

(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2005-503107**(P2005-503107A)**(43) 公表日 **平成17年2月3日(2005.2.3)**

| (51) Int. Cl. ⁷ | F I | テーマコード (参考) |
|----------------------------|------------------------------------|-------------|
| C 1 2 N 15/09 | C 1 2 N 15/00 Z N A A | 2 G O 4 5 |
| A 6 1 K 39/395 | A 6 1 K 39/395 D | 4 B O 2 4 |
| A 6 1 P 13/10 | A 6 1 K 39/395 N | 4 B O 6 3 |
| A 6 1 P 35/00 | A 6 1 P 13/10 | 4 B O 6 4 |
| C O 7 K 14/47 | A 6 1 P 35/00 | 4 B O 6 5 |
| | 審査請求 未請求 予備審査請求 有 (全 173 頁) 最終頁に続く | |

| | | | |
|---------------|------------------------------|----------|--|
| (21) 出願番号 | 特願2002-556752 (P2002-556752) | (71) 出願人 | 301005050 インサイト・ゲノミックス・インコーポレイテッド アメリカ合衆国カリフォルニア州94304・パロアルト・ポータードライブ 3160 |
| (86) (22) 出願日 | 平成14年1月8日 (2002.1.8) | (74) 代理人 | 100089266 弁理士 大島 陽一 |
| (85) 翻訳文提出日 | 平成15年7月7日 (2003.7.7) | (72) 発明者 | レディ、ルーパ アメリカ合衆国カリフォルニア州94086・サニーベイル・#3・ウェストマッキンレーアベニュー 1233 |
| (86) 国際出願番号 | PCT/US2002/000647 | | |
| (87) 国際公開番号 | W02002/055706 | | |
| (87) 国際公開日 | 平成14年7月18日 (2002.7.18) | | |
| (31) 優先権主張番号 | 09/757,781 | | |
| (32) 優先日 | 平成13年1月9日 (2001.1.9) | | |
| (33) 優先権主張国 | 米国 (US) | | |

最終頁に続く

(54) 【発明の名称】 A S I P 関連タンパク質

(57) 【要約】

本発明はASTP関連タンパク質をコードするcDNAを提供する。本発明はまた、cDNAの利用のために、自身の断片、補体、及び変異体を提供し、コードされたタンパク質の利用のために、とりわけ膀胱移行上皮癌のような癌の診断及び治療のために自身の一部及び抗体を提供する。本発明はまた、タンパク質及び遺伝子組換モデルシステムの作成のための発現ベクターおよび宿主細胞をも提供する。

【特許請求の範囲】

【請求項 1】

SEQ ID NO:1もしくはSEQ ID NO:2のアミノ酸配列を有するタンパク質をコードする単離されたcDNA。

【請求項 2】

SEQ ID NO:1のアミノ酸配列を有するタンパク質をコードする単離されたcDNA。

【請求項 3】

SEQ ID NO:2のアミノ酸配列を有するタンパク質をコードする単離されたcDNA。

【請求項 4】

単離されたcDNAであって、

10

- (a) SEQ ID NO:3もしくはSEQ ID NO:20の核酸配列、若しくはその補体と、
- (b) SEQ ID NO:4乃至11から選択されたSEQ ID NO:3の断片、若しくはその補体、または、SEQ ID NO:21乃至39から選択されたSEQ ID NO:20の断片、若しくはその補体と、
- (c) SEQ ID NO:12乃至19から選択されたSEQ ID NO:3の変異体、および、SEQ ID NO:40乃至56から選択されたSEQ ID NO:20の変異体とから選択される単離されたcDNA。

【請求項 5】

前記請求項 1 のcDNA若しくはその補体を有する組成物。

【請求項 6】

前記請求項 1 のcDNAを含むベクター。

【請求項 7】

20

前記請求項 6 のベクターを有する宿主細胞。

【請求項 8】

タンパク質を生成するためにcDNAを用いる方法であって、

- (a) タンパク質発現の条件下で、前記請求項 7 の宿主細胞を培養する過程と、
- (b) 前記宿主細胞の培養から前記タンパク質を回収する過程とを含む方法。

【請求項 9】

サンプル中の核酸の発現を検出するのにcDNAを用いる方法であって、

- (a) 前記サンプルの核酸に前記請求項 5 の組成物をハイブリダイズし、ハイブリダイゼーション複合体を生成する過程と、
- (b) ハイブリダイゼーション複合体形成物と標準物質とを比較する過程であって、前記比較過程が前記サンプル中の前記cDNAの発現を示すことを特徴とする過程とを含む方法。

30

【請求項 10】

さらに、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を増幅させる過程を含むことを特徴とする請求項 9 に記載の方法。

【請求項 11】

前記組成物が基質に取り付けられることを特徴とする請求項 9 の方法。

【請求項 12】

前記cDNAが前記標準物質と比較される場合に差動的に発現され、膀胱移行上皮癌の指標となることを特徴とする請求項 9 の方法。

【請求項 13】

40

複数の分子若しくは混合物をスクリーニングするためにcDNAを用いる方法であって、特異的な結合を可能とするための条件下で、前記請求項 1 のcDNAと複数の分子若しくは混合物とを組み合わせる過程と、特異的な結合を検出し、それにより前記cDNAに特異的に結合する分子若しくは混合物を同定する過程とを含む方法。

【請求項 14】

前記分子若しくは混合物が、DNA分子、RNA分子、ペプチド核酸、人工クロモソーム構造、ペプチド、転写要素、リプレッサー、及び調節分子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】

50

精製されたタンパク質若しくはその一部であって、

- (a) SEQ ID NO:1及びSEQ ID NO:2のアミノ酸配列と、
- (b) SEQ ID NO:1及びSEQ ID NO:2の抗原性エピトープと、
- (c) SEQ ID NO:1及びSEQ ID NO:2の生物学的活性部位とから選択される精製されたタンパク質若しくはその一部。

【請求項 16】

前記請求項 15 のタンパク質を含む組成物。

【請求項 17】

少なくとも一つのリガンドを同定するべく、複数の分子若しくは混合物をスクリーニングするためにタンパク質を用いる方法であって、

特異的な結合を可能とするための条件下で、前記請求項 15 のタンパク質と前記分子若しくは混合物とを結合させる過程と、

特異的な結合を検出し、それにより前記タンパク質に特異的に結合するリガンドを同定する過程とを含む方法。

【請求項 18】

前記分子若しくは混合物が、DNA分子、RNA分子、ペプチド核酸、ペプチド、タンパク質、mimetic、アゴニスト、アンタゴニスト、抗体、免疫グロブリン、阻害剤、及び薬剤から選択されることを特徴とする請求項 17 に記載の方法。

【請求項 19】

SEQ ID NO:1もしくはSEQ ID NO:2 のアミノ酸配列を有する精製された哺乳動物タンパク質。

【請求項 20】

前期請求項 19 のタンパク質に特異的に結合する精製された抗体。

【請求項 21】

前記請求項 20 の抗体の特異性を有する多クローン性抗体を調整し精製する方法であって、

(a) 抗体反応を誘発する条件下で、SEQ ID NO:1もしくはSEQ ID NO:2 のタンパク質で動物を免疫する過程と、

(b) 動物の抗体を単離する過程と、

(c) 前記タンパク質を基質に取りつける過程と、

(d) 前記タンパク質に対する特異的結合が可能となるような条件下で、前記基質と単離した抗体とを接触させる過程と、

(e) 前記抗体を前記タンパク質より解離させ、それによって精製された多クローン性抗体を獲得する過程とを含む方法。

【請求項 22】

前記請求項 21 の方法によって生成された多クローン性抗体。

【請求項 23】

前期請求項 20 の抗体の前記特異性を有する単クローン性抗体を調整するための方法であって、

(a) 抗体反応を誘発する条件下でSEQ ID NO:1もしくはSEQ ID NO:2 のタンパク質で動物を免疫する過程と、

(b) 前記動物より抗体生成細胞を単離する過程と、

(c) 培養内で前記抗体生成細胞と不死化細胞とを融合させ、ハイブリドーマ細胞を生成する単クローン性抗体を形成する過程と、

(d) 前記ハイブリドーマ細胞を培養する過程と、

(e) 培養より単クローン性抗体を単離する過程とを含む方法。

【請求項 24】

前記請求項 23 の方法によって生成された単クローン性抗体。

【請求項 25】

抗体を用いて、タンパク質を免疫精製するための方法であって、

10

20

30

40

50

- (a) 基質に請求項 2 0 の抗体を付着させる過程と、
- (b) 抗体：タンパク質複合体を形成可能とする条件下で、タンパク質を有するサンプルに対して前記抗体を露出させる過程と、
- (c) 前記複合体より前記タンパク質を分離させる過程と、
- (d) 前記精製タンパク質を回収する過程とを含む方法。

【請求項 2 6】

抗体を用いて、サンプル内でタンパク質の発現を検出するための方法であって、

- (a) 前記抗体：タンパク質複合体の形成が可能な条件下で、前記請求項 2 0 の抗体とサンプルとを結合させる過程と、
- (b) 前記サンプル内で前記タンパク質の発現を示すことを特徴とする複合体形成物を検出する過程とを有する方法。

【請求項 2 7】

複合体形成物が基準物質と比較され、また、膀胱移行上皮癌と診断されることを特徴とする請求項 2 6 に記載の方法。

【請求項 2 8】

請求項 2 0 の抗体及び標識部分を含む組成物。

【請求項 2 9】

請求項 2 0 の抗体及び薬剤を含む組成物。

【請求項 3 0】

前記請求項 2 9 の成分の、そのような治療を必要とする患者に対する投与を含む膀胱移行上皮癌の治療方法。

【発明の詳細な説明】

【技術分野】

【0 0 0 1】

本発明は、ASIP関連タンパク質をコードするcDNAに関し、また、とりわけ膀胱移行上皮癌のような癌に於ける診断及び治療でcDNA及びコードされたタンパク質を用いることに関する。

【背景技術】

【0 0 0 2】

生物間系統発生的関係は何回も実証されてきた。そして広範な原核細胞および真核細胞生物の研究により分子、生化学的メカニズム、生理学的メカニズム、代謝経路の多少の漸進的な進化が示されてきた。進化ストレスが異なるが、線虫、ハエ、ラット、ヒトのタンパク質には共通の化学的及び構造的特徴があり、一般的に同じ細胞機能を行う。構造と機能が公知である生物種での核酸とタンパク質配列の比較は、ヒト配列の研究を加速する。またヒトの症状、疾患または障害に対する診断と薬剤をテストするモデル系の発達を可能にする。

【0 0 0 3】

プロテインキナーゼC (PKC) は、脂質誘導セカンドメッセンジャーを介した細胞内の情報伝達の役割を有する (Nishizuka (1995) FASEB J 9:484-496)。PKCの異なるアイソフォームは、異なった組織及び細胞下分布を有し、脂質及びカルシウムに対する特異な反応を示し、また異なった生理学的機能を供給しうる。PKCアイソフォームは、従来型、新型、及び異形の3つに分類される。異形PKC (αPKC) アイソフォームは、アフリカツメガエル卵成熟 (Dominguezら (1992) Mol Cell Biol 12:3776-3783)、proliferation and survival of 線維芽細胞の増殖及び生存 (Berra ら (1993) Cell 74:555-563)、differentiation of PC12及び白血病細胞の分化 (Wootenら (1994) Cell Growth Differ 5:395-403; Waysら (1994) Cell Growth Differ 5:1195-1203)、マイトジェン活性化タンパク質キナーゼの活性化及び遺伝子発現 (Berra ら (1995) EMBO J 14:6157-6163; Lozano ら (1994) J Biol Chem 269:19200-19202)、インシュリン誘発グルコース接種及び原形質膜に対するグルコース輸送体 GLUT4 のトランスロケーション (Kotaniら (1998) Mol Cell Biol 18:6971-6982)、及び細胞極性の確率及び / 又はメンテナンス (Brazil及びHemmings (2000) Curr Biol 10 40 50

生物学的活性化部位を含む群から選択されるタンパク質またはその一部をコードする単離されたタンパク質またはその断片を提供する。本発明はまた、SEQ ID NO:3及びSEQ ID NO:20の核酸配列、SEQ ID NO:4乃至11から選択されたSEQ ID NO:3の断片もしくはSEQ ID NO:21乃至39から選択されたSEQ ID NO:20の断片、及びSEQ ID NO:12-19から選択されたSEQ ID NO:3もしくはSEQ ID NO:40-56から選択されたSEQ ID NO:20の変異体とを含む群から選択される単離されたcDNAまたはその相補体を提供する。更に本発明は、ARP-1もしくはARP-2をコードするcDNAまたはその相補体を含む組成物、基板、及びプローブを提供する。更に本発明は、このようなcDNAを含むベクター、このようなベクターを含む宿主細胞、及びこのようなcDNAを用いてARP-1もしくはARP-2を作製する方法を提供する。本発明はさらにARPをコードするcDNAを有するベクターを備えた遺伝子組み換え株化細胞及び生物

を提供する。また本発明はさらにSEQ ID NO:3-56を含む群より選択された断片もしくはその補体を提供する。或る実施態様では、本発明は少なくとも一つのそれら断片を含む基板を提供する。第2の実施態様では、本発明は、検出法、スクリーニング法、及び精製法に用いることが可能な断片を含むプローブを提供する。更なる実施態様では、プローブは一本鎖相補RNAまたはDNA分子である。

10

20

30

40

50

【0009】

本発明はまた、cDNAを用いて、サンプルの核酸の発現変動を検出する方法であって、前記プローブと前記核酸とを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し、そのハイブリダイゼーション複合体を標準と比較し、その比較によりサンプルにおける前記cDNAの発現変動が分かる方法を提供する。或る実施態様では、この検出方法は、ハイブリダイゼーションの前に前記核酸を増幅するステップを含む。別の実施態様では、この検出方法はcDNAの異なった発現が、とりわけ膀胱移行上皮癌のような癌の診断に用いられることを示している。また別の側面では、cDNA若しくは断片、若しくは補体はエレメント若しくはアレイを含んでも良い。

【0010】

本発明はまた、cDNA、その断片またはその相補体を用いて、ライブラリ即ち複数の分子または化合物をスクリーニングしてそのcDNA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cDNAを前記分子または化合物と結合させるステップと、前記cDNAに対する特異的な結合を検出して、前記cDNAと特異的に結合するリガンドを特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、このような分子または化合物として、アプタマー、DNA分子、RNA分子、ペプチド核酸、人工染色体作製物、ペプチド、転写因子、リプレッサー、及び調節分子が選択される。

【0011】

本発明はまた、SEQ ID NO:1もしくはSEQ ID NO:2のアミノ酸配列と、SEQ ID NO:1もしくはSEQ ID NO:2のアミノ酸配列と95%の同一性を有する変異体と、SEQ ID NO:1もしくはSEQ ID NO:2の抗原エピトープと、SEQ ID NO:1もしくはSEQ ID NO:2の生物学的に活性部分とからなる群から選択される精製された哺乳類タンパク質またはその一部を提供する。本発明はまた、前記精製されたタンパク質、若しくは医薬用担体に関連する前記タンパク質の部分を含む組成物を提供する。本発明は更に、とりわけ膀胱移行上皮癌のような癌を有する被験者を治療するためにARP-1もしくはARP-2を用い、そのような治療を必要とする患者に生成したタンパク質を含む合成物を投与する方法を提供する。本発明は更に、タンパク質を用いてライブラリ即ち複数の分子または化合物をスクリーニングして少なくとも1つのリガンドを特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、そのタンパク質を前記分子または化合物と結合させるステップと、特異的な結合を検出して、前記タンパク質と特異的に結合するリガンドを特定するステップとを含む方法を提供する。ある実施態様においては、分子または化合物は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。別の実施態様では、とりわけ膀胱移行上皮癌のような癌の患者を治療するために前記リガンドを使用する。

【0012】

本発明は、タンパク質を用いて、被検サンプルをそのタンパク質に特異的に結合する抗体に対してスクリーニングする方法であって、前記被検サンプルから抗体を単離するステップと、特異的な結合が許容される条件下で、この単離した抗体を前記タンパク質とを結合させるステップと、結合した前記タンパク質から抗体を解離するステップと、抗体の存在或いは存在量からとりわけ膀胱移行上皮癌のような癌の患者を診断できる抗体の量を既知の標準値と比較するステップとを含むスクリーニング法を提供する。

【0013】

本発明はまた、タンパク質を用いて抗体を作製及び精製する方法であって、抗体反応が起こる条件下でこのタンパク質によって動物を免疫するステップと、動物抗体を単離するステップと、基板にこのタンパク質を付着させるステップと、このタンパク質への特異的な結合が許容される条件下で、この基板を単離した抗体と接触させること、このタンパク質から抗体を分離して精製した抗体を得ることを含む方法を提供する。

10

【0014】

本発明は、とりわけ膀胱移行上皮癌のような癌において発現するタンパク質に特異的に結合する精製された抗体を提供する。本発明はまた、とりわけ膀胱移行上皮癌のような癌を診断する方法であって、抗体を混合して結合抗体量を既知の指標と比較するステップを含み、それによって、とりわけ膀胱移行上皮癌のような癌の存在を証明する診断方法を提供する。本発明は更に、抗体を用いてとりわけ膀胱移行上皮癌のような癌を治療する方法であって、精製された抗体を含む医薬組成物をそのような治療が必要な患者に投与することを含む治療方法を提供する。

20

【0015】

本発明はまた、哺乳動物のゲノムDNAの中にマーカー遺伝子を挿入して、内在性ポリヌクレオチドの発現を阻害する方法を提供する。本発明はまた、cDNAを用いて哺乳動物モデル系を作製する方法であって、SEQ ID NO:2-34から選択されるcDNAを含むベクターを作製するステップと、そのベクターで胚性幹細胞を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、この形質転換された胚性幹細胞を哺乳動物胚盤胞の中に微量注入して、キメラ胚盤胞を形成するステップと、偽妊娠メスにキメラ胚盤胞を導入し、このメスが、生殖細胞系の中にcDNAを含むキメラ子孫を出産し、そのキメラ哺乳動物を交配して同型接合哺乳動物モデル系を作製するステップとを含む哺乳動物モデル系作製方法を提供する。

30

【発明を実施するための最良の形態】

【0016】

(本発明の記載について)

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。

40

【0017】

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。

【0018】

(定義)

「ARP」は、天然、合成、半合成或いは組換え体など任意の哺乳動物種(ウシ、イヌ、マウス、ヒツジ、ブタ、げっ歯類、サルそして好ましくはヒトを含む)から得られる実

50

質的に精製されたタンパク質を指す。

【0019】

「アレイ」は、基板上の少なくとも2つのcDNAの規則正しい配列を指す。cDNAの少なくとも1つが調節または標準を表す。そして他方が目的の診断のcDNA体を表す。基質上の2から約4万のcDNAの構成により、cDNAとサンプル核酸の間で形成される各標識ハイブリダイゼーション複合体の大きさ及びシグナル強度を確実に個別に区別できる。

【0020】

配列表の核酸分子の「相補配列」は、完全長配列に完全に相補的なcDNAを指す。また高いストリンジエンシー条件下でcDNA或いはmRNAにハイブリダイズする。

【0021】

「cDNA」とは、単離したポリヌクレオチド、核酸分子或いはその任意の断片または相補配列を指す。それは組換えまたは合成された二本鎖または一本鎖であり、コード配列及び/または非コード5'及び3'配列を有する。

【0022】

「タンパク質をコードするcDNA」という語は、当分野で周知の分析により同定された保存された領域、モチーフ或いはドメインをコードする配列と密接にアラインメントする核酸配列を指す。これらの分析には、保存された領域内における同一性を特定するBLAST (Basic Local Alignment Search Tool)が含まれる(Altschul (1993) J Mol Evol 36: 290-300、Altschul ら(1990) J Mol Biol 215:403-410)。

【0023】

「誘導体」とは化学修飾されたcDNA或いはタンパク質を指す。cDNAの誘導体化にはクエオシン(queosine)或いはヒポキサンチンなどの類似体等非従来型塩基の置換が含まれ得る。これらの置換は当分野で周知である。タンパク質の誘導体化にはアセチル基、アシル基、アルキル基、アミノ基、ホルミル基またはモルホリン基による水素の置換が含まれる。分子誘導体は天然分子の生物学的活性を保持するが、長い寿命或いは強化された活性などの長所を付与し得る。

【0024】

「発現変動」とは、サンプル中の転写されたメッセンジャーRNA或いは翻訳されたタンパク質の量の存在の有無、もしくは少なくとも二倍の変更により検出される増加即ちアップレギュレーション又は存在、或いは減少即ちダウンレギュレーション又は不在を指す。

【0025】

「障害」とは、とりわけ膀胱移行上皮癌のような癌のような、cDNA及びARPが異なって発現されている場合の症状、疾患、または症候群を指す。

【0026】

「断片」とは長さが約200から約700塩基の連続したヌクレオチド鎖を指す。断片は、関連する核酸分子を同定するためにPCRまたはハイブリダイゼーション技術に、或いはリガンドをスクリーニングするための結合アッセイで使用され得る。核酸とそれらのリガンドはこのような方法で同定され、複製、転写、または翻訳を調節する治療として有用である。

【0027】

「ハイブリダイゼーション複合体」とは、例えば5'-A-G-T-C-3'と3'-T-C-A-G-5'との塩基対などのように1つの分子のプリンが相補的な分子ピリミジンと水素結合して、cDNAとサンプルの核酸との間で形成される。相補性の度合およびヌクレオチド類似体の使用が、ハイブリダイゼーション反応の効率とストリンジエンシーに影響を与える。

【0028】

「リガンド」とは、ポリヌクレオチド或いはcDNA分子上の相補部位、又はタンパク質若しくはエピトープに特異結合するあらゆる物質、分子、または化合物を指す。そのようなリガンドは、ポリヌクレオチドまたタンパク質の活性を安定化或いは調節し、核酸、タンパク質、炭水化物、脂肪、及び脂質を含む無機および/または有機物質から構成され得る。

【0029】

10

20

30

40

50

「オリゴヌクレオチド」とは、長さが約18から約60ヌクレオチドの一本鎖分子を指す。そしてハイブリダイゼーションや増幅技術において、または複製、転写、翻訳の調節において使用され得る。概ね等しい用語としては、アンプライマー、プライマー、及びオリゴマーが挙げられる。

【0030】

「部分」とは、あらゆる目的に使用されるタンパク質の任意の部分を指すが、特にリガンドのスクリーニングまたは抗体の生産に用いられるエピトープを指す。

【0031】

タンパク質の「翻訳後修飾」には、脂質化、グリコシル化、リン酸化、アセチル化、ラセミ化、及び蛋白分解性切断等が含まれ得る。これらのプロセスは、合成或いは生化学的に生じ得る。生化学的修飾は、細胞の位置、細胞型、pH、及び酵素環境などによって異なる。

【0032】

「プローブ」とは、サンプル中の少なくとも1つの核酸にハイブリダイズするcDNAを指す。標的が一本鎖である場合、プローブは相補的一本鎖である。プローブは、サザン法、ノーザン法、*in situ*法、ドットプロット法、及びアレイなどを含むハイブリダイゼーション反応またはスクリーニングアッセイで使用するためにレポーター分子で標識化することができる。

【0033】

「タンパク質」とは、ポリペプチド或いはその任意の部分を指す。タンパク質の「部分」とは、少なくとも1つの生物学的活性を保持し得るアミノ酸配列の長さ、PFAMまたはPRINTS分析によって同定されるドメイン、或いはPROTEAN プログラム(DNASTAR, Madison WI)のKyte-Doolittle アルゴリズムを使用して同定されたタンパク質の抗原エピトープを指す。「オリゴペプチド」は、抗体を産生させるために融合タンパク質の一部として使用される約5残基から約15残基までのアミノ酸配列である。

【0034】

「精製された」とは、自然環境から分離され、自然環境で会合していた他の化合物が約60%から約90%まで分離したあらゆる分子や化合物を指す。

【0035】

「サンプル」は、核酸、タンパク質、及び抗体等を含むとして、その最も広い意味で用いられる。サンプルは体液、細胞調製の可溶性分画、細胞が成長する培地のアリコット、染色体、細胞小器官、或いは細胞から単離または抽出された膜、溶液中のまたは基板に固定されたゲノムDNA、RNA、またはcDNAと、細胞、組織、組織プリント、フィンガープリント、口内細胞、皮膚、または髪等を含み得る。

【0036】

「特異的な結合」とは、構造、特に分子側鎖に依存する2つの分子間での特殊な正確な相互作用を指す。例えば、調節タンパク質のDNA分子の主溝への挿入、2つの一本鎖核酸間の骨格に沿った水素結合、またはタンパク質のエピトープとアゴニスト、アンタゴニストまたは抗体との間の結合がある。

【0037】

配列に適用される「類似性」とは、Smith-Waterman アルゴリズム(Smith及びWaterman (1981) J Mol Biol 147:195-197)或いはBLAST2 (Altschul ら (1997) Nucleic Acids Res 25:3389-3402)などの標準化されたアルゴリズムを用いてアラインメントされた少なくとも2つの配列間で一致する分子或いは残基の定量化(通常は%)を指す。BLAST2は、アラインメントを最適化するために配列の1つでギャップを挿入するまた2つの配列をより有意に比較できる標準化された再現性のある方法で使用され得る。

【0038】

「基板」とは、cDNAまたはタンパク質が結合する任意の固体または半固体の支持体を指すものであって、膜、フィルタ、チップ、スライド、ウエハ、ファイバー、磁性または非磁性ビーズ、ゲル、毛管、またはその他のチューブ、プレート、ポリマー、微小粒子が含ま

れ、穴、溝、ピン、チャネル、孔等、様々な表面形態を有する。

【0039】

「変異体」とは、cDNAまたはそのcDNAがコードするタンパク質の認識できる変異した分子を指す。スプライス変異体は、スコアが少なくとも100、最も好ましいのは少なくとも400であるBLASTスコアにより決定し得る。対立遺伝子変異体は、cDNAに対して高い同一性のパーセントを有し、100塩基に付き約3塩基が異なり得る。「一塩基多型」(SNP)とは、欠失、挿入、または置換による単一塩基による変異を指す。この変異は、保存的(プリンからプリン)或いは非保存的(プリンからピリミジン)であり得、コードされたアミノ酸に変異が起こる可能性がある。

【0040】

(発明)

本発明は、ARPをコードするcDNAの発見と、癌の性質決定、診断、治療における、cDNAまたはその断片、およびタンパク質またはその部分をそのまま使用するか組成物として使用することに基づくものである。

【0041】

本発明のARP-1をコードする核酸は、アミノ酸配列のためのコンピュータサーチを用い、ヒト膀胱腫瘍cDNAライブラリ(BLADTUT04)よりインサイトクローン1555118内で最初に同定された。コンセンサス配列、SEQ ID NO:2は、追従するオーバーラッピング及び/又は拡張された核酸配列(SEQ ID NOs:4-11):インサイトクローン1555118H1(BLADTUT04)、722739IH1(BRAXTDR15)、70158486V1(SG0000039)、70162686V1(SG0000040)、70151326V1(SG0000038)、70154198V1(SG0000038)、2084238T6(UTRSNOT08)、70155923V1(SG0000038)、及びGenBank EST g6661750(SEQ ID NO:57)、及びエディットされたGenscan配列GNN.g10801482 edit(SEQ ID NO:58)に由来した。GNN.g10801482 editのために、コーディング領域はゲノムGenscanDNAの解析によって推定された。g10801482は、Genscanが適用される配列のGenBank同定番号である。

【0042】

ある実施例では、本発明は図1A乃至1Kに示すようにSEQ ID NO:1のアミノ酸配列を含むポリヌクレオチドを包含する。ARP-1は長さが935のアミノ酸である。また、ARP-1は、N332、N502、N516、及びN778に存在する4つの潜在的Nグリコシル化部位、S141及びS745に存在する2つの潜在的環式AMP-若しくは環式GMP-依存タンパク質キナーゼリン酸化部位、T196、S242、T334、T445、T521、S562、S635、S676、S745、S822、及びS927に存在する11の潜在的カゼインキナーゼIIリン酸化部位、S207及びS807に存在する2つの潜在的グリコサミノグリカン吸着部位、S162、S190、S225、S242、S352、S375、T442、S455、S512、T569、T6191、S620、S780及びS801に存在する14個の潜在的プロテインキナーゼCリン酸化部位、Y789に存在する一つの潜在的チロシンキナーゼリン酸化部位、及びG555からS562に存在する一つの潜在的ATP/GTP結合部位モチーフA(P-loop)を有する。PFAM解析は、T203からP290まで、K383からQ469まで、E498からR591までのARP-1領域がPDZドメインに類似することを示している。そのようなドメインは、シグナル伝達分子とのタンパク質-タンパク質インターアクションに参与する。図3A、3B、3C、3D、3E、3F、3G、3H、3I、及び3Jで示されているように、ARMはラットASIP(g3868778; SEQ ID NO:62)及びヒトASIP(gS037915; SEQ ID NO:63)と化学的、構造的類似性を有する。とりわけ、ARP-1とラットASIPとは41%の同一性を有し、ARP-1及びヒトASIPは43%の同一性を有する。S708からF925までの領域は、ラットASIP及びヒトASIPの領域に結合するaPCKに類似する。3つのすべてのタンパク質は、3つの保存PDZドメイン及び9つの潜在的なプロテインキナーゼCリン酸化部位を有する。ARP-1の有用な抗原性エピドープは、R192からS222、K413からS455、K759からQ809に広がっている。また、ARP-1の生物学的に活性な部位は、T203からP290、K383からQ469まで、E498からR591まで、及びS708からF925に広がっている。ARP-1へ特異的に結合する抗体は、とりわけ膀胱移行上皮癌のような癌に対するアッセイにおいて有用である。

【0043】

10

20

30

40

50

本発明のARP-2をコードする核酸は、アミノ酸配列のためのコンピュータサーチを用い、ヒト膀胱腫瘍cDNAライブラリ(KIDNTUT13)よりインサイトクローン2582063内で最初に同定された。コンセンサス配列、SEQ ID NO:2は、追従するオーバーラッピング及び/又は拡張された核酸配列(SEQ ID NOs:4-11):インサイトクローン2582063HI (KIDNTUT13)、7246093HI (PROSTMY01)、7978420H1、55040412HI、2929484F6 (TLYMN0T04)、5627320R8 (PLACFER01)、3209128F6 (BLADN0T08)、349248HI (LVENN0T01)、7019961111 (PANCN0N03)、6303175H2 (TLYMUNT02)、2549906F6 (LUNGTUT06)、1945452M (PITUN0T01)、2549906T6 (LUNGTUT06)、71009002V1 (SG0000308)、7100852IVI (SG0000308)、71010168V1 (SG0000308)、170090181VI (SG0000030)、6833928H1 (BRSTN0N02)、70089663V1 (SG0000030)、及びGenBank EST g6993427 (SEQ DD NO:59)、g5529915 (SEQ ID NO:60)、及びg_1733437 (SEQ ID NO:61)に由来した。

10

【0044】

別の実施例では、本発明は図2A乃至2Sに示すようにSEQ ID NO:2のアミノ酸配列を含むポリヌクレオチドを包含する。ARP-2は長さが1356のアミノ酸である。また、ARP-2は、N112、N159、N208、N265、N391、N594、N608、N922、N1144、及びN1231に存在する10個の潜在的Nグリコシル化部位、S144、T468、S685、S715、S888、及びT1176に存在する6つの潜在的環式AMP-若しくは環式GMP-依存タンパク質キナーゼリン酸化部位、S123、S154、S161、T243、S248、T258、S267、S370、T476、S523、T534、T569、T654、T705、S720、S780、S781、S827、S840、T865、S888、T947、S958、T965、T995、T1048、S1049、S1102、S1116、S1397、S1149、S1245、S1252、及びS1308に存在する34個の潜在的カゼインキナーゼリン酸化部位、S166、S187、S292、S372、S379、S428、T453、T476、S533、S604、T661、S742、S829、T847、S924、S955、S958、T988、T995、T1038、S1116、S1143、S1192、S1228、S1252、及びS1312に存在する26個の潜在的プロテインキナーゼCリン酸化部位、Y199、Y388、Y745、Y933、Y1080、及びY1321に存在する6つの潜在的チロシンキナーゼリン酸化部位、そして、G516からS523及びG647からS654に存在する2つの潜在的ATP/GTP結合部位モチーフA (P-loop)を有する。PFAM解析は、K273からA360まで、N461からQ547まで、及びE590からR683までARP-2領域がPDZドメインに類似することを示している。図3A乃至3Jで示されているように、ARP-2はラットASIP (g3868778; SEQ ID NO:62)及びヒトASIP (gS037915; SEQ ID NO:63)と化学的、構造的類似性を有する。とりわけ、ARP-2とラットASIPとは89%の同一性を有し、ARP-1及びヒトASIPは93%の同一性を有する。R712からK936までのARP-2領域は、ラットASIP及びヒトASIPの領域に結合するaPCKに類似する。3つのすべてのタンパク質は、3つの保存PDZドメイン及び潜在的なプロテインキナーゼCリン酸化部位を有する。ARP-2の有用な抗原性エピドープは、K189からQ236、T895からM1023、R1207からN1267に広がっている。また、ARP-2の生物学的に活性な部位は、T273からA360、N461からQ547まで、E590からR683まで、及びS712からF936に広がっている。ARP-2へ特異的に結合する抗体は、とりわけ膀胱移行上皮癌のような癌に対するアッセイにおいて有用である。

20

30

【0045】

表1は、組織カテゴリーにかかるARP-1及びARP-2の発現を示すものである(実施例8にもリストされている)。表2は、膀胱組織、とりわけ移行上皮癌の患者よりのにおけるARPの発現を示す物である。ARPは、移行上皮癌の患者より摘出した膀胱組織よりのライブラリ(BLADTUT04)と、同一患者のマッチした顕微鏡学的に健康な組織よりのライブラリ(BLADNOT05)とを比較して、過発現を示す。表3は、マイクロアレイ解析によって決定されるように、EGFで処理されたヒトBT20乳癌細胞内のARP-2差動発現を示す物である。ARP-2は、4時間から48時間かけてEGFで処理されたBT20細胞中の、処置されていない細胞と比較して軽減された発現を示す。処置されたBT20細胞における最大に軽減された発現は、36時間五に現れた。故に、ARP-1もしくはARP-2をコードするDNAはとりわけ膀胱移行上皮癌のような癌の診断のアッセイに有用である。約1792から1836のヌクレオチド、約1905から1950のヌクレオチド、及び約2180から2230のヌクレオチドのARP-1をコードするcDNAの断片もまた、診断アッセイにおいて有用である。約585から635のヌクレオチド、約1670から

40

50

1710のヌクレオチドのARP-2をコードするcDNAの断片もまた、診断アッセイにおいて有用である。

【0046】

ARPをコードするcDNAの哺乳動物変異体は、デフォルトのパラメータ及びZOOSEQデータベース(Incyte Genomics)と共にBLAST2を用いて同定された。これら好適な変異体は、以下の表に示すように約80%から約100%の同一性を有する。第1列は、ヒトcDNAのためのSEQ ID(SEQ IDH)を示しており、第2列は、変異体cDNAのためのSEQ ID(SEQ IDvar)を示しており、第3列は変異体cDNAのためのクローン番号(Clonevar)を示しており、第4列は、ライブラリの名前を示しており、第5列は、ヒトcDNA(同じケースでヒトcDNAの異なった領域を含む、変異体cDNAの異領域配列を有する)に対する変異体の配列を示しており、そして第6列はヒトcDNAに対する同一性のパーセンテージを示している。

10

【0047】

| SEQ ID _H | SEQ ID _{var} | クローン _{var} | ライブラリ名 | N _H 配列 | 同一性 |
|---------------------|-----------------------|---------------------|------------|-------------------|------|
| 3 | 12 | 702457609T1 | RACONON05 | 2162-2619 | 84% |
| 3 | 13 | 702458746T1 | RACONON05 | 2152-2579 | 84% |
| 3 | 14 | 701335936H1 | RALINON08 | 1322-1566 | 88% |
| 3 | 15 | 700639694H1 | RATONOT01 | 2361-2619 | 87% |
| 3 | 16 | 700639694F6 | RATONOT01 | 2447-2651 | 85% |
| 3 | 17 | 701191467H1 | RACONON05 | 2063-2333 | 81% |
| 3 | 18 | 702771158H1 | CNLIINOT02 | 615-710 | 87% |
| 3 | 19 | 701266650H1 | MOLUDIT08 | 1289-1544 | 91% |
| 20 | 40 | 702231139H1 | RAFANOT02 | 2500-2951 | 88% |
| 20 | 41 | 700273304F6 | RASJNOT01 | 2705-3074 | 85% |
| 20 | 42 | 700330856H1 | RALINON04 | 1550-1817 | 87% |
| 20 | 43 | 700273304H1 | RASJNOT01 | 2705-2951 | 86% |
| 20 | 44 | 701517518H1 | RAKITXT62 | 2149-2441 | 83% |
| 20 | 45 | 701834089T1 | RAKITXT10 | 2897-3074 | 85% |
| 20 | 46 | 701480437H1 | RALITXT42 | 3679-3824 | 88% |
| 20 | 47 | 701190235H1 | RACONON05 | 2077-2333 | 80% |
| 20 | 48 | 700939688H1 | RALINON07 | 2971-3074 | 88% |
| 20 | 49 | 700939688F6 | RALINON07 | 2971-3074 | 88% |
| 20 | 50 | 702582937T1 | RASJUNN01 | 5276-5319 | 95% |
| 20 | 51 | 700299037F6 | RACONOT02 | 5276-5334 | 91% |
| 20 | 52 | 701246488H1 | RALINOH02 | 2284-2357 | 86% |
| 20 | 53 | 702759912H1 | CNLIUNN01 | 1-566 | 92% |
| 20 | 54 | 700112340H1 | MOOSUNR01 | 121-362 | 88% |
| 20 | 55 | 700827810H1 | MOEMNOT01 | 3559-3701 | 87% |
| 20 | 56 | 700109331H1 | MOOSUNR01 | 2509-2569 | 100% |

20

【0048】

これらcDNAは、ヒトの疾患を形成しそのような疾患の潜在的治療をテストする遺伝子組み換え細胞株若しくは生物体を生成するのにとりわけ有用である。

30

【0049】

遺伝的コード悪化の結果として、ARPをコードする複数のcDNA、既知のcDNAに対最小の類似性を有する複数のベアリング(bearing)、及び自然発生遺伝子が生成されても良いことは当業者が認識している。そのようなわけで、本発明は、可能なコドンに基づいたコンディションを選択することによってなされうるcDNAのすべての可能なバリエーションを意図するものである。これらバリエーションは、自然発生したARPをコードするポリヌクレオチドに適用されたような通常型の三重遺伝子コードに関連してなされ、それらすべてのバリエーションは特異的に開示されていると考えられる。

【0050】

cDNA及びその断片自身(SEQ ID NOs:3-56)は、SEQ ID NO: 3、20およびサンプル内の関連分子間で同定及び識別可能する合成、増幅、及びスクリーニング技術において用いられても良い。ほ乳類のcDNAは、遺伝子組み換え細胞株若しくは生物を生成するのに用いられ得る。それらは、ヒトの癌、とりわけ膀胱移行上皮癌のような癌のモデルシステムであり、潜在的な治療の毒性及び効果がテストされても良い。毒物学の学習、臨床試験、及び被験者/患者の処置のプロフィールは、cDNA、タンパク質、抗体、分子および化合物を用いて実行及びモニターされても良い。前記分子及び化合物は、本発明の分子及びcDNAやタンパク質を用いて識別されたものである。

40

【0051】

(本発明の特徴及び使用)

50

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以降に記載する実施例に示す。コンセンサス配列は、Phrap (P. Green, University of Washington, Seattle WA)及びAUTOASSEMBLERアプリケーション(Applied 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、及び/またはショットガン配列は、クラスタ及び/又はマスタークラスタへと電子的に集まる。

10

【0052】

シーケンシング

核酸をシーケンシングする方法は当分野で周知であり、そのような方法を用いて本発明の任意の実施例を実施することができる。これらの方法は、DNAポリメラーゼIであるクレノウフラグメント、SEQUENASE、Taq DNAポリメラーゼおよび熱耐性T7 DNAポリメラーゼ(Amersham Pharmacia Biotech (APB), Piscataway 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、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373DNAシーケンシングシステム(Applied 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ら(1997; Short Protocols in Molecular Biology, John Wiley & Sons, New York NY, unit 7.7)及びMeyers(1995; Molecular Biology and Biotechnology, Wiley VCH, New York NY, pp. 856-853)に記載されている。

20

【0053】

ショットガン・シーケンシングは、目的の特定クローンインサートの配列を完成させるべく用いられ得る。ショットガン理論は、様々なサイズのセグメントに対するオリジナルインサートの破壊と、それら断片のベクターへのクローニングとをランダムに伴う。断片は、オリジナルインサートの全体配列が既知となるまで重複末端を用いて配列及び再構築される。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cDNA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、組立てが未終了の配列(組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、組立てが未終了の配列を完全な配列に組み立てる。

30

【0054】

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(Applied 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(Molecular Biology Insights, Cascade CO)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。

40

【0055】

ハイブリダイゼーション

50

cDNA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(すなわち、タンパク質の保存された触媒ドメインをコードする5'若しくは3'のヌクレオチド)等のユニークな領域から作製可能であり、ARP、対立遺伝子の変種、若しくは関連分子をコードする天然の分子を同定するためのプロトコルに用いることができる。通常は一本鎖であるDNA若しくはRNAからなるこのプローブは、任意のこの核酸配列、配列番号3 - 56と少なくとも50%の配列同一性を有することが望ましい。ハイブリダイゼーションプローブは、レポータ分子の存在下での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cDNA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APBが販売するキットを用いて行うことができる。

10

【0056】

ハイブリダイゼーションのストリンジェンシー(厳密性)は、プローブのGC含量、塩濃度、及び温度によって決まる。特に、塩濃度を下げる、またはハイブリダイゼーションの温度を上げて、ストリンジェンシーを高めることができる。ある膜系のハイブリダイゼーション用の溶液にホルムアミドなどの有機溶媒を加えて、反応が低い温度で起こるようにすることができる。ハイブリダイゼーションは、低いストリンジェンシーの緩衝液(5×SSC、1%のドデシル硫酸ナトリウム(SDS))で、60℃で行うことができるが、核酸配列間に不適正塩基対を含む複合体の形成を許容し得る。続く洗浄は、45℃(中程度のストリンジェンシー)或いは68℃(高いストリンジェンシー)の何れかの温度、0.2×SSC、0.1% SDSなどの高いストリンジェンシーで行う。高いストリンジェンシーでは、ハイブリダイゼーション複合体は、完全に相補的な核酸分子部分のみが安定して保持される。ある膜系のハイブリダイゼーションにおいて、好ましくは35%、最も好ましくは50%のホルムアミドをハイブリダイゼーション溶液に加えて、ハイブリダイゼーションを行う温度を下げたり、または、SarkosylやTriton X-100(Sigma-Aldrich, St. Louis Mo.)などの界面活性剤及び変性したサケ精子DNAなどのブロッ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

【0057】

当分野で周知の方法でマイクロアレイを準備して分析することができる。オリゴヌクレオチドをマイクロアレイのプローブや標的として用いることができる。マイクロアレイを用いて、同時に極めて多数の遺伝子の発現レベルをモニタリングし、遺伝子変異体、突然変異及びSNP(一塩基多型)を同定することができる。このようなデータを用いて、遺伝子機能の解明や、症状及び疾患、または障害における遺伝子原理の解明や、症状及び疾患、障害の診断または治療、治療薬の開発、並びにこれらの治療薬の活性のモニタリングが可能である(例えば、Brennan他(1995) USPN 5,474,796; Schena他(1996) Proc. Natl. Acad. Sci. 93:10614-10619; Baldeschweiler他(1995) PCT出願 W095/251116; Shalon他(1995) PCT出願 W095/35505; Heller他(1997) Proc. Natl. Acad. Sci. 94:2150-2155; and Heller他(1997) USPN 5,605,662を参照)。

40

【0058】

ハイブリダイゼーションプローブはまた、天然のゲノム配列のマッピングに有用である。このプローブは、(1)特定の染色体、(2)染色体の特定の領域、(3)ヒト人工染色体(HACs)や酵母人工染色体(YACs)、細菌人工染色体(BACs)、細菌P1作製物、若しくは単一クロモソームcDNAライブラリのような人工染色体作製物、にハイブリダイズすること可能である。

【0059】

発現

ARPをコードする複数のcDNAの何れか一つ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

50

フリング (Stemmer and Cramer (1996) USPN5,830,721) や部位特異的変異誘発などの方法によって、新規の制限部位を作り出したり、グリコシル化パターンを変えたり、優先コドンを変えて特定の宿主における発現を増大させたり、スプライス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組み換えDNA技術、合成技術及び/またはin vivo遺伝子組み換え技術を組み合わせて、このベクターにcDNA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

【0060】

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

20

【0061】

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クエンシング、ヘルパーファージによる一本鎖の調整、入れ子状欠失の作製において有用である。

【0062】

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

30

40

【0063】

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

【0064】

50

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)、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。精製に続く分離を容易にするために、タンパク質の分割部位をコードする配列は、タンパク質と異異形部分との間のベクターの一部であって良い。組み換えタンパク質の発現及び精製の方法はAusubel(前出、unit 16)に記載され市販されている。

【0065】

ペプチドの化学合成

タンパク質若しくはその一部は、組み換え方法以外の当分野で周知の化学的方法によって合成することもできる。固相技術を用いるペプチド合成は、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形成する。このアミノ酸残基は、酸不安定Boc(t-butylloxycarbonyl)法若しくは塩基不安定Fmoc(9-fluorenylmethoxycarbonyl)法によって保護されたN-である。保護されたアミノ酸のカルボキシル基をリンカーのアミンに結合して、この残基を固相支持レジンに結合させる。Boc若しくはFmocを用いた場合、トリフルオロ酢酸若しくはピペリジンを用いて保護基を除去する。カップリング試薬若しくは予め活性化されたアミノ酸誘導体を用いて、追加する各アミノ酸を結合された残基に付加してから、レジンを洗浄する。完全長のペプチドは、連続的な保護の停止、即ち誘導体化アミノ酸を結合させて合成し、ジクロロメタン及び/またはN,N-ジメチルホルムアミドで洗浄する。このペプチドは、ペプチドカルボキシル末端とリンカーとの間で切断され、ペプチド酸またはペプチドアミドが作られる (Novabiochem 1997/98 Catalog and Peptide Synthesis Handbook, San Diego CA pp. S1-S20)。ABI 431 A ペプチドシンセサイザー (Applied 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)。

【0066】

抗体の準備及びスクリーニング

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

【0067】

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)。

【0068】

10

20

30

40

50

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。タンパク質のエピトープに対して特異的に結合する部位を含む抗体断片を生産することが可能である。限定するものではないが、このような断片には、例えば、抗体分子のペプシン消化によって作製されたF(ab')₂断片及びこのF(ab')₂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）。

【0069】

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

10

【0070】

アッセイのための分子の標識化

多様なレポータ分子及び接合技術が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、³²P-dCTPまたはCy3-dCTP、Cy5-dCTP（Operon Technologies, Alameda CA）などの標識したヌクレオチドや³⁵Sメチオニン（APB）などのアミノ酸を組み込むための市販のキット（Promega, Madison Wis）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。

20

【0071】

（診断）

本cDNA、断片、オリゴヌクレオチド、相補的なRNA及びDNA分子、PNAを用いて、遺伝子発現の変化やmRNAの過剰な発現の不在/存在を検出及び定量、または治療期間中のmRNAレベルのモニタリングを行うことができる。IPと特異的に結合する同様の抗体は、タンパク質を定量化するのに用いられてもよい。発現変動に関連する疾患には、とりわけ膀胱移行上皮癌のような癌が含まれる。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。

30

【0072】

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。患者のサンプルに於けるコンプレクスフォーメーションが通常もしくは疾患標準値の何れかと著しく異なっている（高いまたは低い）場合は、疾患の存在を示唆する。

40

【0073】

差次的な発現を確立するための基準を設けるために、通常及び疾患発現プロファイルが確立される。この発現プロファイルは、ハイブリダイゼーションまたは好適な条件下で、ヒト若しくは動物の正常被験体から採取したサンプルを、cDNAと結合させることによって確立される。標準的なハイブリダイゼーション複合体は、正常な被験体から得た値と、実質的に精製された配列を所定量用いた実験値とを比較することによって定量することができる。このように求めた標準値を、特定の症状や疾患、または異常症を示す患者のサン

50

ルから得た値と比較することができる。標準値と特定の疾患に関連する値との偏差からその症状を診断する。

【0074】

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。

【0075】

免疫学的方法

特異的なポリクローナル抗体若しくはモノクローナル抗体の何れかを用いるタンパク質の検出及び定量は当分野で周知である。このような技術には、ELISA（酵素結合免疫吸着検定法）及びラジオイムノアッセイ（RIA）、蛍光活性化セルソーター法（FACS）が含まれる。2つの非干渉エピトープに反応するモノクローナル抗体を用いる2部位モノクローナル系イムノアッセイが好ましいが、競合的結合アッセイを用いることもできる（例えば、Coligan 他（1997）Current Protocols in Immunology, Wiley-Interscience, New York NY; および Pound 前出）。

【0076】

（治療）

図3A乃至3Jに示されているように、PDZドメイン及びaPKC結合領域に関連して、ARP-1（1555118; SEQ ID NO: 1）、ARP-2（2582063; SEQ ID NO:2）、ラットASIP（g3868778; SEQ ID NO:62）、及びヒトASIP（g8037915; SEQ ID NO:63）の領域間に化学的及び構造的類似性が存在する。加えて、ARPの差次的発現は、表1及び2に示されているように、消化器系、女性生殖組織、血液及び免疫系、神経系、呼吸器系、及び尿路、及びとりわけ膀胱移行上皮癌のような癌に関連する。ARPはとりわけ膀胱移行上皮癌のような癌に際だった役割を果たす。

【0077】

ARPの増加発現に関する状況の治療では、発現若しくはタンパク質活性を減少させることは不可能である。ある実施例では、タンパク質の抑制遺伝子、アンタゴニスト、及び抗体が被験者に対して投与され、増加した発現及び活性に関するコンディションを治療する。別の実施例では薬剤キャリアに関連する抑制遺伝子、アンタゴニスト、及び抗体を含む薬剤成分が、被験者に対して投与され、内因性タンパク質の増加発現及び活性に関するコンディションを治療する。追加的な実施例では、cDNA若しくはその断片の補体を発現するベクターが、疾患の治療のために被験者に投与される。

【0078】

ARPの減少した発現に関連する症状の治療では、発現若しくはタンパク質活性を増加させることが出来ない。ある実施例では、タンパク質、アゴニスト、若しくはエンハンサが、被験者に投与され、減少した発現若しくは活性に関する症状を治療する。別の実施例では、薬剤のキャリアに関連して、タンパク質、アゴニスト、及びエンハンサを有する薬剤成分が被験者に対して投与され、内因性タンパク質の減少発現若しくは活性に係る症状を治療する。さらなる実施例では、cDNAを発現するベクターが、疾患を治療する目的で被験者に投与される。

【0079】

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

【0080】

10

20

30

40

50

核酸を用いる遺伝子発現の調節

遺伝子の発現は、ARP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、複数のcDNAまたはその断片のライブラリをスクリーニングして、翻訳されない調節配列に特異的に結合する核酸分子または断片を同定することも可能である。

10

【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およびGUU、GUCなどの部位においてヌクレオチド鎖が切断される。このような部位が一旦同定されたら、オリゴヌクレオチドを機能不全にしう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。

【0082】

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

20

【0083】

スクリーニング及び精製のアッセイ

本ARPをコードするcDNAを用いて、分子若しくは混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事が可能である。このライブラリは、生物系においてcDNAの活性、複製、転写、または翻訳を調節するアダプター、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、およびその他のリガンドを含み得る。このアッセイは、特異的な結合が許容される条件下で、cDNA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、一本鎖または適切ならば二本鎖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。

30

【0084】

一実施例において、本発明のcDNAを複数の精製された分子または複合物と共にインキュベートすることが可能であり、電気泳動度移動アッセイ(USPN 6,010,849)または網状赤血球溶解物転写アッセイ等の当分野で公知である方法によって結合活性を測定することが可能である。別の実施例において、cDNAは生検そして/あるいは培養細胞と組織からの核抽出物と共にインキュベートすることが可能である。核抽出物のcDNAと分子あるいは化合物の特異結合は、最初ゲルシフトアッセイにより決定される。そして後にその分子または化合物を回収して、それに対する抗体を作製することによって確認することが可能である。これらの抗体がアッセイに加えられる時、電気泳動度移動アッセイにスーパーシフトが起きる。

