

(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2004-511204

(P2004-511204A)

(43) 公表日 平成16年4月15日(2004.4.15)

(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 2 G O 4 5
A 6 1 K 38/00	A 6 1 K 45/00	4 B O 2 4
A 6 1 K 45/00	A 6 1 K 49/00	A 4 B O 5 0
A 6 1 K 49/00	A 6 1 P 3/06	4 B O 6 3
A 6 1 P 3/06	A 6 1 P 9/00	4 B O 6 4
	審査請求 未請求 予備審査請求 有	(全 282 頁) 最終頁に続く

(21) 出願番号	特願2001-578626 (P2001-578626)	(71) 出願人	301005050
(86) (22) 出願日	平成13年4月20日 (2001. 4. 20)		インサイト・ゲノミックス・インコーポレ
(85) 翻訳文提出日	平成14年10月18日 (2002. 10. 18)		イテッド
(86) 国際出願番号	PCT/US2001/012992		アメリカ合衆国カリフォルニア州9 4 3 0
(87) 国際公開番号	W02001/081555		4・パロアルト・ポータードライブ 3 1
(87) 国際公開日	平成13年11月1日 (2001. 11. 1)		6 0
(31) 優先権主張番号	60/199, 021	(74) 代理人	100089266
(32) 優先日	平成12年4月20日 (2000. 4. 20)		弁理士 大島 陽一
(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	60/200, 226		アメリカ合衆国カリフォルニア州9 4 0 8
(32) 優先日	平成12年4月28日 (2000. 4. 28)		7・サニーベイル・ルイスアベニュー 8
(33) 優先権主張国	米国 (US)		2 6
(31) 優先権主張番号	60/202, 339		
(32) 優先日	平成12年5月5日 (2000. 5. 5)		
(33) 優先権主張国	米国 (US)		

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) と、PKINを同定及びコードするポリヌクレオチドとを提供する。本発明はまた、発現ベクター、宿主細胞、抗体、アゴニスト、及びアンタゴニストを提供する。更に、本発明は、PKINの異常発現に関連する疾患の診断・治療・予防方法を提供する。

【特許請求の範囲】

【請求項 1】

以下の (a) 乃至 (d) を有する群から選択した実質上単離されたポリペプチド。

(a) SEQ ID NO: 1-18 (配列番号 1 乃至 18) 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: 2-18 を有する群から選択したアミノ酸配列と少なくとも 90% が同一であるようなアミノ酸配列を有する天然のポリペプチド

(c) SEQ ID NO: 1 (配列番号 1) のアミノ酸配列と少なくとも 98% が同一であるようなアミノ酸配列を有する天然のポリペプチド

(d) SEQ ID NO: 1-18 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

(e) SEQ ID NO: 1-18 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

【請求項 2】

SEQ ID NO: 1-18 を有する群から選択した請求項 1 に記載の単離されたポリペプチド。

【請求項 3】

請求項 1 のポリペプチドをコードする単離されたポリヌクレオチド。

【請求項 4】

請求項 2 のポリペプチドをコードする単離されたポリヌクレオチド。 20

【請求項 5】

SEQ ID NO: 19-36 (配列番号 19 乃至 36) を有する群から選択した請求項 4 に記載の単離されたポリヌクレオチド。

【請求項 6】

請求項 3 に記載のポリヌクレオチドに機能的に結合したプロモーター配列を含む組換えポリヌクレオチド。

【請求項 7】

請求項 6 に記載の組換えポリヌクレオチドを用いて形質転換した細胞。

【請求項 8】

請求項 6 に記載の組換えポリヌクレオチドを含む遺伝形質転換体。 30

【請求項 9】

請求項 1 のポリペプチドを生産する方法であって、

(a) 前記ポリペプチドの発現に好適な条件下で、請求項 1 のポリペプチドをコードするポリヌクレオチドに機能的に結合されたプロモーター配列を含む組換えポリヌクレオチドで形質転換された細胞を培養するステップと、

(b) そのように発現した前記ポリペプチドを回収する過程とからなり、前記組換えポリヌクレオチドが、請求項 1 に記載の前記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ことを特徴とする方法。

【請求項 10】

請求項 1 に記載のポリペプチドと特異結合するような単離された抗体。 40

【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: 19-36 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: 20-36 (配列番号 20 乃至 36) 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% が同一であるような天然のポリヌクレオチド

(c) 配列番号 19 のポリヌクレオチド配列と少なくとも 98% が同一であるようなポリヌクレオチド配列を有する天然のポリヌクレオチド

(d) (a) のポリヌクレオチドに相補的なポリヌクレオチド

(e) (b) のポリヌクレオチドに相補的なポリヌクレオチド

(f) (c) のポリヌクレオチドに相補的なポリヌクレオチド

(g) (a) ~ (f) の RNA 等価物

【請求項 12】

請求項 11 に記載のポリヌクレオチドの少なくとも 60 の連続したヌクレオチドを含む単離されたポリヌクレオチド。

【請求項 13】

請求項 11 に記載のポリヌクレオチドの配列を有する標的ポリヌクレオチドをサンプル中から検出する方法であって、

(a) 前記サンプル中の前記標的ポリヌクレオチドに相補的な配列を有する少なくとも 20 の連続したヌクレオチドを含むプローブを用いて前記サンプルをハイブリダイズする過程と、 10

(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、前記プローブが前記標的ポリヌクレオチドに特異的にハイブリダイズすることを特徴とする方法。

【請求項 14】

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

【請求項 15】

請求項 11 に記載のポリヌクレオチドの配列を有する標的ポリヌクレオチドをサンプル中から検出する方法であって、

(a) ポリメラーゼ連鎖反応増幅を用いて前記標的ポリヌクレオチドまたはその断片を増幅する過程と、

(b) 前記標的ポリヌクレオチドまたはその断片の存在・不存在を検出し、該標的ポリヌクレオチドまたはその断片が存在する場合にはオプションでその量を検出する過程を含むことを特徴とする方法。

【請求項 16】

請求項 1 のポリペプチドと、薬剤として許容できる賦形剤とを有することを特徴とする成分。 30

【請求項 17】

前記ポリペプチドが、SEQ ID NO: 1-18 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。

【請求項 19】

請求項 1 に記載のポリペプチドのアゴニストとして有効性を確認するために化合物をスクリーニングする方法であって、

(a) 請求項 1 のポリペプチドを含むサンプルを化合物に曝露するステップと、 40

(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。

【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 50

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ステップと、
- (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 2 3】

請求項 2 2 に記載の方法によって同定したアンタゴニスト化合物と、薬剤として許容できる賦形剤とを含むことを特徴とする成分。

【請求項 2 4】

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 2 3 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。 10

【請求項 2 5】

請求項 1 に記載のポリペプチドに特異結合する化合物をスクリーニングする方法であって、

- (a) 適切な条件下で請求項 1 に記載のポリペプチドを少なくとも 1 つの試験化合物に結合させる過程と、

- (b) 請求項 1 に記載のポリペプチドの試験化合物との結合を検出し、それによって請求項 1 に記載のポリペプチドに特異結合する化合物を同定する過程とを含むことを特徴とする方法。

【請求項 2 6】

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、 20

- (a) 請求項 1 に記載のポリペプチドの活性が許容される条件下で、請求項 1 に記載のポリペプチドを少なくとも 1 つの試験化合物に結合させる過程と、

- (b) 請求項 1 に記載のポリペプチドの活性を試験化合物の存在下で算定する過程と、

- (c) 試験化合物の存在下での請求項 1 に記載のポリペプチドの活性を、試験化合物の不存在下での請求項 1 に記載のポリペプチドの活性と比較する過程とを含み、試験化合物の存在下での請求項 1 に記載のポリペプチドの活性の変化が、請求項 1 に記載のポリペプチドの活性を調節する化合物を標示することを特徴とする方法。

【請求項 2 7】

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、 30

- (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ステップと、

- (b) 前記標的ポリヌクレオチドの変異発現を検出する過程と、

- (c) 可変量の前記化合物の存在下と前記化合物の不存在下で、前記標的ポリヌクレオチドの発現を比較する過程とを含むことを特徴とする方法。

【請求項 2 8】

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 1 1 のポリヌクレオチドの少なくとも 2 0 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で、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 1 1 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、 40

- (c) ハイブリダイゼーション複合体の収量を定量するステップと、

- (d) 前記処理された生物学的サンプル中の前記ハイブリタイゼーション複合体の量を、処理されていない生物学的サンプル中の前記ハイブリタイゼーション複合体の量と比較する過程とを含み、前記処理された生物学的サンプル中の前記ハイブリタイゼーション複合体の量の差が、前記試験化合物の毒性を標示することを特徴とする方法。 50

【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) F a b 断片

(d) F (a b ') 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 P K I N 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 P K I N 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、S E Q I D N O : 1 - 18 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またはその免疫抗原性断片を用いて動物を免疫化する過程と、

(b) 前記動物から抗体を単離する過程と、

(c) 前記単離された抗体をポリペプチドでスクリーニングし、それによって、S E Q I D N O : 1 - 18 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、S E Q I D N O : 1 - 18 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) 不死化細胞を用いて前記抗体産出細胞を融合して、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) S E Q I D N O : 1 - 18 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

10

20

30

40

50

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】

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。 10

【請求項 43】

S E Q I D N O : 1 - 1 8 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 - 1 8 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標示することを特徴とする方法。

【請求項 44】

S E Q I D N O : 1 - 1 8 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、 20

(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 - 1 8 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とする方法。

【請求項 45】

配列番号 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】

配列番号 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。 30

【請求項 47】

配列番号 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】

配列番号 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】

配列番号 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。

【請求項 50】

配列番号 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】

配列番号 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。 40

【請求項 52】

配列番号 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】

配列番号 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】

配列番号 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。

【請求項 55】

配列番号 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】

配列番号 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】

配列番号 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】

配列番号 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】

配列番号 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。

【請求項 60】

配列番号 16 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 61】

10

配列番号 17 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 62】

配列番号 18 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 63】

配列番号 19 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 64】

配列番号 20 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 65】

配列番号 21 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 66】

20

配列番号 22 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 67】

配列番号 23 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 68】

配列番号 24 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 69】

配列番号 25 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 70】

配列番号 26 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 71】

30

配列番号 27 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 72】

配列番号 28 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 73】

配列番号 29 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 74】

配列番号 30 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 75】

配列番号 31 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 76】

40

配列番号 32 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 77】

配列番号 33 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 78】

配列番号 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】

配列番号 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】 配列番号 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。

【発明の詳細な説明】

50

【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。可逆的なタンパクの質リン酸化は、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。

10

20

【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関与する。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。

【0004】

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ゼ (PTK) 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 (STK) 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 PTK および STK 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、ATP 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 VI-X I はタンパク質基質と結合し、γ リン酸を ATP 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 ATP 結合共通モチーフを含み、サブドメイン II 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 VI から IX は高度に保存された触媒中心を含む。PTK および STK 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 VI および VII I に固有の配列モチーフを含む。

30

40

【0005】

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

50

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、ATP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)。

【0006】

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 (PTK) は、膜貫通受容体 PTK タンパク質または非膜貫通非受容体 PTK 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 (RTK) 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 PTK に結合する成長因子 (GF) 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様成長因子、神経成長因子、血管内皮成長因子、およびマクロファージコロニー刺激因子が含まれる。

【0007】

非膜貫通型非受容体 PTK 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 PTK 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。

【0008】

多くの PTK は、PTK 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 PTK をコードする。更に、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)。従って、PTK 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。

【0009】

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による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) の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な 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、細胞骨格の調節に関与すると思われる。ERK の活性化は通常一過性であり、細胞は、そのダウンレギュレーション (下方制御) に関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、ERK 活性の上昇がある種の癌に関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、チェックポイントおよび細胞周期キナーゼ、Numb 関連キナーゼ (Nak)、ヒト融合 (hFu)、増殖関連キナーゼ、5'-AMP-活性化プロテインキナーゼ、およびアポトーシスに関与するキナーゼが含まれる。

【0010】

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)

10

20

30

40

50

、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関与し、ホルモン刺激に応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. ら, (1994) Harrison's Principles of Internal Medicine, McGraw-Hill, New York, NY, 416-431ページ, 1887)。

10

【0011】

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上に存在する (Fish, K. J. ら, (1995) J. Biol. Chem. 270:14875-14883.)。

20

【0012】

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 (Cegieliska, A. ら, (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からなる。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。 α 、 β 、 γ 、 δ 、および ϵ の少なくとも5つの哺乳動物イソ型が存在する。Fishらが、ヒト胎盤cDNAライブラリからCKI ϵ を同定した。このCKI ϵ は416のアミノ酸からなる塩基性タンパク質であり、CKIデルタに類似している。組み替え発現によって、CKI ϵ 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 ϵ 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)。

30

【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 ϵ 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 ϵ 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 ϵ 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 ϵ 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 ϵ 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因子であると主張した。従って、CKI ϵ 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. ら, (2000) Science 288:483-491.)。

40

【0014】

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/ス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

50

. G. ら, (2000) *Biochimie* 82:1123-7). HIPKはホメオプロテインと相互作用するドメインから別れた、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よく発現される(Kim, Y. H. ら, (1998) *J. Biol. Chem.* 273:25875-9; Wang, Y. ら, (2001) *Biochim. Biophys. Acta* 1518:168-172). HIPKは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. ら, ., (1999) *Proc. Nat. Acad. Sci. U. S. A.* 96:12350-5)。

10

【0015】

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、および囊胞性線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. ら, (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様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. ら, (1994) *J. Neurosci.* 14:1-13)。

20

30

【0016】

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キナーゼシグナリング経路は哺乳類細胞およびイーストに存在する。MAPキナーゼの経路を活性化する細胞外刺激には、上皮成長因子(EGF)、紫外線、高浸透圧性媒体、熱ショック、リポ多糖(LPS)および腫瘍壊死因子(TNF)やインターロイキン-1(IL-1)のような炎症誘発性サイトカインなどが含まれる。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参与する。

40

【0017】

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

50

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関与する選択されたタンパク質がリン酸化され活性化される。CDK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。

【0018】

NIMA（有糸分裂に決して入らない）関連キナーゼ（Neks）は細胞周期に関連するSTKの別のファミリーである。CDKおよびNeks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関与する（Fry, A. M. ら, (1998) EMBO J. 17: 470-481）。

10

【0019】

チェックポイントおよび細胞周期キナーゼ

細胞分化のプロセスでは、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. ら, (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 ら, (1997) Science 277: 1501-1505.）。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。

20

30

【0020】

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. ら, (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

40

【0021】

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. ら, (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αサブユニットと、このα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βサブユニットおよびγ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

50

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。

【0022】

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S TKがこのプロセスにおいて重要な役割を果たす。Z I P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末端ロイシンジッパードメインを含むS TK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（A T F / C R E B）ファミリーのメンバーである活性化転写因子A T F 4などの転写因子との相互作用を仲介する（S a n j o, H. ら, (1998) J. Biol. Chem., 273:29066-29071）。D R A K 1およびD R A K 2は、死関連プロテインキナーゼ（D A Pキナーゼ）と相同性を共有するS TKであって、インターフェロン γ 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（S a n j o ら 前出）。Z I Pキナーゼと同様に、D A P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。Z I P、D A P、およびD R A Kキナーゼは、N I H 3 T 3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（S a n j o 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。

10

20

【0023】

R I C Kは、死受容体C D 9 5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S TKである（I n o h a r a, N. ら, (1998) J. Biol. Chem., 273:12296-12300）。C D 9 5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（N a g a t a, S. (1997) Cell 88:355-365）。C D 9 5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。R I C Kは、カスパーゼ様ドメインと相互作用するC末端「カスパーゼ補充」ドメインおよびN末端キナーゼ触媒ドメインとからなる。このことから、R I C Kがカスパーゼ-8の動員においてある役割を果たしていると思われる。この解釈は、ヒト293T細胞におけるR I C Kの発現がカスパーゼ-8の活性化を促進し、C D 9 5アポトーシス経路に参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補強される（I n o h a r a ら, 前出）。

30

【0024】

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に対する配列によって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（M P K）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、原始的な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。M P Kは、分枝鎖 α 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（H a r r i s, R. A. ら, (1995) Adv. Enzyme Regul. 34:147-162）。5つのM P Kが同定された。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

40

50

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 α 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 (H a r r i s , R . A . ら, (1997) A d v . E n z y m e R e g u l . 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 (H a r r i s ら, 前出)。

【0025】

(非タンパク質基質に対するキナーゼ)

10

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 (P I) 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 (L e e v e r s , S . ら, (1999) C u r r . O p i n . C e l l . B i o l . 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。イノシトールホスホリピッド (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 (P I) 残基のリン酸化が起こり、それによってP I 残基が二リン酸状態 (P I P₂) に変換される。次にP I P₂ がイノシトール三リン酸 (I P₃) 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、およびトロンビン誘導性血小板凝集がある。

20

【0026】

P I およびその誘導体のD 3位をリン酸化するP I 3-キナーゼ (P I 3 K) 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す。P I 3 K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 (I R S) -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、P I (4, 5) 二リン酸 (P I P₂) をP I (3, 4, 5) P₃ (P I P₃) に変換する。次にP I P₃ が、P K A、プロテインキナーゼB (P K B)、プロテインキナーゼC (P K C)、グリコーゲンシンターゼキナーゼ (G S K) -3、およびP 70リボソームs 6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。P I 3 Kもまた、細胞骨格形成タンパク質であるR a c、r h o、およびc d c 42と直接相互作用する (S h e p h e r d , P . R . ら, (1998) B i o c h e m . J . 333:471-490)。o b e s eマウスおよびf a tマウスなどの糖尿病動物モデルは、P I 3 k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、P I 3 Kが2型糖尿病においてある役割を果たしていると思われる (S h e p h e r d、前出)。

30

40

【0027】

脂質キナーゼリン酸化活性の例は、D-エリトロー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 (S P P) へのリン酸化である。S P P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 (K o h a m a , T . ら (1998) J . B i o l . C h e m . 273:23722-23728)。細胞外では、S P PはGタンパク質結合受容体E D G-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、S P Pは、細胞の成長

50

、生存、運動性、および細胞骨格の変化を調節する。S P Pのレベルは、D-エリトロースフィンゴシンを特異的にリン酸化してS P P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(P D G F)、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S P P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P D G F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(K o h a m aら, 前出)。

【0028】

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(A T P : A M Pホスホトランズフェラーゼ、もしくはA d K)およびグアニル酸キナーゼ(A T P : G M Pホスホトランズフェラーゼ、もしくはG u K)が、ヌクレオチド代謝において重要な役割を果たし、A T PおよびG T P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D N A合成およびR N A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(A T P)であり、シグナル伝達経路(G T P)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(P i l l w e i n, K. ら, (1990) C a n c e r R e s. 50:1576-1579)。

【0029】

A d Kはほとんどの細胞型に見られ、高い速度でA T P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、A d K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、A T Pを生成する代謝プロセスからA T P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A d Kが重要な役割を果たしていることが実証された(Z e l e z n i k a r, R. J. ら, (1995) J. B i o l. C h e m. 270:7311-7319)。従って、A d K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めの候補を提供し得る。別法では、A d K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、A d Kの活性を上昇させて治療できるであろう。

【0030】

R N AおよびD N A合成のためのG T Pの合成における重要なステップを提供するのに加えて、G u Kはまた、G D PおよびG T P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G T P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A M Pが生産される。Gタンパク質に結合するG D Pはこれらのプロセスを阻害する。また、G D PおよびG T Pのレベルによって、細胞増殖の調節および発癌に関与するとして知られるp 2 1^{r a s}などのある種の腫瘍タンパク質の活性が調節される(B o s, J. L. (1989) C a n c e r R e s. 49:4682-4689)。G u Kの抑制によって起こるG D Pに対するG T Pの比率が高いと、p 2 1^{r a s}が活性化され発癌が促進される。G D Pのレベルを上昇させG D Pに対するG T Pのレベルを低下させるためにG u Kの活性を高めることが、発癌を抑制する治療となり得る。

【0031】

G u K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b u c i c l o v i rが含まれる(M i l l e r, W. H. およびM i l l e r R. L. (1980) J. B i o l. C h e m. 255:7204-7207 ; S t e n b e r g, K. ら, (1986) J. B i o l. C h e m. 261:213

10

20

30

40

50

4-2139)。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。

【0032】

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. ら, (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)。

10

【0033】

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。

【0034】

20

(発明の要約)

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」および「PKIN-18」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: 1-18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: 1-18を有する群から選択したアミノ酸配列と少なくとも90%が同一であるアミノ酸配列を有する天然のポリペプチド、(c) SEQ ID NO: 1-18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: 1-18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: 1-18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。

30

【0035】

また、本発明は(a) SEQ ID NO: 1-18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: 1-18を有する群から選択したアミノ酸配列と少なくとも90%が同一であるアミノ酸配列からなる天然のポリペプチド、(c) SEQ ID NO: 1-18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: 1-18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: 1-18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: 19-36を有する群から選択される。

40

本発明は更に、(a) SEQ ID NO: 1-18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: 1-18を有する群から選択したアミノ酸配列と少なくとも90%の相同性のあるアミノ酸配列からなる天然のポリペプチド、(c) SEQ ID NO: 1-18を有する群から選択したアミノ酸配列の生物学的活性

50

断片、または (d) SEQ ID NO: 1-18 を有する群から選択したアミノ酸配列の免疫抗原性断片を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。

【0036】

更に本発明は、(a) SEQ ID NO: 1-18 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: 1-18 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アミノ酸配列からなる天然のポリペプチドと、(c) SEQ ID NO: 1-18 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: 1-18 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。この方法は、(a) このポリペプチドの発現に好適な条件下で、このポリペプチドをコードするポリヌクレオチドと機能的に結合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ステップと、(b) このように発現したポリペプチドを回収するステップとを含む。

10

【0037】

本発明は更に、(a) SEQ ID NO: 1-18 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: 1-18 を有する群から選択したアミノ酸配列と少なくとも90%同一であるアミノ酸配列を有する天然のポリペプチド、(c) SEQ ID NO: 1-18 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: 1-18 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。

20

【0038】

本発明は更に、(a) SEQ ID NO: 19-36 を有する群から選択したポリヌクレオチド配列、(b) SEQ ID NO: 19-36 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%同一であるポリヌクレオチド配列からなる天然のポリヌクレオチド、(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等価物を含む群から選択された実質上単離されたポリヌクレオチドを提供する。一実施態様では、ポリヌクレオチドは少なくとも60の連続したヌクレオチドを有する。

30

【0039】

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し、該標的ポリヌクレオチドは(a) SEQ ID NO: 19-36 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: 19-36 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%同一のポリヌクレオチド配列からなる天然のポリヌクレオチド、(c) (a) のポリヌクレオチドに相補的なポリヌクレオチド、(d) (b) のポリヌクレオチドに相補的なポリヌクレオチド、および(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) サンプル中の標的ポリヌクレオチドに相補的な配列からなる少なくとも20の連続したヌクレオチドを含むプローブを用いて該サンプルをハイブリダイズし、またプローブと標的ポリヌクレオチドまた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、プローブが標的ポリヌクレオチドに特異的に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。

40

【0040】

50

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: 19-36 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: 19-36 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% 同一であるポリヌクレオチド配列からなる天然のポリヌクレオチド、(c) (a) のポリヌクレオチドに相補的なポリヌクレオチド、(d) (b) のポリヌクレオチドに相補的なポリヌクレオチド、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。

10

【0041】

本発明は更に、有効量のポリペプチドと薬剤として許容できる賦形剤とを含む成分を提供し、有効量のポリペプチドは、(a) SEQ ID NO: 1-18 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: 1-18 を有する群から選択したアミノ酸配列と少なくとも 90% 同一であるアミノ酸配列を有する天然のポリペプチド、(c) SEQ ID NO: 1-18 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: 1-18 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: 1-18 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、このような治療を必要としている患者にこの組成物を投与することを含む、機能的 PK IN の発現低下に関連した疾患や症状の治療方法を提供する。

20

【0042】

本発明はまた、(a) SEQ ID NO: 1-18 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: 1-18 を有する群から選択したアミノ酸配列と少なくとも 90% 同一であるアミノ酸配列を含む天然のポリペプチド、(c) SEQ ID NO: 1-18 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の生物学的活性断片、および (d) SEQ ID NO: 1-18 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のアゴニストとしての有効性を確認するために化合物をスクリーニングする方法を提供する。スクリーニング方法は、(a) ポリペプチドを含むサンプルを化合物に暴露する過程と、(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 PK IN 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。

30

【0043】

本発明は更に、(a) SEQ ID NO: 1-18 を有する群から選択したアミノ酸配列、(b) SEQ ID NO: 1-18 を有する群から選択したアミノ酸配列と少なくとも 90% の相同性を有する天然のアミノ酸配列、(c) SEQ ID NO: 1-18 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: 1-18 を有する群から選択したアミノ酸配列の免疫抗原性断片を含む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 PK IN 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。

40

【0044】

50

本発明は更に、(a) SEQ ID NO: 1-18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: 1-18を有する群から選択したアミノ酸配列と少なくとも90%同一のアミノ酸配列を含む天然のポリペプチド、(c) SEQ ID NO: 1-18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の生物学的活性断片、および(d) SEQ ID NO: 1-18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) 試験化合物とポリペプチドとの結合を検出し、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。

10

【0045】

更に本発明は、(a) SEQ ID NO: 1-18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: 1-18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アミノ酸配列からなる天然のポリペプチドと、(c) SEQ ID NO: 1-18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活性断片と、(d) SEQ ID NO: 1-18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群から選択され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。このスクリーニング方法は、(a) このポリペプチドを、その活性が許容される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) この試験化合物の存在下でのこのポリペプチドの活性を評価するステップと、(c) この試験化合物の存在下でのこのポリペプチドの活性と、この試験化合物の不在下でのこのポリペプチドの活性とを比較するステップとを含み、この試験化合物の存在下でのこのポリペプチドの活性の変化が、このポリペプチドの活性を調節する化合物の存在を示唆するという特徴を有する。

20

【0046】

更に本発明は、SEQ ID NO: 19-36からなる一群から選択された配列を含む標的ポリヌクレオチドの発現を変化させるのに効果的な化合物をスクリーニングする方法であって、(a) この標的ポリヌクレオチドを含むサンプルを化合物に曝露するステップと、(b) この標的ポリヌクレオチドの発現の変化を検出するステップとを含む、該スクリーニング方法を提供する。

30

【0047】

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: 19-36を有す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: 19-36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%同一であるポリヌクレオチド配列を含む天然のポリヌクレオチド、(iii) (i) に相補的な配列を有するポリヌクレオチド、(iv) (ii) のポリヌクレオチドに相補的なポリヌクレオチド、(v) (i) ~ (iv) のRNA等価物を含む群から選択したポリヌクレオチドの少なくとも20の連続したヌクレオチドから構成されるプローブを用いて、処理した生物学的サンプルの核酸をハイブリダイズする過程とを含む試験化合物の毒性の算定方法を提供する。

40

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: 19-36を有す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: 19-36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%同一であるポリヌクレオチド配列を含む天然のポリヌクレオチド、(iii) (i) のポリヌクレオチドに相補的なポリヌクレオチド、(iv) (ii) のポリヌクレオチドに相補的なポリヌクレオチド、(v) (i) ~ (iv) のRNA等価物を含む群から選択する。或いは、標的ポリヌクレオチドは、上記(i) ~ (v)を含む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

50

ンプルのハイブリダイゼーション複合体の量を、非処理の生物学的サンプルのハイブリダイゼーション複合体の量と比較する過程を含み、処理した生物学的サンプルのハイブリダイゼーション複合体の量の差は、試験化合物の毒性を示す。

【0048】

(本発明の記載について)

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

10

【0049】

本明細書及び請求の範囲において単数形を表す「或る」、「その（この等）」は、文脈で明確に示していない場合は複数形を含むことに注意されたい。従って、例えば「或る宿主細胞」は複数の宿主細胞を含み、その「抗体」は複数の抗体が含まれ、当業者には周知の等価物なども含まれる。

【0050】

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。

20

【0051】

(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。

【0052】

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。

30

【0053】

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。

40

【0054】

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、或いは置換が起こっても、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

50

。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 P K I N 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 P K I N 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。

【0055】

用語「アミノ酸」及び「アミノ酸配列」は、オリゴペプチド、ペプチド、ポリペプチド、タンパク質配列、或いはそれらの任意の断片を指し、天然の分子及び合成分子を含む。「アミノ酸配列」が天然のタンパク質分子の配列を指す場合、「アミノ酸配列」及び類似の用語は、アミノ酸配列を記載したタンパク質分子に関連する完全で元のままのアミノ酸配列に限定するものではない。

【0056】

用語「増幅」は、核酸配列の複製物を作製することに関連する。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。

【0057】

用語「アンタゴニスト」は、P K I N 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、P K I N に直接相互作用するか、或いは P K I N が関与する生物学的経路の成分と作用して、P K I N 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。

【0058】

用語「抗体」は、抗原決定基と結合可能な F a b 及び F (a b ')₂、及びそれらの断片、F_v断片などの無傷の免疫グロブリン分子を指す。P K I N ポリペプチドと結合する抗体は、抗原を免疫する小ペプチドを含む無傷のポリペプチドまたはその断片を用いて作製可能である。動物（マウス、ラット、ウサギ等）を免疫化するために用いるポリペプチドまたはオリゴペプチドは、翻訳または化学合成された R N A に由来し得るもので、好みに応じて担体タンパク質に接合することも可能である。通常用いられる担体であってペプチドと化学結合するものは、ウシ血清アルブミン、サイログロブリン及びキーホールリンペットヘモシアニン（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。

【0059】

用語「抗原決定基」は、特定の抗体と接触する分子の領域（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、抗体に結合するための無損傷抗原（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。

【0060】

本明細書において「アンチセンス」は、特定の核酸配列のセンス（コーディング）鎖と塩基対を形成し得る任意の組成物を指す。アンチセンス成分には、D N A や、R N A や、ペプチド核酸（P N A）や、ホスホロチオ酸、メチルホスホン酸またはベンジルホスホン酸等の修飾されたバックボーン連鎖を有するオリゴヌクレオチドや、2'-メトキシエチル糖または2'-メトキシエトキシ糖等の修飾された糖類を有するオリゴヌクレオチドや、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ある。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」または「マイナス」という表現はアン

10

20

30

40

50

チセンス鎖であり、「正」または「プラス」という表現はセンス鎖である。

【0061】

用語「生物学的に活性」は、天然分子の構造的、調節的、或いは生化学的な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PKIN、合成のPKIN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。

【0062】

用語「相補的」は、塩基対合によってアニールする2つの一本鎖核酸配列間の関係を指す。例えば、配列「5' A-G-T 3'」は、相補配列「3' T-C-A 5'」に結合する。「所定のポリヌクレオチド配列を含む組成物」または「所定のアミノ酸配列を含む組成物」は広い意味で、所定のポリヌクレオチド配列若しくはアミノ酸配列を含む任意の組成物を指す。この成分には、乾燥製剤または水溶液が含まれ得る。PKIN若しくはPKIN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。

10

【0063】

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

20

【0064】

用語「保存的なアミノ酸置換」は、元のタンパク質の特性を殆ど変えない置換を指す。即ち、置換によってそのタンパク質の構造や機能が大きくは変わらず、そのタンパク質の構造、特にその機能が保存される。下表は、タンパク質中で元のアミノ酸と置換され得るアミノ酸と、保存アミノ酸置換と認められるアミノ酸を示している。

30

【0065】

元の残基	保存的な置換	
Ala	Gly, Set	
Arg	His, Lys	
Asn	Asp, Gln, His	
Asp	Asn, Glu	
Cys	Ala, Ser	10
Gln	Asn, Glu, His	
Glu	Asp, Gln, His	
Gly	Ala	
His	Asn, Arg, Gln, Glu	
Ile	Leu, Val	
Leu	Ile, Val	20
Lys	Arg, Gln, Glu	
Met	Leu, Ile	
Phe	His, Met, Leu, Trp, Tyr	
Ser	Cys, Thr	
Thr	Ser, Val	
Trp	Phe, Tyr	30
Tyr	His, Phe, Trp	
Val	Ile, Leu, Thr	

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

【0066】

用語「欠失」は、1個以上のアミノ酸残基が欠如するアミノ酸配列の変化、或いは1個以上のヌクレオチドが欠如する核酸配列の変化を指す。 40

【0067】

用語「誘導体」は、化学修飾されたポリヌクレオチドまたはポリペプチドを指す。例えば、アルキル基、アシル基、ヒドロキシル基またはアミノ基による水素の置換は、ポリヌクレオチド配列の化学修飾に含まれ得る。ポリヌクレオチド誘導体は、天然分子の生物学的または免疫学的機能を少なくとも1つは保持しているポリペプチドをコードする。ポリペプチド誘導体は、グリコシル化、ポリエチレングリコール化（*pegylation*）、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。 50

【0068】

「検出可能な標識」は、測定可能な信号を生成し得る、ポリヌクレオチドやポリペプチドに共有結合或いは非共有結合するレポーター分子や酵素を指す。

【0069】

「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。

【0070】

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 10
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2
50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 20
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。

【0071】

SEQ ID NO: 19-36の断片は、例えば、この断片を得たゲノム内の他の配列と
は異なる、SEQ ID NO: 19-36を明確に同定する固有のポリヌクレオチド配列
の領域を含む。SEQ ID NO: 19-36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: 19-36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: 19-36
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 30
日常的に測定できる。

【0072】

SEQ ID NO: 1-18のある断片は、SEQ ID NO: 19-36のある断片
によってコードされる。SEQ ID NO: 1-18のある断片は、SEQ ID NO:
1-18を特異的に同定する固有のアミノ酸配列の領域を含む。例えば、SEQ ID
NO: 1-18のある断片は、SEQ ID NO: 1-18を特異的に認識する抗体の開
発における免疫原性ペプチドとして有用である。ある断片と一致するSEQ ID NO
: 1-18の正確な断片の長さや領域は、その断片の目的に基づいて当分野で一般的な技
術によって日常的に決定できる。

【0073】

「完全長」ポリヌクレオチド配列とは、少なくとも1つの翻訳開始コドン（例えばメチオ
ニン）、オープンリーディングフレーム及び翻訳終止コドンを有する配列である。「完全
長」ポリヌクレオチド配列は、「完全長」ポリペプチド配列をコードする。 40

【0074】

「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。

【0075】

ポリヌクレオチド配列についての用語「パーセントの同一性」または「%の同一性」とは
、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る、2つ以上のポリヌクレオチ
ド配列間の一致する残基の百分率のことである。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 50

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。

【0076】

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）（DNASTAR, Madison WI）の一部である。このCLUSTAL Vは、Higgins, D. G. 及び P. M. Sharp (1989) CABIOS 5:151-153、Higgins, D. G. 他 (1992) CABIOS 8:189-191に記載されている。ポリ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。同一性のパーセントは、アラインメントされたポリヌクレオチド配列間の「類似性のパーセント」としてCLUSTAL Vによって報告される。

【0077】

或いは、一般的に用いられ且つ自由に入手できる配列比較アルゴリズム一式が、国立バイオテクノロジー情報センター（NCBI）Basic Local Alignment Search Tool（BLAST）から提供されており（Altschul, S. F. ら (1990) J. Mol. Biol. 215:403-410）、これは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/）を含む幾つかの情報源から入手可能である。このBLASTソフトウェア一式には、既知のポリヌクレオチド配列と様々なデータベースの別のポリヌクレオチド配列とのアラインメントに用いられる「blastn」を含む、様々な配列分析プログラムが含まれる。「BLAST 2 Sequences」と呼ばれるツールが入手可能であり、2つのヌクレオチド配列の対を直接比較するために用いられる。「BLAST 2 Sequences」は、http://www.ncbi.nlm.nih.gov/gorf/bl2.htmlにアクセスして、対話形式で利用ができる。「BLAST 2 Sequences」ツールは、blastn 及び blastp（以下に記載）の両方に用いることができる。BLASTプログラムは、一般的には、ギャップ及びデフォルト設定に設定された他のパラメータと共に用いる。例えば、2つのヌクレオチド配列を比較する場合、ある者は、デフォルトパラメータに設定された「BLAST 2 Sequences」ツールVersion 2.0.12（April-21-2000）でblastnを使用するであろう。デフォルトパラメータの設定例を以下に示す。

【0078】

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

一致率は、完全に画定された（例えば特定の配列番号で画定された）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。

【0079】

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。

【0080】

ポリペプチド配列に用いられる用語「パーセントの同一性」または「%の同一性」とは、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る2つ以上のポリペプチド配列間の一致する残基の百分率のことである。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、置換部位の酸性度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。

10

【0081】

ポリペプチド配列間の一致率は、MEGALIGN version 3.12e 配列アラインメントプログラムに組込まれているようなCLUSTAL Vアルゴリズムのデフォルトのパラメータを用いて決定できる（既に説明したのでそれを参照されたい）。CLUSTAL Vを用い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、Ktupple=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、アラインメントされたポリペプチド配列の対の同一性のパーセントは、「類似性のパーセント」としてCLUSTAL Vによって報告される。

20

【0082】

或いは、NCBI BLASTソフトウェア一式を用いてもよい。例えば、2つのポリペプチド配列を対で比較をする場合、ある者は、デフォルトパラメータで設定された「BLAST 2 Sequences」ツールVersion 2.0.12 (Apr-21-2000)でblastpを使用するであろう。デフォルトパラメータの設定例を以下に示す。

【0083】

Matrix: BLOSUM62
Open Gap: 11 及び Extension Gap: 1 penalties
Gap x drop-off: 50
Expect: 10
Word Size: 3
Filter: on

30

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

【0084】

「ヒト人工染色体（HAC）」は、約6kb（キロベース）～10MbのサイズのDNA配列を含み得る、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。

【0085】

用語「ヒト化抗体」は、もとの結合能力を保持しつつよりヒトの抗体に似せるために、非抗原結合領域のアミノ酸配列が変えられた抗体分子を指す。

【0086】

50

「ハイブリダイゼーション」とは、所定のハイブリダイゼーション条件下で、ある一本鎖ポリヌクレオチドがある相補的な一本鎖と塩基対を形成するアニーリングのプロセスである。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、洗浄過程の後に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。

【0087】

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ら（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。

【0088】

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℃、42℃の温度で行う。SSC濃度は、約0.1%のSDS存在下で、約0.1～2×SSCの範囲で変化し得る。通常は、遮断剤を用いて非特異ハイブリダイゼーションを阻止する。このような遮断剤には、例えば、約100～200μg/mlの変性サケ精子DNAがある。特定条件下で、例えばRNAとDNAのハイブリダイゼーションに有機溶剤、例えば約35～50% v/vの濃度のホルムアミドを用いることもできる。洗浄条件の有用なバリエーションは、当業者には自明であろう。ハイブリダイゼーションは、特に高ストリンジェント条件下では、ヌクレオチド間の進化的な類似性を示唆し得る。このような類似性は、ヌクレオチド及びヌクレオチドにコードされるポリペプチドに対する類似の役割を強く示唆している。

【0089】

用語「ハイブリダイゼーション複合体」は、相補的な塩基間の水素結合によって、形成された2つの核酸配列の複合体を指す。ハイブリダイゼーション複合体は、溶解状態で形成し得る（ C_0t または R_0t 解析等）。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。

【0090】

用語「挿入」或いは「付加」は、1個以上のアミノ酸残基或いはヌクレオチドがそれぞれ追加されるアミノ酸配列或いは核酸配列の変化を指す。

【0091】

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

【0092】

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。

【0093】

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。

【0094】

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固有の指定された位置を有する、ポリヌクレオチド、ポリペプチドまたはその他の化合物を指す。 10

【0095】

用語「調節」は、PKINの活性の変化を指す。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。

【0096】

「核酸」及び「核酸配列」の語は、ヌクレオチド、オリゴヌクレオチド、ポリヌクレオチドまたはこれらの断片を指す。「核酸」及び「核酸配列」の語は、ゲノム起源または合成起源のDNAまたはRNAであって一本鎖または二本鎖であるか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様またはRNA様物質を指すこともある。 20

【0097】

「機能的にリンクした」は、第1の核酸配列と第2の核酸配列が機能的な関係にある状態を指す。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リンクしている。同一のリーディングフレーム内で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、一般に、機能的にリンクしたDNA配列は非常に近接するか、或いは連続し得る。

【0098】

「ペプチド核酸(PNA)」は、末端がリシンで終わるアミノ酸残基のペプチドの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を含む、アンチセンス分子または抗遺伝子剤を指す。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伸長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。 30

【0099】

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まれ得る。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。

【0100】

「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、放射性アイソトープ、リガンド、化学発光試薬及び酵素がある。「プライマー」とは、相補的な塩基対を形成して標的のポリヌクレオチドにアニーリング可能な、通常はDNAオリゴヌクレオチドである短い核酸である。プライマーは次に、DNAポリメラーゼ酵素によって標的DNA鎖に延長し得る。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。 40

【0101】

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 50

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。

【0102】

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. ら (1989) Molecular Cloning: A Laboratory Manual, 第2版, 1-3巻, Cold Spring Harbor Press, Plainview NY、Ausubel, F. M. ら, (1987) Current Protocols in Molecular Biology, Greene Pub. Assoc. & Wiley-Intersciences, New York NY、Innis ら (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。

10

【0103】

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有用である。類似のプライマー選択プログラムには、拡張能力のための追加機能が組込まれている。例えば、PrimOUプライマー選択プログラム(テキサス州ダラスにあるテキサス大学南西部医療センターのゲノムセンターから一般向けに入手可能)は、メガベース配列から特定のプライマーを選択することが可能であり、従ってゲノム全体の範囲でプライマーを設計するのに有用である。Primer3プライマー選択プログラム(Whitehead Institute/MIT Center for Genome Research, Cambridge MAより入手可能)によって、ユーザーは、プライマー結合部位として避けたい配列を指定できる「非プライミングライブラリ(mispriming library)」を入力できる。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、固有であって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プライマーとして、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。

20

30

40

【0104】

本明細書における「組換え核酸」は天然の配列ではなく、2つ以上の配列の離れたセグメントを人工的に組み合わせた配列である。この人為的組合せはしばしば化学合成によって

50

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。

【0105】

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組換え核酸が発現する哺乳動物のワクチン接種に用いるもので、哺乳動物の防御免疫応答を誘導する。

10

【0106】

「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またはウイルスタンパク質と相互作用する。

【0107】

「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。

【0108】

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、窒素性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。

20

【0109】

用語「サンプル」は、その最も広い意味で用いられている。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。

【0110】

用語「特異的結合」及び「特異的に結合する」は、タンパク質若しくはペプチドと、アゴニスト、抗体、アンタゴニスト、小分子、若しくは任意の天然若しくは合成の結合組成物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ことを意味している。例えば、抗体がエピトープ「A」に対して特異的である場合、結合していない標識した「A」及び抗体を含む反応液に、エピトープAを含むポリペプチド或いは結合していない無標識の「A」が存在すると、抗体と結合する標識Aの量が減少する。

30

【0111】

用語「実質的に精製された」は、自然の環境から取り除かれてから、単離或いは分離された核酸配列或いはアミノ酸配列であって、自然に結合している組成物が少なくとも約60%除去されたものであり、好ましくは約75%以上の除去、最も好ましくは90%以上除去されたものを指す。

40

【0112】

「置換」とは、一つ以上のアミノ酸またはヌクレオチドをそれぞれ別のアミノ酸またはヌクレオチドに置き換えることである。

【0113】

用語「基板」は、任意の好適な固体或いは半固体の支持物を指し、膜及びフィルター、チップ、スライド、ウェハ、ファイバー、磁気または非磁気ビード、ゲル、チューブ、プレート、ポリマー、微小粒子、毛細管が含まれる。基質は、壁、溝、ピン、チャネル、孔等、様々な表面形態を有することができ、基質表面にはポリヌクレオチドやポリペプチドが結合する。

50

【0114】

「転写イメージ」は、所定条件下での所定時間における特定の細胞の種類または組織による集合的遺伝子発現のパターンを指す。

【0115】

「形質転換」とは、外来DNAが受容細胞に導入されるプロセスのことである。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、ウイルス感染、電気穿孔法（エレクトロポレーション）、熱ショック、リポフェクション及び微粒子銃を用いる方法がある。「形質転換された」細胞には、導入されたDNAが自律的に複製するプラスミドとして或いは宿主染色体の一部として複製可能である安定的に形質転換された細胞が含まれる。さらに、限られた時間に一時的に導入DNA若しくは導入RNAを発現する細胞も含まれる。

10

【0116】

ここで用いる「遺伝形質転換体」とは任意の有機体であり、限定するものではないが動植物を含み、有機体の1個若しくは数個の細胞が、ヒトの関与によって、例えば本技術分野でよく知られている形質転換技術によって導入された異種核酸を有する。核酸の細胞への導入は、直接または間接的に、細胞の前駆物質に導入することによって、計画的な遺伝子操作によって、例えば微量注射法によって或いは組換えウイルスの導入によって行う。遺伝子操作の語は、古典的な交雑育種或いは *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 ら（1989）等の参考文献に与えられている。

20

【0117】

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（May-07-1999）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。ある変異配列は、例えば、「対立遺伝子」変異配列（上述）または「スプライス」変異配列、「種」変異配列、「多型」変異配列と表すことができる。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。多型変異配列はまた、ポリヌクレオチド配列の1つのヌクレオチドが異なる「1塩基多型性」（SNP）も含み得る。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。

30

40

【0118】

特定のポリペプチド配列の「変異体」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（May-07-1999）を用いる *blastp* によって、ある核酸配列のある長さに対する該特定のポリペプチド配列の同一性が、少なくとも40%と決定されたポリペプチド配列のことである。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

50

%、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。

【0119】

(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。

【0120】

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ポリヌクレオチドコンセンサス配列番号 (IncyteポリヌクレオチドID) によって表示した。

10

【0121】

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および列2はそれぞれ、本発明の各ポリペプチドに対するポリペプチド配列識別番号 (Polypeptide SEQ ID NO :) およびそれに対応するIncyteポリペプチド配列番号 (Incyte Polypeptide ID) を示す。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、GwnBank相同体のアノテーションを示し、更に該当箇所には適当な引用文も示す。これらを引用することを以って本明細書の一部とする。

20

【0122】

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および列2はそれぞれ、本発明の各ポリペプチドのポリペプチド配列識別番号 (SEQ ID NO :) およびそれに対応するIncyteポリペプチド配列番号 (Incyte Polypeptide 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。

30

【0123】

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。例えば、SEQ ID NO: 12は、Basic Local Alignment Search Tool (BLAST) で決定されるようにヒトアデニル酸キナーゼ (Genbank ID g28577) と95%同一である (表2参照)。BLAST確率スコアは 2.4×10^{-112} であり、これは保存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。BLIMPS、MOTIFSおよびPROFILE 分析からのデータ、DOMO およびPRODOM データベースを用いたBLAST 分析からのデータ、PFAM データベースを用いたHMMER 分析からのデータによって、SEQ ID NO: 12 がアデニル酸キナーゼとして分類されることがさらに裏付けられる (表3参照)。別の例において、SEQ ID NO: 13 はBasic Local Alig

40

50

nment Search Tool (BLAST) によって同定されるラットの PCTAIRE 3 (GenBank ID g2257588) に 92% の同一性を有する。BLAST 確率スコアは 1.4×10^{-210} であり、これは保存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 (表 2)。PCTAIRE 1、2 および 3 は主に有糸分裂後細胞に発現される Cdc2 関連キナーゼのサブファミリーからなる。PCTAIRE 2 と PCTAIRE 3 は脳に発現される (Hirose, T. ら, (1997) Eur. J. Biochem. 249: 481-488; Okuda, T. ら, (1992) Oncogene 7: 2249-2258)。SEQ ID NO: 13 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ードメインの PFAM データベースにおいて、統計的に有意な一致を検索して決定された。(表 3 参照) BLIMPS、MOTIFS、及び PROFILESCAN 解析よりのデータは、SEQ ID NO: 13 がキナーゼである、さらに実証的な証拠を提供する。別の例において、SEQ ID NO: 18 は Basic Local Alignment Search Tool (BLAST) によって同定されたヒト細胞周期関連キナーゼ (GenBank ID g4090958) に 86% 同一である。(表 2 参照) BLAST 確率スコアは 9.2×10^{-85} であり、これは保存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: 18 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ードメインの PFAM データベースにおいて、統計的に有意な一致を検索して決定された。(表 3 参照) BLIMPS、MOTIFS、及び PROFILESCAN 解析よりのデータは、SEQ ID NO: 18 が細胞周期関連キナーゼである、さらに実証的な証拠を提供する。SEQ ID NO: 1-11 および SEQ ID NO: 14-17 は同じような方法で解析し、注釈を付けた。SEQ ID NO: 1-18 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。

【0124】

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る構築した。列 1 および列 2 はそれぞれ、本発明の各ポリヌクレオチドのポリヌクレオチド配列識別番号 (Polynucleotide SEQ ID NO:) およびそれに対応する Incyte ポリヌクレオチドコンセンサス配列番号 (Incyte Polynucleotide ID) を示す。列 3 は、塩基対における各ポリヌクレオチド配列の長さを示す。列 4 は、例えば、SEQ ID NO: 19-36 を同定するため、或いは SEQ ID NO: 19-36 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。

【0125】

表 4 の列 5 の識別番号は、特に例えば Incyte cDNA とそれに対応する cDNA ライブラリに照会し得る。例えば、6311370H1 は Incyte cDNA 配列の識別番号であり、NERD TDN03 はそれが由来する cDNA ライブラリの識別番号である。cDNA ライブラリが示されていない Incyte cDNA は、プールされている cDNA ライブラリ (例えば、70518523D1) に由来する。列 5 の識別番号は、完全長ポリヌクレオチド配列の組み立てに用いた GenBank の cDNA すなわち EST (例えば、g1860144) の識別番号の場合もある。或いは列 5 の識別番号は、ゲノム DNA の Genscan 分析により予想されるコード領域に照会し得る。例えば、GNN. g5924006_004 は、Genscan 推定コード配列の識別番号であ

って、g 5 9 2 4 0 0 6 が G e n s c a n 分析によって得られた G e n B a n k の配列の識別番号である。G e n s c a n 予想コード配列は、構築に先立ち編集することができる（実施例 4 を参照）。または列 5 の識別番号は、「エキソンスティッチング（e x o n - s t i t c h i n g）」アルゴリズムにより結び合わせた c D N A 及び G e n s c a n 予想エキソンの両方の集合に照会し得る。または、列 5 の識別番号は、「エキソンスティッチング」アルゴリズムによって c D N A および G e n s c a n 推定エキソンの両方からなる群の場合もある（実施例 5 を参照）。場合によっては、列 5 に示されている配列の範囲と重複する I n c y t e c D N A の範囲が得られ、最終的なコンセンサス配列が決定されるが、それに相当する I n c y t e c D N A の識別番号は示されていない。

【0126】

10

表 5 は、I n c y t e c D N A 配列を用いてアセンブルされた完全長ポリヌクレオチド配列のための代表的な c D N A ライブラリを示している。代表的な c D N A 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及び確認するために用いられる I n c y t e c D N A 配列によって最も頻繁に表される I n c y t e c D N A ライブラリである。c D N A 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。

【0127】

本発明はまた、P K I N の変異体も含む。好適な P K I N の変異体は、P K I N 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 P K I N アミノ酸配列に対して少なくとも約 8 0 %、或いは少なくとも約 9 0 %、更には少なくとも約 9 5 %、或いは少

20

【0128】

本発明はまた、P K I N 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、P K I N をコードする S E Q I D N O : 1 9 - 3 6 からなる一群から選択された配列を含むポリヌクレオチド配列を提供する。配列表に示した S E Q I D N O : 1 9 - 3 6 のポリヌクレオチド配列は、窒素系塩基のチミンがウラシルに置換され、糖鎖のバックボーンがデオキシリボースではなくリボースからなる等価 R N A 配列を含む。

【0129】

本発明はまた、P K I N 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、P K I N をコードするポリヌクレオチド配列と少なくとも 7 0 %、或いは少なくとも 8 5 %、更には少なくとも 9 5 %、なお更には少なくとも 9 8 % 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、S E Q I D N O : 1 9 - 3 6 からなる一群から選択された核酸配列と少なくとも 7 0 %、或いは少なくとも 8 5 %、更に少なくとも 9 5 %、更には少なくとも 9 8 % ものポリヌクレオチド配列同一性を有する S E Q I D N O : 1 9 - 3 6 からなる一群から選択された配列を含むポリヌクレオチド配列の変異配列を提供する。上記したポリヌクレオチド変異配列は何れも、P K I N 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。

30

【0130】

遺伝暗号の縮重により作り出され得る P K I N 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 P K I N 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。

40

【0131】

P K I N 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 P K I N 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含む P K I

50

N 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。

【0132】

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。 10

【0133】

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: 19-36 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。

【0134】

DNA シークエンシングの方法は当分野では公知であり、本発明のいずれの実施例も DNA シークエンシング方法を用いて実施可能である。DNA シークエンシング方法には酵素を用いることができ、例えば DNA ポリメラーゼ I のクレノウ断片、SEQUENASE (US Biochemical, Cleveland OH)、Taq ポリメラーゼ (Applied Biosystems)、熱安定性 T7 ポリメラーゼ (Amersham, Pharmacia Biotech, Piscataway NJ) を用いることができる。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 30
クラー (MJ Research, Watertown MA) 及び ABI CATALYST 800 サーマルサイクラー (Applied Biosystems) 等の装置を用いて配列の準備を自動化する。次に、ABI 373 或いは 377 DNA シークエンシングシステム (Applied Biosystems)、MEGABACE 1000 DNA シークエンシングシステム (Molecular Dynamics, Sunnyvale CA) または当分野でよく知られている他の方法を用いてシークエンシングを行う。結果として得られた配列を当分野でよく知られている種々のアルゴリズムを用いて分析する。 (Ausubel, F. M. (1997) Short Protocols in Molecular Biology, John Wiley & Sons, New York NY, unit 7. 7、Meyers, R. A. (1995) Molecular Biology and Biotechnology, Wiley VCH, New York NY, 856-853 ページ. 等を参照)。 40

【0135】

当分野で周知の PCR 法をベースにした種々の方法で、部分的なヌクレオチド配列を利用して、PKIN をコードする核酸配列を伸長し、プロモーターや調節エレメントなどの上流にある配列を検出する。例えば、使用し得る方法の 1 つである制限部位 PCR 法は、ユニバーサルプライマー及びネステッドプライマーを用いてクローニングベクター内のゲノム DNA から未知の配列を増幅する方法である (Sarkar, G. (1993) *PCR Methods Applic* 2: 318-322 等を参照)。別の方法に逆 PCR 法があり、これは広範な方向に伸長させたプライマーを用いて環状化した鋳型から未知 50

の配列を増幅する方法である。鋳型は、既知のゲノム遺伝子座及びその周辺の配列からなる制限断片から得る (Triglia, T. ら (1988) Nucleic Acids Res 16:8186等を参照)。第3の方法としてキャプチャPCR法があり、これはヒト及び酵母菌人工染色体DNAの既知の配列に隣接するDNA断片をPCR増幅する方法に参与している (Lagerstrom, M. ら (1991) PCR Methods Applic 1:111-119等を参照)。この方法では、PCRを行う前に複数の制限酵素の消化及び連結反応を用いて未知の配列領域内に組換え二本鎖配列を挿入することが可能である。また、未知の配列を検索するために用い得る別の方法については当分野で知られている (Parker, J. D. ら (1991) Nucleic Acids Res. 19:3055-3060等を参照)。更に、PCR、ネステッドプライマー及びPromoterFinder (商標) ライブラリ (Clontech, Palo Alto CA) を用いてゲノムDNAをウォーキングすることができる。この手順は、ライブラリをスクリーニングする必要がなく、イントロン/エキソン接合部を見付けるのに有用である。全てのPCRベースの方法に対して、市販されているソフトウェア、例えばOLIGO 4.06プライマー分析ソフトウェア (National Biosciences, Plymouth MN) 或いは別の好適なプログラムを用いて、長さが約22~30ヌクレオチド、GC含有率が約50%以上、温度約68℃~72℃で鋳型に対してアニーリングするようにプライマーを設計し得る。

10

【0136】

完全長cDNAをスクリーニングする際は、より大きなcDNAを含むようにサイズ選択されたライブラリを用いるのが好ましい。更に、ランダムプライマーライブラリは、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。

20

【0137】

市販の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、発光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。

30

【0138】

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。

40

【0139】

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を介した部位特異的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。

50

【0140】

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ャッフリングして、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を介する組換えを用いて遺伝子変異体のライブラリを生成するプロセスである。ライブラリはその後、その遺伝子変異体を所望の特性に同定するような選択またはスクリーニングにかける。次にこれらの好適な変異体をプールし、更に反復してDNAシャッフリング及び選択／スクリーニングを行ってもよい。従って、人工的な育種及び急速な分子の進化によって多様な遺伝子が作られる。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、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

【0141】

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. ら (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. ら (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

【0142】

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）。

40

【0143】

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、構成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

50

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源とし得る。発現の効率は、用いられる特定の宿主細胞系に好適なエンハンサーを包含することによって高めることができる（Scharf, D. ら（1994）Results Probl. Cell Differ. 20:125-162. 等を参照）。

【0144】

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ある（Sambrook, J. ら（1989）Molecular Cloning. A Laboratory Manual, Cold Spring Harbor Press, Plainview NYの4, 8, 16-17章、Ausubel, F. M. ら（1995）Current Protocols in Molecular Biology, John Wiley & Sons, New York NYの9, 13, 16章等を参照）。

【0145】

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、Bitter, G. A. ら（1987）Methods Enzymol. 153:516-544；Scorer, C. A. ら（1994）Bio/Technology 12:181-184；Engelhard, E. K. ら（1994）Proc. Natl. Acad. Sci. USA 91:3224-3227、Sandig, V. ら（1996）Hum. Gene Ther. 7:1937-1945、Takamatsu, N.（1987）EMBOJ. 6:307-311、Coruzzi, G. ら（1984）EMBOJ. 3:1671-1680、Broglie, R. ら（1984）Science 224:838-843、Winter, J. ら（1991）Results Probl. Cell Differ. 17:85-105、『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. ら（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、ポリヌクレオチド配列を標的器官、組織または細胞集団へ送達することができる（Di Nicola, M. ら（1998）Cancer Gen. Ther. 5(6):350-356、Yu, M. ら（1993）Proc. Natl. Acad. Sci. USA 90(13):6340-6344、Buller, R. M. ら（1985）Nature 317(6040):813-815；McGregor, D. P. ら（

