	670	675
Asn Trp Glu Ala Gln Leu Ala Asp Ile Leu Ser Trp Val Asn Asp		
680	685	690
Glu Lys Val Ser Arg Gly Tyr Leu Gln Ala Leu Ala Thr Lys Met		

WO 02/08399

PCT/US01/23092

	695	700	705
Ala Glu Glu Leu	Glu Ser Leu Arg Asn Val Gly Thr Gln Thr Leu		
710	715	720	
Pro Ala Arg Pro	Leu Lys Met Glu Ala Ser Ala Arg Leu Glu Leu		
725	730	735	
Gln Ser Ala Leu	Glu Ala Glu Ile Arg Ala Lys Gln Gly Leu Gln		
740	745	750	
Glu Arg Leu Thr	Gln Val Gln Glu Ala Gln Leu Gln Ala Glu Arg		
755	760	765	
Arg Leu Gln Glu	Ala Glu Lys Gln Ser Gln Ala Leu Gln Gln Glu		
770	775	780	
Leu Ala Met Leu	Arg Glu Glu Leu Arg Ala Arg Gly Pro Val Asp		
785	790	795	
Thr Lys Pro Ser	Asn Ser Leu Ile Pro Phe Leu Ser Phe Arg Ser		
800	805	810	
Ser Glu Lys Asp	Ser Ala Lys Asp Pro Gly Ile Ser Gly Glu Ala		
815	820	825	
Thr Arg His Gly	Gly Glu Pro Asp Leu Arg Pro Glu Gly Arg Arg		
830	835	840	
Ser Leu Arg Met	Gly Ala Val Phe Pro Arg Ala Pro Thr Ala Asn		
845	850	855	
Thr Ala Ser Thr	Glu Gly Leu Pro Ala Lys Gly Trp Gly Met Gly		
860	865	870	
Pro Trp Glu Ala	Leu Gly Asn Gly Cys Pro Pro Pro Gln Pro Gly		
875	880	885	
Ser His Thr Leu	Arg Pro Arg Ser Phe Pro Ser Pro Thr Lys Cys		
890	895	900	
Leu Arg Cys Thr	Ser Leu Met Leu Gly Leu Gly Arg Gln Gly Leu		
905	910	915	
Gly Cys Asp Ala	Cys Gly Tyr Phe Cys His Thr Thr Cys Ala Pro		
920	925	930	
Gln Ala Pro Pro	Cys Pro Val Pro Pro Asp Leu Leu Arg Thr Ala		
935	940	945	
Leu Gly Val His	Pro Glu Thr Gly Thr Gly Thr Ala Tyr Glu Gly		
950	955	960	
Phe Leu Ser Val	Pro Arg Pro Ser Gly Val Arg Arg Gly Trp Gln		
965	970	975	
Arg Val Phe Ala	Ala Leu Ser Asp Ser Arg Leu Leu Leu Phe Asp		
980	985	990	
Ala Pro Asp Leu	Arg Leu Ser Pro Pro Ser Gly Ala Leu Leu Gln		
995	1000	1005	
Val Leu Asp Leu	Arg Asp Pro Gln Phe Ser Ala Thr Pro Val Leu		
1010	1015	1020	
Ala Ser Asp Val	Ile His Ala Gln Ser Arg Asp Leu Pro Arg Ile		
1025	1030	1035	
Phe Arg Val Thr	Thr Ser Gln Leu Ala Val Pro Pro Thr Thr Cys		
1040	1045	1050	
Thr Val Leu Leu	Leu Ala Glu Ser Glu Gly Glu Arg Glu Arg Trp		
1055	1060	1065	
Leu Gln Val Leu	Gly Glu Leu Gln Arg Leu Leu Leu Asp Ala Arg		
1070	1075	1080	
Pro Arg Pro Arg	Pro Val Tyr Thr Leu Lys Glu Ala Tyr Asp Asn		
1085	1090	1095	
Gly Leu Pro Leu	Leu Pro His Thr Leu Cys Ala Ala Ile Leu Asp		
1100	1105	1110	
Gln Asp Arg Leu	Ala Leu Gly Thr Glu Glu Gly Leu Phe Val Ile		

WO 02/08399

PCT/US01/23092

1115	1120	1125
His Leu Arg Ser Asn Asp Ile Phe Gln Val Gly Glu Cys Arg Arg		
1130	1135	1140
Val Gln Gln Leu Thr Leu Ser Pro Ser Ala Gly Leu Leu Val Val		
1145	1150	1155
Leu Cys Gly Arg Gly Pro Ser Val Arg Leu Phe Ala Leu Ala Glu		
1160	1165	1170
Leu Glu Asn Ile Glu Val Ala Gly Ala Lys Ile Pro Glu Ser Arg		
1175	1180	1185
Gly Cys Gln Val Leu Ala Ala Gly Ser Ile Leu Gln Ala Arg Thr		
1190	1195	1200
Pro Val Leu Cys Val Ala Val Lys Arg Gln Val Leu Cys Tyr Gln		
1205	1210	1215
Leu Gly Pro Gly Pro Gly Pro Trp Gln Arg Arg Ile Arg Glu Leu		
1220	1225	1230
Gln Ala Pro Ala Thr Val Gln Ser Leu Gly Leu Leu Gly Asp Arg		
1235	1240	1245
Leu Cys Val Gly Ala Ala Gly Gly Phe Ala Leu Tyr Pro Leu Leu		
1250	1255	1260
Asn Glu Ala Ala Pro Leu Ala Leu Gly Ala Gly Leu Val Pro Glu		
1265	1270	1275
Glu Leu Pro Pro Ser Arg Gly Gly Leu Gly Glu Ala Leu Gly Ala		
1280	1285	1290
Val Glu Leu Ser Leu Ser Glu Phe Leu Leu Leu Phe Thr Thr Ala		
1295	1300	1305
Gly Ile Tyr Val Asp Gly Ala Gly Arg Lys Ser Arg Gly His Glu		
1310	1315	1320
Leu Leu Trp Pro Ala Ala Pro Met Gly Trp Gly Tyr Ala Ala Pro		
1325	1330	1335
Tyr Leu Thr Val Phe Ser Glu Asn Ser Ile Asp Val Phe Asp Val		
1340	1345	1350
Arg Arg Ala Glu Trp Val Gln Thr Val Pro Leu Lys Lys Val Arg		
1355	1360	1365
Pro Leu Asn Pro Glu Gly Ser Leu Phe Leu Tyr Gly Thr Glu Lys		
1370	1375	1380
Val Arg Leu Thr Tyr Leu Arg Asn Gln Leu Ala Glu Lys Asp Glu		
1385	1390	1395
Phe Asp Ile Pro Asp Leu Thr Asp Asn Ser Arg Arg Gln Leu Phe		
1400	1405	1410
Arg Thr Lys Ser Lys Arg Arg Phe Phe Phe Arg Val Ser Glu Glu		
1415	1420	1425
Gln Gln Lys Gln Gln Arg Arg Glu Met Leu Lys Asp Pro Phe Val		
1430	1435	1440
Arg Ser Lys Leu Ile Ser Pro Pro Thr Asn Phe Asn His Leu Val		
1445	1450	1455
His Val Gly Pro Ala Asn Gly Arg Pro Gly Ala Arg Asp Lys Ser		
1460	1465	1470
Pro Ser Gln Pro Leu Arg Thr Val Thr Gln Gln Ala Pro Glu Glu		
1475	1480	1485
Lys Gly Arg Val Ala Arg Gly Ser Gly Pro Gln Arg Pro His Ser		
1490	1495	1500
Phe Ser Glu Ala Leu Arg Arg Pro Ala Ser Met Gly Ser Glu Gly		
1505	1510	1515
Leu Gly Gly Asp Ala Asp Pro Thr Gly Ala Val Lys Arg Lys Pro		
1520	1525	1530
Trp Thr Ser Leu Ser Ser Glu Ser Val Ser Cys Pro Gln Gly Ser		

WO 02/08399

PCT/US01/23092

	1535		1540		1545									
Leu	Ser	Pro	Ala	Thr	Ser	Leu	Met	Gln	Val	Ser	Glu	Arg	Pro	Arg
	1550		1555		1560									
Ser	Leu	Pro	Leu	Ser	Pro	Glu	Leu	Glu	Ser	Ser	Pro			
	1565		1570											

<210> 21
 <211> 4298
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 2564295CB1

```

<400> 21
gccactgaga ggcccaccag gtcctcttt ctccgactct gcagaccaag gactagegta 60
cggactctgc cttagggttc tccaagagga caaggagcct ctaggagct gagagggct 120
cccagagggc aaggctgtcc acgttctccc ggttcgaggc tgcagaagt taccgcgag 180
atgtcagggc accgggaugc agaattcaga cacagcctc cccagaagcc cgggcgcct 240
ggtgcccct ccggcgtgc agcccactt ggaagaagcc tgtgggtta tcacaactt 300
ctcccagag tcaccggag gagagccgg actggacaca agcccgggt gggacaatgg 360
cagtgcctag tctgtggccc tggggagcat gccctgctgt gatctcttc tctttggat 420
ttggcctgga tacagttagg gtgtgcccc gccctgatat tcgctcagag gtggcagagc 480
ttcgtcagct ggagaactgc agcgtgttgg agggccacct gcagatcctg ctcatgtca 540
cagccaccgg ggaggacttc ccggcctca gcttccctcg cctcaccag gtcaccgact 600
acctgctgct cttccgtgtc taaggactgg agagcctgg ccgacctctc cccaacctag 660
cagtcatccg cgggacgctc ctcttccctg gctatgcaet ggtcatctt gagatggcac 720
atctcgtgga cgtggcactg cctgcacttg gggcctgct gcgtggggct gtgctgtgg 780
agaagaacca ggagctctgc caactctcca ccattgactg gggactgctg cagccagcac 840
ctggcggcaa ccacatcgtg ggcaacaagc tggcgaggga gtgtgtgac gtgtgccctg 900
gtgtgtggg tgcctcgtgt gagccctgtg ccaagaccac cttcagcggg cacactgact 960
acagatcctg gacctccagc cactgcagga gagtgtgccc ctgcccocat gggatggctt 1020
gcacagctag gggcagtgct tgccacaccg aatgcctggg gggctgcagc cagccagaag 1080
acctcgtgc cgtgtgtgct tgcccaccct tctacttcca gggctcctg ctgtgggctt 1140
gccgcgaggg cactaccag tatgagctct ggcctgtgt cacagctgag cgtgtgcca 1200
gcctgcactc tgtcccggc cgtgctcca ccttcggcat acaccagggc agttgctgg 1260
cccagtgccc ttctggcttc acccgttaata gcagcagcat attctgccac aagtgcgagg 1320
ggctgtgccc taagagtgct aaggtaggca ccaagaccat cgaactccatc caggcggcac 1380
aggatcttgt gggctgcagc catgtggagg gaagcctcat cctcaactt cggcaggctt 1440
acaactctga gccacagctg cagcacagcc tgggctggt agaaacatt actgcttcc 1500
tcaaaatcaa gcaactcctt gccctcgtgt ccttgggctt tttcaagaac ctcaactaa 1560
tcgggggaga cgcctcgtgt gatggnaact acactctcta cgtctggac aaccagaacc 1620
tacaacagct agggctcctg gtggccgctg ggtcaccat tccctgggc aagatctact 1680
tcgcttcaa cccgcgctc tcttggaac acactctacc actggaggag gtgacagcca 1740
cgcaggtctg gcgaaacaag gctgagatca acccccgcac caacggagac cgcgcgct 1800
gccagactgc caccctgcgc tctgttcca acgtgacgga gccagaccgc atcctgctac 1860
gtggggagcg ctatgagcca ctggagccc gcgacctgct cagcttcatc gtgtactaca 1920
aggagctccc attccagaac gccacagagc actggtgccc agatgctgt ggaaccaga 1980
cttggaaact gctggatgtg gagctgccc taagcggcac ccaggagcca ggggtgacc 2040
tagctccctc caagccttgg acacagtagc cagtgtttgt gcyggccatc acgctaaca 2100
ctgaggagga cagccctcat caaggagccc agagtccat cgtctacctc cgaacgctgc 2160
ctgcagctcc caggtgccc caagagctca tctccactc caactctctc tcccactctc 2220
tggctgctgc gaagccaccg acccagcgca atgggaaact cacctactac ctggtctgt 2280
ggcagcggct ggcagaggac ggcgacctct acctcaatga ctaetgccac cgcgcttgc 2340
ggtctcccac cagcaacaac gatccgctct tcgacgggga agacggggat cctgaggcgc 2400

```

WO 02/08399

PCT/US01/23092

```

agatggagtc cgaactgtgc ccttgccagc acccaacctc tggtcaggtt ctgccccgc 2460
tggaggcgca agaggcctcg tccagaaga agtttgaaa etttctaac aacggatca 2520
ccatccccat atccccctgg aaggtagcgt ccatccaaa gagccccaa agggactcag 2580
ggcggcaccg cgggacagct gggccccctc ggttgggggg caacagctcg gatttcgaga 2640
tccaggagga caaggtgccc cgtgagcgag cgttgcctgag cggcctgcg cacttcacgg 2700
aataccggat cgaactccat gcctgcaacc acgagggcca caccctggc tcagagcccg 2760
ccaccttgtt ctttgccgcg accatgcccc acagagaggg tgatggtatt ccaggaaagg 2820
tggcctggga ggcctccagc aagaacagt tcttctgcy ctggctcgag ccaccagacc 2880
ccaacggact catcctcaag tacgaaatca agtaccyccg cttgggagag gaggccacag 2940
tgcctgtgtt gtcccgtctt cgaatagcga agtttggggg agtccacctg gccctgtgc 3000
ccctggaaa ctactctgcc aggttaggg caaactcact ggctggcaat ggcctgtga 3060
cagacaggtt gccccttac atccttggcc cagagagaga ggatgctggg ggcctgcatg 3120
tctctctcac tggccacctc gtgggctca cgtgctcat cgttcttctt gccctgggtt 3180
tctctctcac caagaagaga aacagaacc tgtatgcttc tctgaatcca gactacttca 3240
ggcctctgta tatgtatgt cctgatgaat gggaggtgcc tgggagcag atctcgataa 3300
tccgggaact gggccagggc tcttttggga tggatatga ggggctggca cggagactg 3360
agcctggaga ggaagccaca cccctggccc tgaagacgtt gaatgagctg gccagccac 3420
gggaatgcat tgacttctc aaggaagctt ctgtcatgaa agcctcaag tgcacccatg 3480
tggctgctct cctgggtgtg gtatctcagg gccagccaac tctgtctctc atggagttaa 3540
tgaccctgtg ggaactcaag agcactcttc gatctttgcy gccctgagga gagaacaacc 3600
ctgagctccc acagccagca ttgggggaaa tgatccaaat ggctggtgag attgcagag 3660
gcattgctca ccttctgccc aacaagtttg tgcaccgaga tctagcagcc cgaactcgca 3720
tgggtgccc ggaactcaac gtcagatcg ggaacttgg gatgactcgg gacgtgtatg 3780
agacagacta ttaccgcaag ggtggaagg ggtgctgccc cgtgctctg atggccccg 3840
agctccctcaa agatgggabc ttcaccacc actcgatgt ctggtctctt ggcctggtag 3900
tctgggagat tbtgacctg gcagaacac cctaccaggg cctgtccaat gacagagtg 3960
tgaagctgat catggtggc gggctctcgg agggactgga gggctgtccc cttaagctg 4020
agagctgat gagccgtgc tggcagccga acccaagcct cggcccatct ttaacacaa 4080
ttctggacag cacaagggag gactgcccgc cctctctcc cctctctcc ttctactaca 4140
gccccgaatg cggggggccc cgggctccc tgcctaccac cgaatgcagag cctgactctt 4200
caccactcc aagagactgc agcctcaaa atgggggtcc agggcactga gggcactct 4260
atctctggc tggcctccca tggggagaca ggaagggg 4298

```

```

<210> 22
<211> 2863
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Incyte ID No: 2837050CB1

```

```

<400> 22
atgatggaag aattgcatag cctggaccca cgaaggcagg aattattgga ggcagggtt 60
actagagtag gtgttagtaa gggaccactt aatagttagt cttccaacca gagcttgtgc 120
agcgtcggat ccttgagtga taaagaagta gagactccc agaaaaagca gaatgaccag 180
cgaatcggga aaagaaagc tgaaccatat gaaactagcc aagggaaggg caactcctag 240
ggacataaaa ttagtatta ctttgagcga cagtagaac agccctcta tggtttagat 300
ggcagtgctg caaggaggc aacggaggag cagtctgctc tgcacaacct catgtcagtg 360
atgctgcaaa aacctcggct tgacaacagag cacgtggcgc aaaggggagc tggcctctgc 420
ttcacttttg ttccagctca gcaaacagat cctctatcta cgggatctgg caacacagag 480
cattctgca gctcccaaaa acagatctcc atccagcaca gaacagccca gtccgacctc 540
acaatagaaa aatatctgc actagaaac agtaagaatt ctgacttaga gaagaaagg 600
ggagaatag atgatttatt aagagccaac tbtgatttga gacggcagat tgatgaacg 660
caaaagatgc tagagaata caagyaacga ttaaatagat gtgtgacat gagcaagaa 720
ctcttatag aaaagtcaaa acaagagaag atggcgtgta gagataagag catgcaagac 780

