

(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2004-503211  
(P2004-503211A)

(43) 公表日 平成16年2月5日(2004.2.5)

(51) Int. Cl. <sup>7</sup>	F I	テーマコード (参考)
C 1 2 N 15/09	C 1 2 N 15/00	Z N A A
A 6 1 K 38/45	A 6 1 K 48/00	2 G O 4 5
A 6 1 K 48/00	A 6 1 P 25/00	4 B O 2 4
A 6 1 P 25/00	A 6 1 P 35/00	4 B O 5 0
A 6 1 P 35/00	C 1 2 N 9/12	4 B O 6 3
		4 B O 6 4
	審査請求 未請求 予備審査請求 有	(全 121 頁) 最終頁に続く

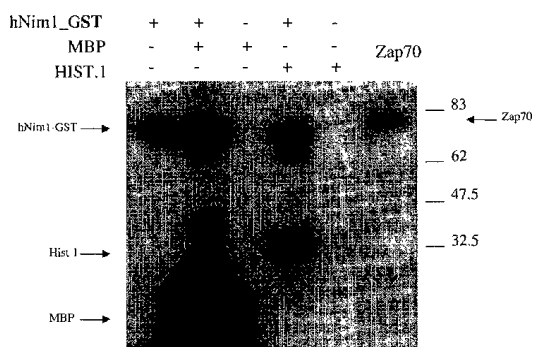
(21) 出願番号	特願2001-567309 (P2001-567309)	(71) 出願人	500014002 インサイト・ファーマスーティカルズ・インコーポレイテッド INCYTE PHARMACEUTICALS INC. アメリカ合衆国カリフォルニア州94304・パロアルト・ポータードライブ 3160
(86) (22) 出願日	平成12年3月22日 (2000.3.22)	(71) 出願人	502334412 モルテーニ、アンジェラ MOLTENI, Angela イタリア国カントゥ市アイー22603・ピアギセツペチッテリーオ 4
(85) 翻訳文提出日	平成14年9月13日 (2002.9.13)	(74) 代理人	100089266 弁理士 大島 陽一
(86) 国際出願番号	PCT/US2000/007715		
(87) 国際公開番号	W02001/068825		
(87) 国際公開日	平成13年9月20日 (2001.9.20)		
(31) 優先権主張番号	09/523,849		
(32) 優先日	平成12年3月13日 (2000.3.13)		
(33) 優先権主張国	米国 (US)		

(54) 【発明の名称】 ヒトNIM1キナーゼ

(57) 【要約】

本発明は、ヒトNIM1キナーゼをコードする核酸分子を提供する。本発明はまた、発現に関連する疾患の特徴付け、診断、評価、治療、または予防における本核酸分子、その断片、その変異体、及びその相補体、更に本タンパク質、その一部、並びにその抗体の利用法を提供する。本発明は更に、本タンパク質を発現させるための発現ベクター及び宿主細胞、並びに遺伝子組換え生物またはモデル系を提供する。

KINASE ASSAY USING hNIM1



## 【特許請求の範囲】

## 【請求項 1】

SEQ ID NO : 2 を含むタンパク質またはその一部をコードする実質的に精製された核酸分子またはその一部。

## 【請求項 2】

請求項 1 のポリヌクレオチド、その断片、またはその相補配列 ( c o m p l e m e n t ) を含む組成物。

## 【請求項 3】

SEQ ID NO : 24 - 30 から選択された請求項 1 の核酸分子の哺乳動物変異体。

## 【請求項 4】

SEQ ID NO : 1 及び SEQ ID NO : 3 - 30 またはそれらの相補配列から選択される少なくとも 18 の連続するヌクレオチド断片。

## 【請求項 5】

請求項 4 の断片を含む基板。

## 【請求項 6】

請求項 4 の断片を含むプローブ。

## 【請求項 7】

請求項 1 の核酸分子を含む発現ベクター。

## 【請求項 8】

請求項 7 の発現ベクターを含む宿主細胞。

## 【請求項 9】

タンパク質を生産する方法であって、

( a ) 前記タンパク質が発現する条件下で、請求項 8 の宿主細胞を培養するステップと、

( b ) 前記宿主細胞から前記タンパク質を回収するステップとを含むことを特徴とするタンパク質生産方法。

## 【請求項 10】

サンプルにおける核酸分子を検出するための方法であって、

( a ) 請求項 4 の断片を前記サンプルの少なくとも 1 つの核酸分子とハイブリダイズさせて、ハイブリダイゼーション複合体を形成させるステップと、

( b ) 前記ハイブリダイゼーション複合体を検出するステップとを含み、

前記ハイブリダイゼーション複合体の存在が、前記サンプルにおける前記核酸分子の存在を示すことを特徴とする核酸分子の検出方法。

## 【請求項 11】

前記断片を脳疾患または癌の治療に用いることを特徴とする請求項 10 に記載の方法。

## 【請求項 12】

前記ハイブリダイゼーションの前に、前記サンプルの前記核酸分子を増幅するステップを更に含むことを特徴とする請求項 10 に記載の方法。

## 【請求項 13】

核酸分子を用いて分子または化合物のライブラリをスクリーニングするための方法であって、

( a ) 特異的な結合が許容される条件下で、請求項 1 の核酸分子を前記分子または化合物のライブラリと結合させるステップと、

( b ) 特異的な結合を検出して、前記核酸分子に特異的に結合する分子または化合物を同定するステップとを含むことを特徴とする方法。

## 【請求項 14】

前記ライブラリが、DNA 分子、RNA 分子、ペプチド核酸、人工の染色体作製物、ペプチド、及びタンパク質から選択されることを特徴とする請求項 13 に記載の方法。

## 【請求項 15】

SEQ ID NO : 2 を含む単離され精製されたタンパク質またはその一部。

## 【請求項 16】

10

20

30

40

50

タンパク質を用いて分子及び化合物のライブラリをスクリーニングする方法であって、  
(a) 特異的な結合が許容される条件下で、請求項 15 のタンパク質を前記分子または化合物のライブラリと結合させるステップと、  
(b) 特異的な結合を検出して、前記タンパク質に特異的に結合する分子または化合物を同定するステップとを含むことを特徴とする方法。

【請求項 17】

前記ライブラリが、DNA 分子、RNA 分子、ペプチド核酸、ペプチド、タンパク質、擬態、アゴニスト、アンタゴニスト、抗体、免疫グロブリン、インヒビター、及び薬剤から選択されることを特徴とする請求項 16 に記載の方法。

【請求項 18】

抗体を準備する方法であって、  
(a) 抗体反応が起こる条件下で、請求項 15 のタンパク質またはそのタンパク質の抗原性を有する部分で動物を免疫化するステップと、  
(b) 動物抗体を単離するステップと、  
(c) 単離した抗体を前記タンパク質でスクリーニングして、前記タンパク質に特異的に結合する抗体を同定するステップとを含むことを特徴とする方法。

【請求項 19】

NIM1 キナーゼに特異的に結合する抗体を用いて疾患を診断する方法であって、  
(a) 特異的な結合が許容される条件下で、前記抗体をサンプルと結合させるステップと、  
(b) 結合した抗体を検出するステップと、  
(c) 知られている標準的な方法で発現を比較して、疾患の存在を確認するステップとを含むことを特徴とする方法。

【請求項 20】

前記疾患が脳疾患または癌であることを特徴とする請求項 19 に記載の方法。

【発明の詳細な説明】

【0001】

(発明の技術分野)

本発明は、ヒト NIM1 キナーゼをコードする核酸分子に関連し、また、脳疾患及び癌、特に乳癌の特徴付け・診断・予防・治療に、この核酸分子及びこの核酸分子がコードするタンパク質を利用することに関連する。

【0002】

(発明の背景)

生物間の系統発生的な関係は何度も実証されており、様々な原核生物及び真核生物の研究結果は、分子、生化学的及び生理学的機構、及び代謝経路の漸進的な進化を示している。異なった進化圧力にもかかわらず、線虫、ハエ、ラット、及びヒトのプロテインキナーゼは、共通の化学的及び構造的特徴を有し、同一の一般的な細胞活性を調節する。構造及び/または機能が分かっている生物の核酸配列及びタンパク質配列を比較することにより、ヒト配列の分析が進み、ヒトの病態、疾患、及び障害のための診断薬及び治療薬を検査するためのモデル系の作製が可能となる。

【0003】

プロテインキナーゼは、タンパク質にリン酸基を付加することにより多くの異なる細胞増殖、細胞分化、及びシグナル伝達を調節する。制御のきかないシグナル伝達は、炎症、癌、動脈硬化、及び乾癬を含む様々な病態に関係することが分かっている。可逆的なタンパク質リン酸化は、真核細胞の活性を調節するための主な方法である。典型的な哺乳動物細胞における 10,000 のタンパク質活性の内の 1,000 以上がリン酸化であると推定される。活性化させる高エネルギーのリン酸は、一般にプロテインキナーゼによりアデノシン三リン酸 (ATP) またはグアノシン三リン酸 (GTP) から特定のタンパク質に導入され、プロテインホスファターゼによりそのタンパク質から除去される。リン酸化は分子スイッチをオンにするのに類似しており、細胞外シグナル (ホルモン、神経伝達物質

10

20

30

40

50

、成長及び分化因子などの分子によって仲介される)、細胞周期チェックポイント、及び環境または栄養のストレスに应答して起こる。スイッチがオンになると、適当なプロテインキナーゼが代謝酵素、調節タンパク質、受容体、細胞骨格タンパク質、イオンチャネルまたはポンプ、または転写因子を活性化させる。

#### 【0004】

プロテインキナーゼは、多様な機能及び特性を有する酵素のスーパーファミリーである最も大きな既知のタンパク質群を構成している。プロテインキナーゼは、通常はそれらの基質、それらの調節分子、または変異表現型の或る特徴によって命名される。基質について述べると、プロテインキナーゼは、チロシン残基をリン酸化する群(プロテインチロシンキナーゼ、PTK)とセリンまたはトレオニン残基をリン酸化する群(セリン/トレオニンキナーゼ、STK)との2つの群に概ね分類される。僅かであるがプロテインキナーゼの中には、トレオニン及びチロシン残基の両方をリン酸化する二重特異性を有するものもある。

10

#### 【0005】

プロテインキナーゼは、触媒ドメインのいずれかの側に位置するアミノ酸残基の違い(通常は5~100残基の間)または触媒ドメインのループ内に挿入されたアミノ酸残基の違いによって複数のファミリーに分類され得る。これらの残基により、標的タンパク質を認識して相互作用する時の各キナーゼの調節が可能となる。殆んど全てのキナーゼは、2つのローブに亘って分布する11のサブドメインを有する250~300のアミノ酸からなる類似した触媒ドメインを含む。サブドメインI-IVを含むN末端ローブは、ATP供与体分子と結合しそのATP供与体分子を方向付ける。サブドメインVIA-XIを含む大きなC末端ローブは、タンパク質基質と結合し、リン酸をATPからセリン、トレオニン、またはチロシン残基の水酸基へ転移させる。サブドメインVは、N末端ローブ及びC末端ローブに亘っている。

20

#### 【0006】

11のサブドメインのそれぞれは、サブドメインの特徴であって高度に保存されている特異的なアミノ酸残基及びモチーフ、またはアミノ酸パターンを含む(Hardie及びHanks(1995) The Protein Kinase Facts Books, Vol I, Academic Press, San Diego CA, pp. 7-20)。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において同定された。第1のシグネチャ配列は、ATP結合に関与する活性部位リシン残基を含み、第2のシグネチャ配列は、触媒活性に重要なアスパラギン酸残基を含む。タンパク質が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可能性はほぼ100%である(MOTIFS 検索プログラム、Genetics Computer Group, Madison WI; Bairochら(1996) Nucleic Acids Res 24:189-196)。

30

#### 【0007】

##### STKファミリー

セカンドメッセンジャー依存性プロテインキナーゼは、サイクリックAMP(cAMP)、サイクリックGMP、イノシトール三リン酸、ホスファチジルイノシトール、3,4,5-三リン酸、サイクリックADPリボース、アラキドン酸、ジアシルグリセロール及びカルシウム-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中心に仲介する。サイクリックAMP依存性プロテインキナーゼ(PKA)はSTKファミリーの重要なメンバーである。cAMPは、研究した全ての真核細胞及び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。このような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、及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、及び心血管疾患に関連する(Isselbacher, K. J.ら, (1994) Harrison's Principles o

40

50

*f Internal Medicine*, McGraw-Hill, New York, NY, pp. 416-431, 1887)。

【0008】

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もまたSTKファミリーのメンバーである。カルモジュリンは、カルシウムの結合にตอบสนองして標的タンパク質に結合することにより、多くのカルシウム調節性プロセスを仲介するカルシウム受容体である。これらのプロセスの主な標的タンパク質はCaM依存性プロテインキナーゼである(CaMK)。CaMK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)、及び神経伝達(CaMKキナーゼI及びCaMKキナーゼII)の調節に参与する。CaMK Iは、神経伝達物質関連タンパク質であるシナプシンI及びII、遺伝子転写調節因子であるCREB、及び囊胞性繊維コンダクタンス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)。CaMK IIもまた、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及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の多くは、CaMに結合するのに加えてリン酸化により活性化される。CaMKは、キナーゼカスケードの一部として他のキナーゼによりリン酸化されたり、自己リン酸化し得る。

10

【0009】

もう1つのリガンド活性化プロテインキナーゼは、5'-AMP-活性化プロテインキナーゼである(AMPK; Gao, G.ら, (1996) *J. Biol. Chem.* 271:17998-17803)。哺乳動物AMPKは、酵素であるアセチル-CoAカルボキシラーゼ及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及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及びATPの枯渇などの細胞内ストレスに対するこれらの経路の応答を仲介する。AMPKは、触媒サブユニット、並びにこのサブユニットの活性を調節すると考えられている2つの非触媒サブユニット及びサブユニットからなるヘテロ三量体複合体である。AMPKのサブユニットは、脳、心臓、脾臓、及び肺などの非脂肪性組織において予想以上に広く分布している。この分布から脂質の代謝の調節以外にもある役割を果たしていると考えられる。

20

【0010】

マイトジェン活性化プロテインキナーゼ(MAPK)もまた、STKファミリーのメンバーである。MAPKは、リン酸化カスケードを介して細胞表面から核へのシグナル伝達を仲介する。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ตอบสนองする(Egan及びWeinberg (1993) *Nature* 365:781-783)。MAPキナーゼシグナル伝達経路は、哺乳動物細胞及び酵母に存在する。哺乳動物経路を活性化する細胞外刺激には、上皮成長因子、紫外線、高浸透圧媒体、熱ショック、内毒素性リポ多糖、腫瘍壊死因子及びインターロイキン-1などの前炎症性サイトカインが含まれる。MAPKの発現の変化は、癌、炎症、免疫疾患、及び成長及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。

30

【0011】

血清/サイトカイン誘導性STKである増殖関連キナーゼ(PRK)は、ヒト巨核細胞における細胞サイクル及び細胞増殖の調節に参与する(Liら (1996) *J Biol Chem* 271:19402-8)。PRKは、細胞分裂に関係するSTKのp10ファミリーに関連する。PRKは肺腫瘍組織においてダウンレギュレートされ、正常組織における抑制の効かないその発現が発癌に繋がるプロトオンコジーンであると思われる。

40

【0012】

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の別の群である。サイクリンは小さな調節タンパク質であって、CDKに結合してCDKを活性化させる。次に、そのCDKが有糸分裂プロセスに参与する選択されたタンパク質をリン酸化及び活性化することにより、細胞周期の様々な相を開始させる。C

50

D Kは、活性化するのに複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、C D Kの活性化には、特定のトレオニン残基のリン酸化及び特定のチロシン残基の脱リン酸が必要である。

【0013】

ヒトN I M 1キナーゼ(h N I M)をコードする核酸分子の発見により、脳疾患及び癌、特に乳癌の特徴付け、診断、予防、治療に有用な新規の組成物を提供する。

【0014】

(発明の要約)

ヒトN I M 1キナーゼをコードする実質的に精製された核酸分子の発見に基づき、脳疾患及び癌、特に乳癌の特徴付け・診断・予防・治療に有用な組成物を提供することで当分野の要望に応える。

10

【0015】

本発明は、S E Q I D N O : 2を含むヒトN I M 1キナーゼをコードする実質的に精製された核酸分子を提供する。本発明はまた、S E Q I D N O : 1、または断片や相補配列を含む組成物を提供する。本発明は更に、S E Q I D N O : 24 - 30から選択される核酸分子の哺乳動物変異体を提供する。本発明はまた、S E Q I D N O : 3 - 30、S E Q I D N O : 1のヌクレオチドの約414 ~ 約1414、またはそれらの相補配列から選択される少なくとも18の連続するヌクレオチド断片を提供する。一実施態様では、本発明はこれらの断片の少なくとも1つを含む基板を提供する。第2の実施態様では、本発明は、検出、スクリーニング、及び精製に用いることができる断片を含むプローブを提供する。更なる実施態様では、このプローブは一本鎖相補R N A分子またはD N A分子である。

20

【0016】

本発明は更に、サンプルにおける核酸分子を検出する方法であって、プローブ即ち相補的な核酸分子をそのサンプルの少なくとも1つの核酸分子とハイブリダイズするステップと、ハイブリダイゼーション複合体を形成するステップと、そのハイブリダイゼーション複合体を検出するステップとを含み、そのハイブリダイゼーション複合体の存在がそのサンプルに核酸分子が存在することを示す方法を提供する。一実施態様では、この方法は更に、ハイブリダイゼーションの前にサンプルの核酸分子を増幅するステップを含む。この核酸分子、その断片、またはその相補配列が、アレイ上のエレメントを構成し得る。

30

【0017】

本発明はまた、核酸配分子、またはその断片や相補配列を用いて、分子または化合物のライブラリをスクリーニングして、その核酸分子と特異的に結合する少なくとも1つのリガンドを同定する方法であって、特異的な結合が許容される条件下で、その核酸分子をその分子または化合物のライブラリと結合させるステップと、その核酸分子に対する特異的な結合を検出して、その核酸分子と特異的に結合するリガンドを同定するステップとを含む方法を提供する。一実施態様では、このような分子及び化合物のライブラリには、D N A分子、R N A分子、ペプチド、P N A、及びタンパク質等が含まれる。

【0018】

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

40

【0019】

本発明はまた、S E Q I D N O : 2を含む単離され精製されたタンパク質またはその一部を提供する。更に、本発明はまた、医薬用担体と共に実質的に精製されたタンパク質またはその一部を含む医薬組成物を提供する。

【0020】

本発明は更に、本タンパク質の少なくとも一部を用いて抗体を産生させる方法を提供する。本発明はまた、タンパク質を用いて分子または化合物のライブラリをスクリーニングし

50

て、そのタンパク質に特異的に結合する少なくとも1つのリガンドを同定する方法を提供する。この方法は、特異的な結合が許容される条件下で、そのタンパク質を分子または化合物のライブラリと結合させるステップと、結合したタンパク質を検出して、そのタンパク質と特異的に結合したリガンドを同定するステップとを含む。このような分子及び化合物のライブラリには、アゴニスト、アンタゴニスト、抗体、DNA分子、RNA分子、免疫グロブリン、薬剤化合物、擬態、ペプチド、医薬品、及びその他のリガンドが含まれる。本発明は更に、本タンパク質を用いてリガンドを精製する類似の方法を提供する。この方法は、特異的な結合が許容される条件下で、本タンパク質をサンプルと結合させステップと、結合したタンパク質を回収するステップと、そのリガンドから本タンパク質を分離して、精製されたリガンドを得るステップとを含む。

10

**【0021】**

本発明は更に、ヒトNIM1キナーゼに対して作製された抗体を用いるスクリーニング法により同定された抗体を提供する。抗体を作製する方法は、抗体反応が起こる条件下で、ヒトNIM1キナーゼ若しくは抗原性を有するその一部で動物を免疫化するステップと、動物抗体を単離するステップと、単離した抗体をNIM1キナーゼでスクリーニングして、Nim1キナーゼに特異的に結合する抗体を同定するステップとを含む。一実施態様では、これらの抗体は、脳疾患及び癌の存在の確認における診断用組成物として有用である。別の実施態様では、この抗体を医薬組成物として、ヒトNIM1キナーゼの過剰な発現に関連する脳疾患及び癌の治療に用い得る。

**【0022】**

本発明は更に、哺乳動物のゲノムDNAの中にマーカー遺伝子を挿入して、その天然の核酸分子の発現を阻止する方法を提供する。本発明はまた、核酸分子を用いて哺乳動物モデル系を作製する方法を提供する。この方法は、SEQ ID NO: 1及びSEQ ID NO: 3 - 30から選択される核酸分子を含むベクターを作製するステップと、そのベクターを胚性幹細胞に形質転換するステップと、形質転換された胚性幹細胞を選択するステップと、この形質転換した胚性幹細胞を哺乳動物胚盤胞の中に微量注入して、キメラ胚盤胞を作製するステップと、偽妊娠メスにキメラ胚盤胞を移植して、このメスから、その生殖細胞系の中にこの核酸分子を含むキメラ子孫が出産されるようにするステップと、そのキメラ哺乳動物を交配してホモ接合性哺乳動物モデル系を作製するステップとを含む。

20

**【0023】**

(本発明の実施方法について)

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

30

**【0024】**

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

40

**【0025】**

(定義)

「NIM1キナーゼ」は、天然、合成、半合成或いは組換え体などの任意の種(ウシ、ヒツジ、ブタ、齧歯類、イヌ、サルを含むが好ましくはヒトである哺乳動物)から得られる実質的に精製された酵素を指す。

**【0026】**

50

「生物学的に活性」は、構造的または免疫学的、調節的、化学的な機能を有する、天然或いは組換え、または合成分子であるタンパク質を指す。

【0027】

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

【0028】

「誘導体」は、化学修飾された核酸分子やタンパク質配列を指す。化学修飾には、分子や配列の生物学的活性または寿命を維持或いは増大する、アルキル基またはアシル基、アミノ基による水素の置換や、グリコシル化、ポリエチレングリコール化、または類似の任意のプロセスが含まれる。 10

【0029】

「断片」は、有用な機能的特性を維持する核酸分子の任意の一部またはインサイト社クロームを指す。有用な断片は、一般にその長さが少なくとも18の連続するヌクレオチドであり、ハイブリダイゼーション、増幅、またはスクリーニング、更に複製、転写、または翻訳の調節に用いられ得るオリゴヌクレオチドを含む。

【0030】

「ハイブリダイゼーション複合体」は、プリン塩基とピリミジン塩基との間の水素結合の形成による2つの核酸分子の複合体を指す。 20

【0031】

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

【0032】

「核酸分子」は、核酸、オリゴヌクレオチド、ヌクレオチド、ポリヌクレオチド、DNA分子、またはそれらの任意の断片や相補配列を指す。また、「核酸分子」は、ゲノム若しくは合成起源の二本鎖若しくは一本鎖のDNA或いはRNAが可能であり、炭水化物または脂質、タンパク質、その他の物質と結合して、形質転換などの特殊な作用を起こしたり、ペプチド核酸(PNA)等の有用な組成物を形成する。核酸分子はまた、非翻訳5'または3'調節領域またはイントロンを含み得る。核酸分子の大きさは、好ましくは約15~10,000ヌクレオチドであり、より好ましくは約60~6,000ヌクレオチドであり、最も好ましくは約400~5000ヌクレオチドである。「オリゴヌクレオチド」は、アンプリマー(amplimer)、プライマー、オリゴマー、エレメント、標的、及びプローブと実質的に同一であり、好ましくは一本鎖であって、その長さは好ましくは約22~25ヌクレオチドである。 30

【0033】

「タンパク質」は、天然或いは合成のアミノ酸配列、オリゴペプチド、ペプチド、ポリペプチド、またはそれらの一部を指す。 40

【0034】

ここで用いる「一部」は、あらゆる目的に用いられるタンパク質の任意の一部を指すが、特に、分子や化合物のライブラリをスクリーニングしてそのタンパク質の一部と特異的に結合する分子を同定するために、または抗体の産生のために用いられる。

【0035】

「レポーター遺伝子」は、ポリヌクレオチド、ポリペプチド、または抗体を標識するために用いられる化学成分または生化学成分である。レポーター遺伝子には、限定するものではないが、放射性核種、酵素、基質、補助因子、インヒビター、蛍光剤、色素剤、化学発光剤、及び磁石粒子等が含まれる。レポーター分子は、特定のポリヌクレオチド、ポリペ 40

プチド、または抗体に特異的に結合させ、その存在を確認し、それらを定量することができる。

【0036】

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

【0037】

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

【0038】

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

【0039】

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

【0040】

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

【0041】

(発明)

本発明は、ヒトNIM1キナーゼをコードする新規の核酸分子の発見に基づき、脳疾患及び癌の特徴付け・診断・治療・予防においてこの核酸分子やその断片、またはタンパク質やその一部を直接、或いは組成として用いることに関する。

【0042】

本発明のNIM1キナーゼをコードする核酸分子は、小脳ライブラリ(CRBLNOT01)のインサイト社クローン670279と受精卵の分裂及び分化の極性にとって重要な線虫の推定上のSTK(G733122)との間のBLOCK I I 相同性一致によってキナーゼとして始めに同定された。完全長の核酸分子であるインサイト社クローン3317608(SEQ ID NO: 1)は、インサイト社LIFESEQ GOLDデータベース(1999年3月公表)テンプレート200700.1、アセンブリ670279CB1及びインサイト社クローン(ライブラリ): 3317608H1 (PROSBPT03)、4313713H1 (BRAFN01)、4617082H1 (BRAYDIT01)、4711644H1 (BRAIHCT01)、2286324H1 (BRAINON01)、2286816H1 (BRAINON01)、2287217H1 (BRAINON01)、2286816R6 (BRAINON01)、22868

16T6 (BRAINON01)、3317608T6 (PROSBPT03)、4201896T6 (BRAITUT29)、4624811T6 (FIBRTXT02)、6559834H1 (BRAFNON02)、670279F1 (CRBLNOT01)、670279H1 (CRBLNOT01)、670279R1 (CRBLNOT01)、670279R6 (CRBLNOT01)、670279T6 (CRBLNOT01)、及び4936446H1 (BRAXNOT03)、即ち、SEQ ID NO : 3 - 23を用いてシーケンシングし、構築した。SEQ ID NO : 2をコードするポリヌクレオチドの有用な断片には、SEQ ID NO : 3 - 23、SEQ ID NO : 1の約414番目のヌクレオチドから約1414番目のヌクレオチド、またはそれらの相補配列から選択された少なくとも18の連続するヌクレオチドの断片が含まれる。

10

#### 【0043】

図1A - 図1Fは、本核酸分子配列及び推論されるアミノ酸に翻訳された配列を示す。ヒトNIM1キナーゼのコード領域を含むインサイト社クローン番号3317608は、Type Culture Collection (ATCC; Manassas VA) に保管され、特許保管指定 (Patent Deposit Designation: P T - 1217) を受けている。

#### 【0044】

図2に示すように、電子的ノーザン分析は、NIM1キナーゼをコードする転写物の神経系における極めて特異的な発現を示す。これらのライブラリの全ては脳に由来し、その内の5つは癌に関連し、3つは発作に関連し、2つはハンチントン病に関連し、1つは癲癇に関連する。この転写物は、1039のライブラリ及び500万を超える配列を含むLIFESEQデータベース (Incyte Pharmaceuticals) において、脳以外のライブラリでは癌性乳房繊維芽細胞FIBRTXT02及び癌性前立腺PROSBPT03の2つのライブラリにおいてのみ発現が確認された。

20

#### 【0045】

図3に示すように、ノーザン分析はまた定量PCRを用いて実験室でも行った。ヒトNIM1キナーゼ転写物の発現は、H460、A2780、A375、HDF、HELA、DU145、MDA-MB231、U87-MG、及びBX-PC3細胞系、並びに脳、結腸、子宮、及び胎盤の組織に見られた。注目すべきは、ヒト乳癌細胞系MDA-MB231及び脳組織はそれぞれ、ヒトNIM1キナーゼ転写物の発現が約100倍及び155倍を示し、脳、子宮頸、結腸、肺、卵巣、及び前立腺の癌を示す細胞系、並びに正常な子宮及び胎盤において5倍以上の発現が見られた。

30

#### 【0046】

SEQ ID NO : 2のアミノ酸配列を含むNIM1キナーゼは、アミノ酸436個の長さであり、M1からA18までの可能性のあるシグナル配列、及びS56、T108、T114、S123、S169、S221、S282、Y288、T311、T332、T349、S359、S420、及びT429残基における潜在的なリン酸化部位を有する。図4に示すように、ヒトNIM1キナーゼを基準として用いて、アラインメントにおけるキナーゼ (SEQ ID NO : 2及びSEQ ID NO : 31 - 36) の触媒領域の保存された残基、モチーフ、及びサブドメインを調べた。分かった保存された残基、モチーフ、及びサブドメインは、G81からV88残基までのサブドメイン1、保存されたA101及びK103残基を有するサブドメイン2、インバリアントなE121残基を有するサブドメイン3、E151からE157残基まで延在するインバリアントな残基M150及びY152を共有するサブドメイン5、H194からN201までの触媒ループであるサブドメイン6B、高度に保存されたトリプレットD214、F215、及びG216を有するサブドメイン7、A238、P239、E240モチーフを含むT229からF242までのサブドメイン8、インバリアントD253を有するサブドメイン9、サブドメイン8のAPEモチーフと相互作用するインバリアントR244を有するサブドメイン11である。触媒ドメインのC末端境界は、H248、A249、及びF250で始まる。PFAM及びPRINTSの両方により、キナーゼドメインが確認された。SEQ ID

40

50

NO: 2のポリペプチドの有用な断片には、SEQ ID NO: 2の約1から約295番目までの残基から選択された少なくとも6個の連続するアミノ酸断片が含まれる。

【0047】

図5は、ヒトNIM1キナーゼアッセイを示す。hNIM1キナーゼ-GSTの自己リン酸化がレーン1、レーン2、及びレーン4に見られ、その基質MBPの自己リン酸化がレーン2に見られ、HIST.1の自己リン酸化がレーン4に見られる。キナーゼが存在しない基質MBPのレーン3及びHIST.1のレーン5は自己リン酸化を示さない。レーン6はポジティブコントロールZAP70を示す。

【0048】

ヒトNIM1キナーゼをコードする核酸分子の哺乳動物変異体を、BLASTまたはBLAST2 (Basic Local Alignment Search Tool: Altschul (1997) Nucleic Acids Res 25: 3389-3402; Altschul (1993) J Mol Evol 36: 290-300; 及びAltschul (1990) J Mol Biol 215: 403-10, with default parameters)を用いて同定し、SEQ ID NO: 1及びSEQ ID NO: 3-23と整列するクローンをLIFESEQまたはZOOSEQデータベース (Incyte Pharmaceuticals)において同定した。哺乳動物変異体は、ZOOSEQデータベーステンプレート (1999年12月に構築) 216150.1 (SEQ ID NO 24) 及びインサイト社クローン: ラットに由来する701925441H1 (RALITXS03)、701910632H1 (RABYUNN02)、701905514H1 (RABYUNS09)、701293826H1 (RABXNOT04)、及び700949543H1 (RASPNON02)、並びにサルに由来する700706950H1 (MNBFFNOT01)、即ちSEQ ID NO: 25-30である。これらの核酸分子は、ヒト疾患モデルとなる遺伝子組替え生物を作成するために特に有用であり、このようなヒト疾患モデル動物を用いてこのような疾患に対する有望な治療方法を研究することが可能である。

【0049】

核酸分子 (SEQ ID NO: 1) 及びそれらの断片 (SEQ ID NO: 3-30) を、ハイブリダイゼーション、増幅、及びスクリーニングに用いて、SEQ ID NO: 1とサンプルの類似分子との間の同定及び区別をすることが可能である。ヒト分子及びそれらの哺乳動物変異体を用いて、ヒト疾患モデルとなる遺伝子組替え生物を作製して、このヒト疾患モデルを用いて有望な治療方法を研究することができる。毒学的研究、臨床トリアル、及び被験動物/患者の治療プロフィールを、本核酸分子、タンパク質、並びに本発明の核酸分子及びタンパク質を用いて同定された抗体、分子、及び化合物を用いて作成し、モニタリングすることができる。

【0050】

(本発明の特徴及び使用)

cDNAライブラリ

ここに開示する特定の実施例では、当分野で周知の方法を用いて哺乳動物の細胞及び組織からmRNAを単離し、これを用いてcDNAライブラリを作製する。上記したインサイト社クローンは、哺乳動物cDNAライブラリから単離された。本発明の代表的な3つのライブラリの作製方法を後述する実施例に示す。コンセンサス配列は、Phrap (P Green, University of Washington, Seattle WA) 及びGELVIEW断片構築システム (Genetics Computer Group, Madison WI)、AUTOASSEMBLERアプリケーション (PE Biosystems, Foster City CA) などのコンピュータプログラムを用いて、インサイト社クローンを含む断片、伸長、及び/またはショットガン配列から化学的かつ/または電子的に構築した。クローン、伸長配列及び/またはショットガン配列は、クラスター及び/または主クラスターの中に電子的に組み入れられた。

【0051】

シーケンシング

核酸をシーケンシングする方法は当分野で周知であり、そのような方法を用いて本発明の任意の実施例を実施することができる。これらの方法は、DNAポリメラーゼIであるクレノウフラグメント、SEQUENASE、Taq DNAポリメラーゼ及び熱耐性T7 DNAポリメラーゼ(Amersham Pharmacia Biotech (APB), Picataway NJ)、或いはELONGASE増幅システム(Life Technologies, Rockville MD)に用いられるような校正エクソヌクレアーゼとポリメラーゼとの組み合わせを用いることができる。配列の準備は、HYDRAマイクロディスペンサー(Robbins Scientific, Sunnyvale CA)、MICROLAB 2200(Hamilton, Reno NV)、及びDNA ENGINEサーマルサイクラー(PTC200; MJ Research, Watertown MA)などの装置を用いて自動的に行うのが望ましい。シーケンシングに用いる装置には、ABI 3700、377または373 DNAシーケンシングシステム(PE Biosystems)、及びMEGABACE 1000 DNAシーケンシングシステム(APB)等がある。当分野で周知の様々なアルゴリズムを用いて、シーケンシングした配列を解析することができる。これらのアルゴリズムは、Ausubel(1997; Short Protocols in Molecular Biology, John Wiley & Sons, New York NY, unit 7.7) 及びMeyers(1995; Molecular Biology and Biotechnology, Wiley VCH, New York NY, pp. 856 - 853)に記載されている。

## 【0052】

ショットガンシーケンシングを用いて、目的の特定のクローニングしたインサートを完全な配列にすることが可能である。ショットガン法は、元のインサートを様々な大きさのセグメントにランダムに分割して、これらの断片をベクターにクローニングする。これらの断片を、元のインサートの全配列が分かるまで重複した端部を用いてシーケンシング即ち再構築する。ショットガンシーケンシング方法は当分野で周知であり、熱耐性DNAポリメラーゼや非熱耐性DNAポリメラーゼ、及び目的の核酸分子に隣接する代表的な領域から選択されたプライマーを用いる。当分野で周知のCONSED(Gordon(1998) Genome Res. 8:195-202)などの様々なアルゴリズムやプログラムを用いて、組み立てが不完全な配列を調べる。ベクターやキメラ配列、または欠失配列を含む汚染配列を除去して、組み立てが不完全な配列を完全な配列に組み立てる。

## 【0053】

核酸配列の伸長

本発明の配列は、当分野で周知の様々なPCR法を用いた方法で伸長することができる。例えば、XL-PCRキット(PE Biosystems)及び入れ子プライマー(nested primer)、市販のcDNAまたはゲノムDNAライブラリ用いてヌクレオチド配列を伸長することが可能である。全てのPCR系の方法に用いることができるように、プライマーは、OLIGO 4.06プライマー分析ソフトウェア(National Biosciences, Plymouth MN)等の市販のソフトウェアを用いて、ヌクレオチドの長さが約22~30個、GC含量が約50%以上、約55~68の温度で標的配列とアニールするように設計することが可能である。調節エレメントを復活させるために配列を伸長する場合は、cDNAライブラリよりゲノムライブラリを用いる方が良い。

## 【0054】

(本哺乳動物核酸分子の使用)

ハイブリダイゼーション

本核酸分子及びその断片は、様々な目的のための様々なハイブリダイゼーション技術に用いることができる。プローブは、5'調節領域や非保存領域(即ち、本タンパク質の保存

された触媒ドメインをコードするヌクレオチドの5'や3')などのユニークな領域から作製可能であり、これらのプローブを、ヒトNIM1キナーゼ、アレル変異体、または関連分子をコードする天然の分子を同定するためのプロトコルに用いることができる。このプローブは、DNAまたはRNAからなる一本鎖が可能であって、任意の核酸配列(SEQ ID NO: 3-30)と少なくとも50%の配列同一性を有すべきである。ハイブリダイゼーションプローブは、レポーター分子の存在下でのPCR増幅、オリゴ標識化、ニックトランスレーション法、または末端標識化を利用して作製することができる。この核酸分子またはその断片を含むベクターを用いて、RNAポリメラーゼ及び標識したヌクレオチドを加えて*in vitro*でmRNAプローブを作製することができる。これらの方法はAmersham Pharmacia Biotech (APB)社が販売するキットを用いて行うことができる。

10

#### 【0055】

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

20

30

#### 【0056】

当分野で周知の方法でアレイを準備して分析することができる。オリゴヌクレオチドをアレイのプローブや標的として用いることができる。アレイを用いて、同時に極めて多数の遺伝子の発現レベルをモニタリングし、遺伝子変異体、突然変異及びSNP(一塩基多型)を同定することができる。このようなデータを用いて、遺伝子機能の解明や、症状及び疾患、または障害における遺伝子原理の解明や、症状及び疾患、障害の診断または治療、治療薬の開発、並びにこれらの治療薬の活性のモニタリングが可能である(例えば、Brennan他(1995) 米国特許第5,474,796号; Schena他(1996) Proc. Natl. Acad. Sci. 93:10614-10619; Baldeschweiler他(1995) PCT出願WO95/251116; Halon他(1995) PCT出願WO95/35505; Heller他(1997) Proc. Natl. Acad. Sci. 94:2150-2155; and Heller他(1997) 5,605,662を参照)。