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。

【0146】

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。ベクターの多数のクローニング部位にPKINをコードする配列をライゲーションするとlacZ遺伝子が破壊され、組換え分子を含む形質転換された細菌の同定のための比色スクリーニング法が可能となる。更にこれらのベクターは、クローニングされた配列における*in vitro*転写、ジデオキシのシーケンシング、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 (Van Heeke, G. and S. M. Schuster (1989) J. Biol. Chem. 264:5503-5509. 等を参照)。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の発現をハイレベルで誘導するベクターが使用できる。例えば、強力な誘導T5バクテリオファージプロモーターまたは誘導T7バクテリオファージプロモーターを含むベクターが使用できる。

【0147】

PKINの発現に酵母の発現系の使用が可能である。α因子、アルコールオキシダーゼ、PGHプロモーター等の構成型或いは誘導型のプロモーターを含む多数のベクターが、酵母菌サッカロミセス・セレビジエまたは*Pichia pastoris*に使用可能である。更に、このようなベクターは、発現したタンパク質を分泌或いは細胞内への保持のいずれかに誘導し、安定した増殖のために外来配列の宿主ゲノムへの組込みを可能にする (前出のAusubel (1995)、前出のBitter、前出のScorer等を参照)。

【0148】

植物系もPKINの発現に使用可能である。PKINをコードする配列の転写は、ウイルスプロモーター、例えば単独或いはTMV (Takamatsu, N. (1987) EMBO J 6:307-311) 由来のオメガリーダー配列と組み合わせて用いられるようなCaMV由来の35S及び19Sプロモーターによって促進される。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 (前出のCoruzzi、前出のBroglie、前出のWinter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。)

【0149】

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。ウイルスのゲノムの非必須のE1またはE3領域への挿入により、感染した宿主細胞にPKINを発現する生ウイルスを得ることが可能である (Logan, J. 及びShenk, T. (1984) Proc. Natl. Acad. Sci. 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。

【0150】

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。治療目的のために約6 kb～10 MbのHAC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（Harrington, J. J. ら（1997）Nat Genet. 15:345-355. 等を参照）。

【0151】

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。

【0152】

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。選択系には、以下のものに限
定はしないが、単純ヘルペスウイルスチミジンキナーゼ遺伝子及びアデニンホスホリボシ
ルトランスフェラーゼ遺伝子が含まれ、それぞれ tk^{-} または apr^{-} 細胞において使用
される。（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. ら（1980）Proc. Natl. Acad. Sci. USA 77:3567-3570、Colbere-Garapin, F.
ら（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. and 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）、 β グルクロニダーゼ及びその基質 β グルクロニド、またはル
シフェラーゼ及びその基質ルシフェリン等を用いてもよい。これらのマーカーを用いて、
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）。

【0153】

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。

【0154】

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法、核酸或いはタンパク

10

20

30

40

50

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。

【0155】

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、酵素に結合した免疫吸着剤検定法(ELISA)、ラジオイムノアッセイ(RIA)、フローサイトメーター(FACS)などがある。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アッセイ(two-site, monoclonal-based immunoassay)が好ましいが、競合の結合アッセイも用いることもできる。これらのアッセイ及びこれ以外のアッセイは、当分野で公知である(Hampton, R. ら(1990) Serological Methods, a Laboratory Manual, APS Press, St Paul, MN, Sect. IV, Coligan, J. E. ら(1997) Current Protocols in Immunology, Greene Pub. Associates and Wiley-Interscience, New York NY, Pound, J. D. (1998) Immunochemical Protocols, Humana Press, Totowa NJ等を参照)。

【0156】

多岐にわたる標識方法及び抱合方法が、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。

【0157】

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。

【0158】

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」形を切断する翻訳後のプロセッシングを利用して、標的タンパク質、折りたたみ及び／または活性を特定することが可能である。翻訳後の活性のための固有の細胞装置及び特徴のある機構を有する種々の宿主細胞(例えばCHO、HeLa、MDCK、MEK293、WI38等)は、American Type Culture Collection(ATCC, Bethesda

， V A）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。

【0159】

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）、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。

【0160】

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カップルさせる。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。

【0161】

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。

【0162】

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（Colligan, 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、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。

【0163】

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。

【0164】

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。

10

【0165】

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を「ノックアウト」する。このような技術は当技術分野において周知であり、ヒト疾患動物モデルの作製に有用である（米国特許第5,175,383号及び第5,767,337号等を参照）。例えば129/SvJ細胞株等のマウスES細胞は初期のマウス胚に由来し、培地で増殖させることができる。このES細胞は、ネオマイシンホスホトランスフェラーゼ遺伝子（neo:Capecchi, M. R. (1989) Science 244:1288-1292）等のマーカー遺伝子で破壊した目的の遺伝子を含むベクターで形質転換する。このベクターは、相同組換えにより宿主ゲノムの対応する領域に組み込まれる。別法では、Cre-loxP系を用いて相同組換えを行い、組織特異的または発生段階特異的に目的遺伝子をノックアウトする（Marth, J. D. (1996) Clin. Invest. 97:1999-2002; Wagner, K. U. 他 (1997) 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。

20

30

【0166】

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）。

40

【0167】

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「ノックイン」ヒト化動物（ブタ）または遺伝子組換え動物（マウスまたはラット）を作製することが可能である。ノックイン技術を用いて、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) Biotechnol. Annu. Rev. 4:55-74）。

50

【0168】

(治療)

P K I N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。さらに、P K I Nの発現は癌、疾患、増殖性、心臓性、腫瘍性、消化組織、脳の変性疾患に密接に関連しており、P K I Nが中枢神経系および神経組織（例えば、脳および脊髄、表6参照）に影響を及ぼす壊死性障害に関与することを示している。したがって、P K I Nは中枢神経系の細胞の維持、またおそらく腫瘍性形質転換にも関与し、また癌、免疫疾患、成長および発達に影響する障害、心臓疾患および脂質障害に関与すると思われる。P K I N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、P K I Nの発現または活性を低下させることが望ましい。また、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の発現または活性を増大させることが望ましい。

【0169】

従って、一実施例において、P K I N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P K I N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（A I D S）、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（A P E C E D）、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（M C T D）、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、W A G R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（S m i t h - M a g e n i s s y n d r o m e）、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-トゥ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（S y n d e n h a m ' s c h o r e a）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 eosinophilic 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_M2 にガングリオシドーシス、セロイドリポフスチン症、無 β -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 α 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。

【0170】

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。

【0171】

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。

【0172】

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。

【0173】

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ティング或いは運搬機構として間接的に用いられ得る。

【0174】

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。

【0175】

別の実施例では、本発明の任意のタンパク質、アンタゴニスト、抗体、アゴニスト、相補配列、またはベクターを、別の好適な治療薬と組み合わせて投与することもできる。併用療法で用いる好適な治療薬は、当業者が従来の医薬原理に従ってを選択し得る。治療薬と

組み合わせることにより、上記した種々の疾患の治療または予防に相乗効果をもたらし得る。この方法を用いることにより少量の各薬剤で医薬効果をあげることが可能となり、それによって副作用の可能性を低減し得る。

【0176】

P K I N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P K I N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P K I Nと特異的に結合するものを同定が可能である。P K I N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、F a bフラグメント、及びF a b発現ライブラリによって作られたフラグメントが含まれる。但し、これらに限定されるものではない。中和抗体（即ち二量体の形成を阻害する抗体）は通常、治療用に好適である。

10

【0177】

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、P K I N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、ブルロニックポリオール、ポリアニオン、ペプチド、油性乳剤、キーホールリンペットヘモシニアン、ジニトロフェノール等の界面活性剤とがある。ヒトに用いられるアジュバントの中では、B C G（カルメット-ゲラン杆菌）及びコリネバクテリウム-パルヴムが特に好ましい。

20

【0178】

P K I N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。P K I Nアミノ酸の短いストレッチは、K L H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。

【0179】

P K I N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E B V-ハイブリドーマ技術がある（K o h l e r, G. ら. (1975) N a t u r e 256: 495-497、K o z b o r, D. ら (1985) . J. I m m u n o l. M e t h o d s 81: 31-42、C o t e, R. J. ら (1983) P r o c. N a t l. A c a d. S c i. U S A 80: 2026-2030、C o l e, S. P. ら (1984) M o l. C e l l B i o l. 62: 109-120等を参照）。

30

【0180】

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、M o r r i s o n, S. L. 他. (1984) P r o c. N a t l. A c a d. S c i. 81-4851-4855; N e u b e r g e r, M. S. 他. (1984) N a t u r e 312: 604-608; T a k e d a, S. ら. (1985) N a t u r e 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、P K I N特異性一本鎖抗体を生成する。関連特異性を有するがイディオタイプ組成が異なるような抗体を、ランダムな組合せの免疫グロブリンライブラリからチェーンシャッフリングによって産生することもできる（B u r t o n D. R. (1991) P r o c. N a t l. A c a d. S c i. U S A 88: 10134-10137等を参照）。

40

50

【0181】

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. ら (1989) Proc. Natl. Acad. Sci. USA 86: 3833-3837、Winter, G. ら (1991) Nature 349: 293-299等を参照)。

【0182】

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る。例えば、限定するものではないがこのような断片には、抗体分子のペプシン消化によって作製される F(ab')₂ 断片と、F(ab')₂ 断片のジスルフィド架橋を還元することによって作製される Fab 断片とがある。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 (Huse, W. D. ら (1989) Science 256: 1275-1281等を参照)。

【0183】

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形成の計測が含まれる。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、競合的結合アッセイも利用することができる (Pound、前出)。

【0184】

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医薬の 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)。

【0185】

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 $1 \sim 2$ mg/ml の特異的な抗体、好ましくは $5 \sim 10$ mg/ml の特異的な抗体を含むポリクローナル抗体医薬は一般に、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 Coligan らの文献等を参照)。

【0186】

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

10

20

30

40

50

片や相補配列が、治療目的で使うことができる。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., 編集（1996）*Antisense Therapeutics*, Humana Press Inc., Totowa NJ等を参照）。

【0187】

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送達することが可能である（Slater, J. E. ら（1998）*J. Allergy Clin. Immunol.* 102（3）：469-475及びScanlon, K. J. ら（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）。その他の遺伝子送達機構には、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. ら（1998）*J. Pharm. Sci.* 87（11）：1308-1315、Morris, M. C. ら（1997）*Nucleic Acids Res.* 25（14）：2730-2736.等を参照）。

【0188】

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. ら（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. ら（1995）*Science* 270：475-480、Bordignon, C. ら（1995）*Science* 270：470-475）、嚢胞性繊維症（Zabner, J. ら（1993）*Cell* 75：207-216；Crystal, R. G. ら（1995）*Hum. Gene Therapy* 6：643-666、Crystal, R. G. ら（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）細胞内の寄生生物（例えばヒト免疫不全ウイルス（HIV）（Baltimore, D.（1988）*Nature* 335：395-396、Poeschbla, E. ら（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等の真菌寄生菌、並びにPlasmodium falciparum及びTrypanosoma cruzi等の原生動物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。

【0189】

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. and 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。 10

【0190】

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β-アクチン遺伝子等)、(ii) 誘導性プロモーター(例えば、市販されているT-REXプラスミド(Invitrogen)に含まれている、テトラサイクリン調節性プロモーター(Gossen, M. and H. Bujard (1992) Proc. Natl. Acad. Sci. U. S. A. 89:5547-5551; Gossen, M. 他 (1995) 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。 20 30

【0191】

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. and A. J. Eb (1973) Virology 52:456-467)若しくは電気穿孔法(Neumann, B. 40 40 (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。

【0192】

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 50

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（VPCCL）において増殖され、VPCCLは、標的細胞上の受容体に対する向性を有するエンベロップ遺伝子またはVSVg等の乱交雑エンベロップタンパク質を発現する（Armentano, D. ら (1987) J. Virol. 61:1647-1650、Bender, M. A. ら (1987) J. Virol. 61:1639-1646、Adam, M. A. and A. D. Miller (1988) J. Virol. 62:3802-3806、Dull, T. ら (1998) J. Virol. 72:8463-8471、Zufferey, R. ら (1998) J. Virol. 72:9873-9880）。RIGGに付与された米国特許第5,910,434号（「Method for obtaining retrovirus packaging cell lines producing high transducing efficiency retroviral supernatant」）において、レトロウイルスパッケージング細胞系を得るための方法が開示されており、引用することをもって本明細書の一部とする。レトロウイルスベクターの増殖、細胞集団（例えばCD4⁺ T細胞）の形質導入、及び形質導入した細胞の患者への戻しは、遺伝子治療の分野では当業者に公知の方法であり、多数の文献に記載されている（Ranga, U. ら. (1997) J. Virol. 71:7020-7029、Bauer, G. ら (1997) Blood 89:2259-2267、Bonyhadi, M. L. (1997) J. Virol. 71:4707-4716、Ranga, U. ら (1998) Proc. Natl. Acad. Sci. U. S. A. 95:1201-1206、Su, L. (1997) Blood 89:2283-2290）。

【0193】

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脾臓の無損傷の脾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、Antinuzzi, P. A. ら (1999) Annu. Rev. Nutr. 19:511-544 及び Verma, I. M. and N. Somia (1997) Nature 18:389:239-242も参照されたい。両文献は、引用することをもって本明細書の一部とする。

【0194】

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、当業者に公知である。複製適格性単純ヘルペスウイルス（HSV）I型系のベクターは、レポーター遺伝子を霊長類の眼に送達するために用いられてきた（Liu, X. ら (1999) Exp. Eye Res. 169:385-395）。HSV-1ウイルスベクターの作製についても、DeLucaに付与された米国特許第5,804,413号（「Herpes simplex virus swains for gene transfer」）が開示されており、該特許の引用をもって本明細書の一部とする。米国特

10

20

30

40

50

許第5, 804, 413号には、ヒト遺伝子治療を含む目的のために好適なプロモーターの制御下において細胞に導入される少なくとも1つの外在性遺伝子を有するゲノムを含む組換えHSV d92についての記載がある。上記特許はまた、ICP4、ICP27及びICP22のために除去される組換えHSV系統の作製及び使用について開示している。HSVベクターについては、Goins, W. F. ら (1999) J. Virol. 73:519-532 及び Xu, H. ら (1994) Dev. Biol. 163:152-161も参照されたい。両文献は、引用をもって本明細書の一部とする。クローン化ヘルペスウイルス配列の操作、巨大ヘルペスウイルスのゲノムの異なった部分を含む多数のプラスミドを形質移入した後の組換えウイルスの産生、ヘルペスウイルスの成長及び増殖、並びにヘルペスウイルスの細胞への感染は、当業者に公知の技術である。 10

【0195】

別法では、αウイルス（正の一本鎖RNAウイルス）ベクターを用いて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αウイルスであるセムリキ森林熱ウイルス（Semliki Forest Virus, SFV）の生物学的研究が広範に行われており、遺伝子導入ベクターがSFVゲノムに基づいていることが分かった（Garoff, H. and K.-J. Li (1998) Curr. Opin. Biotech. 9:464-469）。αウイルスRNAの複製中に、通常はウイルスのキャプシッドタンパク質をコードするサブゲノムRNAが作り出される。このサブゲノムRNAは、完全長のゲノムRNAより高いレベルに複製されるため、酵素活性（例えばプロテアーゼ及びポリメラーゼ）を有するウイルスタンパク質に比べてキャプシッドタンパク質が過剰産生される。同様に、PKINをコードする配列をαウイ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αウイルスの感染が数日以内での細胞溶解に関係する一方で、シンドビスウイルス（SIN）の変異体を有するハムスター正常腎臓細胞（BHK-21）の持続的な感染を確立する能力は、αウイルスの溶解複製を遺伝子治療に適用できるように好適に変更可能であることを示唆している（Dryga, S. A. ら. (1997) Virology 228:74-83）。様々な宿主にαウイルスを導入できることから、様々なタイプの細胞にPKINを導入することできる。或る集団におけるサブセットの細胞の特定形質導入は、形質導入前に細胞の選別を必要とし得る。αウイルスの感染性cDNAクローンの処置方法、αウイルスのcDNA及びRNAの形質移入方法及びαウイルスの感染方法は、当業者に公知である。 20 30

【0196】

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. ら (1994) in: Huber, B. E. and 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の翻訳を阻止するべく設計することができる。 40

【0197】

リボザイムは酵素性RNA分子であり、RNAの特異的切断を触媒するためにリボザイムを用いることもできる。リボザイム作用のメカニズムは、ヌクレオチド鎖切断に先立つ相補的標的RNAへのリボザイム分子の配列特異性ハイブリダイゼーションに参与している。例えば、PKINをコードする配列のヌクレオチド鎖切断を、特異的且つ効果的に触媒 50

する組換え型のハンマーヘッド型リボザイム分子が含まれる。

【0198】

任意のRNA標的内の特異的リボザイム切断部位は、GUA、GUU、GUC配列を含めたりリボザイム切断部位に対する標的分子をスキャンすることによって先ず同定される。一度同定されると、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。

【0199】

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を構成的或いは誘導的に合成するようなこれらcDNA産物を、細胞系、細胞または組織内に導入することができる。

【0200】

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

【0201】

本発明の更なる実施例は、PKIN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、PKIN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、PKIN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、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、PKIN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。

【0202】

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。PKINをコードするポリヌクレオチドを含むサンプルは、少なくとも1つの試験化合物に曝露して得る。サンプルには例えば、無傷細胞、透過化処理した細胞、無細胞再構成系または再構成生

10

20

30

40

50

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

10

【0203】

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

20

【0204】

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。

【0205】

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。

30

【0206】

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。

40

【0207】

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

50

【0208】

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。

【0209】

P K I Nまたはその断片を含む高分子を直接細胞内に送達するべく、特殊な形態に組成物が調製されるのが好ましい。例えば、細胞不透過性高分子を含むリポソーム製剤は、細胞融合及び高分子の細胞内送達を促進し得る。別法では、P K I Nまたはその断片をH I V T a t - 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 (S c h w a r z e , S . R . ら (1999) S c i e n c e 10 285:1569-1572)。

【0210】

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

【0211】

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P K I Nまたはその断片、P K I Nの抗体、P K I N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。薬用有効度及び毒性は、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E D₅₀ (集団の50%の医薬的有效量) またはL D₅₀ (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。毒性効果の治療効果に対する投与量の比は、治療指数であり、L D₅₀ / E D₅₀ 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、E D₅₀ 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。

【0212】

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、患者の通常健康状態、患者の年齢、体重及び性別、投与の時間及び頻度、薬剤の配合、反応感受性及び治療に対する応答等を考慮する。作用期間が長い成分は、特定の製剤の半減期及びクリアランス率によって3～4日毎に1度、1週間に1度、或いは2週間に1度の間隔で投与し得る。

【0213】

通常薬用量は投与の経路によって異なるが、約0.1～100,000 µgまでの最大約1グラムまでである。特定の投与量及び送達方法に関するガイダンスは文献に記載されており、現場の医者は通常それを利用することができる。当業者は、タンパク質またはインヒビターに対する処方とは異なる、ヌクレオチドに対する処方を利用することになる。同様に、ポリヌクレオチドまたはポリペプチドの送達は、特定の細胞、状態、位置等に特異的なものとなる。

【0214】

(診断)

別の実施例では、P K I Nに特異的に結合する抗体が、P K I Nの発現によって特徴付けられる疾患の診断、またはP K I NやP K I N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。P K I Nの

10

20

30

40

50

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。抗体は、修飾して或いは修飾しないで使用し、レポーター分子の共有結合性或いは非共有結合性の接着によって標識化し得る。多様なレポーター分子が本技術分野で知られており、それらを用いることができる。幾つかのレポーター分子については上記した。

【0215】

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。

10

【0216】

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。用いられ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。

20

【0217】

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ェントは、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。

【0218】

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: 19-36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。

30

【0219】

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。

40

【0220】

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。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、

50

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、脾臓、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（A I D S）及びアジソン病、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（A P E C E D）、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、脾炎、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、原発性血小板血症、血小板減少症、潰瘍性大腸炎、ウェルナー症候群、癌合併症、血液透析、体外循環、ウイルス感染症、細菌感染症、真菌感染症、寄生虫感染症、原虫感染症、蠕虫感染症、外傷が含まれ、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（M C T D）、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、脾臓、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、W A G R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（S m i t h - M a g e n i s s y n d r o m e）、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-トゥ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（S y n d e n h a m ' s c h o r e a）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、僧帽弁輪石灰化、僧帽弁脱出、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（p u l m o n a r y e o s i n o p h i l i a b r o n c h i o l i t i s o b l i t e r a n s - o r g a n i z i n g p n e u m o n i a）、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、肺併発膠原血管病併発性肺疾患、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、放射線による肺疾患、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_M2にガングリオシドーシス、セロイドリポフスチン症、無β-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 α リポ蛋白血症、甲状腺症機能低下症、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、サザン法やノーザン法、ドットブロット法、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に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。このような定性方法または定量方法は、当分野で公知である。

10

【0221】

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。

20

【0222】

PKIN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、PKIN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。

30

【0223】

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。

【0224】

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。

40

【0225】

PKINをコードする配列から設計されたオリゴヌクレオチドのさらなる診断への利用には、PCRの利用が含まれ得る。これらのオリゴマーは、化学的に合成するか、酵素により生産するか、或いは *in vitro* で産出し得る。オリゴマーは、好ましくはPKINをコードするポリヌクレオチドの断片、或いはPKINをコードするポリヌクレオチドと相補的なポリヌクレオチドの断片を含み、最適な条件下で、特定の遺伝子や条件を識別するために利用される。また、オリゴマーは、やや緩いストリンジェント条件下で、近縁のDNA或いはRNA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

50

。

【0226】

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

【0227】

PKINの発現を定量するために用いられ得る方法には、ヌクレオチドの放射標識或いはビオチン標識、調節核酸の相互増幅 (coamplification)、及び標準的な曲線に結果が加えられたものが含まれる (例えば、Melby, P. C. ら (1993) J. Immunol. Methods, 159: 235-44; Duplaa, C. ら (1993) Anal. Biochem. 229-236を参照)。目的のオリゴマーが種々の希釈液中に存在し、分光光度法または非色応答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。

【0228】

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

【0229】

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。

【0230】

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、所定の条件下で所定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される（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ィールを提供し得る。

10

【0231】

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。

【0232】

本発明のポリヌクレオチドの発現のプロフィールを作製する転写イメージはまた、工業的または天然の環境化合物の毒性試験のみならず、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. ら、（1999）Mol. Carcinog. 24:153-159、Steiner, S. and N. L. Anderson（2000）Toxicol. Lett. 112-113:467-471、該文献は特に引用することを以って本明細書の一部となす）。試験化合物が、既知の毒性を有する化合物のサインと同一のサインを有する場合には、毒性特性を共有している可能性がある。フィンガープリントまたはサインは、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サインを提供する。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。毒性サインの要素への遺伝子機能を割り当てることが毒性メカニズムの阻止に役立つが、毒性の予測につながるサイン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サインを用いて、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。

20

30

【0233】

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。

40

【0234】

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

50

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 and Anderson）。タンパク質は、通常クーマシーブルーまたはシルバーまたは蛍光染色などの物質を用いてゲルで染色することにより、分散した、独特な位置にある点としてゲル中で可視化される。各タンパク質スポットの光学密度は、通常サンプル中のタンパク質レベルに比例する。異なるサンプル、例えば試験化合物または治療薬で処理または非処理のいずれかの生物学的サンプルから得られるタンパク質スポットの光学密度を比較し、処理に関連するタンパク質スポット密度の変化を同定する。スポット内のタンパク質は、例えば化学的または酵素的切断とそれに続く質量分析を用いる標準的な方法を用いて部分的にシーケンシングする。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。

10

【0235】

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. 1999 Anal. Biochem. 270:103-111、Mendozze, L. G. 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。

20

【0236】

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. and 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。

30

【0237】

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。

40

【0238】

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。

【0239】

50

マイクロアレイは、本技術分野でよく知られている方法を用いて調製し、使用し、そして分析する (Brennan, T. M. ら (1995) の米国特許第5, 474, 796号、Schena, M. ら (1996) Proc. Natl. Acad. Sci. USA 93:10614-10619、Baldeschweiler らの (1995) PCT出願第WO95/251116号、Shalon, D. らの (1995) PCT出願第WO95/35505号、Heller, R. A. ら (1997) Proc. Natl. Acad. Sci. USA 94:2150-2155、Heller, M. J. らの (1997) 米国特許第5, 605, 662号等を参照)。様々なタイプのマイクロアレイが公知であり、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以って本明細書の一部となす。

10

【0240】

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。核酸配列は、特定の染色体、染色体の特定領域または人工形成の染色体、例えば、ヒト人工染色体 (HAC)、酵母人工染色体 (YAC)、細菌人工染色体 (BAC)、細菌P1産物、或いは単一染色体cDNAライブラリに対してマッピングされる (Harrington, J. J. ら (1997) Nat Genet. 15:345-355、Price, C. M. (1993) Blood Rev. 7:127-134、Trask, B. J. (1991) Trends Genet. 7:149-154等を参照)。一度マッピングすると、本発明の核酸配列を用いて、例えば病状の遺伝と特定の染色体領域やまたは制限酵素断片長多型 (RFLP) の遺伝とが相関するような遺伝子連鎖地図を作成可能である (Lander, E. S. and D. Botstein (1986) Proc. Natl. Acad. Sci. USA 83:7353-7357を参照)。

20

【0241】

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相関し得る (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。

30

【0242】

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、特定のヒト染色体の数或いはアームが分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遺伝的疾患を調査する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、該領域に対する任意のマッピングにより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. ら (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。

40

【0243】

50

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。

【0244】

別の薬剤スクリーニング方法は、目的のタンパク質に対して好適な結合親和性を有する化合物を高い処理能力でスクリーニングするために用いられる（Geysen, らの（1984）PCT出願番号WO84/03564等を参照）。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。

10

【0245】

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。

【0246】

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。

20

【0247】

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

【0248】

更に詳細説明をしなくとも、当業者であれば以上の説明を以って本発明を最大限に利用できるであろう。したがって、以下に記載する特定の好適な実施例は、例示目的であって本発明を限定するものではない。

30

【0249】

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60/199,021号、第60/200,226号、第60/202,339号、第60/203,505号、第60/205,654号、第60/207,739号及び第60/208,795号は、言及することをもって本明細書の一部となす。

【0250】

（実施例）

1 cDNAライブラリの作製

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

40

【0251】

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

50

した。場合によっては、DNアーゼでRNAを処理した。殆どのライブラリでは、オリゴd(T)連結常磁性粒子(Promega)、OLIGOTEXラテックス粒子(QIAGEN, Valencia CA)またはOLIGOTEX mRNA精製キット(QIAGEN)を用いて、ポリ(A+) RNAを単離した。別法では、別のRNA単離キット、例えばPOLY(A) PURE mRNA精製キット(Ambion, Austin TX)を用いて組織溶解物からRNAを直接単離した。

【0252】

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

【0253】

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しないで4℃で保管した。

【0254】

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。

【0255】

3 シークエンシング及び分析

実施例2に記載したようにプラスミドから回収したIncyte cDNAを、以下に示すようにシークエンシングした。cDNAのシークエンス反応は、標準的方法或いは高処理装置、例えばABI CATALYST 800 サーマルサイクラー(Applied 50

Biosystems) または PTC-200 サーマルサイクラー (MJ Research) を HYDRA マイクロディスペンサー (Robbins Scientific) または MICROLAB 2200 (Hamilton) 液体転移システムと併用して処理した。cDNA のシーケンス反応は、Amersham Pharmacia Biotech 社が提供する試薬または ABI シーケンシングキット、例えば ABI PRISM BIGDYE Terminator cycle sequencing ready reaction kit (Applied Biosystems) に与えられた試薬を用いて準備した。cDNA のシーケンス反応の電気泳動的分離及び標識したポリヌクレオチドの検出には、MEGABACE 1000 DNA シーケンシングシステム (Molecular Dynamics) か、標準 ABI プロトコル及び塩基対呼び出しソフトウェアを用いる ABI PRISM 373 または 377 シーケンシングシステム (Applied Biosystems) か、或いはその他の本技術分野でよく知られている配列解析システムを用いた。cDNA 配列内のリーディングフレームは、標準的方法 (前出の Ausubel, 1997, unit 7.7 に概説) を用いて決定した。cDNA 配列の幾つかを選択して、実施例 8 に記載した方法で配列を伸長させた。

10

20

30

40

【0256】

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。Incyte の cDNA 配列、またはその翻訳を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して問い合わせた (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 を参照のこと。) 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 Prosite 等のデータベース、PFAM 等の隠れマルコフモデル (HMM) ベースのタンパク質ファミリーデータベースに対する問合せによって完全長ポリペプチド配列を分析した。完全長ポリヌクレオチド配列はまた、MACDNASIS PRO ソフトウェア (日立ソフトウェアエンジニアリング, South San Francisco CA) 及び LASERGENE ソフトウェア (DNASTAR) を用いて分析した。ポリヌクレオチド及びポリペプチド配列アラインメントは、アラインメントした配列と配列の一致率も計算する MEGALIGN マルチシーケ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。

50

【0257】

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 5 の列 1 に、それらの簡単な説明を

列 2 に示す。列 3 は好適な引用文献であり、全ての文献はそっくりそのまま引用を以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）。

【0258】

完全長ポリヌクレオチド配列およびポリペプチド配列の組み立て及び分析に用いる上記のプログラムは、SEQ ID NO: 19-36 のポリヌクレオチド配列断片の同定にも利用できる。ハイブリダイゼーション及び増幅技術に有用である約 20 ～ 約 4000 ヌクレオチドの断片を表 4 の列 4 に示した。

【0259】

10

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に同定された。Genscan は、様々な生物からゲノム DNA 配列を分析する汎用遺伝子同定プログラムである（Burge, C. and 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、メチオニンから終止コドンに及ぶ構築された cDNA 配列を形成する。Genscan の出力は、ポリヌクレオチド及びポリペプチド配列の FASTA データベースである。Genscan が一度に分析する配列の最大範囲は、30 kb 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 モデルにおいてキナーゼについて問合せて分析した（7tm_1、7tm_2、7tm_3、および 7tm_4）。潜在的なヒトキナーゼが、キナーゼとしてアノテーションが付けられたインサイト cDNA 配列に対する相同性を基に同定された。こうして選択された Genscan 予測配列は、次に BLAST 分析により公共データベース gbpri 及び gbhtg と比較した。必要であれば、genpept からヒットしたトップの BLAST と比較することにより Genscan 予測配列を編集し、余分なまたは取り除かれたエキソンなどの Genscan により予測された配列のエラーを修正する。BLAST 分析はまた、任意の Incyte cDNA または Genscan 予測配列の公共 cDNA 適用範囲の発見に用いられるので、転写の証拠を提供する。Incyte cDNA 適用範囲が利用できる場合には、この情報を用いて Genscan 予測配列を修正または確認した。完全長ポリヌクレオチド配列は、実施例 3 に記載された構築プロセスを用いて、Incyte cDNA 配列及び／または公共の cDNA 配列で Genscan 予測コード配列を構築することにより得た。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。

20

30

【0260】

5 ゲノム配列データの cDNA 配列データへの構築

ステッチ配列 (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キソンを用いて伸長させた。実施例 3 に記載されたように構築された部分 cDNA は、ゲノム DNA にマッピングし、関連する cDNA 及び 1 つ若しくは複数のゲノム配列から予測された Genscan エキソンを含むクラスタに分解した。cDNA 及びゲノム情報を統合するべくグラフ理論及び動的プログラミングに基づくアルゴリズムを用いて各クラスタを分析し、引き続いて確認、編集または伸長して完全長配列を産出するような潜在的スプライス変異体を生成した。間隔全体の長さがクラスタ中の 2 以上の配列に存在するような配列を同定し、そのように同定された間隔は推移により等しいと考えられた。例えば、1 つの cDNA 及び 2 つのゲノム配列に間隔が存在する場合、3 つの間隔は全て等しいと考えられる。このプロセスは、無関係であるが連続したゲノム配列を cDNA 配列により結び合わせて架橋し得る。このようにして同定された区間を、親配列 (pare

40

50

nt sequence) に沿って現われるようにステッチアルゴリズムで縫い合わせ、可能な最も長い配列および変異配列を作製する。1種類の親配列に沿って発生した間隔と間隔との連鎖(cDNA-cDNAまたはゲノム配列-ゲノム配列)は、親の種類を変える連鎖(cDNA-ゲノム配列)に優先した。結果として得られるステッチ配列は、BLAST分析により公共データベースgenpept及びgbpriに翻訳されて比較された。Genscanにより予測された不正確なエキソンは、genpeptからヒットしたトップのBLASTと比較することにより修正した。必要な場合には、追加cDNA配列を用いるかゲノムDNAの検査により配列を更に伸長させた。

【0261】

ストレッチ配列 (Stretched Sequence)

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

【0262】

6 ポリヌクレオチドをコードするPKINの染色体マッピング

SEQ ID NO: 19-36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、Incyte LIFESEQデータベース及び公共のドメインデータベースの配列と比較した。SEQ ID NO: 19-36と一致するこれらのデータベースの配列を、Phrap(表7)などの構築アルゴリズムを使用して、連続及び重複した配列のクラスターに組み入れた。スタンフォード・ヒトゲノムセンター(SHGC)、ホワイトヘッド・ゲノム研究所(WIGR)、Genethon等の公的な情報源から入手可能な放射線ハイブリッド及び遺伝地図データを用いて、クラスタ化された配列が前もってマッピングされたかを測定した。マッピングされた配列がクラスタに含まれている結果、個々の配列番号を含めてそのクラスタの全配列が地図上の位置に割り当てられた。

【0263】

地図上の位置は、ヒト染色体の範囲または間隔として表される。センチモルガン間隔の地図上の位置は、染色体のpアームの末端に関連して測定する。(センチモルガン(cM))は、染色体マーカー間の組換え頻度に基づく計測単位である。平均して、1cMは、ヒト中のDNAの1メガベース(Mb)にほぼ等しい。尤も、この値は、組換えのホットスポット及びコールドスポットに起因して広範囲に変化する。cM距離は、配列が各クラスタ内に含まれるような放射線ハイブリッドマーカーに対して境界を提供するようなGenethonによってマッピングされた遺伝マーカーに基づく。NCBI「GeneMap99」(<http://www.ncbi.nlm.nih.gov/genemap>)などの一般個人が入手可能なヒト遺伝子マップおよびその他の情報源を用いて、上記した区間が既に同定されている疾患遺伝子マップ内若しくは近傍に位置するかを決定できる。

【0264】

このようにして、SEQ ID NO: 24 は92.30-103.1センチモルガンの区間内で染色体2にマップされ、SEQ ID NO: 25は104.8から117.9

10

20

30

40

50

センチモルガンの区間内で染色体11にマップされ、SEQ ID NO: 33は25.8から40.3センチモルガンの区間内で染色体8にマップされ、SEQ ID NO: 23は101.20から104.90センチモルガンの区間内で染色体9に、145.20から156.60センチモルガンの区間内で染色体10に、69.90から81.20センチモルガンの区間内で染色体19にマップされた。SEQ ID NO: 23については一つ以上のマップ位置が報告され、異なったマップ位置を持つ配列が、一つのクラスターに構築されたことを示す。この状態は、極めて類似しているが、完全に同一でない配列が一つのクラスターに構築される場合に発生する。

【0265】

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参与している。(前出のSambrook, 7章、同Ausubel, F. M. ら, 4章及び16章等を参照。)

【0266】

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、多数の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どうかを決定するため、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)

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

【0267】

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アセンブルされる。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、肺、筋骨格系、神経系、脾臓、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1つの生物/組織のカテゴリーに分類される。各カテゴリーのライブラリ数を数えて、全カテゴリーの総ライブラリ数で除する。同様に、各ヒト組織は、以下の疾患/病状カテゴリー即ち癌、細胞系、発達、炎症、神経性、外傷、心血管、鬱血、その他の1つに分類される。各カテゴリーのライブラリ数を数えて、全カテゴリーの総ライブラリ数で除する。得られるパーセンテージは、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。

【0268】

8 PKINをコードするポリヌクレオチドの伸長

完全長のポリヌクレオチド配列もまた、完全長分子の適切な断片から設計したオリゴヌクレオチドプライマーを用いて該断片を伸長させて生成した。一方のプライマーは既知の断片の5'伸長を開始するべく合成し、他方のプライマーは既知の断片の3'伸長を開始するべく合成した。開始プライマーの設計は、長さが約22~30ヌクレオチド、GC含有率が50%以上となり、約68~72℃の温度で標的配列にアニーリングするように、OLIGO 4.06ソフトウェア(National Biosciences)或いは別の適切なプログラムを用いて、cDNAから設計した。ヘアピン構造及びプライマー

10

20

30

40

50

プライマー二量体を生ずるようなヌクレオチドの伸長は全て回避した。

【0269】

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。

【0270】

高忠実度の増幅が、当業者によく知られている方法を利用したPCR法によって得られた。PCRは、PTC-200サーマルサイクラー(MJ Research, Inc.)を用いて96穴プレート内で行った。反応混合液は、鋳型DNA及び200 nmolの各プライマー、 Mg^{2+} と $(NH_4)_2SO_4$ と β -メルカプトエタノールを含むバッファー、Taq DNAポリメラーゼ(Amersham Pharmacia Biotech)、ELONGASE酵素(Life Technologies)、Pfu DNAポリメラーゼ(Stratagene)を含む。プライマーの組、PCI AとPCI Bに対して以下のパラメータで増幅を行った。ステップ1: 94℃で3分間 ステップ2: 94℃で15秒 ステップ3: 60度で1分間 ステップ4: 68℃で2分間 ステップ5: ステップ2、3、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 プライマー対T7、SK+に対しては、上記パラメータに代えて以下のパラメータを用いた。ステップ1: 94℃で3分間 ステップ2: 94℃で15秒 ステップ3: 57℃で1分間 ステップ4: 68℃で2分間 ステップ5: ステップ2、3、4を20回繰り返す ステップ6: 68℃で5分間 ステップ7: 4℃で保存。別法では、プライマーの組、T7とSK+に対して以下のパラメータで増幅を行った。ステップ1: 94℃で3分間 ステップ2: 94℃で15秒 ステップ3: 57℃で1分間 ステップ4: 68℃で2分間 ステップ5: ステップ2、3、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。

【0271】

各ウェルのDNA濃度は、1X TE及び0.5 μ lの希釈していないPCR産物に溶解した100 μ lのPICOGREEN定量試薬(0.25 (v/v) PICOGREEN; Molecular Probes, Eugene OR)を不透明な蛍光光度計プレート(Coming Costar, Acton MA)の各ウェルに分配してDNAが試薬と結合できるようにして測定する。サンプルの蛍光を計測してDNAの濃度を定量するべくプレートをFluoroskan II (Labsystems Oy, Helsinki, Finland)でスキャンした。反応混合物の5~10 μ l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。

【0272】

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、pUC 18ベクター(Amersham Pharmacia Biotech)への再連結反応前に音波処理またはせん断した。ショットガン・シークエンシングのために、消化したヌクレオチドを低濃度(0.6~0.8%)のアガロースゲル上で分離し、断片を切除し、寒天をAgar ACE(Promega)で消化した。伸長させたクローンをT4リガーゼ(New England Biolabs, Beverly MA)を用いてpUC 18ベクター(Amersham Pharmacia Biotech)に再連結し、Pfu DNAポリメラーゼ(Stratagene)で処理して制限部位の張出部(overhang)を満たし、大腸菌細胞に形質移入した。形質移入した細胞を選択して抗生物質を含む培地に移し、それぞれのコロニーを切りとってLB/2Xカルベニシリン培養液の384ウェルプレートに37℃で一晩培養した。

【0273】

細胞を溶解して、Taq DNAポリメラーゼ(Amersham Pharmacia

B i o t e c h) 及び P f u DNAポリメラーゼ (S t r a t a g e n e) を用いて以下の手順でDNAをPCR増幅した。ステップ1： 94℃で3分間 ステップ2： 94℃で15秒 ステップ3： 60℃で1分間 ステップ4： 72℃で2分間 ステップ5： ステップ2、3、4を29回繰り返す ステップ6： 72℃で5分間 ステップ7： 4℃で保存。上記したようにP I C O G R E E N 試薬 (M o l e c u l a r P r o b e s) でDNAを定量化した。DNAの回収率が低いサンプルは、上記と同一の条件を用いて再増幅した。サンプルは20%ジメチルスルホキシド (1 : 2 , v / v) で希釈し、D Y E N A M I C e n e r g y t r a n s f e r s e q u e n c i n g p r i m e r 及び D Y E N A M I C D I R E C T k i t (A m e r s h a m P h a r m a c i a B i o t e c h) または A B I P R I S M B I G D Y E T e r m i n a t o r c y c l e s e q u e n c i n g r e a d y r e a c t i o n k i t (A p p l i e d B i o s y s t e m s) を用いてシーケンシングした。

10

【0274】

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。

【0275】

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

S E Q I D N O : 1 9 - 3 6 から導き出されたハイブリダイゼーションプローブを用いて、cDNA、mRNA、またはゲノムDNAをスクリーニングする。約20塩基対からなるオリゴヌクレオチドの標識について特に記載するが、より大きなヌクレオチド断片に対しても事実上同一の手順が用いられる。オリゴヌクレオチドを、O L I G O 4 . 0 6 ソフトウェア (N a t i o n a l B i o s c i e n c e) 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[γ -³²P]アデノシン三リン酸 (A m e r s h a m , C h i c a g o , I L) 及びT4ポリヌクレオチドキナーゼ (D u P o n t N E N , B o s t o n M A) 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、S E P H A D E X G - 2 5 超細分子サイズ排除デキストラン ビードカラム (A m e r s h a m P h a r m a c i a B i o t e c h) を用いて十分に精製する。A s e I、B g l I I、E c o R I、P s t I、X b a I または P v u I I (D u P o n t N E N) 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。

20

30

【0276】

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 (N y t r a n P l u s , S c h l e i c h e r & S c h u e l l , D u r h a m N H) に移す。ハイブリダイゼーションは、40℃で16時間行う。非特異的シグナルを除去するため、例えば0.1×クエン酸ナトリウム食塩水及び0.5%ドデシル硫酸ナトリウムに一致する条件下で、プロットを室温で順次洗浄する。オートラジオグラフィーまたはそれに代わるイメージング手段を用いてハイブリダイゼーションパターンを視覚化し、比較する。

40

【0277】

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 B a l d e s c h w e i l e r 等を参照)、機械的マイクロスポッティング技術及びこれらから派生したものを用いて達成することが可能である。上記各技術において基質は、均一且つ非多孔性の固体とするべきである (S c h e n a (1 9 9 9) 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

50

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 (S ch e n a, M. ら (1995) S c i e n c e 270:467-470、S h a l l o n, D. ら (1996) G e n o m e R e s. 6:639-645、M a r s h a l l, A. a n d J. H o d g s o n (1998) N a t. B i o t e c h n o l. 16:27-31. を参照)。

【0278】

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、レーザ脱着及び質量スペクトロメトリを用いてもハイブリダイゼーションを検出し得る。マイクロアレイ上のエレメントにハイブリダイズする各ポリヌクレオチドの相補性の度合及び相対存在度は、算定し得る。一実施例におけるマイクロアレイの調整及び使用について、以下に詳述する。

10

【0279】

組織または細胞サンプルの準備

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サンプルは、MMLV逆転写酵素、0.05pg/μlのオリゴ(dT)プライマー(21mer)、1×第一鎖合成バッファー、0.03unit/μlのRNアーゼ阻害因子、500μMのdATP、500μMのdGTP、500μMのdTTP、40μMのdCTP、40μMのdCTP-Cy3(BDS)またはdCTP-Cy5(Amersham Pharmacia Biotech)を用いて逆転写する。逆転写反応は、GEMBR IGH Tキット(Incyte)を用いてポリ(A)⁺RNA含有の25体積ml内で行う。特異制御ポリ(A)⁺RNAは、370℃で2時間インキュベートした後、i n v i t r o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、850℃で20分間インキュベートし、反応を停止させてRNAを分解させる。サンプルは、2つの連続するCHROMA SPIN 30ゲル濾過スピンカラム(CLONTECH Laboratories, Inc. (CLONTECH), Palo Alto CA)を用いて精製する。結合後、2つの反応サンプルは、1mlのグリコーゲン(1mg/ml)、60mlの酢酸ナトリウム、及び300mlの100%エタノールでエタノール沈殿させる。サンプルは次に、SpeedVAC(Savant Instruments Inc., Holbrook NY)を用いて乾燥して仕上げ、14μl 5×SSC/0.2% SDS中で再懸濁する。

20

30

【0280】

マイクロアレイの準備

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によって、1~2ngの初期量から5μgを超える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。

40

【0281】

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

50

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 °C の天火で硬化させる。

【0282】

米国特許第5,807,522号に記載されている方法を用いて、コーティングしたガラス基板にアレイエレメントを付加する。この特許に引用することを以て本明細書の一部とする。平均濃度が100 ng/μlのアレイエレメントDNA 1 μlを高速機械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5 nlのアレイエレメントサンプルを加える。

10

【0283】

マイクロアレイには、STRATALINKER UV架橋剤 (Stratagene) を用いてUV架橋する。マイクロアレイは、室温において0.2 % SDSで1度洗浄し、蒸留水で3度洗浄する。リン酸緩衝生理食塩水 (PBS) (Tropix, Inc., Bedford MA) 中の0.2 % カゼイン中において60 °C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。

【0284】

20

ハイブリダイゼーション

ハイブリダイゼーション反応液は、5 × SSC、0.2 % SDSハイブリダイゼーション緩衝液にCy3及びCy5標識したcDNA合成産物を各0.2 μg含む9 μlのサンプル混合体を含めたものである。サンプル混合体は、65 °Cまで5分間加熱し、マイクロアレイ表面上で等分して1.8 cm² 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140 μlの5 × SSCを加えることにより、チェンバー内部を湿度100 %に保持する。アレイを含むチェンバーは、60 °C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 °Cで10分間洗浄し、第2洗浄緩衝液中 (0.1 × SSC) において45 °Cで10分間各々3度洗浄して乾燥させる。

30

【0285】

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488 nm、Cy3の励起のためには632 nmでスペクトル線が発生し得るInnova 70混合ガス10 Wレーザ (Coherent, Inc., Santa Clara CA) を備えた顕微鏡で検出する。20 × 顕微鏡対物レンズ (Nikon, Inc., Melville NY) を用いて、アレイ上に励起レーザ光の焦点を当てる。アレイを含むスライドを顕微鏡のコンピュータ制御のX-Yステージに置き、対物レンズを通過してラスタースキャンする。本実施例で用いた1.8 cm × 1.8 cmのアレイは、20 μmの解像度でスキャンした。

40

【0286】

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つの蛍光体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565 nm、Cy5では650 nm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。

【0287】

50

スキヤンの感度は通常、既知濃度のサンプル混合体に添加されるcDNA対照種により生成されるシグナル強度を用いて較正する。アレイ上の特定の位置には相補的DNA配列が含まれ、その位置におけるシグナルの強度をハイブリダイジング種の重量比1:100,000に相関させる。異なる源泉（例えば試験される細胞及び対照細胞など）からの2つのサンプルを、各々異なる蛍光色素で標識し、他と異なって発現した遺伝子を同定するために単一のアレイにハイブリダイズする場合には、その較正を、2つの蛍光色素で較正するcDNAのサンプルを標識し、ハイブリダイゼーション混合体に各々等量を加えることによって行う。

【0288】

光電子増倍管の出力は、IBMコンパチブルPCコンピュータにインストールされた12ビットRTI-835Hアナログ-ディジタル(A/D)変換ボード(Analog Devices, Inc., Norwood MA)を用いてディジタル化される。ディジタル化されたデータは、青色(低シグナル)から赤色(高シグナル)までの擬似カラー範囲へのリニア20色変換を用いてシグナル強度がマッピングされたようなイメージとして表示される。データは、定量的にも分析される。2つの異なる蛍光色素を同時に励起及び測定する場合には、各蛍光体の発光スペクトルを用いて、データは先ず蛍光色素間の光学的クロストーク(発光スペクトルの重なりに起因する)を補正する。

【0289】

グリッドが蛍光シグナルイメージ上に重ねられ、それによって各スポットからのシグナルはグリッドの各エレメントに集められる。各エレメント内の蛍光シグナルは統合され、シグナルの平均強度に応じた数値が得られる。シグナル分析に用いるソフトウェアは、GEMTOOLS遺伝子発現分析プログラム(Incyte)である。

【0290】

1.1 相補的ポリヌクレオチド

PKIN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PKIN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。Oligo4.06ソフトウェア(National Biosciences)及びPKIN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PKINをコードする転写物に結合するのを阻害する。

【0291】

1.2 PKINの発現

PKIN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PKINが発現するために、抗生物質耐性及びcDNAの転写レベルを高める誘導性のプロモーターを含む好適なベクターにcDNAをサブクローニングする。このようなプロモーターには、lacオペレーター調節エレメントに関連するT5またはT7バクテリオファージプロモーター及びtrp-lac(tac)ハイブリッドプロモーターが含まれるが、これらに限定するものではない。組換えベクターを、BL21(DE3)等の好適な細菌宿主に形質転換する。抗生物質耐性をもつ細菌が、イソプロピルβ-Dチオガラクトピラノシド(IPTG)で誘発されるとPKINを発現する。真核細胞でのPKIN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、PKINをコードするcDNAと置換する。ウイルスの感染力は維持され、強い多角体プロモーターによって高いレベルのcDNAの転写が行われる。組換えバキュロウイルスは、多くの場合はSpodoptera frugiperda(Sf9)昆虫細胞に

感染に用いられるが、ヒト肝細胞の感染にも用いられることもある。後者の感染の場合は、バキュロウイルスの更なる遺伝的変更が必要になる。(Engelhard, E. K. ら (1994) Proc. Natl. Acad. Sci. USA 91:3224-3227, Sandig, V. ら (1996) Hum. Gene Ther. 7:1937-1945. 等を参照)。

【0292】

殆どの発現系では、PKINが、例えばグルタチオンSトランスフェラーゼ(GST)、またはFLAGや6-Hisなどのペプチドエピトープ標識で合成された融合タンパク質となるため、未精製の細胞溶解物からの組換え融合タンパク質の親和性ベースの精製が素早く1回で行うことができる。GSTは日本住血吸虫からの26kDaの酵素であり、タンパク質の活性及び抗原性を維持した状態で、固定化グルタチオン上で融合タンパク質の精製を可能とする(Amersham Pharmacia Biotech)。精製の後、GST部分を特定の操作部位でPKINからタンパク分解的に切断できる。FLAGは8アミノ酸のペプチドであり、市販されているモノクローナル及びポリクローナル抗FLAG抗体(Eastman Kodak)を用いて免疫親和性精製を可能にする。6ヒスチジン残基が連続して伸長した6-Hisは、金属キレート樹脂(QIAGEN)上での精製を可能にする。タンパク質の発現及び精製の方法は、前出のAusubel(1995)10章、16章に記載されている。これらの方法で精製したPKINを直接用いて以下の実施例16、17、及び18のアッセイを行うことができる。

【0293】

13 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)が含まれ、どちらもサイトメガロウイルスプロモーターを有する。リポソーム製剤或いは電気穿孔法を用いて、5~10µgの組換えベクターをヒト細胞株(例えば内皮由来か造血由来のヒト細胞株)に一過的に形質移入する。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。

【0294】

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 I g G か C D 6 4 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 (D Y N A L . L a k e S u c c e s s . N Y) 。 m R N A 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。 P K I N 及び目的の他の遺伝子をコードする m R N A 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。

【0295】

14 P K I N 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 (P A G E ; 例えば、H a r r i n g t o n , M . G . (1 9 9 0) M e t h o d s E n z y m o l . 1 8 2 : 4 8 8 - 4 9 5 を参照) または他の精製技術で実質的に精製された P K I N 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。 10

【0296】

別法では、P K I N アミノ酸配列を L A S E R G E N E ソフトウェア (D N A S T A R) 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 A u s u b e l , 1 9 9 5 , 1 1 章を参照)。

【0297】

通常は、長さ約 1 5 残基のオリゴペプチドを、F m o c ケミストリを用いる A B I 4 3 1 A ペプチドシンセサイザ (A p p l i e d B i o s y s t e m s) 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 (M B S) を用いた反応によって K L H (S i g m a - A l d r i c h , S t . L o u i s M O) に結合させて、免疫抗原性を高める (前出の A u s u b e l , 1 9 9 5 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-K L M 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 P K I N 活性を検査するには、ペプチドまたは P K I N を基板に結合し、1 % B S A 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 I g G と反応させる。 20

【0298】

15 特異的な抗体を用いる天然 P K I N の精製

天然 P K I N 或いは組換え P K I N を、P K I N 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ー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、C N B r - 活性化 S E P H A R O S E (A m e r s h a m P h a r m a c i a B i o t e c h) のような活性化クロマトグラフィー用レジンと抗 P K I N 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。 30

【0299】

P K I N を含む培養液をイムノアフィニティークラムに通し、P K I N 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 P K I N との結合を切るような条件で (例えば、p H 2 ~ 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、P K I N を回収する。 40

【0300】

16 P K I N と相互作用する分子の同定

P K I N または生物学的に活性なその断片を、¹²⁵I ボルトンハンター試薬 (例えば、B o l t o n A . E . 及び W . M . H u n t e r (1 9 7 3) B i o c h e m . J . 1 3 3 : 5 2 9 を参照) で標識する。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 P K I N と共にインキュベートし、洗浄して、標識した P K I N 複合体を有する全てのウェルをアッセイする。様々な P K I N 濃度で得られたデータを用いて、候補分子と結合した P K I N の数量及び親和性、会合についての値を計算する。

【0301】

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。

【0302】

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)。

10

【0303】

17 PKINの活性の実証

通常、プロテインキナーゼ活性は、 γ -標識された $^3\text{ }^2\text{P}$ -ATPの存在で、PKINによりタンパク質基質のリン酸エステル化を定量化することで測定される。PKINはタンパク質基質、 $^3\text{ }^2\text{P}$ -ATP、及び適切なキナーゼバッファと共にインキュベートされる。基質に組み込まれた $^3\text{ }^2\text{P}$ は、電気泳動法で遊離 $^3\text{ }^2\text{P}$ -ATPより分離され、組み込まれた $^3\text{ }^2\text{P}$ はラジオアイソトープカウンタでカウントする。組み込まれた $^3\text{ }^2\text{P}$ 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解析によってなされる。

20

【0304】

或る実施態様では、プロテインキナーゼ活性はガンマ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ガンマ $^3\text{ }^2\text{P}$ -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 $^3\text{ }^2\text{P}$ 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、遊離のガンマ $^3\text{ }^2\text{P}$ -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 $^3\text{ }^2\text{P}$ 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カウントされる。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である (Pearson, J. D. ら、(1991) Methods Enzymol. 200: 62-81)。

30

【0305】

別の実施態様では、PKINのタンパク質活性は、PKIN、50 μl 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1 μg の基質、1mMのDTT、10 μg のATP、及び0.5 μCi の [γ - $^3\text{ }^2\text{P}$] ATPを含むアッセイに於いて*in vitro*で示される。反応液を30°C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 [γ - $^3\text{ }^2\text{P}$] 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°C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。組み込まれた放射活性はPKINの不在下、若しくは不活性キナーゼK38Aの存在下で実行される反応のために修正されても良い。

40

50

【0306】

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、 γ ラジオアイソトープカウンタを用いて、 γ 標識された $^3\text{ }^2\text{ p-A T P}$ の A D P 若しくは G D P 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（ A M P 若しくは G M P ）及びリン酸塩ドナーとしての $^3\text{ }^2\text{ p}$ 標識 A T P と共にインキュベートされる。反応溶液は 37°C 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、カウントされる。回復した放射性活性は酵素の活性に比例する。

【0307】

さらに別の実施例では、 P K I N 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（ S P A ）、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。有用な基質はグルタチオントランスフェラーゼで標識された組みかえタンパク質を含み、ビオチンで標識された合成ペプチド基質を含む。小さな有機分子、タンパク質、及びペプチドのような P K I N の活性の阻害剤は、そのようなアッセイで識別される。

【0308】

18 プロテインキナーゼ活性の強化／抑制

P K I N 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、 P K I N 活性の増加を引き起こし、アンタゴニストは P K I N 活性の減少を引き起こす。

【0309】

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。

【0310】

（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。

【0311】

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

【0312】

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。

【0313】

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 c D N A 及びゲノム D N A 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。

【0314】

表5は、本発明のポリヌクレオチドの代表的な c D N A ライブラリを示す。

【0315】

表6は、表5に示した c D N A 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。

【0316】

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

10

20

30

40

50

【表 1】

表 1

Incyte エクト ID	ポリペプチド SEQ ID NO:	ポリペプチド ID	ポリヌクレオチド SEQ ID NO:	Incyte ポリヌクレオチド ID
2890544	1	2890544CD1	19	2890544CB1
7472693	2	7472693CD1	20	7472693CB1
3107952	3	3107952CD1	21	3107952CB1
5544420	4	5544420CD1	22	5544420CB1
7472832	5	7472832CD1	23	7472832CB1
1551456	6	1551456CD1	24	1551456CB1
2589355	7	2589355CD1	25	2589355CB1
4357117	8	4357117CD1	26	4357117CB1
5511992	9	5511992CD1	27	5511992CB1
7474560	10	7474560CD1	28	7474560CB1
7474602	11	7474602CD1	29	7474602CB1
7475509	12	7475509CD1	30	7475509CB1
7475491	13	7475491CD1	31	7475491CB1
2192119	14	2192119CD1	32	2192119CB1
7474496	15	7474496CD1	33	7474496CB1
1834248	16	1834248CD1	34	1834248CB1
71584520	17	71584520CD1	35	71584520CB1
7475538	18	7475538CD1	36	7475538CB1

【表 2】

10

20

30

40

表 2-1

ポリペプチド SEQ ID NO:	Incyte ポリペプチド ID	GenBank ID NO:	確率スコア	GenBank 相同志体
1	2890544CD1	g5305331	0	ブロテインキナーゼ Myak-L [Mus musculus]
2	7472693CD1	g790790	2.4e-144	cam キナーゼ I [Homo sapiens]
3	3107952CD1	g1403532	2.3e-226	KIS ブロテインキナーゼ [Rattus norvegicus] (Maucuer, A. ら (1997) J. Biol. Chem. 272: 23151-23156)
4	5544420CD1	g205278	6.4e-293	男性生殖細胞関連キナーゼ [Rattus norvegicus] (Matsushime, H. ら (1990) Mol. Cell. Biol. 10: 2261-2268)
5	7472832CD1	g438373	0.0	ブロテインキナーゼ C mu [Homo sapiens] (Johannes, F. J. ら (1994) J. Biol. Chem. 269: 6140-6148)
6	1551456CD1	g4099088	3.8e-26	[Arabidopsis thaliana] SNF1 ファミリー ブロテインキナーゼ
7	2589355CD1	g6760436	1.2e-144	[Gallus gallus] qin 誘導キナーゼ
8	4357117CD1	g6552404	2.4e-199	[Rattus norvegicus] DLG6 アルファ
9	5511992CD1	g971420	3.9e-231	混合系キナーゼ 2 [Homo sapiens]
9	5511992CD1	g12005724	0	Dorow, D. S. ら (1995) Eur. J. Biochem. 234: 492-500 [5' incom] [Homo sapiens] 混合系キナーゼ MLK1
10	7474560CD1	g4691541	7.3e-102	アデニル酸キナーゼ 5 [Homo sapiens] Van Rompay, A. R. ら (1999) Identification of a novel human adenylylate kinase cDNA cloning, expression analysis, chromosome localization and characterization of the recombinant protein, Eur. J. Biochem. 261: 509-516.