```

WO 02/08399

PCT/US01/23092

```

cgcttgagac tgggccaact tactactgtc cgacacggag cctcatttac tgaacagtgg 840
acagatggtt atgcttttca gaatcttatic aagcaacagg aaaggataaa ttccacagagg 900
gaagsgatag aaagacaacg gaaaatgtta gcaaaagcga aacctctctc catgggtcag 960
gcccctctcy caaccaatga gcagaaacag cggaaaagca agaccaatgg agctgaaaa 1020
gaaactttaa cyttagcaga ataccatgaa caagaagaaa tcttcaaatc cagattaggt 1080
catcttaaaa agggagaagc agagatccag gcagagctgg agagactaga aagggttaga 1140
aatctacata tcagggaact aaaaaggata cataatgaag ataattcaca atttaaagat 1200
catccaacgc taatgacag atattgttg ttacatcttt tgggtagagg aggtttcagt 1260
gaagtttaca aggcatttga tctaacagag caagatagc tagctgtgaa aattoaccag 1320
ttaataaaa actggagaga tgagaaaaag gagaattacc acaagcatgc atgtagggaa 1380
taccgatctc ataaagagct ggatcatccc agaattagta agctgtatga ttacttttca 1440
ctggatactg actcgttttg tacagtatta gaatactgtg agggaaatga tctggacttc 1500
tacctgaaac agcacaatc aatgtccagag aaagaggcct ggtccattat catgcagatt 1560
gtgaatgctt taagtactt aatgaaata aaactccca tcatacacta tgacctcaaa 1620
ccaggtataa tctcttttag aatggtaca gtgtgtggag agagaaaaat tacagatttt 1680
ggtctcttca agatcatgga tgatgatagc tacaattcag tgggtggcat ggagtgaca 1740
tcacaagggtg ctggcactta ttggtattta ccaocggagt gtlttgtggt tgagaagaa 1800
ccaccaaga tctcaataaa agttgatgtg tggctggggg gtgtgatctt ctatcaggtg 1860
ctttctggaac ggaagccttt tggccataac cagtctcagc aagacatctt acaagagaat 1920
actattctta aagctctgca agtgcagttc ccgcaaaagc cagttagtaac acctgaagca 1980
aaggcgttta ttgcagcatg ctggcctcac cgaaaggagg actgcatlga tgcccagagc 2040
ctggcctgtg atcctactt gttgctcac atccgaaagt cagctctcac aagtagccct 2100
getggagctg ctattgcatc aacctctggg gcgtccaata acagttcttc taattgagac 2160
tgactccaag gccacaact gttcaacaca cacaagatgg acaaatggcg ttccagagcg 2220
ggtttggaac atagcgaatc tgatggatc tgatgaacc tgaaccaggt gcttttattt 2280
tctgtctttt ttccatcca ctgagcatga cagcatggat tctctttaag gagaacctt 2340
gggcagctcc agcccggcct catagaaaaa ggcocggcat gaggttctgg cgtcaatggc 2400
cactgtgtat gctgctctg agtgaggaaa aaactaaaaa gaaaactggg ttccatgtac 2460
tgtgnaactg aanaactgca gactcacggg ggtctctgat gcaatgcttc agatgaagat 2520
tgtggaactg aanaactgca ctgagagccc gggccacgtg gctcatgctt gtaactcag 2580
cactttggga ggcocaggaa ggtggtatcc aagtcaggga gatcgagacc atcctgctca 2640
acacagtgaa accccgtctc tactaaaaat acaaaaaaat tagccaggct tgggtgtggg 2700
cgcctatagt ccagctact tgggagactg aggcaggaga atgtcgtgaa cccgagaggg 2760
gagacttca gtgagccag atcacgccac tgcactccag cctggggagac agagtgacac 2820
tccgtctcaa aaaaataata tataataaaa taataaaaaa aaa 2863

```

```

<210> 23
<211> 1494
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7474590CE1

```

```

<400> 23
atgtactctg acagcgagga tgagtcatca gagctcagca ctgtgctcag catgtttgag 60
gagaaggagt tcaccaggca gtacacgctc ctgaagacct tgagccagca tggcactact 120
gaagtggagg taigtctcca tcacctcaaa ggtgtcacag ttgctgtcaa agctctgaag 180
taccagaggt ggtgggagcc aaagtttcca gaagtcaaaa tcatgaagat gctcagccac 240
cetaaacttg ttctccttct gcaagtgata gagacagaa agaacattta tctgattatg 300
gaagtggccc aaggcacaca gctacataat ctagtccaag aggetaggtg cctgaaggaa 360
gatgaagcaa gaagcattt tyftcagttg ctcagtccca taggctactg tcatggtgaa 420
ggtgtgttcc acagagacct aaagcctgac aatgtcatag ttgatgagca tggaaatgct 480
aaaattgttg actttgggct aggtgccaga ttcatgctgt ggcagaattt ggaaggctg 540
tgtggagcct tccagttcat tctctcagag atattcttag ggtccctta tgatggccca 600

```

WO 02/08399

PCT/US01/23092

```

aaagtagaca tatgggcett gggggttctt ttgtattata tggtagacagg gatttttcca 660
ttttaggggt ccaccttgtc agaaattagc aaggaagttc tacaaggagg gtatgaaatt 720
ccttataatc tctetaaaga ctttaaggagc atgataggcc tgttattggc aacaaacgca 780
aggcagaggg caactgcaca agacctccta agtcatccat ggcttcagga aggggaaaag 840
actatcacat ttcattccaa tggagacacc agctttccag acctgacat aatggcagcc 900
atgaaaaata ttgggtttca tgtgcaggac attagaaat cattaaaca cagaagattc 960
gatgaaacta tggctacata taacttactg agagctgagg catgtcagga tgatggcaat 1020
tatgttcaaa caaagttaat gaatccaggg atgccacctt tcccttcagt aacagactct 1080
ggagcttttt cctgcccctc taggagaagg gccagtgaac ctctcttaa agtattagtc 1140
tcactactg aagaacatca attaagacaa actgggggga caaatgcccc ttttccacc 1200
aagaaaaaac ccaactatgg cagaagttag aaacagaaac gtgccatgac tggccctgt 1260
atttgtttac tgaagaaacac ttacatagat acagaagaca gcagcttttg cactagctcc 1320
caggcagaaa agacttcaag tgatccagag aaaagtgaga ctccaacttc atgcccctctg 1380
acacctaggg gctggaggaa atggaagaag agaattgtag catgcatcca gcaattgtgt 1440
tgtgacagt tgcctcaaaa aaaatgtccg aggggtgtgc atcccaaaa gtga 1494

```

<210> 24

<211> 2341

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 7474594CB1

<400> 24

```

atgtcagggc tgggtctgat gctggcggcg cgggtgoatt tgggcagctc cccgctctgc 60
cgctgcccgc gccgtgcccc aaggaggatc ggggcccggc cgggcccggg tcatccgggt 120
cgggaaggccg ccgcccggcg agggagcggg tcaccocaag ccgcaactgag ccgcccgcgc 180
cccgcccggg ccccggggga tgcgcccgcc cgagctgctg cctccgcccg cgcgcgagcc 240
gcagcccgag cgggacacaga gcaggtatag ggcacctca gggcagcccc ggccgacacc 300
cctccctctg gctggcggat gcagtgccta cggcccgcct ttaaggacga aaccaacatg 360
agtgggggag gggagcaggg ccacatcctg ccggccaact acgtggtcaa ggatcgctgg 420
aaggtgctga aaaagatcgg gggcgggggc tttggtgaga tctacgagcc catggacctg 480
ctgaccaggg agaattgtgc cctcaaggtg gagtccagcc agcagcccaa ccaggtcctc 540
aagatggagg tggccgtgct caagaagtgg caagggaagg acctatgtgt caggtteatt 600
ggctgtggca ggaacgagaa gtttaactat gtagtgatgc agctccaggg ccggaacctg 660
gccgacctgc gccgtagcca gccgcgaggg accttcacgc tgagcaccac attgcccgtg 720
ggcaagcaga tcttgagtc catcgagggc atccactctg tgggcttctt gcaccgtgac 780
atcaagcctt caaactttgc catgggcagg ctgccctcca cctacaggaat gtctatatg 840
ctggacttgc ggtggcccgc gcagtacacc aacaccaggg gggatgtggc gccccctcgg 900
aatgtggccc ggtttcgagg aacggttcgc tatgctcag tcaatgcca caagaaccgg 960
gagatggccc gccacgacga cctgtgttcc ctctctaca tgcctgttga gtttcagtg 1020
ggccagctgc cctggaggaa gatcaaggac aaggaacagg tagggatgat caaggagaag 1080
tatgagcacc ggtatgctct gaagcacatg ccgtcagagt tccacctctt cctggaccac 1140
attgcccagc tcgactactt caccagccc gactaccagt tgatcatgte agtgtttgag 1200
aacagcatga aggagagggg cattgcccag aatgagccct ttgactggga gaagcagggc 1260
accgatgccc tccgtcccac gagcacctct acccccgcgc agcagaacac ccggcagagc 1320
gcagccatgt ttgggttgtt caatgtgagc ccagtcctgt gggactgtct ccgggagaac 1380
accgagatgt tgcctacaggg agagcaccct agtgaccagg agaatgcacc cccaattctg 1440
cccgggaggg cctctgaggg gctggcccc agtcccacc ttgtcccaca cccgggggtt 1500
cctgagctg agtctggga yagacagat gtcaccyga acaactccg gatcaacac 1560
ggcaagtaaa ctgcccgcag ggcgaagggc gtgggtggcc tttttctca cccccatc 1620
ccagccttgt gcccccgcct tglctccctc aagcaccctg tccccgcca tctgctgct 1680
tgcccgtgct ctgtttcccg gtcctccccc gcactagcct cgctgtgtct tccatcatca 1740
tcactctctg tctccttcc cctgaggaga ccatccgccc acagccgctt catcagcccc 1800

```

WO 02/08399

PCT/US01/23092

```

agctcatggc actccctct cctgcagagc cctgtgtgg aggaggaaca gagccgagcc 1860
atgggggtcc ccagctcccc agtgcgtgccc ccccagact ccccacaaac cccagtcctg 1920
tctctggctt accggagggt gaacagccct gagtcagaaa ggctgtccac ggcggacggg 1980
cgagtggagc taactgagag gagggtgggtc tggggccaag ggcatggttg gggcccaagg 2040
ccctctccgc cttaacgtgg ctgggtctga ggaaaagtta gatgtgtgac ggaagtgagg 2100
agaccctggg aagtgcctgag agggttatac ttgggcctgg ggtcagactc agttggggcc 2160
agagacaggg cctgggagaa ccagtggggg atccagagag gtcccggctc atgccagaaa 2220
acgtaattgg gtgagtgca gctgcaggag gacaggtgg ggcgctggg cccaggaagg 2280
gtgagggggc agttgggttg tgggtgtgtg tgctccaga atctcttctc ctagagacta 2340
a 2341

```

<210> 25

<211> 2552

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 7477585C1

<400> 25

```

cgcggtgtct ggcgctcggg ggggtgtggt gccctagtt tgaggcctgc ccgattacc 60
gcaagaettg ggcagcccc ggccgcctc caccacgac agggaaagga acctaatct 120
catctttaaa ataaggagaa ttaactgagc acctgaagga cctcttccag ctggaagtc 180
tgaactgacc aacactggat gaatttgacc abttcttagg agactggaat gttaaagtt 240
tataaatgaa tgaaccagtt ctctctgttt tggagcaatg ctgaaattcc aagaggcagc 300
taagtgtgtg agtggatcaa cagccatttc cacttatcca aagacctgta ttgcaagaag 360
atcgtgctt caacnaaac ttggcagtyg aagttttgga actgtctatc tggtttccga 420
caagaaagcc aaacgaggag aggaattaaa ggtacttaag gaaatctctg ttggagaact 480
aaatccaaat gaaactgtac aggcacaatt ggaagcccaa ctctctcca agctggacca 540
cccagccatt gtcagttccc atgcaagttt tggggcaga gataattctc gcattatac 600
ggagtaactg gaggggccag atctggagca taaaattcag gaatataaac aagctggaaa 660
aatctttcca gaaaatcaaa taatagaatg gtttatccag ctgctgctgg gagttgacta 720
catgcatgag aggaggatac ttcacgaga cttaaagtca aagaatgtat ttctgaaaa 780
taatctctct aaaaatggag attttggagt ttctcagctt ctaatgggat cctgtgacct 840
ggccacaact ttaactggaa ctccccatta tatgagctct gaggtctgta aacaccaagg 900
ctatgacaca aagtgggaca tctgttctct ggcatgact ttgtatgaga tggctgcat 960
gatcatgca ttgcctggct ccaatttctt atccattgtt ttaaaaattg ttgaaggtga 1020
cacacctctc ctccctgaga gatattcaaa agaactaaat gccatcatgg aaagcatgt 1080
gaacaagaat ccttcattaa gaccatctgc tatcgaatt ttaaaaatcc cttaacctga 1140
tgacagacta cagaacctaa tgtgtagata ttcagaatg actctggaag acaaaaattt 1200
ggattgtcag aaggaggctg ctcatataat taatgccatg caaaaagga tccacctgca 1260
gactctgagg gcactgtcag aagtacagaa aatgaccca agaaaagga tggcgtctgag 1320
gaactcccag gcggctgatg agaaagccag gaagctgaaa aagattgtgg aagaaaaata 1380
tgaagaaaaa agcaaacgaa tgcaagaatt gagatctogg aacttccagc agctgagttg 1440
tgaatgactc catgaaaaaa cacatttaaa agaatggaa gaaaaggagg agcaacctga 1500
gggaagactt tctgtttcac cccagacgca ggaatgaag aggtggcagg gcagggaaga 1560
ggaaatctgat gaaccaactt tagagaacct gctgagctc cagcctatc ctccatgga 1620
ctcccaagaa cttgaaatcaa ttgtagagga tgccacctct gaccttgat accatgagat 1680
cccagaagac ccaactgtgg ctgaaagata ctacgctgat gcahttgatt cctatttgt 1740
agagagtgat gaggaggagg aagaaaatgc gttagaaga ccagagaag aaatcaggaa 1800
tgagggtacc cagcctgctt acsgaaacaaa ccaacaggac agtgatatcg aagcgttggc 1860
caggtgtttg gaaaatgtcc tgggtttcac ttctctagac acaaaagaca tcaccacct 1920
gactgaagac atgtcccag gaccaccaat ttcaacagt gtgatggcca ggaaccaat 1980
gaaacgatg agggatcag ccatgcagaa gctggggaca gaagtattg aagagttota 2040
taattacctc aagagagcaa ggcactcaga tgctagcga gcaagatcc gcgagtttt 2100

```

WO 02/08399

PCT/US01/23092

```

ggaaaaagtg gtgcctcaag ccagcgactg ttttgagtg gaccagctcc tghtacttga 2160
agagcagttg ctgatoacga tgggaaaaa accctactctc cagaaccate tctaggccaac 2220
tatcaaaaaag aagcagaagt tcaagtggac aaatttatgt gaaaattcat ttaacatata 2280
agotgaaactc tattatgggg aatggatata aaagcagagc tcccatcttg accttcaatt 2340
cctcatcaga agtactggct tcttttagaga gtagtaagca tggctgcta tghttggagt 2400
cataagtggt atttggacta taccctgaga taagcttata gatcaagttt ggctcccttg 2460
aaaaagcattt ctctcatgtg egccctcagg gcttccagca ggattgagtc accctgaoga 2520
tgaccgggga gaagccgtgt gctcttcatt at 2552

```

<210> 26

<211> 2176

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 7477587CB1

<400> 26

```

ctcaaccgcc tcccaccag ccccgacatc acaaacgctg ggettctctg ccgtccgaa 60
ttctgctgtg gctcccagt gccacgggcc atcaagcccc aggtlitleak gtgggcccc 120
gggaccgcc agcagggagg acccgaaatg gcacacatcc agaacttoga ggtctabacc 180
tcaagcgcac tgtgggggag aagccccgg aagccccca cccccaccg gcgagagagc 240
ctcagttccc cgtccgagcg gcccccagc ggccgcagcg cgttggctc ggccccgctg 300
cgcacagtc cgcgatgga gcccccggc cggggcggc gcaggagtcg cccctggtc 360
ggccctccc tgcagagccc gggctcggc cggctgtaca ggcggggaa gctgatcgc 420
aagggcgcct tcagccgctg ctacaagctg acagacatg ccaccagcg cgttttccc 480
ctcaagtggt tgcctgtgg cgggctggg gccgggtggc ttgcgccga gggaaaggtg 540
gagcgtgaga ttgcccgtca tagccgctg cgaacccgca acatcgtggc ttccaccga 600
cactttctg accgcgacca cgtgtacatg gtgctggagt actgcagccg ccagtctttg 660
ggccacgtgc tggggcggc gcagatcctg acggagccag aagtgccga ctacctggg 720
ggcctgtgca gggcctgcy ctacctgcac cagcgtgca tccctgaccg cgaactgag 780
ctcagtaact tcttccctaa caagaacatg gagtggaaga ttggagacct gggactggc 840
ggcaaggtgg gggcggggg ccgctgccac aggtacaagg tgcctgactgg caccaccoc 900
ttcattggct caccctgtc ggagatgtac caaacatcc gtgagggcca ctaccgccaa 960
cccgctcacc tgtctgccaa tgcgcgccg ctcactgtgc acctcctagc acccaaccg 1020
ggcagcggc ccagcctgga ccacctgctg caggacgact tcttcacaca gggtttca 1080
ccagaccggc tgcggccca ctctctgcc acgtccccca tcttcgcaat acccccgc 1140
ctgggcagga tctccggaa ggtggccag cggctgctca cccagtgcc gccaccctg 1200
cccttcagc ctaaaagggc ctcggtcca ggagaaggtg ggccagacc tgactccatg 1260
gagtgggacg gcgagagctc cctgtctgcy aaagaggttc cctgctgga agccccatc 1320
cacttggctg cacaagggac cctgcagagt gacctggcg ccacacagga cccccggga 1380
gagcagcagc ccatctctg gggccccaaa tgggtggatt attccagcaa ataccgctt 1440
ggctaccagc tcttgacg gggcgccag ggaaggcacc cacatggccc tgcgacccc 1500
cggaggtatt tattaagcac ctactgtgca cactacagg tgcctcctg ctgcaagt 1560
tgctacatgc ccaactggc gaggctggaa gccttgcgcc tgagggatgt gccggcctg 1620
ctgggcgcca agctggcctg gctgagctc tttgcggct gcctgcggcg gcggctgctg 1680
gaggggggga cctccccac acctgtgcca cctgtggac ccggcctctg cctcctgct 1740
ttcctggcct ctgagcagc cctgtctgct ctgttcagca atgggatgt gcagtgagc 1800
ttcagtgag tccggccca actggtgctg agtggcgagg gtgagggttt gcagctcac 1860
ctctggagc aggggtcccc tggcacctcc tactccctg acgtcccgg gagccagcg 1920
tgcgcccca ccacgggaca gcacctcac cacgcccctc ccatgctgca ggtatctag 1980
tgccccag ggtcagagtg gacctgca tggtagtgc agggaccag gctccattc 2040
cattcctg getccccag agggctgtc ctgggggaga gctgggggc acacgggag 2100
tgggtcttg cctgtggca tgaactgtca acccagact tgcgggac tcttcttt 2160
tcattaaga caattc 2176