40

【0085】

50

別の実施例では、cDNAは、当分野で公知であるアフィニティークロマトグラフィー方法を用いて分子あるいは化合物を精製するために使用される。一実施例では、cDNAは高分子樹脂またはゲル上で臭化シアン基と化学的に反応する。次にサンプルを通し、cDNAと反応あるいは結合させる。cDNAに結合する分子または化合物は、流動媒体の塩の濃度の上昇によりcDNAから放出され、回収される。

【0086】

更なる実施例において、タンパク質あるいはその部分をサンプルのリガンドを精製するために用いる可能性がある。リガンドを精製するためにタンパク質またはその部分を使用する方法には、特異結合を許容する条件下でタンパク質またはその部分をサンプルと結合すること、タンパク質とリガンド間の特異結合を検出すること、結合したタンパク質を回収すること、精製したリガンドからタンパク質を分離するために適切なカオトロピック剤を使用することが含まれる。

10

【0087】

好適な実施例では、ARPまたはその部分を多様なスクリーニングアッセイの任意の複数の分子または化合物をスクリーニングするために用い得る。そのようなスクリーニングで用いられたタンパク質の部分は、溶液中で遊離しているか、非生物的物质または生物的基質に固定させる（例えば細胞表面上に保持される）、あるいは細胞内に位置することになる。例えば1つの方法では、組換え核酸で安定的に形質転換された、生きたまたは固定された原核宿主細胞で、細胞表面にペプチドを発現し局在させる細胞をスクリーニングアッセイで使用することが可能である。細胞を複数のリガンドあるいはリガンドのライブラリに対してスクリーニングする。そして発現したタンパク質とリガンドの間で結合の特異性または複合体の形成を測定し得る。タンパク質と分子間の特異結合を測定してもよい。スクリーニングされたライブラリの種類に応じて、アッセイを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、薬剤及び他のリガンドを同定するために用いる可能性があり、そしてタンパク質を特異結合する。

20

【0088】

或る実施態様において、本発明はUSPN 5,876,946に記載された非常に小さなアッセイ量と微量を使用して試験化合物高い処理能力でスクリーニングするための方法を網羅する。当該文献は特に引用することを以って本明細書の一部となす。この方法は、特異結合により多数の分子と化合物をスクリーニングするために用いられる。別の実施態様において本発明は、タンパク質を結合することができる中和抗体が、タンパク質またはオリゴペプチドまたはその部分を結合するための試験化合物と特異的に競合する、競合的薬剤スクリーニングアッセイの使用も網羅する。スクリーニングによって同定された分子または化合物は、毒性、診断または可能性のある治療薬を評価するために哺乳動物モデル系で用いられ得る。

30

【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、好適な濃度範囲および投与経路を決定する。次に、このような情報を用いて、ヒトへの効果的な投与経路および薬用量を決定する。

40

【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および毒性は、例えば、ED₅₀（集団の50%に医薬的效果がある投与量）および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。

50

高い治療指数を示す医薬組成物が好ましい。細胞培養アッセイおよび動物実験から得られたデータを用いて、ヒトへ適用する薬用量の範囲を決定する。

【0091】

(モデル系)

動物モデル系は生物学的検定法として用い得る。そこではヒトと類似の表現型反応を示し、また暴露条件が人体暴露と関連する哺乳動物は最も一般的なモデルであり、多くの感染体、癌、薬剤および毒性研究が、低費用、利便性、寿命、生殖可能性、豊富な参照文献ゆえにラットまたはマウス等の齧歯類で実施される。近交系また非近交系齧歯類は、目的遺伝子の過小発現と過剰発現の生理作用の結果の研究、および疾患の診断と治療のための方法開発のための便利なモデルを提供する。特定の遺伝子を過剰に発現する哺乳動物近交系（例えば乳汁内に分泌する）は、その遺伝子により発現する便利なタンパク質源となり得る。

10

【0092】

毒性研究

毒性研究は、生物系上での薬剤影響の研究である。多くの毒性研究は、ラットまたはマウスで実施される。生理学、行動、恒常性プロセスにおける定性または定量変更そしてラットまたはマウスの死亡率の観察は、毒性プロファイルを生み出し、薬剤への暴露の後にヒトの健康に及ぼす潜在的な因果関係を評価するために使われる。

【0093】

遺伝毒性研究により、内在性、自発的そして誘導された遺伝突然変異の率への薬剤の影響の同定と分析がなされる。遺伝毒性薬剤は、核酸との相互作用を促進する共通の化学的特性また物理的特性を通常は有する。また染色体の異常が子孫に伝達された時最も有害である。毒性研究により、受胎前に親に、妊娠中に母親にあるいは発達中の生物のいずれかに投与された場合子孫の組織中の構造的また機能的異常の頻度が増大する物質を同定することが可能である。マウスまたはラットは、これらの試験で最も頻繁に使用される。なぜなら生殖周期が短いので統計上の必要にかなう多数の生物の繁殖を可能にするからである。

20

【0094】

急性毒性試験は、症状改善または薬剤の致死率を決定するために被験動物に薬剤を1回投与することに基づいている。3回の試験が実施される。：1) 初回投与量範囲を定める実験、2) 有効量の範囲を狭める実験、3) 用量反応曲線を確定するための最終実験。

30

【0095】

亜慢性毒性試験は、薬剤の連続投与に基づく。ラットとイヌはこれらの研究で一般的に用いられ、様々な属の種からのデータを提供する。発癌を除いて、3 - 4ヶ月の間、高投与量濃度で薬剤を毎日投与することにより、成獣における毒性の多くの形態を明らかにできることにに関して相当の証拠がある。

1年以上の期間の慢性毒性試験により、毒性の不存在か薬剤による発癌の可能性を実証する。ラットによって研究が実行される時、最小限の3つの試験グループに加えて1つの対照グループが用いられる。実験当初および実験中は定期的に動物は調べられ、モニターされる。

40

【0096】

遺伝子組換え動物モデル

目的の遺伝子を過剰発現あるいは過小発現する遺伝子組換え齧歯類は近親交配され、ヒト疾患のモデル、または治療薬や毒性薬剤の試験のために用いられ得る。(USPN 5,175,383、USPN 5,767,337等を参照.)。場合によっては、導入された遺伝子は胎児の発育中または出生後の発育中の特定の時間に特定の組織のタイプで活性化され得る。導入遺伝子の発現は、薬物治療法実験への取り組み前、中、後に遺伝子組換え動物の表現型、組織特異的mRNA発現または血清と組織タンパク質レベルの分析によってモニターされる。

【0097】

胚性幹細胞

齧歯類の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保持する

50

。ES細胞がキャリアーの胚内部に置かれた時、正常な発育を再開して、生きて生まれた動物の組織に寄与する。ES細胞は実験的ノックアウトとノックイン齧歯類の生成に用いられる好適な細胞である。129/SvJ細胞株等のマウスES細胞は初期のマウス胚に由来し、当分野で公知の培養条件下で増殖させることができる。遺伝子組換え類を産生するのに用いられたベクターには、疾患遺伝子候補と標識遺伝子が含まれる。後者は導入された疾患遺伝子の存在を同定することに役立つ。ベクターは当分野で公知の方法でES細胞に形質転換する。そして形質転換したES細胞を同定し、例えばC57BL/6マウス系等から採取したマウス細胞胚盤胞に微量注入する。胚盤胞を偽妊娠メスに外科的に導入し、得られるキメラ子孫の遺伝形質を決め、これを交配させてヘテロ接合性系またはホモ接合性系を作製する。

【0098】

10

ヒト胚盤胞由来のES細胞は、in vitro で少なくとも8つの別々の細胞系統に分化するよう操作することが可能である。これらの系統はin vitro で様々な細胞のタイプと組織の分化を研究するのに用いられ、例えば神経細胞、造血系統及び心筋細胞に分化する内胚葉、中胚葉及び外胚葉の細胞の種類を含む。

【0099】

ノックアウト分析

遺伝子ノックアウト分析では、哺乳動物遺伝子の領域はネオマイシンホスホトランスフェラーゼ遺伝子(neo: Capecchi(1989) Science 244:1288-1292)等の非哺乳動物遺伝子を含むよう酵素処理で修飾される。修飾された遺伝子は、培養したES細胞に形質転換され、相同組換えにより内因性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と翻訳を破壊する。形質転換細胞を齧歯類胚に注入し、その胚を偽妊娠メスに移植する。遺伝子組換え子孫は、哺乳動物遺伝子の機能複製を欠くホモ接合性近交系を得るために交配される。一例では、哺乳動物遺伝子はヒト遺伝子である。

20

【0100】

ノックイン分析

ES細胞を用いて、ヒト疾患の「ノックイン」ヒト化動物(ブタ)または遺伝子組換え動物モデル(マウスまたはラット)を作製することが可能である。ノックイン技術を用いて、ヒト遺伝子の或る領域を動物ES細胞に注入し、注入したヒト配列を動物細胞ゲノムに組み込ませる。形質転換細胞を胚に注入し、胚を上記のように移植する。遺伝子組換え子孫または近交系について研究し、可能性のある医薬品を用いて処理し、類似のヒトの症状の治療に関する情報を得る。これらの方法は、幾つかのヒト疾患をモデル化するために用いられてきた。

30

【0101】

ヒト以外の霊長類モデル

動物実験の分野では、生理学、遺伝学、化学、薬理学、統計学等の基礎科学のデータと方法論を扱う。これらのデータは、ヒトの健康に関連している可能性があるため、ヒト以外の霊長類での治療薬の影響を評価するのに非常に重要である。ワクチンと薬物の評価においてサルがヒトの代用に用いられる。またサルの反応が類似の条件下で人体暴露と関連される。このような調査では、カニクイザルとアカゲザル(それぞれ Macaca fascicularis と Macaca mulatta)、コモンマモセット(Callithrix jacchus)が最も一般的に用いられるヒト以外の霊長類(NHP)である。NHPのコロニーの発達と維持には多額の費用がかかるため、齧歯類のモデルで初期の調査と毒性研究が通常は実施される。薬物常用などの行動測定を利用する研究では、NHPが最初に実験動物に選ばれる。さらに、NHPと個々のヒトは多くの薬剤と毒素に異なる感受性を示す。またこれらの薬剤の「広範囲のメタボライザー」から「代謝不良体質」まで表現型の範囲として分類することが可能性である。

40

【0102】

別の実施例では、将来に開発される分子生物学技術で、現在知られているヌクレオチド配列の特性(限定はされないが、トリプレット遺伝コード、特異的な塩基対相互作用等を含む)に依存しているならば、このタンパク質をコードするcDNAにその新技術を用い得る。

【実施例】

50

【0103】

下記の実施例は、目的の本発明を説明するために提供するが、これは本発明を限定する目的ではない。実施例の目的のために、ヒト膀胱腫瘍 (BLADTUT04)、及び標準化された乳房 (BRSTNON02) ライブラリの準備について記載する。

【0104】

1 cDNAライブラリの作製

膀胱腫瘍ライブラリ

BLADTUT04 cDNAライブラリは、左の膀胱壁に移行上皮癌を有する60歳の白人男性より摘出した膀胱腫瘍組織より作成された。POLYTRONホモジナイザー (Brinkmann Instruments, Westbury NJ) を使用して、凍結組織をホモジナイズして溶解させた。試薬及び抽出手順は、RNA単離キット (Stratagene) 内で供給されるように用いられた。可溶化液は、室温、25000rpm、18時間に渡ってL8-70M超遠心機 (Beckman Coulter, Fullerton CA) 中でSW28ローターを用いることによって、5.7M CsClクッションで遠心分離される。RNAはpH 8.0のフェノールクロロホルムで二回、pH 4.7の酸フェノールで一回抽出され、0.3 Mの酢酸ナトリウム及び2.5倍のエタノールを用いて沈殿され、水中で再懸濁され、また37で15分間に渡りDnase処理される。OLIGOTEX精製キット (QIAGEN, Chatsworth CA) を用いて、RNAが単離され、cDNAライブラリの構築に用いられた。

10

【0105】

これをcDNAライブラリの作製に用いた。このmRNAを、mRNAのポリ(A)尾部で一本鎖cDNAの合成を開始するように設計されたNotIプライマー-アダプターを含むSUPERSCRIPIT プラスミドシステム (Life Technologies) の推奨プロトコルに従って処理した。二本鎖cDNAを平滑化し、EcoRIアダプターに結合し、NotI (New England Biolabs, Beverly MA) で消化した。このcDNAを、SEPHAROSE CL4Bカラム (APB) 上で分画化し、400bpを越える大きさのcDNAを、pINCYプラスミド (Incyte Genomics) に結合させた。このプラスミドpINCYは、DH5 コンピテント細胞 (Life Technologies) に形質転換された。

20

【0106】

標準化乳房

E. coli ストレインDH12Sコンピテント細胞内における、プールされたBRSTNOT34及びBRSTNOT35プラスミドライブラリの約 1.2×10^6 の独立クローンは、エレクトロポレーションによる形質転換に続くカルベニシリン (25 mg/1) 及びメチシリン (1 mg/nil) 選択条件下の液体培養液中で成長した。ライブラリ中の存在レベルに従って過剰なcDNA コピーの数を減少させるべく、cDNAライブラリは、Soaresら (1994; Proc Natl Acad Sci 91:9228-9232)、及びBonaldoらの手順に従って2ラウンドで標準化された。プライマエクステンション反応におけるテンプレート比に対するプライマーは、2:1から300:1へと増加した。アニーリングハイブリダイゼーションは13時間から48時間へと延長された。標準化ライブラリの単鎖DNA環はハイドロキシアパタイトクロマトグラフィーによって精製され、ランダムな初回抗原刺激によって部分的に2重鎖へと変換され、pINCYプラスミド内へと結合され、DH12Sコンピテント細胞へと電気穿孔された (Life Technologies)。

30

【0107】

2 pINCYプラスミドの作成

プラスミドをEcoRI制限酵素 (New England Biolabs, Beverly MA) と共にpSPORT1プラスミド (Life Technologies) を消化して、クレノウ酵素 (New England Biolabs) と2'-デオキシヌクレオチド5'-3リン酸 (dNTP) を用いて張出部 (overhang) 末端を占めるとにより作製した。プラスミドは自己連結反応して、細菌性宿主の大腸菌株JM109に形質転換した。

40

【0108】

バクテリア (pSPORT 1 RI) により産出した仲介プラスミドはEcoRIでは消化されず、Hind III (New England Biolabs) で消化され、張出部 (overhang) 末端は再びクレノウとdNTPで占めた。リンカー配列をリン酸化して、5'平滑末端上に連結反応させ、EcoRIで消化してから、自己連結反応させた。JM109宿主細胞への形質転換後に、プラスミドを単離してEcoRIとの優先的消化性を検査したが、Hind IIIとの優先的消化性を検査しなかった。この

50

基準にかなう単一のコロニーを、pINCYプラスミドと名付けた。

【0109】

NotI、EcoRI 制限酵素を用いて準備したライブラリのcDNAを組み込む能力に関してプラスミドを試験した後に、幾つかのクローンを配列化して、多量のプラスミドを準備するための中から約0.8 kbの挿入を含む単一クローンを選択した。ライブラリ作製で使用するために、NotI、EcoRIで消化した後に、アガロースゲル上でプラスミドを単離して、QIAQUICKカラム(Qiagen)を用いて精製した。

【0110】

3 cDNAクローンの単離とシーケンシング

プラスミドDNAを細胞から放出して、MINIPREPキット (Edge Biosystems, Gaithersburg M D)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、960精製のためリガンドでブロックする96-ウェルから成る。推奨されているプロトコルを下記の変更を除いて使用する。1)細菌を、カルベニシリン25 mg/lと0.4%のグリセロールの入った、殺菌したTERRIFIC BROTH(BD Biosciences, Sparks MD)1 mlで培養した。2) 接種後、細胞を19時間培養し、次いで0.3 mlの溶解バッファで溶解した。3) イソプロパノールによる沈殿の後、プラスミドDNA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後に、サンプルを96-ウェルブロックに移して、4 で保管した。

【0111】

cDNAをDNA ENGINE サーマルサイクラー(MJ Research)と併用してMICROLAB 2200システム(Hamilton)を使用して、シーケンシングのために調製した。cDNAをABI PRISM 377シーケンシングシステム(Applied Biosystems)またはMEGABACE 1000 DNAシーケンシングシステム(APB)を用いてSangerとCoulson(1975; J Mol Biol 94:441-448)の方法で配列化した。多くの単離は、溶液体積0.25x-1.0x濃度の標準ABIプロトコルとキット(Applied Biosystems)に従ってシーケンスされた。別法では、cDNAをAPBの溶液と色素を用いてシーケンスした。

【0112】

4 cDNA配列の伸長

cDNAクローンとオリゴヌクレオチドプライマーを用いて、cDNAを伸長した。一方のプライマーは既知の断片の5'伸長を開始するべく合成し、他方のプライマーは既知の断片の3'伸長を開始するべく合成した。開始プライマーは、OLIGOソフトウェア(Molecular Biology Insights)を用いて、約22個から約30個のヌクレオチドの長さで約50%以上のGC含量を有し、かつ約68~72の温度で標的配列にアニールするように設計した。ヘアピン構造及びプライマー-プライマー二量体が生じないようにヌクレオチドを伸長した。

【0113】

配列を伸長するために、鋳型として選択されたcDNAライブラリを用いた。2段階以上の伸長が必要な場合には、付加的プライマー或いはプライマーのネステッドセットを設計した。好適なライブラリのサイズを選択して、更に遺伝子の5'または上流領域を有する配列を含めるために大きなcDNAとランダムプライマーを含めた。ゲノムライブラリを特に5'プロモーター結合領域への伸長する調節領域を得るために使用する。

【0114】

高忠実度の増幅をUSPN 5,932,451で開示された方法を利用したPCR法によって得た。PCR法をDNA ENGINE サーマルサイクラー(MJ Research.)を用いて96穴プレート内で実施した。反応混合液は、鋳型DNA及び200 nmolの各プライマー、 Mg^{2+} と $(NH_4)_2SO_4$ とメルカプトエタノールを含む反応バッファ、Taq DNAポリメラーゼ(APB)、ELONGASE酵素(Life Technologies)、Pfu DNAポリメラーゼ(Stratagene)を含む。プライマーの組、PCI AとPCI B(Incyte Genomics)に対して以下のパラメータで増幅を行った。:ステップ 1:94で3分間、ステップ 2:94Cで15秒間、ステップ 3:60Cで1分間、ステップ 4:68C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 6:68C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で増幅を行った。ステップ 1:94 で3分間、ステップ 2:94Cで15秒間、ステップ 3:57Cで1分間、ステップ 4:68C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 6:68Cで5分間、ステップ 7:4 で保管。

【0115】

各ウェルのDNA濃度は、1X TE及び0.5µlの希釈していないPCR産物に溶解した100µlのPICOGREEN定量試薬(1x TEの0.25%試薬、v/v; Molecular Probes)を不透明な蛍光光度計プレート(Coming Costar, Acton MA)の各ウェルに分配してDNAが試薬と結合できるようにして測定する。サンプルの蛍光を計測してDNAの濃度を定量するべくプレートをFluoroskan II (Labsystems Oy)でスキャンした。反応混合物のアリコート5~10µlを1%アガロースミニゲル上で電気泳動法によって解析し、どの反応が配列の伸長に成功したかを決定した。

10

【0116】

伸長させたクローンは、脱塩及び濃縮して384穴プレートに移し、CviJIコレラウイルスエンドヌクレアーゼ(Molecular Biology Research, Madison WI)を用いて消化し、pUC18ベクター(APB)への再連結反応前に音波処理またはせん断した。ショットガン・シー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(0.6~0.8%)のアガロースゲル上で分離し、断片を切除し、寒天をAGARACE酵素(Promega)で消化した。伸長させたクローンをT4DNA(New England Biolabs)を用いてpUC18ベクター(APB)に再連結し、Pfu DNAポリメラーゼ(Stratagene)で処理して制限部位の張出部(overhang)を満たし、大腸菌細胞に形質移入した。形質移入した細胞を選択して抗生物質を含む培地に移し、それぞれのコロニーを切りとってLB/2Xカルベニシリン培養液の384ウェルプレートに37 で一晚培養した。

20

【0117】

細胞を溶解して、Taq DNAポリメラーゼ(APB)及びPfu DNAポリメラーゼ(Stratagene)を用いて以下の手順でDNAをPCR増幅した。ステップ 1:94 で3分間、ステップ 2:94Cで15秒間、ステップ 3:60Cで1分間、ステップ 4:72Cで2分間、ステップ 5:ステップ 2、3及び4を29回繰り返す。ステップ 6:72Cで5分間、ステップ 7:4 で保管。上記したようにPICOGREEN量的試薬(Molecular Probes)でDNAを定量化した。DNAの回収率が低いサンプルは、上記と同一の条件を用いて再増幅した。サンプルは20%ジメチルスルホキシド(DMSO; 1:2, v/v)で希釈し、DYENAMIC エネルギートランスファー シーケンシングプライマー、及びDYENAMIC DIRECTサイクルシーケンシングキット(APB)またはABI PRISM BIGDYE ターミネーターサイクル シーケンシング反応キット(Applied Biosystems)を用いてシーケンシングした。

30

【0118】

5 cDNAクローンの相同性検索と推定タンパク質

配列表のcDNAまたは推定アミノ酸配列を用いて、GenBank、SwissProt、BLOCKSなどのデータベースに問合せした。BLASTかBLAST2(前出のAltschulら;前出のAltschul)を用いて前もって同定した、またアノテーションが付けられた配列あるいはドメインを含むこれらのデータベースを検索して、アラインメントを生成し、どの配列が厳密に一致するか、または相同的かを決定した。アラインメントを原核(細菌性)または真核(動物、カビあるいは植物)生物のシーケンシングした。別法として、SmithとSmith(1992, Protein Engineering 5:35-51)で説明されているようなアルゴリズムを用いて、一次配列パターンと二次構造ギャップペナルティで処理することが可能であった。この出願で開示されているすべての配列は、少なくとも49ヌクレオチドの長さを有し、不要な塩基は1.2%に過ぎない(A、C、GまたはTよりもNが記録されている)。

40

【0119】

前出のKarlinで詳説されているように、問合せ配列とデータベース配列の間のBLAST一致を統計的に評価して、ヌクレオチドに 10^{-25} 、ペプチドに 10^{-25} の閾値を満たす時のみ報告した。下記のように計算して積スコアによって、相同性も評価した。BLASTのヌクレオチ

50

ドあるいはアミノ酸の一致率[問い合わせ配列と参照配列間]に%最大限BLASTスコアを乗じ[問い合わせ配列と参照配列の長さに基づく]、次いで100で除する。実験室で用いるハイブリダイゼーション法と比較して、厳密な一致の電子ストリンジェンシーを約40の下限(不要な塩基のために1-2%のエラーがある)から70の100%一致まで設定した。

【0120】

無料で入手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>)には、核酸分子をアラインメントする「blastn」、核酸分子またはアミノ酸分子を対で直接比較するために用いるBLAST 2など多種の配列分析プログラムが含まれる。BLASTプログラムは、一般的には、ギャップ及びデフォルト設定に設定された他のパラメータと共に用いる。例えば: 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; 及び Filter: on同一性は配列全体またはより小さいその部分の長さにより測定する。Brenner ら (1998; Proc Natl Acad Sci 95:6073-6078, 文献は、引用することをもって本明細書の一部とする)は配列同一性によって構造的相同性を同定する能力に関してBLASTを分析した。少なくとも150残基の配列アラインメントには信頼できる閾値は30%の同一性であり、少なくとも70残基のアラインメントに関しては40%であることがわかった。

【0121】

推定上ASIP関連タンパク質は、公共のゲノム配列データベース(例えば、gbpriやgbhtg)においてGenscan遺伝子同定プログラムを実行して初めに同定された。Genscanは、様々な生物に由来するゲノムDNA配列を分析するための汎用遺伝子同定プログラムである(Burge, C. および S. Karlin (1997) J. Mol. Biol. 268 : 78-94, Burge, C. および S. Karlin (1998) Curr. Opin. Struct. Biol. 8 : 346-354を参照)。このプログラムは推定エキソンを連結して、メチオニンから停止コドンまで伸長した組み立てcDNA配列を構築する。Genscanにより得られる配列は、FASTAデータベースのポリヌクレオチド配列およびポリペプチド配列になる。Genscanによって一回で解析できる配列の最大長さは30 kbに設定されている。これらのGenscan推定cDNA配列の内、どの配列がASIP関連タンパク質をコードするかを決定するために、コードされたポリペプチドをPFAMモデルにおいてASIP関連タンパク質について問合せた分析した。潜在的なASIP関連タンパク質が、ASIP関連タンパク質としてアノテーションが付けられたインサイトcDNA配列に対する相同性を基に同定された。次に、これらの選択されたGenscan推定配列を、BLAST解析を用いてgeneptおよびgbpri公共データベースの配列と比較した。必要に応じて、Genscan推定cDNA配列を、geneptにおいてBLASTで最もヒットした配列と比較して、Genscan推定配列における余分なエキソンや省いてしまったエキソンなどのエラーを修正し、編集した。BLAST解析を用いてGenscan推定cDNA配列を含むインサイトcDNAまたは公共のcDNAを見つけ出すことにより、転写の証H拠が得られる。インサイトcDNAがGenscan推定cDNA配列を含む場合、この情報を用いてGenscan推定配列を修正或いは確認できる。完全長ポリヌクレオチド配列は、実施例3に説明した組み立て方法でGenscan推定コード配列とインサイトcDNAおよび/または公共のcDNA配列を組み立てて作製した。別法では、完全長ポリヌクレオチド配列は、その全てが編集した或いは未編集のGenscan推定コード配列から作製した。

【0122】

この出願のcDNAをアセンブルしたコンセンサス配列あるいはLIFESEQ GOLDデータベースで見つけた鋳型と比較した。cDNA、伸長、全長およびショットガン・シーケンシングプロジェクトの構成エレメント配列をPHRED分析にかけて、クオリティスコアを指定した。クオリティスコアが許容できるすべての配列を多様なプリプロセッシングとエディティングの経路にかけ、低質の3'末端、ベクターとリンカー配列、ポリAテイル、Alu繰返し、ミトコンドリア及びリボゾーム配列、細菌汚染配列を除去する。編集した配列は少なくとも長さが50 bpであり、ジヌクレオチド反復、Alu反復などの情報に乏しい配列と反復エレメントを"Ns"あるいはマスクしたものに置換した。

【0123】

10

20

30

40

50

編集した配列は、構築処理にかけられ、配列が遺伝子ピンに割り当てられた。各配列は1つのピンにのみ所属でき、各ピンの配列を構築して、鋳型を作製した。BLASTとCROSSMATCHを用いて、新たにシーケンスした部分を既存のピンに加えた。前記の部分配列をピンに加えるために、150以上のBLAST質のスコアと少なくとも82%の局所同一性のアラインメントを有しなければならない。PHRAPを用いて、各ピンの配列を構築した。DEEP PHRAPを用いて、幾つかの重畳部分配列を有するピンを構築した。部分配列の数と配向に基づいて、各鋳型の配向を決定した。

【0124】

ピンを互いに比較して、少なくとも82%の局所類似性を有するピンを結合して、再構築した。95%以下の局所同一性である鋳型を有するピンを分離した。STITCHER/EXON MAPPERアルゴリズムによって鋳型を分析にかけ、スプライス変異体、或いはスプライスエキソン、スプライス接合部、組織のタイプや疾患状態全体の選択的スプライシングされた遺伝子の示差発現などの存在の確率を分析した。構築処理を周期的に繰り返した。そしてBLASTを用いてGbpriなどのGenBankデータベースに対して鋳型にアノテーションを付けた。厳密な一致について、200の塩基対に対して95%の局所同一性から100の塩基対に対して100%の局所同一性を有すると定義した。また相同一致については、E-value (または確率値) $< 1 \times 10^{-8}$ を有すると定義した。鋳型をGENPEPTに対するフレームシフト型FASTxにもかけた。相同一致をE-value $< 1 \times 10^{-8}$ を有すると定義した。鋳型分析と構築については1999年3月25日に提出したUSSN 09/276,534に記載されている。

10

【0125】

構築に続いて、鋳型をBLAST、モチーフその他の機能的分析にかけて、1997年3月6日に提出したUSSN 08/812,290、USSN 08/811,758、1997年10月9日に提出したUSSN 08/947,845、1998年3月4日に提出したUSSN 09/034,807等に記載されているタンパク質階層に分類した。3つの全フォワードリーディングフレームで各鋳型を翻訳することにより、また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(HMM)を基にした保存されたタンパク質ファミリードメインのPFAMデータベースに対する各翻訳を検索することにより鋳型を分析した。MACDNASIS PROソフトウェア(Hitachi Software Engineering)とLASERGENEソフトウェア(DNASTAR)を用いてcDNAを更に分析した。そして公共のデータベース、例えばGenBankのげっ歯類、哺乳動物、脊椎動物、原核生物、真核生物のデータベースと、SwissProt、BLOCKS、PRINTS、PFAMそしてPrositeに問い合わせた。

20

30

【0126】

6 染色体マッピング

スタンフォード・ヒトゲノムセンター(SHGC)、ホワイトヘッド・ゲノム研究所(WIGR)、Genethon等の公的な情報源から入手可能な放射線ハイブリッド及び遺伝地図データを用いて、配列表に存在するcDNAが前もってマッピングされたかを測定した。マッピングされたKSRCCをコードするcDNAの断片は、すべての関連する調節配列とコード配列を同じ位置に割り当てる。遺伝地図上の位置は、ヒト染色体の範囲または間隔として表される。cM間隔の地図上の位置は、染色体のpアームの末端に関連して測定する(ヒト中のDNAの1メガベース(Mb)にほぼ等しい)。

40

【0127】

7 ハイブリダイゼーション技術と分析

基質上でのcDNAの固定

下記の方法の1つによってcDNAを基質に適用する。ゲル電気泳動法によりcDNAの混合を分画し、キャピラリー転移によりナイロン膜に転移する。別法として、cDNAを個々にベクターに連結反応させ、細菌性宿主細胞に挿入してライブラリを形成する。次に下記の方法の1つによってcDNAを基質に配置する。最初の方法では、個々のコロニーを含む細菌細胞をロボット利用して切りとり、ナイロン膜に配置する。膜を選択培地(用いられたベクターに依存してカルペニシリン、カナマイシン、アンピシリンまたはクロラムフェニコール)を含むLB寒天に置き、37°Cで16時間インキュベートする。寒天から膜を取り除き、コ

50

ロニーを上側にして、10% SDS、変性溶液(1.5 M NaCl, 0.5 M NaOH)、中和溶液(1.5 M NaCl, 1 M Tris, pH 8.0)、2xSSC(2回)に各10分、連続的に膜を置く。次に膜をSTRATALINKER UV架橋剤(Stratagene)でUV照射する。

【0128】

2番目の方法では、インサートに隣接するベクター配列に相補的なプライマーを用いて、cDNAを30サイクルのPCRで細菌ベクターから増幅する。PCR増幅で1~2 ngの初期量の核酸から5 µgより大きい最終量まで増大する。SEPHACRYL-400 ビーズ(APB)を用いて長さが400 bpから約5000 bpまで増幅した核酸を精製する。手であるいはドット/スロットプロット法マニホールドと吸気装置を用いて精製核酸をナイロン膜に置く。そして上述した変性、中和、UV照射により固定する。USPN 5,807,522に記載されている方法を用いて、精製した核酸を、ロボットで配置して、ポリマーコートされたスライドガラス上に固定する。顕微鏡スライドガラス(Corning, Acton MA)を良く洗って0.1%のSDS中で超音波をかけ、4%フッ化水素酸(VWR Scientific Products, West Chester PA)中でエッチングし、95%エタノール中で0.05%アミノプロピルシラン(Sigma Aldrich)を用いてコーティングする、そして110 °Cのオーブンで硬化させることにより、ポリマーコートされたスライドガラスを準備する。スライドガラスを処理前後で蒸留水中で広範囲にわたって洗浄する。核酸をスライドガラス上に置き、次にSTRATALINKER UV架橋剤(Stratagene)を用いてアレイをUV照射に暴露することにより固定する。アレイを室温において0.2% SDSで洗浄し、蒸留水で3度洗浄する。リン酸緩衝生理食塩水(PBS)(Tropix, Inc., Bedford MA)中の0.2%カゼイン中において60 °Cで30分間アレイをインキュベートした後、0.2% SDS及び蒸留水で洗浄することにより、非特異結合部位をブロックする。

10

20

30

40

50

【0129】

膜ハイブリダイゼーションのプロープ調製

配列表のcDNAに由来するハイブリダイゼーションのプロープを使用して、膜系ハイブリダイゼーションでcDNA、mRNAまたはゲノムDNAをスクリーニングする。cDNAを45 µl TEバッファの濃度40-50 ngに希釈し、5分間100 °Cに加熱して変性させ、短時間遠心分離して、プロープを調製する。次に変性したcDNAをREDIPRIME 試験管(APB)を加え、青色が均一に分布するまで軽く混合して、次に短時間遠心分離した。5 µlの³²P]dCTPをその試験管に加え、内容物を37 °Cで10分インキュベートする。5 µlの0.2M EDTAを加えて標識化反応を停止する。そしてPROBEQUANT G-50 microcolumn (APB)を用いてプロープを組み込まれていないヌクレオチドから精製する。精製したプロープを5分間100 °Cに加熱する、そして2分間氷上で急速に冷却して、下記に記載した膜系ハイブリダイゼーションで使用する。

【0130】

ポリマーコートされたスライドハイブリダイゼーションのプロープ調製

サンプルから分離したmRNAに由来するハイブリダイゼーションのプロープを使用して、アレイベースのハイブリダイゼーションで配列表のcDNAをスクリーニングする。9 µl TEバッファで濃度200 ngにmRNAを希釈し、5 µl 5x バッファ、1 µl 0.1 M DTT、3 µl Cy3またはCy5標識化混合液、1 µl RNアーゼ阻害因子、1 µl 逆転写酵素、5 µl 1x 酵母調節mRNAを加え、GEMbrightキット(Incyte Genomics)を用いてプロープを調製する。非コード酵母ゲノムDNAから *in vitro* 転写により酵母調節mRNAを合成する(W. Lei, 未発表)。定量用対照として、サンプルmRNAに0.002 ng、0.02 ng、0.2 ng、2 ngでの対照mRNAの1つの設定をそれぞれ1:100,000、1:10,000、1:1000、1:100 (w/w)の比で逆転写反応混合液に希釈する。mRNA示差発現パターンを調べるために、対照mRNAの2番目の設定を1:3, 3:1, 1:10, 10:1, 1:25, また25:1 (w/w)の比で逆転写反応混合液に希釈する。反応混合液を混合して、37 °Cで2時間インキュベートする。次に反応混合液を20分間、85 °Cでインキュベートする。そして2つの連続するCHROMA SPIN+TE 30 カラム(Clontech, Palo Alto CA)を利用してプロープを精製する。精製したプロープは、プロープをDEPC処理水の90 µlに希釈させ、2 µl 1mg/ml グリコーゲン、60 µl 5 M 酢酸ナトリウム、300 µl 100% エタノールを加えて、エタノール析出させた。プロープを20分間、20,800xgで遠心分離

する。そしてペレットを12 μ lの再懸濁バッファで再懸濁し、5分間65 $^{\circ}$ Cに加熱してから、十分に混合する。プローブを加熱して、以前のように混合して、氷上に保管する。以下に詳述するように、プローブを高密度アレイベースのハイブリダイゼーションにおいて使用する。

【0131】

膜系ハイブリダイゼーション

1% Sarkosylと1x 高リン酸バッファ(0.5 M NaCl, 0.1 M Na₂HPO₄, 5 mM EDTA, pH 7)を含むハイブリダイゼーション溶液で、55 $^{\circ}$ Cで2時間、膜を前もってハイブリダイズする。15 mlの新鮮なハイブリダイゼーション溶液中で希釈されたプローブを次に膜に加える。膜をプローブと共に55 $^{\circ}$ Cで、16時間、ハイブリダイズする。ハイブリダイゼーション後、10

1mM Tris (pH 8.0), 1% Sarkosylで15分間、25 $^{\circ}$ Cで膜を洗浄し、さらに1mM Tris (pH 8.0)で、各回15分間、25 $^{\circ}$ Cで4回洗浄する。ハイブリダイゼーション化合物を検出するために、XOMAT-ARフィルム (Eastman Kodak, Rochester NY)を膜に一晩70 $^{\circ}$ Cで暴露する。そして現像してから、視覚で確認する。

【0132】

ポリマーコートされたスライドハイブリダイゼーション

プローブを5分間65 $^{\circ}$ Cに加熱してから、5415Cマイクロ遠心分離で5分間、9400 rpmで遠心分離する(Eppendorf Scientific, Westbury NY)。そして18 μ lのアリクオットをアレイ表面に置き、カバーガラスで覆う。アレイは、顕微鏡用スライドより僅かに大きい空洞を有する防水チェンバーに移す。チェンバーのコーナーに140 μ lの5 x SSCを加えることにより、チェンバー内部を湿度100%に保持する。このアレイを含むチャンバーを、60 $^{\circ}$ Cで約6.5時間インキュベートする。アッセイを1xSSC, 0.1% SDS中で、45 $^{\circ}$ Cで10分間洗浄し、次に0.1xSSCで、45 $^{\circ}$ Cで10分間の洗浄を3回繰り返した後に、乾燥させた。

【0133】

ハイブリダイゼーション反応を絶対的または示差ハイブリダイゼーションフォーマットで実施する。絶対ハイブリダイゼーションフォーマットでは、1つのサンプルのプローブをアレイエレメントにハイブリダイズし、ハイブリダイゼーション複合体形成後のシグナルを検出する。シグナル強度は、サンプル中のプローブmRNAレベルと相関する。示差ハイブリダイゼーションフォーマットでは、2つの生物サンプルでの遺伝子のセットの示差発現を分析する。2つのサンプルのプローブを調製して、異なる標識成分で標識化する。2つの標識されたプローブの混合液をアレイエレメントにハイブリダイズし、2つの異なる標識からの発光を個々に検出可能な条件下で2つの異なるシグナルを調べる。両方の生物サンプルから由来した実質的に同数のプローブにハイブリダイズするアレイ上のエレメントは、特有の結合した蛍光を出す(Shalon W095/35505)。

【0134】

ハイブリダイゼーション複合体は、Cy3の励起のためには488 nm、Cy5の励起のためには632 nmでスペクトル線を発生し得るInnova 70混合ガス10 Wレーザー(Coherent, Santa Clara CA)を備えた顕微鏡で検出する。20 x顕微鏡対物レンズ(Nikon, Melville NY)を用いて、アレイ上に励起レーザー光の焦点を当てる。アレイを含むスライドを顕微鏡のコンピュータ制御のX-Yステージに置き、20 μ mの解像度で対物レンズを通過してラスター

40

スキャンする。示差ハイブリダイゼーションフォーマットで、レーザーにより2つの蛍光色素を同時に励起する。発光された光は、波長に基づき分離され、2つの蛍光色素に対応する2つの光電子増倍管検出器(PMT R1477, Hamamatsu Photonics Systems, Bridgewater NJ)に送られる。アレイと光電子増倍管間に設置された好適なフィルターを用いて、シグナルをフィルターする。用いる蛍光色素の最大発光の波長は、Cy3では565 nm、Cy5では650 nmである。プローブ混合液に加えた酵母対照mRNAによって生み出されるシグナル強度を用いてスキャンの感度を較正する。アレイ上の特定の位置には相補的DNA配列を含め、その位置におけるシグナルの強度がハイブリダイズする種の重量比1:100,000に相関するようにする。

【0135】

光電子増倍管の出力は、IBMコンパチブルPCコンピュータにインストールされた12ビットRTI-835Hアナログ-デジタル(A/D)変換ボード(Analog Devices, Inc., Norwood MA)を用いてデジタル化される。デジタル化されたデータは、青色(低シグナル)から赤色(高シグナル)までの擬似カラー範囲へのリニア20色変換を用いてシグナル強度がマッピングされたようなイメージとして表示される。データは、定量的にも分析される。2つの異なる蛍光色素を同時に励起及び測定する場合には、各蛍光色素の発光スペクトルを用いて、データは先ず蛍光色素間の光学的クロストーク(発光スペクトルの重なり起因する)を補正する。グリッドが蛍光シグナルイメージ上に重ねられ、それによって各スポットからのシグナルはグリッドの各エレメントに集められる。各エレメント内の蛍光シグナルは統合され、シグナルの平均強度に応じた数値が得られる。シグナル分析に用いるソフトウェアは、GEMTOOLSプログラム(Incyte Genomics)である。

10

【0136】

8 電子分析

BLASTを用いて、GenBankやLifeSeqデータベース(Incyte Genomics)において同一または関連分子を検索した。ヒト及びラットの配列の積スコアを下記のようにして計算した。BLASTスコアにヌクレオチド一致率を乗じ、積を(2つの配列の短い方の長さの5倍)で除し、短い配列の長さへの100%アラインメントが積スコアを100にする。積スコアは2つの配列間の類似性の程度と配列一致の長さの両方を考慮する。例えば、積スコア40では一致はエラーが1%から2%以内の厳密さになり、少なくとも積スコア70では一致は厳密となる。類似または関連する分子は、通常、積スコアが8から40の間を示す分子を選択して同定

20

【0137】

電子ノーザン分析を、図1及び2において示すように積スコア70で実施した。LIFESEQデータベースのすべての配列とcDNAライブラリを系、器官/組織、細胞のタイプにより分類した。分類には、心血管系、結合組織、消化器系、胎芽構造、内分泌系、外分泌腺、女性生殖器、男性生殖器、生殖細胞、血液及び免疫系、肝、筋骨格系、神経系、脾臓、呼吸器系、感覚器、皮膚、顎口腔系、非分類性/混合性または尿路がある。各カテゴリーでは、配列が発現したライブラリの数を数えて、そのカテゴリー内のライブラリの総数を示した。非ノーマライズしたライブラリでは、2つ以上の発現レベルが有意である。

30

【0138】

9 相補的分子

cDNAに相補的な分子、約5(PNA)から5000 bp(cDNAインサートの相補体)を用いて、遺伝子発現を検出あるいは阻害する。これらの分子をOLIGOプライマ解析ソフトウェア(Molecular Biology Insights)を使用して選択した。検出については、実施例7に記載されている。プロモーターが結合するのを妨げることにより転写を阻害するため、相補的分子を最も固有な5'配列に結合し、オープンリーディングフレームの開始コドンの5'UTR上流のヌクレオチドを含むよう設計する。相補的分子は、ゲノム配列(エンハンサーまたはイントロン等)を含み、ポリメラーゼ、転写因子または調節分子の結合のために十分に開くような二重らせんの能力に影響を及ぼす「三重らせん」塩基対の形成に用いられる。翻訳を阻害するためには、相補的分子を設計して、リボソームが哺乳動物タンパク質をコードするmRNAに結合するのを阻害する。

40

【0139】

相補的分子を発現ベクターに置いて、一過性或いは短期治療向けに効果を試験するために細胞株を器官、腫瘍、滑腔また脈管系に使い、あるいは長期或いは安定した遺伝子治療向けに幹細胞、酵素または他の再生システムに形質転換するために用いる。非複製ベクターで一過性発現は1ヶ月以上続く、また形質転換/発現系でベクター複製を誘導する適切なエレメントが用いられる場合は3ヶ月以上続く。

【0140】

相補的分子をコードするベクターを有する適切な分裂する細胞の安定的な形質転換によって、遺伝形質転換した細胞株、組織または生命体を産生する(USPN 4,736,866)。安定した

50

組込みを可能にする十分な量のベクターを同化、複製する細胞も哺乳動物タンパク質をコードするcDNAの活性に影響を及ぼすか完全に除くための十分な相補的分子を生成する。

【0141】

1.0 配列、マイクロアレイ試薬、及び使用の選択

インサイトクローンは、LIFESEQ GOLDの構築 (assembled) されたヒト配列データベース (Incyte Genomics) に由来するテンプレート配列を示すものである。特定のテンプレートに対して1より多いクローンが利用可能であるケースでは、マイクロアレイにテンプレート中の5'-mostクローンが用いられる。HUMAN GENOME GEMシリーズ1-3マイクロアレイ (Incyte Genomics) は、28,626個のアレイエレメントを含み、それは10,068個のannotatedクラスタ及び18,558個のunannotatedクラスタを示す。UNIGEMシリーズのマイクロアレイ (Incyte Genomics) の為には、インサイト社のクローンは非重複性の単一遺伝子クラスタ (Uni gene database (build 46), NCBI; Shuler (1997) J Mol Med 75:694-698) にマッピングされた。また、最も強いBLAST配列 (少なくとも90%同一、及び100 bpのオーバーラップ) を有する5'クローンが、選択され、検証され、またマイクロアレイの合成で用いられる。UNIGEM Vマイクロアレイ (Incyte Genomics) は、7075のアレイエレメントを有し、それは4610個のannotated遺伝子及び2,184個のunannotatedクラスタを示す。

10

20

30

【0142】

マイクロアレイを構築するためには、cDNAは、cDNA挿入断片をフランキングするベクター配列に対するプライマーの相補性を用いることで細菌細胞から増幅された。30サイクルのPCRは、初期量のcDNA、1~2ngから最終量のcDNA5µgより多くなるまで増量された。増幅されたcDNAは、次にSEPHACRYL-400カラム (APB) を用いて精製された。精製されたcDNAは、ポリマーコートされたガラススライド上で固定化された。ガラス顕微鏡のスライド (Corning, Corning N.Y.) は、0.1% SDS及びアセトン中で超音波洗浄され、洗浄前後に多量の蒸留水で洗浄された。スライドガラスは、4%フッ化水素酸 (VWR Scientific Products, West Chester PA) 中でエッチングし、蒸留水中で広範囲にわたって洗浄し、95%エタノール中の0.05%アミノプロピルシラン (Sigma Aldrich) でコーティングする。コーティングしたスライドガラスは、110°Cのオーブンで硬化させる。米国特許第5,807,522号に記載されている方法を用いて、コーティングしたガラス基板にcDNAを付加する。平均濃度が100ng/µlのcDNA1µlを高速機械装置により開放型キャピラリープリンティングエレメント (open capillary printing element) に充填する。次にこの装置が、スライド毎に約5nlのcDNAを分注する。

【0143】

マイクロアレイには、STRATALINKER UVクロスリンカー (Stratagene) を用いてUV架橋する。マイクロアレイは、室温において0.2% SDSで1回洗浄し、蒸留水で3回洗浄する。非特異的な結合部位は、リン酸緩衝生理食塩水 (Tropix, Inc., Bedford MA) における0.2%カゼイン中で60分間マイクロアレイをインキュベートし、その後上述したように0.2% SDS及び蒸留水で洗浄することによってブロックする。

【0144】

1.1 サンプルの調合

ヒトBT20細胞

標本サンプル、ヒトBT20細胞は、*in vitro*での74歳の老人男性より単離した腫瘍の薄いスライスを遊出させた細胞に由来する乳ガンの株化細胞である。BT20細胞は、4、8、12、24、36、及び48時間に渡って濃度50ng/mlのEGFで処理された。すべてのケースで、未処置のBT20細胞からのmRNAが同時に処理された。

40

【0145】

1.2 ARPの発現

哺乳動物細胞発現系か昆虫細胞発現系のどちらかを用いて、タンパク質の発現と精製を達成する。pUB6/V5-His ベクター系 (Invitrogen, Carlsbad CA) を用いて、CHO細胞のCCMを発現する。ベクターは選択可能bsd遺伝子、多数のクローニング部位、ヒトユビキチン C 遺伝子のプロモーター/エンハンサー配列、抗-V5抗体での抗体検出のためのC-末端V5エピ

50

トープ、PROBOND 樹脂 (Invitrogen) 上での急速な精製のためのC-末端ポリヒスジン (6xHis) 配列を含む。形質移入した細胞を選択してプラストサイジンを含む培地に移す。

【0146】

Spodoptera frugiperda (Sf9) 昆虫細胞を組換え *Autographica californica* 核多角体病ウイルス (AcMNPV) で感染させる。相同組換えによって多角体遺伝子をcDNAと置換する。多角体プロモーターによりcDNAの転写が行われる。上述の精製を可能にする6xhisで、融合タンパク質としてタンパク質を合成する。精製したタンパク質を次の活性で用いて、抗体を作製する。

【0147】

1.3 抗体の産生

ポリアクリルアミドゲル電気泳動法を用いてARPを精製して、マウスやウサギを免疫化するために使用する。下記のプロトコルを用いて抗体を産生する。或いは、レーザGENEソフトウェア (DNASTAR) を用いてARPのアミノ酸配列を解析し、免疫原性の高い領域を決定する。通常C-末端付近或いは隣接する親水性領域内で見られる抗原性エピトープを選択して、合成し、抗体を生成するために用いられる。通常は、長さ約15残基のエピトープを、Fmocケミストリを用いるABI 431A ペプチドシンセサイザ (Applied Biosystems) を用いて生成し、N-マレイミドベンゾイル-N-ヒドロキシスクシンイミドエステル (MBS) を用いた反応によってKLH (Sigma-Aldrich) に結合させて、免疫原性を高める。

【0148】

完全フロイントアジュバントにおいてエピトープ-KLH複合体を用いてウサギを免疫化する。不完全フロイントアジュバントにおいて免疫化を間隔を置いて反復する。マウスには最低7週間、ウサギには最低12週間後、抗血清を抽出して、抗ペプチド活性のために検査した。検査には、ペプチドをプラスチックに結合すること、1%のウシ血清アルブミンでブロックすること、ウサギ抗血清と反応させて洗浄すること、さらに放射性ヨウ素標識されたヤギ抗ウサギIgGと反応させることが関係する。当分野で公知の方法を用いて、抗体価と形成複合体の量を決定する。

【0149】

1.4 特異的抗体を用いる天然タンパク質の精製

タンパク質を特異結合する抗体を用いて、イムノアフィニティークロマトグラフィにより天然あるいは組換えタンパク質を精製する。イムノアフィニティークラムは、CNBr-活性化SEPHAROSE樹脂 (APB) に抗体を共有結合させることにより形成する。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、タンパク質を優先的に吸着する条件下 (例えば洗浄剤が存在する高イオン強度緩衝液) でカラムを洗浄する。結合後、そのタンパク質を、抗体とタンパク質との結合を切るために、例えば、pH 2~3のバッファ、或いは高濃度の尿素またはチオシアン酸塩イオンを用いてカラムから溶出させ、タンパク質を回収する。

【0150】

1.5 cDNAまたはタンパク質と特異結合するためのスクリーニング分子

cDNAまたはその断片、タンパク質またはその部分を³²P-dCTP、Cy3-dCTPまたはCy5-dCTP (APB) あるいはBIODIPYかFITC (Molecular Probes, Eugene OR) で標識化する。前もって基質上に配置した候補分子または複合体のライブラリを標識したcDNAまたはタンパク質の存在下でインキュベートする。核酸配列かアミノ酸配列のための条件下でインキュベート後、基質を洗浄し、標識を保持する基質上の、特異結合か複合体成形を示す任意の位置がアッセイされ、リガンドを同定する。異なる濃度の核酸またはタンパク質を用いて得られたデータを使用して、標識された核酸かタンパク質と結合した分子の間の親和性を計算する。

【0151】

1.6 2つのハイブリッドスクリーン

酵母2ハイブリッドシステム、MATCHMAKER LexA 2ハイブリッドシステム (Clontech Laboratories, Palo Alto CA) を使用して、本発明のタンパク質を結合するペプチドをスクリ

10

20

30

40

50

ーニングする。タンパク質をコードするcDNAをpLexAベクター、リガンドの多数のクローニング部位に挿入して、大腸菌に形質転換する。mRNAから調製したcDNAをpB42ADベクターの多数のクローニング部位に挿入して、ライゲーションし、cDNAライブラリを作製するために大腸菌に形質転換する。pLexAプラスミドとpB42AD-cDNAライブラリ作製を大腸菌から単離し、ポリエチレングリコール/リチウムアセテートプロトコルを用いて形質転換受容性をもつ酵母EGY48[p8op-lacZ]細胞に同時形質転換するために比率2:1で使用する。形質転換した酵母細菌を、ヒスチジン(-His)、トリプトファン(-Trp)、ウラシル(-Ura)を含まない合成ドロップアウト(SD)培地で培養する、そしてコロニ-が成長して数えられるまで、30 でインキュベートする。コロニ-を最小限の容積の1x TE (pH 7.5)で集め、2% ガラクトース(Gal), 1% ラフィノース(Raf)、80 mg/ml 5-ブromo-4-クロロ-3-インドリル-d-ガラクトピラノシド(X-Gal)で補われるSD/-His/-Leu/-Trp/-Ura 培地で再培養する。次いで青色コロニ-の成長を検査する。発現したタンパク質とcDNA融合タンパク質間の相互作用により、EGY48のLEU2レポーター遺伝子を活性化し、ロイシン(-Leu)を含まない培地でコロニ-成長を起こす。相互作用はまた、X-Galで成長するコロニ-に青色を産生するp8op-lacZレポーター作成物からの -ガラクトシダーゼの発現を活性化する。