10

20

30

40

表 2-2

ポリペプチド SEQ ID NO:	Incyte ポリペプチド ID	GenBank ID NO:	確率スコア	GenBank 相同志体
11	7474602CD1	g439614	8.9e-145	CaM-様プロテインキナーゼ [Rattus norvegicus]. Cho, F.S. ら (1994) Characterization of a rat cDNA clone encoding calcium/calmodulin-dependent protein kinase I, Biochim. Biophys. Acta 1224:156-160.
12	7475509CD1	g28577	2.4e-112	ヌクレオシド・トリホスフェートアダニル酸キナーゼ [Homo sapiens] Xu, G. ら (1992) Characterization of human adenylylate kinase 3(AK3) cDNA and mapping of the AK3 pseudogene to an intron of the NF1 gene, Genomics 13:537-542
13	7475491CD1	g2257588	1.4e-210	PCTAIRE3 [Rattus rattus]. Hirose, T. ら (1997) PCTAIRE 2, a Cdc2-related serine/threonine kinase, is predominantly expressed in terminally differentiated neurons, Eur. J. Biochem. 249:481-488.
14	2192119CD1	g3880563	2.0e-121	Genefinder を用いた予測。セリン/スレオニンキナーゼに類似、cDNA [Caenorhabditis elegans]. The C.elegans Sequencing Consortium (1998) Science 282:2012-2018.
14	2192119CD1	g10442581	0	105-kDa キナーゼ様プロテイン [Mus musculus] Liu, S.C.H., ら, (2000) Biochim. Biophys. Acta 1517:148-152
15	7474496CD1	g6066585	0.0	GCN2 eIF2 アルファキナーゼ [Mus musculus]
16	1834248CD1	g7595802	1.40E-252	ELKL モチーフキナーゼ 2 ショートフォーム [Mus musculus]

10

20

30

40

表 2-3

ポリペプチド SEQ ID NO:	Incyte ポリペプチド ID	GenBank ID NO:	確率スコア	GenBank 相団体
17	71584520CD1	g3927912	3.9e-157	カルモジュリン結合プロテインキナーゼ[Fugu rubripes] Cottage, A. ら (1999) FEBS Lett. 443: 370-374
18	7475538CD1	g4090958	9.2e-85	細胞周期関連キナーゼ [Homo sapiens]
18	7475538CD1	g9664926	0	CDK-関連プロテインキナーゼ PNQLARE [Mus musculus]

表 3-1

SEQ ID NO:	Incyte ポリペプチド ID	アミノ酸残基数	潜在的リン酸化部位	潜在的グリコシル化部位	シグネチャ配列、ドメインおよびモチーフ	分析方法及びデータベース
1	2890544CD1	1210	S20 T107 T163 S211 T422 T666 S843 S853 T907 S127 T212 T508 S29 S37 T87 S113 S169 S211 S396 T441 T474 T643 S856 S910 T912 T938 S967 Y459 T1057 S1008 S1138 S1187	N57 N111 N133 N149 N262 N471 N566 N570 N1009 N1045	プロテインキナーゼ Atp: L196-K219 プロテインキナーゼ St: L311-L323 真核生物プロテインキナーゼドメイン: Y190-P411 I492-V518 チロシンキナーゼ触媒ドメインシグネチャ PD00109B: K305-L323 ホモボックス DNA 結合セリン/トレオニンタンパク質と相互作用するプロテインキナーゼ核ホモドメイン: PD150874: G1030-L1210 プロテインキナーゼドメイン: DM00004 P14680 371-694: V192-P509 プロテインキナーゼドメイン DM00004 S57347 21-266: E24-C270 プロテインキナーゼ Atp: L29-K52 プロテインキナーゼ St: I140-Y152 シグナル切断: M1-A40 真核生物プロテインキナーゼドメイン p キナーゼ: F23-I279 プロテインキナーゼシグネチャおよびプロファイル: D120-D176 チロシンキナーゼ触媒ドメインシグネチャ PR00109A: N98-V111 PR00109B: Y134-Y152 PR00109D: V202-E224 プロテインキナーゼカルモジュリン結合性 I カルシウム/カルモジュリン依存型 CAM トランスフェラーゼセリン/トレオニンタンパク質リン酸化 PD012137: W278-L322	MOTIFS MOTIFS HMMER_PFAM BLIMPS_PRINTS BLAST_PRODOM BLAST_DOMO BLAST_DOMO MOTIFS MOTIFS SPSCAN HMMER_PFAM PROFILESSCAN BLIMPS_PRINTS BLAST_PRODOM
2	7472693CD1	357	S64 S102 T117 S154 S169 S251 S326 S343 Y116 Y133 S11 T45 T269	N225 N311 N332		

【表 6】

10

20

30

40

表 3-2

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
3	3107952CD1	419	S67 S117 S181 S215 S221 T244 S290 T390	N253	膜貫通ドメイン: V223-L241 真核生物プロテインキナーゼドメイン: W23-F304 チロシンキナーゼ触媒ドメイン PR00109B:F148-W166 PR00109D:V223-V245 PR00109E:P273-A295 セリン/トレオニンプロテインキナーゼ PD153748:S305-V344 スプライシング因子様タンパク質 PD072361:T320-G412 リボ核タンパク質反復 DM00012 A48249 299-407:P319-Y404 プロテインキナーゼドメイン DM00004 P08414 44-285:R74-I293 DM00004 P32485 24-292:S48-P289 DM00004 P49657 101-409:C108-L241	HMMER HMMER_Pfam BLIMPS_PRINTS BLAST_PRODOM BLAST_PRODOM BLAST_DOMO BLAST_DOMO

10

20

30

40

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
4	5544420CD1	624	T274 S424 S436 S496 S41 S179 S204 S315 S406 S451 S466 S500 S530 T6 S53 S161 T218 S388 S391 S420 S463 T475 T581 S587 T594 Y15 Y76 T571	N24 N252 N352 N384 N449 N455 N543 N569	真核生物プロテインキナーゼドメイン Y4-F284 チロシンキナーゼ触媒ドメイン PR00109B:F115-C133 PR00109D:T181-T203 PR00109E:A253-A275 プロテインキナーゼシグネチャおよびプロファイル V101-Q153 プロテインキナーゼ ATP 結合領域: L10-K33 セリン/トレオニンプロテインキナーゼ: F121-C133 ATP/GTP 結合部位モチーフ A(P-loop): G16-S23 キナーゼセリン/トレオニン男性生殖細胞トランスフェラーゼ ATP 結合 PD024663:Q285-R624 キナーゼセリン/トレオニン ATP 結合 II リン酸化カ ゼインアルファ鎖 PD002608:V160-F284 キナーゼトランスフェラーゼ ATP 結合セリン/トレオニンリン 化受容体プロテイン PD000001:P155-F284 プロテインキナーゼドメイン DM00004 Q04859 6-274:T6-A275 DM00004 I48733 6-274:T6-A275 DM00004 P43294 14-281:M7-A275 DM00004 Q00526 6-286:M7-F284	HMMER_Pfam BLIMPS_PRINTS PROFILES SCAN MOTIFS MOTIFS MOTIFS BLAST_PRODOM BLAST_PRODOM BLAST_PRODOM BLAST_DOMO

表 3-4

SEQ ID NO:	Incyte 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	7472832CD1	878	S197 S396 S797 S75 S111 T130 T218 S225 S333 S353 S362 T398 S408 T573 T589 T812 T853 S145 S206 T247 S387 T391 T392 T412 T434 S604 S641 S706 Y87	N187 N431 N432 N454 N473 N729	真核生物プロテインキナーゼドメイン: I551-L807 ホルボールエステル/ジアシルグリセロール結合 H139-C188 H265-C314 PH ドメイン: T398-V478 プロテインキナーゼシグネチャおよびプロファイル L650-S706 ホルボールエステル/ジアシルグリセロール結合 F151-S214 C278-A339 チロシンキナーゼ触媒ドメイン: PR00109B:H664-L682 PR00109D:L732-D754 ジアシルグリセロール/ホルボールエステル-ゼ PR00008B:C152-G161 PR00008C:Q291-C302 PR00008D:H303-L315 キナーゼホルボールエステル C MU セリン/トレオニン NPKCMU ATP 結合 PD039353:D323-V478 PD031784:V45-P138 PD027900:L807-L878 キナーゼホルボールエステル結合トランスフェラーゼ セリン/トレオニン亜鉛 ATP 結合 C 複製 PD000215:H265-C314 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 tyr.prf) : R826-A877 プロテインキナーゼドメイン DM00004 A53215 585-829:P553-V798 DM00004 I48719 591-835:P553-V798 DM04692 A37237 1-676:V556-S771 M04692 P05773 1-672:V556-P845	HMMER_PFAM HMMER_PFAM HMMER_PFAM PROFILESSCAN PROFILESSCAN BLIMPS_PRINTS BLIMPS_PRINTS BLAST_PRODOM BLAST_PRODOM BLAST_DOMO

10

20

30

40

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
					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: H139-C188 H265-C314	MOTIFS
					プロテインキナーゼ ATP 結合領域シグネチャ: L557-K580	MOTIFS
					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: I670-L682	MOTIFS

10

20

30

40

表 3-6

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
6	1551456CD1	440	S11 S175 T74 S85 S230 T297 S361 S365 S376 S396 T426 S11 S19 S131 S155 T214 S251 Y68 Y378		真核生物プロテインキナーゼドメイン: L164-E331 チロシンキナーゼシグネチャ: A192-L210, S260-S282 プロテインキナーゼドメイン DM00004 P34244 82-359: Y167-L324 プロテインキナーゼモチーフ: I198-L210	HMME-PFAM BLIMPS-PRINTS BLAST-DOMO MOTIFS
7	2589355CD1	923	Y104 Y676 T258 S355 T481 S584 T38 S125 S289 T356 S391 T447 S448 T455 T481 S500 S512 T875 S890 T17 S254 S336 T417 S472 S531 S542 S551 S564 S576 S667	N112 N317	プロテインキナーゼドメイン DM00004 P27448 58-297: G25-I259 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プロ テインキナーゼ St: I135-L147 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: S24-M268 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 tyr.prf: Y90-G167 チロシンキナーゼ触媒部位 PR00109: T93-A106, Y129-L147, L195-P217	BLAST-DOMO MOTIFS HMME-PFAM PROFILES SCAN BLIMPS-PRINTS
8	4357117CD1	442	T16 T267 T270 S295 S361 S142 S229 S178 S180 T311 S408 S437 Y294 Y304 Y346	N406	SH3 ドメインシグネチャ PR00452: C147-R159, A115-Q130, D132-I141 プロテインドメイン SH3 キナーゼグアニル酸トランスフ エラーゼ ATP 結合反復 GMP 膜 PD001338: T267- Q360 グアニル酸キナーゼ DM00755 A57653 370-570: P228-P431 PDZ シグナル伝達分子ドメイン PDZ: I3-V83 グアニル酸キナーゼ グアニル酸_kin: T268-Y372 グアニル酸キナーゼプロテイン BL00856: V264-V284, L292-R339	BLAST-DOMO HMME-PFAM HMME-PFAM BLIMPS-BLOCKS

【表 1 1】

10

20

30

40

表 3-7

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	5511992CD1	1046	T72 T112 S118 S223 S276 T294 S603 S707 S770 S781 S814 S880 T913 S531 T777 S89 T135 T363 T394 T395 T436 S599 S606 S636 T644 S808 S821 S834 T884 T919 S923 S966 S972 S983 T984 Y325	N813 N862 N964	受容体チロシンキナーゼ BL00240: E290-V337, V337-I389 受容体チロシンキナーゼ BL00239: E181-P228, L232-I254, W291-R340, L345-I389 受容体チロシンキナーゼ BL00790: I154-C207, S298-W330, L356-M404 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 tyr.pr.f: I232-T294 チロシンキナーゼ触媒部位 PR00109: M210-S223, D248-I266, G301-I311, S320-I342, C364-F386 SH3 ドメインシグネチャ PR00452: P55-A65, D69-K84, D91-N100, R102-R114 SH3 ドメイン: P55-R114 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: L134-L393 ロイシンジッパープロリンを持つキナーゼドメイン SH3 混合系 セリン/トレオニン PD024997: I396-A741 プロテインキナーゼドメイン DM00004 A53800 I19-368: L136-F386 プロテインキナーゼ ATP 結合領域シグネチャプロテインキナーゼ ATP: I140-K161 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プロ テインキナーゼ St: I254-I266	BLIMPS-BLOCKS BLIMPS-BLOCKS BLIMPS-BLOCKS PROFILES CAN BLIMPS-PRINTS BLIMPS-PRINTS HMER-PFAM HMER-PFAM BLAST-PRODOM BLAST-DOMO MOTIFS MOTIFS

表 3-8

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
10	7474560CD1	357	S23, T65, Y103, S124, T140, S142, S181, T185, S200, S211, S213, S245, T288, T304, Y316, T340, T350		キナーゼアデニル酸トランスフェラーゼ, ATP 結合 ATP/AMP トランスボスホリラーゼ, アインザイム タンパク質三次元構造ミ トコンドリア: PD000657; I175-R294. アデニル酸キナーゼ: DM00290 P00570 I-131: E165-R294. アデニル酸キナーゼタンパク質: BL00113A: F174-L190; BL00113B: H198-A241; BL00113C: G247-V261; BL00113D: S298-T328. アデニル酸キナーゼシグネチャ: PR00094A: F174-C187; PR00094B: G202-S216; PR00094C: F252-G268; PR00094D: P300-Y315; PR00094E: R317-Q331. ウリジンキナーゼシグネチャ: PR00988A: C170-C187 シキミ酸キナーゼファミリ: PR01100A: I173-E188. アデニル酸キナーゼ: I9-Q20, F252-Q263. アデニル酸キナーゼ: I11-L88, I175-Q331. アデニル酸キナーゼシグネチャ(アデニル酸キナーゼ pf): V229-A284.	BLAST-PRODOM BLAST-DOMO BLIMPS-BLOCKS BLIMPS-PRINTS BLIMPS-PRINTS MOTIFS HMMER-PFAM PROFILES SCAN
11	7474602CD1	355	S11, T45, S64, S102, Y116, T117, Y133, S154, S169, S251, T269, S326, S343	N225, N311, N332	プロテインキナーゼドメイン: DM00004 S57347 21-266: E24-C270. プロテインキナーゼ ATP 結合領域シグネチャ: L29-K52. 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: I140-Y152. 真核生物プロテインキナーゼドメイン: F23-I279. 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 tyr, prf): D120-D176. チロシンキナーゼ触媒ドメイン: PR00109A: M98-V111; PR00109B: Y134-Y152; PR00109D: V202-E224.	BLAST-DOMO MOTIFS MOTIFS HMMER-PFAM PROFILES SCAN BLIMPS-PRINTS

10

20

30

40

表 3-9

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
12	7475509CD1	224	S16, S75, T118, S124, S132, T201,		プロテインキナーゼ、カルモジュリン結合 I カルシウム/カルモジュリン依存型 CAM トランスフェラーゼ セリン/トレオニンタンパク質リン酸化 PD012137: W278-L322. アダニル酸キナーゼ: DM00290 P27144 I-125; M1-R126. アダニル酸キナーゼシグネチャ: PR00094A: V9-C22; PR00094B: G37-G51; PR00094C: W85-D101; PR00094D: Q159-Y174; PR00094E: D176-V190. シグミ酸キナーゼファミリー: PR01100A: A8-Q23; PR01100E: D107-S124 (p<0.0028). アダニル酸キナーゼモチーフ: W85-Q96. キナーゼアダニル酸トランスフェラーゼ、ATP 結合 ATP/AMP トランスホスホリラーゼアイソザイム、タンパク質三次元構造ミトコンドリア: PD000657: I10-V190. アダニル酸キナーゼ: I10-V190 アダニル酸キナーゼタンパク質: BL00113A: V9-I25; BL00113B: H33-E76; BL00113C: R80-L94; BL00113D: S132-D162. アダニル酸キナーゼシグネチャ(アダニル酸キナーゼ.prf): V64-F116. プロテインキナーゼドメイン: DM000004 Q04899 122-392: V173-A444. キナーゼセリン/トレオニンプロテイントランスフェラーゼ ATP 結合ドメイン-PCTAIRE1, PCTAIRE2, PCTAIRE3, CRK5 代替物質 PD007333: D120-T171 チロシンキナーゼ触媒ドメイン: PR00109B: Y283-I301 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 tyr.prf): D235-A317 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): Y172-F453.	BLAST-PRODOM BLAST-DOMO BLIMPS-PRINTS BLIMPS-PRINTS MOTIFS BLAST-PRODOM HMMER-PFAM BLIMPS-BLOCKS PROFILES SCAN BLAST-DOMO BLAST-PRODOM BLIMPS-PRINTS PROFILES SCAN HMMER-PFAM
13	7475491CD1	502	S12, T20, S24, S60, S64, S87, T104, S122, S126, S156, S160, Y183, T232, T239, T348, T364, T373, T387, S455, S471, S500.			

10

20

30

40

表 3-10

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
					プロテインキナーゼ ATP 結合領域シグネチャ: L178-K201 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: I289-I301	MOTIFS MOTIFS
14	2192119CD1	791	T39, T70, T85, S193, S269, T319, S342, S359, T361, T408, S462, S499, S521, S535, S539, T545, T581, T610, T617, S622, S662, S668, S677, S693, S716, S726, S737, Y184	N447	プロテインキナーゼドメイン DM00004 F52304 27-267: K69-P254 (p > 1.3e-06)	BLAST-DOMO

10

20

30

40

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
15	7474496CD1	1651	T65, T82, S108, S144, S210, T215, S216, S247, S250, T329, T346, T409, S414, T441, S450, T478, S553, S569, T581, S690, S709, T715, S722, S732, S733, S755, S757, T811, T824, S882, S917, T945, S947, S961, T993, T1022, S1038, S1059, T1062, T1182, T1234, T1238, S1397, T1418, S1437, S1457, S1572, T1475, S1547, T1548, T1604, S1641, Y256, Y821, Y840, Y1626	N100, N245, 1057, 1197, 1203, 1248, 1416, 1520, 1602	プロテインキナーゼドメイン: DM00004 P15442 645-902: S732-A994, 真核開始因子キナーゼ(EIF2 アルファ): PD156018: D13-D219 チロシンキナーゼ触媒ドメイン: PR00109B: Y840-L858 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 tyr.prf): R826-A877 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): K335-D446; F592-P664; Y799-L1003 プロテインキナーゼ ATP 結合領域シグネチャ: L598-K621 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: M846-L858 真核生物プロテインキナーゼドメイン: K335-D446; P504-I541; F592-P664; Y799-L1003 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 tyr.prf): R826-A877 チロシンキナーゼ触媒ドメイン: PR00109B: Y840-L858 プロテインキナーゼ ATP 結合領域シグネチャ: L598-K621 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: M846-L858	BLAST-DBOM BLAST-PRODOM BLIMPS-PRINTS PROFILES-SCAN HMMER-PFAM MOTIFS MOTIFS HMMER-PFAM PROFILE-SCAN BLIMPS-PRINTS MOTIFS MOTIFS

【表 1 6】

10

20

30

40

表 3-12

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
16	1834248CD1	752	S139 S2 S210 S23 S27 S34 S354 S399 S423 S441 S458 S48 S494 S661 S666 S710 T127 T281 T300 T323 T332 T344 T507 T511 T515 T536 T564 T620 T624 T81	N395 N532	真核生物プロテインキナーゼドメイン: Y59-I310. チロシンキナーゼ触媒ドメイン PR00109A: M135-V148 PR00109B: Y171-L189 PR00109D: V237-H259 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_tyr.prf: Y132-S210 プロテインキナーゼ ATP 結合領域: プロテインキナーゼ_Atp: I65-K88 セリン/トレオニンプロテインキナーゼ活性部位プロテインキナーゼ St: I177-L189 キナーゼセリン/トレオニントランスフェラーゼ ATP 結合タンパク質 EMK P78 CDC25C PD008571: S412-E632 キナーゼセリン/トレオニントランスフェラーゼ ATP 結合タンパク質 PAR1 KP78 EMK PD005838: I310-R410 キナーゼセリン/トレオニントランスフェラーゼ ATP 結合 KIN1 EMK PAR1 PD004300: E650-L752 キナーゼトランスフェラーゼ ATP 結合セリン/トレオニンリン酸化受容体チロシン PD000001: Y59-Y137 プロテインキナーゼドメイン DM000004 P27448 58-297:L61-L301 DM000004 I48609 55-294:L61-L301 DM000004 Q05512 55-294:L61-L301 DM000004 JC1446 20-261:R60-L301	HMER_PPFAM BLIMPS_PRINTS PROFILES SCAN MOTIFS MOTIFS BLAST_PROD OM BLAST_PROD OM BLAST_PROD OM BLAST_PROD OM BLAST_DOMO

【表 17】

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
17	71584520CD1	501	S118 S138 S292 S341 S364 S482 S483 S495 T103 T21 T276 T422 T46 T470 T51 T7 T91 Y135 Y491		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: E37-I286 チロシンキナーゼ触媒ドメイン PR00109B: Y135-Y153 PR00109D: V201-E223 カルモジュリン結合タンパク質 PD059862: G368-V443 プロテインキナーゼカルモジュリン結合 I カルシウム/カルモジュリン依存型 CAM トランスフェラーゼ セリン/トレオニンタンパク質リン酸化 PD012137: W285-A335 カルモジュリン結合タンパク質 PD050813: M1-E34 プロテインキナーゼドメイン DM00004 S57347 21-266:D25-T276 DM00004 P08414 44-285:I38-T276 DM00004 P11798 15-261:C36-A277 DM00004 A44412 16-262:F35-A277	HMMER_PFAM BLIMPS_PRINTS BLAST_PRODOR BLAST_PRODOR BLAST_PRODOR BLAST_DOMO

表 3-14

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	7475538CD1	346	S241 Y176 Y215		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: Y4-F288, チロシンキナーゼ触媒ドメイン PR00109B: F117-I135 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_tyr.prf: A69-D155 プロテインキナーゼ ATP 結合領域シグネチャプロテインキナーゼ ATP: I10-K33 セリン/トレオニンプロテインキナーゼ活性部位プロテインキナーゼ St: I123-I135 キナーゼトランスフェラーゼ セリン/トレオニン ATP 結合 II リン酸化カゼインアルファ鎖 PD002608: V164-F288 キナーゼ プロテイントランスフェラーゼ ATP 結合 セリン/トレオニンリン酸化受容体チロシン膜貫通 PD000001: Y169-P301 プロテインキナーゼドメイン DM000004 P29620 21-289:I10-A279 DM000004 Q00526 6-286:R9-F288 DM000004 P43450 6-276:R9-A279 DM000004 P23437 6-286:R9-F288	HMMER_PFAM BLIMPS_PRINTS PROFILES SCAN MOTIFS MOTIFS BLAST_PROD OM BLAST_PROD OM BLAST_DOMO

表 4-1

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
19	2890544CB1	4224	1-1378, 4153-4224	6474032H1 (PLACFER01) 71089659V1 71083920V1 71083254V1 71084605V1	545 2981 2936 2263 2351	1204 3638 3576 2853 2959
				7286979H1 (BRAIFER06) 71254276V1	1 1091	408 1695
				5980233F7 (MCLDXT02) 71252928V1	3498 1632	4224 2197
				7086636H1 (BRAUTDR03) 71082857V1	222 1680	702 2356
				7313223H1 (LIVRFE02) 70520498D1	498 760	928 1354
20	7472693CB1	1736	72-149, 1337-1736	609792R6 (COLNNOT01) g1544947 70518493D1	145 1 716	736 385 1182
				70518085D1	1172	1736
21	3107952CB1	1824	1802-1824	7308849H1 (WMLR1DT01) 2925973F6 (TLYMNOT04) 3459433F6 (293TF1T01) 2927552F6 (TLYMNOT04) 3107952F6 (BRSTTUT15)	1215 715 1 1309 507	1807 1215 582 1824 1161
22	5544420CB1	2201	1285-1629, 2128- 2201, 256-376	6922389H1 (PLACFER06) 7739285H1 (THYMNOE01) 5544420F6 (TESTNOG01) 4206166F6 (BRONNOT02) GNN:g5924006_004.edit GBL:g5924006.raw.comp 2512558F6 (LIVRTUT04) g2882961 6909108J1 (PITUDIR01)	1458 646 807 2041 368 1 1617 1644 1138	1597 834 1187 2201 1871 492 1720 2127 1317

10

20

30

40

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
23	7472832CB1	2974	2933-2974, 1-654	6910588J1 (PITUDIR01) 2940460H1 (THYMFET02) 7274331H2 (KIDEXJ01) 60205600U1 60205598U1 6426807H1 (LUNGNO07) 3344032H1 (SPLNNO09) 6811278J1 (SKIRNO01) 6910588H1 (PITUDIR01) 91860144 1732303F6 (BRSTTUT08)	1994 2682 1876 559 690 2411 2708 1 1369 2512 1230	2577 2956 2503 1091 1339 2951 2958 647 2003 2974 1731
24	1551456CB1	3648	1906-2052, 1-752	6830947J1 (SINTNOR01) 6834559H1 (BRSTNO02) 6808991J1 (SKIRNO01) 1389125H1 (EOSINOT01) 1680356F6 (STOMFET01) 6870264H1 (BRAGNO02) 7714085J1 (SINTFEE02) 7746056H1 (ADRETUE04) 3296183F6 (TLYJINT01) 70349297D1 70872871V1 70875814V1 6500833H1 (PROSTUS25)	2357 1610 222 1 3139 3374 298 875 1553 3030 953 2218 2498	3121 2330 921 246 3648 3648 950 1491 2310 3648 1599 2955 3226
25	2589355CB1	4719	1835-1863, 1-1135, 2380-3186	2313925T6 (NGANNO01) 6483636F9 (MIXDUNB01) 70985467V1 7251743H1 (PROSTMV01) 6864125H1 (BRAGNO02) 6056925H1 (BRAENOT04) 6705826H1 (HEAADIR01)	3264 2716 388 810 1323 1322 4080	3885 3270 980 1299 1998 1735 4719

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
				7652463H2 (STOMTDR01)	997	1704
				7039646H1 (UTRSTMR02)	1	566
				7257374H1 (SKIRTD01)	1977	2560
				70402344D1	3105	3674
				2586850F6 (BRAITUT22)	3840	4348
				924331H1 (RATRN02)	4009	4386
26	4357117CB1	1651	1-267	GS.4357117.Fasta	1	1329
				6131509F6 (BMARTX02)	1134	1651
27	5511992CB1	3141	1672-1833, 71-157, 665-828, 1298-1538, 2580- 2878	GS.5511992.Fasta	1	3141
28	7474560CB1	1244	1205-1244, 1-131	6311370H1 (NERDND03)	616	1236
				6853555H1 (BRAIFEN08)	1	683
				6997205H1 (BRAXTDR17)	691	1244
29	7474602CB1	1661	72-149	70518523D1	577	1171
				6772112J1 (BRAUNOR01)	1	640
				34627577 (THYMN02)	1043	1661
				6124350H1 (BRAHNON05)	634	1180
30	7475509CB1	912	881-912	944796H1 (RATRN02)	1	93
				3616204F6 (EPINOT01)	69	699
				71147872V1	193	912
31	7475491CB1	2858	1-269, 2768-2858	7232424H1 (BRAXTDR15)	1831	2429
				7004062H1 (COLNFEC01)	1967	2495
				7677070J1 (NOSETUR01)	1	564
				7716833J1 (SINTFE02)	1293	1877
				754239R6 (BRAITUT02)	1000	1575
				7067423H1 (BRATNOR01)	499	1106
				7428450H1 (UTRMTMR02)	2179	2858
				70680595V1	654	1111
32	2192119CB1	2817	1-249	1687833F6 (PROSTUT10)	2416	2817
				6769461J1 (BRAUNOR01)	221	385

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
				1819105F6 (PROSNOT20)	1052	1721
				7716523H1 (SINTFEE02)	498	1072
				1796441T6 (PROSTUT05)	2186	2793
				1618475T6 (BRAITUT12)	2117	2783
				7675143H1 (NOSETUE01)	1130	1739
				6145288H1 (BRANDIT03)	1543	2170
				3189653H1 (THYNNON04)	1855	2196
				7646563H1 (UTRSTUE01)	367	1049
				96700560	1	393
33	7474496CB1	5305	3463-3753, 1403- 1781, 2384-2464, 1-644, 5126-5305	70886570V1	3332	3875
33				429360R6 (BLADNOT01)	2735	3311
				70885734V1	4100	4744
				1830377F6 (THP1AZT01)	3649	4093
				6488464H1 (MIXDUNB01)	2852	3473
				488190R6 (HNT2AGT01)	2317	2865
				70888476V1	4651	5305
				1832566R6 (BRATNON01)	4034	4547
34	1834248CB1	3269	1754-1773, 1-142, 3072-3269	GNN.g7139831_000025_004	1293	2275
				6146293H1 (BRANDIT03)	1791	2354
				6272238H2 (BRAIFEN03)	2635	3263
				6893004J1 (BRAITDR03)	684	1248
				7663341J1 (UTRSTME01)	873	1492
				4001427R6 (HNT2AZS07)	2927	3264
				60202068B1	2368	2916
				60202069B1	2334	2864
				6954283H1 (BRAITDR02)	1	718
				9810284	2771	3269

10

20

30

40

表 4-5

ポリ ヌクレオチド SEQ ID NO:	Incye 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
35	71584520CB1	3017	2968-3017, 1-49, 1234-1322, 679-736	1287320T6 (BRAINFO11)	2615	2983
				71579751V1	1083	1791
				71580777V1	586	1149
				6764749H1 (BRAUNOR01)	1406	2001
				7581090H1 (BRAIFEC01)	1	577
				1295833H1 (PGANNOT03)	2783	3017
				1414795F6 (BRAINFO12)	2183	2508
				7362189H1 (BRAIFER05)	458	1003
				2157112T6 (BRAINFO09)	2388	2977
				2157112F6 (BRAINFO09)	1963	2450
36	7475538CB1	2168	811-870, 933-1227	6855691H1 (BRAIFEN08)	1500	2168
				70644867V1	613	1261
				70645804V1	515	1249
				70645323V1	1	595
				71564044V1	1277	1993
				71565564V1	1156	1924

10

20

30

40

表 5

ポリヌクレオチド SEQ ID NO:	Incyte プロジェクト ID	代表的ライブラリ
19	2890544CB1	MCLDXTX02
20	7472693CB1	COLNNOT01
21	3107952CB1	TYLMNOT04
22	5544420CB1	SEMVNOT01
23	7472832CB1	SKIRNOR01
24	1551456CB1	LVENNOT03
25	2589355CB1	BRADDIR01
26	4357117CB1	BMARTXT02
27	5511992CB1	SINTFEE02
28	7474560CB1	BRAYDIN03
29	7474602CB1	COLNNOT01
30	7475509CB1	BRAITUT21
31	7475491CB1	SCOMDIT01
32	2192119CB1	PROSTUS23
33	7474496CB1	BRAINON01
34	1834248CB1	BRAITUT22
35	71584520CB1	BRAIFEC01
36	7475538CB1	BRAIFEN08

【表 2 5】

10

20

30

40

表 6-1

ライブラリ	ベクター	ライブラリの説明
BMARTXT02	pINCY	ライブラリは 4 オの白人女子から除去した骨髄神経芽腫腫瘍細胞由来の処理された SH-SY5Y 細胞系から単離した RNA を用いて作製した。細胞はレチノイン酸の存在下で培養した。
BRADDIR01	pINCY	ライブラリは脳血管事故死亡の 57 オの白人男性の脳から除去した側室の罹病脈絡叢から単離した RNA を用いて作製した。
BRAIFEC01	pINCY	この大型分画のライブラリは、妊娠 23 週間で死産の左心低形成症胎児(白人男性)から除去された脳組織から単離した RNA を用いて作製した。
BRAIFEN08	pINCY	この標準化胎児脳組織ライブラリは胎児脳組織ライブラリの 400,000 の独立したクローンから作製した。RNA 源は妊娠 23 週目で死産の左心低形成症胎児(白人男性)から採取した脳組織から作製した。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 回にわたりノーマライズした。
BRAINON01	PSPORT1	ライブラリは脳組織ライブラリからの 488 万の独立型クローンから作製し、ノーマライズした。RNA は 28 オの白人男性の頭蓋形成術および脳髄膜病変の切除時に採取された脳組織から作製した。脳の右前線—頭頂部におけるグレード 4 の少星細胞腫を示す関連腫瘍組織の病理。
BRAITUT21	pINCY	ライブラリは、61 オの白人女性の脳髄膜病変の切除時、正中線前頭葉から採取した脳腫瘍組織から単離した RNA を用いて作製した。病理は非定型を伴わない前頭下髄膜皮髄腫(subfrontal meningotheelial meningioma)を示した。1 つの断骨と粘膜炎組織サンプルは髄膜腫を示した。家族歴には、脳血管疾患、老人性痴呆症、高脂血症、良性の高血圧、アテローム硬化型冠状動脈疾患、鬱血性心不全および乳がんが含まれる。
BRAITUT22	pINCY	ライブラリは、76 オの白人女性の脳髄膜病変の切除時、右前頭葉/頭頂葉から採取した脳腫瘍組織から単離した RNA を用いて作製した。病理は髄膜腫を示した。家族歴に老人性痴呆症が含まれていた。
BRAYDIN03	pINCY	このノーマライズされた胎児脳組織ライブラリは、BRAYDIT01 組織ライブラリの 670,000 の独立型クローンから作製した。開始 RNA は脳血管事故から死亡した 57 オの白人男子から採取された罹病視床下部組織から単離した RNA から作製した。患者の病歴にはハンチントン舞踏病および肺気腫がある。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 を改作した条件を用いて 2 回にわたり標準化した。このライブラリはクローンを含む挿入断片用を選択するために直線化し、再度環状にした。
COLNNOT01	PSPORT1	ライブラリは 75 オの白人男性の半結腸切除時に採取した大腸組織から単離した RNA を用いて作製した。
LIVENNOT03	PSPORT1	ライブラリは 31 オの男性の左心室組織から単離した RNA を用いて作製した。

表 6-2

ライブラリ	ベクター	ライブラリの説明
MCLDXT02	pINCY	ライブラリは、男性から採取し、処理した臍帯血液の樹状細胞から単離された RNA を用いて作製した。この細胞は顆粒球/マクロファージコロニー刺激因子(GM-CSF)、腫瘍壊死因子アルファ(TNF アルファ)、幹細胞因子(SCF)、ホルボールミリスチン酸アセテート(PMA)およびオノマインで処理した。GM-CSF は、時間 0、100 ng/ml で加え、TNF アルファは時間 0、2.5ng/ml で加え、SCF は時間 0、25ng/ml で加えた。PMA およびオノマインは 13 日目に 5 時間加えた。インキュベーション時間は 13 日間であった。
PROSTUS23	pINCY	この引き算された前立腺腫瘍ライブラリは、プールされた前立腺腫瘍ライブラリの 1000 万のクローンと 2 回の引き算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 万のクローンを用いて作製した。引き算開始ライブラリは、58 才 (A)、61 才 (B)、66 才 (C)、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、硬結およびタハ口の癌、ドナー B は PSA の上昇、硬結、前立腺肥厚、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、タハ口/アルコール癌、および C 型肝炎(ヤリア)、ドナー C は、PSA の上昇、硬結およびタハ口の癌、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。サブトラクショ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三人のドナーによる前立腺ストローマからの線維芽細胞から由来する 3 個の組織ライブラリから得た同じ数の cDNA クローンをプールすることによって作製した。引き算ハイブリダイゼーションの条件は Swaroop ら、NAR 19 (1991):1954 および Bonaldo らの Genome Research 6 (1996):791 の方法に基づいたものである。
SCOMDIT01	pINCY	ライブラリは脳血管事故により死亡した 57 才の白人男性の延髄底部から採取した腫瘍脊髄組織から単離した RNA を用いて作製した。患者の病歴にはハンチントン舞踏病および肺気腫がある。
SEMVN0T01	pINCY	ライブラリは 58 才の白人男子の根本的前立腺切除時に採取した精嚢組織から単離した RNA を用いて作製した。関連腫瘍組織の病理は前立腺の腺癌(Gleason グレード 3+2)を示した。腺癌腫瘍性過形成もまた見られた。この患者は前立腺特異抗原(PSA)の上昇も示した。家族歴には悪性乳房腫瘍が含まれていた。
SINTFEB02	PCDNA2.1	この 5' 偏向無作為ライブラリは妊娠 20 週に 7 日(トリソミー)で死亡した白人男胎児から採取した小腸組織から単離した RNA を用いて作製した。血清学は陰生であった。
SKIRNOR01	PCDNA2.1	この無作為ライブラリは 17 才の白人女性の両側性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。
TYMNOT04	pINCY	ライブラリは活性化 Th1 細胞から単離された RNA を用いて作製した。これらの細胞は IL-12 と B7-形質移入 COS 細胞で臍帯 CD4 細胞から分化され、次に、抗 CD3 および抗 CD28 抗体で 6 日間活性化した。

10

20

30

40

表7-1

プログラム	説明	引用文献	パラメータ閾値
ABI FACTURA	ベクター配列を除き、核酸配列のまぎわしい塩基をマスキするプログラム	Applied Biosystems, Foster City, CA.	
ABI/PARACEL FDF	アミノ酸または核酸配列の比較、注釈に有用なFirst Data Finder。	Applied Biosystems, Foster City, CA; Paracel Inc., Pasadena, CA.	不一致<50%
ABI オートアセンブラ	核酸配列の構築を行うプログラム。	Applied Biosystems, Foster City, CA.	
BLAST	アミノ酸および核酸配列の配列類似性検索に有用なパーシクローカルアライメント検索ツール。プラストにはblastp、blastn、tblastnおよびtblastxの5つの機能がある。	Altschul, S.F. 他 (1990) J. Mol. Biol. 215:403-410; Altschul, S.F. 他 (1997) Nucleic Acids Res. 25:3389-3402.	ESTs: 確率値=1.0E-8以下 完全長配列: 確率値1.0E-10 以下
FASTA	一群の同じタイプの配列と問合せ配列との間の類似性について検索するPearson およびLipman アルゴリズム。FASTAには最小5つの機能(: fasta、tfasta、fastx、tfastxおよびsssearch)がある。	Pearson, W.R. 及び D.J. Lipman (1988) Proc. Natl. Acad. Sci. USA 85:2444-2448; Pearson, W.R. (1990) Methods Enzymol. 183:63-98; Smith, T.F. 及び M.S. Waterman (1981) Adv. Appl. Math. 2:482-489.	ESTs: fasta E 値=1.0E-6 構築されたESTs: fasta l同一性=95%以上、一致した長さ=200塩基以上、fastx E value=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、配列相同性および構造的フィンガープリント領域を検索するBLOCKS IMPROVED Searcher。	Henikoff, S. 及び J.G. Henikoff (1991) Nucleic Acids Res. 19:6565-6572; Henikoff, J.G. 及び S. Henikoff (1996) Methods Enzymol. 266:88-105; Attwood, T.K. 他 (1997) J. Chem. Inf. Comput. Sci. 37:417-424.	確率値 = 1.0E-3以下
HMMER	問合せ配列をPFAMのようなタンパク質ファミリーコンセンサス配列の隠れたマルコフモデル(HMM)に基づいたデータベースに対して検索するアルゴリズム。	Krogh, A. 他(1994) J. Mol. Biol. 235:1501-1531; Sonnhammer, E.L.L. 他 (1998) Nucleic Acids Res. 26:320-322; Durbin, R. 他 (1998) Our World View, in a Nutshell, Cambridge Univ. Press, pp. 1-350.	PFAM ヒット: 確率値 =1.0E-3 以下 シグナルペプチドヒット: スコア=0以上

10

20

30

40

表7-2

プログラム	説明	引用文献	パラメータ閾値
ProfileScan	Prosileで定義された配列パターンと一致するタンパク質配列の構造的モチーフおよび配列モチーフを検索するアルゴリズム。	Gribskov, M. 他 (1988) CABIOS 4:61-66; Gribskov, M. 他 (1989) Methods Enzymol. 183:146-159; Bairoch, A. 他 (1997) Nucleic Acids Res. 25:217-221.	標準化された質のスコア≧特定の Prosile モチーフに対するGCC指定” HIGH” 値 通常、スコア=1.4-2.1.
Phred	高度感受性および確立性のある自動シーケンサトレースを調べる塩基読み出しアルゴリズム。	Ewing, B. 他 (1998) Genome Res. 8:175-185; Ewing, B. 及び P. Green (1998) Genome Res. 8:186-194.	
Phrap	Smith-Waterman アルゴリズムの効率的実行に基づいたプログラムである、SWATとCrossMatchを含むPhils Revised 構築プログラムで、配列相同性検索やDNA配列の構築に有用である。	Smith, T.F. 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; Green, P., University of Washington, Seattle, WA.	スコア=120以上 一致した長さ=56以上
Consed	Phrap 構築の表示、編集用グラフィカルツール。	Gordon, D. 他 (1998) Genome Res. 8:195-202.	
SPScan	タンパク質配列を走査し、分泌シグナルペプチドの存在を調べる重み付けマトリックス分析プログラム。	Nielson, H. 他 (1997) Protein Engineering 10:1-6; Claverie, J.M. 及び S. Audic (1997) CABIOS 12:431-439.	スコア=3.5以上
TMAP	重量マトリックスを使ってタンパク質配列上の膜貫通部分を描写し、向きを決定するプログラム。	Persson, B. 及び P. Argos (1994) J. Mol. Biol. 237:182-192; Persson, B. 及び P. Argos (1996) Protein Sci. 5:363-371.	
TMHMMER	隠れたMarkov モデル(HMM)を使ってタンパク質配列上の膜貫通部分を描写し、向きを決定するプログラム。	Sonnhammer, E.L. 他 (1998) Proc. Sixth Intl. Conf. on Intelligent Systems for Mol. Biol., Glasgow 他, eds., The Am. Assoc. for Artificial Intelligence Press, Menlo Park, CA, pp. 175-182.	
Motifs	Prosileで定義されたものと一致するパターンでの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, 9版, Page M51-59, 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
1 November 2001 (01.11.2001)

PCT

(10) International Publication Number
WO 01/81555 A2

(51) International Patent Classification: C12N 9/00

(31) International Application Number: PCT/US01/12992

(22) International Filing Date: 20 April 2001 (20.04.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:

60/199,021	20 April 2000 (20.04.2000)	US
60/200,226	28 April 2000 (28.04.2000)	US
60/202,339	5 May 2000 (05.05.2000)	US
60/203,505	11 May 2000 (11.05.2000)	US
60/205,564	18 May 2000 (18.05.2000)	US
60/207,739	26 May 2000 (26.05.2000)	US
60/208,795	1 June 2000 (01.06.2000)	US

(63) Related by continuation (CON) or continuation-in-part (CIP) to earlier applications:

US	60/200,226 (CIP)
Filed on	28 April 2000 (28.04.2000)
US	60/199,021 (CIP)
Filed on	20 April 2000 (20.04.2000)
US	60/202,339 (CIP)
Filed on	5 May 2000 (05.05.2000)
US	60/203,505 (CIP)
Filed on	11 May 2000 (11.05.2000)
US	60/205,564 (CIP)
Filed on	18 May 2000 (18.05.2000)
US	60/207,739 (CIP)
Filed on	26 May 2000 (26.05.2000)
US	60/208,795 (CIP)
Filed on	1 June 2000 (01.06.2000)

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

(72) Inventors: and

(75) Inventors/Applicants (for US only): YUE, Henry [US/US], 826 Lewis Avenue, Sunnyvale, CA 94087 (US); GANDHI, Ameeta, R. [US/US], 3137 Ruble Avenue, #1, Menlo Park, CA 94025 (US); TRIBOULEY, Catherine, M. [US/US], 1121 Tennessee Street, #5, San Francisco, CA 94107 (US); KEARNEY, Liam [IE/US], 50 Woodside Avenue, San Francisco, CA 94127 (US); GRIFFIN, Jennifer, A. [US/US], 181 Trana Court #11, Belmont, CA 94602 (US); NGUYEN, Daniel, R. [US/US], 1409 Redwood Drive, San Jose, CA 95118 (US); BANDMAN, Olga [US/US], 366 Anna Avenue, Mountain View, CA 94043 (US); LU, Dying, Aina, M. [US/US], 233 Coy Drive, San Jose, CA 95123 (US); LAI, Preeti [IN/US], P.O. Box 5142, Santa Clara, CA 95056 (US); BIRFORD, Neil [GB/US], 105 Wildwood Circle, Durham, CT 06422 (US); KHAN, Farrah, A. [IN/US], 333 Escuela Avenue, #221, Mountain View, CA 94040 (US); WALIA, Narinder, K. [US/US], 890 Davis Street #205, San Leandro, CA 94577 (US); YAO, Monique, G. [US/US], 111 Frederick Court, Mountain View, CA 94043 (US); PATTERSON, Chandra [US/US], 490 Sherwood Way #1, Menlo Park, CA 94025 (US); BURRILL, John, D. [US/US], 650 Emerald Hill Road, Redwood City, CA 94061 (US); MARCUS, Gregory, A. [US/US], 1714 Connecticut Drive, Redwood City, CA 94061 (US); ZINGLER, Kurt, A. [US/US], 723 Ashbury Street #C, San Francisco, CA 94117 (US); RECIPON, Shirley, A. [US/US], 85 Fortuna Avenue, San Francisco, CA 95115 2873 (US); LU, Yan [CN/US], 3885 Cornelia Way, Palo Alto, CA 94304 (US); POLICKY, Jennifer, L. [US/US], 1511 Jarvis Court, San Jose, CA 95118 (US); THORNTON, Michael [US/US], 9 Medway Road, Woodside, CA 94062 2612 (US); TANG, Y., Tom [US/US], 4250 Rumwick Court, San Jose, CA 95118 (US); HAFALIA, April [US/US], 2227 Calle de Primavera, Santa Clara, CA 95054 (US); ELLIOTT, Vicki, S. [US/US], 3770 Politen Place Way, San Jose, CA 95121 (US); BAUGHN, Mariah, R. [US/US], 14244 San Diego Road, San Leandro, CA 94577 (US); WALSH, Roderick, T. [IE/GB], 8 Boundary Court, St. Lawrence Road, Canterbury Kent CT1 3EZ (GB); RAMKUMAR, Jayalaxmi [IN/US], 54359 Maybird Circle, Fremont, CA 94535 (US); BOROWSKY, Mark, J. [US/US], 122 Orchard Avenue, Redwood City, CA 94061 (US); AU-YOUNG, Janice [US/US], 233 Golden Eagle Lane, Brisbane, CA 94005 (US); HILLMAN, Jennifer, L. [US/US], 230 Monroe Drive, #17, Mountain View, CA 94040 (US); GURURAJAN, Rajagopal [US/US], 5591 Dent Avenue, San Jose, CA 95118 (US).

(74) Agents: HAMILL-FOX, 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, CL, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, GR, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LU, LY, MA, MD, ME, MG, MK, MN, MU, MV, MW, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PK, PL, PT, RO, RU, RW, SA, SC, SD, SE, SG, SI, SK, SL, SM, SN, SR, ST, SV, SZ, TD, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, UZ, VC, VE, VG, VI, VN, YU, ZA, ZM, ZW.

[Continued on next page]

(54) Title: HUMAN KINASES

(57) Abstract: The invention provides 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 abnormal expression of PKIN.

WO 01/81555 A2

WO 01/81555 A2



LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MY,
NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM,
TR, TT, UZ, CA, CG, US, UZ, VN, YU, ZA, ZW.

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

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

For two-letter codes and other abbreviations, refer to the "Guid-
ance Notes on Codes and Abbreviations" appearing at the begin-
ning of each regular issue of the PCT Gazette.

WO 01/81555

PCT/US01/12992

HUMAN KINASES**TECHNICAL FIELD**

This invention relates to nucleic acid and amino acid sequences of human kinases and to the 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 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 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 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.

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 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

WO 01/81555

PCT/US01/12992

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

- 5 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.
- 10 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
- 15 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

- 20 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,
- 25 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.
- 30

- Many PTKs were first identified as oncogene products in cancer cells in which PTK 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
- 35

WO 01/81555

PCT/US01/12992

controlling some types of cancer.

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 protein
5 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 proliferation and differentiation, as well as regulation of the cytoskeleton. Activation of ERK is
10 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 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
15 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 messengers such as cyclic AMP (cAMP), cyclic GMP, inositol triphosphate, phosphatidylinositol, 3,4,5-triphosphate, cyclic ADP ribose, arachidonic acid, diacylglycerol and calcineurin-calmodulin. The
20 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 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
25 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 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
30 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 eukaryotic cells, and on the mitotic spindles of mammalian cells (Fish, K.J. et al. (1995) *J. Biol. Chem.* 270:14875-14883).

35 The CKI family members all have a short amino-terminal domain of 9-76 amino acids, a

WO 01/81555

PCT/US01/12992

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). 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 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*).

- 10 The mammalian circadian mutation tau was found to be a semidominant autosomal allele of 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.
- 15 Further, *in vitro*, CKI-epsilon is able to interact with mammalian PERIOD proteins, while the mutant 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
- 20 rhythms, jet-lag and sleep, in addition to other physiologic and metabolic processes under circadian regulation (Lowrey, P.L. et al. (2000) Science 288:493-491).

Homeodomain-interacting protein kinases (HIPKs) are serine/threonine kinases and novel members of the DYRK kinase subfamily (Hofmann, T.G. et al., (2000) Biochimie 82:1123-7). 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, Y.H. et al., (1998) J. Biol. Chem. 273:25875-9; 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 are important in the regulation of cellular protein function (Kim, Y.H. et al., (1999) Proc. Nat. Acad. Sci. U.S.A. 96:12350-5).

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

35 receptor, in response to the concentration of free calcium in the cell. Many CaM kinases are also

WO 01/81555

PCT/US01/12992

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). MAP kinase signaling pathways are present in mammalian cells as well as in yeast. The extracellular stimuli which activate MAP kinase pathways include epidermal growth factor (EGF), ultraviolet light, hyperosmolar medium, heat shock, endotoxic lipopolysaccharide (LPS), and pro-inflammatory cytokines such as tumor necrosis factor (TNF) and interleukin-1 (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.

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

WO 01/81555

PCT/US01/12992

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 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 steroid 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

WO 01/81555

PCT/US01/12992

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 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 DRAK 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*).

30 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

WO 01/81555

PCT/US01/12992

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 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*).

KINASES WITH NON-PROTEIN SUBSTRATES

15 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 (Leervers, 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,

WO 01/81555

PCT/US01/12992

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

5 for PI3K in type-2 diabetes (Shepherd, supra).

An example of lipid kinase phosphorylation activity is the phosphorylation of 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

10 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,

15 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

20 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

25 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

30 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

35 metabolic, muscle-energy disorders that can result in cardiac or respiratory failure and may be

WO 01/81555

PCT/US01/12992

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 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. 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. (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 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 (Amer E.S. and S. Eriksson (1995) Pharmacol. 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 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," "PKIN-7," "PKIN-8," "PKIN-9," "PKIN-10," "PKIN-11," "PKIN-12," "PKIN-13," "PKIN-14,"

WO 01/81555

PCT/US01/12992

"PKIN-15," "PKIN-16," "PKIN-17," and "PKIN-18." 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:1-18, b) a naturally occurring polypeptide comprising an amino acid sequence at least 90% identical to an amino acid sequence selected from

5 the group consisting of SEQ ID NO:1-18, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-18, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-18. In one alternative, the invention provides an isolated polypeptide comprising the amino acid sequence of SEQ ID NO:1-18.

10 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:1-18, b) a naturally occurring polypeptide comprising an amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-18, c) a biologically active fragment of a polypeptide having an amino acid sequence

15 selected from the group consisting of SEQ ID NO:1-18, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-18. In one alternative, the polynucleotide encodes a polypeptide selected from the group consisting of SEQ ID NO:1-18. In another alternative, the polynucleotide is selected from the group consisting of SEQ ID NO:19-36.

20 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:1-18, b) a naturally occurring polypeptide comprising an amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-18, c) a

25 biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-18, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-18. In one alternative, the invention provides a cell transformed with the recombinant polynucleotide. In another alternative, the invention provides a transgenic organism comprising the recombinant polynucleotide.

30 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:1-18, b) a naturally occurring polypeptide comprising an amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-18, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group

35 consisting of SEQ ID NO:1-18, and d) an immunogenic fragment of a polypeptide having an amino

WO 01/81555

PCT/US01/12992

acid sequence selected from the group consisting of SEQ ID NO:1-18. The method comprises a) culturing a cell under conditions suitable for expression of the polypeptide, wherein said cell is transformed with a recombinant polynucleotide comprising a promoter sequence operably linked to a polynucleotide encoding the polypeptide, and b) recovering the polypeptide so expressed.

5 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:1-18, b) a naturally occurring polypeptide comprising an amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-18, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-18, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-18.

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:19-36, b) a naturally occurring polynucleotide comprising a polynucleotide sequence at least 90% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:19-36, 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.

20 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:19-36, b) a naturally occurring polynucleotide comprising a polynucleotide sequence at least 90% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:19-36, 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

WO 01/81555

PCT/US01/12992

of a) a polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:19-36, b) a naturally occurring polynucleotide comprising a polynucleotide sequence at least 90% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:19-36, 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:1-18, b) a naturally occurring polypeptide comprising an amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-18, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-18, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-18, and a pharmaceutically acceptable excipient. In one embodiment, the composition comprises an amino acid sequence selected from the group consisting of SEQ ID NO:1-18. 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:1-18, b) a naturally occurring polypeptide comprising an amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-18, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-18, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-18. 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

WO 01/81555

PCT/US01/12992

an antagonist of a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-18, b) a naturally occurring polypeptide comprising an amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-18, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-18, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-18. 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:1-18, b) a naturally occurring polypeptide comprising an amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-18, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-18, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-18. 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:1-18, b) a naturally occurring polypeptide comprising an amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-18, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-18, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-18. 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

WO 01/81555

PCT/US01/12992

the test compound is indicative of a compound that modulates the activity of the polypeptide.

The invention further provides a method for screening a compound for effectiveness in altering expression of a target polynucleotide, wherein said target polynucleotide comprises a sequence selected from the group consisting of SEQ ID NO:19-36, 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:19-36, ii) a naturally occurring polynucleotide comprising a polynucleotide sequence at least 90% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:19-36, 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:19-36, ii) a naturally occurring polynucleotide comprising a polynucleotide sequence at least 90% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:19-36, iii) a polynucleotide complementary to the polynucleotide of i), iv) a polynucleotide complementary to the polynucleotide of ii), and v) an RNA equivalent of i)-iv). Alternatively, the target polynucleotide comprises a fragment of a polynucleotide sequence selected from the group consisting of i)-v) above; c) quantifying the amount of hybridization complex; and d) comparing the amount of hybridization complex in the treated biological sample with the amount of hybridization complex in an untreated biological sample, wherein a difference in the amount of hybridization complex in the treated biological sample is indicative of toxicity of the test compound.

BRIEF DESCRIPTION OF THE TABLES

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

Table 2 shows the GenBank identification number and annotation of the nearest GenBank homolog for polypeptides of the invention. The probability score for the match between each polypeptide and its GenBank homolog is also shown.

WO 01/81555

PCT/US01/12992

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

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

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

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

10 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
30 the cell lines, protocols, reagents and vectors which are reported in the publications and which might be used in connection with the invention. Nothing herein is to be construed as an admission that the invention is not entitled to antedate such disclosure by virtue of prior invention.

DEFINITIONS

"PKIN" refers to the amino acid sequences of substantially purified PKIN obtained from any
35 species, particularly a mammalian species, including bovine, ovine, porcine, murine, equine, and

WO 01/81555

PCT/US01/12992

human, and from any source, whether natural, synthetic, semi-synthetic, or recombinant.

The term "agonist" refers to a molecule which intensifies or mimics the biological activity of 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
 5 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
 10 allelic variants are generally ascribed to natural deletions, additions, or substitutions of nucleotides. Each of these types of changes may occur alone, or in combination with the others, one or more times in a given sequence.

"Altered" nucleic acid sequences encoding PKIN include those sequences with deletions, insertions, or substitutions of different nucleotides, resulting in a polypeptide the same as PKIN or a
 15 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
 20 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
 25 amino acids may include lysine and arginine. Amino acids with uncharged polar side chains having similar hydrophilicity values may include: asparagine and glutamine; and serine and threonine. Amino acids with uncharged side chains having similar hydrophilicity values may include: leucine, isoleucine, and valine; glycine and alanine; and phenylalanine and tyrosine.

The terms "amino acid" and "amino acid sequence" refer to an oligopeptide, peptide,
 30 polypeptide, or protein sequence, or a fragment of any of these, and to naturally occurring or synthetic molecules. Where "amino acid sequence" is recited to refer to a sequence of a naturally occurring protein molecule, "amino acid sequence" and like terms are not meant to limit the amino acid sequence to the complete native amino acid sequence associated with the recited protein molecule.

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

WO 01/81555

PCT/US01/12992

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
 5 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
 10 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
 15 used to elicit the immune response) for binding to an antibody.

The term "antisense" refers to any composition capable of base-pairing with the "sense" (coding) strand of a specific nucleic acid sequence. Antisense compositions may include DNA; RNA; peptide nucleic acid (PNA); oligonucleotides having modified backbone linkages such as phosphorothioates, methylphosphonates, or benzylphosphonates; oligonucleotides having modified
 25 sugar groups such as 2'-methoxyethyl sugars or 2'-methoxyethoxy sugars; or oligonucleotides having modified bases such as 5-methyl cytosine, 2'-deoxyuracil, or 7-deaza-2'-deoxyguanosine. Antisense molecules may be produced by any method including chemical synthesis or transcription. Once introduced into a cell, the complementary antisense molecule base-pairs with a naturally occurring nucleic acid sequence produced by the cell to form duplexes which block either transcription or
 30 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,
 35 to induce a specific immune response in appropriate animals or cells and to bind with specific

WO 01/81555

PCT/US01/12992

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'.