```

WO 02/08399

PCT/US01/23092

<210> 27
<211> 4277
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7594537CB1

```

<400> 27
tgtgagcaga gttcttgaag ctccactcct cctggggaag ccgagctgtg tgggagcctt 60
cttactgtgc cgggagcgtg tgaattggaa aggatcctga gaactggcta gtcccagttc 120
ctctccggaa aagcagctgg tctcgtctca gagatgcgct cagctttcgc ctgcatcaca 180
ctgcattctt taacacatta ttgaaattac aagcatccaa agcagtttca tggggacaga 240
ttgcataatt tgaagcctg aggtatttta tcatgaaaca tggcatgtgg aatctttgaa 300
gcgatagacc ctgcgcaca cctgaataaa gaatotttta cctggtatgt gacagagctt 360
ctcaccacca ccatgacaaa ccaggaaaaa tgggccacc tcagccttc ggaattttcc 420
caacttcaga aatagctga gtattctaca aagaattaa aggatgttct tgaagaattc 480
catggtaatg gtgtgcttgc aaagtataat cctgaaggga caatagattt tgaaggtttc 540
aaactattca tgaagacatt cctggaagcc gagcttctgt atgatttcaac tgcacacctt 600
ttcatgtcat tttagcaaaa gtttctctat tctagtccaa tggtaaaaag taagcctgct 660
ctctatcagc ggggtctgag aatgaaataa ggtgcoatac cccctcccc aactacttct 720
cctgcaataa cgtgtccccc agaagtaate catctgaagg acattgtctg ttactctgtc 780
ctgcttgaaa gaggagacc ttggagataag cttagattta tgtttcgcct ttatgacacg 840
gatgggaatg gcttctctga cagctcggag ctgaaaaata tcatcagtca gatgatgat 900
gttgagaat acccttgatg ggatgctact gaacttaate caatcctca tgaatgatg 960
gaagaattg actatgatca tgatggaaac gtgtctctgg aggaatggat tcaaggagga 1020
atgcaacaga ttccactctc tgtgtctctg ggtttgaaa ataatgtgaa ggatgatgga 1080
cagcagctgt ggcgactgaa gcacttaac aaactgctc attgcaacct ttgcttgaac 1140
atgtgattg tggctgggaa gcagggcctc tgcgttctct tctgcaagta cacagtccat 1200
gagcctgtg tggctcgagc acctccctct tgcatacaga cctatgtgaa gtccaaaagg 1260
aaactgatg tcatgcacca ttactgggtt gaagtaact gcccaccaa gtgtgataag 1320
tgcacaaaaa ctgttaaatg ttaccagggc ctgacaggac tgcattgtgt ttggtctcag 1380
atcacactgc ataataatg tcttctcat ctaaaacctg aatgtgactg tggacctttg 1440
aaggaccata ttttaccacc cacacaatc tgtccagctg tactcgagac tctgccactt 1500
tcaggagttt cagttctcga ggaagacaaa tcaacagtga aaaaggaaaa gagtggttcc 1560
cagcagccaa acaaaagtat tgacaagaat aaaatgcaaa gagccaactt gtttactgta 1620
gatggacaag gcttcagcgt cactcctgtg cctggtactc acccaacttt agtttttgtg 1680
aacccaaaa gtggtggaaa acaaggagaa cgaatttaca gaaaattcca gtacttatta 1740
aatcctctgc aggtttacag tcttctctga aatggaccaa tggccaggtt aaacttttcc 1800
cgtgatgttc ctgacttcag agtgttagcc tgtggtggag atggaacctg gggctgggtt 1860
ttgattgca tagaaaaagg caatgtaggc aagcatcctc cagttgcgat tctgctctt 1920
gggactggca atgatctagc aagatgctct cgtggggag gagtttaca aggtgagaat 1980
ctgatgaaaa ttctaaaaga cactgaaac agcacagaaa tcatgttga cagtggaag 2040
ttgaaagta taccaatga caaagatgag aaaggagacc cagtgcctta cagtatcctc 2100
aataattact tttccattgg cgtggatgac tccatlgcac acagattcca catcatgaga 2160
gaaaaacacc cagagaatt caacagtaga atgaagaaca aatttttgta ttttgattt 2220
ggcacactcg aaactttctc agccacctgc aagaagctac atgaaatctg agaaatagaa 2280
ataccagca tgcattgagg atccaatctt tggggagagt ctaagaaaag acgaagccat 2340
cgaccgatg agaaaaaagg gtctgacaaa aggaccaccg tcacagatgc caagagttg 2400
aagtttgcga gtcaagatct cagtgcaccg ctgctggagg tggctggctt ggaagagacc 2460
atggagatgg ggcacatata cacagcctg aaaagtctg gccggggctt ggtctagtc 2520
tctgcggtg tcatcaggac gagcaagtct ctgccaatgc aaattgatg ggaagcctgg 2580
atgcagacc catgcacaa aaaaattaca cacaaagacc aagcccaat gctgatggc 2640
cgcctccaa aaacgggttt attctgtctc ctgctcaaaa ggacaagaaa ccgaagcaag 2700

```

WO 02/08399

PCT/US01/23092

```

gaataatcct ggtgtgttcc actcttagaa attgaattag cataatgggg ccatggaaca 2820
catatgctgg aaalccttga accatttcaa gtctcctgct catgcaaat catggaagtg 2880
gtttaacagt tttgtttact aagctaagt aaaattcagc tattagaaaa tttattgtct 2940
cagtttttat aggcactctt gcatagaaga ggcagaagtt tacctgaagt gatctgcat 3000
attttgggt catgcattcc catagatttt tacatctccc acccaactct tcccgaattt 3060
cottttacta aocgtgtgaga aaaaccctgt aacatgaaa aaggaaatac catgggaaac 3120
gtgatttcca gttgtattcc aattattacg aagcaactaat cagtaacgct acaatgatca 3180
taattgcaga ttgotatagc ttcccttttt agaactcagt tatcaactac ctatgacttg 3240
aggagaacct ttaatttoga agattttatt aaatagtga ctacaatacc ttgctatata 3300
tacatagttt ttcttcaaca tcttaactct tctgagtga aataaaaaa tcagcataa 3360
ggttttcca tgctgaaaaa tagaacgagg tttttattt gcttagtttt cttttaatt 3420
ccagaataaa gtgaaaaaat gttactgac agtcaagtgt ggtaaatgg caagcctgt 3480
tcotttctgc atgagaatct aggagagaat tcataaccac accaataacg aaatagaagt 3540
tttaaacat gtgocctaat aatgtgttcc ccaccaaga ttcagaaaac aatgcttgag 3600
agaaatgggt taatgcataa ttaattaacg atgtggagc aaatttaggg ttctgtgat 3660
taattttgtg atgactaaaa tgctggaag caagttagtt gccattaat tatgatata 3720
attctaacct ttcacagaca gacaataagc cagacaacac aatcaagct caatagatga 3780
ttcttctgct tttctagcca tttataata taggtgtaat tttcatgga tcagttaagt 3840
acactggaag gaagtaaatg atgtatcag tttatttcta gtataaatgg gtacctgaa 3900
taatactgag cctctggaag cgaatcatgc atgcaatag ctcctctct ctcactact 3960
ccactcccat ctttatgaca ttcaaatgt ttatttgaa acacacgct agactactgt 4020
tgaaggtgtt catggcatag ttggagtctc tgactgttta aagaatcnc agaacgtac 4080
ttttcttcta gttttcatt aagcctatga tgtaaaatga aatgctctg agcactctg 4140
taatatgttt cactcatatt gacctcatc tcatcattgc atgttttat tttcaaca 4200
tgccataagg aaacgagtg cctgactgc atgatttatt agtttctct cactctgat 4260
taaagtctca atgattt 4277

```

```

<210> 28
<211> 2616
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Incyte ID No: 70467491CB1

```

```

<400> 28
atgtccacta ggacccatt gccaacggtg aatgaacgag acactgaaa cgtgtattg 60
ccgcacacgt cacatggaga tggcgctcaa gaagttacct ctgtaccag ccgtcagga 120
gctcgggtga gaaactctat agcctcctgt gcagatgaac aacctacat cggaaactac 180
agactgttga aaacaatcgg caaggggaat tttgcaaaa taaaattggc aagacatctc 240
cttacaggca gagagaaaaa tgttagaata tccaagaaa ttgataatt tctagaaaa 300
catgacttac caaattaac tctagaaaaa aatcgataca catcagtaac aacagaagt 360
gagaagtag ttaacatatt gccaaacctg gaattcatga ttgaattctt tgagactctac 420
tctataggtg aagtatttga ctatttggtt gcacatggca ggatgaagga aaaagaagca 480
agatctaaat ttgacagat tgtgtctgca gttcaatac gccatcagaa acggatcgt 540
catcgagacc tcaagcctga aaactctatg ttgatgccc atatgaacat taaaatagca 600
gatttcggtt ttgcaatga atttactgtt ggcgtaaac tcgacacgctt ttgtggcagt 660
cctccatcag cagcactcga gctctccag gccagaat atgacgggccc agaagtgat 720
gtgtggagtc tgggggtcat tttatccaca ctagtcatg gctcactcc ctttgatggg 780
caaacctaaa aggaactgag agagagatga ttaagagga aatacagaat tccctctac 840
atgtctacag actgtgaaaa ctttctcaaa cgtttctgg tgctaaatcc aattaaagc 900
ggcactctag agcaaatcat gaagcagagg tggatcaatg cagggcctga agaatgaa 960
ctcaaacat ttgtgaacc agagctgac atctcagacc aaaaaaat agatattatg 1020
gtgggaatgg gatattcaca agaaatctc caagaatctc ttagttaagt gaatacgt 1080
gaaatcacag ctacatattt gttattgggg agaaatctt cagagctgga tgctagtat 1140

```

WO 02/08399

PCT/US01/23092

```

tccagttcta gcagcaatct ttcacttget aaggttaggc cgagcagtga tctcaacaac 1200
agtactggcc agtctectca ccacaaagtg cagagaagtg tttcttcaag ccaaaaagcaa 1260
agacgctaca gtgaccatgc tggaccagct attcttctct ttgtggcgta tccgaaaagg 1320
agtacagacca gcactgcaga tagtgacctc aaagaagatg gaatttctct ccggaaatca 1380
agtggcagtg ctggtggagg aaaggggaatt gctccagcca gtcccatgct tgggaatgca 1440
agtaactcta ataaggcgga tattctcgaa cgcaagaaaa gctccactgt ccttagtagt 1500
aacacagcat ctggtgggaat gacacagcga aatacttatg ttgcaagtga gagaactaca 1560
gctgatagac actcagtgat tcagaaatggc aaagaaaaca gcaactattcc tgatcagaga 1620
actccagttg ctccaacaca cagtatcagt agtgacagca ccccagatcg aatccgcttc 1680
ccaagaggca ctgcccagtc tagcactttc cacggccagc cccgggaacg gcgaaccgca 1740
acataaatg gccctectgc ctctcccagc ctgtcccagc aagccacacc attgtcccag 1800
actcgaagcc gaggctccac taactctttt agtaaatcaa ctcaaaact cacaaggagg 1860
cttccaactg aatatgagag gaacgggaga tatgagggct caagtgcgaa tgtatctgct 1920
gagcaaaaag atgaaaaaca agaagcaaaag cctcgatccc tacgcttccac ctggagcatg 1980
anaaccacta gttcaatgga tcccggggac atgatgcggg aaatccgcaa agtgttgac 2040
gccaataact gcgactatga gcagagggag cgtcttctgc tcttctgctg ccacggagat 2100
gggcagcggg agaacctcgt gcagtgggaa atggaagtgt gcaagctgcc aagactgct 2160
ctgaacgggg tccggttaa cgggatatcg gggacatcca tagccttcaa aaatattgct 2220
tccaataatg ccaatgagct aaagctgtaa cccagtattc atgatgtaa ttaagttagc 2280
attaaagtgt tttcctgaac actgatggaa atgatatagaa taatatttag gcaataactg 2340
ctgcactctc taatatcatga aattaaagtc tgaggacgag agcaaaaaaa aaaaaaagg 2400
gcgcccctcg agccgctcga gcgcaattcg gctcgagatg tcaatgggtg agagggaaga 2460
aggggaggtt ggggggctcc ttcctctcag aactgaaagt tttcccaact gctctctctc 2520
cagtggtctc ccaggtgcc aaccacaaag ctttctccac agtgataccc ttatattttt 2580
acttcccctt gactcatatg ttttaacatg aattttt 2616

```

<210> 29

<211> 1253

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Ineyte ID No: 7478559C81

<400> 29

```

ctggcccctc ctctaccact cccactccct cgcgggacc ccccgccggg gctagcgtct 60
gcgcgggctc cgaggggggtg gggctgctgg gaatggctgt gccccctcg gccctcagc 120
cgcgcgcgtc ctbtcaactg agggagcaca cgccttgccc gcagtgtca tggggcatgg 180
aggagaaggc ggcggccagc gccagctgcc gggagccgcc gggccccccg agggccgccc 240
cgtcgcgta ctctcgcatt tccgtggacc cggagccatc ccttcccggg gccctcggcc 300
tcataccagg gctgcccggc cattggaaac ccgagcaagt tcygaccaag cgtctcatg 360
atggcatcac caacaagctg gtggcctgct atgtgggaga ggacatgca gactcgtgc 420
tggtcgggt gtatggggag cggacggagc tgcgtgggga ccgggagaat gaggtcagaa 480
acttccagct gctcggagca cacagctgtg ccccaaaact ctactgcacc ttcagaatg 540
ggctgtgcta tgagtacatg cagggtgtgg ccttggagcc tgagcacatc cgtgagcccc 600
ggcttttcaa gttaatgcc ttagaatgg caaagatca tactatccac gccaacggca 660
gcctgcccga gccatcctc tggcacaaga tgcacaatta tttcacgctt gtgaagaaag 720
agatcaacc cagcctttct gcagatgccc ctaagtaga ggtgttggaa cgggagctgg 780
cctggctgaa ggaacatctg tcccagctgg agtccccctg ggtgttttgt cacaatgacc 840
tgctctgcaa gaatatcatc tatgacagca tcaaaagtca cgttcgggtc attgactatg 900
aatatgctgg ctacaactac caagcttttg acatggcaca ccattccat gattttcag 960
gcgtgaaatg ggtggattac tgcctgtacc cggcggggga gaccagctg cagtggctgc 1020
actactacct gcagccaca aaggggatgg ccgtgacccc cagggaggtg caaagctct 1080
acgtgcaagt caacaagtt gccctggcgt ctcaactctc ctgggtctc tgggcccctca 1140
tccagaacca gtaactccac atcgaatttg atttctcag gtaogcagtg atccgattca 1200

```

WO 02/08399

PCT/US01/23092

accagctactt caaggtgaag cctcaagcgt cagccttga gatgccaag tga 1253

<210> 30
<211> 1790
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1698381CBI

<400> 30
ttaaagaagg agttccctta taggagatgg aagaaacggc cattaatccg gggacttttt 60
atgctggaaa caaacctgaa ggtacaggtt cggcccggaa gttataccca ccaagagaag 120
tatgtccgga attgtgggtt ctgcagtcac tgacttcaag aatgaagccg cggaccctcg 180
cgytgcagca ttgtactgca agtcaatcga tacaataatt taagtcaact cagctataat 240
ggaaaagtat gaaaaattag ctaagactgg agaaggtctt tatggggttg tattcaaatg 300
cagaaacaaa acctctggac aagtatagc tgttaaaaaa ttgtgggaat ctgaagatga 360
tcctgtgttt aagaaaatag cactaagaga aatcgtatg ttgaagcaat taaaacatcc 420
aaatctgtgt aacctcatcg aggtgttcag gagaagaagg aaaatgcaat tagtttttga 480
atctgtgat catacacttt taaatgagct ggaagaaac ccaaatggag ttgctgatgg 540
agtgatcnaa agcgtattat ggcacaacct tcaagctott aatttctgtc atatacstaa 600
ctgtattcac agagatataa aacctgaaaa tattctaata actaagcaag gaataatcaa 660
gattttgtac tccgggtttg caaaaatctt gattccagga gatgctaca cegattatgt 720
agctacgaga tggaccgag ctcttgaact tcttgtggga gatactcagt atggtctctc 780
agtcatata tggctatttg gttgtgtttt tgcagagctc ctgacagggc agccactgtg 840
gcctggaaaa tcaagatggg acccaactta tctgataatc agaacaactag gaaaattaa 900
cccagacct caatcaact ttaaagtaa cgggtttttc catggcaca gtataactga 960
gccagaagac atggaactc ttgaggaana gttctcagat gttcactctg tggctctgaa 1020
cttcatgaag ggggtctga agatgaatcc agatgacaga ttaacctgtt cccaactcct 1080
ggagagctcc tactttgalt cttttcaaga gycctaaatt aaaagaaaag cacytaatga 1140
aggaagaaac agaagacgcc aacagaatca actgttgcct ctcataccag gaagccaact 1200
ctccccaca cctgatggaa gaaaacaagt cctccagtta aaatttgatc acctccaaa 1260
catttaggaa aatgttcttt caagtcaaaa gtaatttaat atgtacaact ttgtacaag 1320
tgagatagga atctcagtg ttccaatgc aaatgagcca tatgaaaatt aagatgcoct 1380
ctagaattgt ttgtgctctg atcattgctg atcccttccc ccatgctttt acatgccaac 1440
tttatctttt agaatattt ctttaaatgt taaaagcct aaaactgcaac atatggaaga 1500
gacattttca atttcabcg agcagccctc cccagggcta tctatatgga gaatttvtga 1560
gcttataact ggatttatga aaaagattta catgtgtcat ctgtcttcag ctgaccacat 1620
aatttcttaa agcaatatca aatagcctgc ctcaactgtt gtgtaagaaa tgacatatgt 1680
tcctgcatgt gtaattcata cttattgtaa ccaggctctg tgagtattgc tggatatcta 1740
tactgagtaa atatggtgta gaaaggaac tttgaagggc tgcagattcg 1790

<210> 31
<211> 4132
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7474637CBI

<400> 31
ccccgactg tcttgggtgg agaggggact tttattcagc tggaaaccgc cggcagggcc 60
caagtgcttc tggagagatt cggggttcag gaggtggcgg gtgcacccaa ggggtctggg 120
aggaagctcc aggttccat tcttcccag ggtatcggct tgcccctgct cgcgggggta 180