10

【0152】

発現したタンパク質とcDNA融合タンパク質間の陽性の相互作用が、個々の陽性コロニ-を単離して、30 で1から2日間、SD/-Trp/-Ura 液体で成育することにより確認できる。培地のサンプルをSD/-Trp/-Ura培地で培養し、コロニ-が出現するまで30 でインキュベートする。サンプルをSD/-Trp/-UraとSD/-His/-Trp/-Ura プレートで複製培養する。ヒスチジンを含まない培地では成育せず、ヒスチジンを含むSDで成育するコロニ-は、pLexAプラスミドを失っている。ヒスチジンを必要とするコロニ-をSD/Gal/Raf/X-Gal/-Trp/-Uraで成育する。そして白色コロニ-を単離して、増殖する。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cDNA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。

20

【0153】

1.7 ARPアッセイ

ARP活性は、¹²⁵I標識ARPの存在条件下で、aPKC、PKC、及びPKC のような候補リガンド分子を用いてリガンド結合アッセイで決定される。ARPを、¹²⁵Iボルトンハンター試薬で標識する(例えば、Bolton A.E.及びW.M. Hunter (1973) Biochem. J. 133:529-539を参照)。マルチウェルプレートの穴の中に予め配列しておいた候補分子を、標識したARPと共にインキュベートして洗浄し、標識したARP複合体を有する全ての穴をアッセイする。ARP濃度を変えて得たデータを用いて、候補分子とのARPの数、親和性及び会合の値を計算する。

30

【0154】

本明細書において記載した全ての特許出願、刊行物は、言及することをもって本明細書の一部とする。当業者は、本発明の範囲及び精神から逸脱することなく本発明の記載した方法及びシステムの種々の改変を行い得る。本発明について説明するにあたり特定の好適実施例に関連して説明を行ったが、本発明の範囲が、そのような特定の実施例に不当に制限されるべきではないことを理解されたい。実際に、分子生物学または関連分野の専門家には明らかな、本明細書に記載されている本発明の実施方法の様々な改変は、特許請求の範囲内にあるものとする。

40

【0155】

(表の簡単な説明)

表1及び2は、LIFESEQ Goldデータベース(Incyte Genomics, Palo Alto)を用いて作成されたARP-1及びARP-2のノーザン分析を示すものである。表1では第1列が組織カテゴリーを示す。第2列は、組織カテゴリー内のトータルクローン数を示す。第3列は、少なくとも一つの転写物が見られるライブラリの数と全体のライブラリ数との比率を示す。第4列は、転写の絶対クローン発生量を示す。第5列は転写の発生量のパーセンテージを示す。表2は、膀胱組織内のARP発現を示すものである。第1列は、ライブラリ名をリストしている

50

。第2列はライブラリのために配列されたクローンの数を示している。第3列はライブラリが誘導された組織について記載している。第4列は転写の絶対的な発生量を示す。第4列は、転写の発生量のパーセンテージを示す。

【0156】

表3は、マイクロアレイ解析によって決定されるように、未処置細胞と比較したEGF処理ヒトBT20乳癌細胞内のARP-2の差次的発現を示す。第1列は、 $10g_2$ DE (処理細胞/未処理細胞)として示される平均的な差次的発現 (DE) をリストしたものである。第2列は蛍光性の緑色素Cy3でラベルされた未処理の制御サンプルをリストしたものである。第3列は、蛍光性の赤色素Cy5でラベルされたサンプルの処理をリストするものである。

【図面の簡単な説明】

10

【0157】

【図1 - A】図は、cDNA (SEQ ID NO:3) によってコードされたARP-1 (SEQ ID NO: 1) を示している。翻訳は、MACDNASIS PRO software (Hitachi Software Engineering, South San Francisco CA.) によってなされた。

【図1 - B】図は、cDNA (SEQ ID NO:3) によってコードされたARP-1 (SEQ ID NO: 1) を示している。翻訳は、MACDNASIS PRO software (Hitachi Software Engineering, South San Francisco CA.) によってなされた。

【図1 - C】図は、cDNA (SEQ ID NO:3) によってコードされたARP-1 (SEQ ID NO: 1) を示している。翻訳は、MACDNASIS PRO software (Hitachi Software Engineering, South San Francisco CA.) によってなされた。

20

【図1 - D】図は、cDNA (SEQ ID NO:3) によってコードされたARP-1 (SEQ ID NO: 1) を示している。翻訳は、MACDNASIS PRO software (Hitachi Software Engineering, South San Francisco CA.) によってなされた。

【図1 - E】図は、cDNA (SEQ ID NO:3) によってコードされたARP-1 (SEQ ID NO: 1) を示している。翻訳は、MACDNASIS PRO software (Hitachi Software Engineering, South San Francisco CA.) によってなされた。

【図1 - F】図は、cDNA (SEQ ID NO:3) によってコードされたARP-1 (SEQ ID NO: 1) を示している。翻訳は、MACDNASIS PRO software (Hitachi Software Engineering, South San Francisco CA.) によってなされた。

【図1 - G】図は、cDNA (SEQ ID NO:3) によってコードされたARP-1 (SEQ ID NO: 1) を示している。翻訳は、MACDNASIS PRO software (Hitachi Software Engineering, South San Francisco CA.) によってなされた。

30

【図1 - H】図は、cDNA (SEQ ID NO:3) によってコードされたARP-1 (SEQ ID NO: 1) を示している。翻訳は、MACDNASIS PRO software (Hitachi Software Engineering, South San Francisco CA.) によってなされた。

【図1 - I】図は、cDNA (SEQ ID NO:3) によってコードされたARP-1 (SEQ ID NO: 1) を示している。翻訳は、MACDNASIS PRO software (Hitachi Software Engineering, South San Francisco CA.) によってなされた。

【図1 - J】図は、cDNA (SEQ ID NO:3) によってコードされたARP-1 (SEQ ID NO: 1) を示している。翻訳は、MACDNASIS PRO software (Hitachi Software Engineering, South San Francisco CA.) によってなされた。

40

【図1 - K】図は、cDNA (SEQ ID NO:3) によってコードされたARP-1 (SEQ ID NO: 1) を示している。翻訳は、MACDNASIS PRO software (Hitachi Software Engineering, South San Francisco CA.) によってなされた。

【図2 - A】図は、cDNA (SEQ ID NO:20) によってコードされたARP-2 (SEQ ID NO: 2) を示している。翻訳は、MACDNASIS PRO software (Hitachi Software Engineering.) によってなされた。

【図2 - B】図は、cDNA (SEQ ID NO:20) によってコードされたARP-2 (SEQ ID NO: 2) を示している。翻訳は、MACDNASIS PRO software (Hitachi Software Engineering.) によってなされた。

50

なされた。

【図3 - A】図は、ARP-1 (1555118; SEQ ID NO: 1)、ARP-2 (2582063; SEQ ID NO:2)、ラット ASIP (g3868778; SEQ ID NO:62)、及びヒトASIP (g8037915; SEQ ID NO:63)の配列及びドメイン間の化学的、構造的類似性を示している。翻訳は、LASERGENE software のMEGALIGN プログラム (DNASTAR, Madison WI)によってなされた。

【図3 - B】図は、ARP-1 (1555118; SEQ ID NO: 1)、ARP-2 (2582063; SEQ ID NO:2)、ラット ASIP (g3868778; SEQ ID NO:62)、及びヒトASIP (g8037915; SEQ ID NO:63)の配列及びドメイン間の化学的、構造的類似性を示している。翻訳は、LASERGENE software のMEGALIGN プログラム (DNASTAR, Madison WI)によってなされた。

【図3 - C】図は、ARP-1 (1555118; SEQ ID NO: 1)、ARP-2 (2582063; SEQ ID NO:2)、ラット ASIP (g3868778; SEQ ID NO:62)、及びヒトASIP (g8037915; SEQ ID NO:63)の配列及びドメイン間の化学的、構造的類似性を示している。翻訳は、LASERGENE software のMEGALIGN プログラム (DNASTAR, Madison WI)によってなされた。

10

【図3 - D】図は、ARP-1 (1555118; SEQ ID NO: 1)、ARP-2 (2582063; SEQ ID NO:2)、ラット ASIP (g3868778; SEQ ID NO:62)、及びヒトASIP (g8037915; SEQ ID NO:63)の配列及びドメイン間の化学的、構造的類似性を示している。翻訳は、LASERGENE software のMEGALIGN プログラム (DNASTAR, Madison WI)によってなされた。

【図3 - E】図は、ARP-1 (1555118; SEQ ID NO: 1)、ARP-2 (2582063; SEQ ID NO:2)、ラット ASIP (g3868778; SEQ ID NO:62)、及びヒトASIP (g8037915; SEQ ID NO:63)の配列及びドメイン間の化学的、構造的類似性を示している。翻訳は、LASERGENE software のMEGALIGN プログラム (DNASTAR, Madison WI)によってなされた。

20

【図3 - F】図は、ARP-1 (1555118; SEQ ID NO: 1)、ARP-2 (2582063; SEQ ID NO:2)、ラット ASIP (g3868778; SEQ ID NO:62)、及びヒトASIP (g8037915; SEQ ID NO:63)の配列及びドメイン間の化学的、構造的類似性を示している。翻訳は、LASERGENE software のMEGALIGN プログラム (DNASTAR, Madison WI)によってなされた。

【図3 - G】図は、ARP-1 (1555118; SEQ ID NO: 1)、ARP-2 (2582063; SEQ ID NO:2)、ラット ASIP (g3868778; SEQ ID NO:62)、及びヒトASIP (g8037915; SEQ ID NO:63)の配列及びドメイン間の化学的、構造的類似性を示している。翻訳は、LASERGENE software のMEGALIGN プログラム (DNASTAR, Madison WI)によってなされた。

【図3 - H】図は、ARP-1 (1555118; SEQ ID NO: 1)、ARP-2 (2582063; SEQ ID NO:2)、ラット ASIP (g3868778; SEQ ID NO:62)、及びヒトASIP (g8037915; SEQ ID NO:63)の配列及びドメイン間の化学的、構造的類似性を示している。翻訳は、LASERGENE software のMEGALIGN プログラム (DNASTAR, Madison WI)によってなされた。

30

【図3 - I】図は、ARP-1 (1555118; SEQ ID NO: 1)、ARP-2 (2582063; SEQ ID NO:2)、ラット ASIP (g3868778; SEQ ID NO:62)、及びヒトASIP (g8037915; SEQ ID NO:63)の配列及びドメイン間の化学的、構造的類似性を示している。翻訳は、LASERGENE software のMEGALIGN プログラム (DNASTAR, Madison WI)によってなされた。

【図3 - J】図は、ARP-1 (1555118; SEQ ID NO: 1)、ARP-2 (2582063; SEQ ID NO:2)、ラット ASIP (g3868778; SEQ ID NO:62)、及びヒトASIP (g8037915; SEQ ID NO:63)の配列及びドメイン間の化学的、構造的類似性を示している。翻訳は、LASERGENE software のMEGALIGN プログラム (DNASTAR, Madison WI)によってなされた。

40

【0158】

【表1】

表 1

| 組織カテゴリ | クローン カウント | 発見された数 | 絶対 発生量 | 発生量の パーセンテージ |
|---------|--------------|---------|-----------|-----------------|
| 心血管系 | 266190 | 1/68 | 1 | 0.0004 |
| 結合組織 | 144645 | 0/47 | 0 | 0.0000 |
| 消化器系 | 501101 | 4/148 | 4 | 0.0008 |
| 胚構造 | 106713 | 0/21 | 0 | 0.0000 |
| 内分泌系 | 225386 | 0/53 | 0 | 0.0000 |
| 外分泌腺 | 254635 | 1/64 | 1 | 0.0004 |
| 生殖 女性 | 427284 | 2/106 | 2 | 0.0005 |
| 生殖 男性 | 448207 | 1/114 | 1 | 0.0002 |
| 生殖細胞 | 38282 | 1/5 | 1 | 0.0026 |
| 血液及び免疫系 | 680277 | 2/159 | 2 | 0.0003 |
| 肝臓 | 109378 | 0/35 | 0 | 0.0000 |
| 筋骨格系 | 159280 | 0/47 | 0 | 0.0000 |
| 神経系 | 955753 | 3/198 | 5 | 0.0005 |
| 腺臓 | 110207 | 0/24 | 0 | 0.0000 |
| 呼吸器系 | 390086 | 3/93 | 3 | 0.0008 |
| 感覚器 | 19256 | 0/8 | 0 | 0.0000 |
| 皮膚 | 72292 | 0/15 | 0 | 0.0000 |
| 顎口腔系 | 12923 | 0/10 | 0 | 0.0000 |
| 非分類/混合 | 120926 | 2/13 | 2 | 0.0017 |
| 尿管 | 279062 | 3/64 | 5 | 0.0018 |
| 合計 | 5321883 | 23/1292 | 27 | 0.0005 |

10

20

30

40

【 0 1 5 9 】

【 表 2 】

表 2

| クローン カウント | 発見された数 | 絶対 発生量 | 発生量の パーセンテージ |
|--------------|--------|-----------|-----------------|
| 7884 | 0 | 0 | 0.0381 |
| 3776 | 0 | 0 | 0 |
| 7884 | 3 | 3 | 0.0381 |

ライブラリの検索
Bladder tumor, TC CA, 60M, m/BLADN0T05
Bladder, mw/TC CA, CA in situ, 60M, m/BLADTUF04

ライブラリID
BLADTUF04
BLADN0T05

【 0 1 6 0 】

【 表 3 】

表3

| mean log ₂ DE (Cys/Cy3) | Cy3 | Cy5 |
|--|--------------------------------------|--|
| -0.08 | Human, BT20 Line, Unix 4hr, AdenoCA | Human, BT20 Line, \uparrow EGF 50ng/ml 4hr, AdenoCf |
| -0.61 | Human, BT20 Line, Unix 8hr, AdenoCA | Human, BT20 Line, \uparrow EGF 50ng/ml 8hr, AdenoCf |
| -0.61 | Human, BT20 Line, Unix 12hr, AdenoCA | Human, BT20 Line, \uparrow EGF 50ng/ml 12hr, AdenoCf |
| -0.62 | Human, BT20 Line, Unix 24hr, AdenoCA | Human, BT20 Line, \uparrow EGF 50ng/ml 24hr, AdenoCf |
| -1.23 | Human, BT20 Line, Unix 36hr, AdenoCA | Human, BT20 Line, \uparrow EGF 50ng/ml 36hr, AdenoCf |
| -1.04 | Human, BT20 Line, Unix 48hr, AdenoCA | Human, BT20 Line, \uparrow EGF 50ng/ml 48hr, AdenoCf |

10

20

【 1 - A 】

5' A AGG GGC GCT GCC GCG AGC CTC CCG GGC TCA GGG TGT TCC GGG GAG CGG GGC 54
63 CCC GGG TCT CTG GGC CCA CCC GCC CCG GGC CTC CCA GAG TGG GGG CTG GGC 108
117 CCG CGG GGT CMG ACA CCT GGT CCG CCC GGC CCG GCG TGG TCG CCG GGG GCC AAG 162
171 ATG AAA GTG ACC GTG TGC TTC GGC GGG ACC GGC ATC GTG GTC CCC TGC AAG GAG 216
M K V T V C F G R T G I V V P C K E 207
225 GGC CMG GTG GCT GTG GGC GAG CTC ACC CMG CMG GCG CTG CMG GGG TAC CTG AAG 270
G Q U R V G E L T Q A L Q R Y L K 315
279 ACC CGG GAG AAG GGT CCT GGT TAC TGG GTG AAG ATT CAT CAC TTA GAA TAT ACA 324
T R E K G P G Y W V K I H H L E Y T

【 1 - B 】

333 GAT GGA GGA ATC CTG GAT CCA GAT GAT GTC TTC GCA GAT GGT GTT GAA GAT AAA 378
D G G I L D P D D V L A D V V E D K
387 GAC AAG CTG ATT GCT GTG TTT GAA GAA GAA GAA CCA CTC CAC AAG AAG ATT GAG AGC 432
D K L I A V F E E Q E P L H K I E S
441 CCC AGT GGA AAC CCT GCA GAT CCG CMG AGC CCA GAT CCT TTT GAG ACA GAA GTG 486
P S G N P A D R Q S P D A F E T E V
495 GCC GCC CAA CTG GCT GCA TTT AAG CCA ATT GGT GGG GAG ATT GAA GTA ACC CTT 540
A A Q L A A F K P I G G E I E V T P
549 TCT CCT CTA AAA CTA GGC ACT CCA CTG GTG AGG AGA AGC AGT GAC CCA GTC 584
S A L K L G T F L L V R R S S D P V

【 1 - C 】

603 612 621 630 638 648
 CCA GGC CCA CCT GCT GAT ACC CAG CCA AGC GCT TCA CAC CCT GGT GGC CAG AGT
 P G P P A D T Q P S A S H P G G Q S
 657 666 675 684 693 702
 CTG AAA CTG GTT CCA GAT TCC ACG CAG AAC TTG GAA GAC AGA GAA GTT TTG
 L K L V V P D S T Q N L E D R E V L
 711 720 729 738 747 756
 AAT GGT GTA CAG ACA GAA CTA ACT TCG CCA AGA ACT AAG GAC ACA TTG AGT
 N G V Q T E L L T S P R T K D T L S
 765 774 783 792 801 810
 GAT ATG AGS AGA ACA CTS GAG ATT TCT GGG GAA GGA GGC CCA TTG GGA ATG CAT
 D M T R T V E I S G E G G P L G I H
 819 828 837 846 855 864
 GTA GTC CCC TTC TTT TCA TCT CTG AAT GGA AGG AAT CTA GSA CTC TTC ATC CGA
 V V P F F S S L S G R I L G L F I R

【 1 - D 】

873 882 891 900 909 918
 GGC AAT GAA GAC TAC ACC AGG TCC AMG CCG GAG GGA CTA TTT CAC GAA AAT GAA
 G I E D N S R S K R E G L F H E N E
 927 936 945 954 963 972
 TGT ATT GTA AAA ATC AAC AAT GTG GAT CTC GTA GAC AAA ACC TTT GCT CAG GCT
 C I V K I N N V D L V D K T F A Q A
 981 990 999 1008 1017 1026
 CAA GAT GTC TTC CCG CAG GCA ATG AAA TCT CCA AGT GTG CTC CTC CAC GTG GTT
 Q D V F R Q A M K S F S V L L H V L
 1035 1044 1053 1062 1071 1080
 CCT CCA CAA AAC CGT GAA CAG TAT GAA AAG TCA GTC ATT GGC TCT CTT AAC ATT
 P F Q N R E Q Y E K S V I G S L N I
 1089 1098 1107 1116 1125 1134
 TTT GGT AAT AAT GAT GGC GTT TTG AAA ACC AAA GTS CCG CCT CTT GTC CNT GGA
 P G N N D G V L K T K V P P V H G

【 1 - E 】

1143 1152 1161 1170 1179 1188
 AAA TCG GGA CTA AAG ACA GCA AAT CTC ACA ACC GAT AGT CCT GAA ACA GAT
 K S G L K T A N L T G T D S P E T D
 1197 1206 1215 1224 1233 1242
 GCA TCA GCT TTC CTG CAA CAA AAC AAG AGT CCC CGA GTA CCA AGS CTG GGA GGA
 A S A S L Q Q N K S F R V P R L G G
 1251 1260 1269 1278 1287 1296
 AAA CCA TCC TCT CCC TCA CTC TCG CCT CTC ATG GGA TTT GGC AGC AAT AAA AAT
 K P S S P S L S P L M G F G S N K N
 1305 1314 1323 1332 1341 1350
 GCA AAG AAA AAT AAG ATT GAC CTA AAG AAA GGC CCT GAA GCA CTT GGT TTC ACT
 A K K I K I D L K K G P E G L G F T
 1359 1368 1377 1386 1395 1404
 GTG GAT ACC AGA GAC TCT TCC ATA CAT GGT CCC GGT CCC AAT TTT GTA AAA AAC
 V V T R D S S I H G P G P I P V K N

【 1 - F 】

1413 1422 1431 1440 1449 1458
 AAT TTA CCA AAG GGA GCA CCA ATA AAA GAT GGC CCC CTA CAA TCA GGG GAC AGA
 I L P K G A A I K D G R L Q S G D R
 1467 1476 1485 1494 1503 1512
 ATT TTG GAG GTA AAT GGG AGA GAT CTC ACC GGA CCA ACC CAG GAA GAG CTT GAG
 I L E V N G R D V T G R T Q E E L V
 1521 1530 1539 1548 1557 1566
 GCC ATG CTC AAG ACC AAG CAG GGG GAG ACA GCA TCG CTG GTC ATT GCC CCG
 A M L R S T K Q G E T A S L V I A R
 1575 1584 1593 1602 1611 1620
 CAA GAA GGA GAT TTT CTG CCC CGA GAG TTG BAA GSA GSA GAA CCA GAC TCC TGT GCA
 Q E G H F L P R E L K G E P D C C A
 1629 1638 1647 1656 1665 1674
 CTC TCT CTG GAG ACA AGC GAG CAG CTC ACC TTT GAG ATC CCC CTG AAT GAT TCA
 L S L E T S E Q L T F E I P L N D S

【 1 - G 】

1692 1701 1710 1719 1728
 GGT TCT GCT GGC CTC GGG GTG AGC TTA AAA GGG AAC AAA TCC AGA GAA ACT GGA
 G S A G L G V S L K G N K S R E T G
 1737 1746 1755 1764 1773 1782
 ACA GAC TGG GGG ATG TTT ATC AAA TCC ATC ATG GGA GGC GCT GCT TTT TGG
 T D L G I F I K S I I H G G A A F K
 1800 1809 1818 1827 1836
 GAT GGT GGT CTA GAT GAT GAT GAT GAT GAT GAT GAT GAT GAT GAT GAT GAT GAT
 D G R L R M N D Q L I A V N G E S L
 1845 1854 1863 1872 1881 1890
 TTG GGA TAG TCC AAC CAC GAA GCT ATG GAA ACA CTT AGG GCG CCA ATG TCC ATG
 L G K S N H E A M E T L R R P M S M
 1899 1908 1917 1926 1935 1944
 GAG GSA AAC ATC CGA GGG ATG ATC CAG TTG GTG ATT CTG AGG AGS CCA GAG AGA
 E G N I R G M I Q L V I L R R P E R

【 1 - H 】

1953 1962 1971 1980 1989 1998
 CCA ATG GAG GAT CCT GCA GAG TGT GGG GCA TTT TCC AAG CCA TGC TTT GAG AAC
 P M E D P A E C G A F S K P C F E N
 2007 2016 2025 2034 2043 2052
 TCT CAA AAT GCT GTA ACC ACC TCT AGG CCA AAT GAA AAT AGT ATC CTG GAT CCA
 C Q N A V T T S R R N D N S I L H P
 2061 2070 2079 2088 2097 2106
 CTT GGC ACT TGC AGT CCA CAA GAC AAA CAG AAA GGT CTA TTG CTG CCC AAT GAC
 L G T C S P Q D K Q K G L L L P N D
 2115 2124 2133 2142 2151 2160
 GGA TGG GCC GAG AGT GAA GTT CCA CCT TCT CCA ACA CCA GAT TCT GCT CTG GGA
 G W A E S E V P P S P T P H S A L G
 2169 2178 2187 2196 2205 2214
 TTG GGC CTC GAA GAT TAC AGC CAC AGC TCT GGG GTG GAT TCA CCA GTA TAT TTT
 L G L E D Y S S S G V D S A V Y F

【 1 - I 】

2223 2232 2241 2250 2259 2268
 CCA GAT CAG CAC ATC AAC TTC ABA TCT GTG ACA CCG GCC AGS CAG CCT GAA TCA
 P D Q H I N F R S V T P A R Q P E S
 2277 2286 2295 2304 2313 2322
 AAT AAT TTG AAA GCC TCG AAG AGC ATG GAC CTT GTG CCA GAT GAA AGC AAG GTT
 I N L K A S K S M D L V P D E S K V
 2331 2340 2349 2358 2367 2376
 CAC TCA TTG GCT GSA CAA AAA TCG GAA TCT CCA AGC AAA GAT TTT GGT CCA ACT
 H S L A G Q K S E S F S K D F G P T
 2385 2394 2403 2412 2421 2430
 CTG GAT TTG AAA AAG TCC AGC TTC TTG GAG AGT CTG CAG ACT GCA GTG GCC GAG
 L G L K K S S S L E S L Q T A V A E
 2439 2448 2457 2466 2475 2484
 GTC AGG AAG AAT GAC CTT CCC TTT CAC AGS CCC CCG CAG ATG GAT CCA GGC
 V R K N D L P F H R P R F H N V R G

【 1 - J 】

2493 2502 2511 2520 2529 2538
 CGA GGC TGC AAT GAG AGC TTT ABA GCA GCC ATT GAC AAA TCC TAC GAT GSA CCT
 R G C N E S F R A A I D K S Y D G P
 2547 2556 2565 2574 2583 2592
 GAA GAA ATA GAA GCT GAC GGT CTG TCT GAT AAG AGC TCT CAC TCT GGC CAA GGA
 E E I E A D G L S D K S S H S G Q G
 2601 2610 2619 2628 2637 2646
 GCT CTG AAT TGT GAT TCT GCC CCT CAG GGG AAT TCG GAG CTA GAG GAC ATC GAA
 A L N C E S A F Q G N S E L E D M E
 2655 2664 2673 2682 2691 2700
 AAT AAA GCC AGS AAA GTC AAA AAA ACG AAA GAG AAG AAA AAG AAA AAG AAA AAG
 N K A R K V K K T K E K E K K E K
 2709 2718 2727 2736 2745 2754
 GGC AAA TTG AAA GTC AAG GAG AAA AAG CCC AAA GAG GAG AAT GAA GAT CCA GAA
 G K L K V K E K K R K E E N E D P E

【 1 - K 】

2753 2772 2781 2790 2799 2808
 AGG AAA ATX AAG AAG AAG CGC TTC GGC GCC AAG CTG AGG TAT GGG CCT GCT TTG
 R K I K K K G F G A M L R Y G P A L
 2817 2826 2835 2844 2853 2862
 AAG GCA AAG TTG GTT CTC ATT TTG TCT CTC CTG AAA AAA GGC CAC GCT TTT CCT
 K A K L V L I L S L L K K A H A F P
 2871 2880 2889 2907 2916
 CTT CTG CCA AAT GCA TAC GGC TGT CAA TTC TCT GCT CGT TCT CTT TCT GCA
 R L Q P N A Y G S Q P C A R S L S A
 2925 2934 2943 2952 2961 2970
 GAG GAA GAG GAG GAT TTT GGG GAA AAT TAC AGT GAT GAC AGC ACA CTG TCT TAA 3'
 E A E E L F G E S Y S D D R T L S

【 2 - A 】

9 18 27 36 45 54
 ATG AAA ATG ACC GTG TGC TTC GGA GGG ACC GGG GTG GTC GTG CCG TGC GGG GAC
 M K V T V C F G R T R V V V F C D
 63 72 81 90 99 108
 GGC CAC ATG AAA GTT TTC AGC CTC ATC CAG GCG GTG ACC CGC TAC CGS AAG
 G H M K V F S L I Q Q A V T R Y R K
 117 126 135 144 153 162
 GCC ATC CCC AAG GAT CCA AAT TAC TGG ATA CAG GTG CAT CSC TTG GAA CAT GGA
 A I A K D P N Y W I Q V H R L E H G
 171 180 189 198 207 216
 GAT GGA GGA ATA CTA GAT GAT GAT GAT GAT GAT GAT GAT GAT GAT GAT GAT GAT
 D G G I L D L D I L C D V A D D K
 225 234 243 252 261 270
 GAC AGA CTG GTA GCA GTG TTT GAT GAG CAG GAT CCA CAT CAC GGA GGT GAT GGC
 D R L V A V F D E Q D P H H G G D G

【 2 - B 】

279 288 297 306 315 324
 ACC AAT GGC ACT TCC ACG GGT ACC CAG AGC CCA GAG ATA TTT GGT AGT GAG CTT
 T S A S S T G T Q S P E I F G S E L
 333 342 351 360 369 378
 GGC ACC AAC AAT GTC TCA GCC TTT CAG CCT TAC CAA GCA ACA AGT GAA ATG GAG
 G T N N V S A P Q P Y Q A T S E I E
 387 396 405 414 423 432
 GTC ACA CTT TCA GTC CTT CGA GCA AAT ATG OCT CTT CAT GTT CGA CCC AGT AAT
 V T P S V L R A N M P L H V R R S S
 441 450 459 468 477 486
 GAC CCA GCT CTA AAT GGC CTC TCC ACT TCT GTC AGT GAT AAT TTT TCC TCT
 D P A L I G L S T S V S D S N F S S
 495 504 513 522 531 540
 GAA GAG CTT TCA AAG AAA AAT CCC ACA GCG TGG TCA ACA ACA GCT GGC TTC CTC
 E E P S R K N P T R W S T T A G F L

【 2 - C 】

549 558 567 576 585 594
 AAG CAG AAC ACT GCT GGG AGT CCT AAA ACC TTC GAC AGS AAG AAA GAT GAA AAC
 K Q N T A G S P K T C D R K K D E N
 603 612 621 630 639 648
 TAC AGA AGC CTC CCG GAG AAT ACT GAT TCG TCT AAC GAT AAT GAT TTT CAG AGA GAC
 Y R S L P R D T S N W S N Q F Q R D
 657 666 675 684 693 702
 AAT GCT CSC TCG TCT CTG AAT GCC AAT CAC CCA ATG GTS GGC AAG TGG CTG GAG
 N A R S L S A S H P M V G K W L E
 711 720 729 738 747 756
 AAA CAA GAA CAG GAT GAG GAT GGG ACA GAA GAG GAT AAC AAT GAT GAT GAA CCT
 K Q E Q D E D G T E E D N S R V E P
 765 774 783 792 801 810
 GTT GGA CAT GCT GAC ACG GAT TTG GAG CAT AAT CCC AAC TTT TCT CTG GAT GAT
 V G H A D T G L E H I F N F S L D D

【 2 - D 】

819 828 837 846 855 864
 ATG GTA AAG CTC GTA GAA GTC CCC AAC GAT GGA GGG CCT CTG GGA ATC CAT GTA
 M V K L V E V P N D G G P L G I H V
 873 882 891 900 909 918
 GTG CCT TTC AGT GCT CGA GGC AGC AGA ACC CTG GGG TTA TTA GTA AAA CGA TTG
 V P F S A R G G R T L G L L V K R L
 927 936 945 954 963 972
 GAG AAA GGT GGT AAA GCT GAA CAT GAA AAT CTT TTT CCG GAG AAT GAT TGC ATT
 E K G G K A E H E N L F R E N D C I
 981 990 999 1008 1017 1026
 GTC AGS AIT AAT GAT GGC GAC CTT CGA AAT AGA AGA TTT GAA GAA GCA CAT
 V R I N D G D L R N R R F E Q A Q H
 1035 1044 1053 1062 1071 1080
 ATG TTT CCG CAA GCC AAG GGT ACA CCC ATC AAT TGG TTC CAT GTG GTT CCT GCA
 K P R Q A H R T P I I W F H V V P A

【 2 - E 】

1089 1098 1107 1116 1125 1134
 GCA AAT AAA GAG CAG TAT GAA CAA CTA TTC CAA AAT GAG AAG AAC AAT TAC TAT
 A N K E Q Y E Q L S Q S E K N N Y Y
 1143 1152 1161 1170 1179 1188
 TCA AGC GGT TTT AGC CCT GAC ACC CAG TAT ATT GAC AAC AGG AAT GTG AAC AGT
 S S R F S P D S Q Y I D N R S V N S
 1197 1206 1215 1224 1233 1242
 GCA GGG CTT CAC ACG GTG CAG AGA GCA CCC CGA CTG AAC CAC CCG CCT GAG CAG
 A G L H T V Q R A P R L N H P P E Q
 1251 1260 1269 1278 1287 1296
 ATA GAC TGT CAC TCA AGA CTA CCT CRI AGC GCA CAC CCC TCG GGA AAA CCA CCA
 I D S H S R L P H S A H P S G K P
 1305 1314 1323 1332 1341 1350
 TCC GCT CCA GCC TCG GCA CCT CAG AAT GTA TTT AGT ACG ACT GTA ACG AGT GGT
 S A P A S A P Q N V F S T T V S S G

【 2 - F 】

1359 1368 1377 1386 1395 1404
 TAT AAC ACC AAA ATA GGC AAG ACG CTT AAT ATC CAG CTT AAG AAA GGT ACA
 Y N T K K I G K R L N I Q L K K G T
 1413 1422 1431 1440 1449 1458
 GAA GGT TTG GGA TTC ACC ATC ACT TCC AGA GAT GTA ACA ATA GTA GGT GGC TCA GCT
 E G L G F S I T S R D V T I G S A
 1467 1476 1485 1494 1503 1512
 CCA ATC TAT GTG AAT AAC AAT CTC CCC GGG GGG GGC AAT CAG GAT GGC CGA
 P I Y V K N I L P R G A A I Q D G R
 1521 1530 1539 1548 1557 1566
 CTT AAG GCA GGA GAC AGA CTT ATA GAG GTA AAT GGA GTA GAT TTA GTC GCC AAA
 L K A G D R L I E V N G V D L V G K
 1575 1584 1593 1602 1611 1620
 TCC CAA GAG GAA GTT GTT TCG CTG TTG AGA AGC ACC AAG AAG GAA GGA ACT GTG
 S Q E E V V S L L R S T K M E G T V

【 2 - G 】

1629 1638 1647 1656 1665 1674
 AGC CTT CTG GTC TTT CCG CAG GAA GAC CCC TTC CAC CCA ACG GAA CTG AAT GCA
 S L L V F R Q E D A F H P R E L N A
 1683 1692 1701 1710 1719 1728
 GAG CCA AGC CAG AAG CAG AAT CCA AAA GAA ACG AAA CCA GAA GAT GAG GAT AAT
 E P S Q M Q I P K E T K A E D E D I
 1737 1746 1755 1764 1773 1782
 GTT CTT ACA CCT GAT GGC ACC AAG GAA TTT CTG ACA TTT GAA GTC CCA CTT AAT
 V L T P D G T R E F L T F E V P L N
 1791 1800 1809 1818 1827 1836
 GAT TCA GGA TCT GCA GGC CTT GGT GTC AAT GAT AAC CCG TCA AAA GAG
 D S G S A G L G V S V K G N R S K E
 1845 1854 1863 1872 1881 1890
 AAC CAC GCA GAT TTG GGA ATC TTT GTC AAG TCC AAT AAT GGA GGA GCA GCA
 N H A D L G I F V K S I I N G G A

【 2 - H 】

1859 1908 1917 1926 1935 1944
TCT AAA GAT GGA AGG CTT GGG GTG AAT GAT CAA CTG ATA GCA GTA AAT GGA GAA
S K D C R L R V N D Q L I A V N G E
1953 1962 1971 1980 1989 1998
TCC CTG TTG GGC AAG ACA AAC CAA GAT GCC ATG GAA ACC CTA AGA AGG TCT ATG
S L L G R T N Q D A M E T L R R S M
2007 2016 2025 2034 2043
TCT ACT GAA GGC AAT AAA CGA GGA ATG ATC CAG CTT ATT GAT GCA AGG AGA ATA
S T E G N K R G M I Q L I V A R R I
2061 2070 2079 2088 2097 2106
AGC AAG TGC AAT GAG CTG AAG TCA CCT GGG AGC CCC CCT GSA COT GAG CTG CCC
S K C N E L K S P G S P F G F E L P
2115 2124 2133 2142 2151 2160
ATT GAA ACA GCG TTG GAT GAT AAG GAA CGA AGA ATT TCC GAT TCC CTC TAC AAT
I E T A L D D R E R R I S H S L Y S

【 2 - I 】

2169 2178 2187 2196 2205 2214
GGG AAT GAG GGG CTT GAT GAA TCG CCC AGC AGA AAT GCT GCC CTC AGT AGG ATA
G I E G L D E S P S R N A A L S R I
2223 2232 2241 2250 2259 2268
ATG GAT GAG TCA GGT AAA TAC GAG CTG TCC CCT ACA GTG AAT ATG CCC CAA GAT
M G E S G K Y Q L S E T V N M P Q D
2277 2286 2295 2304 2313 2322
GAC ACT GTC AAT ATA GAA GAT GAC AGG TTG CCA GTG CTT CCT CCA CAT CTC TCT
D T V I I E D D R L P V L P F H L S
2331 2340 2349 2358 2367 2376
GAG CAG ACC TCT TCC ACC TCC GAT GAT GTG GGG TTT GTG ACG GCA GAT GCT
D Q S S S S H D D V G F V T A D A
2385 2394 2403 2412 2421 2430
GGT ACT TGG GCC AAG CCA ATC AAT GAT TCA GCC GAC TGC TCT TTG AAT CCA
G T W A K A A I S D S A D C S L S P

【 2 - J 】

2439 2448 2457 2466 2475 2484
GAT GTT GAT CCA GTT CTT GCT TTT CAA GAA GAT GSA GGT CAG AAT ATG
D V D P V L A F Q R E G F G R Q S M
2493 2502 2511 2520 2529 2538
TCA GAA AAA CCG ACA AAG CAA TTT TCA GAT GCC AGT CAA TTG GAT TTC GTT AAA
S E K R T K Q F S D A S Q L D F V K
2547 2556 2565 2574 2583 2592
ACA CGA AAA TCA AAA ACC ATG GAT TTA GAT ATA GCT GAC GAG ACT AAA CTC AAT
T R K S K S M D L G I A D E T K L N
2601 2610 2619 2628 2637 2646
ACA GTG GAT GAC CAG AAA GCA GGT TCT CCC AGC AGA GAT GTG GGT CTT TCC CTG
T V D D Q K A G S P S R D V G P S L
2655 2664 2673 2682 2691 2700
GGT CTG AAG AAG TCA ACC TCG TTG GAG AAT CTG CAG ACC GCA GTT GCC GAG GTG
G L K K S S S L E S L Q T A V A E V

【 2 - K 】

2709 2718 2727 2736 2745 2754
ACT TTG AAT GGG GAT AAT CCT TTC GAT GGT CCA CCG CGG ATA ATC AGA GGC
T L N G D I P F H R P R P R I R G
2763 2772 2781 2790 2799 2808
AGG GGA TGC AAT GAG AGC TTC AGA GCT GCC ATC GAC AAA TCT TAT GAT AAA CCG
R G C N E S F R A A I D K S Y D K P
2817 2826 2835 2844 2853 2862
GCG GTA GAT GAT GAT GAA GGC ATG GAG ACC TTG GAA GAC GAC ACA GAA GAA
A V D D D D E G M E T L E B D T E E
2871 2880 2889 2898 2907 2916
ACT TCA ABA TCA GGG AGA GAG TCT GTA TCC ACA GCC AAT GAT CAG CCT TCC GAC
S S R S G R E S V S T A S D Q P S H
2925 2934 2943 2952 2961 2970
TCT CTG GAG ABA CAA ATG AAT GGA AAC CAA GAG AAA GGT GAT AAG ACT GAT AGA
S L E R Q M N G N Q E K G D K T D R

【 2 - L 】

2979 2988 2997 3006 3015 3024
AAA AAG GAT AAA ACT GGA AAA GAA AAG AAA GAT AGA GAT AAG GAG AAG GAT
K K D K T G K E K K K K K D R D K E K D
3033 3042 3051 3060 3069 3078
AAA ATG AAA GCC AAG AAG GGA ATG CTG AGG GGC TTG GGA GAC AAG TTC AGG TTT
K H K A K K G H L K G L G D M F R F
3087 3096 3105 3114 3123 3132
GGC AAA CAT CGA AAA GAT GAC AAG AAT GAG AAA ACG GGT AAA ATA AAA ATA CAG
G K H R K D D K I E K T G K I K I Q
3141 3150 3159 3168 3177 3186
GAA TCC TTT ACA TCA GAA GAG GAG AGG ATA CGA ATG AAG CAG GAG CAG GAG AGG
E S F T S E E R I R M K Q E Q E R
3195 3204 3213 3222 3231 3240
ATT CGA GCC AAA ACT CGA GAA TTT AGG GAA CGA CAA GCT CGA GAG GGT GAC TAT
I Q A K T R E F R E R Q A R E R D Y

【 2 - M 】

3249 3258 3267 3276 3285 3294
GCT GAA ATT CAA GAT TTT CAT CGG ACA TTT GGC TGT GAT GAT GAT TTA ATG TAT
A E I Q D F H R T F G C D D E L M Y
3303 3312 3321 3330 3339 3348
GGG GGA GTT TCT TCT TAT GAA GGT TCC ATG GCT CTC AAC GCT AGA CCT CAG AGC
G G V S S Y E G S M A L N A R P Q S
3357 3366 3375 3384 3393 3402
CCA CGA GAA GGG CAT ATG ATG GAT GCT TTG TAT GCC CAA GTC AAG AAG CCG CGG
P R E G H M M D A L Y A Q V K K P R
3411 3420 3429 3438 3447 3456
AAT TCC AAA CCC TCA CPT GTA GAC AGT AAC AGA TCA ACT CCT AGC AAT CAT GAT
N S K P S P V D S N R S T P S N H D
3465 3474 3483 3492 3501 3510
CGG ATA CAG CGT CTC AGG CAA GAA TTT CAG CAA GCA AAG CAA GAT GAA GAT GTA
R I Q R L R Q E F Q Q A K Q D E D V

【 2 - N 】

3519 3528 3537 3546 3555 3564
GAA GAT CGT CGG ACC TAT AGT TTT GAG CAA CCC TGG CCG AAC GCA CGG CCG
E D R R R T Y S F E O P W P N A R P
3573 3582 3591 3600 3609 3618
GGG ACG CAG AGC GGG CGA CAC TCG GTG TCC GTG GAG GTG CAG ATG CAG CCG CAG
A T Q S G R H S V S V E V Q M Q R Q
3627 3636 3645 3654 3663 3672
CGG CAG GAG GGC CAG AGC TCC CAG CAG GCC CAG CCG CAG TAC AGC TCT CAG
R Q E E R E S S Q Q A Q R Q Y S S L
3681 3690 3699 3708 3717 3726
CCT CGG CAA GGC AGC AAA AAT GCC AGC TCG TTC CAG GAC TCT TGG GAG CAG
P R Q S R K N A S S V S Q D S W E Q
3735 3744 3753 3762 3771 3780
AAC TAC TCC CCT CGG GAA GGC TTC CAG AGT GCC AAA GAG AAC CCC AGG TAC TCC
N Y S P G E G F Q S A K E N P R Y S

【 2 - O 】

3789 3798 3807 3816 3825 3834
AGC TAC CAA GGC TCC AGG AAC GGC TAC CAG GGA GAT GGC TTC AAC GCC AGG
S Y Q G S R N G Y L G G H G F N A R
3843 3852 3861 3870 3879 3888
GTC ATG CTG GAA ACT CAG GAG CTT CAG CAG GAA CAG AGG CCG AAG GAG CAG
V M L E T Q E L L R Q E Q R R K E Q
3897 3906 3915 3924 3933 3942
CAG ATG AAG AAG CAG CCT CCT TCC GAG GGG CCC AGC AAC TAT GAT TCG TAT AAG
Q M K K Q P P S E G P S N Y D S Y K
3951 3960 3969 3978 3987 3996
AAA GTC CAG GAC CCC ACT TAC GCC CCT CCG AAG GGG CCC TTC CCG CAA GAT GTG
K V Q D P S Y A P P K G P F R Q D V
4005 4014 4023 4032 4041 4050
CCC CCC CCT TCT CAG GTT CCG AGS CTG AAC AGR CTT CAG ACT CCT GAG AAA
P P S P S Q V A R L N R L Q T P E K

【 2 - P 】

4059 4068 4077 4086 4095 4104
 GGG AGG CCC TTC TAT TCC TGA GCA CGC AAA TAA CGS ATG CTT CMT GTC CGC CAA
 G R P F Y S
 4113 4122 4131 4140 4149 4158
 TAA AAG ACA TTT TCC TAT GAA GAC ATG TAT TTT GGG AGT TTT TTT AAA ACC TCG
 4167 4176 4185 4194 4203 4212
 ATG GFA CTA TGG AGT ATT TCT GFT GFT GAT ATC AGT GCC TTT AAG CGS TGT AGG
 4221 4230 4239 4248 4257 4266
 CAA ABA GGA GGG CCG TCA TAA TGT GAT TGG CAC TAT GTC TCA AGT GTC TGT TTC
 4275 4284 4293 4302 4311 4320
 AAG GAA GGA TTT CCC CCG ATG TGA CAA TCA TCT GTT TGA GGT GAT GAT ATG CTC
 4329 4338 4347 4356 4365 4374
 TGC GCA TCT CCA CAG TAC CAG GAA TCT CGS CCC TAC TCA TGA GTT GTC CGC GGC
 4383 4392 4401 4410 4419 4428
 TTG GTT GFA ACA TCC CTG CAC CAC TTG GAG TGA CAA ATT CAC CTG AAG TGG AGG
 4437 4446 4455 4464 4473 4482
 ATG ACG TGC GGC CCT GTT TCT CCC TCT AAG TTC TCT TAG CTA TGG GAT GAC ATC
 4491 4500 4509 4518 4527 4536
 TTA GTC TCT GGT GGA GAA AAG GTG GGC GAC ATA CAC CAA AAA TTG GGG CTT TCT

【 2 - Q 】

4545 4554 4563 4572 4581 4590
 GGT ACT TCA CAG CAC AGC CMT TTT TGG TAC TTT GTC ATC ACT GTS GTT TTC TCT
 4599 4608 4617 4626 4635 4644
 TTC CTT TCT CAG CTC TTT GTS ACC GGA GAG TCG GTC ATC CTA TTA CAG AAG CTA
 4653 4662 4671 4680 4689 4698
 AGC CMT AGT CCA ACA TTG TTT GGT CAC CAT GGG GGT CCT TTT GTA ACT GCC TTA
 4707 4716 4725 4734 4743 4752
 TGA CTC AAC ATT ACC AAT AAA GTG ATG ATC CUG GTC TGC GTT TAT ACA TAC GCT
 4761 4770 4779 4788 4797 4806
 TGT TCG GTC CTG TTC CAG ACA CCG GGG TTG AGT CAC CAC ACG TCT GTS TGG GGA
 4815 4824 4833 4842 4851 4860
 ACG TGG GAG ACA GGA GTG GCT CCT GCC GGG GGA TGG GGC TSC CAT TGG CCC
 4869 4878 4887 4896 4905 4914
 TGT GTC TAT CAT GAG GGG ABA GCT AAG AAA GAA ATT CTC CTA GGA AGA GCT CMT
 4923 4932 4941 4950 4959 4968
 GGC CCA GTA CAT CCT AGT AAT TAT TTT AAT TAG TTT TTG TTG TGA CAG CTT GTC
 4977 4986 4995 5004 5013 5022
 AAG AAG GGC ACA GAA TGG GAC AAG AAT AAA CCA GAC GAC AGT CAT TTT GAT CTG CTC

【 2 - R 】

5040 5049 5058 5067 5076
 TCT ACG GTT TTT CAA GTC AGA GGC AAT TGA TGC TTG TCT AAT GCA TCC ACA CAC
 5085 5094 5103 5112 5121 5130
 TGC ATG TCT GAC TGG CGA TGC CAC GCT CCT AAG TAG TTC TGC CMT GAA ACA TAA
 5139 5148 5157 5166 5175 5184
 AAG ACA AAG GAA AAG CCG TTA CAC ATC ACA CAG ABA ACA TTT TGG GAT CCC ACA
 5193 5202 5211 5220 5229 5238
 GCG GTG GCA GCA AGC TCA CTC TGC CFT CAG TAT TAG AAT GTS TTT GTS GGT
 5247 5256 5265 5274 5283 5292
 CTC GGS GAT CTC GAT GGC TCC CMT CTT TCA TTG TTC TGA ACA TCC TGT ATT
 5301 5310 5319 5328 5337 5346
 GFA AAC CMT GCG TGG GGT GCT AAA GTS CMT GTG AAT CCC GAT GTS GAA AAA GCT
 5355 5364 5373 5382 5391 5400
 GGA GGT GAA AGC TCA GCA TAC CAA GTA TTT ACT TTA AAA ACA GAA AAA AAG ACA
 5409 5418 5427 5436 5445 5454
 TGT AAG GAT ATG TCT ATT TTT TTT TTA TTG GCA CMT TGT AAT TTT TTG TTG ACT
 5463 5472 5481 5490 5499 5508
 TGT TTT TGG AAA TGA TGT GTC CAC ACA CFT ACC CPT GTC TCT TCT GCA TTT CTG

【 2 - S 】

5517 5526 5535 5544 5553 5562
 TGT CMT GGT TCT GTT TCT TAA TCA CFT GCG GCG GTG TCT AAG TGG TGT TAC CAG
 5571 5580 5589 5598 5607 5616
 TGT ACG CCG AGT GAC CTT GGA TGA CAG TGG CTC TTT CTC ACA GCC TCC CCT GAG
 5625 5634 5643 5652 5661 5670
 CTG TGA GAA ACA GCT TTC TCT GFA CMT ATG CAA CTC CTA AAT AAA GGC AAT TTT
 5679 5688
 CTT CCT GTT AAA AAA A A 3'

【 3 - A 】

1 MKVTVCFGRTGIVVVPCKEQLRVRVGEITQQA 1555118
 1 MKVTVCFGTRRVVPCGDDGHMKVFSLIQQA 2582063
 1 MKVTVCFGTRRVVPCGDDGHMKVFSLIQQA 93868778
 1 MKVTVCFGTRRVVPCGDDGHMKVFSLIQQA 98037915
 31 LQRYLKRTRREKGPYVVKIHLHLEYTDGGILD 1555118
 31 VTRYRKAIAKADPNYWIQVHRLEHGDGGILD 2582063
 31 VTRYRKAIAKADPNYWIQVHRLEHGDGGILD 93868778
 31 VTRYRKAIAKADPNYWIQVHRLEHGDGGILD 98037915
 61 PDDVLAJVVVEDKDKLJAVFVEQEPHLEHIES 1555118
 61 LDDILCDDVADDDKDLVAVFDEQDDPHHGGDG 2582063
 61 LDDILCDDVADDDKDLVAVFDEQDDPHHGGDG 93868778
 61 LDDILCDDVADDDKDLVAVFDEQDDPHHGGDG 98037915
 91 PISGNPARDQSSPDNRETVAAQ-LAJKPF-- 1555118
 91 TSSSTGTSQSEIFGSELGTNNVSAFQPYQ 2582063
 91 TSSSTGTSQSEIFGSELGTNNVSAFQPYQ 93868778
 91 TSSSTGTSQSEIFGSELGTNNVSAFQPYQ 98037915
 118 IGTSEIERTVPSAKLGLTLPVRRSSDFFV-- 1555118
 121 ATSEIERTVPSVLRANMPLHVRSSDFFALIG 2582063
 121 ATSEIERTVPSVLRANMPLHVRSSDFFALIG 93868778
 121 ATSEIERTVPSVLRANMPLHVRSSDFFALIG 98037915

【 3 - B 】

145 -----PGLPADTQPISASHPGGQS 1555118
 151 LSTSVSDSNPSSSEEP SRKNPTRWSTTAGFL 2582063
 151 LSTSVSDSNPSSSEEP SRKNPTRWSTTAGFL 93868778
 151 LSTSVSDSNPSSSEEP SRKNPTRWSTTAGFL 98037915
 164 KLVVPPD[S]T----- 1555118
 181 KONTAGSPKTCDRKKDENYRSLPRDPSNWS 2582063
 181 KONTAGSPKTCDRKKDENYRSLPRDPSNWS 93868778
 181 KONTAGSPKTCDRKK----- 98037915
 172 --QNLREDREVLNMGVOTEL--LTTSPRRTKD 1555118
 211 NOFORDNARSLSASHPMVGRWLEKQEQDE 2582063
 211 NOFORDNARSLSASHPMVGRWLEKQEQDE 93868778
 195 -----DE 1555118
 196 -----T[S]D 1555118
 241 DGTEDNSRVEPVGHADTGLEHHPNFSLD 2582063
 241 DGTEDNSRVEPVGHADTGLEHHPNFSLD 93868778
 197 DGTEDNSRVEPVGHADTGLEHHPNFSLD 98037915
 200 MTRTVEISGEGGGLGIHVVPF[S]LSL[S]GRLL 1555118
 271 MVKLVVEVPNDGGPLGIHVVPF-SARGGRTL 2582063
 271 MVKLVVEVPNDGGPLGIHVVPF-SARGGRTL 93868778
 277 MVKLVVEVPNDGGPLGIHVVPF-SARGGRTL 98037915