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.).

10 "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.

20 "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.

25	Original Residue	Conservative Substitution
	Ala	Gly, Ser
	Arg	His, Lys
	Asn	Asp, Gln, His
	Asp	Asn, Glu
30	Cys	Ala, Ser
	Gln	Asn, Glu, His
	Glu	Asp, Gln, His
	Gly	Ala
	His	Asn, Arg, Gln, Glu
35	Ile	Leu, Val
	Leu	Ile, Val
	Lys	Arg, Gln, Glu
	Met	Leu, Ile
	Phe	His, Met, Leu, Trp, Tyr

WO 01/81555

PCT/US01/12992

5	Ser Thr Trp Tyr Val	Cys, Thr Ser, Val Phe, Tyr His, Phe, Trp Asp, Leu, Thr
---	---------------------------------	--

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

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

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

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

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

A "fragment" is a unique portion of 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 fragment may comprise from 5 to 1000 contiguous nucleotides or amino acid residues. A fragment used as a probe, primer, antigen, therapeutic molecule, or for other purposes, may be at least 5, 10, 15, 16, 20, 25, 30, 40, 50, 60, 75, 100, 150, 250 or at least 500 contiguous nucleotides or amino acid residues in length. Fragments may be preferentially selected from certain regions of a molecule. For example, a polypeptide fragment may comprise a certain length of contiguous amino acids selected from the first 250 or 500 amino acids (or first 25% or 50%) of a polypeptide as shown in a certain defined sequence. Clearly these lengths are exemplary, and any length that is supported by the specification, including the Sequence Listing, tables, and figures, may be encompassed by the present embodiments.

WO 01/81555

PCT/US01/12992

A fragment of SEQ ID NO:19-36 comprises a region of unique polynucleotide sequence that specifically identifies SEQ ID NO:19-36, for example, as distinct from any other sequence in the genome from which the fragment was obtained. A fragment of SEQ ID NO:19-36 is useful, for example, in hybridization and amplification technologies and in analogous methods that distinguish
 5 SEQ ID NO:19-36 from related polynucleotide sequences. The precise length of a fragment of SEQ ID NO:19-36 and the region of SEQ ID NO:19-36 to which the fragment corresponds are routinely determinable by one of ordinary skill in the art based on the intended purpose for the fragment.

A fragment of SEQ ID NO:1-18 is encoded by a fragment of SEQ ID NO:19-36. A fragment of SEQ ID NO:1-18 comprises a region of unique amino acid sequence that specifically identifies
 10 SEQ ID NO:1-18. For example, a fragment of SEQ ID NO:1-18 is useful as an immunogenic peptide for the development of antibodies that specifically recognize SEQ ID NO:1-18. The precise length of a fragment of SEQ ID NO:1-18 and the region of SEQ ID NO:1-18 to which the fragment corresponds are routinely determinable by one of ordinary skill in the art based on the intended purpose for the fragment.

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

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

20 The terms "percent identity" and "% identity," as applied to polynucleotide sequences, refer to the percentage of residue matches between at least two polynucleotide sequences aligned using a standardized algorithm. Such an algorithm may insert, in a standardized and reproducible way, gaps in the sequences being compared in order to optimize alignment between two sequences, and therefore achieve a more meaningful comparison of the two sequences.

25 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
 30 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
 35 is provided by the National Center for Biotechnology Information (NCBI) Basic Local Alignment

WO 01/81555

PCT/US01/12992

Search Tool (BLAST) (Altschul, S.F. et al. (1990) *J. Mol. Biol.* 215:403-410), which is available from several sources, including the NCBI, Bethesda, MD, and on the Internet at <http://www.ncbi.nlm.nih.gov/BLAST/>. The BLAST software suite includes various sequence analysis programs including "blastn," that is used to align a known polynucleotide sequence with other polynucleotide sequences from a variety of databases. Also available is a tool called "BLAST 2 Sequences" that is used for direct pairwise comparison of two nucleotide sequences. "BLAST 2 Sequences" can be accessed and used interactively at <http://www.ncbi.nlm.nih.gov/gorf/bl2.html>. The "BLAST 2 Sequences" tool can be used for both blastn and blastp (discussed below). BLAST programs are commonly used with gap and other parameters set to default settings. For example, to compare two nucleotide sequences, one may use blastn with the "BLAST 2 Sequences" tool Version 2.0.12 (April-21-2000) set at default parameters. Such default parameters may be, for example:

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

Percent identity may be measured over the length of an entire defined sequence, for example, as defined by a particular SEQ ID number, or may be measured over a shorter length, for example, over the length of a fragment taken from a larger, defined sequence, for instance, a fragment of at least 20, at least 30, at least 40, at least 50, at least 70, at least 100, or at least 200 contiguous nucleotides. Such lengths are exemplary only, and it is understood that any fragment length supported by the sequences shown herein, in the tables, figures, or Sequence Listing, may be used to describe a length over which percentage identity may be measured.

Nucleic acid sequences that do not show a high degree of identity may nevertheless encode similar amino acid sequences due to the degeneracy of the genetic code. It is understood that changes in a nucleic acid sequence can be made using this degeneracy to produce multiple nucleic acid sequences that all encode substantially the same protein.

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

WO 01/81555

PCT/US01/12992

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

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

Matrix: BLOSUM62

Open Gap: 11 and Extension Gap: 1 penalties

Gap x drop-off: 50

Expect: 10

Word Size: 3

Filter: on

Percent identity may be measured over the length of an entire defined polypeptide sequence, for example, as defined by a particular SEQ ID number, or may be measured over a shorter length, for example, over the length of a fragment taken from a larger, defined polypeptide sequence, for instance, a fragment of at least 15, at least 20, at least 30, at least 40, at least 50, at least 70 or at least 150 contiguous residues. Such lengths are exemplary only, and it is understood that any fragment length supported by the sequences shown herein, in the tables, figures or Sequence Listing, may be used to describe a length over which percentage identity may be measured.

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

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

"Hybridization" refers to the process by which a polynucleotide strand anneals with a complementary strand through base pairing under defined hybridization conditions. Specific hybridization is an indication that two nucleic acid sequences share a high degree of complementarity. Specific hybridization complexes form under permissive annealing conditions and remain hybridized

WO 01/81555

PCT/US01/12992

after the "washing" step(s). The washing step(s) is particularly important in determining the stringency of the hybridization process, with more stringent conditions allowing less non-specific binding, i.e., binding between pairs of nucleic acid strands that are not perfectly matched. Permissive conditions for annealing of nucleic acid sequences are routinely determinable by one of ordinary skill in the art and may be consistent among hybridization experiments, whereas wash conditions may be varied among experiments to achieve the desired stringency, and therefore hybridization specificity. Permissive annealing conditions occur, for example, at 68°C in the presence of about 6 x SSC, about 1% (w/v) SDS, and about 100 µg/ml sheared, denatured salmon sperm DNA.

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

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

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

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

WO 01/81555

PCT/US01/12992

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

5 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.

10 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.

15 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.

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.

20 "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.

25 "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.

30 "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.

35 "Probe" refers to nucleic acid sequences encoding PKIN, their complements, or fragments

WO 01/81555

PCT/US01/12992

thereof, which are used to detect identical, allelic or related nucleic acid sequences. Probes are isolated oligonucleotides or polynucleotides attached to a detectable label or reporter molecule. Typical labels include radioactive isotopes, ligands, chemiluminescent agents, and enzymes.

"Primers" are short nucleic acids, usually DNA oligonucleotides, which may be annealed to a target polynucleotide by complementary base-pairing. The primer may then be extended along the target DNA strand by a DNA polymerase enzyme. Primer pairs can be used for amplification (and identification) of a nucleic acid sequence, e.g., by the polymerase chain reaction (PCR).

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

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

Oligonucleotides for use as primers are selected using software known in the art for such purpose. For example, OLIGO 4.06 software is useful for the selection of PCR primer pairs of up to 100 nucleotides each, and for the analysis of oligonucleotides and larger polynucleotides of up to 5,000 nucleotides from an input polynucleotide sequence of up to 32 kilobases. Similar primer selection programs have incorporated additional features for expanded capabilities. For example, the PrimOU primer selection program (available to the public from the Genome Center at University of Texas South West Medical Center, Dallas TX) is capable of choosing specific primers from megabase sequences and is thus useful for designing primers on a genome-wide scope. The Primer3 primer selection program (available to the public from the Whitehead Institute/MIT Center for Genome Research, Cambridge MA) allows the user to input a "mispriming library," in which sequences to avoid as primer binding sites are user-specified. Primer3 is useful, in particular, for the selection of oligonucleotides for microarrays. (The source code for the latter two primer selection programs may also be obtained from their respective sources and modified to meet the user's specific needs.) The PrimeGen program (available to the public from the UK Human Genome Mapping

WO 01/81555

PCT/US01/12992

Project Resource Centre, Cambridge UK) designs primers based on multiple sequence alignments, thereby allowing selection of primers that hybridize to either the most conserved or least conserved regions of aligned nucleic acid sequences. Hence, this program is useful for identification of both unique and conserved oligonucleotides and polynucleotide fragments. The oligonucleotides and polynucleotide fragments identified by any of the above selection methods are useful in hybridization technologies, for example, as PCR or sequencing primers, microarray elements, or specific probes to identify fully or partially complementary polynucleotides in a sample of nucleic acids. Methods of oligonucleotide selection are not limited to those described above.

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

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

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

"Reporter molecules" are chemical or biochemical moieties used for labeling a nucleic acid, amino acid, or antibody. Reporter molecules include radionuclides; enzymes; fluorescent, chemiluminescent, or chromogenic agents; substrates; cofactors; inhibitors; magnetic particles; and other moieties known in the art.

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

The term "sample" is used in its broadest sense. A sample suspected of containing PKIN, nucleic acids encoding PKIN, or fragments thereof may comprise a bodily fluid; an extract from a cell, chromosome, organelle, or membrane isolated from a cell; a cell; genomic DNA, RNA, or

WO 01/81555

PCT/US01/12992

cDNA, in solution or bound to a substrate; a tissue; a tissue print; etc.

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

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

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

"Substrate" refers to any suitable rigid or semi-rigid support including membranes, filters, chips, slides, wafers, fibers, magnetic or nonmagnetic beads, gels, tubing, plates, polymers, microparticles and capillaries. The substrate can have a variety of surface forms, such as wells,
 15 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
 20 type or tissue under given conditions at a given time.

"Transformation" describes a process by which exogenous DNA is introduced into a recipient cell. Transformation may occur under natural or artificial conditions according to various methods well known in the art, and may rely on any known method for the insertion of foreign nucleic acid sequences into a prokaryotic or eukaryotic host cell. The method for transformation is selected based
 25 on the type of host cell being transformed and may include, but is not limited to, bacteriophage or viral infection, electroporation, heat shock, lipofection, and particle bombardment. The term "transformed cells" includes stably transformed cells in which the inserted DNA is capable of replication either as an autonomously replicating plasmid or as part of the host chromosome, as well as transiently transformed cells which express the inserted DNA or RNA for limited periods of time.

A "transgenic organism," as used herein, is any organism, including but not limited to
 30 animals and plants, in which one or more of the cells of the organism contains heterologous nucleic acid introduced by way of human intervention, such as by transgenic techniques well known in the art. The nucleic acid is introduced into the cell, directly or indirectly by introduction into a precursor of the cell, by way of deliberate genetic manipulation, such as by microinjection or by infection with a recombinant virus. The term genetic manipulation does not include classical cross-breeding, or in
 35

WO 01/81555

PCT/US01/12992

in vitro fertilization, but rather is directed to the introduction of a recombinant DNA molecule. The transgenic organisms contemplated in accordance with the present invention include bacteria, cyanobacteria, fungi, plants and animals. The isolated DNA of the present invention can be introduced into the host by methods known in the art, for example infection, transfection,

5 transformation or transconjugation. Techniques for transferring the DNA of the present invention into such organisms are widely known and provided in references such as Sambrook et al. (1989), *supra*.

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

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

THE INVENTION

35 The invention is based on the discovery of new human kinases (PKIN), the polynucleotides

WO 01/81555

PCT/US01/12992

encoding PKJN, and the use of these compositions for the diagnosis, treatment, or 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 single Incyte project identification number (Incyte Project ID). Each polypeptide sequence is denoted by both a polypeptide sequence identification number (Polypeptide SEQ ID NO:) and an Incyte polypeptide sequence number (Incyte Polypeptide ID) as shown. Each polynucleotide sequence is denoted by both a polynucleotide sequence identification number (Polynucleotide SEQ ID NO:) and an Incyte polynucleotide consensus sequence number (Incyte Polynucleotide ID) as shown.

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

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

Together, Tables 2 and 3 summarize the properties of polypeptides of the invention, and these properties establish that the claimed polypeptides are human kinases. For example, SEQ ID NO:12 is 95% identical to a human adenylate kinase (GenBank ID g28577) as determined by the Basic Local Alignment Search Tool (BLAST). (See Table 2). The BLAST probability score is 2.4e-112, which indicates the probability of obtaining the observed polypeptide sequence alignment by chance. Data from BLIMPS, MOTIFS, and PROFILE analysis; from BLAST analysis using the DDMO and PRODOM databases; and from HMMER analysis using the PFAM database further support the categorization of SEQ ID NO:12 as an adenylate kinase. (See Table 3). In an alternative example,

WO 01/81555

PCT/US01/12992

SEQ ID NO:13 is 92% identical to rat PCTAIRE 3 (GenBank ID g2257588) as determined by the Basic Local Alignment Search Tool (BLAST). The BLAST probability score is $1.4e-210$, which indicates the probability of obtaining the observed polypeptide sequence alignment by chance (Table 2). PCTAIRE 1, 2, and 3 comprise a subfamily of Cdc2-related kinases that are primarily expressed in post-mitotic cells. PCTAIRE 2 and PCTAIRE 3 are expressed in the brain (Hirose, T. *et al.*, (1997) *Eur. J. Biochem.* 249:461-488; Okuda, T. *et al.*, (1992) *Oncogene* 7:2249-2258). SEQ ID NO:13 also contains a kinase active site 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:13 is a kinase. In another alternative example, SEQ ID NO:18 is 86% identical to human cell cycle related kinase (GenBank ID g4090958) as determined by the Basic Local Alignment Search Tool (BLAST). (See Table 2.) The BLAST probability score is $9.2e-85$, which indicates the probability of obtaining the observed polypeptide sequence alignment by chance. SEQ ID NO:18 also contains a 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:18 is a cell cycle related kinase. SEQ ID NO:1-11 and SEQ ID NO:14-17 were analyzed and annotated in a similar manner. The algorithms and parameters for the analysis of SEQ ID NO:1-18 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:19-36 or that distinguish between SEQ ID NO:19-36 and related polynucleotide sequences. Column 5 shows identification numbers corresponding to cDNA sequences, coding sequences (exons) predicted from genomic DNA, and/or sequence assemblages comprised of both cDNA and genomic DNA. These sequences were used to assemble the full length polynucleotide sequences of the invention. Columns 6 and 7 of Table 4 show the nucleotide start (5') and stop (3') positions of the cDNA and 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, 6311370H1 is the

WO 01/81555

PCT/US01/12992

identification number of an Incyte cDNA sequence, and NERDNDN03 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., 70518523D1). Alternatively, the identification numbers in 5 may refer to GenBank cDNAs or ESTs (e.g., g1860144) which contributed to the assembly of the full length polynucleotide sequences. Alternatively, the identification numbers in column 5 may refer to coding regions predicted by Genscan analysis of genomic DNA. For example, GNN.g5924006_004.edit is the identification number of a Genscan-predicted coding sequence, with g5924006 being the GenBank identification number of the sequence to which Genscan was applied. The Genscan-predicted coding sequences may have been edited prior to assembly. (See Example IV.)

10 Alternatively, the identification numbers in column 5 may refer to assemblages of both cDNA and Genscan-predicted exons brought together by an "exon stitching" algorithm. (See Example V.) Alternatively, the identification numbers in column 5 may refer to assemblages of both cDNA and Genscan-predicted exons brought together by an "exon-stretching" algorithm. (See Example V.) In some cases, Incyte cDNA coverage redundant with the sequence coverage shown in column 5 was

15 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

20 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 alternatively at least about 95%, or even at least about 98% amino acid sequence identity to the PKIN amino acid sequence, and which contains

25 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: 19-36, which encodes PKIN. The polynucleotide sequences of SEQ ID NO: 19-36, as presented in the Sequence Listing, embrace the equivalent RNA sequences,

30 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 about 85%, or alternatively at least about 95%, or even at least about 98% polynucleotide

35 sequence identity to the polynucleotide sequence encoding PKIN. A particular aspect of the

WO 01/81555

PCT/US01/12992

invention encompasses a variant of a polynucleotide sequence comprising a sequence selected from the group consisting of SEQ ID NO:19-36 which has at least about 70%, or alternatively at least about 85%, or alternatively at least about 95%, or even at least about 98% polynucleotide sequence identity to a nucleic acid sequence selected from the group consisting of SEQ ID NO:19-36. Any one of the

5 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

10 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.

15 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

20 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.

25 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.

30 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:19-36 and fragments thereof under various conditions of stringency. (See, e.g., Wahl, G.M. and S.L. Berger (1987) *Methods Enzymol.* 152:399-407; Kimmel, A.R. (1987) *Methods Enzymol.* 152:507-511.) Hybridization conditions, including annealing and wash conditions, are described in

35 "Definitions."

WO 01/81555

PCT/US01/12992

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 300 (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., Trigila, T. et al. (1988) *Nucleic Acids Res.* 16:8186.) A third method, capture PCR, involves PCR amplification of DNA fragments adjacent to known sequences in human and yeast artificial chromosome DNA. (See, e.g., Lagerstrom, M. et al. (1991) *PCR Methods Applic.* 1:111-119.) In this method, multiple restriction enzyme digestions and ligations may be used to insert an engineered double-stranded sequence into a region of unknown sequence before performing PCR. Other methods which may be used to retrieve unknown sequences are known in the art. (See, e.g., Parker, J.D. et al. (1991) *Nucleic Acids Res.* 19:3055-3060). Additionally, one may use PCR, nested primers, and PROMOTERFINDER libraries (Clontech, Palo Alto CA) to walk genomic DNA. This procedure avoids the need to screen libraries and is useful in finding intron/exon junctions. For all PCR-based methods, primers may be designed using commercially available software, such as OLIGO 4.06 primer analysis software (National Biosciences, Plymouth MN) or another appropriate program, to be about 22 to 30 nucleotides in length, to have a GC content of about 50% or more, and to anneal to the template at temperatures of about 68°C to 72°C.

WO 01/81555

PCT/US01/12992

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

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

In another embodiment of the invention, polynucleotide sequences or fragments thereof which encode PKIN may be cloned in recombinant DNA molecules that direct expression of PKIN, or fragments or functional equivalents thereof, in appropriate host cells. Due to the inherent degeneracy of the genetic code, other DNA sequences which encode substantially the same or a functionally equivalent amino acid sequence may be produced and used to express 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 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
 25 sites, alter glycosylation patterns, change codon preference, produce splice variants, and so forth.

The nucleotides of the present invention may be subjected to DNA shuffling techniques such as MOLECULAR BREEDING (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.
 30 Biotechnol. 17:259-264; and Cramer, A. et al. (1996) Nat. Biotechnol. 14:315-319) to alter or improve 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
 35 properties. These preferred variants may then be pooled and further subjected to recursive rounds of

WO 01/81555

PCT/US01/12992

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
 5 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.)
 10 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
 15 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
 20 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
 25 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
 30 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, exogenous translational control signals including an in-frame ATG initiation codon should be
 35 provided by the vector. Exogenous translational elements and initiation codons may be of various

WO 01/81555

PCT/US01/12992

origins, both natural and synthetic. The efficiency of expression may be enhanced by the inclusion of enhancers appropriate for the particular host cell system used. (See, e.g., Scharf, D. et al. (1994) *Results Probl. Cell Differ.* 20:125-162.)

Methods which are well known to those skilled in the art may be used to construct expression
 5 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.)
 10 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);
 plant cell systems transformed with viral expression vectors (e.g., cauliflower mosaic virus, CaMV,
 15 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 Hecke, 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
 20 York NY, pp. 191-196; Logan, J. and T. Shenk (1984) *Proc. Natl. Acad. Sci. USA* 81:3655-3659; and
 Harrington, J.J. et al. (1997) *Nat. Genet.* 15:345-355.) Expression vectors derived from retroviruses,
 adenoviruses, or herpes or vaccinia viruses, or from various bacterial plasmids, may be used for
 delivery of nucleotide sequences to the targeted organ, tissue, or cell population. (See, e.g., Di
 Nicola, M. et al. (1998) *Cancer Gen. Ther.* 5(6):350-356; Yu, M. et al. (1993) *Proc. Natl. Acad. Sci.*
 25 *USA* 90(13):6340-6344; Butler, R.M. et al. (1985) *Nature* 317(6040):813-815; McGregor, D.P. et al.
 (1994) *Mol. Immunol.* 31(3):219-226; and Venna, I.M. and N. Sonja (1997) *Nature* 389:239-242.)
 The invention is not limited by the host cell employed.

In bacterial systems, a number of cloning and expression vectors may be selected depending
 upon the use intended for polynucleotide sequences encoding PKIN. For example, routine cloning,
 30 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
 35 in vitro transcription, dideoxy sequencing, single strand rescue with helper phage, and creation of

WO 01/81555

PCT/US01/12992

nested deletions in the cloned sequence. (See, e.g., Vaut 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.

5 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, 10 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.* 15 6:307-311). Alternatively, plant promoters such as the small subunit of RUBISCO or heat shock promoters may be used. (See, e.g., Coruzzi, G. et al. (1984) *EMBO J.* 3:1671-1680; Broglie, R. et al. (1984) *Science* 224:838-843; and Winter, J. et al. (1991) *Results Probl. Cell Differ.* 17:85-105.) These constructs can be introduced into plant cells by direct DNA transformation or pathogen-mediated transfection. (See, e.g., *The McGraw Hill Yearbook of Science and Technology* 20 (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 25 infective virus which expresses PKIN in host cells. (See, e.g., Logan, J. and T. Shenk (1984) *Proc. Natl. Acad. Sci. USA* 81:3655-3659.) In addition, transcription enhancers, such as the Rous sarcoma virus (RSV) enhancer, may be used to increase expression in mammalian host cells. SV40 or EBV-based vectors may also be used for high-level protein expression.

Human artificial chromosomes (HACs) may also be employed to deliver larger fragments of 30 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 35 of PKIN in cell lines is preferred. For example, sequences encoding PKIN can be transformed into

WO 01/81555

PCT/US01/12992

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

Any number of selection systems may be used to recover transformed cell lines. These include, but are not limited to, the herpes simplex virus thymidine kinase and adenine phosphoribosyltransferase genes, for use in *tk⁻* and *aprt⁻* cells, respectively. (See, e.g., Wigler, M. et al. (1977) Cell 11:223-232; Lowy, I. et al. (1980) Cell 22:817-823.) Also, antimetabolite, antibiotic, or herbicide resistance can be used as the basis for selection. For example, *dhfr* confers resistance to methotrexate; *neo* confers resistance to the aminoglycosides neomycin and G-418; and *als* and *pat* confer resistance to chlorsulfuron and phosphinotricin acetyltransferase, respectively. (See, e.g., Wigler, M. et al. (1980) Proc. Natl. Acad. Sci. USA 77:3567-3570; Colbere-Garapin, F. et al. (1981) J. Mol. Biol. 150:1-14.) Additional selectable genes have been described, e.g., *trpB* and *hisD*, which alter cellular requirements for metabolites. (See, e.g., Hartman, S.C. and R.C. Mulligan (1988) Proc. Natl. Acad. Sci. USA 85:8047-8051.) Visible markers, e.g., anthocyanins, green fluorescent proteins (GFP; Clontech), β glucuronidase and its substrate β -glucuronide, or luciferase and its substrate luciferin may be used. These markers can be used not only to identify transformants, but also to quantify the amount of transient or stable protein expression attributable to a specific vector system. (See, e.g., Rhodes, C.A. (1995) Methods Mol. Biol. 55:121-131.)

Although the presence/absence of marker gene expression suggests that the gene of interest is also present, the presence and expression of the gene may need to be confirmed. For example, if the sequence encoding 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. Those 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

WO 01/81555

PCT/US01/12992

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, Torowa 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 case of detection include radionuclides, enzymes, fluorescent, chemiluminescent, or chromogenic agents, as well as substrates, cofactors, inhibitors, magnetic particles, and the like.

Host cells transformed with nucleotide sequences encoding 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 "propro" or "pre" 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

WO 01/81555

PCT/US01/12992

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 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.

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 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 those compounds involves producing appropriate cells which express PKIN, either as a secreted

WO 01/81555

PCT/US01/12992

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.

5 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
10 labeled competitor. Additionally, the assay may be carried out using cell-free preparations, chemical libraries, or natural product mixtures, and the test compound(s) may be free in solution or affixed to a solid support.

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
15 inverse agonists. In one embodiment, an assay is performed under conditions permissive for PKIN 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
20 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.

In another embodiment, polynucleotides encoding PKIN or their mammalian homologs may
25 be "knocked out" in an animal model system using homologous recombination in embryonic stem (ES) cells. Such techniques are well known in the art and are useful for the generation of animal models of human disease. (See, e.g., U.S. Patent Number 5,175,383 and U.S. Patent Number 5,767,337.) For example, mouse ES cells, such as the mouse 129/SvJ cell line, are derived from the early mouse embryo and grown in culture. The ES cells are transformed with a vector containing the
30 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 (Math, J.D. (1996) Clin. Invest. 97:1999-2002; Wagner, K.U. et al. (1997) Nucleic

WO 01/81555

PCT/US01/12992

Acids Res. 25:4323-4330). Transformed ES cells are identified and microinjected into mouse cell blastocysts such as those from the C57BL/6 mouse strain. The blastocysts are surgically transferred to pseudopregnant dams, and the resulting chimeric progeny are genotyped and bred to produce heterozygous or homozygous strains. Transgenic animals thus generated may be tested with potential therapeutic or toxic agents.

Polynucleotides encoding 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. (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 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).

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 cancer, diseased, proliferative, cardiac, tumorous, and digestive tissues, degenerative diseases of the brain, suggesting that PKIN plays a role in necrotic disorders affecting the central nervous system, and neuronal tissues (e.g. brain and spinal cord, see Table 6). Therefore, PKIN appears to play a role in maintenance and potentially the neoplastic transformation of cells of the central nervous system, and 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

WO 01/81555

PCT/US01/12992

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, hypereosinophilia, irritable bowel syndrome, multiple sclerosis, myasthenia gravis, myocardial or pericardial inflammation, osteoarthritis, osteoporosis, pancreatitis, polyarthritis, 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 mucopolysaccharidosis, 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 phlebotrombosis, 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,

WO 01/81555

PCT/US01/12992

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 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 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 storage disorders such as 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, lipomatosis, acute panniculitis, disseminated fat necrosis, adiposis dolorosa, lipoid adrenal hyperplasia, minimal change disease, lipomas, atherosclerosis, hypercholesterolemia, hypercholesterolemia with hypertriglyceridemia, primary hypoparathyroidism, 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 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 provided above.

In still another embodiment, an agonist which modulates the activity of PKIN may be

WO 01/81555

PCT/US01/12992

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 disorders include, but are not limited to, those cancers, 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.

10 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 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 therapeutic agents. Selection of the appropriate agents for use in combination therapy may be made by one of ordinary skill in the art, according to conventional pharmaceutical principles. The combination of therapeutic agents may act synergistically to effect the treatment or prevention of the various disorders described above. Using this approach, one may be able to achieve therapeutic efficacy with lower dosages of each agent, thus reducing the potential for adverse side effects.

20 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

WO 01/81555

PCT/US01/12992

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.

5 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

10 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,

15 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 chain antibodies. Antibodies with related specificity, but of distinct idiotype 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.)

20 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.

25 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.)

30 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

35 reactive to two non-interfering PKIN epitopes is generally used, but a competitive binding assay may

WO 01/81555

PCT/US01/12992

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

The titer and avidity of polyclonal antibody preparations may be further evaluated to determine the quality and suitability of such preparations for certain downstream applications. For example, a polyclonal antibody preparation containing at least 1-2 mg specific antibody/ml, preferably 5-10 mg specific antibody/ml, is generally employed in procedures requiring precipitation of 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*.)

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 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 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)

WO 01/81555

PCT/US01/12992

9(13):1288-1296.) Antisense sequences can also be introduced intracellularly through the use of viral vectors, such as retrovirus and adeno-associated virus vectors. (See, e.g., Miller, A.D. (1990) *Blood* 76:271; Ausubel, supra; Uckert, W. and W. Walther (1994) *Pharmacol. Ther.* 63(3):323-347.) Other gene delivery mechanisms include liposome-derived systems, artificial viral envelopes, and other systems known in the art. (See, e.g., Rossi, J.J. (1995) *Br. Med. Bull.* 51(1):217-225; Beado, 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 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 (Blaise, 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.

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 (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. Réeipon (1998) *Curr. Opin. Biotechnol.* 9:445-450).

Expression vectors that may be effective for the expression of PKIN include, but are not

WO 01/81555

PCT/US01/12992

limited to, the pCDNA 3.1, EPITAG, PROCMV2, PREP, PVAX vectors (Invitrogen, Carlsbad CA), PCMV-SCRIPT, PCMV-TAG, PEGSH/PERV (Stratagene, La Jolla CA), and PTET-OFF, PTET-ON, PTRE2, PTRH2-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 ecdysone-inducible promoter (available in the plasmids PVGRXR and PINO; Invitrogen); the FK506/rapamycin inducible promoter; or the RU486/mifepristone inducible promoter (Rossi, F.M.V. and Blau, H.M. *supra*), or (iii) a tissue-specific promoter or the native promoter of the endogenous gene encoding 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 polynucleotides to target cells in culture and require minimal effort to optimize experimental parameters. In the alternative, transformation is performed using the calcium phosphate method (Graham, F.L. and A.J. Eb (1973) *Virology* 52:456-467), or by electroporation (Neumann, E. et al. (1982) *EMBO J.* 1:841-845). The introduction of DNA to primary cells requires modification of these standardized mammalian transfection protocols.

In another embodiment of the invention, diseases or disorders caused by genetic defects with respect to 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 required for efficient vector propagation. Retrovirus vectors (e.g., PFB and PFBNEO) are commercially available (Stratagene) and are based on published data (Riviere, I. et al. (1995) *Proc. Natl. Acad. Sci. USA* 92:6733-6737), incorporated by reference herein. The vector is propagated in an appropriate vector producing cell line (VPCL) that expresses an envelope gene with a tropism for receptors on the target cells or a promiscuous envelope protein such as VSVg (Armentano, D. et al. (1987) *J. Virol.* 61:1647-1650; Bender, M.A. et al. (1987) *J. Virol.* 61:1639-1646; Adam, M.A. and A.D. Miller (1988) *J. Virol.* 62:3802-3806; Dull, T. et al. (1998) *J. Virol.* 72:8463-8471; Zufferey, R. et al. (1998) *J. Virol.* 72:9873-9880). U.S. Patent Number 5,910,434 to Rigg ("Method for obtaining retrovirus packaging cell lines producing high transducing efficiency retroviral supernatant") discloses a method for obtaining retrovirus packaging cell lines and is hereby incorporated by reference. Propagation of retrovirus vectors, transduction of a population of cells (e.g., CD4⁺ T-

WO 01/81555

PCT/US01/12992

cells), and the return of transduced cells to a patient are procedures well known to persons skilled in the art of gene therapy and have been well documented (Ranga, U. et al. (1997) *J. Virol.* 71:7020-7029; Bauer, G. et al. (1997) *Blood* 89:2259-2267; Bonyhadi, M.L. (1997) *J. Virol.* 71:4707-4716; Ranga, U. et al. (1998) *Proc. Natl. Acad. Sci. USA* 95:1201-1206; Su, L. (1997) *Blood* 89:2283-2290).

In the alternative, an adenovirus-based gene therapy delivery system is used to deliver polynucleotides encoding 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 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 Antinuzzi, P.A. et al. (1999) *Annu. Rev. Nutr.* 19:511-544 and Verma, I.M. and N. Sonja (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

WO 01/81555

PCT/US01/12992

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 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,

WO 01/81555

PCT/US01/12992

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.

5 Complementary ribonucleic acid molecules and ribozymes of the invention may be prepared by any method known in the art for the synthesis of nucleic acid molecules. These include techniques for chemically synthesizing oligonucleotides such as solid phase phosphoramidite chemical synthesis. Alternatively, RNA molecules may be generated by *in vitro* and *in vivo* transcription of DNA sequences encoding 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
10 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
20 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,
25 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
30 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
35 commonly known in the art, including chemical modification of a compound known to be effective in

WO 01/81555

PCT/US01/12992

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 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.

WO 01/81555

PCT/US01/12992

Various formulations are commonly known and are thoroughly discussed in the latest edition of *Remington's Pharmaceutical Sciences* (Mack 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, 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. 5,997,843). 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

WO 01/81555

PCT/US01/12992

lethal to 50% of the population) statistics. The dose ratio of toxic to therapeutic effects is the therapeutic index, which can be expressed as the LD_{50}/ED_{50} ratio. Compositions which exhibit large therapeutic indices are preferred. The data obtained from cell culture assays and animal studies are used to formulate a range of dosage for human use. The dosage contained in such compositions is preferably within a range of circulating concentrations that includes the ED_{50} with little or no toxicity. The dosage varies within this range depending upon the dosage form employed, the sensitivity of the patient, and the route of administration.

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

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

DIAGNOSTICS

In another embodiment, antibodies which specifically bind PKIN may be used for the 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 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 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,

WO 01/81555

PCT/US01/12992

control, and disease samples from biopsied tissues are compared with the standard values. Deviation between standard and subject values establishes the parameters for diagnosing disease.

In another embodiment of the invention, the polynucleotides encoding 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:19-36 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 syndrome, allergies, ankylosing spondylitis, amyloidosis, anemia, asthma, atherosclerosis,

WO 01/81555

PCT/US01/12992

- 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,
- 5 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, polyomyositis, psoriasis, Reiter's syndrome, rheumatoid arthritis, scleroderma, Sjögren's syndrome, systemic anaphylaxis, systemic lupus erythematosus, systemic sclerosis, thrombocytopenic purpura, ulcerative colitis, uveitis, Werner
- 10 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,
- 15 leukemia, lymphoma, melanoma, myeloma, sarcoma, teratocarcinoma, and, in particular, cancers of the adrenal gland, bladder, 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,
- 20 gonadal dysgenesis, WAGR syndrome (Wilms' tumor, aniridia, genitourinary abnormalities, and mental retardation), Smith-Magenis syndrome, myelodysplastic syndrome, hereditary mucopolysaccharidosis, 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
- 25 sensorineural hearing loss; a cardiovascular disease, such as arteriovenous fistula, atherosclerosis, hypertension, vasculitis, Raynaud's disease, aneurysms, arterial dissections, varicose veins, thrombophlebitis and phlebotrombosis, 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,
- 30 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
- 35 anomalies, atelectasis, pulmonary congestion and edema, pulmonary embolism, pulmonary

WO 01/81555

PCT/US01/12992

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 mycoplasma 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 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 storage disorders such as 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, lipomatosis, acute panniculitis, disseminated fat necrosis, adiposis dolorosa, lipoid adrenal hyperplasia, minimal change disease, lipomas, atherosclerosis, hypercholesterolemia, hypercholesterolemia with hypertriglyceridemia, primary hypophospholipoproteinemia, hypothyroidism, renal disease, liver disease, lecithin:cholesterol acyltransferase deficiency, ceroid fetorosis, 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, 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 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 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.

WO 01/81555

PCT/US01/12992

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.

5 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.

10 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.

15 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.

20 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.

25 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 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

WO 01/81555

PCT/US01/12992

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 tSCCP, 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 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.

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

WO 01/81555

PCT/US01/12992

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
 5 transcripts or reverse transcripts of a particular tissue or cell type. In one embodiment, the hybridization takes place in high-throughput format, wherein the polynucleotides of the present invention or their complements comprise a subset of a plurality of elements on a microarray. The resultant transcript image would provide a profile of gene activity.

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

Transcript images which profile the expression of the polynucleotides of the present invention may also be used in conjunction with in vitro model systems and preclinical evaluation of pharmaceuticals, as well as toxicological testing of industrial and naturally-occurring environmental
 15 compounds. All compounds induce characteristic gene expression patterns, frequently termed molecular fingerprints or toxicant signatures, which are indicative of mechanisms of action and toxicity (Nuwaysir, E.F. et al. (1999) Mol. Carcinog. 24:153-159; Steiner, S. and N.L. Anderson (2000) Toxicol. Lett. 112-113:467-471, expressly incorporated by reference herein). If a test compound has a signature similar to that of a compound with known toxicity, it is likely to share
 20 those toxic properties. These fingerprints or signatures are most useful and refined when they contain expression information from a large number of genes and gene families. Ideally, a genome-wide measurement of expression provides the highest quality signature. Even genes whose expression is not altered by any tested compounds are important as well, as the levels of expression of these genes are used to normalize the rest of the expression data. The normalization procedure is useful for
 25 comparison of expression data after treatment with different compounds. While the assignment of gene function to elements of a toxicant signature aids in interpretation of toxicity mechanisms, knowledge of gene function is not necessary for the statistical matching of signatures which leads to prediction of toxicity. (See, for example, Press Release 00-02 from the National Institute of Environmental Health Sciences, released February 29, 2000, available at
 30 <http://www.niehs.nih.gov/oc/news/toxchip.htm>.) Therefore, it is important and desirable in toxicological screening using toxicant signatures to include all expressed gene sequences.

In one embodiment, the toxicity of a test compound is assessed by treating a biological sample containing nucleic acids with the test compound. Nucleic acids that are expressed in the treated biological sample are hybridized with one or more probes specific to the polynucleotides of
 35 the present invention, so that transcript levels corresponding to the polynucleotides of the present

WO 01/81555

PCT/US01/12992

invention may be quantified. The transcript levels in the treated biological sample are compared with levels in an untreated biological sample. Differences in the transcript levels between the two samples are indicative of a toxic response caused by the test compound in the treated sample.

Another particular embodiment relates to the use of the polypeptide sequences of the present invention to analyze the proteome of a tissue or cell type. The term proteome refers to the global pattern of protein expression in a particular tissue or cell type. Each protein component of a proteome can be subjected individually to further analysis. Proteome expression patterns, or profiles, are analyzed by quantifying the number of expressed proteins and their relative abundance under given conditions and at a given time. A profile of a cell's proteome may thus be generated by separating and analyzing the polypeptides of a particular tissue or cell type. In one embodiment, the separation is achieved using two-dimensional gel electrophoresis, in which proteins from a sample are separated by isoelectric focusing in the first dimension, and then according to molecular weight by sodium dodecyl sulfate slab gel electrophoresis in the second dimension (Steiner and Anderson, *supra*). The proteins are visualized in the gel as discrete and uniquely positioned spots, typically by staining the gel with an agent such as Coomassie Blue or silver or fluorescent stains. The optical density of each protein spot is generally proportional to the level of the protein in the sample. The optical densities of equivalently positioned protein spots from different samples, for example, from biological samples either treated or untreated with a test compound or therapeutic agent, are compared to identify any changes in protein spot density related to the treatment. The proteins in the spots are partially sequenced using, for example, standard methods employing chemical or enzymatic cleavage followed by mass spectrometry. The identity of the protein in a spot may be determined by comparing its partial sequence, preferably of at least 5 contiguous amino acid residues, to the polypeptide sequences of the present invention. In some cases, further sequence data may be obtained for definitive protein identification.

A proteomic profile may also be generated using antibodies specific for 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; Mendoza, 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,

WO 01/81555

PCT/US01/12992

N.L. and J. Seilhamer (1997) Electrophoresis 18:533-537), so proteomic 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.

5 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.

10 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.

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

30 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 plasmids (pBACs), etc.

WO 01/81555

PCT/US01/12992

- 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, 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., Greysen, 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.

WO 01/81555

PCT/US01/12992

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.

5 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
10 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.

Without further elaboration, it is believed that one skilled in the art can, using the preceding description, utilize the present invention to its fullest extent. The following preferred specific
15 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. 60/199,021, U.S. Ser. No. 60/200,226, U.S. Ser. No. 60/202,339, U.S. Ser. No. 60/203,505, U.S. Ser. No. 60/205,654, U.S. Ser. No. 60/207,739, and U.S. Ser. No. 60/208,795,
20 are hereby expressly incorporated by reference.

EXAMPLES

I. Construction of cDNA Libraries

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

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

WO 01/81555

PCT/US01/12992

purification kit (Ambion, Austin TX).

In some cases, Stratagene was provided with RNA and constructed the corresponding cDNA libraries. Otherwise, cDNA was synthesized and cDNA libraries were constructed with the UNIZAP vector system (Stratagene) or SUPERSCRIPTM plasmid system (Life Technologies), using the recommended procedures or similar methods known in the art. (See, e.g., Ausubel, 1997, supra, units 5.1-6.6.) Reverse transcription was initiated using oligo d(T) or random primers. Synthetic oligonucleotide adapters were ligated to double stranded cDNA, and the cDNA was digested with the appropriate restriction enzyme or enzymes. For most libraries, the cDNA was size-selected (300-1000 bp) using SEPHACRYL S1000, SEPHAROSE CL2B, or SEPHAROSE CL4B column chromatography (Amersham Pharmacia Biotech) or preparative agarose gel electrophoresis. cDNAs were ligated into compatible restriction enzyme sites of the polylinker of a suitable plasmid, e.g., PBLUESCRIPT plasmid (Stratagene), PSORT1 plasmid (Life Technologies), PCDNA2.1 plasmid (Invitrogen, Carlsbad CA), PBK-CMV plasmid (Stratagene), or pINCY (Incyte Genomics, Palo Alto CA), or derivatives thereof. Recombinant plasmids were transformed into competent *E. coli* cells including XL1-Blue, XL1-BlueMRF, or SOLR from Stratagene or DH15a, DH10B, or ElectroMAX DH10B from Life Technologies.

II. Isolation of cDNA Clones

Plasmids obtained as described in Example I were recovered from host cells by in vivo excision using the UNIZAP vector system (Stratagene) or by cell lysis. Plasmids were purified using at least one of the following: a Magic or WIZARD Minipreps DNA purification system (Promega); an AGTC Miniprep purification kit (Edge Biosystems, Gaithersburg MD); and QIAWELL 8 Plasmid, QIAWELL 8 Plus Plasmid, QIAWELL 8 Ultra Plasmid purification systems or the R.E.A.L. PREP 96 plasmid purification kit from QIAGEN. Following precipitation, plasmids were resuspended in 0.1 ml of distilled water and stored, with or without lyophilization, at 4 °C.

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

III. Sequencing and Analysis

Incyte cDNA recovered in plasmids as described in Example II were sequenced as follows. Sequencing reactions were processed using standard methods or high-throughput instrumentation such as the ABI CATALYST 800 (Applied Biosystems) thermal cycler or the PTC-200 thermal cycler (MJ Research) in conjunction with the HYDRA microdispenser (Robbins Scientific) or the

WO 01/81555

PCT/US01/12992

MICROLAB 2200 (Hamilton) liquid transfer system. cDNA sequencing reactions were prepared using reagents provided by Amersham Pharmacia Biotech or supplied in ABI sequencing kits such as the ABI PRISM BIGDYE Terminator cycle sequencing ready reaction kit (Applied Biosystems). Electrophoretic separation of cDNA sequencing reactions and detection of labeled polynucleotides were carried out using the MEGABACE 1000 DNA sequencing system (Molecular Dynamics); the ABI PRISM 373 or 377 sequencing system (Applied Biosystems) in conjunction with standard ABI protocols and base calling software; or other sequence analysis systems known in the art. Reading frames within the cDNA sequences were identified using standard methods (reviewed in Ausubel, 1997, *supra*, unit 7.7). Some of the cDNA sequences were selected for extension using the techniques disclosed in Example VIII.

The polynucleotide sequences derived from Incyte cDNAs were validated by removing vector, linker, and poly(A) sequences and by masking ambiguous bases, using algorithms and programs based on BLAST, dynamic programming, and dinucleotide nearest neighbor analysis. The Incyte cDNA sequences or translations thereof were then queried against a selection of public databases such as the GenBank primate, rodent, mammalian, vertebrate, and eukaryote databases, and BLOCKS, PRINTS, DOMO, PRODOM, and hidden Markov model (HMM)-based protein family databases such as PFAM. (HMM is a probabilistic approach which analyzes consensus primary structures of gene families. See, for example, Eddy, S.R. (1996) *Curr. Opin. Struct. Biol.* 6:361-365.) The queries were performed using programs based on BLAST, FASTA, BLIMPS, and HMMER. The Incyte cDNA sequences were assembled to produce full length polynucleotide sequences. Alternatively, GenBank cDNAs, GenBank ESTs, stitched sequences, stretched sequences, or Genscan-predicted coding sequences (see Examples IV and V) were used to extend Incyte cDNA assemblies to full length. Assembly was performed using programs based on Phred, Phrap, and Consed, and cDNA assemblies 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.

WO 01/81555

PCT/US01/12992

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:19-36. 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 gbbtg). Genscan is a general-purpose gene identification program which analyzes genomic DNA sequences from a variety of organisms (See Burge, C. and S. Karlin (1997) J. Mol. Biol. 268:78-94, and Burge, C. and S. Karlin (1998) Curr. Opin. Struct. Biol. 8:346-354). The program concatenates predicted exons to form an assembled cDNA sequence extending from a methionine to a stop codon. The output of Genscan is a FASTA database of polynucleotide and polypeptide sequences. The maximum range of sequence for Genscan to analyze at once was set to 30 kb. To determine which of these Genscan predicted cDNA sequences encode human kinases, the encoded polypeptides were analyzed by querying against PFAM models for kinases. Potential human kinases were also identified by homology to Incyte cDNA sequences that had been annotated as kinases. These selected Genscan-predicted sequences were then compared by BLAST analysis to the gonpept and gbpr1 public databases. Where necessary, the Genscan-predicted sequences were then edited by comparison to the top BLAST hit from gonpept to correct errors in the sequence predicted by Genscan, such as extra or omitted exons. BLAST analysis was also used to find any Incyte cDNA or public cDNA coverage of the Genscan-predicted sequences, thus providing evidence for transcription. When Incyte cDNA coverage was available, this information was used to correct or confirm the Genscan predicted sequence. Full length polynucleotide sequences were obtained by assembling Genscan-predicted coding sequences with Incyte cDNA sequences and/or public cDNA sequences using the assembly process described in Example III. Alternatively, full length polynucleotide sequences were derived entirely from edited or unedited Genscan-predicted coding sequences.

V. Assembly of Genomic Sequence Data with cDNA Sequence Data

WO 01/81555

PCT/US01/12992

"Stitched" Sequences

Partial cDNA sequences were extended with exons predicted by the Genscan gene identification program described in Example IV. Partial cDNAs assembled as described in Example III were mapped to genomic DNA and parsed into clusters containing related cDNAs and Genscan exon predictions from one or more genomic sequences. Each cluster was analyzed using an algorithm based on graph theory and dynamic programming to integrate cDNA and genomic information, generating possible splice variants that were subsequently confirmed, edited, or extended to create a full length sequence. Sequence intervals in which the entire length of the interval was present on more than one sequence in the cluster were identified, and intervals thus identified were considered to be equivalent by transitivity. For example, if an interval was present on a cDNA and two genomic sequences, then all three intervals were considered to be equivalent. This process allows unrelated but consecutive genomic sequences to be brought together, bridged by cDNA sequence. Intervals thus identified were then "stitched" together by the stitching algorithm in the order that they appear along their parent sequences to generate the longest possible sequence, as well as sequence variants.

Linkages between intervals which proceed along one type of parent sequence (cDNA to cDNA or genomic sequence to genomic sequence) were given preference over linkages which change parent type (cDNA to genomic sequence). The resultant stitched sequences were translated and compared by BLAST analysis to the *genpept* and *gbpri* public databases. Incorrect exons predicted by Genscan were corrected by comparison to the top BLAST hit from *genpept*. Sequences were further extended with additional cDNA sequences, or by inspection of genomic DNA, when necessary.

"Stretched" Sequences

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

VI. Chromosomal Mapping of PKIN Encoding Polynucleotides

The sequences which were used to assemble SEQ ID NO:19-36 were compared with

WO 01/81555

PCT/US01/12992

sequences from the Incyte LIFSEQ database and public domain databases using BLAST and other implementations of the Smith-Waterman algorithm. Sequences from these databases that matched SEQ ID NO:19-36 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 Génethon 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 Génethon which provide boundaries for radiation hybrid markers whose sequences were included in each of the clusters. Human genome maps and other resources available to the public, such as the NCBI "GeneMap'99" World Wide Web site (<http://www.ncbi.nlm.nih.gov/genemap/>), can be employed to determine if previously identified disease genes map within or in proximity to the intervals indicated above.

In this manner, SEQ ID NO:24 was mapped to chromosome 2 within the interval from 92.30 to 103.1 centiMorgans, SEQ ID NO:25 was mapped to chromosome 11 within the interval from 104.8 to 117.9 centiMorgans, SEQ ID NO:33 was mapped to chromosome 8 within the interval from 25.8 to 40.3 centiMorgans, SEQ ID NO:23 was mapped to chromosome 9 within the interval from 101.20 to 104.90 centiMorgans, to chromosome 10 within the interval from 145.20 to 156.60 centiMorgans, and to chromosome 19 within the interval from 69.90 to 81.20 centiMorgans. More than one map location is reported for SEQ ID NO:23, indicating that sequences having different map locations were assembled into a single cluster. This situation occurs when sequences having strong similarity, but not complete identity, are assembled into a single cluster.

VII. Analysis of Polynucleotide Expression

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

Analogous computer techniques applying BLAST were used to search for identical or related molecules in cDNA databases such as GenBank or LIFSEQ (Incyte Genomics). This analysis is

WO 01/81555

PCT/US01/12992

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:

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

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

Alternatively, polynucleotide sequences encoding PKIN are analyzed with respect to the tissue sources from which they were derived. For example, some full length sequences are assembled, at least in part, with overlapping Incyte cDNA sequences (see Example 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; hematologic 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 LIFESEQ GOLD database (Incyte Genomics, Palo Alto CA).

WO 01/81555

PCT/US01/12992

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 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

WO 01/81555

PCT/US01/12992

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.

- 5 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
- 10 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
- 15 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:19-36 are employed to screen cDNAs, genomic DNAs, or mRNAs. Although the labeling of oligonucleotides, consisting of about 20 base
- 20 pairs, is specifically described, essentially the same procedure is used with larger nucleotide fragments. Oligonucleotides are designed using state-of-the-art software such as OLIGO 4.06 software (National Biosciences) and labeled by combining 50 pmol of each oligomer, 250 µCi of [γ -³²P] adenosine triphosphate (Amersham Pharmacia Biotech), and T₄ polynucleotide kinase (DuPont NEN, Boston MA). The labeled oligonucleotides are substantially purified using a
- 25 SEPHADEX G-25 superfine size exclusion dextran bead column (Amersham Pharmacia Biotech). An aliquot containing 10⁷ 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
- 30 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.

- 35 X. Microarrays

WO 01/81555

PCT/US01/12992

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),

5 *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.,

10 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

15 array elements are hybridized with polynucleotides in a biological sample. The polynucleotides in the biological sample are conjugated to a fluorescent label or other molecular tag for ease of detection. After hybridization, nonhybridized nucleotides from the biological sample are removed, and a fluorescence scanner is used to detect hybridization at each array element. Alternatively, laser desorption and mass spectrometry may be used for detection of hybridization. The degree of

20 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

25 poly(A)⁺ RNA is purified using the oligo-(dT) cellulose method. Each poly(A)⁺ RNA sample is reverse transcribed using MMLV reverse-transcriptase, 0.05 pg/μl oligo-(dT) primer (21mer), 1X first strand buffer, 0.03 units/μl RNase inhibitor, 500 μM dATP, 500 μM dGTP, 500 μM dTTP, 40 μM dCTP, 40 μM dCTP-Cy3 (BDS) or dCTP-Cy5 (Amersham Pharmacia Biotech). The reverse transcription reaction is performed in a 25 ml volume containing 200 ng poly(A)⁺ RNA with

30 GEMBRIGHT kits (Incyte). Specific control poly(A)⁺ RNAs are synthesized by *in vitro* transcription from non-coding yeast genomic DNA. After incubation at 37°C for 2 hr, each reaction sample (one with Cy3 and another with Cy5 labeling) is treated with 2.5 ml of 0.5M sodium hydroxide and incubated for 20 minutes at 85°C to the stop the reaction and degrade the RNA. Samples are purified using two successive CHROMA SPIN 30 gel filtration spin columns (CLONTECH Laboratories, Inc.

WO 01/81555

PCT/US01/12992

(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.

5 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 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

WO 01/81555

PCT/US01/12992

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
 5 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
 10 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
 15 emission maxima of the fluorophores used are 565 nm for Cy3 and 650 nm for Cy5. Each array is
 typically scanned twice, one scan per fluorophore using the appropriate filters at the laser source,
 although the apparatus is capable of recording the spectra from both fluorophores simultaneously.

The sensitivity of the scans is typically calibrated using the signal intensity generated by a
 cDNA control species added to the sample mixture at a known concentration. A specific location on
 20 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
 25 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
 30 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

WO 01/81555

PCT/US01/12992

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 (Jacyte).

XI. Complementary Polynucleotides

Sequences complementary to the PKIN-encoding sequences, or any parts thereof, are used to
 5 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
 10 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
 15 containing an antibiotic resistance gene and an inducible promoter that directs high levels of cDNA transcription. Examples of such promoters include, but are not limited to, the *trp-lac* (*lac*) 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-
 20 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
 25 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 Hingelbard, B.K. et al. (1994) Proc. Natl. Acad. Sci. USA 91:3224-3227; Sandig, V. et al. (1996) Hum. Gene Ther. 7:1937-1945.)

30 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
 35 Pharmacia Biotech). Following purification, the GST moiety can be proteolytically cleaved from

WO 01/81555

PCT/US01/12992

PKIN at specifically engineered sites. FLAG, an 8-amino acid peptide, enables immunaffinity 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, and XVIII 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 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

WO 01/81555

PCT/US01/12992

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-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 anti-peptide 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

WO 01/81555

PCT/US01/12992

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).

- 5 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

- 10 Generally, protein kinase activity is measured by quantifying the phosphorylation of a protein substrate by PKIN in the presence of gamma-labeled ^{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
15 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
20 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
25 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 are as follows: 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
30 ERK (Pearson, J.D. et al. (1991) *Methods Enzymol.* 200:62-81).

In another alternative, protein kinase activity of PKIN is demonstrated *in vitro* in an assay containing PKIN, 50 μl of kinase buffer, 1 μg substrate, such as myelin basic protein (MBP) or synthetic peptide substrates, 1 mM DTT, 10 μg ATP, and 0.5 μCi [γ - ^{32}P]ATP. The reaction is incubated at 30°C for 30 minutes and stopped by pipetting onto P81 paper. The unincorporated [γ -

WO 01/81555

PCT/US01/12992

³²P]ATP is removed by washing and the incorporated radioactivity is measured using a radioactivity scintillation counter. Alternatively, the reaction is stopped by heating to 100°C in the presence of SDS loading buffer and visualized on a 12% SDS polyacrylamide gel by autoradiography.

Incorporated radioactivity is corrected for reactions carried out in the absence of PKIN or in the presence of the inactive kinase, K38A.

In yet another alternative, adenylate kinase or guanylate kinase activity may be measured by the incorporation of ³²P from gamma-labeled ³²P-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 ³²P-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 to gel electrophoresis to separate the mono-, di-, and triphosphonucleotide fractions. The diphosphonucleotide fraction is cut out 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.

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 01/81555

PCT/US01/12992

Table 1

Project ID	Polypeptide Seq ID No:	Protein Seq ID No:	Protein Seq ID No:	Protein Seq ID No:
7472632	1	7472632	1	7472632
7472633	2	7472633	2	7472633
7472634	3	7472634	3	7472634
7472635	4	7472635	4	7472635
7472636	5	7472636	5	7472636
7472637	6	7472637	6	7472637
7472638	7	7472638	7	7472638
7472639	8	7472639	8	7472639
7472640	9	7472640	9	7472640
7472641	10	7472641	10	7472641
7472642	11	7472642	11	7472642
7472643	12	7472643	12	7472643
7472644	13	7472644	13	7472644
7472645	14	7472645	14	7472645
7472646	15	7472646	15	7472646
7472647	16	7472647	16	7472647
7472648	17	7472648	17	7472648
7472649	18	7472649	18	7472649

WO 01/81555

PCT/US01/12992

Table 2

Protein ID	Protein Name	Genebank	Probability	GenBank Homology
1	Protein kinase Msk-1 (Mus musculus)	U03153.1	0.00	Protein kinase Msk-1 (Mus musculus)
2	Cam kinase 1 (Homo sapiens)	U03153.1	2.4e-144	Cam kinase 1 (Homo sapiens)
3	K25 protein kinase (Rattus norvegicus) (Maucuer A. et al. 1997)	U03153.1	2.3e-236	K25 protein kinase (Rattus norvegicus) (Maucuer A. et al. 1997)
4	male germ cell-associated kinase (maki) (Rattus norvegicus) (Matsushima H. et al. 2000)	U03153.1	6.4e-293	male germ cell-associated kinase (maki) (Rattus norvegicus) (Matsushima H. et al. 2000)
5	Protein kinase C- α (Homo sapiens) (Schwartz P. et al. 1994)	U03153.1	0.0	Protein kinase C- α (Homo sapiens) (Schwartz P. et al. 1994)
6	Arabidopsis thaliana SMF1 family protein kinase	U03153.1	3.3e-26	Arabidopsis thaliana SMF1 family protein kinase
7	[Gallus gallus] gdn-induced kinase	U03153.1	1.2e-144	[Gallus gallus] gdn-induced kinase
8	[Rattus norvegicus] p70 α mixed lineage kinase 2 (Homo sapiens)	U03153.1	2.4e-199	[Rattus norvegicus] p70 α mixed lineage kinase 2 (Homo sapiens)
9	Protein kinase C- α (Homo sapiens) (Schwartz P. et al. 1994)	U03153.1	3.9e-231	Protein kinase C- α (Homo sapiens) (Schwartz P. et al. 1994)
10	Adenylate kinase 5 (Homo sapiens) (Van Rompay A.R. et al. 1995)	U03153.1	7.3e-102	Adenylate kinase 5 (Homo sapiens) (Van Rompay A.R. et al. 1995)
11	Identification of a novel human adenylate kinase cDNA cloning. expression analysis, chromosome localization and characterization of the complementary protein	U03153.1	8.9e-145	Identification of a novel human adenylate kinase cDNA cloning. expression analysis, chromosome localization and characterization of the complementary protein

WO 01/81555

PCT/US01/12992

Table 2 (cont.)