40

#### 【0057】

ハイブリダイゼーションプローブはまた、天然のゲノム配列のマッピングに有用である。このプローブを、(1)特定の染色体、(2)染色体の特定の領域、(3)ヒト人工染色体(HAC)、酵母人工染色体(YAC)、細菌人工染色体(BAC)、細菌P1作製物、または単一の染色体DNAライブラリなどの人工染色体作製物にハイブリダイズさせることが可能である。

50

## 【0058】

## 発現

NIM1キナーゼをコードする多数の核酸分子をベクターにクローニングして、このタンパク質若しくはその一部を宿主細胞で発現させることができる。この核酸配列を、DNAシャフリング（米国特許第5,830,721号）や部位特異的変異誘発などの方法によって、新規の制限部位を作り出したり、グリコシル化パターンを変えたり、優先コドンを変えて特定の宿主における発現を増大させたり、スプライスバリエーションを作り出したり、半減期を延長する等の操作が可能である。この発現ベクターは、特定の宿主における各要素の効率に基づいて選択された様々なサンプルに由来する転写及び翻訳調節エレメント（プロモーター及びエンハンサー、特定の開始シグナル、ポリアデニル化3'配列）を含み得る。in vitro組換えDNA技術、合成技術及び/またはin vivo遺伝子組換え技術を組み合わせて、このベクターに核酸配列と調節エレメントをつなぐことができる。このような技術は、当分野で周知であり、Sambrook（前出、ch. 4, 8, 16 and 17）に記載されている。

10

## 【0059】

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

20

## 【0060】

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

30

## 【0061】

長期に渡って組換えタンパク質を産生させるために、同一或いは別のベクター上の選択マーカー遺伝子或いは可視マーカー遺伝子と共にこのベクターを持続的に細胞株に形質転換することができる。形質転換後、細胞を強化培地で約1~2日間増殖させてから選択培地に移す。選択マーカー、代謝拮抗物質、抗生物質、または除草剤耐性遺伝子は、関連する選択薬に対する抵抗性を与え、導入配列を確実に発現する細胞の増殖及び回収が可能となる。アントシアニン、緑色蛍光タンパク質（GFP）、グルクロニダーゼ、ルシフェラーゼなどの可視マーカーの発現によって同定された、或いは選択培地に生存することによって同定された耐性クローンを、培養技術を用いて増殖することができる。また、可視マーカーを用いて、導入された遺伝子によって発現するタンパク質を定量することができる。宿主細胞が目的の哺乳動物核酸分子を含むか否かの決定は、DNA-DNAまたはDNA-RNAハイブリダイゼーション、或いはPCR増幅技術に基づいて行うことができる。

40

## 【0062】

宿主細胞は、組換えタンパク質を目的の形に修飾する能力に基づいて選択することができる。このような修飾には、アセチル化及びカルボキシル化、グリコシル化、リン酸化、脂

50

質化、及びアシル化等が含まれる。「プレプロ」型を切断する翻訳後プロセッシングを利用して、タンパク質のターゲティング、折り畳み及び/または活性を特定することができる。翻訳後活性のための特定の細胞装置及び特徴的な機構を有するATCCから得られる様々な宿主細胞から、組換えタンパク質の適当な修飾及びプロセッシングが確実に行われるように好適な宿主細胞を選択することができる。

#### 【0063】

##### 細胞培地からのタンパク質の回収

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

10

#### 【0064】

##### ペプチドの化学合成

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

20

30

40

#### 【0065】

##### 抗体の準備及びスクリーニング

ヤギ及びウサギ、ラット、マウス、ヒト等を含む様々な宿主は、ヒトNIM1キナーゼ若しくはその任意の一部を注入して免疫することができる。フロイントなどのアジュバント及びミネラルゲルと、リゾレシチン及びpluronic polyol、ポリアニオン、ペプチド、油乳剤、キーホールリンペットヘモシニアン(KLH)、ジニトロフェノールなどの表面活性物質とを用いて免疫反応を高めることができる。オリゴペプチドやペプチド、またはタンパク質の一部を用いて、少なくとも約5個のアミノ酸、より好ましくは10個の天然のタンパク質の一部と同一のアミノ酸を含む抗体を誘発させる。キメラ分子

50

に対する抗体を産生させるために、オリゴヌクレオチドをKLHなどのタンパク質と融合させることができる。

【0066】

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

【0067】

別法では、当分野で周知の方法を用いる上記した一本鎖抗体を生産する技術で、エピトープ特異的一本鎖抗体を生産する。ヒトNIM1キナーゼのエピトープに対して特異的に結合する部位を含む抗体断片を生産することが可能である。限定するものではないが、このような断片には、例えば、抗体分子のペプシン消化によって作製されたF(ab')<sub>2</sub>断片及びこのF(ab')<sub>2</sub>断片のジスルフィド架橋を減少させて作製したFab断片が含まれる。別法では、Fab発現ライブラリを作製して、目的の特異性を有するモノクローナルFab断片の高速かつ容易に同定できるようにする(例えば、Huse他(1989) Science 246:1275-1281を参照)。

【0068】

ヒトNIM1キナーゼまたはその一部を用いて、ファージミドまたはBリンパ球免疫グロブリン・ライブラリをスクリーニングして、目的の特異性を有する抗体を同定する。確立された特異性を有するモノクローナル抗体或いはポリクローナル抗体のいずれか一方を用いる、競合的結合またはイムノアッセイの様々なプロトコルが当分野で周知である。このようなイムノアッセイは通常、このタンパク質とその特異的な抗体との複合体形成の測定を行う。2つの非干渉エピトープに反応するモノクローナル抗体を用いる2部位モノクローナル系イムノアッセイが好ましいが、競合的結合によるアッセイを用いることもできる(Pound(1998) Immunochemical Protocols, Humana Press, Totowa NJ)。

【0069】

アッセイのための分子の標識化

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

【0070】

(診断)

本核酸分子、断片、オリゴヌクレオチド、相補的なRNA及びDNA分子、及びPNAを用いて、遺伝子発現の変化、mRNAの過剰な発現の不在/存在の検出及び定量、または治療期間中のmRNAレベルのモニタリングを行うことができる。同様に、ヒトNIM1キナーゼに特異的に結合する抗体を用いて本タンパク質を定量することもできる。発現の変化に関連する疾患には、静坐不能、アルツハイマー病、健忘症、筋萎縮性側索硬化症、不安、遺伝性運動失調、脳性麻痺、痴呆、皮膚筋炎、ジストニー、ダウン症候群、癩癧、虚血性脳血管障害、小脳網膜血管芽腫(cerebelloretinal heman

g i o b l a s t o m a t o s i s )、ハンチントン病、細菌性及びウイルス性髄膜炎、多発性硬化症、筋ジストロフィー、重症筋無力症、脳腫瘍、神経線維腫症、パーキンソン病、ピック病、多発性筋炎、色素性網膜炎、分裂病、発作、並びに腺癌、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、及び子宮などの癌が含まれる。診断アッセイにハイブリダイゼーションまたは増幅技術を用いて、遺伝子発現の変化を検出するべく、患者からの生体サンプルの遺伝子発現レベルを標準的なサンプルの値と比較する。質的または量的なこのような比較法は当分野で周知である。

#### 【0071】

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

#### 【0072】

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

#### 【0073】

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

#### 【0074】

##### 免疫学的方法

特異的なポリクローナル抗体若しくはモノクローナル抗体の何れかを用いるタンパク質の検出及び定量は当分野で周知である。このような技術には、E L I S A (酵素結合免疫吸着検定法)及びラジオイムノアッセイ(R I A)、蛍光活性化セルソーター法(F A C S)が含まれる。2つの非干渉エピトープに反応するモノクローナル抗体を用いる2部位モノクローナル系イムノアッセイが好ましいが、競合的結合アッセイを用いることもできる(例えば、C o l i g a n 他(1997) C u r r e n t P r o t o c o l s i n I m m u n o l o g y, W i l e y - I n t e r s c i e n c e, N e w Y o r k N Y; 及び P o u n d 前出)。

#### 【0075】

##### (治療)

ヒトN I M 1キナーゼ(S E Q I D N O : 2)の所定の領域と図4 A - 図4 Cに示すキナーゼ(S E Q I D N O : 31 - 35)の所定の領域との間に、キナーゼ触媒ドメインの文脈における化学的及び構造的な類似性が存在する。加えて、遺伝子発現が、図2及び図3に示されているように、脳、前立腺、及び乳房、並びに癌に密接に関連する。ヒトN I M 1キナーゼは、脳及び神経系の疾患、特に脳腫瘍、発作、癲癇、ハンチントン病、並びに脳、子宮頸、結腸、肺、卵巣、及び前立腺の癌において或る役割を果たしていると思われる。脳腫瘍及び乳癌などの発現の大幅な上昇に関連する症状の治療においては、発現

10

20

30

40

50

またはタンパク質活性を低下させることが望ましい。

【0076】

一実施例では、発現若しくは活性の上昇に関連する疾患の治療または予防のために、ヒト N I M 1 キナーゼのインヒビター、アンタゴニスト、または抗体を投与し得る。発現の変化に関連する疾患には、静坐不能、アルツハイマー病、健忘症、筋萎縮性側索硬化症、不安、遺伝性運動失調、脳性麻痺、痴呆、皮膚筋炎、ジストニー、ダウン症候群、癩癩、虚血性脳血管障害、小脳網膜血管芽腫 (cerebelloretinal hemangioblastomatosis)、ハンチントン病、細菌性及びウイルス性髄膜炎、多発性硬化症、筋ジストロフィー、重症筋無力症、脳腫瘍、神経線維腫症、パーキンソン病、ピック病、多発性筋炎、色素性網膜炎、分裂病、発作、並びに腺癌、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、及び子宮などの癌が含まれる。

【0077】

別の実施例では、医薬用担体と共に N I M 1 キナーゼのインヒビター、アンタゴニスト、または抗体を含む医薬組成物を患者に投与して、限定するものではないが上記した疾患を含むこの内在性タンパク質の発現または活性の変化に関連する症状の治療または予防を行うことが可能である。

【0078】

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

【0079】

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

【0080】

核酸を用いる遺伝子発現の調節

遺伝子の発現は、N I M 1 キナーゼをコードする遺伝子の 5' または 3' 調節領域、或いは他の調節領域に対して相補的或いはアンチセンス分子を設計することで調節することが可能である。転写開始部位に対して設計されたオリゴヌクレオチドが好ましい。同様に、ポリメラーゼ、転写因子、または調節分子の結合を阻止する三重螺旋塩基対合で遺伝子発現を阻止することができる (Gee 他 In: Huber and Carr (1994) Molecular and Immunologic Approaches, Futura Publishing, Mt. Kisco NY, pp. 163-177)。また、リボソームと mRNA との結合を阻止して翻訳が行われないように、相補的な分子を設計することも可能である。或るいは、核酸分子またはその断片のライブラリをスクリーニングして、翻訳されない調節配列に特異的に結合する核酸分子または断片を同定することも可能である。

【0081】

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

10

20

30

40

50

## 【0082】

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

10

## 【0083】

スクリーニング及び精製アッセイ

ヒトNIM1キナーゼをコードする核酸分子を用いて、分子または化合物のライブラリをスクリーニングして特異的な結合親和性を調べることが可能である。このライブラリは、生物系において本核酸分子の活性、複製、転写、または翻訳を調節するアプタマー、DNA分子、RNA分子、PNA、ペプチド、転写因子などのタンパク質、エンハンサー、リプレッサー、及びその他のリガンドを含み得る。このアッセイは、特異的な結合が許容される条件下で、本核酸分子またはその断片を分子のライブラリと結合させるステップと、特異的な結合を検出して、一本鎖または二本鎖の本核酸分子と特異的に結合する少なくとも1つの分子を同定するステップとを含む。

20

## 【0084】

一実施例では、本発明のポリヌクレオチドを、単離され精製された分子または化合物のライブラリと共にインキュベートし、当分野で周知の、例えばゲル遅延アッセイ (gel retardation assay) (米国特許第6,010,849号) または網状赤血球溶解産物転写アッセイを用いて結合活性を測定することができる。別の実施例では、本ポリヌクレオチドをバイオブシー及び/または培養細胞及び培養組織に由来する核抽出物でインキュベートしてもよい。本ポリヌクレオチドとこの核抽出物における分子または化合物との間の特異的な結合を、まずゲルシフトアッセイ (gel shift assay) を用いて測定し、後にこれらの分子または化合物に対する抗体のレベルの上昇により確認することが可能である。これらの抗体をアッセイに加えると、ゲル遅延アッセイにおいてスーパーシフトが起こる。

30

## 【0085】

別の実施例では、本ポリヌクレオチドを用いて、当分野で周知のアフィニティクロマトグラフィ法を用いて分子または化合物を生成することが可能である。一実施例では、本ポリヌクレオチドを、ポリマー樹脂またはジェル上で臭化シアン基と化学的に反応させる。次に、サンプルを流して本ポリヌクレオチドと反応すなわち結合させる。本ポリヌクレオチドと結合した分子または化合物を、流す培養液の塩濃度を上昇させて本ポリヌクレオチドから遊離させて回収する。

## 【0086】

更なる実施例では、本タンパク質またはその一部を用いてサンプルからリガンドを精製することが可能である。哺乳動物タンパク質またはその一部を用いてリガンドを精製する方法は、特異的な結合が許容される条件下で、このタンパク質またはその一部をサンプルと結合させ、このタンパク質とリガンドとの間の特異的な結合を検出し、結合したタンパク質を回収し、適当なカオトロピック剤を用いてこのタンパク質を精製したりリガンドから分離することを含む。

40

## 【0087】

好適な実施例では、あらゆるスクリーニングアッセイにおいて、ヒトNIM1キナーゼまたはその一部を用いて、分子または化合物のライブラリをスクリーニングすることが可能である。このようなスクリーニングアッセイに用いるタンパク質の一部は、溶液中に遊離させるか、非生物基板または生物基板 (例えば細胞表面上) に固定させるか、または細胞

50

内に局在化させることができる。例えば、ある方法では、組替え核酸分子で安定的に形質転換され、ポリペプチドを発現してその細胞表面上にポリペプチドが存在する生存可能なわち安定した真核宿主細胞をスクリーニングアッセイに用いることができる。この細胞をリガンドのライブラリすなわち複数のライブラリに対してスクリーニングし、発現ポリペプチドとそのリガンドとの間の複合体の結合すなわち形成の特異性を測定することができる。このタンパク質と分子との間の特異的な結合を測定することができる。スクリーニングするライブラリの種類によって、このタンパク質と特異的に結合するDNA分子、RNA分子、ペプチド核酸、ペプチド、タンパク質、擬態、アゴニスト、アンタゴニスト、抗体、免疫グロブリン、インヒビター、及び薬剤またはその他のあらゆるリガンドをこのアッセイで同定することができる。

10

#### 【0088】

一実施態様では、本発明は、言及することをもって本明細書の一部とする米国特許5,876,946号に記載されているような極めて微量のアッセイ量及び極めて微量の検査化合物を用いるハイスループットのスクリーニング方法を提供する。この方法を用いて、特異的な結合により多数の分子及び化合物をスクリーニングすることができる。別の実施態様では、本発明はまた、競合的薬剤スクリーニングアッセイの方法を提供する。このアッセイでは、本ポリペプチドと結合可能な中和抗体が、本ポリペプチド、そのオリゴペプチド、或いはその一部と結合可能な検査化合物と特異的に競合する。スクリーニングアッセイによって同定された分子または化合物を用いて、哺乳動物モデル系を作成して、それらの毒性、診断、または治療の可能性を評価することができる。

20

#### 【0089】

##### 薬理学

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

#### 【0090】

治療効果のある薬用量とは、症状または病態を改善するタンパク質またはインヒビターの量である。このような薬剤の薬用効果及び毒性は、例えば、ED50（集団の50%に医薬的效果がある投与量）及びLD50（集団の50%に致命的である投与量）などの細胞培養または実験動物における標準的な製薬方法によって決定することができる。或る投与量における毒性効果と治療効果との比率が治療指数となり、LD50/ED50と示すことができる。高い治療指数を示す医薬組成物が好ましい。細胞培養アッセイ及び動物実験から得られたデータを用いて、ヒトへ適用する薬用量の範囲を決定する。

30

#### 【0091】

##### モデル系

動物モデルを用いて、ヒトの暴露に相当する暴露条件にして、動物モデルがヒトに類似の表現型反応を示すバイオアッセイを行うことができる。哺乳動物が最も一般的な動物モデルである。大抵の感染症、癌、薬剤、及び毒物の研究には、ラットやマウスなどの齧歯類が用いられる。これは、低コスト、入手の容易性、寿命、生殖能、及び参考文献の豊富さからである。齧歯類近交系及び非近交系は、目的の遺伝子の過剰或いは過少な発現の生理学的な原因を調査するのに有用なモデルであり、疾患の診断及び治療方法の開発にも有用である。特定の遺伝子を大量に発現する（例えば、乳汁中に分泌される）同系哺乳動物は、その遺伝子によって発現されるタンパク質の便利な供給源となり得る。

40

#### 【0092】

##### 中毒学

中毒学とは、生物系における物質の影響を研究する学問である。殆どの毒物研究はラットまたはマウスを用いて行われる。ラットまたはマウスの生理機能、行動、恒常性プロセス、及び致死率における質的及び量的変化を観察して、毒性プロファイルを作成し、その物

50

質に曝露された後に生じるであろうヒトの健康状態を調べる。

【0093】

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

10

【0094】

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

【0095】

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

20

【0096】

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

【0097】

遺伝子組換え動物モデル

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

30

【0098】

胚性幹細胞

齧歯類胚から単離された胚性幹細胞(ES細胞)は、胚組織を形成する可能性を維持している。ES細胞がキャリアとなる胚の中に導入されると、正常な発生が再開され、生まれる動物の組織の一部を担うことになる。ES細胞は、実験用のノックアウト及びノックイン齧歯類系を作製するのに好適な細胞である。マウス129/SvJ細胞株などのマウスES細胞は、マウスの初期胚から採取されてから当分野で周知の培養条件下で増殖されたものである。遺伝子組換え系の作製に用いるベクターは、疾患候補遺伝子及びマーカー遺伝子配列を含み、このマーカー遺伝子により導入された疾患遺伝子の存在を確認することができる。このベクターを、当分野で周知の方法でES細胞に形質転換され、この形質転換されたES細胞を特定し、C57BL/6マウス株などからのマウス細胞胚盤胞内に微量注入する。この胚盤胞を複数の偽妊娠メスに外科的に導入して、生まれてくるキメラ子孫のそれぞれがその遺伝子型を有するようにして、それらを交配してヘテロ接合系またはホモ接合系を作り出す。

40

【0099】

ヒト胚盤胞に由来するES細胞を、in vitroで操作して8個の異なった細胞系譜

50

に分化することが可能である。これらの細胞系譜を用いて、*i n v i t r o*で、様々な細胞型及び組織の分化を研究する。細胞型には、神経細胞、造血系、及び心筋細胞に分化する内胚葉、中胚葉、外胚葉の細胞型が含まれる。

【0100】

#### ノックアウト分析

遺伝子ノックアウト分析では、哺乳動物遺伝子のある領域が、ネオマイシン・ホスホトランスフェラーゼ遺伝子などの非哺乳動物遺伝子を含むように酵素によって改変される (neo; Capecchi (1989) Science 244:1288-1292)。改変された遺伝子が培養ES細胞に形質転換され、相同組換えにより内在性のゲノムに組み込まれる。挿入された配列は、内在性遺伝子の転写及び翻訳を阻害する。この形質転換細胞を齧歯類胚に注入し、この胚を偽妊娠メスに移植する。この遺伝子組換え子孫をクロス交配して、この哺乳動物遺伝子の機能的な複製物が存在しないホモ接合近交系を作り出す。

10

【0101】

#### ノックイン分析

ES細胞を用いてノックインヒト化動物(ブタ)または遺伝子組換えヒト疾患動物モデル(マウスまたはラット)を作り出すことができる。ノックイン技術を用いて、ヒト遺伝子のある領域を動物ES細胞に注入し、そのヒト配列が動物細胞のゲノムの中に組み込まれるようにする。形質転換ES細胞を胚に注入し、この胚を上記したように注入する。類似のヒトの症状の治療についての情報を収集するべく、遺伝子組換え子孫即ち近交系を有望な医薬品で処置して観察する。これらの方法を用いて、幾つかのヒト疾患モデルを作り出す。

20

【0102】

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

【0103】

#### 【実施例】

本発明は、記載した特定の装置及び物質、方法に限定されるものではないことを理解されたい。特定の実施例について説明するが、同等の実施例を用いて本発明を具現することも可能である。本発明の範囲は、前記請求の範囲によってのみ限定されるものであって、記載した実施例によって限定されるものではない。また、以下に記載の実施例は、本発明を例示するためのものであって本発明を限定するものではない。例示目的で、ヒト小脳ライブラリ(CRBLNOT01)、前立腺ライブラリ(PROSBPT03)、及び標準化された脳ライブラリ(BRAINON01)の作製方法を記載する。

30

【0104】

#### 1 cDNAライブラリの作製

##### 小脳

小脳ライブラリの作成に用いた組織は69歳の白人男性から得た(RT95-05-0301; International Institute for Advanced Medicine, Exton PA)。この凍結組織を、POLYTRONホモジナイザー(PT-3000; Brinkmann Instruments, Westbury NJ)を用いてホモジナイズして溶解した。試薬及び抽出方法は、RNA単離キット(Stratagene)に含まれているものを用いた。この溶解物を、L8-70M超遠心分離器(Beckman Coulter, Fullerton CA)において、SW28ローターを用いて5.7M CsClクッションに対して室温で18時間、毎分25,000回転で遠心分離を行った。RNAをフェノールクロロホルム、pH8.0で2回、酢酸フェノール、pH4.0で1回抽出し、0.3M酢酸ナトリウム及び2.5倍量のエタノールを用いて沈殿させ、水中で再懸濁し、DNアーゼで37で15分間処置

40

50

した。このRNAをOLIGOTEXキット(Qiagen, Chatsworth CA)で単離し、cDNAライブラリの作製のために用いた。

【0105】

#### 前立腺

59歳の白人男性から根治前立腺切除の際に採取した病変前立腺組織を用いてPROSBPT03ライブラリを作製した。病理学的には、この腫瘍は、Gleasonグレード3+3の腺癌であって、左右の周辺に及ぶ微視的な病巣を伴っていた。この腫瘍は限定的で胸膜には及んでいなかった。進行した前立腺上皮内癌が、右側の周囲で確認された。この患者には、前立腺特異的抗原(PSA)の上昇が見られた。家族歴には、両親の脳血管疾患、並びに兄弟の前立腺癌が含まれていた。

10

【0106】

cDNAライブラリを作製するために、SUPERSCRIPTプラスミドシステム(Life Technologies)の推奨プロトコルに従って2.4μgのポリARN Aを用いた。一本鎖cDNAの合成には、オリゴd(T)プライミングを用い、二本鎖の合成には、DNAポリメラーゼI、大腸菌リガーゼ、及びRNアーゼHを組み合わせて用いた。このcDNAをT4ポリメラーゼで平滑化し、このcDNAの平滑末端にSal Iリンカーを付加した。Sal Iが付加された二本鎖cDNAを、Not Iで消化し、SEPHAROSE CL4B カラム(Amersham Pharmacia Biotech)で分画した。

【0107】

400bpを超えるこれらの小脳cDNAをpSPORT Iプラスミドに繋ぎ、このプラスミドをDH5 コンピテント細胞(Life Technologies)に形質転換した。400bpを超えるこれらの前立腺cDNAをpINCYプラスミド(Incyte Pharmaceuticals)のNot I部位及びEcoRI部位に繋ぎ、このプラスミドをDH5 またはELECTROMAX DH10Bコンピテント細胞(Life Technologies)に形質転換した。

20

【0108】

#### 脳ライブラリの標準化

例を示すために、ヒト脳ライブラリ(BRAINON01)の標準化について説明する。大腸菌株DH12Sコンピテント細胞(Life Technologies)におけるBRAINOT03プラスミドライブラリの約 $4.9 \times 10^6$ の独立したクローンを、カルベニシリン(25mg/l)及びメチシリン(1mg/ml)選択下で、培地で成長させ、その後に電気穿孔法により形質転換を行った。ライブラリの存在レベルに従って余剰のcDNA複製物の数を減らすために、cDNAライブラリを以下の変更点を除きSoares他(1994, Proc Natl Acad Sci 91: 9228-9232)の方法に従って1回標準化した。変更点は次の通りである。プライマー伸長反応における鑄型に対するプライマーの比率を2:1から10:1に高めた。反応液のdNTP濃度を、より長いプライマー伸長物(400~1000ヌクレオチド)が形成できるように各dNTPを150μMまで低下させた。アニーリングハイブリダイゼーションを13時間から48時間に延長した。標準化したライブラリの一本鎖DNAサークルを、ヒドロキシアパタイトクロマトグラフィーで精製し、ランダムプライミングにより部分的な二本鎖に変換し、大腸菌株DH10Bコンピテント細胞(Life Technologies)の中に電気穿孔法により導入した。

30

40

【0109】

#### 2 pINCYプラスミドの作製

このプラスミドは、pSPORT1プラスミド(Life Technologies)をEcoRI制限酵素(New England Biolabs, Beverly MA)で消化し、オーバーハングした端部を、クレノウ酵素(New England Biolabs)及び2'-デオキシヌクレオチド5'-三リン酸(dNTPs)を用いて二本鎖に合成した。このプラスミドが自己連結した後、このプラスミドで細菌宿主となる大腸

50

菌株 JM109 を形質転換した。

【0110】

細菌 (pSPORT 1 - RI) によって生成された中間プラスミドは、EcoRI では消化されず、Hind III (New England Biolabs) で消化し、オーバーハングした端部をクレノウ酵素及び dNTPs で二本鎖にした。リンカー配列をリン酸化し、5' 平滑末端に繋ぎ、EcoRI で消化し、自己連結させた。JM109 宿主細胞を形質転換した後、プラスミドを単離し、Hind III ではなく EcoRI で選択的な消化の検査を行った。この基準を満たす 1 つのコロニーを pINCY プラスミドと呼ぶ。

【0111】

NotI 及び EcoRI 制限酵素を用いて作製したライブラリから cDNA を取り込むプラスミドの能力を検査した後、いくつかのクローンをシーケンシングし、約 0.8 kb のインサートを含む 1 つのクローンを選択し、そのクローンから大量のプラスミドを生成した。NotI 及び EcoRI で消化した後、ライブラリの作製に用いるためにこのプラスミドをアガロースゲル上で分画し、QIAQUICK カラム (Qiagen) を用いて精製した。

【0112】

3 cDNA クローンの単離及びシーケンシング

プラスミド DNA を細胞から遊離し、MINIPREP キット (Edge Biosystems, Gaithersburg MD) または REAL Prep 96 プラスミドキット (Qiagen) のいずれかを用いて精製した。このキットには、96 ウェルブロック、及び 960 回の精製を行うための試薬が含まれている。以下の変更点を除き、推奨プロトコルに従った。(1) 細菌を、25 mg/l のカルベニシリン及び 2.4% のグリセロールと共に 1 ml の滅菌 Terrific Broth (Life Technologies) において培養した。(2) 接種後、細胞を 19 時間培養してから、0.3 ml の溶解バッファで溶解した。(3) イソプロパノール沈殿の後、プラスミド DNA のペレットを 0.1 ml の蒸留水で再懸濁した。プロトコルの最終ステップの後、サンプルを 96 ウェルブロックに移し、4 で保管した。

【0113】

MICROLAB 2200 システム (Hamilton, Reno NV) 及び DNA ENGINE サーマルサイクラー (MJ Research) を用いて、シーケンシングするために cDNA を調整した。この cDNA を、ABI PRISM 377 シーケンシングシステム (PE Biosystems) または MEGABASE 1000 DNA シーケンシングシステム (Amersham Pharmacia Biotech) を用いて Sanger and Coulson (1975; J Mol Biol 94: 441-448) の方法でシーケンシングした。単離したもののほとんどを、標準的な ABI プロトコル及びキット (PE Biosystems) に従って、0.25 x ~ 1.0 x の濃度の溶液容量でシーケンシングした。別法では、Amersham Pharmacia Biotech から入手した溶液及び色素を用いて cDNA をシーケンシングした。

【0114】

4 cDNA 配列の伸長

本核酸分子は、cDNA クローン及びオリゴヌクレオチドプライマーを用いて伸長した。一方のプライマーは既知の断片の 5' の伸長を開始するために合成し、他方のプライマーは既知の断片の 3' の伸長を開始するために合成した。開始プライマーは、OLIGO 4.06 ソフトウェア (National Biosciences, Plymouth MN) 若しくは別の好適なプログラムを用いて、約 22 ~ 約 30 個のヌクレオチドの長さ、約 50% 以上の GC 含量で、685 ~ 72 の温度で標的配列にアニールするように設計した。ヘアピン構造及びプライマー-プライマー二量体を形成するヌクレオチドのストレッチは排除した。

10

20

30

40

50

## 【0115】

選択されたcDNAライブラリを鋳型として用いてこの配列を伸長した。2段階以上の伸長が必要な場合は、追加の或いは入れ子状のプライマー(nested primer)を設計した。好適なライブラリは、大きなcDNAを含むように大きさが選択されたライブラリである。また、ランダムプライミングしたライブラリも、遺伝子の5'及び上流領域を備えた多くの配列を含むため好適である。ランダムプライミングライブラリは、オリゴd(T)ライブラリが完全長のcDNAを形成しなかった場合に特に有用である。ゲノムライブラリは、調節エレメントを得るためにプロモーター結合領域の5'を伸長する際に有用である。

## 【0116】

米国特許第5,932,451号に開示されているような方法を用いてPCR法で高い忠実度の増幅を達成した。PCRは、DNA ENGINE サーマルサイクラー(MJ Research)を用いて96ウェルプレートで行った。反応混合液には、鋳型DNA及び200 nmolの各プライマー、反応緩衝液( $Mg^{2+}$ 、 $(NH_4)_2SO_4$ 、 $\beta$ -メルカプトエタノールを含む)、Taq DNAポリメラーゼ(Amersham Pharmacia Biotech)、ELONGASE酵素(Life Technologies)、Pfu DNAポリメラーゼ(Stratagene)を含まれている。プライマーの組、PCI A及びPCI B(Incyte Pharmaceuticals)に対して以下のパラメーターで増幅を行った。

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

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

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

## 【0117】

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

## 【0118】

伸長したヌクレオチド配列を脱塩及び濃縮してから384ウェルプレートに移し、CviJIコレラウィルスエンドヌクレアーゼ(Molecular Biology Research, Madison WI)で消化し、pUC18ベクター(Amersham Pharmacia Biotech)に再連結する前に音波処理または切断した。ショットガンシーケンシングのために、消化したヌクレオチド配列を低濃度(0.6%~0

10

20

30

40

50

．8% ) のアガロースゲル上に分離させ、断片を切断し、ゲルを A G A R A C E 酵素 ( P r o m e g a ) で消化した。T 4 D N A リガーゼ ( N e w E n g l a n d B i o l a b s ) を用いて伸長したクローンを p U C 1 8 ベクター ( A m e r s h a m P h a r m a c i a B i o t e c h ) に再連結し、P f u D N A ポリメラーゼ ( S t r a t a g e n e ) で制限部位の延び出しを処理してから、大腸菌コンピテント細胞に形質転換した。形質転換細胞が抗生物質を含む培地で選択され、それぞれのコロニーを切りとって、L B / 2 X カルベニシリン培養液の入った 3 8 4 ウェルプレートの中で、3 7 で一晩培養した。

#### 【0119】

この細胞を溶解して、T a q D N A ポリメラーゼ ( A m e r s h a m P h a r m a c i a B i o t e c h ) 及び P f u D N A 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 D N A を増幅した。 10

ステップ 1 9 4 で 3 分間

ステップ 2 9 4 で 1 5 秒

ステップ 3 6 0 で 1 分間

ステップ 4 7 2 で 2 分間

ステップ 5 ステップ 2、3、及び 4 を 2 9 回繰り返す

ステップ 6 7 2 で 5 分間

ステップ 7 4 で 保管。

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 D N A を定量した。D N A 回収率の悪いサンプルは、上記した条件で再び増幅した。サンプルを 2 0 % のジメチルサルホサイド ( d i m e t h y s u l p h o x i d e ) ( 1 : 2 , v / v ) で希釈し、D Y E N A M I C エネルギー移動シークエンシングプライマー及び D Y E N A M I C D I R E C T サイクルシークエンシングキット ( 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 キット ( P E B i o s y s t e m s ) を用いてシークエンシングした。 20

#### 【0120】

##### 5 c D N A クローン及びそれらから導き出されたポリペプチドの相同性検索

配列表の核酸分子及びそれらから導き出されたアミノ酸配列を、G e n B a n k、S w i s s P r o t、B L O C K S などのデータベースにおいて検索する。すでに同定されアノテーションの付けられた配列またはドメインを含むこれらのデータベースを、B L A S T または B L A S T 2 ( A l t s c h u l ら、前出； A l t s c h u l、前出) を用いて検索し、アラインメントを作成し、完全に一致する或いは相同な配列を決定する。このアラインメントは、原核生物 ( 細菌 ) または真核生物 ( 動物、真菌、または植物 ) を起源とする配列に対するものである。或いは、S m i t h a n d S m i t h ( 1 9 9 2 , P r o t e i n . E n g i n e e r i n g 5 : 3 5 - 5 1 ) に記載されているようなアルゴリズムを用いて、一次配列パターン及び二次構造ギャップペナルティを処理することが可能である。本明細書に記載した全ての配列は、その長さが少なくとも 4 9 ヌクレオチドであり、不要な塩基 ( A、C、G、或いは T ではなく N と記載 ) は 1 2 % 以下である 40

#### 【0121】

K a r l i n ( 前出 ) に詳述されているように、問い合わせ配列とデータベース配列との間の B L A S T による一致を統計学的に評価し、ヌクレオチドに対しては  $10^{-2.5}$ 、ペプチドに対しては  $10^{-1.4}$  の閾値を満たす一致のみを記録した。相同性はまた、以下のように計算した積スコアによって評価した。まず、B L A S T における核酸またはアミノ酸同一性 ( 問い合わせ配列と参照配列との間の ) のパーセントに、最大可能 B L A S T スコアのパーセント ( 問い合わせ配列及び参照配列の長さに基づく ) を乗じてから、1 0 0 で除した。実験室で用いられるハイブリダイゼーション法と比べ、完全一致に対する電子的なストリンジェンシーを 7 0 と設定し、完全一致に対する保守的な下限を約 4 0 に設 50

定した（不要な塩基対による1～2%のエラーを有する）。

【0122】

無料で利用可能な配列比較アルゴリズムであるBLASTソフトウェア一式（NCBI, Bethesda MD; <http://www.ncbi.nlm.nih.gov/gorf/bl2.html>）は、既知の核酸分子を整列するために用いる“blastn”、及び核酸分子或いはアミノ酸分子のいずれかの直接ペアワイズ比較に用いられるBLAST2を含む様々な配列分析プログラムを含む。BLASTプログラムは、一般に、例えば以下のようにデフォルト設定されたギャップ及びその他のパラメーターで実行される。

【0123】

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

10

同一性或いは類似性を、配列の全長或いは配列の一部に対して測定することができる。Brennerらが（1998; Proc Natl Acad Sci 95: 6073-6078、言及することをもって本明細書の一部とする）、配列同一性により構造の相同性を同定するBLASTの能力を分析した。この分析により、少なくとも150残基の配列アラインメントに対して30%の同一性が信頼できる閾値であり、少なくとも70残基の配列アラインメントに対して40%の同一性が信頼できる閾値である。

20

【0124】

本明細書の哺乳動物核酸分子を、LIFSEQ GOLDデータベースで見出した構築されたコンセンサス配列または鋳型と比較した。cDNA、伸長、完全長、及びショットガンシーケンシングプロジェクトに由来するコンポーネント配列をPHREDで分析し、質のスコアを割り当てた。許容できる質スコアを有する全ての配列に対して様々なプリプロセッシング及び編集を行い、質の低い3'末端、ベクター及びリンカー配列、ポリA尾部、Aluリピート、ミトコンドリア及びリボソーム配列、及び細菌汚染配列を除去した。編集した配列は少なくともその長さが50bpでなければならない。情報の少ない配列、並びにジヌクレオチドリピート、Aluリピートなどの繰り返しエレメントはNで置換するかマスクした。

30

【0125】

編集した配列を、配列が遺伝子ピンに割り当てられる構築処理を行った。それぞれの配列は1つのピンにのみ属し、鋳型を形成するべく各ピンにおける配列を構築した。新規にシーケンシングしたコンポーネントをBLAST及びCROSSMATCHを用いて存在するピンに加えた。ピンに加えるためにはこのコンポーネント配列は、BLAST質スコアが150以上であって、少なくとも82%の局所的な同一性のアラインメントでなければならない。それぞれのピンにおける配列を、PHRAPを用いて構築した。いくつかの重複するコンポーネント配列を有するピンはDEEP PHRAPを用いて構築した。それぞれの鋳型の向きは、そのコンポーネント配列の数及び向きに基づいて決定した。

40

【0126】

それぞれのピンを互いに比較して、局所類似性が82%以上のピンを1つにまとめて再構築した。局所同一性が95%未満の鋳型を有するピンは分けた。鋳型は、STITCHER/EXON MAPPERアルゴリズムで分析した。このSTITCHER/EXON MAPPERアルゴリズムは、スプライスバリエーション、択一的スプライシングエキソン、スプライスジャンクション、並びに組織の種類或いは疾患の状態により発現が異なる択一的にスプライシングされた遺伝子などの存在の確率を分析することができる。構築処理を

