

(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2005-508600

(P2005-508600A)

(43) 公表日 平成17年4月7日(2005.4.7)

(51) Int. Cl. ⁷	F I	テーマコード (参考)
C 1 2 N 15/09	C 1 2 N 15/00 Z N A A	2 G O 4 5
A 6 1 K 38/00	A 6 1 K 39/395 D	4 B O 2 4
A 6 1 K 39/395	A 6 1 K 39/395 N	4 B O 2 9
A 6 1 K 45/00	A 6 1 K 45/00	4 B O 6 3
A 6 1 P 3/04	A 6 1 P 3/04	4 B O 6 4
	審査請求 未請求 予備審査請求 有 (全 175 頁) 最終頁に続く	
(21) 出願番号 特願2002-552156 (P2002-552156)	(71) 出願人 301005050	
(86) (22) 出願日 平成13年12月19日 (2001.12.19)	インサイト・ゲノミックス・インコーポレ	
(85) 翻訳文提出日 平成15年6月18日 (2003.6.18)	イテッド	
(86) 国際出願番号 PCT/US2001/050256	アメリカ合衆国カリフォルニア州9430	
(87) 国際公開番号 W02002/050279	4・パロアルト・ポータードライブ 31	
(87) 国際公開日 平成14年6月27日 (2002.6.27)	60	
(31) 優先権主張番号 60/257,714	(74) 代理人 100089266	
(32) 優先日 平成12年12月21日 (2000.12.21)	弁理士 大島 陽一	
(33) 優先権主張国 米国 (US)	(72) 発明者 ボーグン、マライア・アール	
(31) 優先権主張番号 60/260,081	アメリカ合衆国カリフォルニア州9457	
(32) 優先日 平成13年1月5日 (2001.1.5)	7・サンレアンドロ・サンティアゴロード	
(33) 優先権主張国 米国 (US)	14244	
(31) 優先権主張番号 60/262,302		
(32) 優先日 平成13年1月16日 (2001.1.16)		
(33) 優先権主張国 米国 (US)		最終頁に続く

(54) 【発明の名称】 核酸関連タンパク質

(57) 【要約】

本発明はヒト核酸関連タンパク質 (N A A P) および、N A A P を同定しコードするポリヌクレオチドを提供する。本発明はまた、発現ベクター、宿主細胞、抗体、アゴニストおよびアンタゴニストをも提供する。本発明はまた、N A A P の異常発現に関連する疾患を診断、治療または予防する方法をも提供する。

【特許請求の範囲】

【請求項 1】

以下の (a) 乃至 (e) からなる群から選択した単離されたポリペプチド。

(a) SEQ ID NO:1-16 (配列番号 1 乃至 1 6) からなる群から選択した或るアミノ酸配列を持つポリペプチド

(b) SEQ ID NO:1-2およびSEQ ID NO:4-16からなる群から選択した或るアミノ酸配列と少なくとも90%が同一であるような天然アミノ酸配列を有するポリペプチド

(c) SEQ ID NO:3のアミノ酸配列に対して少なくとも95%が同一であるような天然アミノ酸配列を有するポリペプチド

(d) SEQ ID NO:1-16からなる群から選択したアミノ酸配列を有するポリペプチドの生物学的活性断片 10

(e) SEQ ID NO:1-16からなる群から選択したアミノ酸配列を有するポリペプチドの免疫原性断片

【請求項 2】

SEQ ID NO:1-16からなる群から選択したアミノ酸配列を有する、請求項 1 に記載の単離されたポリペプチド。

【請求項 3】

請求項 1 のポリペプチドをコードする、単離されたポリヌクレオチド。

【請求項 4】

請求項 2 のポリペプチドをコードする、単離されたポリヌクレオチド。 20

【請求項 5】

SEQ ID NO:17-32からなる群から選択したポリヌクレオチド配列を有する、請求項 4 に記載の単離されたポリヌクレオチド。

【請求項 6】

請求項 3 に記載のポリヌクレオチドに機能的に連結したプロモーター配列を有する組換えポリヌクレオチド。

【請求項 7】

請求項 6 に記載の組換えポリヌクレオチドを用いて形質転換した細胞。

【請求項 8】

請求項 6 に記載の組換えポリヌクレオチドを有する遺伝形質転換体。 30

【請求項 9】

請求項 1 のポリペプチドを生産する方法であって、

(a) 前記ポリペプチドの発現に好適な条件下で、請求項 1 のポリペプチドをコードする或るポリヌクレオチドに機能的に連結されたプロモーター配列を持つ組換えポリヌクレオチドで形質転換される細胞を培養する過程と、

(b) そのように発現した前記ポリペプチドを回収する過程、とからなることを特徴とする方法。

【請求項 10】

前記ポリペプチドが、SEQ ID NO:1-16からなる群から選択した或るアミノ酸配列を持つことを特徴とする、請求項 9 に記載の方法。 40

【請求項 11】

請求項 1 に記載のポリペプチドと特異結合するような単離された抗体。

【請求項 12】

以下の (a) 乃至 (g) からなる群から選択した単離されたポリヌクレオチド。

(a) SEQ ID NO:17-32からなる群から選択したポリヌクレオチド配列を有するポリヌクレオチド

(b) SEQ ID NO:17-18、SEQ ID NO:20-32からなる群から選択した或るポリヌクレオチド配列に対して少なくとも90%が同一であるような天然ポリヌクレオチド配列を有するポリヌクレオチド

(c) SEQ ID NO:19のポリヌクレオチド配列に対して少なくとも95%が同一であるよう 50

な天然ポリヌクレオチド配列を有するポリヌクレオチド

(d) (a) のポリヌクレオチドに相補的なポリヌクレオチド

(e) (b) のポリヌクレオチドに相補的なポリヌクレオチド

(f) (c) のポリヌクレオチドに相補的なポリヌクレオチド

(g) (a) ~ (f) のRNA等価物

【請求項13】

請求項12に記載のポリヌクレオチドの少なくとも60の連続したヌクレオチドを持つ単離されたポリヌクレオチド。

【請求項14】

請求項12のポリヌクレオチドの配列を持つ標的ポリヌクレオチドをサンプル中から検出する方法であって、

(a) 前記サンプル中の前記標的ポリヌクレオチドに相補的な配列を有する少なくとも20の連続したヌクレオチド群を持つプローブと前記サンプルとをハイブリダイズし、該プローブが、ハイブリダイゼーション複合体が前記プローブと前記標的ポリヌクレオチドまたはその断片群との間に形成される条件下で、前記標的ポリヌクレオチドへ特異的にハイブリダイズする過程と、

(b) 前記ハイブリダイゼーション複合体の有無を検出し、該複合体が存在する場合にはオプションでその量を検出する過程、とを含む方法。

【請求項15】

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

【請求項16】

請求項12のポリヌクレオチドの配列を持つ標的ポリヌクレオチドをサンプル中から検出する方法であって、

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

(b) 前記の増幅した標的ポリヌクレオチドまたはその断片の有無を検出し、該標的ポリヌクレオチドまたはその断片が存在する場合にはオプションでその量を検出する過程、とを含むことを特徴とする方法。

【請求項17】

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

【請求項18】

前記ポリペプチドが、SEQ ID NO:1-16からなる群から選択した或るアミノ酸配列を持つことを特徴とする、請求項17の組成物。

【請求項19】

機能的なNAAPの発現の低下に関連する疾患や病態の治療方法であって、そのような治療を要する患者への請求項17の組成物の投与を含む治療方法。

【請求項20】

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

(a) 請求項1のポリペプチドを有する或るサンプルを或る化合物に曝す過程と、

(b) 前記サンプルにおけるアゴニスト活性を検出する過程、とを含むことを特徴とするスクリーニング方法。

【請求項21】

請求項20の方法で同定したアゴニスト化合物を有し、薬剤として許容できる賦形剤をも有する組成物。

【請求項22】

機能的なNAAPの発現の低下に関連する疾患や病態の治療方法であって、そのような治療を要する患者への請求項21の組成物の投与を含む治療方法。

10

20

30

40

50

【請求項 23】

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

- (a) 請求項 1 のポリペプチドを有する或るサンプルを或る化合物に曝す過程と、
- (b) 前記サンプルにおけるアンタゴニスト活性を検出する過程、とを含むことを特徴とするスクリーニング方法。

【請求項 24】

請求項 23 の方法で同定したアンタゴニスト化合物と、薬剤として許容できる賦形剤とを有する組成物。

【請求項 25】

機能的な NAAP の過剰な発現に関連する疾患や病態の治療方法であって、そのような治療を要する患者への請求項 24 の組成物の投与を含む治療方法。

10

【請求項 26】

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

- (a) 請求項 1 に記載のポリペプチドを適切な条件下で少なくとも 1 つの試験化合物と混合する過程と、
- (b) 請求項 1 に記載のポリペプチドの試験化合物との結合を検出し、それによって請求項 1 に記載のポリペプチドに特異結合する化合物を同定する過程、とを含むことを特徴とする方法。

【請求項 27】

請求項 1 のポリペプチドの活性をモジュレート（調節）する化合物をスクリーニングする方法であって、

- (a) 請求項 1 に記載のポリペプチドの活性が許容された条件下で、請求項 1 に記載のポリペプチドを少なくとも 1 つの試験化合物と混合する過程と、
- (b) 請求項 1 に記載のポリペプチドの活性を試験化合物の存在下で算定する過程と、
- (c) 試験化合物の存在下での請求項 1 に記載のポリペプチドの活性を、試験化合物の不存在下での請求項 1 に記載のポリペプチドの活性と比較する過程とを含み、試験化合物の存在下での請求項 1 に記載のポリペプチドの活性の変化が、請求項 1 に記載のポリペプチドの活性を調節する化合物を標示することを特徴とする方法。

20

【請求項 28】

請求項 5 の配列を持つ標的ポリヌクレオチドの発現を改変するのに効果的な化合物をスクリーニングする方法であって、

- (a) 前記標的ポリヌクレオチドの発現に好適な諸条件下で、前記標的ポリヌクレオチドを有する或るサンプルを或る化合物に曝露する過程と、
- (b) 前記標的ポリヌクレオチドの、改変された発現を検出する過程と、
- (c) 可変量の前記化合物の存在下と前記化合物の不存在下で、前記標的ポリヌクレオチドの発現を比較する過程、とを含むことを特徴とする方法。

30

【請求項 29】

試験化合物の毒性を算定する方法であって、

- (a) 核酸群を有する或る生体サンプルを前記試験化合物で処理する過程と、
- (b) 処理した前記生体サンプルの核酸群と、請求項 12 のポリヌクレオチドの少なくとも 20 の連続するヌクレオチド群を有する或るプローブをハイブリダイズさせる過程であって、このハイブリダイゼーションが、前記プローブと前記生体サンプルの或る標的ポリヌクレオチドとの間で或る特異的なハイブリダイゼーション複合体が形成される諸条件下で行われ、前記標的ポリヌクレオチドが、請求項 12 のポリヌクレオチドのポリヌクレオチド配列またはその断片を有するポリヌクレオチドである、前記過程と、
- (c) ハイブリダイゼーション複合体の収量を定量する過程と、
- (d) 前記処理された生体サンプル中の前記ハイブリタイゼーション複合体の量を、或る処理されていない生体サンプル中の前記ハイブリタイゼーション複合体の量と比較する過程とを含み、前記処理された生体サンプル中の前記ハイブリタイゼーション複合体の量の

40

50

差が、前記試験化合物の毒性を標示することを特徴とする方法。

【請求項30】

生物学的サンプル中のNAAPの発現に関連する症状または疾患に対する診断試験法であって、

(a) 前記抗体が前記ポリペプチドに結合し、抗体とポリペプチドとの複合体が形成されるのに適した諸条件下で、前記生体サンプルを請求項11に記載の抗体と混合する過程と、
(b) 前記複合体を検出する過程とを含み、前記複合体の存在が、前記生体サンプル中の前記ポリペプチドの存在と相関することを特徴とする方法。

【請求項31】

請求項11の抗体であって、

10

- (a) キメラ抗体
- (b) 単鎖抗体
- (c) Fab断片
- (d) F(ab')₂断片
- (e) ヒト化抗体、のいずれかであることを特徴とする抗体。

【請求項32】

請求項11の抗体と、許容できる賦形剤とを有する組成物。

【請求項33】

被検者におけるNAAPの発現に関連する症状又は疾患の診断方法であって、請求項32に記載の組成物の有効量を前記被検者に投与する過程を含むことを特徴とする方法。

20

【請求項34】

前記抗体が標識されることを特徴とする請求項32の組成物。

【請求項35】

被検者におけるNAAPの発現に関連する症状又は疾患の診断方法であって、請求項34に記載の組成物の有効量を前記被検者に投与する過程を含むことを特徴とする方法。

【請求項36】

請求項11の抗体の特異性を持つポリクローナル抗体を調製する方法であって、

- (a) 抗体応答を誘発する条件下で、SEQ ID NO:1-16を有する群から選択したアミノ酸配列またはその免疫原性断片を含むポリペプチドを用いて或る動物を免疫化する過程と、
- (b) 前記動物から抗体を単離する過程と、
- (c) 前記単離された抗体を該ポリペプチドでスクリーニングし、それによって、SEQ ID NO:1-16からなる群から選択したアミノ酸配列を有するポリペプチドに特異結合するようなポリクローナル抗体を同定する過程とを含むことを特徴とする方法。

30

【請求項37】

請求項36の方法で産生したポリクローナル抗体。

【請求項38】

請求項37のポリクローナル抗体と好適なキャリアとを含む組成物。

【請求項39】

請求項11に記載の抗体の特異性を有するモノクローナル抗体を作製する方法であって、

- (a) 抗体応答を誘発する条件下で、SEQ ID NO:1-16を有する群から選択したアミノ酸配列またはその免疫原性断片を含むポリペプチドを用いて或る動物を免疫化する過程と、
- (b) 前記動物から、抗体を産出する細胞を単離する過程と、
- (c) 不死化した細胞と前記抗体産出細胞とを融合し、モノクローナル抗体を産出するハイブリドーマ細胞を形成する過程と、
- (d) 前記ハイブリドーマ細胞を培養する過程と、
- (e) SEQ ID NO:1-16からなる群から選択したアミノ酸配列を有するポリペプチドに特異結合するようなモノクローナル抗体を前記培養物から単離する過程とを含むことを特徴とする方法。

40

【請求項40】

請求項39に記載の方法で産出したモノクローナル抗体。

50

【請求項41】

請求項40に記載のモノクローナル抗体と適切なキャリアとを含む組成物。

【請求項42】

Fab発現ライブラリをスクリーニングすることにより産出されることを特徴とする請求項11に記載の抗体。

【請求項43】

組換え免疫グロブリンライブラリをスクリーニングすることにより産出されることを特徴とする、請求項11に記載の抗体。

【請求項44】

SEQ ID NO:1-16からなる群から選択したアミノ酸配列を有するポリペプチドをサンプル中に検出する方法であって、

(a) 前記抗体と前記ポリペプチドとの特異結合を許容する条件下で、或るサンプルと共に請求項11に記載の抗体をインキュベートする過程と、

(b) 特異結合を検出する過程とを含み、該特異結合が、SEQ ID NO:1-16からなる群から選択した或るアミノ酸配列を有する或るポリペプチドがサンプル中に存在することを標示することを特徴とする方法。

【請求項45】

或るサンプルからのSEQ ID NO:1-16からなる群から選択したアミノ酸配列を有するポリペプチドを精製する方法であって、

(a) 前記抗体と前記ポリペプチドとの特異結合を許容する条件下で、或るサンプルと共に請求項11に記載の抗体をインキュベートする過程と、

(b) 前記サンプルから前記抗体を分離し、SEQ ID NO:1-16を有する群から選択したアミノ酸配列を有する精製ポリペプチドを得る過程、とを含むことを特徴とする方法。

【請求項46】

マイクロアレイの少なくとも1つのエレメントが請求項13に記載のポリヌクレオチドであることを特徴とするマイクロアレイ。

【請求項47】

ポリヌクレオチドを有するサンプルの発現プロフィールを作製する方法であって、

(a) サンプルのポリヌクレオチドを標識する過程と、

(b) ハイブリダイゼーション複合体の形成に適した諸条件下で請求項46のマイクロアレイのエレメント群をサンプルの標識されたポリヌクレオチド群と接触させる過程と、

(c) 該サンプル中の該ポリヌクレオチド群の発現を定量化する過程、とを含むことを特徴とする方法。

【請求項48】

固体基板上の固有の物理的位置に付着された種々のヌクレオチド分子を含むアレイであって、少なくとも1つの前記ヌクレオチド分子が、標的ポリヌクレオチドの少なくとも30の連続したヌクレオチドと特異的にハイブリダイズ可能な最初のオリゴヌクレオチドまたはポリヌクレオチド配列を有し、前記の標的ポリヌクレオチドが請求項12に記載のポリヌクレオチドであることを特徴とするアレイ。

【請求項49】

請求項48に記載のアレイで、前記のオリゴヌクレオチドまたはポリヌクレオチドの最初の配列が前記の標的ポリヌクレオチドの少なくとも30の連続したヌクレオチドに完全に相補的であることを特徴とするアレイ。

【請求項50】

請求項48に記載のアレイで、前記のオリゴヌクレオチドまたはポリヌクレオチドの最初の配列が前記の標的ポリヌクレオチドの少なくとも60の連続したヌクレオチドに完全に相補的であることを特徴とするアレイ。

【請求項51】

請求項48に記載のアレイで、前記のオリゴヌクレオチドまたはポリヌクレオチドの最初の配列が前記の標的ポリヌクレオチドに完全に相補的であることを特徴とするアレイ。

10

20

30

40

50

【請求項 5 2】

請求項48に記載のアレイで、マイクロアレイであることを特徴とするアレイ。

【請求項 5 3】

請求項48に記載のアレイで、更に前記のオリゴヌクレオチドまたはポリヌクレオチドの最初の配列を有するヌクレオチド分子にハイブリダイズした前記の標的ポリヌクレオチドを有することを特徴とするアレイ。

【請求項 5 4】

請求項48に記載のアレイで、リンカーが少なくとも1つの前記のヌクレオチド分子と前記の固体基板とを連結していることを特徴とするアレイ。

【請求項 5 5】

請求項48に記載のアレイで、基板上の固有の物理的位置の各々が複数のヌクレオチド分子を含み、任意の単一の固有の物理的位置でのその複数のヌクレオチド分子は同一の配列を有し、基板上の固有の物理的位置の各々は、基板上の別の固有の物理的位置でのヌクレオチド分子の配列とは異なる配列を有するヌクレオチド分子を含むことを特徴とするアレイ。

10

【請求項 5 6】

SEQ ID NO:1のアミノ酸配列を有する請求項1に記載のポリペプチド。

【請求項 5 7】

SEQ ID NO:1のアミノ酸配列を有する請求項2に記載のポリペプチド。

【請求項 5 8】

SEQ ID NO:1のアミノ酸配列を有する請求項3に記載のポリペプチド。

20

【請求項 5 9】

SEQ ID NO:1のアミノ酸配列を有する請求項4に記載のポリペプチド。

【請求項 6 0】

SEQ ID NO:1のアミノ酸配列を有する請求項5に記載のポリペプチド。

【請求項 6 1】

SEQ ID NO:1のアミノ酸配列を有する請求項6に記載のポリペプチド。

【請求項 6 2】

SEQ ID NO:1のアミノ酸配列を有する請求項7に記載のポリペプチド。

【請求項 6 3】

SEQ ID NO:1のアミノ酸配列を有する請求項8に記載のポリペプチド。

30

【請求項 6 4】

SEQ ID NO:1のアミノ酸配列を有する請求項9に記載のポリペプチド。

【請求項 6 5】

SEQ ID NO:1のアミノ酸配列を有する請求項10に記載のポリペプチド。

【請求項 6 6】

SEQ ID NO:1のアミノ酸配列を有する請求項11に記載のポリペプチド。

【請求項 6 7】

SEQ ID NO:1のアミノ酸配列を有する請求項12に記載のポリペプチド。

【請求項 6 8】

SEQ ID NO:1のアミノ酸配列を有する請求項13に記載のポリペプチド。

40

【請求項 6 9】

SEQ ID NO:1のアミノ酸配列を有する請求項14に記載のポリペプチド。

【請求項 7 0】

SEQ ID NO:1のアミノ酸配列を有する請求項15に記載のポリペプチド。

【請求項 7 1】

SEQ ID NO:1のアミノ酸配列を有する請求項16に記載のポリペプチド。

【請求項 7 2】

SEQ ID NO:17のポリヌクレオチド配列を有する請求項12に記載のポリヌクレオチド。

【請求項 7 3】

50

SEQ ID NO:18のポリヌクレオチド配列を有する請求項12に記載のポリヌクレオチド。

【請求項74】

SEQ ID NO:19のポリヌクレオチド配列を有する請求項12に記載のポリヌクレオチド。

【請求項75】

SEQ ID NO:20のポリヌクレオチド配列を有する請求項12に記載のポリヌクレオチド。

【請求項76】

SEQ ID NO:21のポリヌクレオチド配列を有する請求項12に記載のポリヌクレオチド。

【請求項77】

SEQ ID NO:22のポリヌクレオチド配列を有する請求項12に記載のポリヌクレオチド。

【請求項78】

SEQ ID NO:23のポリヌクレオチド配列を有する請求項12に記載のポリヌクレオチド。

【請求項79】

SEQ ID NO:24のポリヌクレオチド配列を有する請求項12に記載のポリヌクレオチド。

【請求項80】

SEQ ID NO:25のポリヌクレオチド配列を有する請求項12に記載のポリヌクレオチド。

【請求項81】

SEQ ID NO:26のポリヌクレオチド配列を有する請求項12に記載のポリヌクレオチド。

【請求項82】

SEQ ID NO:27のポリヌクレオチド配列を有する請求項12に記載のポリヌクレオチド。

【請求項83】

SEQ ID NO:28のポリヌクレオチド配列を有する請求項12に記載のポリヌクレオチド。

【請求項84】

SEQ ID NO:29のポリヌクレオチド配列を有する請求項12に記載のポリヌクレオチド。

【請求項85】

SEQ ID NO:30のポリヌクレオチド配列を有する請求項12に記載のポリヌクレオチド。

【請求項86】

SEQ ID NO:31のポリヌクレオチド配列を有する請求項12に記載のポリヌクレオチド。

【請求項87】

SEQ ID NO:32のポリヌクレオチド配列を有する請求項12に記載のポリヌクレオチド。

【発明の詳細な説明】

【0001】

(技術分野)

本発明は、核酸関連タンパク質の核酸配列およびアミノ酸配列に関する。本発明はまた、これらの配列を利用した、細胞増殖異常、神経障害、発達障害、自己免疫/炎症性疾患、および感染症の、診断・治療・予防に関する。本発明は更に、核酸関連タンパク質の核酸配列およびアミノ酸配列の発現における、外来化合物の効果についての算定に関する。

【0002】

(発明の背景)

多細胞生物は、構造及び機能が著しく異なる多様な細胞型から構成されている。細胞の個性は、特徴的な遺伝子発現パターンによって決定される。異なった細胞型は発生の間、重複するが固有の遺伝子群を発現する。遺伝子発現の空間的及び時間的な調節は、生物の発達に寄与する、細胞増殖、細胞分化、アポトーシス、及び他のプロセスの制御にとって極めて重要である。更に、遺伝子発現は、細胞外シグナルに応答して調節される。細胞外シグナルは、細胞間伝達を仲介し、様々な細胞型の作用を協調させる。適切な遺伝子調節はまた、その機能が要求される遺伝子群のみを所定の時期に発現させることによって、細胞を確実に効率よく機能させる。

転写因子

転写調節タンパク質は、遺伝子発現の制御に必須である。これらのタンパク質の一部は転写因子として機能し、遺伝子の転写を開始、活性化、抑制し、または終了させる。転写因子は一般に遺伝子のプロモーター、エンハンサー、または上流調節領域に、配列特異的に

10

20

30

40

50

結合する(ただし、一部の因子は遺伝子コード領域内またはその下流の調節エレメントに結合する)。転写因子はDNAの特定の領域に単独で結合するか、または他のアクセサリー因子との複合体として結合する(概説は Lewin, B. (1990) *Genes IV*, Oxford University Press, New York, NY, および Cell Press, Cambridge, MA, 554570ページを参照)。

【0003】

DNAの二重らせん構造とリピート配列群とがトポロジックおよび化学的な特徴を生み出し、これらの特徴は転写因子によって認識され得る。これらの特徴とは、水素結合の供与基と受容基、疎水性パッチ、主溝と副溝、および、配列の規則的な繰り返しストレッチ群(repeated stretches: ヘリックス内に特有な折れ曲がり群を誘導)である。通常、転写因子が特異的に認識するDNA配列モチーフ群は、約20ヌクレオチドの長さである。複数の、隣接した転写因子結合モチーフ群が遺伝子調節に必要な場合もある。

10

【0004】

多くの転写因子は、DNA結合の構造的モチーフ群を組み込む。これらのモチーフはヘリックスまたはシートからなり、DNAの主溝に結合する。4種の良く特徴付けられたモチーフが、ヘリックス・ターン・ヘリックス、Znフィンガー、ロイシン・ジッパー、およびヘリックス・ループ・ヘリックスである。これらのモチーフを持つタンパク質は、単独で単量体として作用したり、または、ホモ二量体やヘテロ二量体を形成してDNAと相互作用する。

【0005】

ヘリックス・ターン・ヘリックス・モチーフは2つのヘリックスからなり、これらは、決まった角度で、アミノ酸群の或る短鎖によって接続されている。一方のヘリックスが、主溝に結合する。ヘリックス・ターン・ヘリックス・モチーフの1例はホメオボックス・モチーフであり、これはホメオドメイン蛋白質に在る。ホメオドメイン蛋白質は発生期に前後体軸(anterior-posterior body axis)を決めるのに重要であり、動物界全体で保存されている。キイロショウジョウバエのアンテナペディア(Antennapedia)蛋白質およびウルトラバイソラックス(Ultrabithorax)蛋白質は、原型的なホメオドメイン蛋白質である(Pabo, C.O. および R.T. Sauer (1992) *Ann. Rev. Biochem.* 61:1053-1095)。

20

【0006】

マウスHES6は、Hairy/Enhancer-of-split (HES)ファミリーの、塩基性ヘリックス・ループ・ヘリックス転写因子のメンバーである。HES遺伝子は、ノッチ・シグナリングの核内エフェクターとして作用し、数種のNotch標的遺伝子の転写活性を調節する。HES6は全ての神経原基(neurogenic placodes)、その派生物、および脳内に発現され、これらの部位で、前後軸と背腹軸とに沿ってパターン化される。HES6はまた、胚組織でも発現され、胚組織ではNotchシグナリングは細胞運命の決定を制御する。細胞運命としては例えば、体幹、後根神経節、筋節(myotomes)、および胸腺がある。肢芽ではHES6は、骨格筋と予定運命の腱に発現される。HES6はまた、胚の呼吸系、尿系、消化系の、上皮細胞にも発現される(Vasiliauskas, D. および Stern C.D. (2000) *Mech. Dev.* 98:133137; Pissarra, L. 他 (2000) *Mech Dev* 95:275278)。

30

【0007】

Znフィンガーモチーフは亜鉛イオンに結合する。このモチーフは一般に、周期的に配置されたシステイン残基とヒスチジン残基からなる約30アミノ酸のタンデムリピート群を持つ。この配列パターンの2例、C2H2とC3HC4(「RING」フィンガー)については既に記載されている(Lewin, 前出)。Znフィンガータンパク質の各々が1つの逆平行シート及び1つのヘリックスを持ち、それらの近接性とコンフォメーションとは、亜鉛イオンによって維持されている。DNAとの接触は、ヘリックスの前のアルギニン残基によって、並びに、ヘリックスの第2の残基、第3の残基、及び第6の残基によってなされる。Znフィンガーモチーフの変異体の例には、あまり解明されていないシステインリッチモチーフ群があり、これらの変異体モチーフは、亜鉛または他の金属イオンと結合する。これらの変異体モチーフはヒスチジン残基を持たない場合があり、一般に非反復である。タンパク質の各ZnフィンガーのヘリックスがDNA二重らせんの主溝と接触できるように、Znフィン

40

50

ガーモチーフはそのタンパク質内にタンデム型にリピートされ得る。この、タンパク質とDNAとの間の接触の繰り返しによって、強力且つ特異的なDNA-タンパク質相互作用が起こる。この相互作用の強度及び特異性は、タンパク質内のZnフィンガーモチーフの数によって調節され得る。Znフィンガーは、初めはDNAと直接相互作用する領域としてDNA結合タンパク質群において同定されたが、DNAと結合しない種々のタンパク質にも出現する(Lodish, H.他(1995) *Molecular Cell Biology*, Scientific American Books, New York NY, 447-451ページ)。例えば、Galcheva-Gargova, Z. 他((1996) *Science* 272:1797-1802)は、種々のサイトカイン受容体と相互作用するZnフィンガー蛋白質群を同定している。

【0008】

C2H2タイプのZnフィンガーシグネチャモチーフには、或る28アミノ酸の配列があり、この配列には、2つの保存されたCys残基、および2つの保存されたHis残基が、1つのC2C12H3Hタイプのモチーフ内に含まれる。C2H2型モチーフは一般に、多重タンデムリピート内に出現する。或るシステインリッチドメインであってモチーフAsp-His-His-Cysを持つドメイン(DHHC-CRD)は、Znフィンガー蛋白質の独立サブグループとして同定されている。DHHC-CRD領域は、成長と発達とに関与すると思われる。或るDHHC-CRD突然変異体は、Rasの機能欠損を示す。Rasは小さな膜結合型GTP結合タンパク質であり、細胞の成長と分化とを調節する。一方、他のDHHC-CRDタンパク質類は、おそらくRasとは無関係の経路群で機能する(Bartels, D.J. 他(1999) *Mol. Cell Biol.* 19:6775-6787)。

10

【0009】

Znフィンガー転写因子は、しばしば、モジュラー配列モチーフ群を伴う。このモチーフの例は、Kruppel-associated box (KRAB) およびSCANドメインである。例えば、低リポ蛋白血症(hypoalphalipoproteinemia)感受性遺伝子ZNF202は、或るSCANボックスと或るKRABドメイン(8つのC2H2 Znフィンガーモチーフ群が続く)とをコードする(Honer, C. 他(2001) *Biochim. Biophys. Acta* 1517:441448)。SCANドメインは、高度に保存されたロイシンリッチモチーフであり、約60アミノ酸からなり、Znフィンガー転写因子のアミノ末端に見られる。SCANドメインは、最も多くはC2H2 Znフィンガーモチーフに連結しており、この連結は、それらのカルボキシル末端を通じてなされる。生化学的な結合研究は、SCANドメインを選択的なヘテロおよびホモタイプのオリゴマー化ドメインとして確定している。SCANドメインが媒介するタンパク質複合体は、転写因子の生物学的機能を変調させるよう機能するようである(Schumacher, C. 他(2000) *J. Biol. Chem.* 275:17173-17179)。

20

30

【0010】

KRAB (Kruppel-associated box)ドメインは、保存されたアミノ酸配列であって約75アミノ酸にまたがる。KRABドメインは、C2H2 Znフィンガーをコードする300~700種の遺伝子のほぼ3分の1に見られる。KRABドメインは、フィンガーリピートに対してN末端側に見られる。KRABドメインは、一般に2つのエキソンによってコードされる。KRAB-A領域またはボックスが1つのエキソンによってコードされ、KRAB-B領域またはボックスが第2エキソンによってコードされる。KRABドメインの機能は、転写の抑制である。転写抑制を達成するため、KRAB-associated protein-1(1種の転写コリプレッサー)またはKRAB-A相互作用タンパク質のどちらかが動員される。KRABドメインを持つタンパク質は、発生期に調節的な役割を果たすようである(Williams, A.J. 他(1999) *Mol. Cell Biol.* 19:8526-8535)。高度に近縁のヒトKRAB Znフィンガー蛋白質群は全ヒト組織で検出し得るが、或るサブグループは、ヒトTリンパ球に高度に発現される(Bellefroid, E.J. 他(1993) *EMBO J.* 12:13631374)。ZNF85 KRAB Znフィンガー遺伝子はヒトZNF91ファミリーのメンバーであり、正常成人精巣や、精上皮腫(seminoma)や、NT2/D1 奇形癌腫細胞株に高度に発現される(Poncelet, D.A. 他(1998) *DNA Cell Biol.* 17:931943)。

40

【0011】

Kruppelタンパク質は、ショウジョウバエの分節化を調節する。ヒトゲノム全体には、このようなタンパク質をコードする約300種の遺伝子がある。事実、Kruppelマルチフィンガー蛋白質類をコードする100種を超える別々のmRNAが(その殆どは新規である)、ヒト胎盤に発見されている。これらの蛋白質の9種にある106フィンガーリピート群の配列群は

50

、高度に相通的である。ヘリカル構造内には、可変性を示す幾つかの位置がある。研究の示すところでは、この可変性は、これら蛋白質のDNA結合特異性に影響する(Bellefroid, E.J. 他 (1989) DNA 8:377387)。

【 0 0 1 2 】

ZNF143は、ヒトZnフィンガー-Kruppelファミリー蛋白質のGLIタイプである。ZNF143は、アフリカツメガエルのセレノシステイン(selenocysteine) tRNA遺伝子転写活性化因子(Staf)に対して84%同一である。Stafは、核内低分子RNA(snRNA)およびsnRNAタイプの遺伝子の、RNAポリメラーゼII(Pol II)およびIII(Pol III)による転写の増強に関与すると思われる。Stafはまた、Pol II mRNAプロモーターからの発現を刺激する能力を持つ。ZNF143は、近縁のZNF138およびZNF139と共に、染色体領域7q11.2、7q21.3~q22.1、および11p15.3~p15.4に局在する。これらの領域は、Williams症候群、裂手裂足症(split hand and foot disease: SHFD1)、BeckwithWiedemann症候群にそれぞれ伴う欠失および/または転座に関与し、これは、ZNF143遺伝子が発達障害および悪性疾患に関与することを示唆する。ZNF143 mRNAは、多くの正常な成人組織に発現される。発現組織の例は、白血球、結腸、小腸、卵巣、精巣、前立腺、胸腺、および脾臓の組織を含む。更に、マウスのシャペロンをコードするCcta遺伝子は、ZNF143と、別のStafファミリーZnフィンガー転写因子であるZNF76とによって調節され、これは、これらのRNAおよびシャペロン遺伝子が同時制御されて活発な細胞成長期に成熟タンパク質の合成を促進することを示唆する(Tommerup, N. および Vissing, H. (1995) Genomics 27: 259264 ; Myslinski, E. 他 (1998) J. Biol. Chem. 273:219982006; Kubota, H. 他 (2000) J. Biol. Chem. 275:2864128648)。

【 0 0 1 3 】

C4モチーフは、ホルモン調節されるタンパク質に見られる。C4モチーフは一般に、唯2つのリピートのみを持つ。多数の真核生物タンパク質およびウイルス蛋白質には、或る保存されたシステインリッチドメインがあり、これは40~60残基からなり、C3HC4 ZnフィンガーまたはRINGフィンガーと呼ばれる。このドメインは亜鉛2原子と結合する。また、おそらくタンパク質間相互作用の仲介に関与する。亜鉛ライゲーション系の3次元の「十字(crossbrace)」構造は、RINGドメイン独自のものである。このようなドメイン内のシステイン群の間隔は、Cx(2)Cx(9~39)Cx(1~3)Hx(2~3)Cx(2)Cx(4~48)Cx(2)Cである。PHDフィンガーは、或るC4HC3 Znフィンガー様モチーフであって、クロマチンが媒介する転写調節に関与すると思われる核タンパク質に見られる。

【 0 0 1 4 】

GATAタイプの転写因子には1つまたは2つのZnフィンガードメインがあり、これらのZnフィンガードメインは、連続したヌクレオチド配列であるGATAを有するDNAの領域に特異結合する。NMR研究の示すところでは、このZnフィンガーには2枚の不規則な逆平行シートと1つのヘリックスとがあり、続いて、フィンガーのC末端までの長いループがある。(Ominchinski, J.G. (1993) Science 261:438446)。このヘリックスと、2枚のシートを接続するループとがDNAの主溝と接触し、一方、結合の特性を決める部位であるC末端部は、副溝内に巻き付けて入り込む。

【 0 0 1 5 】

LIMモチーフには約60アミノ酸残基があり、7つの保存されたシステイン残基と1つのヒスチジンとが、1つのコンセンサス配列内に含まれる(Schmeichel, K.L. および Beckerle, M.C. (1994) Cell 79:211-219)。LIMファミリーには、発達、分化、および細胞成長に関与し得る、転写因子と細胞骨格タンパク質とを含む。一例のアクチン結合LIMタンパク質は、細胞骨格と細胞形態形成との調節において役割を果たし得る(Roof, D.J. 他 (1997) J. Cell Biol. 138:575-588)。アクチン結合LIMタンパク質のN末端ドメインには、4つの二重Znフィンガーモチーフ群があり、LIMコンセンサス配列を有する。アクチン結合LIMタンパク質のC末端ドメインは、既知のアクチン結合タンパク質(例えばデマチン: dematin およびビリン: villin)との配列類似性を示す。アクチン結合LIMタンパク質は、Fアクチンと結合する。この結合は、デマチン様C末端ドメインを通じて行う。LIMドメインは、他のLIM結合タンパク質とのタンパク質間相互作用を仲介するように思われる。

【0016】

骨髄性細胞の発達は、組織特異的な転写因子群によって制御される。骨髄性Znフィンガー蛋白質(MZF)としては、MZF-1とMZF-2とがある。MZF-1は好中性顆粒球の発生の調節において機能する。或るネズミ相同体MZF2は、骨髄性細胞群、特に好中球系にコミットされた細胞に発現される。MZF-2は、GCSFによって下方調節される。また、好中球発生において独自の機能を持つようである(Murai, K. 他 (1997) *Genes Cells* 2:581591)。

【0017】

ロイシンジッパーモチーフは、ロイシンに富むアミノ酸群の或るストレッチからなり、これは両親媒性のヘリックスを形成できる。この構造は、2つのロイシンジッパー蛋白質の二量体化のための基盤を提供する。ロイシンジッパーに隣接する領域は通常は塩基性であり、タンパク質二量体化すると、主溝に結合するための最適な配置となる。このようなモチーフを持つタンパク質を総称してbZIP転写因子と呼ぶ。ロイシンジッパーモチーフはプロト癌遺伝子であるFosとJunとに見られる。FosとJunとはヘテロ二量体転写因子であるAP1を有する。AP1は、細胞成長と、細胞系統の決定とに関わる(Papavassiliou, A.G. (1995) *N. Engl. J. Med.* 332:45-47)。

10

【0018】

マウスのkreisler (kr)突然変異は、尾側後脳(caudal hindbrain)における分節化異常(segmentation abnormalities)と、内耳発達の欠損とを生じる。kr cDNAは、或る塩基性ドメインロイシンジッパー(bZIP)転写因子をコードする。この転写因子では、全ての既知のbZIPファミリーメンバーのDNA結合ドメイン内に保存された1つのアスパラギン残基の代わりにセリンが置き換わっている。krの同一性、発現、および突然変異表現型は、軸パターン形成(axial patterning)における初期の役割を標示する他、後脳と耳との発達を支配する分子機序と発生学的機序との識見を提供する(Cordes, S.P. および Barsh, G.S. (1994) *Cell* 79:10251034)。

20

【0019】

ヘリックス・ループ・ヘリックス・モチーフ(HLH)は1つの短いヘリックスと、これが1つのループによって接続された長いヘリックスとからなる。このループは柔軟なので、2本のヘリックスは、互いに対して折れ曲がることができ、DNAに結合できる。転写因子であるMyc1は、原型的なHLHモチーフを持つ。

【0020】

NF B/Relシグネチャは、或るファミリーの真核細胞転写因子を規定する。この転写因子は、発癌や、胚の発生、分化、および免疫応答に関与する。Rel 相同性ドメイン(RHD)を持つ殆どの転写因子は、二量体として、或るコンセンサスDNA配列モチーフである Bと結合する。Rel ファミリのメンバーは、高度に保存された300アミノ酸ドメインであるRel 相同性ドメインを共有する。特徴的なRel C末端ドメインは、遺伝子活性化および細胞質固着の機能に関わる。RHDドメインを持つことが知られるタンパク質の例には、脊椎動物核因子NF B(或るDNA結合サブユニットと転写因子p65とのヘテロ二量体)、哺乳類転写因子RelB、および脊椎動物プロト癌遺伝子cre1があり、cre1は、分化とリンパ球新生とに関わるタンパク質である(Kabrun, N., および Enrietto, P.J. (1994) *Semin. Cancer Biol.* 5:103112)。

30

40

【0021】

或るDNA結合モチーフであるARID(ATリッチ相互作用ドメイン)は、或る進化的に保存されたファミリーのタンパク質を特徴づける。約100残基のARID配列は、細胞成長、分化、および組織特異的遺伝子発現の調節への関与が強く示唆される一連のタンパク質に存在する。ARIDタンパク質の例には、Bright(B細胞特異的遺伝子発現の調節因子)、dead ringer(発達に関与)、およびMRF2(サイトメガロウイルスエンハンサーからの発現を抑制)を含む。(Dallas, P.B. 他 (2000) *Mol. Cell Biol.* 20:31373146)。

【0022】

ELM2(Egl27およびMTA1 相同性2)ドメインは、腫瘍転移関連タンパク質MTA1、およびタンパク質ER1に見られる。線虫(*Caenorhabditis elegans*) 遺伝子egl 27は、胚パターン形

50

成に必要である。ヒトMTA1遺伝子は転移性癌で発現が上昇し、この遺伝子は、ヒストン脱アセチル酵素活性と、ヌクレオソームのリモデリング活性とを持つ或るタンパク質複合体の成分である (Solaris, F. 他 (1999) *Development* 126:2483-2494)。ELM2ドメインは通常、myb様DNA結合ドメインのN末端に見られる。ELM2はまた、或るARID DNAに会合しているのが見られる。

【0023】

殆どの転写因子には特徴的なDNA結合モチーフがあり、上記モチーフ群の種々の変異体と新規のモチーフ群とが、これまでに、また現在も特徴付けられている (Faisst, S. および S. Meyer (1992) *Nucl. Acids Res.* 20:3-26)。

クロマチン結合タンパク質

核内で、DNAはクロマチンに詰められる。クロマチンはコンパクトな組織であり、DNAの、転写因子へのアクセス可能性を制限し、遺伝子調節において主要な役割を果たす (Lewin, 前出 409-410ページ)。クロマチンのコンパクトな構造は、クロマチン結合タンパク質によって決定され影響される。このようなタンパク質としては、ヒストン、高移動度グループ (HMG) タンパク質、およびクロモドメイン蛋白質がある。5種のクラスのヒストン、H1、H2A、H2B、H3およびH4があり、全ては高度に塩基性の低分子量タンパク質である。クロマチンの基盤ユニットであるヌクレオソームには、200塩基対のDNAと、H2A、H2B、H3およびH4それぞれ2つずつのコピーとがある。H1は、隣接ヌクレオソーム群を連結する。HMGタンパク質は、低分子量の非ヒストン蛋白質であり、DNAを巻き戻すことと、一本鎖DNAを安定させることにおいて役割を果たすと思われる。クロモドメイン蛋白質は、高度にコンパクト化されたヘテロクロマチンの形成において重要な役割を果たし、ヘテロクロマチンは転写的にはサイレント (不活動) である。

遺伝子調節に関連する疾患および障害

ヒトにおける腫瘍疾患の多くは、不適切な遺伝子発現によって起こり得る。悪性の細胞成長は、腫瘍促進遺伝子の過剰発現、又は腫瘍抑制遺伝子の不十分な発現によって起こり得る (Cleary, M. L. (1992) *Cancer Surv.* 15:89-104)。Znフィンガー型の転写調節因子WT1は、ウィルムス腫瘍を持つ子供では不活性化されている腫瘍抑制タンパク質である。大細胞リンパ腫で重要な役割を果たすオンコジーンbcl-6もZnフィンガータンパク質である (Papavassiliou, A.G. (1995) *N. Engl. J. Med.* 332:45-47)。染色体転座によっても、或る遺伝子のコード配列を別の無関係遺伝子の調節領域と融合させる、キメラ座 (chimeric loci) が生成され得る。このような編成は、おそらく不適切な遺伝子転写を引き起こし、悪性腫瘍を引き起こす可能性もある。例えばパーキット型リンパ腫では、転写因子Mycが免疫グロブリン重鎖遺伝子座に転座され、Myc発現を大きく増強し、急速な細胞成長を生じて白血病をもたらす (Latchman, D. S. (1996) *N. Engl. J. Med.* 334:28-33)。

【0024】

更に、免疫系は、細胞防御機構の漸進的な選択、増幅、及び動員を調整するイベントのカスケードを活性化することにより、感染や外傷に応答する。遺伝子の活性化及び抑制の複雑かつバランスの取れたプログラムが、このプロセスに関与している。しかしながら、遺伝子発現の不適切或いは不十分な調節によって起こる免疫系の過度の活性化は、組織や器官に相当な損傷を与え得る。この損傷は、関節炎、アレルギー、心臓発作、卒中、及び感染に関連する免疫応答についての文献に詳しく記載されている (Isselbacher 他 *Harrison's Principles of Internal Medicine*, 13/e, McGraw Hill, Inc. および Teton Data Systems Software, 1996)。自己免疫性多腺性内分泌不全症カンジダ症外胚葉性ジストロフィ (APECED) の原因遺伝子が最近単離され、2つのPHDタイプZnフィンガーモチーフを持つ或るタンパク質をコードすることが見出された (Bjorses, P. 他 (1998) *Hum. Mol. Genet.* 7:1547-1553)。

【0025】

更に、多細胞生物の産生は発生の適当な段階における細胞分化の誘導及び調整に基づいている。このプロセスの中心は、体中の細胞及び組織に固有性を与える示差的な遺伝子発現である。発生中に遺伝子発現の調節に失敗すると発達障害が起こり得る。Znフィンガー型

10

20

30

40

50

転写調節因子の突然変異に起因するヒトの発達障害には以下が含まれる。泌尿生殖器の発生異常 (WT1に関連) ; Greig頭蓋・多合指趾症候群 (cephalopolysyndactyly)、Pallister-Hall症候群、および軸後性多指症 (postaxial polydactyly) タイプA (GLI3に関連) ; 並びに、肛門、腎臓、手足および耳の異常を特徴とするTownes-Brocks症候群 (SALL1関連) である (Engelkamp, D. および van Heyningen, V. (1996) *Curr. Opin. Genet. Dev.* 6:334-342; Kohlhase, J. 他 (1999) *Am. J. Hum. Genet.* 64:435-445)。

【0026】

核酸の合成

ポリメラーゼ

DNAとRNAとの複製は、細胞の複製と機能とにとって死活的なプロセスである。DNAとRNAとの複製を仲介する酵素がそれぞれDNAポリメラーゼおよびRNAポリメラーゼであり、この仲介は「鋳型取り (templating)」プロセスで行う。このプロセスでは、DNAまたはRNAストランドのヌクレオチド配列が、相補的塩基対合によってDNAまたはRNAどちらかの相補的な核酸配列にコピーされる。しかし、この2つのプロセスには、根本的な相違がある。

【0027】

DNAポリメラーゼが触媒する反応は、或るポリヌクレオチド鎖 (プライマー鎖) の3'-OH末端へのデオキシリボヌクレオチドの逐次付加であり、このプライマー鎖は第二の (鋳型) ストランドと対合される。新たなDNAストランドは、したがって5'から3'方向へ成長する (Alberts, B. 他 (1994) *The Molecular Biology of the Cell*, Garland Publishing Inc., New York, NY, 251-254ページ)。この重合反応の基質は対応するデオキシヌクレオチド三リン酸であり、ポリメラーゼに認識されるためには、鋳型ストランド上の正しいヌクレオチドと塩基対合する必要がある。DNAは二重らせんとして存在するので、二本鎖のそれぞれが、新たな相補的ストランドの形成のための鋳型となり得る。したがって、分裂中の細胞の2個の娘細胞のそれぞれが、1本は古く1本は新しいストランドを持つ新たなDNA二重らせんを受け継ぐ。したがって、DNAは、DNAポリメラーゼによって「半保存的に」複製されると言われる。新たなDNAの合成に加え、DNAポリメラーゼはまた、損傷したDNAの修復にも関わる。これは、下記の「リガーゼ」で述べる。

【0028】

DNAポリメラーゼと対照的に、RNAポリメラーゼは1本のDNA鋳型ストランドを用いてDNAをRNAに「転写」し、これにはリボヌクレオチド三リン酸を基質として用いる。DNA重合と同様、RNA重合は、5'から3'方向に、リボヌクレオシドリン酸の、成長するRNA鎖の3'-OH末端への付加によって進む。DNA転写はメッセンジャーRNA (mRNA)を産生し、mRNAはタンパク質合成のための情報を運ぶ。DNA転写はまた、転移RNA、リボソームRNAや、構造的または触媒的な機能を持つ他のRNAをも産生する。真核細胞では、3種の別々のRNAポリメラーゼが、3タイプのRNAを合成する (Alberts 他、前出 367-368ページ)。RNAポリメラーゼIは大きなリボソームRNAを作り、RNAポリメラーゼIIはタンパク質へと翻訳されるmRNAを作り、RNAポリメラーゼIIIは、種々の小さな安定したRNA (5SリボソームRNAや転移RNA (tRNA) 等) を作る。どの場合も、RNA合成の開始は、RNAポリメラーゼの、DNA上のプロモーター領域への結合によってなされ、合成の開始は、プロモーター内の開始部位で起こる。合成の完了は、DNA内の停止 (終了) シグナルでなされ、ここでポリメラーゼと完成したRNA鎖とが解放される。

リガーゼ

DNA修復のプロセスによって、偶発的な塩基変化 (原因は例えば酸化損傷、加水分解的攻撃、DNAの無秩序なメチル化など) が、DNAの複製または転写が起きる前に修正される。DNA修復プロセスの効率が良いので、偶発的な塩基変化の1000分の1未満のみが突然変異を起こす (Alberts 他、前出 245-249ページ)。殆どのタイプのDNA修復に共通の3つのステップは、(1) 損傷または変容した塩基またはヌクレオチドの、DNAヌクレアーゼによる切除。(2) 切除されたヌクレオチドが残したギャップへの、正しいヌクレオチドの挿入。これはDNAポリメラーゼが、相補的ストランドを鋳型として用いて行う。(3) 挿入されたヌクレオチド (群) と既存DNAストランドとの間に残った切れ目の、DNAリガーゼによ

る埋め合わせ (sealing)、である。最後の反応において、DNAリガーゼはATP加水分解からのエネルギーを用い、切れ目の入ったホスホジエステル結合の5'末端を活性化した後、DNAストランドの3'-OHとの新たな結合を形成する。ブルーム症候群は遺伝性のヒト疾患であるが、この患者はこのDNAライゲーションが部分的に欠損しているため、癌の発症の増加を示す (Alberts 他、前出 247ページ)。

ヌクレアーゼ

ヌクレアーゼ酵素には、DNAを加水分解するDNアーゼと、RNAを加水分解するRNアーゼとがある。2種のヌクレアーゼは、核酸代謝において別々の目的で働く。ヌクレアーゼは、隣接するヌクレオチド間のホスホジエステル結合を加水分解する。エンドヌクレアーゼは内部の位置で加水分解し、エキソヌクレアーゼは3'または5'末端のヌクレオチド位置で加水分解する。例えばDNAポリメラーゼにおける或るDNAエキソヌクレアーゼ活性は、ポリメラーゼによって成長するDNAストランドの3'-OH末端に付着されたヌクレオチドの内、不適切に対合したヌクレオチドを除去するよう働くので、「校正」機能を提供する。上記したように、DNAエンドヌクレアーゼ活性は、DNA修復プロセスの切除ステップに関わる。

10

20

30

40

50

【0029】

RNアーゼ (RNase) もまた、種々の機能を提供する。例えばRNase Pはリボ核タンパク質酵素であり、tRNA前駆体群の5'末端の切断を、それらの成熟プロセスの一部として行う。RNase Hは、或るRNA/DNAハイブリッドのRNAストランドを消化する。このようなハイブリッドはレトロウイルスが侵入した細胞に生じ、RNase Hは、レトロウイルス複製サイクルにおける重要な酵素である。膵臓RNアーゼは膵臓によって腸内に分泌され、摂取した食物中のRNAを加水分解する。血清と細胞抽出物とにおけるRNアーゼ活性の上昇は、種々の癌や感染疾患で見られる (Schein, C.H. (1997) Nat. Biotechnol. 15:529-536)。RNアーゼ活性の調節が、腫瘍血管新生、アレルギー反応、ウイルスの感染および複製、並びに真菌感染を制御する手段として研究されている。

【0030】

核酸の修飾

メチラーゼ

特定ヌクレオチドのメチル化はDNA、RNA双方で起こり、この2種の高分子において別々の機能を提供する。DNAでシトシン残基のメチル化は5-メチルシトシンを形成し、これは、DNA二重らせん内で互いに塩基対合したCG配列群内で特異的に起こる。メチル化のこのパターンは、DNA複製時に世代間で受け渡される。これを行う酵素は「維持メチラーゼ」と呼ばれ、既にメチル化されたCG配列と塩基対合したCG配列上で優先的に作用する。このようなメチル化は、活性遺伝子と非活性遺伝子との区別を、遺伝子を「オンにする (turn on)」調節タンパク質の結合を防ぐ一方、遺伝子を不活化するタンパク質の結合を許すことにより行うようである (Alberts 他、前出 448-451ページ)。RNA代謝では、「tRNAメチラーゼ」がtRNAにおける幾つかのヌクレオチド修飾の1つを起こす。この修飾は、この分子のコンフォメーションと塩基対合とに影響し、適切なmRNAコドンの、特異的tRNA群による認識を促進する。主なメチル化パターンとしてはグアニン残基のジメチル化があり、これはN,N-ジメチルグアニンを形成する。

ヘリカーゼおよび一本鎖結合タンパク質

ヘリカーゼ酵素は、DNA、RNA双方の二重らせん構造を不安定化し、巻き戻す。DNA複製は2本のストランドでほぼ同時に起こるので、2本のストランドは先ず分離して複製「フォーク」を産生し、DNAポリメラーゼが作用できるようにする必要がある。2種の複製タンパク質がこのプロセスに寄与する。DNAヘリカーゼと、一本鎖結合タンパク質とである。DNAヘリカーゼは、ATPを加水分解し、そのエネルギーを用いてDNAストランドを分離させる。一本鎖結合タンパク質 (SSB) が次に、露出したDNAストランドに結合するが、塩基を覆いはせず、DNAポリメラーゼによる鋳型取りのためにそれらを一時的に安定化する (Alberts 他、前出 255-256ページ)。

【0031】

RNAヘリカーゼは、RNAのコンフォメーションと2次構造とを変容させ調節する。DNAヘリ

カーゼと同様、RNAヘリカーゼは、ATP加水分解に由来するエネルギーを利用し、RNA二重鎖を不安定化し巻き戻す。RNAヘリカーゼの最も良く特徴付けられ偏在するファミリーはDEADボックスファミリーであり、その名は、保存されたBタイプATP結合モチーフに由来する。このモチーフは、このファミリーのタンパク質にとって特徴的である。40種を超すDEADボックスヘリカーゼが、細菌、昆虫、酵母、両生類、哺乳類、および植物に及ぶ多様な生物において同定されている。DEADボックスヘリカーゼは、多様なプロセスにおいて機能する。そのプロセスとしては例えば翻訳開始、スプライシング、リボソームの集合、および、RNAの編集、輸送、安定性がある。これらのRNAヘリカーゼの例としては、酵母Drs1タンパク質（リボソームRNAプロセシングに関与）、酵母TIF1およびTIF2並びに哺乳類eIF-4A（これらはRNA翻訳の開始に必須である）、およびヒトp68抗原（細胞の成長および分化を調節）がある(Ripmaster, T.L. 他 (1992) Proc. Natl. Acad. Sci. USA 89:11131-11135; Chang, T.-H. 他 (1990) Proc. Natl. Acad. Sci. USA 87:1571-1575)。これらのRNAヘリカーゼは、強い配列相同性を、約420アミノ酸のストレッチにわたり示す。これらの保存された配列としては、或るATP結合タンパク質のAモチーフのコンセンサス配列、「DEADボックス」配列（ATPアーゼ活性に関与）、配列SAT（実際のヘリカーゼ巻き戻し：actual helicase unwinding領域に関与）、および、RNA結合とATP加水分解とに必要なオクタペプチドコンセンサス配列がある(Pause, A. 他 (1993) Mol. Cell Biol. 13:6789-6798)。これらの保存された領域以外の差異は、個々のタンパク質の機能上の役割における差異を反映すると思われる(Chang, T.H. 他 (1990) Proc. Natl. Acad. Sci. USA 87:1571-1575)。

10

【0032】

20

数種のDEADボックスヘリカーゼは、精子形成と胚形成とにおいて、組織特異的およびステージ特異的な役割を果たす。DEADボックス1タンパク質(DDX1)の過剰発現は、神経芽腫(Nb)と網膜芽腫(Rb)との進行において役割を果たすと思われる(Godbout, R. 他 (1998) J. Biol. Chem. 273:21161-21168)。これらの観察の示唆するところでは、DDX1は、腫瘍形成の促進または増強を、癌細胞内のRNAの正常な二次構造と発現レベルとを改変することによって行う。別のDEADボックスヘリカーゼ群は、腫瘍形成に直接または間接的に関与すると思われる(前出Godboutで考察されている)。例えばネズミp68は紫外線で誘発した腫瘍において突然変異し、ヒトDDX6は、B細胞リンパ腫に関連する或る染色体易切断点(chromosomal breakpoint)に位置する。同様に、DDX10とNUP98(1種の核膜孔タンパク質:nucleoporin)とからなるキメラ蛋白質は、数種の骨髄性悪性腫瘍の病原に関与するようである。

30

トポイソメラーゼ

DNAストランドを複製前に分離する他に、2本のストランドは、DNAヘリカーゼによる分離の前に、互いから「巻き戻す」必要がある。この機能を行うタンパク質は、DNAトポイソメラーゼとして知られる。DNAトポイソメラーゼは、可逆的ヌクレアーゼとして効果的に作用してDNAストランド内のホスホジエステラーゼ結合を加水分解し、二本鎖を互いに対して自由に回転できるようにしてヘリックスを巻き戻した後、二本鎖の間の元のホスホジエステル結合を再結合させる。トポイソメラーゼは、転写、複製、クロマチン形成、組換え、および染色体分離、によって生じたDNAのトポロジー的再編成を担う必須酵素である。DNAへの高次らせんコイルの導入は、前進する酵素(processive enzyme)(例えばRNAポリメラーゼ)の通過によって、または、ヘリカーゼによるDNAストランドの複製前の分離によってなされる。DNAの合成、貯蔵、修復のプロセス中に、結び目(knotting)と連鎖(concatenation)とが生じ得る。全てのトポイソメラーゼの働きは、DNAのリボースリン酸主鎖内のホスホジエステル結合を壊して行う。この酵素にある或る触媒チロシン残基が、切れやすいホスホジエステル結合に求核攻撃することによって反応中間体を生じる。この中間体では、共有結合が、この酵素と、壊れたストランドの1端との間に形成される。或るチロシンDNAホスホジエステラーゼは、偶発するデッドエンドトポイソメラーゼI DNA中間体の中のこの結合を加水分解することでDNAの修復に機能する(Pouliot, J. J. 他(1999) Science 286:552-555)。

40

【0033】

50

2タイプのDNAトポイソメラーゼ、タイプIとタイプIIとがある。タイプIトポイソメラーゼは単量体として働いてDNAの一本鎖に切れ目を作る。一方、タイプIIトポイソメラーゼはホモ二量体として働き、二本鎖の双方を切断する。DNAトポイソメラーゼIは、DNAヘリックス内に一本鎖の切れ目を生じることによってヘリックスの二本鎖が、反対のストランドの中の、残るホスホジエステル結合を軸に回転できるようにする。DNAトポイソメラーゼIIはDNAヘリックスの二本鎖の双方に一時的な切れ目を作り、ここで二本の二重らせんが互いに交差する。このトポイソメラーゼIIは、2つの連環した環状DNAを、効果的に分離できる(Alberts 他、前出 260-262ページ)。タイプIIトポイソメラーゼは、真核生物では増殖する細胞(癌細胞など)にほとんど限られている。このためトポイソメラーゼIIは、抗癌剤の標的である。トポイソメラーゼIIは、多剤耐性(MDR)に関係すると思われる

10

20

30

40

50

【0034】

トポイソメラーゼIファミリーには、トポイソメラーゼI(TopoI)とトポイソメラーゼIII(TopoIII)とを含む。ヒトのトポイソメラーゼIの結晶構造が示唆するところでは、無傷DNAストランドを軸とした回転を、この酵素が部分的に制御する。この「制御された回転(controlled rotation)」モデルでは、タンパク質-DNA相互作用がこの回転を制限する。これはDNA内のねじれひずみ(torsional strain)が駆動する(Stewart, L. 他(1998) Science 379:1534-1541)。構造的にはTopoIは、その活性部位領域内の触媒チロシン残基と、多数の他の保存された残基で認識できる。TopoIは、転写時に機能すると思われる。2種のTopoIIIがヒトでは既知であり、これらは原核生物トポイソメラーゼIに相同で、保存されたチロシンと、このファミリーに特異的な活性部位シグネチャを持つ。TopoIIIは、減数分裂組換え(meiotic recombination)において役割を果たすことが示唆されている。マウスのある種のTopoIIIは精巣組織で高度に発現され、その発現の増加は、太糸期(pachytene)における細胞数の増加に伴う(Seki, T. 他(1998) J. Biol. Chem. 273:28553-28556)。

【0035】

トポイソメラーゼIIファミリーには、2種のイソ酵素(II₁ およびII₂)を含み、2種は別々の遺伝子がコードする。TopoIIは二本鎖DNAの切断を、再現性のある非無作為の方法で、優先的に或るATリッチ領域で行うが、切断部位選択性の原理は未知である。構造上はTopoIIには4つのドメインがある。その最初の2つは構造的に類似し、おそらく真核生物TopoIの類似ドメインとは遠い相同性を持つ。第二ドメインには触媒チロシンと、高度に保存されたペントペプチドを持つ。II₁ イソフォームは、染色体分離時にDNAを解くこと(unlinking)に関係するようである。II₂ を発現するが、II₁ を発現しない細胞株は、II₁ は細胞プロセスに非必須であることを示唆するが、II₂ ノックアウトマウスは、神経発生における不全のため周産期に死亡した。主な異常が主に後期発生イベント(神経形成)時に起きたことは、II₂ が有糸分裂ではなくDNA修復時に必要であることを示唆する(Yang, X. 他(2000) Science 287:131-134)。

【0036】

トポイソメラーゼは多数の病態に関わるとされており、トポイソメラーゼ毒は、数種のヒト悪性腫瘍癌に有効な抗腫瘍薬であることが示されている。TopoIはファンコニ貧血(Fanconi's anemia)では誤った場所に局在する。また、この疾患に見られる染色体切断に関わるとされる(Wunder, E. (1984) Hum. Genet. 68:276281)。毛細血管拡張性運動失調(A-T)細胞での短縮されたTopoIIIの過剰発現は、部分的にA-T表現型を抑制するが、これはおそらく優性阻害機序(dominant negative mechanism)による。これが示唆するところでは、TopoIIIはA-Tでは調節解除される(Fritz, E. 他(1997) Proc. Natl. Acad. Sci. USA 94:4538-4542)。TopoIIIはまた、ブルーム症候群(Bloom's Syndrome)遺伝子産物と相互作用する。また、腫瘍抑制因子としての役割を持つことを示唆されている(Wu, L. 他(2000) J. Biol. Chem. 275:9636-9644)。異常なTopoII活性は、しばしば癌や、増加した癌リスクを伴う。大きく低下したTopoII活性が、全てではないが数種のA-T細胞

株に見られている (Mohamed, R. 他 (1987) *Biochem. Biophys. Res. Commun.* 149:233238)。一方、TopoIIは、DNAの分断を、A-T遺伝子(ATM)の領域で行い得る。ATMは、全てのDNA損傷応答性細胞周期チェックポイントを制御する(Kaufmann, W.K. (1998) *Proc. Soc. Exp. Biol. Med.* 217:327-334)。トポイソメラーゼがDNAを分断する能力は、抗腫瘍薬の基礎として用いられている。トポイソメラーゼ毒は、デッドエンド共有結合DNA-酵素複合体の細胞内の数を増すことによって作用し、最終的には細胞死経路の引き金となる(Fortune, J.M. および N. Osheroff (2000) *Prog. Nucleic Acid Res. Mol. Biol.* 64:221-253; Guichard, S.M. および M.K. Danks (1999) *Curr. Opin. Oncol.* 11:482-489)。TopoIIに対する抗体は全身性強皮症(systemic sclerosis)患者の血清中に見られ、この抗体のレベルは、この疾患に關与する肺のマーカーとして利用できそうである(Diot, E. 他 (1999) *Chest* 116:715-720)。最後に、ヒトのTopoIIのDNA結合領域は、遺伝子治療用のDNA送達伝播体として用いられている(Chen, T.Y. 他 (2000) *Appl. Microbiol. Biotechnol.* 53:558-567)。

リコンビナーゼ

遺伝子組換えは、或る生物のゲノム内でのDNA配列の再編成プロセスであり、環境の変化に応じたその生物の遺伝子変異を提供する。DNA組換えによって、或る個体のゲノム内にある遺伝子の特定の組み合わせにおける変異が可能になり、これらの遺伝子の発現の時期とレベルとにおける変異も可能である(前出Alberts他 263-273ページを参照)。2種の広範なクラスの遺伝子組換え(普遍的組換えと部位特異的組換え)が広く認識されている。普遍的組換えでは、任意の相同的な1対のDNA配列間の遺伝子交換を伴う。これらの配列の位置は通常、同じ染色体の2つのコピー上にある。このプロセスを補助する酵素であるリコンビナーゼは、DNA二重鎖の一方のストランドに幾分無作為的に「ニック(切れ目)を入れ」、他の二重鎖にある相補的ストランドとの交換を可能にする。このプロセスは、通常は染色体内の遺伝子の編成を変えない。部位特異的組換えでは、リコンビナーゼは組み換える分子の一方または双方にある特定のヌクレオチド配列を認識する。塩基対合はこの形の組み換えには關与しないので、組み換える分子の間のDNA相同性は不要である。普遍的組換えとは異なり、部位特異的組み換えは、染色体内のヌクレオチド配列群の相対位置を改変できる。

RNA代謝

リボ核酸(RNA)は直鎖の一本鎖重合体である(4種のヌクレオチドであるATP、CTP、UTP、およびGTPからなる)。殆どの生物では、RNAの転写は、デオキシリボ核酸(DNA)のコピーとしてなされ、DNAがその生物の遺伝子材料である。レトロウイルスでは、DNAでなくRNAが遺伝子材料として働く。遺伝子材料のRNAコピーは、タンパク質をコードするか、もしくは生物内で種々の構造的、触媒的、または調節的な役割を果たす。RNAの分類は、その細胞内局在と機能とに基づく。メッセンジャーRNA(mRNA)は、ポリペプチドをコードする。リボソームRNA(rRNA)は、リボソーム蛋白質と組み合わされてリボソームとなる。リボソームは細胞質粒子であり、mRNAをポリペプチドへ翻訳する。転移RNA(tRNA)はサイトゾルのアダプター分子であり、mRNA翻訳における機能はmRNAコドンと、そのコドンに適合するアミノ酸との双方の認識である。ヘテロ核RNA(hnRNA)としては、mRNA前駆体と、他の種々のサイズの核内RNAとを含む。核内低分子RNA(snRNA)は核内スプライセオソーム複合体の一部である。この複合体は、mRNA前駆体内に介在する非コード配列(イントロン)群を除去し、エキソン群を再結合する。

【0037】

タンパク質は、DNAからのRNAの転写や、RNAプロセッシング、および、タンパク質へのmRNAの翻訳の際に、RNAと会合する。タンパク質はまた、RNAが構造的、触媒的、または調節的な目的で利用される際に、RNAと会合する。

RNAプロセッシング

リボソームRNA(rRNA)は、リボソーム蛋白質と組み合わされてリボソームとなる。リボソームは細胞質粒子であり、メッセンジャーRNA(mRNA)をポリペプチドへ翻訳する。真核生物リボソームは60S(大)サブユニットと40S(小)サブユニットとからなり、両ユニット

が共になって80Sリボソームを形成する。18S、28S、5S、および5.8SのrRNAに加え、リボソームには生物の種類によって50から80種を超す別々のリボソーム蛋白質が含まれる。リボソーム蛋白質の分類は、それらが属するサブユニットによる(すなわち、大きな60Sサブユニットに関連するならL、小さい40SサブユニットならS)。大腸菌リボソームはもっとも詳しく研究されており、50種の蛋白質を持ち、その多くは全ての生物で保存されている。9種のリボソーム蛋白質の構造の解明が3.0D未満の分解能でなされており(すなわちS5、S6、S17、L1、L6、L9、L12、L14、L30)、共通のモチーフ群(例えばb-a- b蛋白質フォールド群)と、b-ストランドの間に位置する酸性および塩基性のRNA結合モチーフ群とを明らかにした。殆どのリボソーム蛋白質はrRNAと直接に接触すると思われる(概説はLiljas, A. および Garber, M. (1995) *Curr. Opin. Struct. Biol.* 5:721-727; また、Woodson, S. A. および Leontis, N.B. (1998) *Curr. Opin. Struct. Biol.* 8:294-300; Ramakrishnan, V. および White, S.W. (1998) *Trends Biochem. Sci.* 23:208-212を参照)。

【0038】

種々の蛋白質が、核内で転写されたRNAのプロセシングに必要である。mRNAプロセシングの前段階としては、5'末端へのメチルグアノシンでのキャッピング、3'末端のポリアダニル化、および、イントロンを除去するためのスプライシングを含む。スプライセオソーム複合体は、5種の低分子核内リボ核タンパク質粒子(snRNP)である、U1、U2、U4、U5、およびU6からなる。それぞれのsnRNP内には、1つの種のsnRNAと、約10種のタンパク質とを含む。数種のsnRNPのRNA成分は、イントロンのコンセンサス配列を認識し塩基対合する。タンパク質成分は、スプライセオソーム構築と、スプライシング反応とを仲介する。snRNPタンパク質に対する抗体は、全身性紅斑性狼瘡患者の血中に見られる(Stryer, L. (1995) *Biochemistry* W.H. Freeman and Company, New York NY, 863ページ)。

【0039】

不均質核内リボ核タンパク質(hnRNP)の役割は、スプライシングや、成熟RNAの細胞質への輸送、および、mRNA翻訳にあることが同定されている(Biamonti, G. 他 (1998) *Clin. Exp. Rheumatol.* 16:317-326)。hnRNPの例としては、酵母タンパク質群である、Hrp1p (RNAの切断と3'末端でのポリアダニル化に関与)、Cbp80p (RNAの5'末端のキャッピングに関与)、および、Npl3p (哺乳類hnRNP A1の相同体であり、mRNAの核外輸送に関与)がある(Shen, E.C. 他 (1998) *Genes Dev.* 12:679-691)。hnRNPは、リウマチ性疾患における自己免疫応答の重要な標的であることが示されている(前出Biamonti)。