Protein ID	Protein Name	GenBank ID	Probability	GeneBank Homology
12	747500CD1	926577	2.4e-112	nucleoside triphosphate-adenylate kinase [Homo sapiens] Xu, G. et al. (1997). Characterization of human adenylylate kinase 3 (AK3) cDNA and mapping of the AK3 pseudogene to an intron of the NF1 gene. Genomics 13:517-542.
13	747549CD1	92257583	1.4e-210	ECTAIR3 [Rattus rattus] Kurose, T. et al. (1997) ECTAIR 2, a novel serine/threonine kinase, is predominantly expressed in terminally differentiated neurons. Eur. J. Biochem. 249:481-488.
14	2132119CD1	93880563	2.0e-121	Predicted using GeneFinder; similar to serine/threonine kinase; cDNA [Caenorhabditis elegans]. Wang, C. et al. (1998) Serine/threonine kinase-18220182018.
14	2132119CD1	910442581	0	105-kDa kinase-like protein [Mus musculus] Lau, S.C.H., et al., (2000)
15	7474496CD1	96066385	0.0	Biochim. Biophys. Acta 1517:148-152
16	1834248CD1	97595902	1.40E-252	GCM2 G1P2alpha kinase [Mus musculus] EMBL motif kinase 2 short form [Mus musculus]
17	71584520CD1	93927912	3.9e-157	cytochrome b5 binding protein kinase [Mus musculus] Cortez, A. et al. (1999)
18	7475538CD1	94090958	9.2e-85	FERB Lett. 443: 370-374 cell cycle related kinase [Homo sapiens]
18	7475538CD1	95664226	0	CDK-related protein kinase FQGLARR [Mus musculus]

Table 3

[illegible]

WO 01/81555

PCT/US01/12992

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 Databases
3	310793CD1	419	S181 S215 S231 S244 S330 T390	S181 S187 S215 S231 S244	Transmembrane domain: V223-L241 Eukaryotic protein kinase domain: W21-F304 Tyrosine Kinase catalytic domain: PR00109B:F148-W156 PR00109D:V223-V245 PR00109E:P273-K295 SERINE/THREONINE PROTEIN KINASE PR00109A:V223-F304 SERINE/THREONINE PROTEIN KINASE PR00109C:V223-F304 RIBONUCLEOPROTEIN REPEAT DN00012:A48249-299-407:P319-7404 PROTEIN KINASE DOMAIN DN00004:P08414 44-285:A74-1293 DN00004:P32485 34-292:846-2289 DN00004:P39657 301-409:C106-2241	HMMER_Pfam BLASTS_PfamMS BLAST_PRODOM BLAST_DOMO BLAST_DOMO

Table 3 (cont.)

[illegible]

Table 3 (cont.)

[illegible]

WO 01/81555

PCT/US01/12952

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 Databases
6	155145CD1	440	S85 T210 T237 S361 S365 S376 S396 T426 S11 S219 S131 S155 T214 S251 Y68 Y378		Eukaryotic protein kinase domain: L164-E331 Tyrosine kinase signature: A192-L210, S260-S282 PROTEIN KINASE DOMAIN: DM00004 P34244 82-359; Y167-L324 Protein Kinase motif: T198-L210	HMMER-PFAM BLIIPS-PRINTS BLAST-DOMO MOTIFS
7	2585155CD1	923	Y106 Y476 T256 S125 T688 S284 T38 S125 S288 T356 S391 T447 S448 T455 T481 S500 S512 T875 S890 T17 S254 S336 T417 S472 S221 S543 S551 S584 S576 S667	M12 N317	PROTEIN KINASE DOMAIN: S25-L259 PROTEIN KINASE DOMAIN: S25-L259 Sectine/Protein kinase active-site signature Protein Kinase St. I: I195-L147 Eukaryotic protein kinase domain Kinase: S24-M268 Protein Kinases signatures and profile Protein Kinase: EY-Prf: Y50-G167 Tyrosine Kinase catalytic site S40-L106 SH3 domain signature: PR00452; C167-S159, A113-Q130, D132-L141 PROTEIN DOMAIN SH3 KINASE GUANYLATE TRANSFERASE ATP-BINDING REPEAT CMP MEMBRANE PD01338: T267-Q360 GUANYLATE KINASE DM00755 A57653 370- 570, P229-P431 P102 signaling molecule domain PD6: C13-L93	BLIIPS-PRINTS BLAST-PROTON BLAST-DOMO HMMER-PFAM BLIIPS-PRINTS
8	435711CD1	442	T16 T267 T278 S395 S361 S142 S229 S178 S190 T311 S408 S437 Y234 Y364 Y346	P406	GUANYLATE KINASE: KIN: G266-Y372 Guanylate kinase protein BL00856: Y264-Y284, L292-F319	HMMER-PFAM BLIIPS-PRINTS

Table 3 (cont.)

[illegible]

Table 3 (cont.)

EBQ ID	Protein ID	Amino Acid Residues	Potential Respiratory Location	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Notes
7470560CD1	357	S53, V65, Y103, S124, T140, S143, S181, T185, S200, S211, S213, S245, Y288, T304, S316, Y340, Y350			KINASE ADENYLATE TRANSFERASE ATP-BINDING AMP/ANP TRANSFEROSHYDROLASE ISOMERING PROTEIN 3D-STRUCTURE MITOCHONDRIUM: P200657; Y175-R294 ADENYLATE KINASE: DMO0290 [P00576]_131; E165-E194 Adenylate kinase protein: BL00113A; P201719; P20180; H198-A241; P201130; C247-V261; BL001130; S298-T328	BIAST-DOMO
					Adenylate kinase signature: PR00094A; F174-C187; PH00094B: G202-S216; PR00094C: F252-G268; PR00094D: P300-Y315; PR00094E: R317-Q331	SLIMPS-BLOCKS
					Uridine kinase signature: PR00388A; C170-C187	SLIMPS-PRINTS
					Serine/threonine kinase family: PR01100A; P201100	SLIMPS-PRINTS
					Adenylate Kinase: I9-Q20, P252-Q261	MOTIFS
					Adenylate kinase: I11-L68, Y172-Q311	BLIMP-PFAM
					Adenylate kinase signature	PROFILESCAN
					(adenylate kinase pfam) Y229-A284	BIAST-DOMO
					PROTEIN KINASE DOMAIN: PR000104 [S5344]/21-266; E24-C270	MOTIFS
					Protein Kinase AMP-binding region Signature: L29-F52.	MOTIFS
					Serine/threonine protein kinases Active site signature: I140-I152. Serine/threonine kinase domain: P23-L279	MOTIFS
					Protein kinases signatures and profile (protein_kinase_bv.fam) Tyrosine kinase catalytic domain: PR00109A: M98-V111; PR00109B: Y134-Y155; PR00109E: V202-E224	PROFILESCAN
					PROTEIN KINASE: CHLARDULREINLING I, CHLARDULREINLING II, CHLARDULREINLING III TRANSFERASE SERINE/THREONINE/PROTEIN PHOSPHORYLATION: PR01217; W276-L322	BIAST-PRODAN

Table 3 (cont.)

SEQ ID NO	Incyte ID	Anticancer peptide	Anticancer peptide	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
12	7475509CD1	24	24	S12, S16, S18, S20, S32, T30L		ADENYLATE KINASE: D600730; P27144 [1-125]; V14-R126 ADENYLATE KINASE SIGNATURE: P00094A4: V9-C72; P00094B: G33-G51; P00094C: M95-D104; P00094D: Q459-Y174; P00094B: D176-Y190 SRLKimate Kinase family: P00100A: AG-Q23; P00100B: D107-S124 [p<0.0028], AG-GLUTE KINASE MOTIF: W85-Q96 KINASE ADENYLATE TRANSFERASE, ATP-BINDING ATP/AMP TRANSFERASE, ISOENZYM, PROTEIN 3D-STRUCTURE NUTRONDRIN: P000657: T10-V190 Adenylate Kinase: T10-V190 Adenylate Kinase protein: E000133a: V9-223; E00133b: R33-R76; E000133c: R000133a: V119-V159; R000133b: V159-223 Nucleotide Kinase signature: S267 ADENYLATE KINASE DEF1: V64-F156 ADENYLATE KINASE DOMAIN: D600014: D04929-122-332; V173-M443 KINASE SERINE/THREONINE RGTAIN TRANSFERASE ATP-BINDING DOMAIN - PCTAIRE1, PCTAIRE2, PCTAIRE3, QKX5 ALTERNATIVE: P0007333; D120-T71 Tyrosine Kinase catalytic domain: P00109B: Y483-T301 Protein Kinase signatures and profile (Protein Kinase TyrDef1): Q235-A317 Eukaryotic protein kinase domain (kinase): V112-F453 Structure: Nucleotide binding region (kinase): V178-K201 Serine/threonine protein kinases active-site signature: I383-I301	BLAST-PMO BLINFS-PRINTS PROTEUSCAN BLAST-IGMO BLAST-PRODOM PCTAIRE1, PCTAIRE2, PCTAIRE3, QKX5 ALTERNATIVE: P0007333; D120-T71 Tyrosine Kinase catalytic domain: P00109B: Y483-T301 Protein Kinase signatures and profile (Protein Kinase TyrDef1): Q235-A317 Eukaryotic protein kinase domain (kinase): V112-F453 Structure: Nucleotide binding region (kinase): V178-K201 Serine/threonine protein kinases active-site signature: I383-I301

Table 3 (cont.)

Seq ID	Protein	Antibody	Potential Respiratory Protein	Potential Signaling Protein	Potential Signatures	Analytical Methods and Database
15	7474436CDL	1651	7555, 782	8100, 8205	PROTEIN KINASE DOMAIN: E000004.P15442.635-902; F732-A994	BLAST-DB
			S108, S144	1207, 1197	EUROPEAN INFLUENZA VIRUS KINASE	BLAST-PROTEIN
			S210, T315	1203, 1248	(EPF ALKIN): P0156018; D13-D219	BLAST-PROTEIN
			S216, S247	1416, 1520	Protein kinase catalytic domain:	BLAST-PROTEIN
			S250, T329		P001098; Y840-L858	PROTEIN-PRINTS
			T346, T409		Protein kinases signatures and profile	PROTEIN-PRINTS
			S14, T441		(protein kinase Tyr-Prf): R326-A877	PROTEIN-PRINTS
			S450, T478		Bakaryotic protein kinase domain	PROTEIN-PRINTS
			S461, S480		(protein kinase Tyr-Prf): R326-A877	PROTEIN-PRINTS
			S709, T115	S722, S732	Protein kinases signatures and profile	PROTEIN-PRINTS
			S733, S755	S739, T115	Protein kinases signatures and profile	PROTEIN-PRINTS
			S757, T811		Protein kinases signatures and profile	PROTEIN-PRINTS
			T824, S882		Protein kinases signatures and profile	PROTEIN-PRINTS
			S917, T945		Protein kinases signatures and profile	PROTEIN-PRINTS
			S937, S961		Protein kinases signatures and profile	PROTEIN-PRINTS
			S973, S997		Protein kinases signatures and profile	PROTEIN-PRINTS
			S1038, S1059		Protein kinases signatures and profile	PROTEIN-PRINTS
			T1062, T1152		Protein kinases signatures and profile	PROTEIN-PRINTS
			T1234, T1318		Protein kinases signatures and profile	PROTEIN-PRINTS
			S1397, T1418		Protein kinases signatures and profile	PROTEIN-PRINTS
			S1437, S1457		Protein kinases signatures and profile	PROTEIN-PRINTS
			S1572, T1475		Protein kinases signatures and profile	PROTEIN-PRINTS
			S1547, T1508		Protein kinases signatures and profile	PROTEIN-PRINTS
			T1604, S1601		Protein kinases signatures and profile	PROTEIN-PRINTS
			T1626, T1652		Protein kinases signatures and profile	PROTEIN-PRINTS
			T1626, T1652		Protein kinases signatures and profile	PROTEIN-PRINTS

Table 3 (cont.)

SEQ ID NO	IncPte ID	Amino Acid Residue	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Methods and databases
16	1834248CD1	752	S339 S2 S210	W395 W532	Extracytic protein kinase domain: Y35-Y339.	BLAST_PRODM
		S23 S27 S34			Tyrosine kinase catalytic domain	BLAST_PRODM
		S354 G399 S423			PRO01094: M135-V148	
		S441 S458 S48			PRO01092: V137-V189	
		S494 S461 S656			PRO01090: V237-H259	
		S710 T127 T281			Protein kinase signatures and profile	PROFILESCAN
		T200 T323 T332			protein_kinase_bty_bf: Y112-S210	
		T344 T507 T511			Protein kinases ATP-binding region:	MOTIFS
		T515 T536 T564			Protein_Kinase_ATP: I65-K88	
		T620 T624 T61			Serine/Threonine protein kinases	MOTIFS
					active-site Protein_Kinase_St:	
					T177-V189	
					KINASE_SERINE/THREONINE_TRANSFERASE	BLAST_PRODM
					ATP_BINDING_PROTEIN_EMR_F78 CDC25C	
					PRO08471: S412-P532	
					KINASE_SERINE/THREONINE_TRANSFERASE	BLAST_PRODM
					ATP_BINDING_PROTEIN_PAM1 AP78 EMR	
					PRO05818: I310-S310	
					KINASE_SERINE/THREONINE_TRANSFERASE	BLAST_PRODM
					ATP_BINDING_KIN1_EMR_PAM1	
					PRO04300: E450-L752	
					KINASE_TRANSFERASE_ATP_BINDING_SERINE/THREONINE_PHOSPHORYLATION_RECEPTOR	BLAST_PRODM
					TYROSINE	
					PRO000001: Y39-V137	
					PROTEIN_KINASE_DOMAIN	BLAST_PRODM
					PRO00004 P27448 18-297 L61-L301	
					DMO0004 I248609 25-294 L61-L301	
					DMO0004 Q05512 25-294 L61-L301	
					DMO0004 JCI446 20-261 P60-L301	

Table 3 (cont.)

Accession NO.	Type	Amino Acid Residues	Potential Polymerization Sites	Potential Phosphorylation Sites	Signature Sequences, Domains and Motifs	Analytical Method and Database
D000004	Polypeptide	501	S118 S138 S292		Eukaryotic protein kinase domain	HMMER_PPAK
			S341 S364 S482		kinase; E37-I286	
			S483 S495 T103		Tyrosine Kinase catalytic domain	HAJMES_PRINTS
			F21 T276 W422		PR00109B; Y135-Y153	
			T46 T470 F51 F77		PR001D9D; V201-E223	
			T91 Y135 Y491		CALMODULIN BINDING PROTEIN	BLAST_PRODOM
					ZD05986Z; G360-V443	
					PROTEIN KINASE CALMODULIN BINDING I	BLAST_PRODOM
					CALCIUM/CALMODULIN DEPENDENT TYPE CAL	
					TRANSFERASE SERINE/THREONINE PROTEIN	
					PHOSPHORYLATION	
					PD012137; W285-A335	
					CALMODULIN BINDING PROTEIN	BLAST_PRODOM
					PD050813; M1-E34	
					PROTEIN KINASE DOMAIN	BLAST_DOMO
					DM000041;S57347 11-266;D25-T276	
					DX000041;P08414 44-285;F38-T276	
					DM000041;P11798 115-261;C36-A277	
					DM000041;A44412 116-262;F35-A277	

Table 3 (cont.)

393Q ID	Isotype ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Database
18	7475538CD1	346	S241 Y176 Y215		Eukaryotic protein kinase domain kinase: Y4-P288. Tyrosine kinase catalytic domain P800109P; F117-I135 Protein kinase signatures and profile protein kinase Tyr pf. A69-D125 Protein Kinase ATP binding region signature Protein_Kinase_Atp. I10-K33 Serine/threonine protein kinases active-site: Protein_Kinase_Sr: T123-T135	ROMER_PPIM
					MOtifs	
					kinase TRANSFERASE SERINE/THREONINE ATP BINDING II PHOSPHORYLATION CASEIN ALPHA CHAIN EP002408: V164-P288	BLAST_PRODOM
					KINASE PROTEIN TRANSFERASE ATP BINDING SERINE/THREONINE PHOSPHORYLATION RECEPTOR TYROSINE TRANSMEMBRANE PD000001: Y169-P301 PROTEIN KINASE DOMAINS	BLAST_PRODOM
					DN00004 P29620/21-289 I10-A279 DN00004 Q00526 I6-286 A9-P288 DN00004 P43450 I6-276 A9-A279 DN00004 P23437 I6-296 I9-P286	BLAST_DOWNO

WO 01/81555

PCT/US01/12992

Table 4

Poly-nucleotide Seq. ID No.	inocyte polynucleotide	Sequence length	Selected Fragment(s)	Sequence Fragments	5' Position	3' Position
19	2690544CBL	4324	L-1372, 4153-4224	6274032GH (P1A222B01) 71085659V1 71083920V1 71083254V1 71084600V1 7286579H1 (BFA-FR06) 71254276V1 598023377 (WCLPTC02) 7086516H1 (BFA-FR03) 71252928V1 7082857V1 713225H1 (LJHFR02) 6056539H1 60529286 (CONNO701) 70544947 70518453D1 7052805D1 7108445H1 (MGR1DT01) 2295923F6 (UYANNV04) 3459433F6 (333NF-T01) 2927552E6 (UYANNV04) 3107952F6 (HSTTUT15) 5922389H1 (PLACFR06) 7739295H1 (P1A222B01) 554420P6 (TESTNOC01) 4205186F6 (HCONNO702) 68N132400G (J04-EB1E) 2211548E6 (UYANNV04) 526829F1 (UYANNV04) 6909108N1 (31WDLR01) 6910588N1 (31WDLR01) 2940465H1 (TAVMSF02) 7274331H2 (R1DSTA01) 50205600V1	545 281 2916 2863 2351 1 1091 3498 222 1680 228 168 183 385 716 1172 1172 1215 715 582 1309 507 1458 646 807 2041 284 1617 1644 1138 1994 2682 1876 559	1204 3638 3576 2853 2959 408 1625 4324 702 2356 226 7394 385 1182 1807 1215 582 1824 1451 1597 834 1157 2201 1821 1720 2129 2577 2956 2503 1091
20	7472693CBL	1736	72-149, 1317-1736			
21	3107952CBL	1824	1802-1824			
22	554420CBL	2201	1245-1629, 2128-2201, 245-376			
23	7472693CBL	2974	2331-2974, 1-654			

100

Polym. ID No.	Incyte Polymerside ID	Sequence Length	Selected Fragment (s)	Sequence Fragments	Position	Position
23				60205598T1 (UMKMN071)	600	1339
				6426807H1 (UMKMN071)	2411	2951
				3343032H1 (SEANNT09)	2708	2958
				5812278H1 (SSIPWR01)	1	647
				6910588H1 (PIIDTR01)	1369	2003
				91860144	2532	2974
24	1551456C01	3640	1906-2052, 1-752	73230359 (BGTFTT08)	1230	1731
				6330947H1 (SPNR01)	2157	3121
				6834599H1 (SPNR02)	1630	3130
				1388125H1 (BSPN01)	1	266
				148035652 (SPCNFT01)	3139	3648
				6870364H1 (SPNFF02)	3174	3648
				711408571 (SPNFF02)	298	950
				7146056H1 (ADRTTB04)	875	1491
				3296163F6 (WTG7INT01)	1553	2310
				70349397D1	3030	3648
				70872871V1	923	1599
				70875814V1	2218	2955
25	2599355C01	4719	1839-1863, 1-1135, 2380-3186	6708033H1 (PROST052)	2498	3226
				211395953 (GKANK001)	3264	3885
				64875653 (MAYAB02)	2719	3270
				70985467H1	388	950
				7351743H1 (PESATW01)	810	1299
				6864125H1 (BBAQCN02)	1323	1998
				6056925H1 (BBAQNT04)	1322	1735
				6701992H1 (HEADTR01)	4080	4719
				7652463H2 (SPNWDTR01)	997	1704
				7035646H1 (PESATR02)	1	566
				7357374H1 (SPNFT01)	1977	2550
				2483256H1 (BBAQNT02)	380	4348
				924333H1 (BBAQNT02)	4009	4386
				GS 4357117 (BBAQNT02)	11	1229
26	4357117C01	1651	1-267	613159966 (BBAQNT02)	1	1451

WO 01/81555

PCT/US01/12992

Table 4 (cont.)

Poly-nucleotide SEQ ID NO.	Incyte Polynucleotide	Sequence Length	Selected Fragment (g)	Sequence Fragments	Position	Position
27	5511992CB1	3141	3572-1833 71-157, 665-828, 1298-1538, 2580-2878	GS-5511992.fasta	1	3141
28	7474560CB1	1244	1203-1244, 1-131	6311370H1 (NEROTD02), 6855555H1 (BRATFEN08), 6997205H1 (BRATFEN08), 7051853J1	616 1 577 1	1236 683 1244 1171
29	7474602CB1	1661	72-149	6772112J1 (BRANDOR01), 3462757 (HYHMOCT02), 6124350H1 (BRANDOR05), 944796H1 (SATENCOT2), 3616204F6 (BRFDMOT01), 71157872V1	1 1043 634 1 69	640 1661 1180 92 699
30	7475509CB1	512	881-912	7332424H1 (BRATFEN08), 7004062H1 (COLMFE01), 7677070J1 (GUSSTF02), 7746933J1 (SINFEE02), 7843286 (BRATFEN02), 7851233H1 (BRANDOR03), 7824460H1 (BRATFEN02), 7060059V1	193 1831 1967 1 1293 1000 1868 2179 654	912 2429 2485 564 1877 1375 1868 2186 1111
32	2192119CB1	2817	1-249	1687835P6 (PROSTUT03), 6769461J1 (BRANDOR01), 2613105P6 (PROSCT20), 7716523H1 (SINFEE02), 1796441P6 (PROSTUT05), 1618475P6 (BRATFEN02), 7673143J1 (GUSSTF02), 6742286H1 (BRANDOR03), 2499633H1 (BRANDOR04), 7846563H1 (UTRSTF01), 8700960	2416 221 1052 498 2186 2117 1130 1543 1855 267	2817 385 1721 1074 2783 2783 1719 2170 2106 1052 385
33	7474496CB1	5305	3563-3753, 1403-1444, 2384-2444, 3164-3305	7086570V1	3352	3875

WO 01/81555

PCT/US01/12992

Table 4 (cont.)

Poly-nucleotide SEQ ID NO.	Incyte polynucleotide ID	Sequence Length	Selected Fragment (s)	Sequence Fragments	Position	Position
33				43932696 (BRADN0201)	2735	3311
				7088573491	4100	4744
				1830377766 (THEIA7001)	3649	4093
				648846481 (MIXDUR01)	2852	3473
				46819085 (HNT245701)	2117	2865
				7088847601	4651	5305
34	1834248CE1	3269	1754-1773, 1-142, 3072-3269	18325606 (BRAIN001)	4034	4547
				GRN 9713981_000025_004	1293	2275
				6146293H1 (BRADIT01)	1791	2354
				67223852 (BSAIFR03)	2635	3263
				69300401 (BSAIFR03)	684	2440
				16334432 (ONSTRE01)	873	2432
				60432786 (ANTP2ES07)	2427	2652
				602020981	2328	2652
				6943283H1 (BSAIFR02)	1	718
35	71564320CE1	3017	2968-3017, 1-43, 1234-1322, 679-736	128732076 (BRAIN011)	2615	2983
				71579751V1	1033	1791
				71580777V1	566	1149
				6764743H1 (BSAIFR01)	1406	2001
				7981090H1 (BSAIFR01)	1	577
				1226833H1 (BSAIFR01)	2783	3017
				7167123H1 (BSAIFR01)	483	2003
				215711256 (BSAIFR02)	2388	2977
36	7475538CE1	2166	811-870, 303-1227	215711256 (BSAIFR02)	1583	2450
				6656591H1 (BSAIFR08)	613	1281
				70644867V1	515	1249
				70645804V1	1	595
				70645323V1	1	1277
				71564044V1	1277	1993
				71565564V1	1156	1924

WO 01/81555

PCT/US01/12992

Table 5

Polynucleotide Seq ID NO:	Incyte Project ID	Representative Library
19	259554CB1	SCOUT002
20	747553CB1	COLMOT01
21	310752CB1	TOYH00704
22	5544470CB1	SEMOOT01
23	747332CB1	SKTNG001
24	155145CB1	LOREN003
25	259315CB1	BRAD0101
26	4357117CB1	BRAD0102
27	531192CB1	SINT002
28	747456CB1	BRAD0103
29	7474502CB1	COLMOT01
30	7475509CB1	BRAD0102
31	7475491CB1	SCOUT001
32	3192119CB1	PROST0523
33	7474596CB1	BRAD0101
34	1534248CB1	BRAD0102
35	7158432CB1	BRAD0101
36	7475538CB1	BRAD0104

WO 01/81555

PCT/US01/12952

Table 6

Library	Vector	Library Description
BRATXT02	P1N1Y	Library was constructed using RNA isolated from treated SH-SY5Y cell line derived from bone marrow neuroblastoma tumor. Cells removed from a 4-year-old Caucasian female, who died from a cerebral aneurysm, were cultured and total RNA was extracted. The library was constructed using total RNA isolated from the brain of a 57-year-old Caucasian male, who died from a cerebrovascular accident.
BRADIP01	P1N1Y	This large size-fractionated library was constructed using RNA isolated from brain tissue removed from a Caucasian male fetus who was stillborn with a hypoplastic left heart at 23 weeks gestation.
BRANFEC01	P1N1Y	This normalized fetal brain tissue library was constructed from 400 thousand independent clones from a fetal brain tissue library. Starting RNA was made from brain tissue removed from a Caucasian male fetus who was stillborn with a hypoplastic left heart at 23 weeks gestation. The library was constructed in 2 rounds using conditions adapted from Soares et al. (1994) 91:9223 and Bonaldo et al., Genome Research (1996) 6:791, except that a significantly longer (43 hours/round) reannealing hybridization was used.
BRANING01	REPORT	Library was constructed and normalized from 4.88 million independent clones from a brain tissue library. RNA was made from brain tissue removed from a 26-year-old Caucasian male during cranioplasty and excision of a cerebral meningial lesion. Pathology for the associated tumor tissue indicated a grade oligodendroglioma in the right frontoparietal part of the brain.
BRATUT01	P1N1Y	Library was constructed using RNA isolated from brain tissue removed from the midline frontal lobe of a 51-year-old Caucasian female during excision of a cerebral meningial lesion. Pathology indicated subfrontal meningothelial meningioma with no atypia. One ethmoid and mucosal tissue sample indicated meningioma. Family history included cerebrovascular disease, senile dementia, hyperlipidemia, benign hypertension, atherosclerotic coronary artery disease, congestive heart failure, and breast cancer.
BRATUT02	P1N1Y	Library was constructed using RNA isolated from brain tumor tissue removed from a 51-year-old Caucasian female during excision of a cerebral meningial lesion. Pathology indicated a meningioma. Family history included senile dementia.

WO 01/81555

PCT/US01/12992

Table 6 (cont.)

Library	Vector	Library Description
BRADNR03	PINCY	This normalized brain tissue library was constructed from 6.7 million independent clones from the Huntington tissue library. Studies have made a link between the Huntington gene and the disease. The library was constructed from a 57-year-old Caucasian male who died from a cerebrovascular accident. Patient history included Huntington's disease and emphysema. The library was normalized in 2 rounds using conditions adapted from Soares et al., EMAS (1994) 91:9228 and Ronaldo et al., Genome Research (1996) 6:791, except that a significantly longer (48 -hours/round) reannealing hybridization was used. The library was linearized and recircularized to select for insert containing clones.
COLANGT01	PSFOR11	Library was constructed using RNA isolated from colon tissue removed from a 75-year-old male. The library was constructed using RNA isolated from a 75-year-old male.
LIVNMG03	PSFOR11	Library was constructed using RNA isolated from the left ventricle tissue of a 11-year-old male.
MCLEXT02	PINCY	Library was constructed using RNA isolated from treated umbilical cord blood dendritic cells removed from a male. The cells were treated with granulocyte/macrophage colony stimulating factor (GM-CSF), tumor necrosis factor alpha (TNF alpha), stem cell factor (SCF), fibroblast growth factor (FGF), and ionomycin. The GM-CSF was added at time 0 at 100 ng/ml, the TNF alpha was added at time 0 at 2.5 ng/ml, the SCF was added at time 0 at 25 ng/ml, and the FGF was added at time 0 at 10 ng/ml. The cells were incubated for five hours. Incubation time was 13 days.

WO 01/81555

PCT/US01/12992

Table 6 (cont.)

Library	Vector	Library Description
PROST023	pUNCY	This subtracted prostate tumor library was constructed using 10 million clones from a pooled prostate tumor library that was subjected to 2 rounds of serial dilution and PCR amplification. The subtracted prostate tumor library was constructed by pooling equal numbers of clones from 4 prostate tumor libraries using NMA (B1, 66 (C), and 68 (D) during prostatectomy with lymph node excision. Pathology indicated adenocarcinoma in all donors. History included elevated PSA, induration and tobacco abuse in donor A; elevated PSA, induration, prostate hyperplasia, renal failure, osteoarthritis, renal artery stenosis, benign prostatic hyperplasia, hyperlipidemia, tobacco/alcohol abuse and benign prostatic hyperplasia in donor B; elevated PSA, hypercholesterolemia, 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 Swaroop et al., <i>Cell</i> 13 (1992), 1127, and Gaudin et al., <i>Gene</i> 199 (1997), 173. Tissue removed from the base of the medulla of a 57-year-old Caucasian male who died from a cerebrovascular accident. Patient history included Huntington's disease and emphysema.
SCOND001	pUNCY	Library was constructed using RNA isolated from seminal vesicle tissue removed from a 58-year-old Caucasian male during radical prostatectomy. Pathology for the associated tumor tissue indicated adenocarcinoma (Gleason grade 3+2) of the prostate. Adenofibromatous hyperplasia was also present. The patient presented with elevated prostate specific antigen (PSA). Family history included prostate cancer.
SINTPE02	PCMA2.1	This 3' biased random primed library was constructed using RNA isolated from small intestine tissue removed from a Caucasian male fetus who died from Patou's syndrome (trisomy 13) at 20-weeks' gestation. Serology was negative.
SKIN00001	PCMA2.1	This 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 hyperplasia. Family history included benign hypertension.
TYM00004	pUNCY	Library was constructed using RNA isolated from activated T41 cells. These cells were obtained from a patient with rheumatoid arthritis. Cells were transferred COS cells, and then activated for six hours with anti-CD3 and anti-CD28 antibodies.

WO 01/81555

PCT/US01/12992

Table 7

Program	Description	Reference	Parameter Threshold
ABI EACTURA	A program that removes vector sequences and masks ambiguous bases in nucleic acid sequences.	Applied Biosystems, Foster City, CA.	
ABIPARACEL.HDF	A Fast Data Finder useful in comparing and annotating amino acid or nucleic acid sequences.	Applied Biosystems, Foster City, CA; Perceci, Inc., Pasadena, CA.	Mismatch <50%
ABI AutoAssembler	A program that assembles nucleic acid sequences.	Applied Biosystems, Foster City, CA.	
BLAST	A Basic Local Alignment Search Tool useful in sequence similarity search for amino acid and nucleic acid sequences. BLAST includes five functions: blastn, blastx, tblastn, and tblastx.	Altschul, S.F. et al. (1990) J. Mol. Biol. 215:403-410; Altschul, S.F. et al. (1997) Nucleic Acids Res. 25:3389-3402.	ESTs: Probability value= 1.0E-8 or less Full Length sequences: Probability value= 1.0E-10 or less
FASTA	A Pearson and Lipman algorithm that searches for similarity between a query sequence and a group of sequences of the same type. FASTA comprises at least five functions: fasta, tfasta, tftstdx, 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.	ESTs: fasta E value=1.0E-6 Assembled ESTs: fasta Identity= 95% or greater and Match length=200 bases or greater; fasta B value=1.0E-8 or less Full Length sequences: fasta score=100 or greater
BLUMPS	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.	Hunkeler, S. and J.G. Hunkeler (1991) Nucleic Acids Res. 19:6565-6572; Hunkeler, J.G. and S. Hunkeler (1996) Methods Enzymol. 266:88-105; and Atwood, T.K. et al. (1997) J. Chem. Inf. Comput. Sci. 37:417-424.	Probability value= 1.0E-3 or less
HMMER	An algorithm for searching a query sequence against a hidden Markov model (HMM)-based databases of protein family consensus sequences, such as PFAM.	Krogh, A. et al. (1994) J. Mol. Biol. 235:1301-1331; Sonnhammer, E.L.L. et al. (1998) Nucleic Acids Res. 26:370-372; Durbin, R. et al. (1998) Our World View, in a Nucleic Acid, Cambridge Univ. Press, pp. 1-550.	PFAM hits: Probability value= 1.0E-3 or less Signal peptide hits: Score=0 or greater

WO 01/81555

PCT/US01/12992

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.	Grisham, M. et al. (1988) CABIOS 4:61-66; Grisham, M. et al. (1989) Methods Enzymol. 183:146-159; Batroch, A. et al. (1997) Nucleic Acids Res. 25:217-221.	Normalized quality scores, GDC-specified "HIGH" value for that particular Prosite motif. Generally, score=1.4-2.1.
Phred	A base-calling algorithm that examines automated sequence traces with high sensitivity and probability.	Ewing, R. et al. (1998) Genome Res. 8:175-185; Ewing, B. and P. Green (1998) Genome Res. 8:186-194.	
Phrap	A Phile Revised Assembly Program including SWAT and CrossMatch, programs based on efficient implementation of the Smith-Waterman algorithm, useful in searching sequence homology and assembling DNA sequences.	Smith, T.F. and M.S. Waterman (1981) Adv. Appl. Math. 2:482-489; Smith, T.F. and M.S. Waterman (1981) J. Mol. Biol. 147:195-197; and Green, P., University of Washington, Seattle, WA.	Score= 120 or greater; Match length= 56 or greater
Consed	A graphical tool for viewing and editing Phrap assemblies.	Gordon, D. et al. (1998) Genome Res. 8:195-202.	
SPScan	A weight matrix analysis program that scans protein sequences for the presence of secretory signal peptides.	Nielson, H. et al. (1997) Protein Engineering 10:1-6; Clavotte, J.M. and S. Audie (1997) CABIOS 12:431-439.	Scores=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. 237:183-192; Persson, B. and P. Argos (1996) Protein Sci. 5:363-371.	
TMMMMR	A program that uses a hidden Markov model (HMM) to delineate transmembrane segments on protein sequences and determine orientation.	Soukhanov, 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.	Batroch, 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 01/81555

PCT/US01/12992

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:1-18,
 - b) a naturally occurring polypeptide comprising an amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:2-18,
 - c) a naturally occurring polypeptide comprising an amino acid sequence at least 98% identical to an amino acid sequence of SEQ ID NO:1,
 - 10 d) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-18, and
 - e) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-18.
- 15 2. An isolated polypeptide of claim 1 selected from the group consisting of SEQ ID NO:1-18.
3. An isolated polynucleotide encoding a polypeptide of claim 1.
- 20 4. An isolated polynucleotide encoding a polypeptide of claim 2.
5. An isolated polynucleotide of claim 4 selected from the group consisting of SEQ ID NO:19-36.
- 25 6. A recombinant polynucleotide comprising a promoter sequence operably linked to a polynucleotide of claim 3.
7. A cell transformed with a recombinant polynucleotide of claim 6.
- 30 8. A transgenic organism comprising a recombinant polynucleotide of claim 6.
9. A method for producing a polypeptide of claim 1, the method comprising:
 - a) culturing a cell under conditions suitable for expression of the polypeptide, wherein said cell is transformed with a recombinant polynucleotide, and said recombinant polynucleotide

WO 01/81555

PCT/US01/12992

comprises a promoter sequence operably linked to a polynucleotide encoding the polypeptide of claim 1, and

b) recovering the polypeptide so expressed.

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

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

- a) a polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:19-36,
 10 b) a naturally occurring polynucleotide comprising a polynucleotide sequence at least 90% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:20-36,
 c) a naturally occurring polynucleotide comprising a polynucleotide sequence at least 98% identical to the polynucleotide sequence of SEQ ID NO:19,
 d) a polynucleotide complementary to a polynucleotide of a),
 15 e) a polynucleotide complementary to a polynucleotide of b),
 f) a polynucleotide complementary to a polynucleotide of c), and
 g) an RNA equivalent of a)-f).

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

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

- a) hybridizing the sample with a probe comprising at least 20 contiguous nucleotides
 25 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.

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

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

WO 01/81555

PCT/US01/12992

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.

5

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

17. A composition of claim 16, wherein the polypeptide has an amino acid sequence selected from the group consisting of SEQ ID NO:1-18.

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.

15

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

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.

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.

30

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

WO 01/81555

PCT/US01/12992

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.

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

- a) combining the polypeptide of claim 1 with at least one test compound under suitable conditions, and
- b) detecting binding of the polypeptide of claim 1 to the test compound, thereby identifying a compound that specifically binds to the polypeptide of claim 1.

10

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

- a) combining the polypeptide of claim 1 with at least one test compound under conditions permissive for the activity of the polypeptide of claim 1,
- b) assessing the activity of the polypeptide of claim 1 in the presence of the test compound, and
- c) comparing the activity of the polypeptide of claim 1 in the presence of the test compound with the activity of the polypeptide of claim 1 in the absence of the test compound, wherein a change in the activity of the polypeptide of claim 1 in the presence of the test compound is indicative of a compound that modulates the activity of the polypeptide of claim 1.

20

27. A method for screening a compound for effectiveness in altering expression of a target polynucleotide, wherein said target polynucleotide comprises a sequence of claim 5, the method comprising:

- a) exposing a sample comprising the target polynucleotide to a compound, under conditions suitable for the expression of the target polynucleotide,
- b) detecting altered expression of the target polynucleotide, and
- c) comparing the expression of the target polynucleotide in the presence of varying amounts of the compound and in the absence of the compound.

30

28. A method for assessing toxicity of a test compound, said method comprising:

- a) treating a biological sample containing nucleic acids with the test compound;
- b) hybridizing the nucleic acids of the treated biological sample with a probe comprising at least 20 contiguous nucleotides of a polynucleotide of claim 11 under conditions whereby a specific hybridization complex is formed between said probe and a target polynucleotide in the biological

35

WO 01/81555

PCT/US01/12992

sample, said target polynucleotide comprising a polynucleotide sequence of a polynucleotide of claim 11 or fragment thereof;

- c) quantifying the amount of hybridization complex; and
- d) comparing the amount of hybridization complex in the treated biological sample with the amount of hybridization complex in an untreated biological sample, wherein a difference in the amount of hybridization complex in the treated biological sample is indicative of toxicity of the test compound.

29. A diagnostic test for a condition or disease associated with the expression of 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
- b) detecting the complex, wherein the presence of the complex correlates with the presence of the polypeptide in the biological sample.

30. The antibody of claim 10, wherein the antibody is:

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

31. A composition comprising an antibody of claim 10 and an acceptable excipient.

32. A method of diagnosing a condition or disease associated with the expression of PKIN in a subject, comprising administering to said subject an effective amount of the composition of claim 31.

33. A composition of claim 31, wherein the antibody is labeled.

34. A method of diagnosing a condition or disease associated with the expression of PKIN in a subject, comprising administering to said subject an effective amount of the composition of claim 33.

WO 01/81555

PCT/US01/12992

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: 1-18, or an immunogenic fragment thereof, under conditions to elicit an antibody response;
- b) isolating antibodies from said animal; and
- c) screening the isolated antibodies with the polypeptide, thereby identifying a polyclonal antibody which binds specifically to a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO: 1-18.

10

36. An antibody produced by a method of claim 35.

37. A composition comprising the antibody of claim 36 and a suitable carrier.

38. A method of making a monoclonal antibody with the specificity of the antibody of claim 10 comprising:

- a) immunizing an animal with a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO: 1-18, 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: 1-18.

25

39. A monoclonal antibody produced by a method of claim 38.

40. A composition comprising the antibody of claim 39 and a suitable carrier.

30

41. The antibody of claim 10, wherein the antibody is produced by screening a Fab expression library.

42. The antibody of claim 10, wherein the antibody is produced by screening a recombinant

WO 01/81555

PCT/US01/12992

immunoglobulin library.

43. A method for detecting a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-18 in a sample, comprising the steps of:

- 5 a) incubating the antibody of claim 10 with a sample under conditions to allow specific binding of the antibody and the polypeptide; and
- b) detecting specific binding, wherein specific binding indicates the presence of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-18 in the sample.

10

44. A method of purifying a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-18 from a sample, the method comprising:

- a) incubating the antibody of claim 10 with a sample under conditions to allow specific binding of the antibody and the polypeptide; and
- 15 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:1-18.

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

20

46. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:2.

47. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:3.

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

25

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

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

30

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.

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

WO 01/81555

PCT/US01/12992

54. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:10.
55. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:11.
- 5 56. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:12.
57. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:13.
- 10 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.
60. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:16.
- 15 61. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:17.
62. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:18.
- 20 63. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:19.
64. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:20.
- 25 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.
- 30 67. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:23.

WO 01/81555

PCT/US01/12992

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
5 NO:25.

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
10 NO:27.

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
15 NO:29.

74. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID
20 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
25 NO:32.

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
30 NO:34.

79. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID

WO 01/81555

PCT/US01/12992

NO:35.

80. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:36.

WO 01/81555

PCT/US01/12992

<110> INCYTE GENOMICS, INC.
 YUE, Henry
 GANDHI, Anshu R.
 TRIBOULEY, Catherine M.
 KEARNEY, Liam
 GRIFFIN, Jennifer A.
 NGUYEN, Daniel B.
 SANDMAN, Olga
 LU, Dzung Aina M.
 LAL, Preeti
 BURFORD, Neil
 KHAN, Farrah A.
 WALIA, Narinder K.
 YAO, Monique G.
 PATTERSON, Chandra
 BURKILL, John D.
 MARCUS, Gregory A.
 ZINGLER, Kurt A.
 RECFON, Shirley A.
 LU, Yan
 POLICKY, Jennifer L.
 THORNTON, Michael
 TANG, Y. Tom
 HAFALIA, April
 ELLIOTT, Vicki S.
 BAUGHN, Mariah R.
 WALSH, Roderick T.
 RAMKUMAR, Jayalaxmi
 BOROWSKY, Mark L.
 AU-YOUNG, Janice
 HILLMAN, Jennifer L.
 GURURAJAN, Rajagopal

<120> HUMAN KINASES

<130> PI-0076 PCT

<140> To Be Assigned

<141> Herewith

<150> 60/199,021; 60/200,226; 60/202,339; 60/203,505; 60/205,654;
 60/207,739; 60/208,795

<151> 2000-04-20; 2000-04-28; 2000-05-05; 2000-05-11; 2000-05-18;
 2000-05-26; 2000-06-01

<160> 36

<170> PERL Program

<210> 1

<211> 1210

<212> FRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 2890544CD1

<400> 1
 Met Ala Ser Gln Leu Gln Val Phe Ser Pro Pro Ser Val Ser Ser
 1 5 10 15
 Ser Ala Phe Cys Ser Ala Lys Lys Leu Lys Ile Glu Pro Ser Gly
 20 25 30

WO 01/81555

PCT/US01/12992

Trp Asp Val Ser Gly Gln Ser Ser Asn Asp Lys Tyr Tyr Thr His
 35 40 45
 Ser Lys Thr Leu Pro Ala Thr Gln Gly Gln Ala Asn Ser Ser His
 50 55 60
 Gln Val Ala Asn Phe Asn Ile Pro Ala Tyr Asp Gln Gly Leu Leu
 65 70 75
 Leu Pro Ala Pro Ala Val Glu His Ile Val Val Thr Ala Ala Asp
 80 85 90
 Ser Ser Gly Ser Ala Ala Thr Ser Thr Phe Gln Ser Ser Gln Thr
 95 100 105
 Leu Thr His Arg Ser Asn Val Ser Leu Leu Glu Pro Tyr Gln Lys
 110 115 120
 Cys Gly Leu Lys Arg Lys Ser Glu Glu Val Asp Ser Asn Gly Ser
 125 130 135
 Val Gln Ile Ile Glu Glu His Pro Pro Leu Met Leu Gln Asn Arg
 140 145 150
 Thr Val Val Gly Ala Ala Ala Thr Thr Thr Thr Val Thr Thr Lys
 155 160 165
 Ser Ser Ser Ser Ser Gly Glu Gly Asp Tyr Gln Leu Val Gln His
 170 175 180
 Glu Ile Leu Cys Ser Met Thr Asn Ser Tyr Glu Val Leu Glu Phe
 185 190 195
 Leu Gly Arg Gly Thr Phe Gly Gln Val Ala Lys Cys Trp Lys Arg
 200 205 210
 Ser Thr Lys Glu Ile Val Ala Ile Lys Ile Leu Lys Asn His Pro
 215 220 225
 Ser Tyr Ala Arg Gln Gly Gln Ile Glu Val Ser Ile Leu Ser Arg
 230 235 240
 Leu Ser Ser Glu Asn Ala Asp Glu Tyr Asn Leu Val Arg Ser Tyr
 245 250 255
 Glu Cys Phe Gln His Lys Asn His Thr Cys Leu Val Phe Glu Met
 260 265 270
 Leu Glu Gln Asn Leu Tyr Asp Phe Leu Lys Gln Asn Lys Phe Ser
 275 280 285
 Pro Leu Pro Leu Lys Tyr Ile Arg Pro Ile Leu Gln Gln Val Ala
 290 295 300
 Thr Ala Leu Met Lys Leu Lys Ser Leu Gly Leu Ile His Ala Asp
 305 310 315
 Leu Lys Pro Glu Asn Ile Met Leu Val Asp Pro Val Arg Gln Pro
 320 325 330
 Tyr Arg Val Lys Val Ile Asp Phe Gly Ser Ala Ser His Val Ser
 335 340 345
 Lys Ala Val Cys Ser Thr Tyr Leu Gln Ser Arg Tyr Tyr Arg Ala
 350 355 360
 Pro Glu Ile Ile Leu Gly Leu Pro Phe Cys Glu Ala Ile Asp Met
 365 370 375
 Trp Ser Leu Gly Cys Val Ile Ala Glu Leu Phe Leu Gly Trp Pro
 380 385 390
 Leu Tyr Pro Gly Ala Ser Glu Tyr Asp Gln Ile Arg Tyr Ile Ser
 395 400 405
 Gln Thr Gln Gly Leu Pro Ala Glu Tyr Leu Leu Ser Ala Gly Thr
 410 415 420
 Lys Thr Thr Arg Phe Phe Asn Arg Asp Pro Asn Leu Gly Tyr Pro
 425 430 435
 Leu Trp Arg Leu Lys Thr Pro Glu Glu His Glu Leu Glu Thr Gly
 440 445 450
 Ile Lys Ser Lys Glu Ala Arg Lys Tyr Ile Phe Asn Cys Leu Asp
 455 460 465
 Asp Met Ala Gln Val Asn Met Ser Thr Asp Leu Glu Gly Thr Asp
 470 475 480
 Met Leu Ala Glu Lys Ala Asp Arg Arg Glu Tyr Ile Asp Leu Leu
 485 490 495
 Lys Lys Met Leu Thr Ile Asp Ala Asp Lys Arg Ile Thr Pro Leu

WO 01/81555

PCT/US01/12992

500	505	510
Lys Thr Leu Asn His Gln Phe Val Thr	Met Thr His Leu Leu Asp	
515	520	525
Phe Pro His Ser Asn His Val Lys Ser	Cys Phe Gln Asn Met Glu	
530	535	540
Ile Cys Lys Arg Arg Val His Met Tyr	Asp Thr Val Ser Gln Ile	
545	550	555
Lys Ser Pro Phe Thr Thr His Val Ala	Pro Asn Thr Ser Thr Asn	
560	565	570
Leu Thr Met Ser Phe Ser Asn Gln Leu	Asn Thr Val His Asn Gln	
575	580	585
Ala Ser Val Leu Ala Ser Ser Ser Thr	Ala Ala Ala Thr Leu	
590	595	600
Ser Leu Ala Asn Ser Asp Val Ser Leu	Leu Asn Tyr Gln Ser Ala	
605	610	615
Leu Tyr Pro Ser Ser Ala Ala Pro Val	Pro Gly Val Ala Gln Gln	
620	625	630
Gly Val Ser Leu Gln Pro Gly Thr Thr	Gln Ile Cys Thr Gln Thr	
635	640	645
Asp Pro Phe Gln Gln Thr Phe Ile Val	Cys Pro Pro Ala Phe Gln	
650	655	660
Thr Gly Leu Gln Ala Thr Thr Lys His	Ser Gly Phe Pro Val Arg	
665	670	675
Met Asp Asn Ala Val Pro Ile Val Pro	Gln Ala Pro Ala Ala Gln	
680	685	690
Pro Leu Gln Ile Gln Ser Gly Val Leu	Thr Gln Gly Ser Cys Thr	
695	700	705
Pro Leu Met Val Ala Thr Leu His Pro	Gln Val Ala Thr Ile Thr	
710	715	720
Pro Gln Tyr Ala Val Pro Phe Thr Leu	Ser Cys Ala Ala Gly Arg	
725	730	735
Pro Ala Leu Val Glu Gln Thr Ala Ala	Val Leu Gln Ala Trp Pro	
740	745	750
Gly Gly Thr Gln Gln Ile Leu Leu Pro	Ser Thr Trp Gln Gln Leu	
755	760	765
Pro Gly Val Ala Leu His Asn Ser Val	Gln Pro Thr Ala Met Ile	
770	775	780
Pro Glu Ala Met Gly Ser Gly Gln Gln	Leu Ala Asp Trp Arg Asn	
785	790	795
Ala His Ser His Gly Asn Gln Tyr Ser	Thr Ile Met Gln Gln Pro	
800	805	810
Ser Leu Leu Thr Asn His Val Thr Leu	Ala Thr Ala Gln Pro Leu	
815	820	825
Asn Val Gly Val Ala His Val Val Arg	Gln Gln Gln Ser Ser Ser	
830	835	840
Leu Pro Ser Lys Lys Asn Lys Gln Ser	Ala Pro Val Ser Ser Lys	
845	850	855
Ser Ser Leu Asp Val Leu Pro Ser Gln	Val Tyr Ser Leu Val Gly	
860	865	870
Ser Ser Pro Leu Arg Thr Thr Ser Ser	Tyr Asn Ser Leu Val Pro	
875	880	885
Val Gln Asp Gln His Gln Pro Ile Ile	Ile Ile Ile Pro Asp Thr Pro Ser	
890	895	900
Pro Pro Val Ser Val Ile Thr Ile Arg	Ser Asp Thr Asp Glu Glu	
905	910	915
Glu Asp Asn Lys Tyr Lys Pro Ser Ser	Ser Gly Leu Lys Pro Arg	
920	925	930
Ser Asn Val Ile Ser Tyr Val Thr Val	Asn Asp Ser Pro Asp Ser	
935	940	945
Asp Ser Ser Leu Ser Ser Pro Tyr Ser	Thr Asp Thr Leu Ser Ala	
950	955	960
Leu Arg Gly Asn Ser Gly Ser Val Leu	Glu Gly Pro Gly Arg Val	
965	970	975

WO 01/81555

PCT/US01/12992

Val Ala Asp Gly Thr Gly Thr Arg Thr Ile Ile Val Pro Pro Leu
 980 985 990
 Lys Thr Gln Leu Gly Asp Cys Thr Val Ala Thr Gln Ala Ser Gly
 995 1000 1005
 Leu Leu Ser Asn Lys Thr Lys Pro Val Ala Ser Val Ser Gly Gln
 1010 1015 1020
 Ser Ser Gly Cys Cys Ile Thr Pro Thr Gly Tyr Arg Ala Gln Arg
 1025 1030 1035
 Gly Gly Thr Ser Ala Ala Gln Pro Leu Asn Leu Ser Gln Asn Gln
 1040 1045 1050
 Gln Ser Ser Ala Ala Pro Thr Ser Gln Glu Arg Ser Ser Asn Pro
 1055 1060 1065
 Ala Pro Arg Arg Gln Gln Ala Phe Val Ala Pro Leu Ser Gln Ala
 1070 1075 1080
 Pro Tyr Thr Phe Gln His Gly Ser Pro Leu His Ser Thr Gly His
 1085 1090 1095
 Pro His Leu Ala Pro Ala Pro Ala His Leu Pro Ser Gln Ala His
 1100 1105 1110
 Leu Tyr Thr Tyr Ala Ala Pro Thr Ser Ala Ala Ala Leu Gly Ser
 1115 1120 1125
 Thr Ser Ser Ile Ala His Leu Phe Ser Pro Gln Gly Ser Ser Arg
 1130 1135 1140
 His Ala Ala Ala Tyr Thr Thr His Pro Ser Thr Leu Val His Gln
 1145 1150 1155
 Val Pro Val Ser Val Gly Pro Ser Ser Leu Thr Ser Ala Ser Val
 1160 1165 1170
 Ala Pro Ala Gln Tyr Gln His Gln Phe Ala Thr Gln Ser Tyr Ile
 1175 1180 1185
 Gly Ser Ser Arg Gly Ser Thr Ile Tyr Thr Gly Tyr Pro Leu Ser
 1190 1195 1200
 Pro Thr Lys Ile Ser Gln Tyr Ser Tyr Leu
 1205 1210

<210> 2

<211> 357

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 7472693CD1

<400> 2

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

WO 01/81555

PCT/US01/12992

```

Leu Tyr Tyr Ser Gln Asp Glu Glu Ser Lys Ile Met Ile Ser Asp
155 160 165
Phe Gly Leu Ser Lys Met Glu Gly Lys Gly Asp Val Met Ser Thr
170 175 180
Ala Cys Gly Thr Pro Gly Tyr Val Ala Pro Glu Val Leu Ala Gln
185 190 195
Lys Pro Tyr Ser Lys Ala Val Asp Cys Trp Ser Ile Gly Val Ile
200 205 210
Ala Tyr Ile Leu Leu Cys Gly Tyr Pro Pro Phe Tyr Asp Glu Asn
215 220 225
Asp Ser Lys Leu Phe Glu Gln Ile Leu Lys Ala Glu Tyr Glu Phe
230 235 240
Asp Ser Pro Tyr Trp Asp Asp Ile Ser Asp Ser Ala Lys Asp Phe
245 250 255
Ile Arg Asn Leu Met Glu Lys Asp Pro Asn Lys Arg Tyr Thr Cys
260 265 270
Glu Gln Ala Ala Arg His Pro Trp Ile Ala Gly Asp Thr Ala Leu
275 280 285
Asn Lys Asn Ile His Glu Ser Val Ser Ala Gln Ile Arg Lys Asn
290 295 300
Phe Ala Lys Ser Lys Trp Arg Gln Ala Phe Asn Ala Thr Ala Val
305 310 315
Val Arg His Met Arg Lys Leu His Leu Gly Ser Ser Leu Asp Ser
320 325 330
Ser Asn Ala Ser Val Ser Ser Ser Leu Ser Leu Ala Ser Gln Lys
335 340 345
Asp Cys Ala Tyr Val Ala Lys Pro Glu Ser Leu Ser
350 355

```

<210> 3

<211> 419

<212> FRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 3107952CD1

<400> 3

```

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

```

WO 01/81555

PCT/US01/12992

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

<210> 4
 <211> 624
 <212> PRT
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 5544420CD1

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

WO 01/81555

PCT/US01/12992

Arg Ser Gln Pro Pro Tyr Thr Asp Tyr Val Ser Thr Arg Trp Tyr
 155 160 165
 Arg Ala Pro Glu Val Leu Leu Arg Ser Ser Val Tyr Ser Ser Pro
 170 175 180
 Ile Asp Val Trp Ala Val Gly Ser Ile Met Ala Glu Leu Tyr Met
 185 190 195
 Leu Arg Pro Leu Phe Pro Gly Thr Ser Glu Val Asp Glu Ile Phe
 200 205 210
 Lys Ile Cys Gln Val Leu Gly Thr Pro Lys Lys Ser Asp Trp Pro
 215 220 225
 Glu Gly Tyr Gln Leu Ala Ser Ser Met Asn Phe Arg Phe Pro Gln
 230 235 240
 Cys Val Pro Ile Asn Leu Lys Thr Leu Ile Pro Asn Ala Ser Asn
 245 250 255
 Glu Ala Ile Gln Leu Met Thr Glu Met Leu Asn Trp Asp Pro Lys
 260 265 270
 Lys Arg Pro Thr Ala Ser Gln Ala Leu Lys His Pro Tyr Phe Gln
 275 280 285
 Val Gly Gln Val Leu Gly Pro Ser Ser Asn His Leu Glu Ser Lys
 290 295 300
 Gln Ser Leu Asn Lys Gln Leu Gln Pro Leu Glu Ser Lys Pro Ser
 305 310 315
 Leu Val Glu Val Glu Pro Lys Pro Leu Pro Asp Ile Ile Asp Gln
 320 325 330
 Val Val Gly Gln Pro Gln Pro Lys Thr Ser Gln Gln Pro Leu Gln
 335 340 345
 Pro Ile Gln Pro Pro Gln Asn Leu Ser Val Gln Gln Pro Pro Lys
 350 355 360
 Gln Gln Ser Gln Glu Lys Pro Pro Gln Thr Leu Phe Pro Ser Ile
 365 370 375
 Val Lys Asn Met Pro Thr Lys Pro Asn Gly Thr Leu Ser His Lys
 380 385 390
 Ser Gly Arg Arg Arg Trp Gly Gln Thr Ile Phe Lys Ser Gly Asp
 395 400 405
 Ser Trp Glu Glu Leu Glu Asp Tyr Asp Phe Gly Ala Ser His Ser
 410 415 420
 Lys Lys Pro Ser Met Gly Val Phe Lys Glu Lys Arg Lys Lys Asp
 425 430 435
 Ser Pro Phe Arg Leu Pro Glu Pro Val Pro Ser Gly Ser Asn His
 440 445 450
 Ser Thr Gly Glu Asn Lys Ser Leu Pro Ala Val Thr Ser Leu Lys
 455 460 465
 Ser Asp Ser Glu Leu Ser Thr Ala Pro Thr Ser Lys Gln Tyr Tyr
 470 475 480
 Leu Lys Gln Ser Arg Tyr Leu Pro Gly Val Asn Pro Lys Lys Val
 485 490 495
 Ser Leu Ile Ala Ser Gly Lys Glu Ile Asn Pro His Thr Trp Ser
 500 505 510
 Asn Gln Leu Phe Pro Lys Ser Leu Gly Pro Val Gly Ala Glu Leu
 515 520 525
 Ala Phe Lys Arg Ser Asn Ala Glu Glu Lys Leu Gly Ser Tyr Ala
 530 535 540
 Thr Tyr Asn Gln Ser Gly Tyr Ile Pro Ser Phe Leu Lys Lys Glu
 545 550 555
 Val Gln Ser Ala Gly Gln Arg Ile His Leu Ala Pro Leu Asn Ala
 560 565 570
 Thr Ala Ser Glu Tyr Thr Trp Asn Thr Lys Thr Gly Arg Gly Gln
 575 580 585
 Phe Ser Gly Arg Thr Tyr Asn Pro Thr Ala Lys Asn Leu Asn Ile
 590 595 600
 Val Asn Arg Ala Gln Pro Ile Pro Ser Val His Gly Arg Thr Asp
 605 610 615
 Trp Val Ala Lys Tyr Gly Gly His Arg