50

繰り返し行い、GBpriなどのGenBankデータベースに対してBLASTを用いて鋳型にアノテーションを付けた。完全一致は、200塩基対に対する95%以上の局所同一性から100塩基対に対する100%の局所同一性を有すると定義し、相一致は、 $1 \times 10^{-8}$ のE値(確率スコア)を有すると定義した。また鋳型を、GENPEPTに対してフレームシフトFASTx分析を行い、相一致を $1 \times 10^{-8}$ のE値を有すると定義した。鋳型の分析及び構築については、1999年3月25日に出願された米国特許出願第09/276,534号に開示されている。

#### 【0127】

構築の後に、鋳型をBLAST、モチーフ、及びその他の機能分析法で分析し、1997年3月6日に出願の米国特許出願第08/812,290号及び同第08/811,758号、1997年10月9日に出願の米国特許出願第08/947,845号、及び1998年3月4日に出願の米国特許出願第09/034,807号に記載されている方法を用いてタンパク質を階層に分類した。次に鋳型を、3つ全ての前方読み枠にそれぞれの鋳型を翻訳して、それぞれの翻訳を、HMMERソフトウェアパッケージ(Washington University School of Medicine, St. Louis MO; <http://pfam.wustl.edu/>)を用いて隠れマルコフモデルをベースにしたタンパク質ファミリー及びドメインのPFAMデータベースにおいて検索した。

10

#### 【0128】

本核酸分子を、MACDNASIS PROソフトウェア(Hitachi Software Engineering)及びLASERGENEソフトウェア(DNA STAR)を用いて更に分析し、GenBankの齧歯類、哺乳類、脊椎動物、原核生物、及び真核生物のデータベース、SwissProt、BLOCKS、PRINTS、PFAM、及びPrositeなどの公共のデータベースに対して問い合わせた。

20

#### 【0129】

### 6 染色体マッピング

Stanford Human Genome Center (SHGC)、Whitehead Institute for Genome Research (WIGR)、及びGenethonなどの公共の情報源から入手可能な放射線ハイブリッド及び遺伝子マッピングデータを用いて、マッピングできる配列表の核酸分子が存在するかを調べた。NIM1キナーゼをコードする核酸分子のマッピングされた全ての断片が、同じ位置にマッピングされた全ての関連する調節及びコード配列に割り当てられた。遺伝子マップ位置は、ヒト染色体の範囲すなわち区間として表される。センチモルガン(cM)(ヒトDNAの100万塩基対に概ね相当する)で表されるマップ位置の範囲は、染色体p腕の末端から測定する。

30

#### 【0130】

### 7 ハイブリダイゼーション技術及び分析

#### 核酸分子の基板への固定

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

40

50

## 【0131】

第2の方法では、インサートに隣接したベクター配列に相補的なプライマーを用いて、PCRを30サイクル行い細菌ベクターから核酸分子を増幅する。PCR増幅により、拡散の濃度を開始時の1~2 ngから最終的に5 µgまで増大させる。約400 bp~約5000 bpの増幅した拡散をSEPHACRYL-400ビーズ(Amersham Pharmacia Biotech)を用いて精製する。精製した拡散を、手動で或いはドット/スロットプロットングマニフォールド及び吸入装置でナイロン膜に配列し、上記した変性、中和、及びUV照射によって固定する。精製した拡散を、米国特許第5,807,522号に開示された方法でポリマーコートスライドガラスに機械的に配列して固定する。ポリマーコートスライドガラスは、顕微鏡スライドガラス(Corning, Acton MA)を0.1% SDS及びアセトンで超音波洗浄した後、4%フッ化水素酸(VWR Scientific Products, West Chester PA)においてエッチングし、次に95%エタノールにおいて0.05%アミノプロピルシランでコーティングし、オープンで110 で硬化させて準備する。処理中及び処理後にこのスライドガラスを蒸留水で十分に洗浄する。次に、核酸分子をスライド上に配列し、STRATALINKER UVクロスリンカー(Stratagene)を用いてこの核酸分子のアレイをUV照射してスライド上に固定する。次にこのアレイを、室温で0.2% SDSで洗浄し、蒸留水で3回すすぐ。非特異的な接合部位を、アレイをリン酸バッファ(PBS:Tropix, Bedford MA)に於ける0.2%カゼインにおいて60 で30分間インキュベートしてブロックし、次にこのアレイを0.2% SDSで洗浄し、前記したように蒸留水で洗い流す。

10

20

## 【0132】

メンブレンを用いるハイブリダイゼーションのためのプローブの準備

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

30

## 【0133】

ポリマーコートスライドを用いるハイブリダイゼーションのためのプローブの準備

サンプルから単離したmRNAに由来するハイブリダイゼーションプローブを用いて、アレイハイブリダイゼーションにおいて本配列表の核酸分子をスクリーニングする。プローブは、GEMbrightキット(Incyte Pharmaceuticals)を用いて、mRNAを9 µl TEバッファにおいて200 ngの濃度に希釈し、5 µlの5xバッファ、1 µlの0.1 M DTT、3 µlのCy3またはCy5標識混合物、1 µlのRNアーゼインヒビター、1 µlの逆転写酵素、及び5 µlの1x酵母コントロールmRNAを加えて準備する。酵母コントロールmRNAは、非コード酵母ゲノムDNA(W. Lei, unpublished)から*in vitro*転写によって合成する。量的コントロールとして、0.002 ng、0.02 ng、0.2 ng、及び2 ngのコントロールmRNAからなるセットを、サンプルmRNAに対してそれぞれ1:100, 000、1:10,000、1:1000、及び1:100 (w/w)の比率で逆転写反応液に希釈する。mRNAの異なった発現パターンを検査するために、第2のセットのコントロールmRNAを、1:3、3:1、1:10、10:1、1:25、及び

40

50

25 : 1 (w/w) の比率で逆転写反応液に希釈する。反応液を混合し、37 で2時間インキュベートする。次に反応液を85 で20分間インキュベートし、2つの連続したCHROMA SPIN+TE 30カラム(Clontech)を用いてプローブを精製する。精製したプローブを、DEPC処理水で90 µlに希釈して、2 µlの1 mg/mlグリコーゲン、60 µlの5 M酢酸ナトリウム、300 µlの100%エタノールを加えてエタノール沈殿させる。このプローブを20,800 × gにおいて20分間遠心分離し、ペレットを12 µlの再懸濁バッファに再懸濁させ、65 で5分間加熱し、完全に混合する。上記したようにプローブを加熱及び混合してから氷上で保管する。プローブを以下に記載するように高密度のアレイハイブリダイゼーションに用いる。

#### 【0134】

##### メンブレンを用いるハイブリダイゼーション

メンブレンを、1% Sarkosyl及び1×高リン酸バッファ(0.5 M NaCl、0.1 M Na<sub>2</sub>HPO<sub>4</sub>、5 mM EDTA、pH 7)を含むハイブリダイゼーション溶液において55 で2時間プレハイブリダイゼーションする。15 mlの新しいハイブリダイゼーション溶液に希釈したプローブをメンブレンに加える。55 で16時間、メンブレンにプローブをハイブリダイズさせる。ハイブリダイゼーションの後、メンブレンを1 mM Tris (pH 8.0)、1% Sarkosylにおいて25 で15分間洗浄し、1 mM Tris (pH 8.0)において25 で15分間づつ4回洗浄する。ハイブリダイゼーション複合体を検出するために、XOMAT-ARフィルム(Eastman Kodak, Rochester NY)をメンブレンに一晚-70 で露出し、現像して目で確認する。

#### 【0135】

##### ポリマーコートスライドを用いるハイブリダイゼーション

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

#### 【0136】

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

#### 【0137】

ハイブリダイゼーション複合体をInnova 70混合ガス10Wレーザーを備えた顕微鏡(Coherent, Santa Clara CA)で検出する。この顕微鏡は、Cy3を励起するために488 nmのスペクトル線が発生し、Cy5を励起するために632 nmのスペクトル線が発生することができる。励起レーザー光を20倍の顕微鏡対物レンズ(Nikon, Melville NY)を用いてアレイに集束させる。アレイを

10

20

30

40

50

含むスライドを顕微鏡上のコンピュータ制御されたXYステージ上に載せ、20 μmの解像度で対物レンズを通してラスタースキャンする。ディフュージョン方式の場合、2つの蛍光物質を順番にレーザーで励起する。波長に応じて放出された光を2つの蛍光物質に対応する2つの光電子増倍管検出器(PMT R1477, Hamamatsu Photonics Systems, Bridgewater NJ)に分割する。アレイと光電子増倍管との間に配置された好適なフィルターを用いてシグナルをフィルタリングする。用いられる蛍光物質の最大発光波長は、Cy3が565 nmであり、Cy5が650 nmである。スキャンの感度は、プローブ混合物に加えられた酵母コントロールmRNAによって生成されるシグナル強度を用いて較正する。アレイのある位置に相補的なDNA配列が含まれていると、その位置におけるシグナルの強度が、ハイブリダイズ種(hybridizing species)の重量比で1:100,000の関係となり得る。

#### 【0138】

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

#### 【0139】

### 8 ノーザン分析

#### 電子的

BLASTに適用するコンピュータ技術を用いて、GenBankまたはLIFESEQデータベース(Incyte Pharmaceuticals)などのヌクレオチドデータベースにおいて同一或いは関連する分子を検索した。ヒト及びラットの配列に対する積スコアを以下のように計算した。BLASTスコアをヌクレオチド同一性のパーセントで乗じて、その値を2つの配列の内の短い方の長さの5倍で除した。したがって、短い方の配列の長さに対する100%のアラインメントが積スコア100となる。積スコアは、2つの配列間の類似性の程度及び一致する配列の長さを考慮している。例えば、積スコアが40では、一致は正確でエラーは1%~2%であり、70以上の積スコアではその一致は完全である。類似分子すなわち関連分子は、通常は積スコアが8~40の範囲の分子を選択することによって同定することができる。

#### 【0140】

70の積スコアで実施した電子的ノーザン分析の結果が図2に示されている。この分析では、cDNAライブラリを系、器官/組織、及び細胞の種類によって幾つかのカテゴリーに分類した。そのカテゴリーには、心血管系、結合組織(特に癌性の乳房線維芽細胞)、消化系、胚、内分泌系、外分泌腺、女性及び男性生殖器(特に癌性の前立腺)、生殖細胞、血液及び免疫系、肝臓、筋骨格系、神経系、膵臓、呼吸器系、感覚器官、皮膚、顎口腔系、未分類/混合、及び尿管が含まれる。各カテゴリーにおいて、実際に本配列を数をカウントし、そのカテゴリーの合計ライブラリ数に対する発現したライブラリ数を示した。

#### 【0141】

#### 定量PCR

定量PCRを用いて、様々な細胞系及び組織におけるヒトNIM1キナーゼの発現を調べた。ATCCなどから得られる細胞系には、H460ヒト非小細胞肺癌、A2780ヒト

卵巣癌系、A375ヒトメラノーマ細胞系、HDFヒト上皮繊維芽細胞、HELAヒト子宮頸癌、DU145アンドロゲン誘導性前立腺癌細胞系、MDA-MB231ヒト乳癌細胞、U87グリア芽腫腫瘍細胞、及びBX-PC3膵臓癌細胞がある。これらの細胞系を培養皿に載置し、10%のウシ胎児血清、及び5%CO<sub>2</sub>における2mMグルタミンを90%のコンフルエントになるまで加えたRPMIにおいて成長させた。組織、脳、結腸、子宮、及び胎盤をClontech社のMTNプロット(Human I、II、III、及びIV)として得た。

#### 【0142】

ノーザン分析は、ABI PRISM 7700シーケンシングシステム及びTAQMANアッセイ試薬(TAQMAN Universal PCR Master mix)を用い、製造者による取扱説明書に従って定量PCRにより行った(すべてPE Biosystems)。全ての反応を3通り行った。

10

#### 【0143】

反応に用いたプライマーには、SEQ ID NO: 37-39が含まれる。相対的な定量は、標準として18s RNAを用いて行った。この標準化方法の直線性は、Spies and Ivel (1999; Biotechniques 26: 46-50、言及することをもって本明細書の一部とする)に記載されている。

#### 【0144】

### 9 相補的な核酸分子

本核酸分子に相補的な配列或いはその断片を用いて、遺伝子の発現を検出したり、低下させたり、阻害することができる。約15~約30個の塩基からなるオリゴヌクレオチドの使用について記載するが、それより小さい或いは大きい配列の断片、またはその誘導體(PNA)の場合でも同じ方法を用いることができる。好適なオリゴヌクレオチドは、Oligo4.06ソフトウェア(National Biosciences)を用いて選択した。プロモーターの結合を阻害して転写を阻止するために、最も好ましくはオープンリーディングフレームの開始コドンの前の約10個のヌクレオチドである最もユニークな5'配列に結合するように相補的なヌクレオチドを設計する。翻訳を阻害するために、本哺乳動物ポリペプチドをコードするmRNAへのリボソームの結合を阻害するべく相補的なオリゴヌクレオチドを設計する。

20

#### 【0145】

転写または翻訳を阻害するために作製したアンチセンス分子を用いるのに加えて、ゲノム配列(例えばエンハンサーやイントロン)或いはトランス作動性調節性遺伝子に対するアンチセンス分子を設計して遺伝子発現を変化させることができる。同様に、三重らせん塩基対形成として知られているHogboom塩基対形成法を用いてアンチセンス阻害を行うことができる。三重らせんの塩基対形成に関与するアンチセンス分子は、ポリメラーゼ、転写因子、または調節分子と結合するために二本鎖の間を広げる二重らせんの能力を損なわせる。

30

#### 【0146】

このようなアンチセンス分子を発現ベクターに導入して、好適な細胞または組織を形質転換する。これには、効果を検査するために細胞系への発現ベクターの導入、一過性或いは短期の治療のための器官、腫瘍、滑液キャビティ、または血管系への発現ベクターの導入、または長期或いは安定した遺伝子治療のための単細胞または他の生殖系への発現ベクターの導入が含まれる。一過性の発現は、ベクターの複製を伴わない場合には1ヶ月以上持続し、ベクターの複製を引き起こす好適なエレメントが形質転換/発現系に用いられている場合には3ヶ月以上持続する。

40

#### 【0147】

アンチセンス分子をコードするベクターで好適な分裂細胞を安定的に形質転換することにより、遺伝子組換え細胞系、遺伝子組換え組織、または遺伝子組換え器官を作製することができる(米国特許第4,736,866号)。ベクターを取り込んで十分な量のベクターを複製するこれらの細胞により、安定した組み込みが可能となり、本哺乳動物タンパク

50

質をコードする核酸分子の活性を弱める或いは完全に消失させる十分なアンチセンス分子を産生させることができる。

#### 【0148】

##### 10 NIM1キナーゼの発現

ヒトNIM1キナーゼの発現は、サル及び昆虫細胞系の発現系を用いて行った。cDNAを、QIAPREPスピニンプレップキット及びPLASMID MAX Iキット(共にQiagen)を用いて取扱説明書に従って精製した。

#### 【0149】

一過的な発現の場合、cDNAをpcDNA3.1(-)/myc-His Bベクター(Invitrogen, Carlsbad CA)にクローニングして、ベクターpcDNA3-Nim1を、CaPO<sup>4</sup>トランスフェクションキット2-463335(Eppendorf-5 Prime, Boulder CO)を用いてCOS 1-1細胞に形質転換した。形質転換する前に、 $4 \times 10^6$ の細胞を10cmの組織培養皿に撒いて24時間培養した。形質転換の日に、500 mlの2×DNA沈殿バッファ、62 mlのM CaCl<sub>2</sub>、10 mg (10 ml)のpcDNA3-Nim1、及び428 mlの水でCaPO<sub>4</sub>-DNAを沈殿させた。この混合液を室温で20分間インキュベートし、次にゆっくりと9 mlの培養液(DMEM、10%ウシ胎児血清、2 mM グルタミン、10 mg/mlのペニシリン、及び10 mg/mlのストレプトマイシン)を加えた。細胞を5% CO<sub>2</sub>で、37 °Cで4時間インキュベートした。培養液を新しい培養液に替え、細胞を更に48時間インキュベートした。

#### 【0150】

Sf21昆虫細胞を、BaculoGoldトランスフェクションキット(BD Pharmingen, San Diego CA)に付属の取扱説明書に従って同時形質転換した。 $2 \times 10^6$ の細胞を6 cmの組織培養皿に蒔き、27 °Cで15分間インキュベートした。4 mg (4 ml)のpVL1392/GST-NIM1発現ベクターを0.5 mg (0.5 ml)のBaculoGold DNAと共に室温で5分間インキュベートした。細胞培養液(TNM-FH)を除去して1 mlのバッファAに替えた。DNA混合物を1 mlのバッファBで希釈し、一滴ずつコンフェクションプレートに加えた。プレートを27 °Cで4時間インキュベートした後、培養液を替え細胞を5日間インキュベートした。組換えウイルスを3回増幅して $10^7$  pfu/mlのウイルスストックを得た。 $1.2 \times 10^7$ のSf21昆虫細胞を10 mlのウイルスストックで感染させ、27 °Cでインキュベートした。3日後に、細胞を溶解してタンパク質を精製した。

#### 【0151】

##### 11 タンパク質の精製

###### His精製

COS-1細胞を毎分1000回転で遠心分離し、4 mlの溶解バッファ(5 mM イミダゾール+0.5 mM NaCl+20 mM Na<sub>2</sub>HPO<sub>4</sub>+Complete protease inhibitor cocktail tablets、(Roche Molecular Biochemicals, Indianapolis IN))において再懸濁し、超音波処理した。毎分10000回転の回転数で、4 °Cで10分間遠心分離して可溶性画分を回収した。組換えNim1キナーゼを、製造者の取扱説明書に従って固定した金属イオンアフィニティークロマトグラフィー(IMAC, Invitrogen)で精製した。精製度を、SDS-ポリアクリルアミドゲル電気泳動法(SDS-PAGE)及びクーマーシー青色染色により測定した。タンパク質の濃度を、Bradford法により測定した。

#### 【0152】

###### GST精製

Sf21細胞を、毎分800回転で遠心分離し、10 mlの溶解バッファ(PBS+1 mM orthovanadate+20 mM DTT+Complete protease inhibitor cocktail tablets (Roche Molec

ular Biochemicals))において再懸濁し、超音波処理した。可溶性画分を、毎分10000回転、4 で10分間遠心分離して回収した。組換えNim1キナーゼを、製造者の取扱説明書に従ってグルタチオン-SEPHAROSEレジン(Amersham Pharmacia Biotech)上でのアフィニティークロマトグラフィーにより精製した。精製濃度を、Bradford法により測定した。

#### 【0153】

ウェスタンブロット分析の場合は、タンパク質をSDS-PAGEにより分離し、標準的な材料及び技術(ECL, Amersham Pharmacia Biotech)を用いて、免疫ブロッティング(GST-HRP結合抗体またはHis-HRP結合抗体(Santa Cruz Biotechnology)の何れかを用いて)により分析した。

#### 【0154】

##### 1.2 本タンパク質の特徴付け

インサイトクローン番号3317608の*in vitro*翻訳は、TN-T7クイック結合転写/翻訳システム(Promega)を用いて製造者の取扱説明書に従って行った。SDS-PAGE分析により、48 kdのタンパク質が存在することが分かった。タンパク質の大きさは、鋳型としてpcDNA3-Nim1を発現させて確認した。

#### 【0155】

先述したように、Nim1キナーゼは哺乳動物細胞及び昆虫細胞の両方において発現した。精製したタンパク質をSDS-PAGEで分析して、50 kdのタンパク質(my-Hisタグを含む)及び80 kdのタンパク質(GSTタグを含む)のそれぞれが確認された。

#### 【0156】

##### 1.3 ヒトNIM1キナーゼアッセイ

*in vitro*キナーゼアッセイを、100 ngの組換えNim1-GSTを10  $\mu$ Ciの[ $\gamma$ -<sup>32</sup>P]-ATP(3000 Ci/mmol; Amersham Pharmacia Biotech)を含む20  $\mu$ lのキナーゼバッファ(50 mM HEPES pH 7.5、3 mM MgCl<sub>2</sub>及びMnCl<sub>2</sub>、10 mM DTT、及び6  $\mu$ M NaOva)において、2  $\mu$ gのミエリン塩基性タンパク質(MBP)またはヒストン(HIST)と共に37 で30分間インキュベートして行った。反応は、サンプルバッファを加え100 で5分間加熱して停止させた。サンプルを、SDS-PAGEにより分析し、ゲルを乾燥させてオートラジオグラフィ分析を行った。

#### 【0157】

##### 1.4 NIM1キナーゼに特異的な抗体の生産

ポリアクリルアミドゲル電気泳動法でNIM1キナーゼを精製し、これを用いてマウスやウサギを免疫する。抗体は、以下の標準的なプロトコルを用いて生産する。別法では、NIM1キナーゼのアミノ酸配列をLASERGENEソフトウェア(DNA STAR)を用いて解析し、免疫原性の高い領域を決定する。通常はC末端付近或いは親水性領域内に存在する免疫原性エピトープを選択して合成し、これを用いて抗体を増大させる。通常は15残基程度の長さのエピトープを、Fmoc法でABI 431Aペプチドシンセサイザー(PE Biosystems)で合成し、これをN-マレイミドベンゾイル-N-ヒドロキシスクシンイミドエステルとの反応でKLH(Sigma-Aldrich)に結合させて免疫原性を高める。

#### 【0158】

エピトープ-KLH複合体をフロイントの完全アジュバントと共に投与してウサギを免疫する。その後、間隔を置いてフロイントの不完全アジュバントで免疫化を繰り返す。マウスの場合は少なくとも7週間、ウサギの場合は少なくとも12週間が経過した後、抗血清を採取して抗ペプチド活性を検査する。この検査には、このペプチドをプラスチックに結合し、1%ウシ血清アルブミン(BSA)でブロック処理し、ウサギ抗血清と反応させ、洗浄し、さらに放射性ヨウ素標識されたヤギ抗ウサギIgGと反応させる。当分野で

周知の方法を用いて抗体力価及び形成された複合体の収量を測定する。

【0159】

15 特異的な抗体を用いる天然タンパク質の精製

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

10

【0160】

16 本核酸分子若しくはタンパク質と特異的に結合する分子のスクリーニング

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

20

【0161】

17 2-ハイブリッドスクリーン

酵母2-ハイブリッド法、即ちMATCHMAKER LexA Two-Hybridシステム(Clontech Laboratories, Palo Alto CA)を用いて、本発明の哺乳動物タンパク質と結合するペプチドをスクリーニングする。本タンパク質をコードする核酸分子をpLexAベクターの複数のクローニング部位に挿入し、連結して大腸菌に形質転換する。mRNAから調整したcDNAを、pB42ADベクターの複数のクローニング部位に挿入して、結合させ、大腸菌に形質転換させてcDNAライブラリを作製した。pLexAプラスミド及びpB42AD-cDNAライブラリ作製物を大腸菌から単離し、これらを2:1の比率で用いて、polyethylene glycol/lithium acetateプロトコルに従ってコンピテント酵母EGY48 [p8op-lacZ]細胞を同時形質転換した。形質転換した酵母細胞をヒスチジン、トリプトファン、及びウラシルを含まない合成ドロップアウト(SD:synthetic dropout)培地(His, Trp, Ura)に移し、コロニーが成長してカウント出来るまで30でインキュベートした。このコロニーを1xTE(pH 7.5)の最小容量にプールし、2%アガロース(Gal)、1%ラフィノース(Raf)、及び80 mg/mlのX-Gal(5-bromo-4-chloro-3-indoryl-b-D-galactoside)を加えたSD培地(His, Leu, Trp, Ura)に移し、青いコロニーの成長を調べた。発現したタンパク質とcDNA融合タンパク質との間の相互作用により、EGY48におけるLEU2レポーター遺伝子の発現が活性化され、ロイシンを含まない培地(Leu)でコロニーが成長した。また、相互作用により、p8op-lacZレポーター作製物からの-ガラクトシダーゼの発現が活性化され、X-Gal上で成長したコロニーにおいて青色コロニーが生成される。

30

40

【0162】

発現タンパク質とcDNA融合タンパク質との間のポジティブな相互作用は、個々のポジティブコロニーを分離して、これらのコロニーをSD液体培地(Trp, Ura)で3

50

0 で2日間成長させるて確認することができる。培養物のサンプルをSD培地( Trp、 Ura )に移して、コロニーが現れるまで30 でインキュベートする。このサンプルを、SDプレート( Trp、 Ura )及びSDプレート( His、 Trp、 Ura )上でレプリカ培養する。ヒスチジン含有SD培地で成長するがヒスチジンを含まない培地では成長しないコロニーは、pLexAプラスミドが欠損している。ヒスチジンを必要とするコロニーをSD( Gal、 Raf、 X-Gal、 Trp、 Ura )上で成長させ、白いコロニーを単離して成長させた。本哺乳動物タンパク質と物理的に相互作用するタンパク質をコードする核酸分子を含むpB42AD-cDNAプラスミドを、酵母細胞から単離して特徴付けることが可能である。

【0163】

10

本明細書に記載した全ての特許及び刊行物に言及することを以って本明細書の一部とする。本発明の範囲及び概念から逸脱することなく本発明の方法及びシステムの種々の改変が可能であることは当業者には明らかであろう。特定の好適な実施例に基づいて本発明を説明したが、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、分子生物学或いは関連する分野の専門家には、本明細書に記載の本発明の実施例の様々な改変は、特許請求の範囲に含まれることが容易に理解できよう。

【図面の簡単な説明】

【図1A】

ヒトNIM1キナーゼ(SEQ ID NO: 2)をコードするヒト核酸分子(SEQ ID NO: 1)を示す。このアラインメントは、MACDNASIS PROソフトウェア(Hitachi Software Engineering, South San Francisco CA)を用いて作成した。

20

【図1B】

ヒトNIM1キナーゼ(SEQ ID NO: 2)をコードするヒト核酸分子(SEQ ID NO: 1)を示す。このアラインメントは、MACDNASIS PROソフトウェア(Hitachi Software Engineering, South San Francisco CA)を用いて作成した。

【図1C】

ヒトNIM1キナーゼ(SEQ ID NO: 2)をコードするヒト核酸分子(SEQ ID NO: 1)を示す。このアラインメントは、MACDNASIS PROソフトウェア(Hitachi Software Engineering, South San Francisco CA)を用いて作成した。

30

【図1D】

ヒトNIM1キナーゼ(SEQ ID NO: 2)をコードするヒト核酸分子(SEQ ID NO: 1)を示す。このアラインメントは、MACDNASIS PROソフトウェア(Hitachi Software Engineering, South San Francisco CA)を用いて作成した。

【図1E】

ヒトNIM1キナーゼ(SEQ ID NO: 2)をコードするヒト核酸分子(SEQ ID NO: 1)を示す。このアラインメントは、MACDNASIS PROソフトウェア(Hitachi Software Engineering, South San Francisco CA)を用いて作成した。

40

【図1F】

ヒトNIM1キナーゼ(SEQ ID NO: 2)をコードするヒト核酸分子(SEQ ID NO: 1)を示す。このアラインメントは、MACDNASIS PROソフトウェア(Hitachi Software Engineering, South San Francisco CA)を用いて作成した。

【図2】

脳組織(74%)、乳癌(結合組織)、及び前立腺(男性生殖器)におけるヒトNIM1

50

キナーゼの高度に特異的な発現の電子的ノーザン分析を示す。電子的ノーザン分析は、L I F E S E Q G o l d データベース ( I n c y t e P h a r m a c e u t i c a l s , P a l o A l t o C A ) を用いて作成した。

【図3】

様々な細胞系及び組織における定量PCRを用いて作成した転写物発現のグラフを示す。x軸は発現の倍数を示し、y軸は、細胞系 ( H 4 6 0、A 2 7 8 0、A 3 7 5、H D F、H E L A、D U 1 4 5、M D A - M B 2 3 1、U 8 7 - M G、及びB X - P C 3 ) 及び組織 ( 脳、結腸、子宮、及び胎盤の組織 ) を示す。

【図4A】

ヒトNIM1キナーゼ ( 3 3 1 7 6 0 8 C D 1 ; S E Q I D N O : 2 )、線虫STK ( g 3 8 7 7 3 2 9 ; S E Q I D N O : 3 1 )、ドブネズミSTK ( g 2 0 5 2 1 8 9 ; S E Q I D N O : 3 2 )、ヒトC-TAK1 ( g 3 0 8 9 3 4 9 ; S E Q I D N O : 3 3 )、ドブネズミ塩誘導性キナーゼ ( g 5 6 7 2 6 7 6 ; S E Q I D N O : 3 4 )、キイロショウジョウバエK78プロテインキナーゼ ( g 2 5 6 4 6 8 0 ; S E Q I D N O : 3 5 )、及びヒトEMK1 ( g 1 7 4 9 7 9 4 ; S E Q I D N O : 3 6 ) の間の保存された化学的及び構造的類似性を実証する。このアラインメントは、L A S E R G E N E ソフトウェアの M E G A L I G N プログラム ( D N A S T A R , M a d i s o n W I ) を用いて作成した。

【図4B】

ヒトNIM1キナーゼ ( 3 3 1 7 6 0 8 C D 1 ; S E Q I D N O : 2 )、線虫STK ( g 3 8 7 7 3 2 9 ; S E Q I D N O : 3 1 )、ドブネズミSTK ( g 2 0 5 2 1 8 9 ; S E Q I D N O : 3 2 )、ヒトCTAK1 ( g 3 0 8 9 3 4 9 ; S E Q I D N O : 3 3 )、ドブネズミ塩誘導性キナーゼ ( g 5 6 7 2 6 7 6 ; S E Q I D N O : 3 4 )、キイロショウジョウバエK78プロテインキナーゼ ( g 2 5 6 4 6 8 0 ; S E Q I D N O : 3 5 )、及びヒトEMK1 ( g 1 7 4 9 7 9 4 ; S E Q I D N O : 3 6 ) の間の化学的及び構造的類似性を実証する。このアラインメントは、L A S E R G E N E ソフトウェアの M E G A L I G N プログラム ( D N A S T A R , M a d i s o n W I ) を用いて作成した。

【図4C】

ヒトNIM1キナーゼ ( 3 3 1 7 6 0 8 C D 1 ; S E Q I D N O : 2 )、線虫STK ( g 3 8 7 7 3 2 9 ; S E Q I D N O : 3 1 )、ドブネズミSTK ( g 2 0 5 2 1 8 9 ; S E Q I D N O : 3 2 )、ヒトCTAK1 ( g 3 0 8 9 3 4 9 ; S E Q I D N O : 3 3 )、ドブネズミ塩誘導性キナーゼ ( g 5 6 7 2 6 7 6 ; S E Q I D N O : 3 4 )、キイロショウジョウバエK78プロテインキナーゼ ( g 2 5 6 4 6 8 0 ; S E Q I D N O : 3 5 )、及びヒトEMK1 ( g 1 7 4 9 7 9 4 ; S E Q I D N O : 3 6 ) の間の化学的及び構造的類似性を実証する。このアラインメントは、L A S E R G E N E ソフトウェアの M E G A L I G N プログラム ( D N A S T A R , M a d i s o n W I ) を用いて作成した。

【図5】

ヒトNIM1 ( h N I M 1 ) キナーゼアッセイを示す。酵素が h N I M 1 - G S T、基質がミエリン塩基性タンパク質 ( M B P ) 及びヒストン ( H I S T . 1 )、ポジティブコントロールが Z A P 7 0 ( 7 0 k d の 鎖 ( T C R ) 関連プロテインキナーゼ ) である。符号の「+」及び「-」はそれぞれ、各レーンにおけるヒトNIMキナーゼ及び基質の存在及び不在を表す。ゲルの右側に沿った大きさの指標はキロダルトンである。

【 1 A 】

5' AGA GAT GAG ATC CCG CAG CAG GGA CGT GGG GGC CTC CCA GGG GCA TTT ACG CAC 10 19 28 37 46 55  
 CAG AGT GCA AGA TTC TCT GGC CAY CAA GGG AAA TGG CAA ACA GAA GCC TTT GTC 64 73 82 91 100 109  
 CTG GGG CAC AGC CAC CTA CCA CAA AGC ATC AGA CTC CAC GTC TGG CCA GAA AGT 118 127 136 145 154 163  
 TTC TGG AGT CCC ATC AGG CCA GAG GGT ATG TAA CMT GTG CCT AAT TGT ACA GCT 172 181 190 199 208 217  
 AGA GCC TGC AAG TTC AAC GTG AGG GAA GGT GGG AAA TGT CTT GAG TGA GGC GAG 226 235 244 253 262 271  
 CAG CTC CTG GCT GGG CTG GGC AGA CTC AGC TAC CAC GTT CAC TGC CTT COT CTC 280 289 298 307 316 325  
 ACT AAA GCC GAG AAG GAG GCT GCT CAG CTC TCA GGA AAA CTC TTT TGA ACC CTG 334 343 352 361 370 379  
 GGC ACC TGC TGT COT CAG TTT GCA TCT CCC ACC CTC TGA GCC TCT TCT GCT CCT 388 397 406 415 424 433

【 1 B 】

GCA CAA CCT GCC TCT TGG CTG AGA TGG ABA CCF GAG CCC CCG TGG ACG ATG ACT 442 451 460 469 478 487 496  
 GCA GTG TAT ATG AAT GGA GGT GGC CTC GTG ATC CCC CAC TAT CCC CGG TGG GAT 496 505 514 523 532 541  
 A V Y M N G G L V N P H Y A R W D  
 CCG CCG GAC AGT GTA GAA AGT GGC TGT CAG ACC GAG AGT AGC AAG GTG GGT GAG 550 559 568 577 586 595  
 R R D S V E S Q C Q T E S K V G E  
 GAG GGA CAG CCC CCG CAG CTG ACG CCA TTC CAG ABA CTG ACA CAG GAG TGG TCC 604 613 622 631 640 649  
 E G Q P R Q L T F E K L I Q D M S  
 CAG GAT GAG AAG GTG GTG AGG GAG ATC AAG CTG GGG AAA CCG ATA GGC TTC TAC 658 667 676 685 694 703  
 Q D E K V V R E I T L G K R I G F Y  
 CGA APT CGA GGG GAA ATC GGA AGT GGA AAC TTC TCC CAA GTG AAG CTT GGG ATT 712 721 730 739 748 757  
 R I R G E I G S G N F S Q V K L G I

【 1 C 】

CAC TTC CTA ACC AAA GAA GAG GTG GCC ATT AAG ATC CTG GAC AAG ACC AAG TTA 766 775 784 802 811  
 H S L T X E K V A I K I L D R T K L  
 GAC CAG AAA ACC CAG AGG CTA CTA TCC CAA ATC TTC TCC AGC ATG GAA AAG CTG 820 829 838 847 856 865  
 D Q K T Q R L L S R E I S E H E K L  
 CAC GAT CCC AAC AUC ATC CCG CTT TAC GAA GTG GAG ACC CTA TCC AAG CAG 874 883 892 901 910 919  
 H H P N I I R L Y E V E T L S K L  
 CAC TUG GTG GAG TAT GCA GGG GGT GGG GAG CTC TTC GGA AAA ATT AGC ACT 928 937 946 955 964 973  
 H L V M E Y A G G E L F C K I S T  
 GAG GGG AAG CTC TCT GAA CCA GAA AGC CTC ATC TTC TCC CAG ATT GTG TCT 982 991 1000 1009 1018 1027  
 E G R L S E F E S R L I F S Q I V S  
 GCC GTG AAG CAC AAG CAA TAC CAA ATT ATT CAA AGA GAT CTC AAA CCA GAA 1036 1045 1054 1063 1072 1081  
 A V K H M H E N Q Y I H R D L K A E

【 1 D 】

AAT GTA TTC GAT ACC AGT AAT ACT TGT GAG AAG GTG GGC GAT TTT GGA TTC AGC 1090 1099 1108 1117 1126 1135  
 N V F Y T S N T C V K V G D F G F S  
 ACA GTA AGC AAA AAT GGT GAA ATG CAG AAC ACT TTC TGT GGG TCT CCF CCC TAC 1144 1153 1162 1171 1180 1189  
 T V S K K G E M L N T F C G S P P Y  
 GGT GCG CTT GAA CTT TTC CCG GAC CAC TAC ATC GGC ATT TAC GTG GAT ATC 1198 1207 1216 1225 1234 1243  
 A A P E L F R D E H Y I G I Y V D I  
 TGG GCC TTG GGG GTG CTT TTG TAC TTC AAG GTG ACT GGC ACC ATG CCA TTT GGG 1252 1261 1270 1279 1288 1297  
 W A L G V L L Y F M V T G T M P F R  
 GCA GAA ACC GTG GCC AAA CTA AAA AAG ACC ATC CTC GAG GGC ACA TAC AGT GTA 1306 1315 1324 1333 1342 1351  
 A E I V A K L K S I L E G T Y S V  
 CCG CCG CAC GTG TCA GAG CCC TGC CAC CGA CTC ATC CGA GGA GTC CTT CAG CAG 1360 1369 1378 1387 1396 1405  
 P P H V S E P C H R L I R G V L Q Q

【 図 1 E 】

1414 1423 1432 1441 1450 1459  
 ATC CCC ACG GAG AGG TAC GGA ATC GAC TGC ATC ATG ATG GAT GAT TGG ATG CAA  
 I P T E R Y G I D C I M N D E W M Q

1468 1477 1486 1495 1504 1513  
 GGG GTG CCA TAC CCT ACA CCT TTG GAA CCT TTC CAA CTG CAT CCC AAA CAT TTG  
 G V P Y P T P L E P F Q L D P K H L

1522 1531 1540 1549 1558 1567  
 TCG GAA ACC AGC ACT CTC AAG GAA GAA AAT GAG GTC AAA TGG NCT TTA GAA  
 S E T S T L K E E E N E Y K S T L E