WO 02/08399

PCT/US01/23092

```

gtctagggca acggaagatg gggggcggg cggggcagg ggttcgggg tccgetggg 240
cagagccac cegtgacca actccggcc cccggcggg cgytgctgtg tccccgagg 300
agtcggagag gatggcaggg gccggaggcc agcaccacc tccggcgcc gctggagag 360
cggccggcgg agccggcgg cgggtcaact cggccgtgc ctggcgggg ccgggagag 420
attcgtctga cagcgaagcg gagcaagagg gacccagaa actgatccg caagtgcta 480
cctcggggca gatccggacc aagcagcta ttaagaggg acagctattg aagcaaacca 540
gttctttcca aaggtggaaa aagcgatact tcaaaactcg aggcggcacc ctttactatg 600
caaggaactc aaagtctctg atatttgatg aagttgacct ctccagatgt agtgtagctg 660
aagcaagcac gaaaaatgct aacaacagct tcacgatcat cactccaatc agaagctca 720
tgcgtgtgct tgagaacaga aaggagatgg aggatggat cagctcaactg aagttctgtac 780
agaccagaga accctacgag gtggccagtg ttaatgtgga acatttctca gggatgcaca 840
actggtacgc ctgcccccac gcccgaccna cctctgttaa cgttgccaga gagagtcttt 900
ctggagtcac ctcccattgg ctgtcctcgg aagtggttaa attcaaggct cacaaaagat 960
gtgcagtga agcaacaat aactgtaat ggactaccct ggctccatc gggaggaca 1020
tlatagaaga tgaagatggc gtccgagtc ctaccagtg gcttgagggc aacctgctg 1080
taagtgcaca ggtgtgctgc tgcgacaaa catgtggcag tgttctccg ctacagagat 1140
ggaaatgct ctggtgtaag caaatgttac acactgctg caagattta taccatcaa 1200
tatgtccact tggcaaatgt aaagtatca tcaatctcc aattgcaact aacagcaccg 1260
attcagatgc tttctgtaga gcaaacatct cgttctgtg tagtctcta ttgtttttg 1320
tcaattctaa gagtggagat aatcagggag taaagtctc cgtctgctt aaacagtgc 1380
caagaccgct ccaggtgttt gatttaatga atgaggtcc tcaattaggt ttaagattat 1440
ttcagaagtt tgacaatttc oggattctg tatgtgggg cgaaggaaat gtaggttggg 1500
ttttgtcaga aatcgataag ctcaacttga ataaacagtg tcaagctgga gtgttgcct 1560
tgggtacag aaatgacct gcccgagttc ttggtcgggg aggttcaatb gacgatgaca 1620
cccagcttc tcagatccta gagaactgg aacgagccag taccaaaatg ttggacaggt 1680
gggtataat gacatagaa ctcaaatgg caccnaaage tccctactc ccaggacctc 1740
cagaagctc tgaagaattt tatatgaca tttatgaaga ctcaagtgca acgatetta 1800
caaaaactc caattctgat gaacatgag tggtaatc ttctgccaag acgatattg 1860
aaactgtaa ggaactcgtt gccaaagtat aaagacgta tgacaaaac ttggaaaatg 1920
ccgttctagc tyatgcctg gccagtaaat gttcagtoct aaacgagag ctgaaacac 1980
tgcctcaggc tttgcacaca gattccagg ctgcgctgt tctccctggc ctgaccctc 2040
tcaattgtga agaagatgct gtggaatcgt ccagtgaaga gtccctggtt gaaagcaaa 2100
agcagcttgg ggatgacgct acaaaactct cctcccagaa agccgtcaaa ccaagggaaa 2160
tcaatgtcgg ggcnaatagt ttaagaaag cagtgaggca agtcaattgag gaagccggaa 2220
aagttatgga tgaccggaca gttccacct gtgaaccagc taatcagttc tctgattatg 2280
acagcacaga sacagatgaa tctaaggagg aagctaaaga tgatggtgcc aaagaatcaa 2340
taactgttaa aactgcacct cgttctccag atgcccgggc aagttatggc cactcccaa 2400
ctgattctgt cctggttcca gctgtggcag ccagcaaaaga aaacctccct gtgctcaata 2460
ccagatcaat ctgccagggt ttaagagcag gaotggctgc ctcaattgct gggagttcga 2520
ttatcaaaa aatgttactg gcaaacattg atccttttgg tgcacgccc tttattgacc 2580
ctgatctaga ttcctgatgat ggatattcag aaaaattgtt catgaacaat tactttggga 2640
ttgpataga tgcaaaaatt tcaattagaat ttaataataa aagagaggag cacctgaaa 2700
aatgcaggag ccgaactaaa aactgatgt ggtatggagt ccttggaaac cgggagttat 2760
tacagagatc gtacaagaat ttagaacaag ggttcaact tgagtgatgat gggcagtaba 2820
ttcctcttc cagcttgcaa ggcataggcg tttgaaatc tccagctat gctggaggca 2880
ctaactttg ggttgaact aaagagatg atatatttgc tgcaccatcc tttgatgaca 2940
agactcggca agttgtagca atatttgata gcatgcaaat ggcagtttca aggtcaatta 3000
aactgcagca tcatgaata gccagtgcc gtacagtgaa aatcaatata tttggtgacg 3060
aaggagtccc agtcaagtg gatgtgagc cgtgggttca gccctcaggg attatcaaaa 3120
ttgtgcacaa aaacagagca caaatgtcaa caagggacag agcctttgag agcaactga 3180
aatcttggga agataagcag aagtgtgatt ctggtaaacc agtctctcga acccaattgt 3240
acatcaata ccacctgac ttggcaacag aagaggtgkc gcagatgca ctatgctccc 3300
aggctcaga ggaactcatt actagabat gtgacgagc caccaattac tgtcttttg 3360
agcaagaact ggccatgct gtgaatgct gctccatgc cctgaataaa gccaaacca 3420
ggtgcccga gactcttaca agagacact ccactgaaat agccatcaat gtgaaggcg 3480
tgtataatga aacagaactc ttgctagttg gcagggttcc tttgagctg gaatgccc 3540

```

WO 02/08399

PCT/US01/23092

```

atgaagagcg agtatccaat gccttacact ctgtggaggt gganitacag aaactgaacg 3600
agattcccttg gctttattat abottacacc caaatgaaga tgaggaacct cctatggatt 3660
gcaccaaaag gaacaacaga agcaaccytaa ttogaatagt gccaaagttt aaaaagppaa 3720
aggttcagaa gcagaagaca agttcacagc ctggatctgg ggataccgaa agtgggtcat 3780
gtgaagcgaa ttctccaggg aattaagagag cttggaagga gcactccaca gtcggaggtg 3840
taatatatt ggtgctattc cttggaagag aagttattgc cacttaatac aaagtccctg 3900
gaagcaagtg gctgtttctt tagttttctg catagataag taagcaccac tgaagcacc 3960
ctgtggcttg atattttgct gtgggtgaaa ttttgatttg aggtattaga aaatattttt 4020
gtgcccgaaca atacattcca cgaagccatt ttctttttgt gcaaacctga catgttcaaa 4080
tatattcaaca atggtataaa ggttaggagga atctgagagc attgcattgt ct 4132

```

```

<210> 32
<211> 1137
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Incyte ID No: 7170260CB1

```

```

<400> 32
atggaggact ttctgetetc caatgggtac cagctgggca agaccattgg ggaaggagcc 60
tactcaaaag tcaaaagaagc attttccaaa aaacaccaaa gaaaagtggc aaltaaagt 120
atagacaaga tgggaggggcc agaagagttt atccagagat tctctcctcg ggagctccaa 180
atcgtccgta cctgggacca caagaacatc atccaggtgt atgagatgct ggagtctgcc 240
gacggaaaaa tctgctgggt gatggagctc gctgaggagc gggatgtctt tgactgctg 300
ctgaatgggg ggcactgccc tgaagccgg gccaaagccc tcttcctca gatggttgag 360
gccatccgct actgocattg ctgtgtgtgt gcccaaggg accccaalg tgagaagccc 420
ttgttgccgg gcttcaacct gaagctgact gactttggct ttgccaaggt gttgcccag 480
tcacaccggg agctgagcca gaacctctgc ggcaglacag cctatgctgc ccccgaggtg 540
ctgcagggca ttccccacga tagcaaaaaa ggtgatgctt ggagcatggg tgtgtctctg 600
tatgtcatgc tctgtgccag cctacctttt gacgacacag acatcccaca gatgtctggt 660
cagcagcaga agggggtgtc cttcccactc catctgagca tctggccgca ttgccaggac 720
ctgctcaaga ggtcctctgga acccgatag atcctccggc ctccaattga agaagttagt 780
tggcatccat ggttagcaag cacttgataa aagcaatggc aagtgtcttc caataaagta 840
gggggagaaa gcaaaaccaa aaaccgctt ctaaaatggt gatataat ttacgcttta 900
agtttaacta tcttaaaact tacctacatc taaccagccc ttactactac tctttctctt 960
tagagatctt catgaaatca aagggcctca ttcagacttc cttttttttt ttaagatct 1020
tgctctgtgc cccaggtggt aatgcagtag cagcattcca gttcactgca actctgcttc 1080
ccaggttcaa gcgatctctc tgcctcagcc tcccagtag ctggctttcc agcacac 1132

```

```

<210> 33
<211> 3365
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Incyte ID No: 1797506CB1

```

```

<400> 33
atgagaaggc cggggatcgg cgaggactcc aggcctgggt tgcaggccca gccaggggcy 60
gacctctctc cgggtcgggc ggggacagag cgtctccttg gaggaccaca gggactggc 120
cagccgtgca gctgcccagc cgtatagcgc agtgcggtca gggggtcgag gccgtggccc 180
cgtctgggcy tccagctcca gttccggcgc ctgctgctcg ggaagctgag tccacaggt 240
catactctca gcccagagaa cctctgctgy gtgtccact tggatggaag totccagca 300

```

WO 02/08399

PCT/US01/23092

```

ctaagcaagc agacagggga cctgaagtgg actctgaggg atgatccgt catcgaagga 360
coaagttaag teacagaaat ggcctttctc tctgaccag cagatggcag cctgtacatc 420
ttggggacc aaaaacaaca gsgattaatg aaactgccat tcaacctccc tgagctggtt 480
catgacctcc cctgcccag ccttgatggg gtcttctaca caggccggaa gcaggatgcc 540
tgglltgggg tggacctga gtcaggggag acccagatga cactgaccac agaggatccc 600
tcacccccc gccctacat tggcgaaca cagtatacgg tcacctgca tgaccaaga 660
gccccagccc tgcgctggaa caccacctac cgcgctact cagccccc catggatggc 720
tcacctggga aatacatgag ccacctggcg tctgogggga tgggctgct gctcaactgt 780
gaccagga aacagggcgt gctgtggaca caggacctgg gcgtgctgt gatggcgtc 840
tacacctggc accaggacgy cctgocccag ctgocccatc tcacctggc tcgagacct 900
ctgcattcc tegocctccg ctggggccac atccgactgc ctgctcagg cccccggac 960
acagcaccct tcttctctac cttggacacc cagctgctaa tgacctgta tgggggaag 1020
gatgaaactg gcttctatgt ctctaaagca ctggtccaca caggagtggc cctgggtcct 1080
cgtggactga ccttgcccc cgcagatggc cccaccacag atgaggtgac actccaagtc 1140
tcaggagagc gagagggtc acccagcact gctgttagat accctcagg cagtgtggcc 1200
ctcccaagcc agtggctgct cattggacac cagagctac cccagctct gcacaccacc 1260
atgctgaggg teatcccac cctggggagt ggaactgcag agacaagacc tccagagaat 1320
accagggccc cagccttctt cttggagcta ttgagctga gccagagaa actttggac 1380
tcacagctgc atccagaaga aaaaactcca gactcttact tggggctggg accccaagc 1440
ctgctggcag ctagcctcac tgcgtcctc ctggggaggt ggattctct tgtgatgag 1500
cagcagcagg agacccccct ggcaactgca gactttgctc acatctcca ggatgccag 1560
teectgcact cgggggcccag ccggagagc cagaagaggg ttcagatcc ctcaactgag 1620
tcaccaacct cctctcccc agctgagcaa ctcaactgag tggggaagat ttecttaat 1680
cccaagagc tgetgggccc cggggcaggc gggactttcg tttcaaggg acagtttgag 1740
ggacggcagc tggctgtcaa cggctcctc cgcagctgtc ttgctcgtt tggcgggaa 1800
gttcaactgc tgcaggatc tgacggacc ccaactgctc tccgctact ctgcaaccag 1860
cggggacccc agttccacta cattyccctg gagctctgcc gggctcctt gcaggaglac 1920
gtagaanaac cggacctgga tccgggggt ctggagccc aggtcgtgct gcagcagctg 1980
atgtctggcc tggcccacct gcaacttta cacatagtc accgggacct gaagccagga 2040
aatattctca tcaccgggccc tgacagccag gycctgggca gagtgggtgt ctcaacttc 2100
ggcctctgca agaagctgcc tgcctggccc tctagcttca gctccactc cggcatccc 2160
ggcagggaa gctggatggc gcccgactt ctgagctcc tggccaccaga cagtccctacc 2220
agcgtctgtg acatcttctc tgcagctgct gtttctact acgtcttcc tgggtggcagc 2280
caccctttg gagacagctc ttatggccag gcaaacatcc tcacaggggc tccctgtctg 2340
gctcactgg aggaagaggt ccacgacaag gtggttgccc gggacctggt tggagccatg 2400
ttgagccacc tgcgcagccc acgccccctc gcccccagc tgetggcca ccccttctt 2460
tggagcagag ccaagcaact ccagttcttc caggacgcca gtactggct ggaagaggag 2520
tcogagcagg agccccgtt gaggccactg gaggcgggag gctgcgagc ggtccgggac 2580
aactggcagc agcacatctc catgcccgtg cagacagatc tgagaaagt ccggtcctat 2640
aaggggacat cagtgcgaga cctgctcctg gctgtgagga acaagaagca ccactacag 2700
gagctcccag ttgaggtgag acaggcactc ggcacagctc ctgaggtct cgtccagtac 2760
ttcacaacc gcttcccag gctgctcctc cacacgcacc gagccatgag gagctgccc 2820
ctgagagccc tcttctgccc ctactaccgc ccagactcag agggcaggag gccatgccct 2880
ggggccacag ggaggtgagg tgggctggat gccacacaga tggctcctg gctggctcac 2940
tgaagagctg agcctgtggc tggcctcaga atcaggctgg gtgactggtc tcacactgt 3000
aatcccagca ttttgggagg ctgagtgaga ggaactctg agctcaggag ttcgagacca 3060
gctggccaa catgcaaca cccattctc acaaaaaatt tgtaaaatta gccagcagc 3120
gtggcgcagc cctgtagtc cagctgcttg ggagctgag gtcggagaat cacttgagcc 3180
caggagttcg aggtgagct gagccagat catgccactg cactccagcc tggccacag 3240
agagacactg tcaccacctc tccccacaa gactggcaga ggcctggcag cctgggctg 3300
atgaagcaga gatgttctct ggaatcccag tctggcaca ctgtaaggaa atacaacgaa 3360
gaggt 3365

```

```

<210> 34
<211> 2049
<212> DNA

```

WO 02/08399

PCT/US01/23092

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 1851973CB1

<400> 34

```

gcgttctttc cggggaagt agttgacatt tacaaggagc agcgcccca aaggtcttta 60
gctgtttttt aaggggagaa cagcctttac cctctttgga cttttcttcc gttttttttt 120
tttttgagaa cggagtttcg ttctttcgcc caggctggcg tacagtggcg cgattctcggc 180
tcactgcaac ctctgctccc cgggttcaag cgattctcct gcctcagcct cggagtagc 240
tgggattaca ggtgcccgcc accagccccc gctgatttcc tcttaagact ttctacagct 300
tccttatgaa atcttctgac tgggecttga gcaataaggc ctcttgctac aatttagtgc 360
tctttctctc acactaaatc gaaaactctc cctgttggtc ctgatctgtt tcagtcaggc 420
aaattacatc ctgggaaaac gtcagatgac aggggagccc actcgtctcc tgctcaacca 480
gtttcgacac ttctgtgtct ttctattagc tccagacctc agccctggcc ctgcctttac 540
tgtacagtca gaactggttt ctacgcctcg cgagggtggg aggtcgtgta tgggaggagg 600
accgctccc accagcctcg ttgggaagcc agggagaatc tctcaaatc ctgcgattca 660
gagtcagatc ccagctgctc ttcttctggt cggcccagaa ctggttggtc ctccctccc 720
atgaggaatg atgtcagctg ggcgcggctc gcccccagc aagagtgtaa ggctgcgaag 780
tcggggcttt ccgcaagccc cctccgtccg cgtctgcgta ggggagtgta cgaggcgagg 840
ggcgcggcgg ggggtgagct cagcgcggcg cgcggcgtgg gcggagcctc actttgaacc 900
cagttggcgg gaatggctgc tgcgggaggg gcagtgtag cggggcggct gtaggctgtc 960
cagcagtgga tcccaccgcg ggaagcaaga agggcctgg agggagcgcg gcgactgagg 1020
agggcgtgaa taggatcgcg gtgccaaac gcctccat tgaggaaatc agcaatgta 1080
agccacttag ccggggcgcc ttgggaaagc tgtatctggg gcgaaaggg gccaaatgt 1140
atcngttaa gggtgttaa aagcagaca tgatcaaca aatatgact catcagctcc 1200
aagctgagag agatgcactg gcaatagca aaagccatt cattgtocat ttgtattatt 1260
cactgcagtc tgcaaacat gtctacttgg taatggaata tcttattggg ggagatgta 1320
agctctctct acatataat ggttatttgg atgaagagat ggcctgtgaa tatattctg 1380
aagttagcact gctcttagac taccttcaca gacatggaat catccacagg gacttgaaac 1440
cggacaatat gcttatttct aatgagggtc atattaact gacggatttt ggcctttcaa 1500
aagttacttt gaatgagatc ataatatga tggatactct tacaacacca tcaatggcaa 1560
aacctagaca agattattca agaaccocag gacaagtgtt atcgtttatc agctcgttgg 1620
gatttaaac accaattgca gaaaaaaatc aagaccctgc aaacatcctt tcagcctgtc 1680
tgtttgaaac atcacagctt tctcaaggac tctgatgccc tatgtctgta gatcaaaagg 1740
acactagcgc ttattctagc aaattactaa aatcatgtct tgaaacagtt gcctcaaac 1800
caggaatgcc tgtgaagtgt ctaacttcta atttactcca gtctaggaaa aggctggcca 1860
catccagtc cagtgtcaa tcccacact tcatatccag tgtgaaatca gaatgccaca 1920
gcagtcccaa atgggaaaaa gattgccagg tttgagggac atttacttta atgaaaatca 1980
attatgtatg tcaaatgaat gtgagaata ttataccttt tcatataaat tccataaaga 2040
aatgaaagg 2049

```

<210> 35

<211> 2962

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 7474604CB1

<400> 35

```

accactgtg cccactgatt atcagcactc ttacttcca ccagcgttcc tgggtgtcca 60
cctctctcgg ccgcccggga aacatgacg aaagcgagg agcagcagcc tctgagttg 120
caaaaagcct tacagcagtg cgaactggtc caaaacatga tagacttgag catctccaac 180