WO 01/81555

PCT/US01/12992

620

<210> 5
 <211> 878
 <212> PRT
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 7472832CD1

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

WO 01/81555

PCT/US01/12992

Thr Thr Arg Lys	380	Ser Ser Thr Thr Leu Arg	385	Glu Gly Trp Val	390
His Tyr Ser Asn Lys	395	Asp Thr Leu Arg Lys	400	Arg His Tyr Trp Arg	405
Leu Asp Cys Lys	410	Cys Ile Thr Leu Phe	415	Gln Asn Asn Thr Thr	420
Arg Tyr Tyr Lys	425	Glu Ile Pro Leu Ser	430	Glu Ile Leu Thr Val	435
Ser Ala Gln Asn Phe	440	Ser Leu Val Pro	445	Gly Thr Asn Pro His	450
Cys Phe Glu Ile	455	Val Thr Ala Asn Ala	460	Thr Tyr Phe Val Gly	465
Met Pro Gly Gly	470	Thr Pro Gly Gly Pro	475	Ser Gly Gln Gly Ala	480
Ala Ala Arg Gly	485	Trp Glu Thr Ala Ile	490	Arg Gln Ala Leu Met	495
Val Ile Leu Gln	500	Asp Ala Pro Ser Ala	505	Pro Gly His Ala Pro	510
Arg Gln Ala Ser	515	Leu Ser Ile Ser Val	520	Ser Asn Ser Gln Ile	525
Glu Asn Val Asp	530	Ile Ala Thr Val Tyr	535	Gln Ile Phe Pro Asp	540
Val Leu Gly Ser	545	Gly Gln Phe Gly Val	550	Val Tyr Gly Gly Lys	555
Arg Lys Thr Gly	560	Arg Asp Val Ala Val	565	Lys Val Ile Asp Lys	570
Arg Phe Pro Thr	575	Lys Gln Glu Ser Gln	580	Leu Arg Asn Glu Val	585
Ile Leu Gln Ser	590	Leu Arg His Pro Gly	595	Ile Val Asn Leu Glu	600
Met Phe Glu Thr	605	Pro Glu Lys Val Phe	610	Val Val Met Glu Lys	615
His Gly Asp Met	620	Leu Glu Met Ile Leu	625	Ser Ser Glu Lys Gly	630
Leu Pro Glu Arg	635	Leu Thr Lys Phe Leu	640	Ile Thr Gln Ile Leu	645
Ala Leu Arg His	650	Leu His Phe Lys Asn	655	Ile Val His Cys Asp	660
Lys Pro Glu Asn	665	Val Leu Leu Ala Ser	670	Ala Asp Pro Phe Pro	675
Val Lys Leu Cys	680	Asp Phe Gly Phe Ala	685	Arg Ile Ile Gly Glu	690
Ser Phe Arg Arg	695	Ser Val Val Gly Thr	700	Pro Ala Tyr Leu Ala	705
Glu Val Leu Leu	710	Asn Gln Gly Tyr Asn	715	Ser Leu Asp Met Trp	720
Ser Val Gly Val	725	Ile Met Tyr Val Ser	730	Leu Ser Gly Thr Phe	735
Phe Asn Glu Asp	740	Glu Asp Ile Asn Asp	745	Gln Ile Gln Asn Ala	750
Phe Met Tyr Pro	755	Ala Ser Pro Trp Ser	760	His Ile Ser Ala Gly	765
Ile Asp Leu Ile	770	Asn Asn Leu Leu Gln	775	Val Lys Met Arg Lys	780
Tyr Ser Val Asp	785	Lys Ser Leu Ser His	790	Pro Trp Leu Gln Glu	795
Gln Thr Trp Leu	800	Asp Leu Arg Glu Leu	805	Gly Lys Met Gly	810
Arg Tyr Ile Thr	815	His Glu Ser Asp Asp	820	Ala Arg Trp Glu Gln	825
Ala Ala Glu His	830	Pro Leu Pro Gly Ser	835	Gly Leu Pro Thr Asp	840
	845		850		855

WO 01/81555

PCT/US01/12992

Asp Leu Gly Gly Ala Cys Pro Pro Gln Asp His Asp Met Gln Gly
 860 865 870
 Leu Ala Glu Arg Ile Ser Val Leu
 875

<210> 6
 <211> 443
 <212> PRT
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 1551456CD1

<400> 6
 Met Ser Lys Leu Arg Met Lys Arg Arg Ala Ser Asp Arg Gly Ala
 1 5 10 15
 Gly Glu Thr Ser Ala Arg Ala Lys Ala Leu Gly Ser Gly Ile Ser
 20 25 30
 Gly Asn Asn Ala Lys Arg Ala Gly Pro Phe Ile Leu Gly Pro Arg
 35 40 45
 Leu Gly Asn Ser Pro Val Pro Ser Ile Val Gln Cys Leu Ala Arg
 50 55 60
 Lys Asp Gly Thr Asp Asp Phe Tyr Gln Leu Lys Ile Leu Thr Leu
 65 70 75
 Glu Glu Arg Gly Asp Gln Gly Ile Glu Ser Gln Glu Glu Arg Gln
 80 85 90
 Gly Lys Met Leu Leu His Thr Glu Tyr Ser Leu Leu Ser Leu Leu
 95 100 105
 His Thr Gln Asp Gly Val Val His His His Gly Leu Phe Gln Asp
 110 115 120
 Arg Thr Cys Glu Ile Val Glu Asp Thr Glu Ser Ser Arg Met Val
 125 130 135
 Lys Lys Met Lys Lys Arg Ile Cys Leu Val Leu Asp Cys Leu Cys
 140 145 150
 Ala His Asp Phe Ser Asp Lys Thr Ala Asp Leu Ile Asn Leu Gln
 155 160 165
 His Tyr Val Ile Lys Glu Lys Arg Leu Ser Glu Arg Glu Thr Val
 170 175 180
 Val Ile Phe Tyr Asp Val Val Arg Val Val Glu Ala Leu His Gln
 185 190 195
 Lys Asn Ile Val His Arg Asp Leu Lys Leu Gly Asn Met Val Leu
 200 205 210
 Asn Lys Arg Thr His Arg Ile Thr Ile Thr Asn Phe Cys Leu Gly
 215 220 225
 Lys His Leu Val Ser Glu Gly Asp Leu Leu Lys Asp Gln Arg Gly
 230 235 240
 Ser Pro Ala Tyr Ile Ser Pro Asp Val Leu Ser Gly Arg Pro Tyr
 245 250 255
 Arg Gly Lys Pro Ser Asp Met Trp Ala Leu Gly Val Val Leu Phe
 260 265 270
 Thr Met Leu Tyr Gly Gln Phe Pro Phe Tyr Asp Ser Ile Pro Gln
 275 280 285
 Glu Leu Phe Arg Lys Ile Lys Ala Ala Glu Tyr Thr Ile Pro Glu
 290 295 300
 Asp Gly Arg Val Ser Glu Asn Thr Val Cys Leu Ile Arg Lys Leu
 305 310 315
 Leu Val Leu Asp Pro Gln Gln Arg Leu Ala Ala Ala Asp Val Leu
 320 325 330
 Glu Ala Leu Ser Ala Ile Ile Ala Ser Trp Gln Ser Leu Ser Ser
 335 340 345
 Leu Ser Gly Pro Leu Gln Val Val Pro Asp Ile Asp Asp Gln Met
 350 355 360

WO 01/81555

PCT/US01/12992

```

Ser Asn Ala Asp Ser Ser Gln Glu Ala Lys Val Thr Glu Glu Cys
365 370 375
Ser Gln Tyr Glu Phe Glu Asn Tyr Met Arg Gln Gln Leu Leu Leu
380 385 390
Ala Glu Glu Lys Ser Ser Ile His Asp Ala Arg Ser Trp Val Pro
395 400 405
Lys Arg Gln Phe Gly Ser Ala Pro Pro Val Arg Arg Leu Gly His
410 415 420
Asp Ala Gln Pro Met Thr Ser Leu Asp Thr Ala Ile Leu Ala Gln
425 430 435
Arg Tyr Leu Arg Lys
440

```

```

<210> 7
<211> 923
<212> FRT
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Incyte ID No: 2589355CD1

```

```

<400> 7
Met Ala Arg Gly Thr Cys Ser Ala Gly Arg Ser Gly Trp Gly Ser
1 5 10 15
Thr Thr Ser Arg Ala Arg Trp Ala Ser Gly Asn Phe Ala Val Val
20 25 30
Lys Leu Gly Arg His Arg Ile Thr Lys Thr Glu Val Ala Ile Lys
35 40 45
Ile Ile Asp Lys Ser Gln Pro Trp Met His Val Asn Leu Glu Lys
50 55 60
Ile Tyr Arg Glu Val Gln Ile Met Lys Met Leu Asp His Pro His
65 70 75
Ile Ile Lys Leu Tyr Gln Val Met Glu Thr Lys Ser Met Leu Tyr
80 85 90
Leu Val Thr Glu Tyr Ala Lys Asn Gly Glu Ile Phe Asp Tyr Leu
95 100 105
Ala Asn His Gly Arg Leu Asn Glu Ser Glu Ala Arg Arg Lys Phe
110 115 120
Trp Gln Ile Leu Ser Ala Val Asp Tyr Cys His Gly Arg Lys Ile
125 130 135
Val His Arg Asp Leu Lys Ala Glu Asn Leu Leu Asp Asn Asn
140 145 150
Met Asn Ile Lys Ile Ala Asp Phe Gly Phe Gly Asn Phe Phe Lys
155 160 165
Ser Gly Glu Leu Leu Ala Thr Trp Cys Gly Ser Pro Pro Tyr Ala
170 175 180
Ala Pro Glu Val Phe Glu Gly Gln Gln Tyr Glu Gly Pro Gln Leu
185 190 195
Asp Ile Trp Ser Met Gly Val Val Leu Tyr Val Leu Val Cys Gly
200 205 210
Ala Leu Pro Phe Asp Gly Pro Thr Leu Pro Ile Leu Arg Gln Arg
215 220 225
Val Leu Glu Gly Arg Phe Arg Ile Pro Tyr Phe Met Ser Glu Asp
230 235 240
Cys Glu His Leu Ile Arg Arg Met Leu Val Leu Asp Pro Ser Lys
245 250 255
Arg Leu Thr Ile Ala Gln Ile Lys Glu His Lys Trp Met Leu Ile
260 265 270
Glu Val Pro Val Gln Arg Pro Val Leu Tyr Pro Gln Glu Gln Glu
275 280 285
Asn Glu Pro Ser Ile Gly Glu Phe Asn Glu Gln Val Leu Arg Leu
290 295 300

```

WO 01/81555

PCT/US01/12992

Met His Ser Leu Gly Ile Asp Gln Gln Lys Thr Ile Glu Ser Leu
 305 310 315
 Gln Asn Lys Ser Tyr Asn His Phe Ala Ala Ile Tyr Phe Leu Leu
 320 325 330
 Val Glu Arg Leu Lys Ser His Arg Ser Ser Phe Pro Val Glu Gln
 335 340 345
 Arg Leu Asp Gly Arg Gln Arg Arg Pro Ser Thr Ile Ala Glu Gln
 350 355 360
 Thr Val Ala Lys Ala Gln Thr Val Gly Leu Pro Val Thr Met His
 365 370 375
 Ser Pro Asn Met Arg Leu Leu Arg Ser Ala Leu Leu Pro Gln Ala
 380 385 390
 Ser Asn Val Glu Ala Phe Ser Phe Pro Ala Ser Gly Cys Gln Ala
 395 400 405
 Glu Ala Ala Phe Met Glu Glu Glu Cys Val Asp Thr Pro Lys Val
 410 415 420
 Asn Gly Cys Leu Leu Asp Pro Val Pro Pro Val Leu Val Arg Lys
 425 430 435
 Gly Cys Gln Ser Leu Pro Ser Asn Met Met Glu Thr Ser Ile Asp
 440 445 450
 Glu Gly Leu Glu Thr Glu Gly Glu Ala Glu Glu Asp Pro Ala His
 455 460 465
 Ala Phe Glu Ala Phe Gln Ser Thr Arg Ser Gly Gln Arg Arg His
 470 475 480
 Thr Leu Ser Glu Val Thr Asn Gln Leu Val Val Met Pro Gly Ala
 485 490 495
 Gly Lys Ile Phe Ser Met Asn Asp Ser Pro Ser Leu Asp Ser Val
 500 505 510
 Asp Ser Glu Tyr Asp Met Gly Ser Val Gln Arg Asp Leu Asn Phe
 515 520 525
 Leu Glu Asp Asn Pro Ser Leu Lys Asp Ile Met Leu Ala Asn Gln
 530 535 540
 Pro Ser Pro Arg Met Thr Ser Pro Phe Ile Ser Leu Arg Pro Thr
 545 550 555
 Asn Pro Ala Met Gln Ala Leu Ser Ser Gln Lys Arg Glu Val His
 560 565 570
 Asn Arg Ser Pro Val Ser Phe Arg Glu Gly Arg Arg Ala Ser Asp
 575 580 585
 Thr Ser Leu Thr Gln Gly Ile Val Ala Phe Arg Gln His Leu Gln
 590 595 600
 Asn Leu Ala Arg Thr Lys Gly Ile Leu Glu Leu Asn Lys Val Gln
 605 610 615
 Leu Leu Tyr Glu Gln Ile Gly Pro Glu Ala Asp Pro Asn Leu Ala
 620 625 630
 Pro Ala Ala Pro Gln Leu Gln Asp Leu Ala Ser Ser Cys Pro Gln
 635 640 645
 Glu Glu Val Ser Gln Gln Gln Glu Ser Val Ser Thr Leu Pro Ala
 650 655 660
 Ser Val His Pro Gln Leu Ser Pro Arg Gln Ser Leu Glu Thr Gln
 665 670 675
 Tyr Leu Gln His Arg Leu Gln Lys Pro Ser Leu Leu Ser Lys Ala
 680 685 690
 Gln Asn Thr Cys Gln Leu Tyr Cys Lys Glu Pro Pro Arg Ser Leu
 695 700 705
 Glu Gln Gln Leu Gln Glu His Arg Leu Gln Gln Lys Arg Leu Phe
 710 715 720
 Leu Gln Lys Gln Ser Gln Leu Gln Ala Tyr Phe Asn Gln Met Gln
 725 730 735
 Ile Ala Glu Ser Ser Tyr Pro Gln Pro Ser Gln Gln Leu Pro Leu
 740 745 750
 Pro Arg Gln Glu Thr Pro Pro Pro Ser Gln Gln Ala Pro Pro Phe
 755 760 765
 Ser Leu Thr Gln Pro Leu Ser Pro Val Leu Glu Pro Ser Ser Glu

WO 01/81555

PCT/US01/12992

```

770      775      780
Gln Met Gln Tyr Ser Pro Phe Leu Ser Gln Tyr Gln Glu Met Gln
785      790      795
Leu Gln Pro Leu Pro Ser Thr Ser Gly Pro Arg Ala Ala Pro Pro
800      805      810
Leu Pro Thr Gln Leu Gln Gln Gln Gln Pro Pro Pro Pro Pro
815      820      825
Pro Pro Pro Pro Arg Gln Pro Gly Ala Ala Pro Ala Pro Leu Gln
830      835      840
Phe Ser Tyr Gln Thr Cys Glu Leu Pro Ser Ala Ala Ser Pro Ala
845      850      855
Pro Asp Tyr Pro Thr Pro Cys Gln Tyr Pro Val Asp Gly Ala Gln
860      865      870
Gln Ser Asp Leu Thr Gly Pro Asp Cys Pro Arg Ser Pro Gly Leu
875      880      885
Gln Glu Ala Pro Ser Ser Tyr Asp Pro Leu Ala Leu Ser Glu Leu
890      895      900
Pro Gly Leu Phe Asp Cys Glu Met Leu Asp Ala Val Asp Pro Gln
905      910      915
His Asn Gly Tyr Val Leu Val Asn
920

```

<210> 3

<211> 442

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 4357117CD1

<400> 3

```

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

```

WO 01/81555

PCT/US01/12992

```

230      235      240
Ser Gly Val Gly Val Asn Glu Leu Arg Arg Gln Leu Ile Glu Phe
245      250      255
Asn Pro Ser His Phe Gln Ser Ala Val Pro His Thr Thr Arg Thr
260      265      270
Lys Lys Ser Tyr Glu Met Asn Gly Arg Glu Tyr His Tyr Val Ser
275      280      285
Lys Glu Thr Phe Glu Asn Leu Ile Tyr Ser His Arg Met Leu Glu
290      295      300
Tyr Gly Glu Tyr Lys Gly His Leu Tyr Gly Thr Ser Val Asp Ala
305      310      315
Val Gln Thr Val Leu Val Glu Gly Lys Ile Cys Val Met Asp Leu
320      325      330
Glu Pro Gln Asp Ile Gln Gly Val Arg Thr His Glu Leu Lys Pro
335      340      345
Tyr Val Ile Phe Ile Lys Pro Ser Asn Met Arg Cys Met Lys Gln
350      355      360
Ser Arg Lys Asn Ala Lys Val Ile Thr Asp Tyr Tyr Val Asp Met
365      370      375
Lys Phe Lys Asp Glu Asp Leu Gln Glu Met Glu Asn Leu Ala Glu
380      385      390
Arg Met Glu Thr Gln Phe Gly Gln Phe Phe Asp His Val Ile Val
395      400      405
Asn Asp Ser Leu His Asp Ala Cys Ala Gln Leu Leu Ser Ala Ile
410      415      420
Gln Lys Ala Gln Glu Glu Pro Gln Trp Val Pro Ala Thr Trp Ile
425      430      435
Ser Ser Asp Thr Glu Ser Gln
440

```

<210> 9

<211> 1046

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 5511992CD1

<400> 9

```

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

```

WO 01/81555

PCT/US01/12992

	170		175		180
Glu	Ala	Lys	Leu	Phe	Ala
Leu	Arg	Gly	Val	Cys	Leu
Glu	Phe	Ala	Arg	Gly	Gly
Arg	Ile	Pro	Pro	Asp	Ile
Arg	Gly	Met	Asn	Tyr	Leu
His	Arg	Asp	Leu	Lys	Ser
Glu	Asn	Gly	Asp	Leu	Ser
Gly	Leu	Ala	Arg	Glu	Trp
Gly	Thr	Tyr	Ala	Trp	Met
Phe	Ser	Lys	Gly	Ser	Asp
Glu	Leu	Leu	Thr	Gly	Glu
Ala	Val	Ala	Tyr	Gly	Val
Pro	Ser	Thr	Cys	Pro	Glu
Trp	Asn	Pro	Asp	Pro	His
Asp	Gln	Leu	Thr	Thr	Ile
Lys	Asp	Ser	Phe	His	Cys
Gln	Glu	Met	Phe	Asp	Gln
Thr	Trp	Glu	Glu	Glu	Leu
Gln	Glu	Glu	Leu	Arg	Arg
Glu	Ile	Asp	Ile	Leu	Glu
Leu	Cys	Gln	Glu	Lys	Pro
Arg	Lys	Ser	Arg	Leu	Lys
Pro	Ser	Gly	Phe	Gln	His
Met	Asp	Lys	Arg	Lys	Ser
Ala	Ser	Pro	Thr	Ile	Ile
Thr	Val	Ser	Lys	Thr	Trp
Glu	Gly	Glu	Glu	Glu	Lys
Thr	Trp	Gly	Pro	Gly	Thr
Asp	Glu	Ser	Leu	Lys	Ser
Ser	Ser	Ala	Pro	Asn	Leu
Leu	Pro	Gly	Phe	Thr	Ser

WO 01/81555

PCT/US01/12992

Asp Glu Asp Ser Glu Gly Pro Gly Ser Gly Glu Ser Arg Leu Gln
 650 655
 His Ser Pro Ser Gln Ser Tyr Leu Cys Ile Pro Phe Pro Arg Gly
 660 665
 Glu Asp Gly Asp Gly Pro Ser Ser Asp Gly Ile His Glu Glu Pro
 670 675
 Thr Pro Val Asn Ser Ala Thr Ser Thr Pro Glu Leu Thr Pro Thr
 680 685
 Asn Ser Leu Lys Arg Gly Gly Ala His His Arg Arg Cys Glu Val
 690 695
 Ala Leu Leu Gly Cys Gly Ala Val Leu Ala Ala Thr Gly Leu Gly
 700 705
 Phe Asp Leu Leu Glu Ala Gly Lys Cys Gln Leu Leu Pro Leu Glu
 710 715
 Glu Pro Glu Pro Pro Ala Arg Glu Glu Lys Lys Arg Arg Glu Gly
 720 725
 Leu Phe Gln Arg Ser Ser Arg Pro Arg Arg Ser Thr Ser Pro Pro
 730 735
 Ser Arg Lys Leu Phe Lys Lys Glu Glu Pro Met Leu Leu Leu Gly
 740 745
 Asp Pro Ser Ala Ser Leu Thr Leu Leu Ser Leu Ser Ser Ile Ser
 750 755
 Glu Cys Asn Ser Thr Arg Ser Leu Leu Arg Ser Asp Ser Asp Glu
 760 765
 Ile Val Val Tyr Glu Met Pro Val Ser Pro Val Glu Ala Pro Pro
 770 775
 Leu Ser Pro Cys Thr His Asn Pro Leu Val Asn Val Arg Val Glu
 780 785
 Arg Phe Lys Arg Asp Pro Asn Gln Ser Leu Thr Pro Thr His Val
 790 795
 Thr Leu Thr Thr Pro Ser Gln Pro Ser Ser His Arg Arg Thr Pro
 800 805
 Ser Asp Gly Ala Leu Pro Ser Pro Ser Arg Asp Pro Gly Glu Phe
 810 815
 Pro Arg Leu Pro Asp Pro Asn Val Val Phe Pro Pro Thr Pro Arg
 820 825
 Arg Trp Asn Thr Gln Gln Asp Ser Thr Leu Glu Arg Pro Lys Thr
 830 835
 Leu Glu Phe Leu Pro Arg Pro Arg Pro Ser Ala Asn Arg Gln Arg
 840 845
 Leu Asp Pro Trp Trp Phe Val Ser Pro Ser His Ala Arg Ser Thr
 850 855
 Ser Pro Ala Asn Ser Ser Ser Thr Glu Thr Pro Ser Asn Leu Asp
 860 865
 Ser Cys Phe Ala Ser Ser Ser Thr Val Glu Glu Arg Pro Gly
 870 875
 Leu Pro Ala Leu Leu Pro Phe Gln Ala Gly Pro Leu Pro Pro Thr
 880 885
 Glu Arg Thr Leu Leu Asp Leu Asp Ala Glu Gly Gln Ser Gln Asp
 890 895
 Ser Thr Val Pro Leu Cys Arg Ala Glu Leu Asn Thr His Arg Pro
 900 905
 Ala Pro Tyr Glu Ile Gln Gln Glu Phe Trp Ser
 910 915
 1040 1045

<210> 10

<211> 357

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 7474560CD1

WO 01/81555

PCT/US01/12992

```

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

<210> 11
<211> 355
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> incyte ID No: 7474602CD1

<400> 11
Met Ala Arg Glu Asn Gly Glu Ser Ser Ser Ser Thr Lys Lys Gln
1      5      10      15
Ala Glu Asp Ile Lys Lys Ile Phe Glu Phe Lys Glu Thr Leu Gly

```

WO 01/81555

PCT/US01/12992

```

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

```

<210> 12
 <211> 224
 <212> PRT
 <213> Homo sapiens

 <220>
 <221> misc_features
 <223> Incyte ID No: 7475509CD1

 <400> 12
 Met Ala Ser Lys Leu Leu Arg Ala Val Ile Leu Gly Pro Pro Gly
 1 5 10 15
 Ser Gly Lys Gly Thr Val Cys Gln Arg Ile Ala Gln Asn Phe Gly
 20 25 30
 Leu Gln His Leu Ser Ser Gly His Phe Leu Arg Glu Asn Ile Lys
 35 40 45
 Ala Ser Thr Glu Val Gly Glu Met Ala Lys Gln Tyr Ile Glu Lys

WO 01/81555

PCT/US01/12992

```

50      55      60
Ser Leu Leu Val Pro Asp His Val Ile Thr Arg Leu Met Met Ser
65      70      75
Glu Leu Glu Asn Arg Arg Gly Gln His Trp Leu Leu Asp Gly Phe
80      85      90
Pro Arg Thr Leu Gly Gln Ala Glu Ala Leu Asp Lys Ile Cys Glu
95      100     105
Val Asp Leu Val Ile Ser Leu Asn Ile Pro Phe Glu Thr Leu Lys
110     115     120
Asp Arg Leu Ser Arg Arg Trp Ile His Pro Pro Ser Gly Arg Val
125     130     135
Tyr Asn Leu Asp Phe Asn Pro Pro His Val His Gly Ile Asp Asp
140     145     150
Val Thr Gly Glu Pro Leu Val Gln Gln Glu Asp Asp Lys Pro Glu
155     160     165
Ala Val Ala Ala Arg Leu Arg Gln Tyr Lys Asp Val Ala Lys Pro
170     175     180
Val Ile Glu Leu Tyr Lys Ser Arg Gly Val Leu His Gln Phe Phe
185     190     195
Arg Asp Arg Arg Arg Thr Lys Ile Trp Pro Tyr Val Tyr Thr Thr
200     205     210
Phe Leu Asn Lys Ile Thr Pro Ile Gln Ser Lys Glu Ala Phe
215     220

```

<210> 13

<211> 502

<212> PCT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 7475491CD1

<400> 13

```

Met Asn Lys Met Lys Asn Phe Lys Arg Arg Phe Ser Leu Ser Val
1      5      10      15
Pro Arg Thr Glu Thr Ile Glu Glu Ser Leu Ala Glu Phe Thr Glu
20      25      30
Gln Phe Asn Glu Leu His Asn Arg Arg Asn Glu Asn Leu Gln Leu
35      40      45
Gly Pro Leu Gly Arg Asp Pro Pro Gln Glu Cys Ser Thr Phe Ser
50      55      60
Pro Thr Asp Ser Gly Glu Glu Pro Gly Gln Leu Ser Pro Gly Val
65      70      75
Gln Phe Gln Arg Arg Gln Asn Gln Arg Arg Phe Ser Met Glu Val
80      85      90
Arg Ala Ser Gly Ala Leu Pro Arg Gln Val Ala Gly Cys Thr His
95      100     105
Lys Gly Val His Arg Arg Ala Ala Ala Leu Gln Pro Asp Phe Asp
110     115     120
Val Ser Lys Arg Leu Ser Leu Pro Met Asp Ile Arg Leu Pro Gln
125     130     135
Glu Phe Leu Gln Lys Leu Gln Met Glu Ser Pro Asp Leu Pro Lys
140     145     150
Pro Leu Ser Arg Met Ser Arg Arg Ala Ser Leu Ser Asp Ile Gly
155     160     165
Phe Gly Lys Leu Glu Thr Tyr Val Lys Leu Asp Lys Leu Gly Glu
170     175     180
Gly Thr Tyr Ala Thr Val Phe Lys Gly Arg Ser Lys Leu Thr Glu
185     190     195
Asn Leu Val Ala Leu Lys Glu Ile Arg Leu Glu His Glu Glu Gly
200     205     210
Ala Pro Cys Thr Ala Ile Arg Glu Val Ser Leu Leu Lys Asn Leu

```

WO 01/81555

PCT/US01/12992

```

215      220      225
Lys His Ala Asn Ile Val Thr Leu His Asp Leu Ile His Thr Asp
230      235      240
Arg Ser Leu Thr Leu Val Phe Glu Tyr Leu Asp Ser Asp Leu Lys
245      250      255
Gln Tyr Leu Asp His Cys Gly Asn Leu Met Ser Met His Asn Val
260      265      270
Lys Ile Phe Met Phe Gln Leu Leu Arg Gly Leu Ala Tyr Cys His
275      280      285
His Arg Lys Ile Leu His Arg Asp Leu Lys Pro Gln Asn Leu Leu
290      295      300
Ile Asn Glu Arg Gly Glu Leu Lys Leu Ala Asp Phe Gly Leu Ala
305      310      315
Arg Ala Lys Ser Val Pro Thr Lys Thr Tyr Ser Asn Glu Val Val
320      325      330
Thr Leu Trp Tyr Arg Pro Pro Asp Val Leu Leu Gly Ser Thr Glu
335      340      345
Tyr Ser Thr Pro Ile Asp Met Trp Gly Val Gly Cys Ile His Tyr
350      355      360
Glu Met Ala Thr Gly Arg Pro Leu Phe Pro Gly Ser Thr Val Lys
365      370      375
Glu Glu Leu His Leu Ile Phe Arg Leu Leu Gly Thr Pro Thr Glu
380      385      390
Glu Thr Trp Pro Gly Val Thr Ala Phe Ser Glu Phe Arg Thr Tyr
395      400      405
Ser Phe Pro Cys Tyr Leu Pro Gln Pro Leu Ile Asn His Ala Pro
410      415      420
Arg Leu Asp Thr Asp Gly Ile His Leu Leu Ser Ser Leu Leu Leu
425      430      435
Tyr Glu Ser Lys Ser Arg Met Ser Ala Glu Ala Ala Leu Ser His
440      445      450
Ser Tyr Phe Arg Ser Leu Gly Glu Arg Val His Gln Leu Glu Asp
455      460      465
Thr Ala Ser Ile Phe Ser Leu Lys Glu Ile Gln Leu Gln Lys Asp
470      475      480
Pro Gly Tyr Arg Gly Leu Ala Phe Gln Gln Pro Gly Arg Gly Lys
485      490      495
Asn Arg Arg Gln Ser Ile Phe
500

```

<210> 14
 <211> 791
 <212> PRT
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 2192119CD1

```

<400> 14
Met Trp Phe Phe Ala Arg Asp Pro Val Arg Asp Phe Pro Phe Glu
1      5      10      15
Leu Ile Pro Glu Pro Pro Glu Gly Gly Leu Pro Gly Pro Trp Ala
20     25     30
Leu His Arg Gly Arg Lys Lys Ala Thr Gly Ser Pro Val Ser Ile
35     40     45
Phe Val Tyr Asp Val Lys Pro Gly Ala Glu Glu Gln Thr Gln Val
50     55     60
Ala Lys Ala Ala Phe Lys Arg Phe Lys Thr Leu Arg His Pro Asn
65     70     75
Ile Leu Ala Tyr Ile Asp Gly Leu Glu Thr Glu Lys Cys Leu His
80     85     90
Val Val Thr Glu Ala Val Thr Pro Leu Gly Ile Tyr Leu Lys Ala

```

WO 01/81555

PCT/US01/12992

95	100	105
Arg Val Glu Ala Gly Gly Leu Lys Glu Leu Glu Ile Ser Trp Gly		
110	115	120
Leu His Gln Ile Val Lys Ala Leu Ser Phe Leu Val Asn Asp Cys		
125	130	135
Ser Leu Ile His Asn Asn Val Cys Met Ala Ala Val Phe Val Asp		
140	145	150
Arg Ala Gly Glu Trp Lys Leu Gly Gly Leu Asp Tyr Met Tyr Ser		
155	160	165
Ala Gln Gly Asn Gly Gly Gly Pro Pro Arg Lys Gly Ile Pro Glu		
170	175	180
Leu Glu Gln Tyr Asp Pro Pro Glu Leu Ala Asp Ser Ser Gly Arg		
185	190	195
Val Val Arg Glu Lys Trp Ser Ala Asp Met Trp Arg Leu Gly Cys		
200	205	210
Leu Ile Trp Glu Val Phe Asn Gly Pro Leu Pro Arg Ala Ala Ala		
215	220	225
Leu Arg Asn Pro Gly Lys Ile Pro Lys Thr Leu Val Pro His Tyr		
230	235	240
Cys Glu Leu Val Gly Ala Asn Pro Lys Val Arg Pro Asn Pro Ala		
245	250	255
Arg Phe Leu Gln Asn Cys Arg Ala Pro Gly Gly Phe Met Ser Asn		
260	265	270
Arg Phe Val Glu Thr Asn Leu Phe Leu Glu Glu Ile Gln Ile Lys		
275	280	285
Glu Pro Ala Glu Lys Gln Lys Phe Phe Gln Glu Leu Ser Lys Ser		
290	295	300
Leu Asp Ala Phe Pro Glu Asp Phe Cys Arg His Lys Val Leu Pro		
305	310	315
Gln Leu Leu Thr Ala Phe Glu Phe Gly Asn Ala Gly Ala Val Val		
320	325	330
Leu Thr Pro Leu Phe Lys Val Gly Lys Phe Leu Ser Ala Glu Glu		
335	340	345
Tyr Gln Gln Lys Ile Ile Pro Val Val Val Lys Met Phe Ser Ser		
350	355	360
Thr Asp Arg Ala Met Arg Ile Arg Leu Leu Gln Gln Met Glu Gln		
365	370	375
Phe Ile Gln Tyr Leu Asp Glu Pro Thr Val Asn Thr Gln Ile Phe		
380	385	390
Pro His Val Val His Gly Phe Leu Asp Thr Asn Pro Ala Ile Arg		
395	400	405
Glu Gln Thr Val Lys Ser Met Leu Leu Leu Ala Pro Lys Leu Asn		
410	415	420
Glu Ala Asn Leu Asn Val Glu Leu Met Lys His Phe Ala Arg Leu		
425	430	435
Gln Ala Lys Asp Glu Gln Gly Pro Ile Arg Cys Asn Thr Thr Val		
440	445	450
Cys Leu Gly Lys Ile Gly Ser Tyr Leu Ser Ala Ser Thr Arg His		
455	460	465
Arg Val Leu Thr Ser Ala Phe Ser Arg Ala Thr Arg Asp Pro Phe		
470	475	480
Ala Pro Ser Arg Val Ala Gly Val Leu Gly Phe Ala Ala Thr His		
485	490	495
Asn Leu Tyr Ser Met Asn Asp Cys Ala Gln Lys Ile Leu Pro Val		
500	505	510
Leu Cys Gly Leu Thr Val Asp Pro Glu Lys Ser Val Arg Asp Gln		
515	520	525
Ala Phe Lys Ala Ile Arg Ser Phe Leu Ser Lys Leu Glu Ser Val		
530	535	540
Ser Glu Asp Pro Thr Gln Leu Glu Glu Val Glu Lys Asp Val His		
545	550	555
Ala Ala Ser Ser Pro Gly Met Gly Gly Ala Ala Ser Trp Ala		
560	565	570

WO 01/81555

PCT/US01/12992

Gly Trp Ala Val Thr Gly Val Ser Ser Leu Thr Ser Lys Leu Ile
 575 580 585
 Arg Ser His Pro Thr Thr Ala Pro Thr Glu Thr Asn Ile Pro Gln
 590 595 600
 Arg Pro Thr Pro Glu Gly His Trp Glu Thr Gln Glu Glu Asp Lys
 605 610 615
 Asp Thr Ala Glu Asp Ser Ser Thr Ala Asp Arg Trp Asp Asp Glu
 620 625 630
 Asp Trp Gly Ser Leu Glu Gln Glu Ala Glu Ser Val Leu Ala Gln
 635 640 645
 Gln Asp Asp Trp Ser Thr Gly Gly Gln Val Ser Arg Ala Ser Gln
 650 655 660
 Val Ser Asn Ser Asp His Lys Ser Ser Lys Ser Pro Glu Ser Asp
 665 670 675
 Trp Ser Ser Trp Glu Ala Glu Gly Ser Trp Glu Gln Gly Trp Gln
 680 685 690
 Glu Pro Ser Ser Gln Glu Pro Pro Pro Asp Gly Thr Arg Leu Ala
 695 700 705
 Ser Glu Tyr Asn Trp Gly Gly Pro Glu Ser Ser Asp Lys Gly Asp
 710 715 720
 Pro Phe Ala Thr Leu Ser Ala Arg Pro Ser Thr Gln Pro Arg Pro
 725 730 735
 Asp Ser Trp Gly Glu Asp Asn Trp Glu Gly Leu Glu Thr Asp Ser
 740 745 750
 Arg Gln Val Lys Ala Glu Leu Ala Arg Lys Lys Arg Glu Glu Arg
 755 760 765
 Arg Arg Glu Met Glu Ala Lys Arg Ala Glu Arg Lys Val Ala Lys
 770 775 780
 Gly Pro Met Lys Leu Gly Ala Arg Lys Leu Asp
 785 790

<210> 15

<211> 1651

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 7474496CD1

<400> 15

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

WO 01/81555

PCT/US01/12992

Ala Glu Arg Arg Arg Glu Gln Ala Gln Arg Glu Ile Leu His
 170 175 180
 Glu Ile Gln Arg Arg Lys Glu Gln Ile Lys Glu Glu Lys Lys Arg
 185 190 195
 Lys Glu Met Ala Lys Gln Glu Arg Leu Glu Ile Ala Ser Leu Ser
 200 205 210
 Asn Gln Asp His Thr Ser Lys Lys Asp Pro Gly Gly His Arg Tyr
 215 220 225
 Ala Ala Ile Leu His Gly Gly Ser Pro Asp Phe Val Gly Asn Gly
 230 235 240
 Lys His Arg Ala Asn Ser Ser Gly Arg Ser Arg Arg Glu Arg Gln
 245 250 255
 Tyr Ser Val Cys Asn Ser Glu Asp Ser Pro Gly Ser Cys Glu Ile
 260 265 270
 Leu Tyr Phe Asn Met Gly Ser Pro Asp Gln Leu Met Val His Lys
 275 280 285
 Gly Lys Cys Ile Gly Ser Asp Glu Gln Leu Gly Lys Leu Val Tyr
 290 295 300
 Asn Ala Leu Glu Thr Ala Thr Gly Gly Phe Val Leu Leu Tyr Glu
 305 310 315
 Trp Val Leu Gln Trp Gln Lys Lys Met Gly Pro Phe Leu Thr Ser
 320 325 330
 Gln Glu Lys Glu Lys Ile Asp Lys Cys Lys Lys Gln Ile Gln Gly
 335 340 345
 Thr Glu Thr Glu Phe Asn Ser Leu Val Lys Leu Ser His Pro Asn
 350 355 360
 Val Val Arg Tyr Leu Ala Met Asn Leu Lys Glu Gln Asp Asp Ser
 365 370 375
 Ile Val Val Asp Ile Leu Val Glu His Ile Ser Gly Val Ser Leu
 380 385 390
 Ala Ala His Leu Ser His Ser Gly Pro Ile Pro Val His Gln Leu
 395 400 405
 Arg Arg Tyr Thr Ala Gln Leu Leu Ser Gly Leu Asp Tyr Leu His
 410 415 420
 Ser Asn Ser Val Val His Lys Val Leu Ser Ala Ser Asn Val Leu
 425 430 435
 Val Asp Ala Glu Gly Thr Val Lys Ile Thr Asp Tyr Ser Ile Ser
 440 445 450
 Lys Arg Leu Ala Asp Ile Cys Lys Glu Asp Val Phe Glu Gln Thr
 455 460 465
 Arg Val Arg Phe Ser Asp Asn Ala Leu Pro Tyr Lys Thr Gly Lys
 470 475 480
 Lys Gly Asp Val Trp Arg Leu Gly Leu Leu Leu Leu Ser Leu Ser
 485 490 495
 Gln Gly Gln Glu Cys Gly Glu Tyr Pro Val Thr Ile Pro Ser Asp
 500 505 510
 Leu Pro Ala Asp Phe Gln Asp Phe Leu Lys Cys Val Cys Leu Asp
 515 520 525
 Asp Lys Glu Arg Trp Ser Pro Gln Gln Leu Leu Lys His Ser Phe
 530 535 540
 Ile Asn Pro Gln Pro Lys Met Pro Leu Val Glu Gln Ser Pro Glu
 545 550 555
 Asp Ser Glu Gly Gln Asp Tyr Val Glu Thr Val Ile Pro Ser Asn
 560 565 570
 Arg Leu Pro Ser Ala Ala Phe Phe Ser Glu Thr Gln Arg Gln Phe
 575 580 585
 Ser Arg Tyr Phe Ile Glu Phe Glu Glu Leu Gln Leu Leu Gly Lys
 590 595 600
 Gly Ala Phe Gly Ala Val Ile Lys Val Gln Asn Lys Leu Asp Gly
 605 610 615
 Cys Cys Tyr Ala Val Lys Arg Ile Pro Ile Asn Pro Ala Ser Arg
 620 625 630
 Gln Phe Arg Arg Ile Lys Gly Glu Val Thr Leu Leu Ser Arg Leu

WO 01/81555

PCT/US01/12992

His His Glu Asn	635	640	645
Ile Val Arg Tyr Tyr Asn	640	Ala Trp Ile Glu Arg	
His Glu Arg Pro	650	655	660
Ala Gly Pro Gly Thr Pro	655	Pro Pro Asp Ser Gly	
Pro Leu Ala Lys	665	670	675
Asp Asp Arg Ala Ala Arg	670	Gly Gln Pro Ala Ser	
680	685	690	
Asp Thr Asp Gly Leu Asp Ser Val Glu	695	Ala Ala Pro Pro Pro	
700	705		
Ile Leu Ser Ser Ser Val Glu Trp Ser Thr	710	Ser Gly Glu Arg Ser	
715	720		
Ala Ser Ala Arg Phe Pro Ala Thr Gly Pro Gly Ser Ser Asp	725	Asp	
730	735		
Glu Asp Asp Asp Glu Asp Glu His Gly Gly Val Phe Ser Gln Ser	740		
745	750		
Phe Leu Pro Ala Ser Asp Ser Glu Ser Asp Ile Ile Phe Asp Asn	755		
760	765		
Glu Asp Glu Asn Ser Lys Ser Gln Asn Gln Asp Glu Asp Cys Asn	770		
775	780		
Glu Lys Asn Gly Cys His Glu Ser Glu Pro Ser Val Thr Thr	785		
790	795		
Ala Val His Tyr Leu Tyr Ile Gln Met Glu Tyr Cys Glu Lys Ser	800		
805	810		
Thr Leu Arg Asp Thr Ile Asp Gln Gly Leu Tyr Arg Asp Thr Val	815		
820	825		
Arg Leu Trp Arg Leu Phe Arg Glu Ile Leu Asp Gly Leu Ala Tyr	830		
835	840		
Ile His Glu Lys Gly Met Ile His Arg Asp Leu Lys Pro Val Asn	845		
850	855		
Ile Phe Leu Asp Ser Asp Asp His Val Lys Ile Gly Asp Phe Gly	855		
860	870		
Leu Ala Thr Asp His Leu Ala Phe Ser Ala Asp Ser Lys Gln Asp	875		
880	885		
Asp Gln Thr Gly Asp Leu Ile Lys Ser Asp Pro Ser Gly His Leu	890		
895	900		
Thr Gly Met Val Gly Thr Ala Leu Tyr Val Ser Pro Glu Val Gln	905		
910	915		
Gly Ser Thr Lys Ser Ala Tyr Asn Gln Lys Val Asp Leu Phe Ser	920		
925	930		
Leu Gly Ile Ile Phe Phe Glu Met Ser Tyr His Pro Met Val Thr	935		
940	945		
Ala Ser Glu Arg Ile Phe Val Leu Asn Gln Leu Arg Asp Pro Thr	950		
955	960		
Ser Pro Lys Phe Pro Glu Asp Phe Asp Asp Gly Glu His Ala Lys	965		
970	975		
Gln Lys Ser Val Ile Ser Trp Leu Leu Asn His Asp Pro Ala Lys	980		
985	990		
Arg Pro Thr Ala Thr Glu Leu Leu Lys Ser Glu Leu Leu Pro Pro	995		
1000	1005		
Pro Gln Met Glu Glu Ser Glu Leu His Glu Val Leu His His Thr	1010		
1015	1020		
Leu Thr Asn Val Asp Gly Lys Ala Tyr Arg Thr Met Met Ala Gln	1025		
1030	1035		
Ile Phe Ser Gln Arg Ile Ser Pro Ala Ile Asp Tyr Thr Tyr Asp	1040		
1045	1050		
Ser Asp Ile Leu Lys Gly Asn Phe Ser Ile Arg Thr Ala Lys Met	1055		
1060	1065		
Gln Gln His Val Cys Glu Thr Ile Ile Arg Ile Phe Lys Arg His	1070		
1075	1080		
Gly Ala Val Gln Leu Cys Thr Pro Leu Leu Leu Pro Arg Asn Arg	1085		
1090	1095		
Gln Ile Tyr Glu His Asn Glu Ala Ala Leu Phe Met Asp His Ser	1100		
1105	1110		

WO 01/81555

PCT/US01/12992

Gly Met Leu Val Met Leu Pro Phe Asp Leu Arg Ile Pro Phe Ala
 1115 1120 1125
 Arg Tyr Val Ala Arg Asn Asn Ile Leu Asn Leu Lys Arg Tyr Cys
 1130 1135 1140
 Ile Glu Arg Val Phe Arg Pro Arg Lys Leu Asp Arg Phe His Pro
 1145 1150 1155
 Lys Glu Leu Leu Glu Cys Ala Phe Asp Ile Val Thr Ser Thr Thr
 1160 1165 1170
 Asn Ser Phe Leu Pro Thr Ala Glu Ile Ile Tyr Thr Ile Tyr Glu
 1175 1180 1185
 Ile Ile Gln Glu Phe Pro Ala Leu Gln Glu Arg Asn Tyr Ser Ile
 1190 1195 1200
 Tyr Leu Asn His Thr Met Leu Leu Lys Ala Ile Leu Leu His Cys
 1205 1210 1215
 Gly Ile Pro Glu Asp Lys Leu Ser Gln Val Tyr Ile Ile Leu Tyr
 1220 1225 1230
 Asp Ala Val Thr Glu Lys Leu Thr Arg Arg Glu Val Glu Ala Lys
 1235 1240 1245
 Phe Cys Asn Leu Ser Leu Ser Ser Asn Ser Leu Cys Arg Leu Tyr
 1250 1255 1260
 Lys Phe Ile Glu Gln Lys Gly Asp Leu Gln Asp Leu Met Pro Thr
 1265 1270 1275
 Ile Asn Ser Leu Ile Lys Gln Lys Thr Gly Ile Ala Gln Leu Val
 1280 1285 1290
 Lys Tyr Gly Leu Lys Asp Leu Glu Glu Val Val Gly Leu Leu Lys
 1295 1300 1305
 Lys Leu Gly Ile Lys Leu Gln Val Leu Ile Asn Leu Gly Leu Val
 1310 1315 1320
 Tyr Lys Val Gln Gln His Asn Gly Ile Ile Phe Gln Phe Val Ala
 1325 1330 1335
 Phe Ile Lys Arg Arg Gln Arg Ala Val Pro Glu Ile Leu Ala Ala
 1340 1345 1350
 Gly Gly Arg Tyr Asp Leu Leu Ile Pro Gln Phe Arg Gly Pro Gln
 1355 1360 1365
 Ala Leu Gly Pro Val Pro Thr Ala Ile Gly Val Ser Ile Ala Ile
 1370 1375 1380
 Asp Lys Ile Ser Ala Ala Val Leu Asn Met Glu Glu Ser Val Thr
 1385 1390 1395
 Ile Ser Ser Cys Asp Leu Leu Val Val Ser Val Gly Gln Met Ser
 1400 1405 1410
 Met Ser Arg Ala Ile Asn Leu Thr Gln Lys Leu Trp Thr Ala Gly
 1415 1420 1425
 Ile Thr Ala Glu Ile Met Tyr Asp Trp Ser Gln Ser Gln Glu Glu
 1430 1435 1440
 Leu Gln Glu Tyr Cys Arg His His Glu Ile Thr Tyr Val Ala Leu
 1445 1450 1455
 Val Ser Asp Lys Glu Gly Ser His Val Lys Val Lys Ser Phe Glu
 1460 1465 1470
 Lys Glu Arg Gln Thr Glu Lys Arg Val Leu Glu Thr Glu Leu Val
 1475 1480 1485
 Asp His Val Leu Gln Lys Leu Arg Thr Lys Val Thr Asp Glu Arg
 1490 1495 1500
 Asn Gly Arg Glu Ala Ser Asp Asn Leu Ala Val Gln Asn Leu Lys
 1505 1510 1515
 Gly Ser Phe Ser Asn Ala Ser Gly Leu Phe Glu Ile His Gly Ala
 1520 1525 1530
 Thr Val Val Pro Ile Val Ser Val Leu Ala Pro Gln Lys Leu Ser
 1535 1540 1545
 Ala Ser Thr Arg Arg Tyr Glu Thr Gln Val Gln Thr Arg Leu
 1550 1555 1560
 Gln Thr Ser Leu Ala Asn Leu His Gln Lys Ser Ser Glu Ile Glu
 1565 1570 1575
 Ile Leu Ala Val Asp Leu Pro Lys Glu Thr Ile Leu Gln Phe Leu

WO 01/81555

PCT/US01/12992

```

1580      1585      1590
Ser Leu Glu Trp Asp Ala Asp Glu Gln Ala Phe Asn Thr Thr Val
1595      1600      1605
Lys Gln Leu Leu Ser Arg Leu Pro Lys Gln Arg Tyr Leu Lys Leu
1610      1615      1620
Val Cys Asp Glu Ile Tyr Asn Ile Lys Val Glu Lys Lys Val Ser
1625      1630      1635
Val Leu Phe Leu Tyr Ser Tyr Arg Asp Asp Tyr Tyr Arg Ile Leu
1640      1645      1650
Phe

```

<210> 16
 <211> 752
 <212> PRT
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 1834248CD1

```

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

```

WO 01/81555

PCT/US01/12992

	305	310	315
Gly Glu Glu Leu Lys Pro Tyr Thr Glu	Pro Glu Glu Asp Phe Gly		
320	325	330	
Asp Thr Lys Arg Ile Glu Val Met Val	Gly Met Gly Tyr Thr Arg		
335	340	345	
Glu Glu Ile Lys Glu Ser Leu Thr Ser	Gln Lys Tyr Asn Glu Val		
350	355	360	
Thr Ala Thr Tyr Leu Leu Leu Gly Arg	Thr Glu Glu Gly Gly		
365	370	375	
Asp Arg Gly Ala Pro Gly Leu Ala Leu	Ala Arg Val Arg Ala Pro		
380	385	390	
Ser Asp Thr Thr Asn Gly Thr Ser Ser	Ser Lys Gly Thr Ser His		
395	400	405	
Ser Lys Gly Gln Arg Ser Ser Ser Ser	Thr Tyr His Arg Gln Arg		
410	415	420	
Arg His Ser Asp Phe Cys Gly Pro Ser	Pro Ala Pro Leu His Pro		
425	430	435	
Lys Arg Ser Pro Thr Ser Thr Gly Glu Ala	Glu Leu Lys Glu Glu		
440	445	450	
Arg Leu Pro Gly Arg Lys Ala Ser Cys Ser	Thr Ala Gly Ser Gly		
455	460	465	
Ser Arg Gly Leu Pro Pro Ser Ser Pro	Met Val Ser Ser Ala His		
470	475	480	
Asn Pro Asn Lys Ala Glu Ile Pro Glu Arg	Arg Lys Asp Ser Thr		
485	490	495	
Ser Thr Pro Asn Asn Leu Pro Pro Ser	Met Met Thr Arg Arg Asn		
500	505	510	
Thr Tyr Val Cys Thr Glu Arg Pro Gly Ala	Glu Arg Pro Ser Leu		
515	520	525	
Leu Pro Asn Gly Lys Glu Asn Ser Ser	Gly Thr Pro Arg Val Pro		
530	535	540	
Pro Ala Ser Pro Ser Ser His Ser Leu Ala	Pro Pro Ser Gly Glu		
545	550	555	
Arg Ser Arg Leu Ala Arg Gly Ser Thr	Ile Arg Ser Thr Phe His		
560	565	570	
Gly Gly Gln Val Arg Asp Arg Arg Ala	Gly Gly Gly Gly Gly Gly		
575	580	585	
Gly Val Gln Asn Gly Pro Pro Ala Ser	Pro Thr Leu Ala His Glu		
590	595	600	
Ala Ala Pro Leu Pro Ala Gly Arg Pro Arg	Pro Thr Thr Asn Leu		
605	610	615	
Phe Thr Lys Leu Thr Ser Lys Leu Thr	Arg Arg Val Ala Asp Glu		
620	625	630	
Pro Glu Arg Ile Gly Gly Pro Glu Val Thr	Ser Cys His Leu Pro		
635	640	645	
Trp Asp Gln Thr Glu Thr Ala Pro Arg	Leu Leu Arg Phe Pro Trp		
650	655	660	
Ser Val Lys Leu Thr Ser Ser Arg Pro	Pro Glu Ala Leu Met Ala		
665	670	675	
Ala Leu Arg Gln Ala Thr Ala Ala Ala	Arg Cys Arg Cys Arg Gln		
680	685	690	
Pro Gln Pro Phe Leu Leu Ala Cys Leu His	Gly Gly Ala Gly Gly		
695	700	705	
Pro Glu Pro Leu Ser His Phe Glu Val	Glu Val Cys Gln Leu Pro		
710	715	720	
Arg Pro Gly Leu Arg Gly Val Leu Phe	Arg Arg Val Ala Gly Thr		
725	730	735	
Ala Leu Ala Phe Arg Thr Leu Val Thr	Arg Ile Ser Asn Asp Leu		
740	745	750	
Glu Leu			

<210> 17

27/46

WO 01/81555

PCT/US01/12992

<211> 501
 <212> PRT
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Inocyte ID No: 71584520CD1

<400> 17
 Met Pro Phe Gly Cys Val Thr Leu Gly Asp Lys Lys Asn Tyr Asn
 1 5 10 15
 Gln Pro Ser Glu Val Thr Asp Arg Tyr Asp Leu Gly Gln Val Ile
 20 25 30
 Lys Thr Glu Glu Phe Cys Glu Ile Phe Arg Ala Lys Asp Lys Thr
 35 40 45
 Thr Gly Lys Leu His Thr Cys Lys Lys Phe Gln Lys Arg Asp Gly
 50 55 60
 Arg Lys Val Arg Lys Ala Ala Lys Asn Glu Ile Gly Ile Leu Lys
 65 70 75
 Met Val Lys His Pro Asn Ile Leu Gln Leu Val Asp Val Phe Val
 80 85 90
 Thr Arg Lys Glu Tyr Phe Ile Phe Leu Glu Leu Ala Thr Gly Arg
 95 100 105
 Glu Val Phe Asp Trp Ile Leu Asp Gln Gly Tyr Tyr Ser Glu Arg
 110 115 120
 Asp Thr Ser Asn Val Val Arg Gln Val Leu Glu Ala Val Ala Tyr
 125 130 135
 Leu His Ser Leu Lys Ile Val His Arg Asn Leu Lys Leu Glu Asn
 140 145 150
 Leu Val Tyr Tyr Asn Arg Leu Lys Asn Ser Lys Ile Val Ile Ser
 155 160 165
 Asp Phe His Leu Ala Lys Leu Glu Asn Gly Leu Ile Lys Glu Pro
 170 175 180
 Cys Gly Thr Pro Glu Tyr Leu Ala Pro Glu Val Val Gly Arg Gln
 185 190 195
 Arg Tyr Gly Arg Pro Val Asp Cys Trp Ala Ile Gly Val Ile Met
 200 205 210
 Tyr Ile Leu Leu Ser Gly Asn Pro Pro Phe Tyr Glu Glu Val Glu
 215 220 225
 Glu Asp Asp Tyr Glu Asn His Asp Lys Asn Leu Phe Arg Lys Ile
 230 235 240
 Leu Ala Gly Asp Tyr Glu Phe Asp Ser Pro Tyr Trp Asp Asp Ile
 245 250 255
 Ser Gln Ala Ala Lys Asp Leu Val Thr Arg Leu Met Glu Val Glu
 260 265 270
 Gln Asp Gln Arg Ile Thr Ala Glu Glu Ala Ile Ser His Glu Trp
 275 280 285
 Ile Ser Gly Asn Ala Ala Ser Asp Lys Asn Ile Lys Asp Gly Val
 290 295 300
 Cys Ala Gln Ile Glu Lys Asn Phe Ala Arg Ala Lys Trp Lys Lys
 305 310 315
 Ala Val Arg Val Thr Thr Leu Met Lys Arg Leu Arg Ala Pro Glu
 320 325 330
 Gln Ser Ser Thr Ala Ala Ala Gln Ser Ala Ser Ala Thr Asp Thr
 335 340 345
 Ala Thr Pro Gly Ala Ala Gly Gly Ala Thr Ala Ala Ala Ser
 350 355 360
 Gly Ala Thr Ser Ala Pro Glu Gly Asp Ala Ala Arg Ala Ala Lys
 365 370 375
 Ser Asp Asn Val Ala Pro Ala Asp Arg Ser Ala Thr Pro Ala Thr
 380 385 390
 Asp Gly Ser Ala Thr Pro Ala Thr Asp Gly Ser Val Thr Pro Ala
 395 400 405