1576 1585 1594 1603 1612 1621  
 CAT TTC GGC ATT ACA GAA GAG CAT ATT CGA AAT AAC CAA GGG AGA GAT GCT CGC  
 H L G I T E E H I R N N Q G R D A E R

1630 1639 1648 1657 1666 1675  
 AGC TCA ATC ACA GGG GTC TAT AGA ATT ATT TTA CAT AGA GTC CAA AGG AAG AAG  
 S S I T G V Y R I I L H R V Q R K K

1684 1693 1702 1711 1720 1729  
 GCT TTC GAA AAT GTC CCA GTC ATG ATG CTA CCA GAC CCT AAA GAA NGA GAC CTC  
 A L E S V P V M M L P D P K E R D L

【 図 1 F 】

1738 1747 1756 1765 1774 1783  
 AAA AAA GCG TCC CGT GTC TAC AGA GGG ATA AGA GAC ACA TCC AAA TTT TGC TCG  
 K K G S R V Y R G I R H T S K F C S

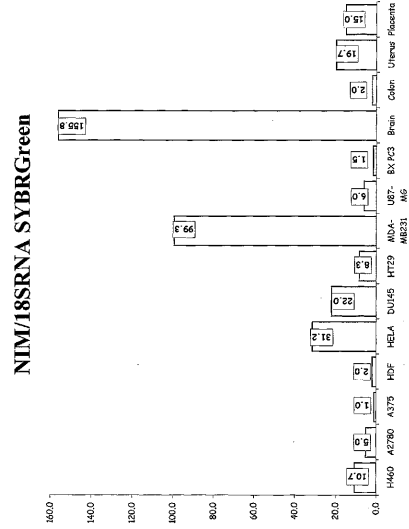
1792 1801 1810 1819 1828 1837  
 ATT TTA TAA ATT GCA CTA GAC TGC TTG TAA CTA ACC AAG ATG ATT GTT GCT GCT  
 I L

1846 1855 1864 1873 1882 1891  
 TCT AAA TTT TTT TCA ACG ACA ACT TGA GTG GAG ACA TTT TTG TAA TTT TTA AAT  
 1900 1909 1918 1927 1936 1945  
 AAA CTT AAA TTT GAG ATA TGC AAA AAA AAA AAA AAA AAA GGC GGC CGA CTA  
 1954 1963 1972 1981 1990 1999  
 GTG AGC TCG TCG ACC CGS GAA TTA ATT CCG GAC CCG TAC CTG CAG CCG TAC CAG  
 2008 2017 2026 2035 2044 2053  
 CTT TCC CTA TAG TGG AGT CCG TAT TAA ACT TGG CCG TAA TCA TCA TCG CAT AAG TTG  
 TTT CCT G 3

【 図 2 】

クローンのヒト組織名	数	ライブラリの割合	存在数
心臓系	247285	0/62	0
結合組織	120592	1/38	1
消化器系	375714	0/311	0
肝	84970	0/16	0
内分泌系	179593	0/43	0
外分泌系	240308	0/59	0
免疫系	344493	0/62	0
生殖系	411485	1/103	1
中枢神経系	643113	0/144	0
血液系/免疫系	78332	0/20	0
肝臓	135466	0/38	0
筋骨格系	709414	10/144	10
神経系	93477	0/21	0
皮膚	335356	0/73	0
呼吸器系	15563	0/8	0
泌尿器系	10993	0/4	0
造血系	89773	0/3	0
未分類/混合	233771	0/53	0
原腸			
合計	4442057	12/1039	12

【 図 3 】



【 図 4 A 】

```

1  M T A V Y M N G G G L V N P H V A R W D R R D S V E S G - - - C O T E S S K V G >317608CD1
2  M S E - - - - - K T Q Y E R A I S O L N N D V - - - V H K E V M A C - >3877329
1  M S T - R P L P T V N E R D D E N H T S - - - H C D D R Q E V T S R T S R - >2052189
1  M S T - R P L P T V N E R D D E N H T S - - - H C D D R Q E V T S R T S R - >3089349
1  M T A M T M K T T K K E M D P P N T S I T T A P P S A A Q N V G C V - >5672676
1  M - - - - - I R G R N S - - - - - >2564680
1  M - - - - - I R G R N S - - - - - >4749794
38  L E G Q R O L T P P E K L T O D M S Q D R K V V R E I T L G R I G F Y E L R >317608CD1
39  - - - - - I V S Y G L - - - - - R K L W F O V A L L G R K M - - - I G Y F L G >3877329
38  - - - - - I V S Y G L - - - - - R K L W F O V A L L G R K M - - - I G Y F L G >2052189
38  - - - - - I V S Y G L - - - - - R K L W F O V A L L G R K M - - - I G Y F L G >3089349
35  - S G - A B C R N S T S A D - D E Q P H - - - - - I G N Y F L L - - - - - >5672676
4  G S S G G S S P K K S Y V M G M C Y - - - - - G K P L - - - - - E V G F Y D V K >2564680
4  G S S G G S S P K K S Y V M G M C Y - - - - - G K P L - - - - - E V G F Y D V K >4749794
8  - - - - - A T S A D E Q P H - - - - - I G N Y F L L - >317608CD1
78  G E I G S G N F S O V K L G I H L S L T K E K V A I K I L D K T K L D O K T O L L >317608CD1
53  K E L I G K G N F S K V K L G V H Q L T K E K V A I K I L D K T K L D O K T O L L >3877329
64  K T I G K G N F A K V K L A R V L L T G R E V A I K I I D K T O L N P T S L O K >2052189
60  K T I G K G N F A K V K L A R V L L T G R E V A I K I I D K T O L N P T S L O K >3089349
31  R T L G K G N F A V V K L A E H R V T Q V A I K I I D K T R T D S S N L E K >5672676
67  K T L G K G N F A K V K L A L L P T G R E V A I K I I D K T A H N T I A R O K >2564680
34  K T L G K G N F A K V K L A R V L L T G R E V A I K I I D K T O L N P T S L O K >4749794

```

【 図 4 B 】

```

118  L S R R E I S M E K L E H P P N I E L I V V V E T L S E K L H L V M E Y A G G E >317608CD1
123  L S R R E I S M E K L E H P P N I E L I V V V E T L S E K L H L V M E Y A G G E >3877329
108  L S R R E I S M E K L E H P P N I E L I V V V E T L S E K L H L V M E Y A G G E >2052189
108  L S R R E I S M E K L E H P P N I E L I V V V E T L S E K L H L V M E Y A G G E >3089349
71  L V R V E T L M K L M H P N I L V L O N I S E S T L Y L V M E Y A S G G E >5672676
107  L V R V E T L M K L M H P N I L V L O N I S E S T L Y L V M E Y A S G G E >2564680
64  L V R V E T L M K L M H P N I L V L O N I S E S T L Y L V M E Y A S G G E >4749794
158  L F G K I T E G K L S E P E S L I F S I V S A V K H M H E M O I I H R D L >317608CD1
133  L Y V H R G K K T B D A L P L P A I I V S A V S H M H S E M I V H R D L >3877329
144  V F D Y L V A H G R K K E A A K P R I V S A V Q Y C H Q C H I V H R D L >2052189
140  V F D Y L V A H G R K K E A A K P R I V S A V Q Y C H Q C H I V H R D L >3089349
111  M F D Y L S N G H L S E N E A R K F W O L I S A M E Y C H N H H I V H R D L >5672676
147  L F N Y L N G R R E R R A R K L P R O L V S A L E Y C H S K S I V H R D L >2564680
104  V E I D Y L V A H G R K K E A A K P R I V S A V Q Y C H Q C H I V H R D L >4749794
168  K A E N V E Y T S N C V K V G D F G S F S K S K G E M I N T F C G S P S Y A >317608CD1
173  K A E N V E Y T S N C V K V G D F G S F S K S K G E M I N T F C G S P S Y A >3877329
184  K A E N L L L D A D M N I K I A D F G F S E F T V G K L D H T F C G S P P Y A >2052189
180  K A E N L L L D A D M N I K I A D F G F S E F T V G K L D H T F C G S P P Y A >3089349
151  K A E N L L L D D G M D I K I A D F G F S E F T V G K L S H T F C G S P P Y A >5672676
187  K A E N L L L D Q O K L I K I A D F G F S E F T V G K L S H T F C G S P P Y A >2564680
144  K A E N L L L D A D M N I K I A D F G F S E F T V G K L D H T F C G S P P Y A >4749794

```

【 図 4 C 】

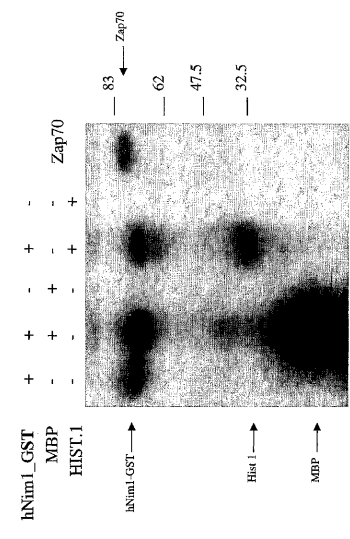
```

238  A P E L F P R D E V I G I Y V D I W A L G V L L V P M V I G T M P F R A E T V A >317608CD1
213  A P E L F Q D T V A G E P P V D V W A L G V L L H E F M L I G V T P P K A B E T V P >3877329
224  A P E L F Q G K K V D G P P V D V W S L G V I L L V T L V S G S L P F D Q N L K >2052189
220  A P E L F Q G K K V D G P P V D V W S L G V I L L V T L V S G S L P F D Q N L K >3089349
191  A P E L F Q G K K V D G P P V D V W S L G V I L L V T L V S G S L P F D G T N L P >5672676
227  A P E L F Q G K K V S G P P V D V W S L G V I L L V T L V S G S L P F D G T N L K >2564680
184  A P E L F Q G K K V D G P P V D V W S L G V I L L V T L V S G S L P F D G T N L K >4749794
278  K L K S T L E C T V S V P E H V S E C H R L F R G V I O O I P T B V G I D >317608CD1
253  K L K S T L E C T V S V P E H V S E C H R L F R G V I O O I P T B V G I D >3877329
260  K L K S T L E C T V S V P E H V S E C H R L F R G V I O O I P T B V G I D >2052189
260  K L K S T L E C T V S V P E H V S E C H R L F R G V I O O I P T B V G I D >3089349
260  K L K S T L E C T V S V P E H V S E C H R L F R G V I O O I P T B V G I D >5672676
231  T L L R O R V L E G R P R I J F F M S O C E T L I R M L L V D P A A K N I T I A >2564680
267  E L L R V L R G K V R V L V V V S I C E S L I K K K P L L V L N P T O K T S L S >4749794
224  E L L R V L R G K V R V L V V V S I C E S L I K K K P L L V L N P T O K T S L S >317608CD1
318  C I N K D F W M G - - - - - >317608CD1
283  S V K K H F W M H - - - - - >3877329
304  O I N K D R W M N V G H - - - - - >2052189
300  O I N K D R W M N V G H - - - - - >3089349
271  O I N K D R W M N V G H - - - - - >5672676
307  A U M A D R M L N G Y - - - - - >2564680
284  Q I M A D R M L N V G H - - - - - >4749794

```

【 図 5 】

hNIM1 を用いるキナーゼアッセイ



【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  
20 September 2001 (20.09.2001)

PCT

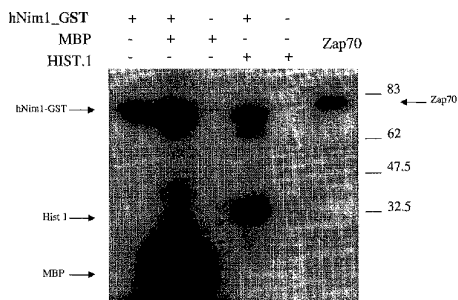
(10) International Publication Number  
**WO 01/68825 A1**

- (51) International Patent Classification: **C12N 9/12**, 15/54, 5/10, C12Q 1/68, C07K 16/40
- (52) International Application Number: PCT/US00/07715
- (22) International Filing Date: 22 March 2000 (22.03.2000)
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data: 09/523,849 13 March 2000 (13.03.2000) US
- (63) Related by continuation (CON) or continuation-in-part (CIP) to earlier application: US 09/523,849 (CIP) Filed on 13 March 2000 (13.03.2000)
- (71) Applicants (for all designated States except US): **INCYTE PHARMACEUTICALS, INC.** [US/US]; 3160 Porter Drive, Palo Alto, CA 94304 (US). **MOLteni**, Angela [IT/IT]; Via Giuseppe Citterio, 4, I-22603 Cantù (IT).
- (72) Inventors; and (75) Inventors/Applicants (for US only): **BANDMAN, Olga** [US/US]; 366 Anna Avenue, Mountain, CA 94043 (US). **MAGNAGHI, Paola** [IT/IT]; Via Padova, 196, I-20132 Milano (IT). **BOSOTTI, Roberta** [IT/IT]; Via Pietro Mascagni, 2, I-20014 Nerviano (IT). **SCACHERI, Emanuela** [IT/IT]; Via Desenzano, 7, I-20146 Milano (IT). **ISACCHI, Antonella** [IT/IT]; Via Montecatini, 14, I-20100 Milano (IT). **HODGSON, David, M.** [US/US]; 567 Addison Avenue, Palo Alto, CA 94304 (US).
- (74) Agents: **MURRY, Lynn, E.** et al.; Incyte Pharmaceuticals, Inc., 3160 Porter Drive, Palo Alto, CA 94304 (US).
- (81) Designated States (national): AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, 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, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZA, ZW.

[Continued on next page]

(54) Title: HUMAN NIM1 KINASE

**KINASE ASSAY USING hNIM1**



(57) Abstract: The invention provides a nucleic acid molecule which encodes the human NIM1 kinase. It also provides for the use of the nucleic acid molecule, fragments, variants and complements thereof and of the protein, portions thereof and antibodies thereto for characterization, diagnosis, evaluation, treatment, or prevention of disorders associated with expression. The invention additionally provides expression vectors and host cells for the production of the protein and a transgenic organism or model system.



WO 01/68825 A1

---

**WO 01/68825 A1**

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

**Published:**  
— with international search report

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

WO 01/68825

PCT/US00/07715

**HUMAN NIMI KINASE****TECHNICAL FIELD**

This invention relates to a nucleic acid molecule which encodes human NIMI kinase and to  
5 the use of the nucleic acid molecule and the protein it encodes in the characterization, diagnosis,  
prevention, and treatment of brain disorders and cancers, particularly breast cancer.

**BACKGROUND OF THE INVENTION**

Phylogenetic relationships among organisms have been demonstrated many times, and studies  
10 from a diversity of prokaryotic and eukaryotic organisms suggest a more or less gradual evolution of  
molecules, biochemical and physiological mechanisms, and metabolic pathways. Despite different  
evolutionary pressures, the protein kinases of nematode, fly, rat, and man have common chemical and  
structural features and modulate the same general cellular activity. Comparisons of the nucleic acid  
and protein sequences from organisms where structure and/or function are known accelerate the  
15 investigation of human sequences and allow the development of model systems for testing diagnostic  
and therapeutic agents for human conditions, diseases, and disorders.

Protein kinases regulate many different cell proliferation, differentiation, and signaling  
processes by adding phosphate groups to proteins. Uncontrolled signaling has been implicated in a  
variety of disease conditions including inflammation, cancer, arteriosclerosis, and psoriasis.  
Reversible protein phosphorylation is the main strategy for controlling activities of eukaryotic cells. It  
20 is estimated that more than 1000 of the 10,000 proteins active in a typical mammalian cell are  
phosphorylated. The high energy phosphate which drives activation is generally transferred from  
adenosine triphosphate (ATP) or guanosine triphosphate (GTP) to a particular protein by protein  
kinases and removed from that protein by protein phosphatases. Phosphorylation is roughly  
analogous to turning on a molecular switch, and it occurs in response to extracellular signals (mediated  
25 by such molecules as hormones, neurotransmitters, growth and differentiation factors), cell cycle  
checkpoints, and environmental or nutritional stresses. When the switch goes on, the appropriate  
protein kinase activates a metabolic enzyme, regulatory protein, receptor, cytoskeletal protein, ion  
channel or pump, or transcription factor.

The protein kinases comprise the largest known protein group, a superfamily of enzymes with  
30 widely varied functions and specificities. They are usually named after their substrate, their regulatory  
molecules, or some aspect of a mutant phenotype. With regard to substrates, the protein kinases may  
be roughly divided into two groups; those that phosphorylate tyrosine residues (protein tyrosine  
kinases, PTK) and those that phosphorylate serine or threonine residues (serine/threonine kinases,  
STK). A few protein kinases have dual specificity and phosphorylate both threonine and tyrosine  
35 residues.

WO 01/68825

PCT/US00/07715

Protein kinases may be categorized into families by the different amino acid residues (generally between 5 and 100 residues) located on either side of, or inserted into loops of, the catalytic domain. These residues allow the regulation of each kinase as it recognizes and interacts with its target protein. Almost all kinases contain a similar 250-300 amino acid catalytic domain with 11 subdomains distributed across two lobes. The N-terminal lobe, which contains subdomains I-IV, binds and orients the ATP donor molecule. The larger C terminal lobe, which contains subdomains VIA-XI, binds the protein substrate and carries out the transfer of the gamma phosphate from ATP to the hydroxyl group of a serine, threonine, or tyrosine residue. Subdomain V spans the N and C terminal lobes.

Each of the 11 subdomains contain specific residues and motifs or patterns of amino acids that are characteristic of that subdomain and are highly conserved (Hardie and Hanks (1995) The Protein Kinase Facts Books, Vol I, Academic Press, San Diego CA, pp. 7-20). In particular, two protein kinase signature sequences have been identified in the kinase domain, the first containing an active site lysine residue involved in ATP binding, and the second containing an aspartate residue important for catalytic activity. If a protein is found to contain the two protein kinase signatures, the probability of that protein being a protein kinase is close to 100% (MOTIFS search program, Genetics Computer Group, Madison WI; Bairoch et al. (1996) *Nucleic Acids Res* 24:189-196).

#### STK Families

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 calcium-calmodulin. The cyclic-AMP dependent protein kinases (PKA) are important members of the STK family. cAMP is an intracellular mediator of hormone action in all prokaryotic and animal cells that have been studied. Such 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 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 et al. (1994) Harrison's Principles of Internal Medicine, McGraw-Hill, New York NY, pp. 416-431 and 1887).

Calcium-calmodulin (CaM) dependent protein kinases are also members of STK family. Calmodulin is a calcium receptor that mediates many calcium regulated processes by binding to target proteins in response to the binding of calcium. The principle target protein in these processes is CaM dependent protein kinases (CaMK). CaMK are involved in regulation of smooth muscle contraction, glycogen breakdown (phosphorylase kinase), and neurotransmission (CaMK I and CaMK II). CaMK I

WO 01/68825

PCT/US00/07715

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 *et al.* (1995) EMBO J 14:3679-86). CaMK II also phosphorylates synapsin at different sites and controls the synthesis of catecholamines in the brain through phosphorylation and activation of tyrosine hydroxylase. Many of the CaMK are activated by phosphorylation in addition to binding to CaM. CaMK may autophosphorylate or be phosphorylated by another kinase as part of a "kinase cascade".

Another ligand-activated protein kinase is 5'-AMP-activated protein kinase (AMPK; Dyck *et al.* (1996) J Biol Chem 271:17998-17803). Mammalian AMPK is a regulator of fatty acid and sterol synthesis through phosphorylation of the enzymes acetyl-CoA carboxylase and hydroxymethylglutaryl-CoA reductase and mediates responses of these pathways to cellular stresses such as heat shock and depletion of glucose and ATP. AMPK is a heterotrimeric complex comprised of a catalytic alpha subunit and two non-catalytic beta and gamma subunits that are believed to regulate the activity of the alpha subunit. Subunits of AMPK have a much wider distribution in non-lipogenic tissues such as brain, heart, spleen, and lung than expected. This distribution suggests that its role may extend beyond regulation of lipid metabolism alone.

The mitogen-activated protein kinases (MAPK) are also members of the STK family, and they regulate intracellular signaling pathways. MAPK mediate signal transduction from the cell surface to the nucleus via phosphorylation cascades. Several subgroups have been identified, and each manifests different substrate specificities and responds to distinct extracellular stimuli (Egan and 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 mammalian pathways include epidermal growth factor, ultraviolet light, hyperosmolar medium, heat shock, endotoxic lipopolysaccharide, and pro-inflammatory cytokines such as tumor necrosis factor and interleukin-1. Altered MAPK expression is implicated in a variety of disease conditions including cancer, inflammation, immune disorders, and disorders affecting growth and development.

Proliferation-related kinase (PRK) is a serum/cytokine inducible STK that is involved in regulation of the cell cycle and cell proliferation in human megakaryotic cells (Li *et al.* (1996) J Biol Chem 271:19402-8). PRK is related to the polo family of STKs implicated in cell division. PRK is downregulated in lung tumor tissue and may be a proto-oncogene whose deregulated expression in normal tissue leads to oncogenic transformation.

The cyclin-dependent protein kinases (CDKs) are another group of STKs that control the progression of cells through the cell cycle. Cyclins are small regulatory proteins that act by binding to and activating CDKs which then trigger various phases of the cell cycle by phosphorylating and activating selected proteins involved in the mitotic process. CDKs are unique in that they require

WO 01/68825

PCT/US00/07715

multiple inputs to become activated. In addition to the binding of cyclin, CDK activation requires the phosphorylation of a specific threonine residue and the dephosphorylation of a specific tyrosine residue.

The discovery of nucleic acid molecules encoding a human NIM1 kinase (hNIM) provides  
5 new compositions which are useful in the characterization, diagnosis, prevention, and treatment of brain disorders and cancers, particularly breast cancer.

#### SUMMARY OF THE INVENTION

The invention is based on the discovery of a substantially purified nucleic acid molecule  
10 encoding a human NIM1 kinase which satisfies a need in the art by providing compositions useful in the characterization, diagnosis, prevention, and treatment of brain disorders and cancers, particularly breast cancer.

The invention provides a substantially purified nucleic acid molecule which encodes the  
human NIM1 kinase comprising SEQ ID NO:2. The invention also provides a composition  
15 comprising SEQ ID NO:1 or a fragment or a complement thereof. The invention further provides a mammalian variant of the nucleic acid molecule selected from SEQ ID NOs:24-30. The invention still further provides a fragment of at least 18 consecutive nucleotides selected from about nucleotide 414 to about 1414 of SEQ ID NO:1, SEQ ID NOs:3-30, or the complements thereof. In one aspect, the  
invention provides a substrate containing at least one of these fragments. In a second aspect, the  
invention provides a probe comprising the fragment which can be used in methods of detection,  
20 screening, and purification. In a further aspect, the probe is a single stranded complementary RNA or DNA molecule.

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

The invention also provides a method for using a nucleic acid molecule or a fragment or a  
30 complement thereof to screen a library of molecules or compounds to identify at least one ligand which specifically binds the nucleic acid molecule, the method comprising combining the nucleic acid molecule with a library of molecules or compounds under conditions allowing specific binding, and detecting specific binding to the nucleic acid molecule, thereby identifying a ligand which specifically binds the nucleic acid molecule. In one aspect, the libraries of molecules and compounds include  
35 DNA and RNA molecules, peptides, PNAs, proteins, and the like.

WO 01/68825

PCT/US00/07715

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

5 The invention provides an isolated and purified protein comprising SEQ ID NO:2 or a portion thereof. Additionally, the invention provides a pharmaceutical composition comprising a substantially purified protein or a portion thereof in conjunction with a pharmaceutical carrier.

The invention provides a method for using at least a portion of the protein to produce antibodies. The invention also provides a method for using a protein to screen a library of molecules  
10 or compounds to identify at least one ligand which specifically binds the protein, the method comprising combining the protein with the library of molecules or compounds under conditions allowing specific binding, and detecting bound protein, thereby identifying a ligand which specifically binds the protein. Such libraries of molecules and compounds include agonists, antagonists, antibodies, DNA and RNA molecules, immunoglobulins, drug compounds, mimetics, peptides,  
15 pharmaceutical agents, and other ligands. The invention further provides an analogous method which uses the protein to purify a ligand. The method involves combining the protein with a sample under conditions to allow specific binding, recovering the bound protein, and separating the protein from the ligand to obtain purified ligand.

The invention additionally provides antibodies identified by screening methods using or  
20 antibodies produced against the human NIM1 kinase. A method of preparing an antibody comprising immunizing an animal with the human NIM1 kinase or an antigenically-effective portion thereof under conditions to elicit an antibody response; isolating animal antibodies; and screening the isolated antibodies with NIM1 kinase thereby identifying an antibody specifically binds Nim1 kinase. In one aspect these antibodies are useful as diagnostic compositions in identification of brain disorders and  
25 cancers. In another aspect, the antibody may be administered as a pharmaceutical composition to treat brain disorders and cancers associated with the overexpression of human NIM1 kinase.

The further provides a method for inserting a marker gene into the genomic DNA of a mammal to disrupt the expression of the natural nucleic acid. The invention also provides a method  
30 for using a nucleic acid molecule to produce a mammalian model system, the method comprising constructing a vector containing the nucleic acid molecule selected from SEQ ID NOs:1 and 3-30; transforming the vector into an embryonic stem cell; selecting a transformed embryonic stem; microinjecting the transformed embryonic stem cell into a mammalian blastocyst, thereby forming a chimeric blastocyst; transferring the chimeric blastocyst into a pseudopregnant dam, wherein the dam gives birth to a chimeric offspring containing the nucleic acid molecule in its germ line; and breeding  
35 the chimeric mammal to produce a homozygous, mammalian model system.

WO 01/68825

PCT/US00/07715

**BRIEF DESCRIPTION OF THE FIGURES AND TABLE**

Figures 1A, 1B, 1C, 1D, 1E, and 1F show the nucleic acid molecule (SEQ ID NO:1) encoding the human NIM1 kinase (SEQ ID NO:2). The alignment was produced using MACDNASIS PRO software (Hitachi Software Engineering, South San Francisco CA).

5 Figure 2 shows electronic northern analysis of the highly specific expression of human NIM1 kinase in brain tissue (74%) and cancers of the breast (connective tissue) and prostate (genitalia, male). The electronic northern was produced using the LIFESEQ Gold database (Incyte Pharmaceuticals, Palo Alto CA).

10 Figure 3 shows a graph of transcript expression produced using quantitative PCR in various cell lines and tissues. The x axis shows fold expression; and the y axis, the cell lines (H460, A2780, A375, HDF, HELA, DU145, MDA-MB231, U87-MG and BX-PC3) and tissues (brain, colon, uterus and placenta tissues).

Figures 4 A, 4B, and 4C demonstrate the conserved chemical and structural similarities among the catalytic domains of human NIM1 kinase (3317608CD1; SEQ ID NO:2), Caenorhabditis elegans 15 STK (g3877329; SEQ ID NO:31), Rattus norvegicus STK (g2052189; SEQ ID NO:32), human C-TAK1 (g3089349; SEQ ID NO:33), Rattus norvegicus salt-inducible kinase(g5672676; SEQ ID NO:34), Drosophila melanogaster K78 protein kinase (g2564680; SEQ ID NO:35), and human EMK1(g1749794; SEQ ID NO:36). The alignment was produced using the MEGALIGN program of LASERGENE software (DNASTAR, Madison WI).

20 Figure 5 shows the human NIM1 (hNIM1) kinase assay. The enzyme, hNIM1-GST; substrates: myelin basic protein (MBP) and histone (HIST.1); and the positive control, ZAP70 (70 kd zeta-chain (TCR) associated protein kinase). The designations + and - indicate the presence or absence of the human NIM kinase and substrates in the lane. Size markers in kilodaltons are shown along the right side of the gel.

**DESCRIPTION OF THE INVENTION**

It is understood that this invention is not limited to the particular machines, materials and methods described. It is also to be understood that the terminology used herein is for the purpose of describing particular embodiments only and is not intended to limit the scope of the present invention which will be limited only by the appended claims. As used herein, the singular forms "a," "an," and 30 "the" include plural reference unless the context clearly dictates otherwise. For example, a reference to "a host cell" includes a plurality of such host cells known to those skilled in the art.

Unless defined otherwise, all technical and scientific terms used herein have the same meanings as commonly understood by one of ordinary skill in the art to which this invention belongs. All publications mentioned herein are cited for the purpose of describing and disclosing the cell lines, 35 protocols, reagents and vectors which are reported in the publications and which might be used in

WO 01/68825

PCT/US00/07715

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**

“NIM1 kinase” refers to a substantially purified enzyme obtained from any mammalian species, including bovine, ovine, porcine, rodent, canine, simian, and preferably the human species, and from any source, whether natural, synthetic, semi-synthetic, or recombinant.

“Biologically active” refers to a protein having structural, immunological, regulatory, or chemical functions of a naturally occurring, recombinant or synthetic molecule

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

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

“Fragment” refers to an Incyte clone or any part of a nucleic acid molecule which retains a usable, functional characteristic. Useful fragments are generally at least 18 consecutive nucleotides in length and include oligonucleotides which may be used in hybridization, amplification or screening technologies or in regulation of replication, transcription or translation.

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

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

“Nucleic acid molecule” refers to a nucleic acid sequence, oligonucleotide, nucleotide, polynucleotide, DNA molecule, or any fragment or complement thereof. It may be DNA or RNA of genomic or synthetic origin, double-stranded or single-stranded, and combined with carbohydrate, lipids, protein or other materials to perform a particular activity such as transformation or form a useful composition such as a peptide nucleic acid (PNA). It may contain untranslated 5' or 3' regulatory regions or introns. Preferably, the nucleic acid molecule has from about 15 to 10,000

WO 01/68825

PCT/US00/07715

nucleotides, more preferably from about 60 to 6,000 nucleotides and most preferably about 400 to 5000 nucleotides. "Oligonucleotide" is substantially equivalent to the terms amplicon, primer, oligomer, element, target, and probe, is preferably single stranded, and preferably is about 15 to 60 nucleotides in length and more preferably, about 22 to 25 nucleotides.

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

"Portion" refers to any part of a protein used for any purpose, but especially for the screening of a library of molecules or compounds to identify molecules which specifically bind to that portion of the protein or for the production of antibodies.

10 "Reporter molecules" are chemical or biochemical moieties used for labeling a polynucleotide, a polypeptide, or an antibody. They include, but are not limited to, radionuclides, enzymes, substrates, cofactors, inhibitors, fluorescent agents, chromogenic agents, chemiluminescent agents, magnetic particles, and the like. Reporter molecules specifically bind, establish the presence of, and allow quantification of a particular polynucleotide, polypeptide, or antibody.

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

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

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

35 "Substantially purified" refers to nucleic acid molecules or proteins that are removed from their natural environment and are isolated or separated, and are at least about 60% free, preferably

WO 01/68825

PCT/US00/07715

about 75% free, and most preferably about 90% free, from other components with which they are naturally associated.

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

#### THE INVENTION

The invention is based on the discovery of a nucleic acid molecule which encodes human NIM1 kinase and on the use of the nucleic acid molecule, or fragments thereof, and protein, or portions thereof, directly or as compositions in the characterization, diagnosis, treatment, or prevention of brain disorders and cancers.

The nucleic acid molecule encoding NIM1 kinase of the present invention was first identified as a kinase by Block II homology match between Incyte Clone 670279 from the cerebellum library (CRBLNOT01) and a putative STK of *C. elegans* (g733122) important for the polarity of zygote division and differentiation. The full length nucleic acid molecule, Incyte Clone 3317608 (SEQ ID NO:1) was sequenced and assembled from the Incyte LIFESEQ GOLD database (Mar 99 release) template 200700.1, the assembly 670279CB1 and Incyte Clones (Library): 3317608H1 (PROSBPT03), 4313713H1 (BRAFN0T01), 4617082H1 (BRAYDIT01), 4711644H1 (BRAIHCT01), 2286324H1 (BRAINON01), 2286816H1 (BRAINON01), 2287217H1 (BRAINON01), 2286816R6 (BRAINON01), 2286816T6 (BRAINON01), 3317608T6 (PROSBPT03), 4201896T6 (BRAITUT29), 4624811T6 (FIBRTXT02), 6559834H1 (BRAFNON02), 670279F1 (CRBLNOT01), 670279H1 (CRBLNOT01), 670279R1 (CRBLNOT01), 670279R6 (CRBLNOT01), 670279T6 (CRBLNOT01), and 4936446H1 (BRAXNOT03) and which are SEQ ID NOs:3-23, respectively. Useful fragments of the polynucleotide that encodes SEQ ID NO:2 include a fragment of at least 18 consecutive nucleotides selected from about nucleotide 414 to about nucleotide 1414 of SEQ ID NO:1, SEQ ID NOs:3-23 or the complements thereof.

Figures 1 A-F show the sequence of the nucleic acid molecule and its deduced translation into amino acids. Incyte Clone number 3317608 which contains the coding region for the human NIM1 kinase has been deposited in the American Type Culture Collection (ATCC; Manassas VA) and has the Patent Deposit Designation: PT-1217.

As shown in Figure 2, electronic northern analysis showed highly differential expression of the transcript encoding NIM1 kinase in the nervous system. All of these libraries were from brain; five were associated with cancer, three, with stroke, two, with Huntington's disease and one, with epilepsy. The transcript was found to be expressed only two other times in the LIFESEQ database (Incyte Pharmaceuticals) which contains 1039 libraries and over 5 million sequences, in cancerous

WO 01/68825

PCT/US00/07715

breast fibroblasts, FIBRTXT02, and in cancerous prostate, PROSBPT03.

As presented in Figure 3, northern analysis was also performed in the laboratory using quantitative PCR. Expression of human NIM1 kinase transcripts is shown in H460, A2780, A375, HDF, HELA, DU145, MDA-MB231, U87-MG and BX-PC3 cell lines and in brain, colon, uterus and placenta tissues. Of particular note, human breast carcinoma cell line, MDA-MB231, and brain tissue show approximately 100 and 155 fold expression of NIM1 kinase transcripts respectively: Greater than 5-fold expression was also seen in cell lines representing carcinomas of the brain, cervix, colon, lung, ovary, and prostate and in normal uterus and placenta.

NIM1 kinase comprising the amino acid sequence of SEQ ID NO:2 is 436 amino acids in length and has a potential signal sequence from M1 to A18, potential phosphorylation sites at residues S56, T108, T114, S123, S169, S221, S282, Y288, T311, T332, T349, S359, S420, and T429. The human NIM1 kinase, as shown in Figure 4, is used as the reference for numbering the conserved residues, motifs, and subdomains of catalytic region of the kinases (SEQ ID NOs:2 and 31-36) in the alignment. The residues, motifs and subdomains are: subdomain 1 extends from G81 to V88; subdomain 2, has conserved residues A101 and K103; subdomain 3 has the invariant E121; subdomain 5 extends from E151 to E157 and shares the invariant residues M150 and Y152; subdomain 6B which represents the catalytic loop, extends from H194 to N201; subdomain 7, has the highly conserved triplet, D214, F215, and G216; subdomain 8, extends from T229 to F242 and contains the A238, P239, E240 motif; subdomain 9 has the invariant D253, subdomain 11 has the invariant R244 which interacts with the APE motif of subdomain 8. The C terminal boundary of the catalytic domain begins with H248, A249 and F250. Both PFAM and PRINTS analyses confirm kinase domains. Useful fragments of the polypeptide of SEQ ID NO:2 include a fragment of at least 6 consecutive amino acids selected from about residue 1 to about residue 295 of SEQ ID NO:2.

Figure 5 shows the human NIM1 kinase assay. hNIM1 kinase-GST autophosphorylated as shown in lanes 1, 2, and 4 and phosphorylated its substrates MBP, lane 2, and HIST.1, lane 4. In the absence of kinase, substrates MBP, lane 3, and HIST.1, lane 5, show no phosphorylation. Lane 6 shows the positive control ZAP70.

Mammalian variants of the nucleic acid molecules encoding the human NIM1 kinase were identified by using BLAST or BLAST2 (Basic Local Alignment Search Tool: Altschul *et al.* (1997) *Nucleic Acids Res* 25:3389-3402; Altschul (1993) *J Mol Evol* 36:290-300; and Altschul *et al.* (1990) *J Mol Biol* 215:403-10, with default parameters), to identify clones in the LIFESEQ or ZOOSEQ databases (Incyte Pharmaceuticals) which aligned with SEQ ID NOs:1 and 3-23. The mammalian variants are ZOOSEQ database template (Dec 99 build) 216150.1 (SEQ ID NO:24) and Incyte clones: 701925441H1 (RALITXS03), 701910632H1 (RABYUNN02), 701905514H1 (RABYUNS09), 701293826H1 (RABXNOT04), and 700949543H1 (RASPNON02) from rat and 700706950H1

WO 01/68825

PCT/US00/07715

(MNBNFNOT01) from monkey; SEQ ID NOs:25-30, respectively. These nucleic acid molecules are particularly useful for producing transgenic organisms which model human disorders and upon which potential therapeutic treatments for such disorders may be tested.

The nucleic acid molecule, SEQ ID NO:1, and fragments thereof (SEQ ID NOs:3-30) may be used in hybridization, amplification, and screening technologies to identify and distinguish among SEQ ID NO:1 and similar molecules in a sample. The human molecules and their mammalian variants may be used to produce transgenic organisms which model the human disorders and upon which potential therapeutic treatments may be tested. Toxicology studies, clinical trials, and subject/patient treatment profiles may be performed and monitored using the nucleic acid molecules, proteins, antibodies and molecules and compounds identified using the nucleic acid molecules and proteins of the present invention.

#### **Characterization and Use of the Invention**

##### cDNA libraries

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

##### Sequencing

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

WO 01/68825

PCT/US00/07715

known in the art and described in Ausubel *et al.* (1997; *Short Protocols in Molecular Biology*, John Wiley & Sons, New York NY, unit 7.7) and Meyers (1995; *Molecular Biology and Biotechnology*, Wiley VCH, New York NY, pp. 856-853).