【 3 - C 】

230 GLFIRGIEDNSRSKRREGLFHEHNCIIVKIIN 1555118
 300 GLVKKRLEKGGKAEHENLFRHENDCIVRIND 2582063
 300 GLVKKRLEKGGKAEHENLFRHENDCIVRIND 93868778
 256 GLVKKRLEKGGKAEHENLFRHENDCIVRIND 98037915
 260 VDLVDRKTFQAQDVFRRQAMKSP[S]VLLHVL 1555118
 330 GDLRNRFRFEQAQHMFRRQAMRTPPIWFFHVVP 2582063
 330 GDLRNRFRFEQAQHMFRRQAMRTPPIWFFHVVP 93868778
 286 GDLRNRFRFEQAQHMFRRQAMRTPPIWFFHVVP 98037915
 290 P[Q]NREOYERKSVIGSLNIFGNN--DGLVK 1555118
 360 AANKEOYELSQ[Q]EKNNYSSRRFSPDSQYI 2582063
 360 AANKEOYELSQ[Q]EKNNYSSRRFSPDSQYI 93868778
 316 AANKEOYELSQ[Q]EKNNYSSRRFSPDSQYI 98037915
 316 TKVPPPVH[Q]ESGLK[Q]ANLTCGTDSPETDAS 1555118
 390 DNRSVNSAGLHTVQRAPKLNHHPPEQIDSH 2582063
 390 ANRSVANSAGLHTVQRAPKLNHHPPEQIDSH 93868778
 346 ANRSVANSAGLHTVQRAPKLNHHPPEQIDSH 98037915
 346 SLQ[Q]NKS[Q]KVRRLGGK[Q]SIS[Q]S--LSPLMG 1555118
 420 RLP[Q]SAH[Q]S[Q]KPP[Q]SA[Q]P[Q]N[Q]V[Q]S[Q]T[Q]S 2582063
 420 RLP[Q]SAH[Q]S[Q]KPP[Q]SA[Q]P[Q]N[Q]V[Q]S[Q]T[Q]S 93868778
 376 RLP[Q]SAH[Q]S[Q]KPP[Q]SA[Q]P[Q]N[Q]V[Q]S[Q]T[Q]S 98037915

【 3 - D 】

373 -FGSN[Q]NA[Q]K[Q]I[Q]H[Q]L[Q]K[Q]E[Q]L[Q]G[Q]F[Q]TVVTRD 1555118
 450 GYNTKIRVGGKRLNIQLKKGTEGLGFSITSRD 2582063
 450 VYNTKIRVGGKRLNIQLKKGTEGLGFSITSRD 93868778
 406 GYNTKIRVGGKRLNIQLKKGTEGLGFSITSRD 98037915
 402 S[Q]H[Q]G[Q]P[Q]I[Q]V[Q]K[Q]N[Q]I[Q]L[Q]P[Q]K[Q]A[Q]I[Q]D[Q]G[Q]R[Q]L[Q]S[Q]G[Q]D 1555118
 480 VTI[Q]G[Q]S[Q]A[Q]P[Q]I[Q]Y[Q]V[Q]K[Q]N[Q]I[Q]L[Q]P[Q]K[Q]A[Q]I[Q]D[Q]G[Q]R[Q]L[Q]S[Q]G[Q]D 2582063
 480 VTI[Q]G[Q]S[Q]A[Q]P[Q]I[Q]Y[Q]V[Q]K[Q]N[Q]I[Q]L[Q]P[Q]K[Q]A[Q]I[Q]D[Q]G[Q]R[Q]L[Q]S[Q]G[Q]D 93868778
 436 VTI[Q]G[Q]S[Q]A[Q]P[Q]I[Q]Y[Q]V[Q]K[Q]N[Q]I[Q]L[Q]P[Q]K[Q]A[Q]I[Q]D[Q]G[Q]R[Q]L[Q]S[Q]G[Q]D 98037915
 432 R[Q]L[Q]E[Q]V[Q]N[Q]G[Q]R[Q]V[Q]T[Q]G[Q]R[Q]T[Q]O[Q]E[Q]R[Q]L[Q]V[Q]A[Q]M[Q]L[Q]R[Q]S[Q]T[Q]K[Q]G[Q]E[Q]T 1555118
 510 R[Q]L[Q]E[Q]V[Q]N[Q]G[Q]R[Q]V[Q]T[Q]G[Q]R[Q]T[Q]O[Q]E[Q]R[Q]L[Q]V[Q]A[Q]M[Q]L[Q]R[Q]S[Q]T[Q]K[Q]G[Q]E[Q]T 2582063
 510 R[Q]L[Q]E[Q]V[Q]N[Q]G[Q]R[Q]V[Q]T[Q]G[Q]R[Q]T[Q]O[Q]E[Q]R[Q]L[Q]V[Q]A[Q]M[Q]L[Q]R[Q]S[Q]T[Q]K[Q]G[Q]E[Q]T 93868778
 466 R[Q]L[Q]E[Q]V[Q]N[Q]G[Q]R[Q]V[Q]T[Q]G[Q]R[Q]T[Q]O[Q]E[Q]R[Q]L[Q]V[Q]A[Q]M[Q]L[Q]R[Q]S[Q]T[Q]K[Q]G[Q]E[Q]T 98037915
 462 A[Q]S[Q]L[Q]V[Q]A[Q]R[Q]Q[Q]E[Q]H[Q]F[Q]L[Q]P[Q]R[Q]E[Q]L[Q]K[Q]E[Q]P[Q]D[Q]C[Q]A[Q]L[Q]S[Q]L[Q]E[Q]T 1555118
 540 V[Q]S[Q]L[Q]V[Q]F[Q]R[Q]Q[Q]E[Q]A[Q]F[Q]H[Q]P[Q]R[Q]E[Q]L[Q]N[Q]A[Q]E[Q]P[Q]S[Q]Q[Q]M[Q]I[Q]P[Q]K[Q]E[Q]T 2582063
 540 V[Q]S[Q]L[Q]V[Q]F[Q]R[Q]Q[Q]E[Q]A[Q]F[Q]H[Q]P[Q]R[Q]E[Q]L[Q]N[Q]A[Q]E[Q]P[Q]S[Q]Q[Q]M[Q]I[Q]P[Q]K[Q]E[Q]T 93868778
 496 V[Q]S[Q]L[Q]V[Q]F[Q]R[Q]Q[Q]E[Q]A[Q]F[Q]H[Q]P[Q]R[Q]E[Q]L[Q]N[Q]A[Q]E[Q]P[Q]S[Q]Q[Q]M[Q]I[Q]P[Q]K[Q]E[Q]T 98037915
 492 SEQ-----LTFPHLP[Q]LNDSSGSA 1555118
 570 KAEDEDIVLTPDGTREFLTFEVP[Q]LNDSSGSA 2582063
 570 KAEDEDIVLTPDGTREFLTFEVP[Q]LNDSSGSA 93868778
 513 KAEDEDIVLTPDGTREFLTFEVP[Q]LNDSSGSA 98037915

【 3 - E 】

508 GLGVSLKGNRSKENHADLGLGFIHGGAI 1555118
 600 GLGVSVKGNRSKENHADLGFVKSIINGGA 2582063
 600 GLGVSVKGNRSKENHADLGFVKSIINGGA 93868778
 543 GLGVSVKGNRSKENHADLGFVKSIINGGA 98037915

538 AFKDDGRLRNDQLIAVNGESLLGKSNHJAM 1555118
 630 ASKDGRLRVNDQLIAVNGESLLGKSNHJAM 2582063
 630 ASKDGRLRVNDQLIAVNGESLLGKSNHJAM 93868778
 573 ASKDGRLRVNDQLIAVNGESLLGKSNHJAM 98037915

568 ETLRRPMSMEGNIRGMIQVILRR- 1555118
 660 ETLRRSMTTEGKRCMIQVILRR- 2582063
 660 ETLRRSMTTEGKRCMIQVILRR- 93868778
 603 ETLRRSMTTEGKRCMIQVILRR- 98037915

592 LKSPGSPFPELPIETALDDRRRSHSLY 1555118
 690 LKSPGSPFPELPIETALDDRRRSHSLY 2582063
 690 LKSPGSPFPELPIETALDDRRRSHSLY 93868778
 633 LKSPGSPFPELPIETALDDRRRSHSLY 98037915

612 NCQNAVTRRRNDNSILHPLCTCSPQDKQK 1555118
 720 SGI EGLDESFSRRAALSRIIMGSKQLSP 2582063
 720 SGI EGLDESFSRRAALSRIIMGSKQLSP 93868778
 663 SGI EGLDESFSRRAALSRIIMGSKQLSP 98037915

【 3 - F 】

642 GLLLPNIGWABSE- 1555118
 750 TVNMPQDDTVI EDDRLLVLPPLSDQSSS 2582063
 750 TVNMPQDDTVI EDDRLLVLPPLSDQSSS 93868778
 690 TVNMPQDDTVI EDDRLLVLPPLSDQSSS 98037915

660 TPIHSA(L)GLLEH- 1555118
 780 SSHDDVGFVTAAGTWAKAATISDSADCSLS 2582063
 780 SSHDDVGFVTAAGTWAKAATISDSADCSLS 93868778
 720 SSHDDVGFVTAAGTWAKAATISDSADCSLS 98037915

676 SG(V)SAVYTPDOHINFR(S)VTPAR- 1555118
 810 PDVDPVLAFOHGFQCSMSEKRTKQFSDA 2582063
 810 PDVDPVLAFOHGFQCSMSEKRTKQFSDA 93868778
 750 PDVDPVLAFOHGFQCSMSEKRTKQFSDA 98037915

699 QPESINLKA(S)KSMDDL-VFDE(S)KVHSLJAG 1555118
 840 SQLDFVKTTRKSKSMDDLGADETKLNTVDDQ 2582063
 840 SQLDFVKTTRKSKSMDDLGADETKLNTVDDQ 93868778
 767 SQLDFVKTTRKSKSMDDLGADETKLNTVDDQ 98037915

727 KISE(S)PSK(D)F(C)F(L)GLKSSSLES(L)QTAVAE 1555118
 870 KAGSPSRDVGPSLGLKSSSLES(L)QTAVAE 2582063
 870 KAGSPSRDVGPSLGLKSSSLES(L)QTAVAE 93868778
 780 KAGSPSRDVGPSLGLKSSSLES(L)QTAVAE 98037915

【 3 - G 】

757 VR-KNDLPPFHRPPHMYRGRGCNESFRAAI 1555118
 900 VTLNGDIPFHRPPRIIRGRGCNESFRAAI 2582063
 900 VTLNGDIPFHRPPRIIRGRGCNESFRAAI 93868778
 810 VTLNGDIPFHRPPRIIRGRGCNESFRAAI 98037915

786 DKSYD(G)PE-EIEAD(G)L- 1555118
 930 DKSYDKPAVDDDDDEGMETLEEDTEESSRS 2582063
 930 DKSYDKPAVDDDDDEGMETLEEDTEESSRS 93868778
 840 DKSYDKPAVDDDDDEGMETLEEDTEESSRS 98037915

809 OGALNCE(S)A.PQGN(S)LEED(N) 1555118
 960 RESVTSADQPSHLSLERMNCDEK(R)DKAE 2582063
 960 RESVTSADQPSHLSLERMNCDEK(R)DKAE 93868778
 870 RESVTSADQPSHLSLERMNCDEK(R)DKAE 98037915

829 NRAR(K)K(R)E(K)K(E)K(L)K(V)K(E) 1555118
 990 RKKDKTKGKRRKDRKDKRKKKGMMLK 2582063
 990 RKKDKTKGKRRKDRKDKRKKKGMMLK 93868778
 900 RKKDKTKGKRRKDRKDKRKKKGMMLK 98037915

854 K(R)K(E)E.NE.D- 1555118
 1020 LGDMFRFGKRRKDDKIEKTKIKIQESFTS 2582063
 1020 LGDMFRFGKRRKDDKIEKTKIKIQESFTS 93868778
 930 LGDMFRFGKRRKDDKIEKTKIKIQESFTS 98037915

【 3 - H 】

863 EFERIRMKOQEQERTQAKTRERFRERQARERD 1555118
 1050 EERIRRMKQEQERTQAKTRERFRERQARERD 2582063
 960 EFERIRMKOQEQERTQAKTRERFRERQARERD 93868778
 960 EFERIRMKOQEQERTQAKTRERFRERQARERD 98037915

872 FGAM- 1555118
 1080 YAEIQDFHRTFGCDELMLYGGVSSYEGSMA 2582063
 1080 YAEIQDFHRTFGCDELMLYGGVSSYEGSMA 93868778
 990 YAEIQDFHRTFGCDELMLYGGVSSYEGSMA 98037915

882 LKAK- 1555118
 1110 LNARPPQSPREGHMMDALYAQQVKKPRMSKPS 2582063
 1110 LNARPPQSPREGHMMDALYAQQVKKPRMSKPS 93868778
 1020 LNARPPQSPREGHMMDALYAQQVKKPRMSKPS 98037915

899 PVR- 1555118
 1140 PVD(S)NRSTPSNHDRIORLROEFQAKQDEDED 2582063
 1140 PVD(S)NRSTPSNHDRIORLROEFQAKQDEDED 93868778
 1050 PVD(S)NRSTPSNHDRIORLROEFQAKQDEDED 98037915

905 VEDRRRTYSFEQEFENARPA.TOSGRHSVSV 1555118
 1170 VEDRRRTYSFEQEFENARPA.TOSGRHSVSV 2582063
 1168 VEDRRRTYSFEQEFENARPA.TOSGRHSVSV 93868778
 1080 VEDRRRTYSFEQEFENARPA.TOSGRHSVSV 98037915

【 3 - I 】

905 ----- 1555118
 1200 EVQMQRQRQEERFSQQAAQRQYSSLPQRSR 2582063
 1198 EVQVQRQRQEERFSQQAAQRQYSSLPQRSR 93868778
 1110 EVQMQRQRQEERFSQQAAQRQYSSLPQRSR 98037915

905 ----- 1555118
 1230 KNASSVSQDSWECNYSPEGGFOSAKENPRY 2582063
 1228 KNASSVSQDSWECNYSPEGGFOSAKENPRY 93868778
 1140 KNASSVSQDSWECNYSPEGGFOSAKENPRY 98037915

905 ----- 1555118
 1260 SYYQSSRKGYLGGHGFNARVM-LETOELLR 2582063
 1258 SYYQSSRNGYLGGHGFNARVM-LETOELLR 93868778
 1170 SYYQSSRNGYLGGHGFNARVM-LETOELLR 98037915

925 ----- G 1555118
 1289 QEORRKEOQMKKQPSEGFPSNYDSYKKVQD 2582063
 1287 QEORRKEOQMKKQPSEGFPSNYDSYKKVQD 93868778
 1199 QEORRKEOQMKKQPSEGFPSNYDSYKKVQD 98037915

926 S Y S D D R T ----- 1555118
 1319 P S Y A F K G P F R Q D V P P S Q V A R L N R L Q T P 2582063
 1306 ----- R G P F R Q D V P P S Q V A R L N R L Q T P 93868778
 1229 P S Y A P K G P F R Q D V P P S Q V A R L N R L Q T P 98037915

【 3 - J 】

934 ----- L S 1555118
 1349 E K G R P P Y S 2582063
 1330 E K G R P P Y S 93868778
 1259 E K G R P P Y S 98037915

【国際公開パンフレット】

(H)60213741551



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

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
18 July 2002 (18.07.2002)

PCT

(10) International Publication Number
WO 02/055706 A2

- (51) International Patent Classification: C12N 15/12, C07K 1447, A61K 31/088, C12Q 1/68, C12N 5/10, A61K 38/17, C07K 16/18, C01N 33/68, A61K 39/395
- (52) International Application Number: PCT/US02/00647
- (22) International Filing Date: 8 January 2002 (08.01.2002)
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data: 9 January 2001 (09.01.2001) US 09/757,781
- (71) Applicant (for all designated States except US): INCYTE GENOMICS, INC. [US/US]; 3160 Porter Drive, Palo Alto, CA 94304 (US).
- (72) Inventors; and
- (73) Inventors/Applicants (for US only): REDDY, Roopa [IN/US]; 1233 W. McKelvey Avenue #3, Sunnyvale, CA 94086 (US). TANG, Y., Tom [US/US]; 4230 Ranwick Court, San Jose, CA 95118 (US). BAUGHN, Mariah, R. [US/US]; 14244 Santiago Road, San Leandro, CA 94577 (US). KRASNOW, Randi [US/US]; 817 Santa Fe Avenue, Stanford, CA 94305 (US).
- (74) Agents: STREETER, David, G. et al.; Incyte Genomics, Inc., 3160 Porter Drive, Palo Alto, CA 94304 (US).
- (81) Designated States (national): AU, AG, AL, AM, AX, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GR, GM, HR, HU, ID, IL, IN, IS, JP, KR, KG, KP, KZ, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.
- (84) Designated States (regional): ARIPO patent (GI, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, UZ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

[Continued on next page]



WO 02/055706 A2

(54) Title: ASIP-RELATED PROTEINS

```

          9       18       27       36       45       54
5'  A  A G G  G G C  G C T  G C C  G C G  A G C  C T C  C G G  G C C  T C A  G G G  T G T  T C C  G G G  G A G  C G G  C G C
          63       72       81       90       99       108
C C C  G G G  T C T  C T G  G G C  C C A  C C C  G C C  C C G  G G C  G T C  C T C  C G A  G A G  T G G  G G G  C T G  C G C
          117      126      135      144      153      162
C C G  C G G  G G T  C A G  A G A  C C T  G T T  C G G  C C C  G G C  C C G  G G G  T G G  T C G  C C G  G G G  C C C  A G G
          171      180      189      198      207      216
A T G  A A A  G T G  A C C  G T G  T G C  T T C  G G C  A G G  A C G  G G C  A T C  G T G  G T G  C C C  T G C  A A G  G A G
          225      234      243      252      261      270
M  K  V  T  V  C  F  G  R  T  G  I  V  V  P  C  K  E
          279      288      297      306      315      324
A C C  C G G  G A G  A A G  G G T  C C T  G G T  T A C  T G G  G T G  A A G  A T T  C A T  C A C  T T A  G A A  T A T  A C A
          T  R  E  K  G  P  G  Y  W  V  K  I  H  H  L  E  Y  T

```

(57) Abstract: The invention provides cDNAs which encode ASIP-related proteins. It also provides for the use of the cDNAs, fragments, recombinants, and variants thereof and of the encoded proteins, portions thereof and antibodies thereto for diagnosis and treatment of cancer, particularly bladder transitional cell carcinoma. The invention additionally provides expression vectors and host cells for the production of the protein and a transgenic model system.

WO 02/055706 A2 

Published:

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

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

WO 02/055706

PCT/US02/00647

ASIP-RELATED PROTEINS

TECHNICAL FIELD

This invention relates to cDNAs which encode ASIP-related proteins and to the use of the
5 cDNAs and the encoded proteins in the diagnosis and treatment of cancer, particularly bladder
transitional cell carcinoma.

BACKGROUND OF THE INVENTION

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

Protein kinase C (PKC) plays roles in intracellular signaling through lipid-derived second
messengers (Nishizuka (1995) *FASEB J* 9:484-496). Different isoforms of PKC have distinct
tissue and subcellular distributions, show differential responses to lipids and calcium, and may
20 serve different physiological functions. The PKC isoforms are divided into three classes,
conventional PKC, novel PKC, and atypical PKC. The atypical PKC (aPKC) isoforms have been
shown to play roles in *Xenopus* oocyte maturation (Dominguez *et al.* (1992) *Mol Cell Biol*
12:3776-3783), proliferation and survival of fibroblasts (Berra *et al.* (1993) *Cell* 74:555-563),
differentiation of PC12 and leukemic cells (Wooten *et al.* (1994) *Cell Growth Differ* 5:395-403;
25 Ways *et al.* (1994) *Cell Growth Differ* 5:1195-1203), activation of mitogen-activated protein
kinase and gene expression (Berra *et al.* (1995) *EMBO J* 14:6157-6163; Lozano *et al.* (1994) *J*
Biol Chem 269:19200-19202), insulin-induced glucose uptake and translocation of the glucose
transporter GLUT4 to the plasma membrane (Kotani *et al.* (1998) *Mol Cell Biol* 18:6971-6982),
and the establishment and/or maintenance of cell polarity (Brazil and Hemmings (2000) *Curr Biol*
30 10:R592-594). Regulation by aPKC involves protein-protein interactions with other signaling
molecules and aPKCs are known to interact in a number of different protein complexes (Moscat
and Diaz-Meco (2000) *EMBO Reports* 1:399-403). The recruitment of aPKC into different
complexes involves specific scaffold or adaptor proteins such as MyD88, TRADD, and p62.

One such adaptor protein may be atypical protein kinase C isotype specific interacting
35 protein (ASIP), which binds to aPKCs including PKC ζ and PKC λ . (Izumi *et al.* (1998) *J Cell Biol*

WO 02/055706

PCT/US02/06647

143:95-106). Rat ASIP is 1337 amino acids in length and contains 3 PDZ domains that may
5 mediate protein-protein interactions involved in the assembly of signaling complexes. The region
of rat ASIP from amino acids 712-936 contains the binding site for aPKC. The sequence of ASIP
is similar to that of the *Caenorhabditis elegans* protein, PAR-3, which is required for asymmetrical
cell division in worm embryos. ASIP may also have a role in establishing and maintaining cell
polarity in mammalian cells. Immunofluorescence microscopy studies show that ASIP and aPKCs
10 colocalize at tight junctions and adherens junctions in epithelial cells where they may be involved
in the formation and function of junctional complexes. Tight junctions maintain cellular
asymmetry by providing a barrier to intramembrane diffusion between apical and basolateral
membrane components as well as to diffusion between cells in an epithelial sheet. In addition,
ASIP may have other roles in cellular signaling involving aPKC. For example, overexpression of
15 ASIP in adipocytes inhibits insulin stimulation of glucose uptake and translocation of the glucose
transporter GLUT4 (Kotani *et al.* (2000) *J Biol Chem* 275:26390-26395). In the presence of
excess ASIP, the amounts of PKC α decrease in the cytosolic fraction and increase in the low
density microsomal and plasma membrane fractions. ASIP perturbs the subcellular distribution of
PKC α and interferes with the function of PKC α in glucose uptake, possibly because of its direct
20 interaction with PKC α .

Regulation of cell proliferation by aPKC involves epidermal growth factor (EGF). EGF
promotes proliferation and differentiation of mesenchymal and epithelial cells. It is a mitogen for
fibroblasts, epithelial and endothelial cells, induces epithelial development, and promotes
angiogenesis (Kim *et al.* (1999) *Histol Histopathol* 14:1175-1182 and Putz *et al.* (1999) *Cancer*
25 *Res* 59:227-233). EGF is highly expressed in breast carcinoma cells and promotes tumor
progression (Artagaveytia *et al.* (1997) *J Steroid Biochem Mol Biol* 60:221-228). EGF binds to
its receptor, the protein tyrosine kinase EGF receptor (EGFR), which is also known as erbB2
(Carpenter (2000) *Bioessays* 22:697-707). Ligation of EGF to EGFR results in the activation of
the tyrosine kinase domain of EGFR and the phosphorylation of multiple substrates. EGF may
30 regulate multimeric signaling complexes associated with aPKCs. In cells stimulated with EGF,
PKC α is activated, shows an increased level of phosphorylation, and increased concentrations in
the cytosol (Akimoto *et al.* (1996) *EMBO J* 15:788-798). EGF also influences the activity of
PKC ζ . In response to EGF, PKC ζ phosphorylates and activates p70 S6 kinase, a regulator of cell
proliferation (Romanelli *et al.* (1999) *Mol Cell Biol* 19:2921-2928).

35 The discovery of cDNAs encoding ASIP-related proteins satisfies a need in the art by
providing compositions which are useful in the diagnosis and treatment of cancer, particularly
bladder transitional cell carcinoma.

SUMMARY OF THE INVENTION

WO 02/055706

PCT/US02/00647

The invention is based on the discovery of cDNAs which encode ARP-related proteins
5 (ARP) which are useful in the diagnosis and treatment of cancer, particularly bladder transitional
cell carcinoma.

The invention provides an isolated cDNA or a fragment thereof encoding a protein or a
portion thereof selected from the group consisting of amino acid sequences of SEQ ID NO:1
(ARP-1) and SEQ ID NO:2 (ARP-2), a variant having at least at least 95% identity to the amino
10 acid sequences of SEQ ID NO:1 or SEQ ID NO:2, an antigenic epitope of SEQ ID NO:1 or SEQ
ID NO:2, and a biologically active portion of SEQ ID NO:1 or SEQ ID NO:2. The invention also
provides an isolated cDNA or the complement thereof selected from the group consisting of a
nucleic acid sequence of SEQ ID NO:3 and SEQ ID NO:20, a fragment of SEQ ID NO:3 selected
from SEQ ID NOs:4-11 or a fragment of SEQ ID NO:20 selected from SEQ ID NOs:21-39, and a
15 variant of SEQ ID NO:3 selected from SEQ ID NOs:12-19 or a variant of SEQ ID NO:20 selected
from SEQ ID NOs:40-56. The invention additionally provides a composition, a substrate, and a
probe comprising the cDNA, or the complement of the cDNA, encoding ARP-1 or ARP-2. The
invention further provides a vector containing the cDNA, a host cell containing the vector and a
method for using the cDNA to make ARP-1 or ARP-2. The invention still further provides a
20 transgenic cell line or organism comprising the vector containing the cDNA encoding ARP. The
invention additionally provides a fragment or the complement thereof selected from the group
consisting of SEQ ID NOs:3-56. 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, screening, and purification. In a further
25 aspect, the probe is a single stranded complementary RNA or DNA molecule.

The invention provides a method for using a cDNA to detect the differential expression of
a nucleic acid in a sample comprising hybridizing a probe to the nucleic acids, thereby forming
hybridization complexes and comparing hybridization complex formation with a standard,
wherein the comparison indicates the differential expression of the cDNA in the sample. In one
30 aspect, the method of detection further comprises amplifying the nucleic acids of the sample prior
to hybridization. In another aspect, the method showing differential expression of the cDNA is
used to diagnose cancer, particularly bladder transitional cell carcinoma. In another aspect, the
cDNA or a fragment or a complement thereof may comprise an element on an array.

The invention additionally provides a method for using a cDNA or a fragment or a
35 complement thereof to screen a library or plurality of molecules or compounds to identify at least
one ligand which specifically binds the cDNA, the method comprising combining the cDNA with
the molecules or compounds under conditions allowing specific binding, and detecting specific
binding to the cDNA, thereby identifying a ligand which specifically binds the cDNA. In one
aspect, the molecules or compounds are selected from aptamers, DNA molecules, RNA

WO 02/055706

PCT/US02/00647

molecules, peptide nucleic acids, artificial chromosome constructions, peptides, transcription factors, repressors, and regulatory molecules.

5 The invention provides a purified protein or a portion thereof selected from the group consisting of an amino acid sequence of SEQ ID NO:1 or SEQ ID NO:2, a variant having at least 95% identity to the amino acid sequences of SEQ ID NO:1 or SEQ ID NO:2, an antigenic epitope of SEQ ID NO:1 or SEQ ID NO:2, and a biologically active portion of SEQ ID NO:1 or SEQ ID
10 NO:2. The invention also provides a composition comprising the purified protein or a portion thereof in conjunction with a pharmaceutical carrier. The invention further provides a method of using the ARP-1 or ARP-2 to treat a subject with cancer, particularly bladder transitional cell carcinoma comprising administering to a patient in need of such treatment the composition containing the purified protein. The invention still further provides a method for using a protein
15 to screen a library or a plurality of molecules or compounds to identify at least one ligand, the method comprising combining the protein with the molecules or compounds under conditions to allow specific binding and detecting specific binding, thereby identifying a ligand which specifically binds the protein. In one aspect, the molecules or compounds are selected from DNA
20 molecules, RNA molecules, peptide nucleic acids, peptides, proteins, mimetics, agonists, antagonists, antibodies, immunoglobulins, inhibitors, and drugs. In another aspect, the ligand is used to treat a subject with cancer, particularly bladder transitional cell carcinoma.

The invention provides a method of using a protein to screen a subject sample for antibodies which specifically bind the protein comprising isolating antibodies from the subject
25 sample, contacting the isolated antibodies with the protein under conditions that allow specific binding, dissociating the antibody from the bound-protein, and comparing the quantity of antibody with known standards, wherein the presence or quantity of antibody is diagnostic of cancer, particularly bladder transitional cell carcinoma.

The invention also provides a method of using a protein to prepare and purify antibodies comprising immunizing a animal with the protein under conditions to elicit an antibody response,
30 isolating animal antibodies, attaching the protein to a substrate, contacting the substrate with isolated antibodies under conditions to allow specific binding to the protein, dissociating the antibodies from the protein, thereby obtaining purified antibodies.

The invention provides a purified antibody which binds specifically to a protein which is expressed in cancer, particularly bladder transitional cell carcinoma. The invention also provides
35 a method of using an antibody to diagnose cancer, particularly bladder transitional cell carcinoma comprising combining the antibody comparing the quantity of bound antibody to known standards, thereby establishing the presence of cancer, particularly bladder transitional cell carcinoma. The invention further provides a method of using an antibody to treat cancer, particularly bladder transitional cell carcinoma comprising administering to a patient in need of

WO 02/055706

PCT/U502/00647

such treatment a pharmaceutical composition comprising the purified antibody;

5 The invention provides a method for inserting a marker gene into the genomic DNA of a mammal to disrupt the expression of the endogenous polynucleotide. The invention also provides a method for using a cDNA to produce a mammalian model system, the method comprising constructing a vector containing the cDNA selected from SEQ ID NOs:3-56, transforming the vector into an embryonic stem cell, selecting a transformed embryonic stem, microinjecting the
10 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 cDNA in its germ line, and breeding the chimeric mammal to produce a homozygous, mammalian model system.

BRIEF DESCRIPTION OF THE FIGURES AND TABLE

15 Figures 1A, 1B, 1C, 1D, 1E, 1F, 1G, 1H, 1I, 1J and 1K show the ARP-1 (SEQ ID NO:1) encoded by the cDNA (SEQ ID NO:3). The translation was produced using MACDNASIS PRO software (Hitachi Software Engineering, South San Francisco CA).

Figures 2A, 2B, 2C, 2D, 2E, 2F, 2G, 2H, 2I, 2J, 2K, 2L, 2M, 2N, 2O, 2P, 2Q, 2R and 2S
20 show the ARP-2 (SEQ ID NO:2) encoded by the cDNA (SEQ ID NO:20). The translation was produced using MACDNASIS PRO software (Hitachi Software Engineering).

Figures 3A, 3B, 3C, 3D, 3E, 3F, 3G, 3H, 3I and 3J demonstrate the conserved chemical and structural similarities among the sequences and domains of ARP-1 (1555118; SEQ ID NO:1), ARP-2 (2582063; SEQ ID NO:2), rat ASIP (g3868778; SEQ ID NO:62), and human ASIP (g8037915; SEQ ID NO:63). The translation was produced using the MEGALIGN program of
25 LASERGENE software (DNASTAR, Madison WI).

Tables 1 and 2 show the northern analysis for ARP-1 and ARP-2 produced using the LIFESEQ Gold database (Incyte Genomics, Palo Alto CA). In Table 1, the first column presents the tissue categories; the second column, the total number of clones in the tissue category; the third column, the ratio of the number of libraries in which at least one transcript was found to the total number of libraries; the fourth column, absolute clone abundance of the transcript; and the
30 fifth column, percent abundance of the transcript. Table 2 shows expression of ARP in bladder tissue. The first column lists the library name, the second column, the number of clones sequenced for that library; the third column, the description of the tissue from which the library was derived; the fourth column, the absolute abundance of the transcript; and the fifth column, the
35 percent abundance of the transcript.

Table 3 shows the differential expression of ARP-2 in human BT20 breast carcinoma cells treated with EGF compared to untreated cells as determined by microarray analysis. Column 1 lists the mean differential expression (DE) values presented as \log_2 DE (treated cells/untreated cells). Column 2 lists the untreated control samples labeled with fluorescent green dye Cy3.

WO 02/055706

PCT/US02/00647

Column 3 lists the treatment for samples labeled with fluorescent red dye Cy5.

5

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 and is not intended to limit the scope of the present invention which will be limited only by the appended claims. As used herein, the singular forms "a", "an", and "the" include plural reference unless the context clearly dictates otherwise. For example, a reference to "a host cell" includes a plurality of such host cells known to those skilled in the art.

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

20 Definitions

"ARP" refers to a substantially purified protein obtained from any mammalian species, including bovine, canine, murine, ovine, porcine, rodent, simian, and preferably the human species, and from any source, whether natural, synthetic, semi-synthetic, or recombinant.

"Array" refers to an ordered arrangement of at least two cDNAs on a substrate. At least one of the cDNAs represents a control or standard sequence, and the other, a cDNA of diagnostic interest. The arrangement of from about two to about 40,000 cDNAs on the substrate assures that the size and signal intensity of each labeled hybridization complex formed between a cDNA and a sample nucleic acid is individually distinguishable.

The "complement" of a cDNA of the Sequence Listing refers to a nucleic acid molecule which is completely complementary over its full length and which will hybridize to the cDNA or an mRNA under conditions of high stringency.

"cDNA" refers to an isolated polynucleotide, nucleic acid molecule, or any fragment or complement thereof. It may have originated recombinantly or synthetically, be double-stranded or single-stranded, represent coding and/or noncoding 5' and 3' sequence.

The phrase "cDNA encoding a protein" refers to a nucleic acid sequence that closely aligns with sequences which encode conserved regions, motifs or domains that were identified by employing analyses well known in the art. These analyses include BLAST (Basic Local Alignment Search Tool; Altschul (1993) J Mol Evol 36: 290-300; Altschul et al. (1990) J Mol Biol 215:403-410) which provides identity within the conserved region.

WO 02/055706

PCT/US02/00647

"Derivative" refers to a cDNA or a protein that has been subjected to chemical modification. Derivatization of a cDNA can involve substitution of a nontraditional base such as queosine or of an analog such as hypoxanthine. These substitutions are well known in the art. Derivatization of a protein involves the replacement of a hydrogen by an acetyl, acyl, alkyl, amino, formyl, or morpholino group. Derivative molecules retain the biological activities of the naturally occurring molecules but may confer advantages such as longer lifespan or enhanced activity.

"Differential expression" refers to an increased, upregulated or present, or decreased, downregulated or absent, gene expression as detected by the absence, presence, or at least two-fold changes in the amount of transcribed messenger RNA or translated protein in a sample.

"Disorder" refers to conditions, diseases or syndromes in which the cDNAs and ARP are differentially expressed such as cancer, particularly bladder transitional cell carcinoma.

"Fragment" refers to a chain of consecutive nucleotides from about 200 to about 700 base pairs in length. Fragments may be used in PCR or hybridization technologies to identify related nucleic acid molecules and in binding assays to screen for a ligand. Nucleic acids and their ligands identified in this manner are useful as therapeutics to regulate replication, transcription or translation.

A "hybridization complex" is formed between a cDNA and a nucleic acid of a sample when the purines of one molecule hydrogen bond with the pyrimidines of the complementary molecule, e.g., 5'-A-G-T-C-3' base pairs with 3'-T-C-A-G-5'. The degree of complementarity and the use of nucleotide analogs affect the efficiency and stringency of hybridization reactions.

"Ligand" refers to any agent, molecule, or compound which will bind specifically to a complementary site on a cDNA molecule or polynucleotide, or to an epitope or a protein. Such ligands stabilize or modulate the activity of polynucleotides or proteins and may be composed of inorganic or organic substances including nucleic acids, proteins, carbohydrates, fats, and lipids.

"Oligonucleotide" refers to a single stranded molecule from about 18 to about 60 nucleotides in length which may be used in hybridization or amplification technologies or in regulation of replication, transcription or translation. Substantially equivalent terms are amplicon, primer, and oligomer.

"Portion" refers to any part of a protein used for any purpose; but especially, to an epitope for the screening of ligands or for the production of antibodies.

"Post-translational modification" of a protein can involve lipidation, glycosylation, phosphorylation, acetylation, racemization, proteolytic cleavage, and the like. These processes may occur synthetically or biochemically. Biochemical modifications will vary by cellular location, cell type, pH, enzymatic milieu, and the like.

"Probe" refers to a cDNA that hybridizes to at least one nucleic acid in a sample. Where

WO 02/055706

PCT/US02/00647

5 targets are single stranded, probes are complementary single strands. Probes can be labeled with reporter molecules for use in hybridization reactions including Southern, northern, *in situ*, dot blot, array, and like technologies or in screening assays.

"Protein" refers to a polypeptide or any portion thereof. A "portion" of a protein refers to that length of amino acid sequence which would retain at least one biological activity, a domain identified by PFAM or PRINTS analysis or an antigenic epitope of the protein identified using
10 Kyte-Doolittle algorithms of the PROTEAN program (DNASTAR, Madison WI). An "oligopeptide" is an amino acid sequence from about five residues to about 15 residues that is used as part of a fusion protein to produce an antibody.

"Purified" refers to any molecule or compound that is separated from its natural environment and is from about 60% free to about 90% free from other components with which it
15 is naturally associated.

"Sample" is used in its broadest sense as containing nucleic acids, proteins, antibodies, and the like. A sample may comprise a bodily fluid; the soluble fraction of a cell preparation, or an aliquot of 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;
20 a tissue; a tissue print; a fingerprint, buccal cells, skin, or hair; and the like.

"Specific binding" refers to a special and precise interaction between two molecules which is dependent upon their structure, particularly their molecular side groups. For example, the intercalation of a regulatory protein into the major groove of a DNA molecule, the hydrogen bonding along the backbone between two single stranded nucleic acids, or the binding between an
25 epitope of a protein and an agonist, antagonist, or antibody.

"Similarity" as applied to sequences, refers to the quantification (usually percentage) of nucleotide or residue matches between at least two sequences aligned using a standardized algorithm such as Smith-Waterman alignment (Smith and Waterman (1981) J Mol Biol 147:195-197) or BLAST2 (Altschul *et al.* (1997) Nucleic Acids Res 25:3389-3402). BLAST2 may be used
30 in a standardized and reproducible way to insert gaps in one of the sequences in order to optimize alignment and to achieve a more meaningful comparison between them.

"Substrate" refers to any rigid or semi-rigid support to which cDNAs 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
35 surface forms including wells, trenches, pins, channels and pores.

"Variant" refers to molecules that are recognized variations of a cDNA or a protein encoded by the cDNA. Splice variants may be determined by BLAST score, wherein the score is at least 100, and most preferably at least 400. Allelic variants have a high percent identity to the cDNAs and may differ by about three bases per hundred bases. "Single nucleotide

WO 02/055706

PCT/US02/00647

polymorphism" (SNP) refers to a change in a single base as a result of a substitution, insertion or deletion. The change may be conservative (purine for purine) or non-conservative (purine to pyrimidine) and may or may not result in a change in an encoded amino acid or its secondary, tertiary, or quaternary structure.

THE INVENTION

The invention is based on the discovery of a cDNA which encodes ARP and on the use of the cDNA, or fragments thereof, and protein, or portions thereof, directly or as compositions in the characterization, diagnosis, and treatment of cancer.

Nucleic acids encoding the ARP-1 of the present invention were first identified in Incyte Clone 1555118 from the human bladder tumor cDNA library (BLADTUT04) using a computer search for amino acid sequence alignments. A consensus sequence, SEQ ID NO:2, was derived from the following overlapping and/or extended nucleic acid sequences (SEQ ID NO:4-11): Incyte Clones 1555118H1 (BLADTUT04), 722739LH1 (BRAXTDR1.5), 70158486V1 (SG0000039), 70162686V1 (SG0000040), 70151326V1 (SG0000038), 70154198V1 (SG0000038), 2084238T6 (UTRSNOT08), 70155923V1 (SG0000038), and GenBank BST g6661750 (SEQ ID NO:57), and edited Genscan sequence GNN.g10801482_004.edit (SEQ ID NO:58). For sequence GNN.g10801482_004.edit, coding regions were predicted by Genscan analysis of the genomic DNA. g10801482 is the GenBank identification number of the sequence to which Genscan was applied.

In one embodiment, the invention encompasses a polypeptide comprising the amino acid sequence of SEQ ID NO:1 as shown in Figures 1A, 1B, 1C, 1D, 1E, 1F, 1G, 1H, 1I, 1J and 1K. ARP-1 is 935 amino acids in length and has four potential N-glycosylation sites at N332, N502, N516, and N778; two potential cyclic AMP- or cyclic GMP-dependent protein kinase phosphorylation sites at S141 and S745; eleven potential casein kinase II phosphorylation sites at T196, S242, T334, T445, T521, S562, S635, S676, S745, S822, and S927; two potential glycosaminoglycan attachment sites at S207 and S807; fourteen potential protein kinase C phosphorylation sites at S162, S190, S225, S242, S352, S375, T442, S455, S512, T569, T619, S620, S780, and S801; one potential tyrosine kinase phosphorylation site at Y789; one potential ATP/GTP-binding site motif A (P-loop) from G555 through S562. PFAM analysis indicates that the regions of ARP-1 from T203 to P290, K383 to Q469, and E498 to R591 are similar to PDZ domains. Such domains participate in protein-protein interactions with signaling molecules. As shown in Figures 3A, 3B, 3C, 3D, 3E, 3F, 3G, 3H, 3I and 3J, ARP-1 has chemical and structural similarity with rat ASIP (g3868778; SEQ ID NO:62) and human ASIP (g8037915; SEQ ID NO:63). In particular, ARP-1 and rat ASIP share about 41% identity and ARP-1 and human ASIP share about 43% identity. The region of ARP-1 from S708 to F925 is similar to the aPCK binding regions of rat ASIP and human ASIP. All three proteins share three conserved PDZ domains and

WO 02/055706

PCT/US02/00647

potential protein kinase C phosphorylation sites. Useful antigenic epitopes of ARP-1 extend from R192 to S222, K413 to S455, and K759 to Q809; and biologically active portions of ARP-1 extend from T203 to P290, K383 to Q469, E498 to R591, and S708 to F925. An antibody which specifically binds ARP-1 is useful in a diagnostic assay for cancer, particularly bladder transitional cell carcinoma.

Nucleic acids encoding the ARP-2 of the present invention were first identified in Incyte Clone 2582063 from the human bladder tumor cDNA library (KIDNTUT13) using a computer search for amino acid sequence alignments. A consensus sequence, SEQ ID NO:2, was derived from the following overlapping and/or extended nucleic acid sequences (SEQ ID NO:4-11): Incyte Clones 2582063H1 (KIDNTUT13), 7246093H1 (PROSTMV01), 7978420H1, 55040412H1, 2929484F6 (TYMN0T04), 5627320R8 (PLACFER01), 3209128F6 (BLADNOT08), 349248H1 (LVBNNOT01), 7019961H1 (PANCN0N03), 6303175H2 (TYMUNT02), 2549906F6 (LUNGTUT06), 1945452H1 (PITUNOT01), 2549906T6 (LUNGTUT06), 71009002V1 (SG0000308), 71008521V1 (SG0000308), 71010168V1 (SG0000308), 70090181V1 (SG0000030), 6833928H1 (BRSTN0N02), 70089663V1 (SG0000030), and GenBank ESTs g6993427 (SEQ ID NO:59), g5529915 (SEQ ID NO:60), and g1733437 (SEQ ID NO:61).

In another embodiment, the invention encompasses a polypeptide comprising the amino acid sequence of SEQ ID NO:2 as shown in Figures 2A, 2B, 2C, 2D, 2E, 2F, 2G, 2H, 2I, 2J, 2K, 2L, 2M, 2N, 2O, 2P, 2Q, 2R and 2S. ARP-2 is 1356 amino acids in length and has ten potential N-glycosylation sites at N112, N159, N208, N265, N391, N594, N608, N922, N1144, and N1231; six potential cyclic AMP- or cyclic GMP-dependent protein kinase phosphorylation sites at S144, T468, S685, S715, S888, and T1176; thirty-four potential casein kinase II phosphorylation sites at S123, S154, S161, T243, S248, T258, S267, S370, T476, S523, T534, T569, T654, T705, S720, S780, S781, S827, S840, T865, S888, T947, S958, T965, T995, T1048, S1049, S1102, S1116, S1139, S1149, S1245, S1252, and S1308; twenty-six potential protein kinase C phosphorylation sites at S166, S187, S292, S372, S379, S428, T453, T476, S533, S604, T661, S742, S829, T847, S924, S955, S958, T988, T995, T1038, S1116, S1143, S1192, S1228, S1252, and S1312; six potential tyrosine kinase phosphorylation sites at Y199, Y388, Y745, Y933, Y1080, and Y1321; two potential ATP/GTP-binding sites motif A (P-loop) from G516 through S523 and from G647 through S654. PFAM analysis indicates that the regions of ARP-2 from K273 to A360, N461 to Q547, and E590 to R683 are similar to PDZ domains. As shown in Figures 3A, 3B, 3C, 3D, 3E, 3F, 3G, 3H, 3I and 3J, ARP-2 has chemical and structural similarity with rat ASIP (g3868778; SEQ ID NO:62) and human ASIP (g8037915; SEQ ID NO:63). In particular, ARP-2 and rat ASIP share about 89% identity and ARP-2 and human ASIP share about 93% identity. The region of ARP-2 from R712 to K936 is similar to the aPCK binding regions of rat ASIP and human ASIP. All three proteins share three conserved PDZ domains and potential protein kinase C

WO 02/055706

PCT/US02/00647

phosphorylation sites. Useful antigenic epitopes of ARP-2 extend from K189 to Q236, I395 to M1023, and R1207 to N1267; and biologically active portions of ARP-2 extend from K273 to A360, N461 to Q547, E590 to R683, and R712 to K936. An antibody which specifically binds ARP-2 is useful in a diagnostic assay for cancer, particularly bladder transitional cell carcinoma.

Table 1 shows expression of ARP-1 and ARP-2 across the tissue categories (also listed in Example VIII). Table 2 shows expression of ARP in bladder tissues, particularly in tissues from patients with transitional cell carcinoma. ARP shows overexpression in a library (BLADTUT04) from bladder tissue from a patient with transitional cell carcinoma compared to a library (BLADNOT05) from matched (m) microscopically normal tissue from the same donor. Table 3 shows the differential expression of ARP-2 in human BT20 breast carcinoma cells treated with EGF as determined by microarray analysis. ARP-2 shows reduced expression in BT20 cells treated with EGF for 4 to 48 hours compared to untreated cells. The largest decrease in expression in treated BT20 cells is observed at 36 hours. Therefore, the cDNAs encoding ARP-1 or ARP-2 are useful in assays to diagnose cancer, particularly bladder transitional cell carcinoma. Fragments of the cDNA encoding ARP-1 from about nucleotide 1792 to about nucleotide 1836, from about nucleotide 1905 to about nucleotide 1950, and from about nucleotide 2180 to about nucleotide 2230 are also useful in diagnostic assays. Fragments of the cDNA encoding ARP-2 from about nucleotide 585 to about nucleotide 635 and from about nucleotide 1670 to about nucleotide 1710 are also useful in diagnostic assays.

Mammalian variants of the cDNA encoding ARP were identified using BLAST2 with default parameters and the ZOOSEQ databases (Incyte Genomics). These preferred variants have from about 80% to about 100% identity as shown in the table below. The first column shows the SEQ ID for the human cDNA (SEQ ID_h); the second column, the SEQ ID for the variant cDNAs (SEQ ID_v); the third column, the clone number for the variant cDNAs (Clone_v); the fourth column, the library name; the fifth column, the alignment of the variant cDNA to the human cDNA (includes the alignment of different regions of the variant cDNA with different regions of the human cDNA in some cases); and the sixth column, the percent identity to the human cDNA.

| | SEQ ID _h | SEQ ID _v | Clone _v | Library Name | N _h Alignment | Identity |
|----|---------------------|---------------------|--------------------|--------------|--------------------------|----------|
| | 3 | 12 | 702457509T1 | RACONOR05 | 2162-2619 | 84% |
| 35 | 3 | 13 | 702458746T1 | RACONOR05 | 2152-2579 | 84% |
| | 3 | 14 | 701335936H1 | RALINOR06 | 1322-1566 | 88% |
| | 3 | 15 | 700639694H1 | RATONOR01 | 2361-2619 | 87% |
| | 3 | 16 | 700639694F6 | RATONOR01 | 2447-2651 | 85% |
| | 3 | 17 | 701191467H1 | RACONOR05 | 2063-2333 | 81% |
| | 3 | 18 | 702771158H1 | CNLIINOT02 | 615-710 | 87% |
| 40 | 3 | 19 | 701266650H1 | MOLINDTC8 | 1289-1544 | 91% |
| | 20 | 40 | 702231139H1 | RAFANOT02 | 2500-2951 | 88% |
| | 20 | 41 | 700273304F6 | RASJNCT01 | 2705-3074 | 85% |
| | 20 | 42 | 700330856H1 | RALINON04 | 1550-1817 | 87% |
| | 20 | 43 | 700273304H1 | RASJNCT01 | 2705-2951 | 86% |

| | | | | | | | |
|----|----|----|--------------|-----------|-----------|----------------|--|
| | | | WO 02/055706 | | | PCT/US02/00647 | |
| | 20 | 44 | 701517518H1 | RALTEXT6Z | 2149-2441 | 83% | |
| 5 | 20 | 45 | 701834089T1 | RAKTEXT10 | 2897-3074 | 85% | |
| | 20 | 46 | 701480437B1 | RALTEXT42 | 3679-3824 | 88% | |
| | 20 | 47 | 701190235H1 | RACONCNO5 | 2077-2333 | 80% | |
| | 20 | 48 | 700939688H1 | RALINCNO7 | 2971-3074 | 88% | |
| | 20 | 49 | 700939688F6 | RALINCNO7 | 2971-3074 | 88% | |
| 10 | 20 | 50 | 702582937T1 | RABYUNND1 | 5276-5319 | 95% | |
| | 20 | 51 | 700299037P6 | RABCONO2 | 5276-5334 | 91% | |
| | 20 | 52 | 701246488E1 | RALINCNO2 | 2284-2357 | 86% | |
| | 20 | 53 | 702759912H1 | CNLUNND1 | 1-566 | 92% | |
| | 20 | 54 | 700112340H1 | MOOSUNRD1 | 121-362 | 88% | |
| 15 | 20 | 55 | 700827810H1 | MOOSUNRD1 | 3559-3701 | 87% | |
| | 20 | 56 | 700109331H1 | MOOSUNRD1 | 2509-2569 | 100% | |

These cDNAs are particularly useful for producing transgenic cell lines or organisms which model human disorders and upon which potential therapeutic treatments for such disorders may be tested.

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

The cDNA and fragments thereof (SEQ ID NOs:3-56) may be used in hybridization, amplification, and screening technologies to identify and distinguish among SEQ ID NO:3, SEQ ID NO:20 and related molecules in a sample. The mammalian cDNAs may be used to produce transgenic cell lines or organisms which are model systems for human cancer, particularly bladder transitional cell carcinoma and upon which the toxicity and efficacy of potential therapeutic treatments may be tested. Toxicology studies, clinical trials, and subject/patient treatment profiles may be performed and monitored using the cDNAs, proteins, antibodies and molecules and compounds identified using the cDNAs 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), and AUTOASSEMBLER application

WO 02/055706

PCT/US02/00647

(Applied Biosystems, Foster City CA). Clones, extension and/or shotgun sequences are

5 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 (APB), Piscataway NJ), or combinations of 10 polymerases and proofreading exonucleases such as those found in the ELONGASE amplification system (Life Technologies, Gaithersburg MD). Preferably, sequence preparation is automated with machines such as the MICROLAB 2200 system (Hamilton, Reno NV) and the DNA ENGINE thermal cycler (MJ Research, Watertown MA). Machines commonly used for 15 sequencing include the ABI PRISM 3700, 377 or 373 DNA sequencing systems (Applied Biosystems), the MEGABACE 1000 DNA sequencing system (APB), and the like. The sequences may be analyzed using a variety of algorithms well known in the art and described in Ausubel *et al.* (1997; *Short Protocols in Molecular Biology*, John Wiley & Sons, New York NY, unit 7.7) and in Meyers (1995; *Molecular Biology and Biotechnology*, Wiley VCH, New York NY, pp. 856- 20 853).