```

WO 02/08399

PCT/US01/23092

```

ctggaagggc ttaggacca atgtgtacc tccaagacc tcaacaaaa agaatccgg 240
acctggaga gcaagctygt gaagtaattc agccggcagc tgtctcgaa aaagaaggta 300
gccttgagag agcgcaacgc ggagctggac gycttcccc agctacggca ctggttcga 360
atcgtcgtat tgcgcaagga ggtcctggag gaaatctccc ccggccagct gagcctggag 420
gaactcttgg agatgacgga tgaacagggt tgcgagacty tggagaaata cggagccaac 480
cgggaggagt gtgccgcct caacgcctcc ctctcctgcc tcaggaaatg ccacatgta 540
ggaggcaacc ttccaacaa agactggacc atccagtggc ccaccacaga gacggggaag 600
gagaacaatc ccgtgtgccc ccggagccc acccgtgga tcgcaccca tctctcccag 660
agccccagg tcccgtccaa gtgcgtccag cactattgtc acaccagccc cactcccgg 720
gccccgtgt acaccacgt ggacagcctt accgtggagc cctaccgggg ctgtgcccgc 780
ccccgccac tggagtccgg ccaccgttcc ctgccccat cggcccggca gcggcaocgc 840
gtccgcacc cgcgcgcac ccccaacatc gtaccacgc tgaccocgc gggcaocgc 900
cccagagga agaagaacaa gctgaagccc ccggggacc caccgccctc ctcccgaaa 960
ctgatacaat tgatcccggg atccaccgct ctgcatcgga gcaaatccca cagattccag 1020
ctggggacc cggggagca ggcaccacgc cccaaagca agaagaagag caaaccttg 1080
aacctcaaga tccacagcag cgttagcagc tgcgagaaca tcccctctca gcagcctcc 1140
ccgctcgtgt ccgagcctcc cctccgctcc ttcttttgg gacacgcacc ttctctgct 1200
tcaaccctcc ctgttcaaac tgagcccaac ttctctgcaa acacactgct agtgccacgc 1260
tggctcccgc agatccctcg cagagatctt ggcaactcca tcaagcacag gttttccacc 1320
aagtaactga tctctcagac gtgcacagtc tgtgggaag ggaatcttt tggcctcaag 1380
tgtaaaaact gcaagttaa gtgccacaac aaatgcacca aagaagcccc accctgcat 1440
ctctgatac tccaccgagg agatccagca aggttagtcc ggacagagtc cgttccgtgt 1500
gacatcaaca accctctacg gaagccacct cgtattccag accctgacat cagtcaagc 1560
ctccccaaa ccaaaaaaat caacaaggac cacatcccty tccctacca gccagatcc 1620
agcagcaacc cctcctccac gaagtctcc acccctcct cgcagcacc cccccctc 1680
cctagtcca cgcgccttc tcccctaac ccttcccac agtgcaacgc gcagcagaag 1740
aactcaacc tgcagcctc ccaactctac aaatacaagc agcagttcat ctcccagat 1800
gtgtgcccgc tgcggagac gcgcaccgc gcgcccagc tcatctgca tccgtgacc 1860
tcgaatccaa tcttggaggg aatccatta ctcaaaty aagtggacc aactcggag 1920
aatgaagagg tccatgatga ggccgaagag tcagaggatg acttcagga gatgaactg 1980
tccctcctc cggcccggag cttcccacgc aaggccagcc agaccagcat ctccctccg 2040
gagtgggaca tccccttga gcagctggag atcggcagc tcattggaaa gggccgctt 2100
gggcaagtgt accacggccg ctggcagtc gaggtygcca tccggctgat tgacattgag 2160
aggacaaagc agaccagct caagccttc aagcgggagc tgatggcta caggcagaca 2220
cggcatgaga acgtgtgtct ttctatgggt gctgcatga gcccgctca cctggccatc 2280
atcaccagcc tctgtaaggc acggagctc tattccgttg tgagggatgc caaaatcgtt 2340
ttgatgtca acaaaaccag gcagattgct caagaattg tgaaggcat gggctacctc 2400
cagccaagg gaatcctaca caaggacctc aagtcaaga agtcttcta tgacaacgc 2460
aaagtgtca tccagcactc tggactcttc agcatttctg ggtgctgca ggtggcagg 2520
cgggaggaca aactgycat ccagaatggc tggctatgcc accctggcacc agagatcacc 2580
cgcagcgtgt ccccgcacac agaggaggat aagctcccc tctccaagca ctctgacgc 2640
tttgccttg gcacaactc gtatgaactc cagcccaggg aatggccttt caagaccaca 2700
ccagcagagg caataactc gcaaatggc acaggcatga aaccacactc cagccagatt 2760
ggcatgggaa aagaactctc ggaattctt ctcttctgct gggccttga acaagaagag 2820
agacctacct taccagact catggacatg ctggagaaac tgcacaagc aaacgtgc 2880
ctgtctacc ctggacattt ctggaagtct gcagagctgt gacctttga catcgggagc 2940
ggcccagct gctgggctc cc 2962

```

- <210> 36
- <211> 3112
- <212> DNA
- <213> Homo sapiens

- <220>
- <221> misc_feature
- <223> Incey ID No: 7474721CB1

WO 02/08399

PCT/US01/23092

```

<400> 36
gggggcatg ctcagcggg ctaggctggc gggccttgg ccgcccggg actgacagct 60
cggctcggc accatggaga cctggccgg tccacaccg ctgcccctc tccctctgcy 120
gatgcagctc tgtctcggc tgcttttgg accctgggg cctgggacc cggaggaagt 180
tatcctcctg gattccaaag cctcccagg cgagctggg tggactgca tgcacaagta 240
tgggtggggg gagatcagc gctggatga acacgacct cccatccga cgtaccaagt 300
gtgcaatgtg ctggagccca eccaggacaa ctggctgcag actggctgga taagccgtgg 360
ccgcccggc gcatctctg tggaaactgca gttcacactc cgtgactgca gcagcatccc 420
tggcggccg ggtacctgca aggagacct caacgtctac taccctgaaa ctgagccga 480
cctggcccg ggccttccc gcctaggcct cagccggccc cgcataatc acacgatcgc 540
ggcggacgg agcttcacg agggcgacct gggctgagcg aagatgaag tgaacacaga 600
ggctggcgg atcggaccgc tccggccgg gggtttccac ctggccttc aggacgtgg 660
cgcctggctg gcgcttgtct cggctggcgt ctactacaag cagtgcggc ccaccgtgg 720
ggcctggcc acgttcccag ccacggcagc cgagagcgc ttctccacac tggctggaag 780
ggcggaaag tgcctggcgc actcggaaag ggagcctggc agccccccac gcactgactg 840
cggccggac ggcagtggt tggctcctg gggccctgc agctgcagc cgggatcca 900
ggagctggg gacatctgg aagcctgtcc cccagggtt tacaaggtg ccccgggcg 960
aagggtctg tcaccgtgc cagagcacag ccggccctg gaaaagcct ccacctctg 1020
cgtgtgccag gacagctatg cggcctacc caccgaccg cctcggctt cctgcaccg 1080
tggcccgccg tccggccgc gggactgca gtcacagct agccctgcg cgtctgtct 1140
ggcactgca tggctggcgc cggccgactc gggagcggc tccgactca cctactcgt 1200
gctgtccctg cgtcggcgc ggcggccccc gggcggccc tgcgagcgt gggggccgc 1260
cgtggccttc ctaccgccc aggcagggct cggggagcga gcccccacc tgcctgact 1320
ggcggccggg ggcgctaca ccgtcggct ggcctgtct aagggcgtc cggcccggc 1380
ggcggccctg gttccggttg cgcctgttc aattaacct ggtacggtg gccctgttc 1440
tgttccggg gttatccgg accgctgga accccagagc gttccctgt cgtggcggg 1500
ggccatccc gccggagccc ctggggcaca tgacagggag tacgagatc gatactaca 1560
gaagtgacg agtgagcaga ctactccat gttgaagaca gggggccca cagtcaact 1620
caccaaactg aagccggcta cccgctact cttcagatc cggggcgtt cccggggcc 1680
atcctgggg gccccaggt taaaccocag cattgaagt cagaccctg gggagcctc 1740
ctcagggcc agggaccaga gcccgccat tctcgtacc gtagtgacca tctcggcct 1800
cctcgtcctg ggtccgtga tgagtgctt ggcatttgg aggagccct gcagctatg 1860
caaaaggaga ggggatgcc atgatgaaga ggagctgat ttccactca aagtcccaac 1920
acgtgcaca ttccgtggacc cccagagctg tggggacctg ctgagcgtg tgcactgtt 1980
cggcaaggaa ctggatgca aaagcgtcac gctggagag agcctggag gaggcgggt 2040
tggggagctg tgcctggct gcttgcagct ccccgctgc caggagctc tctagcct 2100
gcacatgct agggacagc cctccactc acagagctc ggcctcctg ccgagccct 2160
cagcctggc cagttgacc atagccacat cgtcggcct gaggcgttg ttaccgagg 2220
aagcacctg atgattgca ccgagatcat gagccatgg gccctggag ccttccctag 2280
ggcgcacag gggcagctgg tggctgggca actgatggg ttgctgctg ggtggcatc 2340
agccatgag tatctctag agatgggcta cgttcaccg ggcctggcag ctcgccatg 2400
gctggtcagc agcgaactg tctgcaagat ctctggctc gggcggggcc cccgggaccg 2460
atcagagct gctacacca ctatgagtg ccggagccca cgcctatgg ccctcccga 2520
gacactcag tttggccact tcaactctgc cagtacgtg tggagctcg gcactatcat 2580
gtgggggtg atggccttg gggagcggcc ttactgggac atgtctggcc aagcgtgat 2640
caaggctgtg gaggatgct tccggctgcc acccccagg aactgtccta acctctgca 2700
ccgactaat ctcgactgct ggcagaagg cccagtgag cggccaggt tctccagat 2760
ccacagatc ctgagcaaga tggctcagga cccagagccc cccaagtgt cctcactac 2820
ctgtcccag cctcccacc cactagcggc cgttgcctc tccacctcc cctcctttg 2880
ctctgtggc cgtggctgg aggcctcggc cctgtccgc tacaaggaca gcttcgggc 2940
tctgtgctat gggagcctg aggcctggc cagatgact gcccaaggg acctgtgtg 3000
cctaggcctc tctttgctg aacatcgaga ggcctcctc agcgggata cggcctgca 3060
ggcagcagtg ctccagctc agggccagg ggtcaggtg ttagtggacc cc 3112

```

<210> 37
<211> 3650

WO 02/08399

PCT/US01/23092

<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7478815CB1

```

<400> 37
caaacctcca gagtcgtagg agtgaacct gcacaggaat ctctgcccac ctcaagagaa 60
accaaacttg gggaaaaatg ttgcggtcca ctctgatgca ttttacttca gcaagctgaa 120
ggaggaccag atcaagaagg tggacaggtt cctgtatcac atgcccctct ccatgacac 180
ccttttggac atcatgaggc ggttccgggc tgagatggag aaggccctgg caaaggacac 240
caaccccacg gctgcagtga agatgttgcc caccttcgtc agggccattc ccatggttc 300
cgaaaatggg gagttccttt ccctggatct cggagggtcc aagttccgag tctgaaggt 360
gcaagtgcct gaagagggga agcagacagt gcagatggag agtcagtctc acccaagccc 420
caatgaatc atcccgggga acggcacaga gctgtttgaa tatgtagctg actgtctggc 480
agatttcacg aagaccaag atttaagca taagaatg cccttggcc taacttttc 540
ttccccctgt cgacagacta aactggaaga ggtgttccca ctttcgtgga caaaaaagt 600
taaggcacga gtagttcagg acacggatgt ggtgagccgt ctgaccaaaag ccatgagaag 660
acacaggacg atggacgtgg acatcctggc cctggtcaat gacaccgtgg gacatgaa 720
gacctgtgcc tatgacgacc cctactgcga agttggtgtc atcatcgaa ctggcaccaa 780
tgcgtgttac atggaggaca tgacaacat tgacctggtg gaggcgagc agggcaggt 840
gtgacacac acagagtggg gggccttcgg ggcagacggg gccctggagg acattcgac 900
tgagtctgac agggagctgg acctggcttc tctcaacca gaaagcaac tgttcgagaa 960
gatgacaggt ggcctgtacc tgggggagct tgcaggctt atcttctgga agatggccaa 1020
ggctgacctc ctgtttggtg gtgagaatc ttctgctctc cacactaag gcaagatgca 1080
aacacggcac gtggctgcca tggagaagta taaagaaggc ctgtccta caagagat 1140
cctggtggac ctgggtctgg aacctgtgga ggtgactgc atgcccctcc agcatgtctg 1200
tacctcgtc tccctccgct cggccaatct ctgtgcagca gctctggcgg ccactctgac 1260
acgctccgg gagaacaaga agtggaaag gctccggacc acagtggtga tggacygac 1320
cctctcaaa atacaccctc agtaccaaa acgctgacc aagtggtga gaaactggt 1380
cccagctgt gatgtccgct tccctctgtc agagatggc agcaccagg gggccctat 1440
ggtgaccgag gtggcctccc cgtgcaggc ccagcggaa cagatcgaca ggtgtctggc 1500
ttgttccag ctgaccgag agcagctcgt ggcgtgcag gccaagatc gggctgagct 1560
ggagtatggg ctgaaagaaga agcaccagc gctggccag gtcaggatgc tgcaccacta 1620
cgtctgccc gtcgggacg gcacagagaa aggaagtct ctgccctgg atcttgggg 1680
aaaccactc cgggtcctcc tggtgagat cagaagtga cggaggtcag tgcgaatga 1740
caacaagatc ttgcccctcc cctggagat catgcaggc actggtgagg agctcttga 1800
tcacattgtc cagtgcctcc cgcactctc ggactacat ggcctcaagg ggcctcctc 1860
acctttggc ttccattctc catttccctg caggcagatg agcatgaca aggaacact 1920
cataggtgg accaaagtgt tcaaggccc tgactgtgaa ggggagacg tggtgacat 1980
gctcagggaa gccatcaaga ggaagaacga gtttgacctg gacatgttg cagtctgaa 2040
tgatacagtg gggaccatga tgacctgtg ctatgaagat cctaattgtg agatggcct 2100
gattgcagga acaggcagca acatgtgcta catggaggac atgaggaaca tcgagatgt 2160
ggagggggtg gaaggaaga tgtgcatcaa tacagagtg gaggatltg gagacaatg 2220
ctgcatagat gacatcggga cccgatacga cacggagtg gatgaggggt ccttgaatc 2280
tggcaagcag agatacgaga aaatgaccg tggatgttac ttgggggaga ttgtgggca 2340
gatcctgac gacctgacca agcagggtct cctctccga gggcagatt cagagcgtc 2400
ccgaccaggg ggcattctcg aaaccaagt cctgtccag atcgaagcg atcggctggc 2460
cctctccag gtcaggagga ttctgcagca gctggcctg gacagcagc gtgagacag 2520
catcgtgtg aaggaggtgt gcggagccgt gtcocggcg gcggccagc tctgctgtc 2580
tggcctggcc gctatagtg aaaaaagg agaaagacc gggctagac acctgaggt 2640
cactgtgggt gtagcagca cctgtacaa gctgacctc cactttcta gaatttga 2700
ggaaactgtg aaggaactag cccctcgatg tgatgtgaca ttcattctgt cagaagatg 2760
cagtggaaga gggcgagcac tgatcactgc tgtggccag aggttacag aggcacaga 2820
ggagaactag gaacctctg gattggacct gatgcattt ggatactgaa cagctttcc 2880

```

WO 02/08399

PCT/US01/23092

```

tctggcagat cagttggta gagaccaatg ggcaccctcc tggetgacct caccttctgg 2940
atggccgaaa gagaacccca ggttctcggg tactcttagt atctgtactt ggatttgcag 3000
tgacattaca tgacatctct atitgggtata ttggggccaa aatggggccaa cttatgaaat 3060
caaatgtctc gtcctgagag atcccccttc aacacattgt tcaagtgagg cttgagctgt 3120
caattctcta tggcttccag tcttgggtct gggggacttg gaaatatata gaactctccc 3180
atgtggctgg caggctgttt ccccattggg atgcttaagc catctcttat aggggatltg 3240
accctgtact tgtggatgaa cattggagag caagaggaac tcaacttatg aactaggggg 3300
atctcatcfa acttgccttt aactggocat gttgacttca aactgttaa gagaacaaa 3360
actttgaagt atccagcccc aggggtcaga gaggttgatt gccagggagc actgcagaa 3420
tcattgcatg cttaaagcga gttatgtcag caccctgtag gattttgttc cttatgaag 3480
gtgtgccatg tgggggggtg ctgtctgggg catctgtttt tcaatttgcg tgtggttgt 3540
gttcagggtg ttgatagttg ttttaaggat tgttaggtat aggaaatcca gtaaatat 3600
aaaaaaaaat tgattttcca ataaaaaaaa aaaaaaaaaa aaaaaaaaaa 3650

```

```

<210> 38
<211> 7789
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Incyte ID No: 7477141CBI

```

```

<400> 38
cacaccctga aagccggtcc ctggccctgc tggcccccct gcaggacgtg gacgtggggg 60
ccggggagat ggcctgtttt gagtgcctgg tggccgggccc cactgaactg gaggtggat 120
ggctgtgccc tggccgcttg ctgcagcctg cactgctcaa atgcaagatg catttcgatg 180
gccgcaaatg caagctgcta cttacatctg tacatgagga cgacagtggc gtctacacct 240
gcaagctcag caagggccaaa gatgagctga cctgcagtgc ccggtgacc gtgogggccc 300
cgttggcacc cctgttcaaa cggctgctgg aagatgtgga ggtgttgag gcccagactg 360
ccgctttcga ctgcaagatc agtggcacc ccgcccctgt tgttaactgg actcaatttg 420
gctgcccact gaggagagat gagaacttgc ggtcgggca ggacgggggt ctgcaactc 480
tgcaacttgc ccatgtgggc agcagggacc aggggctcta tgogtcaatg cctgttaaca 540
cccattggca gcccactgac tcagcccagc tttatgtaga agagcccgg acagccgccc 600
caggccccag ctgcaagctg gagaagatgc catccattcc gcaggagcca gagcagggtg 660
agctggagcg gctgtccatt cccgacttcc tgcggccact gcaggacctg gaggtgggac 720
tggccaagga ggcctgcta gattgccagg tgaccggcct gccctaccoc accatcagct 780
ggttccacaa tggccaccgc atccagagca gcgacgaccg gcgcatgaca cagtacaggg 840
atgtccatcg cttggtgttc cctgcccgtg ggcctcagca cgcgggtgtc tacaagagcg 900
tcattgccaa caagctgggc aaagctgccc gctatgccca cctgtatgtc acagatgtgg 960
tcccagccc tccagatggc gcccccaggg tggctggctg gacggggagg atgtcaaac 1020
tcacatgaaa cccccccagg agtctggaca tggccatcga cccggactcc ctgactaca 1080
cagtgacaga ccaggtgctg ggtctggacc agtggacggc actggtcaca ggcctgcggg 1140
agccaggggt ggcagccaca gggctgctga aggggtcca gcacatcttc cgggtcctca 1200
gcaccctgtc caagagcagc agcaagccct caccctcttc tgacctgtg cagctgtctg 1260
agcagggccc aaccctggag gagcccctg ccatgtgga caaacccagc atcgtgtatg 1320
tgggtggagg acagcctgcc agcgtcaccg tcaactcaa ccatgtggag gccaggtgc 1380
tctggaggag ctgcccaggg gccctctcag aggcacgggc cggtygtgac gagctgagcc 1440
agccagatga tgaccagtac tgtctcggca tctgcccggg gaggccggcg gacatggggg 1500
ccctcaactg caccgcccga aaccgtcarg gcacaagac ctgctcggte acattggagc 1560
tggcagaggg ccctcggttt gactcaca tggaggactg ggagtgggg gctggggaaa 1620
ctgctcgtct tgcgggtgtg gtcagggaaa aaccactgcc ggcacatcag tggtaacaag 1680
acaggtgctg gctgaccagc agcagccatg tgagcttctg gtacagggag aatgagtgct 1740
ccctgggtgt gctcagcacg ggggcccagg atggaggcgt ctacactgac accgcccaga 1800
acctggcggy tgaggtctcc tgcacaagcag agttggctgt gcattcagct cagacagcta 1860
tggaggtcga gggggtcggg gaggatgagg accatcgagg aaggagactc agcactttt 1920