【0040】

多くのsnRNPタンパク質およびhnRNPタンパク質は、或るRNA認識モチーフ(RRM)によって特徴付けられる(概説はBirney, E. 他 (1993) *Nucleic Acids Res.* 21:5803-5816)。RRMは、約80アミノ酸の長さであり、 γ / δ サンドイッチの形に編成された4本の α ストランドと2本の β ヘリックスとを形成する。RRMは、コアRNP-1オクタペプチドモチーフと、周囲の保存された配列群とを含む。snRNPタンパク質に加え、上記モチーフ群を持つRNA結合タンパク質の例には、新生RNAを安定させるヘテロ核リボ核タンパク質と、選択的スプライシングを調節する因子群とを含む。選択的スプライシング因子には、発生的に調節されるタンパク質が含まれ、その特異的な例が下等真核生物(例えばキイロシヨウジョウバエや線虫 *Caenorhabditis elegans*)において同定されている。これらのタンパク質は、それぞれ発生プロセス(例えばパターン形成や性決定)において重要な役割を果たす(例えばHodgkin, J. 他 (1994) *Development* 120:3681-3689を参照)。

RNAの安定性および分解

RNAヘリカーゼは、RNAのコンフォメーションと二次構造との改変と調節とをATP加水分解に由来するエネルギーを利用して行い、RNA二重鎖を不安定化し巻き戻す。RNAヘリカーゼの最も良く特徴付けられ偏在するファミリーはDEADボックスファミリーであり、その名は、保存されたBタイプATP結合モチーフに由来する。このモチーフは、このファミリーのタンパク質にとって特徴的である。40種を超すDEADボックスヘリカーゼが、細菌、昆虫、酵母、両生類、哺乳類、および植物に及ぶ多様な生物において同定されている。DEADボックスヘリカーゼは、多様なプロセスにおいて機能する。そのプロセスとしては例えば翻訳開始、ス

プライシング、リボソームの集合、および、RNAの編集、輸送、安定性がある。数種のDEADボックスヘリカーゼは、精子形成と胚形成とにおいて、組織特異的およびステージ特異的な役割を果たす。全てのDEADボックスヘリカーゼ内には、約420アミノ酸にわたり広がる、数種の保存された配列モチーフを含む。これらのモチーフとしては、AタイプATP結合モチーフ、DEADボックス/BタイプATP結合モチーフ、セリン/アルギニン/トレオニンのトリペプチド（機能は未知）、および、C末端グリシンリッチモチーフ（可能性のある役割は基質結合および巻き戻し）を含む。また、多岐にわたるDEADボックスヘリカーゼ配列群のアラインメントが示しているところでは37アミノ酸残基がこれらの配列すべてで同一であり、これらの残基の保存がヘリカーゼ機能に重要であることを示唆する（概説はLinder, P. 他 (1989) *Nature* 337:121-122）。

10

【0041】

DEADボックス1タンパク質(DDX1)の過剰発現は、神経芽腫(Nb)と網膜芽腫(Rb)との進行において役割を果たすと思われる。これらの観察の示唆するところでは、DDX1は、腫瘍進行の促進または増強を、癌細胞内のRNAの正常な二次構造と発現レベルとを改変することによって行う。別のDEADボックスヘリカーゼ群は、紫外線で誘発した腫瘍や、B細胞リンパ腫、および骨髄性悪性疾患に、直接または間接的に関与すると思われる（概説はGodbout, R. 他 (1998) *J. Biol. Chem.* 273:21161-21168）。

【0042】

リボヌクレアーゼ(RNアーゼ)はRNA鎖内のホスホジエステル結合の加水分解を、したがってRNAの切断を触媒する。例えばRNase Pはリボ核タンパク質酵素であり、tRNA前駆体群の5'末端の切断を、それらの成熟プロセスの一部として行う。RNase Hは、或るRNA/DNAハイブリッドのRNAストランドを消化する。このようなハイブリッドはレトロウイルスが侵入した細胞に生じ、RNase Hは、レトロウイルス複製サイクルにおける重要な酵素である。RNase Hドメインは、しばしば逆転写に関連するドメインとして見られる。血清と細胞抽出物とにおけるRNアーゼ活性の上昇は、種々の癌や感染疾患で見られる(Schein, C.H. (1997) *Nat. Biotechnol.* 15:529-536)。RNアーゼ活性の調節が、腫瘍血管新生、アレルギー反応、ウイルスの感染および複製、並びに真菌感染を制御する手段として研究されている。

20

【0043】

リボソーム蛋白質は、翻訳を調節するために、翻訳後修飾を受けたり、他のリボソーム関連タンパク質と相互作用することがある。例えば高度に相同な40Sリボソーム蛋白質S6キナーゼであるS6K1とS6K2とは細胞成長の調節において主要な役割を果たすが、これはタンパク質合成装置(例えばリボソーム蛋白質)を作り上げる翻訳成分の生合成を制御することにより行う。S6K1の場合、少なくとも8つのリン酸化部位が、キナーゼ活性化を階層的な手法で仲介すると思われる(Dufner および Thomas (1999) *Exp. Cell. Res.* 253:100109)。数種のリボソーム蛋白質(L1を含む)はまた翻訳抑制因子として機能し、これはリボソーム蛋白質をコードする多シストロン性mRNAと結合して行う(概説は前出Liljas, A. および前出Garber, M.)。

30

【0044】

最近の証拠が示唆するところでは、多数のリボソーム蛋白質が、蛋白質生合成へのこれらの関与とは独立した二次機能を持つ。これらの蛋白質は、細胞増殖の調節因子として、また数例では細胞死の誘発因子として機能する。例えばヒトのリボソーム蛋白質L13aの発現は、アポトーシスを誘発することが示されている。これは、細胞周期のG2/M期に細胞成長を停止させて行う。L13aの発現の抑制は標的細胞内にアポトーシスを誘発し、これが示唆するところでは、この蛋白質は、適量では、細胞生存に必要である。同様の結果が酵母でも得られている。ここでは、L13aの酵母相同体である、rp22とrp23との不活化によって、重度の成長遅滞と死とを招いた。非常に近縁のリボソーム蛋白質であるL7は、G1期の細胞を停止させ、アポトーシスをも誘発する。したがって、或るサブセットのリボソーム蛋白質は、細胞周期チェックポイントとして機能すると思われ、新規ファミリーの細胞増殖調節因子を構成するようである。

40

50

【 0 0 4 5 】

個々のリボソーム蛋白質の、無傷リボソームの表面へのマッピングがなされた。これには3D免疫低温電子顕微鏡 (immunocryoelectronmicroscopy) を用い、これによって特定のリボソーム蛋白質群に対して産生された抗体群が可視化された。L1、L7、およびL12のマッピングに向けて進歩がなされているが、無傷リボソームの構造の解明は、わずかに20~25Dの分解能であり、別々の粗構造 (crude structures) の間には矛盾がある (Frank, J. (1997) *Curr. Opin. Struct. Biol.* 7:266272)。

【 0 0 4 6 】

3種の別々の部位が、リボソーム上で同定されている。アミノアシルtRNA受容部位(A部位)は、チャージした (charged) tRNAを受け取る (例外は開始tRNAである)。ペプチジルtRNA A部位(P部位)は新生ポリペプチドと結合し、A部位からのアミノ酸が、この伸長する鎖に付加される。脱アシル化したtRNA群は、脱出部位(E部位)内に結合した後、リボソームから解放される。リボソームの構造の概説は、Stryer, L. (1995) *Biochemistry* W.H. Freeman and Company, New York NY 888-9081ページ; Lodish, H. 他 (1995) *Molecular Cell Biology* Scientific American Books, New York NY 119-138ページ; および Lewin, B (1997) *Genes VI* Oxford University Press, Inc. New York, NYを見られたい。

【 0 0 4 7 】

種々の蛋白質が、核内で転写されたRNAのプロセッシングに必要である。mRNAプロセッシングの前段階としては、5'末端へのメチルグアノシンでのキャッピング、3'末端のポリアデニル化、および、イントロンを除去するためのスプライシングを含む。DNAからの一次RNA転写物は、エキソン、イントロン双方の配列を有する遺伝子の忠実なコピーであり、イントロン配列がRNA転写物から切り出されることが、蛋白質をコードするmRNAを産出するために必要である。mRNAの「スプライシング」は、核内で、大きな多成分リボ核タンパク質複合体 (スプライセオソームとして知られる) の補助を受けてなされる。スプライセオソーム複合体は、5種の低分子核内リボ核タンパク質粒子 (snRNP) である、U1、U2、U4、U5、およびU6からなる。それぞれのsnRNP内には、1つの種のsnRNAと、約10種のタンパク質とを含む。数種のsnRNPのRNA成分は、イントロンのコンセンサス配列を認識し塩基対合する。タンパク質成分は、スプライセオソーム構築と、スプライシング反応とを仲介する。snRNPタンパク質に対する抗体は、全身性紅斑性狼瘡患者の血中に見られる (Stryer, L. (1995) *Biochemistry* W.H. Freeman and Company, New York NY, 863ページ)。

【 0 0 4 8 】

不均質核内リボ核タンパク質 (hnRNP) の役割は、スプライシングや、成熟RNAの細胞質への移出、および、mRNA翻訳にあることが同定されている (Biamonti, G. 他 (1998) *Clin. Exp. Rheumatol.* 16:317-326)。hnRNPの例としては、酵母タンパク質群である、Hrp1p (RNAの切断と3'末端でのポリアデニル化に関与)、Cbp80p (RNAの5'末端のキャッピングに関与)、および、Npl3p (核からのmRNAの移出に関与する哺乳類hnRNP A1の相同体) がある (Shen, E.C. 他 (1998) *Genes Dev.* 12:679-691)。hnRNPは、リウマチ性疾患における自己免疫応答の重要な標的であることが示されている (前出Biamonti)。

【 0 0 4 9 】

多くのsnRNPタンパク質およびhnRNPタンパク質は、或るRNA認識モチーフ (RRM) によって特徴付けられる (概説はBirney, E. 他 (1993) *Nucleic Acids Res.* 21:5803-5816)。RRMは、約80アミノ酸の長さであり、 / サンドイッチの形に編成された4本のストランドと2本のヘリックスとを形成する。RRM内には、コアのRNP-1オクタペプチドモチーフと、周囲の保存された配列群とを含む。snRNPタンパク質に加え、上記モチーフ群を持つRNA結合タンパク質の例には、新生RNAを安定させるヘテロ核リボ核タンパク質と、選択的スプライシングを調節する因子群とを含む。選択的スプライシング因子には発生的に調節されるタンパク質が含まれ、その特異的な例は下等真核生物 (例えばキイロシヨウジョウバエや線虫 *Caenorhabditis elegans*) において同定されている。これらのタンパク質は、それぞれ発生プロセス (例えばパターン形成や性決定) において重要な役割を果たす (例えばHodgkin, J. 他 (1994) *Development* 120:3681-3689を参照)。

10

20

30

40

50

【0050】

殆どの真核生物 mRNA の 3' 末端は、また、ポリアデニル化による転写後修飾を受ける。ポリアデニル化は、2 種の、酵素的に別々のステップを経て進む。(i) 新生 mRNA のヌクレオチド鎖切断が、3' 非翻訳(非コード)領域内のシス作用ポリアデニル化シグナル群でなされ、(ii) ポリ(A)トラクト (poly(A) tract) が、5' mRNA 断片へ付加される。シス作用 RNA 配列群の存在が、両方のステップに必要である。これらの配列としては、5'-AAUAAA-3' (切断部位の 10~30ヌクレオチド上流に位置) と、あまり保存されていない GUリッチ配列エレメントまたは Uリッチ配列エレメント (切断部位の 10~30ヌクレオチド下流に位置) とを含む。切断刺激因子 (CstF)、切断因子 I (CF I)、および切断因子 II (CF II) は切断反応に関わり、一方、切断およびポリ(A)付加特異性決定因子 (CPSF) とポリ(A)ポリメラーゼ (PAP) とは、切断とポリアデニル化との双方に必要である。付加的酵素であるポリ(A)結合タンパク質 II (PAB II) は、ポリ(A)トラクト伸長を促進する (Ruegsegger, U. 他 (1996) J. Biol. Chem. 271:61076113; およびその参考文献)。

10

翻訳

遺伝子コードの正しい翻訳は、適切な転移 RNA (tRNA) との連結を形成する各アミノ酸に依存する。アミノアシル-tRNA 合成酵素 (aaRS) は、全ての生物に在る本質的なタンパク質である。aaRS は、タンパク質の生合成の最初の段階として、アミノ酸の活性化と、その同族 tRNA との正しい付着とを担う。原核生物は少なくとも 20 種の異なるタイプの aaRS を持ちアミノ酸ごとに 1 種類の aaRS を持つが、真核生物は通常、異なるアミノ酸の各々のために細胞質ゾル形態とミトコンドリア形態の 2 つの aaRS を有する。20 種の aaRS 酵素は 2 つの構造的クラスに分け得る。クラス I の酵素は tRNA の 3' 末端で 2' ヒドロキシルにアミノ酸を加え、クラス II の酵素は tRNA の 3' 末端で 3' ヒドロキシルにアミノ酸を加える。各クラスは触媒的ドメインの特有のトポロジーによって特徴づけられる。クラス I 酵素には、ヌクレオチド結合した「ロスマンフォールド (Rossmann fold)」に基づく触媒的ドメインがある。特に、或るコンセンサステトラペプチドのモチーフは、高度に保存的である (Prosite Document PDOC00161, Aminoacyl-transfer RNA synthetases class-I signature)。クラス I 酵素は、アルギニン、システイン、グルタミン酸、グルタミン、イソロイシン、ロイシン、メチオニン、チロシン、トリプトファンおよびバリンに特異的である。クラス II 酵素は、中心的触媒ドメイン (7 本鎖の逆平行 β -シートドメインからなる) と、N 末端および C 末端調節ドメインを持つ。クラス II 酵素は酵素のヘテロ二量体構造あるいはホモ二量体構造に基づき 2 群に分けられ、後者はさらに N 末端および C 末端調節ドメインの構造によって分けられる (Hartlein, M. および Cusack, S. (1995) J. Mol. Evol. 40:519-530)。クラス II 酵素は、アラニン、アスパラギン、アスパラギン酸、グリシン、ヒスチジン、リジン、フェニルアラニン、プロリン、セリンおよびトレオニンに特異的である。

20

30

【0051】

数種の aaRS には編集 (エディティング) 機能もある。例えば IleRS はバリンを誤って活性化し Val-tRNA^{Ile} を形成する場合があるが、この生成物は、誤ってチャージした生成物を破壊する或る加水分解作用によって取り除かれる。この編集活性は、クラス I 酵素の Rossmann fold ドメイン内の長い挿入配列である接続ポリペプチド 1 領域 (connective polypeptide 1 region) (CP1) で見つかる 2 番目の触媒サイト内に位置する (Schimmel, P. 他 (1998) FASEB J. 12:1599-1609)。AaRS はまた tRNA プロセッシングでも役割を果たす。成熟した tRNA は細胞質に移出される前に、核の中でそれぞれのアミノ酸でチャージされることが示されている。またこのアミノ酸との結合は品質管理機構として機能し、確実に tRNA を機能的に保っている可能性がある (Martinis, S.A. 他 (1999) EMBO J. 18:4591-4596)。

40

【0052】

最適条件下でのポリペプチド合成の進行速度は、約 40 アミノ酸残基/秒である。翻訳時の誤った取り込みの比率は 10^{-4} の位であり、主な原因はアミノアシル-tRNA に間違ったアミノ酸でチャージされたことである。不正にチャージした tRNA は細胞毒性を持つ。その理由は、それらが不正アミノ酸残基を、伸長するポリペプチドに組み込むためである。翻訳

50

の速度は、最適な伸長速度と、翻訳忠実度の必要性との折衷で決まると推測される。数学的計算による予測では、 10^{-4} が、生物系におけるタンパク質合成の最大許容エラー率である(概説は前出 Stryer, L. および Watson, J. 他 (1987) The Benjamin/Cummings Publishing Co., Inc. Menlo Park, CA)。特にエラーの多いアミノアシル-tRNAチャージングイベントは、tRNA^{Gln} のグルタミン (Gln) との結合である。このミスチャージングを補正する機序があり、これはその起源を進化の中に持つようである。Glnはポリペプチド合成に利用され自然界に現れる20種の天然アミノ酸の最後の集団にあった。グラム陽性真正細菌、シアノバクテリア、始原菌 (Archeae)、および真核生物のオルガネラは、Gln-tRNA^{Gln} の合成のための非典型的経路 (noncanonical pathway) を持ち、この経路はGlu-tRNA^{Gln} (Glu-tRNAシンテターゼであるGluRSが合成)のトランスフォーメーションに基づくものであり、酵素であるGlu-tRNA^{Gln} アミドトランスフェラーゼ (Glu-AdT)を用いる。アミド基転移経路に関わる反応は次のとおりである (Curnow, A.W. 他 (1997) Nucleic Acids Symposium 36:2-4)。

10

【0053】

GluRS



Glu-AdT



【0054】

類似の酵素であるAsp-tRNA^{Asn} アミドトランスフェラーゼは始原菌に存在し、Asp-tRNA^{Asn} をAsn-tRNA^{Asn}へ転換する。ホルミラーゼ (formylase) 酵素は真正細菌においてMet-tRNA^{fMet} をfMet-tRNA^{fMet}へ転換し、これは近縁の酵素のようである。或る加水分解活性もまた、ミスチャージしたVal-tRNA^{Ile}を破壊することが同定されている (Schimmel, P. 他 (1998) FASEB J. 12:1599-1609)。原始的な生物形態におけるGlu-AdTの進化の、或る可能性のあるシナリオは、特異的グルタミル-tRNAシンテターゼ (GlnRS)の不在が、Gln-tRNA^{Gln}の合成のための代替経路を必要としたというものである。事実、グラム陽性細菌におけるGlu-AdTオペロンの欠失は致命的である (Curnow, A.W. 他 (1997) Proc. Natl. Acad. Sci. U.S.A. 94:11819-11826)。他の生物におけるGluRS活性の存在が自然界の翻訳機構における高度の保存によって推測されているが、GluRSは、ヒトを含む全ての生物において同定されていない。このような酵素は、翻訳忠実度を確保し、欠陥ポリペプチドの合成を減らすよう働くと思われる。

20

30

【0055】

タンパク質合成における機能に加えて、特定のアミノアシルtRNA合成酵素群は、細胞忠実度、RNA スプライシング、RNA輸送、アポトーシス、および転写と翻訳の調節で役割を果たす。例えばヒトのチロシルtRNA合成酵素はたんぱく質分解により、別々のサイトカイン活性を持つ2つの部分に裂け得る。C末端ドメインは単球と白血球の走化作用を示し、また、ミエロペルオキシダーゼ、腫瘍壊死因子- α 、および組織因子の生成を刺激する。N末端ドメインはインターロイキン-8 タイプA受容体に結合し、インターロイキン8様サイトカインとして機能する。ヒトのチロシルtRNA合成酵素 (TyrRS) はアポトーシス段階の腫瘍細胞から分泌される。また、アポトーシスを加速する可能性がある (Wakasugi, K および Schimmel, P (1999) Science 284:147-151)。ミトコンドリア型のアカパンカビ (*N eurospora crassa*) TyrRS および 出芽酵母 (*S. cerevisiae*) LeuRSはグループリントロンの数種のスプライシング活性に必須であり、ヒトのミトコンドリアLeuRSは酵母のヌル種 (null strain) において酵母LeuRSとして代用できる。数種のバクテリアaaRSは、それ自身の転写または翻訳の調節に関与している (前出Martinis)。数種のaaRSは、細胞増殖、分化、およびアポトーシスにおいて役割を持つシグナル伝達分子の1クラスである、ジアドノシン オリゴホスフェートを合成できる (Kisselev, L.L 他 (1998) FEBS Lett. 427:157-163; Vartanian, A. 他 (1999) FEBS Lett. 456:175-180)。

40

【0056】

アミノアシル-tRNAに対する自己抗体は、リウマチ関節炎、皮膚筋炎、及び多発性筋炎な

50

ど、自己免疫疾患の患者によって産生される。また、複雑な間質性肺疾患 (ILD) と強度に相関する (Freist, W. 他 (1999) Biol. Chem. 380:623-646; Freist, W. 他 (1996) Biol. Chem. Hoppe Seyler 377:343-356)。これらの自己抗体はウイルス感染に応じて産生されるとみられ、コクサッキー ウイルスは動物に実験的ウイルス性筋炎を誘発するために用いられている。

【 0 0 5 7 】

aaRS構造の、人体と病原体とでの比較は、新規の抗生物質の設計に有用となっている (Schimmel, 前出)。遺伝子操作したaaRS群を用いて *in vivo* で、非天然アミノ酸群のタンパク質類への部位特異的取り込みが可能となっている (Liu, D.R. 他 (1997) Proc. Natl. Acad. Sci. USA 94:10092-10097)。

tRNAのチャージ (charging)

蛋白質合成は、適切なtRNAとの連結を形成する各アミノ酸に依存している。アミノアシル-tRNA合成酵素は、アミノ酸の活性化と、その同族tRNAとの正しい付着とを担う。20種のアミノアシル-tRNA合成酵素は、2つの構造的クラスに分け得る。クラスI酵素はtRNAの3'末端で2'ヒドロキシルにアミノ酸を加え、クラスII酵素はtRNAの3'末端で3'ヒドロキシルにアミノ酸を加える。各クラスは触媒的ドメインの特有のトポロジーによって特徴づけられる。クラスI酵素には、ヌクレオチド結合した「ロスマンフォールド」に基づく触媒的ドメインがある。クラスII酵素は中央触媒ドメイン (7本鎖の逆平行 β -シートモチーフからなる) と、N末端およびC末端調節ドメインを持つ。クラスII酵素は酵素のヘテロ二量体構造あるいはホモ二量体構造に基づく2つのグループに分けられ、後者はさらにN末端およびC末端調節ドメインの構造によって分けられる (Hartlein, M. および Cusack, S. (1995) J. Mol. Evol. 40:519-530)。アミノアシル-tRNAに対する自己抗体は、皮膚筋炎及び多発性筋炎の患者によって産生される。また、複雑な間質性肺疾患 (ILD) と強度に相関する。これらの自己抗体はウイルス感染に応じて産生されるとみられ、コクサッキー ウイルスは動物に実験的ウイルス性筋炎を誘発するために用いられている。

tRNAの修飾

修飾されたリボヌクレオシドであるプソイドウリジン (y) は、転移RNA (tRNA)、大小のリボソームRNA (rRNA)、および核内低分子RNA (snRNA) のアンチコドン領域に偏在する。yは、tRNAに存在する最も一般的な修飾ヌクレオシド (すなわちG、A、U、およびC以外) である。タンパク合成に無関係のわずかな酵母tRNAのみが、yを持たない (Cortese, R. 他 (1974) J. Biol. Chem. 249:1103-1108)。ウリジンからyへの転換を担う酵素であるプソイドウリジン合成酵素 (プソイドウリジル酸シンターゼ) は初め、ネズミチフス菌 (*Salmonella typhimurium*) から単離された (Arena, F. 他 (1978) Nuc. Acids Res. 5:4523-4536)。この酵素はそれ以後、多数の哺乳類 (例えば去勢ウシやマウス) から単離されている (Green, C.J. 他 (1982) J. Biol. Chem. 257:3045-3052) 並びに Chen, J. および Patton, J.R. (1999) RNA 5:409-419)。tRNAプソイドウリジン合成酵素は、このファミリーの、最も広く研究されているメンバーである。これらの酵素の最適活性には、チオール供与体 (システインなど) と、一価カチオン (アンモニアやカリウムなど) とが必要である。付加的補因子や高エネルギー分子 (ATPやGTP) は不要である (Green, 前出)。別の真核生物プソイドウリジン合成酵素群が同定されており、rRNAに特異的なようである (概説は Smith, C.M. および Steitz, J.A. (1997) Cell 89:669-672)。また、或る二重特異性酵素は、tRNA基質とrRNA基質との双方を利用することが同定されている (Wrzesinski, J. 他 (1995) RNA 1: 437-448)。yがtRNAのアンチコドンループ内に不在の場合、細菌 (Singer, C.E. 他 (1972) Nature New Biol. 238:727-728) と酵母 (Lecoite, F. (1998) 273:1316-1323) との双方において成長が低下するが、この遺伝子欠損は致死的ではない。

【 0 0 5 8 】

別のリボヌクレオシド修飾が主に真核細胞で起こる。これはグアノシンから N^2, N^2 ジメチルグアノシン (m^2_2G) への転換であり、サイトゾルおよびミトコンドリアのtRNAのDステムの塩基位置26または10で起こる。この転写後修飾は、tRNA構造の安定化を、別のtRNA二次構造および三次構造の形成を防ぐことで行うと思われる。酵母tRNA^{Asp} は、この修飾を持

10

20

30

40

50

たない点の特異である。この修飾は真正細菌には起こらず、理由はおそらく、これらの細胞とオルガネラとにおけるtRNAの構造が配列によって制約され、転写後修飾によって別の構造の形成を防ぐ必要がないためである(Steinberg, S.およびCedergren, R. (1995) RNA 1:886-891、およびその参考文献)。グアノシンから m^2_2G への転換を担う酵素は、63 kDaのS-アデノシルメチオニン(SAM)依存性tRNA N^2,N^2 ジメチルグアノシンメチルトランスフェラーゼである(TRM1 遺伝子産物とも呼ばれ、本明細書ではTRMと呼ぶ)(Edqvist, J. (1995) Biochimie 77:5461)。この酵素は、核とミトコンドリアとの双方に局在する(Li, J.-M. 他(1989) J. Cell Biol. 109:1411-1419)。アフリカツメガエル由来のTRMでの研究に基づくと、修飾されるG26の直前のC11~G24とG10~C25との位置での塩基対合には或る要件があると思われ、tRNAの別の構造的特徴も、G26基質の適切な提示に必要なと思われる(Edqvist, J. 他(1992) Nuc. Acids Res. 20:657581)。酵母での研究が示唆するところでは、或る弱いオーカー-tRNAサプレッサ(sup3i)を有する細胞群は、TRM活性が無ければ翻訳終了を抑制できないことから、TRMが真核細胞において抑制の頻度を調整する役割を持つことを示唆する(Niederberger, C. 他(1999) FEBS Lett. 464:6770)。また、TRMは、より一般的な機能である、tRNAの適切な三次元構造の確保も行うようである。

翻訳開始

翻訳の開始は、3つのステージに分けることができる。第一ステージは、或る開始転移RNA(Met-tRNA_f)と40Sリボソームサブユニットとから43S開始前複合体を形成する。第二ステージは、43S開始前複合体をmRNAと結合させた後、複合体を正しいAUG開始コドンに移動させる。第三ステージは、60Sリボソームサブユニットを40Sサブユニットに導き、開始コドンにおいて80Sリボソームを産生する。翻訳の調節には主に、開始プロセスにおける第一および第二のステージが関わる(V.M. Pain (1996) Eur. J. Biochem. 236:747-771)。

【0059】

数種の開始因子(その多くは複数のサブユニットを持つ)が、開始tRNAと40Sリボソームサブユニットとの複合に関わる。eIF2は1種のグアニンヌクレオチド結合タンパク質であり、開始tRNAを40Sリボソームサブユニットまで動員する。eIF2はGTPと結合した時にのみ開始tRNAと会合する。eIF2Bは1種のグアニンヌクレオチド交換タンパク質であり、eIF2の、GDP結合不活性型からGTP結合活性型への転換を担う。2種の別の因子であるeIF1AとeIF3とは、40Sサブユニットと結合し安定化させ、これは18SリボソームRNAおよび特定のリボソーム構造タンパク質群と相互作用して行う。eIF3はまた、40SリボソームサブユニットとmRNAとの会合にも関わる。Met-tRNA_f、eIF1A、eIF3および40Sリボソームサブユニットは共に、43S開始前複合体を形成する(Pain、前出)。

【0060】

eIF2は、mRNA翻訳における律速段階の維持において中心的な役割を果たす。このステップでは、eIF2はGTPとMet-tRNA_iとを結合し、Met-tRNA_iを40Sリボソームサブユニットへ転移する。開始プロセスの最後に、eIF2と結合していたGTPは、GDPへ加水分解される。eIF2.GDP複合体は、リボソームから解放される。eIF2に結合したGDPのGTPとの交換は、Met-tRNA_iと結合するための前提条件であり、この交換は第二の開始因子であるeIF2B(1種のグアニンヌクレオチド交換因子)が仲介する。eIF2のサブユニットにおけるそのリン酸化は、eIF2を、eIF2Bの基質から競合阻害因子へ転換する。したがって、eIF2のリン酸化は、eIF2.GTP.Met-tRNA_i複合体の形成を効果的に防ぎ、全般的なタンパク合成を阻害する。eIF2のリン酸化が起こる種々の条件としては、ウイルス感染、アポトーシス、栄養欠乏、ヘム欠乏、数種のストレスを含む。C型肝炎ウイルス(HCV)の5'非翻訳領域は、内部リボソーム侵入部位(IRES)として機能し、HCVタンパクの翻訳はここで開始する。eIF2BガンマとeIF2ガンマとは、HCV IRESが媒介する翻訳に関与する細胞内因子である(Kimball, S.R. (1999) Int. J. Biochem. Cell Biol. 31:2529; Webb, B.L.およびProud, C.G. (1997) Int. J. Biochem. Cell Biol. 29:11271131; Kruger M. 他(2000) Proc. Natl. Acad. Sci. U S A 97:85668571)。

【0061】

43S開始前複合体がmRNA分子と結合するには、複数の付加的因子が必要であり、このプロ

セスは、いくつかのレベルで調節される。eIF4Fという複合体は、3種のタンパク質、eIF4E、eIF4A、eIF4Gからなる。eIF4EはmRNAの5'-末端^{m7}GTPキャップを認識して結合し、eIF4Aは二方向性のRNA依存型ヘリカーゼであり、eIF4Gは一種の足場ポリペプチドである。eIF4Gは三つの結合ドメインを持つ。eIF4GのN末端側の三分の一はeIF4Eと相互作用し、中央の三分の一はeIF4Aと相互作用し、C末端の三分の一は43S開始前複合体に結合したeIF3と相互作用する。したがって、eIF4Gは、40SリボソームサブユニットとmRNAとの橋渡しをする(M.W. Hentze (1997) Science 275:500-501)。

【0062】

eIF4Fが43S開始前複合体の結合を開始する能力は、mRNAの構造的特徴によって調節される。mRNA分子は、5'キャップとAUG開始コドンとの間に非翻訳領域(UTR)を持つ。数種のmRNAでは、この領域は43S開始前複合体の結合を妨害する二次構造を形成する。eIF4Aのヘリカーゼ活性は、この二次構造を除去して43S開始前複合体の結合を促進するよう機能すると思われる(Pain、前出)。

10

翻訳伸長

伸長プロセスでは、付加的アミノ酸群が開始メチオニンに結合し、完全ポリペプチド鎖を形成する。伸長因子であるEF1、EF1、およびEF2が、開始後のポリペプチド鎖の伸長に関わる。EF1は、GTP結合タンパク質である。EF1のGTP結合型では、EF1がアミノアシルtRNAをリボソームのA部位に導く。新規に到着したアミノアシルtRNAに付着していたアミノ酸は、ペプチド結合を開始メチオニンとの間に形成する。EF1上のGTPはGDPへ加水分解され、EF1-GDPはリボソームから解離する。EF1がEF1-GDPと結合し、GDPをEF1から解離させることにより、EF1はGTPと結合できるようになり新規サイクルが始まる。

20

【0063】

次々にアミノアシルtRNAがリボソームに導かれ、EF-G(別のGTP結合タンパク質)が、リボソームのA部位からP部位へ、そして最後にはE部位へのtRNAの移動を触媒する。これが翻訳の伸長性(processivity)をもたらす。

翻訳終了

放出因子(release factor)であるeRFが翻訳を終了させる。eRFはmRNA内の停止コドン認識し、ポリペプチド鎖をリボソームから遊離させる。

【0064】

Treacher Collins症候群(TCS)は、最も多いヒト下顎顔面異骨(mandibulofacial dysostosis)障害である。TCSは、常染色体優生遺伝を示し、生児出生5万人に1人の割合で起こり、約60%は新規の突然変異による。TCSの症状は、広い可変性を示す。この疾患の原因は、第一鰓弓と第二鰓弓との発生における干渉であると推論される。TCS遺伝子であるTCOF1は、染色体5q31~33.3に局在する。10種の突然変異がTCOF1では同定され、これにはナンセンス変異、挿入、欠失、またはスプライス変異が含まれ、明らかに、翻訳の未熟な終了をもたらしている。さらに、その全ては、それぞれの患者の家族に特有である。TCOF1は低い複雑度のタンパク質(1411アミノ酸)をコードし、そのタンパク質が持つ繰り返しモチーフは、そのエキソン群の構成の鏡像である。これらのモチーフは、別の種における核小体輸送タンパク質でも共有され、カゼインキナーゼによって高度にリン酸化される。完全長TCOF1タンパク質配列内にはまた、核と核小体との局在化シグナルと、幾つかの多型とを含む。このデータが示唆するところでは、TCSの原因は或る核小体輸送タンパク質の欠損であり、このタンパク質は、ヒトの頭蓋顔面発生に決定的に必要である(Wise, C.A.他(1997) Proc. Natl. Acad. Sci. U.S.A. 94:31103115)。

30

40

乳癌

180,000例を超える新規発症の乳癌が毎年診断されており、乳癌による死亡率は45~54歳の女性の全死亡の10%に近い(Gish, K. (1999) AWIS Magazine 28:7-10)。ただし、限局性乳癌の早期診断に基づく生存率はきわめて高い(97%)。これに対して、疾患のステージが進行し腫瘍が乳房外に進展した場合は22%である。現在の臨床乳癌検診の手法は感度と特異性とを欠いており、乳癌の包括的な遺伝子発現プロファイルを開発する努力がなさ

50

れている。このプロフィールを従来のスクリーニング法とともに用いれば、乳癌の診断と予後診断とを向上し得る (Perou, C.M. 他 (2000) Nature 406:747-752)。

【 0 0 6 5 】

2種の遺伝子、BRCA1とBRCA2との突然変異が、女性の乳癌の大きな素因であることが知られており、この変異は親から子へ受け渡されるようである (前出Gish)。しかし、このタイプの遺伝性乳癌はわずかに乳癌の約5%~9%であり、大多数の乳癌の原因は非遺伝性突然変異であり、乳房上皮細胞での変異である。

【 0 0 6 6 】

上皮成長因子(EGF)とその受容体であるEGFRとの発現と、ヒト乳癌との関係は、特に良く研究されている (この分野の概説には、Khazaie, K. 他 (1993) Cancer and Metastasis Rev. 12:255274、およびその中の参考文献を参照)。EGFRの過剰発現、特にエストロゲン受容体の下方調節と結びついた発現は、乳癌患者の不良な予後のマーカーである。また、乳癌転移におけるEGFR発現は、しばしば原発腫瘍に比して上昇することから、EGFRは腫瘍の進行と転移とに関わることを示唆している。これを支持する証拠が蓄積されている。それは、EGFが、転移の潜在能に関する細胞機能に対する効果を持つという証拠である。この機能とは例えば、細胞運動性、走化性、分泌、および分化である。erbB受容体ファミリー (EGFRはその1つである) の他のメンバーの発現の変化も、乳癌に関わると考えられている。erbB受容体 (例えばHER-2/neu, HER-3およびHER-4) およびそれらのリガンドの乳癌における豊富さは乳癌の病原におけるそれらの機能的な重要性を示しており、したがって乳癌治療の標的を提供しうる (Bacus, S. S. 他 (1994) Am. J. Clin. Pathol. 102:S13-S24)。別の既知の乳癌マーカーとしては、或るヒトの分泌型frizzledタンパク質の mRNA (乳癌では下方調節される)、マトリクスG1aタンパク質 (ヒト乳癌細胞で過剰発現)、Drg1すなわちRTP (この遺伝子の発現は結腸癌、乳癌、前立腺腫瘍で低下)、maspin (この腫瘍抑制遺伝子は侵襲性乳癌で下方調節)、およびCaN19 (S100蛋白ファミリーのメンバーであり、このファミリーは全て、乳癌細胞では正常乳腺上皮細胞に比して下方調節される) を含む (Zhou, Z. 他 (1998) Int. J. Cancer 78:95-99; Chen, L. 他 (1990) Oncogene 5:1391-1395; Ullrich, W. et al (1999) FEBS Lett 455:23-26; Sager, R. 他 (1996) Curr. Top. Microbiol. Immunol. 213:51-64; および Lee, S. W. 他 (1992) Proc. Natl. Acad. Sci. USA 89:2504-2508)。

【 0 0 6 7 】

種々の病期のヒト乳癌の乳腺上皮細胞に由来する細胞株は、悪性転換と腫瘍進行とのプロセスを研究するのに有用なモデルを提供する。その理由は、これらの細胞株が、その親腫瘍の多くの特性を、長い培養期間にわたり保持することが示されているからである (Wistuba, I. I. 他 (1998) Clin. Cancer Res. 4:2931-2938)。このようなモデルが特に有用なのは、ヒト乳腺上皮細胞の表現型と分子との特徴を、種々のステージの悪性転換について比較する場合である。

前脂肪細胞

脂肪組織の最重要の機能は、脂肪を摂食時に貯蔵し空腹時に放出する能力である。白色脂肪組織は過剰エネルギー使用時のための主なエネルギー備蓄部位であり、その第一目的は、エネルギー欠乏時の動員性 (mobilization) である。いかにして種々の分子が、脂肪過剰症と、生理的および病態生理的な諸状況でのエネルギーバランスを調節するかを理解すれば、ヒト肥満症の新規の治療法を開発しうる。脂肪組織はまた、インスリンの重要な標的組織の1つである。II型糖尿病における脂肪生成とインスリン耐性とは関連しており、興味深い関係を示す。II型糖尿病の殆どの患者は肥満であり、肥満は次にインスリン耐性を生じる。

【 0 0 6 8 】

脂肪細胞の生物学の、これまでの研究の大半に、形質転換したマウス前脂肪細胞株が利用されている。マウス前脂肪細胞の分化を刺激する培養条件は、ヒトの初代前脂肪細胞の分化を誘発する条件とは異なる。また、初代細胞は二倍体であるため、in vivo条件を異数体細胞株以上に反映すると思われる。ヒトの脂肪生成期の遺伝子発現プロフィールを理解

すれば、脂肪過多症調節の根本機序を理解できる。更に、脂肪生成の遺伝子発現プロフィールを、正常体重ドナーと肥満ドナーとで比較すれば、非常に重要な遺伝子群、すなわち肥満とII型糖尿病との潜在的な薬物標的の同定が可能となる。

ペルオキシソーム増殖因子活性化受容体 アゴニスト

チアゾリジンジオン (thiazolidinediones) (TZD) は、ペルオキシソーム増殖因子活性化受容体 (PPAR) のアゴニストとして作用する。PPAR は、核ホルモン受容体スーパーファミリのメンバーである。TZDは、高血糖、高インスリン血、および高血圧の低減を、一部にはグルコース代謝を促進して、また糖新生を阻害して行う。PPAR とそのアゴニストとの役割は、広範な病態で実証されている。この病態には例えば糖尿病、肥満、高血圧、アテローム硬化、多嚢胞卵巣症候群、乳癌、前立腺癌、脂肪肉腫、および結腸癌を含む。

10

【 0 0 6 9 】

TZD他のPPAR アゴニストがインスリン感受性を増強する機序は十分に理解されていないが、PPAR が脂肪生成を促進する能力に関わると思われる。培養した前脂肪細胞で異所的に発現されたPPAR は、脂肪細胞分化の強力な誘発因子である。TZDをインスリン他の因子と併用するとまた、培地でのヒト前脂肪細胞の分化を増強できる (Adams 他 (1997) J. Clin. Invest. 100:3149-3153) 。種々のTZDが脂肪生成を *in vitro* で促進する相対的な効力は、それらの *in vivo* でのインスリン感作効果と、PPAR に *in vitro* で結合し活性化するそれらの能力との双方に比例する。興味深いことに、大網脂肪貯蔵 (omental adipose depots) に由来する脂肪細胞は、TZDの効果に不応である。これが示唆するところでは、TZD 20 類のインスリン感作効果は、それらが脂肪生成を皮下脂肪貯蔵において促進する能力の結果である (Adams 他、同上) 。更に、PPAR 遺伝子の優性阻害突然変異 (dominant negative mutations) の同定が、二名の非肥満例でなされている。彼らには重度のインスリン耐性、高血圧、および、顕性の非インスリン依存性糖尿病 (NIDDM) がみられた (Barroso 他 (1998) Nature 402:880-883) 。

20

【 0 0 7 0 】

NIDDMは最も多い型の糖尿病であり、糖尿病は慢性代謝病であって世界に1.43億人の患者がいる。NIDDMの特徴は異常なグルコースと脂質との代謝であり、その原因は、末梢インスリン耐性と、欠陥のあるインスリン分泌とが組み合わさっている。NIDDMは複雑な進行性の病因を持ち、高度の遺伝率をも持つ。多数の合併症が糖尿病に併発し (例えば心臓病、卒中、腎不全、網膜症、および末梢神経障害を含む)、高い罹患率と死亡率との一因となっている。

30

【 0 0 7 1 】

分子レベルでPPAR は、リガンド活性化転写因子として機能する。リガンドの存在下でPPAR はレチノイドX受容体 (RXR) と共にヘテロ二量体を形成する。この二量体は標的遺伝子の転写を活性化する。この標的遺伝子は、或るPPAR 応答エレメント (PPRE) の1つ以上のコピーを持つ。脂質の貯蔵と代謝とに重要な多くの遺伝子がPPREをもち、PPAR 標的として同定されている。これには例えばPEPCK、aP2、LPL、ACS、およびFAT-Pを含む (Auwerx , J. (1999) Diabetologia 42:1033-1049) 。多数のPPAR リガンドが同定されている。このリガンドとしては例えば種々の脂肪酸代謝産物や、合成薬物のうちのTZDクラス (例えばピオグリタゾン : Pioglitazone およびロシグリタゾン : Rosiglitazone (BRL49653)) 並びに、数種の非グリタゾン系チロシン類似体 (例えばGI262570 および GW1929) を含む。プロスタグランジン誘導体である15dPGJ2は、PPAR の強力な内因性リガンドである。

40

【 0 0 7 2 】

PPAR の発現は脂肪内では非常に高いが、骨格筋では殆ど検出されない。骨格筋は体内における、インスリンに刺激されたグルコース処理の主要部位である。PPAR の中程度の発現はまた、大腸、腎臓、肝臓、血管平滑筋、造血細胞、およびマクロファージに見られる。脂肪でのPPAR の高発現が示唆するところでは、TZD類のインスリン感作効果は、脂肪組織における、1つ以上のPPAR 調節遺伝子 (PPAR regulated genes) の発現における変容の結果である。PPAR 標的遺伝子群を同定すれば、糖尿、肥満その他の症状の為のより

50

良い薬物設計と、新規治療戦略の開発とに寄与するであろう。

【0073】

PPAR 標的遺伝子を同定する体系的な試みが、肥満と糖尿との数種のげっ歯類モデルでなされている(Suzuki 他(2000) Jpn. J. Pharmacol. 84:113123; Way 他(2001) Endocrinology 142:1269-1277)。しかし、深刻な欠点がげっ歯類遺伝子発現研究にはある。それは、ヒトとげっ歯類モデルでの、脂肪生成、糖尿、および肥満の、顕著な差異である(Taylor (1999) Cell 97:9-12; Gregoire 他(1998) Physiol.Reviews 78:783-809)。したがって、ヒト組織の初代培養におけるTZD調節遺伝子を同定するための、偏向のない手法が、PPAR 活性に関連する疾患の分子的原理を十分に解明するのに必要である。

【0074】

アレイ技術は、単一の多型遺伝子の発現や、多数の関連遺伝子または無関係の遺伝子の発現プロフィールを探求する簡単な方法を提供し得る。単一遺伝子の発現を試験するときは、アレイを用いて或る特定遺伝子又はその変異体の発現を検出する。発現プロフィールを試験するときは、アレイは次のような試験を行うプラットフォームを提供する。即ちどの遺伝子が組織特異的か、ハウスキーピング機能を行うか、シグナル伝達カスケードの一部であるか、又は、特定の遺伝的素因や、条件、疾患、または障害に、特異的に関連する遺伝子であるかの試験である。遺伝子発現プロフィール作製の潜在的応用は特に、疾患の、診断、予後診断、および治療の向上に関わる。例えば、糖尿病患者に由来の組織に発現されるレベルと配列との双方を、正常組織に発現されるレベルおよび配列と比較しうる。

【0075】

新規の核酸関連タンパク質およびそれらをコードするポリヌクレオチド群の発見により、新規の組成物群を提供することで、当分野の要望に応えることができる。これら新規の組成物は、細胞増殖異常、神経障害、発達障害、自己免疫/炎症疾患、および感染症の、診断・治療・予防において有用であり、また、核酸関連タンパク質の核酸配列およびアミノ酸配列の発現における外来化合物群の効果についての算定にも有用である。

【0076】

(発明の概要)

本発明は、総称して「NAAP」、個別には各々「NAAP-1」、「NAAP-2」、「NAAP-3」、「NAAP-4」、「NAAP-5」、「NAAP-6」、「NAAP-7」、「NAAP-8」、「NAAP-9」、「NAAP-10」、「NAAP-11」、「NAAP-12」、「NAAP-13」、「NAAP-14」、「NAAP-15」、および「NAAP-16」と呼ぶ、核酸関連タンパク質である、精製されたポリペプチドを提供する。或る実施例において本発明は、(a) SEQ ID NO:1-16からなる群から選択した或るアミノ酸配列を持つポリペプチド、(b) SEQ ID NO:1-16からなる群から選択した或るアミノ酸配列との少なくとも90%の同一性を持つ天然アミノ酸配列を有するポリペプチド、(c) SEQ ID NO:1-16からなる群から選択した或るアミノ酸配列を持つポリペプチドの生物学的活性断片、および(d) SEQ ID NO:1-16からなる群から選択した或るアミノ酸配列を持つポリペプチドの免疫原性断片、からなる群から選択した、単離されたポリペプチドを提供する。一実施例で本発明は、SEQ ID NO:1-16のアミノ酸配列を持つ、単離されたポリペプチドを提供する。

【0077】

また、本発明は(a) SEQ ID NO:1-16からなる群から選択した或るアミノ酸配列を持つポリペプチド、(b) SEQ ID NO:1-16からなる群から選択した或るアミノ酸配列との少なくとも90%の同一性を持つ或る天然アミノ酸配列を有するポリペプチド、(c) SEQ ID NO:1-16からなる群から選択した或るアミノ酸配列を持つポリペプチドの生物学的活性断片、または(d) SEQ ID NO:1-16からなる群から選択した或るアミノ酸配列を持つポリヌクレオチドの免疫原性断片、からなる群から選択した或るポリペプチドをコードする、単離されたポリヌクレオチドを提供する。一実施態様では、該ポリヌクレオチドは、SEQ ID NO:1-16からなる群から選択した或るポリペプチドをコードする。別の実施態様では、ポリヌクレオチドはSEQ ID NO:17-32からなる群から選択される。

【0078】

本発明は更に、(a) SEQ ID NO:1-16からなる群から選択した或るアミノ酸配列を持つポリペプチド、(b) SEQ ID NO:1-16からなる群から選択した或るアミノ酸配列との少なくとも90%の同一性を有する或る天然アミノ酸配列を持つポリペプチド、(c) SEQ ID NO:1-16からなる群から選択した或るアミノ酸配列を有するポリペプチドの生物学的活性断片、および(d) SEQ ID NO:1-16からなる群から選択した或るアミノ酸配列を持つポリペプチドの免疫原性断片、からなる群から選択した或るポリペプチドをコードするポリヌクレオチドと機能的に連結したプロモーター配列を持つ組換えポリヌクレオチドを提供する。一実施態様で本発明は、この組換えポリヌクレオチドを用いて形質転換した細胞を提供する。別の実施態様で本発明は、この組換えポリヌクレオチドを持つ遺伝形質転換体を提供する。

10

【0079】

更に本発明は、(a) SEQ ID NO:1-16からなる群から選択した或るアミノ酸配列を持つポリペプチドと、(b) SEQ ID NO:1-16からなる群から選択した或るアミノ酸配列との少なくとも90%以上の同一性を有する或る天然アミノ酸配列を持つポリペプチドと、(c) SEQ ID NO:1-16からなる群から選択した或るアミノ酸配列を持つポリペプチドの生物学的活性断片と、(d) SEQ ID NO:1-16からなる群から選択したアミノ酸配列を持つポリペプチドの免疫原性断片、とからなる群から選択したポリペプチドを製造する一方法を提供する。製造方法は、(a) 或る細胞を該ポリペプチドの発現に適した条件下で培養し、この細胞を組換えポリヌクレオチドを用いて形質転換する過程と、(b) そのように発現したポリペプチドを回収する過程からなる。この組換えポリヌクレオチドは、該ポリペプチド

20

【0080】

本発明は更に、(a) SEQ ID NO:1-16からなる群から選択した或るアミノ酸配列を持つポリペプチド、(b) SEQ ID NO:1-16からなる群から選択した或るアミノ酸配列との少なくとも90%の同一性を有する或る天然アミノ酸配列を持つポリペプチド、(c) SEQ ID NO:1-16からなる群から選択したアミノ酸配列を持つポリペプチドの生物学的活性断片、および(d) SEQ ID NO:1-16からなる群から選択した或るアミノ酸配列を持つポリペプチドの免疫原性断片、からなる群から選択した或るポリペプチドに特異結合する、単離された抗体を提供する。

【0081】

本発明は更に、(a) SEQ ID NO:17-32からなる群から選択したポリヌクレオチド配列を持つポリヌクレオチド、(b) SEQ ID NO:17-32からなる群から選択した或るポリヌクレオチド配列との少なくとも90%の同一性を有する或る天然ポリヌクレオチド配列を持つポリヌクレオチド、(c) (a) に相補的なポリヌクレオチド、(d) (b) に相補的なポリヌクレオチド、および(e) (a) ~ (d) のRNA等価物、からなる群から選択した、単離されたポリヌクレオチドを提供する。一実施態様で該ポリヌクレオチドは、少なくとも60の連続したヌクレオチド群からなる。

30

【0082】

本発明は更に、サンプル中の標的ポリヌクレオチドを検出する一方法を提供する。ここで、標的ポリヌクレオチドは、(a) SEQ ID NO:17-32からなる群から選択した或るポリヌクレオチド配列を持つポリヌクレオチド、(b) SEQ ID NO:17-32からなる群から選択した或るポリヌクレオチド配列との少なくとも90%の同一性を有する天然ポリヌクレオチド配列を持つポリヌクレオチド、(c) (a) のポリヌクレオチドに相補的なポリヌクレオチド、(d) (b) のポリヌクレオチドに相補的なポリヌクレオチド、および(e) (a) ~ (d) のRNA等価物、からなる群から選択した、或るポリヌクレオチドの配列を持つ。検出方法は、(a) サンプル中の上記標的ポリヌクレオチドに相補的な或る配列からなる少なくとも20の連続したヌクレオチド群からなる或るプローブを用いて該サンプルをハイブリダイズする過程と、(b) 該ハイブリダイゼーション複合体の有無を検出し、複合体が存在すればオプションでその量を検出する過程からなる。該プローブと該標的ポリヌクレオチドあるいはその断片との間でハイブリダイゼーション複合体が形成されるよ

40

50

うな条件下で、プローブは、該標的ポリヌクレオチドに対し特異的にハイブリダイズする。一実施態様では、プローブは少なくとも60の連続したヌクレオチド群からなる。

【0083】

本発明はまた、サンプル中の標的ポリヌクレオチドを検出する一方法を提供する。ここで、標的ポリヌクレオチドは、(a) SEQ ID NO: 17-32からなる群から選択した或るポリヌクレオチド配列を持つポリヌクレオチド、(b) SEQ ID NO: 17-32からなる群から選択した或るポリヌクレオチド配列との少なくとも90%の同一性を有する天然ポリヌクレオチド配列を持つポリヌクレオチド、(c) (a)のポリヌクレオチドに相補的なポリヌクレオチド、(d) (b)のポリヌクレオチドに相補的なポリヌクレオチド、および(e) (a) ~ (d)のRNA等価物、からなる群から選択した、或るポリヌクレオチドの配列を持つ。検出方法は、(a)ポリメラーゼ連鎖反応増幅を用いて標的ポリヌクレオチドまたはその断片を増幅する過程と、(b)増幅した標的ポリヌクレオチドまたはその断片の有無を検出し、該標的ポリヌクレオチドまたはその断片が存在すればオプションでその量を検出する過程からなる。

10

【0084】

本発明は更に、或る有効量のポリペプチドと薬剤として許容し得る或る賦形剤とからなる、或る組成物を提供する。有効量のポリペプチドは、(a) SEQ ID NO: 1-16からなる群から選択した或るアミノ酸配列を持つポリペプチド、(b) SEQ ID NO: 1-16からなる群から選択した或るアミノ酸配列との少なくとも90%の同一性を有する天然のアミノ酸配列を含むポリペプチド、(c) SEQ ID NO: 1-16からなる群から選択した或るアミノ酸配列を含むポリペプチドの生物学的活性断片、および(d) SEQ ID NO: 1-16からなる群から選択した或るアミノ酸配列を持つポリペプチドの免疫原性断片、からなる群から選択される。一実施態様では、この組成物はSEQ ID NO: 1-16からなる一群から選択されたアミノ酸配列を持つ。更に、本発明は、機能的NAAPの発現の低下に関連した疾患や症状の治療を必要とする患者にこの組成物を投与することを含む方法を提供する。

20

【0085】

本発明はまた、(a) SEQ ID NO: 1-16からなる群から選択した或るアミノ酸配列を持つポリペプチド、(b) SEQ ID NO: 1-16からなる群から選択した或るアミノ酸配列との少なくとも90%の同一性を有する天然アミノ酸配列を持つポリペプチド、(c) SEQ ID NO: 1-16からなる群から選択したアミノ酸配列を持つポリペプチドの生物学的活性断片、および(d) SEQ ID NO: 1-16からなる群から選択した或るアミノ酸配列を持つポリペプチドの免疫原性断片、からなる群から選択したポリペプチドのアゴニストとしての有効性を確認するために、或る化合物をスクリーニングする方法を提供する。スクリーニング方法は、(a)該ポリペプチドを有するサンプルを或る化合物に曝す過程と、(b)サンプル中のアゴニスト活性を検出する過程からなる。別法では、本発明は、この方法で同定したアゴニスト化合物と許容される医薬用賦形剤とを有する、或る組成物を提供する。更なる別法では、本発明は、機能的NAAPの発現の低下を伴う疾患や症状の治療を要する患者への、この組成物の投与方法を提供する。

30

【0086】

本発明は更に、(a) SEQ ID NO: 1-16からなる群から選択した或るアミノ酸配列を持つポリペプチド、(b) SEQ ID NO: 1-16からなる群から選択した或るアミノ酸配列との少なくとも90%の同一性を有する天然アミノ酸配列を持つポリペプチド、(c) SEQ ID NO: 1-16からなる群から選択したアミノ酸配列を持つポリペプチドの生物学的活性断片、および(d) SEQ ID NO: 1-16からなる群から選択したアミノ酸配列を持つポリペプチドの免疫原性断片、からなる群から選択したポリペプチドのアンタゴニストとしての有効性につき、或る化合物をスクリーニングする方法を提供する。スクリーニング方法は、(a)該ポリペプチドを含むサンプルを或る化合物に曝す過程と、(b)サンプル中のアンタゴニスト活性を検出する過程からなる。別の一実施態様で本発明は、この方法で同定したアンタゴニスト化合物と薬剤として許容し得る賦形剤とを有する組成物を提供する。更なる別法では、本発明は、機能的NAAPの過剰な発現に関連した疾患や症状の治療を必要とする患者へ

40

50

のこの組成物の投与を含む方法を提供する。

【0087】

本発明は更に、(a) SEQ ID NO:1-16からなる群から選択した或るアミノ酸配列を持つポリペプチド、(b) SEQ ID NO:1-16からなる群から選択した或るアミノ酸配列との少なくとも90%の同一性を有する或る天然アミノ酸配列を持つポリペプチド、(c) SEQ ID NO:1-16からなる群から選択したアミノ酸配列を持つポリペプチドの生物学的活性断片、または(d) SEQ ID NO:1-16からなる群から選択したアミノ酸配列を持つポリペプチドの免疫原性断片、からなる群から選択した或るポリペプチドに特異結合する或る化合物をスクリーニングする方法を提供する。スクリーニング方法は、(a) 該ポリペプチドを適切な条件下で少なくとも1つの試験化合物に混合させる過程と、(b) この試験化合物と該ポリペプチドとの結合を検出し、それにより該ポリペプチドに特異結合する化合物を同定する過程からなる。

10

【0088】

本発明は更に、(a) SEQ ID NO:1-16からなる群から選択した或るアミノ酸配列を持つポリペプチド、(b) SEQ ID NO:1-16からなる群から選択したアミノ酸配列との少なくとも90%の同一性を有する或る天然アミノ酸配列を持つポリペプチド、(c) SEQ ID NO:1-16からなる群から選択した或るアミノ酸配列を持つポリペプチドの生物学的活性断片、および(d) SEQ ID NO:1-16からなる群から選択した或るアミノ酸配列を持つポリペプチドの免疫原性断片、からなる群から選択した或るポリペプチドの活性をモジュレートする或る化合物をスクリーニングする方法を提供する。スクリーニング方法は、(a) 該ポリペプチドの活性にとり許容し得る条件下で、該ポリペプチドを少なくとも1つの試験化合物と混合する過程と、(b) 該ポリペプチドの活性をこの試験化合物の存在下で算定する過程と、(c) この試験化合物の存在下での該ポリペプチドの活性をこの試験化合物の不在下での該ポリペプチドの活性と比較する過程からなり、この試験化合物の存在下での該ポリペプチドの活性の変化は、該ポリペプチドの活性をモジュレートする化合物を標示する。

20

【0089】

更に本発明は、SEQ ID NO:17-32からなる群から選択した或るポリヌクレオチド配列を持つ標的ポリヌクレオチドの発現を改変する効果につき、或る化合物をスクリーニングする一方法を提供する。この方法は、(a) この標的ポリヌクレオチドを有するサンプルを或る化合物に曝露する過程と、(b) この標的ポリヌクレオチドの発現の改変を検出する過程と、(c) 可変量のこの化合物の存在下でのこの標的ポリヌクレオチドの発現と、この化合物の不在下での発現とを比較する過程とからなる。

30

【0090】

本発明は更に、試験化合物の毒性の算定方法を提供する。この方法には、以下の過程がある。(a) 核酸群を有する生体サンプルを試験化合物で処理する過程。(b) 処理済み生体サンプルの核酸群をハイブリダイズする過程。この過程には、次のようなプローブを用いる。(i) SEQ ID NO:17-32からなる群から選択した或るポリヌクレオチド配列を持つポリヌクレオチド、(ii) SEQ ID NO:17-32からなる群から選択した或るポリヌクレオチド配列との少なくとも90%の同一性を有する天然ポリヌクレオチド配列を持つポリヌクレオチド、(iii) (i) に相補的な配列を持つポリヌクレオチド、(iv) (ii) のポリヌクレオチドに相補的なポリヌクレオチド、(v) (i) ~ (iv) のRNA等価物、からなる群から選択した或るポリヌクレオチドの少なくとも20の連続したヌクレオチド群からなるプローブである。ハイブリダイゼーションは、上記プローブと生体サンプル中の標的ポリヌクレオチドとの間に特異的ハイブリダイゼーション複合体が形成されるような条件下で生じる。上記標的ポリヌクレオチドは、(i) SEQ ID NO:17-32からなる群から選択した或るポリヌクレオチド配列を持つポリヌクレオチド、(ii) SEQ ID NO:17-32からなる群から選択した或るポリヌクレオチド配列との少なくとも90%の同一性を有する天然ポリヌクレオチド配列を持つポリヌクレオチド、(iii) (i) のポリヌクレオチドに相補的なポリヌクレオチド、(iv) (ii) のポリヌクレオチドに相補的なポリヌクレオチド、(v) (i)

40

50

) ~ (iv) のRNA等価物、からなる群から選択する。あるいは、標的ポリヌクレオチドは、上記(i) ~ (v) からなる群から選択したポリヌクレオチド配列の断片を持つ。毒性の算定方法には更に、以下の過程がある。(c) ハイブリダイゼーション複合体の量を定量する過程と、(d) 処理済み生体サンプル中のハイブリダイゼーション複合体の量を、非処理の生体サンプル中のハイブリダイゼーション複合体の量と比較する過程である。処理済み生体サンプル中のハイブリダイゼーション複合体の量の差異が、試験化合物の毒性を示す。

【0091】

(本発明の記述について)

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

10

【0092】

請求の範囲および明細書中で用いている単数形の「或る」および「その(この)」の表記は、文脈から明らかにそうでないとされる場合を除いて複数のもを指す場合もあることに注意されたい。したがって、例えば「或る宿主細胞」と記されている場合にはそのような宿主細胞が複数あることもあり、「或る抗体」と記されている場合には単数または複数の抗体、および、当業者に公知の抗体の等価物などについても言及している。

20

【0093】

本明細書中で用いる全ての技術用語および科学用語は、特に定義されている場合を除き、当業者に一般に理解されている意味と同じ意味を有する。本明細書で説明するものと類似あるいは同等の任意の装置、材料および方法を用いて本発明の実施または試験を行うことができるが、ここでは好適な装置、材料、方法について説明する。本発明で言及する全ての刊行物は、刊行物中で報告されていて且つ本発明に関係があるであろう細胞、プロトコル、試薬およびベクターについて説明および開示する目的で引用しているものである。本明細書のいかなる開示内容も、本発明が先行技術の効力によってこのような開示に対して先行する権利を与えられていないことを認めるものではない。

【0094】

(定義)

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

30

【0095】

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

【0096】

用語「対立遺伝子変異配列」は、NAAPをコードする別の形の遺伝子を指す。対立遺伝子変異配列群は、核酸配列における少なくとも1つの突然変異から生じ得る他、変容したmRNA群を生じ得る。また、ポリペプチド群を生じ得るが、それらのポリペプチドの構造または機能は、変容することもしないこともある。或る遺伝子は、その天然型の対立遺伝子変異配列を全く持たない場合もあり、1個以上持つこともある。対立遺伝子変異配列を生じさせる通常の突然変異性変化は一般に、ヌクレオチドの自然な欠失、付加または置換による。これら各種の変化は、単独であるいは他の変化と共に、或る配列内で1回以上、生じ得る。

40

【0097】

NAAPをコードする「変容した/改変された」核酸配列に含む配列には、様々なヌクレオチ

50

ドの欠失、挿入、或いは置換を生じた配列であって、NAAPと同一のポリペプチドを、或いは、NAAPの機能的特徴の少なくとも1つを備えるポリペプチドを生じる配列がある。この定義には、NAAPをコードするポリヌクレオチド配列にとり正常な染色体の遺伝子座ではない位置での、対立遺伝子変異配列群への不適当あるいは予期しないハイブリダイゼーションを含み、また、NAAPをコードするポリヌクレオチドの或る特定のオリゴヌクレオチドブロープを用いて容易に検出可能なあるいは検出困難な多型性を含む。コードされるタンパク質も「変容する/改変される」ことがあり、また、サイレント変化を生じた結果、機能的に等価なNAAPとなるような、アミノ酸残基の欠失、挿入または置換を持ち得る。意図的なアミノ酸置換は、生物学的あるいは免疫学的にNAAPの活性が保持される範囲で、残基群の、極性、電荷、溶解度、疎水性、親水性、および/または両親媒性についての、類似性に基いて成され得る。例えば、負に帯電したアミノ酸にはアスパラギン酸およびグルタミン酸があり、正に帯電したアミノ酸にはリジンおよびアルギニンがある。親水性値が近似した非荷電極性側鎖を持つアミノ酸には、アスパラギンとグルタミン、セリンとトレオニンがある。親水性値が近似した非荷電側鎖を持つアミノ酸には、ロイシンとイソロイシンとバリン、グリシンとアラニン、フェニルアラニンとチロシンがある。

10

【0098】

用語「アミノ酸」および「アミノ酸配列」は、オリゴペプチド、ペプチド、ポリペプチド、タンパク質配列、あるいはそれらのいずれかの断片を指し、天然分子または合成分子を指す。「アミノ酸配列」が或る天然タンパク質分子の配列を指す場合、「アミノ酸配列」および類似の用語は、記載したそのタンパク質分子に関連する完全で元のままのアミノ酸配列に限定するものではない。

20

【0099】

用語「増幅」は、或る核酸配列の付加的複製物を作製する行為に関する。増幅には通常、当業者に公知のポリメラーゼ連鎖反応(PCR)技術を用いる。

【0100】

用語「アンタゴニスト」は、NAAPの生物学的活性を阻害或いは減弱する分子を指す。アンタゴニストとしては、NAAPと直接に相互作用するか或いはNAAPが関与する生物学的経路の諸成分に作用してNAAPの活性を調節する、抗体などのタンパク質、核酸、糖質、小分子、任意の他の化合物や組成物を含み得る。

【0101】

用語「抗体」は、エピトープの決定基と結合できる、無傷の免疫グロブリン分子やそれらの断片、例えばFab、F(ab')₂、およびFv断片を指す。NAAPポリペプチド群と結合する抗体類の作製には、免疫化抗原として、無傷ポリペプチド群を用いることができ、または、当該の小ペプチド群を有する断片群を用い得る。マウス、ラット、ウサギなどの動物を免疫化するために用いるポリペプチドまたはオリゴペプチドの由来は、RNAの翻訳の場合や、または化学合成があり得る。また、それらは所望により、キャリアタンパク質に抱合し得る。通常用いられるキャリアであってペプチドと化学結合するキャリアの例は、ウシ血清アルブミン、サイログロブリン、および、スカシガイのヘモシアニン(KLH)などがある。結合したこのペプチドは、前記動物を免疫化するために用いる。

30

【0102】

用語「抗原決定基」は、特定の抗体と接触する、分子の領域(すなわちエピトープ)を指す。タンパク質またはタンパク質断片を用いて宿主動物を免疫化する場合、タンパク質の多数の領域が、抗原決定基(タンパク質の特定の領域または3次元構造)に特異結合する抗体の産生を誘発し得る。或る抗原決定基は、或る抗体への結合について、無傷抗原(すなわち免疫応答を引き出すために用いられる免疫原)と競合し得る。

40

【0103】

用語「アプタマー(apptamer)」は、核酸またはオリゴヌクレオチド分子であって、特定の分子ターゲットに結合する分子を指す。アプタマーは*in vitro*での進化プロセスに由来する(例えば、SELEX(Systematic Evolution of Ligands by EXponential Enrichment)の略、試験管内選択法)、米国特許第5,270,163号に記述)。これは、大規模な組み合わせ

50

ライブラリ群から標的特異的アプタマー配列を選択するプロセスである。アプタマーの構成は二本鎖または一本鎖であり、デオキシリボヌクレオチド、リボヌクレオチド、ヌクレオチド誘導体または他のヌクレオチド様分子を含み得る。アプタマーの各ヌクレオチド成分は修飾された糖基を持つことがあり（例えば、リボヌクレオチドの2'-OH基は2'-Fまたは2'-NH₂によって置換され得る）、これは、ヌクレアーゼへの抵抗性または血中でのより長い寿命など、所望の特性を向上し得る。アプタマーを他の分子（例えば高分子量のキャリア）と抱合させることにより、循環系からのこのアプタマーのクリアランスを遅らせ得る。アプタマー類は、例えば或る架橋剤の光活性化によって、それらの同種リガンド群と特異的に架橋させ得る（例えば、Brody, E.N.およびL. Gold (2000) J. Biotechnol. 74: 5-13を参照）。

10

【0104】

用語「イントラマー (intramer)」は、in vivoで発現されるアプタマーを指す。例えば、ワクシニアウイルスに基づく或るRNA発現系を用いて、白血球の細胞質内で特定のRNAアプタマー類が高レベルに発現されている (Blind, M.他(1999) Proc.Natl Acad. Sci. USA 96:3606-3610)。

【0105】

用語「スピーゲルマー (spiegelmer)」は、アプタマーの内、L-DNA、L-RNAなどの左旋性ヌクレオチド誘導体、または左旋性ヌクレオチド様分子などを指す。左旋性のヌクレオチド群を持つアプタマー類は、右旋性ヌクレオチド群を持つ基質に通常は作用する、天然酵素類による分解に対して耐性がある。

20

【0106】

用語「アンチセンス」は、或る特定の核酸配列の「センス」(コーディング)鎖との塩基対を形成し得る任意の組成物を指す。アンチセンス組成物としては、DNAや、RNAや、ペプチド核酸 (PNA) や、修飾されたバックボーン連結たとえばホスホロチオ酸、メチルホスホン酸またはベンジルホスホン酸などを有するオリゴヌクレオチドや、修飾された糖基たとえば2'-メトキシエチル糖または2'-メトキシエトキシ糖などを有するオリゴヌクレオチドや、あるいは修飾された塩基たとえば5-メチルシトシン、2-デオキシウラシルまたは7-デアザ-2'-デオキシグアノシンなどを有するオリゴヌクレオチドがあり得る。アンチセンス分子は、化学合成または転写など、任意の方法で製造することができる。相補的アンチセンス分子は、細胞に導入されると、細胞が産生した天然核酸配列との塩基対を形成し、二重鎖を形成して転写または翻訳を妨害する。「負」または「マイナス」という表現は、或る参考DNA分子のアンチセンス鎖を指すことがあり、「正」または「プラス」という表現は、或る参考DNA分子のセンス鎖を意味し得る。

30

【0107】

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

【0108】

用語「相補的」は、塩基対合によってアニールする2つの一本鎖核酸配列間の関係を指す。例えば、「5'-AGT-3'」は、その相補配列「3'-TCA-5'」との対を形成する。

40

【0109】

「～のポリヌクレオチド配列を含む(持つ)組成物」または「～のアミノ酸配列を含む(持つ)組成物」は、広い意味で、指定のポリヌクレオチド配列若しくはアミノ酸配列を持つ、任意の組成物を指す。この組成物には、乾燥製剤または水溶液が含まれ得る。NAAPをコード、若しくはNAAPの断片をコードするポリヌクレオチド配列を持つ組成物は、ハイブリダイゼーションプローブとして使用され得る。このプローブは、凍結乾燥状態で保存可能であり、糖質などの安定化剤と結合させることも可能である。ハイブリダイゼーションにおいては、塩(例えばNaCl)、界面活性剤(例えばドデシル硫酸ナトリウム; SDS)およびその他の成分(例えばデンハート液、粉乳、サケの精子のDNAなど)を含む水溶液中

50

にプローブを分散させることができる。

【0110】

「コンセンサス配列」は、不要な塩基を除くためにDNA配列解析を繰り返し行い、XL-PCRキット (Applied Biosystems, Foster City CA) を用いて 5' および / または 3' の方向に伸長され、再度シーケンシングされた核酸配列、または GELVIEW 断片構築システム (GCG, Madison, WI) または Phrap (University of Washington, Seattle WA) などの断片構築用コンピュータプログラムを用いて 1 つあるいはそれ以上のオーバーラップする cDNA や EST、またはゲノム DNA 断片から構築された核酸配列を指す。伸長および構築の両方を行ってコンセンサス配列を産生する配列もある。

【0111】

用語「保存的なアミノ酸置換」は、元のタンパク質の特性を殆ど変えないと予測される置換を指す。すなわち、そのタンパク質の構造、特にその機能が保存され、置換によって大きくは変わらない。下表は、タンパク質内で元のアミノ酸と置換され得るアミノ酸であり、保存的なアミノ酸置換と認められるアミノ酸を示す。