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 25 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 cDNAs 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 30 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 (Applied Biosystems), nested primers, and 35 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 primer analysis software (Molecular Biology Insights, Cascade CO) 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

WO 02/055706

PCT/US02/00647

recover regulatory elements, it is preferable to use genomic, rather than cDNA libraries.

5 **Hybridization**

The cDNA 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 ARP, 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:3-56. 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 cDNA 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 APB.

The stringency of hybridization is determined by G+C content of the probe, salt concentration, and temperature. In particular, stringency can be increased by reducing the concentration of salt or raising the hybridization temperature. In solutions used for some membrane based hybridizations, addition of an organic solvent such as formamide allows the reaction to occur at a lower temperature. Hybridization can be performed at low stringency with buffers, such as 5xSSC with 1% sodium dodecyl sulfate (SDS) at 60C, 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 45C (medium stringency) or 68C (high stringency). At high stringency, hybridization complexes will remain stable only where the nucleic acids 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 (Sigma-Aldrich, St. Louis MO) and a blocking agent such as denatured salmon sperm DNA. Selection of components and conditions for hybridization are well known to those skilled in the art and are reviewed in Ausubel (*supra*) and Sambrook et al. (1989) Molecular Cloning, A Laboratory Manual, Cold Spring Harbor Press, Plainview NY.

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.

WO 02/055706

PCT/US02/00647

(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) an 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

Any one of a multitude of cDNAs encoding ARP 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 create new restriction sites, alter glycosylation patterns, change codon preference to increase expression in a particular host, produce splice variants, extend half-life, and the like. The expression vector may contain transcriptional and translational control elements (promoters, enhancers, specific initiation signals, and polyadenylated 3' sequence) from various sources which have been selected for their efficiency in a particular host. The vector, cDNA, 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.

WO 02/055706

PCT/US02/00647

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), β 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 cDNA is based on DNA-DNA or DNA-RNA hybridizations or PCR amplification techniques.

The host cell may be chosen for its ability to modify a recombinant protein in a desired fashion. Such modifications include acetylation, carboxylation, glycosylation, phosphorylation, lipidation, acylation and the like. Post-translational processing which cleaves a "prepro" form may also be used to specify protein targeting, folding, and/or activity. Different host cells available from the ATCC (Manassas VA) 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.

Recovery of Proteins from Cell Culture

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

Chemical Synthesis of Peptides

Proteins or portions thereof may be produced not only by recombinant methods, but also by using chemical methods well known in the art. Solid phase peptide synthesis may be carried out in a batchwise or continuous flow process which sequentially adds α -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- α -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

WO 02/055706

PCT/US02/00647

to the solid phase support resin. Trifluoroacetic acid or piperidine are used to remove the
5 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
10 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 (Applied Biosystems). A protein or portion thereof may be substantially
purified by preparative high performance liquid chromatography and its composition confirmed by
15 amino acid analysis or by sequencing (Creighton (1984) Proteins, Structures and Molecular
Properties, WH Freeman, New York NY).

Preparation and Screening of Antibodies

Various hosts including goats, rabbits, rats, mice, humans, and others may be immunized
by injection with ARP or any portion thereof. Adjuvants such as Freund's, mineral gels, and
20 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.

25 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)
30 *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 protein may also be
generated. For example, such fragments include, but are not limited to, F(ab')₂ fragments
35 produced by pepsin digestion of the antibody molecule and Fab fragments generated by reducing
the disulfide bridges of the F(ab')₂ fragments. Alternatively, Fab expression libraries may be
constructed to allow rapid and easy identification of monoclonal Fab fragments with the desired
specificity. (See, e.g., Huse et al. (1989) *Science* 246:1275-1281.)

The ARP or a portion thereof may be used in screening assays of phagemid or B-

WO 02/055706

PCT/US02/00647

lymphocyte immunoglobulin libraries to identify antibodies having the desired specificity.

- 5 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).

10 Labeling of Molecules for Assay

A wide variety of reporter molecules 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 commercially available kits (Promega, Madison WI) for incorporation of a labeled nucleotide such as ³²P-dCTP (APB), Cy3-dCTP or Cy5-dCTP (Operon Technologies, Alameda CA), or amino acid such as ³⁵S-methionine (APB). Nucleotides and amino acids may be directly labeled with a variety of substances including fluorescent, chemiluminescent, or chromogenic agents, and the like, by chemical conjugation to amines, thiols and other groups present in the molecules using reagents such as BIODIPY or FITC

- 20 (Molecular Probes, Eugene OR).

DIAGNOSTICS

The cDNAs, fragments, oligonucleotides, complementary RNA and DNA molecules, and PNAs and may be used to detect and quantify differential gene expression, absence/presence vs. excess, expression of mRNAs or to monitor mRNA levels during therapeutic intervention.

- 25 Similarly antibodies which specifically bind ARP may be used to quantitate the protein. Disorders associated with differential expression include cancer, particularly bladder transitional cell carcinoma. 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 differential gene expression. Qualitative or quantitative methods for this comparison are well known in the art.

- For example, the cDNA or probe may be labeled by standard methods and added to a biological sample from a patient under conditions for the formation of hybridization complexes. After an incubation period, the sample is washed and the amount of label (or signal) associated with hybridization complexes, is quantified and compared with a standard value. If complex formation in the patient sample is significantly altered (higher or lower) in comparison to either a normal or disease standard, then differential expression indicates the presence of a disorder.

In order to provide standards for establishing differential expression, normal and disease expression profiles are established. This is accomplished by combining a sample taken from normal subjects, either animal or human, with a cDNA under conditions for hybridization to

WO 02/055706

PCT/US02/00647

5 occur. Standard hybridization complexes 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 sequence is used. Standard values obtained in this manner may be compared with values obtained from samples from patients who were diagnosed with a particular condition, disease, or disorder. Deviation from standard values toward those associated with a particular disorder is used to diagnose that disorder.

10 Such assays may also be used to evaluate the efficacy of a particular therapeutic treatment regimen in animal studies and in clinical trial or to monitor the treatment of an individual patient. Once the presence of a condition is established and a treatment protocol is initiated, diagnostic assays may be repeated on a regular basis to determine if the level of expression in the patient begins to approximate that which is observed in a normal subject. The results obtained from
15 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 antibodies are known in the art. Examples of such techniques include enzyme-linked
20 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 Pound, *supra*.)

25 THERAPEUTICS

Chemical and structural similarity, in the context of the PDZ domains and aPKC binding region, exists between regions of ARP-1 (1555118; SEQ ID NO:1), ARP-2 (2582063; SEQ ID NO:2), rat ASIP (g3868778; SEQ ID NO:62), and human ASIP (g8037915; SEQ ID NO:63) as shown in Figures 3A, 3B, 3C, 3D, 3E, 3F, 3G, 3H, 3I and 3J. In addition, differential expression
30 of ARP is associated with the digestive system, female reproductive tissue, hemic and immune system, nervous system, respiratory system, and urinary tract and with cancer, particularly bladder transitional cell carcinoma as shown in Tables 1 and 2. ARP clearly plays a role in cancer, particularly bladder transitional cell carcinoma.

In the treatment of conditions associated with increased expression of ARP it is desirable
35 to decrease expression or protein activity. In one embodiment, the an inhibitor, antagonist or antibody of the protein may be administered to a subject to treat a condition associated with increased expression or activity. In another embodiment, a pharmaceutical composition comprising an inhibitor, antagonist or antibody in conjunction with a pharmaceutical carrier may be administered to a subject to treat a condition associated with the increased expression or

WO 02/055706

PCT/US02/00647

activity of the endogenous protein. In an additional embodiment, a vector expressing the
5 complement of the cDNA or fragments thereof may be administered to a subject to treat the disorder.

In the treatment of conditions associated with decreased expression of ARP it is desirable to increase expression or protein activity. In one embodiment, the protein, an agonist or enhancer may be administered to a subject to treat a condition associated with decreased expression or
10 activity. In another embodiment, a pharmaceutical composition comprising the protein, an agonist or enhancer in conjunction with a pharmaceutical carrier may be administered to a subject to treat a condition associated with the decreased expression or activity of the endogenous protein. In an additional embodiment, a vector expressing cDNA may be administered to a subject to treat the disorder.

15 Any of the cDNAs, complementary molecules, or fragments thereof, proteins or portions thereof, vectors delivering these nucleic acid molecules or expressing the 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
20 synergistically to affect treatment of a particular disorder 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 ARP. Oligonucleotides designed with reference to the transcription initiation site are preferred.
25 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 or plurality of cDNAs or fragments thereof
30 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
35 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

WO 02/055706

PCT/US02/00647

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

Screening and Purification Assays

The cDNA encoding ARP 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 cDNA in the biological system. The assay involves combining the cDNA 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, if appropriate, double stranded molecule.

In one embodiment, the cDNA of the invention may be incubated with a plurality of 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 cDNA may be incubated with nuclear extracts from biopsied and/or cultured cells and tissues. Specific binding between the cDNA and a molecule or compound in the nuclear extract is initially determined by gel shift assay and may be later confirmed by recovering and 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 cDNA may be used to purify a molecule or compound using affinity chromatography methods well known in the art. In one embodiment, the cDNA 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 cDNA. The molecule or compound which is bound to the cDNA may be released from the cDNA 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 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, ARP or a portion thereof may be used to screen a plurality of

WO 02/055706

PCT/US02/00647

5 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 peptide on their cell surface can be used in screening assays. The cells are screened against a plurality or libraries of ligands and the specificity of binding or formation of complexes between the expressed protein 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.

10 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 protein specifically compete with a test compound capable of binding to the protein 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.

15 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 human use.

WO 02/055706

PCT/US02/00647

Model Systems

5 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

15 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.

20 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,

WO 02/055706

PCT/US02/00647

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

Transgenic Animal Models

10 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.

Embryonic Stem Cells

15 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.

20 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

25 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.

WQ 02/055706

PCT/US02/00647

Knockin Analysis

5 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 inbred lines are studied and treated with potential pharmaceutical agents to
10 obtain information on treatment of the analogous human condition. These methods have been used to model several human diseases.

Non-Human Primate Model

The field of animal testing deals with data and methodology from basic sciences such as physiology, genetics, chemistry, pharmacology and statistics. These data are paramount in
15 evaluating the effects of therapeutic agents on non-human primates as they can be related to human health. Monkeys are used as human surrogates in vaccine and drug evaluations, and their responses are relevant to human exposures under similar conditions. Cynomolgus and Rhesus monkeys (Macaca fascicularis and Macaca mulatta, respectively) and Common Marmosets (Callithrix jacchus) are the most common non-human primates (NHPs) used in these
20 investigations. Since great cost is associated with developing and maintaining a colony of NHPs, early research and toxicological studies are usually carried out in rodent models. In studies using behavioral measures such as drug addiction, NHPs are the first choice test animal. In addition, NHPs and individual humans exhibit differential sensitivities to many drugs and toxins and can be classified as a range of phenotypes from "extensive metabolizers" to "poor metabolizers" of these
25 agents.

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

EXAMPLES

The examples below are provided to illustrate the subject invention and are not included for the purpose of limiting the invention. For purposes of example, preparation of the human bladder tumor (BLADTUT04) and normalized breast (BRSTNON02) libraries will be described.

I cDNA Library Construction35 Bladder tumor library

The BLADTUT04 cDNA library was constructed from bladder tumor tissue obtained from a 60 year-old Caucasian male who had a transitional cell carcinoma in the left bladder wall. The frozen tissue was homogenized and lysed using a POLYTRON homogenizer (Brinkmann Instruments, Westbury NJ). The reagents and extraction procedures were used as supplied in the

WO 02/055706

PCT/US02/00647

RNA Isolation kit (Stratagene). The lysate was centrifuged over a 5.7 M CsCl cushion using an
5 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 4.7; precipitated using 0.3 M sodium acetate and 2.5 volumes of
ethanol; resuspended in 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.

10 The mRNA was handled according to the recommended protocols in the SUPERSCRIPT
plasmid system (Life Technologies) which contains a NotI primer-adaptor designed to prime the
first strand cDNA synthesis at the poly(A) tail of mRNAs. Double stranded cDNA was blunted,
ligated to EcoRI adaptors and digested with NotI (New England Biolabs, Beverly MA). The
cDNAs were fractionated on a SEPHAROSE CL4B column (APB), and those cDNAs exceeding
15 400 bp were ligated into pINCY plasmid (Incyte Genomics). The plasmid pINCY was
subsequently transformed into DH5 α competent cells (Life Technologies).

Normalized breast

About 1.2×10^6 independent clones of the pooled BRSTNOT34 and BRSTNOT35
plasmid libraries in *E. coli* strain DH12S competent cells (Life Technologies) were grown in
20 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 two rounds according to
the procedure of Soares *et al.* (1994; Proc Natl Acad Sci 91:9228-9232) and Bonaldo *et al.* (1996;
Genome Research 6:791-806), with the following modifications. The primer to template ratio in
25 the primer extension reaction was increased from 2:1 to 300:1. 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, ligated into pINCY plasmid and electroporated into DH12S competent cells (Life
Technologies).

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

An intermediate plasmid produced by the bacteria (pSPORT 1- Δ 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

WO 02/055706

PCT/US02/00647

cells, plasmids were isolated and tested for preferential digestibility with EcoRI, but not with Hind

5 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

10 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

15 employed except for the following changes: 1) the bacteria were cultured in 1 ml of sterile TERRIFIC BROTH (BD Biosciences, Sparks MD) 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

20 block for storage at 4C.

The cDNAs were prepared for sequencing using the MICROLAB 2200 system (Hamilton) 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 (Applied Biosystems) or the MEGABACE 1000 DNA sequencing

25 system (APB). Most of the isolates were sequenced according to standard ABI protocols and kits (Applied Biosystems) with solution volumes of 0.25x-1.0x concentrations. In the alternative, cDNAs were sequenced using solutions and dyes from APB.

IV Extension of cDNA Sequences

The cDNAs were extended using the cDNA clone and oligonucleotide primers. One

30 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 primer analysis software (Molecular Biology Insights), 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 structures and primer-primer

35 dimerizations was avoided.

Selected cDNA libraries were used as templates to extend the sequence. If more than one extension was necessary, additional or nested sets of primers were designed. Preferred libraries have been size-selected to include larger cDNAs and random primed to contain more sequences with 5' or upstream regions of genes. Genomic libraries are used to obtain regulatory elements.

WO 02/055706

PCT/US02/00647

especially extension into the 5' promoter binding region.

5 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 Mg²⁺, (NH₄)₂SO₄, and β-mercaptoethanol, Taq DNA polymerase (APB), ELONGASE enzyme (Life Technologies), and Pfu DNA polymerase (Stratagene), with the
10 following parameters for primer pair PCI A and PCI B (Incyte Genomics): 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
15 6: 68C, five min; Step 7: storage at 4C.

The concentration of DNA in each well was determined by dispensing 100 μl PICOGREEN quantitation reagent (0.25% reagent in 1x TE, v/v; Molecular Probes) and 0.5 μl 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
20 Oy) to measure the fluorescence of the sample and to quantify the concentration of DNA. A 5 μl to 10 μl aliquot of the reaction mixture was analyzed by electrophoresis on a 1% agarose mini-gel to determine which reactions were successful in extending the sequence.

The extended clones were desalted, concentrated, transferred to 384-well plates, digested with CviII cholera virus endonuclease (Molecular Biology Research, Madison WI), and sonicated
25 or sheared prior to religation into pUC18 vector (APB). 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 (APB), treated with Pfu DNA polymerase (Stratagene) to fill-in restriction site overhangs, and transfected into E.
30 coli competent cells. Transformed cells were selected on antibiotic-containing media, and individual colonies were picked and cultured overnight at 37C in 384-well plates in LB/2x carbenicillin liquid media.

The cells were lysed, and DNA was amplified using primers, Taq DNA polymerase (APB) and Pfu DNA polymerase (Stratagene) with the following parameters: Step 1: 94C, three
35 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

WO 02/055706

PCT/US02/00647

sequencing primers and the DYNAMIC DIRECT cycle sequencing kit (APB) or the ABI PRISM
5 BIGDYE terminator cycle sequencing kit (Applied Biosystems).

V Homology Searching of cDNA Clones and Their Deduced Proteins

The cDNAs 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
10 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
15 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
20 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
25 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 nucleic acid molecules and BLAST 2 that is used
for direct pairwise comparison of either nucleic or amino acid molecules. BLAST programs are
30 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 drop-off: 50; Expect: 10; Word Size: 11; and Filter: on. Identity is 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
35 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.

Putative ASIP-related proteins were initially identified by running the Genscan gene
identification program against public genomic sequence databases (e.g., gbpr1 and gbhtg).
Genscan is a general-purpose gene identification program which analyzes genomic DNA

WO 02/055706

PCT/US02/00647

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

The cDNAs of this application were compared with assembled consensus sequences or
templates found in the LIFESEQ GOLD database. Component sequences from cDNA, extension,
25 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
30 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
35 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.

WO 02/055706

PCT/US02/00647

Bins were compared to one another and those having local similarity of at least 82% were
5 combined and reassembled. Bins having templates with less than 95% local identity were split.
Templates were subjected to analysis by SPLICER/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
10 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
15 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 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
20 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 cDNA was further analyzed using MACDNASIS PRO software
(Hitachi Software Engineering), and LASERGENE software (DNASTAR) and queried against
25 public databases such as the GenBank rodent, mammalian, vertebrate, prokaryote, and eukaryote
databases, SwissProt, BLOCKS, PRINTS, PFAM, and Prosite.

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),
30 and Génethon are used to determine if any of the cDNAs presented in the Sequence Listing have
been mapped. Any of the fragments of the cDNA encoding ARP 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
35 measured relative to the terminus of the chromosomal p-arm.

VII Hybridization Technologies and Analyses

Immobilization of cDNAs on a Substrate

The cDNAs are applied to a substrate by one of the following methods. A mixture of
cDNAs is fractionated by gel electrophoresis and transferred to a nylon membrane by capillary

WO 02/055706

PCT/US02/00647

transfer. Alternatively, the cDNAs are individually ligated to a vector and inserted into bacterial
5 host cells to form a library. The cDNAs 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
10 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, cDNAs are amplified from bacterial vectors by thirty cycles of PCR
using primers complementary to vector sequences flanking the insert. PCR amplification
15 increases a starting concentration of 1-2 ng nucleic acid to a final quantity greater than 5 µg.
Amplified nucleic acids from about 400 bp to about 5000 bp in length are purified using
SEPHACRYL-400 beads (APB). 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
20 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 (Coming, 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
25 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
30 washed in 0.2% SDS and rinsed in distilled water as before.

Probe Preparation for Membrane Hybridization

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

WO 02/055706

PCT/US02/00647

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

Probe Preparation for Polymer Coated Slide Hybridization

Hybridization probes derived from mRNA isolated from samples are employed for
screening cDNAs of the Sequence Listing in array-based hybridizations. Probe is prepared using
the GEMbright kit (Incyte Genomics) by diluting mRNA to a concentration of 200 ng in 9 μ l TE
10 buffer and adding 5 μ l 5x buffer, 1 μ l 0.1 M DTT, 3 μ l Cy3 or Cy5 labeling mix, 1 μ l RNase
inhibitor, 1 μ l reverse transcriptase, and 5 μ 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
15 diluted into reverse transcription reaction mixture at ratios of 1:100,000, 1:10,000, 1:1000, and
1:100 (w/w) to sample mRNA respectively. To examine mRNA differential expression patterns, a
second set of control mRNAs are diluted into reverse transcription reaction mixture at ratios of
1:3, 3:1, 1:10, 10:1, 1:25, and 25:1 (w/w). The reaction mixture is mixed and incubated at 37C for
two hr. The reaction mixture is then incubated for 20 min at 85C, and probes are purified using
20 two successive CHROMA SPIN+TE 30 columns (Clontech, Palo Alto CA). Purified probe is
ethanol precipitated by diluting probe to 90 μ l in DEPC-treated water, adding 2 μ l 1mg/ml
glycogen, 60 μ l 5 M sodium acetate, and 300 μ l 100% ethanol. The probe is centrifuged for 20
min at 20,800xg, and the pellet is resuspended in 12 μ 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.

25 Membrane-based Hybridization

Membranes are pre-hybridized in hybridization solution containing 1% Sarkosyl and 1x
high phosphate buffer (0.5 M NaCl, 0.1 M Na_2HPO_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
30 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

35 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 μ 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 μ l of 5xSSC in a corner of the chamber. The chamber containing

WO 02/055706

PCT/US02/00647

the arrays is incubated for about 6.5 hr at 60C. The arrays are washed for 10 min at 45C in
5 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
10 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
15 combined fluorescence (Shalon WO95/35505).

Hybridization complexes are detected with a microscope equipped with an Innova 70
mixed gas 10 W laser (Coherent, Santa Clara CA) capable of generating spectral lines at 488 nm
for excitation of Cy3 and at 632 nm for excitation of Cy5. The excitation laser light is focused on
the array using a 20X microscope objective (Nikon, Melville NY). The slide containing the array
20 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
25 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 RTT-835H analog-to-
30 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
35 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.

WQ 02/055706

PCT/US02/00647

The software used for signal analysis is the GEMTOOLS program (Incyte Genomics).

5 VIII Electronic Analysis

BLAST was used to search for identical or related molecules in the GenBank or LIPSESEQ databases (Incyte Genomics). 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
10 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 40.

15 Electronic northern analysis was performed at a product score of 70 as shown in Tables 1 and 2. All sequences and cDNA libraries in the LIPSESEQ database were categorized by system, organ/tissue and cell type. The categories included cardiovascular system, connective tissue, digestive system, embryonic structures, endocrine system, exocrine glands, female and male genitalia, germ cells, hemio/immune system, liver, musculoskeletal system, nervous system,
20 pancreas, respiratory system, sense organs, skin, stomatognathic system, unclassified/mixed, and the urinary tract. For each category, the number of libraries in which the sequence was expressed were counted and shown over the total number of libraries in that category. In a non-normalized library, expression levels of two or more are significant.

IX Complementary Molecules

25 Molecules complementary to the cDNA, from about 5 (PNA) to about 5000 bp (complement of a cDNA insert), are used to detect or inhibit gene expression. These molecules are selected using OLIGO primer analysis software (Molecular Biology Insights). Detection is described in Example VII. To inhibit transcription by preventing promoter binding, the
30 complementary molecule is designed to bind to the most unique 5' sequence and includes nucleotides of the 5' UTR upstream of the initiation codon of the open reading frame. Complementary molecules include genomic sequences (such as enhancers or introns) and are used in "triple helix" base pairing to compromise the ability of the double helix to open sufficiently for the binding of polymerases, transcription factors, or regulatory molecules. To inhibit translation, a complementary molecule is designed to prevent ribosomal binding to the mRNA encoding the
35 protein.

Complementary molecules are placed in expression vectors and used to transform 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, zygote, or other reproducing lineage for long term or stable gene therapy. Transient expression lasts for a month or more with a non-replicating vector and for

WO 02/055706

PCT/US02/00647

three months or more if appropriate elements for inducing vector replication are used in the
5 transformation/expression system.

Stable transformation of appropriate dividing cells with a vector encoding the
complementary molecule produces 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 complementary molecules to compromise or entirely eliminate
10 activity of the cDNA encoding the protein.

X Selection of Sequences, Microarray Preparation and Use

Incyte clones represent template sequences derived from the LIFESEQ GOLD assembled
human sequence database (Incyte Genomics). In cases where more than one clone was available
for a particular template, the 5'-most clone in the template was used on the microarray. The
15 HUMAN GENOME GEM series 1-3 microarrays (Incyte Genomics) contain 28,626 array
elements which represent 10,068 annotated clusters and 18,558 unannotated clusters. For the
UNIGEM series microarrays (Incyte Genomics), Incyte clones were mapped to non-redundant
Unigene clusters (Unigene database (build 46), NCBI; Shuler (1997) J Mol Med 75:694-698), and
the 5' clone with the strongest BLAST alignment (at least 90% identity and 100 bp overlap) was
20 chosen, verified, and used in the construction of the microarray. The UNIGEM V microarray
(Incyte Genomics) contains 7075 array elements which represent 4610 annotated genes and 2,184
unannotated clusters.

To construct microarrays, cDNAs were amplified from bacterial cells using primers
complementary to vector sequences flanking the cDNA insert. Thirty cycles of PCR increased the
25 initial quantity of cDNAs from 1-2 ng to a final quantity of greater than 5 μ g. Amplified cDNAs
were then purified using SEPHACRYL-400 columns (APB). Purified cDNAs were immobilized
on polymer-coated glass slides. Glass microscope slides (Corning, Corning NY) were cleaned by
ultrasound in 0.1% SDS and acetone, with extensive distilled water washes between and after
30 treatments. Glass slides were etched in 4% hydrofluoric acid (VWR Scientific Products, West
Chester PA), washed thoroughly in distilled water, and coated with 0.05% aminopropyl silane
(Sigma Aldrich) in 95% ethanol. Coated slides were cured in a 110°C oven. cDNAs were applied
to the coated glass substrate using a procedure described in USPN 5,807,522. One microliter of
the cDNA at an average concentration of 100 ng/ μ l was loaded into the open capillary printing
element by a high-speed robotic apparatus which then deposited about 5 nl of cDNA per slide.

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

WO 02/055706

PCT/US02/00647

XI Preparation of Samples5 **Human BT20 Cells**

BT-20 is a breast carcinoma cell line derived in vitro from the cells emigrating out of thin slices of a tumor mass isolated from a 74 year old female. BT20 cells were treated with EGF at a concentration of 50 ng/ml for 4, 8, 12, 24, 36 and 48 hours. In all cases mRNA from untreated BT20 cells were prepared in parallel.

10 **XII Expression of ARP**

Expression and purification of the protein are achieved using either a mammalian cell expression system or an insect cell expression system. The pUB6/V5-His vector system (Invitrogen, Carlsbad CA) is used to express ARP in CHO cells. The vector contains the selectable *bsd* gene, multiple cloning sites, the promoter/enhancer sequence from the human ubiquitin C gene, a C-terminal V5 epitope for antibody detection with anti-V5 antibodies, and a C-terminal polyhistidine (6xHis) sequence for rapid purification on PROBOND resin (Invitrogen). Transformed cells are selected on media containing blasticidin.

15 *Spodoptera frugiperda* (SF9) insect cells are infected with recombinant *Autographica californica* nuclear polyhedrosis virus (baculovirus). The polyhedrin gene is replaced with the cDNA by homologous recombination and the polyhedrin promoter drives cDNA transcription. The protein is synthesized as a fusion protein with 6xhis which enables purification as described above. Purified protein is used in the following activity and to make antibodies

XIII Production of Antibodies

25 ARP 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 ARP is analyzed using LASERGENE software (DNASTAR) to determine regions of high antigenicity. An antigenic 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 (Applied Biosystems) using Fmoc-chemistry and coupled to KLH (Sigma-Aldrich) by reaction with N-maleimidobenzoyl-N-hydroxysuccinimide ester to increase antigenicity.

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.

XIV Purification of Naturally Occurring Protein Using Specific Antibodies

WO 02/055706

PCT/US02/00647

Naturally occurring or recombinant protein is purified by immunoaffinity chromatography using antibodies which specifically bind the protein. An immunoaffinity column is constructed by covalently coupling the antibody to CNBr-activated SEPHAROSE resin (APB). 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.

XV Screening Molecules for Specific Binding with the cDNA or Protein

The cDNA, or fragments thereof, or the protein, or portions thereof, are labeled with ³²P-dCTP, Cy3-dCTP, or Cy5-dCTP (APB), 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 cDNA 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.

XVI Two-Hybrid Screen

A yeast two-hybrid system, MATCHMAKER LexA Two-Hybrid system (Clontech Laboratories, Palo Alto CA), is used to screen for peptides that bind the protein of the invention. A cDNA 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 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 are counted. 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 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

WO 02/055706

PCT/US02/00647

days at 30C. A sample of the culture is plated on SD/-Trp/-Ura media and incubated at 30C until colonies 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 cDNA encoding a protein that physically interacts with the protein, is isolated from the yeast cells and characterized.

XVII ARP Assay

ARP activity is determined in a ligand-binding assay using candidate ligand molecules such as the aPKCs, PKC ζ and PKC λ , in the presence of ¹²⁵I-labeled ARP. ARP is labeled with ¹²⁵I Bolton-Hunter reagent (Bolton and Hunter (1973) Biochem J 133:529-539). Candidate ARP molecules, previously arrayed in the wells of a multi-well plate, are incubated with the labeled ARP, washed, and any wells with labeled ARP complex are assayed. Data obtained using different concentrations of ARP are used to calculate values for the number, affinity, and association of ARP with the candidate molecules.

All patents and publications mentioned in the specification are incorporated by reference herein. 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 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.

WG 02/055706

PCT/US02/00647

| Tissue Category | Clone Count | Found In | Abs Abund | Pct Abund |
|-------------------------|----------------|----------|--------------|--------------|
| Cardiovascular System | 266190 | 1/68 | 1 | 0.0004 |
| Connective Tissue | 144645 | 0/47 | 0 | 0.0000 |
| Digestive System | 501101 | 4/148 | 4 | 0.0008 |
| Embryonic Structures | 106713 | 0/21 | 0 | 0.0000 |
| Endocrine System | 225386 | 0/53 | 0 | 0.0000 |
| Exocrine Glands | 254635 | 1/64 | 1 | 0.0004 |
| Reproductive, Female | 427284 | 2/106 | 2 | 0.0005 |
| Reproductive, Male | 448207 | 1/114 | 1 | 0.0002 |
| Germ Cells | 38282 | 1/5 | 1 | 0.0026 |
| Hemic and Immune System | 680277 | 2/159 | 2 | 0.0003 |
| Liver | 109378 | 0/35 | 0 | 0.0000 |
| Musculoskeletal System | 159280 | 0/47 | 0 | 0.0000 |
| Nervous System | 955753 | 3/198 | 5 | 0.0005 |
| Pancreas | 110207 | 0/24 | 0 | 0.0000 |
| Respiratory System | 390086 | 3/93 | 3 | 0.0008 |
| Sense Organs | 19256 | 0/8 | 0 | 0.0000 |
| Skin | 72292 | 0/15 | 0 | 0.0000 |
| Stomatognathic System | 12923 | 0/10 | 0 | 0.0000 |
| Unclassified/Mixed | 120926 | 2/13 | 2 | 0.0017 |
| Urinary Tract | 379062 | 3/64 | 5 | 0.0018 |
| Totals | 5321883 | 23/1292 | 27 | 0.0005 |

TABLE 1

WO 02/055706

PCT/US02/00647

| Library ID | Clone Count | Library Description | Abs Abund | Pct Abund |
|------------|-------------|---|-----------|-----------|
| BHNDPT04 | 7884 | Bladder tumor, TC CA, 60M, m/BHNDPT04 | 3 | 0.0381 |
| BHNDPT05 | 3776 | Bladder, m/TC CA in situ, 60M, m/BHNDPT04 | 0 | 0 |

TABLE 2

WO 02/055706

PCT/US02/00647

| mean log ₂ DB (Cys/C3) | Cys | Cys |
|---|---------------------------------------|---|
| -0.08 | Human, 8720 Uline, Untx 4hr, AdenoCA | Human, 8720 Uline, 7ESF 50ng/ml 4hr, AdenoC2 |
| -0.51 | Human, 8720 Uline, Untx 8hr, AdenoCA | Human, 8720 Uline, 7ESF 50ng/ml 8hr, AdenoC2 |
| -0.61 | Human, 8720 Uline, Untx 12hr, AdenoCA | Human, 8720 Uline, 7ESF 50ng/ml 12hr, AdenoC2 |
| -0.62 | Human, 8720 Uline, Untx 24hr, AdenoCA | Human, 8720 Uline, 7ESF 50ng/ml 24hr, AdenoC2 |
| -1.23 | Human, 8720 Uline, Untx 36hr, AdenoCA | Human, 8720 Uline, 7ESF 50ng/ml 36hr, AdenoC2 |
| -1.04 | Human, 8720 Uline, Untx 48hr, AdenoCA | Human, 8720 Uline, 7ESF 50ng/ml 48hr, AdenoC2 |

TABLE 3

WO 02/055706

PCT/US02/00647

What is claimed is:

- 5 1. An isolated cDNA encoding a protein having the amino acid sequence of SEQ ID NO:1 or SEQ ID NO:2.
2. An isolated cDNA encoding a protein having the amino acid sequence of SEQ ID NO:1.
3. An isolated cDNA encoding a protein having the amino acid sequence of SEQ ID NO:2.
4. An isolated cDNA selected from:
 - 10 a) a nucleic acid sequence of SEQ ID NO:3 or SEQ ID NO:20 or the complement thereof;
 - b) a fragment of SEQ ID NO:3 selected from SEQ ID NOs:4-11 or the complement thereof or a fragment of SEQ ID NO:20 selected from SEQ ID NOs:21-39 or the complement thereof; and
 - c) a variant of SEQ ID NO:3 selected from SEQ ID NOs:12-19 or a variant of SEQ ID NO:20 selected from SEQ ID NOs:40-56.
- 15 5. A composition comprising the cDNA or the complement of the cDNA of claim 1.
6. A vector comprising the cDNA of claim 1.
7. A host cell comprising the vector of claim 6.
8. A method for using a cDNA to produce a protein, the method comprising:
 - a) culturing the host cell of claim 7 under conditions for protein expression; and
 - 20 b) recovering the protein from the host cell culture.
9. A method for using a cDNA to detect expression of a nucleic acid in a sample comprising:
 - a) hybridizing the composition of claim 5 to nucleic acids of the sample, thereby forming hybridization complexes; and
 - b) comparing hybridization complex formation with a standard, wherein the comparison
 - 25 indicates expression of the cDNA in the sample.
10. The method of claim 9 further comprising amplifying the nucleic acids of the sample prior to hybridization.
11. The method of claim 9 wherein the composition is attached to a substrate.
12. The method of claim 9 wherein the cDNA is differentially expressed when compared with the standard and diagnostic of bladder transitional cell carcinoma.
- 30 13. A method of using a cDNA to screen a plurality of molecules or compounds, the method comprising:
 - a) combining the cDNA of claim 1 with a plurality of molecules or compounds under conditions to allow specific binding; and
 - 35 b) detecting specific binding, thereby identifying a molecule or compound which specifically binds the cDNA.
14. The method of claim 13 wherein the molecules or compounds are selected from DNA molecules, RNA molecules, peptide nucleic acids, artificial chromosome constructions, peptides,

WQ 02/055706

PCT/US02/00647

- transcription factors, repressors, and regulatory molecules.
- 5 15. A purified protein or a portion thereof selected from:
- a) an amino acid sequence of SEQ ID NO:1 or SEQ ID NO:2;
 - b) an antigenic epitope of SEQ ID NO:1 or SEQ ID NO:2; and
 - c) a biologically active portion of SEQ ID NO:1 or SEQ ID NO:2.
16. A composition comprising the protein of claim 15.
- 10 17. A method for using a protein to screen a plurality of molecules or compounds to identify at least one ligand, the method comprising:
- a) combining the protein of claim 15 with the molecules or compounds under conditions to allow specific binding; and
 - b) detecting specific binding, thereby identifying a ligand which specifically binds the
- 15 protein.
18. The method of claim 17 wherein the molecules or compounds are selected from DNA molecules, RNA molecules, peptide nucleic acids, peptides, proteins, mimetics, agonists, antagonists, antibodies, immunoglobulins, inhibitors, and drugs.
19. A purified mammalian protein comprising an amino acid sequence of SEQ ID NO:1 or SEQ
- 20 ID NO:2.
20. A purified antibody which binds specifically to the protein of claim 19.
21. A method for preparing and purifying a polyclonal antibody with the specificity of the antibody of claim 20 comprising:
- a) immunizing a animal with a protein of SEQ ID NO:1 or SEQ ID NO:2 under conditions to
- 25 elicit an antibody response;
- b) isolating animal antibodies;
 - c) attaching the protein to a substrate;
 - d) contacting the substrate with isolated antibodies under conditions to allow specific binding
- 30 to the protein;
- e) dissociating the antibodies from the protein, thereby obtaining purified polyclonal antibodies.
22. A polyclonal antibody produced by the method of claim 21.
23. A method for preparing a monoclonal antibody with the specificity of the antibody of claim 20 comprising:
- 35 a) immunizing a animal with a protein of SEQ ID NO:1 or SEQ ID NO:2 under conditions to elicit an antibody response;
- b) isolating antibody-producing cells from the animal;
 - c) fusing the antibody-producing cells with immortalized cells in culture to form monoclonal

WQ 02/055706

PCT/US02/00647

- antibody producing hybridoma cells;
- 5 d) culturing the hybridoma cells; and
- e) isolating monoclonal antibodies from culture.
24. A monoclonal antibody produced by the method of claim 23.
25. A method for using an antibody to immunopurify a protein comprising:
- 10 a) attaching an antibody of claim 20 to a substrate,
- b) exposing the antibody to a sample containing protein under conditions to allow antibody:protein complexes to form,
- c) dissociating the protein from the complex, and
- d) collecting the purified protein.
26. A method for using an antibody to detect expression of a protein in a sample, the method
- 15 comprising:
- a) combining the antibody of claim 20 with a sample under conditions which allow the formation of antibody:protein complexes; and
- b) detecting complex formation, wherein complex formation indicates expression of the protein in the sample.
- 20 27. The method of claim 26 wherein complex formation is compared with standards and is diagnostic of a bladder transitional cell carcinoma.
28. A composition comprising an antibody of claim 20 and a labeling moiety.
29. A composition comprising an antibody of claim 20 and a pharmaceutical agent.
30. A method for treating a bladder transitional cell carcinoma comprising administering to a
- 25 subject in need of such treatment the composition of claim 29.

WO 02/055706

1/40

PCT/US02/00647

5' 9 18 27 36 45 54
 A AGG GGC GCT GCC GGG AGC CTC CGG GCC TCA GGG TGT TCC GGG GAG CGG CGC
 63 72 81 90 99 108
 CCC GGG TCT CTG GGC CCA CCC GGC CCG GGC CTC CGA GAG TGG GGG CTG CGC
 117 126 135 144 153 162
 CCG CGG GGT CAG ACA CCT GTT CGG CCC GGC CGG TGG TCG CCG GGG GCC AGG
 171 180 189 198 207 216
 ATG AAA GTG ACC GTG TGC TTC GGC AGG ACG GGC ATC GTG GTG CCC TGC AAG GAG
 M K V T V C F G R T G I V V P C K E
 225 234 243 252 261 270
 GGC CAG CTG CGC GTC GGC GAG CTC ACC CAG CAG GCG CTG CAG CGG TAC CTG AAG
 G Q L R V G E L T Q Q A L Q R Y L K
 279 288 297 306 315 324
 ACC CGG GAG AAG GGT CCT GGT TAC TGG GTG AAG ATT CAT CAC TTA GAA TAT ACA
 T R E K G P G Y W V K I H H L E Y T

FIGURE 1A

```

333 342 351 360 369 378
GAT GGA ATC CTC GAT CCA GAT GAT GTC TTG GCA GAT GAT GTT GAA GAT AAA
D G G I L D P D D V L A D V V E D K
387 396 405 414 423 432
GAC AAG CTG ATT GCT GTG TTT GAA GAA CAA GAA CCA CTC CAC AAG ATT GAG AGC
D K L I A V F E E Q E P L H K I E S
441 450 459 468 477 486
CCC AGT GGA AAC CCT GCA GAT CGG CAG AGC CCA GAT GCT TTT GAG ACA GAA GTG
P S G N P A D R Q S P D A F E T E V
495 504 513 522 531 540
GCC GCC CAA CTG GCT GCA TTT AAG CCA ATT GGT GGG GAG ATT GAA GTA ACC CCT
A A Q L A A F K P I G G E I E V T P
549 558 567 576 585 594
TCT GCT CTA AAA CTA GGC ACT CCA CTG CTG GTG AGG AGA AGC AGT GAC CCA GTG
S A L K L G T P L L V R R S S D P V

```

FIGURE 1B

WO 02/055706

3/40

PCT/US02/00647

```

603      612      621      630      639      648
CCA GGC CCA CCT GCT GAT ACC CAG CCA AGC GCT TCA CAC CCT GGT GGC CAG AGT
-----
P G P P A D T Q P S A S H P G G Q S
-----
657      666      675      684      693      702
CTG AAA CTG GTT CCA GAT TCC ACG CAG AAC TTG GAA CAC AGA GAA GTT TTG
-----
L K L V V P D S T Q N L E D R E V L
-----
711      720      729      738      747      756
AAT GGT GTA CAG ACA GAA CTA CTA ACT TCG CCA AGA ACT AAG GAC ACA TTG AGT
-----
N G V Q T E L L T S P R T K D T L S
-----
765      774      783      792      801      810
GAT ATG ACA AGA ACA GTG GAG ATT TCT GGG GAA GGA GGC CCA TTG GGA ATA CAT
-----
D M T R T V E I S G E G G P L G I H
-----
819      828      837      846      855      864
GTA GTG CCC TTC TTT TCA TCT CTG AGT GGA AGG ATT CTA GGA CTC TTC ATC CGA
-----
V V P F F S S L S G R I L G L F I R

```

FIGURE 1C

WO 02/055706

4/40

PCT/US02/00647

```

873      882      891      900      909      918
GGC AAT GAA GAC AAC AGC AGG TCC AAG CCG GAG GGA CTA TTT CAC GAA AAT GAA
-----
G I E D N S R S K R E G L F H E N E
-----
927      936      945      954      963      972
TGT ATT GTA AAA ATC AAC AAT GTG GAT CTC GTA GAC AAA ACC TTT GCT CAG GCT
-----
C I V K I N N V D L V D K T F A Q A
-----
981      990      999      1008      1017      1026
CAA GAT GTC TTC CGC CAG GCA ATG AAA TCT CCA AGT GTG CTC CTC CAC GTG CTT
-----
Q D V F R Q A M K S P S V L L H V L
-----
1035      1044      1053      1062      1071      1080
CCT CCA CAA AAC CGT GAA CAG TAT GAA AAG TCA GTC ATT GGC TCT CTT AAC ATT
-----
P P Q N R E Q Y E K S V I G S L N I
-----
1089      1098      1107      1116      1125      1134
TTT GGT AAT AAT GAT GGC GTT TTG AAA ACC AAA GTG CCG CCT CCT GTC CAT GGA
-----
F G N N D G V L K T K V P P P V H G
-----

```

FIGURE 1D

WO 02/055706

5/40

PCT/US02/00647

```

1143      1152      1161      1170      1179      1188
AAA TCG GGA CTA AAG ACA GCA AAT CTC ACA GGA ACC GAT AGT CCT GAA ACA GAT
-----
K S G L K T A N L T G T D S P E T D
-----
1197      1206      1215      1224      1233      1242
GCA TCA GCT TCC CTG CAA CAA AAC AAG AGT CCC CGA GTA CCA AGG CTG GGA GGA
-----
A S A S L Q Q N K S P R V P R L G G
-----
1251      1260      1269      1278      1287      1296
AAA CCA TCC TCT CCC TCA CTC TCG CCT CTC ATG GGA TTT GGC AGC AAT AAA AAT
-----
K P S S P S L S P L M G F G S N K N
-----
1305      1314      1323      1332      1341      1350
GCA AAG AAA ATT AAG ATT GAC CTA AAG AAA GGC CCT GAA GGA CTT GGT TTC ACT
-----
A K K I K I D L K K G P E G L G F T
-----
1359      1368      1377      1386      1395      1404
GTG GTT ACC AGA GAC TCT TCC ATA CAT GGT CCC GGT CCC ATT TTT GTA AAA AAC
-----
V V T R D S S I H G P G P I F V K N
-----

```

FIGURE 1E

WO 02/055706

6/40

PCT/US02/00647

```

1413      1422      1431      1440      1449      1458
ATT TTA CCA AAG GGA GCA GCA ATA AAA GAT GGC CGC CTA CAA TCA GGG GAC AGA
-----
I L P K G A A I K D G R L Q S G D R
-----
1467      1476      1485      1494      1503      1512
ATT TTG GAG GTA AAT GGG AGA GAT GTC ACC GGA ACC CAG GAA GAG CTT FTG
-----
I L E V N G R D V T G R T Q E E L V
-----
1521      1530      1539      1548      1557      1566
GCC ATG CTC AGG AGC ACC AAG CAG GGG CAG ACA GCA TCG CTG GTC ATT GCC CGC
-----
A M L R S T K Q G E T A S L V I A R
-----
1575      1584      1593      1602      1611      1620
CAA GAA GGA CAT TTT CTG CCC CGA GAG TTG AAA GGA GAA CCT GAC TGC TGT GCA
-----
Q E G H F L P R E L K G E P D C A
-----
1629      1638      1647      1656      1665      1674
CTC TCT CTG GAG ACA AGC GAG CAG CTC ACC TTT GAG ATC CCC CTG AAT GAT TCA
-----
L S L E T S B Q L T F E I P L N D S
-----

```

FIGURE 1F

WO 02/055706

7/40

PCT/US02/00647

```

1683      1692      1701      1710      1719      1728
GGT TCT GCT GGC CTC GGG GTG AGC TTA AAA GGG AAC AAA TCC AGA GAA ACT GGA
-----
G S A G L G V S L K G N K S R E T G
1737      1746      1755      1764      1773      1782
ACA GAC TTG GGG ATT TTT ATC AAA TCC ATC ATT CAT GGA GGC GCT GCT TTT AAG
-----
T D L G I F I K S I I H G G A A F K
1791      1800      1809      1818      1827      1836
GAT GGT CGT CTG CGA ATG AAT GAC CAG CTG ATT GCA GTT AAT GGG GAA TCT CTT
-----
D G R L R M N D Q L I A V N G E S L
1845      1854      1863      1872      1881      1890
TTG GGA AAG TCC AAC CAC GAA GCT ATG GAA ACA CTT AGG CGG CCA ATG TCC ATG
-----
L G K S N H E A M E T L R R P M S M
1899      1908      1917      1926      1935      1944
GAG GGA AAC ATC CGA GGG ATG ATC CAG TTG GTG ATT CTG AGG AGG CCA GAG AGA
-----
E G N I R G M I Q L V I L R R P E R

```

FIGURE 1G

| | | | | |
|---|------|------|------|------|
| 1953 | 1962 | 1971 | 1980 | 1998 |
| CCA ATG GAG GAT CCT CCA GAG TGT GGG GCA TTT TCC AAG CCA TGC TTT GAG AAC | | | | |
| P M E D P A E C G A F S K P C F E N | | | | |
| 2007 | 2016 | 2025 | 2034 | 2052 |
| TGT CAA AAT GCT GTA ACC ACC TCT AGG CGA AAT GAT AAT AGT ATC CTG CAT CCA | | | | |
| C Q N A V T T S R R N D N S I L H P | | | | |
| 2061 | 2070 | 2079 | 2088 | 2097 |
| CTT GGC ACT TGC AGT CCA CAA GAC AAA CAG AAA GGT CTA TTG CTG CCC AAT GAC | | | | |
| L G T C S P Q D K Q K G L L L P N D | | | | |
| 2115 | 2124 | 2133 | 2142 | 2151 |
| GGA TGG GCC GAG AGT GAA GTT CCA CCT TCT CCA ACA CCA CAT TCT GCT CTG GGA | | | | |
| G W A E S E V P P S P T P H S A L G | | | | |
| 2169 | 2178 | 2187 | 2196 | 2205 |
| TTG GGC CTC GAA GAT TAC AGC CAC AGC TCT GGG GTG GAT TCA GCA GTA TAT TTT | | | | |
| L G L E D Y S H S S G V D S A V Y F | | | | |

FIGURE 1H

WO 02/055706

9/40

PCT/US02/00647

2223 2232 2241 2250 2259 2268
 CCA GAT CAG CAC ATC AAC TTC AGA TCT GTG ACA CCG GCC AGG CAG CCT GAA TCA

 P D Q H I N F R S V T P A R Q P E S

 2277 2286 2295 2304 2313 2322
 ATT AAT TTG AAA GCC TCG AAG AGC ATG GAC CTT GTG CCA GAT GAA AGC AAG GTT

 I N L K A S K S M D L V F D E S K V

 2331 2340 2349 2358 2367 2376
 CAC TCA TTG CCT GGA CAA AAA TCG GAA TCT CCA AGC AAA GAT TTT GGT CCA ACT

 H S L A G Q K S E S P S K D F G P T

 2385 2394 2403 2412 2421 2430
 CTG GGT TTG AAA AAG TCC AGC TCC TTG GAG AGT CTG CAG ACT GCA GTG GCC GAG

 L G L K K S S L E S L Q T A V A E

 2439 2448 2457 2466 2475 2484
 GTC AGG AAG AAT GAC CTT CCC TTT CAC AGG CCC CGG CCG CAC ATG GTT CGA GGC

 V R K N D L P F H R R P R P H M V R G

FIGURE 11

WO 02/055706

10/40

PCT/US02/00647

```

2493      2502      2511      2520      2529      2538
CGA GGC TGC AAT GAG AGC TTT AGA GCA GCC AAT GAC AAA TCC TAC GAT GGA CCT
-----
R G C N E S F R A A I D K S Y D G P
-----
2547      2556      2565      2574      2583      2592
GAA GAA ATA GAA GCT GAC GGT CTG TCT GAT AAG AGC TCT CAC TCT GGC CAA GGA
-----
E E I E A D G L S D K S S H S G Q G
-----
2601      2610      2619      2628      2637      2646
GCT CTG AAT TGT GAG TCT GCC CCT CAG GGG AAT TCG GAG CTA GAG GAC ATG GAA
-----
A L N C E S A P Q G N S E L E D M E
-----
2655      2664      2673      2682      2691      2700
AAT AAA GCC AGG AAA GTC AAA AAA ACG AAA GAG AAG GAG AAG AAA AAG GAA AAG
-----
N K A R K V K K T K E K E K K E K
-----
2709      2718      2727      2736      2745      2754
GGC AAA TTG AAA GTC AAG GAG AAA AAG CGC AAA GAG GAG AAT GAA GAT CCA GAA
-----
G K L K K V K K E K R K E E N E D P E

```

FIGURE 1J

WO 02/055706

11/40

PCT/US02/00647

```

2763      2772      2781      2790      2799      2808
AGG AAA ATA AAG AAG GGC TTC GGC GCC ATG CTG AGG TAT GGG CCT GCT TTG
-----
R K I K K K G F G A M L R Y G P A L
2817      2826      2835      2844      2853      2862
AAG GCA AAG TTG GTT CTC ATT TTG TCT CTC CTG AAA AAA GCG CAC GCT TTT CCT
-----
K A K L V L I L S L L K K A H A F P
2871      2880      2889      2898      2907      2916
CGT CTT CAG CCA AAT GCA TAC GGC TCT CAA TTC TGT GCT CGT TCT CTT TCT GCA
-----
R L Q P N A Y G S Q F C A R S L S A
2925      2934      2943      2952      2961      2970
GAG GCA GAG GAG CTT TTT GGG GAA AGT TAC AGT GAT GAC AGG ACA CTG TCT TAA 3'
-----
E A E E L F G E S Y S D D R T L S

```

FIGURE 1K

```

5'  ATG AAA GTG ACC GTG TGC TTC GGA ACC CGG ACC CGG GTG GTC GTG CCG TGC GGG GAC 54
    M K V T V C F G R T R V V V P C G D
    -----
    63  CAC ATG AAA GTT TTC AGC CTC ATC CAG CAG GCG GTG ACC CGC TAC CGG AAG 108
    G H M K V F S L I Q Q A V T R Y R K
    -----
    117  ATC GCC AAG GAT CCA AAC TAC TGG ATA CAG GTG CAT CGC TTG GAA CAT GGA 162
    A I A K D P N Y W I Q V H R L E H G
    -----
    171  GGA GGA ATA CTA GAC CTT GAT GAC AIT CTT TGT GAT GTA GCA GAC GAT AAA 216
    D G G I L D L D I L C D V A D D K
    -----
    225  AGA CTG GTA GCA GTG TTT GAT GAG CAG GAT CCA CAT CAC GGA GGT GAT GGC 270
    D R L V A V F D E Q D F H H G G D G
  
```

FIGURE 2A

WO 02/055706

13/40

PCT/US02/00647

```

279 ACC AGT GCC AGT TCC ACG GGT ACC CAG AGC CCA GAG ATA TTT GGT AGT GAG CTT 324
-----
T S A S S T G T Q S P E I F G S E L
-----
333 GGC ACC AAC AAT GTC TCA GCC TTT CAG CCT TAC CAA GCA ACA AGT GAA AAT GAG 378
-----
G T N N V S A F Q P Y Q A T S E I E
-----
387 GTC ACA CCT TCA GTC CTT CGA GCA AAT ATG CCT CTT CAT GTT CGA CGC AGT AGT 432
-----
V T P S V L R A N M P L H V R R S S
-----
441 GAC CCA GCT CTA AAT GGC CTC TCC ACT TCT GTC AGT GAT AGT AAT TTT TCC TCT 486
-----
D P A L I G L S T S V S D S N F S S
-----
495 GAA GAG CCT TCA AGG AAA AAT CCC ACA CGC TGG TCA ACA ACA GCT GGC TTC CTC 540
-----
E E P S R K N P T R W S T T A G F L