Shotgun sequencing may also be used to complete the sequence of a particular cloned insert of interest. Shotgun strategy involves randomly breaking the original insert into segments of various sizes and cloning these fragments into vectors. The fragments are sequenced and reassembled using overlapping ends until the entire sequence of the original insert is known. Shotgun sequencing methods are well known in the art and use thermostable DNA polymerases, heat-labile DNA polymerases, and primers chosen from representative regions flanking the nucleic acid molecules of interest. Incomplete assembled sequences are inspected for identity using various algorithms or programs such as CONSED (Gordon (1998) *Genome Res* 8:195-202) which are well known in the art. Contaminating sequences including vector or chimeric sequences or deleted sequences can be removed or restored, respectively, organizing the incomplete assembled sequences into finished sequences.

#### Extension of a Nucleic Acid Sequence

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

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

##### Hybridization

The nucleic acid molecule and fragments thereof can be used in hybridization technologies for various purposes. A probe may be designed or derived from unique regions such as the 5' regulatory region or from a nonconserved region (i.e., 5' or 3' of the nucleotides encoding the conserved catalytic domain of the protein) and used in protocols to identify naturally occurring molecules encoding the human NIM1 kinase, allelic variants, or related molecules. The probe may be DNA or RNA, may be single stranded and should have at least 50% sequence identity to any of the nucleic acid sequences, SEQ ID NOs:3-30. Hybridization probes may be produced using oligolabeling, nick translation, end-labeling, or PCR amplification in the presence of a reporter molecule. A vector containing the nucleic acid molecule or a fragment thereof may be used to produce an mRNA probe *in vitro* by addition of an RNA polymerase and labeled nucleotides. These procedures may be conducted using commercially available kits such as those provided by Amersham Pharmacia Biotech.

WO 01/68825

PCT/US00/07715

The stringency of hybridization is determined by G+C content of the probe, salt concentration, and temperature. In particular, stringency can be increased by reducing the concentration of salt or raising the hybridization temperature. In solutions used for some membrane based hybridizations, addition of an organic solvent such as formamide allows the reaction to occur at a lower temperature.

5 Hybridization can be performed at low stringency with buffers, such as 5xSSC with 1% sodium dodecyl sulfate (SDS) at 60° C, which permits the formation of a hybridization complex between nucleic acid sequences that contain some mismatches. Subsequent washes are performed at higher stringency with buffers such as 0.2xSSC with 0.1% SDS at either 45° C (medium stringency) or 68° C (high stringency). At high stringency, hybridization complexes will remain stable only where the

10 nucleic acid molecules are completely complementary. In some membrane-based hybridizations, preferably 35% or most preferably 50%, formamide can be added to the hybridization solution to reduce the temperature at which hybridization is performed, and background signals can be reduced by the use of other detergents such as Sarkosyl or Triton X-100 and a blocking agent such as denatured salmon sperm DNA. Selection of components and conditions for hybridization are well known to those skilled in the art and are reviewed in Ausubel (*supra*) and Sambrook *et al.* (1989) *Molecular Cloning, A Laboratory Manual*, Cold Spring Harbor Press, Plainview NY.

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

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

#### Expression

A multitude of nucleic acid molecules encoding NIM1 kinase may be cloned into a vector and used to express the protein, or portions thereof, in host cells. The nucleic acid sequence can be engineered by such methods as DNA shuffling (*USPN* 5,830,721) and site-directed mutagenesis to

35 create new restriction sites, alter glycosylation patterns, change codon preference to increase

WO 01/68825

PCT/US00/07715

expression in a particular host, produce splice variants, extend half-life, and the like. The expression vector may contain transcriptional and translational control elements (promoters, enhancers, specific initiation signals, and polyadenylated 3' sequence) from various sources which have been selected for their efficiency in a particular host. The vector, nucleic acid molecule, and regulatory elements are combined using *in vitro* recombinant DNA techniques, synthetic techniques, and/or *in vivo* genetic recombination techniques well known in the art and described in Sambrook (*supra*, ch. 4, 8, 16 and 17).

A variety of host systems may be transformed with an expression vector. These include, but are not limited to, bacteria transformed with recombinant bacteriophage, plasmid, or cosmid DNA expression vectors; yeast transformed with yeast expression vectors; insect cell systems transformed with baculovirus expression vectors; plant cell systems transformed with expression vectors containing viral and/or bacterial elements, or animal cell systems (Ausubel *supra*, unit 16). For example, an adenovirus transcription/translation complex may be utilized in mammalian cells. After sequences are ligated into the E1 or E3 region of the viral genome, the infective virus is used to transform and express the protein in host cells. The Rous sarcoma virus enhancer or SV40 or EBV-based vectors may also be used for high-level protein expression.

Routine cloning, subcloning, and propagation of nucleic acid sequences can be achieved using the multifunctional pBLUESCRIPT vector (Stratagene, La Jolla CA) or pSPORT1 plasmid (Life Technologies). Introduction of a nucleic acid sequence into the multiple cloning site of these vectors disrupts the *lacZ* gene and allows colorimetric screening for transformed bacteria. In addition, these vectors may be useful for *in vitro* transcription, dideoxy sequencing, single strand rescue with helper phage, and creation of nested deletions in the cloned sequence.

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

The host cell may be chosen for its ability to modify a recombinant protein in a desired fashion. Such modifications include acetylation, carboxylation, glycosylation, phosphorylation,

WO 01/68825

PCT/US00/07715

lipidation, acylation and the like. Post-translational processing which cleaves a "prepro" form may also be used to specify protein targeting, folding, and/or activity. Different host cells available from the ATCC which have specific cellular machinery and characteristic mechanisms for post-translational activities may be chosen to ensure the correct modification and processing of the recombinant protein.

5 Recovery of Proteins from Cell Culture

Heterologous moieties engineered into a vector for ease of purification include glutathione S-transferase (GST), calmodulin binding peptide (CBP), 6-His, FLAG, MYC, and the like. GST, CBP, and 6-His are purified using commercially available affinity matrices such as immobilized glutathione, calmodulin, and metal-chelate resins, respectively. FLAG and MYC are purified using commercially available monoclonal and polyclonal antibodies. A proteolytic cleavage site may be located between the desired protein sequence and the heterologous moiety for ease of separation following purification. Methods for recombinant protein expression and purification are discussed in Ausubel (supra, unit 16) and are commercially available.

10 Chemical Synthesis of Peptides

15 Proteins or portions thereof may be produced not only by recombinant methods, but also by using chemical methods well known in the art. Solid phase peptide synthesis may be carried out in a batchwise or continuous flow process which sequentially adds  $\alpha$ -amino- and side chain-protected amino acid residues to an insoluble polymeric support via a linker group. A linker group such as methylamine-derivatized polyethylene glycol is attached to poly(styrene-co-divinylbenzene) to form the support resin. The amino acid residues are N- $\alpha$ -protected by acid labile Boc (t-butylloxycarbonyl) or base-labile Fmoc (9-fluorenylmethoxycarbonyl). The carboxyl group of the protected amino acid is coupled to the amine of the linker group to anchor the residue to the solid phase support resin.

Trifluoroacetic acid or piperidine are used to remove the protecting group in the case of Boc or Fmoc, respectively. Each additional amino acid is added to the anchored residue using a coupling agent or pre-activated amino acid derivative, and the resin is washed. The full length peptide is synthesized by sequential deprotection, coupling of derivitized amino acids, and washing with dichloromethane and/or N, N-dimethylformamide. The peptide is cleaved between the peptide carboxy terminus and the linker group to yield a peptide acid or amide. (Novabiochem 1997/98 Catalog and Peptide Synthesis Handbook, San Diego CA pp. S1-S20). Automated synthesis may also be carried out on machines such as the ABI 431A peptide synthesizer (PE Biosystems). A protein or portion thereof may be substantially purified by preparative high performance liquid chromatography and its composition confirmed by amino acid analysis or by sequencing (Creighton (1984) Proteins, Structures and Molecular Properties, WH Freeman, New York NY).

20 Preparation and Screening of Antibodies

35 Various hosts including goats, rabbits, rats, mice, humans, and others may be immunized by

WO 01/68825

PCT/US00/07715

injection with human NIM1 kinase or any portion thereof. Adjuvants such as Freund's, mineral gels, and surface active substances such as lysolecithin, pluronic polyols, polyanions, peptides, oil emulsions, keyhole limpet hemacyanin (KLH), and dinitrophenol may be used to increase immunological response. The oligopeptide, peptide, or portion of protein used to induce antibodies should consist of at least about five amino acids, more preferably ten amino acids, which are identical to a portion of the natural protein. Oligopeptides may be fused with proteins such as KLH in order to produce antibodies to the chimeric molecule.

Monoclonal antibodies may be prepared using any technique which provides for the production of antibodies by continuous cell lines in culture. These include, but are not limited to, the hybridoma technique, the human B-cell hybridoma technique, and the EBV-hybridoma technique. (See, e.g., Kohler *et al.* (1975) *Nature* 256:495-497; Kozbor *et al.* (1985) *J. Immunol Methods* 81:31-42; Cote *et al.* (1983) *Proc Natl Acad Sci* 80:2026-2030; and Cole *et al.* (1984) *Mol Cell Biol* 62:109-120.)

Alternatively, techniques described for the production of single chain antibodies may be adapted, using methods known in the art, to produce epitope specific single chain antibodies. Antibody fragments which contain specific binding sites for epitopes of the human NIM1 kinase may also be generated. For example, such fragments include; but are not limited to, F(ab)<sub>2</sub> fragments produced by pepsin digestion of the antibody molecule and Fab fragments generated by reducing the disulfide bridges of the F(ab)<sub>2</sub> 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 *et al.* (1989) *Science* 246:1275-1281.)

The human NIM1 kinase or a portion thereof may be used in screening assays of phagemid or B-lymphocyte immunoglobulin libraries to identify antibodies having the desired specificity. Numerous protocols for competitive binding or immunoassays using either polyclonal or monoclonal antibodies with established specificities are well known in the art. Such immunoassays typically involve the measurement of complex formation between the protein and its specific antibody. A two-site, monoclonal-based immunoassay utilizing monoclonal antibodies reactive to two non-interfering epitopes is preferred, but a competitive binding assay may also be employed (Pound (1998) *Immunochemical Protocols*, Humana Press, Totowa NJ).

#### Labeling of Molecules for Assay

A wide variety of labels and conjugation techniques are known by those skilled in the art and may be used in various nucleic acid, amino acid, and antibody assays. Synthesis of labeled molecules may be achieved using Promega (Madison WI) or Amersham Pharmacia Biotech kits for incorporation of a labeled nucleotide such as <sup>32</sup>P-dCTP, Cy3-dCTP or Cy5-dCTP or amino acid such as <sup>35</sup>S-methionine. Nucleotides and amino acids may be directly labeled with a variety of substances

WO 01/68825

PCT/US00/07715

including fluorescent, chemiluminescent, or chromogenic agents, and the like, by chemical conjugation to amines, thiols and other groups present in the molecules using reagents such as BIODIPY or FITC (Molecular Probes, Eugene OR).

#### DIAGNOSTICS

5 The nucleic acid molecules, fragments, oligonucleotides, complementary RNA and DNA molecules, and PNAs and may be used to detect and quantify altered gene expression, absence/presence vs. excess, expression of mRNAs or to monitor mRNA levels during therapeutic intervention. Similarly antibodies which specifically bind the human NIM1 kinase may be used to  
10 quantitate the protein. Disorders associated with altered expression include akathisia, Alzheimer's disease, amnesia, amyotrophic lateral sclerosis, anxiety, hereditary ataxias, cerebral palsy, dementia, dermatomyositis, dystonias, Down's syndrome, epilepsy, ischemic cerebrovascular disease, cerebelloretinal hemangioblastomatosis, Huntington's disease, bacterial and viral meningitis, multiple sclerosis, muscular dystrophy, myasthenia gravis, cerebral neoplasms, neurofibromatosis, Parkinson's  
15 disease, Pick's disease, polymyositis, retinitis pigmentosa, schizophrenia, stroke, and cancers 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,  
20 penis, prostate, salivary glands, skin, spleen, testis, thymus, thyroid, and uterus. The diagnostic assay may use hybridization or amplification technology to compare gene expression in a biological sample from a patient to standard samples in order to detect altered gene expression. Qualitative or quantitative methods for this comparison are well known in the art.

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

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

WO 01/68825

PCT/US00/07715

particular condition is used to diagnose that condition.

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

#### Immunological Methods

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

#### **THERAPEUTICS**

Chemical and structural similarity, in the context of the kinase catalytic domain, exists  
between regions of human NIM1 kinase (SEQ ID NO:2) and the kinases shown in Figures 4A-4C  
(SEQ ID NOs: 31-35. In addition, gene expression is highly associated with brain, prostate, and breast  
20 and with cancers as shown in Figures 2 and 3. The human NIM1 kinase appears to play a role in disorders of the brain and nervous system, particularly brain cancer, stroke, epilepsy and Huntington's disease, and cancers of the brain, cervix, colon, lung, ovary, and prostate. In the treatment of conditions associated with highly increased expression such as brain and breast cancers, it is desirable to decrease expression or protein activity.

In one embodiment, an inhibitor, antagonist or antibody of human NIM1 kinase may be  
administered to a subject to treat or prevent a condition associated with increased expression or activity. Examples of such conditions include, but are not limited to, akathisia, Alzheimer's disease,  
25 amnesia, amyotrophic lateral sclerosis, anxiety, hereditary ataxias, cerebral palsy, dementia, dermatomyositis, dystonias, Down's syndrome, epilepsy, ischemic cerebrovascular disease, cerebelloretinal hemangioblastomatosis, Huntington's disease, bacterial and viral meningitis, multiple sclerosis, muscular dystrophy, myasthenia gravis, cerebral neoplasms, neurofibromatosis, Parkinson's disease, Pick's disease, polymyositis, retinitis pigmentosa, schizophrenia, stroke, and cancers such as adenocarcinoma, leukemia, lymphoma, melanoma, myeloma, sarcoma, teratocarcinoma, and, in  
30 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,

WO 01/68825

PCT/US00/07715

penis, prostate, salivary glands, skin, spleen, testis, thymus, thyroid, and uterus.

In another embodiment, a pharmaceutical composition comprising an inhibitor, antagonist or antibody of NIM1 kinase in conjunction with a pharmaceutical carrier may be administered to a subject to treat or prevent a condition associated with altered expression or activity of the endogenous protein including, but not limited to, those provided above.

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

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

#### Modification of Gene Expression Using Nucleic Acids

Gene expression may be modified by designing complementary or antisense molecules (DNA, RNA, or PNA) to the control, 5', 3', or other regulatory regions of the gene encoding NIM1 kinase. Oligonucleotides designed with reference to the transcription initiation site are preferred. Similarly, inhibition can be achieved using triple helix base-pairing which inhibits the binding of polymerases, transcription factors, or regulatory molecules (Gee *et al.* In: Huber and Carr (1994) Molecular and Immunologic Approaches, Futura Publishing, Mt. Kisco NY, pp. 163-177). A complementary molecule may also be designed to block translation by preventing binding between ribosomes and mRNA. In one alternative, a library of nucleic acid molecules or fragments thereof may be screened to identify those which specifically bind a regulatory, nontranslated sequence.

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

Complementary nucleic acids and ribozymes of the invention may be prepared via recombinant expression, *in vitro* or *in vivo*, or using solid phase phosphoramidite chemical synthesis. In addition, RNA molecules may be modified to increase intracellular stability and half-life by

WO 01/68825

PCT/US00/07715

addition of flanking sequences at the 5' and/or 3' ends of the molecule or by the use of phosphorothioate or 2'-O-methyl rather than phosphodiesterase linkages within the backbone of the molecule. Modification is inherent in the production of PNAs and can be extended to other nucleic acid molecules. Either the inclusion of nontraditional bases such as inosine, queosine, and wybutosine, and or the modification of adenine, cytidine, guanine, thymine, and uridine with acetyl-, methyl-, thio- groups renders the molecule less available to endogenous endonucleases.

#### Screening and Purification Assays

The nucleic acid molecule encoding the human Nim1 kinase may be used to screen a library of molecules or compounds for specific binding affinity. The libraries may be aptamers, DNA molecules, RNA molecules, PNAs, peptides, proteins such as transcription factors, enhancers, repressors, and other ligands which regulate the activity, replication, transcription, or translation of the nucleic acid molecule in the biological system. The assay involves combining the nucleic acid molecule or a fragment thereof with the library of molecules under conditions allowing specific binding, and detecting specific binding to identify at least one molecule which specifically binds the single stranded or double stranded nucleic acid molecule.

In one embodiment, the polynucleotide of the invention may be incubated with a library of isolated and purified molecules or compounds and binding activity determined by methods well known in the art, e.g., a gel-retardation assay (USPN 6,010,849) or a reticulocyte lysate transcriptional assay.

In another embodiment, the polynucleotide may be incubated with nuclear extracts from biopsied and/or cultured cells and tissues. Specific binding between the polynucleotide and a molecule or compound in the nuclear extract is initially determined by gel shift assay and may be later confirmed by raising antibodies against that molecule or compound. When these antibodies are added into the assay, they cause a supershift in the gel-retardation assay.

In another embodiment, the polynucleotide may be used to purify a molecule or compound using affinity chromatography methods well known in the art. In one embodiment, the polynucleotide is chemically reacted with cyanogen bromide groups on a polymeric resin or gel. Then a sample is passed over and reacts with or binds to the polynucleotide. The molecule or compound which is bound to the polynucleotide may be released from the polynucleotide by increasing the salt concentration of the flow-through medium and collected.

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

In a preferred embodiment, human NIM1 kinase or a portion thereof may be used to screen

WO 01/68825

PCT/US00/07715

libraries of molecules or compounds in any of a variety of screening assays. The portion of the protein employed in such screening may be free in solution, affixed to an abiotic or biotic substrate (e.g. borne on a cell surface), or located intracellularly. For example, in one method, viable or fixed prokaryotic host cells that are stably transformed with recombinant nucleic acids that have expressed and positioned a polypeptide on their cell surface can be used in screening assays. The cells are screened against libraries or a plurality of ligands and the specificity of binding or formation of complexes between the expressed polypeptide and the ligand may be measured. Specific binding between the protein and molecule may be measured. Depending on the kind of library being screened, the assay may be used to identify DNA molecules, RNA molecules, peptide nucleic acids, peptides, proteins, mimetics, agonists, antagonists, antibodies, immunoglobulins, inhibitors, and drugs or any other ligand, which specifically binds the protein.

In one aspect, this invention contemplates a method for high throughput screening using very small assay volumes and very small amounts of test compound as described in USPN 5,876,946, incorporated herein by reference. This method is used to screen large numbers of molecules and compounds via specific binding. In another aspect, this invention also contemplates the use of competitive drug screening assays in which neutralizing antibodies capable of binding the polypeptide specifically compete with a test compound capable of binding to the polypeptide or oligopeptide or portion thereof. Molecules or compounds identified by screening may be used in a mammalian model system to evaluate their toxicity, diagnostic, or therapeutic potential.

#### 20 Pharmacology

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

A therapeutically effective dose refers to that amount of protein or inhibitor which ameliorates the symptoms or condition. Therapeutic efficacy and toxicity of such agents may be determined by standard pharmaceutical procedures in cell cultures or experimental animals, e.g., ED50 (the dose therapeutically effective in 50% of the population) and LD50 (the dose lethal to 50% of the population). The dose ratio between toxic and therapeutic effects is the therapeutic index, and it may be expressed as the ratio, LD50/ED50. Pharmaceutical compositions which exhibit large therapeutic indexes are preferred. The data obtained from cell culture assays and animal studies are used in formulating a range of dosage for

WO 01/68825

PCT/US00/07715

human use.

#### Model Systems

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

#### Toxicology

Toxicology is the study of the effects of agents on living systems. The majority of toxicity studies are performed on rats or mice. Observation of qualitative and quantitative changes in physiology, behavior, homeostatic processes, and lethality in the rats or mice are used to generate a toxicity profile and to assess potential consequences on human health following exposure to the agent. Genetic toxicology identifies and analyzes the effect of an agent on the rate of endogenous, spontaneous, and induced genetic mutations. Genotoxic agents usually have common chemical or physical properties that facilitate interaction with nucleic acids and are most harmful when chromosomal aberrations are transmitted to progeny. Toxicological studies may identify agents that increase the frequency of structural or functional abnormalities in the tissues of the progeny if administered to either parent before conception, to the mother during pregnancy, or to the developing organism. Mice and rats are most frequently used in these tests because their short reproductive cycle allows the production of the numbers of organisms needed to satisfy statistical requirements.

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

Subchronic toxicity tests are based on the repeated administration of an agent. Rat and dog are commonly used in these studies to provide data from species in different families. With the exception of carcinogenesis, there is considerable evidence that daily administration of an agent at high-dose concentrations for periods of three to four months will reveal most forms of toxicity in adult animals.

Chronic toxicity tests, with a duration of a year or more, are used to demonstrate either the absence of toxicity or the carcinogenic potential of an agent. When studies are conducted on rats, a

WO 01/68825

PCT/US00/07715

minimum of three test groups plus one control group are used, and animals are examined and monitored at the outset and at intervals throughout the experiment.

#### Transgenic Animal Models

Transgenic rodents that over-express or under-express a gene of interest may be inbred and used to model human diseases or to test therapeutic or toxic agents. (See, e.g., USPN 5,175,383 and USPN 5,767,337.) In some cases, the introduced gene may be activated at a specific time in a specific tissue type during fetal or postnatal development. Expression of the transgene is monitored by analysis of phenotype, of tissue-specific mRNA expression, or of serum and tissue protein levels in transgenic animals before, during, and after challenge with experimental drug therapies.

#### 10 Embryonic Stem Cells

Embryonic (ES) stem cells isolated from rodent embryos retain the potential to form embryonic tissues. When ES cells are placed inside a carrier embryo, they resume normal development and contribute to tissues of the live-born animal. ES cells are the preferred cells used in the creation of experimental knockout and knockin rodent strains. Mouse ES cells, such as the mouse 129/SvJ cell line, are derived from the early mouse embryo and are grown under culture conditions well known in the art. Vectors used to produce a transgenic strain contain a disease gene candidate and a marker gene, the latter serves to identify the presence of the introduced disease gene. The vector is transformed into ES cells by methods well known in the art, and 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.

ES cells derived from human blastocysts may be manipulated *in vitro* to differentiate into at least eight separate cell lineages. These lineages are used to study the differentiation of various cell types and tissues *in vitro*, and they include endoderm, mesoderm, and ectodermal cell types which differentiate into, for example, neural cells, hematopoietic lineages, and cardiomyocytes.

#### Knockout Analysis

In gene knockout analysis, a region of a mammalian gene is enzymatically modified to include a non-mammalian gene such as the neomycin phosphotransferase gene (neo; Capecchi (1989) Science 244:1288-1292). The modified gene is transformed into cultured ES cells and integrates into the endogenous genome by homologous recombination. The inserted sequence disrupts transcription and translation of the endogenous gene. Transformed cells are injected into rodent blastulae, and the blastulae are implanted into pseudopregnant dams. Transgenic progeny are crossbred to obtain homozygous inbred lines which lack a functional copy of the mammalian gene. In one example, the mammalian gene is a human gene.

#### 35 Knockin Analysis

WO 01/68825

PCT/US00/07715

ES cells can be used to create knockin humanized animals (pigs) or transgenic animal models (mice or rats) of human diseases. With knockin technology, a region of a human gene is injected into animal ES cells, and the human sequence integrates into the animal cell genome. Transformed cells are injected into blastulae and the blastulae are implanted as described above. Transgenic progeny or  
5 inbred lines are studied and treated with potential pharmaceutical agents to obtain information on treatment of the analogous human condition. These methods have been used to model several human diseases.

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

#### EXAMPLES

It is to be understood that this invention is not limited to the particular machines, materials and methods described. Although particular embodiments are described, equivalent embodiments may be  
15 used to practice the invention. The described embodiments are not intended to limit the scope of the invention which is limited only by the appended claims. The examples below are provided to illustrate the subject invention and are not included for the purpose of limiting the invention. For purposes of  
20 example, preparation of the human cerebellum (CRBLNOT01), prostate (PROSBPT03), and normalized brain (BRAINON01) libraries will be described.

#### 20 I cDNA Library Construction

##### Cerebellum

The tissue used for cerebellum library construction was obtained from a 69 year-old, Caucasian male (RT95-05-0301; International Institute for Advanced Medicine, Exton PA). The frozen tissue was homogenized and lysed using a POLYTRON homogenizer (PT-3000; Brinkmann  
25 Instruments, Westbury NJ). The reagents and extraction procedures were used as supplied in the RNA Isolation kit (Stratagene). The lysate was centrifuged over a 5.7 M CsCl cushion using an SW28 rotor in an L8-70M ultracentrifuge (Beckman Coulter, Fullerton CA ) for 18 hr at 25,000 rpm at ambient temperature. The RNA was extracted twice with phenol chloroform, pH 8.0, and once with acid phenol, pH 4.0; precipitated using 0.3 M sodium acetate and 2.5 volumes of ethanol; resuspended in  
30 water; and treated with DNase for 15 min at 37C. The RNA was isolated with the OLIGOTEX kit (Qiagen, Chatsworth CA) and used to construct the cDNA library.

##### Prostate

Diseased prostate tissue removed from a 59-year-old Caucasian male during a radical prostatectomy was used to construct the PROSBPT03 library. Pathology for the tumor indicated  
35 adenocarcinoma, Gleason grade 3+3, with microscopic foci involving the right and left sides

WO 01/68825

PCT/US00/07715

peripherally. The tumor was confined and did not involve the capsule. High-grade prostatic intraepithelial neoplasia was identified on the right side peripherally. The patient presented with elevated prostate specific antigen (PSA). Family history included cerebrovascular disease in both parents and prostate cancer in a sibling.

5 To construct the cDNA library, 2.4 micrograms of polyA RNA was used according to the recommended protocols in the SUPERSCRIPT Plasmid system (Life Technologies). First strand cDNA synthesis was accomplished using oligo d(T) priming, and second strand synthesis was performed using a combination of DNA polymerase I, *E. coli* ligase, and RNase H. The cDNA was blunted with T4 polymerase, and a Sal I linker was added to the blunt end of the cDNA. The Sal I adapted, double stranded cDNAs were digested with Not I and fractionated on a SEPHAROSE CL4B column (Amersham Pharmacia Biotech).

10 Those cerebellum cDNAs exceeding 400 bp were ligated into pSPORT I plasmid which was subsequently transformed into DH5 $\alpha$  competent cells (Life Technologies). Those prostate cDNAs exceeding 400 bp were ligated into the NotI and EcoRI sites of the pINCY plasmid (Incyte Pharmaceuticals) which was subsequently transformed into competent DH5 $\alpha$  or ELECTROMAX DH10B competent cells (Life Technologies).

#### Normalized Brain

For purposes of example, the normalization of the human brain library (BRAINON01) is described. About  $4.9 \times 10^6$  independent clones of the BRAINOT03 plasmid library in *E. coli* strain 20 DH12S competent cells (Life Technologies) were grown in liquid culture under carbenicillin (25 mg/l) and methicillin (1 mg/ml) selection following transformation by electroporation. To reduce the number of excess cDNA copies according to their abundance levels in the library, the cDNA library was normalized in a single round according to the procedure of Soares *et al.* (1994, Proc Natl Acad Sci 91:9228-9232), with the following modifications. The primer to template ratio in the primer extension 25 reaction was increased from 2:1 to 10:1. The dNTP concentration in the reaction was reduced to 150  $\mu$ M for each dNTP to allow the generation of longer (400 to 1000 nt) primer extension products. The reannealing hybridization was extended from 13 to 48 hr. The single stranded DNA circles of the normalized library were purified by hydroxyapatite chromatography and converted to partially double-stranded by random priming, followed by electroporation into *E. coli* strain DH10B competent cells 30 (Life Technologies).

#### II Construction of pINCY Plasmid

The plasmid was constructed by digesting the pSPORT1 plasmid (Life Technologies) with EcoRI restriction enzyme (New England Biolabs, Beverly MA) and filling the overhanging ends using Klenow enzyme (New England Biolabs) and 2'-deoxynucleotide 5'-triphosphates (dNTPs). The

WO 01/68825

PCT/US00/07715

plasmid was self-ligated and transformed into the bacterial host, *E. coli* strain JM109.

An intermediate plasmid produced by the bacteria (pSPORT 1- $\Delta$ RI) showed no digestion with EcoRI and was digested with Hind III (New England Biolabs) and the overhanging ends were again filled in with Klenow and dNTPs. A linker sequence was phosphorylated, ligated onto the 5' blunt end, digested with EcoRI, and self-ligated. Following transformation into JM109 host cells, plasmids were isolated and tested for preferential digestibility with EcoRI, but not with Hind III. A single colony that met this criteria was designated pINCY plasmid.

After testing the plasmid for its ability to incorporate cDNAs from a library prepared using NotI and EcoRI restriction enzymes, several clones were sequenced; and a single clone containing an insert of approximately 0.8 kb was selected from which to prepare a large quantity of the plasmid. After digestion with NotI and EcoRI, the plasmid was isolated on an agarose gel and purified using a QIAQUICK column (Qiagen) for use in library construction.

### III Isolation and Sequencing of cDNA Clones

Plasmid DNA was released from the cells and purified using either the MINIPREP Kit (Edge Biosystems, Gaithersburg MD) or the REAL Prep 96 plasmid kit (Qiagen). This kit consists of a 96-well block with reagents for 960 purifications. The recommended protocol was employed except for the following changes: 1) the bacteria were cultured in 1 ml of sterile Terrific Broth (Life Technologies) with carbenicillin at 25 mg/l and glycerol at 0.4%; 2) after inoculation, the cells were cultured for 19 hours and then lysed with 0.3 ml of lysis buffer; and 3) following isopropanol precipitation, the plasmid DNA pellet was resuspended in 0.1 ml of distilled water. After the last step in the protocol, samples were transferred to a 96-well block for storage at 4C.

The cDNAs were prepared for sequencing using the MICROLAB 2200 system (Hamilton, Reno NV) in combination with the DNA ENGINE thermal cyclers (MJ Research). The cDNAs were sequenced by the method of Sanger and Coulson (1975; *J Mol Biol* 94:441-448) using an ABI PRISM 377 sequencing system (PE Biosystems) or the MEGABASE 1000 DNA sequencing system (Amersham Pharmacia Biotech). Most of the isolates were sequenced according to standard ABI protocols and kits (PE Biosystems) with solution volumes of 0.25x - 1.0x concentrations. In the alternative, cDNAs were sequenced using solutions and dyes from Amersham Pharmacia Biotech.

### IV Extension of cDNA Sequences

The nucleic acid molecules were extended using a cDNA clone and oligonucleotide primers. One primer was synthesized to initiate 5' extension of the known fragment, and the other, to initiate 3' extension of the known fragment. The initial primers were designed using OLIGO 4.06 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 target sequence at temperatures of about 68C to about 72C. Any stretch of nucleotides that would result in hairpin

WO 01/68825

PCT/US00/07715

structures and primer-primer dimerizations was avoided.

Selected cDNA libraries were used as templates to extend the sequence. If more than one extension was necessary or desired, additional or nested sets of primers were designed. Preferred libraries are ones that have been size-selected to include larger cDNAs. Also, random primed libraries are preferred because they will contain more sequences with the 5' and upstream regions of genes. A randomly primed library is particularly useful if an oligo d(T) library does not yield a full length cDNA. Genomic libraries are useful for extension 5' of the promoter binding region in order to obtain regulatory elements.

High fidelity amplification was obtained by PCR using methods such as that taught in USPN 5,932,451. PCR was performed in 96-well plates using the DNA ENGINE thermal cycler (MJ Research). The reaction mix contained DNA template, 200 nmol of each primer, reaction buffer containing  $Mg^{2+}$ ,  $(NH_4)_2SO_4$ , and  $\beta$ -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 (Incyte Pharmaceuticals): Step 1: 94C, three min; Step 2: 94C, 15 sec; Step 3: 60C, one min; Step 4: 68C, two min; Step 5: Steps 2, 3, and 4 repeated 20 times; Step 6: 68C, five min; Step 7: storage at 4C. In the alternative, the parameters for primer pair T7 and SK+ (Stratagene) were as follows: Step 1: 94C, three min; Step 2: 94C, 15 sec; Step 3: 57C, one min; Step 4: 68C, two min; Step 5: Steps 2, 3, and 4 repeated 20 times; Step 6: 68C, five min; Step 7: storage at 4C.

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

The extended nucleotide sequences were desalted, concentrated, transferred to 384-well plates, digested with CviII cholera virus endonuclease (Molecular Biology Research, Madison WI), and sonicated or sheared prior to religation into pUC18 vector (Amersham Pharmacia Biotech). For shotgun sequences, the digested nucleotide sequences were separated on low concentration (0.6 to 0.8%) agarose gels, fragments were excised, and the agar was digested with AGARACE enzyme (Promega). Extended clones were religated using T4 DNA ligase (New England Biolabs) into pUC18 vector (Amersham Pharmacia Biotech), treated with Pfu DNA polymerase (Stratagene) to fill-in restriction site overhangs, and transfected into *E. coli* competent cells. Transformed cells were selected on antibiotic-containing media, and individual colonies were picked and cultured overnight at

WO 01/68825

PCT/US00/07715

37C in 384-well plates in LB/2x carbenicillin liquid media.

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

#### V Homology Searching of cDNA Clones and Their Deduced Polypeptides

The nucleic acid molecules of the Sequence Listing or their deduced amino acid sequences were used to query databases such as GenBank, SwissProt, BLOCKS, and the like. These databases that contain previously identified and annotated sequences or domains were searched using BLAST or BLAST 2 (Altschul et al. *supra*; Altschul, *supra*) to produce alignments and to determine which sequences were exact matches or homologs. The alignments were to sequences of prokaryotic (bacterial) or eukaryotic (animal, fungal, or plant) origin. Alternatively, algorithms such as the one described in Smith and Smith (1992, Protein Engineering 5:35-51) could have been used to deal with primary sequence patterns and secondary structure gap penalties. All of the sequences disclosed in this application have lengths of at least 49 nucleotides, and no more than 12% uncalled bases (where N is recorded rather than A, C, G, or T).

As detailed in Karlin (*supra*), BLAST matches between a query sequence and a database sequence were evaluated statistically and only reported when they satisfied the threshold of  $10^{-25}$  for nucleotides and  $10^{-14}$  for peptides. Homology was also evaluated by product score calculated as follows: the % nucleotide or amino acid identity [between the query and reference sequences] in BLAST is multiplied by the % maximum possible BLAST score [based on the lengths of query and reference sequences] and then divided by 100. In comparison with hybridization procedures used in the laboratory, the electronic stringency for an exact match was set at 70, and the conservative lower limit for an exact match was set at approximately 40 (with 1-2% error due to uncalled bases).

The BLAST software suite, freely available sequence comparison algorithms (NCBI, Bethesda MD; <http://www.ncbi.nlm.nih.gov/gorf/bl2.html>) includes various sequence analysis programs including "blastn", that is used to align a known nucleic acid molecules, BLAST 2 that is used for direct pairwise comparison of either nucleic or amino acid molecules. BLAST programs are commonly used with gap and other parameters set to default settings, e.g.: Matrix: BLOSUM62; Reward for match: 1; Penalty for mismatch: -2; Open Gap: 5 and Extension Gap: 2 penalties; Gap x

WO 01/68825

PCT/US00/07715

drop-off: 50; Expect: 10; Word Size: 11; and Filter: on. Identity or similarity may be measured over the entire length of a sequence or some smaller portion thereof. Brenner *et al.* (1998; Proc Natl Acad Sci 95:6073-6078, incorporated herein by reference) analyzed the BLAST for its ability to identify structural homologs by sequence identity and found 30% identity is a reliable threshold for sequence alignments of at least 150 residues and 40%, for alignments of at least 70 residues.

The mammalian nucleic acid molecule of this application were compared with assembled consensus sequences or templates found in the LIFESEQ GOLD database. Component sequences from cDNA, extension, full length, and shotgun sequencing projects were subjected to PHRED analysis and assigned a quality score. All sequences with an acceptable quality score were subjected to various pre-processing and editing pathways to remove low quality 3' ends, vector and linker sequences, polyA tails, Alu repeats, mitochondrial and ribosomal sequences, and bacterial contamination sequences. Edited sequences had to be at least 50 bp in length, and low-information sequences and repetitive elements such as dinucleotide repeats, Alu repeats, and the like, were replaced by "Ns" or masked.

Edited sequences were subjected to assembly procedures in which the sequences were assigned to gene bins. Each sequence could only belong to one bin, and sequences in each bin were assembled to produce a template. Newly sequenced components were added to existing bins using BLAST and CROSSMATCH. To be added to a bin, the component sequences had to have a BLAST quality score greater than or equal to 150 and an alignment of at least 82% local identity. The sequences in each bin were assembled using PHRAP. Bins with several overlapping component sequences were assembled using DEEP PHRAP. The orientation of each template was determined based on the number and orientation of its component sequences.

Bins were compared to one another and those having local similarity of at least 82% were combined and reassembled. Bins having templates with less than 95% local identity were split. Templates were subjected to analysis by STITCHER/EXON MAPPER algorithms that analyze the probabilities of the presence of splice variants, alternatively spliced exons, splice junctions, differential expression of alternative spliced genes across tissue types or disease states, and the like. Assembly procedures were repeated periodically, and templates were annotated using BLAST against GenBank databases such as GBpri. An exact match was defined as having from 95% local identity over 200 base pairs through 100% local identity over 100 base pairs and a homolog match as having an E-value (or probability score) of  $\leq 1 \times 10^{-8}$ . The templates were also subjected to frameshift FASTx against GENPEPT, and homolog match was defined as having an E-value of  $\leq 1 \times 10^{-8}$ . Template analysis and assembly was described in USSN 09/276,534, filed March 25, 1999.