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

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

【0112】

用語「欠失」は、1 個以上のアミノ酸残基が欠如するアミノ酸配列内の变化、あるいは 1 個以上のヌクレオチドが欠如するヌクレオチド配列内の变化を指す。

【0113】

用語「誘導体」は、化学修飾されたポリヌクレオチドまたはポリペプチドを指す。例えば、アルキル基、アシル基、ヒドロキシル基またはアミノ基による水素の置換は、ポリヌクレオチドの化学修飾に含まれ得る。ポリヌクレオチド誘導体は、天然分子の生物学的または免疫学的機能を少なくとも 1 つは保持しているポリペプチドをコードする。ポリペプチド誘導体は、次のようなプロセスによって修飾されたポリペプチドである。すなわち、グリコシル化、ポリエチレングリコール化 (pegylation)、あるいは任意の同様なプロセスであって、誘導元のポリペプチドの少なくとも 1 つの生物学的若しくは免疫学的機能を保持するプロセスである。

【0114】

10

20

30

40

50

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

【0115】

「示差発現」は、少なくとも2例のサンプルを比較して判定する、増加（上方調節）、あるいは減少（下方調節）、または欠損した、遺伝子またはタンパク質の発現を指す。このような比較は例えば、処理済サンプルと不処理サンプル、または病態サンプルと健常サンプルとの間で行われ得る。

【0116】

「エキソンシャッフリング」は、異なるコード領域（エキソン）群の組換えを意味する。或るエキソンは、コードするタンパク質の1つの構造的または機能的ドメインを提示し得るため、安定したサブストラクチャー群の新たな組み合わせによって新規のタンパク質群を構築することが可能であり、新たなタンパク質機能の進化を促進できる。

10

【0117】

用語「断片」は、NAAPの又はNAAPをコードするポリヌクレオチドの固有の部分であって、その親配列（parent sequence）と配列は同一であるが親配列より長さが短いものを指す。断片は、定義された配列の全長から1ヌクレオチド/アミノ酸残基を差し引いた長さよりも短い長さを有し得る。例えば或る断片は、5～1000の連続したヌクレオチドまたはアミノ酸残基を有し得る。プローブ、プライマー、抗原、治療用分子として、あるいはその他の目的のために用いられる或る断片は、少なくとも長さ5、10、15、20、25、30、40、50、60、75、100、150、250若しくは500の、連続したヌクレオチドあるいはアミノ酸残基であり得る。断片は、或る分子の特定領域から優先的に選択し得る。例えば、或るポリペプチド断片は、定義された或る配列内に見られるような或るポリペプチドの最初の250または500アミノ酸（または最初の25%または50%）から選択した、或る長さの連続したアミノ酸を持ち得る。これらの長さは明らかに例として挙げているものであり、本発明の実施態様では、配列表、表および図面を含む本明細書が支持する任意の長さであり得る。

20

【0118】

SEQ ID NO:17-32の或る断片は、或る固有のポリヌクレオチド配列領域を持つ。この領域は、SEQ ID NO:17-32を特異的に同定するものであり、例えばこの断片を得たゲノム中の他のいかなる配列とも異なる。SEQ ID NO:17-32の或る断片は、例えば、ハイブリダイゼーションや増幅技術に、またはSEQ ID NO:17-32を関連ポリヌクレオチド配列から区別する類似の方法に有用である。SEQ ID NO:17-32の或る断片の正確な長さは、また、その断片に対応するSEQ ID NO:17-32の領域は、その断片に意図した目的に基づき当業者が慣例的に決定し得る。

30

【0119】

SEQ ID NO:1-16の或る断片は、SEQ ID NO:17-32の或る断片によってコードされる。SEQ ID NO:1-16の或る断片は、SEQ ID NO:1-16を特異的に同定する固有のアミノ酸配列の領域を含む。例えば、SEQ ID NO:1-16の或る断片は、SEQ ID NO:1-16を特異的に認識する抗体の開発における免疫原性ペプチドとして有用である。SEQ ID NO:1-16の或る断片の正確な長さは、また、この断片に対応するSEQ ID NO:1-16の領域は、その断片の目的に基づいて当業者が慣例的に判定し得る。

40

【0120】

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

【0121】

「相同性」の語は、2つ以上のポリヌクレオチド配列または2つ以上のポリペプチド配列の、配列類似性、互換性、または配列同一性を意味する。

【0122】

50

ポリヌクレオチド配列に適用される「一致率」または「%一致」の語は、標準化されたアルゴリズムを用いてアラインメントされた、2つ以上のポリヌクレオチド配列間で一致する残基の割合を意味する。標準化アルゴリズムは、2配列間のアラインメントを最適化するため、標準化された再現性のある方法で比較対象の2配列内にギャップ群を挿入し得るので、2つの配列をより有意に比較できる。

【0123】

ポリヌクレオチド配列間の一致率を判定するには、MEGALIGN version 3.12e配列アラインメントプログラムなどに組込まれているCLUSTAL Vアルゴリズムの、デフォルトのパラメータ群を用い得る。このプログラムは、LASERGENE ソフトウェアパッケージ（一組の分子生物学的分析プログラム）（DNASTAR, Madison WI）の一部である。このCLUSTAL Vは、Higgins, D.G.およびP.M. Sharp (1989) CABIOS 5:151-153、Higgins, D.G. 他 (1992) CABIOS 8:189-191に記載されている。ポリヌクレオチド配列をペアワイズアラインメントする際のデフォルトパラメータは、Ktuple=2、gap penalty=5、window=4、「diagonals saved」=4と設定される。「weighted」残基重み付け表がデフォルトで選択される。CLUSTAL Vによる一致率の報告は、アラインメントされたポリヌクレオチド配列間の「percent similarity (類似性パーセント)」としてなされる。

10

【0124】

あるいは、米国国立バイオテクノロジー情報センター（NCBI）のBasic Local Alignment Search Tool（BLAST）が、一般的に用いられ、且つ、無料で利用可能な配列比較アルゴリズム一式を提供している（Altschul, S.F. 他 (1990) J. Mol. Biol. 215:403-410）。BLASTアルゴリズムは、幾つかの情報源から入手可能であり、メリーランド州ベセスダにあるNCBIおよびインターネット（<http://www.ncbi.nlm.nih.gov/BLAST/>）からも入手可能である。このBLASTソフトウェア一式には、既知のポリヌクレオチド配列と、様々なデータベースの別のポリヌクレオチド配列とのアラインメントに用いる「blastn」を含む、様々な配列分析プログラムが含まれる。「BLAST 2 Sequences」と呼ばれるツールも入手可能であり、これは2つのヌクレオチド配列を直接のペアワイズで比較するために用いられる。「BLAST 2 Sequences」は、<http://www.ncbi.nlm.nih.gov/gorf/bl2.html>にアクセスして、対話形式で利用できる。「BLAST 2 Sequences」ツールは、blastn および blastp（以下に記載）の両方に用い得る。BLASTプログラムは、一般的には、ギャップ（gap）などのパラメータをデフォルト設定にセットして用いる。例えば、2つのヌクレオチド配列を比較するには、デフォルトパラメータに設定した「BLAST 2 Sequences」ツールVersion 2.0.12（2000年4月21日）を用いてblastnを実行し得る。デフォルトパラメータの設定例を以下に示す。

20

30

【0125】

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

40

一致率は、ある定義された配列の全長（例えば特定のSEQ IDナンバーで定義された配列）で測定し得る。あるいは、より短い長さ、例えば、定義された、より大きな配列から得た断片（例えば少なくとも20、30、40、50、70、100または少なくとも200の連続したヌクレオチドの断片）の長さの一致率も測定し得る。ここに挙げた長さは単なる例示的なものに過ぎず、表、図および配列リストを含めた本明細書に記載された配列が支持する任意の断片長を用いて、一致率を測定し得る或る長さを説明し得ることを理解されたい。

【0126】

50

高度の同一性を示さない核酸配列が、それにもかかわらず遺伝子コードの縮重が原因で、類似のアミノ酸配列をコードする場合がある。この縮重を利用して核酸配列内で変化を生じさせて、全ての核酸配列が実質上同一のタンパク質をコードするような多数の核酸配列を生成し得ることを理解されたい。

【0127】

ポリペプチド配列に用いられる用語「一致率」または「%一致」とは、標準化されたアルゴリズムを用いてアラインメントされる2つ以上のポリペプチド配列間の一致する残基の百分率のことである。ポリペプチド配列アラインメントの方法は公知である。保存的アミノ酸置換を考慮するアラインメント方法もある。既に詳述したこのような保存的置換は通常、置換部位の電荷および疎水性を保存するので、ポリペプチドの構造を（したがって機能も）保存する。

10

【0128】

ポリペプチド配列間の一致率は、MEGALIGN version 3.12e配列アラインメントプログラムに組み込まれているようなCLUSTAL Vアルゴリズムのデフォルトのパラメータ群を用いて判定できる（参照先と共に上述）。CLUSTAL Vを用いて、ポリペプチド配列をペアワイズアラインメントする際のデフォルトパラメータは、Ktuple=1、gap penalty=3、window=5、「diagonals saved」=5と設定される。PAM250マトリクスが、デフォルトの残基重み付け表として選択される。ポリヌクレオチドのアラインメントと同様に、アラインメントされたポリペプチド配列の対の一致率は、CLUSTAL Vによって「percent similarity（類似性パーセント）」として報告される。

20

【0129】

あるいは、NCBI BLASTソフトウェア一式を用い得る。例えば、2つのポリペプチド配列をペアワイズで比較する場合、「BLAST 2 Sequences」ツールVersion 2.0.12（2000年4月21日）のblastpをデフォルトパラメータに設定して用い得る。デフォルトパラメータの設定例を以下に示す。

【0130】

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

30

一致率は、ある定義されたポリペプチド配列の全長（例えば特定のSEQ IDナンバーで定義された配列）で測定し得る。あるいは、より短い長さ、例えば、定義された、より大きなポリペプチド配列から得た断片（例えば少なくとも15、20、30、40、50、70、または少なくとも150の連続した残基の断片）の長さの一致率も測定し得る。ここに挙げた長さは単なる例示的なものに過ぎず、表、図および配列リストを含めた本明細書に記載された配列が支持する任意の断片長を用いて、一致率を測定し得る或る長さを説明し得ることを理解されたい。

【0131】

「ヒト人工染色体（HAC）」は直鎖状の微小染色体であり、約6 kb ~ 10 MbのサイズのDNA配列を持ち得る。また、染色体の複製、分離および維持に必要な、全てのエレメントを持つ。

40

【0132】

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

【0133】

「ハイブリダイゼーション」とは、所定のハイブリダイゼーション条件下で、或る一本鎖ポリヌクレオチドが或る相補的な一本鎖と塩基対を形成するアニーリングのプロセスである。特異的ハイブリダイゼーションは、2つの核酸配列が高い相補性を共有することの指

50

標である。特異的ハイブリダイゼーション複合体は許容されるアニーリング条件下で形成され、1回以上の「洗浄」ステップ後もハイブリダイズされたままである。洗浄ステップは、ハイブリダイゼーションプロセスのストリンジェンシーを決定する際に特に重要であり、よりストリンジェントな条件では、非特異結合（すなわち完全には一致しない核酸鎖対間の結合）が減少する。核酸配列のアニーリングに関する許容条件は、当業者が慣例的に決定できる。アニール条件はどのハイブリダイゼーション実験でも一定であり得るが、洗浄条件は、所望のストリンジェンシー、したがってハイブリダイゼーション特異性を得るように、実験ごとに変更し得る。許容されるアニーリング条件は、例えば、温度が68で、約6×SSC、約1%（w/v）のSDS、並びに約100μg/mlの、せん断して変性したサケ精子DNAの存在下である。

10

【0134】

一般に、ハイブリダイゼーションのストリンジェンシーは或る程度、洗浄ステップを実行する温度を基準にして表すことができる。このような洗浄温度は通常、所定のイオン強度およびpHにおける特定配列の融点（ T_m ）より約5～20低くなるように選択する。この T_m は、所定のイオン強度およびpHの条件下で、完全一致プローブに標的配列の50%がハイブリダイズする温度である。 T_m を計算する式および核酸のハイブリダイゼーション条件はよく知られており、Sambrook, J. 他（1989）Molecular Cloning: A Laboratory Manual, 第2版, 1-3巻, Cold Spring Harbor Press, Plainview NYに記載されており、特に2巻の9章を参照されたい。

【0135】

本発明のポリヌクレオチド間の高ストリンジェンシー条件のハイブリダイゼーションには、約0.2×SSCおよび約0.1%のSDSの存在下、68で1時間の洗浄条件を含む。別法では、温度は約65、60、55、または42を用い得る。SSC濃度は、約0.1%のSDS存在下で、約0.1～2×SSCの範囲で変更し得る。通常は、ブロッキング剤を用いて非特異ハイブリダイゼーションを阻止する。このようなブロッキング剤の例として、約100～200μg/mlの、せん断して変性サケ精子DNAがある。特定条件下で、例えばRNAとDNAのハイブリダイゼーションでは、有機溶剤、例えば約35～50%v/vの濃度のホルムアミドを用いることもできる。洗浄条件の有用なバリエーションは、当業者には自明であろう。ハイブリダイゼーションは、特に高ストリンジェント条件下では、ヌクレオチド間の進化的な類似性を示唆し得る。進化的類似性は、ヌクレオチド群、およびヌクレオチドがコードするポリペプチド群について、或る同様の役割を強く示唆する。

20

30

【0136】

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

【0137】

用語「挿入」あるいは「付加」は、1個以上のアミノ酸残基あるいはヌクレオチドがそれぞれ追加される、アミノ酸配列あるいは核酸配列における変化を指す。

40

【0138】

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

【0139】

用語「免疫原性断片」は、生物（例えば哺乳類）に導入すると免疫応答を引き起こし得る、NAAPのポリペプチド断片またはオリゴペプチド断片を指す。用語「免疫原性断片」はまた、本明細書で開示するまたは当分野で周知のあらゆる抗体生産方法に有用な、NAAPの任

50

意のポリペプチド断片またはオリゴペプチド断片をも指す。

【0140】

用語「マイクロアレイ」は、或る基板上の複数のポリヌクレオチド、ポリペプチドまたはその他の化合物の構成を指す。

【0141】

用語「エレメント」または「アレイエレメント」は、マイクロアレイ上に固有の指定された位置を有する、ポリヌクレオチド、ポリペプチドまたはその他の化合物を指す。

【0142】

用語「モジュレート」または「活性を調節」は、NAAPの活性を変化させることを指す。例えば、モジュレートによって、NAAPのタンパク質活性の、或いは結合特性の、またはその他の生物学的特性、機能的特性或いは免疫学的特性の変化が起き得る。

10

【0143】

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

【0144】

「機能的に連結した」は、第1の核酸配列と第2の核酸配列が機能的な関係にある状態を指す。例えば、或るプロモーターが或るコード配列の転写または発現に影響を及ぼす場合には、そのプロモーターはそのコード配列に機能的に連結している。機能的に連結したDNA配列群は非常に近接するか連続的に隣接することがあり、また、2つのタンパク質コード領域を結合するために必要な場合は同一リーディングフレーム内にあり得る。

20

【0145】

「ペプチド核酸(PNA)」は、末端がリシンで終わるアミノ酸残基のペプチドのバックボーンに結合した、少なくとも約5ヌクレオチドの長さのオリゴヌクレオチドを持つ、アンチセンス分子または抗遺伝子剤を指す。末端のリシンは、この組成に溶解性を与える。PNAは、相補的一本鎖DNAまたはRNAに優先的に結合して転写の伸長を停止させる。ポリエチレングリコール化することにより、細胞におけるPNAの寿命を延長し得る。

【0146】

NAAPの「翻訳後修飾」としては、脂質化、グリコシル化、リン酸化、アセチル化、ラセミ化、蛋白分解的切断、及びその他の当分野で既知の修飾を含み得る。これらのプロセスは、合成的或いは生化学的に生じ得る。生化学的修飾は、NAAPの酵素環境に依存し、細胞の種類によって異なることとなる。

30

【0147】

「プローブ」とは、核酸配列の内、NAAPやそれらの相補配列、またはそれらの断片をコードし、同一配列や対立遺伝子核酸配列、または関連する核酸配列の検出に用いる配列を指す。プローブは、単離されたオリゴヌクレオチドまたはポリヌクレオチドであって、検出可能な標識またはレポーター分子に結合した配列である。典型的な標識には、放射性アイソトープ、リガンド、化学発光試薬および酵素がある。「プライマー」とは、相補的な塩基対を形成して標的ポリヌクレオチドにアニーリング可能な、通常はDNAオリゴヌクレオチドである短い核酸である。プライマーは次に、DNAポリメラーゼ酵素により、標的DNA鎖に沿って伸長され得る。プライマー対は、例えばポリメラーゼ連鎖反応(PCR)による、核酸配列の増幅(および同定)に用い得る。

40

【0148】

本発明に用いるプローブおよびプライマーは通常、既知の配列の、少なくとも15の連続したヌクレオチド群からなる。特異性を高めるため、長めのプローブおよびプライマー、例えば開示した核酸配列の少なくとも20、25、30、40、50、60、70、80、90、100または少なくとも150の連続したヌクレオチドからなるようなプローブおよびプライマーも用い得る。これよりもかなり長いプローブおよびプライマーもある。

50

表、図面および配列リストを含む本明細書が支持する、任意の長さのヌクレオチドを用い得るものと理解されたい。

【0149】

プローブおよびプライマーの調製および使用方法については、Sambrook, J. 他 (1989) Molecular Cloning: A Laboratory Manual, 第2版, 1-3巻, Cold Spring Harbor Press, Plainview NY、Ausubel, F.M. 他 (1987) Current Protocols in Molecular Biology, Greene Publ. Assoc. & Wiley-Intersciences, New York NY、Innis, M. 他 (1990) PCR Protocols, A Guide to Methods and Applications, Academic Press, San Diego CAなどを参照されたい。PCRプライマー対を既知の配列から得るには、例えば、そのためのコンピュータプログラム、例えばPrimer (Version 0.5, 1991, Whitehead Institute for Biomedical Research, Cambridge MA) を用い得る。

10

【0150】

プライマーとして用いるオリゴヌクレオチドの選択は、そのような目的のために本技術分野でよく知られているソフトウェアを用いて行う。例えばOLIGO 4.06ソフトウェアは、各100ヌクレオチドまでのPCRプライマー対の選択に有用であり、オリゴヌクレオチドおよび、最大5,000までの大きめのポリヌクレオチドであって32キロボースまでのインプットポリヌクレオチド配列から得た配列を分析するのにも有用である。類似のプライマー選択プログラムには、拡張能力のための追加機能が組込まれている。例えば、PrimOUプライマー選択プログラム (テキサス州ダラスにあるテキサス大学南西部医療センターのゲノムセンターから一般向けに入手可能) は、メガベース配列から特定のプライマーを選択することが可能であり、したがってゲノム全体の範囲でプライマーを設計するのに有用である。Primer3プライマー選択プログラム (Whitehead Institute/MIT Center for Genome Research (マサチューセッツ州ケンブリッジ) より入手可能) を用いれば、プライマー結合部位として避けたい配列を指定できる「非プライミングライブラリ (mispriming library)」を入力できる。Primer3は特に、マイクロアレイのためのオリゴヌクレオチドの選択に有用である (後二者のプライマー選択プログラムのソースコードは、それぞれの情報源から得てユーザー固有のニーズを満たすように修正し得る)。PrimerGenプログラム (英国ケンブリッジ市の英国ヒトゲノムマッピングプロジェクト-リソースセンターから一般向けに入手可能) は、多数の配列アラインメントに基づいてプライマーを設計し、それによって、アラインメントされた核酸配列の最大保存領域または最小保存領域のいずれかとハイブリダイズするようなプライマーの選択を可能にする。したがって、このプログラムは、独自のものであれ保存されたものであれ、オリゴヌクレオチドとポリヌクレオチド断片との同定に有用である。上記選択方法のいずれかによって同定したオリゴヌクレオチドおよびポリヌクレオチド断片は、ハイブリダイゼーション技術において、例えばPCRまたはシーケンシングプライマーとして、マイクロアレイエレメントとして、あるいは核酸のサンプルにおいて完全または部分的相補的ポリヌクレオチドを同定する特異プローブとして有用である。オリゴヌクレオチドの選択方法は、上記の方法に限定されるものではない。

20

30

【0151】

「組換え核酸」は、天然の配列ではなく、配列の、2つ以上の離れたセグメントを人工的に組み合わせた配列である。この人為的組合せはしばしば化学合成によって達成するが、より一般的には核酸の単離セグメントの人為的操作によって、例えばSambrookの文献 (前出) に記載されているような遺伝子工学的手法によって達成する。組換え核酸の語は、単に核酸の一部が付加、置換または欠失により改変された核酸も含む。しばしば組換え核酸には、プロモーター配列に機能的に連結した核酸配列が含まれる。このような組換え核酸は、例えばある細胞を形質転換するために使用されるベクターの一部と成し得る。

40

【0152】

あるいはこのような組換え核酸は、ウイルスベクターの一部と成すことができ、ベクターは例えばワクシニアウイルスに基づくものであり得る。そのようなベクターは哺乳類に接種され、その組換え核酸が発現されて、その哺乳類内で防御免疫応答を誘導するように使

50

用することができる。

【0153】

「調節エレメント」は、或る遺伝子の非翻訳領域に通常は由来する核酸配列であり、エンハンサー、プロモーター、イントロンおよび5'および3'の非翻訳領域(UTR)を含む。調節エレメントは、転写、翻訳またはRNA安定性を制御する宿主タンパク質またはウイルスタンパク質と相互作用する。

【0154】

「レポーター分子」は、核酸、アミノ酸または抗体の標識に用いる、化学的または生化学的成分である。レポーター分子には、放射性核種、酵素、蛍光剤、化学発光剤、発色剤、基質、補助因子、阻害因子、磁気粒子およびその他の当分野で既知の成分がある。

10

【0155】

或るDNA配列に対する「RNA等価物」は、基準となるDNA配列と同じ直鎖のヌクレオチド配列からなるが、生じる全ての窒素性塩基のチミンがウラシルに置換され、糖鎖のバックボーンがデオキシリボースではなくリボースからなる。

【0156】

用語「サンプル」は、その最も広い意味で用いられている。NAAP、NAAPをコードする核酸群、またはその断片群を含むと推定されるサンプルとしては、体液と、細胞や細胞から単離した染色体や細胞内小器官(オルガネラ)や膜からの抽出物と、細胞と、溶液中に存在するまたは基板に固定されたゲノムDNA、RNA、cDNAと、組織と、組織プリントなどがあり得る。

20

【0157】

用語「特異結合」および「特異的に結合する」は、タンパク質若しくはペプチドの、アゴニスト、抗体、アンタゴニスト、小分子、若しくは任意の天然若しくは合成の結合組成物との間の相互作用を指す。この相互作用は、タンパク質の特定の構造(例えば抗原決定基すなわちエピトープ)であって結合分子が認識する構造の有無に依存する。例えば、或る抗体がエピトープ「A」に対して特異的である場合、結合していない標識した「A」と抗体とを含む或る反応の中に、エピトープAを持つポリペプチドが、あるいは結合していない無標識の「A」が存在すると、抗体と結合する標識Aの量が減少する。

【0158】

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

30

【0159】

「置換」とは、1つ以上のアミノ酸残基またはヌクレオチドを、それぞれ別のアミノ酸残基またはヌクレオチドに置き換えることである。

【0160】

用語「基板」は、任意の好適な固体あるいは半固体の支持物を指し、膜およびフィルター、チップ、スライド、ウエハ、ファイバー、磁気または非磁気ビーズ、ゲル、チューブ、プレート、ポリマー、微小粒子、毛細管が含まれる。基板は、ウェル、溝、ピン、チャンネル、孔など、様々な表面形態を有することができ、基板表面にはポリヌクレオチドやポリペプチドが結合する。

40

【0161】

「転写イメージ」または「発現プロファイル」は、所定条件下での所定時間における、或る特定の細胞タイプまたは組織による、遺伝子発現の集合的パターンを指す。

【0162】

「形質転換」とは、外来性DNAが、或る受容細胞に導入されるプロセスを言う。形質転換は、本技術分野で知られている種々の方法に従って自然条件または人工条件下で生じ得るものであり、外来性の核酸配列を原核宿主細胞または真核宿主細胞に挿入する、任意の既知の方法を基にし得る。形質転換の方法は、形質転換する宿主細胞の種類によって選択す

50

る。限定するものではないが形質転換方法には、ウイルス感染、電気穿孔法（エレクトロポレーション）、熱ショック、リポフェクションおよび微粒子銃を用いる方法がある。「形質転換された細胞」には、導入されたDNAが、自律的に複製するプラスミドとしてあるいは宿主染色体の一部として複製可能である、安定的に形質転換された細胞が含まれる。更に、限られた期間に一過的に導入DNA若しくは導入RNAを発現する細胞も含まれる。

【0163】

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

10

【0164】

或る特定核酸配列の「変異体/変異配列」とは、核酸配列1本の或る長さ全体について、該特定核酸配列に対し少なくとも40%の配列同一性を有する核酸配列として決定された配列である。決定には、デフォルトパラメータに設定した「BLAST 2 Sequences」ツールVersion 2.0.9（1999年5月7日）を用いてblastnを実行する。このような核酸対は、所定の長さに対して、例えば少なくとも50%、60%、70%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%またはそれ以上の配列同一性を示し得る。変異配列は、例えば、「対立遺伝子」変異配列（上述）または「スプライス」変異配列、「種」変異配列、「多型」変異配列と記述され得る。スプライス変異配列は参照分子との顕著な同一性を有し得るが、mRNAプロセッシング中のエキソン群の選択的スプライシングによって通常、より多数またはより少数のポリヌクレオチドを有することになる。対応するポリペプチドは、追加機能ドメイン群を有するか、あるいは参照分子には存在するドメイン群が欠落していることがある。種変異配列は、種によって異なるポリヌクレオチド配列である。結果的に生じるポリペプチドは通常、相互に有意のアミノ酸同一性を持つ。多型性変異配列は、或る種の個体間における、或る特定遺伝子のポリヌクレオチド配列内での変異である。多型性変異配列にはまた、ポリヌクレオチド配列の1つのヌクレオチド塩基が異なる「1塩基多型性」（SNP）も含み得る。SNPの存在は、例えば特定の個体群、病状または病状性向を示し得る。

20

30

【0165】

或る特定ポリペプチド配列の「変異配列」とは、ポリペプチド配列1本の或る長さ全体について、該特定ポリペプチド配列に対し少なくとも40%の配列同一性を有するポリペプチド配列として決定された配列である。決定には、デフォルトパラメータに設定した「BLAST 2 Sequences」ツールVersion 2.0.9（1999年5月7日）を用いてblastpを実行する。このようなポリペプチド対は、そのポリペプチドの一方の所定の長さに対して、例えば少なくとも50%、60%、70%、80%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%またはそれ以上の配列同一性を示し得る。

40

【0166】

（発明）

本発明は、新規のヒト核酸関連タンパク質群（NAAP）および、NAAPをコードするポリヌクレオチド群の発見に基づく。また、これらの組成物を利用した、細胞増殖異常、神経障害、発達障害、自己免疫/炎症性障害、および感染症の、診断、治療、および予防の開発に基づく。

【0167】

50

表 1 は、本発明の完全長ポリヌクレオチド配列およびポリペプチド配列の命名の概略である。各ポリヌクレオチドおよびその対応するポリペプチドは、1つのIncyteプロジェクト識別番号 (IncyteプロジェクトID) に関連する。各ポリペプチド配列は、ポリペプチド配列識別番号 (ポリペプチドSEQ ID NO:) とIncyteポリペプチド配列番号 (IncyteポリペプチドID) によって表示した。各ポリヌクレオチド配列は、ポリヌクレオチド配列識別番号 (ポリヌクレオチドSEQ ID NO:) とIncyteポリヌクレオチドコンセンサス配列番号 (IncyteポリヌクレオチドID) によって表示した。

【0168】

表 2 は、GenBankタンパク質 (genpept) データベースとPROTEOMEデータベースとに対するBLAST分析で同定した、本発明のポリペプチド群に相同な配列群を示す。列 1 および列 2 はそれぞれ、本発明の各ポリペプチドに対するポリペプチド配列識別番号 (ポリペプチドSEQ ID NO:) と、それに対応するIncyteポリペプチド配列番号 (IncyteポリペプチドID) を示す。列 3 は、最も近いGenBank相同体のGenBank識別番号 (Genbank ID NO) を示す。列 4 は、各ポリペプチドとその相同体 1 つ以上との間の一致に関する確率スコアを示す。列 5 は、該当箇所には適当な引用を示すとともにGenBank相同体 1 つ以上の注釈 (annotation) を示し、これらはすべて本明細書では参考文献に含まれる。

【0169】

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

【0170】

表 2 および 3 は共に、本発明の各々のポリペプチドの特性を要約しており、それら特性が、請求の範囲に記載されたポリペプチドが核酸関連タンパク質であることを確立している。

【0171】

例えばSEQ ID NO:2は、アミノ酸残基210から768までが、インフルエンザ菌 (*Haemophilus influenzae*) (Rd) 転写アクセサリ蛋白質 (tex) (GenBank ID g1573555) との38%の同一性を持つことが、Basic Local Alignment Search Tool (BLAST) によって判定された (表 2 参照)。BLAST確率スコアは $7.9e-105$ であり、これは観察されたポリペプチド配列アラインメントが偶然に得られる確率を示している。SEQ ID NO:2はまた、1つのS1 RNA結合ドメインを有するが、これは、隠れマルコフモデル (HMM) を基にした保存されたタンパク質ファミリードメインのPFAMデータベースにおいて、統計的に有意な一致を検索して決定された (表 3 参照)。BLIMPS解析および付加的BLAST解析からのデータは、SEQ ID NO:2 が転写調節因子であることを更に確証する証拠を提供する。

【0172】

別の例においてSEQ ID NO:3は、ゼブラフィッシュのemx2 hemeoprotein (GenBank ID g1089816) に対して92%が同一であると、Basic Local Alignment Search Tool (BLAST) で判定された (表 2 参照)。BLAST確率スコアは $3.2e-126$ であり、これは保存されたポリペプチド配列アラインメントが偶然に得られる確率を示している。SEQ ID NO:3はまた、1つのホメオボックスドメインを有するが、これは、隠れマルコフモデル (HMM) を基にした保存されたタンパク質ファミリードメイン群のPFAMデータベースにおいて、統計的に有意な一致を検索して決定された (表 3 参照)。BLIMPS、MOTIFS、及びPROFILESCAN解析よりのデータは、SEQ ID NO:3がホメオボックス蛋白質である、さらに実証的な証拠を提供する。

10

20

30

40

50

【0173】

別の例においてSEQ ID NO:4は、シロイヌナズナのブロモドメイン (bromodomain) 含有転写因子 (GenBank ID g6850321) に対して38%が同一であることがBLAST分析によって判定され、確率スコアは $4.8e-52$ である。SEQ ID NO:4はまたブロモドメインを有することが、隠れマルコフモデル (HMM) を基にした、保存されたタンパク質ファミリドメイン群のPFAMデータベースにおいて、統計的に有意な一致を検索して決定された。BLIMPS及びPROFILESCAN解析よりのデータは、SEQ ID NO:4 がブロモドメイン蛋白質である、さらに実証的な証拠を提供する。

【0174】

別の例においてSEQ ID NO:5はヒトRETフィンガー蛋白質様3 (GenBank ID g3417319) に対して58%が同一であるとBLAST分析で判定され、確率スコアは $2.1e-83$ である。SEQ ID NO:5はまた1つのRINGフィンガードメイン (転写因子の特徴である) を持つことが、隠れマルコフモデル (HMM) を基にした、保存されたタンパク質ファミリドメイン群のPFAMデータベースにおいて、統計的に有意な一致を検索して判定された。

【0175】

別の例においてSEQ ID NO:6は或るヒト酸性フィンガー蛋白質 (C3HC4タイプZnフィンガー蛋白質に類似する) (GenBank ID g563127) に対して34%が同一であるとBLAST分析で判定され、確率スコアは $1.3e-23$ である。SEQ ID NO:6はまたZnフィンガードメインをもつことが、隠れマルコフモデル (HMM) を基にした、保存されたタンパク質ファミリドメイン群のPFAMデータベースにおいて、統計的に有意な一致を検索して判定された。PROFILESCAN解析およびMOTIFS解析よりのデータは、SEQ ID NO:6 がZnフィンガー蛋白質である、さらに実証的な証拠を提供する。

【0176】

別の例においてSEQ ID NO:7はヒトのCAGH44蛋白質 (GenBank ID g2565057) に対して52%が同一であるとBLAST分析で判定され、確率スコアは $2.7e-46$ である。SEQ ID NO:7はまた1つのフォークヘッド (fork-head) ドメインを持つことが、隠れマルコフモデル (HMM) を基にした、保存されたタンパク質ファミリドメイン群のPFAMデータベースにおいて、統計的に有意な一致を検索して判定された。BLIMPS、MOTIFS及びPROFILESCAN分析から得たデータは、SEQ ID NO:7がフォークヘッドドメインおよびZnフィンガードメイン (これはDNA結合タンパク質の特徴である) を持つことをさらに裏づける証拠を提供する。

【0177】

別の例においてSEQ ID NO:8は、ネズミのOASIS CREB/ATFファミリ転写因子 (GenBank ID g4519621) に対して91%が同一であるとBasic Local Alignment Search Tool (BLAST) で判定された (表2参照)。BLAST確率スコアは $9.8e-251$ であり、これは観察されたポリペプチド配列アラインメントが偶然に得られる確率を示している。SEQ ID NO:8はまた、bZIP転写因子ドメインを有し、これは、隠れマルコフモデル (HMM) を基にした保存されたタンパク質ファミリドメインのPFAMデータベースにおいて、統計的に有意な一致を検索して判定された (表3参照)。BLIMPS及びMOTIFS解析よりのデータは、SEQ ID NO:8が転写因子であることをさらに確認する証拠を提供する。

【0178】

別の例においてSEQ ID NO:11は、或るヒトKruppelタイプZnフィンガー蛋白質 (GenBank ID g4519270) に対して58%が同一であると、Basic Local Alignment Search Tool (BLAST) で判定された (表2参照)。BLAST確率スコアは $2.0e-213$ であり、これは観察されたポリペプチド配列アラインメントが偶然に得られる確率を示している。SEQ ID NO:11はまた、1つのKRABボックスと、1つのC2H2タイプZnフィンガードメインとを持つと、隠れマルコフモデル (HMM) を基にした保存されたタンパク質ファミリドメインのPFAMデータベースにおいて、統計的に有意な一致を検索して判定された (表3参照)。BLIMPS解析およびMOTIFS解析よりのデータは、SEQ ID NO:11 がZnフィンガー蛋白質である、さらに実証的な証拠を提供する。

【0179】

10

20

30

40

50

別の例としてSEQ ID NO:14は、キイロショウジョウバエのプロテアーゼ(GenBank ID g2791289)に対して64%が同一であると、Basic Local Alignment Search Tool (BLAST)で判定された(表2参照)。BLAST確率スコアは0.0であり、これは観測されたポリペプチド配列アラインメントが偶然に得られる確率を示している。SEQ ID NO:14はまた、1つの逆転写酵素(RNA依存性DNAポリメラーゼ)ドメインと、1つの「ファージ」インテグラーゼファミリーシグネチャ配列と、1つのインテグラーゼ コアドメインとを有し、これは、隠れマルコフモデル(HMM)を基にした保存されたタンパク質ファミリードメイン群のPFAMデータベースにおいて、統計的に有意な一致を検索して判定された(表3参照)。BLIMPS解析および付加的BLAST解析からのデータは、SEQ ID NO:14がレトロウイルス関連ポリ蛋白であることを更に確認する証拠を提供する。

10

【0180】

SEQ ID NO:1、SEQ ID NO:9、SEQ ID NO:10、SEQ ID NO:12、SEQ ID NO:13、SEQ ID NO:15 およびSEQ ID NO:16については、同様の方法で分析し、注釈を付けた。SEQ ID NO:1-16の解析のためのアルゴリズムとパラメータとを表7に記述した。

【0181】

表4に示すように、本発明の完全長ポリヌクレオチド配列は、cDNA配列またはゲノムDNA由来のコード(エキソン)配列を用いて、あるいはこれら2種類の配列を任意に組み合わせて構築した。列1は本発明の各ポリヌクレオチドに対するポリヌクレオチド配列識別番号(ポリヌクレオチドSEQ ID NO)および対応するIncyteポリヌクレオチドコンセンサス配列番号(Incyte ID)、および塩基対の各ポリヌクレオチド配列の長さを示している。列2は、本発明の完全長ポリヌクレオチド配列を構築するのに用いたcDNA配列および/またはゲノム配列の、また、例えばSEQ ID NO:17-32を同定するため、或いはSEQ ID NO:17-32と、関連するポリヌクレオチド配列群とを区別するためのハイブリダイゼーション技術または増幅技術に有用なポリヌクレオチド配列の断片の、開始ヌクレオチド(5')位置および終了ヌクレオチド(3')位置を示す。

20

【0182】

表4の列2に記載したポリヌクレオチド断片は、特に例えば組織特異的cDNAライブラリ群やプールしたcDNAライブラリ群に由来するIncyte cDNA群を指す場合もある。或いは列2に記載したポリヌクレオチド断片は、完全長ポリヌクレオチド配列の構築(アセンブリ)に寄与したGenBank cDNAまたはESTを指す場合もある。更に、列2に記載したポリヌクレオチド断片は、ENSEMBL(The Sanger Centre、英国ケンブリッジ)データベースに由来する配列を同定する場合もある(すなわち「ENST」の命名を含む配列)。あるいは、列2に記載したポリヌクレオチド断片は、NCBI RefSeq Nucleotide Sequence Records データベースに由来する場合もあり(すなわち「NM」または「NT」の命名を含む配列)、またNCBI RefSeq Protein Sequence Recordsに由来する場合もある(すなわち「NP」の命名を含む配列)。または列2に記したポリヌクレオチド断片は、「エキソンスティッチング(exon-stitching)」アルゴリズムにより結び合わせたcDNAおよびGenscan予測エキソン群の両方からなる集合を指す場合がある。例えば、ポリヌクレオチド配列の内、FL_XXXXXX_N₁_N₂_YYYYY_N₃_N₄として同定される配列は「スティッチされた」配列であり、その内、XXXXXXは該アルゴリズムが適用される配列のクラスターの識別番号であり、YYYYYは該アルゴリズムが作成する予測の数であり、N_{1,2,3...}がある場合には、解析中に手動で編集された特定のエキソン群を表す(実施例5参照)。または、列2のポリヌクレオチド断片は「エキソンストレッチング(exon-stretching)」アルゴリズムにより結び合わせたエキソンの集合を指す場合もある。例えば、ポリヌクレオチド配列の内、FL_XXXXXX_gAAAAA_gBBBBB_1_Nとして同定される配列は「ストレッチ」配列の識別番号である。ここでXXXXXXはIncyteプロジェクト識別番号、gAAAAAは「エキソンストレッチング」アルゴリズムを適用したヒトゲノム配列のGenBank識別番号、gBBBBBは一番近いGenBankタンパク質相同体のGenBank識別番号、またはNCBI RefSeq識別番号、Nは特定のエキソンを指す(実施例5を参照)。或るRefSeq配列が「エキソンストレッチング」アルゴリズムのためのタンパク質相同体として使用された場合では、「NM」、「NP」、または「NT」によって表されるRefSeq識

30

40

50

別子が、GenBank識別子（すなわち、gBBBBB）の代わりに使用され得る。

【0183】

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

接頭コード	解析タイプやプログラムの例	10
GNN	例えば、GENSCAN (Stanford University, CA, USA) または	
GFG	FGENES (Computer Genomics Group, The Sanger Centre,	
ENST	Cambridge, UK) を用いた、ゲノム配列群からのエキソン予測	20
GBI	手動で編集された、ゲノム配列群の解析	
FL	スティッチまたはストレッチゲノム配列 (<u>実施例5</u> 参照)	
INCY	ゲノムへのEST配列群のマッピングからの、完全長転写物とエキソンとの予測。エキソン群と、生じる転写物とを予測するため、ゲノム位置とEST構成のデータが組み合わされる。	30

【0184】

場合によっては、最終コンセンサポリヌクレオチド配列を確認するために、表4に示すような配列カバレッジと重複するIncyte cDNAカバレッジが得られたが、該当するIncyte cDNA識別番号は示さなかった。 40

【0185】

表5は、Incyte cDNA配列を用いて構築された完全長ポリヌクレオチド配列のための代表的なcDNAライブラリを示している。代表的なcDNAライブラリとはIncyte cDNAライブラリであり、これは、最も頻繁にはIncyte cDNA配列群によって代表されるが、これらは、上記のポリヌクレオチド配列を構築および確認するために用いられた。cDNAライブラリを複製するために用いた組織およびベクターを表5に示し、表6で説明している。

【0186】

本発明には、NAAPの変異体も含まれる。好適なNAAP変異配列は、NAAPの機能的或いは構造 50

的特徴の少なくとも1つを有し、かつ該NAAPアミノ酸配列に対して少なくとも約80%のアミノ酸配列同一性、或いは少なくとも約90%のアミノ酸配列同一性、更には少なくとも約95%のアミノ酸配列同一性を有する配列である。

【0187】

本発明はまた、NAAPをコードするポリヌクレオチドを提供する。特定の実施態様において、本発明は、NAAPをコードする、SEQ ID NO:17-32からなる一群から選択された1配列を持つポリヌクレオチド配列を提供する。SEQ ID NO:17-32のポリヌクレオチド配列には、配列表に示したように等価RNA配列をも含むが、そこでは窒素塩基チミンの出現はウラシルに置換され、糖のバックボーンはデオキシリボースではなくリボースで構成される。

【0188】

本発明にはまた、NAAPをコードするポリヌクレオチド配列の変異配列を含む。詳細には、このような変異体ポリヌクレオチド配列は、NAAPをコードするポリヌクレオチド配列との、少なくとも約70%のポリヌクレオチド配列同一性、あるいは少なくとも約85%のポリヌクレオチド配列同一性、更には少なくとも約95%ものポリヌクレオチド配列同一性を持つこととなる。本発明の或る実施態様では、SEQ ID NO:17-32からなる群から選択されたアミノ酸配列と少なくとも約70%、或いは少なくとも約85%、または少なくとも約95%もの一致率を有するようなSEQ ID NO:17-32からなる群から選択された配列を有するポリヌクレオチド配列の変異配列を含む。上記の任意のポリヌクレオチドの変異体は、NAAPの機能的若しくは構造的特徴を少なくとも1つ有するアミノ酸配列をコードし得る。

【0189】

更に別の例では、本発明の或るポリヌクレオチド変異配列は、NAAPをコードするポリヌクレオチド配列のspray変異配列である。或るspray変異配列はNAAPをコードするポリヌクレオチド配列との顕著な配列同一性を持つ部分複数を有し得るが、mRNAプロセッシング中のエキソン群の選択的sprayingによって生ずる、配列の数ブロックの付加または欠失により、通常、より多数またはより少数のポリヌクレオチドを有することになる。或るspray変異配列には、約70%未満、または約60%未満、あるいは約50%未満のポリヌクレオチド配列同一性が、NAAPをコードするポリヌクレオチド配列との間で全長に渡って見られるが、このspray変異配列のいくつかの部分には、NAAPをコードするポリヌクレオチド配列の各部との、少なくとも約70%、あるいは少なくとも約85%、または少なくとも約95%、なおまたは100%の、ポリヌクレオチド配列同一性を有することとなる。例えば、SEQ ID NO:29の配列を持つ或るポリヌクレオチドは、SEQ ID NO:25の配列を持つポリヌクレオチドのspray変異体である。上記のspray変異配列は何れも、NAAPの機能的或いは構造的特徴の少なくとも1つを有するアミノ酸配列をコードすることができる。

【0190】

遺伝暗号の縮重により、NAAPをコードする種々のポリヌクレオチド配列が作り出され、中には、既知のいかなる天然遺伝子のポリヌクレオチド配列群とも最小の類似性しか有しない配列もあることは、当業者には理解されよう。したがって本発明には、可能コドン選択に基づく組合せの選択によって産出し得るあらゆる可能なポリヌクレオチド配列のバリエーションを網羅し得る。これらの組み合わせは、天然NAAPのポリヌクレオチド配列に適用される標準的なトリプレット遺伝暗号を基に作られ、全ての変異が明確に開示されていると考慮されたい。

【0191】

NAAPとその変異配列とをコードするヌクレオチド配列は一般に、好適に選択されたストリンジェンシー条件下で天然NAAPのヌクレオチド配列とハイブリダイズ可能であるが、非天然コドン群を含めるなどの実質的に異なるコドン使用を有する、NAAP或いはその誘導体をコードするヌクレオチド配列を作り出すことは、有益であり得る。宿主が特定コドンを利用する頻度に基づいて、特定の真核宿主または原核宿主に発生するペプチドの発現率を高めるようにコドンを選択し得る。コードされるアミノ酸配列を改変せずに、NAAP及びその

10

20

30

40

50

誘導体をコードするヌクレオチド配列を実質的に改変する別の理由には、天然配列から作られる転写物より例えば長い半減期など好ましい特性を備えるRNA転写物を作ることもある。

【0192】

本発明には、また、NAAPとその誘導体群とをコードする、DNA配列群またはその断片群を、完全に合成化学によって作り出す過程をも含む。作製後、当分野で公知の試薬類を用いて、この合成配列を任意の多くの入手可能な発現ベクター及び細胞系中に挿入し得る。更に、合成化学を用いて、NAAPまたはその任意の断片をコードする配列に突然変異を誘発し得る。

【0193】

更に本発明には、種々のストリンジェンシー条件下で、請求項に記載されたポリヌクレオチド配列、特にSEQ ID NO:17-32およびそれらの断片へのハイブリダイズが可能なポリヌクレオチド配列を含む（例えばChiez, R.M.およびF.Z. Regnier (1987) *Methods Enzymol.* 152:399-407; Kimmel, A.R. (1987) *Methods Enzymol.* 152:507-511)。アニーリングおよび洗浄条件を含むハイブリダイゼーションの条件は、「定義」に記載した。

【0194】

DNAシーケンシングの方法は当分野では公知であり、本発明のいずれの実施例も、DNAシーケンシング方法を用い得る。DNAシーケンシング方法には酵素を用い得る。例えばDNAポリメラーゼIのクレノウ断片、SEQUENASE (US Biochemical, Cleveland OH)、Taqポリメラーゼ (Applied Biosystems)、熱安定性T7ポリメラーゼ (Amersham, Pharmacia Biotech, Piscataway NJ) を用い得る。あるいは、例えばELONGASE増幅システム (Life Technologies, Gaithersburg MD) において見られるように、ポリメラーゼと校正エキソヌクレアーゼとを併用し得る。好適には、MICROLAB2200液体転移システム (Hamilton, Reno, NV)、PTC200サーマルサイクラー (MJ Research, Watertown MA) およびABI CATALYST 800サーマルサイクラー (Applied Biosystems) などの装置を用いて配列の準備を自動化する。次に、ABI 373あるいは377 DNAシーケンシングシステム (Applied Biosystems)、MEGABACE 1000 DNAシーケンシングシステム (Molecular Dynamics, Sunnyvale CA) または当分野で公知の他のシステムを用いてシーケンシングを行う。結果として得た配列を、当分野で周知の種々のアルゴリズムを用いて分析する（例えば、Ausubel, F.M. (1997) *Short Protocols in Molecular Biology*, John Wiley & Sons, New York NY, 7.7ユニット、Meyers, R.A. (1995) *Molecular Biology and Biotechnology*, Wiley VCH, New York NY, 856-853ページを参照）。

【0195】

当分野で周知の、PCR法をベースにした種々の方法と、部分的ヌクレオチド配列とを利用して、NAAPをコードする核酸配列を伸長し、プロモーターや調節エレメントなど、上流にある配列を検出し得る。例えば、使用し得る方法の1つである制限部位PCR法は、ユニバーサルプライマーおよびネステッドプライマーを用いてクローニングベクター内のゲノムDNAからの未知の配列を増幅する方法である（例えば、Sarkar, G. (1993) *PCR Methods Appl. 2:318-322*を参照）。別の方法に逆PCR法があり、これは多岐の方向に伸長するプライマー群を用いて、環状化した鋳型から未知の配列を増幅する方法である。鋳型は、或る既知のゲノム遺伝子座およびその周辺の配列群からなる制限酵素断片群から得る（例えば Triglia, T. 他(1988) *Nucleic Acids Res.* 16:8186を参照）。第3の方法としてキャプチャPCR法があり、これにはヒトおよび酵母人工染色体DNAの既知の配列群に隣接するDNA断片群をPCR増幅する方法を含む（例えば Lagerstrom, M. 他(1991) *PCR Methods Appl. 1:111-119*を参照）。この方法では、PCRを行う前に、複数の制限酵素の消化およびライゲーション反応を用い、未知の配列領域内に組換え二本鎖配列を挿入し得る。未知の配列群を検索するために用い得る他の複数の方法も当分野で公知である（例えば Parker, J.D. 他(1991) *Nucleic Acids Res.* 19:30553060を参照）。更に、PCR、ネステッドプライマー類、およびPROMOTERFINDERライブラリ類 (Clontech, Palo Alto CA) を用いてゲノムDNAをウォーキングし得る。この手順は、ライブラリ類をスクリーニングする必要がなく、イ

10

20

30

40

50

ントロン/エキソン接合部の発見に有用である。全てのPCRベースの方法では、市販ソフトウェア、例えばOLIGO 4.06プライマー分析ソフトウェア (National Biosciences, Plymouth MN) 或いは別の好適なプログラムを用いて、長さが約22~30ヌクレオチド、GC含有率が約50%以上で、温度約68~72で鋳型に対してアニーリングするように、プライマー群を設計し得る。

【0196】

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

10

【0197】

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

20

【0198】

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

【0199】

種々の目的で、NAAPをコードする配列群を改変するために、当分野で一般的に既知の複数の方法を用いて、本発明のヌクレオチド配列群を組換えることができる。この組換えの多様な目的には、遺伝子産物のクローニングの、あるいはプロセッシングおよび/または発現のモディフィケーションが含まれるが、これらに限定されるものではない。遺伝子断片と合成オリゴヌクレオチドとの、ランダムなフラグメンテーションおよびPCR再アセンブリによるDNAシャッフリングを用い、ヌクレオチド配列を組み換え得る。例えば、オリゴヌクレオチドを介した部位特異的変異誘導を利用して、新規な制限部位の作製、グリコシル化パターンの改変、コドン優先の変更、スプライス変異配列の生成などを起こす突然変異を導入し得る。

30

【0200】

本発明のヌクレオチドは、MOLECULARBREEDING (Maxygen Inc., Santa Clara CA; 米国特許第5,837,458号; Chang, C.-C. 他 (1999) Nat. Biotechnol. 17:793-797; Christians, F.C. 他 (1999) Nat. Biotechnol. 17:259-264; Cramer, A. 他 (1996) Nat. Biotechnol. 14:315-319)などのDNAシャッフリング技術の対象となり得る。これにより、生物活性または酵素活性、あるいは他の分子や化合物と結合する能力など、NAAPの生物学的特性を改変あるいは改良し得る。DNAシャッフリングは、PCRを介する遺伝子断片組換えを用いて遺伝子変異配列のライブラリを作製するプロセスである。ライブラリはその後、所望の特性を持つ遺伝子変異配列群を同定する、選択またはスクリーニングの手順を経る。続いて、これら好適な変異配列をプールし、更に反復してDNAシャッフリングおよび選択/スクリーニングを行い得る。かくして、「人工的な」育種および急速な分子の進化によって、多様な遺伝子が作られる。例えば、ランダムポイント突然変異を持つ単一の遺伝子の断片を組み換えて、スクリーニングし、その後所望の特性が最適化されるまでシャッフリング

40

50

し得る。あるいは、所定の遺伝子の断片群を同種または異種のいずれかから得た同一遺伝子ファミリーの相同遺伝子の断片と組み換え、それによって天然に存在する複数の遺伝子の遺伝多様性を、方向付けした制御可能な方法で最大化させることができる。

【0201】

別の実施態様によれば、NAAPをコードする配列は、当分野で周知の化学的方法を用いて、全体あるいは一部が合成可能である（例えば、Caruthers, M.H.他(1980) Nucl. Acids Res. Symp. Ser 7:215-223; 及びHorn, T.他(1980) Nucl. Acids Res. Symp. Ser.7:225-232を参照）。別法として、化学的方法を用いてNAAP自体またはその断片を合成することが可能である。例えば、種々の液相または固相技術を用いてペプチド合成を行い得る（例えばCreighton, T. (1984) Proteins, Structures and Molecular Properties, WH Freeman, New York NY,55-60;およびRoberge, J.Y. 他(1995) Science 269:202-204等を参照）。自動合成はABI 431Aペプチドシンセサイザ（Applied Biosystems）を用いて達成し得る。更に、NAAPのアミノ酸配列または任意のその一部を、直接合成の際に改変することにより、及び/または他のタンパク質からの配列群または任意のその一部と組み合わせることにより、或る天然ポリペプチドの配列を有するポリペプチドまたは変異体ポリペプチドが作製され得る。

【0202】

このペプチドは、分離用高速液体クロマトグラフィーを用いて実質的に精製し得る（例えばChiez, R.M.およびF.Z. Regnier (1990) Methods Enzymol. 182:392-421を参照）。この合成ペプチドの組成は、アミノ酸分析またはシーケンシングによって確認できる（例えば前出のCreighton, 28-53ページを参照）。

【0203】

生物学的に活性なNAAPを発現させるために、NAAPをコードするヌクレオチド配列またはその誘導体を好適な発現ベクターに挿入し得る。好適な発現ベクターとは、好適な宿主に挿入されたコーディング配列の転写および翻訳の制御に必要なエレメント群を持つベクターである。必要なエレメントとしては、該ベクターと、NAAPをコードするポリヌクレオチド配列群における調節配列群（エンハンサー、構成型及び発現誘導型のプロモーター、5'及び3'の非翻訳領域など）が含まれる。必要なエレメント群は、特長および特異性が様々である。特定の開始シグナル類を用いて、NAAPをコードする配列群の、より効果的な翻訳を達成できる。開始シグナルの例には、ATG開始コドンと、コザック配列など近傍の配列とが含まれる。NAAPをコードする配列群、その開始コドン、および上流の調節配列群が、好適な発現ベクターに挿入された場合は、更なる転写制御シグナルや翻訳制御シグナルは必要ないこともある。しかしながら、コーディング配列あるいはその断片のみが挿入された場合は、インフレームATG開始コドンなど外来性の翻訳制御シグナルが発現ベクターに含まれるようにすべきである。外来性の翻訳エレメント群および開始コドン群は、様々な天然物および合成物を起源とし得る。用いられる特定の宿主細胞系に好適なエンハンサーを含めることで発現の効率を高めることが可能である（例えば Scharf, D. 他(1994) Results Probl.Cell Differ. 20:125162を参照）。

【0204】

当業者に周知の方法を用いて、NAAPをコードする配列と、好適な転写および翻訳制御エレメントとを持つ発現ベクターを作製し得る。これらの方法としては、in vitro組換えDNA技術、合成技術、およびin vivo遺伝子組換え技術がある（例えば Sambrook, J. 他(1989) Molecular Cloning, A Laboratory Manual, Cold Spring Harbor Press, Plainview NY, 4, 8,および16-17章; Ausubel, F.M. 他(1995) Current Protocols in Molecular Biology, John Wiley & Sons, New York NY, 9, 13, および 16章を参照）。

【0205】

種々の発現ベクター/宿主系を利用して、NAAPをコードする配列群の保持及び発現ができる。限定するものではないがこのような発現ベクター/宿主系には、組換えバクテリオファージ、プラスミドまたはコスミドDNA発現ベクターで形質転換させた細菌や、酵母菌発現ベクターで形質転換させた酵母菌など微生物や、ウイルス発現ベクター（例えばバキュ

ロウイルス)に感染した昆虫細胞系や、ウイルス発現ベクター(例えばカリフラワーモザイクウイルス、CaMVまたはタバコモザイクウイルス、TMV)または細菌発現ベクター(例えばTiプラスミドまたはpBR322プラスミド)で形質転換させた植物細胞系、動物細胞系がある(例えば前出 Sambrook; 前出 Ausubel; Van Heeke, G.およびS.M. Schuster (1989) J. Biol. Chem. 264:55035509; Engelhard, E.K. 他(1994) Proc. Natl. Acad. Sci. USA 91:3224-3227; Sandig, V. 他(1996) Hum. Gene Ther. 7:1937-1945; Takamatsu, N. (1987) EMBO J. 6:307311; 『マグローヒル科学技術年鑑』(The McGraw Hill Yearbook of Science and Technology) (1992) McGraw Hill New York NY, 191-196; Logan, J.およびT. Shank (1984) Proc. Natl. Acad. Sci. USA 81:36553659; およびHarrington, J.J. 他(1997) Nat. Genet. 15:345-355を参照)。レトロウイルス、アデノウイルス、ヘルペスウイルスまたはワクシニアウイルス由来の発現ベクター、または種々の細菌性プラスミド由来の発現ベクターを用いて、ヌクレオチド配列を標的器官、組織または細胞集団へ送達することができる(例えばDi Nicola, M. 他(1998) Cancer Gen. Ther. 5(6):350-356; Yu, M. 他(1993) Proc. Natl. Acad. Sci. USA 90(13):6340-6344; Buller, R.M. 他(1985) Nature 317(6040):813-815; McGregor, D.P. 他(1994) Mol. Immunol. 31(3):219-226; およびVerma, I.M.およびN. Somia (1997) Nature 389:239-242を参照)。本発明は使用される宿主細胞によって限定されるものではない。

10

【0206】

細菌系では、多数のクローニングベクターおよび発現ベクターが、NAAPをコードするポリヌクレオチド配列の使用目的に応じて選択可能である。例えば、NAAPをコードするポリヌクレオチド配列の慣例的なクローニング、サブクローニング、増殖には、PBLUESCRIPT (Stratagene, La Jolla CA) またはPSPORT1 プラスミド (Life Technologies) などの多機能の大腸菌ベクターを用いることができる。ベクターのマルチクローニング部位にNAAPをコードする配列をライゲーションするとlacZ 遺伝子が破壊され、組換え分子を含む形質転換された細菌の同定のための比色スクリーニング法が可能となる。更にこれらのベクターは、クローニングされた配列における *in vitro* 転写、ジデオキシのシーケンシング、ヘルパーファージによる一本鎖のレスキュー、入れ子欠失の生成にも有用であろう(例えば、Van Heeke, G.およびS.M. Schuster (1989) J. Biol. Chem. 264:55035509を参照)。例えば、抗体の産生のためなどに多量のNAAPが必要な場合は、NAAPの発現をハイレベルで誘導するベクターが使用できる。例えば、強力な誘導SP6バクテリオファージプロモーターまたは誘導T7バクテリオファージプロモーターを持つベクターが使用できる。

20

30

【0207】

酵母の発現系を使用してNAAPを産出することもできる。因子、アルコールオキシダーゼ、PGHプロモーターなど、構成型あるいは誘導型のプロモーターを持つ多数のベクターが、出芽酵母菌 (*Saccharomyces cerevisiae*) またはピキア酵母 (*Pichia pastoris*) 内で使用可能である。更に、このようなベクターは、発現したタンパク質の分泌か細胞内での保持かのどちらかを指示し、安定した増殖のために宿主ゲノムの中に外来配列を組み込むことを可能にする(例えば前出 Ausubel, 1995; Bitter, G.A. 他(1987) Methods Enzymol. 153:516544; およびScorer, C.A. 他(1994) Bio/Technology 12:181-184を参照)。

【0208】

植物系もNAAPの発現に使用可能である。NAAPをコードする配列の転写は、ウイルスプロモーター、例えば単独或いはTMV (Takamatsu, N. (1987) EMBO J 6:307-311) 由来のオメガリーダー配列と組み合わせるようなCaMV由来の35S及び19Sプロモーターによって促進される。あるいは、RuBisCOの小サブユニットなどの植物プロモーター、または熱ショックプロモーターを用い得る(例えばCoruzzi, G. 他(1984) EMBO J. 3:16711680; Bröglie, R. 他(1984) Science 224:838843; およびWinter, J. 他(1991) Results Probl. Cell Differ. 17:85105を参照)。これらの構成物は、直接DNA形質転換により、または病原体を媒介とする形質移入により、植物細胞内に導入し得る(例えば『マグローヒル科学技術年鑑』(The McGraw Hill Yearbook of Science and Technology) (1992) McGraw Hill New York NY, 191-196ページを参照)。

40

50

【0209】

哺乳類細胞においては、多数のウイルスベースの発現系を利用し得る。アデノウイルスが発現ベクターとして用いられる場合、後期プロモーターと3連リーダー配列とからなるアデノウイルス転写/翻訳複合体に、NAAPをコードする配列群をライゲーションし得る。アデノウイルスゲノムの非必須E1またはE3領域へ挿入することにより、宿主細胞内でNAAPを発現する感染ウイルスを得ることができる(例えばLogan, J.およびT. Shenk (1984) Proc. Natl. Acad. Sci. USA 81:36553659を参照)。更に、ラウス肉腫ウイルス(RSV)エンハンサーなどの転写エンハンサーを用いて、哺乳動物宿主細胞における発現を増大させ得る。SV40またはEBVをベースにしたベクターを用いてタンパク質を高レベルで発現させることもできる。

10

【0210】

また、ヒト人工染色体(HAC)類を用いて、或るプラスミドに含まれ得る断片やプラスミドから発現し得る断片より大きなDNAの断片群を送達し得る。治療のために約6 kb~10 MbのHACが作製され、従来の送達方法(リポソーム、ポリカチオンアミノポリマー、またはベシクル)で送達されている(例えばHarrington, J.J. 他(1997) Nat. Genet. 15:345-355を参照)。

【0211】

哺乳動物系内の組換えタンパク質の長期にわたる産生のためには、細胞株におけるNAAPの安定した発現が望ましい。例えば、NAAPをコードする配列を細胞株に形質転換するために、発現ベクター類と、同じベクター上の或いは別のベクター上の選択可能マーカー遺伝子とを用い得る。用いる発現ベクターは、ウイルス起源の複製要素、および/または内因性の発現エレメント群を持ち得る。ベクターの導入後、選択培地に移す前に強化培地で約1~2日間、細胞を増殖させ得る。選択可能マーカーの目的は選択剤への抵抗性を与えることであり、選択可能マーカーの存在により、導入した配列をうまく発現するような細胞の成長および回収が可能となる。安定的に形質転換された細胞の耐性クローンは、その細胞型に適した組織培養技術を用いて増殖可能である。

20

【0212】

任意の数の選択系を用いて、形質転換細胞株を回収できる。限定するものではないがこのような選択系には、 tk^{-} 細胞のために用いられる単純ヘルペスウイルスのチミジンキナーゼ遺伝子と、 apr^{-} 細胞のために用いられる単純ヘルペスウイルスのアデニンホスホリボシルトランスフェラーゼ遺伝子がある(例えばWigler, M. 他(1977) Cell 11:223-232; Lowy, I. 他(1980) Cell 22:817-823を参照)。また、選択の基礎として代謝拮抗物質、抗生物質あるいは除草剤への耐性を用いることができる。例えばdhfrはメトトレキセートに対する耐性を与え、neoはアミノグリコシッドネオマイシンおよびG-418に対する耐性を与え、alsはクロルスルフロンに対する耐性を、patはホスフィノトリシンアセチルトランスフェラーゼに対する耐性を各々与える(例えばWigler, M. 他(1980) Proc. Natl. Acad. Sci. USA 77:35673570; ColbereGarapin, F. 他(1981) J. Mol. Biol. 150:114を参照)。その他の選択可能な遺伝子、例えば、代謝のための、細胞の必要条件を改変するtrpBおよびhisDも、文献に記載がある(例えばHartman, S.C.およびR.C. Mulligan (1988) Proc. Natl. Acad. Sci. USA 85:80478051を参照)。可視マーカー類、例えばアントシアニンや、

30

40

【0213】

マーカー遺伝子発現の有無によって目的の遺伝子の存在が示唆されても、その遺伝子の存在および発現の確認が必要な場合もある。例えば、NAAPをコードする配列がマーカー遺伝子配列の中に挿入された場合、NAAPをコードする配列を持つ形質転換された細胞は、マーカー遺伝子機能の欠落により同定可能である。または、単一プロモーターの制御下で、或

50

るマーカー遺伝子を、NAAPをコードする1配列とタンデムに配置することもできる。誘導または選択に反応したマーカー遺伝子の発現は通常、タンデム遺伝子の発現も示す。

【0214】

一般に、NAAPをコードする核酸配列を含み且つNAAPを発現する宿主細胞は、当業者によく知られている種々の方法を用いて同定することが可能である。限定するものではないが当業者によく知られている方法には、DNA-DNA或いはDNA-RNAハイブリダイゼーションと、PCR増幅とがあり、また、核酸配列或いはタンパク質配列の検出、定量、或いはその両方を行うための、膜系、溶液ベース或いはチップベースの技術を含む、タンパク質の生物学的検定法または免疫学的検定法もある。

【0215】

特異的ポリクローナル抗体または特異的モノクローナル抗体を用いてNAAPの発現の検出と計測とを行うための免疫学的方法は、当分野で既知である。このような技術の例としては、酵素結合イムノソルベントアッセイ(ELISA)、ラジオイムノアッセイ(RIA)、蛍光活性化細胞選別(FACS)などが挙げられる。NAAP上の2つの非干渉エピトープに反応するモノクローナル抗体群を用いた、2部位のモノクローナルベースイムノアッセイ(two-site, monoclonal-based immunoassay)が好ましいが、競合結合試験を用いることもできる。これらのアッセイおよび他のアッセイは、当分野で周知である(例えばHampton, R. 他(1990) Serological Methods, a Laboratory Manual, APS Press, St. Paul MN, Sect. IV; Coligan, J.E. 他(1997) Current Protocols in Immunology, Greene Pub. Associates and Wiley-Interscience, New York NY; および Pound, J.D. (1998) Immunochemical Protocols, Humana Press, Totowa NJを参照)。

【0216】

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

【0217】

NAAPをコードするヌクレオチド配列で形質転換された宿主細胞は、細胞培地でのこのタンパク質の発現および回収に好適な条件下で培養される。形質転換細胞から製造されたタンパク質が分泌されるか細胞内に留まるかは、使用される配列、ベクター、あるいはその両者に依存する。NAAPをコードするポリヌクレオチド群を持つ発現ベクター類は、原核細胞膜または真核細胞膜を透過してのDMEの分泌を指示するシグナル配列群を持つように設計できることは、当業者には理解されよう。

【0218】

更に、宿主細胞株の選択は、挿入した配列の発現をモジュレートする能力、または発現したタンパク質を所望の形に処理する能力によって行い得る。限定するものではないがこのようなポリペプチドの修飾には、アセチル化、カルボキシル化、グリコシル化、リン酸化、脂質化およびアシル化がある。タンパク質の「プレプロ」または「プロ」形を切断する翻訳後のプロセッシングを利用して、タンパク質の標的への誘導、折りたたみ、および/または活性を特定することが可能である。翻訳後の活性のための、特定の細胞装置および特

10

20

30

40

50

徴的な機序を持つ、種々の宿主細胞（例えばCHO、HeLa、MDCK、MEK293、WI38など）は、American Type Culture Collection (ATCC, Manassas, VA) から入手可能であり、外来タンパク質の正しい修飾および処理を確実にするように選択し得る。

【0219】

本発明の別の実施態様では、NAAPをコードする、天然核酸配列、修飾された核酸配列、または組換え核酸配列を、或る異種配列に結合させることにより、上記した任意の宿主系内で、或る融合タンパク質の翻訳を生じ得る。例えば、市販の抗体によって認識できる異種部分を含むキメラNAAPタンパク質は、NAAP活性のインヒビターに対するペプチドライブラリのスクリーニングを促進し得る。また、異種タンパク質部分および異種ペプチド部分も、市販されている親和性マトリックスを用いて融合タンパク質の精製を促進し得る。限定されるものではないがこのような部分には、グルタチオンSトランスフェラーゼ (GST)、10 マルトース結合タンパク質 (MBP)、チオレドキシシン (Trx)、カルモジュリン結合ペプチド (CBP)、6-His、FLAG、c-myc、赤血球凝集素 (HA) がある。GSTは固定化グルタチオン上で、MBPはマルトース上で、Trxはフェニルアルシンオキシド上で、CBPはカルモジュリン上で、そして6-Hisは金属キレート樹脂上で、同族の融合タンパク質の精製を可能にする。FLAG、c-mycおよび赤血球凝集素 (HA) は、これらのエピトープ標識を特異的に認識する市販されているモノクローナル抗体およびポリクローナル抗体を用いた、融合タンパク質の免疫親和性精製を可能にする。また、融合タンパク質が、NAAPをコードする配列と異種タンパク質配列との間にタンパク質分解切断部位を持つように遺伝子操作すると、20 精製後にNAAPが異種部分から切断され得る。融合タンパク質の発現および精製方法は、前出のAusubel (1995) 10章に記載されている。市販されている種々のキットを用いて融合タンパク質の発現および精製を促進することもできる。

【0220】

本発明の更なる実施態様では、TNTウサギ網状赤血球可溶化液またはコムギ胚芽抽出系 (Promega) を用いて *in vitro* で、放射能標識したNAAの合成が可能である。これらの系は、T7、T3またはSP6プロモーターと機能的に連結したタンパク質コード配列の、転写と翻訳とを結合する。翻訳は、例えば³⁵Sメチオニンのような放射能標識したアミノ酸前駆体の存在下で起こる。

【0221】

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

【0222】

或る実施態様では、このように同定された化合物は、NAAPの天然リガンドに密接に関連し、例えばリガンドやその断片であり、または天然基質や、構造的または機能的な擬態物質 (mimetic)、あるいは自然結合パートナーである (例えばColigan, J.E. 他 (1991) Current Protocols in Immunology 1(2):Chapter 5を参照)。同様に、この化合物は、NAAPが結合する天然受容体に、或いは例えばリガンド結合部位など少なくともこの受容体の或る断片に、密接に関連し得る。何れの場合も、既知の技術を用いてこの化合物を合理的に設計20 することができる。或る実施例では、これら化合物のためのスクリーニングには、分泌タンパク質として、あるいは細胞膜上でNAAPを発現する、好適な細胞群の作製が含まれる。好適な細胞には、哺乳類、酵母、ショウジョウバエ、または大腸菌からの細胞が含まれる。NAAPを発現する細胞またはNAAPを含有する細胞膜分画を試験化合物と接触させて、NAAPまたはこの化合物のどちらかの結合、刺激、または活性の阻害を分析する。

【0223】

或るアッセイは、単に試験化合物をポリペプチドに実験的に結合させ、蛍光色素、放射性同位体、酵素抱合体またはその他の検出可能な標識によりその結合を検出することができる。例えば、このアッセイは、少なくとも1つの試験化合物を、溶液中の、あるいは固体支持物に固定されたNAAPと混合するステップと、NAAPとこの化合物との結合を検出するス30

10

20

30

40

50

チップを含み得る。別法では、標識された競合物の存在下での試験化合物の結合の検出および測定を行うことができる。更にこのアッセイは、無細胞再構成標本、化学ライブラリ、または、天然産物の混合物を用いて実施でき、試験化合物（群）は、溶液中で遊離させるか固体支持体に固定する。

【0224】

本発明のNAAPまたはその断片を用いて、NAAPの活性をモジュレートする化合物をスクリーニングすることが可能である。このような化合物の例には、アゴニスト、アンタゴニスト、部分的アゴニストまたは逆アゴニストなどが含まれる。或る実施例では、NAAPが少なくとも1つの試験化合物と結合する、NAAPの活性が許容される条件下でアッセイを実施し、試験化合物の存在下でのNAAPの活性と、試験化合物不在下でのNAAPの活性とを比較する。試験化合物の存在下でのNAAPの活性の変化は、NAAPの活性をモジュレートする化合物の存在を示唆する。別法では、試験化合物を、NAAPの活性に適した条件下で、NAAPを含む *in vitro* または細胞遊離系と結合させてアッセイを実施する。これらアッセイの何れにおいても、NAAPの活性を調節する試験化合物は間接的に調節する場合があり、その際は試験化合物と直接接触する必要がない。少なくとも1つ、または複数の試験化合物をスクリーニングすることができる。