```

FIGURE 2B

WO 02/055706

14/40

PCT/US02/00647

```

549      558      567      576      585      594
AAG CAG AAC ACT GCT GGG AGT CCT AAA ACC TGC GAC AGG AAG AAA GAT GAA AAC
-----
K  Q  N  T  A  G  S  P  K  T  C  D  R  K  K  D  E  N
-----
603      612      621      630      639      648
TAC AGA AGC CTC CCG GAT ACT AGT AAC TGG TCT AAC CAA TTT CAG AGA GAC
-----
Y  R  S  L  P  R  D  T  S  N  W  S  N  Q  F  Q  R  D
-----
657      666      675      684      693      702
AAT GCT CGC TCG TCT CTG AGT GCC AGT CAC CCA AYG GTG GGC AAG TGG CTG GAG
-----
N  A  R  S  S  L  S  A  S  H  P  M  V  G  K  W  L  E
-----
711      720      729      738      747      756
AAA CAA GAA CAG GAT GAG ACA GAA GAG GAT AAC AGT CGT GTT GAA CCT
-----
K  Q  E  Q  D  E  D  G  T  E  E  D  N  S  R  V  E  P
-----
765      774      783      792      801      810
GTT GGA CAT GCT GAC ACG GGT TTG GAG CAT ATA CCC AAC TTT TCT CTG GAT GAT
-----
V  G  H  A  D  T  G  L  E  H  I  P  N  F  S  L  D  D

```

FIGURE 2C

WO 02/055706

15/40

PCT/US02/00647

```

819      828      837      846      855      864
ATG GTA AAG CTC GTA GAA GTC CCC AAC GAT GGA GGG CCT CTG GGA ATC CAT GTA
-----
M V K L V E V P N D G G P L G I H V
-----
873      882      891      900      909      918
GTG CCT TTC AGT GCT CGA GGC GGC AGA ACC CTG GGG TTA TTA GTA AAA CGA TTG
-----
V P F S A R G G R T L G L L V K R L
-----
927      936      945      954      963      972
GAG AAA GGT GGT AAA GCT GAA CAT GAA AAT CTT TTT CGT GAG AAT GAT TGC ATT
-----
E K G G K A E H E N L F R E N D C I
-----
981      990      999      1008      1017      1026
GTC AGG AIT AAT GAT GGC GAC CTT CGA AAT AGA AGA TTT GAA CAA GCA CAA CAT
-----
V R I N D G D L R N R R F E Q A Q H
-----
1035     1044     1053     1062     1071     1080
ATG TTT GGC CAA GCC ATG CGT ACA CCC ATC AAT TGG TTC CAT GTG GAT CCT GCA
-----
M F R Q A M R T P I I W F H V V P A
-----

```

FIGURE 2D

```

1089      1098      1107      1116      1125      1134
GCA AAT AAA GAG CAG TAT GAA CAA CTA TCC CAA AGT GAG AAG AAC AAT TAC TAT
-----
A N K E Q Y E Q L S Q S E K N N Y Y
1143      1152      1161      1170      1179      1188
TCA AGC CGT TTT AGC CCT GAC AGC CAG TAT ATT GAC AAC AGG AGT GTG AAC AGT
-----
S S R F S P D S Q Y I D N R S V N S
1197      1206      1215      1224      1233      1242
GCA GGG CTT CAC ACG GTG CAG AGA GCA CCC CGA CTG AAC CAC CCG CCT GAG CAG
-----
A G L H T V Q R A P R L N H P P E Q
1251      1260      1269      1278      1287      1296
ATA GAC TCT CAC TCA AGA CTA CCT CAT AGC GCA CAC CCC TCG GGA AAA CCA CCA
-----
I D S H S R L P H S A H P S G K P P
1305      1314      1323      1332      1341      1350
TCC GCT CCA GCC TCG GCA CCT CAG AAT GTA TTT AGT ACG ACT GTA AGC AGT GGT
-----
S A P A S A P Q N V F S T T V S S G

```

FIGURE 2E

WO 02/055706

17/40

PCT/US02/00647

```

1359      1368      1377      1386      1395      1404
TAT AAC ACC AAA ATA GGC AAG AGG CTT AAT ATC CAG CTT AAG AAA GGT ACA
-----
Y N T K K I G K R L N I Q L K K G T
1413      1422      1431      1440      1449      1458
GAA GGT TTG GCA TTC AGC ATC ACT TCC AGA GAT GTA ACA ATA GGT GGC TCA GCT
-----
E G L G F S I T S R D V T I G G S A
1467      1476      1485      1494      1503      1512
CGA ATC TAT GTG AAA AAC AAT CTC CCC CGG GGG GCG GCC AAT CAG GAT GGC CGA
-----
P I Y V K N I L P R G A A I Q D G R
1521      1530      1539      1548      1557      1566
CTT AAG GCA GGA GAC AGA CTT ATA CAG GTA AAT GGA GTA GAT TTA GTG GGC AAA
-----
L K A G D R L I E V N G V D L V G K
1575      1584      1593      1602      1611      1620
TCC CAA GAG GAA GTT GTT TCG CTG TTG AGA AGC ACC AAG ATG GAA GGA ACT GTG
-----
S Q E E V V S L L R S T K M E G T V

```

FIGURE 2F

WO 02/055706

18/40

PCT/US02/00647

```

1629      1638      1647      1656      1665      1674
AGC CTT CTG GTC TTT CGC CAG GAA GAC GCC TTC CAC CCA AGG GAA CTG AAT GCA
-----
S  L  L  V  F  R  Q  E  D  A  F  H  F  R  E  L  N  A
-----
1683      1692      1701      1710      1719      1728
GAG CCA AGC CAG ATG CAG ATT CCA AAA GAA ACG AAA GCA GAA GAT GAG GAT ATT
-----
E  P  S  Q  M  Q  I  P  K  E  T  K  A  E  D  E  D  I
-----
1737      1746      1755      1764      1773      1782
GTT CTT ACA CCT GAT GGC ACC AGG GAA TTT CTG ACA TTT GAA GTC CCA CTT AAT
-----
V  L  T  P  D  G  T  R  E  F  L  T  F  E  V  P  L  N
-----
1791      1800      1809      1818      1827      1836
GAT TCA GGA TCT GCA GGC CTT GGT GTC AGT GTC AAA GGT AAC CGG TCA AAA GAG
-----
D  S  G  S  A  G  L  G  V  S  V  K  G  N  R  S  K  E
-----
1845      1854      1863      1872      1881      1890
AAC CAC GCA GAT TTG GGA ATC TTT GTC AAG TCC ATT ATT AAT GGA GGA GCA GCA
-----
N  H  A  D  L  G  I  F  V  K  S  I  I  N  G  G  A  A

```

FIGURE 2G

WO 02/055706

19/40

PCT/US02/00647

| | | | | | |
|---|------|------|------|------|------|
| 1899 | 1908 | 1917 | 1926 | 1935 | 1944 |
| TCT AAA GAT GGA AGG CTT CGG GTG AAT GAT CAA CTG ATA GCA GTA AAT GGA GAA | | | | | |
| ----- | | | | | |
| S K D G R L R V N D Q L I A V N G E | | | | | |
| ----- | | | | | |
| 1953 | 1962 | 1971 | 1980 | 1989 | 1998 |
| TCC CTG TTG GGC AAG ACA AAC CAA GAT GCC ATG GAA ACC CTA AGA AGG TCT ATG | | | | | |
| ----- | | | | | |
| S L L G K T N Q D A M E T L R R S M | | | | | |
| ----- | | | | | |
| 2007 | 2016 | 2025 | 2034 | 2043 | 2052 |
| TCT ACT GAA GGC AAT AAA CGA GGA ATG ATC CAG CTT AAT GTT GCA AGG AGA ATA | | | | | |
| ----- | | | | | |
| S T E G N K R G M I Q L I V A R R I | | | | | |
| ----- | | | | | |
| 2061 | 2070 | 2079 | 2088 | 2097 | 2106 |
| AGC AAG TCC AAT GAG CTG AAG TCA CCT GGG AGC CCC CCT GGA CCT GAG CTG CCC | | | | | |
| ----- | | | | | |
| S K C N E L K S P G S P P G P E L P | | | | | |
| ----- | | | | | |
| 2115 | 2124 | 2133 | 2142 | 2151 | 2160 |
| ATT GAA ACA CCG TTG GAT GAT AGA GAA CGA AGA ATT TCC CAT TCC CTC TAC AGT | | | | | |
| ----- | | | | | |
| I E T A L D D R E R I S H S L Y S | | | | | |

FIGURE 2H

```

2169      2178      2187      2196      2205      2214
GGG ATT GAG GGG CTT GAT GAA TCG CCC AGC AGA AAT GCT GCC CTC AGT AGG ATA
-----
G I E G L D E S P S R N A A L S R I
-----
2223      2232      2241      2250      2259      2268
ATG GGT GAG TCA GGT AAA TAC CAG CTG TCC CCT ACA GTG AAT ATG CCC CAA GAT
-----
M G E S G K Y Q L S P T V N M P Q D
-----
2277      2286      2295      2304      2313      2322
GAC ACT GTC ATT ATA GAA GAT GAC AGG TTG CCA GTG CTT CCT CCA CAT CTC TCT
-----
D T V I I E D D R L P V L P P H L S
-----
2331      2340      2349      2358      2367      2376
GAC CAG TCC TCT TCC AGC TCC CAT GAT GAT GTG GGG TTT GTG ACG GCA GAT GCT
-----
D Q S S S S H D D V G F V T A D A
-----
2385      2394      2403      2412      2421      2430
GGT ACT TGG GCC AAG GCT GCA ATC AGT GAT TCA GCC GAC TGC TCT TTG AGT CCA
-----
G T W A K A A I S D S A D C S L S P
-----

```

FIGURE 2I

```

2439      2448      2457      2466      2475      2484
GAT GTT GAT CCA GTT CTT GCT TTT CAA CGA GAA GGA TTT GGA CGT CRG AGT ATG
-----
D V D P V L A F Q R E G F G R Q S M
-----
2493      2502      2511      2520      2529      2538
TCA GAA AAA CGC ACA AAG CAA TTT TCA GAT GCC AGT CAA TTG GAT TTC GTT AAA
-----
S E K R T K Q F S D A S Q L D F V K
-----
2547      2556      2565      2574      2583      2592
ACA CGA AAA TCA AAA AGC ANG GAT TTA GGT ATA GCT GAC GAG ACT AAA CTC AAT
-----
T R K S K S M D L G I A D E T K L N
-----
2601      2610      2619      2628      2637      2646
ACA GTG GAT GAC CAG AAA GCA GGT TCT CCC AGC AGA GAT GTG GGT CCT TCC CTG
-----
T V D D Q K A G S P S R D V G P S L
-----
2655      2664      2673      2682      2691      2700
GGT CTG AAG AAG TCA AGC TCG TTG GAG AGT CTG CAG ACC GCA GTT GCC GAG GTG
-----
G L K K S S S L E S L Q T A V A E V

```

FIGURE 2J

```

2709 2718 2727 2736 2745 2754
ACT TTG AAT GGG GAT AAT CCT TTC CAT CGT CCA CGG CCG CGG ATA ATC AGA GGC
-----
T L N G D I P F H R P R P R I I R G
2763 2772 2781 2790 2799 2808
AGG GGA TGC AAT GAG AGC TTC AGA GCT GCC ATC GAC AAA TCT TAT GAT AAA CCC
-----
R G C N E S F R A A I D K S Y D K P
2817 2826 2835 2844 2853 2862
GCG GTA GAT GAT GAT GAA GGC ATG GAG ACC TTG GAA GAA GAC ACA GAA GAA
-----
A V D D D D E G M E T L E E D T E E
2871 2880 2889 2898 2907 2916
AGT TCA AGA TCA GGG AGA GAG TCT GTA TCC ACA GCC AGT CAT CAG CCT TCC CAC
-----
S S R S G R E S V S T A S D Q P S H
2925 2934 2943 2952 2961 2970
TCT CTG GAG AGA CAA ATG AAT GGA AAC CAA GAG AAA GGT GAT AAG ACT GAT AGA
-----
S L E R Q M N G N Q E K G D K T D R

```

FIGURE 2K

```

2979      2988      2997      3006      3015      3024
AAA AAG GAT AAA ACT GGA AAA GAA AAG AAG AAA GAT AAG GAG AAG GAT
-----
K K D K T G K E K K K D R D K E K D
-----
3033      3042      3051      3060      3069      3078
AAA ATG AAA GCC AAG AAG GGA ATG CTG AAG GGC TTG GGA GAC ATG TTC AGG TTT
-----
K M K A K K G M L K G L G L G D M F R F
-----
3087      3096      3105      3114      3123      3132
GGC AAA CAT CGA AAA GAT GAC AAG ATT GAG AAA ACG GGT AAA ATA AAA ATA CAG
-----
G K H R K D D K I E K T G K I K I Q
-----
3141      3150      3159      3168      3177      3186
GAA TCC TTT ACA TCA GAA GAG AAG ATA CGA ATG AAG AAG CAG GAG CAG GAG AGG
-----
E S F T S E E E R I R M K Q E Q E R
-----
3195      3204      3213      3222      3231      3240
ATT CAA GCC AAA ACT CGA GAA TTT AGG GAA CGA CAA GCT CGA GAG CGT GAC TAT
-----
I Q A K T R E F R E R Q A R E R D Y

```

FIGURE 2L

WO 02/055706

24/40

PCT/US02/00647

```

3249      3258      3267      3276      3285      3294
GCT GAA ATT CAA GAT TTT CAT CGG ACA TTT GGC TGT GAT GAT GAG TTA ATG TAT
-----
A E I Q D F H R T F G C D D E L M Y
-----
3303      3312      3321      3330      3339      3348
GGG GGA GTT TCT TCT TAT GAA GGT TCC ATG GCT CTC AAC GCT AGA CCT CAG AGC
-----
G G V S S Y E G S M A L N A R P Q S
-----
3357      3366      3375      3384      3393      3402
CCA CGA GAA GGG CAT ATG ATG GAT GCT TTG TAT GCC CAA GTC AAG AAG CCG CGG
-----
P R E G H M M D A L Y A Q V K K P R
-----
3411      3420      3429      3438      3447      3456
AAT TCC AAA CCC TCA CCT GTA GAC AGT AAC AGA TCA ACT CCT AGC AAT CAT GAT
-----
N S K F S P V D S N R S T P S N H D
-----
3465      3474      3483      3492      3501      3510
CGG ATA CAG CGT CTG AGG CAA GAA TTT CAG CAA GCA AAG CAA GAT GAA GAT GTA
-----
R I Q R L R Q E F Q Q A K Q D E D V

```

FIGURE 2M

WO 02/055746

25/40

PCT/US02/00647

```

3519      3528      3537      3546      3555      3564
GAA GAT CGT CGG ACC TAT AGT TTT GAG CAA CCC TGG CCG AAC GCA CGG CCG
-----
E D R R R T Y S F E Q P W P N A R P
3573      3582      3591      3600      3609      3618
GCG ACG CAG AGC GGG CGA CAC TCG GTG TCC GTG GAG GTG CAG ATG CAG CGG CAG
-----
A T Q S G R H S V S V E V Q M Q R Q
3627      3636      3645      3654      3663      3672
CGG CAG GAG GAG CGC GAG AGC TCC CAG CAG GCC CAG CGC CAG TAC AGC TCT CTG
-----
R Q E E R E S S Q Q A Q R Q Y S S L
3681      3690      3699      3708      3717      3726
CCT CGG CAA AGC AGG AAA AAT GCC AGC TCG GTC TCC CAG GAC TCT TGG GAG CAG
-----
P R Q S R R K N A S S V S Q D S W E Q
3735      3744      3753      3762      3771      3780
AAC TAC TCC CCT GGG GAA GGC TTC CAG AGT GCC AAA GAG AAC CCC AGG TAC TCC
-----
N Y S P G E G F Q S A K E N P R Y S

```

FIGURE 2N

WO 02/055706

26/40

PCT/US02/00647

```

3789      3798      3807      3816      3825      3834
AGC TAC CAA GGC TCC AGG AAC GGC TAC CTG GGA GGA CAT GGC TTC AAC GCC AGG
S Y Q G S R N G Y L G G H G F N A R
-----
3843      3852      3861      3870      3879      3888
GTC ATG CTG GAA ACT CAG GAG CTC CTT CGC CAG GAA CAG AGG CGG AAG GAG CAG
V M L E T Q E L L R Q E Q R R K E Q
-----
3897      3906      3915      3924      3933      3942
CAG ATG AAG AAG CAG CCT CCT TCC GAG GGG CCC AGC AAC TAT GAC TCG TAT AAG
Q M K K Q P P S E G P S N Y D S Y K
-----
3951      3960      3969      3978      3987      3996
AAA GTC CAG GAC CCC AGT TAC GCC CCT CCC AAG GGG CCC TTC CGG CAA GAT GTG
K V Q D P S Y A P P K G P F R Q D V
-----
4005      4014      4023      4032      4041      4050
CCC CCC TCC CCT TCT CAG GTT GCG AGG CTG AAC AGA CTT CAG ACT CCT GAG AAA
P P S P S Q V A R L N R L Q T P E K

```

FIGURE 20

WO 02/055706

27/40

PCT/US02/00647

```

4059      4068      4077      4086      4095      4104
GGG AGG CCC TTC TAT TCC TGA GCA CGC AAA TAA CGG ATG CTT CAT GTC GCG CAA
---
G   R   P   F   Y   S
4113      4122      4131      4140      4149      4158
TAA AAG ACA TTT TCC TAT GAA GAC TTG TAT TTT GGG AGT TTT TTT AAA ACC TCG
4167      4176      4185      4194      4203      4212
ATG GTA CTA TGG AGT ATT TCT GTT GTT GGT ATC AGT GCC TTT AAG CGG TGT AGG
4221      4230      4239      4248      4257      4266
CAA AGA AAT GGA AGG CCT TAA TGT CTT TGC CAC TAT GTC TCA AGT GTC TGT TTC
4275      4284      4293      4302      4311      4320
ATG GAA GGA TTT CCC ACC CTG TGA CAA TCA TCT GTT TGA GGT GTT CAT ATG CTC
4329      4338      4347      4356      4365      4374
TGC GCC TCT CCA CAG TAC CAG GAA TCT CGG CCC TAC TCA TGA GTT GTC CGC GGC
4383      4392      4401      4410      4419      4428
TTG GTT GTA ACA TCC CTG CAC CAC TTG CAG TGA CAA ATT CAC CTG AAG TGG AGG
4437      4446      4455      4464      4473      4482
ATG ACG TGC GGC CCT GTT TCT CCC TCT AAG TTC TCT TAG CTA TGG GAT GAC ATC
4491      4500      4509      4518      4527      4536
TTA GTC TCT GGT GGA GGA AAA GTG GGC GAC ATA CAC CAA AAA TTG GGG CTT TCT

```

FIGURE 2P

WO 02/055706

28/40

PCT/US02/00647

4545 4554 4563 4572 4581 4590
 GGT ACT TCA CAG CAC AGC CAT TTG TCG TAC TTT GTC ATC ACT GTG GTT TTC TCT
 4599 4608 4617 4626 4635 4644
 TTC CTT TCT CAG CTC TTT GTG ACG GGA GAG TCG GTC ATC CTA TTA CAG AAG CTA
 4653 4662 4671 4680 4689 4698
 AGC CAT AGT CCA ACA TTG TTT GGT CAC CAT GGG GGT CCT TTT GTA ACT GCC TTA
 4707 4716 4725 4734 4743 4752
 TGA CTC AAC ATT ACC AAT AAA GTG ATG ATC CTG GTC TGC GTT TAT ACA TAC GCT
 4761 4770 4779 4788 4797 4806
 TGT TCG GTC CTG TTC CTG ACA CGT GGG TTG AGT CAC CAC AGC TCT GTG TGG GGA
 4815 4824 4833 4842 4851 4860
 ACG TGG GAG ACA GGA GTG GCT CCT GCC GGG GGA AGC TGG GCC TGC CAT TGG CCC
 4869 4878 4887 4896 4905 4914
 TGT GTC TAT CAT GAG GGG AGA GCT ANG AAA GAA ATT CTC CTA GGA AGA GCT CAT
 4923 4932 4941 4950 4959 4968
 GGC CCA GTA CAT CCT AGT AAT TAT TTT AAT TAG TTT TTG TTC TGA CAG CTT GTC
 4977 4986 4995 5004 5013 5022
 AGG AAG GGC ACA GAA TGG GAC AGA CAT AAA CCA GAC AGT CAT TTT GAT CTG CTC

FIGURE 2Q

WO 02/055706

29/40

PCT/US02/00647

5031 TCT ACG GTT TTT CAA GTC AGA 5049 GGC AAT TGA TGC TTG TCT AAT GCA TCC ACA CAC 5076
 5085 5094 5103 5112 5121 5130
 TGC ATG TCT GAC TGG CGA TGC CAC GCT CCT AAG TAG TTC TGC CAT GAA ACA TAA
 5139 5148 5157 5166 5175 5184
 AAG ACA AAG GAA AAG CCG TTA CAC ATC ACA CAG AGA ACA TTT TCG GGT CCC ACA
 5193 5202 5211 5220 5229 5238
 GCG GTG GTG GCA GGA AGC TCA CTC TCG CGT CAG TAT TAG AGT GTG TGT GTG GGT
 5247 5256 5265 5274 5283 5292
 CTC GGG GAT CTC GGT GGC TCC CAT CTT CCT TCA TTG TTC TGA ACA TCC TGT ATT
 5301 5310 5319 5328 5337 5346
 GTA AAC CAT GGC TGG GGT GCT AAA GTG CCT GTG AAT CCC GAT GTG GAA AAA GCT
 5355 5364 5373 5382 5391 5400
 GGA GGT GAA AGC TCA GCA TAC CAT GTA TTT ACT TTA AAA ACA GAA AAA AAG ACA
 5409 5418 5427 5436 5445 5454
 TGT ATG GAT ATG TCT ATT TTT TTT TTA TTG GCA CAT TGT ATT TTT GTG TTG ACT
 5463 5472 5481 5490 5499 5508
 TGT TTT TAG AAA TGA TGT GTC CAC ACA CGT ACC CGT GTC TCT TCT GCA TTT CTG

FIGURE 2R

WO 02/055706

30/40

PCT/US02/00647

5517 5526 5535 5544 5553 5562
TGT CAT GGT TCT GTT TCT TAA TCA CGT GCG GTG TCT AAG TGG TGT TAC CAG
5571 5580 5589 5598 5607 5616
TGT ACG CGC AGT GAC CTT GGA TGA CAG TGG CTC TTT CTC ACA GCC TCC CCT GAG
5625 5634 5643 5652 5661 5670
CTG TGA GAA ACA GCT TTC TCT GTA CAT ATG CAA CTC CTA ATA AAA GGC ATA TTT
5679 5688
CTT CCT GTT AAA AAA AAA A 3'

FIGURE 2S

WO 02/055706

31/40

PCT/US02/00647

| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
|-----|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|----------|----------|
| 1 | M | K | V | T | V | C | F | G | R | T | G | I | V | V | P | C | K | E | G | Q | L | R | V | G | E | L | T | Q | Q | A | 1555118 |
| 1 | M | K | V | T | V | C | F | G | R | T | R | V | V | V | P | C | G | D | G | H | M | K | V | F | S | L | I | Q | Q | A | 2582063 |
| 1 | M | K | V | T | V | C | F | G | R | T | R | V | V | V | P | C | G | D | G | R | M | K | V | F | S | L | I | Q | Q | A | 93868778 |
| 1 | M | K | V | T | V | C | F | G | R | T | R | V | V | V | P | C | G | D | G | H | M | K | V | F | S | L | I | Q | Q | A | 98037915 |
| 31 | L | Q | R | Y | L | K | T | R | E | K | G | P | G | Y | W | V | K | I | H | L | E | Y | T | D | G | G | I | L | D | 1555118 | |
| 31 | V | T | R | Y | R | K | A | I | A | K | D | P | N | Y | W | I | Q | V | H | R | L | E | H | G | D | G | I | L | D | 2582063 | |
| 31 | V | T | R | Y | R | K | A | V | A | K | D | P | N | Y | W | I | Q | V | H | R | L | E | H | G | D | G | I | L | D | 93868778 | |
| 31 | V | T | R | Y | R | K | A | I | A | K | D | P | N | Y | W | I | Q | V | H | R | L | E | H | G | D | G | I | L | D | 98037915 | |
| 61 | P | D | D | V | L | A | D | V | E | D | K | D | K | L | I | A | V | F | E | E | Q | E | P | L | H | K | I | E | S | 1555118 | |
| 61 | L | D | D | I | L | C | D | V | A | D | D | K | D | R | L | V | A | V | F | D | E | Q | D | P | H | H | G | D | G | 2582063 | |
| 61 | L | D | D | I | L | C | D | V | A | D | D | K | D | R | L | V | A | V | F | D | E | Q | D | P | H | H | G | D | G | 93868778 | |
| 61 | L | D | D | I | L | C | D | V | A | D | D | K | D | R | L | V | A | V | F | D | E | Q | D | P | H | H | G | D | G | 98037915 | |
| 91 | F | S | G | N | P | A | D | R | Q | S | F | D | A | F | E | T | E | V | A | A | Q | - | L | A | A | F | K | P | - | - | 1555118 |
| 91 | T | S | A | S | S | T | G | T | Q | S | P | E | I | F | G | S | E | L | G | T | N | N | V | S | A | F | Q | P | Y | Q | 2582063 |
| 91 | T | S | A | S | S | T | G | T | Q | S | P | E | I | F | G | S | E | L | G | T | N | N | V | S | A | F | R | P | Y | Q | 93868778 |
| 91 | T | S | A | S | S | T | G | T | Q | S | P | E | I | F | G | S | E | L | G | T | N | N | V | S | A | F | Q | P | Y | Q | 98037915 |
| 118 | I | G | G | E | I | E | V | T | P | S | A | L | K | L | G | T | P | L | L | V | R | R | S | S | D | P | V | - | - | 1555118 | |
| 121 | A | T | S | E | I | E | V | T | P | S | V | L | R | A | N | M | P | L | H | V | R | R | S | S | D | P | A | L | I | G | 2582063 |
| 121 | T | S | E | I | E | V | T | P | S | V | L | R | A | N | M | P | L | H | V | R | R | S | S | D | P | A | L | I | G | 93868778 | |
| 121 | A | T | S | E | I | E | V | T | P | S | V | L | R | A | N | M | P | L | H | V | R | R | S | S | D | P | A | L | I | G | 98037915 |

FIGURE 3A

WO 02/055706

33/40

PCT/US02/00647

230 G L F I R G I E D N S R S K R E G L F H E N E C I V K I M N 1555118
300 G L L V K R L E K G K A E H E N L F R E N D C I V R I N D 2582063
300 G L L V K R L E K G K A E Q E N L F H E N D C I V R I N D 93868778
256 G L L V K R L E K G K A E H E N L F R E N D C I V R I N D 98037915

260 V D L V D K K T F A Q A Q D V F R Q A M K S P S V L L H V L P 1555118
330 G D L R N R R F E Q A Q H M F R Q A M R T P I I W F H V V P 2582063
330 G D L R N R R F E Q A Q H M F R Q A M R A R V I W F H V V P 93868778
286 G D L R N R R F E Q A Q H M F R Q A M R T P I I W F H V V P 98037915

290 P Q N R E Q Y E K S V I G S L N I F G N N - - - D G V L K 1555118
360 A A N K E Q Y E Q L S Q S E K N N Y Y S S R F S P D S Q Y I 2582063
360 A A N K E Q Y E Q L S Q R E M N N Y S P G R F S P D S H C V 93868778
316 A A N K E Q Y E Q L S Q S E K N N Y Y S S R F S P D S Q Y I 98037915

316 T K V P P P V H G K S G L K T A N L T G T D S P E T D A S A 1555118
390 D N R S V N S A G L H T V Q R A P R L N H P P E Q I D S H S 2582063
390 A N R S V A N N A P O A L P R A P R L S Q P P E Q L D A H P 93868778
346 D N R S V N S A G L H T V Q R A P R L N H P P E Q I D S H S 98037915

346 S L Q Q N K S P R V P R L G G K P S P S - - L S P L M G - 1555118
420 R L P H S A H P S G K P P S A P A S A P Q N V F S T V S 2582063
420 R L P H S A H P S T K P P T A P A P A P N V L S T S V G S 93868778
376 R L P H S A H P S G K P P S A P A S A P Q N V F S T V S 98037915

FIGURE 3C

WO 02/055706

34/40

PCT/US02/00647

| | | |
|-----|---|----------|
| 373 | - F G S N K N A K K I K I D L K K K G P E G L G F F T V V T R D | 1555118 |
| 450 | G Y N T K K I G K R L N I Q L K K G T E G L G F F S I T S R D | 2582063 |
| 450 | V Y N T K R V G K R L N I Q L K K G T E G L G F F S I T S R D | 93868778 |
| 406 | G Y N T K K I G K R L N I Q L K K G T E G L G F F S I T S R D | 98037915 |
| 402 | S S I H G P G P I F V K N I L P K G A A I K D G R L Q S I G D | 1555118 |
| 480 | V T I G G S A P I Y V K N I L P R G A A I Q D G R L K A G D | 2582063 |
| 480 | V T I G G S A P I Y V K N I L P R G A A I Q D G R L K A G D | 93868778 |
| 436 | V T I G G S A P I Y V K N I L P R G A A I Q D G R L K A G D | 98037915 |
| 432 | R L L E V N G R D V T G R T Q E E L V A M L R S T K I Q G E T | 1555118 |
| 510 | R L I E V N G V D L V G K S Q E E V S L L R S T K M E G T | 2582063 |
| 510 | R L I E V N G V D L A G K S Q E E V S L L R S T K M E G T | 93868778 |
| 466 | R L I E V N G V D L V G K S Q E E V S L L R S T K M E G T | 98037915 |
| 462 | A S L V I A R Q E G H F L P R E L K G E P D C C A L S L E T | 1555118 |
| 540 | V S L L V F R Q E D A F H P R E L N A E P S Q M Q I I P K E T | 2582063 |
| 540 | V S L L V F R Q E E A F H P R E M N A E P S Q M Q S I P K E T | 93868778 |
| 496 | V S L L V F R Q E D A F H P R E L - - - - - | 98037915 |
| 492 | S E Q - - - - - L T F E I P L N D S G S A | 1555118 |
| 570 | K A E D E D I V L T P D G T R E F L T F E V P L N D S G S A | 2582063 |
| 570 | K A E D E D I V L T P D G T R E F L T F E V P L N D S G S A | 93868778 |
| 513 | K A E D E D I V L T P D G T R E F L T F E V P L N D S G S A | 98037915 |

FIGURE 3D

WO 02/055706

35/40

PCT/US02/00647

508 GLGVSLKGNKSR^ETGTDLGLGIFIKSIIHGGGA 1555118
600 GLGVSVKGNRSKENHADLGLFVKSSIIINGGA 2582063
600 GLGVSVKGNRSKENHADLGLFVKSSIIINGGA 93868778
543 GLGVSVKGNRSKENHADLGLFVKSSIIINGGA 98037915

538 A^EKDGR^LR^MNDQLIAVNGESLLGK^SNI^HE^AM 1555118
630 ASKDGRLRVNDQLIAVNGESLLGK^TNQ^DAM 2582063
630 ASKDGRLRVNDQLIAVNGESLLGK^ANQ^EAM 93868778
573 ASKDGRLRVNDQLIAVNGESLLGK^TNQ^DAM 98037915

568 ETLRR^PM^SEGN^IIRGM^IQL^VIL^LRR⁻ 1555118
660 ETLRRSMSTEGNKRGM^IQL^IVARR^ISK^CNE 2582063
660 ETLRRSMSTEGNKRGM^IQL^IVARR^ISR^CNE 93868778
603 ETLRRSMSTEGNKRGM^IQL^IVARR^ISK^CNE 98037915

592 - - - - - P^EIR^PM^EDP - AECGAFSKPCFE 1555118
690 LKSPGSP^PG^PEL^PIE^TALDDRRR^ISH^SLY 2582063
690 L^ESPGSP^APEL^PIE^TELDDRRR^ISH^SLY 93868778
633 LKSPGSP^PG^PEL^PIE^TALDDRRR^ISH^SLY 98037915

612 NCQNAV^TIS^RRRNDNSILH^PL^GTCS^PQDK^QK 1555118
720 SGIEGLDES^PSR^NAALS^RIMGES^GK^YQL^SSP 2582063
720 SGIEGLDES^PIR^NAALS^RIMGES^GK^CQL^SSP 93868778
663 SGIEGLDES^PSR^NAALS^RIM^G- - -KYQL^SSP 98037915

FIGURE 3E

WO 02/055706

36/40

PCT/US02/00647

642 G L L L P N D G W A E S E - - - - - V P P - - - - - S P 1555118
750 T V N M P Q D D T V I E D D R L P V L P P H L S D Q S S S 2582063
750 T V N M P Q D D T V M I E D D R L P V L P P H L S D Q S S S 93868778
690 T V N M P Q D D T V I I E D D R L P V L P P H L S D Q S S S 98037915

660 T P H S A L G L G L E D - - - - - Y S H S 1555118
780 S S H D D V G F V T A D A G T W A K A A I S D S A D C S L S 2582063
780 S S H D D V G F I M T E A G T W A K A T I S D S A D C S L S 93868778
720 S S H D D V G F V T A D A G T W A K A A I S D S A D C S L S 98037915

676 S G V D S A V Y F P D Q H I N F R S V T P A R - - - - - 1555118
810 P D V D P V L A F Q R E G F G R Q S M S E K R T K Q F S D A 2582063
810 P D V D P V L A F Q R E G F G R Q S M S E K R T K Q F S N A 93868778
750 P D V D P V L A F Q R E G F G R Q - - - - - 98037915

699 - Q P E S I N L K A S K S M D L - V P D E I S R V H S L A G Q 1555118
840 S Q L D F V K T R K S K S M D L G I A D E T K L N T V D D Q 2582063
840 S Q L D F V K T R K S K S M D L G I A D E T K L N T V D D Q 93868778
767 - - - - - I A D E T K L N T V D D Q 98037915

727 K S E S P S K D F G P T L G L K K S S S L E S L Q T A V A E 1555118
870 K A G S P S R D V G P S L G L K K S S S L E S L Q T A V A E 2582063
870 R A G S P M R D V G P S L G L K K S S S L E S L Q T A V A E 93868778
780 K A G S P S R D V G P S L G L K K S S S L E S L Q T A V A E 98037915

FIGURE 3F

WO 02/055706

37/40

PCT/US02/00647

757 V R - K N D L P F H R P R P H M V R G R G C N E S F R A A I 1555118
900 V T L N G D I P F H R P R P R I I R G R G C N E S F R A A I 2582063
900 V T L N G N I P F H R P R P R I I R G R G C N E S F R A A I 93868778
810 V T L N G D I P F H R P R P R I I R G R G C N E S F R A A I 98037915

786 D K S Y D I G P E - E I E A D G L - - - - S D K S S H S G 1555118
930 D K S Y D K P A V D D D D E G M E T L E E D T E E S S R S G 2582063
930 D K S Y D K P M V D D D D E G M E T L E E D T E E S S R S G 93868778
840 D K S Y D K P A V D D D D E G M E T L E E D T E E S S R S G 98037915

809 Q G A L N C E S A P Q G N S E L E D M - - - - - E 1555118
960 R E S V S T A S D Q P S H S L E R Q M N G N O E K G D K T D 2582063
960 R E S V S T S S D Q P S Y S L E R Q M N G D E E K R D K A E 93868778
870 R E S V S T A S D Q P S H S L E R Q M N G N O E K G D K T D 98037915

829 N K A R K V K K T K E K K K E K G K L K V K E - - - - 1555118
990 R K K D K T G K E K K K D R D K E K D K K A K K G M L K G 2582063
990 K K K D K A G K D K K K D R E K E K D K L K A K K G M L K G 93868778
900 R K K D K T G K E K K K D R D K E K D K M K A K K G M L K G 98037915

854 - - - - - K R K E N E D - - - - - 1555118
1020 L G D M F R F G K H R K D D K I E K T G K I K I O E S F T S 2582063
1020 L G D M F R F G K H R K D D K M E K M G R I K I O D S F T S 93868778
930 L G D M F R F G K H R K D D K I E K T G K I K I O E S F T S 98037915

FIGURE 3G

WO 02/055706

38/40

PCT/US02/00647

```

863 -----P-----ERKIKKKG 1555118
1050 EERRIRMKQEQERIQAKTREFFERQARERD 2582063
1050 EEDVIRMKQEERIQAKTREFFERQARERD 93868778
960 EERRIRMKQEQERIQAKTREFFERQARERD 98037915

872 FGAM-----LRYCP-----A 1555118
1080 YAEIQDFHRTFGCDDDELMYGGVSSYEGSMA 2582063
1080 YAEIQDFHRTFGCDDDELMYGGVSSYDGLA 93868778
990 YAEIQDFHRTFGCDDDELMYGGVSSYEGSMA 98037915

882 LKAK-----LVLLLSLKLKKAHA----- 1555118
1110 LNARPPQSPREGHMMDALYAQQVKKKPRNSKPS 2582063
1110 LNARPPQSPREGHMMDALYAQQVKKKPRNSKPS 93868778
1020 LNARPPQSPREGHMMDALYAQQVKKKPRNSKPS 98037915

899 -----FRLQP----- 1555118
1140 EVDSNRS T P S N H D R I Q R L R Q E F Q Q A K Q D E D 2582063
1140 --DSNRS T P S N H D R I Q R L R Q E F Q Q A K Q D E D 93868778
1050 FVDSNRS T P S N H D R I Q R L R Q E F Q Q A K Q D E D 98037915

905 ----- 1555118
1170 VEDRRRTYSFEEQPWPNA R P A T Q S G R H S V S V 2582063
1168 VEDRRRTYSFEEQWS S R P A S Q S G R H S V S V 93868778
1080 VEDRRRTYSFEEQPWPNA R P A T Q S G R H S V S V 98037915

```

FIGURE 3H

WO 02/055706

PCT/US02/00647

```

<110> INCYTE GENOMICS, INC.
      REDDY, Roopa
      TANG, Y. Tom
      BAUGHN, Mariah R.
      KRASNOW, Randi E.

<120> ASIP-RELATED PROTEINS

<130> FC-0032 PCT

<140> To Be Assigned
<141> Herewith

<150> 09/757,781
<151> 2001-01-09

<160> 63

<170> PERL Program

<210> 1
<211> 935
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1555118CD1

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

```

WO 02/055706

PCT/US02/00647

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

WO 02/055706

PCT/US02/00647

```

710          715          720
His Ser Leu Ala Gly Gln Lys Ser Glu Ser Pro Ser Lys Asp Phe
725          730          735
Gly Pro Thr Leu Gly Leu Lys Lys Ser Ser Ser Leu Glu Ser Leu
740          745          750
Gln Thr Ala Val Ala Glu Val Arg Lys Asn Asp Leu Pro Phe His
755          760          765
Arg Pro Arg Pro His Met Val Arg Gly Arg Gly Cys Asn Glu Ser
770          775          780
Phe Arg Ala Ala Ile Asp Lys Ser Tyr Asp Gly Pro Glu Glu Ile
785          790          795
Glu Ala Asp Gly Leu Ser Asp Lys Ser Ser His Ser Gly Gln Gly
800          805          810
Ala Leu Asn Cys Glu Ser Ala Pro Gln Gly Asn Ser Glu Leu Glu
815          820          825
Asp Met Glu Asn Lys Ala Arg Lys Val Lys Lys Thr Lys Glu Lys
830          835          840
Glu Lys Lys Lys Glu Lys Gly Lys Leu Lys Val Lys Glu Lys Lys
845          850          855
Arg Lys Glu Glu Asn Glu Asp Pro Glu Arg Lys Ile Lys Lys Lys
860          865          870
Gly Phe Gly Ala Met Leu Arg Tyr Gly Pro Ala Leu Lys Ala Lys
875          880          885
Leu Val Leu Ile Leu Ser Leu Leu Lys Lys Ala His Ala Phe Pro
890          895          900
Arg Leu Gln Pro Asn Ala Tyr Gly Ser Gln Phe Cys Ala Arg Ser
905          910          915
Leu Ser Ala Glu Ala Glu Glu Leu Phe Gly Glu Ser Tyr Ser Asp
920          925          930
Asp Arg Thr Leu Ser
935

```

```

<210> 2
<211> 1356
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2582063CD1

```

```

<400> 2
Met Lys Val Thr Val Cys Phe Gly Arg Thr Arg Val Val Val Pro
1          5          10          15
Cys Gly Asp Gly His Met Lys Val Phe Ser Leu Ile Gln Gln Ala
20          25          30
Val Thr Arg Tyr Arg Lys Ala Ile Ala Lys Asp Pro Asn Tyr Trp
35          40          45
Ile Gln Val His Arg Leu Glu His Gly Asp Gly Gly Ile Leu Asp
50          55          60
Leu Asp Asp Ile Leu Cys Asp Val Ala Asp Asp Lys Asp Arg Leu
65          70          75
Val Ala Val Phe Asp Glu Gln Asp Pro His His Gly Gly Asp Gly
80          85          90
Thr Ser Ala Ser Ser Thr Gly Thr Gln Ser Pro Glu Ile Phe Gly
95          100          105
Ser Glu Leu Gly Thr Asn Asn Val Ser Ala Phe Gln Pro Tyr Gln
110          115          120
Ala Thr Ser Glu Ile Glu Val Thr Pro Ser Val Leu Arg Ala Asn
125          130          135
Met Pro Leu His Val Arg Arg Ser Ser Asp Pro Ala Leu Ile Gly
140          145          150
Leu Ser Thr Ser Val Ser Asp Ser Asn Phe Ser Ser Glu Glu Pro

```

WO 02/055706

PCT/US02/00647

155 160 165
 Ser Arg Lys Asn Pro Thr Arg Trp Ser Thr Thr Ala Gly Phe Leu
 170 175 180
 Lys Gln Asn Thr Ala Gly Ser Pro Lys Thr Cys Asp Arg Lys Lys
 185 190 195
 Asp Glu Asn Tyr Arg Ser Leu Pro Arg Asp Thr Ser Asn Trp Ser
 200 205 210
 Asn Gln Phe Gln Arg Asp Asn Ala Arg Ser Ser Leu Ser Ala Ser
 215 220 225
 His Pro Met Val Gly Lys Trp Leu Glu Lys Gln Glu Gln Asp Glu
 230 235 240
 Asp Gly Thr Glu Glu Asp Asn Ser Arg Val Glu Pro Val Gly His
 245 250 255
 Ala Asp Thr Gly Leu Glu His Ile Pro Asn Phe Ser Leu Asp Asp
 260 265 270
 Met Val Lys Leu Val Glu Val Pro Asn Asp Gly Gly Pro Leu Gly
 275 280 285
 Ile His Val Val Pro Phe Ser Ala Arg Gly Gly Arg Thr Leu Gly
 290 295 300
 Leu Leu Val Lys Arg Leu Glu Lys Gly Gly Lys Ala Glu His Glu
 305 310 315
 Asn Leu Phe Arg Glu Asn Asp Cys Ile Val Arg Ile Asn Asp Gly
 320 325 330
 Asp Leu Arg Asn Arg Arg Phe Glu Gln Ala Gln His Met Phe Arg
 335 340 345
 Gln Ala Met Arg Thr Pro Ile Ile Trp Phe His Val Val Pro Ala
 350 355 360
 Ala Asn Lys Glu Gln Tyr Glu Gln Leu Ser Gln Ser Glu Lys Asn
 365 370 375
 Asn Tyr Tyr Ser Ser Arg Phe Ser Pro Asp Ser Gln Tyr Ile Asp
 380 385 390
 Asn Arg Ser Val Asn Ser Ala Gly Leu His Thr Val Gln Arg Ala
 395 400 405
 Pro Arg Leu Asn His Pro Pro Glu Gln Ile Asp Ser His Ser Arg
 410 415 420
 Leu Pro His Ser Ala His Pro Ser Gly Lys Pro Pro Ser Ala Pro
 425 430 435
 Ala Ser Ala Pro Gln Asn Val Phe Ser Thr Thr Val Ser Ser Gly
 440 445 450
 Tyr Asn Thr Lys Lys Ile Gly Lys Arg Leu Asn Ile Gln Leu Lys
 455 460 465
 Lys Gly Thr Glu Gly Leu Gly Phe Ser Ile Thr Ser Arg Asp Val
 470 475 480
 Thr Ile Gly Gly Ser Ala Pro Ile Tyr Val Lys Asn Ile Leu Pro
 485 490 495
 Arg Gly Ala Ala Ile Gln Asp Gly Arg Leu Lys Ala Gly Asp Arg
 500 505 510
 Leu Ile Glu Val Asn Gly Val Asp Leu Val Gly Lys Ser Gln Glu
 515 520 525
 Glu Val Val Ser Leu Leu Arg Ser Thr Lys Met Glu Gly Thr Val
 530 535 540
 Ser Leu Leu Val Phe Arg Gln Glu Asp Ala Phe His Pro Arg Glu
 545 550 555
 Leu Asn Ala Glu Pro Ser Gln Met Gln Ile Pro Lys Glu Thr Lys
 560 565 570
 Ala Glu Asp Glu Asp Ile Val Leu Thr Pro Asp Gly Thr Arg Glu
 575 580 585
 Phe Leu Thr Phe Glu Val Pro Leu Asn Asp Ser Gly Ser Ala Gly
 590 595 600
 Leu Gly Val Ser Val Lys Gly Asn Arg Ser Lys Glu Asn His Ala
 605 610 615
 Asp Leu Gly Ile Phe Val Lys Ser Ile Ile Asn Gly Gly Ala Ala
 620 625 630

WO 02/055706

PCT/US02/00647

Ser Lys Asp Gly Arg Leu Arg Val Asn Asp Gln Leu Ile Ala Val
635 640 645
Asn Gly Glu Ser Leu Leu Gly Lys Thr Asn Gln Asp Ala Met Glu
650 655 660
Thr Leu Arg Arg Ser Met Ser Thr Glu Gly Asn Lys Arg Gly Met
665 670 675
Ile Gln Leu Ile Val Ala Arg Arg Ile Ser Lys Cys Asn Glu Leu
680 685 690
Lys Ser Pro Gly Ser Pro Pro Gly Pro Gln Leu Pro Ile Glu Thr
695 700 705
Ala Leu Asp Asp Arg Glu Arg Arg Ile Ser His Ser Leu Tyr Ser
710 715 720
Gly Ile Glu Gly Leu Asp Glu Ser Pro Ser Arg Asn Ala Ala Leu
725 730 735
Ser Arg Ile Met Gly Glu Ser Gly Lys Tyr Gln Leu Ser Pro Thr
740 745 750
Val Asn Met Pro Gln Asp Asp Thr Val Ile Ile Glu Asp Asp Arg
755 760 765
Leu Pro Val Leu Pro Pro His Leu Ser Asp Gln Ser Ser Ser Ser
770 775 780
Ser His Asp Asp Val Gly Phe Val Thr Ala Asp Ala Gly Thr Trp
785 790 795
Ala Lys Ala Ala Ile Ser Asp Ser Ala Asp Cys Ser Leu Ser Pro
800 805 810
Asp Val Asp Pro Val Leu Ala Phe Gln Arg Glu Gly Phe Gly Arg
815 820 825
Gln Ser Met Ser Glu Lys Arg Thr Lys Gln Phe Ser Asp Ala Ser
830 835 840
Gln Leu Asp Phe Val Lys Thr Arg Lys Ser Lys Ser Met Asp Leu
845 850 855
Gly Ile Ala Asp Glu Thr Lys Leu Asn Thr Val Asp Asp Gln Lys
860 865 870
Ala Gly Ser Pro Ser Arg Asp Val Gly Pro Ser Leu Gly Leu Lys
875 880 885
Lys Ser Ser Ser Leu Glu Ser Leu Gln Thr Ala Val Ala Glu Val
890 895 900
Thr Leu Asn Gly Asp Ile Pro Phe His Arg Pro Arg Pro Arg Ile
905 910 915
Ile Arg Gly Arg Gly Cys Asn Glu Ser Phe Arg Ala Ala Ile Asp
920 925 930
Lys Ser Tyr Asp Lys Pro Ala Val Asp Asp Asp Asp Glu Gly Met
935 940 945
Glu Thr Leu Glu Glu Asp Thr Glu Glu Ser Ser Arg Ser Gly Arg
950 955 960
Glu Ser Val Ser Thr Ala Ser Asp Gln Pro Ser His Ser Leu Glu
965 970 975
Arg Gln Met Asn Gly Asn Gln Glu Lys Gly Asp Lys Thr Asp Arg
980 985 990
Lys Lys Asp Lys Thr Gly Lys Glu Lys Lys Lys Asp Arg Asp Lys
995 1000 1005
Glu Lys Asp Lys Met Lys Ala Lys Lys Gly Met Leu Lys Gly Leu
1010 1015 1020
Gly Asp Met Phe Arg Phe Gly Lys His Arg Lys Asp Asp Lys Ile
1025 1030 1035
Glu Lys Thr Gly Lys Ile Lys Ile Gln Glu Ser Phe Thr Ser Glu
1040 1045 1050
Glu Glu Arg Ile Arg Met Lys Gln Glu Gln Glu Arg Ile Gln Ala
1055 1060 1065
Lys Thr Arg Glu Phe Arg Glu Arg Gln Ala Arg Glu Arg Asp Tyr
1070 1075 1080
Ala Glu Ile Gln Asp Phe His Arg Thr Phe Gly Cys Asp Asp Glu
1085 1090 1095
Leu Met Tyr Gly Gly Val Ser Ser Tyr Glu Gly Ser Met Ala Leu

WO 02/055706

PCT/US02/00647

1100 1105 1110
 Asn Ala Arg Pro Gln Ser Pro Arg Glu Gly His Met Met Asp Ala
 1115 1120 1125
 Leu Tyr Ala Gln Val Lys Lys Pro Arg Asn Ser Lys Pro Ser Pro
 1130 1135 1140
 Val Asp Ser Asn Arg Ser Thr Pro Ser Asn His Asp Arg Ile Gln
 1145 1150 1155
 Arg Leu Arg Gln Glu Phe Gln Gln Ala Lys Gln Asp Glu Asp Val
 1160 1165 1170
 Glu Asp Arg Arg Thr Tyr Ser Phe Gln Gln Pro Trp Pro Asn
 1175 1180 1185
 Ala Arg Pro Ala Thr Gln Ser Gly Arg His Ser Val Ser Val Glu
 1190 1195 1200
 Val Gln Met Gln Arg Gln Arg Gln Glu Glu Arg Glu Ser Ser Gln
 1205 1210 1215
 Gln Ala Gln Arg Gln Tyr Ser Ser Leu Pro Arg Gln Ser Arg Lys
 1220 1225 1230
 Asn Ala Ser Ser Val Ser Gln Asp Ser Trp Glu Gln Asn Tyr Ser
 1235 1240 1245
 Pro Gly Glu Gly Phe Gln Ser Ala Lys Glu Asn Pro Arg Tyr Ser
 1250 1255 1260
 Ser Tyr Gln Gly Ser Arg Asn Gly Tyr Leu Gly Gly His Gly Phe
 1265 1270 1275
 Asn Ala Arg Val Met Leu Glu Thr Gln Glu Leu Leu Arg Gln Glu
 1280 1285 1290
 Gln Arg Arg Lys Glu Gln Gln Met Lys Lys Gln Pro Pro Ser Glu
 1295 1300 1305
 Gly Pro Ser Asn Tyr Asp Ser Tyr Lys Lys Val Gln Asp Pro Ser
 1310 1315 1320
 Tyr Ala Pro Pro Lys Gly Pro Phe Arg Gln Asp Val Pro Pro Ser
 1325 1330 1335
 Pro Ser Gln Val Ala Arg Leu Asn Arg Leu Gln Thr Pro Glu Lys
 1340 1345 1350
 Gly Arg Pro Phe Tyr Ser
 1355

<210> 3
 <211> 2968
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 1555118CB1