```

WO 02/08399

PCT/US01/23092

```

atgacatcca ccaggagatc ggcaggggtg ctttctecta ctgcccgcg atagtgagc 1980
gtagctccgg cctggagttt gcggccaagt tcaltcccag ccagcccaag ccaaaagcat 2040
cagccgctcg gaggcccgg ctgctggcca ggtcccagca cgaactgttc ctctacttec 2100
atgaggcctt cagagggcgc cggggactgg tcattgtcac cgaactctgc acagaggagc 2160
tgcctggagc aatcgccagg aaaccaccag tctgtgagtc tgagatccgg gctatatatc 2220
ggcagggtgt agaggaata cactacctgc accagagcca cgtgctgca cctgatgtca 2280
agcctgagaa cctgctggtg tgggatggtg ctgcccggca gcaagcagtg cggatctgtg 2340
acttgggaaa tgcaccaggag ctgactccag gagagcccga gtactgccag tatggcacac 2400
ctgagtttgt agcaccaggag atgtccaatc agagcccagt gtctggagtc actgacatc 2460
ggcctgtggg tgttgttggc ttctctctgc tgacaggaat ctcccgttt gtggggaaa 2520
atgaccggac aacattgatg aacatccgaa actacaactt ggccttcgag gagaccacat 2580
tcctgagcct gagcaggagg gcccccggct tcctcatcaa agtgttggtg caggaccggc 2640
tgagacctac cgcagaagag accctagaac atccttggtt caaaactcag gcaagggcgg 2700
cagaggtgag cacggtatcac ctgaagctat tcctctccgg cgggaggtgg cagcctccc 2760
agatcagcta caaatgccac ctggtgctgc gccccatccc cgaactgtct cgggcccccc 2820
cagagcgggt gtgggtgacc atgccagaaa gcccaccccc cagtgggggg ctctcatcct 2880
cctcggatlc tgaagaggaa gactgggaa agctgocctc agtcccgcgc ccaactcagc 2940
ccgagttctc tggctcccgg gtgctccca cagacatbcc cactgaggat gaggccctgg 3000
ggaccaccga gactggggct gccaccacca tggactggca ggagcaggga aggtctcct 3060
ctcaggacca gggagctccc agcccagagg cctcccctcc cccagggcag gaggcccgag 3120
ctggggctag ccccagggcg gggagagctc gcagggggcag ctcgctgag agcgcctgc 3180
cccggccggg gcgcggggag ctggccgggg gctgcacaaa ggcggcgtct gtggagctgc 3240
cgagcccgcg gagccccggc ccgggagcca ccgcctggc ccggggaggc ctgggtgagg 3300
cgagtatgac ccagaggctg caggccctgc cccagcggct gctcggggga ggcctcgagg 3360
atggcaagtc cagcggcctc aggggtcccc tgcctggagag actggggggc cgtctcggg 3420
accocgggat ggcagagact gctcccagcg aggcagcgcc ccaccaccag ccccactcgc 3480
agaaaccggg cctgcaaaag agcagcagct tctcccaggg tgaggcggag ccccgggggc 3540
ggcaccggcg agcggggggc cccctcgaga tcccgtggc caggtctggg gcccgtaggc 3600
taccaggatc tccctccctg tctgcccba gggaggccca gccatccagc cctgcacggc 3660
ccagcggccc caaaaccagt accctaaagt ctgcagaacc tctgtccacc acaactagtg 3720
atgctccgca gcccccgcga ccccagctgt cccaagataa ggtcccagag cccagggcag 3780
aacccagtcg agctcccag cctgcaacc ccccaccagg cctgcaaac cttagcgtgc 3840
ccctcaacc ctatgtctag atcattcagt cctcccagct gtcaggccac gccaccggcc 3900
cctgcagagg cctgcccgg cggccttcag agcccaagcc ccacgtgct gtctttgcca 3960
gggtggcctc cccacctcgg ggagccccg agaagcggct gccctcagcc gggggctccc 4020
cgtgtctagc cgaagaagcc cagtttccca cgtgtccccc caggccaggc agcagttcca 4080
gtagcagcat cgaaaacttg gactcggagg cgtgttcca ggccaagttc aagcgcagcc 4140
gctagctccc cctgtcgttg gggctcgggc tgcctgagcc ttccgctcg gaggagcggc 4200
gcccctcccg tgggcccag gaggaggatg gcatataccg gccagcccg cgggggaccc 4260
cgtcggagct ggtgcagcgg cctgagcgtc cacgtcggct gcagagctc agggctgtcg 4320
gagaccctgg cctgtccgc cgcctctcgc tgtcactgtc ccagcggctg cggcggaccc 4380
ctcccgcgca gcgccaaccg gccctgggagg cccgcggcgg ggacggagag agctcggagg 4440
gcccggagctc ggccgggggc tcccggctgc tggcgtatgc caggcggctg agcttccccc 4500
tggagcggct gtcacccgca ttgcagcgca tggcagcag cagagactcg gggggcggct 4560
cgggcccag caaccggctg ttccgagcgc ttccgagggc cacgtccag gggagagatc 4620
tgcggcctc tggccttccg caaaaccagt tggccgcca ggcggcggcc accacgctt 4680
cgcggagtc cctgggtccc gagccagcg ccaactgggt ctctcagcc ccaggggaaa 4740
gccgaagcgg gctcggctgg ggtctctctc ggcccggaaa ggacaagggg ttatcgccac 4800
caaacctctc tgcacggctc caggaggagt tgggtcaca gtactgtgc agtgagtcag 4860
acttcccccc agtcttccca atcaaaactca aggaccaggt gctcgtggag ggggagtcag 4920
caaccctgct ctgctcccca cggcctcgcc ctgcccgcca catctcctgg atgaagaca 4980
agaaagctct gaggtaagag cctcagtgca tcatcgtgtc ctgcaaaagt gggcggcagc 5040
tgtcagcat cccccggcg gccaaagcgc acgcccgtct ctatagatc tccgcccaca 5100
acgtactggg cagcatcacc agctcctgta ccgtggctgt gcccagatc ccaggaaaagc 5160
tagctcctcc agaggtaccc cagaccatcc agaacagcgc gctgggtgctg tggagccgg 5220
gagacagcgg ggcacctgc acgtatacc tggagcggcg agtggatggg gactctgtgt 5280

```

WO 02/08399

PCT/US01/23092

```

ggcaccctgt gagctcagge atccccgaet gttactaca ogtgacccac ctgccagttg 5340
ggtgactgt gaggttccgt gtggcctgtg ccaaccgtgc tgggcagggg cocttcagca 5400
actctctetga gaaggtcttt gtccagggta ctcaagattc ttcagctgtg coactctctg 5460
cccccaaga ggcocctgtc acotcaaggc cagccagggc ccggcctcct gactctccta 5520
ctcactggc cccaccctta gctcctgtg cccccacac ccogtcagtc actgtcagcc 5580
ctcactctcc cccaccacct cctagccagg cctgtctctc gctcaaggct gtgggtccac 5640
cccccaaac cctccacga agacacaggg gctgtcagge tgcccgcca ggggagcca 5700
ccctaccag taccacgyc accccaagtg agcccaagcc tttgtctct gacactggga 5760
ccccgatccc agcotccact cctcaagggg ttaaccagt gctctctct actcctgtgt 5820
atgtggtgac ttctttgtg tctgcaccac cagccctga gcccccagcc cctgagcccc 5880
ctcctgagcc taccaggtg actgtgcaga gctcagccc gggcaaggag gtgtcagct 5940
ccccggggc cagtccccga agctctccca ggcctgaggg taccactct cgacagggtc 6000
ccccccaga accctacacc ttctggagg agaaagccag gggccgcttt ggtgtgtgc 6060
gagcgtgccc ggaagaatgcc acggggcgaa cgttcgtggc caagatcgt cctatgtctg 6120
ccgagggcaa gcgcggggtc ctgcagagat acgaggtgct ggggacctg caccacgagc 6180
ggatcatgtc cctgcacgag gctcatatca cccctcgta cctcgtctc attgctgaga 6240
gctgtggcaa ccgggaactc ctctgtgggc tcagtgacag gttccggtat tctgagatg 6300
acgtggccac ttactgggtg cagctgtcac aagcctgga ctacctcac ggcaccacg 6360
tgctccacct agacatcaag ccagacaacc tgcctgtggc ccctgacaat gccctcaaga 6420
ttgtggactt tggcagtgcc cagcctaca acccccagge ccttagyccc cttggccacc 6480
gncogggcac gctggagttc atggctccgg agatggtgaa gggganaacc atcgctctg 6540
ccogggacat ctggggagcg ggtgtgctca ctacattat gctcagtgga cgtccccgt 6600
tctatgagcc agaccocccag gaaagggagg ctggatgtt gggggccgc tttgatgct 6660
tccagctgta ccccaataca tcccagagcg ccacctctt ctgggaaag gttctctctg 6720
tactctccgt gagccggccc tccctgaggg actgcctggc ccacctggg ttgcaggagc 6780
cctacctgat gaagctggcg cgcagagcg tcacctcac caccacccg ctcaaggagt 6840
tccctgggca gacggggggc cgcggggctg aggtctccac ccggcacaag gtgtgtctgc 6900
gctcctacc tggcgccccc tagagggcag gaccacagcc aggcctcgg ctcaactgg 6960
ggttcccacc aatgccacgg gacattccag ggcaccgct gagccagggc ggcctggggc 7020
ttcggttacc accagcagca acatctggct gggctcttac ctcatagacc ttcaaggaca 7080
gagaccocag ggcctggacc tgatgccacc ccagggcaaa gccagagtg gagaccatt 7140
ggtcaggctc agcaggggtg gaacagggcag agggacaaga ggggaatgga gaagtgaga 7200
ggaaaaggaa tccaggggaca ggaaggggga ggtcttagga agttctggg ttgggggtca 7260
gtgcatctca ggggaacca aggaaggtg gcatggctgg agaggaggaa aaggagagag 7320
ccccaggtg caggcgagta ggcctggagt cagtgtgcca aagcggggc aggcacagca 7380
taccagtgca gggggccagg gctgggacat gagagaagcc agcagggcg cagagggaga 7440
agagaggact cagggtgag tgggtgggt cagctgtcag catcctcag aggaagaatg 7500
tggagagctg gaggccagca gtcactcaca ctgcctctgt cctcctgtcc agtggaatac 7560
gccctgggcy ctctgctggc ccaagatgt cccactgccc cctccatggc cttggcctt 7620
cttccattc atatttattt atttatgac ttttatgaag ttccccctc catccgatcc 7680
ctactgccc tgtgtcctg accactcctc ccagccatcc agctgtctgt ctgtctgcca 7740
caaggaaata aaaatggcaa gcagcataaa aaaaaaaaa aaaaaaaaa 7789

```

```

<210> 39
<211> 1937
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Incyte ID No: 2190612CB1

```

```

<400> 39
gtgggtggc tgcagtgag agtcccaac aaggctacgc agaagaacc ccttgactga 60
agcaatggag ggggtccag ctgtctgtg ccaggatcct cgggcagagc tggtagaacg 120
ggtggcagcc atcgatgtga ctcaactgga ggaagcagat ggtggcccag agcctactag 180

```

WO 02/08399

PCT/US01/23092

```

aaacgggtgtg gacccccac caccgggocag agctgcctct gtgatccctg gcagtacttc 240
aaagactgctc ccagcccygc ctagcctctc agccaggaaag ctttccctac aggagcggcc 300
agcaggaagc tatctggagg cgcaggctgg gccttatgcc acggggcctg ccagccacat 360
ctccccccgg gctcggcggg gcccaccat cagtcaccac cactggcca tctcagatgc 420
agaggactgc gtgcagctga accagtacaa gctgcagagt gagatggca aggtgccta 480
cgggtgtggg aggtcggcct acaacgaaag tgaagacaga cactatgcaa tgaagtcct 540
ttccaaaaag aagttactga agcagtatgg ctttccacgt cggcctcccc cgagagggtc 600
ccaggtgccc cagggaggac cagccaagca gctgctgccc ctggagcggg tgtaccagga 660
gattgcctcc ctgaagaagc tggaccacgt gaatgtggtc aaactgatcg agtccctgga 720
tgaccacagc gaggacaacc tctatttggg gtttgacctc ctgagaaagg ggcctcctat 780
ggaagtgcct tgtgacaagc ccttctcggg ggagcaagct cgcctctacc tgcgggacgt 840
catcctgggc ctgcagtaact tgcactgcca gaagatgctc cacagggaca tcaagccatc 900
caacctgctc ctgggggatg atgggcacgt gaagatgccc gactttggcg tcagcaacca 960
gitttggggg aacgacgctc agctgtccag caccggggga accccagcat tcatggcccc 1020
cgaggccatt tetgattccg gccagagctt cagtgggaag gccctggatg tatgggccac 1080
tggcgtcacg tbtactgctc ttgtctatag gaagtgcocg ttcctgacg atttcaatct 1140
ggccctccac aggaagatca agaattgagc cgtggtgttt cctgaggagc cagaatcag 1200
cgaggagctc aaggacctga tccctgaagt gttagacaag aatcccagaa cgagaattgg 1260
ggtgccagac atcaagttgc accttgggt gaccaagaac ggggaggagc cccttctctc 1320
ggaggaggag cactgcagcg tggtaggagt gaccagagag gaggttaaga actcagtcag 1380
gtcaccctcc agctggacca cgtgatcctt ggtgaagtcc atgctgagga agccttctct 1440
tggaaaccgg tttgagcccc aagcacggag ggaagacgca tccatgtctg ctccagaaaa 1500
cctaatggty aaagaaggty ttggtgagg gggcaagagc ccagagctcc ccggcgtcca 1560
ggaagacgag gctgcactct gaccctctat gcaacaccag ggcaccocgg cagcaacttc 1620
atcccggccc tccagaggcc caccctctat gcaacaccag ccccgcagg cagggggctg 1680
gggactgcag cccactctcc gccctctccc catcgtctg catgactccc acgcaagcag 1740
gtccaggagc agactggaat gtatgtcatt tggggtcttg ggggcagggc tcccacagg 1800
ccatctctct ctcttggac ctctctggcc tggaccatcc tgtggggaaa ccgggtgccc 1860
atgagcctcc agaaatgaca cccggctggt tggcatggcc tggggcagga ggcagaggca 1920
ggagaccag atggcag

```

```

<210> 40
<211> 5373
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Incyte ID No: 7477549CB1

```

```

<400> 40
atggagcggc ggctgcgcgc gctggagcag ctggcggggg gcgaggccgg cggctgccc 60
gggctcgacg gccctcctaga tctgctgctg gcgctgcacc acgagctcag cagcggcccc 120
ctacggcggg agcgacagct ggccagttcc ctgagctggg ccagccccct cgtatcaaa 180
gtgaaagaac tgcgtctgca gagagatgac tttgagatct tgaagtgat cggccgagga 240
gcccttgggg aggtcacctg ggtgaggcag agggacactg ggcagatttt tgccatgaaa 300
atgctgcaca agtgggagat gctgaagagg gctgagacag cctgcttccg ggaggagcgg 360
gatgtgctcg tgaaggggga cagccgttgg gtgaccactc tgcaactatg cttccaagac 420
gagagatacc tgtacctgtg gatggactac tatgctggtg gggacctcct gacgctgcty 480
agccgcttcg aggaccgtct cccgcocgag ctggcccagt tctacctgpc tgagatggtg 540
ctgoccatoc actcgtgca ccagctgggt tatgtocaca gggatgtcaa gccagacaac 600
gtcctgctgg atgtgaacgg gccacattgc ctggctgaet tgggntcctg cctgctctc 660
aacaccacag cactggtgga ttcactcagt gcagtaggga cggccgacta tatctccct 720
ggatccctgc agcccatgga ggaggcaag ggcactacg gccacagtg tgactggtgg 780
tcgctggag tctgyccta tgagctgctc tttggggaga cggccttcta tctgagtc 840
tbtgtgaaa cctacggcaa gatcatgaac caccaggacc acctgcagtt ccccocggac 900