10

【0225】

別の実施態様では、胚性幹細胞（ES細胞）における相同組換えを用いて動物モデル系内で、NAAPまたはその哺乳類相同体をコードするポリヌクレオチドを「ロックアウト」する。このような技術は当技術分野において周知であり、ヒト疾患動物モデルの産生に有用である（米国特許第5,175,383号および第5,767,337号などを参照）。例えば129/SvJ細胞系などのマウスES細胞は初期のマウス胚に由来し、培地で増殖させることができる。このES細胞は、ネオマイシンホスホトランスフェラーゼ遺伝子（neo: Capecchi, M.R. (1989) Science 244:1288-1292）などのマーカー遺伝子で破壊した、目的の遺伝子を含むベクターで形質転換する。このベクターは、相同組換えにより、宿主ゲノムの対応する領域に組み込まれる。別法では、Cre-loxP系を用いて相同組換えを行い、組織特異的または発生段階特異的に目的遺伝子をロックアウトする（Marth, J.D. (1996) Clin. Invest. 97:1999-2002; Wagner, K.U. 他(1997) Nucleic Acids Res. 25:4323-4330）。形質転換したES細胞を同定し、例えばC57BL/6マウス株などから採取したマウス細胞胚盤胞に微量注入する。これらの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特定し、これらを交配させてヘテロ接合性系またはホモ接合性系を作製する。このようにして産生した遺伝子組換え動物は、潜在的治療薬や毒性薬剤で検査することができる。

20

30

【0226】

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

【0227】

NAAPをコードするポリヌクレオチドを用いて、ヒト疾患をモデルとした「ロックイン」ヒト化動物（ブタ）または遺伝子組換え動物（マウスまたはラット）を作製することもできる。ロックイン技術を用いて、NAAPをコードするポリヌクレオチドの或る領域を動物ES細胞に注入し、注入した配列を動物細胞ゲノムに組み込ませる。形質転換細胞を胞胚に注入し、胞胚を上記のように移植する。遺伝子組換え子孫または近交系について試験し、潜在的医薬品を用いて処理し、ヒトの疾患の治療に関する情報を得る。別法では、例えばNAAPを乳汁内に分泌するなどLMMを過剰に発現する哺乳動物近交系は、便利なタンパク質源となり得る（Janne, J. 他(1998) Biotechnol. Annu. Rev. 4:55-74）。

40

【0228】

（治療）

NAAPの領域と核酸関連タンパク質の領域との間には、例えば配列及びモチーフの文脈にお

50

ける、化学的及び構造的類似性が存在する。また、NAAPを発現する組織は、前立腺腫瘍および肺腫瘍の組織と密接に関連し、また、臍帯血および臍帯血樹状細胞、および脳組織と密接に関連する。これらの組織の例は、表6を見られたい。このように、NAAPは、細胞増殖異常、神経障害、発達障害、自己免疫/炎症疾患、および感染症において、或る役割を果たすと考えられる。NAAPの発現または活性の増大に関連する疾患の治療においては、NAAPの発現または活性を低下させることが望ましい。また、NAAPの発現または活性の低下に関連する疾患の治療においては、NAAPの発現または活性を増大させることが望ましい。

【0229】

従って、一実施態様では、NAAPの発現または活性の低下に関連した疾患の治療または予防のために、被験者にNAAPまたはその断片や誘導体が投与され得る。限定するものではないが、このような疾患には細胞増殖異常、神経障害、発達障害、自己免疫/炎症疾患、および感染症が含まれ、細胞増殖異常には日光角化症、動脈硬化、アテローム硬化、滑液包炎、硬変、肝炎、混合性結合組織病(MCTD)、骨髄線維症、発作性夜間ヘモグロビン尿症、真性多血症、乾癬、原発性血小板血症、並びに、腺癌、白血病、リンパ腫、黒色腫、骨髄腫、肉腫、および奇形癌など癌、具体的には、副腎、膀胱、骨、骨髄、脳、乳房、頸部、胆嚢、神経節、消化管、心臓、腎臓、肝臓、肺、筋肉、卵巣、膵臓、副甲状腺、陰茎、前立腺、唾液腺、皮膚、脾臓、精巣、胸腺、甲状腺、および子宮の癌が含まれ、神経障害には、癲癇、虚血性脳血管障害、脳卒中、脳腫瘍、アルツハイマー病、ピック病、ハンチントン病、痴呆、パーキンソン病およびその他の錐体外路障害、筋萎縮性側索硬化およびその他の運動ニューロン障害、進行性神経性筋萎縮症、網膜色素変性症(色素性網膜炎)、遺伝性運動失調、多発性硬化症および他の脱髄疾患、細菌性およびウイルス性髄膜炎、脳膿瘍、硬膜下膿瘍、硬膜外膿瘍、化膿性頭蓋内血栓性静脈炎、脊髄炎および神経根炎、ウイルス性中枢神経系疾患と、プリオン病(クールー、クロイツフェルト ヤコブ病、およびGerstmann-Straussler-Scheinker症候群を含む)、致死性家族性不眠症、神経系性栄養病および代謝病、神経線維腫症、結節硬化症、小脳網膜血管腫症(cerebelloretinal hemangioblastomatosis)、脳3叉神経血管症候群、中枢神経系性精神遅滞および他の発達障害、脳性麻痺、神経骨格異常症、自律神経系障害、脳神経障害、脊髄疾患、筋ジストロフィー他の神経筋障害、末梢神経疾患、皮膚筋炎および多発性筋炎と、遺伝性、代謝性、内分泌性、および中毒性ミオパシーと、重症筋無力症、周期性四肢麻痺、精神障害(気分性、不安性の障害、統合失調症/分裂病)、季節性感情障害(SAD)、静座不能、健忘症、緊張病、糖尿病性ニューロパシー、遅発性ジスキネジア、ジストニー、パラノイド精神病、帯状疱疹後神経痛、トゥレット病が含まれ、発達障害には尿細管性アシドーシス、貧血、クッシング症候群、軟骨形成不全性小人症、デュシェンヌ/ベッカー型筋ジストロフィー、癲癇、性腺形成異常、WAGR症候群(ウィルムス腫瘍、無虹彩症、尿生殖器異常、精神遅滞)、スミス マジェニス症候群(Smith-Magenis syndrome)、骨髄異形成症候群、遺伝性粘膜上皮異形成、遺伝性角皮症や、シャルコーマリーツース病および神経線維腫症などの遺伝性神経病、甲状腺機能低下症、水頭症や、Sydenham's chorea)および脳性小児麻痺などの発作障害、二分脊椎、無脳症、頭蓋脊椎披裂、先天性緑内障、白内障、感音難聴が含まれ、自己免疫/炎症疾患には、後天性免疫不全症候群(AIDS)、アジソン病(慢性原発性副腎機能不全)、成人呼吸窮迫症候群、アレルギー、強直性脊椎炎、アミロイド症、貧血、喘息、アテローム性動脈硬化、自己免疫性溶血性貧血、自己免疫性甲状腺炎、自己免疫性多腺性内分泌カンジダ症外胚葉ジストロフィー(APECED)、気管支炎、胆嚢炎、接触皮膚炎、クローン病、アトピー性皮膚炎、皮膚筋炎、糖尿病、肺気腫、リンパ球傷害因子性偶発性リンパ球減少症、新生児溶血性疾患(胎児赤芽球症)、結節性紅斑、萎縮性胃炎、糸球体腎炎、グッドパスチャー症候群、痛風、グレーブス病、橋本甲状腺炎、過好酸球増加症、過敏性腸症候群、多発性硬化症、重症筋無力症、心筋または心膜炎、骨関節炎、骨粗しょう症、膵炎、多発性筋炎、乾癬、ライター症候群、リウマチ様関節炎、強皮症、シェーグレン症候群、全身性アナフィラキシー、全身性エリテマトーデス、全身性強皮症、血小板減少性紫斑病、潰瘍性大腸炎、ぶどう膜炎、ウェルナー症候群、癌合併症、血液透析、体外循環、原虫感染症、蠕虫感染症、外傷が含ま

れ、感染症には、アデノウイルス、アレナウイルス、ブンヤウイルス (bunyavirus)、カリチウイルス (calicivirus)、コロナウイルス、フィロウイルス、ヘパドナウイルス、ヘルペスウイルス、フラビウイルス、オルソミクソウイルス、パルボウイルス、パポーバウイルス、パラミキソウイルス、ピコルナウイルス、ポックスウイルス、レオウイルス、レトロウイルス、ラドウイルス、トガウイルスに分類されるウイルス病原体による感染や、細菌による感染 (肺炎球菌感染症、ブドウ球菌、連鎖球菌、バシラス菌、コリネバクテリウム、クロストリジウム、髄膜炎菌、淋菌、リステリア、モラクセラ、キングラ、ヘモフィルス、レジオネラ、百日咳菌や、シゲラ、サルモネラ、カンピロバクターを含むグラム陰性腸内細菌、シュードモナス、ビブリオ、ブルセラ、野兔病菌、エルシニア、バルトネラ、nocardium、放線菌、ミコバクテリウム、spirochaetale、リケッチア、クラミジア、マイコプラズマに分類)、真菌感染 (分類はアスペルギルス、プラストミセス、皮膚糸状菌、クリプトコッカス、コクシジオイデス、malassezia、ヒストプラズマ、または他の真菌症起因菌)、寄生虫感染 (分類はプラスモディウムすなわちマラリア原虫、寄生性アメーバ、リーシュマニア、トリパノソーマ、トキソプラズマ、ニューモシスチスカリニ、腸内原虫 (ジアルジアなど)、トリコモナス、組織線虫 (旋毛虫など)、腸管寄生線虫 (回虫など)、リンパ管フィラリア線虫、吸虫 (住血吸虫など)、および糸虫 (サナダムシなど)) が含まれる。

10

【0230】

別の実施態様では、限定するものではないが上に列記した疾患を含む、NAAPの発現または活性の低下に関連した疾患の治療または予防のために、NAAPまたはその断片や誘導体を発現し得るベクターを被験者に投与し得る。

20

【0231】

更に別の実施態様では、限定するものではないが上に列記した疾患を含む、NAAPの発現または活性の低下に関連した疾患の治療または予防のために、実質的に精製されたNAAPを有する組成物を、好適な医薬用キャリアと共に被験者に投与し得る。

【0232】

更に別の実施態様では、限定するものではないが上に列記した疾患を含む、NAAPの発現または活性の低下に関連した疾患の治療または予防のために、NAAPの活性を調節するアゴニストを被験者に投与し得る。

【0233】

更なる実施例では、NAAPの発現または活性の増大に関連した疾患の治療または予防のために、被験者にNAAPのアンタゴニストを投与し得る。限定するものではないが、このような疾患の例には、上記した細胞増殖異常、神経障害、発達障害、自己免疫/炎症性疾患、および、感染症が含まれる。一実施態様では、NAAPと特異的に結合する抗体が直接アンタゴニストとして、或いはNAAPを発現する細胞または組織に薬剤を運ぶターゲティング機構或いは送達機構として間接的に用いられ得る。

30

【0234】

別の実施例では、NAAPをコードするポリヌクレオチドの相補配列を発現するベクターを被験者に投与して、限定するものではないが上記した疾患を含む、NAAPの発現または活性の増大に関連した疾患の治療または予防を成し得る。

40

【0235】

別の実施態様では、本発明の任意のタンパク質、アンタゴニスト、抗体、アゴニスト、相補配列、またはベクターを、別の好適な治療薬と組み合わせることもできる。併用療法で用いる好適な治療薬は、当業者が従来 of 医薬原理に従って選択し得る。治療薬と組み合わせることにより、上記した種々の疾患の治療または予防に相乗効果をもたらし得る。この方法により、少量の各薬剤で医薬効果をあげることが可能となり、それによって副作用の可能性を低減し得る。

【0236】

NAAPのアンタゴニストは、本技術分野で一般的に知られている方法を用いて製造し得る。具体的には、精製されたNAAPを用いて抗体を作るか、治療薬のライブラリをスクリーニン

50

グして、NAAPと特異結合するものを同定することが可能である。NAAPへの抗体も、本技術分野で公知の方法を用いて産生され得る。限定するものではないがこのような抗体には、ポリクローナル抗体、モノクローナル抗体、キメラ抗体、一本鎖抗体、Fab断片、およびFab発現ライブラリによって作られた断片が含まれ得る。中和抗体（すなわち二量体の形成を阻害する抗体）は通常、治療用に好適である。

【0237】

抗体の産生のためには、ヤギ、ウサギ、ラット、マウス、ヒトその他の種々の宿主が、NAAPまたは任意の断片の注入、または免疫原性の特性を備えるそのオリゴペプチドの注入によって免疫化され得る。宿主の種に応じて、種々のアジュバントを用いて免疫応答を高めることもできる。限定するものではないがこのようなアジュバントには、フロイントアジュバントと、水酸化アルミニウムなどのミネラルゲルアジュバントと、リゾレシチン、プルロニックポリオール、ポリアニオン、ペプチド、油性乳剤、スカシガイのヘモシアニン（KLH）、ジニトロフェノールなどの界面活性剤とがある。ヒトに用いられるアジュバントの中では、BCG（カルメット ゲラン杆菌）およびコリネバクテリウム パルバム（*Corynebacterium parvum*）が特に好ましい。

10

【0238】

NAAPに対する抗体を誘導するために用いるオリゴペプチド、ペプチドまたは断片は、少なくとも約5個のアミノ酸からなるアミノ酸配列を持つものが好ましく、一般的には約10個以上のアミノ酸からなるものとなる。これらのオリゴペプチド、ペプチドまたは断片は、天然タンパク質のアミノ酸配列の一部と同一であることが望ましい。NAAPアミノ酸類の短いストレッチ群は、KLHなど他のタンパク質の配列と融合され、このキメラ分子に対する抗体群が産生され得る。

20

【0239】

NAAPに対するモノクローナル抗体は、培地内の連続した細胞株によって、抗体分子を産生する任意の技術を用いて作製することが可能である。限定するものではないがこのような技術には、ハイブリドーマ技術、ヒトB細胞ハイブリドーマ技術、およびEBV-ハイブリドーマ技術がある（例えばKohler, G. 他(1975) Nature 256:495497; Kozbor, D. 他(1985) J. Immunol. Methods 81:31-42; Cote, R.J. 他(1983) Proc. Natl. Acad. Sci. USA 80:20262030; およびCole, S.P. 他(1984) Mol. Cell Biol. 62:109-120を参照）。

【0240】

更に、「キメラ抗体」作製のために発達した、ヒト抗体遺伝子にマウス抗体遺伝子をスプライシングするなどの技術が、好適な抗原特異性および生物学的活性を備える分子を得るために用いられる（例えばMorrison, S.L. 他(1984) Proc. Natl. Acad. Sci. USA 81:68516855; Neuberger, M.S. 他(1984) Nature 312:604608; およびTakeda, S. 他(1985) Nature 314:452-454を参照）。別法では、当分野で周知の方法を用い、一本鎖抗体の産生のための記載された技術を適用して、NAAP特異的一本鎖抗体を生成する。関連した特異性を有するガイデオタイプ組成が異なるような抗体を、ランダムな組合せの免疫グロブリンライブラリ類からチェーンシャッフリングによって産生することもできる（例えばBurton D.R. (1991) Proc. Natl. Acad. Sci. USA 88:10134-10137を参照）。

30

【0241】

抗体類の産生は、リンパ球集団における *in vivo* 産生の誘導によって、或いは文献に開示されているように非常に特異的な結合試薬の免疫グロブリンのライブラリまたはパネルのスクリーニングによっても行い得る（例えばOrlandi, R. 他(1989) Proc. Natl. Acad. Sci. USA 86:38333837; Winter, G. 他(1991) Nature 349:293-299を参照）。

40

【0242】

NAAPに対する特異的な結合部位を含む抗体断片も産生され得る。例えば、限定するものではないが、このような断片には、抗体分子のペプシン消化によって産生されるF(ab')₂断片と、F(ab')₂断片のジスルフィド架橋を還元することによって作製されるFab断片とがある。あるいは、Fab発現ライブラリを作製することによって、所望の特異性を持つモノクローナルFab断片を迅速且つ容易に同定することが可能となる（例えばHuse, W.D.

50

他 (1989) Science 246:1275-1281を参照)。 (1989) Science 246:1275-1281等を参照)。

【0243】

種々の免疫学的検定 (イムノアッセイ) を用いてスクリーニングし、所望の特異性を有する抗体を同定することができる。確立された特異性を有するポリクローナル抗体またはモノクローナル抗体の何れかを用いる免疫放射定量測定法または競合結合アッセイに対する数々のプロトコルは、本技術分野において公知である。このようなイムノアッセイは通常、NAAPとその特異抗体との間の複合体形成の計測を含む。2つの非干渉性NAAPエピトープに対して反応性を持つモノクローナル抗体群を用いる、2部位モノクローナルベースのイムノアッセイが一般に利用されるが、競合結合試験も利用できる (Pound、前出)。

【0244】

ラジオイムノアッセイ技術と共にScatchard分析などの様々な方法を用いて、NAAPに対する抗体の親和性を評価し得る。親和性を結合定数Kaで表すが、このKaは、平衡状態下の、NAAP抗体複合体のモル濃度を、遊離抗体と遊離抗原とのモル濃度で除して得られる値である。ポリクローナル抗体類は多様なNAAPエピトープに対する親和性が不均一であり、或るポリクローナル抗体試薬に関して判定したKaは、NAAP抗体群の平均の親和性または結合活性を表す。モノクローナル抗体は或る特定のNAAPエピトープに対して単一特異的であり、モノクローナル抗体の或る試薬について判定したKaは、親和性の真の測定値を表す。Ka値が $10^9 \sim 10^{12}$ L/molの高親和性抗体試薬は、NAAP-抗体複合体が激しい操作に耐えなければならないイムノアッセイに用いるのが好ましい。Ka値が $10^6 \sim 10^7$ liter/molの低親和性抗体試薬は、NAAPが抗体から最終的に活性化状態で解離する必要がある免疫精製 (immunopurification) 及び類似の処理に用いるのが好ましい (Catty, D. (1988) Antibodies, Volume I: A Practical Approach. IRL Press, Washington, DC; Liddell, J. E. 及び Cryer, A. (1991) A Practical Guide to Monoclonal Antibodies, John Wiley & Sons, New York NY)。

【0245】

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

【0246】

本発明の別の実施例では、NAAPをコードするポリヌクレオチド、またはその任意の断片や相補配列を、治療目的で使用することができる。或る実施例では、NAAPをコードする遺伝子のコーディング領域や調節領域に相補的な配列やアンチセンス分子 (DNA、RNA、PNA、または修飾したオリゴヌクレオチド) を設計して遺伝子発現をモディフィケーションし得る。このような技術は当分野では周知であり、アンチセンスオリゴヌクレオチドまたはより大きな断片を、NAAPをコードする配列の制御領域、またはコード領域に沿った、さまざまな位置から設計可能である (例えばAgrawal, S., 編集 (1996) Antisense Therapeutics, Humana Press Inc., Totawa NJを参照)。

【0247】

治療に用いる場合、アンチセンス配列を適切な標的細胞に導入するのに好適な、任意の遺伝子送達系を用いることができる。アンチセンス配列は、転写時に標的タンパク質をコードする細胞配列の少なくとも一部に相補的な配列を作製する発現プラスミドの形で、細胞内に送達し得る (例えばSlater, J.E. 他 (1998) J. Allergy Clin. Immunol. 102 (3) :469-475 および Scanlon, K.J. 他 (1995) 9 (13) :1288-1296を参照)。アンチセンス配列はまた、例えばレトロウイルスやアデノ随伴ウイルスベクターなどのウイルスベクターを用いて細胞内に導入することもできる (例えばMiller, A.D. (1990) Blood 76:271、前出のAusubel, Uckert, W. および W. Walther (1994) Pharmacol. Ther. 63 (3) :323-347

10

20

30

40

50

を参照)。その他の遺伝子送達機構には、リポソーム系、人工的なウイルスエンベロープおよび当分野で公知のその他のシステムが含まれる(例えばRossi, J.J. (1995) Br. Med. Bull. 51(1):217-225; Boado, R.J.他(1998) J. Pharm. Sci. 87(11):1308-1315、Morris, M.C.他(1997) Nucleic Acids Res. 25(14):2730-2736を参照)。

【0248】

本発明の別の実施態様では、NAAPをコードするポリヌクレオチドを、体細胞若しくは生殖細胞の遺伝子治療に用いることが可能である。遺伝子治療により、(i) 遺伝子欠損症を治療し(例えばX染色体連鎖遺伝(Cavazzana-Calvo, M.他(2000) Science 288:669-672)により特徴付けられる重症複合型免疫不全(SCID)-X1病の場合)、先天性アデノシンデアミナーゼ(ADA)欠損症に関連する重症複合型免疫不全症候群(Blaese, R.M. 他(1995) Science 270:475-480、Bordignon, C.他(1995) Science 270:470-475)、嚢胞性繊維症(Zabner, J.他(1993) Cell 75:207-216; Crystal, R.G.他(1995) Hum. Gene Therapy 6:643-666、Crystal, R.G.他(1995) Hum. Gene Therapy 6:667-703)、サラセミア(thalassaemia)、家族性高コレステロール血症や、第VIII因子若しくは第IX因子欠損に起因する血友病(Crystal, R.G. (1995) Science 270:404-410、Verma, I.M. および N. Somia (1997) Nature 389:239-242)、(ii) 条件的致死性遺伝子産物を発現させ(例えば制御不能な細胞増殖に起因する癌の場合)、(iii) 細胞内の寄生生物、例えばヒト免疫不全ウイルス(HIV)(Baltimore, D. (1988) Nature 335:395-396、Poeschla, E.他(1996) Proc. Natl. Acad. Sci. USA. 93:11395-11399)などヒトレトロウイルス、B型若しくはC型肝炎ウイルス(HBV、HCV)、*Candida albicans*および*Paracoccidioides brasiliensis*等の寄生真菌、並びに熱帯熱マラリア原虫およびクルーズトリパノソーマ等の寄生原虫に対する防御機能を有するタンパク質を発現させることができる。NAAPの発現若しくは調節に必要な遺伝子の欠損が疾患を引き起こす場合、導入した細胞の好適な集団からNAAPを発現させて、遺伝子欠損によって起こる症状の発現を緩和することが可能である。

10

20

【0249】

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

30

【0250】

NAAPの発現に影響を及ぼし得る発現ベクターには、限定するものではないが、PCDNA 3.1、EPITAG、PRCCMV2、PREP、PVAX、PCR2-TOPOTAベクター(Invitrogen, Carlsbad CA)、PCM V-SCRIPT、PCMV-TAG、PEGSH/PERV (Stratagene, La Jolla CA)、PTET-OFF、PTET-ON、PTRE2、PTRE2-LUC、PTK-HYG (Clontech, Palo Alto CA)が含まれる。DMEを発現させるために、(i) 構成的に活性なプロモーター(例えば、サイトメガロウイルス(CMV)、ラウス肉腫ウイルス(RSV)、SV40ウイルス、チミジンキナーゼ(TK)、若しくは - アクチンの遺伝子など)、(ii) 誘導性プロモーター(例えば、市販されているT-REXプラスミド(Invitrogen)に含まれている、テトラサイクリン調節性プロモーター(Gossen, M. および H. Bujard (1992) Proc. Natl. Acad. Sci. U.S.A. 89:5547-5551; Gossen, M. 他(1995) Science 268:1766-1769; Rossi, F.M.V. および H.M. Blau (1998) Curr. Opin. Biotechnol. 9:451-456))、エクジソン誘導性プロモーター(市販されているプラスミドPVGR XRおよびPINDに含まれている: Invitrogen)、FK506/ラパマイシン誘導性プロモーター、またはRU486/ミフェプリストーン誘導性プロモーター(Rossi, F.M.V. および H.M. Blau, 前出)、または(iii) 正常な個体に由来する、NAAPをコードする内因性遺伝子の天然プロモーター若しくは組織特異的プロモーターを用い得る。

40

【0251】

50

市販のリポソーム形質転換キット（例えばInvitrogen社のPerFect Lipid Transfection Kit）を用いれば、当業者は実験の各パラメータを最適化する努力をさほど要さず、ポリヌクレオチド群を、培養中の標的細胞群に送達し得る。別法では、リン酸カルシウム法（Graham, F.L. および A.J. Eb (1973) *Virology* 52:456-467）若しくは電気穿孔法（Neuman, B. 他(1982) *EMBO J.* 1:841845）。初代培養細胞にDNAを導入するためには、標準化された哺乳類の形質移入プロトコルの修正が必要である。

【0252】

本発明の別の実施例では、NAAPの発現に関連する遺伝子欠損によって起こる疾患や障害を、(i)レトロウイルス長末端反復配列(LTR)プロモーター若しくは或る独立プロモーターのコントロール下でNAAPをコードするポリヌクレオチドと、(ii)好適なRNAパッケージングシグナル群と、(iii)追加のレトロウイルス・シス作用性RNA配列と、効率的なベクター増殖に必要なコーディング配列とを伴うRev応答性エレメント(RRE)と、からなるレトロウイルスベクターを作製して治療することができる。レトロウイルスベクター（例えばPFBおよびPFBNE0）はStratagene社から市販されており、刊行データ(Riviere, I. 他(1995) *Proc. Natl. Acad. Sci. USA* 92:6733-6737)に基づく。該文献は特に引用を以て本明細書の一部となす。このベクターは、好適なベクター産生細胞株(VPCL)において増殖される。VPCLは、各標的細胞上の受容体への親和性を持つエンベロープ遺伝子を、またはVSVgなど汎親和性エンベロープタンパク質を発現する(Armentano, D. 他(1987) *J. Virol.* 61:1647-1650、Bender, M.A. 他(1987) *J. Virol.* 61:1639-1646、Adam, M.A. および A.D. Miller (1988) *J. Virol.* 62:3802-3806、Dull, T. 他(1998) *J. Virol.* 72:8463-8471、Zufferey, R. 他(1998) *J. Virol.* 72:9873-9880)。Riggに付与された米国特許第5,910,434号（「Method for obtaining retrovirus packaging cell lines producing high transducing efficiency retroviral supernatant」）において、レトロウイルスパッケージング細胞株群を得る方法が開示されており、引用を以て本明細書の一部とする。レトロウイルスベクターの増殖、細胞集団（例えばCD4⁺ T細胞）の形質導入、及び形質導入した細胞の患者への戻しは、遺伝子治療の分野では当業者に公知の方法であり、多数の文献に記載されている(Ranga, U. 他(1997) *J. Virol.* 71:7020-7029; Bauer, G. 他(1997) *Blood* 89:2259-2267; Bonyhadi, M.L. (1997) *J. Virol.* 71:4707-4716; Ranga, U. 他(1998) *Proc. Natl. Acad. Sci. USA* 95:1201-1206; Su, L. (1997) *Blood* 89:2283-2290)。

【0253】

別法では、アデノウイルス系遺伝子治療の或る送達系を用いて、NAAPの発現に関連する1つ以上の遺伝子異常を有する細胞群に、NAAPをコードするポリヌクレオチド群を送達する。アデノウイルス系ベクター類の作製およびパッケージングについては、当業者に周知である。複製欠損型アデノウイルスベクター類は、種々の免疫調節タンパク質をコードする遺伝子群を、無損傷の臍島内にインポートする目的で多様に利用し得ることが証明された(Csete, M.E. 他(1995) *Transplantation* 27:263-268)。潜在的に有用なアデノウイルスベクター類はArmentanoに付与された米国特許第5,707,618号（「Adenovirus vectors for gene therapy」）に記載されており、引用することを以て本明細書の一部とする。アデノウイルスベクター類については、Antinozzi, P.A. 他(1999) *Annu. Rev. Nutr.* 19:511-544 並びに、Verma, I.M. および N. Somia (1997) *Nature* 18:389:239-242も参照されたい。両文献は、引用を以て本明細書の一部とする。

【0254】

別法では、ヘルペス系遺伝子治療の送達系を用いて、NAAPの発現に関連する1つ或いは複数の遺伝子異常を有する標的細胞に、NAAPをコードするポリヌクレオチドを送達する。単純ヘルペスウイルス(HSV)系ベクター類は、HSVが親和性を持つ中枢神経細胞にNAAPを導入する際に、特に有益に思える。ヘルペス系ベクター類の作製およびパッケージングは、当業者に公知である。或る複製適格性単純ヘルペスウイルス(HSV) I型系のベクターが、或るレポーター遺伝子の、霊長類の眼への送達に用いられている(Liu, X. 他(1999) *Exp. Eye Res.* 169:385395)。HSV-1ウイルスベクターの作製についても、DeLucaに付与された

10

20

30

40

50

米国特許第5,804,413号(「Herpes simplex virus strains for gene transfer」)に詳細に開示されており、該特許の引用を以て本明細書の一部とする。米国特許第5,804,413号は、ヒト遺伝子治療などの目的で、好適なプロモーターの制御下において細胞に導入される少なくとも1つの外来遺伝子を有するゲノムを持つ、組換えHSV d92の利用について記載する。上記特許はまた、ICP4、ICP27、およびICP22を欠失した組換えHSV系統の、作製および使用について開示している。HSVベクター類については、Goins, W.F. 他(1999) *J. Virol.* 73:519-532 および Xu, H. 他(1994) *Dev. Biol.* 163:152-161も参照されたい。両文献は、引用を以て本明細書の一部とする。クローン化したヘルペスウイルス配列群の操作、大ヘルペスウイルスゲノム(large herpesvirus genomes)の様々なセグメントを持つ複数のプラスミドを形質移入した後の組換えウイルスの産生、ヘルペスウイルスの成長および増殖、並びに細胞群へのヘルペスウイルスの感染は、当業者に公知の技術である。

10

【0255】

別法では、或るウイルス(正の一本鎖RNAウイルス)ベクターを用いて、NAAPをコードするポリヌクレオチド群を標的細胞群に送達する。プロトタイプのウイルスであるセムリキ森林熱ウイルス(SFV)の生物学的研究が広範に行われており、遺伝子導入ベクター類は、SFVゲノムに基づく(Garoff, H. および K.-J. Li (1998) *Curr. Opin. Biotechnol.* 9:464-469)。ウイルスRNAの複製中に、ウイルスのカプシドタンパク質群を通常はコードする、サブゲノムRNAが産生される。このサブゲノムRNAは、完全長ゲノムRNAよりも高レベルに複製するので、酵素活性(例えばプロテアーゼおよびポリメラーゼ)を持つウイルスタンパク質に比べてカプシドタンパク質群が過剰産生される。同様に、NAAPをコードする配列をウイルスゲノムのカプシドをコードする領域に導入することによって、ベクター導入細胞において多数のNAAPをコードするRNAが産生され、高いレベルでNAAPが合成される。通常、ウイルスの感染は、数日以内の細胞溶解を伴う。一方、シンドビスウイルス(SIN)の或る変異体を有するハムスター正常腎臓細胞(BHK-21)群が持続的な感染を確立する能力は、ウイルス類の溶解複製を、遺伝子治療に応用し得るように好適に改変可能であることを示唆する(Dryga, S.A. 他(1997) *Virology* 228:74-83)。様々な宿主にウイルスを導入できるので、様々なタイプの細胞にNAAPを導入することができる。或る集団における或るサブセットの細胞群の特異的形質導入は、形質導入前に細胞の選別を必要とし得る。ウイルスの感染性cDNAクローンの処置方法、ウイルスのcDNAおよびRNAの形質移入方法およびウイルスの感染方法は、当業者に公知である。

20

30

【0256】

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

40

【0257】

リボザイム類は、酵素性RNA分子である。RNAの特異的切断を触媒するためにリボザイムを用い得る。リボザイム作用のメカニズムは、相補的標的RNAへのリボザイム分子の配列特異的ハイブリダイゼーションと、その後起こる内ヌクレオチド鎖切断に関与している。例えば、人工的に作製されたハンマーヘッド型リボザイム分子が、NAAPをコードする配列の、内ヌクレオチド鎖切断を特異的且つ効果的に触媒できる可能性がある。

【0258】

任意の潜在的RNA標的内の特異的リボザイム切断部位を先ず同定するため、GUA、GUU、GUC配列などリボザイム切断部位に対して標的分子をスキャンする。一度同定されると、切断

50

部位を持つ標的遺伝子の領域に対応する15～20リボヌクレオチドの短いRNA配列が、そのオリゴヌクレオチドを機能不全にするような2次構造の特徴をもっていないかを評価することが可能になる。候補標的の適合性の評価も、リボヌクレアーゼ保護アッセイを用いて相補的オリゴヌクレオチド群とのハイブリダイゼーションへのアクセス可能性をテストすることによって行い得る。

【0259】

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

10

【0260】

細胞内の安定性を高め、半減期を長くするためにRNA分子を修飾し得る。限定するものではないが可能な修飾としては、分子の5'末端、3'末端、あるいはその両方において隣接配列群を追加することや、分子の主鎖内においてホスホジエステラーゼ結合ではなくホスホリボチオエートまたは2'-O-メチルを使用することが含まれる。この概念は、本来はPNA群の産出におけるものであるが、これら全ての分子に拡大することができる。そのためには、内因性エンドヌクレアーゼによって容易には認識されない、アデニン、シチジン、グアニン、チミン、およびウリジンにアセチル-、メチル-、チオ-および同様の修飾をしたものや、非従来型塩基、例えばイノシン、クエオシン(queosine)、ワイプトシン(wybutosine)を加える。

20

【0261】

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

30

【0262】

或る特定ポリヌクレオチドの発現を改変する際の有効性に対して、少なくとも1個または複数個の試験化合物をスクリーニングし得る。試験化合物を得るには、当分野で公知の任意の方法を用い得る。取得方法としては、以下の場合に有効な既知化合物の化学修飾がある。ポリヌクレオチドの発現を改変する場合と、既存の、商用または専用の、天然または非天然の化合物のライブラリから選択する場合と、標的ポリヌクレオチドの化学的および/または構造的特性に基づく化合物を合理的にデザインする場合と、組合せ的にまたは無作為に生成した化合物のライブラリから選択する場合とである。NAAPをコードするポリヌクレオチドを持つサンプルを、このようにして得た試験化合物の少なくとも1つに曝露する。サンプルは例えば、無傷細胞、透過化処理した細胞、*in vitro*無細胞系すなわち再構成生化学系があり得る。NAAPをコードするポリヌクレオチドの発現における改変は、当分野で周知の任意の方法でアッセイする。通常は、NAAPをコードするポリヌクレオチドの配列に相補的なヌクレオチド配列を有するプローブを用いたハイブリダイゼーションにより、特定のヌクレオチドの発現を検出する。ハイブリダイゼーション量を定量することにより、1つ以上の試験化合物に曝露される、または曝露されないポリヌクレオチドの発現の

40

50

比較のための基礎を形成し得る。試験化合物に曝露されるポリヌクレオチドの発現における変化の検出は、ポリヌクレオチドの発現を改変する際に試験化合物が有効であることを示す。或る特定ポリヌクレオチドの改変発現に有効な化合物のためのスクリーニングを実行でき、例えば分裂酵母 (*Schizosaccharomyces pombe*) 遺伝子発現系 (Atkins, D. 他 (1999) 米国特許第5,932,435号、Arndt, G.M. 他 (2000) *Nucleic Acids Res.*28:E15) またはヒト細胞株、例えばHeLa細胞 (Clarke, M.L. 他 (2000) *Biochem. Biophys. Res. Commun.* 268:8-13)を用いる。本発明の特定の実施例は、特異的ポリヌクレオチド配列に対するアンチセンス活性のためのオリゴヌクレオチド (デオキシリボヌクレオチド、リボヌクレオチド、ペプチド核酸、修飾オリゴヌクレオチド) の組み合わせライブラリをスクリーニングすることに関与している (Bruce, T.W. 他 (1997) 米国特許第5,686,242号、Bruce, T.W. 他 (2000) 米国特許第6,022,691号)。

【0263】

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

【0264】

上記の治療方法はいずれも、例えば、ヒト、イヌ、ネコ、ウシ、ウマ、ウサギ、サルなどの哺乳類を含めて治療が必要な全ての被験体に適用できる。

【0265】

本発明の或る更なる実施例は、薬剤として許容できる賦形剤と共に処方される活性成分を一般に持つ、或る組成物の投与に関する。賦形剤には例えば、糖、でんぷん、セルロース、ゴムおよびタンパク質がある。様々な処方が広く知られており、詳細は *Remington's Pharmaceutical Sciences* (Maack Publishing, Easton PA) の最新版に記載されている。このような組成物は、NAAP、NAAPの抗体、擬態物質、アゴニスト、アンタゴニスト、またはインヒビターなどからなる。

【0266】

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

【0267】

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

【0268】

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

【0269】

NAAPを有する、またはその断片群を有する高分子群を直接、細胞内に送達すべく、特殊な種々の形状の組成物群が調製され得る。例えば、細胞不透過性高分子を含むリポソーム製剤は、その高分子の細胞融合と細胞内送達とを促進し得る。別法では、NAAPまたはその断片を、HIV Tat-1タンパク質の陽イオンN末端部に結合することもできる。このようにして産生された融合タンパク質類は、或るマウスモデル系の、脳を含む全ての組織の細胞群に

形質導入することがわかっている (Schwarze, S.R. 他(1999) Science 285:1569-1572)。

【0270】

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

【0271】

治療有効投与量とは、症状や容態を回復させる、たとえばNAAPまたはその断片、NAAPの抗体、NAAPのアゴニストまたはアンタゴニスト、インヒビターなど活性処方成分の量を指す。治療有効度および毒性は、細胞培養または動物実験における標準的な薬学手法によって、例えばED₅₀ (集団の50%の治療有効量) またはLD₅₀ (集団の50%の致死量) 統計を計算するなどして判定できる。毒性効果の、治療効果に対する投与量比が治療指数であり、LD₅₀/ED₅₀比として表すことができる。高い治療指数を示すような組成物が望ましい。細胞培養アッセイと動物実験とから得られたデータは、ヒトに用いる投与量の範囲での調剤に用いられる。このような組成物に含まれる投与量は、毒性を殆どあるいは全く持たず、ED₅₀を含むような血中濃度の範囲にあることが好ましい。用いられる投与形態、患者の感受性および投与の経路によって、投与量はこの範囲内で様々に変わる。

10

【0272】

正確な投与量は、治療が必要な被験者に関する要素を考慮して、現場の医者が決定することになる。十分なレベルの活性成分を与え、あるいは所望の効果を維持すべく、用法および用量を調整する。被験者に関する要素としては、疾患の重症度、患者の一般的な健康状態、患者の年齢、体重および性別、投与の時間および頻度、薬剤の配合、反応感受性および治療に対する応答などを考慮し得る。作用期間が長い組成物は、特定の製剤の半減期およびクリアランス率によって3~4日毎に1度、1週間に1度、あるいは2週間に1度の間隔で投与し得る。

20

【0273】

通常の投与量は、投与の経路にもよるが約0.1~100,000 μgであり、合計で約1gまでとする。特定の投与量および送達方法に関するガイダンスは文献に記載されており、現場の医者は通常それを利用することができる。当業者は、タンパク質またはそれらのインヒビター類についての処方とは異なる、ヌクレオチドについての処方を利用することになる。同様に、ポリヌクレオチドまたはポリペプチドの送達は、特定の細胞、症状、部位などに特異的なものとなる。

30

【0274】

(診断)

別の実施例では、NAAPに特異的に結合する抗体を、NAAPの発現によって特徴付けられる障害の診断に、または、NAAPやNAAPのアゴニストまたはアンタゴニスト、インヒビターで治療を受けている患者をモニターするためのアッセイに用い得る。診断目的に有用な抗体は、上記の治療の箇所に記載した方法と同じ方法で調合される。NAAPの診断アッセイには、ヒトの体液から、あるいは細胞や組織の抽出物から、抗体および標識を用いてNAAPを検出する方法が含まれる。この抗体は修飾されたものもされていないものも可能であり、レポーター分子との共有結合または非共有結合で標識化できる。レポーター分子としては広くさまざまな種類が本分野で知られており、また使用可能であるが、そのうちのいくつかは上記で説明されている。

40

【0275】

NAAPを測定するための様々なプロトコル、例えばELISA、RIA、FACS等が当分野では周知であり、変化した、あるいは異常なレベルのNAAPの発現を診断するための基盤を提供する。正常あるいは標準的なNAAPの発現の値は、複合体の形成に適した条件下で、正常な哺乳動物、例えばヒトなどの被験体から採取した体液または細胞抽出物と、NAAPに対する抗体とを混合させることによって確定する。標準複合体形成量は、種々の方法、例えば測光法で

50

定量できる。被験体、対照、および、生検組織からの疾患サンプルでの、NAAP発現の量を標準値と比較する。標準値と被験体との偏差が、疾患を診断するパラメータを確定する。

【0276】

本発明の別の実施態様によれば、NAAPをコードするポリヌクレオチドを診断のために用いることもできる。用い得るポリヌクレオチドには、オリゴヌクレオチド配列、相補的RNAおよびDNA分子、そしてPNAが含まれる。これらのポリヌクレオチドを用いて、NAAPの発現が疾患と相関し得る生検組織における遺伝子発現を検出し定量し得る。この診断アッセイを用いて、NAAPの存在の有無、更には過剰な発現を判定し、治療時のNAAPレベルの調節を監視する。

【0277】

或る実施形態では、NAAPまたは近縁の分子をコードする、ゲノム配列などポリヌクレオチド配列を検出可能なPCRプローブ類とのハイブリダイゼーションを、NAAPをコードする核酸配列を同定するために用いることができる。プローブが高度に特異的な領域（例えば5'調節領域）から作られている、或いはやや特異性の低い領域（例えば保存されたモチーフ）から作られているかにかかわらず、プローブの特異性と、ハイブリダイゼーション或いは増幅のストリンジェンシーとによって、そのプローブがNAAPをコードする天然配列のみを同定するかどうか、或いは対立遺伝子変異配列や関連配列を同定するかどうかが決まることとなる。

【0278】

プローブはまた、関連する配列の検出に利用でき、また、NAAPをコードする任意の配列との少なくとも50%の配列同一性を有し得る。目的の本発明のハイブリダイゼーションプローブには、DNAあるいはRNAが可能であり、SEQ ID NO:17-32の配列、或いはNAAP遺伝子のプロモーター、エンハンサー、イントロンを含むゲノム配列に由来し得る。

【0279】

NAAPをコードするDNAに対して特異的なハイブリダイゼーションプローブの作製方法には、NAAPをコードまたはNAAP誘導体をコードするポリヌクレオチド配列をmRNAプローブの作製のためのベクターにクローニングする方法がある。mRNAプローブ作製用ベクター類は、当業者に知られており、市販されており、好適なRNAポリメラーゼと好適な標識されたヌクレオチドとを加えることによって、*in vitro*でRNAプローブを合成するために用い得る。ハイブリダイゼーションプローブ類は、種々のレポーターの集団によって標識され得る。レポーター集団の例としては、³²Pまたは³⁵Sなどの放射性核種が、あるいはアビジン/ビオチン結合系を介してプローブに結合されたアルカリホスファターゼなど酵素標識が挙げられる。

【0280】

NAAPをコードするポリヌクレオチド配列を、NAAPの発現に関係する疾患の診断に用い得る。限定するものではないが、このような疾患には細胞増殖異常、神経障害、発達障害、自己免疫/炎症疾患、および感染症が含まれ、細胞増殖異常には日光角化症、動脈硬化、アテローム硬化、滑液包炎、硬変、肝炎、混合性結合組織病（MCTD）、骨髄線維症、発作性夜間ヘモグロビン尿症、真性多血症、乾癬、原発性血小板血症、並びに、腺癌、白血病、リンパ腫、黒色腫、骨髄腫、肉腫、および奇形癌など癌、具体的には、副腎、膀胱、骨、骨髄、脳、乳房、頸部、胆嚢、神経節、消化管、心臓、腎臓、肝臓、肺、筋肉、卵巣、膵臓、副甲状腺、陰茎、前立腺、唾液腺、皮膚、脾臓、精巣、胸腺、甲状腺、および子宮の癌が含まれ、神経障害には、癲癇、虚血性脳血管障害、脳卒中、脳腫瘍、アルツハイマー病、ピック病、ハンチントン病、痴呆、パーキンソン病およびその他の錐体外路障害、筋萎縮性側索硬化およびその他の運動ニューロン障害、進行性神経性筋萎縮症、網膜色素変性症（色素性網膜炎）、遺伝性運動失調、多発性硬化症および他の脱髄疾患、細菌性およびウイルス性髄膜炎、脳膿瘍、硬膜下膿瘍、硬膜外膿瘍、化膿性頭蓋内血栓性静脈炎、脊髄炎および神経根炎、ウイルス性中枢神経系疾患と、プリオン病（クールー、クロイツフェルト ヤコブ病、およびGerstmann-Straussler-Scheinker症候群を含む）、致死性家族性不眠症、神経系性栄養病および代謝病、神経線維腫症、結節硬化症、小脳網膜血管腫症

10

20

30

40

50

(cerebelloretinal hemangioblastomatosis)、脳3叉神経血管症候群、中枢神経系性精神遅滞および他の発達障害、脳性麻痺、神経骨格異常症、自律神経系障害、脳神経障害、脊髄疾患、筋ジストロフィー他の神経筋障害、末梢神経疾患、皮膚筋炎および多発性筋炎と、遺伝性、代謝性、内分泌性、および中毒性ミオパシーと、重症筋無力症、周期性四肢麻痺、精神障害(気分性、不安性の障害、統合失調症/分裂病)、季節性感情障害(SAD)、静座不能、健忘症、緊張病、糖尿病性ニューロパシー、遅発性ジスキネジア、ジストニー、パラノイド精神病、帯状疱疹後神経痛、トゥーレット病が含まれ、発達障害には尿細管性アシドーシス、貧血、クッシング症候群、軟骨形成不全性小人症、デュシェンヌ/ベッカー型筋ジストロフィー、癲癇、性腺形成異常、WAGR症候群(ウィルムス腫瘍、無虹彩症、尿生殖器異常、精神遅滞)、スミス マジェニス症候群(Smith-Magenis syndrome)、骨髄異形成症候群、遺伝性粘膜上皮異形成、遺伝性角皮症や、シャルコーマリーツース病および神経線維腫症などの遺伝性神経病、甲状腺機能低下症、水頭症や、Sydenham舞蹈病(Sydenham's chorea)および脳性小児麻痺などの発作障害、二分脊椎、無脳症、頭蓋脊椎披裂、先天性緑内障、白内障、感音難聴が含まれ、自己免疫/炎症疾患には、後天性免疫不全症候群(AIDS)、アジソン病(慢性原発性副腎機能不全)、成人呼吸窮迫症候群、アレルギー、強直性脊椎炎、アミロイド症、貧血、喘息、アテローム性動脈硬化、自己免疫性溶血性貧血、自己免疫性甲状腺炎、自己免疫性多腺性内分泌カンジダ症外胚葉ジストロフィー(APECED)、気管支炎、胆嚢炎、接触皮膚炎、クローン病、アトピー性皮膚炎、皮膚筋炎、糖尿病、肺気腫、リンパ球傷害因子性偶発性リンパ球減少症、新生児溶血性疾患(胎児赤芽球症)、結節性紅斑、萎縮性胃炎、糸球体腎炎、グッドパスチャー症候群、痛風、グレーブス病、橋本甲状腺炎、過好酸球増加症、過敏性腸症候群、多発性硬化症、重症筋無力症、心筋または心膜炎、骨関節炎、骨粗しょう症、脾炎、多発性筋炎、乾癬、ライター症候群、リウマチ様関節炎、強皮症、シェーグレン症候群、全身性アナフィラキシー、全身性エリテマトーデス、全身性強皮症、血小板減少性紫斑病、潰瘍性大腸炎、ぶどう膜炎、ウェルナー症候群、癌合併症、血液透析、体外循環、原虫感染症、蠕虫感染症、外傷が含まれ、感染症には、アデノウイルス、アレナウイルス、ブンヤウイルス(bunyavirus)、カリチウイルス(calicivirus)、コロナウイルス、フィロウイルス、ヘパドナウイルス、ヘルペスウイルス、フラビウイルス、オルソミクソウイルス、パルボウイルス、パポバウイルス、パラミキソウイルス、ピコルナウイルス、ポックスウイルス、レオウイルス、レトロウイルス、ラブドウイルス、トガウイルスに分類されるウイルス病原体による感染や、細菌による感染(肺炎球菌感染症、ブドウ球菌、連鎖球菌、バシラス菌、コリネバクテリウム、クロストリジウム、髄膜炎菌、淋菌、リステリア、モラクセラ、キングラ、ヘモフィルス、レジオネラ、百日咳菌や、シゲラ、サルモネラ、カンピロバクターを含むグラム陰性腸内細菌、シュードモナス、ビブリオ、ブルセラ、野兔病菌、エルシニア、バルトネラ、nocardium、放線菌、ミコバクテリウム、spirochaetale、リケッチア、クラミジア、マイコプラズマに分類)、真菌感染(分類はアスペルギルス、プラストミセス、皮膚糸状菌、クリプトコッカス、コクシジオイデス、malassezia、ヒストプラズマ、または他の真菌症起因菌)、寄生虫感染(分類はプラスモディウムすなわちマラリア原虫、寄生性アメーバ、リーシュマニア、トリパノソーマ、トキソプラズマ、ニューモシスチスカリニ、腸内原虫(ジアルジアなど)、トリコモナス、組織線虫(旋毛虫など)、腸管寄生線虫(回虫など)、リンパ管フィラリア線虫、吸虫(住血吸虫など)、および条虫(サナダムシなど))が含まれる。NAAPをコードするポリヌクレオチド配列は、変容したNAAP発現を検出するために患者から採取した体液或いは組織を利用する、サザーン法やノーザン法、ドットプロット法、或いはその他の膜系の技術、PCR法や、ディップスティック(dipstick)、ピン(pin)、およびマルチフォーマットのELISA式アッセイ、およびマイクロアレイに使用可能である。このような定性方法または定量方法は、当分野で公知である。

【0281】

或る特定の態様では、NAAPをコードするヌクレオチド配列群は、関連する障害、特に上記した障害を検出するアッセイ類において有用であり得る。NAAPをコードするヌクレオチド

配列を、標準的な方法で標識化し、ハイブリダイゼーション複合体の形成に好適な条件下、患者から採取した体液或いは組織のサンプルに加えることができる。好適なインキュベーション期間が経過したらサンプルを洗浄し、シグナルを定量して標準値と比較する。患者サンプル中のシグナルの量が、対照サンプルと較べて著しく変化している場合は、該サンプル内の、NAAPをコードするヌクレオチド配列群のレベル変化の存在が、関連する障害の存在を標示する。このようなアッセイは、動物実験、臨床試験における特定の治療効果を推定するため、あるいは個々の患者の治療をモニターするために用いることもできる。

【0282】

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

【0283】

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

【0284】

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

【0285】

NAAPをコードする配列群から設計したオリゴヌクレオチド群の更なる診断的利用には、PCRの利用を含み得る。これらのオリゴマーは、化学的に合成するか、酵素により産生するか、あるいは*in vitro*で産出し得る。オリゴマーは、好ましくはNAAPをコードするポリヌクレオチドの断片、或いはNAAPをコードするポリヌクレオチドと相補的なポリヌクレオチドの断片を含み、最適化した条件下で、特定の遺伝子や条件を識別するために利用される。また、オリゴマーは、やや緩いストリンジェンシー条件下で、近縁のDNAあるいはRNA配列の検出、定量、あるいはその両方のため用いることが可能である。

【0286】

或る実施態様においては、NAAPをコードするポリヌクレオチド配列群に由来するオリゴヌクレオチドプライマー類を用いて、一塩基多型（SNP）を検出し得る。SNPは、多くの場合にヒトの先天性または後天性遺伝病の原因となるような置換、挿入および欠失である。限定するものではないがSNPの検出方法には、SSCP（single-stranded conformation polymorphism）および蛍光SSCP（fSSCP）がある。SSCPでは、NAAPをコードするポリヌクレオチド配列群に由来のオリゴヌクレオチドプライマー類とポリメラーゼ連鎖反応法（PCR）とを用いた、DNAの増幅を行う。このDNAは例えば、病変組織または正常組織、生検サンプル、体液などに由来し得る。DNA内のSNPは、一本鎖形のPCR生成物の2次および3次構造に差異を生じさせる。差異は非変性ゲル中でのゲル電気泳動法を用いて検出可能である。fSSCPでは、オリゴヌクレオチドプライマー類を蛍光的に標識することにより、DNAシーケンシング機などの高処理機器でアンプリマー（amplimer）類の検出が可能になる。更に、インシリコSNP（*in silico* SNP, isSNP）と呼ばれる配列データベース分析法は、一般的なコンセンサス配列に配列されるような個々のオーバーラップするDNA断片群の配列を比

較することにより、多型性を同定し得る。これらのコンピュータベースの方法は、DNAの実験室での調製に、または統計モデルとDNA配列クロマトグラムの自動分析とを用いたシーケンシングのエラーに起因する配列変異を、フィルタリングして除去する。別の態様では、例えば高処理MASSARRAYシステム (Sequenom, Inc., San Diego CA) を用いた質量分析によりSNPを検出し、特徴付ける。

【0287】

NAAPの発現を定量するために用い得る方法には、ヌクレオチドの放射標識またはビオチン標識、対照核酸の共増幅 (coamplification)、および標準曲線から得た結果の補間もある (例えばMelby, P.C. 他(1993) J. Immunol. Methods 159:235244; Duplaa, C. 他(1993) Anal. Biochem. 212:229236を参照)。複数サンプルの定量速度を加速するには、目的のオリゴマーまたはポリヌクレオチドを種々の希釈液中に置き、分光光度法または比色反応によって定量が迅速になるような高処理フォーマットでのアッセイを行い得る。

10

【0288】

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

20

【0289】

別の実施例では、NAAP、NAAPの断片、NAAPに特異的な抗体を、マイクロアレイ上のエレメントとして用いることができる。マイクロアレイを用いて、上記のようなタンパク質-タンパク質相互作用、薬剤-標的相互作用および遺伝子発現プロファイルをモニターまたは測定することが可能である。

【0290】

或る実施態様は、或る組織または細胞タイプの転写イメージを作製する、本発明のポリヌクレオチドの使用に関連する。転写イメージは、特定の組織または細胞タイプによる、遺伝子発現の包括的パターンを表す。包括的遺伝子発現パターンは、所与の条件下で所与の時間に発現した遺伝子の数および相対存在量を定量することにより分析し得る (Seilhame r 他(1993) 米国特許第5,840,484号「Comparative Gene Transcript Analysis」を参照。該特許は特に引用することを以って本明細書の一部となす)。したがって、特定の組織または細胞タイプの転写物または逆転写物全体に、本発明のポリヌクレオチドまたはそれらの相補体をハイブリダイズすることにより、転写イメージを作製し得る。或る実施態様では、本発明のポリヌクレオチドまたはそれらの相補体が1マイクロアレイ上に複数エレメントの1サブセットを持つような高処理フォーマットでハイブリダイゼーションを発生させる。結果として得られる転写イメージは、遺伝子活性のプロファイルを提供し得る。

30

40

【0291】

転写イメージは、組織、細胞株、生検またはその生体サンプルから単離した転写物を用いて作製し得る。転写イメージはしたがって、組織または生検サンプルの場合には in vivo、細胞株の場合には in vitroでの遺伝子発現を反映する。

【0292】

本発明のポリヌクレオチドの発現のプロファイルを作製する転写イメージはまた、in vitroモデル系および薬剤の前臨床評価に、あるいは工業的または天然の環境化合物の毒性試験に関連して使用し得る。全ての化合物は、作用および毒性のメカニズムを標示し、しばしば分子フィンガープリントまたは毒性シグネチャと称される、特徴的な遺伝子発現パ

50

ーンを惹起する (Nuwaysir, E.F. 他 (1999) Mol. Carcinog. 24:153-159、Steiner, S. および N.L. Anderson (2000) Toxicol. Lett. 112-113:467-471、該文献は特に引用を以って本明細書の一部となす)。試験化合物が、既知毒性を有する化合物のシグネチャと同一のシグネチャを有する場合には、毒性特性を共有している可能性がある。フィンガープリントまたはシグネチャは、多数の遺伝子および遺伝子ファミリーからの発現情報を含んでいる場合に、最も有用且つ正確である。理想的には、ゲノム全域にわたる発現の測定が、最高品質のシグネチャを提供する。たとえ、発現が任意の試験された化合物によって変容しない遺伝子があったとしても、それらの発現レベルを残りの発現データを正規化することに使用できるため、それらの遺伝子は重要である。正規化手順は、種々の化合物で処理した後の発現データの比較に有用である。或る毒性シグネチャのエレメント群に遺伝子機能を割り当てることは毒性機序の解釈に役立つが、毒性の予測につながる、シグネチャ群の統計的マッチングには、遺伝子機能の知識は必要でない (例えば2000年2月29日に米国国立環境健康科学研究所 (National Institute of Environmental Health Sciences) より発行されたPress Release 00-02を参照されたい。これについては<http://www.niehs.nih.gov/oc/news/toxchip.htm>で入手可能である)。したがって、毒性シグネチャを用いる中毒学的スクリーニングの際に、全ての発現した遺伝子配列を含めることは、重要且つ望ましい。

10

【0293】

或る実施例では、核酸を有する生体サンプルを試験化合物で処理することにより、この試験化合物の毒性を算定する。処理した生体サンプル中で発現した核酸は、本発明のポリヌクレオチドに特異的な1つ以上のプローブでハイブリダイズし、それによって本発明のポリヌクレオチドに対応する転写レベルを定量し得る。処理した生体サンプル中の転写レベルを、未処理生体サンプル中のレベルと比較する。両サンプルの転写レベルの差は、処理済サンプル中で試験化合物が引き起こす毒性反応を標示する。

20

【0294】

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

30

40

【0295】

プロテオームのプロフィールは、NAAPに特異的な抗体を用いてNAAP発現レベルを定量することによっても作成可能である。或る実施態様では、マイクロアレイ上のエレメントとしてこれら抗体を用い、マイクロアレイをサンプルに曝して各アレイエレメントへのタンパク質結合レベルを検出することにより、タンパク質発現レベルを定量する (Lueking, A.

50

他(1999) Anal. Biochem. 270:103-111; Mendozze, L.G. 他(1999) Biotechniques 27:778-788)。検出は当分野で既知の様々な方法で行うことができ、例えば、チオール反応性またはアミノ反応性蛍光化合物とサンプル中のタンパク質を反応させ、各アレイエレメントにおける蛍光結合の量を検出し得る。

【0296】

プロテオームレベルでの毒性シグネチャも中毒学的スクリーニングに有用であり、転写レベルでの毒性シグネチャと並行に分析するべきである。いくつかの組織のいくつかのタンパク質については、転写物の存在量とタンパク質の存在量との相関が乏しいので(Anderson, N.L. および J. Seilhamer (1997) Electrophoresis 18:533-537)、転写イメージにはそれ程影響しないがプロテオームのプロフィールを改変するような化合物の分析において、プロテオーム毒性シグネチャは有用たり得る。更に、体液中の転写物の分析はmRNAの急速な分解のために困難なので、プロテオームのプロフィール作成はこのような場合により信頼性が高いと思われ、情報価値があり得る。

10

【0297】

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

20

【0298】

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

【0299】

マイクロアレイは、本技術分野で既知の方法で調製し、使用し、分析する(例えばBrennan, T.M. 他(1995)米国特許第5,474,796号、Schena, M. 他(1996) Proc. Natl. Acad. Sci. USA 93:10614-10619、Baldeschweiler 他(1995) PCT出願第W095/251116号、Shalon, D. 他(1995) PCT出願第W095/35505号、Heller, R.A. 他(1997) Proc. Natl. Acad. Sci. USA 94:2150-2155、Heller, M.J. 他(1997)米国特許第5,605,662号を参照)。様々なタイプのマイクロアレイが公知であり、詳細については、DNA Microarrays: A Practical Approach, M.Schena, 編集。(1999) Oxford University Press, Londonに記載されている。該文献は、特に引用することを以って本明細書の一部となす。

30

【0300】

本発明の別の実施態様ではまた、NAAPをコードする核酸配列群を用いて、天然ゲノム配列をマッピングするのに有用なハイブリダイゼーションプローブ群を産生し得る。コード配列または非コード配列のいずれかを用いることができ、いくつかの例では、コード配列より非コード配列が好ましい。例えば、多重遺伝子族のメンバー内でのコード配列の保存により、染色体マッピング中に望ましくないクロスハイブリダイゼーションが生じる可能性がある。配列は、或る特定の染色体に、または或る染色体の或る特定領域に、または人為形成の染色体、例えば、ヒト人工染色体(HAC)、酵母人工染色体(YAC)、細菌人工染色体(BAC)、細菌P1産物、あるいは単一染色体cDNAライブラリ群に対してマッピングされる(例えばHarrington, J.J. 他(1997) Nat. Genet. 15:345-355; Price, C.M. (1993) Blood Rev. 7:127134; Trask, B.J. (1991) Trends Genet. 7:149154を参照)。一度マッピングすると、本発明の核酸配列群を用いて、例えば或る病状の遺伝を特定染色体領域の遺伝とまたは制限酵素断片長多型(RFLP)と関連させるような遺伝子連鎖地図を開発し得る

40

50

(例えば、Lander, E.S.およびD. Botstein (1986) Proc. Natl. Acad. Sci. USA 83:7353-7357を参照)。

【0301】

蛍光原位ハイブリッド形成法 (FISH) は、他の物理的および遺伝的地図データと関連し得る (例えばHeinz-Ulrich, 他 (1995) in Meyers, 前出、965-968ページを参照)。遺伝的地図データの例は、種々の科学雑誌あるいはOnline Mendelian Inheritance in Man (OMIM) のウェブサイトに見られる。物理的染色体地図上の、NAAPをコードする遺伝子の位置と、特定の疾患との相関性、あるいは特定の疾患に対する素因との相関性は、この疾患と関連するDNAの領域の決定に役立ち得るため、位置を決定するクローニングの作業を促進し得る。

10

【0302】

確定した染色体マーカー類を用いた連鎖分析などの物理的マッピング技術、および染色体標本の原位ハイブリッド形成法を用いて、遺伝的地図を拡張することができる。例えばマウスなど別の哺乳類の染色体上に遺伝子を配置することにより、正確な染色体上の遺伝子座が未知でも、関連するマーカー類をしばしば明らかにし得る。この情報は、位置クローニングなどの遺伝子発見技術を用いて疾患遺伝子を探る研究者にとって価値がある。いったん疾患または症候群に關与する遺伝子 (群) が、血管拡張性失調症の11q22-23領域など、特定のゲノム領域への遺伝的連鎖によって、大まかに位置決めがなされると、該領域にマップされる任意の配列は、更なる調査のための関連遺伝子あるいは調節遺伝子を提示している可能性がある (例えばGatti, R.A. 他 (1988) Nature 336:577-580を参照)。転座、反転などに起因する、健常者、保有者、罹病者の三者間における染色体位置の相違を検出する場合にも、本発明のヌクレオチド配列を用い得る。

20

【0303】

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

【0304】

別の薬剤スクリーニング方法は、目的のタンパク質に対して好適な結合親和性を有する化合物を高い処理能力でスクリーニングするために用いられる (例えばGeysen, 他 (1984) PC T 出願番号 W084/03564を参照)。この方法においては、多数の様々な低分子の試験用化合物を固体基板上で合成する。試験用化合物は、NAAP、或いはその断片と反応してから洗浄される。次に、結合したNAAPを本技術分野で周知の方法で検出する。精製したNAAPはまた、上記した薬剤スクリーニング技術に用いるプレート上に直接コーティングできる。別法では、非中和抗体を用いてペプチドを捕捉し、ペプチドを固体支持物に固定することもできる。

30

【0305】

別の実施例では、NAAPと特異結合可能な中和抗体がNAAPとの結合について試験化合物と競合する、競合的薬剤スクリーニングアッセイを用いることができる。この方法では、抗体を用いて、1つ以上の抗原決定基をNAAPと共有するどのペプチドの存在をも検出できる。

40

【0306】

別の実施例では、NAAPをコードするヌクレオチド配列を、将来に開発される分子生物学技術であり、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) に依存する新技術に用い得る。

【0307】

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

【0308】

50

本明細書において開示した全ての特許、特許出願及び刊行物、特に米国特許出願第60/257,714号、第60/260,081号、第60/262,302号、第60/266,088号、出願 [代理人整理番号 (Attorney Docket No.) PF-1249 P, ファイル日は2001年10月29日]、および出願第60/263,823号は、言及することをもって本明細書の一部となす。

【0309】

(実施例)

1 cDNAライブラリの作製

Incyte cDNA群の由来は、LIFESEQ GOLDデータベース (Incyte Genomics, Palo Alto CA) に記載されたcDNAライブラリ群である。幾つかの組織はホモジナイズしてグアニジニウムイソチオシアネート溶液に溶解し、他の組織はホモジナイズしてフェノールにまたは変性剤群の好適な混合液に溶解した。混合液の1例であるTRIZOL (Life Technologies) は、フェノールとグアニジンイソチオシアネートとの単相溶液である。結果として得た溶解物は、塩化セシウムのクッション液の上に重層して遠心分離するか、クロロフォルムで抽出した。イソプロパノールか、酢酸ナトリウムとエタノールか、いずれか一方、あるいは別の慣例的方法で、溶解物からRNAを沈殿させた。

10

【0310】

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

20

【0311】

場合によってはStratagene社にRNAを提供し、対応するcDNAライブラリ群をStratagene社が作製した。そうでない場合は、UNIZAPベクターシステム (Stratagene) またはSUPERSCRIPTプラスミドシステム (Life Technologies) を用いて本技術分野で既知の推奨方法または類似の方法でcDNAを合成し、cDNAライブラリを作製した (前出のAusubel, 1997, 5.1-6.6ユニットなどを参照)。逆転写は、オリゴd(T) またはランダムプライマーを用いて開始した。合成オリゴヌクレオチドアダプターを二本鎖cDNAに連結反応させ、好適な制限酵素または酵素群でcDNAを消化した。殆どのライブラリに対して、cDNAのサイズ選択 (300 ~ 1000 bp) は、SEPHACRYL S1000、SEPHAROSE CL2BまたはSEPHAROSE CL4Bカラムクロマトグラフィー (Amersham Pharmacia Biotech)、あるいは調製用アガロースゲル電気泳動法を用いて行った。cDNAは好適なプラスミドのポリリンカーの、適合する制限酵素部位にライゲーションされた。好適なプラスミドは、例えばPBLUESCRIPTプラスミド (Stratagene)、PSPORT1プラスミド (Life Technologies) PCDNA2.1プラスミド (Invitrogen, Carlsbad CA)、PBK-CMVプラスミド (Stratagene)、PCR2-TOPOTAプラスミド (Invitrogen)、PCMV-ICISプラスミド (Stratagene)、pIGEN (Incyte Genomics, Palo Alto CA)、pRARE (Incyte Genomics)、またはpINCY (Incyte Genomics)、またはこれらの誘導体である。組換えプラスミドは、Stratagene社のXL1-Blue、XL1-BlueMRFまたはSOLR、あるいはLife Technologies社のDH5、DH10BまたはElectroMAX DH10Bなど適格な大腸菌細胞に形質転換した。

30

40

【0312】

2 cDNAクローンの単離

実施例1のようにして得たプラスミドの、宿主細胞からの回収は、UNIZAPベクターシステム (Stratagene) を用いた *in vivo* 切除によって、あるいは細胞溶解によって行った。プラスミドを精製する方法は、MagicまたはWIZARD Minipreps DNA精製システム (Promega)、AGTC Miniprep精製キット (Edge Biosystems, Gaithersburg MD)、QIAGEN社のQIAWELL 8 Plasmid、QIAWELL 8 Plus PlasmidおよびQIAWELL 8 Ultra Plasmid 精製システム、R.E.A.L. Prep 96プラスミド精製キットの中から少なくとも1つを用いた。プラスミドは、沈殿させた後、0.1mlの蒸留水に再懸濁して、凍結乾燥して或いは凍結乾燥しないで4

50

で保管した。

【0313】

別法では、高処理フォーマットで直接結合PCR法を用い、宿主細胞溶解物からプラスミドDNAを増幅した(Rao, V.B. (1994) Anal. Biochem. 216:1-14)。宿主細胞の溶解および熱サイクリング過程は、単一反応混合液中で行った。サンプルを加工し、それを384穴プレート内で保管し、増幅したプラスミドDNAの濃度をPICOGREEN色素(Molecular Probes, Eugene OR)およびFLUOROSKANII蛍光スキャナ(Labsystems Oy, Helsinki, Finland)を用いて蛍光分析的に定量した。

【0314】

3 シークエンシングおよび分析

実施例2に記載したようにプラスミドから回収したIncyte cDNAを、以下に示すようにシークエンシングした。cDNAのシークエンス反応は、標準的方法あるいは高処理装置、例えばABI CATALYST 800 サーマルサイクラー(Applied Biosystems)またはPTC-200 サーマルサイクラー(MJ Research)を、HYDRAマイクロディスペンサー(Robbins Scientific)またはMICROLAB 2200(Hamilton)液体転移システムと併用して処理した。cDNAのシークエンス反応は、Amersham Pharmacia Biotech社が提供する試薬、またはABIシークエンシングキット、例えばABI PRISM BIGDYE Terminator cycle sequencing ready reaction kit(Applied Biosystems)の試薬を用いて準備した。cDNAのシークエンス反応の電気泳動的分離には、また、標識したポリヌクレオチドの検出には、MEGABACE 1000 DNAシークエンシングシステム(Molecular Dynamics)か、標準ABIプロトコルと塩基対呼び出しソフトウェアとを用いるABI PRISM 373または377シークエンシングシステム(Applied Biosystems)か、あるいはその他の本技術分野で既知の配列解析システムを用いた。cDNA配列内のリーディングフレームは、標準的方法(前出のAusubel, 1997, 7.7ユニットに概説)を用いて同定した。いくつかのcDNA配列を選択し、実施例8に開示する技術で配列を伸長させた。

【0315】

Incyte cDNAに由来するポリヌクレオチド配列は、ベクター、リンカーおよびポリ(A)配列を除去し、あいまいな塩基をマスクすることによって有効性を確認した。その際、BLASTと、動的プログラミングと、隣接ジヌクレオチド頻度分析とに基づく、アルゴリズムとプログラムとを用いた。次に、Incyte cDNA配列またはそれらの翻訳の問い合わせを、以下のデータベース群に対して行った。すなわち、選抜した公共のデータベース群(例えばGenBankの霊長類、げっ歯類、哺乳類、脊椎動物、真核生物のデータベースと、BLOCKS、PRINTS、DOMO、PRODOM)と、ヒト、ラット、マウス、線虫(Caenorhabditis elegans)、出芽酵母(Saccharomyces cerevisiae)、分裂酵母(Schizosaccharomyces pombe)および驚口瘡カンジダ(Candida albicans)からの配列群を持つPROTEOMEデータベース群(Incyte Genomics, Palo Alto CA)、および、隠れマルコフモデル(HMM)ベースのタンパク質ファミリーデータベース群(PFAM等)である(HMMは、遺伝子ファミリーのコンセンサス1次構造を分析する確率的アプローチである。例えば、Eddy, S.R. (1996) Cuff. Opin. Struct. Biol. 6:361-365等を参照)。問合せは、BLAST、FASTA、BLIMPS、およびHMMERに基づくプログラムを用いて行った。Incyte cDNA配列を構築し、完全長のポリヌクレオチド配列を産出した。あるいは、GenBank cDNA群、GenBank EST群、ステッチされた配列群、ストレッチされた配列群、またはGenscan予測コード配列群(実施例4および5を参照)を用い、Incyte cDNAの集団を完全長まで伸長させた。PhredとPhrapとConsedとに基づくプログラムを用いて構築し、GenMarkとBLASTとFASTAとに基づくプログラムを用いて、cDNAの集団を、オープンリーディングフレームについてスクリーニングした。完全長ポリヌクレオチド配列を翻訳し、対応する完全長ポリペプチド配列を誘導した。あるいは、本発明のポリペプチドは、完全長翻訳ポリペプチドの任意のメチオニン残基で開始し得る。完全長ポリペプチド配列群の続いた分析としての問い合わせを、GenBankタンパク質データベース群(genpept)、SwissProt、PROTEOMEデータベース群、BLOCKS、PRINTS、DOMO、PRODOMおよびProsites等のデータベースや、PFAM等の隠れマルコフモデル(HMM)ベー