<400> 3
 aaggggcgct gccgcgagcc tcggggccct aggggtgtcc ggggagcggc gccccgggctc 60
 tctgggccca cccggcccgg gcgtctcccg agagtgggg ctgcgccgcg agggctcagc 120
 acctgttcgg cccggcccgg cgtgtctgcc gggggccagg atgaaagtga ccgtgtcctt 180
 cggcaggacg ggcctcgttg tgccctgcaa ggaggccagg ctgcctctcg gcgagctcac 240
 ccagcagcgc ctgcagcggc acctgaagac cgggggaaag ggtcctggtt actgggtgaa 300
 gattcatcac ttagaatata cagatggagg aatcctggat ccagatgatg tcttggcaga 360
 tgttgttgaa gataaagaca agctgattgc tgtgtttgaa gaacaagaac cactccacaa 420
 gattgagagc cccagtgaaa accttgacga tcggcagagc ccagatgctt ttgagacaga 480
 agtgcgccgc caactgctg catttaagcc aattggtggg gagattgaag caaccctctc 540
 tgcctlaaaa ctaggcactc cactgctggt gaggagagc agtgaccagc tgcacagccc 600
 acctgctgat acccagccaa ggccttcaca ccttgggtgc cagagtctga aactggtgtg 660
 tccagattcc acgcagaact tggagacagc agaagttttg aatggtgtac agacagaact 720
 actaacttcg ccaagaacta agaacacatt gagtgatag acagaacag tggagatttc 780
 tggggaagga ggcaccattg gaatacatgt agtccctctc ttttcatctc tgalggaag 840
 gattctagga ctcttcatcc gaggcattga agacaacagc aggtccaagc gggagggact 900
 atttcacgaa aatgaatgta ttgtaaaaat caacaatgtg gatctogtag acaaaaacct 960
 tgcctcagct caagatgctt tccgcccagg aatgaaatct ccaagtgctc tctccacagt 1020

WO 02/055706

PCT/US02/00647

```

ggttctccca ccaaacccgt aacagatga aaagtcagtc attggctctc ttaacatttt 1080
tggtataaat gatggcgctt tgaaaaccaa agtgcgcct cctgtccatg gaaatcggg 1140
actaaagaca gcaaatctca caggaaccga tagtctgaa acagatgat cagcttccct 1200
gcaacaaac aagagctccc gagtaccag gctgggagga aaacctcct ctcctcact 1260
ctgcctctc atgggatttg gcagcaataa naatgcaag aaaaataaga ttgacctaaa 1320
gaaaggccct gaaggacttg gtttcaactg ggttaccaga gaactctcca taratggtc 1380
cgttccatt tttgtaaaaa acattttacc aaaggagca gcaataaag atggccgct 1440
acaatcagg gcacgaattt tggaggtaaa tgggagagat gtcaccggac gaaccaggga 1500
agagcttctg gcaatgctca gggcaaccaa gcaagggag acagcatgc tggctattgc 1560
ccgcaagaa ggaatttttc tgcctcgaa gttgaaagg aacctgact gctgtgact 1620
ctctctggag acaagcgagc agtcaactt tgagatccc ctgagatgatt caggttctgc 1680
tggctcggg gtgagcttaa aegggacaa atccagaga actggaacg acttgggat 1740
ttttatcaaa tccatcttc atggggcgc tcttttaag gatgtctgc tgcgaatgaa 1800
tgacagctg attgcagta atgggggac tcttttggg aagtcacac acgaagctat 1860
gaaacactt aggcggcaa tgcctatgga ggaacatc cgagggatg tccsgtggg 1920
gattctgagg aggcagaga gaccaatgga ggatcctgca gagtggggg cattttcaa 1980
gcaatgctt gagaactgtc aaaaatgctt aaccactct aggcgaatg ataattgat 2040
ctctgcalca ctggcaactt gcagtccaca agacaacag aaaggtctat tgcctgcca 2100
tgagatgg gtcgagatg aagttccacc ttctccaca caacattctg ctctgggatt 2160
gggctctgaa gattccagc acagctctgg ggtgatcca gcagataat ttccagatca 2220
gcacatcac ttcaagctg tgaccacggc cagggagct gaatcaata atttgaagc 2280
ctcgaagagc atggacottg tgccagatga aagcaagtt caatcattg ctggcaaaa 2340
atcgaaact ccaagcaag attttggtc aactctgggt ttgaaaaagt ccagctcct 2400
ggagatctg cagactcgag tggcagaggt caggaagat gaccttccc ttccagagcc 2460
ccggccgacc atggtctgag gccagagctg caatgagagc tttagagcag ccattgaca 2520
atctcaagat ggaactgaag aaatgaagc tgacggtctg tctgataaga gctctcact 2580
tggccaagga gctctgaatt gtgagctctg cctcagggg aattcggagc tagagacat 2640
gnaaataaa gccagaaag tcaaaaaac gaaagagag gaaagaaaa aggaaaagg 2700
caaatgaaa gtaagggaga aaagcggcaa agagagaaat gaagatccag aaagaaat 2760
aaagaaagag ggtctcggcg ccatgctgag gtatgggct gctttgaag caaagtggg 2820
tctcaatttg tctctctgta aaaaagcgca cgttttctc cgtcttcag caaatgcata 2880
cggctctcaa tctctgctc gttctcttc tgcagagca gaggagcttt ttgggaaag 2940
ttacagtgat gacaggaac tgtcttaa 2968

```

```

<210> 4
<211> 194
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Incyte ID No: 1555118H1

```

```

<400> 4
caacacaca ttctgctctg ggttgggccc tcaagatta cagccacagc tctggggtgg 60
attcagcagt atattttca gacacagaca tcaactccag alctgtgaca ccggccaggc 120
agcctgaac mattaatttg aaagcctcga agagcctgga ccttctgcca gatgaaagca 180
aggttcactc attg

```

```

<210> 5
<211> 533
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Incyte ID No: 7227391H1

```

```

<400> 5
ggaatctgga acaaccagtt tcagactctg gccaccaggt gtgaagcgt tggctgggta 60
tcagcaggtg ggcctggcac tgggtcaact ctctctctca ccagcagtg agtgcctagt 120
tttagagcag aggggtttac ttcaatctcc ccaccaatg gcttcaatgc agccagttgg 180
ggcgcactt ctgtctcaaa agcatctggg cctctcagat ctgcaggggt tccactgggg 240

```

WO 02/055706

PCT/US02/00647

```

ctctcaatct tgtggagtgg ttcttgttct tcaaacacag caatcagctt gtctttatct 300
tcaacaacat ctgccaagac atcatctgga tccaggattc ctccatctgt atattotaag 360
tgatgaatct tcaaccagta accaggaccc ttctcccggg tcttcaggta ccgctggaac 420
gcctgtctgg tgagctcgcc gacgcgaagc tggccctctc tgcagggaac caccgatccc 480
gtcctgcccga agcacacggt cacttctcat ctggccctcg gcgaccacgc cgg 533

```

```

<210> 6
<211> 588
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Incyte ID No: 70150486V1

```

```

<400> 6
ccatggacag gagcggcac ttgggtttc aaaaagccat cattattacc aaaaatgta 60
agagagccaa tgactgaatt tcaactctgt tcaaggttct gtaggagaaq caccgtggagg 120
agcacacttg gagatttcat tgcctggcgg aagacatctt gagcctgagc aaaggtttg 180
tctacagagat caacatltgt gatttttaca atacattcat ttctgtgaaa tagtccctcc 240
cgcttggacc tgctgttgtc tccaatgcct cggatgaaga gtccctagaat ccttccactc 300
agagtgaaa agaaagggcac tacatgtatt cccaatgggc ctccctccco agaaatctcc 360
actgttcttg tcaatacaat caatgtgtcc ttagtctctg gcgaagttag tagtctctgc 420
tgtacacacat tcaaaacttc tctgtcttcc aagttctcgg tggaaatcgg aacaaccagt 480
ttcagactcl ggccaacagg gtgtgaagcg cttggctggg taccagcagg tgggctctgc 540
atgggggtca tgcttctcat cacaagcagt ggagtgccta gttttaga 588

```

```

<210> 7
<211> 481
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Incyte ID No: 70162686V1

```

```

<400> 7
gaccatgtat ggaagagtct ctggtaadca cagtgaaca agtccttca gggcctttct 60
ttaggtcaat cftaatttct ttgcatcttt tattgctgcc aaatccratg agagcgaga 120
gtgagggaga ggaatgtttt cctcccagcc ttgtactcgc gggactcttg tttttgtgca 180
ggaaagctga tgcgtctgtt tcaggactat cggttctcgt gagatttctc gtctttagtc 240
ccgattttcc atggacagga ggccgcaact tggttttcaa aacgccatca ttattaccaa 300
aaatgttaag agggccaatg actgactttt catactgttc acggttttct ggaggaaagc 360
acgtggagga gcaactctgg agatttcatt gcttggcggc agcatctctg agcctgagca 420
aaggttttct ctacagatc cacattgttg atttttcaa taacttcatt ttoggaaat 480
a 481

```

```

<210> 8
<211> 355
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Incyte ID No: 70151326V1

```

```

<400> 8
accagagact cttccataca tggctccggc cccatttttg taaaaaacat ttaccagnag 60
ggagcagcaa taaaagatgg ccgctctaaa tcaaggggaca gaattttgga ggtaaatggg 120
agagatgtca ctggacgaac ccaggaaagc cttgtggcca tgcctcaggag caccaagcag 180
ggggagacag cactcgtggt cattyccccc caagaaggac attttctgco ccagagattg 240
aaaggagaa ctagctctgg tgcactctct ctggagaca gcgagcagct cacttttag 300
aatccccctg gatgattcag gttctgctgg cctcggggtg agctlaaaag ggaac 355

```

WO 02/055706

PCT/US02/00647

<210> 9
<211> 417
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 70154198V1

<400> 9
gctgtaatct tggaggccca atccagagc agaatgtggt gttgggagag gtggaacttc 60
actctggccc catccgtcat tgggcagcaa tagaccttc tgtttgtctt atggacbgca 120
agtccaagat gnatgcagga tactattatc atttcgcta gaggtggtta cagatcttg 180
acagttctca aagcatggct tggaaatgc cccacctct gcaagatcct ccattggctc 240
ctctggcttc ctcaagatca ccaactggat catccctcgg atgttccctt ccattggcat 300
tggccgctca agtgtttcca tagcttcctg gttggacttc ccaaaagaga ttcccatta 360
cggcaatcac tggtcattca ttgcagagc accatccttt aaagcagcgc cttcacy 417

<210> 10
<211> 537
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2084238T6

<400> 10
ccttgacttt caattggccc tttctctttt tcttctcctt ctcttttgtt tttttgactt 60
tcttggcttt atlttccatg tctctctagt cagaattccc ctgaggggca gactcaaat 120
tcagagcttc ttggccagag tggagctctt tctcagacag accgtcagct tctattctct 180
caggtccatc gtaggatttg tcaatggctg ctctaaagct ctcaatggag cctcggcttc 240
gaacatgctg cggccggggc ctgtgaaagg gaaggtcatt ctctctgacc tcggccactg 300
cagcttgagc actctccaag gagctggact ttttcaaacc cagagtggca ccaaaatctt 360
tgettgagca tcccgatttt tgtccagcca atgagtgaac cttgcttcca tctggcaca 420
ggtccatgct cttcagagct tccaatttaa ttgattcagg ctgctggcgg gtgtcacaga 480
tctgaagtgg atgctgtgat ctggaaaata tactgtctgaa tccaaccagc agtgtgg 537

<210> 11
<211> 498
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 70155923V1

<400> 11
cctttctgga tcttcattct cctcttggcg ctttttctcc ttgaacttea atttgccctt 60
tcccctttcc tctctctctt cttctgtttt tttgaacttc ctggttaat ttcccagtc 120
ctctagctcc gaattccctt gagggcgaga ctccaatcc agagctcctt gcccagatg 180
aaagctctta tcaagacagc cgtcagcttc aattctttaa ggtccatcgt aggatttgc 240
aatggctgct ctaaaagctt cattgaagcc tggcctcaga accatgtgag gccggggcct 300
ttgaaaggya aggtcattct tectgaactc gggcaactgca gtttgagac tctccaagga 360
gctggacttt tctaaaccca gagttggccc aaaatctttg cttggaastl ccgatttttg 420
tccagccaat gactgaact tgccttcttc tggcaccagg tccatgccct ccggggcttt 480
caaatcaatt ggattcag

<210> 12
<211> 498
<212> DNA
<213> Rattus norvegicus

WO 02/055706

PCT/US02/00647

<220>
 <221> misc_feature
 <223> Incyte ID No: 702457609T1

<400> 12
 ttggctttat ttccacaca catctagctc aggytttocc tgaggggcag actcacaatt 60
 cagaactctg tggcccaggt gagagctctt atcagacaga ccatcagctt ctgcctcttc 120
 tggcccgctg taagatttgt caatggctgc tcgagaagct ctcaattgca cccggccgt 180
 cgcaccatgt gtggtctggg tctgtggaag ggcagatcat tcttcttgac ttcagccaca 240
 gaagtctgta ggtctccaa ggaactgac tttttcagge ccaggsttg accaaaaatc 300
 ttgcctggag agtctgacct gcatcagccc agtgactgga ctctgcttc gtctgaca 360
 aggtccatgc tottgagggc ttccaggtta atgattcgg gctgcccagc ctggtgtcac 420
 aattctgaaa ttgacatgtt gatctgaaa atatcctggt gaatccactc tagagctgtg 480
 actgagctc tcaaggcc 498

<210> 13
 <211> 460
 <212> DNA
 <213> Rattus norvegicus

<220>
 <221> misc_feature
 <223> Incyte ID No: 702458746T1

<400> 13
 gtttggcccg aagtggagac tcttatcaga cagaccatca gcttotgect cttctggccc 60
 gtctaaagat ttgtcaatgg ctgctggaaa gctctcattg cagcccggcc ctgcaccatc 120
 gtgtgtcttg ggtctgtgga agggcaactc attctctcgg acctcagaca cagcagctcg 180
 taggcctctc aaggcaactg actttttcag gcccaagggt ggcaccaaat ctttgcctgg 240
 agagctcagc ctggatcag ccagtgactg gactttgctt tctgttgga caaggtccat 300
 gctcttgagc gcttccaggt taattgatcc ggcctgccc actggtgtca caattctgaa 360
 attgacatgt tgatctgaaa aatatcctgt ggaatccact ctgagactgt gactgaagtc 420
 tccaaggccc cattccagac tgggatgtgg tggcggggac 460

<210> 14
 <211> 245
 <212> DNA
 <213> Rattus norvegicus

<220>
 <221> misc_feature
 <223> Incyte ID No: 701335936H1

<400> 14
 aaaggcccag aaggacttgg attcactgtg ggaactagag attcttggat acatggacct 60
 ggtcccattt ttgtaaaaaa taccctacca aaggagcag cagttaagga tggccgacct 120
 caatcaggag acagaatttt agaggttaat ggcagagatg tccaggaag accccaggaa 180
 gaacttggc ccatgctgag ggcactaag caggagaga cggatcactc ggtcatgtcc 240
 cgcca 245

<210> 15
 <211> 260
 <212> DNA
 <213> Rattus norvegicus

<220>
 <221> misc_feature
 <223> Incyte ID No: 700639694H1

<400> 15
 attttggctc aaccctgggc ctgaaaaagt ccagttcctt ggagagccta cagactgctg 60
 tggctgaagt caggaagaat gatctgacct tccacagacc cagaccacac atggtgcag 120
 gccggggctg caatgagagc ttccagagag ccattgacaa atcttaagac gggccagaag 180

WO 02/055706

PCT/US02/00647

aggcagaagc tgatggtctg tctgataaga gctctcgctc gggccacaca gctctgaatt 240
gtgagtctgc cctcaggga 260

<210> 16
<211> 211
<212> DNA
<213> Rattus norvegicus

<220>
<221> misc_feature
<223> Incyte ID No: 700639694F6

<400> 16
gatctgccct tccacagacc cagaccacac atggtgcgag gccggggctg caatgagagc 60
ttccgagcag coattgacaa atcttaacgac gggccagaag aggcagaagc tgatggtctg 120
tctgataaga gctctgctc gggccacaca gctctgaatt gtgagtctgc cctcaggga 180
aacctgagc tagatgatgt ggaataataa g 211

<210> 17
<211> 276
<212> DNA
<213> Rattus norvegicus

<220>
<221> misc_feature
<223> Incyte ID No: 701191467H1

<400> 17
ctatagtcca caagataaaa ggaagacct attgctccc agtbatgggt gggctgagaa 60
tgaagtaccg cgttccccgc caccacatcc cagtctgaa tggggccttg aagacttccg 120
tccagctctt agagtggatt cacaggatatt ttccagatc accatgtaa ttccgaatt 180
gtgacaccag tggggcagcc cgaatcaatt aacctgaaag cctccaggag catgacctt 240
gtgccagagc aaagcaaat cagtcactg gotgat 276

<210> 18
<211> 555
<212> DNA
<213> Canis familiaris

<220>
<221> misc_feature
<223> Incyte ID No: 702771158H1

<400> 18
aagatgttcc accttccct cctcagacc acagagtggc ctaccaggaa atgggtgagc 60
caggcccca gggggcagcc cagaccagta cccctaccgc gccccaggatc ccaggcagaa 120
gaaacctatg actgagcag tctagctgaa taccacagag ctccagccag cccagaaggg 180
cgaactctga catcaacttc gccctcccta gactcttaag gcttctctc tgtccagaag 240
tctccatggt acagataggt ttgtctcacc gaggttgcaa cacttgactg ctgaccagag 300
gggaaaagga ggggcagga ggggtgggga gaaaggacag gactcaca gacagcactg 360
cctgggattt gaaatatgtt tagaatctct cagttagggg caagtggcag ttgagcagcc 420
aaataccat ggaacatag cacacggggc ctccctggcg tacaccattc ataacttctc 480
catctctaa gttctgtgtg gaatctagaa taacagggtt cagactctgg ccactggggt 540
gtgagcact tggct 555

<210> 19
<211> 257
<212> DNA
<213> Mus musculus

<220>
<221> misc_feature
<223> Incyte ID No: 701266650H1

WO 02/055706

PCT/US02/00647

```

<400> 19
gaaaaatgca aaaaaaatta agattgacct aaagaaggcg cctgagggag ttggattcac 60
tgtggaacca gagattcttc tatacatggg cotggctcca ttctgtcaa aacatctta 120
ccnaaggag cagcagtaaa ggatggccgc ctacaatnag gggacagaat ttggaggtta 180
aatggcagag atgttacagg aagaacccaa gaagaactcg tggccatggt aagagacacc 240
aagcaggag agacagt
257

<210> 20
<211> 5689
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2582061cb1

<400> 20
atgaaagtga cngtgtggett cggacggacc cgggtggctg tgcctgccc ggaagggcac 60
atgaaagtgt tcaagctcat ccagcagcgg ctgacccggc agcgggaagg cctcggccaa 120
gatcccaact actggataca ggtgcctgdc ttggaacctg gagaaggag aactactagc 180
cttgatgaca ttcttltgtg tgtagcagac gataaagaca gactggtagc agtgtttgat 240
gagcaggatc cacatcaggg aggtgatggc accagtgcca gttccagggg taccagagc 300
ccagagatat ttggtagtga gcttggcaac aacaatgtct cagccttcca gcttaccaca 360
gcaacaagtg naattgaggt cacaccttca gtccttcgag caaatatgcc tcttcatggt 420
tctctcgagc gtgancaccg tetaattggc ctctccactt ctgtcagtga tagtaatttt 480
tctctcgagc agccttcaag gaaaaatccc acacgctggt caacaacagc tggcttccct 540
angagaaca ctgctggggg tcttaaaacc tggcagcagg agaagatga aactacaga 600
agcctccccc gggatctag taactgtct aaccaattc agaaaccaga acaaggtgag 720
tctctgagtg ccagtcacc aatggtgggc aagtggctgg gecatgctga caggggtttg 780
gatggagcag aagaggtata cagtcgtggt gaacctgttg gecatgctga caggggtttg 780
gagcatatcc ccaacttttc tctggatgat atgtaaaag tctgagaagt ccccaaccat 840
ggagggcctc tggaaatcca tgtagtccct ttcagtgcct gagggggcag aacctgggg 900
ttattagtaa aacgattgga gaaagtggt aaagctgaac atgaaaatc tttctctgag 960
aatgattgca ttgtcaggat taatgatggc gaacctgaaa atagaagatt tgaacaagca 1020
caacatagt ttgcacaagc catgctgaca cccatcattt ggttccatgt ggttccctga 1080
gcaaaataag agcagtatga acaactatcc caaagtgaga agaacaatta ctattcaagc 1140
cgttttagcc ctgacagcca gtaacttggc aacaggggtg tgaacagtgc agggcttca 1200
acggtgcaag gagcccccgc actgaaccac ccgctggagc agtatagctc tcaatcaaga 1260
ctacctcata ggcacacccc ctccggaaaa ccaactccg ctccagctcc ggaacctcag 1320
aatgattata gtacagctgt aagcagtggt tataacacca aaaaatagg caagaggctt 1380
aatatccagg ttaagaaagg tacagaaggt ttgggatcca goatcacttc cagagatga 1440
acaatagggt gctcagctcc aatctatgtg aaaaacatc tccccgggg ggcggccatt 1500
caggtggccc gacttaaggc aggagaacga cttatagagg taatggagt agatttagt 1560
ggcaaatccc aagaggaagt tgttctgctg ttgagaagca ccaagatgga agaaactgt 1620
agcctctctg tctttggcca ggaagcggcc ttcccacca gggactgca tgcagagcca 1680
agccagatgc agattccaaa agaaacgaaa gcagaagctg agtatatgt tcttaccact 1740
gatggcacca gggatattct gacattttaa gttccactta atgattccag atctgacagg 1800
cttgggttca gtgtcaagg taaccggcca aaagagaacc accagattt gggatcttt 1860
gtcaagtcct tattaatgg aggagcagca tctaaagatg gaaggttcc ggtgatgat 1920
caactgatag cagtaaatgg afaatccctg ttggccaaga caaaccaaga tgcattgga 1980
accctaagaa ggtctatgtc tactgaagggc aataaacgag gaatgatcca gcttatgt 2040
gcaagagaa taagcaagtg caatgagctg aagtcacctg ggagccccc tggacctgag 2100
ctgcccattg aaacagcgtt ggatgataga gaacgaagaa ttcccattc cctctacagt 2160
ggatltgggg gccttgatga atcggccagg agaaatgct cctccagtag gatactgggt 2220
gagtaaggtt aatccagct gtcccotaca gtgaatatgc ccaaatgga cactgcaat 2280
atcagagatg acaggttggc agtgcttctc ccacatctct ctgaccagtc ccttccagc 2340
tcccatgatg atgtggggtt tgtgagggca gatgctgga cttgggcca gcttgcact 2400
agtgaattcag ccgactgctc tttgatcca gatgttgatc cagtctttgc ttttcaagc 2460
gaagattttg gacgtcaagag tatgtcagaa aaacgccaaa agcaatttcc agatgocagt 2520
caattggatt tctttaaacc acgaaaatca aaagcatgg atttaggtat agctgacag 2580
actaaactca atcaagtgga tgaccagaaa gcaggttctc ccagcagaga tctgggtcct 2640
tccctgggtc tgaagaagtc aagtcctgtg gagagcttgc agaccagat tgcagaggt 2700
actttgagtg gggatattcc ttccatctgt ccaagggccc ggaatantcg aggcagggga 2760

```

WO 02/055706

PCT/US02/00647

```

tgcattgaga gcttcagagc tgcattgaga aatatttatg ataaaccocg ggtagatgat 2820
gatgataaag gcatggagac ctgggaagaa gacacagaag aaagttcaag atcaggaga 2880
gactctgtat ccacccagag tgatcagcct tcccactctc tggagagaca satgaatgga 2940
aaaccagaga aaggtgataa gactgataga aanaaggata aaactggaaa agaaaagaa 3000
aaagatagag ataaggagaa ggataaaatg aaagcacaaga agggatgct gaagggcttg 3060
ggagacatgt tcaagtttgg caaacatcga aaagatgaca agattgagaa aacgggtaaa 3120
ataaaaaatc agaatccctt tacatcagaa gagagagaga tacgaatgaa gcaggagaa 3180
gagagatttc aagccaaaac tccagaattt agggacagac aagctcgaga gcgtgactat 3240
gtgaaattc agattttca tgggaacatt ggcctgtgat atgagttaat gtatggggga 3300
gtttctctct acgaaggttc catggctctc aacgttagac ctccagagccc acgagaaggg 3360
catatgatgg atgctttgta tggccaagtc aagaagccgc ggaattocaa accctcaact 3420
gtagacagta aagatcagac tcttagcaat catgacaga tacagagctt gaagcaagaa 3480
tttccagcaag caaagcaaga tgaagatgta gaagatcgtc gggcgaacta tagttttgag 3540
caaccctggc cgaagcagac gcccggcagc cagagccggc gacactcggg gtcctggag 3600
gtgcaatgac agggagagc gcagagagag cgcgagagct cccagcagcc ccagggcag 3660
tacagctctc tgcctcggca aagcagaaa aatgccagct cggctcccaa ggactcttg 3720
gagcagaact acctccctgg ggaaggtctc cagagtccca aagagaacc caggtaactc 3780
agctaccaag gctccaggaa gggctacctg ggaggaacat gcttcaacgc cagggtcatg 3840
ctggcaactc agaggtcctc tggccaagaa cagagccgga aggagcagca gatgaagag 3900
cagctcctt cgaaggggac cagaactat gactgtata aagaagctcc ggaaccagc 3960
taccocctc ccaaggggac ctccggcaa gatgtgccct cctcccttc tcaagtgag 4020
aggtcgaaca gacttccagc tcttagaaa gggagccctc tctattcctg agcagcaaa 4080
taacggatgc ttcattgtgc gcaataaag acatttctct atgaagact gtatttggg 4140
gcttttttca aaacctgat ggtcatgatg agtattctgt ttgttgatc cagtgcctt 4200
aagcgtgta ggaagaaa tggagggctc taatgtctt gccactatg ctcaaggtc 4260
gtttctatg aagattctc caccctgta caatcactg tttgaggtg tcatatgctc 4320
tgccctctc caacgtacca ggaattctgg cctactcat gactgtccg cggctggtt 4380
gtacatccc tgcaccactt gcagtgcaaa attcactga agtggaggt gactgaggg 4440
cctgtttctc cctcaagtt ccttagctc tgggatgaca tcttagctc tggtyggga 4500
aaagtggcg acatacaaca aaattgggg ctttctgta cttaacaga cagccattg 4560
tgctacttg tcatcactgt ggtttctct tctctctc agctcttct gacggagag 4620
teggctatcc tattacagaa gctaaagcct agtccaaact ttttggta ccatggggg 4680
cctttgtaa ctgcttatg actcaactt acaataaag tgatgatct ggtctcggt 4740
tatacatacg ctgttctggt cctgttctg acacgtgggt tgagtcaaa cagctctgt 4800
tgggaaact gggagacag agtgcctct gcgggggaa gctgggctg ccattggcc 4860
tgttctatc atgagggag agctaagaaa gaaattctc taggaagag tcatggcca 4920
gtacatccta gtaattatt taattagttt ttgttctgac agcttgcag gaagggcaa 4980
gaaaggaca gagaataacc agcagctat ttgatctgc tctcaggt tttcaagtc 5040
agagcaat gatgctgtc taatgcatc acacactca tgtctgact gcatgcaac 5100
gctcctaagt agtctgcaca tgaacataa aagacaagg aaagccgtt acactata 5160
cagacaacat ttccgggtcc caacgggtg gtggcagaa gctcactc gctcagat 5220
tagagtgtg gtgtgggtct cgggatctc ggtggctccc atctctctc atgtttota 5280
acatcctgta ttgtaaaaca tggctgggt gctaaagtc ctgtgaatc cagtgtgaa 5340
aaagctggag gtgaaagctc agcatcact gtatttact taaaaacaga aaaaagaca 5400
tgtatggata tgtctatatt tttttattg gcacattgta ttttgtgt gactgttt 5460
tagaatgat ggtccacac acgtaccggt gctctctctc cattctctg tcatggtct 5520
gtttctaat cngtgggc ggtcttagg tgggttacc agtgaacgc cagtgaact 5580
ggatgacagt ggtctttct cccgctctc cctgagctg gaaaaacag tttctctgta 5640
catatgcac tcttaataaa aggcataatt ctctctgta aaaaaaaa 5680

```

```

<210> 21
<211> 249
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2582063H1

```

```

<400> 21
caggggtggt ggcaggaagc tcaactctgc gtcagtatta gactgtgtgt ggggtctc 60
gggatctagg ggcctcccat ctctctctat gttctgaaac atcctgtatt gtaaacctg 120
gctgggggag taaagtccct gtagaatccg atgtggaaaa agctggaggt gaaagctcag 180

```

WO 02/055706

PCT/US02/00647

cataccatgt atttacttta aaaacagaaa aaaagacatg tatggatag tctatttttt 240
ttttattgg 249

<210> 22
<211> 549
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7246093H1

<400> 22
cccgggtggt cgtgccgtgc ggggacggcc acatgaaagt ttccagcctc atccagcagg 60
cggtagcccg ctaccggaag gccatcgcca aggatccaaa ctactggata caggtagcgc 120
gcttggaaaca tggagatgga ggaatactag accttgatga cattctttgt gatgtagcag 180
acgataaaga cagactggta gcagtgtttg atgagcagga tccacatcac gggagtgatg 240
gcaccagtgc cagttccaag ggtaccacga gccnagagat atttggtagt gggcttgcca 300
ccaacaatgt ctaagccttt cagccttacc aagcaacaag tgaattgag gtcacacctt 360
cagcctctcg agcaaatatg cctctctcag ttogaagag tagtgaccac gctcaatctg 420
gcctctccac ttctgtcagt gatgtaatt ttctctctga agacacctca aggaasaatc 480
ccaacagctg gtcaacaaca gctggcttcc tcaagcagaa cactgctggg agtcctaaa 540
cctgcgaca 549

<210> 23
<211> 502
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7978420H1

<400> 23
ggggagocca gagatatttg gtatgagctc tggcaccac aatgtotoag cctttcagcc 60
ttaccgaagca acaagtgaat ttgaggctac accttcagtc cttcgagcaa atatgcctct 120
tcwtgttoga cgcagtatga acccagctct aattggcctc tccaactctg tcagtgtatg 180
taatttttcc tctgagagag ctccagggaa aattcccaca cgtcgttcca caccagctgg 240
cttccccaag cagaacactg ctgggagctc taacaactgc gacaggaaga aagatgaaa 300
ctacaagaagc ctcccgggg atactagttaa ctggctaac caatttcaga gagcacatgc 360
tcgctcgtct ctgagtcca gtcacccaat ggtgggcaag tggctggaga aacaagaaca 420
ggataggatc gggacagaag aggataacag tcgtgtttaa cctgttggac atgctgacac 480
gggtttggag catataccca ac 502

<210> 24
<211> 611
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 55040412H1

<400> 24
gctgtgcata aagcaggga agaaaggat aacagctgtc ttgaacctgt tggcaatgct 60
gacacgggtt tggagcatab acccaacttt tctctgatg atatgttaa gctcgcagaa 120
gtccccaacg atggagggcc tctgggaatc catgtactgc ctttcagtgc tcaagcggc 180
agaacctgg ggttattagt aaaaagcttg gagaagctg gtaaacctga acatgaaat 240
ctttttcgtg aqaatgattg cattgtcagg ataatgatg gcgaacctcg aeatgaaaga 300
tttgaacaag cacacatata gtttgcacaa gccatgcgta caccatcat ttggttccat 360
gtggttctct cagcaaataa agagcagtat gaacaactat cccaagtga gaagaacat 420
tactattoa gccgttttag cctgacacgc cagtatattg acaacaggag tctgaacagt 480
gagggctgc acacggtgca gagagcaacc cgaactgaacc acccgcctga gcagatagac 540

WO 02/055706

PCT/US02/00647

tctactcaa gactacctca tngcgcacac cctcgggaa aaccaacato cgtccatcc 600
tcctggacag c 611

<210> 25
<211> 462
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2929484F6

<220>
<221> unsure
<222> 405, 441
<223> a, t, c, g, or other

<400> 25
gggcaaccg actgaaaccac cggcctgagc agatagactc tcactcaaga ctactcata 60
gggcaaccac ctgggaaaaa ccaccatccg ctccagcctc ggcaactcag aatgattta 120
gtacgactgt sagcagtggt tataaaccca aaaaatagg caagaggctt aatatccagc 180
ttaaagaagg tacagaaggt ttgggattca gcataccttc cagagatgta acatagagtg 240
gtcagctcc aatctatgtg aaaaacatc tccccgggg ggccggcatt caggatggcc 300
gacttaaggc aggagacaga ctatatagg taatagggt agatttagtg gccaatccc 360
aagaggaagt tgtttcgtg ttggaagca ccaagatga aggantgta gttctgggc 420
tttgccagg aagaccctc naaccaagg aactgaatgc ag 462

<210> 26
<211> 375
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 5627320R8

<400> 26
acactcggg gttctctttt gaccggttac ctttgacat gacaccaagg cctgcagatc 60
ctgaaatcatt aagtgggact tcaatgtca gaattcccl ggtgcatca ggtgtasga 120
caatctctc atctctcgtc ttctgttctt ttggaatctg catctggctt ggatctgat 180
tcagttccct tgggtggaag gcgtctctct ggcgaagac cagaaggctc acagtctctt 240
ccatcttggg gcttctcaac agcgaacaaa ctctctcttg ggatttgccc actaaatcta 300
ctccatttac ctctataagt ctgtctctg acttaagtcg gccatcctga atggccgccc 360
cggggggaga atgtt 375

<210> 27
<211> 543
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 3209128F6

<400> 27
cgcctccac ccaagggaa cgaagcaga agatgaggat attgtctta caectgatg 60
caccagaaa tttctgact ttgaagtcc actaatgat tcaggatctg caggccttg 120
gtccagtgc aaagtaacc ggtcaaaaga gaaccacga gatttggaa tctttgcaa 180
gtccattatt aatggaggag cagcatctaa agatggaag ctctgggtga atgatcaact 240
gatagcgtta aatggagaat cctgttggg caagacaac caagatgcca tggaaacct 300
aagaaggtct atgtctactg aagyaataa acgaggaat atccagctta ttgttgaag 360
gagaaacagc aagtcaatg agctgaagtc accggggag ccccctggac ctgagctgcc 420
cattgaaca gccttggatg atagasaacg aagaattcc cttctctct acagtggat 480

WO 02/055706

PCT/US02/00647

tgaggggctt gatgatcgg ccagcagaaa tgctggcctc agtaggataa tgggtgagtc 540
agg 543

<210> 28
<211> 220
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 349249H1

<220>
<221> unsure
<222> 167
<223> a, t, c, g, or other

<400> 28
aatatgccc aagatgacac tgtcattata gaagatgaca ggttgccagt gcttctccca 60
cattctctg accagtccct tttcagctcc catgatgatg tgggtttgt gaccgcagat 120
gctggtaact ggccaagc tgcaatcagt gattcagccg actgctnttt gactccagat 180
gttgatccag ttcttgcttt tcaacgagan ggatttgac 220

<210> 29
<211> 613
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7019961H1

<400> 29
gtgcttctc cacatctctc tgaccagtc tttccagct cccatgatga tgtggggttt 60
gtgacggcag atgctggatc ttgggccaag gctgcaatca gtgattcagc cgactgctct 120
ttgagtcacg atgttgatcc agttcttgct tttcaacgag aaggatttgg acgtcagagt 180
atgtcagaaa aacgcaaaa gcaattttca gatgccagtc aattggattt cgtlaaaaa 240
cgaaatcna aagcatgga tttaggtata gctgacgaga ctaaactcaa tacagtggat 300
gaccagaag cagttctcc cagcagagat gtgggtcct cctgggtct gaagaatca 360
agctcgttgg acagtctgca gaccgcagtt gccaggtga ctttgaatga gstatctct 420
ttccatcctc cagggccgag gataatcaga ggcagggat gcaatgagag cttcagagct 480
gccatgaca aatcttatga taaaccgcg gtgagatgat atgatgagg catggagacc 540
ttggaagag acacagaaga cagttcagc tcaggagag agtctgtatc cccagccagc 600
atcaggcttc cac 613

<210> 30
<211> 249
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 6303175H2

<400> 30
tcggatcta gaacaagaa aagaagaag atagacataa ggagaaggat aaatgatag 60
ccaagaagg aatgctgaag gcttggggg acatgthcag gtttggcaa catcgaaag 120
atgacaagat tgagaaaag ggtaaaataa aaatacagga atcctttaca tcagaagagg 180
agaggatacg atgaagcag gagcagaga ggattcagc caaaactcga gaatttagg 240
aacgacaag 249

<210> 31
<211> 501

WO 02/055706

PCT/US02/00647

```

<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2549906F6

<220>
<221> unsure
<222> 137, 164, 463
<223> a, t, c, g, or other

<400> 31
aggagaagga taaatgaaa gccagaagg gaatgctgaa gggcttggga gscatgttca 60
ggtttggcaa acatcyaaaa gatgacaaga ttgagaaac gggtaaaata aaatcacag 120
aatcccttac atcagangag gagaggatc gaatgaagca ggancaggag aggattcaag 180
cnaaactcg agaatttagg gaacgacaag ctgagagcg tgactatgct gaaattcaag 240
attttcatcg gacatttggc tgtgatgatg agttaatgta tgggggagtt tcttcttatg 300
aggtttcaat ggtctcaaac gctagacctc agagcccacg agaagggcat atgatggatg 360
ctttgtatgc ccaagtcnaag aaaccgagga attcaaac ctccacctgta gcaagtaaca 420
gatcaactcc tagcaatcat gatcggatc agcgtctgag gcnagaattt cagcaagcna 480
agcaagatga agatgtgaaa g
501

<210> 32
<211> 265
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1945452H1

<400> 32
gttccatggc tctcaacgct agacctcaga gccaccgaga agggcatatg atggatgctt 60
tgtatgccca agtcaagag cgcggaatt ccaaacctc acctgtagc agtaacagat 120
caactcctag caatcatgat cggatcacgc gctgaggca agatttcag caagcaaac 180
aagatgaaga tgtagangat cgtcggcgga cctatagttt tgagcaacc tggccgaaac 240
caaggccgga gacgaagagc gggcg
265

<210> 33
<211> 469
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2549906F6

<220>
<221> unsure
<222> 77, 194, 196, 441, 466
<223> a, t, c, g, or other

<400> 33
gttatttggc tgctcaggaa tagaaggccc tccctttctc agsagtctga agtctgttca 60
gctcgcacac ctgagangga gagggggga catcttgccc gaaggcccc ttggagggg 120
cgtaaactggg gctctggact ttcttatacg agtcatagtt gctggcccc tgggaaggag 180
gotgcttctt catnanctgc tccctccgcc tetgttccct gcaagagagc tctctgattt 240
ccagcatgac cctggcgttg aagccatgtc ctcccaggta gccgttccct gagccttggt 300
agcggagata cctggggttc tctttggcac tctggaagcc tcccaccagg gactagttct 360
gttcccaga gtctctggag accgagctgg catttttccct gctttgccga ggcagagagc 420
tgtactggcg ctgggcctgt ngggagtctc ggcctctctc tgccgntgc
469

```

WO 02/055706

PCT/US02/00647

<210> 34
<211> 558
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 71009002V1

<400> 34
cgggccacgc gccagtcacg ctctctgect cggcaaaagc ggaaaaatgc cagctcggtc 60
tcccaggact cttaggagca gaactactcc cctggggaag gcttccagag tgcctaaagag 120
aaccccaggc actccagcta ccaaggctcc aggaaggcct acctggggag acatggcttc 180
aacgccaggc catgctggaa actcaggagc tccttcgcca ggaacagagc cggagggagc 240
agcagatgaa gaagcagcct ccttcaggag ggcctcagca ctatgactcg tataagaagag 300
tcccaggacc cagttaagcc cctcccaagg ggcctctccg gcaagatgtg cccctctccc 360
cttctcaggt tgcgaggctg aacagacttc agactctcga gaaagggagg cctctctatt 420
cctgagcagc caaataacgg atgcttcctg tcgcgcaata aaagacattt tccctatgag 480
acttgatttc cgggagtttt ttaaaaacct cgtatggtact atggagtata ctggcctgag 540
tatcagtgcc ttaagcgc 558

<210> 35
<211> 632
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 71008521V1

<220>
<221> unsure
<222> 605
<223> a, t, c, g, or other

<400> 35
ttcccacac agagctgtgg tgaactcaacc cacctgtcag gaacaggacc gaacaagcgt 60
atgtataaac gcaagcaagg atcaccactt tattggtgat gttgagtcac aagggcagta 120
caaaaggacc cccatggtga ccaaacatg ttggactatg gcttagcttc tgtaatagga 180
tgaccgactc tcccctcaca aagagctgag aagggaaaga gaaaccacac ctgatgaca 240
agtcagaaa atggctgtgc tgtgaagtc cagaagccc caattttgg tglatgtcgc 300
caacttttcc tccaccagag actaagatgt catcccatag ctaagagaac tttagagggag 360
aaacagggcc gcacgtcacc ctccacttca ggtgaatttg tcaactgcaag tggtagcagg 420
atgttacacc caagccgcgg acaactcatg agtagggccc agattcctag tactgtggag 480
agggcagag catatgaaca cctcaaacag atgattgtca cagggtggga aatccttcca 540
tgaaacagac acttgagaca tagtggcaaa gacatttagg ccttccattt ctttgctacc 600
accgntana ggcactgata ccaacaacag aa 632

<210> 36
<211> 646
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 71010168V1

<400> 36
cttctcagga gaatttcttt cttagctctc cctcctatgat agacacaggg ccaatggcag 60
gcccagcttc ccccgagag agccactcct gtctcccacg ttcccacac agagctgtgg 120
tgactcaacc cccgtgtcag gaacaggacc gaacaaagct atgtataac gcagaccagg 180
atcatcaact tattggtaat gttgagtcac aaggcagtta caaaaggacc cccatggtga 240
ccaaaacatg ttggactatg gcttagcttc tgtaatagga tgaccgactc tcccgtcaca 300

WO 02/055706

PCT/US02/00647

```

aagagctgag aaaggaaga gaaaaccaca gtgatgacaa agtacgacaa atggctgtgc 360
tggTgaagta ccagaaagcc ccaatttttg gtgatgtcg cccacttttc ctccacaga 420
gactaagatg tcatcccata gctaagagaa cttagaggya gaacagggg cgcacgtcat 480
ctccacttc aggtgaattt gtoactgcaa gtgggtccag gatgttcaa ccaagccgcg 540
gacaactcat gagtgggccc gagattctcg gtactgtgga gagggcaga gcatatgac 600
aactcaaca gatgatgtcc cagggtggga aatccttcca tgaaac 646

```

```

<210> 37
<211> 498
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Incyte ID No: 70090181V1

```

```

<400> 37
agctttcacc tccagctttt tccacatcgg gattcacagc caatttagca ccccajccat 60
ggtttacaat acaggatgtt cagaacaatg aaggaagatg ggagccaccg agatccccga 120
gaccacacaa cacactctaa taactgagcg agagtgagct tctctgcacc accgctgtgg 180
gaccggaaaa tgttctctgt gtgatgtgta aeggettttc ctbtgtcttc tatgtttcat 240
ggcagaacta cttaggagcg tggcatcgcc agtcagacat gcagtgctgt gatgcattag 300
acaagacatca attgcctctg acttgaaaaa ccgtagagag cagatcaaaa tgactgtctg 360
gtttactctct gtccactctt gtgcccttcc tgacaagctg tcaagaacaaa aactaattaa 420
aataattact aggatgtact gggccatgag ctcttcttag gagaattct ttcttagctc 480
tcccctcatg atagaaac 498

```

```

<210> 38
<211> 572
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Incyte ID No: 6833928H1

```

```

<400> 38
cttgtcagga agggcacaga atgggacaga gataaacag acagtcattt gatctgtctc 60
ctaaggtttt tcaagtccga gcccaattgat gcttgtctaa tgcattccaa cactgcattg 120
ctgactggcg atgccacgct cctaagtagt tctgccatga aanaataaag acaaaggaaa 180
agccgtbaca catcaacacg agaacatttt cgggtccacc agcgggtggt gcaggaaact 240
cactctcccg taagtattag agtctgtctg tgggtctcgg gcatctcgtt ggtctccacc 300
ttccttcatt gttctgaaca tctgtattg taaccatggt ctgggtgct aaagtgcctc 360
tgaatccaga tgtggaaaaa gctggaggtg aaagctcagc ataccatgta tttactttaa 420
aaacagaaaa aaagacatgt atggatatgt ctattttttt tttatggca catgttattt 480
ttgtgtgac ttgttttag aatgatgtg tccacacag taccctgtc tcttctgeat 540
tctgtgtca tggctctggt tcttaattcac gt 572

```

```

<210> 39
<211> 550
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Incyte ID No: 70089663V1

```

```

<400> 39
gctcataagt agttctgcca tgaaacataa aagacaaagg aaaagccgtt acacatcaca 60
cagaaacat tttcgggtcc cacagcgggt gtggcaggaa gctcactctc gcgtcaat 120
tagatgtgt gtgtgggtct cggggatctc ggtggctccc atcttctctc attgtctgaa 180
acatctctgt tgttaacca tggctggggt gctaaagtgc ctgtgaatcc cgtatgtaa 240
aaaactggag gtgaaagctc agcataccat gtatttactt taaaaacaga aaaaaagaca 300

```

WO 02/055706

PCT/US02/00647

```

tgtatggata tgtctatgtt ttttttattg gaacattgta tttttgtgtt gacttgtttt 360
tagaagatgat gtgtccacac acgtaccctg gtctctctctg catttctgtg tcaatggttt 420
gtttcttaat cacgtgcccg ggtgtctaag tgggtttacc agtgcacggc cagtgaactt 480
ggatgacagt ggctcttctg cacagcctcc cctgagctgt gagacacage tttctctgta 540
catatgcaac
550

```

```

<210> 40
<211> 514
<212> DNA
<213> Rattus norvegicus

```

```

<220>
<221> misc_feature
<223> Incyte ID No: 702231139H1

```

```

<400> 40
aagcaatttt caaatgccag tcaattggat ttctttaaaa cagcaaaaac aaaaagcatg 60
gatttaggta tagctgacga gaactaaact caatacagtg gatgaccaga ggcagggctc 120
ccccataga gatgtgggac cctccttggg totgagaaaa tccagctctt tagaaaagtct 180
gcagacggct gtgtctgagc tgaccctgaa tggpacatt ccttccacc gccacggccc 240
agcaatcacc ccaggaaggg gctgcaacga gagcttcaga gccgccattg acagctccta 300
cgataagccc atggtggatg acgacacaga aggcattggag acctggaaag aagcacacaga 360
agaagttca aggtcagggg gggagctcct gtccacgtcc agtgcacagc ctctctattc 420
tctggagaga caaatgcatg gagaccaga gaagagggac aaggaagaga agaaaaggaa 480
caagccggga aaggataaga agaagaccg agag
514

```

```

<210> 41
<211> 544
<212> DNA
<213> Rattus norvegicus

```

```

<220>
<221> misc_feature
<223> Incyte ID No: 700273304F6

```

```

<400> 41
cctgatggg aacatctctt tccaccgcc acggccaaga atcattccag gaaggggctg 60
caacgagagc tbcagagccg ccattgacaa gtctctcagc aggcacatgg cggatgacga 120
cgaagagacc atggagacct tggagagaga cacagaagaa agttcaagtt cagggagaga 180
gtccgtgtcc acgtccagct atcagccttc ctattctctg gagagacaaa tgaatggaga 240
cccagagag agggacaagc cagagaagaa aaaggacaaa gccggaaggy ataaagagaa 300
agaccgagag aaggagaagc ataaactgaa agccaagag gggatgctga aagccttggg 360
ggacatgttc agcctggcca aactgagacc ggagaagaga tgaacagcat gccagactca 420
aactgtcttg gagacacaaa gttgcacaat tgttttttaa aagcacgggt cctgggctgt 480
ggctcagctt agagtgcctg cctgggtgac acaaaagcct gggctcaatc ccagcacc 540
tata
544

```

```

<210> 42
<211> 272
<212> DNA
<213> Rattus norvegicus

```

```

<220>
<221> misc_feature
<223> Incyte ID No: 700330856H1

```

```

<400> 42
tagatccago gggcaagctc caggaggaag ttgtttccct gttgagaag accaagatgg 60
aggggacagt gagccttctg gtctttctgc aagaagagggc ttccagcca agggcaatga 120
atcoggaacc cagccagatg cagagtccaa aagaacgaa agccgaagac gaggacattg 180
ttctcaacc tgaccgtacc agggagttc tgactttbga agttccactg aatgactcag 240
ggctctcagg gctttgtgtc agcgtcaagg gg
272

```

WO 02/055706

PCT/US02/00647

```

<210> 43
<211> 300
<212> DNA
<213> Rattus norvegicus

<220>
<221> misc_feature
<223> Incyte ID No: 700273304H1

<400> 43
actggaatggg aacattcctt tccaccgcc accgccacga atcatccag gaaggggobg 60
caaccgagagc ttcagagccg ccattgacaa gtccacgat aagcccatgg tggatgacga 120
cgaagaaggc atgagagcct tggagaaga cacagaaga agttcaaggt caggagggga 180
gtccgtgtcc acgtccagtg atcagcttc ctatctctg gagagcaaa tgaatggaga 240
cccaaggaag agggacaagg cagagaagaa aaaggacuaa gccggaagg ataagaagaa 300

<210> 44
<211> 300
<212> DNA
<213> Rattus norvegicus

<220>
<221> misc_feature
<223> Incyte ID No: 701517518H1

<400> 44
caggatctta cactccctct acagtgggat cgaggggctg gatgagtctc ccaccaggaa 60
tgcgcacttc agcaggataa tgggtaaatg ccagctctcc ccaaccgtga acatgccaca 120
tgaatgaaact gtcagtattg aagatgacag gctgctctg ctccctcttc acctctctga 180
ccagtcctcc tccagctccc atgatgacgt gggattcaba atgacagaag caggoacgtg 240
ggccaaagct accatcagtg actcagccga ctgctcattg actccagatg ttgatccggt 300

<210> 45
<211> 544
<212> DNA
<213> Rattus norvegicus

<220>
<221> misc_feature
<223> Incyte ID No: 701834089T1

<220>
<221> unsure
<222> 11-12, 17, 130-131
<223> a, t, c, g, or other

<400> 45
aaacctgagt nnccttnaca acccaaaagta aatttattgt ttggatttta aaaaaacttt 60
cttgagaca cgtttcgtgt atcccaggct ggcctcgaac actacgtarg caggatgacc 120
tgaacttca nnnnnnnnn nnnnnnnnn nnnnnnnnn nnnnnnnnn nnnnnnnnn 180
nnnnnnnnn ntataggggt ctggggattg agcccacggc ttgtgtaca ccaggcaggc 240
actctagaact gagccacagc ccagacaccg tgccttttaa aaacaattgt gcaactgtg 300
ctgtccasga cagtttgagt ctggcatget gttcatctct tctccggctt caytttgcc 360
agctcaaca tgtccocaa gctttccagc atcccctctc tggctttcag tttatccttc 420
tctctctctc gttctttctt ctatcccttt ccggctttgt cctttttctt ctctgcttg 480
tccctctctc ctgggtctcc attcatttgt ctctccagag aataggaagg ctgatcactg 544
gaag 544

<210> 46
<211> 196
<212> DNA
<213> Rattus norvegicus

```

WO 02/055706

PCT/US02/00647

```

<220>
<221> misc_feature
<223> Incyte ID No: 701480437H1

<400> 46
ctctctctct ctcctccttg actgactaac ttctttgctt tattgccaga caaagcagga 60
agaatgccag ctctgtata caggattcct gggaacagaa ctacgcccc ggtgaaggct 120
tccagagtgc caagagaaac cccaggattt ccagttacca gggctccagg aacggctatc 180
taggtggcca tggcct 196

<210> 47
<211> 273
<212> DNA
<213> Rattus norvegicus

<220>
<221> misc_feature
<223> Incyte ID No: 701190235H1

<400> 47
gcaatgttaa cgagttgctg tctcctggga gccccyctgc acccgatctg ccatacaaa 60
cagagttgga tgacagacaa cgcaggtatc cacacrcctc ctacagtggg atcgatggcc 120
tggatgagtc tcccaccagg aatgccgcac tcagcaggat aatgggtaaa tgcagctctc 180
ccccaacctg gaacatgcta catgatgaca ctgtcatgat tgaagatgac aggtcctctg 240
tgctcactcc tcactctctc gaccagctcc cct 273

<210> 48
<211> 248
<212> DNA
<213> Rattus norvegicus

<220>
<221> misc_feature
<223> Incyte ID No: 700939688H1

<400> 48
cagagaagag ggacaaggca gagaagaaaa aggacaagc cggaaaggat aagaagaaag 60
accgagagaa gggagaggat aaactgaaag ccaagaaggc gatgctgaaa gcttggggg 120
acatgttcag cctggccaaa ctgaagccgg agaagagatg aacagcatgc cagactcaaa 180
ctgtcttggg cagcacaagt tgcacaattg ttttttaaaa gcacgggtgc tggctgtg 240
ctcagctc 248

<210> 49
<211> 351
<212> DNA
<213> Rattus norvegicus

<220>
<221> misc_feature
<223> Incyte ID No: 700939688F6

<220>
<221> unsure
<222> 337
<223> a, t, c, g, or other

<400> 49
cagagaagag ggacaaggca gagaagaaaa aggacaagc cggaaaggat aagaagaaag 60
accgagagaa gggagaggat aaactgaaag ccaagaaggc gatgctgaaa gcttggggg 120
acatgttcag cctggccaaa ctgaagccgg agaagagatg aacagcatgc cagactcaaa 180
ctgtcttggg cagcacaagt tgcacaattg ttttttaaaa gcacgggtgc tggctgtg 240
ctcagctcag aagatgcctg cctggctgta cacaagagcc agtggagctc aagctcccag 300
acagccctat agaacagagc tctggtagac acatgcnctg tcattccagc a 351