```

WO 02/08399

PCT/US01/23092

```

gtgectgagc tgcagccag cgcacaagc ctgatccgc agctgtgtg tgcagcgsa 960
gagcggctag gccctggtg gctggatgac ttccggaacc atcctttctt cgaaggcgtg 1020
gactgggagc ggctggcgag cagcacagcc ccctalattc ctgagctgcg gggcccctag 1080
gacacctcca accttgatgt ggatgagcac accctcaacc atccagggac cctgccaccg 1140
ccctcccaag gggcctcttc cggccatcac ctgccattcg tgggtttcac ctacacctca 1200
ggcagtcaca gtctgagag cagctctgag gcttgggctg cctggagcg gaagctccag 1260
tgtctggagc aggagaaggt ggagctgagc aggaagcacc aagaggcctt gcaagcctcc 1320
acagaccatc gggagctgga gcactacggy aaggaggtgc agactctcg gacagggctg 1380
ccagagatgc tgagggacaa ggctctattg tccagacggy atgggcccc agctggtagc 1440
ccaggtcagg acagtgcact acggcaggag cttgaccgac ttccaccgga gctggccgag 1500
ggctcgggag ggctgcaggc tcaggagcag gagctctgca gggcccagg gcaagcaggag 1560
gagctgcttc agaggtcaca ggagggccag gagagagagg cggccacagc tagccagacc 1620
cggccctctc gctccagctt ggaggaagcc cgggctgccc agagggagct ggggcccag 1680
gtgtctctcc tgagccggca ggtgagcagc ctgcagggac agtgggagca acgcttgag 1740
gagtgctccc aggcacaagc caaccacaca gctctgaga ccaacgggat gggacccct 1800
gaggtgtggc ctcaggaggg ccaactgagg aaggaggtgg ccgcccctgc agagcagctg 1860
gagcaggccc acagccacag gccgagtggt aaggaggagg ctctgtgcca gctgcaggag 1920
gaaaaccgga ggctgagccg ggagcaggag cggctagaag cagagctggc ccaggagcag 1980
gagagcaagc agcggctgga ggtgagcgg cgggagacgg agagcaactg ggggcccag 2040
ctgcgcgaca tctctagctg ggtgaatgat gagaaggtct caagaggcta cctgcaggcc 2100
ctggccacca agatggcaga ggagctggag tcttgaggca acgtaggcac ccagacgctc 2160
ctgcccggc cactgaagat ggagcctcg ccagagctgg agctgcagtc agcctggag 2220
gccagatcc gcgccaagca ggccctcgag gagcggctga cacagtgca ggggcccag 2280
ctcagagctg agccgctct cagagagggc gcaagcagca gccagccctt gcaacaggag 2340
ctgcacatgc tgcgggagga gctgcccggc cgaggcccag tggacacca gccctaaac 2400
tcctctgatt ccttctgttc cttccggagc tcagagaagg attctgcca ggaacctgga 2460
atctcaggag aggcacaag agctggagga gagccagatc tgaggccgga gggccgagc 2520
agcctgggca tggggcgtgt gtlccccaga gcaaccactg ccaacacagc ctctacagaa 2580
ggtctctctg ctaagggatg gggcatgggg cctggggagg ccttgggtaa tggctgtccc 2640
cctcccagc cggctcaca caagctgctg ccccgagct tcccctccc gaccaggtg 2700
ctccctgca cctgctgat gctgggctg ggcggcagg gcctgggtt tgatgctgc 2760
ggctactttt gtcacacaa ctytgccca caggcccac cctgcccctt gccctctgac 2820
ctctctcga cagccctggg agtacacccc gaaacaggca caggcactgc ctatgagggc 2880
ttttctgctg tgcggggccc ctcaaggtgc cggcggggct ggcagcctg gtttctgccc 2940
ctgagtgact cagcctgct gctgtttgac gccctgacc tgaggctcag ccgcccagt 3000
ggggcctcc tgcaggtct agatctgagg gaccccagt tctggctac cctgtctg 3060
gcctctgat tttaccatgc ccaactcagg gaactgccc gcatcttag ggtgacaacc 3120
tcccagctgg cagtgccgcc caccacgtgc actgtgctg tgcctggcaga gagcagggg 3180
gagcgggaac gctggctgca ggtgctgggt gagctgcagc ggtgctgctt ggcagcggg 3240
ccaagacccc ggcctgtgta cacactcaag gaggctaac acaacggctt gccctgctg 3300
cctcacacgc tctgctgca catcctcagc caggatcgac ttgccttgg caccgaggag 3360
gggctctttg tcatccatct gcgcagcaac gacatctcc aggtggggga gtgcccggc 3420
gtgcagcagc tgaccttgag ccccagtgca ggcctgctg tctgtctgtg tggccgggc 3480
cccagcgtgc gctctcttgc cctggcggag ctggagaaca tgaggttagc aggtgccaag 3540
atcccagat ctgagggctg ccaggtgctg gcagctgaa gcatctgca ggcggcacc 3600
cgggtgctc ggtgagcct caagcgcag gctgctgct accagctggg cccggccct 3660
gggcccctgg agcggccat ccgtgagct caggcacctg ccaactgca gagcctggg 3720
ctgctggggg accgctatg tgtggggccc gccggtggct ttgcaactca ccgctgctc 3780
aacgagctg cgcctgtgca gctggggccc ggtttggtgc ctgaggagct gccaccatc 3840
cgcggggccc tgggtgagc actgggtgcc gttgagctta gcctcagcga gttctgcta 3900
ctcttaccac ctgctgcat ctacgtggat gccgcagcc gcaagtctg tggcccagag 3960
ctgtgtggc cagcagcgc catgggtgg ggtalggc cccctacct gacagtttc 4020
agcagaact ccatcgatgt gtttgcagct aggagggcag aatgggtgca gaccgtgccc 4080
ctcaagaag tgcggccct caatccagag gctccctgt tctctacgg caccgagaag 4140
gtccgctga ctaactcag gaaccagctg gcagagaag acgagttca catcccggac 4200
ctaccggca acagcggcg ccagctgttc gacccaaga gcaagcggc cttcttttc 4260

```

WO 02/08399

PCT/US01/23092

```

cgcgtgtcgg aggagcagca gaagcagcag cgcagggaga tgcgaaagga cccctttgtg 4320
cgctccaagc tcactctgcc gcctaccaac tccaaccaac tagtacaegt gggccctgac 4380
aaagggcggc cgggcgccag ggacaagtcc ccgtcccagc cctcccgac tgcacccaa 4440
caggctcccg aagagaaggg ccgagttgcc cgggctccg gccccacagc gccccacagc 4500
ttctccyagg cyttgcggcg cccagcctcc atggggcagcg aaggcctcgg tggagacgca 4560
gaccccactg gacgagtgaa gaggaaacc ttgacatccc tgtcccagca gtctgtgtcc 4620
tgccccaggg gatcgetgag ccctgcaacc tccctaatgc aggtctcaga acggcccaga 4680
agcctcccc tgtccctga attgagagc tctcttgat gccctctgtt agggcccacc 4740
ccaatcccag ggcaagaagga catgagggag caaagagctt gaggaatgcc atactccggc 4800
tggctccggg catggaatt cggactcagg gaggaccgg gctgggcaat gactgggaga 4860
cttccctggg ttcccaggac ttgggggtcc tgcctccag cctccatcct gccctadccc 4920
tctgttccca gccccagcct ttctaagcca ttgggaatag aatggcccct ttgttctgg 4980
tgtccagggg tgattgtgcc aaagctctta ttccagtgcc caagcccga gaggtctgta 5040
agagttggga tgagggatgg agagggactg ggtctctggg aacaggttgg aggtcttacc 5100
tgtggactgt ctgactccca gctgaggcca agatggggca tgtcccctc tctgcttagc 5160
gtctgggtga gaaaaacag ctgtgatcca gaagaagga agatagagaa ggagggaag 5220
gatgtaggcg aaggaggtga gagacaggat aggaggaagg aagtggagga ggaggtgsta 5280
ggaattggaa ggaggtagaa gccgtgcaga ggaagagggg agagggacga aggaggagcg 5340
atgaagaaga ggaggagac aaaaaaggg aag 5373

```

【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
31 January 2002 (31.01.2002)

PCT

(10) International Publication Number
WO 02/008399 A3(51) International Patent Classification: C12N 15/00,
15/54, C07K 16/40, C12N 15/63, 9/12, A61K 38/45,
C12Q 1/48, 1/68

(21) International Application Number: PCT/US01/23092

(22) International Filing Date: 20 July 2001 (20.07.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
50/220,038 21 July 2000 (21.07.2000) US
50/222,112 28 July 2000 (28.07.2000) US
50/222,831 4 August 2000 (04.08.2000) US
50/224,729 11 August 2000 (11.08.2000) US(71) Applicant (for all designated States except US): INCYTE
GENOMICS, INC. [US/US]; 3160 Porter Drive, Palo
Alto, CA 94304 (US).

(71) Applicant and

(72) Inventor: THORNTON, Michael [US/US]; 9 Medway
Road, Woodside, CA 94062 (US).

(72) Inventors; and

(75) Inventors/Applicants (for US only): YUE, Henry
[US/US]; 826 Lois Avenue, Sunnyvale, CA 94087 (US).
KHAN, Farrah, A. [IN/US]; 3617 Central Road #102,
Glenview, IL 60025 (US). GURURAJAN, Rajagopal
[IN/US]; 5591 Dent Avenue, San Jose, CA 95118 (US).
HAFALIA, April, J., A. [US/US]; 2227 Calle de Primavera,
Santa Clara, CA 95054 (US). CHAWLA, Narinder
K. [US/US]; 33 Union Square, #712, Union City, California
94587 (US). ARVIZU, Chandra S. [US/US];
1706 Monrocco Drive, San Jose, California 95125 (US).
RAMKUMAR, Jayalaxmi [IN/US]; 34359 Maybird
Circle, Fremont, CA 94555 (US). GANDHI, Ameena, R.
[US/US]; 837 Roble Avenue, #1, Menlo Park, CA 94025
(US). POLICKY, Jennifer, L. [US/US]; 1511 Jarvis
Court, San Jose, CA 95118 (US). BAUGHN, Mariah,
R. [US/US]; 14244 Santiago Road, San Leandro, CA
94577 (US). TRIBOULEY, Catherine, M. [US/US];
1121 Tennessee Street, #5, San Francisco, CA 94107 (US).
BANDMAN, Olga [US/US]; 366 Anna Avenue, MI,
View, CA 94043 (US). NGUYEN, Dannie, B. [US/US];
1403 Ridgewood Drive, San Jose, CA 95118 (US). LU,
Yan [CN/US]; 3885 Corina Way, Palo Alto, CA 94303
(US). BURFORD, Neil [US/US]; 105 Wildwood Circle,Durham, CT 06422 (US). LAL, Preeti [US/US]; P.O. Box
5142, Santa Clara, CA 95056 (US). DING, Li [CN/US];
3353 Alma Street #146, Palo Alto, CA 94306 (US). XAO,
Monique, G. [US/US]; 111 Frederick Court, Mountain
View, CA 94043 (US). ELLIOTT, Vicki, S. [US/US];
3770 Polton Place Way, San Jose, CA 95121 (US). RE-
CIPON, Shirley, A. [US/US]; 85 Fortuna Avenue, San
Francisco, CA 94115 (US). KEARNEY, Liam [IE/US];
50 Woodside Avenue, San Francisco, CA 94127 (US).
LU, Dyung, Aina, M. [US/US]; 233 Coy Drive, San Jose,
CA 95123 (US). GREENWALD, Sara, R. [US/US]; 21
Bucareli Drive, San Francisco, CA 94132 (US). TANG,
Y., Tom [US/US]; 4230 Ranwick Court, San Jose, CA
95118 (US). XU, Yuming [US/US]; 1739 Walnut Drive,
Mountain View, CA 94040 (US). WALSH, Roderick,
T. [IE/GB]; 8 Boundary Court, St. Lawrence Road,
Canterbury, Kent CT1 3EZ (GB). GIETZEN, Kimberly,
J. [US/US]; 691 Los Hincos Drive, San Jose, CA 95123
(US). YANG, Junming [CN/US]; 7125 Bark Lane, San
Jose, CA 95129 (US). JACKSON, Jennifer, L. [US/US];
1826 Rina Court, Santa Cruz, California 95062 (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 (GI, 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, GQ, GW, ML, MR, NE, SN, TD,
TG).Published:
with international search report(88) Date of publication of the international search report:
28 August 2003For 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 02/008399 A3

(54) Title: HUMAN KINASES

(57) Abstract: The invention provides human human kinases (PKIN) and polynucleotides which identify and encode PKIN. 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 PKIN.

【 国際調査報告 】

INTERNATIONAL SEARCH REPORT		International Application No. PCT/US 01/23092
A. CLASSIFICATION OF SUBJECT MATTER IPC 7 C12N15/00 C12N15/54 C07K16/40 C12N15/63 C12N9/12 A61K38/45 C12Q1/48 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		
Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched		
Electronic data base consulted during the international search (name of data base and, where practical, search terms used)		
C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	SHIER P AND WATT V M: "Primary structure of a putative receptor for a ligand of the insulin family" JOURNAL OF BIOLOGICAL CHEMISTRY, AMERICAN SOCIETY OF BIOLOGICAL CHEMISTS, BALTIMORE, MD, US, 'Online! vol. 264, no. 25, 5 September 1989 (1989-09-05), pages 14605-14608, XP002154780 ISSN: 0021-9258 page 14607; figure 1 -& DATABASE EMBL 'Online! "insulin receptor-related receptor" Database accession no. p14616 XP002213066 --- -/--	1-19, 21, 22, 24-45, 65
<input checked="" type="checkbox"/>	Further documents are listed in the continuation of box C.	<input checked="" type="checkbox"/> Patent family members are listed in annex.
* Special categories of cited documents : *A* document defining the general state of the art which is not considered to be of particular relevance *E* earlier document but published on or after the international filing date *L* document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified) *O* document referring to an oral disclosure, use, exhibition or other means *P* document published prior to the international filing date but later than the priority date claimed ** later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention *X* document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone *Y* document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art. *Z* document member of the same patent family		
Date of the actual completion of the international search		Date of mailing of the international search report
11 September 2002		07.03.03
Name and mailing address of the ISA European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 851 epo nl, Fax: (+31-70) 340-3015		Authorized officer Seroz, T

INTERNATIONAL SEARCH REPORT		Int'l Application No. PCT/US 01/23092
C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	<p>HAENZE J ET AL: "CLONING AND SEQUENCING OF THE COMPLETE CDNA ENCODING THE HUMAN INSULIN RECEPTOR RELATED RECEPTOR" HORMONE AND METABOLIC RESEARCH, THIEME-STRATTON, STUTTGART, DE, vol. 31, no. 2/3, 1999, pages 77-79, XP000944669 ISSN: 0018-5043 page 77, right-hand column, last paragraph -page 78, left-hand column, paragraph 1 page 78, left-hand column, line 10-13,32-34</p>	1-19,21, 22, 24-45,65
X	<p>WO 00 14212 A (ACTON SUSAN ;MILLENNIUM PHARM INC (US)) 16 March 2000 (2000-03-16) page 42, line 11 -page 43, line 20 page 56, line 14-17 page 62, line 24-29 page 63, line 28 -page 91, line 16; claims 1-26; examples 1-5</p>	1-19,21, 22, 24-45,65
X	<p>SCHULTZ S J ET AL: "IDENTIFICATION OF 21 NOVEL HUMAN PROTEIN KINASES, INCLUDING 3 MEMBERS OF A FAMILY RELATED TO THE CELL CYCLE REGULATOR NIMA OF ASPERGILLUS NIDULANS" CELL GROWTH AND DIFFERENTIATION, THE ASSOCIATION, PHILADELPHIA, PA, US, vol. 4, 1 October 1993 (1993-10-01), pages 821-830, XP000564042 ISSN: 1044-9523 the whole document</p>	1-19,21, 22, 24-45,65

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US 01/23092**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. 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. 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:

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. 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.
1-19(partially), 21(partially), 22(partially), 24-44(partially), 45(completely), 65(completely)

Remark on Protest

- The additional search fees were accompanied by the applicant's protest.
 No protest accompanied the payment of additional search fees.

International Application No. PCT/US 01 23092

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

1. Claims: 1-19 (partially), 21 (partially), 22 (partially),
24-44 (partially), 45 (completely),
65 (completely)

Human kinase comprising SEQ ID No 1 and
a polynucleotide comprising SEQ ID No 21 which encodes and
identifies said kinase. Expression vectors, host cells,
antibodies. Methods for diagnosing and treating or
preventing disorders associated with aberrant expression of
PKIN. Method for screening compounds that modulates the
activity of the kinase.

2. Claims: 1-19 (partially), 21 (partially), 22 (partially),
24-44 (partially), 46 (completely),
66 (completely)

Human kinase comprising SEQ ID No 2 and
a polynucleotide comprising SEQ ID No 22 which encodes and
identifies said kinase. Expression vectors, host cells,
antibodies. Methods for diagnosing and treating or
preventing disorders associated with aberrant expression of
PKIN. Method for screening compounds that modulates the
activity of the kinase.

3. Claims: 1-19 (partially), 21 (partially), 22 (partially),
24-44 (partially), 47 (completely),
67 (completely)

Human kinase comprising SEQ ID No 3 and
a polynucleotide comprising SEQ ID No 23 which encodes and
identifies said kinase. Expression vectors, host cells,
antibodies. Methods for diagnosing and treating or
preventing disorders associated with aberrant expression of
PKIN. Method for screening compounds that modulates the
activity of the kinase.

4. Claims: 1-19 (partially), 21 (partially), 22 (partially),
24-44 (partially), 48 (completely),
68 (completely)

Human kinase comprising SEQ ID No 4 and
a polynucleotide comprising SEQ ID No 24 which encodes and
identifies said kinase. Expression vectors, host cells,
antibodies. Methods for diagnosing and treating or
preventing disorders associated with aberrant expression of
PKIN. Method for screening compounds that modulates the
activity of the kinase.

International Application No. PCT/US 01 23092

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

5. Claims: 1-19 (partially), 21 (partially), 22 (partially),
24-44 (partially), 49 (completely),
69 (completely)

Human kinase comprising SEQ ID No 5 and a polynucleotide comprising SEQ ID No 25 which encodes and identifies said kinase. Expression vectors, host cells, antibodies. Methods for diagnosing and treating or preventing disorders associated with aberrant expression of PKIN. Method for screening compounds that modulates the activity of the kinase.

6. Claims: 1-19 (partially), 21 (partially), 22 (partially),
24-44 (partially), 50 (completely),
70 (completely)

Human kinase comprising SEQ ID No 6 and a polynucleotide comprising SEQ ID No 26 which encodes and identifies said kinase. Expression vectors, host cells, antibodies. Methods for diagnosing and treating or preventing disorders associated with aberrant expression of PKIN. Method for screening compounds that modulates the activity of the kinase.

7. Claims: 1-19 (partially), 21 (partially), 22 (partially),
24-44 (partially), 51 (completely),
71 (completely)

Human kinase comprising SEQ ID No 7 and a polynucleotide comprising SEQ ID No 27 which encodes and identifies said kinase. Expression vectors, host cells, antibodies. Methods for diagnosing and treating or preventing disorders associated with aberrant expression of PKIN. Method for screening compounds that modulates the activity of the kinase.

8. Claims: 1-19 (partially), 21 (partially), 22 (partially),
24-44 (partially), 52 (completely),
72 (completely)

Human kinase comprising SEQ ID No 8 and a polynucleotide comprising SEQ ID No 28 which encodes and identifies said kinase. Expression vectors, host cells, antibodies. Methods for diagnosing and treating or preventing disorders associated with aberrant expression of PKIN. Method for screening compounds that modulates the activity of the kinase.

International Application No. PCT/US 01 23092

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

9. Claims: 1-19 (partially), 21 (partially), 22 (partially),
24-44 (partially), 53 (completely),
73 (completely)

Human kinase comprising SEQ ID No 9 and a polynucleotide comprising SEQ ID No 29 which encodes and identifies said kinase. Expression vectors, host cells, antibodies. Methods for diagnosing and treating or preventing disorders associated with aberrant expression of PKIN. Method for screening compounds that modulates the activity of the kinase.

10. Claims: 1-19 (partially), 21 (partially), 22 (partially),
24-44 (partially), 54 (completely),
74 (completely)

Human kinase comprising SEQ ID No 10 and a polynucleotide comprising SEQ ID No 30 which encodes and identifies said kinase. Expression vectors, host cells, antibodies. Methods for diagnosing and treating or preventing disorders associated with aberrant expression of PKIN. Method for screening compounds that modulates the activity of the kinase.