10

20

30

40

50

スのタンパク質ファミリーデータベース群に対し行った。完全長ポリヌクレオチド配列はまた、MACDNASIS PROソフトウェア（日立ソフトウェアエンジニアリング, South San Francisco CA）およびLASERGENEソフトウェア（DNASTAR）を用いて分析した。ポリヌクレオチドとポリペプチドとの配列アラインメントは、MEGALIGNマルチシーケンシングアラインメントプログラム（DNASTAR）に組み込まれたCLUSTALアルゴリズムが指定する、デフォルトパラメータを用いて作製する。これは、アラインメントした配列間の一致率も計算する。

【0316】

表7は、Incyte cDNAと完全長配列との分析と構築とに利用したツールとプログラムとアルゴリズムとの概略と、適用可能な説明、参照文献、閾値パラメータを示す。用いたツール、プログラムおよびアルゴリズムを表7の列1に、それらの簡単な説明を列2に示す。列3は好適な参照文献であり、全ての文献は引用を以って全文を本明細書の一部となす。適用可能な場合には、列4は、2つの配列が一致する強さを評価するために用いた、スコア、確率値などのパラメータを示す（スコアが高いほど、または確率値が低いほど、2配列間の同一性が高い）。

10

【0317】

完全長ポリヌクレオチド配列およびポリペプチド配列の構築および分析に用いる上記のプログラム群は、SEQ ID NO:17-32のポリヌクレオチド配列断片の同定にも利用できる。ハイブリダイゼーション及び増幅技術に有用である約20～約4000ヌクレオチドの断片群を、表4の列2に示した。

【0318】

20

4 ゲノムDNAからのコード配列の同定および編集

推定上の核酸関連タンパク質は、先ず公共のゲノム配列データベース（例えばgbpriやgbhtg）に対しGenscan遺伝子同定プログラムを実行して同定した。Genscanは汎用遺伝子同定プログラムであり、様々な生物からのゲノムDNA配列を分析する（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. 8:346-354参照）。このプログラムは予測されたエキソン群を連結し、メチオニンから終止コドンに及ぶ、構築されたcDNA配列を形成する。Genscanの出力は、ポリヌクレオチドおよびポリペプチド配列のFASTAデータベースである。Genscanが一度に分析する配列の最大範囲は、30 kbに設定した。これらのGenscan予測cDNA配列の内、どの配列が核酸関連分子をコードするかを決定するために、コードされるポリペプチドを、PFAMモデル群に対し核酸関連分子について問合せて分析した。潜在的な核酸関連タンパク質はまた、既に核酸関連タンパク質としてアノテーションが付けられたIncyte cDNA配列に対する相同性を基に同定された。これら選択したGenscan予測配列を、次にBLAST分析により、公共データベースgenpeptおよびgbpriと比較した。必要であれば、genpeptからのトップBLASTヒットと比較することによりGenscan予測配列を編集し、余分なまたは省略されたエキソンなど、Genscanが予測した配列におけるエラーを補正した。BLAST分析はまた、Genscan予測配列の、いかなるIncyte cDNAまたは公共cDNAカバレッジ（coverage）の発見にも用いられ、したがって転写の証拠を提供した。Incyte cDNAカバレッジが利用できた場合には、この情報を用いてGenscan予測配列を補正または確認した。完全長ポリヌクレオチド配列は、実施例3に記載した構築プロセスを用いて、Incyte cDNA配列および/または公共cDNA配列でGenscan予測コード配列を構築して得た。あるいは、完全長ポリヌクレオチド配列は、編集後または非編集のGenscan予測コード配列に、完全に由来する。

30

40

【0319】

5 ゲノム配列データのcDNA配列データとの統合

スティッチ配列（Stitched Sequence）

部分cDNA配列群を伸長させるため、実施例4に記載のGenscan遺伝子同定プログラムが予測したエキソン群を用いた。実施例3に記載したように構築した部分cDNA群を、ゲノムDNAにマッピングし、また、関連するcDNA群と、1つ以上のゲノム配列からGenscan予測されたエキソン群とを有するクラスター群に分解した。cDNAとゲノムとの情報を統合すべく、

50

グラフ理論および動的プログラミングに基づく或るアルゴリズムを用いて各クラスターを分析し、潜在的スプライス変異配列群を産生した。配列群を続いて確認、編集または伸長し、完全長配列を創出した。区間の全長が、2つ以上の配列に在るような配列区間群をクラスター内で同定し、そのように同定された区間群は、推移性により、等しいと考えた。例えば、或る区間が、1つのcDNAと2つのゲノム配列とに在る場合、3つの区間は全て等しいと考えた。このプロセスにより、無関係だが連続したゲノム配列群を、cDNA配列で結び合わせて架橋し得る。このようにして同定した区間を、それらの親配列 (parent sequences) に沿って現われる順にスティッチアルゴリズムで「縫い合わせ」、可能な限り最長の配列と変異配列群とを産生した。1種類の親配列に沿って続く区間間の連鎖 (cDNA - cDNAまたはゲノム配列 - ゲノム配列) は、親の種類を変える連鎖 (cDNA - ゲノム配列) より優先した。結果として得たスティッチ配列群を翻訳し、BLAST分析で公共データベース genpeptおよびgbpriと比較した。Genscanが予測した不正確なエキソン群は、genpeptからのトップBLASTヒットとの比較により補正した。必要な場合、追加cDNA配列群を用いるかゲノムDNAの検査により、配列群を更に伸長させた。

【0320】

ストレッチ配列 (Stretched Sequence)

部分DNA配列群を、BLAST分析に基づく1アルゴリズムにより完全長まで伸長した。先ず、BLASTプログラムを用いて、GenBankの霊長類、げっ歯類、哺乳類、脊椎動物および真核生物のデータベースなどの公共データベースに対し、実施例3に記載されたように構築された部分cDNAを問い合わせた。次に、最も近いGenBankタンパク質相同体を、BLAST分析により、Incyte cDNA配列または実施例4に記載のGenScanエキソン予測配列のいずれかと比較した。結果として得られる高スコアリングセグメント対 (HSP) を用いてキメラタンパク質を産生し、翻訳した配列をGenBankタンパク質相同体上にマッピングした。元のGenBankタンパク質相同体に対し、キメラタンパク質内では挿入または欠失が起こり得る。GenBankタンパク質相同体、キメラタンパク質またはその両方をプローブとして用い、公共のヒトゲノムデータベースから相同ゲノム配列を検索した。このようにして、部分的なDNA配列を、相同ゲノム配列の付加によりストレッチすなわち伸長した。結果として得られるストレッチ配列を検査し、完全遺伝子を含んでいるか否かを決定した。

【0321】

6 NAAPをコードするポリヌクレオチドの染色体マッピング

SEQ ID NO:17 - 32を構築するために用いた配列群を、BLAST他のSmith-Watermanアルゴリズムの実装群を用いて、Incyte LIFESEQデータベースおよび公共ドメインデータベース群の配列群と比較した。SEQ ID NO:17 - 32と一致するこれらのデータベースの配列を、Phrapなどの構築アルゴリズムを使用して、連続的配列及びオーバーラップした配列のクラスターに構築した (表7)。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などの公的な情報源から入手可能な放射線ハイブリッドおよび遺伝地図データを用いて、いずれかのクラスター化された配列が既にマッピングされているかを判定した。マッピングされた配列が或るクラスターに含まれている場合、そのクラスターの全配列が、個々の配列番号と共に、地図上の位置に割り当てられた。

【0322】

地図上の位置は、ヒト染色体の範囲または間隔として表される。センチモルガン間隔の地図上の位置は、染色体のpアームの末端に対して測定する (センチモルガン (cM) は、染色体マーカー間の組換え頻度に基づく計測単位である。平均して、1 cMは、ヒトDNAの1メガベース (Mb) にほぼ等しい。尤も、この値は、組換えのホットスポットおよびコールドスポットに起因して広範囲に変化する)。cM距離は、各クラスター内に配列が含まれる放射線ハイブリッドマーカー類に対して境界を提供するGenethonによってマッピングされた遺伝マーカー群に基づく。NCBI「GeneMap'99」 (<http://www.ncbi.nlm.nih.gov/genemap/>) など一般個人が入手可能なヒトゲノム地図などの資源を用いて、既に同定された疾患遺伝子類が、上記の区間内若しくは近傍にマップされているかを判定できる。このようにして、SEQ ID NO:20は、8番染色体の125.80から140.60センチモルガンの区間内にマップ

され、またSEQ ID NO:28は、19番染色体の41.7から49.4センチモルガンの区間内にマップされた。

【0323】

7 ポリヌクレオチド発現の分析

ノーザン分析は、遺伝子の転写物の存在を検出するために用いられる実験技術であり、特定の細胞種または組織からのRNAが結合している膜への、標識されたヌクレオチド配列のハイブリダイゼーションに關与する（例えば前出のSambrook, 7章、同Ausubel (1995) 4章および16章を参照）。

【0324】

BLASTを応用した類似のコンピュータ技術を用いて、GenBankやLIFESEQ (Incyte Genomics) などのcDNAデータベースにおいて、同一または関連分子を検索した。ノーザン分析は、いくつかの膜系ハイブリダイゼーションよりも非常に速い。更に、任意の特定の一致を厳密なあるいは相同的なものとして分類するか否かを決定するため、コンピュータ検索の感度を修正することができる。検索の基準は積スコアであり、次式で定義される。

【0325】

【数1】

(BLAST スコア×配列一致率)

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

【0326】

積スコアは、2つの配列間の類似度と、配列が一致する長さとの両方を考慮している。積スコアは、0 ~ 100の正規化された値であり、次のようにして求める。BLASTスコアにヌクレオチドの配列一致率を乗じ、その積を2つの配列の短い方の長さの5倍で除する。BLASTスコアを計算するには、或る高スコアリングセグメント対(HSP)内の一致する各塩基に+5のスコアを割り当て、各不一致塩基に-4を割り当てる。2つの配列は、2以上のHSPを共有し得る(ギャップにより隔離される)。2以上のHSPがある場合には、最高BLASTスコアのセグメント対を用いて積スコアを計算する。積スコアは、断片的オーバーラップとBLASTアラインメントの質とのバランスを表す。例えば積スコア100は、比較した2つの配列の短い方の長さ全体にわたって100%一致する場合のみ得られる。積スコア70は、一端が100%一致し、70%オーバーラップしているか、他端が88%一致し、100%オーバーラップしているかのいずれかの場合に得られる。積スコア50は、一端が100%一致し、50%オーバーラップしているか、79%一致し、100%オーバーラップしているかのいずれかの場合に得られる。

【0327】

或いは、NAAPをコードするポリヌクレオチド配列を、由来する組織に対して分析する。例えば幾つかの完全長配列は、少なくとも一部は、オーバーラップするIncyte cDNA配列群を用いて構築される(実施例3を参照)。各cDNA配列は、ヒト組織から作製されたcDNAライブラリに由来する。各ヒト組織は、以下の臓器/組織カテゴリーの1つに分類される。すなわち心血管系、結合組織、消化器系、胎芽構造、内分泌系、外分泌腺、女性生殖器、男性生殖器、生殖細胞、血液および免疫系、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/混合性または尿路である。各カテゴリーのライブラリ数を数えて、全カテゴリーの総ライブラリ数で除する。同様に、各ヒト組織は、以下の疾患/状況カテゴリーすなわち癌、細胞株、発達、炎症、神経性、外傷、心血管、プール、その他の1つに分類される。各カテゴリーのライブラリ数を数えて、全カテゴリーの総ライブラリ数で除する。得られるパーセンテージは、NAAPをコードするcDNAの、組織特異的または疾患特異的な発現を反映する。cDNA配列およびcDNAライブラリ/組織の情報は、LIFESEQ GOLD データベース(Incyte Genomics, Palo Alto CA)から得ることができる。

【0328】

8 NAAPをコードするポリヌクレオチドの伸長

10

20

30

40

50

完全長のポリヌクレオチド配列はまた、完全長分子の適切な断片から設計したオリゴヌクレオチドプライマー群を用いて該断片を伸長させて産生した。或る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、別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。開始プライマー群の設計にはOLIGO 4.06ソフトウェア (National Biosciences) 或いは別の適切なプログラムを用い、長さが約22~30ヌクレオチド、GC含有率が約50%以上となり、約68~約72の温度で標的配列にアニーリングするようにした。ヘアピン構造およびプライマー-プライマー二量体を生ずるようなヌクレオチド群の伸長は、全て回避した。

【0329】

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

【0330】

高忠実度の増幅が、当業者によく知られている方法を利用したPCR法によって得られた。PCRは、PTC-200 サーマルサイクラー (MJ Research, Inc.) を用いて96穴プレート内で行った。反応混合液は、鋳型DNAを有し、また、200 nmolの各プライマーを有する。また、 Mg^{2+} と $(NH_4)_2SO_4$ と2-メルカプトエタノールとを含有する反応バッファート、Taq DNAポリメラーゼ (Amersham Pharmacia Biotech) と、ELONGASE酵素 (Life Technologies) と、Pfu DNAポリメラーゼ (Stratagene) とを含有する。プライマー対であるPCI AとPCI Bとに対し、以下のパラメータで増幅した。

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

別法では、プライマー対であるT7とSK+に対し、以下のパラメータで増幅した。

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

各ウェルのDNA濃度は、 $1 \times TE$ 及び $0.5 \mu l$ の希釈していないPCR産物に溶解した $100 \mu l$ のPICOGREEN定量試薬(0.25(v/v) PICOGREEN; Molecular Probes, Eugene OR)を不透明な蛍光度計プレート(Corning 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した。サンプルの蛍光を計測してDNAの濃度を定量すべく、プレートをFluoroskan II (Labsystems Oy, Helsinki, Finland)でスキャンした。反応混合物のアリコット5~10 μl を1%アガロース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判定した。

【0331】

伸長したヌクレオチドは、脱塩および濃縮して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、音波処理またはせん断し、pUC 18ベクター (Amersham Pharmacia Biotech) への再連結を行った。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を低濃度(0.6~0.8%)のアガロースゲル上で分離し、断片を切除し、寒天をAgar ACE (Promega) で消化した。伸長させたクローンをT4リガーゼ (New England Biolabs, Beverly MA) を用いてpUC 18ベクター (Amersham Pharmacia Biotech) に再連結し、Pfu DNAポリメ

ラーゼ (Stratagene) で処理して制限部位のオーバーハングを満たし、適格な大腸菌細胞に形質移入した。形質移入した細胞群の選択を抗生物質を含む培地で行い、それぞれのコロニーを採取し、LB/2Xカルベニシリン培養液中の384ウェルプレート群に37で一晚培養した。

【0332】

細胞を溶解し、Taq DNAポリメラーゼ (Amersham Pharmacia Biotech) およびPfu DNAポリメラーゼ (Stratagene) を用いて以下の手順でDNAをPCR増幅した。

- ステップ1 94 で3分間
- ステップ2 94 で15秒間
- ステップ3 60 で1分間
- ステップ4 72 で2分間
- ステップ5 ステップ2、3、及び4を29回繰り返す
- ステップ6 72 で5分間
- ステップ7 4 で保存

10

DNAの定量化には、上記したようにPICOGREEN試薬 (Molecular Probes) を用いた。DNAの回収率が低いサンプルは、上記と同一の条件を用いて再増幅した。サンプルは20%ジメチルスルホキシド (1:2, v/v) で希釈し、DYENAMIC エネルギートランスファー シークエンシングプライマー、およびDYENAMIC DIRECT kit (Amersham Pharmacia Biotech) またはABI PRISM BIGDYE ターミネーターサイクル シークエンシングレディ反応キット (Terminator cycle sequencing ready reaction kit) (Applied Biosystems) を用いてシークエンシングした。

20

【0333】

同様に、上記手順を用いて完全長ポリヌクレオチド配列を検証した。あるいは、完全長ポリヌクレオチドを用い、上記手順で、そのような伸長のために設計したオリゴヌクレオチド類と、或る適切なゲノムライブラリとを用いて5'調節配列を得た。

【0334】

9 個々のハイブリダイゼーションプローブの標識及び使用

SEQ ID NO:17 - 32由来のハイブリダイゼーションプローブ群を用いて、cDNA、mRNA、またはゲノムDNAをスクリーニングする。約20塩基対からなるオリゴヌクレオチドの標識について特に記載するが、より大きなヌクレオチド断片に対しても本質的に同一の手順が用いられる。オリゴヌクレオチドは、OLIGO 4.06ソフトウェア (National Biosciences) 等の最新ソフトウェアを用いて設計し、各オリゴマー50 pmolと、[$^{-32}$ P]アデノシン3リン酸 (Amersham Pharmacia Biotech) 250 μ Ciと、T4ポリヌクレオチドキナーゼ (DuPont NEN, Boston MA) とを化合させることにより標識する。標識したオリゴヌクレオチドは、SEPHADEX G-25超細繊維分子サイズ排除デキストラン ピーズカラム (Amersham Pharmacia Biotech) を用いて実質的に精製する。Ase I、Bgl II、Eco RI、Pst I、XbaIまたはPvu II (DuPont NEN) のいずれか1つのエンドヌクレアーゼで消化されたヒトゲノムDNAの、典型的な膜ベースのハイブリダイゼーション解析において、毎分 10^7 カウントの標識されたプローブを含むアリコットを用いる。

30

【0335】

各消化物から得たDNAは、0.7%アガロースゲル上で分画してナイロン膜 (Nytran Plus, Schleicher & Schuell, Durham NH) に移す。ハイブリダイゼーションは、40で16時間行う。非特異的シグナル群を除去するため、最大で例えば0.1 \times クエン酸ナトリウム食塩水および0.5%ドデシル硫酸ナトリウムの条件下で、プロット群を室温で順次洗浄する。オートラジオグラフィーまたはそれに代わるイメージング手段を用いてハイブリダイゼーションパターンを視覚化し、比較する。

40

【0336】

10 マイクロアレイ

或るマイクロアレイ上でのアレイエレメント群の結合または合成は、フォトリソグラフィ、圧電印刷 (インクジェット印刷、前出のBalteschweilerなどを参照)、機械的マイクロ

50

スポットティング技術およびこれらから派生した技術を用いて達成し得る。上記各技術において基板は、均一な、非多孔性の表面を持つ固体とすべきである (Schena (1999) 前出)。推奨する基板には、シリコン、シリカ、スライドガラス、ガラスチップおよびシリコンウエハがある。あるいは、ドットプロット法またはスロットプロット法に類似した手順を利用し、熱的、紫外線的、化学的または機械的結合手順を用いて基板の表面にエレメントを配置および結合させてもよい。通常のアレイは、利用可能な、当業者に公知の方法と機械とを用いて作製でき、任意の適正数のエレメントを有し得る (例えば Schena, M. 他 (1995) Science 270:467-470、Shalon, D. 他 (1996) Genome Res. 6:639-645、Marshall, A. および J. Hodgson (1998) Nat. Biotechnol. 16:27-31を参照)。

【0337】

完全長cDNA、発現配列タグ (EST)、またはそれらの断片またはオリゴマーが、マイクロアレイのエレメントと成り得る。ハイブリダイゼーションに好適な断片またはオリゴマーを、LASERGENEソフトウェア (DNASTAR) など本技術分野で公知のソフトウェアを用いて選択することが可能である。アレイエレメント群を、生体サンプル中のポリヌクレオチド群とハイブリダイズする。生体サンプル中のポリヌクレオチドは、検出を容易にするために蛍光標識などの分子タグに抱合させる。ハイブリダイゼーション後、生体サンプルからのハイブリダイズされていないヌクレオチドを除去し、蛍光スキャナを用いて各アレイエレメントでのハイブリダイゼーションを検出する。あるいは、レーザ脱離および質量スペクトロメトリを用いてもハイブリダイゼーションを検出し得る。マイクロアレイ上の或るエレメントにハイブリダイズする各ポリヌクレオチドの、相補性の度合と相対存在度とを算定し得る。一実施態様におけるマイクロアレイの調整および使用について、以下に詳述する。

【0338】

組織または細胞サンプルの調製

全RNAを組織サンプル群から単離するためグアニジニウム チオシアネート法を用い、ポリ (A)⁺RNAを精製するためオリゴ (dT) セルコース法を用いる。各ポリ (A)⁺RNAサンプルを逆転写するため、MMLV逆転写酵素を用い、また、0.05pg/ μ lのオリゴ (dT) プライマー (21mer)、1 \times 第一鎖バッファー、0.03unit/ μ lのRNアーゼ阻害因子、500 μ MのdATP、500 μ MのdGTP、500 μ MのdTTP、40 μ MのdCTP、40 μ MのdCTP-Cy3 (BDS) またはdCTP-Cy5 (Amersham Pharmacia Biotech) を用いる。逆転写反応は、GEMBRIGHTキット (Incyte) を用い、200ngのポリ (A)⁺RNAを含有する体積25mlで行う。特異的対照ポリ (A)⁺RNA群は、非コード酵母ゲノムDNAから *in vitro* 転写により合成する。37 $^{\circ}$ Cで2時間インキュベートした後、各反応サンプル (1つはCy3、もう1つはCy5標識) は、2.5mlの0.5M水酸化ナトリウムで処理し、85 $^{\circ}$ Cで20分間インキュベートし、反応を停止させてRNAを分解させる。サンプル群の精製には、2つの連続するCHROMA SPIN 30ゲル濾過スピナラム (CLONTECH Laboratories, Inc. (CLONTECH), Palo Alto CA) を用いる。混合後、2つの反応サンプルのエタノール沈殿を、1mlのグリコーゲン (1mg/ml)、60mlの酢酸ナトリウム、および300mlの100%エタノールで行う。サンプルは次に、SpeedVAC (Savant Instruments Inc., Holbrook NY) を用いて完全に乾燥させ、14 μ lの5 \times SSC / 0.2% SDS中で再懸濁する。

【0339】

マイクロアレイの調製

本発明の配列を用いて、アレイエレメントを作製する。各アレイエレメントは、クローン化したcDNAインサート群により、ベクター含有細菌細胞群から増幅する。PCR増幅は、cDNAインサートの側面に位置するベクター配列に相補的なプライマーを用いる。30サイクルのPCRによって、1~2ngの初期量から5 μ gを超える最終量まで、アレイエレメントを増幅する。増幅したアレイエレメントは、SEPHACRYL-400 (Amersham Pharmacia Biotech) を用いて精製する。

【0340】

精製したアレイエレメントは、ポリマーコートされたスライドガラス上に固定する。顕微

10

20

30

40

50

鏡スライドガラス (Corning) は、0.1%のSDSおよびアセトン中で超音波洗浄し、処理中および処理後に多量の蒸留水で洗浄する。スライドガラスは、4%フッ化水素酸 (VWR Scientific Products Corporation (VWR), West Chester PA) 中でエッチングし、多量の蒸留水中で洗浄し、95%エタノール中で0.05%アミノプロピルシラン (Sigma) を用いてコーティングする。コーティングしたスライドは、110のオーブンで硬化させる。

【0341】

米国特許第5,807,522号に記載されている方法を用いて、コーティングしたガラス基板にアレイエレメント群を付加する。この特許は引用を以って本明細書の一部とする。平均濃度100 ng/ μ lのアレイエレメントDNA 1 μ lを、高速ロボット装置 (robotic apparatus) 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。装置はここで、スライド毎に約5 nlのアレイエレメントサンプルを置く。

10

【0342】

マイクロアレイをUV架橋するため、STRATALINKER UV架橋剤 (Stratagene) を用いる。マイクロアレイの洗浄を、室温において、0.2% SDSで1回、蒸留水で3回行う。非特異結合部位をブロックするため、マイクロアレイのインキュベートをリン酸緩衝生理食塩水 (PBS) (Tropix, Inc., Bedford MA) 中の0.2%カゼイン中において60で30分間行った後、前に行ったように0.2% SDSおよび蒸留水で洗浄する。

【0343】

ハイブリダイゼーション

ハイブリダイゼーション反応に用いる9 μ lのサンプル混合体には、Cy3またはCy5で標識したcDNA合成産物群の各0.2 μ gを、5 \times SSC, 0.2% SDSハイブリダイゼーション緩衝液中に含む。サンプル混合体は、65まで5分間加熱し、マイクロアレイ表面上で等分して1.8 cm^2 のカバーガラスで覆う。アレイを、顕微鏡用スライドよりわずかに大きい空洞を有する防水チェンバーに移す。チェンバー内を湿度100に保持するため、140 μ lの5 \times SSCをチェンバーの1コーナーに加える。アレイを入れたチェンバーは、60で約6.5時間インキュベートする。アレイは、第1洗浄緩衝液中 (1 \times SSC, 0.1% SDS) において45で10分間洗浄し、第2洗浄緩衝液中 (0.1 \times SSC) において45で10分間ずつ3回洗浄し、乾燥させる。

20

30

【0344】

検出

レポーター標識したハイブリダイゼーション複合体を検出するには、Cy3の励起のために488 nm、Cy5の励起のために632 nmでスペクトル線を発生し得るInnova70混合ガス10Wレーザ (Coherent, Inc., Santa Clara CA) を備えた顕微鏡を用いる。励起レーザ光の焦点をアレイ上に置くため、20 \times 顕微鏡対物レンズ (Nikon, Inc., Melville NY) を用いる。アレイを含むスライドを、顕微鏡の、コンピュータ制御のX-Yステージに置き、対物レンズを通してラスタースキャンする。本実施例で用いる1.8 $\text{cm} \times 1.8 \text{ cm}$ のアレイは、解像度20 μm でスキャンする。

【0345】

異なる2回のスキャンで、混合ガスマルチラインレーザは2つの蛍光色素を連続的に励起する。発光された光は、波長に基づき分離され、2つの蛍光色素に対応する2つの光電子増倍管検出器 (PMT R1477, Hamamatsu Photonics Systems, Bridgewater NJ) に送られる。好適なフィルター群をアレイと光電子増倍管との間に設置して、シグナルをフィルターする。用いる蛍光色素の最大発光の波長は、Cy3では565 nm、Cy5では650 nmである。装置は両方の蛍光色素からのスペクトルを同時に記録し得るが、レーザ源において好適なフィルターを用いて、蛍光色素1つにつき1度スキャンし、各アレイを通常2度スキャンする。

40

【0346】

通常、各スキャンの感度を較正するため、cDNA対照種を或る既知濃度でサンプル混合体に

50

添加し、対照種が発生するシグナル強度を用いる。アレイ上の或る特定の位置には或る相補的DNA配列が含まれ、その位置におけるシグナルの強度をハイブリダイゼーション種の重量比1:100,000で相関させる。異なる源泉(例えば代表的な試験細胞および対照細胞など)からの2つのサンプルを、各々異なる蛍光色素で標識し、異なる発現をする遺伝子群を同定するために単一のアレイにハイブリダイズする場合には、較正するcDNAのサンプルを2種の蛍光色素で標識し、ハイブリダイゼーション混合液に各々等量を加えることによって較正を行う。

【0347】

光電子増倍管の出力をデジタル化するため、IBMコンパチブルPCコンピュータにインストールした12ビットRTI-835Hアナログデジタル(A/D)変換ボード(Analog Devices, Inc., Norwood MA)を用いる。デジタル化したデータは、青色(低シグナル)から赤色(高シグナル)までの擬似カラースケールへのリニア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および測定する場合には、各蛍光体の発光スペクトルを用いて、データは先ず蛍光色素間の光学的クロストーク(発光スペクトルの重なりに起因する)を補正される。

【0348】

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、GEMT 00LS遺伝子発現分析プログラム(Incyte)である。

【0349】

1.1 相補的ポリヌクレオチド

NAAPをコードする配列あるいはその任意の一部に対して相補的な配列は、天然NAAPの発現を検出、低減または阻害するために用いられる。約15~30塩基対を含むオリゴヌクレオチドの使用について記すが、これより小さなあるいは大きな配列の断片の場合でも、本質的に同じ手順を用いる。適切なオリゴヌクレオチド群を設計するため、Oligo 4.06ソフトウェア(National Biosciences)および、NAAPのコーディング配列を用いる。転写を阻害するためには、最も独特な5'配列から相補的オリゴヌクレオチドを設計し、これを用いて、プロモーターがコーディング配列に結合するのを防止する。翻訳を阻害するには、相補的なオリゴヌクレオチドを設計して、NAAPをコードする転写物にリボソームが結合するのを防ぐ。

【0350】

1.2 NAAPの発現

NAAPの発現及び精製は、細菌若しくはウイルスを基にした発現系を用いて行うことができる。細菌内でATRSを発現させるには、抗生物質耐性遺伝子と、cDNA転写レベルを高める誘導性プロモーターとを有する好適なベクターにcDNAをサブクローニングする。このようなプロモーターとしては、lacオペレーター調節エレメントと併用するT5またはT7バクテリオファージプロモーター、およびtrp-lac(tac)ハイブリッドプロモーターが含まれるが、これらに限定するものではない。組換えベクターを、BL21(DE3)などの好適な細菌宿主に形質転換する。抗生物質耐性細菌にNAAPを発現させるには、イソプロピル-Dチオガラクトピラノシド(IPTG)で誘発する。真核細胞でのNAAPの発現は、昆虫細胞株または哺乳動物細胞株に、一般にバキュロウイルスとして知られるAutographica californica核多角体病ウイルス(AcMNPV)の組換え型を感染させて行う。バキュロウイルスの非必須ポリヘドリン遺伝子をNAAPをコードするcDNAと置換するには、相同組換えを行うか、或いは、トランスファープラスミドの媒介を伴う、細菌の媒介による遺伝子転移を行う。ウイルスの感染力は維持され、強力な多角体プロモーターによって高レベルのcDNA転写が行われる。組換えバキュロウイルスは、多くの場合は夜蛾の1種Spodoptera frugiperda(Sf9)昆虫細胞への感染に用いるが、ヒト肝細胞への感染に用いることもある。後者の感染の場合には、バキュロウイルスへの更なる遺伝的修飾が必要になる(Engelhard, E.K. 他(1994) Proc

. Natl. Acad. Sci. USA 91:3224-3227; Sandig, V. 他 (1996) Hum. Gene Ther. 7:1937-1945を参照)。

【0351】

殆どの発現系では、NAAPが、例えばグルタチオンSトランスフェラーゼ(GST)と、またはFLAGや6-Hisなどのペプチドエピトープ標識と合成された融合タンパク質となるため、未精製の細胞溶解物からの組換え融合タンパク質の親和性ベースの精製を、迅速に1ステップで行い得る。GSTは日本住血吸虫からの26kDaの酵素であり、タンパク質の活性および抗原性を維持した状態で、固定化したグルタチオン上での融合タンパク質の精製を可能とする(Amersham Pharmacia Biotech)。精製後、GST要素を、操作した特定の部位においてNAAPからタンパク質的に切断することが可能である。FLAGは8アミノ酸のペプチドであり、市販されているモノクローナルおよびポリクローナル抗FLAG抗体(Eastman Kodak)を用いた免疫親和性精製を可能にする。6ヒスチジン残基が連続して伸長した6-Hisは、金属キレート樹脂上での精製を可能にする(QIAGEN)。タンパク質の発現および精製の方法は、前出のAusubel(1995)10章、16章に記載されている。これらの方法で精製したNAAPを直接用いて、以下の実施例16、17、18及び19の、適用可能なアッセイを行い得る。

【0352】

1.3 機能的アッセイ

NAAPの機能は、哺乳動物細胞培養系において生理的に高められたレベルでの、NAAPをコードする配列の発現によって算定する。cDNAを高いレベルで発現する強いプロモーターを持つ哺乳動物発現ベクターにcDNAをサブクローニングする。選択されるベクターとしては、PCMV SPORT(Life Technologies)およびPCR3.1(Invitrogen, Carlsbad CA)があり、どちらもサイトメガロウイルスプロモーターを持つ。リポソーム製剤あるいは電気穿孔法を用いて、5~10 μ gの組換えベクターをヒト細胞株、例えば内皮由来または造血由来の細胞株に、一過的に形質移入する。更に、標識タンパク質をコードする配列を含む1~2 μ gのプラスミドを同時に形質移入する。標識タンパク質の発現により、形質移入細胞と非形質移入細胞を区別する手段が与えられる。また、標識タンパク質の発現によって、組換えベクターからのcDNA発現を正確に予想できる。標識タンパク質は、例えば緑色蛍光タンパク質(GFP; Clontech)、CD64またはCD64-GFP融合タンパク質から選択できる。自動化された、レーザー光学に基づく技術であるフローサイトメトリー(FCM)を用いて、GFPまたはCD64-GFPを発現する形質移入された細胞を同定し、それらの細胞のアポトーシス状態や他の細胞特性を評価する。FCMは、細胞死に先行するかあるいは同時に発生する現象を診断する蛍光分子の取込を検出して計量する。このような現象として挙げられるのは、ヨウ化プロピジウムによるDNA染色によって計測される核DNA含量の変化、前方散乱光と90°側方散乱光によって計測される細胞サイズと顆粒性の変化、プロモデオキシウリジンの取込量の低下によって計測されるDNA合成の下方調節、特異抗体との反応性によって計測される細胞表面および細胞内におけるタンパク質の発現の変容、および、フルオレセイン抱合したアネキシンVタンパク質の細胞表面への結合によって計測される原形質膜組成の変容とがある。フローサイトメトリー法については、Ormerod, M.G.(1994) Flow Cytometry, Oxford, New York NYに記述がある。

【0353】

遺伝子発現におけるNAAPの影響は、NAAPをコードする配列と、CD64またはCD64-GFPのどちらかが形質移入された、高度に精製された細胞集団を用いて評価することができる。CD64またはCD64-GFPは、形質転換された細胞表面で発現し、ヒト免疫グロブリンG(IgG)の保存された複数の領域に結合する。形質転換された細胞と形質転換されない細胞とは、ヒトIgGがCD64に対する抗体のどちらかで被覆された磁気ビーズを用いて効率的に分離できる(DYNAL, Lake Success NY)。mRNAは、当業者に周知の方法で細胞から精製できる。NAAPと、目的とする他の遺伝子とをコードするmRNAの発現は、ノーザン分析やマイクロアレイ技術で分析できる。

【0354】

10

20

30

40

50

1.4 NAAPに特異的な抗体の作製

実質的に精製されたNAAPをポリアクリルアミドゲル電気泳動法 (PAGE; 例えばHarrington, M.G. (1990) *Methods Enzymol.* 182:488-495を参照) または他の精製技術で作製し、これを用いて標準的なプロトコルでウサギを免疫化して抗体を作り出す。

【0355】

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

10

【0356】

通常は、長さ約15残基のオリゴペプチドを、FMOCケミストリを用いるABI 431A ペプチドシンセサイザ (Applied Biosystems) を用いて合成し、N-マレイミドベンゾイル-N-ヒドロキシスクシンイミドエステル (MBS) を用いた反応によってKLH (Sigma-Aldrich, St. Louis MO) に結合させて、免疫原性を高める (例えば前出のAusubel, 1995を参照)。完全フロイントアジュバントにおいて、オリゴペプチド-KLH複合体を用いてウサギを免疫化する。得られた抗血清の抗ペプチド活性及び抗NAAP活性を試験するには、ペプチドまたはATRSを基板に結合し、1% BSAを用いてブロッキング処理し、ウサギ抗血清と反応させて洗浄し、さらに放射性ヨウ素標識されたヤギ抗ウサギIgGと反応させる。

20

【0357】

1.5 特異抗体を用いた天然NAAPの精製

天然NAAPあるいは組換えNAAPを実質的に精製するため、NAAPに特異的な抗体群を用いるイムノアフィニティークロマトグラフィーを行う。イムノアフィニティークラムは、CNBr-活性化したSEPHAROSE (Amersham Pharmacia Biotech) のような活性化クロマトグラフィー用レジンと抗NAAP抗体とを共有結合させることにより形成する。結合後に、製造者の使用説明書に従ってこのレジンをブロックし、洗浄する。

【0358】

NAAPを有する培養液をイムノアフィニティークラムに通し、NAAPを優先的に吸着できる条件で (例えば、界面活性剤の存在下で高イオン強度のバッファーで) そのカラムを洗浄する。そのカラムを、抗体とNAAPとの結合を切るような条件で (例えば、或るpH 2~3のバッファー、あるいは高濃度の、例えば尿素またはチオシアン酸イオンなどのカオトロップで) 溶出させ、NAAPを回収する。

30

【0359】

1.6 NAAPと相互作用する分子の同定

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

40

【0360】

別法では、NAAPと相互作用する分子を、Fields, S.およびO. Song (1989, *Nature* 340:245-246) に記載の酵母2-ハイブリッドシステム (yeast two-hybrid system) を、または、MATCHMAKERシステム (Clontech) など2-ハイブリッドシステムに基づく市販キットを用いて分析する。

【0361】

NAAPはまた、ハイスループット型の酵母2ハイブリッドシステムを使用するPATHCALLINGプロセス (CuraGen Corp., New Haven CT) に用いて、遺伝子の2大ライブラリによってコードされるタンパク質間の全ての相互作用を判定できる (Nandabalan, K. 他 (2000) 米国特許第6,057,101号)。

50

【0362】

1.7 NAAP活性の実証

NAAP活性の測定を、レポーター遺伝子の転写を刺激するNAAPの能力によって行う(Liu, H. Y. 他(1997) EMBO J. 16:5289-5298)。このアッセイは、十分に特徴付けられたレポーター遺伝子作成物であるLexA_{op}-LacZを用いる。このLexA_{op}-LacZは、大腸菌LacZ酵素をコードする配列に融合されたLexA DNA転写調節エレメント(LexA_{op})からなる。融合遺伝子の作成及び発現、細胞への融合遺伝子の導入、及び、LacZ酵素活性の測定方法は、当分野で周知である。NAAPをコードする配列を、LexA転写因子に由来するDNA結合ドメインとNAAPとからなる融合タンパク質、LexA-NAAPの合成を誘導するプラスミドにクローニングする。LexA-NAAP融合タンパク質をコードする得られたプラスミドを、LexA_{op}-LacZレポーター遺伝子を持つプラスミドと共に酵母細胞内に導入する。対照細胞と比較したLexA-NAAP形質転換細胞に関連するLacZ酵素活性の量が、NAAPによって刺激された転写の量に比例する。

10

【0363】

あるいはNAAPの活性は、亜鉛に結合する能力として測定される。5~10 μMのサンプル溶液(2.5 mMの酢酸アンモニウム溶液中(pH 7.4))を0.05 Mの硫酸亜鉛溶液(Aldrich, Milwaukee WI)と、100 μMのジチオスレイトール(10%メタノールを添加)の存在下で混合する。サンプルと硫酸亜鉛溶液とのインキュベーションを20分間おこなう。反応溶液をVYCACカラム(約300 Åのポアサイズ(bore size)、5 μmの粒径)(Grace Vydac, Hesperia CA)に通す。これにより、亜鉛-サンプル複合体を溶液から単離し、次に質量分析計(PE Sciex, Ontario, Canada)にかける。サンプルに結合した亜鉛の定量には、機能的原子量である63.5 Daを用いる。この原子量を観測したのはWhittall, R. M. 他((2000) Biochemistry 39:8406-8417)である。

20

【0364】

或いは、NAAPの核酸結合活性を判定する或る方法は、ポリアクリルアミドゲル泳動移動度シフトアッセイを伴う。このアッセイの調製では、NAAPは、NAAP cDNAを有する真核生物発現ベクターを用い、COS7、HeLa、若しくはCHOなど、哺乳類細胞株を形質転換することで発現される。形質転換の後、この細胞株がNAAPを発現し蓄積するのに適した条件下で、これらの細胞を48~72時間インキュベーションする。可溶化した蛋白質を有する抽出物の調製は、NAAPを発現する細胞から行いする。この方法は当分野で周知である。NAAPを有する抽出物の各部分を、³²P標識したRNAまたはDNAに添加する。放射性核酸の合成は、*in vitro*で、当分野で周知の技術でなしうる。混合物のインキュベーションを25℃、RNase阻害剤とDNase阻害剤との存在下、緩衝した条件下で5-10分間おこなう。インキュベーション後、サンプルの分析を、ポリアクリルアミドゲル電気泳動とその後のオートラジオグラフィとで行う。バンドがオートラジオグラム上に存在すれば、NAAPと放射性転写物との複合体の形成を示す。同様の移動度のバンドは、非形質転換細胞から調製した対照抽出物を用いて準備したサンプル中には見られないであろう。

30

【0365】

或いはNAAPのメチラーゼ活性を判定する或る方法では、供与体基質と受容体基質との間での放射標識メチル基の転移を測定する。反応混合液(50 μlの最終容量)には、15 mMのHEPES(pH 7.9)、1.5 mMのMgCl₂、10 mMジチオスレイトール、3%ポリビニルアルコール、1.5 μCi [メチル-³H]AdoMet(0.375 μMのAdoMet)(DuPont-NEN)、0.6 μgのNAAP、および受容体基質(例えば0.4 μgの³⁵S]RNAまたは6-メルカプトプリン(6-MP)、最大1 mMの最終濃度)を含む。反応混液を30℃で30分間インキュベーションし、次に65℃で5分間インキュベーションする。

40

【0366】

[メチル-³H]RNAの分析法は次のとおりである。(1) 50 μlの2×ローディングバッファ(20 mMのTris-HCl(pH 7.6)、1 MのLiCl、1 mMのEDTA、1%ドデシル硫酸ナトリウム(SDS))および50 μlオリゴd(T)-セルロース(1×ローディングバッファ中に10 mg/ml)を反応混液に添加し、インキュベーションを室温で振盪しながら30分間おこなう。(2) 反応混液を96穴ろ過

50

プレート(真空装置に接続)に移す。(3)各サンプルを逐次洗浄する。これには3つの2.4 ml アリコットの1×オリゴd(T)ローディングバッファ(0.5% SDSまたは0.1% SDSを有するものと、SDSを有しないもの)を用いる。(4) RNAの溶出を、300 µlの水で96ウェル収集プレートへ行き、シンチレーションバイアル(液体シンチラントを含有)に移し、放射能を判定する。

【0367】

[メチル-³H]6-MPの分析法は次のとおりである。(1) 500 µlの0.5 M ホウ酸バッファ(pH 10.0)を、次に20% (v/v)でトルエンに希釈したイソアミルアルコールの2.5 mlを、反応混液に添加する。(2) サンプルの混合を、10秒間、ボルテックスで良く攪拌しておこなう。(3) 遠心分離を700gで10分間おこなった後、1.5 mlの有機相をシンチレーションバイアル(0.5 ml 無水エタノールと液体シンチラントとを含有)に移し、放射能を判定する。(4) 結果の補正を、6-MPの、有機相への抽出について行う(約41%)。

10

【0368】

或いは、NAAPのタイプIIトポイソメラーゼ活性のアッセイを、或るスーパーコイルDNA基質の弛緩(relaxation)に基づき行う。NAAPのインキュベートをその基質と共に、Mg²⁺とATPとを欠くバッファ中で行い、反応を終了させ、生成物をアガロースゲルにロードする。変容したトポアイソマーの、スーパーコイル基質からの弁別は、電気泳動でできる。このアッセイはタイプIIトポイソメラーゼ活性に特異的である。その理由は、Mg²⁺とATPとが、タイプIIトポイソメラーゼに必要な補因子だからである。

【0369】

NAAPのタイプIIトポイソメラーゼ活性のアッセイを、或るキネトプラストDNA(KDNA)基質の脱連環(decatenation)に基づいて行い得る。NAAPのインキュベートをKDNAと共に行い、反応を終了させ、生成物をアガロースゲルにロードする。単量体環状KDNAの、連環したKDNAとの弁別は、電気泳動でできる。タイプIIトポイソメラーゼ活性とタイプIIトポイソメラーゼ活性とを測定する市販キットはTopogen (Columbus OH)で得られる。

20

【0370】

NAAPのATP依存性RNAヘリカーゼ巻き戻し活性は、ZhangおよびGrosse (1994; Biochemistry 33:3906-3912)に記載された方法で測定できる。RNA巻き戻し用の基質には³²P-標識RNAを含み、このRNAは二本のRNAストランド(194および130ヌクレオチドの長さ)からなり、17塩基対のduplex領域を持つ。RNA基質のインキュベートを、ATP、Mg²⁺、および可変量NAAPと共に、Tris-HClバッファ中(pH 7.5)、37°Cで30分間おこなう。次に一本鎖RNA産物を二本鎖RNA基質から分離するための電気泳動を10% SDSポリアクリルアミドゲルでおこない、オートラジオグラフィで定量する。回収した一本鎖RNAの量は、調製物中のNAAP量に比例する。

30

【0371】

あるいはNAAP機能のアッセイは、哺乳動物細胞培養系において生理的に高められたレベルでの、NAAPをコードする配列の発現によって算定する。cDNAを高いレベルで発現する強いプロモーターを持つ哺乳動物発現ベクターにcDNAをサブクローニングする。選択されるベクターとしては、pCMV SPORT (Life Technologies) およびpCR3.1 (Invitrogen Corporation, Carlsbad CA)があり、どちらもサイトメガロウイルスプロモーターを持つ。リポソーム製剤あるいは電気穿孔法を用いて、5~10 µgの組換えベクターをヒト細胞株、好適には内皮由来または造血由来の細胞株に、一過的に形質移入する。更に、標識タンパク質をコードする配列を含む1~2 µgのプラスミドを同時に形質移入する。

40

【0372】

標識タンパク質の発現により、形質移入細胞と非形質移入細胞を区別する手段が与えられる。また、標識タンパク質の発現によって、組換えベクターからのcDNA発現を正確に予想できる。標識タンパク質は、例えば緑色蛍光タンパク質(GFP; CLONTECH)、CD64またはCD64-GFP融合タンパク質から選択できる。自動化された、レーザー光学に基づく技術であるフローサイトメトリー(FCM)を用いて、GFPまたはCD64-GFPを発現する形質移入された細胞を同定し、それらの細胞のアポトーシス状態や他の細胞特性を評価する。

50

【0373】

FCMは、細胞死に先行するかあるいは同時に発生する現象を診断する蛍光分子の取込を検出して計量する。このような現象として挙げられるのは、ヨウ化プロピジウムによるDNA染色によって計測される核DNA含量の変化、前方散乱光と90°側方散乱光によって計測される細胞サイズと顆粒性の変化、プロモデオキシウリジンの取込量の低下によって計測されるDNA合成の下方調節、特異抗体との反応性によって計測される細胞表面および細胞内におけるタンパク質の発現の変容、および、フルオレセイン抱合したアネキシンVタンパク質の細胞表面への結合によって計測される原形質膜組成の変容とがある。フローサイトメトリー法については、Ormerod, M. G. (1994) *Flow Cytometry*, Oxford, New York NY に記述がある。

10

【0374】

遺伝子発現に与えるNAAPの影響は、NAAPをコードする配列と、CD64またはCD64-GFPのどちらかだけが形質移入された、高度に精製された細胞集団を用いて評価することができる。CD64またはCD64-GFPは、形質転換された細胞表面で発現し、ヒト免疫グロブリンG(IgG)の保存された複数の領域に結合する。形質転換された細胞と形質転換されない細胞とは、ヒトIgGがCD64に対する抗体のどちらかで被覆された磁気ビーズを用いて効率的に分離できる(Dynal Biotech, Lake Success NY)。mRNAは、当業者に周知の方法で細胞から精製できる。NAAPと、目的とする他の遺伝子とをコードするmRNAの発現は、ノーザン分析やマイクロアレイ技術で分析できる。

【0375】

NAAPのプロソドウリジン合成酵素活性のアッセイには、トリチウム(³H)放出アッセイを、Nurse 他((1995) RNA 1:102112)を改作して行い、RNA中の³H-放射標識Uが異性化されてプロソドウリジン(y)になるときに、ウリジル酸(U)のピリミジン成分のC₅位置からの³Hの放出を測定する。典型的な500 μlアッセイ混液の内容は、50 mMのHEPESバッファ(pH 7.5)、100 mM酢酸アンモニウム、5 mMジチオスレイトール、1 mMのEDTA、30単位のRNase阻害剤、および0.1~4.2 μMの [³H]tRNA (約1 μCi/nmolのtRNA)である。反応を開始するには、5 μlより少量のNAAPの濃縮溶液(またはNAAP含有サンプル)を添加し、インキュベートを5分間37 °Cで行う。NAAPを加えた後、反応混液の一部を様々な経過時間(最長30分まで)に採取し、NoritSA3 (12% w/v)を含有する1 mlの0.1 MのHClに希釈して反応を停止する。停止した反応混液の遠心分離を5分間、最大速度のマイクロ遠心機で行い、上澄みのろ過を、グラスウールの詰め物を通して行う。ペレットを、1 mlの0.1 M HCl中で2回、再懸濁して洗浄した後、遠心分離する。洗浄後の上澄みの各々を、別々にグラスウールの詰め物に通し、元のろ過物と混合する。混合したろ過物の一部を、シンチレーション液(最大10 ml)と混合し、シンチレーションカウンタでカウントする。RNAから放出された³Hの量、また可溶性ろ液に存在する³Hの量が、サンプルにおけるプロソドウリジン合成酵素活性の量に比例する(Ramamurthy, V. (1999) J. Biol. Chem. 274:2222522230)。

20

30

【0376】

あるいはNAAPのプロソドウリジン合成酵素活性のアッセイは、30 °C ~ 37 °Cで、次の混液で行う。すなわち100 mMのTrisHCl (pH 8.0)、100 mM 酢酸アンモニウム、5 mMのMgCl₂、2 mMジチオスレイトール、0.1 mMのEDTA、および1~2 fmolの [³²P]放射標識ランオフ(run off)転写物(産生は*in vitro*で、適切なRNAポリメラーゼすなわちT7またはSP6で行う)を基質として含む混液である。NAAPを加えて反応を開始させるか、反応液に加えずに対照サンプルとする。インキュベート後、RNAをフェノール-クロロホルムで抽出し、エタノール中に沈殿させる。また、完全に加水分解して3-ヌクレオチドリン酸にするためRNase T₂を用いる。水解物の分析には2次元薄層クロマトグラフィを用い、³²P放射ラベルの量(yMPおよびUMPスポット中に存在する量)の評価を、薄層クロマトグラフィプレートフィルムまたはPHOSPHORIMAGERスクリーン(Molecular Dynamics)に曝した後に行う。基質RNA中のウリジル酸残基の相対数を考慮してyMPとUMPとの相対量を判定し、この量を用いてtRNA分子の量に対するyの相対量を計算する(モルy / tRNAモル、またはモルy / tRNAモル/分で表す)。この量が、NAAPサンプル中のプロソドウリジン合成酵素活性の量に一致する(Leco

40

50

inte, F. 他(1998) J. Biol. Chem. 273:13161323)。

【0377】

NAAPの N^2, N^2 ジメチルグアノシン転移酵素(m^2_2G)メチルトランスフェラーゼ)活性の測定は、160 μ l 反応混液に以下を含有して行う。すなわち100 mMのTris-HCl (pH 7.5)、0.1 mMのEDTA、10 mMの $MgCl_2$ 、20 mMの NH_4Cl 、1mMジチオスレイトール、6.2 μ MのS-アデノシル-L-[メチル- 3H]メチオニン(30~70 Ci/mM)、8 μ gの m^2_2G -欠損tRNAまたは野生型tRNA(酵母由来)および約100 μ gの精製NAAPまたはNAAP含有サンプルを含む。反応液を30 で90分間インキュベートし、氷上で冷却する。各反応液の一部を、100 μ g BSAを含む1 mlの水に希釈する。1 mlの2 M HClを各サンプルに加え、酸性不溶性産物を氷上で20分間沈殿させた後、収集のためのろ過をガラス繊維フィルターで行う。収集した物質の洗浄を数回、塩酸で行い、液体シンチレーションカウンターを用いて定量する。 3H が、 m^2_2G を欠損した酸性不溶性tRNAに取り込まれる量が、NAAPサンプル中の N^2, N^2 ジメチルグアノシン転移酵素活性の量に比例する。基質tRNAを含まない反応、または既に修飾された野生型tRNAを含む反応は対照反応として役立ち、この反応では酸性不溶性 3H -標識産物は得られないはずである。

10

【0378】

NAAPのポリアデニル化活性の測定には、in vitro のポリアデニル化反応を用いる。反応混合液の作製は氷上で行い、以下を含む。10 μ l の 5 mM ジチオスレイトール、0.025% (v/v) Nonidet P40界面活性剤、50 mMクレアチンリン酸、6.5% (w/v)ポリビニルアルコール、0.5ユニット/ μ lのRNAGUARD (Amersham Pharmacia Biotech)、0.025 μ g/ μ lクレアチン・キナーゼ、1.25 mM コルジセピン5'三リン酸、および3.75 mMの $MgCl_2$ で、総容量25 μ lとする。60 fmolのCstF、50 fmolのCPSF、240 fmolのPAP、4 μ lの未精製CF II、又は部分的に精製したCF IIと、可変量CF Iを次に反応混液に加える。容量を調整して23.5 μ lにするため、以下を有するバッファを加える。50 mMのTrisHCl (pH 7.9)、10% (v/v)グリセロール、および0.1 mMのNaEDTAである。最終的な硫酸アンモニウム濃度は20 mM未満とする。反応を氷上で開始するため、15 fmol の ^{32}P 標識mRNA前駆体テンプレートと、2.5 μ gの無標識tRNAを含む1.5 μ lの水を加える。反応液を次にインキュベートする。これは30 で75~90分間行い、停止するには75 μ l (約2倍量)のプロテイナーゼKミックス(0.2 MのTris-HCl (pH 7.9)、300 mMのNaCl、25 mMのNaEDTA、2% (w/v) SDS)、1 μ lの10 mg/mlプロテイナーゼK、0.25 μ lの20 mg/mlグリコーゲンおよび23.75 μ lの水を加える。インキュベート後、RNAをエタノールで沈殿させ、分析を6% (w/v)ポリアクリルアミド、8.3 M尿素シーケンシングゲル上で行う。乾燥したゲルの感光を、オートラジオグラフィで、またはホスホイメジャー(phosphoimager)で行う。切断活性を判定するには、切断産物の量を、mRNA前駆体テンプレートの量と比較する。いずれかのポリペプチド成分を反応から除くことと、NAAPを代替することとは、mRNA前駆体ポリアデニル化におけるNAAPの特定の生物機能を同定する上で有用である(Ruegsegger, U. 他(1996) J. Biol. Chem. 271:6107-6113、およびその参考文献)。

20

30

【0379】

tRNA合成酵素活性は、 $[^{14}C]$ -標識されたアミノ酸の存在下で基質tRNAのアミノアシル化として測定される。NAAPは緩衝液中で、 $[^{14}C]$ -標識されたアミノ酸と好適な同族tRNA(例えば、 $[^{14}C]$ アラニンおよび tRNA^{ala})と共にインキュベートされる。 ^{14}C -標識の産物は遊離 $[^{14}C]$ アミノ酸からクロマトグラフィーによって分離され、取り込まれた $[^{14}C]$ はシンチレーションカウンターで測定される。 ^{14}C -標識産物の量が、このアッセイのNAAPの活性に比例する。

40

【0380】

あるいはNAAP活性を測定するため、NAAP含有サンプルのインキュベートを、以下を含む溶液で行う。1 mMのATP、5 mMのHepes-KOH (pH 7.0)、2.5 mMのKCl、1.5 mM塩化マグネシウム、および0.5 mMのDTTであり、また誤ってアシル化した $[^{14}C]$ -GluRnAGIn(例えば1 μ M)および同様の濃度の無標識Lグルタミンをも含む。反応を停止させるため3 M酢酸ナトリウム(pH 5.0)を加えた後、混液の抽出を同量の水 飽和フェノールで行い、水相と有機相

50

とを分離する遠心分離を15,000 × g、室温で1分間おこなう。水相の沈殿を3倍量のエタノール、-70 で15分間おこなう。沈殿したアミノアシルtRNAを回収する遠心分離を15,000 × g、4 で15分間おこなう。ペレットの再懸濁を25 mMのKOHで行い、脱アシル化を65 で10分間おこない、中和を0.1 MのHClで行い(最終pH 6~7)、真空乾燥する。乾燥ペレットの再懸濁を水中で行い、セルロースTLCプレート上にスポットする。プレートの感光を、イソプロパノール/蟻酸/水またはアンモニア/水/クロロホルム/メタノール中で行う。イメージを濃度計分析し、GluとGlnとの相対量の計算を、スポットのRf値と相対強度とに基づき行う。NAAP活性の計算を、GluがGlu-tRNA^{Gln}としてアシル化され転換した結果のGlnの量に基づき行う(Curnow, A.W. 他(1997) Proc. Natl. Acad. Sci. U. S. A. 94:11819-26から適用)。

10

【0381】

18 NAAPアゴニスト及びアンタゴニストの同定

NAAPの活性化若しくは抑制のアゴニスト若しくはアンタゴニストを、実施例17で記述したアッセイを用いてテストし得る。アゴニストはNAAP活性の増加を引き起こし、アンタゴニストはNAAP活性の減少を引き起こす。

【0382】

19 NAAP分泌アッセイ

或る高処理アッセイを用いて、真核細胞内に分泌されるポリペプチドを同定しうる。このようなアッセイの1例では、ポリペプチド発現ライブラリを作製するため、5'に偏向したcDNAを、或るリーダーのない ラクタマーゼ遺伝子の5'末端に融合する。ラクタマーゼは便利な遺伝子レポーターである。理由は、この酵素が高い信号/ノイズ比と、低い内因性バックグラウンド活性とを示し、他の蛋白質に融合しても活性を保持するからである。或る2重プロモーター系によって、ラクタマーゼ融合ポリペプチドの発現を、細菌又は真核細胞でなしうる。これにはlacまたはCMVプロモーターを各々用いる。

20

【0383】

ライブラリは先ず細菌(大腸菌など)に形質転換され、真核生物系で分泌されうる融合ポリペプチドをコードするライブラリメンバーが同定される。哺乳類シグナル配列は、ラクタマーゼ融合ポリペプチドが細菌のペリプラズムへ移動するよう指示する。ペリプラズムでこれはカルベニシリンへの抗生物質耐性を授ける。カルベニシリン選択された細菌の単離を固体培地で行い、個々のクローンを液体培地で成長させ、生じた培養物を用いてライブラリメンバープラスミドDNAを単離する。

30

【0384】

哺乳類細胞(293細胞など)の播種を96ウェル組織培養プレート上に、約40,000細胞/ウェルの密度で、100 μlフェノールレッドを含まないDMEに10%ウシ胎児血清(FBS) (Life Technologies, Rockville, MD)を加えて行う。次の日に、精製プラスミドDNA(カルベニシリン耐性細菌から単離)を、15 μlのOPTI-MEM 1培養液(Life Technologies)で、形質移入すべき細胞の各ウェルに25 μlの容量まで希釈する。別のプレートで、1 μlのLF2000 Reagent (Life Technologies)の希釈を25 μl/ウェルOPTI-MEM 1中に行う。25 μl希釈LF2000 Reagentを次に25 μl希釈DNAと混合し、短時間混ぜ合わせ、インキュベートを20分間、室温で行う。生じたDNA-LF2000試薬複合体を次に、直接、293細胞の各ウェルに加える。細胞の形質移入はまた、適切な対照プラスミド(野生型 ラクタマーゼ、リーダーのない ラクタマーゼ、又は例えばCD4融合したリーダーのない ラクタマーゼのいずれかを発現するプラスミド)で行う。形質移入の24時間後、約90 μlの細胞培地のアッセイを37°Cで、100 μMのNitrocefin (ニトロセフィン、Calbiochem, San Diego CA)及び0.5 mMのオレイン酸(Sigma, St. Louis, MO)を用い、10 mMリン酸バッファー(pH 7.0)中でおこなう。ニトロセフィンはラクタマーゼの基質であり、加水分解すると黄色から赤へ顕著に変色する。ラクタマーゼ活性のモニターを、20分間、マイクロタイタープレートリーダーにおいて486nmで行う。486nmでの色吸収の増加が、形質移入した細胞培地でのラクタマーゼ融合ポリペプチドの分泌に一致する。これは、融合ポリペプチドにおける真核生物シグナル配列の存在の結果である。対応するライブラリメンバープラスミドDNAのポリヌク

40

50

レオチド配列分析を次に用いて、シグナル配列をコードするcDNAを同定する(記載は米国特許出願09/803,317号、ファイル日は2001年3月9日)。

【0385】

当業者には、本発明の要旨および精神から逸脱しない範囲での、記載した本発明の方法およびシステムの、種々の修正および変更の手段は自明であろう。本発明について説明するにあたり幾つかの実施例に関連して説明を行ったが、本発明の請求の範囲が、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。分子生物学または関連分野の専門家には明らかな、本明細書に記載する本発明の実施方法の様々な修正は、明確に特許請求の範囲内にあるものとする。

【0386】

10

(表の簡単な説明)

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

【0387】

表2は、本発明のポリペプチド群のGenBank識別番号と、最も近いGenBank相同体の注釈(annotation)と、PROTEOMEデータベース識別番号と、PROTEOMEデータベース相同体群の注釈とを示す。また、各ポリペプチドとその相同体(1つ以上)が一致する確率スコアも併せて示す。

【0388】

表3は、予測されるモチーフおよびドメインなど、本発明のポリヌクレオチド配列の構造的特徴を、ポリペプチドの分析に用いる方法、アルゴリズムおよび検索可能なデータベースと共に示す。

20

【0389】

表4は、本発明のポリヌクレオチド配列を構築するために用いたcDNA断片やゲノムDNA断片を、ポリヌクレオチド配列の選択した断片と共に示す。

【0390】

表5は、本発明のポリヌクレオチドの代表的なcDNAライブラリを示す。

【0391】

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

30

【0392】

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

【表1】

表 1

Incyte プロジェクト ID	ボリベブチド SEQ ID NO:	Incyte ボリベブチド N ID	ボリベブチド SEQ ID NO:	Incyte ボリベブチド N ID
2530775	1	2530775CD1	17	2530775CB1
926296	2	926296CD1	18	926296CB1
1322761	3	1322761CD1	19	1322761CB1
7472664	4	7472664CD1	20	7472664CB1
7473124	5	7473124CD1	21	7473124CB1
7473171	6	7473171CD1	22	7473171CB1
7477026	7	7477026CD1	23	7477026CB1
6428773	8	6428773CD1	24	6428773CB1
2749402	9	2749402CD1	25	2749402CB1
118539	10	118539CD1	26	118539CB1
4005918	11	4005918CD1	27	4005918CB1
5435937	12	5435937CD1	28	5435937CB1
7503560	13	7503560CD1	29	7503560CB1
7472378	14	7472378CD1	30	7472378CB1
1306049	15	1306049CD1	31	1306049CB1
3187174	16	3187174CD1	32	3187174CB1

【表 2】

表 2-1

ポリペプチド SEQ ID NO:	Incyte ポリ ペプチド ID	GenBank ID NO:または PROTEOME ID NO:	塩基スコア	注釈
1	2530775CD1	5780548	5.80E-29	[出芽酵母 (<i>Saccharomyces cerevisiae</i>)] Rrlp. West, R.W. 他(2000) RLR1 (THO2), required for expressing lacZ fusions in yeast, is conserved from yeast to humans and is a suppressor of SIN4. Gene 243, 195-205.
2	926296CD1	91573555	7.90E-105	[インフルエンザ菌 (<i>Haemophilus influenzae</i>) RCI] 転写アクセサリ蛋白(tex)
3	1322761CD1	91089816	3.20E-126	Emx2 ホメオ蛋白 (homeoprotein) [ゼブラフィッシュ (<i>Danio rerio</i>)]. Morita, T. 他(1995) Differential expression of two zebrafish emx homeoprotein mRNAs in the developing brain. Neurosci. Lett. 198:131-134.
4	7472664CD1	96850321	4.80E-52	YTA7 ATPase 遺伝子 (出芽酵母 gblX8(072 由来) に類似、プロモドメイン PF00439, AAA PF00004, および Sigma-54 PF00158 転写因子ドメインを含む) [シロイヌナズナ]
5	7473124CD1	93417319	2.10E-83	RET フインガン-蛋白質様 3 [ヒト] Seroussi, E. 他(1999) Duplications on human chromosome 22 reveal a novel Ret Finger Protein-like gene family with sense and endogenous antisense transcripts. Genome Res. 9:803-814.
6	7473171CD1	9563127	1.30E-23	酸性フインガン-蛋白質 [ヒト] Chau, T.W. 他(1995) Cloning of a new 'finger' protein gene (ZNF173) within the class I region of the human MHC. Genomics 29:229-239.
7	7477026CD1	92565057	2.70E-46	CAGH44 [ヒト] Margolis, R.L. 他(1997) cDNAs with long CAG trinucleotide repeats from human brain. Hum. Genet. 100:114-122.
8	6428773CD1	94519621	9.80E-251	OASIS 蛋白質 [ハツカネズミ] Homma, Y. 他(1999) Identification of a novel gene, OASIS, which encodes for a putative CREB/ATF family transcription factor in the long-term cultured astrocytes and glial tissue. Brain Res. Mol. Brain Res. 69:93-103.