Following assembly, templates were subjected to BLAST, motif, and other functional analyses and categorized in protein hierarchies using methods described in USSN 08/812,290 and USSN

WO 01/68825

PCT/US00/07715

08/811,758, both filed March 6, 1997; in USSN 08/947,845, filed October 9, 1997; and in USSN 09/034,807, filed March 4, 1998. Then templates were analyzed by translating each template in all three forward reading frames and searching each translation against the PFAM database of hidden Markov model-based protein families and domains using the HMMER software package (Washington University School of Medicine, St. Louis MO; <http://pfam.wustl.edu/>).

The nucleic acid molecule was further analyzed using MACDNASIS PRO software (Hitachi Software Engineering), and LASERGENE software (DNASTAR) and queried against public databases such as the GenBank rodent, mammalian, vertebrate, prokaryote, and eukaryote databases, SwissProt, BLOCKS, PRINTS, PFAM, and Prosite.

#### 10 VI Chromosome Mapping

Radiation hybrid and genetic mapping data available from public resources such as the Stanford Human Genome Center (SHGC), Whitehead Institute for Genome Research (WIGR), and Généthon are used to determine if any of the nucleic acid molecules presented in the Sequence Listing have been mapped. Any of the fragments of the nucleic acid molecule encoding NIMI kinase that have been mapped result in the assignment of all related regulatory and coding sequences mapping to the same location. The genetic map locations are described as ranges, or intervals, of human chromosomes. The map position of an interval, in cM (which is roughly equivalent to 1 megabase of human DNA), is measured relative to the terminus of the chromosomal p-arm.

#### VII Hybridization Technologies and Analyses

##### 20 Immobilization of Nucleic acid molecules on a Substrate

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

In the second method, nucleic acid molecules are amplified from bacterial vectors by thirty cycles of PCR using primers complementary to vector sequences flanking the insert. PCR amplification increases a starting concentration of 1-2 ng nucleic acid to a final quantity greater than 5 µg. Amplified nucleic acids from about 400 bp to about 5000 bp in length are purified using

WO 01/68825

PCT/US00/07715

SEPHACRYL-400 beads (Amersham Pharmacia Biotech). Purified nucleic acids are arranged on a nylon membrane manually or using a dot/slot blotting manifold and suction device and are immobilized by denaturation, neutralization, and UV irradiation as described above. Purified nucleic acids are robotically arranged and immobilized on polymer-coated glass slides using the procedure described in USPN 5,807,522. Polymer-coated slides are prepared by cleaning glass microscope slides (Corning, Acton MA) by ultrasound in 0.1% SDS and acetone, etching in 4% hydrofluoric acid (VWR Scientific Products, West Chester PA), coating with 0.05% aminopropyl silane (Sigma Aldrich) in 95% ethanol, and curing in a 110C oven. The slides are washed extensively with distilled water between and after treatments. The nucleic acids are arranged on the slide and then immobilized by exposing the array to UV irradiation using a STRATALINKER UV-crosslinker (Stratagene). Arrays are then washed at room temperature in 0.2% SDS and rinsed three times in distilled water. Non-specific binding sites are blocked by incubation of arrays in 0.2% casein in phosphate buffered saline (PBS; Tropic, Bedford MA) for 30 min at 60C; then the arrays are washed in 0.2% SDS and rinsed in distilled water as before.

15 Probe Preparation for Membrane Hybridization

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

20 Probe Preparation for Polymer Coated Slide Hybridization

Hybridization probes derived from mRNA isolated from samples are employed for screening nucleic acid molecules of the Sequence Listing in array-based hybridizations. Probe is prepared using the GEMbright kit (Incyte Pharmaceuticals) by diluting mRNA to a concentration of 200 ng in 9  $\mu$ l TE buffer and adding 5  $\mu$ l 5x buffer, 1  $\mu$ l 0.1 M DTT, 3  $\mu$ l Cy3 or Cy5 labeling mix, 1  $\mu$ l RNase inhibitor, 1  $\mu$ l reverse transcriptase, and 5  $\mu$ l 1x yeast control mRNAs. Yeast control mRNAs are synthesized by *in vitro* transcription from noncoding yeast genomic DNA (W. Lei, unpublished). As quantitative controls, one set of control mRNAs at 0.002 ng, 0.02 ng, 0.2 ng, and 2 ng are diluted into reverse transcription reaction mixture at ratios of 1:100,000, 1:10,000, 1:1000, and 1:100 (w/w) to sample mRNA respectively. To examine mRNA differential expression patterns, a second set of control

WO 01/68825

PCT/US00/07715

mRNAs are diluted into reverse transcription reaction mixture at ratios of 1:3, 3:1, 1:10, 10:1, 1:25, and 25:1 (w/w). The reaction mixture is mixed and incubated at 37C for two hr. The reaction mixture is then incubated for 20 min at 85C, and probes are purified using two successive CHROMA SPIN+TE 30 columns (Clontech, Palo Alto CA). Purified probe is ethanol precipitated by diluting probe to 90  $\mu$ l in DEPC-treated water, adding 2  $\mu$ l 1mg/ml glycogen, 60  $\mu$ l 5 M sodium acetate, and 300  $\mu$ l 100% ethanol. The probe is centrifuged for 20 min at 20,800xg, and the pellet is resuspended in 12  $\mu$ l resuspension buffer, heated to 65C for five min, and mixed thoroughly. The probe is heated and mixed as before and then stored on ice. Probe is used in high density array-based hybridizations as described below.

10 Membrane-based Hybridization

Membranes are pre-hybridized in hybridization solution containing 1% Sarkosyl and 1x high phosphate buffer (0.5 M NaCl, 0.1 M  $\text{Na}_2\text{HPO}_4$ , 5 mM EDTA, pH 7) at 55C for two hr. The probe, diluted in 15 ml fresh hybridization solution, is then added to the membrane. The membrane is hybridized with the probe at 55C for 16 hr. Following hybridization, the membrane is washed for 15 min at 25C in 1mM Tris (pH 8.0), 1% Sarkosyl, and four times for 15 min each at 25C in 1mM Tris (pH 8.0). To detect hybridization complexes, XOMAT-AR film (Eastman Kodak, Rochester NY) is exposed to the membrane overnight at -70C, developed, and examined visually.

Polymer Coated Slide-based Hybridization

Probe is heated to 65C for five min, centrifuged five min at 9400 rpm in a 5415C microcentrifuge (Eppendorf Scientific, Westbury NY), and then 18  $\mu$ l is aliquoted onto the array surface and covered with a coverslip. The arrays are transferred to a waterproof chamber having a cavity just slightly larger than a microscope slide. The chamber is kept at 100% humidity internally by the addition of 140  $\mu$ l of 5xSSC in a corner of the chamber. The chamber containing the arrays is incubated for about 6.5 hr at 60C. The arrays are washed for 10 min at 45C in 1xSSC, 0.1% SDS, and three times for 10 min each at 45C in 0.1xSSC, and dried.

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

35 Hybridization complexes are detected with a microscope equipped with an Innova 70 mixed

WO 01/68825

PCT/US00/07715

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

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

#### 25 VIII Northern Analysis

##### Electronic

Computer techniques applying BLAST were used to search for identical or related molecules in nucleotide databases such as GenBank or LIFESEQ databases (Incyte Pharmaceuticals). The product score for human and rat sequences was calculated as follows: the BLAST score is multiplied by the % nucleotide identity and the product is divided by (5 times the length of the shorter of the two sequences), such that a 100% alignment over the length of the shorter sequence gives a product score of 100. The product score takes into account both the degree of similarity between two sequences and the length of the sequence match. For example, with a product score of 40, the match will be exact within a 1% to 2% error, and with a product score of at least 70, the match will be exact. Similar or related molecules are usually identified by selecting those which show product scores between 8 and

WO 01/68825

PCT/US00/07715

40.

The results of the electronic northern performed at a product score of 70 are shown in Figure 2. Analysis involved the categorization of cDNA libraries by system, organ/tissue and cell type. The categories included cardiovascular system, connective tissue (specifically cancerous breast fibroblasts), digestive system, embryonic structures, endocrine system, exocrine glands, female and male genitalia (specifically cancerous prostate), germ cells, hemic and immune system, liver, musculoskeletal system, nervous system, pancreas, respiratory system, sense organs, skin, stomatognathic system, unclassified/mixed, and the urinary tract. For each category, the number of libraries expressing the sequence was counted and shown over the total number of libraries in that category.

#### Quantitative PCR

Quantitative PCR was used to examine expression of human NIM1 kinase in various cell lines and tissues. The cell lines, obtained from sources such as ATCC, were H460 human non-small cell lung carcinoma, A2780 human ovarian carcinoma line, A375 human melanoma cell line, HDF human dermal fibroblasts, HELA human cervix carcinoma, DU145 androgen-independent prostate carcinoma cell line, MDA-MB231 human mammary tumor cells, U87 glioblastoma tumor cells, and BX-PC3 pancreatic cancer cells. The cell lines were plated in culture dishes and grown in RPMI supplemented with 10% fetal bovine serum, 2 mM glutamine at 37C in 5% CO<sub>2</sub> until they were 90% confluent. The tissues, brain, colon, uterus and placenta, were obtained as MTN blots (Human I, II, III, and IV) from Clontech.

Northern analysis was by quantitative PCR as performed using the ABI PRISM 7700 Sequencing system and TAQMAN assay reagents (TAQMAN Universal PCR Master mix) and technology according to manufacturer instructions (all PE Biosystems). All reactions were performed in triplicate.

The primers used in the reaction included SEQ ID NOs:37-39. Relative quantification was done using 18s RNA as the standard. This linearity of this standardization procedure was described in Spiess and Ivell (1999; Biotechniques 26:46-50, incorporated herein by reference).

#### IX Complementary Nucleic Acid Molecules

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

WO 01/68825

PCT/US00/07715

complementary oligonucleotide is designed to prevent ribosomal binding to the mRNA encoding the mammalian polypeptide.

In addition to using antisense molecules constructed to interrupt transcription or translation, modifications of gene expression can be obtained by designing antisense molecules to genomic sequences (such as enhancers or introns) or even to trans-acting regulatory genes. Similarly, antisense inhibition can be achieved using Hogeboom base-pairing methodology, also known as "triple helix" base pairing. Antisense molecules involved in triple helix pairing compromise the ability of the double helix to open sufficiently for the binding of polymerases, transcription factors, or regulatory molecules.

Such antisense molecules are placed in expression vectors and used to transform preferred cells or tissues. This may include introduction of the expression vector into a cell line to test efficacy; into an organ, tumor, synovial cavity, or the vascular system for transient or short term therapy; or into a stem cell or other reproducing lineage for long term or stable gene therapy. Transient expression may last for a month or more with a non-replicating vector and for three months or more if appropriate elements for inducing vector replication are used in the transformation/expression system.

Stable transformation of appropriate dividing cells with a vector encoding the antisense molecule can produce a transgenic cell line, tissue, or organism (USPN 4,736,866). Those cells that assimilate and replicate sufficient quantities of the vector to allow stable integration also produce enough antisense molecules to compromise or entirely eliminate activity of the nucleic acid molecule encoding the mammalian protein.

#### **X Expression of NIM1 Kinase**

Expression of the human NIM1 kinase was achieved using monkey and insect cell-based expression systems. The cDNA was purified using QIAPREP spin miniprep kit and PLASMID MAXI kit (both Qiagen) accordingly to the manufacturer's instructions.

For transient expression, the cDNA was cloned into the pcDNA3.1(-)/myc-His B vector (Invitrogen, Carlsbad CA), and the vector pcDNA3-Nim1 was transformed into COSI-1 cells using CaPO<sub>4</sub> transfection kit 2-463335 (Eppendorf - 5 Prime, Boulder CO). 4x10<sup>6</sup> cells were seeded in 10 cm tissue culture dishes 24 hr before transformation. On the day of transformation, CaPO<sub>4</sub>-DNA precipitate was obtained using 500 ml 2x DNA precipitation buffer, 62 ml M CaCl<sub>2</sub>, 10 mg (10 ml) pcDNA3-Nim1, and 428 ml water. The mixture was incubated at room temperature for 20 min and then slowly added to 9 ml of culture medium (DMEM, 10% fetal calf serum, 2mM glutamine, 10 mg/ml penicillin and 10 mg/ml streptomycin). Cells were incubated at 37C at 5% CO<sub>2</sub> for 4 h. The medium was replaced with fresh medium, and cells were incubated for 48 additional hours.

Sf21 insect cells were cotransformed according instructions supplied with the BaculoGold transfection kit (BD Pharmingen, San Diego CA). 2 x 10<sup>6</sup> cells were seeded in one 6 cm tissue culture

WO 01/68825

PCT/US00/07715

dish and incubated at 27C for 15 min. 4 mg (4 ml) of pVL1392/GST-NIM1 expression vector were combined with 0.5 mg (0.5 ml) of BaculoGold DNA and incubated at room temperature for 5 min. The cells' medium (TNM-FH) was removed and replaced with 1 ml of Buffer A. The DNA mixture was diluted in 1ml of Buffer B and added drop by drop to the cotransfection plate. After the plate was incubated at 27C for 4 hours, the medium was replaced, and the cells were incubated for 5 days. The recombinant viruses were subjected to three cycles of amplification to obtain a 10<sup>7</sup> pfu/ml viral stock. 1.2 x 10<sup>7</sup> Sf21 insect cells were infected with 10 ml of viral stock and incubated at 27C. After 3 days, the cells were lysed, and the protein was purified.

#### XI Protein Purification

##### 10 His-purification

COS-1 cells were centrifuged at 1000 rpm, resuspended in 4 ml of lysis buffer (5 mM imidazole + 0.5 mM NaCl + 20 mM Na<sub>2</sub>HPO<sub>4</sub> + Complete protease inhibitor cocktail tablets, (Roche Molecular Biochemicals, Indianapolis IN)) and sonicated. The soluble fraction was recovered by centrifugation at 10000 rpm for 10 min at 4C. The recombinant Nim1 kinase was purified by immobilized metal ion affinity chromatography (IMAC, Invitrogen) according to manufacturer instructions. Purity was determined by SDS-polyacrylamide gel electrophoresis (SDS-PAGE) and Coomassie blue staining. Protein concentration was determined using the Bradford method.

##### 15 GST-purification

Sf21 cells were centrifuged at 800 rpm, resuspended in 10 ml of lysis buffer (PBS + 1 mM orthovanadate + 20 mM DTT + Complete protease inhibitor cocktail tablets (Roche Molecular Biochemicals)), and sonicated. The soluble fraction was recovered by centrifugation at 10000 rpm for 10 min at 4C. The recombinant Nim1 kinase was purified by affinity chromatography on glutathione-SEPHAROSE resin (Amersham Pharmacia Biotech) following the manufacturer instructions. Purity was determined using SDS-PAGE followed by Coomassie blue staining. Protein concentration was determined by the Bradford method.

For Western blot analysis, proteins were separated by SDS-PAGE and analyzed by immunoblotting (with either GST-HRP conjugated antibody or His-HRP conjugated antibody, Santa Cruz Biotechnology) using standard materials and techniques (ECL, Amersham Pharmacia Biotech).

#### XII Characterization of the Protein

30 In vitro translation of Incyte Clone number 3317608 was accomplished using the TNT T7 quick coupled transcriptional translational system (Promega) and the manufacturer's instructions. SDS-PAGE analysis revealed the presence of a 48 kd protein. The size of the protein was confirmed by expressing pcDNA3-Nim1 as a template.

As previously described, Nim1 kinase was expressed in both mammalian and insect cells. The purified protein was subjected to SDS-PAGE and revealed a 50 kd protein (including myc-His tag) and

WO 01/68825

PCT/US00/07715

an 80 kd protein (including GST tag) respectively.

#### XIII Human NIM1 Kinase Assay

An *in vitro* kinase assay was performed by incubating 100 ng of recombinant Nim1-GST with 2 µg of myelin basic protein (MBP) or histone (HIST) in 20 µl of kinase buffer (50 mM HEPES pH 7.5, 3 mM MgCl<sub>2</sub> and MnCl<sub>2</sub>, 10 mM DTT and 6 µM NaOAc) containing 10 µCi of [ $\gamma$ -<sup>32</sup>P]-ATP (3000 Ci/mmol; Amersham Pharmacia Biotech) for 30 min at 37°C. The reaction was stopped by adding sample buffer and heating to 100°C for 5 min. Samples were analyzed by SDS-PAGE, the gels were dried and subjected to autoradiographic analysis.

#### XIV Production of NIM1 kinase Specific Antibodies

NIM1 kinase is purified using polyacrylamide gel electrophoresis and used to immunize mice or rabbits. Antibodies are produced using the protocols below. Alternatively, the amino acid sequence of NIM1 kinase is analyzed using LASERGENE software (DNASTAR) to determine regions of high immunogenicity. An immunogenic epitope, usually found near the C-terminus or in a hydrophilic region is selected, synthesized, and used to raise antibodies. Typically, epitopes of about 15 residues in length are produced using an ABI 431A peptide synthesizer (PE Biosystems) using Fmoc-chemistry and coupled to KLH (Sigma-Aldrich) by reaction with N-maleimidobenzoyl-N-hydroxysuccinimide ester to increase immunogenicity.

Rabbits are immunized with the epitope-KLH complex in complete Freund's adjuvant. Immunizations are repeated at intervals thereafter in incomplete Freund's adjuvant. After a minimum of seven weeks for mouse or twelve weeks for rabbit, antisera are drawn and tested for antipeptide activity. Testing involves binding the peptide to plastic, blocking with 1% bovine serum albumin, reacting with rabbit antisera, washing, and reacting with radio-iodinated goat anti-rabbit IgG. Methods well known in the art are used to determine antibody titer and the amount of complex formation.

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

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

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

The nucleic acid molecule, or fragments thereof, or the protein, or portions thereof, are labeled

WO 01/68825

PCT/US00/07715

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

#### XVII Two-Hybrid Screen

10 A yeast two-hybrid system, MATCHMAKER LexA Two-Hybrid system (Clontech Laboratories, Palo Alto CA), is used to screen for peptides that bind the mammalian protein of the invention. A nucleic acid molecule encoding the protein is inserted into the multiple cloning site of a pLexA vector, ligated, and transformed into *E. coli*. cDNA, prepared from mRNA, is inserted into the multiple cloning site of a pB42AD vector, ligated, and transformed into *E. coli* to construct a cDNA library. The pLexA plasmid and pB42AD-cDNA library constructs are isolated from *E. coli* and used  
15 in a 2:1 ratio to co-transform competent yeast EGY48[p8op-lacZ] cells using a polyethylene glycol/lithium acetate protocol. Transformed yeast cells are plated on synthetic dropout (SD) media lacking histidine (-His), tryptophan (-Trp), and uracil (-Ura), and incubated at 30C until the colonies have grown up and can be counted.  
20 The colonies are pooled in a minimal volume of 1x TE (pH 7.5), replated on SD/-His/-Leu/-Trp/-Ura media supplemented with 2% galactose (Gal), 1% raffinose (Raf), and 80 mg/ml 5-bromo-4-chloro-3-indolyl β-d-galactopyranoside (X-Gal), and subsequently examined for growth of blue colonies. Interaction between expressed protein and cDNA fusion proteins activates expression of a LEU2 reporter gene in EGY48 and produces colony growth on media lacking leucine (-Leu). Interaction also  
25 activates expression of β-galactosidase from the p8op-lacZ reporter construct that produces blue color in colonies grown on X-Gal.

Positive interactions between expressed protein and cDNA fusion proteins are verified by isolating individual positive colonies and growing them in SD/-Trp/-Ura liquid medium for 1 to 2 days at 30C. A sample of the culture is plated on SD/-Trp/-Ura media and incubated at 30C until colonies  
30 appear. The sample is replica-plated on SD/-Trp/-Ura and SD/-His/-Trp/-Ura plates. Colonies that grow on SD containing histidine but not on media lacking histidine have lost the pLexA plasmid. Histidine-requiring colonies are grown on SD/Gal/Raf/X-Gal/-Trp/-Ura, and white colonies are isolated and propagated. The pB42AD-cDNA plasmid, which contains a nucleic acid molecule encoding a protein that physically interacts with the mammalian protein, can be isolated from the yeast  
35 cells and characterized.

WO 01/68825

PCT/US00/07715

All patents and publications mentioned in the specification are herein incorporated by reference. Various modifications and variations of the described method and system 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 specific preferred embodiments, it  
5 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 that are obvious to those skilled in the field of molecular biology or related fields are intended to be within the scope of the following claims.

WO 01/68825

PCT/US00/07715

What is claimed is:

1. A substantially purified nucleic acid molecule or a fragment thereof encoding the protein comprising SEQ ID NO:2 or a portion thereof.
2. The composition comprising the polynucleotide of claim 1 or a fragment or a complement thereof.
3. A mammalian variant of the nucleic acid molecule of claim 1 selected from SEQ ID NO:24-30.
4. A fragment of at least 18 consecutive nucleotides selected from SEQ ID NOs:1 and 3-30 or the complements thereof.
5. A substrate containing the fragment of claim 4.
6. A probe comprising the fragment of claim 4.
7. An expression vector comprising the nucleic acid molecule of claim 1.
8. A host cell containing the expression vector of claim 7.
9. A method for producing a protein, the method comprising the steps of:
  - a) culturing the host cell of claim 8 under conditions for the expression of the protein; and
  - b) recovering the protein from the host cell culture.
10. A method for detecting a nucleic acid molecule in a sample, the method comprising the steps of:
  - a) hybridizing the fragment of claim 4 to at least one nucleic acid molecule in the sample, thereby forming a hybridization complex; and
  - b) detecting the hybridization complex, wherein the presence of the hybridization complex indicates the presence of the nucleic acid molecule in the sample.
11. The method of claim 10 wherein the fragment is used to diagnose a brain disorder or cancer.
12. The method of claim 10 further comprising amplifying the nucleic acid molecules of the sample prior to hybridization.
13. A method of using a nucleic acid molecule to screen a library of molecules or compounds, the method comprising:
  - a) combining the nucleic acid molecule of claim 1 with a library of molecules or compounds under conditions to allow specific binding; and
  - b) detecting specific binding, thereby identifying a molecule or compound which specifically binds the nucleic acid molecule.
14. The method of claim 13 wherein the library is selected from DNA molecules, RNA molecules, peptide nucleic acids, artificial chromosome constructions, peptides, and proteins.

WO 01/68825

PCT/US00/07715

15. An isolated and purified protein comprising SEQ ID NO:2 or a portion thereof.
16. A method for using a protein to screen a library of molecules and compounds, the method comprising:
- a) combining the protein of claim 15 with the library of molecules or compounds under conditions to allow specific binding; and
  - b) detecting specific binding, thereby identifying a molecule or compound which specifically binds the protein.
17. The method of claim 16 wherein the library is selected from DNA molecules, RNA molecules, peptide nucleic acids, peptides, proteins, mimetics, agonists, antagonists, antibodies, immunoglobulins, inhibitors, and drugs.
18. A method of preparing an antibody comprising:
- a) immunizing an animal with the protein of claim 15 or an antigenically-effective portion thereof under conditions to elicit an antibody response;
  - b) isolating animal antibodies; and
  - c) screening the isolated antibodies with the protein thereby identifying an antibody which specifically binds the protein.
19. A method of using an antibody which specifically binds NIMI kinase to diagnose a disease, the method comprising:
- a) combining the antibody with a sample under conditions to allow specific binding;
  - b) detecting the bound antibody;
  - c) comparing expression with that of known standards, thereby establishing the presence of the disease.
20. The method of claim 19 wherein the disease is a brain disorder or cancer.

WO 01/68825

PCT/US00/07715

1/12

10 19 28 37 46 55  
 5' AGA GAT GAG ATC CCG CAG CAG GGA CGT GGG GGC CTC CCA GGG GCA TTT ACG CAC  
 64 73 82 91 100 109  
 CAG AGT GCA AGA TTC TCT GGC CAT CAA GGG AAA TAG CAA ACA GAA GCC TTT GTC  
 118 127 136 145 154 163  
 CTG GGG CAC AGC CAC CTA CCA CAA AGC ATC AGA CTC CAC GTC TGG CCA GAA AGT  
 172 181 190 199 208 217  
 TCC TGG AGT CCC ATC AGG CCA GTG GGT ATG TAA CAT GTG CCT AAT TGT ACA GCT  
 226 235 244 253 262 271  
 AGA GCC TGC AAG TTC AAC GTG AGG GAA GGT GGG AAA TGT CTT GAG TGA GGC GAG  
 280 289 298 307 316 325  
 CAG CTC CTG GCT GGG CTG GGC AGA CTC AGC TAC CAC GTT CAC TGC CTT CCT CTC  
 334 343 352 361 370 379  
 ACT AAA GCC GAG AGG GAG GCT GCT CAG CTC TCA GGA AAA CTC TTT TGA ACC CTG  
 388 397 406 415 424 433  
 GGC ACC TGC TGT CCT CAG TTG GCA TCT CCC ACC CTC TGA GCC TCT TCT GCT CCT

FIGURE 1A

WO 01/68825

PCT/US00/07715

2/12

442 GCA CAA CCT GCC TCT TCG CTG AGA TGG AGA CGT GAG CCC CCG TGG ACG ATG ACT 460 469 478 487  
 M T  
 496 GCA GTG TAT ATG AAT GGA GGT GGC CTG GTG AAC CCC CAC TAT GCC CGG TGG GAT 514 523 532 541  
 A V Y M N G G G L V N P H Y A R W D  
 550 CCG CGC GAC AGT GTA GAA AGT GGC TGT CAG ACC GAG AGT AGC AAG GTG GGT GAG 568 577 586 595  
 R R D S V E S G C Q T E S S K V G E  
 604 GAG GGA CAG CCC CGC CAG CTG ACG CCC TTC GAG AAA CTG ACA CAG GAC ATG TCC 613 622 631 640 649  
 E G Q P R Q L T P F E K L T Q D M S  
 658 CAG GAT GAG AAG GTG GTG AGG GAG ATC ACG CTG GGG AAA CGG ATA GGC TTC TAC 676 685 694 703  
 Q D E K V V R E I T L G K R I G F Y  
 712 CGA ATT CGA GGG GAA ATC GGA AGT GGA AAC TTC TCC CAA GTG AAG CTT GGG APT 721 730 739 748 757  
 R I R G E I G S G N F S Q V K L G I

FIGURE 1B

766 CAC TCC CTA ACC AAA GAA AAG GTG GCC ATT AAG ATC CTG GAC AAG ACC AAG TTA 811  
 H S L T K E K V A I K I L D K T K L  
 820 GAC CAG AAA ACC CAG AGG CTA CTA TCC CGA GAA ATC TTC AGC ATG GAA AAG CTG 865  
 D Q K T Q R L L S R E I S S M E K L  
 874 CAC CAT CCC AAC ATC ATC CGC CTT TAC GAA GTG GTG GAG ACC CTA TCC AAG CTG 919  
 H H P N I I R L Y E V E T L S K L  
 928 CAC TTG GTG ATG CAG TAT GCA GGG GGT GGG GAG CTC TTC GGA AAA ATT AGC ACT 973  
 H L V M E Y A G G G E L F G K I S T  
 982 GAG GGG AAG CTC TCT GAA CCA GAA AGC AAG CTC ATC TTC TCC CAG ATT GTG TCT 1027  
 E G K L S E P E S K L I F S Q I V S  
 1036 GCC GTG AAG CAC ATG CAT GAA AAC CAA ATT ATT CAT AGA GAT CTG AAA GCA GAA 1081  
 A V K H M H E N Q I I H R D L K A E

FIGURE 1C

1090 1099 1108 1117 1126 1135  
 AAT GTA TTC TAT ACC AGT AAT ACT TGT GTG AAG GTG GGC GAT TTT GGA TTC AGC  
 N V F Y T S N T C V K V G D F G F S  
  
 1144 1153 1162 1171 1180 1189  
 ACA GTA AGC AAA AAA GGT GAA ATG CTG AAC ACT TTC TGT GGG TCT CCT CCC TAC  
 T V S K K G E M L N T F C G S P P Y  
  
 1198 1207 1216 1225 1234 1243  
 GCT GCG CCT GAA CTC TTC CCG GAC GAG CAC TAC ATC GGC ATT TAC GTG GAT ATC  
 A A P E L F R D E H Y I G I Y V D I  
  
 1252 1261 1270 1279 1288 1297  
 TGG GCC TTG GGG GTG CTT TTG TAC TTC ATG GTG ACT GGC ACC ATG CCA TTT CGG  
 W A L G V L L Y F M V T G T M P F R  
  
 1306 1315 1324 1333 1342 1351  
 GCA GAA ACC GTG GCC AAA CTA AAA AAG AGC ATC CTC GAG GGC ACA TAC AGT GTA  
 A E T V A K L K K S I L E G T Y S V  
  
 1360 1369 1378 1387 1396 1405  
 CCG CCG CAC GTG TCA GAG CCC TGC CAC CGA CTC ATC CGA GGA GTC CTT CAG CAG  
 P P H V S E P C H R L I R G V L Q Q

FIGURE 1D

WO 01/68825

PCT/US00/07715

5/12

1414 1423 1432 1441 1450 1459  
 ATC CCC ACG GAG AGG TAC GGA ATC GAC TGC ATC ATG AAT GAT GAA TGG ATG CAA  
 I P T E R Y G I D C I M N D E W M Q

1468 1477 1486 1495 1504 1513  
 GGG GTG CCA TAC CCT ACA CCT TTG GAA CCT TTC CAA CTG GAT CCC AAA CAT TTG  
 G V P Y P T P L E P F Q L D P K H L

1522 1531 1540 1549 1558 1567  
 TCG GAA ACC AGC ACT CTC AAG GAA GAA AAT GAG GTC AAA AGC ACT TTA GAA  
 S E T S T L K E E N E V K S T L E

1576 1585 1594 1603 1612 1621  
 CAT TTG GGC ATT ACA GAA GAG CAT ATT CGA AAT AAC CAA GGG AGA GAT GCT CGC  
 H L G I T E E H I R N N Q G R D A R

1630 1639 1648 1657 1666 1675  
 AGC TCA ATC ACA GGG GTC TAT AGA ATT ATT TTA CAT AGA GTC CAA AGG AAG AAG  
 S S I T G V Y R I I L H R V Q R K K

1684 1693 1702 1711 1720 1729  
 GCT TTG GAA AGT GTC CCA GTC ATG ATG CTA CCA GAC CCT AAA GAA AGA GAC CTC  
 A L E S V P V M M L P D P K E R D L

FIGURE 1E

WO 01/68825

PCT/US00/07715

6/12

1738 1747 1756 1765 1774 1783  
 AAA AAA GGG TCC CGT GTC TAC AGA GGG ATA AGA CAC ACA TCC AAA TTT TGC TCG  
 K K G S R V Y R G I R H T S K F C S  
 1792 1801 1810 1819 1828 1837  
 ATT TTA TAA ATT GCA CTA GAC TGC TTG TAA CTA ACC AAG ATG ATT GTT GCT GCT  
 I L  
 1846 1855 1864 1873 1882 1891  
 TCT AAA TTT TTT TCA AGG ACA ACT TGA GTG GAG ACA TTT TTG TAA TTT TTA AAT  
 1900 1909 1918 1927 1936 1945  
 AAA CTT AAA TTT GAG ATA TGC AAA AAA AAA AAA AAA AAG GGC CGC CGA CTA  
 1954 1963 1972 1981 1990 1999  
 GTG AGC TCG TCG ACC CGG GAA TTA ATT CCG GAC CGG TAC CTG CAG GCG TAC CAG  
 2008 2017 2026 2035 2044 2053  
 CTT TCC CTA TAG TGG AGT CCG TAT TAA ACT TGG CCG TAA TCA TGG CAT AAC TTG  
 TTC CCT G 3'

FIGURE 1F

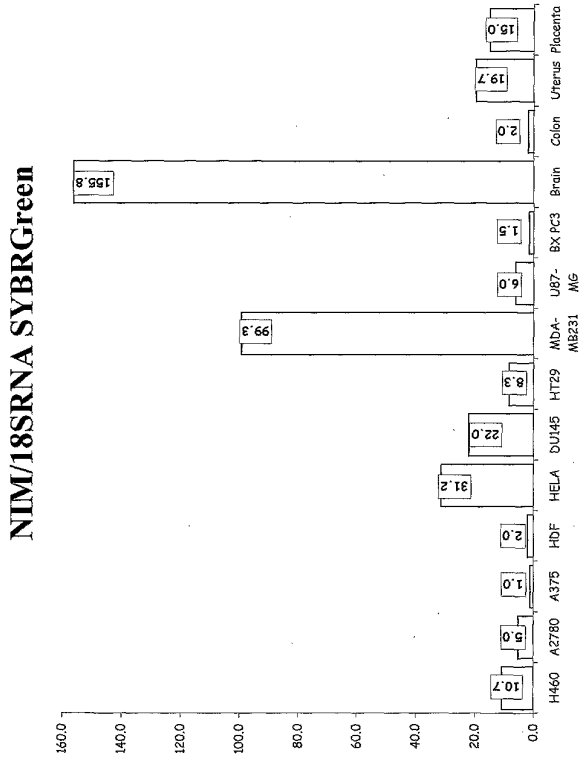
WO 01/68825

PCT/US00/07715

7/12

Human Tissue Category	Clone	Count	Absolute	
			Found	Abundance
Cardiovascular System		247285	0/62	0
Connective Tissue		120592	1/38	1
Digestive System		375714	0/111	0
Embryonic Structures		84970	0/16	0
Endocrine System		179593	0/43	0
Exocrine Glands		240308	0/59	0
Genitalia, Female		344493	0/82	0
Genitalia, Male		411465	1/103	1
Germ Cells		28289	0/4	0
Hemic and Immune System		643113	0/144	0
Liver		78332	0/20	0
Musculoskeletal System		135466	0/38	0
Nervous System		709414	10/144	10
Pancreas		93477	0/21	0
Respiratory System		335356	0/73	0
Sense Organs		19263	0/8	0
Skin		60390	0/13	0
Stomatognathic System		10993	0/4	0
Unclassified/Mixed		89773	0/3	0
Urinary Tract		233771	0/53	0
<b>Totals</b>		<b>4442057</b>	<b>12/1039</b>	<b>12</b>

FIGURE 2



**FIGURE 3**

```

1 M T A V Y M N G G G L V N P H Y A R W D R R D S V E S G >3317608CD1
1 M S E - - - - - K T Q Y E A I L Q L N N D P V - - - V H R E V W A C - - >g3877329
1 M S A - R T P L P T V N E R D T E N H T S V D G Y T E T H I P P T K S S R - - >g2052189
1 M S T - R T P L P T V N E R D T E N H T S - - - H G D G R Q E V T S R T S R - - >g3089349
1 M V I M - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - >g5672676
1 M T A N T N K T T D K E N D P G P N T S I S T T A T P P S A A A Q N V G C V >g2564680
1 M - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - >g1749794

38 E E G O P R O L T P F E K L T O D M S O D E K V V R E I T L G K R I G F Y R I R >3317608CD1
28 - - - - - I V S Y G A - - - - - R L M F Q V A L G R I G F Y R L G >g3877329
38 - Q N I P R C R N S I T S A T D E Q P H - - - - - I G N Y R L Q >g2052189
35 - S G - A R C R N S I A S C A D E Q P H - - - - - I G N Y R L L >g3089349
5 - - - - - S E F S A V P T G T G Q G Q - - - - - Q K P L - - - R V G F Y D V E >g5672676
41 G S S G G R S S P K F X S Y V N G N G Y - - - - - - - - - - - G V Y K I I >g2564680
8 - - - - - A T S A D E Q P H - - - - - - - - - - - I G N Y R L L >g1749794

78 G E I G S G N F S Q V K L G I H S L T K E K V A I K I L D K T K L D Q K T Q R L >3317608CD1
53 K E L L G A G N F S K V K L G V H O L T K E K V A V K I I M D K A K M D A K A O K L >g3877329
64 K T I G K G N F A K V K L A R H V L T G R E V A I K I I D K T Q L N P T S L Q K >g2052189
31 R T L G K G N F A V V K L A R H V T K T O V A I K I I D K T R L D S S N L E K >g3089349
67 K T L G K G N F A K V K L A R H L P T G R E V A I K I I D K T A L N T I A R Q K >g5672676
24 K T I G K G N F A K V K L A R H I L T G K E V A V K I I D K T Q L N S S L Q K >g1749794

```

FIGURE 4A

118 L S R E I S S M E K L L H H P N I I R L Y E V V E T L S K L H L V M E Y A G G E >3317608CD1  
 93 L S R E J Q A M E E M N H P N I V K L F E V V E T L T R V H L L V I E Y A S G G E >g3877329  
 104 L F R E V R I M K I L N H P N I V K L F E V I E T E K T L Y L L Y M E Y A S G G E >g2052189  
 100 L F F R E V R I M K I L N H P N I V K L F E V I E T E K T L Y L L I M E Y A S G G E >g3089349  
 71 I Y R E V Q L M K L L N H P N I I K L Y Q V M E T K D M L Y I V T E F A K N G E >g5672676  
 107 L Y R E V N I M K K L N H P N I V R L L Q V I E S E R T L Y L V M E Y V S G G E >g2564680  
 64 L F R E V R I M K V L N H P N I V K L F E V I E T E K T L Y L V M E Y A S G G E >g1749794  
  
 158 L F G K I S T E G K L S E P E S K L I F S Q I V S A V K H M H E N O I I H R D L >3317608CD1  
 133 L Y T Y V H E R G K L T E G D A K P L F A Q I V S A V S H M H S R R I V H R D L I >g3877329  
 144 V F D Y L V A H G R M K E K E A R A K F R Q I V S A V Q Y C H Q K C I V H R D L >g2052189  
 140 V F D Y L V A H G R M K E K E A R S K F R Q I V S A V Q Y C H Q K R I V H R D L >g3089349  
 111 M F D Y L T S N G H L S E N E A R K K F W Q I L S A V E Y C H N H H I V H R D L >g5672676  
 147 L F N Y L V K N G R M R E R D A R V L F R Q L V S A I E Y C H S K S I V H R D L >g2564680  
 104 V E D Y L V A H G R M K E K E A R A K F R Q I V S A V Q Y C H Q K F I V H R D L >g1749794  
  
 198 K A E N V F Y T S N T C V K V G D F G F S T V S K K G E M L N T F C G S P P Y A >3317608CD1  
 173 K A E N V M F S S P N T V K L V D F G F S C L V D R E Q M L R T F C G S P P Y A >g3877329  
 184 K A E N L L D A D M N I K I A D F G F S N E F T V G N K L D T F C G S P P Y A >g2052189  
 180 K A E N L L D A D M N I K I A D F G F S N E F T V G G K L D T F C G S P P Y A >g3089349  
 151 K T E N L L D G N M D I K L A D F G F G N F Y K P G E P L S T M C G S P P Y A >g5672676  
 187 K A E N L L D Q Q M K L K I A D F G F S T F F E P K A P L E T F C G S P P Y A >g2564680  
 144 K A E N L L D A D M N I K I A D F G F S N E F T F G N K L D T F C G S P P Y A >g1749794

FIGURE 4B

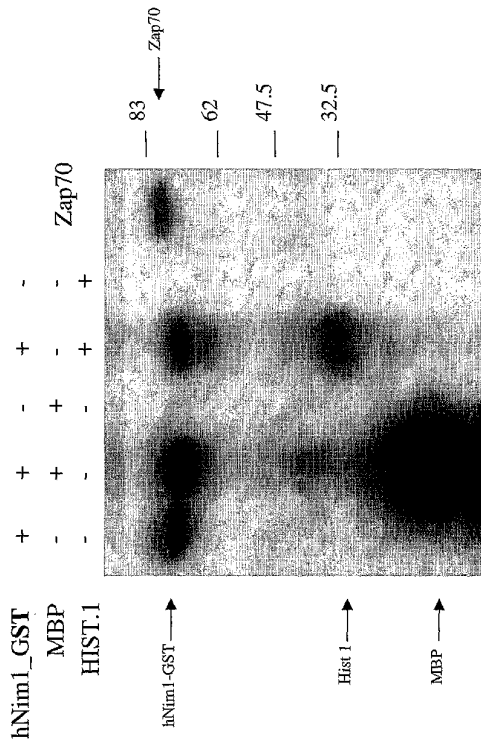


WO 01/68825

PCT/US00/07715

12/12

**KINASE ASSAY USING hNIMI**



**FIGURE 5**

WO 01/68825

PCT/US00/07715

SEQUENCE LISTING

<110> INCYTE PHARMACEUTICALS, INC.  
 BANDMAN, Olga  
 MOLTENI, Angela  
 MAGNAGHI, Paola  
 BOSOTTI, Roberta  
 SACCHERI, Emanuela  
 ISACCHI, Antonella  
 HODGSON, David M.