WO 01/81555

PCT/US01/12992

Thr Asp Gly Ser Ile Thr Pro Ala Thr Asp Gly Ser Val Thr Pro
 410 415 420
 Ala Thr Asp Arg Ser Ala Thr Pro Ala Thr Asp Gly Arg Ala Thr
 425 430 435
 Pro Ala Thr Glu Glu Ser Thr Val Pro Thr Thr Gln Ser Ser Ala
 440 445 450
 Met Leu Ala Thr Lys Ala Ala Ala Thr Pro Glu Pro Ala Met Ala
 455 460 465
 Gln Pro Asp Ser Thr Ala Pro Glu Gly Ala Thr Gly Gln Ala Pro
 470 475 480
 Pro Ser Ser Lys Gly Glu Glu Ala Ala Gly Tyr Ala Gln Glu Ser
 485 490 495
 Gln Arg Glu Glu Ala Ser
 500

<210> 48

<211> 366

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 7475538CD1

<400> 18

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

WO 01/81555

PCT/US01/12992

Gln Tyr Phe Phe Thr Ala Pro Leu Pro Ala His Pro Ser Glu Leu
 290 295 300
 Pro Ile Pro Gln Arg Leu Gly Gly Pro Ala Pro Lys Ala His Pro
 305 310 315
 Gly Pro Pro His Ile His Asp Phe His Val Asp Arg Pro Leu Glu
 320 325 330
 Glu Ser Leu Leu Asn Pro Glu Leu Ile Arg Pro Phe Ile Leu Glu
 335 340 345
 Gly

<210> 19

<211> 4224

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 2890544CB1

<400> 19

```

ccccggacgg cagcaggagg ccgagggcggg agcggcgggg gctgaggcgg 60
cgccggcgcc ggccgggaagg ggcctgggtgg tggagcgggg gggaggctga ggaggtccc 120
cctgcggggac gggcgcgggg acggctcdgg gggcgggggg ccggggcccg ggccggctcg 180
cgcgggggggt atgatgccc ggcctggcggg ccccgagatct cgtctctctc gccgcctcct 240
cagggcagcc ccagctggcg gctgagaacc agacacaccc ggggtatggc atcacagctg 300
caagtgtttt cgcctccatc agtgcgtcg agtgccttct gcagtcgcaa gaactgaaa 360
atagagccct ctggctggga tgtttcagga cagagtcgca acgacaaata ttatctccac 420
agcaaaaccc tcccgagcac acagggcgaa gccaaactcct ctcaacaggt agcaaatte 480
aacatccctg cttaagacca gggcctctct ctcccagntc ctgagtgga gcatattgt 540
gtacacgctg ctgatagctc gggcagtgct gctacatcaa ccttccaaag cagccagacc 600
ctgactcaaa gaagcaacgt ttcttggctt gagccatctc aaaaatgttg attgaaacga 660
aaagctgagg aagttgacgg caacggtagt gtgcagatca tagcaaaaca tcccccttc 720
atgctgcaaa acaggagctg ggtgggtgct gctgcacaaa caaccactgt gacacaaag 780
agtacagttt ccagcgagga aggggaltac cagctggctc agcatgagat cctttgtctc 840
atgacaaata gctatgaagt ctgggaattc ctggccgggg gacatttgg acgggtggct 900
aagtgctgga agggagcac caaggaaatt gtggctatta aaatcttgaa gaaccaccc 960
tctatgcca gacagagaca gattgaagtg agctctcttt ccgccttaag cagtgaatc 1020
gctgatgagt ataatttgtt ccgttcatac gagtgccttc agcataagaa tccacctgc 1080
ctgttttttt aaatgttggg ccgaacttta tatgattttc taaagcaaaa caaatttagc 1140
ccactgccac tcaagtacat cagaccaalc ttgcagcagg tggccacagc cttgatgaag 1200
ctcaagagtc ttggtctgat ccacgtgac cttaagcctg aaaaacatct cgtggttgat 1260
ccagttgcgc agccctaccg agtgaaggtc attgactttg gttctgctg tcaagtttcc 1320
aaagctgigt gctcaaccta cttaacgtca cgttactaca gagctcctga aactattctt 1380
gggttaccat ttgtgaagg taltgalatg tggctactgg gctgtgtgat agctgagctg 1440
ttcttgggat ggctcttta tcttggctct tgaagatals atcagattcg ctatatttca 1500
caaacacaa gcttgccagg tgaatatctt ctacgtgcgg gaacaaaac aaccagttt 1560
ttcaacagag atcctaattt ggggtaccca ctgtggaggg ttaagacac tgaagaacat 1620
gaactggaga ctggaataaa atcaaaaaga cctgggaagt acattttta ttgcttagat 1680
gacatggctc aggtgaatat gcttaccgac ctggaggaaa cagacatggt ggcagagaag 1740
gcagaccgaa gagaatacat tgatctgtta aagaaatgc tcacattga tgcagataag 1800
agaaataccc ctctaaaaac tettaacct cagtttgtga caatgacta ccttttggat 1860
tttccacata gcaatcatgt taagtcttgt ttccagaaca tggagatctg caacggagg 1920
gttccacatg atgatacagc gagtcagatc aagagtcctt tcaactacaa tgttgcccca 1980
aatcaacgca caaatctaac catgagcttc agcaatcagc tcaatacag gcacatcag 2040
gccgtgttcc tagcttccag ttctactgca gcagctgcta ctctttctct ggctaatcca 2100
gagtgttccac taactaaacta ccagtcagct ttgtaccat catctgctgc accagttcct 2160
ggagttyccc agcagggtgt ttccctgcag cctggaacca cccagatttg caactagaca 2220
gacccattcc aacagacatt tatagtagct ccaactgcgt ttcaacttgg actacagaca 2280
acacaaagg attctggatt cctgtgagg atgatacag ctgtaacctg tgaacccag 2340
gcaccagctg ctgagcaact acagattcag taaggagttc tcaacgagg aagctgtaca 2400
ccactaatgg tagcaactct ccaactctca gtacacacca tcaacccga gtatgcgtg 2460
ccctttactc tgagctgcgc agcggcgcg ccggcgctgg ttgaacagac tgcgctgta 2520

```

WO 01/81555

PCT/US01/12992

```

ctgcaggcgt ggccctggagg gactcagcaa attctcctgc ctccaacttg gcaacagttg 2580
cctggggtag cctacacaa cctctgccag cccacagcaa tgattccaga ggccatgggg 2640
agtgggacagc agctagctga ctggagggaat gcccaactctc atggcacaac gtacagcact 2700
alcatgcagc agccatcctt gctgactaac catgtgacat tggccactgc tcagcctctg 2760
aatgttgggt ttgcccatgt tgcagacaa caacaatcca gtccctccc ttccagaaga 2820
aataagcagt cagctccagt cctctccag tccctccag atgttctgcc ttcccaagtc 2880
tattctctgg ttggagcag tccctccgc accaatctt ctctaatc ctgtgtcct 2940
gtccaaagtc agctcagcc catcatcatt ccagatactc ccagccctcc ctgtagctgt 3000
atcactatcc gaagtgcac tcatgggaa gaggacaaca aatacaagg ccagtactct 3060
ggactgaagc caaggtctat tctctcagt tatgtcactg tcaatgattc tccagactct 3120
gactctctct tgagcagcc tbatctcact gatccctga gtgctctcc agccaatagt 3180
ggatccgttt tggggggcc tggcagagt ttggcagatg gcactggcac ccgactatc 3240
attgtgcctc cactgaaac tcagcttggt gactgcactg tagcaacca ggctcaggt 3300
ctcctgagca ataagactaa gccagtcgct tcagtgagtg ggcagtcctc tggatgctgt 3360
atcaccacca cagggctatg agctcaacgc gggggggaca gtgacagaca accactcaat 3420
cttagcagca accagcagtc atggggggt ccaacctcac agggagagag cagcaacca 3480
gcccccgcga ggcagcaggc gtttgtggcc cctctctcc aagcccca caactccag 3540
catggcagcc cgtacactc gacagggcac ccaactctg cccggcccc tgcctactg 3600
ccaaagccag ctcatctgta tctglatgct gcccgactt ctgctgtgc actgggctca 3660
accagctcca ttgtctact tctctccca cgggttctc caggctatg tgcagctat 3720
accactcaac ctgagacttt ggtgcacccg gtccctgcca gtgtgggc cagctctc 3780
actttgcca cgttgccccc tgcctagta caacaccagt ttgctccca atctacat 3840
gggtcttccc gaggctcaac aattacact ggtacccgc tgaactctac cagatcagc 3900
cagtatctct acttatagtt ggtgacatg agggaggag tatcatggct acctctctc 3960
ggcctcgtct tcttatcttt gggctatgga gactctctc ttctccctct tgaactctc 4020
tagccagcaa cttgtctgc agggcccaac tgaagcaga ggtttttctc tgggggaa 4080
tgtctcagtg ttgactgcat tctgtagtc ttctcaagt tggcctatt tttaataca 4140
tattttgtg acgtactctg gtactggaga gtcaatgcca ttctgctc cagcagaagt 4200
gctcagctcc ctgaatttgc cctg 4224

```

<210> 20

<211> 1736

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 7472693CB1

<400> 20

```

cgagagcgag agaggaact gcagaggagg aagctgcgcc gcagcccgag 60
ccgcccggca tcccgcgc cctctgcccc ggcgcgcgcc cccgcgcgcc cctcccagc 120
gcgcccgcgc cgcctctcc gcgcgcgcct cgtcgcccat ggcccggag aacggcgaga 180
gcagctcctc ctggaaaaag caagctgaag acatcaagaa gatcttcgag ttcaaaagaa 240
ccctcggaac cggggcttt tccgaagtg ttttagctga agagaaggca actgggaagc 300
tcttgcctgt gaagtgatc cctcaagag ccctgaagg ccaggaagc agcatagaga 360
atgagatagc cgtctcgaga aagattcaag atgaanaat tatgtccctg gaagacatt 420
atgaagcccc aattcaactg tacttggta tgcagctggt gtccggtga gactgtttg 480
accgagtagt ggaagaaggg ttttatcag agaaggtgc cagcactctc atccgccaag 540
tcttgagcgc cgtgtactat ctccacagaa tgggcatcgt ccacagagac ctcaagcccg 600
aaaatctctt gtactacagt caagatgagg agtccaaat aatgatcagt gactttggat 660
tgtcaaaaat ggaagycaaa ggagatgtga tgtccactgc ctgtggaact ccaggtatg 720
tcgctcctga agtctctgc cagaacactt acagcaagc cgttgactga tggctcatg 780
gagtgattgc ctacatcttg cctctggggt accctcctt ttatgatgaa aatgactcca 840
agctcttga gcagatcttc aaggcggaat atgagttga ctctccctac tgggagtaca 900
tctccgactc tgcaaaagac ttcattcgga acctgatgga gaaggacccg aataaagat 960
acacptgtga gcaggcagct cggcacccat ggaatcctg tgcacagcc ctcaacaaa 1020
acatccatga gtccgtcagc gccagatcc ggaaaaact tgcacagcc aaatggagac 1080
aagcctttta tgcacggcc gtcgtgagac acatgagaa actacacctc ggcagcgcc 1140
tggcagcttc aaatgcaggt gttctgagca gccctcagct agccagcaa aagactgtg 1200
cgtatgtacc aaacacagaa tccctcagct gacactgaag accagcctgg ggtggagag 1260
agggagccgc cabctgcaga gccactcctg tttgccagcc gctttctata cttaatacca 1320
tgtcatgca cctagagact ttttttaaca tgaatcact gggccgggtg cagtgtcca 1380

```

WO 01/81555

PCT/US01/12992

```

cgccgtgaat cccaacactt tgggaggttg aggcaggagg actgttttag ttcaggagtt 1440
ctaaagccag cctgacccac atgggtgaac ccaatctcta ctaaaatata aaaattagcc 1500
gggtgtgttg gcgagccact gtaatgtcag ctacttggga ggctgaggca ggagaatcac 1560
ttgaaccag gaagcggagg ttgaatgtag ctgagatcac accactgccc tccagcctgg 1620
gtgacagatt gagactccct ctcaaaaaaa aaagggaat catggaacac tcgtgggaac 1680
ctaggtattg catattccat ttacgggttg ggaatccagg gctcaagtc tcgcag 1736

```

<210> 21

<211> 1824

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 3107952CH1

<400> 21

```

aacggcttcc ggtgtcatgg ctgcttgaag tccggggagt cggtagggcg gctgcaggtc 60
cctccctcgg gagcccgctg tccggctggc ggagctgtga ccgcgggccc ggccggcctg 120
cctcaggcgt ggcgtcagct cccgtgtccg tgccttcaac ccacacccat ggccggatcc 180
gggtgtgctt ggggcgggga gccgcgcggt ttcttgaggc ccttcgggag gctgtggcag 240
gtacagagcc gtctgggtag cggctctccc gctcgggtgt atcgggttg ctgtggggc 300
aacctctggt cgcctccggg cgcctcaag cagtctcttc cgcagggaac caccgggct 360
gcgcctcttg ccgccagata tgggttccgc aaagagcggg cggcgttga acagttgcag 420
ggctcacgaa acatcgtgac ttgttatgga gtgtttacaa tccacttttc tccaaatg 480
ccatcacgct gtctgttgc tgaactcttg gatgtcagtg ttccggaatt gctcttat 540
tccagtcacc aggtgtgttc catgtggatg atacagcatt gtcccgaga tgttttggag 600
ggccttgctt ttcttcatca tgggggttat gtccatggcg acctcaaac acgtaacata 660
ttgtggagtg cagagaatga atgttttaaa ctccattgact ttggacttag cttcaagaa 720
ggcaatcagg atgtaaatga tattcagaca gaagggtatc gggctcaga agcagaattg 780
caaaattgct tggccaggc tggcctgcag agtgatacag aatgtacote agctgtgat 840
ctgtggagcc taggaatcat ttactgga atgttttcag gaatgaact gaacataca 900
gtcagatctc agaatgga ggcacacagt tctgtcttta ttgatacat atttgcag 960
aaagagtggt tgaatggcga aatccagcc tatcaactaa gagacatct caaagcag 1020
cttcagtggt atccagcag aagaattcct gctgaatgg catgtgcag cccatttttt 1080
agcaatctct ttgccctca tattgaagat ctgtctatgc ttcacctcc agtgctaga 1140
ctgtgaatg tgcctgatga tgattatctt gagaatgaag agaatatga agatgttga 1200
gaagatgtaa aagaggagtg tcaaaaatnt ggcacagtg tatctctact tgttccaaag 1260
ggaaatcctg gcagagga atgttttgtt gagtatgcaa atgctgtga ttccaaagct 1320
gcgcagaaat tactgactgg aagatgttt gatgggaagt ctgttgtgc taccattcac 1380
ccgctgtgag octacaagag gggatctct tatcaaacct tgccttaac agtaacctaa 1440
ggactgttct ctltttctcc tcttccattt ctgtgggtat tccacatat aatgcaggag 1500
taccctctta ccattttaag aaggtacttt atacatttat ttaactctac taatgtgcag 1560
ccattgcaca agcagtgaat gegtgcata catttggcac tgaataggac aagacctctc 1620
agctctacat tgaagggttt tagagcatcc atgtgggcaa cccctttttg tgcgggagag 1680
caggtgttgc tcttcagttt gttagcctaaa aaactcttaa ttatttcag gatcatgag 1740
caagatgaa taatatcatg tcttggtaaa tactaacaaa ttgttaggt ttgtgacat 1800
cattacagg ttattctct atgt 1824

```

<210> 22

<211> 2201

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 5544420CB1

<400> 22

```

atgaaccgat acacaacat gagacagttg ggggacggca cgtatgggag tgtgattatg 60
ggccagarta atgaatccg gpagctggtg gccatcaaaa gaatgaagag aagttctat 120
tcttgggttg aatgcattga ctgagagaa gtttaagctc tgaagaact taatcatgcc 180
aagtttata aattgaaga agttatcaga gaaaatgacc atctttattt tatatttga 240

```


WO 01/81555

PCT/US01/12992

```

tatatgaag aaaaactcta tcaattaatg aaagacagga acaagttgtt cootgaatca 300
gicacacagaa acattatgta tcaaatattg caagggtgtg cttttatoca taacacatgt 360
ttttttcata gggacatgaa aacagaaaac ttgcttttga tgggtccaga gcttgtgaaa 420
attgctgtatt ttggaactgc aagagaatta aggtcacagc caccatcac tgaattatga 480
tctacacagat ggtatcgtgc ccttgaagtt ttactbgagat cttcagttta tagttctcnc 540
attgatgtgt gggctgttgg aagtatcatg gctgaactct atatgttaag gccacttttc 600
ccagggtcaa gtaggttga tgaactctt aaactctgac aagttttagg gactccaaa 660
aaaagtgaat ggcagagagg atacagctg gcatcctata tgaactctg ttttccccag 720
tggtttcta taaacttcaa aactcttatt ccaatgccca gtaatgaagc tattcagctc 780
atgaacgana tgttgaattg ggtatccaaag aacagatcca cagcaagcca ggcattgaaa 840
aacacatatt ttcaagttgg tcaagttatta ggcctttcgt caaatcatct ggaatcaaaa 900
cagtttttaa ataagcagct gcaacacatta gaatcaaaagc catctttagt tgggttagag 960
cctaaagctc tgcaggatatt aatcagatnag gttgttggaac aacccagccc aaaaactagc 1020
cagcagcdaa tgcagcccat tcaagccgcca cagaacctga gcgtccagca accccaag 1080
caacagagtc agggagaaac gccacaaaac ctattcccca gcatctcca aacacatcca 1140
actaaagcaa atggcacact gactcataaa agtggtagga ggcgttgggg tcaagctatc 1200
ttcaagctcg gtagtagctg ggaagagttg gaggactatg atttcggagc ctccatctcc 1260
aagagaccaa gcatgggtgt ttttaagaa aaaaggaaaa aagattctcc atttcggctt 1320
ccagagccag taccctcagg ctccacccac tcagcagggg aaaaacagag cttactgtct 1380
gttacttccc taaaatctga ttccgaattg tcaactgtct caacctctaa acagactaac 1440
ttgaacaaat caagatattc tccaggtgta aatcccaaga aggttctct gatagccagt 1500
ggaagagaaa taaaccccca caattggagc aaccagttat tcccaagtc actgggagtc 1560
gttggtccag aacttgcctt caaagagagc aatgcgaag aaaaacttgg aagtatgct 1620
acttcaatc agtccagata tattctcttc ttctcaaaa aagaagtga gtcagctggt 1680
cagagggatc acttagcacc ttctaatgca acggtctcag aataacctg gaacacaaaa 1740
actgtctggg ggaagtttto aggaagtagt tataatccta cagcaaaaaa cctaataatt 1800
gtgaacgtg cagcagcccat tccctcagtg cagcggagga cagactgggt ggcgaagtab 1860
ggagccaccc ggtagggatc tatgtgtga aacctacagc cattgtccg tagagtagct 1920
gcaagttcct tgaacctggg aaatgtctac aaatgtctat tctactctag tctgtgaga 1980
aataagcaaa agtgggtact tggaaaggca aaaaatcacc ctattttac ttatttccaa 2040
gaatgcatt ttcttagcat cattgcacac agtgttgata tatgggtagg atgttacaaa 2100
gtattgata aactattgc caaagtatga agtattgat ctacattta aataatagta 2160
aatcaataa gaaccttaa aaaaaacaa ctccagaaa a 2201

```

<210> 23
 <211> 2974
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 7472832C1

```

<400> 23
gcagatcccc ggggatctag aacccagtg ccccccgggg ccccggggc ggtccgggt 60
gggtccaggg cggcaggtcc cgggctctcc ccatggcca ccgcccctc ttatccggc 120
gggtctactg gctctccggg gcgggggtct cctccggccc ccggggctc agagctcag 180
tcgcgcgcac cgtactgccc ccagatccc gcccggggt cggggctct ctttcaatc 240
cagatcgggc tgaccccgca gttcgtgctg ttgcgccgc cctccagag ggtctatgtg 300
aagcagctgg cctgttccat cgtggaccag aagtctcccg agtgtggct ctacggcctt 360
tacgacanga tctgtctttt caaacatgac cccagctgg ccaactcct gtcagctggtg 420
cgctgtctcg gagaatccca ggaaggcgac ctggttgagg tggctgctgc ggcctcgcc 480
accttcaggg acttcagat ccgcccgcac gccctcacgg tgcactcta tgggggcct 540
gccttctgtg atcactcggg gsgatgctc ttggccttag tgcgcaagg cotcaagtg 600
gatggtctgg ggtgaacta ccaaaagcgc tgrgcttcca gcatcccaa ctaacttagt 660
ggggcccgca aacggcgctt gtcaccaaag tctctggcca gtggccactc ggtgcgctc 720
ggcactcccg agtccctgcc ctgcacgggt gaagagctga gccgtagcac caccgaactc 780
ctgctctggc gtcocccgtc atcctcttcc tccctctctg cctcactgta taaggcgcc 840
ccacttgagc tggacaagat gctgctctcc aaggtcaagg tgcctcacac ctctctatc 900
cacagctata cccggccccc cgtttgcag gcttgcaaga aactcctca gggcctctc 960
cgccagggcc tgaatgcaa agactgcag tttactgtc aaaaagctg cgcacccgc 1020
gtccctaatg actgocctgg gaggccctt atcaatggag atgtgccat gggagggcc 1080
acagatttca gggaggtga caagagccc ctcagggatg agtcaagaga ctccggctc 1140

```

WO 01/81555

PCT/US01/12992

```

atccctggct cccactcaga gaatgcgctc caccgccagt agggaggagg agggcaggga 1200
ggcaaggccc agagctccct ggggtacatc cccctaataga aggtgggtga atcggtgcga 1260
cacacgacgc ggaatccag caccacgctg cgggagggtt ggggtggtta ttacagcaac 1320
aaggacacgc tgagaaagcg gcaactatgg cgcctggact gcaagtgtat cagctcttc 1380
cagaacaaac cgaacaaacg atactataag gaaatccgc tgtcagaat cctcacggtg 1440
gagtcggccc agaactccag ccttgtgcgc cggggcccca accacactg ctttggagtc 1500
gtactgcga atgcacacta cttcgtgggc gaggatgcct ccagacctc ggttgggcca 1560
agtggtcagc gggctgagcg cggccgggac tgagagacag ccattcccca ggcctgatg 1620
ccgctcatcc ttccggacgc acccagcgcc ccaggccacg cggcccaacg acagcttct 1680
ctgagcactc ctgtgtccaa cagtccagatc caaggcaatg tggacattgc cactgtctac 1740
cagatcttcc ctgacgaagt gctgggctca gggcagttt gagtgtctta tggagggaaa 1800
caccggagga caggccggga cgtggcagtt aagctcattg acaaactgcg cttccctacc 1860
aagcaggaga gccagctccg gaatgaagt ggcattctgc agagcctgcg gcatccggg 1920
atcggtgaac tggagtgcat gttccgagac cctgagaaag tgttttgtgt gatggagaag 1980
ctgcattggg acatgttggg gatgatcctg tccagtgaag agggccgggt gcttgagcgc 2040
ctcaccaggt tctctcatcc ccagatcctg gtgtgtttga gacacctca cttcaagaa 2100
atgttccact gtgacttgaa accagaaaac gttgtgctgg catcagcaga cccatttct 2160
caggtgaagc tgtgtgactt tggctttgct cgcactatcg gcgagaagt gttccgcgc 2220
ttagtggtag gcaagccggc ctacttgcca ccgaggtgc tgcataacca gggctacaa 2280
cgtcgtctgg acatgtgttc cgtggcggtg atcatgtacg ttagcctcag cggcacttc 2340
cctttcaacg aggtatgaga cgtcaatgac cagatccaga agccgcctt catgtacgc 2400
ccagccctcc gggcccat ctccagctga gccattgacc tcatcaaca cctgtgcag 2460
gtcaagatgc gcaaacgcta cagcctggac aaatctctca gccacccctg gttacaggag 2520
taccagactg gcttggaact ccgagagctg gaggggagga atacatcacg 2580
catggagagc acgacgcgcg ctggggagcg tttggagcag agcatccgt gcttgggtct 2640
gggctggcca cggacaggga tctcgggtgg gctcgtccac caagagaca cgaatgcag 2700
gggctggggc agcgcctcag tgtctcttga ggtcctgtgc cctcgtccag ctctgctct 2760
ccacagcgtt tcttcacagg atcccagcaa tgaactgttc taggaaagt ggtctctgc 2820
ccaaactggc tgggacacgt ggggagtggt gtggggggag ctatttccaa ggcctcccc 2880
tgtttcccca gcaatcaaaa cggactcctc tctgcccact ggccttgatc tcaaaaaaa 2940
aaaaaaaacg aaaaaaacg cpgtgcgcag ctta 2974

```

<210> 24
 <211> 3548
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Inyte ID No: 1551456CB1

```

<400> 24
gggtacgagc cggatcacta ttacggcgca gtgtgctgga aagtgcctca agagatctta 60
agctggagcg accgggtctg aattccagac tcttcccac caccacact tccactccaa 120
ctggagctat accacagacc ctttcaggga ggttggcgga ctcttcaccc tggacagttc 180
cttactgtat gtcaaacctg aaatgagagc ggaacccatc agacagagga gcttggggaa 240
cgtccggccg ggcacaggct ctgagaaagt ggtttcttgg aaataatgca aagagagctg 300
gacattctat ccttggctcc cgtctggcca actcaccgtt gccaaagata gtgcagttk 360
tggcgaggga agatggcagc gatgacttct atcagctgaa gatcctgacc ctggagagga 420
ggggggacca aggtatagag agccagggaag agcggcaggc caagatgctg ctgcacaccg 480
agtactcaat gctgtctctc ctgcacacgc aggatggcgt ggtgcacac caccgctct 540
tccaggaccg caactgtgaa atcgttgagc acacagaatc cagccggatg gttagaaga 600
tgaaggaagc catctgcctc gtcttggact gctctgtgto tcatgacttc agcgtataga 660
ccgttgacct catcaacctg cagcactaag tcatcaagga gaagaggctc agcagagagg 720
agacttggtt aatcttctac gacgttgttc gctgtgttga ggccttgca cagaataata 780
tcgtgcacag agaccgtgag ctgggggaaca tgggtgtcaa caagaggaca catcggataa 840
ccatcaccaa ctctgccttc ggggaagcgc tggtagcga gggggacctg ctgaaggacc 900
agagggggag ccttgctcac atcagtcctg agtgtctcag cggccggccg taccgtggca 960
agccagtgta catgtgggac ctgggggtag tgcctctcac catgctgtat ggcagttccc 1020
cctctctaga cagatcccg caggagctct tccgcaaat caagctgccc aggtatcca 1080
tctctgagga tggacgggtt tctgagaaca cgtgtgtct catccggag ctgctgtctc 1140
ttaccccca cgcgcgctg gccgcgcgc agtctctgga ggcctcagt gccactattg 1200
catcttgga gtccctgca tctctgagtg ggcctttgca agtgttctt gacattgatg 1260

```

WO 01/81555

PCT/US01/12992

```

acaaatgag caatgaggat agctccacag aggcgaaggt gacggagag tgctccagat 1320
acgagtttga gaactacatg cgtcagcagc tgctgctggc cgaggagag agtccatcc 1380
atgacgccc gagctgggta cccaaggggc agttcggcag cgcaccccg gtgcgacggc 1440
tggggcacga cgcacagccc atgacctcct tggacacggc catcctggcg cagcgctacc 1500
tggggaaata acagcctcag cgggggcccac cagcactgct gccacttctt ccagccccc 1560
ccaaagggct ggctgtccgg gctgggccc gtatgtctgg abctctcccg gccacatag 1620
ggacaggga ggaacaggga cagccacaggt cccacgtggg gtccagcag gtaccagaa 1680
gttacctttt gggtgattg ctogattgtt tgggttttaa abctgagag cctagataac 1740
taactctgtt ttaatacaga tgttttaata taccctgttc totttaaca tctgtctct 1800
ggactgagca agaggaggga gggagcctgc tccacccact ccaggccctt cccagcggc 1860
oaddaactga odtggggggc tgctccccc agtccaaata agctgaaagt gcagctcgt 1920
gcaggcccca gagcgagctt cccctccctc ctgctctccc aggcctctgc cagagcctct 1980
ttcctgctct ctctttctga tccagggccc ttagtccag ctttggaaaa ccttcacctc 2040
atcttaaac gaactcaaat atatttattt ttttaacata caaatttctc tccatctct 2100
aggtggctca gtccatggcc actccctgct cccagcctgg ctggacagca aggaatccac 2160
agccacacag tgaactcctt cctcccccag aggcgggaa gccctctctg ccagtcctct 2220
tcccttttca gccacacagt cctctctctg tgcgggtgat gggaggcctt tctagacctg 2280
gtctttcttc tccgctctca gtggtctctc tggaggtctg tacacggcg ttaacctgtt 2340
ccctttctta tccctcccg tggtaactgag ctacagtgga ctccagtgct gaaggggccc 2400
atgggtggg ctgcagggct ggccgtgagc gggggctgct tgcagctctc cctagcctac 2460
tctgtgtttt aggggtgggt ggggaactat cccagtgccc ttgctctcta atagatgag 2520
tgactctccc ggtagacact agcaagggtc ctccatgtgt gtgaggact caggagatt 2580
gtaggagatt ggggacctg cctgcttggc ttgagacag cctgctgct cttttgagcc 2640
gagattttga agtggatgct cgtcttgcca gaattgctgt tctccacaga atgcacctc 2700
cccttgctct tactggaact ggccctgctt gatgccaagc aagacacctt cccagaggc 2760
ctaccccca tatgtctca gagggctga gtgtccctc caggcagta tgggacctga 2820
ggccctctct gcttggtctt gctccacagt ggggagtgta ctgtgttct cagagtgta 2880
agctgtctcc tccctcaaaa aagctgactc actgtgagt accttggga agttcccaa 2940
cctctctgtg cctcagtttc cccatctgga aaaaatgggg ccaactcttg ccagcagtag 3000
cagggtgccc cagccctctt tctcccatg cccatccag cacttggcg actcatgct 3060
ctgctctagt ggcctctgtg gactcactg gagccagca cttaactccc ctgagcagc 3120
agcttgctc tgttcagct gtccagcgt gaggggctg gtctgtgct gtggggctg 3180
ggatgacct ctctctata ttattttat agaaagtct ctgcgggag ggaatgag 3240
tccggcttag gctccacag ccttgactgt cctactgga ggcctgag ctgggagc 3300
gcccgtctga cgggagccc ctctcagcag cccacgggt cctccagc tgctgccc 3360
gcgtgctct tctctctct ttcaagaaa tagcccgag gtctgcaag cctctcaga 3420
cagactgggc ccttccagg tcaagcctg tgtctgatga catlcttgt gaagcaagg 3480
agggaggat gggtagccc tcaactgggt tccacacagt agagaagtcc tatgttaag 3540
aaacggaaaa agtcacaaaa aagtttgtat aagacatat tttgtacta catggggact 3600
ctctctgcat gtcagcaata aaacttctct atctggaaaa aaaaaaaa 3648

```

<210> 75
 <211> 4719
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 2589355CB1

```

<400> 25
ggggagcggc cgtcgcacaa gccaaagcgc gctgcaaac ctcggcccg ccggcgctcc 60
tgctccgctgt gtctagcagc ggggcccagc atggtcatgg cggatggccc gaggcacttg 120
cagcgccggc cgttcgggtt ggggtttctac gacatcgagg gcacgctggg caagtggcaa 180
cttcgctgtg gtgaagctgg ggcggcacag gatccacagc agggaggtgg caatanaaat 240
aatcgataag tctcagcctt ggtatgcatgt gaaccttgag aaaaactacc gagagataca 300
aataatgaaa atgttagacc accctcacat aatcaaaact tatcaggtua tggagaccaa 360
aggttatgtt taacttctga cagaatctgc caaaaatgga gaaatttttg actatcttgc 420
taatactgca cgttcaaatg agtctgagc caggcgcaaa tctgtctgac cctgtctgac 480
tgtgtatgat tgtctgtgtt ggaagattgt gcaactgac ctcaagctg aaaaactcct 540
gctgtgatac aacatgata tcaaatagc agatttcggt ttggaaatt totttaaag 600
tggtaactgt ctgcaacat ggtgtggcag ccccccctat gcagccccc aggtctttga 660
agggcagcag tatgaaggac cacagctgga catctggagt atggagttg ttctttatgt 720

```

WO 01/81555

PCT/US01/12992

```

ccttgctgtt ggagctctgc cctttgatgg accgactctt ccaattttga ggcagagggg 780
tcctgaagga agattccgga ttccgtattt catgtcagaa gattgcagag accttatccg 840
aaggatgttg gtccctagacc catccaasog gctaaccata gcccaastca aggcagatata 900
atggatgctc atagaagttc ctgtccagag acctgttctc talccacaag agcaagaaaa 960
tgagccatcc atcggagagt ttaalgagca ggttctgcga ctgatgcaca gctttggaat 1020
agctccagag aaacacattg agtctttgca gaacaagagc tatcaacct ctgctggcat 1080
ttattttctg ttgggtggagc gctgaatc acatcggagc agtttccagc tggagcagag 1140
acttgatggc cgcacggctc ggcctagcac catgtctgag caaacagttg ccaggtccac 1200
gaatgtgggg ctccacgtga coatgnattc accgaacatg aggtctgtgc gatctgacct 1260
ctctcccccag ccatccaacg tggagccttt ttctattcca gcatctggct ctacggcgga 1320
agctcctatt atggaagaag agtgtgtgga cactccaaag gtcaatggct gtctgcttga 1380
ccctgtgctc cctgtcctgg tgccgaaggg atgcagatca ctgcccagca acatgatgga 1440
gaactccatt gacgaagggc tggagacaga agagagggcc gaggagayco ccgtcctatc 1500
ctttgagcca ttctagtcca cagcagcggc gacagagcgg caactctgtc cagaagtgc 1560
caatcaactg gtctgtatgc ctggggcagg gaaaattttc tccatgaatg acagcccttc 1620
ccttgacagt gtgactctg agtatgat atgggtctgtt cagagggacc tgactttct 1680
ggagacaaac ccttccctta aggacatcat gttagccaat cagcttccac ccgcctatgc 1740
atctgctctc ataagcttga gactaccaa cccagccatg caggctctga gctccagaa 1800
acagagggcc cacaacaggt ctccagtgag ctccagagag ggcgcagag catcagatgc 1860
ctctccaccc cagggaattg tagctattag acaacactct cagaactctg ctgagaccaa 1920
aggaattcta gatttgaaac aagtgcagtc gttgtatgaa caactagcac cggagcagga 1980
ccctaaactc cgcgcggcgg ctctcagct ccaggacctt gctagcagct gctctcagga 2040
agaahtttct cagcagcagg aaagcgtctc cactctccct gccagcgtgc atccccagct 2100
gtccccctgg daagcctgag agacccagta cctgcagcac agactccaga agcccgact 2160
tttgctaaag gccacagaca cctgtcagct ttattgcgaa gaacacacgc ggagccttga 2220
gcagcagctg caggaacata ggtctcagca gaagcgcact ttctttcaga agcagctcca 2280
actgcagccc tatttbaatc agatgcagat agcagagagc tcttaaccac agccaaagta 2340
cgactgcccc ctcccccgcc agggagctcc accgctcttc nagaagggcc aacgttccag 2400
cctgacccag cccctagacc ccgtctctga cgcctctctc gagcagatgc aatcagccgc 2460
ttctctcagc cagtaccag agatgcagct tcagccctgc cctccactt ccggtctccg 2520
ggctgctctc cctctgcccc cgcagctaca gcagcagcag ccgccacccg caaccacccc 2580
tccacacca cagcagccag gactgtcccc agccccctta cagttctctt atcagacttg 2640
tgagctgcca agcgtgctt cccctgcgcc agactatccc acctcctgtc agtatcctgt 2700
gcatggagcc cagcagagcc acctaacggg gccagactgt cccagagccc caggactgca 2760
agaaactccc tccagctaac accactagc cctctctgag ctactctgac tcttttattg 2820
tgaaatgcta gacgtgtgg atccacacaa caacgggtat gtcctgttga attagttcca 2880
gcaaggaat tgaggtgggt caggtgaagg aayagtgcat gttctatttt ttattccagc 2940
ottttaaatt taagcttat ttctctgccc tctctctaac ggggagaaat cagcagccc 3000
aacctggaatc agaggtctct gctgggtgg atgtgtctc ctctgtgtc tgccccacca 3060
caagttttc tgtggcaagt gctggaacat agttgtatgc tgaggtctct gccctctggg 3120
caggtggagc aagctctcga gggcagcact gacaaatgt tctctaaaga gacattcaga 3180
cccaggtctt atgcaggatt acatccgttt attatcaagg gcaaccttgg tgaagcaga 3240
aaggtgtgtt gctattgcat atatatgggg gaaagggcaa tatatttttc actgaagctg 3300
agcaccacca tattgtccca aggaacatca agnagacatc aggaactcag atgcacagga 3360
aataaaggaa agctgtgctt tgtctatgaa tctcaagttc ctactgctg atgcagattg 3420
tcccccagg ccatacaaaa gcagtgggc atgagctgtg ttccagggc cactaaata 3480
cagctgttgc tgacccagaa aaccgcttcc atctccatcc ggaagcaggt gacacacccc 3540
ttcaagagct gccctgggtt gccgagtgc aqaatatac caggactcca gaggtgtcac 3600
acgtggaaat gacaggagac ccgcccctgt ggaagcaggg gcccaagaaac tcaagaaagc 3660
atcaagagca ccagccctgg gccagggag acaggtcttt cctgcagttt ctogtgagca 3720
ctgctggctt ggggcagtc ggtctccagg gtaactgttt tctcttttcc gatgtataaa 3780
ctactttgac cttaacatct atgttgtat tagtttattg agctttgtat atttggacag 3840
ttctatatag ggtctagaga ttttaaggac atgataaatg aacttttctg tcccatgtga 3900
agtggtatg cgtgtccttt cccccagatc atgctttaat tcttttcttt ctgtagaaac 3960
caacagtttc catttatgtc aatgtatat ccaagtcac ttcaagtttt gttttccacc 4020
atgtgggaat cagcatcttt aatttcgkta aagttttgac ttgtaatgaa atgttccagt 4080
attacagcaa tattcaaga aagaaacaca gatgtgttaa caatttcaagc agatcatctg 4140
ccaaacatta tattcaaac aaaaactaac caacactaac aattcagtcn tcaagtagag 4200
taaaaattag atctcacagc tatctaactg tatccctaac attttgctac aatttgctac 4260
ttggacagtt aacatccagg ttttacaaag tcaaggttaa ttctaaagat gatcatctct 4320
gccttttaga atgcttgttc ccatcagcag atgaatgtgt taagcacaaa gcattctctc 4380
taagcacaaa agagaggagc taactgatgc tgcacttaca aaacaccttt aagttgctct 4440
tctattttgt agttagcgtt caggcaggtg acgtgtggaa agtctagggg gttccactct 4500

```

WO 01/81555

PCT/US01/12992

```

ggccatggga gccagctcc tcccaacgtc ggtaacttga gcagtccttg ttgctggcca 4560
ggagctgctt ggtccgcagc gctcaacatg ggtgccagga tgcttcgcag aggaactgtg 4620
ctcacgggtg gaacttggtgt cagtcgggaa ggccagtggt gggactgtca ctttttgtat 4680
ttataacaca cgtgaaatc aggaagaatg aataagctt 4719

```

<210> 26
 <211> 1651
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 4357117CB1

```

<400> 26
atgaggattg ttgttttagt gaaaaccaca cagccctgg gagccaccat caagcccaac 60
gagatgacag gggacatctt ggtggccagg atcatccacg gtgggtggc ggagagaagt 120
ggtttgctat atgctggaga caaactggta gaagtgaatg gagtttcagt tgaggggactg 180
gacctgaaac aagtgtatca tattctggcc atgtctcgag gcacaatcat gttcaagggtg 240
gttccagttct ctgaacctcc tgtgaatagc cagcagatgg tatatgtcgt tgccatgact 300
gagtactggc ccaagaggga tcccgacatc cctgcgatgg acgttggatt gcctttccag 360
aagggggaca tcttcacgat tgtggaccag aatgatgccc tctgttggca ggcccgaaaa 420
atctcagac ctgtacactg cgtggggttt gttccctcta accactctct gaagagggaag 480
caacgggaat tctggtgttc taaagcgtac cagcctcaaa cctgcctcaa gtcaactcaa 540
gcaagaggag agttttgttg ctacggctag aagtcttcta taggtaggtt cagcccgctg 600
catgccagtg tgtgtgtcac cggcagctgc taccgtgcag tgggtgccc ttacggaggag 660
gtgggtgaggt accagcgacg ccttcagac aagtaccgac tcatagtgtc cataggacc 720
tctgggtgtg gagtaaatga gctcagaaga caacttattg aatttaatcc cagccatttt 780
caaatgtctg tgcacacac tactcgtact aaaaagagtt acgaatgaa tggcgctgag 840
tatcactatg tgtccaaagg aacatttgaa aacctcatat atagtcacag gatgctggag 900
tatggtgagt acaagggcca cctgtatggc actagtgtgg atgctgttca aacagtctt 960
gtcgaaggaa agatctgtgt catggaacta gagcctcagg atattoaaag ggttcgaacc 1020
catgaactga agccctatgt catatttata aagccatcga atatgaggtg tatgaaacaa 1080
tctcggaaa atgcgaaggt tattactgac tactatgtgt acatgaagtt caaggtgaa 1140
gaactacag agatggaaa tttagcccaa agaatggaaa ctcaatttgg caatttttt 1200
gatcatgta ttgtgaatga cagcttgac gatgcatgtg ccagttgtt gtctgcata 1260
cagaaggctc aggaagagcc tcagtgggta cagcaacat ggaattccct agatactgag 1320
tctcaatgag actctgtgt taatgttga gttttaacac tgaacccttg atacagcgt 1380
ccatgtttgc aatctaaaa aacagtattt gacctattt aatgtgtaca actttaaaag 1440
tgcagcaatt tattaattaa tcttatttga aaaaaattt tattgtatgg ttatgtggt 1500
acctatttta acttaatttt ttttcttcta cctcatatgc agctgtggtt gaatatgaa 1560
taatgttaag tcaatgagta tgagaacctt tcgcagattt cacatgatc ttttaagatt 1620
taaaacaaa gctttcttaa aaaaaaaa a 1651

```

<210> 27
 <211> 3141
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 5511892CB1

```

<400> 27
atggagccct ccagagcgct tctcggtgc ctacgagcgg ccgccgctgc cggcccgccg 60
ggggaggatg gagcaggggc cggggccgag gagggaggag aggagaggga ggaaggcgcg 120
gcggcggttg gcccggggga gctgggtgc gaacgcgcgc tgcctactg gaaggcggtg 180
ttcgagtaca agcgggcggg cagggacgag ctgacctgcg ggcctgggca cgttggtggag 240
gtcgtgtcca aggaactgca ggtgtccggc gacgagggct ggtggaccgg cagctgaac 300
cagcggggtg gcatcttccc cagcaactac gcaacccggc gcaagcgcct ctcacagcgc 360
tgcagccccc gggggagaga agaaattgat ttggcgagag tcaaccttga agagattatt 420
ggcaccgggg gctttgggaa ggtctatgt gctttctgga taggggagta ggttgcgtg 480
aaagcagctc gccacgacc tgaatgggac atcagccaga ccatagagaa ggttgcgca 540

```

WO 01/81555

PCT/US01/12992

```

gagggcaagc tcttggccat gctgaagcac ccccaacaca ttgcccatag aggggtatgt 600
ctgaggagag ccaacccctg ctgtgtcatg gagtttgctc gtggaggacc ttggaataga 660
gtgttatctg ggaagaaggat tccccagac atccctgtga attgggctgt gcagattgac 720
agaggatga actacttact tgatgaggca attgttccca tcabccaccg cgaacttaag 780
tccagcaaca tatgatccct ccagaagggt gagaatggag acctgagcaa caagattctg 840
aagatccactg attttggcct ggctggggaa tggcaaccga ccaaccaagt gagtgcggca 900
gggagctatg ctggagtggt acccgaagtc atccgggcct ccattgttct caaaggcagt 960
gatgtgtgga gctatgggggt gctactttgg gagtgtgtga ctggtgaggt gcccttccga 1020
ggcatttgatg gotttagcgt cgttatggga gtggccatga acaaactcgc ccttccattt 1080
ccttctactg gcccaagacc ttttgcacaa ctctgtgaag actgtggaa tcttgatccc 1140
cacctacgac catctttcac gaatatccctg gacccgctaa ccaacctaga ggaactcgtt 1200
ttctttgaaa tgcaccaaggc ctcttctcac tgcctgcagg acaactggaa acagagattt 1260
caggagatgt ttgaccaact cagggccaaa gaaaaggaa ttcgcacctg ggaggaggag 1320
ctgacgcggg ctgactgca cgaagaagac caggaggaa tctgtcggcg tcggagcag 1380
gagctggcgg agcaggagat tgacatccctg gaacgggagc tcaacatcat catccaccag 1440
ctgtgccagg agaagccccc ggtgaagaaa cgaagggca agttcaggaa gaggcggctg 1500
aagctcaagg atggaacccg catcagctc ccttctggtt tccagccaa gttcagggtg 1560
caggccctccc ctaccatgga taaagggaag agttctatca acagccgtc cagtccctct 1620
gcaagcccca ccatcatccc tgcctctgga gccatccagt gtgagactgt tagcaaaccc 1680
tggggcaggc gctcagctgt cccaaaggag gaagggggag aggaggagaa gaggggccca 1740
aagaaagaga caggagctgt ggggcagggg agccttggltc agagagagct tgcctcggga 1800
gatgaagctc tcaagctcgt ggttagatga tatcaagagt ggtcgtccag tgcctccaac 1860
ctggtgaagg gcccaaggag tagcccgccc ctgcacaggt taccacagct tatggagatg 1920
ggttaagltta caggagatga ggaactgtaa gggccagggg gtggagagag tgcgtcacag 1980
catctaccca gccagctcta cctctgtatc ccatctcttc gtggagagga tgcagctggc 2040
ccctccagtg atggaatcca tgaggagccc accccagtca actcggccac gagtaccctt 2100
cagctgacgc caaccaacag cctcaagcgg ggcgggtccc accaccgcgc ctgagagggt 2160
gctctgctcg gctgtggggc tgttttggca gccacaggcc tagggtttga ctgtcgtgaa 2220
gcttgccagt gccagctgct tcccctggag gacccctgag caccagcccg ggaggagag 2280
aaagcaaggg aggytctttt tcaagagtc acccgtcttc gtggagcac cagcccccga 2340
tccggaagcg ttttcaagaa ggaaggagccc atgctgttgc tagggagacc ccttgcctcc 2400
ctgagctgctc tctccctctc ctccatctcc gagtgcacct ccaacgcctc cctgtctgca 2460
tccagtagcg atgaatttgt cgtgtatgag atgcccagta gccacgtcga gggccctccc 2520
ctgagctcat gtaaccacca ccccctggtc atgtctcagg tagagagctt caaacagag 2580
ctcaacccat ctctgactcc caccatgttc accctaccca cccctcgca gcccaagatt 2640
caccgcgaga ctcttctctg tggggccctt cccagtcaca gccagagccc aggtgaattc 2700
cccctctctc ctgaccacca tgtgtctctt cccccaaccc caagcgctg gaactctag 2760
caggactcta ccttgaggag acccaagact ctggagtttc tgcctcggcc cgtctcttct 2820
gccaacggcg aacggctgga ccttgggtgg tttgtgtccc ccagccatgc ccgagagacc 2880
tcccagacca acagctccag caccagagcg cccagcaacc tggactctgt ccttgcctag 2940
agttagagca ctgttaggga gcggcctgga ctccacccc tgcctccggt ccaggcaggg 3000
ccgctgcccc cgaactgagcg gacgctcctg gacctggatg cagaggggca gagtccagg 3060
agcaacgtgc cgtgtgtcag agcgggaactg aacacacaca ggcctgccc tttatgagatc 3120
cagcaggagt tctggtctta g 3141

```

<210> 28

<211> 1244

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 7474560CB1

<400> 28

```

gtcttattgc caagataatt acaactggag aattggcccc acaggaaaca acattaccg 60
agataaaaca aaatttgatg caaataccctg atgaaggagg cattgttatt gatggatttc 120
caagagatgt tggccaggct ctatcttttg aggaacaaat ctgtacccc gatttggtgg 180
tattcttggc ttgtgttaat cagagactca aagaagattt actgagcgt gcagacagc 240
agggccagac agagacaaat gtaagagcta ccaaggag actaatgaac ttacagcaga 300
atgctgctcc atttggttaa tacttccagg aagaaggggt catcatgac ttgtgtgag 360
acggcgatga ggaatgggtt ttctatgaca tcaagcattg agttgacac aagttatttc 420
caaacaaaga ggctgcagca ggttcaagtg accttgatcc ttcatgata ttggacactg 480

```

WO 01/81555

PCT/US01/12992

```

gagagatcat tgatacagga tctgattatg aagatcaggg tcatgacag ttaactgtat 540
ttggagagga cactatggga ggtttcatgg aagatttgag aaagtgtaaa attattttca 600
taatttggtg tcttggtctt ggcaaaagcca cacagtgtga aagcttggtg gaaaaaatatg 660
gatttacaca tcttcaact ggcaagctcc tgcgtgagga actggcatca gaactctgaa 720
gaagcaatatt gatcagagac attatggaaac gtggagacct ggtgacctca ggaactcgttt 780
tggagctcct gaaggagggc atggttgcca gctcggggga ccccaagggc ttcttgattg 840
acggtatctc tgggggggtg aagcaagggg aagagttcgg acgcaagatt ggagaccctac 900
agtttggtgat ctgtatggac tgcctggcag caaacacct ctcacaagga 960
gcggagcag cctgctctgt gacgacacca ccaagacct agccaagcgc ctgagagct 1020
actacggagc gtccatcccc gtgatcgct actacggagc aabaaacag ctacacaga 1080
taaatgcaga gggaaaccca gaggaagctt ttcttcaact ctgcacagct attgactcta 1140
ttattttctg aaggcaaaaa tgcattgttag aatggaaca gaaaaacatt aaaaagtcca 1200
tccgttaac acaatgtttc aagttaaacc ttttgttcca ccgc 1244

```

<210> 29

<211> 1661

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 7474602C51

<400> 29

```

gcggcgccgg cgagagcgaa agaggaaact gcagaggagg aagctgcgcc gcagcccgag 60
ccggccggga tcccgccgcc ctctgcgcc gcggcgccgc cccggcgccc cctcccagc 120
gcggcccgcc ccgctcctcc gcggcgccgc ctctgcgccat ggccggggag aacggcgaga 180
gcggtctctc ctggaaaaag caagctgaag acatcaagaa gatcttcgag ttcaaaagaa 240
ccctcggaac cggggccttt tccgaagtgg ttttagctga agagaaggca actggcagc 300
tctttgctgt gaagtgatc cctaagaagg cgttgaagg caaggaaaag agcatagaga 360
atgagatagc cgtctcgaga aagattaaag ctgaaaaatc tgttgccctg gaagacattt 420
atgaaagccc aaatcacctg tacttggtca tgcagctggc gtcgggtaga gagctggttg 480
acgggataag ggagaagggg ttttatcacg agaaggatgc cagcaactctg atccggcag 540
tcttgaaag cgtgactct ctccacagaa tgggactcgt cccacagagc ctccagcccg 600
aaatctctct gtaactacgt caagatgagc agtccaaaat aatgatcagt gactttggat 660
tgtcaaaaat ggggggcata ggaatgtga tgtccactgc ctgtggaact ccaggctatg 720
tcgctcctga agtccctgcc cagaacctt acagcaaaag cgttgactgc tggctccatg 780
gagtgattgc ctacatcttg ctctgcgctt accctccttt ttatgatga aatgactcca 840
agctctttga gcagatcttc aaggcggaat atgagtttga ctctccctac tgggatgaca 900
tctccgactc tgcaaaagac ttcatctgga acctgatgga gaaggacccg aataaaagat 960
acacgtgtga gcagggagct cggcacccat ggaatcgtgg tgacacagcc ctcaacaaaa 1020
acatccacga gtccgtcagc gccacagatc ggaaaaactc tgcacagagc aaatggagac 1080
aagcatltaa tgcacggcc gtctgagac atatgagaaa aotacaacto ggcagcagcc 1140
tggacagttc aatgcagct gtttcagaca gccacagttt ggcacagcaa aaagactgtg 1200
cgtctggcac ctccacgct ctgtagtttc atttctctt cgtcgggggt ctccaggatt 1260
ggagccgagc ggagacccag gccacccact gtgacggcag tgcactctgg aagcaagtga 1320
ctggcctgag agtggggcc cgggtctggc gctggggag ggagaccca ggttcggcag 1380
agccggagac cactccagcg agacccccc ttgagtggtg ccccttcctg cataggaact 1440
gaagcccgaa gtttttttat ggccatattt tctactgcac ttctgaagtg ttctattctc 1500
acaaactgta ctgactcgag gggcgttgat tctataggat ctggtgctgt atatacgaat 1560
cttgcaagc tctaactgaa cggaccttct tattctcttc tcttaacacc atcgtttcca 1620
ctcttctcag tctaggtaac cgtctatggt gtgtctttca t 1661

```

<210> 30

<211> 912

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 7475509C51

<400> 30

WO 01/81555

PCT/US01/12992

```

cggcggtgg gctcgggtcc ggcggtgggc tgaggggggg ggttgtctta aaagtctctc 60
cttccccctg tagggggcggc cggcgagttcc cagtgagagc ggaggggtgcc agaggttaggg 120
ggcgagagaaa caaagtttccc ggggcttctct cgggggcgcgc ggtcgggggtc ggcggtttga 180
ccgccccctc cctcgcgaag gcaatggctt ccaaacctct gcgcgcggtc atctctcggc 240
cgccgggtct ggcaaggggc accgtgtgccc agaggatcgc ccagaaactt ggtctccagc 300
atctctctag cggccacttc ttgggggaga acatcaaggg ccagaccgaa gttaggtaga 360
tggcaagga gtaataagag aaagtccttt tgggtccaga ccatgtgac acagcctaa 420
tgatgtccaa gttggagagc aggcctggcc agcactggct ccttgatggt ttctctagga 480
cattaggaac agccgaagcc ctggacaata tctgtgagc ggcctctagt atcagttta 540
atatctcatt tgaacacctt aaagatcgtc tcaagccgcg ttggattcac cctctagcg 600
gaagggctata taadctggac ttcaatccac ctactgtaca tggattgat gacgtcactg 660
gtgaacgctt agtcagcagc gaggatgata aacccgaagc agttgctgcc aggtcaagc 720
agtaacaaga cgtggcaagc ccagtcattg aattatacaa gagccggaga gtgctccacc 780
aatttttccc gaaccggaga cgaacaaaaa tctggcctta cgtttacaca actttctcca 840
acaagatcac acctattcag tccaggaag dattttgacc tgcccatgga gaccggaatg 900
ggcctcactc ct 912

```

<210> 31

<211> 2858

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 7475491CB1

<400> 31

```

ggtagggccc ggaatcactg taaggcgagc tgtgtctggc agactctctc cttccctccc 60
ccttctactc cctcctcggc cagctcaggt tgcagcttct ctgggggaact gctcaacttt 120
ccggagcagg ggaagctgccc ccgtgcctgg gaggagagcgc gcgcacccgc gcccccagga 180
cacggcgtgt gactccgcgc gcggtgcgcc ttggaggaag ggggaaggtcg gaggaggggg 240
cactcgccgc cggggtatga ggaacaaagg acctggctgc ccagtcctct atgactatga 300
acaaatgaa gaactttaag cgcgctttct cctgtcagtc gccccgcact gagaaccttg 360
aagaatcctt ggtgaatbc ccgagcaaat tcaacagact ccaacacccg ccgaatgaga 420
acttgcagct cgttctctct ggcaagagac ccccgccagg gtgcagcacc ttctcccaaa 480
cagaacagcg ggagagccgc ggccagctct cccctggcgt gcagttccag cggcgccaga 540
accagccgcg cttctccatg gaggtgaggg cctctggagc tctgcccgcg ccggtgagcag 600
gatgcacgca caagggtgtg cccagggagg cagctgcctt acagccagac tttagctca 660
gcagaggctc cttctgtccc atggatatcc gcttgcccca ggaattccta cagaagctac 720
agatggagag cccagatctg cccaagccgc tcagcccgcat gtcgccccgc gctcctctgt 780
cagacattgg ctttgggaaa ctggaaacat acgtgaaact ggaacaaact ggagagggca 840
cctatgcacc agtcttcaaa gggcgacgaa aactgacgga gaaccttctg gccctgaag 900
agatccggct ggagacagag gaggagagcg cctgcactgc catccgagag gtgtctctgc 960
tgaagaacct gaagacagcc aatattgtga cctgcactga cctcatccac acagatcgg 1020
ccctacccct gggtttggag taactggaca gtgacctgaa ccagtatctg gacccctgtg 1080
ggacccctcat gacatgcac aactcaaga ttttctggtt ccagctgctc cgggcccctg 1140
cctactgtca ccaacgcagc atcctgcacc gggacctgaa gccccagAAC ctactcatca 1200
ccgagagggg ggagctgaag ctggccgact ttggactggc cagggccagc tcaagtccca 1260
caagagatta ctcaatgag gtggtgaccc tgtggtacag gcccccagc gtgctgtgtg 1320
gatccacaga gtaactccac cccattgata tgtggggcgt gggtgtgact caetacagga 1380
tggccacagg gaggcccttc ttcccgggt cccagtgcaa ggagagctg caetcatct 1440
ttcgcctctc cgggaccccc acagaagaga cgtggcccg cgtgacccgc ttctctgagt 1500
tcgcgaacta cagcttcccc tgcctacctc cgaagccgct catcaacac gcccccaggt 1560
tggatcaaga tggcatccac ctcttgagca cctctgctct gtabgaatcc aagagtgcga 1620
tgtcagcaga ggtgcctct agtcatctct acttccgctc tctggggagc cgtgtgcacc 1680
agcttgaaag cactgcctcc atcttctccc tgaaggagat ccagctccag aaggaccag 1740
gctacccagg ctgggcttc cagcagccag gacaggggaa gaacaggcg ccagcatct 1800
tctgagctac gccaccttg ctggtggcaa ggaacaaag atcacataga gcaacatc 1860
ggtagagatg gacgtgtgtt ggcctctgga ggaactgaga ccaggggtcg acagccagc 1920
tggagagccg cttggcagcc cttctgga cgtgtgtt tctctgtg tcccgtgtg 1980
cttccccagt agcctcacc tgaataccaa cccctcttt acccaagtt gggtggcat 2040
aagctgtctc cctgagagga catgaggggg gggcggtctc cgtacccct cccacccctg 2100
tgtttgggca cctgcgtggg atgcacagc atgacagag caaggcgcca ggtgggac 2160