```

WO 02/055706

PCT/US02/00647

<210> 50
<211> 571
<212> DNA
<213> Rattus norvegicus

<220>
<221> misc_feature
<223> Incyte ID No: 702582937T1

<400> 50
cactttatga acccagcctt ggtttacaat acaggatgtt cagaccnaca gatgaacggc 60
gggaacaagg agggcctcgt gccacagcca tgcacggaga atgggactcc cgggtctcag 120
agggacatcg acaggctcct gagggggatg gctctctctc tgtttgtgaa taacacagcg 180
agtnactcag taatgttggc ctgctcaggt cgggacatgg tatgaggata tagggagcca 240
aatcctgact gcaacctcaa aagctgtgtt gaggttgatt ctacagaatcc caagtcactg 300
acctttttcc ttgatccccc totgtgctcc ccttgacaac ctacggtgac acgaagtaaa 360
gtaaggactg gatagaccgg cctaagctcc tccagagagt cttccctcag actcctatct 420
ctctcctcgg ggtgcgtaca catgggccac tcccatgccc cttgttcccg agtgcctgca 480
gtgactgaaa ctgaacgcat gactctatag aaccactgac tagaaatcgg ctgtagatat 540
gggtgggtgg agttataaag ggcagttgga a 571

<210> 51
<211> 694
<212> DNA
<213> Rattus norvegicus

<220>
<221> misc_feature
<223> Incyte ID No: 700299037P6

<400> 51
ctctagggtg gtagtgaaga agctaagcca taggcccagt ccggttggtc tggggggtga 60
gggttaacttt ccaactgcct tataactcca ccaccatatt cacagcgatt ctagtcaagt 120
tttataagta catgcttca gtttcaagtc ctcatgacac tggggaacaa gggcctatgg 180
agtgcccact ggttacgca cccgaggaag gatataggag ctgagggaa actctctgga 240
ggagcttagg ccggctctac cagctcttac tttactctgt gtcaccgtag gttgtcaagg 300
gagggcaga ctgggacag gaaaaggtcc agtcaactgg gattctgaga atcaacctca 360
acnagccttt tgggttgc gtcagatgg gttctctata tctcctacc atgtcccgac 420
ctgacggggg ccacattact gagtgcctct gctgtttatt caccaacaga aggagagcca 480
tcccatctga ggacctgtg atgtccctct gagcaccggg agtccattc tccgtgatg 540
cctgtggccc gaggctccg byttcccgcc gttcctcgt tggctotgac atcctgtatg 600
taaaccaagg ctgggttgc aaagtccctg agaatctoga tataaaaaac aaaaaacaa 660
aaaatccttg gggcaaaagc tcagagtacc atgt 694

<210> 52
<211> 110
<212> DNA
<213> Rattus norvegicus

<220>
<221> misc_feature
<223> Incyte ID No: 701246488H1

<400> 52
gattgaagat gacaggctgc ctgtctccc tccgacctc totgaccagg cgtctccag 60
ctccatgat gacgtgggat tcaaatgac agaagcaggc acgtggcca 110

<210> 53
<211> 578
<212> DNA
<213> Canis familiaris

<220>

WQ 02/055706

PCT/US02/00647

```

<221> misc_feature
<223> Incyte ID No: 702759912H1

<400> 53
atgaaagtga cegtgtgctt cgggcggacc cgggtggctg tgcctgctgg gacgggac 60
atgaaagtgt tcagcctcat ccagcaagcg gtgaccgctt accggaaggc catcgccaag 120
gatccaaact actggataca ggtgcaccga ctgaaacalg gagatggagg aatactagac 180
cttgatgaca cccctctgtga tgtagcagat gataaagaca gactggttag agtgtttgat 240
gagcaagatc cacatcatgg aggtgatggc accagtgcca gctccacagg taccagagt 300
ccagagata tggcagtga gcttggcacc acaatgtttt cagccttcca gccttatcaa 360
gbcacaagtg aattgaggtt cacacttca gttctctgtg caaatatgoc tcttcabgtc 420
cgacgaagca gtgaccgggc tttaatggc cttttaaactt ccatcagtga cactaatttt 480
ccttctgaag agccttcacg gaagaacccc acagtttgtt caacacacgc tggctttctg 540
aagcaaaaac ctgctggcag ccctaatact gtgcaaaa 578

<210> 54
<211> 293
<212> DNA
<213> Mus musculus

<220>
<221> misc_feature
<223> Incyte ID No: 700112340H1

<400> 54
gggaatttga ctgagatgtc ccaaaggtgc ctattggaag agcattatga tccaaactac 60
tggatcacgg tgcactgctt ggagcatgga gatggaggga ttctagacct gpatgacatc 120
ctctgtgacg ttctgtatga caasgacaga ctggtagcag tatttgatga acaggtctcc 180
caccatggag gagatggtao cagggccagc tccacgggaa cccagagtcc agagatattc 240
gpcagtggc tgggcaccaa caatgtttct gcttttcagc cttatcaagc ctc 293

<210> 55
<211> 233
<212> DNA
<213> Mus musculus

<220>
<221> misc_feature
<223> Incyte ID No: 700627610H1

<400> 55
cgcggccggc atcgcaagt ggtcggcact cgggtctcct ggaggttcaa gtacaacggc 60
agccccagga ggagcgagag agcttccagc aggccacgoc ccagtacagc tcactgcca 120
gacaaagcag gaagaatgcc agctccatat cacaggattc ctgggacag aagagtgaag 180
aaactcttgg gcaagtatgg cctagcagt gtgagagaca ccacaggaag tgg 233

<210> 56
<211> 222
<212> DNA
<213> Mus musculus

<220>
<221> misc_feature
<223> Incyte ID No: 700109331H1

<400> 56
gggcacttca atgcaagga aaactaatct ttttgccaaa ttgacacttt gtaaarrtat 60
ttctctattt gctaaaaaa aaatagacat gtgtttggga ccttgagctc catcccgag 120
catccgaacc ttactcaag aatcatggag attgtactca ctacctsaat ccatgctttt 180
tgattttcgt gttttaaaga aatccaattg actggcatct ga 222

<210> 57
<211> 369

```

WO 02/055706

PCT/US02/00647

```

<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: g6661750

<400> 57
aaggggagct gccgcagacc tccgggectc aggggtgttc ggggagcggc gccccgggctc 60
tctggccaca ccgcgcccgq gqgtccctcc agestlssgg ctgcgcccgc ggggtccagac 120
acctgtttcg ccsgcccccq cgtggtctgc gggggccagg atgaaagtga ccgtgtcctt 180
cggcaggacg ggcacatctgg tgcctctcaa ggaaggccag ctgcgctcg gggagctcac 240
ccagcaggcg ctgcagcggc acctgagac ccgggagaag ggtcctgatt actgggtgaa 300
gattcacac ttagaatata cagatggagg aatcctggat ccagatgatg tcttggcaga 360
tgtgttga                                     369

<210> 58
<211> 511
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: GNN.g10801482_004.edit

<400> 58
gcccccggccg cacatggttc gaggccgagg ctgcaatgag agcttttagag cagccattga 60
caatcctac gatgacctg aagaataga agctgaggt ctgtctgata agagctotca 120
ctctggccaa gsgctctga atgtggatc tgcctctcag ggaattcgg agctgagga 180
catggaaaat aagccaggga aagtcaaaa aacgaaagag aaggagaaga aaaaagaaa 240
gggcaaatg aagtccaag aaaaaaacg caaagaggag aatgaagatc cagaagoga 300
aataaagaag aaggcttcg ggcacctgc gaggtatgg cctgtttga agcacaagt 360
ggtctcatt ttgtctctc tgaaaaaag gcacgctttt cctgtcttc agccaaatgc 420
atcggctct caattctgq ctgctctct tctcgacag gcagaggagc tttttgggga 480
aagttacagt gatcacagga cactgtctta a                                     511

<210> 59
<211> 389
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: g6993427

<400> 59
ccagtttcat cctttttct atcagtetta tcaactttct atgggtttcc attcatttgt 60
ctctccagcg agtgggaag ctgatactg gctgtggata cgaactctct cctgatctt 120
gaactttct ctgtctctc ttccaagtc tccatgctt catcatcat atctaccag 180
ggtttatcat aagatttgc gatggcagct ctgaagctc cattgcctc cctgctctg 240
attatccgq gctgtggag atggaagga atatccccat tcaaaagcac ctgccaact 300
gqgctctgca gactctcaa cgagcttgc ttcttcagac ccagggaagg acccaactc 360
ctgctggag aacctgttt ctggtcatcc actgtattga gtttagtct gtcagctata 420
ctbaatcaa tctttctga tttctgtct taaagaaat ccaattgact ggcactgaa 480
aattgcttg tgggttttc tgaatctc tgaagcaaa atcctctctg ttgaaaagca 540
egaactggat caacatctg actcaagag cagtcgggtg aatcactgat t                                     591

<210> 60
<211> 389
<212> DNA
<213> Homo sapiens

<220>

```

WO 02/055706

PCT/US02/00647

<221> misc_feature
<223> Incyte ID No: g5529915

<400> 60
tttttttca gttttatcct tttttctatc agttttatca cttttctett ggtttccatt 60
catttgtctc tccagagagt gggagggctg atcaactggct gtggatacag actctctccc 120
tgatcttgaa cttttctctg tgtttctctc caaggtctcc atgctctcat cactcctc 180
taccgggat ttatcctaag atttctgat ggcagctctg aagctctcat tgcctccct 240
gctctgat atccgggccc gtggagatg gaaaggnata tcccattca agtccactc 300
ggcaactgg gtctggagac tctccaacga gcttgactc ttcagaccca gggagagacc 360
cactcctctg ctgggagaac ctgctttct 389

<210> 61
<211> 367
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: g1733437

<400> 61
ggaaccacag gaaagggtga taagactgat agaaaggagg atnaaactgg aaaagaaaag 60
aagaangata gagataagga gaaggataaa atgaaagcca agaagggat gctgaagggc 120
ttgggagaca tgttcaggtt tggcaaacat cgaagagatg acaagattga gaaacggggc 180
aaaataaaaa tacaggatc ctttacatca gaagaggaga gatacagaat gaagcaggag 240
caggagagga tccaagcca aactcgagaa ttagggaac cgaacaagctc gagagcgtga 300
ctatgctgaa attcaagatt ttcacggagc atttggctgt gatgatgagt taatgatgg 360
ggaggtt 367

<210> 62
<211> 1337
<212> PRT
<213> Rattus norvegicus

<220>
<221> misc_feature
<223> Incyte ID No: g3868778

<400> 62
Met Lys Val Thr Val Cys Phe Gly Arg Thr Arg Val Val Val Pro
1 5 10 15
Cys Gly Asp Gly Arg Met Iys Val Phe Ser Leu Ile Gln Gln Ala
20 25 30
Val Thr Arg Tyr Arg Lys Ala Val Ala Lys Asp Pro Asn Tyr Trp
35 40 45
Ile Gln Val His Arg Leu Glu His Gly Asp Gly Ile Leu Asp
50 55 60
Leu Asp Asp Ile Leu Cys Asp Val Ala Asp Asp Lys Asp Arg Leu
65 70 75
Val Ala Val Phe Asp Glu Gln Asp Pro His His Gly Gly Asp Gly
80 85 90
Thr Ser Ala Ser Ser Thr Gly Thr Gln Ser Pro Glu Ile Phe Gly
95 100 105
Ser Glu Leu Gly Thr Asn Asn Val Ser Ala Phe Arg Pro Tyr Gln
110 115 120
Thr Thr Ser Glu Ile Glu Val Thr Pro Ser Val Leu Arg Ala Asn
125 130 135
Met Pro Leu His Val Arg Arg Ser Ser Asp Pro Ala Leu Thr Gly
140 145 150
Leu Ser Thr Ser Val Ser Asp Asn Asn Phe Ser Ser Glu Glu Pro
155 160 165
Ser Arg Lys Asn Pro Thr Arg Trp Ser Thr Thr Ala Gly Phe Leu

WO 02/055706

PCT/US02/00647

| | | | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Lys | Gln | Asn | Thr | Thr | Gly | Ser | Pro | Lys | Thr | Cys | Asp | Arg | Lys | Lys | 170 | 175 | 180 |
| | | | | | | | | | | | | | | | 185 | 190 | 195 |
| Asp | Glu | Asn | Tyr | Arg | Ser | Leu | Pro | Arg | Asp | Pro | Ser | Ser | Trp | Ser | 200 | 205 | 210 |
| Asn | Gln | Phe | Gln | Arg | Asp | Asn | Ala | Arg | Ser | Ser | Leu | Ser | Ala | Ser | 215 | 220 | 225 |
| His | Pro | Met | Val | Asp | Arg | Trp | Leu | Glu | Lys | Gln | Glu | Gln | Asp | Glu | 230 | 235 | 240 |
| Glu | Gly | Thr | Glu | Glu | Asp | Ser | Ser | Arg | Val | Glu | Pro | Val | Gly | His | 245 | 250 | 255 |
| Ala | Asp | Thr | Gly | Leu | Glu | Asn | Met | Pro | Asn | Phe | Ser | Leu | Asp | Asp | 260 | 265 | 270 |
| Met | Val | Lys | Leu | Val | Gln | Val | Pro | Asn | Asp | Gly | Gly | Pro | Leu | Gly | 275 | 280 | 285 |
| Ile | His | Val | Val | Pro | Phe | Ser | Ala | Arg | Gly | Gly | Arg | Thr | Leu | Gly | 290 | 295 | 300 |
| Leu | Leu | Val | Lys | Arg | Leu | Glu | Lys | Gly | Gly | Lys | Ala | Glu | Gln | Glu | 305 | 310 | 315 |
| Asn | Leu | Phe | His | Glu | Asn | Asp | Cys | Ile | Val | Arg | Ile | Asn | Asp | Gly | 320 | 325 | 330 |
| Asp | Leu | Arg | Asn | Arg | Arg | Phe | Glu | Gln | Ala | Gln | His | Met | Phe | Arg | 335 | 340 | 345 |
| Gln | Ala | Met | Arg | Ala | Arg | Val | Ile | Trp | Phe | His | Val | Val | Pro | Ala | 350 | 355 | 360 |
| Ala | Asn | Lys | Glu | Gln | Tyr | Glu | Gln | Leu | Ser | Gln | Arg | Glu | Met | Asn | 365 | 370 | 375 |
| Asn | Tyr | Ser | Pro | Gly | Arg | Phe | Ser | Pro | Asp | Ser | His | Cys | Val | Ala | 380 | 385 | 390 |
| Asn | Arg | Ser | Val | Ala | Asn | Asn | Ala | Pro | Gln | Ala | Leu | Pro | Arg | Ala | 395 | 400 | 405 |
| Pro | Arg | Leu | Ser | Gln | Pro | Pro | Glu | Gln | Leu | Asp | Ala | His | Pro | Arg | 410 | 415 | 420 |
| Leu | Pro | His | Ser | Ala | His | Ala | Ser | Thr | Lys | Pro | Pro | Thr | Ala | Pro | 425 | 430 | 435 |
| Ala | Leu | Ala | Pro | Pro | Asn | Val | Leu | Ser | Thr | Ser | Val | Gly | Ser | Val | 440 | 445 | 450 |
| Tyr | Asn | Thr | Lys | Arg | Val | Gly | Lys | Arg | Leu | Asn | Ile | Gln | Leu | Lys | 445 | 450 | 455 |
| Lys | Gly | Thr | Glu | Gly | Leu | Gly | Phe | Ser | Ile | Thr | Ser | Arg | Asp | Val | 470 | 475 | 480 |
| Thr | Ile | Gly | Gly | Ser | Ala | Pro | Ile | Tyr | Val | Lys | Asn | Ile | Leu | Pro | 485 | 490 | 495 |
| Arg | Gly | Ala | Ala | Ile | Gln | Asp | Gly | Arg | Leu | Lys | Ala | Gly | Asp | Arg | 500 | 505 | 510 |
| Leu | Ile | Glu | Val | Asn | Gly | Val | Asp | Leu | Ala | Gly | Lys | Ser | Gln | Glu | 515 | 520 | 525 |
| Glu | Val | Val | Ser | Leu | Leu | Arg | Ser | Thr | Lys | Met | Glu | Gly | Thr | Val | 530 | 535 | 540 |
| Ser | Leu | Leu | Val | Phe | Arg | Gln | Glu | Glu | Ala | Phe | His | Pro | Arg | Glu | 545 | 550 | 555 |
| Met | Asn | Ala | Glu | Pro | Ser | Gln | Met | Gln | Ser | Pro | Lys | Glu | Thr | Lys | 560 | 565 | 570 |
| Ala | Glu | Asp | Glu | Asp | Ile | Val | Leu | Thr | Pro | Asp | Gly | Thr | Arg | Glu | 575 | 580 | 585 |
| Phe | Leu | Thr | Phe | Glu | Val | Pro | Leu | Asn | Asp | Ser | Gly | Ser | Ala | Gly | 590 | 595 | 600 |
| Leu | Gly | Val | Ser | Val | Lys | Gly | Asn | Arg | Ser | Lys | Glu | Asn | His | Ala | 605 | 610 | 615 |
| Asp | Leu | Gly | Ile | Phe | Val | Lys | Ser | Ile | Ile | Asn | Gly | Gly | Ala | Ala | 620 | 625 | 630 |
| Ser | Lys | Asp | Gly | Arg | Leu | Arg | Val | Asn | Asp | Gln | Leu | Ile | Ala | Val | 635 | 640 | 645 |

WO 02/055706

PCT/US02/00647

Asn Gly Glu Ser Leu Leu Gly Lys Ala Asn Gln Glu Ala Met Glu
650 655 660
Thr Leu Arg Arg Ser Met Ser Thr Glu Gly Asn Lys Arg Gly Met
665 670 675
Ile Gln Leu Ile Val Ala Arg Arg Ile Ser Arg Cys Asn Glu Leu
680 685 690
Arg Ser Pro Gly Ser Pro Ala Ala Pro Glu Leu Pro Ile Glu Thr
695 700 705
Glu Leu Asp Asp Arg Glu Arg Arg Ile Ser His Ser Leu Tyr Ser
710 715 720
Gly Ile Glu Gly Leu Asp Glu Ser Pro Thr Arg Asn Ala Ala Leu
725 730 735
Ser Arg Ile Met Gly Glu Ser Gly Lys Cys Gln Leu Ser Pro Thr
740 745 750
Val Asn Met Pro His Asp Asp Thr Val Met Ile Glu Asp Asp Arg
755 760 765
Leu Pro Val Leu Pro Pro His Leu Ser Asp Gln Ser Ser Ser Ser
770 775 780
Ser His Asp Asp Val Gly Phe Ile Met Thr Glu Ala Gly Thr Trp
785 790 795
Ala Lys Ala Thr Ile Ser Asp Ser Ala Asp Cys Ser Leu Ser Pro
800 805 810
Asp Val Asp Pro Val Leu Ala Phe Gln Arg Glu Gly Phe Gly Arg
815 820 825
Gln Ser Met Ser Glu Lys Arg Thr Lys Gln Phe Ser Asn Ala Ser
830 835 840
Gln Leu Asp Phe Val Lys Thr Arg Lys Ser Lys Ser Met Asp Leu
845 850 855
Gly Ile Ala Asp Glu Thr Lys Leu Asn Thr Val Asp Asp Gln Arg
860 865 870
Ala Gly Ser Pro Asn Arg Asp Val Gly Pro Ser Leu Gly Leu Lys
875 880 885
Lys Ser Ser Ser Leu Glu Ser Leu Gln Thr Ala Val Ala Glu Val
890 895 900
Thr Leu Asn Gly Asn Ile Pro Phe His Arg Pro Arg Pro Arg Ile
905 910 915
Ile Arg Gly Arg Gly Cys Asn Glu Ser Phe Arg Ala Ala Ile Asp
920 925 930
Lys Ser Tyr Asp Lys Pro Met Val Asp Asp Asp Asp Glu Gly Met
935 940 945
Glu Thr Leu Glu Glu Asp Thr Glu Glu Ser Ser Arg Ser Gly Arg
950 955 960
Glu Ser Val Ser Thr Ser Ser Asp Gln Pro Ser Tyr Ser Leu Glu
965 970 975
Arg Gln Met Asn Gly Asp Pro Glu Lys Arg Asp Lys Ala Glu Lys
980 985 990
Lys Lys Asp Lys Ala Gly Lys Asp Lys Lys Lys Asp Arg Glu Lys
995 1000 1005
Glu Lys Asp Lys Leu Lys Ala Lys Lys Gly Met Leu Lys Gly Leu
1010 1015 1020
Gly Asp Met Phe Arg Phe Gly Lys His Arg Lys Asp Asp Lys Met
1025 1030 1035
Glu Lys Met Gly Arg Ile Lys Ile Gln Asp Ser Phe Thr Ser Glu
1040 1045 1050
Glu Asp Arg Val Arg Met Lys Glu Glu Gln Glu Arg Ile Gln Ala
1055 1060 1065
Lys Thr Arg Glu Phe Arg Glu Arg Gln Ala Arg Glu Arg Asp Tyr
1070 1075 1080
Ala Glu Ile Gln Asp Phe His Arg Thr Phe Gly Cys Asp Asp Glu
1085 1090 1095
Leu Leu Tyr Gly Gly Met Ser Ser Tyr Asp Gly Cys Leu Ala Leu
1100 1105 1110
Asn Ala Arg Pro Gln Ser Pro Arg Glu Gly His Leu Met Asp Thr

WO 02/055706

PCT/US02/00647

1115 1120 1125
 Leu Tyr Ala Gln Val Lys Lys Pro Arg Ser Ser Lys Pro Gly Asp
 1130 1135 1140
 Ser Asn Arg Ser Thr Pro Ser Asn His Asp Arg Ile Gln Arg Leu
 1145 1150 1155
 Arg Gln Glu Phe Gln Gln Ala Lys Gln Asp Glu Asp Val Glu Asp
 1160 1165 1170
 Arg Arg Arg Thr Tyr Ser Phe Gln Gln Ser Trp Ser Ser Ser Arg
 1175 1180 1185
 Pro Ala Ser Gln Ser Gly Arg His Ser Val Ser Val Glu Val Gln
 1190 1195 1200
 Val Gln Arg Gln Arg Gln Glu Glu Arg Glu Ser Phe Gln Gln Ala
 1205 1210 1215
 Gln Arg Gln Tyr Ser Ser Leu Pro Arg Gln Ser Arg Lys Asn Ala
 1220 1225 1230
 Ser Ser Val Ser Gln Asp Ser Trp Glu Gln Asn Tyr Ala Pro Gly
 1235 1240 1245
 Glu Gly Phe Gln Ser Ala Lys Glu Asn Pro Arg Tyr Ser Ser Tyr
 1250 1255 1260
 Gln Gly Ser Arg Asn Gly Tyr Leu Gly Gly His Gly Phe Asn Ala
 1265 1270 1275
 Arg Val Met Leu Glu Thr Gln Glu Leu Leu Arg Gln Glu Gln Arg
 1280 1285 1290
 Arg Lys Glu Gln Gln Leu Lys Lys Gln Pro Pro Ala Asp Gly Val
 1295 1300 1305
 Arg Gly Pro Phe Arg Gln Asp Val Pro Pro Ser Pro Ser Gln Val
 1310 1315 1320
 Ala Arg Leu Asn Arg Leu Gln Thr Pro Glu Lys Gly Arg Pro Phe
 1325 1330 1335
 Tyr Ser

<210> 63
 <211> 1266
 <212> PRF
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: g8037915

<400> 63
 Met Lys Val Thr Val Cys Phe Gly Arg Thr Arg Val Val Val Pro
 1 5 10 15
 Cys Gly Asp Gly His Met Lys Val Phe Ser Leu Ile Gln Gln Ala
 20 25 30
 Val Thr Arg Tyr Arg Lys Ala Ile Ala Lys Asp Pro Asn Tyr Trp
 35 40 45
 Ile Gln Val His Arg Leu Glu His Gly Asp Gly Gly Ile Leu Asp
 50 55 60
 Leu Asp Asp Ile Leu Cys Asp Val Ala Asp Asp Lys Asp Arg Leu
 65 70 75
 Val Ala Val Phe Asp Glu Gln Asp Pro His His Gly Gly Asp Gly
 80 85 90
 Thr Ser Ala Ser Ser Thr Gly Thr Gln Ser Pro Glu Ile Phe Gly
 95 100 105
 Ser Glu Leu Gly Thr Asn Asn Val Ser Ala Phe Gln Pro Tyr Gln
 110 115 120
 Ala Thr Ser Glu Ile Glu Val Thr Pro Ser Val Leu Arg Ala Asn
 125 130 135
 Met Pro Leu His Val Arg Arg Ser Ser Asp Pro Ala Leu Ile Gly
 140 145 150
 Leu Ser Thr Ser Val Ser Asp Ser Asn Phe Ser Ser Glu Glu Pro

WO 02/055706

PCT/US02/00647

| | | | |
|---|-----|-----|-----|
| Ser Arg Lys Asn | 155 | 160 | 165 |
| Pro Thr Arg Trp Ser | 170 | 175 | 180 |
| Thr Thr Ala Gly Phe Leu | 185 | 190 | 195 |
| Lys Gln Asn Thr Ala Gly Ser Pro Lys | 200 | 205 | 210 |
| Thr Cys Asp Arg Lys Asp | 215 | 220 | 225 |
| Glu Asp Gly Thr Glu Glu Asp Asn Ser | 230 | 235 | 240 |
| Arg Val Glu Pro Val Gly | 245 | 250 | 255 |
| His Ala Asp Thr Gly Leu Glu His Ile | 260 | 265 | 270 |
| Pro Asn Phe Ser Leu Asp | 275 | 280 | 285 |
| Asp Met Val Lys Leu Val Glu Val Pro | 290 | 295 | 300 |
| Asn Asp Gly Gly Pro Leu | 305 | 310 | 315 |
| Gly Ile His Val Val Pro Phe Ser Ala | 320 | 325 | 330 |
| Arg Gly Gly Arg Thr Leu | 335 | 340 | 345 |
| Gly Leu Leu Val Lys Arg Leu Glu Lys | 350 | 355 | 360 |
| Gly Gly Lys Ala Glu His | 365 | 370 | 375 |
| Glu Asn Leu Phe Arg Glu Asn Asp Cys | 380 | 385 | 390 |
| Ile Val Arg Ile Asn Asp | 395 | 400 | 405 |
| Asn Asp Ser Thr Thr Thr Val Ser Ser | 410 | 415 | 420 |
| Gly Asp Leu Arg Asn Arg Arg Phe Glu | 425 | 430 | 435 |
| Gln Ala Gln His Met Phe | 440 | 445 | 450 |
| Arg Gln Ala Met Arg Thr Pro Ile Ile | 455 | 460 | 465 |
| Trp Phe His Val Val Pro | 470 | 475 | 480 |
| Ala Ala Asn Lys Glu Gln Tyr Glu Gln | 485 | 490 | 495 |
| Leu Ser Gln Ser Glu Lys | 500 | 505 | 510 |
| Asn Asn Tyr Tyr Ser Ser Arg Phe Ser | 515 | 520 | 525 |
| Pro Asp Ser Gln Tyr Ile | 530 | 535 | 540 |
| Asp Asn Arg Ser Val Asn Ser Ala Gly | 545 | 550 | 555 |
| Leu His Thr Val Gln Arg | 560 | 565 | 570 |
| Ala Pro Arg Leu Asn His Pro Pro Glu | 575 | 580 | 585 |
| Gln Ile Asp Ser His | 590 | 595 | 600 |
| Arg Leu Pro His Ser Ala His Pro Ser | 605 | 610 | 615 |
| Gly Lys Pro Pro Ser Ala | 620 | 625 | 630 |
| Pro Ala Ser Ala Pro Gln Asn Val Phe | | | |
| Ser Thr Thr Val Ser Ser | | | |
| Gly Tyr Asn Thr Lys Lys Ile Gly Lys Arg | | | |
| Leu Asn Ile Gln Leu | | | |
| Lys Lys Gly Thr Glu Gly Leu Gly Phe | | | |
| Ser Ile Thr Ser Arg Asp | | | |
| Val Thr Ile Gly Gly Ser Ala Pro Ile | | | |
| Tyr Val Lys Asn Ile Leu | | | |
| Pro Arg Gly Ala Ala Ile Gln Asp Gly | | | |
| Arg Leu Lys Ala Gly Asp | | | |
| Arg Leu Ile Glu Val Asn Gly Val Asp | | | |
| Leu Val Gly Lys Ser Gln | | | |
| Glu Glu Val Val Ser Leu Leu Arg Ser | | | |
| Thr Lys Met Glu Gly Thr | | | |
| Val Ser Leu Leu Val Phe Arg Gln Glu | | | |
| Asp Ala Phe His Pro Arg | | | |
| Glu Leu Lys Ala Glu Asp Glu Asp Ile | | | |
| Val Leu Thr Pro Asp Gly | | | |
| Thr Arg Glu Phe Leu Thr Phe Glu Val | | | |
| Pro Leu Asn Asp Ser Gly | | | |
| Ser Ala Gly Leu Gly Val Ser Val Lys | | | |
| Gly Asn Arg Ser Lys Glu | | | |
| Asn His Ala Asp Leu Gly Ile Phe Val Lys | | | |
| Ser Ile Ile Asn Gly | | | |
| Gly Ala Ala Ser Lys Asp Gly Arg Leu | | | |
| Arg Val Asn Asp Gln Leu | | | |
| Ile Ala Val Asn Gly Glu Ser Leu Leu | | | |
| Gly Lys Thr Asn Gln Asp | | | |
| Ala Met Glu Thr Leu Arg Arg Ser Met | | | |
| Ser Thr Glu Gly Asn Lys | | | |
| Arg Gly Met Ile Gln Leu Ile Val Ala | | | |
| Arg Arg Ile Ser Lys Cys | | | |

WO 02/055706

PCT/US02/00647

Asn Glu Leu Lys Ser Pro Gly Ser Pro Pro Gly Pro Glu Leu Pro 635 640 645
Ile Glu Thr Ala Leu Asp Asp Arg Glu Arg Arg Ile Ser His Ser 650 655 660
Leu Tyr Ser Gly Ile Glu Gly Leu Asp Glu Ser Pro Ser Arg Asn 665 670 675
Ala Ala Leu Ser Arg Ile Met Gly Lys Tyr Gln Leu Ser Pro Thr 680 685 690
Val Asn Met Pro Gln Asp Asp Thr Val Ile Ile Glu Asp Asp Arg 695 700 705
Leu Pro Val Leu Pro Pro His Leu Ser Asp Gln Ser Ser Ser Ser 710 715 720
Ser His Asp Asp Val Gly Phe Val Thr Ala Asp Ala Gly Thr Trp 725 730 735
Ala Lys Ala Ala Ile Ser Asp Ser Ala Asp Cys Ser Leu Ser Pro 740 745 750
Asp Val Asp Pro Val Leu Ala Phe Gln Arg Glu Gly Phe Gly Arg 755 760 765
Gln Ile Ala Asp Glu Thr Lys Leu Asn Thr Val Asp Asp Gln Lys 770 775 780
Ala Gly Ser Pro Ser Arg Asp Val Gly Pro Ser Leu Gly Leu Lys 785 790 795
Lys Ser Ser Ser Leu Glu Ser Leu Gln Thr Ala Val Ala Glu Val 800 805 810
Thr Leu Asn Gly Asp Ile Pro Phe His Arg Pro Arg Pro Arg Ile 815 820 825
Ile Arg Gly Arg Gly Cys Asn Glu Ser Phe Arg Ala Ala Ile Asp 830 835 840
Lys Ser Tyr Asp Lys Pro Ala Val Asp Asp Asp Asp Glu Gly Met 845 850 855
Glu Thr Leu Glu Glu Asp Thr Glu Glu Ser Ser Arg Ser Gly Arg 860 865 870
Glu Ser Val Ser Thr Ala Ser Asp Gln Pro Ser His Ser Leu Glu 875 880 885
Arg Gln Met Asn Gly Asn Gln Glu Lys Gly Asp Lys Thr Asp Arg 890 895 900
Lys Lys Asp Lys Thr Gly Lys Glu Lys Lys Lys Asp Arg Asp Lys 905 910 915
Glu Lys Asp Lys Met Lys Ala Lys Lys Gly Met Leu Lys Gly Leu 920 925 930
Gly Asp Met Phe Arg Phe Gly Lys His Arg Lys Asp Asp Lys Ile 935 940 945
Glu Lys Thr Gly Lys Ile Lys Ile Gln Glu Ser Phe Thr Ser Glu 950 955 960
Glu Glu Arg Ile Arg Met Lys Gln Glu Gln Glu Arg Ile Gln Ala 965 970 975
Lys Thr Arg Glu Phe Arg Glu Arg Gln Ala Arg Glu Arg Asp Tyr 980 985 990
Ala Glu Ile Gln Asp Phe His Arg Thr Phe Gly Cys Asp Asp Glu 995 1000 1005
Leu Met Tyr Gly Gly Val Ser Ser Tyr Glu Gly Ser Met Ala Leu 1010 1015 1020
Asn Ala Arg Pro Gln Ser Pro Arg Glu Gly His Met Met Asp Ala 1025 1030 1035
Leu Tyr Ala Gln Val Lys Lys Pro Arg Asn Ser Lys Pro Ser Pro 1040 1045 1050
Val Asp Ser Asn Arg Ser Thr Pro Ser Asn His Asp Arg Ile Gln 1055 1060 1065
Arg Leu Arg Gln Glu Phe Gln Gln Ala Lys Gln Asp Glu Asp Val 1070 1075 1080
Glu Asp Arg Arg Arg Thr Tyr Ser Phe Glu Gln Pro Trp Pro Asn 1085 1090 1095
Ala Arg Pro Ala Thr Gln Ser Gly Arg His Ser Val Ser Val Glu

WO 02/055706

PCT/US02/00647

| | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|------|
| Val | Gln | Met | Gln | Arg | Gln | Arg | Gln | Glu | Glu | Arg | Glu | Ser | Ser | Gln | 1100 |
| | | | | | | | | | | | | | | | 1105 |
| | | | | | | | | | | | | | | | 1110 |
| Gln | Ala | Gln | Arg | Gln | Tyr | Ser | Ser | Leu | Pro | Arg | Gln | Ser | Arg | Lys | 1115 |
| | | | | | | | | | | | | | | | 1120 |
| | | | | | | | | | | | | | | | 1125 |
| Asn | Ala | Ser | Ser | Val | Ser | Gln | Asp | Ser | Trp | Glu | Gln | Asn | Tyr | Ser | 1130 |
| | | | | | | | | | | | | | | | 1135 |
| | | | | | | | | | | | | | | | 1140 |
| Pro | Gly | Glu | Gly | Phe | Gln | Ser | Ala | Lys | Glu | Asn | Pro | Arg | Tyr | Ser | 1145 |
| | | | | | | | | | | | | | | | 1150 |
| | | | | | | | | | | | | | | | 1155 |
| Ser | Tyr | Gln | Gly | Ser | Arg | Asn | Gly | Tyr | Leu | Gly | Gly | His | Gly | Phe | 1160 |
| | | | | | | | | | | | | | | | 1165 |
| | | | | | | | | | | | | | | | 1170 |
| Asn | Ala | Arg | Val | Met | Leu | Glu | Thr | Gln | Glu | Leu | Leu | Arg | Gln | Glu | 1175 |
| | | | | | | | | | | | | | | | 1180 |
| | | | | | | | | | | | | | | | 1185 |
| Gln | Arg | Arg | Lys | Glu | Gln | Gln | Met | Lys | Lys | Gln | Pro | Pro | Ser | Glu | 1190 |
| | | | | | | | | | | | | | | | 1195 |
| | | | | | | | | | | | | | | | 1200 |
| Gly | Pro | Ser | Asn | Tyr | Asp | Ser | Tyr | Lys | Lys | Val | Gln | Asp | Pro | Ser | 1205 |
| | | | | | | | | | | | | | | | 1210 |
| | | | | | | | | | | | | | | | 1215 |
| Tyr | Ala | Pro | Pro | Lys | Gly | Pro | Phe | Arg | Gln | Asp | Val | Pro | Pro | Ser | 1220 |
| | | | | | | | | | | | | | | | 1225 |
| | | | | | | | | | | | | | | | 1230 |
| Pro | Ser | Gln | Val | Ala | Arg | Leu | Asn | Arg | Leu | Gln | Thr | Pro | Glu | Lys | 1235 |
| | | | | | | | | | | | | | | | 1240 |
| | | | | | | | | | | | | | | | 1245 |
| Gly | Arg | Pro | Phe | Tyr | Ser | | | | | | | | | | 1250 |
| | | | | | | | | | | | | | | | 1255 |
| | | | | | | | | | | | | | | | 1260 |
| | | | | | | | | | | | | | | | 1265 |

【国際公開パンフレット(コレクトバージョン)】

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

(19) World Intellectual Property
Organization
International Bureau(43) International Publication Date
18 July 2002 (18.07.2002)

PCT

(10) International Publication Number
WO 2002/055706 A3(51) International Patent Classification: C12N 15/12, (74) Agents: STREETER, David, G. et al.; Incyte Genomics,
C07K 14/47, A61K 31/7088, C12Q 1/68, C12N 5/10, Inc., 3160 Porter Drive, Palo Alto, CA 94304 (US).
A61K 38/17, C07K 16/18, G01N 33/68, A61K 39/395(21) International Application Number:
PCT/US2002/000647

(22) International Filing Date: 8 January 2002 (08.01.2002)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
09/757,781 9 January 2001 (09.01.2001) US(71) Applicant (for all designated States except US): INCYTE
GENOMICS, INC. [US/US]; 3160 Porter Drive, Palo
Alto, CA 94304 (US).

(72) Inventors; and

(75) Inventors/Applicants (for US only): REDDY, Roopa
[IN/US]; 1233 W. McKinley Avenue #3, Sunnyvale, CA
94086 (US); TANG, Y., Tom [US/US]; 4230 Ranwick
Court, San Jose, CA 95118 (US); BAUGHN, Mariah, R.
[US/US]; 14244 Santiago Road, San Leandro, CA 94577
(US); KRASNOW, Randi [US/US]; 817 Santa Fe Avenue,
Stanford, CA 94305 (US).(81) Designated States (national): AE, AG, AL, AM, AT, AU,
AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU,
CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH,
GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC,
LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW,
MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK,
SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA,
ZW.(84) Designated States (regional): ARIPO patent (GH, GM,
KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW),
Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM),
European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR,
GB, GR, IE, IT, LI, LU, MC, NL, PT, SE, TR), OAPI patent
(BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR,
NE, SN, TD, TG).

Published:

— with international search report

(88) Date of publication of the international search report:
15 April 2004For 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 2002/055706 A3

(54) Title: ASIP-RELATED PROTEINS

(57) Abstract: The invention provides cDNAs which encode ASIP-related proteins. It also provides for the use of the cDNAs, fragments, complements, and variants thereof and of the encoded proteins, portions thereof and antibodies thereto for diagnosis and treatment of cancer, particularly bladder transitional cell carcinoma. The invention additionally provides expression vectors and host cells for the production of the protein and a transgenic model system.

【 国際調査報告 】

| INTERNATIONAL SEARCH REPORT | | International Application No. PCT/US 02/00647 |
|--|---|--|
| A. CLASSIFICATION OF SUBJECT MATTER IPC 7 C12N15/12 C07K14/47 A61K31/7088 C12Q1/68 C12N5/10 A61K38/17 C07K16/18 G01N33/68 A61K39/395 | | |
| 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) | | |
| Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched | | |
| Electronic data base consulted during the international search (name of data base and, where practical, search terms used) | | |
| EPO-Internal, EMBL, BIOSIS, WPI Data, PAJ, EMBASE | | |
| C. DOCUMENTS CONSIDERED TO BE RELEVANT | | |
| Category * | Citation of document, with indication, where appropriate, of the relevant passages | Relevant to claim No. |
| X | IZUMI Y ET AL: "AN ATYPICAL PKC DIRECTLY ASSOCIATES AND COLOCALIZES AT THE EPITHELIAL TIGHT JUNCTION WITH ASIP, A MAMMALIAN HOMOLOGUE OF CAENORHABDITIS ELEGANS POLARITY PROTEIN PAR-3" THE JOURNAL OF CELL BIOLOGY, ROCKEFELLER UNIVERSITY PRESS, US, vol. 143, no. 1, November 1998 (1998-11), pages 95-106, XP002933131 ISSN: 0021-9525 page 96, right-hand column, paragraph 3; figure 1 --- -/-- | 15-29 |
| <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 applicant 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 |
| 5 June 2003 | | 23.09.03 |
| Name and mailing address of the ISA European Patent Office, P.B. 5818 Patentlaan 2 NL - 2200 HV Rijswijk Tel: (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016 | | Authorized officer Seroz, T |

Form PCT/ISA/210 (second sheet) (July 1999)

INTERNATIONAL SEARCH REPORT

International Application No.
PCT/US 02/00647

| C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT | | |
|--|--|-----------------------|
| Category* | Citation of document, with indication, where appropriate, of the relevant passages | Relevant to claim No. |
| X | DATABASE EMBL 'Online! 1 December 2000 (2000-12-01) STRAUSBERG ET AL.: Database accession no. BF434006 XP002243449 the whole document | 1,2,4-29 |

Form PCT/ISA210 (continuation of second sheet) (July 1992)

| INTERNATIONAL SEARCH REPORT | national application No. PCT/US 02/00647 |
|--|--|
| 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. <input checked="" type="checkbox"/> Claims Nos.: because they relate to subject matter not required to be searched by this Authority, namely: Although claim 30 is directed to a method of treatment of the human/animal body, the search has been carried out and based on the alleged effects of the composition. | |
| 2. <input type="checkbox"/> Claims Nos.: because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically: | |
| 3. <input type="checkbox"/> Claims Nos.: because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a). | |
| Box II Observations where unity of invention is lacking (Continuation of Item 2 of first sheet) | |
| This International Searching Authority found multiple inventions in this International application, as follows: | |
| see additional sheet | |
| 1. <input type="checkbox"/> As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims. | |
| 2. <input type="checkbox"/> As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee. | |
| 3. <input type="checkbox"/> As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.: | |
| 4. <input checked="" type="checkbox"/> No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.: 1 (partially), 2 (completely), 4-30 (all partially) | |
| Remark on Protest | <input type="checkbox"/> The additional search fees were accompanied by the applicant's protest. <input type="checkbox"/> No protest accompanied the payment of additional search fees. |

International Application No. PCT/US 02/0647

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

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

1. Claims: Claims: 1 (partially), 2 (completely),
4-30 (all partially)

ASIP-related protein comprising SEQ ID No 1 and a polynucleotide comprising SEQ ID No 3 which encodes and identifies said ASIP-related protein. Expression vector, host cells, antibodies. Methods for detecting the expression of the ASIP-related protein coding gene and method for treating a bladder transitional cell carcinoma. Method for screening compounds that bind to the ASIP-related protein.

2. Claims: Claims: 1 (partially), 3 (completely),
4-30 (all partially)

ASIP-related protein comprising SEQ ID No 2 and a polynucleotide comprising SEQ ID No 20 which encodes and identifies said ASIP-related protein. Expression vector, host cells, antibodies. Methods for detecting the expression of the ASIP-related protein coding gene and method for treating a bladder transitional cell carcinoma. Method for screening compounds that bind to the ASIP-related protein.

フロントページの続き

| (51) Int.Cl. ⁷ | F I | テーマコード(参考) |
|---------------------------|----------------|------------|
| C 0 7 K 16/18 | C 0 7 K 14/47 | 4 C 0 8 5 |
| C 1 2 N 1/15 | C 0 7 K 16/18 | 4 H 0 4 5 |
| C 1 2 N 1/19 | C 1 2 N 1/15 | |
| C 1 2 N 1/21 | C 1 2 N 1/19 | |
| C 1 2 N 5/10 | C 1 2 N 1/21 | |
| C 1 2 N 15/02 | C 1 2 P 21/02 | C |
| C 1 2 P 21/02 | C 1 2 P 21/08 | |
| C 1 2 P 21/08 | 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 | D |
| G 0 1 N 33/53 | G 0 1 N 33/53 | M |
| G 0 1 N 33/566 | G 0 1 N 33/566 | |
| | C 1 2 N 5/00 | A |
| | C 1 2 N 15/00 | C |

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

(72) 発明者 タング、ワイ・トム

アメリカ合衆国カリフォルニア州 9 5 1 1 8 ・ サンノゼ ・ ランウィックコート 4 2 3 0

(72) 発明者 ボーグン、マライア・アール

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

(72) 発明者 クラスノー、ランディ

アメリカ合衆国カリフォルニア州 9 4 3 0 5 ・ スタンフォード ・ サンタフェアベニュー 8 1 7

F ターム(参考) 2G045 AA34 AA35 AA40 BA11 BB50 DA12 DA13 DA14 DA36 FB02

FB03

4B024 AA01 AA11 BA10 CA04 DA02 DA05 DA06 DA11 DA12 EA04

GA11 HA12

4B063 QA18 QA19 QQ08 QQ43 QR08 QR42 QR56 QS25 QS34 QX02

QX07

4B064 AG26 AG27 CA10 CA20 CC24 CE12 DA01 DA13

4B065 AA01X AA57X AA72X AA90X AA93Y AB01 BA02 CA29 CA44 CA46

4C085 AA13 AA14 BB11 CC02 CC22 CC23 GG01 GG08 GG10

4H045 AA10 AA11 AA20 AA30 BA10 CA41 DA76 DA89 EA28 EA51

FA72 FA74

| | | | |
|----------------|---|---------|------------|
| 专利名称(译) | ASIP相关蛋白 | | |
| 公开(公告)号 | JP2005503107A | 公开(公告)日 | 2005-02-03 |
| 申请号 | JP2002556752 | 申请日 | 2002-01-08 |
| [标]申请(专利权)人(译) | 洞察Genomics公司 | | |
| 申请(专利权)人(译) | 洞察基因组公司 | | |
| [标]发明人 | レディルーパ タングワイトム ボーグンマライアール クラスノーランディ | | |
| 发明人 | レディ、ルーパ タング、ワイトム ボーグン、マライアール クラスノー、ランディ | | |
| IPC分类号 | G01N33/50 A61K38/00 A61K39/395 A61P13/10 A61P35/00 C07K14/47 C07K16/18 C12N1/15 C12N1/19 C12N1/21 C12N5/10 C12N15/02 C12N15/09 C12N15/12 C12P21/02 C12P21/08 C12Q1/68 G01N33/15 G01N33/53 G01N33/566 | | |
| CPC分类号 | A61K38/00 A61P13/10 A61P35/00 C07K14/47 C12N2799/026 | | |
| FI分类号 | C12N15/00.ZNA.A A61K39/395.D A61K39/395.N A61P13/10 A61P35/00 C07K14/47 C07K16/18 C12N1/15 C12N1/19 C12N1/21 C12P21/02.C C12P21/08 C12Q1/68.A G01N33/15.Z G01N33/50.Z G01N33/53.D G01N33/53.M G01N33/566 C12N5/00.A C12N15/00.C | | |
| F-TERM分类号 | 2G045/AA34 2G045/AA35 2G045/AA40 2G045/BA11 2G045/BB50 2G045/DA12 2G045/DA13 2G045/DA14 2G045/DA36 2G045/FB02 2G045/FB03 4B024/AA01 4B024/AA11 4B024/BA10 4B024/CA04 4B024/DA02 4B024/DA05 4B024/DA06 4B024/DA11 4B024/DA12 4B024/EA04 4B024/GA11 4B024/HA12 4B063/QA18 4B063/QA19 4B063/QQ08 4B063/QQ43 4B063/QR08 4B063/QR42 4B063/QR56 4B063/QS25 4B063/QS34 4B063/QX02 4B063/QX07 4B064/AG26 4B064/AG27 4B064/CA10 4B064/CA20 4B064/CC24 4B064/CE12 4B064/DA01 4B064/DA13 4B065/AA01X 4B065/AA57X 4B065/AA72X 4B065/AA90X 4B065/AA93Y 4B065/AB01 4B065/BA02 4B065/CA29 4B065/CA44 4B065/CA46 4C085/AA13 4C085/AA14 4C085/BB11 4C085/CC02 4C085/CC22 4C085/CC23 4C085/GG01 4C085/GG08 4C085/GG10 4H045/AA10 4H045/AA11 4H045/AA20 4H045/AA30 4H045/BA10 4H045/CA41 4H045/DA76 4H045/DA89 4H045/EA28 4H045/EA51 4H045/FA72 4H045/FA74 | | |
| 优先权 | 09/757781 2001-01-09 US | | |
| 外部链接 | Espacenet | | |

摘要(译)

本发明提供了编码ASIP相关蛋白的cDNA。它还提供了cDNA，其片段，补体和变体以及编码的蛋白质，其部分和其抗体在诊断和治疗癌症特别是膀胱移行细胞癌中的用途。本发明还提供用于产生蛋白质的表达载体和宿主细胞以及转基因模型系统。

| SEQ ID _{ref} | SEQ ID _{var} | クローン _{var} | ライブラリ名 | N _{ref} 配列 | 同一性 |
|-----------------------|-----------------------|---------------------|-----------|---------------------|------|
| 3 | 12 | 702457609T1 | RACOMON05 | 2162-2619 | 84% |
| 3 | 13 | 702458746T1 | RACOMON05 | 2152-2379 | 84% |
| 3 | 14 | 701335936H1 | RALINON08 | 1322-1566 | 88% |
| 3 | 15 | 700639694H1 | RATONOT01 | 2361-2619 | 87% |
| 3 | 16 | 700639694F6 | RATONOT01 | 2447-2851 | 85% |
| 3 | 17 | 701191467H1 | RACOMON05 | 2063-2333 | 81% |
| 3 | 18 | 702771158H1 | CMLINOT02 | 615-710 | 87% |
| 3 | 19 | 701266650H1 | MOLUDITO8 | 1289-1544 | 91% |
| 20 | 40 | 702231139H1 | RAFANOT02 | 2500-2951 | 88% |
| 20 | 41 | 700273304F6 | RASJNOT01 | 2705-3074 | 85% |
| 20 | 42 | 700330856H1 | RALINON04 | 1550-1817 | 87% |
| 20 | 43 | 700273304H1 | RASJNOT01 | 2705-2951 | 86% |
| 20 | 44 | 701517518H1 | RAKITXT62 | 2149-2441 | 83% |
| 20 | 45 | 701834089T1 | RAKITXT19 | 2897-3074 | 85% |
| 20 | 46 | 701480437H1 | RALITXT42 | 2679-3024 | 88% |
| 20 | 47 | 701190235H1 | RACOMON05 | 2077-2333 | 80% |
| 20 | 48 | 700939688H1 | RALINON07 | 2971-3074 | 88% |
| 20 | 49 | 700939688F6 | RALINON07 | 2971-3074 | 88% |
| 20 | 50 | 702582937T1 | RASJUNN01 | 5276-5319 | 95% |
| 20 | 51 | 700299037F6 | RAFONOT02 | 5276-5334 | 91% |
| 20 | 52 | 701246488H1 | RALINON02 | 2284-2357 | 86% |
| 20 | 53 | 702759912H1 | CMLIUNN01 | 1-565 | 92% |
| 20 | 54 | 700112340H1 | MOOSUNR01 | 121-362 | 88% |
| 20 | 55 | 700827810H1 | MOBONOT01 | 3359-3701 | 87% |
| 20 | 56 | 700109331H1 | MOOSUNR01 | 2509-2569 | 100% |