<120> HUMAN NIM1 KINASE

<130> PC-0009 PCT

<140> To Be Assigned  
 <141> Herewith

<150> Unassigned  
 <151> 2000-03-13

<160> 39

<170> PERL Program

<210> 1  
 <211> 2060  
 <212> DNA  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte Clone No: 3317608CBI

<400> 1  
 cagagatgag atcccgcagc agggacgtgg gggcctccca ggggcattta cgcaccagag 60  
 tgcaagattc tctggccatc aagggaaata gcaaacagaa gcctttgtcc tggggcacag 120  
 ccaactacca caaagcatca gactccactg ctggccagaa agttcctgga gtcccacag 180  
 gccagtggtt atgtaacatg tgcctaactg taccagctaga gcctgcaagt tcaaccgtgag 240  
 ggaagtggtg aatgtctctg agtgagggga gtagctctct gctgggctgg gcagactcag 300  
 ctaccacgtt caactgcttc ctctcactaa agccagagag gaggtgctc agctctcag 360  
 aaaactcttt tgaaccttgg gcaoctgctg tctcagtttg gcatctccca cccctgagc 420  
 ctcttctgct cctgcacaac ctgacctctc gctgagatgg agacgtgag ccccgtagc 480  
 gatgactgca gtgtatatga atggaggtgg cctggtgaac ccccaactat ccccgtagc 540  
 tcggccgcag agttagaaaa gtggctgtca gaccgagagt agcaaggtgg gtgaggagg 600  
 acagccccgc cagctcagc ccttcgagaa actgacacag gacatgtccc aggatgagaa 660  
 ggtggtgagg gagatcacgc tggggaacg gataggcttc taccgaatc gaggagaaat 720  
 cggaaagtga aacttctccc aagtgaagct tgggattcac tccctaacca aagaaaagt 780  
 ggccatlaag atcctggaca agaccaagtt agaccagaaa acccagagcc tactatccc 840  
 agaaatctcc agcatgaaa agctgcacca tccaacatc atccgcttt acgaaagtgt 900  
 ggagacccta tccaagctgc acttggtgat ggagatgca ggggtggg agctcttgg 960  
 aaaaattagc actgagggga agctctctga accagaaagc aagctcatct tctccagat 1020  
 tgtgtctgcc gtgaagcaca tgcataaaaa ccaaatatt catagatc tgaagcaga 1080  
 aaatgtattc tatccagta atacttgtgt gaagtgggc gattttgat tcagcacagt 1140  
 aagcaaaaaa ggtgaaatgc tgaacacttt ctgtggctct ctccctacg ctgcctcga 1200  
 acttctccgg gacgagcact acatcgcat ttactggat atctgggct tgggggtgt 1260  
 tttgtacttc atgtgactg gcaocatgcc atttcgggca gaaaccgtgg ccaactaaa 1320  
 aaagagcacc ctgcaggcca catacaggt acccgcgcac gtgtcagag cctgccacc 1380  
 actcatccga ggagtcttc agcagatccc caccgagag tacggaatcg actgcatcat 1440  
 gaatgatgaa tggatgcaag ggttgccata ccttacacct ttggaacct tccaactgga 1500  
 tcccaaacat ttgtcggaaa ccagcactct caaggaagaa gaaaatgagg tcaaaagcac 1560  
 tttagaacat ttgggcatca cagaagagca tattcgaat aaccaagga gagatgctc 1620  
 cagctcaatc acaggtctct atagaattat ttacataga gtccaaagga agaaggttt 1680  
 ggaagtgtc ccagctatga tgcaccaga cctaagaga agagacctca aaaaaggctc 1740  
 ccgtgtctac agagggataa gacacacatc caaattttgc tcatatttat aaatgcatc 1800  
 agactgcttg taactaacca agatgattgt tgcctctctt aaattttttt caaggacaac 1860

WO 01/68825

PCT/US00/07715

```

ttgagtggag acatttttgt aatttttaaa taaacttaaa ttgagatat gcaaaaaaaaa 1920
aaaaaaaaag ggcggccgccc gactagttag ctctctgacc cgggaattaa ttccggaccg 1980
gtactgtcag ggcgaaccagc tttccctata gtggagtccg tattaacctt ggcgtaacc 2040
atggcataac ttgtccctg 2060

```

```

<210> 2
<211> 436
<212> PRT
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Incyte ID No: 3317608CD1

```

```

<400> 2
Met Thr Ala Val Tyr Met Asn Gly Gly Gly Leu Val Asn Pro His
1 5 10 15
Tyr Ala Arg Trp Asp Arg Arg Asp Ser Val Glu Ser Gly Cys Gln
20 25 30
Thr Glu Ser Ser Lys Val Gly Glu Glu Gly Gln Pro Arg Gln Leu
35 40 45
Thr Pro Phe Glu Lys Leu Thr Gln Asp Met Ser Gln Asp Glu Lys
50 55 60
Val Val Arg Glu Ile Thr Leu Gly Lys Arg Ile Gly Phe Tyr Arg
65 70 75
Ile Arg Gly Glu Ile Gly Ser Gly Asn Phe Ser Gln Val Lys Leu
80 85 90
Gly Ile His Ser Leu Thr Lys Glu Lys Val Ala Ile Lys Ile Leu
95 100 105
Asp Lys Thr Lys Leu Asp Gln Lys Thr Gln Arg Leu Leu Ser Arg
110 115 120
Glu Ile Ser Ser Met Glu Lys Leu His His Pro Asn Ile Ile Arg
125 130 135
Leu Tyr Glu Val Val Glu Thr Leu Ser Lys Leu His Leu Val Met
140 145 150
Glu Tyr Ala Gly Gly Glu Leu Phe Gly Lys Ile Ser Thr Glu
155 160 165
Gly Lys Leu Ser Glu Pro Glu Ser Lys Leu Ile Phe Ser Gln Ile
170 175 180
Val Ser Ala Val Lys His Met His Glu Asn Gln Ile Ile His Arg
185 190 195
Asp Leu Lys Ala Glu Asn Val Phe Tyr Thr Ser Asn Thr Cys Val
200 205 210
Lys Val Gly Asp Phe Gly Phe Ser Thr Val Ser Lys Lys Gly Glu
215 220 225
Met Leu Asn Thr Phe Cys Gly Ser Pro Pro Tyr Ala Ala Pro Glu
230 235 240
Leu Phe Arg Asp Glu His Tyr Ile Gly Ile Tyr Val Asp Ile Trp
245 250 255
Ala Leu Gly Val Leu Leu Tyr Phe Met Val Thr Gly Thr Met Pro
260 265 270
Phe Arg Ala Glu Thr Val Ala Lys Leu Lys Lys Ser Ile Leu Glu
275 280 285
Gly Thr Tyr Ser Val Pro Pro His Val Ser Glu Pro Cys His Arg
290 295 300
Leu Ile Arg Gly Val Leu Gln Gln Ile Pro Thr Glu Arg Tyr Gly
305 310 315
Ile Asp Cys Ile Met Asn Asp Glu Trp Met Gln Gly Val Pro Tyr
320 325 330
Pro Thr Pro Leu Glu Pro Phe Gln Leu Asp Pro Lys His Leu Ser
335 340 345
Glu Thr Ser Thr Leu Lys Glu Glu Glu Asn Glu Val Lys Ser Thr
350 355 360
Leu Glu His Leu Gly Ile Thr Glu Glu His Ile Arg Asn Asn Gln
365 370 375
Gly Arg Asp Ala Arg Ser Ser Ile Thr Gly Val Tyr Arg Ile Ile
380 385 390

```

WO 01/68825

PCT/US00/07715

Leu	His	Arg	Val	Gln	Arg	Lys	Lys	Ala	Leu	Glu	Ser	Val	Pro	Val
				395					400					405
Met	Met	Leu	Pro	Asp	Pro	Lys	Glu	Arg	Asp	Leu	Lys	Lys	Gly	Ser
				410					415					420
Arg	Val	Tyr	Arg	Gly	Ile	Arg	His	Thr	Ser	Lys	Phe	Cys	Ser	Ile
				425					430					435
Leu														

<210> 3  
 <211> 1051  
 <212> DNA  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte Template No: 200700.1

```

<400> 3
gctgcacat cccaacatca tccgccttta cgaagtgtg gagaccctat ccaagctgca 60
cttggtagt gagtatgcag ggggtgggga gctcttcgga aaaattaga ctaggggaa 120
gctctctgaa ccgaaaagca agctcatctt ctccagatt gtgtctgcc tgaagcacat 180
cagtgaaaac caattattc atagagatct gaaagcagaa aatgtattct ataccagtaa 240
tacttgtgt aaggtggcg attttgatt cagcacagta agcaaaaag gtgaaatgct 300
gaaacatttc tgtgggtctc ctccctacgc tgcgctgaa ctcttcggg acggcacta 360
catggcatt tacgtggata tctggccctt gggggtgctt ttgtaactca tggtgactgg 420
caccatgcca ttccgggcag aaaccgtggc caaactaaaa aagagcatcc tgggggac 480
atacagtgta ccgcgcagc tgtcagagcc ctgcaccca ctatccgag gaggctcta 540
gcagatcccc ccggagaggt acggaatcga ctgcacatg aatgatgat ggatgcaagg 600
ggtgccatcc ctacacatt tggaaacttt ccaactggat cccaacatt tgcggaaac 660
cagcatcttc aaggaagaag aaaatgaggt caaagcact ttagaacatt tgggcattc 720
agaagacat attcgaata accaaggagc agatgctgc agctcaatca cagggtcta 780
tagaattatt ttacatagag tccaaggaa gaaggcttgg gaaagtgtcc cagtcgatg 840
gctaccagac ctaaaagaaa gagacctcaa aaaaggttcc cgtgtctaca gagggataag 900
accacatccc aaattttgct cgattttata aattgacta gactgctgt aactaaccaa 960
gatgattgtt gctgcttcta aattttttc aaggacaact tgagtgaga catttttga 1020
atttttaaat aaacttaaat ttgagatag c

```

<210> 4  
 <211> 1507  
 <212> DNA  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte Clone No: 670279CB1

```

<400> 4
gagcctcttc tgcctctgca caacctgect cttcgctgag atggagagct gagccccct 60
ggacgatgac tgcagtgtat atgaatggag gtggcctggg gaacccccca ctatgcccgg 120
tggatcggc gcgacaggtt agaaagtggc tgcagacccg agagtageaa gttgggtgag 180
gagggacagc ccgcccagct gacgcccctt gagaactga cacaggacat gtcccaggat 240
gagaaggtgg tggaggagat cagcctgggg aaacggatag gcttctaccg aattcgaggg 300
gaaatcggaa gtggaaactt ctcccaagtg aagcttggga ttactctccc aaccaagaa 360
aaggtggcca ttaagatcct ggacaagacc aagttagacc agaaaacca gaggctacta 420
tcccagaaaa tctccagcat ggaaaagctg caccatccca acatcatccg cctttacgaa 480
gtgtgggaga cccatccaa gctgcacttg gtgatggagt atgcagggg tggggagctc 540
ttcggaaaaa ttgacactga ggggaagctc ctgaaaccag aaagcaagct catctctccc 600
cagattgtgt ctgcccgaaa gacatgcat gaaaaccaaa ttaticatag agatctgaaa 660
ccagaaaaatg tattctatc cagtaatact tgttgaagg tgggcgattt tggattcagc 720
acagtaagca aaaaagtgta aatgctgaac actttctgtg ggtctctccc ctacgtggcg 780
cctgaactct tccgggacga gcactacatc ggcatttacc tggatatctg ggccctgggg 840
gtgcttttgt acttcatggt gactggcacc atgccaattc gggcagaaac cgtggccaaa 900
ctaaaaaaga gcatctctga gggcacatcc agtataccgc cgcacgtgtc agagccctgc 960
caccgactca tccgaggagt ccttcagcag atccccaccg agaggtacgg aatcgactgc 1020
atcatgaatg atgaatggat gcaaggggtg ccataacccta caccttggaa accttccaa 1080
ctggtccca acatttttc ggaaccgnc actctcaagg aagagaaaa tggagtcaaa 1140
agcacttag aacatttgg cattacagaa gagcatatto gaataacca agggagagat 1200

```

WO 01/68825

PCT/US00/07715

```

gctcgcagct caatcacagg ggtctataga attattttac atagagtcca aaggaagaag 1260
gcttctggaaa gtgtcccagc catgatgcta ccagacccta aagaagaga cctcaaaaaa 1320
gggtcccgtg tctacagggg gataagaac acatccaat ttgtctgat ttataaatt 1380
gacatagact gcttctaact aaccaagatg attgttgctg cttctaaatt tttttcaag 1440
acaacttgag tggagacatt tttgtaattt ttaaataaac ttaatttga gatatgcaa 1500
aaaaaaa 1507

```

```

<210> 5
<211> 258
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Incyte Clone No: 3317608H1

```

```

<400> 5
cagagatgag atccccgagc agggcagctgg gggcctccca ggggcattta cgcaccagag 60
tgcaagattc tctggccatc aagggaaata gcaaacagaa gctttgttcc tggggcacag 120
ccacctacca caaagatca gactccacgt ctggccagaa agttcctgga tgcccatcag 180
gcccagtggg atgtaacatg tgccataattg tacagctaga gcctgcaagt tcaacgtgag 240
ggaagtgagg aaatgtct 258

```

```

<210> 6
<211> 290
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Incyte Clone No: 4313713H1

```

```

<400> 6
ggagtatgca ggggggtggg agctctctcg aaaaattagt actgagggga agctctctga 60
accagaagac aagctcatct tctcccagat tgtgtctgcc gtgaagcaca tgcattgaaa 120
ccaaattatt catagagatc tgaaagcaga aaatgtttct ataccagtaa tacttgtgtg 180
aaggtggcgc attttggatt cagcacagta agcaaaaaag gtgaaatgct gaacatttct 240
gtgggtctcc tcctacgct gcyctgaact cttccgggga cgagcattac 290

```

```

<210> 7
<211> 256
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Incyte Clone No: 4617082H1

```

```

<220>
<221> unsure
<222> 194, 241
<223> a, t, c, g, or other

```

```

<400> 7
agcaaaaaag gtgaaatgct gaacactttc tgtgggtctc ctccctacgc tgcgcctgaa 60
ctcttccggg acgagactca catcggcatt tacgtggata tctgggcctt gggggtgctt 120
ttgtacttca tggtaactgg caccatgcca ttctggggag aaaccgtgac caactaaaa 180
agagcatcc tcgngggcac atacagtgtc ccgcgcacg tctcagagcc ctgccaccga 240
ntcatccgag gactct 256

```

```

<210> 8
<211> 250
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature

```

WO 01/68825

PCT/US00/07715

```

<223> Incyte Clone No: 4711644H1
<220>
<221> unsure
<222> 228
<223> a, t, c, g, or other
<400> 8
ctcactaaag cagagagga ggctgctcag ctctcaggaa aactcttttg aacctgggc 60
acctgctgct ctcagttggc atctcccacc ctctgagcct ctctgctcc tgcacaacct 120
gctctctgct tgagatggag acgtgagccc ccgtggagca tgactgcagt gtatatgaat 180
ggaggtggcc tgggtgaacc ccactatgcc cgggtgggatc ggcgcganag tgtagaaagt 240
ggctgtcaga 250
<210> 9
<211> 243
<212> DNA
<213> Homo sapiens
<220>
<221> misc_feature
<223> Incyte Clone No: 2286324H1
<400> 9
gaattathtt acatagagtc caaaggaaga aggcctttgga aagtgtcca gtcctgatgc 60
taccagacc ctaaagaaga gacctcaaaa aagggtcccg tgtctacaga gggataagac 120
acaacatcaa attttgctcg attttataaa ttgcaactaga ctgcttgtaa ctaaccaaga 180
tgattgttgc tgcttctaaa tttttttcaa ggacaacttg agtgagagaca tttttgtaat 240
ttt 243
<210> 10
<211> 232
<212> DNA
<213> Homo sapiens
<220>
<221> misc_feature
<223> Incyte Clone No: 2286816H1
<400> 10
gaattathtt acatagagtc caaaggaaga aggcctttgga aagtgtcca gtcctgatgc 60
taccagacc ctaaagaaga gacctcaaaa aagggtcccg tgtctacaga gggataagac 120
acaacatcaa attttgctcg attttataaa ttgcaactaga ctgcttgtaa ctaaccaaga 180
tgattgttgc tgcttctaaa tttttttcaa ggacaacttg agtgagagaca tt 232
<210> 11
<211> 88
<212> DNA
<213> Homo sapiens
<220>
<221> misc_feature
<223> Incyte Clone No: 2287217H1
<220>
<221> unsure
<222> 4
<223> a, t, c, g, or other
<400> 11
caanagaag aaacctcaaa aaagggtccc gtgtctacag agggataaga cacacatcca 60
aattttgctc gattttataa attgcact 88
<210> 12
<211> 270
<212> DNA

```

WO 01/68825

PCT/US00/07715

```

<213> Homo sapiens
<220>
<221> misc_feature
<223> Incyte Clone No: 2286816R6

<220>
<221> unsure
<222> 76
<223> a, t, c, g, or other

<400> 12
gaattatattt acatagagtc caaaggaaga aggctttgga aagtgtccca gtcgatgatgc 60
taccagaccoc taaagnaaga gacctcaaaa aagggtcccg tgtctacaga gggataagac 120
acacatccaa attttgctcg attttataaa ttgcactaga ctgcttataa ctaaccaaga 180
tgattgttgc tgcttcaaaa tttttttcaa ggacaacttg agtggagaca tttttgtaat 240
tttttaataa aacttaaatt tgagatgatgc 270

<210> 13
<211> 230
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte Clone No: 2286816R6

<400> 13
tgtctccact caagttgtcc ttgaaaaaaa tttagaagca gcaacaatca tcttgggttag 60
ttacaagcag tctagtgcaa tttataaaat cgagcaaaat ttggatgtgt gtcttatccc 120
tctgtagaca cgggaccctt ttttgaggtc tctttcttta gggctctgta gcatcatgac 180
tgggacactt tccaaagcct tcttcctttg gactctatgt aaaataatc 230

<210> 14
<211> 468
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte Clone No: 3317608R6

<220>
<221> unsure
<222> 393
<223> a, t, c, g, or other

<400> 14
tctccactca agttgtcctt gaaaaaaatt tagaagcagc aacaatcatc ttggttagtt 60
acaagcagtc tagtgaatt tataaaatcg agcaaaatct ggaatgtgtg ctatccctc 120
tgtagacacg ggaacctttt ttgaggtctc ttcttttagg gtctggttag atcatgactg 180
ggacactttc caaagccttc ttccctttgga ctctatgtaa aataattcta tagaccctg 240
tgattgagct gcgagcatct ctcccttggg tatttcgaa atgctcttct gtaatgccca 300
aatgttctaa agtgcttttg acctcatttt ctctctcctt gagagtgctg gtttcgcaca 360
aatgtttggg atccagtttg aaagggctca aangtgaag gtatggcacc ccttgcacc 420
atctcatatt catgatgcag tcatctcctt aactctcctg gggggatc 468

<210> 15
<211> 485
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte Clone No: 4201896R6

<400> 15

```

WO 01/68825

PCT/US00/07715

```

ttcactaatt acaaaaatgt ctccactcaa gttgtccttg acaaaaattt agaagcagca 60
acaatcatct tggtagtla caagcagctc agtgcattt ataaaaatga gcaaaatttg 120
gatgtgtgtc ttatccctct gtagacaagg gaaccttttt tgggtctct ttctttaggg 180
tctgttagca tcatgactgg gacactttcc aaagcctctc tcccttggac tctatgtaaa 240
ataattctat agaccctctg gattgagctg cagacatctc tcccttggtt atttogaata 300
tgctctctcg taatgcccaa atgttctaaa gtgcttttga cctcatttc ttcttctctg 360
agagtgtcgg ttccgacca atgtttggga tccagttgga aaggttccaa aggtgtaggg 420
tatggcaacc cttagcctca tcatcattc ctgatgccgt cgattccgta cctctccgtg 480
gggat

```

```

<210> 16
<211> 423
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Incyte Clone No: 4624811T6

```

```

<400> 16
aaatgtctcc actcaagttg tccctgaaaa aaatttagaa gcagcaacaa tcatcttgg 60
tagttacaag cagtctagtg caatttataa aatcgagcaa aatttggatg tgtgtcttat 120
ccctctgtag acacgggacc cttttttgag gtctctttct ttagggtctg gtagcatcat 180
gactlgggaca ctttccaaag cctctctctc ttggactcta tgtaaaaata ttctatagac 240
ccctgtgatt gagctgcgag catctctccc ttggttatit cgaatgatct cttctgtaat 300
gccccaaatg tctaaagtgc tttgacctc atttctctct tccctggagag tgcctgtctc 360
cgacaatgt tgggatcca gttggaaagg ttccaaagggt gtaggtatg gaccccttg 420
cat

```

```

<210> 17
<211> 564
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Incyte Clone No: 6559834H1

```

```

<400> 17
attgagggcg cgggaatcca aaagggaaac aaacgaaacc agatgaaaaa aagtgtattc 60
aagtcgaagt ttcaaatgga aatctaaaga aaggtggcca ttaagatcct ggaacaagac 120
aagtttagac agaaaaccca gaggtaacta tcccgagaaa tctccagcat ggaaaagctg 180
cacatctcca acatcatccg cctttacgaa gtggtggaga cctatccaa gctgcaattg 240
gtgatggagt atgcaggggg tggggagctc ttcggaaaaa ttgactctga ggggaagctc 300
tctgaaccag aaagcaagct catctctccc cagatttgtt ctgccgtgaa gcacatgat 360
gaaaaccaaa ttattcatag agatctgaaa gcagaaaatg tattctatac cagtaatact 420
tgtgtgaagg tggcgattt tggattcagc acagtaagca aaaaagtgaa aatgctgaac 480
actttctgtg ggtctctctc ctacgctgcg cctgaactct tccgggacga gcaactacac 540
ggcatttacg tggatatctg ggcc

```

```

<210> 18
<211> 598
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Incyte Clone No: 670279F1

```

```

<220>
<221> unsure
<222> 227
<223> a, t, c, g, or other

```

```

<400> 18
gcatactca aatttaagtt tatttaaaaa ttacaaaat gtctccaac aagttgtcct 60
tgaaaaaat ttagaagcag caacaatcat cttggttagt tacaagcag ctagtgaact 120

```

WO 01/68825

PCT/US00/07715

```

ttatanaatc gagcaaaatt tggatgtgtg tcttatccct ctgtagacac gggacccttt 180
tttgaggctc ctttctttag ggtctggtag catcatgaact gggacanttt ccaaagcctt 240
cttcccttgg actctatgta aaataatctc atagaccctc gtgaltgagc tgcgagacc 300
tctcccttgg ttatttcgaa tatgctcttc tghtaatgcc aatgtttcta agtgctttt 360
gacctcaatt tcttctctc tgaagatgct ggtttccgac aatgttttg gatccagtgt 420
gaaaggttcc aaaggtgtag ggtatggcac cccttgcatc cattcatcat tcatgatgca 480
gtcgattccg tacctctccg tggggatctg ctgaaggact cctcggatga gtcggtggca 540
gggctctgac acgtgcgccg gtacactgta tgtgccctcg aggatgctct tttttagt 598

<210> 19
<211> 288
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte Clone No: 670279H1

<220>
<221> unsure
<222> 19, 153, 268
<223> a, t, c, g, or other

<400> 19
gctgcaccat cccaacatna tccgccttta cgaagtgtgt gagaccctat ccaagctgca 60
cttggatgag gagtatgcag ggggtgggga gctcttcgga aaattagca ctgaggggaa 120
gctctctgaa ccagaaagca agctcatctt ctccagatt gtgtctcgg tgaagcacat 180
gcattgaaac caaattattc atagagatct gaaagcagaa aatgtattct ataccagtaa 240
tacttgtgtg aagtggtggc attttgntt cagcacagta agcaaaaa 288

<210> 20
<211> 546
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte Clone No: 670279R1

<220>
<221> unsure
<222> 534
<223> a, t, c, g, or other

<400> 20
gctgcaccat cccaacatca tccgccttta cgaagtgtgt agaccctatc caagctgca 60
ttggtgatgg agtatgcag ggggtgggag ctcttcggaa aaattagca tgaggggaag 120
ctctctgaac cagaagcaaa gctcatcttc tcccagattt tgtctgctg gaagcaaatg 180
catgaaacc aaattatbca tagagatctg aaagcagaaa atgtatbcta taccagtaat 240
acttgtgtga aggtgggcca ttttggatcc agcacagtaa gcaaaaagg tgaatgctg 300
aacactttct gtgggtctcc tccctacgct gcgctgaa tcttccggga cgagcactac 360
atcggatctt acgtggatat ctgggctctg ggggtgcttt tgtacttcat ggtgactggc 420
accatgcatc ttcgggcaga aaccgtggcc aaactaaaaa agagcatcct cgagggcaca 480
tacagtgatc cgccgcacgt gtcagagccc tgccaccgac tcatccgagg agtntctcag 540
cagatc

<210> 21
<211> 383
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte Clone No: 670279R6

<220>
<221> unsure

```

WO 01/68825

PCT/US00/07715

```

<222> 81, 147
<223> a, t, c, g, or other

<400> 21
gctgcaccat cccaacatca tccgccttta cgaatgggtg agaccctatc caagctgcac 60
ttggatgagg agtatgcagg nggtggggag ctcttcggaa aaattagcac tgagggaag 120
ctctctgaac cagaaagcaa gctcatnttc tcccagattg tgtctgccgt gaagcacatg 180
catgaaaacc aaattattca tagagatctg aaagcagaaa atgtattcta taccagtaat 240
acttgggtga aggtggggoga ttttggattc agcacagtaa gcaaaaaagg tgaaatgctg 300
aacactttct gtgggtctcc tccctacgct ggcctgaaac tcttccggga ctagcactac 360
atcgggcatt tacgtgggat atc 383

<210> 22
<211> 537
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte Clone No: 670279T6

<220>
<221> unsure
<222> 504
<223> a, t, c, g, or other

<400> 22
aatgtctcca ctcaagttgt cctgaaaaa aatttagaag cagcaacaat catcttggtt 60
agttacaagc agtctagtgc aattataaaa atcgagcaaa atttggatgt gtgtcttacc 120
cctctgtaga caggggacc ttttttgagg tctctttctt taggtctggt tagcatcatg 180
actgggacac ttccaagaac ctctctcctt tggactctat gtaaaataat tctatagacc 240
cctgtgattg agctgcgagc atctctccct tggttatttc gaatatgctc ttctgtaatg 300
cccaaatgct ctaaagtgtc tttagacctca tttctctctt ccttgagagt gctgggttcc 360
gacaatggt tgggaltcag ttgaaaggt tccaaaggt taggtatgg caaccttctg 420
atccatcat cttcatgat gcaatcgatt cagtaacctc ccgtgggat ctgctgaagg 480
actctcoga tgagtcggtg gcanggcctg acacgtggcg cgttacctg tatgtgc 537

<210> 23
<211> 279
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte Clone No: 4936446H1

<220>
<221> unsure
<222> 84
<223> a, t, c, g, or other

<400> 23
gcatcatgaa tgatgaatg atgcaagggg tgccatatacc tacacctttg gaacctttcc 60
aactggatcc caaacatttg tcgnaaacca gcaactctcaa ggaagaagaa aatgaggtca 120
aaagcacttt agaacatttg ggcattacag aagagcatat tcgaaataac caagggagag 180
atgctcgcag ctcaatcaca ggggtctata gaattatttt acatagagt caaaggaaga 240
aggctttgga aagtgctcca gtcgatgatc taccagacc 279

<210> 24
<211> 488
<212> DNA
<213> Rattus norvegicus

<220>
<221> misc_feature
<223> Incyte Template No: 216150.1

```

WO 01/68825

PCT/US00/07715

```

<400> 24.
aaaagttgca ccatcccaac attgtccgtc tttatgaagt cgtggagacc ctgtccaagc 60
tcocactagt gatgggtat gcaggagtg gggagctctt tgggaaatt agcaaccgag 120
ggaaactttc tgaaccgaa agcaagctca tcttctccca gatcgtgtct gccgtgaagc 180
aacggcatga gaaccaaat atccacagag atctgaaagc agaaaacgtc ttctatacca 240
gtagcacttg tgtgaagtg ggggattttg gattcagcac cgtaaagtaag aaaggtgaga 300
tctgaacac cttctgtggg tctccgccct acgctgcacc ggaactcttc cgtgacgagc 360
actatgttg cytttatgtg gatatctggg ccttgggtgt ccttttgtac ttcatggtga 420
ctggtacgat gccatttcca gcagaaaccg tggccaacct gaaaaagagc atcctcgtatg 480
gtgcctac

<210> 25
<211> 272
<212> DNA
<213> Rattus norvegicus

<220>
<221> misc_feature
<223> Incyte Clone No: 701925441H1

<400> 25
agccgcgccca gctgacacct ttcgagaac tgactcagga catgtgccaa gatgagaagg 60
tggtagggga gatcacgctg gggaaacyca taggcttcta tcgaattcga ggggagatcg 120
gaagcggaaa cttttccag gtaagctgg gaattcactc cctaaccaaa gaaaaggtgg 180
ccattaaagt tctggacaaa accaagttag accagaaaac ccaaaggctg ttatccagag 240
aaattccag cagccggaaa tgcaccatcc ca

<210> 26
<211> 303
<212> DNA
<213> Rattus norvegicus

<220>
<221> misc_feature
<223> Incyte Clone No: 701910632H1

<220>
<221> unsure
<222> 183, 190, 287, 291
<223> a, t, c, g, or other

<400> 26
aagttgcacc atcccaacat tgtccgtctt tatgaagtcg tggagaccct gtccaagctc 60
cacctagtga tggagtatgc aggagtgagg gagctctttg gaaaattag caccgagggg 120
aaactttctg aaccgaaag caagctcatc ttctcccaga tcgtgtctgc cgtgaagcaa 180
atnccatgan aaccaaaata tccacagaga tctgaaagca gaaaagctct tctataccag 240
tagcacttgt gtgaaggtgg gggattttgg attcagcacc gtaagtnaga naggtgagat 300
gct

<210> 27
<211> 299
<212> DNA
<213> Rattus norvegicus

<220>
<221> misc_feature
<223> Incyte Clone No: 70190514H1

<220>
<221> unsure
<222> 3, 29
<223> a, t, c, g, or other

<400> 27
ganaacaaa ttatccacag agatctgana gcagaaaacy tcttctatac cagtagcact 60
tgtgtgaagg tggggattt tggattcagc accgtaagta agaaagtgga gatgctgaac 120
accttctgtg ggtctccgcc ctacgctgca ccggaactct tccgtgacga gcactatggt 180

```

WO 01/68825

PCT/US00/07715

```

ggcgtttatg tggatatctg ggcctgggtg tccttttga cttcatggtg actggtacga 240
tgcattcga gcagaaaccg tggcctaaact gaaaaagagc atcctcgagg gtgcctaca 299

<210> 28
<211> 258
<212> DNA
<213> Rattus norvegicus

<220>
<221> misc_feature
<223> Incyte Clone No: 701293826H1

<400> 28
aatggatgcg aggggtgccc taccctccc ctctggaacc tttccaactg gatcctaaac 60
atttgcgga aactagcacc ctcaaagaag aagaaaacga ggtgaaaagc actttagagc 120
acttgggat cacagacgaa catatccgga ataaccaag gagagacgct cgaagctcta 180
tcacgggggt ctatagaatc attttacatc gagtgcaaa gaaaaagcc tggaaagtgt 240
gccaatggcg acactacc 258

<210> 29
<211> 113
<212> DNA
<213> Rattus norvegicus

<220>
<221> misc_feature
<223> Incyte Clone No: 700949543H1

<220>
<221> unsure
<222> 60, 102, 106
<223> a, t, c, g, or other

<400> 29
tacctggcag gcgtaccagc tttccctata agtggagtcg tattaggagc ttgggcgtan 60
atccatgggt coataggctg tttccctgtg tggaaattgt tnatnecgt cca 113

<210> 30
<211> 300
<212> DNA
<213> Macaca fascicularis

<220>
<221> misc_feature
<223> Incyte Clone No: 700706950H1

<220>
<221> unsure
<222> 191, 227, 247-248, 252, 255-256, 269, 285, 287, 290, 293, 296
<223> a, t, c, g, or other

<400> 30
gaagtcgtgg agaccctatc taagctgcac ttggtgatgg aatatgcagg ggggtggggag 60
ctgttcgga aaattagcac tgaggggaag ctctctgaac cagaaagcaa gctcatcttc 120
teccagattg tgtctgcccgt gaagcactgc atgaaccaa ttattcacga gtctgaagca 180
gaaatgtatc natacaglat actgtgtgaa gtggcgattt ggatcancag taattaaaa 240
gtgaatmaa cnttmggtg ctctccatnt gacgaacctc ggcgmctan cgnatnctg 300

<210> 31
<211> 334
<212> FRT
<213> Caenorhabditis elegans

<220>
<221> misc_feature
<223> GenBank Accession No: g3877329

```

WO 01/68825

PCT/US00/07715

```

<400> 31
Met Ser Glu Lys Thr Gln Tyr Glu Arg Ala Ile Leu Gln Leu Asn
1 5 10 15
Asn Asp Pro Val Val His Lys Glu Val Trp Ala Cys Ile Val Ser
20 25 30
Tyr Gly Lys Arg Lys Leu Trp Phe Gln Val Ala Leu Gly Arg Arg
35 40 45
Ile Gly Phe Tyr Arg Leu Gly Lys Glu Leu Gly Ala Gly Asn Phe
50 55 60
Ser Lys Val Lys Leu Gly Val His Gln Leu Thr Lys Glu Lys Val
65 70 75
Ala Val Lys Ile Met Asp Lys Ala Lys Met Asp Ala Lys Ala Gln
80 85 90
Lys Leu Leu Ser Arg Glu Ile Gln Ala Met Glu Glu Met Asn His
95 100 105
Pro Asn Ile Val Lys Leu Phe Glu Val Val Glu Thr Leu Thr Arg
110 115 120
Val His Leu Val Ile Glu Tyr Ala Ser Gly Gly Glu Leu Tyr Thr
125 130 135
Tyr Val His Glu Arg Gly Lys Leu Thr Glu Gly Asp Ala Lys Pro
140 145 150
Leu Phe Ala Gln Ile Val Ser Ala Val Ser His Met His Ser Arg
155 160 165
Asn Ile Val His Arg Asp Ile Lys Ala Glu Asn Val Met Phe Ser
170 175 180
Ser Pro Asn Thr Val Lys Leu Val Asp Phe Gly Phe Ser Cys Leu
185 190 195
Val Asp Arg Glu Gln Met Leu Arg Thr Phe Cys Gly Ser Pro Pro
200 205 210
Tyr Ala Ala Pro Glu Leu Phe Gln Asp Thr Ser Tyr Ala Gly Glu
215 220 225
Leu Val Asp Val Trp Ala Leu Gly Val Leu Leu Phe Phe Met Leu
230 235 240
Ile Gly Val Thr Pro Phe Lys Ala Glu Thr Val Pro Asp Met Lys
245 250 255
Val Leu Ile Thr Ala Gly Lys Tyr Gln Ile Pro Asp Tyr Val Ser
260 265 270
Leu Leu Ala Thr Glu Leu Ile Lys Ser Met Leu Lys Thr Asp Thr
275 280 285
Gly Gln Arg Ala Asp Ile Asp Ser Val Lys Lys His Phe Trp Met
290 295 300
Arg Asp Cys Arg Phe Thr Lys Ser Tyr Leu Ser Ile Lys Ala Thr
305 310 315
Ala Lys Ile Asp Asn Glu Glu Glu Lys Lys Ala Ile Asp Asp Lys
320 325 330
Val Ile Phe Val