【表 3】

10

20

30

表 2-2

ポリペプチド SEQ ID NO:	Incyte ポリ ペプチド ID	GenBank ID NO:または PROTEOME ID NO:	確率スコア	注釈
9	2749402CD1	9625042	1.20E-135	基本的ドメイン/ロイシンジッパー転写因子[ハツカネズミ] Cordes, S.P. 他(1994) The mouse segmentation gene kr encodes a novel basic domain-leucine zipper transcription factor. Cell 79:1025-1034.
10	118539CD1	92674195	9.70E-38	ポリメラゼ I 転写物放出因子、PTRF[ハツカネズミ] Jansa, P. 他(1998) Cloning and functional characterization of PTRF, a novel protein which induces dissociation of paused ternary transcription complexes. EMBO J. 17:2855-2864.
11	4005918CD1	9186774	2.60E-44	[ヒト] リンパ関連 Zn フィンガー蛋白質 Bellefroid, E.I. 他(1993) EMBO J 12:1363-1374.
12	5435937CD1	94519270	2.00E-213	[ヒト] 造血性アポトーシス関連 Kruppel タイプ Zn フィンガー蛋白質 Katoh, O. 他(1998) Biochem. Biophys. Res. Commun. 249:595-600.
13	7503560CD1	9625042	2.70E-79	[マウス] 基本的ドメイン/ロイシンジッパー転写因子 Cordes, S.P. および Baish, G.S. (1994) The mouse segmentation gene kr encodes a novel basic domain-leucine zipper transcription factor. Cell 79:1025-1034.
13		421895 Maf b	2.40E-80	[マウス][転写因子] Maf ファミリー転写因子のメンバー、発生と細胞分化と を制御すると思われる。
14	7472378CD1	92791289	0	[キイロショウジョウバエ] プロテアーゼ Leblanc, P. 他(1997) EMBO J. 16:7521-7531.
15	1306049CD1	99651711	1.60E-61	[ハツカネズミ] 亜硫酸還元 RNA 関連タンパク質

10

20

30

表 3-1

Eq ID No.	Incyte ポリペプチド ID/	アミノ酸 残基数	潜在的リン酸化部位	潜在的グリコシ 化部位	シグネチャ配列、 ドメインおよびモチーフ	分析方法及びデータベース
1	2530775CD1	1096	S67 S103 S229 S292 S330 S331 S339 S471 S492 S533 S558 S678 S720 S725 S726 S765 S827 S856 S893 S930 S938 S989 S1019 S1039 S1047 S1050 S1064 S1079 S1092 T60 T207 T225 T312 T354 T391 T400 T472 T535 T580 T730 T751 T791 T824 T926 T1000 Y529 Y563	N32 N290 N474 N513 N755 N1007 N1034	横貫通ドメイン: T610-R649, L638-T658 N 末端はサイトソル内	TMAP
					コスミド C16A3 RLRI タンパク質 転写調節 PD044185;Y384-R679 PD044186;N4-E121	BLAST_PRODUM
					タンパク質 トポイソメラーゼ I DNA イソメラーゼ リビーン DNA 結合 中間体 フィラメント ヘプタッド (HEPTAD) PD000422;K770-S1039	BLAST_PRODUM
					do 神経フィラメント; トリプレット; DM04498 P12036 434-1019; S712-K1074 DM04498 P19246 716-1085; K692-G1059	BLAST_DOMO

【表 5】

10

20

30

表 3-2

Eq ID NO.	Incyte ポリペプチド ID/	アミノ酸 残基数	潜在的リン酸化部位	潜在的グリコシル化部位	シグネチャ配列、ドメインおよびモチーフ	分析方法及びデータベース
2	926296CD1	995	S23 S29 S31 S39 S50 S143 S203 S219 S260 S406 S407 S488 S492 S511 S697 S794 S810 S861 S890 S964 T96 T110 T113 T126 T130 T139 T164 T204 T222 T267 T304 T361 T449 T846 T899 T923 T959	N88 N150 N202 N453 N490 N589 N597 N928 N950	S1 RNA 結合ドメイン: D915-I992	HMMER_PFAW
					膜貫通ドメイン: V926-N950 N 末端はサイトソル内にはない リボソーム蛋白質 S1 シグネチャ PR00681: F932-V951, V975-R993 タンパク質 YDCI RNA 結合 TEX 領域 GREBFOA 遺伝子間 HI0568 E PD011817:I233-S697 タンパク質 RNA 結合 TEX 領域 YDCI 遺伝子間 GREBFOA HI0568 E コンピタンス PD006092:Q683-I766 核内局在化シグナルドメイン DM07019 P34703 1-1454:R602-E782	TMAP BLIMPS_PRINTS BLAST_PRODUM
3	132276 CD1	252	S25, S65, T139, S192, T228, S242, T249	N137	シグナル切断: M1-S65	BLAST_DOMO SPSCAN

【表 6】

10

20

30

表 3-3

EQ ID NO:	Incyte ポリペプチド ID/	アミノ酸 残基数	潜在的リン酸化部位 潜在的グリコシル化部位	シグネチャ配列, ドメインおよびモチーフ	分析方法及びデータベース
3				ホメオボックスドメイン: K155-R211 ホメオボックスポリペプチド; アンテナペディア (antennapedia)タイプ蛋白質: BL00032:L108-L130, R158-T196, Q197-L214 ホメオボックスドメインシグネチャおよびプロファイル (homeobox.prf): R168-N232 ホメオボックスシグネチャ: PR00024:N176-L187, L191-W201, W201-K210 ホメオ蛋白質 ホメオボックス DNA 結合 核内タンパク質 EMX1 EMX2:PD15275:M1-P154; PD153752:K213-D252 ホメオボックス DM00009 F18488 385-450:R152-K213, DM00009 A46305 134-196:K153-Q212 ホメオボックスドメインシグネチャ (Homeobox_1): L187-K210	HMMER-PFAM BLIMPS-BLOCKS PROFILES BLIMPS-PRINT BLAST-PRDOM BLAST-DMO MOTIFS
4	7472664CDI	602	S2, T48, S49, T84, S108, S159, T183, T197, T209, S248, T257, S319, S327, S351, S382, T388, S396, S412, T420, S433, S434, S447, S467, T471, S514	シグナル切断:M1-G18	SPSCAN

10

20

30

表 3-4

EPQ ID NO:	Incyte ポリペプチド ID/	アミノ酸 残基数	潜在的リン酸化部位	潜在的グリコシル化部位	シグネチャ配列、ドメインおよびモチーフ	分析方法及びデータベース
					プロモドメイン (Bromodomain) :I201-D278	HMMER-PFAM
					プロモドメイン蛋白質:BL00633:P231-Y255, D264-N276	BLIMPS-BLOCKS
					プロモドメインシグネチャおよびプロファイル (bromodomain.prf):P224-T296	PROFLESCAN
4					プロモドメインシグネチャ PRO0503:D264-D283, K216-E229, V230-V246, V246-D264	BLIMPS-PRINTS
					蛋白質 TAT-結合相合体 ATP-結合 プロモドメイン C31G5.19 染色体 I.F11A10.1_PD151474:MI-P156	BLAST-PRODOM
5	7473124C01	280	S46, S65, T152, S157, S209, S242, S254, T268, S269	N240	プロモドメイン DM00265 P55201 618-733:I203-E303 (p = 2.2e-10) Znフィンガー-, C3HC4 タイプ(RING フィンガー):C20-C61	BLAST-DMO HMMER-PFAM
					RET フィンガー_蛋白質様 RPPLAL_PD152257:M10-P72	BLAST-PRODOM
					RFP トランスフォーミング蛋白質 DM01944 P14373 368-492:S157-C279	BLAST-DMO
6	7473171C01	221	T12, S107, T197, S202		シグナル切断:MI-C29	SPSCAN
					シグナルペプチド:MI-C29	HMMER
					B-ボックス Zn フィンガー:A125-L166	HMMER-PFAM
					Zn フィンガー-, C3HC4 タイプ(RING フィンガー):C29-I53, F74-C81	HMMER-PFAM

10

20

30

表 3-5

BQ ID NO:	Incyte ポリペプチド ID/	アミノ酸 残基数	潜在的リン酸化部位	潜在的グリコシル化部位	シグネチャ配列、ドメインおよびモチーフ	分析方法及びデータベース
					Zn フィンガー, C3HC4 タイプ(RING フィンガー), シグネチャ (zinc_finger_c3hc4.prf):A27-Q89	PROFILESCAN
					RFP トランスプオートミニング蛋白質 DM02346 P14373 61-366:L90-S208	BLAST-DMO
7	7477026CD1	668	S5, T9, T42, T48, S52, S57, S244, T286, S287, S291, S292, S293, T316, S360, S387, S391, S430, S431, T477, T496, T498, T525, S592	N90, N506, N559, N606	Zn フィンガー, C3HC4 タイプ(RING フィンガー), シグネチャ C44-I53; フォークヘッドドメイン:R455-R534	MOTIFS HMMER-PPAM
					フォークヘッドドメイン蛋白質:BL00657;R455-T496, KS00-S542	BLIMPS-BLOCKS
					フォークヘッドドメインシグネチャおよびプロフィール (fork_head_1.prf):S391-E480	PROFILESCAN
					フォークヘッドドメインシグネチャ:PR00053;R455-I468, I476-R493, W499-V516	BLIMPS-PRINTS
					蛋白質転写因子 核内 DNA 結合 調節 フォークヘッド フォークヘッドドメイン:PD000425;R455-D527	BLAST-PRODOM

【表 9】

表 3-6

BQ ID NO.	Incyte ポリペプチド ID/	アミノ酸 残基数	潜在的リン酸化部位 潜在的グリコシル化部位	シグネチャ配列、 ドメインおよびモチーフ	分析方法及びデータベース
8	642873CD1	519	S18, S57, T108, T197, S214, S221, T263, T284, S303, S307, T356, T403, S474, T478, T494, T515	フォークヘッド DNA-結合ドメイン DM00381 Q91167 8-329:K378-P611 Znフィンガ、C2H2 タイプ、ドメイン: C309-H332 bZIP 転写因子: E288-K352	BLAST-DMO MOTIFS HMWER-PPAM
8				CAMP 応答エレメント結合 (CREB) 蛋白質シグネチャ: PR00041:Q305-F325 bZIP 転写因子: BL00036:K299-K311 蛋白質 DNA 結合 核内 転写因子 調節 活性化因子 結合 リン酸化 前駆体: PD000088: E287-Q354 (p = 8.0e-10) BZIP 転写因子 塩基性ドメイン DM00107 P29747 360-509:H212-A369 ロイシンジッパーバータール: L332-L353 bZIP 転写因子 塩基性ドメイン シグネチャ: R295-K310	BLIMFS-PRINTS BLIMFS-BLOCKS BLAST-PRODOM BLAST-DMO MOTIFS MOTIFS

【表 10】

10

20

30

表 3-7

Eq ID NO.	Incyte ポリペプチド ID/	アミノ酸 残基数	潜在的リン酸化部位	潜在的グリコシ ル化部位	シグネチャ配列、 ドメインおよびモチーフ	分析方法及びデ ータベース
9	2749402CD1	1256	S15, S55, T85, T119, S122, T150, S163, S187, T214, T236, S252, T284, S312, S321, S332, T369, S399, S459, S567, Y580, T620, S685, T686, T724, S782, S794, T813, T822, T853, S860, S909, T923, T930, S934, T949, S950, S1026, S1043, T1048, S1077, S1180, T1098, Y1111, T1163	N83, N114, N148, N179, N250, N287, N376, N457, N551, N607, N614, N642, N684, N699, N729, N759, N829, N1009, N1056, N1129	EF-ハンド:NI009-I1037 (p = 7.5e-2)	HMMER-PFAM
					Notch (DSL) ドメイン: K448-S469, N500-G536 蛋白質 PFBB RFBC/D CPSY SABC カプセル遺伝子複合 体 UPD グルコース 4 エピメララーゼ GALE PD011737:S319-Y421 ロイシンジッパー-バタール:L1033-L1068 EF-ハンドカルシウム-結合ドメイン:D1018-I1030	HMMER-PFAM BLAST-PRODOM MOTIFS MOTIFS

【表 1 1】

10

20

30

表 3-8

Eq ID No.	Incyte ポリペプチド ID/	アミノ酸 残基数	潜在的リン酸化部位	潜在的グリコシル化部位	シグネチャ配列、ドメインおよびモチーフ	分析方法及びデータベース
10	118539CD1	364	S8, S20, T22, T35, S42, S50, T93, S97, S172, S173, S184, S190, S273, T281, S315, T343	N4, N152	蛋白質 ポリメラーゼ放出因子 ロイシンジッパー 欠乏症 倉 1 転写物 PD013178:D26-K149	BLAST_PRODOM
11	4005918CD1	331	S38 S47 S87 S165 S203 T110 T196		KRAB-ボックス:V37-A99	HMMER_PPFAM
					Zn フィンガー, C2H2 タイプ: H238-H260, F210-H232, H182-H204, Y266-H288	HMMER_PPFAM
					C2H2 タイプ Zn フィンガー シグネチャ PR00048:P237-R250, L253-G262	BLIMPS_PRINTS
					C2H2 タイプ Zn フィンガー シグネチャ BL00028: C240-H256	BLIMPS_BLOCKS
					蛋白質 Zn フィンガー	BLIMPS_PRODOM
					PD01066:F39-G77 PD00016:H256-C268	BLAST_PRODOM
					Zn フィンガー 金属結合 DNA 結合 PD169183:K264-G331 PD001562:V37-W96	BLAST_PRODOM
					KRAB-ボックス ドメイン DM00605	BLAST_DOMQ
					P52738 3-77:G35-W96	
					P52736 1-72:V37-P109	
					P51786 24-86:G35-W96	
					Zn フィンガー, C2H2 タイプ, ドメイン DM00002 Q05481 831-885: C215-P265	BLAST_DOMO

10

20

30

【表 1 2】

表 3-9

SEQ ID NO.	Incyte ポリペプチド ID/	アミノ酸 残基数	潜在的リン酸化部位	潜在的グリコシ ル化部位	シグネチャ配列、 ドメインおよびモチーフ	分析方法及びデ ータベース
					ATP/GTP 結合部位モチーフ A(P-loop) G183-S190	MOTIFS
12	5435937CD1	670	S24 S162 S237 S321 S330 S362 S405 S517 T14 T36 T95 Y106 Y134 Y475 Y480	N401 N458 N513	Znフィンガー、 C2H2 タイプ、ドメイン; C184-H204, C212-H232, C240-H260, C268-H288 KRAB-ボックス;V4-K74	MOTIFS HMMER_PFM
					Znフィンガー、 C2H2 タイプ: Y307-H329, Y447- H469, Y643-H665, Y251-H273, Y335-H357, Y587-H609, Y195-H217, Y503-H525, Y139-H161, Y475-H497, Y223-H245, Y167-H189, Y531-H553, Y419-H441, H363-H385, Y559-H581, Y391-H413, Y279-H301, Y615-H637 C2H2 タイプ Znフィンガー-シグネチャ PR00048;P222- S235, I630-G639	HMMER_PFM
					蛋白質 Znフィンガー 亜鉛 PD01066:F6-G44 PD00066:H213-C225	BLIMPS_PRODUM
					Znフィンガー PD017719:C247-F484, G191-F428, K383-F624, G219-K473, G443-K644, G163-F400, G415-F652, E133-F372, P474-H665 PD149061:K252-H441 PD149420:V105-Y447, E220-Y531, C141-R331, E416-G648 PD000072:K221-C284, K277-C340, K585-G648, K389-C452	BLAST_PRODUM

10

20

30

表 3-10

Eq ID NO.	Incyte ポリペプチド ID/	アミノ酸 残基数	潜在的リン酸化部位	潜在的グリコシル化部位	シグネチャ配列、ドメインおよびモチーフ	分析方法及びデータベース	
12					KRAB-ボックス ドメイン DM00605 P52737 1-76:MI-E75 I49636 10-85:D2-R64 Znフィンガー、C2H2 タイプ、ドメイン DM00002 Q05481 789-829:K214-E255 Q05481 831-885:C228-Q283 Znフィンガー、C2H2 タイプ、ドメイン C141-H161 C169-H189 C197-H217 C225-H245 C253-H273 C281-H301 C309-H329 C337-H357 C365-H385 C393-H413 C421-H441 C449-H469 C505-H525 C533-H553 C561-H581 C589-H609 C617-H637 C645-H665		BLAST_DOM0 BLAST_DOM0 MOTIFS
13	7503560CDI	1196	S15 S55 S122 S163 S187 S252 S312 S321 S399 S507 S625 S722 S734 S800 S849 S874 S890 S966 S983 S1017 S1120 T85 T119 T150 T214 T236 T284 T369 T560 T626 T793 T863 T870 T889 T988 T1038 T1103 Y520 Y1051	N83 N114 N148 N179 N250 N287 N376 N491 N547 N554 N582 N624 N639 N669 N769 N949 N996 N1069	Signal_cleavage:MI-A34	SPSCAN	
					サイトゾル内ドメイン:MI-G19, K1176-V1196 膜貫通ドメイン:L20-I42, H1153-L1175 非サイトゾル内ドメイン:E43-T1152	TMHMMER	

10

20

30

表 3 - 1 1

BQ ID NO.	Incyte ポリペプチド ID/	アミノ酸 残基数	潜在的リン酸化部位	潜在的グリコシ 化部位	シグネチャ配列、 ドメインおよびモチーフ	分析方法及びヒテ ータベース
14	7472378CD1	2138	S2 S55 S59 S110N57 S167 S168 S249 S297 S380 S407 S437 S536 S587 S600 S634 S699 S721 S796 S860 S863 S953 S965 S972 S1039 S1048 S1070 S1122 S1335 S1362 S1591 S1632 S1853 S2102 T2073 Y126 Y295 Y606 Y1182 Y1193	N185 N212 N280 N284 N341 N413 N476 N483 N534 N896 N970 N1091 N1126 N1258 N1399 N1441 N1655 N1745 N1794 N1840 N1852 N1888 N1904 N1938	蛋白質 RFEC/D CFSY SACB カプセル遺伝子複合体 UP D-グルコース 4-エピメラゼ GALE PD011737;S319-Y421; F871-P947; P74-K120; C1089-I1118; I497-P504 EF_HandID958-1970	BLAST-PRODOM
					ロイジンジッパー; L973-L994, L980-L1001, L987-L1008 「フアーゼ」インテグラーゼ ファミリ;S1945-A2117	MOTIFS
						HMMER_PFAM

10

20

30

【表 1 5】

表 3-12

EQ ID NO.	Incyte ポリペプチド ID/	アミノ酸 残基数	潜在的リン酸化部位	潜在的グリコシル化部位	シグネチャ配列、 ドメインおよびモチーフ	分析方法及びデータベース
14			T12 T13 T23 T81 T100 T109 T130 T133 T181 T214 T272 T286 T400 T425 T453 T509 T615 T621 T624 T677 T697 T739 T830 T889 T998 T1171 T1260 T1275 T1343 T1386 T1426 T1437 T1443 T1467 T1485 T1566 T1687 T1739 T1747 T1817 T1842 T1920 T1925 T1936 T2031 T2034 T2043 T2046		インテグラーゼ コア ドメイン:L1238-N1392	
					逆転写酵素 (RNA-依存性 DNA ポリメラーゼ):D656-V837	
					随在ドメイン:M771-E799 Q1969-L1990 N 末端はサイトゾル内にならない	TMAP
					RNアーゼ H PF00075:P678-N694, V785-T795, I1209-C1219	BLIMPS_PFAM
					「ファージ」インテグラーゼ ファミリー PF00589:P2071-Y2084, M2096-Y2107	

10

20

30

【表 16】

表 3-13

Eq ID No.	Incyte ID/	アミノ酸残基数	潜在的リン酸化部位	潜在的グリコシル化部位	シグネチャ配列、ドメインおよびモチーフ	分析方法及びデータベース
14					蛋白質 ジプシー (GYPSY) エンペローブ レトロウイルス関連 ENV ポリタンポソン 糖タンパク コー トランスポザザブル PD018328:V1512-R1993 蛋白質 GAG GYPSY レトロウイルス関連 ポリタンポソ ンスポゾン コア トランスポザザブル エレメント PRGAGPOL PD009899:565-Y265 レトロウイルス関連 POL ポリタンポソ ン含有:プロテアーゼ 逆転写酵素 エンドヌクレアーゼ PD031814:1484-M652 ポリタンポソ プロテアーゼ DNA 逆転写酵素 ヒドローラーゼ RNA 指令性 (RNA-DIRECTED) アスバルチル ポリメラ ーゼ 含有:PD000970:R875-1999	BLAST_PRODOM
15	I306049CD1	257	S197 S209 S219 S233 S235 S248 T97 T141 T178	N253	POL ポリタンポソ DM00140 P04323 208-577:P629-K1004 POL ポリタンポソ DM00140 P20825 206-576:P629-K1004 POL ポリタンポソ DM00140 P10401 181-564:P629-K1004 POL ポリタンポソ DM00140 P10394 315-689:I630-P1005 ANI-様 zn フィンガー:A103-S142, C10-I52	BLAST_DOMO HAMMER_PPAM

10

20

30

【表 17】

表 3 - 1 4

Eq ID No:	Incyte ポリペプチド ID/	アミノ酸 残基数	潜在的リン酸化部位	潜在的グリコシル化部位	シグネチャ配列、ドメインおよびモチーフ	分析方法及びデータベース
16	3187174CD1	517	S54 S59 S113 S115 S121 S240 S244 S248 S390 S400 S503 T39 T203 T249 T273 T353 T450 Y136 Y489		signal_cleavage.M1-A40	SPSCAN
16					RNA 認識モチーフ (別名 RRM, RBD, または RNP ドメイン): I422-V498, L29-V97 T12M4.6 PD184535:A3-D251, N370-K484, E242-Q267 細胞接着配列 R24-D26 真核生物 推定上 RNA 結合領域 RNP-1 シグネチャ K466-L473	HMER_PFAM BLAST_PRODOM MOTIFS MOTIFS

【 表 1 8 】

10

20

30

表 4-1

ボリヌクレオチド SEQ ID NO/ incyte 求りヌクレオチド ID/ 配列長	配列断片
17/2530775CB1/4456	1-553, 27-67, 33-67, 33-82, 33-89, 40-89, 51-391, 126-238, 126-340, 126-2955, 239-340, 239-425, 341-425, 341-519, 383-972, 426-519, 490-676, 490-949, 490-1025, 490-1066, 490-1084, 490-1140, 514-1137, 520-600, 520-697, 600-788, 601-697, 601-815, 627-905, 627-1045, 637-936, 698-815, 698-995, 722-1039, 726-1155, 747-1000, 747-1001, 754-1389, 767-1039, 770-1238, 788-1272, 796-1163, 802-1457, 816-995, 816-1160, 818-1249, 941-1562, 954-1437, 991-1248, 996-1160, 996-1436, 1046-1577, 1072-1319, 1086-1683, 1088-1367, 1116-1367, 1116-1400, 1117-1595, 1119-1330, 1119-1721, 1156-1546, 1161-1436, 1161-1578, 1166-1804, 1246-1795, 1355-1983, 1381-1976, 1437-1578, 1466-1923, 1533-2131, 1554-1814, 1579-1736, 1621-1897, 1666-2182, 1668-1954, 1669-2068, 1685-2037, 1709-1959, 1709-2145, 1733-1941, 1735-1937, 1737-1865, 1737-1997, 1761-2006, 1761-2016, 1787-2401, 1866-1997, 1866-2182, 1920-2198, 1931-2285, 1931-2339, 1932-2175, 1932-2412, 1972-2217, 1996-2238, 1998-2182, 1998-2357, 2003-2617, 2023-2602, 2083-2337, 2101-2292, 2101-2635, 2181-2315, 2182-2476, 2182-2658, 2182-2665, 2183-2357, 2183-2464, 2203-2657, 2205-2661, 2216-2656, 2230-2436, 2286-2727, 2351-2653, 2357-2819, 2362-2464, 2362-2894, 2370-2731, 2372-2634, 2375-2618, 2380-2627, 2383-2689, 2384-2797, 2428-2669, 2445-2705, 2452-2688, 2465-2894, 2465-2955, 2483-3035, 2534-2955, 2736-3364, 2781-2943, 2781-2944, 2791-3004, 2791-3075, 2791-3098, 2828-3089, 2828-3120, 2837-3103, 2844-3353, 2895-2955, 2953-3054, 2954-3101, 2994-3023, 2994-3128, 2996-3219, 2996-3438, 3020-3288, 3024-3128, 3024-3197, 3042-3301, 3042-3310, 3092-3365, 3094-3389, 3132-3197, 3153-3260, 3192-3437, 3208-3479, 3209-3758, 3231-3357, 3263-3480, 3367-3885, 3391-3960, 3400-3856, 3418-3669, 3425-4060, 3468-3700, 3468-4028, 3551-4084, 3584-3839, 3584-3855, 3767-4228, 3924-4169, 3967-4212, 3967-4456

10

20

30

【表 19】

表 4 - 2

ボリスクレオチド SEQ ID NO/ Inocyte ポリヌクレオチド ID/ 配列長	配列断片
18/926296CBI/ 3662	1-261, 1-471, 10-294, 12-243, 12-274, 12-438, 60-331, 87-387, 216-773, 250-742, 339-617, 376-524, 442-848, 481-803, 489-893, 546-687, 619-1137, 631-1153, 666-1152, 881-1403, 926-1546, 1109-1501, 1109-1505, 1223-1476, 1223-1616, 1289-1930, 1339-1554, 1381-1634, 1436-1679, 1488-1799, 1493-2000, 1512-2128, 1513-1712, 1513-1752, 1513-2033, 1690-1988, 1727-2106, 1728-2134, 1821-2021, 1829-2093, 1835-2106, 1877-2158, 1879-2068, 1993-2205, 2053-2378, 2053-2431, 2078-2578, 2096-2604, 2120-2366, 2122-2295, 2122-2353, 2122-2584, 2183-2441, 2183-2455, 2217-2501, 2229-2493, 2348-2540, 2280-2501, 2288-2560, 2293-2550, 2304-2564, 2313-2978, 2323-2913, 2379-2633, 2397-3058, 2408-2822, 2422-2673, 2422-2740, 2540-2840, 2576-3017, 2613-2773, 2709-2969, 2714-2980, 2714-3073, 2714-3164, 2714-3173, 2714-3219, 2718-2985, 2719-3223, 2760-3017, 2760-3019, 2765-3226, 2794-3045, 2920-3148, 2928-3155, 2931-3410, 2994-3198, 3030-3560, 3041-3474, 3057-3480, 3069-3611, 3072-3324, 3090-3535, 3098-3536, 3113-3617, 3162-3375, 3165-3640, 3169-3616, 3179-3344, 3179-3655, 3186-3662, 3194-3433, 3215-3616, 3233-3662, 3238-3546, 3250-3648, 3267-3604, 3267-3654, 3285-3655, 3300-3613, 3314-3648, 3354-3614, 3356-3598, 3357-3655
19/1322761CBI/ 2201	1-116, 1-158, 1-527, 2-694, 41-111, 166-801, 274-500, 305-516, 375-1019, 384-1038, 424-726, 466-1145, 575-758, 619-877, 632-1267, 797-1060, 886-1136, 915-1174, 937-1189, 997-1417, 1000-1272, 1016-1545, 1071-1439, 1158-1457, 1203-1749, 1216-1816, 1218-1433, 1218-1542, 1314-1597, 1395-1739, 1418-1563, 1418-1880, 1539-1784, 1562-1816, 1587-1825, 1591-1924, 1690-2175, 1722-2201, 1724-1960, 1724-2197, 1727-2118, 1760-2182, 1761-2184, 1763-2182, 1767-2182, 1767-2183, 1768-2187, 1777-2186, 1784-2182, 1797-2182, 1798-2182, 1821-2201, 1869-2182, 1899-2186, 1903-2186, 1993-2186, 2038-2182, 2038-2195

【表 20】

表 4 - 3

ボリヌクレオチド SEQ ID NO/ ヌクレオチド ID/ 配列長	ボリヌクレオチド SEQ ID NO/ ヌクレオチド ID/ 配列長	配列断片
20/7472664CBL/ 3188		1-225, 1-505, 1-552, 1-614, 141-612, 183-756, 232-818, 232-843, 238-830, 243-598, 276-890, 283-602, 295-943, 419-801, 420-973, 431-541, 431-937, 450-541, 451-541, 452-541, 456-973, 460-541, 464-541, 464-1029, 465-946, 466-948, 466-973, 466-1017, 550-1106, 564-1058, 866-1273, 925-1273, 939-1525, 949-1603, 971-1448, 972-1165, 974-1162, 974-1273, 974-1351, 1005-1417, 1037-1598, 1086-1394, 1096-1277, 1134-1418, 1163-1273, 1163-1351, 1274-1351, 1274-1447, 1274-1448, 1296-1568, 1354-1448, 1397-1491, 1405-1762, 1431-2028, 1489-1734, 1489-2051, 1489-2073, 1489-2116, 1491-1738, 1505-2112, 1509-2033, 1511-1773, 1528-2096, 1528-2138, 1528-2139, 1581-1691, 1581-1694, 1583-2191, 1590-1854, 1615-2191, 1648-2086, 1656-2221, 1669-2217, 1672-1920, 1672-2169, 1750-2221, 1791-2073, 1816-2449, 1835-2447, 1839-2067, 1839-2407, 1841-2293, 1856-2274, 1880-2139, 1885-2123, 1913-2555, 1917-2467, 1919-2527, 1924-2495, 1926-2185, 1928-2496, 1941-2530, 1950-2570, 1951-2487, 1960-2501, 1961-2467, 1961-2554, 1967-2454,
		1981-2576, 2005-2352, 2018-2672, 2021-2666, 2034-2627, 2039-2713, 2041-2123, 2049-2671, 2055-2618, 2071-2561, 2082-2563, 2085-2206, 2094-2670, 2098-2321, 2108-2671, 2129-2345, 2133-2345, 2138-2807, 2148-2775, 2160-2731, 2190-2715, 2196-2579, 2209-2735, 2231-2798, 2240-2724, 2258-2904, 2259-2775, 2292-2572, 2300-2913, 2304-2898, 2331-2866, 2333-2913, 2337-2624, 2358-2968, 2366-3024, 2439-3061, 2441-2973, 2445-2516, 2447-3171, 2459-2854, 2462-3029, 2468-2950, 2490-3039, 2492-3030, 2501-2733, 2516-3007, 2547-3108, 2575-3090, 2576-3142, 2618-3188, 2621-3171, 2698-3188, 2725-3172, 2729-3188, 2741-3167, 2743-3188, 2763-3158, 2784-3188, 2836-3188, 2851-3188, 2854-3184, 2981-3098
21/7473124CBL/ 843		1-312, 1-843, 28-837, 196-651, 313-843, 493-837, 799-829
22/7473171CBL/ 1102		1-55, 1-222, 1-288, 3-55, 51-432, 68-589, 68-685, 96-432, 425-929, 426-728, 426-827, 443-926, 445-876, 448-883, 448-930, 484-939, 590-675, 590-685, 590-1010, 590-1059, 595-891, 595-1078, 595-1102, 649-1099, 684-1097, 692-1076, 816-1102, 870-1097

【表 2 1】

表 4-4

ボリスクレオチド SEQ ID NO/ Incyte ボリスクレオチド ID/ 配列長	配列断片
23/7477026CB1/ 2481	1-660, 22-221, 108-389, 136-311, 140-311, 414-910, 460-1037, 571-862, 571-1033, 635-915, 786-1107, 944-1523, 976-1642, 1071-1333, 1071-1609, 1276-1787, 1315-1914, 1356-1969, 1446-1651, 1467-2028, 1474-1871, 1556-1824, 1636-1967, 1718-2375, 1737-1986, 1786-2162, 1795-2241, 1802-2481, 1855-2123, 1866-2470, 1912-2384, 1964-2229, 2023-2305, 2073-2436, 2085-2426, 2098-2327, 2137-2392, 2168-2317
24/6428773CB1/ 2308	1-248, 1-263, 1-268, 1-277, 1-551, 10-251, 19-280, 19-309, 19-466, 19-552, 28-551, 45-528, 52-550, 62-262, 63-300, 63-551, 80-551, 81-533, 107-551, 111-528, 113-551, 118-551, 127-528, 130-542, 137-528, 139-551, 169-528, 170-551, 172-456, 186-552, 187-470, 187-546, 213-526, 236-442, 243-482, 275-552, 342-911, 370-1004, 431-551, 544-797, 544-918, 544-986, 544-1031, 546-973, 563-806, 563-1050, 608-871, 618-800, 633-864, 633-1040, 662-1140, 669-918, 748-1151, 762-1027, 817-1127, 857-1221, 878-1146, 899-1150, 899-1497, 948-1238, 953-1050, 971-1226, 983-1249, 1021-1300, 1026-1294, 1026-1512, 1050-1484, 1055-1310, 1076-1359, 1087-1344, 1091-1271, 1092-1273, 1101-1369, 1104-1405, 1107-1354, 1111-1386, 1120-1315, 1120-1571, 1164-1392, 1176-1412, 1180-1468, 1181-1461, 1199-1554, 1211-1490, 1239-1513, 1253-1542, 1257-1409, 1257-1517, 1257-1585, 1257-1657, 1258-1532, 1315-1582, 1320-1571, 1323-1636, 1325-1777, 1341-1584, 1342-1823, 1404-1936, 1408-1933, 1439-1754, 1446-2064, 1446-2080, 1464-1715, 1468-1790, 2027-2291, 2028-2302, 2029-2308, 2036-2304, 2055-2296, 2061-2298, 2062-2307, 2062-2308, 2067-2308, 2069-2303,

10

20

30

【表 2 2】

表 4 - 5

ボリヌクレオチド SEQ ID NO/ Incyte ボリヌクレオチド ID/ 配列長	配列断片
	1468-2100, 1479-1715, 1504-1777, 1507-1711, 1550-2116, 1550-2160, 1555-2086, 1555-2217, 1575-1846, 1581-2086, 1581-2160, 1594-1911, 1614-1812, 1614-1818, 1614-2099, 1614-2143, 1628-1794, 1629-1942, 1629-2151, 1629-2160, 1630-1907, 1638-1869, 1658-2064, 1658-2160, 1660-1842, 1660-1843, 1662-1920, 1664-2308, 1690-1943, 1726-2295, 1731-2191, 1732-1990, 1732-2002, 1732-2185, 1732-2294, 1756-2280, 1758-1958, 1767-2029, 1767-2289, 1776-2160, 1777-1973, 1778-2308, 1784-2273, 1800-2307, 1801-2308, 1822-2289, 1827-2288, 1829-2308, 1841-2080, 1841-2291, 1845-2308, 1846-2060, 1852-2085, 1855-2302, 1861-2273, 1867-2307, 1867-2308, 1872-2123, 1874-2308, 1890-2167, 1890-2302, 1897-2289, 1918-2190, 1920-2305, 1922-2165, 1922-2202, 1923-2195, 1923-2217, 1939-2305, 1941-2256, 1944-2230, 1957-2301, 1958-2301, 1967-2190, 1967-2281, 1969-2302, 1978-2302, 1986-2220, 1989-2198, 2000-2301, 2000-2308, 2001-2304, 2003-2305, 2004-2291, 2005-2301, 2008-2268, 2024-2305, 2025-2304, 2026-2295, 2074-2308, 2076-2288, 2094-2308, 2096-2289, 2120-2301, 2124-2308, 2129-2289, 2146-2303, 2158-2308, 2160-2308, 2162-2289, 2216-2291, 2239-2289, 2243-2308
25/2749402CB1/4369	1-337, 1-400, 1-433, 1-437, 1-470, 1-610, 20-281, 20-336, 125-436, 125-437, 128-806, 130-336, 150-425, 160-336, 265-425, 359-451, 431-1116, 546-1110, 565-904, 587-1275, 632-904, 667-957, 677-944, 746-1242, 812-1451, 817-1240, 854-1518, 958-1200, 958-1229, 1083-1433, 1186-1463, 1240-1745, 1250-1532, 1275-1557, 1297-1562, 1307-1622, 1383-1674, 1448-2015, 1491-1745, 1554-1745, 1606-1745, 1649-1745, 1651-1745, 1654-1886, 1654-2051, 1661-1745, 1662-1745, 1671-1809, 1671-2173, 1671-2206, 1671-2213, 1671-2365, 1731-1950, 1757-2003, 1770-2057, 1808-2641, 1939-2213, 1983-2025, 2083-2389, 2137-2734, 2149-2645, 2180-2451, 2224-2735, 2301-2848, 2420-2848, 2478-2719, 2478-2993, 2556-3191, 2594-2869, 2594-3056, 2604-3076, 2647-3169, 2671-2938, 2683-2890, 2702-2970, 2765-3027, 2857-3484, 2924-3415, 2943-3175, 2968-3226, 3058-3342, 3118-3375, 3118-3630, 3164-3441, 3210-3350, 3216-3440, 3244-3564, 3246-3668, 3269-3382, 3276-3533, 3284-3568, 3412-3919, 3448-3703, 3460-3735, 3487-3775, 3566-3755, 3567-3735, 3568-3702, 3568-3735, 3600-4057, 3647-4317, 3668-4239, 3677-3998, 3735-4003, 3736-3826, 3736-4280, 3745-4236, 3762-4369, 3766-3961, 3771-4027

【表 2 3】

表 4 - 6

ポリスクレオチド SEQ ID NO/ Incyte ポリ スクレオチド ID/ 配列長	配列断片
26/118539CBI/ 1445	1-427, 19-289, 112-397, 161-508, 169-341, 169-721, 250-765, 250-781, 361-1028, 375-1007, 703-1245, 776-1050, 776-1240, 900-1445, 945- 1120, 1022-1300, 1083-1316, 1149-1444
27/4005918CBI/ 1608	1-556, 4-163, 4-488, 49-407, 160-702, 163-677, 203-761, 223-496, 227-435, 236-488, 287-919, 351-958, 489-702, 510-745, 531-702, 625- 881, 635-1205, 676-971, 680-1259, 739-920, 739-1373, 873-1487, 882- 1354, 906-1523, 992-1318, 1148-1451, 1149-1608, 1151-1402, 1205- 1435, 1251-1548, 1327-1572, 1333-1608, 1354-1592
28/5435937CBI/ 2275	1-214, 1-215, 1-221, 1-237, 38-237, 43-235, 63-713, 73-612, 76-471, 77-391, 102-407, 102-667, 102-741, 103-635, 277-551, 284-588, 284- 812, 284-328, 304-636, 401-914, 504-745, 782-977, 782-1408, 782- 1516, 840-1232, 840-1437, 859-1535, 861-1457, 897-1410, 897-1509, 916-1585, 918-1588, 1014-1559, 1056-1725, 1064-1331, 1064-1607, 1094-1350, 1094-1752, 1149-1763, 1163-1833, 1290-1565, 1291-1532, 1366-1972, 1380-1956, 1446-2275, 1606-2110, 1637-2165, 1638-1320, 1677-1979, 1678-2275, 1689-2275, 1761-2275, 1770-2275, 1971-2275, 2092-2275
29/7503560CBI/ 4277	1-477, 2-407, 2-443, 2-444, 8-617, 8-4193, 27-288, 141-843, 141-859, 141-907, 438-1123, 553-1117, 597-1282, 674-964, 684-951, 753-1249, 819-1442, 824-1247, 843-1440, 965-1207, 965-1236, 965-1486, 986- 1424, 1062-1289, 1119-1394, 1481-1868, 1506-1636, 1506-2000, 1506- 2033, 1506-2042, 1506-2192, 1584-1830, 1597-1894, 1635-2468, 1726- 2310, 1766-2040, 1766-2267, 1771-2366, 1872-2366, 1910-2316, 1948- 2466, 1964-2361, 2001-2513, 2007-2278, 2053-2562, 2089-2500, 2127- 2875, 2239-2744, 2247-2676, 2305-2546, 2305-2820, 2367-2514, 2371- 2839, 2382-2839, 2383-3018, 2421-2696, 2421-2883, 2431-2903, 2474- 2996, 2498-2765, 2510-2817, 2529-2797, 2593-2854, 2641-3547, 2684- 3311, 2751-3242, 2770-3002, 2795-3053, 2876-3093, 2885-3169, 2945- 3202, 2945-3457, 2966-3733, 2991-3268, 3037-3177, 3043-3267, 3063- 3733, 3073-3495, 3103-3360, 3149-3409, 3169-3728, 3186-3891, 3239- 3746, 3256-3526, 3275-3530, 3307-3562, 3316-3602, 3355-3396, 3395- 3797, 3427-3883, 3474-4147, 3504-3825, 3563-3830, 3572-4063, 3599-4189, 3594-3788, 3598-3854, 3641-3903, 3652-3862, 3652-3936, 3672-3934, 3672-4248, 3685-4050, 3723-4159, 3742-4172, 3760-4257, 3812-4274, 3812-4277, 3840-4106, 3840-4118, 3870-4148, 3878-4140, 3936-4196

【表 2 4】

表 4 - 7

ボリスグレオチドSEQ ID NO/ Inceyte ボリスグレオチド ID/ 配列数	配列断片
30/7472378CBI/ 6417	1-6017, 157-6417
31/1306049CBI/ 1194	1-224, 1-236, 1-288, 1-579, 15-280, 19-277, 19-294, 22-259, 22-268, 22-283, 23-492, 24-224, 24-263, 25-241, 25-267, 28-331, 30-271, 31-217, 31-229, 31-269, 31-548, 31-627, 32-550, 33-156, 34-141, 34-323, 34-369, 34-492, 35-309, 39-393, 39-488, 39-503, 56-188, 56-197, 56-202, 56-208, 56-221, 56-237, 56-253, 56-256, 56-280, 56-289, 56-300, 56-347, 56-401, 56-461, 56-482, 56-532, 56-567, 62-226, 74-337, 75-319, 75-441, 85-225, 86-290, 86-302, 86-515, 108-383, 115-224, 129-225, 136-350, 148-409, 153-418, 155-467, 161-434, 167-445, 190-448, 213-560, 231-466, 231-725, 243-527, 247-501, 247-669, 259-470, 259-472, 269-861, 264-478, 294-326, 303-326, 308-534, 308-609, 308-735, 308-901, 309-483, 354-743, 356-599, 367-657, 370-605, 378-614, 381-647, 385-524, 395-664, 413-662, 418-601, 429-695, 429-710, 429-759, 429-898, 451-669, 452-685, 452-686, 452-693, 452-756, 452-764, 460-706, 461-724, 475-727, 475-1008, 485-804, 487-1062, 492-1173, 500-1168, 513-601, 515-601, 520-1173, 522-1171, 556-1182, 573-816, 583-804, 592-883, 600-668, 600-839, 600-842, 604-897, 606-1177, 607-1194, 611-730, 611-742, 623-730, 630-878.
	631-1173, 632-1173, 645-730, 662-730, 670-1194, 673-944, 692-985, 703-1186, 709-1194, 725-792, 725-799, 725-800, 725-804, 728-1188, 731-804, 735-793, 738-804, 740-804, 740-990, 743-799, 744-1194, 757-1186, 762-1186, 775-799, 775-804, 780-1194, 798-1186, 802-870, 802-892, 802-945, 802-950, 802-951, 802-958, 802-988, 802-1010, 802-1012, 802-1016, 802-1026, 802-1046, 802-1053, 802-1070, 802-1075, 802-1087, 802-1120, 802-1150, 802-1159, 802-1166, 802-1169, 802-1171, 802-1173, 802-1180, 802-1184, 802-1186, 802-1194, 803-830, 803-878, 804-882, 804-983, 804-1194, 805-995, 805-1181, 805-1194, 806-1090, 806-1092, 808-926, 833-1188, 846-1180, 849-1164, 853-1084, 856-942, 861-1123, 861-1175, 861-1186, 897-1046, 906-1186, 931-1185, 965-1186, 965-1193, 973-1127, 974-1081, 1034-1186, 1045-1159, 1072-1184, 1124-1186

10

20

30

【表 2 5】

表 4 - 8

ポリヌクレオチド SEQ ID NO/ Incyte ポリヌクレオチド ID/ 配列長	配列断片
32/3187174CB1/2801	1-231, 1-250, 1-464, 1-569, 12-623, 17-329, 61-716, 62-711, 94-346, 303-481, 311-867, 372-754, 394-835, 395-641, 395-879, 395-927, 396-958, 405-650, 419-954, 456-727, 488-921, 496-945, 524-803, 531-943, 548-937, 594-1081, 646-936, 647-893, 651-1089, 678-937, 679-1338, 766-990, 768-1304, 912-1129, 1181-1927, 1203-1455, 1265-1839, 1351-1584, 1470-1804, 1476-1708, 1528-2049, 1548-2095, 1559-2317, 1690-2375, 1701-1924, 1726-2000, 1737-2366, 1750-2228, 1776-2447, 1807-1987, 1809-1965, 1865-2116, 1910-2382, 1912-2496, 1974-2230, 1974-2601, 2001-2158, 2002-2323, 2045-2267, 2045-2340, 2152-2801, 2153-2420, 2220-2448, 2237-2798, 2264-2623, 2274-2515, 2333-2458

【表 2 6】

10

20

30

表 5

ホリスケレオチド SEQ ID NO:	Incyte プロジェクト ID	代表的ライブラリ
17	2530775CB1	PROSNOT26
18	926296CB1	LUNGTUT09
19	1322761CB1	KIDNOT05
20	7472684CB1	MYEPTX02
22	7473171CB1	TONSNOT03
23	7477026CB1	SKIRNOR01
24	6428773CB1	FIBPFER06
25	2749402CB1	MCLDXT02
26	118539CB1	LATRNOT01
27	4005918CB1	LUNGNOM07
28	5435937CB1	SPLNNOT17
29	7503560CB1	MCLDXT02
31	1306049CB1	UCMCL5T01
32	3187174CB1	BRAITUT13

10

20

【表 27】

表 6 - 1

ライブラリ BRATTUT13	ベクター pINCY	ライブラリの説明 ライブラリは、68才の白人男性の脳腫瘍の切除時、左前頭葉から採取した脳腫瘍組織から単離した RNA を用いて作製した。病理は左前頭葉に髄膜腫を示した。
FIBPFEN06	pINCY	標準化された前立腺癌細胞系ライブラリは前立腺癌細胞系ライブラリからの156万個の独立性クロームから作製した。開始 RNA は妊娠26週間後死亡の男子胎児から採取した前立腺癌質の線維芽細胞から作製した。このライブラリは、極めて長時間(48時間/1回)の再アニーリングによるハイブリダイゼーションを用いた他は、Soares 他, PNAS (1994) 91:9228 および Bonaldo 他, Genome Research (1996) 6:791 を改作した条件を用いて2回にわたり標準化した。このライブラリは次に、下記のようにクロームを有する挿入断片用に選択するために直線化し、再度線状にした。プラスミド DNA の調製は、約100万のクローム (由来は正統化した前立腺癌質線維芽細胞ライブラリ) から行い、ソフトアガー形質転換後に行った。
KIDNUT05	PSPORT1	ライブラリは、脳腫瘍で死亡した2才のヒスバニツク系女子の腎組織から単離された RNA を用いて作製された。家族歴には、先天性心疾患がある。
LATRN0F01	PEL/ESCRIP1	ライブラリは、頭蓋内出血で死亡した51才白人女性の左心房から単離された RNA を用いて作製された。
LUNGNO07	pINCY	この標準化した肺組織ライブラリは、或る肺組織ライブラリの510万個の独立性クロームから作製した。開始 RNA の作製は、肺組織から単離した RNA から行った。このライブラリは、極めて長時間(48時間/1回)の再アニーリングによるハイブリダイゼーションを用いた他は、Soares 他, PNAS (1994) 91:9228-9232 および Bonaldo 他, Genome Research (1996) 6:791 を適用した条件を用いて2回にわたりノーモライズした。
LUNGUT09	pINCY	ライブラリは、部分的肺切除時に68才白人男性から取り除かれた肺腫瘍組織から単離された RNA を用いて作製された。病理学の示すところでは侵襲性グレード3の扁平上皮癌および転移性腫瘍である。患者の病歴には、II型糖尿病、甲状腺腫瘍、鬱病、高脂血症、食道潰瘍と喫煙が含まれる。
MCLDTXT02	pINCY	ライブラリは、男性から採取し、処理した臍帯血液の樹状細胞から単離された RNA を用いて作製した。この細胞は顆粒球/マクロファージコロニー刺激因子(GM-CSF)、腫瘍壊死因子アルファ(TNFアルファ)、幹細胞因子(SCF)、ホルホルミリスチンチン酸アセテート(PMA)およびイノマイシンで処理した。GM-CSFは、時間0に、100 ng/ml で加え、TNFアルファは時間0に、2.5ng/ml で加え、SCFは時間0に、25ng/ml で加えた。PMA およびイノマイシンは13日に5時間加えた。インキュベーション時間は13日間であった。
MYEFTXT02	pINCY	ライブラリは53才の1女性から採取された慢性骨髄性白血病患者細胞由来の処理された K-562 細胞系から単離した RNA を用いて作製した。細胞は1 mM の PMA で96時間処理した。

【表 2 8】

表 6-2

ライブラリ	ベクター	ライブラリの説明
PROSNOT26	pJNCY	ライブラリは、根治的前立腺摘除時に65才白人男性から取り除かれた前立腺組織から単離されたRNAを用いて作製された。随伴する腫瘍組織の病理学検査では、腺癌が見られた。この患者は前立腺特異抗原(PSA)の上昇も示した。家族歴には悪性胃新生物が含まれていた。
SKIRNOR01	PCDNA2.1	このランダムプライムされたライブラリは、17才の白人女性の両側小乳房形成術時に採取した皮膚組織から単離したRNAを用いて作製した。患者の病歴には乳房肥大が含まれる。家族歴には良性高血圧が含まれている。
SPLANNOT17	pJNCY	ライブラリは、脳無糖素症で死亡した2才ヒスバニック系男子の脾臓組織から単離されたRNAを用いて作製された。
TONSNOT03	pJNCY	ライブラリは6才の白人男子の扁桃腺アデノイド切除中に除去した腫瘍した左扁桃腺組織から単離したRNAを用いて作製した。病理学検査では反応性リンパ組織増殖症が両側に見られた。家族歴には、良性高血圧症、心筋梗塞およびアテローム性冠動脈疾患が含まれている。
UCMCL5T01	PBLUESCRIPT	ライブラリは、12個体の臍帯血から得た単核細胞から単離されたRNAを用いて作製した。細胞の培養を12日間、IL-5を用いて行った後、RNAをプールのライゼットから得た。

【表 2 9】

10

20

30

【表 30】

表 7-1

プログラム	説明	参照文献	パラメータ閾値
ABI FACTURA	ベクター配列を除き、核糖配列のまぎらわしい塩基をマスクするプログラム。	Applied Biosystems, Foster City, CA.	
ABI/PARACEL FDF	アミノ酸または核糖配列の比較、注釈に有用なファーストデーターアラインダ。	Applied Biosystems, Foster City, CA. Paracel Inc., Pasadena, CA.	不一致<50%
ABI AutoAssembler	核糖配列の構築を行うプログラム。	Applied Biosystems, Foster City, CA.	
BLAST	アミノ酸および核糖配列の配列類似性検索に有用なベクタローカルアライメント検索ツール (Basic Local Alignment Search Tool)。BLASTには5つの機能がある: blastp、blastn、blastx、tblastnおよびtblastx。	Altschul, S.F. 他 (1990) J. Mol. Biol. 215:403-410; Altschul, S.F. 他 (1997) Nucleic Acids Res. 25:3389-3402.	EST: 確率値=1.0E-8 以下 完全長配列: 確率値=1.0E-10以下
FASTA	一群の同じタイプの配列と問合せ配列との間の類似性について検索する Pearson および Lipman アルゴリズム。FASTAは最低5つの機能からなる: fasta、tfasta、fastx、tfastx および ssearch。	Pearson, W. R. および D. J. Lipman (1988) Proc. Natl. Acad. Sci. USA 85:2444-2448; Pearson, W. R. (1990) Methods Enzymol. 183:63-98. Smith, T.F. および M.S. Waterman (1981) Adv. Appl. Math. 2:482-489.	EST: fasta E 値=1.0E-6 構築された EST: fasta n 性=95%以上 一致した長さ=200塩基以上 fasta E 値=1.0E-8以下 完全長配列: fastx スコア=100以上
BLIMPS	配列をBLOCKS、PRINTS、DOMO、PRODOM および PFAM データベースの配列と対応させて遺伝子ファミリー、配列相同性および構造的アイソマープリント領域を検索するBLOCKS IMPROVED Searchet。	Henikoff, S. および J. G. Henikoff (1991) Nucleic Acids Res. 19:6565-6572; Henikoff, J. G. および S. Henikoff (1996) Methods Enzymol. 266:88-105. Attwood, T.K. 他 (1997) J. Chem. Inf. Comput. Sci. 37:417-424.	確率値=1.0E-3以下
HMMER	問合せ配列をPFAMのようなタンパク質ファミリーコンセンサス配列の隠れマルコフモデル (HMM) に基づいたデータベースに対して検索するアルゴリズム。	Krogh, A. 他 (1994) J. Mol. Biol. 235:1501-1531; Sonnhammer, E.L.L. 他 (1998) Nucleic Acids Res. 26:320-322; Durbin, R. 他 (1998) Our World View, in a Nutshell, Cambridge Univ. Press, 1-350 ページ。 1-350.	PFAM ヒット 確率値=1.0E-3 以下 シグナル ペプチド ヒット: スコア=0 以上

10

20

30

表 7-2

プログラム	説明	参照文献	パラメータ閾値
ProfileScan	Prositeで定義された配列パターンと一致するタンパク質配列の構造的モチーフおよび配列モチーフを探索するアルゴリズム。	Gribskov, M. 他 (1988) CABIOS 4:61-66; Gribskov, M. 他 (1989) Methods Enzymol. 183:146-159; Bairoch, A. 他 (1997) Nucleic Acids Res. 25:217-221.	標準化された質スコア≧特定のPrositeモチーフに対するGCG指定「HIGH」値 通常、スコア=1.4~2.1.
Phred	高度の感受性および確率のある自動シーケンサートレースを調べる塩基コーリングアルゴリズム。	Ewing, B. 他 (1998) Genome Res. 8:175-185; Ewing, B. および P. Green (1998) Genome Res. 8:186-194.	
Plucap	Smith-Waterman アルゴリズムの効率的実行に基づいたプログラムである、SWATとCrossMatchを含むPhils改訂構築プログラムで、配列同源性検索やDNA配列の構築に有用である。	Smith, T. F. および M. S. Waterman (1981) Adv. Appl. Math. 2:482-489; Smith, T. F. および M. S. Waterman (1981) J. Mol. Biol. 147:195-197; Green, P., University of Washington, Seattle, WA.	スコア=120以上 一致した長さ=56以上
Consed	Phrap 構築の表示、編集用グラフィカルツール。	Gordon, D. 他 (1998) Genome Res. 8:195-202.	
SPScan	タンパク質配列をスキャンし、分泌シグナルペプチドの存在を調べる重量マトリックス分析プログラム。	Nielson, H. 他 (1997) Protein Engineering 10:1-6; Claverie, J.-M. および S. Audic (1997) CABIOS 12:431-439.	スコア=3.5以上
TMAP	重量マトリックスを使ってタンパク質配列上の模写通部分を描写し、方向を決定するプログラム。	Persson, B. および P. Argos (1994) J. Mol. Biol. 237:182-192; Persson, B. および P. Argos (1996) Protein Sci. 5:363-371.	
TREBMER	隠れマルコフモデル(HMM)を使ってタンパク質配列上の模写通部分を描写し、方向を決定するプログラム。	Sonnhammer, E.L. 他 (1998) Proc. Sixth Intl. Conf. on Intelligent Systems for Mol. Biol., Glasgow 他, 編集, The Am. Assoc. for Artificial Intelligence Press, Menlo Park, CA, 175-182ページ。 175-182.	
Motifs	Prositeで定義されたものと一致するパターンのアミノ酸配列を検索するプログラム。	Bairoch, A. 他 (1997) Nucleic Acids Res. 25:217-221; Wisconsin Package Program Manual,バージョン9, M51-59ページ, Genetics Computer Group, Madison, WI	

10

20

30

【配列表】

<110> INCYTE GENOMICS, INC.
 BAUGHN, Mariah R.
 LU, Yan
 ARVIZU, Chandra
 RAMKUMAR, Jayalaxmi
 YAO, Monique G.
 POLICKY, Jennifer L.
 WALIA, Narinder K.
 TRIBOULEY, Catherine M.
 YUE, Henry
 BATRA, Sajeev
 DING, Li
 LAL, Preeti G.
 BOROWSKY, Mark L.
 LU, Dyung Aina M.
 GANDHI, Ameena R.
 GRIFFIN, Jennifer A.
 XU, Yuming
 AZIMZAI, Yalda
 GIETZEN, Kimberly J.
 TANG, Y. Tom
 WARREN, Bridget A.
 MASON, Patricia M.
 BURFORD, Neil
 HAFALIA, April J.A.
 LEE, Ernestine A.
 YANG, Junming
 GORVAD, Ann E.
 EMERLING, Brooke M.
 MARQUIS, Joseph P.
 LEE, Soo Yeun
 SWARNAKAR, Anita
 REDDY, Reddy

<120> NUCLEIC ACID-ASSOCIATED PROTEINS

<130> PF-0869 PCT

<140> To Be Assigned
 <141> Herewith

<150> 60/257,714; 60/260,081; 60/262,302; 60/263,823; 60/266,088; Unassigned
 <151> 2000-12-21; 2001-01-05; 2001-01-16; 2001-01-23; 2001-02-02; 2001-10-26

<160> 32

<170> PERL Program

<210> 1
 <211> 1096
 <212> PRT
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 2530775CD1

<400> 1
 Met Asp Cys Asn Ala Cys Met Ser Glu Glu Leu Trp Gly Met Phe
 1 5 10 15
 Lys Thr Phe Pro Tyr Gln His Arg Tyr Arg Leu Tyr Gly Gln Trp
 20 25 30
 Lys Asn Glu Thr Tyr Asn Ser His Pro Leu Leu Val Lys Val Lys

10

20

30

40

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

10

20

30

40

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

10

20

30

40

```

          980          985          990
Asn Asp Arg Glu Val Pro Pro Asp Leu Thr Lys Arg Arg Lys Glu
          995          1000          1005
Glu Asn Gly Thr Met Gly Val Ser Lys His Lys Ser Glu Ser Pro
          1010          1015          1020
Cys Glu Ser Pro Tyr Pro Asn Glu Lys Asp Lys Glu Lys Asn Lys
          1025          1030          1035
Ser Lys Ser Ser Gly Lys Glu Lys Gly Ser Asp Ser Phe Lys Ser
          1040          1045          1050
Glu Lys Met Asp Lys Ile Ser Ser Gly Gly Lys Lys Glu Ser Arg
          1055          1060          1065
His Asp Lys Glu Lys Ile Glu Lys Lys Glu Lys Arg Asp Ser Ser
          1070          1075          1080
Gly Gly Lys Glu Glu Lys Lys His His Lys Ser Ser Asp Lys His
          1085          1090          1095
Arg

```

10

```

<210> 2
<211> 995
<212> PRT
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Incyte ID No: 926296CD1

```

```

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

```

20

30

40

Leu	Arg	Ala	Val	260	Ala	Lys	Lys	Val	His	265	Ser	Thr	Ile	Gln	Lys	Ile	270
				275						280							285
Lys	Lys	Glu	Gly	290	Lys	Met	Ser	Glu	Cys	295	Leu	Leu	Lys	Ala	Met	Leu	300
Asn	Cys	Lys	Thr	305	Phe	Glu	Glu	Leu	Glu	310	His	Val	Ser	Ala	Pro	Tyr	315
Lys	Thr	Gly	Ser	320	Lys	Gly	Thr	Lys	Ala	325	Gln	Arg	Ala	Arg	Gln	Leu	330
Gly	Leu	Glu	Gly	335	Ala	Arg	Ala	Leu	Leu	340	Glu	Lys	Pro	Gly	Glu		345
Leu	Ser	Leu	Leu	350	Ser	Tyr	Ile	Arg	Pro	355	Asp	Val	Lys	Gly	Leu	Ser	360
Thr	Leu	Gln	Asp	365	Ile	Glu	Ile	Gly	Val	370	Gln	His	Ile	Leu	Ala	Asp	375
Met	Ile	Ala	Lys	380	Asp	Lys	Asp	Thr	Leu	385	Asp	Phe	Ile	Arg	Asn	Leu	390
Cys	Gln	Lys	Arg	395	His	Val	Cys	Ile	Gln	400	Ser	Ser	Leu	Ala	Lys	Val	405
Ser	Ser	Lys	Lys	410	Val	Asn	Glu	Lys	Asp	415	Val	Asp	Lys	Phe	Leu	Leu	420
Tyr	Gln	His	Phe	425	Ser	Cys	Asn	Ile	Arg	430	Asn	Ile	His	His	His	Gln	435
Ile	Leu	Ala	Ile	440	Asn	Arg	Gly	Glu	Asn	445	Leu	Lys	Val	Leu	Thr	Val	450
Lys	Val	Asn	Ile	455	Ser	Asp	Gly	Val	Lys	460	Asp	Glu	Phe	Cys	Arg	Trp	465
Cys	Ile	Gln	Asn	470	Arg	Trp	Arg	Pro	Arg	475	Ser	Phe	Ala	Arg	Pro	Glu	480
Leu	Met	Lys	Ile	485	Leu	Tyr	Asn	Ser	Leu	490	Asn	Asp	Ser	Phe	Lys	Arg	495
Leu	Ile	Tyr	Pro	500	Leu	Leu	Cys	Arg	Glu	505	Phe	Arg	Ala	Lys	Leu	Thr	510
Ser	Asp	Ala	Glu	515	Lys	Glu	Ser	Val	Met	520	Met	Phe	Gly	Arg	Asn	Leu	525
Arg	Gln	Leu	Leu	530	Leu	Thr	Ser	Pro	Val	535	Pro	Gly	Arg	Thr	Leu	Met	540
Gly	Val	Asp	Pro	545	Gly	Tyr	Lys	His	Gly	550	Cys	Lys	Leu	Ala	Ile	Ile	555
Ser	Pro	Thr	Ser	560	Gln	Ile	Leu	His	Thr	565	Asp	Val	Val	Tyr	Leu	His	570
Cys	Gly	Gln	Gly	575	Phe	Arg	Glu	Ala	Glu	580	Lys	Ile	Lys	Thr	Leu	Leu	585
Leu	Asn	Phe	Asn	590	Cys	Ser	Thr	Val	Val	595	Ile	Gly	Asn	Gly	Thr	Ala	600
Cys	Arg	Glu	Thr	605	Glu	Ala	Tyr	Phe	Ala	610	Asp	Leu	Ile	Met	Lys	Asn	615
Tyr	Phe	Ala	Pro	620	Leu	Asp	Val	Val	Tyr	625	Cys	Ile	Val	Ser	Glu	Ala	630
Gly	Ala	Ser	Ile	635	Tyr	Ser	Val	Ser	Pro	640	Glu	Ala	Asn	Lys	Glu	Met	645
Pro	Gly	Leu	Asp	650	Pro	Asn	Leu	Arg	Ser	655	Ala	Val	Ser	Ile	Ala	Arg	660
Arg	Val	Gln	Asp	665	Pro	Leu	Ala	Glu	Leu	670	Val	Lys	Ile	Glu	Pro	Lys	675
His	Ile	Gly	Val	680	Gly	Met	Tyr	Gln	His	685	Asp	Val	Ser	Gln	Thr	Leu	690
Leu	Lys	Ala	Thr	695	Leu	Asp	Ser	Val	Val	700	Glu	Glu	Cys	Val	Ser	Phe	705
Val	Gly	Val	Asp	710	Ile	Asn	Ile	Cys	Ser	715	Glu	Val	Leu	Leu	Arg	His	720
Ile	Ala	Gly	Leu	725	Asn	Ala	Asn	Arg	Ala	730	Lys	Asn	Ile	Ile	Glu	Trp	735

10

20

30

40

Arg Glu Lys Asn Gly Pro Phe Ile Asn Arg Glu Gln Leu Lys Lys
 740 745 750
 Val Lys Gly Leu Gly Pro Lys Ser Phe Gln Gln Cys Ala Gly Phe
 755 760 765
 Ile Arg Ile Asn Gln Asp Tyr Ile Arg Thr Phe Cys Ser Gln Gln
 770 775 780
 Thr Glu Thr Ser Gly Gln Ile Gln Gly Val Ala Val Thr Ser Ser
 785 790 795
 Ala Asp Val Glu Val Thr Asn Glu Lys Gln Gly Lys Lys Lys Ser
 800 805 810
 Lys Thr Ala Val Asn Val Leu Leu Lys Pro Asn Pro Leu Asp Gln
 815 820 825
 Thr Cys Ile His Pro Glu Ser Tyr Asp Ile Ala Met Arg Phe Leu
 830 835 840
 Ser Ser Ile Gly Gly Thr Leu Tyr Glu Val Gly Lys Pro Glu Met
 845 850 855
 Gln Gln Lys Ile Asn Ser Phe Leu Glu Lys Glu Gly Met Glu Lys
 860 865 870
 Ile Ala Glu Arg Leu Gln Thr Thr Val His Thr Leu Gln Val Ile
 875 880 885
 Ile Asp Gly Leu Ser Gln Pro Glu Ser Phe Asp Phe Arg Thr Asp
 890 895 900
 Phe Asp Lys Pro Asp Phe Lys Arg Ser Ile Val Cys Leu Glu Asp
 905 910 915
 Leu Gln Ile Gly Thr Val Leu Thr Gly Lys Val Glu Asn Ala Thr
 920 925 930
 Leu Phe Gly Ile Phe Val Asp Ile Gly Val Gly Lys Ser Gly Leu
 935 940 945
 Ile Pro Ile Arg Asn Val Thr Glu Ala Lys Leu Ser Lys Thr Lys
 950 955 960
 Lys Arg Arg Ser Leu Gly Leu Gly Pro Gly Glu Arg Val Glu Val
 965 970 975
 Gln Val Leu Asn Ile Asp Ile Pro Arg Ser Arg Ile Thr Leu Asp
 980 985 990
 Leu Ile Arg Val Leu
 995

10

20

<210> 3
 <211> 252
 <212> PRT
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 1322761CD1

<400> 3
 Met Phe Gln Pro Ala Pro Lys Arg Cys Phe Thr Ile Glu Ser Leu
 1 5 10 15
 Val Ala Lys Asp Ser Pro Leu Pro Ala Ser Arg Ser Glu Asp Pro
 20 25 30
 Ile Arg Pro Ala Ala Leu Ser Tyr Ala Asn Ser Ser Pro Ile Asn
 35 40 45
 Pro Phe Leu Asn Gly Phe His Ser Ala Ala Ala Ala Ala Gly
 50 55 60
 Arg Gly Val Tyr Ser Asn Pro Asp Leu Val Phe Ala Glu Ala Val
 65 70 75
 Ser His Pro Pro Asn Pro Ala Val Pro Val His Pro Val Pro Pro
 80 85 90
 Pro His Ala Leu Ala Ala His Pro Leu Pro Ser Ser His Ser Pro
 95 100 105
 His Pro Leu Phe Ala Ser Gln Gln Arg Asp Pro Ser Thr Phe Tyr
 110 115 120

30

40

Pro Trp Leu Ile His Arg Tyr Arg Tyr Leu Gly His Arg Phe Gln
 125 130 135
 Gly Asn Asp Thr Ser Pro Glu Ser Phe Leu Leu His Asn Ala Leu
 140 145 150
 Ala Arg Lys Pro Lys Arg Ile Arg Thr Ala Phe Ser Pro Ser Gln
 155 160 165
 Leu Leu Arg Leu Glu His Ala Phe Glu Lys Asn His Tyr Val Val
 170 175 180
 Gly Ala Glu Arg Lys Gln Leu Ala His Ser Leu Ser Leu Thr Glu
 185 190 195
 Thr Gln Val Lys Val Trp Phe Gln Asn Arg Arg Thr Lys Phe Lys
 200 205 210
 Arg Gln Lys Leu Glu Glu Gly Ser Asp Ser Gln Gln Lys Lys
 215 220 225
 Lys Gly Thr His His Ile Asn Arg Trp Arg Ile Ala Thr Lys Gln
 230 235 240
 Ala Ser Pro Glu Glu Ile Asp Val Thr Ser Asp Asp
 245 250

10

<210> 4
 <211> 602
 <212> PRT
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 7472664CD1

<400> 4
 Met Ser Phe Arg Pro Arg Ile Leu Ile Val Gly Glu Pro Gly Phe
 1 5 10 15
 Gly Gln Gly Ser His Leu Ala Pro Ala Val Ile His Ala Leu Glu
 20 25 30
 Lys Phe Thr Val Tyr Thr Leu Asp Ile Pro Val Leu Phe Gly Val
 35 40 45
 Ser Thr Thr Ser Pro Glu Glu Thr Cys Ala Gln Val Ile Arg Glu
 50 55 60
 Ala Lys Arg Thr Ala Pro Ser Ile Val Tyr Val Pro His Ile His
 65 70 75
 Val Trp Trp Glu Ile Val Gly Pro Thr Leu Lys Ala Thr Phe Thr
 80 85 90
 Thr Leu Leu Gln Asn Ile Pro Ser Phe Ala Pro Val Leu Leu Leu
 95 100 105
 Ala Thr Ser Asp Lys Pro His Ser Ala Leu Pro Glu Glu Val Gln
 110 115 120
 Glu Leu Phe Ile Arg Asp Tyr Gly Glu Ile Phe Asn Val Gln Leu
 125 130 135
 Pro Asp Lys Glu Glu Arg Thr Lys Phe Phe Glu Asp Leu Ile Leu
 140 145 150
 Lys Gln Ala Ala Lys Pro Pro Ile Ser Lys Lys Lys Ala Val Leu
 155 160 165
 Gln Ala Leu Glu Val Leu Pro Val Ala Pro Pro Pro Glu Pro Arg
 170 175 180
 Ser Leu Thr Ala Glu Glu Val Lys Arg Leu Glu Glu Gln Glu Glu
 185 190 195
 Asp Thr Phe Arg Glu Leu Arg Ile Phe Leu Arg Asn Val Thr His
 200 205 210
 Arg Leu Ala Ile Asp Lys Arg Phe Arg Val Phe Thr Lys Pro Val
 215 220 225
 Asp Pro Asp Glu Val Pro Asp Tyr Val Thr Val Ile Lys Gln Pro
 230 235 240
 Met Asp Leu Ser Ser Val Ile Ser Lys Ile Asp Leu His Lys Tyr
 245 250 255

20

30

40

```

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

```

10

20

30

```

<210> 5
<211> 280
<212> PRT
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Incyte ID No: 7473124CD1

```

```

<400> 5
Met Pro Ile Tyr Ser Gln Thr Val Ala Met Ala Glu His Phe Lys
1 5 10 15
Gln Ala Ser Ser Cys Pro Ile Cys Leu Asp Tyr Leu Glu Asn Pro
20 25 30

```

40

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

10

20

<210> 6
 <211> 221
 <212> PRT
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 7473171CD1

<400> 6
 Met Ala Ala Val Gly Pro Arg Thr Gly Pro Gly Thr Gly Ala Glu
 1 5 10 15
 Ala Leu Ala Leu Ala Ala Glu Leu Gln Gly Glu Ala Thr Cys Ser
 20 25 30
 Ile Cys Leu Glu Leu Phe Arg Glu Pro Val Ser Val Glu Cys Gly
 35 40 45
 His Ser Phe Cys Arg Ala Cys Ile Gly Arg Cys Trp Glu Arg Pro
 50 55 60
 Gly Ala Gly Ser Val Gly Ala Ala Thr Arg Ala Pro Pro Phe Pro
 65 70 75
 Leu Pro Cys Pro Gln Cys Arg Glu Pro Ala Arg Pro Ser Gln Leu
 80 85 90
 Arg Pro Asn Arg Gln Leu Ala Ala Val Ala Thr Leu Leu Arg Arg
 95 100 105
 Phe Ser Leu Pro Ala Ala Ala Pro Gly Glu His Gly Ser Gln Ala
 110 115 120
 Ala Ala Ala Arg Ala Ala Ala Ala Arg Cys Gly Gln His Gly Glu
 125 130 135

30

40

Pro Phe Lys Leu Tyr Cys Gln Asp Asp Gly Arg Ala Ile Cys Val
 140 145 150
 Val Cys Asp Arg Ala Arg Glu His Arg Glu His Ala Val Leu Pro
 155 160 165
 Leu Asp Glu Ala Val Gln Glu Ala Lys Glu Leu Leu Glu Ser Arg
 170 175 180
 Leu Arg Val Leu Lys Lys Glu Leu Glu Asp Cys Glu Val Phe Arg
 185 190 195
 Ser Thr Glu Lys Lys Glu Ser Lys Glu Leu Leu Val Ser Gln Ala
 200 205 210
 Pro Ala Gly Pro Pro Trp Asp Ile Thr Glu Ala
 215 220

<210> 7
 <211> 668
 <212> PRT
 <213> Homo sapiens

10

<220>
 <221> misc_feature
 <223> Incyte ID No: 7477026CD1

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

20

30

40

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

10

20

30

<210> 8
 <211> 519
 <212> PRT
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 6428773CD1

<400> 8
 Met Asp Ala Val Leu Glu Pro Phe Pro Ala Asp Arg Leu Phe Pro
 1 5 10 15

40

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

10

20

30

40

				350						355				360
Asp	Asn	Pro	Arg	Val	Thr	Ile	Val	Thr	His	Gln	Asp	Val	Phe	Arg
				365					370					375
Asn	Leu	Ser	His	Leu	Pro	Thr	Phe	Ser	Ser	Pro	Ala	Ile	Glu	Ser
				380					385					390
His	Ile	His	Arg	Ile	Glu	Gly	Leu	Ser	Gln	Lys	Phe	Ile	Tyr	Leu
				395					400					405
Asn	Asp	Asp	Val	Met	Phe	Gly	Lys	Asp	Val	Trp	Pro	Asp	Asp	Phe
				410					415					420
Tyr	Ser	His	Ser	Lys	Gly	Gln	Lys	Val	Tyr	Leu	Thr	Trp	Pro	Val
				425					430					435
Pro	Asn	Cys	Ala	Glu	Gly	Cys	Pro	Gly	Ser	Trp	Ile	Lys	Asp	Gly
				440					445					450
Tyr	Cys	Asp	Lys	Ala	Cys	Asn	Asn	Ser	Ala	Cys	Asp	Trp	Asp	Gly
				455					460					465
Gly	Asp	Cys	Ser	Gly	Asn	Ser	Gly	Gly	Ser	Arg	Tyr	Ile	Ala	Gly
				470					475					480
Gly	Gly	Gly	Thr	Gly	Ser	Ile	Gly	Val	Gly	Gln	Pro	Trp	Gln	Phe
				485					490					495
Gly	Gly	Gly	Ile	Asn	Ser	Val	Ser	Tyr	Cys	Asn	Gln	Gly	Cys	Ala
				500					505					510
Asn	Ser	Trp	Leu	Ala	Asp	Lys	Phe	Cys	Asp	Gln	Ala	Cys	Asn	Val
				515					520					525
Leu	Ser	Cys	Gly	Phe	Asp	Ala	Gly	Asp	Cys	Gly	Gln	Asp	His	Phe
				530					535					540
His	Glu	Leu	Tyr	Lys	Val	Ile	Leu	Leu	Pro	Asn	Gln	Thr	His	Tyr
				545					550					555
Ile	Ile	Pro	Lys	Gly	Glu	Cys	Leu	Pro	Tyr	Phe	Ser	Phe	Ala	Glu
				560					565					570
Val	Ala	Lys	Arg	Gly	Val	Glu	Gly	Ala	Tyr	Ser	Asp	Asn	Pro	Ile
				575					580					585
Ile	Arg	His	Ala	Ser	Ile	Ala	Asn	Lys	Trp	Lys	Thr	Ile	His	Leu
				590					595					600
Ile	Met	His	Ser	Gly	Met	Asn	Ala	Thr	Thr	Ile	His	Phe	Asn	Leu
				605					610					615
Thr	Phe	Gln	Asn	Thr	Asn	Asp	Glu	Glu	Phe	Lys	Met	Gln	Ile	Thr
				620					625					630
Val	Glu	Val	Asp	Thr	Arg	Glu	Gly	Pro	Lys	Leu	Asn	Ser	Thr	Ala
				635					640					645
Gln	Lys	Gly	Tyr	Glu	Asn	Leu	Val	Ser	Pro	Ile	Thr	Leu	Leu	Pro
				650					655					660
Glu	Ala	Glu	Ile	Leu	Phe	Glu	Asp	Ile	Pro	Lys	Glu	Lys	Arg	Phe
				665					670					675
Pro	Lys	Phe	Lys	Arg	His	Asp	Val	Asn	Ser	Thr	Arg	Arg	Ala	Gln
				680					685					690
Glu	Glu	Val	Lys	Ile	Pro	Leu	Val	Asn	Ile	Ser	Leu	Leu	Pro	Lys
				695					700					705
Asp	Ala	Gln	Leu	Ser	Leu	Asn	Thr	Leu	Asp	Leu	Gln	Leu	Glu	His
				710					715					720
Gly	Asp	Ile	Thr	Leu	Lys	Gly	Tyr	Asn	Leu	Ser	Lys	Ser	Ala	Leu
				725					730					735
Leu	Arg	Ser	Phe	Leu	Met	Asn	Ser	Gln	His	Ala	Lys	Ile	Lys	Asn
				740					745					750
Gln	Ala	Ile	Ile	Thr	Asp	Glu	Thr	Asn	Asp	Ser	Leu	Val	Ala	Pro
				755					760					765
Gln	Glu	Lys	Gln	Val	His	Lys	Ser	Ile	Leu	Pro	Asn	Ser	Leu	Gly
				770					775					780
Val	Ser	Glu	Arg	Leu	Gln	Arg	Leu	Thr	Phe	Pro	Ala	Val	Ser	Val
				785					790					795
Lys	Val	Asn	Gly	His	Asp	Gln	Gly	Gln	Asn	Pro	Pro	Leu	Asp	Leu
				800					805					810
Glu	Thr	Thr	Ala	Arg	Phe	Arg	Val	Glu	Thr	His	Thr	Gln	Lys	Thr
				815					820					825