11. Claims: 1-19 (partially), 21 (partially), 22 (partially),
24-44 (partially), 55 (completely),
75 (completely)

Human kinase comprising SEQ ID No 11 and a polynucleotide comprising SEQ ID No 31 which encodes and identifies said kinase. Expression vectors, host cells, antibodies. Methods for diagnosing and treating or preventing disorders associated with aberrant expression of PKIN. Method for screening compounds that modulates the activity of the kinase.

12. Claims: 1-19 (partially), 21 (partially), 22 (partially),
24-44 (partially), 56 (completely),
76 (completely)

Human kinase comprising SEQ ID No 12 and a polynucleotide comprising SEQ ID No 32 which encodes and identifies said kinase. Expression vectors, host cells, antibodies. Methods for diagnosing and treating or preventing disorders associated with aberrant expression of PKIN. Method for screening compounds that modulates the

International Application No. PCT/US 01 23092

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

activity of the kinase.

13. Claims: 1-19 (partially), 21 (partially), 22 (partially),
24-44 (partially), 57 (completely),
77 (completely)

Human kinase comprising SEQ ID No 13 and a polynucleotide comprising SEQ ID No 33 which encodes and identifies said kinase. Expression vectors, host cells, antibodies. Methods for diagnosing and treating or preventing disorders associated with aberrant expression of PKIN. Method for screening compounds that modulates the activity of the kinase.

14. Claims: 1-19 (partially), 21 (partially), 22 (partially),
24-44 (partially), 58 (completely),
78 (completely)

Human kinase comprising SEQ ID No 14 and a polynucleotide comprising SEQ ID No 34 which encodes and identifies said kinase. Expression vectors, host cells, antibodies. Methods for diagnosing and treating or preventing disorders associated with aberrant expression of PKIN. Method for screening compounds that modulates the activity of the kinase.

15. Claims: 1-19 (partially), 21 (partially), 22 (partially),
24-44 (partially), 59 (completely),
79 (completely)

Human kinase comprising SEQ ID No 15 and a polynucleotide comprising SEQ ID No 35 which encodes and identifies said kinase. Expression vectors, host cells, antibodies. Methods for diagnosing and treating or preventing disorders associated with aberrant expression of PKIN. Method for screening compounds that modulates the activity of the kinase.

16. Claims: 1-19 (partially), 21 (partially), 22 (partially),
24-44 (partially), 60 (completely),
80 (completely)

Human kinase comprising SEQ ID No 16 and a polynucleotide comprising SEQ ID No 36 which encodes and identifies said kinase. Expression vectors, host cells, antibodies. Methods for diagnosing and treating or preventing disorders associated with aberrant expression of

International Application No. PCT/US 01 23092

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

PKIN. Method for screening compounds that modulates the activity of the kinase.

17. Claims: 1-19 (partially), 21 (partially), 22 (partially), 24-44 (partially), 61 (completely), 81 (completely)

Human kinase comprising SEQ ID No 17 and a polynucleotide comprising SEQ ID No 37 which encodes and identifies said kinase. Expression vectors, host cells, antibodies. Methods for diagnosing and treating or preventing disorders associated with aberrant expression of PKIN. Method for screening compounds that modulates the activity of the kinase.

18. Claims: 1-19 (partially), 21 (partially), 22 (partially), 24-44 (partially), 62 (completely), 82 (completely)

Human kinase comprising SEQ ID No 18 and a polynucleotide comprising SEQ ID No 38 which encodes and identifies said kinase. Expression vectors, host cells, antibodies. Methods for diagnosing and treating or preventing disorders associated with aberrant expression of PKIN. Method for screening compounds that modulates the activity of the kinase.

19. Claims: 1-19 (partially), 21 (partially), 22 (partially), 24-44 (partially), 63 (completely), 83 (completely)

Human kinase comprising SEQ ID No 19 and a polynucleotide comprising SEQ ID No 39 which encodes and identifies said kinase. Expression vectors, host cells, antibodies. Methods for diagnosing and treating or preventing disorders associated with aberrant expression of PKIN. Method for screening compounds that modulates the activity of the kinase.

20. Claims: 1-19 (partially), 21 (partially), 22 (partially), 24-44 (partially), 64 (completely), 84 (completely)

Human kinase comprising SEQ ID No 20 and a polynucleotide comprising SEQ ID No 40 which encodes and identifies said kinase. Expression vectors, host cells, antibodies. Methods for diagnosing and treating or

International Application No. PCT/US 01 23092

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

preventing disorders associated with aberrant expression of PKIN. Method for screening compounds that modulates the activity of the kinase.

International Application No. PCT/US 01 23092

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

Continuation of Box I.1

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.

Although claim 18, 21, 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 composition.

Continuation of Box I.2

Claims Nos.: 20, 23

Present claims 20, 23 relate to a compound defined by reference to a desirable characteristic or property, namely agonist and antagonist. The claims cover all compounds having this characteristic or property, whereas the application does not provide support within the meaning of Article 6 PCT and disclosure within the meaning of Article 5 PCT. In the present case, the claims so lack support, and the application so lacks disclosure, that a meaningful search over the whole of the claimed scope is impossible. Consequently, no search has been carried out for those 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 SEARCH REPORT

In International Application No
PCT/US 01/23092

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 0014212 A	16-03-2000	US 6183962 B	06-02-2001
		AU 5817799 A	27-03-2000
		CA 2342311 A	16-03-2000
		EP 1112354 A	04-07-2001
		JP 2002524073 T	06-08-2002
		US 6043040 A	28-03-2000
		US 6146841 A	14-11-2000
		US 6180358 B	30-01-2001
		US 6153417 A	28-11-2000
		US 6146832 A	14-11-2000
		US 6190874 B	20-02-2001
		US 6121030 A	19-09-2000
		US 6200770 B	13-03-2001
		US 2002094559 A	18-07-2002
		US 6214597 B	10-04-2001

フロントページの続き

(51) Int.Cl. ⁷	F I	テーマコード(参考)
A 6 1 P 9/00	A 6 1 P 9/10	4 B 0 6 5
A 6 1 P 9/10	A 6 1 P 35/00	4 C 0 8 4
A 6 1 P 35/00	A 6 1 P 37/00	4 H 0 4 5
A 6 1 P 37/00	A 6 1 P 43/00	1 1 1
A 6 1 P 43/00	C 0 7 K 16/40	
C 0 7 K 16/40	C 1 2 N 1/15	
C 1 2 N 1/15	C 1 2 N 1/19	
C 1 2 N 1/19	C 1 2 N 1/21	
C 1 2 N 1/21	C 1 2 N 9/12	
C 1 2 N 5/10	C 1 2 P 21/08	
C 1 2 N 9/12	C 1 2 Q 1/02	
C 1 2 P 21/08	C 1 2 Q 1/68	A
C 1 2 Q 1/02	G 0 1 N 33/15	Z
C 1 2 Q 1/68	G 0 1 N 33/50	Z
G 0 1 N 33/15	G 0 1 N 33/53	D
G 0 1 N 33/50	G 0 1 N 33/53	M
G 0 1 N 33/53	G 0 1 N 33/566	
G 0 1 N 33/566	C 1 2 N 5/00	A
	A 6 1 K 37/02	

(31)優先権主張番号 60/224,729

(32)優先日 平成12年8月11日(2000.8.11)

(33)優先権主張国 米国(US)

(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, GQ, 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, S D, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW

(72)発明者 カーン、ファラ・エイ

アメリカ合衆国イリノイ州60025・グレンビュー・#102・セントラルロード 3617

(72)発明者 ガルラジャン、ラジャゴバル

アメリカ合衆国カリフォルニア州95118・サンノゼ・デントアベニュー 5591

(72)発明者 ハファリア、エープリル・ジェイ・エイ

アメリカ合衆国カリフォルニア州95054・サンタクララ・コーレデプリマベラ 2227

(72)発明者 チョーラ、ナリンダー・ケイ

アメリカ合衆国カリフォルニア州94587・ユニオンシティ・#712・ユニオンスクエア 33

(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 ランクマール、ジャヤラクシミ

アメリカ合衆国カリフォルニア州94555・フレモント・メイバードサークル 34359

(72)発明者 ガンディー、アミーナ・アール

アメリカ合衆国カリフォルニア州94025・メンロパーク・#1・ローブルアベニュー 837

(72)発明者 ポリッキー、ジェニファー・エル

アメリカ合衆国カリフォルニア州95118・サンノゼ・ジャービスコート 1511

- (72)発明者 ボーゲン、マライア・アール
アメリカ合衆国カリフォルニア州9 4 5 7 7・サンレアンドロ・サンティアゴロード 1 4 2 4 4
- (72)発明者 トリボレー、キャサリン・エム
アメリカ合衆国カリフォルニア州9 4 1 0 7・サンフランシスコ・# 5・テネシーストリート 1
1 2 1
- (72)発明者 バンドマン、オルガ
アメリカ合衆国カリフォルニア州9 4 0 4 3・マウンテンビュー・アンナアベニュー 3 6 6
- (72)発明者 ニュエン、ダニエル・ビー
アメリカ合衆国カリフォルニア州9 5 1 1 8・サンノゼ・リッジウッドドライブ 1 4 0 3
- (72)発明者 リュ、ヤン
アメリカ合衆国カリフォルニア州9 4 3 0 3・パロアルト・コリーナウェイ 3 8 8 5
- (72)発明者 パーフォード、ニール
アメリカ合衆国コネチカット州0 6 4 2 2・ダラム・ワイルドウッドサークル 1 0 5
- (72)発明者 ラル、ブリーティ
アメリカ合衆国カリフォルニア州9 5 0 5 6・サンタクララ・ピーオーボックス 5 1 4 2
- (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9 4 3 0 6・パロアルト・# 1 4 6・アルマストリート 3 3 5
3
- (72)発明者 ヤオ、モニック・ジー
アメリカ合衆国カリフォルニア州9 4 0 4 3・マウンテンビュー・フレデリックコート 1 1 1
- (72)発明者 エリオット、ビッキー・エス
アメリカ合衆国カリフォルニア州9 5 1 2 1・サンノゼ・ポルトンブレイスウェイ 3 7 7 0
- (72)発明者 レシボン、シャーリー・エイ
アメリカ合衆国カリフォルニア州9 5 1 1 5・サンフランシスコ・フォーチュナアベニュー 8 5
- (72)発明者 キーニー、ライアム
アメリカ合衆国カリフォルニア州9 4 1 2 7・サンノゼ・ウッドサイドアベニュー 5 0
- (72)発明者 リュ、デュング・アイナ・エム
アメリカ合衆国カリフォルニア州9 5 1 2 3・サンノゼ・コイドライブ 2 3 3
- (72)発明者 グリーンワールド、サラ・アール
アメリカ合衆国カリフォルニア州9 4 1 3 2・サンフランシスコ・ブカレリドライブ 2 1
- (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9 5 1 1 8・サンノゼ・ランウィックコート 4 2 3 0
- (72)発明者 スー、ユーミング
アメリカ合衆国カリフォルニア州9 4 0 4 0・マウンテンビュー・ウォルナットドライブ 1 7 3
9
- (72)発明者 ウォルシュ、ロドリック・ティー
イギリス国ケント州・シーティー1 3 イージー・カンタベリー・セントローレンスロード・バ
ウンダリーコート 8
- (72)発明者 ギーツェン、キンバリー・ジェイ
アメリカ合衆国カリフォルニア州9 5 1 2 3・サンノゼ・ロスウエコスドライブ 6 9 1
- (72)発明者 ヤング、ジュンミン
アメリカ合衆国カリフォルニア州9 5 1 2 9・サンノゼ・パークレーン 7 1 2 5
- (72)発明者 ジャクソン、ジェニファー・エル
アメリカ合衆国カリフォルニア州9 5 0 6 2・サンタクルス・リナコート 1 8 2 6

F ターム(参考) 2G045 AA34 AA35 AA40 BA11 BB50 DA12 DA13 DA36 FB02 FB03
4B024 AA01 AA11 BA10 CA01 CA04 CA09 DA02 DA05 DA11 EA02
EA03 EA04 FA02 GA01 GA11 HA01 HA03 HA12 HA15
4B050 CC01 CC03 DD11 LL01 LL03
4B063 QA01 QA18 QQ05 QQ13 QQ27 QQ42 QQ52 QR08 QR33 QR42
QR55 QR59 QR62 QR74 QR80 QR82 QS05 QS25 QS36 QX02

4B064 AG27 CA10 CA20 CC24 DA01 DA13
4B065 AA01X AA57X AA90X AA93Y AB01 AB02 BA01 BA08 CA25 CA29
CA44 CA46
4C084 AA01 AA02 AA17 BA01 BA08 BA22 BA23 CA18 DC25 NA14
ZA361 ZA451 ZB071 ZB261 ZC031 ZC331 ZC421
4H045 AA11 AA20 AA30 CA40 DA76 EA20 EA50 FA72 FA74

专利名称(译)	人类激酶		
公开(公告)号	JP2004527209A	公开(公告)日	2004-09-09
申请号	JP2002513883	申请日	2001-07-20
[标]申请(专利权)人(译)	迈克尔·桑顿		
申请(专利权)人(译)	桑顿, 迈克尔·		
[标]发明人	ソーントンマイケル ユエヘンリー カーンファラエイ ガルラジャンラジャゴバル ハファリアエープリルジェイエイ チョーラナリンダーケイ アービズチャンドラエス ランクマールジャヤラクシミ ガンディーアミーナアール ポリッキージェニファーエル ボーグンマライアアール トリボレーキャサリーンエム バンドマンオルガ ニュエンダニエルビー リュヤン バーフォードニール ラルプリーティ デイングリー ヤオモニークジー エリオットビッキーエス レシボンシャーリーエイ キーニーライアム リュデュングアイナエム グリーンワールドサラアール タングワイトム スーユーミング ウォルシュロドリックティー ギーツエンキンバリージェイ ヤングジュンミング ジャクソンジェニファーエル		
发明人	ソーントン、マイケル ユエ、ヘンリー カーン、ファラ・エイ ガルラジャン、ラジャゴバル ハファリア、エープリル・ジェイ・エイ チョーラ、ナリンダー・ケイ アービズ、チャンドラ・エス ランクマール、ジャヤラクシミ ガンディー、アミーナ・アール ポリッキー、ジェニファー・エル ボーグン、マライア・アール トリボレー、キャサリーン・エム バンドマン、オルガ ニュエン、ダニエル・ビー		

リュ、ヤン
 バーフォード、ニール
 ラル、プリーティ
 ディング、リー
 ヤオ、モニーク・ジー
 エリオット、ビッキー・エス
 レシボン、シャーリー・エイ
 キーニー、ライアム
 リュ、デュング・アイナ・エム
 グリーンワールド、サラ・アール
 タング、ワイトム
 スー、ユーミング
 ウォルシュ、ロドリック・ティー
 ギーツエン、キンバリー・ジェイ
 ヤング、ジュンミング
 ジャクソン、ジェニファー・エル

IPC分类号	G01N33/50 A61K38/00 A61K45/00 A61P3/06 A61P5/06 A61P9/00 A61P9/10 A61P35/00 A61P37/00 A61P43/00 C07K16/40 C12N1/15 C12N1/19 C12N1/21 C12N5/10 C12N9/12 C12N15/09 C12P21/08 C12Q1/02 C12Q1/68 G01N33/15 G01N33/53 G01N33/566
CPC分类号	C12N9/1205
FI分类号	C12N15/00.ZNA.A A61K45/00 A61P3/06 A61P5/06 A61P9/00 A61P9/10 A61P35/00 A61P37/00 A61P43/00.111 C07K16/40 C12N1/15 C12N1/19 C12N1/21 C12N9/12 C12P21/08 C12Q1/02 C12Q1/68.A G01N33/15.Z G01N33/50.Z G01N33/53.D G01N33/53.M G01N33/566 C12N5/00.A A61K37/02
F-TERM分类号	2G045/AA34 2G045/AA35 2G045/AA40 2G045/BA11 2G045/BB50 2G045/DA12 2G045/DA13 2G045/DA36 2G045/FB02 2G045/FB03 4B024/AA01 4B024/AA11 4B024/BA10 4B024/CA01 4B024/CA04 4B024/CA09 4B024/DA02 4B024/DA05 4B024/DA11 4B024/EA02 4B024/EA03 4B024/EA04 4B024/FA02 4B024/GA01 4B024/GA11 4B024/HA01 4B024/HA03 4B024/HA12 4B024/HA15 4B050/CC01 4B050/CC03 4B050/DD11 4B050/LL01 4B050/LL03 4B063/QA01 4B063/QA18 4B063/QQ05 4B063/QQ13 4B063/QQ27 4B063/QQ42 4B063/QQ52 4B063/QR08 4B063/QR33 4B063/QR42 4B063/QR55 4B063/QR59 4B063/QR62 4B063/QR74 4B063/QR80 4B063/QR82 4B063/QS05 4B063/QS25 4B063/QS36 4B063/QX02 4B064/AG27 4B064/CA10 4B064/CA20 4B064/CC24 4B064/DA01 4B064/DA13 4B065/AA01X 4B065/AA57X 4B065/AA90X 4B065/AA93Y 4B065/AB01 4B065/AB02 4B065/BA01 4B065/BA08 4B065/CA25 4B065/CA29 4B065/CA44 4B065/CA46 4C084/AA01 4C084/AA02 4C084/AA17 4C084/BA01 4C084/BA08 4C084/BA22 4C084/BA23 4C084/CA18 4C084/DC25 4C084/NA14 4C084/ZA361 4C084/ZA451 4C084/ZB071 4C084/ZB261 4C084/ZC031 4C084/ZC331 4C084/ZC421 4H045/AA11 4H045/AA20 4H045/AA30 4H045/CA40 4H045/DA76 4H045/EA20 4H045/EA50 4H045/FA72 4H045/FA74
優先権	60/220038 2000-07-21 US 60/222112 2000-07-28 US 60/222831 2000-08-04 US 60/224729 2000-08-11 US
外部リンク	Espacenet

摘要(译)

本发明提供的多核苷酸鉴定的人激酶 (PKIN) 和PKIN，代码。本发明还提供了表达载体，宿主细胞，抗体，也激动剂和拮抗剂来提供。本发明还诊断有PKIN的异常表达相关的疾病，还提供了治疗或预防的方法。

検索ワード	分析対象となるプログラムの例
CNY(CR)ENST	例えば CENSAN(Stanford University CA,USA) または RENES(Computer Circuits Group, The Sanger Centre, Cambridge, UK)を用いたゲーム翻訳のためのエディタ系ツール
CH	ゲーム翻訳の重編翻訳
EL	ステイッチまたはストレッチされたゲーム翻訳 (詳細は参 照)
INCY	BSI 翻訳プログラムへのマッピングからの全単語とエディタ の対照 エディタと単語を対照するため、ゲーム位置と BSI 構造データの紐付けされる。