```

<210> 32  
<211> 793  
<212> PRT  
<213> Rattus norvegicus

```

<220>
<221> misc_feature
<223> GenBank Accession No: g2052189

```

```

<400> 32
Met Ser Ala Arg Thr Pro Leu Pro Thr Val Asn Glu Arg Asp Thr
1 5 10 15
Glu Asn His Thr Ser Val Asp Gly Tyr Thr Glu Thr His Ile Pro
20 25 30
Pro Thr Lys Ser Ser Ser Arg Gln Asn Ile Pro Arg Cys Arg Asn
35 40 45
Ser Ile Thr Ser Ala Thr Asp Glu Gln Pro His Ile Gly Asn Tyr
50 55 60
Arg Leu Gln Lys Thr Ile Gly Lys Gly Asn Phe Ala Lys Val Lys
65 70 75

```

WO 01/68825

PCT/US00/07715

Leu Ala Arg His Val Leu Thr Gly Arg Glu Val Ala Val Lys Ile  
80 85 90  
Ile Asp Lys Thr Gln Leu Asn Pro Thr Ser Leu Gln Lys Leu Phe  
95 100 105  
Arg Glu Val Arg Ile Met Lys Ile Leu Asn His Pro Asn Ile Val  
110 115 120  
Lys Leu Phe Glu Val Ile Glu Thr Glu Lys Thr Leu Tyr Leu Val  
125 130 135  
Met Glu Tyr Ala Ser Gly Gly Glu Val Phe Asp Tyr Leu Val Ala  
140 145 150  
His Gly Arg Met Lys Glu Lys Glu Ala Arg Ala Lys Phe Arg Gln  
155 160 165  
Ile Val Ser Ala Val Gln Tyr Cys His Gln Lys Cys Ile Val His  
170 175 180  
Arg Asp Leu Lys Ala Glu Asn Leu Leu Leu Asp Ala Asp Met Asn  
185 190 195  
Ile Lys Ile Ala Asp Phe Gly Phe Ser Asn Glu Phe Thr Val Gly  
200 205 210  
Asn Lys Leu Asp Thr Phe Cys Gly Ser Pro Pro Tyr Ala Ala Pro  
215 220 225  
Glu Leu Phe Gln Gly Lys Lys Tyr Asp Gly Pro Glu Val Asp Val  
230 235 240  
Trp Ser Leu Gly Val Ile Leu Tyr Thr Leu Val Ser Gly Ser Leu  
245 250 255  
Pro Phe Asp Gly Gln Asn Leu Lys Glu Leu Arg Glu Arg Val Leu  
260 265 270  
Arg Gly Lys Tyr Arg Val Pro Phe Tyr Met Ser Thr Asp Cys Glu  
275 280 285  
Asn Leu Leu Lys Lys Leu Leu Val Leu Asn Pro Ile Lys Arg Gly  
290 295 300  
Ser Leu Glu Gln Ile Met Lys Asp Arg Trp Met Asn Val Gly His  
305 310 315  
Glu Glu Glu Glu Leu Lys Pro Tyr Ser Glu Pro Glu Leu Asp Leu  
320 325 330  
Asn Asp Ala Lys Arg Ile Asp Ile Met Val Thr Met Gly Phe Ala  
335 340 345  
Arg Asp Glu Ile Asn Asp Ala Leu Val Ser Gln Lys Tyr Asp Glu  
350 355 360  
Val Met Ala Thr Tyr Ile Leu Leu Gly Arg Lys Pro Pro Glu Phe  
365 370 375  
Glu Gly Gly Glu Ser Leu Ser Ser Gly Asn Leu Cys Gln Arg Ser  
380 385 390  
Arg Pro Ser Ser Asp Leu Asn Asn Ser Thr Leu Gln Ser Pro Ala  
395 400 405  
His Leu Lys Val Gln Arg Ser Ile Ser Ala Asn Gln Lys Gln Arg  
410 415 420  
Arg Phe Ser Asp His Ala Gly Pro Ser Ile Pro Pro Ala Val Ser  
425 430 435  
Tyr Thr Lys Arg Pro Gln Ala Asn Ser Val Glu Ser Glu Gln Lys  
440 445 450  
Glu Glu Trp Asp Lys Asp Thr Ala Arg Arg Leu Gly Ser Thr Thr  
455 460 465  
Val Gly Ser Lys Ser Glu Val Thr Ala Ser Pro Leu Val Gly Pro  
470 475 480  
Asp Arg Lys Lys Ser Ser Ala Gly Pro Ser Asn Asn Val Tyr Ser  
485 490 495  
Gly Gly Ser Met Thr Arg Arg Asn Thr Tyr Val Cys Glu Arg Ser  
500 505 510  
Thr Asp Arg Tyr Ala Ala Leu Gln Asn Gly Arg Asp Ser Ser Leu  
515 520 525  
Thr Glu Met Ser Ala Ser Ser Met Ser Ser Thr Gly Ser Thr Val  
530 535 540  
Ala Ser Ala Gly Pro Ser Ala Arg Pro Arg His Gln Lys Ser Met  
545 550 555  
Ser Thr Ser Gly His Pro Ile Lys Val Thr Leu Pro Thr Ile Lys  
560 565 570  
Asp Gly Ser Glu Ala Tyr Arg Pro Gly Thr Ala Gln Arg Val Pro

WO 01/68825

PCT/US00/07715

```

575          580          585
Ala Ala Ser Pro Ser Ala His Ser Ile Ser Ala Ser Thr Pro Asp
590          595          600
Arg Thr Arg Phe Pro Arg Gly Ser Ser Ser Arg Ser Thr Phe His
605          610          615
Gly Glu Gln Leu Arg Glu Arg Arg Ser Ala Ala Tyr Ser Gly Pro
620          625          630
Pro Ala Ser Pro Ser His Asp Thr Ala Ala Leu Ala His Ala Arg
635          640          645
Arg Gly Thr Ser Thr Gly Ile Ile Ser Lys Ile Thr Ser Lys Phe
650          655          660
Val Arg Arg Asp Pro Ser Glu Gly Glu Ala Ser Gly Arg Thr Asp
665          670          675
Thr Ala Arg Gly Ser Ser Gly Glu Pro Lys Asp Lys Glu Glu Gly
680          685          690
Lys Glu Ala Lys Pro Arg Ser Leu Arg Phe Thr Trp Ser Met Lys
695          700          705
Thr Thr Ser Ser Met Asp Pro Asn Asp Met Val Arg Glu Ile Arg
710          715          720
Lys Val Leu Asp Ala Asn Thr Cys Asp Tyr Glu Gln Arg Glu Arg
725          730          735
Phe Leu Leu Phe Cys Val His Gly Asp Ala Arg Gln Asp Ser Leu
740          745          750
Val Gln Trp Glu Met Glu Val Cys Lys Leu Pro Arg Leu Ser Leu
755          760          765
Asn Gly Val Arg Phe Lys Arg Ile Ser Gly Thr Ser Ile Ala Phe
770          775          780
Lys Asn Ile Ala Ser Lys Ile Ala Asn Glu Leu Lys Leu
785          790          795
<210> 33
<211> 729
<212> PRT
<213> Homo sapiens
<220>
<221> misc_feature
<223> GenBank Accession No: g3089349
<400> 33
Met Ser Thr Arg Thr Pro Leu Pro Thr Val Asn Glu Arg Asp Thr
1          5          10          15
Glu Asn His Thr Ser His Gly Asp Gly Arg Gln Glu Val Thr Ser
20          25          30
Arg Thr Ser Arg Ser Gly Ala Arg Cys Arg Asn Ser Ile Ala Ser
35          40          45
Cys Ala Asp Glu Gln Pro His Ile Gly Asn Tyr Arg Leu Leu Lys
50          55          60
Thr Ile Gly Lys Gly Asn Phe Ala Lys Val Lys Leu Ala Arg His
65          70          75
Ile Leu Thr Gly Arg Glu Val Ala Ile Lys Ile Ile Asp Lys Thr
80          85          90
Gln Leu Asn Pro Thr Ser Leu Gln Lys Leu Phe Arg Glu Val Arg
95          100          105
Ile Met Lys Ile Leu Asn His Pro Asn Ile Val Lys Leu Phe Glu
110          115          120
Val Ile Glu Thr Glu Lys Thr Leu Tyr Leu Ile Met Glu Tyr Ala
125          130          135
Ser Gly Gly Glu Val Phe Asp Tyr Leu Val Ala His Gly Arg Met
140          145          150
Lys Glu Lys Glu Ala Arg Ser Lys Phe Arg Gln Ile Val Ser Ala
155          160          165
Val Gln Tyr Cys His Gln Lys Arg Ile Val His Arg Asp Leu Lys
170          175          180
Ala Glu Asn Leu Leu Leu Asp Ala Asp Met Asn Ile Lys Ile Ala
185          190          195
Asp Phe Gly Phe Ser Asn Glu Phe Thr Val Gly Gly Lys Leu Asp
200          205          210

```

WO 01/68825

PCT/US00/07715

Thr Phe Cys Gly Ser Pro Pro Tyr Ala Ala Pro Glu Leu Phe Gln  
 215 220 225  
 Gly Lys Lys Tyr Asp Gly Pro Glu Val Asp Val Trp Ser Leu Gly  
 230 235 240  
 Val Ile Leu Tyr Thr Leu Val Ser Gly Ser Leu Pro Phe Asp Gly  
 245 250 255  
 Gln Asn Leu Lys Glu Leu Arg Glu Arg Val Leu Arg Gly Lys Tyr  
 260 265 270  
 Arg Ile Pro Phe Tyr Met Ser Thr Asp Cys Glu Asn Leu Leu Lys  
 275 280 285  
 Arg Phe Leu Val Leu Asn Pro Ile Lys Arg Gly Thr Leu Glu Gln  
 290 295 300  
 Ile Met Lys Asp Arg Trp Ile Asn Ala Gly His Glu Glu Asp Glu  
 305 310 315  
 Leu Lys Pro Phe Val Glu Pro Glu Leu Asp Ile Ser Asp Gln Lys  
 320 325 330  
 Arg Ile Asp Ile Met Val Gly Met Gly Tyr Ser Gln Glu Glu Ile  
 335 340 345  
 Gln Glu Ser Leu Ser Lys Met Lys Tyr Asp Glu Ile Thr Ala Thr  
 350 355 360  
 Tyr Leu Leu Leu Gly Arg Lys Ser Ser Glu Leu Asp Ala Ser Asp  
 365 370 375  
 Ser Ser Ser Ser Ser Asn Leu Ser Leu Ala Lys Val Arg Pro Ser  
 380 385 390  
 Ser Asp Leu Asn Asn Ser Thr Gly Gln Ser Pro His His Lys Val  
 395 400 405  
 Gln Arg Ser Val Ser Ser Ser Gln Lys Gln Arg Arg Tyr Ser Asp  
 410 415 420  
 His Ala Gly Pro Ala Ile Pro Ser Val Val Ala Tyr Pro Lys Arg  
 425 430 435  
 Ser Gln Thr Ser Thr Ala Asp Gly Asp Leu Lys Glu Asp Gly Ile  
 440 445 450  
 Ser Ser Arg Lys Ser Ser Gly Ser Ala Val Gly Gly Lys Gly Ile  
 455 460 465  
 Ala Pro Ala Ser Pro Met Leu Gly Asn Ala Ser Asn Pro Asn Lys  
 470 475 480  
 Ala Asp Ile Pro Glu Arg Lys Lys Ser Ser Thr Val Pro Ser Ser  
 485 490 495  
 Asn Thr Ala Ser Gly Gly Met Thr Arg Arg Asn Thr Tyr Val Cys  
 500 505 510  
 Ser Glu Arg Thr Thr Ala Asp Arg His Ser Val Ile Gln Asn Gly  
 515 520 525  
 Lys Glu Asn Ser Thr Ile Pro Asp Gln Arg Thr Pro Val Ala Ser  
 530 535 540  
 Thr His Ser Ile Ser Ser Ala Ala Thr Pro Asp Arg Ile Arg Phe  
 545 550 555  
 Pro Arg Gly Thr Ala Ser Arg Ser Thr Phe His Gly Gln Pro Arg  
 560 565 570  
 Glu Arg Arg Thr Ala Thr Tyr Asn Gly Pro Ala Ser Pro Ser  
 575 580 585  
 Leu Ser His Glu Ala Thr Pro Leu Ser Gln Thr Arg Ser Arg Gly  
 590 595 600  
 Ser Thr Asn Leu Phe Ser Lys Leu Thr Ser Lys Leu Thr Arg Ser  
 605 610 615  
 Arg Asn Val Ser Ala Glu Gln Lys Asp Glu Asn Lys Glu Ala Lys  
 620 625 630  
 Pro Arg Ser Leu Arg Phe Thr Trp Ser Met Lys Thr Thr Ser Ser  
 635 640 645  
 Met Asp Pro Gly Asp Met Met Arg Glu Ile Arg Lys Val Leu Asp  
 650 655 660  
 Ala Asn Asn Cys Asp Tyr Glu Gln Arg Glu Arg Phe Leu Leu Phe  
 665 670 675  
 Cys Val His Gly Asp Gly His Ala Glu Asn Leu Val Gln Trp Glu  
 680 685 690  
 Met Glu Val Cys Lys Leu Pro Arg Leu Ser Leu Asn Gly Val Arg  
 695 700 705  
 Phe Lys Arg Ile Ser Gly Thr Ser Ile Ala Phe Lys Asn Ile Ala

WO 01/68825

PCT/US00/07715

```

Ser Lys Ile Ala Asn Glu Leu Lys Leu
710
715
720
<210> 34
<211> 776
<212> PRT
<213> Rattus norvegicus

<220>
<221> misc_feature
<223> GenBank Accession No: g5672676

<400> 34
Met Val Ile Met Ser Glu Phe Ser Ala Val Pro Thr Gly Thr Gly
1 5 10 15
Gln Gly Gln Gln Lys Pro Leu Arg Val Gly Phe Tyr Asp Val Glu
20 25 30
Arg Thr Leu Gly Lys Gly Asn Phe Ala Val Val Lys Leu Ala Arg
35 40 45
His Arg Val Thr Lys Thr Gln Val Ala Ile Lys Ile Ile Asp Lys
50 55 60
Thr Arg Leu Asp Ser Ser Asn Leu Glu Lys Ile Tyr Arg Glu Val
65 70 75
Gln Leu Met Lys Leu Leu Asn His Pro Asn Ile Ile Lys Leu Tyr
80 85 90
Gln Val Met Glu Thr Lys Asp Met Leu Tyr Ile Val Thr Glu Phe
95 100 105
Ala Lys Asn Gly Glu Met Phe Asp Tyr Leu Thr Ser Asn Gly His
110 115 120
Leu Ser Glu Asn Glu Ala Arg Lys Lys Phe Trp Gln Ile Leu Ser
125 130 135
Ala Val Glu Tyr Cys His Asn His His Ile Val His Arg Asp Leu
140 145 150
Lys Thr Glu Asn Leu Leu Leu Asp Gly Asn Met Asp Ile Lys Leu
155 160 165
Ala Asp Phe Gly Phe Gly Asn Phe Tyr Lys Pro Gly Glu Pro Leu
170 175 180
Ser Thr Trp Cys Gly Ser Pro Pro Tyr Ala Ala Pro Glu Val Phe
185 190 195
Glu Gly Lys Glu Tyr Glu Gly Pro Gln Leu Asp Ile Trp Ser Leu
200 205 210
Gly Val Val Leu Tyr Val Leu Val Cys Gly Ser Leu Pro Phe Asp
215 220 225
Gly Pro Asn Leu Pro Thr Leu Arg Gln Arg Val Leu Glu Gly Arg
230 235 240
Phe Arg Ile Pro Phe Phe Met Ser Gln Asp Cys Glu Thr Leu Ile
245 250 255
Arg Arg Met Leu Val Val Asp Pro Ala Lys Arg Ile Thr Ile Ala
260 265 270
Gln Ile Arg Gln His Arg Trp Met Gln Ala Asp Pro Thr Leu Leu
275 280 285
Gln Gln Asp Asp Pro Ala Phe Ser Met Gln Gly Tyr Thr Ser Asn
290 295 300
Leu Gly Asp Tyr Asn Glu Gln Val Leu Gly Ile Met Gln Ala Leu
305 310 315
Gly Ile Asp Arg Gln Arg Thr Val Glu Ser Leu Gln Asn Ser Ser
320 325 330
Tyr Asn His Phe Ala Ala Ile Tyr Tyr Leu Leu Glu Arg Leu
335 340 345
Arg Glu His Arg Ser Thr Gln Pro Ser Ser Arg Ala Thr Pro Ala
350 355 360
Pro Ala Arg Gln Pro Gln Leu Arg Asn Ser Asp Leu Ser Ser Leu
365 370 375
Glu Val Pro Gln Glu Ile Leu Pro Cys Asp Pro Phe Arg Pro Ser
380 385 390
Leu Leu Cys Pro Gln Pro Gln Ala Leu Ala Gln Ser Val Leu Gln
395 400 405

```

WO 01/68825

PCT/US00/07715

Ala Glu Ile Asp Cys Asp Leu His Ser Ser Leu Gln Pro Leu Phe  
 410 415 420  
 Phe Pro Leu Asp Thr Asn Cys Ser Gly Val Phe Arg His Arg Ser  
 425 430 435  
 Ile Ser Pro Ser Ser Leu Leu Asp Thr Ala Ile Ser Glu Glu Ala  
 440 445 450  
 Arg Gln Gly Pro Ser Leu Glu Glu Glu Gln Glu Val Gln Glu Pro  
 455 460 465  
 Leu Pro Gly Ser Thr Gly Arg Arg His Thr Leu Ala Glu Val Ser  
 470 475 480  
 Thr His Phe Ser Pro Leu Asn Pro Pro Cys Ile Ile Val Ser Ser  
 485 490 495  
 Ser Ala Ala Val Ser Pro Ser Glu Gly Thr Ser Ser Asp Ser Cys  
 500 505 510  
 Leu Pro Phe Ser Ala Ser Glu Gly Pro Ala Gly Leu Gly Gly Gly  
 515 520 525  
 Leu Ala Thr Pro Gly Leu Leu Gly Thr Ser Ser Pro Val Arg Leu  
 530 535 540  
 Ala Ser Pro Phe Leu Gly Ser Gln Ser Ala Thr Pro Val Leu Gln  
 545 550 555  
 Ser Gln Ala Gly Leu Gly Ala Thr Val Leu Pro Pro Val Ser Phe  
 560 565 570  
 Gln Glu Gly Arg Arg Ala Ser Asp Thr Ser Leu Thr Gln Gly Leu  
 575 580 585  
 Lys Ala Phe Arg Gln Leu Arg Lys Asn Ala Arg Thr Lys Gly  
 590 595 600  
 Phe Leu Gly Leu Asn Lys Ile Lys Gly Leu Ala Arg Gln Val Cys  
 605 610 615  
 Gln Ser Ser Ile Arg Gly Ser Arg Gly Gly Met Ser Thr Phe His  
 620 625 630  
 Thr Pro Ala Pro Ser Ser Gly Leu Gln Gly Cys Thr Ala Ser Ser  
 635 640 645  
 Arg Glu Gly Arg Ser Leu Leu Glu Glu Val Leu His Gln Gln Arg  
 650 655 660  
 Leu Leu Gln Leu Gln His His Ser Ala Val Ser Ser Asp Tyr Gln  
 665 670 675  
 Gln Ala Pro Gln Leu Ser Pro Val Pro Tyr Val Leu Thr Pro Cys  
 680 685 690  
 Asp Gly Leu Leu Val Ser Gly Ile Pro Leu Leu Pro Thr Pro Leu  
 695 700 705  
 Leu Gln Pro Gly Met Ser Pro Val Ala Ser Ala Ala Gln Leu Leu  
 710 715 720  
 Asp Ala His Leu His Ile Ser Ala Gly Pro Val Ala Leu Pro Thr  
 725 730 735  
 Gly Pro Leu Pro Gln Cys Leu Thr Arg Leu Ser Pro Ser Cys Asp  
 740 745 750  
 Pro Ala Gly Leu Pro Gln Gly Asp Cys Glu Met Glu Asp Leu Thr  
 755 760 765  
 Ser Gly Gln Arg Gly Thr Phe Val Leu Val Gln  
 770 775  
 <210> 35  
 <211> 604  
 <212> PRT  
 <213> *Drosophila melanogaster*  
 <220>  
 <221> misc\_feature  
 <223> GenBank Accession No: g2564680  
 <220>  
 <221> unsure  
 <222> 52, 291, 354  
 <223> unknown or other  
 <400> 35  
 Met Thr Ala Ala Asn Thr Asn Lys Thr Thr Asp Lys Glu Asn Asp  
 1 5 10 15

WO 01/68825

PCT/US00/07715

Pro Gly Pro Asn Thr Ser Ile Ser Thr Thr Ala Thr Pro Pro Ser  
20 25 30  
Ala Ala Ala Gln Asp Val Gly Gly Cys Val Gly Ser Ser Gly Gly  
35 40 45  
Arg Ser Ser Pro Lys Phe Xaa Ser Tyr Val Asn Gly Asn Gly Tyr  
50 55 60  
Gly Val Tyr Lys Ile Ile Lys Thr Leu Gly Lys Gly Asn Phe Ala  
65 70 75  
Lys Val Lys Leu Ala Ile His Leu Pro Thr Gly Arg Glu Val Ala  
80 85 90  
Ile Lys Leu Ile Asp Lys Thr Ala Leu Asn Thr Ile Ala Arg Gln  
95 100 105  
Lys Leu Tyr Arg Glu Val Asn Ile Met Lys Lys Leu Asn His Pro  
110 115 120  
Asn Ile Val Arg Leu Gln Val Ile Glu Ser Glu Arg Thr Leu  
125 130 135  
Tyr Leu Val Met Glu Tyr Val Ser Gly Gly Glu Leu Phe Asn Tyr  
140 145 150  
Leu Val Lys Asn Gly Arg Met Arg Glu Arg Asp Ala Arg Val Leu  
155 160 165  
Phe Arg Gln Leu Val Ser Ala Ile Glu Tyr Cys His Ser Lys Ser  
170 175 180  
Ile Val His Arg Asp Leu Lys Ala Glu Asn Leu Leu Leu Asp Gln  
185 190 195  
Gln Met Lys Leu Lys Ile Ala Asp Phe Gly Phe Ser Thr Thr Phe  
200 205 210  
Glu Pro Lys Ala Pro Leu Glu Thr Phe Cys Gly Ser Pro Pro Tyr  
215 220 225  
Ala Ala Pro Glu Leu Phe Lys Gly Lys Lys Tyr Ser Gly Pro Glu  
230 235 240  
Val Asp Ser Trp Ser Leu Gly Val Val Leu Tyr Thr Leu Val Ser  
245 250 255  
Gly Ser Leu Pro Phe Asp Gly Thr Asn Leu Lys Glu Leu Arg Asp  
260 265 270  
Arg Val Leu Arg Gly Lys Tyr Arg Val Pro Tyr Tyr Val Ser Ile  
275 280 285  
Glu Cys Glu Ser Leu Xaa Arg Lys Phe Leu Val Leu Asn Pro Thr  
290 295 300  
Gln Arg Thr Ser Leu Ser Ala Val Met Ala Asp Arg Trp Ile Asn  
305 310 315  
Met Gly Tyr Glu Gln Gly Asn Gly Leu Arg Pro Phe Gln Glu Lys  
320 325 330  
Pro Met Asp Leu His Asp Val Asn Arg Leu Ser Leu Leu Ser Asn  
335 340 345  
Met Gly His Lys Pro Arg Asp Val Xaa Gln Ser Leu Lys Asn Gln  
350 355 360  
Lys Phe Asp Asp Ile Tyr Cys Ala Tyr Met Leu Leu Asp Val Ala  
365 370 375  
Lys Pro Arg Ser Thr Ala Cys Ser Glu Lys Ser Gly Ser Ser Phe  
380 385 390  
Arg Glu Thr Pro Thr Ala Met Pro Gly Ser Ser Arg Ile Pro Val  
395 400 405  
Pro Ile Ala Ala Pro Asn Val Thr Ile Ser Gln Val Thr Phe Ala  
410 415 420  
Leu Asp Lys Ser Thr Pro Asn Arg Pro Gly Ala Thr Ser Ile Arg  
425 430 435  
Pro Met Ala Pro Arg Leu Ala Asn Ala Leu Thr Pro Leu Pro Leu  
440 445 450  
Thr Pro Pro Pro Lys Lys Tyr Ile Cys Cys Ser Ala Ser Lys Ala  
455 460 465  
Ala Asn Pro Arg Arg Ser Glu Pro Ser Ser Ile Pro Gln Ser Ala  
470 475 480  
Met Pro Lys Lys Gly Val Gly Ser Pro Val Asp Val Lys Thr Thr  
485 490 495  
Leu Leu Ser Ala Gln Arg Lys Leu Ala Val Asn His Lys Leu Thr  
500 505 510  
Ser Ala Ser His Gln Ile Arg Ser Pro Ile Thr Gln Ser Ser Ser

WO 01/68825

PCT/US00/07715

Gln	Ala	Ser	Glu	Cys	Thr	Arg	Thr	Pro	520	Thr	His	Phe	Glu	Met	525
									535						540
Leu	Asp	Ser	Thr	Ser	Thr	Pro	Leu	Lys	Val	Leu	Lys	Leu	Val	Ala	545
									550						555
Ser	Asn	Ser	Gln	Thr	Pro	Pro	Ser	Thr	Glu	Asn	Ile	Asn	Arg	Pro	560
									565						570
Thr	Arg	Val	Gly	Phe	Phe	Ser	Lys	Leu	Ser	Ala	Arg	Phe	Val	Arg	575
									580						585
Arg	Ser	Leu	His	Lys	Gly	Glu	Lys	Asp	Ile	Ser	Glu	Gln	Gly	Arg	590
									595						600
Asn	Leu	Thr	Lys												

<210> 36

<211> 745

<212> PRT

<213> Homo sapiens

<220>

<221> misc\_feature

<223> GenBank Accession No: g1749794

<400> 36

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

WO 01/68825

PCT/US00/07715

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

<210> 37  
 <211> 28  
 <212> DNA  
 <213> Homo sapiens  
  
 <220>  
 <221> misc\_feature  
 <223> primer: forward  
  
 <400> 37

WO 01/68825	PCT/US00/07715
ggttagttac aagcagtcta gtgcaatt	28
<210> 38	
<211> 25	
<212> DNA	
<213> Homo sapiens	
<220>	
<221> misc_feature	
<223> primer: reverse	
<400> 38	
tgctaccaga ccctaagaa agaga	25
<210> 39	
<211> 28	
<212> DNA	
<213> Homo sapiens	
<220>	
<221> misc_feature	
<223> Primer: taqman	
<400> 39	
ccctctgtag acacgggacc cttttttg	28

## 【 国際調査報告 】

## INTERNATIONAL SEARCH REPORT

International Application No  
PCT/US 00/07715

A. CLASSIFICATION OF SUBJECT MATTER		
IPC 7	C12N9/12	C12N15/54 C12N5/10 C12Q1/68 C07K16/40
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		
Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched		
Electronic data base consulted during the international search (name of data base and, where practical, search terms used)		
C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	DATABASE GENEMBL [Online] 26 April 1999 (1999-04-26) STRAUSBERG, R.: "tf82g08.x1 NCI CGAP Brn23 Homo sapiens cDNA clone IMAGE:2105822 3?, mRNA sequence" XP002158330 Accession AI609751 --- -/-	1-10, 12-18
<input checked="" type="checkbox"/> Further documents are listed in the continuation of box C. <input type="checkbox"/> Patent family members are listed in annex.		
* Special categories of cited documents:		
*A* document defining the general state of the art which is not considered to be of particular relevance *E* earlier document but published on or after the international filing date *L* document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified) *O* document referring to an oral disclosure, use, exhibition or other means *P* document published prior to the international filing date but later than the priority date claimed *T* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention *X* document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone *Y* document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art. *Z* document member of the same patent family		
Date of the actual completion of the international search	Date of mailing of the international search report	
29 January 2001	06.06.01	
Name and mailing address of the ISA European Patent Office, P.O. 5018 Patentlaan 2 NL-2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Authorized officer ALCONADA RODRIG., A	

Form PCT/ISA/210 (second sheet) (July 1992)

## INTERNATIONAL SEARCH REPORT

 International Application No  
 PCT/US 00/07715

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT		
Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	DATABASE GENEMBL [Online] 25 April 1997 (1997-04-25) DREWES ET AL.: "R.norvegicus mRNA for serine/threonine kinase MATK2" XP002158331 Accession Z83869 -& DREWES ETAL.: "MARK, a novel family of protein kinases that phosphorylate microtubule-associated proteins and trigger microtubule disruption" CELL, vol. 89, no. 2, 18 April 1997 (1997-04-18), pages 297-308, XP002105681 figures 1,4,6 ---	15-20
A	DATABASE SWISSPROT [Online] 1 January 1998 (1998-01-01) THOMAS, K.: "F49C5.4 protein bearing protein kinase motifs" XP002158332 Accession O17879 ---	15-20
A	BALDIN VERONIQUE ET AL: "Evidence for a mammalian NIM1-like kinase pathway acting at the G0-1/S transition." BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS, vol. 236, no. 1, 1997, pages 130-134, XP000979000 ISSN: 0006-291X the whole document -----	1-20

INTERNATIONAL SEARCH REPORT

International application No.  
PCT/US 00/07715

**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:
  
- 2.  Claims Nos.:  
because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:
  
- 3.  Claims Nos.:  
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

**Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)**

This International Searching Authority found multiple inventions in this international application, as follows:

- 1.  As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.
  
- 2.  As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
  
- 3.  As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:
  
- 4.  No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:  
1, 2, 7-9, 13-20 compl. and 3-6, 10-12 part.

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 00/07715

FURTHER INFORMATION CONTINUED FROM PCT/SA/ 210

1. Claims: 1,2, 7-9, 13-20 (complete) and 3-6, 10-12 (partially)

A purified nucleic acid molecule corresponding to SEQ ID NO:1, which encodes the polypeptide the human NIM1 kinase of SEQ ID NO:2, mammalian variants thereof (selected from SEQ ID NO:24-28 and 30) and fragments thereof (SEQ ID NOs: 3-23); a fragment of at least 18 consecutive nucleotides from said sequences; expression vectors, host cells, a method for producing the recombinant protein, a method of detecting a nucleic acid molecule in a sample comprising the steps of hybridizing said nucleic acid molecule and detecting the hybridization complex; said method used to diagnose a brain disorder or cancer; said method involving amplifying the nucleic acid molecules of the sample prior to hybridization; said method used to diagnose a brain disorder or cancer, a method of using a nucleic acid molecule to screen a library of molecules or compounds; an isoated protein comprising SEQ ID NO:2 or a portion thereof; a method to screen a library of molecules or compounds using the polypeptide of SEQ ID NO:2; a method of preparing an antibody comprising immunizing an animal with the polypeptide of SEQ ID NO:2; a method of using an antibody which specifically binds NIM1 kinase to diagnose a disease, preferably, when the disease is a brain disorder or cancer.

2. Claims: 3,4-6,10-12 (partially)

A fragment of at least 18 consecutive nucleotides from SEQ ID NO:29 and the complements thereof; a substrate containing said fragment; a probe containing said fragment; a method of detecting a nucleic acid molecule in a sample comprising the steps of hybridizing said nucleic acid molecule and detecting the hybridization complex; said method used to diagnose a brain disorder or cancer; said method involving amplifying the nucleic acid molecules of the sample prior to hybridization.

## フロントページの続き

(51) Int.Cl. <sup>7</sup>	F I	テーマコード(参考)
C 1 2 N 9/12	C 1 2 P 21/08	4 C 0 8 4
C 1 2 P 21/08	C 1 2 Q 1/48	Z
C 1 2 Q 1/48	C 1 2 Q 1/68	A
C 1 2 Q 1/68	G 0 1 N 33/15	Z
G 0 1 N 33/15	G 0 1 N 33/50	Z
G 0 1 N 33/50	G 0 1 N 33/53	M
G 0 1 N 33/53	G 0 1 N 33/566	
G 0 1 N 33/566	G 0 1 N 33/573	A
G 0 1 N 33/573	G 0 1 N 33/574	A
G 0 1 N 33/574	A 6 1 K 37/52	

(81) 指定国 AP(GH,GM,KE,LS,MW,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), OA(BF,BJ,CF,CG,CI,CM,GA,GN,GW,ML,MR,NE,SN,TD,TG), AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, 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, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZA, ZW

- (72) 発明者 バンドマン、オルガ  
アメリカ合衆国カリフォルニア州 9 4 0 4 3 ・ マウンテンビュー ・ アンナアベニュー 3 6 6
- (72) 発明者 マニャギ、パオラ  
イタリア国ミラノ市アイ - 2 0 1 3 2 ・ ビアパドーバ 1 9 6
- (72) 発明者 ボソッティ、ロベルタ  
イタリア国ネルビアアーノ市アイ - 2 0 0 1 4 ・ ビアピエトロマスカーニ 2
- (72) 発明者 スカケーリ、エマニュエラ  
イタリア国ミラノ市アイ - 2 0 1 4 6 ・ ビアデセンツァーノ 7
- (72) 発明者 イサッキ、アントーネラ  
イタリア国ミラノ市アイ - 2 0 1 0 0 ・ ビアモンテカティーニ 1 4
- (72) 発明者 ホジソン、デイビッド・エム  
アメリカ合衆国カリフォルニア州 9 4 3 0 4 ・ パロアルト ・ アディソンアベニュー 5 6 7

F ターム(参考) 2G045 AA25 AA26 AA29 AA40 BA13 BB20 CB01 CB09 CB16 CB17  
DA12 DA13 DA14 DA20 DA77 DA78 FB01 FB02 FB03 FB07  
4B024 AA12 BA10 CA04 DA02 EA02 EA04 GA11  
4B050 CC03 DD11 LL01 LL03  
4B063 QA19 QQ08 QQ43 QR32 QR55 QS34  
4B064 AG27 CA10 CA19 CA20 CC24 DA14  
4C084 AA13 DC25 NA14 ZA01 ZB26

专利名称(译)	人类NIM1激酶		
公开(公告)号	<a href="#">JP2004503211A</a>	公开(公告)日	2004-02-05
申请号	JP2001567309	申请日	2000-03-22
[标]申请(专利权)人(译)	莫提了两个安吉拉 MOLTENI ANGELA		
申请(专利权)人(译)	洞察制药公司 MOLTENI, 安吉拉		
[标]发明人	バンドマンオルガ マニャギパオラ ボソッティロベルタ スカケーリエマニュエラ イサッキアントーネラ ホジソンデイビッドエム		
发明人	バンドマン、オルガ マニャギ、パオラ ボソッティ、ロベルタ スカケーリ、エマニュエラ イサッキ、アントーネラ ホジソン、デイビッド・エム		
IPC分类号	G01N33/50 A61K38/45 A61K48/00 A61P25/00 A61P35/00 C12N9/12 C12N15/09 C12N15/54 C12P21/08 C12Q1/48 C12Q1/68 G01N33/15 G01N33/53 G01N33/566 G01N33/573 G01N33/574		
CPC分类号	A61P25/00 C12N9/1205		
FI分类号	C12N15/00.ZNA.A A61K48/00 A61P25/00 A61P35/00 C12N9/12 C12P21/08 C12Q1/48.Z C12Q1/68.A G01N33/15.Z G01N33/50.Z G01N33/53.M G01N33/566 G01N33/573.A G01N33/574.A A61K37/52		
F-TERM分类号	2G045/AA25 2G045/AA26 2G045/AA29 2G045/AA40 2G045/BA13 2G045/BB20 2G045/CB01 2G045/CB09 2G045/CB16 2G045/CB17 2G045/DA12 2G045/DA13 2G045/DA14 2G045/DA20 2G045/DA77 2G045/DA78 2G045/FB01 2G045/FB02 2G045/FB03 2G045/FB07 4B024/AA12 4B024/BA10 4B024/CA04 4B024/DA02 4B024/EA02 4B024/EA04 4B024/GA11 4B050/CC03 4B050/DD11 4B050/LL01 4B050/LL03 4B063/QA19 4B063/QQ08 4B063/QQ43 4B063/QR32 4B063/QR55 4B063/QS34 4B064/AG27 4B064/CA10 4B064/CA19 4B064/CA20 4B064/CC24 4B064/DA14 4C084/AA13 4C084/DC25 4C084/NA14 4C084/ZA01 4C084/ZB26		
优先权	09/523849 2000-03-13 US		
外部链接	<a href="#">Espacenet</a>		

#### 摘要(译)

本发明提供了人NIM1激酶，编码核酸分子和与蛋白质特异性结合的抗体。本发明还提供了这些组合物在表征，诊断和治疗与NIM1激酶表达有关的疾病中的用途。

# KINASE ASSAY USING hNIM1

hNim1_GST	+	+	-	+	-	
MBP	-	+	+	-	-	Zap70
HIST.1	-	-	-	+	+	