10

20

30

40

Ile Gly Gly Asn Val Thr Lys Glu Lys Pro Pro Ser Leu Ile Val
 830 835 840
 Pro Leu Glu Ser Gln Met Thr Lys Glu Lys Lys Ile Thr Gly Lys
 845 850 855
 Glu Lys Glu Asn Ser Arg Met Glu Glu Asn Ala Glu Asn His Ile
 860 865 870
 Gly Val Thr Glu Val Leu Leu Gly Arg Lys Leu Gln His Tyr Thr
 875 880 885
 Asp Ser Tyr Leu Gly Phe Leu Pro Trp Glu Lys Lys Lys Tyr Phe
 890 895 900
 Gln Asp Leu Leu Asp Glu Glu Glu Ser Leu Lys Thr Gln Leu Ala
 905 910 915
 Tyr Phe Thr Asp Ser Lys Asn Thr Gly Arg Gln Leu Lys Asp Thr
 920 925 930
 Phe Ala Asp Ser Leu Arg Tyr Val Asn Lys Ile Leu Asn Ser Lys
 935 940 945
 Phe Gly Phe Thr Ser Arg Lys Val Pro Ala His Met Pro His Met
 950 955 960
 Ile Asp Arg Ile Val Met Gln Glu Leu Gln Asp Met Phe Pro Glu
 965 970 975
 Glu Phe Asp Lys Thr Ser Phe His Lys Val Arg His Ser Glu Asp
 980 985 990
 Met Gln Phe Ala Phe Ser Tyr Phe Tyr Tyr Leu Met Ser Ala Val
 995 1000 1005
 Gln Pro Leu Asn Ile Ser Gln Val Phe Asp Glu Val Asp Thr Asp
 1010 1015 1020
 Gln Ser Gly Val Leu Ser Asp Arg Glu Ile Arg Thr Leu Ala Thr
 1025 1030 1035
 Arg Ile His Glu Leu Pro Leu Ser Leu Gln Asp Leu Thr Gly Leu
 1040 1045 1050
 Glu His Met Leu Ile Asn Cys Ser Lys Met Leu Pro Ala Asp Ile
 1055 1060 1065
 Thr Gln Leu Asn Asn Ile Pro Pro Thr Gln Glu Ser Tyr Tyr Asp
 1070 1075 1080
 Pro Asn Leu Pro Pro Val Thr Lys Ser Leu Val Thr Asn Cys Lys
 1085 1090 1095
 Pro Val Thr Asp Lys Ile His Lys Ala Tyr Lys Asp Lys Asn Lys
 1100 1105 1110
 Tyr Arg Phe Glu Ile Met Gly Glu Glu Glu Ile Ala Phe Lys Met
 1115 1120 1125
 Ile Arg Thr Asn Val Ser His Val Val Gly Gln Leu Asp Asp Ile
 1130 1135 1140
 Arg Lys Asn Pro Arg Lys Phe Val Cys Leu Asn Asp Asn Ile Asp
 1145 1150 1155
 His Asn His Lys Asp Ala Gln Thr Val Lys Ala Val Leu Arg Asp
 1160 1165 1170
 Phe Tyr Glu Ser Met Phe Pro Ile Pro Ser Gln Phe Glu Leu Pro
 1175 1180 1185
 Arg Glu Tyr Arg Asn Arg Phe Leu His Met His Glu Leu Gln Glu
 1190 1195 1200
 Trp Arg Ala Tyr Arg Asp Lys Leu Lys Phe Trp Thr His Cys Val
 1205 1210 1215
 Leu Ala Thr Leu Ile Met Phe Thr Ile Phe Ser Phe Phe Ala Glu
 1220 1225 1230
 Gln Leu Ile Ala Leu Lys Arg Lys Ile Phe Pro Arg Arg Arg Ile
 1235 1240 1245
 His Lys Glu Ala Ser Pro Asn Arg Ile Arg Val
 1250 1255

10

20

30

<210> 10
 <211> 364
 <212> PRT
 <213> Homo sapiens

40

<220>
<221> misc_feature
<223> Incyte ID No: 118539CD1

<400> 10

```

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

```

10

20

30

<210> 11
<211> 331
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature

40

<223> Incyte ID No: 4005918CD1

<400> 11

Met Thr Asp Pro Ser Leu Gly Leu Thr Val Pro Met Ala⁵ Pro Pro
 1 5 10 15
 Leu Ala Pro Leu Pro Pro Arg Asp Pro Asn Gly Ala Gly Ser Glu
 20 25 30
 Trp Arg Lys Pro Gly Ala Val Ser Phe Ala Asp Val Ala Val Tyr
 35 40 45
 Phe Ser Arg Glu Glu Trp Gly Cys Leu Arg Pro Ala Gln Arg Ala
 50 55 60
 Leu Tyr Arg Asp Val Met Arg Glu Thr Tyr Gly His Leu Gly Ala
 65 70 75
 Leu Gly Val Gly Gly Ser Lys Pro Ala Leu Ile Ser Trp Val Glu
 80 85 90
 Glu Lys Ala Glu Leu Trp Asp Pro Ala Ala Gln Asp Pro Glu Val
 95 100 105
 Ala Lys Cys Pro Thr Glu Ala Asp Pro Ala Asp Ser Arg Asn Lys
 110 115 120
 Glu Glu Glu Arg Gln Arg Glu Gly Thr Gly Ala Leu Glu Lys Pro
 125 130 135
 Asp Pro Val Ala Ala Gly Ser Pro Gly Leu Lys Ala Pro Gln Ala
 140 145 150
 Pro Phe Ala Gly Leu Glu Gln Leu Ser Lys Ala Arg Arg Arg Ser
 155 160 165
 Arg Pro Arg Phe Phe Ala His Pro Pro Val Pro Arg Ala Asp Gln
 170 175 180
 Arg His Gly Cys Tyr Val Cys Gly Lys Ser Phe Ala Trp Arg Ser
 185 190 195
 Thr Leu Val Glu His Ile Tyr Ser His Arg Gly Glu Lys Pro Phe
 200 205 210
 His Cys Ala Asp Cys Gly Lys Gly Phe Gly His Ala Ser Ser Leu
 215 220 225
 Ser Lys His Arg Ala Ile His Arg Gly Glu Arg Pro His Arg Cys
 230 235 240
 Pro Glu Cys Gly Arg Ala Phe Met Arg Arg Thr Ala Leu Thr Ser
 245 250 255
 His Leu Arg Val His Thr Gly Glu Lys Pro Tyr Arg Cys Pro Gln
 260 265 270
 Cys Gly Arg Cys Phe Gly Leu Lys Thr Gly Met Ala Lys His Gln
 275 280 285
 Trp Val His Arg Pro Gly Gly Glu Gly Arg Arg Gly Arg Arg Pro
 290 295 300
 Gly Gly Leu Ser Val Thr Leu Thr Pro Val Arg Gly Asp Leu Asp
 305 310 315
 Pro Pro Val Gly Phe Gln Leu Tyr Pro Glu Ile Phe Gln Glu Cys
 320 325 330
 Gly

10

20

30

<210> 12

<211> 670

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 5435937CD1

<400> 12

Met Asp Ser Val Ala Phe Glu Asp Val Ala Val Thr Phe Thr Gln
 1 5 10 15
 Glu Glu Trp Ala Leu Leu Asp Pro Ser Gln Lys Asn Leu Cys Arg

40

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

10

20

30

40

Thr His Thr Gly Gly Lys Thr Tyr Glu Cys Lys Gln Cys Gly Arg
 500 505
 Ser Phe Asn Cys Ser Ser Ser Phe Arg Tyr His Gly Arg Thr His
 515 520
 Thr Gly Glu Lys Pro Tyr Glu Cys Lys Gln Cys Gly Lys Ala Phe
 530 535
 Arg Ser Ala Ser Gln Leu Gln Ile His Gly Arg Thr His Thr Gly
 545 550
 Glu Lys Pro Tyr Glu Cys Lys Gln Cys Gly Lys Ala Phe Gly Ser
 560 565
 Ala Ser His Leu Gln Met His Gly Arg Thr His Thr Gly Glu Lys
 575 580
 Pro Tyr Glu Cys Lys Gln Cys Gly Lys Ser Phe Gly Cys Ala Ser
 590 595
 Arg Leu Gln Met His Gly Arg Thr His Thr Gly Glu Lys Pro Tyr
 605 610
 Lys Cys Lys Gln Cys Gly Lys Ala Phe Gly Cys Pro Ser Asn Leu
 620 625
 Arg Arg His Gly Arg Thr His Thr Gly Glu Lys Pro Tyr Lys Cys
 635 640
 Asn Gln Cys Gly Lys Val Phe Arg Cys Ser Ser Gln Leu Gln Val
 650 655
 His Gly Arg Ala His Cys Ile Asp Thr Pro
 665 670

<210> 13
 <211> 1196
 <212> PRT
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 7503560CD1

<400> 13
 Met Leu Phe Lys Leu Leu Gln Arg Gln Thr Tyr Thr Cys Leu Ser
 1 5 10 15
 His Arg Tyr Gly Leu Tyr Val Cys Phe Leu Gly Val Val Val Thr
 20 25 30
 Ile Val Ser Ala Phe Gln Phe Gly Glu Val Val Leu Glu Trp Ser
 35 40 45
 Arg Asp Gln Tyr His Val Leu Phe Asp Ser Tyr Arg Asp Asn Ile
 50 55 60
 Ala Gly Lys Ser Phe Gln Asn Arg Leu Cys Leu Pro Met Pro Ile
 65 70 75
 Asp Val Val Tyr Thr Trp Val Asn Gly Thr Asp Leu Glu Leu Leu
 80 85 90
 Lys Glu Leu Gln Gln Val Arg Glu Gln Met Glu Glu Glu Gln Lys
 95 100 105
 Ala Met Arg Glu Ile Leu Gly Lys Asn Thr Thr Glu Pro Thr Lys
 110 115 120
 Lys Ser Glu Lys Gln Leu Glu Cys Leu Leu Thr His Cys Ile Lys
 125 130 135
 Val Pro Met Leu Val Leu Asp Pro Ala Leu Pro Ala Asn Ile Thr
 140 145 150
 Leu Lys Asp Leu Pro Ser Leu Tyr Pro Ser Phe His Ser Ala Ser
 155 160 165
 Asp Ile Phe Asn Val Ala Lys Pro Lys Asn Pro Ser Thr Asn Val
 170 175 180
 Ser Val Val Val Phe Asp Ser Thr Lys Asp Val Glu Asp Ala His
 185 190 195
 Ser Gly Leu Leu Lys Gly Asn Ser Arg Gln Thr Val Trp Arg Gly
 200 205 210

10

20

30

40

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

10

20

30

40

Gln	Ala	Ile	Ile	Thr	Asp	Glu	Thr	Asn	Asp	Ser	Leu	Val	Ala	Pro	680	685	690
															695	700	705
Gln	Glu	Lys	Gln	Val	His	Lys	Ser	Ile	Leu	Pro	Asn	Ser	Leu	Gly	710	715	720
Val	Ser	Glu	Arg	Leu	Gln	Arg	Leu	Thr	Phe	Pro	Ala	Val	Ser	Val	725	730	735
Lys	Val	Asn	Gly	His	Asp	Gln	Gly	Gln	Asn	Pro	Pro	Leu	Asp	Leu	740	745	750
Glu	Thr	Thr	Ala	Arg	Phe	Arg	Val	Glu	Thr	His	Thr	Gln	Lys	Thr	755	760	765
Ile	Gly	Gly	Asn	Val	Thr	Lys	Glu	Lys	Pro	Pro	Ser	Leu	Ile	Val	770	775	780
Pro	Leu	Glu	Ser	Gln	Met	Thr	Lys	Glu	Lys	Lys	Ile	Thr	Gly	Lys	785	790	795
Glu	Lys	Glu	Asn	Ser	Arg	Met	Glu	Glu	Asn	Ala	Glu	Asn	His	Ile	800	805	810
Gly	Val	Thr	Glu	Val	Leu	Leu	Gly	Arg	Lys	Leu	Gln	His	Tyr	Thr	815	820	825
Asp	Ser	Tyr	Leu	Gly	Phe	Leu	Pro	Trp	Glu	Lys	Lys	Lys	Tyr	Phe	830	835	840
Gln	Asp	Leu	Leu	Asp	Glu	Glu	Glu	Ser	Leu	Lys	Thr	Gln	Leu	Ala	845	850	855
Tyr	Phe	Thr	Asp	Ser	Lys	Asn	Thr	Gly	Arg	Gln	Leu	Lys	Asp	Thr	860	865	870
Phe	Ala	Asp	Ser	Leu	Arg	Tyr	Val	Asn	Lys	Ile	Leu	Asn	Ser	Lys	875	880	885
Phe	Gly	Phe	Thr	Ser	Arg	Lys	Val	Pro	Ala	His	Met	Pro	His	Met	890	895	900
Ile	Asp	Arg	Ile	Val	Met	Gln	Glu	Leu	Gln	Asp	Met	Phe	Pro	Glu	905	910	915
Glu	Phe	Asp	Lys	Thr	Ser	Phe	His	Lys	Val	Arg	His	Ser	Glu	Asp	920	925	930
Met	Gln	Phe	Ala	Phe	Ser	Tyr	Phe	Tyr	Tyr	Leu	Met	Ser	Ala	Val	935	940	945
Gln	Pro	Leu	Asn	Ile	Ser	Gln	Val	Phe	Asp	Glu	Val	Asp	Thr	Asp	950	955	960
Gln	Ser	Gly	Val	Leu	Ser	Asp	Arg	Glu	Ile	Arg	Thr	Leu	Ala	Thr	965	970	975
Arg	Ile	His	Glu	Leu	Pro	Leu	Ser	Leu	Gln	Asp	Leu	Thr	Gly	Leu	980	985	990
Glu	His	Met	Leu	Ile	Asn	Cys	Ser	Lys	Met	Leu	Pro	Ala	Asp	Ile	995	1000	1005
Thr	Gln	Leu	Asn	Asn	Ile	Pro	Pro	Thr	Gln	Glu	Ser	Tyr	Tyr	Asp	1010	1015	1020
Pro	Asn	Leu	Pro	Pro	Val	Thr	Lys	Ser	Leu	Val	Thr	Asn	Cys	Lys	1025	1030	1035
Pro	Val	Thr	Asp	Lys	Ile	His	Lys	Ala	Tyr	Lys	Asp	Lys	Asn	Lys	1040	1045	1050
Tyr	Arg	Phe	Glu	Ile	Met	Gly	Glu	Glu	Glu	Ile	Ala	Phe	Lys	Met	1055	1060	1065
Ile	Arg	Thr	Asn	Val	Ser	His	Val	Val	Gly	Gln	Leu	Asp	Asp	Ile	1070	1075	1080
Arg	Lys	Asn	Pro	Arg	Lys	Phe	Val	Cys	Leu	Asn	Asp	Asn	Ile	Asp	1085	1090	1095
His	Asn	His	Lys	Asp	Ala	Gln	Thr	Val	Lys	Ala	Val	Leu	Arg	Asp	1100	1105	1110
Phe	Tyr	Glu	Ser	Met	Phe	Pro	Ile	Pro	Ser	Gln	Phe	Glu	Leu	Pro	1115	1120	1125
Arg	Glu	Tyr	Arg	Asn	Arg	Phe	Leu	His	Met	His	Glu	Leu	Gln	Glu	1130	1135	1140
Trp	Arg	Ala	Tyr	Arg	Asp	Lys	Leu	Lys	Phe	Trp	Thr	His	Cys	Val	1145	1150	1155

10

20

30

40

Leu Ala Thr Leu Ile Met Phe Thr Ile Phe Ser Phe Phe Ala Glu
 1160 1165 1170
 Gln Leu Ile Ala Leu Lys Arg Lys Ile Phe Pro Arg Arg Arg Ile
 1175 1180 1185
 His Lys Glu Ala Ser Pro Asn Arg Ile Arg Val
 1190 1195

<210> 14
 <211> 2138
 <212> PRT
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 7472378CD1

10

<400> 14
 Met Ser Lys Lys Phe Val Tyr Asn Leu Arg Lys Thr Thr Arg Ser
 1 5 10 15
 Val Val Gly Val Pro Pro Asn Thr Asn Arg Pro Pro His Pro Val
 20 25 30
 Arg Arg Pro Asp Ser Leu Leu Pro Ile Ser Glu Glu Pro Lys Ser
 35 40 45
 Ile Ser Ser Gln Thr Pro Asn Met Asp Ser Gly Asn Asp Ser Ala
 50 55 60
 Arg Pro Thr Pro Ser Pro Leu Ala Pro Thr Val Ser Gly Ile Ser
 65 70 75
 Ser Leu Ile Ser Thr Thr Phe Lys Pro Lys Asp Ile Met Ala Phe
 80 85 90
 Val Glu His Leu Pro Thr Phe Asp Gly Thr Pro Arg Leu Leu Asp
 95 100 105
 Arg Phe Ile Thr Ser Val Glu Glu Ile Leu Met Leu Ile Arg Gly
 110 115 120
 Ala Asp Gln Thr Pro Tyr Gly Leu Leu Thr Leu Arg Thr Ile Arg
 125 130 135
 Asn Lys Ile Ile Asp Arg Ala Asp Glu Ala Leu Glu Leu Ala Asn
 140 145 150
 Thr Pro Leu Val Trp Asp Glu Ile Lys Ser Asn Leu Ile Arg Leu
 155 160 165
 Tyr Ser Ser Lys Lys Ser Glu Ala Asn Leu Leu Ser Glu Leu Asn
 170 175 180
 Thr Phe Ser Asp Asn Leu Thr Leu Gly Gln Leu Phe Phe Gly Ile
 185 190 195
 Ser Lys Val Arg Ser Gln Leu Phe Ser Ile Leu Lys Asn Ser Glu
 200 205 210
 His Asn Asn Thr Val Val Asp Ala Lys Lys Val Val Tyr Asn Glu
 215 220 225
 Val Cys Leu Asn Ala Phe Met Thr Gly Leu Lys Glu Pro Leu Lys
 230 235 240
 Thr Phe Val Arg Ile Lys Ser Pro Ser Thr Leu Glu Gln Ala Tyr
 245 250 255
 Glu Gln Cys Gln Ile Glu Gln Thr Leu Tyr Arg Ala Gln Asn Lys
 260 265 270
 Arg Thr Asn Arg Pro Glu Gln Gly Pro Asn Gly Ser Asp Asn Lys
 275 280 285
 Thr Tyr Arg Asn Ser Tyr Asp Ser Asn Tyr Arg Ser Gly Arg Asn
 290 295 300
 Asp Arg Asn Asp Arg Arg Gly Pro Tyr Ser Asn Ser Asn Ser Asn
 305 310 315
 Ser Asn Ser Gly Gln Asn Arg Pro Phe Asn Ser His Asn Arg Thr
 320 325 330
 Pro Gln Ser Gly Thr Lys Asp Asn Arg Ala Asn Thr Ser Asn Pro
 335 340 345

20

30

40

Phe Arg Ala Pro Ser His Ser Leu Asn Asn Ile Glu Glu Asn Pro
 350 355 360
 Gln Pro Asp Ser Asn Phe Gln Gln Thr Ala Ser Gly Asn Gln Gln
 365 370 375
 Gly Ser Thr His Ser Phe Ile Asp Pro Lys Tyr Val Asp Pro Arg
 380 385 390
 Asn Cys Val Thr Leu Asp Thr Pro Ile Thr Leu Lys Thr Ala Leu
 395 400 405
 Asn Ser Phe Lys Ile Tyr Gln Asn Val Ser Ile Pro Phe Pro Pro
 410 415 420
 Glu Phe Gln Ile Thr Gly Lys Met Thr Leu Leu Pro Phe Lys Phe
 425 430 435
 His Ser Tyr Phe Asp Gly Leu Ile Gly Met Asp Leu Leu Ser Tyr
 440 445 450
 Leu Lys Thr Glu Ile Asp Leu Leu Asn Leu Asn Leu Lys Thr Pro
 455 460 465
 Ser Thr Ile Ile Pro Leu Trp Thr His Ser Asn Ser Thr Ser Asn
 470 475 480
 Val Phe Asn Ile Ser Gly His Tyr Lys Thr Ile Leu Pro Leu Pro
 485 490 495
 Val Glu Thr Lys Gln Gly Asp Phe Tyr Ile Asp Ser Ile Thr Ile
 500 505 510
 Asn Asp Asp Leu Ile Ile Ser Asp Gly Ile Tyr Asn Ala Gln Asn
 515 520 525
 Asn Ile Ala Asn Phe Val Ile Thr Asn Tyr Ser Glu Arg Asp Gln
 530 535 540
 Leu Leu Tyr Leu Glu Ser Pro Ile Lys Gly Met Pro Tyr Ser Thr
 545 550 555
 Ala Asn Asn Val Glu Leu Phe Ser Ile Thr Ser Asp Thr Pro Gln
 560 565 570
 Pro Gln Asn Ser Ala Ala Ser Leu Gln Ala Leu Gly Val Asp His
 575 580 585
 Leu Ser Ser Glu Glu Lys Gln Ser Leu Leu Ser Leu Cys Lys Ser
 590 595 600
 Tyr Leu Asp Ile Phe Tyr Asn Glu Asp Lys Ser Leu Thr Phe Thr
 605 610 615
 Asn Lys Ile Thr His Thr Ile Lys Thr Thr Asp Asp Thr Pro Ile
 620 625 630
 His Thr Lys Ser Tyr Arg Tyr Pro Tyr Ile His Lys Glu Glu Val
 635 640 645
 Lys Lys Gln Ile Glu Ala Met Leu Asn Gln Asp Ile Ile Lys Ser
 650 655 660
 Ser Tyr Ser Pro Trp Ser Ala Pro Val Trp Val Val Pro Lys Lys
 665 670 675
 Ile Thr Pro Thr Gly Glu Gln Lys Trp Arg Leu Val Ile Asp Tyr
 680 685 690
 Arg Lys Leu Asn Glu Lys Thr Ile Ser Asp Arg Tyr Pro Ile Pro
 695 700 705
 Asn Ile Ala Asp Ile Leu Asp Arg Leu Gly Lys Ala Lys Tyr Phe
 710 715 720
 Ser Thr Leu Asp Leu Ala Ser Gly Phe His Gln Ile Glu Met Asn
 725 730 735
 Pro Asp Asp Thr Pro Lys Thr Ala Phe Thr Val Glu Gly Gly His
 740 745 750
 Tyr Glu Phe Ile Arg Met Pro Phe Gly Leu Lys Asn Ala Pro Ala
 755 760 765
 Thr Phe Gln Arg Val Met Asp Asn Ile Phe Gly Asp Leu Ile Gly
 770 775 780
 Thr Ile Cys Leu Val Tyr Leu Asp Asp Ile Ile Ile Phe Ser Thr
 785 790 795
 Ser Leu Gln Glu His Phe Ile His Leu Lys Thr Ile Phe Gly Arg
 800 805 810
 Leu Arg Ser Ala Asn Phe Lys Val Gln Leu Thr Lys Ser Tyr Phe

10

20

30

40

Leu Arg Arg Glu Thr	815	820	825
Glu Phe Leu Gly His		Ile Val Ser Gln Glu	
830	835	840	
Gly Val Arg Pro Asn	845	850	855
Pro Asn Lys Ile Glu		Ala Ile Lys Asn Phe	
860	865	870	
Pro Cys Pro His Ser		Lys Lys Ser Ile Lys	
875	880	885	
Leu Gly Tyr Tyr Arg		Lys Phe Ile Arg Asp	
890	895	900	
Gln Pro Met Thr Gln		Lys Leu Arg Gly Asn	
905	910	915	
Ile Asp Asp Glu Phe		Lys Lys Ala Phe Glu	
920	925	930	
Leu Ser Asn Asp Pro		Ile Leu Gln Tyr Pro	
935	940	945	
Phe Thr Leu Thr Thr		Asp Ala Ser Asn Phe	
950	955	960	
Leu Ser Gln Gly Pro		Val His Ser Asp Arg	
965	970	975	
Ser Arg Thr Leu Ser		Ala Ala Glu Thr Asn	
980	985	990	
Lys Glu Met Leu Ala		Ile Ile Trp Ala Val	
995	1000	1005	
Tyr Leu Phe Gly Arg		Arg Phe Thr Ile Ile	
1010	1015	1020	
Leu Thr Trp Leu Met		Asn Phe Lys Gln Pro	
1025	1030	1035	
Arg Trp Arg Leu Gln		Leu Gln Glu Tyr Asp	
1040	1045	1050	
Lys Lys Gly Ser Gln		Asn Val Ile Ala Asp	
1055	1060	1065	
Glu Ala Ser Val Asn		His Asn Glu Ala Leu	
1070	1075	1080	
Val Cys Pro Ile Ser		Glu Lys Pro Leu Asn	
1085	1090	1095	
Leu Leu Phe Lys Ile		Thr Pro Asp Thr Asn	
1100	1105	1110	
Pro Phe Lys His Lys		Leu Arg Arg Glu Phe	
1115	1120	1125	
Gln Tyr Asp Asp Val		Val Cys Ile Leu Arg	
1130	1135	1140	
Asn Lys Thr Cys Ala		Val Phe Ala Pro Asp	
1145	1150	1155	
Val Glu Gln Ala Tyr		Gln Thr Tyr Phe Ser	
1160	1165	1170	
Lys Leu Ile Arg Cys		Leu Ile Phe Leu Pro	
1175	1180	1185	
Thr Glu Ile Glu Lys		Ile Ile Thr Asp Tyr	
1190	1195	1200	
His Arg Gly Ile Asp		Glu Thr Tyr Leu His	
1205	1210	1215	
Phe Phe Pro His Met		Lys Glu Arg Ile Thr	
1220	1225	1230	
Cys Glu Thr Cys Leu		Lys Leu Lys Tyr Asp	
1235	1240	1245	
Ile Thr Tyr Gln Ile		Ser Glu Leu Pro Ser	
1250	1255	1260	
Leu His Ile Asp Ile		Tyr Thr Ile Asn Lys	
1265	1270	1275	
Ile Ile Asp Lys Phe		Ser Lys Phe Ala Ala	
1280	1285	1290	
Asn Arg Asn Cys Ile		Asn Val Val Lys Ala	
		Leu Lys His Phe Ile	

10

20

30

40

Ser Gln Phe Gly Ile Pro Lys Lys Leu Ile Tyr Asp Gln Gly Ala
 1295 1300 1305
 Glu Phe Ala Ser Asp Met Phe Asn Lys Phe Cys Thr Gln Phe Asn
 1310 1315 1320
 Ile Asp Leu His Val Thr Ser Phe Gln Gln Ser Ser Ser Asn Ser
 1325 1330 1335
 Pro Val Glu Arg Leu His Ser Thr Leu Thr Glu Ile Tyr Arg Ile
 1340 1345 1350
 Ile Leu Asp Val Arg Lys Gln Gln Lys Leu Ser Ser Glu His Asp
 1355 1360 1365
 Glu Ile Met Ser Glu Thr Leu Ile Thr Tyr Asn Asn Ala Ile His
 1370 1375 1380
 Ser Ala Thr Lys His Thr Pro Phe Glu Leu Phe Asn Gly Arg Thr
 1385 1390 1395
 His Ile Phe Asn Gln Thr Ile Gln Phe Asn Asn Glu His Asp Tyr
 1400 1405 1410
 Leu Thr Lys Leu Asn Glu Phe Arg Glu Lys Leu Tyr Pro Leu Ile
 1415 1420 1425
 Thr Asp Lys Leu Ser Asn Asp Val Val Arg Arg Thr Leu Lys Leu
 1430 1435 1440
 Asn Glu Thr Arg Thr Asp Pro Val Asp Leu Gln Pro Asp Thr Leu
 1445 1450 1455
 Val Leu Arg Lys Glu Asn Arg Arg Asn Lys Ile Thr Pro Arg Phe
 1460 1465 1470
 Ser Ile His Lys Val Lys His Asp Lys Gly His Thr Leu Ile Thr
 1475 1480 1485
 Ala Arg Asn Gln Lys Leu His Lys Ser Lys Ile Arg Lys Thr Val
 1490 1495 1500
 Leu Lys Lys Asp Lys Ser Asn Asn Ala Ile His Val His Tyr Leu
 1505 1510 1515
 Asn Asp Asn Ala Pro Ile Ala Lys Ile Glu Leu Gly Lys Ala Leu
 1520 1525 1530
 Leu Ile Glu Arg Tyr Lys Ile Ile Ser His Val Ile Asn Leu Gln
 1535 1540 1545
 Asp Tyr Ser Arg Cys Met Glu Gln Phe His Leu Thr Ile Asn Lys
 1550 1555 1560
 Phe Asn Pro Asp Ser Thr Leu Thr Asp Ser Val Thr Ile Leu Lys
 1565 1570 1575
 Thr Lys Leu Thr Gln Ala Gln Val Lys Leu Lys Ala Leu Thr Pro
 1580 1585 1590
 Ser Tyr Arg Asn Lys Arg Gly Leu Ile Asn Gly Leu Gly Ser Leu
 1595 1600 1605
 Val Lys Val Val Thr Gly Asn Met Asp Ala Asn Asp Asn Lys Glu
 1610 1615 1620
 Ile His Glu Glu Leu Asp Asn Ile Lys Lys Asn Ser Glu Val Ser
 1625 1630 1635
 Asn Asp Asn Leu Gln Lys Gln Val Met Phe Asn Asn Glu Ile Leu
 1640 1645 1650
 Ile Arg Phe Glu Asn Ile Thr Asp His Ile Asn Asn Glu Gln Ile
 1655 1660 1665
 Leu Ile Ser Lys Phe Phe Asp Thr Ser Gln Asn Lys Ile Tyr Lys
 1670 1675 1680
 His Leu Asn Leu Gln Asp Thr Leu Leu Glu Glu Ile Gln Tyr Leu
 1685 1690 1695
 Asn Arg Ile Asn Tyr Asn Ile Glu Leu Phe Ile Asn His Leu Asn
 1700 1705 1710
 Asp Ile Thr Glu Ser Met Leu Leu Ala Lys Ile Asn Ile Ile Pro
 1715 1720 1725
 Lys Phe Ile Leu Asn Glu Gln Glu Met Asp Lys Ile Lys Thr Ile
 1730 1735 1740
 Leu Glu Lys Gln Asn Ile Thr Val Lys Asn Glu Gln Ser Ile Tyr
 1745 1750 1755
 Asn Phe Leu Gln Met Asn Thr Leu Asn Tyr Glu Gln Lys Ile Ile

10

20

30

40

	1760		1765		1770									
Phe	Asn	Ile	Lys	Val	Pro	Ile	Phe	Lys	Gln	Pro	Phe	His	Thr	Leu
	1775		1780		1785									
Ala	Arg	Leu	Val	Pro	Leu	Pro	Ile	Asn	Asn	Thr	Tyr	Phe	Val	Ile
	1790		1795		1800									
Thr	Pro	Asn	Tyr	Leu	Ala	Tyr	Asn	Ile	Asn	Asn	Lys	Lys	Phe	His
	1805		1810		1815									
Met	Thr	Arg	Lys	Cys	Pro	Lys	Leu	Asp	Asn	Thr	Phe	Leu	Cys	Asp
	1820		1825		1830									
Glu	Asn	Phe	Tyr	Val	Asp	Thr	Pro	Gln	Asn	Asn	Thr	Cys	Leu	Glu
	1835		1840		1845									
His	Leu	Leu	Asn	Gly	Glu	Asn	Ser	Ser	Cys	Asp	Val	Arg	Glu	Thr
	1850		1855		1860									
Gly	Pro	Ile	Thr	Asp	Val	Phe	Glu	Ala	Glu	Arg	Gly	Tyr	Ile	Phe
	1865		1870		1875									
Ala	Phe	Asn	Val	Asn	Lys	Leu	Lys	Val	Ser	Leu	Thr	Asn	Gly	Ser
	1880		1885		1890									
Glu	Leu	Ser	Ile	Met	Gly	Ser	Ala	Ile	Ile	Arg	Tyr	Ile	Asn	Glu
	1895		1900		1905									
Thr	Ile	Gln	Ile	Asn	Gly	Ile	Asp	Tyr	Asp	Gly	Thr	Val	Asp	Thr
	1910		1915		1920									
Phe	Pro	Glu	Gln	Thr	Asp	Phe	Asp	Leu	Pro	Pro	Met	Arg	Lys	Val
	1925		1930		1935									
Thr	Arg	Asn	Thr	Thr	Ile	Thr	Val	Leu	Ser	Leu	Glu	Lys	Leu	His
	1940		1945		1950									
Leu	Glu	Ala	Thr	Gln	Thr	Met	Asp	Lys	Ile	Leu	Ala	Val	His	His
	1955		1960		1965									
Asn	Thr	Ile	Gln	His	Thr	Trp	Thr	Leu	Tyr	Thr	Leu	Leu	Gly	Leu
	1970		1975		1980									
Val	Thr	Phe	Leu	Ala	Val	Ile	Leu	Trp	Leu	His	Arg	Arg	Thr	Lys
	1985		1990		1995									
His	Ile	Val	His	Ile	His	Glu	Asp	His	His	Val	Leu	Thr	Pro	Ile
	2000		2005		2010									
Ala	Lys	Ser	Lys	Gly	Lys	Val	Leu	Ile	Arg	Asp	Lys	Leu	Met	Glu
	2015		2020		2025									
Arg	Arg	Asn	Arg	Arg	Thr	Glu	Arg	Thr	Glu	Lys	Ala	Arg	Ile	Trp
	2030		2035		2040									
Glu	Val	Thr	Asp	Arg	Thr	Val	Arg	Thr	Trp	Ile	Gly	Glu	Ala	Val
	2045		2050		2055									
Ala	Ala	Ala	Ala	Ala	Asp	Gly	Val	Thr	Phe	Ser	Val	Pro	Val	Thr
	2060		2065		2070									
Pro	His	Thr	Phe	Arg	His	Ser	Tyr	Ala	Met	His	Met	Leu	Tyr	Ala
	2075		2080		2085									
Gly	Ile	Pro	Leu	Lys	Val	Leu	Gln	Ser	Leu	Met	Gly	His	Lys	Ser
	2090		2095		2100									
Ile	Ser	Ser	Thr	Glu	Val	Tyr	Thr	Lys	Val	Phe	Ala	Leu	Asp	Val
	2105		2110		2115									
Ala	Ala	Arg	His	Arg	Val	Gln	Phe	Ala	Met	Pro	Glu	Ser	Asp	Ala
	2120		2125		2130									
Val	Ala	Met	Leu	Lys	Gln	Leu	Ser							
	2135													

<210> 15
 <211> 257
 <212> PRT
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 1306049CD1

<400> 15
 Met Glu Phe Pro Asp Leu Gly Ala His Cys Ser Glu Pro Ser Cys

10

20

30

40

```

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

```

10

20

```

<210> 16
<211> 517
<212> PRT
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Incyte ID No: 3187174CD1

```

```

<400> 16
Met Ala Ala Pro Glu Gln Pro Leu Ala Ile Ser Arg Gly Cys Thr
  1           5           10           15
Ser Ser Ser Ser Leu Ser Pro Pro Arg Gly Asp Arg Thr Leu Leu
      20           25           30
Val Arg His Leu Pro Ala Glu Leu Thr Ala Glu Glu Lys Glu Asp
      35           40           45
Leu Leu Lys Tyr Phe Gly Ala Gln Ser Val Arg Val Leu Ser Asp
      50           55           60
Lys Gly Arg Leu Lys His Thr Ala Phe Ala Thr Phe Pro Asn Glu
      65           70           75
Lys Ala Ala Ile Lys Ala Leu Thr Arg Leu His Gln Leu Lys Leu
      80           85           90
Leu Gly His Thr Leu Val Val Glu Phe Ala Lys Glu Gln Asp Arg
      95           100          105
Val His Ser Pro Cys Pro Thr Ser Gly Ser Glu Lys Lys Lys Arg
      110          115          120
Ser Asp Asp Pro Val Glu Asp Asp Lys Glu Lys Lys Glu Leu Gly

```

30

40

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

10

20

30

<210> 17
 <211> 4456
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 2530775CB1

40

<400> 17

```

caaaaacgcg atgtcgatct gacggatcac tcataggaca gctatgacgt cgcgtgctcg 60
cgtgcggttta ctccggtcttc ggtgcgagag aagataaaga aaaaacggaa gttatcctta 120
gtggtttgct tgcattact gaccaggtac tacttccatc tctttctttg atggactgca 180
atgcttgtat gctcgaggaa ctatggggaa tgtttaaac atttccatat cagcatagat 240
atcgtctgta tggccagtg aagaatgaaa cttataacag tccccactt ttagtaaaag 300
ttaaagctca acaaatgac agagccaaat atatcatgaa ggcctaac aaggaaaatg 360
tgaagccttc tggaaagaca attgggaagt atgacacaac catctcaagc attttgtttg 420
attatatctt gtcacaaata cagaagtatg ataacttaat aacacctgta gtagattcat 480
tgaataacct cacttcactg aattatgatg tcttggccta ttgtatcatt gaagctttag 540
ctaaccaga atgaacatg atgaacatg atgacacaac catctcaagc tggcttcaga 600
gtctggctag tttctgtggt gcagtttttc gtaaatatcc aattgatctt gctggctctc 660
ttcagtatgt tgccaatcag ctaaaggcgg gcaaaagttt tgacctgctt atattgaaag 720
aagtgttaca aaaaatggca ggaatagaaa ttacagagga aatgacaatg gagcaactag 780
aggctatgac tgggtggag cagctaaaag ctgagggtgg ttattttggg cagatcagaa 840
acactaaaaa atcctctcag agattaaagg atgctctatt ggacctgat ctgcccctc 900
ctctctgtct gcttatggct cagcagagaa atggggtaat ctttcaggaa ggtggagaga 960
aacatttgaa acctgtggga aagctctatg accagtgtca tgataacctg gtgcagttt 1020
gtgggttttt agcatccta ctgagccacg atgattatat aaagcgagt ccttcaattg 1080
atgtactctg taatgaattt catacaccct atgatgcagc atttttctctg tctaggccaa 1140
tgtatgcccc tcatatctcg tcaaagtatg atgaacttaa aaaatcagaa aagggaagta 1200
aacagcaaca taaagttcat aagtacatta catcatgtga gatggtgatg ggcctgtctc 1260
atgaagcagt ggtctcctta catgtttcca aagctctggga tgacatcagc cctcaattct 1320
atgctacatt ctggctcatt acaatgtatg accttgcagt tccacacacc agctatgaac 1380
gagaagtcaa taaacttaaa gtccagatga aagcaattga tgacaatcag gaaatgcccc 1440
caaataaaaa gaaaaaagag aaggagcgtg gtaactgcct tcaggacaag ctcttgaag 1500
aagaaaagaa acagatggaa catgtacaga gagttctaca gagattgaaa ctggaaaagg 1560
acaactggct ttagcaaaa tctaccaaaa atgagaccat cacaaaattt ctacagctgt 1620
gtatatttcc tcatgtatt ttttcagcaa ttgatgctgt ttactgtgct cgttttgttg 1680
aattggtaaca tcaacagaaa actccaaatt tttccacact ctttggctat gatcgagttt 1740
tctctgacat aatttacaca gttgcaagct gtactgaaaa tgaagccagt cgatacggaa 1800
ggtttctttg ctgcatgtta gagactgtga ccagtgggca tagtgataga gccacatatg 1860
aaaaggaatg tggaaactat ccaggattcc ttaccatatt acgggcaact ggatttgatg 1920
gtggaaataa gctgatcaa ttagactatg aaaattttcg acatggtgta cataaatggc 1980
attacaaact aaccaaggca tgggtacatt gccttgaaac aggcgaatat actcacatca 2040
ggaatacctt gattgtgcta acaaaaatac ttccttggta cccaaaagt ttgaaatctg 2100
gtcaagcttt ggaagaaga gtacacaaaa tctgccaaaga agaaaaagag aagaggccag 2160
atctatatgc attggtatg ggtactctcg ggcagttgaa aagttagaaag tcatacatga 2220
tacctgaaaa tgagtttcat cacaaagacc cccctccgag gaatgcagtt gccagtgctc 2280
aaaatgggccc tgggtgggg ccttcttcat catcaatagg aagtgcactc aaatcggatg 2340
aaagcagtac tgaggagact gataaatcaa gggagagatc tcagtgtggt gtgaaagctg 2400
ttaaataaagc tctagtacc acacctaagg ggaattcaag caatggaaat agtggctcta 2460
acagcaacaa agctgttaaa gaaaatgaca aagaaaaagg gaaagagaaa gaaaaagaga 2520
aaaaagaaaa gactccagct actactccag aggccagggt acttggtaaa gatggtaaa 2580
aaaaaccaaa ggaagagcgg ccaataaag atgaaaaagc aagagagacc aaggaaagaa 2640
cgccgaagtc tgacaaagag aaagaaaaat tcaagaagga agaaaaagct aaagatgaga 2700
aatttaagac cactgtcccc aacgcagaat caaaaatcac tcaagaaagg gaaagagaga 2760
aggagccatc cagagaaaga gatatagcaa aggaaatgaa atcaaaaggaa aatgttaaag 2820
gaggagaaaa aacaccagtt tctgggtcct tgaatcacc tgttcccaga tcagatattc 2880
cagagcctga aagggacaaa aaacgcccga aaattgatac tcccccttct ccatcacatt 2940
cctccacagt aaaggacagt ctcatcgaac tcaaggaaatc ttcagcaaag ctctacatta 3000
atcatactcc tccaccactg tccaagagta aggagagaga aatggacaag aaagatttg 3060
acaagtcaag ggaagatcc agagaagag agaaaaaaga tgaaaaggac aggaaagagc 3120
ggaaaaggga tcaactcaaac aacgaccgtg aagtgccacc ggacttaacc aagagacgta 3180
aagaggagaa tggacaatg ggggtttcaa aacataaaag tgaaagtcct tgtgaaatctc 3240
cttatccaaa tgagaaagac aaggaaaaaa ataagtcaaa atcttcaggc aaagaaaaag 3300
gcagtgattc atttaaatct gagaagatgg ataaaaatctc ctccgggtggc aaaaaggagt 3360
ccaggcatga taaagaaaag atagaaaaga aagagaaacg ggacagttca ggaggaagg 3420
aagagaagaa acatcataag tccctcggaca agcacagata atgaagactt tccatcaagg 3480
ctatggacag acccctaag ccgaaagtct cccagaggag gatgccaaag ctccgtgct 3540
cactctcaga atgatatctc ccataattaa gagcttctgg tctgtccat ttaatgggaa 3600
tctgctttaa gtcagaagat gactacacac caccatctct ttgacgagac attccaagt 3660
cactgatttt cagaacatta ttttcacctc caggcagttt cttacagcaa ggtccctgtg 3720

```

10

20

30

```

tatacatttt ttactcgaga tacaccatcc agaatcagca tctaatagaaa atttcaactga 3780
gttttagttc attcttcttt ctctgaaagg agaggaaatc gctcccagga accagttcct 3840
taatgacgta gatagggcgt ctttbtgtgt gaactcctat acctttgaca ttactgggat 3900
gattatattg cttgaggatt ttggcttcta gtaaaaattt taatttcagt tcctggggaa 3960
gacgttcttg acgtgtttta acagcaacag caattttatc ctttaatgtg accttaaatg 4020
tgtcaccaaa atttcccttg cctcttgtaa cgtggcacct ttatgattga gaaccatttt 4080
cttattctcc taatggccat actgtgatac catgatgtcc ttaattggaa cattgacttt 4140
tttttttttt tgcaatttaa acaattgaga taaaattcat ataacatcaa atttaccttt 4200
cacttttttt ttttcttttg gggggcaaaag tcccgtgtt gtccccggg tgggagtcaa 4260
ggggggagat cggggccac tgggaccccc gccccgggg ttaaagaaat tccccctccc 4320
aagcaatttg gggggctggg gcccgaaaat tccttaaac cggggggttg gggttccctt 4380
ggcccaaaat ggggccaatg cctttaaag ggggggaaaa aaaaccctgc tcaaaaaaag 4440
aattttccct ggtggc 4456

```

```

<210> 18
<211> 3662
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Incyte ID No: 926296CB1

```

```

<400> 18
acggaatctc gggcttctg acgtgccggg cgggaagatg tcatcattgc caagaagagc 60
gaaagtacag gtcaggatg ttggtactgaa agatgaattt tcttcattct ctgagttatc 120
atctgcctct gaagaagatg acaaggaaga tagtgcctgg gagcccaaa agaagttcc 180
cagaagccgt aaacagcccc ctccaagga atccaacca aagaggatgc ctgggtgaa 240
gaagaatgcc ccacagatca gtgatggctc agaagtcgtt gttgttaagg aggagctgaa 300
tagctctgtg gctattgtct atactgcctt agaagacaga aaaaataaat tggatactgt 360
acagactctg aaaacagcca agacaaaaca gaaatgtgca gcgagccac atacagttcg 420
aaggactaaa aagctgaaag ttgaagaaga aaccagcaa gccagcaact tagaggggtg 480
gagtaatagt tcagagacac catccacaag cactgtgtgg ggaggtaacat gcaagaagga 540
agagaatgac gatgacttta catttggctc gtccgcttta aagaaaatca agactgagac 600
atctctcag gggcagcctg tcaagtttcc agcaaatgcc aatagtacta aagaggaggt 660
ggaaatgaat tgggacatgg tacaggtttt atctgagaga actaatatg aaccttgggt 720
ttgtgccaac atcattcgtc tctttaatga tgataacaca attcccttca ttatacgtta 780
tagaaaagag ctatctaata accttgatgc tgattccttg agagaagttc agcaaccct 840
agaagagcta cgggctgttg caaagaaagt tcatagtaca atccagaaa ttaagaagga 900
aggggaagatg tctgagtgct tgtttaaagc catgctgaat tgtaaaactt ttgaagaact 960
agaacacgtg tctgctccat ataaaactgg aagcaaaagg actaaagccc agagagcaag 1020
acagttgggc ttagaagag cagccagggc actgcttgag aaaccagggg agctcagtc 1080
gctatcgtac attaggcctg acgttaaagg gctttcaacg cttcaggata ttgaaatagg 1140
agtgcagcat attttagcag atatgattgc taaagacaaa gacacgctg acttcattcg 1200
gaacttgtgc cagaagagac atgtttgtat ccagtcactc ctggcaaaaag tatcctcaa 1260
aaaggtaaat gagaaagatg ttgataagtt tctgctctac cagcattttt cctgcaacat 1320
aagaaacatt caccatcacc agattctggc aattaaccgt ggagaaaatt tgaagggtact 1380
gacggttaag gtcaatattt ctgatggagt gaaggatgaa ttctgtaggt ggtgcatcca 1440
aaacaggtgg agaccagta gctttgcaag gccagagtta atgaagatct tatataattc 1500
actgaatgat tctttaaac gcttattta tctcttctc tgtagagaat tcagagccaa 1560
actaacatca gatgcagaga aggaatoagt aatgatgttt ggacggaacc ttcgtcagct 1620
ccttttaaca agccctgttc cagggcgcac cttaatggga tgggatcctg gttataaaca 1680
tggttgcaaa ttagctataa tttctctac tagtcagata cttcactctg atgtggttta 1740
cttgcatgtg ggacaaggct tccgagagcc ggagaaaata aagacacttt tgcgtaattt 1800
caactgcagc acagtagtga ttggaatgg aactgcctgc agggaaacag aagcttactt 1860
tgctgacctg ataatgaaga attattttgc accactggat gttgtttact gtatcgtcag 1920
tgaagcagga gcatcaatct acagtgctcag ccctgaagct aacaaagaga tgcaggggct 1980
ggaccctaatt ttgagaagtg cagtttccat agcaaggcgt gtacaagatc cattagctga 2040
gctagtgaata attgagccaa agcacattgg agttggaatg taccagcatg acgtatccca 2100
gactttactc aaggcaacac tggacagtg tgtagaagaa tgtgtcagct ttgtgggagt 2160
ggatattaac atctgttcag aagttttgtt aaggcatatt gcaggactca atgccaacag 2220
ggccaaaaat attattgaat ggcgagagaa aatggacc tttatcaacc gagaacagct 2280
gaagaaagtg aaagggctgg gcccaaaatc ctccaacag tbtgctggct tcatcagaat 2340

```

10

20

30

```

caaccaggat tatatccgaa cgttttgcag tcagcaaact gaaacttcag gccaaattca 2400
aggagttgct gtgacatctt cagcagacgt tgaggtcaca aatgagaagc agggcaaaaa 2460
gaagagcaaa actgcagtgat atgttttact gaagccaaat cctttggacc aaacttgtat 2520
tcatccagaa tcatatgaca tagcaatgag gtttttgcag tccattggag ggacactgta 2580
tgaggttggg aagcctgaaa tgcaacaaaa aataaattca ttctctgaaa aggaaggaat 2640
ggagaaaaat gcagaaagat tgcaacaacac agtacacacc ttacaggtca tcatagatgg 2700
tctcagccag cctgaaagct ttgactttcg aacagatttt gataaacctg atttcaagag 2760
aagcatagta tgctggaag atctgcagat tgggacagtt cttacaggca aagttgagaa 2820
tgccactctc tttggaattt ttgtggatat aggagtgggg aaatctgggc tgattcccat 2880
acgaaatgta acagaagcaa aactttcaaa aacaaagaag agaagaagcc ttggactggg 2940
ccccggagaa agagtgggag tccaagtact caacattgac atccccgat ctaggattac 3000
ctcggactct attcgggtgt tatgagtatc ccacgaaggc cagacgctga ttttattttc 3060
tcatttccac agattgacaa ggataagtca gttgtttgta aactctaggt agcagatgag 3120
aaataattca ctaaatatca gaaatatttt ccaaacactt tctcttattt tttcttctga 3180
ataaatagaa aacgaacagt ttgatttcct tttcccttaa aggaaacaac taatacacat 3240
tcttatatgg ctttatgtag taatagtttt ctgactaaaa tttttgtttt tattttttgt 3300
aatttatctt taactccttt tgcattttgt ataacagatt gcttaacttc tacttgccaa 3360
catctgcctt gctggacttg tatgggattg tottcttgat ttgaaattgta cegtctttgt 3420
tgacacagta gggtgggca gtgtttaatc cttccatttt atagattttt ttttaatcag 3480
gccttttggg cttcattcat aattttgcag taatctcttt tcccttgtca tgcaagccaa 3540
aatatacca gtaaaacaga ttctgacgtg tttgtagtta tcaaatgaat ggctcgaaac 3600
acttctcaaa aggatatacg tattgacccc aacaataaat gtttgtggct agtgaaaaaa 3660
aa

```

10

```

<210> 19
<211> 2201
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Incyte ID No: 1322761CB1

```

20

```

<400> 19
ccccaaaca acgagtcccc aattctcgtc cgtcctcgcc gcgggcagcg ggcggcggag 60
gcagcgtgcg gcggtcgcca ggagctggga gccagggcg cccgctctc gccgcagcat 120
gttcagccg gcgcccagc gctgcttcac catcgagtcg ctgggtggca aggacagtcc 180
cctgcccgcc tcgctcgcg aggaccccat ccgtcccgcg gcactcagct acgctaactc 240
cagccccata aatccgttcc tcaacggctt ccaactcgcc gccgcccgg ccgccggtag 300
gggctctac tcaaacccgg acttggtgtt cgccgaggcg gctctgcacc cccccacc 360
cgccgtgcca gtgcaccggc tgcccgccgc gcacgcctg gccgccacc ccctaccctc 420
ctgcacctcg ccacaccoc tattcgctc gcagcagcgg gatccgtcca cctctacc 480
ctggctcatc caccgctacc gatactggg tcatcgctc caaggaacg aactagccc 540
cgagagtttc cttttgcaca acgcgtggc cggaaagccc aagcggatcc gaaccgctt 600
ctccccgtcc cagcttctaa ggctggaaca cgccttgag aagaatcact acgtgggtgg 660
cgccgaaagg aagcagctgg cacacagcct cagcctcacg gaaactcagg taaaagtatg 720
gtttcagaac cgaagaacaa agttcaaaag gcagaagctg gaggaagaag gctcagattc 780
gcaacaaaag aaaaaaggga cgcaccatat taaccggtgg agaatcgcca ccaagcaggc 840
gagtcgggag gaaatagacg tgacctcaga tgattaaaaa cataaaccta accccacaga 900
aacggacaac atggagcaaa agagacaggg agaggtggag aaggaaaaaa cctacaaaa 960
caaaaacaaa ccgcatacac gttcaccgag aaaggagag ggaatcggag ggagcagcgg 1020
aatgcccgga agactctgga cagcaggggc acaggtccc aaaccgagg ccgccaaga 1080
tggcagagga tggaggtcc tcatcaaca agcgaccctc gctaaagag gcagctgagt 1140
gagagacaca gagagaagga gaaagagggg gggagagaga gaaagagaga gaaagagaga 1200
gagagagaga gagaaagctg aactgacact ctgacaaggg gagctgtcaa tcaaacacca 1260
aacgggggag acaagatgat tggcaggtat tccgtttatc acagtcact taaaaaatga 1320
tgatgatgat aaaaaccacg acccaaccag gcacaggact tttttgtttt ttgacttcg 1380
ctgtgtttcc ccccatctt taaaaataat tagtaataaa aaacaaaaat tccatatcta 1440
gccccatccc acacctgtt caaatccttg aatgcattg agcagttgtt gggcgaatgg 1500
tgtttaaga ccgaaaatga attgtaattt tcttttctt ttaaagacag gttctgtgtg 1560
ctttttattt tgatttttt tccaagaaa tgtgcagtct gtaaacactt tttgatacct 1620
tctgatgtca aagtattgt gcaagctaaa tgaagtaggc tcagcagatg tggctctctt 1680
acagagaaac ggggagcagg acgacggggg ggctgggggt gcggggggag ggtgccca 1740

```

30

```

aaaagaatca ggacttgtac tgggaaaaaa acccctaact taattatatt tcttggacat 1800
tccctttcct aacatcctga ggctbaaaac cctgatgcaa acttctcctt tcagtgggtg 1860
gagaaattgg ccgagttcaa ccattcactg caatgcctat tccaaacttt aaatctatct 1920
attgcaaaac ctgaaggact gtagttagcg gggatgatgt taagtgtggc caagcgcacg 1980
gcggaagtt ttcaagcact gagtttctat tccaagatca tagacttact aaagagagtg 2040
acaaatgctt ccttaatgtc ttctatacca gaatgtaaat atttttgtgt tttgtgttaa 2100
ttgttagaa ttctaacaca ctatatactt ccaagaagta tgtcaatgtc aatattttgt 2160
caataaagat ttatcaatat gccctcaaaa aaaaaaaaaa a 2201

```

```

<210> 20
<211> 3188
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Incyte ID No: 7472664CB1

```

10

```

<400> 20
gtgaatctag tctgtgtgat tttaaaactc ttgcttttga cctctgtgct atacttagtg 60
cttgtgtatt gcaactctaa aagttgattg attgttgatg gccattacta cctataaagt 120
ttactgggta aaagagattt tggctcttatt actcttaaat atttgctaac tatattccga 180
cgctgatttt taattttttt ttttaacctg gtcatgcctaa tagagttttg tgtttccttg 240
tttggctggc tggttccttt cttccctgg ttgcttttga taggatactg tggagcagat 300
attaatcaa tatgtgctga agctgcttta tgtgctttac gacgacgcta cccacagatc 360
tataccacta gtgagaaaat gcagttggat ctctcttcaa ttaatatctc agctaaggat 420
ttcgaggtag ctatgcaaaa gatgatacca gcctcccaaa gagctgtgac atcacctggg 480
caggcactgt ccaccgttgt gaaaccactc ctgcaaaaac ctgttgacaa gatttttagaa 540
gccctgcaga gactatctcc acatgcagaa ttcagaacaa ataaaacatt agactcaggt 600
ataaaacttg attgaccatt taaagttagt ttttataatc agtaaagtga aatgtttcat 660
acataccaga aatattttat tgtaaaagaa tctcttgaaa agaatttggg aatgctttac 720
ctgggatatt tgaaggaaat ttactccatg gattgccttt aatttaactt attgtaagtt 780
tggggtgcat ttttaccatt acttattaat tcatactctac acaaatatth ttgattcctt 840
gaaaaaaatt taattgaagt atgtgactt ggttaactgaa gcaataaaga tttgctttgc 900
aatccaaaag ttactaatgt acaatttttt ggtttttag atatttcttg tctctgcta 960
gaaagtgact tggcttacag tgatgatgat gttccatcag ttatgaaaa tggactttct 1020
cagaaatctt ctcataaggc aaaagacaat ttttaatttc ttcaattgaa tagaaatgct 1080
tgttaccaac ctatgtcttt tcgaccaaga atattgatag taggagaacc aggatttggg 1140
caaggttctc acttggcacc agctgtcatt catgctttgg aaaagtttac tgtatataca 1200
ttagacattc ctgttctttt tggagttagt actacatccc ctgaagaaac atgtgccag 1260
gtgattcgtg aagctaagag aacagcacca agtatagtgt atgttctca tatccacgtg 1320
tgggtggaaa tagttggacc gacacttaaa gccacattta ccacattatt acagaatatt 1380
ccttcatthg ctccagtttt actacttgca acttctgaca aacccattc cgctttgcca 1440
gaagaggtag aagaattggt tatccgtgat tatggagaga tttttaatgt ccagttaccg 1500
gataaagaag aacggacaaa attttttgaa gatttaattc taaaacaagc tgctaagcct 1560
cctatatcaa aaaagaaagc agttttgcag gctttggagg tactcccagt agcaccacca 1620
cctgagccaa gatcactgac agcagaagaa gtgaaacgac tagaagaaca agaagaagat 1680
acatttagag aactgaggat tttcttaaga aatgttacac ataggcttgc tattgacaag 1740
cgattccgag tgtttactaa gcctgttgac cctgatgagg ttcttgatta tgtcactgta 1800
ataaagcaac caatggacct ttcatctgta atcagtaaaa ttgatctaca caagtatctg 1860
actgtgaaag actatthtgg agatattgat ctaatctgta gtaatgcctt agaatacaat 1920
ccagatagag atcctggaga tctgtcttatt aggcataagag cctgtgcttt aagagatact 1980
gcctatgcca taattaaga agaactgat gaagactttg agcagctctg tgaagaaatt 2040
caggaatcta gaaagaaaag aggtgttagc tcctccaaat atgcccctgc ttactacat 2100
gtgatgccaa agcaaaatbc cactcttggt ggtgataaaa gatcagacc agagcagaat 2160
gaaaagctaa agacaccgag tactcctgtg gcttgacgca ctctgtctca gttgaagagg 2220
aaaattcgca aaaagtcaaa ctgttactta ggcaccataa aaaagcgaag gaagatttca 2280
caggcaagg atgatagcca gaatgccata gatcacaana ttgagagtga tacagaggaa 2340
actcaagaca caagtgtaga tcataatgag accggaacaa caggagagtc ttcgggtgaa 2400
gaaaatgaaa aacagcaaaa tgctctgaa agcaaaactgg aattgagaaa taattcaaat 2460
acttgtaata tagagaatga gcttgaagac tctaggaaga ctacagcatg tacagaattg 2520
agagacaaga ttgcttgtaa tggagatgct tctagctctc agataataca tatttctgat 2580
gaaaatgaag gaaaagaaat gtgtgttctg cgaatgactc gagctagacg ttcccaggta 2640

```

20

30

```

gaacagcagc agctcatcac tgttgaaaag gctttggcaa ttctttctca gectacacc 2700
tcacttgttg tggatcatga gcgattaaaa aatcttttga agactggtgt taaaaaaagt 2760
caaaactaca acatatttca gttggaaaat ttgtatgcag taatcagcca atgtatttat 2820
cggcatcgca aggaccatga taaaacatca cttattcaga aaatggagca agaggtagaa 2880
aacttcagtt gttccagatg atgatgtcat ggtatcgagt atcttttata ttcagttcct 2940
atttaagtca tttttgtcat gtccgcctaa ttgatgtagt atgaaaccct gcacttttaa 3000
ggaaaagatt aaaaatagtaa aataaaaagta tttaaacttt cctgatattt atgtacata 3060
taagataaat gtcattgtga agataactga taaatattgg aactttgcta gaacaagacc 3120
ctgtagtaat agtaataata gttgaagttt ggccaactct taataaagtt attttggtaa 3180
ctaaaaaa 3188

```

```

<210> 21
<211> 843
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Incyte ID No: 7473124CB1

```

```

<400> 21
atgcctatct actcccagac agtggccatg gctgaacact ttaaacaagc aagcagttgt 60
cctatctgcc tggattatct tgaaaacccc acgcacctga aatgtggata catctggtgc 120
ctccgatgca tgaactcact gcgaaagggg cccgatggga aggggggtgct gtgccctttc 180
tgccctgtgg tctctcagaa aaatgacatc aggcccgtg cccagctggg ggcgctgggt 240
tccaagatca aggaactaga gcccaaggtg agagctgttc tgcagatgaa tccaaggatg 300
agaaagtccc aagtggatg gaccttggat gtggacacag ccaacaacga tctcatcggt 360
tctgaagacc tgaggcgtgt ccgatgtggg aatttcagac agaataggaa ggagcaagct 420
gagaggttcg acactgccct gtgctcctg ggcacccctc gcttcacttc cggccgccat 480
tactgggagg agggcggtgg caccagccaa gtgtgggatg tgggcgtgtg caaggaatct 540
gtgaaccgac agtgggaacgt tgtactctct tcagaactcg gcttctggac tgtgggtttg 600
agacaaggac agatctactt tgccagcact aagcctgtga cgggtctctg ggtgagctca 660
ggtctacacc gagtggggat ttacctggat ataaaaacga gggccatttc cttctataat 720
gtcagtgata ggtcacatat cttcacattc acgaaaattt ctgctactga gccactgcgc 780
ccatgttttg ctcatgcaga tacaagtcgt gatgatcag gatacttgag tgtgtgtgtg 840
taa 843

```

```

<210> 22
<211> 1102
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Incyte ID No: 7473171CB1

```

```

<400> 22
gtccgcggtt ctgtgaggtc cagtggccgc ccaggcgcga ccagatctgg gtgcccggag 60
agcgcgcatg gcggtgtggt gaccgcggac cggcccggga accggcgcgc aggcctctagc 120
gtggcgggca gagctgcagg gcgagggcag gtgctccatc tgcttagagc tctttcgtga 180
gccggtgtcc gtcgagtgcg gccacagctt ctgccgcgcc tgcatagggc gctgctggga 240
gogcccgggc goggggtctg ttggggccgc caccgcgcgc ccccccttc cactgccctg 300
tccgcagtgc cgcgagcccg cgcgccccag tcagctgcgg cccaaccgpc agctggcggc 360
agtggccaag ctcttgcggc gcttcagcct gcccgcggt gccccgggag agcacgggtc 420
tcaggcggcc gcggcccggg cagcggctgc cgcctgcggg cagcatggcg aacccttcaa 480
gtctacttgc caggacgacg gacgcgccat ctgctgtggt tgcgaccgpc cccgcgagca 540
ccgcgagcac gccgtgctgc cgttgacga ggccgtgcag gaggccaagg agctcttggg 600
gtccaggctg aggttcttga agaaggaact ggaggactgt gaggtgttcc ggtccacgga 660
aaagaaggag agcaaggagc tgctggtgag ccaggcacc gcaggcccc cgtgggacat 720
tacagaggcc tgagaactca gcaccagggc tcggtgtgtg tgggtgttgg gtgtgtgcta 780
tggaaaccgca gaatcgattt cagaaagata atagagtcca tattatatag ggtgtccaca 840
taattgttgt acaaaccaga gctttttaa gtgaaaagca gtgctaaaat aattattgca 900
aaacaactgg cttaaactgg agctgtccca gcgaatcagg acgctcagtc actctgatat 960

```

10

20

30

```

tacgtaacat accagttagg gcctgcggaa gcatcttgta atggaacaca ttactatttc 1020
tgcagagaaa catggatatt caataagtgg gaatataat acaataaaga gcctcatggc 1080
atgttttgtc aacaaaaaaaa aa 1102

```

```

<210> 23
<211> 2481
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Incyte ID No: 7477026CB1

```

```

<400> 23
gcagcccaca aagtgcgatg ccccgggctg cgaccggagc gagccccatg ccccgcgcg 60
gctgacgggc cggagcccgc accgagcggc cgggcyggac aggtaccgct agagcgacat 120
gatgggtgaa tctgcctcgg agacaatcag gtcggctcca tctggtcaga atggcgtgg 180
cagcctctct gggcaagccg atggcagcag cggcggggcc acaggggaca ctgcaagtgg 240
cacgggcagg gaagtgacca cgggtgcaga cagcaatggt gagatgagtc ccgcagagct 300
gctgcacttc cagcagcaac aggcctctcca agtggcccgg cagttcctgc tgcagcaggc 360
ctcaggcctg agctccccag ggaacaatga cagcaaacag tctgcctctg ctgtgcagg 420
gcctgtgtcg gtggccatga tgtcgcggca gatgcttacc ccgcaacaga tgcagcagat 480
cctgtgcagg ccagcagctg aggccttgct ccagcagcag caagccctc tgctccagca 540
gctacaggag tactacaaga agcagcagga gcagctccac ctgcagctcc tcaccagca 600
gcaggctggg aaaccgcagc ccaaagaggc actggggaac aagcagctgg ccttccagca 660
gcagctcctg caaatgcaac agttgcagca gcagcacctg ctcaacctgc agaggcagg 720
gctgtcagc ctgcagccca accaagcctc ggggcccctc cagacccttc cgcaagcagc 780
tgtttgcca acagacctgc cccagctgtg gaagggcgag ggtgcccccg ggcagcctgc 840
cgaggacagc gtaagcagg aggggctgga cctcactggc acggccgcca ccgtacctc 900
gtttgcccgt cccccaagc tctcaccctc cctctcccac cataacctgc ccaacggaca 960
gctactgtg ctcacatctc ggagagacag ctcttcccac gaggagacc ccggctcca 1020
ccccctgtac ggacacggag agtgcagtg gccaggctgt gagaccctgt gtaagacct 1080
gggccagttt atcaaacacc tcaaacacaga gcacgcctgt gatgaccgga gtacagcca 1140
gtgccgggta cagatgcagg tggtcagca catggggccc tcggagccca agccttcag 1200
gaggctgag gccatgatg cccacctgca catggggccc tcggagccca agccttcag 1260
ccagccagtg accgtctctg cagcagactc attcccagat ggtctcgtgc acccccagc 1320
ctcggccgca gccctgtca cccctctacg gcccctggc ctgggctctg cctccctgca 1380
tggtaggggc caagcccgtc ggagaagcag tgacaagttc tgcctcccca tctcctcaga 1440
gctggcccag aatcatgagt tctacaagaa cgccgacgtc cggccccctc tcacctacg 1500
ctccctcacc cgcaggccca tccctgaaac ccctgacagg cagctgacc tgatgagat 1560
ctataactgg ttaccagga tgttcgccta tttccgcaga aacctgcca cctggaagaa 1620
cgccgtgcgc cacaacctca gcctgcacaa gtgcttcgtc cgctgggaga acgtcaagg 1680
tgccgtgtgg actgtggacg agcgggagta tcagaagcgg agaccgcaa agatgacag 1740
gagccccacc ctggtgaaga acatgatctc tggcctcagc tatggagcac ttaatgccag 1800
ctaccaggcc gcctggccg agagcagctt cccctcctc aacagccctg gcatgctgaa 1860
ccctggctcc gccagcagc tgetgcccct cagccaagat gacgtgggtg ccccggtgga 1920
gcccgtgccc agcaacggca gcagcagccc tccctcgcctc tccccgccc agtacagcca 1980
ccaggtgag gtgaaggagg agccagcaga ggcagaggaa gacaggcagc ccgggcctcc 2040
cctgggcgcc cctaacccca gcgcctcggg gcctccggaa gacagggacc tggaggagga 2100
gctgccggga gaagaactgt cctaagggcc tgtagtgacc ggcagggtg ggtgagacc 2160
cctcccttcc agaatccagg ccccatctcc cccaactcca cagcccctcc cgagcctcaa 2220
ggcaagtcca gactcagac cggggaggcc cgggcccagca gctcccagtg tgacctgaca 2280
aaaacacgta ggggcaggga cggctcccac ccccayggac acaaccctg gtcttgacc 2340
agtagaggac acggagggtt cagaccctc ctcagacct cccacatct gaaactgcct 2400
cccccaacc accagcagca gcagggccct cctccccac cagctctccc cacagggcc 2460
ctcagcatca tggagaccg c 2481

```

10

20

30

```

<210> 24
<211> 2308
<212> DNA
<213> Homo sapiens

```

```

<220>

```

<221> misc_feature
<223> Incyte ID No: 6428773CB1

```

<400> 24
cgtttgccag cgctcaggca ggagctctgg actgggcgcg ccgcccctt ggagtgaggg 60
aagcccagtg gaaggggggt cggggagccg gctgcgatgg acgcccgtct ggaacccctc 120
ccggccgaca ggtgttccc cggatccagc ttcctggact tgggggatct gaacgagtcg 180
gacttcctca acaatgcgca ctttcctgag cacctggacc actttacgga gaacatggag 240
gactttctca atgacctgtt cagcagcttc ttgatgacc ctgtgctgga tgagaagagc 300
cctctatggy acatggaact ggactcccct acgccaaggca tccaggcggg gcacagctac 360
tccttgagcg gcgactcagg gcccagagc ccccttgtgc ccatcaagat ggaggacacc 420
acccaagatg cagagcatgg agcatgggcg ctgggacaca aactgtgctc catcatgggtg 480
aagcaggagc agagcccgga gctgcccgtg gaccctctgg ctgcccctc ggccatggct 540
gpcgcccggc ccatggccac caccocgctg ctgggctca gccccctgtc caggctgccc 600
atccccacc aggcccggg agagatgact cagctgccag tgatcaaagc agagcctctg 660
gaggtgaacc agttcctcaa agtgacaccg gaggacctgg tgcagatgcc tccgacgccc 720
cccagcagcc atggcagtg cagcagcggc tcccagagtc cccgctctct gccccctcc 780
agccctgtca ggccatggc gcgctcctcc acggccatct ccacctccc actcctcact 840
gcccctcaca aattacaggg gacatcaggg ccaactgctc tgacagagga ggagaagcgg 900
accctgattg ctgagggcta ccccacccc acaaaactcc cctcaccaa agccgaggag 960
aagcccttga agagagtcgg gaggaaaatc aagaaacaga tctcagccca ggagagccgt 1020
cgtaagaaga aggagtatgt ggagtgtcta gaaaagaagg tggagacatt tacatctgag 1080
aacaatgaac tgtggaagaa ggtggagacc ctggagaatg ccaacaggac cctgctccag 1140
cagctgcaga aactccagac tctggtcacc aacaagatct ccagacctta caagatggcc 1200
gccaccacaga ctgggacctg cctcatgggt gcagocctgt gctttgttct ggtgctgggc 1260
tccctctgtc cctgcccctc cgagttctcc tccggctccc agactgtgaa ggaagacccc 1320
ctggccgagc acgcccgtca cagcggcagg cagatgccc cccgaagcct cctattctac 1380
gatgacgggg caggcttatg ggaagatggc cgcagcacc tgctgcccac ggagcccca 1440
gatggctggg aatcaacccc cggggggccg gcagagcagc ggccccggga ccacctgcag 1500
catgatcacc tggacagcac ccacgagacc accaagtacc tgagtggggc ctggcctaaa 1560
gacgggtgaa acggcaccag ccccgacttc tcccactcca aggagtgggt ccacgacag 1620
gatctgggccc ccaacaccac catcaaaact tcctaggcca tgccaagacc caggacatag 1680
gacggacccc tggtacccag aagaggagtt cttgtcact aaccgggac cgctcgtgc 1740
ccctgctcc cttgagcttc cattccagga gaaaaggct cacttcccag ccttccctg 1800
cccctggcat ttggactct ccttgggccc gaccactctg ttctcattct ccttcccacc 1860
aacatccatc cgtccttctc agacaaacca ctactgggt accccacctc ctctctcata 1920
tgcccaacac gaccactgcc tccctgcccc cacacctgca cccaaacaga cacatcaacg 1980
caccaccctc acagacaccc cttaccccac cccactgta cagagaccaa gaacagaat 2040
tgtttgtaaa taatgaacct tattttttat tattgccaat ccctaagat attgtatatt 2100
acaaatctcc ctctcccct cgcccctccc ttgttttata ttttatgaag ttagtgccgg 2160
ctttgctgct ccttgcccca ggaaagaggg actaactgac cctcacctgg cacccccctg 2220
ctgctgcca agccgctgg cctttttaat tgccaaactg ctctcttcat cagctcagca 2280
catgctttaa gaaagcaaaa ccaaaaaa 2308