```


WO 01/81555

PCT/US01/12992

```

tctgcccctg atacaggctc taccctctcc ccccgagacc tgcctagtgc cagtttggta 2220
gtcccccttt ctggcccctt gtagccacac caggtttcat ctttttcccc tctgagagca 2280
agaaagagac tggcatgttc tctgggaccc tggaaaccta ggtaccacac tgtgtgucac 2340
agcttacccc acctggcagg tctcccacag caacagaagg aatagtagtc cccactcttt 2400
ccatccagcc taccctaccc tcatctcccg aacccctctg gcttgaacac tggctgagca 2460
gtaccggcat acgctttgct ggcakgcttg gatcccaagg tctgtccaga ggtggcctgg 2520
gaccgcagct tgaacgctg ccacctcagc cagcccccgc ccagctcctc agtctgaatg 2580
gagttgcctt aaattggcag gtggtacgct actcaatgcc cttggagctg tgacggctc 2640
ctgctgtctc accccttccc gaggtggctc ctgcttaact tatcaltcca ggcctctgat 2700
taaccaggcc tggctagggt cctggggacg gcacccagat atgcagagtc accctgacac 2760
tgggtccagg ctgacctcag ctcccgaggg ctgcacagc ctcccctccc ttccttccca 2820
gcctttgtgg ctctgtccac ctgabcacaa taacagct 2858

```

<210> 32

<211> 2817

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 2192119CB1

<400> 32

```

ttctctctat ctgaaagctg tctctctccc actcccgaca caetgggagc ccccctatcc 60
ccgcccctct ccccccctc ccttgtcccc gcccttcccc gccccaacac aggtttctgc 120
cccccctccc ccaattccag tctgtctccg ccccgacgta ggcctccccc tggcccctgg 180
gccccctccc cgtctgacgc aggcctccgc cctctctccc cccgcccagg ctggggcggc 240
cggagagacc ggagctaaag cgcctgaacc cgcgcggcgg gtggggagca tctggtttct 300
tgcctccggac ccgctccggg actttccgct cagctcctac ccggagcccc cagagggcgg 360
ccttgcctgg ccctgggccc tgcaacggcg ccgcaagaaag gccacaggca gccctgtgtc 420
catctctgtc tatgtgtgta agctctggcg gaagagagag acccaggtag ccaaaagctgc 480
cttcaagcgc ttcaaaactc taagggaacc caaatctctg gcttactatg atgagctgga 540
gaacgaaaaa tgcctccacg tctgacacga ggtgtgacc cgttgggaa tatacctcaa 600
ggcggagctg gaggctggtg gctgaagga gctggagatc tcttgggggc tacacccaga 660
cgtgaagacc ctcaagcttc tggtaacgca ctgcaagctc atccacaaca atgtctgcat 720
ggcggcctgt ttcgtggacc gagctggcga gtggaagctt gggggcctgg actaatgta 780
ttcggcccag ggcaacgggt ggggaacctc ccgcaagggg acccccagag ttgagcagta 840
tgaccccccg gagttggctg acagaagtg gtaggtgttc agagagaagt ggtcagcaga 900
catgtggcgc ttgggtgtgc tcatthtggg agtcttcaat gggccctac ctccgggagc 960
agcctctacg aacccctggg agatccccaa aacgctgtgt cccctattct gtgagctggt 1020
gggaacaaac cccaaggctg gtcccaaccc agcccgcttc ctgcagacct gccgggcacc 1080
tggttgcttc atgagcaacc gctttgtaga aaccaacctc ttctggagg agattcagat 1140
caaaagagca gccgagaacc aaaaattctt ccaggagctg agcaaggacc tggagcattt 1200
ccttgaggtat ttctgtggcg acaaggctgt gaccagagct ctgacagcct tggagttcgg 1260
caatgtctgg gcgctgtgct tcaagcctct ctccaaggtg ggaaggttcc tgagcgttga 1320
ggagtatcag cagaagatca tccctgtggt ggtcaagatg ttctctacca ctgacggggc 1380
catgcgcatc cgcctctctg agcagatgga gcagttctac cagtaccttg accgacccac 1440
agtcacacac cagatctctc cccacgtcgt acatggcttc ctggacacca accctgcaat 1500
ccggagagac acggtcaagt ccatgtctgt cctggcccca aagctgaacg aggcacaact 1560
caatgtggag ctgataagac actttgcaag gctacaggcc aaggatgaac agggcccaat 1620
ccgtgtcaac accacagtct ccttgggcaa aatcggtctc taacctcagt ctgacccacg 1680
acacagggct cttaacctct ccttaagcgc agccactagg gaacctttg caacgtcccg 1740
ggttgcgggt gtcttgggct ttgtgtccac ccacaaacct tactcaatga acgactgtgc 1800
ccagaagatc ctgctgtgct tctgtggtct cactgtagat cctgagaant ccgtgcgaga 1860
ccaggccttc aaggccattc ggagcttctt gtccaaattg gactctgtgt cggaggagcc 1920
gacccagctg gaggaagtgg agaaagtgtt ccatgcagcc tccagccctg gcatgggagg 1980
agccagagct agctgagcag gctgggctgt gaccggggtc tctctctcca cctccaaagt 2040
gactcgttgg caacaaacca ctgcccacac agaaacccac attcccccaa gaacacagc 2100
tgagggccac tggagagcgc agggagagga caaggacaca gcagagagca gcagcactgc 2160
tgacagatgg gacgaggaag actggggcag cctggagcag gaggccagat ctgtgtgctg 2220
ccagagggac gactggagca ccgggggcca agtgagcgtt gctagtgaag tcaagcaact 2280
cgaccacaaa tctccaaat cccagagtc cgaactggag agctgggag ctgagggctc 2340
ctgggaacag gctgtggagg agccaagctc ccaggagcca cctcctgacg gtacagggct 2400

```

WO 01/81555

PCT/US01/12992

```

ggccagcgag tataactggg gtggccaga gtccagcgac aaggcgacc ccttcgtac 2460
cctgtctgca cgtcccgagc cccagccgag gccagactct tggggtgag acaactggga 2520
gggcccgcag actgacagtc gacaggtcaa ggtgagctg gcccggaaga agcgcgagga 2580
ggcgccgcgg gagatggagg ccaaacgcgc cgagaggaag gtggccaaag gccccatgaa 2640
gctggggagcc cggaaagctgg actgaacgct ggcggtggcc ctcccggtc gcygagagcc 2700
cgccccacag atgtatttat tgtacaaacc atgtgagccc ggccggccca gccaggccat 2760
ctcagctga cataatcaga gccacaataa attctatttc acaaaaaaa aaaaaaa 2820

<210> 33
<211> 5305
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7474456CB1

<400> 33
atggctgggg gccctggggc ccccgggcgc ggcggggagc agcctccgga gagctacccg 60
caacgacagg accacgagct acaggccctg gaggccatct acggyccgga ctcccaagac 120
ctggcgccgg agccttgcgg accggtcaaa gagcccccctg aatcaattt agttctgac 180
ctccaaaggcc taactgtgtg agaagtatat gtaaaagtgg atttgaggg taastgccca 240
cctacactatc cagatgtagt tcttgaataa gagttaaaaa abgccaaagg tctatcaaat 300
gaagtgctca atttgttaaa atctcgccct gaagaactgg ccaagaacaa ctgtggggag 360
gtatgagatc ttgaactggc ttaccacgtg cagtcatttc tcagcgagca taacagagcc 420
ctccccaagt cttttcattg agaatgctg gaaaggcggg ctcaagggga gcaacagagg 480
ctgttggagc cccaaagcgg aagaagaaga gacgagggac aacaacgtga aatcctgcat 540
gaggttcaga gaaggaaaga agagataaaa gaagagaaaa aaaggaaaga aatgggctaa 600
caggacgctt tggaaattgc tagtttgcga aaccagatc atacctctaa gaaggcccca 660
gggggacaca gaacggctgc cattctacat ggaggctctc ctgactttgt aggaatttgt 720
aaacatcggg aaactcctc aggaaggtct agcgagagac gtcagctatc tgtatgtaat 780
agtgaagctt ctctggctc ttgtgaatt ctgacttcca atatggggag tctgattcag 840
ctcaaggctc caaaagggaa atgtattggc agtgaagagc aacttggaaa attagtctac 900
aatgcttttg aaacggccac tgggtgcttt gtctgtgtgt atgagtgggt ccttcagttg 960
cggaananaa tgggtccctt ccttaccagt caagaaaaag agaagattga taagtgcata 1020
aagcagatcc aaggaaacaga aacagaatcc aactccttg taasattgag ccatccaat 1080
gtagtacgtc accttgcaat gaatctcaaa gagcaagccg actccattgt ggtggacatt 1140
ttagtggagc acattagttg ggtctctctt gctgcacacc tgagcgaact agggcccatc 1200
cctgtgctc agcttcgagc gtacacagct cagctcctgt caggccttga ttatctgac 1260
agcaattctg tgggtcctaa ggtcctgagt gcatctaatg tcttgggtga tgcagaaggg 1320
accgtcaaga ttacggacta tagcatttct aagcgctctg cagacatttg caaggaggat 1380
gtgtttgagc aaacccgagt tctgtttagt gacaatgctc tgccttataa aaaggggag 1440
aaaggagatg ttggcgctc tggcctctg ctgctgtccc tcagccaaag acaggaaagt 1500
ggagagtgac ctgtgacat ccttagtgac ttaccagctg actttcaaga ttttctaaag 1560
tptgtgtgct tgggtgcaa ggaagatgg agtccccgc agttgttga acacagcttt 1620
ataakccc agctcaaaat gctctagtg gacaaagatc ctgaagatc tgaagggcaa 1680
gatttatatt agactgttat tcttagcaac cggctaccca gtgtgactt ctttagtgag 1740
acacagagac agtttccccc atactctatt gagtgtgaag aattacaact tcttggtaaa 1800
ggagcttttg gagtgtcat caaggtgcag aacaagtgtg acggctgctg ctacgcagtg 1860
aagcgctacc ccatcaaccc ggccagccgg cagltccgca ggaatcaagg cgaagtgaac 1920
ctgctgtcac ggtgacacca tgagaacatt gtgcgctact acaacgctg gatcagagcg 1980
cacagagcgc cggcgggacc ggggacgcg ccccggact cggggccctt ggccaaagga 2040
gaccagagctg caacggggca gccggcgagc gacacagagc gctgggacag cgttagagcc 2100
gccgctgcgc caaccatctc cagcagctcg gtggagtga gcaacttcgg cgaagcctcg 2160
gocagtgcce gtttccccgc caccggcccg ggctccagcg atcagagga cgaacgagag 2220
gacgagcagc gtggcgctct ctcccagtc ttcctgctg cttcagatc tgaagtgtat 2280
attatcttgg acaatgaaga tgagaacagt aaaaagtcaga atcaggatga agattgcaat 2340
gaaagaatg gctgcaatga aagtgggcca tcaagtgaaga ctgaggtgtg gcaactaccta 2400
taactcaga tgggtatctg tgagaagagc actttacagc acacccttga ccagggactg 2460
tatcagagaa cgttcagact ctggaggttt ttccagagga ttctgtatg attagctctc 2520
atcctatgaa aaggaatgat tcaccgggat ttgaagcctg tcaacatttt ttgtgattct 2580
gatgacatg tgaatagag tggtttttgt ttggcgccag accatctaga cttttctgtc 2640
gacgcaaac aagacgata gacagagagc ttgattaagt cagaaccttc aggtcaccta 2700

```

WO 01/81555

PCT/US01/12992

```

actgggagtg ttggcactgc tctctatgta agcccoagagg tcccaaggag caccsaatct 2750
gcatacaaco agaaagtggg tctcttcaag ctgggaatta tctcttttga gatgtccat 2820
caccoccatgg tcaaggtctc agaaaggatc ttgttttcca accaactcag agatccact 2880
tcgctcaagt ttccagaga ctttgacgat ggagagcatg caaagcagaa atcagtcac 2940
tcctggctgt tgaaccacga tccagcaaaa cggccacag ccacagaact gctcagagt 3000
gagctgtctc cccaccccca gatggaggag tcaagcttgc atgaagtgt gcaacacag 3060
ctgaccaagc tggatgggaa ggcctaccgc accatgatgg ccagatctt ctgcagctc 3120
abctccctgc ccatcgatta cacttatgac agcgaatcac tgaaggagaa ctctccaac 3180
cgtacagccc agatgcagca gcatgtgtgt gaaacacaca tccgctctt taaagacat 3240
ggagctgttc agltgtgtac tccactactg ctcccccaa acagacaaat atatgacac 3300
aaagagctgc cctattcat ggcacacagc gggatgctgg tgnatgcttc ttttgactg 3360
cggatccctt ttgcagata tctggcaga aataatata tgaatttaa acgactactg 3420
atagacagtg tttccaggcc ggcgaagtta gatcgattc atcccaaga actctggag 3480
tgtgaatttg atattgtcac ttctaccacc aacagcttc tgcacactgc tgaattatc 3540
tacctatct atgaatcat ccagagttt ccagcacttc aggaagaaa ttacagtat 3600
tatttgaacc ataccatgtt attgaagca atactcttac actgtggg atccagagat 3660
aaactcagtc aagttacat tattctgtat gatgctgtga cagagaagc gacgagaga 3720
gaagtggag ctaaatttg taatctgtct ttgtcttcta alagtctgt tgcactcac 3780
aagtttatg aacgaaggg agatttgcaa gatottatgc caacataaa ttcattata 3840
aaacagaaea cagttatgc acagttggg aagtatggct taaaagacct agaggaggt 3900
gttggactgt tgaagaact cggatcaag ttacaggtct tgaatattt gggctggt 3960
tacaaggctc agcagcaaa tggaaatcac tccagcttgc caaagagag 4020
caaaaggctc tacctgaat cctgcagct ggagcgagat atgaactgt gattccccc 4080
tttagaggcg caaaagctct gggcgagtt cccactgcca ttggggcag catagctata 4140
gacagatct ctgctgctgt cctcaacatg gaggaaatct ttacaataag ctttgtgac 4200
ctcctggttg taagtgttg ccagatgtct atgtccagg ccactcaact aaccagaaa 4260
ctctggagc caggcatcac agcagaatc atgtacgact ggtcacagtc ccaagagaa 4320
ttacaagagt actgcagca tcatgaatc acctatgttg ccttgrctc ggaataaga 4380
ggagccatg tcaaggttaa gctcttcag aaggaaagcc agcagagaa cgtgtgtgt 4440
ggactgaac ttgtggaca tgtactgag aaactgagga ctaagtcac tgaatgaag 4500
aatggcagag aagcttccga taactctgca gtgcacatc tgaagggtc atttctaat 4560
gcttcagggt ttttgaat ccatggagca acagtgttc ccattgtgag tgtgctagc 4620
cgggagaagc tptcagccg cactaggag cgtatgaaa ctcaggtaca aatcagctt 4680
cagactccc ttgcacact acatcagaa agcagtgaa ttgaattct ggttggtat 4740
ctcccaag aacaaatatt acagtttta tcatcagat gggatgtgga tgaacagga 4800
tttaacaaa ctgtgaagca gctgcttca cgcctgccaa agcaagata cctcaatta 4860
gtctgtgat aaatttata catcaagta gaaaaaagg tctctgtgt atttctgtac 4920
agctatagag atgactacta cagatcttta ttttaacct aaagaactgt cgttaactc 4980
atcaaacag adagaggtt atactggaat aatggatgt tgtacattc tcataattt 5040
aaattcaatt ctaagaagag gctgggtgca gtggctcaca cctttaatc cagcatttg 5100
ggaagccaag gaaggaagac tgcctgaac caggagtttg agaccagct gacnacana 5160
gcagacccc atctctata aaactaaaa aattagtttg gcatgggtg acatgctgt 5220
agtcacagct actccagag gtgagatgga tcatctgag ctcagaggt tgaggctga 5280
gttaactgga cggcggggg atcca 5305

```

<210> 34
 <211> 3269
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> incyte ID No: 1834248CB1

```

<400> 34
cgggagcccg gagaagatgt cttcgggac ggtgctggcc cggggaacy atcggaactc 60
ggacacgcac ggcacacttg gcagtgcccg ctctcggac aaaggcccg cctggctcag 120
cgtctcactg ggtgcccggt gcgggaactc catcgctccc tgcctcgagg agcagcccca 180
cgtgggcaac taccgctgc tgaagacat tgggaaggcc aactttgca aagtcagct 240
ggtcgggac atctcactg gtcgggaagt tgcacacag atkatgaca aaactcagct 300
gaatccagc agctcgagc agctgttccg gaagtcgc atcatgagg gctcaaaa 360
cccacactc gtgaagctct ttgagtgat tgagactgag aagacgctg acctggtat 420
gggtacgca agtctggag aatgtttga ctacctctg tgcagtgcc gactgaagga 480

```

WO 01/81555

PCT/US01/12992

```

gaaggaaagt cagcccaagt tccgacagat tgttccggct gtgcactatt gtcaccagaa 540
aaatattgta cacagggaac tgaaggctga gaaccccttg ctggatgccc aggcacacat 600
caagattgct gactttggct tcagcaacga gttcaagctg ggaatcgaagc tggacacgtt 660
ctggcgggagc ccccatatg cgcgcccgga gctgtttcag ggcacgaagt acgacgggco 720
ggaggtggac atctggagcc tgggagtcac cctgtacacc ctgctcagcg gctccctgoc 780
cttcagcggg cacacacctca aggaagctgc ggaagcagta ctacagggga agtacccggg 840
cccttuctac atgtcaacga actgtgagag catctgtcgg agatttttg tctgtacac 900
agctaaacgc tgtactctcg accaatctat gaaggaacaa tgggtcaaca tgggtctatga 960
ggctcagggg ttgaagccat acacagagcc cgaaggagac ttcggggaca caagagaaat 1020
tgaggtgatg gtgggtatgg gctacacacc ggaagaaalc aaagagtcct tgcacagcca 1080
gaagtacaac gaagtgaacc ccacctacct cctgctgggc aggaagactg aggagggtag 1140
ggaccggggc gccccagggc tggccctggc accggtgccc gcgccacagc acaccaccaa 1200
cggaacaaat tccagcaaa gacccagcca cagcaaaagg cagcggagtt cctcttccaa 1260
ctaccaddgc cagcagcagg atagcattt ctgtggccca tccctgcac cctgcaccc 1320
caaacgcagc ccgacagaca cgggggaggg ggaagctgaag gaggagcggc tgcagggcgg 1380
gaaggcgagc tgcagcagcc cggggagtg ggtcagggg ctgccccctt ccagcccat 1440
ggctcagcgc gccacacacc ccaacaagcc agagatccca gagcggcgga aggacagcac 1500
gagcaccacc acaacctcc ctcctagcat gatgacccgc agaaacacat acgtttgcac 1560
agaagccccc ggggctgagc gccctgccct gttgccaaat gggaagaaa acagctcagg 1620
cacccacagg gtgcacctg cctcccccct cagtcacagc ctggcaccct catcaggga 1680
ggggagccgc ctggaacagg gtttccacat ccgcagcacc ttcacatggg gccaggtccg 1740
ggaccggcgg gcagggggct ggggtggtgg ggggtgtgag aatgggccc ctgctctcc 1800
cacactggcc catgaggctg caccctgccc cgcggggcgg ccccgcccca ccaacacct 1860
cttcacacag ctgacctcca aactgacccc aagggttcga gacgaacctg agagaactcg 1920
gggacactgg gtcacaaagt gccatctacc ttgggataaa accgaaaccc ccccccggct 1980
gctccgattc cctggagtg tgaagctgac cagctggccc cctctgagg cctctgatgc 2040
agctctggcc caggccacag cagccggccc ctgctgctgc cgcacagcac agccgttctt 2100
gttggcctgc ctgacgggg gtgcggggcg gcccgagccc ctgtccact tcgaagtga 2160
ggctctgccc ctgccccggc caggcttgcc gggagttctc tcccgccgct tggcgggccc 2220
cgcctctggc tcccgacccc tggtaacccc catctccaaa gacctcagag cctgagccac 2280
caggttccca gggcccttac tcttctcttc cctgtctgcc ttcacttcta caggagggga 2340
agggggccagg gaagggaatt cctctttatc atcactctag ttccctgaa ttatatttgy 2400
gggcaagatc tgtccctctt gctgttctct ggggcggctc agcacagagc aggatgag 2460
gggtcagcg ggggagctg gcaactctct ggggcttcca gccagctctg tctctctcg 2520
ccctaccag agggcactg aggaacttt ggggacaggc caggggcagg gagggaact 2580
gaagaaatct tccattctct ccaacagctc aaacttaggc ctggggcagg ggcagggaga 2640
cctgctgagc ctacagactg gagaatctgg gggactggga gtgggggtca gagaggcaga 2700
ttcttctccc tccggtcccc tcaagctcaa accccacatt cctgcccag gctgcccgg 2760
ggacttttgt acaaatctct gtaaatcccc caaacctccc cctctgcaaa ggtctcttga 2820
ggagctggcg ctgtaccta cggtttttaa gttattacac ccagacccct cctctgtcag 2880
ccccctcac tgcagcctgt tgcacaataa atttaggaga gtccccccct ccccaatgct 2940
gaccttagga ttttctctcc ctgcccctac ctgcacaaat gttaaaagag aggcgttgga 3000
atccaggcag tggtttttcc tttcggagcc tcgggtttct catctgcaga atgggagcgg 3060
tgggggtggg aaggtaagga tggctgtgga agaaggcagg atggaaactg gctcatccc 3120
cagggcccca gtccctatat cgggcccccc attcatccac tcaactcccc agccacatg 3180
ttacactgga ccttaagcca ctcttctctc cagtagtaaa ttatttcaat aaacatcat 3240
tgaccatga aaaaaaaaaa aaaaaaaaaa 3269

```

<210> 35

<211> 3017

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 71584520CB1

<400> 35

```

gggcagcgg ctgacgatgc tggaggttgc ctacgcgaag cggctgcato tggcgcggcg 60
tctgcccggc gtgctcggag cggatcttgc ccgcggctcc cggagccctc gggcgcggcg 120
tggcccccgc gataccttcc tccctgtgac caaccgggccc tgcaggttag agccttgcaa 180
tgcctatttg gtgtgtgact ctgggcgaca agaagaacta taaccagcca tcggagggtg 240
ctgacagata tgattttgga caggctatca agactgagga gttttgtgaa atcttccggg 300

```

WO 01/81555

PCT/US01/12992

```

ccaaggacaa gacgacaggg aagctgcaca cctgcacgaa gttccagag cyggacggcc 360
gcaaggtgcg gaaagctgcc aagaacgaga taggcaccc caagatgggt aagcatccca 420
acatccctaca gctgggtggt gtgtctgtga cccgcaagga gtactttatc ttcttgagac 480
tggccacggg gaggaggggt ttlgactgga tctgggacca gggctactac tggagcgag 540
acacaaagca cgtggtacgg caagtccctgg aggcctgtgg ctatttgca cactcaaga 600
tcgtgcacag gaattctcaag ctggagaccc tggtttacta caaccggctg aagaacctga 660
agattgtcat cagtgactic catctgacta agctagaaaa tggctcactc aaggagcct 720
gtggagacccc cgaatctctg gccccagagg tggtagggctg gaaaggglat ggaacgctg 780
tggactgctg gccattgga gtcatctgt acatctgct ttcaggcaac ccactctct 840
atgagggaggt ggaagagat gattatgaga accatgataa gaaatctctt cgaagatcc 900
tggctgtgtg ctatgagttt gactctcoat atthggatga tatttgcag gacgcaag 960
acctggtcac aaggctgatg gaggctggagc aagaccagcg gatcactgca gaagagacca 1020
tctcccatga gtgatttct ggcaatgctg ctctgataa gaacatcaag gatggtgtct 1080
gtgcccagat tgaanaaac ttggccaggg ccaagtggaa gaagctgtct caggtgacca 1140
ccctaatgaa acggtccggg gaacnagagc agtccagcac ggctgctgct cagtggcct 1200
cagccacaga acgtgccacc cccggggctg cagggtgggg caccagtgtc gctgaggtg 1260
gagctacctc agccctcgag ggtgatgctg ctgctgtctg aaagagtgat aatgtggccc 1320
cccgagacgg tagtgccacc ccagccacag atgggaagtgc caccccagcc actgatgcca 1380
gtgtaccccc agccaccgat ggaagcatca ctccagccac tgaaggaggt gtcaccccag 1440
cactgagag gaggctact ccagcactg atggggagag caccagagcc aagaagaga 1500
gaactgtgca caccacccaa agcagtgcca tgggtggcac caaggcagct gccacccctg 1560
aggcgtctat gggccagcgg gacagccag cccagagggg cggccagggc caggctccc 1620
cccttagtaa aggggaagag gctgctggtt atgcctagga gctcxaag gaggagacca 1680
gctgagttag cagcctggtg agggcgggca gggatgggg agagaggtgg gaggatggat 1740
gaggggcttc tcactgtaca tagagtcact ggcctgatgc cctgcctccc ccattccccc 1800
acatcccatg ggggcataac taggggtcac gggagagcag tctctctccc ggtgtgtatg 1860
tgtgtgagtg gtgggcaggg cagtggcagg gcgggdonca gccctgcat ggttctctg 1920
tggcttttct gctctttgct agcttcaaca gttctgttct ctgtgtggat gctgctctag 1980
ggatactcag ggggtctctg ctctctctcc ccttcccttc tlgcctcacc attcccttag 2040
gcaggccctg caggctccac actctcccag gccctaaact tggggcgctt tgcctcaga 2100
gctgtctctc cagcgaggcc ctgtcagcgg tcttaggtct ctgcacatga aggtgtgtg 2160
ctgtgtgtg tgggtgtctc taggagcaga taccagctgg tatagaggat gcaagaaggt 2220
agggcagat gtttaagtc aacttgcca catggctagg gatcctgctc actagctgtg 2280
gaggtctctc ggggtggaga gaatgagtag gaggcgaga gcttccctt ttgtctctc 2340
taagaccctg ttatttgtgt tatttctctg ctttccaggt cctgcaagt gctgctctg 2400
acactgaacc tcaatgagct ctaagggaag gagggaacaa ttaggacgtg gcaatgagac 2460
ctggcaggcg agagtacaag cccagcacc agtctccag ccttactgg tcttacctt 2520
gggcacacaa gggggggctg atacctctt gcttctcta gatgccacc tctacacct 2580
tcagccacaa agtctctctc accctagggg gcttctgca tggcaatac tcaatctg 2640
atttgaggt ttgccttta caggggcaga ttttctgct agttcaacaa tgaatgag 2700
aggaactccc tcttctaca gctcacttct atcagaggcc cagggtctc agagccacat 2760
tgggtgtctt ttctgggat gagggaagtag ggttaacct cccagtctc tggggaggg 2820
tctgacaggg tgcctttgt cagaccctac cagagcctgg ataggcagcc acattggctc 2880
tcgccccttg tggccctcc gtgtgtgtcc tgcctctct cctgcatgcc tptgggctg 2940
ctctgtgtg tgaaggtcgg tgggttaact yltgtctctc tgaacctggc aaataacat 3000
caccctgcaa agccaaa 3017

```

<210> 36
 <211> 2168
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Inocyte ID No: 7475538CB1

```

<400> 36
gctgcatcgt ttccgtgggc cctgcttgcg ggcacgctct cggcgcatgc gttttttatg 60
ggggatcaag cttgctgtcg ctgacagcg gaggctagg aaaaagcaca gggggcccg 120
gagctgtcac cctgctctcg agcagcttc cgttctctg gtgactggtg ctacaggaac 180
cgcccaggtg gtacagctgc gcctgttgc taggaacacg cgtacacact cagatcagcg 240
tgggttgagg gagaagtggg gtttggaggt tcaaggcgac agggggcacc gccccagact 300
cagcgggggt ggaacagtag tgcattctgg gccgcatcgg ggaaggcgcc cagggaatcg 360

```

WO 01/81555

PCT/US01/12992

```

tctcaaggc caagcagctg gagactggcg agatagttgc cctcaagaag gtggccctaa 420
ggcggttggg agacggcttc cctaacagg cctgcgggga gattaaaggt ctgcaggaga 480
tggaggacaa tgaatgatgt gtacaactga aggtctgttt cccacacggt gggaggtttg 540
tgctggcctt tgagttcatg ctgtcggatc tggccagagt ggtgcgccat gccagagggc 600
cactagccca ggcacagggtc aagagctaac tgcagatgct gctcaagggt gtcgctttct 660
gcatgcctaa caacattgta catcggggac tgaacctgc caacctgctc atcagcgctc 720
caggccagct caagatagcg gaatttggcc tggctcgagt ctlttcccca gacggcagcc 780
gctctacac acaacagggt gccaccaggt ggtaccgggc ccccgagctc ctgtatgggt 840
ccgcgcagta tgaccagggt gtcgatctgt ggtctatgga ctgcacatg gggagctat 900
tgatgggtc ccccttttc ccgggcaaga aggatattga acagctttgc tatgtgttc 960
gcacttggg caaccacac cctcaagtct ggcggagct cactgagctg ccggactaca 1020
acaagatctc ctttaaggag caggtgccca tgccttgga ggaagtgtg cctgagctc 1080
ctcccgagc attgatctg ctgggtcaat tcttctcta cctctctac cagcgcatc 1140
cagcttccaa ggtctctct catcagtaet tcttccagc tccctgcat gccatccat 1200
ctgagctgac gattctctag cgtctagggt gacctgcac caaggcccat ccaggccccc 1260
ccacatcca tgacttccac gtggaccggc ctcttgagga gtcgctgtt aaaccagagc 1320
tgatttggcc ctctatctcg gagggtgtg agtttgccc tggctccgtc tgcctgctcc 1380
tcaggaccac tcagtccacc tgttctctg ccactgcctt ggttccacc tccagggctc 1440
ccccatggc acagtgggccc cacaccacac cctgcacctt agcctctgc aggttgggc 1500
tcaggccaga ggtcatgttc ccagccaaga gtatgagAAC atccagtcga gccagggaga 1560
ttcatggcct gtgctcgggt agccttaact tctgtatgct actgaggtac ccacaggaa 1620
agtgagctct gtcgcagtc aagccttgca tatgagagat gacgatgctt gcttgggtc 1680
tgcctcccg agtgcctgct cctggtcaag gagaagtgca gagagtaagg gtccttatg 1740
tgggaaactc aagtggaggt agatttgggt ctggttttat tctcagagcc attaaacct 1800
agttaagtat gtgagatata gattctaaaa acctcaggtg gctctgcctt atgtctgtc 1860
ctccttcatt tcttcaagg gaattggcta aggtggcatt gctctatgc tctcgtttt 1920
ggggtcaggt gggaggtagc accaggcata gccactttt cctcaggga ctcctgtgt 1980
cttcacatca ctgagcactc atttagaagt gaggagagaa gaagtctagg cccagggatg 2040
gctcagctg gggatccagc agggagccct ctgcacatga ggtgtgttta ccaacatcta 2100
ctcctcagg atgagcgtga gccagaagca gctgtgtatt taaggaaaca agcgttctg 2160
gaattaat 2168

```

【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
1 November 2001 (01.11.2001)

PCT

(10) International Publication Number
WO 01/81555 A3(51) International Patent Classification: C12N 15/54
912, C12Q 1/68, A61K 38/45, A01K 67/027, G01N
33/53, 33/68

(21) International Application Number: PCT/US01/12992

(22) International Filing Date: 20 April 2001 (20.04.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/199,021 20 April 2000 (20.04.2000) US
60/200,226 28 April 2000 (28.04.2000) US
60/202,339 5 May 2000 (05.05.2000) US
60/203,505 11 May 2000 (11.05.2000) US
60/205,564 18 May 2000 (18.05.2000) US
60/207,739 26 May 2000 (26.05.2000) US
60/208,795 1 June 2000 (01.06.2000) US

(63) Related by continuation (CON) or continuation-in-part (CIP) to earlier applications:

US 60/200,226 (CIP)
Filed on 28 April 2000 (28.04.2000)
US 60/199,021 (CIP)
Filed on 20 April 2000 (20.04.2000)
US 60/202,339 (CIP)
Filed on 5 May 2000 (05.05.2000)
US 60/203,505 (CIP)
Filed on 11 May 2000 (11.05.2000)
US 60/205,564 (CIP)
Filed on 18 May 2000 (18.05.2000)
US 60/207,739 (CIP)
Filed on 26 May 2000 (26.05.2000)
US 60/208,795 (CIP)
Filed on 1 June 2000 (01.06.2000)

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

(72) Inventors; and

(75) Inventors/Applicants (for US only): YUE, Henry [US/US]; 826 Lois Avenue, Sunnyvale, CA 94087 (US); GANDHI, Aruna, R. [US/US]; 837 Roble Avenue, #1, Menlo Park, CA 94025 (US); TRIBOULET, Catherine, M. [FR/US]; 1121 Tennessee Street, #5, San Francisco, CA 94107 (US); KEARNEY, Liam [IE/US]; 50 Woodside Avenue, San Francisco, CA 94127 (US); GRIFFIN, Jennifer,

A. [US/US]; 181 Irene Court #11, Belmont, CA 94002 (US); NGUYEN, Daniel, B. [US/US]; 140 Ridgewood Drive, San Jose, CA 95118 (US); BANDMAN, Olga [US/US]; 366 Alma Avenue, Mountain View, CA 94043 (US); LU, Dyang, Aina, M. [US/US]; 243 Cay Drive, San Jose, CA 95123 (US); LAL, Prati [IN/US]; P.O. Box 5142, Santa Clara, CA 95056 (US); BURFORD, Neil [GB/US]; 105 Wildwood Circle, Durham, CT 06422 (US); KILIAN, Farrah, A. [IN/US]; 433 Locust Avenue, #221, Mountain View, CA 94040 (US); WALIA, Narinder, K. [US/US]; 890 Davis Street #205, San Leandro, CA 94577 (US); YAO, Manique, G. [US/US]; 111 Frederick Court, Mountain View, CA 94043 (US); PATTERSON, Chandra [US/US]; 490 Sherwood Way #1, Menlo Park, CA 94025 (US); BURRILL, John, D. [US/US]; 650 Emerald Hill Road, Redwood City, CA 94061 (US); MARCUS, Gregory, A. [US/US]; 1714 Comstock Drive, Redwood City, CA 94061 (US); ZINGLER, Kurt, A. [US/US]; 723 Ashbury Street #C, San Francisco, CA 94117 (US); RECIPON, Shirley, A. [US/US]; 85 Fortuna Avenue, San Francisco, CA 95115-3873 (US); LU, Yan [CN/US]; 3885 Cornia Way, Palo Alto, CA 94303 (US); POLICKY, Jennifer, L. [US/US]; 1511 Jarvis Court, San Jose, CA 95118 (US); THORNTON, Michael [US/US]; 9 Medway Road, Woodside, CA 94062-2612 (US); YANG, Y. Tom [US/US]; 4280 Ramewick Court, San Jose, CA 95116 (US); HAFALIA, April [US/US]; 2327 Calle de Primavera, Santa Clara, CA 95054 (US); ELLIOTT, Vicki, S. [US/US]; 3770 Polton Place Way, San Jose, CA 95121 (US); BAUGHIN, Mariak, R. [US/US]; 14244 Santiago Road, San Leandro, CA 94577 (US); WALSH, Roderick, T. [IE/GB]; 8 Boundary Court, St. Lawrence Road, Canterbury Kent CT1 3EZ (GB); RAMKUMAR, Jayalaxmi [IN/US]; 34359 Maybird Circle, Fremont, CA 94555 (US); BOROWSKY, Mark, L. [US/US]; 122 Orchard Avenue, Redwood City, CA 94061 (US); AE-YOUNG, Junice [US/US]; 233 Golden Eagle Lane, Brisbane, CA 94005 (US); HILLMAN, Jennifer, L. [US/US]; 250 Monroe Drive, #17, Mountain View, CA 94040 (US); GURURAJAN, Rajagopal [US/US]; 5591 Dent Avenue, San Jose, CA 95118 (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, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, GR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR,

[Continued on next page]

WO 01/81555 A3

(54) Title: HUMAN KINASES

(57) Abstract: The invention provides 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 01/81555 A3



LS, LE, LC, LY, MA, MD, MG, MK, MN, MW, MX, MY,
NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM,
TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

Published:

- with international search report
- before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments

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

(88) **Date of publication of the international search report:**
27 June 2002

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

【国際調査報告】

INTERNATIONAL SEARCH REPORT		International Application No. PCT/US 01/12992
A. CLASSIFICATION OF SUBJECT MATTER IPC 7 C12N15/54 C12N9/12 C12Q1/68 A61K38/45 A01K67/027 G01N33/53 G01N33/68		
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 C12Q A61K G01N		
Documentation searched other than minimum documentation to the extent that such document(s) are included in the fields searched		
Electronic data base consulted during the international search (name of data base and, where practical, search terms used)		
D. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	DATABASE EMBL [Online] Entry/Acc. no. AW166113, 13 November 1999 (1999-11-13) STRAUSBERG, R.: "xf45d06.x1 NCI CGAP Brn50 Homo sapiens cDNA clone IMAGE:2521083 3' similar to TR:088904 088904 HOMEODOMAIN-INTERACTING PROTEIN KINASE 1. ; mRNA sequence." XP002186883 the whole document --- -/-	1-19, 21, 22, 24-45, 62
☒ Further documents are listed in the continuation of box C. ☒ 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 on publication or of date the international filing date. "L" document which may show obstacles on priority claims or which is cited to establish the publication date of another claim or other procedural matters (e.g. prior art). "O" document relating to an oral disclosure, use, exhibition or other means. "P" documents published prior to the international filing date but later than the priority date claimed. "T" later document published after the international filing date as priority date and not in conflict with the application but which is understood 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(s) is taken alone. "Y" document of particular relevance: the claimed invention cannot be considered to involve an inventive step when the document(s) 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 latest convention of the international search		Date of mailing of the international search report
9 January 2002		25 04 2002
Name and mailing address of the ISA European Patent Office, P.O. Box 5516, 6500 LA Haye, NL-6500 HA The Hague, The Netherlands Tel: (+31 70) 340 2540, Telex: 31 651 eppo nl, Fax: (+31 70) 340 0316		Authorized officer Smaitt, R

Form PCT/JP/2010 (second sheet) July 2002

INTERNATIONAL SEARCH REPORT

International Application No.
PCT/US 01/12992

G.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT		
Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	ISHIKAWA K-I ET AL: "PREDICTION OF THE CODING SEQUENCES OF UNIDENTIFIED HUMAN GENES. X. THE COMPLETE SEQUENCES OF 100 NEW CDNA CLONES FROM BRAIN WHICH CAN CODE FOR LARGE PROTEINS IN VITRO" DNA RESEARCH, UNIVERSAL ACADEMY PRESS, JP, vol. 5, no. 3, 30 June 1998 (1998-06-30), pages 169-176, XP000980105 ISSN: 1340-2838 see KIAA0630 ---	1-19,21, 22, 24-45,62
X	WO 99 57132 A (GENETICS INST) 11 November 1999 (1999-11-11) see passages relating to Seq.ID's 31 and 32, and clone col55_12. ---	11,12
X	DATABASE EMBL [Online] Entry/Acc.no. AC027074, 27 March 2000 (2000-03-27) BIRREN, B. ET AL.: "Homo sapiens chromosome 1 clone RP11-214B13 map 1, WORKING DRAFT SEQUENCE, 25 unordered pieces" XP002186884 see nt. 27175-27943 and nt. 114343-115421 ---	11,12
A	YOUNG HO KIM ET AL: "HOMEODOMAIN-INTERACTING PROTEIN KINASES, A NOVEL FAMILY OF CO-REPRESSORS FOR HOMEODOMAIN TRANSCRIPTION FACTORS" JOURNAL OF BIOLOGICAL CHEMISTRY, AMERICAN SOCIETY OF BIOLOGICAL CHEMISTS, BALTIMORE, MD, US, vol. 273, no. 40, 2 October 1998 (1998-10-02), pages 25875-25879, XP002943633 ISSN: 0021-9258 see HPK1 gene. ---	
A	WO 99 25843 A (LUYTEN WALTER H M L ;JANSSEN PHARMACEUTICA NV (BE); PARKER ANDREW) 27 May 1999 (1999-05-27) the whole document ---	
A	EP 0 861 896 A (DADE BEHRING MARBURG GMBH) 2 September 1998 (1998-09-02) the whole document ---	
A	WO 00 06180 A (MARKEN JOHN S ;ANDERSON DIRK M (US); BIRD TIMOTHY A (US); IMMUNEX) 17 February 2000 (2000-02-17) the whole document ---	

INTERNATIONAL SEARCH REPORT

International Application No.
PCT/US 01/12992

C (Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indications, where appropriate, of the relevant passages	Relevant to claim No.
E	WO 01 90365 A (CURTIS RORY A J ;WEICH NADINE (US); MILLENNIUM PHARM INC (US)) 29 November 2001 (2001-11-29) see seq.ID's 1-3 -----	1-19,21, 22, 24-45,62

INTERNATIONAL SEARCH REPORT	International application No. PCT/US 01/12992
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: Although claims 18, 21, and 24 are directed to a method of treatment of the human/animal body, the search has been carried out and based on the alleged effects of the compound/composition.	
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:	
see additional sheet	
1. ☐ As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.	
2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.	
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:	
4. ☒ 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.: 1-44 (partially); 45, 63 (complete)	
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 01/12992

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

Continuation of Box I.2

Claims Nos.: 20,23

Present claims 20 and 23 relate to a product defined by reference to a desirable characteristic or property, namely it acting (ant)agonistically towards the protein(s) of claim 1.

The application provides no support within the meaning of Article 6 PCT and/or disclosure within the meaning of Article 5 PCT for any such products. In the present case, the claims so lack support, and the application so lacks disclosure, that a meaningful search for said claims is impossible. Independent of the above reasoning, the claims also lack clarity (Article 6 PCT). An attempt is made to define the product by reference to a result to be achieved. Again, this lack of clarity in the present case is such as to render a meaningful search impossible. Consequently, the present search report does not extend to said claims.

The applicant's attention is drawn to the fact that claims, or parts of claims, relating to inventions in respect of which no international search report has been established need not be the subject of an international preliminary examination (Rule 66.1(e) PCT). The applicant is advised that the EPO policy when acting as an International Preliminary Examining Authority is normally not to carry out a preliminary examination on matter which has not been searched. This is the case irrespective of whether or not the claims are amended following receipt of the search report or during any Chapter II procedure.

International Application No. PCT/US 01/12992

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

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

1. Claims: Invention 1: claims 45,63,
and claims 1-44 partially

Human kinase polypeptide according to seq.ID.1, homologues with at least 98% homology thereto, immunogenic fragments thereof, nucleic acids encoding them, expression vector comprising said nucleic acid, host comprising said vector, method for producing said polypeptide, antibody directed against said polypeptide, method for detecting the nucleic acid, compositions of the polypeptide or the antibody, method for identifying (ant)agonists and/or modulators of activity or expression of the polypeptide/nucleic acid, method for diagnosis using levels of said polypeptide/nucleic acid.

2. Claims: Invention 2: claims 46,64,
and claims 1-44 partially

Human kinase polypeptide according to seq.ID.2, homologues with at least 90% homology thereto, immunogenic fragments thereof, nucleic acids encoding them, expression vector comprising said nucleic acid, host comprising said vector, method for producing said polypeptide, antibody directed against said polypeptide, method for detecting the nucleic acid, compositions of the polypeptide or the antibody, method for identifying (ant)agonists and/or modulators of activity or expression of the polypeptide/nucleic acid, method for diagnosis using levels of said polypeptide/nucleic acid.

3. Claims: Inventions 3-18: claims 46-62 and 64-80 as
applicable, and claims 1-44 partially

Subject matter as defined for invention 2, but limited to the respective polypeptide sequences 3-18, and nucleic acids encoding them, whereby invention 3 corresponds to seq.ID.3, ..., and invention 18 corresponds to seq.ID.18.

For the sake of conciseness, the first subject matter is explicitly defined, the other subject matters are defined by analogy thereto.

INT NATIONAL SEARCH REPORT
 Information on patent family members

 International Application No.
PCT/US 01/12992

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9957132	A	11-11-1999	AU 4071199 A
			23-11-1999
			CA 2327457 A1
			11-11-1999
			EP 1077991 A1
			28-02-2001
WO 9925843	A	27-05-1999	WO 9957132 A1
			11-11-1999
			AU 5475199 A
			06-03-2000
			EP 1112285 A1
			04-07-2001
EP 0861896	A	02-09-1998	WO 0009551 A1
			24-02-2000
			AU 1232299 A
			10-05-1999
			AU 2612699 A
			07-06-1999
WO 0008180	A	17-02-2000	CA 2306492 A1
			27-05-1999
			CA 2308013 A1
			29-04-1999
			WO 9925843 A2
			27-05-1999
WO 0190365	A	29-11-2001	WO 9920747 A2
			29-04-1999
			EP 1025236 A2
			09-08-2000
			EP 1025237 A2
			09-08-2000
WO 0190365	A	29-11-2001	JP 2001520037 T
			30-10-2001
			JP 2001523461 T
			27-11-2001
			DE 19708173 A1
			03-09-1998
WO 0190365	A	29-11-2001	CA 2224404 A1
			28-08-1998
			EP 0861896 A2
			02-09-1998
			JP 10248566 A
			22-09-1998
WO 0190365	A	29-11-2001	US 6326181 B1
			04-12-2001
			AU 5254399 A
			28-02-2000
			CA 2305308 A1
			17-02-2000
WO 0190365	A	29-11-2001	EP 1032687 A2
			06-09-2000
			WO 0008180 A2
			17-02-2000
			AU 6328701 A
			03-12-2001
WO 0190365	A	29-11-2001	WO 0190365 A2
			29-11-2001

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 35/00	4 B 0 6 5
A 6 1 P 35/00	A 6 1 P 37/00	4 C 0 8 4
A 6 1 P 37/00	C 0 7 K 1/22	4 C 0 8 5
C 0 7 K 1/22	C 0 7 K 16/40	4 H 0 4 5
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 Q 1/68	A
C 1 2 N 9/12	G 0 1 N 33/15	Z
C 1 2 N 15/02	G 0 1 N 33/50	Z
C 1 2 Q 1/68	G 0 1 N 33/53	D
G 0 1 N 33/15	G 0 1 N 33/53	M
G 0 1 N 33/50	G 0 1 N 33/53	N
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
// C 1 2 P 21/08	C 1 2 N 15/00	C
	A 6 1 K 37/02	
	C 1 2 P 21/08	

- (31)優先権主張番号 60/203,505
 (32)優先日 平成12年5月11日(2000.5.11)
 (33)優先権主張国 米国(US)
 (31)優先権主張番号 60/205,654
 (32)優先日 平成12年5月18日(2000.5.18)
 (33)優先権主張国 米国(US)
 (31)優先権主張番号 60/207,739
 (32)優先日 平成12年5月26日(2000.5.26)
 (33)優先権主張国 米国(US)
 (31)優先権主張番号 60/208,795
 (32)優先日 平成12年6月1日(2000.6.1)
 (33)優先権主張国 米国(US)

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

- (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 9 4 0 2 5 ・ メンロパーク ・ # 1 ・ ローブルアベニュー 8 3 7
 (72)発明者 トリボレー、キャサリーン・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 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 ・ サンフランシスコ ・ ウッドサイドアベニュー 5 0
 (72)発明者 グリフィン、ジェニファー・エイ

- アメリカ合衆国カリフォルニア州94002・ベルモント・#11・アイリーンコート 181
- (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95118・サンノゼ・リッジウッドドライブ 1403
- (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94043・マウンテンビュー・アンナアベニュー 366
- (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95123・サンノゼ・コイドライブ 233
- (72)発明者 ラル、プリーティ
- アメリカ合衆国カリフォルニア州95056・サンタクララ・ピーオーボックス 5142
- (72)発明者 バーフォード、ニール
- アメリカ合衆国コネチカット州06422・ダラム・ワイルドウッドサークル 105
- (72)発明者 カーン、ファラ・エイ
- アメリカ合衆国カリフォルニア州94040・マウンテンビュー・#221・エスケラアベニュー 333
- (72)発明者 ワライヤ、ナリンダー・ケイ
- アメリカ合衆国カリフォルニア州94577・サンレアンドロ・#205・デービスストリート 890
- (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94043・マウンテンビュー・フレデリックコート 111
- (72)発明者 パターソン、チャンドラ
- アメリカ合衆国カリフォルニア州94025・メンロパーク・#1・シャーウッドウェイ 490
- (72)発明者 バリル、ジョン・ディー
- アメリカ合衆国カリフォルニア州94061・レッドウッドシティ・エメラルドヒルロード 650
- (72)発明者 マーカス、グレゴリー・エイ
- アメリカ合衆国カリフォルニア州94061・レッドウッドシティ・コネチカットドライブ 1714
- (72)発明者 ジングラー、カート・エイ
- アメリカ合衆国カリフォルニア州94117・サンフランシスコ・#シー・アッシュベリーストリート 723
- (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95115-3873・サンフランシスコ・フォーチュナアベニュー 85
- (72)発明者 リュ、ヤン
- アメリカ合衆国カリフォルニア州94303・パロアルト・コリーナウェイ 3885
- (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 ソーントン、マイケル
- アメリカ合衆国カリフォルニア州94062-2612・ウッドサイド・メッドウェイロード 9
- (72)発明者 タング、ワイ・トム
- アメリカ合衆国カリフォルニア州95118・サンノゼ・ランウィックコート 4230
- (72)発明者 ハファリア、エープリル
- アメリカ合衆国カリフォルニア州95054・サンタクララ・コーレデブリマベール 2227
- (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95121・サンノゼ・ポルトンブレイスウェイ 3770
- (72)発明者 ボーグン、マライア・アール
- アメリカ合衆国カリフォルニア州94577・サンレアンドロ・サンティアゴロード 14244
- (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・バウダリーコート 8

- (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 9 4 5 5 5 ・ フレモント・メイバードサークル 3 4 3 5 9
- (72)発明者 ボロースキー、マーク・エル
アメリカ合衆国カリフォルニア州 9 4 0 6 1 ・ レッドウッドシティー・オーチャードアベニュー
1 2 2
- (72)発明者 オウーヤング、ジャニス
アメリカ合衆国カリフォルニア州 9 4 0 0 5 ・ ブリスベーン・ゴールデンイーグルレーン 2 3 3
- (72)発明者 ヒルマン、ジェニファー・エル
アメリカ合衆国カリフォルニア州 9 4 0 4 0 ・ マウンテンビュー・# 1 7 ・ モンロードライブ 2
3 0
- (72)発明者 ガルラジャン、ラジャゴパル
アメリカ合衆国カリフォルニア州 9 5 1 1 8 ・ サンノゼ・デントアベニュー 5 5 9 1

F ターム(参考) 2G045 AA34 AA35 BB01 BB07 BB14 BB20 BB46 BB48 BB50 BB51
DA13 DA36 FA16 FB02 FB03 FB08 FB12 GC15
4B024 AA01 AA11 BA10 BA44 CA04 DA02 DA05 DA06 DA11 DA12
EA02 EA04 GA11 HA12 HA15
4B050 CC03 LL01 LL03
4B063 QA19 QA20 QQ08 QQ44 QR08 QR42 QR56 QS25 QS34 QX02
4B064 AG27 CA02 CA05 CA06 CA10 CA19 CC24 DA01 DA13
4B065 AA01X AA57X AA72X AA90X AB01 BA02 CA25 CA27 CA44 CA46
4C084 AA02 AA07 AA17 BA01 BA08 BA22 BA23 CA18 MA13 MA52
MA56 MA57 MA59 MA60 MA63 MA66 NA14 ZA362 ZB052 ZB262
ZC332
4C085 KA03 KA04 LL05 LL07 LL15 LL18
4H045 AA10 AA11 AA20 AA30 BA10 CA40 DA76 DA89 EA20 EA50
FA72 FA74 GA26

专利名称(译)	人类激酶		
公开(公告)号	JP2004511204A	公开(公告)日	2004-04-15
申请号	JP2001578626	申请日	2001-04-20
[标]申请(专利权)人(译)	洞察Genomics公司		
申请(专利权)人(译)	洞察基因组公司		
[标]发明人	ユエヘンリー ガンディーアミーナアール トリボレーキャサリーンエム キーニーライアム グリフィンジェニファーエイ ニュエンダニエルビー バンドマンオルガ リュデュングアイナエム ラルプリーティ バーフォードニール カーンファラエイ ワライヤナリンダーケイ ヤオモニークジー パターソンチャンドラ バリルジョンディー マーカスグレゴリーエイ ジングラーカートエイ レシボンシャーリーエイ リュヤン ポリッキージェニファーエル ソーントンマイケル タングワイトム ハファリアエープリル エリオットビッキエス ボーグンマライアアール ウォルシュロドリックティー ランクマールジャヤラクシミ ボロースキーマークエル オウヤングジャニス ヒルマンジェニファーエル ガルラジャンラジャゴバル		
发明人	ユエ、ヘンリー ガンディー、アミーナ・アール トリボレー、キャサリーン・エム キーニー、ライアム グリフィン、ジェニファー・エイ ニュエン、ダニエル・ビー バンドマン、オルガ リュ、デュング・アイナ・エム ラル、プリーティ バーフォード、ニール カーン、ファラ・エイ ワライヤ、ナリンダー・ケイ ヤオ、モニーク・ジー		

パターソン、チャンドラ
バリル、ジョン・ディー
マーカス、グレゴリー・エイ
ジングラー、カート・エイ
レシボン、シャーリー・エイ
リュ、ヤン
ポリッキー、ジェニファー・エル
ソーントン、マイケル
タング、ワイ・トム
ハファリア、エープリル
エリオット、ビッキ・エス
ボーグン、マライア・アール
ウォルシュ、ロドリック・ティー
ランクマール、ジャヤラクシミ
ボロースキー、マーク・エル
オウ・ヤング、ジャニス
ヒルマン、ジェニファー・エル
ガルラジャン、ラジャゴパル

IPC分类号	G01N33/50 A61K38/00 A61K45/00 A61K49/00 A61P3/06 A61P9/00 A61P35/00 A61P37/00 C07K1/22 C07K16/40 C12N1/15 C12N1/19 C12N1/21 C12N5/10 C12N9/12 C12N15/02 C12N15/09 C12P21/08 C12Q1/68 G01N33/15 G01N33/53 G01N33/566
FI分类号	C12N15/00.ZNA.A A61K45/00 A61K49/00.A A61P3/06 A61P9/00 A61P35/00 A61P37/00 C07K1/22 C07K16/40 C12N1/15 C12N1/19 C12N1/21 C12N9/12 C12Q1/68.A G01N33/15.Z G01N33/50.Z G01N33/53.D G01N33/53.M G01N33/53.N G01N33/566 C12N5/00.A C12N15/00.C A61K37/02 C12P21/08
F-TERM分类号	2G045/AA34 2G045/AA35 2G045/BB01 2G045/BB07 2G045/BB14 2G045/BB20 2G045/BB46 2G045/BB48 2G045/BB50 2G045/BB51 2G045/DA13 2G045/DA36 2G045/FA16 2G045/FB02 2G045/FB03 2G045/FB08 2G045/FB12 2G045/GC15 4B024/AA01 4B024/AA11 4B024/BA10 4B024/BA44 4B024/CA04 4B024/DA02 4B024/DA05 4B024/DA06 4B024/DA11 4B024/DA12 4B024/EA02 4B024/EA04 4B024/GA11 4B024/HA12 4B024/HA15 4B050/CC03 4B050/LL01 4B050/LL03 4B063/QA19 4B063/QA20 4B063/QQ08 4B063/QQ44 4B063/QR08 4B063/QR42 4B063/QR56 4B063/QS25 4B063/QS34 4B063/QX02 4B064/AG27 4B064/CA02 4B064/CA05 4B064/CA06 4B064/CA10 4B064/CA19 4B064/CC24 4B064/DA01 4B064/DA13 4B065/AA01X 4B065/AA57X 4B065/AA72X 4B065/AA90X 4B065/AB01 4B065/BA02 4B065/CA25 4B065/CA27 4B065/CA44 4B065/CA46 4C084/AA02 4C084/AA07 4C084/AA17 4C084/BA01 4C084/BA08 4C084/BA22 4C084/BA23 4C084/CA18 4C084/MA13 4C084/MA52 4C084/MA56 4C084/MA57 4C084/MA59 4C084/MA60 4C084/MA63 4C084/MA66 4C084/NA14 4C084/ZA362 4C084/ZB052 4C084/ZB262 4C084/ZC332 4C085/KA03 4C085/KA04 4C085/LL05 4C085/LL07 4C085/LL15 4C085/LL18 4H045/AA10 4H045/AA11 4H045/AA20 4H045/AA30 4H045/BA10 4H045/CA40 4H045/DA76 4H045/DA89 4H045/EA20 4H045/EA50 4H045/FA72 4H045/FA74 4H045/GA26
優先権	60/199021 2000-04-20 US 60/200226 2000-04-28 US 60/202339 2000-05-05 US 60/203505 2000-05-11 US 60/205654 2000-05-18 US 60/207739 2000-05-26 US 60/208795 2000-06-01 US
外部リンク	Espacenet

摘要(译)

本发明提供人激酶 (PKIN) 和鉴定和编码PKIN的多核苷酸。本发明还提供了表达载体，宿主细胞，抗体，激动剂和拮抗剂。此外，本发明提供了诊断，治疗和预防与PKIN异常表达相关的疾病的方法。

Inocle プロジ エクト ID	ホバツド SSQ ID No:	ホバツド ID	ホバツド SSQ ID No:	Inocle ホバツド ID
2890544	1	2890544001	19	2890544001
7472693	2	7472693001	20	7472693001
3107952	3	3107952001	21	3107952001
5544420	4	5544420001	22	5544420001
7472832	5	7472832001	23	7472832001
1551456	6	1551456001	24	1551456001
2589355	7	2589355001	25	2589355001
4357117	8	4357117001	26	4357117001
5511992	9	5511992001	27	5511992001
7474560	10	7474560001	28	7474560001
7474602	11	7474602001	29	7474602001
7475509	12	7475509001	30	7475509001
7475491	13	7475491001	31	7475491001
2192119	14	2192119001	32	2192119001
7474496	15	7474496001	33	7474496001
1834248	16	1834248001	34	1834248001
71584520	17	7158452001	35	7158452001
7475538	18	7475538001	36	7475538001