```

10

20

<210> 25
<211> 4369
<212> DNA
<213> Homo sapiens

30

<220>
<221> misc_feature
<223> Incyte ID No: 2749402CB1

```

<400> 25
gcaatgagcg gcgcccggag gctgtgacct gcgcgcgcg gcccgaccgg ggcccctgaa 60
tggcggctcg ctgagggcgc ggcgcgggcg gcggctcagg ctccctcggg cgtggcgtgg 120
cgggtgaaggg gtgatgctgt tcaagctcct gcagagacag acctatacct gcctgtcca 180
caggataggg ctctacgtgt gcttcttggg cgtcgttgtc accatcgtct ccgcttcca 240
gttcggagag gtggttctgg aatggagccg agatcaatac catgttttgt ttgattccta 300
tagagacaat attgctggaa agtcctttca gaatcggtt tgtctgcca tgccgattga 360
cgttgtttac acctgggtga atgacacaga tcttgaacta ctgaaggaac tacagcaggt 420
cagagaacag atggaggagg agcagaaagc aatgagagaa atccttggga aaaacacaac 480
ggaacctact aagaagagtg agaagcagtt agagtgtttg ctaacacact gcattaaggt 540

```

```

gccaatgctt gtcctggacc cagccctgcc agccaacatc accctgaagg acctgccatc 600
tctttatcct tcttttcatt ctgccagtga ctttttcaat gttgcaaac caaaaaaccc 660
ttctaccaat gtctcagttg ttgtttttga cagtactaag gatgttgaag atgccactc 720
tggactgctt aaaggaata gcagacagac agtatggagg ggctacttga caacagataa 780
agaagtcctt ggattagtgc taatgcaaga tttggcttcc ctgagtggat tccaccaac 840
attcaaggaa acaaatcaac taaaaacaaa attgccagaa aatctttcct ctaaagtcaa 900
actgtttgcag ttgtattcag aggcctgtgt agcgcttcta aaactgaata accccaagga 960
ttttcaagaa ttgtaataag aaactaagaa gaacatgacc attgatggaa aagaactgac 1020
cataagtcct gcataatatt tatgggatct gagcgccatc agccagtcta agcaggatga 1080
agacatctct gccagtcggt ttgaagataa cgaagaactg aggtactcat tgcgatctat 1140
cgagaggcat ttcagatggg ttcggaatat tttcatgtgc accaacgggc agattccatc 1200
ctggctgaac ctgacaactc ctcgagtgc aatagtaaca caccaggatg tttttcgaaa 1260
tttgagccac ttgctactct ttagttcacc tgctatbtgaa agtcacattc atcgcatcga 1320
agggctgtcc cagaagttta tttacctaaa tgatgatgtc atgtttggga aggatgtctg 1380
gccagatgat ttttacagtc actccaagg ccagaagggt tttttgacat ggctgtgccc 1440
aaactgtgcc gagggctgcc caggttcctg gattaaggat ggctattgtg acaaggcttg 1500
taataattca gctcgcgatt gggatgggtg ggattgctct gaaacacatg gaggggatcg 1560
ctatattgca ggaggtggag gtactggggg tattggagtt ggacagccct ggagtttgg 1620
tggagggaata taagtgatct cttactgtaa tcagggaatg cgaattcct ggctcgtga 1680
taagttctgt gaccaagcat gcaatgtctt gtcctgtggg tttgatgctg gcgactgtgg 1740
gcaagatcat tttcatgaat tgtataaagt gatccttctc ccaaaccaga ctactatat 1800
tattccaaaa ggtgaatgcc tgccttattt cagctttgca gaagtacca aaagaggagt 1860
tgaagtggcc catgtgaca atccaataat tcgacatgct tctattgcca acaagtggaa 1920
aaccatccac ctcataatgc acagtggaat gaatgccacc acaatacat ttaatctcac 1980
gtttcaaaat acaaacgatg aagagttaa aatgcagata acagtggagg tggacacaag 2040
ggagggacca aaactgaatt ctacggccca gaagggttac gaaaatttag ttagtcccc 2100
aacacttctt taagtgatg aaactgcagc aaactcttct cccaaagaaa aacgcttccc 2160
gaagtttaag agacatgatg ttaactcaac aaggagagcc caggaagagg tgaaaattcc 2220
cctggtaaat atttcaactc ttccaaaaga cgcccagttg agtctcaata ccttggattt 2280
gcaactgaaa catggagaca caactttgaa aggatacaat ttgtccaagt cagccttgct 2340
gagatcattt ctgatgaact cacagcatgc taaaaataaaa aatcaagcta taataacaga 2400
tgaacaacaa gatcagtttg tggctccaca ggaaaaacag gttcataaaa gcactttgcc 2460
aacagctta ggagtgtctg aaagattgca gaggttgact tttcctgcag tgagtgtaaa 2520
agtgaatggt catgaccagg gtcagaatcc acccctggac ttggagacca cagcaagatt 2580
tagagtggaa actcacacc aaaaaacat aggcggaaat gtgacaaaag aaaagcccc 2640
atctctgatt gttccactgg aaagccagat gacaaaagaa aagaaaatca cagggaaaga 2700
aaaagagaac agtagaatgg aggaaaatgc tgaanaatcac atagggctta ctgaagtgtt 2760
acttggaaaga aagctgcagc attacacaga tagttacttg ggctttttgc catgggagaa 2820
aaaaaagtat ttccaagatc ttctcgacga agaagagtca ttgaagacac aattggcata 2880
cttcaactgat agcaaaaaa ctgggaggca actaaaagat acatttgcag attccctcag 2940
atatgtaaat aaaattctaa atagcaagtt tggattcaca tcgggaaag tccctgctca 3000
catgctcac atgattgacc gtcattgctt gcaagaactg caagatatgt tccctgaga 3060
atttgacaag acgtcatttc acaaagtgcg ccattctgag gatatgcagt ttgccttctc 3120
ttatttttat tatctcatga gtgcagtgca gccactgaa atactctcaag tctttgatga 3180
agttgataca gatcaatctg gtgtcttgtc tgacagagaa atccgaacac tggctaccg 3240
aattcacgaa ctgcccgttaa gtttgcagga tttgacaggt ctggaacaca tgctaataaa 3300
ttgctcaaaa atgcttccctg ctgatatcac gcagctaaat aatattccac caactcagga 3360
atcctactat gatcccaacc tgccaccggt cactaaaagt ctagtaacaa actgtaaacc 3420
agtaactgac aaaatccaca aagcatataa ggacaaaaac aaatataggt ttgaaatcat 3480
gggagaagaa gaaatcgctt ttaaaatgat tcgtaccaac gtttctcatg tggttggcca 3540
gttggatgac ataagaaaa accctaggaa gtttgtttgc ctgaatgaca acattgacca 3600
caatcataaa gatgctcaga cagtgaaggc tgttctcagg gacttctatg aatccatgtt 3660
cccatacct tcccaatttg aactgccaag agagtatcga aaccgtttcc ttcataatga 3720
tgagctgcag gaatggaggg cttatcgaga caaattgaag ttttggacc attgtgtact 3780
agcaacatgt attatgttta ctatattctc attttttgcg gaggagttaa ttgcaactaa 3840
gcggaagata ttccacagaa ggaggataca caaagaagct agtcccaatc gaatcagagt 3900
atagaagatc ttcatttgaa aacctctac ctgagcattt actgagcatt ttaaaactca 3960
gcttcacaga gatgtctttg tgatgtgatg cttagcagtt tggcccgaag aaggaaaata 4020
tccagtacca tgcgtttttg tggcatgaat atagcccact gaccaggaat tatttaacca 4080
accactgaa aactgtgtgt ttgagcagct ctgaactgat tttactttta aagaatttgc 4140
tcatggacct taactccttt ttataaaaag gctcactgac aagagacagc tgtaatttc 4200
ccacagcaat cattgcagac taactttatt agggagaagcc tatgccagct gggagtgtat 4260
gctaagagge tccagctctt gcattccaaa gccttttgct aaagttttgc actttttcct 4320

```

10

20

30

ttttcatttc ccattttcaa gtagttacct aagtttaacta gttacttccg 4369

<210> 26
<211> 1445
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 118539CB1

```

<400> 26
aacagtcctcc ggtgctttgc aagggttcag attgcagaac agtcacatat gggatattgt 60
agttataaat acttgaccgt tctctagggg tggggtttcc aactctttaa acaaccctgt 120
tgccctgttat caagctgact ttttgctcca agtgccagaa ttgattgtgg ttaggacatc 180
ttacgctgag aaataaataa aatggaacat aatgggtctg cttcaaatgc tgataaaatc 240
caccagaatc gcctgtcgag tgttacagaa gatgaagacc aagacgctgc tcttaccatt 300
gtgactgtgc tggacaaagt agcctccatc gtggacagtg tgcaggcaag ccagaagaga 360
atagaagaga gacacagga aatggaaaat gccataaaat ccgtccagat tgacctggtg 420
aagctttcac agtcgcatag caatacaggg catatcatta acaaattggt tgagaaaacc 480
cgaaaaagtta gtgctccatc taaagatgtg aaagcccggg tggagaagca acaaattcat 540
gttaaaaaag ttgaagtcaa gcaagaggaa ataatgaaga aaaacaaatt ccgcgtggta 600
atattccagg agaagtttgc gtgtccgaca tccctgtctg ttgttaaaga cagaacctga 660
actgagaacc aagaagagga tgatgatgat atctttgatc ccccagtaga tctgtcttcc 720
gatgaagaat attatgttga agaaagcaga tctgccaggc ttaggaagtc aggcaggag 780
cacattgata atatcaagaa ggcattttcc aaagaaaaca tgcagaagac acggcagaat 840
cttgacaaga aagtgaacag aattagaact agaatagtga cccgggagag gagagagagg 900
ctaaggcagt caggagagag gctgagacag tcaggggaga ggctgagaca gtcaggggag 960
aggtttaaga aatctatttc taatgcagct ccctcaaagg aagcttttaa gatgocgagc 1020
ctcaggaaag gtaaggaccg aacagtggtc gaaggtgagg aatgtgccag ggagatgggt 1080
gtggacatca ttgccaggag cgagtcctcg ggccccatca gtgagctcta ctctgatgag 1140
ctcagtgaac cagaacacga ggcagccagg ccggtgtatc ctccccatga aggaagggaa 1200
atccccacc cagagccttt aaaagttact tttaaatctc aggtgaaagt agaggatgat 1260
gaatctcttt tgttagattt aaagcaactca tcgtaaagag gaattaagta tatcctaact 1320
atgaatctcc taatcatgca gttttagttt gaatagtgtg gtcgtctaca tttctgtgcc 1380
atgtaggaaa acataaatgt attttttttc ttatatatta aatcttgaag ataataataa 1440
tattt 1445

```

10

20

<210> 27
<211> 1608
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 4005918CB1

```

<400> 27
ggcagcgggg ggcgccccgg gggggcgtcg tcgtgggtac aatbgcgcag gggcaaaggt 60
cgagaggtcg cgtgcccgc gttttatttg aagacatcgt ccagttctga ccatggactc 120
gcagccatcg gcccttagtt tccatcccct ctagtgggcc ttcgggggct ctactgacgt 180
ccctccttcc cttggtaccg ggccggggaa gtgttctcgg gcgccccgag ttccgcatgc 240
ccaggcctgg ccaaggggaga tgaccgatcc gtcgctgggg ctgacagtcc ccatggcgc 300
gcctctggcc ccgctccctc cccgggacc aaacggggcg ggatccgagt ggagaaagcc 360
cggggccgtg agcttcgccc acgtggccgt gtacttctcc cgggaggagt gggctgcct 420
gcccccccg cagagggccc tgtaccggga cgtgatgcgg gagacctacg gccacctggg 480
cgcgctcgg gtcggaggca gcaagccggc gctcatctcc tgggtggagg agaaggccga 540
actgtgggat ccggtgccc aggatccgga ggtggcgaag tgtccgacag aagcggacc 600
agcagattcc agaacaaggg aagagggaaag acaaagggaa gggacgggag ccctggagaa 660
gccccaccct gtggccgccc ggtctcctgg gctgaaggct cccaagccc cctttgccgg 720
gtttggagcag ctgtccaagg ccggcgccc gagtcgcccc cgcttttttg cccaccccc 780
tgtccccga gctgaccagc gtcacggctg ctacgtgtgc gggagagct tcgctggcg 840
ctccacactg gtggagcaca tttacagcca cagggcgag aagcccttc actgcccaga 900

```

30

```

ctgcccgaag ggccttcggcc acgcttcttc cctgagcaaa caccgggcca tccatcgtgg 960
ggagcggccc caccgctgtc ccgagtgtgg tcgggccttc atgcgcgcga cggcgtgac 1020
ttctcacctg cgcgttcaca ctggcgagaa gccctaccgc tgcccgcagt gtggccgctg 1080
cttcggcctg aagaccggca tggccaagca ccaatgggtc catcggcccg gggcgaggg 1140
gcgtaggggc cggcgcctcg gggggctgtc tgtgaccctg actcctgtcc gcggggacct 1200
ggaccccctt gttggccttc agctgtatcc agagatattc caggaatgtg ggtgacggcc 1260
taaaaagtga ccatctagac attgtgggcg gcccgagatg ggctcagggg cccgaacctc 1320
tgcagcggcc tgcagggagg tcccagaatc caccgcaaga gctggcctgg ggtgaggaca 1380
gtctgatctt gggctctcag cagcctcttc tgccagcacc ttgctcccog ctgccctggg 1440
ctctccaagg ccccttttgc tgaggcaggg ctgaggtgag aacccccag acctccatc 1500
agggaaagcaa aagctgtttc tccctccaga gatgctaaga ggattgaggt agagaagaac 1560
cttgttttct ctggtgtctt tttcttttta cttttttaat tttttgag 1608

```

```

<210> 28
<211> 2275
<212> DNA
<213> Homo sapiens

```

10

```

<220>
<221> misc_feature
<223> Incyte ID No: 5435937CB1

```

```

<400> 28
ggctgagggg ggggggcccg gcgcttcgtc caatcagggg cgctgggagg ggcttcccga 60
gtgtccatcg gggagggcga aggcctcgct ctcctcgggt cccggctcca ggcggcgagc 120
tgaggttggg agcctggcct tcccctccga gagggttcag gtgcctctgc catagcttct 180
gtcggcctgt ctgtgacccg cactggctgt gggagtcacc tgaaaggcaa gaaatggatt 240
cagtgccctt tgaggatgtg gctgtgacct tcacccaaga ggagtgggct ttgctggatc 300
cttcccagaa aaatctctgt agagatgtga tgcaagaaac cttcaggaac ctggcctcta 360
tagggaaaaa atggaaaacc cagaacatat atgtagagta cgaaaatcta aggagaaacc 420
taagaattgt gggagagaga ctctttgaaa gtaaagaagg tcatcagcat ggagaaattt 480
tgacccaggt tcagatgac atgctgaaga aaacaactac tggagtaaaa tcatgcaaaa 540
gcagtggtga tggagaagta ggcagtgctc attcatctct taataggcac atcagagatg 600
acactggaca caaggcatal gagtatcaag aatatggaca gaaaccataa aatgtaaat 660
actgtaaaaa acctttcaac tgtctctcct ctgttcagac acatgaaagg gctcatagt 720
gaaggaact ctatgtttgt gaggatgag gaaaaacatt tatttcccat tcaaaccttc 780
aaagacacag gataatgcac cgtggagatg gacctataa gtgtaattt tgtgggaaag 840
ccttgatgtt tctcagtttg tatcttatct acaaacgaac tcacactgga gagaaacct 900
atcaatgtaa acagtggtgt aaagccttta gtcattctag tagccttoga atacatgaaa 960
gaactcacac tggggagaaag ccttataaat gtaatgaatg tgggaaagca tccatagtt 1020
ccacatgcct tcagctcat aaaaagaact acactgggga gaagccatat gaatgtaaac 1080
agtgtgggaa agccttcagc tcttcccatt cctttcaaat acatgaaaga actcacacgg 1140
gggagaagcc atatgaatgt aaggaatgtg gaaaagcatt caagtgtccc agttctgttc 1200
gcagacatga aagaacccac tctaggaaaa aaccctatga atgtaaacat tgtgggaaag 1260
tattatctta tcttaccagc tttcaaaacc acttgggaat gcacactgga gagatatctc 1320
ataaatgtaa gatatgtggg aaagcctttt attctcccag ttcacttcaa acacatgaaa 1380
aaactcacac tggagagaaa ccctataaat gcaaccaatg tggtaagacc ttaattctt 1440
ccagttcctt ccgatatcat gaaagaactc acactggaga gaaaccttac gagtgtaac 1500
aatgtgggaa agccttcaga tctgcctcac tcttcaaac acatggtagg actcacacgg 1560
gagagaaacc ctatgcatgt aaggaatgtg gaaaaccatt tagtaatttc tctttcttcc 1620
aaatacatga aaggatgcac agagaagaga agccttatga atgtaagggt tatgggaaaa 1680
cattcagttt gccagttta tttcatagac atgaaaggac tcacactgga ggaaaaacct 1740
atgaatgcaa gcagtggtgc agatccttca actgttcgag ctcctttcga tatcatgaa 1800
ggactcacac tggagagaaa ccctatgaat gcaagcaatg tggaaaagcc ttcagatctg 1860
cctcacagct tcaaatctat ggaaggactc acactggaga gaaaccttat gaatgtaagc 1920
agtgtgggaa agccttttga tctgcctcac accttcaaat gcatggaagg actcacactg 1980
gagagaaacc ctatgaatgt aagcagtgty ggaagtcttt tggatgtgcc tcgagacttc 2040
aaatgcatgg aaggactcac actggagaga aaccgtataa atgtaagcaa tgtgggaaag 2100
cttttgatg tccctcaaac cttcgaaggc atggaaggac tcacactgga gagaaacct 2160
ataaatgtaa ccaatgtggt aaagtcttta gatgttcttc acaacttcaa gtgcatgaa 2220
gggctcactg catagacacc ccataacccc aggccttagg aggctgaggt ggggg 2275

```

20

30

<210> 29

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

<220>
<221> misc_feature
<223> Incyte ID No: 7503560CB1

<220>
<221> unsure
<222> 4266
<223> a, t, c, g, or other

```

<400> 29
gggagctgca atgagcggcg cccggaggct gtgacctgcg cgcggcggcc cgaccggggc 60
ccctgaatgg cggctcgctg agggggcggc gggggcggcg gctcaggctc ctcggggct 120
ggcgtgcccgg tgaaggggtg atgctgttca agctcctgca gagacagacc tatacctgcc 180
tgtcccacag gtaaggggctc tacgtgtgct tcttgggctg cgttgtcacc atcgtctccg 240
ccttccagtt cggagaggtg gttctggaat ggagccgaga tcaataccat gttttggttg 300
attcctatag agacaatatt gctggaaagt cctttcagaa tcggccttgt ctgccatgc 360
cgattgacgt tgtttacacc tgggtgaatg gcacagatct tgaactactg aaggaactac 420
agcaggctcag agaacagatg gaggaggagc agaaagcaat gagagaaatc cttgggaaaa 480
acacaacgga actcactaag aagagtgaga agcagttaga gtggttgcta acacactgca 540
ttaagggtgcc aatgcttgtc ctggaccagc ccctgccagc caacatcacc ctgaaggacc 600
tgccatctct ttatccttct tttcattctg ccagtgcacat tttcaatggt gcaaaaccaa 660
aaaacccttc taccaatatt tcagttgttg tttttgacag tactaaggat gttgaagatg 720
cccactctgg actgcttaaa ggaaatagca gacagacagt atggaggggc tacttgacaa 780
cagataaaga agtccctgga ttagtgctaa tgcaagatth ggctttcctg agtggatttc 840
caccaacatt caaggaaaca aatcaactaa aaacaaaatt gccagaaaat ctttctctta 900
aagtcaaaact gttgcagttg tattcagagg ccagtgtagc gcttctaata ctgaataacc 960
ccaaggatth tcaagaattg aataagcaaa ctaagaagaa catgaccatt gatggaaaag 1020
aactgaccat aagtcctgca tatttattat gggatctgag cgccatcagc cagtctaagc 1080
aggatgaaga catctctgcc agtcgttttg aagataacga agaactgagg tactcattgc 1140
gatctatcga gaggcatgca ccctgggttc ggaatatttt cattgtcacc aacgggcaga 1200
ttccatcctg gctgaacctt gacaatcctc gaggtagaat agtaacacac caggatgtht 1260
ttcgaaatth gagccacttg cctaccttta gttcacctgc tattgaaagt cacattcatc 1320
gcatcgaagg gctgtcccag aagtttattt acctaaatga tgatgtcatg tttgggaagg 1380
atgtctggcc aagtgatttt tacagtcact ccaaaggcca gaaggthtat ttgacatgpc 1440
cttttggtgg aggaataaac agtgtctctt actgtaatca gggatgtgpc aattcctggc 1500
tcgctgataa gttctgtgac caagcatgca atgtctgttc ctgtgggttt gatgctggcg 1560
actgtgggca agatcatttt catgaattgt ataaagtgat ccttctccca aaccagactc 1620
actatattat tccaaaagggt gaatgcctgc cttatctcag ctttgacaga gtacccaaaa 1680
gaggagttga aggtgcctat agtgacaatc caataattcg acatgcttct attgccaaca 1740
agtggaaaac catccacctc ataatgcaca gtggaatgaa tgcccaccaca atacatttta 1800
atctcacgth tcaaaaataca aacgatgaag agttcaaat gcagataaca gtggagggtg 1860
acacaaggga gggaccaaaa ctgaattcta cagcccagaa gggttacgaa aatttagtta 1920
gtcccataac acttcttcca gaggcggaaa tcctttttga ggatattccc aaagaaaaac 1980
gcttcccga gtttaagaga catgatgtha actcaacaag gagagcccag gaagagggtg 2040
aaattccctt ggtaaaatatt tcactccttc caaaagacgc ccagttgagt ctcaatacct 2100
tggatttgca actggaacat ggagacatca ctttgaaagg atacaatttg tccaagtcag 2160
ccttgctgag atcatttctg atgaactcac agcatgctaa aataaaaaat caagctataa 2220
taacagatga aacaaatgac agtttgggtg ctccacagga aaaacaggth cataaaagca 2280
tcttgccaaa cagcttagga gtgtctgaaa gattgcagag gttgacttht cctgcagtha 2340
gtgtaaaagt gaatggatcat gaccagggtc agaatccacc cctggacttg gagaccacag 2400
caagatttag agtggaaact cacacccaaa aaacatagc cggaaatgtg acaaaagaaa 2460
agccccatc tctgattgth ccactggaaa gccagatgac aaaagaaaag aaaatcacag 2520
ggaaagaaaa agagaacagt agaatggagg aaaaatgctga aatatcata ggcgttactg 2580
aagtgttact tggagaaaag ctgcagcatt acacagatag ttacttgggc tttttgcat 2640
gggagaaaaa aaagtatttc caagatcttc tcgacgaaga agagtcattg aagacacaat 2700
tggcatactt cactgatagc aaaaatactg ggaggcaact aaaagataca tttgcagatt 2760
ccctcagata tgtaaataaa attctaataa gcaagtttgg attcacatcg cggaaagtcc 2820
ctgctcacat gcctcacatg attgaccgga ttgttatgca agaactgcaa gatatgttcc 2880
ctgaagaatt tgacaagacg tcaattcaca aagtrcgcca ttctgaggat atgcagtttg 2940

```

10

20

30

```

ccttctctta tttttattat ctcatgagtg cagtgcagcc actgaatata tctcaagtct 3000
ltgatgaagt tgatacagat caatctgggtg tcttgtctga cagagaaatc cgaacactgg 3060
ctaccagaat tcacgaactg ccgttaagtt tgcaggattt gacaggctcg gaacacatgc 3120
taataaattg tcaaaaaatg ctctctgctg atatcacgca gctaaataat attccacca 3180
ctcaggaatc ctactatgat cccaacctgc caccggtcac taaaagtcta gtaacaaact 3240
gtaaaccagt aactgacaaa atccacaaag catataagga caaaaacaaa tatagggtttg 3300
aaatcatggg agaagaagaa atcgccttta aaatgattcg taccaactgt tctcatgttg 3360
ttggccagtt ggatgacata agaaaaaac ctaggaaagt tgtttgcctg aatgacaaca 3420
ttgaccacaa tcataaagat gctcagacag tgaaggctgt tctcagggac ttctatgat 3480
ccatgttccc cataccttcc caatttgaac tgccaagaga gtatcgaaac cgtttccttc 3540
atgtgcatga gctgcaggaa tggagggtct atcgagacaa attgaagttt tggaccatt 3600
gtgtactagc aacattgatt atgtttacta tatttctcatt ttttgcctgag cagttaattg 3660
cacttaagcg gaagatattt cccagaagga ggatacacia agaagctagt cccaatcgaa 3720
tcagagtata gaagatcttc atttgaaaac catctacctc agcatttact gagcatttta 3780
aaactcagct tcacagagat gtccttgtga tgtgatgctt agcagtttgg cccgaagaag 3840
gaaaataatc agtaccatgc tgttttgtgg catgaatata gccactgac caggaaattat 3900
ttaaccaacc cactgaaaac ttgtgtgttg agcagctctg aactgatttt acttttaaag 3960
aatttgctca tggacctgtc atccttttta taaaaaggct cactgacaag agacagctgt 4020
taatttcca cagcaatcat tgcagactaa ctttattagg agaagcctat gccagctggg 4080
agtgattgct aagaggctcc agtctttgca ttccaaagcc ttttgctaaa gttttgact 4140
ttttttttt catttcccat ttttaagtag ttactaagtt aactagttat tcttgcctct 4200
gagtataacg aattgggatg tctaaaccct atttttatag atgttattta aataatgcag 4260
catttncacc tcttttt 4277

```

10

```

<210> 30
<211> 6417
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Incyte ID No: 7472378CB1

```

20

```

<400> 30
atgtcaaaaa agttcgttta caaccttagg aaaactacac gttcagttgt tggagtcca 60
cmetaacta ataggcccc acatcccgtt agacgtcctg actcccttct cccgatttcc 120
gaagaaccca atcaaatatc ttcccaaac cccaatatgg actcgggaaa cgattctgcc 180
cgcccactc ctcccctct ggcggccact gtcagtggta ttagctcctt aattcaact 240
acgttcaagc ctaaaagatc catggcattt gttgagcatt tgccaacctt tgatggta 300
cctcgtctat tggacaggtt tatcactagc gtagaagaaa tctgatgct catcagggga 360
gctgaccaaa caccgtatgg cctgcttact ctgaggacca tcaggaacaa aatcattgat 420
agggccgacg aagccttggg actggcaaat accccttgg tttgggatga gattaaga 480
aatctcatcc gcctctactc gagcaagaaa agcggagcca acttgtaag cgagcttaac 540
acattttcgg acaacctgac cttgggcaa ctgttctttg gtatatcaa ggtgagaagc 600
caactcttct ccatactcaa aaacagcgaa cacaacaaca ctgtttaga tgcaaaaaag 660
gttgtctaca acgaggtttg tctcaatgct tttatgactg gtttgaagga acctctcaag 720
actttcgtca ggataaagtc cccttctaca cttgaacagg cgtacgagca atgccaata 780
gagcagacct tatatagggc acaaaaacag cgaaccaaca gaccagagca gggaccat 840
ggatcagaca ataaaacctc ccgaaatagc tacgacagca attaccgag cggacgtaac 900
gaccgaaatg accgtagggg acctactct aactctaact ctaactctaa ctctggccea 960
aatagaccat ttaattcaca caatcgaca cccaatccg gcaccaagga caaccgggcc 1020
aatacatcaa acccctttcg agcacttca catagtttga ataatataga ggagaacct 1080
caacctgatt cgaattttca gcaaacggcc tcgggaaacc acagggctc tacactcc 1140
ttcatcgacc caaaaatagt cgacctagg aactgtgtga ccttagatac gccataaca 1200
ctcaaacag cctgaaacag ttttaaaata tatcaaacg tctctatacc atttccaccg 1260
gaattccaaa tcacgggcaa aatgacctt ctaccttca agttccactc ttattttgac 1320
ggattgatag gaatggactt attatcttac ctaaaaacag aaatagatt acttaacct 1380
aatctaaaaa cccaagtac cattataccc ttatggacc acagtaactc aacttcaaac 1440
gtatttaata tctctggaca tacgaaaact attttgcac taccagtgga aaccaaacag 1500
ggcgacttct acatcgatc aattacaatc aatgatgact taataatc agacgggat 1560
tataatgcc aaaacaatat tgctaatttc gttatcacia actatagcga gagggatcag 1620
ttattgtacc tcgagagccc gataaaaggg atgccatct ccacggccea caatgttga 1680
cttttcagta tcacttcaga caccocacag ccccaaacct ccgagcgtc gttacaagcc 1740

```

30

```

cttggcgtcg atcacctctc ctctgaagag aaacaaagcc tacttttact ttgcaaaagt 1800
tatctagaba tcttctacaa tgaagacaaa tcattgacct tcaccaacaa gattacacac 1860
acgattaana ccacggacga cacccccatt catacaaaat cttatagata tccttacatt 1920
cataaagagg aaagtcaaaa acaaatagag gcaatgttaa atcaggacat tatcaaatcc 1980
agttattccc cgtggagcgc ccccgctctgg gtctgtccca agaaaatcac tcctacggga 2040
gagcaaaaaat ggcgtctagt tatcgattat agaaaaactca acgagaagac tatatccgat 2100
agatatccaa tacctaacat cgcggatatt ttagacagat tgggcaaagc caaatatttc 2160
tcacacactg atctggcaag tggattccat cagatagaaa tgaatcccga cgacacaccc 2220
aaaactgcat ttacagtaga ggggggccac tacgagttca ttagaatgcc gtttggcctc 2280
aaaaatgccc cagccacatt ccaaaggggtg atggacaata ttttgggaga ccttatcgga 2340
actatctgct tagtttacct agatgatata ataattttct caacctcctt acaagaacac 2400
tcatacactc tgaaaactat ttttgggaaga ctcagatctg ccaactttaa agtccaactc 2460
acaaaatcct acttctctag gcgggagaca gaattccttg gccacatcgt ttcacaagaa 2520
gggtgttaggc caaatcccaa taagatcgaa gctataaaaa actttccatg tcccacagat 2580
aaaaagtc aaagtcttt cctaggcttg tgggatatt acagaaaatt tatcagagat 2640
tttgcgagac ttaccaaac catgacacaa aaattaaggg gaaacaataa atcgtacata 2700
atagatgatg aattcaaaaa ggcccttgaa tattgcaaaa ccttactgtc taacgacca 2760
atcctccaat acccggactt tacaaaaacct ttcacactaa ccacggacgc aagtaatttc 2820
gcaataggag ctgtcctatc ccaaggtccg gtgcatagtg ataggcccgt atgttttgc 2880
agtagaacct tgtcggctgc ggaaacaaat tattccacaa ttgagaagga aatgctggcc 2940
atttatatgg cgttccaata cttcagaccc tacctctttg gacggagatt cactataatc 3000
accgatcaca aaccactaac ttggttaatg aatttcaaac aaccaaattc taaaatagtt 3060
aggtggagac tccagcttca gggatcgat ttcgaagtgc tctacaagaa aggcctctca 3120
aatgtaattg ctgatgctct cagtagacca gaggcctctg tcaaccataa cgaagcccta 3180
tcaattctct aaaaatgtttg ccccatctca gagaaacccc ttaatgattt taatattcag 3240
ctcctgttca aaataacccc agatacaaat aacgccacac tgaccccggt taaacacaaa 3300
cttagggagg aattctgtta acccaatttt cagtatgacg acgtagtttg cattcttagg 3360
cagtcgttaa aaccocaaac gacatgcgcg gtatttggcc ccgaccacat ttttcaaatg 3420
gtggaacaag cctaccaaac ctactctctca gccacacgct aattttaaact cattagatgt 3480
ttgatcttcc tcccggaaat tactgatagt acggagatcg aaaaaattat aaccgactat 3540
cactataata gtaaccatcg agggatcgat gaaacatatt tacacataaa acgacaacag 3600
ttcttcccac atatgaagga gagaataact cagttaattc gaaaatgtga aacatgttta 3660
aaattaaat acgacagaca acctcaaaaag atcacttacc aaatatccga actaccttca 3720
aaaccgttgg acatcttaca tatagacatt tatactatta acaaaaatta taaccttact 3780
attatcgata aattttctaa atttgcggct gcctaccta taactaatag gaattgcatt 3840
aacgtagtta aagccttaaa acatltcatt tcccatttg gtattcccaa aaagctgatc 3900
tatgatcagg gagcagaatt cgctagcgat atgttcaata agttctgcac tcaatttaac 3960
attgacctac acgttacgtc ctttcaacaa tccctctagta actctcccg tgaacggctt 4020
cactcgacac taactgagat ttacagaata atacttgacg tcaggaacaa acagaaactc 4080
agtagcagac atgacgagat aatgtccgaa accctaatac catataataa cgctattcat 4140
tctgcaacta aacatacccc ctttgaacta ttttaacggac gtactcatab attcaaccaa 4200
acaatccagt tcaataacga acacgactac ttaacgaaat taaatgaatt tccgagaag 4260
ttgtaccccc tcatcacgga caaactttca aatgacgtag ttagggagaac cctaaaatta 4320
aatgaaaccc gaacagaccc cgtagacctt caaccagaca ctttagtcoct taggaaggaa 4380
aacagacgta ataagattac acccaggttt tcgattcaca aagtcaaaac cgacaaaggt 4440
catacattga taactgctag gaatcaaaaa ctacacaaat caaaaattcg aaaaacagtt 4500
ttgaaaaaag acaaaagcaa caacgctatc catgtccatt atttaaata taacgcccc 4560
atagccaaga tagaactagg gaaagcctta ctaattgaga ggtacaaaat aattagtc 4620
gtaatcaacc tacaagacta cagcagatgt atggaacaat tccatctgac cattaataa 4680
tttaaccccg attccacggt gacggactcc gtcacaattt taaaaaccaa attaacccaa 4740
gcccaagtaa agctcaaaag ccttacacct tcatatagaa acaaacgggg tttgattaac 4800
ggattgggga gtctagtaaa ggtggttacc ggcaacatgg atgccaaag caataaagaa 4860
atcacatgaag aacttgacaa tataaagaaa aattccgaag tcagtaacga caatctccaa 4920
aaacaagtaa tgtttaacaa cgaaataact atccggttcg aaaaatcac ggaccatata 4980
aataatgaac aaattttgat aagtaaatcc tttgatacct cacaaaacaa aatatacaaa 5040
cacttaactt tacaagatac ccttctggaa gaaatacaat atttaaata gattaattat 5100
aacatagaat tattcattaa ccacctaacc gacataacag aaagtatgct attggcgaaa 5160
ataaatataa ttcccaggtt catcctaact gaacaagaaa tggataaaat aaaaacaata 5220
ctggaaaaaac aaaatatcac agtcaaaaat gaacaaagta tatacaattt cctacaaaatg 5280
aatacactaa attacgaaca aaagattatt ttaataatca aagtcccaat ttttaacaa 5340
ccttttcata cctcggccag actagttcca ttaccaataa ataacacata ttttghtaata 5400
accccaattt acctagctta taatattaat aataagaaat ttcatatgac ccgtaaatgc 5460
cccaactggg ataatacatt cttgtgcgac gagaacttct acgttgatac accacagaac 5520

```

10

20

30

```

aacacatgcc tggaacacct tttgaacgga gaaaacagtt cctgcgatgt acgggaaacc 5580
ggccccatca cggacgtggt cgaggcagag agaggttaca tcttcgcatt caacgtgaac 5640
aaactgaagg tatccctaac aaacggctcc gagctctcaa taatggggtc agccatcatc 5700
agatacatta acgaaacaat acagattaac ggtatcgatt acgacggcac ggttgacacg 5760
ttccctgaac agacggatct tgatcttccc cccatgcgaa aagtaactag gaataccact 5820
attacggtac taagcctaga aaaactgcac ctggaagcca cccaaacaat ggataaaatc 5880
ctggccgtcc atcacaatac tatacagcac acctggacac tctacactct gctcggattg 5940
gtaacgttcc tagcagtcac cttatggctg caccgacgaa cgaaacacat cgtccacatc 6000
cacgaggatc atcacgtttt gaccccaatt gcaaaaagta aaggtaaggt attaattcgc 6060
gataagctca tggagcggcg taaccgtcgc acagaaagga cagagaaagc gcgatctg 6120
gaagtgaagg acagaacggc caggaccctgg attggggagg cggttgcgcg cgtcgtctgc 6180
gacgggtgga cgttctctgt tccggtcaca ccacatacgt tccgccatc cttatgcgatg 6240
cacatgctgt atgcccgtat accgctgaaa gttctgcaa gcctgatggg acataagtcc 6300
atcagttcaa cggaggtcta cacgaagggt tttgcgctgg atgtggctgc ccggcaccgg 6360
gtgcagtttg cgtgcccgga gtctgatgcg gttgcgatgc tgaacaatt atcctga 6417

```

```

<210> 31
<211> 1194
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Incyte ID No: 1306049CB1

```

```

<400> 31
gcccgcgcgc cggagggagg agcggggcgc gggggccggc tggcgcgggg gctccgacc 60
tgcccggctt ggcgatggag tttccggacc tgggcgctca ctgttcggag ccgagctgtc 120
agcgccttga tttctgcgc ctttaagtgt atgcctgctc aggcattctc tgcgcagacc 180
atgtggccta cggccagcat cactgtggat ctgcttacca aaaggatata caggtacctg 240
tgtgccctct ctgtaatgtg cctgtgcctg tggccagagg ggagccccc gaccgtgctg 300
tgggagagca caltgacaga gactgtcgc ctgatccagc acagcaaaaa cgtaaatctc 360
tcaccaataa gtgtgaacgc gctggtcgc ggcagcagaa aatgatgaaa ctgacctgtg 420
aacgctgtag ccgaaacttc tgcatacagc accggcatcc actggaccat gattgctctg 480
gggaggggca cccaaccagc cgggcaggac ttgctgccat ctccagagca caagctgtg 540
cttctacaag cactgtcccc agcccagtc aaaccatgcc ttctgttacc tctcccagca 600
gagccacaac ccgatctccg tccctggacag cccctccagt gattgctttg cagaatggcc 660
tgagtgaagg tgaagctctg cagcgggccc tggaaatgtc cctggcagaa accaaacccc 720
aggttccaag ttgtcaggag gaagaagacc tagctttagc acaagcactg tcagccagtg 780
aggcagaata ccagcggcag caggcccaga gccgcagctc gaagccctcc aactgcagcc 840
tgtgctaggg cctcgggctt ggggagggag gttcacctga ggaggactgt ggcctcaca 900
cctctagggg acacagggag aggagggccc gagcacctcg gagggcagag acaagcggga 960
gtgatgtgga ggtcgcctcg ggagcctctg gaaggccttg ctagtctctc agctgcatgg 1020
aagagagcgg ctgaccaactg ttccctgggt gggccctcag tggatgctgg ccaggcccta 1080
ctcttagccc cttcatcatg tcatctccct tatgctggag ctgcccgatg gttgagtggt 1140
caggaagggg cctggaaaaa ataaaggatc ttggcagttg ataaaaaaa aaaa 1194

```

```

<210> 32
<211> 2801
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Incyte ID No: 3187174CB1

```

```

<400> 32
accttcttcc cctgaagtgg ctggggttcc tgtttctctc tttgattgac aacttgtgtt 60
aacccctcga catctctggy ccaatttttg cttgaaaatg gcagctcccg agcagccgct 120
tgcgataatc aggggatgca cgagctcctc ctgcctttcc ccgcctcggg gcgaccgaac 180
ccttctggtc aggcacctgc cggctgagct tactgctgag gagaagagg acttgctgaa 240
gtacttcggg gctcagctctg tgcgggtcct gtcagataag gggcgactga aacatacagc 300
ttttgccaca ttccctaagt aaaaagcagc tataaaggca ttgacaagac tccatcaact 360

```

10

20

30

```

gaaactttta ggtcatactt tagtcgttga atttgcaaaa gagcaagatc gagttcactc 420
cccatgtccc acttcagggtt ctgaaaaaaa aaaagggtct gatgaccctg tcgaagatga 480
taagaaaaaa aaagaacttg gttatттаac agtagaaaaat ggaattgcac caaacocatgg 540
gctgactttt cttttaaatt catgcctcaa gtatatgtac ccaccacctt ccagcacaat 600
cctagcaaac attgtaaaatg ccttggcaag cgtgcctaag ttctatgtac aggtccttca 660
tcttatgaat aaaatgaatt tgcccacacc ttttgacca attactgcgc gacctcccat 720
gtatgaagac tatatgccat tgcatgcacc tcttccacc acatctcctc agccacctga 780
ggaacctcct ttgccagacc aggatgagga attatctagt gaagaatcag aatatgaaag 840
cactgatgat gaggaccgac agagaatgaa caaatтаatg gaactagcaa atcttcagcc 900
caaaagacct aaaacaataa agcagcgcca tgtgagaaaa aaaaaaaaaa aaaaggatat 960
gttgaataca cctttgtgtc cttcacacag cagtttcaat ccagtgtctg taccttcaga 1020
tgtatttgac caaccacaac ctgtaggtaa caaaagaatt gaatccata tatctaccga 1080
catgccagct gcatttaaga aagatttaga aaaggaaaca aattgtgagg aaaaaaatca 1140
tgatttacct gctactgaag ttgatgcac caatatagga tttggaaaaa tcttcccaa 1200
agctaatttg gacatcacag aggagattaa agaagactct gatgaaatgc cttcagaatg 1260
tatttctaga agggaattgg aaaaggcgag aatttctaga gaagaaatgg aaacctttc 1320
agttttcaga agttatgaac cgggtgaacc aaactgtaga atttatgtaa agaatttagc 1380
taaacatggt caagaaaagg accttaataa tatttttgya agatattgtg acttttcac 1440
agaaacacag cggatcatgt ttgatatacg tttgatgaaa gaaggtcgtg tgaaaggaca 1500
agctttcatt ggacttccta atgaaaaagc agcagcaaaa gccttaaagg aagctaattg 1560
atatgtgctt tttggaaaaa ccatggtggt tcagtttgct cgatctgcta gaccaaaaca 1620
agatcctaag gaaggaaaaa gaaagtgtta aaaattaata aagaagcatt octgcgtaaa 1680
tactgctgta atactgtcat gcaaagtgtg tcctttcttg tcgtatcctt tttggggcag 1740
tgtttttttg ttttttttct agaaaatgttt gtcttctccc cacctgttga tccaggttaa 1800
ggaatacttt tttacacttt attcaaatga aatatttcta aaattatttg tatagactga 1860
acagatcttt tatgtgtttt tagatttgtt gttgaatttt ctgtgctgtc ctttatataa 1920
ttttttgagg gaaagttagt gaatcaggtc aacttactta gagaatgtgt tcatttactt 1980
taaccagaa tacagtcttg tttctctat ttgtatgttt cetaaaccta attcaataac 2040
atatgctttc tgtttgttaa tatatctggt ttaggtattt ataatgtgtt taaaatttgg 2100
gcaaagggaa tgtttttctt ttaaaaagta cttacattga aaattaagat gtctggatta 2160
ctatgtaaat tctagagagt agcagacctc tcatctgaag tcttagtgaa tctcttttga 2220
catagatagc aatagaagta tctttctctc ttccccttc tttttctaaa caagagaaga 2280
aaagcgtaat agaggggaga acacataatg cccactaagg gtatgtcatt aaggaaaaac 2340
agtcttggca ggtatatagg aatagtggtt tccagactgg ttgatgaccg taatcaccaa 2400
gaacagtggg tctcagctct ggctgcacat tgcagtgatc tggaaactta atactaattc 2460
taaaagggtg cagtggctca tacctgtaat ccagcactt tgcaagtccg agatgggaga 2520
atcacttgag ccagaggatt tgagaccagc ctgggcaatg tagggagacc ctgtccctac 2580
aaaaataca aaaattagcc aagtgtggtg gcttgcacct ctggtctcag ctacttggga 2640
tgctaaggca ggaggattac ttgagcccca gaggttgagg ttgcagtgaa ccatgatcac 2700
accactgcat tctagcctgg gtgacagagt gagacctct cctcctaaa aaaatcctta 2760
agaaatatat tgatgcttgg ttcctttggt cagaattttg a 2801

```

10

20

【国際調査報告】

INTERNATIONAL SEARCH REPORT		International Application No PCT/US 01/50256
A. CLASSIFICATION OF SUBJECT MATTER IPC 7 C12N15/12 C07K14/47 A01K67/027 C12N5/10 C07K16/18 C12Q1/68 G01N33/68		
According to International Patent Classification (IPC) or to both national classification and IPC		
B. FIELDS SEARCHED Minimum documentation searched (classification system followed by classification symbols) IPC 7 C07K		
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, BIOSIS, EMBL, SEQUENCE SEARCH, MEDLINE, WPI Data, PAJ		
C. 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! 29 June 1998 (1998-06-29) BIRD C: "Human DNA sequence from clone 1189B24" Database accession no. AL030996 XP002221027 the whole document	1-20, 23, 26-56, 72
X	-& DATABASE SWALL 'Online! 1 November 1998 (1998-11-01) BIRD C: "dJ1189B24.4 (novel putative protein similar to hypothetical proteins S. pombe C22F3.14C and C. elegans C16A3.8) (Fragment)" Database accession no. 075984 XP002221028 the whole document	1-20, 23, 26-56, 72
X	-& WEST R W ET AL: "RLR1 (TH02), required for expressing lacZ fusions in yeast, is conserved from yeast to humans and is a -/--	1-20, 23, 26-56, 72
<input checked="" type="checkbox"/> Further documents are listed in the continuation of box C. <input checked="" type="checkbox"/> Patent family members are listed in annex.		
* Special categories of cited documents : "A" document defining the general state of the art which is not considered to be of particular relevance "E" earlier document but published on or after the international filing date "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified) "O" document referring to an oral disclosure, use, exhibition or other means "P" document published prior to the international filing date but later than the priority date claimed "T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art. "8" document member of the same patent family		
Date of the actual completion of the international search 19 November 2002		Date of mailing of the international search report 11. 04. 03
Name and mailing address of the ISA European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016		Authorized officer Morawetz, R

INTERNATIONAL SEARCH REPORT

International Application No.
PCT/US 01/50256

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT		
Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
	<p>suppressor of SIN4" GENE, ELSEVIER BIOMEDICAL PRESS. AMSTERDAM, NL, vol. 243, no. 1-2, February 2000 (2000-02), pages 195-205, XP004187689 the whole document ---</p>	
X	<p>DATABASE EMBL 'Online! 22 February 2000 (2000-02-22) STRAUSBERG R.: "Homo sapiens cDNA FLJ10896 fis, clone NT2RP5003461, weakly similar to RLR1 PROTEIN" Database accession no. AK001758 XP002221029 the whole document ---</p>	1-20,23, 26-56,72
X	<p>DATABASE EMBL 'Online! 29 September 2000 (2000-09-29) ISOGAI T. ET AL.: "Homo sapiens cDNA FLJ13597 fis, clone PLACE1009798, weakly similar to RLR1 PROTEIN" Database accession no. AK023659 XP002221030 the whole document ---</p>	1-20,23, 26-56,72
X	<p>WO 00 50593 A (YOO TAI JUNE ;UNIV TENNESSEE AT MEMPHIS RESE (US)) 31 August 2000 (2000-08-31) see SEQ ID NO: 13 ---</p>	1-20,23, 26-56,72
A	<p>JANSA ET AL: "cloning and functional characterization of PTRF, a novel protein which induces dissociation of paused ternary transcription complexes" EMBO JOURNAL, OXFORD UNIVERSITY PRESS, SURREY, GB, vol. 17, no. 10, 15 May 1998 (1998-05-15), pages 2855-2864, XP002129532 ISSN: 0261-4189 the whole document ---</p>	
	-/--	

INTERNATIONAL SEARCH REPORT

International Application No
PCT/US 01/50256

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT		
Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
P,X	<p>DATABASE GENESEQ 'Online! 9 March 2001 (2001-03-09) ROSEN C.A., ET AL.,: "Human pancreatic cancer antigen nucleotide sequence SEQ ID NO:341." Database accession no. AAC99113 XP002221031 the whole document</p>	1-20,23, 26-56,72
P,X	<p>-& DATABASE GENESEQ 'Online! 9 March 2001 (2001-03-09) ROSEN C.A., ET AL.,: "Human pancreatic cancer antigen protein sequence SEQ ID NO:800." Database accession no. AAB54348 XP002221032 the whole document</p>	1-20,23, 26-56,72
X	<p>& WO 00 55320 A (HUMAN GENOME SCIENCES INC ;ROSEN CRAIG A (US); RUBEN STEVEN M (US)) 21 September 2000 (2000-09-21) SEQ ID NO: 800</p>	1-20,23, 26-56,72
P,X	<p>WO 01 57190 A (CAO YICHENG ;CHEN RUI HONG (US); GOODRICH RYLE (US); HYSEQ INC (US)) 9 August 2001 (2001-08-09) SEQ ID NO: 2109; SEQ ID NO: 3093</p>	1-20,23, 26-56,72
T	<p>STRASSER KATJA ET AL: "TRES is a conserved complex coupling transcription with messenger RNA export." NATURE (LONDON), vol. 417, no. 6886, 2002, pages 304-308, XP002221026 16 May, 2002 ISSN: 0028-0836</p>	

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

INTERNATIONAL SEARCH REPORT

international application No. PCT/US 01/50256
--

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:
Although claim 19 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 compound/composition.
2. Claims Nos.: 21,22,24,25
because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:
see FURTHER INFORMATION sheet PCT/ISA/210
3. Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

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

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

see additional sheet

1. As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.
2. As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3. As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:
4. No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:
1-55 (all partially), 56 and 72 (all completely)

Remark on Protest

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

INTERNATIONAL SEARCH REPORT

International Application No. PCT/US 01 50256

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

Continuation of Box I.2

Claims Nos.: 21,22,24,25

Present claims 21 and 24 relate to a composition comprising a compound defined by reference to a desirable characteristic or property, namely an agonist compound and an antagonist compound, respectively. The claims cover all compounds having this characteristic or property, whereas the application provides support within the meaning of Article 6 PCT and/or disclosure within the meaning of Article 5 PCT for none of such compounds. In the present case, the claims so lack support, and the application so lacks disclosure, that a meaningful search over the whole of the claimed scope is impossible. Independent of the above reasoning, the claims also lack clarity (Article 6 PCT). An attempt is made to define the compound by reference to a result to be achieved. Again, this lack of clarity in the present case is such as to render a meaningful search over the whole of the claimed scope impossible. Claims 22 and 25 relate to the use of the compositions of claims 21 and 24, respectively and the same objections as set out above, apply.

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

INTERNATIONAL SEARCH REPORT

International Application No. PCT/US 01 50256

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

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

Invention 1: claims 1-55 (all partially),
56 and 72 (all completely)

An isolated polypeptide comprising the amino acid sequence of SEQ ID NO: 1; and subject-matter related thereto.

Inventions 2 - 16: claims 1-55 (all partially); 57
and 73 (completely) - 71 and 87 (completely)

An isolated polypeptide comprising the amino acid sequence of SEQ ID NO: 2 - SEQ ID NO: 16; and subject-matter related thereto.

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No.

PCT/US 01/50256

Patent document cited in search report		Publication date	Patent family member(s)	Publication date
WO 0050593	A	31-08-2000	AU 3605700 A	14-09-2000
			EP 1157107 A1	28-11-2001
			WO 0050593 A1	31-08-2000
WO 0157190	A	09-08-2001	US 2002128187 A1	12-09-2002
			AU 2591801 A	31-07-2001
			AU 3128801 A	14-08-2001
			AU 3297101 A	07-08-2001
			AU 3300301 A	07-08-2001
			AU 3329301 A	14-08-2001
			AU 3484701 A	14-08-2001
			AU 3484801 A	14-08-2001
			AU 3486501 A	14-08-2001
			AU 3494401 A	14-08-2001
			AU 3665801 A	14-08-2001
			AU 3666001 A	14-08-2001
			AU 3666301 A	14-08-2001
			AU 3672101 A	14-08-2001
			AU 4314201 A	14-08-2001
			EP 1261735 A2	04-12-2002
			EP 1254265 A1	06-11-2002
			EP 1254266 A1	06-11-2002
			EP 1254268 A1	06-11-2002
			EP 1254269 A1	06-11-2002
			EP 1254270 A1	06-11-2002
			WO 0153326 A1	26-07-2001
			WO 0155334 A2	02-08-2001
			WO 0155335 A2	02-08-2001
			WO 0157255 A1	09-08-2001
			WO 0157260 A1	09-08-2001
			WO 0157175 A2	09-08-2001
			WO 0157261 A1	09-08-2001
			WO 0157262 A1	09-08-2001
			WO 0157187 A2	09-08-2001
			WO 0157265 A1	09-08-2001
			WO 0157188 A2	09-08-2001
			WO 0157266 A1	09-08-2001
WO 0157267 A1	09-08-2001			
WO 0157190 A2	09-08-2001			
US 2002137044 A1	26-09-2002			
US 2002127199 A1	12-09-2002			

フロントページの続き

(51) Int.Cl. ⁷	F I	テーマコード(参考)
A 6 1 P 3/10	A 6 1 P 3/10	4 B 0 6 5
A 6 1 P 9/10	A 6 1 P 9/10	1 0 1
A 6 1 P 9/12	A 6 1 P 9/12	4 C 0 8 4
A 6 1 P 15/00	A 6 1 P 15/00	4 C 0 8 5
A 6 1 P 25/00	A 6 1 P 25/00	4 H 0 4 5
A 6 1 P 29/00	A 6 1 P 29/00	
A 6 1 P 31/00	A 6 1 P 31/00	
A 6 1 P 35/00	A 6 1 P 35/00	
A 6 1 P 35/02	A 6 1 P 35/02	
A 6 1 P 37/06	A 6 1 P 37/06	
A 6 1 P 43/00	A 6 1 P 43/00	1 0 5
C 0 7 K 1/22	A 6 1 P 43/00	1 1 1
C 0 7 K 14/47	C 0 7 K 1/22	
C 0 7 K 16/18	C 0 7 K 14/47	
C 1 2 M 1/00	C 0 7 K 16/18	
C 1 2 N 1/15	C 1 2 M 1/00	A
C 1 2 N 1/19	C 1 2 N 1/15	
C 1 2 N 5/10	C 1 2 N 1/19	
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 37/00	G 0 1 N 37/00	1 0 2
	C 1 2 N 5/00	A
	C 1 2 N 15/00	C
	C 1 2 N 15/00	F
	A 6 1 K 37/02	

(31)優先権主張番号 60/263,823

(32)優先日 平成13年1月23日(2001.1.23)

(33)優先権主張国 米国(US)

(31)優先権主張番号 60/266,088

(32)優先日 平成13年2月2日(2001.2.2)

(33)優先権主張国 米国(US)

(31)優先権主張番号 60/348,442

(32)優先日 平成13年10月29日(2001.10.29)

(33)優先権主張国 米国(US)

(81)指定国 AP(GH,GM,KE,LS,MW,MZ,SD,SL,SZ,TZ,UG,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,R O,RU,SD,SE,SG,SI,SK,SL,TJ,TM,TR,TT,TZ,UA,UG,US,UZ,VN,YU,ZA,ZW

- (72)発明者 リュ、ヤン
アメリカ合衆国カリフォルニア州9 4 3 0 3・パロアルト・コリーナウェイ 3 8 8 5
- (72)発明者 アービズ、チャンドラ
アメリカ合衆国カリフォルニア州9 4 0 2 5・メンロパーク・# 1・シャーウッドウェイ 4 9 0
- (72)発明者 ランクマール、ジャヤラクシミ
アメリカ合衆国カリフォルニア州9 4 5 5 5・フレモント・メイバードサークル 3 4 3 5 9
- (72)発明者 ヤオ、モニック・ジー
アメリカ合衆国インディアナ州4 6 0 3 3・カーメル・ウッドゲートドライブ 1 1 8 9
- (72)発明者 ポリッキー、ジェニファー・エル
アメリカ合衆国カリフォルニア州9 5 1 1 8・サンノゼ・ジャービスコート 1 5 1 1
- (72)発明者 チョーラ、ナリンダー・ケイ
アメリカ合衆国カリフォルニア州9 4 5 8 7・ユニオンシティ・# 7 1 2・ユニオンスクエア 3
3
- (72)発明者 トリボレー、キャサリン・エム
アメリカ合衆国カリフォルニア州9 4 1 0 7・サンフランシスコ・# 5・テネシーストリート 1
1 2 1
- (72)発明者 ユエ、ヘンリー
アメリカ合衆国カリフォルニア州9 4 0 8 7・サニーベイル・ルイスアベニュー 8 2 6
- (72)発明者 バトラ、サジーブ
アメリカ合衆国カリフォルニア州9 4 0 8 7・サニーベイル・# 7 0 9・エルカミーノリアル 5
5
- (72)発明者 ディング、リー
アメリカ合衆国カリフォルニア州9 4 3 0 6・パロアルト・# 1 4 6・アルマストリート 3 3 5
3
- (72)発明者 ラル、プリーティ・ジー
アメリカ合衆国カリフォルニア州9 5 0 5 6・サンタクララ・ピーオーボックス 5 1 4 2
- (72)発明者 ボロースキー、マーク・エル
アメリカ合衆国カリフォルニア州9 4 0 6 1・レッドウッドシティ・オーチャードアベニュー 1
2 2
- (72)発明者 リュ、デュング・アイナ・エム
アメリカ合衆国カリフォルニア州9 5 1 2 3・サンノゼ・コイドライブ 2 3 3
- (72)発明者 ガンディー、アミーナ・アール
アメリカ合衆国カリフォルニア州9 4 1 1 8・サンフランシスコ・フィフスアベニュー 7 0 5
- (72)発明者 グリフィン、ジェニファー・エイ
アメリカ合衆国カリフォルニア州9 4 5 5 5・フレモント・メローウェイ 3 3 6 9 1
- (72)発明者 スー、ユーミング
アメリカ合衆国カリフォルニア州9 4 0 4 0・マウンテンビュー・ウォルナットドライブ 1 7 3
9
- (72)発明者 アジムザイ、ヤルダ
アメリカ合衆国カリフォルニア州9 4 5 2 2・カストロバレー・ボールダーキャニオンドライブ
5 5 1 8
- (72)発明者 ギエツェン、キンバリー・ジェイ
アメリカ合衆国カリフォルニア州9 5 1 2 3・サンノゼ・ロスウエコスドライブ 6 9 1
- (72)発明者 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 ワレン、ブリジット・エイ
アメリカ合衆国カリフォルニア州9 5 0 1 4・クーペルティノー・# 2・パークウッドドライブ
1 0 1 3 0
- (72)発明者 メイソン、パトリシア・エム
アメリカ合衆国カリフォルニア州9 5 0 1 4・モーガンヒル・クラークレーン 3 6 0

- (72)発明者 バーフォード、ニール
アメリカ合衆国コネチカット州06422・ダラム・ワイルドウッドサークル 105
- (72)発明者 ハファリア、エープリル、ジェイ・、エイ
アメリカ合衆国カリフォルニア州95054・サンタクララ・コーレデプリマベータ 2227
- (72)発明者 リー、アーンステーション・エイ
アメリカ合衆国カリフォルニア州94706・アルパニー・ケインズストリート 624
- (72)発明者 ヤング、ジュンミン
アメリカ合衆国カリフォルニア州95129・サンノゼ・パークレーン 7125
- (72)発明者 ゴーバッド、アン・イー
アメリカ合衆国カリフォルニア州94550・リバモア・マリーコモン 369
- (72)発明者 エマーリング、ブルック・エム
アメリカ合衆国カリフォルニア州94303・パロアルト・#71・ウッドランドアベニュー 1735
- (72)発明者 マーキス、ジョーゼフ・ピー
アメリカ合衆国カリフォルニア州95135・サンノゼ・モーリンレーン 3299
- (72)発明者 リー、ソー・ユーン
アメリカ合衆国カリフォルニア州94015・ダリーシティ・ウエストデールアベニュー 40
- (72)発明者 スウォーナカール、アニータ
アメリカ合衆国カリフォルニア州94122・サンフランシスコ・#5ディー・ロックスリーアベニュー 8
- (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
- (72)発明者 ジヤング、シン
アメリカ合衆国カリフォルニア州95070・サラトガ・エルバアベニュー 14371
- (72)発明者 ジャクソン、アラン・エイ
アメリカ合衆国カリフォルニア州95032・ロスガトス・エルウッドドライブ 1541

F ターム(参考) 2G045 AA25 AA40 BB03 BB20 CB01 CB21 DA12 DA13 DA14 DA36
DA37 DA77 FB02 FB03
4B024 AA01 AA11 BA44 BA80 CA04 CA09 DA02 DA06 EA02 FA03
FA04 GA11 HA12 HA15
4B029 AA23 BB20 CC03 CC08
4B063 QA18 QA19 QQ42 QQ52 QQ79 QR55 QR59 QR62 QR69 QR80
QR82 QS24 QS25 QS28 QS34
4B064 AG01 AG27 CA10 CA19 CA20 CC24 CE12 DA01 DA13
4B065 AA26X AA90X AA93Y AB01 AC14 CA24 CA44 CA46
4C084 AA02 AA17 BA01 BA08 BA22 CA18 CA53 DC50 NA14 ZA01
ZA42 ZA45 ZA70 ZA81 ZB08 ZB11 ZB21 ZB26 ZB27 ZB31
ZC02 ZC19 ZC20 ZC35 ZC41 ZC42
4C085 AA13 AA14 AA16 AA19 BB07 BB11 BB22 BB41 BB43 CC07
CC08 CC21 CC31 DD62 EE01
4H045 AA10 AA11 AA20 AA30 BA10 CA40 DA76 DA86 EA21 EA22
EA28 EA29 EA50 FA72 FA74 GA26

專利名称(译)	核酸相关蛋白		
公开(公告)号	JP2005508600A	公开(公告)日	2005-04-07
申请号	JP2002552156	申请日	2001-12-19
[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	ボーグンマライアアール リュヤン アービズチャンドラ ランクマールジャヤラクシミ ヤオモニークジー ポリッキージェニファーエル チョーラナリンダーケイ トリボレーキャサリーンエム ユエヘンリー バトラサジープ デイングリー ラルプリーティジー ボロースキーマークエル リュデュングアイナエム ガンディーアミーナアール グリフィンジェニファーエイ スーユーミング アジムザイヤルダ ギエツエンキンバリージェイ タングワイトム ワレンブリジットエイ メイソンパトリシアエム バーフォードニール ハファリアエープリルジェイエイ リーアーンステイーンエイ ヤングジュンミング ゴーバッドアンイー エマーリングブルックエム マーキスジョーゼフピー リーソーユーン スウォーナカールアニータ レディルーパ ジアングシン ジャクソンアランエイ		
发明人	ボーグン、マライア・アール リュ、ヤン アービズ、チャンドラ ランクマール、ジャヤラクシミ ヤオ、モニーク・ジー ポリッキー、ジェニファー・エル チョーラ、ナリンダー・ケイ トリボレー、キャサリーン・エム ユエ、ヘンリー バトラ、サジープ		

ディング、リー
 ラル、プリーティ・ジー
 ボロースキー、マーク・エル
 リュ、デュング・アイナ・エム
 ガンディー、アミーナ・アール
 グリフィン、ジェニファー・エイ
 スー、ユーミング
 アジムザイ、ヤルダ
 ギエツェン、キンバリー・ジェイ
 タング、ワイ・トム
 ワレン、ブリジット・エイ
 メイソン、パトリシア・エム
 バーフォード、ニール
 ハファリア、エープリル、ジェイ、エイ
 リー、アーンステイーン・エイ
 ヤング、ジュンミング
 ゴーバッド、アン・イー
 エマーリング、ブルック・エム
 マーキス、ジョーゼフ・ピー
 リー、ソー・ユーン
 スウォーナカール、アニータ
 レディ、ルーパ
 ジアング、シン
 ジャクソン、アラン・エイ

IPC分類号	G01N33/50 A61K38/00 A61K39/395 A61K45/00 A61P3/04 A61P3/10 A61P9/10 A61P9/12 A61P15/00 A61P25/00 A61P29/00 A61P31/00 A61P35/00 A61P35/02 A61P37/06 A61P43/00 C07K1/22 C07K14/47 C07K16/18 C12M1/00 C12N1/15 C12N1/19 C12N5/10 C12N15/02 C12N15/09 C12N15/12 C12P21/02 C12P21/08 C12Q1/68 G01N33/15 G01N33/53 G01N33/68 G01N37/00
CPC分類号	A01K2217/05 A61K38/00 A61P3/04 A61P3/10 A61P9/10 A61P9/12 A61P15/00 A61P25/00 A61P29/00 A61P31/00 A61P35/00 A61P35/02 A61P37/06 A61P43/00 C07K14/47 G01N33/6875 G01N2500/00 Y02A50/58
FI分類号	C12N15/00.ZNA.A A61K39/395.D A61K39/395.N A61K45/00 A61P3/04 A61P3/10 A61P9/10.101 A61P9/12 A61P15/00 A61P25/00 A61P29/00 A61P31/00 A61P35/00 A61P35/02 A61P37/06 A61P43/00.105 A61P43/00.111 C07K1/22 C07K14/47 C07K16/18 C12M1/00.A C12N1/15 C12N1/19 C12P21/02.C C12P21/08 C12Q1/68.A G01N33/15.Z G01N33/50.Z G01N33/53.D G01N33/53.M G01N37/00.102 C12N5/00.A C12N15/00.C C12N15/00.F A61K37/02
F-TERM分類号	2G045/AA25 2G045/AA40 2G045/BB03 2G045/BB20 2G045/CB01 2G045/CB21 2G045/DA12 2G045/DA13 2G045/DA14 2G045/DA36 2G045/DA37 2G045/DA77 2G045/FB02 2G045/FB03 4B024/AA01 4B024/AA11 4B024/BA44 4B024/BA80 4B024/CA04 4B024/CA09 4B024/DA02 4B024/DA06 4B024/EA02 4B024/FA03 4B024/FA04 4B024/GA11 4B024/HA12 4B024/HA15 4B029/AA23 4B029/BB20 4B029/CC03 4B029/CC08 4B063/QA18 4B063/QA19 4B063/QQ42 4B063/QQ52 4B063/QQ79 4B063/QR55 4B063/QR59 4B063/QR62 4B063/QR69 4B063/QR80 4B063/QR82 4B063/QS24 4B063/QS25 4B063/QS28 4B063/QS34 4B064/AG01 4B064/AG27 4B064/CA10 4B064/CA19 4B064/CA20 4B064/CC24 4B064/CE12 4B064/DA01 4B064/DA13 4B065/AA26X 4B065/AA90X 4B065/AA93Y 4B065/AB01 4B065/AC14 4B065/CA24 4B065/CA44 4B065/CA46 4C084/AA02 4C084/AA17 4C084/BA01 4C084/BA08 4C084/BA22 4C084/CA18 4C084/CA53 4C084/DC50 4C084/NA14 4C084/ZA01 4C084/ZA42 4C084/ZA45 4C084/ZA70 4C084/ZA81 4C084/ZB08 4C084/ZB11 4C084/ZB21 4C084/ZB26 4C084/ZB27 4C084/ZB31 4C084/ZC02 4C084/ZC19 4C084/ZC20 4C084/ZC35 4C084/ZC41 4C084/ZC42 4C085/AA13 4C085/AA14 4C085/AA16 4C085/AA19 4C085/BB07 4C085/BB11 4C085/BB22 4C085/BB41 4C085/BB43 4C085/CC07 4C085/CC08 4C085/CC21 4C085/CC31 4C085/DD62 4C085/EE01 4H045/AA10 4H045/AA11 4H045/AA20 4H045/AA30 4H045/BA10 4H045/CA40 4H045/DA76 4H045/DA86 4H045/EA21 4H045/EA22 4H045/EA28 4H045/EA29 4H045/EA50 4H045/FA72 4H045/FA74 4H045/GA26

優先権 60/257714 2000-12-21 US
 60/260081 2001-01-05 US
 60/262302 2001-01-16 US
 60/263823 2001-01-23 US
 60/266088 2001-02-02 US
 60/348442 2001-10-29 US

外部リンク Espacenet

摘要(译)

本发明提供了人核酸相关蛋白 (NAAP) 和鉴定和编码NAAP的多核苷酸。本发明还提供了表达载体，宿主细胞，抗体，激动剂和拮抗剂。本发明还提供了诊断，治疗或预防与NAAP异常表达相关的疾病的方法。

表1

Incipit プロトコル セグメント ID	請求項 ID NO.	Incipit プロトコル ID NO.	請求項 ID NO.	Incipit プロトコル ID NO.	請求項 ID NO.	Incipit プロトコル ID NO.	請求項 ID NO.
2500775	1	2500775	17	2500775	17	2500775	17
926296	2	926296	18	926296	18	926296	18
1322761	3	1322761	19	1322761	19	1322761	19
7472664	4	7472664	20	7472664	20	7472664	20
7473124	5	7473124	21	7473124	21	7473124	21
7473171	6	7473171	22	7473171	22	7473171	22
7477026	7	7477026	23	7477026	23	7477026	23
6428773	8	6428773	24	6428773	24	6428773	24
2749402	9	2749402	25	2749402	25	2749402	25
118533	10	118533	26	118533	26	118533	26
4005910	11	4005910	27	4005910	27	4005910	27
5453917	12	5453917	28	5453917	28	5453917	28
7503560	13	7503560	29	7503560	29	7503560	29
7472376	14	7472376	30	7472376	30	7472376	30
1306049	15	1306049	31	1306049	31	1306049	31
3187074	16	3187074	32	3187074	32	3187074	32